PROKARYOTIC COMMUNICATIONS: FROM MACROMOLECULAR INTERDOMAIN TO INTERCELLULAR TALKS (RECOGNITION) AND BEYOND

EDITED BY: Chew Chieng Yeo, Manuel Espinosa and Tatiana Venkova

PUBLISHED IN: Frontiers in Molecular Biosciences



frontiers Research Topics

Frontiers eBook Copyright Statement

The copyright in the text of individual articles in this eBook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this eBook is the property of Frontiers.

Each article within this eBook, and the eBook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this eBook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or eBook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-88966-850-2
DOI 10.3389/978-2-88966-850-2

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

PROKARYOTIC COMMUNICATIONS: FROM MACROMOLECULAR INTERDOMAIN TO INTERCELLULAR TALKS (RECOGNITION) AND BEYOND

Topic Editors:
Chew Chieng Yeo, Sultan Zainal Abidin University, Malaysia
Manuel Espinosa, Spanish National Research Council, Spain
Tatiana Venkova, Fox Chase Cancer Center, United States



Image caption: "Roads to communicate. Roads that go beyond mountains. Roads that lead to knowledge. For ever, roads." Image: San Bernardino Mountains, California, United States. Image courtesy of Steve Kendall, www.purplekitephoto.com.

Citation: Yeo, C. C., Espinosa, M., Venkova, T., eds. (2021). Prokaryotic Communications: From Macromolecular Interdomain to Intercellular Talks (Recognition) and Beyond. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-88966-850-2

Table of Contents

05 ***Editorial: Prokaryotic Communications: From Macromolecular Interdomain to Intercellular Talks (Recognition) and Beyond***
Chew Chieng Yeo, Manuel Espinosa and Tatiana Venkova

09 ***Crosstalk Between Type VI Secretion System and Mobile Genetic Elements***
Arancha Peñil-Celis and M. Pilar Garcillán-Barcia

16 ***Naringenin Inhibition of the* Pseudomonas aeruginosa *Quorum Sensing Response Is Based on Its Time-Dependent Competition With* N-*(3-Oxo-dodecanoyl)-L-homoserine Lactone for LasR Binding***
Sara Hernando-Amado, Manuel Alcalde-Rico, Teresa Gil-Gil, José R. Valverde and José L. Martínez

27 **In vitro *DNA Inversions Mediated by the PsrA Site-Specific Tyrosine Recombinase of* Streptococcus pneumoniae**
Jingwen Li, Juanjuan Wang, Sofía Ruiz-Cruz, Manuel Espinosa, Jing-Ren Zhang and Alicia Bravo

41 ***ComWΔ6 Stimulates Transcription of Pneumococcal Competence Genes* in vitro**
Nicole L. Inniss and Donald A. Morrison

50 ***Genome Segregation by the Venus Flytrap Mechanism: Probing the Interaction Between the ParF ATPase and the ParG Centromere Binding Protein***
Marisa Caccamo, Aneta Dobruk-Serkowska, Fernando Rodríguez-Castañeda, Cecilia Pennica, Daniela Barillà and Finbarr Hayes

62 ***Temperature Dependent Control of the R27 Conjugative Plasmid Genes***
Marta Gibert, Sonia Paytubi, Cristina Madrid and Carlos Balsalobre

72 ***Role of Bacterial Cytoskeleton and Other Apparatuses in Cell Communication***
Divya Singhi and Preeti Srivastava

86 **Bacillus subtilis *PcrA Couples DNA Replication, Transcription, Recombination and Segregation***
María Moreno-del Alamo, Rubén Torres, Candela Manfredi, José A. Ruiz-Masó, Gloria del Solar and Juan Carlos Alonso

102 ***Gene Duplications in the Genomes of Staphylococci and Enterococci***
José Francisco Sanchez-Herrero, Manuel Bernabeu, Alejandro Prieto, Mário Hüttener and Antonio Juárez

116 ***Intermolecular Communication in* Bacillus subtilis*: RNA-RNA, RNA-Protein and Small Protein-Protein Interactions***
Inam Ul Haq, Peter Müller and Sabine Brantl

133 ***Conjugative Coupling Proteins and the Role of Their Domains in Conjugation, Secondary Structure and* in vivo *Subcellular Location***
Itxaso Álvarez-Rodríguez, Begoña Ugarte-Uribe, Igor de la Arada, José Luis R. Arrondo, Carlos Garbisu and Itziar Alkorta

148 ***Type IV Coupling Proteins as Potential Targets to Control the Dissemination of Antibiotic Resistance***
Itxaso Álvarez-Rodríguez, Lide Arana, Begoña Ugarte-Uribe, Elena Gómez-Rubio, Sonsoles Martín-Santamaría, Carlos Garbisu and Itziar Alkorta

167 ***Formation of Complexes Between O Proteins and Replication Origin Regions of Shiga Toxin-Converting Bacteriophages***
Katarzyna Kozłowska, Monika Glinkowska, Lidia Boss, Lidia Gaffke, Jakub Deptuła and Grzegorz Węgrzyn

183 ***Role of Regulated Proteolysis in the Communication of Bacteria With the Environment***
Sarah Wettstadt and María A. Llamas

196 ***Translation Initiation Control of RNase E-Mediated Decay of Polycistronic* gal *mRNA***
Heung Jin Jeon, Changjo Kang, Monford Paul Abishek N, Yonho Lee, Xun Wang, Dhruba K. Chattoraj and Heon M. Lim

212 ***Identification of Novel Thermosensors in Gram-Positive Pathogens***
Pilar Fernández, Alejandra Raquel Díaz, María Florencia Ré, Lucía Porrini, Diego de Mendoza, Daniela Albanesi and María Cecilia Mansilla

224 ***Metabolic Modeling to Interrogate Microbial Disease: A Tale for Experimentalists***
Fabrice Jean-Pierre, Michael A. Henson and George A. O'Toole

235 ***Genome-Wide Analysis of Targets for Post-Transcriptional Regulation by Rsm Proteins in* Pseudomonas putida**
Óscar Huertas-Rosales, Manuel Romero, Kok-Gan Chan, Kar-Wai Hong, Miguel Cámara, Stephan Heeb, Laura Barrientos-Moreno, María Antonia Molina-Henares, María L. Travieso, María Isabel Ramos-González and Manuel Espinosa-Urgel

249 ***Acidic pH Decreases the Endonuclease Activity of Initiator RepB and Increases the Stability of the Covalent RepB-DNA Intermediate while Has Only a Limited Effect on the Replication of Plasmid pMV158 in* Lactococcus lactis**
Rafael Valdelvira, Lorena Bordanaba-Ruiseco, Cristina Martín-Huestamendía, José Angel Ruiz-Masó and Gloria del Solar

265 ***Multiple Layered Control of the Conjugation Process of the* Bacillus subtilis *Plasmid pLS20***
Wilfried J. J. Meijer, D. Roeland Boer, Saúl Ares, Carlos Alfonso, Fernando Rojo, Juan R. Luque-Ortega and Ling Juan Wu

frontiers in Molecular Biosciences

EDITORIAL
published: 21 April 2021
doi: 10.3389/fmolb.2021.670572

Editorial: Prokaryotic Communications: From Macromolecular Interdomain to Intercellular Talks (Recognition) and Beyond

Chew Chieng Yeo[1]*, *Manuel Espinosa*[2] *and Tatiana Venkova*[3]

[1] Centre for Research in Infectious Diseases and Biotechnology (CeRIDB), Faculty of Medicine, Universiti Sultan Zainal Abidin, Kuala Terengganu, Malaysia, [2] Margarita Salas Center for Biological Research, Spanish National Research Council, Madrid, Spain, [3] Fox Chase Cancer Center, Philadelphia, PA, United States

Keywords: prokaryotic communications, quorum sensing, intermolecular communication, intramolecular communication, molecular recognition

OPEN ACCESS

Edited by:
Ray Luo,
University of California, Irvine, United States

Reviewed by:
María Del Pilar Garcillán-Barcia, University of Cantabria, Spain
Jose L. Martinez, Consejo Superior de Investigaciones Científicas (CSIC), Spain

***Correspondence:**
Chew Chieng Yeo
chewchieng@gmail.com

Specialty section:
This article was submitted to Molecular Recognition, a section of the journal Frontiers in Molecular Biosciences

Received: *21 February 2021*
Accepted: *26 March 2021*
Published: *21 April 2021*

Citation:
Yeo CC, Espinosa M and Venkova T (2021) Editorial: Prokaryotic Communications: From Macromolecular Interdomain to Intercellular Talks (Recognition) and Beyond. Front. Mol. Biosci. 8:670572. doi: 10.3389/fmolb.2021.670572

Editorial on the Research Topic

Prokaryotic Communications: From Macromolecular Interdomain to Intercellular Talks (Recognition) and Beyond

INTRODUCTION

Communication is integral to all life on Earth. From complex microbial communities to the regulation of fundamental cellular processes such as DNA replication and transcription, intra- and inter-molecular communications, macromolecular cross-talks, and cell-cell communication lies at the heart of these processes. The concept of intercellular communication in the prokaryotic world came from discoveries made around 50 years ago when Tomasz reported that competence for natural transformation in *Streptococcus pneumoniae* requires a "hormone-like activator" that synchronize the bacterial population for entry into the competent state (Tomasz, 1965). Five years later in 1970, it was reported that the bioluminescent marine bacteria, *Vibrio fischeri* and *Vibrio harveyi* only produced light at high cell density but not when the bacterial cultures were diluted (Nealson et al., 1970). It took another 10 years or so before acyl-homoserine lactone (AHL) was identified as the component, termed the autoinducer, that was responsible for the stimulation of light production in the vibrios (Eberhard et al., 1981) and another 10–15 years before the concept of quorum sensing was firmly established along with its ever-growing lexicon of chemical languages as outlined in the seminal review by (Bassler and Losick, 2006). Besides quorum sensing, other forms of intercellular communications play important roles in microbial interactions. These include secretion systems such as the type VI secretion system (T6SS) that is central to predator-prey interactions (Cherrak et al., 2019), extracellular or membrane vesicles that transfer a multitude of biomolecules including DNA and RNA (Bose et al., 2021) and appendages such as conjugative pili that mediates the horizontal transfer of DNA through conjugation (Virolle et al., 2020). This Research Topic not only covers intercellular communication among bacteria but also the gamut of processes that govern intracellular communication within each bacterial cell.

OVERVIEW OF MANUSCRIPTS IN THIS RESEARCH TOPIC

T6SS is a macromolecular multiprotein complex that functions to deliver killer toxins from the donor bacterium (predator) into the cytoplasm of the target cell (prey) which can either be another bacterium or a eukaryotic cell. Most T6SS are dependent on cell-cell contact and, in some cases, are regulated by quorum sensing networks. T6SS are also influenced by bacterial conjugation and associated genetic elements including plasmids and integrative and conjugative elements (ICEs) and this is the subject of a mini-review by Peñil-Celis and Garcillán-Barcia.

Conjugation is one of the most widespread mechanisms of horizontal gene transfer in bacteria (Virolle et al., 2020) and one of its consequences is the widespread dissemination of antimicrobial resistance genes. The World Health Organization has recognized antibiotic resistance as one of the major threats to public health (Tacconelli et al., 2018). Due to the vital nature of type IV coupling proteins (T4CPs) in the conjugative process, they have become appealing targets in the quest to inhibit bacterial conjugation and, in parallel, to limit the spread of antimicrobial resistance. Two papers were contributed by Itziar Alkorta's research group on T4CPs: Álvarez-Rodríguez, Ugarte-Uribe et al. studied the functionality of T4CPs, especially its transmembrane domain (TMD), in plasmid transfer, their secondary structure, thermal stability, and subcellular localization. Álvarez-Rodríguez, Arana et al. also reviewed our current state of knowledge on T4CPs and explore the possibility of using T4CP inhibitors to impede the spread of antimicrobial resistance.

The IncH1 conjugative plasmids, of which the R27 plasmid is its prototype, have been associated with multidrug resistance in pathogens such as *Salmonella enterica* and *Escherichia coli* (Holt et al., 2011). An interesting feature of the IncH1 plasmids is the temperature-dependence of their conjugation whereby higher conjugative frequencies were observed at low temperatures of between 22 and 30°C. Gibert et al. utilized a transcriptomics approach to characterize the effect of temperature on the expression of R27-encoded genes leading to the finding that the HtdA-TrhR/TrhY regulatory circuit is a likely mediator for the environmental regulation of R27 gene expression. The different regulatory circuits involved in the control of conjugation in the *Bacillus subtilis* plasmid pLS20 are reviewed by Meijer et al. who focuses on the proteins that control the activity of the main promoter, Pc, located upstream of the conjugation operon.

Two integral membrane thermosensor histidine kinases from the Gram-positive pathogens *Staphylococcus aureus* and *Bacillus anthracis* were identified by Fernández et al. who also showed that these histidine kinases likely control the expression of putative ATP-binding cassette (ABC) transporters and were regulated by environmental temperature.

Another Gram-positive pathogen, *Streptococcus pneumoniae*, is well-known to develop competence for the uptake of DNA by natural transformation and this ability enables the pneumococci to have highly diverse genomes, making them well-equipped for adaptation (Andam and Hanage, 2015). SigX (σ^X) is the only known alternative σ-factor in streptococci and is the master competence regulator that enables transcription from competence-specific combox promoters. However, the pneumococcal transformation also requires ComW and Innis and Morrison show that a DNA-binding ComW variant, ComWΔ6, can stimulate transcription from σ^X promoters *in vitro*.

In another paper on *S. pneumoniae*, Li et al. demonstrated that the PsrA tyrosine recombinase controls DNA inversions that are mediated by three inverted repeats in three DNA methyltransferase *hsdS* genes of the type I restriction-modification *cod* locus. This leads to diversification of genomic DNA methylation patterns and phenotypic phase variation in colony opacity, which in turn, leads to adaptation of the pneumococci in colonization and virulence. Interestingly, PsrA activity relies on Mg^{2+} ions, unlike other site-specific tyrosine recombinases.

Gene duplications are characteristic of bacterial genomes and it is another way of generating functional diversity and increasing genomic complexity in prokaryotes. Sanchez-Herrero et al. analyzed the extent of gene duplications in the genomes of three pathogens, *Staphylococcus aureus*, *Enterococcus faecium* and *Enterococcus faecalis*, and found an irregular distribution of duplications in the genomes of the analyzed strains. Although mobile DNA accounted for the majority of the duplicated genes, duplication of core genes was also discovered.

The mechanism of segregation of the low copy number multidrug resistance plasmid TP228 in *Escherichia coli* was explored by Caccamo et al. who presented a unique "Venus flytrap" mechanism involving interactions between the ParF ATPase and the ParG centromere-binding protein. Replication of the Shiga toxin-converting bacteriophages (Stx phages), which are responsible for the virulence of enterohemorrhagic *E. coli* strains, was investigated by Kozłowska et al. An essential step in replication initiation is the binding of the phage-encoded O protein to iterons in the *origin* of replication region of which Stx phages has six as compared to four iterons in phage λ. Kozłowska et al. described the binding of the Stx-encoded O protein to their cognate iterons and compared them to those that occur during the formation of replication complexes in phage λ. The effects of pH on the activity of the Gram-positive plasmid pMV158-encoded RepB replication initiation protein was examined by Valdelvira et al. who showed that despite acidic pH impairing the endonuclease activity of RepB *in vitro*, the ability of pMV158 to replicate in *Lactococcus lactis in vivo* was largely unaffected at extracellular pHs that ranged from 5.0 to 7.0.

The important roles of small regulatory RNAs (sRNAs) and small proteins in intercellular communication and the control of various cellular processes have only been elucidated recently (Wagner and Romby, 2015). Ul Haq et al. reviews our current knowledge of interactions involving sRNA-mRNA, sRNA-protein and small protein-protein interactions in the Gram-positive *Bacillus subtilis*. In another article on the molecular interactions in *B. subtilis*, Moreno-del Alamo et al. showed that the PcrA protein, which is crucial in cell viability, functions at the interface of DNA replication, transcription, recombination, and

segregation by working in tandem with recombination and repair proteins at stalled DNA polymerase/RNA polymerase complexes to facilitate replication beyond any conflict points.

Jeon et al. investigated the RNase E-mediated decay of polycistronic mRNA in *E. coli*, a major mechanism for the regulation of gene expression. A global analysis was performed by Huertas-Rosales et al. to identify RNA sequences bound *in vivo* by three post-transcriptional regulators of the RsmA/CsrA family in *Pseudomonas putida* KT2440 leading to an overview of the network of genes and processes subjected to post-transcriptional regulation in a bacterium that harbors three Rsm homologs.

Hernando-Amado et al. investigated the molecular basis of the plant flavonoid naringenin as an anti-quorum sensing compound and showed that naringenin directly binds to the quorum-sensing regulator LasR of *Pseudomonas aeruginosa*.

The role of the bacterial cytoskeleton in cell-cell communication along with cellular apparatuses that function in contact-dependent signaling is extensively reviewed by Singhi and Srivastava. Regulated proteolysis is a fundamental process used in the transmission of extracellular signals through the cell membrane into the cytosol. The role of regulated proteolysis in the communication of bacteria with the environment through modulation of bacterial signal transduction systems is reviewed by Wettstadt and Llamas.

Finally, in a mini-review, Jean-Pierre et al. present the latest advances in metabolic modeling, a computational method used to predict metabolic capabilities and interactions from individual microorganisms to complex polymicrobial communities in efforts to better understand microbial community function.

WHERE DO WE GO FROM HERE?

The manuscripts that are compiled under the aegis of this Research Topic gave us an indication of the breadth and depth of research that has been carried out on prokaryotic communications in the post-genomic era—either intracellular, cell-cell, cell-host, or cell-environment. From the intricate intra- and intermolecular dance that characterizes the fundamental cellular processes of DNA replication, segregation, transcription, translation, and post-translational modifications to the complex but elegant interplay of the various components of the cellular regulatory networks, the wonders of the living prokaryotic cell never cease to amaze. The dictionary of the chemical language of bacterial cells continues to grow and we are constantly finding ways to exploit our knowledge of them to mitigate the spread of multidrug-resistant pathogens. The success of this Research Topic led us to launch a second volume where we hope to further explore the nooks, the crannies, the various alleyways, paths, avenues, and roads that lie ahead of us in our quest to understand how the most abundant living organisms on planet Earth talk.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

While this Editorial was written, the authors were funded by the following research Grants: FRGS/1/2019/SKK11/UNISZA/02/1 from the Malaysian Ministry of Higher Education and UniSZA/LABMAT/2018/09 from Universiti Sultan Zainal Abidin (to CCY), PID2019-104553RB-C21, Spanish Ministry of Science and Innovation (to ME).

ACKNOWLEDGMENTS

Our heartfelt thanks are due to all the generous contributors and reviewers to this Research Topic as well as the gracious and helpful support of the journal's Editorial Team.

REFERENCES

Andam, C. P., and Hanage, W. P. (2015). Mechanisms of genome evolution of Streptococcus. *Infect. Genet. Evol.* 33, 334–342. doi: 10.1016/j.meegid.2014.11.007

Bassler, B. L., and Losick, R. (2006). Bacterially speaking. *Cell* 125, 237–246. doi: 10.1016/j.cell.2006.04.001

Bose, S., Aggarwal, S., Singh, D. V., and Acharya, N. (2021). Extracellular vesicles: an emerging platform in gram-positive bacteria. *Microb. Cell* 7, 312–322. doi: 10.15698/mic2020.12.737

Cherrak, Y., Flaugnatti, N., Durand, E., Journet, L., and Cascales, E. (2019). Structure and activity of the type VI secretion system. *Microbiol. Spectr.* 7:PSIB-0031-2019. doi: 10.1128/9781683670285.ch26

Eberhard, A., Burlingame, A. L., Eberhard, C., Kenyon, G. L., Nealson, K. H., and Oppenheimer, N. J. (1981). Structural identification of autoinducer of *Photobacterium fischeri* luciferase. *Biochemistry* 20, 2444–2449. doi: 10.1021/bi00512a013

Holt, K. E., Phan, M. D., Baker, S., Duy, P. T., Nga, T. V. T., Nair, S., et al. (2011). Emergence of a globally dominant IncHI1 plasmid type associated with multiple drug resistant typhoid. *PLoS Negl. Trop. Dis.* 5:e1245. doi: 10.1371/journal.pntd.0001245

Nealson, K. H., Platt, T., and Hastings, J. W. (1970). Cellular control of the synthesis and activity of the bacterial luminescent system. *J. Bacteriol.* 104, 313–322. doi: 10.1128/JB.104.1.313-322.1970

Tacconelli, E., Carrara, E., Savoldi, A., Harbarth, S., Mendelson, M., Monnet, D. L., et al. (2018). Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. *Lancet Infect. Dis.* 18, 318–327. doi: 10.1016/S1473-3099(17)30753-3

Tomasz, A. (1965). Control of the competent state in pneumococcus by a hormone-like cell product: an example for a new type of regulatory mechanism in bacteria. *Nature* 208, 155–159. doi: 10.1038/208155a0

Virolle, C., Goldlust, K., Djermoun, S., Bigot, S., and Lesterlin, C. (2020). Plasmid transfer by conjugation in gram-negative bacteria: from the cellular to the community level. *Genes* 11:1239. doi: 10.3390/genes11111239

Wagner, E. G. H., and Romby, P. (2015). Small RNAs in bacteria and archaea: who they are, what they do, and how they do it. *Adv. Genet.* 90, 133–208. doi: 10.1016/bs.adgen.2015.05.001

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Yeo, Espinosa and Venkova. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Molecular Biosciences

MINI REVIEW
published: 13 November 2019
doi: 10.3389/fmolb.2019.00126

Crosstalk Between Type VI Secretion System and Mobile Genetic Elements

*Arancha Peñil-Celis and M. Pilar Garcillán-Barcia**

Instituto de Biomedicina y Biotecnología de Cantabria, Universidad de Cantabria—Consejo Superior de Investigaciones Científicas, Santander, Spain

Many bacterial processes require cell-cell contacts. Such are the cases of bacterial conjugation, one of the main horizontal gene transfer mechanisms that physically spreads DNA, and the type VI secretion systems (T6SSs), which deploy antibacterial activity. Bacteria depend on conjugation to adapt to changing environments, while T6SS killing activity could pose a threat to mating partners. Here we review the experimental evidences of overlapping and interaction between the T6SSs, bacterial conjugation, and conjugative genetic elements.

Keywords: type VI secretion system, bacterial conjugation, horizontal gene transfer, cell-cell communication, T6SS regulation, mobile genetic elements, plasmids, integrative and conjugative elements

OPEN ACCESS

Edited by:
Tatiana Venkova, Fox Chase Cancer Center, United States

Reviewed by:
Elisabeth Grohmann, Beuth Hochschule für Technik Berlin, Germany
Luis Rey, Polytechnic University of Madrid, Spain

****Correspondence:***
M. Pilar Garcillán-Barcia
garcilmp@unican.es

Specialty section:
This article was submitted to Molecular Recognition, a section of the journal Frontiers in Molecular Biosciences

Received: *23 September 2019*
Accepted: *28 October 2019*
Published: *13 November 2019*

Citation:
Peñil-Celis A and Garcillán-Barcia MP (2019) Crosstalk Between Type VI Secretion System and Mobile Genetic Elements. Front. Mol. Biosci. 6:126. doi: 10.3389/fmolb.2019.00126

INTRODUCTION

Microbes often exist in complex multispecies communities where interaction is essential for keeping a balanced microbial ecosystem. Bacterial survival depends on the ability to succeed in the competition against rival bacterial cells. Critical to this goal is the type VI secretion system (T6SS), a macromolecular multiprotein complex dedicated to the delivery of toxins across the cell envelope of the donor bacteria (predator) into the cytoplasm of a target cell (prey), either another bacterium, or an eukaryotic cell [for recent reviews on T6SS structure, assembly, and activity, see (Cianfanelli et al., 2016; Nazarov et al., 2018; Nguyen et al., 2018; Cherrak et al., 2019; Wang et al., 2019). Only a few phyla of Gram-negative microorganisms encode T6SSs (Abby et al., 2016), suggesting a role for horizontal gene transfer (HGT) in its origin. Four T6SS phylogenetic subtypes are distinguished (Russell et al., 2014; Böck et al., 2017). Several protein complexes compose the T6SS most common subtype (T6SS^{i}): a membrane complex (Durand et al., 2012, 2015; Rapisarda et al., 2019), which encompasses proteins TssJ, TssL, and TssM, a baseplate [TssE, TssF, TssG, and TssK (Nazarov et al., 2018; Park et al., 2018; Liebl et al., 2019)], a syringe (Wang et al., 2017), which is in turn composed of an inner tube that is secreted by the functional system [Hcp (Douzi et al., 2014)], a needle spike [VgrG (Uchida et al., 2014) and PAAR (Shneider et al., 2013)], and a sheath [TssB and TssC (Kudryashev et al., 2015)]. Furthermore, a sheath assembly protein, TssA (Schneider et al., 2019), and a sheath disassembly ATPase, TssH (also known as ClpV) (Pietrosiuk et al., 2011) are part of the system. A plethora of T6SS protein effectors and their corresponding immunity proteins have been identified (Durand et al., 2014; Kostiuk et al., 2017; Lien and Lai, 2017).

A claimed common evolutionary origin for T6SSs and contractile tailed-bacteriophages is supported by the conservation of some of their components (Hcp, VgrG, TssC, TssF, TssA, TssK, TssG, and TssE (Pukatzki et al., 2007; Leiman et al., 2009; Lossi et al., 2011; Planamente et al., 2016), and the similarities in the assembly mechanisms between the T6SS and the bacteriophage tail tube-sheath (Veesler and Cambillau, 2011; Lossi et al., 2013; Brunet et al., 2014). Accretion of DotU/IcmH and IcmF homologs of the Dot/Icm T4SS of *Legionella pneumophila* (Christie, 2016), TssL and TssM, which mediate the polar targeting of such T4SS (Ghosal et al., 2019), provides a docking station for the phage-like T66S device (Durand et al., 2015).

Most T6SSs are functional in a cell-cell contact-dependent way. Cell-cell communication mechanisms are thus expected to have an impact on T6SS activity. Several examples support the T6SS regulation by quorum sensing (QS) networks (Pena et al., 2019). A phosphorylation cascade that mediates a positive regulatory loop entwining T6SS and QS was found in *Vibrio alginolyticus* (Yang et al., 2018). In *Vibrio cholerae*, expression of Hcp showed to be growth phase-dependent, positively and negatively regulated by the QS regulators HapR and LuxO, respectively (Ishikawa et al., 2009; Zheng et al., 2010). Biofilm formation and T6SS expression in *Burkolderia cenocepacia* were upregulated by QS (Aubert et al., 2013), while expression of T6SS effector and immunity genes of *Burkholderia thailandensis* was also activated by QS (Majerczyk et al., 2016). In the case of *Pseudomonas aeruginosa,* its *three* T6SSs were differentially regulated by QS: whereas QS regulators LasR and MvfR suppressed the expression of H1-T6SS, they positively regulated that of H2- and H3-T6SS (Lesic et al., 2009). On the other hand, although cell-cell contact is not a requirement for natural transformation, in *Vibrio cholerae,* T6SS is part of the competence regulon and functions as a transformation enhancer for DNA acquisition (Borgeaud et al., 2015; Veening and Blokesch, 2017; Matthey et al., 2019). Competence-mediated induction of T6SS released DNA and made it accessible for HGT in *Vibrio cholerae*. Besides, competence-induced T6SS-mediated killing increased the natural transformation efficiency and boosted the acquisition of large genomic regions from killed neighbors (Borgeaud et al., 2015), including novel functional T6SS effector-immunity pairs (Thomas et al., 2017).

The above results suggest that T6SS and cell-cell communication processes are interwoven. Here we will focus on the contribution of another intercellular mechanism, bacterial conjugation, and its associated players, plasmids and integrative and conjugative elements (ICEs), to the activity of T6SS.

T6SS CONTROL BY PLASMID-ENCODED REGULATORS

T6SSs are implicated in a wide range of functions and regulated by a large diversity of mechanisms (Bernard et al., 2010; Alteri and Mobley, 2016). Transcriptional regulators of T6SS loci have been detected in plasmids. The first report on a T6SS regulation by a plasmid-encoded regulator was that of Sci-2 (**Figure 1A**). Sci-2 T6SS is encoded in entero-aggregative, avian-pathogenic and Shiga toxin-producing *Escherichia coli* strains (Journet and Cascales, 2016). It confers a growth advantage to entero-aggregative *E. coli* (EAEC) by causing non-immune *E. coli* killing (Brunet et al., 2013). The prototype EAEC strain 042 carries the pAA2 plasmid, which encodes a transcriptional regulator of the AraC family, the *aggR* gene (Nataro et al., 1994). AggR plays a central role in modulating adherence of EAEC 042 by activating plasmid-borne genes, such as the attachment adherence fimbriae *aafDA* (Elias et al., 1999), the anti-aggregative protein dispersin *aap* (Sheikh et al., 2002), and the T1SS *aatPABCD* for Aap transport (Nishi et al., 2003). EAEC chromosomal genes located in the pathogenicity island PAI-1 are also included in the AggR regulon (Dudley et al., 2006; Morin et al., 2013; Yasir et al., 2019). Among these genes, a cluster designated *aaiA-P* (Sci-2) was found expressed at least 2-fold higher in the wild-type strain 042 than in the *aggR*- derivative using a microarray approach (Dudley et al., 2006). Two proteins of this cluster were identified by mass spectrometry analysis in the whole-cell proteome of the wild-type strain, while they were absent in the *aggR*- derivative. One of

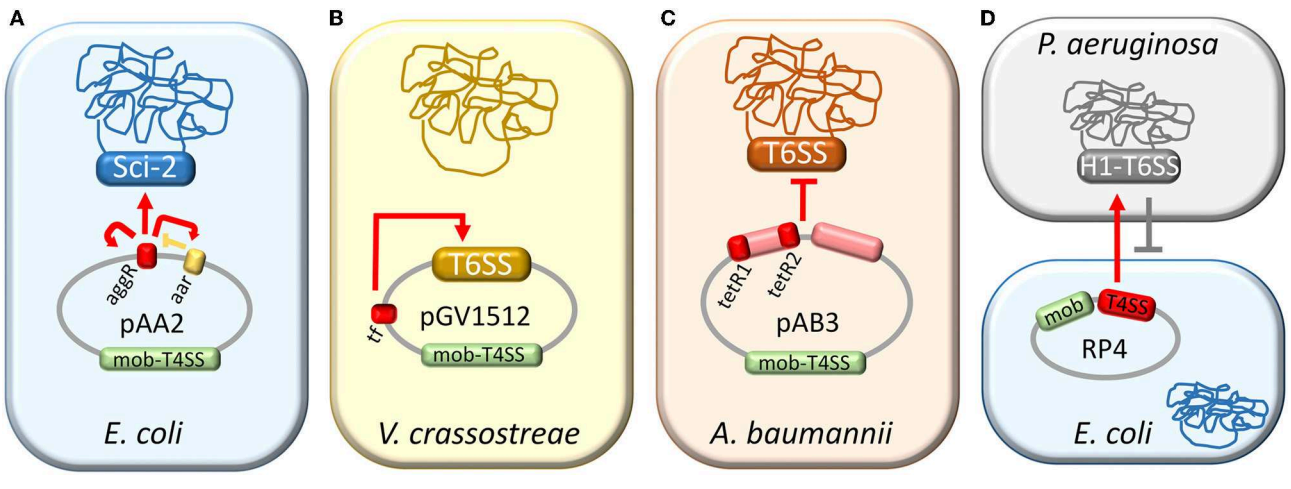


FIGURE 1 | Control of T6SS activity by plasmids. Plasmid-mediated regulation of T6SSs in *E. coli* **(A)**, *V. crassostreae* **(B)**, *A. baumannii* **(C)**, and *P. aeruginosa* **(D)**. Plasmid regions involved in T6SS regulation are depicted in red. **(A)** pAA2-encoded auto-regulator AggR activates the Sci-2 T6SS (red arrows) and is negatively regulated by Aar (yellow lines). **(B)** Transcriptional factor (TF) of pGV1512 activates the plasmid-encoded T6SS (red arrow). **(C)** Two regions of the pAB3 plasmid, one of them including *tetR*-like regulators, repress the T6SS of *A. baumannii* (red lines) **(D)** Cell-cell interactions mediated by the T4SS of plasmid RP4 trigger H1-T6SS counterattack (red arrow), resulting in a decreased survival specifically in plasmid-bearing cells (gray lines).

these proteins, AaiC (Hcp) was also detected in the supernatant of exponentially-growing cells cultured in DMEM (Dudley et al., 2006), a condition that induces AggR production (Sheikh et al., 2001). An AggR-dependent promoter was localized in the region comprised from 466 bp upstream to 300 bp downstream of the predicted *aaiA* translational start site (Dudley et al., 2006). In the wild-type strain, the expression from this promoter was approximately 3-fold over that seen in the *aggR* mutant.

AggR was found to autoactivate its expression (Morin et al., 2010). By primer extension analysis, the transcriptional start site of the *aggR* promoter (P*aggR*) was located 40 nucleotides upstream of the translational start (Morin et al., 2010). DNA footprinting experiments revealed the presence of two AggR-binding sites: one upstream of the transcriptional start site and one downstream. A consensus sequence resembling sites for the Rns regulator from enterotoxigenic *E. coli* (Munson, 2013) was found for AggR binding: ANNNNNNTATC (Morin et al., 2010). EAEC plasmid and chromosomal AggR-regulated genes identified by RNAseq were found to contain a WWWWWWWTATC (where W means A or T) sequence spaced 21–23 bp upstream of the −10 promoter elements (Yasir et al., 2019). AggR is negatively controlled by the plasmid-encoded Aar repressor protein through direct binding (Santiago et al., 2016), while the *aar* gene is in turn positively regulated by AggR (Santiago et al., 2014). It remains unknown what environmental signal is sensed by AggR (Morin et al., 2013).

Another example of T6SS regulation by an AraC-like activator is provided by pGV1512, a conjugative plasmid that invaded *Vibrio crassostreae*, turning it from a benign oyster commensal into a pathogen (Bruto et al., 2017; **Figure 1B**). This plasmid encodes a T6SS organized into two divergently-transcribed operons. A plasmid region located between T6SS and T4SS, Px3, was found to be necessary for *V. crassostreae* virulence (Bruto et al., 2017). Px3 encodes a transcriptional regulator (TF) of the AraC family, which activated the transcription of plasmid-encoded T6SS genes, as shown by RNAseq analysis, and restored the virulence to a ΔPx3 mutant (Piel et al., 2019). Two functional promoters were located in the intergenic region between T6SS operons. Deletion of a palindromic sequence of six nucleotides spaced by five nucleotides, located between both promoters, altered the induction capacity of TF (Piel et al., 2019).

Negative transcriptional regulation exerted by a plasmid to a chromosomally-encoded T6SS has been also documented in *Acinetobacter baumannii* (**Figure 1C**). T6SSs are conserved and syntenic among *A. baumannii* strains (Weber et al., 2013). *A. baumannii* ATCC 17978 constitutively produces and secretes Hcp, a hallmark of an active T6SS, but initial efforts to observe a predator phenotype failed. There were no significant differences in the survival of any of the preys (*E. coli, A. baumannii, A. nosocomialis, A. baylyi*) in different experimental conditions tested (Weber et al., 2013). Nevertheless, when individual colonies from a clinical isolate of *A. baumannii* (strain Ab 04) were analyzed for Hcp secretion by ELISA, two contrasting profiles were found (Weber et al., 2015). Some colonies behaved coherently with a T6SS- and others with a T6SS+ phenotype. Whole-genome sequencing of both Ab 04 subpopulations revealed that their genomes encoded a T6SS, though they differed in the carriage of a multidrug-resistance conjugative plasmid (pAB04-1), absent in the cells exhibiting the T6SS+ phenotype.

This plasmid was not fully stable in *A. baumannii*, so a subpopulation lost it, rendering colonies that produced and secreted Hcp while becoming susceptible to antimicrobials. Colonies displaying the ±T6SS phenotypes were also isolated for *A. baumannii* strains 17978 and 1438 (Weber et al., 2015). In all cases, T6SS+ cells efficiently killed *E. coli* in competition assays, in contrast to their T6SS- counterparts. This killing was dependent on a functional T6SS, as verified by using mutants of essential T6SS components, which did not kill *E. coli*. Mating experiments in which the transmissible plasmid was transferred from donor *A. baumannii* cells with the T6SS- phenotype to T6SS+ recipient cells rendered transconjugants deficient in bacterial killing. These experiments provided evidence that pAB04-1 encoded the genetic determinants for silencing the *A. baumannii* T6SS activity.

The pAB04-1 backbone is commonly found in other *A. baumannii* strains, and the plasmid variants mainly differ in their antimicrobial-resistance cargoes (Weber et al., 2015). *A. baumannii* prototype strain, ATCC 17978, harbors a pAB04-1-like plasmid, pAB3. Transferred to either a clonally-unrelated *A. baumannii* strain, *A. baylyi* or *A. nosocomialis*, the transconjugants did not secrete Hcp, indicating that the T6SSs of these strains were also susceptible to repression by the plasmid and thus pointing to a common repression mechanism in a broad range of *Acinetobacter* species (Di Venanzio et al., 2019b). Another pAB04-1-like plasmid, pAB5, from *A. baumannii* strain UPAB1, was shown to regulate the expression of multiple chromosomally-encoded virulence factors, including T6SS (Di Venanzio et al., 2019a). An UPAB1 derivative lacking pAB5 displayed increased susceptibility to multiple antibiotics and activation of the T6SS, as observed by secretion of the T6SS-associated protein Hcp. Two *tetR*-like regulators were found to be encoded only in pAB04-1-like plasmids (Weber et al., 2015). The deletion of either pAB3 *tetR* gene or both increased Hcp expression but did not restore Hcp secretion (Di Venanzio et al., 2019b). The deletion of the *tetR*-like repressor genes was thus not enough to trigger *per se* the T6SS activity. Nevertheless, when these regulators were overexpressed in a T6SS+ phenotype ATCC 17978 population, Hcp expression and secretion drastically decreased, and this transformed population was impaired as a predator in its *E. coli* killing ability (Weber et al., 2015).

T6SS ACTIVITY AFFECTS BACTERIAL CONJUGATION

P. aeruginosa senses exogenous attacks by the T6SS of akin and non-akin bacteria and post-translationally activates its H1-T6SS at the precise location of the initial strikes (Basler and Mekalanos, 2012; Basler et al., 2013). Other perturbations of the cell envelope, such as the presence of the Gram-negative bacterial

membrane disruptor polymyxin B, also increased the H1-T6SS activity (Ho et al., 2013). These results suggested that membrane perturbations trigger the H1-T6SS activity.

Bacterial conjugation is an HGT mechanism that involves contact between the cell envelopes of the mating-pair partners, i.e., donor and recipient cells. It is essentially mediated by plasmids and integrative and conjugative elements (ICEs) (Smillie et al., 2010; Guglielmini et al., 2011). These genomic platforms encode the genetic requirements for their transfer: a *mob* region containing an origin of transfer (the single element strictly required *in cis*), a relaxase (Smillie et al., 2010; Zechner et al., 2017; Guzmán-Herrador and Llosa, 2019), and a type IV coupling protein (Gomis-Rüth et al., 2004; Peña and Arechaga, 2013); and a mating-pair bridge composed of a type IV secretion system (T4SS) (Cabezón et al., 2015; Bergé et al., 2017; Grohmann et al., 2018; Li et al., 2019).

Donor *E. coli* strains bearing broad host-range conjugative plasmids RP4/RK2 or pKM101, but not the narrow host-range plasmid F, were more sensitive to killing by a T6SS+ *P. aeruginosa* recipient than plasmid-lacking donors (Ho et al., 2013; **Figure 1D**). Furthermore, RP4-containing *E. coli* donors were selectively killed by *P. aeruginosa* in a mixed culture with RP4-lacking *E. coli* cells. *P. aeruginosa* mutants defective in the attack-sensing pathway genes *tagT* and *pppA* exhibited greater conjugation efficiency as recipient strains and did not kill RP4-bearing *E. coli* donors. Transposon mutagenesis of RP4 rendered T4SS mutants with impaired ability to transfer and induce T6SS donor-directed killing response in *P. aeruginosa.* RP4 mutants in the relaxosome components genes *traI* and *traJ* and the coupling protein gene *traG* were also defective in DNA conjugation but showed increased donor-directed T6SS response compared to wild type RP4, suggesting that successful DNA transfer was not necessary to activate a T6SS attack by *P. aeruginosa*. So, mating donors can trigger a T6SS response in the recipient, which in turn causes a decrease in their survival.

In the case of *A. baumannii*, conjugative dissemination of the pAB4-01/pAB3 plasmids relies on the repression of the T6SS encoded in the donor strain. To evaluate the impact of the plasmid-mediated T6SS repression in plasmid dissemination through conjugation, pAB3 derivatives were used in mating assays (Di Venanzio et al., 2019b). One of them, pAB3Δ*tetR1,2* lacked genes ACX60_RS18875-ACX60_RS18900 (GenBank Acc. No. NZ_CP012005.1), a region that contains the *tetR*-like genes *tetR1* and *tetR2*. On this mutant, a second deletion comprising genes ACX60_RS18760-ACX60_RS18795 was introduced, producing plasmid pAB3*. Wild-type and mutant plasmids were efficiently transferred between isogenic *A. baumannii* ATCC 17,978 strains (T6SS-resistant). When a non-immunogenic and thus T6SS-susceptible ATCC 17978 derivative was used as a recipient, pAB3 and pAB3Δ*tetR1,2* transfer efficiencies were not altered. Nevertheless, the transfer of the mutant plasmid pAB3* and that of mobilizable plasmids relying on pAB3* as a helper were practically abolished due to recipient killing mediated by the donor's T6SS. Congruently, in competition experiments pAB3 overcame the mutant plasmid pAB3* at invading a T6SS-susceptible population. On the other hand, the efficiency of plasmid conjugation was also deeply affected when the recipient strain was T6SS-proficient and non-isogenic to the donor. pAB3-like plasmids would thus seem to guarantee their transmission by preventing T6SS-mediated killing of non-isogenic recipient strains, depending on their capacity to repress T6SS activity in the donor cells.

T6SSs ENCODED IN MOBILE PLATFORMS

The above results point to T6SS and conjugation as incompatible processes. Plasmids and ICEs are mobile genetic platforms that rely on bacterial conjugation for their dissemination, and thus do not seem *a priori* good platforms for T6SS. Plasmids have a predominant role as genetic couriers (Halary et al., 2010). Twenty-nine plasmids encoding T6SS genes are listed in the T6SS database SecReT6 (http://db-mml.sjtu.edu.cn/SecReT6/) (Li et al., 2015). Experimental evidence on these T6SSs is available only for that encoded in the 2.1 Mb megaplasmid pGMI1000MP of the plant pathogen *Ralstonia solanacearum* GMI1000. A strain containing a *tssB* mutant of this plasmid was reported to be impaired in Hcp secretion and biofilm formation (Zhang et al., 2014). It had also significantly attenuated its virulence on tomato plants. Megaplasmids from other *R. solanacearum* strains also encode T6SS genes: RCFBPv3_mp, CMR15_mp, FQY_4 megaplasmid, Po82 megaplasmid, and mpPSI07. Plasmids pESA3 (*Cronobacter sakazakii*) and pEA320 (*Pantoea ananatis*) are also recorded in SecReT6. The existence of plasmids encoding T6SS genes in these species does not seem anecdotal. Using PCR probes based on the T6SS of pESA3, at least a partial T6SS cluster was detected in 175 out of 177 plasmid-harboring *C. sakazakii* strains analyzed (Franco et al., 2011) and comparative genomics of T6SSs in strains of *P. ananatis* from different environments revealed a plasmid-borne T6SS in a third of the analyzed strains (Shyntum et al., 2014). This T6SS type was restricted to strains of *P. ananatis* isolated from symptomatic plant material, suggesting the possibility of an association between the plasmid-borne T6SS and either pathogenicity or host specificity.

Out of the SecReT6 database, and besides the above-mentioned in pGV1512, a functional T6SS was reported in the *Rizobium etli* Mim1 megaplasmid pRETMiM1f (Salinero-Lanzarote et al., 2019). Immunodetection of Hcp protein indicated that this T6SS was active at high cell densities, in the presence of root exudates, and in bean nodules. *R. etli* T6SS-mutants produced plants with lower dry weight and smaller nodules than the wild-type strain, indicating for the first time that the T6SS played a positive role in *Rhizobium*-legume symbiosis.

ICEs are the most abundant conjugative elements in practically all prokaryotic clades (Guglielmini et al., 2011). Genome analysis strongly suggested transfer of the T6SS^{iii} subtype mediated by an ICE between Bacteroidales strains within the human gut ecosystem (Coyne et al., 2014). T6SS^{iii} has been exclusively found in Bacteroidetes (Abby et al., 2016) and functions in a mechanistically similar manner to T6SS^{i} to target competitor bacteria (Russell et al., 2014). T6SS^{iii} loci were found in more than half of human gut Bacteroidales strains, and they segregated into three evolutionarily-distinct genetic

architectures, two of which were located on ICEs, and one of them had been transferred among co-resident Bacteroidales species in the human gut (Coyne et al., 2016). Low diversity in the effector-immunity pairs of the ICE-encoded T6SS[iii] was detected in human microbiome samples, an indicator of neighbor compatibility likely facilitated by HGT through ICE conjugation (Verster et al., 2017).

CONCLUSIONS

The antibacterial activity of T6SS affects bacterial conjugation by either killing plasmid-bearing donors or non-immune recipients and thus could undermine the DNA spreading. In turn, conjugative platforms (plasmids and ICEs) encode regulatory elements for controlling the T6SS activity. More intriguing is the fact that T6SS is not confined to the chromosome, but also present in these mobile platforms. How plasmids and ICEs have found their ways to deal with this weaponry carriage without sacrificing their transfer potential remains to be researched.

AUTHOR CONTRIBUTIONS

MG-B conceived the study. AP-C and MG-B wrote the manuscript.

FUNDING

This work was supported by Consejo Superior de Investigaciones Científicas (201820I143 to MG-B). We acknowledge support of the publication fee by the CSIC Open Access Publication Support Initiative through its Unit of Information Resources for Research (URICI).

REFERENCES

Abby, S. S., Cury, J., Guglielmini, J., Néron, B., Touchon, M., and Rocha, E. P. C. (2016). Identification of protein secretion systems in bacterial genomes. *Sci. Rep.* 6:23080. doi: 10.1038/srep23080

Alteri, C. J., and Mobley, H. L. T. (2016). The versatile type VI secretion system. *Microbiol. Spectr.* 4, 1–26. doi: 10.1128/microbiolspec.VMBF-0026-2015

Aubert, D. F., O'Grady, E. P., Hamad, M. A., Sokol, P. A., and Valvano, M. A. (2013). The Burkholderia cenocepacia sensor kinase hybrid AtsR is a global regulator modulating quorum-sensing signalling. *Environ. Microbiol.* 15, 372–385. doi: 10.1111/j.1462-2920.2012.02828.x

Basler, M., Ho, B. T., and Mekalanos, J. J. (2013). Tit-for-tat: type VI secretion system counterattack during bacterial cell-cell interactions. *Cell* 152, 884–894. doi: 10.1016/j.cell.2013.01.042

Basler, M., and Mekalanos, J. J. (2012). Type 6 secretion dynamics within and between bacterial cells. *Science* 337:815. doi: 10.1126/science.1222901

Bergé, C., Waksman, G., and Terradot, L. (2017). Structural and molecular biology of type IV secretion systems. *Curr. Top. Microbiol. Immunol.* 413, 31–60. doi: 10.1007/978-3-319-75241-9_2

Bernard, C. S., Brunet, Y. R., Gueguen, E., and Cascales, E. (2010). Nooks and crannies in type VI secretion regulation. *J. Bacteriol.* 192, 3850–3860. doi: 10.1128/JB.00370-10

Böck, D., Medeiros, J. M., Tsao, H.-F., Penz, T., Weiss, G. L., Aistleitner, K., et al. (2017). *In situ* architecture, function, and evolution of a contractile injection system. *Science* 357, 713–717. doi: 10.1126/science.aan7904

Borgeaud, S., Metzger, L. C., Scrignari, T., and Blokesch, M. (2015). The type VI secretion system of *Vibrio cholerae* fosters horizontal gene transfer. *Science* 347, 63–67. doi: 10.1126/science.1260064

Brunet, Y. R., Espinosa, L., Harchouni, S., Mignot, T., and Cascales, E. (2013). Imaging type VI secretion-mediated bacterial killing. *Cell Rep.* 3, 36–41. doi: 10.1016/j.celrep.2012.11.027

Brunet, Y. R., Hénin, J., Celia, H., and Cascales, E. (2014). Type VI secretion and bacteriophage tail tubes share a common assembly pathway. *EMBO Rep.* 15, 315–321. doi: 10.1002/embr.201337936

Bruto, M., James, A., Petton, B., Labreuche, Y., Chenivesse, S., Alunno-Bruscia, M., et al. (2017). Vibrio crassostreae, a benign oyster colonizer turned into a pathogen after plasmid acquisition. *ISME J.* 11, 1043–1052. doi: 10.1038/ismej.2016.162

Cabezón, E., Ripoll-Rozada, J., Peña, A., de la Cruz, F., and Arechaga, I. (2015). Towards an integrated model of bacterial conjugation. *FEMS Microbiol. Rev.* 39, 81–95. doi: 10.1111/1574-6976.12085

Cherrak, Y., Flaugnatti, N., Durand, E., Journet, L., and Cascales, E. (2019). Structure and activity of the type VI secretion system. *Microbiol. Spectr.* 7, 1–11. doi: 10.1128/microbiolspec.PSIB-0031-2019

Christie, P. J. (2016). The mosaic type IV secretion systems. *EcoSal Plus* 7, 1–34. doi: 10.1128/ecosalplus.ESP-0020-2015

Cianfanelli, F. R., Monlezun, L., and Coulthurst, S. J. (2016). Aim, load, fire: the type VI secretion system, a bacterial nanoweapon. *Trends Microbiol.* 24, 51–62. doi: 10.1016/j.tim.2015.10.005

Coyne, M. J., Roelofs, K. G., and Comstock, L. E. (2016). Type VI secretion systems of human gut bacteroidales segregate into three genetic architectures, two of which are contained on mobile genetic elements. *BMC Genomics* 17:58. doi: 10.1186/s12864-016-2377-z

Coyne, M. J., Zitomersky, N. L., McGuire, A. M., Earl, A. M., and Comstock, L. E. (2014). Evidence of extensive DNA transfer between bacteroidales species within the human gut. *MBio* 5, e01305–e01314. doi: 10.1128/mBio.01305-14

Di Venanzio, G., Flores-Mireles, A. L., Calix, J. J., Haurat, M. F., Scott, N. E., Palmer, L. D., et al. (2019a). Urinary tract colonization is enhanced by a plasmid that regulates uropathogenic Acinetobacter baumannii chromosomal genes. *Nat. Commun.* 10:2763. doi: 10.1038/s41467-019-10706-y

Di Venanzio, G., Moon, K. H., Weber, B. S., Lopez, J., Ly, P. M., Potter, R. F., et al. (2019b). Multidrug-resistant plasmids repress chromosomally encoded T6SS to enable their dissemination. *Proc. Natl. Acad. Sci. U.S.A.* 116, 1378–1383. doi: 10.1073/pnas.1812557116

Douzi, B., Spinelli, S., Blangy, S., Roussel, A., Durand, E., Brunet, Y. R., et al. (2014). Crystal structure and self-interaction of the type VI secretion tail-tube protein from enteroaggregative *Escherichia coli*. *PLoS ONE* 9:e86918. doi: 10.1371/journal.pone.0086918

Dudley, E. G., Thomson, N. R., Parkhill, J., Morin, N. P., and Nataro, J. P. (2006). Proteomic and microarray characterization of the AggR regulon identifies a pheU pathogenicity island in enteroaggregative *Escherichia coli*. *Mol. Microbiol.* 61, 1267–1282. doi: 10.1111/j.1365-2958.2006.05281.x

Durand, E., Cambillau, C., Cascales, E., and Journet, L. (2014). VgrG, Tae, Tle, and beyond: the versatile arsenal of Type VI secretion effectors. *Trends Microbiol.* 22, 498–507. doi: 10.1016/j.tim.2014.06.004

Durand, E., Nguyen, V. S., Zoued, A., Logger, L., Spinelli, S., Desmyter, A., et al. (2015). Biogenesis and structure of a type VI secretion membrane core complex. *Nature* 523, 555–560. doi: 10.1038/nature14667

Durand, E., Zoued, A., Spinelli, S., Watson, P. J. H., Aschtgen, M.-S., Journet, L., et al. (2012). Structural characterization and oligomerization of the TssL protein, a component shared by bacterial type VI and type IVb secretion systems. *J. Biol. Chem.* 287, 14157–14168. doi: 10.1074/jbc.M111.338731

Elias, W. P., Czeczulin, J. R., Henderson, I. R., Trabulsi, L. R., and Nataro, J. P. (1999). Organization of biogenesis genes for aggregative adherence fimbria II defines a virulence gene cluster in enteroaggregative *Escherichia coli*. *J. Bacteriol.* 181, 1779–85.

Franco, A. A., Hu, L., Grim, C. J., Gopinath, G., Sathyamoorthy, V., Jarvis, K. G., et al. (2011). Characterization of putative virulence genes on the related RepFIB plasmids harbored by *Cronobacter* spp. *Appl. Environ. Microbiol.* 77, 3255–3267. doi: 10.1128/AEM.03023-10

Ghosal, D., Jeong, K. C., Chang, Y.-W., Gyore, J., Teng, L., Gardner, A., et al. (2019). Molecular architecture, polar targeting and biogenesis of the Legionella Dot/Icm T4SS. *Nat. Microbiol.* 4, 1173–1182. doi: 10.1038/s41564-019-0427-4

Gomis-Rüth, F. X., Solà, M., de la Cruz, F., and Coll, M. (2004). Coupling factors in macromolecular type-IV secretion machineries. *Curr. Pharm. Des.* 10, 1551–1565. doi: 10.2174/1381612043384817

Grohmann, E., Christie, P. J., Waksman, G., and Backert, S. (2018). Type IV secretion in Gram-negative and Gram-positive bacteria. *Mol. Microbiol.* 107, 455–471. doi: 10.1111/mmi.13896

Guglielmini, J., Quintais, L., Garcillán-Barcia, M. P., de la Cruz, F., and Rocha, E. P. C. (2011). The repertoire of ICE in prokaryotes underscores the unity, diversity, and ubiquity of conjugation. *PLoS Genet.* 7:e1002222. doi: 10.1371/journal.pgen.1002222

Guzmán-Herrador, D. L., and Llosa, M. (2019). The secret life of conjugative relaxases. *Plasmid* 104:102415. doi: 10.1016/j.plasmid.2019.102415

Halary, S., Leigh, J. W., Cheaib, B., Lopez, P., and Bapteste, E. (2010). Network analyses structure genetic diversity in independent genetic worlds. *Proc. Natl. Acad. Sci. U.S.A.* 107, 127–132. doi: 10.1073/pnas.0908978107

Ho, B. T., Basler, M., and Mekalanos, J. J. (2013). Type 6 secretion system-mediated immunity to type 4 secretion system-mediated gene transfer. *Science* 342, 250–253. doi: 10.1126/science.1243745

Ishikawa, T., Rompikuntal, P. K., Lindmark, B., Milton, D. L., and Wai, S. N. (2009). Quorum sensing regulation of the two hcp alleles in *Vibrio cholerae* O1 strains. *PLoS One* 4, e6734. doi: 10.1371/journal.pone.0006734

Journet, L., and Cascales, E. (2016). The type VI secretion system in *Escherichia coli* and related species. *EcoSal Plus* 7, 1–20. doi: 10.1128/ecosalplus.ESP-0009-2015

Kostiuk, B., Unterweger, D., Provenzano, D., and Pukatzki, S. (2017). T6SS intraspecific competition orchestrates *Vibrio cholerae* genotypic diversity. *Int. Microbiol.* 20, 130–137. doi: 10.2436/20.1501.01.294

Kudryashev, M., Wang, R. Y.-R., Brackmann, M., Scherer, S., Maier, T., Baker, D., et al. (2015). Structure of the type VI secretion system contractile sheath. *Cell* 160, 952–962. doi: 10.1016/j.cell.2015.01.037

Leiman, P. G., Basler, M., Ramagopal, U. A., Bonanno, J. B., Sauder, J. M., Pukatzki, S., et al. (2009). Type VI secretion apparatus and phage tail-associated protein complexes share a common evolutionary origin. *Proc. Natl. Acad. Sci. U.S.A.* 106, 4154–4159. doi: 10.1073/pnas.0813360106

Lesic, B., Starkey, M., He, J., Hazan, R., and Rahme, L. G. (2009). Quorum sensing differentially regulates *Pseudomonas aeruginosa* type VI secretion locus I and homologous loci II and III, which are required for pathogenesis. *Microbiology* 155, 2845–2855. doi: 10.1099/mic.0.029082-0

Li, J., Yao, Y., Xu, H. H., Hao, L., Deng, Z., Rajakumar, K., et al. (2015). SecReT6: a web-based resource for type VI secretion systems found in bacteria. *Environ. Microbiol.* 17, 2196–2202. doi: 10.1111/1462-2920.12794

Li, Y. G., Hu, B., and Christie, P. J. (2019). Biological and structural diversity of type IV secretion systems. *Microbiol. Spectr.* 7, 1–15. doi: 10.1128/microbiolspec.PSIB-0012-2018

Liebl, D., Robert-Genthon, M., Job, V., Cogoni, V., and Attrée, I. (2019). Baseplate component TSSK and spatio-temporal assembly of T6SS in *Pseudomonas aeruginosa*. *Front. Microbiol.* 10:1615. doi: 10.3389/fmicb.2019.01615

Lien, Y.-W., and Lai, E.-M. (2017). Type VI secretion effectors: methodologies and biology. *Front. Cell. Infect. Microbiol.* 7:254. doi: 10.3389/fcimb.2017.00254

Lossi, N. S., Dajani, R., Freemont, P., and Filloux, A. (2011). Structure-function analysis of HsiF, a gp25-like component of the type VI secretion system, in *Pseudomonas aeruginosa*. *Microbiology* 157, 3292–3305. doi: 10.1099/mic.0.051987-0

Lossi, N. S., Manoli, E., Förster, A., Dajani, R., Pape, T., Freemont, P., et al. (2013). The HsiB1C1 (TssB-TssC) complex of the *Pseudomonas aeruginosa* type VI secretion system forms a bacteriophage tail sheathlike structure. *J. Biol. Chem.* 288, 7536–7548. doi: 10.1074/jbc.M112.439273

Majerczyk, C., Schneider, E., and Greenberg, E. P. (2016). Quorum sensing control of Type VI secretion factors restricts the proliferation of quorum-sensing mutants. *Elife* 5:e14712. doi: 10.7554/eLife.14712

Matthey, N., Stutzmann, S., Stoudmann, C., Guex, N., Iseli, C., and Blokesch, M. (2019). Neighbor predation linked to natural competence fosters the transfer of large genomic regions in *Vibrio cholerae*. *Elife* 8:e48212. doi: 10.7554/eLife.48212

Morin, N., Santiago, A. E., Ernst, R. K., Guillot, S. J., and Nataro, J. P. (2013). Characterization of the AggR regulon in enteroaggregative *Escherichia coli*. *Infect. Immun.* 81, 122–132. doi: 10.1128/IAI.00676-12

Morin, N., Tirling, C., Ivison, S. M., Kaur, A. P., Nataro, J. P., and Steiner, T. S. (2010). Autoactivation of the AggR regulator of enteroaggregative *Escherichia coli in vitro* and *in vivo*. *FEMS Immunol. Med. Microbiol.* 58, 344–355. doi: 10.1111/j.1574-695X.2009.00645.x

Munson, G. P. (2013). Virulence regulons of enterotoxigenic *Escherichia coli*. *Immunol. Res.* 57, 229–236. doi: 10.1007/s12026-013-8453-4

Nataro, J. P., Yikang, D., Yingkang, D., and Walker, K. (1994). AggR, a transcriptional activator of aggregative adherence fimbria I expression in enteroaggregative *Escherichia coli*. *J. Bacteriol.* 176, 4691–4699. doi: 10.1128/jb.176.15.4691-4699.1994

Nazarov, S., Schneider, J. P., Brackmann, M., Goldie, K. N., Stahlberg, H., and Basler, M. (2018). Cryo-EM reconstruction of Type VI secretion system baseplate and sheath distal end. *EMBO J.* 37:e97103. doi: 10.15252/embj.201797103

Nguyen, V. S., Douzi, B., Durand, E., Roussel, A., Cascales, E., and Cambillau, C. (2018). Towards a complete structural deciphering of Type VI secretion system. *Curr. Opin. Struct. Biol.* 49, 77–84. doi: 10.1016/j.sbi.2018.01.007

Nishi, J., Sheikh, J., Mizuguchi, K., Luisi, B., Burland, V., Boutin, A., et al. (2003). The export of coat protein from enteroaggregative *Escherichia coli* by a specific ATP-binding cassette transporter system. *J. Biol. Chem.* 278, 45680–45689. doi: 10.1074/jbc.M306413200

Park, Y.-J., Lacourse, K. D., Cambillau, C., DiMaio, F., Mougous, J. D., and Veesler, D. (2018). Structure of the type VI secretion system TssK-TssF-TssG baseplate subcomplex revealed by cryo-electron microscopy. *Nat. Commun.* 9:5385. doi: 10.1038/s41467-018-07796-5

Peña, A., and Arechaga, I. (2013). Molecular motors in bacterial secretion. *J. Mol. Microbiol. Biotechnol.* 23, 357–369. doi: 10.1159/000351360

Pena, R. T., Blasco, L., Ambroa, A., González-Pedrajo, B., Fernández-García, L., López, M., et al. (2019). Relationship between quorum sensing and secretion systems. *Front. Microbiol.* 10:1100. doi: 10.3389/fmicb.2019.01100

Piel, D., Bruto, M., James, A., Labreuche, Y., Lambert, C., Janicot, A., et al. (2019). Selection of Vibrio crassostreae relies on a plasmid expressing a type 6 secretion system cytotoxic for host immune cells. *Environ. Microbiol.* doi: 10.1111/1462-2920.14776. [Epub ahead of print].

Pietrosiuk, A., Lenherr, E. D., Falk, S., Bönemann, G., Kopp, J., Zentgraf, H., et al. (2011). Molecular basis for the unique role of the AAA+ chaperone ClpV in type VI protein secretion. *J. Biol. Chem.* 286, 30010–30021. doi: 10.1074/jbc.M111.253377

Planamente, S., Salih, O., Manoli, E., Albesa-Jové, D., Freemont, P. S., and Filloux, A. (2016). TssA forms a gp6-like ring attached to the type VI secretion sheath. *EMBO J.* 35, 1613–1627. doi: 10.15252/embj.201694024

Pukatzki, S., Ma, A. T., Revel, A. T., Sturtevant, D., and Mekalanos, J. J. (2007). Type VI secretion system translocates a phage tail spike-like protein into target cells where it cross-links actin. *Proc. Natl. Acad. Sci. U.S.A.* 104, 15508–15513. doi: 10.1073/pnas.0706532104

Rapisarda, C., Cherrak, Y., Kooger, R., Schmidt, V., Pellarin, R., Logger, L., et al. (2019). *In situ* and high-resolution cryo-EM structure of a bacterial type VI secretion system membrane complex. *EMBO J.* 38:e100886. doi: 10.15252/embj.2018100886

Russell, A. B., Wexler, A. G., Harding, B. N., Whitney, J. C., Bohn, A. J., Goo, Y. A., et al. (2014). A type VI secretion-related pathway in Bacteroidetes mediates interbacterial antagonism. *Cell Host Microbe* 16, 227–236. doi: 10.1016/j.chom.2014.07.007

Salinero-Lanzarote, A., Pacheco-Moreno, A., Domingo-Serrano, L., Durán, D., Ormeño-Orrillo, E., Martínez-Romero, E., et al. (2019). The Type VI secretion system of Rhizobium etli Mim1 has a positive effect in symbiosis. *FEMS Microbiol. Ecol.* 95:fiz054. doi: 10.1093/femsec/fiz054

Santiago, A. E., Ruiz-Perez, F., Jo, N. Y., Vijayakumar, V., Gong, M. Q., and Nataro, J. P. (2014). A large family of antivirulence regulators modulates the effects of transcriptional activators in Gram-negative pathogenic bacteria. *PLoS Pathog.* 10:e1004153. doi: 10.1371/journal.ppat.1004153

Santiago, A. E., Yan, M. B., Tran, M., Wright, N., Luzader, D. H., Kendall, M. M., et al. (2016). A large family of anti-activators accompanying XylS/AraC family regulatory proteins. *Mol. Microbiol.* 101, 314–332. doi: 10.1111/mmi.13392

Schneider, J. P., Nazarov, S., Adaixo, R., Liuzzo, M., Ringel, P. D., Stahlberg, H., et al. (2019). Diverse roles of TssA-like proteins in the assembly of bacterial type VI secretion systems. *EMBO J.* 38:e100825. doi: 10.15252/embj.201800825

Sheikh, J., Czeczulin, J. R., Harrington, S., Hicks, S., Henderson, I. R., Le Bouguénec, C., et al. (2002). A novel dispersin protein in enteroaggregative *Escherichia coli*. *J. Clin. Invest.* 110, 1329–1337. doi: 10.1172/JCI16172

Sheikh, J., Hicks, S., Dall'Agnol, M., Phillips, A. D., and Nataro, J. P. (2001). Roles for Fis and YafK in biofilm formation by enteroaggregative *Escherichia coli*. *Mol. Microbiol.* 41, 983–997. doi: 10.1046/j.1365-2958.2001.02512.x

Shneider, M. M., Buth, S. A., Ho, B. T., Basler, M., Mekalanos, J. J., and Leiman, P. G. (2013). PAAR-repeat proteins sharpen and diversify the type VI secretion system spike. *Nature* 500, 350–353. doi: 10.1038/nature12453

Shyntum, D. Y., Venter, S. N., Moleleki, L. N., Toth, I., and Coutinho, T. A. (2014). Comparative genomics of type VI secretion systems in strains of Pantoea ananatis from different environments. *BMC Genomics* 15:163. doi: 10.1186/1471-2164-15-163

Smillie, C., Garcillán-Barcia, M. P., Francia, M. V., Rocha, E. P. C., and de la Cruz, F. (2010). Mobility of plasmids. *Microbiol. Mol. Biol. Rev.* 74, 434–452. doi: 10.1128/MMBR.00020-10

Thomas, J., Watve, S. S., Ratcliff, W. C., and Hammer, B. K. (2017). Horizontal gene transfer of functional type VI killing genes by natural transformation. *MBio* 8::e00654–17. doi: 10.1128/mBio.00654-17

Uchida, K., Leiman, P. G., Arisaka, F., and Kanamaru, S. (2014). Structure and properties of the C-terminal β-helical domain of VgrG protein from *Escherichia coli* O157. *J. Biochem.* 155, 173–182. doi: 10.1093/jb/mvt109

Veening, J.-W., and Blokesch, M. (2017). Interbacterial predation as a strategy for DNA acquisition in naturally competent bacteria. *Nat. Rev. Microbiol.* 15, 621–629. doi: 10.1038/nrmicro.2017.66

Veesler, D., and Cambillau, C. (2011). A common evolutionary origin for tailed-bacteriophage functional modules and bacterial machineries. *Microbiol. Mol. Biol. Rev.* 75, 423–33. doi: 10.1128/MMBR.00014-11

Verster, A. J., Ross, B. D., Radey, M. C., Bao, Y., Goodman, A. L., Mougous, J. D., et al. (2017). The landscape of type VI secretion across human gut microbiomes reveals its role in community composition. *Cell Host Microbe* 22, 411–9.e4. doi: 10.1016/j.chom.2017.08.010

Wang, J., Brackmann, M., Castaño-Díez, D., Kudryashev, M., Goldie, K. N., Maier, T., et al. (2017). Cryo-EM structure of the extended type VI secretion system sheath-tube complex. *Nat. Microbiol.* 2, 1507–1512. doi: 10.1038/s41564-017-0020-7

Wang, J., Brodmann, M., and Basler, M. (2019). Assembly and subcellular localization of bacterial type VI secretion systems. *Annu. Rev. Microbiol.* 73, 621–638. doi: 10.1146/annurev-micro-020518-115420

Weber, B. S., Ly, P. M., Irwin, J. N., Pukatzki, S., and Feldman, M. F. (2015). A multidrug resistance plasmid contains the molecular switch for type VI secretion in Acinetobacter baumannii. *Proc. Natl. Acad. Sci. U.S.A.* 112, 9442–9447. doi: 10.1073/pnas.1502966112

Weber, B. S., Miyata, S. T., Iwashkiw, J. A., Mortensen, B. L., Skaar, E. P., Pukatzki, S., et al. (2013). Genomic and functional analysis of the type VI secretion system in Acinetobacter. *PLoS ONE* 8:e55142. doi: 10.1371/journal.pone.0055142

Yang, Z., Zhou, X., Ma, Y., Zhou, M., Waldor, M. K., Zhang, Y., et al. (2018). Serine/threonine kinase PpkA coordinates the interplay between T6SS2 activation and quorum sensing in the marine pathogen Vibrio alginolyticus. *Environ. Microbiol.* 20, 903–919. doi: 10.1111/1462-2920.14039

Yasir, M., Icke, C., Abdelwahab, R., Haycocks, J. R., Godfrey, R. E., Sazinas, P., et al. (2019). Organization and architecture of AggR-dependent promoters from enteroaggregative *Escherichia coli*. *Mol. Microbiol.* 111, 534–551. doi: 10.1111/mmi.14172

Zechner, E. L., Moncalián, G., and de la Cruz, F. (2017). Relaxases and plasmid transfer in gram-negative bacteria. *Curr. Top. Microbiol. Immunol.* 413, 93–113. doi: 10.1007/978-3-319-75241-9_4

Zhang, L., Xu, J., Xu, J., Zhang, H., He, L., and Feng, J. (2014). TssB is essential for virulence and required for type VI secretion system in *Ralstonia solanacearum*. *Microb. Pathog.* 74, 1–7. doi: 10.1016/j.micpath.2014.06.006

Zheng, J., Shin, O. S., Cameron, D. E., and Mekalanos, J. J. (2010). Quorum sensing and a global regulator TsrA control expression of type VI secretion and virulence in *Vibrio cholerae*. *Proc. Natl. Acad. Sci. U.S.A.* 107, 21128–21133. doi: 10.1073/pnas.1014998107

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Peñil-Celis and Garcillán-Barcia. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 28 February 2020
doi: 10.3389/fmolb.2020.00025

Naringenin Inhibition of the *Pseudomonas aeruginosa* Quorum Sensing Response Is Based on Its Time-Dependent Competition With *N*-(3-Oxo-dodecanoyl)-L-homoserine Lactone for LasR Binding

Sara Hernando-Amado[1†], Manuel Alcalde-Rico[1,2,3†], Teresa Gil-Gil[1], José R. Valverde[4] and José L. Martínez[1]*

[1] Departamento de Biotecnología Microbiana, Centro Nacional de Biotecnología, Consejo Superior de Investigaciones Científicas, Madrid, Spain, [2] Laboratorio de Genética e Inmunología Molecular, Instituto de Biología, Facultad de Ciencias, Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile, [3] Millennium Nucleus for Collaborative Research on Bacterial Resistance, Valparaíso, Chile, [4] Servicio de Computación Científica, Centro Nacional de Biotecnología, Consejo Superior de Investigaciones Científicas, Madrid, Spain

Bacterial quorum sensing (QS) is a cell-to-cell communication system that governs the expression of a large set of genes involved in bacterial–host interactions, including the production of virulence factors. Conversely, the hosts can produce anti-QS compounds to impair virulence of bacterial pathogens. One of these inhibitors is the plant flavonoid naringenin, which impairs the production of QS-regulated *Pseudomonas aeruginosa* virulence factors. In the present work, we analyze the molecular basis for such inhibition. Our data indicate that naringenin produces its effect by directly binding the QS regulator LasR, hence competing with its physiological activator, *N*-(3-oxo-dodecanoyl)-L-homoserine lactone (3OC12-HSL). The *in vitro* analysis of LasR binding to its cognate target DNA showed that the capacity of naringenin to outcompete 3OC12-HSL, when the latter is previously bound to LasR, is low. By using an *E. coli* LasR-based biosensor strain, which does not produce 3OC12-HSL, we determined that the inhibition of LasR is more efficient when naringenin binds to nascent LasR than when this regulator is already activated through 3OC12-HSL binding. According to these findings, at early exponential growth phase, when the amount of 3OC12-HSL is low, naringenin should proficiently inhibit the *P. aeruginosa* QS response, whereas at later stages of growth, once 3OC12-HSL concentration reaches a threshold enough for binding LasR, naringenin would not efficiently inhibit the QS response. To test this hypothesis, we analyze the potential effect of naringenin over the QS response by adding naringenin to *P. aeruginosa* cultures at either time zero (early inhibition) or at stationary growth phase (late inhibition). In early inhibitory conditions, naringenin inhibited the expression of QS-regulated genes, as well as the production of the QS-regulated virulence factors, pyocyanin and elastase. Nevertheless, in late inhibitory conditions, the *P. aeruginosa* QS response was not

OPEN ACCESS

Edited by:
Manuel Espinosa, Center for Biological Research (CSIC), Spain

Reviewed by:
Gloria Soberón-Chávez, National Autonomous University of Mexico, Mexico
Kurt Henry Piepenbrink, University of Nebraska–Lincoln, United States
Manuel Espinosa-Urgel, Experimental Station of Zaidín (EEZ), Spain

***Correspondence:**
José L. Martínez
jlmtnez@cnb.csic.es

[†] These authors have contributed equally to this work

Specialty section:
This article was submitted to Molecular Recognition, a section of the journal Frontiers in Molecular Biosciences

Received: *19 November 2019*
Accepted: *07 February 2020*
Published: *28 February 2020*

Citation:
Hernando-Amado S, Alcalde-Rico M, Gil-Gil T, Valverde JR and Martínez JL (2020) Naringenin Inhibition of the Pseudomonas aeruginosa Quorum Sensing Response Is Based on Its Time-Dependent Competition With N-(3-Oxo-dodecanoyl)-L-homoserine Lactone for LasR Binding. Front. Mol. Biosci. 7:25. doi: 10.3389/fmolb.2020.00025

inhibited by naringenin. Therefore, this time-dependent inhibition may compromise the efficiency of this flavonoid, which will be effective just when used against bacterial populations presenting low cellular densities, and highlight the importance of searching for QS inhibitors whose mechanism of action does not depend on the QS status of the population.

Keywords: quorum sensing, virulence inhibitor, LasR, naringenin, *Pseudomonas aeruginosa*

INTRODUCTION

Quorum sensing (QS) is a cell-to-cell communication system that enables to coordinate a global response of a bacterial population when a certain cell density is reached (Nealson et al., 1970; Fuqua et al., 1994). This response is based on the production of intercellular communication molecules, which function as autoinducer signals (AIs) and are constantly synthesized and released by the bacterial population. The whole population perceives these AIs through their binding to their cognate receptors, which are then activated to globally regulate the QS response. The progressive accumulation of the regulator-AIs complexes promotes the AIs synthesis, thus providing a positive feedback to the system, and the subsequent modification of the expression of a large set of QS-regulated genes (Williams et al., 2007). Consequently, while at low cell densities the concentration of the AIs is very low, at higher cell densities, the signal concentration exceeds a given threshold and the QS response is exponentially activated.

The molecular mechanisms underlying the QS response have been deeply studied in *Pseudomonas aeruginosa*. In this opportunistic pathogen, the QS regulatory network is based on the synthesis and perception of the AIs *N*-acyl-homoserine lactones (AHLs) and 2-alkyl-4-quinolones (AQs) (Williams et al., 2000, 2007). In particular, *N*-(3-oxo-dodecanoyl)-L-homoserine lactone (3OC12-HSL) and *N*-butanoyl-L-homoserine lactone (C4-HSL) AHL signals are synthesized by the AHL synthases LasI and RhlI, respectively. The *lasR* and *rhlR* genes, which encode the receptors of 3OC12-HSL and C4-HSL respectively, are located upstream the *lasI* and *rhlI* genes. These two LuxR-type transcriptional regulators contain both, a DNA binding domain and a ligand binding domain (Lamb et al., 2003; Bottomley et al., 2007) and control the expression of hundreds of genes in *P. aeruginosa* (Wagner et al., 2003; Schuster and Greenberg, 2007). It is worth noting that several biological processes with relevance for *P. aeruginosa* infection, such as motility, production of virulence factors and biofilm formation, are regulated by the QS response (Parsek and Greenberg, 1999; Erickson et al., 2002; Bjarnsholt and Givskov, 2008). Indeed, it has been described that the disruption of the QS response reduces *P. aeruginosa* virulence (Tan et al., 1999; Diggle et al., 2007) and impairs biofilm formation (Wagner and Iglewski, 2008). This indicates that QS is a fundamental element in the success of this opportunistic pathogen for colonizing and infecting its hosts, from plants to humans. For those reasons, the QS system transcriptional regulators, such as LasR, have become attractive targets to be inhibited (Williams et al., 2000; Hurley et al., 2012; Kalia et al., 2017).

Plants have evolved different strategies to deal with bacterial infections, being inhibition of the QS response one of their potential defense mechanisms (Diggle et al., 2007). Therefore, plants have been widely screened in the search of compounds able to interfere with the QS response of bacterial pathogens (Rudrappa and Bais, 2008; Smyth et al., 2010; Vandeputte et al., 2011; Rasamiravaka et al., 2013; Corral-Lugo et al., 2016; Fernandez et al., 2018), particularly for those, as *P. aeruginosa*, able of infecting both humans and plants (Rahme et al., 1995). Among those potential inhibitors, naringenin seems to be particularly interesting (Salehi et al., 2019). This plant-produced flavonoid inhibits the inflammatory response (Jin et al., 2017). In addition, studies using animal models have shown that naringenin might have therapeutic potential for the treatment of different inflammation-related diseases, including sepsis and endotoxic shock, hepatitis, pulmonary fibrosis, radiation-induced lung injury, atherosclerosis, obesity, diabetes, or cancer (Zeng et al., 2018). Besides these effects in the host response, it has also been described that naringenin reduces the production of AHLs in *P. aeruginosa*, as well as the expression of QS-regulated genes, thus inhibiting the production of QS-regulated virulence factors, such as pyocyanin or elastase (Vandeputte et al., 2011; Rasamiravaka et al., 2013). Nevertheless, the mechanisms underlying naringenin-mediated inhibition of the QS response are still unknown. Herein, we analyze the interaction of naringenin with LasR, a master regulator of QS in *P. aeruginosa* and find out that this flavonoid competes with 3OC12-HSL for the binding of the nascent LasR protein. However, the inhibitory capacity of naringenin is limited when 3OC12-HSL has already bound the regulator. These results shed light on the mechanisms of QS inhibition based on targeting the LasR regulator in *P. aeruginosa* populations and may help in the design and development of efficient QS-inhibitors.

MATERIALS AND METHODS

Bacterial Strains, Plasmids, and Growth Conditions

Bacterial strains and plasmids used during this work are listed in **Table 1**. Bacteria were routinely grown at 37°C in Lysogeny Broth, Lennox (LB) (Pronadisa) at 250 rpm. For the culture of the LasR-based reporter strain (**Table 1**), tetracycline 5 μg/mL was added to LB medium, while ampicillin 100 μg/mL was added in the case of the SHA011 LasR producing *E. coli* strain. When necessary, naringenin and/or 3OC12-HSL were added to *E. coli*

TABLE 1 | Bacterial strains and plasmids used in this work.

Bacterial strains		**Description**	**Reference/origin**
E. coli strains	BL21(DE3)pLysS	*hsdS gal (lcIts857 ind1 S am7 nin5 lacUV5-T7 gene 1)*	Sambrook and Russell, 2001
	SHA010	BL21(DE3)pLysS strain carrying pGEX6p	This work
	SHA011	BL21(DE3)pLysS strain carrying pSHA-LasR	This work
	JM109 (pSB1142)	LasR-based reporter strain	Miguel Cámara laboratory
P. aeruginosa strains	PAO1	Wild-type *P. aeruginosa* strain	Laboratory collection
Plasmids		**Description**	**Reference/origin**
pGEMT-Easy Vector		Amp^R, commercial vector	Promega
pGEMT-*lasR*		pGEMT-Easy Vector carrying *lasR* gene from *P. aeruginosa* PAO1 strain.	This work
pGEX6p		Expression vector, Amp^R	GE Healthcare
pSHA-LasR		pGEX6p vector carrying *lasR* gene from *P. aeruginosa* PAO1 strain.	This work
pSB1142		Reporter plasmid carrying the *lasR* gene and the *luxCDABE* operon under the control of the *lasI* promoter.	Miguel Cámara Laboratory

and *P. aeruginosa* cultures. Both compounds were purchased from Sigma Aldrich and dissolved in dimethyl sulfoxide (DMSO).

Overexpression and Purification of LasR

The *lasR* open reading frame was amplified with the Expand Long Template System (Roche). PCR was carried out using the primers lasR-pGEX.fw and lasR-pGEX.rv (**Table 2**) with the following conditions: 2 min at 94°C, 35 cycles of 15 s at 94°C, 30 s at 55°C, 1 min at 68°C and 7 min of final elongation at 68°C. The DNA fragments amplified were purified following the supplier's instructions with the QIAquick Purification PCR Kit (Qiagen) and cloned into pGEMT-Easy vector. The cloned *lasR* was sequenced to verify that no mutations were introduced during PCR process. This plasmid (pGEMT-*lasR*), containing the gene amplified from lasR-pGEX.fw/lasR-pGEX.rv, was digested with *Eco*RI-*Bam*HI and the DNA product was cloned into the polylinker site of pGEX6p, immediately after the GST gene (pSHA-LasR). The resulting plasmid was introduced by transformation in *E. coli* BL21. The resulting clone, which encodes a GST–LasR fusion protein, was dubbed *E. coli* SHA011, and the respective one containing pGEX6p without LasR, SHA010 (**Table 1**). For the induction of LasR production, cells were grown with 1 mM IPTG at 17°C overnight. LasR was purified as a GST-fusion protein according to the manufacturer's instructions for the GST gene fusion system (Amersham Pharmacia Biotech, Europe, GmbH) using an immobilized Glutathione Sepharose column. After elution, the GST affinity tag was removed in solution with PreScission protease (GE Healthcare). Separation of LasR and the GST was done using a Glutathione Sepharose column. After dialysis against TBS (20 mM Tris-HCl, 200 mM NaCl, pH8), the purified native protein appeared as a single band on SDS-PAGE. The verification that the purified protein was LasR was performed by matrix-assisted laser desorption ionization-time of flight (MALDI-TOF) mass spectrometry at the Proteomics Service of the Centro Nacional de Biotecnología.

In silico Prediction of the Interaction of Naringenin With LasR

The LasR structures available on PDB were downloaded and confirmed incomplete. To avoid missing any potential interaction of the few missing amino acids, a full model of the structure of LasR was built using homology modeling with RaptorX (Kallberg et al., 2012) and used for subsequent docking. The structure of naringenin was optimized using Semiempirical Quantum Mechanics with OpenMopac (Stewart, 2008) using PM7 in an

TABLE 2 | Primers used in this study.

Primer	**Sequence**	**Utilization**
rplU.fw	5′-CGCAGTGATTGTTACCGGTG-3′	Real time PCR for *rplU* and for checking DNA in RNA samples
rplU.rv	5′-AGGCCTGAATGCCGGTGATC-3′	
lasA.fw	5′-ATGGACCAGATCCAGGTGAG-3′	Real time PCR for *lasA*
lasA.rv	5′-CGTTGTCGTAGTTGCTGGTG-3′	
lasB.fw	5′-ATCGGCAAGTACACCTACGG-3′	Real time PCR for *lasB*
lasB.rv	5′-ACCAGTCCCGGTACAGTTTG-3′	
lasI.fw	5′-TGCTCTGATCTTTTCGGA-3′	Amplication and labeling of *lasI**
lasI.rv	5′-CGATCATCTTCACTTCCT-3′	probe used in EMSA
lasR-pGEX.fw	5′-CCGGATCCATGGCCTTGGTTGACGGTTT-3′	LasR amplification for cloning in pGEX6p expression vector
lasR-pGEX.rv	5′-CCGAATTCTCAGAGAGTAATAAGACCCA-3′	

implicit water moiety and analyzed to assign partial charges prior to docking.

An initial docking step was carried out using PatchDock (Schneidman-Duhovny et al., 2005) with naringenin against the whole surface of the full-sequence LasR model to obtain 100 poses. After inspection, the best 20 were selected for further analysis and the complexes were subjected to Molecular Mechanics optimization using UCSF Chimera (Pettersen et al., 2004) with the AMBER ff14SB (Maier et al., 2015) force field allowing up to 1000 Newton-Raphson minimization steps and 100 conjugate-gradient steps to account for full flexibility in the interaction of LasR with naringenin.

Affinity was estimated using X-Score (Wang et al., 2002) and DSX (Neudert and Klebe, 2011), and UCSF Chimera was used to calculate the H-bonds, contacts, and clashes. Pose selection was based on docking scores, affinity estimates, counts of contacts and H-bonds, and absence of clashes. PDB structure 3ix3 was used to calculate the predicted affinity of LasR and 3OC12-HSL for comparison with naringenin.

The best interactions were selected to identify the preferential binding site, and this was used to perform a higher quality, restricted docking using AutoDock/Vina (Morris et al., 2009). The resulting poses were optimized and analyzed as described.

Electrophoretic Mobility Shift Assays

The complete intergenic region upstream of *lasI* was amplified by using the fluorescein labeled oligonucleotides lasI.fw and lasI.rv (**Table 2**). The DNA (90 ng) was incubated with 20 μg of protein extracts of *E. coli* SHA011 cells or with 5 μg of purified LasR in the binding buffer (20 mM Tris-HCl pH 7.4, 1 mM EDTA, 150 mM KCl, 1 mM DTT, 10 μM 3OC12-HSL or DMSO, 100 μg/mL BSA, 50 ng/μL Poly (dIdC) and 10% glycerol). The inhibition assay was carried out by adding naringenin (10, 50, and 500 μM), or DMSO, to the binding reaction buffer during 30 min at room temperature. DNA-protein complexes were subjected to electrophoresis on 6% native polyacrylamide gels and visualized by using a Typhoon phosphorimager.

In vivo Analysis of the Competition Between 3OC12-HSL and Naringenin Using a LasR-Based Reporter Strain

In vivo analysis of the potential competition between naringenin and 3OC12-HSL for their binding to LasR was performed using the *E. coli* JM109 (pSB1142) LasR-based reporter strain carrying the plasmid pSB1142 (**Table 1**). This plasmid contains the gene encoding LasR and the *luxCDABE* operon under control of the LasR-dependent *lasI* promoter. The reporter strain produces LasR, but light is emitted only when the protein is in its active form, in presence of the exogenously added effector 3OC12-HSL. Only when the effector is bound to LasR, the regulator is active, being able to recognize the *lasI* promoter, and activates the LasR-dependent expression of the *luxCDABE* operon. This system hence allows to get direct information on the *in vivo* activation of LasR through the binding of its effector and, consequently on the *in vivo* inhibition of such binding by potential 3OC12-HSL competitors as naringenin. To perform the assay, overnight biosensor cultures were diluted to 0.01 OD600 in LB into Flat white 96-well plates with optical bottom (Thermo Scientific Nunc) adding then the molecule(s) of interest. The assay was carried out by growing the LasR-based reporter strain in the presence of 1 μM of 3OC12-HSL, 1 μM of naringenin or both (competition assay). The capacity of each molecule to displace the other, already bound to LasR, was analyzed by adding one of the molecules at time zero and the other one with a delay of 2 h. Bioluminescence and OD600 were monitored at 37°C using a TECAN Infinite M200 plate reader. Luminescence was represented as relative light units normalized by growth (OD_{600}).

Gene Expression Analysis by Real Time PCR

Pseudomonas aeruginosa cells were grown in LB containing either 1.5 mM of naringenin or its solvent, DMSO. Independent cultures in which naringenin was added either at time 0 or 4 h after inoculation (early stationary phase) were grown. Cells were harvested, after 4 h of naringenin treatment, by centrifuging the cultures at 7000 rpm for 20 min at 4°C. The cell pellets were stored at −80°C and total RNA was extracted using the RNeasy minikit (Qiagen), according to the manufacturer's instructions. Samples were treated with Turbo DNA-free (Ambion) to eliminate any remaining DNA. RNA integrity was verified on a 1% agarose gel, and the absence of DNA was verified by PCR using primers rplU.fw and rplU.rv (**Table 2**). cDNA was synthesized from RNA using a High Capacity cDNA reverse transcription kit (Applied Biosystems). Real time PCR was performed using a Power SYBR green kit (Applied Biosystems) as indicated by the manufacturer. The primers (**Table 2**) were used at a concentration of 400 nM. 50 ng of cDNA were used in each reaction. A denaturation step (95°C for 10 min), followed by 40 cycles of amplification (95°C for 15 s and 60°C for 1 min) was followed for amplification and quantification. Primers rplU.fw and rplU.rv were used to amplify the housekeeping gene *rplU*. All the primers were designed using Primer3 Input software; their specificity was tested by BLAST alignment against *P. aeruginosa* PAO1 genome from Pseudomonas Genome Database (Winsor et al., 2016) and their efficiency was analyzed by real time PCR using serial dilutions of cDNA. Differences in the relative amounts of mRNA for the different genes were determined according to the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001). In all cases, the mean values for relative mRNA expression obtained in three independent triplicate experiments were considered.

Elastase Activity and Pyocyanin Production

Pseudomonas aeruginosa cells were cultured in 20 ml of LB, at 37°C, containing either 4 mM of naringenin or its solvent, DMSO, as previously described (Vandeputte et al., 2011), with some modifications. Naringenin or DMSO was added to independent cultures either at time 0 or 4 h after inoculation (early stationary phase). Eight hours after inoculation (late stationary phase), 2 ml samples were collected from each culture and centrifuged for 10 min at 7000 rpm, and the

supernatants were filtered using 0.2 μm filters (Whatman). Pyocyanin production was determined by measuring the OD_{690} of filtered supernatants. The elastase assay was adapted from Kessler and Safrin (2014); 1 ml of Congo Red elastin (Sigma-Aldrich) was added to 100 μl of each sample, and the mixture was incubated at 37°C and 250 rpm for 2 h. Subsequently, samples were centrifuged (10 min, 7000 rpm) and the OD_{495} of the filtered supernatants was determined. Three replicates of each condition were included in the analyses.

RESULTS

The Transcriptional Regulatory Activity of LasR Depends on 3OC12-HSL Presence During Protein Expression

To analyze the putative capacity of naringenin to inhibit the binding of LasR to the *lasI* promoter, whose activity is directly dependent of LasR binding, we started by establishing the purification conditions required to obtain an active LasR protein. For that, the open reading frame of *lasR* was cloned into the inducible expression vector pGEX6p (**Table 1**), next to a glutathione *S*-transferase (GST) tag, therefore encoding a LasR-GST fusion protein. The resulting clones, containing either pGEX6p or pGEX6p-LasR (pSHA-LasR plasmid, see **Table 1**) were dubbed SHA010 and SHA011, respectively (**Table 1**). Although previous works had suggested that LasR purification requires the presence of its cognate signal effector 3OC12-HSL to obtain a fully active protein (Kiratisin et al., 2002; Schuster et al., 2004; Bottomley et al., 2007), biochemical studies would be far simpler using a purified protein without the bound activator. Therefore, we firstly expressed LasR from *E. coli* SHA011, either in presence or in absence of 20 μM of 3OC12-HSL. Protein extracts obtained in each of the conditions were used for performing electrophoretic mobility shift assays (EMSAs) using the complete intergenic region upstream of *lasI,* containing the LasR-binding site (Schuster et al., 2004), as the DNA probe (hereafter dubbed as *lasI**). As shown in **Figure 1**, only the protein extracts obtained in the presence of 3OC12-HSL presented the capacity to bind *lasI**. Moreover, addition of 3OC12-HSL to the binding reaction containing the inactive protein extracts, obtained from cultures grown without 3OC12-HSL, did not rescue the DNA-binding capability of LasR. This finding is in agreement with previous information regarding the conditions required to purify LasR and reasserts the requirement of 3OC12-HSL during LasR production to stabilize the conformation of the nascent protein and to permit its dimerization (Schuster et al., 2004; Bottomley et al., 2007), hence preventing the formation of insoluble aggregates and the subsequent degradation by endogenous proteases (Zhu and Winans, 2001; Kiratisin et al., 2002). Further, our results support that the 3OC12-HSL effector is unable to turn on the activity of an inactive LasR protein. This information is in agreement with those reported in a recent publication, which states that, as other LuxR-type regulators (Zhu and Winans, 2001; Oinuma and Greenberg, 2011), LasR is only soluble when ligand is present

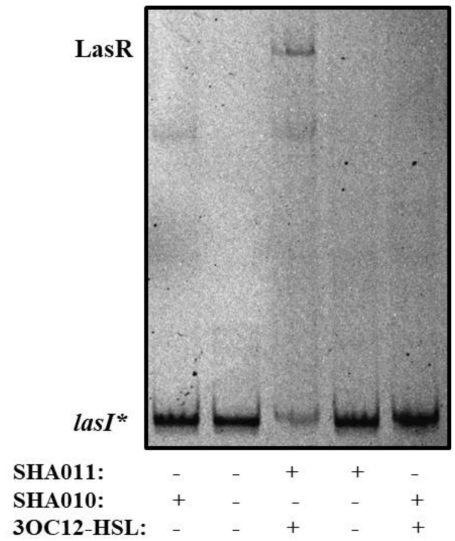


FIGURE 1 | LasR DNA-binding capacity requires 3OC12-HSL during protein expression. *E. coli* SHA011 LasR producing strain was grown in the presence or absence of 20 μM 3OC12-HSL. As a control, *E. coli* SHA010, LasR non-producing strain was grown in same conditions. Protein extracts were tested for their capacity to bind to the *lasI** probe by EMSA assays in the presence of 10 μM 3OC12-HSL. As shown, the presence of 3OC12-HSL in the culture is essential for purifying an active form of LasR (able to bind to the *lasI** probe) from the *E. coli* SHA011 LasR producing strain. A non-specific retarded band was also observed in extracts obtained either from the control strain that does not produce LasR (SHA010) or from the SHA011 LasR-producing strain.

(Mccready et al., 2019). Further, it has been also shown that extensive dialysis of active LasR against signal-free buffer does not allow to remove such signal from the protein, hence impeding to have purified LasR without any bound effector site (Schuster et al., 2004). Taking into account this situation, LasR purified in the presence of its effector, which forms the LasR-3OC12-HSL complex, was used in further biochemical assays.

Naringenin Inhibits the Binding of LasR to the *lasI* Promoter

The structural basis of the molecular interactions of 3OC12-HSL with LasR have been studied by analyzing the co-crystal structure of a C-terminally truncated form of LasR bound to this effector (Bottomley et al., 2007). Previous efforts to get soluble forms of LasR were not successful in line with the reported requirement of 3OC12-HSL during LasR production for stabilizing the protein (Schuster et al., 2004). Further work allowed a better delimitation of the structure of the effector binding site at 1.4 Å of resolution (Zou and Nair, 2009). To address whether or not naringenin could be a competitor of 3OC12-HSL, we analyzed *in silico* the capability of LasR to accommodate the flavonoid in its binding pocket.

To explore this issue, we obtained a full-sequence homology model of LasR using Raptor-X (see section "Materials and Methods") to avoid ignoring any potential interaction between LasR and naringenin, and docked naringenin to the model.

Docking with PatchDock (see section "Materials and Methods") revealed a number of potential binding sites for naringenin all over the surface of LasR. To evaluate their relative preference, we optimized the best poses using Molecular Mechanics to allow full flexibility on both molecules (see section "Materials and Methods"). Affinity estimations revealed similar rankings with both DSX and X-Score, with the first five poses consistently showing a significantly larger score than the rest and binding the active pocket of LasR. Docking with AutoDock/Vina specifically to the identified preferred binding site, resulted in higher quality conformers (**Figure 2**) according to the subsequent analysis. The affinity estimated for the best naringenin binding pose to LasR is similar to the affinity estimated from the structure of 3OC12-HSL bound to LasR, determined experimentally by X-ray crystallography (−9.30 kcal/mol for naringenin and −9.25 kcal/mol for 3OC12-HSL, according to X-Score). Despite seemingly showing some freedom to position itself in the binding pocket, our analysis revealed that, in any case, naringenin would be unable to interact with all the amino acids required for the formation of an oligomerizable LasR monomer.

To verify this prediction, we purified LasR from *E. coli* SHA011 grown in the presence of 20 μM 3OC12-HSL, condition where LasR presented an active conformation when protein extracts were analyzed (**Figure 1**). The protein was purified using a GST-affinity column and the LasR protein was collected by removing the GST-tag as described in Section "Materials and Methods." The purified non-tagged protein appears as a single band when analyzed by polyacrylamide gel electrophoresis (PAGE) and was identified as LasR by MALDI-TOF mass spectrometry. Afterward, we measured the capacity of purified LasR for binding to *lasI** by EMSA and determined the capacity of increasing concentrations of naringenin to inhibit this binding. As shown in **Figure 3**, LasR purified from *E. coli* SHA011 grown in presence of 3OC12-HSL is able to bind its cognate DNA operator. Moreover, the addition of 10 μM 3OC12-HSL to the binding reaction buffer does not increase LasR-DNA binding capacity, supporting that the protein is fully active (**Figure 3**). To analyze the capacity of naringenin to displace 3OC12-HSL from LasR, we added increasing naringenin concentrations, from 10 through 50 and reaching 500 μM, to the EMSA binding reaction. As **Figure 3** shows, the addition of 50 μM naringenin to the binding buffer reduces the amount of the DNA-LasR complex indicating that the flavonoid partially displaces 3OC12-HSL from LasR. Nevertheless, concentrations of naringenin as high as 500 μM do not further reduce the amount of LasR-DNA complex. This finding indicates that the ability of naringenin to outcompete 3OC12-HSL is limited when the latter is previously bound to the LasR protein, a feature in agreement with the *in silico* analysis, which predicts that the affinities of LasR for naringenin and O3C12-HSL are similar.

Perception of 3OC12-HSL by LasR Is Inhibited by Naringenin

Our *in vitro* results indicate that, once 3OC12-HSL is bound to LasR, naringenin competes just partially for such binding

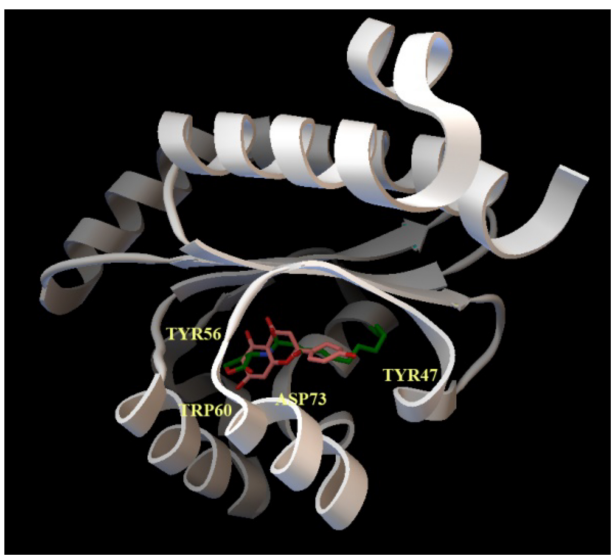


FIGURE 2 | Docking of naringenin and 3OC12-HSL to LasR. The cartoon representation shows the best binding conformation for the analyzed ligands, naringenin (orange) and 3OC12-HSL (green), in the ligand binding pocket of the protein. In yellow, the aminoacid residues Tyr56, Trp60 and Asp73, which interact with the polar homoserine lactone head group of the autoinducer and with the hydrocarbon chain of the autoinducer (Tyr47), are shown. The prediction model supports that naringenin may be accommodated in the binding pocket where 3OC12-HSL binds, suggesting a possible competition of both molecules for LasR binding.

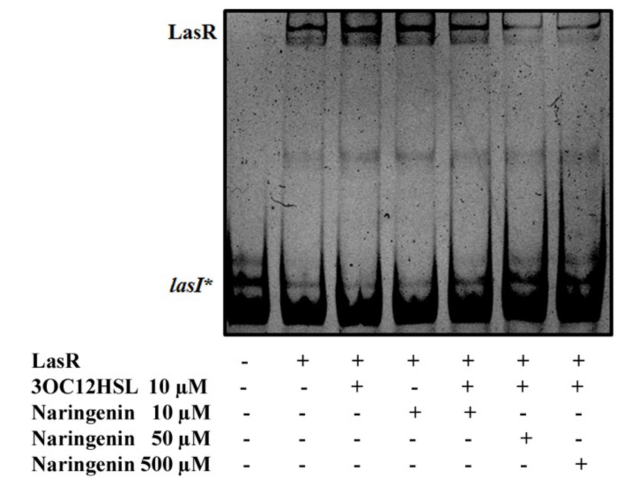


FIGURE 3 | Partial inhibition of LasR binding activity by naringenin. The EMSA assays were carried out using a purified active protein obtained from SHA011 strain grown in the presence of 3OC12-HSL. The capacity of naringenin to displace 3OC12-HSL from LasR was analyzed by adding increasing concentrations of the flavonoid (10, 50, 500 μM) to the binding reaction. The figure shows that neither 3OC12-HSL (10 μM) nor naringenin (10 μM) addition to the binding buffer alters the LasR-DNA binding complex. However, increased concentrations of naringenin (50 and 500 μM) in the binding buffer were able to partially inhibit the LasR-DNA interaction, even in the presence of 3OC12-HSL (10 μM), evidencing that the flavonoid may partially displace the 3OC12-HSL from LasR.

(**Figure 3**), being unable to completely displace the AHL. To analyze the capability of both molecules to displace each other when the first molecule is bound to LasR, the LasR-based reporter strain, *E. coli* JM109 (pSB1142) (**Table 1**), was used. This is a 3OC12-HSL non-producer strain that harbors the pSB1142 plasmid containing the *lasR* gene, which encodes the QS response regulator LasR, and a *luxCDABE* reporter operon under the control of the LasR-dependent *lasI* promoter. This biosensor strain produces LasR, but the protein is active just in presence of 3OC12-HSL. Only when this effector in bound to LasR, the protein binds to the *lasI* promoter and triggers the transcription of the *lux* system. Therefore, luminescence emitted from LasR-based biosensor strain provides a direct *in vivo* quantification of LasR activation by its effector and, conversely, when a potential inhibitor as naringenin is added, provides a direct measurement of the competition for LasR of the activator and the inhibitor. To note here that, since LasR cannot be purified in its inactive form, in the absence of an effector, biosensors as the one herein used are regularly utilized for studying the interaction of LasR with potential effectors (Mccready et al., 2019).

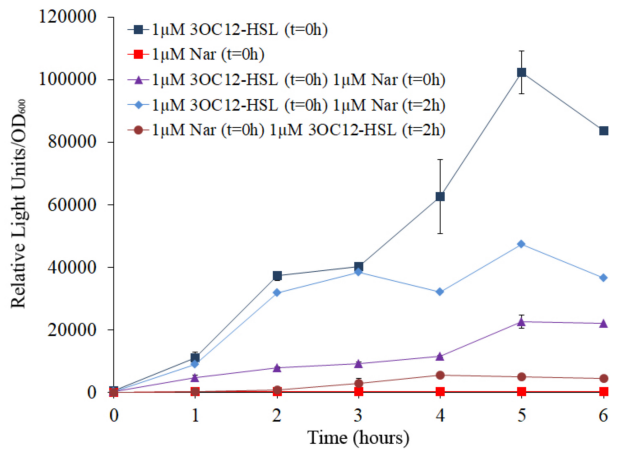


FIGURE 4 | Analysis of naringenin and 3OC12-HSL competition for the binding of LasR using a biosensor strain of *E. coli*. The *E. coli* JM109 (pSB1142) LasR-based reporter strain was used to analyze the capacity of naringenin and 3OC12-HSL to displace each other when the first molecule has already bound the LasR regulator. The luminescence produced by *E. coli* JM109 (pSB1142) was recorded and it is represented as relative light units normalized by growth (OD_{600}). 3OC12-HSL or naringenin were added in combination or separately, at different times. No differences in growth were observed between the analyzed conditions. As shown, the highest or the lowest luminescence were emitted in presence of 1 μM of 3OC12-HSL or 1 μM naringenin, respectively, when were independently added at time zero (t = 0 h). Moreover, intermediate luminescence was observed when the two compounds were added at time zero, demonstrating that there is a competition between them for LasR binding. Addition of 3OC12-HSL after 2 h of incubation with naringenin (added at time = 0 h) was unable to increase LasR activity, while addition of naringenin after 2 h of incubation with 3OC12-HSL (added at time = 0 h) was enough to partially decrease the LasR-dependent luminescence emitted by *E. coli* JM109 (pSB1142). Altogether, these results suggest a more efficient displacement of 3OC12-HSL by naringenin compared with the displacement of naringenin by 3OC12-HSL. Error bars represent standard error of three independent replicates.

As could be expected, when *E. coli* JM109 (pSB1142) cells were grown in the presence of 3OC12-HSL, induction of the *luxCDABE* reporter was observed, while no induction was observed when *E. coli* JM109 (pSB1142) cells were grown in the presence of naringenin (**Figure 4**). It is important to remark that the concentrations of naringenin used in the assay do not alter the growth of the reporter *E. coli* strain. When, *E. coli* JM109 (pSB1142) cells grown in presence of equimolar concentrations of naringenin and 3OC12-HSL, showed a relevant reduction of the expression of the reporter operon, compared to the luminescence levels reached when cells were grown with 3OC12-HSL alone. This result confirms that naringenin can inhibit the activity of LasR *via* its competition with 3OC12-HSL. However, our *in vitro* results (**Figure 3**) suggested that naringenin just partially outcompetes 3OC12-HSL when this AHL is already bound to LasR protein. To analyze this issue, the *E. coli* JM109 (pSB1142) strain was grown in presence of 3OC12-HSL and, with a delay of 2 h, naringenin was added to the cultures. As shown in **Figure 4**, the inhibitory effect of naringenin, when added to cultures already containing 3OC12-HSL, was lower than when both compounds were simultaneously added, while no induction of the LasR-driven response was observed when 3OC12-HSL was added with 2 h of delay respect to naringenin addition. In agreement with our *in vitro* analyses (**Figure 3**), these results suggest that, although a partial shift between 3OC12-HSL and naringenin is possible when the first is bound to LasR, the efficiency of this displacement is low. Altogether, our results support that QS inhibition mediated by naringenin is more efficient when naringenin binds to the nascent LasR protein than when 3OC12-HSL has already bound to the LasR regulator. This is consistent with the similar affinities estimated for both molecules during modeling.

The Expression of QS-Regulated Genes and the Production of QS-Controlled Virulence Factors in *P. aeruginosa* Are Inhibited in a Time-Dependent Manner by Naringenin

The previous analyses show that the pre-activation of LasR by 3OC12-HSL may limit the inhibitory capacity of naringenin. These analyses were made *in vitro* and in a heterologous host that does not produce QS signals, in order to fully control the concentrations and timing of both, naringenin and 3OC12-HSL addition.

Nevertheless, since 3OC12-HSL is naturally produced by *P. aeruginosa* depending on the cellular status of the population, the functional verification of these results in this organism, the host bacterium of the studied QS network, is needed. Consequently, the expression of the QS-regulated genes *lasA* and *lasB,* that encode a staphylolysin and a pseudolysin respectively (Storey et al., 1998), was measured using early or late QS inhibitory conditions in *P. aeruginosa* to determine if the inhibitory capacity of naringenin is also time-dependent in this microorganism. For that purpose, *P. aeruginosa* PAO1 was grown in the presence or in the absence of naringenin, added at time zero (early inhibition), when 3OC12-HSL concentration is low,

or after 4 h of growth (late inhibition), when the population is at early stationary phase and cells have accumulated 3OC12-HSL. In both inhibitory assays, gene expression was measured by real time RT-PCR after 4 h of naringenin treatment. As shown in **Figure 5A**, naringenin was able to reduce halfway the expression of both *lasA* and *lasB* genes in early inhibitory conditions. However, late inhibition of the expression of these QS-regulated genes was not observed.

It has been described that naringenin is able to inhibit the production of QS-regulated virulence factors in *P. aeruginosa* (Rasamiravaka et al., 2013). However, this work neither analyze the molecular basis of such inhibition nor its time-dependent efficacy. According to our results, this should be also relevant for efficient QS inhibition, because the inhibition of the QS response may depend on the QS status of the population. Consequently, we analyzed the effect that either early (t = 0 h) or late (t = 4 h) addition of naringenin may have on the production of two known QS-regulated virulence factors, pyocyanin and elastase, at the late stationary phase. As shown in **Figure 5B**, early addition of naringenin inhibited pyocyanin and elastase production, up to 5- and 7-times, respectively. However, when naringenin was added later, inhibition of the production of these virulence factors was not detected (**Figure 5B**).

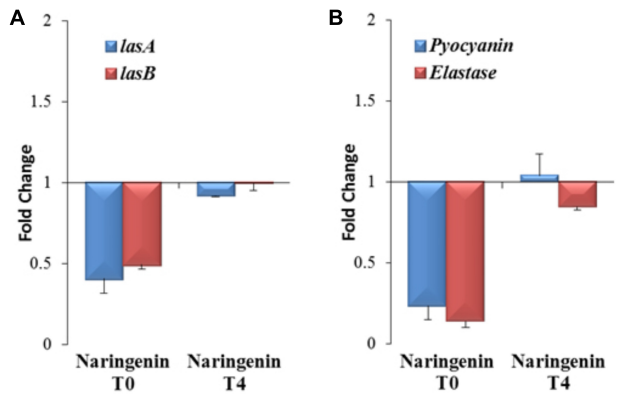


FIGURE 5 | Analysis of early and late inhibition of the expression of the QS-regulated genes and virulence factors by naringenin in *P. aeruginosa* PAO1. **(A)** The expression of the *lasA* and *lasB* QS-regulated genes was analyzed by real time PCR in cells treated with naringenin at t = 0 h (early inhibition) or at t = 4 h (late inhibition) after 4 h of growth. Results are represented as the fold change in the level of expression compared to same strain grown in presence of the solvent of naringenin, DMSO. The results show that inhibition of LasR-dependent regulation by naringenin is more efficient when this flavonoid is added at t = 0 h. Contrary, there is no inhibition by naringenin of *lasA* and *lasB* expression when this flavonoid was added 4 h after incubation (early stationary phase), when 3OC12-HSL production by *P. aeruginosa* is high. **(B)** Elastase activity and pyocyanin production were determined at the late stationary phase (OD_{600} = 4.4; t = 8 h) in cells treated, at t = 0 h (early inhibition) or at t = 4 h (late inhibition), with naringenin or its solvent, DMSO. These graphs show that inhibition of LasR-dependent regulation and QS-dependent production of virulence factors by naringenin is more efficient when this flavonoid is added at t = 0 h, a cellular status where the production of 3OC12-HSL is practically zero. Contrary, there is no inhibition by naringenin of QS-regulated genes or of virulence factors production when this flavonoid was added 4 h after incubation (early stationary phase), when 3OC12-HSL production by *P. aeruginosa* is high and naringenin cannot displace the AI from LasR. In both figures, error bars represent standard error of three independent replicates.

Altogether, these results indicate that early inhibition of the QS response mediated by naringenin is possible in *P. aeruginosa*, but its effectiveness may be compromised when 3OC12-HSL is already bound to the LasR regulator.

DISCUSSION

Flavonoids play important physiological functions in plants as well as in plant-bacteria interactions, in particular in the natural ecosystem of the rhizosphere, where microorganisms live in contact with plants' roots. In this environment, plant-produced flavonoids confer to them protection against microbial attacks (Farhadi et al., 2019). One of such protection mechanism has been described in the case of *Pseudomonas syringae* infections; flavonoids reduce bacterial virulence by inhibiting the GacS/GacA two-component system (Chatterjee et al., 2003; Vargas et al., 2013). Meanwhile, in an example of Red-Queen adaptive co-evolution, this bacterium is able to overcome this inhibition by the flavonoid-dependent induction of the multidrug efflux system, MexAB-OprM (Vargas et al., 2011; Alcalde-Rico et al., 2016), which is able to extrude flavonoids and is needed for the efficient colonization of tomato plants by this bacterium. Thus, taking into consideration the intense co-evolution of plants and bacteria, exemplified by the above described flavonoid-mediated plant/bacteria interactions, it is not difficult to think on the existence of plant molecules with capacity to inhibit one of the most relevant circuits regulating bacterial virulence, the QS regulatory network, a field currently under study (Ta et al., 2014; Bacha et al., 2016; Dettweiler et al., 2019).

It has been described that naringenin is a flavonoid with therapeutic potential (Salehi et al., 2019), particularly in inflammation-related diseases (Zeng et al., 2018), as well as an inhibitor of the production of *P. aeruginosa* QS-regulated virulence factors (Vandeputte et al., 2011; Rasamiravaka et al., 2013). However, the molecular mechanism by which naringenin inhibits QS in *P. aeruginosa* has not been determined yet. The modeling of naringenin binding to LasR predicts that this flavonoid binds preferentially to the same site as 3OC12-HSL, with an affinity that seems similar to that of 3OC12-HSL, and that it would be unable to establish contact with all the amino acid residues needed for the formation of the structure of a functional monomer that can oligomerize (**Figure 2**), so the nascent LasR chain that binds naringenin would be unable to oligomerize, and 3OC12-HSL would not be able to form active LasR complexes in the presence of this flavonoid (**Figure 4**). Further, the similarity in affinities could explain why each of the two molecules has difficulty in displacing the other once bound to LasR, since it has been stated that only ligands presenting higher affinities can displace from LasR the originally bound ligand (Paczkowski et al., 2017). Although a direct effect of naringenin on LasR stability cannot be fully disregarded, our experimental findings support that naringenin inhibits *P. aeruginosa* QS response through its binding to the LasR regulator by a time-dependent model of competition with 3OC12-HSL. This model implies

that, when *P. aeruginosa* populations present low densities (and hence low 3OC12-HSL concentration), naringenin can bind to nascent LasR, inhibiting its 3OC12-HSL-dependent activation. However, in dense *P. aeruginosa* population, presenting a high accumulation of 3OC12-HSL, inhibition of the QS response by naringenin would not be achieved. This time-dependent model of inhibition might compromise the efficiency of this flavonoid as a QS inhibitor, hence making difficult its application in clinics, but also opens the door for seeking more efficient analogs. This issue will be relevant when dense populations, whose QS response is already triggered, as those forming biofilms (Costerton, 2001; Winstanley et al., 2016), are confronted with this type of QS inhibitors.

Deepening into the mechanisms of QS inhibition may help in the development of such inhibitors, which have been proposed as suitable anti-infective drugs (Fleitas Martinez et al., 2018). Based in our results, we therefore propose that screenings for QS inhibitors should not only focus on finding molecules with an overall inhibitory effect of virulence phenotypes, but also should analyze their effectiveness over dense populations and select those whose mechanisms of action do not depend on the QS status of populations. This would increase their potential application against *P. aeruginosa* infections, where QS response is frequently triggered.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/supplementary material.

AUTHOR CONTRIBUTIONS

SH-A, MA-R, and TG-G performed the experimental work. JV performed the computational predictive work. SH-A, MA-R, TG-G, JV, and JM designed the study, contributed to the interpretation of the results and in writing the article.

FUNDING

Work has been supported by grants from the Instituto de Salud Carlos III (Spanish Network for Research on Infectious Diseases [RD16/0016/0011]), from the Spanish Ministry of Economy, Industry and Competitivity (BIO2017-83128-R) and from the Autonomous Community of Madrid (B2017/BMD-3691). The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

ACKNOWLEDGMENTS

Thanks are given to Alicia Sánchez-Gorostiaga for help in the design of the EMSA experiments and to Mónica Chagoyen, from the Sequence Analysis and Structure Prediction facility of the Centro Nacional de Biotecnología for her help in preliminary *in silico* docking studies. TG-G is the recipient of a FPI fellowship from MINECO.

REFERENCES

Alcalde-Rico, M., Hernando-Amado, S., Blanco, P., and Martinez, J. L. (2016). Multidrug efflux pumps at the crossroad between antibiotic resistance and bacterial virulence. *Front. Microbiol.* 7:1483. doi: 10.3389/fmicb.2016.01483

Bacha, K., Tariku, Y., Gebreyesus, F., Zerihun, S., Mohammed, A., Weiland-Brauer, N., et al. (2016). Antimicrobial and anti-Quorum sensing activities of selected medicinal plants of Ethiopia: implication for development of potent antimicrobial agents. *BMC Microbiol.* 16:139. doi: 10.1186/s12866-016-0765-9

Bjarnsholt, T., and Givskov, M. (2008). Quorum sensing inhibitory drugs as next generation antimicrobials: worth the effort? *Curr. Infect. Dis. Rep.* 10, 22–28. doi: 10.1007/s11908-008-0006-y

Bottomley, M. J., Muraglia, E., Bazzo, R., and Carfi, A. (2007). Molecular insights into quorum sensing in the human pathogen *Pseudomonas aeruginosa* from the structure of the virulence regulator LasR bound to its autoinducer. *J. Biol. Chem.* 282, 13592–13600. doi: 10.1074/jbc.m700556200

Chatterjee, A., Cui, Y., Yang, H., Collmer, A., Alfano, J. R., and Chatterjee, A. K. (2003). GacA, the response regulator of a two-component system, acts as a master regulator in *Pseudomonas syringae* pv. tomato DC3000 by controlling regulatory RNA, transcriptional activators, and alternate sigma factors. *Mol. Plant Microbe Interact.* 16, 1106–1117. doi: 10.1094/mpmi.2003.16.12.1106

Corral-Lugo, A., Daddaoua, A., Ortega, A., Espinosa-Urgel, M., and Krell, T. (2016). Rosmarinic acid is a homoserine lactone mimic produced by plants that activates a bacterial quorum-sensing regulator. *Sci. Signal.* 9:ra1. doi: 10.1126/scisignal.aaa8271

Costerton, J. W. (2001). Cystic fibrosis pathogenesis and the role of biofilms in persistent infection. *Trends Microbiol.* 9, 50–52. doi: 10.1016/s0966-842x(00)01918-1

Dettweiler, M., Lyles, J. T., Nelson, K., Dale, B., Reddinger, R. M., Zurawski, D. V., et al. (2019). American Civil War plant medicines inhibit growth, biofilm formation, and quorum sensing by multidrug-resistant bacteria. *Sci. Rep.* 9:7692. doi: 10.1038/s41598-019-44242-y

Diggle, S. P., Crusz, S. A., and Camara, M. (2007). Quorum sensing. *Curr. Biol.* 17, R907–R910.

Erickson, D. L., Endersby, R., Kirkham, A., Stuber, K., Vollman, D. D., Rabin, H. R., et al. (2002). *Pseudomonas aeruginosa* quorum-sensing systems may control virulence factor expression in the lungs of patients with cystic fibrosis. *Infect. Immun.* 70, 1783–1790. doi: 10.1128/iai.70.4.1783-1790.2002

Farhadi, F., Khameneh, B., Iranshahi, M., and Iranshahy, M. (2019). Antibacterial activity of flavonoids and their structure-activity relationship: an update review. *Phytother. Res.* 33, 13–40. doi: 10.1002/ptr.6208

Fernandez, M., Corral-Lugo, A., and Krell, T. (2018). The plant compound rosmarinic acid induces a broad quorum sensing response in *Pseudomonas aeruginosa* PAO1. *Environ. Microbiol.* 20, 4230–4244. doi: 10.1111/1462-2920.14301

Fleitas Martinez, O., Rigueiras, P. O., Pires, A. D. S., Porto, W. F., Silva, O. N., De La Fuente-Nunez, C., et al. (2018). Interference with quorum-sensing signal biosynthesis as a promising therapeutic strategy against multidrug-resistant pathogens. *Front. Cell Infect. Microbiol.* 8:444. doi: 10.3389/fcimb.2018.00444

Fuqua, W. C., Winans, S. C., and Greenberg, E. P. (1994). Quorum sensing in bacteria: the LuxR-LuxI family of cell density-responsive transcriptional regulators. *J. Bacteriol.* 176, 269–275. doi: 10.1128/jb.176.2.269-275.1994

Hurley, M. N., Camara, M., and Smyth, A. R. (2012). Novel approaches to the treatment of *Pseudomonas aeruginosa* infections in cystic fibrosis. *Eur. Respir. J.* 40, 1014–1023.

Jin, L., Zeng, W., Zhang, F., Zhang, C., and Liang, W. (2017). Naringenin ameliorates acute inflammation by regulating intracellular cytokine degradation. *J. Immunol.* 199, 3466–3477. doi: 10.4049/jimmunol.1602016

Kalia, M., Singh, P. K., Yadav, V. K., Yadav, B. S., Sharma, D., Narvi, S. S., et al. (2017). Structure based virtual screening for identification of potential quorum sensing inhibitors against LasR master regulator in *Pseudomonas aeruginosa*. *Microb. Pathog.* 107, 136–143. doi: 10.1016/j.micpath.2017.03.026

Kallberg, M., Wang, H., Wang, S., Peng, J., Wang, Z., Lu, H., et al. (2012). Template-based protein structure modeling using the RaptorX web server. *Nat. Protoc.* 7, 1511–1522. doi: 10.1038/nprot.2012.085

Kessler, E., and Safrin, M. (2014). Elastinolytic and proteolytic enzymes. *Methods Mol. Biol.* 1149, 135–169. doi: 10.1007/978-1-4939-0473-0_13

Kiratisin, P., Tucker, K. D., and Passador, L. (2002). LasR, a transcriptional activator of *Pseudomonas aeruginosa* virulence genes, functions as a multimer. *J. Bacteriol.* 184, 4912–4919. doi: 10.1128/jb.184.17.4912-4919.2002

Lamb, J. R., Patel, H., Montminy, T., Wagner, V. E., and Iglewski, B. H. (2003). Functional domains of the RhlR transcriptional regulator of *Pseudomonas aeruginosa*. *J. Bacteriol.* 185, 7129–7139. doi: 10.1128/jb.185.24.7129-7139.2003

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2-DDCT Method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., and Simmerling, C. (2015). ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* 11, 3696–3713. doi: 10.1021/acs.jctc.5b00255

Mccready, A. R., Paczkowski, J. E., Henke, B. R., and Bassler, B. L. (2019). Structural determinants driving homoserine lactone ligand selection in the *Pseudomonas aeruginosa* LasR quorum-sensing receptor. *Proc. Natl. Acad. Sci. U.S.A.* 116, 245–254. doi: 10.1073/pnas.1817239116

Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., et al. (2009). AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J. Comput. Chem.* 30, 2785–2791. doi: 10.1002/jcc.21256

Nealson, K. H., Platt, T., and Hastings, J. W. (1970). Cellular control of the synthesis and activity of the bacterial luminescent system. *J. Bacteriol.* 104, 313–322. doi: 10.1128/jb.104.1.313-322.1970

Neudert, G., and Klebe, G. (2011). DSX: a knowledge-based scoring function for the assessment of protein-ligand complexes. *J. Chem. Inf. Model.* 51, 2731–2745. doi: 10.1021/ci200274q

Oinuma, K., and Greenberg, E. P. (2011). Acyl-homoserine lactone binding to and stability of the orphan *Pseudomonas aeruginosa* quorum-sensing signal receptor QscR. *J. Bacteriol.* 193, 421–428. doi: 10.1128/JB.01041-10

Paczkowski, J. E., Mukherjee, S., Mccready, A. R., Cong, J. P., Aquino, C. J., Kim, H., et al. (2017). Flavonoids suppress *Pseudomonas aeruginosa* virulence through allosteric inhibition of quorum-sensing receptors. *J. Biol. Chem.* 292, 4064–4076. doi: 10.1074/jbc.M116.770552

Parsek, M. R., and Greenberg, E. P. (1999). Quorum sensing signals in development of *Pseudomonas aeruginosa* biofilms. *Methods. Enzymol.* 310, 43–55. doi: 10.1016/s0076-6879(99)10005-3

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., et al. (2004). UCSF chimera - a visualization system for exploratory research and analysis. *J. Comput. Chem.* 25, 1605–1612. doi: 10.1002/jcc.20084

Rahme, L. G., Stevens, E. J., Wolfort, S. F., Shao, J., Tompkins, R. G., and Ausubel, F. M. (1995). Common virulence factors for bacterial pathogenicity in plants and animals. *Science* 268, 1899–1902. doi: 10.1126/science.7604262

Rasamiravaka, T., Jedrzejowski, A., Kiendrebeogo, M., Rajaonson, S., Randriamampionona, D., Rabemanantsoa, C., et al. (2013). Endemic malagasy dalbergia species inhibit quorum sensing in *Pseudomonas aeruginosa* PAO1. *Microbiology* 159, 924–938. doi: 10.1099/mic.0.064378-0

Rudrappa, T., and Bais, H. P. (2008). Curcumin, a known phenolic from curcuma longa, attenuates the virulence of *Pseudomonas aeruginosa* PAO1 in whole plant and animal pathogenicity models. *J. Agric. Food Chem.* 56, 1955–1962. doi: 10.1021/jf072591j

Salehi, B., Fokou, P. V. T., Sharifi-Rad, M., Zucca, P., Pezzani, R., Martins, N., et al. (2019). The Therapeutic potential of naringenin: a review of clinical trials. *Pharmaceuticals* 12:E11.

Sambrook, J., and Russell, D. W. (2001). *Molecular Cloning. A Laboratory Manual*, 3rd Edn, Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Schneidman-Duhovny, D., Inbar, Y., Nussinov, R., and Wolfson, H. J. (2005). PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Res.* 33, W363–W367.

Schuster, M., and Greenberg, E. P. (2007). Early activation of quorum sensing in *Pseudomonas aeruginosa* reveals the architecture of a complex regulon. *BMC Genom.* 8:287. doi: 10.1186/1471-2164-8-287

Schuster, M., Urbanowski, M. L., and Greenberg, E. P. (2004). Promoter specificity in *Pseudomonas aeruginosa* quorum sensing revealed by DNA binding of purified LasR. *Proc. Natl. Acad. Sci. U.S.A.* 101, 15833–15839. doi: 10.1073/pnas.0407229101

Smyth, A. R., Cifelli, P. M., Ortori, C. A., Righetti, K., Lewis, S., Erskine, P., et al. (2010). Garlic as an inhibitor of *Pseudomonas aeruginosa* quorum sensing in cystic fibrosis–a pilot randomized controlled trial. *Pediatr. Pulmonol.* 45, 356–362. doi: 10.1002/ppul.21193

Stewart, J. J. P. (2008). *MOPAC2009, Stewart Computational Chemistry*. Colorado Springs, CO: MOPAC.

Storey, D. G., Ujack, E. E., Rabin, H. R., and Mitchell, I. (1998). *Pseudomonas aeruginosa* lasR transcription correlates with the transcription of lasA, lasB, and toxA in chronic lung infections associated with cystic fibrosis. *Infect. Immun.* 66, 2521–2528. doi: 10.1128/iai.66.6.2521-2528.1998

Ta, C. A., Freundorfer, M., Mah, T. F., Otarola-Rojas, M., Garcia, M., Sanchez-Vindas, P., et al. (2014). Inhibition of bacterial quorum sensing and biofilm formation by extracts of neotropical rainforest plants. *Planta Med.* 80, 343–350. doi: 10.1055/s-0033-1360337

Tan, M. W., Rahme, L. G., Sternberg, J. A., Tompkins, R. G., and Ausubel, F. M. (1999). Pseudomonas aeruginosa killing of *Caenorhabditis elegans* used to identify *P. aeruginosa virulence factors*. *Proc. Natl. Acad. Sci. U.S.A.* 96, 2408–2413. doi: 10.1073/pnas.96.5.2408

Vandeputte, O. M., Kiendrebeogo, M., Rasamiravaka, T., Stevigny, C., Duez, P., Rajaonson, S., et al. (2011). The flavanone naringenin reduces the production of quorum sensing-controlled virulence factors in *Pseudomonas aeruginosa* PAO1. *Microbiology* 157, 2120–2132. doi: 10.1099/mic.0.049338-0

Vargas, P., Farias, G. A., Nogales, J., Prada, H., Carvajal, V., Baron, M., et al. (2013). Plant flavonoids target *Pseudomonas* syringae pv. *tomato DC*3000 flagella and type III secretion system. *Environ. Microbiol. Rep.* 5, 841–850. doi: 10.1111/1758-2229.12086

Vargas, P., Felipe, A., Michan, C., and Gallegos, M. T. (2011). Induction of *Pseudomonas* syringae pv. tomato DC3000 MexAB-OprM multidrug efflux pump by flavonoids is mediated by the repressor PmeR. *Mol. Plant Microbe Interact.* 24, 1207–1219. doi: 10.1094/MPMI-03-11-0077

Wagner, V. E., Bushnell, D., Passador, L., Brooks, A. I., and Iglewski, B. H. (2003). Microarray analysis of *Pseudomonas aeruginosa* quorum-sensing regulons: effects of growth phase and environment. *J. Bacteriol.* 185, 2080–2095. doi: 10.1128/jb.185.7.2080-2095.2003

Wagner, V. E., and Iglewski, B. H. (2008). *Pseudomonas aeruginosa* biofilms in CF infection. *Clin. Rev. Allergy Immunol.* 35, 124–134. doi: 10.1007/s12016-008-8079-9

Wang, R., Lai, L., and Wang, S. (2002). Further development and validation of empirical scoring functions for structure-based binding affinity prediction. *J. Comput. Aided Mol. Des.* 16, 11–26.

Williams, P., Camara, M., Hardman, A., Swift, S., Milton, D., Hope, V. J., et al. (2000). Quorum sensing and the population-dependent control of virulence. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 355, 667–680.

Williams, P., Winzer, K., Chan, W. C., and Camara, M. (2007). Look who's talking: communication and quorum sensing in the bacterial world. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 362, 1119–1134. doi: 10.1098/rstb.2007.2039

Winsor, G. L., Griffiths, E. J., Lo, R., Dhillon, B. K., Shay, J. A., and Brinkman, F. S. (2016). Enhanced annotations and features for comparing thousands of *Pseudomonas* genomes in the *Pseudomonas genome* database. *Nucleic Acids Res.* 44, D646–D653. doi: 10.1093/nar/gkv1227

Winstanley, C., O'brien, S., and Brockhurst, M. A. (2016). *Pseudomonas aeruginosa* evolutionary adaptation and diversification in cystic fibrosis chronic lung infections. *Trends Microbiol.* 24, 327–337. doi: 10.1016/j.tim.2016.01.008

Zeng, W., Jin, L., Zhang, F., Zhang, C., and Liang, W. (2018). Naringenin as a potential immunomodulator in therapeutics. *Pharmacol. Res.* 135, 122–126. doi: 10.1016/j.phrs.2018.08.002

Zhu, J., and Winans, S. C. (2001). The quorum-sensing transcriptional regulator TraR requires its cognate signaling ligand for protein folding, protease resistance, and dimerization. *Proc. Natl. Acad. Sci. U.S.A.* 98, 1507–1512. doi: 10.1073/pnas.98.4.1507

Zou, Y., and Nair, S. K. (2009). Molecular basis for the recognition of structurally distinct autoinducer mimics by the *Pseudomonas aeruginosa* LasR quorum-sensing signaling receptor. *Chem. Biol.* 16, 961–970. doi: 10.1016/j.chembiol.2009.09.001

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Hernando-Amado, Alcalde-Rico, Gil-Gil, Valverde and Martínez. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 19 March 2020
doi: 10.3389/fmolb.2020.00043

In vitro DNA Inversions Mediated by the PsrA Site-Specific Tyrosine Recombinase of *Streptococcus pneumoniae*

Jingwen Li[1†], *Juanjuan Wang*[1†], *Sofía Ruiz-Cruz*[2‡], *Manuel Espinosa*[2], *Jing-Ren Zhang*[1*] *and Alicia Bravo*[2*]

[1] Department of Basic Medical Science, Center for Infectious Disease Research, School of Medicine, Tsinghua University, Beijing, China, [2] Centro de Investigaciones Biológicas Margarita Salas, Consejo Superior de Investigaciones Científicas, Madrid, Spain

OPEN ACCESS

Edited by:
Brian M. Baker,
University of Notre Dame,
United States

Reviewed by:
Andrew Benjamin Herr,
Cincinnati Children's Hospital Medical Center, United States
Victoria L. Robinson,
University of Connecticut,
United States

***Correspondence:**
Jing-Ren Zhang
zhanglab@tsinghua.edu.cn
Alicia Bravo
abravo@cib.csic.es

[†]These authors have contributed equally to this work

[‡]**Present address:**
Sofía Ruiz-Cruz,
School of Microbiology, University College Cork, Cork, Ireland

Specialty section:
This article was submitted to Molecular Recognition, a section of the journal Frontiers in Molecular Biosciences

Received: 17 October 2019
Accepted: 28 February 2020
Published: 19 March 2020

Citation:
Li J, Wang J, Ruiz-Cruz S, Espinosa M, Zhang J-R and Bravo A (2020) In vitro DNA Inversions Mediated by the PsrA Site-Specific Tyrosine Recombinase of Streptococcus pneumoniae. Front. Mol. Biosci. 7:43. doi: 10.3389/fmolb.2020.00043

Site-specific recombination is a DNA breaking and reconstructing process that plays important roles in various cellular pathways for both prokaryotes and eukaryotes. This process requires a site-specific recombinase and direct or inverted repeats. Some tyrosine site-specific recombinases catalyze DNA inversions and regulate subpopulation diversity and phase variation in many bacterial species. In *Streptococcus pneumoniae*, the PsrA tyrosine recombinase was shown to control DNA inversions in the three DNA methyltransferase *hsdS* genes of the type I restriction-modification *cod* locus. Such DNA inversions are mediated by three inverted repeats (IR1, IR2, and IR3). In this work, we purified an untagged form of the PsrA protein and studied its DNA-binding and catalytic features. Gel retardation assays showed that PsrA binds to linear and supercoiled DNAs, containing or not inverted repeats. Nevertheless, DNase I footprinting assays showed that, on linear DNAs, PsrA has a preference for sites that include an IR1 sequence (IR1.1 or IR1.2) and its boundary sequences. Furthermore, on supercoiled DNAs, PsrA was able to generate DNA inversions between specific inverted repeats (IR1, IR2, and IR3), which supports its ability to locate specific target sites. Unlike other site-specific recombinases, PsrA showed reliance on magnesium ions for efficient catalysis of IR1-mediated DNA inversions. We discuss that PsrA might find its specific binding sites on the bacterial genome by a mechanism that involves transitory non-specific interactions between protein and DNA.

Keywords: tyrosine recombinase PsrA, *Streptococcus pneumoniae*, type I restriction-modification *cod* locus, site-specific DNA inversions, inverted repeats

INTRODUCTION

Site-specific recombination (SSR) is a DNA breaking and reconstructing process widely distributed in both prokaryotes and eukaryotes, in which a specialized enzyme catalyzes reciprocal strand exchange at specific target sites. According to this fundamental definition, two critical elements participate in the process, namely a site-specific recombinase and a pair of inversely or directly repeated sequences (Grindley et al., 2006; Rajeev et al., 2009). Based on the amino acid residue involved in the catalysis, site-specific recombinases have been classified into two families: tyrosine (Tyr) and serine (Ser) recombinases, both of them employing the C-terminal OH group of the

active residue to perform the nucleophilic attack for DNA sequence exchange (Grindley et al., 2006). In addition to the difference in the catalytic residues, the catalytic mechanisms of Tyr- and Ser-recombinases are different. In the former, after the nucleophilic attack, the catalytic Tyr residue remains covalently linked to the 3′-end of the DNA strand, generating a 3′-phospho-tyrosyl bond and leaving a free 5′-OH group. Besides, Tyr-recombinases cleave just one DNA strand every time, and only after the cleaved strand is rejoined at the crossover site forming a unique holiday junction (HJ) intermediate, the uncleaved strand can be cut by its partner within the same dimer. In contrast, Ser-recombinases form a 5′-phosphor-serine bond generating free 3′-OH intermediates. Further, the enzyme always cleaves both strands at a time and does not generate the HJ intermediate. Unlike general homologous recombination event that requires large-size homologous DNA segments, DNA replication and high-energy cofactors, the specific repeated sequences recognized by site-specific recombinases are usually 20–40 bp in length, and no DNA synthesis or energy factors are required (Grindley et al., 2006).

In bacteria, SSR is an efficient and feasible way to generate genetic rearrangements, leading to regulation of subpopulation diversity, or phase variation and cellular adaption to various environments at physiological and phenotypic level (Darmon and Leach, 2014). As one of the typical forms of SSR, DNA inversion is an important regulatory mechanism to introduce genetic and phenotypic variation in a clonal population, also referred to intercellular heterogeneity, which is well-characterized in several pathogenic bacterial species (Darmon and Leach, 2014). In particular, promoter inversion is one of the most commonly used strategies to regulate surface antigen expression, as in the cases of the *fim* switch of *Escherichia coli* (Abraham et al., 1985; Olsen and Klemm, 1994) and the flagellar switch in *Salmonella typhimurium* (Zieg and Simon, 1980; Johnson et al., 1986). In addition, genetic rearrangement to switch gene contents is also a practical strategy widely adopted in bacteria, such as the VlsE lipoprotein variation in *Borrellia burgdorferi* (Zhang et al., 1997; Norris, 2014) and the V-1 surface antigen variation in *Mycoplasma pulmonis* (Bhugra et al., 1995; Shen et al., 2000).

Studies from our (Feng et al., 2014; Li et al., 2016; Li and Zhang, 2019) and other laboratories (Manso et al., 2014; De Ste Croix et al., 2017; de Ste Croix et al., 2019) revealed that DNA inversions in the *hsdS* genes of the *cod* locus of *Streptococcus pneumoniae* (the pneumococcus) lead to diversification of genomic DNA methylation patterns, and phenotypic variation in colony opacity. As a type I restriction-modification (RM) system, the *cod* locus is composed of six genes, including *hsdR* (restriction enzyme), *hsdM* (DNA methyltransferase), three target recognition subunit *hsdS* homologs, and *psrA* (site-specific tyrosine recombinase) (**Figure 1**). Among the three *hsdS* genes, only the $hsdS_A$ gene, which is co-transcribed with *hsdR* and *hsdM*, is functional to recognize the specific methylation motif, whereas $hsdS_B$ and $hsdS_C$ genes are serving as DNA sources for sequence switching with $hsdS_A$ by inversions (see also **Figure 1**). Three pairs of inverted repeats (IRs), termed 15-bp IR1, 298-bp IR2, and 85-bp IR3, were identified to drive sequence exchange between $hsdS_A$ and $hsdS_B$/$hsdS_C$ (Li et al., 2016). Further investigation focusing on the molecular mechanisms of site-specific *hsdS* inversions in the *cod* locus revealed that pneumococci employ multiple pathways to catalyze *hsdS* inversions between the three inverted repeats, including dominant PsrA-driven inversions and less frequent inversions that don't require PsrA (Li et al., 2019). Particularly, the Tyr-recombinase PsrA plays an essential role in catalyzing all three inversion reactions mediated by IR1, IR2, and IR3. While IR1 inversions are strictly dependent on PsrA catalysis, IR2- and IR3-inversions are dramatically reduced in *psrA* mutant strains. Sequence replacement of the IR1 repeat and its boundary sequences with irrelevant sequences demonstrated that efficient catalysis by PsrA *in vivo* depends on the sequence specificity of the 15-bp IR1 repeat and its immediately upstream sequence, which are supposed to be the target sequence where PsrA performs a nucleophilic attack or interacts directly. In addition, PsrA-mediated IR2- and IR3-inversions rely on the presence of IR1-like sequences in the IR2 and IR3 repeats. However, the precise interaction of PsrA with the IR1 repeat and the molecular mechanism of its catalytic activity remain to be elucidated.

In the present work, we have purified an untagged form of the pneumococcal PsrA protein and studied its DNA-binding and catalytic features. Our results suggest that PsrA might locate its preferred target sites (specific interactions) by a mechanism that involves transient binding at random DNA sites (non-specific interactions). Different from other Tyr-recombinases, protein PsrA exhibited a marked dependence on Mg^{2+} cations for efficient catalysis *in vitro*.

RESULTS

Purification of PsrA

To overproduce an untagged version of the pneumococcal PsrA site-specific recombinase, the *psrA* gene was inserted into the *E. coli* inducible expression vector pET24b, which is based on the phi10 promoter recognized by the T7 RNA polymerase. The recombinant plasmid (pTH12647) was then introduced into the *E. coli* BL21 (DE3) strain that carries the T7 RNA polymerase-encoding gene fused to the *lacUV5* promoter (Studier and Moffatt, 1986). This strain carries also a chromosomal copy of the *lacI* repressor gene. Thus, expression of the T7 RNA polymerase-encoding gene, and consequently expression of the *psrA* gene, is induced when isopropyl β-D-1-thiogalactopyranoside (IPTG) is added to the bacterial culture (see Materials and Methods) (**Figure 2**, lane 2). The procedure used for large-scale purification of PsrA involved essentially the following steps (**Figure 2**): (i) precipitation of nucleic acids with polyethyleneimine (PEI) (0.2%) in the presence of NaCl (300 mM). The ionic strength at which PEI precipitation was done was high enough to keep PsrA from binding to DNA (lane 6); (ii) after PEI precipitation, proteins in the supernatant (including PsrA) were precipitated with 70% saturated ammonium sulfate; and (iii) chromatography

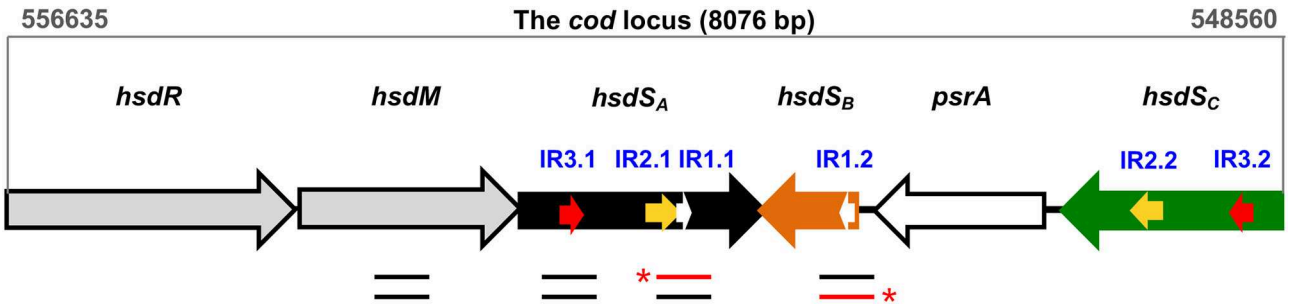


FIGURE 1 | Genetic organization of the *cod* locus. The *cod* locus of the pneumococcal ST556 genome (coordinates 556635–548560) consists of six genes: *hsdR* (restriction enzyme), *hsdM* (DNA methyltransferase), three *hsdS* homologs ($hsdS_A$, $hsdS_B$, $hsdS_C$), and *psrA* (tyrosine recombinase). Three pairs of inverted repeats, which mediate site-specific DNA inversions, are indicated with white arrows (IR1.1/IR1.2), yellow arrows (IR2.1/IR2.2), and red arrows (IR3.1/IR3.2), respectively. The location of DNA fragments used in EMSA and/or DNase I footprinting assays is indicated with lines below genes. The red lines indicate the DNA strands that were radiolabeled at the 5′-end (asterisks).

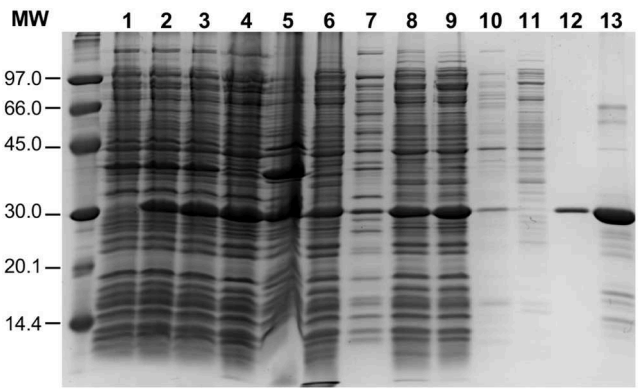


FIGURE 2 | Purification of PsrA. PsrA was overproduced using *E. coli* BL21 (DE3)/pTH12647 cells. Lane 1: whole-cell extract before adding IPTG. Lane 2: whole-cell extract from cells treated with IPTG (25 min). Lane 3: whole-cell extract from cells treated with IPTG (25 min) and then with rifampicin (60 min). The latter whole-cell extract (French Pressure Cell) was centrifuged: supernatant (lane 4; cleared cell lysate) and pellet (lane 5, cell debris). Precipitation of nucleic acids with PEI: supernatant (lane 6) and pellet (lane 7). Lane 8: Proteins recovered by precipitation with ammonium sulfate; this fraction was further dialyzed against buffer S containing 300 mM NaCl and centrifuged: supernatant (lane 9) and pellet (lane 10). The protein preparation shown in lane 9 was loaded onto a heparin affinity column: flow-through fraction (lane 11) and PsrA-containing fraction (lane 12). Lane 13: PsrA preparation after heparin chromatography. Proteins were analyzed by SDS-PAGE (12%). Gels were stained with Coomassie Blue. Molecular weight (MW) standards (in kDa) were run in the same gel (LMW Marker, GE Healthcare).

on a heparin column using a linear gradient of NaCl (300–800 mM) for protein elution. Protein PsrA was eluted at 675 mM NaCl. Purified PsrA was analyzed by SDS-polyacrylamide (PAA) (12%) gel electrophoresis. The PsrA preparation obtained after heparin chromatography was about 90% pure (lane 13). The 260/280 absorbance ratio for the purified protein was 0.52. PsrA migrated slightly above the 30 kDa reference band, which is consistent with the molecular weight of the PsrA monomer (31.2 kDa) calculated from the predicted amino acid sequence.

Binding of PsrA to Linear Double-Stranded DNAs

To analyse the DNA-binding properties of PsrA, we performed EMSA experiments using different linear double-stranded DNAs (dsDNAs). First, we used a radioactively labeled 265-bp DNA fragment (coordinates 550883-550619 of the pneumococcal ST556 genome). This fragment (here named IR1.2 DNA) contains the IR1.2 repeat (**Figure 1**). The labeled IR1.2 DNA (4 nM) was incubated with increasing concentrations of PsrA (60–400 nM) in the absence of competitor DNA. Free DNA and bound DNA were separated by native PAA (6%) gel electrophoresis (**Figure 3A**). As the protein concentration was increased, free IR1.2 DNA disappeared gradually and IR1.2-PsrA complexes, which did not enter the gel, were generated. The PsrA concentration required to bind half the DNA was determined by measuring the decrease in free DNA rather than the increase in complexes, which gives an indication of the approximate magnitude of the dissociation constant, K_d. Such a concentration was about 145 nM (**Figure 3B**). However, this value would underestimate the affinity of PsrA for its primary binding site if multiple protein units bind to the same DNA molecule. IR1.2-PsrA complexes unable to enter the gel were also observed when the binding reactions were analyzed by electrophoresis on agarose (1%) gels (**Figure S1**). By EMSA, we also analyzed the effect of NaCl on the binding reaction. PsrA (400 nM) was incubated with the labeled IR1.2 DNA in the presence of different concentrations of NaCl (125–500 mM) (**Figure 3C**). As the salt concentration was increased, the amount of free IR1.2 DNA increased, indicating that the formation of IR1.2-PsrA complexes is impaired at high NaCl concentration.

Next, the labeled IR1.2 DNA (4 nM, 0.65 μg/ml) was incubated with PsrA (400 nM) in the presence of different concentrations of non-labeled competitor calf thymus DNA (1–10 μg/ml) (**Figure 4A**). In the absence of competitor DNA, no free IR1.2 DNA was visualized. As the concentration of competitor DNA was increased, the amount of free IR1.2 DNA increased. However, only 20% of the IR1.2 DNA (0.13 μg/ml) moved as free DNA in the presence of 4 μg/ml of competitor DNA, which indicated that PsrA binds preferentially to the

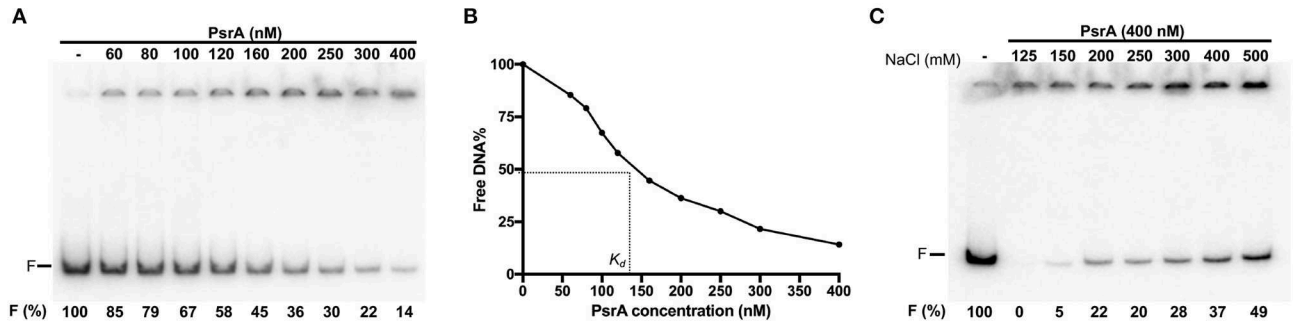


FIGURE 3 | Formation of IR1.2-PsrA complexes in the absence of competitor DNA. **(A)** Binding of PsrA (60–400 nM) to the ^{32}P-labeled IR1.2 DNA (4 nM). Reactions were loaded onto a native gel (6% PAA). DNA was visualized using a Fujifilm Image Analyzer FLA-3000. The band corresponding to free DNA (F) and the percentage of free DNA in each reaction (F%) are indicated. **(B)** Free DNA from the experiment shown in A was quantified using the Quantity One software (Bio-Rad), and the percentage of free DNA was plotted against the concentration of PsrA. K_d: dissociation constant. **(C)** Effect of increasing concentrations of NaCl on the binding of PsrA (400 nM) to the ^{32}P-labeled IR1.2 DNA (2 nM).

IR1.2 DNA. We also performed dissociation experiments using non-labeled competitor calf thymus DNA (**Figure 4B**). In this case, PsrA was first incubated with the labeled IR1.2 DNA for 20 min at room temperature. Under these conditions, free IR1.2 DNA molecules were not detected (formation of IR1.2-PsrA complexes). Then, different concentrations of competitor DNA were added to the reactions. After 5 min, reaction mixtures were loaded onto a native PAA (6%) gel. At 8 μg/ml of competitor DNA, only 16% of the IR1.2 DNA (0.10 μg/ml) moved as free DNA. Thus, PsrA dissociated from the IR1.2-PsrA complexes at high concentrations of competitor DNA, which can be taken as an indication of complex stability. IR1.2-PsrA interactions were also disrupted when heparin was used as a competitor (**Figure S2**).

Further EMSA experiments were performed using other linear dsDNA fragments from the pneumococcal ST556 genome: (i) a 265-bp DNA fragment that contains the IR1.1 repeat (here named IR1.1 DNA, coordinates 551690-551954; **Figure 1**); (ii) a 260-bp DNA fragment that contains the IR3.1 repeat (coordinates 552497-552756; **Figure 1**); and (iii) a 265-bp DNA fragment (coordinates 553471-553735) amplified from the *hsdM* gene, which lacks IRs. In all cases, DNA-PsrA complexes that did not enter the gel were generated (**Figure S3**). Moreover, we analyzed the binding of PsrA to a 26-bp dsDNA obtained by annealing of complementary oligonucleotides. These oligonucleotides do not contain IRs (Solano-Collado et al., 2013). In this case, DNA-PsrA complexes moving slower than free DNA were visualized by agarose (2%) gel electrophoresis (**Figure 5**). However, such complexes did not enter native PAA (8%) gels (data not shown). Thus, large DNA-PsrA complexes were also generated with the 26-bp dsDNA.

Taken all the above results together, we can conclude that gel retardation assays showed that PsrA binds to linear dsDNAs in a non-sequence-specific manner generating higher-order complexes. Similar results were obtained when supercoiled plasmid DNAs, containing or not IR1 repeats, were used as target DNAs (**Figure S4**).

DNase I Footprinting Analysis of IR1-PsrA Complexes

In vivo studies showed that PsrA is necessary and sufficient for IR1-mediated inversions. Moreover, PsrA requires both the 15-bp inverted repeats (IR1.1 and IR1.2) and the immediately upstream sequences of both repeats (Li et al., 2019). Here, we analyzed the interaction of PsrA with the IR1.1 and IR1.2 DNA fragments by DNase I footprinting experiments (**Figures 6**, **7**). The 265-bp IR1.1 DNA (see **Figure 1**) was radioactively labeled at the 5′-end of the coding strand (coordinate 551954; **Figure 6**). Labeled DNA (6 nM) was incubated with increasing concentrations of PsrA. At 45 nM, changes in DNase I sensitivity (diminished cleavages) were observed from coordinate 551835 to 551808, from 551792 to 551788, and from position 551761 onwards. Thus, the region protected against DNase I digestion includes the IR1.1 repeat (15-bp), the bordering regions (6-bp) and sequences located downstream of IR1.1 (**Figure 6**). When PsrA concentration was increased to 180 nM, DNase I cleavage was highly reduced, indicating the formation of higher-order IR1.1-PsrA complexes, in which the DNA was not anymore accessible to the DNase I attack.

The 265-bp IR1.2 DNA fragment was labeled at the 5′-end of the coding strand (coordinate 550619) (see **Figure 1**). Labeled DNA (4 nM) was incubated with increasing concentrations of PsrA. At 30 nM, diminished cleavages were observed within the region spanning coordinates 550704 and 550789 (**Figure 7**). Thus, PsrA recognizes a region that includes the IR1.2 repeat and its adjacent sequences. At higher PsrA concentrations (90 nM), IR1.2-PsrA complexes insensitive to DNase I cleavage were generated (**Figure 7**).

PsrA Catalyzes the Formation of IR1-Mediated Inversions *in vitro*

We further investigated whether PsrA was able to generate IR1 inversions *in vitro*. To this end, we constructed the recombinant plasmids pTH13166 (IR1-For) and pTH13170 (IR1-Rev), harboring S1 and S2 configurations of the *cod* locus,

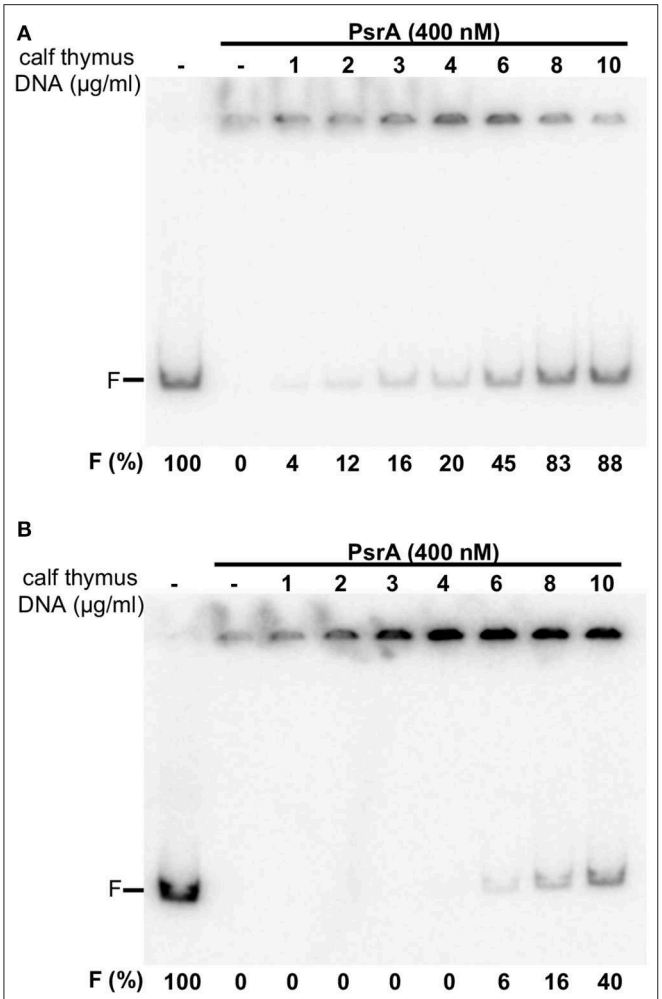


FIGURE 4 | Formation of IR1.2-PsrA complexes in the presence of competitor DNA. **(A)** Effect of increasing concentrations of calf thymus DNA on the binding of PsrA (400 nM) to the ^{32}P-labeled IR1.2 DNA (4 nM). Reactions were loaded onto a native gel (6% PAA). The band corresponding to free DNA (F) and the percentage of free DNA in each reaction (F%) are indicated **(B)** Dissociation of IR1.2-PsrA complexes. The indicated concentration of non-labeled calf thymus DNA was added to pre-formed IR1.2-PsrA complexes.

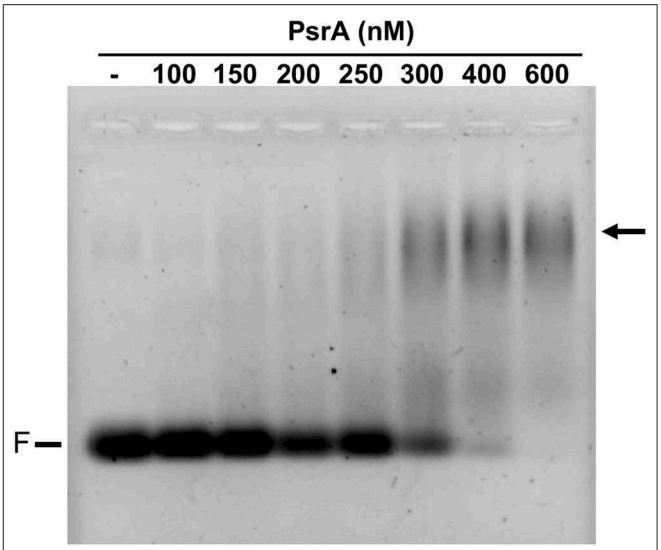


FIGURE 5 | Binding of PsrA to a 26-bp DNA fragment. The indicated concentration of PsrA was mixed with 50 nM of DNA. Binding reactions were analyzed by agarose (2%) gel electrophoresis. DNA was stained with GelRed (Biotium). F: free DNA. The position of DNA-PsrA complexes is indicated with an arrow.

respectively (**Figure 8A**). In both plasmids, the IR2.2 and IR3.2 repeats were removed to avoid the impact of IR2- and IR3-mediated inversions. Plasmid pTH13166 was used to examine IR1 inversions from "For state" to "Rev state," whereas the employment of plasmid pTH13170 allowed us to examine IR1 inversions from "Rev state" to "For state." To detect the occurrence of IR1 inversions, both PCR and qPCR assays were carried out using primers P1/P2 for pTH13166 and P1/P3 for pTH13170 (**Figure 8A**). PCR results were analyzed by agarose gel electrophoresis (upper part in **Figures 8B,C**). The qPCR data were used to calculate the relative inversion frequency (bar graphs in **Figures 8B,C**). In a first approach, PsrA (125 nM) was mixed with plasmid DNA (1 nM) in the presence of NaCl (125 mM) and $MgCl_2$ (5 mM). Reaction mixtures were incubated at different temperatures for 1 h (**Figures 8B,C**). IR1 inversions were detected in both plasmids only in the presence of PsrA. Thus, we could demonstrate that protein PsrA generates IR1 inversions *in vitro*. The optimum reaction temperature was 37°C in plasmid pTH13166 (**Figure 8B**) and 37–40°C in plasmid pTH13170 (**Figure 8C**), corresponding to the optimal temperature for pneumococci to grow.

Next, we determined the relative frequency of IR1 inversions in the presence of PsrA as a function of the incubation time. In both plasmids (pTH13166 and pTH13170), the maximum frequency of inversions was observed after 2 h incubation (bar graphs in **Figure S5**). Moreover, we examined the effect of pH and NaCl concentration on the frequency of PsrA-mediated IR1 inversions. The highest inversion frequencies were observed at a pH value of 7.6–8.0 and 150 mM NaCl in plasmid pTH13166 (**Figures S6A,C**) and at a pH value of 8.0 and 125 mM NaCl in plasmid pTH13170 (**Figures S6B,D**).

PsrA Relies on Mg^{2+} to Achieve Efficient IR1 Inversion Reactions

Current understanding of the mechanisms of site-specific recombination, based on *in vitro* studies, proposes that site-specific recombinases do not require high-energy factors and/or divalent cations for catalysis (Mack et al., 1992; Grindley et al., 2006; Fan, 2012). However, here we found that this does not seem to be the case for the pneumococcal tyrosine recombinase PsrA. First, we analyzed the effect of various divalent cations (Mg^{2+}, Mn^{2+}, Ca^{2+}, Ba^{2+}, Fe^{2+}, and Zn^{2+}) on the efficiency of the PsrA-mediated IR1 inversion reactions. Compared to the reaction without divalent cations, the highest frequency of DNA inversions was observed in the presence of 5 mM $MgCl_2$ (∼335-fold; plasmid pTH13166) (**Figure 9A**). Similar

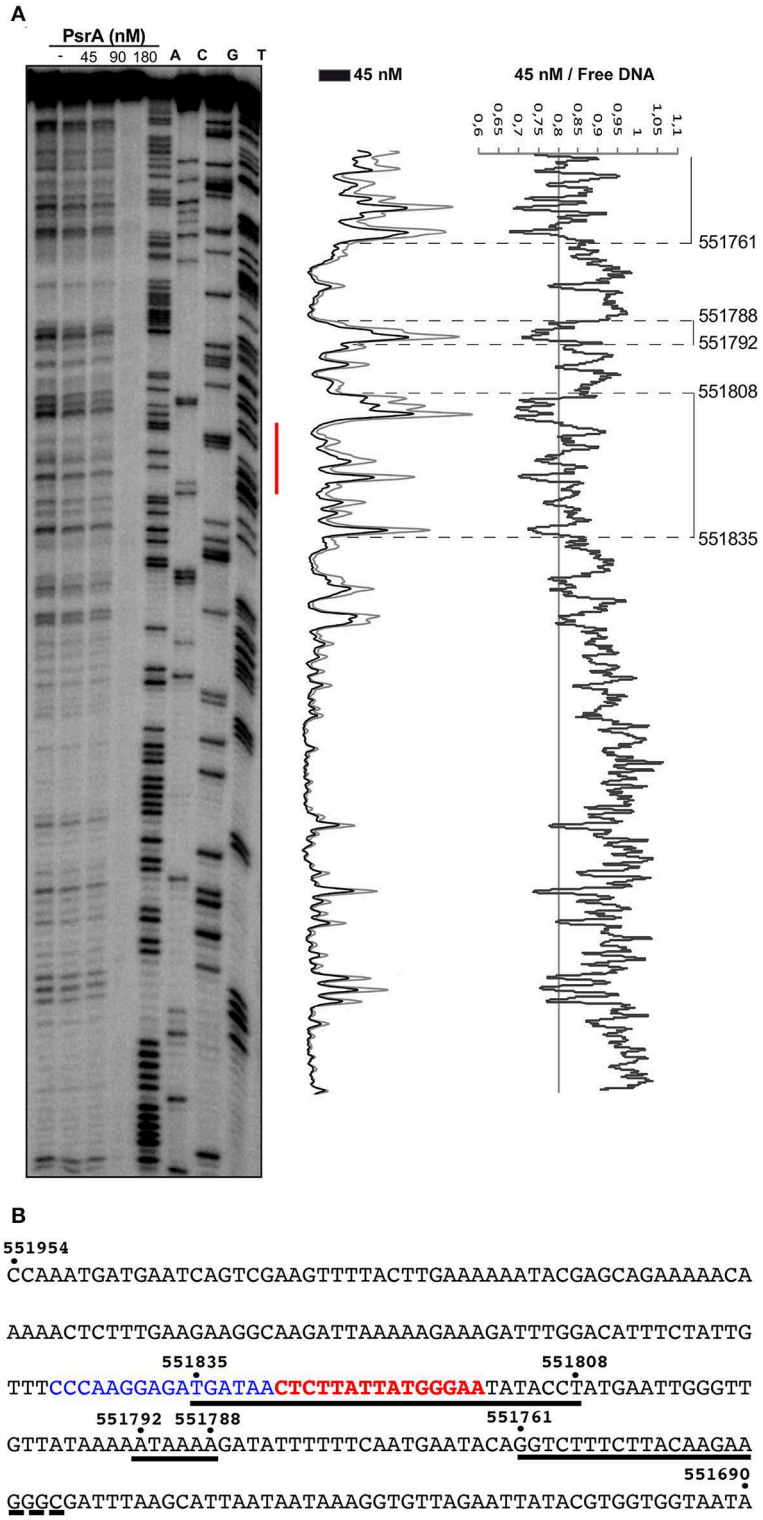


FIGURE 6 | DNase I footprints of complexes formed by PsrA on the IR1.1 DNA. **(A)** ^{32}P-labeled IR1.1 DNA (6 nM) was incubated with the indicated concentrations of PsrA (formation of complexes) and then digested with DNase I. Dideoxy-mediated chain termination sequencing reactions were run in the same gel (lanes A, C, G, T). The red bar indicates the position of the IR1.1 repeat. Densitometer scans corresponding to DNA with PsrA (45 nM; black line) and DNA without PsrA (gray line) are shown. The ratio of intensities between both scans is also shown. **(B)** Nucleotide sequence of the IR1.1 DNA fragment (coordinates 551954-551690 of the ST556 genome). Regions protected against DNase I digestion in the presence of PsrA (45 nM) are indicated with black lines. The IR1.1 repeat sequence (in red) and the upstream adjacent sequence (in blue) are required for PsrA-mediated IR1 inversions *in vivo* (Li et al., 2019).

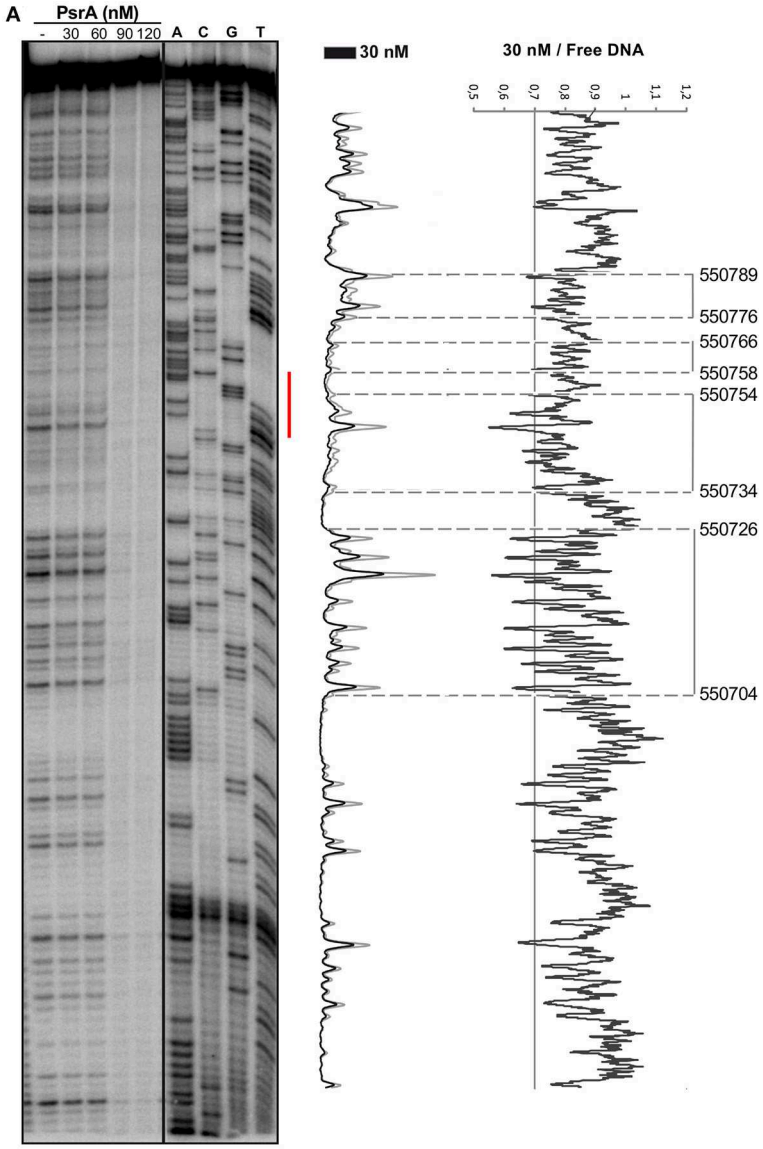


FIGURE 7 | DNase I footprints of complexes formed by PsrA on the IR1.2 DNA. **(A)** ^{32}P-labeled IR1.2 DNA (4 nM) was incubated with the indicated concentrations of PsrA (formation of complexes) and then digested with DNase I. Dideoxy-mediated chain termination sequencing reactions were run in the same gel (lanes A, C, G, T). The red bar indicates the position of the IR1.2 repeat. Densitometer scans corresponding to DNA with PsrA (30 nM; black line) and DNA without PsrA (gray line) are shown. The ratio of intensities between both scans is also shown. **(B)** Nucleotide sequence of the IR1.2 DNA fragment (coordinates 550619–550883 of the ST556 genome). Regions protected against DNase I digestion in the presence of PsrA (30 nM) are indicated with black lines. The IR1.2 repeat sequence (in red) and the upstream adjacent sequence (in blue) are required for PsrA-mediated IR1 inversions *in vivo* (Li et al., 2019).

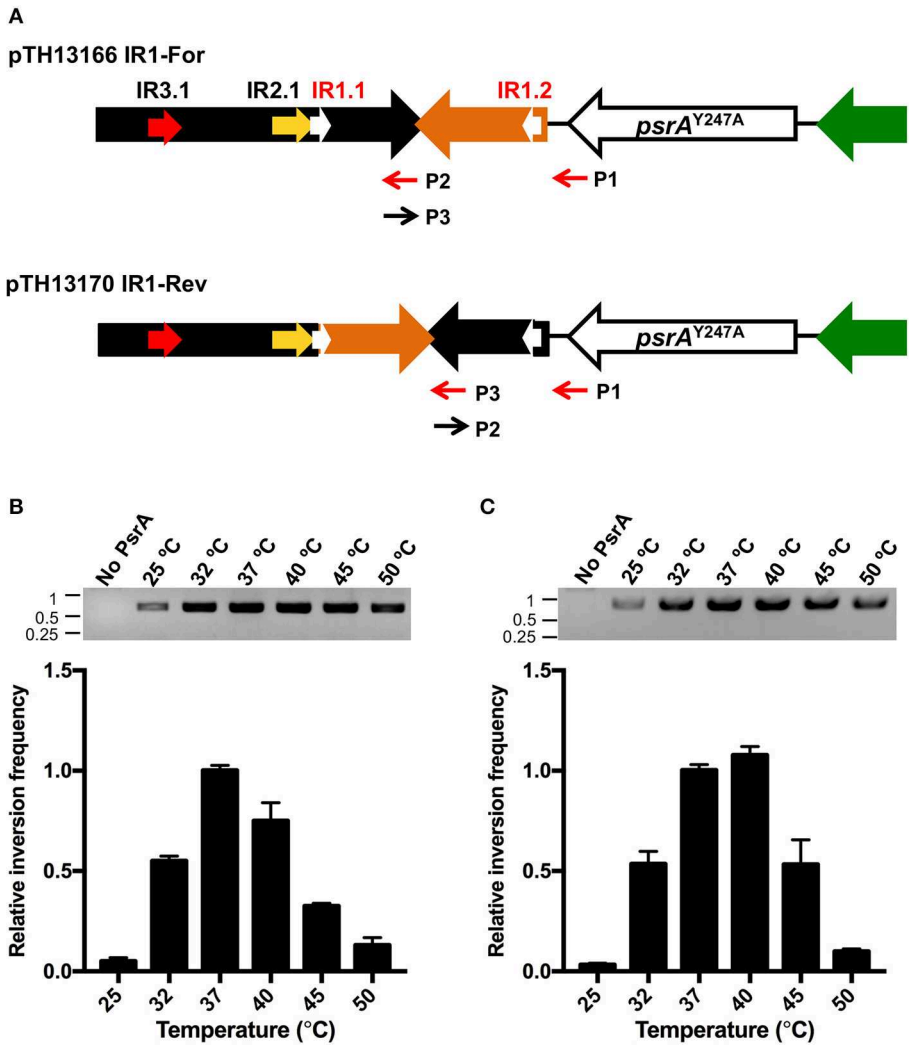


FIGURE 8 | PsrA generates IR1 inversions *in vitro*. **(A)** Relevant features of the plasmids used to analyse PsrA-mediated inversions between IR1.1 and IR1.2. Primers used to detect IR1 inversions are indicated with red arrows. Primers P2 and P3 have complementary sequences. **(B)** Effect of temperature on the formation of IR1 inversions (from "For" to "Rev") using plasmid pTH13166. PCR assays (agarose gel) and qPCR assays (bar graph) were carried out using the P1/P2 primers. The position of DNA molecular weight markers (in kb) is indicated on the left of the gel. **(C)** Effect of temperature on the formation of IR1 inversions (from "Rev" to "For") using plasmid pTH13170. PCR assays (agarose gel) and qPCR assays (bar graph) were performed using the P1/P3 primers.

results were obtained with plasmid pTH13170 (**Figure S7**). The frequency of IR1 inversions increased also to some extent in the presence of 5 mM $CaCl_2$, 5 mM $BaCl_2$, and 0.5 mM $FeCl_2$ (14-, 37-, and 102-fold in pTH13166, and 2-, 3-, and 13-fold in pTH13170) (**Figure 9A** and **Figure S7**A). We also examined the effect of various concentrations of $MgCl_2$ (0.5–50 mM) on the efficiency of the PsrA-mediated IR1 inversion reaction. The maximum frequency of inversions was observed at 7.5 mM of $MgCl_2$ in both plasmids (**Figure 9B** and **Figure S7B**). By EMSA experiments, we found that neither $MgCl_2$ (5 mM) nor $FeCl_2$ (0.5 mM) influences the binding of PsrA to DNA fragments that contain the IR1.2 repeat (**Figure S8**). Thus, Mg^{2+} might play a significant role in PsrA catalysis.

PsrA Is Able to Generate IR2- and IR3-Mediated Inversions *in vitro*

To investigate whether PsrA was able to generate IR2- and IR3-mediated inversions *in vitro*, we constructed four recombinant plasmids (**Figure 10A**). Plasmids pTH13341 (IR2-For) and pTH13344 (IR2-Rev) lack the IR1.2 and IR3.2 repeats. They were used to detect and quantify the occurrence of IR2 inversions by PCR and qPCR using primers P4/P5 (from "For" to "Rev," pTH13341) or P5/P6 (from "Rev" to "For," pTH13344). Plasmids pTH13337 (IR3-For) and pTH13339 (IR3-Rev) lack the IR1.2 and IR2.2 repeats. They allowed us to examine the occurrence of IR3 inversions using primers P7/P8 (from "For" to "Rev," pTH13337) or P8/P9 (from "Rev" to "For," pTH13339). In the inversion reactions shown in **Figures 10B** (PCR assays) and **10C**

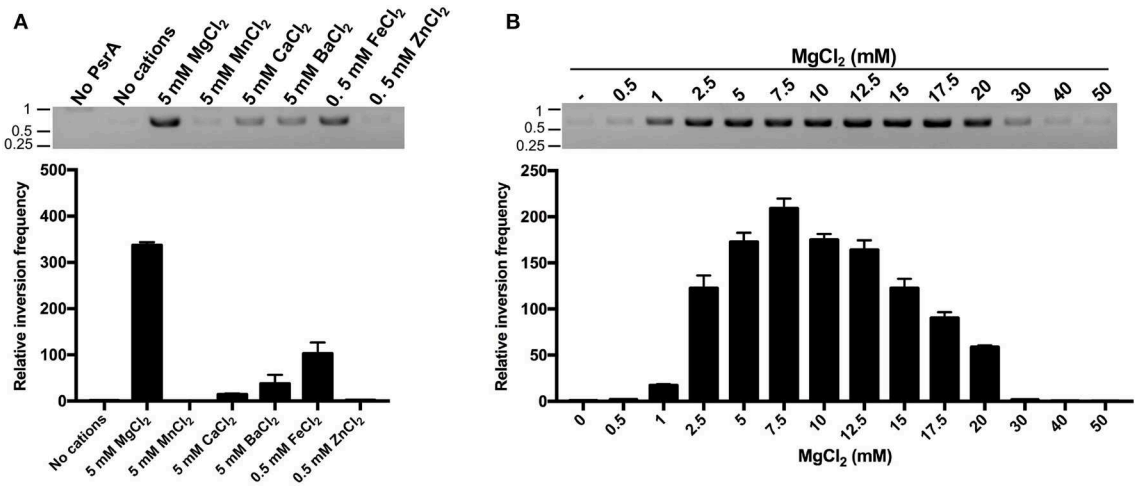


FIGURE 9 | Mg^{2+} enhances the frequency of PsrA-mediated IR1 inversions. **(A)** Relative frequency of PsrA-mediated IR1 inversions in the presence of different divalent cations (bar graph). Plasmid pTH13166 (1 nM) was incubated with PsrA (125 nM) at 37°C for 1 h. IR1 inversions were detected by PCR assays (agarose gel) and by qPCR assays (bar graph) using the P1/P2 primers (see **Figure 8A**). The position of DNA molecular weight markers (in kb) is indicated on the left of the gel. **(B)** Relative frequency of PsrA-mediated IR1 inversions using plasmid pTH13166 and the indicated concentrations of $MgCl_2$ (bar graph). IR1 inversions were detected by PCR assays (agarose gel) and by qPCR assays (bar graph).

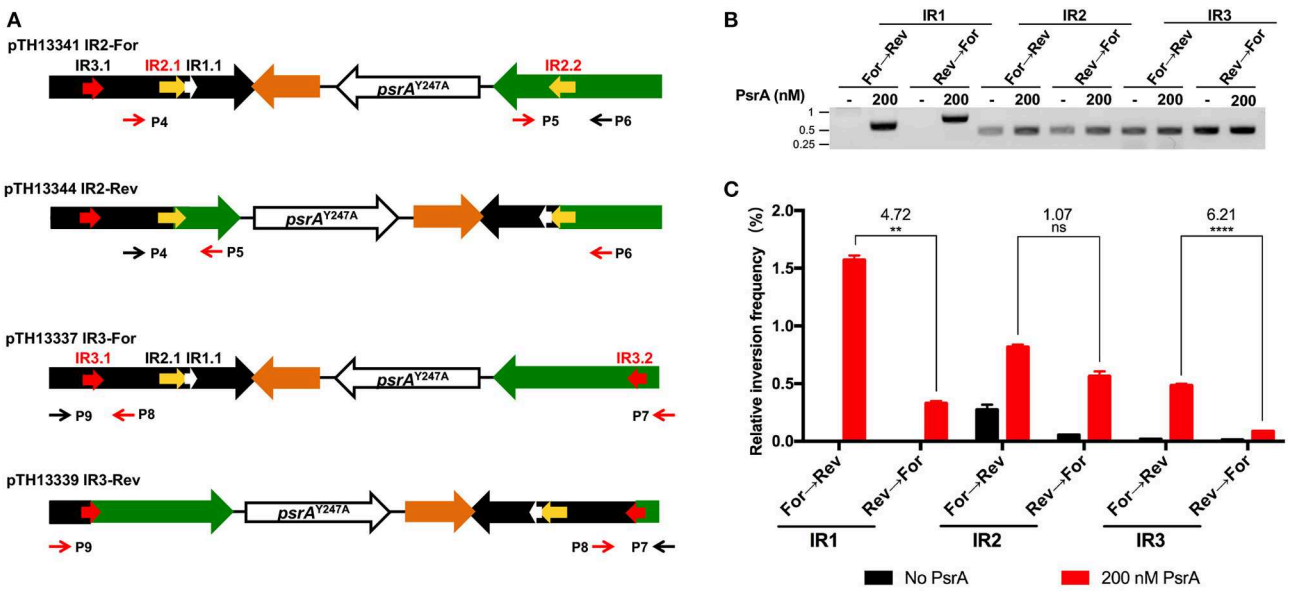


FIGURE 10 | PsrA generates IR2 and IR3 inversions *in vitro*. **(A)** Relevant features of the plasmids used to analyse PsrA-mediated inversions between IR2.1 and IR2.2 and between IR3.1 and IR3.2. Primers P4/P5 and primers P5/P6 were used to detect IR2 inversions in pTH13341 and pTH13344, respectively. Primers P7/P8 and primers P8/P9 were used to detect IR3 inversions in pTH13337 and pTH13339, respectively. **(B)** IR1, IR2, and IR3 inversions detected by PCR assays (agarose gel). Plasmids used to analyse IR1 inversions are shown in **Figure 8A**. Plasmid DNA (1 nM) was incubated with PsrA (200 nM) at 37°C for 2 h in the presence of 7.5 mM $MgCl_2$. The position of DNA molecular weight markers (in kb) is indicated on the left of the gel. **(C)** Relative frequency of IR1, IR2, and IR3 inversions determined by qPCR assays (bar graph). The result of inversion frequency was analyzed using two-tailed unpaired Student's *t* tests (means ± SEM). Statistical significance is defined by *P* values of <0.01 (**) or <0.0005 (****).

(qPCR assays), we used 200 nM of PsrA since this concentration led to the highest frequency of IR1 inversions when 1 nM plasmid DNA, 125 mM NaCl and 7.5 mM $MgCl_2$ were used (**Figure S9**). According to our previous *in vivo* studies (Li et al., 2019), PsrA-independent IR2- and IR3-mediated inversions may occur in *E. coli* with low frequency. Thus, the abovementioned plasmids isolated from *E. coli* were expected to have both configurations, "For" and "Rev," as it was further confirmed when *in vitro* inversion reactions were carried out in the absence of PsrA (**Figure 10B**). Nevertheless, in the presence of PsrA (200 nM),

we found a significant increase in the frequency of IR2- and IR3-mediated inversions (**Figure 10C**). Moreover, like in PsrA-mediated IR1 inversions (**Figure 10C**), the frequency of IR3 inversions from "For state" to "Rev sate" was higher than from "Rev state" to "For state," whereas the frequency of IR2 inversions was similar in both directions (**Figure 10C**). These results demonstrated that PsrA catalyzes *in vitro* the formation of IR2- and IR3-mediated inversions in both directions.

DISCUSSION

Site-specific inversions in the *cod* locus of *S. pneumoniae* play an important role in the bacterial lifestyle by generation of epigenetic diversity and phenotypic phase variation in colony opacity that, in turn, leads to the adaptation of the pneumococci in colonization and virulence processes (Manso et al., 2014; Li et al., 2016; De Ste Croix et al., 2017). The gene encoding the PsrA Tyr-recombinase is included within the *cod* locus. *in vivo* studies showed that PsrA plays an essential role in promoting *hsdS* inversions between a pair of IR1 or IR1-like inverted repeats, with the exception of low-frequency sequence- and RecA-independent "spontaneous" inversions (Li et al., 2019). In this work, we have purified an untagged form of the PsrA protein and analyzed its DNA-binding capacities using different target DNAs. Furthermore, we have designed an approach to detect and quantify the DNA inversions generated by PsrA *in vitro*. This assay has allowed us to define the optimal conditions required for PsrA-mediated IR1 inversions.

In DNA-binding proteins, the location of their target sites on the genome is often preceded by an extensive search process, which results in transitory non-specific interactions between protein and DNA (Halford and Marko, 2004; Marcovitz and Levy, 2013). In the facilitated diffusion model (Berg et al., 1981; von Hippel and Berg, 1989), the protein first collides with the DNA at a random site (non-specific interactions) and then either locates its target site (specific interactions) or dissociates from the DNA. According to this model, sliding, and hopping events performed by the protein along the DNA regions surrounding the random site contribute to the search of the specific binding site. Our results suggest that PsrA might find its preferred target sites by a mechanism that involves initial binding at random DNA sites. First, by gel retardation assays, we found that PsrA binds to both linear dsDNAs and supercoiled DNAs containing or not inverted repeats. This behavior differs from other well-known Tyr-recombinases, such as Cre (Andrews et al., 1986; de Vargas et al., 1988; McCusker et al., 2007). However, further DNase I footprinting experiments performed with linear dsDNAs (265-bp) showed that PsrA binds to regions that contain either the IR1.1 or the IR1.2 repeat. On binding to these regions, PsrA generates higher-order complexes in which the DNA is not accessible to DNase I cleavage. Such nucleoprotein complexes do not enter the native PAA or agarose gels used in our EMSA experiments. Thus, to catalyze IR1-mediated inversions, PsrA could recognize the IR1 sequences and/or particular IR1-mediated structures. This conclusion is in agreement with our previous *in vivo* studies, which showed that PsrA-mediated IR1 inversions require both the 15-bp inverted repeats (IR1.1 and IR1.2) and the upstream adjacent sequences of both repeats (Li et al., 2019). Some reports support that site-specific recombinases should interact directly with inversely repeated sequences, and the position of the nucleophilic attack should be within the repeated sequence (Tong et al., 2014).

We have set up an *in vitro* procedure to detect PsrA-mediated DNA inversions. The procedure involves the use of an untagged form of PsrA and supercoiled plasmid DNA as target DNA. Using this *in vitro* assay, we have demonstrated that PsrA is able to catalyze three independent DNA inversion events mediated by IR1, IR2, and IR3 repeats, respectively. These findings support that PsrA locates specific binding sites on supercoiled DNAs and thus promote DNA inversions. The optimal conditions required for IR1-mediated inversions were found to be 37°C, pH 7.6, 125 mM NaCl, and 7.5 mM $MgCl_2$. Interestingly, the frequency of PsrA-mediated IR1 inversions increased greatly in the presence of Mg^{2+}. This cation could be partially replaced by Ba^{2+}, Ca^{2+}, and Fe^{2+} but not by Mn^{2+} or Zn^{2+}. We hypothesize that Mg^{2+} could play a critical role in the catalytic reaction, although we cannot rule out the possibility of an effect on the PsrA oligomerization state and/or on the binding of PsrA to supercoiled plasmid DNAs. Usually, the active form of the Tyr-recombinases is a homodimer (Guo et al., 1997; Grindley et al., 2006). In the case of Cre, Mg^{2+} was shown to induce dimer formation and to stabilize Cre-Cre interactions in solution (Abremski and Hoess, 1984), even though the efficiency of the Cre-mediated catalytic reaction was Mg^{2+}-independent (Ghosh et al., 2007). Additionally, structural studies have shown that mono- and multivalent ions influence the configuration of DNA molecules, and cations like Mg^{2+} can induce DNA condensation (Baumann et al., 1997; Bloomfield, 1997).

The three-dimensional structures of various recombinase-DNA complexes have been solved and no cations have been identified in the interface of the active site (Guo et al., 1997; Chen et al., 2000; Biswas et al., 2005). In addition to the catalytically active Tyr residue, five highly conserved amino acids participate in the catalytic center, namely one Lys, two Arg, and two His residues (or one His and one Trp in Cre and Flp). These residues generate a positively charged pocket that accommodates the negatively charged scissile phosphate group (Chen and Rice, 2003). Furthermore, several reports have shown that 'classical' Tyr-recombinases do not need a cation for catalysis (Ghosh et al., 2007; Fan, 2012). Thus, and to the best of our knowledge, the pneumococcal PsrA protein would represent the first instance of a Mg^{2+}-dependent Tyr-recombinase. We speculate that Mg^{2+} could be placed either at the active center or at another domain of the protein; alternatively, Mg^{2+} could favor a conformational change in the DNA target to increase the catalytic efficacy of PsrA. Finally, whether accessory host factor(s) participates in the DNA inversions mediated by PsrA *in vivo* and, if so, how would they work is still unknown to us. Previous reports have shown that the two scenarios (independence or dependence of host factors) do exist: Cre recombinase from phage P1 and Flp recombinase from the 2 micron plasmid of *S. cerevisiae* do not need accessory factors (Grindley et al., 2006), whereas λ integrase relies on host factors like IHF, Fis, and Xis

to promote recombination and to determine the directionality of the process (Landy, 1989). This would be also a matter of future exploration.

MATERIALS AND METHODS

Bacterial Strains and Chemical Reagents

All strains used are listed in **Table S1**. *Escherichia coli* strains were grown in Luria-Bertani (LB) broth, or on LB agar plates, at 37°C under aerobic conditions. To overproduce PsrA, *E. coli* cells were grown in TY (tryptone-yeast extract) medium. Kanamycin (30 μg/ml) was added when needed. All primers were synthesized by Sangon Biotech (Beijing, China) and are listed in **Table S2**. PrimeSTAR® HS DNA Polymerase (TaKaRa Bio Inc, Kyoto, Japan) was used for polymerase chain reaction (PCR). PCR products were purified with the QIAquick PCR purification kit (QIAGEN, Hilden, Germany). All restriction enzymes were purchased from New England Biolabs (NEB; Beijing, China). All chemical reagents used in this work were purchased from Sigma (Shanghai, China) unless otherwise mentioned. Sanger sequencing data were obtained from Ruibiotech Co (Beijing, China) or from Secugen CIB (Madrid, Spain).

Construction of Recombinant Plasmids

To overproduce PsrA, gene *psrA* was cloned into the inducible expression vector pET24b (Novagen, Beijing, China) harbored by *E. coli* BL21 (DE3) strain (a gift of F. W. Studier). This strain has the λDE3 lysogen that carries the gene for T7 RNA polymerase under control of the *lacUV5* promoter (Studier and Moffatt, 1986), which is inducible by isopropyl β-D-1-thiogalactopyranoside (IPTG). Plasmid pET24b carries the Ø10 promoter that is recognized by the T7 RNA polymerase. Gene *psrA* was amplified with primers pr15159 and pr15160 from ST606, digested with *Nde*I and *Xho*I, and inserted into the multiple cloning site (MCS) of pET24b to generate pTH12647 (pET24b::*psrA*).

Recombinant plasmids used for *in vitro* DNA inversion assays were constructed as described in **Table S3**. A pair of plasmids containing either the "For" or the "Rev" state were constructed for each pair of inverted repeats (IR1, IR2, and IR3) in a blunt-end cloning vector pEASY-blunt zero (TransGen Biotech, Beijing, China), which was then introduced into *E. coli* DH5α competent cells by heat shock as described before (Li et al., 2019). In each recombinant plasmid, only one pair of IRs was preserved. For instance, IR2.2 and IR3.2 were removed from the inserted S region of the *cod* locus in the pneumococcal strain TH6552 ($psrA^{Y247A}$) (Li et al., 2019) to obtain TH8563 ($psrA^{Y247A}$ΔIR2.2IR3.2) via JC1-insertion and counter selection. Briefly, upstream and downstream fragments were amplified with pr9803/pr11594 and pr10106/pr11595, digested with *Xba*I and *Xho*I, respectively, and then ligated with *Xba*I/*Xho*I-digested JC1 (modified Janus cassette, amplified with pr9840/pr1098) (Li et al., 2016); this generated element was introduced into *S. pneumoniae* ST606 by natural transformation to obtain TH8551 ($psrA^{Y247A}$ΔIR2.2IR3.2::JC1) as previously described (Li et al., 2019); next, the two amplicons were both digested with *Xba*I, then ligated together, and transformed into TH8551 to obtain TH8563, whose chromosomal DNA was used as the template to construct pTH13166 (IR1-For) and pTH13170 (IR1-Rev), in which the orientations of the S region between IR1.1 and IR1.2 were opposite, represented as S1 and S2 configurations. Similarly, pTH13341 (IR2-For) and pTH13344 (IR2-Rev) were used for testing IR2 inversions, with IR1.2 and IR3.2 deleted, and S configurations were S1 and S3; pTH13337 (IR3-For) and pTH13339 (IR3-Rev), used for IR3 inversions, have S1 and S4 configurations, respectively, with IR1.2 and IR2.2 deleted. Pneumococcal chromosomal DNA was isolated by the DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany) as previously described (Feng et al., 2014). Plasmid DNA was prepared with HiPure Plasmid Plus Midi Kit (Magen, Beijing, China) according to manufacturer's instructions. All insertion elements were verified by Sanger sequencing.

Overproduction and Purification of PsrA

To overproduce PsrA, *E. coli* BL21 (DE3) cells harboring plasmid pTH12647 were grown in TY medium supplemented with kanamycin (30 μg/ml) at 37°C and with aeration to mid-log phase [optical density at 600 nm (OD_{600}) of 0.4–0.5]. Expression of the *psrA* gene was induced with IPTG (0.25 mM). After 25 min, bacterial cells were incubated with rifampicin (0.2 mg/ml) for 1 h. Rifampicin inhibits *E. coli* RNA polymerase but not T7 RNA polymerase. Cells were harvested by centrifugation, washed twice with ice-cold buffer S (10 mM Tris-HCl, pH 7.6, 1 mM DTT, 5% glycerol) containing 300 mM NaCl, and stored at −80°C. To purify PsrA, we designed a procedure based on previous studies (Solano-Collado et al., 2013; Ruiz-Cruz et al., 2018). Essentially, bacterial cells were concentrated (40x) in buffer S containing 300 mM NaCl and a protease inhibitor cocktail (Roche, Shanghai, China). Cells were disrupted using a French Pressure Cell, and the whole-cell extract was centrifuged to sediment cell debris. Nucleic acids were precipitated with 0.2% polyethyleneimine (PEI) (on ice for 30 min). After centrifugation, proteins recovered in the supernatant were precipitated with 70% saturated ammonium sulfate (on ice for 30 min). Precipitated proteins were collected by centrifugation, dissolved in buffer S containing 300 mM NaCl, and dialyzed against the same buffer at 4°C. Proteins were then loaded onto a HiTrap™ Heparin HP column (GE Healthcare Bio-Sciences Corp., NJ, USA), and proteins with an affinity for heparin were eluted using a linear gradient of NaCl (300–800 mM). Fractions containing PsrA were identified by SDS-polyacrylamide (12%) gel electrophoresis (SDS-PAGE), pooled and dialyzed against buffer S containing 300 mM NaCl. PsrA preparations were then concentrated with Microsep™ Advance Centrifugal Device (3K MWCO) (Pall Life Sciences, NY, USA). Protein concentration was determined with the BCA assay kit (Solarbio, Beijing) and using a NanoDrop ND-2000 Spectrophotometer (Thermo Scientific, USA).

Radiolabeling of DNA Fragments

Primers pr14684 and pr14687 were radioactively labeled at the 5′ end using T4 polynucleotide kinase (T4 PNK, NEB) and [γ-^{32}P]-ATP (PerkinElmer Inc, USA). Reactions (25 μl)

contained 25 pmol primer, 10 units T4 PNK, 50 pmol [γ-^{32}P]-ATP (3,000 Ci/mmol, 10 mCi/ml), and kinase buffer (1x), which was provided by the supplier. After 30 min at 37°C, 10 units T4 PNK was added and reactions were incubated again at 37°C for 30 min. T4 PNK was inactivated (65°C, 20 min) and non-incorporated [γ-^{32}P]-ATP was removed by passing through an Illustra MicroSpin™ G-25 column (GE Healthcare). 5′-radiolabeled pr14684 and non-labeled pr14685 were used to generate the radiolabeled IR1.1 dsDNA fragment by PCR amplification; 5′-radiolabeled pr14687 and non-labeled pr14686 were used to generate the radiolabeled IR1.2 dsDNA fragment. The labeled pr14684 and pr14687 primers were also used to perform sequencing reactions with a 1067-bp IR1.1 fragment (pr14682/pr14685) and a 1009-bp IR1.2 fragment (pr14686/pr15159), respectively, by USB® Sequenase Version 2.0 DNA Sequencing Kit (Affymetrix Inc, Ohio, USA).

Electrophoretic Mobility Shift Assays (EMSA)

The DNA fragments used for EMSA were amplified with the following primer pairs: IR3.1-pr14682/pr14683, IR1.1-pr14684/pr14685, IR1.2-pr14686/pr14687, and *hsdM*-pr14691/pr14692 from ST606. Short-length dsDNA fragments of 26-bp (pr15163/pr15164) were obtained by oligonucleotide annealing (Solano-Collado et al., 2013). EMSA reactions were basically performed in the following conditions: 30 mM Tris-HCl, pH 7.6, 1 mM DTT, 0.25 mM EDTA, 125 mM NaCl, 0.5 mg/ml bovine serum albumin (BSA), 10 mM $MgCl_2$, 1 mM $CaCl_2$, 1.25% glycerol, 10 nM non-labeled DNA or 2 nM ^{32}P-labeled IR1.2 DNA, and various concentrations of PsrA. Reactions were incubated at room temperature for 20 min, followed by mixing with 10 × BXGE loading buffer (0.25% bromophenol blue, 0.25% xylene cyanol, 60% glycerol, and 10 mM EDTA). Reactions with non-labeled DNA were analyzed by electrophoresis on native PAA (6%) gels in TBE (Tris-borate-EDTA, pH 8.3) buffer or on agarose (0.6% to 2%) gels in TAE (Tris-acetate-EDTA, pH 8.3) buffer. Non-labeled DNA was visualized with Gel-Red (Biotium) staining under a ChemiDoc™ XRS+ System (Bio-Rad) and quantified with ImageLab software (Bio-Rad). Reactions with radiolabeled DNA were also analyzed by electrophoresis on native PAA (6%) gels, but DNA was visualized using a Fujifilm Image Analyzer FLA-3000 and quantified using the Quantity One software (Bio-Rad) (Solano-Collado et al., 2016). Competitive EMSA reactions were performed with 2 nM ^{32}P-labeled IR1.2 DNA and increasing concentrations of non-labeled calf thymus DNA (Thermo Fisher Scientific, USA). Dissociation of PsrA-IR1.2 complexes was achieved by adding non-labeled IR1.2 or a 266-bp control DNA (amplified with pr15161 and pr15162 from *E. faecalis* OG1RF (a gift from Ana Moreno-Blanco at CIB, CSIC, Spain), and incubating for another 5 min before analyzing by 6% PAGE.

DNase I Footprinting Assays

PsrA-DNA binding reactions (8 μl) contained 30 mM Tris-HCl, pH 7.6, 1 mM DTT, 0.25 mM EDTA, 125 mM NaCl, 0.5 mg/ml BSA, 10 mM $MgCl_2$, 1 mM $CaCl_2$, 1.25% glycerol, 4 nM radiolabeled IR1.2 DNA or 6 nM radiolabeled IR1.1 DNA, and different concentrations of PsrA. Reactions were incubated at room temperature for 20 min. Then, protein-DNA complexes were treated with DNase I (0.01 units) for 5 min at the same temperature. DNase I digestion was stopped by adding 1 μl of 250 mM EDTA. Four microlitres of loading buffer (80% formamide, 1 mM EDTA, 10 mM NaOH, 0.1% bromophenol blue, and 0.1% xylene cyanol) was added to the reaction mixtures. Samples were heated at 95°C for 5 min, and immediately chilled on ice before loading onto 8 M urea-6% polyacrylamide gels. After running, gels were dried and exposed to a Phosphorimager screen (Fujifilm, Japan). The radioactive intensity was visualized by a Fujifilm Image Analyser FLA-3000 and was quantified using the Quantity One software (Bio-Rad) (Solano-Collado et al., 2013; Ruiz-Cruz et al., 2018).

In vitro DNA Inversion Assays

The basic components of the *in vitro* DNA inversion reactions are listed as followed: 25 mM Tris-HCl, pH 7.6, 1 mM DTT, 1 mM EDTA, 125 mM NaCl, 5 mM $MgCl_2$, 1 nM DNA, and 125 nM PsrA, referring to previous studies on several classical tyrosine recombinases (Parsons et al., 1990; Ringrose et al., 1998; Sandal et al., 2001; Tong et al., 2014). Reactions (20 μl) were started by adding PsrA protein, and performed at 37°C for 1 h. 80 μl ddH_2O was then added to stop the reaction and reactions were 20-fold diluted in ddH_2O for subsequent analysis and quantification by PCR and quantitative real-time PCR (q-PCR), respectively. Primer pairs used for testing DNA inversions of different substrates were as followed: P1/P2-pTH13366; P1/P3-pTH13370; P4/P5-pTH13341; P5/P6-pTH13344; P7/P8-pTH13337, and P8/P9-pTH13339. PCR was performed with the primer pairs mentioned above and analyzed by 1% agarose gel electrophoresis. qPCR was used to calculate the relative inversion frequency by normalizing each ΔC_T value (C_T of inversion reaction minus C_T of reference gene *era*, amplified with primers pr7932 and pr10129) with ΔC_T of the reaction at 37°C for 1 h containing basic components.

To calculate absolute inversion frequency of IR1, IR2, and IR3-mediated inversions *in vitro*, two pairs of primers were used to amplify "For" and "Rev" states, respectively. pTH13366: For-P1/P3, Rev-P1/P2; pTH13370: For-P1/P2, Rev-P1/P3; pTH13341: For-P5/P6, Rev-P4/P5; pTH13344: For-P4/P5, Rev-P5/P6; pTH13337: For-P8/P9, Rev-P7/P8; pTH13339: For-P7/P8, Rev-P8/P9. Like in qPCR to calculate relative inversion frequency, individual C_T values of For and Rev state were first normalized by subtracting the C_T of the reference gene *era* to obtain ΔC_T. ΔC_T values of For and Rev states were normalized again with the For state to obtain $\Delta\Delta C_{T-For}$ and $\Delta\Delta C_{T-Rev}$. The final amplification level of For and Rev states was $2^{-\Delta\Delta CT-For}$ and $2^{-\Delta\Delta CT-Rev}$, and the absolute inversion frequency was the ratio of $2^{-\Delta\Delta CT-Rev}$ in the total value of $2^{-\Delta\Delta CT-For}$ and $2^{-\Delta\Delta CT-Rev}$.

Statistical Analysis

Statistical analysis was performed using the GraphPad Prism 7.0 software (San Diego, CA). The qPCR data of *in vitro* inversion frequency were also analyzed using two-tailed unpaired

Student's *t* tests (means ± standard error of the mean). Statistical significance is defined by *P* values of < 0.05 (*), < 0.01 (**), < 0.001(***), or < 0.0005 (****).

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

This study was designed by ME, AB, and J-RZ. Laboratory work was performed by JL, JW, and SR-C. The results were analyzed and discussed between all authors. The first draft of the manuscript was elaborated by JL, JW, ME, and AB. Several copies of the different versions of the manuscript were circulated among all the authors. The final version of the manuscript was done by AB and J-RZ. All authors read and approved the final manuscript.

FUNDING

Work supported by grants from National Natural Science Foundation of China (Nos. 31728002, 81671972, 31530082, and 31820103001) to J-RZ and Grant BIO2016-76412-C2-2-R AEI/FEDER from the Ministry of Science, Innovation and Universities, Spain, to AB.

ACKNOWLEDGMENTS

We thank Ana Moreno-Blanco for her help in several stages of the experiments, and Dr. Virtu Solano-Collado for providing the 26-bp DNA fragment.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2020.00043/full#supplementary-material

REFERENCES

Abraham, J. M., Freitag, C. S., Clements, J. R., and Eisenstein, B. I. (1985). An invertible element of DNA controls phase variation of type 1 fimbriae of *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 82, 5724–5727. doi: 10.1073/pnas.82.17.5724

Abremski, K., and Hoess, R. (1984). Bacteriophage P1 site-specific recombination. Purification and properties of the Cre recombinase protein. *J. Biol. Chem.* 259, 1509–1514.

Andrews, B. J., McLeod, M., Broach, J., and Sadowski, P. D. (1986). Interaction of the FLP recombinase of the *Saccharomyces cerevisiae* 2 micron plasmid with mutated target sequences. *Mol. Cell. Biol.* 6, 2482–2489. doi: 10.1128/MCB.6.7.2482

Baumann, C. G., Smith, S. B., Bloomfield, V. A., and Bustamante, C. (1997). Ionic effects on the elasticity of single DNA molecules. *Proc. Natl. Acad. Sci. U.S.A.* 94, 6185–6190. doi: 10.1073/pnas.94.12.6185

Berg, O. G., Winter, R. B., and von Hippel, P. H. (1981). Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. models and theory. *Biochemistry* 20, 6929–6948. doi: 10.1021/bi00527a028

Bhugra, B., Voelker, L. L., Zou, N., Yu, H., and Dybvig, K. (1995). Mechanism of antigenic variation in *Mycoplasma pulmonis*: interwoven, site-specific DNA inversions. *Mol. Microbiol.* 18, 703–714. doi: 10.1111/j.1365-2958.1995.mmi_18040703.x

Biswas, T., Aihara, H., Radman-Livaja, M., Filman, D., Landy, A., and Ellenberger, T. (2005). A structural basis for allosteric control of DNA recombination by lambda integrase. *Nature* 435, 1059–1066. doi: 10.1038/nature03657

Bloomfield, V. A. (1997). DNA condensation by multivalent cations. *Biopolymers* 44, 269–282. doi: 10.1002/(SICI)1097-0282(1997)44:3<269::AID-BIP6>3.0.CO;2-T

Chen, Y., Narendra, U., Iype, L. E., Cox, M. M., and Rice, P. A. (2000). Crystal structure of a Flp recombinase-holliday junction complex: assembly of an active oligomer by helix swapping. *Mol. Cell* 6, 885–897. doi: 10.1016/S1097-2765(00)00086-1

Chen, Y., and Rice, P. A. (2003). New insight into site-specific recombination from Flp recombinase-DNA structures. *Annu. Rev. Biophys. Biomol. Struct.* 32, 135–159. doi: 10.1146/annurev.biophys.32.110601.141732

Darmon, E., and Leach, D. R. (2014). Bacterial genome instability. *Microbiol. Mol. Biol. Rev.* 78, 1–39. doi: 10.1128/MMBR.00035-13

de Ste Croix, M., Chen, K. Y., Vacca, I., Manso, A. S., Johnston, C., Polard, P., et al. (2019). Recombination of the phase-variable spnIII locus is independent of all known pneumococcal site-specific recombinases. *J. Bacteriol.* 201:e00233–19. doi: 10.1128/JB.00233-19

De Ste Croix, M., Vacca, I., Kwun, M. J., Ralph, J. D., Bentley, S. D., Haigh, R., et al. (2017). Phase-variable methylation and epigenetic regulation by type I restriction-modification systems. *FEMS Microbiol. Rev.* 41(Supp. 1), S3–S15. doi: 10.1093/femsre/fux025

de Vargas, L. M., Pargellis, C. A., Hasan, N. M., Bushman, E. W., and Landy, A. (1988). Autonomous DNA binding domains of λ integrase recognize two different sequence families. *Cell* 54, 923–929. doi: 10.1016/0092-8674(88)90107-9

Fan, H. F. (2012). Real-time single-molecule tethered particle motion experiments reveal the kinetics and mechanisms of Cre-mediated site-specific recombination. *Nucleic Acids Res.* 40, 6208–6222. doi: 10.1093/nar/gks274

Feng, Z., Li, J., Zhang, J. R., and Zhang, X. (2014). qDNAmod: a statistical model-based tool to reveal intercellular heterogeneity of DNA modification from SMRT sequencing data. *Nucleic Acids Res.* 42, 13488–13499. doi: 10.1093/nar/gku1097

Ghosh, K., Guo, F., and van Duyne, G. D. (2007). Synapsis of loxP sites by Cre recombinase. *J. Biol. Chem.* 282, 24004–24016. doi: 10.1074/jbc.M703283200

Grindley, N. D., Whiteson, K. L., and Rice, P. A. (2006). Mechanisms of site-specific recombination. *Annu. Rev. Biochem.* 75, 567–605. doi: 10.1146/annurev.biochem.73.011303.073908

Guo, F., Gopaul, D. N., and van Duyne, G. D. (1997). Structure of Cre recombinase complexed with DNA in a site-specific recombination synapse. *Nature* 389, 40–46. doi: 10.1038/37925

Halford, S. E., and Marko, J. F. (2004). How do site-specific DNA-binding proteins find their targets? *Nucleic Acids Res.* 32, 3040–3052. doi: 10.1093/nar/gkh624

Johnson, R. C., Bruist, M. F., and Simon, M. I. (1986). Host protein requirements for *in vitro* site-specific DNA inversion. *Cell* 46, 531–539. doi: 10.1016/0092-8674(86)90878-0

Landy, A. (1989). Dynamic, structural, and regulatory aspects of lambda site-specific recombination. *Ann. Rev. Biochem.* 58, 913–941. doi: 10.1146/annurev.bi.58.070189.004405

Li, J., Li, J. W., Feng, Z., Wang, J., An, H., Liu, Y., et al. (2016). Epigenetic switch driven by DNA inversions dictates phase variation in *Streptococcus pneumoniae*. *PLoS Pathog.* 12: e1005762. doi: 10.1371/journal.ppat.1005762

Li, J., and Zhang, J. R. (2019). Phase variation of *Streptococcus pneumoniae*. *Microbiol. Spectr.* 7:GPP3-0005-2018. doi: 10.1128/microbiolspec.GPP3-0005-2018

Li, J. W., Li, J., Wang, J., Li, C., and Zhang, J. R. (2019). Molecular mechanisms of hsdS inversions in the cod locus of *Streptococcus pneumoniae*. *J. Bacteriol.* 201:e00581. doi: 10.1128/JB.00581-18

Mack, A., Sauer, B., Abremski, K., and Hoess, R. (1992). Stoichiometry of the Cre recombinase bound to the lox recombining site. *Nucleic Acids Res.* 20, 4451–4455. doi: 10.1093/nar/20.17.4451

Manso, A. S., Chai, M. H., Atack, J. M., Furi, L., de Ste Croix, M., Haigh, R., et al. (2014). A random six-phase switch regulates pneumococcal virulence via global epigenetic changes. *Nat. Commun.* 5:5055. doi: 10.1038/ncomms6055

Marcovitz, A., and Levy, Y. (2013). Obstacles may facilitate and direct DNA search by proteins. *Biophys. J.* 104, 2042–2050. doi: 10.1016/j.bpj.2013.03.030

McCusker, M. P., Turner, E. C., and Dorman, C. J. (2007). DNA sequence heterogeneity in Fim tyrosine-integrase recombinase-binding elements and functional motif asymmetries determine the directionality of the fim genetic switch in *Escherichia coli* K-12. *Mol. Microbiol.* 67, 171–187. doi: 10.1111/j.1365-2958.2007.06037.x

Norris, S. J. (2014). *vls* antigenic variation systems of lyme disease *Borrelia:* eluding host immunity through both random, segmental gene conversion and framework heterogeneity. *Microbiol. Spectr.* 2:MDNA3-0038-2014. doi: 10.1128/microbiolspec.MDNA3-0038-2014

Olsen, P. B., and Klemm, P. (1994). Localization of promoters in the fim gene cluster and the effect of H-NS on the transcription of fimB and fimE. *FEMS Microbiol. Lett.* 116, 95–100. doi: 10.1111/j.1574-6968.1994.tb06681.x

Parsons, R. L., Evans, B. R., Zheng, L., and Jayaram, M. (1990). Functional analysis of Arg-308 mutants of Flp recombinase. Possible role of Arg-308 in coupling substrate binding to catalysis. *J. Biol. Chem.* 265, 4527–4533.

Rajeev, L., Malanowska, K., and Gardner, J. F. (2009). Challenging a paradigm: the role of DNA homology in tyrosine recombinase reactions. *Microbiol. Mol. Biol. Rev.* 73, 300–309. doi: 10.1128/MMBR.00038-08

Ringrose, L., Lounnas, V., Ehrlich, L., Buchholz, F., Wade, R., and Stewart, A. F. (1998). Comparative kinetic analysis of FLP and cre recombinases: mathematical models for DNA binding and recombination. *J. Mol. Biol.* 284, 363–384. doi: 10.1006/jmbi.1998.2149

Ruiz-Cruz, S., Moreno-Blanco, A., Espinosa, M., and Bravo, A. (2018). DNA-binding properties of MafR, a global regulator of *Enterococcus faecalis*. *FEBS Lett.* 592, 1412–1425. doi: 10.1002/1873-3468.13032

Sandal, T., Lillehaug, J., SO, D.S., and Ahlgren, R. (2001). Establishment of okadaic acid resistant cell clones using CDNA library expression cloning. *Sci. World J.* 1:42. doi: 10.1100/tsw.2001.149

Shen, X., Gumulak, J., Yu, H., French, C. T., Zou, N., and Dybvig, K. (2000). Gene rearrangements in the vsa locus of *Mycoplasma pulmonis*. *J. Bacteriol.* 182, 2900–2908. doi: 10.1128/JB.182.10.2900-2908.2000

Solano-Collado, V., Huttener, M., Espinosa, M., Juarez, A., and Bravo, A. (2016). MgaSpn and H-NS: two unrelated global regulators with similar DNA-binding properties. *Front. Mol. Biosci.* 3:60. doi: 10.3389/fmolb.2016.00060

Solano-Collado, V., Lurz, R., Espinosa, M., and Bravo, A. (2013). The pneumococcal MgaSpn virulence transcriptional regulator generates multimeric complexes on linear double-stranded DNA. *Nucleic Acids Res.* 41, 6975–6991. doi: 10.1093/nar/gkt445

Studier, F. W., and Moffatt, B. A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* 189, 113–130. doi: 10.1016/0022-2836(86)90385-2

Tong, W., Warren, D., Seah, N. E., Laxmikanthan, G., van Duyne, G. D., and Landy, A. (2014). Mapping the lambda integrase bridges in the nucleoprotein holliday junction intermediates of viral integrative and excisive recombination. *Proc. Natl. Acad. Sci. U.S.A.* 111, 12366–12371. doi: 10.1073/pnas.1413007111

von Hippel, P. H., and Berg, O. G. (1989). Facilitated target location in biological systems. *J. Biol. Chem.* 264, 675–678.

Zhang, J. R., Hardham, J. M., Barbour, A. G., and Norris, S. J. (1997). Antigenic variation in lyme disease borreliae by promiscuous recombination of VMP-like sequence cassettes. *Cell* 89, 275–285. doi: 10.1016/S0092-8674(00)80206-8

Zieg, J., and Simon, M. (1980). Analysis of the nucleotide sequence of an invertible controlling element. *Proc. Natl. Acad. Sci. U.S.A.* 77, 4196–4200. doi: 10.1073/pnas.77.7.4196

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Li, Wang, Ruiz-Cruz, Espinosa, Zhang and Bravo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 06 May 2020
doi: 10.3389/fmolb.2020.00061

ComWΔ6 Stimulates Transcription of Pneumococcal Competence Genes *in vitro*

*Nicole L. Inniss and Donald A. Morrison**

Department of Biological Sciences, The University of Illinois at Chicago, Chicago, IL, United States

The alternative streptococcal σ-factor and master competence regulator, σ^X, stimulates transcription from competence promoters, *in vitro.* As the only known alternative σ-factor in streptococci, σ^X expression is tightly controlled in each species and has a specific physiological role. Pneumococcal transformation also requires the DNA binding activity of ComW, a known σ^X activator and stabilizer. Mutations to the housekeeping σ factor, σ^A, partially alleviate the ComW requirement, suggesting that ComW is a key player in the σ factor swap during the pneumococcal competence response. However, there is no evidence of a direct ComW – RNA polymerase interaction. Furthermore, if and how ComW functions directly at combox promoters is still unknown. Here we report that a DNA-binding ComW variant, ComΔ6, can stimulate transcription from σ^X promoters *in vitro*.

Keywords: alternative sigma factor, *S. pneumoniae* competence, genetic transformation, ComW, transcription

OPEN ACCESS

Edited by:
Manuel Espinosa, Margarita Salas Center for Biological Research (CSIC), Spain

Reviewed by:
Marco Rinaldo Oggioni, University of Leicester, United Kingdom
Gee W. Lau, University of Illinois at Urbana-Champaign, United States

***Correspondence:**
Donald A. Morrison
damorris@uic.edu

Specialty section:
This article was submitted to Molecular Recognition, a section of the journal Frontiers in Molecular Biosciences

Received: *05 February 2020*
Accepted: *24 March 2020*
Published: *06 May 2020*

Citation:
Inniss NL and Morrison DA (2020) ComWΔ6 Stimulates Transcription of Pneumococcal Competence Genes in vitro. Front. Mol. Biosci. 7:61. doi: 10.3389/fmolb.2020.00061

INTRODUCTION

Streptococci are Gram-positive commensal cocci for a wide range of animal hosts, including humans. *Streptococcus pneumoniae* (pneumococcus), an inhabitant of the human nasopharynx, is a causative agent of pneumonia and meningitis in children, the elderly, and immune-compromised individuals. Natural genetic transformation (NGT) was discovered in the pneumococcus (Griffith, 1928; Avery and MacLeod, 1944), and subsequently the genes required for transformation have been found in the genomes of streptococci from all species groups (Håvarstein, 2010). Their genomes encode transcriptional regulators and apparatuses that enable exogenous DNA uptake and genetic recombination. This ability yields highly diverse genomes, making streptococci well equipped for adaptation, as seen with frequent capsular switching and the rapid spread of genes that mediate antibiotic resistance (Andam and Hanage, 2015).

Pneumococci primed for transformation must develop competence, a transient state marked by a shift in both transcriptomic and proteomic profiles, including two waves of new gene transcription, yielding early and late competence gene groups (Peterson et al., 2000; Fontaine et al., 2010; Gao et al., 2014; Khan et al., 2016, 2017). The link between early gene expression and late gene expression is the production of the alternative σ-factor, SigX (σ^X), a member of the σ^{70} family of proteins (Lee and Morrison, 1999, 2000; Luo and Morrison, 2003). All bacteria produce the principal σ factor σ^A, responsible for most gene transcription, and many produce multiple alternative σ factors that promote specific cellular responses, however, σ^X is the only alternative σ factor produced in pneumococcus (Håvarstein, 2010), and *sigX* expression is strictly linked to competence development.

Streptococci utilize tightly regulated quorum-sensing systems to coordinate σ^X- mediated competence, as well as other group behaviors (Cheng et al., 1997; Peterson et al., 2004; Mashburn-Warren et al., 2010, 2012). Species of the angionsus and mitis groups use the ComCDE pathway (Cheng et al., 1997), an auto-regulated two-component signal transduction system (TCSTS), that responds to the Competence Stimulating Peptide (CSP) (Håvarstein et al., 1995). Much of our understanding of streptococcal transformation comes from initial work done with the ComCDE system in pneumococcus.

During exponential growth, the σ^A-RNA polymerase complex (σ^A holo-enzyme) basally transcribes *comC* leading to production of the 41 aa pro-peptide, ComC, 41 (Håvarstein et al., 1995). The ComC pro-peptide is simultaneously cleaved and exported to form mature CSP by the ABC-transporter, ComAB (*comAB*). Extracellular CSP is sensed by the histidine-kinase receptor, ComD, resulting in auto-phosphorylation and activation of its cognate response regulator, ComE, via a phospho-relay event (Håvarstein et al., 1996; Martin et al., 2010, 2012). Activated ComE promotes transcription of *comCDE* and other early competence genes (Ween et al., 1999). ComE-mediated transcription triggers robust competence among the cell population in a positive-feedback auto-regulatory loop. This response culminates in ComE dependent production of σ^X (Ween et al., 1999; Peterson et al., 2004). During competence, the σ^X – RNA polymerase (σ^X holo-enzyme) transcribes from the competence specific combox promoter, directly linking ComCDE quorum sensing to the expression of the transformation regulon (Peterson et al., 2004). Genes under control of the σ^X holo-enzyme include all genes that are required for DNA uptake and recombination.

Although σ^X is the master competence regulator required for late gene expression in pneumococcus, it is not sufficient for high transformation efficiencies. An additional ComE dependent early gene product, ComW, a 9.5 kDa DNA-binding protein (Luo et al., 2004; Inniss et al., 2019), is required for the strongest transformation phenotype possible. *sigX* is expressed independently of ComW (Luo et al., 2004; Piotrowski et al., 2009) and σ^X activity is dependent on the presence of ComW (Sung and Morrison, 2005). SigX's requirement for ComW is supported by observations that mutants lacking ComW (Δ*comW*) transcribe late genes and transform at levels 10- and 10,000-fold below that of wild type cells, respectively (Tovpeko and Morrison, 2014). Furthermore, Δ*comW* cells have decreased σ^X levels (Sung and Morrison, 2005), and ComW and σ^X weakly interact in yeast-2-hybrid assays (Tovpeko et al., 2016), suggesting that they physically interact as protein partners in the cell. Interestingly, mutations to σ^A that presumably decrease its affinity for RNA polymerase (RNAP) alleviate the ComW requirement, indicating that ComW functions on or with the σ^X holo-enzyme. Beyond this, there is currently no evidence of a physical interaction between ComW and RNAP.

A structural model of ComW suggests that is structurally similar to known σ factors, and ComW binds to DNA, non-specifically (Inniss et al., 2019). DNA binding does not require residues ^{73}RGFISC^{78}, as demonstrated with the truncated variant, ComWΔ6, although ComWΔ6 producing cells transform at levels slightly below wild type cells (Inniss et al., 2019). Instead, the conserved ^{38}LxYYLxR^{44} motif is important for DNA binding, and L42A and R44A mutations in the ^{38}LxYYLxR^{44} motif decrease transformation efficiency in pneumococci (Inniss et al., 2019).

Although much is known about how σ^X functions in streptococcal species the details of how it facilitates combox promoter recognition and melting are not known. Moreover, how and why σ^X-mediated transcription depends on ComW in unknown. Here we show that ComW stimulates transcription of competence genes, *in vitro*.

MATERIALS AND METHODS

Bacterial Strains, Culture Media, and Plasmids

Bacterial strains and plasmids used are listed in **Table 1**. Plasmids pNLI60, pNLI115, and pNLI116 were synthesized by GenScript. *E. coli* strains DH5α and BL21De3 were hosts for plasmid isolation and protein expression, respectively. For plasmid introduction, *E. Coli* strains were chemically transformed according to (Mandel and Higa, 1970). *E. coli* strains were cultured in lysogeny broth (LB) (Bertani, 1951). LB was prepared from 5 g bacto tryptone (Difco), 5 g NaCl (Fisher) 2.5 g of yeast extract (Difco) in 1 L H_2O, sterilized for 20 min at 121°C, and supplemented with appropriate antibiotics and 1.5% agar, as needed. Ampicillin was used at 100 μg/mL, for growth of *E. coli* strains. Antibiotics were purchased from Sigma-Aldrich.

Expression and Purification of Pneumococcal Proteins From *E. coli*

Pneumococcal ComWΔ6 was expressed from pNLI60 and purified according to (Inniss et al., 2019). Pneumococcal σ^X and σ^A were expressed from pNLI94 and pNLI114, respectively. *E. coli* transformed with pNLI94 (SigX-V5H6, pI ∼8.42) or with pNLI114 (σ^A-V5H6, pI ∼5.02) were cultured in 3 L of LB medium at 37°C, 200 rpm to an OD_{600} 0.5. To prepare a sample for SDS-PAGE, 2 mL of uninduced cells were collected and pelleted for 10 min at room temperature, and then boiled at 95°C for 2 min in a 3:1 mixture of *E. coli* resuspension buffer (50 mM Tris-HCl, pH 8.0, 500 mM NaCl) and Laemmli's buffer (Bio Rad), respectively, then stored at 4°C. The large culture was induced for protein expression by addition of $[1\ mM]_f$ isopropyl β-D-1-thiogalactopyranoside (IPTG) (Gold Biotechnology). Induced cultures were incubated overnight at 20°C, 200 rpm. The next day, 2 mL of induced culture was prepped for analysis by SDS PAGE, and the large cultured collected at 5,000 rpm, 4°C, for 30 min. Cells were resuspended in 50 mL of *E. coli* resuspension buffer supplemented with 25 mM imidazole, 50 mM $MgCl_2$, 100 μg/mL DNase, and a protease inhibitor tablet (Roche). The cell resuspension was lysed using an Emulsiflex-C3 (Avestin). The

TABLE 1 | Bacterial strains, plasmids, and primers.

Bacterial strains (strain, genotype)		
E. coli DH5α	F- *φ80lacZ*Δ*M16* Δ(*lacZYA-argF*)U169 *recA1 endA1 hsdR17* (rk-mk +) *phoA supE44*λ- *thi-1 gyrA96 relA1*	
E. coli BL21(DE3)	F- *ompT hsdSB* (rB-mB-) *gal dcm* (DE3)	
Plasmids (name, description)		
pNLI60	pET22b+, Sp *comW*Δ*6*, C terminal fusion to V5His6 tag, under IPTG inducible promoter, Amp^R	
pNLI94	pET22b+, Sp *sigX*, C-terminal fusion to V5His6 tag, under, IPTG inducible promoter, Amp^R	
pNLI114	pET22b+, Sp *rpoD*, C-terminal fusion to V5His6 tag, under IPTG inducible promoter, Amp^R	
pNLI115	pET22b+, partial Sp *comEA*, partial Sp *amiA,*	
pNLI116	pET22b+, partial Sp *ssbB*	
Primers (name, sequence, gene target)		
NL228	CGACGGTTGACAGCGATAGTTGC	*ssbB*
NL229	CAGATATGACCATTATGGCCAATCAACAG	*ssbB*
NL295	CATGAAAAAGGCCGAATCGTGACAAGAGTT	*comEA*
NL296	CCTGCAAATTCGTCTCTTTGACAGGTGTTT	*comEA*
NL299	CATTTTACTGTATGTCTTCCTAAACTCCAAAG	*amiA*
NL300	CATTTAACCCCTTTACGAATCTTATAAGTGTAG	*amiA*

Sp, Streptococcus pneumoniae; V5H6, V5, and Hisx6 epitopes.

soluble fraction was separated by centrifugation at 30,000 × g for 30 min at 4°C.

For σ^X purification, the soluble fraction was discarded. The pellet was washed in 15 mL of cold, filtered 20 mM L-arginine, 1 mM EDTA, 5% glycerol (RGE buffer, final pH, 10.95) supplemented with 1 mM βME and 2% w/v sodium deoxycholate (NaDOC) (Sigma Aldrich) with a homogenizer, followed by sonication for 2 min at 65% amplitude (0.8/0.2, on/off cycles). The washed pellet was spun was 30,000 × g for 30 min at 4°C. Washing, sonication, and spinning were repeated. Then the pellet was resuspended and dissolved in 15 mL of RGE supplemented with 1 mM βME and 0.6% N-lauroyl sarcosine sodium salt (Sigma-Aldrich) at 4°C with slow stirring for up to 2 days. The dissolved inclusion body was collected by centrifugation at 30,000 × g for 30 min at 4°C. The supernatant was saved and dialyzed for 8 h at room temperature in 2 L of 25 mM Tris-HCl, pH 8.3, 500 mM NaCl, 30 mM L-arginine, 10% glycerol, and 5 mM βME (CB4), and then overnight at 4°C in 2 L of CB4. The soluble supernatant was passed through a 5 mL Ni-NTA resin (Thermo Fisher Scientific) on a column equilibrated with CB4 (no βME) via gravity flow. The column was washed with 500 mL of CB4 supplemented with 5 mM βME and 40 mM imidazole. SigX-V5H6 was eluted from the column with 10 mL of CB4 supplemented with 5 mM βME and 275 mM imidazole. Subsequently, the flow through was saved and concentrated using a Vivaspin 15R 2,000 MWCO column concentrator (Sartorius). The concentrated solution was applied to a pre-packed Superdex 200 column (GE) on an AKTA (Amersham Biosciences). SigX-V5H6 was eluted at 0.3 mL/min in CB4 supplemented with 1 mM βME and collected in 1 ml fractions for analysis on SDS-PAGE gels. The fractions were pooled and concentrated again using a VivaSpin column concentrator. Protein concentration was measured at 5.6 mg/mL using a Nanodrop. Purified SigX-V5H6 was diluted to 1.2 mg/mL in CB4 supplemented with 1 mM EDTA, 1 mM βME and 20% glycerol for storage at −20°C until use.

For purification of σ^A, the soluble fraction was separated by centrifugation at 30,000 × g for 30 min at 4°C. The soluble supernatant was passed through a 5 mL Ni-NTA resin (Thermo Fisher Scientific) on a column equilibrated with Column Buffer (50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 10% glycerol) via gravity. The column was washed with 500 mL of Column Buffer supplemented with 5 mM βME and 40 mM imidazole. σ^A was eluted from the column with 10 mL of column buffer supplemented with 5 mM βME and 275 mM imidazole. The eluate was concentrated using a Vivaspin 15R concentrator column into 50 mM Tris-HCl pH 8.0, 20 mM L-arginine, 500 mM NaCl, 1 mM EDTA, 10% glycerol, and 1 mM βME (DBR buffer). Protein concentration was measured as 1.85 mg/mL using a Nanodrop. SigA-V5H6 was diluted to 1 mg/mL in DBR supplemented with 20% glycerol and stored at −20°C until use.

Amplification of Linear Templates for *in vitro* Transcription

Primers used in this study are listed in **Table 1**. Gene sequences for *amiA*, *comEA*, and *ssbB* were taken from the genome of *S. pneumoniae* R6 (NC_003098.1). Oligonucleotides were synthesized by IDT (Coralville, Iowa). Equal amounts (1 μM) of forward and reverse primers were used to amplify *amiA* (NL299/NL300), *comEA* (NL295/NL296), and *ssbB* (NL228/NL229) from 50 ng of CP2137 genomic DNA with 0.2 mM dNTP (Thermo Fisher Scientific) and 1 μL of Phire Hot Start DNA Polymerase II (Thermo Fisher Scientific) in 50 μL reactions in a thermocycler. PCR products were purified with Zymo kits, eluted in diethyl pyrocarbonate (DEPC)

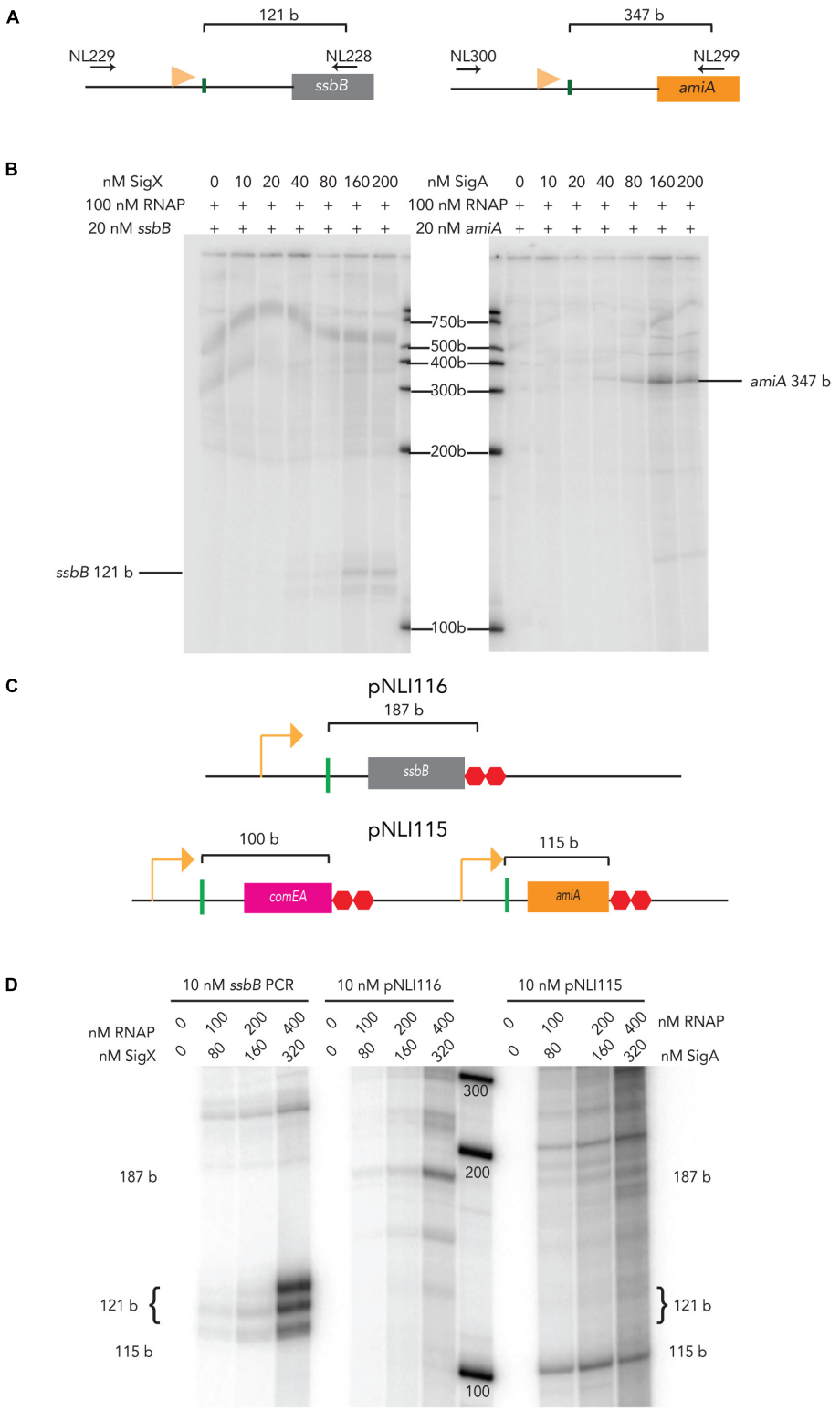


FIGURE 1 | *E. coli* RNAP activation by pneumococcal sigma factors on linear and plasmid templates. **(A)** Schematic of *ssbB* (left) and *amiA* (right) PCR templates amplified from CP2137 genomic DNA using gene specific primers (their positions are indicated by the horizontal black arrows). The horizontal orange flags mark the positions of the combox promoter for *ssbB* and the Pribnow promoter box for *amiA*. The vertical green lines mark the transcription start sites for each gene, and the brackets indicate the size of the expected mRNA product (Aprianto et al., 2018). **(B)** The mRNA products from titration of pneumococcal σ^X (left blot) or σ^A with (right

(Continued)

FIGURE 1 | Continued
blot) with 100 nM of *E. coli* RNA polymerase and 20 nM of *ssbB* (left) or *amiA* (right) PCR templates. Pneumococcal σ-factors were titrated up to 200 nM. **(C)** (Top) Schematic of relevant region of plasmid pNLI116 containing pneumococcal combox promoter upstream of an *ssbB* gene fragment. (Bottom) Schematic of the relevant region of plasmid pNLI115, containing pneumococcal combox promoter upstream of *comEA*, followed the promoter upstream of pneumococcal *amiA*. The promoter upstream of each gene fragment is indicated by an orange arrow, and the vertical green lines mark the transcription start sites, with brackets that indicate the size of the expected mRNA product (Aprianto et al., 2018), and two red octagons indicate one copy of a pneumococcal termination signal (Aprianto et al., 2018) and a *ltR2* terminator. **(D)** mRNA products that result from *in vitro* transcription with pneumococcal σ^X or σ^A and *E. coli* RNA polymerase from 10 nM of PCR or plasmid template with *ssB* or *amiA* gene fragments, respectively. The sizes of the mRNA products produced are indicated on the side of the gel, and were estimate based on Aprianto et al. (2018).

H_2O and visualized on a 1% agarose gel prior to storage at −20°C.

In vitro Transcription From DNA Templates

In vitro transcription from linear templates was carried out using 100 nM of *E. coli* RNAP obtained from NEB, pneumococcal σ^X and σ^A purified from *E. coli*, and ComWΔ6 purified from *E. coli*, when needed. To determine the minimum concentration of each σ-factor required to stimulate RNAP, initiation complexes were formed with 100 nM of *E. coli* RNAP, 20 nM of *ssbB* or *amiA*, and 0–200 nM of σ^X or σ^A, in 50 mM Tris-acetate, 50 mM sodium-acetate, 12.5% glycerol, 0.25 mM EDTA, 15 mM magnesium-acetate, and 35 mM βME (4X TGA buffer). Transcription from plasmid templates was carried out using 0–400 nM of *E. coli* RNAP and 0–320 nM of σ^A or σ^X. Solutions were incubated at 37°C for 15 min in a thermocycler, followed by addition of 1 μL of a master mix containing (120 nM) ATP, CTP, and GTP (Thermo Fisher Scientific) with $^{\alpha\ P32}$UTP at 0.5 μCi/μL (3000 ci/mmol, Perkin Elmer). Runoff transcription was allowed to proceed for 6 min at 37°C in a thermocycler, followed by addition of 20 nM cold UTP (0.5 μL of a 200 nM stock) and a final elongation step at 37°C for 10 min in a final volume of 5.5 μL. Reactions were stopped by addition of one volume of formamide dye (80% formamide, 10 mM EDTA pH 8.0, 1 μg/mL xylene cyonal, and 1 μg/mL bromophenol blue) and incubation at 65°C for 2 min in a thermocycler. The samples were placed on ice until gel loading.

To measure transcription stimulation by ComWΔ6 on PCR templates, 100 nM of *E. coli* RNAP, 80 nM σ^X or σ^A, 20 nM of *comEA*, *ssbB*, or *amiA*, and 0–160 nM of ComWΔ6 were mixed in 50 mM Tris-acetate, 50 mM sodium-acetate, 12.5% glycerol, 0.25 mM EDTA, 15 mM magnesium-acetate, and 35 mM βME (4X TGA buffer). Solutions were incubated at 37°C for 15 min in a thermocycler, followed by addition of 1 μL of a master mix containing (120 nM) ATP, CTP, and GTP (Thermo Fisher Scientific) with $^{\alpha\ P32}$UTP at 0.5 μCi/μL (3000 ci/mmol, Perkin Elmer). Runoff transcription was allowed to proceed for 6 min at 37°C in a thermocycler, followed by addition of 20 nM cold UTP (0.5 μL of a 200 nM stock) and a final elongation step at 37°C for 10 min in a final volume of 5.5 μL. Reactions were stopped by addition of one volume of formamide dye (80% formamide, 10 mM EDTA pH 8.0, 1 μg/mL xylene cyonal, and 1 μg/mL bromophenol blue) and incubation at 65°C for 2 min in a thermocycler. The samples were placed on ice until gel loading.

To visualize mRNA transcripts, 2 μL of each reaction was run on a 20 × 20, 0.375 mm thick, 6% polyacrylamide urea gel at 20 watts for 45 min. Gels were dried for 40 min at 80°C and incubated on a phosphorimager and imaged on a Typhoon 3000 (GE). The expected mRNA transcript sizes were determined using the transcriptional start sites (TSS) for each gene as determined by Aprianto et al. (2018). Experimental transcripts were compared to markers generated by *in vitro* transcription using the MAXIscript T7 Transcription Kit (Invirtogen) and the RNA Century-Plus Marker Templates (Invitrogen).

RESULTS

Pneumococcal σ-Factors Stimulate *Escherichia coli* Core RNA Polymerase, *in vitro*

In vitro transcription assays confirmed that σ^X, in molar excess to RNAP, promotes transcription from combox promoters (Luo et al., 2003). However σ^X-mediated late gene transcription and transformation requires ComW *in vivo* (Sung and Morrison, 2005). Previous work concluded that pneumococcal RNAP exists at ∼2000 molecules per cell, and that σ^X and ComW peak at ∼3000 and ∼500 molecules per cell, respectively (Luo and Morrison, 2003; Piotrowski, 2010). Although not confirmed, σ^X might have a lower affinity for RNAP than σ^A, a trait shared among multiple alternative σ-factors (Gruber and Gross, 2003), so excess σ^X could increase the amount of σ^X holo-enzyme formation. Interestingly, levels of σ^A decrease 20–40% during the pneumococcal competence response (Luo and Morrison, 2003; Luo et al., 2003), and mutations to σ^A, that presumably decrease its affinity for RNAP, can suppress Δ*comW* phenotypes (Tovpeko and Morrison, 2014; Tovpeko et al., 2016). These data hint that σ^X access to RNAP and/or competence promoters is an important step in efficient competence gene expression. In addition, we have shown that ComW, a likely σ^X binding partner (Tovpeko et al., 2016), interacts with DNA and this interaction is important for transformation (Inniss et al., 2019). We hypothesize that ComW functions at or near competence promoters to boost σ^X transcriptional activity. We tested this theory with *in vitro* transcription (IVT) assays.

Luo and Morrison used a ∼1.7 molar excess of σ^X to pneumococcal RNAP ($RNAP_{Sp}$) in their IVT assays, making the number of molecules of σ^X higher than that of $RNAP_{Sp}$ (roughly 5.12×10^{12} vs. 3.13×10^{12} molecules, respectively, and based on the average molar ratio of σ^X:$RNAP_{Sp}$ in cells as determined by Luo and Morrison, 2003). Thus, IVT from every competence template tested was successful

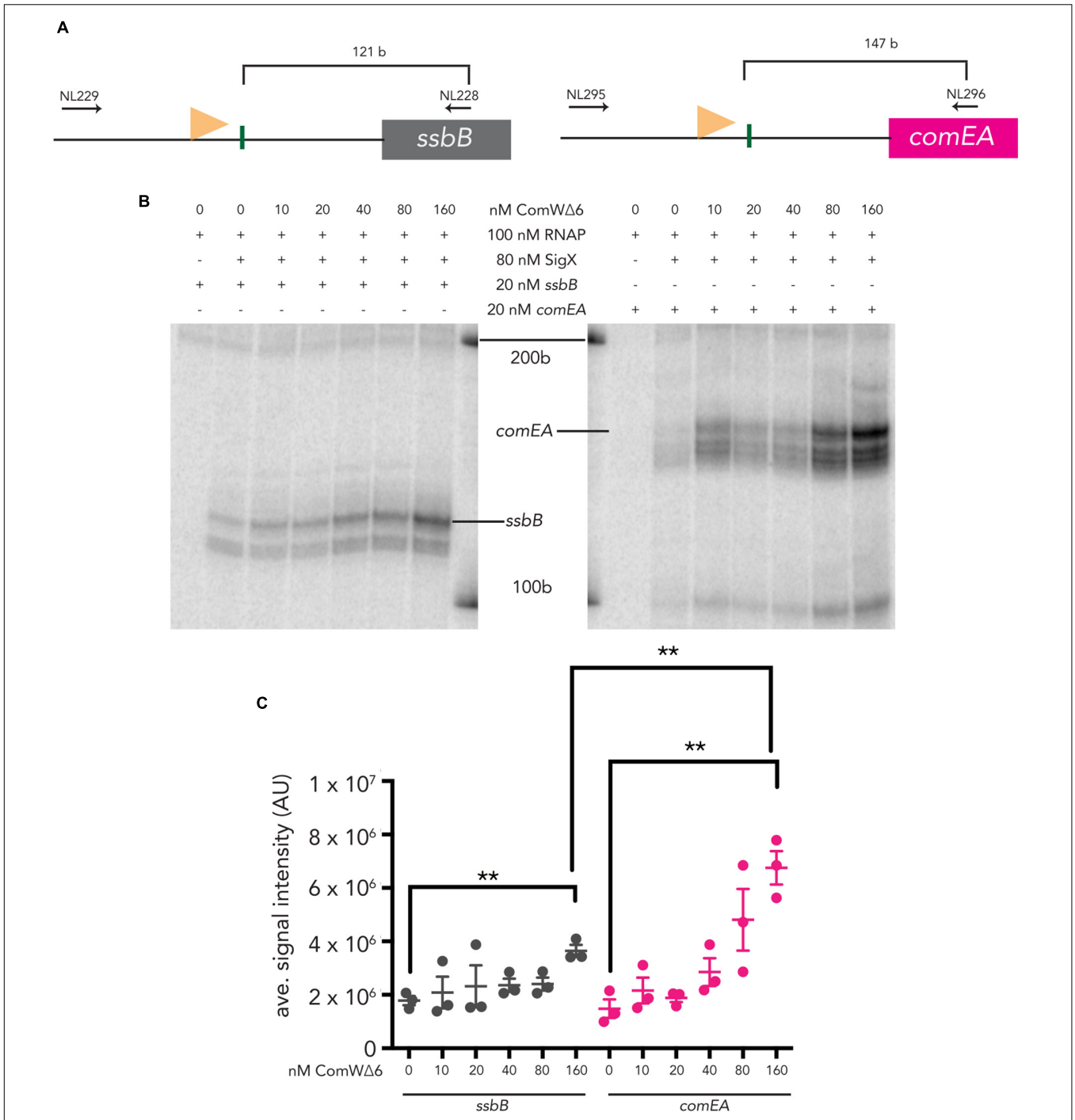


FIGURE 2 | *In vitro* transcription from two pneumococcal late competence promoters using ComWΔ6. **(A)** A schematic of pneumococcal *ssbB* (left) and *comEA* (right) templates. A light orange flag represents each promoter and vertical green lines mark the transcription start sites. Horizontal arrows mark the positions of the primers used to amplify each template. **(B)** An image of *ssbB* (121 b, left) and *comEA* (147 b, right) mRNA products produced by σ^X holoenzymes with increasing amounts of ComWΔ6. The positions of the 100 b and 200 b standard bands are indicated. **(C)** Quantification of the signal intensities of mRNA transcripts from three *in vitro* transcription experiments. Asterisks mark statistically significant differences between samples with 0 nM ComWΔ6 and 160 nM ComWΔ6, and between samples with *ssbB* and *comEA* templates.

and did not require ComW. As RNAP is conserved in prokaryotes (Murakami and Darst, 2003), we titrated pneumococcal σ^X and σ^A to determine the minimal amount of each σ-factor required to target 100 nM of *E. coli* RNAP ($RNAP_{Ecoli}$) to the pneumococcal combox and housekeeping promoters, *in vitro*.

The promoter regions of pneumococcal *ssbB* and *amiA* were PCR amplified for use in *in vitro* transcription assays (**Figure 1A**). At 40 nM σ^X and 100 nM $RNAP_{Ecoli}$, a 121-base (b) mRNA product from 20 nM of an *ssbB* linear template was produced, although the amount of mRNA produced (**Figure 1B**). The amount of *ssbB* transcript produced increased as σ^X was titrated up to 160 nM. At 80 nM, there are approximately 1.93×10^{11} molecules of σ^X versus 2.41×10^{11} molecules of $RNAP_{Ecoli}$ at 100 nM in the reaction. This molar ratio (80 nM σ^X: 100 nM $RNAP_{Ecoli}$) allowed for observable, yet sub-optimal transcription from 20 nM of linear *ssbB* DNA template containing a combox promoter, without ComW. In addition, 80 nM of σ^X + 100 nM $RNAP_{Ecoli}$ did not produce an mRNA product when incubated with 20 nM of linear *amiA* template, a gene under control of the primary σ-factor, σ^A (not shown), demonstrating that σ^X transcription specifically initiates from the combox promoter.

In contrast at 20 nM, pneumococcal σ^A produced detectable levels of a 347-base mRNA product from the housekeeping promoter upstream of *amiA* when incubated with $RNAP_{Ecoli}$ (**Figure 1B**). The amount of *amiA* transcript produced increased as σ^A was titrated up to 160 nM, with a detectable *amiA* transcript at 80 nM. Interestingly, we did observe production of an *ssbB* transcript with 80 nM of σ^A (not shown). This might result from an orphan -10 promoter region σ^A binding site that is present 29 bp downstream of the *ssbB* transcription start site, according to Aprianto et al. (2018). The nature of this mRNA product must be examined more thoroughly. Nonetheless, these results demonstrate that both pneumococcal σ-factors direct $RNAP_{Ecoli}$ to pneumococcal promoters, *in vitro*. In addition, as it appeared that more σ^X was required to stimulate $RNAP_{Ecoli}$ activity at combox promoters than σ^A at housekeeping promoters, these results might indicate that more σ^X is needed to stimulate $RNAP_{Sp}$. As 80 nM of σ^X and σ^A are below saturating concentrations but still produced visible *ssbB* and *amiA* transcripts, respectively, we used 80 nM of each σ-factor in subsequent reactions.

To further explore *in vitro* transcription with pneumococcal σ factor, we tested transcription from plasmid templates. Supercoiling can affect bacterial gene transcription (Pruss and Drlica, 1989; Zhi et al., 2017), thus we asked whether or not our chimeric holo-enzymes could transcribe the *ssbB* and *amiA* genes from these plasmid templates. We designed two plasmids that carry pneumococcal promoters upstream of pneumococcal gene fragments, pNLI115 and pNLI116 (**Figure 1C**). pNLI116 contained only the combox promoter followed by an *ssbB* gene fragment, similar to our original PCR *ssbB* IVT template. pNLI115 contained the combox promoter region upstream of competence gene fragment *comEA*, and the housekeeping promoter region upstream of *amiA*, to serve as an internal control for promoter specific IVT from plasmid assays. After isolation from *E. coli,* each plasmid preparation contained predominantly supercoiled species, as observed on DNA agarose gels.

We titrated holo-enzymes into IVT reactions with 10 nM of either PCR DNA templates or plasmid templates. As expected, the amount of the 121 b mRNA transcript produced from linear DNA increased as the amount of σ^X holo-enzyme increased. Similar results were observed from pNLI116, as increasing amounts of a 187 b mRNA transcript was produced as the amount of σ^X holo-enzyme was titrated onto the plasmid (**Figure 1D**, left). In addition, a 115 b mRNA transcript was produced from pNLI115 when σ^A-$RNAP_{E.\ coli}$ complex was titrated onto the plasmid (**Figure 1D**, right). Together these data show that our chimeric holo-enzymes can be targeted to specific promoters on either linear or plasmid DNA templates, *in vitro*.

ComWΔ6 Stimulates Transcription From σ^X Promoters, *in vitro*

Both *ssbB* and *comEA* are induced by CSP and require ComW for robust transcription (Sung and Morrison, 2005). To determine the effect of ComW on transcription from combox promoters upstream of these two genes, we titrated ComWΔ6 into IVT reactions with σ^X holo-enzyme. Transcripts from both linear *ssbB* and *comEA* templates (**Figure 2**) were produced when ComWΔ6 was added to IVT reactions. Specifically, the σ^X holo-enzyme incubated with ComWΔ6 at 160 nM (2X molar excess of ComW to σ^X) produced 2X as much mRNA signal from the *sbbB* template compared to reactions with 0 nM ComWΔ6, an increase that was statistically significant (**Figure 2**). Similarly, the σ^X holo-enzyme incubated with ComWΔ6 at 160 nM produced 4.5X as much mRNA signal from the *comEA* template compared

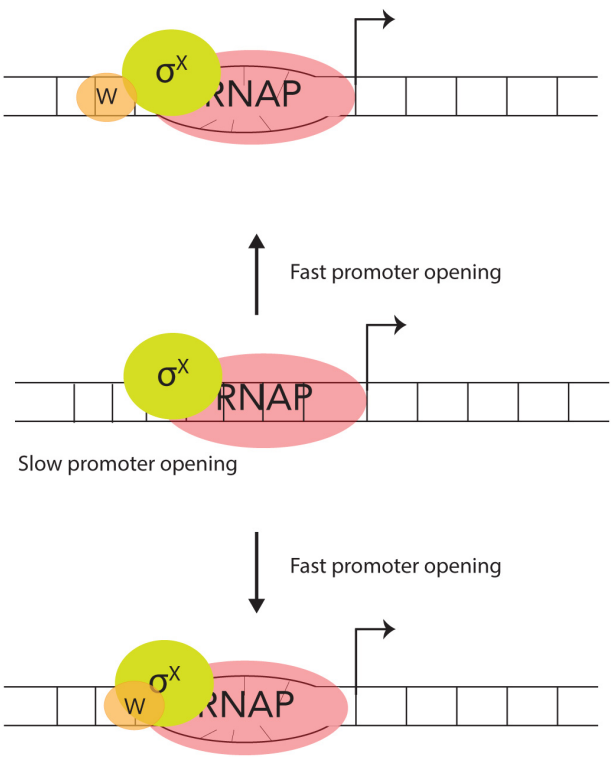


FIGURE 3 | Two proposed mechanisms for ComW-dependent promoter melting. In the absence of ComW the pneumococcal σ^X-holoenzyme can recognize and bind to combox containing DNA, but is slow to melt the promoter (middle). ComW is brought to combox promoters via interaction with σ^X (top) and/or RNAP (bottom). ComW uses its non-specific DNA binding function to aid in melting promoter DNA.

to σ^X holo-enzyme with no ComWΔ6, also an increase that was statistically significantly (**Figure 2**). The 4.5X increase in *comEA* mRNA signal production versus the 2X increase in *ssbB* signal production also represented a statistically significant difference. These data demonstrate that ComW is required for robust transcription when the number of σ^X molecules is below the number of core RNAP molecules, *in vitro*. Furthermore, the difference in transcription stimulation suggests that the requirement for ComW at combox promoters differs. Lastly, as ComWΔ6 can stimulate σ^X-mediate transcription, *in vitro*, it is possible that residues ^{73}RGFISC^{78} are dispensable for transcription activation in pneumococci.

DISCUSSION

Anginosus and mitis group streptococci like the pneumococcus use the DNA binding protein, ComW, to regulate σ^X-mediated transcription. However, previous data show that σ^X transcribes late genes in the absence of ComW, *in vitro* (Luo and Morrison, 2003). As we hypothesized that ComW is an active and important member of the competence-specific holo-enzyme, we examined how ComWΔ6 affected IVT from σ^X promoters. Remarkably, two σ^X targets, *ssbB* and *comEA* appeared to respond differently to the addition of ComWΔ6 to reactions; although transcription stimulation was observed for both genes, the level of stimulation appears greater for *comEA*. This is common in other naturally transformable species. For example, the *V. cholerae* competence regulator, QstR, binds upstream of some competence genes, but not others (Jaskólska et al., 2018), and the number of AT boxes that lie upstream of competence promoters in *B. subtilis* can differ from gene to gene, likely reflecting the difference in ComK activity at these promoters (Hamoen et al., 2003). Thus, at this moment, we cannot rule out that ComW is specific for transcription of some pneumococcal late competence genes, and might be dispensable for the transcription of others.

Interestingly, it appears that σ^A better stimulates *E. coli* RNAP than σ^X. This might suggest that pneumococcal RNAP responds similarly to these two σ factors, and the specifics of pneumococcal σ – RNAP interactions must be determined. This could be another reason as to why σ^X – mediated transcription requires ComW in the pneumococcus. Furthermore, the σ^X – ComW pair appear similar to a two-part σ factor, σ^O – RsoA, identified in *B. subtilis* (Maclellan et al., 2009). In this system RsoA aids in σ^O – RNAP open complex formation to promote transcription of genes required during growth in acidic conditions. Like σ^O, σ^X likely provides promoter recognition, and ComW, like RsoA, might aid in stabilizing the σX complex at the promoter (Maclellan et al., 2009). However, the specifics of a probable σX – ComW or ComW – RNAP interaction are yet to be determined.

The DNA – binding protein, ComW does stimulate σ^X – mediated transcription, *in vitro*. Although more effort is required to fully work out the details, we posit that there are likely two possible roles for how ComW functions (**Figure 3**). (A) As ComW interacts non-specifically with DNA, it is likely brought to competence promoters by σ^X, and increases promoter melting. (B) ComW might function as a component of a two-part σ factor with σ^X and RNAP as an active member of the holo-enzyme.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

AUTHOR CONTRIBUTIONS

NI conceived, conducted experiments and wrote the manuscript. DM edited the manuscript.

FUNDING

Funding was provided by The University of Illinois at Chicago through the Abraham Lincoln Fellowship, as well as an R03 grant from the National Institute of Allergy and Infectious Diseases of the National Institutes of Health (1R03AI128228).

REFERENCES

Andam, C. P., and Hanage, W. P. (2015). Mechanisms of genome evoluation of *Streptococcus*. *Infect. Genet. Evol.* 33, 334–342. doi: 10.1016/j.meegid.2014.11.007

Aprianto, R., Slager, J., Holsappel, S., and Veening, J.-W. (2018). High-resolution analysis of the pneumococcal transcriptome under a wide range of infection-relevant conditions. *Nucleic Acids Res.* 17:1133. doi: 10.1093/nar/gky750

Avery, O. T., and MacLeod, C. M. (1944). Studies on the chemical nature of the substance inducing transformation of pneumococcal types. *J. Exp. Med.* 79, 137–158. doi: 10.1084/jem.149.2.297

Bertani, G. (1951). Studies on lysogenesis. *J. Bacteriol.* 62, 293–300. doi: 10.1128/jb.62.3.293-300.1951

Cheng, Q., Campbell, E. A., Naughton, A. M., Johnson, S., and Masure, H. R. (1997). The com locus controls genetic transformation in *Streptococcus pneumoniae*. *Mol. Microbiol.* 23, 683–692. doi: 10.1046/j.1365-2958.1997.2481617.x

Fontaine, L., Boutry, C., de Frahan, M. H., Delplace, B., Fremaux, C., Horvath, P., et al. (2010). A novel pheromone quorum-sensing system controls the development of natural competence in *Streptococcus thermophilus* and *Streptococcus salivarius*. *J. Bacteriol.* 192, 1444–1454. doi: 10.1128/JB.01251-09

Gao, X.-Y., Zhi, X.-Y., Li, H.-W., Klenk, H.-P., and Li, W.-J. (2014). Comparative genomics of the bacterial genus *Streptococcus* illuminates evolutionary implications of species groups. *PLoS One* 9:e101229. doi: 10.1371/journal.pone.0101229.s005

Griffith, F. (1928). The significance of pneumococcal types. *J. Hygiene* 27, 1–47.

Gruber, T. M., and Gross, C. A. (2003). Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu. Rev. Microbiol.* 57, 441–466. doi: 10.1146/annurev.micro.57.030502.090913

Hamoen, L. W., Venema, G., and Kuipers, O. P. (2003). Controlling competence in *Bacillus subtilis*: shared use of regulators. *Microbiology* 149, 9–17. doi: 10.1099/mic.0.26003-0

Håvarstein, L. S. (2010). Increasing competence in the genus *Streptococcus*. *Mol. Microbiol.* 78, 541–544. doi: 10.1111/j.1365-2958.2010.07380.x

Håvarstein, L. S., Coomaraswamy, G., and Morrison, D. A. (1995). An unmodified heptadecapeptide pheromone induced competence for genetic transformation in *Streptococcus pneumoniae*. *Proc. Natl. Acad. Sci. U.S.A.* 92, 11140–11144. doi: 10.1073/pnas.92.24.11140

Håvarstein, L. S., Gaustad, P., Nes, I. F., and Morrison, D. A. (1996). Identification of the streptococcal competence- pheromone receptor. *Mol. Microbiol.* 21, 863–869. doi: 10.1046/j.1365-2958.1996.521416.x

Inniss, N. L., Prehna, G., and Morrison, D. A. (2019). The pneumococcal σX activator, ComW, is a DNA-binding protein critical for natural transformation. *J. Biol. Chem.* 294, 11101–11118. doi: 10.1074/jbc.RA119.007571

Jaskólska, M., Stutzmann, S., Stoudmann, C., and Blokesch, M. (2018). QstR-dependent regulation of natural competence and Type VI secretion in *Vibrio cholerae*. *Nucleic Acid Res.* 46, 10619–10634. doi: 10.1093/nar/gky717

Khan, R., Junges, R., Mdal, H. A. A. X., Chen, T., Morrison, D. A., and Petersen, F. C. (2017). A positive feedback loop mediated by Sigma X enhances expression of the streptococcal regulator ComR. *Sci. Rep.* 7:5984. doi: 10.1038/s41598-017-04768-5

Khan, R., Rukke, H. V., Høvik, H., Åmdal, H. A., Chen, T., Morrison, D. A., et al. (2016). Comprehensive transcriptome profiles of *Streptococcus mutans* UA159 map core streptococcal compentence genes. *mSystems* 1:e00038-15. doi: 10.1128/mSystems.00038-15

Lee, M. S., and Morrison, D. A. (1999). Identification of a new regulator in *Streptococcus pneumoniae* linking qurum sensing to competence for genetic transformation. *J. Bacteriol.* 181, 5004–5016. doi: 10.1128/jb.181.16.5004-5016.1999

Lee, M. S., and Morrison, D. A. (2000). Regulation of competence for genetic transformation in *Streptococcus pneumoniae*: a link between quorum sensing and DNA processing genes. *Res. Microbiol.* 151, 445–451. doi: 10.1016/s0923-2508(00)00171-6

Luo, P., and Morrison, D. A. (2003). Transient association of an alternative sigma factor, ComX, with RNA polymerase during the period of competence for genetic transformation in *Streptococcus pneumoniae*. *J. Bacteriol.* 185, 349–358. doi: 10.1128/JB.185.1.349-358.2003

Luo, P., Li, H., and Morrison, D. A. (2003). ComX is a unique link between multiple quorum sensing outputs and competence in *Streptococcus pneumoniae*. *Mol. Microbiol.* 50, 623–633. doi: 10.1046/j.1365-2958.1999.01528.x

Luo, P., Li, H., and Morrison, D. A. (2004). Identification of ComW as a new component in the regulation of genetic transformation in *Streptococcus pneumoniae*. *Mol. Microbiol.* 54, 172–183. doi: 10.1111/j.1365-2958.2004.04254.x

Maclellan, S. R., Guariglia-Oropeza, V., Gaballa, A., and Helmann, J. D. (2009). A two-subunit bacterial σ-factor activates transcription in *Bacillus subtilis*. *Proc. Natl. Acad. Sci.* 106, 21232–21328.

Mandel, M., and Higa, A. (1970). Calcium-dependent bacteriophage DNA infection. *J. Mol. Biol.* 53, 159–162. doi: 10.1016/0022-2836(70)90051-3

Martin, B., Granadel, C., Campo, N., Hénard, V., Prudhomme, M., and Claverys, J.-P. (2010). Expression and maintenance of ComD-ComE, the two-component signal-transduction system that controls competence of *Streptococcus pneumoniae*. *Mol. Microbiol.* 75, 1513–1528. doi: 10.1111/j.1365-2958.2010.07071.x

Martin, B., Soulet, A.-L., Mirouze, N., Prudhomme, M., Mortier-Barrière, I., Granadel, C., et al. (2012). ComE/ComE P interplay dictates activation or extinction status of pneumococcal X-state (competence). *Mol. Microbiol.* 87, 394–411. doi: 10.1046/j.1365-2958.1999.01528.x

Mashburn-Warren, L., Morrison, D. A., and Federle, M. J. (2010). A novel double-tryptophan peptide pheromone controls competence in *Streptococcus* spp. via an Rgg regulator. *Mol. Microbiol.* 78, 589–606. doi: 10.1111/j.1365-2958.2010.07361.x

Mashburn-Warren, L., Morrison, D. A., and Federle, M. J. (2012). The cryptic competence pathway in *Streptococcus pyogenes* is controlled by a peptide pheromone. *J. Bacteriol.* 194, 4589–4600. doi: 10.1128/JB.00830-12

Murakami, K. S., and Darst, S. A. (2003). Bacterial RNA polymerases: the wholo story. *Curr. Opin. Struct. Biol.* 13, 31–39. doi: 10.1016/S0959-440X(02)00005-2

Peterson, S., Cline, R. T., Tettelin, H., Sharov, V., and Morrison, D. A. (2000). Gene expression analysis of the *Streptococcus pneumoniae* competence regulons by use of DNA microarrays. *J. Bacteriol.* 182, 6192–6202. doi: 10.1128/jb.182.21.6192-6202.2000

Peterson, S. N., Sung, C. K., Cline, R., Desai, B. V., Snesrud, E. C., Luo, P., et al. (2004). Identification of competence pheromone responsive genes in *Streptococcus pneumoniae* by use of DNA microarrays. *Mol. Microbiol.* 51, 1051–1070. doi: 10.1046/j.1365-2958.2003.03907.x

Piotrowski, A. (2010). *Regulation of the Activities of ComX and ComW During Competence Development in Streptococcus Pneumoniae*. Ph.D. thesis, Univerisity of Illinois at Chicago, Chicago. 1–169.

Piotrowski, A., Luo, P., and Morrison, D. A. (2009). Competence for genetic transformation in *Streptococcus pneumoniae*: termination of activity of the alternative sigma factor ComX is independent of proteolysis of ComX and ComW. *J. Bacteriol.* 191, 3359–3366. doi: 10.1128/JB.01750-08

Pruss, G. J., and Drlica, K. (1989). DNA supercoiling and prokaryotic transcription. *Cell* 56, 521–523. doi: 10.1016/0092-8674(89)90574-6

Sung, C. K., and Morrison, D. A. (2005). Two distinct functions of ComW in stabilization and activation of the alternative sigma factor ComX in *Streptococcus pneumoniae*. *J. Bacteriol.* 187, 3052–3061. doi: 10.1128/JB.187.9.3052-3061.2005

Tovpeko, Y., Bai, J., and Morrison, D. A. (2016). Competence for genetic transformation in *Streptococcus pneumoniae*: mutations in σA bypass the ComW requirement for late gene expression. *J. Bacteriol.* 198, 2370–2378. doi: 10.1128/JB.00354-16

Tovpeko, Y., and Morrison, D. A. (2014). Competence for genetic transformation in *Streptococcus pneumoniae*: mutations in a bypass the comW requirement. *J. Bacteriol.* 196, 3724–3734. doi: 10.1128/JB.01933-14

Ween, O., Gaustad, P., and Håvarstein, L. S. (1999). Identification of DNA binding sites for ComE, a key regulator of natural competence in *Streptococcus pneumoniae*. *Mol. Microbiol.* 33, 817–827. doi: 10.1046/j.1365-2958.1999.01528.x

Zhi, X., Dages, S., Dages, K., Yingting, L., Hua, Z., Makemson, J., et al. (2017). Transient and dynamic DNA supercoiling potently stimulates the leu-500 promoter in *Escherichia coli*. *J. Biol. Chem.* 292, 14566–14575. doi: 10.1074/jbc.M117.794628

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Inniss and Morrison. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 16 June 2020
doi: 10.3389/fmolb.2020.00108

Genome Segregation by the Venus Flytrap Mechanism: Probing the Interaction Between the ParF ATPase and the ParG Centromere Binding Protein

Marisa Caccamo[1†], *Aneta Dobruk-Serkowska*[1†], *Fernando Rodríguez-Castañeda*[2†], *Cecilia Pennica*[2], *Daniela Barillà*[2] *and Finbarr Hayes*[1*]

[1] Faculty of Biology, Medicine and Health, The University of Manchester, Manchester, United Kingdom, [2] Department of Biology, University of York, York, United Kingdom

The molecular events that underpin genome segregation during bacterial cytokinesis have not been fully described. The tripartite segrosome complex that is encoded by the multiresistance plasmid TP228 in *Escherichia coli* is a tractable model to decipher the steps that mediate accurate genome partitioning in bacteria. In this case, a "Venus flytrap" mechanism mediates plasmid segregation. The ParG sequence-specific DNA binding protein coats the *parH* centromere. ParF, a ParA-type ATPase protein, assembles in a three-dimensional meshwork that penetrates the nucleoid volume where it recognizes and transports ParG-*parH* complexes and attached plasmids to the nucleoid poles. Plasmids are deposited at the nucleoid poles following the partial dissolution of the ParF network through a combination of localized ATP hydrolysis within the meshwork and ParG-mediated oligomer disassembly. The current study demonstrates that the conformation of the nucleotide binding pocket in ParF is tuned exquisitely: a single amino acid change that perturbs the molecular arrangement of the bound nucleotide moderates ATP hydrolysis. Moreover, this alteration also affects critical interactions of ParF with the partner protein ParG. As a result, plasmid segregation is inhibited. The data reinforce that the dynamics of nucleotide binding and hydrolysis by ParA-type proteins are key to accurate genome segregation in bacteria.

Keywords: multidrug resistance, plasmid partition, ParF, ParG, ParA, *Escherichia coli*, segregation

OPEN ACCESS

Edited by:
Tatiana Venkova,
Fox Chase Cancer Center,
United States

Reviewed by:
Barbara Funnell,
University of Toronto, Canada
Preeti Srivastava,
Indian Institute of Technology
Delhi, India

***Correspondence:**
Finbarr Hayes
finbarr.hayes@manchester.ac.uk

†Present address:
Marisa Caccamo,
Syneos Health Clinical Solutions,
Saronno, Italy
Aneta Dobruk-Serkowska,
Government Center for Security,
Warsaw, Poland
Fernando Rodríguez-Castañeda,
Department of Biochemistry,
Universidad del Valle de Guatemala,
Guatemala, Guatemala

Specialty section:
This article was submitted to
Molecular Recognition,
a section of the journal
Frontiers in Molecular Biosciences

Received: *20 March 2020*
Accepted: *08 May 2020*
Published: *16 June 2020*

Citation:
Caccamo M, Dobruk-Serkowska A, Rodríguez-Castañeda F, Pennica C, Barillà D and Hayes F (2020) Genome Segregation by the Venus Flytrap Mechanism: Probing the Interaction Between the ParF ATPase and the ParG Centromere Binding Protein. Front. Mol. Biosci. 7:108. doi: 10.3389/fmolb.2020.00108

INTRODUCTION

Accurate genome segregation is an essential cellular process that guarantees the stable transmission of genetic material during cytokinesis (Bloom and Joglekar, 2010; Hirano, 2015). Despite the importance of bacteria in maintenance of the biosphere, in nutrient recycling, and in human and animal health, and disease, knowledge of the mechanisms that underpin genome maintenance in bacteria is very fragmented (Badrinarayanan et al., 2015). Nevertheless, low copy number plasmids that encode dedicated partition complexes have proven to be highly tractable and informative elements to probe the events leading to accurate bacterial genome segregation (Hayes and Barillà, 2006, 2010; Gerdes et al., 2010; Sengupta and Austin, 2011; Baxter and Funnell, 2014; Oliva, 2016).

The plasmid segregation complex (segrosome) typically comprises three components: a centromere-like site that is composed of repeat sequences whose numbers, lengths, and sequences differ between plasmids; a centromere binding protein that is specific for the cognate centromere; and, a motor protein that interacts with the DNA binding factor and which drives plasmid movement. The most widely-distributed motor proteins that mediate plasmid segregation are those of the ParA superfamily of Walker-type ATP-binding proteins that also are encoded widely by bacterial chromosomes. Despite their prevalence, the mechanism by which ParA proteins facilitate DNA segregation remains unclear. Numerous ParA homologs from diverse sources form higher-order structures upon ATP binding. Oligomerization is modulated by the cognate centromere binding protein and/or by DNA (Barillà et al., 2005, 2007; Leonard et al., 2005; Lim et al., 2005; Ebersbach et al., 2006; Bouet et al., 2007; Hatano et al., 2007; Machón et al., 2007; Pratto et al., 2008; Batt et al., 2009; Ringgaard et al., 2009; Ditkowski et al., 2010, 2013; Hui et al., 2010; Banigan et al., 2011; Dobruk-Serkowska et al., 2012; Kalliomaa-Sanford et al., 2012). A cycle of ParA assembly and disassembly within the nucleoid may promote the directed movement of plasmid DNA prior to cell division (Barillà et al., 2005; Ringgaard et al., 2009; McLeod et al., 2017). Another model suggests that the nucleoid provides a substructure for plasmid segregation by a diffusion-ratchet mechanism in which a gradient of ParA protein across the surface of the nucleoid promotes plasmid movement (Hatano and Niki, 2010; Vecchiarelli et al., 2010, 2014; Hwang et al., 2013). A variation of this model proposes that plasmids are segregated by recruitment to high-density regions within the nucleoid via interactions mediated by the ParA protein (Le Gall et al., 2016). Alternatively, a DNA-relay mechanism may utilize the elastic dynamics of the chromosome to convey the segregation complex between DNA regions via a ParA gradient (Lim et al., 2014).

The multidrug resistance plasmid TP228 replicates stably at low copy number in *Escherichia coli*. The TP228 segrosome comprises the *parH* centromere site; the ParG centromere binding protein; and the ParA homolog, ParF (Hayes, 2000; Barillà and Hayes, 2003). The *parH* site consists of an array of degenerate tetramer boxes interspersed by AT-rich spacers (Wu et al., 2011). The site is coated by ParG dimers each of which includes a C-terminal ribbon-helix-helix domain that is formed by the intertwining of a pair of monomers into a 2-fold symmetrical structure (Golovanov et al., 2003; Saeed et al., 2015). The ribbon-helix-helix domain mediates sequence-specific DNA binding (Carmelo et al., 2005; Zampini et al., 2009; Wu et al., 2011; Zampini and Hayes, 2012). This domain is linked to flexible N-terminal extensions that modulate DNA binding by ParG (Wu et al., 2011). The flexible tails in ParG also promote the formation of higher-order structures by ParF, and in addition harbor arginine finger-like motifs that enhance ATP hydrolysis by ParF (Carmelo et al., 2005; Barillà et al., 2007).

The structures of ParF in complex with ADP and the non-hydrolyzable ATP analog AMPPCP in both cases comprise a single domain that consists of a central seven-stranded twisted β-sheet surrounded on each side by four α-helices (Schumacher et al., 2012). The protein is monomeric within the ParF-ADP complex which correlates with observations that ADP inhibits the production of higher-order structures by ParF (Barillà et al., 2005; Dobruk-Serkowska et al., 2012). In contrast with the ADP bound form of the protein, the ParF-AMPPCP structure is dimeric with the nucleotide molecules sandwiched between monomer subunits (**Figure 1**). Significantly, these subunits pack into dimer-of-dimer assemblies that in turn are organized into higher-order structures (Schumacher et al., 2012). Importantly, mutations at the interfaces of the ParF dimer-of-dimer concomitantly disrupted oligomerisation and DNA segregation which affirmed that plasmid partitioning by ParF necessitates the formation of higher-order structures by the protein (Schumacher et al., 2012).

Cycles of ATP binding and hydrolysis promote the oscillation of ParF between the poles of the nucleoid. The tethered ParG-*parH* complex is co-transported with ParF as it relocates. The tempo of this oscillatory pattern is key to effective segregation (McLeod et al., 2017). Moreover, in contrast with the diffusion-ratchet segregation model in which a gradient of ParA protein over the surface of the nucleoid has been suggested to promote plasmid transport, the dynamic ParF matrix instead invades the interior of the nucleoid. This three-dimensional meshwork acts as a "Venus flytrap" that captures and transports ParG-*parH* complexes, and therefore the attached plasmids, to the nucleoid poles. The plasmids are deposited at the poles as a result of localized dissolution of the ParF meshwork via ParG-promoted disassembly (McLeod et al., 2017). Here, we probe the role in plasmid segregation of a critical secondary structure element in the ParF protein. The data reveal that a single amino acid change that perturbs the molecular arrangement of the bound nucleotide moderates ATP hydrolysis by ParF, as well as the interaction of ParG *in trans* with the nucleotide binding pocket.

MATERIALS AND METHODS

Bacterial Strains and Growth Conditions

E. coli DH5α (F^- *endA1 hsdR17* (r_K^- m_K^+) *supE44 thi-1 recA1 gyrA96* (Nal^r) *relA1 deoR* Φ80*lacZ*ΔM15 Δ(*lacZYA-argF*) U169) (Hanahan, 1985) was used for plasmid propagation and gene cloning. High level gene expression and protein purification was performed with *E. coli* BL21 (DE3) (F^- *omp hsdSB* ($r_B^-m_B^-$) *gal dcm*) (Studier and Moffatt, 1986). Strain BR825 is a *polA* mutant (Ludtke et al., 1989) that was used in plasmid segregation assays. *E. coli* SP850 (λ^- *e14*$^-$ *spoT1* Δ(*cya1400::kan*) *thi-1*) carries a 200 bp deletion within the *cya* gene (Shah and Peterkofsky, 1991) and was used in two-hybrid assays. Strains were propagated in Luria-Bertani (LB) broth or agar medium at 37°C unless otherwise stated. Appropriate antibiotics for plasmid selection were included when required (ampicillin, 100 μg/ml; chloramphenicol, 10 μg/ml).

Plasmids

Plasmid pFH547 consists of the *parFGH* cassette of multiresistance plasmid TP228 cloned in the plasmid partition vector pFH450. The latter is a derivative of pALA136 into which a multiple cloning site was inserted (Martin et al., 1987;

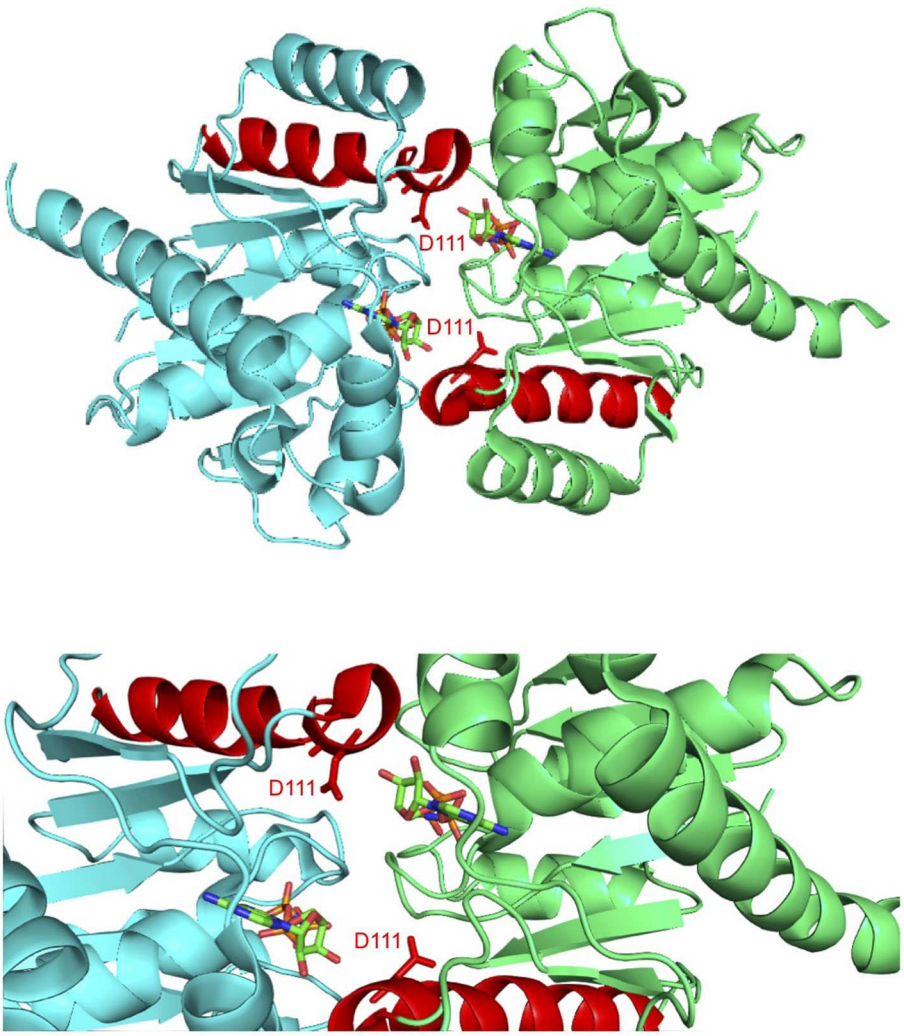


FIGURE 1 | Co-crystal structure of the ParF dimer with AMPPCP (Schumacher et al., 2012). α-5 is highlighted in red with residue Asp-111, which is located at the C-terminal end of this helix, and AMPPCP shown as sticks. The *top* image shows the overall ParF dimer structure and the *bottom* image shows a zoom of the region that contains Asp-111. The images were made using PyMol (Delano, 2001).

Hayes, 2000). Site-specific mutations in *parF* were introduced into pFH547 by overlap extension PCR (Higuchi et al., 1988). Mutated genes were amplified from pFH547-based plasmids and cloned in pT18 and pT25 two-hybrid vectors (Karimova et al., 1998) and in the pET-22b(+) expression vector (Novagen) as outlined previously (Barillà and Hayes, 2003). All mutations were verified by sequencing. Plasmids comprising the *parG* gene cloned in two-hybrid vectors and in expression vectors were described elsewhere (Barillà and Hayes, 2003).

Plasmid Partition Assays

Plasmid partition vector pFH450 includes the moderate copy number pMB1 replicon and the low copy number P1 replicon as well as a chloramphenicol resistance determinant. The pMB1 replicon allows facile propagation of the plasmid and its derivatives in strain DH5α, whereas the plasmid switches to the low copy number P1 replicon in the BR285 strain that lacks the *polA* gene product (DNA polymerase I) that is essential for replication by the pMB1 replicon. The FH450 plasmid is segregationally unstable in this background which is detectable as rapid plasmid loss in the absence of chloramphenicol selective pressure. The vector is stabilized by the insertion of the *parFGH* cassette (Hayes, 2000). Partition assays were performed as outlined previously (Martin et al., 1987; Hayes, 2000). In brief, pFH450-based plasmids were transformed into strain BR825 with selection on LB agar containing chloramphenicol. Eight colonies from this transformation were streaked on chloramphenicol plates to ensure plasmid establishment. Following ∼16 h of growth, one colony from each of these eight streaks was inoculated on LB plates without antibiotic to allow for plasmid retention or loss during the ∼20–25 generations required for subsequent colony formation. One colony from each of these

streaks was again streaked on non-selective plates to separate cells that had retained or lost the plasmid during the first round of non-selective growth. Eight colonies from each of the eight streaks were tested for plasmid retention by replica plating onto LB agar with and without chloramphenicol. Growth was assessed following ∼16 h incubation. Assays were performed at least in triplicate and are expressed as mean values with typical standard deviations of ∼10%.

Two-Hybrid Assays

Protein-protein interaction studies *in vivo* utilized the two-hybrid assay developed by Karimova et al. (1998) that has been used previously to monitor ParF self-association and the ParF-ParG interaction (Barillà and Hayes, 2003; Carmelo et al., 2005; Barillà et al., 2007; Dobruk-Serkowska et al., 2012). Test genes were fused in-frame with genes that specify the T18 and T25 fragments of *Bordetella pertussis* adenylate cyclase. When the encoded test proteins interact, the T18 and T25 fragments are brought in sufficiently close proximity to reconstitute adenylate cyclase activity in a *cya* mutant of *E. coli*. Adenylate cyclase function is monitored indirectly by assaying β-galactosidase activity (Karimova et al., 1998). Plasmid pairs producing fusions of the T18 or T25 subunits with either wild-type or mutated ParF or with ParG were cotransformed in strain SP850. Self-association of ParF or the interaction with ParG was assessed by β-galactosidase assays on cultures grown at 30°C for ∼16 h. Results are averages of at least three independent tests with typical standard deviations <10%.

Protein Purification and Biochemical Analysis

Wild-type and mutant ParF proteins and ParG protein were overproduced using pET-22b(+) expression plasmids and were purified as His-tagged derivatives as described previously (Barillà and Hayes, 2003). ATPase assays were conducted with [α-^{35}S]ATP, analyzed by thin layer chromatography, and quantified as detailed elsewhere (Barillà et al., 2005). In brief, ParF in the absence or presence of ParG at the concentrations indicated in figure legends, was incubated with 1 μCi of [α-^{35}S]ATP (1250 Ci/mmol) and unlabeled ATP at the concentrations indicated in figure legends in 30 mM Tris-HCl pH 7.5, 100 mM KCl, 5 mM $MgCl_2$, 2 mM DTT in a final volume of 16 μl at 30°C for 60 min. Aliquots (2.5 μl) were applied to polyethyleneimine cellulose plates that had been prerun in water, dried and then subjected to TLC with 0.5 M KH_2PO_4 (pH 3.5) as buffer. The plates were dried and exposed to Kodak BioMax MR film.

Higher-order assembly of wild-type and mutant ParF proteins was assessed by sedimentation assays as described previously (Barillà et al., 2005). Briefly, ParF with or without ParG at the concentrations indicated in figure legends, was incubated in 30 mM Tris-HCl pH 8.0, 100 mM KCl, 2 mM DTT, 10% glycerol in a volume of 60 μl in the absence or presence of nucleotides (2 mM) and $MgCl_2$ (5 mM) for 10 min at 30°C. Reactions were centrifuged for 30 min at 4°C at 20,800 × *g*. 20 μl of the supernatant was collected for SDS-PAGE analysis, 10 μl were retained for Bradford quantitation, and the remaining supernatant was carefully aspirated. The pellet was resuspended in 15 μl of water. The supernatant (20 μl; 33%) and pellet (15 μl; 100%) fractions were analyzed by SDS-PAGE and Coomassie Blue staining. Protein bands were quantitated with ImageJ (National Institutes of Health, Bethesda, MD, USA).

Biophysical Measurements

The ATP binding activity of wild-type and mutated ParF proteins was determined by anisotropy measurements of the fluorescent ATP analog MANT-ATP as outlined previously (Dobruk-Serkowska et al., 2012). Briefly, fluorescence anisotropy measurements were made with a Jovin-Yvon Horiba Fluoromax-3 spectrofluorimeter in a quartz microcuvette in a total volume of 150 μl in a buffer comprising 20 mM HEPES, 150 mM NaCl, 1 mM $MgCl_2$, pH 7.0. The excitation wavelength (λ_{ex}) and emission wavelengths (λ_{em}) were 356 and 442 nm, respectively. The ParF concentration was increased from 0.25 to ∼5 μM whereas the MANT-ATP concentration was 0.9 μM. Ten measurements of fluorescence anisotropy were taken for each protein increment and the average value was plotted against ParF concentration. Due to potential hydrolysis of MANT-ATP by ParF and mutant proteins during measurements (Dobruk-Serkowska et al., 2012), binding assays also were done with the non-hydrolyzable MANT-ATPγS analog using the same conditions as for MANT-ATP.

Circular dichroism (CD) spectroscopy was performed with ParF solutions diluted to ∼10 μM in 10 mM NaCl, 5 mM sodium phosphate pH 7.0. Ellipticity was determined in a quartz cell with a 0.5 cm path length (Fothergill et al., 2005). CD measurements were performed with a Jasco J-810 spectropolarimeter scanning from 190 to 260 nm at 25°C. Scans are the averages of at least three accumulations and were corrected against buffer-only spectra. Data were analyzed on the DichroWeb website (Lobley et al., 2002; Whitmore and Wallace, 2004) using CONTIN and CDSSTR software (Provencher et al., 1981; Compton and Johnson, 1986; Sreerama and Woody, 2000).

RESULTS

Alanine Scanning Mutagenesis of Helix α5 in ParF

ParF co-structures with ADP and with the non-hydrolyzable AMPPCP nucleotides are available. The nucleotides in both co-structures are embedded within a surface-exposed cavity of ParF. This niche is formed by residues from the Walker A nucleotide binding motif and other amino acids which make tight contacts with the nucleotide and with a hexacoordinated magnesium ion (Schumacher et al., 2012). However, the ADP and AMPPCP ribose moieties adopt different puckers in the ParF structures: the ribose of the bound AMPPCP molecule assumes a C3′-*endo* conformation whereas a C2′-*endo* conformation is evident in the ParF-ADP structure. The C3′ conformation is not viable in the latter as this configuration would result in steric clash. The distance between adjacent phosphorus atoms and the orientation of these atoms relative to the sugar and bases differ significantly in these two ribose configurations. Thus, the distinctive nucleotide sugar puckers in the ADP- and AMPPCP-bound states may be

a critical factor in determining whether the assembly of higher-order structures by ParF is antagonized by ADP or promoted by ATP (Schumacher et al., 2012).

Alpha-helix 5 (α5; residues 111–123) is of particular significance in the ParF structure as residue Asp-111, which is positioned at the N-terminal end of this element, makes key contacts with the ribose O3′ hydroxyl that stabilizes the C3′-*endo* conformation when ParF is bound to AMPPCP (**Figure 1**). Asp-111 is highly conserved among ParF homologs which attests further to the importance of this residue (Schumacher et al., 2012). Moreover, the C-terminal end of α5 forms a major part of one of two interfaces that underpin the ParF oligomer observed in complexes with AMPPCP. This interface is formed by interactions between residues 61–71 in α3 with 117–129 partly in α5, and 2-fold related contacts between residues 87–98 in α4 (Schumacher et al., 2012). In view of the importance of α5, and particularly of residue Asp-111, to the structure and function of ParF, the contribution of this element was probed further here.

Amino acids 111–123 that comprise α5 were subjected to alanine scanning mutagenesis (Cunningham and Wells, 1989). Thus, the codon for each residue was changed independently by site-directed mutagenesis to an alanine codon, except at positions 114 and 115 both of which specify alanine in the wild-type protein. The effects of these alterations on ParF-mediated plasmid segregation were assessed (**Figure 2A**). The wild-type segregation module conferred ∼70% plasmid retention in the assay whereas the vector without a partition cassette showed <5% retention. These values correspond to those described previously (Hayes, 2000; Fothergill et al., 2005; Barillà et al., 2007; Dobruk-Serkowska et al., 2012; Schumacher et al., 2012; Saeed et al., 2015). The S117A and V119A changes in α5 of ParF were innocuous and mutations at positions 113, 118, 120, 122, and 123 conferred partial segregation defects. In contrast, the D111A, F112A, G116A, and V121A mutations exerted potent effects and reduced plasmid maintenance levels to those of the empty vector (**Figure 2A**). These data support the functional importance of α5 in ParF-directed segregation. As the mutation of Asp-111 imparted a strong partitioning defect and as this residue makes vital interactions with the ribose moiety that stabilizes the C3′-*endo* conformation when ParF is in contact with the ATP analog AMPPCP (**Figure 1**), the ParF-D111A protein was purified and characterized further. CD analysis revealed that the gross structure of the protein was not disrupted appreciably by the D111A mutation (**Figure 2B**).

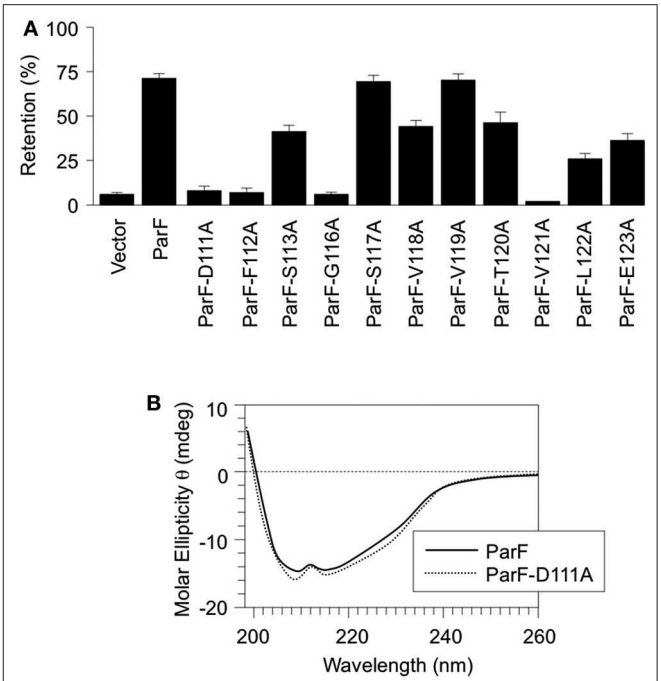


FIGURE 2 | Alanine scanning mutagenesis of α5 in ParF. **(A)** Segregation assays using pFH547-based plasmids that contain the *parFGH* cassette cloned in the pFH450 partition assay vector. Plasmids encoded either wild-type ParF or ParF derivatives with the indicated site-specific mutations. Partition assays are means of at least three independent tests with typical standard deviations of <10%. **(B)** Far UV CD spectra of wild-type ParF (solid line) and ParF-D111A (dashed line). The differences in the spectra between the wild-type and mutated proteins are within the reproducibility limits of the technique. The CD spectrum of wild-type ParF was reported previously (Dobruk-Serkowska et al., 2012).

The D111A Mutation Perturbs ATP Hydrolysis by ParF

ATP binding elicits the assembly of ParF into higher-order assemblies. The formation of these structures is regulated by the protein's weak ATPase activity ($K_{0.5}$ ∼100 μM) that generates an ADP-bound form of ParF which antagonizes assembly (Barillà et al., 2005). Mutations in ParF that either inhibit or enhance ATP binding or hydrolysis interfere both with the generation of higher-order ParF assemblies and with plasmid maintenance which confirm the crucial role of nucleotide binding in ParF meshwork formation and in segregation (Barillà et al., 2005; Dobruk-Serkowska et al., 2012). The ATPase kinetics of ParF-D111A were compared with those of the wild-type protein under *in vitro* conditions that were used previously to characterize ParF and other mutant derivatives. First, the proteins were tested at fixed concentrations (4 μM) with titrations of ATP up to 400 μM (**Figure 3A**). Wild-type ParF displayed the hyperbolic curve previously noted with maximal activity at ∼250 μM ATP (Barillà et al., 2005; Dobruk-Serkowska et al., 2012). However, ATP hydrolysis was impaired in ParF-D111A: the protein showed a modest increase in activity at ATP concentrations up to ∼100 μM but, unlike wild-type ParF, hydrolysis was not enhanced at higher nucleotide concentrations (**Figure 3A**). Second, a fixed ATP concentration (5 μM) was used in reactions with up to 10 μM ParF or ParF-D111A (**Figure 3B**). Hydrolysis was enhanced with increasing concentrations of the wild-type protein and tended toward a plateau at ∼6 μM ParF. ParF-D111A displayed weaker ATPase activity than the wild-type protein using the same protein concentration range with maximal ADP production less than half that observed with wild-type ParF (**Figure 3B**).

The nucleotide binding dynamics of ParF and ParF-D111A were monitored by fluorescence anisotropy using the fluorescent ATP analog MANT-ATP (0.9 μM) titrated with the two proteins

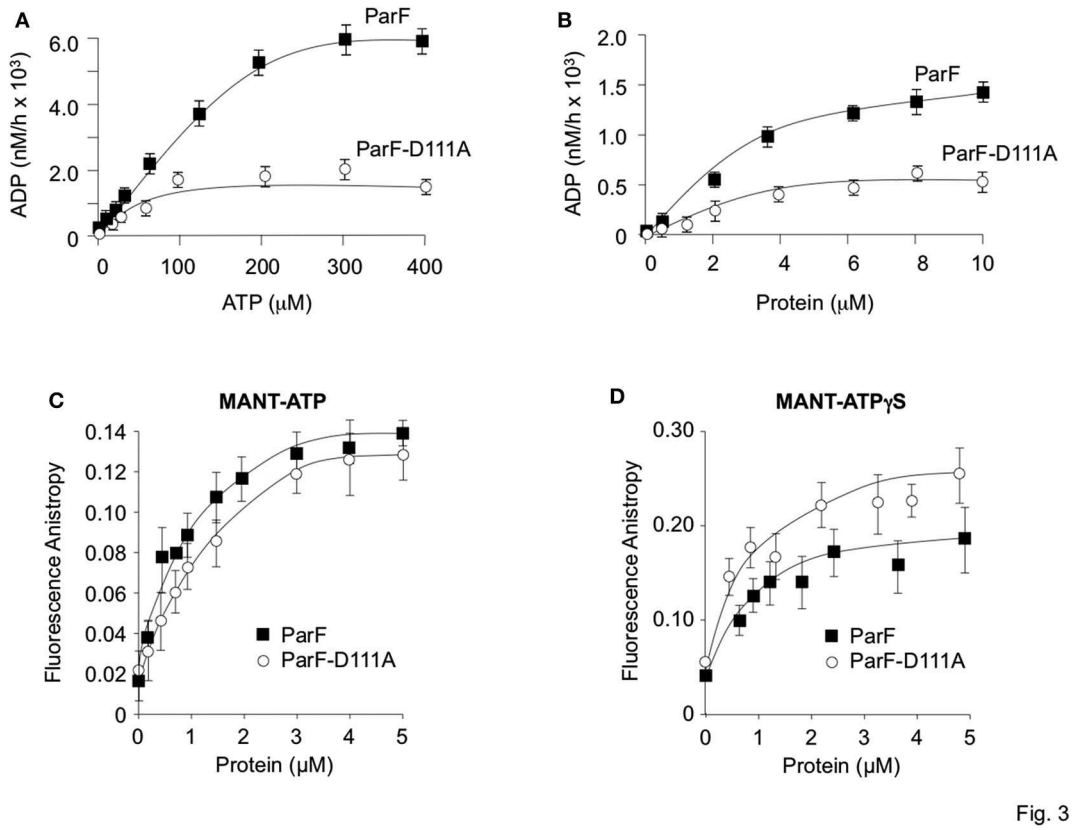


FIGURE 3 | Comparative ATP hydrolysis and binding assays of wild-type ParF and ParF-D111A. **(A)** ATP hydrolysis is plotted with proteins (4 μM) at 0–400 μM ATP concentrations. **(B)** ATP hydrolysis is plotted with ATP (5 μM) at 0–10 μM ParF or ParF-D111A concentrations. **(C)** Anisotropy changes when MANT-ATP (0.9 μM) was titrated with increasing concentrations of wild-type ParF (filled squares) and ParF-D111A (open circles) proteins. **(D)** Anisotropy changes when MANT-ATPγS (0.9 μM) was titrated with increasing concentrations of wild-type ParF (filled squares) and ParF-D111A (open circles) proteins. The average fluorescence anisotropy values for 10 measurements for each point are shown in **(C,D)**.

(**Figure 3C**). At low concentrations of the wild-type protein, MANT-ATP anisotropy elevated rapidly followed by a more gradual increase with an apparent K_d of ∼0.5 μM which matches the value determined previously in anisotropy experiments with ParF (Dobruk-Serkowska et al., 2012; Schumacher et al., 2012). The ParF-D111A protein showed a very similar binding curve (**Figure 3C**). Fluorescence anisotropy studies also were performed with the non-hydrolyzable analog MANT-ATPγS (**Figure 3D**). As there essentially is no hydrolysis within the time scale of the experiment, MANT-ATPγS anisotropy traces represent the true ligand binding curves. Wild-type ParF and ParF-D111A each elicited a rapid initial increase in MANT-ATPγS anisotropy. Values with ParF leveled off at concentrations ∼2 μM, yielding an apparent K_d of ∼0.4 μM which is similar to that determined for MANT-ATP (**Figure 3C**) and to that established previously for ParF with MANT-ATPγS (Dobruk-Serkowska et al., 2012). Although the binding curve for ParF-D111A with MANT-ATPγS was slightly more shallow than the trace for ParF, the patterns were very similar with an apparent K_d (∼0.5 μM) for ParF-D111A that was very close to that for the wild-type protein. Overall the fluorescence anisotropy experiments showed that both ParF and ParF-D111A bind MANT-ATP and non-hydrolyzable MANT-ATPγS similarly. Thus, the impaired ATPase activity of ParF-D111A is not due principally to a defect in nucleotide binding, but instead arises from a decreased rate of ATP hydrolysis which correlates with the role of Asp-111 in stabilizing the C3′-*endo* conformation when ParF is bound to AMPPCP (Schumacher et al., 2012).

The D111A Mutation Poisons the Function of a ParF Hyperactive ATPase Mutant

Residue Pro-104 is conserved in ParF and related members of the ParA superfamily, but is located distantly in the primary sequence from canonical nucleotide interaction motifs (Hayes, 2000). The residue is part of a proline-rich patch in ParF that inserts into a niche near to the ATP binding pocket of the neighboring subunit in ParF dimers (Schumacher et al., 2012). Although the role of this patch awaits elucidation, the P104A mutation induces ATPase hyperactivity in ParF that results from reorganization of the catalytic pocket. The mutation is thought

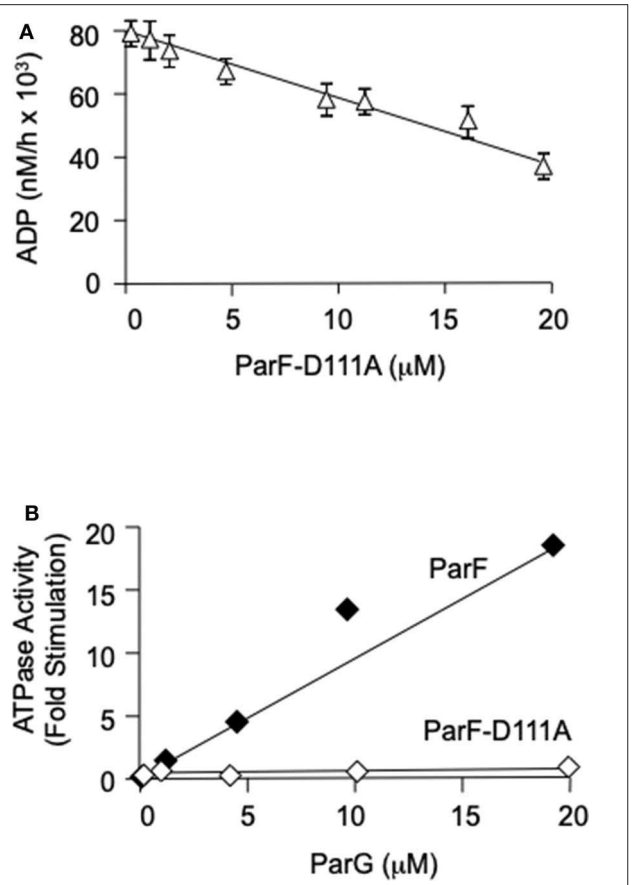


FIGURE 4 | (A) Poisoning of the ATPase hyperactivity of ParF-P104A by D111A using 200 μM ATP. ParF-104A (4 μM) and ParF-D111A (0–20 μM) were added sequentially to the reaction buffer followed by the immediate addition of nucleotide. The results are averages of at least three replicates. **(B)** The ATPase activity of ParF-D111A is not stimulated by ParG. Levels of ATP hydrolysis driven by ParF ParF-D111A as a function of ParG concentration are shown. ParF proteins were used at 1 μM and ParG at 0-20 μM with 5 μM ATP. The data are expressed as fold stimulation of ATPase activity compared with basal activity without added ParG.

to modify nucleotide access and stability in the pocket (Dobruk-Serkowska et al., 2012). Mixing experiments were conducted in which purified ParF-P104A (4 μM) was co-incubated with increasing concentrations of ParF-D111A (0–20 μM) in the presence of 200 μM ATP (**Figure 4A**). The ParF-P104A protein alone showed the characteristic ATPase hyperactivity described previously which manifested as an ∼20-fold increase in ADP production compared to wild-type ParF (Dobruk-Serkowska et al., 2012). This activity was ablated modestly by ParF-D111A which reduced nucleotide hydrolysis by ParF-P104A by approximately half at the highest concentration of ParF-D111A that was tested. The data demonstrate that ParF-D111A interacts with ParF-P104A to impair the hyperactive ATPase properties of the latter, potentially through the formation of assorted dimers. Moreover, the results confirm that the D111A mutation does not grossly perturb the self-association properties of ParF.

The D111A Mutation Impairs the Interaction of ParF With ParG

Self-association of ParF and the interaction between ParF and the partner segregation protein ParG can be monitored in two-hybrid assays *in vivo* (Barillà and Hayes, 2003; Fothergill et al., 2005; Barillà et al., 2007). The proteins under study are fused with the T18 and T25 polypeptide fragments that form the catalytic domain of the adenylate cyclase protein of *B. pertussis*. If the test proteins interact, the T18 and T25 fragments are brought into sufficiently close spatial proximity that adenylate cyclase enzymatic activity is reconstituted and therefore cyclic AMP synthesis is restored in an *E. coli cya* mutant (Karimova et al., 1998). Cyclic AMP triggers the expression of numerous catabolic operons, including the lactose operon. Therefore, the interaction between the test proteins can be assessed semi-quantitatively by measuring β-galactosidase levels (Karimova et al., 1998). The unfused adenylate cyclase polypeptides or ParF fused to one fragment, but not to the second domain, produced <55 β-galactosidase units. In contrast, self-association of wild-type ParF fused to the T18 and T25 polypeptides generated ∼700 β-galactosidase units (**Figure 5A**) as noted previously (Dobruk-Serkowska et al., 2012). The D111A mutation decreased β-galactosidase levels by approximately one-third when the alteration was included in both the ParF-T18 and ParF-T25 fusions. However, heterodimerization of wild-type ParF with ParF-D111A when fused to T18 and T25, respectively, or vice versa, resulted in similar β-galactosidase values to that observed for homodimerization of the wild-type protein (**Figure 5A**).

The effect of the D111A mutation in ParF on the interaction with ParG also was examined in two-hybrid assays. In contrast to the modest effect that the mutation exerted on ParF self-association, the D111A change abolished detectable association of ParF with ParG *in vivo* (**Figure 5B**). Thus, the two-hybrid analyses reveal that ParF-D111A is specially impaired in interaction with the partner segregation protein ParG. Thus, in parallel with the perturbation in nucleotide binding described above, the D111A mutation may disrupt the nucleotide binding pocket in ParF sufficiently that the interaction with the arginine-like finger in ParG is compromised thereby leading to loss of detectable interaction with the latter in two-hybrid tests (**Figure 5B**). Accordingly, whereas the ATPase activity of wild-type ParF was stimulated up to ∼20-fold by ParG as observed previously (Barillà et al., 2005, 2007; Dobruk-Serkowska et al., 2012), the ParF-D111A protein was unresponsive to ParG stimulation (**Figure 4B**).

Self-Association Properties of ParF-D111A *in vitro*

Purified ParF self-associates modestly in the absence of added nucleotide *in vitro*. Self-association is enhanced strongly by ATP, but is inhibited by ADP, which indicates that the cycle of ParF assembly and disassembly is regulated in part by hydrolysis of bound nucleotide (Barillà et al., 2005, 2007; Machón et al., 2007; Dobruk-Serkowska et al., 2012; Schumacher et al., 2012). The formation of higher-order ParF complexes is monitored by sedimentation assays in which the protein is incubated with

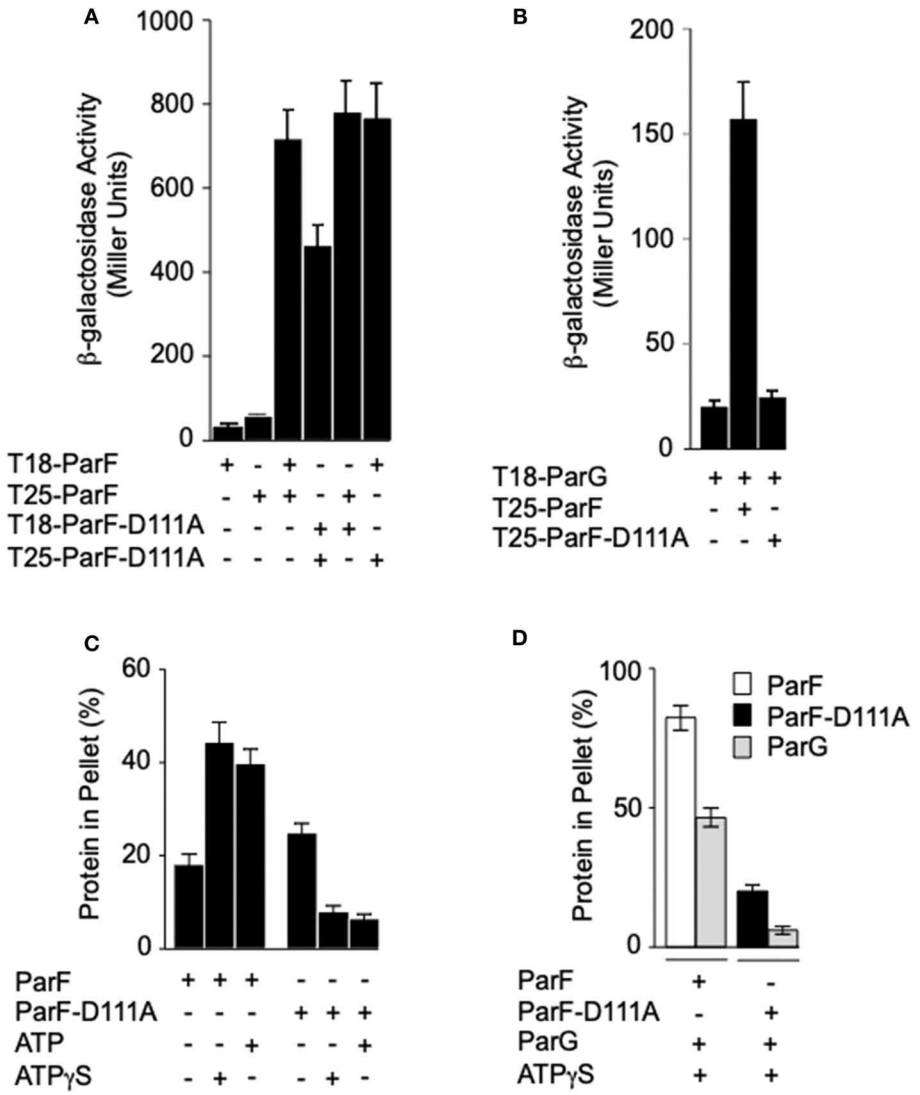


FIGURE 5 | Effects of the D111A mutation in ParF on protein-protein interactions and ParF assembly into higher-order structures. **(A)** Two-hybrid analysis of ParF-D111A self-association and interaction with wild-type ParF. Fusions of wild-type or ParF-D111A to the T18 and T25 subunits of adenylate cyclase were tested in *E. coli* SP850. **(B)** Two-hybrid analysis of ParF-D111A interaction with ParG. Fusions of wild-type ParF or ParF-D111A to the T25 subunit of adenylate cyclase were tested with a fusion of ParG to the T18 subunit in *E. coli* SP850. Results in panels A and B are means of at least three independent tests with typical standard deviations <10%. **(C)** Kinetics of wild-type ParF and ParF-D111A assembly into higher-order species. Sedimentation assays were performed in which proteins (4–8 μM) were incubated in the absence (–) or presence of nucleotides (2 mM) for 10 min at 30°C, and the reactions were then centrifuged. In all, 100 and 33%, respectively, of the pellet and supernatant fractions were resolved on a 12% SDS gel and stained with Coomassie Blue (Barillà et al., 2005). The percentages of ParF and ParF-D111A proteins detected in the pellet fractions are shown. **(D)** Co-sedimentation assays of wild-type ParF and ParF-D111A with ParG. Sedimentation assays were performed in which ParF proteins (4-8 μM) and ParG (~5 μM) were incubated in the absence (–) or presence of ATPγS (2 mM) for 10 min at 30°C, and the reactions were then centrifuged. In all, 100 and 33%, respectively, of the pellet and supernatant fractions were resolved on a 12% SDS gel and stained with Coomassie Blue. The percentages of ParF, ParF-D111A, and ParG proteins detected in the pellet fractions are shown.

or without added nucleotide, and separated by centrifugation into pellet, and supernatant fractions that harbor oligomeric, and non-oligomeric protein species, respectively. The samples are analyzed by SDS-PAGE and the concentrations of ParF in the two fractions are determined (Barillà et al., 2005). Here, the addition of ATP or non-hydrolyzable ATPγS increased the levels of higher-order assemblies of wild-type ParF in the pellet fraction ~2-fold compared to the association levels in the absence of nucleotides (**Figure 5C**) which correlates with previous observations (Barillà et al., 2005, 2007; Machón et al., 2007; Dobruk-Serkowska et al., 2012; Schumacher et al., 2012). The concentrations of wild-type ParF and ParF-D111A in the pellet fractions were similar without added nucleotide. In contrast to observations with wild-type protein, both ATP

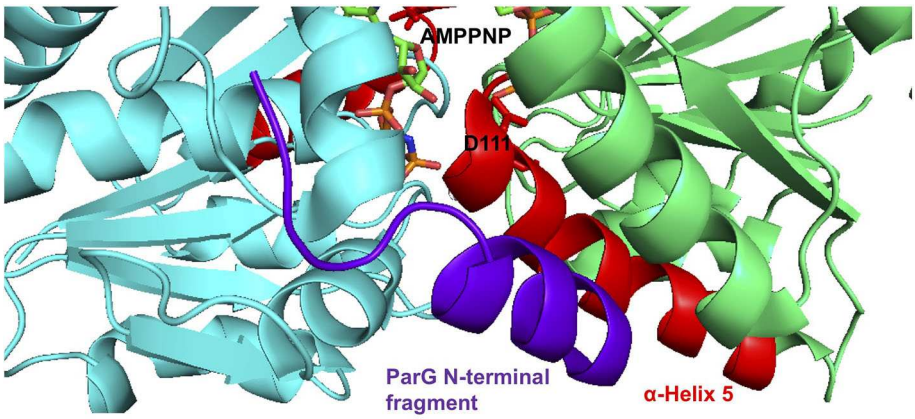


FIGURE 6 | Co-crystal structure of the ParF dimer with AMPPNP and a 15 residue fragment from the N-terminal tail of ParG (Zhang and Schumacher, 2017). α-5 is highlighted in red with residue Asp-111 and AMPPNP shown as sticks. The ParG fragment is shown in purple. The images were made using PyMol (Delano, 2001).

and ATPγS markedly decreased the levels of ParF-D111A that entered the pellet. The ATPase activity of the mutant protein is compromised at the hydrolysis step, but not in nucleotide binding, as detailed above. As a consequence, the ineffective interaction with ATP may lock ParF-D111A into a conformation that is refractory to form higher-order structures.

ParG promotes the assembly of ParF bundles in the absence of added nucleotide, and bundling is stimulated further in the presence of ATP. ParG co-sediments partly with ParF in these reactions, but ParG alone fails to enter the pellet fraction (Barillà et al., 2005). Analogous results were observed again here: ∼80% of wild-type ParF and ∼50% of ParG were found in the pellet fraction, when the proteins were coincubated with ATPγS. In contrast, ParG neither stimulated the assembly of ParF-D111A nor entered into the pellet fraction when coincubated with the mutant protein (**Figure 5D**). These observations correlate with the impaired interaction between ParG and ParF-D111A that was evident in two-hybrid analysis (**Figure 5B**) and ATPase stimulation assays (**Figure 4B**).

DISCUSSION

ParF is a multifunctional ParA-type segregation protein encoded by the TP228 multidrug resistance plasmid. The protein binds and hydrolyzes ATP, assembles into higher-order species upon ATP binding, and interacts with the partner protein ParG which modulates the ATPase activity and multimerization properties of ParF (Barillà et al., 2005, 2007; Dobruk-Serkowska et al., 2012). These and other characteristics of ParF and ParG form the basis of the "Venus flytrap" mechanism of plasmid segregation: ParF forms a three-dimensional meshwork that penetrates the nucleoid interior *in vivo* where it recognizes and transports ParG-*parH* complexes and the attached plasmids to the nucleoid poles. Plasmids are deposited at the poles following the partial dissolution of the ParF network through a combination of ParG-mediated bundle disassembly and localized ATP hydrolysis within the meshwork (McLeod et al., 2017).

The "Venus flytrap" model of plasmid segregation and the action of ParA proteins more generally are contingent on the binding and hydrolysis of ATP. Residue Asp-111, which is positioned at the N-terminal end of α5 in ParF, makes crucial contacts with the ribose O3′ hydroxyl that stabilize the C3′-*endo* conformation when ParF is bound to the ATP analog, AMPPCP. Ser-108 in the preceding loop region also participates in this stabilization (Schumacher et al., 2012). Moreover, Asp-111 is part of a proline-rich region (residues 102–112) in ParF. This region in one subunit inserts into a side pocket near the nucleotide-binding pocket of the adjacent ParF subunit. These cross-contacts by the proline-rich region are proposed to stabilize the ParF-ATP sandwich dimer (Schumacher et al., 2012). The significance of Asp-111 was explored further here. First, the D111A mutation dramatically reduced plasmid segregation *in vivo*. Mutation of other residues, notably Phe-112, Gly-116, and Val-121, in α5 also negatively impacted plasmid partitioning (**Figure 2A**). Although the role(s) of these amino acids awaits further investigation, this triad of residues may be important in ParF subunit interactions. Second, the D111A mutation modestly affected nucleotide binding, but impaired ATP hydrolysis more conspicuously (**Figure 3**) which broadly agrees with a role for Asp-111 in stabilizing the C3′-*endo* pucker of the ribose ring. The sugar may undergo a switch to a less favorable C2′-*endo* conformation when alanine substitutes for aspartic acid at position 111. This physicochemical alteration in the structure of the ribose ring may influence the cleavage of the γ phosphate group in the bound ATP (Kobayashi et al., 2013). It remains to be determined whether the C2′-endo conformation is compatible with ATP binding and/or how this conformation might interfere with ATP hydrolysis by wild-type ParF. Third, ATP binding promotes ParF assembly in oligomeric bundles (Barillà et al., 2005). Nevertheless, although ParF-D111A still bound ATP effectively, formation of higher-order structures was inhibited fully *in vitro* (**Figure 5C**). Thus, the perturbed conformation of the nucleotide within the binding pocket of the ParF-D111A variant blocks the assembly of ParF into higher

order species. Despite this inhibition, ParF-D111A dimerized in two-hybrid analysis, albeit not as effectively as the wild-type protein (**Figure 5A**). As a suite of amino acids are involved in dimer formation by ParF (Schumacher et al., 2012), mutation of Asp-111 alone apparently is sufficient to weaken but not abolish the two-hybrid interaction.

The ParG protein plays multiple roles in plasmid segregation. The protein interacts both with the *parH* centromere during segrosome assembly and with the operator site upstream of the *parFG* genes to exert transcriptional repression (Barillà and Hayes, 2003; Carmelo et al., 2005; Zampini et al., 2009; Wu et al., 2011). DNA binding is influenced by a pair of unstructured N-terminal tails in the ParG dimer (Carmelo et al., 2005; Wu et al., 2011). The N-terminal tails also modulate both ATP hydrolysis and assembly of ParF. For the former, arginine finger-like motifs in the ParG flexible tails enhance nucleotide hydrolysis by ParF *in trans*, possibly by stabilization of the transition state through neutralization of the negative charge that develops during phosphoryl transfer (Barillà et al., 2007). However, the arginine finger-like motif in ParG did not enhance ATP hydrolysis by the ParF-D111A mutant protein (**Figure 4B**). The effect of the mutation on the conformation of the ribose ring in the bound nucleotide may preclude formation of the appropriate contacts by the ParG arginine finger when inserted in the ATP binding pocket. The determinants for stimulation of ParF oligomerisation by the N-terminal tails of ParG have not been defined, but are separable from the arginine finger-like motif. One or more residues within the tails are thought to play an architectural role that assists in organizing the assembly of ParF subunits into higher-order species (Barillà et al., 2007). The D111A mutation in the nucleotide binding pocket abolished the stimulation of ParF assembly in bundles by ParG (**Figure 5D**). This observation reveals that the block to ParF oligomerisation that is caused by the mutation is not alleviated by the stimulatory effect of the partner protein, and confirms that the conformation of the ATP binding site is critical for effective ParF assembly into higher-order structures. The profound effect that the D111A mutation has on the interaction of ParF with ParG is reflected further in the lack of detectable interaction between ParF-D111A and ParG in two-hybrid assays (**Figure 5B**). NMR investigation showed that residues 17–23 of the ParG N-terminal flexible tail are characterized by limited flexibility and suggested that this region might have some α -helical content (Golovanov et al., 2003). These studies also highlighted that the limited flexibility of the 17–23 region might result in entropic benefits if this fragment was to become configured rigidly when interacting with ligands, thus overall suggesting that the region might contain interaction sites. Recently, a cocrystal structure of ParF bound to a short, ∼15 residue fragment of the N-terminal tail of ParG in the presence of AMPPNP was solved (Zhang and Schumacher, 2017). Although the structure has a low 3.65 Å resolution, it revealed that the ParG fragment (amino acids 8–22) folded into a short helix that was positioned next to α-helix five of ParF within a cleft at the base of the ParF dimer interface. These structural data provide evidence that ParF α-helix five is indeed part of one interface in ParF-ParG interaction (**Figure 6**). The absence of association observed between ParF-D111A and ParG is in agreement with the structural results and indicates that the replacement of a single residue is sufficient to disrupt the interaction interface to the point of ablating the protein-protein contact.

In conclusion, the "Venus flytrap" model of plasmid segregation necessitates the binding and hydrolysis of ATP by the ParA-type protein, ParF. The current study demonstrates that the conformation of the nucleotide binding pocket in ParF is tuned exquisitely: a single amino acid alteration that perturbs the molecular arrangement of the bound nucleotide affects not only ATP hydrolysis, but also the interaction with the partner protein ParG. The conformation of the ATP binding site in ParF profoundly impacts not only the kinetics of nucleotide hydrolysis and the capacity of an arginine finger-like motif in the N-terminal flexible tails of the ParG protein to stimulate nucleotide hydrolysis *in trans*, but also the ParF higher order dynamics that drive the segregation process. Thus, the dynamics of nucleotide binding and hydrolysis by ParA-type proteins are key to accurate genome segregation in bacteria.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation, to any qualified researcher.

AUTHOR CONTRIBUTIONS

FH and DB conceived and designed the project with input from all authors and wrote the manuscript with input from all authors. MC, AD-S, FR-C, and CP performed the experiments and analyzed the data.

FUNDING

This work was supported by Biotechnology and Biological Sciences Research Council Grants G003114 (to FH) and by BB/F012004/1 (to DB). This work also benefitted from Wellcome Trust Value in People Award 092403 (to FH and MC). CP was supported by Biotechnology and Biological Sciences Research Council White Rose Doctoral Training Partnership BB/J014443/1.

REFERENCES

Badrinarayanan, A., Le, T. B., and Laub, M. T. (2015). Bacterial chromosome organization and segregation. *Annu. Rev. Cell. Dev. Biol.* 31, 171–199. doi: 10.1146/annurev-cellbio-100814-125211

Banigan, E. J., Gelbart, M. A., Gitai, Z., Wingreen, N. S., and Liu, A. J. (2011). Filament depolymerization can explain chromosome pulling during bacterial mitosis. *PLoS Comput. Biol.* 7:e1002145. doi: 10.1371/journal.pcbi.1002145

Barillà, D., Carmelo, E., and Hayes, F. (2007). The tail of the ParG DNA segregation protein remodels ParF polymers and enhances ATP hydrolysis via

an arginine finger-like motif. *Proc. Natl. Acad. Sci. U. S. A.* 104, 1811–1816. doi: 10.1073/pnas.0607216104

Barillà, D., and Hayes, F. (2003). Architecture of the ParF-ParG protein complex involved in prokaryotic DNA segregation. *Mol. Microbiol.* 49, 487–499. doi: 10.1046/j.1365-2958.2003.03564.x

Barillà, D., Rosenberg, M. F., Nobbmann, U., and Hayes, F. (2005). Bacterial DNA segregation dynamics mediated by the polymerizing protein ParF. *EMBO J.* 24, 1453–1464. doi: 10.1038/sj.emboj.7600619

Batt, S. M., Bingle, L. E., Dafforn, T. R., and Thomas, C. M. (2009). Bacterial genome partitioning: N-terminal domain of IncC protein encoded by broad-host-range plasmid RK2 modulates oligomerisation and DNA binding. *J. Mol. Biol.* 385, 1361–1374. doi: 10.1016/j.jmb.2008.12.016

Baxter, J. C., and Funnell, B. E. (2014). Plasmid partition mechanisms. *Microbiol. Spectrum*. 2:PLAS-0023-2014. doi: 10.1128/microbiolspec.PLAS-0023-2014

Bloom, K., and Joglekar, A. (2010). Towards building a chromosome segregation machine. *Nature* 463, 446–456. doi: 10.1038/nature08912

Bouet, J. Y., Ah-Seng, Y., Benmeradi, N., and Lane, D. (2007). Polymerization of SopA partition ATPase: regulation by DNA binding and SopB. *Mol. Microbiol.* 63, 468–481. doi: 10.1111/j.1365-2958.2006.05537.x

Carmelo, E., Barillà, D., Golovanov, A. P., Lian, L. Y., Derome, A., and Hayes, F. (2005). The unstructured N-terminal tail of ParG modulates assembly of a quaternary nucleoprotein complex in transcription repression. *J. Biol. Chem.* 280, 28683–28691. doi: 10.1074/jbc.M501173200

Compton, L. A., and Johnson, W. C. Jr. (1986). Analysis of protein circular dichroism spectra for secondary structure using a simple matrix multiplication. *Anal. Biochem.* 155, 155–167. doi: 10.1016/0003-2697(86)90241-1

Cunningham, B. C., and Wells, J. A. (1989). High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis. *Science* 244, 1081–1085. doi: 10.1126/science.2471267

Delano, W. L. (2001). *The PyMOL Molecular Graphics System. San* Carlos, CA: *DeLano Scientific LLC*.

Ditkowski, B., Holmes, N., Rydzak, J., Donczew, M., Bezulska, M., Ginda, K., et al. (2013). Dynamic interplay of ParA with the polarity protein, Scy, coordinates the growth with chromosome segregation in *Streptomyces coelicolor*. *Open Biol.* 3:130006. doi: 10.1098/rsob.130006

Ditkowski, B., Troć, P., Ginda, K., Donczew, M., Chater, K. F., Zakrzewska-Czerwinska, J., et al. (2010). The actinobacterial signature protein ParJ (SCO1662) regulates ParA polymerization and affects chromosome segregation and cell division during *Streptomyces* sporulation. *Mol. Microbiol.* 78, 1403–1415. doi: 10.1111/j.1365-2958.2010.07409.x

Dobruk-Serkowska, A., Caccamo, M., Rodríguez-Castañeda, F., Wu, M., Bryce, K., Ng, I., et al. (2012). Uncoupling of nucleotide hydrolysis and polymerization in the ParA protein superfamily disrupts DNA segregation dynamics. *J. Biol. Chem.* 287, 42545–42553. doi: 10.1074/jbc.M112.410324

Ebersbach, G., Ringgaard, S., Møller-Jensen, J., Wang, Q., Sherratt, D. J., and Gerdes, K. (2006). Regular cellular distribution of plasmids by oscillating and filament-forming ParA ATPase of plasmid pB171. *Mol. Microbiol.* 61, 1428–1442. doi: 10.1111/j.1365-2958.2006.05322.x

Fothergill, T. J. G., Barillà, D., and Hayes, F. (2005). Protein diversity confers specificity in plasmid segregation. *J. Bacteriol.* 187, 2651–2661. doi: 10.1128/JB.187.8.2651-2661.2005

Gerdes, K., Howard, M., and Szardenings, F. (2010). Pushing and pulling in prokaryotic DNA segregation. *Cell* 141, 927–942. doi: 10.1016/j.cell.2010.05.033

Golovanov, A. P., Barillà, D., Golovanova, M., Hayes, F., and Lian, L. Y. (2003). ParG, a protein required for active partition of bacterial plasmids, has a dimeric ribbon-helix-helix structure. *Mol. Microbiol.* 50, 1141–1153. doi: 10.1046/j.1365-2958.2003.03750.x

Hanahan, D. (1985). *DNA Cloning, Vol. 1.* ed D. M. Glover (Oxford: IRL Press), 109–135.

Hatano, T., and Niki, H. (2010). Partitioning of P1 plasmids by gradual distribution of the ATPase ParA. *Mol. Microbiol.* 78, 1182–1198. doi: 10.1111/j.1365-2958.2010.07398.x

Hatano, T., Yamaichi, Y., and Niki, H. (2007). Oscillating focus of SopA associated with filamentous structure guides partitioning of F plasmid. *Mol. Microbiol.* 64, 1198–1213. doi: 10.1111/j.1365-2958.2007.05728.x

Hayes, F. (2000). The partition system of multidrug resistance plasmid TP228 includes a novel protein that epitomizes an evolutionarily-distinct subgroup of the ParA superfamily. *Mol. Microbiol.* 37, 528–541. doi: 10.1046/j.1365-2958.2000.02030.x

Hayes, F., and Barillà, D. (2006). The bacterial segrosome: a dynamic nucleoprotein machine for DNA trafficking and segregation. *Nat. Rev. Microbiol.* 4, 133–143. doi: 10.1038/nrmicro1342

Hayes, F., and Barillà, D. (2010). "Extrachromosomal components of the nucleoid: recent developments in deciphering the molecular basis of plasmid segregation," in: *Bacterial Chromatin*. eds. C. J. Dorman and R. T. Dame (Dordrecht: Springer Publishing), 49–70. doi: 10.1007/978-90-481-3473-1_4

Higuchi, R., Krummel, B., and Saiki, R. (1988). A general method of *in vitro* preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. *Nucleic Acids Res.* 16, 7351–7367. doi: 10.1093/nar/16.15.7351

Hirano, T. (2015). Chromosome dynamics during mitosis. *Cold Spring Harb. Perspect. Biol.* 7:a015792. doi: 10.1101/cshperspect.a015792

Hui, M. P., Galkin, V. E., Yu, X., Stasiak, A. Z., Stasiak, A., Waldor, M. K., et al. (2010). ParA2, a *Vibrio cholerae* chromosome partitioning protein, forms left-handed helical filaments on DNA. *Proc. Natl. Acad. Sci. U. S. A.* 107, 4590–4595. doi: 10.1073/pnas.0913060107

Hwang, L. C., Vecchiarelli, A. G., Han, Y. W., Mizuuchi, M., Harada, Y., Funnell, B. E., et al. (2013). ParA-mediated plasmid partition driven by protein pattern self-organization. *EMBO J.* 32, 1238–1249. doi: 10.1038/emboj.2013.34

Kalliomaa-Sanford, A. K., Rodriguez-Castañeda, F. A., McLeod, B. N., Latorre-Roselló, V., Smith, J. H., Reimann, J., et al. (2012). Chromosome segregation in archaea mediated by a hybrid DNA partition machine. *Proc. Natl. Acad. Sci. U. S. A.* 109, 3754–3759. doi: 10.1073/pnas.1113384109

Karimova, G., Pidoux, J., Ullmann, A., and Ladant, D. (1998). A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proc. Natl. Acad. Sci. U. S. A.* 95, 5752–5756. doi: 10.1073/pnas.95.10.5752

Kobayashi, E., Yura, K., and Nagai, Y. (2013). Distinct conformation of ATP molecule in solution and on protein. *Biophysics* 9, 1–12. doi: 10.2142/biophysics.9.1

Le Gall, A., Cattoni, D. I., Guilhas, B., Mathieu-Demazière, C., Oudjedi, L., Fiche, J. B., et al. (2016). Bacterial partition complexes segregate within the volume of the nucleoid. *Nat. Commun.* 7:12107. doi: 10.1038/ncomms12107

Leonard, T. A., Butler, P. J., and Lowe, J. (2005). Bacterial chromosome segregation: structure and DNA binding of the Soj dimer - a conserved biological switch. *EMBO J.* 24, 270–282. doi: 10.1038/sj.emboj.7600530

Lim, G. E., Derman, A. I., and Pogliano, J. (2005). Bacterial DNA segregation by dynamic SopA polymers. *Proc. Natl. Acad. Sci. U. S. A.* 102, 17658–17663. doi: 10.1073/pnas.0507222102

Lim, H. C., Surovtsev, I. V., Beltran, B. G., Huang, F., Bewersdorf, J., and Jacobs-Wagner, C. (2014). Evidence for a DNA-relay mechanism in ParABS-mediated chromosome segregation. *Elife* 3:e02758. doi: 10.7554/eLife.02758

Lobley, A., Whitmore, L., and Wallace, B. A. (2002). DICHROWEB: an interactive website for the analysis of protein secondary structure from circular dichroism spectra. *Bioinformatics* 18, 211–212. doi: 10.1093/bioinformatics/18.1.211

Ludtke, D. N., Eichorn, B. G., and Austin, S. J. (1989). Plasmid partition functions of the P7 prophage. *J. Mol. Biol.* 209, 393–406. doi: 10.1016/0022-2836(89)90005-3

Machón, C., Fothergill, T. J. G., Barillà, D., and Hayes, F. (2007). Promiscuous stimulation of ParF protein polymerization by heterogeneous centromere binding factors. *J. Mol. Biol.* 374, 1–8. doi: 10.1016/j.jmb.2007.09.025

Martin, K. A., Friedman, S. A., and Austin, S. J. (1987). Partition site of the P1 plasmid. *Proc. Natl. Acad. Sci. U. S. A.* 84, 8544–8547. doi: 10.1073/pnas.84.23.8544

McLeod, B. N., Allison-Gamble, G. E., Barge, M. T., Tonthat, N. K., Schumacher, M. A., Hayes, F., et al. (2017). A three-dimensional ParF meshwork assembles through the nucleoid to mediate plasmid segregation. *Nucleic Acids Res.* 45, 3158–3171. doi: 10.1093/nar/gkw1302

Oliva, M. A. (2016). Segrosome complex formation during DNA trafficking in bacterial cell division. *Front. Mol. Biosci.* 3:51. doi: 10.3389/fmolb.2016.00051

Pratto, F., Cicek, A., Weihofen, W. A., Lurz, R., Saenger, W., and Alonso, J. C. (2008). *Streptococcus pyogenes* pSM19035 requires dynamic assembly of ATP-bound ParA and ParB on *parS* DNA during plasmid segregation. *Nucleic Acids Res.* 36, 3676–3689. doi: 10.1093/nar/gkn170

Provencher, S. W., and, J., Glockner (1981). Estimation of protein secondary structure from circular dichroism. *Biochemistry* 20, 33–37. doi: 10.1021/bi00504a006

Ringgaard, S., van Zon, J., Howard, M., and Gerdes, K. (2009). Movement and equipositioning of plasmids by ParA filament disassembly. *Proc. Natl. Acad. Sci. U. S. A.* 106, 19369–19374. doi: 10.1073/pnas.0908347106

Saeed, S., Jowitt, T. A., Warwicker, J., and Hayes, F. (2015). Breaking and restoring the hydrophobic core of a centromere-binding protein. *J. Biol. Chem.* 290, 9273–9283. doi: 10.1074/jbc.M115.638148

Schumacher, M. A., Ye, Q., Barge, M. T., Zampini, M., Barillà, D., and Hayes, F. (2012). Structural mechanism of ATP-induced polymerization of the partition factor ParF: implications for DNA segregation. *J. Biol. Chem.* 287, 26146–26154. doi: 10.1074/jbc.M112.373696

Sengupta, M., and Austin, S. (2011). Prevalence and significance of plasmid maintenance functions in the virulence plasmids of pathogenic bacteria. *Infect. Immun.* 79, 2502–2509. doi: 10.1128/IAI.00127-11

Shah, S., and Peterkofsky, A. (1991). Characterization and generation of *Escherichia coli* adenylate cyclase deletion mutants. *J. Bacteriol.* 173, 3238–3242. doi: 10.1128/JB.173.10.3238-3242.1991

Sreerama, N., and Woody, R. W. (2000). Estimation of protein secondary structure from CD spectra: comparison of CONTIN, SELCON and CDSSTR methods with an expanded reference set. *Anal. Biochem.* 287, 252–260. doi: 10.1006/abio.2000.4880

Studier, F. W., and Moffatt, B. A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* 189, 113–130. doi: 10.1016/0022-2836(86)90385-2

Vecchiarelli, A. G., Han, Y. W., Tan, X., Mizuuchi, M., Ghirlando, R., Biertümpfel, C., et al. (2010). ATP control of dynamic P1 ParA-DNA interactions: a key role for the nucleoid in plasmid partition. *Mol. Microbiol.* 78, 78–91. doi: 10.1111/j.1365-2958.2010.07314.x

Vecchiarelli, A. G., Neuman, K. C., and Mizuuchi, K. (2014). A propagating ATPase gradient drives transport of surface-confined cellular cargo. *Proc. Natl. Acad. Sci. U. S. A.* 111, 4880–4885. doi: 10.1073/pnas.1401025111

Whitmore, L., and Wallace, B. A. (2004). DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res.* 32, W668–673. doi: 10.1093/nar/gkh371

Wu, M., Zampini, M., Bussiek, M., Hoischen, C., Diekmann, S., and Hayes, F. (2011). Segrosome assembly at the pliable *parH* centromere. *Nucleic Acids Res.* 39, 5082–5097. doi: 10.1093/nar/gkr115

Zampini, M., Derome, A., Bailey, S. E. S., Barillà. D., and Hayes, F. (2009). Recruitment of the ParG segregation protein to different affinity DNA sites. *J. Bacteriol.* 191, 3832–3841. doi: 10.1128/JB.01630-08

Zampini, M., and Hayes, F. (2012). Combinatorial targeting of ribbon-helix-helix artificial transcription factors to chimeric recognition sites. *Nucleic Acids Res.* 40, 6673–6682. doi: 10.1093/nar/gks314

Zhang, H., and Schumacher, M. A. (2017). Structures of partition protein ParA with nonspecific DNA and ParB effector reveal molecular insights into principles governing Walker-box DNA segregation. *Genes. Dev.* 31, 481–492. doi: 10.1101/gad.296319.117

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Caccamo, Dobruk-Serkowska, Rodríguez-Castañeda, Pennica, Barillà and Hayes. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 10 July 2020
doi: 10.3389/fmolb.2020.00124

Temperature Dependent Control of the R27 Conjugative Plasmid Genes

Marta Gibert, Sonia Paytubi†, Cristina Madrid and Carlos Balsalobre**

Department of Genetics, Microbiology and Statistics, University of Barcelona, Barcelona, Spain

Conjugation of R27 plasmid is thermoregulated, being promoted at 25°C and repressed at 37°C. Previous studies identified plasmid-encoded regulators, HtdA, TrhR and TrhY, that control expression of conjugation-related genes (*tra*). Moreover, the nucleoid-associated protein H-NS represses conjugation at non-permissive temperature. A transcriptomic approach has been used to characterize the effect of temperature on the expression of the 205 R27 genes. Many of the 35 *tra* genes, directly involved in plasmid-conjugation, were upregulated at 25°C. However, the majority of the non-*tra* R27 genes—many of them with unknown function—were more actively expressed at 37°C. The role of HtdA, a regulator that causes repression of the R27 conjugation by counteracting TrhR/TrhY mediated activation of *tra* genes, has been investigated. Most of the R27 genes are severely derepressed at 25°C in an *htdA* mutant, suggesting that HtdA is involved also in the repression of R27 genes other than the *tra* genes. Interestingly, the effect of *htdA* mutation was abolished at non-permissive temperature, indicating that the HtdA-TrhR/TrhY regulatory circuit mediates the environmental regulation of R27 gene expression. The role of H-NS in the proposed model is discussed.

Keywords: plasmid conjugation, temperature-dependent control, TrhR/TrhY-HtdA, transcriptional regulation, R27

OPEN ACCESS

Edited by:
Manuel Espinosa, Margarita Salas Center for Biological Research (CSIC), Spain

Reviewed by:
Elisabeth Grohmann, Beuth Hochschule für Technik Berlin, Germany
Fabian Lorenzo-Diaz, University of La Laguna, Spain

****Correspondence:***
Cristina Madrid
cmadrid@ub.edu
Carlos Balsalobre
cbalsalobre@ub.edu

†Present address:
Sonia Paytubi, Infections and Cancer Laboratory, Cancer Epidemiology Research Program (CERP), Institut Català d'Oncologia (ICO)–IDIBELL, Barcelona, Spain

Specialty section:
This article was submitted to Molecular Recognition, a section of the journal Frontiers in Molecular Biosciences

Received: *30 April 2020*
Accepted: *02 June 2020*
Published: *10 July 2020*

Citation:
Gibert M, Paytubi S, Madrid C and Balsalobre C (2020) Temperature Dependent Control of the R27 Conjugative Plasmid Genes. Front. Mol. Biosci. 7:124. doi: 10.3389/fmolb.2020.00124

INTRODUCTION

Horizontal gene transfer (HGT) is a process of genetic exchange that highly contributes to evolution and adaptation of bacteria to new niches by promoting acquisition of genes coding for different metabolic pathways, toxins, adhesins, or antimicrobial resistances. Plasmid conjugation is one of the main HGT mechanisms responsible for the rapid dissemination of antibiotic resistances between pathogenic strains (Bennett, 2008; Smillie et al., 2010). In order to develop strategies to avoid multidrug spread, a detailed knowledge of the conjugation process and the mechanisms regulating that transfer process is needed.

In our research group we have focused in the study of the IncHI conjugative plasmids, which are associated with multidrug resistance in several pathogens, like *Salmonella enterica* and *Escherichia coli* (Holt et al., 2011). The conjugation of this incompatibility group plasmid is thermosensitive, showing higher conjugation frequencies at low temperatures, between 22 and 30°C (Taylor and Levine, 1980). This feature suggests that the dissemination of resistances among pathogenic species by IncHI plasmids is enhanced in water and soil environments (Maher and Taylor, 1993). In addition to temperature, the physiological state of the cell is critical to define the conjugative potential of the donor cells during the transfer process of R27 plasmid, the prototype of IncHI1 plasmids (Gibert et al., 2016). The possible role of other environmental factors in the conjugation of R27, such as osmolarity, anaerobiosis, quorum sensing and acidity, has been tested (Alonso et al., 2005). No significant influence of these environmental parameters was found on R27 transfer frequency.

The genes involved in the conjugation process of the R27 plasmid, denominated *tra* (transfer) genes, are clustered in two separated regions, Tra1 and Tra2 (Sherburne et al., 2000). Each region contains three operons: R, F and H operons in the Tra1 region (Lawley et al., 2002) and AC, Z and AN operons in the Tra2 region (Rooker et al., 1999; Lawley et al., 2003). We described a regulatory circuit, highly conserved among the IncHI plasmids, composed of three R27-encoded elements, HtdA, TrhR and TrhY, which is crucial in the temperature control of conjugation of R27 plasmid (Gibert et al., 2013, 2014). TrhR and TrhY are both required simultaneously to activate transcriptional expression of the F, H, AC and Z operons and consequently to promote R27 conjugation. HtdA has an overall negative role by counteracting the stimulatory activity of TrhR and TrhY (Gibert et al., 2013, 2014, 2016).

Most regulatory studies of R27 conjugation focused on the effect of different environmental or genetic factors on the expression of one or several *tra* genes by using gene fusions and/or mRNA quantification. In a recent report, we demonstrated that, other genetic loci located outside the *tra* regions seem to be coregulated together with the *tra* genes and, by their putative functions, might be also involved in promoting an efficient R27 conjugation (Gibert et al., 2016).

To further characterize the thermoregulation of R27 conjugation, the variations on the expression pattern of the totality of genes of the R27 plasmid at permissive (25°C) and non-permissive (37°C) temperatures was determined. Having in consideration that the HtdA-TrhR/TrhY regulatory circuit described seems to play a crucial role in the environmental control of the R27 conjugation, the transcriptomic studies have been extended to an *htdA* derivative of the R27 plasmid.

MATERIALS AND METHODS

Bacterial Strains, Plasmids, and Growth Conditions

The strains used are MG1655 (Guyer et al., 1981) and its Δ*lac* derivative AAG1 (Aberg et al., 2008) harboring either R27 (Tc^R, Grindley et al., 1972) or drR27 (R27*htdA*$^-$ Tc^R, Gibert et al., 2013) plasmids. Strain AAG1-F contains a fusion between the F operon promoter and the *lacZ* gene, located at the *attB* chromosomal locus (Gibert et al., 2013). Plasmid pBAD*trhRY* (Gibert et al., 2014) was used to overexpress the TrhR/TrhY proteins. Bacteria were routinely grown in LB (10 g/L NaCl, 10 g/L tryptone, 5 g/L yeast extract). For conjugation experiments strains were grown in PB medium (1.5 g/L meat extract, 1.5 g/L yeast extract, 5 g/L peptone, 1 g/L glucose, 3.5 g/L NaCl, 1.32 g/L KH_2PO_4, 4.82 g/L $K_2HP_4{\cdot}3H_2O$). M9 minimal media plates with the following composition: 1 x M9 salts, 0.2 % lactose, 10 μM thiamine and 1.5 % bactoagar, were used to differentiate donor from transconjugant cells in conjugation experiments, as described previously (Gibert et al., 2014). When needed, tetracycline (Tc) and arabinose were added at the concentration of 15 μg/mL and 0.02%, respectively. Unless indicated, all cultures were grown at either 25 or 37°C and under shaking (200 rpm) conditions.

Strain Constructions

lacZ fusions with the intergenic regions containing putative promoter sequences were constructed (fragments a, b and c in **Figure 4B**). An intragenic region within R0009 gene was used as a negative control (fragment d in **Figure 4B**). PCR amplification of those regions was performed using primers pairs described in **Table S1**. Primers were designed to incorporate either *Eco*RI (forward primers) or *Bam*HI (reverse primers) restriction sites. The PCR-amplified fragments were cloned in plasmid pGEM-T and subsequently in pRS551, either in *Bam*HI or *Eco*RI-*Bam*HI sites. The resulting constructs were transferred to the *attB* chromosomal locus of the AAG1 strain using previously described protocols (Simons et al., 1987). Controls to confirm single gene fusion in the *attB* locus were performed for all fusions. All genetic constructions were confirmed by DNA sequencing.

Conjugation Experiments

Mating experiments, using strains AGG1 R27 and MG1655 as donor and recipient strains, respectively, were performed as described previously (Gibert et al., 2014).

β-Galactosidase Assay

β-Galactosidase assays were performed as described previously (Miller, 1992). Data are given as mean values from at least three independent experiments, plotted with standard deviations.

Total RNA Isolation

The RNA used in microarray experiments was purified from three independent cultures grown in LB under shaking conditions. The temperature of incubation was either 25 or 37°C and samples were taken at mid logarithmic phase (log, OD_{600nm} of 0.4). The RNA was purified using an SV Total RNA Isolation System (Promega) according to the manufacturer's instructions. The RNA was DNase treated with TURBO DNAse (Ambion). After concentration, using a RNeasy Minielute Clean-up kit (Qiagen), purity and quality of the purified RNA was tested by Bionalyzer 2100 (Agilent Technologies). All samples show a RNA integrity number (RIN) over 8.0.

Microarray Analysis

Transcriptomic analysis was performed on a custom-designed DNA microarray engineered by NimbleGen, containing two replicates of seven selected probes for each of the 205 annotated genes of the R27 plasmid (NC_002305), as previously described (Paytubi et al., 2014). In this work we compare the microarray data obtained from cultures of strains harboring R27 or drR27 plasmid, grown at 25 or 37°C to mid log phase of growth. Microarray data of R27 genes expression at 25°C have been previously described in a comparative study between logarithmic and stationary phase cultures (Gibert et al., 2016). The complete dataset has been deposited under the accession number E-MTAB-9150 at http://www.ebi.ac.uk/arrayexpress.

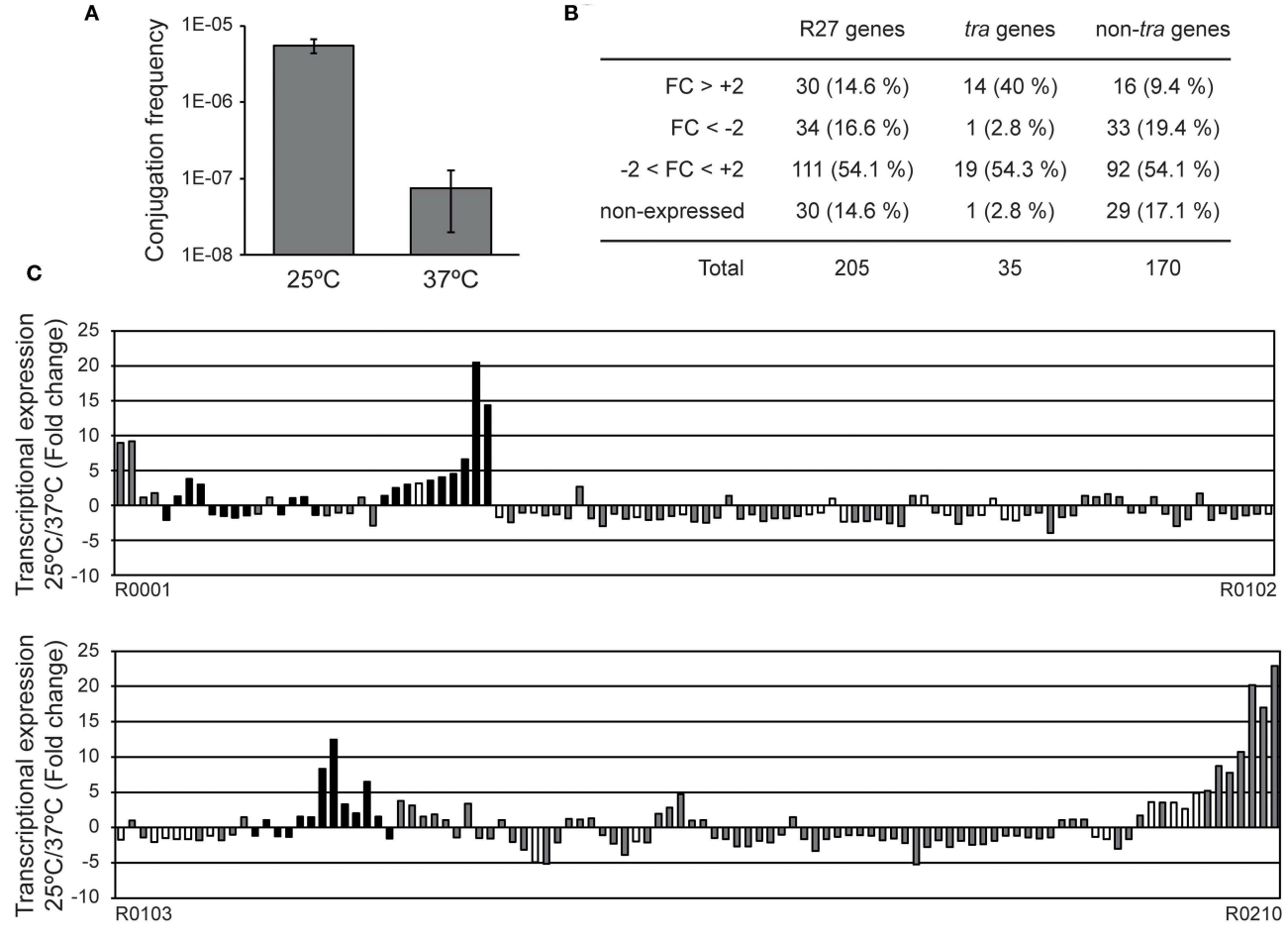
Characterization of Transcripts

To describe the transcriptional organization of the AN operon, total RNA was isolated from cultures of strain AAG1(R27) grown in LB at 25°C to an OD_{600nm} of 2.0. To characterize the 5′ and 3′ end of the transcripts generated from the AN operon, circularized

RNA was obtained as follows. CircRNA samples containing processed transcripts (monophosphated in the original RNA samples) were obtained after direct ligation of 10 μg of total RNA with CircLigase RNA Ligase (Epicentre). To obtain circRNA samples carrying primary transcripts (triphosphated in the original RNA samples), prior ligation, 10 μg of total RNA was treated with phosphatase (CIAP, Invitrogen) and subsequent RNA 5′ polyphosphatase (Epicentre). CircRNA samples were retrotranscribed to cDNA using the AMV RT (Promega) with specific primers of the AN operon (**Table S2**). cDNA sequences generated from the junction between the 5′- and 3′- ends were PCR amplified, using two subsequent rounds of PCR with the indicated primers (**Table S2**). The PCR products were further purified and sequenced with the primers used in the second round PCR by a Sanger approach using the BigDye™ Terminator v3.1 Cycle Sequencing Kit.

The 5′RACE was performed using the FirstChoice RLM-RACE kit (Ambion) and following manufacturer's instructions. After cDNA synthesis, two rounds of PCR were performed using the primer pairs outer primer/R1-trhU and inner primer/R2-trhU. The outer and inner primers are supplied by the manufacturer. The amplicons generated were purified and sequenced.

To determine cotranscription among the different transcripts generated from the AN operon, cDNA was obtained from total RNA samples using primers R1-0009, R1-trhP, and R1-trhU and the reverse transcriptase AMV RT (Promega). PCR amplification using the following primer pairs was used to detect cotranscription: F1-htdK/R1-0009, F1-0009/R1-trhP, and F1-trhW/R1-trhU. As control of DNA contamination, not retrotranscribed samples were used.



	R27 genes	*tra* genes	non-*tra* genes
FC > +2	30 (14.6 %)	14 (40 %)	16 (9.4 %)
FC < -2	34 (16.6 %)	1 (2.8 %)	33 (19.4 %)
-2 < FC < +2	111 (54.1 %)	19 (54.3 %)	92 (54.1 %)
non-expressed	30 (14.6 %)	1 (2.8 %)	29 (17.1 %)
Total	205	35	170

FIGURE 1 | Expression profile of the R27 genes at permissive and non-permissive temperatures. **(A)** Conjugation rates were calculated using cultures of both donor [AAG1(R27)] and recipient (MG1655) strains grown at either 25 or 37°C. Average values and standard deviation of three independent experiments is shown. **(B)** Summary of the genes with an altered transcriptional expression in cultures of the strain AAG1(R27) grown at either 25 or 37°C. R27 genes are also divided in *tra* and non-*tra* genes. Attending to the fold change (FC), genes are classified in derepressed at 25°C (FC> +2), repressed at 25°C (FC<-2), no thermoregulated (-2<FC< +2) and genes with signal values lower than 100 fluorescence units in both culture conditions were arbitrarily considered as non-expressed. **(C)** Fold change expression of the 205 ORFs encoded in R27 plasmid between AAG1(R27) cells grown at either 25 or 37°C. Expressed genes were classified as *tra* genes (black bars) and non-*tra* genes (gray bars). Non-expressed genes are indicated as white bars.

TABLE 1 | Expression level of the R27 genes with altered expression at either 25 or 37°C in cultures grown up to mid-logarithmic phase.

		Gene	25°C	37°C	FC	Gene name and predicted function
tra genes	FC>+2	R0007	1172	309	3.79	*trhW* (AN), pilus assembly
		R0008	299	101	2.97	*trhP* (AN), peptidase
		R0025	322	128	2.52	*trhV* (AC), lipoprotein
		R0027	528	176	2.99	*htdT* (AC)
		R0029	664	185	3.59	*trhB* (AC), Mpf formation
		R0030	558	138	4.04	*htdO* (AC)
		R0031	465	102	4.56	*trhK* (AC), pilus assembly
		R0032	1337	201	6.64	*trhE* (AC), pilus assembly
		R0033	1424	70	20.44	*trhL* (AC), pilus assembly
		R0034	1536	107	14.35	*trhA* (AC), pilus major subunit
		R0121	1207	145	8.32	(H)
		R0122	511	41	12.45	*traH* (H), relaxosome protein
		R0123	1669	510	3.27	*trhR* (R), regulator
		R0126	129	20	6.47	*trhF* (F), pilus assembly
	FC<-2	R0005	190	393	−2.06	*trhN* (AN), Mpf complex stability
non-*tra* genes	FC> +2	R0001	1306	147	8.91	
		R0002	172	19	9.18	
		R0042	6132	2278	2.69	partition protein
		R0129	169	45	3.76	
		R0130	183	59	3.13	*bfpH,* trbN-like protein
		R0135	1433	425	3.37	dsbA, outer membrane protein
		R0153	692	242	2.86	partition protein
		R0154	1714	362	4.73	cythosine methylase
		R0200	129	36	3.58	
		R0204	403	77	5.25	
		R0205	352	40	8.72	
		R0206	197	26	7.74	
		R0207	921	86	10.71	
		R0208	268	13	20.23	
		R0209	310	18	17.00	
		R0210	1201	52	22.90	
	FC<-2	R0023	62	181	−2.92	
		R0036	585	1422	−2.43	*repHIA*, replication protein
		R0044	42	124	−2.95	
		R0048	562	1190	−2.12	
		R0052	154	359	−2.33	
		R0053	115	282	−2.46	
		R0058	176	395	−2.25	
		R0067	98	218	−2.23	
		R0069	47	118	−2.54	
		R0070	558	1671	−3.00	
		R0075	67	177	−2.65	
		R0083	231	914	−3.95	*tetC*, transcriptional regulator
		R0094	1306	3893	−2.98	*insB1,* IS1 transposase
		R0095	6229	12567	−2.02	*insA,* IS1 transposase
		R0097	191	397	−2.08	*yigE*
		R0139	233	475	−2.03	
		R0140	119	378	−3.18	
		R0142	1126	5777	−5.13	
		R0143	372	779	−2.09	
		R0148	428	993	−2.32	Transposase
		R0149	58	224	−3.89	
		R0151	189	401	−2.12	
		R0159	725	1932	−2.66	
		R0160	122	325	−2.66	
		R0162	298	637	−2.13	
		R0166	259	867	−3.34	
		R0175	399	893	−2.24	
		R0176	154	807	−5.24	
		R0177	2327	6441	−2.77	*insB,* transposase
		R0180	3294	9075	−2.76	*insB,* transposase
		R0182	117	289	−2.47	*hha*, regulator
		R0183	1567	3694	−2.36	
		R0195	1684	5041	−2.99	*insD,* transposase

The fold change of the fluorescence arbitrary units and the predicted function is indicated. The genes are divided in tra and non-tra genes. Attending to the fold change (FC), genes are classified in derepressed at 25°C (FC> +2) and repressed at 25°C (FC<-2).

RESULTS AND DISCUSSION

Transcription Profile of the R27 Plasmid at Permissive and Non-permissive Temperatures

The conjugation of the R27 plasmid is tightly regulated by temperature (**Figure 1A**) (Taylor and Levine, 1980; Forns et al., 2005). Moreover, R27 transfer is promoted at temperatures lower than 30°C whereas is strongly repressed at 37°C. In order to characterize the effect of temperature on the expression of the R27 genes, the transcriptional expression of the 205 genes was monitored by using specific microarrays.

RNA samples extracted from AAG1 (a Δ*lac* derivative of MG1655) carrying the R27 plasmid were analyzed. Cultures were grown in LB at permissive (25°C) and non-permissive (37°C) temperatures. Previous studies showed that R27 conjugation occurs more efficiently in cells growing exponentially as compared with cells in stationary phase of growth (Gibert et al., 2016). For this reason, cultures were grown up to mid logarithmic phase (OD_{600nm} 0.4). Only expression values higher than 100 (arbitrary units of intensity of fluorescence), in at least one of the two conditions compared, were considered as significant expression and thus, the fold-change expression between conditions was calculated. We arbitrarily defined the fold-change threshold for the differential expression of a gene to 2-fold or higher. Therefore, genes with a fold-change $\geq +2$ are induced at 25°C whereas genes with a fold-change ≤ -2

are induced at 37°C. A summary of the number of genes with a temperature-dependent expression (**Figure 1B**) and the fold change of all R27 genes (**Figure 1C** and **Table S3**) are shown. Overall, 64 genes have the expression altered by the temperature under the experimental conditions used, representing a 31.2% over the total number of R27 genes. Amongst these genes, approximately half (30) are induced at low temperature whereas the other half (34) are repressed at 25°C. No expression was detected in 30 genes, representing the 14.6% of all R27 genes. When discriminating between *tra* and non-*tra* genes, the distribution of affected genes showed different patterns. Up to 42.8% of the *tra* genes are affected by temperature. Most of them, 14 of the 15 genes affected, were stimulated at 25°C consistent with the clear increase in the conjugation ratio detected at this temperature. On the other hand, only 28.8% of the non-*tra* genes were affected, being most of them (33 of 49) repressed at 25°C. In **Table 1**, all genes with altered expression at low temperature are detailed. Genes from all *tra* operons, except for the Z operon coding for the entry-exclusion system, were found significantly induced at 25°C. Consistent with the increased conjugation frequency detected at low temperature, several genes from the AC operon, coding mostly for proteins required for mating pair formation, were found among the genes with the greatest induction. For instance, the gene *trhA*, coding for the major subunit of the putative conjugative pilus, is induced more than 14-fold. The expression of the R operon, coding for the regulators TrhR and TrhY, that are required for activation of *tra* operons expression, was induced at 25°C (3.2 and 1.9-fold for *trhR* and *trhY*, respectively) consistent with previous transcriptional studies (Gibert et al., 2014). Most of the non-*tra* R27 genes have not been characterized and their predicted function has been assigned attending to protein homology studies. Among the non-*tra* genes induced at 25°C, some of them might contribute to promote plasmid transfer, such as the muramidase (R0130) and a protein involved in the turnover of disulphide bonds (R0135) (Elton et al., 2005; Zahrl et al., 2005). Interestingly, all those non-*tra* genes were also found to be induced under physiological conditions promoting R27 conjugation (Gibert et al., 2016). A cluster of 9 genes, from ORF R0204 to ORF R0002, are importantly induced at low temperature. The function of the predicted proteins is not known.

A

htdA/WT 25ºC

	R27 genes	*tra* genes	non-*tra* genes
F > +2	139 (67.8 %)	33 (94.3 %)	106 (62.3 %)
F < -2	1 (0.5 %)	0	1 (0.6 %)
-2 < F < +2	41 (20 %)	2 (5.7 %)	39 (23 %)
non-expressed	24 (11.7 %)	0	24 (14.1 %)
Total	205	35	170

B

htdA/WT 37ºC

	R27 genes	*tra* genes	non-*tra* genes
F > +2	11 (5.4 %)	4 (11.4 %)	7 (4.1 %)
F < -2	43 (21 %)	6 (17.1 %)	37 (21.8 %)
-2 < F < +2	107 (52.2 %)	21 (60 %)	86 (50.6 %)
non-expressed	44 (21.4 %)	4 (11.4 %)	40 (21.4 %)
Total	205	35	170

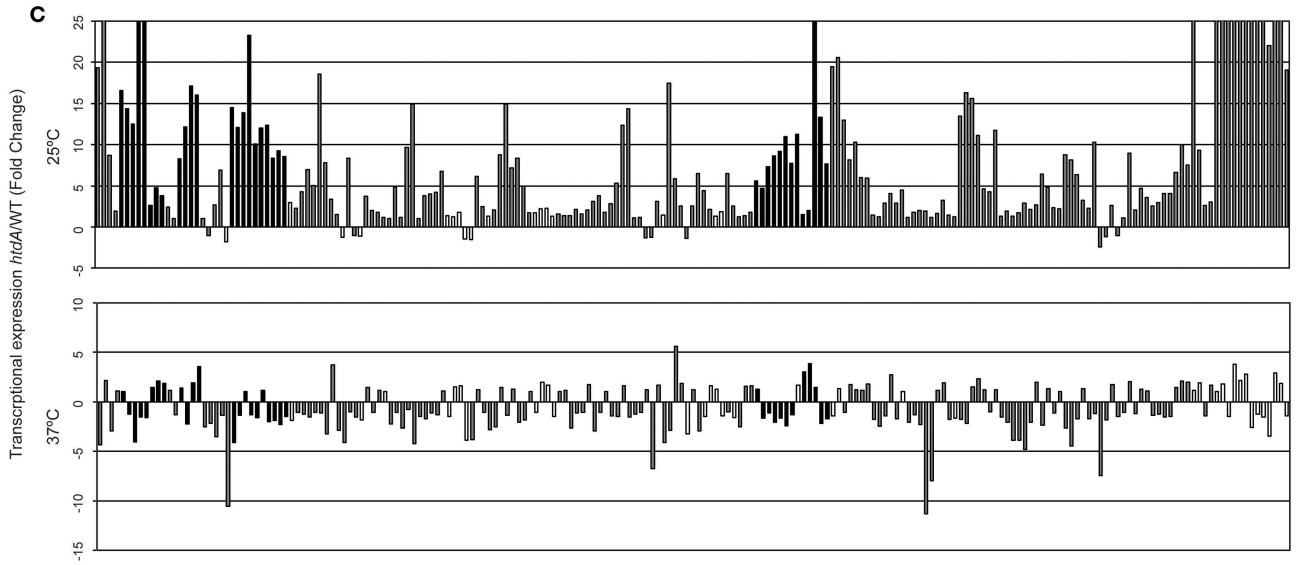


FIGURE 2 | HtdA represses expression of R27 genes primarily at 25°C. Summary of the genes showing an altered transcriptional expression in cultures of the strains AAG1(drR27) and AAG1(R27) grown at either 25°C **(A)** or 37°C **(B)**. R27 genes are also divided in *tra* and non-*tra* genes. Attending to the fold change (FC), genes are classified in derepressed at (FC> +2), repressed (FC<-2), no thermoregulated (-2<FC< +2) and genes with signal values lower than 100 fluorescence units in both culture conditions were arbitrarily considered as non-expressed. **(C)** Fold change expression of the 205 ORFs encoded in R27 plasmid between AAG1(drR27) cells and AAG1(R27) cells grown at either 25°C (upper panel) or 37°C (lower panel). Expressed genes were classified as *tra* genes (black bars) and non-*tra* genes (gray bars). Non-expressed genes are indicated as white bars.

Regarding the genes overexpressed at 37°C, only one belongs to the *tra* group of genes, *trhN*, coding for a predicted protein involved in mating pair formation complex stability (Lawley et al., 2003). The *trhN* gene was over 2-fold overexpressed at 37°C. Only a few genes with a higher expression at non-permissive temperature encoded predicted proteins with homologies to an annotated gene (**Table 1**). We found the *repHIA* gene (R0036), involved in plasmid replication; *tetC* (R0083), the tetracycline repressor; *insA* (R0095) and three *insB* genes (R0094, R0177 and R0180) that encode putative transposases of the insertion element IS1; the ORF R0195, encoding a IS2 transposase; and another putative transposase, R0148. Interestingly, the *hha* gene, coding for a transcriptional regulator that acts in combination with H-NS, is more expressed at 37°C, in concordance with its repressor role in the regulation of the conjugation process at non-permissive temperature described for H-NS/Hha proteins (Forns et al., 2005).

Altogether, our data clearly indicate that at 25°C there is a higher expression of the *tra* genes, consistent with the promoted R27 conjugation at low temperatures, whereas at 37°C the expression of mobile elements and transposase genes is induced. The results at high temperature might suggest that within the host (37°C), the mobilization of IS elements is induced promoting the transfer of material from the plasmid to the chromosome.

Evidently, many R27 genes were not thermoregulated. Some of R27 genes were highly expressed (higher than 600 arbitrary units of intensity of fluorescence) at both temperatures. These results could be expected, since many of these genes are involved in global processes that may take place at any temperature. Among these genes we found the tetracycline operon (*tetR, tetA,* and *tetD)*, genes involved in partitioning (*parA*/R0020 and *parB*/R0019), replication (*repHIB*) or transposition (R0046, R0076, and R0085).

HtdA Causes Repression of R27 Genes Expression Only at Permissive Temperature

HtdA is a R27 encoded regulator that represses conjugation since R27 plasmid transfer is strongly promoted in an *htdA* mutant (Whelan et al., 1994; Gibert et al., 2013). HtdA is involved in the repression of the transcriptional regulation of four *tra* operons (F, Z, AC, and H) by counteracting the activation mediated by TrhR/TrhY (Gibert et al., 2014). Moreover, variations in the cellular levels of HtdA seem to play a crucial role in the growth phase dependency of R27 conjugation (Gibert et al., 2016). Microarray analyses using a strain carrying a drR27 plasmid (*htdA* mutant) were performed to monitor the effect of HtdA in the transcriptional expression of the R27 genes at both permissive and non-permissive temperatures (**Figure 2**, **Table S4**). At permissive temperature (**Figures 2A,C**), the *htdA* mutation causes a generalized derepression of R27 plasmid genes. Under the applied experimental conditions, 139 out of 181 expressed genes were upregulated in the *htdA* mutant strain whereas only one was downregulated (R0176). The downregulated gene is a non-*tra* gene with unpredicted function. Among the *tra* genes, 33 out of the 35 genes are upregulated in the *htdA* mutant. Only the R operon, containing the *trhR* and *trhY* genes, is not upregulated more than two-fold. This is consistent with previous transcriptional data from a *lacZ* fusion with the R operon promoter, indicating that the R operon was not regulated by the HtdA protein (Gibert et al., 2013). Remarkably, TrhR and TrhY are defined as activators of the *tra* genes whose activity is counteracted at the protein level by HtdA (Gibert et al., 2014). The fact that most non-*tra* genes (106 of 170) are derepressed in an *htdA* mutant strain clearly indicates that HtdA, initially described as a regulator of the *tra* genes, controls directly or indirectly the expression of most R27 genes. Interestingly, among the genes derepressed in the *htdA* mutant, we found *hns* and *hha* which are also involved in the temperature regulation of R27 conjugation (Forns et al., 2005). An earlier study showed that H-NS silences *trhR* transcriptional expression at non-permissive temperature whereas no relevant repression was detected at 25°C (Gibert et al., 2014). The HtdA-mediated repression of H-NS/Hha at 25°C might contribute to promote expression of *trhR* and *trhY* at low temperature.

Interestingly, among the non-*tra* genes induced in the *htdA* mutant we found genes *bfph*, *dsbA* and two partition genes (R0042 and R00153), which have been found induced in R27 (htdA+) under all conditions known to promote R27 conjugation such as low temperature (**Table S3**) and log phase (Gibert et al., 2016).

Remarkably, all the 39 non-*tra* genes HtdA-independent were not induced at permissive temperature as compared to 37°C in the R27 plasmid ($htdA^+$) (**Tables S3**, **S4**). Among these genes,

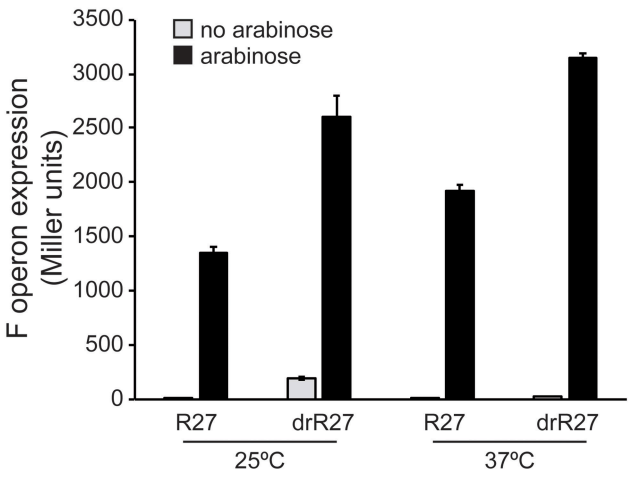
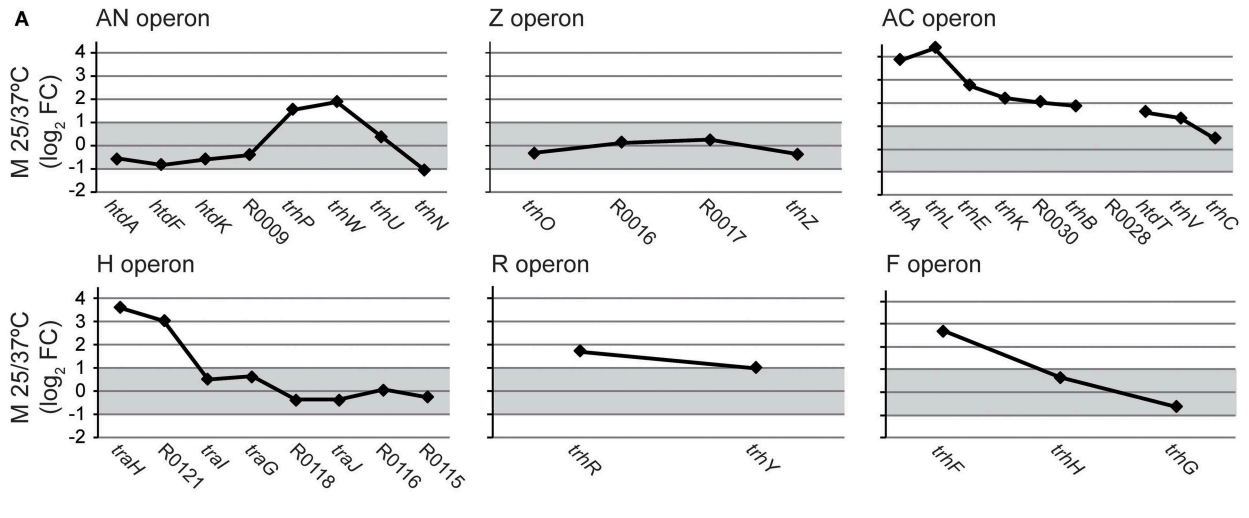
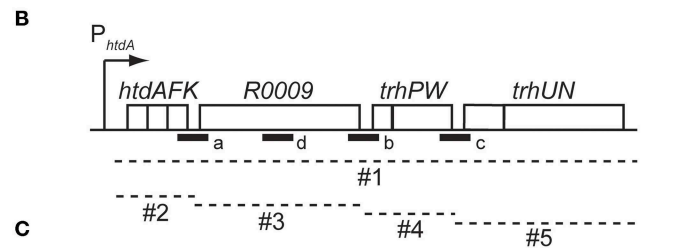


FIGURE 3 | TrhR and TrhY overexpression causes induction of the F operon even at non-permissive temperature. Transcriptional expression of the F operon of R27 was monitored in cultures of the strain AAG1-F, carrying a chromosomal *lacZ* fusion with the promoter sequence of the F operon. The effect of the pBAD*trhRY* plasmid in the presence of either R27 or its *htdA* derivative drR27 was assessed. Cultures were grown at either 25 or 37°C to mid-logarithmic phase in the absence (gray bars) or presence (black bars) of arabinose (0.02%). The β-galactosidase activity (Miller units) was determined in three independent cultures and mean values with standard deviations are plotted.

8 are involved in transposition (R0046, R0076, R0085, R0094, R0095, R0177, R0180, and R0181). We also found 2 partition genes (*parA*/R0020 and *parB*/R0019), 2 genes involved in citrate transport (*citA* and *citB*), part of the Tc operon (*tetR*, *tetA* and *tetC*) and 2 genes involved in UV protection (*mucA* and *mucB*) (**Tables S3**, **S4**). All these genes are involved in processes that presumably are independent of the conjugation process of the R27 plasmid.

The effect of the *htdA* mutation at non-permissive temperature was tested (**Figures 2B,C**) and a completely different pattern was observed. Most of the expressed genes did not show altered expression in the *htdA* mutant strain (107 of 161). Only 11 genes were derepressed by the *htdA* mutation as compared with the 139 genes detected at permissive temperature. On the other hand, at 37°C, 43 genes have lower expression in the *htdA* mutant than in the strain harboring the wt plasmid whereas at 25°C only one gene showed a decrease in the expression. Among the *tra* genes, 4 genes are derepressed in the *htdA* mutant, *htdF* from AN operon, *trhZ* from the Z operon and the 2 genes from the R operon; whereas 6 genes

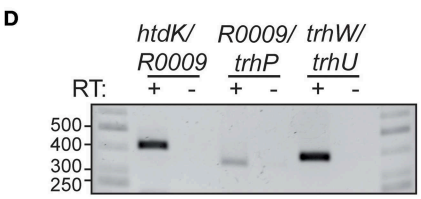




C

	5-end	P-5'	3-end	P-3'
#2	CCCAT**A**AGATC	+ 96	GGTTT**A**TGTTT	- 137
#3	ATTCT**A**ATTCT	+ 107	TTTTT**T**GTTTT	- 48
#4	GAAAG**T**TACTT	+ 190	GCAGA**G**GAGAA	+ 10
#5	AGTGT**T**TTAAC	+ 170	n.d.	
	TTACT**A**AATAG	+ 122		
	AAGAA**G**CTAAA	+ 97		

P-5': Position of the 5'end, using as a reference the first codon of the upstream ORF of the indicated transcript.
P-3': Position of the 3'end, using as a reference the stop codon of the downstream ORF of the indicated transcript.
n.d.: not detected



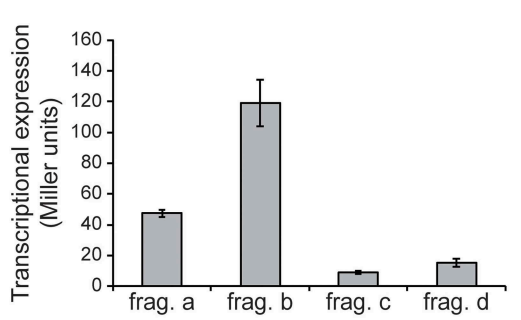


FIGURE 4 | Transcriptional organization of the AN operon. **(A)** M-plots of the temperature dependent expression of the *tra* operons. The M value (log_2FC 25/37°C) for the expressed genes of the different *tra* operons is depicted. An M value between +1 and−1, equivalent to a FC of +2 and−2, is considered no alteration in the expression and is shadowed in gray. **(B)** Representation of the AN operon. Dashed lines indicate the different putative transcripts (#1 to #5) derived from the AN operon. A thick line indicates the fragments (a to d) cloned in pRS551 to construct the *lacZ* fusions used. **(C)** The positions of 5′- and 3′ ends of the transcripts #2, #3 and #4, labeled in bold, were determined by circRNA. The putative 5′ ends of transcript #5 identified by 5′RACE assays. **(D)** RT-PCR using primers within the ORFs flanking the intergenic regions indicated. In all cases RNA not-RT AMV-treated was used as negative control. **(E)** Transcriptional expression of the *lacZ* fusions with the indicated fragments of AN operon was determined in cultures grown at 25°C to mid-logarithmic phase. The β-galactosidase activity (Miller units) was determined in three independent cultures and mean values with standard deviations are plotted.

are repressed *trhH* (F operon), *traI* and R0118 (H operon), *trhW* (AN operon), R0016 (Z operon) and *trhC* (AC operon). Among the non-*tra* genes, only 7 genes were depressed in the *htdA* mutant, all of them were also repressed by HtdA at permissive temperature.

Overall our data suggest that the repressor role of HtdA at permissive temperature is not limited to the *tra* genes and its regulatory role vanishes at 37°C.

Overexpression of TrhR/TrhY Induces Expression of the F Operon Even at Non-permissive Temperature

Our data demonstrate a powerful repressor role of HtdA on the expression of R27 genes, especially among the *tra* genes, which is temperature dependent. Having in consideration that HtdA acts by counteracting the activation mediated by TrhR/TrhY, the role of the activators in the temperature mediated regulation was studied. F operon expression has been systematically used as a reporter to study HtdA-mediated regulation of the *tra* genes (Gibert et al., 2013, 2014, 2016). F operon is repressed by HtdA at permissive temperature but not at non-permissive temperature (**Table S4**). The effect of overexpressing TrhR/TrhY on F operon expression was monitored at permissive and non-permissive temperature in both $htdA^+$ and $htdA^-$ genetic backgrounds. Overexpression of trhR/trhY using the arabinose inducible promoter from pBAD18 causes a robust upregulation of F operon independently of the temperature (**Figure 3**). These results suggest that TrhR/TrhY activity is crucial for the temperature-dependent regulation of R27 conjugation and provides a rational for the mechanisms of HtdA in the thermoregulation.

New Insights in the Transcriptional Expression of the an Operon

In **Figure 4A**, the expression pattern of the six *tra* operons is depicted using the M value ($\log_2$FC) between 25 and 37°C of the different genes. Genes with no altered expression are defined by having a M value between +1 and−1, equivalent to$-2 >$ FC $> +2$. A closer look to the gene expression pattern of the *tra* operon in response to the temperature reveals three different patterns. Pattern 1, unresponsiveness, as shown by the Z operon, where the expression of the three genes is not altered by temperature. Pattern 2, operons with induced expression at 25°C and the temperature responsiveness is greater among the proximal genes than among the distal ones. This is the most common pattern, shared by the operons AC, H, R, and F. The decrease in the expression of downstream genes in the same operon is defined as transcriptional polarity and it is a common feature among polycistronic operons. Pattern 3, the unusual pattern shown by the AN operon. The 4 proximal genes (*htdA*, *htdF*, *htdK* and R0009) do not respond to temperature, the 4 following genes show a pattern similar to those described in pattern 2. *trhP* and *trhW*, are induced at 25°C (2.9 and 3.7-fold) and the 2 distal genes (*trhU* and *trhN*) are again unresponsive to temperature. This expression profile within a polycistronic operon suggests the presence of complex regulatory mechanisms. Previous studies on the transcriptional organization of the *tra* genes revealed that the AN operon contains 8 genes, spanning from *htdA* to *trhN* which were apparently cotranscribed (Alonso et al., 2005). However, those studies did not rule out the occurrence of several transcripts arising from different events such as partial termination, mRNA processing and/or the possible presence of internal promoters. Interestingly, we found 3′ intergenic regions in the AN operon, located between *htdK* and R0009 (197 bp), R0009 and *trhP* (233 bp), and *trhW* and *trhU* (249 bp). Promoter search with the BPROM software (Solovyev and Salamov, 2011) predicts two putative promoters located upstream of the R009 gene and the *trhP* gene. Transcriptional terminators searches, using the ARNold software (Naville et al., 2011) identify a putative Rho-independent terminator downstream of the *htdK* gene. To characterize the transcripts generated from the AN operon, total RNA was isolated from cultures of strain AAG1(R27) grown at 25°C. Circularized RNA was obtained from either the isolated RNA (monophosphated mRNA, processed transcripts) or after 5′polyphosphatase RNA treatment (originally triphosphated, real transcriptional starts). The circRNA samples were used as template to obtain cDNA using specific primers (**Table S2**). Sequencing of the cDNA allowed us to clearly identify three transcripts, labeled as #2, #3 and #4 in **Figures 4B,C**. The fact that these three transcripts have been detected provides a feasible explanation to the particular expression profile of the AN operon independently of the mechanisms involved in its generation. Interestingly, transcript #2 (*htdA*-*htdK*) was the only transcript detected after polyphosphatase treatment, indicating a triphosphated mRNA and thus, a transcript with a real transcriptional start. Transcript #3 was detected from RNA untreated (processed transcript). Remarkably, its 5′ end overlaps with the 3′ end of transcript #2, by 46 bp. A possible explanation of this result is that transcript #3 was generated from a secondary promoter different from the promoter upstream of *htdA* and once is generated this transcript would be further processed to the final transcript #3 detected. Transcript #4 has a 5′ end located downstream of the 3′ end of transcript #3, therefore it could be generated from processing of a pre-existing polycistronic operon. After several trials, a transcript containing the sequences of the distal genes *trhU-trhN* was not identified by circRNA. However, 5′RACE experiments identified several transcripts *trhU-trhN* with distinctive 5′ ends (transcript #5). Further studies will be required to identify the exact location of the 3′ end of this transcript.

RT-PCR assays (**Figure 4D**) using primers from the internal sequences of the transcripts #2, #3, #4, and #5 let us detect cotranscription among the different transcripts as previously described in Alonso et al. (2005) suggesting the presence of a polycistronic transcript, named #1.

The presence of putative promoter sequences in the intergenic regions (IR) upstream of R0009, *trhP* and *trhU* was assessed. The three IR and a R0009 intragenic region used as a negative control (named a, b, c, and d in **Figure 4B**) were cloned in pRS551 vector, upstream of the promoterless *lacZ* gene. The *lacZ* fusion constructs generated were transferred to the *attB* site present in the *E. coli* chromosome. Transcriptional expression was measured by a β-galactosidase assay. The results (**Figure 4E**)

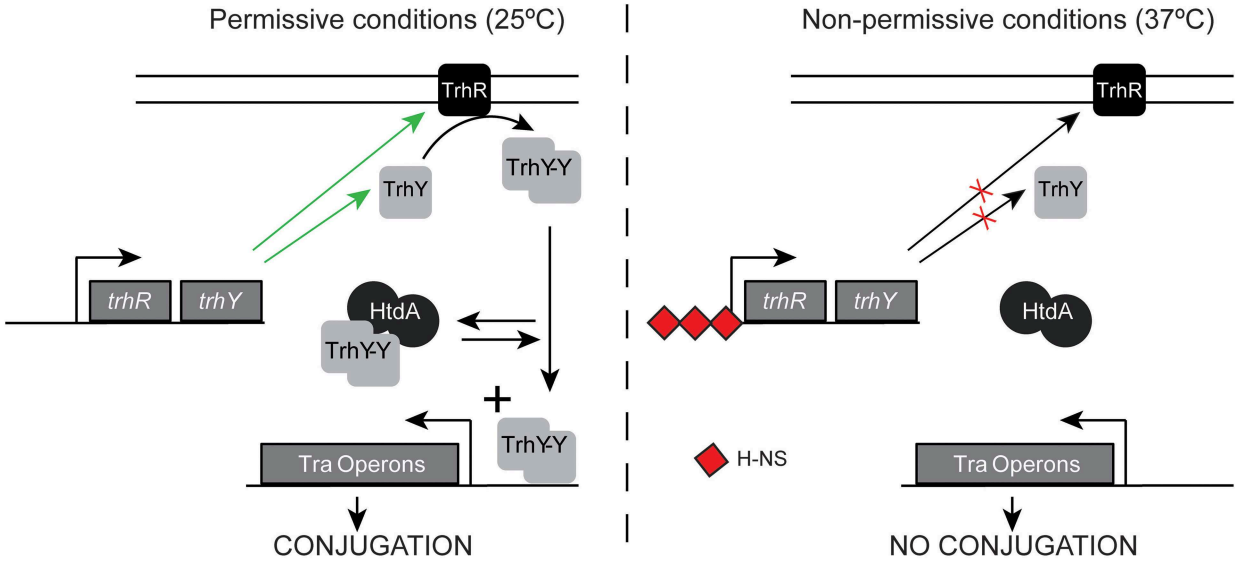


FIGURE 5 | Schematic model of regulation of the *tra* operons at permissive and non-permissive temperatures. Green arrows indicate expression is promoted, whereas red crosses indicate expression is repressed.

suggest the presence of putative active promoter sequences upstream the R0009 and *trhP* genes. No activity was detected in the third IR between *trhW* and *trhU*.

Further extensive experiments will be required to fully describe the transcripts generated from the AN operon and the mechanisms involved in its generation. We would like to highlight that our data identify different mRNA species that provide a feasible explanation to the differential expression from the AN operon genes as revealed by the microarray data. The results obtained suggest that different transcriptional and posttranscriptional events may participate in the control of AN operon expression, such as partial termination, processing and transcription from secondary promoters. As reported by Conway et al. (2014), a large proportion (36%) of operons in *E. coli* are complex, with internal promoters or terminators that generate multiple transcription units. For 43% of operons, differential expression of polycistronic genes was observed, despite being in the same operons, indicating that *E. coli* operon architecture allows fine-tuning of gene expression.

CONCLUDING REMARKS

The transcriptional expression of R27 genes involved in plasmid transfer was induced at 25°C, consistent with R27 conjugation being optimal at this temperature. R27 conjugation is modulated by a regulator circuit composed by HtdA and TrhR/TrhY. Our data indicate that HtdA is not thermoregulated, while TrhR and TrhY are slightly more expressed at 25°C. Moreover, TrhR and TrhY expression is not regulated by HtdA at permissive temperature. These two proteins are responsible, at least partially, of the induction of the *tra* genes observed under these conditions (**Figure 3**).

A common expression profile shared among many *tra* and non-*tra* genes is exemplified by the *trhA* gene, coding for the pilin. A dramatic upregulation at 25°C in the absence of HtdA, but similar expression either in the presence or in the absence of the repressor at 37°C was observed (**Table S4**). This profile suggests that HtdA represses primarily at low temperature whereas at non-permissive temperature the presence or the absence of some specific regulators might avoid upregulation even in the absence of HtdA. HtdA acts by counteracting the activity of TrhR/TrhY and it has previously been described that H-NS silences *tra* genes expression at 37°C, by repressing *trhRY* expression (**Figure 5**). Having that in mind, we suggest that HtdA can exert its regulatory function only when H-NS is not blocking the expression of TrhR/TrhY. The TrhR/TrhY-HtdA mediated regulation is not restricted to the *tra* operon genes since many other R27 genes were also thermoregulated in a HtdA-dependent manner. These results also suggest that the effect of the H-NS/Hha proteins is more general as regulators of the *tra* operons than previously reported.

DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in the ArrayExpress repository, under the accession number E-MTAB-9150 (http://www.ebi.ac.uk/arrayexpress).

AUTHOR CONTRIBUTIONS

MG contributed in the investigation. SP contributed in the investigation and writing the manuscript. CM and CB contributed in the conceptualization, investigation, formal

analysis, and writing the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Spanish Ministry of Economy and Competitiveness (grant AGL2013-45339R) Spanish Ministry of Science, Innovation and Universities (grant PGC2018-096958-B-I00), and the Catalonian government (grant 2017SGR499). MG was recipient of a FPU grant from the Spanish Ministry of Science and Innovation.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2020.00124/full#supplementary-material

REFERENCES

Aberg, A., Shingler, V., and Balsalobre, C. (2008). Regulation of the fimB promoter: a case of differential regulation by ppGpp and DksA *in vivo*. *Mol. Microbiol.* 67, 1223–1241. doi: 10.1111/j.1365-2958.2008.06115.x

Alonso, G., Baptista, K., Ngo, T., and Taylor, D. E. (2005). Transcriptional organization of the temperature-sensitive transfer system from the IncHI1 plasmid R27. *Microbiology* 151, 3563–3573. doi: 10.1099/mic.0.28256-0

Bennett, P. M. (2008). Plasmid encoded antibiotic resistance: acquisition and transfer of antibiotic resistance. *Br. J. Pharmacol.* 153, S347–357. doi: 10.1038/sj.bjp.0707607

Conway, T., Creecy, J. P., Maddox, S. M., Grissom, J. E., Conkle, T. L., Shadid, T. M., et al. (2014). Unprecedented high-resolution view of bacterial operon architecture revealed by RNA sequencing. *MBio* 5:e01442-14. doi: 10.1128/mBio.01442-14

Elton, T. C., Holland, S. J., Frost, L. S., and Hazes, B. (2005). F-like type IV secretion systems encode proteins with thioredoxin folds that are putative DsbC homologues. *J. Bacteriol.* 187, 8267–8277. doi: 10.1128/JB.187.24.8267-8277.2005

Forns, N., Baños, R. C., Balsalobre, C., Juárez, A., and Madrid, C. (2005). Temperature-dependent conjugative transfer of R27: role of chromosome- and plasmid-encoded Hha and H-NS proteins. *J. Bacteriol.* 187, 3950–3959. doi: 10.1128/JB.187.12.3950-3959.2005

Gibert, M., Juárez, A., Madrid, C., and Balsalobre, C. (2013). New insights in the role of HtdA in the regulation of R27 conjugation. *Plasmid* 70, 61–68. doi: 10.1016/j.plasmid.2013.01.009

Gibert, M., Juárez, A., Zechner, E. L., Madrid, C., and Balsalobre, C. (2014). TrhR, TrhY and HtdA, a novel regulatory circuit that modulates conjugation of the IncHI plasmids. *Mol. Microbiol.* 94, 1146–1161. doi: 10.1111/mmi.12823

Gibert, M., Paytubi, S., Beltrán, S., Juárez, A., Balsalobre, C., and Madrid, C. (2016). Growth phase-dependent control of R27 conjugation is mediated by the interplay between the plasmid-encoded regulatory circuit TrhR/TrhY-HtdA and the cAMP regulon. *Environ. Microbiol.* 18, 5277–5287. doi: 10.1111/1462-2920.13579

Grindley, N. D., Grindley, J. N., and Anderson, E. S. (1972). R factor compatibility groups. *Mol. Gen. Genet.* 119, 287–297. doi: 10.1007/BF00272087

Guyer, M. S., Reed, R. R., Steitz, J. A., and Low, K. B. (1981). Identification of a sex-factor-affinity site in *E. coli* as gamma delta. *Cold Spring Harb. Symp. Quant. Biol.* 45(Pt 1), 135–140. doi: 10.1101/SQB.1981.045.01.022

Holt, K. E., Phan, M. D., Baker, S., Duy, P. T., Nga, T. V. T., Nair, S., et al. (2011). Emergence of a globally dominant IncHI1 plasmid type associated with multiple drug resistant typhoid. *PLoS Negl. Trop. Dis.* 5:e1245. doi: 10.1371/journal.pntd.0001245

Lawley, T. D., Gilmour, M. W., Gunton, J. E., Standeven, L. J., and Taylor, D. E. (2002). Functional and mutational analysis of conjugative transfer region 1 (Tra1) from the IncHI1 plasmid R27. *J. Bacteriol.* 184, 2173–2180. doi: 10.1128/JB.184.8.2173-2180.2002

Lawley, T. D., Gilmour, M. W., Gunton, J. E., Tracz, D. M., and Taylor, D. E. (2003). Functional and mutational analysis of conjugative transfer region 2 (Tra2) from the IncHI1 plasmid R27. *J. Bacteriol.* 185, 581–591. doi: 10.1128/JB.185.2.581-591.2003

Maher, D., and Taylor, D. E. (1993). Host range and transfer efficiency of incompatibility group HI plasmids. *Can. J. Microbiol.* 39, 581–587. doi: 10.1139/m93-084

Miller, J. H. (1992). *A Short Course in Bacterial Genetics : A Laboratory Manual and Handbook for Escherichia coli and Related Bacteria*. New York, NY: Cold Spring Harbor Laboratory Press.

Naville, M., Ghuillot-Gaudeffroy, A., Marchais, A., and Gautheret, D. (2011). ARNold: a web tool for the prediction of Rho-independent transcription terminators. *RNA Biol.* 8, 11–13. doi: 10.4161/rna.8.1.13346

Paytubi, S., Aznar, S., Madrid, C., Balsalobre, C., Dillon, S. C., Dorman, C. J., et al. (2014). A novel role for antibiotic resistance plasmids in facilitating Salmonella adaptation to non-host environments. *Environ. Microbiol.* 16, 950–962. doi: 10.1111/1462-2920.12244

Rooker, M. M., Sherburne, C., Lawley, T. D., and Taylor, D. E. (1999). Characterization of the Tra2 region of the IncHI1 plasmid R27. *Plasmid* 41, 226–239. doi: 10.1006/plas.1999.1396

Sherburne, C. K., Lawley, T. D., Gilmour, M. W., Blattner, F. R., Burland, V., Grotbeck, E., et al. (2000). The complete DNA sequence and analysis of R27, a large IncHI plasmid from *Salmonella typhi* that is temperature sensitive for transfer. *Nucleic Acids Res.* 28, 2177–2186. doi: 10.1093/nar/28.10.2177

Simons, R. W., Houman, F., and Kleckner, N. (1987). Improved single and multicopy lac-based cloning vectors for protein and operon fusions. *Gene* 53, 85–96. doi: 10.1016/0378-1119(87)90095-3

Smillie, C., Garcillán-Barcia, M. P., Francia, M. V., Rocha, E. P. C., and de la Cruz, F. (2010). Mobility of plasmids. *Microbiol. Mol. Biol. Rev.* 74, 434–452. doi: 10.1128/MMBR.00020-10

Solovyev, V., and Salamov, A. (2011). "Automatic annotation of microbial genomes and metagenomics sequences," in *Metagenomics and its Applications in Agriculture, Biomedicine and Environmental Studies*, ed R. W. Li (New York, NY: Nova Science Publishers), 61–78.

Taylor, D. E., and Levine, J. G. (1980). Studies of temperature-sensitive transfer and maintenance of H incompatibility group plasmids. *J. Gen. Microbiol.* 116, 475–484. doi: 10.1099/00221287-116-2-475

Whelan, K. F., Maher, D., Colleran, E., and Taylor, D. E. (1994). Genetic and nucleotide sequence analysis of the gene htdA, which regulates conjugal transfer of IncHI plasmids. *J. Bacteriol.* 176, 2242–2251. doi: 10.1128/JB.176.8.2242-2251.1994

Zahrl, D., Wagner, M., Bischof, K., Bayer, M., Zavecz, B., Beranek, A., et al. (2005). Peptidoglycan degradation by specialized lytic transglycosylases associated with type III and type IV secretion systems. *Microbiology* 151, 3455–3467. doi: 10.1099/mic.0.28141-0

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Gibert, Paytubi, Madrid and Balsalobre. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Molecular Biosciences

REVIEW
published: 16 July 2020
doi: 10.3389/fmolb.2020.00158

Role of Bacterial Cytoskeleton and Other Apparatuses in Cell Communication

*Divya Singhi and Preeti Srivastava**

Department of Biochemical Engineering and Biotechnology, Indian Institute of Technology Delhi, New Delhi, India

The bacterial cytoskeleton is crucial for sensing the external environment and plays a major role in cell to cell communication. There are several other apparatuses such as conjugation tubes, membrane vesicles, and nanotubes used by bacterial cells for communication. The present review article describes the various bacterial cytoskeletal proteins and other apparatuses, the physical structures they form and their role in sensing environmental stress. The implications of this cellular communication in pathogenicity are discussed.

Keywords: MreB, membrane vesicles, curli, nanotube, tubular spinae 3

OPEN ACCESS

Edited by:
Tatiana Venkova,
Fox Chase Cancer Center,
United States

Reviewed by:
Grzegorz Wegrzyn,
University of Gdańsk, Poland
Miguel Angel Cevallos,
National Autonomous University of Mexico, Mexico

***Correspondence:**
Preeti Srivastava
preeti@dbeb.iitd.ac.in;
preetisrivastava@hotmail.com

Specialty section:
This article was submitted to Molecular Recognition, a section of the journal Frontiers in Molecular Biosciences

Received: *30 April 2020*
Accepted: *24 June 2020*
Published: *16 July 2020*

Citation:
Singhi D and Srivastava P (2020) Role of Bacterial Cytoskeleton and Other Apparatuses in Cell Communication. Front. Mol. Biosci. 7:158. doi: 10.3389/fmolb.2020.00158

INTRODUCTION

In order to survive in the changing environmental conditions, microorganisms have evolved two broadly described complex communication systems- the contact-independent quorum sensing (QS) and the contact-dependent signaling mechanisms (Bassler and Losick, 2006; Blango and Mulvey, 2009). Cellular communications are usually mediated through synthesis, secretion and detection of signaling molecules commonly known as the inducers, which are released in the environment directly or through various cellular apparatuses (Kaprelyants and Kell, 1996). The mechanism of cell-cell communication that allows bacteria to share information about cell density and adjust gene expression accordingly through chemical signaling is known as QS (Miller and Bassler, 2001; Schertzer and Whiteley, 2011). As a part of a symbiotic relationship, QS was first observed in a bioluminescent bacterium *Vibrio fischeri* which lived on the light producing organ (photophore) of the bobtail squid (Bassler, 2002; Waters and Bassler, 2005). These interactions can be intra and/or inter-species or even inter-kingdoms and allowed co-existence of both the bacterium and its host. These communication systems are also the means through which the pathogens communicate and control their virulence traits. The phenomenon of QS has been extensively described in recent reviews (Papenfort and Bassler, 2016; Abisado et al., 2018; Mukherjee and Bassler, 2019).

The present review describes the role of bacterial cytoskeleton and contact dependent signaling through cellular apparatuses which play a major role in cellular communication. Bacteria have well-defined physical structures and cytoskeleton system which play key role in cellular communication.

As community behavior is highly complex, it can involve various genetic loci encoding extracellular factors and structures that promote surface sensing, cell-to-cell contact and surface colonization. These structures are usually the tubular cellular extensions such as curli, pili, flagella, and fimbriae, etc which help bacteria to communicate with their environment (Van Houdt and Michiels, 2005; Kline et al., 2009). These structures are often assembled with the help of secretion systems of the cell. In bacteria, till date there are nine major groups of bacterial

secretion machinery reported from type I to type IX along with some additional categorized groups such as fimbrial chaperone-usher pathway (CU), Curli pathway, outer membrane vesicle secretion system (OMVSS), etc (Abby et al., 2016; Behzadi, 2019). These systems and pathways are critically important for the survival or death of the bacterial cells because they are involved in wide range of cellular activities including invasion, virulence, pathogenesis, genetic material exchange, immunological and biological interactions, antimicrobial resistance, colonization, and biofilm formation, etc (Abby et al., 2016; Behzadi, 2019). It has been reported that secretion systems assemble various surface structures that interact with both prokaryotic as well as eukaryotic target cells to deliver DNA or protein effectors to modulate cell physiology and growth (Hayes et al., 2010). A schematic representation showing assembly of bacterial surface appendages by the three secretion systems and how these appendages mediate cell-to-cell communication along with the help of cytoskeleton protein is shown in **Figure 1**. Based upon the mechanistic studies of these secretion systems, various bacterial cell surface structures have been identified that play a role in contact-mediated signaling.

Some of the bacterial surface structures that are directly or indirectly involved in cell signaling are shown in **Figure 2** and described in **Table 1**.

SURFACE APPENDAGES IN BACTERIA

Curli Fibers

Curli are the highly aggregated extracellular amyloid fibers expressed by many Gram-negative enteric bacteria such as *E. coli* and *Salmonella* spp. (Chapman et al., 2002). They are the part of bacterial extracellular matrix through which the neighboring cells contact each other and form cell aggregates. With the studies conducted on *E. coli* curli, it was revealed that these fibers are produced either by a specialized secretion pathway commonly known as nucleation-precipitation mechanism or by well-defined Type VIII secretion system. These surface associated fibers are commonly 4–6 nm wide, varying in length up to several micrometers (Barnhart and Chapman, 2006). The structural and assembly components of these long amyloid (protein aggregates) fibers are encoded by seven curli-specific genes (*csg*) present in two different operons *csgBAC* and *csgDEFG*, respectively (Van Gerven et al., 2015). Curli plays an important role in cell aggregation and biofilm development, majorly during the attachment phase. The direct role of curli in causing pathogenesis has not been demonstrated but there are studies which suggest its role in infection process (attachment and invasion). Curli expressing bacterial cells attach better to the host cells as compared to those without it. Their invasion has been shown to result in activation of immune system causing host inflammatory response due to their interaction with host proteins (Barnhart and Chapman, 2006).

Flagella

Being motile is one of the major advantages for the bacteria to adjust in adverse environmental stresses. Also, motility is considered as an important virulence factor of the pathogenic bacteria as it provides cell-to-surface contact required in the initial phase of pathogenicity. The most extensively studied mode of motility in bacteria is flagellum-mediated (Duan et al., 2013). Flagellum (plural-flagella) helps the bacterial cell to move toward the favorable environment. The structure of flagellum is well known consisting of three parts: the basal rotary motor, the hook joint and the helical propeller filament composed majorly of flagellin protein (Khan and Scholey, 2018). On the basis of number of flagella on the cell and their arrangements, motility and virulence of the bacteria dramatically varies. Monotrichous or single polar flagellum confers swimming motility to the bacteria in liquid (in *Vibrio* spp. and *Pseudomonas* spp. during planktonic growth) similar to the propeller on a boat and swarming on surfaces. Amphitrichous or one flagellum on each pole provides darting mobility to the bacteria (e.g., *Campylobacter jejuni*). In peritrichous arrangement or multiple flagella across the entire cell surface, run and tumble motility is adopted by the cell (Yang et al., 2016). This is observed in *Salmonella eneterica*, *E. coli* and *Bacillus subtilis*. Pathogenicity can be caused by bacteria only when it reaches the target site, colonize, invade the host tissue and proliferate. This makes motility an advantage for infectious bacteria. In the case of aflagellate mutants of various pathogens (*Borrelia burgdorferi*, *Campylobacter jejuni, Clostrodium chauvoei*, *Proteus mirabilis*, *Salmonella typhimurium*, etc.) it has been observed that their ability to cause a disease reduced drastically suggesting strong link between flagella and virulence (Penn and Luke, 1992; Moens and Vanderleyden, 1996). In addition to motility, flagella aid the bacteria even in rotating which is essential for adhesion in many cases. Flagella confers bacterial pathogenicity not only by providing propulsion or motility to the cell but also plays multiple functions such as surface sensing which is required for colonization through biofilm formation, adhesion or invasion to host epithelial cells, secretion of virulence factor, chemotaxis, and triggering a pro-inflammatory response of the host eukaryotic cells altering their immune system response mechanism (Duan et al., 2013). Besides all these, the sense of reaching a surface is also signaled by flagella. Once the bacteria reach a surface or come in contact with another cell, it attaches to the surface and provides signals to the cell about the surface contact which is actually the outcome of hindered rotation or motility (Kimkes and Heinemann, 2019). In *E. coli* and *P. mirabilis* with the variation in viscosity of their fluid environment, mechanical load varies which ultimately results in obstruction or inhibition of their flagellar rotation (Chawla et al., 2017). Similar results were observed with *B. subtilis*, *Caulobacter crescentus* and *Vibrio parahaemolyticus* in case of contact with surfaces, which provided evidence for the role of flagella in mechanism associated with surface-sensing and initiation of surface-dependent behavior (Lele et al., 2013; Gordon and Wang, 2019). This provides evidence that flagella are the main surface sensor in various bacteria and thus play an important role in pathogen-host interaction.

Pili (pilus)

One of the mechanisms for microbial cell-to-cell interactions is through physical contact between the cells. These physical

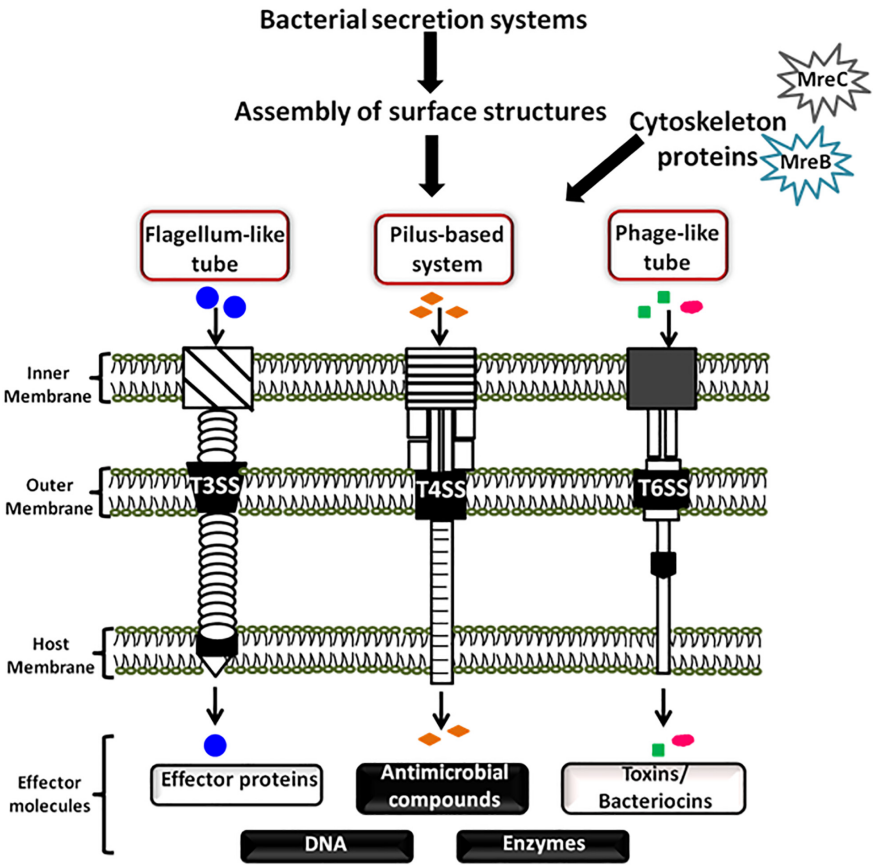


FIGURE 1 | Schematic diagram showing bacterial cell communication mediated by surface appendages and cytoskeleton proteins. Bacteria have evolved two broad ways of communication systems- the contact-independent and the contact-dependent signaling mechanisms. The figure is a schematic showing cellular communication mediated by the bacterial cytoskeleton proteins and contact-dependent signaling. The cellular apparatuses which play a major role in cellular communication through contact-dependent signaling are assembled by the secretion systems. These appendages transfer the secreted effector molecules (proteins, DNA, antimicrobial compounds, toxins, and enzymes) by intra and/or inter-species or even inter-kingdom interactions.

contacts or proximity between the cells are necessary for various function such as to assist the exchange of chemical signals and communicate, to develop multicellular structures. Moreover, for host-bacterium interactions and even for intracellular invasions by bacteria, physical contact becomes an essential requirement. One of the most important outcomes of physical interactions is the genetic exchange which leads to direct transfer of information from one cell to another cell. Such cell-to-cell interactions either between the cells of same species or others are often mediated by non-flagellar surface appendages known as pili.

Based upon the morphology and function, pili are classified into various sub-forms. There are long conjugative pili and short adhesive or attachment pili (Kuehn, 1997). Conjugative (F or sex) pili facilitate the transfer of genetic material between two bacterial cells. The two most extensively characterized pili of Gram-negative bacteria are Type I and Type IV pili. These pili are thin (2–8 nm diameters), several micrometers long and are non-covalently homopolymerized with their pilin subunits. The type I pili are common in the family Enterobacteriaceae and pathogenic *E. coli* strains. Apart from imparting adherence properties to the bacteria, they also play role in biofilm formation. The type IV pili are found in large variety of Gram-negative bacteria such as *Neisseria gonorrhoeae*, *Neisseria meningitides*, *Pseudomonas aeruginosa*, *Salmonella enterica*, *Vibrio cholerae*, etc (Craig et al., 2004). These pili besides playing key role in host cell adhesion, phage transduction, DNA uptake during transformation, most importantly provide "twitching motility" to the bacteria. The ability of type IV pili structures to undergo repetitive cycles of extension and retraction supports their pivotal roles in adhesion, natural transformation and motility (Adams et al., 2019). Motility is driven by interaction (polymerization) and retraction (depolymerization) of the major pilin subunit (PilA) in the presence of two cytoplasmic ATPase, PilF/B and PilT, respectively (Proft and Baker, 2009).

The pili observed in *Agrobacterium tumefaciens* are found to be similar to bacterial conjugative pili on the basis of their assembly and function but are morphologically distinct. These pili are observed on the surface of bacterium only under the induced expression of *vir* genes. Homologs of *vir* genes are even observed in bacterial species of *Bordetella* and *Helicobacter* in their toxin export systems. The adhesive pili are the general

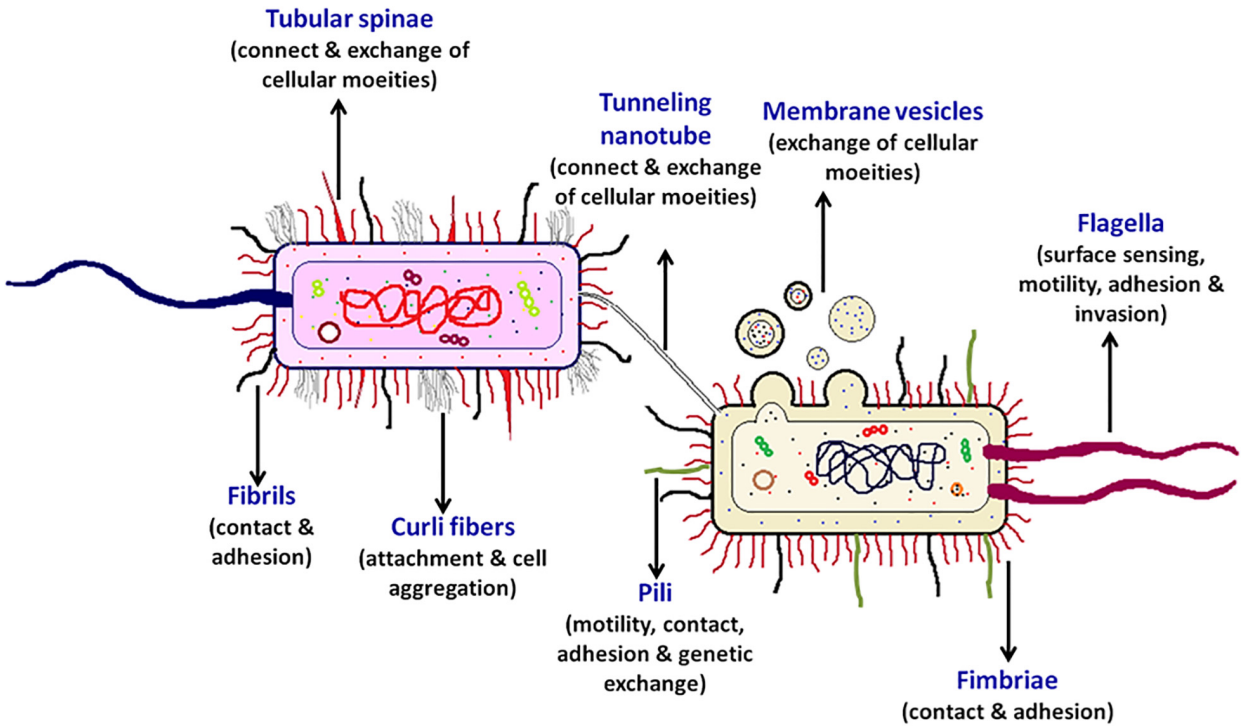


FIGURE 2 | Bacterial cell-surface appendages. Various proteinaceous tubular or fibrous structures are found extending from the surface of bacterial cell wall. These structures help the bacterium in diverse functions such as in locomotion, adhesion, surface attachment, signaling, exchange of effector proteins, and genetic material, etc. Every cell surface appendage plays a specialized role and eventually helps in mediating bacterial communication. The strings like tubular structure flagellum, protruding outward from a bacterium is majorly known for providing motility, surface sensing, adhesion, and invasion. Pili helps in providing twitching motility to the bacterium, helps in adhesion and developing contacts for genetic exchange. The tunneling nanotubes are the tubular structures which help in exchange of cellular moieties in the adjacent cells. Whereas, the spherical shaped membrane vesicles help in exchange and delivery of various cellular moieties even to the far away cells. The curli fibers helps the bacterium in attachment and cell aggregation, the two filamentous surface appendages fimbriae and fibril help in contact and adhesion, the tubular spinae connect the cells and helps in the exchange of cellular substances.

characteristic of pathogenic bacteria and are involved in initial adhesion of the bacterium to the host cells.

Pili are even detected on the surfaces of Gram-positive bacteria. There are various reports on characterization of pili in Gram-positive pathogenic bacteria such as *Clostridia*, *Corynebacterium*, *Enterococcus*, *Ruminococcus*, and *Streptococcus* (Telford et al., 2006). Apart from these, pili are also reported in various members of extensively studied Gram-positive Lactic acid bacteria (LAB) belonging to genera *Lactococcus* and *Lactobacillus*. Pili appendages aid in adhesion of these probiotic strains to intestinal epithelial cells and thus provide significant contribution in exerting the response of host immune system (Chapot-Chartier and Kulakauskas, 2014).

Broadly, two types of pili are identified in Gram-positive bacteria by electron microscopy. One form is of the thin (1–2 nm diameters) and short rods (70–500 nm length) whereas the other form is comparatively thicker (3–10 nm diameters), longer (0.3–3 μm length) and more flexible. The main building blocks of these pili are the subunits of pilin (pilus protein) present in multiple copies along with some minor pilus proteins or accessory pilins to form pilus shaft. These subunits are covalently connected and are anchored with the aid of sortase enzyme to the cell wall. The pili reported in the Gram-positive pathogens such as *Corynebacterium diphtheriae*, *Streptococcus pneumoniae* (pili detected in some strains not all), *Streptococcus pyogenes*, *Mycobacterium tuberculosis* and *Actinomyces naeslundii* are used as one of the most powerful means of cellular attachment and colonization (Proft and Baker, 2009).

It has been reported that various pathogenic bacteria, for instance some members belonging to genera *Pseudomonas* and *Xanthomonas* possess genes for hypersensitive reaction and pathogenicity (*hrp*). These *hrp* genes (mainly *hrpS* and *hrpC*) produces filamentous surface appendage known as Hrp pilus, responsible for bacterial entry into targeted host cell. The main structural protein of Hrp pilus is HrpA protein and in *hrpA* mutant, the bacterium becomes incapable of forming Hrp pilus and thus losses its ability to cause disease in plant (Roine et al., 1997). Similar to this, formation of *virB*-dependent pilus is reported in *A. tumefaciens* for transfer of T-DNA into plant cells (He, 1996). These reports suggest formation of surface appendages as a widespread attribute of both plant and animal pathogenic bacteria for causing infection (Dangl, 2013). Moreover, for surface-associated *P. aeruginosa*, pilus motors (type-IV pili) are found to be involved in mechanosensing. Mechanical tension generated either by inhibiting the retraction of type-IV pili motors or due to shearing and friction associated

TABLE 1 | Bacterial surface appendages, composition and the list of microorganisms.

S. No.	Surface appendages	Composition	Microorganisms	References
1.	Curli fibers	Amyloid protein	*E. coli*, *Salmonella* spp., *Shigella*, *Citrobacter*, *Enterobacter* spp., *Pseudomonas* spp., *B. subtilis*, *S. coelicolor*, *S. aureus*, etc	Römling et al., 1998; Barnhart and Chapman, 2006; Nicastro and Tükel, 2019; Taglialegna et al., 2020
2.	Fibrils	Protein and polysaccharide	*M. xanthus*, *A. tumefaciens,S. typhimurium*, etc	Matthysse, 1983; Dworkin, 1999
3.	Fimbriae (SEF 14, SEF 21, Myf, pH6, F1, Type I, Type IV)	Protein (fimbrin)	*E. coli*, *K. pneumonia*, *Salmonella* spp., *Y. enterocolitica*, etc	Müller et al., 1991; Rehman et al., 2019
4.	Flagella	Globular protein (flagellin)	*E. coli, Vibrio* spp., *Pseudomonas* spp., *C. crescentus, S. eneterica*, *B. subtilis,* etc	Moens and Vanderleyden, 1996; Haiko and Westerlund-Wikström, 2013; Subramanian and Kearns, 2019
5.	Membrane vesicles (MVs, OMV, O-IMV)	Lipopolysaccharides, outer membrane proteins, phospholipids and periplasmic proteins	*P. aeruginosa, B. subtilis, Acinetobacter baumannii*, *Neisseria gonorrhoeae, Shewanella vesiculosa, Myxococcus xanthus, Flavobacterium* spp*., Vibrio* spp.	Deatherage et al., 2009; Turnbull et al., 2016
6.	Pili (Type I, Type IV)	Fibrous protein (pilin)	*E. coli, P. aeruginosa*, *S. enterica*, *V. cholera, A. tumefacians, Corynebacterium* spp., *Enterococcus* spp., *Ruminococcus* spp., *Streptococcus* spp., etc	Sauer et al., 2000; Kline et al., 2010; Lukaszczyk et al., 2019
7.	Tubular spinae	Protein (spinin)	*S. maritime, Rosebacter* spp., *Agrobacterium* spp., *Chlorobium* spp., etc	Easterbrook and Coombs, 1976; Bayer and Easterbrook, 1991; Bernadac et al., 2012
8.	Tunneling nanotubes (TNTs/nanopods/nanowires)	Membranous segments/lipid bilayer/outer membrane and periplasmic extensions	*E. coli, B. subtilis*, *S. typhimurium, C. acetobutylicum, D. vulgaris, S. aureus*, etc	Dubey et al., 2016; Baidya et al., 2018

with the twitching motility and surface adhesion of the bacterium acted as the signals for biological responses. As a resultant, cAMP-dependent upregulation of virulence and an increase in the production of c-di-GMP involved in biofilm formation is observed in the bacterium (Rodesney et al., 2017). Furthermore, the crucial role of type IV pili has also been observed in *N. gonorrhoeae*. It has been reported that attachment of bacteria to the host cell and cell-cell interaction driven assembly of microcolonies on the surfaces are primarily mediated by type IV pili (Pönisch et al., 2017). In *Thermus thermophilus,* with the help of cryo-electron microscopy and mass spectrometry it is revealed that the bacterium possesses two different types of type IV pili- wide and narrow. Both not only differ on the basis of protein composition and structure but have different functions as well. The wide pili composed of the PilA4 pilin protein is majorly required by the bacterium during DNA uptake (natural transformation) whereas the narrow pili composed of the PilA5 pilin protein is required for the twitching motility (Neuhaus et al., 2020). Moreover, in *Geobacter sulfurreducens* contact-dependent mechanism of communication using pili as nanowires has been observed. It has been demonstrated that pili serves as the electric conduits for electron transfer in *G. sulfurreducens* fuel cells and on Fe (III) oxide surfaces to form biofilms (Reguera et al., 2007).

In general, DNA-uptake by pili can be studied in two steps; first step includes mechanistic possibilities of DNA-binding to the pili and second step could be the DNA path through cell's initial barriers. Various mechanistic hypotheses for the same have been given but all vary between different bacteria (Piepenbrink, 2019).

Tunneling Nanotubes/Intercellular Nanotubes

The tunneling nanotube (TNTs) connects bacteria of the same or different species and enables transfer of intra and intercellular contents (Marzo et al., 2012). Studies done on *B. subtilis*, *Staphylococcus aureus* and *E. coli* proposed TNTs as one of the major form of bacterial communication method for exchange of cellular molecules between and within the species (Dubey and Ben-Yehuda, 2011). Bacterial nanotubes are categorized into two basic forms: (1) The thin nanotubes that connect the nearby neighboring cells and (2) thick nanotubes connecting the distal cells. With the help of these TNTs, cell bridges communication with neighboring bacteria for exchange of cytoplasmic constituents, transmission of DNA and plasmids and also to acquire non-hereditary resistance to antibiotics from the nearby cells (Dubey and Ben-Yehuda, 2011).

Interestingly, nanotube connections have been demonstrated in various other bacterial species as well. Under nutritional stress, the induced cell-to-cell interactions through nanotubular structures resulted in bidirectional exchange between Gram-positive bacterium *Clostridium acetobutylicum* and Gram-negative bacterium *Desulfovibrio vulgaris* (Benomar et al., 2015). Likewise, *Acinetobacter baylyi* and *E. coli* showed connection through membrane-derived nanotubes for exchange of nutrients (Pande et al., 2015). An intermediate structure (outer membrane tubes) between nanotube and chain-like structure has also been reported in *Myxococcus xanthus* for exchange of cellular moieties on cell-cell contact (Wei et al., 2014). Thus, nanotube connections have shown to help in distributing metabolic functions across the various microbial communities. Moreover such tubular connections are also been observed between *S. typhimurium* cells and its host eukaryotic cells resulting in a host-pathogen interaction (Galkina et al., 2011).

Membrane Vesicles

Many bacteria release spherical, extracellular vesicles packaged with specific molecules involved in diverse functions. These membrane vesicles (MVs) have similar composition to that of the outer membrane consisting of lipopolysaccharides (LPS), outer membrane proteins, phospholipids and even some periplasmic proteins. In most cases, vesicles have only one membrane but two membranes i.e., outer and inner membrane vesicles (O-IMVs) are also observed in few bacterial species such as *Acinetobacter baumannii*, *N. gonorrhoeae*, *P. aeruginosa* and *Shewanella vesiculosa* (Pérez-Cruz et al., 2015).

The formation of MVs was earlier believed to take place by controlled blebbing of the outer membrane. Both Gram-negative and Gram-positive bacteria have different cell wall structure, it is likely that they posses different mechanisms for MVs formation (Toyofuku, 2019). Extensively studied Gram-negative bacterium *P. aeruginosa* demonstrated the formation of MVs through explosive cell lysis (Turnbull et al., 2016). However, Gram-positive bacterium such as *Bacillus subtilis* produces MVs through bubbling cell death. *Bacillus* cells did not explode rather died (ghost cells) but retained their cell morphology while releasing MVs (Toyofuku et al., 2019). Variation to these synthesis processes is the genesis of membrane tubes from the outer membrane as an intermediate before vesicle formation. An abiotic morphogenic process known as pearling transform these unstable intermediate membrane tubes into stabilized chains of interconnected vesicles. These chains of interconnected vesicles enlarge the cell surface of bacteria resulting into increased surface enzymes per cell volume (Bar-Ziv and Moses, 1994). These vesicle chains are well reported in various species of Flavobacteria such as *F. columnare*, *F. psychrophilum* and also in strain Hel3_A1_48 of a marine flavobacterium (Fischer et al., 2019). Apart from these, *Francisella novicida*, *Myxococcus xanthus*, and *Shewanella oneidensis* also showed chains of vesicles on their cell surfaces (Subramanian et al., 2018; Fischer et al., 2019).

These released membrane vesicles form one of the ways implied for prokaryotic communication. The packaged MVs travel and fuse with distant cells and thus facilitate exchange of various cellular molecules such as those involved in the process of QS, transfer of factors responsible for development of antimicrobial resistance, delivery of toxins, and even exchange of the genetic material (Dubey and Ben-Yehuda, 2011). MVs mediate communications between interspecies and also between different cells of interkingdom (Mashburn-Warren and Whiteley, 2006). Recent researches have shown that MVs fusion, their transmission to specific target cells in microbial and host-microbial interactions is self-guided. MVs have been demonstrated as the carrier of various signaling molecules involved in bacterial cell-to-cell communication. In pathogens like *P. aeruginosa,* MVs transport the interbacterial signaling molecule *pseudomonas* quinolone signal; PQS (2-heptyl-3-hydroxy-4-quinolone) to the bacterial population for dealing with the hostile environmental conditions (Häussler and Becker, 2008). Similarly, in marine pathogen *Vibrio harveyi*, outer membrane vesicles (OMVs) package the long-chain amino-ketone CAI-1 QS signaling molecule triggering the QS phenotype not only in CAI-1 non-producing *V. harveyi* but also in *V. cholerae* cells (Brameyer et al., 2018). Moreover, the coral-associated pathogen, bacterium *Vibrio shilonii* also releases OMVs containing N-acyl homoserine lactones (AHLs) signaling molecules, alkaline phosphatase, chitinase, and lipase (Li et al., 2016). Furthermore, in the soil bacterium *Paracoccus denitrificans* PD1222, the long-chain AHLs (C16 - N-(hexadecanoyl)- L-homoserine lactone) associated with cell-to-cell communication are majorly released in the population through MVs (Toyofuku et al., 2017).

Tubular Spinae

The long, hollow, and tubular appendages known as spinae have been reported in various Gram-negative bacteria. Spinae are non-prosthecate (echinuliform) appendages which do not have any connection with the cytoplasm. They are about 3 μm in length, 50–70 nm in diameter and observed randomly on the cell surface (Kim, 2017). In Gram-negative pseudomonad *Spinomonas maritime,* production of long tubular surface appendages (spinae) is controlled by various growth parameters such as osmolarity, temperature and pH. It has been observed that at slightly elevated temperature of about 34°C with pH 7.4 and relatively low osmolarity of the growth medium, highest yield of spinae was obtained whereas no or very less number of spinae were observed at lower temperature, pH, and higher salt concentration (Easterbrook and Sperker, 1982). Apart from playing major role in cell protection from protozoan predators and cell sedimentation, there are reports which suggest spinae as long-distance cell-to-cell connectors. Examination of the *S. maritime* cells grown in low osmolarity medium through scanning and transmission electron microscopy showed various cells connected to each other through tubular spinae over distances of several micrometers (Bayer and Easterbrook, 1991). These surface appendages apart from permitting cell-to-cell signal exchange could also be a way of uniting similar or single type of cells in the present diverse forms of multicellular organisms.

Fibrils

Another mechanism of cell-to-cell interaction is mediated through extracellular appendages known as fibrils. These

appendages are reported in various bacteria but have been extensively studied only in the Gram-negative myxobacterium *Myxococcus xanthus* (Dworkin, 1999). They are filamentous organelles, 15–30 nm in diameter composed of equal amounts of polysaccharides and proteins. Fibrils are differentiated from pili based upon size as they are thicker and longer as compared to pili. They may be arranged either densely as clumps or tufts together or even can be sparsely distributed all over the bacterial cell surface. Similar to pili, fibrils are also necessary for the social behavior of the cells. They assist in maintaining the physical contact between the cells and even in between the cells and their substratum. Apart from this, there are various other examples of fibrils, morphologically similar to those described for *M. xanthus*. Fibrils also serve an important function of attaching the bacterial cell to their host targets as seen in the case of cellulose fibrils synthesized by the *Agrobacterium tumefaciens* (Matthysse, 1983). The fibrils thus synthesized anchor the bacteria to host cells and aid in production of crown gall tumors. Similarly, the short, stubby fibrillar appendages formed due to the contact between *S. typhimurium* and cultured epithelial cells have been shown to be necessary for internalization of the bacteria. Another important example is of an extracellular filamentous appendage haemagglutinin of the virulent *Bordetella* which in conjunction with the pertussis toxin causes pathogenicity.

Fimbriae

Adhesion to the host cell is considered as the initial and one of the most crucial steps in pathogenesis. Functionally similar to fibril, another potential adhesin which mediates attachment of bacterial cell to the host receptor molecule is fimbriae. They are proteinaceous filamentous surface appendages consisting of helically arranged fimbrin (protein) monomers (Müller et al., 1991). Fimbriae though functionally seem similar to fibrils but are ultrastructurally as well as biochemically different entities. Fimbriae structures ranges in between 0.5 and 10 μm length and 2–8 nm width whereas fibrils lengths vary in different strains (Handley, 1990; Rehman et al., 2019). Presence of fimbriae in some cases provides strong indication toward bacterial virulence. They facilitate adhesion and are also involved in bacterial aggregation, colonization and biofilm formation (Müller et al., 1991). In host, these extracellular bacterial appendages play important role in interaction with macrophages and intestinal persistence. Initially, based upon their morphology and hemagglutination patterns (ability of the mannose monosaccharide to inhibit the adhesion of fimbriae to erythrocytes i.e., mannose sensitive or mannose resistant) fimbriae were categorized into 7 types ranging from Type I to VI and Type F (Clegg and Gerlach, 1987). These surface appendages types are common to various bacteria and have been extensively studied in various members of Enterobacteriaceae family such as in *E. coli*, *Klebsiella pneumonia*, *Salmonella* etc. But later, with the help of serological tests, genetic relatedness of fimbrial antigens was identified and fimbriae were further classified into 3 different types based upon their assembly pathways. The three assembly pathways which lead to multiple types of fimbriae are chaperon–usher (CU), nucleation–precipitation (N/P), and Type IV fimbriae (Rehman et al., 2019). SEF 14 and SEF 21 are the two comprehensively studied Type I fimbriae of *Salmonella enteritidis* (Müller et al., 1991). Thus, affinity provided by fimbriae to bind with host cell receptors helps in host-pathogen interaction and cause severe disease and infections like salmonellosis.

These surface appendages help in contact-dependent signaling, mechanosensing or inhibition of the receiver cell.

Mechanism of Contact-Dependent Cellular Communication

The highly specialized secretion systems of bacteria secrete a variety of molecules including proteins and DNA which have significant role in bacterial communication. Based upon the type of secretion system, these secreted molecules have three possible outcomes, first is either they remain anchored to the outer membrane of the producer cell or second they get released into the extracellular space or third they are directly injected into the targeted bacterial or eukaryotic cell (Costa et al., 2015).

The type III, IV, V, and VI secretion systems (T3SSs, T4SSs, T5SSs, and T6SSs) provide important example of contact-sensor mechanism. Out of these systems, the effector molecules from T3SS, T4SS, and T6SS generally crosses all the three phospholipid membranes (two of producer bacterial cell and one of the hosts) and reaches the cytosol of host with the help of assembled surface structure (**Figure 1**). The T4SSs are pili-based secretion systems which majorly require cell-to-cell contact between the emitter and receiver cells. This contact is mediated by the cell surface adhesins and/or pili structures. In Gram-negative pathogens, it has been reported that the pili of T4SS interacts with various proteins of host cell surface (Hayes et al., 2010). Various other well reported pili such as the F-pilus from *E. coli* serve as a gripping hook that elongates and retracts from the cell surface to bring both donor and the recipient cell together for direct cell-to-cell interaction. Similarly, the P-pili of the conjugation systems act as the adhesive structure for binding together the mating cells. Moreover, in *A. tumefaciens* it has been shown that the primary component of the T-pilus VirB2 pilin protein is responsible for cell-cell interactions, assisted along with the VirB5 protein localized at the tip of the T-pilus (Aly and Baron, 2007; Backert et al., 2008).

In addition to this, the type III secretion systems (T3SSs) of various Gram-negative pathogens report the use of flagellum-like apparatus known as injectisome (needle complex) for the delivery of effector proteins to the target host cells (Cornelis, 2006). This apparatus has a similar structural organization to flagella and also shares amino acid sequences homology with several proteins (Hayes et al., 2010). Moreover, the component proteins (FlhA, FlhB, FliO, FliP, FliQ, FliR FliH, FliI, and FliJ) of the flagellar export apparatus shares considerable sequence similarities with the pathogenic bacteria possessing type III secretion system (Minamino, 2014).

Likewise, studies on *Campylobacter jejuni* report the importance of cell surface appendages in developing the contact with host cells evading their defense systems. It has been shown that in order to possess antigenic diversity and provide a competitive advantage to *C. jejuni*, the surface structures are subjected to various modifications such as *O*- and

N-linked glycosylation of the flagella filament. These flagellar modifications ultimately affect the pathogen- host interactions. Thus flagella are important for locomotion as well as for the pathogenesis of the bacterium (Cullen et al., 2012). Interestingly, another important example of bacterial cross-talk has been reported in *S. typhimurium*. The bacterium after sensing the signals from host assembles the surface appendages for further signaling. These surface-associated appendages help the bacterium in developing contact with the host cell to initiate bacterial uptake and thus demonstrates a significant example of two-way biochemical signaling for the adaptation of a pathogen by its host (Ginocchio et al., 1994).

There are multiple contact-dependent mechanisms by which bacteria can communicate; the major ones are depicted in **Figure 3**. Another example is of the syntrophic relationship between *Pelotomaculum thermopropionicum* and *Methanothermobacter thermautotrophicus*. The bacterium *P. thermopropionicum* liberates hydrogen gas by fermentation and *M. thermautotrophicus* is a methanogen that uses H_2 to reduce CO_2 (Ishii et al., 2005). The flagellum expressed by the *P. thermopropionicum* specifically bind to *M. thermautotrophicus* to trap it within range for H_2 diffusion and enhance methanogenesis (Shimoyama et al., 2009; Hayes et al., 2010). In case of *M. xanthus*, the cell-cell signaling for the formation of fruiting bodies within which they sporulate is observed through direct physical interactions. A contact-dependent cue known as C-signal is exchanged when two cells comes in contact end-to-end (Kaiser, 2004; Visick and Fuqua, 205). Whereas, the other mechanisms involve cell surface appendages which develops the contact between two adjacent for even between two far-away cells to further mediate signaling. Therefore, it is observed that every cell surface appendage plays a specialized role in developing bacterial communication. The Curli fibers and type IV pili helps in recognition, adherence and even in invasion into the target cells whereas have no role in transport (Fronzes et al., 2008). In contrast to this, the tunneling nanotubes, tubular spinae connect the cell, and transport signaling molecules.

Contact-Dependent Inhibition

In order to survive in the vast microbial community, bacteria exhibit various cooperative and competitive interaction mechanisms. One such contact-based competitive mechanism is Contact-dependent growth inhibition (CDI). CDI systems are evolved in bacteria as a mechanism to inhibit growth or kill the neighboring outcompeting cells. To execute its operation, CDI system requires a direct contact between the producer and the targeted cell. These systems deliver polymorphic proteinaceous toxins into the cytoplasm of adjacent competitors unless they produce a corresponding antidote protein. Thus, CDI system also mediates cooperative communication between the cells which produce identical toxin-antitoxin pairs (Garcia et al., 2016; Roussin et al., 2019).

The most widely studied contact-dependent growth inhibition systems present in various bacterial pathogens comprises of Cdi toxins (CdiA and CdiB). Mostly, the sec complex of the type V secretion system (T5SS) is used for translocation and anchoring of the CdiA and CdiB toxin from inner to the outer membrane of the producer cell and then directly deliver to the target cell. CdiA protein also promotes cell-cell adhesion needed for contact-dependent inhibition (García-Bayona et al., 2017). In addition to T5SS structures, the surface-associated appendages of T6SS and T7SS also deliver toxin directly to the adjacent cells (García-Bayona et al., 2017; Xiong et al., 2018). The T6SS which is often known for its antagonism utilizes a bacteriophage-like subassembly to secrete effector proteins to both prokaryotic and eukaryotic target cells. The system shares both structural features as well as protein sequences with phages. In addition to the role of T6SS in mediating interbacterial competition, the system also has potential role in signaling and virulence (Hayes et al., 2010; Russell et al., 2014).

Interestingly, a contact-dependent inhibition mediated by glycine zipper proteins (Cdz) system involving a T1SS is reported in *C. crescentus* that enables the bacterium to kill the neighboring cells with contact-inhibition (García-Bayona et al., 2017).

Contact-Dependent Mechanosensing

One of the important aspects of all the living cells is to sense and respond to various signals. Similarly, bacteria communicate with their environment through sensing and responding to chemical, biological and physical signals. In bacteria, surface-sensing and attachment dependent behavior indicate toward their physical-sensing or mechanosensing properties. These two steps are important and significant in the biofilm mode of bacterial growth. When a planktonic bacterium senses and reaches a surface to become sessile, it undergoes through various substantial changes. These involve sensing: (a) physicochemical changes, (b) attachment of cellular appendages, and then (c) attachment of the bacterial cell body to the surface (Kimkes and Heinemann, 2020). As the microenvironment near bacteria and surface is not same, difference is observed in them on the basis of ionic strength, nutrient availability, pH and osmolarity (Berne et al., 2018). To sense these differences, a two-component signal transduction system is usually used by bacteria. It comprises of a membrane-bound histidine kinase and a cytoplasmic response regulator to sense the stimulus and thus mediate the cellular response, respectively. For example in *E. coli* the major systems involved in sensing physicochemical changes and thus resulting in downstream regulation for biofilm formation are CpxAR, EnvZ/OmpR and RcsCDB (Gordon and Wang, 2019; Kimkes and Heinemann, 2020).

Next to this occurs the adhesion of bacterial cell appendages to the surface. Studies suggest that bacterial motility appendages such as type-IV pili, flagella and even the envelope proteins are likely to be the potential mechanosensory elements involved in adhesion (Gordon and Wang, 2019). As the bacterium attaches its appendages to the surface, due to the change in mechanical properties of bacterium's environment, biological responses such as phenotypic changes are observed. Thus, with initial adhesion impairment in motility is observed and the movement of cell appendage is hindered (Stones and Krachler, 2016). This generates and transmits signals to the cell indicating its attachment to the surface.

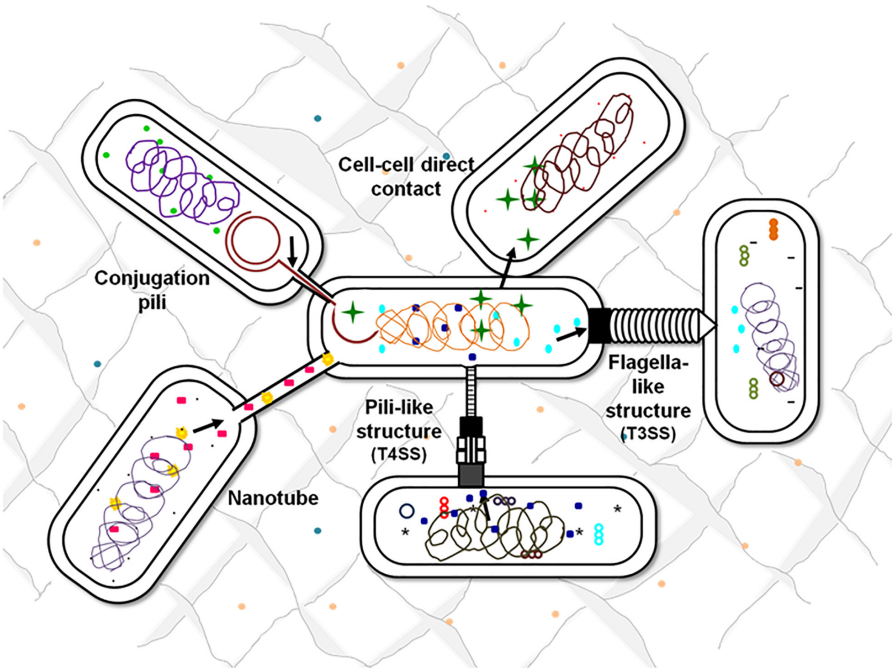


FIGURE 3 | Contact-dependent signaling mechanisms. Effector molecules can be transferred from one cell to another using contact-mediated signaling methods. Exchange of metabolites can be either through the surface appendages or can be even by direct surface contact. Flagella-like structures assembled by T3SSs, pili-based structures by T4SSs, intercellular nanotubes, and conjugation pili are the appendages-based methods that facilitate the exchange of cellular contents.

Further adhesion of bacterial cell body to surface is mediated by interactions of various forces such as long-range Van der Waals, short-range repulsive electrostatic forces and acid-base interactions (Berne et al., 2018). This contact can be additionally supported by cell appendages or by the long O-antigen part of lipopolysaccharides (LPS). Moreover, production of adhesins also plays crucial role in attachment of cell body with surface in both polar and flat orientation. For example, *A. tumefaciens* and *C. crescentus* have polar orientation during attachment due to synthesis of polar adhesins at the cell poles whereas in *P. aeruginosa* with the production of an exopolymeric matrix component- Pel polysaccharide, a transition is observed from polar adhesion to a flat orientation. Following this entire process bacteria attaches to the surfaces, sticks together and promotes bacterial aggregation, colonization, and biofilm formation (Li et al., 2012; Cooley et al., 2013; Kimkes and Heinemann, 2020).

Studies related to contact-dependent signaling for communication conducted on various microorganisms such as on *E. coli*, *B. subtilis*, *M. xanthus*, and *Lactobacilli* suggests that physical contact apart from being a direct way of bacterial communication also helps in various mechanisms through which bacteria can optimize the use of quorum-sensing molecules. Thus physical contact mediated either through direct cell-to-cell contact or with the help of surface appendages helps in all the prevailing mechanism of bacterial communication (Harapanahalli et al., 2015).

As stated by Stacy and coworkers, that every biological interaction cannot be considered as genuine communication rather it can be a response to a cue or coercion or sometimes just a contact which eventually does not lead to any molecular exchange (Stacy et al., 2012). Thus, every contact developed through the surface appendages and cytoskeleton proteins do not necessarily lead to such communications in which both the emitter and receiver organism gains benefits but are actually contact-mediated mode of establishing connections that can eventually assist in communication or even in cue and coercion.

BACTERIAL CYTOSKELETON IN CELL COMMUNICATION

Besides surface appendages, bacterial cytoskeleton is also involved in cell communication. Interestingly, counterparts of all three-known eukaryotic cytoskeletal proteins (actin, tubulin, and Intermediate filaments) have been found in eubacteria that form filamentous structures and show cytoskeletal properties. In general, the three major functions carried out by cytoskeleton are: (1) spatially organizing the contents of the cells, (2) connects the cell to the external environment both biochemically and physically and (3) plays key role in cell motility, shape change and cell division. All these functions are performed by the dynamic and coordinated activities of the cytoskeletal proteins (Shih and Rothfield, 2006).

Interestingly, a study conducted on *Salmonella* with actin-like proteins, revealed that cytoskeleton proteins influence motility and colonization of the bacteria (Bulmer et al., 2012). In prokaryotes various actin-like proteins include MreB, MreB-like proteins (Mbl and MreBH), FtsA, ParM (StbA), ActA, etc.

In *Salmonella*, MreB was found to be an essential protein. In Δ*mreC* mutant, the flagella system and the expression of virulence factors were found to be down regulated. Moreover, *mre* operon had shown to play an important role in colonization of the bacterium during infection. The study suggested the strong connection between the bacterial cytoskeleton and pathogenicity, as expression of virulence genes were observed to be in direct coordination with the cytoskeletal integrity (Bulmer et al., 2012).

Similarly, in *Helicobacter pylori* the key role of MreB cytoskeletal protein was to maintain the enzymatic activity of a virulence factor urease rather than maintaining cell shape (Waidner et al., 2009). Furthermore, a different aspect of the cytoskeleton protein MreB has also been demonstrated in *P. aeruginosa*. In this bacterium, MreB protein was found to be essential for the production and polar localization of type IV pili. The type IV pili are significant for virulence and even for providing antibiotic resistance to the bacterium by biofilm formation during chronic infections (Cowles and Gitai, 2010).

Motility is an important feature by which even far away cells can reach each other and communicate. Gliding motility is one of the motility type by which bacteria actively move over the surfaces without the involvement of flagella (McBride, 2001). *Myxococcus xanthus*, uses gliding motility to move along the solid surfaces even without the aid of type IV pili. But *M. xanthus* requires MreB, the bacterial actin for its motility. It has been observed that there is interdependency between the movement machineries of gliding motors and MreB filaments. This is analogous to the movement of myosin motors and actin in eukaryotic cells (Fu et al., 2018). Similar to this, gliding motility is observed in various species of mycoplasma such as in *M. genitalium*, *M. pulmonis*, and *M. mobile,* etc. A well-defined cytoskeleton comprising of 25 different proteins is reported in mycoplasma. It has been observed that the cytoskeleton of mycoplasma plays important role in its gliding motility as unusual or irregular shaped cells were observed in non-motile mutants of *M. mobile* (Miyata et al., 2000).

TARGETTING CELL COMMUNICATION

A report states that biofilm formation is involved in about two-thirds of the human infections. This includes infections of skin (integumentary), ears (auditory), urinary tract, respiratory tract, reproductive organs, digestive system, uncontrolled dental plaque, and fouling of various implants and even infection of contact lenses (Otto, 2019; Vestby et al., 2020). It becomes difficult to deal with such infections as biofilms provide more tolerance to bacteria against the antimicrobial treatments and even toward the host immune responses and defense mechanisms (Kimkes and Heinemann, 2019).

The detailed understanding of the bacterial surface associated structures and related cytoskeleton proteins involved in cell communication will not only aid in designing of potential regulators or inhibitors in the form of novel therapeutics but will also have direct application in preventing biofilm formation one of the major causes of pathogenesis. Various other applications of targeting bacterial cell communication are illustrated in the **Figure 4**. There are various potential targets at different levels that can be explored for designing of therapeutics. On the basis of knowledge about biogenesis of various tubular extensions, the key genes responsible for the structural functionality of the surface appendages can be targeted. Production of exopolysaccharides and adhesive molecules responsible for irreversible attachment of pathogen to the host can also be blocked. There are various reports which demonstrate that even with blocking or hindering the flagellar rotation, biofilm formation can be controlled (Wood et al., 2006; Yoshihara et al., 2015). In a recent study conducted on *S. typhimurium* and *E. coli,* it was observed that the bacterial flagella strongly associated with the host cell

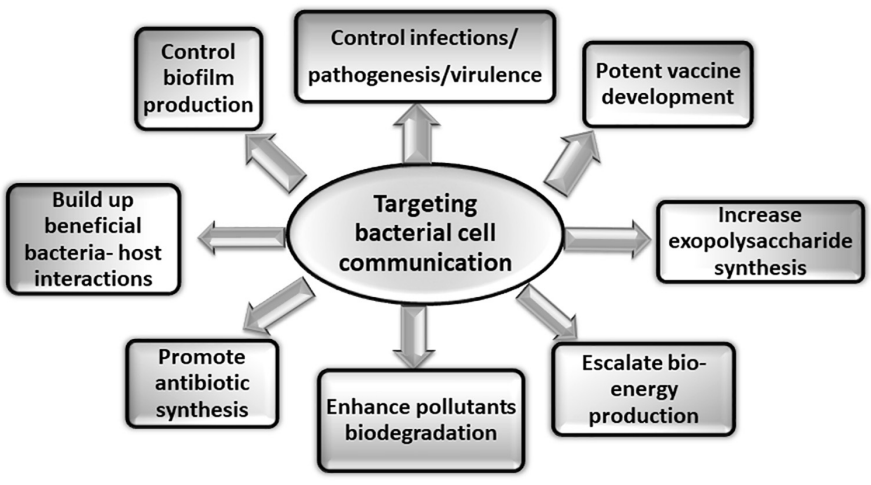


FIGURE 4 | Applications of targeting bacterial cell communication. Studies targeting the various aspects of bacterial cell communication system are of immense importance. The figure shows some of the applications that can be explored by addressing bacterial communication systems.

membrane and disrupted it with flagellar rotation. It was seen that the bacterial flagella showed affinity toward actin and actin-binding proteins and binding was observed even during *in vitro* conditions. This phenomenon suggests the existence of molecular mechanism which connects both cytoskeletal dynamics and bacterial colonization (Wolfson et al., 2020). Vaccines against the building blocks of various apparatuses such as flagellin of flagella and pilin of pili can also be an alternative approach for preventing pathogenesis. As depletion of various cytoskeletal proteins have also shown to modulate the pathogenicity of many microorganisms either directly or indirectly, so identification of more such proteins and understanding their roles in pathogenic bacteria could be useful in hampering cell-to-cell contact, growth and ultimately survival of the pathogenic bacteria. Thus, targeting and blocking bacterial communication can be one of the most advanced way of preventing infectious diseases. Other advantageous aspects of targeting apparatuses involved in cell communication could be manipulating and monitoring biofilm formation for the betterment of mankind. Not all the biofilms are problematic whereas there are naturally occurring biofilms in the environment that biodegrades various pollutants, industrial effluents and also helps in waste water treatment (Bryers, 1990; Edwards and Kjellerup, 2013). So, by promoting cell communication in the respective microbial species *in situ* or *ex situ* (at the places where bioremediation is aimed), biofilm formation can be enhanced resulting in degradation of toxic compounds. This could be an approach for minimizing the build-up of pollutants at various places (Kumar and Anand, 1998). Extracellular polymeric substances (EPS) form the architecture of bacterial biofilm matrix. It provides various properties to the biofilm so as to shield and protect it from antibiotics and different forms of stress (Vu et al., 2009; Nwodo et al., 2012). Thus, it is primary attribute of the microbial community to produce EPS in order to exist in the form of biofilm. This property of the microbial cells can be exploited. Furthermore, the production of commercially important EPS, medicines, drugs, and antibiotics can be increased by implementing such techniques of controlled biofilm formation (Ortega-Morales et al., 2010). Moreover, similar to the healthy colonization of lactic acid bacteria (LAB) in the gastrointestinal tract, more such beneficial bacteria-host interactions can be established in the forms of consumable probiotics (Kumar and Anand, 1998). To support the tremendously increasing energy demand, production of bio-energy using micofluidic devices and microbial fuel cell is the current solution to the energy crises (Singh and Verma, 2015). This strategy can also explore the importance of microbial cell communication to address the issue more efficiently.

IMPORTANCE AND FUTURE PROSPECTS

Currently, one of the major ongoing global concerns is the upsurge of anti-microbial resistance leading to the emergence of various antibiotic-resistant pathogens along with the enhanced rate of microbial evolution. In contrast to this, the development of novel antibiotics is lagging as it is being developed comparatively at a very slower rate. In this era of antibiotic-resistance, to combat with abruptly erupting deadly infectious diseases there is an urgent need of developing alternative strategies and innovative therapeutics (Njoroge and Sperandio, 2009). One such anti-virulence strategy is exploiting bacterial cell-to-cell communication. This approach could serve as a way of targeting and obstructing the outbreak of diseases. The bacterial pathogens cause harm not only to humans but equally affect other living forms such as animal health and agricultural productivity. These adverse effects of the bacterial pathogens can be controlled by hindering bacterial cell communication.

Such basic studies which aim to increase our knowledge about the bacterial systems and proteins involved in cell communication when combined with present-time modern technologies can definitely lead to development of various novel therapeutics. Research conducted on bacterial cytoskeletal system and other related apparatuses such as surface appendages which assist in communication are of great importance and hold immense expectations with their application in addressing issues related to bacterial infections and diseases.

AUTHOR CONTRIBUTIONS

DS and PS wrote and edited the manuscript.

FUNDING

This work was supported by the Department of Science and Technology, Government of India.

REFERENCES

Abby, S. S., Cury, J., Guglielmini, J., Néron, B., Touchon, M., and Rocha, E. P. (2016). Identification of protein secretion systems in bacterial genomes. *Sci. Rep.* 6, 1–14.

Abisado, R. G., Benomar, S., Klaus, J. R., Dandekar, A. A., and Chandler, J. R. (2018). Bacterial quorum sensing and microbial community interactions. *MBio* 9:e2331-17.

Adams, D. W., Pereira, J. M., Stoudmann, C., Stutzmann, S., and Blokesch, M. (2019). The type IV pilus protein PilU functions as a PilT-dependent retraction ATPase. *PLoS Genet.* 15:e1008393. doi: 10.1371/journal.pgen.1008393

Aly, K. A., and Baron, C. (2007). The VirB5 protein localizes to the T-pilus tips in *Agrobacterium tumefaciens*. *Microbiology* 153, 3766–3775. doi: 10.1099/mic.0.2007/010462-0

Backert, S., Fronzes, R., and Waksman, G. (2008). VirB2 and VirB5 proteins: specialized adhesins in bacterial type-IV secretion systems? *Trends Microbiol.* 16, 409–413. doi: 10.1016/j.tim.2008.07.001

Baidya, A. K., Bhattacharya, S., Dubey, G. P., Mamou, G., and Ben-Yehuda, S. (2018). Bacterial nanotubes: a conduit for intercellular molecular trade. *Curr. Opin. Microbiol.* 42, 1–6. doi: 10.1016/j.mib.2017.08.006

Barnhart, M. M., and Chapman, M. R. (2006). Curli biogenesis and function. *Annu. Rev. Microbiol.* 60, 131–147. doi: 10.1146/annurev.micro.60.080805.142106

Bar-Ziv, R., and Moses, E. (1994). Instability and" pearling" states produced in tubular membranes by competition of curvature and tension. *Phys. Rev. Lett.* 73:1392. doi: 10.1103/physrevlett.73.1392

Bassler, B. L. (2002). Small talk: cell-to-cell communication in bacteria. *Cell* 109, 421–424. doi: 10.1016/s0092-8674(02)00749-3

Bassler, B. L., and Losick, R. (2006). Bacterially speaking. *Cell* 125, 237–246. doi: 10.1016/j.cell.2006.04.001

Bayer, M., and Easterbrook, K. (1991). Tubular spinae are long-distance connectors between bacteria. *Microbiology* 137, 1081–1086. doi: 10.1099/00221287-137-5-1081

Behzadi, P. (2019). Classical chaperone-usher (CU) adhesive fimbriome: uropathogenic *Escherichia coli* (UPEC) and urinary tract infections (UTIs). *Folia Microbiol.* 65, 45–65. doi: 10.1007/s12223-019-00719-x

Benomar, S., Ranava, D., Cárdenas, M. L., Trably, E., Rafrafi, Y., Ducret, A., et al. (2015). Nutritional stress induces exchange of cell material and energetic coupling between bacterial species. *Nat. Commun.* 6, 1–10.

Bernadac, A., Wu, L.-F., Santini, C.-L., Vidaud, C., Sturgis, J., Menguy, N., et al. (2012). Structural properties of the tubular appendage spinae from marine bacterium Roseobacter sp. *strain YSCB*. *Sci. Rep.* 2:950.

Berne, C., Ellison, C. K., Ducret, A., and Brun, Y. V. (2018). Bacterial adhesion at the single-cell level. *Nat. Rev. Microbiol.* 16, 616–627. doi: 10.1038/s41579-018-0057-5

Blango, M. G., and Mulvey, M. A. (2009). Bacterial landlines: contact-dependent signaling in bacterial populations. *Curr. Opin. Microbiol.* 12, 177–181. doi: 10.1016/j.mib.2009.01.011

Brameyer, S., Plener, L., Müller, A., Klingl, A., Wanner, G., and Jung, K. (2018). Outer membrane vesicles facilitate trafficking of the hydrophobic signaling molecule CAI-1 between Vibrio harveyi cells. *J. Bacteriol.* 200:e740-17.

Bryers, J. D. (1990). "Biofilms in biotechnology," in *Biofilms*, eds W. G. Characklis and K. C. Marshall (New York, NY: Wiley-Interscience), 733–773.

Bulmer, D. M., Kharraz, L., Grant, A. J., Dean, P., Morgan, F. J., Karavolos, M. H., et al. (2012). The bacterial cytoskeleton modulates motility, type 3 secretion, and colonization in *Salmonella*. *PLoS Pathog.* 8:e1002500. doi: 10.1371/journal.ppat.1002500

Chapman, M. R., Robinson, L. S., Pinkner, J. S., Roth, R., Heuser, J., Hammar, M., et al. (2002). Role of *Escherichia coli* curli operons in directing amyloid fiber formation. *Science* 295, 851–855. doi: 10.1126/science.1067484

Chapot-Chartier, M.-P., and Kulakauskas, S. (2014). "Cell wall structure and function in lactic acid bacteria," in *Microbial Cell Factories*, eds D. Sharma and B. Singh Saharan (Berlin: Springer), S9.

Chawla, R., Ford, K. M., and Lele, P. P. (2017). Torque, but not FliL, regulates mechanosensitive flagellar motor-function. *Sci. Rep.* 7, 1–9.

Clegg, S., and Gerlach, G. (1987). Enterobacterial fimbriae. *J. Bacteriol.* 169:934. doi: 10.1128/jb.169.3.934-938.1987

Cooley, B. J., Thatcher, T. W., Hashmi, S. M., L'Her, G., Le, H. H., Hurwitz, D. A., et al. (2013). The extracellular polysaccharide Pel makes the attachment of *P. aeruginosa* to surfaces symmetric and short-ranged. *Soft. Matter.* 9, 3871–3876.

Cornelis, G. R. (2006). The type III secretion injectisome. *Nat. Rev. Microbiol.* 4, 811–825. doi: 10.1038/nrmicro1526

Costa, T. R., Felisberto-Rodrigues, C., Meir, A., Prevost, M. S., Redzej, A., Trokter, M., et al. (2015). Secretion systems in Gram-negative bacteria: structural and mechanistic insights. *Nat. Rev. Microbiol.* 13, 343–359. doi: 10.1038/nrmicro3456

Cowles, K. N., and Gitai, Z. (2010). Surface association and the MreB cytoskeleton regulate pilus production, localization and function in *Pseudomonas aeruginosa*. *Mol. Microbiol.* 76, 1411–1426. doi: 10.1111/j.1365-2958.2010.07132.x

Craig, L., Pique, M. E., and Tainer, J. A. (2004). Type IV pilus structure and bacterial pathogenicity. *Nat. Rev. Microbiol.* 2, 363–378. doi: 10.1038/nrmicro885

Cullen, T. W., Madsen, J. A., Ivanov, P. L., Brodbelt, J. S., and Trent, M. S. (2012). Characterization of unique modification of flagellar rod protein FlgG by *Campylobacter jejuni* lipid A phosphoethanolamine transferase, linking bacterial locomotion and antimicrobial peptide resistance. *J. Biol. Chem.* 287, 3326–3336. doi: 10.1074/jbc.m111.321737

Dangl, J. L. (2013). *Bacterial Pathogenesis of Plants and Animals: Molecular and Cellular Mechanisms*. Berlin: Springer.

Deatherage, B. L., Lara, J. C., Bergsbaken, T., Barrett, S. L. R., Lara, S., and Cookson, B. T. (2009). Biogenesis of bacterial membrane vesicles. *Mol. Microbiol.* 72, 1395–1407.

Duan, Q., Zhou, M., Zhu, L., and Zhu, G. (2013). Flagella and bacterial pathogenicity. *J. Basic Microbiol.* 53, 1–8. doi: 10.1002/jobm.201100335

Dubey, G. P., and Ben-Yehuda, S. (2011). Intercellular nanotubes mediate bacterial communication. *Cell* 144, 590–600. doi: 10.1016/j.cell.2011.01.015

Dubey, G. P., Mohan, G. B. M., Dubrovsky, A., Amen, T., Tsipshtein, S., Rouvinski, A., et al. (2016). Architecture and characteristics of bacterial nanotubes. *Dev. Cell* 36, 453–461. doi: 10.1016/j.devcel.2016.01.013

Dworkin, M. (1999). Fibrils as extracellular appendages of bacteria: their role in contact-mediated cell–cell interactions in *Myxococcus xanthus*. *Bioessays* 21, 590–595. doi: 10.1002/(sici)1521-1878(199907)21:7<590::aid-bies7>3.0.co;2-e

Easterbrook, K., and Coombs, R. (1976). Spinin: the subunit protein of bacterial spinae. *Can. J. Microbiol.* 22, 438–440. doi: 10.1139/m76-067

Easterbrook, K., and Sperker, S. (1982). Physiological controls of bacterial spinae production in complex medium and their value as indicators of spina function. *Can. J. Microbiol.* 28, 130–136. doi: 10.1139/m82-014

Edwards, S. J., and Kjellerup, B. V. (2013). Applications of biofilms in bioremediation and biotransformation of persistent organic pollutants, pharmaceuticals/personal care products, and heavy metals. *Appl. Microbiol. Biotechnol.* 97, 9909–9921. doi: 10.1007/s00253-013-5216-z

Fischer, T., Schorb, M., Reintjes, G., Kolovou, A., Santarella-Mellwig, R., Markert, S., et al. (2019). Biopearling of interconnected outer membrane vesicle chains by a marine flavobacterium. *Appl. Environ. Microbiol.* 85:e829-19.

Fronzes, R., Remaut, H., and Waksman, G. (2008). Architectures and biogenesis of non-flagellar protein appendages in Gram-negative bacteria. *EMBO J.* 27, 2271–2280. doi: 10.1038/emboj.2008.155

Fu, G., Bandaria, J. N., Le Gall, A. V., Fan, X., Yildiz, A., Mignot, T., et al. (2018). MotAB-like machinery drives the movement of MreB filaments during bacterial gliding motility. *Proc. Natl. Acad. Sci. U.S.A.* 115, 2484–2489. doi: 10.1073/pnas.1716441115

Galkina, S. I., Romanova, J. M., Bragina, E. E., Tiganova, I. G., Stadnichuk, V. I., Alekseeva, N. V., et al. (2011). Membrane tubules attach *Salmonella* typhimurium to eukaryotic cells and bacteria. *FEMS Immunol. Med. Microbiol.* 61, 114–124. doi: 10.1111/j.1574-695x.2010.00754.x

Garcia, E. C., Perault, A. I., Marlatt, S. A., and Cotter, P. A. (2016). Interbacterial signaling via Burkholderia contact-dependent growth inhibition system proteins. *Proc. Natl. Acad. Sci. U.S.A.* 113, 8296–8301. doi: 10.1073/pnas.1606323113

García-Bayona, L., Guo, M. S., and Laub, M. T. (2017). Contact-dependent killing by *Caulobacter crescentus* via cell surface-associated, glycine zipper proteins. *Elife* 6:e24869.

Ginocchio, C. C., Olmsted, S. B., Wells, C. L., and Galán, J. E. (1994). Contact with epithelial cells induces the formation of surface appendages on *Salmonella* typhimurium. *Cell* 76, 717–724. doi: 10.1016/0092-8674(94)90510-x

Gordon, V. D., and Wang, L. (2019). Bacterial mechanosensing: the force will be with you, always. *J. Cell Sci.* 132, jcs227694. doi: 10.1242/jcs.227694

Haiko, J., and Westerlund-Wikström, B. (2013). The role of the bacterial flagellum in adhesion and virulence. *Biology* 2, 1242–1267. doi: 10.3390/biology2041242

Handley, P. S. (1990). Structure, composition and functions of surface structures on oral bacteria. *Biofouling* 2, 239–264. doi: 10.1080/08927019009378148

Harapanahalli, A. K., Younes, J. A., Allan, E., van der Mei, H. C., and Busscher, H. J. (2015). Chemical signals and mechanosensing in bacterial responses to their environment. *PLoS Pathog.* 11:e1005057. doi: 10.1371/journal.ppat.1005057

Häussler, S., and Becker, T. (2008). The *Pseudomonas* quinolone signal (PQS) balances life and death in *Pseudomonas aeruginosa* populations. *PLoS Pathog.* 4:e1000166. doi: 10.1371/journal.ppat.1000166

Hayes, C. S., Aoki, S. K., and Low, D. A. (2010). Bacterial contact-dependent delivery systems. *Annu. Rev. Genet.* 44, 71–90. doi: 10.1146/annurev.genet.42.110807.091449

He, S. Y. (1996). Elicitation of plant hypersensitive response by bacteria. *Plant Physiol.* 112:865. doi: 10.1104/pp.112.3.865

Ishii, S., Kosaka, T., Hori, K., Hotta, Y., and Watanabe, K. (2005). Coaggregation facilitates interspecies hydrogen transfer between Pelotomaculum thermopropionicum and Methanothermobacter thermautotrophicus. *Appl. Environ. Microbiol.* 71, 7838–7845. doi: 10.1128/aem.71.12.7838-7845.2005

Kaiser, D. (2004). Signaling in myxobacteria. *Annu. Rev. Microbiol.* 58, 75–98. doi: 10.1146/annurev.micro.58.030603.123620

Kaprelyants, A. S., and Kell, D. B. (1996). Do bacteria need to communicate with each other for growth? *Trends Microbiol.* 4, 237–242. doi: 10.1016/0966-842x(96)10035-4

Khan, S., and Scholey, J. M. (2018). Assembly, functions and evolution of archaella, flagella and cilia. *Curr. Biol.* 28, R278–R292.

Kim, K. W. (2017). Electron microscopic observations of prokaryotic surface appendages. *J. Microbiol.* 55, 919–926. doi: 10.1007/s12275-017-7369-4

Kimkes, T. E., and Heinemann, M. (2019). How bacteria recognise and respond to surface contact. *FEMS Microbiol. Rev.* 44, 106–122. doi: 10.1093/femsre/fuz029

Kimkes, T. E., and Heinemann, M. (2020). How bacteria recognise and respond to surface contact. *FEMS Microbiol. Rev.* 44, 106–122. doi: 10.1093/femsre/fuz029

Kline, K. A., Dodson, K. W., Caparon, M. G., and Hultgren, S. J. (2010). A tale of two pili: assembly and function of pili in bacteria. *Trends Microbiol.* 18, 224–232. doi: 10.1016/j.tim.2010.03.002

Kline, K. A., Fälker, S., Dahlberg, S., Normark, S., and Henriques-Normark, B. (2009). Bacterial adhesins in host-microbe interactions. *Cell Host Microbe* 5, 580–592. doi: 10.1016/j.chom.2009.05.011

Kuehn, M. J. (1997). Establishing communication via gram-negative bacterial pili. *Trends Microbiol.* 5, 130–132. doi: 10.1016/s0966-842x(96)30045-0

Kumar, C. G., and Anand, S. (1998). Significance of microbial biofilms in food industry: a review. *Int. J. Food Microbiol.* 42, 9–27. doi: 10.1016/s0168-1605(98)00060-9

Lele, P. P., Hosu, B. G., and Berg, H. C. (2013). Dynamics of mechanosensing in the bacterial flagellar motor. *Proc. Natl. Acad. Sci. U.S.A.* 110, 11839–11844. doi: 10.1073/pnas.1305885110

Li, G., Brown, P. J., Tang, J. X., Xu, J., Quardokus, E. M., Fuqua, C., et al. (2012). Surface contact stimulates the just-in-time deployment of bacterial adhesins. *Mol. Microbiol.* 83, 41–51. doi: 10.1111/j.1365-2958.2011.07909.x

Li, J., Azam, F., and Zhang, S. (2016). Outer membrane vesicles containing signalling molecules and active hydrolytic enzymes released by a coral pathogen *Vibrio shilonii* AK1. *Environ. Microbiol.* 18, 3850–3866. doi: 10.1111/1462-2920.13344

Lukaszczyk, M., Pradhan, B., and Remaut, H. (2019). "The Biosynthesis and Structures of Bacterial Pili," in *Bacterial Cell Walls and Membranes*, ed. A. Kuhn (Berlin: Springer), 369–413. doi: 10.1007/978-3-030-18768-2_12

Marzo, L., Gousset, K., and Zurzolo, C. (2012). Multifaceted roles of tunneling nanotubes in intercellular communication. *Front. Physiol.* 3:72. doi: 10.3389/fphys.2012.00072

Mashburn-Warren, L. M., and Whiteley, M. (2006). Special delivery: vesicle trafficking in prokaryotes. *Mol. Microbiol.* 61, 839–846. doi: 10.1111/j.1365-2958.2006.05272.x

Matthysse, A. (1983). Role of bacterial cellulose fibrils in *Agrobacterium tumefaciens* infection. *J. Bacteriol.* 154, 906–915. doi: 10.1128/jb.154.2.906-915.1983

McBride, M. J. (2001). Bacterial gliding motility: multiple mechanisms for cell movement over surfaces. *Annu. Rev. Microbiol.* 55, 49–75. doi: 10.1146/annurev.micro.55.1.49

Miller, M. B., and Bassler, B. L. (2001). Quorum sensing in bacteria. *Annu. Rev. Microbiol.* 55, 165–199.

Minamino, T. (2014). Protein export through the bacterial flagellar type III export pathway. *Biochim. Biophys. Acta (BBA)-Mol. Cell Res.* 1843, 1642–1648. doi: 10.1016/j.bbamcr.2013.09.005

Miyata, M., Yamamoto, H., Shimizu, T., Uenoyama, A., Citti, C., and Rosengarten, R. (2000). Gliding mutants of Mycoplasma mobile: relationships between motility and cell morphology, cell adhesion and microcolony formation. *Microbiology* 146, 1311–1320. doi: 10.1099/00221287-146-6-1311

Moens, S., and Vanderleyden, J. (1996). Functions of bacterial flagella. *Crit. Rev. Microbiol* 22, 67–100. doi: 10.3109/10408419609106456

Mukherjee, S., and Bassler, B. L. (2019). Bacterial quorum sensing in complex and dynamically changing environments. *Nat. Rev. Microbiol.* 17, 371–382. doi: 10.1038/s41579-019-0186-5

Müller, K., Collinson, S. K., and Kay, W. (1991). Type 1 fimbriae of *Salmonella* enteritidis. *J. Bacteriol.* 173, 4765–4772. doi: 10.1128/jb.173.15.4765-4772.1991

Neuhaus, A., Selvaraj, M., Salzer, R., Langer, J. D., Kruse, K., Kirchner, L., et al. (2020). Cryo-electron microscopy reveals two distinct type IV pili assembled by the same bacterium. *Nat. Commun.* 11, 1–13.

Nicastro, L., and Tükel, Ç (2019). Bacterial amyloids: the link between bacterial infections and autoimmunity. *Trends Microbiol.* 27, 954–963. doi: 10.1016/j.tim.2019.07.002

Njoroge, J., and Sperandio, V. (2009). Jamming bacterial communication: new approaches for the treatment of infectious diseases. *EMBO Mol. Med.* 1, 201–210. doi: 10.1002/emmm.200900032

Nwodo, U. U., Green, E., and Okoh, A. I. (2012). Bacterial exopolysaccharides: functionality and prospects. *Int. J. Mol. Sci.* 13, 14002–14015. doi: 10.3390/ijms131114002

Ortega-Morales, B. O., Chan-Bacab, M. J., De la Rosa, S. D. C., and Camacho-Chab, J. C. (2010). Valuable processes and products from marine intertidal microbial communities. *Curr. Opin. Biotechnol.* 21, 346–352. doi: 10.1016/j.copbio.2010.02.007

Otto, M. (2019). Staphylococcal biofilms. *Curr. Top. Microbiol. Immunol.* 322, 207–228. doi: 10.1007/978-3-540-75418-3_10

Pande, S., Shitut, S., Freund, L., Westermann, M., Bertels, F., Colesie, C., et al. (2015). Metabolic cross-feeding via intercellular nanotubes among bacteria. *Nat. Commun.* 6, 1–13.

Papenfort, K., and Bassler, B. L. (2016). Quorum sensing signal–response systems in Gram-negative bacteria. *Nat. Rev. Microbiol.* 14:576. doi: 10.1038/nrmicro.2016.89

Penn, C. W., and Luke, C. J. (1992). Bacterial flagellar diversity and significance in pathogenesis. *FEMS Microbiol. Lett.* 100, 331–336. doi: 10.1111/j.1574-6968.1992.tb05723.x

Pérez-Cruz, C., Delgado, L., López-Iglesias, C., and Mercade, E. (2015). Outer-inner membrane vesicles naturally secreted by gram-negative pathogenic bacteria. *PLoS ONE* 10:e0116896. doi: 10.1371/journal.pone.0116896

Piepenbrink, K. H. (2019). DNA uptake by type IV filaments. *Front. Mol. Biosci.* 6:1. doi: 10.3389/fmolb.2019.00001

Pönisch, W., Weber, C. A., Juckeland, G., Biais, N., and Zaburdaev, V. (2017). Multiscale modeling of bacterial colonies: how pili mediate the dynamics of single cells and cellular aggregates. *New J. Phys.* 19:015003. doi: 10.1088/1367-2630/aa5483

Proft, T., and Baker, E. (2009). Pili in Gram-negative and Gram-positive bacteria—structure, assembly and their role in disease. *Cell. Mol. Life Sci.* 66:613. doi: 10.1007/s00018-008-8477-4

Reguera, G., Pollina, R. B., Nicoll, J. S., and Lovley, D. R. (2007). Possible nonconductive role of Geobacter sulfurreducens pilus nanowires in biofilm formation. *J. Bacteriol.* 189, 2125–2127. doi: 10.1128/jb.01284-06

Rehman, T., Yin, L., Latif, M. B., Chen, J., Wang, K., Geng, Y., et al. (2019). Adhesive mechanism of different *Salmonella* fimbrial adhesins. *Microb. Pathog.* 37:103748. doi: 10.1016/j.micpath.2019.103748

Rodesney, C. A., Roman, B., Dhamani, N., Cooley, B. J., Katira, P., Touhami, A., et al. (2017). Mechanosensing of shear by *Pseudomonas aeruginosa* leads to increased levels of the cyclic-di-GMP signal initiating biofilm development. *Proc. Natl. Acad. Sci. U.S.A.* 114, 5906–5911. doi: 10.1073/pnas.1703255114

Roine, E., Wei, W., Yuan, J., Nurmiaho-Lassila, E.-L., Kalkkinen, N., Romantschuk, M., et al. (1997). Hrp pilus: an hrp-dependent bacterial surface appendage produced by *Pseudomonas* syringae pv. *tomato DC*3000. *Proc. Natl. Acad. Sci.* 94, 3459–3464. doi: 10.1073/pnas.94.7.3459

Römling, U., Bian, Z., Hammar, M., Sierralta, W. D., and Normark, S. (1998). Curli fibers are highly conserved between *Salmonella* typhimurium and *Escherichia coli* with respect to operon structure and regulation. *J. Bacteriol.* 180, 722–731. doi: 10.1128/jb.180.3.722-731.1998

Roussin, M., Rabarioelina, S., Cluzeau, L., Cayron, J., Lesterlin, C., Salcedo, S. P., et al. (2019). Identification of a contact-dependent growth Inhibition (CDI) system that reduces biofilm formation and host cell adhesion of *Acinetobacter baumannii* DSM30011 strain. *Front. Microbiol.* 10:2450. doi: 10.3389/fmicb.2019.02450

Russell, A. B., Peterson, S. B., and Mougous, J. D. (2014). Type VI secretion system effectors: poisons with a purpose. *Nat. Rev. Microbiol.* 12, 137–148. doi: 10.1038/nrmicro3185

Sauer, F. G., Mulvey, M. A., Schilling, J. D., Martinez, J. J., and Hultgren, S. J. (2000). Bacterial pili: molecular mechanisms of pathogenesis. *Curr. Opin. Microbiol.* 3, 65–72. doi: 10.1016/s1369-5274(99)00053-3

Schertzer, J. W., and Whiteley, M. (2011). Microbial communication superhighways. *Cell* 144, 469–470. doi: 10.1016/j.cell.2011.02.001

Shih, Y.-L., and Rothfield, L. (2006). The bacterial cytoskeleton. *Microbiol. Mol. Biol. Rev.* 70, 729–754.

Shimoyama, T., Kato, S., Ishii, S. I., and Watanabe, K. (2009). Flagellum mediates symbiosis. *Science* 323, 1574–1574. doi: 10.1126/science.1170086

Singh, S., and Verma, N. (2015). Fabrication of Ni nanoparticles-dispersed carbon micro-nanofibers as the electrodes of a microbial fuel cell for bio-energy production. *Int. J. Hydrogen Energy* 40, 1145–1153. doi: 10.1016/j.ijhydene.2014.11.073

Stacy, A. R., Diggle, S. P., and Whiteley, M. (2012). Rules of engagement: defining bacterial communication. *Curr. Opin. Microbiol.* 15, 155–161. doi: 10.1016/j.mib.2011.11.007

Stones, D. H., and Krachler, A. M. (2016). Against the tide: the role of bacterial adhesion in host colonization. *Biochem. Soc. Trans.* 44, 1571–1580. doi: 10.1042/bst20160186

Subramanian, P., Pirbadian, S., El-Naggar, M. Y., and Jensen, G. J. (2018). Ultrastructure of Shewanella oneidensis MR-1 nanowires revealed by electron cryotomography. *Proc. Natl. Acad. Sci. U.S.A.* 115, E3246–E3255.

Subramanian, S., and Kearns, D. B. (2019). Functional regulators of bacterial flagella. *Annu. Rev. Microbiol.* 73, 225–246. doi: 10.1146/annurev-micro-020518-115725

Taglialegna, A., Matilla-Cuenca, L., Dorado-Morales, P., Navarro, S., Ventura, S., Garnett, J. A., et al. (2020). The biofilm-associated surface protein Esp of *Enterococcus faecalis* forms amyloid-like fibers. *NPJ Biofilms Microbiomes* 6, 1–12.

Telford, J. L., Barocchi, M. A., Margarit, I., Rappuoli, R., and Grandi, G. (2006). Pili in gram-positive pathogens. *Nat. Rev. Microbiol.* 4:509.

Toyofuku, M. (2019). Bacterial communication through membrane vesicles. *Biosci. Biotechnol. Biochem.* 83, 1599–1605. doi: 10.1080/09168451.2019.1608809

Toyofuku, M., Morinaga, K., Hashimoto, Y., Uhl, J., Shimamura, H., Inaba, H., et al. (2017). Membrane vesicle-mediated bacterial communication. *ISME J.* 11, 1504–1509. doi: 10.1038/ismej.2017.13

Toyofuku, M., Nomura, N., and Eberl, L. (2019). Types and origins of bacterial membrane vesicles. *Nat. Rev. Microbiol.* 17, 13–24. doi: 10.1038/s41579-018-0112-2

Turnbull, L., Toyofuku, M., Hynen, A. L., Kurosawa, M., Pessi, G., Petty, N. K., et al. (2016). Explosive cell lysis as a mechanism for the biogenesis of bacterial membrane vesicles and biofilms. *Nat. Commun.* 7, 1–13.

Van Gerven, N., Klein, R. D., Hultgren, S. J., and Remaut, H. (2015). Bacterial amyloid formation: structural insights into curli biogensis. *Trends Microbiol.* 23, 693–706. doi: 10.1016/j.tim.2015.07.010

Van Houdt, R., and Michiels, C. W. (2005). Role of bacterial cell surface structures in *Escherichia coli* biofilm formation. *Res. Microbiol.* 156, 626–633. doi: 10.1016/j.resmic.2005.02.005

Vestby, L. K., Grønseth, T., Simm, R., and Nesse, L. L. (2020). Bacterial biofilm and its role in the pathogenesis of disease. *Antibiotics* 9:59. doi: 10.3390/antibiotics9020059

Visick, K. L., and Fuqua, C. (205). Decoding microbial chatter: cell-cell communication in bacteria. *J. Bacteriol.* 187, 5507–5519. doi: 10.1128/jb.187.16.5507-5519.2005

Vu, B., Chen, M., Crawford, R. J., and Ivanova, E. P. (2009). Bacterial extracellular polysaccharides involved in biofilm formation. *Molecules* 14, 2535–2554. doi: 10.3390/molecules14072535

Waidner, B., Specht, M., Dempwolff, F., Haeberer, K., Schaetzle, S., Speth, V., et al. (2009). A novel system of cytoskeletal elements in the human pathogen *Helicobacter* pylori. *PLoS Pathog.* 5:e1000669. doi: 10.1371/journal.ppat.1000669

Waters, C. M., and Bassler, B. L. (2005). Quorum sensing: cell-to-cell communication in bacteria. *Annu. Rev. Cell Dev. Biol.* 21, 319–346. doi: 10.1146/annurev.cellbio.21.012704.131001

Wei, X., Vassallo, C. N., Pathak, D. T., and Wall, D. (2014). Myxobacteria produce outer membrane-enclosed tubes in unstructured environments. *J. Bacteriol.* 196, 1807–1814. doi: 10.1128/jb.00850-13

Wolfson, E. B., Elvidge, J., Tahoun, A., Gillespie, T., Mantell, J., McAteer, S. P., et al. (2020). Bacterial flagella disrupt host cell membranes and interact with cytoskeletal components. *BioRxiv [Preprint]* doi: 10.1101/2020.02.12.945204

Wood, T. K., Barrios, A. F. G., Herzberg, M., and Lee, J. (2006). Motility influences biofilm architecture in *Escherichia coli*. *Appl. Microbiol. Biotechnol.* 72, 361–367. doi: 10.1007/s00253-005-0263-8

Xiong, L., Cooper, R., and Tsimring, L. S. (2018). Coexistence and pattern formation in bacterial mixtures with contact-dependent killing. *Biophys. J.* 114, 1741–1750. doi: 10.1016/j.bpj.2018.02.012

Yang, D. C., Blair, K. M., and Salama, N. R. (2016). Staying in shape: the impact of cell shape on bacterial survival in diverse environments. *Microbiol. Mol. Biol. Rev.* 80, 187–203. doi: 10.1128/mmbr.00031-15

Yoshihara, A., Nobuhira, N., Narahara, H., Toyoda, S., Tokumoto, H., Konishi, Y., et al. (2015). Estimation of the adhesive force distribution for the flagellar adhesion of *Escherichia coli* on a glass surface. *Colloids Surf. B Biointerf.* 131, 67–72. doi: 10.1016/j.colsurfb.2015.04.038

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Singhi and Srivastava. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 21 July 2020
doi: 10.3389/fmolb.2020.00140

Bacillus subtilis PcrA Couples DNA Replication, Transcription, Recombination and Segregation

María Moreno-del Alamo[1], Rubén Torres[1], Candela Manfredi[1†], José A. Ruiz-Masó[2], Gloria del Solar[2] and Juan Carlos Alonso[1]*

[1] Department of Microbial Biotechnology, Centro Nacional de Biotecnología, CNB-CSIC, Madrid, Spain, [2] Centro de Investigaciones Biológicas Margarita Salas, CIB-CSIC, Madrid, Spain

Bacillus subtilis PcrA abrogates replication-transcription conflicts *in vivo* and disrupts RecA nucleoprotein filaments *in vitro*. Inactivation of *pcrA* is lethal. We show that PcrA depletion lethality is suppressed by *recJ* (involved in end resection), *recA* (the recombinase), or *mfd* (transcription-coupled repair) inactivation, but not by inactivating end resection (*addAB* or *recQ*), positive and negative RecA modulators (*rarA* or *recX* and *recU*), or genes involved in the reactivation of a stalled RNA polymerase (*recD2*, *helD*, *hepA,* and *ywqA*). We also report that *B. subtilis* mutations previously designated as *recL16* actually map to the *recO* locus, and confirm that PcrA depletion lethality is suppressed by *recO* inactivation. The *pcrA* gene is epistatic to *recA* or *mfd*, but it is not epistatic to *addAB*, *recJ*, *recQ*, *recO16*, *rarA*, *recX*, *recU*, *recD2*, *helD*, *hepA,* or *ywqA* in response to DNA damage. PcrA depletion led to the accumulation of unsegregated chromosomes, and this defect is increased by *recQ*, *rarA,* or *recU* inactivation. We propose that PcrA, which is crucial to maintain cell viability, is involved in different DNA transactions.

Keywords: replication fork stalling, RNA polymerase backtracking, replication-transcription conflict, RecL16, Rep, UvrD

OPEN ACCESS

Edited by:
Chew Chieng Yeo, Sultan Zainal Abidin University, Malaysia

Reviewed by:
Arijit Dutta, The University of Texas Health Science Center at San Antonio, United States
Harshad Ghodke, University of Wollongong, Australia

****Correspondence:***
Juan Carlos Alonso
jcalonso@cnb.csic.es

†Present address:
Candela Manfredi, Dept of Pediatrics, Emory University School of Medicine, Atlanta, GA, United States

Specialty section:
This article was submitted to Molecular Recognition, a section of the journal Frontiers in Molecular Biosciences

Received: *28 April 2020*
Accepted: *10 June 2020*
Published: *21 July 2020*

Citation:
Moreno-del Alamo M, Torres R, Manfredi C, Ruiz-Masó JA, del Solar G and Alonso JC (2020) Bacillus subtilis PcrA Couples DNA Replication, Transcription, Recombination and Segregation. Front. Mol. Biosci. 7:140. doi: 10.3389/fmolb.2020.00140

INTRODUCTION

Homologous recombination is the major pathway to circumvent a replicative stress, a replication fork collapse and for the elimination of DNA double-strand breaks (DSBs) induced by endogenous or exogenous stress. Super-family 1 (SF1) DNA helicases, which are conserved motor proteins that couple nucleoside triphosphate hydrolysis to the unwinding of duplex DNA, play crucial roles in repair-by-recombination and in coping with replication-transcription conflicts (RTCs) (Wu and Hickson, 2006; Singleton et al., 2007). The prototype of bacterial SF1 helicases that translocate with 3′ → 5′ direction is UvrD (Singleton et al., 2007; Dillingham, 2011). This enzyme shares a significant degree of structural similarity with Rep, which is restricted to the γ-Proteobacteria Class, PcrA and yeast Srs2 DNA helicases (Wu and Hickson, 2006; Marini and Krejci, 2010). Rep and PcrA play essential roles in the replication of extrachromosomal elements, whereas UvrD and PcrA participate in the resolution of RTCs by poorly understood mechanisms (Boubakri et al., 2010; Bruning et al., 2014; Epshtein et al., 2014; Merrikh et al., 2015). *In vitro* studies reveal that PcrA and UvrD interact with the RNA polymerase (RNAP), Rep interacts with the replicative DNA helicase (DnaB in Proteobacteria) and Srs2 physically interacts with Rad51 and with the PCNA sliding clamp (ortholog of

bacterial DnaN) among other proteins (Antony et al., 2009; Guy et al., 2009; Kaniecki et al., 2017; Sanders et al., 2017). Absence of *Escherichia coli* Rep and UvrD renders cells inviable when grown in rich medium (Taucher-Scholtz et al., 1983), but lack of *Bacillus subtilis* PcrA renders cells inviable even when grown in minimal medium (Petit et al., 1998; Merrikh et al., 2015). To gain insight into the crucial steps carried on by PcrA, a comparative analysis with UvrD and Rep was undertaken. From the comparative analysis of *E. coli* (best-characterized representative of the Proteobacteria Phylum) and *B. subtilis* (best-characterized from the Firmicutes Phylum), which are evolutionarily separated by more than 2,000 million years, a genetic divergence larger than that between human and paramecium, we expect to understand the role of the SF1 UvrD-like DNA helicases.

When the DNA of a single inert mature haploid non-replicating chromosome of a *B. subtilis* spore is damaged, the RTCs of the single spore genome are compounded during the rapid outgrowth. In the absence of end resection (AddAB and/or RecJ-RecQ [RecS]) cells remain recombination proficient and apparently are as capable of repairing damaged template bases as the wild type (*wt*) control (Vlasic et al., 2014). These spores, which lack an intact homologous template, require RecA mediators (RecO, RecR), RecA itself, and positive (RecF) and negative (RecX, RecU) RecA modulators, with PcrA facilitating DNA replication through transcription units (Vlasic et al., 2014; Merrikh et al., 2015; Raguse et al., 2017). It is likely that PcrA is implicated in the processing of damaged replication forks and/or recombination intermediates formed at damaged forks and in circumventing DNA lesions in the absence of an intact homologous template, without generating a fork breakage that should be lethal for the revival of a haploid spore (Raguse et al., 2017). Then, a nucleoprotein RecA filament aids to overcome RTCs *via* different DNA damage tolerance pathways and to reactivate replication by recruiting the damage checkpoint DisA and pre-primosome DnaD proteins (Million-Weaver et al., 2015; Torres et al., 2019). This is consistent with the observation that the lethality of *B. subtilis* Δ*pcrA* cells is suppressed by *recF*17, *recL*16, Δ*recO,* or Δ*recR* mutations (Petit and Ehrlich, 2002).

In *E. coli* cells, the synthetic lethality of Δ*uvrD* Δ*rep* is partially suppressed by *recJ*, *recQ*, *recO*, *recR,* or *recF* inactivation in minimal medium, but it is only marginally suppressed by *recA* inactivation (Lestini and Michel, 2008; Guy et al., 2009). Inactivation of *recJ* or *recQ* provides very limited suppression of Δ*rep* Δ*uvrD* rich medium lethality, but *recA* inactivation does not (Petit and Ehrlich, 2002; Veaute et al., 2005; Lestini and Michel, 2008; Guy et al., 2009). *in vivo* assays reveal that PcrA expression in *E. coli* can substitute several functions of UvrD, but antagonizes the function of Rep, providing a heterologous dominant negative phenotype (Petit et al., 1998). These data suggest that in the absence of UvrD and Rep, toxic RecA nucleoprotein filaments and/or RecA-mediated recombination intermediates or DNA structures can accumulate. This hypothesis is supported by the following observations: (i) in budding yeast, suppressors of Srs2 mutations map in the Rad51 gene (Aboussekhra et al., 1992); and (ii) RecA- or Rad51-mediated DNA strand exchange is actively prevented by the PcrA, UvrD, or Srs2 DNA helicase *in vitro*, considered as a paradigmatic anti-recombinase activity (Krejci et al., 2003; Veaute et al., 2003, 2005; Anand et al., 2007; Park et al., 2010; Fagerburg et al., 2012; Petrova et al., 2015; Kaniecki et al., 2017).

Other studies have demonstrated that the rich-medium synthetic lethality of Δ*uvrD* Δ*rep* cells is caused primarily by RTCs, with partial reduction of transcription or translation rates across heavily transcribed genes in the opposite orientation relative to the replication forks, to compensate the impact of transcription on DNA replication (Guy et al., 2009; Baharoglu et al., 2010; Kamarthapu et al., 2016; Myka et al., 2017). Indeed, rich-medium synthetic lethality of Δ*uvrD* Δ*rep* cells is fully suppressed by reducing transcription, as mutations in different RNAP subunits (rpo* [*rpoB* and *rpoC* point mutants]), or reducing translation elongation, such as by mutations in a tRNA gene (AspRS, *aspT*), in an aminoacyl tRNA synthetase, in a translation factor needed for efficient formation of proline-proline bonds (EF-P), and *spoT*1 mutation [encoding a (p)ppGpp pyrophosphorylase-defective SpoT] (Guy et al., 2009; Baharoglu et al., 2010; Kamarthapu et al., 2016; Myka et al., 2017). Transcription and translation are coupled with the leading ribosome pushing RNAP forward; however, when these two processes become uncoupled, RNAP *rpo** mutants are prone to pausing with (p)ppGpp promoting UvrD-mediated RNAP backtracking (Kamarthapu et al., 2016; Myka et al., 2017). These observations altogether open the question about the primary cause of PcrA lethality. (Unless stated otherwise, indicated genes and products are of *B. subtilis* origin. The nomenclature used to denote the origin of proteins from other bacteria is based on the bacterial genus and species abbreviation [*e.g., E. coli* UvrD is referred to as $UvrD_{Eco}$]).

The physiological causes of PcrA/UvrD-Rep_{Eco} lethality is/are poorly understood. To understand the role of PcrA in rich medium exponentially growing *B. subtilis* cells, we have studied the genetic linkage of PcrA depletion (Merrikh et al., 2015) with mutations in genes acting at the presynaptic (Δ*recJ*, Δ*recQ,* Δ*addAB*, *recL*16, Δ*rarA*, Δ*recX*, Δ*recU*) and synaptic (Δ*recA*) stages, as well as in genes that contribute to bypass RTCs or facilitate RNAP backtracking or removal (Δ*helD,* Δ*recD*2 [absent in *E. coli*], Δ*hepA* [also termed *yqhH*or *rapA*], Δ*ywqA* or Δ*mfd*). We show that PcrA depletion reduced cell viability by >4,000-fold, and survival in the presence of limiting H_2O_2 or methyl methanesulfonate (MMS) concentrations, suggesting that PcrA is involved in repair-by-recombination. The *recL*16 mutations were mapped in the *recO* gene and the mutation was termed *recO*16. The PcrA depletion lethality is suppressed by *recO16*, *recJ*, *recA,* or *mfd* inactivation, but not by *recQ*, *rarA*, *recX*, *recU*, *addAB*, *helD*, *hepA,* or *ywqA* inactivation. The *pcrA* gene is not epistatic to *recO16*, *recJ*, *rarA*, *recQ*, *recX*, *recU*, *recD*2, *addAB*, *helD*, *hepA,* or *ywqA* in response to MMS-induced DNA damage, but it is epistatic to *recA* or *mfd*. Absence of PcrA promotes a net accumulation of unsegregated nucleoids, and this defect is increased in the absence of RarA, RecQ, and RecU. We conclude that PcrA contributes to untangle branched intermediates and works at the interface of DNA replication, transcription, recombination and segregation.

MATERIALS AND METHODS

Bacterial Strains and Plasmids

All *B. subtilis* strains used derived from BG214, and are listed in **Table 1**. The gene to be characterized was deleted by gene replacement with the *six-cat-six* (SCS) cassette flanked by appropriate homologous regions up- and downstream. The SCS cassette, composed of two directly oriented β-recombinase cognate sites (*six* sites) and the *cat* gene, confers chloramphenicol resistance (Cm^R) (Rojo and Alonso, 1995). Natural competent cells were transformed with the SCS cassette flanked by homologous regions to the gene to be deleted with selection for Cm^R. Integration of the SCS cassette, through double crossover recombination, replaced the gene under characterization. The β recombinase promoter, which maps within the six site, can read the downstream gene to overcome any potential polar effect (Rojo et al., 1994). A plasmid-borne β-recombinase gene was moved into the background, and followed by β site-specific recombinase-mediated excision between the two directly oriented six sites, leading to the deletion of the *cat* gene and one six site (Rojo et al., 1994). The final outcome of this strategy is that the gene to be characterized is replaced by a single six site (Sanchez et al., 2005). Accuracy of deletions was confirmed by complementation with a plasmid-borne gene (Sanchez et al., 2005). The IPTG-inducible *sspB* cassette encoding the SspB adaptor protein (a gift from Houra and Christopher Merrikh, Vanderbilt University, USA) was ectopically integrated into the *amy* locus by natural transformation (Alonso et al., 1988; Torres et al., 2017). The null *hepA* or *ywqA* mutations (a gift from Marie-Agnès Petit, Université Paris-Saclay, France) were moved into the BG214 background (**Table 1**) by SPP1-mediated transduction (Valero-Rello et al., 2017).

The *ssrA* degradation tag fused to the 3′-end of the *pcrA* gene (*pcrA-ssrA*) replaces the *pcrA* gene (Merrikh et al., 2015; Torres et al., 2017). The *pcrA-ssrA* gene was moved by SPP1-mediated transduction into *recO*16 *sspB*, Δ*recJ sspB*, Δ*recD*2 *sspB*, Δ*recQ sspB*, Δ*recX sspB*, Δ*recU sspB*, Δ*rarA sspB*, Δ*addAB sspB*, Δ*helD sspB*, Δ*hepA sspB*, Δ*ywqA sspB,* or Δ*mfd sspB* strain as well as in the otherwise *wt* background (*rec*$^+$ *sspB* strain). The Δ*recA* mutation was mobilized by SPP1-mediated transduction into the *pcrA-ssrA sspB* ($pcrA_T$) context.

To segregate the *recO* mutations and to reconstruct a new BG107 strain (BG107-1), limiting concentrations of chromosomal DNA of the original BG107 (*recL*16) strain and 0.1 μg/ml of pHP14 DNA were added to competent BG214 cells, with selection for the plasmid-borne Cm^R marker. From those transformants it was expected that by co-transformation of any unlinked markers (congression) 0.1 to 1% of the Cm^R cells should receive the *recL*16 MMS sensitive (MMS^S) phenotype (Alonso et al., 1988). Selection of MMS^S was performed by streaking colonies on agar plate containing or lacking MMS (Alonso et al., 1988). Five selected MMS^S clones were sequenced by high-throughput sequence analyzer (Illumina) technology using standard sequencing libraries and filtered sequence data (Beijing Genomics Institute [BGI]), of ∼1 gigabases per sample, followed by whole-genome comparison with the nucleotide sequencing of the MMS^R parental BG214 strain.

TABLE 1 | *Bacillus subtilis* strains.

Strains[a]	Relevant genotype	Source/references
BG214	Wild type	Laboratory strain
BG107[b]	+ *recL*16 *hem*E *pks*L *dea*D	Alonso et al., 1987
BG107-1	+ *recO*16	This work
BG1873	+ Δ*recA*	This work
BG763	+ Δ*recJ*	Sanchez et al., 2006
BG705	+ Δ*recQ*	Sanchez et al., 2006
BG633	+ Δ*recU*	Fernández et al., 1998
BG1065	+ Δ*recX*	Cárdenas et al., 2012
BG551	+ Δ*helD*	Carrasco et al., 2001
BG1067	+ Δ*rarA*	Romero et al., 2019b
BG1887	+ Δ*mfd*	This work
BG1779	+ Δ*recD2*	Torres et al., 2017
BG1337	+ Δ*addAB*	Torres et al., 2017
BG1525	+ *pcrA-ssrA sspB*	Torres et al., 2017
BG1583	+ *pcrA-ssrA sspB* Δ*recD*2	Torres et al., 2017
BG1861	+ Δ*hepA* (also termed *yqhH*)	This work
BG1862	+ Δ*ywqA*	This work
BG1715	+ *pcrA-ssrA sspB recO*16	This work
BG1877	+ *pcrA-ssrA sspB* Δ*recA*	This work
BG1731	+ *pcrA-ssrA sspB* Δ*recJ*	This work
BG1713	+ *pcrA-ssrA sspB* Δ*recQ*	This work
BG1709	+ *pcrA-ssrA sspB* Δ*recX*	This work
BG1711	+ *pcrA-ssrA sspB* Δ*recU*	This work
BG1823	+ *pcrA-ssrA sspB* Δ*rarA*	This work
BG1869	+ *pcrA-ssrA sspB* Δ*addAB*	This work
BG1859	+ *pcrA-ssrA sspB* Δ*helD*	This work
BG1875	+ *pcrA-ssrA sspB* Δ*mfd*	This work
BG1839	+ *pcrA-ssrA sspB* Δ*hepA*	This work
BG1857	+ *pcrA-ssrA sspB* Δ*ywqA*	This work
Plasmid	Relevant genotype	Source/reference
pQE1-pcrA	AmpR, ori_{Eco}	This work
pHP14	CmR, EmR ori_{Eco} ori_{Bsu}	de la Hoz et al., 2000
pCB1133	pHP14 + *pcrA* K37A	This work
pCB1119	pHP14 + *pcrA* Q254A (E224V)	This work
pCB1225	pHP14 + *pcrA* T65I (pcrA3)	This work

[a]*All strains are derivatives of B. subtilis BG214 (trpCE metA5 amyE1 ytsJ1 rsbV37 xre1 xkdA1 att*^{SPβ} *att*^{ICEBs1}*) strain.* [b]*The phenotype of the original BG107 strain is defined elsewhere (C.M., M.C.G., M.M.C. and J.C.A.) (Supplementary Material,* ***Annex 1****).*

The K37A, T65I, or Q254A *pcrA* variants were generated by means of mutation site directed mutagenesis (QuickChange Kit, Stratagene) using the pQE-1-borne *wt pcrA* plasmid as a template. Unexpectedly, all the Q254A mutants analyzed also contained the unselected E224V mutation. K37A, T65I, or Q254A-E224V *pcrA* genes were amplified by PCR and were cloned into *Xma*I-*Bam*HI-cleaved pHP14. The resultant recombinant plasmids were used to transform competent *B. subtilis* BG214 cells.

Survival Assays and Colony Size

Plating exponentially growing *pcrA-ssrA sspB* ($pcrA_T$) cells in rich medium onto agar plates containing isopropyl-β-D

thiogalactopyranoside (IPTG) induced SspB expression from a regulated promoter, which then bound the SsrA peptide tag and rapidly delivered the tagged PcrA-SsrA protein to the *B. subtilis* ClpXP protease for degradation (PcrA degron [$pcrA_T$] strain) (Keiler et al., 1996; Griffith and Grossman, 2008; Merrikh et al., 2015). PcrA degron cultures were grown to $OD_{560} = 0.4$. The cultures were divided and aliquots plated in LB agar plates alone or with 500 μM IPTG (Calbiochem). The percentage of colony forming units (CFUs) in LB agar plates containing IPTG was measured. The mean and SEM were calculated using Prism 6 software (GraphPad), and a Student's *t*-test, with $P < 0.01$, was performed to denote the threshold of significance.

Cell sensitivity to chronic MMS (Sigma Aldrich) or H_2O_2 (Sigma Aldrich) exposure was determined by growing cultures to $OD_{560} = 0.4$ and plating appropriate dilutions on rich LB agar plates containing IPTG (500 μM) and MMS (1.3 mM) or H_2O_2 (0.2 mM) as described (Sánchez et al., 2007). Cells grew in rich LB medium with a doubling time of 28–35 min. Plates were incubated overnight (16–18 h, 37°C) and the number of CFUs determined. Experiments were conducted independently at least four times. Fractional survival data are shown as mean ± SEM. Statistical analysis was performed with a two-tailed Student's *t*-test. For experiments involving more than two groups, one-way analysis of variance (ANOVA) was performed. For all tests, a $P < 0.1$ was considered significant (**Supplementary Material**). All statistical analyses were performed using Prism 6 software.

B. subtilis cells form round smooth colonies that raised above the agar. Colony size on Petri dishes was calculated via the diameter of a hypothetical circular colony. After overnight incubation, Petri dishes pictures were acquired and analyzed with the aid of a BioRad ChemiDoc™ imaging system equipped with the QuantityOne software (BioRad). The relative mean colony diameter of ∼50 isolated blind scoring colonies from $pcrA_T$ vs. $pcrA_T$ cells bearing a second mutation were measured from the pictures using ImageJ software (NIH). Upon colony magnification, the relative mean colony size was calculated using the formula for the area of a circle. Average and standard deviation were calculated using Prism 6 software. The colony area was compared by analysis of variance or Student's *t*-test, with $P < 0.01$ as the threshold of significance.

Nucleotide Sequence Analysis

The samples of genomic DNA from *B. subtilis wt* (BG214, Reference strain) and the Test *recL16* (BG107) strains were analyzed using the first step of comparative genome sequencing, a service provided by NimbleGen Systems, Inc. as described earlier (Albert et al., 2005). Briefly, the genome of *B. subtilis wt* was tiled on custom-designed "mapping" microarrays with 29 base oligonucleotide pieces (probes) and a 7 or 8 base spacing from the start of one probe to the start of the next. The genomic DNA from the Reference BG214 strain was labeled with Cy5 and of the Test BG107 strain with Cy3. The ratios of hybridization intensities (reference/test) were calculated and plotted against genome position. If the test sample contains no mutation, the reference/test ratio equals 1. This analysis gives a high-resolution map of possible mutation sites, in which each mutation is localized in a window of 29 bases (the length of the reporting probe). A custom algorithm, based on hybridization intensity ratios significantly above background that have good agreement between results from corresponding probes from both strands, allows the identification of likely sites of mutation. The data were analyzed graphically using the SignalMap software provided by NimbleGen. The regions around the probes with a reference/test ratio significantly >1 were tested for mutations by direct DNA sequencing. Sequencing was performed with Big Dye (from Operon Technologies Inc.) using the protocols of the University of Wisconsin Biotechnology Center, after PCR amplification of the target genomic DNA sequence.

The samples of genomic DNA from *B. subtilis wt* (BG214, Reference strain) and five MMS^S clones from the newly constructed. Test strain were re-sequenced by high-throughput sequence analyzer (Illumina) technology using standard sequencing libraries and filtered sequence data, a service provided by BGI, of ∼1 gigabase per sample. The sequencing data were used to conduct paired-end nucleotide sequencing with the rec^+ BG214 reference strain and the five MMS^S clones from the newly constructed BG107-1 sample as described (Quail et al., 2008).

Fluorescence Microscopy and Data Analysis

For chromosome segregation analyses (**Figure 2**), cells were fixed and stained as described (Carrasco et al., 2004). To obtain exponentially growing cells, overnight cultures were inoculated in LB rich medium. The *recO*16, Δ*recA*, Δ*rarA*, Δ*recJ*, Δ*recQ*, Δ*recX,* or Δ*recU* mutants in the *pcrA-ssrA sspB* ($pcrA_T$) context (**Table 1**) were grown unperturbed in LB medium to $OD_{560} = 0.2$ (37°C). IPTG (500 μM) was added to half of the culture, and both cultures were incubated (60 min, 37°C). Then, cells were collected, subjected to fixation with 2% formaldehyde, and finally stained with 4′,6′-diamino-2-phenylindole (DAPI) (1 μg/ml). Samples were visualized and photographed by fluorescence microscopy with a Hamamatsu 3CCD Digital Camera C7780 coupled to a BX61 Olympus fluorescence microscope, equipped with an 100x immersion oil lens and a DAPI filter (U-MNU2).

The ImageJ software (NIH) was used to merge the phase contrast and DAPI-fluorescence images, which allowed us to distinguish the septum, and thus determine the filamentation event and was also used to determine the cell length. Blind scoring was performed on captured images as described (Carrasco et al., 2004).

RESULTS AND DISCUSSION

Experimental Rationale

The phenotypes associated with the absence of PcrA are exceedingly complex and reflect an involvement in several aspects of DNA metabolism, including DNA replication, transcription, RTCs, repair-by-recombination and chromosomal segregation. The PcrA depletion lethality was attributed to the accumulation of toxic RecA-mediated intermediates or SOS toxicity. In fact, inactivation of positive RecA mediators (*recO*, *recR*), modulator (*recF*) (**Table S1**) or an unknown function (*recL*) suppresses the Δ*pcrA* lethality of cells grown

in minimal medium (Petit and Ehrlich, 2002). In addition, a mutation in *recO*, *recR*, *recF,* or *recL* reduces and delays the SOS induction (Gassel and Alonso, 1989). However, the lethality in the Δ*pcrA recA*1 (a leaky *recA* mutation) challenged such hypothesis (Petit and Ehrlich, 2002). This observation was revisited and extended using *rec*-deficient strains (**Table 1**).

PcrA-like enzymes contribute to release stalled RTCs with the subsequent recruitment of repair factors or to modulate the re-initiation of DNA replication and transcription (Guy et al., 2009; Boubakri et al., 2010; Merrikh et al., 2015). Several *B. subtilis* translocases, namely the SF1 DNA helicases (as PcrA, HelD, RecD2) and the SF2 enzymes (Mfd, HepA [also termed YqhH] and YwqA) have been shown to interact with the RNAP (Selby and Sancar, 1993; Muzzin et al., 1998; Sukhodolets et al., 2001; Shaw et al., 2008; Yawn et al., 2009; Wiedermannova et al., 2014; Sanders et al., 2017; Le et al., 2018). The poorly characterized HepA and YwqA enzymes, which belong to the Swi2/Snf2 family of translocases, share ∼30% identity to $HepA_{Eco}$. These translocases actively process a RNAP backwards as in the case of a RNAP stalled elongation complex (backtracking) or remove RNAP from the DNA template to resolve RTCs (**Table S2**) (Muzzin et al., 1998; Sukhodolets et al., 2001; Shaw et al., 2008; Yawn et al., 2009; Wiedermannova et al., 2014; Sanders et al., 2017; Le et al., 2018). The contribution of these functions to the inviability of PcrA depletion will be tested.

Finally, the $UvrD_{Eco}$ or PcrA helicase/translocase plays a crucial role in nucleotide excision repair (Mendonca et al., 1993; Petit and Ehrlich, 2002; Epshtein, 2015), but only the former contributes to DNA mismatch repair (Lenhart et al., 2016). In this report, we show that PcrA also contributes to the repair of non-bulky lesions of oxidative nature generated upon exposure to MMS or H_2O_2 (**Figure 1**). Upon PcrA depletion, cells were exposed to limiting MMS or H_2O_2 concentrations, whose lesions are specifically removed by base excision repair (Sedgwick, 2004; Fu et al., 2012). If damaged template bases escape specialized repair, because they are in ssDNA regions, the offending lesion halts DNA polymerase (DNAP) or RNAP progression, and the lesion-containing gap is circumvented *via* damage avoidance pathways (template switching, fork reversal, translesion synthesis). When the damage is bypassed and the lesion is present in duplex DNA, it is removed by specialized pathways (Sedgwick, 2004; Fu et al., 2012).

PcrA ATP Binding Mutants Render Cells Unviable

$PcrA3_{Sau}$ (T61I), defective in ATPase and helicase activities, and the $PcrA_{Sau}$ ($PcrAH^-$) mutant variant (K33A Q250A), which lacks ATPase and helicase activities, can displace $RecA_{Eco}$ from ssDNA, and inhibit $RecA_{Eco}$-mediated DNA strand exchange *in vitro* (Anand et al., 2007), suggesting that ATP binding and hydrolysis and the DNA helicase activity of PcrA are dispensable for its anti-recombinase activity. To assess whether the ATPase and/or translocase activity of *B. subtilis* PcrA are essential for cell proliferation, plasmid-borne *pcrA* genes with T65I, K37A, or Q254A mutations, equivalent to *S. aureus* mutations, were constructed and used to replace the *wt pcrA* gene. The single *pcrA* Q254A mutant gene also contained the unselected E224V mutation. The crystal structure of $PcrA_{Bst}$ DNA helicase with ADP suggests that residue Q254 is in hydrogen bonding distance to the γ-phosphate group of ATP, the E224 and a water molecule at the nucleotide binding pocket (Subramanya et al., 1996; Dillingham et al., 1999). The Q254 residue directly contacts the γ-phosphate of ATP, and the E224 forms hydrogen bonds with two water molecules that are hydrogen bonded to γ-phosphate oxygen atoms (Dillingham et al., 1999). Residue E224 is in an ideal position for activating the nucleophilic water molecule during hydrolysis. Mutations in residue Q254 alter the coupling between ATPase and helicase activities, and a transient interaction of Q254 with the γ-phosphate of ATP modifies the protein DNA binding site. Same mutations in residue Q254 are toxic for *E. coli* cells (Dillingham et al., 1999), and the untargeted E224V mutation might overcome such toxicity. The plasmid-borne *pcrA* Q254A (E224V) gene was also tested.

Monomeric plasmid DNA is inactive for transformation, but if it shares homology with recipient is activated upon interaction with the homologous region in the chromosome *via* RecA-mediated gene conversion, and the information present in the plasmid is transferred to the host chromosome [see Canosi et al. (1981)]. Monomeric DNA of plasmid-borne *pcrA* T65I (pCB1225), *pcrA* K37A (pCB1133), or *pcrA* Q254A (E224V) (pCB1119) was used to transform *B. subtilis* BG214 cells with selection for the plasmid marker (**Table 1**). Transformants carrying the established low copy unstable pCB1225, pCB1133, or pCB1119 plasmids were grown in the absence of selective pressure, and the plasmid-less segregants were subjected to nucleotide sequence analysis. We confirmed that in all transformants that lost pCB1225, the *pcrA* T65I gene had replaced the chromosomal *wt pcrA* gene by gene conversion (Petit et al., 1998). In only ∼20% of the transformants that lost pCB1119, *pcrA* Q254A (E224V) mutant gene had replaced the *wt* gene, but all sequenced *pcrA* Q254A clones also contained the unselected E224V mutation. However, in all transformants that loss pCB1133, nucleotide sequence analyses revealed the presence of the *wt* gene, thus we have failed in the attempt to observe the replacement of the *wt* gene by the *pcrA* K37A mutant gene, suggesting that the K37A mutation, which impairs ATP binding, renders cells non-viable. This failure, however, could also be attributed to different reasons, for example that the mutation indirectly impacts in the expression of the downstream essential *ligA* gene (http://www.subtiwiki.uni-goettingen.de). Since the $PcrAH^-_{Sau}$ mutant variant can promote disassembly of heterologous $RecA_{Eco}$ from ssDNA (Anand et al., 2007), but $PcrAH^-$ fails to overcome RTCs, and depletion of *wt* PcrA in a background expressing $PcrAH^-$ renders cell inviable (Merrikh et al., 2015), we have dropped the plasmid segregation approach and moved our analysis to the condition in which PcrA is selectively depleted by the *pcrA-ssrA sspB* degron ($pcrA_T$) strain (Merrikh et al., 2015) to answer these puzzling observations.

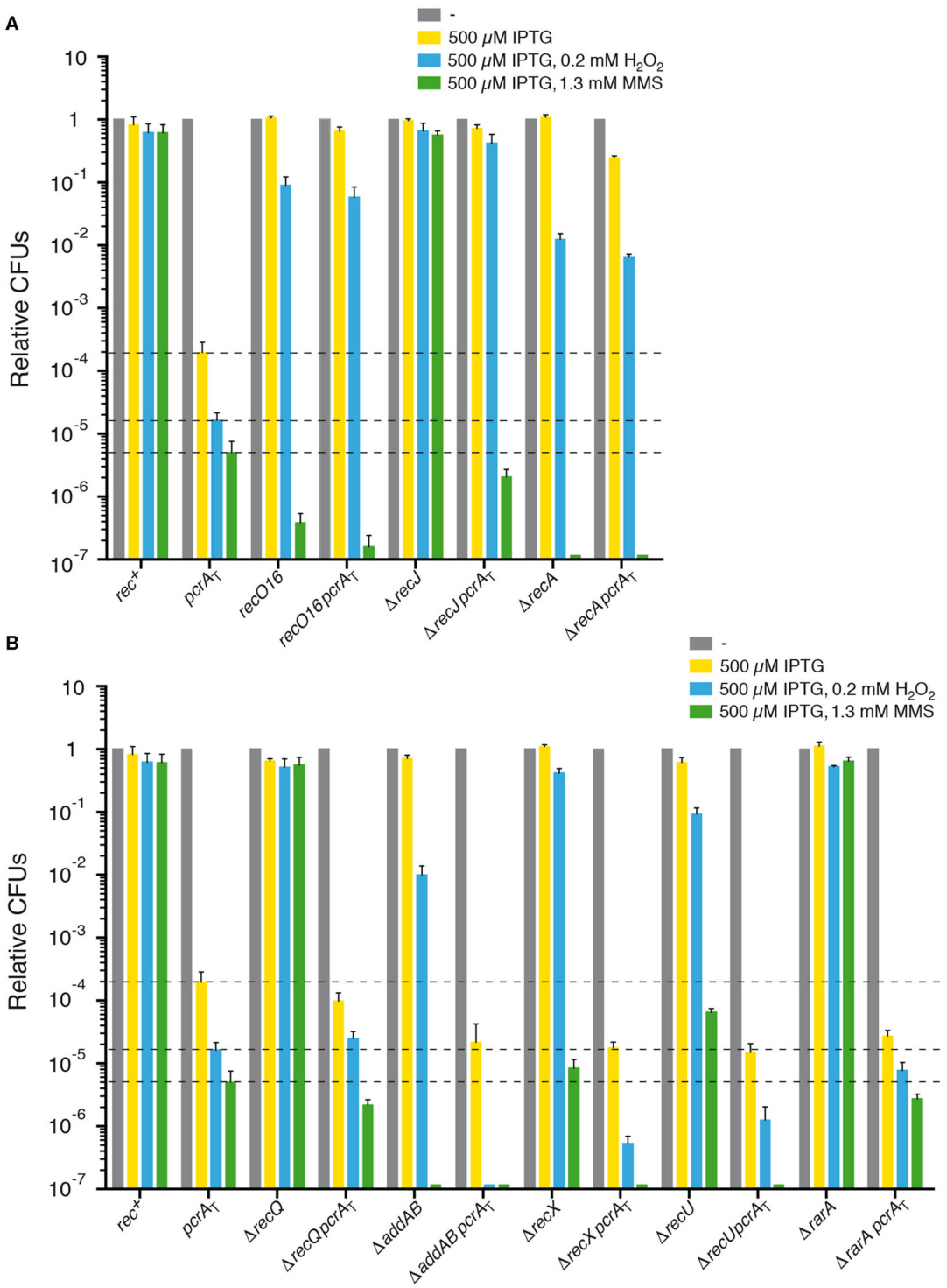


FIGURE 1 | PcrA lethality is suppressed by *recJ*, *recO,* or *recA* inactivation, but not by *addAB*, *recQ*, *rarA*, *recX,* or *recU* inactivation. Log phase cultures of *wt*, single and double mutant strains were diluted and plated on LB agar containing 500 μM IPTG (yellow bars) or lacking it (gray bars). Lethality assays showing cell viability upon PcrA depletion in *recJ*, *recO,* or *recA* cells **(A)** and in *addAB*, *recQ*, *rarA*, *recX,* or *recU* cells **(B)**. Log phase cultures of indicated strains were diluted and plated on LB agar containing IPTG and 0.2 mM H_2O_2 (blue bars) or IPTG and 1.3 mM MMS (green bars). Experiments were performed at least four times. The dotted lines mark the survival rate upon PcrA depletion. Data are shown as mean fractional survival ± SEM.

PcrA Depletion Inviability Requires RecJ and RecO, but Not RecQ, RarA, RecU, RecX, or AddAB

First, to confirm the reduction of cell survival following PcrA depletion, the *pcrA-ssrA sspB* degron ($pcrA_T$) strain (**Table 1**) was exponentially grown to an $OD_{560} = 0.4$ (at 37° C) in LB medium and then, appropriate dilutions were plated on LB agar plates lacking or containing 500 μM IPTG (Materials and methods). In the absence of IPTG, the viability of the *pcrA-ssrA sspB* ($pcrA_T$) degron strain was slightly compromised (1.3-fold) (**Figure 1A**, gray bar [- IPTG condition]) under the experimental conditions used (see below), perhaps due to noise introduced by *sspB gene* expression (see below). The plating efficiency of the $pcrA_T$ strain was reduced ∼5,000-fold upon plating onto LB agar plates containing IPTG when compared to the $pcrA^+$ control (rec^+) (**Figure 1A**, yellow bar [+ IPTG condition]). This is consistent with the earlier observation that $pcrA_T$ cell viability was reduced by >1,000-fold when plated onto 100 μM IPTG-containing plates (Merrikh et al., 2015). We have observed a linear decrease in the number of viable cells with increasing IPTG concentration, but it saturates above 500 μM IPTG (data not shown). Upon IPTG addition the expression of the SspB adaptor increases, and SspB interacts with the SsrA moiety of the PcrA-SsrA protein to deliver it to the ClpXP protease for PcrA degradation [see Keiler et al. (1996), Griffith and Grossman (2008)]. It is worth mentioning that no fitness cost to *B. subtilis* cells was observed at IPTG concentrations as high as 5 mM IPTG. After 15 min of IPTG addition to the $pcrA_T$ degron strain, 60 to 90% of PcrA is degraded (Merrikh et al., 2015).

To gain insight into PcrA contribution to repair-by-recombination, the $pcrA_T$ degron strain was exposed to limiting H_2O_2 (0.2 mM) or MMS (1.3 mM) concentrations. The survival of *pcrA-ssrA sspB* cells in the presence of 500 μM IPTG and 0.2 mM H_2O_2 (blue bar [+ IPTG and H_2O_2]) or 1.3 mM MMS (green bar [+ IPTG and MMS]) was significantly decreased (by ∼12- and ∼40-fold, respectively) when compared to the absence of H_2O_2 or MMS (**Figures 1A,B**), suggesting that depletion of PcrA renders cells sensitive to both DNA damaging agents.

To elucidate whether PcrA prevents unscheduled RecA loading or dismantles its cognate recombinase assembled at or behind a stalled fork and which function(s) may counteract the PcrA antirecombinase activity, null mutants in presynaptic functions, namely end resection (Δ*addAB*, Δ*recJ*, Δ*recQ*), RecA mediation (*recO*16), and positive (Δ*rarA*) or negative (Δ*recX*, Δ*recU*) modulators (**Table S1**), were assessed in the *pcrA-ssrA sspB* ($pcrA_T$) context (**Table 1**) (Sanchez et al., 2006; Cárdenas et al., 2012; Romero et al., 2020).

B. subtilis Δ*pcrA* lethality is suppressed in the *recL*16 background (Petit and Ehrlich, 2002). Before testing the causes of the suppression of *B. subtilis* Δ*pcrA* lethality in the *recL*16 context, we must understand the function(s) impaired in the *recL*16 strain. As described in Supplementary Material, **Annex 1**, the *B. subtilis* mutations previously designated as *recL*16 actually map to the *recO* locus (C. M, Marielle C. Gruenig, Michael M. Cox and J.C.A., to be published elsewhere). To simplify the analysis, the MMS^S phenotype was transferred by gene congression to competent BG214 cells. Five of the resulting MMS^S clones were selected for whole genome sequencing along with the isogenic BG214 rec^+ isogenic strain (Supplementary Material, **Annex 1**). One of the MMS^S clones showed a TGA Opal stop triplet at codon 37 of *recO* and was designated recO16 (BG107-1 strain) and selected for further analysis (**Table 1**).

In the presence of IPTG, the survival of the *recO*16, Δ*addAB*, Δ*recX,* or Δ*recU* strain decreased ∼12-, ∼65-, ∼3-, and ∼7-fold upon addition of 0.2 mM H_2O_2 (blue bar [+ IPTG and H_2O_2]), and ∼2.5 × 10^6-, ∼6 × 10^6-, ∼1 × 10^5,- and ∼1 × 10^4-fold upon addition of 1.3 mM MMS (green bar [+ IPTG and MMS]), respectively (**Figure 1**). Cell survival when comparing the single Δ*recJ*, Δ*recQ,* and Δ*rarA* mutant strains and the rec^+ strain was not statistically significant decreased under the H_2O_2 and MMS concentrations used in this assay (**Figure 1**), although cell survival was significantly decreased in the Δ*recJ* and Δ*recQ* background at higher H_2O_2 and MMS concentrations, and in the Δ*rarA* context at higher H_2O_2 concentrations (Sanchez et al., 2006; Romero et al., 2019a). No significant differences were observed when the survival of rec^+ and *rec*–deficient strains in the absence of IPTG was compared to the presence of IPTG.

Next, the contribution of the absence of the AddAB helicase-nucleases complex (counterpart of $RecBCD_{Eco}$), the RecJ 5′ → 3′ ssDNA exonuclease, or the RecQ DNA helicase (**Table S1**) in the $pcrA_T$ context was analyzed. In *E. coli* cells, the RecBCD complex during the end resection process loads RecA onto naked ssDNA, to generate a 3'-overhang coated by RecA (Kowalczykowski, 2015). If AddAB works in a similar fashion it is expected that PcrA depletion inviability requires AddAB. The presence of IPTG in rich medium agar plates did not significantly affected cell viability of the Δ*recJ* $pcrA_T$ (**Figure 1A**), and thus the lethality upon PcrA depletion was suppressed by inactivating *recJ*, but the colony area was ∼9-fold smaller than in the absence of IPTG. In contrast, PcrA depletion was not suppressed by *addAB or* Δ*recQ* inactivation (**Figure 1B**). It is likely that the inviability upon PcrA depletion required the RecJ, but not the AddAB or RecQ functions. Furthermore, the viability of Δ*addAB* $pcrA_T$ cells was significantly decreased (∼10-fold), but not that of Δ*recQ* $pcrA_T$ cells (∼3-fold) when compared with the $pcrA_T$ degron strain (**Figure 1B**).

Exponentially growing Δ*recJ* $pcrA_T$, Δ*addAB* $pcrA_T$, or Δ*recQ* $pcrA_T$ cells were then plated on LB agar plates containing IPTG and H_2O_2 or MMS. The Δ*recJ* mutation significantly increased the survival of $pcrA_T$ cells on plates containing IPTG and H_2O_2 ($P < 0.001$). Unexpectedly, cell survival decreased by ∼3-fold on plates containing IPTG and MMS when compared with $pcrA_T$ cells (**Figure 1A**), suggesting that addition of IPTG and MMS renders Δ*recJ* $pcrA_T$ cells extremely sensitive ($P < 0.001$) when compared to the only IPTG condition (**Figure 1A**, yellow vs. green bar). Addition of IPTG and H_2O_2 or MMS to the Δ*recQ* $pcrA_T$ strain did not significantly affect cell survival, when compared with $pcrA_T$ cells (**Figure 1B**). Cell survival in the Δ*addAB* $pcrA_T$ strain was significantly decreased (by ∼150- and ∼50-fold) on plates containing IPTG and H_2O_2 or MMS, respectively, when compared with $pcrA_T$ cells (**Figure 1B**). It is likely that *pcrA* is not epistatic to *recJ* or *addAB* in response to

MMS-induced DNA damage, and that PcrA is important possibly for backup pathways for single-strand gap and DSB repair.

Here, we have observed that there are different host requirements for the suppression of lethality between *E. coli* and *B. subtilis* cells impaired in end resection: first, the minimal medium synthetic lethality of *E. coli* Δ*rep* Δ*uvrD* cells is suppressed by *recJ* or *recQ* inactivation (Lestini and Michel, 2008), although lack of RecQ or RecJ provides very limited suppression of Δ*rep* Δ*uvrD* rich medium lethality (Guy et al., 2009). On the other hand, in *B. subtilis*, the lethality upon PcrA depletion is suppressed by *recJ* inactivation of cells grown in rich medium (**Figure 1A**) and it will be of significant interest to understand the molecular basis of the small colony size of Δ*recJ* $pcrA_T$ upon PcrA depletion. Second, the lethality upon PcrA depletion is not suppressed by *recQ* inactivation, and survival of the Δ*recQ* $pcrA_T$ cells was not significantly affected and reduced in the presence of H_2O_2- and MMS-induced DNA damage, respectively (**Figure 1B**). Unlike *E. coli*, two RecQ-like enzymes (RecQ and RecS) are present in *B. subtilis* (Fernández et al., 1998). RecQ, which is 591 amino acid long, shares ∼43% identity with RecS (496 amino acid in length) if the first 346 residues containing the DExH helicase domains are used for the alignment; in short, RecS lacks the zinc-finger, the winged-helix, and the RNaseD C-terminal domains (Fernández et al., 1998; Bernstein et al., 2003). We cannot rule out that the partial genetic redundancy, exerted by RecS and RecQ, might mask the phenotype. However, since the viability of the Δ*recJ* $pcrA_T$ or Δ*recQ* $pcrA_T$ strain in the presence of IPTG and MMS was similar (**Figures 1A,B**), we have to assume that PcrA depletion and Δ*recJ* or Δ*recQ* inactivation might have different host requirements.

Next, the contribution of the *recO*16 mutation in the survival of the *recO*16 $pcrA_T$ strain was evaluated. The presence of IPTG did not significantly affect cell viability of the *recO*16 $pcrA_T$ strain grown in rich medium (**Figure 1A**). This is in good agreement with the observation that *pcrA* inactivation inviability requires RecO when grown in synthetic minimal medium (Petit and Ehrlich, 2002). To test whether PcrA works in a similar or different pathway than RecO, exponentially growing *recO*16 $pcrA_T$ cells were plated on LB agar plates containing 500 μM IPTG and H_2O_2 or MMS. The survival of the *recO*16 $pcrA_T$ strain significantly increased (>3,500-fold), when compared to the $pcrA_T$ strain upon plating in IPTG and H_2O_2 containing plates. Unexpectedly, the survival of the *recO*16 $pcrA_T$ strain was significantly decreased (>30-fold) when compared to the $pcrA_T$ strain on plates containing IPTG and MMS (**Figure 1A**). Addition of IPTG and MMS rendered *recO*16 $pcrA_T$ cells extremely sensitive ($P < 0.001$) when compared to just the addition of IPTG alone (**Figure 1A**, yellow vs. green bar). RecO has two activities: to load RecA onto SsbA-coated ssDNA in concert with RecR and to mediate DNA strand annealing independently of RecR (Kidane et al., 2004; Manfredi et al., 2008, 2010; Lenhart et al., 2014). Since Δ*pcrA* inviability requires RecO and RecR (Petit and Ehrlich, 2002), we assumed that the inactivation of *recO* compromises RecA loading at or behind a stalled fork, and thereby avoids unscheduled or unwanted RecA-mediated recombination during DNA replication in the context of PcrA depletion.

The contribution of mutants in the modulation of RecA filament growth in the inviability of PcrA depleted cells was also assessed. RarA has at least two activities: to control the loading of pre-primosomal proteins at a stalled fork and to positively modulate RecA filament growth (Carrasco et al., 2018; Romero et al., 2019a, 2020). PcrA depletion lethality was not suppressed by *rarA* inactivation (**Figure 1B**). After IPTG addition, the viability of PcrA depleted cells was significantly decreased (by ∼7-fold) in the Δ*rarA* $pcrA_T$ context when compared with $pcrA_T$ cells (**Figure 1B**).

The absence of RarA renders cells sensitive to H_2O_2, but not to MMS (Romero et al., 2019a). Cell survival in the Δ*rarA* $pcrA_T$ background was not significantly reduced on plates containing IPTG and H_2O_2 or MMS, when compared to the $pcrA_T$ strain (**Figure 1B**). However, this observation is not consistent with the proposal that in *E. coli*, UvrD prevents RecA binding to ssDNA, possibly by counteracting RarA (Lestini and Michel, 2007).

RecX is a negative modulator of RecA filament growth (Cárdenas et al., 2012; Le et al., 2017), whereas RecU has two activities: to negatively modulate RecA filament growth and to cleave Holliday junctions (HJs) at a cognate site in concert with the RuvAB branch migration translocase (Ayora et al., 2004; Carrasco et al., 2005; Cárdenas et al., 2012; Cañas et al., 2014; Serrano et al., 2018). The absence of RecX or RecU did not restore viability of PcrA depleted cells grown in rich medium agar plates containing IPTG. The viability of PcrA depleted cells was significantly reduced (by ∼12- and ∼13-fold) in the Δ*recX* $pcrA_T$ or Δ*recU* $pcrA_T$ context, respectively, when compared to the $pcrA_T$ strain (**Figure 1B**). The presence of IPTG and H_2O_2 or MMS strongly reduced cell survival in the Δ*recX* $pcrA_T$ (∼30- and ∼40-fold) and Δ*recU* $pcrA_T$ (∼13- and ∼40-fold) backgrounds, respectively, when compared to the $pcrA_T$ strain (**Figure 1B**), suggesting that the *pcrA* gene is likely not epistatic to *recX* or *recU* in response to non-bulky DNA lesions of oxidative nature. In *E. coli*, Δ*uvrD* Δ*ruvC* (counterpart of *B. subtilis* RecU) cells are inviable (Magner et al., 2007).

PcrA Depletion Inviability Requires RecA

Rich medium synthetic lethality of *E. coli* Δ*uvrD* Δ*rep* cells is not suppressed by *recA* inactivation (Veaute et al., 2005; Guy et al., 2009). Similarly, *B. subtilis pcrA* lethality is not suppressed when the leaky *recA*1 (formerly termed *recE*1) mutation is moved into the background (Petit and Ehrlich, 2002), but the RecA1 activities present in the background may mask the outcome (e.g., competent *recA*1 cells are marginally affected, whereas null *recA* cells are blocked in chromosomal transformation) (Alonso et al., 1988). To address whether *recA* inactivation suppresses cell inviability upon PcrA depletion, a *recA* null mutant allele (Δ*recA*) was moved, *via* SPP1-mediated generalized transduction, onto the $pcrA_T$ strain (see **Table 1**). Absence of RecA reduced survival of the parental strain by ∼80-fold in response to H_2O_2, and strongly reduced cell survival (∼6×10^6-fold) in response to MMS, when compared to the *wt* strain (**Figure 1A**).

The lethality of PcrA depleted cells was significantly suppressed ($P < 0.001$) by *recA* inactivation upon IPTG addition

to rich LB agar plates (**Figure 1A**), but the colony area was ∼17-fold smaller than in the absence of IPTG. Addition of IPTG and H_2O_2 or MMS did not significantly reduce cell survival in the Δ*recA* $pcrA_T$ background when compared to the Δ*recA* strain (**Figure 1A**, blue vs. green bar). To re-evaluate the results in the Δ*recA* $pcrA_T$ background, the Δ*recA* $pcrA_T$ strain was reconstructed, and similar results were observed. The apparent contradiction between PcrA depleted cells in the Δ*recA* context (**Figure 1A**) with the Δ*pcrA recA*1 condition (Petit and Ehrlich, 2002), can be attributed either to the leaky *recA*1 mutation or to background differences, since there are extrachromosomal elements (as the conjugative element ICE*Bs1* and the prophage SPβ) in the Δ*pcrA recA*1 context (Petit and Ehrlich, 2002) that are absent in the Δ*recA* $pcrA_T$ background (see **Table 1**).

It is likely, therefore, that: (i) inviability upon PcrA depletion could not be associated with the inability to promote auto-proteolysis of the transcriptional repressor LexA, because inactivation of *recA* prevents SOS induction, but suppresses $pcrA_T$ lethality in the presence of IPTG (**Figure 1A**); and (ii) PcrA depletion inviability requires RecA, and *pcrA* is epistatic to *recA* in response to non-bulky DNA lesions of oxidative nature.

PcrA Depletion Leads to Unsegregated Chromosomes

As previously proposed for eukaryotic Srs2 (Marini and Krejci, 2010), the role of PcrA might be to prevent the accumulation of crossovers (CO) by promoting synthesis-dependent strand annealing (SDSA), leading to the exclusive accumulation of non-crossover (NCO) products. In circular chromosomes, the outcome of CO and NCO will be a dimeric or two monomeric chromosomes, respectively. Dimeric chromosomes are deleterious and need to be processed before cell division. To test whether PcrA prevents CO accumulation, the chromosome segregation of PcrA depleted cells was studied.

Branched DNA structures can be processed by different pathways. First, the RecU HJ resolvase, in concert with the RuvAB translocase, cleaves the outside or the inside strands of a HJ, followed by religation to produce CO and NCO products, respectively (Carrasco et al., 2004; Cañas et al., 2014). Second, a HJ can be dissolved by the RecQ helicase in concert with a Type I DNA topoisomerase to produce NCO products (Kowalczykowski, 2015). Finally, PcrA may dismantle RecA nucleoprotein filaments from branched structures and may promote SDSA by unwinding the elongated invading strand, a step that is followed by annealing to the ssDNA of the other end of the break, an event associated with NCO products (Marini and Krejci, 2010).

To evaluate whether PcrA depletion provokes a chromosomal segregation defect, the nucleoid (supercoiled and compacted chromosome bound by proteins) of the $pcrA_T$, *recO*16 $pcrA_T$, Δ*recJ* $pcrA_T$, Δ*recQ* $pcrA_T$, Δ*rarA* $pcrA_T$, Δ*recX* $pcrA_T$, Δ*recU* $pcrA_T$, or Δ*recA* $pcrA_T$ cells was DAPI-stained and analyzed by fluorescence microscopy. As controls we have used the Δ*recU* and Δ*recA* strains. In the absence of any external DNA damage and at mid-exponential phase, Δ*recU* or Δ*recA* mutations reduce the number of CFUs by ∼5- and ∼10-fold (**Figure 1**), and ∼30 and ∼40% of cells are filamented, respectively (**Figure 2**) (Carrasco et al., 2004), suggesting that in the absence of RecU or RecA, a cell subpopulation undergoes a death-by-recombination phenotype.

Previously it has been shown that after 15 min (37°C) of 100 μM IPTG addition to *B. subtilis* cells 60 to 90% of PcrA is degraded (Merrikh et al., 2015). In this study, cells were grown in rich medium under unperturbed conditions until they reached $OD_{560} = 0.2$ (37°C). IPTG (500 μM) was added to half of the culture, and both cultures were incubated (60 min, 37°C) before harvesting, fixing and staining the cells with DAPI. During vegetative growth, net accumulation of anucleated cells, unsegregated chromosomes and filamented cells was rare in rec^+ cells in the absence (**Figure 2**) or presence of IPTG (data not shown). In this scenario, cells displayed an average length of 4–6 μm, and exhibited a bimodal distribution of nucleoid positioning with ∼35% of total cells having two nucleoids, and ∼65% of total cells containing only one nucleoid with about twice the fluorescence signal as judged by eye (Carrasco et al., 2004). This suggests that the former class were replicated cells with segregated chromosomes and the latter were replicated cells with yet-unsegregated chromosomes.

In the absence of IPTG, ∼97% of *pcrA-ssrA* cells appeared normal compared to the rec^+ control (∼100%) (**Figure 2**). Absence of DAPI staining (anucleated cells) significantly increased (by ∼20-fold) and the fraction of cells with aberrant chromosomes by ∼6-fold when compared to the rec^+ control, whereas upon IPTG addition absence of DAPI staining was significantly decreased (**Figure 2**), suggesting that the *pcrA-ssrA* fusion or noise from *sspB* gene expression affects chromosomal segregation of unperturbed exponentially growing $pcrA_T$ cells. In the presence of IPTG, PcrA dropped, the proportion of cells with an incompletely separated nucleoid or aberrant chromosomes increased by ∼50-fold when compared to the condition without IPTG, and the average cell length was >8 μm in ∼20% of total $pcrA_T$ cells (**Figure 2**). Since those elongated cells contained a single nucleoid it was assumed that they were filamented cells. It is likely that, upon PcrA depletion, NCO shifted toward CO products. In other words, PcrA might suppress COs or might directly contribute to the formation of NCO products, as was shown for its yeast homolog Srs2 (Marini and Krejci, 2010).

In the absence of IPTG, the proportion of anucleated cells and cells with aberrant chromosomes was marginally affected in the *recO*16 $pcrA_T$, Δ*recJ* $pcrA_T$, and Δ*recA* $pcrA_T$ strains when compared to the single mutant strain, but in the Δ*recA* or Δ*recA* $pcrA_T$ condition 40–45% of total cells were filamented (**Figure 2**). In the presence of IPTG, the absence of DAPI staining and the proportion of unsegregated nucleoids were not significantly affected (<2-fold) when compared to the parental control (rec^+ $pcrA_T$) strain (**Figure 2**). Likewise, in the *recO*16 $pcrA_T$, Δ*recJ* $pcrA_T$, and Δ*recA* $pcrA_T$ conditions a significant proportion of cells formed filaments, with <2-fold increase in the proportion of filamented cells when compared to the most affected parental strain (**Figure 2**). It is likely that the PcrA pro-SDSA function requires RecJ, RecO, or RecA.

The remaining strains were classified into three different classes. First, for RarA and RecQ, which have two activities each

Mutant strain	% of normal cells (no. of cells counted) 	% of anucleated cells 	% of unsegregated nucleoids 	% filamentation
rec^{+} [a]	99.8 (3676)	< 0.1	< 0.1	< 1
pcrA$_{T}$ [a]	97.2 (1002)	2.2	0.6	< 1
pcrA$_{T}$ [b]	70.8 (1137)	0.7	28.5	19.6
recO16 [a]	96.7 (998)	0.6	2.7	< 1
recO16 pcrA$_{T}$ [a]	94.2 (1321)	1.1	4.7	< 1
recO16 pcrA$_{T}$ [b]	63.2 (1189)	1.2	35.6	20.3
Δ*recJ* [a]	96.8 (976)	0.6	2.6	< 1
Δ*recJ pcrA*$_{T}$ [a]	95 (999)	0.9	4.1	< 1
Δ*recJ pcrA*$_{T}$ [b]	68.9 (1123)	0.9	30.2	33.3
Δ*recQ* [a]	98.7 (1084)	0.2	1.1	19.6
Δ*recQ pcrA*$_{T}$ [a]	96.6 (963)	0.3	3.1	18.1
Δ*recQ pcrA*$_{T}$ [b]	44.1 (1064)	0.8	55.1	47.3
Δ*rarA* [a]	96.8 (976)	1	2.2	< 1
Δ*rarA pcrA*$_{T}$ [a]	93.2 (1031)	0.4	6.4	< 1
Δ*rarA pcrA*$_{T}$ [b]	53.8 (1096)	0.8	45.4	20.3
Δ*recX* [a]	98.8 (1225)	0.4	0.8	< 1
Δ*recX pcrA*$_{T}$ [a]	95.5 (1031)	0.4	4.1	< 1
Δ*recX pcrA*$_{T}$ [b]	74.8 (2021)	0.8	24.6	< 1
Δ*recU* [a]	79.9 (2533)	5.2	15.1	31.5
Δ*recU pcrA*$_{T}$ [a]	77.1 (1347)	5.5	17.4	33.2
Δ*recU pcrA*$_{T}$ [b]	31.4 (1201)	7.4	61.2	45.7
Δ*recA* [a]	82.4 (2543)	0.7	16.9	39.7
Δ*recA pcrA*$_{T}$ [a]	76.7 (1143)	1.1	22.2	44.2
Δ*recA pcrA*$_{T}$ [b]	52 (985)	1.3	46.7	61.5

FIGURE 2 | Chromosome segregation in the absence of presynaptic or synaptic functions. Cells were grown in LB medium to OD_{560} = 0.2; after 60 min, cells were harvested, prepared for DAPI DNA-fluorescence microscopy, and the percentage of anucleate and unsegregated nucleoids determined (condition *a*). In parallel, at OD_{560} = 0.2, IPTG (500 μM) was added and after 60 min, cells were harvested, DAPI-stained, and the percentage of anucleate and unsegregated nucleoids determined (condition *b*). Representative fluorescent images of two dividing DAPI-treated cells (DNA stain, light blue) are shown. The pictures are taken at the same amplification, two none separated cells (four nucleoids) are presented under normal conditions. The mean of at least three independent experiments is shown.

(see above): inactivation of *recQ* or *rarA* revealed a marginal chromosome segregation defect, but ∼20% of total Δ*recQ* cells formed filaments (**Figure 2**). In the presence of IPTG, the proportion of unsegregated nucleoids significantly increased in Δ*recQ pcrA*$_{T}$ or Δ*rarA pcrA*$_{T}$ when compared to the *pcrA*$_{T}$ control, but absence of DAPI staining was not significantly affected (**Figure 2**). Upon PcrA depletion the proportion of filamentous cells was similar in *pcrA*$_{T}$ and Δ*rarA pcrA*$_{T}$ cells, but significantly increased in the Δ*recQ pcrA*$_{T}$ context. Second, RecX negatively modulates RecA filament growth (Cárdenas et al.,

2012; Le et al., 2017). In the presence of IPTG, the proportion of unsegregated nucleoids in Δ*recX pcrA*$_T$ was similar to the *pcrA*$_T$ control, but counteracted the formation of filamented cells (**Figure 2**). Third, for RecU, which has two activities (see above): in the absence of RecU, ∼5% of total cells were anucleated as previously described (Carrasco et al., 2004), suggesting that cell division occurred in regions that had not received a nucleoid, and unsegregated nucleoids accounted up to ∼15% of total cells. In the presence of IPTG, most of Δ*recU pcrA*$_T$ cells (∼60%) had unsegregated nucleoids, and ∼45% of cells were present as "filaments" (**Figure 2**), suggesting that PcrA may promote SDSA prior to the formation of a double-HJ that can be resolved to NCO and CO by the RecU HJ resolvase in concert with the RuvAB branch migration translocase (Ayora et al., 2004; Cañas et al., 2014).

Altogether, the data presented in **Figures 1**, **2** revealed certain paradoxes. First, the inviability of PcrA depletion requires RecJ, RecO, and RecA, but under these conditions a chromosomal segregation defect was observed, suggesting that PcrA processes branched DNA structures formed at replication or at replication-transcription stalled forks, but with the help of accessory proteins (e.g., RecJ, RecO) a formed RecA nucleoprotein filament may be dismantled by PcrA. Alternatively, PcrA removes proteins bound to stalled forks to indirectly allow the formation of branched structures. Second, PcrA depletion halts cell proliferation, initiates accumulation of unprocessed branched intermediates, and additively reduces repair-by-recombination in the Δ*recQ* or Δ*rarA* context. Finally, PcrA depletion exacerbates the segregation defect of Δ*recU* cells, with only ∼30% having normal chromosomal segregation, and in the absence of both negative RecA modulators there is a synergistic repair-by-recombination defect.

PcrA Inviability Requires Mfd, but Not RecD2, HelD, HepA, or YwqA

Enzymes of the UvrD family of translocases provide different solutions to cope with a replicative stress and/or RTCs. UvrD$_{Eco}$ and PcrA can interact with and backtrack RNAP *in vivo* and *in vitro*, that is a crucial step for minimizing RTCs and for the repair of lesions occluded by a stalled RNAP, which become a major obstacle to DNA replication (Epshtein et al., 2014). Other DNA helicases/translocases of SF1, namely HelD and RecD2, and of SF2, such as Mfd, HepA (YqhH), and YwqA, also interact with RNAP *via* a conserved domain (**Table S2**) (Muzzin et al., 1998; Sukhodolets et al., 2001; Deaconescu et al., 2006; Shaw et al., 2008; Guy et al., 2009; Boubakri et al., 2010; Jin et al., 2011; Epshtein et al., 2014; Wiedermannova et al., 2014; Liu et al., 2015; Sanders et al., 2017; Le et al., 2018). These enzymes have been also implicated: (i) in preventing or mitigating the impact of protein-DNA complexes or spontaneous non-bulky DNA lesions of oxidative nature that halt transcription or replication, (ii) in avoiding the conflicts generated by the collision between the replication and transcription machineries, and (iii) in promoting RNAP recycling, sliding backward along the template (backtracking or retreating) or RNAP removal, or when replication forks are arrested by the formation of R-loops (Ayora et al., 1996; Komissarova and Kashlev, 1997; Sukhodolets et al., 2001; Trautinger et al., 2005; Deaconescu et al., 2006; Guy et al., 2009; Yawn et al., 2009; Boubakri et al., 2010; Gupta et al., 2013; Bruning et al., 2014; Wiedermannova et al., 2014; Merrikh et al., 2015; Sanders et al., 2017; Ho et al., 2018; Le et al., 2018).

To study whether the lack of PcrA destabilization of transcription complexes is the primary cause of inviability, the Δ*helD pcrA*$_T$, Δ*recD2 pcrA*$_T$, Δ*hepA pcrA*$_T$, Δ*ywqA pcrA*$_T$, and Δ*mfd pcrA*$_T$ strains were constructed as described (Materials and methods). PcrA depletion, which decreases cell survival in the Δ*recD*2 context (Torres et al., 2017), was used as control. Except in the Δ*mfd* strain, the H_2O_2 or MMS concentrations used were not sufficient to reveal a reduced viability phenotype (**Figure 3**).

Upon addition of IPTG, the lethality observed following PcrA depletion was not suppressed by *recD*2 inactivation (**Figure 3**). In the presence of IPTG, the viability was significantly decreased (by ∼25-fold) in Δ*recD*2 *pcrA*$_T$ when compared to the *pcrA*$_T$ strain. Addition of both IPTG and H_2O_2 or MMS significantly reduced cell survival (by ∼8- and ∼5-fold, respectively) in the Δ*recD*2 *pcrA*$_T$ cells when compared to the *pcrA*$_T$ strain (**Figure 3**), suggesting that *pcrA* is not epistatic to *recD*2 in response to H_2O_2- or MMS-induced DNA damage.

Inactivation of the *helD* helicase partially suppresses the DNA repair defect of *recF*15, Δ*recO,* and Δ*recR* cells (Carrasco et al., 2001). The absence of HelD did not suppress cell inviabilily upon PcrA depletion; in this background, cell viability was not significantly decreased when compared to the *pcrA*$_T$ control upon addition of IPTG (**Figure 3**, yellow bar [+ IPTG]). Upon addition of IPTG and MMS, cell survival was significantly decreased (by ∼4-fold), when compared to the *pcrA*$_T$ strain (**Figure 3**, green bar [+ IPTG and MMS]). In contrast, in the absence of *E. coli* HelD and UvrD cells remain recombination proficient and apparently are as capable of repairing MMS-induced DNA damage as the *wt* control (Mendonca et al., 1993). Addition of IPTG and H_2O_2 marginally decreased cell survival (by ∼2-fold) (**Figure 3**, blue bar [+ IPTG and H_2O_2]).

The absence of HepA or YwqA did not suppress cell inviability upon PcrA depletion (**Figure 3**). Cell viability was significantly decreased (by ∼8-fold) in the Δ*ywqA pcrA*$_T$ when compared to the *pcrA*$_T$ control, but not in the Δ*hepA pcrA*$_T$ (decreased by ∼3-fold) when compared to the *pcrA*$_T$ control (**Figure 3**).

Addition of IPTG and H_2O_2 did not significantly reduce survival in Δ*hepA pcrA*$_T$ or Δ*ywqA pcrA*$_T$ cells when compared to the *pcrA*$_T$ strain (**Figure 3**). The presence of IPTG and MMS significantly reduced cell survival in the Δ*hepA pcrA*$_T$, but marginally reduced cell survival in Δ*ywqA pcrA*$_T$ cells when compared to the *pcrA*$_T$ strain (**Figure 3**). These data altogether suggest that decreasing the probability of backtracking events contributes to maintaining genome stability, but not to suppress the lethality of PcrA depleted cells. Unlike in *E. coli* cells (Shaw et al., 2008; Jin et al., 2011; Liu et al., 2015), we have little information about how the *B. subtilis* HepA or YwqA ATPase propels backward translocation of the RNAP along the DNA template or release a sequestered RNAP.

E. coli cells lacking Mfd show a weak sensitivity to UV irradiation (Witkin, 1969). *E. coli* Mfd recruits UvrA to the site of a roadblock that stalls RNAP. This activity is crucial

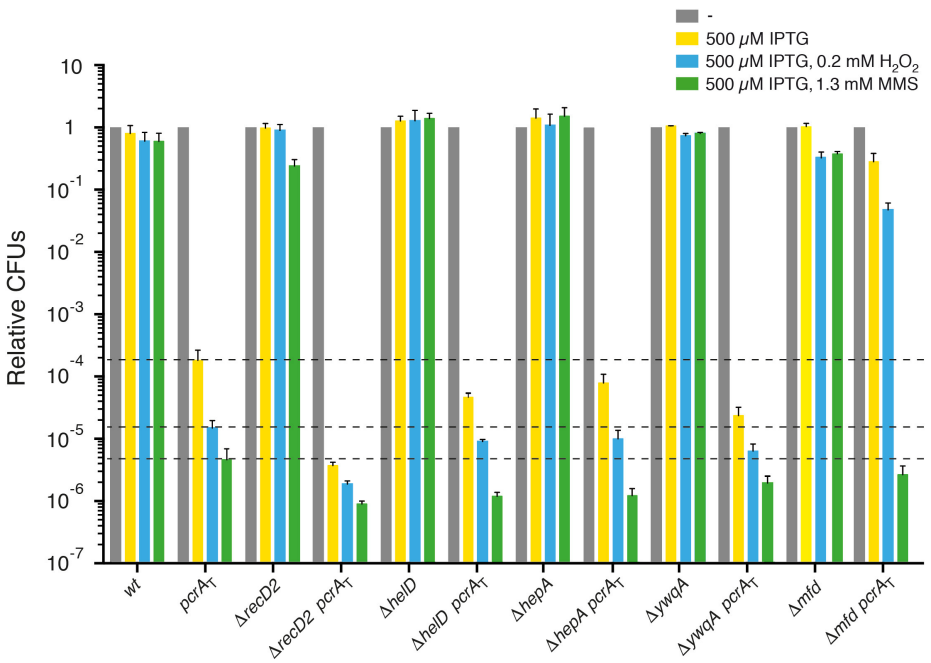


FIGURE 3 | PcrA lethality is suppressed by *mfd* inactivation, but not by *recD2*, *helD*, *hepA, ywqA* inactivation. Log phase cultures of *wt*, single or double mutant strains were diluted and plated on LB agar containing 500 μM IPTG (yellow bars) or lacking it (gray bars). Lethality assays showing cell viability upon PcrA depletion in *recD2*, *helD*, *hepA, ywqA,* or *mfd* cells. Log phase cultures of indicated strains were diluted and plated on LB agar containing IPTG and 0.2 mM H_2O_2 (blue bars) or IPTG and 1.3 mM MMS (green bars). Experiments were performed at least four times. The dotted lines mark the survival rate upon PcrA depletion. Data are shown as mean fractional survival ± SEM.

for the recognition and removal of a stalled RNAP, but Mfd is subsequently displaced by UvrB to initiate transcription coupled repair (Selby and Sancar, 1993, 1994; Ayora et al., 1996; Manelyte et al., 2010; Epshtein, 2015; Ho et al., 2018, 2020; Le et al., 2018). In contrast, inactivation of *B. subtilis mfd* renders cells significantly sensitive to the UV mimetic 4-nitroquinoline-1-oxide and also to oxidative non-bulky lesions as those generated by exposure to MMS (Ayora et al., 1996). Our results suggested that Mfd is also required in *B. subtilis* cells to repair non-bulky lesions (**Figure 3**). It is likely that in *B. subtilis*: (i) bulky and non-bulky DNA lesions stall RNAP; and (ii) Mfd interacts with and dislodges RNAP from the damaged DNA template.

Upon addition of IPTG, inactivation of *mfd* significantly suppressed ($P < 0.001$) the lethality induced by PcrA depletion (**Figure 3**), but the colonies were minute and with an area ∼19-fold smaller than in the absence of IPTG, suggesting that Mfd and PcrA play a crucial role in response to a replicative stress. The Δ*mfd* mutation significantly suppressed the sensitivity of Δ*mfd* $pcrA_T$ cells to H_2O_2 (addition of both IPTG and H_2O_2), but Δ*mfd* $pcrA_T$ cells showed a non-significant decrease when compared to the $pcrA_T$ strain in the presence of both IPTG and MMS (**Figure 3**). It is likely that the *pcrA* gene is epistatic to *mfd* in response to MMS-induced DNA lesions. Inactivation of *E. coli mfd* partially suppresses the sensitivity to UV irradiation in the *uvrD* context (Epshtein et al., 2014).

CONCLUSIONS

We show that PcrA depletion lethality is suppressed by *recJ*, *recO*16, or *recA* inactivation, but not by *addAB*, *recQ*, *rarA*, *recX*, or *recU* inactivation when cells are grown in rich medium (**Figures 1A,B**). These data suggest that PcrA depleted cells primarily die due to their inability to resuscitate replisomes blocked by a RecA-ssDNA complex. Indeed, RecO loads RecA onto SsbA-coated ssDNA and a RecA nucleoprotein filament downregulates initiation of PriA-dependent DNA replication *in vitro* (Vlasic et al., 2014), and PcrA depletion inviability requires RecA for replication re-start (Million-Weaver et al., 2015). This is consistent with the observation that PcrA depletion inviability also requires RecO, which loads RecA onto ssDNA (Carrasco et al., 2015), but not AddAB. In a minimal synthetic medium, the *pcrA* inactivation lethality is also suppressed by inactivation of the *recO* or *recR* positive mediators or a leaky mutation in the positive *recF*17 modulator (Petit and Ehrlich, 2002).

In *E. coli* cells, the synthetic lethality of *uvrD* and *rep* mutations is partially suppressed by *recJ* or *recQ* inactivation in minimal medium (Lestini and Michel, 2008). We can envision that the discrepancies observed between *E. coli* and *B. subtilis* cells are related to genetic differences between these genetically distant bacteria. First, *B. subtilis* cells have two RecQ-like helicases, RecQ and RecS, with the latter potentially masking the outcome, whereas *E. coli* has only RecQ. Second, *E. coli* cells have

two proteins (UvrD and RecX) to actively dismantle a RecA nucleoprotein filament (Petrova et al., 2015; Le et al., 2017), whereas *B. subtilis* cells have four different proteins (PcrA, RecX, RecU, and RecD2) to do this job (Anand et al., 2007; Le et al., 2017; Torres et al., 2017; Serrano et al., 2018). PcrA was also found to be necessary to survive DNA damage. The *pcrA* gene is not epistatic to genes involved in end resection (*addAB*, *recJ*, *recQ*), RecA mediators (*recO*16), or negative RecA modulators (*recX*, *recU*) in response to MMS- or H_2O_2-induced DNA damage, but it is epistatic to the *recA* gene, suggesting that PcrA also contributes to repair-by-recombination *via* poorly understood mechanisms. The role of the positive RecA modulator RarA upon PcrA depletion requires further studies.

As it has been seen previously for Srs2 (Marini and Krejci, 2010), PcrA might play a putative role in promoting SDSA, which does not entail the generation of COs. Depletion of PcrA leads to additive, in *recJ*, *recO* and *recA*, and to synergic accumulation of unsegregated chromosomes in *recQ*, *rarA recU* backgrounds. The dual activities of these proteins (see above) mask the interpretation of our results.

Finally, we show that PcrA depletion lethality is suppressed by *mfd* inactivation, but not by *recD*2, *helD*, *hepA,* or *ywqA* inactivation (**Figure 3**). We show that *pcrA* is not epistatic to *recD*2, *helD,* or *hepA* in response to non-bulky DNA damage, but it is epistatic to *mfd*. The role of the poorly characterized YwqA ATPase upon PcrA depletion requires further studies.

PcrA and Mfd might act on RTCs, both dependent (Ayora et al., 1996) and independently of the nucleotide excision repair pathway (**Figure 3**). The PcrA and Mfd translocases physically interact with stalled RNAPs at lesions on the DNA template. PcrA is a pro-backtracking factor by promoting forward RNAP translocation, and Mfd might be an anti-backtracking that dislodges a stalled RNAP, as previously postulated for the isolated protein *in vitro* (Selby and Sancar, 1993; Ayora et al., 1996; Park et al., 2002; Deaconescu et al., 2006; Epshtein et al., 2014; Sanders et al., 2017; Ho et al., 2018; Le et al., 2018). It is likely that when damaged template bases interfere with RNAP progression, it backtracks and becomes transiently arrested. Under this condition, Mfd and PcrA, which interact with UvrA and UvrB, respectively, are crucial factors involved in mitigating RTCs in the presence of DNA lesions that are or not targeted by transcription coupled repair. Thus, we propose that PcrA is crucial to remove a stalled RNAP that would otherwise hinder DNA replication even in the presence of DNA lesions that are not targeted by transcription coupled repair. In other words, in the absence of PcrA, RNAP may not be evicted from the damage site by Mfd, leading to a harmful genotoxic stress that induces lethality. PcrA allows genome duplication to occur concurrently with other essential DNA transactions (replication, transcription, repair, segregation). Partial PcrA depletion sensitizes cells to severe DNA transactions and its major role is to work in concert with recombination and repair proteins at stalled DNAP/RNAP complexes to facilitate replication progression beyond the conflict point.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

MM, RT, and JA designed the experiments and drafted the manuscript. MM, RT, CM, JR-M, and GS performed the experiments. JA coordinated the research and wrote the manuscript. MM, RT, CM, JR-M, GS, and JA interpreted the data. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Ministerio de Ciencia e Innovación, Agencia Estatal de Investigación (MCIU/AEI)/FEDER PGC2018-097054-B-I00 to JA.

ACKNOWLEDGMENTS

We are grateful to Marie A. Petit for the Δ*yqhH* (Δ*hepA*) and Δ*ywqA* strains, to Houra and Christopher Merrikh for the *pcrA-ssrA sspB* degron strain, and to Marielle C. Gruenig and Michael M. Cox for the early steps of *recL*16 mapping. We thank C. Marchisone for technical assistance and S. Ayora for comments on the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2020.00140/full#supplementary-material

REFERENCES

Aboussekhra, A., Chanet, R., Adjiri, A., and Fabre, F. (1992). Semidominant suppressors of Srs2 helicase mutations of *Saccharomyces cerevisiae* map in the RAD51 gene, whose sequence predicts a protein with similarities to procaryotic RecA proteins. *Mol. Cell. Biol.* 12, 3224–3234. doi: 10.1128/MCB.12.7.3224

Albert, T. J., Dailidiene, D., Dailide, G., Norton, J. E., Kalia, A., Richmond, T. A., et al. (2005). Mutation discovery in bacterial genomes: metronidazole resistance in *Helicobacter pylori*. *Nat. Methods* 2, 951–953. doi: 10.1038/nmeth805

Alonso, J. C., Tailor, R. H., and Luder, G. (1988). Characterization of recombination-deficient mutants of *Bacillus subtilis*. *J. Bacteriol.* 170, 3001–3007. doi: 10.1128/JB.170.7.3001-3007.1988

Alonso, J. C., Viret, J. F., and Tailor, H. R. (1987). Plasmid maintenance in *Bacillus subtilis* recombination-deficient mutants. *Mol. Gen. Genet.* 208, 349–352. doi: 10.1007/BF00330464

Anand, S. P., Zheng, H., Bianco, P. R., Leuba, S. H., and Khan, A. S. (2007). DNA helicase activity of PcrA is not required for the displacement of RecA protein from DNA or inhibition of RecA-mediated strand exchange. *J. Bacteriol.* 189, 4502–4509. doi: 10.1128/JB.00376-07

Antony, E., Tomko, E. J., Xiao, Q., Krejci, L., Lohman, T. M., and Ellenberger, T. (2009). Srs2 disassembles Rad51 filaments by a protein-protein interaction triggering ATP turnover and dissociation of Rad51 from DNA. *Mol. Cell.* 35, 105–115. doi: 10.1016/j.molcel.2009.05.026

Ayora, S., Carrasco, B., Doncel, E., Lurz, R., and Alonso, C. J. (2004). *Bacillus subtilis* RecU protein cleaves Holliday junctions and anneals single-stranded DNA. *Proc. Natl. Acad. Sci. U.S.A.* 101, 452–457. doi: 10.1073/pnas.2533829100

Ayora, S., Rojo, F., Ogasawara, N., Nakai, S., and Alonso, C. J. (1996). The Mfd protein of *Bacillus subtilis* 168 is involved in both transcription-coupled DNA repair and DNA recombination. *J. Mol. Biol.* 256, 301–318. doi: 10.1006/jmbi.1996.0087

Baharoglu, Z., Lestini, R., Duigou, S., and Michel, B. (2010). RNA polymerase mutations that facilitate replication progression in the *rep uvr*D *rec*F mutant lacking two accessory replicative helicases. *Mol. Microbiol.* 77, 324–336. doi: 10.1111/j.1365-2958.2010.07208.x

Bernstein, D. A., Zittel, M. C., and Keck, L. J. (2003). High-resolution structure of the *E. coli* RecQ helicase catalytic core. *EMBO* 22, 4910–4921. doi: 10.1093/emboj/cdg500

Boubakri, H., de Septenville, A. L., Viguera, E., and Michel, B. (2010). The helicases DinG, Rep and UvrD cooperate to promote replication across transcription units *in vivo*. *EMBO J*. 29, 145–157. doi: 10.1038/emboj.2009.308

Bruning, J. G., Howard, J. L., and McGlynn, P. (2014). Accessory replicative helicases and the replication of protein-bound DNA. *J. Mol. Biol.* 426, 3917–3928. doi: 10.1016/j.jmb.2014.10.001

Cañas, C., Suzuki, Y., Marchisone, C., Carrasco, B., Freire-Beneitez, V., Takeyasu, K., et al. (2014). Interaction of branch migration translocases with the Holliday junction-resolving enzyme and their implications in Holliday junction resolution. *J. Biol. Chem.* 289, 17634–17646. doi: 10.1074/jbc.M114.552794

Canosi, U., Iglesias, A., and Trautner, A. T. (1981). Plasmid transformation in *Bacillus subtilis*: effects of insertion of *Bacillus subtilis* DNA into plasmid pC194. *Mol. Gen. Genet.* 181, 434–440. doi: 10.1007/BF00428732

Cárdenas, P. P., Carrasco, B., Defeu Soufo, C., César, C. E., Herr, K., Kaufenstein, M., et al. (2012). RecX facilitates homologous recombination by modulating RecA activities. *PLoS Genet.* 8:e1003126. doi: 10.1371/journal.pgen.1003126

Carrasco, B., Ayora, S., Lurz, R., and Alonso, C. J. (2005). *Bacillus subtilis* RecU Holliday-junction resolvase modulates RecA activities. *Nucleic Acids Res.* 33, 3942–3952. doi: 10.1093/nar/gki713

Carrasco, B., Cozar, M. C., Lurz, R., Alonso, J. C., and Ayora, S. (2004). Genetic recombination in *Bacillus subtilis* 168: contribution of Holliday junction processing functions in chromosome segregation. *J. Bacteriol.* 186, 5557–5566. doi: 10.1128/JB.186.17.5557-5566.2004

Carrasco, B., Fernández, S., Petit, M. A., and Alonso, C. J. (2001). Genetic recombination in *Bacillus subtilis* 168: effect of DeltahelD on DNA repair and homologous recombination. *Bacteriol. J.* 183, 5772–5777. doi: 10.1128/JB.183.19.5772-5777.2001

Carrasco, B., Seco, E. M., López-Sanz, M., Alonso, J. C., and Ayora, S. (2018). *Bacillus subtilis* RarA modulates replication restart. *Nucleic Acids Res.* 46, 7206–7220. doi: 10.1093/nar/gky541

Carrasco, B., Yadav, T., Serrano, E., and Alonso, C. J. (2015). *Bacillus subtilis* RecO and SsbA are crucial for RecA-mediated recombinational DNA repair. *Nucleic Acids Res.* 43, 5984–5997. doi: 10.1093/nar/gkv545

de la Hoz, B., Ayora, S., Sitkiewicz, I., Fernández, S., Pankiewicz, R., Alonso, J. C., et al. (2000). Plasmid copy-number control and better-than-random segregation genes of pSM19035 share a common regulator. *Proc. Natl. Acad. Sci. U.S.A.* 97, 728–733. doi: 10.1073/pnas.97.2.728

Deaconescu, A. M., Chambers, A. L., Smith, A. J., Nickels, B. E., Hochschild, A., Savery, N. J., et al. (2006). Structural basis for bacterial transcription-coupled DNA repair. *Cell* 124, 507–520. doi: 10.1016/j.cell.2005.11.045

Dillingham, M. S. (2011). Superfamily I helicases as modular components of DNA-processing machines. *Biochem. Soc. Trans.* 39, 413–423. doi: 10.1042/BST0390413

Dillingham, M. S., Soultanas, P., and Wigley, B. D. (1999). Site-directed mutagenesis of motif III in PcrA helicase reveals a role in coupling ATP hydrolysis to strand separation. *Nucleic Acids Res.* 27, 3310–3317. doi: 10.1093/nar/27.16.3310

Epshtein, V. (2015). UvrD helicase: an old dog with a new trick: how one step backward leads to many steps forward. *Bioessays* 37, 12–19. doi: 10.1002/bies.201400106

Epshtein, V., Kamarthapu, V., McGary, K., Svetlov, V., Ueberheide, B., Proshkin, S., et al. (2014). UvrD facilitates DNA repair by pulling RNA polymerase backwards. *Nature* 505, 372–377. doi: 10.1038/nature12928

Fagerburg, M. V., Schauer, G. D., Thickman, K. R., Bianco, P. R., Khan, S. A., Leuba, S. H., et al. (2012). PcrA-mediated disruption of RecA nucleoprotein filaments–essential role of the ATPase activity of RecA. *Nucleic Acids Res.* 40, 8416–8424. doi: 10.1093/nar/gks641

Fernández, S., Sorokin, A., and Alonso, C. J. (1998). Genetic recombination in *Bacillus subtilis* 168: effects of *rec*U and *rec*S mutations on DNA repair and homologous recombination. *J. Bacteriol.* 180, 3405–3409. doi: 10.1128/JB.180.13.3405-3409.1998

Fu, D., Calvo, J. A., D., and Samson, L. (2012). Balancing repair and tolerance of DNA damage caused by alkylating agents. *Nat. Rev. Cancer* 12, 104–120. doi: 10.1038/nrc3185

Gassel, M., and Alonso, C. J. (1989). Expression of the recE gene during induction of the SOS response in *Bacillus subtilis* recombination-deficient strains. *Mol. Microbiol.* 3, 1269–1276. doi: 10.1111/j.1365-2958.1989.tb00277.x

Griffith, K. L., and Grossman, D. A. (2008). Inducible protein degradation in *Bacillus subtilis* using heterologous peptide tags and adaptor proteins to target substrates to the protease ClpXP. *Mol. Microbiol.* 70, 1012–1025. doi: 10.1111/j.1365-2958.2008.06467.x

Gupta, M. K., Guy, C. P., Yeeles, J. T., Atkinson, J., Bell, H., Lloyd, R. G., et al. (2013). Protein-DNA complexes are the primary sources of replication fork pausing in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 110, 7252–7257. doi: 10.1073/pnas.1303890110

Guy, C. P., Atkinson, J., Gupta, M. K., Mahdi, A. A., Gwynn, E. J., Rudolph, C. J., et al. (2009). Rep provides a second motor at the replisome to promote duplication of protein-bound DNA. *Mol. Cell.* 36, 654–666. doi: 10.1016/j.molcel.2009.11.009

Ho, H. N., van Oijen, A. M., and Ghodke, H. (2018). The transcription-repair coupling factor Mfd associates with RNA polymerase in the absence of exogenous damage. *Nat. Commun.* 9:1570. doi: 10.1038/s41467-018-03790-z

Ho, H. N., van Oijen, A. M., and Ghodke, H. (2020). Single-molecule imaging reveals molecular coupling between transcription and DNA repair machinery in live cells. *Nat. Commun.* 11:1478. doi: 10.1038/s41467-020-15182-3

Jin, D. J., Zhou, Y. N., Shaw, G., and Ji, X. (2011). Structure and function of RapA: a bacterial Swi2/Snf2 protein required for RNA polymerase recycling in transcription. *Biochim. Biophys. Acta* 1809, 470–475. doi: 10.1016/j.bbagrm.2011.03.003

Kamarthapu, V., Epshtein, V., Benjamin, B., Proshkin, S., Mironov, A., Cashel, M., et al. (2016). ppGpp couples transcription to DNA repair in *E. coli*. *Science* 352, 993–996. doi: 10.1126/science.aad6945

Kaniecki, K., De Tullio, L., Gibb, B., Kwon, Y., Sung, P. C., and Greene, E. C. (2017). Dissociation of Rad51 presynaptic complexes and heteroduplex DNA joints by tandem assemblies of Srs2. *Cell Rep.* 21, 3166–3177. doi: 10.1016/j.celrep.2017.11.047

Keiler, K. C., Waller, P. R., and Sauer, T. R. (1996). Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RN. *Science* 271, 990–993. doi: 10.1126/science.271.5251.990

Kidane, D., Sánchez, H., Alonso, J. C., and Graumann, L. P. (2004). Visualization of DNA double-strand break repair in live bacteria reveals dynamic recruitment of *Bacillus subtilis* RecF, RecO and RecN proteins to distinct sites on the nucleoids. *Mol. Microbiol.* 52, 1627–1639. doi: 10.1111/j.1365-2958.2004.04102.x

Komissarova, N., and Kashlev, M. (1997). Transcriptional arrest: *Escherichia coli* RNA polymerase translocates backward, leaving the 3' end of the RNA intact and extruded. *Proc. Natl. Acad. Sci. U.S.A.* 94, 1755–1760. doi: 10.1073/pnas.94.5.1755

Kowalczykowski, S. C. (2015). An overview of the molecular mechanisms of recombinational DNA repair. *Cold Spring Harb. Perspect. Biol.* 7:a016410. doi: 10.1101/cshperspect.a016410

Krejci, L., Van Komen, S., Li, Y., Villemain, J., Reddy, M. S., Klein, H., et al. (2003). DNA helicase Srs2 disrupts the Rad51 presynaptic filament. *Nature* 423, 305–309. doi: 10.1038/nature01577

Le, S, Serrano, E., Kawamura, R., Carrasco, B., Yan, J., and Alonso, C. J. (2017). *Bacillus subtilis* RecA with DprA-SsbA antagonizes RecX function during natural transformation. *Nucleic Acids Res.* 45, 8873–8885. doi: 10.1093/nar/gkx583

Le, T. T., Yang, Y., Tan, C., Suhanovsky, M. M., Fulbright, R. M. Jr., Inman, J. T., et al. (2018). Mfd Dynamically regulates transcription via a release and catch-up mechanism. *Cell* 172, 344–357 e15. doi: 10.1016/j.cell.2017.11.017

Lenhart, J. S., Brandes, E. R., Schroeder, J. W., Sorenson, R. J., Showalter, H. D., and Simmons, A. L. (2014). RecO and RecR Are Necessary for RecA loading in response to DNA damage and replication fork stress. *J. Bacteriol.* 196, 2851–2860. doi: 10.1128/JB.01494-14

Lenhart, J. S., Pillon, M. C., Guarne, A., Biteen, J. S., A., and Simmons, L. (2016). Mismatch repair in Gram-positive bacteria. *Res. Microbiol.* 167, 4–12. doi: 10.1016/j.resmic.2015.08.006

Lestini, R., and Michel, B. (2007). UvrD controls the access of recombination proteins to blocked replication forks. *EMBO J.* 26, 3804–3814. doi: 10.1038/sj.emboj.7601804

Lestini, R., and Michel, B. (2008). UvrD and UvrD252 counteract RecQ, RecJ, and RecFOR in a rep mutant of *Escherichia coli*. *J. Bacteriol.* 190, 5995–6001. doi: 10.1128/JB.00620-08

Liu, B., Zuo, Y., and Steitz, A. T. (2015). Structural basis for transcription reactivation by RapA. *Proc. Natl. Acad. Sci. U.S.A.* 112, 2006–2010. doi: 10.1073/pnas.1417152112

Magner, D. B., Blankschien, M. D., Lee, J. A., Pennington, J. M., Lupski, J. R., and Rosenberg, M. S. (2007). RecQ promotes toxic recombination in cells lacking recombination intermediate-removal proteins. *Mol. Cell*. 26, 273–286. doi: 10.1016/j.molcel.2007.03.012

Manelyte, L., Kim, Y. I., Smith, A. J., Smith, R. M., and Savery, J. N. (2010). Regulation and rate enhancement during transcription-coupled DNA repair. *Mol. Cell*. 40, 714–724. doi: 10.1016/j.molcel.2010.11.012

Manfredi, C., Carrasco, B., Ayora, S., and Alonso, C. J. (2008). *Bacillus subtilis* RecO nucleates RecA onto SsbA-coated single-stranded DNA. *J. Biol. Chem.* 283, 24837–24847. doi: 10.1074/jbc.M802002200

Manfredi, C., Suzuki, Y., Yadav, T., Takeyasu, K., and Alonso, C. J. (2010). RecO-mediated DNA homology search and annealing is facilitated by SsbA. *Nucleic Acids Res.* 38, 6920–6929. doi: 10.1093/nar/gkq533

Marini, V., and Krejci, L. (2010). Srs2: the "Odd-Job Man" in DNA repair. *DNA Repair* 9, 268–275. doi: 10.1016/j.dnarep.2010.01.007

Mendonca, V. M., Kaiser-Rogers, K., and Matson, W. S. (1993). Double helicase II (uvrD)-helicase IV (helD) deletion mutants are defective in the recombination pathways of *Escherichia coli*. *J. Bacteriol.* 175, 4641–4651. doi: 10.1128/JB.175.15.4641-4651.1993

Merrikh, C. N., Brewer, B. J., and Merrikh, H. (2015). The *B. subtilis* accessory helicase PcrA facilitates DNA replication through transcription units. *PLoS Genet*. 11:e1005289. doi: 10.1371/journal.pgen.1005289

Million-Weaver, S., Samadpour, A. N., and Merrikh, H. (2015). Replication restart after replication-transcription conflicts requires RecA in *Bacillus subtilis*. *J. Bacteriol.* 197, 2374–2382. doi: 10.1128/JB.00237-15

Muzzin, O., Campbell, E. A., Xia, L., Severinova, E., Darst, S. A., and Severinov, K. (1998). Disruption of Escherichia coli hepA, an RNA polymerase-associated protein, causes UV sensitivity. *J. Biol. Chem.* 273, 15157–15161. doi: 10.1074/jbc.273.24.15157

Myka, K. K., Hawkins, M., Syeda, A. H., Gupta, M. K., Meharg, C., Dillingham, M. S., et al. (2017). Inhibiting translation elongation can aid genome duplication in *Escherichia coli*. *Nucleic Acids Res.* 45, 2571–2584. doi: 10.1093/nar/gkw1254

Park, J., Myong, S., Niedziela-Majka, A., Lee, K. S., Yu, J., Lohman, T. M., et al. (2010). PcrA helicase dismantles RecA filaments by reeling in DNA in uniform steps. *Cell* 142, 544–555. doi: 10.1016/j.cell.2010.07.016

Park, J. S., Marr, M. T., and Roberts, W. J. (2002). *E. coli* transcription repair coupling factor (Mfd protein) rescues arrested complexes by promoting forward translocation. *Cell* 109, 757–767. doi: 10.1016/S0092-8674(02)00769-9

Petit, M. A., Dervyn, E., Rose, M., Entian, K. D., McGovern, S., Ehrlich, S. D., et al. (1998). PcrA is an essential DNA helicase of *Bacillus subtilis* fulfilling functions both in repair and rolling-circle replication. *Mol. Microbiol.* 29, 261–273. doi: 10.1046/j.1365-2958.1998.00927.x

Petit, M. A., and Ehrlich, D. (2002). Essential bacterial helicases that counteract the toxicity of recombination proteins. *EMBO J.* 21, 3137–3147. doi: 10.1093/emboj/cdf317

Petrova, V., Chen, S. H., Molzberger, E. T., Tomko, E., Chitteni-Pattu, S., Jia, H., et al. (2015). Active displacement of RecA filaments by UvrD translocase activity. *Nucleic Acids Res.* 43, 4133–4149. doi: 10.1093/nar/gkv186

Quail, M. A., Kozarewa, I., Smith, F., Scally, A., Stephens, P. J., Durbin, R., et al. (2008). A large genome center's improvements to the Illumina sequencing system. *Nat. Methods* 5, 1005–1010. doi: 10.1038/nmeth.1270

Raguse, M., Torres, R., Seco, E. M., Gándara, C., Ayora, S., Moeller, R., et al. (2017). *Bacillus subtilis* DisA helps to circumvent replicative stress during spore revival. *DNA Repair.* 59, 57–68. doi: 10.1016/j.dnarep.2017.09.006

Rojo, F., and Alonso, C. J. (1995). The β recombinase of plasmid pSM19035 binds to two adjacent sites, making different contacts at each of them. *Nucleic Acids Res.* 23, 3181–3188. doi: 10.1093/nar/23.16.3181

Rojo, F., C., and Alonso, J. (1994). The β recombinase from the Streptococcal plasmid pSM 19035 represses its own transcription by holding the RNA polymerase at the promoter region. *Nucleic Acids Res.* 22, 1855–1860. doi: 10.1093/nar/22.10.1855

Romero, H., Rosch, T. C., Hernández-Tamayo, R., Lucena, D., Ayora, S., Alonso, J. C., et al. (2019b). Single molecule tracking reveals functions for RarA at replication forks but also independently from replication during DNA repair in *Bacillus subtilis*. *Sci. Rep.* 9:1997. doi: 10.1038/s41598-018-38289-6

Romero, H., Serrano, E., Hernández-Tamayo, R., Carrasco, B., Cárdenas, P. P., Ayora, S., et al. (2020). *Bacillus subtilis* RarA acts as a positive RecA accessory protein. *Front. Microbiol.* 11:92. doi: 10.3389/fmicb.2020.00092

Romero, H., Torres, R., Hernández-Tamayo, R., Carrasco, B., Ayora, S., Graumann, P. L., et al. (2019a). *Bacillus subtilis* RarA acts at the interplay between replication and repair-by-recombination. *DNA Repair* 78, 27–36. doi: 10.1016/j.dnarep.2019.03.010

Sánchez, H., Carrasco, B., Cozar, M. C., and Alonso, C. J. (2007). *Bacillus subtilis* RecG branch migration translocase is required for DNA repair and chromosomal segregation. *Mol. Microbiol.* 65, 920–935. doi: 10.1111/j.1365-2958.2007.05835.x

Sanchez, H., Kidane, D., Cozar, M. C., Graumann, P. L., and Alonso, C. J. (2006). Recruitment of *Bacillus subtilis* RecN to DNA double-strand breaks in the absence of DNA end processing. *J. Bacteriol.* 188, 353–360. doi: 10.1128/JB.188.2.353-360.2006

Sanchez, H., Kidane, D., Reed, P., Curtis, F. A., Cozar, M. C., Graumann, P. L., et al. (2005). The RuvAB branch migration translocase and RecU Holliday junction resolvase are required for double-stranded DNA break repair in *Bacillus subtilis*. *Genetics* 171, 873–883. doi: 10.1534/genetics.105.045906

Sanders, K., Lin, C. L., Smith, A. J., Cronin, N., Fisher, G., Eftychidis, V., et al. (2017). The structure and function of an RNA polymerase interaction domain in the PcrA/UvrD helicase. *Nucleic Acids Res.* 45, 3875–3887. doi: 10.1093/nar/gkx074

Sedgwick, B. (2004). Repairing DNA-methylation damage. *Nat. Rev. Mol. Cell Biol.* 5, 148–157. doi: 10.1038/nrm1312

Selby, C. P., and Sancar, A. (1993). Molecular mechanism of transcription-repair coupling. *Science* 260, 53–58. doi: 10.1126/science.8465200

Selby, C. P., and Sancar, A. (1994). Mechanisms of transcription-repair coupling and mutation frequency decline. *Microbiol. Rev.* 58, 317–329. doi: 10.1128/MMBR.58.3.317-329.1994

Serrano, E., Carrasco, B., Gilmore, J. L., Takeyasu, K., and Alonso, C. J. (2018). RecA regulation by RecU and DprA during *Bacillus subtilis* natural plasmid transformation. *Front. Microbiol.* 9:1514. doi: 10.3389/fmicb.2018.01514

Shaw, G., Gan, J., Zhou, Y. N., Zhi, H., Subburaman, P., Zhang, R., et al. (2008). Structure of RapA, a Swi2/Snf2 protein that recycles RNA polymerase during transcription. *Structure* 16, 1417–1427. doi: 10.1016/j.str.2008.06.012

Singleton, M. R., Dillingham, M. S., and Wigley, B. D. (2007). Structure and mechanism of helicases and nucleic acid translocases. *Annu. Rev. Biochem.* 76, 23–50. doi: 10.1146/annurev.biochem.76.052305.115300

Subramanya, H. S., Bird, L. E., Brannigan, J. A., and Wigley, B. D. (1996). Crystal structure of a DExx box DNA helicase. *Nature* 384, 379–383. doi: 10.1038/384379a0

Sukhodolets, M. V., Cabrera, J. E., Zhi, H., and Jin, J. D. (2001). RapA, a bacterial homolog of SWI2/SNF2, stimulates RNA polymerase recycling in transcription. *Genes Dev.* 15, 3330–3341. doi: 10.1101/gad.936701

Taucher-Scholtz, G., Abdel-Monem, M., and Hoffmann-Berling, H. (1983). "Functions of helicases in *E. coli*," in *Mechanisms of DNA Replication and Recombination*, ed N. R. Cozzarelli (New York, NY: Alan R. Liss Inc.), 65–76.

Torres, R., Carrasco, B., Gandara, C., Baidya, A. K., Ben-Yehuda, S., and Alonso, C. J. (2019). *Bacillus subtilis* DisA regulates RecA-mediated DNA strand exchange. *Nucleic Acids Res.* 47, 5141–5154. doi: 10.1093/nar/gkz219

Torres, R., Romero, H., Rodríguez-Cerrato, V., C., and Alonso, J. (2017). Interplay between *Bacillus subtilis* RecD2 and the RecG or RuvAB helicase in recombinational repair. *DNA Repair* 55, 40–46. doi: 10.1016/j.dnarep.2017.05.004

Trautinger, B. W., Jaktaji, R. P., Rusakova, E., and Lloyd, G. R. (2005). RNA polymerase modulators and DNA repair activities resolve conflicts between DNA replication and transcription. *Mol. Cell.* 19, 247–258. doi: 10.1016/j.molcel.2005.06.004

Valero-Rello, A., Lopez-Sanz, M., Quevedo-Olmos, A., Sorokin, A., and Ayora, S. (2017). Molecular mechanisms that contribute to horizontal transfer of plasmids by the bacteriophage SPP1. *Front. Microbiol.* 8:1816. doi: 10.3389/fmicb.2017.01816

Veaute, X., Delmas, S., Selva, M., Jeusset, J., Le Cam, E., Matic, I., et al. (2005). UvrD helicase, unlike Rep helicase, dismantles RecA nucleoprotein filaments in *Escherichia coli*. *EMBO J.* 24, 180–189. doi: 10.1038/sj.emboj.7600485

Veaute, X., Jeusset, J., Soustelle, C., Kowalczykowski, S. C., Le Cam, E., and Fabre, F. (2003). The Srs2 helicase prevents recombination by disrupting Rad51 nucleoprotein filaments. *Nature* 423, 309–312. doi: 10.1038/nature01585

Vlasic, I., Mertens, R., Seco, E. M., Carrasco, B., Ayora, S., Reitz, G., et al. (2014). *Bacillus subtilis* RecA and its accessory factors, RecF, RecO, RecR and RecX, are required for spore resistance to DNA double-strand break. *Nucleic Acids Res.* 42, 2295–2307. doi: 10.1093/nar/gkt1194

Wiedermannova, J., Sudzinova, P., Koval, T., Rabatinova, A., Sanderova, H., Ramaniuk, O., et al. (2014). Characterization of HelD, an interacting partner of RNA polymerase from *Bacillus subtilis*. *Nucleic Acids Res.* 42, 5151–5163. doi: 10.1093/nar/gku113

Witkin, E. M. (1969). Ultraviolet-induced mutation and DNA repair. *Annu. Rev. Microbiol.* 23, 487–514. doi: 10.1146/annurev.mi.23.100169.002415

Wu, L. D., and Hickson, I. (2006). DNA helicases required for homologous recombination and repair of damaged replication forks. *Annu. Rev. Genet.* 40, 279–306. doi: 10.1146/annurev.genet.40.110405.090636

Yawn, B., Zhang, L., Mura, C., and Sukhodolets, V. M. (2009). RapA, the SWI/SNF subunit of *Escherichia coli* RNA polymerase, promotes the release of nascent RNA from transcription complexes. *Biochemistry* 48, 7794–7806. doi: 10.1021/bi9004123

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Moreno-del Alamo, Torres, Manfredi, Ruiz-Masó, del Solar and Alonso. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 23 July 2020
doi: 10.3389/fmolb.2020.00160

Gene Duplications in the Genomes of Staphylococci and Enterococci

*José Francisco Sanchez-Herrero[1,2,3], Manuel Bernabeu[1], Alejandro Prieto[1], Mário Hüttener[1] and Antonio Juárez[1,4]**

[1] Department of Genetics, Microbiology and Statistics, University of Barcelona, Barcelona, Spain, [2] Biodiversity Research Institute (IRBio), University of Barcelona, Barcelona, Spain, [3] High Content Genomics and Bioinformatics Unit, Germans Trias i Pujol Research Institute (IGTP), Campus Can Ruti, Badalona, Spain, [4] Institute for Bioengineering of Catalonia, The Barcelona Institute of Science and Technology, Barcelona, Spain

Gene duplications are a feature of bacterial genomes. In the present work we analyze the extent of gene duplications in the genomes of three microorganisms that belong to the Firmicutes phylum and that are etiologic agents of several nosocomial infections: *Staphylococcus aureus*, *Enterococcus faecium*, and *Enterococcus faecalis*. In all three groups, there is an irregular distribution of duplications in the genomes of the strains analyzed. Whereas in some of the strains duplications are scarce, hundreds of duplications are present in others. In all three species, mobile DNA accounts for a large percentage of the duplicated genes: phage DNA in *S. aureus*, and plasmid DNA in the enterococci. Duplicates also include core genes. In all three species, a reduced group of genes is duplicated in all strains analyzed. Duplication of the *deoC* and *rpmG* genes is a hallmark of *S. aureus* genomes. Duplication of the gene encoding the PTS IIB subunit is detected in all enterococci genomes. In *E. faecalis* it is remarkable that the genomes of some strains encode duplicates of the *prgB* and *prgU* genes. They belong to the *prgABCU* cluster, which responds to the presence of the peptide pheromone cCF10 by expressing the surface adhesins PrgA, PrgB, and PrgC.

Keywords: gene duplication, *Staphylococcus aureus, Enterococcus faecium, Enterococcus faecalis*, bacterial genomics

OPEN ACCESS

Edited by:
Tatiana Venkova, Fox Chase Cancer Center, United States

Reviewed by:
Juan Carlos Alonso, National Center for Biotechnology (CNB), Spain
Miguel Angel Cevallos, National Autonomous University of Mexico, Mexico

***Correspondence:**
Antonio Juárez
ajuarez@ub.edu

Specialty section:
This article was submitted to Molecular Recognition, a section of the journal Frontiers in Molecular Biosciences

Received: *19 May 2020*
Accepted: *24 June 2020*
Published: *23 July 2020*

Citation:
Sanchez-Herrero JF, Bernabeu M, Prieto A, Hüttener M and Juárez A (2020) Gene Duplications in the Genomes of Staphylococci and Enterococci. Front. Mol. Biosci. 7:160. doi: 10.3389/fmolb.2020.00160

INTRODUCTION

Gene duplication is an event in which one gene gives rise to two genes that cannot be operationally distinguished from each other. The duplicated genes remain in the same genome. Gene duplications are among the oldest and perhaps the most frequent of mutation types (Lynch et al., 2008; Lipinski et al., 2011). A duplicated gene provides a greater chance for natural selection to shape a novel function (Long et al., 2003). Gene duplication occurs both in eukaryotes and prokaryotes, and significantly impact their gene repertoires, generating functional diversity and increasing the genome complexity (Zhang, 2003; Conant and Wolfe, 2008; Serres et al., 2009; Innan and Kondrashov, 2010; Gao et al., 2017). Duplication events are highly relevant from a biological point of view because, whenever cellular growth is restricted, escape from these growth restrictions can occur by duplication events that resolve the selective problem. In turn, novel duplication events may facilitate subsequent genetic change by allowing cells to proliferate, hence increasing the probability for subsequent adaptive mutations to occur either in the amplified genes or in unrelated ones (Andersson and Hughes, 2009).

In the bacterial kingdom, gene duplication has been associated with survival in extreme or fluctuating conditions, including exposure to antimicrobial compounds or growth on poor nutrient sources, and may have a role in the coevolution between host and pathogens (Romero and Palacios, 1997; Riehle et al., 2001; Duvernay et al., 2011; Kondrashov, 2012; Sun et al., 2012; Toussaint et al., 2017). Several examples correlating gene duplication with bacterial adaptation to the environment are available. For instance, when high gene dosage confers selective benefits, bacteria maintain tandem arrays of duplicated genes (as previously reviewed Romero and Palacios, 1997; Andersson and Hughes, 2009). There is a high natural frequency of bacterial gene duplication, which exceeds the rate of spontaneous point mutation by several orders of magnitude (Andersson and Hughes, 2009). Recent studies indicate that more than 20% of cells in a population contain duplications in some genomic region despite the absence of any evident selection for such duplications (Anderson and Roth, 1981; Hooper and Berg, 2003; Treangen and Rocha, 2011; Elliott et al., 2013).

In most studies the presence of gene duplications is restricted to specific genes or genomic regions, and a global view of the impact of gene duplications in the bacterial genomes is missing. *Escherichia coli* is an example. Previous studies in *E. coli* had shown that some genes such as *flu*, which encodes the adhesin Ag43, can be present in several copies in different strains (van der Woude and Henderson, 2008; Elliott et al., 2013; Arun et al., 2016), but until recently the extent of gene duplications in the genomes of the different types of pathogenic *E. coli* has not been available (Bernabeu et al., 2019). Most pathogenic *E. coli* strains harbor between 80 and 100 duplicated genes. Despite the high genomic diversity of *E. coli*, a group of about 25 genes is duplicated in most of the virulent *E. coli* strains, irrespective of the pathotype to which they belong (Bernabeu et al., 2019). Most of those genes code for proteins of unknown function and, as they are absent from the genomes of commensal strains, their gene products likely play a role in virulence.

In the present report we have undertaken a whole-genome analysis of gene duplications in the genomes of some of the most clinically relevant Gram-positive cocci, namely *Staphylococcus aureus, Enterococcus faecium*, and *Enterococcus faecalis*.

S. aureus and E. faecium are the Gram-positive representatives of the ESKAPE group, which includes microorganisms that are frequent causes of life-threatening nosocomial infections and display multiple antibiotic resistance phenotypes (Murray, 2000; Naimi et al., 2001; Torell et al., 2005). Staphylococcal and enterococcal bacteremia are prevalent in hospitalized patients, and are associated with significant morbidity and mortality (Bartash and Nori, 2017). The emergence of *S. aureus* strains resistant to many antibiotics, including methicillin-resistance (MRSA), poses a serious threat to human health even in countries with well-developed health surveillance systems. Some *E. faecium* and *E. faecalis* isolates account for about 15% of hospital acquired infections in Europe and the US (Werner et al., 2008; Zarb et al., 2012). *E. faecium* infections are nowadays of major concern because of their multidrug resistance phenotypes, including resistance to vancomycin (VRE) and ampicillin. Strains of *E. faecalis* are commensals of the gut microbiota, but under some circumstances they can be pathogenic. Pathogenic strains of *E. faecalis* are increasingly recognized as serious clinical threats due to both the acquisition of multiple antibiotic resistance determinants and to their capacity to disseminate resistance and virulence features by horizontal gene transfer (HGT) mechanisms (Kao and Kline, 2019). Coinfection of MRSA with VRE can occur, being VRE able to transfer VR to the staphylococci (Kos et al., 2012; McGuinness et al., 2017; Cong et al., 2020).

The genomic analysis performed in this work highlights the importance of some genes in the physiology of staphylococci and enterococci. Some of the identified duplicates likely play a role in virulence and hence can be considered as targets of antimicrobial therapies designed to combat infections caused by these pathogens.

MATERIALS AND METHODS

Bacterial Strains and Data Retrieval

We retrieved and analyzed data (genomic fasta, genbank format file, and the translated coding sequences) of all *S. aureus* (n = 473), *E. faecium* (n = 133), and *E. faecalis* (n = 40) complete assembled genomes from NCBI Refseq (**Supplementary Table S1**).

Strategy Used for the Analysis of Duplicates

For each of the strains studied, irrespective of the species, we downloaded the data and analyzed the extent of gene duplications within its genome. Once each strain was analyzed, we summarized the results obtained and generated the corresponding analysis at the level of species. For each of the three species analyzed we selected a reference strain and only for these we created visual representations of their duplicates. Then we analyzed the gene duplications shared with other strains of the same species. We also analyzed further specific characteristics of each of the species analyzed for further interpretation of the data.

Gene Duplication Within Strain

For the analysis of gene duplications we performed an all-vs.-all BLASTp (Altschul et al., 1990) protein similarity search using the translated coding sequence regions and filtering the results with a similarity cutoff $>85\%$, an alignment length between pairs $>85\%$, bit-score >50, and an e-value $<10\text{-}10$. We discarded auto hits and grouped duplicates accordingly.

Analysis of Inserted Phages

We analyzed the putative insertion of phages in the bacterial genomic sequences by using the PhiSpy tool (Akhter et al., 2012) and the genbank format file (gbk) for each of the strains. We used the appropriate training set for each strain according to their species and additional default parameters.

Phylogenetic Reconstruction

We generated a phylogenetic and clusterization analysis using all strains for each of the species analyzed. We used Mash (Ondov et al., 2016) and Sourmash (Titus and Irber, 2016) as they extend a dimensionality-reduction technique to include a pairwise mutation distance enabling the efficient clustering of massive sequence collections. We employed the genomic FASTA files downloaded for each strain and default parameters. For each species, the order of the strains in the different tables follows its phylogenetic relationship.

Exploratory Analysis

We summarized and plotted using R[1] the duplicated count of groups and genes identified and the number of proteins encoding for transposases, selected within the functional annotation associated. We also analyzed the correlation between the amount of duplicated genes and some variables of interest: phages inserted and duplicated proteins annotated as transposases, hypothetical proteins or proteins of unknown function.

Duplicate Coordinates Visualization

For the visualization of duplicates in the corresponding genomes we retrieved, for each duplicate, genomic features such as the start and end coordinates and strand harboring the coding sequence [either from the genomic feature format file (gff) or from the genomic features within the fasta sequence header]. By using R package BioCircos (Cui et al., 2016) we created a circular representation of the duplicate coordinates along the main chromosome and plasmid sequences if any.

Gene Duplications Shared Between Strains

For the analysis of gene duplications shared with other strains of the same species, we selected a single sequence within each duplicated group and we employed BLASTp with filtering parameters as above (similarity >85%, alignment >85%, bit-score >50, and *e*-value <10-10).

Analysis of the Genes Associated to the Generation of Small Colony Variants (SCV)

The analysis of the presence of SCV-associated genes (*deoC*, *sstD*, *plsY*, and *eap)* was done in several staphylococcal strains both coagulase + (*S. aureus*) and coagulase - (*S. epidermidis*, *S. carnosus*, and *S. xylosus*). We downloaded from NCBI Refseq the complete assembled genomes for *S. epidermidis* (n = 29) and any available genome for *S. carnosus* (n = 10) and *S. xylosus* (n = 57) (**Supplementary Table S2**). Homology search was done by using BLASTp and the translated coding sequences of these selected genes (**Supplementary Table S3**) and *Staphylococcus* spp. proteomes. In this case, we used less stringent filtering parameters, a similarity cutoff >50%, following protein homology guidelines (Pearson, 2013), and other parameters as above (alignment >85%, bit-score >50, and *e*-value <10–10).

Analysis of Methicillin and Vancomycin Resistance

Methicillin and vancomycin resistance determinants were searched for each strain by using the Comprehensive Antibiotic Resistance Database (CARD) (Alcock et al., 2020). They were clustered in operons as previously reported (Courvalin, 2006; Shore and Coleman, 2013). We also used BLASTp to identify the presence of each determinant (similarity >80%, alignment >85%, bit-score >50, and *e*-value <10-10) and manually curated the results. The presence of *mecA* or *mecC* genes in *S. aureus* (CARD IDs ARO: 3000617 and 3001209, respectively) was considered to confer the methicillin resistance phenotype (MR) and the presence in *Enterococcus* of *vanA*, *vanB*, and/or *vanG* genes (CARD IDs ARO: 3000010, 3000013, and 3002909, respectively) was considered to confer the vancomycin resistance phenotype (VRE).

Code Availability

The bioinformatics scripts employed for the analysis were deposited and are available at the github website: https://github.com/molevol-ub/BacterialDuplicates.

RESULTS

Gene Duplications in *Staphylococcus* and *Enterococcus* Genomes

We searched first for the presence of duplications in the overall number of 473 *S. aureus* genomes available at the Refseq NCBI database (**Supplementary Table S1**) performing an all-vs.-all protein search by BLASTp for each strain (**Figure 1A** and **Supplementary Table S4**). We also looked for the presence of methicillin resistance determinants and the putative insertion of phages within each genome. From the 473 genomes analyzed, some contain more than 50 groups of duplicates, with up to 190 duplicates. Duplications range from 6 to 84 groups, with more than 50% of the strains encoding more than 26 groups of duplicates. No clear correlation was identified between the total number of duplicates and the number of duplicated transposases annotated (R^2 = 0.332, pval-adj = 1.894e-43). On the other hand, a slight correlation was identified between the total number of duplicates and duplicated hypothetical or proteins of unknown function (R^2 = 0.8, pval-adj = 4.735e-167 and R^2 = 0.532, pval-adj = 7.439e-80, respectively). We also explored the distribution of duplicates under a cutoff of phages inserted. Those strains with >2 phages inserted contain, on average, more duplicates (pval = 0.024).

With regard to *E. faecium,* we searched for the presence of gene duplications within each strain in all 133 genomes available at the NCBI Refseq database (**Supplementary Table S1**) by using the same BLAST strategy described above (**Figure 1B** and **Supplementary Table S5**). From the 133 genomes analyzed some contain more than 50 groups of duplicates, with up to

[1] R-project.org

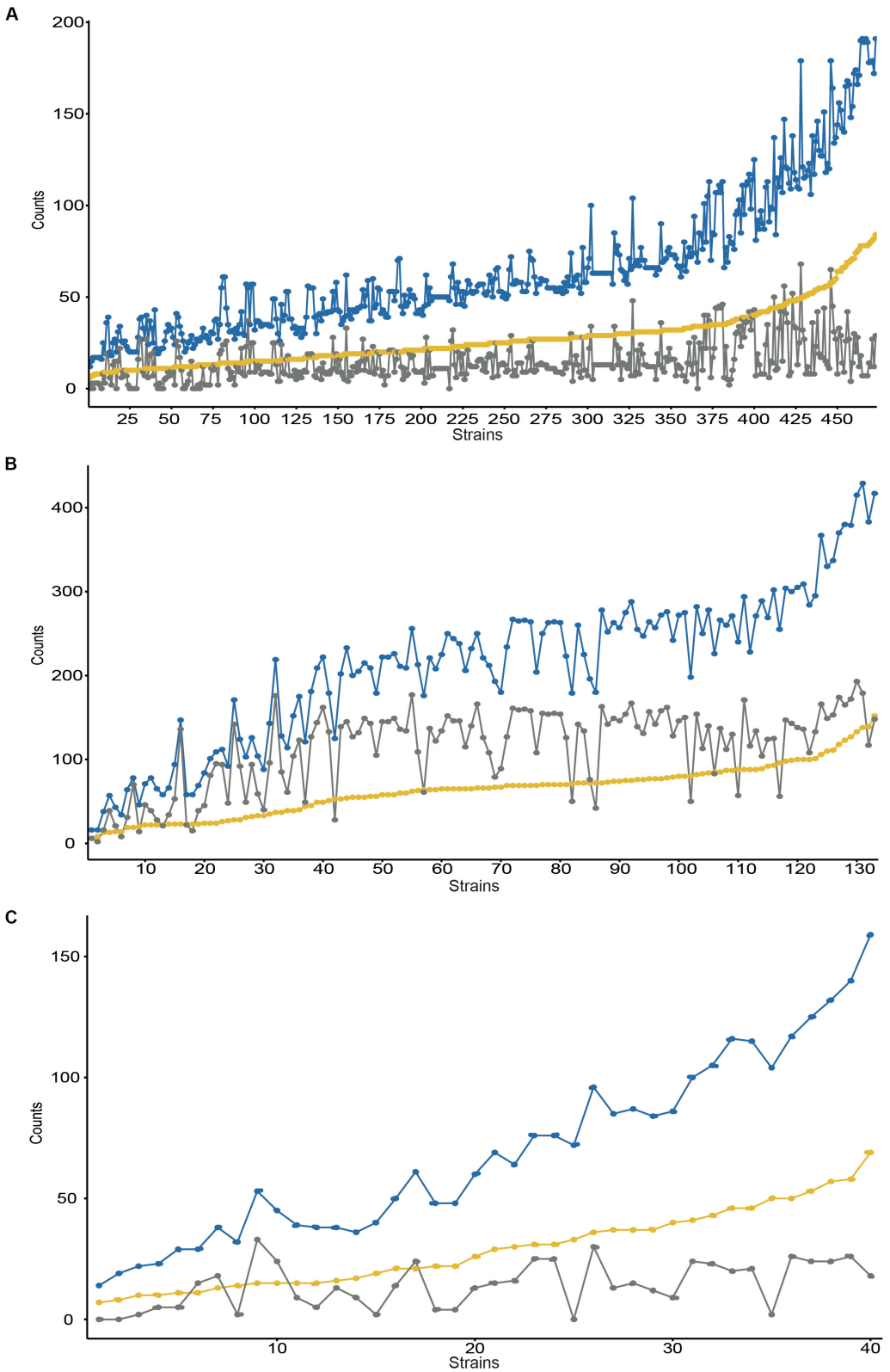


FIGURE 1 | Number of gene duplications identified at the complete assembled genomes for NCBI Refseq entries for *S. aureus* **(A)**, *E. faecalis* **(B)**, and *E. faecium* **(C)**. The Y-axis contains the number of counts for the number of duplicated groups (yellow), the number of gene duplications (blue), and the duplicated transposases (gray). Results are ordered by duplicated groups in increasing order for each strain (axis X).

429 duplicates. Duplications range from 6 to 152 groups, with more than 50% of the strains encoding more than 66 groups of duplicates. A half correlation was identified between the total number of duplicates and the number of duplicated transposases annotated (R^2 = 0.708, pval-adj = 4.828e-37) and with the duplicated hypothetical proteins annotated (R^2 = 0.721, pval-adj = 2.0439e-38). We also explored the distribution of duplicates under a cutoff of phages inserted. Those strains with >4 phages inserted contain, on average, more duplicates (p-val = 0.0017).

To complete our survey of duplications, we searched for the presence of duplications in all 40 *E. faecalis* genomes available (**Supplementary Table S1**) by using the same strategy described above (**Figure 1C** and **Supplementary Table S6**). From the 40 genomes analyzed, some contain more than 50 groups of duplicates, with up to 159 duplicates. Duplications range from 14 to 69 groups, with more than 50% of the strains encoding more than 26 groups of duplicates. No clear correlation was identified between the total number of duplicates and the number of duplicated transposases annotated (R^2 = 0.320, pval-adj = 8.4656e-05). A medium correlation was identified with the duplicated hypothetical proteins annotated (R^2 = 0.787, pval-adj = 1.435e-14). We also explored the distribution of duplicates under a cutoff of phages inserted. Those strains with >4 phages inserted contain, on average, more duplicates (p-val = 0.0014).

Duplications in *S. aureus* Strain Newman

To further study the gene duplications in *S. aureus* genomes, we decided to analyze them in a well-characterized strain such as *S. aureus* Newman. It was isolated in 1952 from a human infection (Duthie and Lorenz, 1952) and has been commonly used as a model strain both for studying *S. aureus* pathogenesis (Richardson et al., 2008; Alonzo et al., 2013) and for the assessment of the therapeutic efficacy of antimicrobial compounds designed to threat *S. aureus* infections (Thammavongsa et al., 2013; Zhang et al., 2014). Its genome sequence has been available since 2008 (Baba et al., 2008).

We analyzed the extent of gene duplications in strain Newman (GCA_000010465.1), and mapped along the Newman genome those genes that are present in two or more copies (**Figure 2** and **Supplementary Table S7**). A total number of 78 genes are duplicated in that strain. Most of the duplicated genes are located in two main regions (**Supplementary Table S7**). The insertion phage analysis identified five putative phages within the chromosome of this strain. Four of this phage coordinates match quite well with four prophages previously described in the Newman strain (Bae et al., 2006; Baba et al., 2008). Many of the genes that are duplicated in this strain are in the same coordinates as these phages (**Figure 2**). Specifically, several duplicates correspond to genes of phages ΦMN4, ΦMN2, and ΦMN1. Some ΦMN4 genes are present in both ΦMN2 and ΦMN1 and hence are present as triplicates. Other ΦMN2 genes are also present in ΦMN1, and are therefore present as duplicates (**Supplementary Table S7**). A small percentage of the duplicates maps outside the phage genomes (**Figure 2**).

With regard to the gene functions of the duplicated genes, 33% correspond to hypothetical proteins, 20% of proteins of unknown function, 19% to phage proteins, and the rest of proteins display miscellaneous functions (**Supplementary Table S7**).

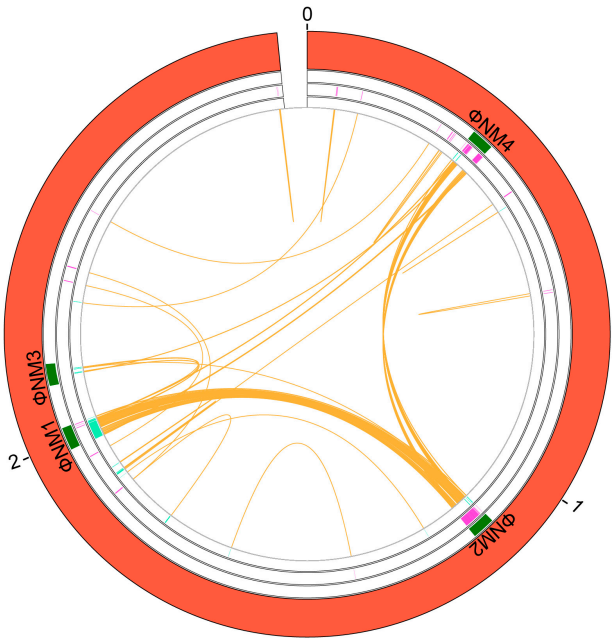


FIGURE 2 | Genes duplicated in the *S. aureus* strain Newman. A circular map of the chromosome showing the duplications is shown. From the outside in, the outer circle represents the chromosome (red). The second circle includes the position of the identified phages that are inserted in the chromosome (green). Next circle shows the duplicated genes in the (+) strand at each coordinate (purple). The innermost circle shows the duplicated genes in the (–) strand at each coordinate (turquoise). The orange color shows the connection between duplicates. The size is shown in Mb.

Duplicated Genes From Strain Newman That Are Also Duplicated in Other *S. aureus* Strains

We addressed next the question as to whether the existing duplicates in this *S. aureus* strain are strain-specific or, on the contrary, they were generated in some putative ancestor and are also present in many other *S. aureus* strains. We used the 473 *S. aureus* genomes to check the shared duplicated genes (see Material and Methods for details). A representative summary of the results obtained is presented in **Figure 3**. The complete analysis is detailed in **Supplementary Table S8**.

As expected, other sequenced Newman strains [strains 412 (GCA_002310435.1), 414 (GCA_002310395.1), and 415 (GCA_900092595.1)] show the same duplication pattern than that obtained with the strain used for the analysis of duplications in the genome [strain 413 (GCA_000010465.1)]. It is also remarkable that strain NCTC13140 [strain 411 (GCA_900474725.1)] shows a gene duplication pattern quite similar to that of the Newman strain. This suggests a close phylogenetic relationship between this strain and the different Newman strains.

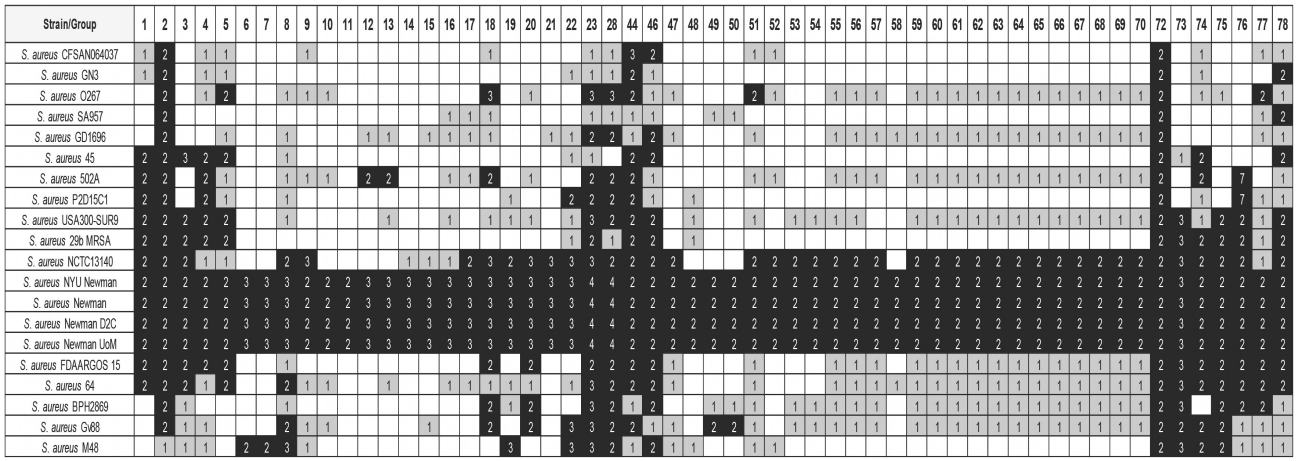


FIGURE 3 | Distribution of some of the detected groups of duplicated genes in strain Newman, in 20 selected *S. aureus* strains. The white, gray, and black colors indicate, respectively, gene absent; gene present in a single copy; gene present in two or more copies. The numbers show the copy number of each gene. The order of the strains follows its phylogenetic relationship. Analysis of the whole set of 473 genomes is shown in **Supplementary Table S8**. Details of the duplicated groups are referred in **Supplementary Table S7**.

The analysis performed also shows that two of the duplicates in strain Newman are also duplicated in most of the strains analyzed (**Table 1**). These genes are *deoC* (group 2) and *rpmG* (group 72). Mutations in *deoC* (codes for the enzyme deoxyribose phosphate aldolase), *prkC* (codes for the serine/threonine-protein kinase PrkC), *plsY* (codes for the enzyme glycerol-3-phosphate acyltransferase), *eap* (codes for an extracellular adherence protein), and *sstD* (codes for an iron-binding protein belonging to an ABC uptake transporter) have been shown to result in the generation of small colony variants (SCV) of *S. aureus* (Chen et al., 2018). We analyzed therefore whether there also existed duplicates of other genes related to the formation of SCVs in all staphylococci (**Supplementary Table S2**). For this analysis, a lower similarity cutoff (>50%) was used in order to detect duplicates with lower similarity (**Supplementary Table S9**). Out of *deoC,* the rest of the genes that have been associated to the generation of SCVs are not duplicated in the genus *Staphylococcus.* Interestingly, *deoC* is duplicated in all *S. aureus* strains analyzed, but not in other catalase negative staphylococci.

The *rpmG* gene codes for the ribosomal protein L33. The existence of duplicates of the genes coding, among other ribosomal proteins, for the ribosomal protein L33 was already reported in some Gram-positive microorganisms (i.e., *Bacillus subtilis*, *B. anthracis*, and *Lactococcus lactis*) as well as in some mycoplasma (Makarova et al., 2001; R Development Core Team, 2008; Kandari et al., 2018).

Out of *deoC* and *rpmG*, another set of genes are duplicated in a group of about 250 strains that are phylogenetically related (**Figure 3** and **Supplementary Table S8**). From these genes, (groups 1, 3, 4, 5, 8, 18–23, 28, 44, 46, 73–78; **Supplementary Table S7**) eight are phage genes (groups: 8, 18–23, 28). From the rest, the function of some is known (**Table 1**). They code, respectively, for a lipoprotein (group 1, gene csa1A), for the M subunit of a restriction/modification system (group44, gene *hsdM*), for the SplF serine protease (group 74, gene *splF*), for a lantibiotic of the gallidemin/nisin family (group 75, gene *bsaA2*), and for a gene belonging to the v*raDEH* operon (group 78, gene vraH), associated to the *S. aureus* resistance to antimicrobial peptides and to cells survival in an infection model (Popella et al., 2016). *S. aureus* Spl proteases are believed to induce allergic reactions (Stentzel et al., 2017).

TABLE 1 | Details of the selected duplicated genes of the strain *Staphylococcus aureus* Newman that are also duplicated in other *S. aureus* strains.

Group	Locus tag 1	Symbol	Description	Percentage
1	NWMN_RS00250	*csa1A*	Tandem type lipoprotein	50.95%
2	NWMN_RS00465	*deoC*	Deoxyribose phosphate aldolase	98.94%
44	NWMN_RS02255	*hsdM*	Type I restriction-modification system subunit M	76.74%
72	NWMN_RS07040	*rpmG*	50S ribosomal protein L33	100.00%
74	NWMN_RS09560	*splF*	Serine protease	53.07%
75	NWMN_RS09635	*bsaA2*	Gallidermin/nisin family lantibiotic	36.79%
78	NWMN_RS14900	*vraH*	Peptide resistance ABC transporter activity modulator	37.00%

Percentage of the analyzed strains that contain the duplicated gene is shown.

Duplications in *E. faecium* Strain 6E6

We used for the study strain 6E6 (GCA_001518735.1), a vancomycin resistant isolate from the University of Minnesota (Geldart and Kaznessis, 2017), that contains a large number of duplicates (n = 337). We identified and mapped those genes that are present in two or more copies in strain 6E6 (see section "Materials and Methods" for details; **Figure 4**

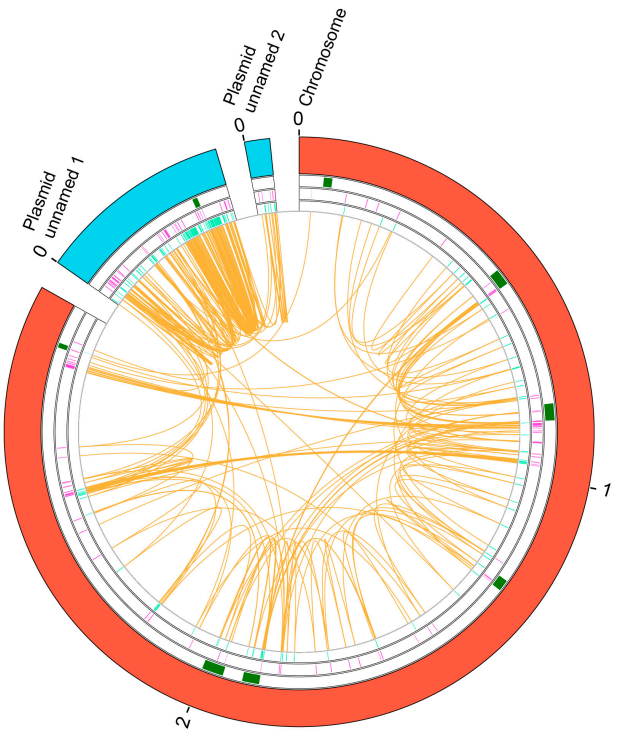


FIGURE 4 | Genes duplicated in the *E. faecium* 6E6 strain. A circular map of the genome of the *E. faecium* 6E6 strain showing the duplicates in the chromosome and in two plasmids is shown. From the outside in, the outer circle represents the chromosome (red) and the plasmids (blue). The second circle includes the position of the identified phages that are inserted in the chromosome (green). Next circle shows the duplicated genes in the (+) strand at each coordinate (purple). The innermost circle shows the duplicated genes in the (–) strand at each coordinate (turquoise). The orange color shows the connection between duplicates. The size is shown in Mb.

and **Supplementary Table S10**). A total number of 111 gene groups are duplicated in that strain. Some of them (23%) correspond to transposases, which are present in three or more copies (**Supplementary Table S10**). Several of the duplicates contain at least one of the copies in a plasmid (**Figure 4**). With regard to the gene functions of the duplicated genes, 41% code for hypothetical proteins, 4.5% code for proteins of unknown function, and the rest for proteins with varying functions (**Supplementary Table S10**). The high percentage of duplicates that code for proteins of unknown function can be correlated with the HGT origin of these genes.

Duplicated Genes From Strain 6E6 That Are Also Duplicated in Other *E. faecium* Strains

Once we determined the gene duplications in the *E. faecium* 6E6 genome, we analyzed whether the existing duplicates are also present in other *E. faecium* strains. We used the 133 *E. faecium* genomes to check the shared duplicated genes. Among the genes that are duplicated in strain 6E6 (**Supplementary Table S10**), a set of 26 genes are also duplicated in most if not all of the strains. A summary of representative data is presented in **Figure 5** (see **Supplementary Table S11** for the complete analysis). These duplicates can be divided in two groups. The first group (18 genes) includes either transposases or transposon-associated genes. At least one of the corresponding copies maps in a plasmid. The second group (eight genes) comprises four other genes that code either for phage proteins or for hypothetical proteins (groups 105–108), and four genes that code for proteins, which display well-characterized physiological functions (**Table 2**). They code, respectively, for a GlsB-like protein (group 103), for a LysM-like protein (group 104), and for two proteins of the PTS

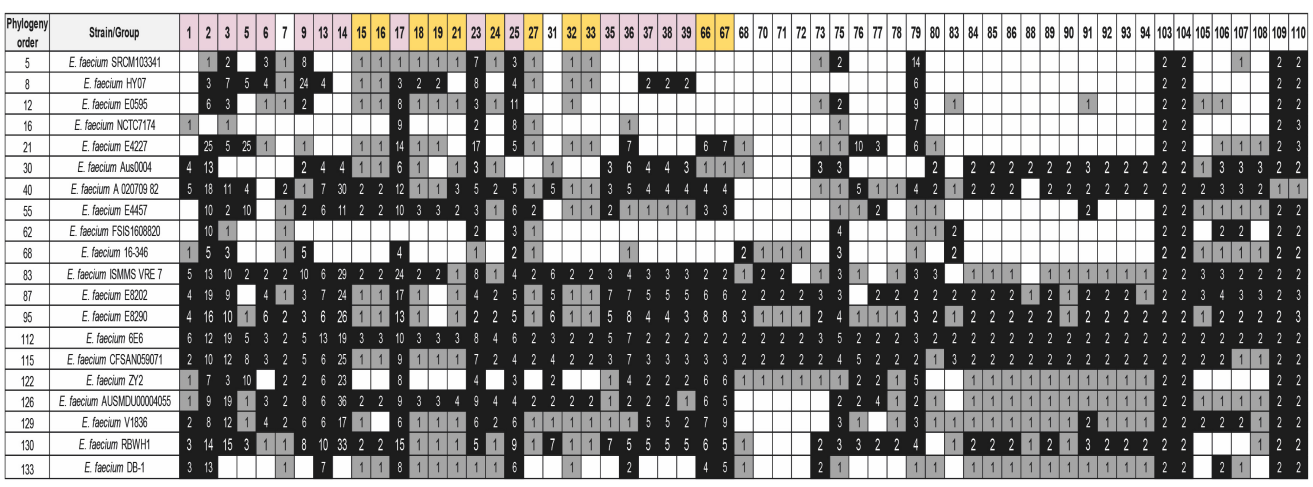


FIGURE 5 | Distribution of some of the detected groups of duplicated genes of strain 6E6 in 20 selected *E. faecium* strains. BLASTp analysis was used for the study. The white, gray, and black colors indicate, respectively, gene absent; gene present in a single copy; gene present in two or more copies. Purple color corresponds to genes with at least one copy in the chromosome and one copy in a plasmid. The gold color corresponds to plasmid genes. The numbers show the copy number of each gene. The order of the strains follows its phylogenetic relationship. Analysis of the whole set of 133 genomes together with their phylogenetic relationship is shown in the **Supplementary Table S11**. Details of the duplicated groups are shown in **Supplementary Table S10**.

TABLE 2 | Details of the selected duplicated genes of the strain *Enterococcus faecium* 6E6 that are also duplicated in other *E. faecium* strains.

Group	Locus Tag 1	Description	Percentage
103	AWJ25_RS06350	GlsB/YeaQ/YmgE family stress response membrane protein	99.25%
104	AWJ25_RS07455	LysM peptidoglycan binding domain containing protein	99.25%
109	AWJ25_RS09645	PTS lactose/cellobiose transporter subunit IIA	95.49%
110	AWJ25_RS09650	PTS sugar transporter subunit IIB	78.20%

Percentage of the analyzed strains that contain the duplicated gene is showed.

system: the lactose/cellobiose IIA and IIB subunits (groups 109 and 110, respectively).

Duplications in the *E. faecalis* Strain V583

We selected for this study as a reference strain *E. faecalis* V583 (GCA_000007785.1), a VanB-type vancomycin-resistant virulent isolate that is a model strain for *E. faecalis* studies (Paulsen et al., 2003).

We analyzed the gene duplications in strain V583 (see Materials and Methods for details), and mapped along the V583 genome those genes that are present in two or more copies (**Figure 6** and **Supplementary Table S12**). A total number of 52 gene groups are duplicated in that strain. Some of them (5.7%) correspond to transposases (**Supplementary Table S12**). As it happened with *E. faecium* strain 6E6, a large number of duplicates (50.9%) are plasmid genes. With regard to the gene functions of the duplicated genes, 18 duplicated genes code for hypothetical proteins (35%), and the rest code for proteins with different functions (**Supplementary Table S12**).

Duplicated Genes From Strain V583 That Are Also Duplicated in Other *E. faecalis* Strains

We also analyzed if the existing duplicates in strain V583 are also present in many other *E. faecalis* strains. We used the 40 *E. faecalis* genomes (**Supplementary Table S1**) to check the shared duplicated genes with strain V583 (see Material and Methods for details) (**Figure 7** and **Supplementary Table S13**). In contrast to *E. faecium*, only a small group of duplicates (five) is present in almost all the strains analyzed. All of them are chromosomal duplicates. The *dgaEF* genes (groups 41 and 42) are required for microbial growth on glucose aminoate (Miller et al., 2013). The genes from group 43 code for the subunit IIB of the PTS system. This gene is also duplicated in *E. faecium*. The genes from group 44 (*galE*) code for the UDP glucose 4 epimerase that

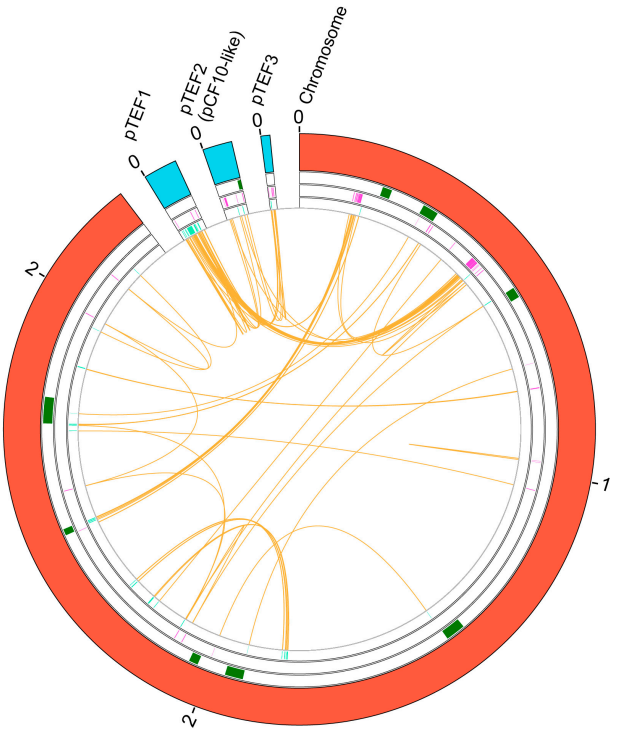


FIGURE 6 | Genes duplicated in the *E. faecalis* V583 strain. A circular map of the genome showing the duplicates in the chromosome and in three plasmids is shown. From the outside in, the outer circle represents the chromosome (red) and plasmids (blue). The second circle shows the position of the identified phages that inserted in the chromosome (green). The next circle shows the duplicated genes in the (+) strand at each coordinate (purple). The innermost circle shows the duplicated genes in the (–) strand at each coordinate (turquoise). The orange color shows the connection between duplicates. The size is shown in Mb.

catalyzes the last step of the Leloir pathway for the assimilation of galactose. The genes from Group 52 code for a protein containing a LPXTG cell wall anchor domain (**Table 3**).

Two other groups of duplicates are present in a significant number of the strains analyzed: groups 1–9, encoded in plasmids, and groups 36–40, encoded in the chromosome. From groups 1–9, it is relevant to mention here the *prgB* and *prgU* genes (groups 4 and 8, respectively). *E. faecalis* strains harboring plasmid pCF10 respond to the presence of the peptide pheromone cCF10 by expressing three surface adhesins: PrgA, PrgB, and PrgC. They play a relevant role in host tissue attachment and biofilm formation (Gilmore et al., 2014). Overexpression of PrgB can be highly toxic to *E. faecalis* cells, and PrgU mitigates toxicity by downregulating PrgB synthesis (Bhatty et al., 2017). It was already reported that strain V583 contains several copies of the *prgU* gene (Bhatty et al., 2017). We show here that this duplication is present in several other *E. faecalis* strains.

With respect to groups 36–40, they code, respectively, for a holin, for a protein containing a LysM peptide-binding domain, for a transposase, for a protein that participates in pectin degradation (the *kdu* gene product) and for the cold shock protein CspA (**Table 3**).

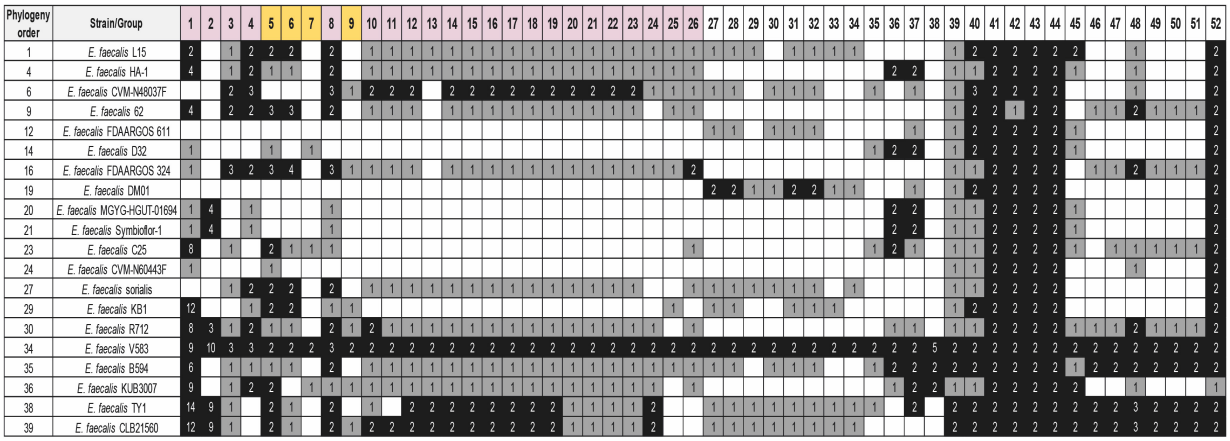


FIGURE 7 | Distribution of the detected groups of duplicated genes of strain V583 in 20 selected *E. faecalis* strains. BLASTp analysis was used for the study. The white, gray, and black colors indicate, respectively, gene absent; gene present in a single copy; gene present in two or more copies. Purple color corresponds to genes with at least one copy in the chromosome and one copy in a plasmid. The gold color corresponds to genes with all copies in plasmids. The numbers show the copy number of each gene. The order of the strains follows its phylogenetic relationship. Analysis of the whole set of 40 genomes together with their phylogenetic relationship is shown in **Supplementary Table S13**. Details of the duplicated groups are shown in **Supplementary Table S12**.

TABLE 3 | Details of the selected duplicated genes of the strain *Enterococcus faecalis* V583 that are also duplicated in other *E. faecalis* strains.

Group	Locus Tag 1	Symbol	Description	Percentage
4	EF_RS00670	*prgB*	LPXTG anchored aggregation substance	32.50%
8	EF_RS02380	*prgU*	Pheromone response system. RNA binding regulator	30.00%
21	EF_RS02400	*prgC*	Cell wall surface anchor family protein	7.50%
26	EF_RS00665	*sea1/prgA*	Surface exclusion protein	5.00%
37	EF_RS01780		LysM peptidoglycan binding domain containing protein	35.00%
39	EF_RS02105	*kduI*	5-dehydro-4-deoxy-D-glucuronate isomerase	17.50%
40	EF_RS03725	*cspA*	Cold shock protein	55.00%
41	EF_RS03995	*dgaE*	Pyridoxal phosphate dependent ammonia lyase family protein	100.00%
42	EF_RS04000	*dgaF*	KDGP aldolase family protein	97.50%
43	EF_RS04815	*celA*	PTS sugar transporter subunit IIB	100.00%
44	EF_RS05170	*galE*	UDP glucose-4-epimerase	100.00%

Percentage of the analyzed strains that contain the duplicated gene is showed.

Strains V583, VE18379, VE14089, and VE18395 that appear to share a set of duplicates, are closely related. Strain VE14089 is plasmid-free V583. Strains VE18379 and VE18395 are derivatives from strain VE14089.

Genomic Context of Some Core Genes That Are Duplicated in These Species

A relevant question to be addressed is whether the identified duplicated genes that code for core functions result either from ancient duplications and are located in fixed points of the chromosome, or have been generated because of some of these genes being flanked by IS elements and jumping to different positions in the chromosome. To assess this, we analyzed the genomic context of both copies of the *deoC* gene in representative strains of *S. aureus*, and of the *celA* gene in representative strains of *E. faecium* and E. *faecalis* (**Figure 8**). The genomic context of the two alleles of the *deoC* and *celA* genes is the same in the different strains analyzed. They are not surrounded by IS elements, but by other core genes.

DISCUSSION

It is apparent that in the Gram-positive cocci studied in this work, mobile DNA elements encode a significant part of the duplicates present in their genomes. The statistical analysis performed correlates duplicates both with phages inserted in the chromosome and with genes encoding either hypothetical or proteins of unknown function (much more common in mobile DNA than in the core genome). In *S. aureus*, most of the duplicated genes are of phage origin. This is not surprising because of the relevant role that phages have in the biology of this microorganism (Xia and Wolz, 2014; Ingmer et al., 2019). In contrast, plasmids are the predominant mobile elements, which encode a significant part of the duplicates in the enterococci. In *E. faecium* there exists a significant correlation between duplicates and transposases. This is also shown when the existing duplicates are identified by BLAST analysis. The question to be addressed is the biological significance of the duplication of genes encoded in mobile DNA in these microorganisms.

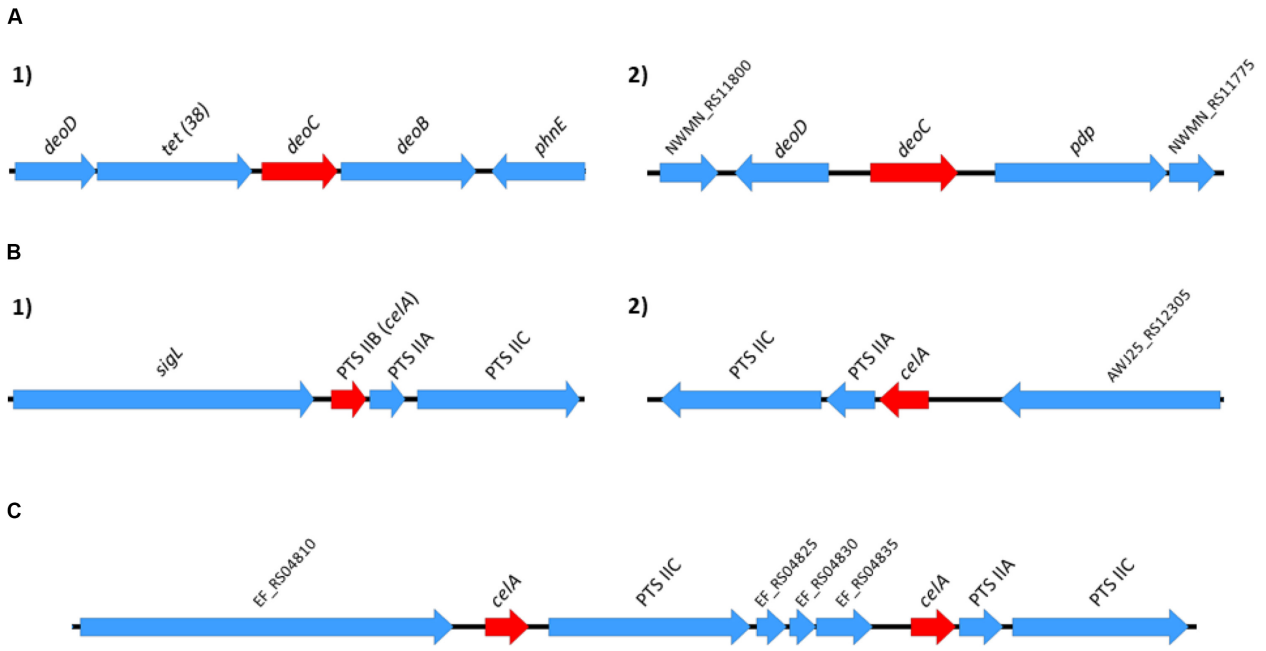


FIGURE 8 | Genomic context of the *S. aureus deoC* duplicated genes **(A)** and of the *E. faecium* and *E. faecalis celA* duplicated genes (**B,C**, respectively). The genomic context was analyzed in the following representative strains: *S. aureus* Newman, *S. aureus* N315, *S. aureus* USA300, and *S. aureus* MW2; *E. faecium* 6E6, *E. faecium* Aus0004 and *E. faecium* DO; *E. faecalis* V583, *E. faecalis* OG1RF and *E. faecalis* Symbioflor. Red arrows in positions 1 and 2 correspond to the alleles of the duplicated genes. The genomic context for the *deoC* and *celA* genes was identical in all *S. aureus* and *E. faecium/E. faecalis* strains analyzed.

Although duplicates located in mobile DNA predominate in the microorganisms studied here, others are located in the chromosome. Some of these latter duplicates are widespread among all the strains of the same species analyzed. In *S. aureus,* two duplicates are present in almost all strains analyzed: the *rpmG* and the *deoC* genes. The former codes for the ribosomal protein L33. Duplications of the gene encoding the L33 protein appear to be a hallmark of several Gram-positive genera. As a general rule, one of the *rmpG* paralogs codes for a protein that contains a Zn-binding motif comprising a two pair of conserved CXXC stretch (CC form), which is absent in the other (C- form). In strain Newman, two copies are C- (NWMN_RS07040, and NWMN_RS08205, respectively), and the third is CC [NWMN_0496.1 (this latter shows less than 85% identity)]. In addition to their role in translation, ribosomes also serve as reservoirs for zinc in the cell (Moore and Helmann, 2005). The zinc-responsive regulator Zur has been shown to repress the C- form (Akanuma et al., 2006; Gabriel and Helmann, 2009). Under zinc-depleted conditions, the Zur mediated repression of the genes encoding the C- forms of the ribosomal proteins is alleviated. These C- forms then replace the corresponding CC forms from the ribosomes, resulting in exoneration of zinc, which can then be used by other metalloproteins. This enables the bacterial cell to survive in zinc limiting environments (Moore et al., 2005; Akanuma et al., 2006).

The *deoC* gene product is the deoxyribose phosphate aldolase, which enables bacterial cells to grow on deoxyribonucleosides as the carbon source. As commented above, mutations in the *S. aureus deoC* gene have been associated with the generation of SCVs (Chen et al., 2018). Interestingly, *deoC* mutations were associated with alterations in the response to extracellular signaling in *E. coli* (Joloba and Rather, 2003). It can be hypothesized that, as the *deoC* gene is duplicated in *S. aureus* but not in other catalase negative cocci, its gene product can play a role in *S. aureus* virulence.

In addition to these widespread duplicates, another group of duplicates is present in a subset of the *S. aureus* genomes. The strains containing these duplicates are phylogenetically related. The reported functions for these genes (i.e., *vraH*, *splF*) are also related to virulence, and they can be considered as virulence markers of that group of *S. aureus* strains. A question to be addressed is whether the duplication of these genes confers specific virulence features to *S. aureus*.

In *E. faecium* genomes, a significant part of the duplicates (58%) are located (at least one of the copies) in plasmids. Several of these duplicates code either for transposases or for hypothetical proteins. Some of them are shared by most of the strains analyzed. In addition to these genes of plasmid origin, four chromosomal genes are also duplicated in most of the *E. faecium* strains analyzed. Expression of GlsB proteins has been associated with virulence and bile salt stress (Choudhury et al., 2011; Zhang et al., 2013a). Proteins containing a LysM domain have been shown to be induced under infection conditions of a mammalian host (Cacaci et al., 2018). Although the proteins of the lactose/cellobiose PTS system IIA and IIB have not been hitherto described as relevant elements in *E. faecium* virulence, the relevance of the PTS system for the ability of *E. faecium* to colonize the host has been previously reported. Deletion of the

pstD gene, which is predicted to encode the enzyme IID subunit of a PTS system, influenced *E. faecium* virulence (Zhang et al., 2013b). Furthermore, insertional inactivation of the *bepA* gene, coding for putative a PTS permease, was found to be relevant for *E. faecium* pathogenesis (Paganelli et al., 2016). The fact that, as shown in this report, other PTS specific components are duplicated in *E. faecium* strains further highlights the role of the phosphotransferase system in *E. faecium* physiology and hence, in the ability of virulent strains to colonize their hosts.

As it happens in *E. faecium,* about half of the *E. faecalis* duplicates are located (at least one of the copies) in a plasmid. Nevertheless, in contrast to *E. faecium*, few of the *E. faecalis* duplicates that are located in plasmids are transposases. A cluster of duplicated plasmid genes (*prgABCU*) are of special relevance. Its gene products enable *E. faecalis* to respond to the presence of the peptide pheromone cCF10 by expressing the surface adhesins PrgA, PrgB, and PrgC (Gilmore et al., 2014). It has been suggested that *prgU* expression controls *prgB* expression, avoiding that excess of the PrgB protein can be deleterious for the cell. *prgB* and *prgU* are present in several copies in strain V583, and there exists a genetic linkage between both genes (Bhatty et al., 2017). We show in this work that gene duplication occurs predominantly with both the *prgB* and *prgU* genes, and not with *prgA* and *prgC* genes. Our data are hence consistent with the genetic linkage of *prgB a*nd *prgU* (Bhatty et al., 2017). *prgU* genes are widely distributed on plasmids and chromosomes of *E. faecalis* and other enterococci, and it has been suggested that the *prgB-prgU* genetic linkage might have evolved to ensure the controlled synthesis of PrgB-like adhesins (Bhatty et al., 2017). Accordingly, we show here that the genetic linkage of *prgB-prgU* also involves gene duplication. In accordance with the rule that we suggested previously for *E. coli* (Bernabeu et al., 2019), the duplication of the regulated gene (*prgB*) correlates with the duplication of its modulator (*prgU*).

In *E. faecalis* there also exists a group of duplicates located in the chromosome that is shared by all the strains analyzed. Three of them code for proteins playing a role in cell metabolism, including the duplication of the gene coding for of the subunit IIB of the lactose/cellobiose PTS system. Different components of the PTS system have also been reported as relevant for *E. faecalis* colonization (Paulsen et al., 2003), and they have deserved special attention in the last years (Ruiz-Cruz et al., 2015; Sauvageot et al., 2017; Grand et al., 2019).

Both for *S. aureus* and *Enterococcus*, the genomic context of the duplicates coding for core genes is similar among strains, and corresponds to other core genes. This fact, together with the widespread distribution of these genes among all strains analyzed suggests that these duplications correspond to ancient events that have been positively selected in the course of evolution.

Genomic analysis is powerful to gain insight into several aspects of the biology of organisms. We show here that the analysis of the pattern of gene duplications in microorganisms can provide relevant information that can be useful for both establishing phylogenetic relationships between strains, and for the identification of genes that can play relevant roles in, among other processes, bacterial virulence.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

JS-H, MB, and AJ designed this study. JS-H and MB performed the *in silico* work. JS-H, MB, AP, MH, and AJ wrote the manuscript. MH and AJ did the final version of the manuscript. All authors analyzed and discussed the results, read and approved the final manuscript.

FUNDING

JS-H was beneficiary of a contract by the Instituto de Salud Carlos III through the Accion Estrategica en Salud 2018 (Co-funded by the European Regional Development Fund/European Social Fund; "A way to make Europe"/"Investing in your future"). MB was the recipient of a FI Fellowship from the Generalitat de Catalunya. This work was supported by grant BIO2016-76412-C2-1-R (AEI/FEDER, UE) from the Ministerio de Economía, Industria y Competitividad, Fundació La Marató TV3, Spain (Project ref. 201818 10), and the CERCA Programme/Generalitat de Catalunya to AJ.

ACKNOWLEDGMENTS

We thank the IGTP High Content Genomics & Bioinformatics Unit Core Facility and staff (Dr. Lauro Sumoy) for their contribution to this publication. We additionally thank the Evolutionary Genomics and Bioinformatics Group and Dr. Julio Rozas particularly at the University Barcelona for providing us with computational resources for this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2020.00160/full#supplementary-material

TABLE S1 | List of the complete assembled genomes corresponding to NCBI Refseq entries for *S. aureus*, *E. faecalis*, and *E. faecium* used for the gene duplication analysis.

TABLE S2 | List of the NCBI Refseq entries for *S. epidermidis*, *S. carnosus*, and *S. xylosus* used for the analysis of the duplications of genes involved in the generation of small colony variants (SCV).

TABLE S3 | Details of the genes involved in the generation of SCVs used for the duplication analysis.

TABLE S4 | Gene duplication data obtained for the *S. aureus* strains analyzed. Indicated are: (i) Phylogeny order: according to the clusterization analysis using mash and sourmash tools. (ii) Groups: counts of groups of duplicated genes. (iii) Dups: total count of duplicated genes. (iv) Transpo: counts of genes within duplicates annotated as "transposase". (v) Hypo: counts of gene names within duplications annotated or containing "hypothetical" functions. (vi) Phages: Number

of putative phages identified with PhiSpy inserted in the genome. (vii) MRSA: presence of methicillin resistance determinants *mecA* and/or *mecC*. The order of the strains follows its phylogenetic relationship.

TABLE S5 | Gene duplication data obtained for the *E. faecium* strains analyzed. Indicated are: (i) Phylogeny order: according to the clusterization analysis using mash and sourmash tools. (ii) Groups: counts of groups of duplicated genes. (iii) Dups: total count of duplicated genes. (iv) Transpo: counts of genes within duplicates annotated as "transposase". (v) Hypo: counts of gene names within duplications annotated or containing "hypothetical" functions. (vi) Phages: Number of putative phages identified with PhiSpy inserted in the genome. (vii) VRE: presence of vancomycin resistance determinants *vanA* (blue), *vanB* (yellow), *vanA*/*vanB* (green), or *vanG* (red). The order of the strains follows its phylogenetic relationship.

TABLE S6 | Gene duplication data obtained for the *E. faecalis* strains analyzed. Indicated are: (i) Phylogeny order: according to the clusterization analysis using mash and sourmash tools. (ii) Groups: counts of groups of duplicated genes. (iii) Dups: total count of duplicated genes. (iv) Transpo: counts of genes within duplicates annotated as "transposase". (v) Hypo: counts of gene names within duplications annotated or containing "hypothetical" functions. (vi) Phages: Number of putative phages identified with PhiSpy inserted in the genome. (vii) VRE: presence of vancomycin resistance determinants *vanA* (blue) or *vanB* (yellow). The order of the strains follows its phylogenetic relationship.

TABLE S7 | Locus tag and description of the duplicates identified in *S. aureus* strain Newman. Red, blue and green colors correspond to genes located in the ΦMN4, ΦMN2, and ΦMN1 phages, respectively.

TABLE S8 | Distribution of the detected duplicated genes of strain Newman in the genomes of the other 472 *S. aureus* strains analyzed. BLASTp analysis was used for the study. The white, gray, and black colors indicate, respectively: gene absent; gene present in a single copy; gene present in two or more copies. The numbers show the copy number of each gene. The order of the strains follows its phylogenetic relationship.

TABLE S9 | Analysis of the duplications of genes *deoC*, *sstD*, *plsY*, and *eap* in *S. aureus*, *S. epidermidis*, *S. carnosus*, and *S. xylosus*.

TABLE S10 | Locus tag and description of the duplicates identified in *E. faecium* strain 6E6. The purple and gold colors correspond, respectively, to genes with at least one copy in chromosome and one in plasmids, and to genes with all copies in plasmids.

TABLE S11 | Distribution of the detected duplicated genes of strain 6E6 in the genomes of the other 132 *E. faecium* strains analyzed. BLASTp analysis was used for the study. The white, gray, and black colors indicate, respectively: gene absent; gene present in a single copy; gene present in two or more copies. Purple color corresponds to genes with at least one copy in the chromosome and one copy in a plasmid. The gold color corresponds to genes with all copies in plasmids. The numbers show the copy number of each gene. The order of the strains follows its phylogenetic relationship.

TABLE S12 | Locus tag and description of the duplicates identified in *E. faecalis* strain V583. The purple and gold colors correspond, respectively, to genes with at least one copy in chromosome and one in plasmids, and to genes with all copies in plasmids.

TABLE S13 | Distribution of the detected duplicated genes of strain V583 in the genomes of the other 39 *E. faecalis* strains analyzed. BLASTp analysis was used for the study. The white, gray, and black colors indicate, respectively: gene absent; gene present in a single copy; gene present in two or more copies. Purple color corresponds to genes with at least one copy in the chromosome and one copy in a plasmid. The gold color corresponds to genes with all copies in plasmids. The numbers show the copy number of each gene. The order of the strains follows its phylogenetic relationship.

REFERENCES

Akanuma, G., Nanamiya, H., Natori, Y., Nomura, N., and Kawamura, F. (2006). Liberation of zinc-containing L31 (RpmE) from ribosomes by its paralogous gene product, YtiA, in *Bacillus subtilis*. *J. Bacteriol.* 188, 2715–2720. doi: 10.1128/JB.188.7.2715-2720.2006

Akhter, S., Aziz, R. K., and Edwards, R. A. (2012). PhiSpy: a novel algorithm for finding prophages in bacterial genomes that combines similarity- and composition-based strategies. *Nucleic Acids Res.* 40, e126. doi: 10.1093/nar/gks406

Alcock, B. P., Raphenya, A. R., Lau, T. T. Y., Tsang, K. K., Bouchard, M., Edalatmand, A., et al. (2020). CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res.* 48, D517–D525. doi: 10.1093/nar/gkz935

Alonzo, F., Kozhaya, L., Rawlings, S. A., Reyes-Robles, T., DuMont, A. L., Myszka, D. G., et al. (2013). CCR5 is a receptor for *Staphylococcus aureus* leukotoxin ED. *Nature* 493, 51–55. doi: 10.1038/nature11724

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2

Anderson, P., and Roth, J. (1981). Spontaneous tandem genetic duplications in *Salmonella* Typhimurium arise by unequal recombination between rRNA (rrn) cistrons. *Proc. Natl. Acad. Sci. U.S.A.* 78, 3113–3117. doi: 10.1073/pnas.78.5.3113

Andersson, D. I., and Hughes, D. (2009). Gene amplification and adaptive evolution in bacteria. *Annu. Rev. Genet.* 43, 167–195. doi: 10.1146/annurev-genet-102108-134805

Arun, P. V. P. S., Miryala, S. K., Chattopadhyay, S., Thiyyagura, K., Bawa, P., Bhattacharjee, M., et al. (2016). Identification and functional analysis of essential, conserved, housekeeping and duplicated genes. *FEBS Lett.* 590, 1428–1437. doi: 10.1002/1873-3468.12192

Baba, T., Bae, T., Schneewind, O., Takeuchi, F., and Hiramatsu, K. (2008). Genome sequence of Staphylococcus aureus strain Newman and comparative analysis of staphylococcal genomes: polymorphism and evolution of two major pathogenicity islands. *J. Bacteriol.* 190, 300–310. doi: 10.1128/jb.01000-07

Bae, T., Baba, T., Hiramatsu, K., and Schneewind, O. (2006). Prophages of *Staphylococcus aureus* Newman and their contribution to virulence. *Mol. Microbiol.* 62, 1035–1047. doi: 10.1111/j.1365-2958.2006.05441.x

Bartash, R., and Nori, P. (2017). Beta-lactam combination therapy for the treatment of *Staphylococcus aureus* and *Enterococcus* species bacteremia: a summary and appraisal of the evidence. *Int. J. Infect. Dis.* 63, 7–12. doi: 10.1016/j.ijid.2017.07.019

Bernabeu, M., Sánchez-Herrero, J. F., Huedo, P., Prieto, A., Hüttener, M., Rozas, J., et al. (2019). Gene duplications in the *E. coli* genome: common themes among pathotypes. *BMC Genomics* 20:313. doi: 10.1186/s12864-019-5683-4

Bhatty, M., Camacho, M. I., Gonzalez-Rivera, C., Frank, K. L., Dale, J. L., Manias, D. A., et al. (2017). PrgU: a suppressor of sex pheromone toxicity in Enterococcus faecalis. *Mol. Microbiol.* 103, 398–412. doi: 10.1111/mmi.13563

Cacaci, M., Giraud, C., Leger, L., Torelli, R., Martini, C., Posteraro, B., et al. (2018). Expression profiling in a mammalian host reveals the strong induction of genes encoding LysM domain-containing proteins in *Enterococcus faecium*. *Sci. Rep.* 8:12412. doi: 10.1038/s41598-018-30882-z

Chen, H., Wang, Q., Yin, Y., Li, S., Niu, D.-K., and Wang, H. (2018). Genotypic variations between wild-type and small colony variant of *Staphylococcus aureus* in prosthetic valve infectious endocarditis: a comparative genomic and transcriptomic analysis. *Int. J. Antimicrob. Agents* 51, 655–658. doi: 10.1016/j.ijantimicag.2017.12.006

Choudhury, T., Singh, K. V., Sillanpää, J., Nallapareddy, S. R., and Murray, B. E. (2011). Importance of two *Enterococcus faecium* loci encoding Gls-like proteins for in vitro bile salts stress response and virulence. *J. Infect. Dis.* 203, 1147–1154. doi: 10.1093/infdis/jiq160

Conant, G. C., and Wolfe, K. H. (2008). Turning a hobby into a job: how duplicated genes find new functions. *Nat. Rev. Genet.* 9, 938–950. doi: 10.1038/nrg2482

Cong, Y., Yang, S., and Rao, X. (2020). Vancomycin resistant Staphylococcus aureus infections: a review of case updating and clinical features. *J. Adv. Res.* 21, 169–176. doi: 10.1016/j.jare.2019.10.005

Courvalin, P. (2006). Vancomycin resistance in gram-positive cocci. *Clin. Infect. Dis.* 42(Suppl. 1), S25–S34. doi: 10.1086/491711

Cui, Y., Chen, X., Luo, H., Fan, Z., Luo, J., He, S., et al. (2016). BioCircos.js: an interactive Circos JavaScript library for biological data visualization on web applications. *Bioinformatics* 32, 1740–1742. doi: 10.1093/bioinformatics/btw041

Duthie, E. S., and Lorenz, L. L. (1952). Staphylococcal coagulase; mode of action and antigenicity. *J. Gen. Microbiol.* 6, 95–107.

Duvernay, C., Coulange, L., Dutilh, B., Dubois, V., Quentin, C., and Arpin, C. (2011). Duplication of the chromosomal blaSHV-11 gene in a clinical hypermutable strain of *Klebsiella pneumoniae*. *Microbiology (Reading, Engl.)* 157, 496–503. doi: 10.1099/mic.0.043885-0

Elliott, K. T., Cuff, L. E., and Neidle, E. L. (2013). Copy number change: evolving views on gene amplification. *Future Microbiol.* 8, 887–899. doi: 10.2217/fmb.13.53

Gabriel, S. E., and Helmann, J. D. (2009). Contributions of Zur-controlled ribosomal proteins to growth under zinc starvation conditions. *J. Bacteriol.* 191, 6116–6122. doi: 10.1128/jb.00802-09

Gao, Y., Zhao, H., Jin, Y., Xu, X., and Han, G.-Z. (2017). Extent and evolution of gene duplication in DNA viruses. *Virus Res.* 240, 161–165. doi: 10.1016/j.virusres.2017.08.005

Geldart, K., and Kaznessis, Y. N. (2017). Characterization of class iia bacteriocin resistance in enterococcus faecium. *Antimicrob. Agents Chemother.* 61: e2033-16.

Gilmore, M. S., Clewell, D. B., Ike, Y., Shankar, N., Clewell, D. B., Weaver, K. E., et al. (2014). *Extrachromosomal and Mobile Elements in Enterococci: Transmission, Maintenance, and Epidemiology*. Boston, MA: Massachusetts Eye and Ear Infirmary.

Grand, M., Aubourg, M., Pikis, A., Thompson, J., Deutscher, J., Hartke, A., et al. (2019). Characterization of the gen locus involved in β-1,6-oligosaccharide utilization by *Enterococcus faecalis*. *Mol. Microbiol.* 112, 1744–1756. doi: 10.1111/mmi.14390

Hooper, S. D., and Berg, O. G. (2003). Duplication is more common among laterally transferred genes than among indigenous genes. *Genome Biol.* 4:R48. doi: 10.1186/gb-2003-4-8-r48

Ingmer, H., Gerlach, D., and Wolz, C. (2019). Temperate phages of *Staphylococcus aureus*. *Microbiol Spectr* 7:GPP-3-0058-2018.

Innan, H., and Kondrashov, F. (2010). The evolution of gene duplications: classifying and distinguishing between models. *Nat. Rev. Genet.* 11, 97–108. doi: 10.1038/nrg2689

Joloba, M. L., and Rather, P. N. (2003). Mutations in deoB and deoC alter an extracellular signaling pathway required for activation of the gab operon in *Escherichia coli*. *FEMS Microbiol. Lett.* 228, 151–157. doi: 10.1016/s0378-1097(03)00754-7

Kandari, D., Gopalani, M., Gupta, M., Joshi, H., Bhatnagar, S., and Bhatnagar, R. (2018). Identification, functional characterization, and regulon prediction of the zinc uptake regulator (zur) of *Bacillus anthracis* – An insight into the zinc homeostasis of the pathogen. *Front. Microbiol.* 9:3314. doi: 10.3389/fmicb.2018.03314

Kao, P. H. N., and Kline, K. A. (2019). Dr. Jekyll and Mr. Hide: how *Enterococcus faecalis* subverts the host immune response to cause infection. *J. Mol. Biol.* 431, 2932–2945. doi: 10.1016/j.jmb.2019.05.030

Kondrashov, F. A. (2012). Gene duplication as a mechanism of genomic adaptation to a changing environment. *Proc. Biol. Sci.* 279, 5048–5057. doi: 10.1098/rspb.2012.1108

Kos, V. N., Desjardins, C. A., Griggs, A., Cerqueira, G., Van Tonder, A., Holden, M. T. G., et al. (2012). Comparative genomics of vancomycin-resistant Staphylococcus aureus strains and their positions within the clade most commonly associated with Methicillin-resistant S. *aureus* hospital-acquired infection in the United States. *MBio* 3:e112.

Lipinski, K. J., Farslow, J. C., Fitzpatrick, K. A., Lynch, M., Katju, V., and Bergthorsson, U. (2011). High spontaneous rate of gene duplication in *Caenorhabditis elegans*. *Curr. Biol.* 21, 306–310. doi: 10.1016/j.cub.2011.01.026

Long, M., Betrán, E., Thornton, K., and Wang, W. (2003). The origin of new genes: glimpses from the young and old. *Nat. Rev. Genet.* 4, 865–875. doi: 10.1038/nrg1204

Lynch, M., Sung, W., Morris, K., Coffey, N., Landry, C. R., Dopman, E. B., et al. (2008). A genome-wide view of the spectrum of spontaneous mutations in yeast. *Proc. Natl. Acad. Sci. U.S.A.* 105, 9272–9277. doi: 10.1073/pnas.0803466105

Makarova, K. S., Ponomarev, V. A., and Koonin, E. V. (2001). Two C or not two C: recurrent disruption of Zn-ribbons, gene duplication, lineage-specific gene loss, and horizontal gene transfer in evolution of bacterial ribosomal proteins. *Genome Biol.* 2:RESEARCH0033. doi: 10.1186/gb-2001-2-9-research0033

McGuinness, W. A., Malachowa, N., and DeLeo, F. R. (2017). Vancomycin resistance in *Staphylococcus aureus*. *Yale J Biol Med* 90, 269–281.

Miller, K. A., Phillips, R. S., Mrázek, J., and Hoover, T. R. (2013). *Salmonella* utilizes D-glucosaminate via a mannose family phosphotransferase system permease and associated enzymes. *J. Bacteriol.* 195, 4057–4066. doi: 10.1128/jb.00290-13

Moore, C. M., Gaballa, A., Hui, M., Ye, R. W., and Helmann, J. D. (2005). Genetic and physiological responses of Bacillus subtilis to metal ion stress. *Mol. Microbiol.* 57, 27–40. doi: 10.1111/j.1365-2958.2005.04642.x

Moore, C. M., and Helmann, J. D. (2005). Metal ion homeostasis in *Bacillus subtilis*. *Curr. Opin. Microbiol.* 8, 188–195. doi: 10.1016/j.mib.2005.02.007

Murray, B. E. (2000). Vancomycin-resistant enterococcal infections. *N. Engl. J. Med.* 342, 710–721. doi: 10.1056/NEJM200003093421007

Naimi, T. S., LeDell, K. H., Boxrud, D. J., Groom, A. V., Steward, C. D., Johnson, S. K., et al. (2001). Epidemiology and clonality of community-acquired methicillin-resistant *Staphylococcus aureus* in Minnesota, 1996–1998. *Clin. Infect. Dis.* 33, 990–996. doi: 10.1086/322693

Ondov, B. D., Treangen, T. J., Melsted, P., Mallonee, A. B., Bergman, N. H., Koren, S., et al. (2016). Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biol.* 17:132. doi: 10.1186/s13059-016-0997-x

Paganelli, F. L., Huebner, J., Singh, K. V., Zhang, X., van Schaik, W., Wobser, D., et al. (2016). Genome-wide screening identifies phosphotransferase system permease BepA to be involved in *Enterococcus faecium* endocarditis and biofilm formation. *J. Infect. Dis.* 214, 189–195. doi: 10.1093/infdis/jiw108

Paulsen, I. T., Banerjei, L., Myers, G. S. A., Nelson, K. E., Seshadri, R., Read, T. D., et al. (2003). Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*. *Science* 299, 2071–2074. doi: 10.1126/science.1080613

Pearson, W. R. (2013). An introduction to sequence similarity ("homology") searching. *Curr. Protoc. Bioinformat. Chapter* 3:Unit3.1. doi: 10.1002/0471250953.bi0301s42

Popella, P., Krauss, S., Ebner, P., Nega, M., Deibert, J., and Götz, F. (2016). VraH is the third component of the *Staphylococcus aureus* VraDEH system involved in gallidermin and daptomycin resistance and pathogenicity. *Antimicrob. Agents Chemother.* 60, 2391–2401. doi: 10.1128/aac.02865-15

R Development Core Team (2008). *R Development Core Team R-project.org*. Available online at: http://www.R-project.org.

Richardson, A. R., Libby, S. J., and Fang, F. C. (2008). A nitric oxide-inducible lactate dehydrogenase enables *Staphylococcus aureus* to resist innate immunity. *Science* 319, 1672–1676. doi: 10.1126/science.1155207

Riehle, M. M., Bennett, A. F., and Long, A. D. (2001). Genetic architecture of thermal adaptation in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 98, 525–530. doi: 10.1073/pnas.021448998

Romero, D., and Palacios, R. (1997). Gene amplification and genomic plasticity in prokaryotes. *Annu. Rev. Genet.* 31, 91–111. doi: 10.1146/annurev.genet.31.1.91

Ruiz-Cruz, S., Espinosa, M., Goldmann, O., and Bravo, A. (2015). Global regulation of gene expression by the MafR protein of *Enterococcus faecalis*. *Front. Microbiol.* 6:1521. doi: 10.3389/fmicb.2015.01521

Sauvageot, N., Mokhtari, A., Joyet, P., Budin-Verneuil, A., Blancato, V. S., Repizo, G. D., et al. (2017). *Enterococcus faecalis* uses a phosphotransferase system permease and a host colonization-related ABC transporter for maltodextrin uptake. *J. Bacteriol.* 199:e00878-16.

Serres, M. H., Kerr, A. R. W., McCormack, T. J., and Riley, M. (2009). Evolution by leaps: gene duplication in bacteria. *Biol. Direct.* 4:46. doi: 10.1186/1745-6150-4-46

Shore, A. C., and Coleman, D. C. (2013). Staphylococcal cassette chromosome mec: recent advances and new insights. *Int. J. Med. Microbiol.* 303, 350–359. doi: 10.1016/j.ijmm.2013.02.002

Stentzel, S., Teufelberger, A., Nordengrün, M., Kolata, J., Schmidt, F., van Crombruggen, K., et al. (2017). Staphylococcal serine protease-like proteins are pacemakers of allergic airway reactions to *Staphylococcus aureus*. *J. Allergy Clin. Immunol.* 139, 492.e–500.e. doi: 10.1016/j.jaci.2016.03.045

Sun, S., Ke, R., Hughes, D., Nilsson, M., and Andersson, D. I. (2012). Genome-wide detection of spontaneous chromosomal rearrangements in bacteria. *PLoS ONE* 7:e42639. doi: 10.1371/journal.pone.0042639

Thammavongsa, V., Missiakas, D. M., and Schneewind, O. (2013). Staphylococcus aureus degrades neutrophil extracellular traps to promote immune cell death. *Science* 342, 863–866. doi: 10.1126/science.1242255

Titus, B. C., and Irber, L. (2016). Sourmash: a library for MinHash sketching of DNA. *J. Open Source Softw.* 1:27. doi: 10.21105/joss.00027

Torell, E., Molin, D., Tano, E., Ehrenborg, C., and Ryden, C. (2005). Community-acquired pneumonia and bacteraemia in a healthy young woman caused by methicillin-resistant *Staphylococcus aureus* (MRSA) carrying the genes encoding Panton-Valentine leukocidin (PVL). *Scand. J. Infect. Dis.* 37, 902–904. doi: 10.1080/00365540500348945

Toussaint, J.-P., Farrell-Sherman, A., Feldman, T. P., Smalley, N. E., Schaefer, A. L., Greenberg, E. P., et al. (2017). Gene duplication in *Pseudomonas aeruginosa* improves growth on adenosine. *J. Bacteriol.* 199:e261-17.

Treangen, T. J., and Rocha, E. P. C. (2011). Horizontal transfer, not duplication, drives the expansion of protein families in prokaryotes. *PLoS Genet.* 7:e1001284. doi: 10.1371/journal.pgen.1001284

van der Woude, M. W., and Henderson, I. R. (2008). Regulation and function of Ag43 (flu). *Annu. Rev. Microbiol.* 62, 153–169. doi: 10.1146/annurev.micro.62.081307.162938

Werner, G., Coque, T. M., Hammerum, A. M., Hope, R., Hryniewicz, W., Johnson, A., et al. (2008). Emergence and spread of vancomycin resistance among enterococci in Europe. *Euro. Surveill.* 13:19046.

Xia, G., and Wolz, C. (2014). Phages of *Staphylococcus aureus* and their impact on host evolution. *Infect. Genet. Evol.* 21, 593–601. doi: 10.1016/j.meegid.2013.04.022

Zarb, P., Coignard, B., Griskeviciene, J., Muller, A., Vankerckhoven, V., Weist, K., et al. (2012). The european centre for disease prevention and control (ECDC) pilot point prevalence survey of healthcare-associated infections and antimicrobial use. *Eur. Surveill.* 17:20316. doi: 10.2807/ese.17.46.20316-en

Zhang, J. (2003). Evolution by gene duplication: an update. *Trends Ecol. Evol.* 18, 292–298. doi: 10.1016/s0169-5347(03)00033-8

Zhang, J., Liu, H., Zhu, K., Gong, S., Dramsi, S., Wang, Y.-T., et al. (2014). Antiinfective therapy with a small molecule inhibitor of *Staphylococcus aureus* sortase. *Proc. Natl. Acad. Sci. U.S.A.* 111, 13517–13522. doi: 10.1073/pnas.1408601111

Zhang, X., Bierschenk, D., Top, J., Anastasiou, I., Bonten, M. J. M., Willems, R. J. L., et al. (2013a). Functional genomic analysis of bile salt resistance in *Enterococcus faecium*. *BMC Genomics* 14:299. doi: 10.1186/1471-2164-14-299

Zhang, X., Top, J., de Been, M., Bierschenk, D., Rogers, M., Leendertse, M., et al. (2013b). Identification of a genetic determinant in clinical Enterococcus faecium strains that contributes to intestinal colonization during antibiotic treatment. *J. Infect. Dis.* 207, 1780–1786. doi: 10.1093/infdis/jit076

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Sanchez-Herrero, Bernabeu, Prieto, Hüttener and Juárez. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Molecular Biosciences

REVIEW
published: 07 August 2020
doi: 10.3389/fmolb.2020.00178

Intermolecular Communication in *Bacillus subtilis*: RNA-RNA, RNA-Protein and Small Protein-Protein Interactions

*Inam Ul Haq, Peter Müller and Sabine Brantl**

Matthias-Schleiden-Institut, AG Bakteriengenetik, Friedrich-Schiller-Universität Jena, Jena, Germany

In bacterial cells we find a variety of interacting macromolecules, among them RNAs and proteins. Not only small regulatory RNAs (sRNAs), but also small proteins have been increasingly recognized as regulators of bacterial gene expression. An average bacterial genome encodes between 200 and 300 sRNAs, but an unknown number of small proteins. sRNAs can be *cis*- or *trans*-encoded. Whereas *cis*-encoded sRNAs interact only with their single completely complementary mRNA target transcribed from the opposite DNA strand, *trans*-encoded sRNAs are only partially complementary to their numerous mRNA targets, resulting in huge regulatory networks. In addition to sRNAs, uncharged tRNAs can interact with mRNAs in T-box attenuation mechanisms. For a number of sRNA-mRNA interactions, the stability of sRNAs or translatability of mRNAs, RNA chaperones are required. In Gram-negative bacteria, the well-studied abundant RNA-chaperone Hfq fulfils this role, and recently another chaperone, ProQ, has been discovered and analyzed in this respect. By contrast, evidence for RNA chaperones or their role in Gram-positive bacteria is still scarce, but CsrA might be such a candidate. Other RNA-protein interactions involve tmRNA/SmpB, 6S RNA/RNA polymerase, the dual-function aconitase and protein-bound transcriptional terminators and antiterminators. Furthermore, small proteins, often missed in genome annotations and long ignored as potential regulators, can interact with individual regulatory proteins, large protein complexes, RNA or the membrane. Here, we review recent advances on biological role and regulatory principles of the currently known sRNA-mRNA interactions, sRNA-protein interactions and small protein-protein interactions in the Gram-positive model organism *Bacillus subtilis*. We do not discuss RNases, ribosomal proteins, RNA helicases or riboswitches.

Keywords: sRNA, small regulatory RNA, *Bacillus subtilis*, type I toxin-antitoxin system, T-box riboswitch, RNA chaperones, small proteins, antiterminators

OPEN ACCESS

Edited by:
Manuel Espinosa, Margarita Salas Center for Biological Research (CSIC), Spain

Reviewed by:
Jörg Stülke, University of Göttingen, Germany
Elena Evguenieva-Hackenberg, University of Giessen, Germany

****Correspondence:***
Sabine Brantl
sabine.brantl@uni-jena.de

Specialty section:
This article was submitted to Molecular Recognition, a section of the journal Frontiers in Molecular Biosciences

Received: *28 April 2020*
Accepted: *09 July 2020*
Published: *07 August 2020*

Citation:
Ul Haq I, Müller P and Brantl S (2020) Intermolecular Communication in Bacillus subtilis: RNA-RNA, RNA-Protein and Small Protein-Protein Interactions. Front. Mol. Biosci. 7:178. doi: 10.3389/fmolb.2020.00178

Abbreviations: aa, amino acid; bp, base pair; MW, molecular weight; nt, nucleotide; RBS, ribosome binding site, sRNA, small RNA; SU, subunit; TA, toxin-antitoxin; TSS, transcription start site.

INTRODUCTION

The first regulatory RNAs were identified and intensively studied as control elements in replication, conjugation and maintenance of bacterial plasmids as well as in transposition and transduction (rev. in Wagner and Romby, 2015). However, only after the publication of a variety of regulatory RNAs from intergenic regions of *E. coli* by two groups (Argaman et al., 2001; Wassarman et al., 2001), knowledge on bacterial regulatory RNAs started to expand. This was due to systematic bioinformatic approaches combined with experimental studies as well as RNA sequencing, which revealed that an average bacterial genome encodes between 200 and 300 small regulatory RNAs (sRNAs). The majority of sRNA targets are mRNAs. Therefore, sRNA-mRNA interactions play an important role in virtually all bacterial cells. Furthermore, uncharged tRNA can interact with the T-box in 5′ UTRs of various mRNAs encoding amino-acid related genes resulting in transcriptional read-through. Bacterial cells contain a number of RNA chaperones that stabilize sRNAs, modulate mRNA translation or promote mRNA-sRNA interactions. In general, Hfq (rev. in Kavita et al., 2018), ProQ (rev. in Olejniczak and Storz, 2017) and CsrA (rev. in Vakulskas et al., 2015) are known and have been investigated in great detail in this respect. Interestingly, Gram-positive bacteria do not encode ProQ, and Hfq does not seem to play a comparable role as in Gram-negative bacteria. Furthermore, CsrA has not been found to sequester small RNAs in Gram-positives. Other RNA-protein interactions involve tmRNA, the ubiquitous RNase P, 6S RNA that interacts with RNA polymerase, RNA antiterminators and terminators that are bound and regulated by proteins as well as the aconitase, a TCA enzyme that moonlights at iron limiting conditions as RNA binding protein. In addition, interactions between small proteins and larger proteins or protein complexes play an important role in gene expression control. Only a few interactions have been elucidated and investigated in some detail in this new field, e.g., modulation of the function of individual proteins or activation/inactivation of multiprotein complexes by small proteins, interactions of small proteins with membrane proteins or the membrane itself.

Here, we present an overview of recent advances in sRNA/tRNA-mRNA, RNA-protein and small protein-protein interactions in the Gram-positive model organism *Bacillus subtilis*. However, we will not include RNases, RNA helicases, ribosomal proteins or riboswitches.

RNA-RNA INTERACTIONS

Two main groups of interactions will be reviewed here: small regulatory RNAs (sRNAs) that interact with their target mRNAs and tRNAs that interact with the 5′ UTR of mRNAs in the so-called T-box riboswitches.

sRNAs can be classified in *cis*- and *trans*-encoded sRNAs. The first sRNAs discovered 35–40 years ago on plasmids and transposons are *cis*-encoded, i.e., transcribed convergently to their RNA targets from the complementary DNA strand and able to form complete duplexes with them (rev. in Brantl, 2012; Wagner and Romby, 2015). Therefore, each *cis*-encoded sRNA has only one target RNA. By contrast, *trans*-encoded sRNAs are encoded in another region of the genome and are only partially complementary to their broad variety of target RNAs. Whereas for the majority of the currently known ≈150 *E. coli* sRNAs targets have been identified and mechanisms of action elucidated, only a dozen of the 108 confirmed *B. subtilis* sRNAs (Rasmussen et al., 2009; Irnov et al., 2010) have been investigated in great detail. Here, we focus on sRNAs with known targets.

Cis-Encoded sRNAs

The largest group of currently known *B. subtilis cis*-encoded sRNAs are type I antitoxins that interact with their target toxin mRNAs either at their 5′ or 3′ end by a base-pairing mechanism. In addition, the targets for a few other *cis*-encoded sRNAs have been identified and characterized (see below). **Table 1** provides an overview of all currently known *cis*-encoded sRNAs in *B. subtilis*.

Antitoxins in Type I Toxin-Antitoxin Systems

Type I toxin-antitoxin (TA) systems are composed of two elements, a hydrophobic toxin and an RNA antitoxin that neutralizes toxin action by interacting with the toxin mRNA to affect its stability and/or translation. Out of 14 predicted type I TA systems of *B. subtilis* (Durand et al., 2012a), four have been verified experimentally and analyzed in some detail: *txpA*/RatA (Silvaggi et al., 2005), *bsrG*/SR4 (Jahn et al., 2012), *bsrE*/SR5 (Meißner et al., 2015; Müller et al., 2016), and *yonT*/SR6 (Durand et al., 2012b; Reif et al., 2018). The majority of them are located on prophage elements or phage remnants in the chromosome.

The ***txpA*/RatA** module is encoded on the chromosomal *skin* element. In *ratA* knockout strains, the toxin TxpA (59 aa) causes cell lysis on agar plates after 5 days. The 220 nt long RatA overlaps the 3′ end of *txpA* mRNA by ∼120 nt. The interaction between both RNAs promotes the degradation of *txpA* mRNA by RNase III, which is required for the viability of *B. subtilis* and makes RNase III an essential enzyme in this bacterium (Durand et al., 2012b; **Figure 1A**). The secondary structures of RatA and *txpA* RNA as well as their complex have been experimentally determined. The *SD* sequence of *txpA* is sequestered in a 5 bp double-stranded region. RatA binding does not alter the *txpA* structure around the *SD* sequence, i.e., it has no influence on its accessibility to the ribosomal 30S SU. None of the interacting loops contains a U-turn motif (see below).

The ***bsrG*/SR4** module is located on the chromosomal SPβ prophage. In *sr4* knockout strains, the hydrophobic toxin BsrG (38 aa) causes cell lysis on agar plates after overnight incubation at 37°C (Jahn et al., 2012, 2015). The 294 nt long *bsrG* mRNA and the 180 nt long antitoxin SR4 interact at their 3′ ends resulting in degradation of *bsrG* mRNA by RNase III 8 nt downstream from the stop codon. RNase III is, however, not involved in the degradation of either *bsrG* RNA or SR4 alone. Endoribonuclease Y and the 3′-5′ exoribonuclease R are responsible for further degradation of both RNAs. PNPase processes SR4 precursors into the mature RNA. In contrast to *txpA*/RatA, RNase III is not essential for the *bsrG*/SR4 system because a Δ*rnc* strain

TABLE 1 | Overview of *cis*-encoded sRNA/mRNA systems from *B. subtilis*.

Sense/as gene pair	sRNA length	sRNA action	Target function	Biological role/regulation in	Regulation peculiarity
txpA/RatA	222 nt	RD	Toxin (59 aa)	Skin maintenance	Glucose dependent, III**
bsrG/SR4	180 nt	TI + RD	Toxin (38 aa)	SPβ maintenance?	Temperature dependent
bsrE/SR5	163 nt	RD	Toxin (30 aa)	Anaerobic stress	Multistress-responsive
yonT/yoyJ/SR6	100/215* nt	RD *yonT* TI *yoyJ*	Toxin (58 aa) Toxin (83 aa)	n. d.	Multistress-responsive, SR6 very stable, III**
bsrH/as-bsrH	200 nt	RD	Toxin (29 aa)	n. d.	Multistress-responsive J1 cleaves mRNA
yabE/S25	1350 nt#	n. d.	Autolysin	Regulation of autolysin expression	sRNA under σ^X and σ^M control
gdpP/S1559	667 nt#	TI?	c-di-AMP PD	Regulation of c-diAMP phosphodiesterase	sRNA under σ^D control
cwlO/S1326	700–2200 nt	n. d.	Endopeptidase type autolysin	Exit from stationary phase	antisense sRNA under σ^B control
rpsD/S1136+	185–600 nt	T interference	Ribosomal protein S4	Ribosome reduction under ethanol stress	sRNA under σ^B control dual-function sRNA
opuB/S1290.	300–3800 nt	T interference	Choline transporter	Delayed osmoprotection under salt stress	sRNA under σ^B control

*RD, promotion of mRNA degradation; PD, phosphodiesterase; TI, translational inhibition; T intereference, transcriptional interference; ?, proposed but not experimentally shown. *Longer SR6 species due to read-through of SR6 terminator. #Length has not been determined in Northern blots, but was estimated from transcriptome data; III**, RNase III is essential for inhibition by the antitoxin; n.d., not determined; +Has been renamed SR7 (Ul Haq et al., 2020) (see bellow). For further details see text.*

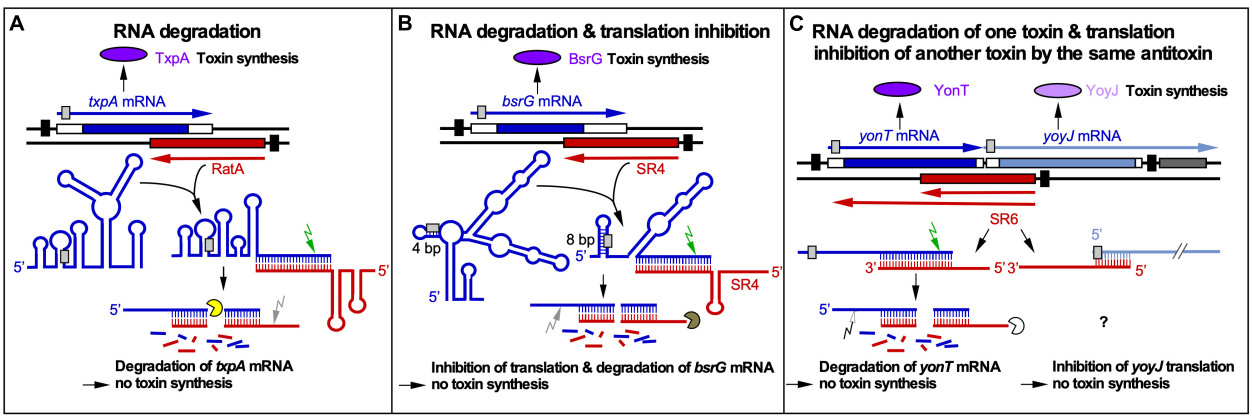


FIGURE 1 | RNA-RNA interactions of *cis*-encoded sRNAs from *B. subtilis*. Black bars denote promoters. Toxins are drawn in purple or light purple, toxin mRNAs in blue or gray blue and antitoxins in red. Toxin ORFs are represented by blue and gray blue bars. Light gray boxes depict RBS. Arrows symbolize endoribonucleases (RNase III, green; RNase Y, gray; unknown RNase, white) and circular segments 3′-5′ exoribonucleases (PNPase, yellow; RNase R, brown; unknown RNase, white). **(A)** Promotion of RNA degradation. The antitoxin RatA and its *txpA* toxin mRNA base-pair at their 3′ ends. **(B)** RNA degradation and translation inhibition. The antitoxin SR4 and the corresponding *bsrG* toxin mRNA interact at their 3′ ends. SR4 binding to *bsrG* mRNA induces a conformational alteration that extends the region sequestering the SD sequence from 4 to 8 bp which inhibits *bsrG* translation. Additionally, the SR4/*bsrG* mRNA interaction facilitates toxin mRNA decay by an initial RNase III cleavage followed by subsequent RNase R and RNase Y degradation. **(C)** One antitoxin inhibits two toxins by different mechanisms. Antitoxin SR6 and *yonT* toxin mRNA base-pair at their 3′ ends, which promotes *yonT* mRNA decay by an initial RNase III cleavage that is followed by degradation by so far unidentified RNases. Furthermore, SR6 interacts with *yoyJ* toxin mRNA by base-pairing at the 5′ ends, which does not promote *yoyJ* mRNA degradation, but prevents *yoyJ* overexpression, possibly by translational inhibition. (Adapted with permission from Brantl and Müller, 2019).

neither lysed on agar plates nor had mutations in the *bsrG* ORF. Furthermore, the RNA chaperone Hfq is not required for the function of the *bsrG*/SR4 system, since a Δ*hfq* strain does neither lyse on agar plates nor displays changed half-lives of SR4 or *bsrG* mRNA (Jahn et al., 2012). The secondary structures of SR4 and *bsrG* mRNA and their complex were experimentally determined (Jahn and Brantl, 2013). SR4 induces structural alterations around the SD sequence of *bsrG* by extending a 4 bp double-stranded region that makes the SD barely accessible into an 8 bp region which inhibits *bsrG* translation. SR4 is, therefore, the first dual-function type I antitoxin: It facilitates *bsrG* mRNA degradation and prevents *bsrG* translation by impairing ribosome access to the *bsrG* SD (**Figure 1B**).

Complex formation assays with *bsrG* RNA and SR4 yielded an apparent binding constant k_{app} of 6.5×10^5 M^{-1} s^{-1} (Jahn and Brantl, 2013). The elucidation of the binding pathway of *bsrG*

mRNA and SR4 (Jahn and Brantl, 2013; **Figure 2A**) revealed that binding initiates with a single loop-loop contact between loop L3 of *bsrG* RNA and loop L4 of the SR4 terminator stem-loop. Subsequently, base-pairing progresses via the single-stranded region between L4 and L3 toward L3 of SR4, and, finally to L2 that pairs with the *bsrG* terminator-stem-loop. However, the latter interaction is not essential for efficient binding. A 5′ YUNR motif present in L3 of *bsrG* RNA might form a U-turn (see Heidrich and Brantl, 2003) to provide a scaffold for the efficient initial interaction with SR4. An excess of antitoxin over toxin mRNA is obtained by the 6- to 10-fold higher promoter strength of the *sr4* promoter compared to the *bsrG* promoter (Jahn et al., 2012).

The multistress-responsive ***bsrE*/SR5** module is encoded on the prophage-like element P6. In contrast to *txpA*/RatA and *bsrG*/SR4, deletion of the *sr5* promoter to prevent antitoxin expression does not lead to cell lysis on agar plates. Only *bsrE* overexpression from a multicopy plasmid inhibits cell growth and causes lysis on agar plates indicating that BsrE (30 aa) is a much weaker toxin than BsrG (Müller et al., 2016). SR5 (163 nt) and *bsrE* mRNA (256 nt) interact by 114 complementary bp at their 3′ ends, and this interaction causes degradation of the toxin mRNA by RNase III at all three *bsrE* stop codon positions (Meißner et al., 2015). The secondary structures of *bsrE* mRNA, the antitoxin SR5 and their complex were mapped and are similar to those of *bsrG* RNA and SR4 (Meißner et al., 2015). Complex formation between toxin mRNA and RNA antitoxin were studied and showed a similar k_{app} value of $1–3 \times 10^6$ M^{-1} s^{-1}. Likewise, the SR5/*bsrE* mRNA interaction pathway (**Figure 2B**) closely resembles that of *bsrG* RNA and SR4 (see above). The main three differences between both pathways are: (i) the presence of only one U-turn motif in loop L3 of *bsrG* mRNA but two U-turn motifs, one in L4 of *bsrE* RNA and another in terminator loop L4 in SR5 which are engaged in the initial contact, (ii) two stem-loops of SR5, SL2 and SL4, are fundamental for formation of a stable duplex with *bsrE* mRNA, whereas only one stem-loop and a single stranded region of SR4 sufficed and (iii) SR5 binding did not trigger a structural change around the toxin mRNA RBS, whereas SR4 binding did. Therefore, the antitoxin SR5 is monofunctional: It promotes toxin mRNA degradation but does not inhibit *bsrE* translation directly. Interestingly, both *bsrE* mRNA and SR5 respond to different stress factors: Whereas *bsrG* mRNA is only heat-shock sensitive due to its refolding at 48–55°C which results in rapid degradation by RNases Y and J1 (Jahn and Brantl, 2016), *bsrE* mRNA is sensitive to heat-shock, ethanol stress and alkaline pH (Müller et al., 2016). SR5 amounts were influenced by iron limitation, acid, alkaline and anaerobic stress. Oxygen deficiency was the only stress that altered the SR5/*bsrE* mRNA ratio from 9: 1 to 0.5: 1, i.e., has the potential to induce cell lysis (Müller et al., 2016).

The multistress-responsive ***yonT*/*yoyJ*/SR6** module is the second type I TA system encoded on the SPβ prophage. The peculiarity in this system is that antitoxin SR6 interacts with two toxin mRNAs encoding YonT (59 aa) and YoyJ (83 aa). The 3′ end of the 100 nt long SR6 interacts with the 3′ end of *yonT* mRNA to promote its degradation by RNase III. On the other hand, the 5′ end of SR6 binds to the 5′ end *yoyJ* mRNA, but neither affects the amount nor half-life of *yoyJ* RNA. Instead, it seems to inhibit *yoyJ* translation (Reif et al., 2018; **Figure 1C**). The latter interaction can only be postulated as there is only indirect evidence for YoyJ being a weak toxin: A *yoyJ* overexpression plasmid could only be established by transformation in a *B. subtilis* strain that carries the *sr6* gene in the chromosome (Reif et al., 2018). In contrast to TxpA, BsrG or BsrE, YonT is a very strong toxin, as neither SR6 could be deleted from the chromosome nor *yonT* overexpressed from a medium copy plasmid in *E. coli* or *B. subtilis*. So far, no secondary structures were probed or complex formation studies performed. However, calculations of the *yonT* mRNA and SR6 amounts revealed only in minimal CSE medium with glucose at stationary phase an excess of *yonT* mRNA over SR6, i.e., under these conditions the toxin could be expressed.

Similarly to *bsrE*/SR5, *yonT*/*yoyJ*/SR6 responds to a number of stress factors: After ethanol stress, *yonT* mRNA disappeared as *bsrE* mRNA within 0.5 min, and heat-shock caused a 4-fold decrease of *yonT* RNA levels. However, in contrast to SR4 or SR5, SR6 levels also decreased about 4-fold after heat-shock. Vancomycin (cell-wall stress) treatment led to a decrease of both *yonT* mRNA and SR6.

Another potential type I TA system is *bsrH*/anti-bsrH (Durand et al., 2012a,b) located on the *skin* prophage. *bsrH* mRNA (285 nt) and anti-bsrH (200 nt) can interact at their 3′ ends and RNase J1 is involved in cleavage of the complex. But so far, it has not been shown that BsrH (29 aa) acts as a toxin in *B. subtilis* and that anti-bsrH neutralizes toxin action.

Other *Cis*-Encoded sRNAs in *B. subtilis*

At least five other *cis*-encoded sRNAs (antisense RNAs) have been reported: S25, gdpPas/S1559, S1326, S1136-1134 and S1290. **S25** (≈1350 nt) is expressed under control of extracytoplasmic sigma factors σ^X and σ^M and regulates an autolysin encoded by *yabE* (Eiumphungporn and Helmann, 2009). **S1559/gdpPas** (Luo and Helmann, 2012) is transcribed under control of s^D from the middle of the *gdpP* gene encoding the phosphodiesterase responsible for degradation of cyclic di-AMP. Neither sense nor antisense RNA were detectable in Northern blots. The presence of S1559/gdpPas reduced the level of FLAG-tagged GdpP about 2.5 to three-fold, but could not be associated with any phenotype which makes the biological role of this regulation elusive. Antisense RNA **S1326** is transcribed from a σ^B dependent promoter located downstream of the *cwlO* gene that encodes a D,L-endopeptidase type autolysin (Noone et al., 2014). S1326 has a heterogenous length (between 700 and 2200 nt) and was detectable under phosphate stress. The largest S1326 transcript extends beyond the *cwlO* promoter. The *cwlO* mRNA is highly unstable due to an RNase Y cleavage in the leader region which seems to be decisive for regulation of this RNA. S1326 influenced weakly the exit of the cells from stationary phase.

The σ^B dependent sRNA **S1136-1134** transcribed *in cis* to *rpsD* mRNA encoding the ribosomal primary binding protein S4 (22.7 kD) was reported to decrease the amount of *rpsD* mRNA and, consequently, the amount of the small ribosomal subunit ∼1.5 fold under ethanol-stress (Mars et al., 2015a). S1136-1134 could only act *in cis*, which suggests that its mechanism of action is transcriptional interference.

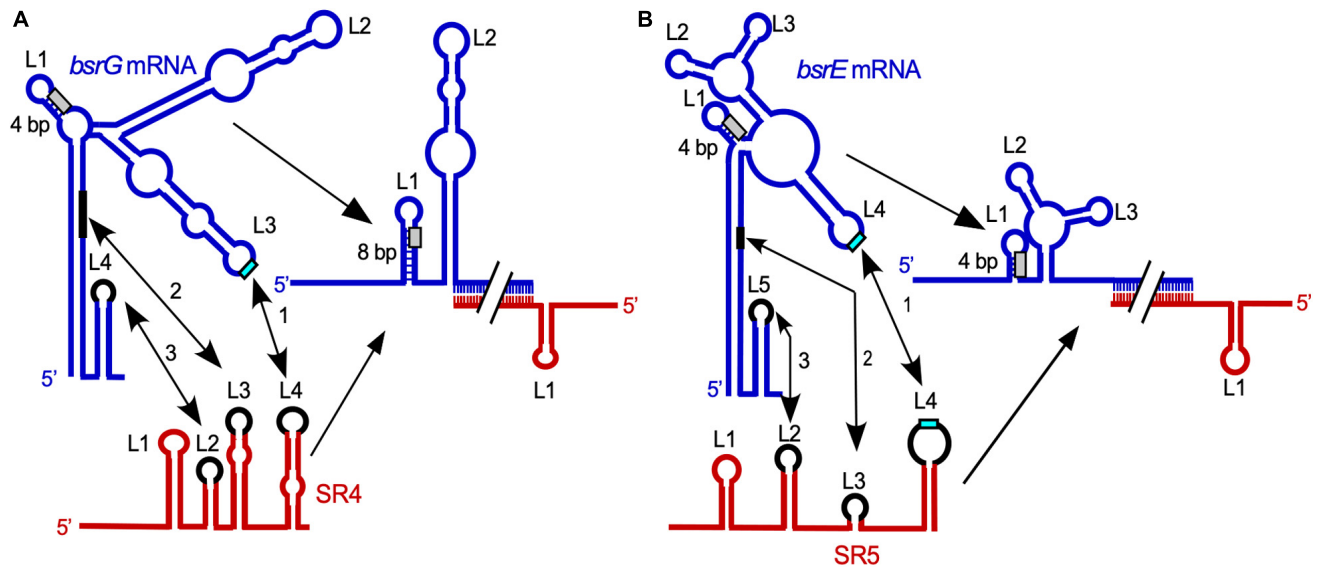


FIGURE 2 | Comparison of the SR4/*bsrG* RNA **(A)** and SR5/*bsrE* RNA **(B)** interaction pathways. Blue, toxin mRNAs; red, RNA antitoxins. U-turn motifs are indicated in turquoise and RBS by light gray boxes. The interaction chronology is designated by 1–3. L, loop. **(A)** The initial contact between SR4 and *bsrG* RNA occurs between L4 of SR4 and L3 of *bsrG* RNA (1). It is followed by helix progression to an interaction between SR4 L3 and the 3′ part of helix P1 of *bsrG* RNA (2) and finally reaches L2 of SR4 that binds terminator loop L4 of *bsrG* RNA (3). The latter interaction is not essential. **(B)** The binding pathway of SR5 and *bsrE* RNA comprises three similar subsequent interactions. Schematic secondary structures are based on experimentally probed structures in Jahn and Brantl (2013) and Meißner et al. (2015). (Adapted with permission from Brantl and Jahn, 2015; Brantl and Müller, 2019).

The same holds true for σ^B-dependent **S1290** that is transiently transcribed in response to salt stress. S1290 acts solely *in cis* and independent of RNase III on its convergently transcribed target *opuB* mRNA encoding a choline transporter (Rath et al., 2020). It causes a time-delayed osmotic induction of *opuB* mRNA, since under acute salt stress, *B. subtilis* initially relies on the promiscuous OpuC transporter to import pre-formed compatible solutes as proline- or glycine betaine before employing OpuB to import choline that has to be converted into glycine betaine. In addition to S1290, *opuB* expression is dependent on the degree of the imposed osmotic stress.

For neither of these antisense RNAs, secondary structures or – in case they act through complex formation with their target RNAs – those of the antisenseRNA/target RNA complexes were determined.

Trans-Encoded sRNAs in *B. subtilis*

Currently, four *trans*-encoded sRNAs are known for which targets have been identified: SR1 (Licht et al., 2005; Heidrich et al., 2006), FsrA (Gaballa et al., 2008), RoxS (Durand et al., 2015, 2017), and RnaC (Mars et al., 2015b; **Figure 3**). Among them, SR1 is the only dual-function sRNA which acts as a base-pairing sRNA and as mRNA encoding a small peptide, SR1P (see below). **Table 2** provides an overview of all currently known *trans*-encoded sRNAs in *B. subtilis*.

SR1 (205 nt) has been discovered using a computational search for sRNAs in intergenic regions of the *B. subtilis* genome and later confirmed by Northern blotting (Licht et al., 2005). Its transcription is repressed under glycolytic conditions ≈20–30-Fold by CcpN binding upstream of and overlapping the *sr1* promoter p_{sr1} and, to a minor extent, by CcpA binding to a site ∼260 bp upstream of the transcription start site. CcpN needs ATP and a slightly acidic pH to exert its effect (Licht et al., 2008) and interacts with the α-subunit of the RNA polymerase to prevent promoter escape (Licht and Brantl, 2009). The first identified primary target of SR1 is *ahrC* mRNA (Heidrich et al., 2006) encoding the transcription activator of the arginine catabolic operons *rocABC* and *rocDEF* and the transcription repressor of the arginine biosynthesis genes (Czaplewski et al., 1992). In contrast to many other sRNAs, SR1 neither affects the half-life nor the amount of *ahrC* mRNA (Heidrich et al., 2006). Binding of SR1 and *ahrC* mRNA via their 7 complementary regions (A to G in SR1, A′ to G′ in *ahrC* mRNA) results in inhibition of *ahrC* translation initiation by a novel mechanism: induction of structural changes 20 to 40 nt downstream of the RBS, although SR1 binds ∼100 nt downstream of the *ahrC* RBS (Heidrich et al., 2007; **Figure 3A**). The interaction between SR1 and *ahrC* mRNA initiates at region G/G′ and is the crucial interaction, but the other complementary regions also contribute to complex formation and hence, translation inhibition (Heidrich et al., 2007). Recently, it was shown that the abundant RNA chaperone CsrA promotes the interaction between SR1 and *ahrC* mRNA: It binds both RNAs and induces a slight structural alteration in *ahrC* mRNA which liberates region G' for a more efficient interaction with SR1 (Müller et al., 2019; **Figure 3A**).

SR1 is not only a *trans*-encoded base-pairing sRNA, but also an mRNA encoding a 39 aa protein, SR1P (Gimpel et al., 2010). SR1P interacts with the glyceraldehyde-3P dehydrogenase

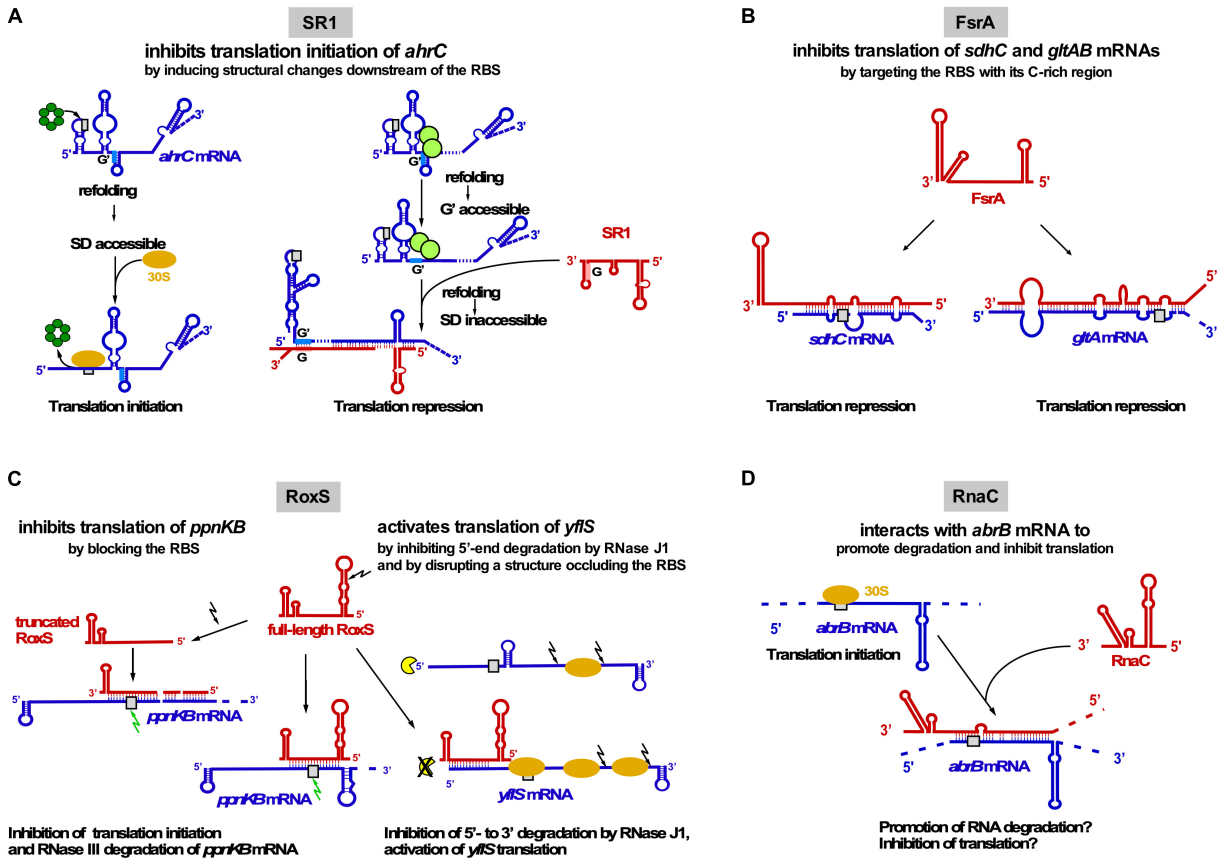


FIGURE 3 | Interactions of *trans*-encoded sRNAs with their target mRNAs **(A)** SR1 interacts with *ahrC* mRNA about 100 nt downstream of the RBS which induces structural changes around the *ahrC* RBS that inhibit translation initiation. Left: Hfq (dark-green) binds immediately upstream of the *ahrC* RBS (gray rectangle) to make it accessible to the 30S SU. Right: By binding to GGA motives 1–3 of *ahrC* mRNA, CsrA (light-green) induces a slight structural change that makes region G′ accessible to complementary region G of SR1. SR1 binding induces a structural change that renders the *ahrC* RBS inaccessible to 30S binding, thereby inhibiting translation initiation (Brantl and Brückner, 2014). **(B)** Interaction of FsrA with *sdhC* and *gltAB* mRNAs. FsrA has numerous targets, but probably for all of them (see text) inhibits translation by targeting the RBS or an adjacent region. **(C)** Interaction of RoxS with two target mRNAs. Whereas RoxS and its truncated derivative repress *ppnKB* translation, full-length RoxS activates *yflS* translation, mainly by inhibiting RNase J1 degradation from the 5′ end. **(D)** Interaction of RnaC with *abrB* mRNA which might promote RNA degradation and inhibit *abrB* translation. Red, sRNAs; blue, target mRNAs; beige oval, 30S SU; gray box, RBS; green arrow, RNase III; black arrow, RNase Y; yellow, RNase J1.

TABLE 2 | Overview of trans-encoded sRNAs from *B. subtilis.*

sRNA (nt)	Target RNA(s)	Biological function	Mechanism of action	Control of expression/specific characteristic
SR1 (205)	*ahrC*	Arginine catabolism	TI	CcpN, CcpA, sporulation
FsrA (84)	*sdhCAB,*	Iron sparing response	TI	Fur, iron;
	citB	Aconitase		some targets need FbpA, B or C
	gltA	Glutamate synthase (iron-sulfur)		
	lutABC	Iron-sulfur oxidase		
	dctP	Dicarboxylate permease		
	leuCD	Leucin biosynthesis		
RoxS (115)	*ppnKB, sucC, yflS*	Redox regulation, TCA cycle	TI, RD	ResD (NO), Rex (malate)
RnaC (125?)	*abrB*	Transition state	RD?,TI?	Growth phase

TI, translational inhibition; RD, promotion of mRNA degradation; ?, proposed but not experimentally shown. For further details see text.

GapA, thereby affecting RNA degradation (Gimpel and Brantl, 2016, 2017, see below and **Figure 4**). Both functions of SR1 are remarkably conserved over more than one billion years of evolution (Gimpel et al., 2012).

In 2008, **FsrA** (84 nt), a functional homolog of *E. coli* RyhB, was discovered (Gaballa et al., 2008). FsrA is under transcriptional control of the iron-dependent Fur repressor. When iron is scarce, FsrA inhibits translation of target

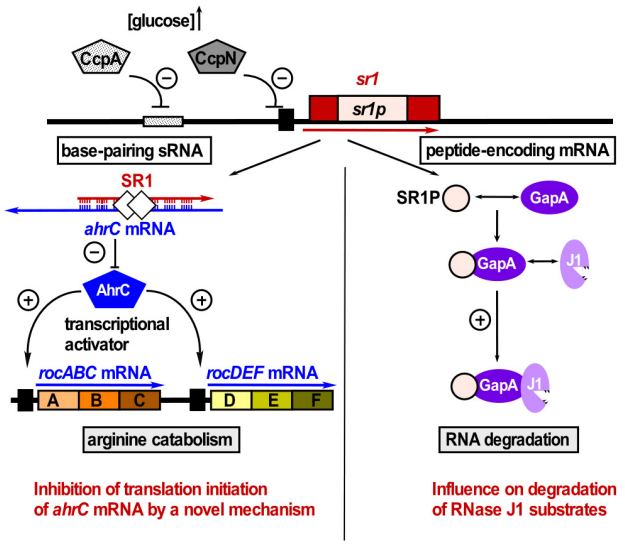


FIGURE 4 | SR1 is a dual-function sRNA. Left: SR1 acts as base-pairing sRNA in arginine catabolism by inhibiting translation of *ahrC* mRNA encoding the transcription activator of the arginine catabolic operons. SR1 transcription is inhibited by repressors CcpA and CcpN. CsrA (white diamonds) promotes the SR1-*ahrC* mRNA interaction (Müller et al., 2019). Right: SR1 is an mRNA encoding SR1P that plays a role in RNA degradation. SR1P interacts with GapA thereby promoting the GapA-RNase J1 interaction and the enzymatic activity of RNase J1 (adapted with permission from Gimpel and Brantl, 2017).

mRNAs involved in iron metabolism and storage, e.g., succinate dehydrogenase *sdhCAB* (**Figure 3B**), *citB* (aconitase) and *lutABC* (oxidases important for growth on lactate as sole C source). Subsequent studies revealed that FsrA is a global regulator with a broad variety of targets, among them *gltAB* encoding the iron-sulfur containing enzyme glutamate synthase, *dctP* (dicarboxylate transporter), *resA* and *qcrA*. By predicting complementary regions with RNAhybrid and M-fold the authors suggest that a C-rich single-stranded region of FsrA base-pairs with the RBS of its target mRNAs (Smaldone et al., 2012a). So far, *in vivo* base-pairing has only been confirmed between FsrA and the 5′ UTR of *gltAB* (Smaldone et al., 2012a; **Figure 3B**). *In vitro* base-pairing was demonstrated by EMSA for FsrA and *sdhC* mRNA, whereas for the other targets, only effects on mRNA and protein levels were reported (Gaballa et al., 2008). In contrast to the functionally related Fur-regulated sRNA RyhB from *E. coli*, which requires the RNA chaperone Hfq, FsrA cooperates with one, two or three Fur-regulated small basic proteins FbpA, FbpB, and FbpC suggested to be RNA chaperones. Under iron-deplete growth conditions, FsrA and FbpB (48 aa) inhibit *lutABC* to allow the direction of iron to higher-priority target proteins. Both target mRNA and protein levels were affected about two-fold by FsrA and FbpB with FbpB having a stronger effect on RNA levels. Whereas FsrA might directly inhibit translation, FbpB might facilitate FsrA/*lutABC* RNA pairing or recruit the RNA degradation machinery (Smaldone et al., 2012b). Currently, there is neither experimental evidence for FbpB binding RNA nor for its mechanism of action. Neither for *sdhCAB* mRNA nor for *citB* mRNA a contribution of one of the Fbp proteins to FsrA-mediated repression was found (Gaballa et al., 2008).

RoxS (originally termed RsaE, 115 nt) is the only base-pairing sRNA conserved between *B. subtilis* and *S. aureus*. RoxS transcription is activated by nitric oxide depending on the two-component system ResDE (Durand et al., 2015) and repressed by the NADH sensitive Rex (Durand et al., 2017). Repression is released by malate. RoxS helps to restore the NAD^+/NADH balance by temporarily turning down part of the TCA cycle. Experimentally confirmed direct RoxS targets are *ppnKB* mRNA (NAD^+/NADH kinase), *sucCD* mRNA (succinyl-CoA synthase) and *yflS* (one of four malate transporters; Durand et al., 2015, 2017). RoxS inhibits translation and promotes degradation of *ppnKB* and *sucC* mRNAs (**Figure 3C**). RNase III cleaves the RoxS/*ppnKB* RNA duplex and RNase Y cleaves *ppnK* mRNA both RoxS-dependently but also -independently. For translation regulation, one of four CCC rich regions of RoxS (CCR3) suffices. RoxS itself is also subject of RNase Y cleavage at nt + 20 yielding truncated RoxS(Y) which can efficiently regulate *ppnKB* and is required for translation inhibition of *sucCD* operon RNA. In contrast to *ppnKB* and *sucCD*, *yflS* is positively regulated: RoxS activates *yflS* translation by disrupting a structure at the 5′ UTR that impedes ribosome recruitment. In addition, RoxS stabilizes *yflS* mRNA by preventing 5′-3′ exonucleolytic degradation by RNase J1 (**Figure 3C**). Both effects are independent and require RoxS region CCR3. Hfq is not involved in RoxS-dependent regulation of *ppnKB* mRNA (Durand et al., 2015).

RnaC/S1022 was first identified in a microarray screen of *B. subtilis* intergenic regions, but has so far not been detected in Northern blots. It is transcribed exclusively under control of σ^D during logarithmic growth in LB medium (Schmalisch et al., 2010). RnaC modulates the cellular level of transition state regulator AbrB via base-pairing with *abrB* mRNA at the RBS and the first six codons (Mars et al., 2015b; **Figure 3D**). The Hfq-independent RnaC/*abrB* RNA interaction seems to promote degradation of *abrB* mRNA and might also inhibit its translation. However, only minimal alterations (≈33%) of mRNA and protein levels were observed in Δ*rnaC* strains. RnaC increases the cell-to-cell variation of the AbrB levels, which leads to growth rate heterogeneity of cells within one population during exponential phase. This heterogeneity is physiologically relevant, since slowly growing bacterial cells are less susceptible to antibiotics and environmental stress (Mars et al., 2015b).

The T-box Riboswitch

Another example of regulatory RNA-RNA interactions is the T-box riboswitch, first discovered in the *B. subtilis tyrS* gene (Henkin et al., 1992). It is an example for transcription attenuation mechanisms that are based on the ability of an mRNA leader region to fold into two mutually exclusive structures, a transcription terminator or an antiterminator, depending on ribosome movement, binding of proteins, an antisense RNA, small ligands or – in the T-box– the loading state of a tRNA (rev. in Brantl, 2004).

The T-box riboswitch is mainly found in Firmicutes to control the expression of amino acid related genes (aa biosynthesis or transport, aminoacyl-tRNA synthetases). Grundy et al. (1994)

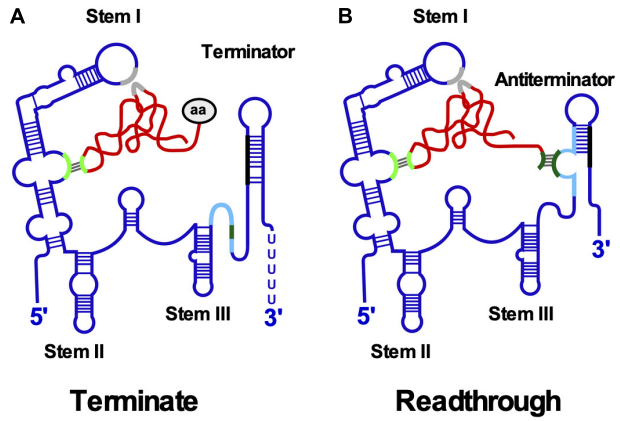


FIGURE 5 | T-box riboswitch. **(A)** Interaction between T-box leader RNA and tRNA occurs at two positions. Aminoacylated tRNA binds the specifier loop (light green) in stem I through base-pairing and the stem I platform (gray) by stacking interactions. The aa at the 3′ end prevents binding of the acceptor arm to the terminator helix. The more stable terminator helix (black-blue) forms and transcription terminates. **(B)** Uncharged tRNA interacts with specifier loop and stem I platform. The free acceptor arm of tRNA (dark green) binds the antiterminator (light blue) and stabilizes it allowing transcription to read through the termination site into the downstream gene. Conserved structural domains including stems I-III and the mutually exclusive terminator and antiterminator helices are labeled. The tRNA is shown in red and the aa in gray.

showed that a UAC to UUC mutation in the *tyrS* 5′ UTR suffices to induce expression under phenylalanine instead of tyrosine limiting conditions indicating that a tRNA was involved in regulation.

T-box leader RNAs are composed of stems I, II and III and a pseudoknot (Stem IIA/B) present upstream of the competing terminator and antiterminator helices (**Figure 5**). Stem I harbors an internal loop with the specifier sequence that base-pairs with the tRNA anticodon. The stem I distal region comprises conserved sequence motifs and a terminal loop which interact to form a contact surface with the tRNA D-/T-loops through stacking interactions (Grigg and Ke, 2013; Zhang and Ferré-D'Amaré, 2013). Both an E-loop and a pseudoknot directly downstream of size-variable stem II are required for efficient antitermination. Stem III upstream of the antiterminator varies significantly in sequence and length (rev. in Kreuzer and Henkin, 2018).

Downstream gene expression depends on the ratio between charged and uncharged tRNAs (**Figure 5**): Both tRNAs bind to the specifier loop and stem I platform. However, the presence of the charged tRNA aa prevents the interaction of the acceptor end with the antiterminator bulge, facilitating formation of a more stable terminator helix causing premature transcription termination. By contrast, the acceptor arm of uncharged tRNA base-pairs with the antiterminator bulge thereby stabilizing the antiterminator, allowing read-through into the downstream ORF. Binding of uncharged tRNA induces structural changes throughout the leader RNA. An smFRET analysis of *B. subtilis glyQS* (Zhang et al., 2018) demonstrated a two-step T-box-tRNA interaction comprising the interaction with the anticodon followed by sensing of the 3′ NCCA end. The association rate constant for the first binding step was with 5.0×10^5 M^{-1} s^{-1} comparable to that of *cis*-encoded antisense RNAs (see above).

A cryo-EM structure of the *B. subtilis* T-box riboswitch-tRNA complex (Li et al., 2019) revealed a 66 nt functional unit – T-box discriminator – that selectively binds uncharged tRNA. It is formed of stem III with flanking purines and the adjacent antiterminator. The T-box adopts a U-shaped molecular vice that clamps the tRNA. The discriminator captures the tRNA 3′ end with nanomolar affinity and uses a steric filter fashioned from a G-U wobble bp to determine its aminoacylation state. When the tRNA is uncharged, the T-box clutches it and forms a continuously stacked central spine allowing transcriptional readthrough.

RNA-PROTEIN INTERACTION

RNA Chaperones

RNA chaperones are proteins binding to RNAs and affecting their structure or stability and therefore influencing both translation of mRNAs and regulatory functions of sRNAs. *B. subtilis* encodes two major RNA chaperones: Hfq and CsrA, which are both well investigated in Gram-negative bacteria (rev. in Vakulskas et al., 2015; Kavita et al., 2018) and have a broad impact on gene regulation.

Hfq

As in other Gram-positive bacteria, the role for Hfq in *B. subtilis* seems to be minor. Only around 150 transcripts could be identified to bind Hfq and even less were affected in their abundance (Dambach and Winkler, 2013; Hämmerle et al., 2014). The structure of *B. subtilis* Hfq is similar to that of *E. coli* Hfq, but it prefers binding to AG-repeats instead of ARN-repeats (Someya et al., 2012). So far, no case has been described where *B. subtilis* Hfq facilitates the interaction between two RNAs, a major task of *E. coli* and *Salmonella* Hfq with numerous examples and a pleiotropic phenotype in its absence. In contrast, in a huge screen with more than 2000 growth conditions no growth differences between a wild-type and an isogenic *hfq*-deficient *B. subtilis* strain were detected (Rochat et al., 2015). The only known phenotypes in the absence of *hfq* are a decreased long-term survival of cells in stationary phase, independent of sporulation (Rochat et al., 2015) and a drastically impaired motility and chemotaxis (Jagtap et al., 2016).

Remarkably some toxin mRNAs as well as the corresponding antitoxins of type I TA systems are bound by Hfq (Dambach and Winkler, 2013) and are less abundant in its absence (Hämmerle et al., 2014). However, this had no impact on growth or survival (Rochat et al., 2015). Similarly, Hfq had no effect on the type I TA system *bsrE*/SR5, although it stabilizes SR5 (Müller et al., 2016). The expression of *hfq* is upregulated in stationary phase and under several stress conditions (Dambach and Winkler, 2013; Rochat et al., 2015; Jagtap et al., 2016). Together this rather implies a role for Hfq as a general stationary phase fine-tuning regulator than as a globally acting RNA-RNA matchmaker in *B. subtilis.* Interestingly Hfq binds *in vitro* near the SD sequence

of *ahrC* mRNA, refolds the 5′ UTR to allow access to the 30S SU and is required for *ahrC* translation *in vivo*, but dispensable for its inhibition by SR1 (Heidrich et al., 2007, see above). This confirms that Hfq is a chaperone in *B. subtilis*, but its role may be restricted to mRNAs.

CsrA

The second RNA chaperone, CsrA, also differs from its *E. coli* relative. Despite the highly similar general structure of the *B. subtilis* CsrA dimer (Altegoer et al., 2016) as well as the shared affinity for GGA motifs in single-stranded and looped-out RNA structures, only one major function in motility regulation in *B. subtilis* was described in detail (Yakhnin et al., 2007): The major flagellin protein Hag, the regulatory protein FliW, CsrA and the *hag* mRNA form a regulatory circuit to limit the intracellular Hag concentration. At high concentrations, Hag binds FliW, and CsrA is free to block Hag synthesis by *hag* mRNA binding, otherwise FliW sequesters CsrA and *hag* is expressed (Mukherjee et al., 2011). Nevertheless, the interaction with FliW does not exclude RNA binding by CsrA in general (Altegoer et al., 2016; Mukherjee et al., 2016), and further functions of CsrA are conceivable.

Only recently, a novel mechanism was discovered, where CsrA refolds the *ahrC* mRNA to facilitate binding of the regulatory sRNA SR1 (Müller et al., 2019, **Figure 3A**). This also raises the question of a more global role of CsrA in *B. subtilis*.

tmRNA

Trans-translation is an omnipresent bacterial mechanism to rescue ribosome stalling on mRNAs without stop-codon, to tag the nascent peptide and to initiate its degradation (rev. in Himeno et al., 2014, 2015). The functional complex comprises the tmRNA (*ssrA*) and the protein SmpB (*ssrB*). Together the complex mimics the structure of $tRNA^{Ala}$, is bound by EF-Tu-GTP and can be loaded onto stalled ribosomes. The globular domain of SmpB acts as a substitute of the anticodon-loop and occupies the decoding site, while the C-terminal tail of SmpB probes the vacant mRNA channel of the 30S SU (rev. in Miller and Buskirk, 2014). In contrast to normal tRNAs, the tmRNA has neither an anticodon-loop nor a D-loop. SmpB compensates for these structures and allows recognition by the ribosome and also fine-tunes the tmRNA structure (Miller and Buskirk, 2014). An additional looped out coding strand of tmRNA is used by the ribosome as mRNA template that encodes the species-specific tag sequence, in *B. subtilis* (A)GKTNSFNQNVALAA. Finally, the ribosome is released and the truncated and tagged protein degraded by Clp proteases that recognize the ALAA-motif (Wiegert and Schumann, 2001; Kolkman and Ferrari, 2004, **Figure 6**).

Trans-translation is important under specific stress conditions to maintain error-free gene expression and rescue stalled ribosomes. In *B. subtilis* both *ssrA* and *ssrB* are important for growth at temperature extremes and under ethanol and heavy metal ion stress (Muto et al., 2000; Shin and Price, 2007). They are encoded in one operon – together with the gene for RNase R that degrades the aberrant mRNA – and upregulated under stress conditions by a σ^B-dependent promoter (Muto et al., 2000).

Surprisingly, *trans*-translation is not only involved in the prevention of defective expression under stress conditions, but also part of gene regulation in general (rev. in Keiler, 2015). In *B. subtilis*, the system is essential in the absence of a second ribosome rescue system, BrfA/PrfB, which uses a distinct mechanism and is directly downregulated by *ssrA-smpB*-dependent *trans*-translation (Shimokawa-Chiba et al., 2019). The expression of several genes depends on the tmRNA-SmpB-mechanism (Fujihara et al., 2002). It is also involved in carbon catabolite repression: CcpA binding to *cre*-sites causes a stall of transcription and subsequently translation, which is finally solved

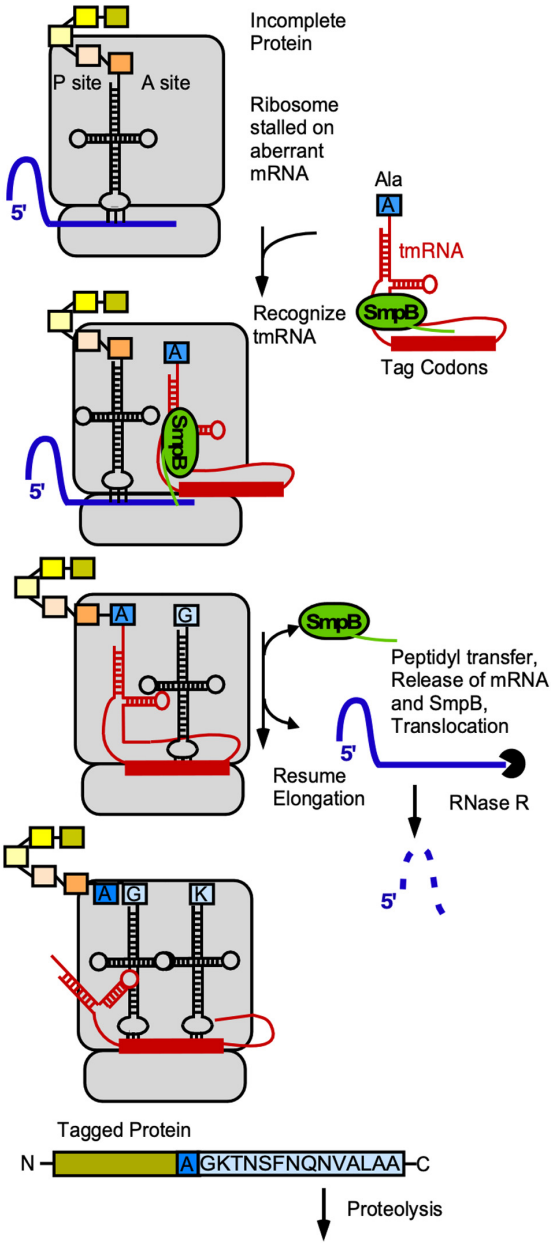


FIGURE 6 | Role of tmRNA and SmpB in *trans*-translation. Gray, ribosome; blue, mRNA, red, tmRNA; yellow/orange-colored rectangles, aa in peptide chain associated with last tRNA; blue rectangle, alanine with which tRNA was charged by alanyl-tRNA synthetase; light-blue, tag encoded by tmRNA. For details see text.

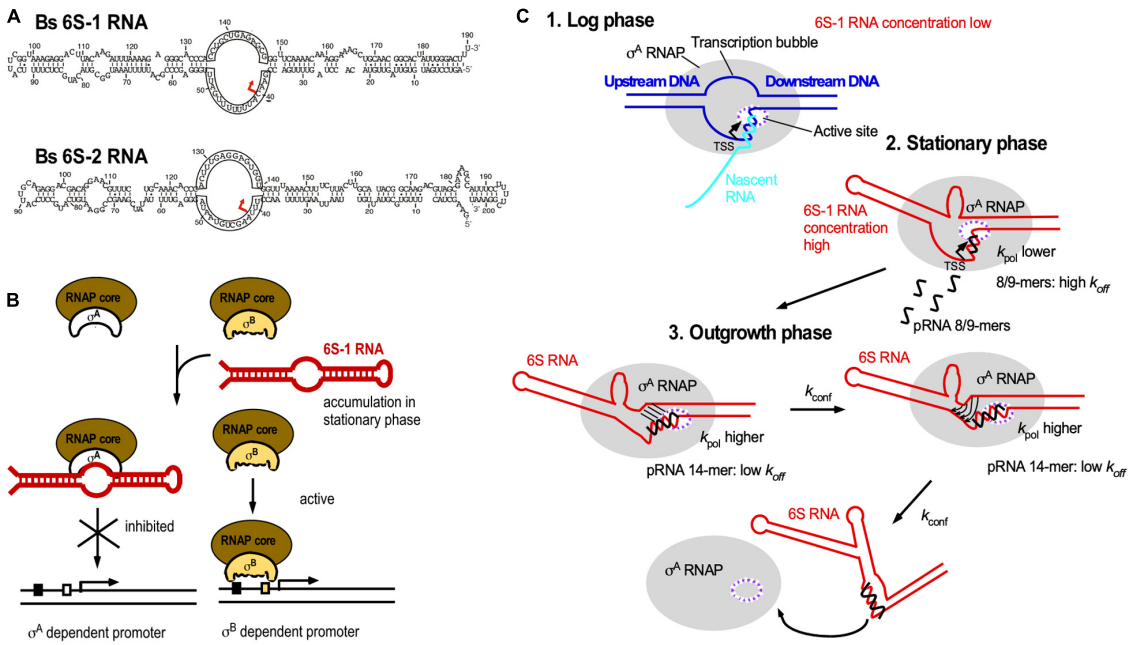


FIGURE 7 | 6S RNA. **(A)** *B. subtilis* encodes two 6S RNAs, 6S-1 and 6S-2. Red arrow, TSS for pRNAs. **(B)** Biological function of 6S RNA. By mimicking an open promoter, 6S-1 RNA, which is abundant in stationary phase, can sequester σ^A-RNAP thereby preventing transcription initiation at σ^A-dependent promoters. σ^B-RNAP is not bound by 6S-1 RNA, thus allowing transcription from σ^B-dependent promoters. **(C)** Mechanism of action of pRNAs. Black, pRNA; k_{pol}, rate constant of polymerization (synthesis) of pRNAs; k_{off}, dissociation rate constant of pRNAs; k_{conf}, rate constant for refolding of 6S RNA. The ratio between k_{pol}, k_{off}, and k_{conf} determines whether or not pRNAs dissociate or 6S-1 RNA refolds. For detailed explanations see text. **(C)** is modified based on Beckmann et al., 2012, with permission.

by *trans*-translation (Ujiie et al., 2008). Furthermore, *trans*-translation is important for efficient sporulation: In Δ*ssrA* cells, the formation of *sigK* – coding for a sporulation dependent sigma factor – by a DNA rearrangement between *spoIIIC* and *spoIVCB* is impaired, leading to drastically reduced sporulation efficiency (Abe et al., 2008). It is also likely that other physiological processes could be affected by *trans*-translation in general.

6S RNA

The ≈200 nt long 6S RNA is ubiquitous in bacteria, and *B. subtilis* has even two 6S RNAs, 6S-1 and 6S-2. All 6S RNAs have a characteristic secondary structure with a central single-stranded loop flanked by two irregular double-stranded stems, which are interrupted by small bulges (**Figure 7A**). In 2000, it was discovered that 6S RNA forms a stable complex with *E. coli* RNA polymerase (RNAP) to regulate its activity (Wassarman and Storz, 2000). It interacts with the RNAP β/B' subunit and σ^{70}, but not with the stationary phase σ^S. Therefore, it can in stationary phase, when its amount increases to 10.000 molecules/cell, repress transcription at vegetative promoters (**Figure 7B**). Crystal structure analysis revealed that *E. coli* 6S RNA mimics B-form DNA (Chen et al., 2017) allowing it to mimic an open promoter thus interfering with the formation of transcription initiation complexes (**Figure 7B**). 6S RNA acts as template for the synthesis of 14–22 nt pRNAs that are required to relieve 6S-dependent transcription inhibition and thus, recovery of cells from stationary phase in *B. subtilis* and *E. coli* (Beckmann et al., 2011; Cavanagh et al., 2012). Both *B. subtilis* 6S-1 and 6S-2 are *in vitro* and *in vivo* templates for pRNA synthesis (Burenina et al., 2014; Hoch et al., 2016). 6S-1 is highest expressed in stationary phase, allows growth adaptation and prevents premature sporulation (Cavanagh and Wassarman, 2013) whereas 6S-2 is more abundant in logarithmic phase, and its function is still unknown. In 2012, the mechanism of action of *B. subtilis* 6S-1 RNA was elucidated (Beckmann et al., 2012; **Figure 7C**): As soon as the newly synthesized pRNA has formed a sufficiently stable duplex with 6S-1 RNA, it induces *in cis* a refolding of the 6S RNA that involves a base-pairing between the 5′ part of the central bulge and nucleotides that are available due to pRNA invasion. This rearrangement decreases the affinity of 6S-1 to the RNAP. Only 12–14 nt long pRNAs, but not shorter ones, are stable enough to induce a refolding. The ratio between the rate constants for polymerization (k_{pol}), pRNA:6S RNA-dissociation (k_{off}) and refolding (k_{conf}) determines whether or not pRNAs dissociate or refold the 6S-1 RNA.

Future investigations will reveal the function of 6S-2 and show which mechanisms mediate promoter-specific transcription regulation by both 6S RNAs.

RNA Binding Proteins Inducing Terminator or Antiterminator Formation

Examples for transcription attenuation (see above and **Figure 5**) depending on protein binding involve *B. subtilis*

TRAP (*trp* RNA binding attenuation protein) and various antitermination proteins.

TRAP (Babitzke and Yanofsky, 1993; Otridge and Gollnick, 1993) is composed of 11 subunits stabilized through 11 intersubunit β-sheets to form a β-wheel with a large central hole of 80 Å diameter (Antson et al., 1999). L-tryptophan binding in clefts between adjacent β-sheets induces conformational changes. The unstructured *trp* operon 5′ UTR contains 11 U/GAG repeats spaced by 2–3 nt and forms a matching circle around Trp-bound TRAP using mostly specific protein-base interactions (**Figure 8A**) to cause transcription termination (rev. in Babitzke et al., 2019). When tryptophan is scarce, uncharged ^{trp}tRNA allows expression of Anti-TRAP that binds TRAP to inhibit its activity. The RNA refolds into an antiterminator and transcription of *trpEDCFBA* continues (rev. in Babitzke et al., 2019; **Figure 8A**). Interestingly, TRAP-mediated termination is neither intrinsic nor Rho-dependent. Instead, a region around aa E60 which is not involved in Trp or RNA binding (McAdams et al., 2017) allows TRAP to induce forward translocation of the RNAP, which is required for efficient termination (Potter et al., 2011). Both the *trpE* and *trpG* genes are regulated by TRAP exclusively at translational level (rev. in Babitzke et al., 2019).

The *pyr* operon contains 10 genes encoding all enzymes for UMP *de novo* synthesis under control of one constitutive promoter. **PyrR**, one of the two *B. subtilis* uracil-phosphoribosyltransferases, moonlights as RNA binding attenuation protein in regulation of this operon in three identical attenuation mechanisms (Turner et al., 1994, **Figure 8B**). An UMP-bound PyrR dimer stabilizes the anti-antiterminators comprising ARUCCAGAGAGGYU to allow formation of downstream terminators (rev. in Turnbough and Switzer, 2008; Turnbough, 2019). PyrR recognizes only conserved RNA sequences that are properly positioned in the correct secondary structure: terminal loop, top of the upper stem and a purine-rich internal bulge are crucial for efficient PyrR binding (Bonner et al., 2001) which involves a basic concave (Savacool and Switzer, 2002). The crystal structure of *B. caldolyticus* PyrR, which is active in *B. subtilis pyr* operon regulation (Chander et al., 2005) revealed an unexpected specific GMP binding antagonistic to RNA binding which suggests cross-regulation of the *pyr* operon by purines. Interestingly, in each of the three attenuation sites, NusA stabilizes *in vitro* pausing of the RNA polymerase at a major pause site to prevent formation of a complete antiterminator thus promoting formation of a PyrR binding loop (Zhang and Switzer, 2003). This might also play a role in termination of *pyr* transcription *in vivo*.

Antitermination proteins **GlcT, SacY, SacT**, and **LicT** contain an N-terminal RNA binding domain and regulatory domains PRD1 and PRD2 which are reversibly phosphorylated in response to the cognate carbon source (rev. in Stülke, 2002). In the absence of the carbon source, the cognate EII transporter of the phosphotransferase system phosphorylates PRD1 preventing dimerization. In its presence, EII dephosphorylates PRD1, while HPr phosphorylates PRD2, which allows dimerization and antiterminator binding (**Figure 8C**).

Crystal structures of inactive LicT revealed a wide swing movement of PRD2 causing a dimer opening that brings the phosphorylation sites to the protein surface (Graille et al., 2005). LicT interacts with two bulges in the antiterminator and the minor groove of the stem between them (Yang et al., 2002, **Figure 8C**). In GlcT, arrangement of the PRDs is under selective pressure to ensure a proper regulatory output (Himmel et al., 2012). For SacT, SacY, and LicT, specificity domains

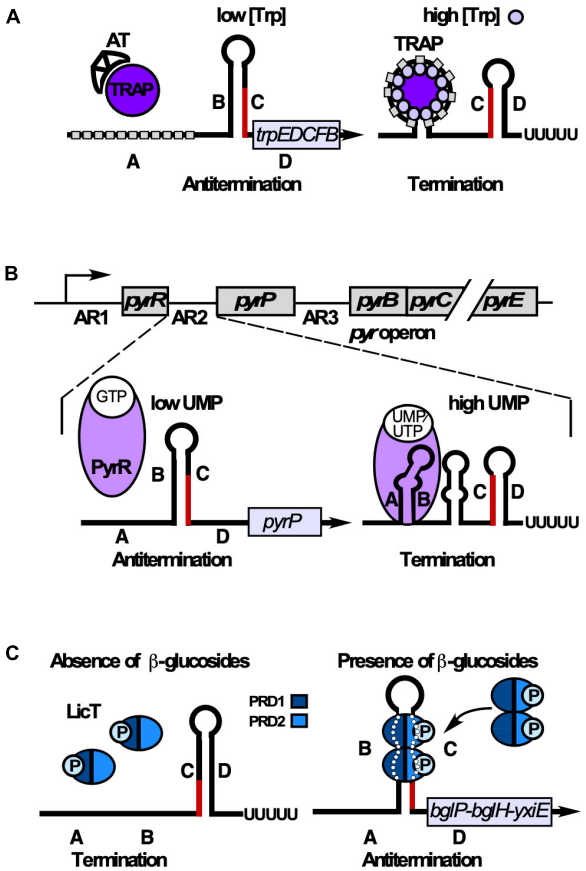


FIGURE 8 | RNA-binding proteins that induce termination or antitermination. Alternative folding occurs by base-pairing between regions B and C (antiterminator) or C and D (terminator). B and C overlap**. (A)** Control of tryptophan (Trp) biosynthesis. When Trp is scarce, anti-TRAP (AT) is expressed and sequesters TRAP. Regions B and C base-pair preventing terminator formation. At excess Trp, 11-mer TRAP binds Trp enabling each subunit to bind one G(U)AG triplet (gray rectangle) through contacts with aa K37, K56, and R58. This results in RNA wrapping around TRAP in regions A and B. Therefore, B cannot base-pair with C, allowing terminator formation. **(B)** Top: The 5′ end of the *pyr* operon mRNA contains 3 attenuation regions (AR1-3) upstream of the *pyrR,* the *pyrP* and the *pyrB* ORF, respectively. Bottom: Control of UMP synthesis. At low UMP/UTP but high GMP levels, PyrR binds GTP and is unable to bind RNA. B and C base-pair preventing formation of the terminator and thus, termination of the *pyr* operon. At high UMP/UTP levels, PyrR binds UMP or UTP, but not GTP and stabilizes the A-B antianti-terminator, thus preventing of B-C-antiterminator formation. Instead, the terminator stem-loop forms resulting in premature transcription termination. **(C)** Control of utilization of β-glucans/β-glucosides. In the absence of salicin (β-glucoside), PRD1 is phosphorylated by BglP, causing LicT inactivation (monomers). In the presence of salicin, BglP dephosphorylates PRD1 and transfers the phosphate to incoming salicin, and HPr phosphorylates PRD2. This allows LicT to dimerize and bind/stabilize the otherwise unstable antiterminator.

were identified that prevent a cross-talk between these systems (Hübner et al., 2011).

Glycerol-3P-bound **GlpP** binds the antiterminator and additionally stabilizes *glpD* mRNA (rev. in Stülke, 2002). Hexameric HutP (histidine utilization protein) prevents terminator formation in the *hut* operon. Coordination of L-histidine and Mg^{2+} activates HutP to bind two clusters of three NAG repeats spaced by 20 nt without undergoing further structural rearrangements (rev. in Kumarevel, 2007).

Aconitase CitB, a Metabolic Enzyme That Moonlights as RNA Binding Protein

Aconitase is an enzyme that operates in the tricarboxylic acid cycle (TCA) of all three kingdoms of life to convert citrate into isocitrate. For this activity, it needs a saturated iron-sulfur (FeS) cluster. Under iron limitation, the FeS cluster disassembles causing enzyme inactivation and, therefore, TCA shutdown. The inactive enzyme adopts an alternative conformation to bind at iron-response elements (IREs) located in the 5′ or 3′ UTR of mRNAs to repress translation or RNA degradation, respectively. The first bacterial aconitase discovered to bind RNA was *B. subtilis* CitB (Alén and Sonnenshein, 1999): Monomeric CitB binds *in vitro* at an IRE in the 3′ UTR of *qoxD* (cytochrome oxidase subunit) and another IRE between *feuA* and *feuB* (iron uptake). In addition, it binds at a stem-loop in the 3′ UTR of *gerE* mRNA encoding a transcription factor in sporulation, which might stabilize the RNA to ensure proper timing of spore coat formation (Serio et al., 2006). However, half-life measurements of *qoxD, feuAB* or *gerE* mRNAs have not yet been performed to confirm stabilization of these RNAs by CitB under iron limitation. Binding of CitB to the 5′ UTR of *citZ* mRNA resulted in destabilization of the three transcripts originating at the *citZ* promoter which is in line with prevention of translation (Pechter et al., 2013; **Figure 9**). Surprisingly, CitB bound *in vitro* both *citZ* mRNA and the negative control *hag* mRNA. To date, only the structure of the dual function human iron-regulatory protein 1 in complex with ferritin RNA has been solved (Walden et al., 2006). Under Fe^{2+} limitation, aconitase adopts a more open conformation with structural domains 3 and 4 extending perpendicularly from the central core composed of domains 1 and 2 (**Figure 9**). Direct contacts between the exposed IRE loop residues AGU and the domain 2/domain 3 interface provide specificity and stability to the interaction (Walden et al., 2006). Protein binding to the lower stem of the IRE is centered on a – sometimes bulged out – C nucleotide, which inserts into a pocket on the inner face of domain 4. Due to the high conservation of aconitase, similar structural changes most likely occur in *B. subtilis* CitB as well (**Figure 9**), although only in the case of *feuAB*, the IRE loop comprises an AGU motif.

SMALL PROTEIN-PROTEIN INTERACTIONS

Small proteins encoded by small ORFs comprise less than 50 aa and are involved in the regulation of a number of cellular functions including morphogenesis, cell division, enzymatic activities and stress response (rev. in Storz et al., 2014). Until recently, small ORFs had escaped the researchers' attention as their detection was difficult. With the advent of new technologies, thousands of translated small ORFs have been recently identified in prokaryotes and eukaryotes. However, the identification and characterization of small proteins is still challenging. Here, we review *B. subtilis* small proteins that interact with larger proteins (**Figure 10**).

Sda Interacts With KinA and Affects Signal Transduction

An example for small proteins localized at the inner membrane that interact with transmembrane sensor kinases is the 46 aa Sda. It binds to and inhibits KinA, the first histidine kinase in the phosphorelay that regulates *B. subtilis* sporulation. Upon starvation and stress, kinases KinA and KinB autophosphorylate and transfer their phosphates via Spo0F and Spo0B to the central regulator Spo0A. The *sda* gene was identified in a *dnaA1* suppressor mutant which was able to sporulate. Its promoter region contains multiple DnaA binding sites, and mutations in these regions or in *dnaA* affected *sda* expression (Burkholder et al., 2001). Structural studies showed that the Sda monomer is composed of two α-helices held together by an interhelix loop forming a helical hairpin (Rowland et al., 2004). The Sda-KinA interaction surface was mapped: Sda uses a hydrophobic surface portion formed by L21 and F25 to interact with the KinA dimerization/phosphotransfer (DHp) domain (Rowland et al., 2004). Several mechanisms for KinA inhibition by Sda were proposed: Sda might prevent KinA autophosphorylation by interrupting the phosphate transfer between the ATP binding site in the catalytic and the DHp domain (Rowland et al., 2004). Sda could also induce conformational changes in the DHp domain which in turn interferes with the KinA-Spo0F-phosphotransfer (Whitten et al., 2007). A later study suggested that Sda directly blocks this phosphotransfer as its binding site on KinA overlaps with that of Spo0F (Cunningham and Burkholder, 2009). KinA regulation through Sda is advantageous for *B. subtilis*, as it contributes to modulate sporulation initiation at multiple levels in response to unfavorable conditions.

MciZ Regulates Cell Division

Bacterial cytokinesis is initiated when the cell division machinery, the so called divisome, assembles at midcell. The divisome comprises about ten core proteins which help in cell membrane attachment and constriction (Lutkenhaus et al., 2012). Among them is FtsZ, which polymerizes into the Z ring in a GTP-dependent manner and exerts the actual constriction force (Shapiro et al., 2009). A number of small proteins interacting with FtsZ or other divisome components were identified. They also include MciZ (mother cell inhibitor of FtsZ) which was discovered in 2008 in a yeast two-hybrid screening using FtsZ as a bait protein (Handler et al., 2008). MciZ (40 aa) is produced under control of σ^E during sporulation (Handler et al., 2008). It impedes Z-ring formation by inhibiting FtsZ polymerization under moderate to low concentration. The crystal structure of the FtsZ-MciZ complex revealed a binding pocket for MciZ

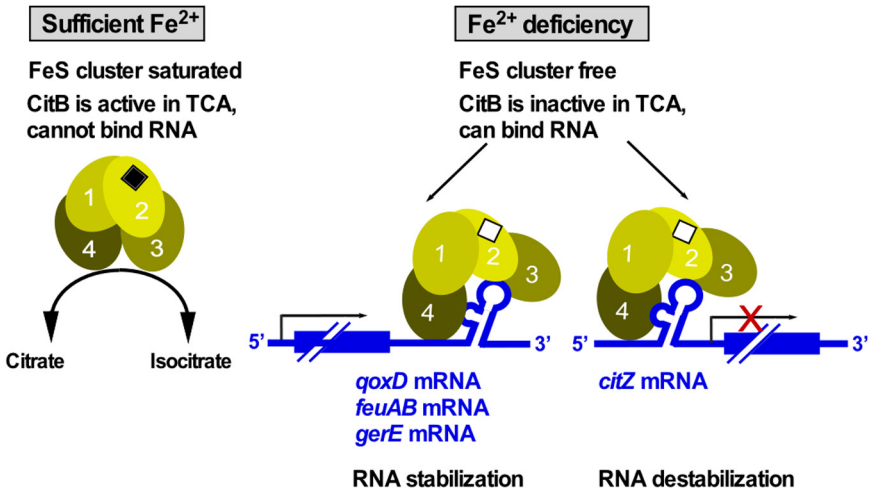


FIGURE 9 | CitB – a dual-function enzyme that moonlights in RNA binding. Under iron-rich conditions, the FeS cluster of CitB is saturated allowing it to function as metabolic enzyme in the TCA. Under iron limitation, the FeS cluster is unsaturated resulting in a conformational change that allows CitB to bind RNA at iron-response elements (IREs) or stem-loops in 3′ UTRs to stabilize the corresponding RNAs or at the 5′ UTR to destabilize the RNA. Depicted conformational changes and domain numbers are based on the crystal structure of the human iron regulatory protein 1 complexed with ferritin IRE-RNA (see text, Walden et al., 2006).

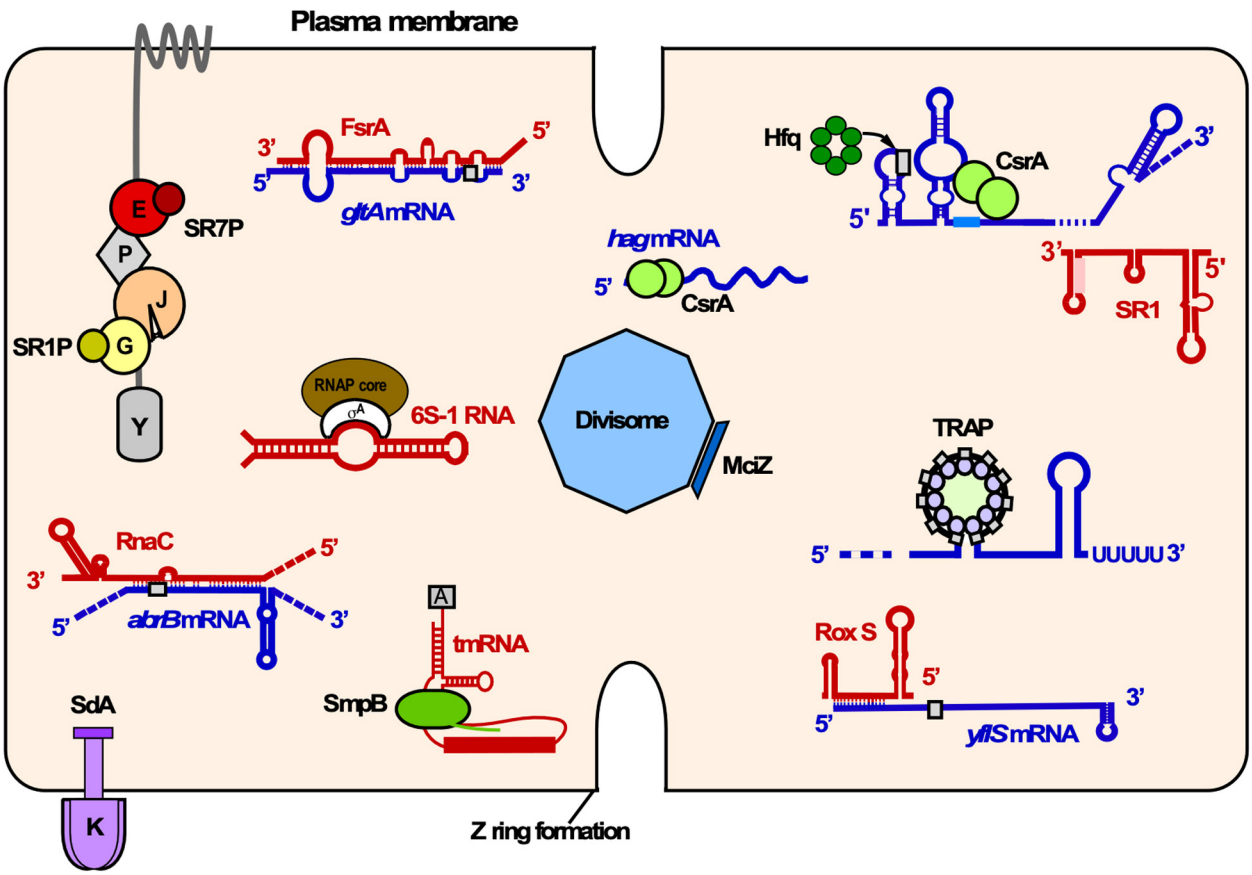


FIGURE 10 | Overview of RNA-RNA, RNA-protein and small protein-protein interactions in *B. subtilis*. Shown is the cytosol of *B. subtilis* bounded by the plasma membrane. Proteins associated with various cellular functions are labeled as follows: K, KinA; E, enolase; P, PfkA; J, RNase J1; G, GapA. Small proteins are depicted as rectangles or solid circles. Only components of the degradosome mentioned in the text are illustrated. Red, sRNAs and other regulatory RNAs; blue, mRNAs; gray rectangles, RBS; different shades of green, RNA binding proteins. Other colors, protein-small protein interactions. Y, RNase Y.

on one of the polymerization surfaces at the FtsZ C-terminus. MciZ binding inhibits FtsZ polymerization by steric hindrance (Bisson-Filho et al., 2015). Whereas the N-terminus of free MciZ is unstructured, the interaction with FtsZ induces the formation of two β-sheets (Bisson-Filho et al., 2015). At high concentration, MciZ sequesters FtsZ, while at sub-stoichiometric concentration, it acts by filament capping. Moreover, it can also cause filamentation *in vitro*. A recent study showed that MciZ can affect *B. subtilis* sporulation: Excessive amounts of MciZ produced intracellularly or added exogenously can not only decrease spore formation efficiency but also inhibit spore germination (Araújo-Bazán et al., 2019).

Two Small Proteins Modulate the Degradosome

The proposed *B. subtilis* degradosome is composed of major endoribonuclease RNase Y, RNases J1 and J2, PNPase, helicase CshA, the glycolytic enzymes enolase and phosphofructokinase and under certain circumstances, GapA (Commichau et al., 2009; Gimpel and Brantl, 2016). So far, two small proteins – SR1P and SR7P – have been identified which interact with degradosome components.

SR1P is a 39 aa protein encoded by the dual-function sRNA SR1 (see above). It interacts with GapA, one of the two glyceraldehyde-3P-dehydrogenases (GAPDHs) in *B. subtilis*. SR1P stabilizes *gapA* mRNA by preventing its rapid degradation under gluconeogenesis (Gimpel et al., 2010). In addition to its role in glycolysis, GapA has a moonlighting property: It binds both RNases J1 and Y. SR1P promotes GapA binding to RNase J1 and enhances RNase J1 activity *in vitro* on at least two substrates, SR5 and threonyl-tRNA. In addition, it affects SR5 stability *in vivo* (Gimpel and Brantl, 2016). This means, SR1P modifies the moonlighting activity of GapA. Using peptide mutants in co-elution experiments complemented by SR1P functionality tests in Northern blotting, the SR1P-GapA interaction surface was determined. SR1P attaches to GapA using aa contacts in a binding pocket formed by the C-terminal GapA helix 14. In addition, SR1P contacts the N-terminal GapA helix 1 (Gimpel et al., 2017). Interestingly, SR1 is transcribed under gluconeogenic conditions, when GapB is active (Fillinger et al., 2000) and the glycolytic activity of GapA is not needed. Newman et al. proposed that RNA degradation is connected to the metabolic state of the cell (Newman et al., 2012). Similarly, it was hypothesized that under nutrient limitation, glycolytic enzymes like enolase, phosphofructokinase that are part of the *B. subtilis* degradosome as well as GapA might sense nutritional stress and transfer this signal to the RNA degradation machinery (Gimpel and Brantl, 2016, 2017). This could modulate the global RNA turnover rate to conserve energy for other important cellular functions.

SR7P (39 aa) previously known as S1136 (Mars et al., 2015a) is encoded by the dual-function antisense RNA SR7 (Ul Haq et al., 2020). The *sr7* gene is located in the intergenic region between *tyrS* and *rpsD* and controlled by a σ^B-dependent promoter. SR7P is synthesized under five different stress conditions from the σ^B-dependent SR7 as well as constitutively from a *tyrS* mRNA processing product. SR7P interacts with enolase present in the degradosome. This interaction in turn improves enolase binding to RNase Y. The SR7P-Eno-RNase Y interaction is not bridged by RNA. The activity of RNase Y is significantly higher in the tri-component SR7P/Eno/RNase Y complex than in the Eno/RNase Y complex alone. This modulating effect of SR7P was demonstrated both *in vivo* and *in vitro*. In addition, SR7P impacts cell survival under selective stress conditions suggesting that it might play a specific role in stress response.

CONCLUSION AND PERSPECTIVES

Although a variety of intermolecular interactions have been investigated in *B. subtilis*, major cavities in our knowledge still exist. Only for four of 108 *trans*-encoded sRNAs and for 8 small proteins, interaction partners have been identified and biological functions elucidated. Furthermore, a global approach is required to ascertain if CsrA plays a similar role in *B. subtilis* as Hfq and ProQ in Gram-negatives. The function of the putative RNA chaperones FbpA, FbpB and FbpC has to be unraveled. Targets of the multitude of sRNAs have to be identified and their – perhaps sometimes novel – mechanisms of action elucidated. A considerable increase in the number of dual-function sRNAs can be anticipated. sRNAs might be found that bind enzymes or act directly on the genome like some eukaryotic siRNAs. Structures of more antiterminator/protein complexes and CitB have to be solved and the function of 6S-2 RNA uncovered. The further study of small proteins which is still in its infancy, will allow insights into regulatory networks comprising not only small proteins modulating large proteins or the membrane but also those binding small ligands.

AUTHOR CONTRIBUTIONS

SB wrote the sections on RNA-RNA interactions, 6S RNA, terminators/antiterminators, CitB., and designed **Figures 1–4**, **6–9**. IU wrote the sections on T-box riboswitch, small protein-protein interactions, and designed **Figure 5**. PM wrote the sections on RNA chaperones and tmRNA. SB and IU designed **Figure 10**. All authors approved the final version of the manuscript.

FUNDING

Work on this topic was supported by the grants BR1552/6-1 to 6-3, 7-1, 7-2, 8-1, 10-1, and 11-1 from the Deutsche Forschungsgemeinschaft (to SB).

ACKNOWLEDGMENTS

We thank all former coworkers of the AG Bakteriengenetik for their substantial experimental contributions. Furthermore, we apologize for not having mentioned all valuable contributions from colleagues in the field due to space limitations.

REFERENCES

Abe, T., Sakaki, K., Fujihara, A., Ujiie, H., Ushida, C., Himeno, H., et al. (2008). tmRNA-dependent *trans*-translation is required for sporulation in *Bacillus subtilis*. *Mol. Microbiol.* 69, 1491–1498. doi: 10.1111/j.1365-2958.2008.06381.x

Alén, C., and Sonnenshein, A. L. (1999). *Bacillus subtilis* aconitase is an RNA-binding protein. *Proc. Natl. Acad. Sci. U.S.A.* 96, 10412–10417. doi: 10.1073/pnas.96.18.10412

Altegoer, F., Rensing, S. A., and Bange, G. (2016). Structural basis for the CsrA-dependent modulation of translation initiation by an ancient regulatory protein. *Proc. Natl. Acad. Sci. U.S.A.* 113, 10168–10173. doi: 10.1073/pnas.1602425113

Antson, A. A., Dodson, E. J., Dodson, G., Greaves, R. B., Chen, X., and Gollnick, P. (1999). Structure of the *trp* RNA-binding attenuation protein, TRAP, bound to RNA. *Nature* 401, 235–242. doi: 10.1038/45730

Araújo-Bazán, L., Huecas, S., Valle, J., Andreu, D., and Andreu, J. M. (2019). Synthetic developmental regulator MciZ targets FtsZ across *Bacillus* species and inhibits bacterial division. *Mol. Microbiol.* 111, 965–980. doi: 10.1111/mmi.14198

Argaman, L., Hershberg, R., Vogel, J., Bejerano, G., Wagner, E. G., Margalit, H., et al. (2001). Novel small RNA-encoding genes in the intergenic regions of *Escherichia coli*. *Curr. Biol.* 11, 941–950. doi: 10.1016/s0960-9822(01)00270-6

Babitzke, P., Lai, Y. J., Renda, A. J., and Romeo, T. (2019). Posttranscription initiation control of gene mediated by bacterial RNA-binding proteins. *Annu. Rev. Microbiol.* 73, 43–67. doi: 10.1146/annurev-micro-020518-115907

Babitzke, P., and Yanofsky, C. (1993). Reconstitution of *Bacillus subtilis trp* attenuation in vitro with TRAP, the trp RNA-binding attenuation protein. *Proc. Natl. Acad. Sci. U.S.A.* 90, 133–137. doi: 10.1073/pnas.90.1.133

Beckmann, B. M., Burenina, O. Y., Hoch, P. G., Kubareva, E. A., Sharma, C. M., and Hartmann, R. K. (2011). In vivo and *in vitro* analysis of 6S RNA-templated short transcripts in *Bacillus subtilis*. *RNA Biol.* 8, 839–849. doi: 10.4161/rna.8.5.16151

Beckmann, B. M., Philipp, G. H., Marz, M., Willkomm, D. K., Salas, M., and Hartmann, R. K. (2012). pRNA-induced structural rearrangements triggers 6S-1 RNA release from RNA polymerase in *Bacillus subtilis*. *EMBO J.* 31, 1727–1738. doi: 10.1038/emboj.2012.23

Bisson-Filho, A. W., Discola, K. F., Castellen, P., Blasios, V., Martins, A., Sforça, M. L., et al. (2015). FtsZ filament capping by MciZ, a developmental regulator of bacterial division. *Proc. Natl. Acad. Sci. U.S.A.* 12, 2130–2138.

Bonner, E. R., D'Elia, J. N., Billips, B. K., and Switzer, R. L. (2001). Molecular recognition of *pyr* mRNA by the *Bacillus subtilis* attenuation regulatory protein PyrR. *Nucleic Acids Res.* 29, 4851–4165. doi: 10.1093/nar/29.23.4851

Brantl, S. (2004). Bacterial gene regulation: from transcription attenuation to riboswitches and ribozymes. *Trends Microbiol.* 12, 473–475. doi: 10.1016/j.tim.2004.09.008

Brantl, S. (2012). Bacterial type I toxin-antitoxin systems. *RNA Biol.* 9, 1488–1490. doi: 10.4161/rna.23045

Brantl, S. (2012). Acting antisense: plasmid- and chromosome-encoded sRNAs from Gram-positive bacteria. *Future Microbiol.* 7, 853–871. doi: 10.2217/fmb.12.59

Brantl, S., and Brückner, R. (2014). Small regulatory RNAs from low-GC Gram-positive bacteria. *RNA Biol.* 11, 443–456. doi: 10.4161/rna.28036

Brantl, S., and Jahn, N. (2015). sRNAs in bacterial type I and type III toxin/antitoxin systems. *FEMS Microbiol. Rev.* 39, 413–427. doi: 10.1093/femsre/fuv003

Brantl, S., and Müller, P. (2019). Toxin antitoxin systems in *Bacillus subtilis*. *Toxins (Basel)* 11:E262.

Burenina, O. Y., Hoch, P. G., Damm, K., Salas, M., Zatsepin, T. S., Lechner, M., et al. (2014). Mechanistic comparison of *Bacillus subtilis* 6S-1 and 6S-2 RNAs – commonalities and differences. *RNA* 20, 348–359. doi: 10.1261/rna.042077.113

Burkholder, W. F., Kurtser, I., and Grossman, A. D. (2001). Replication initiation proteins regulate a developmental checkpoint in *Bacillus subtilis*. *Cell* 104, 269–279. doi: 10.1016/s0092-8674(01)00211-2

Cavanagh, A. T., Sperger, J. M., and Wassarman, K. M. (2012). Regulation of 6S RNA by pRNA synthesis is required for efficient recovery from stationary phase in *E. coli* and *B. subtilis*. *Nucleic Acids Res.* 40, 2234–2246. doi: 10.1093/nar/gkr1003

Cavanagh, A. T., and Wassarman, K. M. (2013). 6S-1 RNA function leads to a delay in sporulation in *Bacillus subtilis*. *J. Bacteriol.* 195, 2079–2086. doi: 10.1128/jb.00050-13

Chander, P., Halbig, K. M., Miller, J. K., Fields, C. J., Bonner, H. K. S., Grabner, G. K., et al. (2005). Structure of the nucleotide complex of PyrR, the Pyr attenuation protein from *Bacillus caldolyticus*, suggests dual regulation by pyrimidine and purine nucleotides. *J. Bacteriol.* 187, 1773–1782. doi: 10.1128/jb.187.5.1773-1782.2005

Chen, J., Wassarman, K. M., Feng, S., Leon, K., Feklistov, A., Winkelman, J. T., et al. (2017). 6S RNA mimics B-form DNA to regulate *Escherichia coli* RNA polymerase. *Mol. Cell* 68, 388–397.

Commichau, F. M., Rothe, F. M., Herzberg, C., Wagner, E., Hellwig, D., Lehnik-Habrink, M., et al. (2009). Novel activities of glycolytic enzymes in *Bacillus subtilis*: interactions with essential proteins involved in mRNA processing. *Mol. Cell Proteomics.* 8, 1350–1360. doi: 10.1074/mcp.m800546-mcp200

Cunningham, K. A., and Burkholder, W. F. (2009). The histidine kinase inhibitor Sda binds near the site of autophosphorylation and may sterically hinder autophosphorylation and phosphotransfer to Spo0F. *Mol. Microbiol.* 71, 659–677. doi: 10.1111/j.1365-2958.2008.06554.x

Czaplewski, L. G., North, A. I., Smith, M. C., Baumberg, S., and Stockley, P. G. (1992). Purification and initial characterization of AhrC: the regulator of arginine metabolism genes in *Bacillus subtilis*. *Mol. Microbiol.* 6, 267–275. doi: 10.1111/j.1365-2958.1992.tb02008.x

Dambach, M. C., and Winkler, W. C. (2013). Association of RNAs with *B. subtilis* Hfq. *PLoS One* 8:e55156. doi: 10.1371/journal.pone.0055156

Durand, S., Braun, F., Helfer, A. C., Romby, P., and Condon, C. (2017). sRNA-mediated activation of gene expression by inhibiton of 5'-3' exonulceolytic mRNA degradation. *Elife* 17:e23602.

Durand, S., Braun, F., Lioliou, E., Romilly, C., Helfer, A. C., Kuhn, L., et al. (2015). A nitric oxide regulated small RNA controls expression of genes involved in redox homeostasis in *Bacillus subtilis*. *PLoS Genet.* 11:e1004957. doi: 10.1371/journal.pgen.1004957

Durand, S., Gilet, L., and Condon, C. (2012a). The essential function of *B. subtilis* RNase III is to silence foreign toxin genes. *PLoS Genet*. 8:e1003181. doi: 10.1371/journal.pgen.1003181

Durand, S., Jahn, N., Condon, C., and Brantl, S. (2012b). Type I toxin-antitoxin systems in *Bacillus subtilis*. *RNA Biol.* 9, 1491–1497. doi: 10.4161/rna.22358

Eiumphungporn, W., and Helmann, J. D. (2009). Extracytoplasmic function sigma factors regulate expression of the *Bacillus subtilis yabE* gene via a cis-acting antisense RNA. *J. Bacteriol.* 191, 1101–1105. doi: 10.1128/jb.01530-08

Fillinger, S., Boschi-Muller, S., Azza, S., Dervyn, E., Branlant, G., and Aymerich, S. (2000). Two glyceraldehyde-3-phosphate dehydrogenases with opposite physiological roles in a nonphotosynthetic bacterium. *J. Biol. Chem.* 275, 14031–14037. doi: 10.1074/jbc.275.19.14031

Fujihara, A., Tomatsu, H., Inagaki, S., Tadaki, T., Ushida, C., Himeno, H., et al. (2002). Detection of tmRNA-mediated *trans*-translation products in *Bacillus subtilis*. *Genes Cells* 7, 343–350. doi: 10.1046/j.1365-2443.2002.00523.x

Gaballa, A., Antelmann, H., Aguilar, C., Khakh, S. K., Song, K. B., Smaldone, G. T., et al. (2008). The *Bacillus subtilis* iron-sparing response is mediated by a Fur-regulated small RNA and three small, basic proteins. *Proc. Natl. Acad. Sci. U.S.A.* 105, 11927–11932. doi: 10.1073/pnas.0711752105

Gimpel, M., and Brantl, S. (2016). Dual-function sRNA encoded peptide SR1P modulates moonlighting activity of *B. subtilis* GapA. *RNA Biol.* 13, 916–926. doi: 10.1080/15476286.2016.1208894

Gimpel, M., and Brantl, S. (2017). Dual-function small regulatory RNAs in bacteria. *Mol. Microbiol.* 103, 387–397. doi: 10.1111/mmi.13558

Gimpel, M., Heidrich, N., Mäder, U., Krügel, H., and Brantl, S. (2010). A dual-function sRNA from *Bacillus subtilis:* SR1 acts as a peptide-encoding mRNA on the *gapA* operon. *Mol. Microbiol.* 76, 990–1009. doi: 10.1111/j.1365-2958.2010.07158.x

Gimpel, M., Maiwald, C., Wiedemann, C., Görlach, M., and Brantl, S. (2017). Characterization of the interaction between the small RNA-encoded peptide SR1P and GapA from *Bacillus subtilis*. *Microbiology* 163, 1248–1259. doi: 10.1099/mic.0.000505

Gimpel, M., Preis, H., Barth, E., Gramzow, L., and Brantl, S. (2012). SR1 – a small RNA with two remarkably conserved functions. *Nucleic Acids Res.* 40, 11659–11672. doi: 10.1093/nar/gks895

Graille, M., Zhou, C. Z., Brechot, V. R., Collinet, B., Declerck, N., and van Tilbeurgh, H. (2005). Activation of the LicT transcriptional antiterminator involves a domain swing/lock mechanism provoking massive structural changes. *J. Biol. Chem.* 280, 14780–14789. doi: 10.1074/jbc.m414642200

Grigg, J. C., and Ke, A. (2013). Structural determinants for geometry and information decoding of tRNA by T box leader RNA. *Cell* 21, 2025–2032. doi: 10.1016/j.str.2013.09.001

Grundy, F. J., Rollins, S. M., and Henkin, T. M. (1994). Interaction between the acceptor end of tRNA and the T box stimulates antitermination in the *Bacillus subtilis tyrS* gene: a new role for the discriminator base. *J. Bacteriol*. 176, 4518–4526. doi: 10.1128/jb.176.15.4518-4526.1994

Hämmerle, H., Amman, F., Veèerek, B., Stülke, J., Hofacker, I., and Bläsi, U. (2014). Impact of Hfq on the *Bacillus subtilis* transcriptome. *PLoS One* 9:e98661. doi: 10.1371/journal.pone.0098661

Handler, A. A., Lim, J. E., and Losick, R. (2008). Peptide inhibitor of cytokinesis during sporulation in *Bacillus subtilis*. *Mol. Microbiol.* 68, 588–599. doi: 10.1111/j.1365-2958.2008.06173.x

Heidrich, N., and Brantl, S. (2003). Antisense-RNA mediated transcriptional attenuation: importance of a U-turn loop structure in the target RNA of plasmid pIP501 for inhibition by the antisense RNA. *J. Mol. Biol.* 153, 420–427. doi: 10.1099/mic.0.2006/002329-0

Heidrich, N., Chinali, A., Gerth, U., and Brantl, S. (2006). The small untranslated RNA SR1 from the *B. subtilis* genome is involved in the regulation of arginine catabolism. *Mol. Microbiol.* 62, 520–536. doi: 10.1111/j.1365-2958.2006.05384.x

Heidrich, N., Moll, I., and Brantl, S. (2007). *In vitro* analysis of the interaction between the small RNA SR1 and its primary target *ahrC* mRNA. *Nucleic Acids Res.* 35, 4331–4346. doi: 10.1093/nar/gkm439

Henkin, T. M., Glass, B. L., and Grundy, F. J. (1992). Analysis of the *Bacillus subtilis tyrS* gene: conservation of a regulatory sequence in multiple tRNA synthetase genes. *J. Bacteriol*. 174, 1299–1306. doi: 10.1128/jb.174.4.1299-1306.1992

Himeno, H., Kurita, D., and Muto, A. (2014). tmRNA-mediated *trans*-translation as the major ribosome rescue system in a bacterial cell. *Front. Genet*. 5:66.

Himeno, H., Nameki, N., Kurita, D., Muto, A., and Abo, T. (2015). Ribosome rescue systems in bacteria. *Biochimie* 114, 102–112. doi: 10.1016/j.biochi.2014.11.014

Himmel, S., Zschiedrich, C. P., Becker, S., Hsiao, H. H., Wolff, S., Diethmaier, C., et al. (2012). Determinants of interaction specificity of the *Bacillus subtilis* GlcT antitermination protein: functionality and phosphorylation specificity depend on the arrangement of the regulatory domains. *J. Biol. Chem*. 287, 27731–27742. doi: 10.1074/jbc.m112.388850

Hoch, P. G., Schlereth, J., Lechner, M., and Hartmann, R. K. (2016). *Bacillus subtilis* 6S-2 RNA serves as a template for short transcripts *in vivo*. RNA 22, 614–622. doi: 10.1261/rna.055616.115

Hübner, S., Declerck, N., Diethmaier, C., Le Coq, D., Aymerich, S., and Stülke, J. (2011). Prevention of cross-talk in conserved regulatory systems: identification of specificity determinants in RNA-binding anti-termination proteins of the BglG family. *Nucleic Acids Res.* 39, 4360–4372. doi: 10.1093/nar/gkr021

Irnov, I., Sharma, C. M., Vogel, J., and Winkler, W. C. (2010). Identification of regulatory RNAs in *Bacillus subtilis*. *Nucleic Acids Res.* 38, 6637–6651. doi: 10.1093/nar/gkq454

Jagtap, C. B., Kumar, P., and Rao, K. K. (2016). *Bacillus subtilis* Hfq: a role in chemotaxis and motility. *J. Biosci*. 41, 347–358. doi: 10.1007/s12038-016-9618-9

Jahn, N., and Brantl, S. (2013). One antitoxin – two functions: SR4 controls toxin mRNA decay and translation. *Nucleic Acids Res.* 41, 9870–9880. doi: 10.1093/nar/gkt735

Jahn, N., and Brantl, S. (2016). Heat shock induced refolding entails rapid degradation of *bsrG* toxin mRNA by RNases Y and J1. *Microbiology* 162, 590–599. doi: 10.1099/mic.0.000247

Jahn, N., Brantl, S., and Strahl, H. (2015). Against the mainstream: the membrane associated type I toxin BsrG from *Bacillus subtilis* interferes with cell envelope biosynthesis without increasing membrane permeability. *Mol. Microbiol.* 98, 651–666. doi: 10.1111/mmi.13146

Jahn, N., Preis, H., Wiedemann, C., and Brantl, S. (2012). BsrG/SR4 from *Bacillus subtilis* – the first temperature-dependent type I toxin-antitoxin system. *Mol. Microbiol.* 83, 579–598. doi: 10.1111/j.1365-2958.2011.07952.x

Kavita, K., de Mets, F., and Gottesman, S. (2018). New aspects of RNA-based regulation by Hfq and its partner sRNAs. *Curr. Opin. Microbiol.* 42, 53–61. doi: 10.1016/j.mib.2017.10.014

Keiler, K. C. (2015). Mechanisms of ribosome rescue in bacteria. *Nat. Rev. Microbiol*. 13, 285–296.

Kolkman, M. A. B., and Ferrari, E. (2004). The fate of extracellular proteins tagged by the SsrA system of *Bacillus subtilis*. *Microbiology* 150, 427–436. doi: 10.1099/mic.0.26388-0

Kreuzer, K. D., and Henkin, T. M. (2018). The T-Box riboswitch: tRNA as an effector to modulate gene regulation. *Microbiol. Spectr*. 6:1128.

Kumarevel, T. (2007). Structural insights of HutP-mediated regulation of transcription of the *hut* operon in *Bacillus subtilis*. *Biophys. Chem.* 128, 1–12. doi: 10.1016/j.bpc.2007.03.003

Li, S., Su, Z., Lehmann, J., Stamatopoulou, V., Giarimoglou, N., Henderson, F. E., et al. (2019). Structural basis of amino acid surveillance by higher-order tRNA-mRNA interactions. *Nat. Struct. Mol. Biol.* 26, 1094–1105. doi: 10.1038/s41594-019-0326-7

Licht, A., and Brantl, S. (2009). The transcriptional repressor CcpN from *B. subtilis* uses different repression mechanisms at different promoters. *J. Biol. Chem*. 284, 30032–30038. doi: 10.1074/jbc.m109.033076

Licht, A., Golbik, R., and Brantl, S. (2008). Identification of ligands affecting the activity of the transcriptional repressor CcpN from *Bacillus subitlis*. *J. Mol. Biol.* 380, 17–30. doi: 10.1016/j.jmb.2008.05.002

Licht, A., Preis, S., and Brantl, S. (2005). Implication of CcpN in the regulation of a novel untranslated RNA (SR1) in *B. subtilis*. *Mol. Microbiol.* 58, 189–206. doi: 10.1111/j.1365-2958.2005.04810.x

Luo, Y., and Helmann, J. D. (2012). A sigma D-dependent antisense transcript modulates expression of the cyclic-di-AMP hydrolase GdpP in *Bacillus subtilis*. *Microbiology* 158, 2732–2741. doi: 10.1099/mic.0.062174-0

Lutkenhaus, J., Pichoff, S., and Du, S. (2012). Bacterial cytokinesis: from Z ring to divisome. *Cytoskeleton* 69, 778–790. doi: 10.1002/cm.21054

Mars, R. A., Mendonça, K., Denham, E. L., and van Dijl, J. M. (2015a). The reduction in small ribosomal subunit abundance in ethanol-stressed cells of *Bacillus subtilis* is mediated by a SigB-dependent antisense RNA. *Biochim. Biophys. Acta* 1853, 2553–2559. doi: 10.1016/j.bbamcr.2015.06.009

Mars, R. A., Nicolas, P., Ciccolini, M., Reilman, E., Reder, A., Schaffer, M., et al. (2015b). Small regulatory RNA-induced growth rate heterogeneity of *Bacillus subtilis*. *PLoS Genet*. 11:e1005046. doi: 10.1371/journal.pgen.1005046

McAdams, N. M., Patterson, A., and Gollnick, P. (2017). Identification of a residue (Glu60) in TRAP required for inducing efficient transcription termination at the trp attenuator independent of binding tryptophan and RNA. *J. Bacteriol.* 199, e710–e716.

Meißner, C., Jahn, N., and Brantl, S. (2015). In vitro characterization of the type I toxin-antitoxin system *bsrE*/SR5 from *Bacillus subtilis*. *J. Biol. Chem*. 291, 560–571. doi: 10.1074/jbc.m115.697524

Miller, M. R., and Buskirk, A. R. (2014). The SmpB C-terminal tail helps tmRNA to recognize and enter stalled ribosomes. *Front. Microbiol.* 5:462.

Mukherjee, S., Oshiro, R. T., Yakhnin, H., Babitzke, P., and Kearns, D. B. (2016). FliW antagonizes CsrA RNA binding by a noncompetitive allosteric mechanism. *Proc. Natl. Acad. Sci. U.S.A.* 113, 9870–9875. doi: 10.1073/pnas.1602455113

Mukherjee, S., Yakhnin, H., Kysela, D., Sokoloski, J., Babitzke, P., and Kearns, D. B. (2011). CsrA-FliW interaction governs flagellin homeostasis and a checkpoint on flagellar morphogenesis in *Bacillus subtilis*. *Mol. Microbiol.* 82, 447–461.

Müller, P., Gimpel, M., Wildenhain, T., and Brantl, S. (2019). A new role for CsrA: promotion of complex formation between an sRNA and its mRNA target in *Bacillus subtilis*. *RNA Biol.* 16, 972–987. doi: 10.1080/15476286.2019.1605811

Müller, P., Jahn, N., Ring, C., Maiwald, C., Neubert, R., Meißner, C., et al. (2016). A multistress responsive type I toxin-antitoxin system: *bsrE*/SR5 from the *B. subtilis* chromosome. *RNA Biol.* 13, 511–523. doi: 10.1080/15476286.2016.1156288

Muto, A., Fujihara, A., Ito, K. I., Matsuno, J., Ushida, C., and Himeno, H. (2000). Requirement of transfer-messenger RNA for the growth of *Bacillus subtilis* under stresses. *Genes Cells*. 5, 627–635. doi: 10.1046/j.1365-2443.2000.00356.x

Newman, J. A., Hewitt, L., Rodrigues, C., Solovyova, A. S., Harwood, C. R., and Lewis, R. J. (2012). Dissection of the network of interactions that links RNA processing with glycolysis in the *Bacillus subtilis* degradosome. *J Mol. Biol.* 416, 121–136. doi: 10.1016/j.jmb.2011.12.024

Noone, D., Salzberg, L. I., Botella, E., Basell, K., Becher, D., Antelmann, H., et al. (2014). A highly unstable transcript makes CwlO D,L endopeptidase expression responsive to growth conditions in *Bacillus subtilis*. *J. Bacteriol.* 196, 237–247. doi: 10.1128/jb.00986-13

Olejniczak, M., and Storz, G. (2017). ProQ/FinO-domain proteins: another ubiquitous family of RNA matchmakers. *Mol. Microbiol.* 104, 905–915. doi: 10.1111/mmi.13679

Otridge, J., and Gollnick, P. (1993). MtrB from *Bacillus subtilis* binds specifically to *trp* leader RNA in a tryptophan dependent manner. *Proc. Natl. Acad. Sci. U.S.A.* 90, 128–132. doi: 10.1073/pnas.90.1.128

Pechter, K. B., Meyer, F. M., Serio, A. W., Stülke, J., and Sonenshein, A. L. (2013). Two roles for aconitase in the regulation of tricarboxylic acid branch gene expression in *Bacillus subtilis*. *J. Bacteriol.* 195, 1525–1537. doi: 10.1128/jb.01690-12

Potter, K., Merlino, N. M., Jacobs, T., and Gollnick, P. (2011). TRAP binding to the *Bacillus subtilis trp* leader region RNA causes efficient transcription termination at a weak intrinsic terminator. *Nucleic Acids Res.* 39, 2092–2102. doi: 10.1093/nar/gkq965

Rasmussen, S., Nielsen, H. B., and Jarmer, H. (2009). Transcriptionally active regions in the genome of *Bacillus subtilis*. *Mol. Microbiol.* 73, 1043–1057. doi: 10.1111/j.1365-2958.2009.06830.x

Rath, H., Reder, A., Hoffmann, T., Hammer, E., Seubert, A., Bremer, E., et al. (2020). Management of osmoprotectant uptake hierarchy in *Bacillus subtilis* via a SigB-dependent antisense RNA. *Front. Microbiol.* 11:622.

Reif, C., Löser, C., and Brantl, S. (2018). *Bacillus subtilis* type I antitoxin SR6 promotes degradation of toxin *yonT* mRNA and is required to prevent toxin *yoyJ* overexpression. *Toxins (Basel)* 10:E74.

Rochat, T., Delumeau, O., Figueroa-Bossi, N., Noirot, P., Bossi, L., Dervyn, E., et al. (2015). Tracking the elusive function of *Bacillus subtilis* Hfq. *PLoS One* 10:e0124977. doi: 10.1371/journal.pone.0124977

Rowland, S. L., Burkholder, W. F., Cunningham, K. A., Maciejewski, M. W., Grossman, A. D., and King, G. F. (2004). Structure and mechanism of action of Sda, an inhibitor of the histidine kinases that regulate initiation of sporulation in *Bacillus subtilis*. *Mol. Cell* 13, 689–701. doi: 10.1016/s1097-2765(04)00084-x

Savacool, H. K., and Switzer, R. L. (2002). Characterization of the interaction of *Bacillus subtilis* PyrR with *pyr* mRNA by site-directed mutagenesis of the protein. *J. Bacteriol.* 184, 2521–2528. doi: 10.1128/jb.184.9.2521-2528.2002

Schmalisch, M., Maiques, E., Nikolov, L., Camp, A. H., Chevreux, B., Muffler, A., et al. (2010). Small genes under sporulation control in the *Bacillus subtilis* genome. *J. Bacteriol.* 192, 5402–5412. doi: 10.1128/jb.00534-10

Serio, A. W., Pechter, K. B., and Sonenshein, A. L. (2006). *Bacillus subtilis* aconitase is required for efficient late-sporulation gene expression. *J. Bacteriol.* 188, 6396–6405. doi: 10.1128/jb.00249-06

Shapiro, L., McAdams, H. H., and Losick, R. (2009). Why and how bacteria localize proteins. *Science* 326, 1225–1228. doi: 10.1126/science.1175685

Shimokawa-Chiba, N., Müller, C., Fujiwara, K., Beckert, B., Ito, K., Wilson, D. N., et al. (2019). Release factor-dependent ribosome rescue by BrfA in the Gram-positive bacterium *Bacillus subtilis*. *Nat. Commun.* 10:5397.

Shin, J. H., and Price, C. W. (2007). The SsrA-SmpB ribosome rescue system is important for growth of *Bacillus subtilis* at low and high temperatures. *J. Bacteriol.* 189, 3729–3737. doi: 10.1128/jb.00062-07

Silvaggi, J. M., Perkins, J. B., and Losick, R. (2005). Small untranslated RNA antitoxin in *Bacillus subtilis*. *J. Bacteriol.* 187, 6641–6650. doi: 10.1128/jb.187.19.6641-6650.2005

Smaldone, G. T., Revelles, O., Gaballa, A., Sauer, U., Antelmann, H., and Helmann, J. D. (2012a). A global investigation of the *Bacillus subtilis* iron-sparing response identifies major changes in metabolism. *J. Bacteriol.* 194, 2594–2605. doi: 10.1128/jb.05990-11

Smaldone, G. T., Antelmann, H., Gaballa, A., and Helmann, J. D. (2012b). The FsrA sRNA and FbpB protein mediate the iron-dependent induction of the *Bacillus subtilis lutABC* iron-sulfur-containing oxidases. *J. Bacteriol.* 194, 2586–2593. doi: 10.1128/jb.05567-11

Someya, T., Baba, S., Fujimoto, M., Kawai, G., Kumasaka, T., and Nakamura, K. (2012). Crystal structure of Hfq from *Bacillus subtilis* in complex with SELEX-derived RNA aptamer: insight into RNA-binding properties of bacterial Hfq. *Nucleic Acids Res.* 40, 1856–1867. doi: 10.1093/nar/gkr892

Storz, G., Wolf, Y. I., and Ramamurthi, K. S. (2014). Small proteins can no longer be ignored. *Annu. Rev. Biochem.* 83, 753–777. doi: 10.1146/annurev-biochem-070611-102400

Stülke, J. (2002). Control of transcription termination in bacteria by RNA-binding proteins that modulate RNA structures. *Arch. Microbiol.* 177, 433–440.

Turnbough, C. L. Jr. (2019). Regulation of bacterial gene expression by transcription attenuation. *Microbiol. Mol. Biol. Rev.* 83:e00019-19.

Turnbough, C. L. Jr., and Switzer, R. L. (2008). Regulation of pyrimidine biosynthetic gene expression bacteria: repression without repressors. *Microbiol. Mol. Biol. Rev.* 72, 266–300. doi: 10.1128/mmbr.00001-08

Turner, R. J., Lu, Y., and Switzer, R. L. (1994). Regulation of the *Bacillus subtilis* pyrimidine biosynthetic (*pyr*) gene cluster by an autogenous transcriptional attenuation mechanism. *J. Bacteriol.* 176, 3708–3722. doi: 10.1128/jb.176.12.3708-3722.1994

Ujiie, H., Matsutani, T., Tomatsu, H., Fujihara, A., Ushida, C., Miwa, Y., et al. (2008). Trans-translation is involved in the CcpA-dependent tagging and degradation of TreP in *Bacillus subtilis*. *J Biochem.* 145, 59–66. doi: 10.1093/jb/mvn143

Ul Haq, I., Müller, P., and Brantl, S. (2020). SR7– a dual function antisense RNA from *Bacillus subtilis*. *RNA Biol.* 5, 1–14. doi: 10.1080/15476286.2020.1798110

Vakulskas, C. A., Potts, A. H., Babitzke, P., Ahmer, B. M., and Romeo, T. (2015). Regulation of bacterial virulence by Csr (Rsm) systems. *Microbiol. Mol. Biol. Rev.* 79, 193–224. doi: 10.1128/mmbr.00052-14

Wagner, E. G., and Romby, P. (2015). Small RNAs in bacteria and archaea: who they are, what they do, and how they do it. *Adv. Genet.* 909, 133–208.

Walden, W. E., Selezneva, A. I., Dupuy, J., Volbeda, A., Fontecilla-Camps, J. C., Theil, E., et al. (2006). Structure of dual function iron regulatory protein 1 complexed with ferritin IRE-RNA. *Science* 314, 1903–1908. doi: 10.1126/science.1133116

Wassarman, K. M., Repoila, F., Rosenow, C., Storz, G., and Gottesman, S. (2001). Identification of novel small RNAs using comparative genomics and microarrays. *Genes Dev.* 15, 1637–1651. doi: 10.1101/gad.901001

Wassarman, K. M., and Storz, G. (2000). 6S RNA regulates *E. coli* RNA polymerase activity. *Cell* 101, 613–623. doi: 10.1016/s0092-8674(00)80873-9

Whitten, A. E., Jacques, D. A., Hammouda, B., Hanley, T., King, G. F., Guss, J. M., et al. (2007). The structure of the KinA-Sda complex suggests an allosteric mechanism of histidine kinase inhibition. *J. Mol. Biol.* 368, 407–420. doi: 10.1016/j.jmb.2007.01.064

Wiegert, T., and Schumann, W. (2001). SsrA-mediated tagging in *Bacillus subtilis*. *J. Bacteriol.* 183, 3885–3889. doi: 10.1128/jb.183.13.3885-3889.2001

Yakhnin, H., Pandit, P., Petty, T. J., Baker, C. S., Romeo, T., and Babitzke, P. (2007). CsrA of *Bacillus subtilis* regulates translation initiation of the gene encoding the flagellin protein (*hag*) by blocking ribosome binding. *Mol. Microbiol.* 64, 1605–1620. doi: 10.1111/j.1365-2958.2007.05765.x

Yang, Y., Declerck, N., Manival, X., Aymerich, S., and Kochoyan, M. (2002). Solution structure of the LicT RNA antitermination complex: CAT clamping RAT. *EMBO J.* 21, 1987–1997. doi: 10.1093/emboj/21.8.1987

Zhang, H., and Switzer, R. L. (2003). Transcriptional pausing in the *Bacillus subtilis pyr* operon in vitro: a role in transcriptional attenuation? *J. Bacteriol.* 185, 4764–4771. doi: 10.1128/JB.185.16.4764-4771.2003

Zhang, J., Chetnani, B., Cormack, E. D., Alonso, D., Liu, W., Mondragón, A., et al. (2018). Specific structural elements of the T-box riboswitch drive the two-step binding of the tRNA ligand. *eLife* 7:e39518.

Zhang, J., and Ferré-D'Amaré, A. R. (2013). Co-crystal structure of a T-box riboswitch stem I domain in complex with its cognate tRNA. *Nature* 500, 363–366. doi: 10.1038/nature12440

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Ul Haq, Müller and Brantl. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 11 August 2020
doi: 10.3389/fmolb.2020.00185

Conjugative Coupling Proteins and the Role of Their Domains in Conjugation, Secondary Structure and *in vivo* Subcellular Location

*Itxaso Álvarez-Rodríguez[1,2], Begoña Ugarte-Uribe[1†], Igor de la Arada[1,2], José Luis R. Arrondo[1,2], Carlos Garbisu[3] and Itziar Alkorta[1,2]**

[1] Department of Biochemistry and Molecular Biology, University of the Basque Country (UPV/EHU), Leioa, Spain, [2] Instituto Biofisika (UPV/EHU, CSIC), University of the Basque Country (UPV/EHU), Spanish Research Council (CSIC), Leioa, Spain, [3] NEIKER, Soil Microbial Ecology Group, Department of Conservation of Natural Resources, Derio, Spain

OPEN ACCESS

Edited by:
Tatiana Venkova,
Fox Chase Cancer Center,
United States

Reviewed by:
Elisabeth Grohmann,
Beuth Hochschule für Technik Berlin,
Germany
Graciela Castro Escarpulli,
National School of Biological
Sciences, National Polytechnic
Institute of Mexico (IPN), Mexico

***Correspondence:**
Itziar Alkorta
itzi.alkorta@ehu.eus

†Present address:
Begoña Ugarte-Uribe,
EMBL Heidelberg Cell Biology
and Biophysics, Heidelberg, Germany

Specialty section:
This article was submitted to
Molecular Recognition,
a section of the journal
Frontiers in Molecular Biosciences

Received: 28 May 2020
Accepted: 14 July 2020
Published: 11 August 2020

Citation:
Álvarez-Rodríguez I,
Ugarte-Uribe B, de la Arada I,
Arrondo JLR, Garbisu C and Alkorta I
(2020) Conjugative Coupling Proteins
and the Role of Their Domains
in Conjugation, Secondary Structure
and in vivo Subcellular Location.
Front. Mol. Biosci. 7:185.
doi: 10.3389/fmolb.2020.00185

Type IV Coupling Proteins (T4CPs) are essential elements in many type IV secretion systems (T4SSs). The members of this family display sequence, length, and domain architecture heterogeneity, being the conserved Nucleotide-Binding Domain the motif that defines them. In addition, most T4CPs contain a Transmembrane Domain (TMD) in the amino end and an All-Alpha Domain facing the cytoplasm. Additionally, a few T4CPs present a variable domain at the carboxyl end. The structural paradigm of this family is $TrwB_{R388}$, the T4CP of conjugative plasmid R388. This protein has been widely studied, in particular the role of the TMD on the different characteristics of $TrwB_{R388}$. To gain knowledge about T4CPs and their TMD, in this work a chimeric protein containing the TMD of $TraJ_{pKM101}$ and the cytosolic domain of $TrwB_{R388}$ has been constructed. Additionally, one of the few T4CPs of mobilizable plasmids, $MobB_{CloDF13}$ of mobilizable plasmid CloDF13, together with its TMD-less mutant MobBΔTMD have been studied. Mating studies showed that the chimeric protein is functional *in vivo* and that it exerted negative dominance against the native proteins $TrwB_{R388}$ and $TraJ_{pKM101}$. Also, it was observed that the TMD of $MobB_{CloDF13}$ is essential for the mobilization of CloDF13 plasmid. Analysis of the secondary structure components showed that the presence of a heterologous TMD alters the structure of the cytosolic domain in the chimeric protein. On the contrary, the absence of the TMD in $MobB_{CloDF13}$ does not affect the secondary structure of its cytosolic domain. Subcellular localization studies showed that T4CPs have a unipolar or bipolar location, which is enhanced by the presence of the remaining proteins of the conjugative system. Unlike what has been described for $TrwB_{R388}$, the TMD is not an essential element for the polar location of $MobB_{CloDF13}$. The main conclusion is that the characteristics described for the paradigmatic $TrwB_{R388}$ T4CP should not be ascribed to the whole T4CP family. Specifically, it has been proven that the mobilizable plasmid-related $MobB_{CloDF13}$ presents different characteristics regarding the role of its TMD. This work will contribute to better understand the T4CP family, a key element in bacterial conjugation, the main mechanism responsible for antibiotic resistance spread.

Keywords: coupling proteins, type IV secretion systems, bacterial conjugation, membrane proteins, antibiotic resistance spread

INTRODUCTION

Type IV coupling proteins (T4CPs) are essential elements in the conjugative type IV secretion systems (T4SSs) and are also key elements in many pathogenic T4SSs. The members of this family display a high sequence, length, and domain architecture heterogeneity being the Nucleotide-Binding Domain (NBD) the only conserved domain in all T4CPs. For this reason, they are classified according to the different domain architectures: (i) VirD4-type subfamily that are integral membrane proteins; (ii) TraG-J pairs, which are also integral membrane proteins but additionally present a physical and functional association with another membrane protein of the T4SS; (iii) T4CPs without Transmembrane Domain (TMD), which could or could not interact with other T4SS membrane proteins creating a VirD4-type complex, like the pair $TraJ_{pIP501}$ and $TraI_{pIP501}$; (iv) FtsK-like T4CPs; and (v) Archaeal T4CPs (Llosa and Alkorta, 2017).

The structural paradigm of this family is the T4CP of conjugative plasmid R388, $TrwB_{R388}$. It is a VirD4-type protein composed of a TMD at the N-terminus (consisting of two transmembrane α-helices connected through a small periplasmic loop) and a bulky globular cytosolic domain (CD). $TrwB_{R388}$ is the only full-length T4CP that has been successfully purified to date (Hormaeche et al., 2002; Redzej et al., 2017), while trials for purifying other membrane T4CPs have not rendered the sufficient amounts of high quality protein for performing *in vitro* assays (Chen et al., 2008). For this reason most of the *in vitro* studies of T4CPs have been achieved using deletion mutant proteins that lack the TMD (Schroder and Lanka, 2003; Tato et al., 2007; Larrea et al., 2017). In this regard, the TMD deletion mutant protein of $TrwB_{R388}$, TrwBΔN70, was resolved by X-ray crystallography, showing that the CD of $TrwB_{R388}$ contains an NBD with the Walker A and Walker B motifs and a small membrane-distal All-Alpha Domain (AAD) (Gomis-Rüth et al., 2001).

Comparative studies of the properties of $TrwB_{R388}$ and TrwBΔN70 showed significant differences regarding biological activity (such as *in vivo* function, *in vitro* nucleotide-binding properties, and *in vitro* ATPase activity), oligomerization pattern, subcellular location, and stability (Moncalián et al., 1999; Vecino et al., 2010, 2011; Hormaeche et al., 2002, 2004, 2006; Segura et al., 2013, 2014). For this reason it has been concluded that the TMD of $TrwB_{R388}$ accomplishes a role beyond the anchorage of the protein to the membrane, influencing the location, stability, and activity of this protein.

To delve into the role of the TMD in T4CPs two different strategies have been followed. On the one hand, we have constructed a chimeric protein named $TMD_{TraJ}CD_{TrwB}$ composed of the TMD of $TraJ_{pKM101}$, the phylogenetically closest T4CP to $TrwB_{R388}$ from the conjugative plasmid pKM101 (Paterson et al., 1999; Alvarez-Martinez and Christie, 2009) and the CD of $TrwB_{R388}$. This strategy has already been used for the study of components of T4SSs, showing interesting results (Bourg et al., 2009). Specifically, through a chimeric protein approach the function of the AAD of $VirD4_{At}$ from the T-plasmid of *Agrobacterium tumefaciens* (Whitaker et al., 2016) and of the N-terminal HUH domain of $TrwC_{R388}$ (Agúndez et al., 2018) have been analyzed. On the other hand, the T4CP from the mobilizable plasmid CloDF13, $MobB_{CloDF13}$ and its deletion protein lacking the TMD, MobBΔTMD, have been constructed and studied. It is an interesting system to characterize since it is part of the rare MOB_{C1} plasmid family, which are mobilizable plasmids that encode their T4CP (Smillie et al., 2010). Additionally, $MobB_{CloDF13}$ has been described as an atypical T4CP, due to its dual role in DNA transfer, since it acts as an accessory protein in CloDF13 relaxation process and also as a T4CP (Nuñez and de la Cruz, 2001).

We studied the functionality of these proteins in plasmid transfer, secondary structure, thermal stability, and subcellular localization. Our findings indicate that the TMD plays different roles in conjugative plasmid related and mobilizable plasmid related T4CPs. Specifically while the TMD could play a regulatory role in $TrwB_{R388}$ this cannot be inferred from the results about $MobB_{CloDF13}$.

MATERIALS AND METHODS

Materials

n-dodecyl β-D maltoside (DDM) was purchased from Anatrace (Santa Clara, CA, United States). Mouse anti-His (C-term) monoclonal antibody and Alexa Fluor goat anti-mouse antibody were purchased from Invitrogen (Carlsbad, CA, United States) and Molecular Probes (Eugene, OR, United States), respectively. All buffers employed in this work are shown in **Supplementary Table S1**.

Bacterial Strains and Bacterial Growth Conditions

E. coli DH5α strain was used as host for plasmid constructions. This strain also served as the donor for mating experiments by hosting the plasmids of interest. *E. coli* UB1637 served as the recipient for mating experiments. *E. coli* Lemo21 (DE3), *E. coli* BL21 (DE3), and *E. coli* BL21C41 (DE3) strains were used for protein production, purification, and *in vivo* localization.

E. coli strains were grown in LB medium and when necessary antibiotics were added at the following final concentrations: ampicillin (100 μg/mL), streptomycin (50 μg/mL), kanamycin (50 μg/mL), chloramphenicol (12.5–25 μg/mL), and thrimethoprim (10 μg/mL).

Plasmids

The plasmids and oligonucleotides used in this study are listed in **Tables 1**, **2**, respectively.

pSU4814 and pSU4833 plasmids were kindly provided by Fernando de la Cruz. pOPINE (Addgene plasmid # 26043; RRID: Addgene_26043)[1] and pOPINE-3C-eGFP (Addgene plasmid # 41125; RRID: Addgene_41125)[2] plasmids were a gift from Ray Owens. $pKM101Spc^r$_Δ*traJ* was kindly provided by Peter J. Christie. $pKM101Amp^r$_Δ*traJ* plasmid was obtained by cleavage

[1]http://n2t.net/addgene:26043
[2]http://n2t.net/addgene:41125

TABLE 1 | Plasmids employed in this work.

Plasmid	Description	Phenotype	References
pET24a (+)	Expression vector	Kan^R, C-terminal His tag	Novagen
R388	Natural plasmid	Tmp^R, TRA_W, IncW	Datta and Hedges, 1972
pKM101	Natural plasmid	Amp^R, TRA_N, IncN	Langer and Walker, 1981
pKM101Δ*traJ*	pKM101:Δ*traJ*	Amp^R, TRA_N, IncN, $TraJ^-$	This work
pKM101Spc^r_Δ*traJ*	pKM101:Δ*traJ* and Spt^R resistance cassette	Spt^R, TRA_N, IncN, $TraJ^-$	Whitaker et al., 2016
pOPINE	Expression vector	Amp^R, C-terminal His tag	Berrow et al., 2007
pOPINE-3C-eGFP	Expression vector	Amp^R	OPPF-UK (Addgene)
pOPINE-3C-eGFP-*mobB*	pOPINE-3C-eGFP:*mobB*	pOPINE-3C-eGFP:$MobB_{CloDF13}$GFP expression under T7 promoter	This work
pOPINE-3C-eGFP-*MobB*Δ*TMD*	pOPINE-3C-eGFP:*MobB*Δ*TMD*	pOPINE-3C-eGFP:MobBΔTMDGFP expression under T7 promoter	This work
pOPINE-*mobB*	pOPINE:*mobB*	pOPINE:$MobB_{CloDF13}$ expression under T7 promoter	This work
pOPINE-*MobB*Δ*TMD*	pOPINE:*MobB*Δ*TMD*	pOPINE:MobBΔTMD expression under T7 promoter	This work
pSU1456	R388:Δ*trwB*	Su^R, Tmp^R, TRA_W, IncW, $TrwB^-$	Llosa et al., 1994
pSU4814	pSU19:$mob_{CloDF13}$	Chl^r, Rep (p15A), MOB (CloDF13)	Nuñez and de la Cruz, 2001
pSU4833	pSU4814:Δ*mobB*	Chl^R, Rep (p15A), $MobB^-$	Nuñez and de la Cruz, 2001
pUB9	pWaldo-GFPe:*trwB-GFP*	Kan^R, TrwB-TEV-GFP-H_8, expression under T7 promoter	Segura et al., 2014
pUBQ4	pET24a (+):*traJTMD-trwBCD*	Kan^R, $TMD_{TraJ}CD_{TrwB}$-H_6, expression under T7 promoter	This work
pUBQ4 (K142T)	pET24a (+):*traJTMD-trwBCD (K142T)*	Kan^R, $TMD_{TraJ}CD_{TrwB}$-H_6 containing the K142T mutation in the CD_{TrwB} domain, expression under T7 promoter	This work
pUBQ4GFP	pWaldo-GFPe:*traJTMD-trwBCD-GFP*	Kan^R, $TMD_{TraJ}CD_{TrwB}$-TEV-GFP-H_8, expression under T7 promoter	This work
pWaldo-GFPe	Expression vector	Kan^R, C-Terminal GFP and His tag	Waldo et al., 1999

TABLE 2 | Oligonucleotides used in this work.

Construct	Template	Protein		Sequence (5′ → 3′)	Cloning technology
pOPINE-*mobB*	pSU4814	$MobB_{CloDF13}$	F:	AGGAGATATACCATGTTTAATACGGATTCGCTTGCCTGGCAGTGG	In-fusion[a]
			R:	GTGATGGTGATGTTTGTACAGCCCCGCAAAATCATCATCACCCG	
pOPINE-*MobB*Δ*TMD*	pSU4814	MobBΔTMD	F:	AGGAGATATACCATGGACGAGGCGGTGAACGCGAAAC	In-fusion[a]
			R:	GTGATGGTGATGTTTGTACAGCCCCGCAAAATCATCATCAC	
pOPINE-3C-eGFP-*mobB*	pSU4814	$MobB_{CloDF13}$GFP	F:	AGGAGATATACCATGTTTAATACGGATTCGCTTGCCTGGCAGTGG	In-fusion[a]
			R:	CAGAACTTCCAGTTTGTACAGCCCCGCAAAATCATCATCACCCG	
pOPINE-3C-eGFP-*MobB*Δ*TMD*	pSU4814	MobBΔTMDGFP	F:	AGGAGATATACCATGGACGAGGCGGTGAACGCGAAAC	In-fusion[a]
			R:	CAGAACTTCCAGTTTGTACAGCCCCGCAAAATCATCATCAC	
pUBQ4 (K142T)	pUBQ4	$TMD_{TraJ}CD_{TrwB}$ (K142T)	F:	AAGCAACACCGATGTACCCGTACCAGTGGCA	Site-Directed mutagenesis[b]
			R:	GTGCCACTGGTACGGGTACATCGGTGTTGCTT	
pUBQ4GFP	pUBQ4	$TMD_{TraJ}CD_{TrwB}$GFP	F:	CCG**CTCGAG**ATGGACGATAGAGAAAGAGG	Restriction enzymes (*Xho*I and *Bam*HI)
			R:	CGC**GGATCC**GATAGTCCCCTCAAC	

[a]*Berrow et al. (2007).* [b]*QuikChange II Site-Directed Mutagenesis (Stratagene, San Diego, CA, United States).*

of the *spc cassette of* pKM101Spc^r_Δ*traJ* plasmid using *Eco*RI restriction enzyme.

Cloning of T4CPs

To construct the chimeric protein $TMD_{TraJ}CD_{TrwB}$ and the transmembrane deletion mutant protein of $MobB_{CloDF13}$, MobBΔTMD, the sequences of $TrwB_{R388}$, $TraJ_{pKM101}$, and $MobB_{CloDF13}$, were analyzed through bioinformatics tools. First, the different characteristics of the proteins, such as molecular weight and isoelectric point, were analyzed using *ProtParam*.[3] Second, the topology of the membrane proteins was studied using *Topcons*[4] (Tsirigos et al., 2015).

[3]https://web.expasy.org/protparam/

[4]http://topcons.cbr.su.se/

Then different constructions were achieved as follows:

$TMD_{TraJ}CD_{TrwB}$ chimeric protein consists of amino acids M_1-D_{76} from $TraJ_{pKM101}$ followed by amino acids L_{71}-I_{507} from $TrwB_{R388}$. The *tmd*$_{TraJ}$*cd*$_{TrwB}$ sequence was synthesized *de novo* and inserted it into pET24a (+) plasmid vector using *Nde*I and *Xho*I restriction sites, rendering pUBQ4 plasmid to produce the chimeric protein. This construction was made by TOP Gene Technologies, Inc. (Saint-Lauren, QC-Canada). To study the role of the conserved lysine of the Walker A motif (K_{142}), this residue was substituted with a threonine using the QuikChange II Site-Directed Mutagenesis Kit from Stratagene (San Diego, CA, United States) to obtain the $TMD_{TraJ}CD_{TrwB}$ (K142T) protein. Additionally, to clone the *tmd*$_{TraJ}$*cd*$_{TrwB}$-*eGFP* gene *tmd*$_{TraJ}$*cd*$_{TrwB}$ sequence was inserted into the pWaldo-GFPe plasmid vector using *Xho*I and *Bam*HI restriction sites (Segura et al., 2014).

To clone $MobB_{CloDF13}$-related proteins, the cloning of *mobB, MobBΔTMD, mob-eGFP*, and *MobBΔTMD-eGFP* genes was performed in the Oxford Protein Production Facility (OPPF-UK) using the High-throughput protocol described by Bird (2011). Specifically, MobBΔTMD soluble mutant protein consists of amino acids D_{185}-Y_{653} of $MobB_{CloDF13}$ obtained after deletion of amino acids M_1-A_{184} from wild type $MobB_{CloDF13}$.

All the oligonucleotides employed in the aforementioned cloning experiments are specified in **Table 2**.

Overexpression and Purification

The same purification protocol was followed for $TMD_{TraJ}CD_{TrwB}$ and $MobB_{CloDF13}$ proteins. Briefly, *E. coli* BL21C41 (DE3) cells freshly transformed with plasmids pUBQ4 for $TMD_{TraJ}CD_{TrwB}$ and pOPINE-*mobB* for $MobB_{CloDF13}$ were grown overnight in LB (8 flasks of 10 mL) supplemented with the corresponding antibiotics at 37°C with continuous shaking. Then, cells were diluted 1:50 (v/v) with fresh LB supplemented with antibiotics (8 flasks of 500 mL) and were grown at 37°C with continuous shaking until an OD_{600} of 0.4–0.5 was achieved. Next, overexpression was induced by the addition of 1 mM isopropyl α-D-thiogalactopyranoside (IPTG) and cells were grown with continuous shaking at 25°C overnight. Cells were harvested by centrifugation at 8,000 *g* for 15 min at 4°C. The pellet was suspended in 80 mL of *Cell buffer*, frozen with liquid N_2 and stored at −80°C.

For purification, cells were thawed at 37°C and 0.02 mg/mL DNase I, 1 mM dithiothreitol (DTT), 0.07% (w/v) lysozyme, 1 mM $MgCl_2$, 1 mM phenylmethanesulfonyl fluoride (PMSF) and two tablets of cOmplete™ EDTA-free Protease Inhibitor Cocktail from Sigma-Aldrich (San Luis, MO, United States) were added. From this point onward, the whole process was performed at 4°C to avoid aggregation of the proteins. After 45 min of incubation with agitation, cells were disrupted by sonication and centrifuged at 8,000 *g* for 15 min to remove non-lysed cell. The supernatant, containing the broken cells, was centrifuged at 138,000 *g* for 45 min to pellet the membrane fraction which was subsequently carefully resuspended in 30 mL of *Cell buffer*. Then DDM and NaCl were added to final concentrations of 19.6 mM and 600 mM, respectively, and the volume was adjusted to 40 mL. The mixture was incubated for 90 min with continuous stirring and then centrifuged at 138,000 *g* for 1 h.

The supernatant containing the protein to be purified was mixed 1:1 (v/v) with *MP1 buffer* to decrease the concentration of DDM and NaCl to 8.3 mM and 300 mM, respectively. Then, the sample was supplemented with 50 mM imidazole and loaded onto a 5 mL *HisTrap™ FF* (GE Life Sciences; Marlborough, MA, United States) column previously equilibrated with *MP2 buffer*. To increase the binding of the protein, the sample was left recirculating overnight. Next, it was connected to an *ÄKTA-FPLC system* and it was washed with 50 mL of *MP2 buffer* at a flow rate of 2 mL/min until the absorbance at 280 nm reached the baseline. Bound proteins were eluted with *MP3 buffer*, at a flow rate of 1.5 mL/min and fractions of 1 mL were collected. Obtained samples were analyzed by SDS-PAGE and the ones containing each target protein were pulled-down and concentrated using a centrifugal filter with a MWCO of 100 kDa. Then, 5 mL of the resulting samples were separately loaded onto a *Superdex 200 HR 16/60* column and the size-exclusion chromatography (SEC) was performed in *MP purification buffer* at 0.5 mL/min. The fractions corresponding to each target protein were pulled-down and concentrated as explained before. Glycerol was added to a final concentration of 20% (v/v) and protein concentration was determined by measuring absorption at 280 nm. Finally, aliquots were stored at −80°C.

When $TMD_{TraJ}CD_{TrwB}$ was purified with the aim of performing infrared spectroscopy (IR) assays the DDM and NaCl concentrations of the MP purification buffer were changed to the ones described in the previously published purification protocols of $TrwB_{R388}$ and TrwBΔN50 (i.e., 0.2 mM DDM and 200 mM NaCl instead of 0.6 mM DDM and 300 mM NaCl) (Vecino et al., 2011, 2012).

For MobBΔTMD purification, *E. coli* Lemo21 (DE3) cells freshly transformed with pOPINE-*MobBΔTMD* plasmid were grown in 4 L of LB supplemented with ampicillin at 37°C with continuous shaking until an OD_{600} of 0.5–0.6 was reached. Expression was induced with 1 mM IPTG and performed for 20 h at 25°C. Cells were harvested and stored as explained previously.

For protein purification, the cell suspension was thawed and the lysis protocol previously described for $TMD_{TraJ}CD_{TrwB}$ and $MobB_{CloDF13}$ was followed. Then, the sample was centrifuged at 138,000 *g* for 45 min to pellet the membrane fraction and the inclusion bodies. The supernatant with the soluble proteins was supplemented with 50 mM imidazole and loaded onto a 5 mL *HisTrap™ FF* (GE Life Sciences; Marlborough, MA, United States) column, previously equilibrated with *MobBΔTMD1 buffer*. Affinity chromatography was performed as previously described for $MobB_{CloDF13}$, but using *MobBΔTMD1 buffer* for washing and *MobBΔTMD2 buffer* for elution. Fractions containing the target protein were pulled-down and concentrated using a *Centricon YM-30* to a final volume of 600 μL. The resulting sample was centrifuged to remove aggregates and loaded onto a *Superdex 200 HR 10/30* column. The SEC was performed in *Cell buffer* at 0.3 mL/min and 0.5 mL fractions were collected. Fractions that contained the protein of interest were pulled-down and the sample was centrifuged to discard the aggregates.

Glycerol was added to a final concentration of 20% (v/v) and aliquots were made.

Mating Assays

Mating assays were performed as described by Llosa et al. (2003) with small modifications. Briefly, donors (*E. coli* DH5α co-transformed with the appropriate plasmids) and recipient cells (*E. coli* UB1637) were grown in LB supplemented with the corresponding antibiotics overnight at 37°C. For each mating assay 100 μL of the donor and the recipient cells were mixed, centrifuged, resuspended in 50 μL LB and placed onto a GS Millipore filter (0.22 μm pore size) settled on a pre-warmed LB-agar plate. After 1 h incubation at 37°C bacteria were washed from the filters in 2 mL LB by shaking at 450 rpm for 20 min and vortexing for 30 s. Then, 100 μL of the appropriate dilutions were plated on selective media for donors and transconjugants. The plates were incubated overnight at 37°C and the colonies were counted, normalizing the conjugation frequency as the number of transconjugants per donor cell.

Infrared Spectroscopy

To accomplish IR studies of different T4CPs and their variants (i.e., $TMD_{TraJ}CD_{TrwB}$, $MobB_{CloDF13}$, and MobBΔTMD) purification of each protein was carried out as previously described. The H-D exchange protocol was performed at 4°C and adapted for each protein. Briefly, $TMD_{TraJ}CD_{TrwB}$ was diluted with the IR buffer, dialyzed against the same buffer using a *D-Tube™ Dialyzer Midi* (MWCO 3.5 kDa) (Merck; Darmstadt, Germany), diluted again in IR buffer and concentrated using an Amicon Ultra-0.5 mL centrifugal filter (MWCO: 100 kDa). A similar process was followed for $MobB_{CloDF13}$ except for the dialysis step. Regarding MobBΔTMD, sample was diluted, dialyzed and concentrated as described for $TMD_{TraJ}CD_{TrwB}$ but using in MobBΔTMD IR buffer. Final protein samples were always above 1 mg/mL protein concentration.

Infrared spectra were recorded in a Thermo Nicolet Nexus 5700 (Thermo Fisher Scientific; Waltham, MA, United States) spectrometer equipped with a liquid nitrogen-refrigerated mercury-cadmium-telluride detector using a Peltier-based temperature controller (TempComp™, BioTools; Wauconda, IL, United States), and a 25 μm optical path. Typically 370 scans for each, background and sample, were collected at 2 cm^{-1} resolution and averaged after each minute. Spectra were collected with OMNIC software (Nicolet) and data processing was performed with OMNIC and SpectraCalc, following previously resolved methods (Arrondo et al., 1993; Arrondo and Goñi, 1999).

The information about the secondary structure and about the thermal denaturation of T4CPs and their derivatives was obtained by IR spectroscopy through analysis of the infrared amide I band that corresponds mainly to the C = O stretching vibrations of the peptide bonds and which is located between 1700 and 1600 cm^{-1} region of the IR spectrum. Amide I band is conformationally sensitive and can be used to monitor the protein secondary structure composition and changes induced by thermal denaturation. Secondary structure studies were made at 20°C and band decomposition of the amide I was performed as previously reported (Vecino et al., 2011, 2012). For thermal stability studies samples were heated from 20 to 80°C at a rate of 1°C/min.

Subcellular Location of T4CPs

Subcellular location of different T4CPs and their variants (i.e., $TrwB_{R388}$, $TMD_{TraJ}CD_{TrwB}$, $MobB_{CloDF13}$, and MobBΔTMD) was achieved by confocal fluorescence microscopy. To do so, two different approaches were used: (i) eGFP-labeling (Cormack et al., 1996) and (ii) immunofluorescence. Since the eGFP moiety only emits fluorescence when properly folded (Drew et al., 2005), the eGFP based approach allowed visualizing only properly folded proteins.

Prior to localization assays, the *in vivo* activity of the T4CP-eGFP fusion-proteins was proved by mating assays as previously described (**Supplementary Table S2**). Afterward, T4CP-eGFP fusion-proteins were expressed in *E. coli* BL21C41 (DE3) strain, except for MobBΔTMD that was expressed in BL21 (DE3) strain. To do so cells were transformed with pUBQ4, by induction with 1 mM IPTG at OD_{600} 0.4–0.5 for the membrane proteins and OD_{600} 0.5–0.6 for MobBΔTMD. Protein expression was performed for 4 and 20 h at 25°C. Additionally, the subcellular location of $TrwB_{R388}$-related proteins was determined in the presence of pSU1456 plasmid, which codes for all the R388 conjugative proteins except $TrwB_{R388}$. Similarly, $MobB_{CloDF13}$ was also expressed in the presence of plasmid pSU1456 and to mimic the *in vivo* transfer of CloDF13, its location was additionally studied in the presence of plasmids pSU1456 (R388 plasmid that lacks $TrwB_{R388}$ protein) and pSU4833 (CloDF13 plasmid that contains its mobilization region except for $MobB_{CloDF13}$ protein). Sample handling was performed as described by Segura et al. (2014). The images were acquired in a Leica TCS SP5 confocal fluorescence microscope, with a 60× oil immersion objective. Sample excitation was performed with 488 nm wavelength, while fluorescence emission was measured between 505 and 525 nm. The images were analyzed using Huygens and ImageJ softwares. To ease the counting process and better distinguish the different locations the images were treated with the preset ICE filter of ImageJ software; in this manner five different locations for the T4CP-eGFP fusion-proteins were described (**Supplementary Figure S1**).

For immunofluorescence assays protein expression was performed as with the eGFP fusion-proteins. Sample collection and handling was performed as described by Segura et al. (2014). Cells were immunostained with mouse anti-His (C-term) monoclonal antibody as primary antibody, and Alexa Fluor goat anti-mouse as secondary antibody. Image acquisition was performed in an Olympus FluoviewTM 500 confocal fluorescence microscope at the "Analytical and high-resolution microscopy in biomedicine" facility (SGIker, UPV/EHU).

RESULTS

Bacterial conjugation is one of the main processes responsible for the horizontal dissemination of antibiotic resistance genes

among bacteria. One of the essential proteins in this process is the T4CP, which is ubiquitous in all conjugative systems. Despite its importance, the only widely studied T4CP is $TrwB_{R388}$. Given its central role in bacterial conjugation, detailed knowledge of the T4CP family could contribute to the development of new strategies against the spread of antibiotic resistance among bacteria.

Previously published papers have highlighted the role of the TMD on different characteristics of $TrwB_{R388}$, such as plasmid conjugation (Moncalián et al., 1999), subcellular localization (Segura et al., 2014), nucleotide-binding (Hormaeche et al., 2006), hexamerization (Hormaeche et al., 2002; Matilla et al., 2010), protein stability (Hormaeche et al., 2004), interaction with other proteins of the T4SS of R388 (Segura et al., 2013), and ATP hydrolase activity (Tato et al., 2005, 2007). From all these studies it was inferred that the TMD of $TrwB_{R388}$ has a role beyond the mere anchorage in the membrane.

To gain more knowledge about different T4CPs, and in particular about the role of their TMD in T4CP features, in this work a $TrwB_{R388}$ chimeric protein that combines its CD with the TMD of its phylogenetically closest T4CP, $TraJ_{pKM101}$, has been studied. Also one of the few T4CPs of mobilizable plasmids, $MobB_{CloDF13}$, and its TMD deletion mutant protein, MobBΔTMD, have been studied. Plasmid transfer, secondary structure, thermal stability, and subcellular location studies have been carried out to shed light on the functioning of this protein family and in the role of their TMD.

Cloning of Soluble Mutant and Chimeric Proteins

The membrane protein topologies obtained after the bioinformatic analysis performed with Topcons software of $TrwB_{R388}$, $TraJ_{pKM101}$, and $MobB_{CloDF13}$ are shown in **Figure 1A**. $TrwB_{R388}$ and $TraJ_{pKM101}$ have similar size and organization of their TMDs that consist of about 70 residues and contain two α-helixes connected by a small periplasmic loop. In contrast, the TMD of $MobB_{CloDF13}$ is larger (about 150 amino acids) and is organized into three α-helixes. This information was used to design the chimeric and mutant proteins studied in this work (**Figure 1B**). The chimeric protein $TMD_{TraJ}CD_{TrwB}$ was made by combination of the TMD of the T4CP $TraJ_{pKM101}$ and the CD of $TrwB_{R388}$. In addition, in this work the T4CP of the mobilizable plasmid CloDF13, $MobB_{CloDF13}$, and its TMD-deletion protein MobBΔTMD, were constructed (**Figure 1B**). The theoretical molecular weights of these proteins, necessary for their purification process, were calculated using *ProtParam*[3] bioinformatic tool. The estimated molecular weights were 58.28, 73.95, and 53.13 kDa for $TMD_{TraJ}CD_{TrwB}$, $MobB_{CloDF13}$, and MobBΔTMD, respectively. Finally, the eGFP fusion-proteins (i.e., $TMD_{TraJ}CD_{TrwB}$-eGFP, $MobB_{CloDF13}$-eGFP, and MobBΔTMD-eGFP) were constructed and since they emitted a fluorescent signal, it was deduced that they were correctly folded (Drew et al., 2006).

Functionality and Dominance Experiments

Through mating assays two different properties of $TMD_{TraJ}CD_{TrwB}$ were analyzed: (i) its capacity to complement the conjugative process in the absence of another T4CP (*functionality studies*) and (ii) its effect on each native conjugative system (R388 or pKM101 plasmids), being the corresponding T4CP present ($TrwB_{R388}$ or $TraJ_{pKM101}$, respectively) (*dominance studies*). Results obtained in mating assays are summarized in **Table 3**.

Our results showed that $TMD_{TraJ}CD_{TrwB}$ efficiently complemented the Δ*trwB* mutation in R388 transfer but to a lower rate than native R388 (0.21 vs. 1.82 10^{-4} transconjugants per donor, respectively). On the contrary, $TMD_{TraJ}CD_{TrwB}$ was unable to complement the Δ*traJ* mutation in pKM101 transfer (**Table 3**). These results are in agreement with the necessary specific interactions between the CD of the T4CP and its cognate relaxase for transfer to happen as reported previously (Cabezón et al., 1997; Hamilton et al., 2000; Llosa et al., 2003).

It has been reported that mutation of the conserved lysine in the Walker A motif rendered a transfer deficient mutant protein TrwBK136T (Hormaeche et al., 2006). Similarly, an equivalent mutant of the soluble protein TrwBΔN70, TrwBΔN70 (K136T), lacked ATPase activity (Moncalián et al., 1999), underlying the essential role of this amino acid in the activity of $TrwB_{R388}$. Here we studied the effect of the equivalent point mutation in the Walker A domain, $TMD_{TraJ}CD_{TrwB}$ (K142T), on the transfer capacity of chimeric protein. As expected, $TMD_{TraJ}CD_{TrwB}$ (K142T) was unable to complement the Δ*trwB* mutation in R388 plasmid transfer (**Table 3**). This is in agreement with the crucial role of the K residue as it has been reported with homologous mutants in other T4CPs (Moncalián et al., 1999; Kumar and Das, 2002; Gunton et al., 2005).

Next, to accomplish dominance assays, the transfer frequencies of plasmids R388 or pKM101 in the presence of the cognate T4CP and the chimeric protein were measured. It was observed that $TMD_{TraJ}CD_{TrwB}$ reduced the transfer frequency of R388 or pKM101 by an order of magnitude (**Table 3**).

Mobilization Experiments

CloDF13 mobilization experiments were achieved to know whether the TMD of $MobB_{CloDF13}$ was essential for the mobilization of the plasmid as happens with $TrwB_{R388}$ or not as it has been described for $TcpA_{pcW3}$, whose TMD-less mutant can perform conjugation although at a lower frequency (Parsons et al., 2007). First of all, the transfer frequency of the mobilizable region of CloDF13 (plasmid pSU4814) mediated by the T4SS of R388 was analyzed. Afterwards, the complementation experiments were performed in the presence of both pSU4833 (the mobilizable region of CloDF13 without functional $MobB_{CloDF13}$) and pSU1456 plasmid (encoding for R388 conjugative system except for $TrwB_{R388}$). Results obtained in mobilization assays are summarized in **Table 4**. It was observed that cloned $MobB_{CloDF13}$ was functional when the T4SS of R388 was used, in agreement with what has been previously published for $TrwB_{R388}$ (Hormaeche et al., 2006). Similarly, the deletion

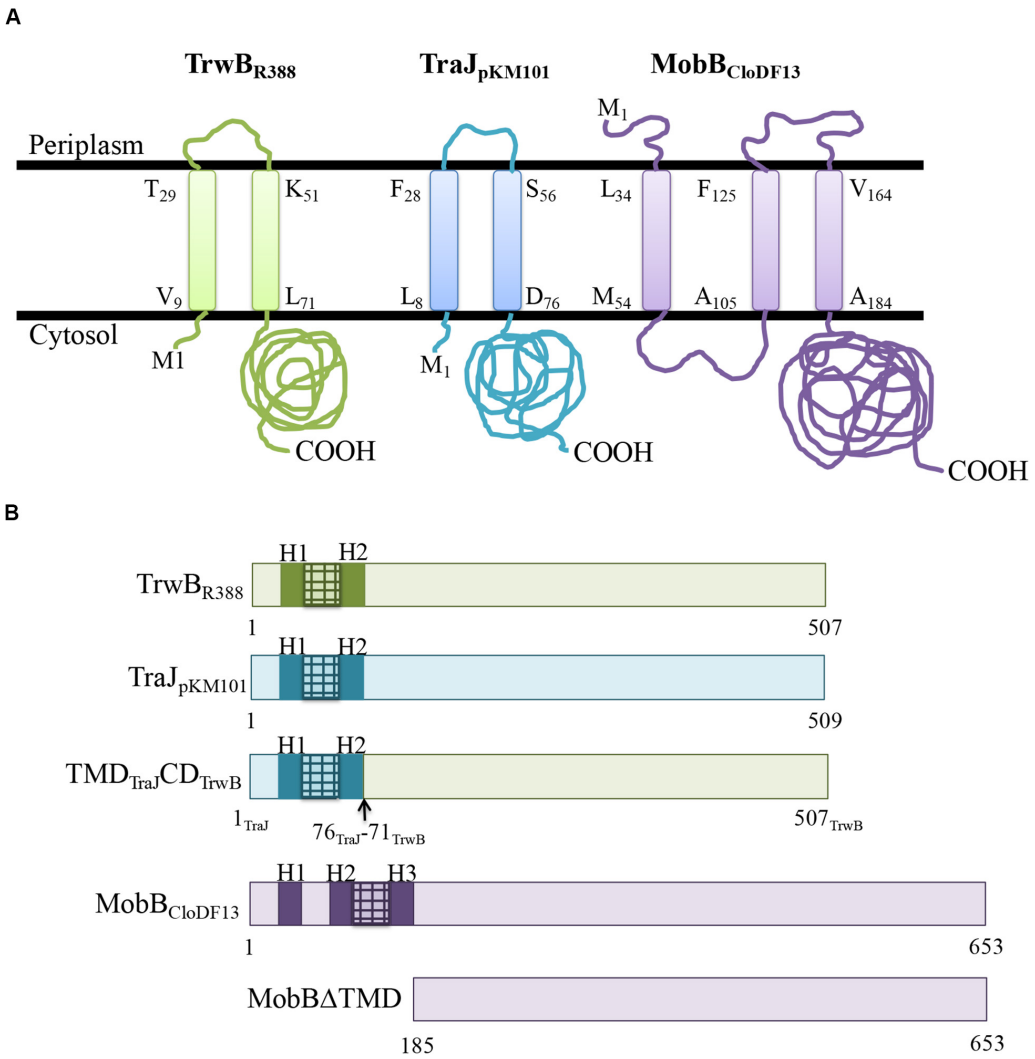


FIGURE 1 | (A) Predicted membrane topology of TrwB$_{R388}$, TraJ$_{pKM101}$, and MobB$_{CloDF13}$ proteins. Membrane topology of the different T4CPs was predicted using Topcons software. The black lines represent the inner bacterial membrane. M1, amino-terminus; COOH, carboxy-terminus. The first and last residues of each transmembrane helix are shown indicating their position in the sequence. Proteins from R388, pKM101, and CloDF13 plasmids are shown in green, blue, and purple, respectively. **(B)** Schematic representation of the different T4CPs and their variants used in the present study. Proteins from R388, pKM101, and CloDF13 plasmids are shown in green, blue and purple, respectively. The transmenbrane α-helices (H) and the small periplasmic loops connecting α-helices are indicated in dark boxes and stripped boxes, respectively.

of the TMD, MobBΔTMD mutant, rendered a non-functional phenotype as it occurs with other T4CP mutants that lack the TMD, such as TrwBΔN70 and PcfCΔN103 (Moncalián et al., 1999; Chen et al., 2008).

Secondary Structure of T4CPs and Their Variants

Analysis of the secondary structure components of TMD$_{TraJ}$CD$_{TrwB}$, MobB$_{CloDF13}$, and MobBΔTMD were performed by IR spectroscopy through analysis of the IR amide I band.

The secondary structure of TMD$_{TraJ}$CD$_{TrwB}$ was compared to those of the native TrwB$_{R388}$ and its mutants TrwBΔN50 and TrwBΔN70 (Vecino et al., 2012). **Figure 2A** shows the original spectra and the curve-fitting decomposition corresponding to TMD$_{TraJ}$CD$_{TrwB}$ purified in the presence of detergent. Band position, percentage area, and structure assignation corresponding to the deconvolved spectrum of the amide I region are summarized in **Table 5** together with those previously reported of TrwB$_{R388}$, TrwBΔN50, and TrwBΔN70 (Vecino et al., 2012).

The spectrum of TMD$_{TraJ}$CD$_{TrwB}$ purified in the presence of detergent exhibited four bands related to protein structure at 1671, 1654, 1640, and 1626 cm^{-1} (**Table 5**). Interpretation of the results must be done taking into account that band assignation is not always a straightforward process since its position can be altered by the environment (Arrondo and Goñi, 1999).

TABLE 3 | Conjugation and dominance experiments with $TMD_{TraJ}CD_{TrwB}$. Transfer frequencies of plasmids pSU1456 and pKM101 Δ*traJ* complemented with $TMD_{TraJ}CD_{TrwB}$ protein have been studied.

Plasmids in donors	T4CP	Transfer frequency
R388 (+)	Wild type $TrwB_{R388}$	[a]$2.1\ 10^{-1}$
R388Δ*trwB* (pSU1456) (–)	Ø	$< 10^{-8}$
pKM101 (+)	Wild type $TraJ_{pKM101}$	[b]$2.2\ 10^{-1}$
pKM101Δ*traJ* (–)	Ø	$< 10^{-8}$
R388Δ*trwB* (pSU1456) pUBQ4	$TMD_{TraJ}CD_{TrwB}$	$1.82\ 10^{-4}$
pKM101Δ*traJ* pUBQ4	$TMD_{TraJ}CD_{TrwB}$	$< 10^{-8}$
R388Δ*trwB* (pSU1456) pUBQ4 (K142T)	$TMD_{TraJ}CD_{TrwB}$ (K142T)	$< 10^{-8}$
R388 pUBQ4	Wild type $TrwB_{R388}$ $TMD_{TraJ}CD_{TrwB}$	$1.13\ 10^{-2}$
pKM101 pUBQ4	Wild type $TraJ_{pKM101}$ $TMD_{TraJ}CD_{TrwB}$	$1.10\ 10^{-2}$

Additionally, the effect of the chimeric protein in the conjugative process of R388 and pKM101 has been analyzed (dominance assays, shaded in gray). E. coli DH5α and UB1637 strains were used as donor and recipient cells, respectively. Transfer frequencies were normalized to the number of transconjugants per donor and are the mean value of at least five independent experiments. (+) Positive control; (–) negative control; Ø: no T4CP. [a]Data from Segura et al. (2013). [b]Data from Whitaker et al. (2016).

TABLE 4 | Mobilization experiments.

Plasmids in donors	T4CP	Transfer frequency
R388Δ*trwB* (pSU1456) pSU4814 (+)	Wild type $MobB_{CloDF13}$	$2.42\ 10^{-2}$
R388Δ*trwB* (pSU1456) $MOB_{CloDF13}$Δ*mobB* (pSU4833) (–)	Ø	$<10^{-8}$
R388Δ*trwB* (pSU1456) $MOB_{CloDF13}$Δ*mobB* (pSU4833) pOPINE-*mobB*	$MobB_{CloDF13}$	1.75×10^{-2}
R388Δ*trwB* (pSU1456) pSU4833	Ø	$<10^{-8}$

Transfer frequency of pSU4833 plasmid complemented with $MobB_{CloDF13}$ or MobBΔTMD mediated by the secretion channel of R388 conjugative system has been studied. E. coli DH5α and UB1637 strains were used as donor and recipient cells, respectively. Transfer frequencies were normalized to the number of transconjugants per donor and are the mean value of at least five independent experiments. (+) Positive control; (–) Negative control; Ø: no T4CP.

The component at 1654 cm^{-1} was assigned to α-helix, the bands at 1626 and 1671 cm^{-1} were associated with the low and high-frequency vibrations of β-sheet, respectively, although it should be noted that the later is also assigned to β-turns. And the band at 1640 cm^{-1} was assigned to flexible, non-periodic elements.

When these results were compared to previously reported ones (i.e., $TrwB_{R388}$ TrwBΔN50, and TrwBΔN70) (Vecino et al., 2012) it can be observed that the proportion of the α-helix (35%) is lower than the one observed in the native protein $TrwB_{R388}$ purified in detergent (41%) and higher than the one shown in the deletion mutant proteins (26%). This result can be directly associated to the presence of a TMD both in $TrwB_{R388}$ and $TMD_{TraJ}CD_{TrwB}$, even if in the later belongs to another T4CP such as $TraJ_{pMK101}$. Regarding bands associated to β-sheet and β-turns, it is remarkable the absence of a band around 1661–1665 cm^{-1} in $TMD_{TraJ}CD_{TrwB}$, as it was observed in $TrwB_{R388}$, TrwBΔN50, and TrwBΔN70. Nevertheless the total proportion of the different bands associated to β-sheet and β-turns of $TMD_{TraJ}CD_{TrwB}$ (31%) is similar to the proportion seen in the soluble mutant proteins and significantly lower than that of $TrwB_{R388}$ (59%). Finally, a sizeable proportion (35%) of the structure of $TMD_{TraJ}CD_{TrwB}$ gave off a signal centered at 1640 cm^{-1} (assigned to flexible, non-periodic elements) as seen in the deletion mutant proteins TrwBΔN50 and TrwBΔN70 but not in $TrwB_{R388}$ (Vecino et al., 2012). Previous studies about TrwBΔN70 and TrwBΔN50 showed that this band at 1640 cm^{-1} also had a β-sheet component (Vecino et al., 2012). And it was published that at higher temperatures the band at 1640 cm^{-1} split showing a β-sheet related band, which would not happen if the band was purely composed of unordered structures (Andraka et al., 2017). To elucidate if this also happened in $TMD_{TraJ}CD_{TrwB}$, IRS experiments at different temperatures were performed (20, 40, 60, and 80°C). With the increase of temperature, the 1640 cm^{-1} band shifts to 1645 cm^{-1} and that at 1626 cm^{-1} increases both its contribution and width, suffering also a shift to higher wavenumbers. This behavior indicates that although those two not resolved bands included in that at 1640 cm^{-1} do not directly split into two bands, there is a transfer of the β-sheet contribution to the 1626 cm^{-1} only β-sheet band, confirming that the 1640 cm^{-1} band in the chimeric protein was composed of both unordered and β-sheet elements.

With the aim of further studying if the deletion of the TMD has the same effect in all T4CPs, the same study was carried out with $MobB_{CloDF13}$ and its deletion mutant MobBΔTMD to be compared with $TrwB_{R388}$ and its variants. **Figure 2A** shows the original spectra and their curve-fitting decomposition corresponding to $MobB_{CloDF13}$ purified in the presence of detergent and its deletion mutant MobBΔTMD. In this case, both proteins exhibited four main bands. Band position, percentage area, and structure assignation corresponding to the deconvolved spectrum of the amide I region are summarized in **Table 5**. It can be observed that the component at 1653 and 1652 cm^{-1} was assigned to α-helix in $MobB_{CloDF13}$ and MobBΔTMD, respectively. But it should be pointed out that the proportion of the α-helix in MobBΔTMD (29%) was lower than in native protein (35%), as expected taking into account the three transmembrane α-helices postulated for the native protein (**Figure 1A**). The bands at 1666 or 1665 cm^{-1} were associated with β-turns which were also observed in the proteins studied in Vecino et al. (2012), but not in $TMD_{TraJ}CD_{TrwB}$. Nevertheless, proportions were slightly different (13 and 22% for $MobB_{CloDF13}$ and MobBΔTMD, respectively). On the contrary, the signal centered at 1677 or 1682 cm^{-1} assigned to low-frequency vibrations of β-sheet and β-turns showed similar proportions in $MobB_{CloDF13}$, MobBΔTMD (7 and 3%, respectively) and in the proteins studied in Vecino et al. (2012) but not in $TMD_{TraJ}CD_{TrwB}$ where the proportion of this component was 14%. Finally, as observed in $TMD_{TraJ}CD_{TrwB}$, TrwBΔN50 and TrwBΔN70, both $MobB_{CloDF13}$ and MobBΔTMD had

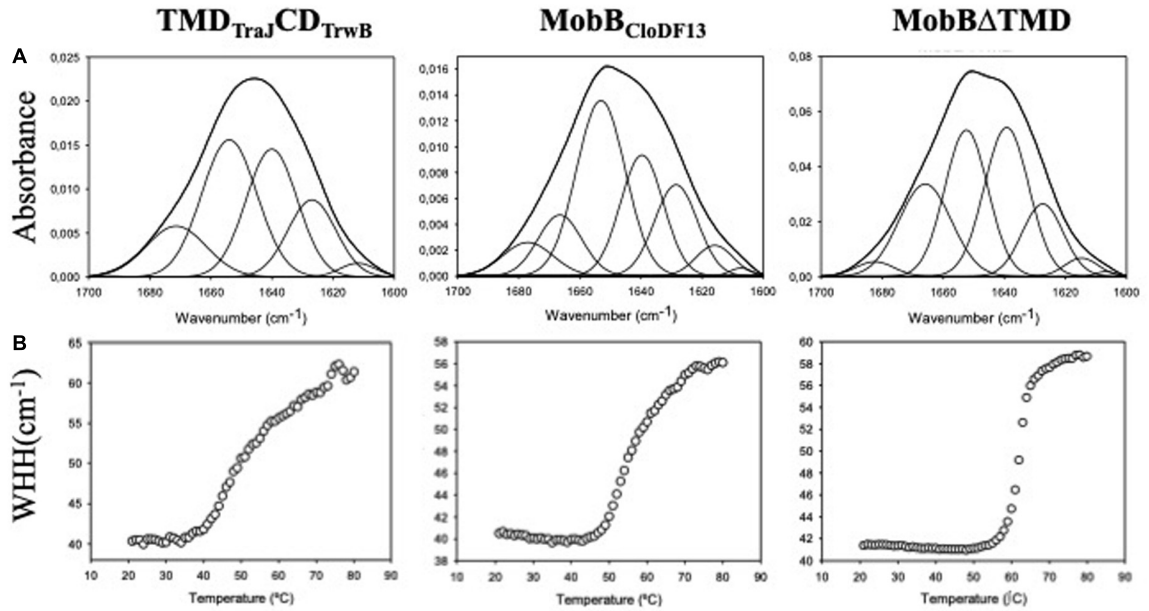


FIGURE 2 | (A) Amide I region of the infrared spectra of $TMD_{TraJ}CD_{TrwB}$, $MobB_{CloDF13}$, and MobBΔTMD. Proteins were purified, dialyzed against the corresponding buffer in D_2O and analyzed by IR spectroscopy as explained in "Materials and Methods" section. Obtained spectra were curve-fitted to show the different secondary structure components as detailed in **Table 5**. **(B)** Thermal denaturation of $TMD_{TraJ}CD_{TrwB}$, $MobB_{CloDF13}$, and MobBΔTMD as seen by IR spectroscopy. The widths at half-height (WHH) of the amide I bands are plotted as a function of temperature (°C). Thermal denaturation is marked by an abrupt increase in bandwidth. Mid-point denaturation temperature (Tm) values corresponding to each protein are detailed in **Table 5**.

TABLE 5 | Secondary structure components and mid-point denaturation temperatures (Tm) of $TrwB_{R388}$, $TMD_{TraJ}CD_{TrwB}$, TrwBΔN50, TrwBΔN70, $MobB_{CloDF13}$, and MobBΔTMD.

	$TrwB_{R388}$[a]		$TMD_{TraJ}CD_{TrwB}$		TrwBΔN50[a]		TrwBΔN70[a]		$MobB_{CloDF13}$		MobBΔTMD	
Assignment	**Position (cm^{-1})**	**Area (%)**	**Position (cm^{-1})**	**Area (%)**	**Position (cm^{-1})**	**Area (%)**	**Position (cm^{-1})**	**Area (%)**	**Position (cm^{-1})**	**Area (%)**	**Position (cm^{-1})**	**Area (%)**
β-T + β-S	1676	5	1671	14	1676	2	1675	8	1677	7	1682	3
β-T	1661	27	/	/	1665	11	1665	11	1666	13	1665	22
α-H	1647	41	1654	35	1654	27	1653	26	1653	35	1652	29
Unordered + β-S	/	/	1640	30	1641	37	1640	39	1639	29	1639	30
β-S	1631	27	1626	17	1628	17	1627	12	/	/	/	/
Tm (°C)	48		51		48		63[b]		56		62	

*Secondary structure components were obtained from band decomposition of the amide I band of IR spectra in D_2O medium. The spectra are shown in **Figure 2**. Only areas above 5% are shown. For details on assignments, see the text. [a]Data from Vecino et al. (2011, 2012). [b]Unpublished data from our group. β-T, β-turns; β-S, β-sheet; α-H, α-helix; Tm, mid-point denaturation temperature.*

a band at 1639 cm^{-1} assigned to unordered structures that represented 30% of the structure, a band that is missing in $TrwB_{R388}$.

Thermal Stability of T4CPs and Their Variants

The information regarding the denaturation of $TMD_{TraJ}CD_{TrwB}$, $MobB_{CloDF13}$, and MobBΔTMD was obtained through analysis of the amide I band (**Figure 2B**). Specifically, two bands appear at 1615–1620 cm^{-1} and 1680 cm^{-1} when the protein aggregates. The appearance of these bands allows monitoring the denaturation process of the protein and the calculation of the mid-point denaturation temperature (Tm) (**Table 5**). Data treatment and band decomposition of the original amide I have been described previously (Vecino et al., 2011, 2012).

The thermal stability of $TMD_{TraJ}CD_{TrwB}$ purified in detergent was compared to $TrwB_{R388}$, TrwBΔN50, and TrwBΔN70 (Vecino et al., 2012). As observed in **Figure 2B**, the thermal denaturation of $TMD_{TraJ}CD_{TrwB}$ started at 35°C showing a mid-point denaturation temperature of 51.1°C at the tested conditions. This result is similar to $TrwB_{R388}$ and to the TrwBΔN50 (Vecino et al., 2012; **Table 5**).

Thermal stability of $MobB_{CloDF13}$ and MobBΔTMD was also studied by IR spectroscopy. As depicted in **Figure 2B**, the denaturation of the native protein starts at 47°C, achieving its Tm

at 56°C, while the denaturation of MobBΔTMD starts at 55°C, achieving its Tm at 62°C (**Table 5**).

Subcellular Location

Studies about subcellular location of different T4CPs reported up to now have rendered ambiguous results (Kumar and Das, 2002; Gunton et al., 2005; Segura et al., 2014). To gain a deeper knowledge of this matter, specifically regarding the role of the TMD, we have studied the subcellular location of $TMD_{TraJ}CD_{TrwB}$, $MobB_{CloDF13}$, and MobBΔTMD under different experimental conditions. Subcellular location was analyzed by confocal fluorescence microscopy using eGFP-labeling and immunofluorescence techniques. Since eGFP proteins are only fluorescent when they are properly folded, their visualization ensures the analysis of functional proteins, excluding those that are denatured or included in inclusion bodies (Drew et al., 2006). In both studies similar results were obtained but eGFP-labeling rendered better quality images (**Supplementary Figure S2**).

To study the effect of expression times in the absence of the rest of T4SS proteins, the subcellular location of each protein was visualized after induction with 1 mM IPTG for 4 or 20 h. Different patterns were observed in the location of each protein at different expression times (**Figure 3** and **Table 6**). After 4 h of expression, $TrwB_{R388}$ and $TMD_{TraJ}CD_{TrwB}$ were mainly located along the whole cell membrane and switched to a single-pole after 20 h, being this change less pronounced in the case of the chimeric protein. $MobB_{CloDF13}$ was predominantly located at both poles both at 4 and 20 h, showing a little increase in one pole location after 20 h (**Figure 4** and **Table 6**). On the contrary, most of the cells (95%) showed MobBΔTMD located on a single pole in the cytosol at both tested times.

Since the interaction with other conjugative proteins of the T4SS could modify the subcellular location pattern (Segura et al., 2014), the proteins that were active *in vivo* (i.e., $TrwB_{R388}$, $TMD_{TraJ}CD_{TrwB}$, and $MobB_{CloDF13}$) were also observed in the presence of a T4SS lacking a functional T4CP that could interfere with the studied eGFP variant. On the one hand, $TrwB_{R388}$ and $TMD_{TraJ}CD_{TrwB}$ were analyzed in the presence of the remaining conjugative proteins of R388 (i.e., in the presence of plasmid pSU1456 that lacks functional $TrwB_{R388}$ but contains the remaining conjugative proteins). On the other hand, since CloDF13 needs the T4SS of a co-resident conjugative plasmid to be mobilized, the subcellular location of $MobB_{CloDF13}$ was studied in the presence of R388 lacking functional $TrwB_{R388}$ (plasmid pSU1456) and also in the presence of both, R388 lacking functional $TrwB_{R388}$ (plasmid pSU1456) and the mobility region of CloDF13 lacking functional $MobB_{CloDF13}$ (plasmid pSU4833) (**Table 6**). In these experiments (**Figures 3**, **4**) the location pattern of each protein was the same in the absence or presence of a conjugative system (**Table 6**). However, although $MobB_{CloDF13}$ kept two poles as its main location in all the tested conditions, in the presence of R388 lacking functional $TrwB_{R388}$ it did not partially switch to a single-pole after 20 h, as in the absence of it.

The general observed pattern for all studied T4CPs was that the presence of T4SS enhanced the percentage of cells with the T4CP at the predominant location shown in the absence of the T4SS for each protein after 20 h (**Figures 3**, **4**). Specifically, the percentage of cells showing $TrwB_{R388}$ and $TMD_{TraJ}CD_{TrwB}$ at a single pole increased and so did the percentage of cells showing $MobB_{CloDF13}$ at both poles. Moreover, regarding $MobB_{CloDF13}$, the additional presence of its cognate mobilization region enhanced the effect produced by $T4SS_{R388}$.

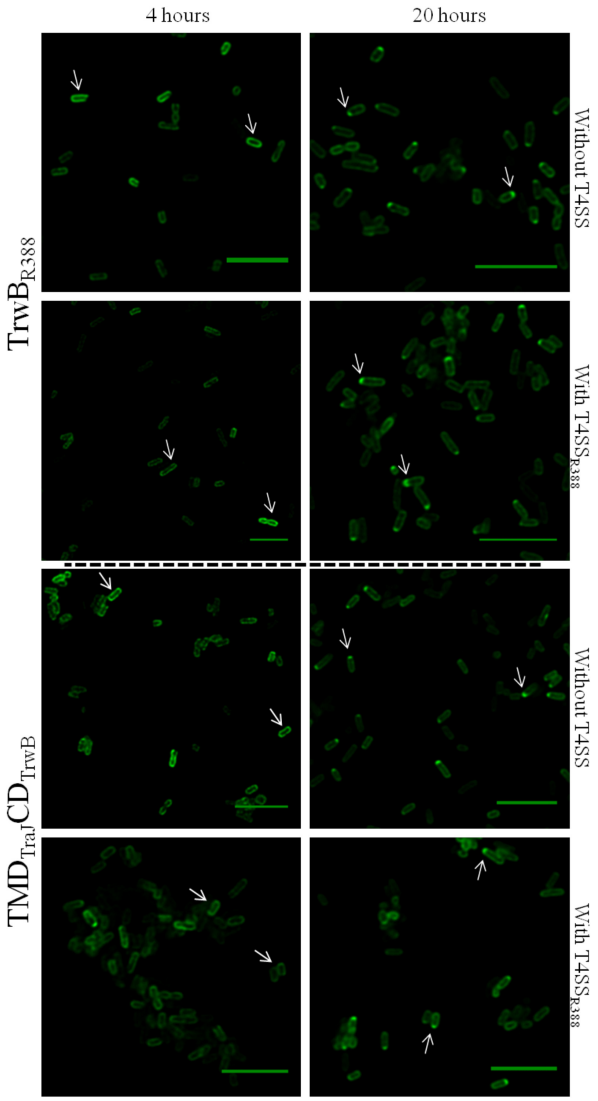


FIGURE 3 | Subcellular location of $TrwB_{R388}$GFP and $TMD_{TraJ}CD_{TrwB}$GFP fusion-proteins by confocal fluorescence microscopy. Proteins were expressed in *E. coli* BL21C41 (DE3) strain by induction with 1 mM IPTG for 4 (left panels) and 20 h (right panels) at 25°C. Subcellular location of $TrwB_{R388}$GFP and $TMD_{TraJ}CD_{TrwB}$GFP was determined in *E. coli* strains containing plasmid pSU1456 that expresses all R388 conjugative proteins except $TrwB_{R388}$, or without plasmid pSU1456. The images were acquired in a *Leica TCS SP5* confocal fluorescence microscope, with a 60× oil immersion objective. Sample excitation was performed with 488 nm wavelength, while fluorescence emission was measured between 505 and 525 nm. The images were analyzed using Huygens and ImageJ softwares. Arrrowheads indicate the eGFP fluorescence through the periphery (1st column) and at a single cell pole (2nd column). Scale bar: 5 μm.

TABLE 6 | Subcellular location of different T4CP-eGFP fusion-proteins at different expression times in the absence or presence of $T4SS_{R388}$.

	Expression time (h)	Without T4SS (%)			$T4SS_{R388}$ (%)			$T4SS_{R388}$ and $MOB_{CloDF13}$ (%)		
		M	1P	2P	M	1P	2P	M	1P	2P
$TrwB_{R388}$	4	**97 (77)**	11 (9)	18 (14)	**34 (84)**	3 (8)	3 (8)	–		
	20	31 (15)	**155 (75)**	29 (10)	10 (6)	**138 (90)**	6 (4)			
$TMD_{TraJ}CD_{TrwB}$	4	**85 (83)**	11 (11)	6 (6)	**55 (76)**	13 (18)	4 (6)	–		
	20	49 (40)	**62 (51)**	10 (8)	18 (14)	**85 (67)**	23 (18)			
$MobB_{CloDF13}$	4	18 (18)	8 (8)	**73 (74)**	22 (19)	16 (14)	**80 (68)**	7 (5)	19 (13)	**122 (82)**
	20	20 (18)	24 (21)	**69 (61)**	16 (15)	16 (15)	**74 (70)**	7 (4)	43 (22)	**143 (74)**
MobBΔTMD[a]	4	0	**40 (95)**[b]	2 (5)[b]	N.d.			N.d.		
	20	0	**35 (95)**[b]	2 (5)[b]						

The eGFP variants of the proteins in the first column were expressed in E. coli BL21C41 (DE3) strain by induction with 1 mM IPTG for 4 and 20 h. The numbers represent the amount of cells showing each location. Total number of cells counted for each sample was between 40 and 200. Numbers between parentheses represent the percentages calculated up to the total amount of cells counted for each sample. The main location for each condition is marked in bold. M, location through the perimeter of the cell membrane; 1P, location at a single-pole; 2P, location at both poles. –, without relevance for this study. N.d., not determined. [a]Protein expressed in E. coli BL21 (DE3). [b]Polar location in the cytosol.

DISCUSSION

The increase of multidrug-resistant bacteria has become one of the major health concerns in our society (World Health Organization [WHO], 2019), being bacterial conjugation one of the key mechanisms responsible for the spread of antibiotic resistance genes among bacteria (Bello-López et al., 2019). This process is performed through a T4SS, a multiprotein complex that transfers the nucleoprotein substrate from a donor into a recipient bacterium (Waksman, 2019). T4CPs are essential proteins during conjugation, as they connect the substrates to be transferred in the cytosol with the secretion machinery in the membrane (Gomis-Ruth et al., 2005). Despite their importance, as membrane proteins are challenging to be studied, their characterization has been mostly accomplished using mutants that lack their TMD (Schroder and Lanka, 2003; Tato et al., 2007; Larrea et al., 2017). However, several studies performed with $TrwB_{R388}$, the full-length T4CP of the conjugative plasmid R388, have proven that the TMD is more than a mere anchor to the membrane and that it has a role in protein activity, stability, oligomerization and subcellular localization (Moncalián et al., 1999; Hormaeche et al., 2002, 2004, 2006; Vecino et al., 2010, 2011; Segura et al., 2013, 2014).

The aim of this work has been to provide new data about different T4CPs that will contribute to infer general conclusions on their functioning to develop strategies to inhibit them and control the spread of antibiotic resistance genes. To do so, we studied the *in vivo* functionality, secondary structure, thermal stability, and subcellular location of $TrwB_{R388}$, its chimeric protein $TMD_{TraJ}CD_{TrwB}$, and the T4CP of the mobilizable plasmid CloDF13, $MobB_{CloDF13}$, and its TMD-less mutant, MobBΔTMD.

In vivo Functionality

$TMD_{TraJ}CD_{TrwB}$ was able to complement the Δ*trwB* mutation for R388 transfer, although with a lower transfer frequency than $TrwB_{R388}$. This decrease in conjugation frequency may be due to a combination of effects such as conformational changes in the CD of $TrwB_{R388}$ due to its chimeric nature that render a less active protein and the heterologous interaction between TMD_{TraJ} and the T4SS of R388 (Llosa et al., 2003). Thus, $TMD_{TraJ}CD_{TrwB}$-mediated transfer of R388 would occur through a specific interaction between CD_{TrwB} with its cognate relaxase, $TrwC_{R388}$, and an unspecific interaction of TMD_{TraJ} with the heterologous T4SS from R388. This could explain why $TMD_{TraJ}CD_{TrwB}$ did not complement the Δ*traJ* mutation (pKM101Δ*traJ*) since CD_{TrwB} could not recognize the heterologous $TraI_{pKM101}$ relaxase, even if TMD_{TraJ} interacted with its cognate $T4SS_{pKM101}$.

Moreover, $TMD_{TraJ}CD_{TrwB}$ showed negative dominance in the presence of the native T4CPs $TrwB_{R388}$ or $TraJ_{pKM101}$. Since it has been reported that TMD_{TraJ} can interact with both $T4SS_{R388}$ and $T4SS_{pKM101}$ (Llosa et al., 2003; De Paz et al., 2010; Celaya et al., 2017), a possible explanation for the observed negative dominance could be that competition for the T4SS occurred. According to this hypothesis, $TMD_{TraJ}CD_{TrwB}$ would have interacted with the secretion channel, sequestering it from interacting with the native T4CPs and reducing the conjugative rate of the wild type system. This hypothesis comes into agreement with the fact that the transfer of pKM101 using $TraJ_{pKM101}$ and $T4SS_{R388}$ is one order of magnitude lower than using $T4SS_{pKM101}$ (Llosa et al., 2003). Another explanation compatible with dominance experiments with both systems would be that non-functional heteroligomers were made between the native and the chimera proteins, competing for the conjugative machinery and therefore lowering the transfer frequency. Any of these alternatives or a combination of them would have caused a decrease in the plasmid transfer rate, as observed in the dominance mating assays. It must be underlined that a point mutation in the Walker A domain, $TMD_{TraJ}CD_{TrwB}$ (K142T), resulted in a non-functional phenotype. Hence, the K mutation totally arrested transfer capacity as previously reported for other T4CPs (Moncalián et al., 1999; Kumar and Das, 2002; Gunton et al., 2005), suggesting that concerning its NBD $TMD_{TraJ}CD_{TrwB}$ is functionally similar to $TrwB_{R388}$ despite the results obtained in the complementation and dominance studies.

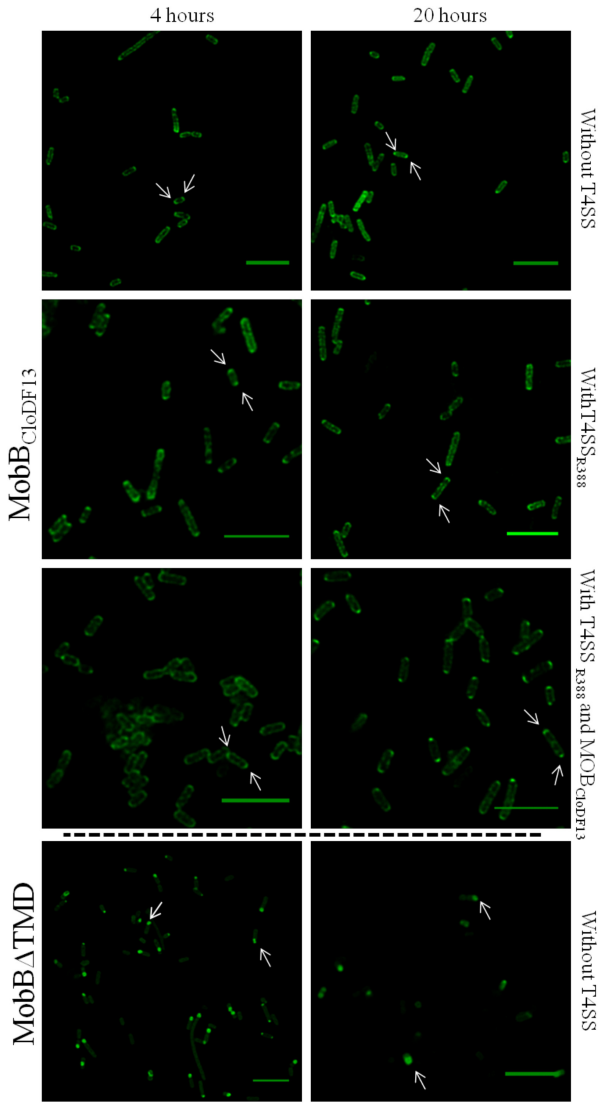


FIGURE 4 | Subcellular location of MobB$_{CloDF13}$GFP and MobBΔTMDGFP fusion-proteins by confocal fluorescence microscopy. MobB$_{CloDF13}$GFP and MobBΔTMDGFP proteins were expressed in *E. coli* BL21C41 (DE3) and *E. coli* BL21 (DE3) strains, respectively. Subcellular location of these proteins was determined by induction with 1 mM IPTG after 4 (left panels) and 20 h (right panels) at 25°C. Additionally, subcelular location was analyzed in the presence of pSU1456 plasmid, which expresses all R388 conjugative proteins except TrwB$_{R388}$, or without plasmid pSU1456. MobB$_{CloDF13}$ was also expressed in the presence of both pSU1456 and pSU4833 that codes for the mobilization proteins of plasmid CloDF13 (MOB$_{CloDF13}$), except for MobB$_{CloDF13}$. The images were acquired in a *Leica TCS SP5* confocal fluorescence microscope, with a 60× oil immersion objective. Sample excitation was performed with 488 nm wavelength, while fluorescence emission was measured between 505 and 525 nm. The images were analyzed using *Huygens* and *ImageJ* software. Arrrowheads indicate the eGFP fluorescence foci at both cell poles (MobB$_{CloDF13}$) and at a single cell pole (MobBΔTMD). Scale bar: 5 μm.

Regarding CloDF13 system, it was observed that the deletion of the TMD rendered a non-functional MobBΔTMD as it occurs with other TMD-less mutants such as TrwBΔN70 and PcfCΔN103 (Chen et al., 2008). However, a TcpA mutant lacking the TMD, but not the N-term cytosolic residues, TcpA$_{\Delta 46-104}$, was able to perform conjugation, although at a frequency lower than the wild type plasmid. In this regard, unpublished experiments with TrwBΔN8, which lacks the N-term cytosolic eight residues, showed a decrease in the transfer rate of more than three orders of magnitude (Vecino, 2009). Taken together these results, it seems that not only the TMD but also the N-term cytosolic residues have an important role in the transfer capacity of T4CPs. In this context, the importance of this small region in specific interactions with the relaxosome has already been described (Llosa et al., 2003; Schroder and Lanka, 2003). In the case of MobB$_{CloDF13}$ its N-term is located in the periplasm (**Figure 1A**), which could be an important feature for recognition and interaction with conjugative secretion channels.

Secondary Structure and Thermal Stability

As it has been described that the TMD influences the secondary structure of the CD and the thermal stability of TrwB$_{R388}$ (Hormaeche et al., 2004; Vecino et al., 2011, 2012), in this work we have studied whether this behavior can be observed in TMD$_{TraJ}$CD$_{TrwB}$ and in MobB$_{CloDF13}$.

Regarding TMD$_{TraJ}$CD$_{TrwB}$, one of the most important differences in comparison with TrwB$_{R388}$ was the appearance of a band at 1640 cm^{-1}, mainly assigned to flexible structures (non-periodic elements) related to a less compact overall structure (Echabe et al., 1998; Agopian et al., 2016). As the crystal structure of TrwBΔN70 shows flexible loops (Gomis-Rüth et al., 2002), the presence of the 1640 cm^{-1} band present at similar percentages (30–37%) in TMD$_{TraJ}$CD$_{TrwB}$, TrwBΔN50, and TrwBΔN70 mutant proteins but missing in TrwB$_{R388}$, could be explained as the loss of the compact structure of TrwB$_{R388}$ due to the deletion of its cognate TMD (Hormaeche et al., 2004).

Taking into account the values related to all the bands associated to β-sheet elements (1671, 1640, and 1626 cm^{-1}) it could be concluded that the total percentage of β-sheets in TMD$_{TraJ}$CD$_{TrwB}$ is just slightly smaller to that of TrwB$_{R388}$. Concerning β-turns, the band assigned to them in the TrwB$_{R388}$ related proteins, 1661–1665 cm^{-1}, was not observed in TMD$_{TraJ}$CD$_{TrwB}$. However, it can be postulated that the increase seen in TMD$_{TraJ}$CD$_{TrwB}$ of the band at 1671 cm^{-1}, was partially related to the β-turns component seen in the mutants at 1665 cm^{-1}. These would imply that the β-turns component of the chimeric protein is lower than the one of the native protein and similar to the TMD deletion mutants. Finally, TMD$_{TraJ}$CD$_{TrwB}$ shows a decrease in α-helix percentage (35%) in comparison with TrwB$_{R388}$ (41%), but an increase in comparison to the mutants (26%).

Therefore it seems that TMD$_{TraJ}$CD$_{TrwB}$ presents qualitative and quantitative features in between the native protein and the deletion mutants. These results suggest that the presence of a heterologous full-length TMD does provide a more compact and ordered structure to the T4CP in comparison to the TMD-less mutants, even if it does not reach the level of the native

protein. This result comes in agreement with the transfer capacity reduction of the chimeric protein that could be explained partially by the observed structural changes reported here.

To test if the results described for $TrwB_{R388}$ could be extrapolated to other T4CPs, the secondary structures of $MobB_{CloDF13}$ and MobBΔTMD were studied (**Table 5**). Surprisingly, MobBΔTMD presented an IR spectrum quite similar to the one obtained for $MobB_{CloDF13}$. It presented smaller helical structure percentages and higher β-turns percentages but both showed similar unordered and β-sheet percentages. This suggests that in $MobB_{CloDF13}$ the presence of the TMD does not have an effect on the structure of the CD as it does in $TrwB_{R388}$. As $MobB_{CloDF13}$ has to interact with heterologous T4SSs, it could be that its TMD has to interact with heterologous T4SSs without altering the structure of its CD where the specific interaction with its cognate relaxosome occurs.

Concerning the thermal denaturation of $TMD_{TraJ}CD_{TrwB}$, its Tm was similar to the ones described for $TrwB_{R388}$ and the mutant lacking the first transmembrane helix, TrwBΔN50 (**Table 5**). This result suggests that although the secondary structures of the mutants differ from that of the native protein, their overall thermal stability is similar. Additionally, when comparing the results between both studied systems, as expected due to the high instability of purified membrane proteins (González Flecha, 2017) the soluble proteins showed higher mid-point denaturation temperatures than the full-length proteins. Specifically, MobBΔTMD and TrwBΔN70 showed similar Tm values (Tm 62 and 63°C, respectively); on the contrary, $MobB_{CloDF13}$ was more stable than $TrwB_{R388}$ against thermal denaturation (Tm 56 and 48°C, respectively). This could be related to different buffer compositions that had to be used when $MobB_{CloDF13}$ and $TrwB_{R388}$ were analyzed.

Subcellular Location

The polar location of proteins in bacteria underlines their sophisticated internal organization, being important in many processes like chemotaxis and cellular division (Howard, 2004). Similarly, subcellular location has been considered important in bacterial conjugation (Chen et al., 2008; Leonetti et al., 2015). Sequence analysis, cell fractionation, and protein purification experiments proved that $TrwB_{R388}$, $TMD_{TraJ}CD_{TrwB}$, and $MobB_{CloDF13}$ are located in the bacterial membrane, while MobBΔTMD is located in the cytosol (data not shown). As studies in the literature do not show a consensus pattern either in the location of the T4CPS nor in the role of each domain in this property (Kumar and Das, 2002; Bauer et al., 2011; Segura et al., 2014), the subcellular location of $TrwB_{R388}$, $TMD_{TraJ}CD_{TrwB}$, $MobB_{CloDF13}$, and MobBΔTMD was investigated.

In this work we observed $TrwB_{R388}$ located along the membrane after 4 h of induction and only after 20 h it focused at single pole. These results differ from our previous results where the polar location of $TrwB_{R388}$ was observed after 4 h (Segura et al., 2014). However, the induction OD_{600} values were different (0.4 vs. 0.7, this work and previous work, respectively), probably rendering populations at different growth phase. On the contrary $MobB_{CloDF13}$ was located at both poles in the membrane at 4 and 20 h after induction.

Previous studies have reported that in the absence of other conjugative proteins, T4CPs that lacked the whole TMD or even the periplasmic loop located in the cytosol or at the membrane periphery, respectively (Kumar and Das, 2002; Segura et al., 2014). Moreover, the TMD alone of $TrwB_{R388}$ located at the membrane poles without the need for the CD (Segura et al., 2014), suggesting a leading role for the TMD in the subcellular location of the T4CPs. Surprisingly, MobBΔTMD located at a single cell pole in the cytosol even in the absence of other conjugative proteins. These results suggest that mobilizable plasmid-related T4CPs could use different mechanisms than VirD4-type T4CPs for subcellular location. It has been speculated that the polar location of T4CPs could be related to interactions with the cardiolipin enriched membrane poles (Mileykovskaya and Dowhan, 2009; Segura et al., 2014), but at the same time mediated by complex and dynamic changes in transduction, cytoskeleton proteins, etc. (Shapiro et al., 2002). Since interactions between mobilizable plasmid-related T4CPs and T4SSs are not specific, these T4CPs could have evolved to develop different mechanisms to interact with the membrane and ensure their polar location.

Previous studies (Kumar and Das, 2002; Gunton et al., 2005; Segura et al., 2014) have reported that the location of native T4CPs is independent of the presence of the rest of the conjugative proteins. Our results suggest that the presence of a complete conjugative system (i.e., mobilizable region and secretion channel) seems to enhance the polar location of wild type T4CPs. Moreover, although the location of $MobB_{CloDF13}$ at both poles was enhanced in the presence of $T4SS_{R388}$, it was further enhanced when $MOB_{CloDF13}$ was also present.

Taking all together, it seems that that no universal location patterns can be attributed to T4CPs. Nevertheless, three conclusions can be undertaken regarding subcellular location: (i) T4CPs localize either at a single pole or both poles, depending on the system; (ii) the presence of a TMD is not essential for the polar location of a mobilizable plasmid associated T4CP, and (iii) the presence of a conjugative system enhances the polar location of full-length T4CPs.

To sum up, the comparative study between the conjugative system related $TrwB_{R388}$ and the mobilizable plasmid-related $MobB_{CloDF13}$ and their variants has underlined that the characteristics described for the paradigmatic conjugative plasmid related VirD4-type T4CPs and their TMDs should not be ascribed to the whole T4CP family.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation, to any qualified researcher.

AUTHOR CONTRIBUTIONS

IÁ-R, CG, and IAl contributed to the design of the work (text and figures) and the acquisition of the data, writing and revision of the content, approval of the

last version, and ensuring accuracy and integrity of the work. IAr and JA contributed to acquisition of the data, revision of the content, and approval of the last version of the work. BU-U contributed to writing, revision of the content, and approval of the last version. All authors contributed to the article and approved the submitted version.

FUNDING

This work was in part supported by grants from the University of the Basque Country (GIU18/229 and COLAB19/08) and the Industry Department of the Basque Government (ELKARTEK 2020 KK-2020/00007). IÁ-R was a pre-doctoral student supported by the Basque Government.

ACKNOWLEDGMENTS

We would like to sincerely thank Dr. Lide Arana for her invaluable help and contribution to this work, to Unai Lorenzo for his help with the confocal microscope and image treatment, and to Dr. Louise Bird for her help with the high-throughput cloning at OPPF-UK.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2020.00185/full#supplementary-material

REFERENCES

Agopian, A., Quetin, M., and Castano, S. (2016). Structure and interaction with lipid membrane models of Semliki Forest virus fusion peptide. *Biochim. Biophys. Acta Biomembr.* 1858, 2671–2680. doi: 10.1016/j.bbamem.2016.07.003

Agúndez, L., Zárate-Pérez, F., Meier, A. F., Bardelli, M., Llosa, M., Escalante, C. R., et al. (2018). Exchange of functional domains between a bacterial conjugative relaxase and the integrase of the human adeno-associated virus. *PLoS One* 13:e0200841. doi: 10.1371/journal.pone.0200841

Alvarez-Martinez, C. E., and Christie, P. J. (2009). Biological diversity of prokaryotic type IV secretion systems. *Microbiol. Mol. Biol. Rev.* 73, 775–808. doi: 10.1128/mmbr.00023-09

Andraka, N., Sánchez-Magraner, L., García-Pacios, M., Goñi, F. M., and Arrondo, J. L. R. (2017). The conformation of human phospholipid scramblase 1, as studied by infrared spectroscopy. Effects of calcium and detergent. *Biochim. Biophys. Acta Biomembr.* 1859, 1019–1028. doi: 10.1016/j.bbamem.2017.02.015

Arrondo, J. L., and Goñi, F. M. (1999). Structure and dynamics of membrane proteins as studied by infrared spectroscopy. *Prog. Biophys. Mol. Biol.* 72, 367–405. doi: 10.1016/s0079-6107(99)00007-3

Arrondo, J. L., Muga, A., Castresana, J., and Goñi, F. M. (1993). Quantitative studies of the structure of proteins in solution by Fourier-transform infrared spectroscopy. *Prog. Biophys. Mol. Biol.* 59, 23–56.

Bauer, T., Rösch, T., Itaya, M., and Graumann, P. L. (2011). Localization pattern of conjugation machinery in a gram-positive bacterium. *J. Bacteriol.* 193, 6244–6256. doi: 10.1128/jb.00175-11

Bello-López, J. M., Cabrero-Martínez, O. A., Ibáñez-Cervantes, G., Hernández-Cortez, C., Pelcastre-Rodríguez, L. I., Gonzalez-Avila, L. U., et al. (2019). Horizontal gene transfer and its association with antibiotic resistance in the genus *Aeromonas* spp. *Microorganisms* 7:363. doi: 10.3390/microorganisms7090363

Berrow, N. S., Alderton, D., Sainsbury, S., Nettleship, J., Assenberg, R., Rahman, N., et al. (2007). A versatile ligation-independent cloning method suitable for high-throughput expression screening applications. *Nucleic Acids Res.* 35, e45. doi: 10.1093/nar/gkm047

Bird, L. E. (2011). High throughput construction and small scale expression screening of multi-tag vectors in *Escherichia coli*. *Methods* 55, 29–37. doi: 10.1016/j.ymeth.2011.08.002

Bourg, G., Sube, R., O'Callaghan, D., and Patey, G. (2009). Interactions between Brucella suis VirB8 and its homolog TraJ from the plasmid pSB102 underline the dynamic nature of type IV secretion systems. *J. Bacteriol.* 191, 2985–2992. doi: 10.1128/jb.01426-08

Cabezón, E., Sastre, J. I, and de la Cruz, F. (1997). Genetic evidence of a coupling role for the TraG protein family in bacterial conjugation. *Mol. Gen. Genet.* 254, 400–406. doi: 10.1007/s004380050432

Celaya, G., Perales-Calvo, J., Muga, A., Moro, F., and Rodriguez-Larrea, D. (2017). Label-free, multiplexed, single-molecule analysis of protein-DNA complexes with nanopores. *ACS Nano* 11, 5815–5825. doi: 10.1021/acsnano.7b01434

Chen, Y., Zhang, X., Manias, D., Yeo, H.-J., Dunny, G. M., and Christie, P. J. (2008). Enterococcus faecalis PcfC, a Spatially Localized Substrate Receptor for Type IV Secretion of the pCF10 Transfer Intermediate. *J. Bacteriol.* 190, 3632–3645. doi: 10.1128/jb.01999-07

Cormack, B. P., Valdivia, R. H., and Falkow, S. (1996). *FACS-Optimized Mutants of the Green Fluorescent Protein (GFP)*. Amsterdam: Elsevier, 33–38.

Datta, N., and Hedges, R. W. (1972). Trimethoprim resistance conferred by W plasmids in Enterobacteriaceae. *J. Gen. Microbiol.* 72, 349–355. doi: 10.1099/00221287-72-2-349

De Paz, H. D., Larrea, D., Zunzunegui, S., Dehio, C., De La Cruz, F., and Llosa, M. (2010). Functional dissection of the conjugative coupling protein TrwB. *J. Bacteriol.* 192, 2655–2669. doi: 10.1128/jb.01692-09

Drew, D., Lerch, M., Kunji, E., Slotboom, D.-J., and de Gier, J.-W. (2006). Optimization of membrane protein overexpression and purification using GFP fusions. *Nat. Methods* 3, 303–313. doi: 10.1038/nmeth0406-303

Drew, D., Slotboom, D.-J., Friso, G., Reda, T., Genevaux, P., Rapp, M., et al. (2005). A scalable, GFP-based pipeline for membrane protein overexpression screening and purification. *Protein Sci.* 14, 2011–2017. doi: 10.1110/ps.051466205

Echabe, I., Prado, A., Goñi, F. M., Arrondo, J. L. R., and Dornberger, U. (1998). Topology of sarcoplasmic reticulum Ca^{2+}-ATPase: an infrared study of thermal denaturation and limited proteolysis. *Protein Sci.* 7, 1172–1179. doi: 10.1002/pro.5560070511

Gomis-Ruth, F., Sola, M., Cruz, F., and Coll, M. (2005). Coupling factors in macromolecular type-IV secretion machineries. *Curr. Pharm. Des.* 10, 1551–1565. doi: 10.2174/1381612043384817

Gomis-Rüth, F. X., Moncalián, G., de la Cruz, F., and Coll, M. (2002). Conjugative plasmid protein TrwB, an integral membrane type IV secretion system coupling protein. *J. Biol. Chem.* 277, 7556–7566. doi: 10.1074/jbc.m110462200

Gomis-Rüth, F. X., Moncalián, G., Pérez-Luque, R., González, A., Cabezón, E., de la Cruz, F., et al. (2001). The bacterial conjugation protein TrwB resembles ring helicases and F1-ATPase. *Nature* 409, 637–641. doi: 10.1038/35054586

González Flecha, F. L. (2017). Kinetic stability of membrane proteins. *Biophys. Rev.* 9, 563–572. doi: 10.1007/s12551-017-0324-0

Gunton, J. E., Gilmour, M. W., Alonso, G., and Taylor, D. E. (2005). Subcellular localization and functional domains of the coupling protein, TraG, from IncHI1 plasmid R27. *Microbiology* 151, 3549–3561. doi: 10.1099/mic.0.28255-0

Hamilton, C. M., Lee, H., Li, P.-L., Cook, D. M., Piper, K. R., von Bodman, S. B., et al. (2000). TraG from RP4 and TraG and VirD4 from Ti plasmids confer relaxosome specificity to the conjugal transfer system of pTiC58. *J. Bacteriol.* 182, 1541–1548. doi: 10.1128/jb.182.6.1541-1548.2000

Hormaeche, I., Alkorta, I., Moro, F., Valpuesta, J. M., Goni, F. M., and De La Cruz, F. (2002). Purification and properties of TrwB, a hexameric, ATP-binding integral membrane protein essential for R388 plasmid conjugation. *J. Biol. Chem.* 277, 46456–46462. doi: 10.1074/jbc.m207250200

Hormaeche, I., Iloro, I., Arrondo, J. L. R., Goñi, F. M., De La Cruz, F., and Alkorta, I. (2004). Role of the transmembrane domain in the stability of TrwB, an integral protein involved in bacterial conjugation. *J. Biol. Chem.* 279, 10955–10961. doi: 10.1074/jbc.m310422200

Hormaeche, I., Segura, R. L., Vecino, A. J., Goñi, F. M., de la Cruz, F., and Alkorta, I. (2006). The transmembrane domain provides nucleotide binding specificity to the bacterial conjugation protein TrwB. *FEBS Lett.* 580, 3075–3082. doi: 10.1016/j.febslet.2006.04.059

Howard, M. (2004). A mechanism for polar protein localization in bacteria. *J. Mol. Biol.* 335, 655–663. doi: 10.1016/j.jmb.2003.10.058

Kumar, R. B., and Das, A. (2002). Polar location and functional domains of the *Agrobacterium tumefaciens* DNA transfer protein VirD4. *Mol. Microbiol.* 43, 1523–1532. doi: 10.1046/j.1365-2958.2002.02829.x

Langer, P. J., and Walker, G. C. (1981). Restriction endonuclease cleavage map of pKM101: relationship to parental plasmid R46. *Mol. Gen. Genet.* 182, 268–272. doi: 10.1007/BF00269669

Larrea, D., de Paz, H. D., Matilla, I., Guzmán-Herrador, D. L., Lasso, G., de la Cruz, F., et al. (2017). Substrate translocation involves specific lysine residues of the central channel of the conjugative coupling protein TrwB. *Mol. Genet. Genomics* 292, 1037–1049. doi: 10.1007/s00438-017-1331-3

Leonetti, C. T., Hamada, M. A., Laurer, S. J., Broulidakis, M. P., Swerdlow, K. J., Lee, C. A., et al. (2015). Critical components of the conjugation machinery of the integrative and conjugative element ICEBs1 of *Bacillus subtilis*. *J. Bacteriol.* 197, 2558–2567. doi: 10.1128/jb.00142-15

Llosa, M., and Alkorta, I. (2017). Coupling proteins in type IV secretion. *Curr. Top. Microbiol. Immunol.* 413, 143–168. doi: 10.1007/978-3-319-75241-9_6

Llosa, M., Bolland, S., and de la Cruz, F. (1994). Genetic organization of the conjugal DNA processing region of the IncW plasmid R388. *J. Mol. Biol.* 235, 448–464. doi: 10.1006/jmbi.1994.1005

Llosa, M., Zunzunegui, S., and de la Cruz, F. (2003). Conjugative coupling proteins interact with cognate and heterologous VirB10-like proteins while exhibiting specificity for cognate relaxosomes. *Proc. Natl. Acad. Sci. U.S.A.* 100, 10465–10470. doi: 10.1073/pnas.1830264100

Matilla, I., Alfonso, C., Rivas, G., Bolt, E. L., de la Cruz, F., and Cabezon, E. (2010). The conjugative DNA translocase TrwB is a structure-specific DNA-binding protein. *J. Biol. Chem.* 285, 17537–17544. doi: 10.1074/jbc.m109.084137

Mileykovskaya, E., and Dowhan, W. (2009). Cardiolipin membrane domains in prokaryotes and eukaryotes. *Biochim. Biophys. Acta Biomembr.* 1788, 2084–2091. doi: 10.1016/j.bbamem.2009.04.003

Moncalián, G., Cabezón, E., Alkorta, I., Valle, M., Moro, F., Valpuesta, J. M., et al. (1999). Characterization of ATP and DNA binding activities of TrwB, the coupling protein essential in plasmid R388 conjugation. *J. Biol. Chem.* 274, 36117–36124. doi: 10.1074/jbc.274.51.36117

Nuñez, B., and de la Cruz, F. (2001). Two atypical mobilization proteins are involved in plasmid CloDF13 relaxation. *Mol. Microbiol.* 39, 1088–1099. doi: 10.1046/j.1365-2958.2001.02308.x

Parsons, J. A., Bannam, T. L., Devenish, R. J., and Rood, J. I. (2007). TcpA, an FtsK/SpoIIIE homolog, is essential for transfer of the conjugative plasmid pCW3 in *Clostridium perfringens*. *J. Bacteriol.* 189, 7782–7790. doi: 10.1128/jb.00783-07

Paterson, E. S., Moré, M. I., Pillay, G., Cellini, C., Woodgate, R., Walker, G. C., et al. (1999). Genetic analysis of the mobilization and leading regions of the IncN plasmids pKM101 and pCU1. *J. Bacteriol.* 181, 2572–2583. doi: 10.1128/jb.181.8.2572-2583.1999

Redzej, A., Ukleja, M., Connery, S., Trokter, M., Felisberto-Rodrigues, C., Cryar, A., et al. (2017). Structure of a VirD4 coupling protein bound to a VirB type IV secretion machinery. *EMBO J.* 36, 3080–3095. doi: 10.15252/embj.201796629

Schroder, G., and Lanka, E. (2003). TraG-like proteins of type IV secretion systems: functional dissection of the multiple activities of TraG (RP4) and TrwB (R388). *J. Bacteriol.* 185, 4371–4381. doi: 10.1128/jb.185.15.4371-4381.2003

Segura, R. L., Águila-Arcos, S., Ugarte-Uribe, B., Vecino, A. J., De La Cruz, F., Goñi, F. M., et al. (2013). The transmembrane domain of the T4SS coupling protein TrwB and its role in protein-protein interactions. *Biochim. Biophys. Acta Biomembr.* 1828, 2015–2025. doi: 10.1016/j.bbamem.2013.05.022

Segura, R. L., Águila-Arcos, S., Ugarte-Uribe, B., Vecino, A. J., de la Cruz, F., Goñi, F. M., et al. (2014). Subcellular location of the coupling protein TrwB and the role of its transmembrane domain. *Biochim. Biophys. Acta Biomembr.* 1838, 223–230. doi: 10.1016/j.bbamem.2013.08.016

Shapiro, L., McAdams, H. H., and Losick, R. (2002). Generating and exploiting polarity in bacteria. *Science* 298, 1942–1946. doi: 10.1126/science.1072163

Smillie, C., Garcillán-Barcia, M. P., Francia, M. V., Rocha, E. P. C., and de la Cruz, F. (2010). Mobility of plasmids. *Microbiol. Mol. Biol. Rev.* 74, 434–452.

Tato, I., Matilla, I., Arechaga, I., Zunzunegui, S., de la Cruz, F., and Cabezon, E. (2007). The ATPase activity of the DNA transporter TrwB is modulated by protein TrwA: implications for a common assembly mechanism of DNA translocating motors. *J. Biol. Chem.* 282, 25569–25576.

Tato, I., Zunzunegui, S., de la Cruz, F., and Cabezon, E. (2005). TrwB, the coupling protein involved in DNA transport during bacterial conjugation, is a DNA-dependent ATPase. *Proc. Natl. Acad. Sci. U.S.A.* 102, 8156–8161. doi: 10.1073/pnas.0503402102

Tsirigos, K. D., Peters, C., Shu, N., Käll, L., and Elofsson, A. (2015). The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. *Nucleic Acids Res.* 43, W401–W407.

Vecino, A. J. (2009). *La Reconstitución en Proteoliposomas de TrwB, un Nanomotor que Transporta DNA, Revela la Importancia del Dominio Transmembrana de la Proteína*. Ph.D. thesis. País Vasco: University of the Basque Country.

Vecino, A. J., de la Arada, I., Segura, R. L., Goñi, F. M., de la Cruz, F., Arrondo, J. L. R., et al. (2011). Membrane insertion stabilizes the structure of TrwB, the R388 conjugative plasmid coupling protein. *Biochim. Biophys. Acta* 1808, 1032–1039. doi: 10.1016/j.bbamem.2010.12.025

Vecino, A. J., de Segura, R. L., de la Arada, I., de la Cruz, F., Goñi, F. M., Arrondo, J. L., et al. (2012). Deletion of a single helix from the transmembrane domain causes large changes in membrane insertion properties and secondary structure of the bacterial conjugation protein TrwB. *Biochim. Biophys. Acta Biomembr.* 1818, 3158–3166. doi: 10.1016/j.bbamem.2012.08.015

Vecino, A. J., Segura, R. L., Ugarte-Uribe, B., Aguila, S., Hormaeche, I., de la Cruz, F., et al. (2010). Reconstitution in liposome bilayers enhances nucleotide binding affinity and ATP-specificity of TrwB conjugative coupling protein. *Biochim. Biophys. Acta* 1798, 2160–2169. doi: 10.1016/j.bbamem.2010.07.005

Waksman, G. (2019). From conjugation to T4S systems in Gram-negative bacteria: a mechanistic biology perspective. *EMBO Rep.* 20:e47012.

Waldo, G. S., Standish, B. M., Berendzen, J., and Terwilliger, T. C. (1999). Rapid protein-folding assay using green fluorescent protein. *Nat. Biotechnol.* 17, 691–695. doi: 10.1038/10904

Whitaker, N., Berry, T. M., Rosenthal, N., Gordon, J. E., Gonzalez-Rivera, C., Sheehan, K. B., et al. (2016). Chimeric coupling proteins mediate transfer of heterologous type IV effectors through the *Escherichia coli* pKM101-encoded conjugation machine. *J. Bacteriol.* 198, 2701–2718. doi: 10.1128/jb.00378-16

World Health Organization [WHO] (2019). *Thirteenth General Programme of Work 2019–2023*. Geneva: World Health Organization [WHO].

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Álvarez-Rodríguez, Ugarte-Uribe, de la Arada, Arrondo, Garbisu and Alkorta. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Molecular Biosciences

REVIEW
published: 12 August 2020
doi: 10.3389/fmolb.2020.00201

Type IV Coupling Proteins as Potential Targets to Control the Dissemination of Antibiotic Resistance

Itxaso Álvarez-Rodríguez[1], Lide Arana[1], Begoña Ugarte-Uribe[1†], Elena Gómez-Rubio[2], Sonsoles Martín-Santamaría[2], Carlos Garbisu[3] and Itziar Alkorta[1]*

[1] Department of Biochemistry and Molecular Biology, University of the Basque Country, Leioa, Spain, [2] Department of Structural and Chemical Biology, Centro de Investigaciones Biológicas Margarita Salas, CIB-CSIC, Madrid, Spain, [3] Department of Conservation of Natural Resources, Soil Microbial Ecology Group, NEIKER - Basque Institute for Agricultural Research and Development, Basque Research and Technology Alliance (BRTA), Parque Científico y Tecnológico de Bizkaia, Derio, Spain

OPEN ACCESS

Edited by:
Tatiana Venkova, Fox Chase Cancer Center, United States

Reviewed by:
Ronnie Per-Arne Berntsson, Umeå University, Sweden
Fabian Lorenzo-Diaz, University of La Laguna, Spain

****Correspondence:***
Itziar Alkorta
itzi.alkorta@ehu.eus

†Present address:
Begoña Ugarte-Uribe, EMBL Heidelberg Cell Biology and Biophysics, Heidelberg, Germany

Specialty section:
This article was submitted to Molecular Recognition, a section of the journal Frontiers in Molecular Biosciences

Received: *29 May 2020*
Accepted: *27 July 2020*
Published: *12 August 2020*

Citation:
Álvarez-Rodríguez I, Arana L, Ugarte-Uribe B, Gómez-Rubio E, Martín-Santamaría S, Garbisu C and Alkorta I (2020) Type IV Coupling Proteins as Potential Targets to Control the Dissemination of Antibiotic Resistance. Front. Mol. Biosci. 7:201. doi: 10.3389/fmolb.2020.00201

The increase of infections caused by multidrug-resistant bacteria, together with the loss of effectiveness of currently available antibiotics, represents one of the most serious threats to public health worldwide. The loss of human lives and the economic costs associated to the problem of the dissemination of antibiotic resistance require immediate action. Bacteria, known by their great genetic plasticity, are capable not only of mutating their genes to adapt to disturbances and environmental changes but also of acquiring new genes that allow them to survive in hostile environments, such as in the presence of antibiotics. One of the major mechanisms responsible for the horizontal acquisition of new genes (*e.g.*, antibiotic resistance genes) is bacterial conjugation, a process mediated by mobile genetic elements such as conjugative plasmids and integrative conjugative elements. Conjugative plasmids harboring antibiotic resistance genes can be transferred from a donor to a recipient bacterium in a process that requires physical contact. After conjugation, the recipient bacterium not only harbors the antibiotic resistance genes but it can also transfer the acquired plasmid to other bacteria, thus contributing to the spread of antibiotic resistance. Conjugative plasmids have genes that encode all the proteins necessary for the conjugation to take place, such as the type IV coupling proteins (T4CPs) present in all conjugative plasmids. Type VI coupling proteins constitute a heterogeneous family of hexameric ATPases that use energy from the ATP hydrolysis for plasmid transfer. Taking into account their essential role in bacterial conjugation, T4CPs are attractive targets for the inhibition of bacterial conjugation and, concomitantly, the limitation of antibiotic resistance dissemination. This review aims to compile present knowledge on T4CPs as a starting point for delving into their molecular structure and functioning in future studies. Likewise, the scientific literature on bacterial conjugation inhibitors has been reviewed here, in an attempt to elucidate the possibility of designing T4CP-inhibitors as a potential solution to the dissemination of multidrug-resistant bacteria.

Keywords: antibiotic resistance, bacterial conjugation, coupling proteins, drug discovery, novel conjugation inhibitors

INTRODUCTION

Since their discovery, antibiotics have undoubtedly been one of the biggest, if not the biggest, medical revolutions. Regrettably, the abuse and misuse of antibiotics (mainly, for medicinal and veterinary purposes) has led to the emergence and dissemination of antibiotic resistant bacteria (ARB). The dissemination of antibiotic resistance is mainly driven by the horizontal transfer of antibiotic resistance genes (ARGs) using a variety of mobile genetic elements (MGEs), such as conjugative plasmids and integrative conjugative elements (ICEs) (Chan, 2015). In particular, biofilms represent a hot spot for horizontal gene transfer (HGT) by bacterial conjugation (Águila-Arcos et al., 2017).

Antibiotic resistance is responsible for the death of more than 33,000 people per year only in Europe, as well as a medical expenditure of around 1.5 billion euros (Cassini et al., 2019). It has been estimated that, by the year 2050, 10 million human lives might be at risk if no solution is found (Cassini et al., 2019). About 70% of hospital diseases are currently caused by multi-resistant bacteria (bacteria resistant to several antibiotics), further complicating their treatment. Not surprisingly, antibiotic resistance is at present one of the major threats to modern medicine and public health (World Health Organization, 2015, 2019a). In order to both inform and sensitize governments and the society in general, the World Health Organization presented a list of the most critical antibiotic-resistant pathogen species (Tacconelli et al., 2018). Although inevitably much more attention has been paid to the magnitude of this problem in hospital settings, the role of the environment on the emergence and dissemination of antibiotic resistance has been relatively recently recognized (Garbisu et al., 2018; Urra et al., 2019a,b,c).

Alas, the magnitude of this alarming problem has increased during the last years and decades, among other factors, due to the scarce presence of innovative proposals and treatments to fight ARB, as well as insufficient resources to carry out the required research to solve, or at least mitigate, the current antibiotic resistance crisis (Högberg et al., 2010; World Health Organization, 2019b).

At this point, it is essential to devise and develop new effective strategies in the fight against the emergence and dissemination of multi-resistant bacteria. Bacterial conjugation is one of the main mechanisms by which bacteria can acquire ARGs. Since type IV coupling proteins (T4CPs) are essential for the conjugation to take place, *a priori* they are attractive therapeutic targets for the inhibition of bacterial conjugation and, concomitantly, the limitation of antibiotic resistance dissemination. Therefore, in this review, in addition to presenting the existing knowledge on T4CPs, the latest advances in the search for inhibitors of bacterial conjugation will be described and discussed in an attempt to lay the foundations to facilitate the much needed work in the search for molecules that can efficiently inhibit T4CPs, thus opening the door to novel approaches and methodologies to control the spread of antibiotic resistance among bacteria.

HORIZONTAL DNA TRANSFER

Bacteria can modify their genetic material through three main mechanisms: (i) mutations that alter the DNA sequence; (ii) genetic rearrangements; and (iii) acquisition of new genetic material via HGT. Nowadays, it is a well-known fact that a high proportion of the bacterial genome corresponds to horizontally acquired genetic material. This fact highlights the key role of HGT for bacterial genome plasticity (Thomas and Nielsen, 2005; Johnsborg et al., 2007). Actually, bacteria can horizontally acquire genetic material through three different mechanisms: transformation, transduction, and conjugation (the latter is the most widespread mechanism of HGT) (Bello-López et al., 2019). A fourth, still scarcely understood, mechanism of HGT is mediated by bacterial membrane vesicles (Johnson and Grossman, 2015).

Mobile Genetic Elements

The HGT processes of transduction and conjugation are mediated by MGEs. Mobile genetic elements have genes that encode the proteins necessary to move genetic material between different bacteria (intercellular transfer) or within the same bacteria among different DNA molecules (intracellular transfer) (Frost et al., 2005). Due to the complexity of the interactions among them, it is not easy to describe and unambiguously classify the variety of currently known MGEs. In any case, in general terms, the following groups can be differentiated: (i) bacteriophages; (ii) conjugative and mobilizable plasmids; (iii) genomic islands; (iv) transposable elements; and (v) integrons.

Bacteriophages (i.e., viruses that infect bacteria) can pack DNA segments of the host bacterium and transfer them to a new host bacterium, where they can be incorporated into its chromosome by recombination (Frost et al., 2005).

Conjugative and mobilizable plasmids are stable, self-replicating DNA molecules that carry genes not required for essential cellular functions but potentially useful in harsh conditions (Frost et al., 2005). Both types of plasmids can transfer genes between bacteria by means of using a type IV secretion system (T4SS) as conjugative machinery. But while conjugative plasmids code for all the proteins needed for self-transfer, mobilizable plasmids instead need a co-resident conjugative plasmid for them to be transferred by conjugation (Smillie et al., 2010).

Genomic islands are clusters of genes integrated into the bacterial chromosome through which bacteria can acquire advantageous functions such as, for instance, the ability to become infectious or resistant to antibiotics. Genomic islands are passively propagated during chromosomal replication, segregation and cell division (Bellanger et al., 2014). Two types of genomic islands can be distinguished: (i) ICEs, also known as conjugative transposons, contain a T4SS and, in consequence, they have the capacity to be transferred to other bacteria (Johnson and Grossman, 2015); and (ii) integrative mobilizable elements (IMEs) which encode their own excision, recircularization and integration sequences but lack some of the genes necessary for conjugative transfer and, hence, they need a co-resident ICE

or conjugative plasmid for them to be transferred by conjugation (Bellanger et al., 2014).

Transposable elements are DNA sequences capable of moving themselves into different sites of the bacterial genome by (i) excision from the original site and insertion into a new site; or (ii) the generation of a new copy which will then be moved to a new site. This group of transposable elements includes (i) transposons (Tn) that, in addition to the genes specifically needed for transposition, encode other genes; and (ii) insertion sequences (IS) which only contain the genes required for transposition (Siguier et al., 2014).

Integrons can capture gene cassettes using site-specific recombination mediated by an integron-encoded integrase. Furthermore, they contain a promoter necessary for efficient transcription and expression of the gene cassettes. Relevantly, integrons can be embedded in different MGEs which they use for intra- and intercellular mobilization (Gillings et al., 2008).

For many years, it has been thought that conjugative plasmids were the main cause of HGT. Nevertheless, recent research has proven the relevance of ICEs (integrative conjugative elements) and IMEs (integrative and mobilizable elements) for HGT (Guédon et al., 2017). In any case, since conjugative plasmids have genes that encode all the elements required for conjugation, they have been studied thoroughly and accurately.

Type IV Secretion Systems

Type IV secretion systems (T4SS) are macromolecular assemblies that can transport DNA and/or proteins (Christie, 2016). This broad superfamily, in terms of function and structure, of macromolecular complexes can be found in both Gram-negative and Gram-positive bacteria (Christie et al., 2017). Different classifications, based on different criteria, have been proposed for T4SSs (**Table 1**). Although they do not cover all the cases, due to the great variability observed within this superfamily, two main classifications are regularly used. The first classification is based on the T4SS function and differentiates two types (de Paz et al., 2005): T4SS that take part in conjugation (cT4SS) and T4SS that take part in pathogenic processes (pT4SS). In any case, some T4SS, such as the one present in *Agrobacterium tumefaciens*, show a dual nature (Li and Christie, 2018). The second classification is based on the protein components (Christie et al., 2017; Grohmann et al., 2018). In this case, T4SS are similarly divided into two groups: type IV A (T4ASS) and type IV B (T4BSS), whose paradigmatic systems are the VirB/VirD system of *A. tumefaciens* and the Dot/Icm system of *Legionella pneumophila*, respectively. T4ASS are composed of approximately 12 proteins related to the VirB-D proteins of *A. tumefaciens* (Christie et al., 2005). Although T4ASS proteins are termed VirB1-11, there is a tendency to use TivB1-11 nomenclature, instead of VirB, to avoid confusion (Thomas et al., 2017). T4ASS are associated with the dissemination of ARGs. On the other hand, T4BSS are composed of 25 proteins, of which only a few are homologous to the VirB system (Voth et al., 2012).

Gram-negative T4ASS evolved as supramolecular protein structures formed by three different functional units or modules (**Figure 1**). The first module, formed by three ATPases [VirD4 (T4CP), VirB4, VirB11], is an energy center located in the cytoplasmic side at the entrance of the secretion channel (Ripoll-Rozada et al., 2013). These proteins interact with the second module, a larger infrastructure called the inner membrane complex (IMC) (Low et al., 2014), formed by VirB3, VirB6, VirB8, and the N-terminus of VirB10. The IMC, responsible for substrate translocation through the inner membrane (IM), is connected to the outer membrane core complex (OMCC), i.e., the third module. The OMCC, formed by VirB7, VirB9, and the C-terminus of VirB10, is responsible for substrate translocation through the periplasm and the outer membrane (OM) (Gordon et al., 2017). Finally, the conjugative pilus, a structure essential for the direct contact between bacterial cells (Lawley et al., 2003), is composed of VirB2 and VirB5 proteins. Recent structural studies on the F-plasmid have provided new insights into the different types of channel structures and conjugative pili that can be present in donor bacteria (Hu et al., 2019a).

In Gram-positive bacteria, the OMCC is not needed and, consequently, T4ASS are minimized. Actually, in Gram-positive bacteria, the OMCC is composed of only six VirB/VirD4 proteins and seems to lack a pilus-like structure. Instead, Gram-positive bacteria use adhesins for cell-to-cell contact (Grohmann et al., 2017). Recently, it has been reported (González-Rivera et al., 2019) that the Gram-negative pKM101 plasmid can also use adhesin-like structures instead of a conjugative pilus, a mechanism until now observed only in Gram-positive bacteria.

Conjugative and Mobilizable Plasmids

Plasmids are autonomously replicating circular DNA molecules located in the bacterial cytoplasm. It has been estimated (Smillie et al., 2010) that a quarter of the plasmids are conjugative plasmids, another quarter corresponds to mobilizable plasmids and, finally, half of all plasmids are non-mobilizable plasmids. Together, conjugative and mobilizable plasmids constitute the so-called transmissible plasmids. In general, plasmids provide advantageous properties and capacities for bacterial survival such as, for instance, resistance to bactericidal and bacteriostatic antibiotics. Plasmids are classified according to Incompatibility Groups (Inc) so that two plasmids of the same group cannot coexist at the same time within the same bacterial cell. Some Inc groups (e.g., IncN, IncP, and IncW) can be found in many different bacterial species; others (e.g., IncF and IncI), in contrast, can only be stably kept in specific bacterial species (Kittell and Helinski, 1993). Conjugative and mobilizable plasmids contain all or some of the genes that take part in the conjugation process, respectively.

Conjugation-related genes are divided into two modules: MOB (mobility) and MPF (mating pair formation) genes (Smillie et al., 2010; **Figure 2**). Conjugative plasmids have both modules and, therefore, harbor genes that encode all the proteins needed for HGT. The MOB module, formerly called DNA Transfer and Replication (Bhatty et al., 2013), consists of three sequences necessary for substrate processing. First, the *oriT* (origin of transfer) sequence, which marks the starting point for DNA processing. Second, the sequence that encodes the relaxase which specifically recognizes the *oriT* sequence of its plasmid. Relaxase is the only conserved protein in all transmissible plasmids. The third protein encoded in the MOB module is T4CP, which

TABLE 1 | Classifications of T4SSs.

Criterion	Sub-families	Reference
Pilus type	F-like, P-type, I-type	Lawley et al., 2003
Relaxase phylogeny	MOB(F), MOB(H), MOB(Q), MOB(C), MOB(P), MOB(V)	Garcillán-Barcia et al., 2009
Substrate	Relaxase and ssDNA (conjugative systems), effector translocator systems, DNA release and uptake systems	Alvarez-Martinez and Christie, 2009
VirB4 phylogeny	MPF_I, MPF_C, MPF_G, MPF_T, MPF_F, MPF_B, MPF_{FATA}, MPF_{FA}	Guglielmini et al., 2013
Function	cT4SS, pT4SS	de Paz et al., 2005
Protein components	T4ASS, T4BSS	Christie et al., 2017; Grohmann et al., 2018

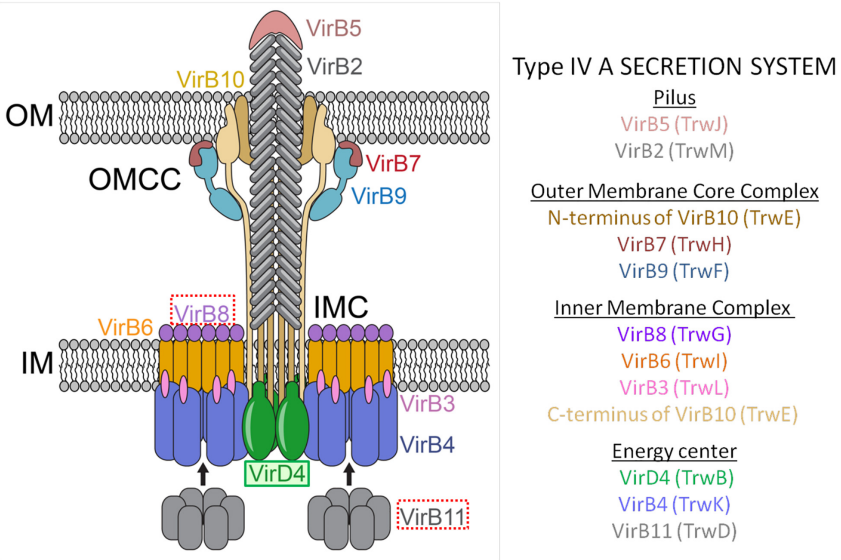


FIGURE 1 | Molecular architecture of the T4ASS showing the location of the T4CP. Different components of the pilus, OMCC, IMC, and the ATPase energy center are shown. To show equivalence between VirB/VirD4 and R388 systems, proteins corresponding to R388 are shown in brackets. Arrows represent dynamic interactions of VirB11 protein through the conjugative process. VirD4 is highlighted in green. Red boxes highlight the proteins selected as targets for drug discovery (VirB8 and VirB11). OMCC, outer membrane core complex; IMC, inner membrane complex. The schematic representation is based on the structure described by Redzej et al. (2017). The image has been adapted from Waksman (2019).

recognizes the relaxosome in the cytosol and connects it with the secretion channel located in the membrane. Likewise, accessory proteins that take part in the processing of the relaxosome can also be found in the MOB module. On the other hand, the MPF module encodes the proteins that build the T4SS (VirB1-11 in the case of T4ASS).

Mobilizable plasmids need a co-resident conjugative plasmid for them to be transferred because they lack the MPF module and contain only some of the sequences described above for the MOB module. In particular, all mobilizable plasmids contain the *oriT* and the relaxase which are specific to each plasmid. Some mobilizable plasmids also contain a T4CP. Mobilizable plasmids that encode their own T4CP form the MOBC1 plasmid subfamily, with plasmid CloDF13 as the prototype of this subfamily (Garcillán-Barcia et al., 2009).

Conjugation

Bacterial conjugation was initially described as a mechanism of sexual reproduction among bacteria (Lederberg and Tatum, 1946). This mechanism of HGT originally emerged in ancient proteobacteria and then expanded to all prokaryotes. Thanks to the new horizontally acquired genes, the recipient bacterium can acquire novel traits that allow it to survive in challenging environments such as, for instance, in the presence of antibiotics.

Bacterial conjugation requires direct physical contact between the two bacterial cells so that the ssDNA-protein complex can be transferred from the donor to the recipient bacterium. This process is usually divided into three steps: (i) the formation of the relaxosome; (ii) the reception of the substrate to the secretion channel; and (iii) the translocation of the substrate from the donor to the recipient bacterial cell (Alvarez-Martinez and Christie, 2009). However, recent studies have suggested additional steps: first, to provide the required cell-to-cell contact that initiates the transfer of the substrate, before the formation of the relaxosome, the secretion channel has to be established and the pilus formed. Also, during substrate translocation, the machinery has to switch from a protein transfer mode to a ssDNA transfer mode (Waksman, 2019). All these steps reflect the highly complex nature of the conjugative process, with many actions happening simultaneously.

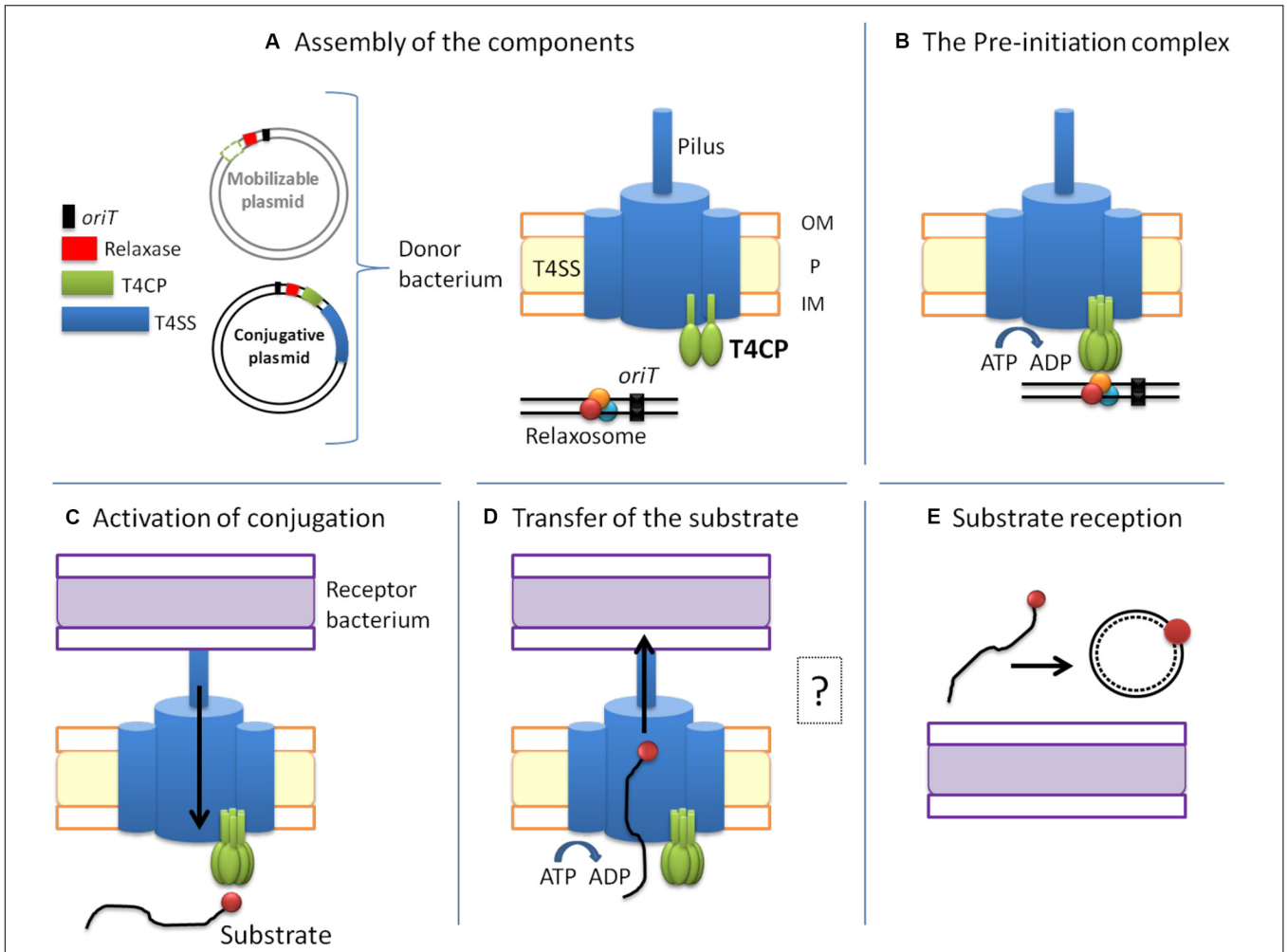


FIGURE 2 | The conjugative process. Conjugative plasmids encode all the elements necessary for conjugation, while mobilizable plasmids lack a cognate T4SS and most of them also lack the T4CP. The conjugative process can be divided into five main steps: **(A)** Assembly of the components. **(B)** Formation of the pre-initiation complex. The T4CP performs the reception of the relaxosome to the secretion channel. This interaction triggers the hexamerization of the T4CP, the ATP hydrolysis by T4CP and VirB11, and finally the activation of the secretion channel. Both conjugative and mobilizable plasmids employ the T4CP of a conjugative plasmid, except for the few mobilizable plasmids that encode their cognate T4CP. **(C)** Activation of the conjugative process and processing of the substrate. When the pilus recognizes a receptor bacterium, a signal is transmitted to the relaxosome and DNA is processed. Each relaxase accomplishes the processing of its cognate plasmid through the specific recognition of the *oriT* sequence. **(D)** Transfer of the substrate through the T4SS. This step is always accomplished *via* the T4SS encoded in the conjugative plasmid. The question mark indicates that many questions remain open. **(E)** Substrate reception and end of conjugation. Once the substrate enters the recipient bacterium, the ssDNA is recircularized and replicated. Color code: *oriT*, black; relaxase, red; T4CP, green; T4SS and pilus, blue; donor bacterium, orange; receptor bacterium, purple.

A high amount of energy is required for conjugation to occur, which explains the presence of different ATPases within the conjugative machinery. In order to perform the transfer of the substrate with the proper energy expenditure, the conjugative process is highly regulated. Although there are certainly still many unclear and unresolved aspects, in this review we have divided the conjugative process into five main steps (**Figure 2**) which apply to conjugation in Gram-negative bacteria (the conjugation process is slightly different in Gram-positive bacteria) (Grohmann et al., 2018):

A. Assembly of the T4SS channel, the conjugative pilus and the relaxosome. Many unknown signals trigger the formation of the T4SS secretion channel. In the canonical T4SS of *A. tumefaciens*, *virB* gene expression is activated through the sensing of an external factor from the milieu, which could also be the case for other systems (Li and Christie, 2018). For the establishment of the physical contact between donor and recipient cells, the pilus (a polymer of VirB2 subunits) is then assembled. Additionally, the relaxosome (composed of the relaxase, DNA and accessory proteins) has to be formed in the cytoplasm (**Figure 2A**).

B. Formation of the pre-initiation complex. Once the pilus, the T4SS and the relaxome are formed, they all interact to form a pre-initiation complex which stays latent until conjugation begins (Waksman, 2019). The relaxosome is recruited by the T4CP through a variety of interactions. Although relaxases are

known to possess translocation signals in their sequence for T4CP recognition (Christie, 2004), they are not the main responsible elements for the abovementioned interaction. Actually, the T4CP establishes the strongest interactions with the accessory proteins (Llosa and Alkorta, 2017). The structural basis of these interactions has not been resolved to date. Therefore, much further research is needed to better understand this mechanism (Li et al., 2019). It has been postulated (Waksman, 2019) that, as a consequence of those interactions, the T4CP either hexamerizes or creates heterohexamers with VirB4, which is attached to the secretion channel. The binding of DNA to the T4CP and the subsequent interaction with the T4SS appear to trigger the hydrolysis of ATP by the T4CP and the VirB11-like protein. These processes induce a conformational change in VirB10, thus activating the secretion channel (Cascales and Christie, 2004a; Cascales et al., 2013; **Figure 2B**).

C. Activation of the conjugative process. Although intracellular signals caused by DNA-binding and ATP-hydrolysis are necessary for substrate transfer, an extracellular signal is also required. This extracellular signal is generated when the pilus recognizes a recipient cell (Zechner et al., 2012). Once the physical contact is established, the pilus retracts (F-type) or falls (P-type), forming the junction between the two bacterial cells (Lawley et al., 2003; Schröder and Lanka, 2005). Moreover, the recognition signal is transduced through the T4SS to the cytoplasm, thanks to a series of interactions between VirB10-like proteins and T4CP (Llosa and Alkorta, 2017). As a consequence, the relaxase nicks one of the DNA strands and gets covalently attached to the 5′ end. Further, DNA replication takes place as a rolling circle replication, forming the DNA strand from the free 3′ end (Llosa et al., 2002; **Figure 2C**).

D. Substrate transport through the T4SS. The substrate is composed of the relaxase covalently bound to ssDNA. To transfer a complex of such different characteristics (i.e., protein and nucleic acid), the T4SS needs to undergo substantial changes during conjugation. Many questions remain unanswered on this matter, e.g., Which component is transferred first? How is the relaxase unfolded and transferred? According to the transfer DNA immunoprecipitation (TrIP) assay, the ssDNA is transferred from T4CP to VirB11 without energy cost. Then, the substrate is transferred to VirB6 and VirB8 proteins. At this moment, the ATPases need to be active to generate the required energy. Next, the substrate is transferred from VirB8 to VirB9 and, finally, to VirB2 in the pilus (Cascales and Christie, 2004b). However, the substrate pathway solved by the TrIP does not match the existing structural models. In consequence, this matter should be further explored according to new pieces of evidence (Waksman, 2019). Although the interactions between the substrate and the different proteins have already been defined and described, the translocation mechanism is not well established yet. Originally, one-step and two-step models were proposed. In the first model (one-step), the ssDNA substrate would be transferred directly to the T4SS whereas, in the second model (two-step), the ssDNA substrate would be transferred from

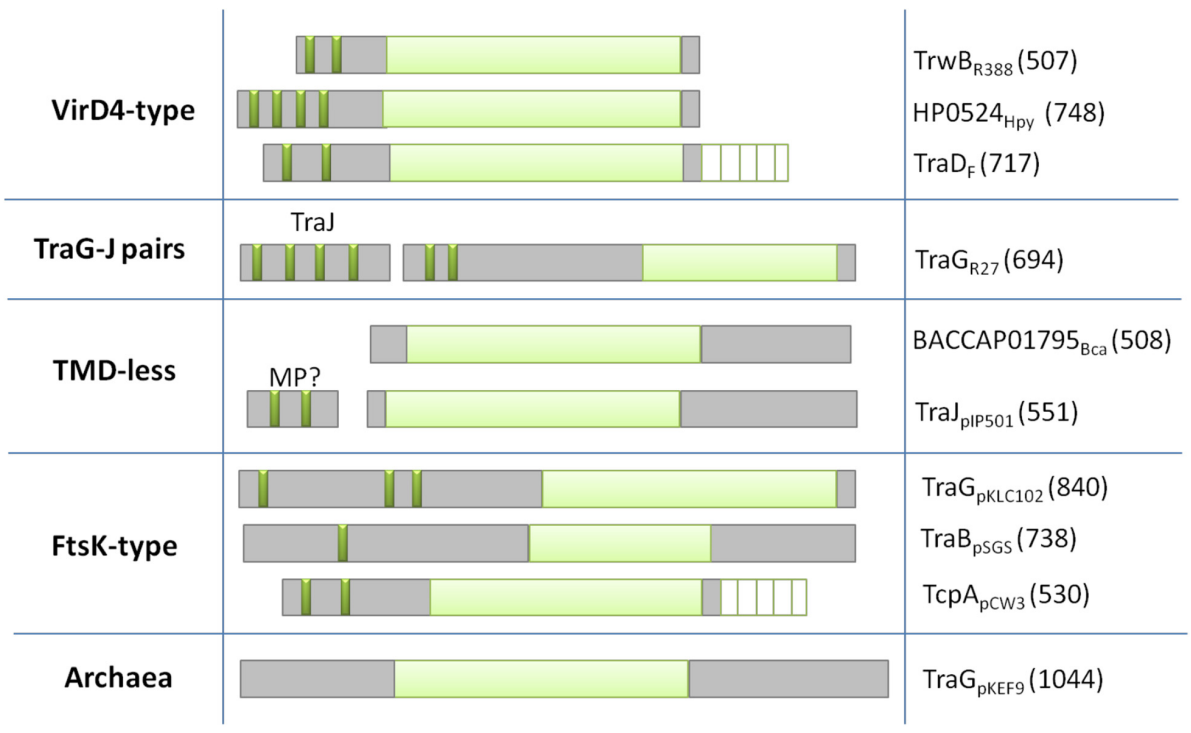


FIGURE 3 | Classification of T4CPs according to their molecular architecture. Each line represents one of the five subfamilies. Examples of each family are shown in the second (molecular architecture) and third (protein name and sequence length) column. Dark green, transmembrane helices; light green, nucleotide-binding domain; white lined squares, C-terminal domain. MP?, membrane protein postulated to interact with the TMD-less T4CP to perform its *in vivo* activity.

the T4CP channel to the periplasm, and from there to the T4SS (Llosa et al., 2002). In any case, this second model has been discarded due to the high content of nucleases present in the periplasm. Instead, it has been proposed (Christie, 2016) that the T4CP is placed at the entrance of the T4SS, translocating the substrate directly to the T4SS. Three possible translocation pathways have been proposed for the ssDNA substrate: (i) translocation could occur, through the channel of the T4CP hexamer, to the T4SS; (ii) after its uptake by the T4CP, the substrate could pass through the VirB4 hexamer and from there it would be translocated to the T4SS; or (iii) after its interaction with the ATPases, the substrate could be translocated directly through a channel formed by IMC subunits (Christie et al., 2017). In any case, many open questions remain to be answered about the conjugative process (**Figure 2D**).

E. Substrate reception and end of conjugation. Although substrate transport through the T4SS has been relatively well studied, it is currently not clear how the substrate crosses the outer and inner membrane of the recipient cell. Once in the recipient cell, the relaxase circularizes the lineal ssDNA, which is then used as substrate for the synthesis of the complementary strand (to obtain the dsDNA plasmid), a process in which the relaxase itself may also take part (Draper et al., 2005; **Figure 2E**).

Thanks to the conjugative process, the recipient cell benefits from the characteristics encoded in the acquired plasmid. In particular, if the recipient cell has acquired a conjugative plasmid, it gains the ability to transfer it to other bacteria, thus becoming a donor bacterial cell.

COUPLING PROTEINS

Type IV coupling proteins are essential for most T4SSs. As a matter of fact, T4CPs are (i) key players in all the conjugative systems; and (ii) present in many pathogenic processes that translocate effector proteins. As indicated above, T4CPs are (i) part of the MOB module in transmissible plasmids; and (ii) normally encoded in juxtaposition with relaxases. The phylogeny of T4CPs shows strong similarity with the phylogeny of relaxases, suggesting that substrates might have evolved along with their receptors. Concerning phylogeny, T4CPs are good markers of conjugative systems because of their ubiquity and size (Smillie et al., 2010). The most studied T4CPs are those associated with T4ASS of Gram-negative bacteria (e.g., VirD4$_{At}$, TrwB$_{R388}$, and TraD$_{F}$). Nonetheless, studies have been extended to T4CPs from Gram-positive bacteria (e.g., PcFC$_{pCF10}$ and TcpA$_{pCW3}$) (Parsons et al., 2007; Chen et al., 2008) and T4BSSs of Gram-negative bacteria, such as DotL$_{Lp}$ (Sutherland et al., 2012; Meir et al., 2018).

The study of T4CPs has a double purpose. On the one hand, taking into account their essential role in conjugation, they have become interesting potential targets for the development of conjugation inhibitors, as it is the case of other T4SS components (Shaffer et al., 2016). In this manner, the spread of ARGs could be controlled (similarly, in the case of pT4SSs, pathogenic processes could be prevented). On the other hand, the specific substrate transport carried out by T4CPs offers a wide range of possibilities in the field of biotechnology, including the targeted transport of desired molecules through heterologous systems (Guzmán-Herrador et al., 2017).

Classification

Type IV coupling proteins are a family of heterogeneous proteins that show lower sequence similarity than expected taking into account their phylogenetic proximity. The motif that defines the members of this family is the conserved Nucleotide-Binding Domain (NBD). Nonetheless, different domain architectures can be found among T4CPs. For instance, most T4CPs are membrane proteins that show a transmembrane domain (TMD) at the N-terminal end, as well as a cytoplasm-oriented all-alpha domain (AAD), but, in some cases, they can also present a variable domain in the C-terminal domain (CTD).

Based on the architecture and nature of these domains, T4CPs were classified into six subfamilies (Alvarez-Martinez and Christie, 2009) that were later reduced to five (Llosa and Alkorta, 2017; **Figure 3**).

VirD4-Type

This is the first described and most studied subfamily of T4CPs (Christie, 2016). Its name is given by the T4CP of the *A. tumefaciens* T-plasmid (i.e., VirD4$_{At}$). TrwB$_{R388}$, the prototype of the T4CP family, is included within this VirD4-type subfamily (Christie, 2004). The VirD4-type subfamily includes membrane proteins of 500–750 residues which show, at least, two transmembrane helices at the N-terminal end. They have a low sequence identity among themselves (around 15–20%). However, it has been observed (Cabezón et al., 1994; Llosa et al., 2003) that some of these proteins can be interchanged to perform the conjugation of mobilizable plasmids *in vivo*.

TraG-J Pairs

These are membrane proteins with a similar structure to VirD4-type T4CPs, a length of around 700 residues, and shorter periplasmic domains. They present higher sequence similarity among themselves and, like VirD4-type T4CPs, they can be functionally interchanged. They need to interact with a membrane protein, i.e., TraJ, encoded next to the T4CP gene, in order to perform their function (Gunton et al., 2007). TraJ-like proteins have a length of approximately 200 residues and up to five transmembrane helices.

T4CP Lacking TMD

The members of this subfamily do not have a TMD (this TMD is essential for the function of most T4CPs). However, some of these proteins interact with small membrane proteins, forming complexes similar to VirD4-type T4CPs. Two of the most studied proteins within this subfamily are TraJ$_{pIP501}$ (formerly called Orf10$_{pIP501}$) and TraI$_{pIP501}$ (formerly called Orf9$_{pIP501}$) (Grohmann et al., 2016).

FtsK-Type T4CP

It has been reported (Gomis-Rüth and Coll, 2006) that the structure of T4CPs shows certain similarities with the structure of translocases from ssDNA SpoIIIE/FtsK. This similarity is

increased in the case of FtsK-type T4CP. FtsK-type T4CPs have more than 750 residues and several TMDs separated from the NBD by large linker sequences. Although they do not show any AAD, they usually have an additional domain for ATP-binding and hydrolysis (Parsons et al., 2007).

Archaea T4CP

In the Archaea domain, some T4CPs have been identified by predicting sequences based on NBDs (Alvarez-Martinez and Christie, 2009).

Structure

As mentioned before, almost all T4CPs are composed of a TMD and a cytosolic domain. The cytosolic domain contains an NBD, an AAD and, in some cases, a CTD. The prototype of these T4CPs is $TrwB_{R388}$. The structure of its soluble mutant protein (i.e., TrwBΔN70) has been resolved by X-ray crystallography (Gomis-Rüth et al., 2001).

Domains

Transmembrane domain

For VirD4-type T4CPs, the TMD is essential for *in vivo* conjugation (Chen et al., 2008; Segura et al., 2013). This domain is involved in the interaction between the T4CP and VirB10 protein (Llosa et al., 2003; Segura et al., 2013), which is crucial for the activation of the T4SS. Besides, the TMD is critical for protein oligomerization as it contains key residues for such process (Redzej et al., 2017). The TMD is formed by α-helices and contains periplasmic loops essential for T4SS interactions and T4CP functions (Kumar and Das, 2002; Gunton et al., 2005).

The TMD of $TrwB_{R388}$ is the most extensively studied. Actually, it is the only T4CP that has been purified in its native version. In those studies, it has been demonstrated that (i) a progressive elimination of TMD elements leads to a reduction in the frequency of DNA transfer; and (ii) the mutant protein lacking the TMD (TrwBΔN70) has no DNA transfer activity (Segura et al., 2013). Likewise, the effect of the TMD on the protein structure was analyzed by infrared spectroscopy (Hormaeche et al., 2004; Vecino et al., 2011), observing that $TrwB_{R388}$ is more stable than TrwBΔN70 under denaturing conditions (i.e., low ionic force, presence of chaotropic agents, high temperature), thus suggesting that the TMD contributes to a more compact and ordered folding of the cytosolic domain. The TMD also has an effect on nucleotide-binding activity, even though this activity does not happen in this TMD. The wild-type protein presents a lower binding affinity to all nucleotides but a higher specificity to purine nucleotide triphosphates (Hormaeche et al., 2006). It seems that the presence of the TMD reduces the accessibility of the NBD, thus becoming more specific for ATP. Still, TrwBΔN70 has DNA-dependent ATPase activity, which is enhanced in the presence of $TrwA_{R388}$ (this activity has not been detected in the native protein) (Tato et al., 2007). All these results suggest that the TMD could have a regulatory function in the cytosolic domain. Finally, the native protein, as well as the TMD by itself, is located in the cell poles, whereas the soluble mutant protein, in the absence of other conjugative proteins, is found in the cytoplasm. By contrast, in the presence of T4SS proteins, TrwBΔN70 is evenly distributed along the cell periphery, suggesting that the TMD leads the polar location but interactions between the cytosolic domain and T4SS proteins contribute to the membrane location of $TrwB_{R388}$ (Segura et al., 2014).

Despite its crucial role, the TMD shows the most heterogeneous sequence among T4CPs. Although critical for the activity of VirD4-type T4CPs, T4CPs that lack a TMD are capable of accomplishing the transfer of their cognate plasmids. Moreover, the TMD-lacking mutant protein of FtsK-type $TcpA_{pcW3}$ T4CP, i.e., $TcpA_{\Delta 46-104}$, can certainly perform conjugation, although at a lower frequency than the wild-type protein (Parsons et al., 2007). Comparative studies between $TrwB_{R388}$ and TrwBΔN70 have shed light on the role of the TMD, which goes beyond being a simple anchor of $TrwB_{R388}$ to the membrane, something that may be also true for other T4CPs. The study of membrane proteins is a challenging task. In consequence, most studies on T4CPs have been performed using soluble mutant versions. Accordingly, more comparative studies using wild-type and soluble-mutant proteins from different members of the T4CP family are highly required.

Cytosolic domain

The cytosolic domain can be divided into two subdomains: NBD and AAD. Occasionally, a third domain, CTD, is also present.

Nucleotide-binding domain. The NBD, the most conserved domain, is shared by all T4CPs. This domain has two motifs. The first motif, i.e., P-loop or Walker A, binds β and γ phosphates of nucleotides (Walker et al., 1982). This P-loop motif, usually represented as G-x(4)-GK-(TS), where x can be any amino acid, is rich in glycine and has a conserved lysine followed by a serine or a threonine. The P-loop motif is present in many ATP- or GTP-binding proteins such as, for instance, the α and β subunits of ATP synthases, myosin, adenylate cyclase, and the ABC protein family (Geourjon et al., 2001). The second motif is called Walker B and its sequence is represented as (RK)xxxGxxx-LhhhDE, where x refers to any amino acid and h is a hydrophobic amino acid (Walker et al., 1982; Hanson and Whiteheart, 2005). Topologically, the NBD is located near the membrane while, structurally, it is formed by a twisted β-sheet surrounded by α-helices (Gomis-Rüth and Coll, 2006). This conformation resembles that of the β-domain of FtsK protein, a membrane protein motor that drives DNA across the membrane during cell division (Aussel et al., 2002). The NBD is essential for the *in vivo* activity of T4CPs. In fact, Walker A mutants cannot accomplish conjugation (Moncalián et al., 1999; Kumar and Das, 2002; Gunton et al., 2005), possibly due to the lack of ATPase activity (Tato et al., 2005). This lack of ATPase activity does not seem to affect other characteristics of the mutant proteins such as cell location, nucleotide-binding affinity, and interaction with the substrate (Moncalián et al., 1999; Chen et al., 2008). These findings suggest that T4CP function could take place after nucleotide- and substrate-binding events.

All-alpha domain

The AAD is oriented toward the cytoplasm and presents a variable size, sequence and folding pattern. This domain is also

present in many proteins from the RecA family (Gomis-Rüth et al., 2001). Furthermore, the AAD domain of $TrwB_{R388}$ presents structural similarities to the N-terminal domain of the XerD recombinase, a protein that interacts with the previously named FtsK resolving the DNA knots formed after replication. As the XerD's N-terminal domain is located over the NBD, it blocks DNA-binding and, therefore, a conformational rearrangement is needed to interact with the substrate. A similar situation could happen in T4CPs, where the AAD would block the NBD and, after a conformational change, could take part in DNA-binding (Gomis-Ruth et al., 2005). Besides, the six α-helices present in AAD of $TrwB_{R388}$ show similar topological characteristics to the DNA-binding domain of $TraM_F$, a component of the relaxosome of the F-plasmid that interacts with T4CPs. No homologous of $TraM_F$ has been found in the R388 plasmid, suggesting that $TrwB_{R388}$ has incorporated the structure of this interacting domain for DNA-binding. Whitaker et al. (2015) reported two essential roles of the AAD in bacterial conjugation: (i) it activates the secretion channel; and (ii) it specifically recognizes its cognate relaxosome and accessory proteins. Mobile genetic elements that lack a cognate T4CP must then overcome this specificity in order to be transferred through a heterologous T4SS.

C-terminal domain

Some T4CPs, such as VirD4-type (e.g., $VirD4_{At}$), FtsK-type (e.g., $TcpA_{pCW3}$), and Archaea-type (e.g., $TraG_{pKEF9}$), acquired a new CTD (Alvarez-Martinez and Christie, 2009). CTMs present high heterogeneity in terms of sequence and size, but all of them have glutamate- and aspartate-enriched regions.

The CTDs of different conjugative systems can present different functions. Many DNA-binding proteins show acidic C-terminal sequences that take part in the DNA-binding process (Lee and Thomas, 2000; Wang et al., 2007; Harrison et al., 2016; Basu et al., 2020), which appears to be their main function in T4CPs. For example, in the case of $TraD_F$ T4CP, the interaction of its CTD with one of the components of the relaxosome, $TraM_F$, is essential for recognition and plasmid transfer (Lu and Frost, 2005; Lu et al., 2008). In a similar manner, the CTD domain of $TcpA_{pCW3}$ triggers a threefold enhancement of conjugation frequency via its interaction with other T4SS proteins (Steen et al., 2009).

Some T4CPs are related with pathogenic systems that contain a CTD. On the one hand, Whitaker et al. (2016) reported that the relevance of the CTM in $VirD4_{At}$ depends on the substrate. These authors concluded that the CTD of VirD4-type T4CPs (related to pT4SS) evolved to (i) widen the effector repertoire; and (ii) improve the space-time regulation of the effector presentation. On the other hand, it has been described (Sutherland et al., 2012; Meir et al., 2018) that the acidic tail of protein $DotL_{Lp}$ from *L. pneumophila* is involved in the reception of the effector proteins from the IcmSW system.

Structure of $TrwB_{R388}$

Up to date, only two crystal structures related to T4CPs have been achieved: (i) the crystal structure of the TMD-less variant of $TrwB_{R388}$, TrwBΔN70 (Gomis-Rüth et al., 2001); and (ii) the C-terminal extension of $DotL_{Lp}$, as part of the Dot/Icm holocomplex from the T4BSS of *L. pneumophila* (Kwak et al., 2017). Since $TrwB_{R388}$ is the most studied T4CP, it is considered the archetypal of this protein family. The obtained structure is similar to $TrwK_{R388}$ structure (a homolog of VirB4 in the R388 system), which underlines the structural homology between VirD4 and VirB4-like proteins, despite their poor sequence similarity, and reinforces their phylogenetic relationship (Walldén et al., 2012; Guglielmini et al., 2013). In fact, the X-ray crystallographic structure of TrwBΔN70 (PDB ID 1E9R) was used to construct, by homology modeling, a 3D model for VirB4 (Middleton et al., 2005). These authors suggested that VirB4 may act as a docking and pumping agent at the entrance of the T4SS channel, in concert with other T4SS components, to transport substrates through the T4SS. Subsequent structural and mechanistic studies have demonstrated that T4CPs are located in between the VirB4 ATPases (Redzej et al., 2017).

TrwBΔN70 has been crystallized without ligands or in the presence of nucleotides and sulfate ions (Gomis-Rüth et al., 2001, 2002b). Its structure shows similarities with other protein families, such as AAA+ ATPases, RNA and DNA helicases, and proteins from the FtsK/SpoIIIE family. TrwBΔN70 has a homohexameric globular structure (110 Å in diameter and 90 Å in height) with an internal channel (ICH) that goes across the hexamer (a 22 Å wide in the membrane side and 7 Å wide in the cytosolic side). This narrowing from 22 to 7 Å in the cytosolic side suggests the existence of conformational changes to enable the substrate to enter into the transport system (Schröder and Lanka, 2005). By means of electrostatic potential calculations, it has been described (Larrea et al., 2017) that there are three electrostatic regions inside the channel: (i) a short hydrophobic region in the cytosolic side; (ii) a highly electronegative segment in the middle part of the channel; and (iii) an electropositive part in the second half of the channel. The functional relevance of the different amino acids has been revealed by the analysis of mutants, concluding that amino acids at both ends of the ICH regulate the opening of the channel and that amino acids in the middle part of the channel define the translocation pathway of the substrate (Larrea et al., 2017).

Regarding the primary structure of $TrwB_{R388}$ (**Figure 4A**), the first amino acids (M_1 to K_8) are related to a small cytosolic domain, followed by the TMD (V_9 to V_{69}). According to its sequence, this region is formed by two α-helices linked by a loop that is proposed to be located in the periplasmic region, perhaps interacting with the T4SS and with periplasmic regions of other subunits of the hexamer. Walker A and Walker B domains are located from L_{71} to K_{183} and from D_{298} to I_{507} amino acids, respectively. The AAD is located between both Walker domains (G_{184}-G_{297}) (**Figure 4A**). The region between R_{437}-T_{444} and T_{450}-E_{455} amino acids, related to the membrane-proximal edge of β9 and β10 strands, has not been solved in the soluble mutant protein structure, probably due to its interaction with the excised TMD (**Figure 4B**). Modeling of this region rendered β-sheets that, together with the α-helices of the TMD, could create the right channel for substrate transport through the membrane (**Figure 4C**).

Each $TrwB_{R388}$ monomer looks like an orange slice and can be divided into the NBD and AAD domains (**Figure 4B**).

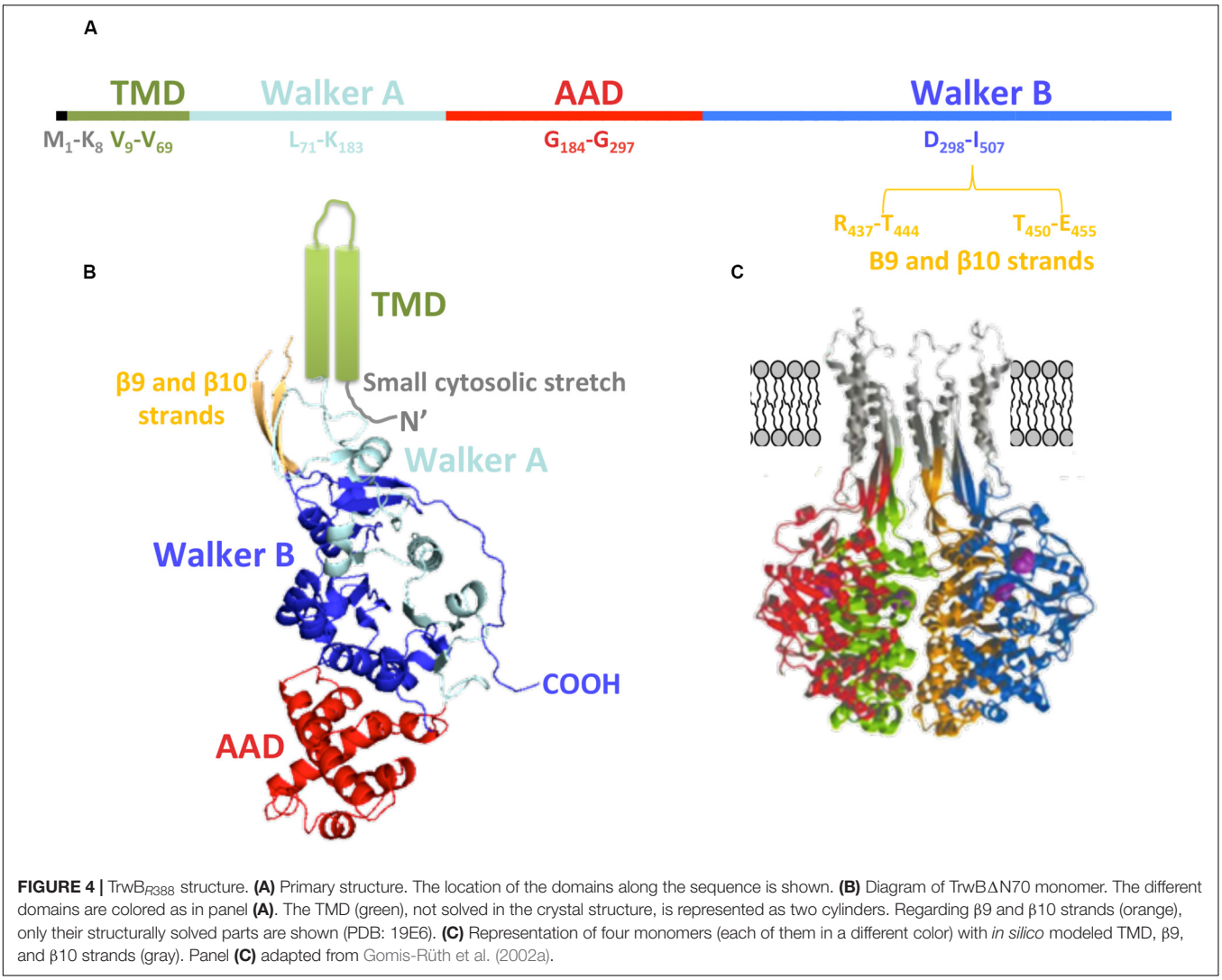


FIGURE 4 | TrwB$_{R388}$ structure. **(A)** Primary structure. The location of the domains along the sequence is shown. **(B)** Diagram of TrwBΔN70 monomer. The different domains are colored as in panel **(A)**. The TMD (green), not solved in the crystal structure, is represented as two cylinders. Regarding β9 and β10 strands (orange), only their structurally solved parts are shown (PDB: 19E6). **(C)** Representation of four monomers (each of them in a different color) with *in silico* modeled TMD, β9, and β10 strands (gray). Panel **(C)** adapted from Gomis-Rüth et al. (2002a).

Nucleotide-binding pockets are located in the intermonomeric contact surface (these contact sites represent 25% of the total surface). Therefore, in the hexameric conformation, 50% of the surface of each subunit is in contact with neighboring subunits. The crystal structures solved in the presence of nucleotides (Gomis-Rüth et al., 2002b) show conformational movements in the active site which extend through the ICH. Then, nucleotide binding or hydrolysis can trigger conformational changes in the T4CP, thus facilitating interactions with the substrate (Gomis-Rüth et al., 2002b).

Negative-stain electron microscopy images (NS-EM) (Hormaeche et al., 2002) of the complete TrwB$_{R388}$ reflect a strong structural similarity to TrwBΔN70, as well as a 25 Å-sized appendix that corresponds to the TMD (**Figure 5A**). Similarly, Redzej et al. (2017) studied TrwB$_{R388}$, along with the T4SS$_{R388}$ complex, by NS-EM, observing that two TrwB$_{R388}$ dimers are localized between two TrwK$_{R388}$ hexamers. It must be taken into account that fixed images provided by Ns-EM cannot properly represent all the steps involved in the conjugative process, where interactions between substrates and T4SSs are highly dynamic and complex. It is possible that the abovementioned NS-EM images represent only one of the steps of the transfer process. In this regard, some authors (Li et al., 2019) have claimed that the hexameric structure solved for TrwBΔN70 does not correspond to the real *in vivo* oligomeric and functional state.

Oligomerization

The functional structure of T4CPs is postulated to be a homohexamer that presents wide contact regions in the cytoplasmic part of the monomers. However, the results presented by Redzej et al. (2017) suggest that the *in vivo* functional state might be dynamic in the sense of possible switching from dimers to hexamers (**Figure 5B**). Also, from combined electron microscopy and docking calculations, it has been postulated (Peña et al., 2012; Waksman, 2019) that its active form could consist of heterohexamers with VirB4-like proteins. This interesting hypothesis requires further exploration but opens up a road for computational modeling and biophysical studies regarding the possibility of (hetero)-dimerization interfaces as crucial sites for the inhibition of

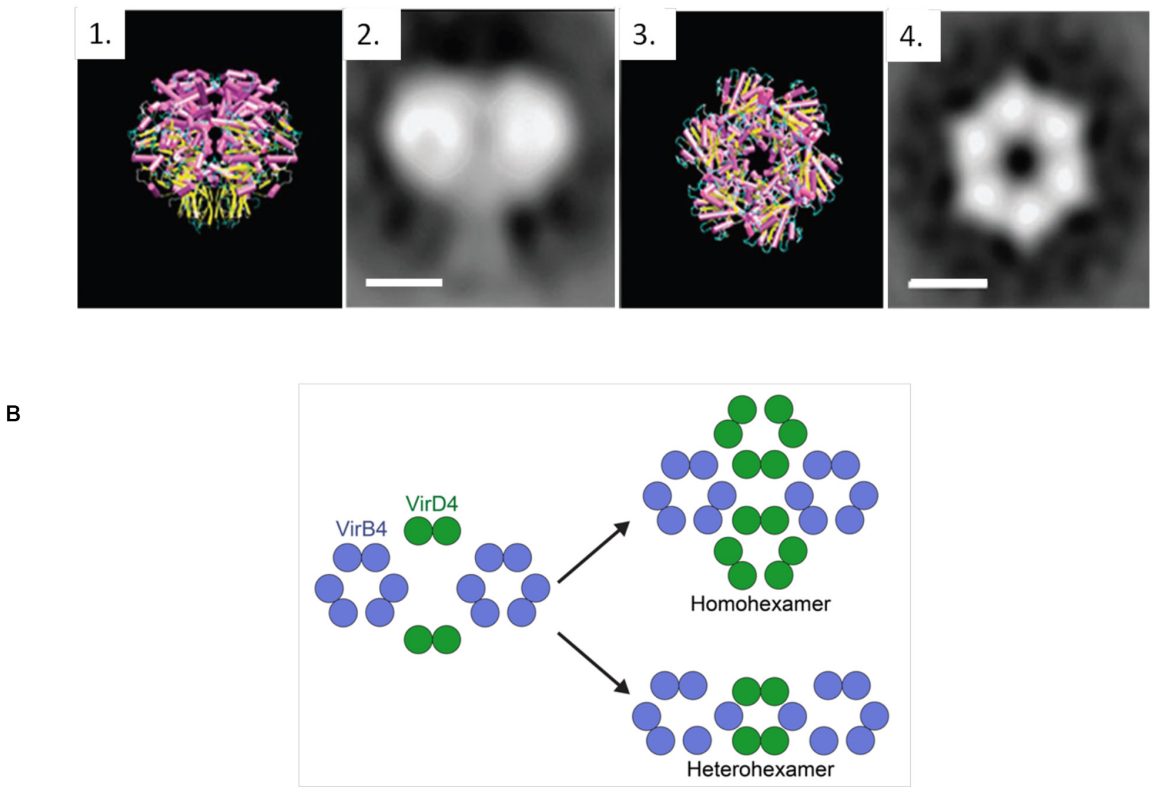


FIGURE 5 | (A) TrwBΔN70 and $TrwB_{R388}$ hexamers. 1 and 3 represent the side and top view of TrwBΔN70, respectively, solved by X-ray diffraction. 2 and 4 represent the side and top view of $TrwB_{R388}$, respectively, obtained by electron microscopy. Scale bar: 50 Å. Figure adapted from Hormaeche et al. (2002). **(B)** Representation of the possible functional oligomeric states of T4CPs. It has been described that T4CPs interact with the T4SS as dimers (Redzej et al., 2017). As their functional structure is postulated to be a hexamer, they could perform their role as (i) homohexamers, composed of six T4CP subunits; or (ii) heterohexamers, composed of four T4CP subunits and two VirB4 subunits.

conjugation. Additionally, through electron cryotomography, a technique most useful for the analysis of secretion systems (Oikonomou and Jensen, 2019), density images of T4CPs from Dot/IcmT4SS and CagT4SS ($DotL_{Lp}$ and $Cag\beta_{Hp}$, respectively) have been obtained (Chetrit et al., 2018; Hu et al., 2019b; Li et al., 2019). The discrepancies shown by the *in vivo* oligomeric states of these T4CPs underline the need for further studies on this matter.

Nevertheless, the requirement of the TMD for oligomerization has been highlighted in different studies (Hormaeche et al., 2002; Schroder and Lanka, 2003; Mihajlovic et al., 2009). Precisely, it seems that interactions in the TMD are held by key amino acids located in the second helix of the TMD (Segura et al., 2013), specifically the GXXXXG motif which is related to the oligomerization of many membrane proteins (Teese and Langosch, 2015). In fact, the mutant protein TrwBΔN70 cannot be purified as a hexamer and only *in vitro* assays in which TrwBΔN70 interacts with G-quadruplex DNA (G4-DNA) show oligomers (Matilla et al., 2010). Other studies (Haft et al., 2007) indicate that interaction with DNA is not necessary for *in vivo* oligomerization of $TraD_F$. According to this $TraD_F$ oligomerization model, the T4CP enters the membrane as a monomer and then establishes contact with the TMD of another monomer. Thus, homodimers are formed by CTM-mediated interactions. Finally, accessory proteins encoded in the F-plasmid recognize N-terminal sequences and favor the creation of trimers of dimers, thus obtaining membrane-stabilized hexamers. In the absence of the TMD, the cytosolic domains of $TraD_F$ monomers do not interact, highlighting the relevance of the TMD for oligomerization. In any case, the oligomerization mechanism of T4CPs lacking TMD, such as $TraJ_{pIP501}$, is unknown. Despite some evidence for their dimerization, the hexameric form of $TraJ_{pIP501}$ has not been observed so far (Abajy et al., 2007).

Molecular Interactions

Type IV coupling proteins interact with both T4SS-proteins and the substrate to be transferred, creating a complex interaction network which is thought to be highly dynamic during the conjugative process itself, possibly owing to the presence of different signals (Christie, 2016; Redzej et al., 2017). As an example, it is a well-known fact that the cytosolic domain interacts with the substrate and/or the accessory proteins of the relaxosome. Interactions with the relaxase are indeed specific, but

less tight than those with the accessory proteins of the relaxosome or molecular chaperones (Sastre et al., 1998; Hamilton et al., 2000; Llosa et al., 2003). Additionally, highly regulated interactions between the cytosolic domain of T4CPs and DNA have been described both *in vivo* and *in vitro*. These interactions are independent of the DNA sequence (Moncalián et al., 1999; Abajy et al., 2007). *In vivo* assays suggest that they depend on DNA processing and MOB protein activity (Cascales and Christie, 2004a; Chen et al., 2008). Moreover, the DNA structure seems a relevant factor for interaction specificity, as TrwBΔN70 shows higher affinity for G4-DNA than for DNA molecules lacking this particular structure (Matilla et al., 2010). On the other hand, interactions between $TraG_{RP4}$ and DNA can be inhibited *in vivo* with ATP, ADP, and Mg^{2+} (Schroder and Lanka, 2003). T4CPs of T4BSS have also shown binding to DNA, although they are only implicated in effector protein transport (Schroder et al., 2002).

In addition, TMD and cytosolic domains are involved in interactions with the T4SS. The cytosolic side of T4CPs interacts with other ATPases that form the energy center (i.e., VirB4 and VirB11-like proteins) (Llosa and Alkorta, 2017) and TMD interacts with VirB10-like proteins (Segura et al., 2013). This non-specific interaction occurs between the periplasmic loops (Llosa et al., 2003; Gunton et al., 2005). Moreover, the T4CP can interact with heterologous T4SSs, creating an opportunity to form chimeric conjugation or effector-translocator systems (de Paz et al., 2005). Finally, in Gram-negative bacteria, T4CPs can interact with peptidoglycan hydrolases (Alvarez-Martinez and Christie, 2009).

Function

There is no consensus regarding the possible roles of the T4CP during conjugation. When the crystal structure of $TrwB_{R388}$ was first solved, it was proposed (Gomis-Rüth et al., 2001) that, apart from having a role in the coupling between the substrate and the T4SS, it could also function as a molecular motor. Nevertheless, since then, four additional roles have been proposed highlighting the importance of T4CPs for conjugation: (i) substrate reception to the T4SS; (ii) energy source for conjugation; (iii) signal transduction; and (iv) substrate translocation. All domains appear to be relevant; actually, the lack of any domain or the presence of mutations have important effects on conjugation frequency (De Paz et al., 2010; Cascales et al., 2013; Whitaker et al., 2015). Specifically, mutants lacking the TMD or NBD cannot accomplish conjugation, underlining the relevance of the TMD (i.e., subcellular location, regulation, interactions with the T4SS) and the ATPase activity (Moncalián et al., 1999; Kumar and Das, 2002; Gunton et al., 2005; Chen et al., 2008; Lang and Zechner, 2012). Indeed, the ATPase activity of the T4CP is essential for conjugation (Cascales et al., 2013). What is more, as mentioned before, punctual mutations on the NBD can inhibit conjugation (Tato et al., 2005). However, so far, this ATPase activity has not been described in the native protein, probably because of the highly complex regulatory network existing in the bacterial cell which can include accessory proteins, interactions with lipids, and unknown signal-mediated conformational changes. Along these lines, it has been reported (Cascales et al., 2013) that T4CPs take part in the complex signaling network happening during conjugation through their interactions with the relaxosome and the channel subunits.

Pertaining to the essential role of the T4CP in substrate transfer, it has been demonstrated (Draper et al., 2005; De Paz et al., 2010) that, in the absence of DNA transfer, the ATPase activity of the T4CP is required for the transport of the relaxase. On the other hand, TrIP studies have demonstrated that, upon the activation of conjugation, the T4CP is the first protein to interact with the ssDNA substrate prior to its delivery to VirB11 (Cascales and Christie, 2004b).

Finally, considering the structural homology of T4CPs with other DNA transfer motors, as well as their interactions with the substrate, it has been suggested (Cabezon and de la Cruz, 2006) that T4CPs could act as motors that pump the ssDNA through its ICH, taking advantage of their ATPase activity, Nonetheless, no key amino acids have been clearly identified to perform such ssDNA transport (Larrea et al., 2017).

INHIBITION OF T4CPs TO CONTROL ANTIBIOTIC RESISTANCE DISSEMINATION

Inhibition of Conjugation

Bacterial conjugation is the main process by which bacteria acquire (i) resistance to antibiotics; and (ii) the ability to transfer such resistance to other bacteria, thus contributing to the spread of antibiotic resistance. Therefore, a possible strategy to control the spread of antibiotic resistance is to inhibit the activity of conjugative proteins. However, inhibition of bacterial conjugation has received little attention as the search for new antibiotics has captured most of the scientific interest (Kwapong et al., 2019).

Since the T4SS has, at least, 12 different proteins, the possibilities for finding therapeutic targets (inhibitors of conjugative proteins and, hence, conjugation itself) are broad. A detailed biochemical and biophysical study of the key elements in the conjugative process is required to be able to properly assess the potential of such strategy. In the last decade, a lot of research (see below) has been conducted in the search for specific inhibitors of key conjugative proteins, without interfering with other processes present in human cells or the environment. In particular, the search for conjugation inhibitors has focused on two key proteins: (i) the VirB8-like proteins and (ii) VirB11-like proteins (one of the ATPases of the T4SS).

VirB8-like proteins are essential assembly factors that play a key role in protein-protein interactions in all T4SS. They are needed for substrate translocation from the cytoplasm to the recipient bacterium. Therefore, their specific inhibition could control the dissemination of antibiotic resistance in a targeted way (Baron, 2006). Based on structural data of $TraE_{pKM101}$ (the VirB8-like protein of plasmid pKM101), the surface groove located between the α-helical and the β-sheet faces was used as docking site to identify and then develop small molecules capable of inhibiting the dimerization of $TraE_{pKM101}$ and the conjugative transfer of plasmid pKM101 (Casu et al., 2016, 2017).

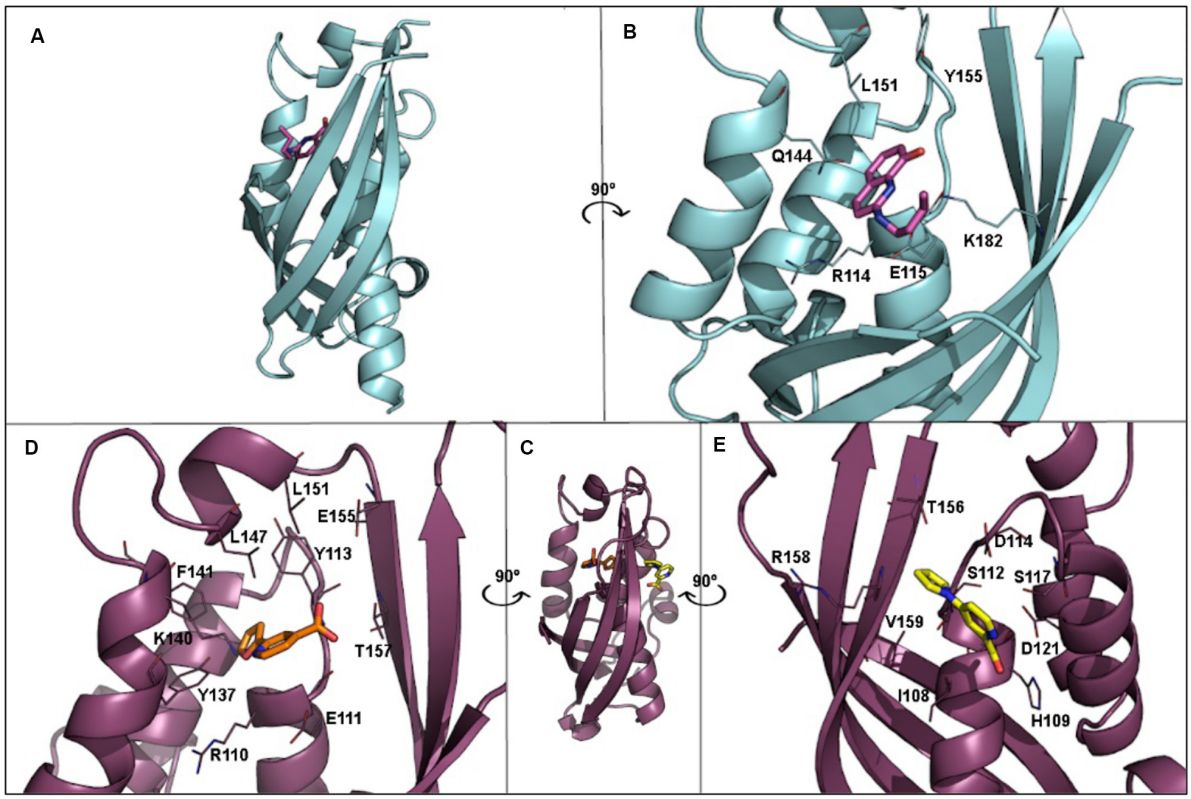


FIGURE 6 | (A) 3D structure of VirB8 from *Brucella suis* (cyan) in complex with inhibitor B8I-1 (2-(butylamino)-8-quinolinol, magenta) (PDB ID 4AKY). **(B)** Detail of key residues at the binding site from PDB ID 4AKY. **(C)** 3D structure of $TraE_{pKM101}$ from *Escherichia coli* (plum) in complex with inhibitor 239852 (orange) bound at the groove (PDB ID 5WIP) and superimposed to the 3D structure of *E. coli* $TraE_{pKM101}$ (also in plum) in complex with inhibitor 105055 (yellow) bound to the α-helical region (PDB ID 5WIO). **(D)** Detail of key residues at the binding site from PDB ID 5WIP. **(E)** Detail of key residues at the binding site from PDB ID 5WIO.

This work proved the possibility of finding conjugation inhibitors by using structural information of T4SS proteins. In any event, it is still highly necessary to identify molecules that can inhibit most members of this VirB8-like protein family, regardless of their particular conjugative system.

Regarding VirB11-like proteins, two members of this family have been considered as drug targets in the search for specific conjugation inhibitors: Cagα and $TrwD_{R388}$. Compounds that inhibit both the *in vitro* ATPase activity of Cagα (an essential protein for *Helicobacter pylori* virulence) and the transfer of the CagA effector protein have been identified (Hilleringmann et al., 2006). Additionally, based on virtual high throughput screening (HTS), a series of 8-amino imidazo(1,2-a)pyrazine derivatives that inhibit the ATPase of Cagα have been identified (Sayer et al., 2014). Unfortunately, in both studies (Hilleringmann et al., 2006; Sayer et al., 2014), there are concerns about the specificity of these compounds since they may also inhibit other bacterial or mammalian ATPases. Alternatively, an approach that does not specifically target Cagα ATPase activity has been tested (Arya et al., 2019), finding molecules that bind and stabilize Cagα, some of which inhibited Cagα ATPase activity and also dissociated Cagα hexamers.

Unsaturated and alkynoic fatty acid derivatives (Fernandez-Lopez et al., 2005; Getino et al., 2015) have also been found to effectively inhibit bacterial conjugation. Similarly, tanzawaic acids have potential as conjugation inhibitors (Getino et al., 2016). In particular, $TrwD_{R388}$, the VirB11 analog of R388 plasmid, has been the target of the abovementioned fatty acids and some derivatives, such as 2-bromopalmitic acid (Ripoll-Rozada et al., 2016; García-Cazorla et al., 2018).

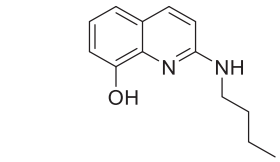
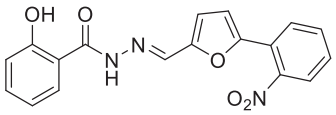
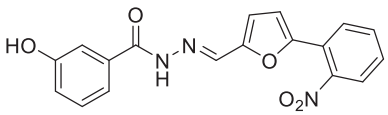
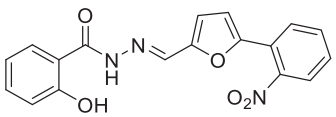
All these studies support the possibility of using specific inhibitors of key conjugative proteins in order to control the spread of antibiotic resistance among bacteria. Considering the well-known current magnitude of the antibiotic resistance problem, we must urgently test the widest possible repertoire of therapeutic targets in the search for specific inhibitors of bacterial conjugation. In that line of action, here we have described (i) the key role played by T4CPs; (ii) the different members of this family; and (iii) the functional and structural characteristics of those members (most relevantly, $TrwB_{R388}$) on which there is information.

Computational Approaches for the Identification and Design of T4SS Protein Inhibitors

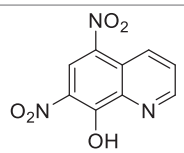
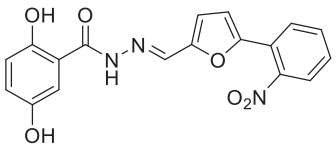
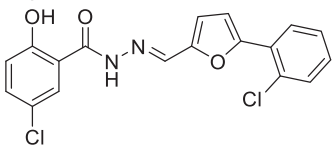
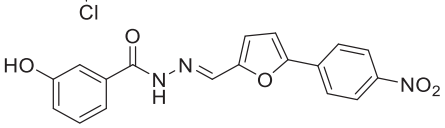
Computational approaches applied to tackle biological problems encompass multiple techniques, such as molecular docking,

TABLE 2 | Chemical structure of ligands reported as T4SS protein inhibitors.

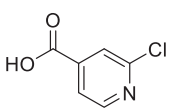
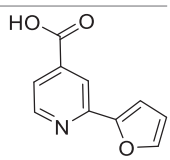
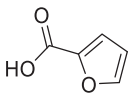
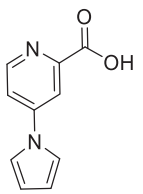
Target: VirB8 (*Brucella suis*) (Smith et al., 2012)

Compound	Structure
B8I-1	
B8I-2	
UM0125823	
UM0125848	

Target: TraE$_{pKM101}$ (Casu et al., 2016)

Compound	Structure
B8I-16	
BAR-072	
BAR-073	
UM-024	

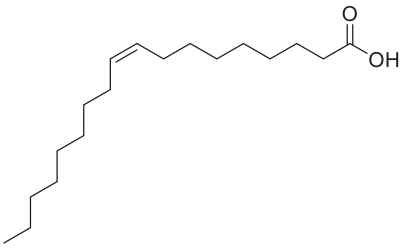
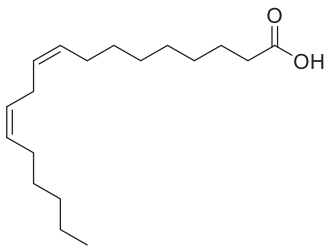
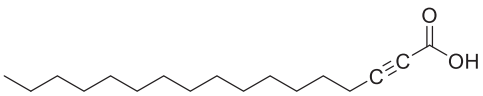
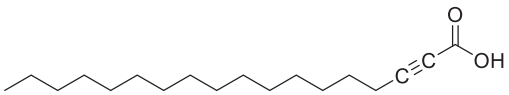
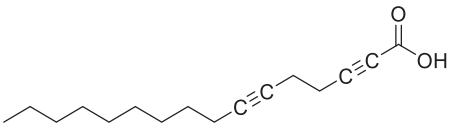
Target: TraE$_{pKM101}$ (Casu et al., 2017)

Compound	Structure	Compound	Structure
4H10		239852	
1E6		105055	

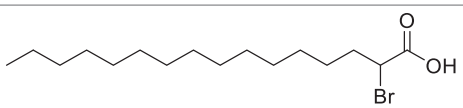
(Continued)

TABLE 2 | Continued

Target: TrwD (*Escherichia coli*) (Ripoll-Rozada et al., 2016)

Compound[a]	Structure
Oleic acid	
Linoleic acid	
2-HAD	
2-ODA	
2,6-HDA	

Target: TrwD (*Escherichia coli*) (García-Cazorla et al., 2018)

Compound[a]	Structure
2-BP	

[a]*IUPAC names. oleic acid, (Z)-octadec-9-enoic acid; linoleic acid, (9Z,12Z)-octadeca-9,12-dienoic acid; 2-HDA, hexadec-2-ynoic acid; 2-ODA, octadec-2-ynoic acid; 2,6-HDA, hexadeca-2,6-diynoic acid; 2-BP, 2-bromo-hexadecanoic acid.*

virtual screening, and molecular dynamics simulations, among others (Martín-Santamaría, 2018). These techniques have proven to be very useful for the study of molecular recognition processes, as well as for the identification and design of a vast variety of compounds with biological activity (Leelananda and Lindert, 2016). Computational methods are especially helpful in the early stages of the search for novel modulators, e.g., prior to HTS, but also to understand the functioning of biological macromolecules, thus providing clues for the identification and design of novel molecules with potential as therapeutic agents.

T4SS proteins have been studied by computational methods in order to (i) better understand their mechanism of action; and (ii) obtain inhibitors to modulate their functions. In this respect, salicylidene acylhydrazide-type compounds were

described (Smith et al., 2012) as inhibitors of *Brucella suis* VirB8 upon binding to a pocket away from the dimerization interface. The binding of these compounds prevented dimerization in both bacterial two-hybrid (BTH) assays and fluorescence resonance energy transfer-based *in vitro* assays, suggesting an allosteric mechanism of inhibition (Smith et al., 2012). The identification of the binding pocket was assisted by molecular docking of ligands B8I-1 to B8I-20 (**Figure 6** and **Table 2**). The binding takes place in the pocket formed by residues R_{114}, E_{115}, Q_{144}, L_{151}, Y_{155}, and K_{182} (**Figures 6A,B**). Based on these results, B8I-2 derivatives were synthesized and two of them (UM0125823 and UM0125848, **Table 2**) were found to be active against VirB8 dimerization. Docking prediction for these derivatives was in agreement with experimental data. Some of these compounds (B8I-1 to B8I-16 and UM-023 to UM-050) were later tested against $TraE_{pKM101}$, along with other novel BI8-2 derivatives (Casu et al., 2016). The binding site (mainly, R_{110}, E_{111}, K_{140}, and T_{157} residues; **Figures 6C,D**), a cavity homologous to the groove present in VirB8, was identified by means of molecular docking. The predicted binding for these compounds agreed with *in vitro* and *in vivo* experimental data: B8I-16, BAR-072, BAR-073, and UM-024 (**Table 2**) showed reduced TraE dimerization and pKM101 transfer.

Other VirB8-type protein inhibitors have been observed to inhibit $TraE_{pKM101}$ upon fragment-based screening using a differential scanning fluorimetry assay. Possible binding sites for the fragments were identified by means of molecular docking and X-ray crystallography (Casu et al., 2017). Docking calculations showed that many fragments bind to the surface groove previously reported (Smith et al., 2012; Casu et al., 2016), but also to a novel binding site located at the α-helical region (H_{109}, S_{112}, D_{121}, and V_{159} residues; **Figure 6E**) near to the dimerization interface. For two of the fragments, 4H10 and 1E6 (**Table 2**), X-ray structures, in complex with TraE, were obtained (PDB ID 5WII and 5WIC), showing that 1E6 binds to both binding sites while 4H10 binds near to the groove. Based on these fragments, derivatives 105055 and 239852 (**Table 2**) were found and docked, predicting improved affinity. The prediction was confirmed by X-ray crystallography (PDB ID 5WIO, in complex with 105055, and 5WIP, with 239852), which showed binding to the surface groove for 239852 (**Figure 6D**) and to the α-helical region of TraE for 105055 (**Figure 6E**). Further biological assays confirmed the inhibition of conjugative transfer of pKM101.

Some fatty acids, e.g., linoleic and oleic acid (Fernandez-Lopez et al., 2005), and 2-alkynoic fatty acid derivatives (**Table 2**), such as 2-hexadecynoic acid (2-HDA), 2-octadecynoic acid (2-ODA) and 2,6-hexadecadiynoic acid (2,6-HDA) (Getino et al., 2015), have been reported to inhibit bacterial conjugation. These fatty acids have been identified to target the VirB11-like protein $TrwD_{R388}$ of *Escherichia coli* (Ripoll-Rozada et al., 2016), inhibiting TrwD function without competing with its endogenous substrates. To understand the inhibition mechanisms of these fatty acids, a homology model of TrwD was built with using *B. suis* VirB11 as a template (Ripoll-Rozada et al., 2016). Subsequent docking studies were performed, showing that the fatty acids could bind the N-terminal domain (residues 37–54) and the linker region (residues 118–125), thus affecting the plasticity of the N-terminal domain and, hence, avoiding the conformational changes needed for ATPase activity. The docking predictions also reported the same binding site for palmitic acid. Further studies (García-Cazorla et al., 2018) have shown that the fatty acids abovementioned as TrwD inhibitors, as well as a novel inhibitor, 2-bromopalmitic acid (2-BP, **Table 2**), can be incorporated to the bacterial membrane and affect TrwD binding upon competition with palmitic acid. Docking studies for 2-BP showed the same binding site than that for linoleic, oleic and palmitic acid, but with a different binding mode.

CONCLUDING REMARKS

Bacterial conjugation, the major mechanism responsible for the acquisition of new genes by bacteria, is one of the main processes responsible for the dissemination of antibiotic resistance among bacteria. Conjugation is performed by T4SSs, which provide substrate processing and its transfer to the recipient bacterium through a macromolecular complex composed of at least 12 different proteins. The T4CP, one of the essential proteins of this process, is present in all conjugative systems. Despite their importance, few T4CPs have been studied in detail, being $TrwB_{R388}$ the structural and functional paradigm of the family. Along with $TrwB_{R388}$, in the scientific literature, there are data on the structure, oligomerization, function and interactions with other T4SS proteins for a few members of the T4CP family, which have been described here.

The worrying increase in the emergence and dissemination of antibiotic multi-resistant bacteria, together with the loss of effectiveness of current antibiotics, highlights the great urgency to find effective strategies to control antibiotic resistance spread. The search for inhibitors of key conjugative proteins, using *in vitro* HTS and computational approaches, has proved to be extremely useful to identify potential inhibitors of bacterial conjugation. This review aims to compile the progress made so far in the knowledge of T4CPs, from a structural and molecular viewpoint, as well as the work performed to identify and design conjugation inhibitors through the specific inhibition of key conjugative proteins. The combination of computational approaches with the existing body of knowledge on T4CPs will facilitate the development of conjugation inhibitors targeting T4CPs, with the ultimate goal of selectively controlling the spread of antibiotic resistance. To this purpose, further biophysical and computational studies on T4CP function, interactions between T4CPs and T4SS proteins, and dimerization/oligomerization among T4CP homologues are required.

AUTHOR CONTRIBUTIONS

IÁ-R, SM-S, CG, and IA: design of the work (text and figures) and literature review, writing and revision of the manuscript, and approval of the last version. EG-R, and SM-S: literature review, writing and design of figures and tables, and approval of the last version. BU-U and LA: writing and design of figures, revision of the manuscript, and approval of the last version. All authors contributed to the article and approved the submitted version.

FUNDING

This work has been supported by the University of the Basque Country (GIU18/229 and COLAB19/08), the Department of Economic Development and Infrastructures of the Basque Government (ELKARTEK 2020 KK-2020/00007), and the Spanish Ministry of Science and Innovation (CTQ2017-88353-R and PID2019-110055RB-C21). IÁ-R was a pre-doctoral student supported by the Basque Government. Part of this work was done by IA in a sabbatical leave at the CIB-CSIC with support from the Ikerbasque Foundation.

ACKNOWLEDGMENTS

We sincerely thank Jon Ander Nieto-Garai for his help with the design of the figures. We also thank Euskampus Fundazioa for support of the JRL on Environmental Antibiotic Resistance.

REFERENCES

Abajy, M. Y., Kopeć, J., Schiwon, K., Burzynski, M., Döring, M., Bohn, C., et al. (2007). A type IV-secretion-like system is required for conjugative DNA transport of broad-host-range plasmid pIP501 in gram-positive bacteria. *J. Bacteriol.* 189, 2487–2496. doi: 10.1128/jb.01491-06

Águila-Arcos, S., Álvarez-Rodríguez, I., Garaiyurrebaso, O., Garbisu, C., Grohmann, E., and Alkorta, I. (2017). Biofilm-forming clinical Staphylococcus isolates harbour horizontal transfer and antibiotic resistance genes. *Front. Microbiol.* 8:2018. doi: 10.3389/fmicb.2017.02018

Alvarez-Martinez, C. E., and Christie, P. J. (2009). Biological diversity of prokaryotic type IV secretion systems. *Microbiol. Mol. Biol. Rev.* 73, 775–808. doi: 10.1128/mmbr.00023-09

Arya, T., Oudouhou, F., Casu, B., Bessette, B., Sygusch, J., and Baron, C. (2019). Fragment-based screening identifies inhibitors of ATPase activity and of hexamer formation of Cagα from the *Helicobacter* pylori type IV secretion system. *Sci. Rep.* 9:6474.

Aussel, L., Barre, F. X., Aroyo, M., Stasiak, A., Stasiak, A. Z., and Sherratt, D. (2002). FtsK is a DNA motor protein that activates chromosome dimer resolution by switching the catalytic state of the XerC and XerD recombinases. *Cell* 108, 195–205. doi: 10.1016/s0092-8674(02)00624-4

Baron, C. (2006). VirB8: a conserved type IV secretion system assembly factor and drug target. *Biochem. Cell Biol.* 25, 890–899. doi: 10.1139/o06-148

Basu, U., Lee, S.-W., Deshpande, A., Shen, J., Sohn, B.-K., Cho, H., et al. (2020). The C-terminal tail of the yeast mitochondrial transcription factor Mtf1 coordinates template strand alignment, DNA scrunching and timely transition into elongation. *Nucleic Acids Res.* 48, 2604–2620. doi: 10.1093/nar/gkaa040

Bellanger, X., Payot, S., Leblond-Bourget, N., and Guédon, G. (2014). Conjugative and mobilizable genomic islands in bacteria: evolution and diversity. *FEMS Microbiol. Rev.* 38, 720–760. doi: 10.1111/1574-6976.12058

Bello-López, J. M., Cabrero-Martínez, O. A., Ibáñez-Cervantes, G., Hernández-Cortez, C., Pelcastre-Rodríguez, L. I., Gonzalez-Avila, L. U., et al. (2019). Horizontal gene transfer and its association with antibiotic resistance in the genus aeromonas spp. *Microorganisms* 7:363. doi: 10.3390/microorganisms7090363

Bhatty, M., Gomez, J. A. L., and Christie, P. J. (2013). The expanding bacterial type IV secretion lexicon. *Res. Microbiol.* 164, 620–639. doi: 10.1016/j.resmic.2013.03.012

Cabezon, E., and de la Cruz, F. (2006). TrwB: an F1-ATPase-like molecular motor involved in DNA transport during bacterial conjugation. *Res. Microbiol.* 157, 299–305. doi: 10.1016/j.resmic.2005.12.002

Cabezón, E., Lanka, E., and de la Cruz, F. (1994). Requirements for mobilization of plasmids RSF1010 and ColE1 by the IncW plasmid R388: trwB and RP4 traG are interchangeable. *J. Bacteriol.* 176, 4455–4458. doi: 10.1128/jb.176.14.4455-4458.1994

Cascales, E., Atmakuri, K., Sarkar, M. K., and Christie, P. J. (2013). DNA substrate-induced activation of the agrobacterium VirB/VirD4 Type IV secretion system. *J. Bacteriol.* 195, 2691–2704. doi: 10.1128/jb.00114-13

Cascales, E., and Christie, P. J. (2004a). *Agrobacterium* VirB10, an ATP energy sensor required for type IV secretion. *Proc. Natl. Acad. Sci. U.S.A.* 101, 17228–17233. doi: 10.1073/pnas.0405843101

Cascales, E., and Christie, P. J. (2004b). Definition of a bacterial type IV secretion pathway for a DNA substrate. *Science* 304, 1170–1173. doi: 10.1126/science.1095211

Cassini, A., Högberg, L. D., Plachouras, D., Quattrocchi, A., Hoxha, A., Simonsen, G. S., et al. (2019). Attributable deaths and disability-adjusted life-years caused by infections with antibiotic-resistant bacteria in the EU and the European Economic Area in 2015: a population-level modelling analysis. *Lancet Infect. Dis.* 19, 56–66.

Casu, B., Arya, T., Bessette, B., and Baron, C. (2017). Fragment-based screening identifies novel targets for inhibitors of conjugative transfer of antimicrobial resistance by plasmid pKM101. *Sci. Rep.* 7:14907.

Casu, B., Smart, J., Hancock, M. A., Smith, M., Sygusch, J., and Baron, C. (2016). Structural Analysis and Inhibition of TraE from the pKM101 Type IV Secretion System. *J. Biol. Chem.* 291, 23817–23829. doi: 10.1074/jbc.m116.753327

Chan, M. (2015). WHO Library Cataloguing-in-Publication Data Global Action Plan on Antimicrobial Resistance, ed. World Health Organization. doi: 10.1074/jbc.m116.753327

Chen, Y., Zhang, X., Manias, D., Yeo, H.-J., Dunny, G. M., and Christie, P. J. (2008). *Enterococcus faecalis* PcfC, a spatially localized substrate receptor for type iv secretion of the pcf10 transfer intermediate. *J. Bacteriol.* 190, 3632–3645. doi: 10.1128/jb.01999-07

Chetrit, D., Hu, B., Christie, P. J., Roy, C. R., and Liu, J. (2018). A unique cytoplasmic ATPase complex defines the *Legionella pneumophila* type IV secretion channel. *Nat. Microbiol.* 3, 678–686. doi: 10.1038/s41564-018-0165-z

Christie, P. J. (2004). Type IV secretion: the agrobacterium VirB/D4 and related conjugation systems. *Biochim. Biophys. Acta Mol. Cell Res.* 1694, 219–234. doi: 10.1016/j.bbamcr.2004.02.013

Christie, P. J. (2016). The mosaic type IV secretion systems. *EcoSal Plus* 7:10.1128/ecosalplus.ESP-0020-2015. doi: 10.1128/ecosalplus.ESP-0020-2015

Christie, P. J., Atmakuri, K., Krishnamoorthy, V., Jakubowski, S., and Cascales, E. (2005). Biogenesis, architecture, and function of bacterial type IV secretion systems. *Annu. Rev. Microbiol.* 59, 451–485. doi: 10.1146/annurev.micro.58.030603.123630

Christie, P. J., Gomez Valero, L., and Buchrieser, C. (2017). Biological diversity and evolution of Type IV secretion systems. *Curr. Top. Microbiol. Immunol.* 413, 1–30. doi: 10.1007/978-3-319-75241-9_1

De Paz, H. D., Larrea, D., Zunzunegui, S., Dehio, C., De La Cruz, F., and Llosa, M. (2010). Functional dissection of the conjugative coupling protein TrwB. *J. Bacteriol.* 192, 2655–2669. doi: 10.1128/jb.01692-09

de Paz, H. D., Sangari, F. J., Bolland, S., García-Lobo, J. M., Dehio, C., de la Cruz, F., et al. (2005). Functional interactions between type IV secretion systems involved in DNA transfer and virulence. *Microbiology* 151, 3505–3516. doi: 10.1099/mic.0.28410-0

Draper, O., César, C. E., Machón, C., de la Cruz, F., and Llosa, M. (2005). Site-specific recombinase and integrase activities of a conjugative relaxase in recipient cells. *Proc. Natl. Acad. Sci. U.S.A.* 102, 16385–16390. doi: 10.1073/pnas.0506081102

Fernandez-Lopez, R., Machón, C., Longshaw, C. M., Martin, S., Molin, S., Zechner, E. L., et al. (2005). Unsaturated fatty acids are inhibitors of bacterial conjugation. *Microbiology* 151, 3517–3526.

Frost, L. S., Leplae, R., Summers, A. O., and Toussaint, A. (2005). Mobile genetic elements: the agents of open source evolution. *Nat. Rev. Microbiol.* 3, 722–732. doi: 10.1038/nrmicro1235

Garbisu, C., Garaiyurrebaso, O., Lanzén, A., Álvarez-Rodríguez, I., Arana, L., Blanco, F., et al. (2018). Mobile genetic elements and antibiotic resistance in mine soil amended with organic wastes. *Sci. Total Environ.* 621, 725–733. doi: 10.1016/j.scitotenv.2017.11.221

García-Cazorla, Y., Getino, M., Sanabria-Ríos, D. J., Carballeira, N. M., de la Cruz, F., Arechaga, I., et al. (2018). Conjugation inhibitors compete with palmitic acid for binding to the conjugative traffic ATPase TrwD, providing a mechanism to inhibit bacterial conjugation. *J. Biol. Chem.* 293, 16923–16930. doi: 10.1074/jbc.ra118.004716

Garcillán-Barcia, M. P., Francia, M. V., and de La Cruz, F. (2009). The diversity of conjugative relaxases and its application in plasmid classification. *FEMS Microbiol. Rev.* 33, 657–687. doi: 10.1111/j.1574-6976.2009.00168.x

Geourjon, C., Orelle, C., Steinfels, E., Blanchet, C., Deléage, G., Di Pietro, A., et al. (2001). A common mechanism for ATP hydrolysis in ABC transporter and helicase superfamilies. *Trends Biochem. Sci.* 26, 539–544. doi: 10.1016/s0968-0004(01)01907-7

Getino, M., Fernández-López, R., Palencia-Gándara, C., Campos-Gómez, J., Sánchez-López, J. M., Martínez, M., et al. (2016). Tanzawaic acids, a chemically novel set of bacterial conjugation inhibitors. *PLoS One* 11:148098. doi: 10.1371/journal.pone.0148098

Getino, M., Sanabria-Ríos, D. J., Fernández-López, R., Campos-Gómez, J., Sánchez-López, J. M., Fernández, A., et al. (2015). Synthetic fatty acids prevent plasmid-mediated horizontal gene transfer. *mBio* 6:e0032-15.

Gillings, M., Boucher, Y., Labbate, M., Holmes, A., Krishnan, S., Holley, M., et al. (2008). The evolution of class 1 integrons and the rise of antibiotic resistance. *J. Bacteriol.* 190, 5095–5100. doi: 10.1128/jb.00152-08

Gomis-Ruth, F., Sola, M., Cruz, F., and Coll, M. (2005). Coupling factors in macromolecular Type-IV secretion machineries. *Curr. Pharm. Des.* 10, 1551–1565. doi: 10.2174/1381612043384817

Gomis-Rüth, F. X., and Coll, M. (2006). Cut and move: protein machinery for DNA processing in bacterial conjugation. *Curr. Opin. Struct. Biol.* 16, 744–752. doi: 10.1016/j.sbi.2006.10.004

Gomis-Rüth, F. X., de la Cruz, F., and Coll, M. (2002a). Structure and role of coupling proteins in conjugal DNA transfer. *Res. Microbiol.* 153, 199–204. doi: 10.1016/s0923-2508(02)01313-x

Gomis-Rüth, F. X., Moncalián, G., De La Cruz, F., and Coll, M. (2002b). Conjugative plasmid protein TrwB, an integral membrane type IV secretion system coupling protein. Detailed structural features and mapping of the active site cleft. *J. Biol. Chem.* 277, 7556–7566. doi: 10.1074/jbc.m110462200

Gomis-Rüth, F. X., Moncalián, G., Pérez-Luque, R., González, A., Cabezón, E., and de la Cruz, F. (2001). The bacterial conjugation protein TrwB resembles ring helicases and F1-ATPase. *Nature* 409, 637–641. doi: 10.1038/35054586

González-Rivera, C., Khara, P., Awad, D., Patel, R., Li, Y. G., Bogisch, M., et al. (2019). Two pKM101-encoded proteins, the pilus-tip protein TraC and Pep, assemble on the *Escherichia coli* cell surface as adhesins required for efficient conjugative DNA transfer. *Mol. Microbiol.* 111, 96–117. doi: 10.1111/mmi.14141

Gordon, J. E., Costa, T. R. D., Patel, R. S., Gonzalez-Rivera, C., Sarkar, M. K., Orlova, E. V., et al. (2017). Use of chimeric type IV secretion systems to define contributions of outer membrane subassemblies for contact-dependent translocation. *Mol. Microbiol.* 105, 273–293. doi: 10.1111/mmi.13700

Grohmann, E., Christie, P. J., Waksman, G., and Backert, S. (2018). Type IV secretion in Gram-negative and Gram-positive bacteria. *Mol. Microbiol.* 107, 455–471. doi: 10.1111/mmi.13896

Grohmann, E., Goessweiner-Mohr, N., and Brantl, S. (2016). DNA-binding proteins regulating pIP501 transfer and replication. *Front. Mol. Biosci.* 3:42.

Grohmann, E., Keller, W., and Muth, G. (2017). Mechanisms of conjugative transfer and Type IV secretion-mediated effector transport in gram-positive bacteria. *Curr. Top. Microbiol. Immunol.* 413, 115–141. doi: 10.1007/978-3-319-75241-9_5

Guédon, G., Libante, V., Coluzzi, C., Payot, S., and Leblond-Bourget, N. (2017). The obscure world of integrative and mobilizable elements, highly widespread elements that pirate bacterial conjugative systems. *Genes* 8:337. doi: 10.3390/genes8110337

Guglielmini, J., de la Cruz, F., and Rocha, E. P. C. (2013). Evolution of conjugation and type IV secretion systems. *Mol. Biol. Evol.* 30, 315–331. doi: 10.1093/molbev/mss221

Gunton, J. E., Gilmour, M. W., Alonso, G., and Taylor, D. E. (2005). Subcellular localization and functional domains of the coupling protein, TraG, from IncHI1 plasmid R27. *Microbiology* 151, 3549–3561. doi: 10.1099/mic.0.28255-0

Gunton, J. E., Gilmour, M. W., Baptista, K. P., Lawley, T. D., and Taylor, D. E. (2007). Interaction between the co-inherited TraG coupling protein and the TraJ membrane-associated protein of the H-plasmid conjugative DNA transfer system resembles chromosomal DNA translocases. *Microbiology* 153, 428–441. doi: 10.1099/mic.0.2006/001297-0

Guzmán-Herrador, D. L., Steiner, S., Alperi, A., González-Prieto, C., Roy, C. R., and Llosa, M. (2017). DNA delivery and genomic integration into mammalian target cells through Type IV A and B secretion systems of human pathogens. *Front. Microbiol.* 8:1503.

Haft, R. J. F., Gachelet, E. G., Nguyen, T., Toussaint, L., Chivian, D., and Traxler, B. (2007). In vivo oligomerization of the F conjugative coupling protein TraD. *J. Bacteriol.* 189, 6626–6634. doi: 10.1128/jb.00513-07

Hamilton, C. M., Lee, H., Li, P.-L., Cook, D. M., Piper, K. R., von Bodman, S. B., et al. (2000). TraG from RP4 and TraG and VirD4 from Ti Plasmids confer relaxosome specificity to the conjugal transfer system of pTiC58. *J. Bacteriol.* 182, 1541–1548. doi: 10.1128/jb.182.6.1541-1548.2000

Hanson, P. I., and Whiteheart, S. W. (2005). AAA+ proteins: have engine, will work. *Nat. Rev. Mol. Cell Biol.* 6, 519–529. doi: 10.1038/nrm1684

Harrison, M. L., Desaulniers, M. A., Noyce, R. S., and Evans, D. H. (2016). The acidic C-terminus of vaccinia virus I3 single-strand binding protein promotes proper assembly of DNA-protein complexes. *Virology* 489, 212–222. doi: 10.1016/j.virol.2015.12.020

Hilleringmann, M., Pansegrau, W., Doyle, M., Kaufman, S., MacKichan, M. L., Gianfaldoni, C., et al. (2006). Inhibitors of *Helicobacter* pylori ATPase Cagα block CagA transport and cag virulence. *Microbiology* 152, 2919–2930. doi: 10.1099/mic.0.28984-0

Högberg, L. D., Heddini, A., and Cars, O. (2010). The global need for effective antibiotics: challenges and recent advances. *Trends Pharmacol. Sci.* 31, 509–515. doi: 10.1016/j.tips.2010.08.002

Hormaeche, I., Alkorta, I., Moro, F., Valpuesta, J. M., Goni, F. M., and De La Cruz, F. (2002). Purification and properties of TrwB, a hexameric, ATP-binding integral membrane protein essential for R388 plasmid conjugation. *J. Biol. Chem.* 277, 46456–46462. doi: 10.1074/jbc.m207250200

Hormaeche, I., Iloro, I., Arrondo, J. L. R., Goñi, F. M., De La Cruz, F., and Alkorta, I. (2004). Role of the transmembrane domain in the stability of TrwB, an integral protein involved in bacterial conjugation. *J. Biol. Chem.* 279, 10955–10961. doi: 10.1074/jbc.m310422200

Hormaeche, I., Segura, R. L., Vecino, A. J., Goñi, F. M., de la Cruz, F., and Alkorta, I. (2006). The transmembrane domain provides nucleotide binding specificity to the bacterial conjugation protein TrwB. *FEBS Lett.* 580, 3075–3082. doi: 10.1016/j.febslet.2006.04.059

Hu, B., Khara, P., and Christie, P. J. (2019a). Structural bases for F plasmid conjugation and F pilus biogenesis in *Escherichia coli*. *Proct. Nat. Acad. Sci. U.S.A.* 116, 14222–14227. doi: 10.1073/pnas.1904428116

Hu, B., Khara, P., Song, L., Lin, A. S., Frick-Cheng, A. E., Harvey, M. L., et al. (2019b). In situ molecular architecture of the *Helicobacter pylori* cag type IV secretion system. *mBio* 10:e00849-19.

Johnsborg, O., Eldholm, V., and Håvarstein, L. S. (2007). Natural genetic transformation: prevalence, mechanisms and function. *Res. Microbiol.* 158, 767–778. doi: 10.1016/j.resmic.2007.09.004

Johnson, C. M., and Grossman, A. D. (2015). Integrative and conjugative elements (ICEs): what they do and how they work. *Annu. Rev. Genet.* 49, 577–601. doi: 10.1146/annurev-genet-112414-055018

Kittell, B. L., and Helinski, D. R. (1993). "Plasmid incompatibility and replication control," in *Bacterial Conjugation*, ed. D. B. Clewell (Boston, MA: Springer), 223–242. doi: 10.1007/978-1-4757-9357-4_8

Kumar, R. B., and Das, A. (2002). Polar location and functional domains of the *Agrobacterium tumefaciens* DNA transfer protein VirD4. *Mol. Microbiol.* 43, 1523–1532. doi: 10.1046/j.1365-2958.2002.02829.x

Kwak, M.-J., Kim, J. D., Kim, H., Kim, C., Bowman, J. W., Kim, S., et al. (2017). Architecture of the type IV coupling protein complex of *Legionella pneumophila*. *Nat. Microbiol.* 2:17114.

Kwapong, A. A., Stapleton, P., and Gibbons, S. (2019). Inhibiting plasmid mobility: the effect of isothiocyanates on bacterial conjugation. *Int. J. Antimicrob. Agents* 53, 629–636. doi: 10.1016/j.ijantimicag.2019.01.011

Lang, S., and Zechner, E. L. (2012). General requirements for protein secretion by the F-like conjugation system R1. *Plasmid* 67, 128–138. doi: 10.1016/j.plasmid.2011.12.014

Larrea, D., de Paz, H. D., Matilla, I., Guzmán-Herrador, D. L., Lasso, G., de la Cruz, F., et al. (2017). Substrate translocation involves specific lysine residues of the

central channel of the conjugative coupling protein TrwB. *Mol. Genet. Genomics* 292, 1037–1049. doi: 10.1007/s00438-017-1331-3

Lawley, T., Klimke, W., Gubbins, M., and Frost, L. (2003). F factor conjugation is a true type IV secretion system. *FEMS Microbiol. Lett.* 224, 1–15. doi: 10.1016/s0378-1097(03)00430-0

Lederberg, J., and Tatum, E. L. (1946). Gene recombination in *Escherichia coli*. *Nature* 158, 558. doi: 10.1038/158558a0

Lee, K. B., and Thomas, J. O. (2000). The effect of the acidic tail on the DNA-binding properties of the HMG1,2 class of proteins: Insights from tail switching and tail removal. *J. Mol. Biol.* 304, 135–149. doi: 10.1006/jmbi.2000.4206

Leelananda, S. P., and Lindert, S. (2016). Computational methods in drug discovery. *Beilstein J. Org. Chem.* 12, 2694–2718.

Li, Y. G., and Christie, P. J. (2018). "The Agrobacterium VirB/VirD4 T4SS: mechanism and architecture defined through in vivo mutagenesis and chimeric systems," in *Agrobacterium Biology. Current Topics in Microbiology and Immunology*, Vol. 418, ed. S. Gelvin (Cham: Springer), 233–260. doi: 10.1007/82_2018_94

Li, Y. G., Hu, B., and Christie, P. J. (2019). Biological and structural diversity of type IV secretion systems. *Microbiol. Spectr.* 7:10.1128/microbiolsec.SIB-0012-2018.

Llosa, M., and Alkorta, I. (2017). Coupling proteins in type IV secretion. *Curr. Top. Microbiol. Immunol.* 413, 143–168. doi: 10.1007/978-3-319-75241-9_6

Llosa, M., Gomis-Rüth, F. X., Coll, M., and de la Cruz Fd, F. (2002). Bacterial conjugation: a two-step mechanism for DNA transport. *Mol. Microbiol.* 45, 1–8. doi: 10.1046/j.1365-2958.2002.03014.x

Llosa, M., Zunzunegui, S., and de la Cruz, F. (2003). Conjugative coupling proteins interact with cognate and heterologous VirB10-like proteins while exhibiting specificity for cognate relaxosomes. *Proc. Natl. Acad. Sci. U.S.A.* 100, 10465–10470. doi: 10.1073/pnas.1830264100

Low, H. H., Gubellini, F., Rivera-Calzada, A., Braun, N., Connery, S., Dujeancourt, A., et al. (2014). Structure of a type IV secretion system. *Nature* 508, 550–553. doi: 10.1038/nature13081

Lu, J., and Frost, L. S. (2005). Mutations in the C-terminal region of TraM provide evidence for in vivo TraM-TraD interactions during F-plasmid conjugation. *J. Bacteriol.* 187, 4767–4773. doi: 10.1128/jb.187.14.4767-4773.2005

Lu, J., Wong, J. J. W., Edwards, R. A., Manchak, J., Frost, L. S., and Glover, J. N. M. (2008). Structural basis of specific TraD-TraM recognition during F plasmid-mediated bacterial conjugation. *Mol. Microbiol.* 70, 89–99. doi: 10.1111/j.1365-2958.2008.06391.x

Martín-Santamaría, S. (2018). *Computational Tools for Chemical Biology*. London: Royal Society of Chemistry.

Matilla, I., Alfonso, C., Rivas, G., Bolt, E. L., de la Cruz, F., and Cabezon, E. (2010). The Conjugative DNA translocase TrwB is a structure-specific DNA-binding protein. *J. Biol. Chem.* 285, 17537–17544. doi: 10.1074/jbc.m109.084137

Meir, A., Chetrit, D., Liu, L., Roy, C. R., and Waksman, G. (2018). Legionella DotM structure reveals a role in effector recruiting to the Type 4B secretion system. *Nat. Commun.* 9:507.

Middleton, R., Sjölander, K., Krishnamurthy, N., Foley, J., and Zambryski, P. (2005). Predicted hexameric structure of the Agrobacterium VirB4 C terminus suggests VirB4 acts as a docking site during type IV secretion. *Proc. Natl. Acad. Sci. U.S.A.* 102, 1685–1690. doi: 10.1073/pnas.0409399102

Mihajlovic, S., Lang, S., Sut, M. V., Strohmaier, H., Gruber, C. J., Koraimann, G., et al. (2009). Plasmid R1 conjugative DNA processing is regulated at the coupling protein interface. *J. Bacteriol.* 191, 6877–6887. doi: 10.1128/jb.00918-09

Moncalián, G., Cabezón, E., Alkorta, I., Valle, M., Moro, F., Valpuesta, J. M., et al. (1999). Characterization of ATP and DNA binding activities of TrwB, the coupling protein essential in plasmid R388 conjugation. *J. Biol. Chem.* 274, 36117–36124. doi: 10.1074/jbc.274.51.36117

Oikonomou, C. M., and Jensen, G. J. (2019). "Electron cryotomography of bacterial secretion systems," in *Protein Secretion in Bacteria*, eds M. Sandkvist, E. Cascales, and P. Christie (American Society for Microbiology), 1–12.

Parsons, J. A., Bannam, T. L., Devenish, R. J., and Rood, J. I. (2007). TcpA, an FtsK/SpoIIIE homolog, is essential for transfer of the conjugative plasmid pCW3 in *Clostridium perfringens*. *J. Bacteriol.* 189, 7782–7790.

Peña, A., Matilla, I., Martín-Benito, J., Valpuesta, J. M., Carrascosa, J. L., De La Cruz, F., et al. (2012). The hexameric structure of a conjugative VirB4 protein ATPase provides new insights for a functional and phylogenetic relationship with DNA translocases. *J. Biol. Chem.* 287, 39925–39932.

Redzej, A., Ukleja, M., Connery, S., Trokter, M., Felisberto-Rodrigues, C., Cryar, A., et al. (2017). Structure of a VirD4 coupling protein bound to a VirB type IV secretion machinery. *EMBO J.* 36, 3080–3095.

Ripoll-Rozada, J., García-Cazorla, Y., Getino, M., Machón, C., Sanabria-Ríos, D., de la Cruz, F., et al. (2016). Type IV traffic ATPase TrwD as molecular target to inhibit bacterial conjugation. *Mol. Microbiol.* 100, 912–921.

Ripoll-Rozada, J., Zunzunegui, S., de la Cruz, F., Arechaga, I., and Cabezón, E. (2013). Functional interactions of VirB11 traffic ATPases with VirB4 and VirD4 molecular motors in type IV secretion systems. *J. Bacteriol.* 195, 4195–4201.

Sastre, I., Cabezón, E., and de la Cruz, F. (1998). The carboxyl terminus of protein TraD adds specificity and efficiency to F-plasmid conjugative transfer. *J. Bacteriol.* 180, 6039–6042.

Sayer, J. R., Walldén, K., Pesnot, T., Campbell, F., Gane, P. J., Simone, M., et al. (2014). 2- and 3-substituted imidazo[1,2-a]pyrazines as inhibitors of bacterial type IV secretion. *Bioorgan. Med. Chem.* 22, 6459–6470.

Schroder, G., Krause, S., Zechner, E. L., Traxler, B., Yeo, H.-J., Lurz, R., et al. (2002). TraG-like proteins of DNA transfer systems and of the *Helicobacter* pylori Type IV secretion system: inner membrane gate for exported substrates? *J. Bacteriol.* 184, 2767–2779.

Schroder, G., and Lanka, E. (2003). TraG-Like proteins of type IV secretion systems: functional dissection of the multiple activities of TraG (RP4) and TrwB (R388). *J. Bacteriol.* 185, 4371–4381.

Schröder, G., and Lanka, E. (2005). The mating pair formation system of conjugative plasmids-A versatile secretion machinery for transfer of proteins and DNA. *Plasmid* 54, 1–25.

Segura, R. L., Águila-Arcos, S., Ugarte-Uribe, B., Vecino, A. J., De La Cruz, F., Goñi, F. M., et al. (2013). The transmembrane domain of the T4SS coupling protein TrwB and its role in protein-protein interactions. *Biochim. Biophys. Acta Biomembr.* 1828, 2015–2025.

Segura, R. L., Águila-Arcos, S., Ugarte-Uribe, B., Vecino, A. J., de la Cruz, F., Goñi, F. M., et al. (2014). Subcellular location of the coupling protein TrwB and the role of its transmembrane domain. *Biochim. Biophys. Acta Biomembr.* 1838, 223–230.

Shaffer, C. L., Good, J. A. D., Kumar, S., Krishnan, K. S., Gaddy, J. A., Loh, J. T., et al. (2016). Peptidomimetic small molecules disrupt type IV secretion system activity in diverse bacterial pathogens. *mBio* 7:e00221-16.

Siguier, P., Gourbeyre, E., and Chandler, M. (2014). Bacterial insertion sequences: their genomic impact and diversity. *FEMS Microbiol. Rev.* 38, 865–891.

Smillie, C., Garcillán-Barcia, M. P., Francia, M. V., Rocha, E. P. C., and de la Cruz, F. (2010). Mobility of plasmids. *Microbiol. Mol. Biol. Rev.* 74, 434–452.

Smith, M. A., Coinon, M., Paschos, A., Jolicoeur, B., Lavallée, P., Sygusch, J., et al. (2012). Identification of the binding site of Brucella VirB8 interaction inhibitors. *Chem. Biol.* 19, 1041–1048.

Steen, J. A., Bannam, T. L., Teng, W. L., Devenish, R. J., and Rood, J. I. (2009). The putative coupling protein TcpA interacts with other pCW3-encoded proteins to form an essential part of the conjugation complex. *J. Bacteriol.* 191, 2926–2933.

Sutherland, M. C., Nguyen, T. L., Tseng, V., and Vogel, J. P. (2012). The Legionella IcmSW complex directly interacts with DotL to mediate translocation of adaptor-dependent substrates. *PLoS Pathog.* 8:e1002910.

Tacconelli, E., Carrara, E., Savoldi, A., Harbarth, S., Mendelson, M., Monnet, D. L., et al. (2018). Discovery, research and development of new antibiotics: the who priority list of antibiotic-resistant bacteria and tuberculosis. *Lancet Infect. Dis.* 18, 318–327.

Tato, I., Matilla, I., Arechaga, I., Zunzunegui, S., de la Cruz, F., and Cabezon, E. (2007). The ATPase activity of the DNA transporter TrwB is modulated by protein TrwA: implications for a common assembly mechanism of DNA translocating motors. *J. Biol. Chem.* 282, 25569–25576.

Tato, I., Zunzunegui, S., de la Cruz, F., and Cabezon, E. (2005). TrwB, the coupling protein involved in DNA transport during bacterial conjugation, is a DNA-dependent ATPase. *Proc. Natl. Acad. Sci. U.S.A.* 102, 8156–8161.

Teese, M. G., and Langosch, D. (2015). Role of GxxxG motifs in transmembrane domain interactions. *Biochemistry* 54, 5125–5135.

Thomas, C. M., and Nielsen, K. M. (2005). Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nat. Rev. Microbiol.* 3, 711–721.

Thomas, C. M., Thomson, N. R., Cerdeño-Tárraga, A. M., Brown, C. J., Top, E. M., and Frost, L. S. (2017). Annotation of plasmid genes. *Plasmid* 91, 61–67.

Urra, J., Alkorta, I., and Garbisu, C. (2019a). Potential benefits and risks for soil health derived from the use of organic amendments in agriculture. *Agronomy* 9:542. doi: 10.3390/agronomy9090542

Urra, J., Alkorta, I., Lanzén, A., Mijangos, I., and Garbisu, C. (2019b). The application of fresh and composted horse and chicken manure affects soil quality, microbial composition and antibiotic resistance. *Appl. Soil Ecol.* 135, 73–84.

Urra, J., Alkorta, I., Mijangos, I., Epelde, L., and Garbisu, C. (2019c). Application of sewage sludge to agricultural soil increases the abundance of antibiotic resistance genes without altering the composition of prokaryotic communities. *Sci. Total Environ.* 647, 1410–1420.

Vecino, A. J., de la Arada, I., Segura, R. L., Goñi, F. M., de la Cruz, F., Arrondo, J. L. R., et al. (2011). Membrane insertion stabilizes the structure of TrwB, the R388 conjugative plasmid coupling protein. *Biochim. Biophys. Acta* 1808, 1032–1039.

Voth, D. E., Broederdorf, L. J., and Graham, J. G. (2012). Bacterial Type IV secretion systems: versatile virulence machines. *Future Microbiol.* 7, 241–257.

Waksman, G. (2019). From conjugation to T4S systems in Gram-negative bacteria: a mechanistic biology perspective. *EMBO Rep.* 20:e47012.

Walker, J. E., Saraste, M., Runswick, M. J., and Gay, N. J. (1982). Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* 1, 945–951.

Walldén, K., Williams, R., Yan, J., Lian, P. W., Wang, L., Thalassinos, K., et al. (2012). Structure of the VirB4 ATPase, alone and bound to the core complex of a type IV secretion system. *Proc. Natl. Acad. Sci. U.S.A.* 109, 11348–11353.

Wang, Q., Zeng, M., Wang, W., and Tang, J. (2007). The HMGB1 acidic tail regulates HMGB1 DNA binding specificity by a unique mechanism. *Biochem. Biophys. Res. Commun.* 360, 14–19.

Whitaker, N., Berry, T. M., Rosenthal, N., Gordon, J. E., Gonzalez-Rivera, C., Sheehan, K. B., et al. (2016). Chimeric coupling proteins mediate transfer of heterologous type IV effectors through the *Escherichia coli* pKM101-encoded conjugation machine. *J. Bacteriol.* 198, 2701–2718.

Whitaker, N., Chen, Y., Jakubowski, S. J., Sarkar, M. K., Li, F., and Christie, P. J. (2015). The All-alpha domains of coupling proteins from the *Agrobacterium tumefaciens* VirB/VirD4 and *Enterococcus faecalis* pCF10-encoded Type IV secretion systems confer specificity to binding of cognate DNA substrates. *J. Bacteriol.* 197, 2335–2349.

World Health Organization (2015). *Global Action Plan On Antimicrobial Resistance*. Geneva: World Health Organization.

World Health Organization (2019a). *Thirteenth General Programme Of Work 2019-2023*. Geneva: World Health Organization.

World Health Organization (2019b). *2019 Antibacterial Agents In Clinical Development: An Analysis Of The Antibacterial Clinical Development Pipeline*. Geneva: World Health Organization.

Zechner, E. L., Lang, S., and Schildbach, J. F. (2012). Assembly and mechanisms of bacterial type IV secretion machines. *Philos. Trans. R. Soc. B Biol. Sci.* 367, 1073–1087.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Álvarez-Rodríguez, Arana, Ugarte-Uribe, Gómez-Rubio, Martín-Santamaría, Garbisu and Alkorta. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 19 August 2020
doi: 10.3389/fmolb.2020.00207

Formation of Complexes Between O Proteins and Replication Origin Regions of Shiga Toxin-Converting Bacteriophages

*Katarzyna Kozłowska[1], Monika Glinkowska[2], Lidia Boss[2], Lidia Gaffke[1], Jakub Deptuła[1] and Grzegorz Węgrzyn[1]**

[1] Department of Molecular Biology, Faculty of Biology, University of Gdańsk, Gdańsk, Poland, [2] Department of Bacterial Molecular Genetics, Faculty of Biology, University of Gdańsk, Gdańsk, Poland

OPEN ACCESS

Edited by:
Chew Chieng Yeo, Sultan Zainal Abidin University, Malaysia

Reviewed by:
Radoslaw Pluta, Institute for Research in Biomedicine, Spain
Lanying Zeng, Texas A&M University, United States
Stephen M. Kwong, The University of Sydney, Australia

***Correspondence:**
Grzegorz Węgrzyn
grzegorz.wegrzyn@biol.ug.edu.pl; wegrzyn@biotech.ug.gda.pl

Specialty section:
This article was submitted to Molecular Recognition, a section of the journal Frontiers in Molecular Biosciences

Received: *09 April 2020*
Accepted: *28 July 2020*
Published: *19 August 2020*

Citation:
Kozłowska K, Glinkowska M, Boss L, Gaffke L, Deptuła J and Węgrzyn G (2020) Formation of Complexes Between O Proteins and Replication Origin Regions of Shiga Toxin-Converting Bacteriophages. Front. Mol. Biosci. 7:207. doi: 10.3389/fmolb.2020.00207

Shiga toxin-converting bacteriophages (or Stx phages) are responsible for virulence of enterohemorrhagic *Escherichia coli* strains. Although they belong to the group of lambdoid phages, which have served as models in studies on DNA replication mechanisms, details of regulation of replication of Stx phage genomes are poorly understood. Despite high similarity of their replication regions to that of phage lambda, considerable differences occur between them. Here, we present a comparison of origins of replication and O proteins of lambda and selected Stx phages (phages P27 and 933W). Stx initiator proteins, similarly to the lambda O protein, exist in the form of dimers. Only 4 iteron sequences are strongly bound *in vitro* by the O proteins, despite the presence of 6 such fragments in the Stx *ori*, while the function of the other two iterons is still crucial for transformation of *E. coli* wild-type strain by the P27-derived lambdoid plasmid. As these sequences are found in the gene coding for Stx O proteins, the sequences of these proteins themselves are also extended compared to lambda phage. Therefore, proteins O of Stx phages P27 and 933W have 13 additional amino acids. They can act as a space barrier, thus affecting the lesser packing of the O-some Stx complex compared to the structure found in lambda. Such structure of the DNA replication initiation complex may determine its lesser dependence on the processes occurring in the host cell, including transcriptional activation of the origin. Differences between molecular processes occurring during formation of replication complexes in lambda and Stx phages may indicate the specialization of the latter phages and their adaptation to specific environmental conditions where quick genetic switches are crucial.

Keywords: Shiga toxin-converting bacteriophages, DNA replication, replication initiator protein, replication complex, protein-DNA interactions

INTRODUCTION

Escherichia coli constitutes crucial component of the natural microbiota of the human and warm-blooded animals intestine (Actis, 2014). Although the vast majority of *E. coli* strains are commensals, some of them can also be pathogenic to humans (Leimbach et al., 2013). Shiga toxin-producing *E. coli* (STEC) strains, synthesizing one of the strongest toxins, are among the most dangerous pathogens (Paletta et al., 2020). Infections with these strains are particularly dangerous, because they can cause severe complications, while available antibacterial therapies are often not effective against those pathogens (Karmali, 2018). The latter problem is caused by two factors: firstly, STEC strains are often resistant to many antibiotics, secondly, the use of antimicrobials may be counterproductive, as some of them may enhance expression of Shiga toxins genes. Thus, even if a STEC strain is susceptible to a particular therapeutic its application may cause production of sufficiently high level of toxins to evoke hemolytic uremic syndrome in patients, resulting in significant morbidity and mortality (Kakoullis et al., 2019; Mir and Kudva, 2019). There are examples of many outbreaks caused by STEC which resulted in severe complications in thousands of patients and relatively many deaths (Devleesschauwer et al., 2019; Kintz et al., 2019).

Shiga toxins, the main virulence factors of STEC, are encoded by the genes (called *stx* genes) present in genomes of Shiga toxin-converting bacteriophages (or Stx bacteriophages) rather than in bacterial chromosome (Łoś et al., 2011). Stx phages occur as prophages in STEC genomes, and *stx* genes, like vast majority of prophage genes, are silent in the lysogenic state (Krüger and Lucchesi, 2015). Under various stress conditions, transition of the phage from lysogenic into lytic development is possible which involves the excision of its genome and replication of phage DNA, followed by progeny virion production and cell lysis (Łoś et al., 2011; Licznerska et al., 2016). Simultaneously, extensive production of Shiga toxin occurs, while inhibition of phage DNA replication impairs expression of *stx* genes (coding for the toxins) (Nejman-Faleńczyk et al., 2012; Nowicki et al., 2013; Balasubramanian et al., 2019). Therefore, detailed knowledge on the regulation of DNA replication of Stx phages may contribute to better understanding of the mechanisms of pathogenicity of STEC strains.

Stx phages belong to the group of lambdoid bacteriophages which include the model virus of molecular biology – bacteriophage λ (Łoś et al., 2011). DNA replication of this phage has been well understood, but its regulation is still not clear (Węgrzyn et al., 2012). It was described as a complex process involving both phage elements and protein machinery of its bacterial host (Weigel and Seitz, 2006). The key role in the regulation of initiation of this process is played by the bacteriophage-encoded O protein and transcriptional activation of the *origin* (Taylor and Wegrzyn, 1995). λ DNA replication requires binding of the phage-encoded O protein to four DNA sequences in the *origin* of replication region, named iterons, and formation of the nucleoprotein structure called O-some. This is the crucial step in the initiation event, and provides the basis for formation of the replication complex (Taylor and Wegrzyn, 1995; Weigel and Seitz, 2006).

It was demonstrated that the replication regions in genomes of Stx bacteriophages are similar to that of bacteriophage λ (Nejman et al., 2009). The key elements in this region are the p_R promoter, and *O* and *P* genes (coding for replication proteins) with the *origin* sequence (consisting of O-binding sequences, called iterons, and the AT-rich tract) located in the middle of the former gene (**Figure 1**). Only a few alterations were found between replication regions of λ and Stx phages, but they are responsible for significant differences in the DNA replication regulation. Particularly, various requirements for host-encoded factors, like DnaA and DksA proteins, and different dependence on the transcriptional activation of the *origin* (the process demonstrated previously as crucial for the control of the replication initiation; see Szalewska-Pałasz et al., 1994; Szambowska et al., 2011) were reported (Nejman et al., 2011). These differences result in various responses of intracellularly developing λ and Stx phages to environmental conditions, including starvation (Nejman et al., 2011).

Replication regions of Stx phages contain 6 iterons, contrary to 4 iterons present in λ, and a few amino acid alterations can be find in O and P proteins between these phages (**Figure 1**), which nevertheless cause significant differences in the phage DNA replication regulation (Nejman et al., 2011; Nowicki et al., 2015; Kozłowska et al., 2017). Recently, an improved method for overproduction and purification of O proteins of Stx phages has been reported which allowed to identify their specific binding to the *origin* regions (Kozłowska et al., 2017). Nevertheless, details of molecular interactions between these O proteins and iterons remained unknown, especially in the light of the above described differences between λ and Stx phages. Therefore, in this work, we aimed to determine details of formation of complexes between O proteins and *origin* regions of Stx phages and compare them to those of λ.

MATERIALS AND METHODS

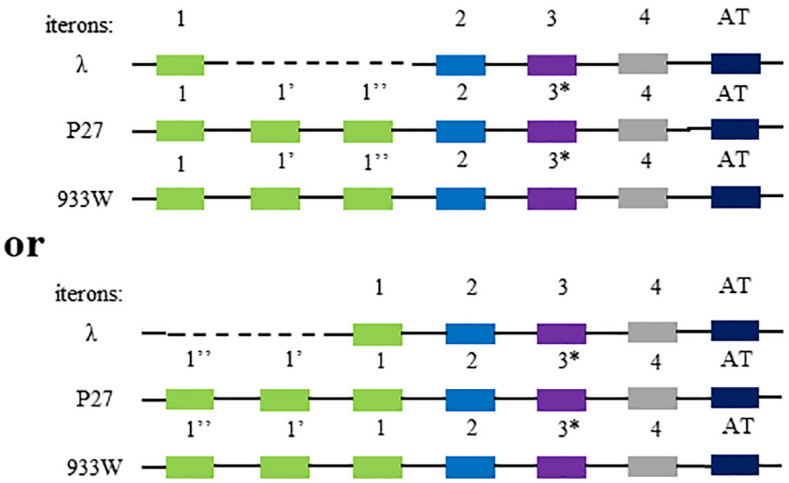
Proteins

Overproduction and purification of O proteins derived from phages λ, P27 and 933W used in this study was described previously (Kozłowska et al., 2017).

Bacterial Strains and Source Plasmids

E. coli DH5α (F^- Φ80*lacZ*ΔM15 Δ(*lacZYA-argF*) U169 *recA1 endA1 hsdR17*(r^-, m^+) *phoA supE44 thi-1 gyrA96 relA1* λ^-) strain (Taylor et al., 1993; Chen et al., 2018), and plasmids pCB104cmr (λ phage-derived), p27cmr (P27 phage-derived) (Nejman et al., 2009), pET24a_wtO_histag, pET24a_P27O_histag, pET24a_933WO_histag, pUC18_oriλ, and pUC18_oriStx (Kozłowska et al., 2017), and pUC18 (Yanisch-Perron et al., 1985) were used for production of proteins and DNAs for *in vitro* experiments. *E. coli* MG1655 and its *dnaA46* and Δ*seqA* derivatives (Nejman et al., 2009)

A



B

Iteron	Sequence
1	ATCCCTCAAAACGAGGGA
1'	T----CT----T------
1''	ATCCCTCAAAACGAGGGA
2	-----CT----------
3	----------TT-G----
3*	----------TTAG----
4	----------AG----

C

	O protein			
Phage	residue 17 (λ)	residue 37 (λ)	residue 282 (λ)	Total aa residues
λ	Gly	Leu	Ser	299
P27	Glu	Ile	Ser	312
933W	Glu	Ile	Gly	312

FIGURE 1 | Differences in the nucleotide sequences of viral genome fragments that are involved in the initiation of DNA replication of bacteriophage λ and Stx phages (P27 and 933). Compositions of iterons and the AT-rich fragment are shown in **(A)** (since sequences of iterons 1 and 1″ are identical, two alternative ways of comparison between λ and Stx *origin* regions are possible, with dashed line indicating the missing fragment; the upper scheme presents the comparison as originally proposed by Nejman et al. (2009), and the lower scheme presents the alternative option). The scheme is not drawn to scale. Nucleotide sequences of iterons are shown in **(B)**, with (–) indicating bases identical to those occurring in iteron 1 and 1″, and letters indicating substitutions. Differences in amino acid sequences of O proteins of bacteriophage λ and Stx phages (P27 and 933W) are depicted in **(C)**.

were used for transformation efficiency analysis, performed as described previously (Nejman et al., 2009).

Construction of DNA Templates and Plasmids

O protein binding to single iterons was tested by EMSA assay, employing DNA fragments obtained by single stranded oligonucleotides hybridization. Respective primer sequences are presented in **Table 1** (oligonucleotides 8–19). Single-stranded DNA fragments (synthetic oligonucleotides) for hybridization were ordered from Sigma Aldrich. Oligonucleotides were used at a final concentration 1 μg/ml. Hybridization buffer was prepared in a volume of 1 ml. Synthetic oligonucleotide hybridization buffer consisted of 0.1 ml of 1 M Na_3PO_4, 0.03 ml of 5 M NaCl and 0.002 ml of 0.5 M EDTA at pH 8. The hybridization was carried out in a thermocycler starting from temperature 98°C for 2 min, then 2 min at 80°C, another 2 min at 65°C and slow cooling down to 20°C. Following hybridization, double stranded DNA fragments were ligated with plasmid

TABLE 1 | Oligonucleotides.

No.	Oligonucleotide name	Sequence 5′–> 3′	Assay	PCR template
1.	Cy5_ori_EMSA_ATrich_for	[Cyanine 5]TCAAGCAGCAAGGCGGCATGTTTG	EMSA, forward	pUC18_oriλ and pUC18_oriStx
2.	ori_EMSA_ATrich_for	CACTTGCTGCCGCTCTGAATTGC	EMSA, reverse	
3.	ATrich_57GC_rev	GGATTCGCCAGAATTCTCGGACGAAGCAT CGCCTCAGCTCGCTTCGGTACTAGTGTCC GATGTGTCCC	EMSA, reverse with AT-rich region replaced with region of 57% GC content	
4.	Cy5_pUC18_iterony_for	[Cyanine 5]GGGCTGGCTTAACTATGCGGCATC	PCR primer, forward	
5.	pUC18_iterony_rev	GTTAGCTCACTCATTAGGCACCC	PCR primer, reverse	
6.	pUC18_liniowy_footprint_for	CTTCCGGCTGGCTGGTTTATTG	Linearization of plasmid for footprinting	
7.	pUC18_liniowy_footprint_rev	CTTTATCCGCCTCCATCCAGTC	Linearization of plasmid for footprinting	
8.	Iteron1_pUC18_for	GATCCCGTAGCAGTCGCTGACATCCCT CAAAACGAGGGACGTAGCAGT CGCTGACA	Hybridization, forward	Not applicable
9.	it1_rev	AGCTTGTCAGCGACTGCTACGTCCCTC GTTTTGAGGGATGTCAGCGAC TGCTACGG	Hybridization, reverse	
10.	Iteron1×_pUC18_for	GATCCCGTAGCAGTCGCTGAGTTCCC CTAAAATGAGGGACGTAGCAGT CGCTGACA	Hybridization, forward	
11.	it1′_rev	AGCTTGTCAGCGACTGCTACGTCCCTC ATTTTAGGGGAACTCAGCGAC TGCTACGG	Hybridization, reverse	
12.	Iteron2_pUC18_for	GATCCCGTAGCAGTCGCTGAAATCCCCT AAAACGAGGGACGTAGCAGT CGCTGACA	Hybridization, forward	
13.	it2_rev	AGCTTGTCAGCGACTGCTACGTCCCTCG TTTTAGGGGATTTCAGCGAC TGCTACGG	Hybridization, reverse	
14.	Iteron3_pUC18_for	GATCCCGTAGCAGTCGCTGACATCCCTC AAATTGGGGGACGTAGCAGT CGCTGACA	Hybridization, forward	
15.	it3_rev	AGCTTGTCAGCGACTGCTACGTCCCCC AATTTGAGGGATGTCAGCGAC TGCTACGG	Hybridization, reverse	
16.	Iteron3*_pUC18_for	GATCCCGTAGCAGTCGCTGACATCCCT CAAATTAGGGGACGTAGCAGT CGCTGACA	Hybridization, forward	
17.	it3*_rev	AGCTTGTCAGCGACTGCTACGTCCCC TAATTTGAGGGATGTCAGCGAC TGCTACGG	Hybridization, reverse	
18.	Iteron4_pUC18_for	GATCCCGTAGCAGTCGCTGAATCCCTC AAAACAGGGGGACGTAGCAGT CGCTGACA	Hybridization, forward	
19.	it4_rev	AGCTTGTCAGCGACTGCTACGTCCCC CTGTTTTGAGGGATTCAGCGAC TGCTACGG	Hybridization, reverse	
20.	EMSA_for	[Cyanine 5]TGCTGCAAGGCGA	EMSA with single iteron sequence, forward	pUC18 with single iterons sequences (depending on iteron)
21.	EMSA_rev	CACTTTATGCTTCCGGCTCGTAT	EMSA with single iteron sequence, reverse	
22.	Footrinting for	AAGCAGCAAGGCGGCATGTTTG	Footprinting	pUC18_oriλ and pUC18_oriStx
23.	Footrinting rev	GCTGGTCAGAGGATTCGC	Footprinting	
24.	Fret	[Cyanine 5]CTGTTCTATTGTGATCTCTTATTAG [Cyanine 3]	FRET	Not applicable
25.	Fret_Donor_Cy3	CTGTTCTATTGTGATCTCTTATTAG[Cyanine 3]	FRET	

(Continued)

TABLE 1 | Continued

No.	Oligonucleotide name	Sequence 5′–> 3′	Assay	PCR template
26.	FRET_methylation	CT(mp)G(mp) T(mp)TC(mp) T(mp)A(mp)T T(mp)G(mp)T(mp) GA(mp)T CTC TTA TTA G	EMSA with methylated ssDNA	
27.	Mut1Fd	GGACCAAATAAAAACATCTCAGAATGGTGC ATTCCTCAAAACGAAGGAGGTTCCCCTAAAATG	Site-directed mutagenesis of iteron sequence 1 (AT**C**CCTCAAAAACGA **G**GGAAT**T**CCTCAAAACGA**A**GGA)	p27R6K
28.	Mut1Rev	GATGTCCCTCATTTTAGGGGAACCTCCTTC GTTTTGAGGAATGCACCATTCTGAGATG	Site-directed mutagenesis of iteron sequence 1 (AT**C**CCTCAAAA ACGA**G**GGAAT**T**CCTCAAAACGA**A**GGA)	p27R6K
29.	Mut1 + 1′Fd	CATCTCAGAATGGTGCATTCCTCAAAACG AAGGAGGTTCTCCTAAAATGAGTGACATC CCTCAAAACGAGGG	Site-directed mutagenesis of iteron sequences 1 and 1′ (AT**C**CCT CAAAAACGA**G**GGAAT**T**CCTCAAAA CGA**A**GGA and TTC**C**CCTAAA ATGAG**G**GATTC**T**CCTAAAATGAG**T**GA)	p27mut1
30.	Mut1 + 1′Rev	GATGTTTTATCCCTCGTTTTAGGGGATTTT CCCTCGTTTTGAGGGATGTCACTCATTTT AGGAGAACCTCCT	Site-directed mutagenesis of iteron sequences 1 and 1′ (AT**C**CCTCA AAAACGA**G**GGAAT**T**CCTCAAAACGA **A**GGA and TTC**C**CCTAAAATGAG **G**GATTC**T**CCTAAAATGAG**T**GA respectively)	p27mut1
31.	Mut1′Fd	CATCTCAGAATGGTGCATCCCTCAAAACG AGGGAGGTTCTCCTAAAATGAGTGACATC CCTCAAAACGAGGG	Site-directed mutagenesis of iteron sequence 1′ (TTC**C**CCTAAAATGAG **G**GATTC**T**CCTAAAATGAG**T**GA)	p27R6K
32.	Mut1′Rev	GATGTTTTATCCCTCGTTTTAGGGGATTTTC CCTCGTTTTGAGGGATGTCACTCATTTTAGGA GAACCTCCC	Site-directed mutagenesis of iteron sequence 1′ (TTC**C**CCTAAAATGAG **G**GATTC**T**CCTAAAATGAG**T**GA)	p27R6K
33.	p27_IT1_Fd	[Cyanine5]GTGCATCCCTCAAAACGAGGGAGGTT	EMSA with wt and mutant variants in iteron 1	
34.	p27_IT1_Rev	AACCTCCCTCGTTTTGAGGGATGCAC	EMSA with wt variant in iteron 1 of P27	
35.	p27_IT1_prim_Fd	[Cyanine5]GAGGTTCCCCTAAAATGAGGGACATC	EMSA with wt variant in iteron 1′ of P27	
36.	p27_IT1_prim_Rev	[Cyanine5]GTGCATTCCTCAAAACGAAGGAGGTT	EMSA with wt variant in iteron 1′ of P27	
37.	p27_IT1mut_Rev	AACCTCCTTCGTTTTGAGGAATGCAC	EMSA with the mutant variant in iteron 1 of P27	
38.	p27_IT1mut_Rev	AACCTCCTTCGTTTTGAGGAATGCAC	EMSA with the mutant variant in iteron 1 of P27	
39.	p27_IT1_prim_mut_Fd	[Cyanine5]GAGGTTCTCCTAAAATGAGTGACATC	EMSA with the mutant variant in iteron 1′ of P27	
40.	p27IT1_prim_mut_Rev	GATGTCACTCATTTTAGGAGAACCTC	EMSA with the mutant variant in iteron 1′ of P27	

pUC18, previously linearized with *Bam*HI and *Hin*dIII (Thermo Scientific™), and used for transformation of *E. coli* DH5α competent cells. Obtained plasmids were isolated from bacterial cells using plasmid isolation kit (Sigma Aldrich) and sequenced to confirm their proper construction. DNA fragments for EMSA or footprinting tests were obtained by PCR technique with use of specific primers (see **Table 1** for details).

Derivatives of the p27cmr plasmid, bearing additional *origin* of replication of the R6K plasmid, were constructed as follows. The PCR fragment bearing *origin* R6K was obtained after the reaction with pCAH56 plasmid and primers *OriR6KFd* (5′ – GCC GAG CTC CAT CCC TGG CTT GTT GTC C) and *OriR6KFd* (5′ – GAT GAG CTC GAT CCG GCC ACG ATG CGT C). The reaction product was cut with *Sac*I, and ligated with the *Sac*I fragment of the p27cmr plasmid. Changes in iteron sequences 1 and/or 1′ of p27R6K plasmids, which did not alter the O amino acid sequence, were introduced by site-directed mutagenesis, producing p27mut1, p27mut1′, and p27mut1 + 1′ plasmids, respectively, using suitable primers (**Table 1**, primers 27–32). *E. coli* BW25142 (*lacI*^{q} *rrnB3 lacZ4787 hsdR* D(*araBAD*) D(*rhaBAD*) *phoBR580 rph-1 galU95 endA9 uid*A:*pir-116 recA1*) strain was used for plasmid replication after site-directed mutagenesis.

Electrophoretic Mobility Shift Assay (EMSA)

All electrophoretic mobility shift assays (EMSA) were performed in the same manner. Indicated amounts of purified O proteins were incubated with 100 ng of different DNA fragments (1 μg in case of ssDNA fragment – oligonucleotides 26 in **Table 1**) in a freshly prepared O incubation buffer (0.2 M sodium phosphate

pH 7.4, 0.1 M NaCl, 0.1 mM EDTA, 10 mM $MgCl_2$, 1 mM DTT, 5% glycerol) containing 2 μg of poly dI:dC (final volume of the reaction mixture was 20 μl). Incubation was performed for 15 min at 30°C, and it was stopped by placing samples on ice, and addition 1 μl of 40% saccharose. 1% agarose gel electrophoresis was run in 0.5 × TBE at 4°C, 80 V, for 3.5 h. Gels were visualized with the use of Typhoon FLA 7000 scanner (GE Healthcare).

Dimethyl Sulfide (DMS) Footprinting

Labeling of Oligonucleotides With ^{32}P Isotope

Labeling of oligonucleotides for the primer extension reaction with a ^{32}P phosphorus isotope was performed using polynucleotide kinase (PNK) (Thermo Scientific). Reagents (oligonucleotides, Buffer A, γ^{32}P-ATP, T4 PNK – in amounts added according to the PNK manufacturer's instruction) were mixed and incubated for 25 min at 37°C. Then, 10 μl of 50 mM EDTA were added to the mixture and incubation was continued for 10 min at 75°C. Following incubation, the solution was applied to Sephadex G-50 (Sigma Aldirch) resin. The labeled oligonucleotides were purified on the resin by spinning for 5 min at 800 rpm. The efficiency of the labeling reaction was assessed by measuring the radiation of 1 μl of purified oligonucleotide sample with a Geiger counter (minimum 1000 /μl).

Obtaining DNA Template for Primer Extension Reaction

In order to prepare a modified plasmid DNA template for primer extension reactions, 1 μg of plasmid DNA (pCB104cmr or P27cmr) were incubated with proteins λ O, P27 O and 933W O for 15 min at 30°C (to achieve different molar ratios: 1:2, 1:4, 1:8, 1:16, and 1:32) in the incubation buffer [see section "Electrophoretic Mobility Shift Assay (EMSA)"]. Following addition of 1 mM dimethyl sulfide (DMS) to the mixture, the incubation was continued for 5 min at 30°C. Reaction was halted by addition of 100 μl ice-cold stopping buffer (3 M ammonium acetate, 1 M β-mercaptoethanol, 20 mM EDTA) and 300 μl of 96% ethanol. DNA samples were then incubated at −80°C for 30 min. Samples were centrifuged for 30 min at 15,000 rpm at 4°C. The pellets were washed gently with 300 μl of 70% ethanol and centrifuged again under above described conditions. Precipitates were dried using a Vacufuge Concentrator (Eppendorf). Pellets were suspended in 100 μl of 1 M piperidine solution and incubated for 30 min at 90°C to break methylated DNA bonds. 100 μl of the solution were applied to Sephacryl S-500 resin (Sigma Aldirch), and centrifuged for 5 min at 500 rpm.

DNA Primer Extension Reaction

Primer extension reaction with Taq polymerase (Thermo Fisher Scientific) was carried out using an isotope-labeled oligonucleotide and modified plasmid DNA template under the conditions specified by Taq polymerase manufacturer. Following incubation scheme was applied: 3 min at 92°C, then 40 cycles of denaturation for 30 s at 92°C, annealing for 50 s at 50°C, and extension for 1 min at 72°C, and then final extension for 2 min at 72°C.

Sanger Sequencing Reaction

The sequencing reaction was carried out using the DNA Cycle Sequencing Kit (Jena Bioscience), according to the manufacturer's instructions.

DNA Electrophoresis

The 8% polyacrylamide gel was prepared using 50 ml of TBE buffer, 100 ml 40% polyacrylamide 19:1 mixture, 240 g urea and 150 ml ultrapure water. The mixture was slightly heated to dissolve urea and degassed using a vacuum pump. To 80 ml of the mixture 640 μl of 10% APS and 30 μl of TEMED were added, and the mixture was poured between glass plates of the electrophoresis apparatus. The gel polymerization lasted 1 h at room temperature. DNA sample after the primer extension reaction and sequencing was dried in a heating block at 99°C for 20–30 min. Obtained pellet was suspended in 12 μl of Loading Buffer (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol FF) and 5 μl of the solution were applied onto the polyacrylamide gel. Electrophoresis was carried out in a previously heated (45–50°C) gel, at 80 W, 1900 V, and 60 mA for 3.0–3.5 h. Following electrophoresis, gel was rinsed for 5 min in water, then 5 min in 10% acetic acid and again for 5 min in water. The gel was then dried, and sealed with an ionizing radiation capture screen (GE Healthcare). The exhibition was carried out overnight. Screen signal was visualized in the Typhoon FLA 7000 scanner (GE Healthcare).

FRET Assay

15 pM of ssDNA fragment with donor (Cy3) and acceptor (Cy5) was incubated with 500 ng of purified O proteins in Incubation buffer (0.2 M sodium phosphate pH 7.4, 0.1 M NaCl, 10 mM $MgCl_2$, 1 mM DTT, 0.1 mM EDTA, 5% glycerol) in a final volume of 20 μl for 5 min at room temperature. Excitation was conducted at 530 nm (Cy3) and fluorescence (550–700 nm wavelength spectrum) was detected with the use of EnSpire™ Multimode Plate Reader (Perkin Elmer).

RESULTS

The *origin* of replication of lambdoid bacteriophages, located in the middle of the *O* gene, consists of sequences which bind the O replication initiator protein, called iterons, and the region containing large proportion of AT base pairs (the AT-rich region). As indicated in **Figure 1A**, the *origin* of bacteriophage *λ* DNA replication, *oriλ*, contains 4 iterons, while *origins* of the Stx phages (exemplified in this work by phages P27 and 933W) contain 6 iterons. They are numbered as 1, 1′, 1″, 2, 3, and 4, where iterons 1, 2, 3, and 4 are present in *λ*, and iterons 1′ and 1″ occur only in Stx phages. There are a few nucleotide alterations between cognate iterons (**Figure 1B**). The sequence alignment is presented in **Supplementary Figure S1**. All iterons are almost the same in all tested *origin* regions (when compared between phages), except one nucleotide difference in iteron 3 of Stx phages relative to that of *λ*. Thus, this Stx-specific iteron has been named 3*. However, since iterons 1 and 1″ have identical sequences, there are two possible ways of comparing

the replication regions of λ and Stx phages (**Figure 1A**). The presence of 2 additional iterons in DNA of Stx phages results in appearance of 13 additional amino acid residues in the O protein, relative to phage λ *O* gene product (**Figure 1C**). Moreover, the sequence of the O protein of phage P27 contains two alterations relative to λ O protein (Gly17Glu and Leu37Ile), and that of phage 933W contains three such alterations (Gly17Glu, Leu37Ile, and Ser282Gly) (**Figure 1C**). The O protein sequence alignment with secondary structure assignment is shown in **Supplementary Figure S2**. Moreover, potential secondary DNA structures formed at the iteron region are presented in **Supplementary Figure S3**. Considering significant differences in the regulation of DNA replication between λ and Stx phages (Nejman et al., 2009, 2011; Nowicki et al., 2015), we aimed to investigate the effects of the above mentioned alterations on formation of the nucleoprotein complexes by the O proteins of lambdoid phages and their *origin* regions.

The O proteins of phages λ, P27, and 933W were purified as described previously (Kozłowska et al., 2017). Using EMSA assay with DNA fragments containing the replication *origin* regions of phages λ and P27 (with 4 and 6 iterons, respectively), we found that all tested O proteins formed multiple complexes with both templates (**Figure 2**). However, binding efficiency to *origin* regions varied among the O proteins variants. It was also different for the two Stx O proteins dependent on whether the origin fragment containing 4 or 6 iterons was present in the binding reactions, whereas λO seemed to bind similarly to both origins. For the four iteron λ *ori* both Stx O proteins formed nucleoprotein complexes of lower mass slightly more efficiently, at lower protein concentrations, than the λO protein. However, as the protein concentration was increased, λO formed DNA-protein complexes of higher mass more readily than both the P27 and 933W O proteins. On the other hand, the two latter O proteins formed consecutive nucleoprotein complexes with their cognate origin with lower efficiency than the λO protein. The difference in binding efficiency in comparison to λO was higher for the 933 O protein than for its counterpart from the phage P27 (**Figure 2**). These results strongly suggest that amino acid alterations between various O proteins result in differences in efficiency of formation of their complexes with the *origin* region. However, it was difficult to verify in the EMSA test if all iterons in the Stx *ori* are bound by the O proteins at higher protein concentrations, since nucleoprotein complexes of high mass often formed diffused bands, migrating at variable positions, reflecting their lower stability, various conformations or protein aggregation. Discrete bands representing 4, 5, or 6 dimers of the O proteins bound could not be distinguished.

The differences in interactions of O proteins derived from λ and Stx phages to *origin* regions might result from either various binding cooperativity properties of the initiator proteins or lower affinity of O proteins of Stx phages to iterons. Therefore, we tested efficiency of binding of O proteins to individual iterons, using templates containing single iterons: 1/1″, 1′, 2, 3, 3*, 4. We found that all tested O proteins formed complexes with each tested iteron, however, efficiency of their formation were slightly different. Phage λ O protein forms two different complexes with each iteron while O proteins of Stx phages form predominantly one complex (**Figure 3**). Again, binding of the λ O protein was the strongest among tested proteins. Moreover, iterons 2 and 3* were bound with the highest efficiency by all tested O proteins, while formation of the nucleoprotein complexes was the weakest for iterons 1′ and 4 (**Figure 3**). Nevertheless, the differences between efficiency of binding of O proteins by individual iterons were not dramatic.

More detailed information about binding of the O proteins to iterons were obtained in footprinting experiments. Results of experiments with both types of *origin* regions (with 4 and 6 iterons) confirmed the efficiency of binding of different variants of the O protein, with λ O being the most effective and the 933W O protein the least effective in interactions with iterons (**Figures 4**, **5** for results of experiments with the templates containing 4 and 6 iterons, respectively). In experiments with the template containing four iterons, all O proteins bound iterons 2 and 3 with the highest efficiency, then iteron 1, and iteron 4 with the lowest efficiency (**Figure 4**). Protection of the DNA sequence, indicating the contact sites with the O protein, is evident on both sides of the template, and includes two distal G residues in the G-tracks present in each iteron. Enhancement of the signal, indicating DNA bending, can be seen also at both template sides, between O-bound residues and in the middle of each iteron (**Figure 4**). These results are in agreement with previously published observations (Zahn and Blattner, 1985). Importantly, when the template containing six iterons was used, only four of them were bound efficiently by O proteins in each experimental system (i.e., with the O protein of phages λ, P27 and 933W). Sequences of two proximal iterons were not clearly protected by any of the tested O proteins, though weak interactions of iteron 1′ with O at the highest protein concentrations could be observed (**Figure 5**). Therefore, we conclude that the arrangement of iterons in the Stx phage *origin* relative to λ should be represented as indicated in the lower part of **Figure 1A**. At higher protein concentrations, λO protein bound iteron 1′, although with lower efficiency than the four iterons proximal to the AT-rich region, and very weakly interacted with the most distant iteron 1″. On the other hand the P27 and W933 O proteins showed even weaker footprint on iteron 1′, with very low level of protection and moderate enhancement at the correct nucleotides. The signal of binding was very weak at protein concentrations which resulted in full occupancy of the consecutive iterons 1/1″, 2, 3, and 4 by both proteins (**Supplementary Figure S4**). Similar, but even weaker pattern was observed with respect to all O proteins at the 1/1″ iteron situated at the edge of the origin.

Iterons are intrinsically bent sequences (Zahn and Blattner, 1985) and their anisotropy could affect binding parameters of the O protein. To assess whether curvature could underlie slight differences in binding of the O proteins to individual iterons, we performed an analysis of iterons geometry using AA Wedge model (Goodsell and Dickerson, 1994). The data obtained with DNA Curvature Analysis online application[1] showed that iterons 1/1″, 1′, 2, and 4 have identical local bending pattern. On the other hand, iterons 3 and 3*, which contain a track of three A, instead of four present in the remaining iterons, have

[1]https://www.lfd.uci.edu/~gohlke/dnacurve/

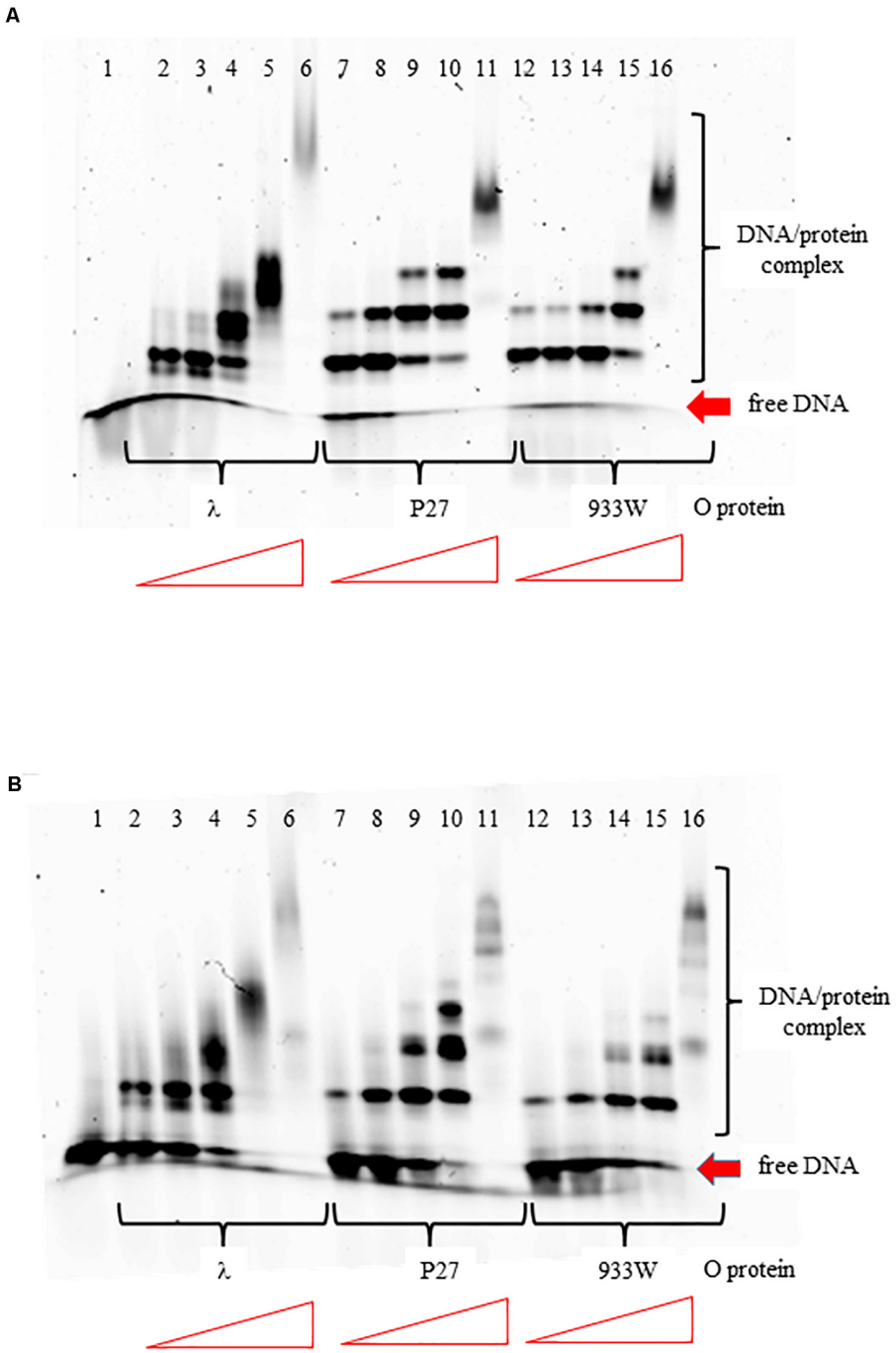


FIGURE 2 | Interactions of O proteins of phages λ, P27 and 933W with DNA templates containing four panel **(A)** or six panel **(B)** iterons, as assessed by EMSA. Increasing molar ratios of DNA:O proteins, indicated in particular wells, were as follows: 1:2, 1:4, 1:8, 1:16, and 1:32. Lane no. 1 contains no O protein.

slightly different curvature from the rest but indistinguishable from each other (**Supplementary Figure S5**). This suggests that differences in intrinsic DNA bending between individual iterons do not account for the observed preferences of the O proteins to particular sequences present in the origin context.

Using the tool mentioned above, we also analyzed DNA curvature of the entire *origin* regions, both containing 4 and 6 iterons. The results of the analysis (**Supplementary Figures S6, S7**) indicate that both types of origins contain a significantly curved region encompassing iterons 1–3. Both iteron 4 and especially iterons 1′–1″ are separated by regions of low curvature from iterons 1–3. The DNA structure of phage origins could be related to the observed hierarchy of iterons occupation by the O

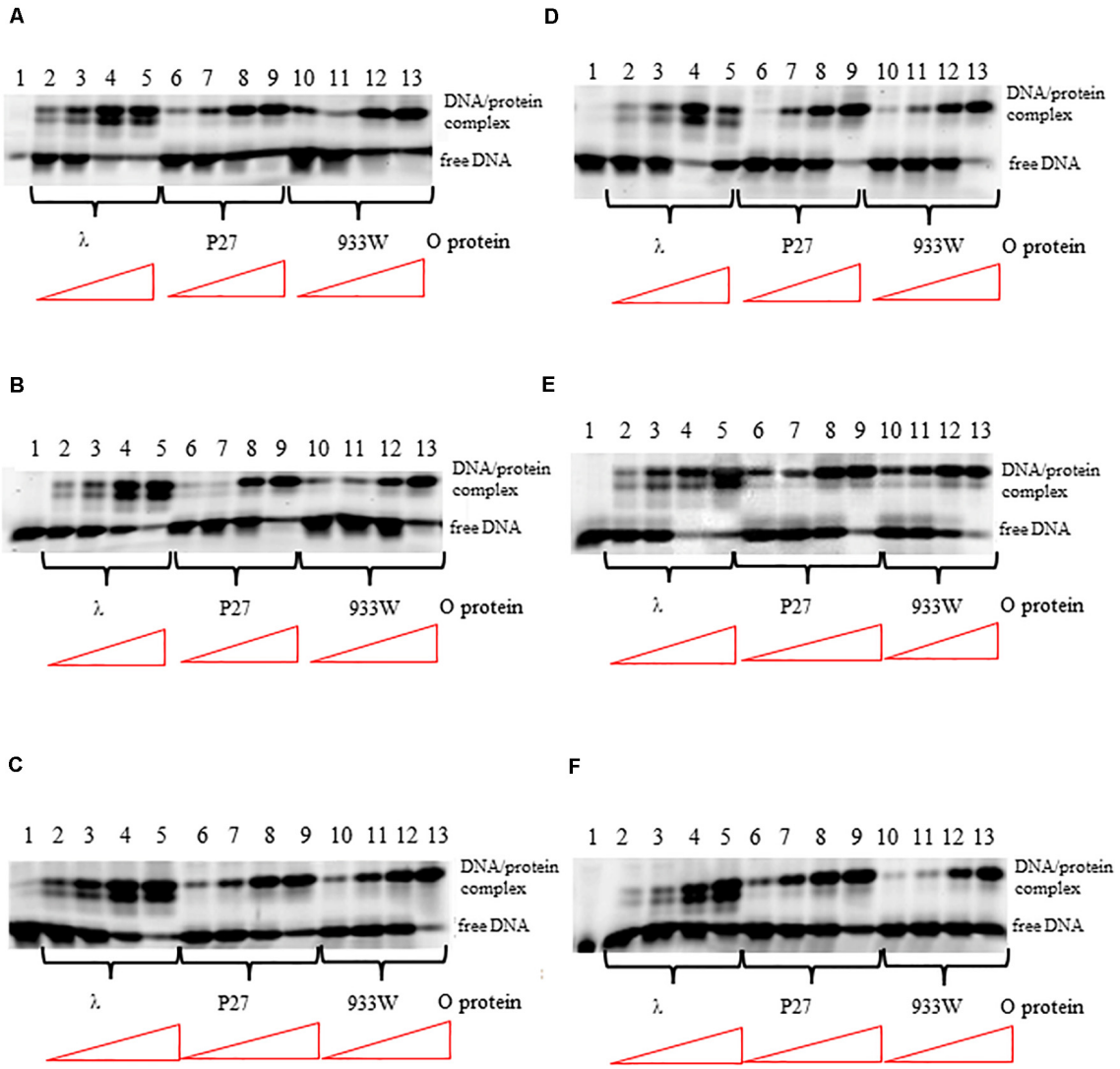


FIGURE 3 | Interactions of O proteins of phages λ, P27 and 933W with DNA templates containing single iterons: iteron 1/1″ **(A)**, iteron 1′ **(B)**, iteron 2 **(C)**, iteron 3 **(D)**, iteron 3* **(E)**, and iteron 4 **(F)**. Increasing molar ratios of DNA:O proteins, indicated in particular wells, were as follows: 1:1, 1:2, 1:4, and 1:8. Lane no. 1 contains no O protein. Note that lanes 10/11 in panel A, and lanes 4/5 in panel D, are swapped.

replication initiators, with iterons 4 and 1′–1″ being the least preferred.

To test if the additional iterons occurring in the *origin* regions of Stx phages are functionally significant, we have constructed plasmids bearing disrupted iteron 1 and/or 1′ sequence(s). DNA templates bearing such mutations bound λ O only very weakly at the highest used amounts of the protein, and did not bind the P27 O protein at all (**Figure 6**). Some differences in results of experiments with wild-type iterons may be due to differences in the DNA fragment lengths (55 vs. 28 bps). The presence of two bands of free DNA for wild-type iterons but only one band for mutated ones is apparently because of formation of bending in wild-type sequence which disappears when the mutation is present. To test functionality of the iteron 1′, double-*origin* plasmids have been constructed which contained lambdoid (P27-derived) replication region and *origin* of plasmid R6K, active only in the presence of the active *pir* gene. Sequences of iterons 1 and 1′ were changed as described above, and the alterations of base pairs were chosen to obtain disrupted O-binding properties and unchanged amino acid sequences of the O proteins, encoded by the *O* genes (as mentioned, the *origin* is located in the middle of the *O* gene). Replication of plasmids containing either wild-type or mutated iterons was possible in the *E. coli* strain expressing the *pir* gene due to activity of the *ori* R6K, irrespective of the activity of lambdoid *origin*. However, transformation of *E. coli* cells devoid of the *pir* gene was possible only in the presence of the active lambdoid *origin*. We have tested efficiency of transformation by lambdoid plasmids bearing either wild-type iterons or those with mutations in iteron 1 and/or 1′ of *E. coli* wild-type host, as well as *dnaA46,* Δ*seqA* or

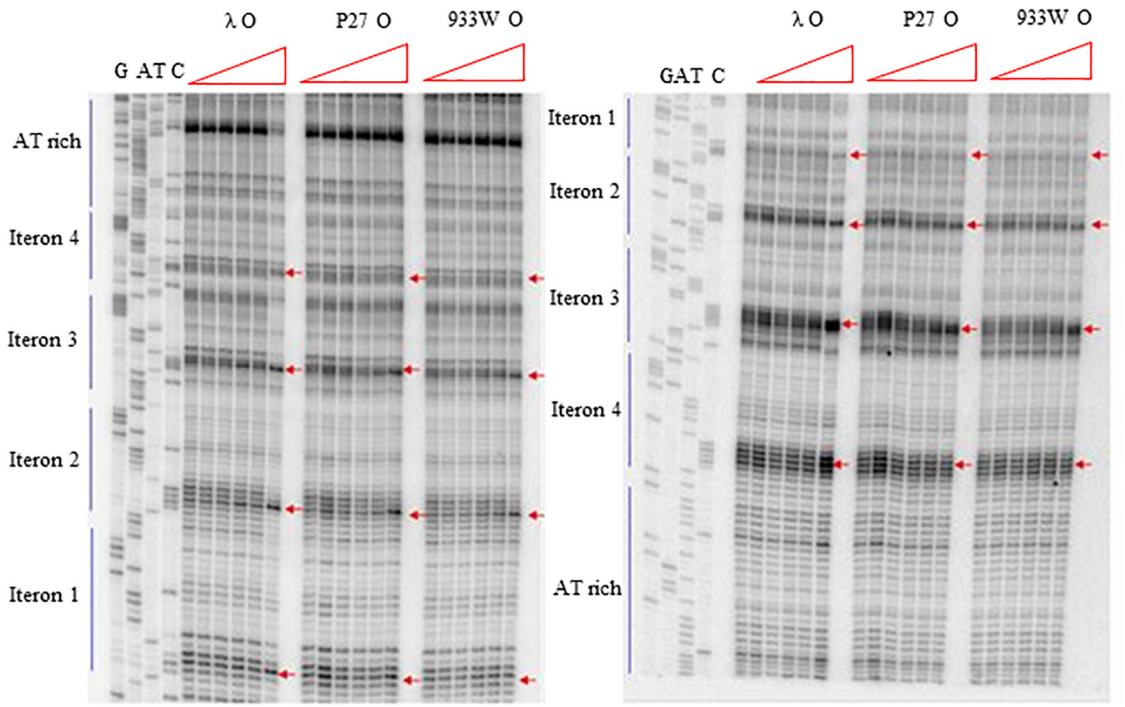


FIGURE 4 | Interactions of O proteins of phages λ, P27 and 933W with DNA templates containing four iterons, as assessed by DMS footprinting. Increasing molar ratios of DNA:O proteins, indicated in particular wells, were as follows: 1:0 (no O protein), 1:2, 1:4, 1:8, 1:16, and 1:32. Arrows indicate major changes (enhancement) in bands' intensities.

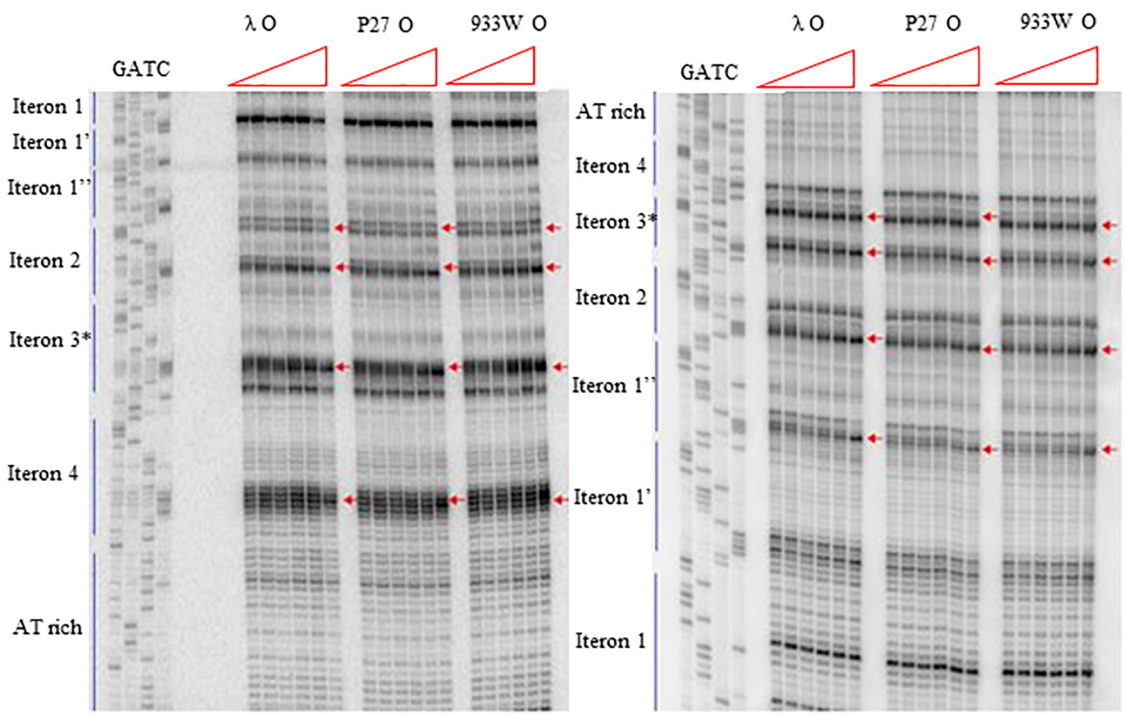


FIGURE 5 | Interactions of O proteins of phages λ, P27 and 933W with DNA templates containing six iterons, as assessed by DMS footprinting. Increasing molar ratios of DNA:O proteins, indicated in particular wells, were as follows: 1:0 (no O protein), 1:2, 1:4, 1:8, 1:16, and 1:32. Arrows indicate major changes (enhancement) in bands' intensities.

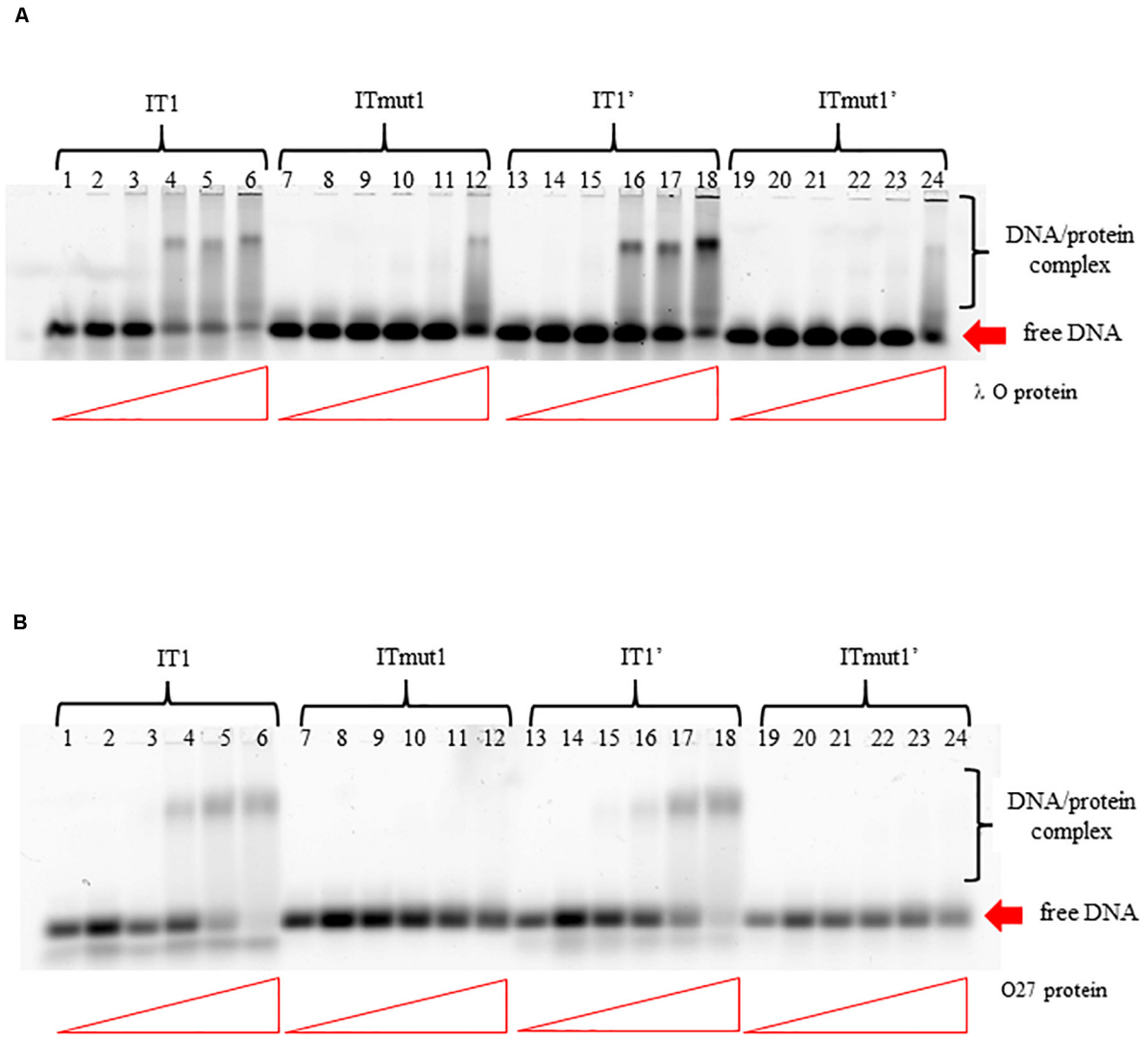


FIGURE 6 | Interactions of O proteins of phages λ panel **(A)** and P27 panel **(B)** with DNA templates containing various versions of iterons (IT), as assessed by EMSA. Templates with wild-type iterons 1 or 1′ (IT1 or IT1′, respectively) and with mutant iterons 1 or 1′, unable to bind O proteins (ITmut1 or ITmut1′, respectively) were used. Increasing molar ratios of DNA:O proteins, indicated in particular wells, were as follows: 1:2, 1:4, 1:8, 1:16, and 1:32. Lanes no. 1, 7, 13, and 19 contain no O protein.

double *dnaA46* Δ*seqA* mutants. These mutants were used since it was previously demonstrated that: (i) incompatibility between λ plasmids and *dnaA* mutations arises from both inefficient stimulation of transcription from the p_R promoter, responsible for transcriptional activation of the *origin*, by the mutant DnaA protein and the competition between the lambda P protein and the host DnaA and DnaC proteins for DnaB helicase, (ii) both mechanisms must be operative for the incompatibility, and (iii) the incompatibility is abolished in the absence of the *seqA* gene function (Glinkowska et al., 2001) as the *seqA* gene product modulates p_R activity (Słomińska et al., 2003a) and interplays with DnaA at this promoter (Słomińska et al., 2003b). We found that disruption of either interon 1 or iteron 1′ or both abolished ability of lambdoid plasmid to transform *E. coli* wild-type strain. The presence of mutation(s) in *dnaA, seqA* or both these genes did not change the plasmid inability to transform host cells (**Table 2**). We conclude, that both iteron 1 and iteron 1′ are necessary for efficient replication of lambdoid plasmids derived from Stx phages, despite only weak interactions of the latter iteron with the O protein.

In the lambdoid phage *origin* of replication, iterons are flanked by the AT-rich region (see **Figure 1A**). To test if the AT-rich sequence influence binding of the O protein to iterons, we have replaced the original AT-rich region with that containing 57% GC base pairs. We found that such a change did not influence efficiency of binding of O protein to iterons (**Figure 7**). Patterns of formed nucleoprotein complexes were also not influenced by replacement of the AT-rich region with a neutral sequence. This

TABLE 2 | Transformation efficiency of E. coli wild-type, dnaA, seqA, and dnaA seqA hosts by derivatives of p27cmr plasmids bearing mutated iteron sequences 1 and 1′ at 30°C. Results are presented as mean values of three independent experiments ± SD.

Plasmid	Efficiency of transformation (transformants/μg of plasmid DNA)			
	dnaA$^+$seqA$^+$	dnaA46 seqA$^+$	dnaA$^+$ ΔseqA	dnaA46 ΔseqA
p27R6K	$5.38 \pm 1.37 \times 10^4$	$3.79 \pm 0.39 \times 10^4$	$<10^1$	$7.00 \pm 3.49 \times 10^2$
p27mut1	$<10^1$	$<10^1$	$<10^1$	$<10^1$
p27mut1′	$<10^1$	$<10^1$	$<10^1$	$<10^1$
p27mut1 + 1′	$<10^1$	$<10^1$	$<10^1$	$<10^1$

was true for both templates, containing either four or six iterons (**Figures 7A,B**, respectively). The presence of two bands for free DNA (**Figure 7B**, lines 1–14) and the presence of six protein-DNA bands (line 26) can be explained by formation of secondary structured by ssDNA fragments.

Binding of phage λ O protein to single stranded DNA (ssDNA) has been demonstrated previously (Learn et al., 1997). Therefore, we have asked if all tested O proteins can interact with ssDNA of the AT-rich region. The analysis was performed using FRET technique. ssDNA was labeled with Cy3 and Cy5 fluorescent dyes forming the donor-acceptor pair. For the free ssDNA, donor excitation resulted in effective energy transfer and an increase in acceptor fluorescence emission concomitant with a decrease of the donor fluorescence. Binding of the O proteins caused reduction of FRET efficiency observed as an increase of the donor signal and diminution of the acceptor fluorescence. Results of these experiments indicated that O proteins of phages *λ*, P27 and 933W can efficiently interact with ssDNA, indeed (**Figure 8**). The fluorescence signals were similar in all experiments in which O proteins (from *λ*, P27 and 933W) were present (**Figure 8**), suggesting that the investigated mechanism does not differ between tested phages.

DISCUSSION

Since DNA replication of Stx phages appears to be necessary for effective production of Shiga toxins by STEC strains (Nejman-Faleńczyk et al., 2012; Nowicki et al., 2013; Balasubramanian et al., 2019), understanding of the regulation of this process is crucial not only for basic science but also for applications related to development of novel anti-STEC treatments. Although replication regions of genomes of Stx phages are highly similar to that of bacteriophage λ, there are important differences (summarized in **Figure 1** and **Supplementary Figures S1–S3**) which result in significantly altered regulation of Stx phage DNA replication relative to λ (Nejman et al., 2009, 2011; Nowicki et al., 2015). Therefore, we decided to investigate details of molecular interactions between the initiator O proteins of Stx phages with the *origin* regions, using purified proteins and nucleic acids.

Previously published results (Zahn and Blattner, 1985, 1987; Schnos et al., 1988) indicated that the 19 bp (imperfect) inverted repeat of the iteron is the minimal sequence required for efficient binding by the *λ* O protein. At low protein concentrations, only the inner two iterons were bound, while binding to all four iterons occurred at higher protein concentrations. The O protein binds to the replication origin in both linear and circular (relaxed) forms, but *origin* unwinding requires the negatively supercoiled DNA form introduced by the host gyrase. Results of our experiments indicated that O proteins of Stx bacteriophages bind specifically to iterons in the *origin* regions (**Figure 2**; see also Kozłowska et al., 2017), however, these interactions are weaker relative to binding of bacteriophage λ O protein to its specific DNA target sequences (**Figure 2**). Apparently, amino acid alterations in O proteins of Stx phages relative to the λ O protein, listed in **Figure 1C** and indicated in **Supplementary Figure S2**), are responsible for these differences in their affinity to iterons. Although each individual iteron can be bound by any of tested O proteins (encoded by λ, P27 or 933W phages) (**Figure 3**), when the whole *origin* region was considered, only four iterons were involved in effective interactions with O protein molecules, irrespective of the presence of either four or six such DNA sequences in the template (**Figures 4**, **5**). Nevertheless, two additional iterons (1′ and 1″) in *origin* of Stx phage can be bound by the O protein and are necessary for effective DNA replication (**Figure 6**, **Table 2**, and **Supplementary Figure S4**).

The AT-rich region appears to be dispensable for formation of the O-some structure (a specific nucleoprotein structure containing the O protein and iterons) (**Figure 7**). In fact, it was demonstrated (Dodson et al., 1986) that two iterons and the adjacent AT-rich region form the minimal DNA fragment competent for unwinding. Moreover, for *in vitro* replication of *ori*λ-based plasmids, only the two iterons proximal to the AT-rich region were required (Wickner and McKenney, 1987). Since this AT-rich tract is crucial for DNA unwinding during replication initiation, and all tested O protein variants revealed similar ability to interact with single stranded DNA rich in A and T residues (**Figure 8**), one may suggest that the mechanism of DNA stretching is the same for all tested phages.

The presence of additional two iterons in *origin* sequences of Stx phages relative to λ results also in appearance of additional 13 amino acid residues in corresponding O proteins, as iterons are located in the middle of the *O* gene (**Figure 1**). One might ask if such additional protein fragment can influence interaction between O and its DNA target. The secondary structure prediction of O proteins (**Supplementary Figure S2**) suggests that the additional protein fragments are not involved in interactions with DNA. Regarding the role of these two additional iterons in Stx phages, on one hand, it was reported that three iterons are sufficient for efficient replication of phage λ, as suggested by the normal growth and burst size of the mutant phage (Moore et al., 1981). On the other hand, additional iterons were found here to be necessary for transformation of P27-derived plasmids (**Table 2**). Therefore, one might propose that iterons 1′ and 1″ ensure either elevation of the local concentration of the O protein or stabilization of the O-some which can be crucial due to lower affinity of Stx O proteins to iterons, relative to λ O. Such a model is compatible with the already described mechanism of DNA-protein interactions which usually requires

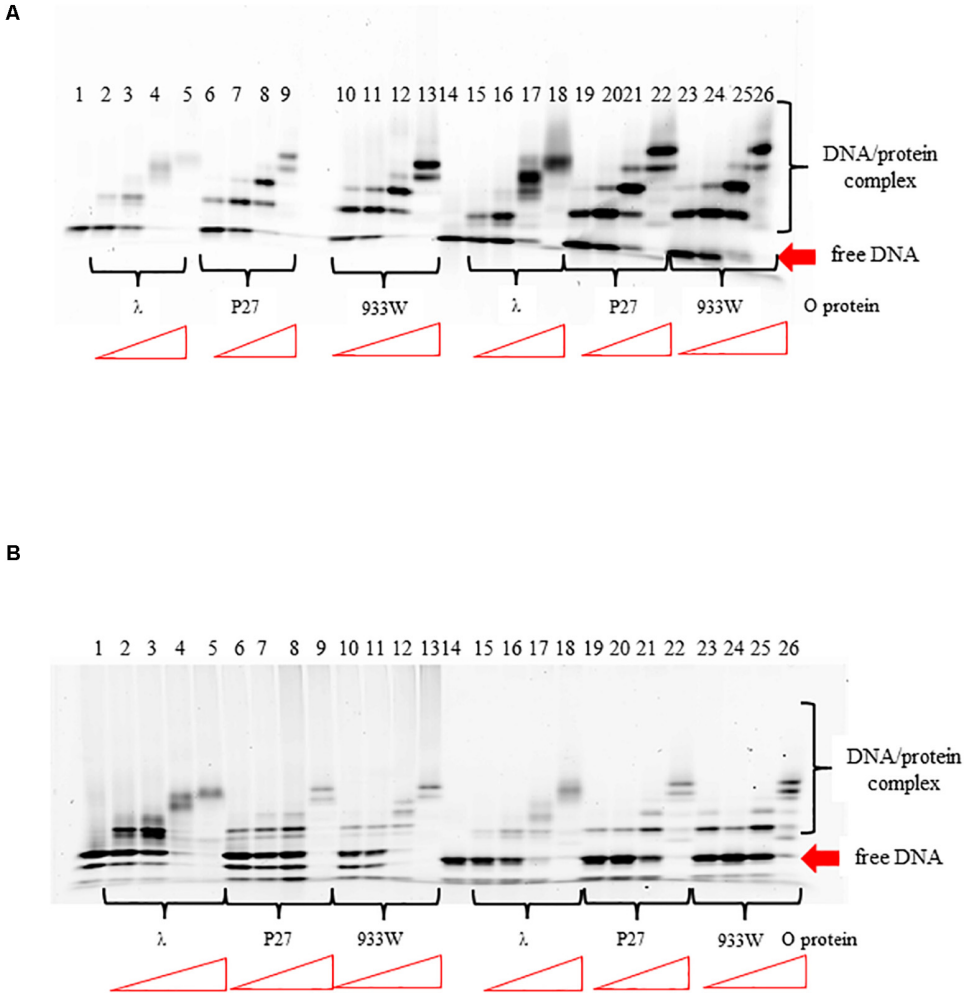


FIGURE 7 | Interactions of O proteins of phages λ., P27 and 933W with DNA templates containing four panel **(A)** or six panel **(B)** iterons together with the AT-rich region (lanes 1–13) or the AT-rich region replaced with a sequence contained 57% of GC base pairs (lanes 14–26), as assessed by EMSA. Increasing molar ratios of DNA:O proteins, indicated in particular wells, were as follows: 1:2, 1:4, 1:8, and 1:16. Lanes no. 1 and 14 contain no O protein.

a proximity of A residue stretch in DNA (Rohs et al., 2009), and such a motif is actually present in iterons located in the *origin* of lambdoid phage replication (see **Figure 1B**).

There is also a question about potential roles of hairpin and cruciform structures, potentially formed by the replication region of lambdoid phages. Roles of such structures in DNA replication have been discussed previously (Bikard et al., 2010). In fact, each iteron in lamboid *origin* is an inverted repeat, and in the sequences of the entire gene O, the biggest and most stable hairpin structures are encoded by the iterons 2 and 4, 1′ and 4, and 2 and 3 (**Supplementary Figure S3**). It appears that duplication of the iterons in Stx pahges may influence folding of the potential hairpin/cruciform structures. This may influence efficiency of binding of the O protein and formation of the O-some, and in fact, Stx replication regions are bound less efficiently than that of λ. Therefore changes in secondary DNA structures might be an additional role of additional iterons in Stx phages.

Recent studies indicated that DNA replication may significantly influence the major developmental decision of bacteriophage λ, whether to produce progeny phages and lyse the host cell or to lysogenize it (Shao et al., 2018). Since such decision is crucial for all lambdoid phages, one can speculate that λ and Stx phages evolved in somewhat different ways to produce optimal regulatory systems functioning in environments they occur in. Less stringent binding of the O protein to iterons, and the presence of additional iterons which can modulate formation of the O-some, possibly in response to various environmental conditions, might be profitable for Stx phages which have to quickly change the developmental pathway. Gastrointestinal tract of ruminants is their primary habitat in which STEC strains developed a specific Stx phage-dependent strategy to fight unicellular eukaryotic predators, like ciliates (Łoś et al., 2013). Therefore, ability to adjust the genetic switch between lysogenic and lytic developmental options rapidly may require more flexible regulation of DNA replication initiation, ensured

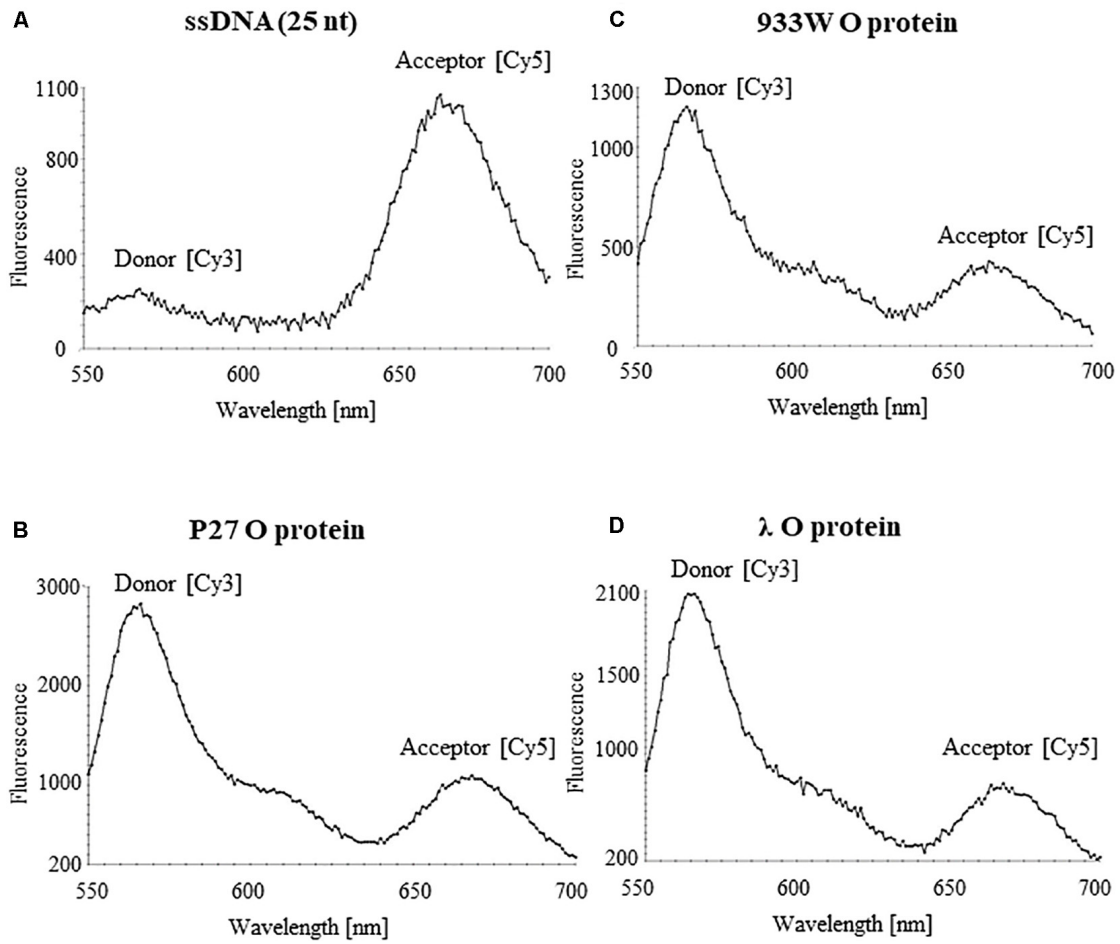


FIGURE 8 | Interactions of O proteins of phages P27 **(B)**, 933W **(C)**, and λ **(D)**, with ssDNA templates containing the AT-rich region as assessed by FRET. Results obtained with the sample containing only the ssDNA template are shown in **(A)**.

by modulatory effects of additional iterons and weaker binding of the O protein to them. On the other hand, phage λ that was discovered as a parasite of the *E. coli* K12 strain (Lederberg and Lederberg, 1953), a bacterium originally isolated from human stool and known to occur in human and animal intestine as well as in natural waters (Kuhnert et al., 1995; Na et al., 2006). Hence, this phage might prefer more strict developmental decisions which can be effectively made when employing stronger interactions between the O initiator protein and the replication *origin*. This strategy may ensure more unequivocal lysis-vs.-lysogenization decision which is compatible with the proposal of Shao et al. (2018) about the coupling of phage DNA replication to the λ developmental choice.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation, to any qualified researcher.

AUTHOR CONTRIBUTIONS

KK participated in planning experiments, performed most of them, participated in data interpretation, and drafting the manuscript. MG participated in planning experiments, data analysis, and revising the manuscript. LB participated in constructing of plasmids, performing EMSA, conducting experiments with mutant plasmids, performing DMS footprinting experiments, and revising the manuscript. LG participated in data analysis, preparation of figures, and revising the manuscript. JD assisted in EMSA experiments. GW was principal investigator of the project, supervised the work, and drafted the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Science Center (Poland) (project grant no. 2011/02/A/NZ1/00009).

ACKNOWLEDGMENTS

We are grateful to Anna Zawilak-Pawlik and Paweł Donczew for consultations and advices regarding DMS footprinting experiments.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2020.00207/full#supplementary-material

REFERENCES

Actis, G. C. (2014). The gut microbiome. *Inflamm. Allergy Drug Targets*. 13, 217–223.

Balasubramanian, S., Osburne, M. S., Brin-Jones, H., Tai, A. K., and Leong, J. M. (2019). Prophage induction, but not production of phage particles, is required for lethal disease in a microbiome-replete murine model of enterohemorrhagic *E. coli* infection. *PLoS Pathog.* 15:e1007494. doi: 10.1371/journal.ppat.1007494

Bikard, D., Loot, C., Baharoglu, Z., and Mazel, D. (2010). Folded DNA in action: hairpin formation and biological functions in prokaryotes. *Microbiol. Mol. Biol. Rev.* 74, 570–588. doi: 10.1128/MMBR.00026-10

Chen, J., Li, Y., Zhang, K., and Wang, H. (2018). Whole-genome sequence of phage-resistant strain *Escherichia coli* DH5α. *Genome Announc.* 6, : e00097–e-18.

Devleesschauwer, B., Pires, S. M., Young, I., Gill, A., Majowicz, S. E., and the study team. (2019). Associating sporadic, foodborne illness caused by Shiga toxin-producing *Escherichia coli* with specific foods: a systematic review and meta-analysis of case-control studies**. *Epidemiol. Infect.* 147, :e235. doi: 10.1017/S0950268819001183

Dodson, M., Echols, H., Wickner, S., Alfano, C., Mensa-Wilmot, K., Gomes, B., et al. (1986). Specialized nucleoprotein structures at the origin of replication of bacteriophage lambda: localized unwinding of duplex DNA by a six protein reaction. *Proc. Natl. Acad. Sci. U.S.A.* 83, 7638–7642. doi: 10.1073/pnas.83.20.7638

Glinkowska, M., Konopa, G., Węgrzyn, A., Herman-Antosiewicz, A., Weigel, C., Seitz, H., et al. (2001). The double mechanism of incompatibility between lambda plasmids and *Escherichia coli* dnaA(ts) host cells. *Microbiology* 147, 1923–1928. doi: 10.1099/00221287-147-7-1923

Goodsell, D. S., and Dickerson, R. E. (1994). Bending and curvature calculations in B-DNA. *Nucleic Acids Res.* 22, 5497–5503. doi: 10.1093/nar/22.24.5497

Kakoullis, L., Papachristodoulou, E., Chra, P., and Panos, G. (2019). Shiga toxin-induced haemolytic uraemic syndrome and the role of antibiotics: a global overview. *J. Infect.* 79, 75–94. doi: 10.1016/j.jinf.2019.05.018

Karmali, M. A. (2018). Factors in the emergence of serious human infections associated with highly pathogenic strains of shiga toxin-producing *Escherichia coli*. *Int. J. Med. Microbiol.* 308, 1067–1072. doi: 10.1016/j.ijmm.2018.08.005

Kintz, E., Byrne, L., Jenkins, C., McCarthy, N., Vivancos, R., and Hunter, P. (2019). Outbreaks of Shiga toxin-producing *Escherichia coli* linked to sprouted seeds, salad, and leafy greens: a systematic review. *J. Food Prot.* 82, 1950–1958. doi: 10.4315/0362-028X.JFP-19-014

Kozłowska, K. I., Tymecka-Mulik, J., and Węgrzyn, G. (2017). Purified Stx and λ phage initiator O proteins bind specifically to two different origins of replication in vitro. *Protein Expr. Purif.* 131, 16–26. doi: 10.1016/j.pep.2016.11.002

Krüger, A., and Lucchesi, P. M. (2015). Shiga toxins and stx phages: highly diverse entities. *Microbiology* 161, 451–462. doi: 10.1099/mic.0.000003

Kuhnert, P., Nicolet, J., and Frey, J. (1995). Rapid and accurate identification of *Escherichia coli* K-12 strains. *Appl. Environ. Microbiol.* 61, 4135–4139. doi: 10.1128/aem.61.11.4135-4139.1995

Learn, B. A., Um, S. J., Huang, L., and McMacken, R. (1997). Cryptic single-stranded-DNA binding activities of the phage lambda P and *Escherichia coli* DnaC replication initiation proteins facilitate the transfer of *E. coli* DnaB helicase onto DNA. *Proc. Natl. Acad. Sci. U.S.A.* 94, 1154–1159. doi: 10.1073/pnas.94.4.1154

Lederberg, E. M., and Lederberg, J. (1953). Genetic studies of lysogenicity in *Escherichia coli*. *Genetics* 38, 51–64.

Leimbach, A., Hacker, J., and Dobrindt, U. (2013). E. coli as an all-rounder: the thin line between commensalism and pathogenicity. *Curr. Top. Microbiol. Immunol.* 358, 3–32. doi: 10.1007/82_2012_303

Licznerska, K., Nejman-Faleńczyk, B., Bloch, S., Dydecka, A., Topka, G., Gąsior, T., et al. (2016). Oxidative stress in Shiga toxin production by enterohemorrhagic *Escherichia coli*. *Oxid. Med. Cell. Longev.* 2016, :3578368. doi: 10.1155/2016/3578368

Łoś, J. M., Łoś, M., Węgrzyn, A., and Węgrzyn, G. (2013). Altruism of Shiga toxin-producing *Escherichia coli*: recent hypothesis versus experimental results. *Front. Cell. Infect. Microbiol.* 2:166. doi: 10.3389/fcimb.2012.00166

Łoś, J. M., Łoś, M., and Węgrzyn, G. (2011). Bacteriophages carrying Shiga toxin genes: genomic variations, detection and potential treatment of pathogenic bacteria. *Future Microbiol.* 6, 909–924. doi: 10.2217/FMB.11.70

Mir, R. A., and Kudva, I. T. (2019). Antibiotic-resistant Shiga toxin-producing *Escherichia coli*: an overview of prevalence and intervention strategies. *Zoonoses Public Health*. 66, 1–13. doi: 10.1111/zph.12533

Moore, D. D., Denniston, K. J., and Blattner, F. R. (1981). Sequence organization of the origins of DNA replication in lambdoid coliphages. *Gene* 14, 91–101. doi: 10.1016/0378-1119(81)90151-7

Na, S. H., Miyanaga, K., Unno, H., and Tanji, Y. (2006). The survival response of *Escherichia coli* K12 in a natural environment. *Appl. Microbiol. Biotechnol.* 72, 386–392. doi: 10.1007/s00253-005-0268-3

Nejman, B., Łoś, J. M., Łoś, M., Węgrzyn, G., and Węgrzyn, A. (2009). Plasmids derived from lambdoid bacteriophages as models for studying replication of mobile genetic elements responsible for the production of Shiga toxins by pathogenic *Escherichia coli* strains. *J. Mol. Microbiol. Biotechnol.* 17, 211–220. doi: 10.1159/000242447

Nejman, B., Nadratowska-Wesołowska, B., Szalewska-Pałasz, A., Węgrzyn, A., and Węgrzyn, G. (2011). Replication of plasmids derived from Shiga toxin-converting bacteriophages in starved *Escherichia coli*. *Microbiology* 157, 220–233. doi: 10.1099/mic.0.042820-0

Nejman-Faleńczyk, B., Golec, P., Maciąg, M., Wegrzyn, A., and Węgrzyn, G. (2012). Inhibition of development of Shiga toxin-converting bacteriophages by either treatment with citrate or amino acid starvation. *Foodborne Pathog. Dis.* 9, 13–19. doi: 10.1089/fpd.2011.0980

Nowicki, D., Bloch, S., Nejman-Faleńczyk, B., Szalewska-Pałasz, A., Węgrzyn, A., and Węgrzyn, G. (2015). Defects in RNA polyadenylation impair both lysogenization by and lytic development of Shiga toxin-converting bacteriophages. *J. Gen. Virol.* 96, 1957–1968. doi: 10.1099/vir.0.000102

Nowicki, D., Kobiela, W., Węgrzyn, A., Wegrzyn, G., and Szalewska-Pałasz, A. (2013). ppGpp-dependent negative control of DNA replication of Shiga toxin-converting bacteriophages in *Escherichia coli*. *J. Bacteriol.* 195, 5007–5015. doi: 10.1128/JB.00592-13

Paletta, A. C. C., Castro, V. S., and Conte-Junior, C. A. (2020). Shiga toxin-producing and enteroaggregative *Escherichia coli* in animal, foods, and humans: pathogenicity mechanisms, detection methods, and epidemiology. *Curr. Microbiol.* 77, 612–620. doi: 10.1007/s00284-019-01842-1

Rohs, R., West, S. M., Sosinsky, A., Liu, P., Mann, R. S., and Honig, B. (2009). The role of DNA shape in protein-DNA recognition. *Nature* 461, 1248–1253. doi: 10.1038/nature08473

Schnos, M., Zahn, K., Inman, R. B., and Blattner, F. R. (1988). Initiation protein induced helix destabilization at the lambda origin: a prepriming step in DNA replication. *Cell* 52, 385–395. doi: 10.1016/s0092-8674(88)80031-x

Shao, Q., Cortes, M. G., Trinh, J. T., Guan, J., Balázsi, G., and Zeng, L. (2018). Coupling of DNA replication and negative feedback controls gene expression for cell-fate decisions. *iScience* 6, 1–12. doi: 10.1016/j.isci.2018.07.006

Słomińska, M., Konopa, G., Barańska, S., Wegrzyn, G., and Wegrzyn, A. (2003b). Interplay between DnaA and SeqA proteins during regulation of bacteriophage lambda pR promoter activity. *J. Mol. Biol.* 329, 59–68. doi: 10.1016/s0022-2836(03)00378-4

Słomińska, M., Konopa, G., Ostrowska, J., Kedzierska, B., Wegrzyn, G., and Wegrzyn, A. (2003a). SeqA-mediated stimulation of a promoter activity by facilitating functions of a transcription activator. *Mol. Microbiol.* 47, 1669–1679. doi: 10.1046/j.1365-2958.2003.03392.x

Szalewska-Pałasz, A., Wegrzyn, A., Herman, A., and Wegrzyn, G. (1994). The mechanism of the stringent control of lambda plasmid DNA replication. *EMBO J.* 13, 5779–5785. doi: 10.1002/j.1460-2075.1994.tb06916.x

Szambowska, A., Pierechod, M., Wegrzyn, G., and Glinkowska, M. (2011). Coupling of transcription and replication machineries in λ DNA replication initiation: evidence for direct interaction of *Escherichia coli* RNA polymerase and the λO protein. *Nucleic Acids Res.* 39, 168–177. doi: 10.1093/nar/gkq752

Taylor, K., and Wegrzyn, G. (1995). Replication of coliphage lambda DNA. *FEMS Microbiol. Rev.* 17, 109–119. doi: 10.1111/j.1574-6976.1995.tb00192.x

Taylor, R. G., Walker, D. C., and McInnes, R. R. (1993). E. coli host strains significantly affect the quality of small scale plasmid DNA preparations used for sequencing. *Nucleic Acids Res.* 21, 1677–1678. doi: 10.1093/nar/21.7.1677

Węgrzyn, G., Licznerska, K., and Węgrzyn, A. (2012). Phage λ - new insights into regulatory circuits. *Adv. Virus Res.* 82, 155–178. doi: 10.1016/B978-0-12-394621-8.00016-9

Weigel, C., and Seitz, H. (2006). Bacteriophage replication modules. *FEMS Microbiol. Rev.* 30, 321–381. doi: 10.1111/j.1574-6976.2006.00015.x

Wickner, S., and McKenney, K. (1987). Deletion analysis of the DNA sequence required for the in vitro initiation of replication of bacteriophage lambda. *J. Biol. Chem.* 262, 13163–13167.

Yanisch-Perron, C., Vieira, J., and Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* 33, 103–119. doi: 10.1016/0378-1119(85)90120-9

Zahn, K., and Blattner, F. R. (1985). Binding and bending of the lambda replication origin by the phage O protein. *EMBO J.* 4, 3605–3616. doi: 10.1002/j.1460-2075.1985.tb04124.x

Zahn, K., and Blattner, F. R. (1987). Direct evidence for DNA bending at the lambda replication origin. *Science* 236, 416–422. doi: 10.1126/science.2951850

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Kozłowska, Glinkowska, Boss, Gaffke, Deptuła and Węgrzyn. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Molecular Biosciences

REVIEW
published: 15 October 2020
doi: 10.3389/fmolb.2020.586497

Role of Regulated Proteolysis in the Communication of Bacteria With the Environment

*Sarah Wettstadt and María A. Llamas**

Department of Environmental Protection, Estación Experimental del Zaidín-Consejo Superior de Investigaciones Científicas, Granada, Spain

For bacteria to flourish in different niches, they need to sense signals from the environment and translate these into appropriate responses. Most bacterial signal transduction systems involve proteins that trigger the required response through the modification of gene transcription. These proteins are often produced in an inactive state that prevents their interaction with the RNA polymerase and/or the DNA in the absence of the inducing signal. Among other mechanisms, regulated proteolysis is becoming increasingly recognized as a key process in the modulation of the activity of these signal response proteins. Regulated proteolysis can either produce complete degradation or specific cleavage of the target protein, thus modifying its function. Because proteolysis is a fast process, the modulation of signaling proteins activity by this process allows for an immediate response to a given signal, which facilitates adaptation to the surrounding environment and bacterial survival. Moreover, regulated proteolysis is a fundamental process for the transmission of extracellular signals to the cytosol through the bacterial membranes. By a proteolytic mechanism known as regulated intramembrane proteolysis (RIP) transmembrane proteins are cleaved within the plane of the membrane to liberate a cytosolic domain or protein able to modify gene transcription. This allows the transmission of a signal present on one side of a membrane to the other side where the response is elicited. In this work, we review the role of regulated proteolysis in the bacterial communication with the environment through the modulation of the main bacterial signal transduction systems, namely one- and two-component systems, and alternative σ factors.

Keywords: bacteria, proteolysis, signaling, gene regulation, transcription factor, sigma factor, two-component system

OPEN ACCESS

Edited by:
Chew Chieng Yeo,
Sultan Zainal Abidin University, Malaysia

Reviewed by:
Iain Lamont,
University of Otago, New Zealand
Pierre Cornelis,
Vrije University Brussel, Belgium

***Correspondence:**
María A. Llamas
marian.llamas@eez.csic.es

Specialty section:
This article was submitted to Molecular Recognition, a section of the journal Frontiers in Molecular Biosciences

Received: *23 July 2020*
Accepted: *22 September 2020*
Published: *15 October 2020*

Citation:
Wettstadt S and Llamas MA (2020) Role of Regulated Proteolysis in the Communication of Bacteria With the Environment. Front. Mol. Biosci. 7:586497. doi: 10.3389/fmolb.2020.586497

INTRODUCTION

Proteolysis was long considered solely a mechanism of protein degradation to recycle amino acids in a slow and somewhat non-selective way. However, recent advances in the field have evidenced that in biological systems, proteolysis serves both as a cellular housekeeper (general proteolysis) and a modulator of regulatory pathways (regulated proteolysis) (Ehrmann and Clausen, 2004; Dougan, 2013). General proteolysis is important for the removal of misfolded or damaged proteins in a relatively non-specific manner, helping to preserve cell physiology. On the contrary, regulated proteolysis can produce the specific cleavage but also the complete degradation of selected proteins

in response to intra- or extracellular signals. Regulated protein degradation is also referred to as processive proteolysis and allows a cell to get rid of a given protein, while protein cleavage is referred to as non-processive proteolysis or processing and produces a defined change in the activity of that protein (Jenal and Hengge-Aronis, 2003). Many proteins, both in eukaryotes and prokaryotes, including receptors, kinases, transcription factors, and structural components, are proteolytically altered to gain activity or to modify their functions (Brown et al., 2000; Gur et al., 2011; Konovalova et al., 2014). Proteolysis is a fast process able to alter the function of a protein instantaneously without the delay associated with activation or inhibition of transcription or translation mechanisms. Furthermore, proteolysis is able to eliminate a protein from a cell when it is no longer needed on a much faster time scale than for example the dilution of the protein as consequence of cell division would allow. As such a fast mechanism, regulation of protein activity by proteolysis is advantageous when a fast response is needed. This is especially the case during signal transduction that requires a quick response to a given signal in order for the cell to adapt and survive in the surrounding environment. In addition to mediating a fast response, regulated proteolysis is a sophisticated tool to overcome the biological problem of signal transfer between two cellular compartments. The field of cell signaling was expanded by the discovery that transmembrane proteins can be cleaved on either side or within the plane of the membrane to liberate a cytosolic domain or protein able to modulate gene transcription. This regulatory paradigm know as regulated intramembrane proteolysis (RIP) is commonly found in compartmental membranes in eukaryotes and the cytoplasmic membrane of prokaryotes, and allows the immediate transfer of a signal from an extracytosolic compartment to the cytosol (Brown et al., 2000). We review in this work the role of regulated proteolysis in the modulation of the main bacterial signal transduction systems, including one- and two-component systems, and alternative σ factors.

PROTEASES INVOLVED IN REGULATED PROTEOLYSIS IN BACTERIA

Proteolysis is typically achieved by proteases (or peptidases), a group of enzymes that hydrolyze peptide bonds and thus breakdown proteins or peptides. Proteases are classified according to their active site residue or ion that carries out catalysis, and they include serine, threonine, cysteine, glutamic, asparagine, aspartic, and metallo- proteases (a comprehensive classification of proteases is available through the MEROPS database[1] Rawlings et al., 2014). Proteases do not attack their substrates at random but display high degrees of specificity that depend on a variety of factors, such as co-localization of the protease and the substrate, the regulation of protease activity by activation/inhibition processes, and the accessibility of the substrate cleavage site for the protease (Ehrmann and Clausen, 2004). Notably, specific cues often modulate these factors. For example, specific signals can modify protease activity by changing the structural properties of the active site either exposing or hiding it, or by activating adaptor proteins that feed the substrate to the protease. The accessibility of the protease to the cleavage site of the substrate can be also modified in response to specific signals that for example produces the unfolding of the substrate or the loss of an interaction partner.

General and regulated processive proteolysis of cytoplasmic proteins is often carried out by chaperone-protease complexes (reviewed in Kirstein et al., 2009; Gur et al., 2011). The chaperone belongs to the AAA + protein family and uses ATP hydrolysis to deliver the substrate to the protease. Examples of chaperones include ClpA, ClpX, ClpC, and HslU, which associate with ClpP and HslV serine and threonine proteases, respectively (**Table 1**). These complexes (e.g., ClpAP, ClpXP, ClpCP, HslUV) are known as ATP-dependent proteases. This group of proteases also includes the Lon and FtsH serine and metallo- proteases, respectively (**Table 1**), that contain the AAA + ATPase and the protease functions in the same polypeptide.

The proteases involved in regulated non-processive proteolysis are more diverse and often substrate-specific. Increasingly recognized have been the intramembrane cleaving proteases (known as I-Clips) for their unique abilities to cleave peptide bonds within cellular membranes and their function in RIP of signaling pathways (reviewed in Urban and Freeman, 2002; Makinoshima and Glickman, 2006). Bacterial I-Clips are divided into aspartyl, serine, and zinc metallo- proteases. Aspartyl intramembrane proteases are exemplified by the eukaryotic presenilin (**Table 1**), which has been linked to Alzheimer's disease. Serine intramembrane proteases include the rhomboid proteases, which contain six transmembrane domains and an active site cavity that opens to the periplasm, and are represented by *Escherichia coli*

TABLE 1 | Proteases discussed in this review.

Protease families[a]	Subfamilies/ examples	MEROPS family
Aspartyl proteases		
Membrane-inserted endopeptidases	Presenilin	A22
Aspartyl endopeptidase	SpoIIGA	A36
Metalloproteases		
ATP-dependent metallo-endopeptidases	FtsH	M41
Intramembrane metallo-endopeptidases	RseP, SpoIVFB, RasP	M50
Membrane-bound bacterial endopeptidases	BlaR1, MecR1	M56
Metallo-endopeptidases	IrrE	M78
Serine proteases		
Serine endopeptidases	DegS, DegP	S1
Clp endopeptidases	ClpP	S14
ATP-dependent serine endopeptidases	Lon	S16
Signal peptidases	SipS, SipT, SipT	S26
C-terminal processing peptidases	Prc	S41
Membrane-bound serine endopeptidases	Rhomboid (e.g., GlpG)	S54
Threonine proteases		
Component peptidases of the proteasome	HslV	T1

[a] *According to the MEROPS database (https://www.ebi.ac.uk/merops/).*

[1] https://www.ebi.ac.uk/merops/

GlpG (**Table 1**; Wang et al., 2006). The zinc intramembrane metalloprotease family includes the so-called site-2 proteases (S2P) represented in bacteria by *E. coli* RseP and *Bacillus subtilis* SpoIVFB (**Table 1**) (reviewed in Kroos and Akiyama, 2013). S2P proteases share a conserved core domain containing at least three transmembrane domains, with the first and the third transmembrane segments containing the HExxH and LDG active site motifs, respectively (Kinch et al., 2006). Importantly, S2P proteases play a key role in RIP of several signaling pathways in bacteria, including one-component systems and those involving alternative σ factors (see below). These RIP cascades usually involve a site-1 protease that produces the substrate for the S2P protease. A site-1 cleavage is likely required because SP2 proteases, like RseP, contain PDZ domains that act as size-exclusion filters that only allow truncated, and thus smaller forms of the substrate, to enter the catalytic site of the protease (Hizukuri et al., 2014). Identified site-1 proteases involved in RIP of bacterial signaling pathways include the DegS and DegP serine endoproteases, SipS and SipT serine signal peptidases, and the Prc C-terminal processing serine protease (**Table 1**) (described below).

REGULATED PROTEOLYSIS IN MODULATING THE ACTIVITY OF ONE-COMPONENT SYSTEMS AND TRANSCRIPTION FACTORS

One-component systems are the simplest structures for sensing cues and the predominant signaling mechanisms in bacteria (Ulrich et al., 2005). Canonical one-component systems are composed of a single cytoplasmic protein that carries out both signal recognition and response. These functions are performed by two different domains within the protein: the signal is recognized by the signal sensing or input domain, and the response is performed by the output domain, which often is a helix-turn-helix (HTH) domain able to bind to DNA thus activating or repressing gene transcription (Ulrich et al., 2005). Therefore, one-component proteins are also referred to as transcription factors or transcriptional regulators. According to their cytosolic location, one-component proteins mainly respond to signals produced in the cytosol or small molecules able to diffuse through the membrane. Signal input usually induces a conformational change that exposes the output domain allowing its binding to DNA. Activity of one-component proteins can be modulated by regulated proteolysis, which thus controls transcription of target genes in a fast but irreversible manner.

Both regulated degradation and RIP of one-component proteins have been identified. Regulated degradation in response to specific signals usually occurs in the cytosol by the action of ATP-dependent proteolytic complexes (i.e., Clp). A well-studied example is the *Bacillus subtilis* Spx transcription factor, which controls thiol homeostasis, heat and oxidative response, and competence through both transcriptional activation and repression. Activity and concentration of Spx after exposure to stress is controlled by both the redox state of Spx and its ClpXP-mediated regulated proteolysis (Zuber, 2004; **Figure 1A**). The latter occurs through the activity of the YjbH adaptor protein, which in absence of stress binds to the C-terminus of Spx and delivers Spx to the ClpXP protease complex for degradation (Chan et al., 2014; **Figure 1A**). Since YjbH is highly prone to misfolding, under heat stress YjbH quickly aggregates rendering it unable to bind and deliver Spx to ClpXP (Engman and von Wachenfeldt, 2015). Moreover, under oxidative and disulfide stress, disulfide bond formation occurs in Spx, ClpXP and YjbH, leading to structural changes that renders Spx highly active, but inactivates ClpXP and promotes YjbH aggregation, thus avoiding Spx degradation (Rojas-Tapias and Helmann, 2018; **Figure 1A**). Similarly, the transcription factor FixK2 of the nitrogen fixing soybean endosymbiont *Bradyrhizobium japonicum* undergoes Clp-mediated regulated proteolysis in response to intracellular oxygen concentrations. Under microoxic conditions, $FixK_2$ promotes transcription of genes required to adapt to low oxygen concentrations. In oxidative stress conditions, a cysteine near the DNA-binding domain of $FixK_2$ is oxidized producing a conformational change that exposes the C-terminal 12 amino acids and renders $FixK_2$ prone to ClpAP-mediated proteolysis (Bonnet et al., 2013).

Interestingly, some transcription factors are coupled with proteases that contain sensor domains and these two proteins form signal transduction pathways that modify gene transcription through regulated proteolysis. Such a signaling circuit is present in *Deinococcus* species, an extreme radiation-tolerant bacterium that employs the DrdO transcriptional repressor and the IrrE metalloprotease to respond to radiation and desiccation (Ludanyi et al., 2014). After exposure to radiation, signal transduction within IrrE occurs likely via its C-terminal GAF-like sensor domain resulting in activation of its N-terminal zinc-coordinating peptidase-like domain (Vujicic-Zagar et al., 2009). The peptidase domain then cleaves the DrdO repressor thus allowing transcription of *recA* and other DNA repair machine genes to adapt to radiation stress (Ludanyi et al., 2014).

Coupling can also occur between cytosolic transcription factors and membrane-bound proteases. In this case, the protease contains an extracytosolic signal sensing domain, which allows the one-component protein to respond to extracellular cues. As such, the activities of the transcriptional repressors BlaI of *Staphylococcus aureus* and MecI of methicillin-resistant *S. aureus* (MRSA) are controlled by the membrane-bound proteases BlaR1 and MecR1, respectively, in response to β-lactam antibiotics present on the outer surface of the bacterium (Zhang et al., 2001). BlaR1 and MecR1 are transmembrane proteins with a cytosolic N-terminal zinc metalloprotease domain and a C-terminal extracellular domain that resembles class-D β-lactamases and functions as a β-lactam sensor (Alexander et al., 2020). Presence of β-lactam antibiotics trigger acetylation of an active site serine located in the C-terminal sensor domains of BlaR1 and MecR1. Acetylation is transduced to the cytoplasmic domain leading to the activation of the metalloprotease domain and the degradation of the BlaI/MecI repressor. However, the mechanism behind this process is still unclear but it is likely similar for both repressors. It has been shown that acetylation of BlaR1

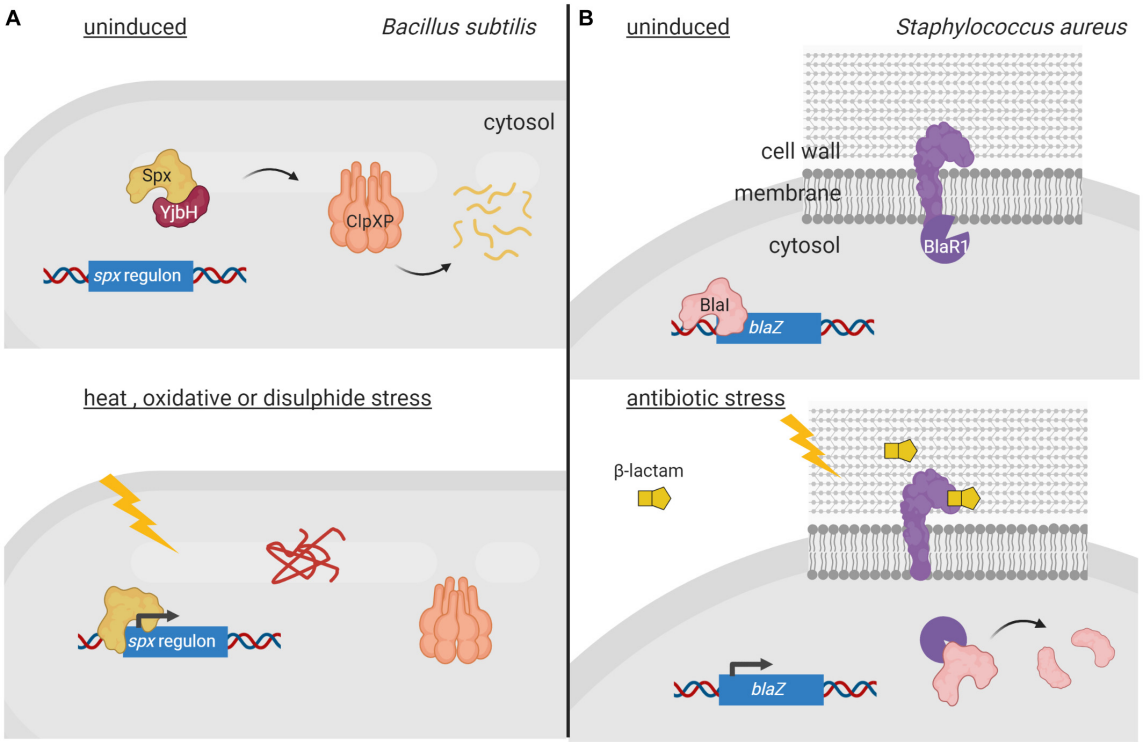


FIGURE 1 | Regulated proteolysis of one-component proteins and transcription factors. **(A)** In non-stressed conditions, the transcription factor Spx (orange) is sequestered by YjbH (red) which delivers it to ClpXP (pink) for degradation. Upon heat, oxidative or disulfide stress, YjbH aggregates and releases Spx. Conformation change within ClpXP renders the protease unable to bind and digest Spx, resulting in transcription of the *spx* regulon. **(B)** In non-induced conditions, the membrane-bound protease BlaR1 (purple) is sequestered to the membrane while the repressor BlaI (orange) inhibits gene transcription. Extracellular presence of β-lactams triggers BlaR1 autocleavage and the release of its protease domain into the cytosol. BlaR1 cleaves BlaI thus terminating its inhibiting activity. This allows transcription of *blaZ*, among other genes, and thus production of β-lactamase that degrades the antibiotic. Figure produced with BioRender (BioRender.com).

results in its autocleavage and release of the metalloprotease domain into the cytoplasm where it degrades BlaI (Llarrull et al., 2011; **Figure 1B**). A more recent model suggests that acetylation of MecR1 produces sterical overlap within the sensor domain and dislodgement of the last transmembrane helix that likely opens the cavity of the metalloprotease domain enabling substrate accessibility and catalysis (Belluzo et al., 2019). Nevertheless, degradation of the BlaI and MecI repressors by the proteases allows the transcription of (among others) the β-lactamase-encoding *blaZ* and the penicillin-binding protein PBP2a-encoding *mecA* genes, respectively, that confer resistance to β-lactam antibiotics (Llarrull et al., 2011; Belluzo et al., 2019; **Figure 1B**).

There are also membrane-bound one-component proteins that contain an extracytosolic signal-sensing domain and a cytosolic output domain, and regulated proteolysis is essential in activating these regulatory proteins. As such, the activities of the membrane-bound transcription activators ToxR and TcpP of the human pathogen *Vibrio cholerae* are controlled by both RIP and cytosolic regulated degradation. Under virulence—promoting conditions, e.g., presence of bile salts from the host, ToxR and TcpP, together with their respective interacting partners ToxS and TcpH, form a complex in the membrane that promotes the transcription of the *toxT* gene (Häse and Mekalanos, 1998). This gene encodes the ToxT transcriptional regulator, which is the direct activator of the two main virulence factors of *V. cholerae*, the cholera toxin and the toxin-co-regulated pilus (DiRita et al., 1991). Under conditions unfavorable for virulence gene activation, ToxR and TcpP undergo RIP. The RIP cascade of TcpP occurs in response to yet unknown signals that disrupt the TcpH/TcpP interaction. This exposes the periplasmic domain of TcpP to the C-terminal processing protease Prc that performs the site-1 cleavage of TcpP, which is followed by a site-2 transmembrane cleavage by the RseP protease (Matson and DiRita, 2005; Teoh et al., 2015). ToxR undergoes RIP in response to nutrient limitation at alkaline pH, which results in the reduction of two periplasmic cysteines of ToxR and the release of its interacting partner ToxS (Almagro-Moreno et al., 2015b; Lembke et al., 2018). This exposes ToxR to the action of the periplasmic site-1 serine proteases DegP and DegS, which produce the substrate for the site-2 protease RseP (Almagro-Moreno et al., 2015a,b). RIP of ToxR promotes the entrance of *V. cholerae* into a dormant environmentally persistent state (Almagro-Moreno et al., 2015a). Another membrane-bound one-component protein subjected to RIP is CadC of the pathogen *Salmonella enterica sv.* Typhimurium. CadC is an acid-sensing regulator that activates transcription of genes that contribute to the acid tolerance response of this pathogen (Lee et al., 2008).

In response to low pH and lysine, the CadC periplasmic signal sensing domain is degraded by a yet unknown protease resulting in accumulation of the N-terminal DNA-binding domain in the cytoplasm where it can bind its target promoters. This leads to the activation of the acid stress response with production of the outer membrane porins OmpC and OmpF, among others (Lee et al., 2008).

MODULATION OF TWO-COMPONENT SYSTEM ACTIVITY BY REGULATED PROTEOLYSIS

Two-component systems (TCSs) represent very efficient signal transduction pathways that allow bacteria to sense environmental changes and adapt to them in an appropriate manner. A canonical TCS is composed of a cytoplasmic membrane-bound histidine kinase that recognizes a specific stimulus, and a cytosolic response regulator that mediates the response (Stock et al., 2000). Activation of this signaling system in response to the inducing signal occurs via a phosphorelay cascade that leads to a His autophosphorylation of the histidine kinase and a subsequent Asp phosphorylation of the response regulator. Response regulators are DNA-binding proteins that modify gene transcription upon phosphorylation, thus translating the signal sensed by the histidine kinase into a response. Chemical ligands, environmental cues, phosphatases and auxiliary proteins directly adjust the phosphorylation levels of histidine kinases (Mascher et al., 2006). All these regulatory modes are reversible in controlling histidine kinase activity; however, irreversible proteolytic modification of histidine kinases also occurs. For example, in *Xanthomonas campestris*, proteolysis of the sensor histidine kinase VgrS is the signal required to modulate the DNA-binding activity of the cognate response regulator VgrR. In response to osmotic stress, the periplasmic PDZ-domain containing protease Prc directly binds to VgrS and cleaves its N-terminal periplasmic domain between residues Ala9 and Gln10 (Deng et al., 2018; **Figure 2A**). This terminates VgrS autokinase activity and halts VgrR phosphorylation. Unphosphorylated VgrR triggers transcription of stress-response genes required for the bacterium to resist osmostress. VgrR also promotes transcription of the *prc* gene, which results in a positive feedback loop within the regulatory cascade (Deng et al., 2018). The *X. campestris* VgrRS TCS is unusual for two reasons. First, the unphosphorylated form of VgrR binds to DNA with considerably higher affinity than the phosphorylated form (Deng et al., 2018), in contrast to most response regulators that are able to bind their target DNA only in their phosphorylated form (Stock et al., 2000). Second, while Prc is considered a C-terminal processing protease in *E. coli* and

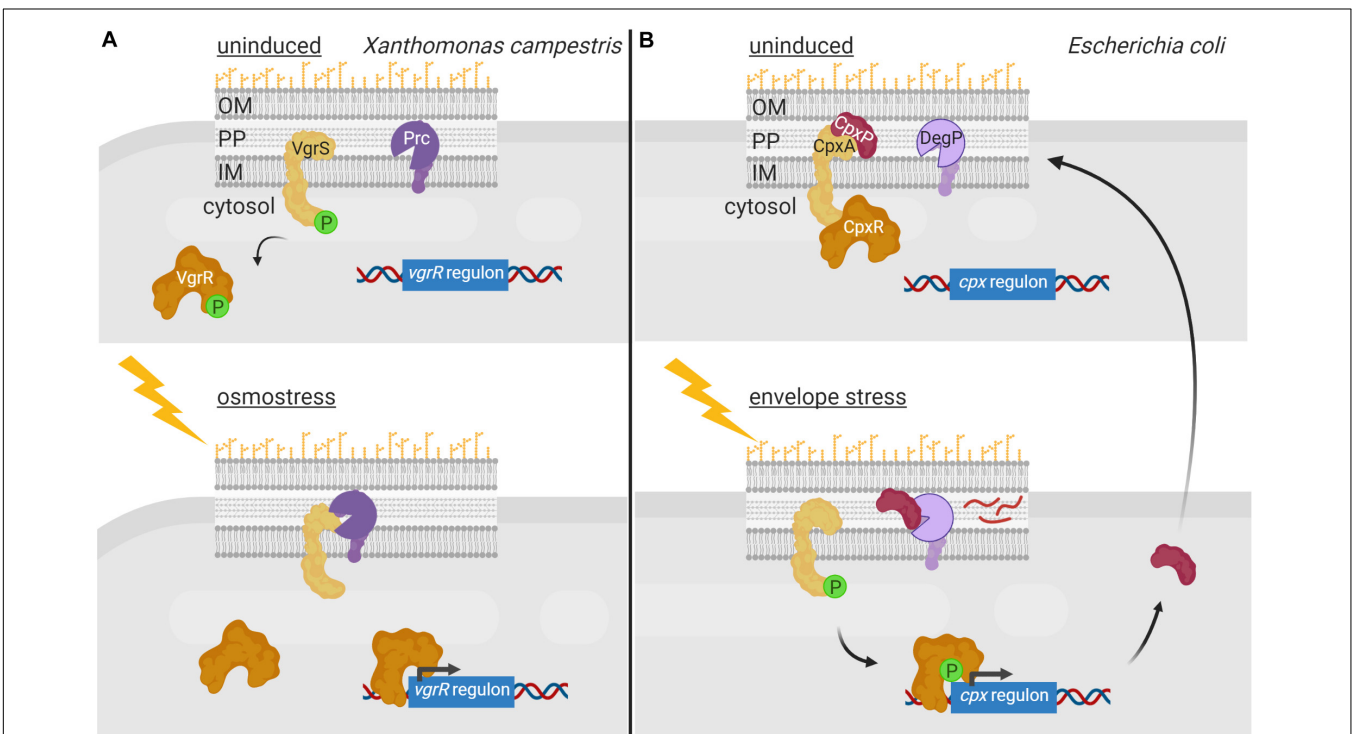


FIGURE 2 | Regulated proteolysis of two-component systems. **(A)** In the uninduced state, the histidine kinase VgrS (yellow) is active and phosphorylates its cognate response regulator VgrR (orange) diminishing its affinity for the DNA. Upon osmostress, VgrS is cleaved by the periplasmic protease Prc (purple), which inhibits VgrS autokinase activity and VgrR phosphorylation. This increases the affinity of VgrR for its target promoters resulting in transcription of genes required to cope with osmostress. **(B)** In non-stressed condition, the histidine kinase CpxA (yellow) is bound to periplasmic regulator CpxP (red), which inhibits its kinase activity. Envelope stress triggers the degradation of CpxP by DegP (purple) and the activation of CpxA, which phosphorylates the response regulator CpxR (orange) triggering a transcription program to combat the envelope stress. Figure produced with BioRender (BioRender.com).

other Gram-negative bacteria (Chueh et al., 2019), it processes VgrS N-terminally (Deng et al., 2018). It is at present unclear how Prc is activated under osmostress in *X. campestris*. In *E. coli*, the PDZ domain of Prc stimulates protease activity by directly interacting with the inducing substrate (Chueh et al., 2019). However, in *X. campestris*, the Prc PDZ domain does not interact with the VgrS sensor domain (Deng et al., 2018). This difference could be the reason why Prc^{Eco} recognizes and cleaves its substrates C-terminally, while Prc^{Xca} cleaves within the N-terminal end of VgrS.

Regulated proteolysis also modulates activity of the *E. coli* CpxAR TCS that responds to cell envelope stress. The activity of this signaling system is controlled by the periplasmic regulator CpxP, which interacts with the CpxA histidine kinase preventing the activation of the CpxAR TCS pathway (Raivio and Silhavy, 1997; **Figure 2B**). In high-salt conditions, CpxP dissociates from CpxA while in conditions that lead to envelope stress or elevated pH, CpxP is degraded by the periplasmic protease DegP (Isaac et al., 2005; Tschauner et al., 2014; **Figure 2B**). CpxA liberation as well as changes in the lipid bilayer properties of the cytoplasmic membrane, activates CpxA autokinase function (Keller et al., 2015). This results in phosphorylation and thus activation of the CpxR response regulator (Isaac et al., 2005; **Figure 2B**). CpxR triggers the transcription of the *cpx* regulon, which includes gene products necessary to combat envelope stress (Isaac et al., 2005). Furthermore, the *cpxP* gene is part of this regulon and CpxP production allows shutting down the CpxAR pathway at the same time the cell is being relieved from the envelope stress. Similarly, proteolysis of the periplasmic regulator ExoR of the Gram-negative *Sinorhizobium meliloti* modulates activity of the ExoS/ChvI TCS, which controls the switch from free-living to invading cells. ExoR retains the membrane-bound sensor kinase ExoS in an off state by protecting it from degradation (Lu et al., 2012; Wiech et al., 2015). As soon as ExoR is cleaved within its N-terminus by yet unknown proteases, the ExoS histidine kinase domain is activated leading to ChvI phosphorylation (Lu et al., 2012; Wiech et al., 2015). This results in expression of ExoR as well as genes to shift toward a free-living lifestyle.

REGULATED PROTEOLYSIS IN THE MODULATION OF ALTERNATIVE σ FACTOR ACTIVITY

Another important modulator of gene transcription in bacteria is the σ subunit of the RNA polymerase (RNAP). This dissociable subunit contains most promoter recognition determinants and directs the RNAP to the promoter region of the genes to be transcribed (Ishihama, 2000). Most bacteria contain a primary σ factor to transcribe general functions and a number of alternative σ factors to transcribe functions required only under specific conditions (Paget, 2015). Thus, the first step in the regulation of gene expression to shift toward a specific cell response often occurs through the substitution of the σ subunit of the RNAP (Ishihama, 2000). There are four different structural and functional groups of bacterial σ factors, with group I comprising primary σ factors and groups II to IV alternative σ factors (Paget, 2015). Group IV contains the so-called extracytoplasmic function σ factors (σ^{ECF}), which is the largest and more diverse group, and is considered the third signaling mechanism in bacteria (Mascher, 2013; Sineva et al., 2017). Activity of alternative σ factors, especially that of σ^{ECF} factors, is often controlled through an inhibitory interaction with an anti-σ factor (Paget, 2015). Anti-σ factors keep the σ factor sequestered preventing its binding to the RNAP and only releasing it in response to a specific signal. There are cytosolic anti-σ factors that respond to intracellular signals although most anti-σ factors are single-pass transmembrane proteins that respond to extracytosolic signals. Importantly, regulated proteolysis plays a central role in the modulation of the activity of alternative σ factors by controlling the amount of σ factor in the cell or by allowing its release from the anti-σ factor in response to specific cues. The alternative σ factor, its cognate anti-σ factor or both can be subjected to regulated proteolysis.

Regulated Proteolysis of Alternative σ Factors

Regulated proteolysis of alternative σ factors is for example used to prevent competition with the primary σ factor for binding to the RNAP during conditions in which the alternative σ factor is not necessary. This occurs in *E. coli* with the group II σ factor σ^{RpoS} that is the master regulator of the general stress response and essential for bacterial survival in the stationary phase as well as under a variety of stress conditions. σ^{RpoS} is always produced in the cell at basal levels but in exponentially growing bacteria it is hardly detectable due to proteolytic degradation by the ClpXP complex (Hengge, 2009). However, when cells enter the stationary phase, σ^{RpoS} accumulates and competes with the primary σ^{70} factor for binding to RNAP. σ^{RpoS} accumulation depends on different regulatory mechanisms acting at different levels, such as induction of *rpoS* gene transcription, increase of *rpoS* translation, and reduction of σ^{RpoS} proteolysis (Hengge, 2009). Besides the ClpXP proteolytic machine, σ^{RpoS} proteolysis requires the response regulator RssB that, when phosphorylated, functions as an anti-σ factor adaptor that interacts with σ^{RpoS} preventing its binding to the RNAP and delivering it to ClpXP for degradation (Stüdemann et al., 2003; **Figure 3A**). RssB is an orphan response regulator but it is phosphorylated by the ArcB histidine kinase in response to the energy state of the cell. When ArcB senses a lack of energy, its sensor kinase is inactivated (Malpica et al., 2004). This halts RssB phosphorylation decreasing its affinity for σ^{RpoS} and thus the proteolysis of this σ factor (**Figure 3A**). In these conditions, σ^{RpoS} can interact with the RNAP and promote transcription of the general stress response genes (**Figure 3A**).

In *B. subtilis,* proteolysis of alternative σ factors plays a key role during sporulation, a tightly regulated process that is activated when the nutrient availability is limited in the environment. Two different cellular compartments are formed during sporulation, the forespore and the mother cell, and differential gene expression in these compartments is governed by the successive appearance of four group III σ factors, namely σ^{SigF} and σ^{SigG} in the forespore, and σ^{SigE} and σ^{SigK} in the mother

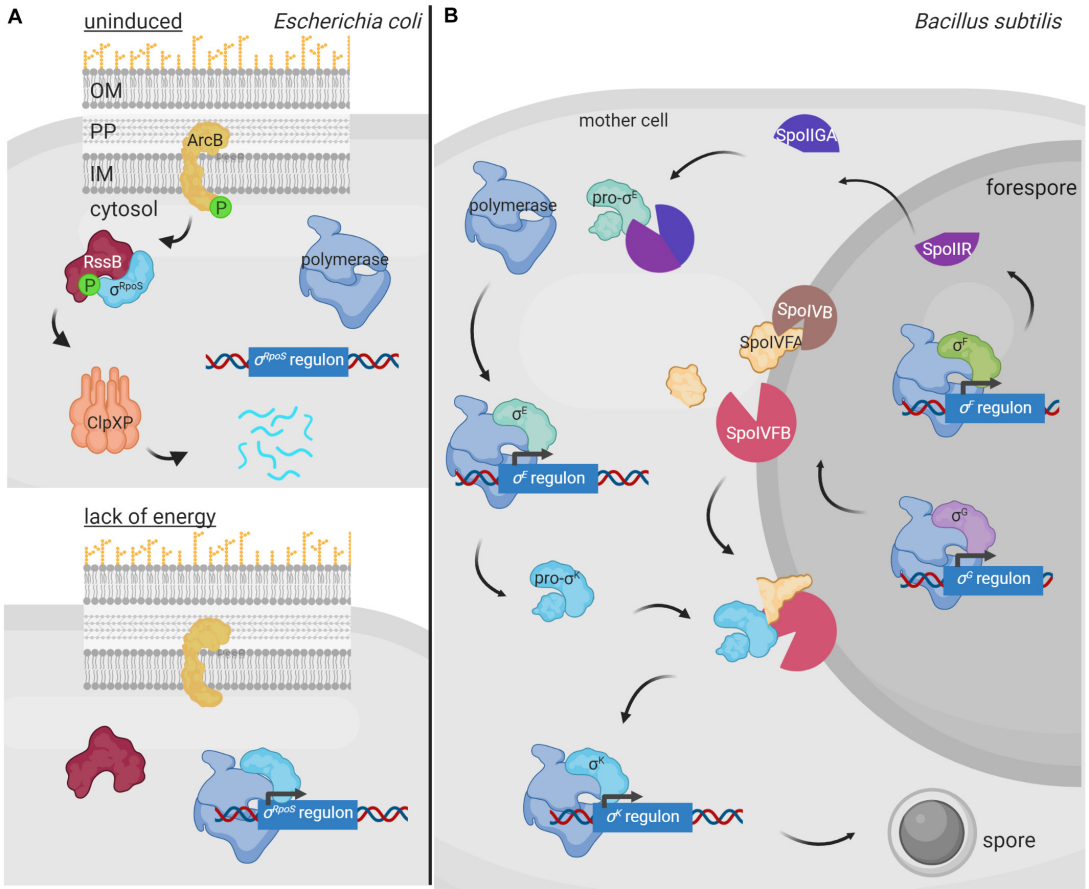


FIGURE 3 | Regulated proteolysis in the modulation of alternative σ factor activity. **(A)** During *E. coli* exponential growth, the histidine kinase ArcB (yellow) phosphorylates RssB (red), which in turn binds to the σ^{RpoS} factor (cyan) and delivers it to the ClpXP proteolytic complex (pink) for degradation. Lack of energy during the stationary phase of growth blocks the kinase activity of ArcB and thus RssB phosphorylation. Unphosphorylated RssB cannot bind σ^{RpoS}, and the σ factor is then able to bind the RNA polymerase (blue) and promotes transcription of the σ^{RpoS} regulon. **(B)** During spore formation in *B. subtilis*, a differential gene expression program is activated in the mother cell and the forespore by the action of the σ^{SigF}, σ^{SigE}, σ^{SigG} and σ^{SigK} factors. In the forespore, σ^{SigF} (σ^{F}) promotes the transcription of SpoIIR (bright purple), which translocates into the mother cell and forms an active protease complex with SpoIIGA (dark purple) that produces the cleavage of the σ^{SigE} precursor pro-σ^{SigE} (pro-σ^{E}) (green). Active σ^{SigE} initiates transcription of its regulon, including the σ^{SigK} precursor σ factor pro-σ^{SigK} (pro-σ^{K}) (blue). Simultaneously, σ^{SigG} (σ^{G}) promotes transcription of some of the proteases required to cleave pro-σ^{SigK}, which includes the SpoIVB protease (brown) that cleaves the SpoIVFA protease (yellow) allowing it to form a complex with SpoIVFB (pink) able to cleave and activate σ^{SigK}. Active σ^{SigK} triggers the final steps of spore formation. Figure produced with BioRender (BioRender.com).

cell (Fimlaid and Shen, 2015). σ^{SigE} and σ^{SigK} are produced as inactive precursor peptides pro-σ^{SigE} and pro-σ^{SigK} factors and are activated by regulated proteolysis at specific points during sporulation (**Figure 3B**). Sporulation begins with the activation of the transcription factor Spo0A by phosphorylation through a phosphorelay cascade involving several histidine kinases sensing nutrient starvation conditions (Burbulys et al., 1991; Jiang et al., 2000). Phosphorylated Spo0A promotes transcription of about 120 sporulation genes, including those encoding the pro-σ^{SigE} and σ^{SigF} factors (Molle et al., 2003). The latter is synthesized together with a sequestering anti-σ factor SpoIIAB and gets activated after polar septum formation (Duncan et al., 1995). Active σ^{SigF} triggers the transcription of a DNA translocase complex that pumps the chromosome into the forespore, and that of the *sigG* σ factor and *spoIIR* genes (Fujita and Losick, 2002). SpoIIR is secreted into the intermembrane space and activates the aspartic protease SpoIIGA, which subsequently cleaves pro-σ^{SigE} activating this σ factor (Imamura et al., 2008). In the mother cell, σ^{SigE} promotes transcription of the gene encoding the membrane-associated pro-σ^{SigK} factor and that of a hydrolase complex required for the assembly of a channel or "feeding tube" between the mother cell and the forespore that is also necessary for σ^{SigG} activity (Camp and Losick, 2009; Morlot et al., 2010). Activated σ^{SigG} promotes transcription of determinants needed in the proteolytic cascade that leads to the σ^{SigK} activation. In this way, σ^{SigK} activation in the mother cell is coupled to σ^{SigG} activity in the forespore. Cleavage and activation of pro-σ^{SigK} requires the S2P intermembrane metalloprotease SpoIVFB, which gets itself activated through the proteolysis of its inhibitory factor SpoIVFA (Campo and Rudner, 2006; **Figure 3B**). Active σ^{SigK} triggers a

gene expression program that results in complete sporulation formation (Ramirez-Guadiana et al., 2018).

Regulated Proteolysis of Anti-σ Factors

Regulated proteolysis of an anti-σ factor in response to a signal is often used to liberate and activate the alternative σ factor, especially in the case of σ^{ECF} associated anti-σ factors. Cytosolic anti-σ factors are often processed by ATP-dependent proteases (i.e., Clp, Lon) (**Table 1**). For instance, the *Streptomyces coelicolor* σ^{SigT} factor is sequestered and protected by its cytosolic anti-σ factor RstA and both undergo regulated proteolysis by cytosolic proteases at different time points of cell growth (Mao et al., 2013). Regulated proteolysis of RstA by yet unknown proteases during the onset of secondary metabolism liberates σ^{SigT}, which leads to a ClpP1/P2-dependent degradation of this σ factor (Mao et al., 2013). Because σ^{SigT} negatively regulates cell differentiation and *clpP1P2* expression, degradation of this σ factor results in production of secondary metabolites and the ClpP1/P2 proteolytic complex, which further accelerate σ^{SigT} degradation (Mao et al., 2013). The two-step proteolysis of σ^{SigT} by degrading first its cognate anti-σ factor RstA likely allows σ^{SigT} liberation when its function is required while subsequent σ^{SigT} degradation shuts off gene expression as soon as those conditions passed.

Membrane-bound anti-σ factors are usually processed by RIP in response to an inducing signal (Heinrich and Wiegert, 2009). Indeed, the periplasmic site-1 protease of a RIP pathway sometimes even functions as the sensor protein that triggers the activation of the σ^{ECF}/anti-σ factor signaling cascade. Generally, such protease contains one or multiple PDZ domains that occlude the catalytic site in absence of the signal and undergo conformational changes upon signal sensing exposing the catalytic site. The first and best characterized examples of such signaling pathway are the pathways activating the RpoE-like σ^{ECF} factors σ^{E} of *E. coli* and σ^{AlgU} of *P. aeruginosa* (reviewed in Ades, 2008; Chevalier et al., 2019; Otero-Asman et al., 2019b; **Figure 4A**). In absence of cell envelope stress, these σ factors are sequestered by their cognate membrane-bound anti-σ factors RseA and MucA. These inhibitions are enhanced by the periplasmic regulators RseB and MucB that bind to the periplasmic domains of RseA and MucA, respectively, protecting them from proteolysis. Activation of the σ^{E}/RseA and σ^{AlgU}/MucA signaling pathways occurs in response to cell envelope stress that leads to the accumulation of misfolded outer membrane proteins. Unfolded periplasmic peptides bind

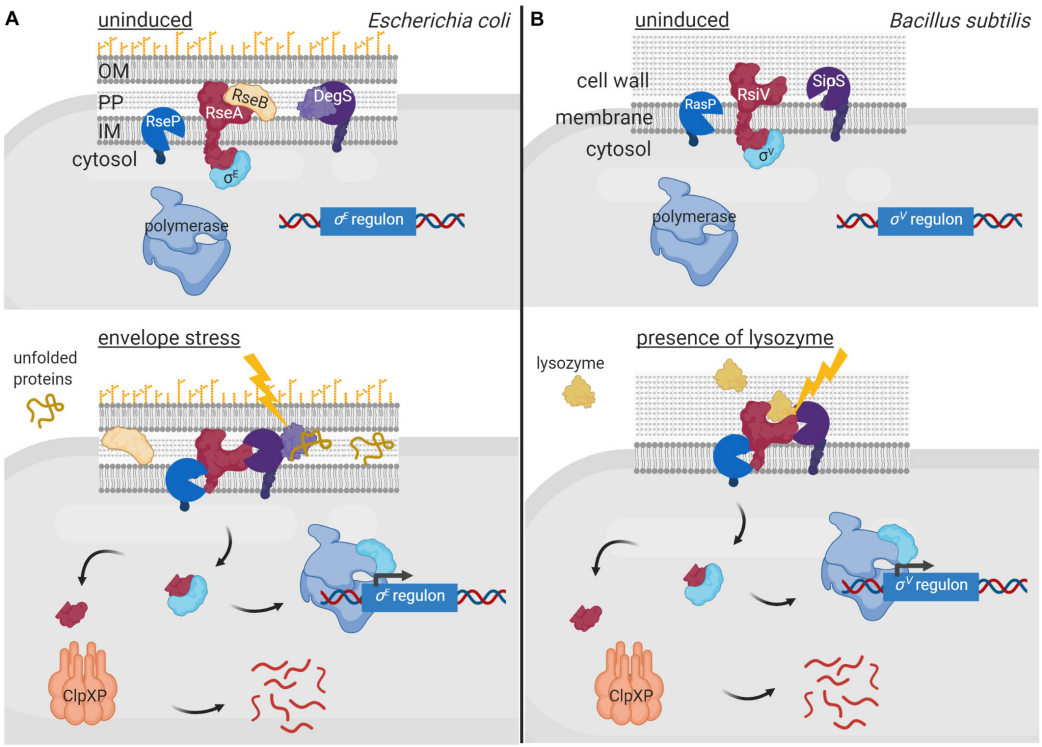


FIGURE 4 | Regulated proteolysis in the activation of σ^{ECF} factors. In uninduced conditions, the membrane bound anti-σ factors RseA of *E. coli* **(A)** and RsiV of *B. subtilis* **(B)** (red) sequester and keep inactive their cognate σ^{ECF} factors σ^{E} **(A)** and σ^{V} **(B)** (cyan). Sequestration of σ^{E} by RseA is enhanced by RseB (yellow), which binds to and impedes RseA proteolysis. Activation of the σ^{E} and σ^{V} factors occurs in response to envelope stress and presence of lysozyme, respectively, by RIP and cytosolic proteolysis of the anti-σ factors. The DegS protease **(A)** and the SipS peptidase **(B)** (purple) perform the site-1 cleavage of the RIP cascade while the RseP **(A)** and RasP **(B)** proteases (blue) the site-2 cleavage. The cytosolic proteolysis is carried out by the ClpXP complex (pink). The signal activating the proteolytic cascade of RseA is unfolded periplasmic peptides and is sensed by the DegS site-1 protease **(A)**, while that of RsiV is lysozyme and is sensed by the anti-σ factor itself **(B)**. Figure produced with BioRender (BioRender.com).

to the PDZ domain of the site-1 protease DegS (also known as AlgW in *P. aeruginosa*), being thus the signal that initiates the RIP pathway (**Figure 4A**). Signal binding reorientates the PDZ domain of DegS exposing its catalytic site (Walsh et al., 2003; Wilken et al., 2004). Simultaneously, LPS accumulation in the periplasm displaces the regulators RseB and MucB from RseA and MucA, respectively (Lima et al., 2013), allowing the binding of DegS to the anti-σ factors and their subsequent site-1 cleavage (**Figure 4A**). This cleavage generates the substrate for the site-2 protease of the RIP pathway, the metalloprotease RseP (also named MucP in *P. aeruginosa*) (**Figure 4A**). RseP cleaves within the transmembrane domain of the RseA/MucA anti-σ factors releasing their cytosolic N-domains bound to the σ^{E}/σ^{AlgU} factor into the cytosol (**Figure 4A**). The RseA/MucA N-domain is subsequently degraded by the ClpXP protease, which completely liberate σ^{E} and σ^{AlgU} allowing their interaction with the RNAP and the transcription of the σ^{E}/σ^{AlgU} regulon genes (Flynn et al., 2004; Qiu et al., 2008; **Figure 4A**). Importantly, the site-1 protease DegS does not only act as the sensor protein in these signaling pathways, but its activity is also the rate-limiting step in the regulatory cascade controlling σ^{E} and σ^{AlgU} activity (Chaba et al., 2007). Active σ^{E} and σ^{AlgU} factors promote transcription of genes to adapt to and protect the bacteria against periplasmic stress (Rhodius et al., 2006; Wood and Ohman, 2012). Besides, σ^{AlgU} is recognized for promoting the production of the exopolysaccharide alginate, which is responsible for the clinically relevant mucoid phenotype of *P. aeruginosa* that contributes to the persistence of this human pathogen in chronic infections (Ramsey and Wozniak, 2005).

In other membrane-associated σ^{ECF}/anti-σ signaling systems, the anti-σ factor is the sensor protein and undergoes a conformational change upon signal recognition that allows the site-1 cleavage. The best characterized example of such signal transduction network is the σ^{SigV}/RsiV system of *Bacillus subtilis* (reviewed in Ho and Ellermeier, 2019; **Figure 4B**). The membrane-bound protein RsiV does not only function as an anti-σ factor but also as a receptor for lysozyme (Hastie et al., 2016), which is an important component of the innate immune system of many organisms (Ragland and Criss, 2017). Binding of lysozyme to RsiV triggers RsiV degradation through a RIP cascade that leads to the activation of σ^{SigV} and the transcription of genes that confer resistance to lysozyme (**Figure 4B**). Importantly, the site-1 cleavage of RsiV is not carried out by a specific protease but by signal peptidases, especially SipS and SipT (Castro et al., 2018), which are the major signal peptidases of *B. subtilis*. The activity of signal peptidases is not regulated but constitutive, and RsiV cleavage in absence of lysozyme is prevented by two amphipathic helices that occlude the cleavage site (Lewerke et al., 2018). It has been hypothesized that binding of lysozyme to RsiV pulls the amphipathic helix into a β−sheet conformation, which exposes the cleavage site for the signal peptidases (Ho and Ellermeier, 2019). The site-2 cleavage of RsiV is performed by RasP, a membrane-embedded metalloprotease homologous to *E. coli* RseP (**Figures 4A,B**). RasP cleaves within the transmembrane domain of RsiV releasing the RsiV N-terminal cytosolic domain bound to the σ^{SigV} factor into the cytosol. Similarly to RseA and MucA (**Figure 4A**), it is assumed that the N-domain of RsiV is degraded by cytosolic proteases although they have not been identified yet (Ho and Ellermeier, 2019).

An additional layer of complexity in signal sensing and proteolytic activation of σ^{ECF}/anti-σ factor systems occurs in Gram-negative bacteria, as *Pseudomonas*, in which several σ^{ECF}/anti-σ proteins are functionally associated with an outer membrane receptor. These three proteins form a signal transduction system known as cell-surface signaling (CSS) that is activated by regulated proteolysis in response to a signal (reviewed in Llamas et al., 2014; Otero-Asman et al., 2019b; **Figure 5**). In these pathways, the outer membrane receptor, which belongs to the TonB-dependent transporter (TBDT) family, is the sensor protein and thus these systems are able to sense and respond to extracellular signals. CSS receptors function both as transporters of the inducing signal (often an iron-chelating compound, e.g., a siderophore) and as signal transfer proteins. This last function resides in the periplasmic N-terminal signaling domain (SD) of the CSS receptor that interacts with the periplasmic domain of its cognate membrane-bound anti-σ factor. In the current CSS model, signal binding at the outside of the receptor triggers a conformational change that modifies the SD/anti-σ factor interaction. This likely exposes the cleavage site of the anti-σ factor allowing its site-1 cleavage by a yet unknown periplasmic protease (**Figure 5**). The C-terminal processing periplasmic protease Prc seems to mediate the site-1 cleavage of a unique CSS protein of *P. putida*, named IutY, in which the σ^{ECF} and anti-σ factor functions are fused in a single protein (Bastiaansen et al., 2014; Bastiaansen et al., 2017). However, although required for CSS activation (Llamas et al., 2014; Otero-Asman et al., 2019a), whether Prc is directly or indirectly involved in cleaving CSS anti-σ factor proteins that are not fused to σ^{ECF} proteins is still unknown. Nevertheless, it is clear that a site-1 cleavage is necessary to generate the substrate for the site-2 protease, which for CSS anti-σ factors is the metalloprotease RseP (Draper et al., 2011; Bastiaansen et al., 2014; **Figure 5**). RseP cleavage leads to the release of the N-terminal cytosolic domain of the CSS anti-σ factor bound to its cognate σ^{ECF} factor into the cytosol (**Figure 5**). In some CSS pathways, this anti-σ factor domain is proteolytically removed by cytoplasmic proteases (e.g., ClpP). The best studied example of a CSS anti-σ factor subjected to cytosolic degradation is FpvR of *P. aeruginosa*. In response to the binding of the siderophore pyoverdine to the CSS receptor FpvA, FpvR undergoes regulated proteolysis liberating the σ^{FpvI} and σ^{PvdS} factors into the cytosol (Beare et al., 2003). RIP of FpvR is carried out by an unknown site-1 protease and by the site-2 protease RseP, and its cytosolic proteolysis by the ClpP protease (Draper et al., 2011; Bishop et al., 2017). However, in several pathways the N-terminal cytosolic domain of the anti-σ factor is not degraded and is required for CSS σ^{ECF} activity, having thus pro-σ activity (**Figure 5**). The best studied examples of CSS anti-σ factors with pro-σ activity include FoxR, FiuR and HxuR, which undergo RIP in response to the binding of the iron-chelating compounds ferrioxamine, ferrichrome and heme, respectively, to the CSS receptors FoxA, FiuA and HxuA, respectively (Llamas et al., 2006; Mettrick and Lamont, 2009; Bastiaansen et al., 2015; Otero-Asman et al., 2019a). Cleavage

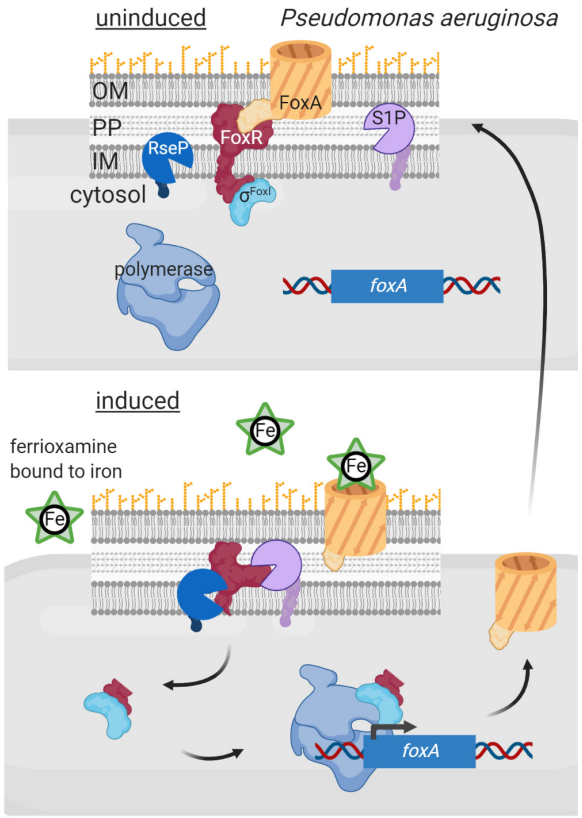


FIGURE 5 | Regulated intramembrane proteolysis of CSS anti-σ factors. In uninduced conditions, the CSS anti-σ factor FoxR (red) keeps σ^{FoxI} (cyan) sequestered and inactive. In this situation RIP of FoxR is likely prevented through the interaction with the CSS outer membrane receptor FoxA (orange). Binding of the siderophore ferrioxamine to FoxA likely modifies the FoxR/FoxA interactions surface, which exposes the anti-σ factor to the action of the proteases of the RIP cascade. FoxR is cleaved by a still unknown site-1 protease (purple) and by the site-2 protease RseP (blue), so that the N-terminal domain of FoxR is released into the cytosol bound to σ^{FoxI}. This complex interacts with the RNA polymerase (blue) and triggers transcription of the σ^{FoxI} regulon, including the *foxA* gene to increase the amount of the receptor in the outer membrane. Figure produced with BioRender (BioRender.com).

of these anti-σ factors by an unknown site-1 and the site-2 protease RseP releases their N-terminal cytosolic domains bound to their cognate σ^{ECF} factors, σ^{FoxI}, σ^{FiuI} and σ^{HxuI}, respectively (**Figure 5**). This anti-σ factor domain is thought to be bound to the σ^{ECF}-RNAP holoenzyme during the transcription process (Llamas et al., 2014; Otero-Asman et al., 2019b; **Figure 5**).

CONCLUDING REMARKS AND PROSPECTS

To communicate with the surrounding environment, bacteria contain tightly regulated and complex signaling pathways that allow them to perceive changes and perform adequate responses. While modulation of the activity of signal transduction systems by post-translational modifications like phosphorylation or methylation has been long known, the involvement of proteolysis in this process has become evident more recently. In fact, the idea of proteolysis as a regulatory mechanism has long been rejected because of the waste associated with destroying a protein synthesized at an energy cost. However, it has become evident that energy costs are less of a concern for a cell in the context of the regulation of its metabolic and signaling pathways. A major advantage of proteolysis as a regulatory post-transcriptional modification is that proteolysis is irreversible, which allows for an immediate response. It is therefore not surprising that the activity of many signal transduction proteins, which need to respond quickly to a stimulus (e.g., sudden envelope stress), is controlled by proteolysis. Moreover, the fact that it is an irreversible process allows for a longer duration of a response, which may be beneficial in developmental processes (e.g., sporulation). As outlined here, the signaling protein subject to regulated proteolysis can be the signal response protein directly or a signal transfer protein. Also, dependent on the signal transduction system, the activity of the regulatory protease can have stimulating or inhibiting effects on downstream gene transcription. Proteolysis of signal response proteins occurs either in the absence of the inducing stimulus to prevent the activity of a transcriptional activator, as exemplified by σ^{RpoS} (**Figure 3A**), or in response to the signal to inhibit the activity of a transcriptional repressor e.g., BlaI (**Figure 1B**) or to activate an inactive precursor protein as occurs with pro-σ^{SigE} and pro-σ^{SigK} (**Figure 3B**). Proteolysis of signal transfer proteins or that of interaction partners of these proteins usually occurs in response to the signal leading to the activation of a transcriptional activator. This is exemplified by the RIP of anti-σ factors that produces the activation of σ^{ECF} factors (**Figures 4**, **5**) or the proteolysis of histidine kinase interaction partners leading to the activation of the cognate response regulator (**Figure 2**).

An important aspect of the regulation of signaling systems by proteolysis is the involvement of regulatory proteases in signal sensing, as described for BlaR1 (**Figure 1B**) and DegS (**Figure 4A**). BlaR1 is actually a fusion protein with a penicillin-binding and a zinc metalloprotease domain. The BlaR1 sensor domain resembles class-D β-lactamases and contains the three catalytic motifs found in all penicillin binding proteins (Alexander et al., 2020). A key difference between the BlaR1/MecR1 sensor domains and β-lactamases is their deacetylation rates, which is considerably slower in the proteases and likely necessary for proper signal transmission (Alexander et al., 2020). In contrast, the sensor domain of DegS serine protease is the PDZ domain located in the periplasmic domain of the protease. PDZ is a modular domain of about 80–100 amino acids found in proteins of all organisms that often recognize short amino acid motifs at the C-termini of target proteins (Lee and Zheng, 2010). C-terminal peptides of unassembled outer membrane proteins are known to bind to the PDZ domain of DegS, initiating a steric clash between the PDZ domain and the L3 loop of DegS that breaks the L3-mediated autoinhibition of the proteolytic active site (De Regt et al., 2015). This allows the allosteric activation of DegS by the binding of peptides. Other proteases involved in signaling also contain PDZ domains, e.g., C-terminal processing proteases (Chueh et al., 2019).

It would be interesting to determine whether this domain is involved in signal sensing in these proteases and their mechanism of action.

In all, regulated proteolysis is a powerful tool to allow a bacterium to respond to environmental signals facilitating these simple life forms to thrive in highly diverse environments. Importantly, inhibition of signal transduction mechanisms is an interesting strategy for drug development that would prevent pathogens to detect and respond to the host environment. Because proteases are known to be druggable proteins, regulatory proteases involved in modulation of the activity of signal transduction systems required for pathogen's survival represent excellent drug targets. In fact, an inhibitor of the site-2 RseP protease has already shown to considerably decrease *E. coli* survival (Konovalova et al., 2018). The identification of new regulatory proteases involved in critical bacterial processes thus holds promise for the development of novel antibacterials.

AUTHOR CONTRIBUTIONS

SW and ML wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was funded by the FEDER and the Spanish Ministry of Economy with grant BIO2017-83763-P.

ACKNOWLEDGMENTS

We acknowledge support of the publication fee by the CSIC Open Access Publication Support Initiative through its Unit of Information Resources for Research (URICI).

REFERENCES

Ades, S. E. (2008). Regulation by destruction: design of the σE envelope stress response. *Curr. Opin. Microbiol.* 11, 535–540. doi: 10.1016/j.mib.2008.10.004

Alexander, J. A. N., Radaeva, M., King, D. T., Chambers, H. F., Cherkasov, A., Chatterjee, S. S., et al. (2020). Structural analysis of avibactam-mediated activation of the bla and mec divergons in methicillin-resistant *Staphylococcus aureus*. *J. Biol. Chem.* 295, 10870–10884. doi: 10.1074/jbc.ra120.013029

Almagro-Moreno, S., Kim, T. K., Skorupski, K., and Taylor, R. K. (2015a). Proteolysis of virulence regulator ToxR is associated with entry of *Vibrio cholerae* into a dormant state. *PLoS Genet.* 11:e1005145. doi: 10.1371/journal.pgen.1005145

Almagro-Moreno, S., Root, M. Z., and Taylor, R. K. (2015b). Role of ToxS in the proteolytic cascade of virulence regulator ToxR in *Vibrio cholerae*. *Mol. Microbiol.* 98, 963–976. doi: 10.1111/mmi.13170

Bastiaansen, K. C., Civantos, C., Bitter, W., and Llamas, M. A. (2017). New insights into the regulation of cell-surface signaling activity acquired from a mutagenesis screen of the *Pseudomonas putida* IutY sigma/anti-sigma factor. *Front. Microbiol.* 8:747. doi: 10.3389/fmicb.2017.00747

Bastiaansen, K. C., Ibanez, A., Ramos, J. L., Bitter, W., and Llamas, M. A. (2014). The Prc and RseP proteases control bacterial cell-surface signalling activity. *Environ. Microbiol.* 16, 2433–2443. doi: 10.1111/1462-2920.12371

Bastiaansen, K. C., Otero-Asman, J. R., Luirink, J., Bitter, W., and Llamas, M. A. (2015). Processing of cell-surface signalling anti-sigma factors prior to signal recognition is a conserved autoproteolytic mechanism that produces two functional domains. *Environ. Microbiol.* 17, 3263–3277. doi: 10.1111/1462-2920.12776

Beare, P. A., For, R. J., Martin, L. W., and Lamont, I. L. (2003). Siderophore-mediated cell signalling in *Pseudomonas aeruginosa*: divergent pathways regulate virulence factor production and siderophore receptor synthesis. *Mol. Microbiol.* 47, 195–207. doi: 10.1046/j.1365-2958.2003.03288.x

Belluzo, B. S., Abriata, L. A., Giannini, E., Mihovilcevic, D., Dal Peraro, M., and Llarrull, L. I. (2019). An experiment-informed signal transduction model for the role of the *Staphylococcus aureus* MecR1 protein in beta-lactam resistance. *Sci. Rep.* 9:19558.

Bishop, T. F., Martin, L. W., and Lamont, I. L. (2017). Activation of a cell surface signaling pathway in *Pseudomonas aeruginosa* requires ClpP protease and new sigma factor synthesis. *Front. Microbiol.* 8:2442. doi: 10.3389/fmicb.2017.02442

Bonnet, M., Kurz, M., Mesa, S., Briand, C., Hennecke, H., and Grutter, M. G. (2013). The structure of *Bradyrhizobium japonicum* transcription factor FixK2 unveils sites of DNA binding and oxidation. *J. Biol. Chem.* 288, 14238–14246. doi: 10.1074/jbc.m113.465484

Brown, M. S., Ye, J., Rawson, R. B., and Goldstein, J. L. (2000). Regulated intramembrane proteolysis: a control mechanism conserved from bacteria to humans. *Cell* 100, 391–398.

Burbulys, D., Trach, K. A., and Hoch, J. A. (1991). Initiation of sporulation in *B. subtilis* is controlled by a multicomponent phosphorelay. *Cell* 64, 545–552. doi: 10.1016/0092-8674(91)90238-t

Camp, A. H., and Losick, R. (2009). A feeding tube model for activation of a cell-specific transcription factor during sporulation in Bacillus subtilis. *Genes Dev.* 23, 1014–1024.

Campo, N., and Rudner, D. Z. (2006). A branched pathway governing the activation of a developmental transcription factor by regulated intramembrane proteolysis. *Mol. Cell* 23, 25–35. doi: 10.1016/j.molcel.2006.05.019

Castro, A. N., Lewerke, L. T., Hastie, J. L., and Ellermeier, C. D. (2018). Signal peptidase is necessary and sufficient for site 1 cleavage of RsiV in Bacillus subtilis in response to lysozyme. *J. Bacteriol.* 200:e00663-17.

Chaba, R., Grigorova, I. L., Flynn, J. M., Baker, T. A., and Gross, C. A. (2007). Design principles of the proteolytic cascade governing the σE-mediated envelope stress response in *Escherichia coli*: keys to graded, buffered, and rapid signal transduction. *Genes Dev.* 21, 124–136. doi: 10.1101/gad.1496707

Chan, C. M., Hahn, E., and Zuber, P. (2014). Adaptor bypass mutations of Bacillus subtilis spx suggest a mechanism for YjbH-enhanced proteolysis of the regulator Spx by ClpXP. *Mol. Microbiol.* 93, 426–438. doi: 10.1111/mmi.12671

Chevalier, S., Bouffartigues, E., Bazire, A., Tahrioui, A., Duchesne, R., Tortuel, D., et al. (2019). Extracytoplasmic function sigma factors in *Pseudomonas aeruginosa*. *Biochim. Biophys. Acta Gene Regul. Mech.* 1862, 706–721. doi: 10.1016/j.bbagrm.2018.04.008

Chueh, C. K., Som, N., Ke, L. C., Ho, M. R., Reddy, M., and Chang, C. I. (2019). Structural basis for the differential regulatory roles of the PDZ domain in C-terminal processing proteases. *mBio* 10:e01129-19.

De Regt, A. K., Baker, T. A., and Sauer, R. T. (2015). Steric clashes with bound OMP peptides activate the DegS stress-response protease. *Proc. Natl. Acad. Sci. U.S.A.* 112, 3326–3331. doi: 10.1073/pnas.1502372112

Deng, C. Y., Zhang, H., Wu, Y., Ding, L. L., Pan, Y., Sun, S. T., et al. (2018). Proteolysis of histidine kinase VgrS inhibits its autophosphorylation and promotes osmostress resistance in *Xanthomonas campestris*. *Nat. Commun.* 9:4791.

DiRita, V. J., Parsot, C., Jander, G., and Mekalanos, J. J. (1991). Regulatory cascade controls virulence in *Vibrio cholerae*. *Proc. Natl. Acad. Sci. U.S.A.* 88, 5403–5407. doi: 10.1073/pnas.88.12.5403

Dougan, D. A. (2013). *Regulated Proteolysis in Microorganisms*. Dordrecht: Springer.

Draper, R. C., Martin, L. W., Beare, P. A., and Lamont, I. L. (2011). Differential proteolysis of sigma regulators controls cell-surface signalling in *Pseudomonas aeruginosa*. *Mol. Microbiol.* 82, 1444–1453. doi: 10.1111/j.1365-2958.2011.07901.x

Duncan, L., Alper, S., Arigoni, F., Losick, R., and Stragier, P. (1995). Activation of cell-specific transcription by a serine phosphatase at the site of asymmetric division. *Science* 270, 641–644. doi: 10.1126/science.270.5236.641

Ehrmann, M., and Clausen, T. (2004). Proteolysis as a regulatory mechanism. *Annu. Rev. Genet.* 38, 709–724. doi: 10.1146/annurev.genet.38.072902.093416

Engman, J., and von Wachenfeldt, C. (2015). Regulated protein aggregation: a mechanism to control the activity of the ClpXP adaptor protein YjbH. *Mol. Microbiol.* 95, 51–63. doi: 10.1111/mmi.12842

Fimlaid, K. A., and Shen, A. (2015). Diverse mechanisms regulate sporulation sigma factor activity in the Firmicutes. *Curr. Opin. Microbiol.* 24, 88–95. doi: 10.1016/j.mib.2015.01.006

Flynn, J. M., Levchenko, I., Sauer, R. T., and Baker, T. A. (2004). Modulating substrate choice: the SspB adaptor delivers a regulator of the extracytoplasmic-stress response to the AAA+ protease ClpXP for degradation. *Genes Dev.* 18, 2292–2301. doi: 10.1101/gad.1240104

Fujita, M., and Losick, R. (2002). An investigation into the compartmentalization of the sporulation transcription factor σE in Bacillus subtilis. *Mol. Microbiol.* 43, 27–38. doi: 10.1046/j.1365-2958.2002.02732.x

Gur, E., Biran, D., and Ron, E. Z. (2011). Regulated proteolysis in Gram-negative bacteria—how and when? *Nat. Rev. Microbiol.* 9, 839–848. doi: 10.1038/nrmicro2669

Häse, C. C., and Mekalanos, J. J. (1998). TcpP protein is a positive regulator of virulence gene expression in *Vibrio cholerae*. *Proc. Natl. Acad. Sci. U.S.A.* 95, 730–734. doi: 10.1073/pnas.95.2.730

Hastie, J. L., Williams, K. B., Bohr, L. L., Houtman, J. C., Gakhar, L., and Ellermeier, C. D. (2016). The anti-sigma factor RsiV is a bacterial receptor for lysozyme: co-crystal structure determination and demonstration that binding of lysozyme to RsiV is required for σV activation. *PLoS Genet.* 12:e1006287. doi: 10.1371/journal.pgen.1006287

Heinrich, J., and Wiegert, T. (2009). Regulated intramembrane proteolysis in the control of extracytoplasmic function sigma factors. *Res. Microbiol.* 160, 696–703. doi: 10.1016/j.resmic.2009.08.019

Hengge, R. (2009). Proteolysis of σS (RpoS) and the general stress response in *Escherichia coli*. *Res. Microbiol.* 160, 667–676. doi: 10.1016/j.resmic.2009.08.014

Hizukuri, Y., Oda, T., Tabata, S., Tamura-Kawakami, K., Oi, R., Sato, M., et al. (2014). A structure-based model of substrate discrimination by a noncanonical PDZ tandem in the intramembrane-cleaving protease RseP. *Structure* 22, 326–336. doi: 10.1016/j.str.2013.12.003

Ho, T. D., and Ellermeier, C. D. (2019). Activation of the extracytoplasmic function σ factor σV by lysozyme. *Mol. Microbiol.* 112, 410–419. doi: 10.1111/mmi.14348

Imamura, D., Zhou, R., Feig, M., and Kroos, L. (2008). Evidence that the Bacillus subtilis SpoIIGA protein is a novel type of signal-transducing aspartic protease. *J. Biol. Chem.* 283, 15287–15299. doi: 10.1074/jbc.m708962200

Isaac, D. D., Pinkner, J. S., Hultgren, S. J., and Silhavy, T. J. (2005). The extracytoplasmic adaptor protein CpxP is degraded with substrate by DegP. *Proc. Natl. Acad. Sci. U.S.A.* 102, 17775–17779. doi: 10.1073/pnas.0508936102

Ishihama, A. (2000). Functional modulation of *Escherichia coli* RNA polymerase. *Annu. Rev. Microbiol.* 54, 499–518.

Jenal, U., and Hengge-Aronis, R. (2003). Regulation by proteolysis in bacterial cells. *Curr. Opin. Microbiol.* 6, 163–172. doi: 10.1016/s1369-5274(03)00029-8

Jiang, M., Shao, W., Perego, M., and Hoch, J. A. (2000). Multiple histidine kinases regulate entry into stationary phase and sporulation in *Bacillus subtilis*. *Mol. Microbiol.* 38, 535–542. doi: 10.1046/j.1365-2958.2000.02148.x

Keller, R., Arioz, C., Hansmeier, N., Stenberg-Bruzell, F., Burstedt, M., Vikstrom, D., et al. (2015). The *Escherichia coli* envelope stress sensor CpxA responds to changes in lipid bilayer properties. *Biochemistry* 54, 3670–3676. doi: 10.1021/acs.biochem.5b00242

Kinch, L. N., Ginalski, K., and Grishin, N. V. (2006). Site-2 protease regulated intramembrane proteolysis: sequence homologs suggest an ancient signaling cascade. *Protein Sci* 15, 84–93. doi: 10.1110/ps.051766506

Kirstein, J., Moliere, N., Dougan, D. A., and Turgay, K. (2009). Adapting the machine: adaptor proteins for Hsp100/Clp and AAA+ proteases. *Nat. Rev. Microbiol.* 7, 589–599. doi: 10.1038/nrmicro2185

Konovalova, A., Grabowicz, M., Balibar, C. J., Malinverni, J. C., Painter, R. E., Riley, D., et al. (2018). Inhibitor of intramembrane protease RseP blocks the σE response causing lethal accumulation of unfolded outer membrane proteins. *Proc. Natl. Acad. Sci. U.S.A.* 115, E6614–E6621.

Konovalova, A., Sogaard-Andersen, L., and Kroos, L. (2014). Regulated proteolysis in bacterial development. *FEMS Microbiol. Rev.* 38, 493–522. doi: 10.1111/1574-6976.12050

Kroos, L., and Akiyama, Y. (2013). Biochemical and structural insights into intramembrane metalloprotease mechanisms. *Biochim. Biophys. Acta* 1828, 2873–2885. doi: 10.1016/j.bbamem.2013.03.032

Lee, H. J., and Zheng, J. J. (2010). PDZ domains and their binding partners: structure, specificity, and modification. *Cell Commun. Signal.* 8:8.

Lee, Y. H., Kim, J. H., Bang, I. S., and Park, Y. K. (2008). The membrane-bound transcriptional regulator CadC is activated by proteolytic cleavage in response to acid stress. *J. Bacteriol.* 190, 5120–5126. doi: 10.1128/jb.00012-08

Lembke, M., Pennetzdorfer, N., Tutz, S., Koller, M., Vorkapic, D., Zhu, J., et al. (2018). Proteolysis of ToxR is controlled by cysteine-thiol redox state and bile salts in *Vibrio cholerae*. *Mol. Microbiol.* 110, 796–810. doi: 10.1111/mmi.14125

Lewerke, L. T., Kies, P. J., Muh, U., and Ellermeier, C. D. (2018). Bacterial sensing: a putative amphipathic helix in RsiV is the switch for activating σV in response to lysozyme. *PLoS Genet.* 14:e1007527. doi: 10.1371/journal.pgen.1007527

Lima, S., Guo, M. S., Chaba, R., Gross, C. A., and Sauer, R. T. (2013). Dual molecular signals mediate the bacterial response to outer-membrane stress. *Science* 340, 837–841. doi: 10.1126/science.1235358

Llamas, M. A., Imperi, F., Visca, P., and Lamont, I. L. (2014). Cell-surface signaling in *Pseudomonas*: stress responses, iron transport, and pathogenicity. *FEMS Microbiol. Rev.* 38, 569–597. doi: 10.1111/1574-6976.12078

Llamas, M. A., Sparrius, M., Kloet, R., Jimenez, C. R., Vandenbroucke-Grauls, C., and Bitter, W. (2006). The heterologous siderophores ferrioxamine B and ferrichrome activate signaling pathways in *Pseudomonas aeruginosa*. *J. Bacteriol.* 188, 1882–1891. doi: 10.1128/jb.188.5.1882-1891.2006

Llarrull, L. I., Toth, M., Champion, M. M., and Mobashery, S. (2011). Activation of BlaR1 protein of methicillin-resistant *Staphylococcus aureus*, its proteolytic processing, and recovery from induction of resistance. *J. Biol. Chem.* 286, 38148–38158. doi: 10.1074/jbc.m111.288985

Lu, H. Y., Luo, L., Yang, M. H., and Cheng, H. P. (2012). Sinorhizobium meliloti ExoR is the target of periplasmic proteolysis. *J. Bacteriol.* 194, 4029–4040. doi: 10.1128/jb.00313-12

Ludanyi, M., Blanchard, L., Dulermo, R., Brandelet, G., Bellanger, L., Pignol, D., et al. (2014). Radiation response in *Deinococcus deserti*: IrrE is a metalloprotease that cleaves repressor protein DdrO. *Mol. Microbiol.* 94, 434–449. doi: 10.1111/mmi.12774

Makinoshima, H., and Glickman, M. S. (2006). Site-2 proteases in prokaryotes: regulated intramembrane proteolysis expands to microbial pathogenesis. *Microbes Infect.* 8, 1882–1888. doi: 10.1016/j.micinf.2006.02.021

Malpica, R., Franco, B., Rodriguez, C., Kwon, O., and Georgellis, D. (2004). Identification of a quinone-sensitive redox switch in the ArcB sensor kinase. *Proc. Natl. Acad. Sci. U.S.A.* 101, 13318–13323. doi: 10.1073/pnas.0403064101

Mao, X. M., Sun, N., Wang, F., Luo, S., Zhou, Z., Feng, W. H., et al. (2013). Dual positive feedback regulation of protein degradation of an extra-cytoplasmic function sigma factor for cell differentiation in Streptomyces coelicolor. *J. Biol. Chem.* 288, 31217–31228. doi: 10.1074/jbc.m113.491498

Mascher, T. (2013). Signaling diversity and evolution of extracytoplasmic function (ECF) σ factors. *Curr. Opin. Microbiol.* 16, 148–155. doi: 10.1016/j.mib.2013.02.001

Mascher, T., Helmann, J. D., and Unden, G. (2006). Stimulus perception in bacterial signal-transducing histidine kinases. *Microbiol. Mol. Biol. Rev.* 70, 910–938. doi: 10.1128/mmbr.00020-06

Matson, J. S., and DiRita, V. J. (2005). Degradation of the membrane-localized virulence activator TcpP by the YaeL protease in *Vibrio cholerae*. *Proc. Natl. Acad. Sci. U.S.A.* 102, 16403–16408. doi: 10.1073/pnas.0505818102

Mettrick, K. A., and Lamont, I. L. (2009). Different roles for anti-sigma factors in siderophore signalling pathways of *Pseudomonas aeruginosa*. *Mol. Microbiol.* 74, 1257–1271. doi: 10.1111/j.1365-2958.2009.06932.x

Molle, V., Fujita, M., Jensen, S. T., Eichenberger, P., Gonzalez-Pastor, J. E., Liu, J. S., et al. (2003). The Spo0A regulon of *Bacillus subtilis*. *Mol. Microbiol.* 50, 1683–1701. doi: 10.1046/j.1365-2958.2003.03818.x

Morlot, C., Uehara, T., Marquis, K. A., Bernhardt, T. G., and Rudner, D. Z. (2010). A highly coordinated cell wall degradation machine governs spore morphogenesis in Bacillus subtilis. *Genes Dev.* 24, 411–422. doi: 10.1101/gad.1878110

Otero-Asman, J. R., Garcia-Garcia, A. I., Civantos, C., Quesada, J. M., and Llamas, M. A. (2019a). *Pseudomonas aeruginosa* possesses three distinct systems for sensing and using the host molecule haem. *Environ. Microbiol.* 21, 4629–4647. doi: 10.1111/1462-2920.14773

Otero-Asman, J. R., Wettstadt, S., Bernal, P., and Llamas, M. A. (2019b). Diversity of extracytoplasmic function sigma (σECF) factor-dependent signaling in *Pseudomonas*. *Mol. Microbiol.* 112, 356–373. doi: 10.1111/mmi.14331

Paget, M. S. (2015). Bacterial sigma factors and anti-sigma factors: structure, function and distribution. *Biomolecules* 5, 1245–1265. doi: 10.3390/biom5031245

Qiu, D., Eisinger, V. M., Head, N. E., Pier, G. B., and Yu, H. D. (2008). ClpXP proteases positively regulate alginate overexpression and mucoid conversion in *Pseudomonas aeruginosa*. *Microbiology* 154, 2119–2130. doi: 10.1099/mic.0.2008/017368-0

Ragland, S. A., and Criss, A. K. (2017). From bacterial killing to immune modulation: recent insights into the functions of lysozyme. *PLoS Pathog.* 13:e1006512. doi: 10.1371/journal.ppat.1006512

Raivio, T. L., and Silhavy, T. J. (1997). Transduction of envelope stress in *Escherichia coli* by the Cpx two-component system. *J. Bacteriol.* 179, 7724–7733. doi: 10.1128/jb.179.24.7724-7733.1997

Ramirez-Guadiana, F. H., Rodrigues, C. D. A., Marquis, K. A., Campo, N., Barajas-Ornelas, R. D. C., Brock, K., et al. (2018). Evidence that regulation of intramembrane proteolysis is mediated by substrate gating during sporulation in Bacillus subtilis. *PLoS Genet.* 14:e1007753. doi: 10.1371/journal.pgen.1007753

Ramsey, D. M., and Wozniak, D. J. (2005). Understanding the control of *Pseudomonas aeruginosa* alginate synthesis and the prospects for management of chronic infections in cystic fibrosis. *Mol. Microbiol.* 56, 309–322. doi: 10.1111/j.1365-2958.2005.04552.x

Rawlings, N. D., Waller, M., Barrett, A. J., and Bateman, A. (2014). MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res.* 42, D503–D509.

Rhodius, V. A., Suh, W. C., Nonaka, G., West, J., and Gross, C. A. (2006). Conserved and variable functions of the σE stress response in related genomes. *PLoS Biol.* 4:e2. doi: 10.1371/journal.pbio.0040002

Rojas-Tapias, D. F., and Helmann, J. D. (2018). Induction of the Spx regulon by cell wall stress reveals novel regulatory mechanisms in Bacillus subtilis. *Mol. Microbiol.* 107, 659–674. doi: 10.1111/mmi.13906

Sineva, E., Savkina, M., and Ades, S. E. (2017). Themes and variations in gene regulation by extracytoplasmic function (ECF) sigma factors. *Curr. Opin. Microbiol.* 36, 128–137. doi: 10.1016/j.mib.2017.05.004

Stock, A. M., Robinson, V. L., and Goudreau, P. N. (2000). Two-component signal transduction. *Annu. Rev. Biochem.* 69, 183–215.

Stüdemann, A., Noirclerc-Savoye, M., Klauck, E., Becker, G., Schneider, D., and Hengge, R. (2003). Sequential recognition of two distinct sites in σS by the proteolytic targeting factor RssB and ClpX. *EMBO J.* 22, 4111–4120. doi: 10.1093/emboj/cdg411

Teoh, W. P., Matson, J. S., and Dirita, V. J. (2015). Regulated intramembrane proteolysis of the virulence activator TcpP in *Vibrio cholerae* is initiated by the tail-specific protease (Tsp). *Mol. Microbiol.* 97, 822–831. doi: 10.1111/mmi.13069

Tschauner, K., Hornschemeyer, P., Muller, V. S., and Hunke, S. (2014). Dynamic interaction between the CpxA sensor kinase and the periplasmic accessory protein CpxP mediates signal recognition in E. *coli*. *PLoS One* 9:e107383. doi: 10.1371/journal.pone.0107383

Ulrich, L. E., Koonin, E. V., and Zhulin, I. B. (2005). One-component systems dominate signal transduction in prokaryotes. *Trends Microbiol.* 13, 52–56. doi: 10.1016/j.tim.2004.12.006

Urban, S., and Freeman, M. (2002). Intramembrane proteolysis controls diverse signalling pathways throughout evolution. *Curr. Opin. Genet. Dev.* 12, 512–518. doi: 10.1016/s0959-437x(02)00334-9

Vujicic-Zagar, A., Dulermo, R., Le Gorrec, M., Vannier, F., Servant, P., Sommer, S., et al. (2009). Crystal structure of the IrrE protein, a central regulator of DNA damage repair in Deinococcaceae. *J. Mol. Biol.* 386, 704–716. doi: 10.1016/j.jmb.2008.12.062

Walsh, N. P., Alba, B. M., Bose, B., Gross, C. A., and Sauer, R. T. (2003). OMP peptide signals initiate the envelope-stress response by activating DegS protease via relief of inhibition mediated by its PDZ domain. *Cell* 113, 61–71. doi: 10.1016/s0092-8674(03)00203-4

Wang, Y., Zhang, Y., and Ha, Y. (2006). Crystal structure of a rhomboid family intramembrane protease. *Nature* 444, 179–180. doi: 10.1038/nature05255

Wiech, E. M., Cheng, H. P., and Singh, S. M. (2015). Molecular modeling and computational analyses suggests that the Sinorhizobium meliloti periplasmic regulator protein ExoR adopts a superhelical fold and is controlled by a unique mechanism of proteolysis. *Protein Sci.* 24, 319–327. doi: 10.1002/pro.2616

Wilken, C., Kitzing, K., Kurzbauer, R., Ehrmann, M., and Clausen, T. (2004). Crystal structure of the DegS stress sensor: how a PDZ domain recognizes misfolded protein and activates a protease. *Cell* 117, 483–494.

Wood, L. F., and Ohman, D. E. (2012). Identification of genes in the σ22 regulon of *Pseudomonas aeruginosa* required for cell envelope homeostasis in either the planktonic or the sessile mode of growth. *mBio* 3:e00094-12.

Zhang, H. Z., Hackbarth, C. J., Chansky, K. M., and Chambers, H. F. (2001). A proteolytic transmembrane signaling pathway and resistance to beta-lactams in Staphylococci. *Science* 291, 1962–1965. doi: 10.1126/science.1055144

Zuber, P. (2004). Spx-RNA polymerase interaction and global transcriptional control during oxidative stress. *J. Bacteriol.* 186, 1911–1918. doi: 10.1128/jb.186.7.1911-1918.2004

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Wettstadt and Llamas. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 06 November 2020
doi: 10.3389/fmolb.2020.586413

Translation Initiation Control of RNase E-Mediated Decay of Polycistronic *gal* mRNA

Heung Jin Jeon[1†], Changjo Kang[1], Monford Paul Abishek N[1], Yonho Lee[1], Xun Wang[2], Dhruba K. Chattoraj[3] and Heon M. Lim[1]*

[1] Department of Biological Sciences, College of Biological Sciences and Biotechnology, Chungnam National University, Daejeon, South Korea, [2] State Key Laboratory of Agricultural Microbiology, College of Life Science and Technology, Huazhong Agricultural University, Wuhan, China, [3] Basic Research Laboratory, Center for Cancer Research, National Cancer Institute, Bethesda, MD, United States

In bacteria, mRNA decay is a major mechanism for regulating gene expression. In *Escherichia coli*, mRNA decay initiates with endonucleolytic cleavage by RNase E. Translating ribosomes impede RNase E cleavage, thus providing stability to mRNA. In transcripts containing multiple cistrons, the translation of each cistron initiates separately. The effect of internal translation initiations on the decay of polycistronic transcripts remains unknown, which we have investigated here using the four-cistron *galETKM* transcript. We find that RNase E cleaves a few nucleotides (14–36) upstream of the translation initiation site of each cistron, generating decay intermediates *galTKM*, *galKM*, and *galM* mRNA with fewer but full cistrons. Blocking translation initiation reduced stability, particularly of the mutated cistrons and when they were the 5′-most cistrons. This indicates that, together with translation failure, the location of the cistron is important for its elimination. The instability of the 5′-most cistron did not propagate to the downstream cistrons, possibly due to translation initiation there. Cistron elimination from the 5′ end was not always sequential, indicating that RNase E can also directly access a ribosome-free internal cistron. The finding in *gal* operon of mRNA decay by cistron elimination appears common in *E. coli* and *Salmonella*.

Keywords: mRNA decay, RNase E, mRNA stability, translation initiation, polycistronic gal mRNA

OPEN ACCESS

Edited by:
Chew Chieng Yeo, Sultan Zainal Abidin University, Malaysia

Reviewed by:
Miguel Angel Cevallos, National Autonomous University of Mexico, Mexico
Xiongfeng Dai, Central China Normal University, China

****Correspondence:***
Heon M. Lim
hmlim@cnu.ac.kr

†Present address:
Heung Jin Jeon, Infection Control Convergence Research Center, College of Medicine, Chungnam National University, Daejeon, South Korea

Specialty section:
This article was submitted to Molecular Recognition, a section of the journal Frontiers in Molecular Biosciences

Received: *23 July 2020*
Accepted: *28 August 2020*
Published: *06 November 2020*

Citation:
Jeon HJ, Kang C, N MPA, Lee Y, Wang X, Chattoraj DK and Lim HM (2020) Translation Initiation Control of RNase E-Mediated Decay of Polycistronic gal mRNA. Front. Mol. Biosci. 7:586413. doi: 10.3389/fmolb.2020.586413

INTRODUCTION

In bacteria, mRNA concentration is modulated to cope with a rapidly changing environment. The modulation is mediated by varying the synthesis as well as the decay rates of mRNA. In *Escherichia coli*, mRNA decay is primarily initiated by the endoribonuclease RNase E (Ono and Kuwano, 1979; Mudd et al., 1990; Babitzke and Kushner, 1991; Melefors and von Gabain, 1991; Taraseviciene et al., 1991; Carpousis et al., 2009). RNase E can target mRNAs for the initial cleavage either by first sensing a 5′ monophosphate group (5′ end-dependent pathway) or by bypassing this requirement (direct access pathway) (Mackie, 2013; Hui et al., 2014).

The 5′ end-dependent pathway begins with RNase E binding to the mono-phosphorylated 5′ end of transcripts (Celesnik et al., 2007) and cleavage at an internal site rich in AU sequences (McDowall et al., 1994). The N-terminal sensor domain of RNase E allows it to bind preferentially to the mono-phosphorylated rather than to the tri- or di-phosphorylated 5′ end of mRNA (Mackie, 1998; Jiang and Belasco, 2004; Callaghan et al., 2005). Thus, the 5′ end-dependent access requires prior conversion of transcripts with 5′ tri- or di-phosphorylated end

to mono-phosphorylated end by RNA pyrophosphohydrolase (RppH) (Celesnik et al., 2007; Deana et al., 2008). RNase E can also bypass the 5′ end binding and access a cleavage site internal to the message directly (Joyce and Dreyfus, 1998; Baker and Mackie, 2003). In either case, RNase E cleavage results in the generation of the 5′ portion of mRNA with a new 3′-OH end and the 3′ portion of mRNA with a new 5′ end. Generally, the 5′ portion mRNA is degraded immediately by 3′→ 5′ exonuclease digestion (Mott et al., 1985; Belasco et al., 1986; Hui et al., 2014; Wang et al., 2019). The 3′ portion mRNA could be subjected to decay by further rounds of cleavage by RNase E by either of the pathways (Joyce and Dreyfus, 1998; Spickler et al., 2001; Hui et al., 2014).

There are several mechanisms that modulate RNase E activity. Secondary structures at the 5′ end can extend the stability of mRNA because these structures interfere with RNase E initial cleavage (Bouvet and Belasco, 1992; Emory et al., 1992). Translation has a substantial effect on mRNA decay since translating ribosomes tend to inhibit RNase E cleavage (Braun et al., 1998; Deana and Belasco, 2005; Dreyfus, 2009). By the same token, translation initiation also promotes the stability of mRNA (Arnold et al., 1998; Joyce and Dreyfus, 1998; Baker and Mackie, 2003). In view of these modulating factors, it is unsurprising that polycistronic mRNAs show a differential decay of component cistrons (Belasco et al., 1985; Newbury et al., 1987; Adhya, 2003; Dar and Sorek, 2018).

Several models to account for the decay process of polycistronic mRNA in relation to translating ribosomes have been proposed (Alifano et al., 1994a). In polycistronic mRNA, the decay process is expected to be complex because, other than the secondary structures on RNA that interfere with RNase E and/or exoribonuclease activities, translation initiation at the internal start sites of each of the comprising cistrons could also be a compounding factor. Here we have addressed the decay process of a polycistronic mRNA in relation to translation initiation using the well-studied *gal* operon of *E. coli*. The operon harbors four genes–*galE*, *galT*, *galK*, and *galM* (about 1 kb each) – and produces the full-length mRNA *galETKM* of 4.3-kb size with the structure: 5′-g*alE–galT–galK–galM*-terminator hairpin-3′ (**Figure 1A**; Adhya, 2003). The terminator hairpin causes Rho-independent transcription termination and blocks 3′→ 5′ exoribonuclease digestion initiated from the free 3′-OH end of the transcript, providing stability to the *galETKM* mRNA (**Figure 1B**; Wang et al., 2019). Interestingly, the operon also produces two other mRNA species, *galETK* and *galTKM*, both about 3.3 kb. The *galETK* mRNA, 5′-g*alE–galT–galK*-3′, is generated by Rho-dependent transcription termination (Lee et al., 2008; Wang et al., 2014, 2015). How the *galTKM* mRNA, 5′-*galT–galK–galM*-3′, is generated is not clear.

In *E. coli*, mRNA decay is generally believed to proceed in 5′ to 3′ direction without any known 5′ to 3′ exoribonuclease (Apirion, 1973). There are also a number of well-documented examples of mRNAs whose degradation begins with cleavage far from the 5′ end (von Gabain et al., 1983; Belasco et al., 1985; Newbury et al., 1987; Braun et al., 1998). In addition to *galTKM* mRNA, the *gal* operon also produces *galKM* (formerly known as mK2; Lee et al., 2008). Most intermediate *gal* mRNA species are stable enough to be detected by northern blot and harbor full open reading frames (ORF) of the remaining cistrons. In other words, *gal* operon produces mRNA species that are cistron(s)-shorter from the 5′ end of the full-length mRNA. Here we have addressed whether the intermediates *galTKM* and *galKM* mRNA are generated by cistron elimination of *galETKM* from the 5′ end.

In this study, using northern blot and 5′RACE assays, we found that the decay of polycistronic *galETKM* mRNA by and large proceeds by eliminating nearly full-length cistron(s) from the 5′ end, generating *galTKM*, *galKM*, and *galM* mRNA species as decay intermediates, but the elimination is not always sequential. Translation initiation at the newly generated 5′ end of the intermediate mRNA contributes to the stability of the decay intermediates. A sequence analysis of the 5′ end of mRNA generated as a result of RNase E cleavage (Chao et al., 2017) shows that eliminating cistrons from the 5′ end is the common theme in polycistronic mRNA decay in *E. coli* as well as in *Salmonella*.

MATERIALS AND METHODS

Extraction of RNA From *E. coli* Cells

Total RNA was prepared from 2×10^8 *E. coli* cells grown to OD_{600} of 0.6 in Luria broth (LB) as described previously (Wang et al., 2014). The primers used in this study are listed in **Supplementary Table S1**.

5′RACE

To determine the 5′ end of the *gal* mRNA, we ligated the 3′ end of the 5S rRNA in the RNA preparation from *E. coli* cells to 5′ ends of RNA in the same RNA preparation (Lee et al., 2008). Later, the *gal* mRNA of interest was amplified using RT-PCR. First, RNA ligation reaction was performed in a 15-μl volume containing 2.5 μg of total RNA, 5 U T4 RNA ligase (Ambion, United States), and 10 U rRNasin (Promega, United States) at 37°C for 3 h. One microgram of the ligated RNA was reverse-transcribed at 37°C for 2 h in a 20-μl reaction volume containing 4 U reverse transcriptase (Qiagen, Germany), 0.5 mM each of dNTP, 10 μM random hexamer primer (Takara, Japan), and 10 U rRNasin. Two microliters of the reverse transcription reaction was used as template for PCR amplification of the *gal* cDNA of interest in the next step in a total volume of 20 μl using 1 U of HotStar *Taq* DNA polymerase (Qiagen, Germany), with a forward primer complementary to the 3′ end of the *E. coli* 5S rRNA and a reverse primer specific to a *gal* cDNA of interest (Lee et al., 2008) (listed in **Supplementary Table S1**). The 5′ end of the *gal* cDNA (thus, *gal* mRNA) of interest was assayed by extending a ^{32}P-labeled DNA primer bound to a specific region of the amplified *gal* cDNA of interest. This "primer extension" reaction was performed in a 20-μl volume containing 10 μl of amplified *gal* cDNA reaction, 0.75 μl ^{32}P-labeled primer, 0.15 mM each of dNTP, and 1 U *Taq* polymerase with 25 cycles of amplification. The exact location of the 5′ ends of the *gal* mRNA was identified as follows: The extended-primer DNAs were resolved on 8% polyacrylamide–urea sequencing gel, and radioactive bands were visualized on an X-ray film. We measured the number of nucleotides of the extended-primer DNAs using the DNA sequencing ladder as a

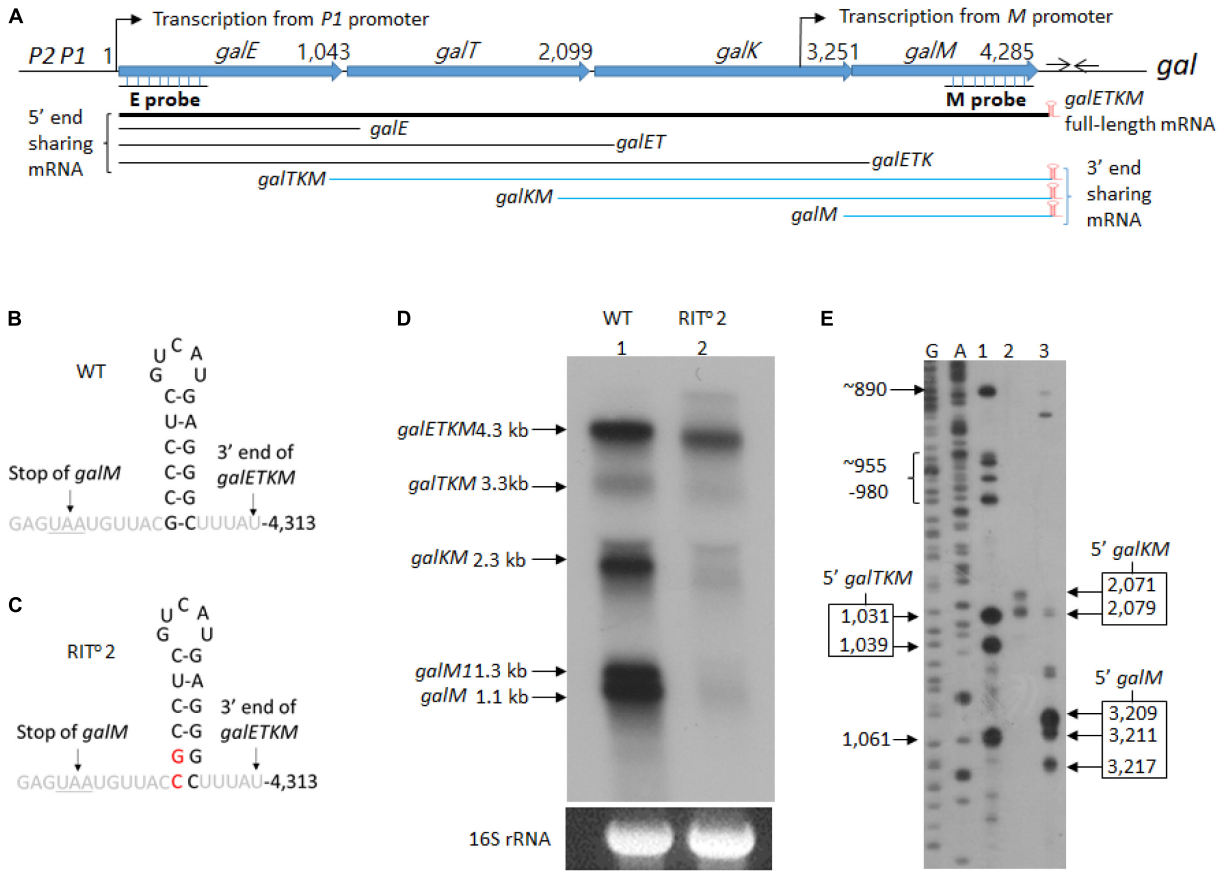


FIGURE 1 | mRNA intermediates of the *gal* operon that share the 3′ end with the full-length *gal* mRNA. **(A)** Map of *galETKM* operon comprising the genes *galE*, *galT*, *galK,* and *galM* (Lewis and Adhya, 2015). Numbers indicate *gal* nucleotide residue coordinate, where 1 is the start of transcription from the *galP1* promoter. The number at the end of each gene belongs to the last nucleotide of the stop codon of that gene. The *gal* operon transcription from the *M* promoter initiates at 3,073. At six nucleotides downstream of the stop codon of *galM* gene, there is a 17-nucleotide inverted repeat sequence (head-to-head arrows). It forms the terminator hairpin (red) that terminates transcription and protects *galETKM* mRNA from 3′→5′ exonuclease digestion (McLaren et al., 1991; Wang et al., 2019). Two probes, E and M (underlined), were used to identify intermediates that share the 5′ and the 3′ ends with the full-length *galETKM* mRNA. The 3′-end-sharing mRNA shares the same 3′ end at 4,313 and have 5′ ends at the beginning of each comprising gene. **(B)** Nucleotide sequence of the terminator hairpin. **(C)** Nucleotide sequence of the terminator hairpin in the RIT^{0}2 *gal* mutant. The nucleotide changes are indicated in red. **(D)** Northern blot of wild type (WT; lane 1) and RIT^{0}2 *gal* mutant (lane 2). *gal* mRNA prepared from cells grown up to OD_{600} of 0.6 in Luria broth with 0.5% galactose. The blot was probed with the M probe complementary to the 3′ half of the *galM* gene of the *gal* operon **(A)**. **(E)** 5′RACE of WT *gal* mRNA prepared under the same condition as in **(D)**. The lanes show the 5′ ends found near the beginning of either *galT* (lane 1) or *galK* (lane 2) or *galM* (lane 3) genes. DNA sequencing ladders used as molecular weight markers are in lanes marked G and A. Note that the 5′RACE procedure depends on a ligation step that requires a mono-phosphorylated 5′ end. Thus, 5′RACE detects specifically 5′ ends of mRNA that are mono-phosphorylated.

ruler. By subtracting the number of nucleotides in the primer DNA from the number of nucleotides of the extended-primer DNA, we located the exact position of the 5′ end of the *gal* mRNA in the nucleotide residue coordinates of the *gal* operon.

Northern Blot

Typically, 10 μg of total RNA (after staining with 1 μg/ml ethidium bromide) was resolved by electrophoresis using 1.2% (wt/vol) formaldehyde–agarose gel at 5 V/cm for 4 h. RNA on the gel was then transferred overnight to a positively charged nylon membrane (Ambion, United States) using a downward transfer system (TurboBlotter, Whatman, United Kingdom). The nylon membrane was baked at 80°C for 1 h. M3 probe was prepared by PCR using a pair of primers complementary to 3,751 and 4,285 in *gal* coordinates. Probe DNA was labeled with ^{32}P. Hybridization procedures followed the manufacturer's instructions (Ambion, United States). RNA bands were quantified using the software JMOL (NIH).

RESULTS

mRNA Intermediates of the *gal* Operon That Share the 3′ End With the Full-Length *gal* mRNA

We performed northern blot of total RNA prepared from wild-type *E. coli* cells (MG1655) grown in LB supplemented with 0.5% galactose to OD_{600} of 0.6. The blots were analyzed using the

M-probe, which hybridizes to the distal half of the last gene of the *gal* operon, *galM* (**Figure 1A**). We detected five mRNA species with sizes of 4.3, 3.3, 2.3, 1.3, and 1.1 kb (lane 1 in **Figure 1D**). We named these mRNA species *galETKM*, *galTKM*, *galKM, galM1*, and *galM*, respectively, as we show later that these mRNA species have their 5′ ends at the beginning of each cistron but have the same 3′ end, which is at the end of the operon (**Figure 1A**). We called these mRNA species "3′-end-sharing" mRNA.

The 1.3-kb mRNA named *galM1* turned out to be the result of transcription initiation at an internal promoter in front of the *galM* gene, which terminated at the end of *galM* (**Supplementary Figure S1**). We designated the internal promoter as *M* promoter. We excluded the *galM1* and *galETKM* mRNA species from the 3′-end-sharing group of decay intermediates because these are primary transcripts (see below).

To test if the 3′-end-sharing mRNA species harbor the terminator hairpin structure of the full-length *galETKM* 3′ end (**Figures 1A,B**; Wang et al., 2019), we performed northern blot on RNA prepared from the culture of MG1655Δ*gal* strain, where the chromosomal *gal* operon is entirely deleted. The strain harbored a pBAC-derived low-copy-number plasmid that carries a mutant version of the *gal* operon, RIT°2 (Wang et al., 2014). In the RIT°2 mutant, two bases from the bottom of the stem of the terminator hairpin are substituted to their complementary bases (**Figure 1C**). In the RIT°2 mutant, the function of the terminator hairpin in the blockage of 3′→ 5′ exoribonuclease was impaired (Wang et al., 2019). When RNA from the RIT°2 mutant was analyzed using the M-probe, we confirmed that the mutations caused the shortening of the *galETKM* mRNA by about 100 nucleotides and also decreased the *galETKM* amount to about 40% of the wild type (WT; lane 2 in **Figure 1D**; Wang et al., 2019). The northern blot also showed that *galTKM*, galKM, and *galM1* also got similarly shorter than their corresponding WT counterparts (lane 2, **Figure 1D**), and the amount of these mRNA species decreased to less than 10% of their amounts in the WT. The *galM* mRNA was hardly observed in the RIT°2 mutant (lane 2, **Figure 1D**). These results demonstrate that the 3′-end-sharing mRNA species and the primary transcript *galM1* harbor the terminator hairpin structure of the full-length *galETKM*.

To determine the 5′ end sequences of the 3′-end-sharing mRNA, we used 5′RACE. The assay showed that the 5′ ends of the *galTKM* mRNA were clustered around *gal* coordinates 1,031 and 1,039 (lane 1, **Figure 1E**), those of the *galKM* mRNA at 2,071 and 2,079 (lane 2, **Figure 1E**), and those of the *galM* mRNA at 3,209, 3,211, and 3,217 (lane 3, **Figure 1E**). The *gal* nucleotide coordinates, starting from the *P1*-transcription initiation site, are as in **Figure 1A**. These results indicate that the 5′ ends of the 3′-end-sharing mRNA are generated by endoribonucleolytic cleavages occurring near cistron junctions, mostly within the last few codons of the respective upstream ORF.

The 3′-End-Sharing mRNA Species Are the Intermediate Products of the RNase E-Mediated Decay of *galETKM*

RNase E is the major endoribonuclease in *E. coli*. To test whether this nuclease is responsible for cleavages near cistron junctions, generating the 3′-end-sharing mRNA species, we assayed for their presence in GW20 (*ams1*^{ts}), a strain temperature-sensitive for RNase E activity (Wachi et al., 1997). We subjected the total RNA prepared from GW20 to northern blot analysis. Cells were grown to OD_{600} of 0.2 at 30°C in LB supplemented with 0.5% galactose, and the culture was divided into two halves. One half was shifted to 44°C, a non-permissive temperature, whereas the other half continued to grow at 30°C until they reached OD_{600} of 0.6. The northern blot results showed that, at non-permissive temperature, when GW20 cells were deprived of functional RNase E, *galTKM*, *galKM*, and *galM* mRNA species decreased drastically (**Figure 2A**). These results support the view that the 3′-end-sharing mRNA species, *galTKM*, *galKM*, and *galM*, are the products of RNase E cleavage.

To confirm that the 5′ ends of the 3′-end-sharing mRNA (shown in **Figure 1E**) are generated by RNase E cleavage, we assayed the 5′ ends by 5′RACE on total RNA from GW20 cells grown at permissive and non-permissive temperatures. The results showed that, at the non-permissive temperature of 44°C, most 5′ ends of *galTKM* (lane 2, **Figure 2B**), *galKM* (lane 4, **Figure 2B**), and *galM* (lane 6, **Figure 2B**) were less than 10% of their amounts at the permissive temperature of 30°C (lanes 1, 3, and 5, **Figure 2B**). These results demonstrate that the 5′ ends of the 3′-end-sharing mRNA species are generated primarily by RNase E cleavage.

It is noteworthy that, in contrast to the decreasing amounts of 3′-end-sharing mRNA at the non-permissive temperature, the level of the primary transcripts, *galETKM* and *galM1*, increased by about twofold (lane 2, **Figure 2A**). The increased level of primary transcripts is to be expected if the 3′-end-sharing mRNA species are the RNase E-mediated decay products of *galETKM*.

The 5′ Portion mRNA Generated by RNase E Cleavage Is Degraded Rapidly

So far, we have analyzed only the 3′ portions of the *galETKM* mRNA after the RNase E cleavages at 1,031 and 1,039, 2,071 and 2,079, and 3,209, 3,211, and 3,217 that generated the 3′-end-sharing mRNA species *galTKM*, *galKM*, and *galM*, respectively. The 5′ portion of the *galETKM* mRNA after the RNase E cleavages should be in sizes of 1,031–1,039, 2,071–2,079, and 3,209–3,217 nucleotides long, respectively.

To detect these species, we performed northern blot of *gal* mRNA from wild-type and RIT°2 mutant cells, this time using the E-probe (**Figure 1A**) that hybridizes to the first half of the *galE* gene (**Figure 1A**). The northern blot of the wild-type cells showed *galETKM* (4.3 kb), *galETK* (3.3 kb), *galET* (2.2 kb), *galE1* (1.8 kb), and *galE* (1.2 kb) (**Figure 3**). These have been presented as "5′-end-sharing" mRNA in our previous reports (**Figure 1A**; Lee et al., 2008; Wang et al., 2014, 2015). In the RIT°2 mutant, where the terminator hairpin is no longer functioning (Wang et al., 2019), *galETKM* is the only mRNA species affected; *galETKM* reduced in size by ∼100 nucleotides and decreased in amount to about 25% of the WT, but all the other mRNA species were present as much as they were in wild-type cells (**Figure 3**). These results demonstrate that the *galETKM* mRNA is *not* a precursor of all other *gal* mRNA species shown in

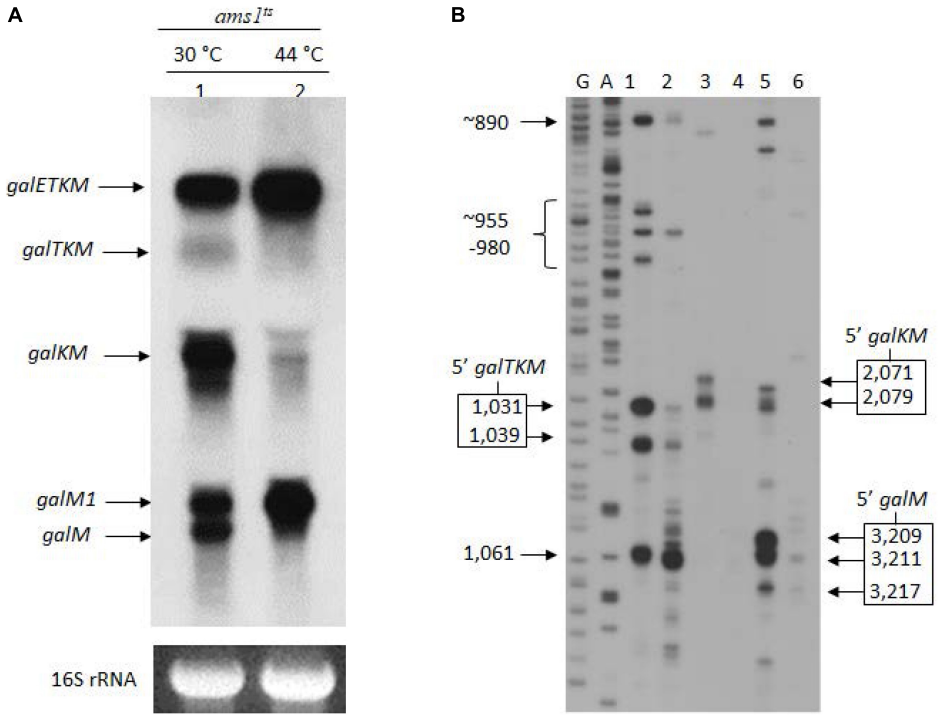


FIGURE 2 | The 3′-end-sharing mRNA species are the intermediate products of the RNase E-mediated decay of *galETKM*. **(A)** Northern blot of the *gal* mRNA in GW20 (*ams1*^{ts}) cells grown at either the permissive temperature of 30°C (lane 1) or the non-permissive temperature of 44°C (lane 2). The blot was probed with the M probe complementary to the 3′ half of the *galM* gene of the *gal* operon. *galETKM* and *galM1* are the primary transcripts, and *galTKM*, *galKM*, and *galM* are the decay intermediates. **(B)** 5′RACE of *gal* mRNA prepared from GW20 (*ams1*^{ts}) cells as in **(A)**. The lanes show the 5′ ends of RNA at the permissive and the non-permissive temperatures found near the beginning of either the *galT* gene (lanes 1 and 2) or the *galK* gene (lanes 3 and 4) or the *galM* gene (lanes 5 and 6). DNA sequencing ladders used as molecular weight markers are in lanes marked G and A. An exception was the 5′ end at 1,061, which was unchanged at the two temperatures, indicating that this 5′ end was not generated by RNase E.

Figure 3. These mRNA species are known to be generated as a result of premature transcription termination by Rho (Adhya, 2003; Wang et al., 2015). Indeed bicyclomycin (BCM), the Rho-inhibitor, inhibited the production of all the 5′-sharing mRNA species (**Supplementary Figure S2**).

These results thus suggest that the 5′ portion of the *galETKM* mRNA after the RNase E cleavages is removed immediately, possibly by 3′→5′ exoribonucleases, as has been demonstrated in many previous studies (Mott et al., 1985; Belasco et al., 1986; Hui et al., 2014; Wang et al., 2019).

Temporal Order of Cistron Elimination From the *galETKM* mRNA

To test whether the cistrons are eliminated one at a time from the 5′ end of the *galETKM* mRNA, we performed a "temperature shift-down" (TSD) experiment and followed the order of appearance of the decay intermediates with time. We used the GW20 (*ams1*^{ts}) strain, expecting that RNase E would regain its function and initiate the decay process as the temperature of the GW20 culture is brought down from 44 to 30°C. If the cistrons are eliminated one at a time from the 5′ end of *galETKM*, we expect that the order of appearance of the intermediates would be *galTKM*, *galKM*, and finally *galM*.

We transferred a culture of GW20 grown to OD_{600} of 0.6 at 44 to 30°C and continued its growth. We took aliquots of the culture after 5, 10, and 20 min of TSD, prepared RNA, and subjected the RNA to northern blot analysis using the M-probe. The levels of decay intermediates 20 min after TSD (lane 5, **Figure 4A**), when the temperature of the culture reached 31°C, appeared as much as their levels when the GW20 culture was grown continuously to OD_{600} of 0.6 at 30°C (lane 1, **Figure 4A**). This indicates that, 20 min after, TSD suffices to restore the full function of RNase E. The same inference could also be made from the appearance of the 16S rRNA band presented as a loading control in **Figure 4A**. We observed that an RNA band with a molecular weight higher than that of the 16S rRNA appears at the non-permissive temperature (thick arrow, lane 2, **Figure 4A**). The intensity of this RNA band decreases and that of the 16S rRNA band increases during the TSD period. By 20 min, the ratio of the two bands became comparable to the ratio seen at the permissive temperature. Considering that the RNase E cleavage of the precursor 17S RNA generates 16S rRNA (Li et al., 1999), these results corroborate the abovementioned finding that the function of RNase E is restored by 20 min of TSD.

The sequence of appearance of the decay intermediates during TSD period could be sequential from the 5′ end since, at the

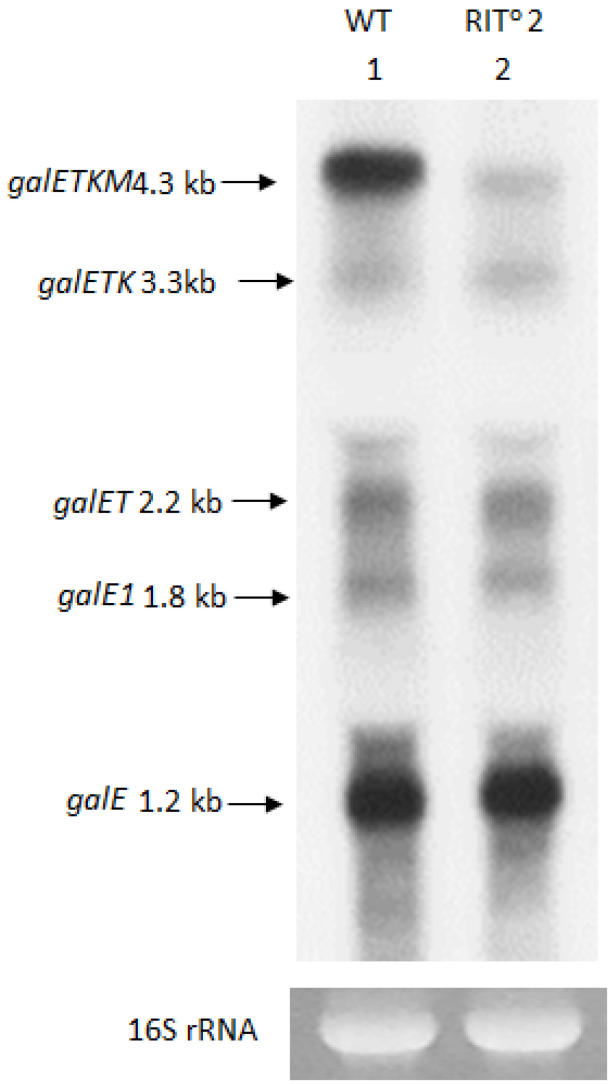


FIGURE 3 | The 5′ portion mRNA generated by RNase E cleavage is degraded rapidly. Northern blot of *gal* mRNA probed with the E probe. *gal* mRNA was prepared from wild-type cells (lane 1), and the RIT°2 *gal* mutant cells (lane 2) were grown up to OD_{600} of 0.6 in Luria broth with 0.5% galactose. The blot was probed with the E probe complementary to the 3′ half of the *galE* gene of the *gal* operon (**Figure 1A**).

earliest time point (5 min) after TSD, the amount of *galTKM* increased the most, followed by the amounts of *galKM* and *galM*, the increase in the latter being hardly detectable (**Figure 4A**). This suggests that *galTKM* and *galKM* are generated before *galM*. At 10 min after TSD, the amounts of *galTKM* and *galKM* increased 2.7- and 1.6-fold, respectively, relative to their amounts at time-zero (**Figure 4C**). A slight increase in *galM* amount (about 1.1-fold) was also detected. At 20 min after TSD, the *galM* amount increased 1.8-fold, while the amounts of the other two intermediates, *galTKM* and *galKM*, decreased (**Figure 4C**). These results further show that *galM* is formed subsequent to the formation of *galTKM* and *galKM*, and their decrease suggests that these longer intermediates could be serving as precursors of *galM*.

During the TSD period, the amounts of both primary transcripts *galETKM* and *galM1* decreased (**Figure 4D**). Considering that the function of RNase E restores gradually during the TSD period, the gradual decrease in the amounts of *galETKM* mRNA is consistent with the notion that RNase E cleavage of the *galETKM* gives rise to the decay intermediates. The decrease in *galM1* indicates that it is also a substrate of RNase E.

To test if there is a difference in the timing of appearance between *galTKM* and *galKM*, we analyzed the decay intermediates every minute during the first 5-min period after TSD. Temperature gradually fell from 44 to 37°C during the first 5 min (**Figure 4B**). *galTKM* and *galKM* appeared essentially simultaneously, while the *galM* mRNA amount hardly changed (**Figures 4B,E**). The simultaneous appearance of *galTKM* and *galKM* indicates that the latter is not produced from the former and that both are independently produced by the RNase E cleavage of *galETKM*. Thus, the appearance of the decay intermediates may not be strictly sequential.

RNase E Also Cleaves Upstream of the Terminator Hairpin of the *gal* Operon

All the decay intermediates and the primary transcripts *galETKM* and *galM1* have the same terminator hairpin stem-loop structure at the 3′ end (**Figure 1B**). In the RIT°2 mutant, where the function of the stem-loop is impaired, production of these mRNA species was greatly impaired (**Figure 1D**). The 3′ terminal stem-loop is known to enhance the half-life of mRNA by serving as a road-block to the 3′→5′ exoribonuclease digestion (McLaren et al., 1991; Regnier and Hajnsdorf, 1991; Carpousis et al., 2009). We argued that, to complete the decay of the *galETKM* mRNA, RNase E should cleave upstream of the terminator hairpin of *galM* mRNA to allow 3′→ 5′ exoribonuclease access for digestion of *galM* mRNA.

To test if there are any RNase E-generated 5′ ends upstream of the terminator hairpin, we performed 5′RACE around the terminator hairpin in GW20 ($ams1^{ts}$) cells. We found two clusters of 5′ ends at permissive temperature. In one cluster, the ends were at coordinates 4,273, 4,269, and 4,266, and in the other, the ends were at 4,241, 4,237, and 4227 (lane 1, **Figure 5A**). These 5′ end products became nearly undetectable at non-permissive temperature, except for the product ending at 4,227 (lane 2, **Figure 5A**). We infer that, barring the end at 4,227, the other five ends were generated by RNase E. The RNase E cleavage sites were at 20–40 nucleotides upstream of the hairpin, which should allow exoribonuclease processing unencumbered by the hairpin (**Figure 5B**). The origin of the 5′ end at 4,227 remains to be studied.

Effect of Translation Initiation on the Production of mRNA Intermediates

RNase E cleavage is known to have moderate sequence specificity with a degenerate consensus sequence 5′-R(A/G)NW(A/U)UU-3′ (R = A or G, N = any, W = A or U) (Belasco, 2017; Chao et al., 2017). In this sequence, RNase E cleaves between the N and the W residues, leaving the W residue at the 5′ end (Belasco, 2017; Chao

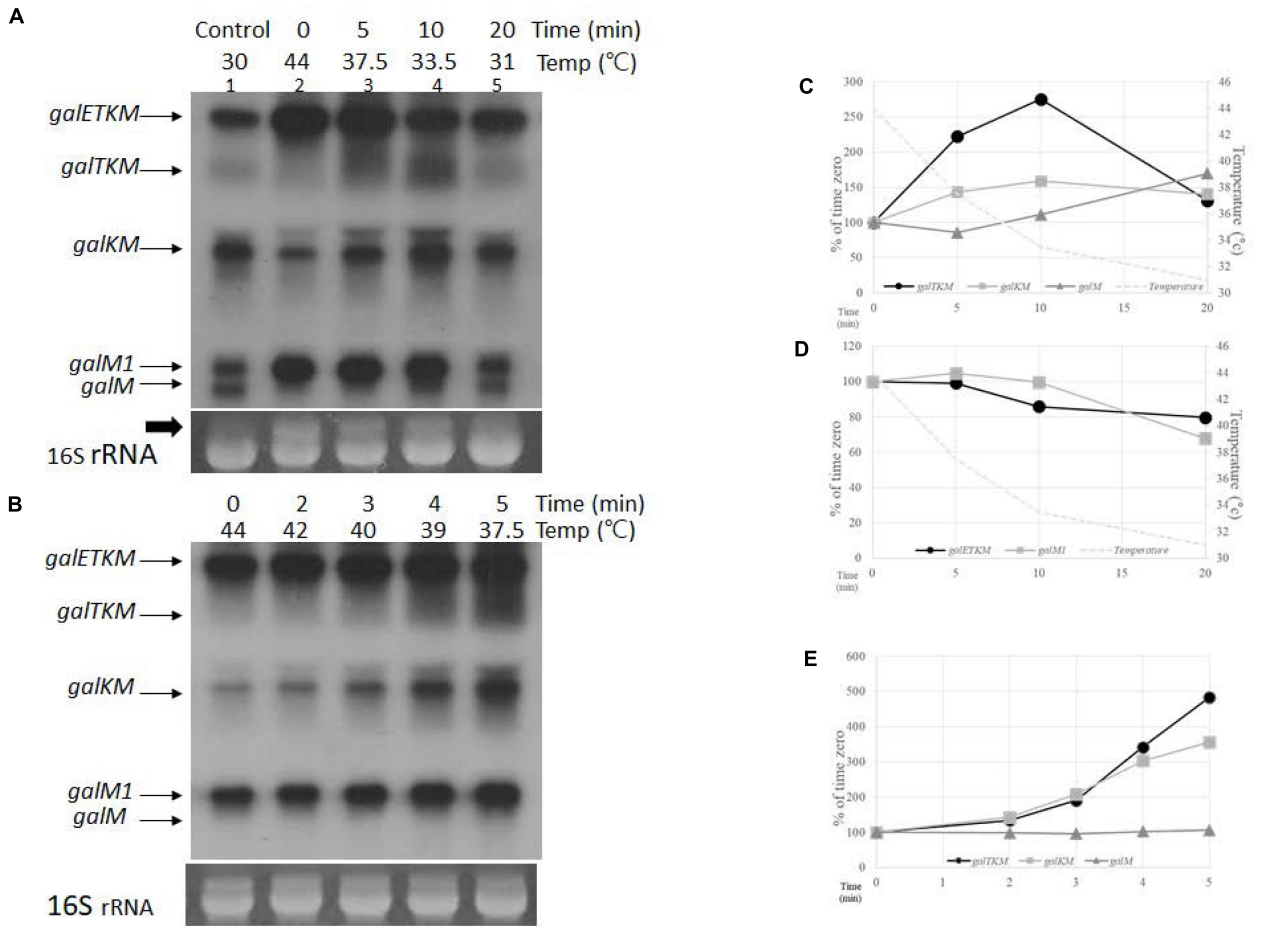


FIGURE 4 | Temporal order of cistron elimination from the *galETKM* mRNA. **(A)** Northern blot (M-probed) of *gal* mRNA prepared from GW20 (*ams1*^{ts}) cells taken at various times after temperature shift-down. The culture of GW20 cells grown up to A_{600} of 0.6 in Luria broth (LB) at 44°C was transferred to 30°C, and growth was allowed to continue. Aliquots of cells were taken at 0, 5, 10, and 20 min after transfer (lanes 2–5). The temperature of the culture at each time point is indicated. Lane 1 is *gal* mRNA from GW20 cells grown continuously to A_{600} of 0.6 in LB at 30°C. The arrow indicates the putative 17S rRNA, a precursor of 16S rRNA (Li et al., 1999), which were used here as loading controls. **(B)** Northern blot (M-probed) of *gal* mRNA from GW20 cells taken at 0, 2, 3, 4, and 5 min after temperature shift-down from 44 to 30°C. Temperature of the culture at each time point is indicated. **(C)** Changes in the amount of decay intermediate mRNA species after 5, 10, and 20 min of temperature shift-down. Amounts relative to that of time-zero were graphed; *galTKM* (dark circle), *galKM* (gray square), and *galM* (gray triangle). Temperature changes are indicated by a dotted line. **(D)** Changes in the amount of primary transcripts after 5, 10, and 20 min of temperature shift-down. Amounts relative to that of time-zero were graphed; *galETKM* (dark circle) and *galM1* (gray square). Temperature changes are indicated by a dotted line. **(E)** Changes in the amount of decay intermediate mRNA species after 2, 3, 4, and 5 min of temperature shift-down. Amounts relative to that of time-zero were graphed; *galTKM* (dark circle), *galKM* (gray square), and *galM* (gray triangle).

et al., 2017). We found that all the 5′ ends of decay intermediates indeed had the W residue (downward arrows) of the consensus RNase E cleavage sequence (underlined) (**Figures 6A–C**). The 5′ ends of *galTKM* mRNA at 1,031 and 1,039 are 13 and 5 nucleotides upstream from the putative Shine–Dalgarno (SD) sequence of the *galT* gene (**Figure 6A**). Similarly, in the *galKM* mRNA, the 5′ end is 12 or 20 nucleotides upstream from the putative SD sequence of the *galK* gene (**Figure 6B**), and in the *galM* mRNA, the 5′ end is 16, 22, or 24 nucleotides upstream from the putative SD sequence of the *galM* gene (**Figure 6C**).

These results suggest that translation initiation at the beginning of each cistron in *galETKM* could be responsible for the location of the observed RNase E cleavage sites. To test the effect of translation initiation on the observed RNase E cleavage sites, we mutated the putative SD sequences of *galT*, *galK*, and *galM* (bold letters) to their complementary sequences (**Figures 6A–C**), one SD sequence at a time, and assayed the production of decay intermediates as well as the primary transcripts. These SD mutations were generated in plasmid, p*gal*. The resulting mutant plasmids were called p*galT-SD*, p*galK-SD*, and p*galM-SD*, respectively. We performed northern blot of *gal* mRNA from the three SD mutant plasmid-carrying cells using the M-probe.

Results showed that the primary transcript *galETKM* production decreased gradually, farther down where the SD mutations were in the operon (**Figure 6D**). The *galETKM*

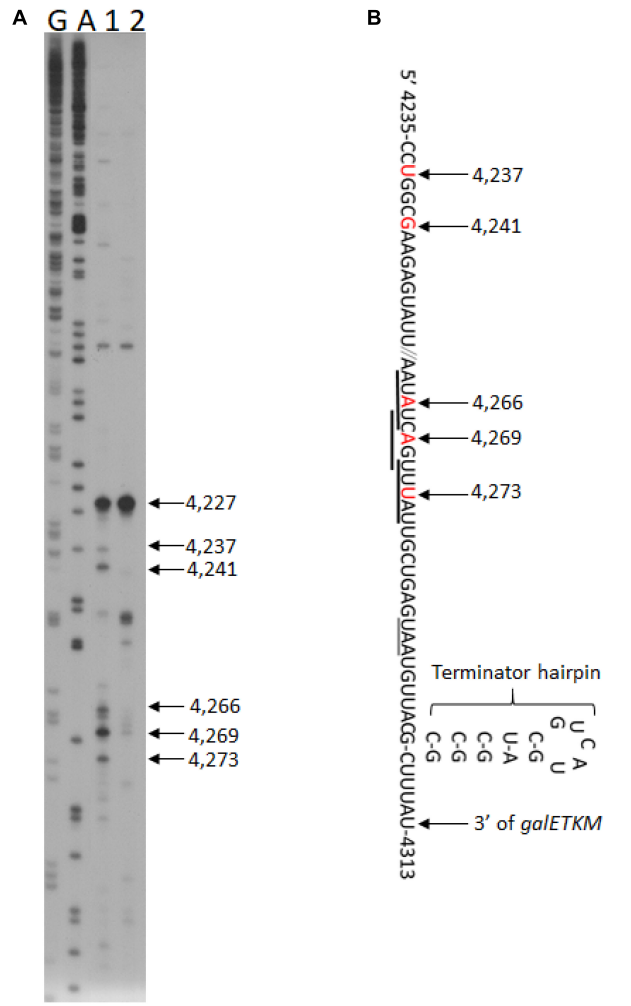


FIGURE 5 | RNase E cleavage upstream of the terminator hairpin of the *gal* operon. **(A)** 5′RACE of RNA in front of the terminator hairpin of *gal* operon isolated from GW20 (*ams1*^{ts}) cells grown either at the permissive temperature of 30°C (lane 1) or the non-permissive temperature of 44°C (lane 2). The 5′ end at 4,227 is not generated by RNase E cleavage since its intensity did not decrease at the non-permissive temperature. Lanes G and A are DNA sequencing ladders, as in **Figure 1E**. We note that the 5′ ends located between 4,273 and 4,266 also fall within the consensus sequence for the RNase E cleavage (Belasco, 2017; Chao et al., 2017), but the ends located at 4,241–4,237 seem to deviate from it. Nevertheless, both cleavage clusters would serve the purpose of removing the hairpin, allowing 3′ to 5′ exoribonuclease digestion. **(B)** Nucleotide sequence of mRNA at the end of *galM* presented together with their *gal* coordinates. The nucleotide residue and its *gal* coordinate that becomes the 5′ end after RNase E cleavage are indicated in red. The RNase E cleavage is indicated with a downward arrow. Matches to the consensus sequence for RNase E cleavage are underlined with thick horizontal lines. The translation stop codon of the *galM* gene is also underlined (thin line). Also shown is the terminator hairpin. The cytosine residue at 4,313 is the last residue of the full-length mRNA, *galETKM* (Wang et al., 2019).

was 60, 55, and 17% of the WT in *galT-SD*, *galK-SD*, and *galM-SD* mutants, respectively. It has been well documented that disruption of the transcription–translation coupling leads to premature Rho-dependent transcription termination (Adhya and Gottesman, 1978). This notion has been rigorously supported in a recent study (Zhu et al., 2019). We reasoned that the premature Rho-dependent transcription termination caused by the SD mutations was responsible for the decreased production of *galETKM*. To test, we performed northern blot of RNA from the SD mutant plasmid-carrying cells that have been treated with the Rho-inhibitor BCM (20 μg/ml for 10 min). The results showed that, upon BCM treatment, *galETKM* increased 20% in cells with p*galT-SD*, 50% in cells with p*galK-SD*, and 250% in cells with p*galM-SD* (**Supplementary Figure S3**). Thus, in the presence of BCM, SD mutants produced more *galETKM*, suggesting that Rho-dependent transcription termination could be primarily responsible for the decreased production of *galETKM*.

The northern blot of RNA from the SD mutant plasmid-carrying cells (**Figure 6D**) additionally showed that:

In p*galT-SD*-harboring cells (lane 2, **Figure 6D**), *galTKM* mRNA decreased to less than 20% of the WT level, while the other two intermediate mRNAs, *galKM* and *galM*, decreased to about 60–70% of the WT level (**Figure 6E**).
In p*galK-SD*-harboring cells (lane 3, **Figure 6D**), *galKM* mRNA decreased to 12% of the WT level, while the other two intermediate mRNAs, *galTKM* and *galM*, decreased to about 30–40% of the WT level (**Figure 6E**).
In p*galM-SD*-harboring cells (lane 4, **Figure 6D**), *galM* mRNA decreased less than 5% of the WT level, while the other two intermediate mRNAs, *galTKM* and *galKM*, decreased to about 10% of the WT level (**Figure 6E**).

These results show that, when an intermediate mRNA harbors SD mutation at its 5′ end, its amount is reduced the most than the other two intermediate mRNAs in every case (**Figure 6E**). The level of decrease of the other two intermediates appeared roughly similar to that of *galETKM* in all of the SD mutants, suggesting that the SD mutations, by decreasing the *galETKM* amount indirectly, caused an overall decrease of the intermediates. We have not correlated mRNA level with protein level, which remains an important task for the future.

Translation Initiation at the 5′ End of Decay Intermediates Promotes Their Stability

To test the effect of blocking translation initiation further, we measured the half-lives of intermediate mRNA species from cells with the WT or the SD-mutated p*gal* plasmids. MG1655Δ*gal* cells harboring WT or the mutant p*gal* plasmids were grown in LB to OD_{600} of 0.6, at which time rifampicin was added at 100 μg/ml to block new RNA synthesis. Cells were harvested every minute or two, and total RNA was subjected to northern blot analysis using the M-probe. From these blots, we measured the half-lives of the intermediates as well as the full-length *galETKM* mRNA (**Figure 7** and **Table 1**).

The results are as follows:

galETKM: Its half-life in all the SD mutant plasmid-carrying cells was nearly the same as in WT p*gal*-carrying cells, which was 1.82 ± 0.11 min (**Figure 7A** and **Table 1**).

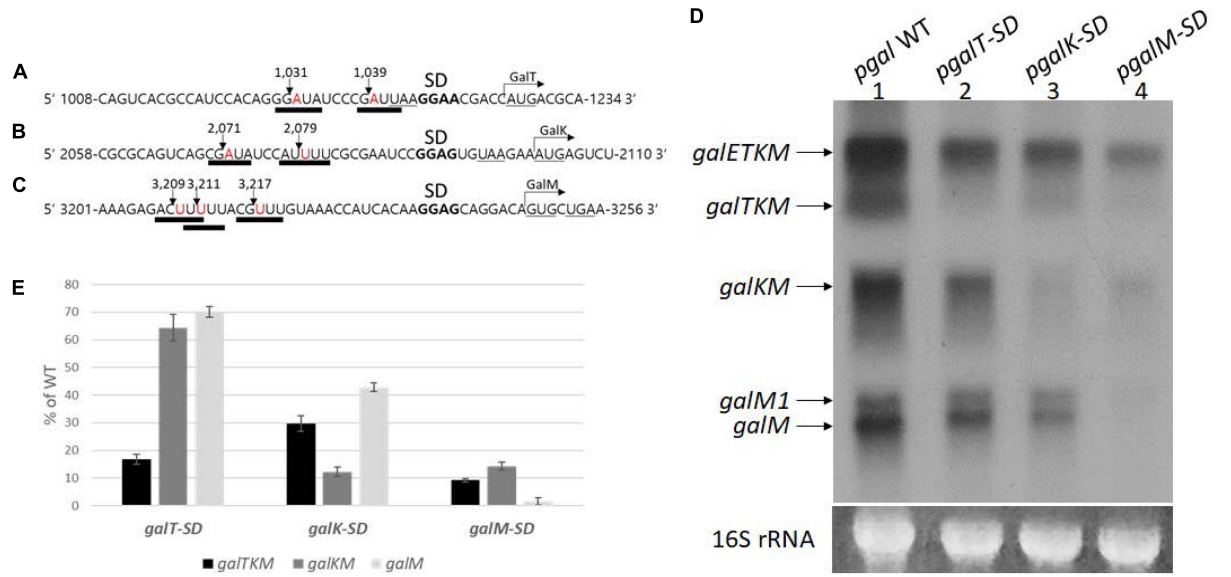


FIGURE 6 | Effect of translation initiation on the production of mRNA intermediates. **(A–C)** Sequence features at cistron boundaries in *gal* mRNA, showing the correspondence of the cleavage positions (inverted arrows) of the samples in **Figure 2A** to the RNase E consensus cleavage sites (thick underlines). Also highlighted are the putative Shine–Dalgarno sequence (in bold) and initiation (AUG or GUG) and termination (UAA or UGA) codons (underlined). **(D)** Northern blot of *gal* mRNA prepared from MG1655Δ*gal* cells harboring p*gal* (lane 1) or its mutant derivative plasmids, p*galT-SD* (lane 2), p*galK-SD* (lane 3), and p*galM-SD* (lane 4). The blot was probed with the M probe. **(E)** Quantification of each decay intermediate mRNA from **(D)** using the software ImageJ (NCBI, NIH). Note that when mRNA was extracted from cells with p*galT-SD* (lane 2 in **D**), of the three intermediates, *galTKM*, *galKM*, and *galM*, *galTKM* mRNA decreased the most, compared to their levels in cells with the wild-type plasmid p*gal*. Similarly, when the plasmid was p*galK-SD* (lane 3 in **D**), the *galKM* mRNA decreased the most, and the same trend was seen when the plasmid was p*galM-SD* (lane 4 in **D**).

galTKM: Its half-life in p*galK-SD*- and p*galM-SD*-carrying cells was nearly the same as in WT p*gal*-carrying cells, which was 1.50 ± 0.08 min (**Figure 7B** and **Table 1**). However, when the SD mutation was in *galT*, the half-life of *galTKM* decreased to an unmeasurably low value.

galKM: Its half-life in p*galT-SD*- and p*galM-SD*-carrying cells was nearly the same as in the WT p*gal*-carrying cells, which was 2.82 ± 0.14 min (**Figure 7C** and **Table 1**). However, when the SD mutation was in *galK*, the half-life of *galKM* decreased to 1.31 ± 0.07, significantly below its value in WT p*gal*-carrying cells of 2.82 ± 0.14 (**Table 1**).

galM: Its half-life in p*galT-SD*- and p*galK-SD*-carrying cells was nearly the same as in the WT p*gal*-carrying cells, which was 1.86 ± 0.37 min (**Figure 7D** and **Table 1**). However, when the SD mutation was in *galM*, the half-life of *galM* decreased to an unmeasurably low value.

Since the *galETKM* mRNA half-life did not change significantly when the RNA had a defective translation initiation signal in any one of its internal cistrons, *galT*, *galK*, or *galM* (**Figure 7A**), this suggests that internal translation initiation does not affect RNase E cleavage significantly and thus the decay process of *galETKM*.

Studies with the mutant plasmids revealed that, unlike the SD mutations internal to the operon, when the mutations are present at the 5′ end of an intermediate mRNA, its half-life is specifically reduced (**Figures 7B,D**). These results suggest that translation initiation at the 5′ end of an intermediate mRNA could be one of the factors that confer stability to the intermediate mRNA. It is also clear that the instability due to SD mutations does not propagate to downstream cistrons, consistent with the view that translation initiation in those cistrons confers stability to them.

RNase E Cleavage That Results in *galTKM* Generation Requires 5′ Mono-Phosphate

Di-phosphorylated mRNA 5′ ends are abundant in *E. coli*. RppH removes the β-phosphate from di-phosphorylated mRNA 5′ ends and generates mono-phosphorylated mRNA 5′ ends (Luciano et al., 2017). Thus, in the absence of RppH, the percentage of full-length primary transcripts that are monophosphorylated drops to almost zero as measured by 5′RACE (Deana et al., 2008). To see the generation of the decay intermediates in a situation where most *galETKM* 5′ ends are tri- or di-phosphorylated, we performed northern blot using the M-probe of MG1655Δ*rppH* cells (where the gene for RppH is deleted from the chromosome) grown to OD_{600} of 0.6. The result showed that, while *galETKM* decreased to only 90% of its level in WT, *galTKM* and *galKM* dropped to undetectable and 40% of their level in WT, respectively. The results also showed that *galM* decreased to 80% of WT, but *galM1* hardly at all (**Figures 8A,C**). Thus, in a situation where most *galETKM* and *galM1* 5′ ends are tri- or di-phosphorylated, mainly the generation of *galTKM* was affected.

Using 5′RACE, we assayed the 5′ end of the *gal* transcript initiated from the two promoters of *gal*, *P1* and *P2* (Adhya and

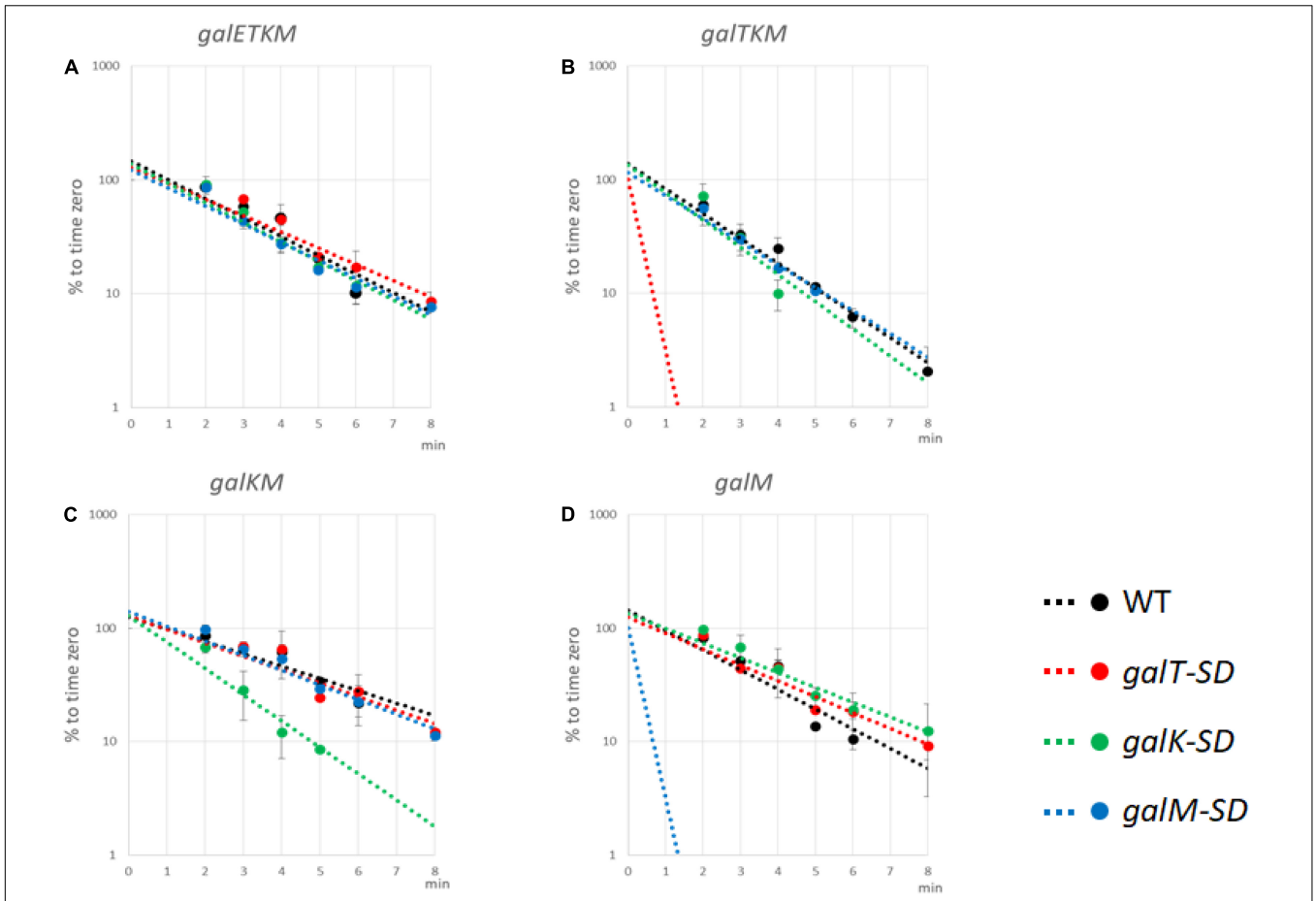


FIGURE 7 | Translation initiation at 5′ end of decay intermediates promotes their stability. The Shine–Dalgarno (SD) mutants as well as wild-type (WT) cells were grown in Luria broth with 0.5% galactose up to OD_{600} of 0.6. Rifampicin (100 μg/ml, final concentration) was added, and thereafter aliquots of the culture were taken at times indicated in the abscissa of the graphs. RNA prepared from the culture samples were subjected to northern blot and probed with M-probe. RNA bands were quantified and plotted against the time of their sampling. Northern blot analyses were repeated at least two times. The decay rate of *galETKM* **(A)**, *galTKM* **(B)**, *galKM* **(C)**, and *galM* **(D)** in WT (black), *galT-SD* (red), *galK-SD* (green), and *galM-SD* (blue) is presented. Note that the decay rate of the primary transcript, *galETKM*, did not change in any of the SD mutants, but the decay rate of an intermediate mRNA decreased significantly only when it harbored an SD mutation at its 5′ end. The dotted lines for *galTKM* in the *galT-SD* mutant [red in **(B)**] and for *galM* in the *galM-SD* mutant [blue in **(D)**] without any experimental points indicate that those RNAs were undetectable even at the earliest sampling time (2 min).

TABLE 1 | Half-life of decay intermediate mRNA (min).

Strain	*galETKM*	*galTKM*	*galKM*	*galM*
Wild type	1.82 ± 0.11	1.50 ± 0.08	2.82 ± 0.14	1.86 ± 0.37
galT-SD	1.77 ± 0.10	ND	2.10 ± 0.24	1.89 ± 0.01
galK-SD	1.76 ± 0.06	1.14 ± 0.10	1.31 ± 0.07	2.29 ± 0.25
galM-SD	1.78 ± 0.07	1.52 ± 0.09	2.56 ± 0.38	ND

Half-life was measured from two independent northern blots.
ND, not determined.

Miller, 1979), in MG1655Δ*rppH*. The 5′RACE results showed that the 5′ ends of *gal* transcripts that are monophosphorylated drastically reduces in MG1655Δ*rppH* compared to their levels in the WT (**Figure 8B**). Since the 5′RACE assay requires mono-phosphorylated mRNA 5′ end, the 5′RACE assay results indicate that most *galETKM* in MG1655Δ*rppH* are tri- or di-phosphorylated at the 5′ end. Taken together, these results suggest that the reason for the lack of generation of *galTKM* and, to a significant extent, *galKM* is because of the lack of mono-phosphorylation of the 5′ end of *galETKM*. In other words, the 5′ mono-phosphorylated *galETKM* appears to be the precursor of *galTKM* and *galKM*, as was suggested by the results of the TSD experiments (**Figure 4**). Since *galKM* is still generated in Δ*rppH* cells, albeit in reduced amounts, this indicates that there

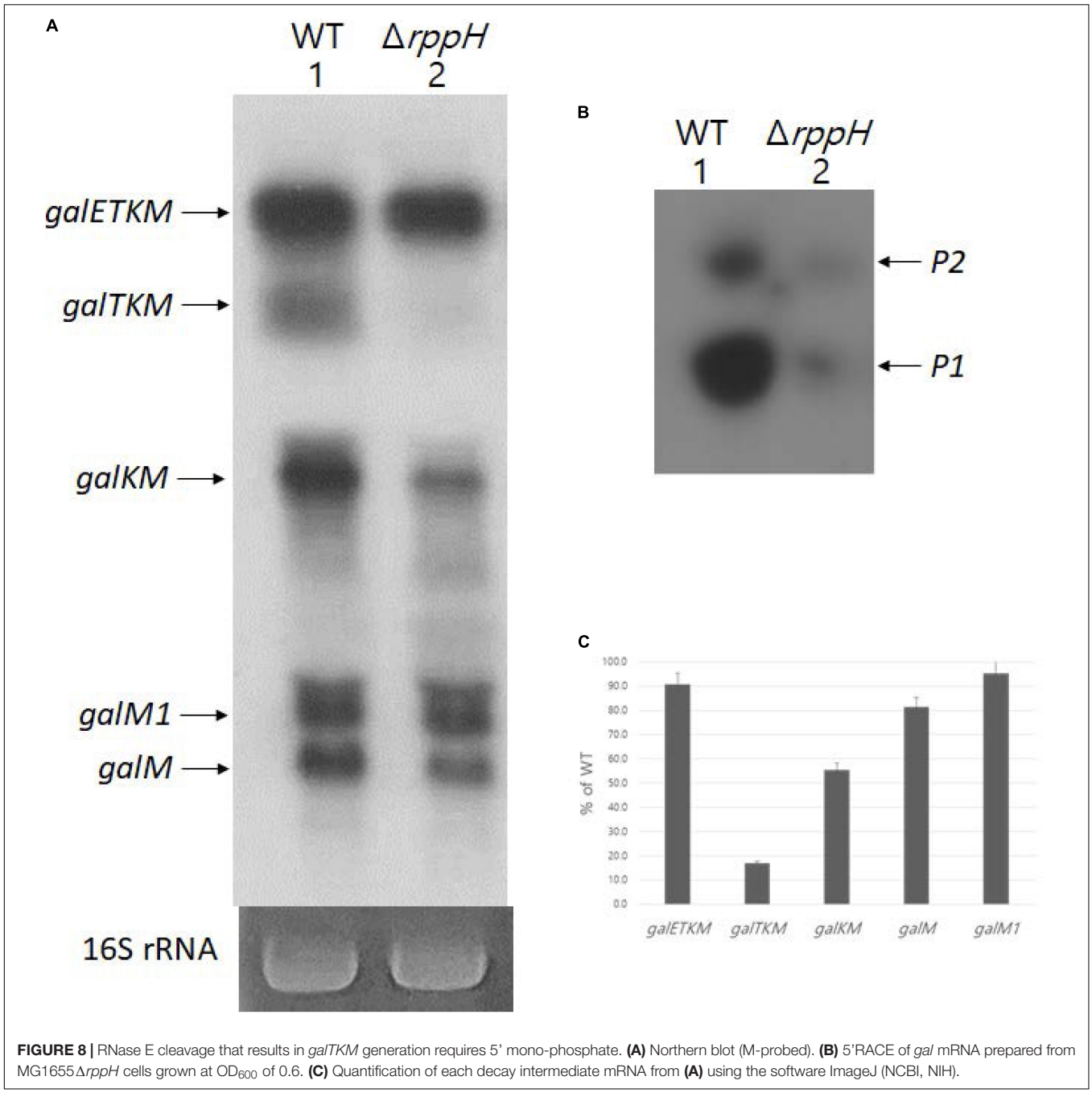


FIGURE 8 | RNase E cleavage that results in *galTKM* generation requires 5' mono-phosphate. **(A)** Northern blot (M-probed). **(B)** 5'RACE of *gal* mRNA prepared from MG1655Δ*rppH* cells grown at OD_{600} of 0.6. **(C)** Quantification of each decay intermediate mRNA from **(A)** using the software ImageJ (NCBI, NIH).

is a second pathway of generating them which does not depend on a mono-phosphorylated 5′ end.

DISCUSSION

The *gal* Operon Produces Two Classes of mRNA: 5′-End-Sharing and 3′-End-Sharing

Northern blot analyses revealed that the *gal* operon produces two kinds of mRNA species that are 5′-end-sharing or 3′-end-sharing, as depicted in **Figure 1A**. Northern blot of WT *E. coli* cells probed with the E-probe, which hybridizes to the first half of the first gene *galE*, shows the 5′-end-sharing mRNA (lane 1 in **Figure 3**). However, if the same blot is probed with the M-probe, which hybridizes to the second half of the last gene *galM*, it shows the 3′-end-sharing mRNA (lane 1 in **Figure 1D**). We investigated how the 5′-end-sharing mRNA species are generated in all of our previous studies (Lee et al., 2008; Wang et al., 2014, 2015, 2019). In this study, we focused on how the 3′-end-sharing mRNA species are generated.

RNase E Cleavage at Cistron Boundaries

RNase E cleavage at cistron boundaries has been reported in many different operons. Representative examples are in between *dnaG* and *rpoD* (Burton et al., 1983) in the polycistronic *rxc* mRNA (Belasco et al., 1985), in between *papB* and *papA* in the *pap* operon (Baga et al., 1988), and between cistrons of the *his* operon (Alifano et al., 1994b). Recently, using transcriptome analysis, Dar and Sorek demonstrated that RNase E cleavage at the cistron boundaries occur in many different operons in *E. coli* (Dar and Sorek, 2018).

Using the sequence information downstream of the 5′ end of 22,000 RNA generated by RNase E cleavage in *Salmonella* (Chao et al., 2017), we searched for *Salmonella* operons harboring the RNase E-generated 5′ end sequences 40 nucleotides upstream from the initiator codon of internal cistrons. The results showed that 182 operons (out of 881 operons investigated) have RNase E-cleaved 5′ end at cistron boundaries, a few codons upstream from the putative SD sequence of the downstream genes (**Supplementary Figure S4**). Many operons have RNase E-cleaved 5′ end at multiple cistron boundaries in the same operon (**Supplementary Figure S4**). Thus, the RNase E cleavages at the cistron boundaries shown in this study appear to be a common practice of mRNA decay of polycistronic mRNA in *E. coli* as well as in *Salmonella*. Of the two activities of RNase E, processing and degradation, generation of stable decay intermediates can also be thought of as a processing event rather than a destructive event.

Polycistronic *galETKM* mRNA Decays Preferentially Through Cistron Elimination From the 5′ End

Since the initial proposal by Apirion (1973) that mRNA decay in *E. coli* proceeds in 5′→3′ direction, in spite of the lack of any 5′ to 3′ exoribonucleases in *E. coli* (Deutscher, 1985), it is generally believed that mRNA decay in *E. coli* initiates from the 5′ end and proceeds toward the 3′ end (Mackie, 2013; Hui et al., 2014; Mohanty and Kushner, 2016). This is because RNase E has a preference to initiate decay by binding first to the 5′ mono-phosphorylated end. From our limited kinetic analysis of precursor-to-product formation, most of the data could be interpreted as sequential elimination of cistrons from the 5′ end of *galETKM*. A clear exception was also found: About 40% of *galKM* that was dependent on RNase E, but not on 5′ monophosphate formation, was produced (**Figures 2A**, **4E**, **8A**). It appears that RNase E can also directly access internal sites within a polycistronic message. As we elaborate below, although overall there is a preference for elimination of cistron from the 5′ end, the process may not be always sequential.

Use of Two Modes by RNase E in the Elimination of Cistrons From *galETKM* mRNA

Based on the data presented in this study, we propose the following model for the decay of *galETKM* mRNA (**Figure 9**).

Progressive Elimination of *galE* From *galETKM* That Generates *galTKM*

RNase E is a large tetramer and forms a multi-protein complex called the degradosome (Py et al., 1996; Callaghan et al., 2005). Its binding to the 5′ end of mRNA causes steric hindrance to nearby translation initiation (Deana and Belasco, 2005). Inhibition of new translation initiation will make the already initiated ribosomes to retreat away from the 5′ end and expose RNase E cleavage sites which were otherwise to be occluded by translating ribosomes (**Figure 9A**). When RNase E cleaves the first time, it would generate a new 5′mono-phosphate end, to which RNase E can bind. This cleavage and binding can repeat until the cleavage site approaches the *galT* translation initiation site. The RNase E cleavage labeled as circle 1 would generate *galTKM* (**Figure 9A**).

In our temperature shift-down experiments, *galTKM* appears to be produced by the gradual shortening of the full-length *galETKM* mRNA during the 5- to 20-min time period (**Figure 4A**). The trailing of the *galTKM* band (the material between the bands labeled *galETKM* and *galTKM*) is most conspicuous in lane 3 of **Figure 4A**. The results of 5′RACE indicate that there are two major 5′ ends of *galTKM* located at 1,031 and 1,039, but there are also longer-length species located at ∼890 and 955–980 (**Figure 1E**), which can be interpreted to represent progressive RNase E cleavage on the *galE* part of *galETKM* during the *galTKM* formation process. We did find a consensus of RNase E cleavage sequences near 890, but none at 955–980. This mechanism is analogous to RNase E scanning activity on RNA from the 5′ end before it cleaves the message (Richards and Belasco, 2019). The generation of *galTKM* was also drastically reduced in Δ*rppH* mutant (**Figure 8**) in support of the 5′ end-dependent pathway. Thus, the 5′ end-dependent pathway appears to be used exclusively to generate *galTKM* from *galETKM* (**Figure 9A**).

One-Step Elimination of *galET* From 5′-(P)PP–*galETKM* Generates *galKM*

In comparison to *galTKM*, the *galKM* band was less smeary in northern blots (**Figure 4A**). Only two closely spaced ends of *galKM* at 2,071 or 2,079 were also revealed by 5′RACE (**Figure 1E**). The appearance of *galTKM* and *galKM* at the same time and the RppH-independent accumulation of *galKM* through direct-entry processing of a triphosphorylated *galETKM* precursor suggest that RNase E generates *galKM* directly from full-length *galETKM* and not from intermediate length precursors, as observed for *galTKM*. Since *galTKM* and *galKM* appear at the same time, they are most likely produced independently from different *galETKM* RNAs. Alternatively, it is also possible that, once it has formed, *galTKM* is rapidly cleaved to produce *galKM* in a step that is not rate limiting.

Since a significant amount of *galKM* is generated in Δ*rppH* cells (**Figure 8**), it indicates that the direct access pathway was responsible for the generation. Since the level of production was different in the WT and Δ*rppH* cells, this indicates that both the direct access (by RNase E cleavage labeled as circle 2 in **Figure 9B**) and the 5′ end-dependent (by the circle 3 RNase E cleavage in **Figure 9A**) pathways are employed to generate *galKM*. Note that,

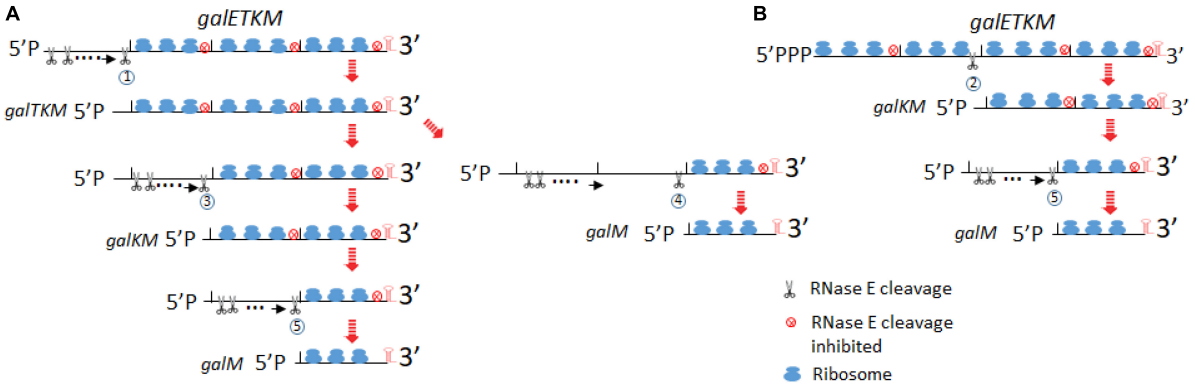


FIGURE 9 | Model for decay of the polycistronic mRNA, *galETKM*. The model depicts two different ways to eliminate a cistron(s) from the 5′ end of *galETKM*. In **(A)**, the *galE* gene is progressively eliminated from the *galETKM* mRNA by the 5′ end-dependent pathway. We propose that RNase E binds to the mono-phosphorylated 5′ end of the *galETKM* mRNA. The RNase E binding causes a cascade of events: (1) inhibition of translation initiation of *galETKM*, (2) receding of already initiated ribosomes from the 5′ end, which would reveal RNase E cleavage sites in the ribosome-free regions of mRNA, and (3) cleavage. Because the newly generated 5′ ends will be mono-phosphorylated, RNase E can again bind and cleave, and the cycle can continue until near the end of the cistron at 1,031 or 1,039, generating *galTKM* by RNase E cleavage labeled 1. By the same process, *galKM* and *galM* (RNase E cleavage labeled 3 and 4, respectively) could be generated. *galM* could also be generated from *galKM* by RNase E cleavage labeled 5. RNase E cleavage 6 would complete the decay of *galETKM*. In **(B)**, we depict the simultaneous elimination of *galE* and *galT* from the *galETKM* mRNA by the direct access pathway. We also propose that the tri-phosphorylated *galETKM* mRNA can stochastically become ribosome-free in the internal cistrons. For example, when the *galE–galT* part becomes ribosome-free, RNase E cleavage labeled 2 at 2,071 or 2,079 directly generates *galKM*. The results in this study suggested that *galKM* could be generated from *galTKM* by RNase E cleavage 3. *galM* could be generated by RNase E cleavage 4 and 5 from *galTKM* and *galKM*, respectively.

in *E. coli*, ∼35–50% of each mRNA is di-phosphorylated, and the rest is mono-phosphorylated at the 5′ end (Luciano et al., 2017), and the direct access pathway of mRNA decay has been demonstrated as a major pathway for mRNA decay in *E. coli* (Clarke et al., 2014).

Elimination of a Cistron(s) to Generate *galM*

From the temperature shift-down experiment, we propose the following pathways for *galM* mRNA generation. By the sequence of appearance, *galM* could be generated either by the simultaneous removal of the *galTK* part from *galTKM* (the RNase E cleavage labeled as circle 4 in **Figure 9A**) or by removing the *galK* part from *galKM* (the cleavage labeled as circle 5 in **Figure 9A**). The decrease of *galTKM* and the increase of *galM* from 10 to 20 min time points in one hand and relatively no change in *galKM* amount on the other may suggest that *galM* is derived directly from *galTKM* by the cleavage 4, but the alternate pathway, indicated by the cleavage 5, such that *galM* is primarily derived from *galKM*, which is derived from *galTKM*, cannot be eliminated (**Figure 9A**). *galKM* amounts are not expected to change if the rate of decay of *galTKM* to *galKM* and *galKM* to *galM* is similar. The generation of *galM* in Δ*rppH* cells does not necessarily indicate that it is produced by direct access if the precursors are *galTKM* and/or *galKM*, which have a mono-phosphorylated 5′ end. This indicates that *galM* is derived from the 5′ end-dependent pathway. The smearing RNA below the *galKM* band in the first 5 min (**Figure 4B**) or at 10 min (**Figure 4A**) of TSD suggests that the *galK* part of the *galKM* mRNA could be progressively removed from the 5′ end (as indicated in the cleavage 5). However, the smearing does not extend all the way down to the *galM* band (**Figure 4B**), suggesting that the progressive elimination of *galK* proceeds to a certain point of *galK* before RNase E cleaves at 3,209–3,217 (**Figure 2B**) to generate *galM*.

Elimination of *galM*

Elimination of *galM* is the last step in the sequence of cistron elimination, completing the decay process of *galETKM* mRNA. The *galM* band also smears to lower-size species to some extent, indicating that it decays progressively (**Figure 4B**). However, RNase E could also directly access the cleavage sites at 4,266–4,273.

Failure in Translation Initiation at the 5′ End of Polycistronic mRNA Initiates mRNA Decay

Our results here demonstrate that interfering with translation initiation at the 5′ end of each intermediate mRNA greatly increases the message decay rate (**Figure 7** and **Table 1**). Because translating ribosomes inhibit RNase E cleavage (Deana and Belasco, 2005) and RNase E cleaves mRNA regions free of ribosome (Iost and Dreyfus, 1995), the "absence of ribosome" seems obligatory for the cleavage. Inhibiting translation initiation seems the most reasonable way to generate ribosome-free regions within a message. Creating ribosome-free regions within an ORF undergoing translation other than by inhibiting translation initiation would lead to the production of truncated proteins, which would be wasteful (Keiler et al., 1996; Muto et al., 1996; Hayes et al., 2002). Thus, RNase E cleavage on *galTKM* or *galKM* to generate *galM* is most likely preceded by inhibition of translation initiation at the 5′ end.

The same decay strategy could be applied to the primary (full-length) polycistronic mRNA. Considering that both of the

5′ end-dependent and direct access pathways work on the primary mRNA, *galETKM*, it is likely that the phosphorylation status of 5′ end influences which pathway RNase E would take to initiate the cleavage. Aside from occasional stochastic failure to initiate translation, there could be formation of a stem-loop structure as in the 5′ UTR (Arnold et al., 1998), other secondary structures, or sRNA binding that could inhibit translation initiation at the 5′ end of many mRNAs. For example, sRNA Spot 42 binding to the 5′ end of one of the intermediate *gal* mRNA, *galKM* (mK2), expedites the decay of *galKM* (Wang et al., 2015). Even proteins such as RNase E itself could inhibit translation initiation (Deana and Belasco, 2005). We propose that it is these translation initiation-inhibiting factors at the 5′ end of polycistronic mRNA that induce cleavage by making the RNase E cleavage sites ribosome-free. Our experiments also revealed that the cistron elimination did not proceed to the downstream cistrons possibly because of their translation initiation signals.

In all the cartoons of **Figure 9**, a cistron was eliminated when it was the 5′-most cistron and ribosome-free. This was evident in experiments of **Figure 6** where translation initiation signals were mutated. For example, in the case of the *galT-SD* mutant, whatever pathway generates *galTKM* from *galETKM*, the intermediate is eliminated selectively if it did not have a translation initiation signal of *galT*. However, *galTKM* was not preferentially eliminated, although it was the 5′-most cistron when it had the *galK-SD* mutation (**Figure 6**). Direct entry of RNase E into the ribosome-free *galK* region should have eliminated the *galTKM* intermediate together with *galKM*, but apparently that did not happen. In other words, the absence of translation, although necessary, is not sufficient for decay.

In sum, even in a four-cistron mRNA, there can be progressive and one-step elimination of cistrons and no elimination of internal cistrons even in the absence of translation. Although the picture is complex, vulnerability of the 5′-most cistrons appears to be the most consistent theme of the present study, in line with the general thinking that decay proceeds in the 5′→3′ direction.

Future Studies on *gal* mRNA Decay

Intermediates of *gal* mRNA other than those described here are also produced by agents such as Spot 42 small RNA and transcription terminator factor. For example, in MG1655Δ*spf* cells, where the gene for Spot 42 is deleted from the chromosome, the production of *galKM* mRNA is reduced compared to that seen in the WT, and this is because Spot 42 expedites the decay of the *galKM* mRNA (Wang et al., 2015). Based on these results, we propose that Spot 42 binding might expedite the decay of those *gal* mRNA species, such as *galETKM*, *galTKM*, and *galKM*, that harbor the Spot 42 binding site at the *galT–galK* cistron junction. We are currently testing this prediction. We are also currently investigating whether in the generation of the 5′ end of *galKM* there is a relationship between the RNase P cleavage at 1,764 and 1,777 (Wang et al., 2014) and the RNase E cleavage at 2,071 and 2,079 (**Figure 1E**).

DATA AVAILABILITY STATEMENT

All datasets presented in this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

HL designed the research. HJ, CK, MN, and YL performed the research. HJ, CK, MN, YL, XW, DC, and HL analyzed the data. DC and HL wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by a grant (2017-1400-01) of the Basic Science Research Program of the National Research Foundation of Korea (NRF). This work was partially supported by the Intramural Research Program of the Center for Cancer Research, NCI, NIH.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Taeok Bae (Indiana University, United States) for his critical reading of the manuscript and editorial suggestions.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2020.586413/full#supplementary-material

Supplementary Figure 1 | Newly found internal promoter *M* and the transcript *galM1*.

Supplementary Figure 2 | Effect of bicyclomycin, the Rho-inhibitor, on the production of *gal* mRNA.

Supplementary Figure 3 | Production of *galETKM* in the Shine–Dalgarno mutants in the presence of bicyclomycin.

Supplementary Figure 4 | Search of *Salmonella* operon harboring the 5′ end of mRNA at cistron boundaries.

Supplementary Table 1 | A list of primers used in this study.

REFERENCES

Adhya, S. (2003). Suboperonic regulatory signals. *Sci. STKE* 2003:e22.

Adhya, S., and Gottesman, M. (1978). Control of transcription termination. *Annu. Rev. Biochem.* 47, 967–996. doi: 10.1146/annurev.bi.47.070178.004535

Adhya, S., and Miller, W. (1979). Modulation of the two promoters of the galactose operon of *Escherichia coli*. *Nature* 279, 492–494. doi: 10.1038/279492a0

Alifano, P., Bruni, C. B., and Carlomagno, M. S. (1994a). Control of mRNA processing and decay in prokaryotes. *Genetica* 94, 157–172. doi: 10.1007/bf01443430

Alifano, P., Rivellini, F., Piscitelli, C., Arraiano, C. M., Bruni, C. B., and Carlomagno, M. S. (1994b). Ribonuclease E provides substrates for ribonuclease P-dependent processing of a polycistronic mRNA. *Genes Dev.* 8, 3021–3031. doi: 10.1101/gad.8.24.3021

Apirion, D. (1973). Degradation of RNA in *Escherichia coli*. A hypothesis. *Mol. Gen. Genet.* 122, 313–322. doi: 10.1007/bf00269431

Arnold, T. E., Yu, J., and Belasco, J. G. (1998). mRNA stabilization by the ompA 5′ untranslated region: two protective elements hinder distinct pathways for mRNA degradation. *RNA* 4, 319–330.

Babitzke, P., and Kushner, S. R. (1991). The Ams (altered mRNA stability) protein and ribonuclease E are encoded by the same structural gene of *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 88, 1–5. doi: 10.1073/pnas.88.1.1

Baga, M., Goransson, M., Normark, S., and Uhlin, B. E. (1988). Processed mRNA with differential stability in the regulation of *E. coli* pilin gene expression. *Cell* 52, 197–206. doi: 10.1016/0092-8674(88)90508-9

Baker, K. E., and Mackie, G. A. (2003). Ectopic RNase E sites promote bypass of 5′-end-dependent mRNA decay in *Escherichia coli*. *Mol. Microbiol.* 47, 75–88. doi: 10.1046/j.1365-2958.2003.03292.x

Belasco, J. G. (2017). Ribonuclease E: chopping knife and sculpting tool. *Mol. Cell* 65, 3–4. doi: 10.1016/j.molcel.2016.12.015

Belasco, J. G., Beatty, J. T., Adams, C. W., Von Gabain, A., and Cohen, S. N. (1985). Differential expression of photosynthesis genes in R. capsulata results from segmental differences in stability within the polycistronic rxcA transcript. *Cell* 40, 171–181. doi: 10.1016/0092-8674(85)90320-4

Belasco, J. G., Nilsson, G., Von Gabain, A., and Cohen, S. N. (1986). The stability of E. coli gene transcripts is dependent on determinants localized to specific mRNA segments. *Cell* 46, 245–251. doi: 10.1016/0092-8674(86)90741-5

Bouvet, P., and Belasco, J. G. (1992). Control of RNase E-mediated RNA degradation by 5′-terminal base pairing in *E. coli*. *Nature* 360, 488–491. doi: 10.1038/360488a0

Braun, F., Le Derout, J., and Regnier, P. (1998). Ribosomes inhibit an RNase E cleavage which induces the decay of the rpsO mRNA of *Escherichia coli*. *EMBO J.* 17, 4790–4797. doi: 10.1093/emboj/17.16.4790

Burton, Z. F., Gross, C. A., Watanabe, K. K., and Burgess, R. R. (1983). The operon that encodes the sigma subunit of RNA polymerase also encodes ribosomal protein S21 and DNA primase in *E. coli* K12. *Cell* 32, 335–349. doi: 10.1016/0092-8674(83)90453-1

Callaghan, A. J., Marcaida, M. J., Stead, J. A., Mcdowall, K. J., Scott, W. G., and Luisi, B. F. (2005). Structure of *Escherichia coli* RNase E catalytic domain and implications for RNA turnover. *Nature* 437, 1187–1191. doi: 10.1038/nature04084

Carpousis, A. J., Luisi, B. F., and Mcdowall, K. J. (2009). Endonucleolytic initiation of mRNA decay in *Escherichia coli*. *Prog. Mol. Biol. Transl. Sci.* 85, 91–135. doi: 10.1016/s0079-6603(08)00803-9

Celesnik, H., Deana, A., and Belasco, J. G. (2007). Initiation of RNA decay in *Escherichia coli* by 5′ pyrophosphate removal. *Mol. Cell* 27, 79–90. doi: 10.1016/j.molcel.2007.05.038

Chao, Y., Li, L., Girodat, D., Forstner, K. U., Said, N., Corcoran, C., et al. (2017). In vivo cleavage map illuminates the central role of RNase E in coding and non-coding RNA pathways. *Mol. Cell* 65, 39–51. doi: 10.1016/j.molcel.2016.11.002

Clarke, J. E., Kime, L., Romero, A. D., and Mcdowall, K. J. (2014). Direct entry by RNase E is a major pathway for the degradation and processing of RNA in *Escherichia coli*. *Nucleic Acids Res.* 42, 11733–11751. doi: 10.1093/nar/gku808

Dar, D., and Sorek, R. (2018). Extensive reshaping of bacterial operons by programmed mRNA decay. *PLoS Genet* 14:e1007354. doi: 10.1371/journal.pgen.1007354

Deana, A., and Belasco, J. G. (2005). Lost in translation: the influence of ribosomes on bacterial mRNA decay. *Genes Dev.* 19, 2526–2533. doi: 10.1101/gad.1348805

Deana, A., Celesnik, H., and Belasco, J. G. (2008). The bacterial enzyme RppH triggers messenger RNA degradation by 5′ pyrophosphate removal. *Nature* 451, 355–358. doi: 10.1038/nature06475

Deutscher, M. P. (1985). E. coli RNases: making sense of alphabet soup. *Cell* 40, 731–732. doi: 10.1016/0092-8674(85)90330-7

Dreyfus, M. (2009). Killer and protective ribosomes. *Prog Mol Biol Transl Sci* 85, 423–466. doi: 10.1016/s0079-6603(08)00811-8

Emory, S. A., Bouvet, P., and Belasco, J. G. (1992). A 5′-terminal stem-loop structure can stabilize mRNA in *Escherichia coli*. *Genes Dev.* 6, 135–148. doi: 10.1101/gad.6.1.135

Hayes, C. S., Bose, B., and Sauer, R. T. (2002). Stop codons preceded by rare arginine codons are efficient determinants of SsrA tagging in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 99, 3440–3445. doi: 10.1073/pnas.052707199

Hui, M. P., Foley, P. L., and Belasco, J. G. (2014). Messenger RNA degradation in bacterial cells. *Annu. Rev. Genet* 48, 537–559. doi: 10.1146/annurev-genet-120213-092340

Iost, I., and Dreyfus, M. (1995). The stability of *Escherichia coli* lacZ mRNA depends upon the simultaneity of its synthesis and translation. *EMBO J.* 14, 3252–3261. doi: 10.1002/j.1460-2075.1995.tb07328.x

Jiang, X., and Belasco, J. G. (2004). Catalytic activation of multimeric RNase E and RNase G by 5′-monophosphorylated RNA. *Proc. Natl. Acad. Sci. U.S.A.* 101, 9211–9216. doi: 10.1073/pnas.0401382101

Joyce, S. A., and Dreyfus, M. (1998). In the absence of translation, RNase E can bypass 5′ mRNA stabilizers in *Escherichia coli*. *J. Mol. Biol.* 282, 241–254. doi: 10.1006/jmbi.1998.2027

Keiler, K. C., Waller, P. R., and Sauer, R. T. (1996). Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. *Science* 271, 990–993. doi: 10.1126/science.271.5251.990

Lee, H. J., Jeon, H. J., Ji, S. C., Yun, S. H., and Lim, H. M. (2008). Establishment of an mRNA gradient depends on the promoter: an investigation of polarity in gene expression. *J. Mol. Biol.* 378, 318–327. doi: 10.1016/j.jmb.2008.02.067

Lewis, D. E., and Adhya, S. (2015). Molecular mechanisms of transcription initiation at gal promoters and their multi-level regulation by GalR, CRP and DNA Loop. *Biomolecules* 5, 2782–2807. doi: 10.3390/biom5042782

Li, Z., Pandit, S., and Deutscher, M. P. (1999). RNase G (CafA protein) and RNase E are both required for the 5′ maturation of 16S ribosomal RNA. *EMBO J.* 18, 2878–2885. doi: 10.1093/emboj/18.10.2878

Luciano, D. J., Vasilyev, N., Richards, J., Serganov, A., and Belasco, J. G. (2017). A Novel RNA Phosphorylation State Enables 5′ End-Dependent Degradation in *Escherichia coli*. *Mol. Cell* 67, 44–54e46.

Mackie, G. A. (1998). Ribonuclease E is a 5′-end-dependent endonuclease. *Nature* 395, 720–723. doi: 10.1038/27246

Mackie, G. A. (2013). RNase E: at the interface of bacterial RNA processing and decay. *Nat. Rev. Microbiol.* 11, 45–57. doi: 10.1038/nrmicro2930

McDowall, K. J., Lin-Chao, S., and Cohen, S. N. (1994). A+U content rather than a particular nucleotide order determines the specificity of RNase E cleavage. *J. Biol. Chem.* 269, 10790–10796.

McLaren, R. S., Newbury, S. F., Dance, G. S., Causton, H. C., and Higgins, C. F. (1991). mRNA degradation by processive 3′-5′ exoribonucleases in vitro and the implications for prokaryotic mRNA decay in vivo. *J. Mol. Biol.* 221, 81–95. doi: 10.1016/0022-2836(91)90806-h

Melefors, O., and von Gabain, A. (1991). Genetic studies of cleavage-initiated mRNA decay and processing of ribosomal 9S RNA show that the *Escherichia coli* ams and rne loci are the same. *Mol. Microbiol.* 5, 857–864. doi: 10.1111/j.1365-2958.1991.tb00759.x

Mohanty, B. K., and Kushner, S. R. (2016). Regulation of mRNA decay in bacteria. *Annu. Rev. Microbiol.* 70, 25–44. doi: 10.1146/annurev-micro-091014-104515

Mott, J. E., Galloway, J. L., and Platt, T. (1985). Maturation of *Escherichia coli* tryptophan operon mRNA: evidence for 3′ exonucleolytic processing after rho-dependent termination. *EMBO J.* 4, 1887–1891. doi: 10.1002/j.1460-2075.1985.tb03865.x

Mudd, E. A., Krisch, H. M., and Higgins, C. F. (1990). RNase E, an endoribonuclease, has a general role in the chemical decay of *Escherichia coli* mRNA: evidence that rne and ams are the same genetic locus. *Mol. Microbiol.* 4, 2127–2135. doi: 10.1111/j.1365-2958.1990.tb00574.x

Muto, A., Sato, M., Tadaki, T., Fukushima, M., Ushida, C., and Himeno, H. (1996). Structure and function of 10Sa RNA: trans-translation system. *Biochimie* 78, 985–991. doi: 10.1016/s0300-9084(97)86721-1

Newbury, S. F., Smith, N. H., and Higgins, C. F. (1987). Differential mRNA stability controls relative gene expression within a polycistronic operon. *Cell* 51, 1131–1143. doi: 10.1016/0092-8674(87)90599-x

Ono, M., and Kuwano, M. (1979). A conditional lethal mutation in an *Escherichia coli* strain with a longer chemical lifetime of messenger RNA. *J. Mol. Biol.* 129, 343–357. doi: 10.1016/0022-2836(79)90500-x

Py, B., Higgins, C. F., Krisch, H. M., and Carpousis, A. J. (1996). A DEAD-box RNA helicase in the *Escherichia coli* RNA degradosome. *Nature* 381, 169–172. doi: 10.1038/381169a0

Regnier, P., and Hajnsdorf, E. (1991). Decay of mRNA encoding ribosomal protein S15 of *Escherichia coli* is initiated by an RNase E-dependent endonucleolytic cleavage that removes the 3′ stabilizing stem and loop structure. *J. Mol. Biol.* 217, 283–292. doi: 10.1016/0022-2836(91)90542-e

Richards, J., and Belasco, J. G. (2019). Obstacles to scanning by RNase E govern bacterial mRNA lifetimes by hindering access to distal cleavage sites. *Mol. Cell* 74, 284–295e285.

Spickler, C., Stronge, V., and Mackie, G. A. (2001). Preferential cleavage of degradative intermediates of rpsT mRNA by the *Escherichia coli* RNA degradosome. *J. Bacteriol.* 183, 1106–1109. doi: 10.1128/jb.183.3.1106-1109.2001

Taraseviciene, L., Miczak, A., and Apirion, D. (1991). The gene specifying RNase E (rne) and a gene affecting mRNA stability (ams) are the same gene. *Mol. Microbiol.* 5, 851–855. doi: 10.1111/j.1365-2958.1991.tb00758.x

von Gabain, A., Belasco, J. G., Schottel, J. L., Chang, A. C., and Cohen, S. N. (1983). Decay of mRNA in *Escherichia coli*: investigation of the fate of specific segments of transcripts. *Proc. Natl. Acad. Sci. U.S.A.* 80, 653–657. doi: 10.1073/pnas.80.3.653

Wachi, M., Umitsuki, G., and Nagai, K. (1997). Functional relationship between *Escherichia coli* RNase E and the CafA protein. *Mol. Gen. Genet* 253, 515–519. doi: 10.1007/s004380050352

Wang, X., Ji, S. C., Jeon, H. J., Lee, Y., and Lim, H. M. (2015). Two-level inhibition of galK expression by Spot 42: degradation of mRNA mK2 and enhanced transcription termination before the galK gene. *Proc. Natl. Acad. Sci. U.S.A.* 112, 7581–7586. doi: 10.1073/pnas.1424683112

Wang, X., Ji, S. C., Yun, S. H., Jeon, H. J., Kim, S. W., and Lim, H. M. (2014). Expression of each cistron in the gal operon can be regulated by transcription termination and generation of a galk-specific mRNA, mK2. *J. Bacteriol.* 196, 2598–2606. doi: 10.1128/jb.01577-14

Wang, X., Monford Paul Abishek, N., Jeon, H. J., Lee, Y., He, J., Adhya, S., et al. (2019). Processing generates 3′ ends of RNA masking transcription termination events in prokaryotes. *Proc. Natl. Acad. Sci. U.S.A.* 116, 4440–4445. doi: 10.1073/pnas.1813181116

Zhu, M., Mori, M., Hwa, T., and Dai, X. (2019). Disruption of transcription-translation coordination in *Escherichia coli* leads to premature transcriptional termination. *Nat Microbiol.* 4, 2347–2356. doi: 10.1038/s41564-019-0543-1

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Jeon, Kang, N, Lee, Wang, Chattoraj and Lim. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 26 November 2020
doi: 10.3389/fmolb.2020.592747

Identification of Novel Thermosensors in Gram-Positive Pathogens

Pilar Fernández[1†], Alejandra Raquel Díaz[2], María Florencia Ré[1], Lucía Porrini[3], Diego de Mendoza[1,3], Daniela Albanesi[1,3†] and María Cecilia Mansilla[1,3†]*

[1] Instituto de Biología Molecular y Celular de Rosario (IBR-CONICET), Rosario, Argentina, [2] Departamento de Biología, Bioquímica y Farmacia, Universidad Nacional del Sur, Centro de Recursos Naturales Renovables de la Zona Semi-árida (CERZOS-CONICET), Bahía Blanca, Argentina, [3] Departamento de Microbiología, Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario, Rosario, Argentina

OPEN ACCESS

Edited by:
Manuel Espinosa,
Spanish National Research Council, Spain

Reviewed by:
Dmitry A. Los,
Timiryazev Institute of Plant Physiology (RAS), Russia
Andrea Villarino,
Universidad de la República, Uruguay
Ivo Konopasek,
Charles University, Czechia

***Correspondence:**
María Cecilia Mansilla
mansilla@ibr-conicet.gov.ar

†ORCID:
Pilar Fernández
orcid.org/0000-0001-7589-6559
Daniela Albanesi
orcid.org/0000-0003-4380-9152
María Cecilia Mansilla
orcid.org/0000-0002-1444-8661

Specialty section:
This article was submitted to Molecular Recognition, a section of the journal Frontiers in Molecular Biosciences

Received: 08 August 2020
Accepted: 28 October 2020
Published: 26 November 2020

Citation:
Fernández P, Díaz AR, Ré MF, Porrini L, de Mendoza D, Albanesi D and Mansilla MC (2020) Identification of Novel Thermosensors in Gram-Positive Pathogens. Front. Mol. Biosci. 7:592747. doi: 10.3389/fmolb.2020.592747

Temperature is a crucial variable that every living organism, from bacteria to humans, need to sense and respond to in order to adapt and survive. In particular, pathogenic bacteria exploit host-temperature sensing as a cue for triggering virulence gene expression. Here, we have identified and characterized two integral membrane thermosensor histidine kinases (HKs) from Gram-positive pathogens that exhibit high similarity to DesK, the extensively characterized cold sensor histidine kinase from *Bacillus subtilis*. Through *in vivo* experiments, we demonstrate that SA1313 from *Staphylococcus aureus* and BA5598 from *Bacillus anthracis*, which likely control the expression of putative ATP binding cassette (ABC) transporters, are regulated by environmental temperature. We show here that these HKs can phosphorylate the non-cognate response regulator DesR, partner of DesK, both *in vitro* and *in vivo*, inducing in *B. subtilis* the expression of the *des* gene upon a cold shock. In addition, we report the characterization of another DesK homolog from *B. subtilis*, YvfT, also closely associated to an ABC transporter. Although YvfT phosphorylates DesR *in vitro*, this sensor kinase can only induce *des* expression in *B. subtilis* when overexpressed together with its cognate response regulator YvfU. This finding evidences a physiological mechanism to avoid cross talk with DesK after a temperature downshift. Finally, we present data suggesting that the HKs studied in this work appear to monitor different ranges of membrane lipid properties variations to mount adaptive responses upon cooling. Overall, our findings point out that bacteria have evolved sophisticated mechanisms to assure specificity in the response to environmental stimuli. These findings pave the way to understand thermosensing mediated by membrane proteins that could have important roles upon host invasion by bacterial pathogens.

Keywords: thermosensor, gram positive pathogen, ABC transporter, two component system, signalling

INTRODUCTION

Organisms permanently sense and respond to environmental changes in order to adapt and survive. For this, bacteria employ a large number of two-component systems (TCS) consisting of pairs of sensor histidine kinases (HKs) and response regulators (RRs). TCSs perceive specific external signals including extreme pH, high salinity, antibiotics and temperature, among others, and usually

give a transcriptional response to the initial stimulus. Up to date, only few TCS responsible for low temperature perception in different organisms have been characterized (Jin et al., 1993; Suzuki et al., 2000; Braun et al., 2008; Chakraborty et al., 2010; Steele et al., 2012; Najnin et al., 2016). For more than 20 years, our group has been studying the DesK/DesR TCS responsible for cold adaptation in the Gram-positive model bacterium *Bacillus subtilis* (for a review see de Mendoza, 2014), turning it into a paradigmatic transduction pathway. This TCS is conserved among Firmicutes, including Gram-positive pathogens, where temperature adaptive strategies play a significant role. Although the thermosensing ability of these paralog proteins has not yet been addressed, they emerge as promising candidates to accomplish this function.

The HK DesK lacks an extracellular domain and works as a homodimer, with each protomer contributing with five transmembrane (TM) helices to the sensory domain. It perceives changes in cytoplasmic membrane properties associated to a temperature downshift, like fluidity and thickness (Cybulski et al., 2002, 2010; Porrini et al., 2014; Saita et al., 2015). Upon a cold shock, DesK adopts a kinase-ON state phosphorylating DesR, its cognate RR. DesR-P is able to bind to the promoter region of the *des* gene and induces its expression (Aguilar et al., 2001; Cybulski et al., 2004). *des* encodes a Δ5-desaturase (Δ5-Des), which introduces a *cis* double bond into the acyl-chain of phospholipids increasing cytoplasmic membrane fluidity (Altabe et al., 2003). In the context of a fluid membrane, DesK adopts its phosphatase catalytic state, dephosphorylating DesR-P and turning off *des* expression (Albanesi et al., 2004). Crystal structures of the DesKC:DesR complex, trapped in the phosphatase and the phosphotransferase functional states allowed to uncover the structural bases of HK phosphatase/phosphotransferase control. The DesK:DesR structures revealed that both partners contribute residues to build the catalytic centers for TCS-mediated phosphotransfer and dephosphorylation reactions. Moreover, it was established that the HK is the component that drives the construction of a given active site by selecting a conformation of the RR that is prone for transfer reactions on its catalytic aspartate. The transition from the phosphatase to the phosphotransferase state of the HK, switches the distance and orientation of the catalytic histidine side chain with respect to the RR aspartate, allowing for the distinct reaction mechanisms to take place and ensuring the unidirectional flow of information (Trajtenberg et al., 2016).

B. subtilis contains a paralog of the DesK/DesR TCS of unknown function, named YvfT/YvfU. The amino acid identity between YvfT and DesK is 37% while the identity between YvfU and DesR is 58% (**Figure 1** and **Supplementary Figure S1**). The *des* gene, target of DesK/DesR, is located right upstream of the *desK/desR* operon while upstream of *yvfT/yvfU*, there are two genes, *yvfR* and *yvfS*, coding for subunits of a putative ATP binding cassette (ABC) transporter (**Figure 2A**). YvfR is homologous to the nucleotide-binding domain of the transporter, while YvfS resembles the permease. The four-gene cluster, coding for the ABC transporter system and the TCS, is conserved among Firmicutes, including important human pathogens, like *Staphylococcus aureus*, *Staphylococcus haemolyticus*, *Staphylococcus epidermidis*, *Streptococcus pneumoniae*, *Streptococcus agalactiae*, *Bacillus cereus*, and *Bacillus anthracis* (**Figure 2B**). Interestingly, in spite of encoding proteins with complete different biological functions, the promoter regions of these ABC transporter coding genes and *des* are quite similar. All these promoters have three RR binding regions, with an inverted and a direct repeat similar to that described for the *des* promoter (P*des*) (**Figure 2C**; Najle et al., 2009).

In this work, we investigated the functionality of three TCS homologous to DesK/DesR: YvfT/YvfU from *B. subtilis*, BA5598/BA5597 from *B. anthracis* and SA1313/SA1314 from *S. aureus*. Using a *B. subtilis desK* knockout reporter strain, we demonstrated that the HKs YvfT, BA5598 and SA1313 are able to perceive a downshift in growth temperature. We also analyze the ability of these HKs to respond to changes in the cytoplasmic membrane physical properties, due to variations in lipid composition, at a constant temperature. We report that each HK responds to a particular set of stimuli with different sensitivity, probably because they belong to bacteria that inhabit different ecosystems. The *in vivo* approach shown here uncovered novel temperature responsive HKs found in Gram-positive pathogens.

MATERIALS AND METHODS

Bacterial Strains and Growth Conditions

Bacterial strains and plasmids used in this study are listed in **Table 1**. *Escherichia coli* and *B. subtilis* strains were routinely grown in Lysogeny Broth (LB) at 37°C. When necessary, *B. subtilis* was propagated in Spizizen Salts (Spizizen, 1958), supplemented with 0.5% of glycerol, 0.01% tryptophan, 0.01% phenylalanine and trace elements (Harwood and Cutting, 1990). This medium was designated SMM. Isoleucine (0.01%) was added when indicated.

Antibiotics were used at the following final concentrations: lincomycin (Lm) 12.5 μg mL^{-1}; erythromycin (Erm) 0.5 μg mL^{-1}; chloramphenicol (Cm) 5 μg mL^{-1}; kanamycin (Km) 5 μg mL^{-1}; cerulenin (Cer) 2 μg mL^{-1}; ampicillin (Amp) 100 μg mL^{-1}.

For the experiments involving expression of the HKs under the control of the inducible P*xylA* promoter, xylose was added to the media as indicated in each case.

Plasmids and Strain Construction

The genes encoding the different HKs and RRs were PCR-amplified using chromosomal DNA from *B. anthracis* 34F2, *S. aureus* RN4220 or *B. subtilis* JH642, and the corresponding oligonucleotides indicated in **Table 2**. In all cases Q5® high-fidelity DNA polymerase (New England Biolabs) was used and the identity and correct sequence of the cloned fragments were corroborated by sequencing.

BA_RS26385 was first cloned into the *Cla*I and *Sal*I sites of pGES40, downstream of *xylR*-P*xylA* (Albanesi et al., 2004). The resulting plasmid was then digested with *Bam*HI and *Sal*I to release the fragment containing *xylR-PxylA-BA_RS26385*

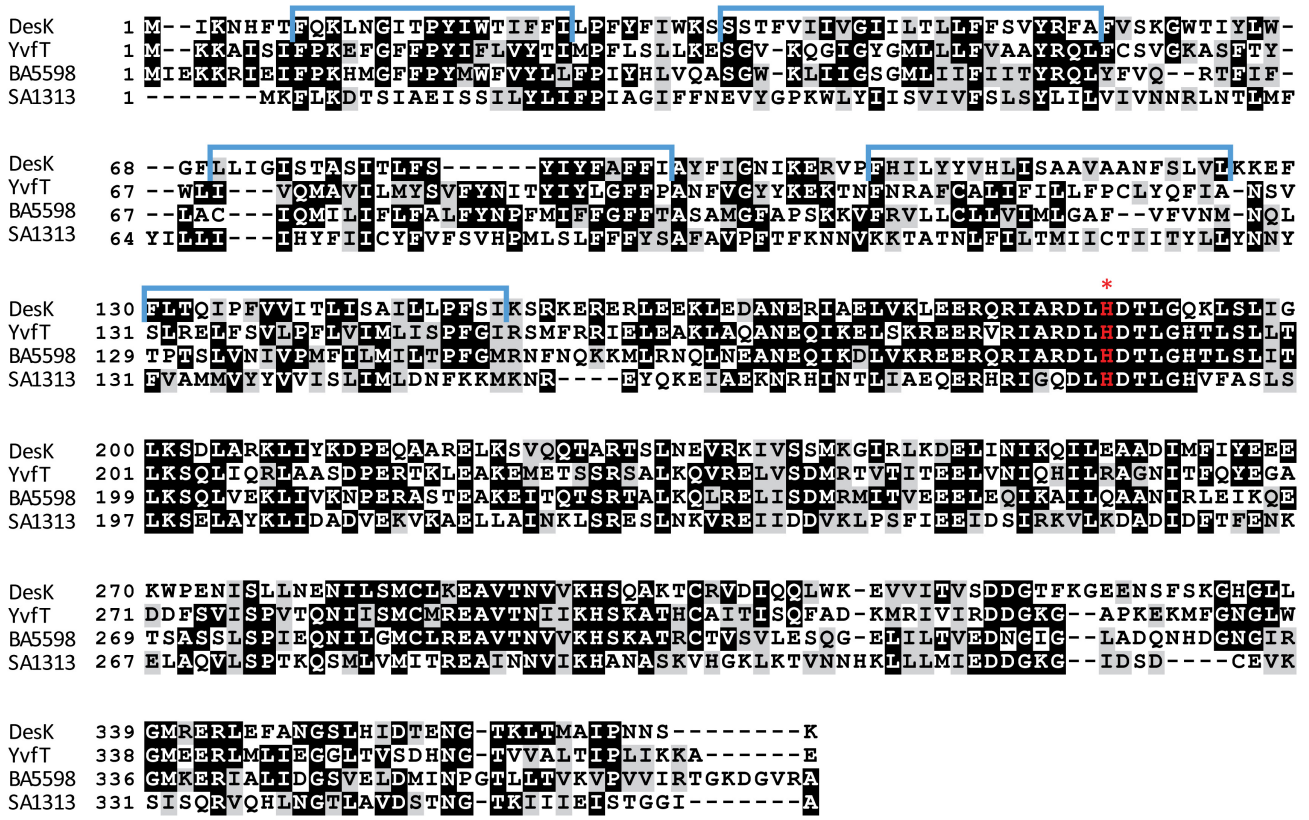


FIGURE 1 | Amino acid sequence alignment of HKs, DesK, YvfT, BA5598, and SA1313. Identical amino acids are shown highlighted in black and similar ones are highlighted in gray. TMS are marked above the DesK sequence in blue (Cybulski et al., 2010). The red asterisk indicates the phosphorylatable His.

which was cloned into pHPKS digested with the same enzymes, obtaining plasmid pH5598. In this work we refer to the product of *BA_RS26385* gene, as BA5598 (previous gene name).

SAOUHSC_01313, was cloned into the *Bam*HI and *Pst*I of pARD7 (**Table 1**). The resulting plasmid was called pPxyl-HKsa. In this work we refer to the product of *SAOUHSC_01313* as SA1313.

yvfT-yvfU TCS was first cloned into pCR-bluntII-TOPO, then digested with *Sal*I and the resulting 5′-overhang filled with DNA polymerase I, Klenow fragment (Thermo Fisher Scientific). This construction was digested with *Spe*I, a restriction site contained in the vector, and the resulting fragment containing *yvfT-yvfU* was cloned into the *Sma*I and *Spe*I sites of pLarC1 (**Table 1**), yielding pLUP161. *yvfT* was amplified and cloned into pCR-bluntII-TOPO. The resulting plasmid was digested with *Spe*I and *Xba*I enzymes, and the 1.1 Kb fragment was cloned into pLarC1 digested with the same enzymes, yelding pPF23.

For overexpression in *E. coli*, the ORFs coding for the cytoplasmic domains of BA5598 (BA5598C), SA1313 (SA1313C), and YvfT (YvfTC), were amplified using the pair of oligonucleotides indicated in **Table 2**. The PCR fragments were purified and cloned into the *Bam*HI and *Pst*I sites of pQE30 or pQE32, yielding pQE5598c, pPF172, and pPF46 (**Table 2**).

To express the *B. anthracis* RR and YvfU as a maltose binding proteins (MBP) fusion protein to the N-terminus, the *BA_RS26380* gene, encoding BA5597 and *yvfU* were amplified and cloned into the *Bam*HI site of pMAL-c2X expression vector, to obtain pM5597 and pPF28, respectively.

To obtain the *S. aureus* RR with a His tag, the *SAOUHSC_01314*, encoding the RR SA1314, was cloned in *Bam*HI and *Pst*I sites of pQE30, yielding pPF177.

In all cases, amplicons were first cloned into pJet1.2 blunt (Invitrogen) or pCR-bluntII-TOPO, and sequenced.

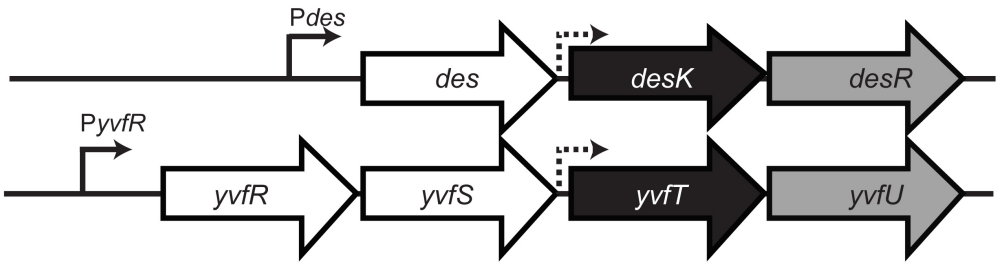
P*des-lacZ* Induction Analysis

B. subtilis DAK3, CM21, and CM28 strains (**Table 1**), all harboring P*des-lacZ* fusions, were transformed with the plasmids expressing the different HKs under the inducible *PxylA* promoter and, in parallel, with pARD7 (pHPKS-*PxylA*) to be used as a control. Strains were streaked on LB agar plates supplemented with 0.05% xylose and 60 μg mL^{-1} X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) by duplicate. Plates were incubated for 5 h at 37°C and then one plate was transferred to 25°C overnight.

Protein Overexpression and Purification

YvfTC, SA1313C, BA5598C, DesKC, and the RR SA1314 were expressed in *E. coli* M15/pREP4. Cells expressing the His-tagged recombinant proteins were grown in LB at 37°C with agitation to DO_{600} = 0.5. Then the culture was shifted to 20°C, and

A



B

Species	ABC-NBD	ABC-TMD	HK	RR
B. anthracis	***BA_RS26395***	***BA_RS26390***	***BA_RS26385***	***BA_RS26380***
B. cereus	*EJ379_27360*	*EJ379_27355*	*EJ379_27350*	*EJ379_53745*
S. aureus	***SAOUHSC_01311***	***SAOUHSC_01312***	***SAOUHSC_01313***	***SAOUHSC_01314***
S haemolyticus	*SH1589*	*SH1588*	*SH1587*	*SH1586*
S. epidermidis	*B4U56_RS07360*	*B4U56_RS07355*	*B4U56_RS07350*	*B4U56_RS07345*
S. pneumoniae	*SPR_RS09125*	*SPR_RS09120*	*SPR_RS09115*	*SPR_RS09110*
S. agalactiae	*GBSCOH1_RS01795*	*GBSCOH1_RS01800*	*GBSCOH1_RS01805*	*GBSCOH1_RS01810*

C



FIGURE 2 | (A) Chromosomal organization of the *B. subtilis* homologous TCS regions. Sensor HK genes are shaded in black, RR in gray and upstream genes in white. *des*: desaturase; *yvfR* ABC transporter nucleotide binding domain (ABC-NBD), *yvfS* ABC transporter TM domain (ABC-TMD). Continuous arrows represent promoters and dashed arrows putative promoters. **(B)** Conserved clusters among Gram positive pathogens. The gene names of different species were extracted from GeneBank. *B. anthracis* str. Ames, *B. cereus* ATCC 14579, *Staphylococcus aureus* RN4220, *S. haemolyticus* JCSC1435, *S. epidermidis* 1457, *Streptococcus pneumonia* R6, *S. agalactiae* COH1. **(C)** Alignment of promoter regions. Direct and inverted repeats involved in RR binding in P*des* promoter (Najle et al., 2009) are marked in boxes. Sequence differences respect to P*des* are shown in bold.

protein expression was induced by adding 0.3 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG). After 16 h, cells were harvested, resuspended in lysis buffer [50 mM Tris–HCl pH 8, 500 mM NaCl and 1 mM phenylmethylsulfonyl fluoride (PMSF)] and treated with 1 mg mL^{-1} lysozyme for 30 min at 37°C. Subsequently, 5 μg ml^{-1} DNase plus 5 mM $MgCl_2$ were added and after 20 min of incubation at room temperature cells were disrupted with an Emulsiflex^{TM} C3 homogenizer (AVESTIN). Cell debris were separated by centrifugation at 12000 × *g* for 20 min. The supernatant was recovered and incubated with pre-equilibrated Ni^{+}-Nitrilotriacetic acid agarose resin (Qiagen), in the presence of 10 mM imidazole, for 3 h. The His_6-tagged proteins were eluted with increasing concentrations of imidazole (**Supplementary Figure S2**) and dialyzed against 50 mM Tris pH 8, 300 mM NaCl and 10% glycerol. Proteins were conserved at −80°C for further experiments. Coomassie-stained SDS–PAGE of purification steps or final stored protein fractions are shown in **Supplementary Figure S2**.

Maltose binding protein-YvfU and MBP-BA5597 were overexpressed in *E. coli* XL1Blue. Cells expressing the MBP-tagged recombinant proteins were grown in LB at 37°C to a DO_{600} = 0.5 and then shifted to 30°C. Protein expression was induced by adding 0.3 mM IPTG. After 4 h, cells were harvested, resuspended in lysis buffer (20 mM Tris–HCl

TABLE 1 | Strains and plasmids used in this study.

Strain or plasmid	Relevant characteristics	Source
Strains		
RN4220	*S. aureus*	Laboratory stock
34F2	*B. anthracis* Sterne (pXO1$^+$ pXO2$^-$)	Laboratory stock
JH642	*B. subtilis trpC2 phe*A1	Laboratory stock
DAK3	JH642 *desKR::Kn* P*xylA-desR amyE::*P*des-lacZ*	Saita et al., 2015
CM21	JH642 *desKR::Kn thrC::*P*xylA-desR amyE::*P*des-lacZ*	Albanesi et al., 2004
CM28	JH642 *desKR::Kn amyE::*P*des-lacZ*	This work
Plasmids		
pHPKS	*B. subtilis* replicative vector of low copy number	Johansson and Hederstedt, 1999
pGES40	P*xylA* cloned into pBluescript SK(+) (Stratagene)	Diez et al., 2012
pARD7	pHPKS derivative with P*xylA* cloned in *Sac*I	Díaz et al., 2019
pLarC1	pHPKS derivative with P*xylA* cloned in *Hin*dIII and *Eco*RI sites	Trajtenberg et al., 2014
pLUP124	pHPKS-*PxylA-desK*	Porrini et al., 2014
pPF23	pHPKS *PxylA-yvfT*	This work
pLUP161	pHPKS *PxylA-yvfTU*	This work
pH5598	pHPKS *PxylA-BA_RS26385*	This work
pPxyl-HKsa	pHPKS-*PxylA-SAOUHSC_01313*	This work
pQE30-32	*E. coli* expression vector (fusion to His tag)	QIAGEN
pMAL-c2X	*E. coli* expression vector (fusion to MBP tag)	New England Biolabs
pDA327	pQE32-*desKC*	Albanesi et al., 2004
pPF46	pQE30-*yvfTC*	This work
pQE5598c	pQE32-*BA_RS26385C*	This work
pPF172	pQE30-*SAOUHSC_01313C*	This work
pPF177	pQE30-*SAOUHSC_01314*	This work
pPF28	pMAL-c2X-*yvfU*	This work
pM5597	pMAL-c2X-*BA_RS26380*	This work
pCMdesR	*desR* cloned into pETGEXCT	Albanesi et al., 2004

TABLE 2 | Oligonucleotides used in this study.

Target gene	Name	Sequence
yvfT	yvfTBamUp	AGGATCCAAGAAGCGGTGTAATACATGA
	IBR	AATCGGTACCTGGATCCTCCAATCATTCTGC
yvfTC	yvfTCBamUp	GGATCCCGGAGGATTGAGCTTGAAG
	IBR	AATCGGTACCTGGATCCTCCAATCATTCTGC
yvfU	yvfUBamUp	GGATCCATTCGTCTGTTTATTGCTGAG
	yvfUDw	GGTCTGCAGTATATGGTTAGATCCAGC
yvfT-yvfU	yvfT-RBS	CCGTCGACAAGAAGCGGTGTAATACATG
	yvfUDw	GGTCTGCAGTATATGGTTAGATCCAGC
SAOUHSC_01313	1313Up	TGGATCCGGGCGGAATAAAATATG
	1313Dw	GGCTGCAGTAATTAAAGATGTCATGC
SAOUHSC_01313C	1313CUp	GGATCCGAATATCAAAAAGAAATAGC
	1313Dw	GGCTGCAGTAATTAAAGATGTCATGC
SAOUHSC_01314	SA1314ATG UP	ATGGATCCACATCTTTAATTATTGCAG
	SA59 DW	TCTGCAGTTTGTATTTAGATCCAGCC
BA_RS26385	Ba2552	AATCGATAGGGAGGCGTATAAATGATAGA GAAGAAG
	BAC3	AGGTCGACTTCGAATCATGCCCTCACC
BA_RS26385C	BAC5	CGGGATCCGTATGCGCAACTTTAATCAG
	BAC3	AGGTCGACTTCGAATCATGCCCTCACC
BA_RS26380	5597Up	TCGGGATCCGAAAACCTGTATTTTCAGGA TGATTCGAATTATTATTGC
	5597BxDw	AGGGGATCCTCTAGATCATATCCATCCCT TTTCTTC

Restriction sites are underlined.

pH 7.4, 200 mM NaCl and 1 mM PMSF) and treated with 1 mg mL^{-1} lysozyme for 30 min at 37°C. Subsequently, 5 μg ml^{-1} DNase plus 5 mM $MgCl_2$ were added. After 20 min of incubation at 37°C cells were disrupted with an Emulsiflex^{TM} C3 homogenizer (AVESTIN). Cell debris were separated by centrifugation at 12000 × *g* for 20 min. The supernatant was recovered and incubated with pre-equilibrated Amylose resin (New England Biolabs) for 1 h. The MBP fusion proteins were eluted with 10 mM maltose and dialyzed against 50 mM Tris pH 8, 300 mM NaCl and 10% glycerol. Proteins were conserved at −80°C for further experiments. Coomassie-stained SDS–PAGE of purification steps or final stored protein fractions are shown in **Supplementary Figure S2**.

GST-DesR was purified as previously described (Albanesi et al., 2004). Briefly, BL21/λDE3 strain transformed with pCMdesR was grown for 1 h at 37°C in LB and then 0.1 mM IPTG was added. Growth was continued for 4 h, cells harvested at 4°C and washed with 1× phosphate saline buffer, PBS (0.14 M NaCl, 0.27 mM KCl, 10.1 mM Na_2HPO_4, 1.8 mM KH_2PO_4). Cells were disrupted by sonication, centrifuged and the supernatant incubated with 50% glutathione-agarose resin (SIGMA), pre-equilibrated with 1 × PBS. GST-DesR was eluted with 5 mM Glutathione.

In vitro Phosphorylation Assays

For the autokinase assay, purified YvfTC-Hisx6, BA5598C-Hisx6, SA1313C-Hisx6 at a final concentration of 10 μM were incubated in R Buffer (50 mM Tris pH 8, 200 mM NaCl, 1 mM DTT, 20% glycerol, 50 mM KCl, 1 mM $MgCl_2$) containing 25 μM ATP in the presence of 0.25 μCi of [γ-^{32}P]ATP/μl at room temperature. Reactions were initiated by addition of the corresponding HK.

For the phosphotransfer assays, each HK (10 μM) was allowed to autophosphorylate for 30 min in 150 μl of R buffer at room temperature. An aliquot of 15 μl was withdrawn (considered as time zero) and then the phosphotransfer assay was initiated by adding a volume of the corresponding RR to reach an equimolar final concentration with the HK. In all cases, aliquots were withdrawn at various time points, mixed with 5× sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) loading buffer plus 50 mM EDTA to stop the reaction and kept on ice until loading.

Bioinformatics

Protein sequences were analyzed with the program BLASTP (Altschul et al., 1990). Sequence alignments were performed using T-Coffee (Notredame et al., 2000) and drawn using

Boxshade[1]. Molecular modeling was carried out with Swiss-Model[2] (Waterhouse et al., 2018) and analyzed with PyMOL (PyMOL Molecular Graphic system, 2.0, Schrodinger, LLC, available at https://pymol.org/).

RESULTS

In vitro Characterization of the DesK/DesR TCS Homologs

DesK is a multifunctional enzyme with autokinase, DesR phosphotransfer and phospho-DesR phosphatase activities that can assume different signaling states depending on the input signal. In the absence of the TM sensor domain, its cytosolic domain (DesKC) exhibits the three activities *in vitro* at 25°C (Albanesi et al., 2004). According to the two-component signal transduction paradigm, the HKs would first undergo autophosphorylation in a conserved histidine residue and then the phosphate would be transferred to the conserved aspartate residue in the receiver (REC) domain of its cognate transcriptional regulator. His188 has been identified unequivocally as the autophosphorylation site in DesK. A His188Val mutant was unable to receive the phosphate *in vitro* and to induce *des* transcription after a cold shock *in vivo* (Albanesi et al., 2004). His188 is located in a conserved domain called H-box, included in the dimerization and histidine phosphotransfer (DHp) domain, that defines the active site of phosphorylation (Trajtenberg et al., 2010). The alignment in **Figure 1** clearly shows that the HKs under study contain this H-box, conserved in the HisKA_3 subfamily.

To analyze the catalytic activities of YvfT, BA5598 and SA1313 HKs, we expressed and purified the corresponding cytoplasmic soluble domains, YvfTC, BA5598C, and SA1313C, from *E. coli* (see section "Materials and Methods"). In all cases, the recombinant proteins included the complete DHp domain, containing the phospho-receiving His (His189, His187, and His185, respectively) and the ATP-binding domains (ABDs), since autophosphorylation involve specific contacts between these interacting domains (Trajtenberg et al., 2010). As shown in **Figure 3A**, YvfTC, BA5598C, and SA1313C are able to autophosphorylate in the presence of [γ-^{32}P] ATP. YvfTC-P can be detected after 1 min of reaction, while detection of SA1313C-P and BA5598C-P requires longer incubation times. In all cases, the level of phosphorylation increased with time, reaching a plateau at 30 min for YvfTC and SA1313C (**Supplementary Figure S3A**).

To study the activation by phosphorylation of the RRs by their cognate kinases we produced the recombinant proteins in *E. coli*. As YvfU and BA5597 have similar molecular weight to the cytosolic domains of their corresponding HKs, YvfTC, and BA5598C, respectively, these RRs were fused to the MBP to be able to distinguish both partners in SDS-PAGE runs. As previously described, there is no detrimental effect of GST or MBP tags on RR recognition and phosphorylation (Albanesi et al., 2004; Tsai, unpublished). Besides, substrate specificity in HK-RR pairs is really high. Cognate pairs can identify each other in bacterial cells, where up to 300 HK-RR pairs are present. Such specificity is determined by a "specificity code" present in the contact surfaces: DHp in HKs and REC domain in RR (Skerker et al., 2008; Capra et al., 2012; Buschiazzo and Trajtenberg, 2019). The interacting surface between DesK and DesR in the phosphotransfer and dephosphorylation reactions has been described, and concerning DesR mainly involves its α1-helix and the β5α5 loop, with a few more contacts from the N-terminus of α5 and the beginning of the β4α4 loop. All these regions are conserved in the RRs under study in this manuscript (**Supplementary Figure S1**).

YvfTC, BA5598C, and SA1313C were allowed to autophosphorylate for 30 min in R buffer and then an equimolar amount of the cognate RR was added, promoting their phosphorylation (**Figure 3B**). For the YvfTC/MBP-YvfU pair, the level of RR-P was maximum after 1 min of reaction and then the radioactive label from both proteins decreased. After 15 min of incubation, MBP-YvfU-P could not be longer detected, while YvfTC-P could still be observed after 30 min of reaction (**Figure 3B** and **Supplementary Figure S3B**). Phosphotransfer from SA1313C-P to SA1314 was fast and very efficient. SA1313C-P could not be detected after 30 s of reaction while the level of SA1314-P reached a maximum within this incubation interval. Then, the level of the phosphorylated regulator decreased and was negligible after 15 min of reaction (**Figure 3B** and **Supplementary Figure S3B**). Finally, the phosphotransfer from BA5598C-P to MBP-BA5597 was slower and a longer time was necessary to observe a full transfer of the radioactive label. The level of MBP-BA5597-P increased during 60 min of incubation and no decay of the label was detected in these assay conditions (**Figure 3B** and **Supplementary Figure S3B**). To verify that the decrease of the level of YvfTC-P and SA1313C-P was due to specific phosphotransfer to MPB-YvfU and SA1314, respectively, the HKs were allowed to autophosphorylate for 30 min in R buffer and then incubated in the absence of their cognate RR, in the same reaction conditions. As shown in **Supplementary Figure S5**, YvfTC and SA1313C remained phosphorylated even after 45 min of incubation, confirming that the phosphorylated HKs are stable and do not lose the radioactive label spontaneously.

Cross-Phosphorylation Between the Different HKs and DesR *in vitro*

Due to the high sequence similarity of the RRs (**Supplementary Figure S1**), we decided to investigate if YvfTC, SA1313C, and BA5598C are capable of recognizing and phosphorylating the non-cognate RR DesR. To this end, YvfTC, SA1313C, and BA5598C were allowed to autophosphorylate for 30 min and then an equimolar amount of purified GST-DesR was added. This tagged DesR version has been used successfully to test DesKC phosphotransfer and phosphatase activities (Albanesi et al., 2004). As shown in **Figure 4**, the phosphotransfer to GST-DesR was detected in all cases. GST-DesR was quickly phosphorylated upon incubation with YvfTC-P, the label on the RR could be detected after 30 s reaching a maximum after 5 min

[1]https://embnet.vital-it.ch/software/BOX_form.html

[2]https://swissmodel.expasy.org/

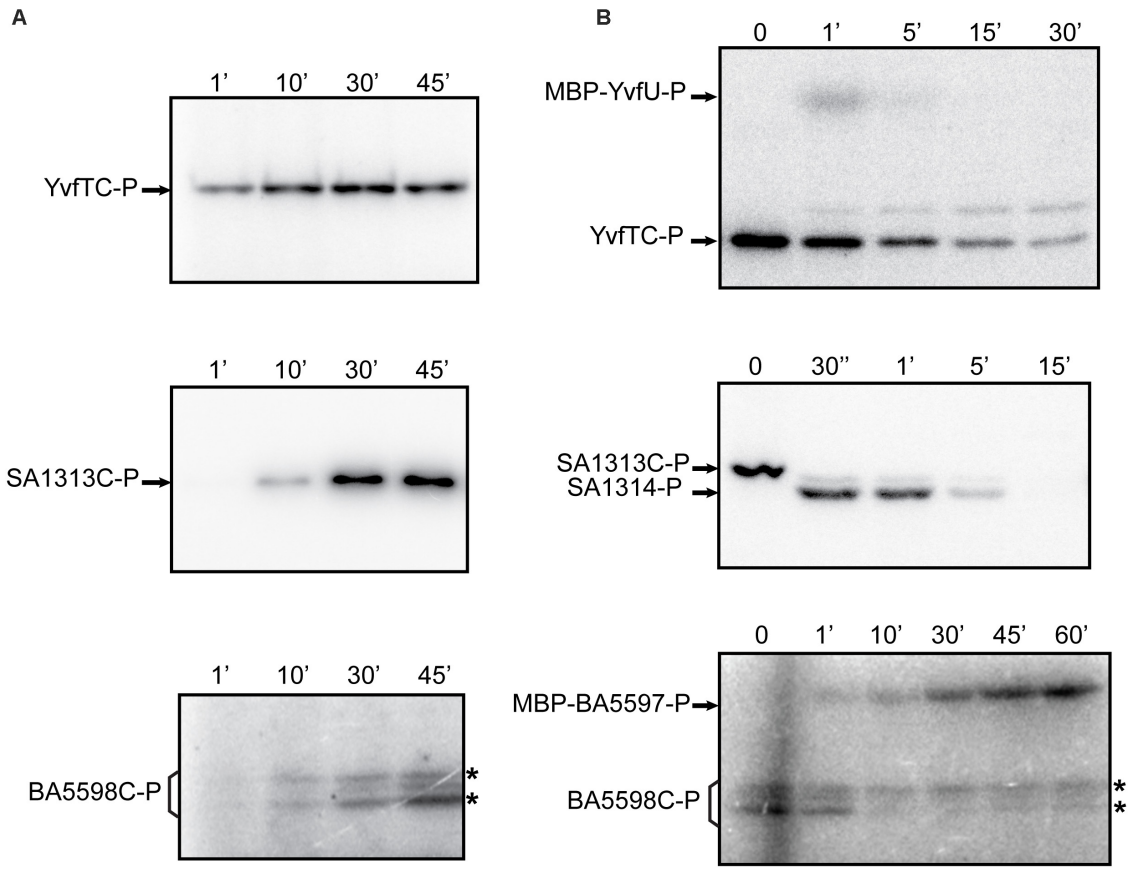


FIGURE 3 | (A) Autophosphorylation assay of YvfTC, BA5598C, and SA1313C. The purified cytosolic domains were incubated with R buffer at room temperature. Reaction was initiated by the addition of [$\gamma - {}^{32}$P] ATP. In the case of BA5598C, western blot analysis with anti-DesKC polyclonal antibodies showed that both bands marked with * correspond to the protein, while anti-His antibodies indicated that only the upper one possesses the 6xHis-tag (data not shown) **(B)** HKs phosphotransfer assay to cognate RRs. After 30 min of autophosphorylation at room temperature, each HK was mixed with an equimolar amount of its each cognate RR. After SDS-PAGE, gels were dried, exposed to a radioactive storage screen and analyzed using an Amersham Biosciences Molecular Dynamics Typhoon^{TM} FLA 7000 scanner. Pictures are representative of at least three independent experiments.

of reaction. Afterward, the level of GST-DesR-P and YvfTC-P decreased, being very low after 15 min of incubation (**Figure 4** and **Supplementary Figure S4**). SA1313C-P and BA5598C-P were less efficient to catalyze the phosphotransfer to DesR and significant levels of GST-DesR-P could be detected only after 10 min of reaction. In both cases, the amount of GST-DesR-P remained almost constant after reaching the maximum, showing that dephosphorylation of the RR by the non-cognate HK was rather inefficient (**Figure 4** and **Supplementary Figure S4**).

In vivo Analysis of HKs Temperature Sensing Capacity

The TM domain of DesK is essential for signal detection and modulation of its catalytic activities (Albanesi et al., 2004; Martín et al., 2009; Cybulski et al., 2010). Although the TM helices of DesK, YvfT, SA1313 and BA5598 have low amino acid sequence similarity, they share the same architecture (**Figure 1**). Thus, we decided to investigate whether YvfT, SA1313 and BA5598 are also able to detect temperature changes. Taking advantage that these HKs recognize and phosphorylate DesR *in vitro* (**Figure 3**), we analyzed their functional properties in a *B. subtilis* host. With this aim, we engineered a DesK-less host containing the β-galactosidase reporter gene fused to the desaturase promoter (P*des-lacZ*). In this strain, (CM21, **Table 1**), the levels of β-galactosidase activities rely on activation of P*des*, that ultimately depends on the flux of phosphate from the HK to DesR (Albanesi et al., 2004). The induction of transcription of the reporter gene by the HKs expressed in this strain can be easily monitored on solid media containing X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside), where the colonies turn blue if the HK is in the kinase-ON state and remain clear if the HK is in a kinase-OFF state. As shown in **Figure 5**, expression of BA5598 and SA1313 from a xylose-inducible promoter (P*xyl*) in CM21 rendered clear colonies at 37°C and blue colonies at 25°C. These results show that, similarly to DesK, both HKs specifically adopt a kinase-ON state at low temperature and phosphorylate DesR, stimulating transcription from P*des*. *In vivo* phosphorylation of DesR by the non-cognate HKs agrees with the *in vitro* experiments. Nevertheless, expression of YvfT in CM21

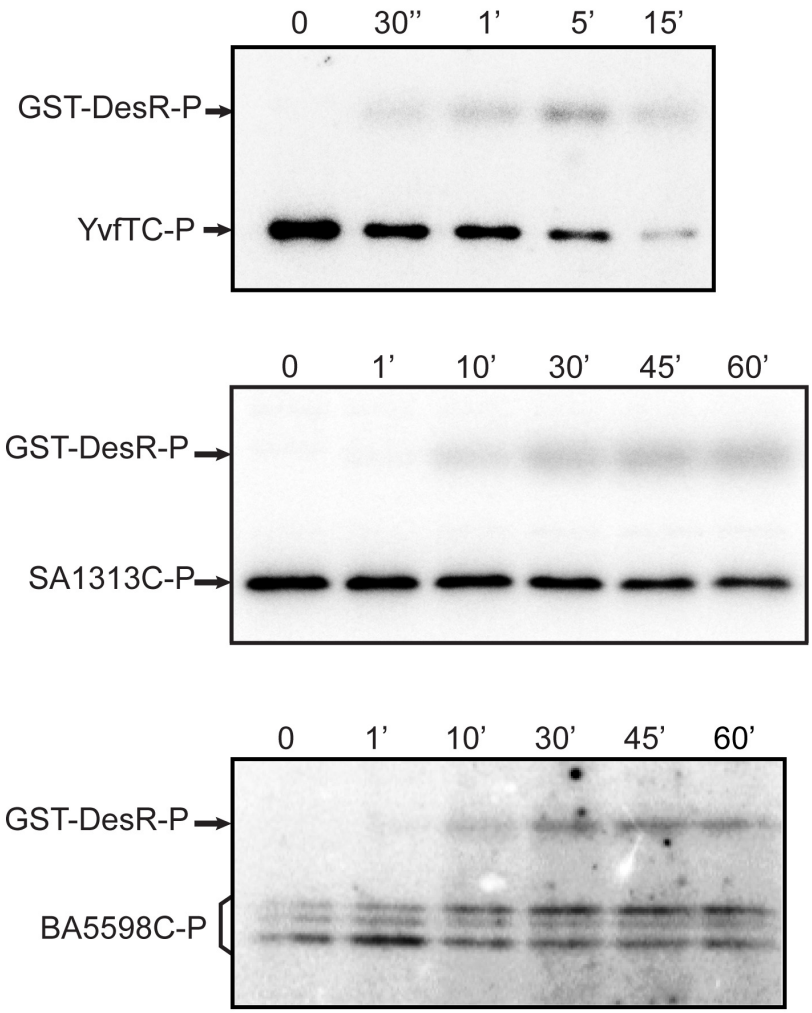


FIGURE 4 | HKs phosphotransfer assay to DesR. After 30 min of autophosphorylation at room temperature, HKs were mixed with an equimolar amount of DesR. In all cases aliquots were taken at the indicated time points. Reactions were stopped by the addition of 5× sample buffer and loaded in 12% polyacrylamide gels. After SDS-PAGE, gels were dried, exposed to a radioactive storage screen and analyzed using an Amersham Biosciences Molecular Dynamics Typhoon^{TM} FLA 7000 scanner. Pictures are representative of at least three independent experiments.

produced clear colonies both, at 37 and 25°C (**Figure 5**). This result suggests that this HK does not perceive low temperature stimulus, or alternatively, a mechanism to avoid cross talk between YvfT/YvfU and DesK/DesR systems could explain this observation. To test this possibility, we decided to co-overexpress YvfT together with YvfU from P*xyl* in either CM21 or in a *B. subtilis desKR* knock-out mutant strain harboring the P*des-lacZ* reporter fusion (CM28, **Table 1**). In both cases, we observed blue colonies at 25°C and clear colonies at 37°C, indicating that YvfT is able to adopt a kinase-ON state after a temperature downshift phosphorylating YvfU to activate expression of P*des-lacZ* (**Figure 5**). It should be noted that expression from P*des* is stimulated by YvfT/YvfU only when this TCS is overexpressed (**Figure 5**). Indeed, P*des-lacZ* induction was not detected in a *desK*$^{-}$ *or desKR*$^{-}$ strains that maintain their wild type single chromosomal copy of the *yvfTU* operon (Aguilar et al., 2001; Saita et al., 2015). These results directly prove that both signaling pathways are not redundantly regulating *des* transcription under physiological expression conditions.

Isothermal Perception of Changes in the Physical State of the Membrane

Activation of SA1313, BA5598, and YvfT upon cooling suggests that these HKs could be sensing changes in membrane properties through their TM helices. Based in previous experiments with DesK (Cybulski et al., 2002; Porrini et al., 2014), we analyzed whether fluctuations in membrane fatty acid (FA) composition at a constant temperature affect the sensory function of these homologous HKs.

Changes in the Proportion of Iso- and Anteiso-Branched Chain FAs

B. subtilis adjust membrane fluidity at low growth temperatures by increasing the proportion of low melting point FAs, like



FIGURE 5 | Complementation test of *desK* or *desK/desR* null mutants. CM21 strain (*desK:Kn thrC*:P*xylA-desR* P*des-lacZ*, white background) was transformed with plasmids expressing DesK, BA5598, SA1313, YvfT, or YvfT/YvfU under a xylose inducible promoter. CM28 (*desKR:Kn* P*des-lacZ*, shaded background) was transformed with a plasmid expressing YvfT/YvfU under a xylose inducible promoter. All transformants were grown in LB plates containing X-gal 60 μg mL^{-1} and xylose 0.05%. One plate was incubated at 37°C and the other was kept at 37°C for 5 h and then transferred to 25°C for 24 h.

unsaturated (UFAs) and branched chain fatty acids (BCFAs). The biosynthetic precursors of anteiso (ai)-BCFA are α-ketoacids derived from isoleucine (Ile), which is synthetized from threonine (Thr). If the culture media lacks Ile or Thr, JH642 strain cannot synthesize enough ai-BCFAs and the membrane becomes less fluid (Klein et al., 1999). Under such growth conditions, Δ5-Des expression is induced and UFAs are synthesized, even at 37°C, leading to membrane fluidity optimization. This induction is mediated by the DesK/DesR TCS demonstrating that DesK detects changes in the physical state of the cytoplasmic membrane rather than fluctuation in temperature to control *des* expression (Cybulski et al., 2002). Hence, we decided to test if the HKs homologous to DesK are also capable of perceiving isothermal membrane fluidity changes. To this end, we used a Thr prototroph *desK* knockout mutant strain, harboring the P*des-lacZ* reporter fusion, named DAK3 (**Table 1**). This strain was transformed with plasmids expressing the genes encoding the HKs SA1313 and BA5598 or the YvfT/YvfU TCS from a xylose inducible promoter. Transformants were grown in SMM agar plates containing X-gal and 0.05% xylose, in the presence or in the absence of isoleucine, at 37°C. As shown in **Figure 6A**, BA5598 was able to promote *lacZ* expression from P*des*, indicating that this *B. anthracis* HK is sensitive to the isothermal fluidity change. However, neither YvfT nor SA1313 were activated by isoleucine depletion (**Figure 6A**).

Effect of FA Chain Shortening by the Action of Cerulenin

We have proposed that membrane thickening is the main driving force for signal sensing by DesK (Cybulski et al., 2010; Martín and de Mendoza, 2013; Saita et al., 2015). A thicker membrane favors DesK kinase activity while its phosphatase state is promoted when membrane thickness is reduced. This regulation can be detected at a constant temperature *in vivo* using the antibiotic cerulenin. This compound is an inhibitor of the elongation of FAs and, in sub-inhibitory concentrations, favors the production and incorporation of short FAs into *B. subtilis* membrane phospholipids (Porrini et al., 2014). Lipids with shorter acyl chains generate thinner and less viscous membranes (de Mendoza, 2014). Thus, after cerulenin treatment of *B. subtilis* cells DesK adopts a kinase-OFF/phosphatase-ON state, even at low temperature, so transcription from P*des* is not induced at 25°C and hence, UFAs are not synthetized (Porrini et al., 2014). Therefore, we analyzed *in vivo* the response to low temperature of SA1313, BA5598 and YvfT in the presence of cerulenin. To this end, the reporter strain DAK3 expressing SA1313, BA5598 or YvfT/YvfU was grown in LB agar plates containing X-gal and xylose, supplemented or not with 2 μg/mL of cerulenin, at 25°C. As shown in **Figure 6B**, under these conditions, SA1313 does not promote P*des-lacZ* expression, indicating that, in the presence of cerulenin, it adopts a kinase-OFF conformation at low temperature. This result suggests that SA1313 senses the narrowing of the membrane caused by incorporation of shorter acyl chains. In contrast, BA5598 and YvfT exhibit a kinase-ON state, phosphorylating DesR or YvfU, respectively and inducing P*des-lacZ* transcription at 25°C, in spite of the presence of cerulenin. These results indicate that the decrease in membrane thickness that results from the treatment with cerulenin was insufficient to trigger the switch of these two HKs to the kinase-OFF state.

Altogether, these results demonstrate that, even though all the studied HKs have similar primary structure and respond to temperature changes, they differ in their sensitivity to changes in membrane physical properties. At a constant temperature, BA5598 can perceive fluidity changes promoted by a modification in the proportion of ai-BCFA while SA1313 can sense fluidity changes provoked by variation in the length of acyl-chains of membrane phospholipids. On the other hand, YvfT is insensitive to both stimuli.

DISCUSSION

Temperature is an important parameter that free-living organisms monitor constantly. In the case of bacterial pathogens, the expression of cold shock, heat shock and some virulence

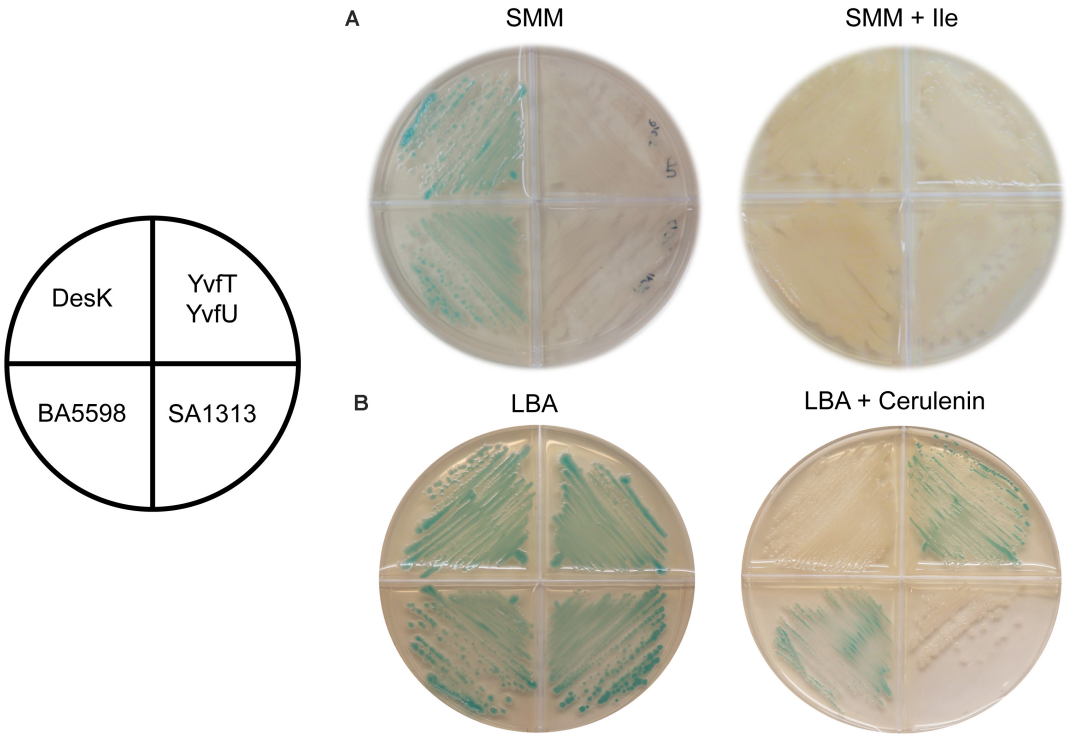


FIGURE 6 | Response of homologous HKs to changes in membrane fatty acid composition. DAK3 strain (*desK:Kn* P*xylA-desR* P*des-lacZ*) transformed with plasmids expressing DesK, BA5598, SA1313, or YvfT/YvfU under a xylose inducible promoter were grown in SMM or LB plates containing X-gal 60 μg ml^{-1}, xylose 0.05%, and isoleucine 0.01% or cerulenin 2 μg ml^{-1}, as indicated. Plates were incubated at 37°C for 16 h **(A)**, or incubated at 37°C for 5 h and then transferred to 25°C for 24 h **(B)**.

genes is coordinated in response to temperature changes. This behavior highlights temperature sensing as a key event for a fine-tuned orchestration of bacterial virulence and pathogenesis. TCSs are one of the mechanisms that bacteria use to sense and respond to variations in temperature, so understanding the stimulus perception of HKs and their transmission to the RRs would be a way to control their growth and spread.

In this work, we studied three HKs homologous to DesK: BA5598 from *B. anthracis*, SA1313 from *S. aureus* and YvfT from *B. subtilis*. Using *in vivo* assays, we demonstrated that these HKs are able to perceive a downshift in temperature, shifting its activity toward the kinase-ON state to promote transcription from P*des*, when expressed in *B. subtilis*. We found that BA5598 and SA1313 could recognize and phosphorylate the non-cognate RR DesR both *in vitro* and *in vivo*, while YvfT could only phosphorylate and activate DesR *in vitro*. We also determined that the co-overexpression of YvfT and YvfU in *B. subtilis* led to the induction of P*des-lacZ* expression upon a cold shock, without the assistance of DesR. This finding demonstrates that YvfU is capable of recognizing the operator sequences of the *des* promoter to activate transcription when temperature decreases. Under our experimental conditions, time course experiments demonstrated a similar phosphotransfer rate between YvfT-P and both MBP-YvfU or GST-DesR (**Figures 3B**, **4**). However, in physiological conditions, YvfT or YvfT/YvfU do not cross-talk with the Des pathway since in *desK* or *desK/desR B. subtilis* null mutants the *des* gene is not turned on upon a cold shock (Aguilar et al., 2001; Saita et al., 2015). Interestingly, the molecular model of the YvfT:DesR complex shows that most of the molecular recognition determinants involved in cognate pair DesK:DesR recognition are present (**Supplementary Figure S6**). These results highlight that the cellular context is essential to avoid the cross talk between the YvfT/YvfU and DesK/DesR systems.

Furthermore, we determined that all the studied sensor proteins are activated by a temperature decrease, although they behave differently toward changes in the cytoplasmic membrane properties provoked under isothermal conditions. Changes in membrane fluidity caused by increasing the proportion of shorter FAs into phospholipids prompt DesK and SA1313 to adopt a phosphatase-ON state even at 25°C. However, this stimulus is not enough to turn off the kinase state of BA5598 and YvfT at this temperature. The motif conformed by three hydrophilic amino acids (Gln9, Lys10, and Asn12) located near the DesK N-terminal end of its first TM helix, called the sunken buoy, is critical to sense membrane thickness (Cybulski et al., 2010). At low temperatures, the membrane becomes thicker due to an increase in the lipid order, thereby trapping this hydrophilic motif inside the hydrophobic membrane environment, promoting the stretching of TM1 with the concomitant generation of the signal that finally promotes the kinase activity of DesK (Cybulski et al., 2010; Saita et al., 2015). In BA5598 and YvfT there is a Pro residue

instead Gln9 in TM1, which could introduce a kink into this amphipathic domain changing the sensitivity of these HKs to membrane thickness variations.

Apart from temperature, BA5598 also responds to a decrease in membrane fluidity caused by the reduction of ai-BCFA in membrane phospholipids, adopting a kinase-ON state at 37°C. However, YvfT and SA1313 were unable to respond to this stimulus. Previous studies of P*des-lacZ* transcription at different temperatures, in the presence or absence of ai-BCFA precursors, revealed that each stimuli triggers a variation in membrane fluidity of a particular extent (Cybulski et al., 2002). Our results suggest that this set of HKs might have particular thresholds to acquire the kinase-ON state exhibiting different sensitivity to membrane fluidity changes.

In *toto*, we demonstrate here that YvfT, SA1313 and BA5598 respond to changes in environmental temperature adopting a kinase-ON state upon a cold shock, sensing different dynamic changes of membrane lipids. Finally, considering the similarity of the promoters of the putative ABC transporters and P*des*, as well as the high sequence identity between DesR and the studied RRs, we postulate that YvfT/YvfU, SA1313/SA1314, and BA5598/BA5597 control the expression of their cognate transporters. As the translocated molecules in each case are unknown, it is still unclear which is the response triggered by the cells upon a decrease in temperature and/or changes in the cytoplasmic membrane properties. Identifying which are the molecules exported by the ABC transporters would help to elucidate the function of these related TCSs and dissect their roles in bacterial pathogenesis.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

PF, ARD, MFR, and LP performed the experiments. PF and MCM designed the study and conceived the experiments. PF, DA, and MCM wrote the manuscript, with input from ARD and DdM. All authors analyzed the data, have read, and approved the final version.

FUNDING

This work was supported by grants from Universidad Nacional del Sur (PGI 24/ZB53, ARD), Universidad Nacional de Rosario (1BIO367, MCM), Consejo Nacional de Investigaciones Científicas y Técnicas (PUE-2016-0039, MCM), and Agencia Nacional de Promoción Científica y Tecnológica (PID-2013-0043, DdM).

ACKNOWLEDGMENTS

We thank Marta Perego for the gift of chromosomal DNA from *B. anthracis* 34F2, and Florencia Torresi for purifying SA1314-His. We are grateful to Marina Avecilla, Viviana Villalba, and Marina Perozzi for excellent technical assistance in the laboratory. PF, MFR, and LP were fellows, ARD is a staff member and DdM, DA, and MCM are Career Investigators of Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2020.592747/full#supplementary-material

REFERENCES

Aguilar, P. S., Hernandez-Arriaga, A. M., Cybulski, L. E., Erazo, A. C., and de Mendoza, D. (2001). Molecular basis of thermosensing: a two-component signal transduction thermometer in *Bacillus subtilis*. *EMBO J.* 20, 1681–1691. doi: 10.1093/emboj/20.7.1681

Albanesi, D., Mansilla, M. C., and de Mendoza, D. (2004). The membrane fluidity sensor DesK of *Bacillus subtilis* controls the signal decay of its cognate response regulator. *J. Bacteriol.* 186, 2655–2663. doi: 10.1128/jb.186.9.2655-2663.2004

Altabe, S. G., Aguilar, P. S., Caballero, G. M., and de Mendoza, D. (2003). The *Bacillus subtilis* acyl lipid desaturase is a Δ5 desaturase. *J. Bacteriol.* 185, 3228–3231. doi: 10.1128/jb.185.10.3228-3231.2003

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-80362

Braun, Y., Smirnova, A. V., Schenk, A., Weingart, H., Burau, C., Muskhelishvili, G., et al. (2008). Component and protein domain exchange analysis of a thermoresponsive, two-component regulatory system of *Pseudomonas syringae*. *Microbiology* 154, 2700–2708. doi: 10.1099/mic.0.2008/018820-18820

Buschiazzo, A., and Trajtenberg, F. (2019). Two-Component sensing and regulation: how do histidine kinases talk with response regulators at the molecular level? *Annu. Rev. Microbiol.* 73, 507–528. doi: 10.1146/annurev-micro-091018-54627

Capra, E. J., Perchuk, B. S., Skerker, J. M., and Laub, M. T. (2012). Adaptive mutations that prevent crosstalk enable the expansion of paralogous signaling protein families. *Cell* 150, 222–232. doi: 10.1016/j.cell.2012.05.033

Chakraborty, S., Li, M., Chatterjee, C., Sivaraman, J., Leung, K. Y., and Mok, Y.-K. (2010). Temperature and Mg2+ sensing by a novel PhoP-PhoQ two-component system for regulation of virulence in *Edwardsiella tarda*. *J. Biol. Chem.* 285, 38876–38888. doi: 10.1074/jbc.M110.179150

Cybulski, L. E., Albanesi, D., Mansilla, M. C., Altabe, S., Aguilar, P. S., and de Mendoza, D. (2002). Mechanism of membrane fluidity optimization: isothermal control of the *Bacillus subtilis* acyl-lipid desaturase. *Mol. Microbiol.* 45, 1379–1388. doi: 10.1046/j.1365-2958.2002.03103.x

Cybulski, L. E., del Solar, G., Craig, P. O., Espinosa, M., and de Mendoza, D. (2004). *Bacillus subtilis* DesR functions as a phosphorylation-activated switch to control membrane lipid fluidity. *J. Biol. Chem.* 279, 39340–39347. doi: 10.1074/jbc.M405150200

Cybulski, L. E., Martín, M., Mansilla, M. C., Fernández, A., and de Mendoza, D. (2010). Membrane thickness cue for cold sensing in a bacterium. *Curr. Biol.* 20, 1539–1544. doi: 10.1016/j.cub.2010.06.074

de Mendoza, D. (2014). Temperature sensing by membranes. *Annu. Rev. Microbiol.* 68, 101–116. doi: 10.1146/annurev-micro-091313-103612

Díaz, A. R., Porrini, L., de Mendoza, D., and Mansilla, M. C. (2019). A genetic screen for mutations affecting temperature sensing in *Bacillus subtilis*. *Microbiology* 165, 90–101. doi: 10.1099/mic.0.000741

Diez, V., Schujman, G. E., Gueiros-Filho, F. J., and de Mendoza, D. (2012). Vectorial signalling mechanism required for cell-cell communication during sporulation in *Bacillus subtilis*. *Mol. Microbiol.* 83, 261–274. doi: 10.1111/j.1365-2958.2011.07929.x

Harwood, C. R., and Cutting, S. M. (1990). *Molecular Biological Methods for Bacillus*. Chichester: Wiley.

Jin, S., Song, Y. N., Deng, W. Y., Gordon, M. P., and Nester, E. W. (1993). The regulatory VirA protein of *Agrobacterium tumefaciens* does not function at elevated temperatures. *J. Bacteriol.* 175, 6830–6835. doi: 10.1128/JB.175.21.6830-6835.1993

Johansson, P., and Hederstedt, L. (1999). Organization of genes for tetrapyrrole biosynthesis in gram–positive bacteria. *Microbiology* 145, 529–538. doi: 10.1099/13500872-145-3-529

Klein, W., Weber, M. H., and Marahiel, M. A. (1999). Cold shock response of *Bacillus subtilis*: isoleucine-dependent switch in the fatty acid branching pattern for membrane adaptation to low temperatures. *J. Bacteriol.* 181, 5341–5349. doi: 10.1128/JB.181.17.5341-5349.1999

Martín, M., Albanesi, D., Alzari, P. M., and de Mendoza, D. (2009). Functional in vitro assembly of the integral membrane bacterial thermosensor DesK. *Protein Expr. Purif.* 66, 39–45. doi: 10.1016/j.pep.2009.02.006

Martín, M., and de Mendoza, D. (2013). Regulation of *Bacillus subtilis* DesK thermosensor by lipids. *Biochem. J.* 451, 269–275. doi: 10.1042/BJ20121825

Najle, S. R., Inda, M. E., de Mendoza, D., and Cybulski, L. E. (2009). Oligomerization of *Bacillus subtilis* DesR is required for fine tuning regulation of membrane fluidity. *Biochim. Biophys. Acta* 1790, 1238–1243. doi: 10.1016/j.bbagen.2009.07.002

Najnin, T., Siddiqui, K. S., Taha, Elkaid, N., Kornfeld, G., Curmi, P. M. G., et al. (2016). Characterization of a temperature-responsive two component regulatory system from the *Antarctic archaeon. Methanococcoides burtonii*. *Sci. Rep.* 6:24278. doi: 10.1038/srep24278

Notredame, C., Higgins, D. G., and Heringa, J. (2000). T-coffee: a novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* 302, 205–217. doi: 10.1006/jmbi.2000.4042

Porrini, L., Cybulski, L. E., Altabe, S. G., Mansilla, M. C., and de Mendoza, D. (2014). Cerulenin inhibits unsaturated fatty acids synthesis in *Bacillus subtilis* by modifying the input signal of DesK thermosensor. *Microbiologyopen* 3, 213–224. doi: 10.1002/mbo3.154

Saita, E., Abriata, L. A., Tsai, Y. T., Trajtenberg, F., Lemmin, T., Buschiazzo, A., et al. (2015). A coiled coil switch mediates cold sensing by the thermosensory protein DesK. *Mol. Microbiol.* 98, 258–271. doi: 10.1111/mmi.13118

Skerker, J. M., Perchuk, B. S., Siryaporn, A., Lubin, E. A., Ashenberg, O., Goulian, M., et al. (2008). Rewiring the specificity of two-component signal transduction systems. *Cell* 133, 1043–1054. doi: 10.1016/j.cell.2008.04.040

Spizizen, J. (1958). Transformation of biochemically deficient strains of *Bacillus subtilis* by deoxyribonucleate. *Proc. Natl. Acad. Sci. U S A.* 44, 1072–1078. doi: 10.1073/pnas.44.10.1072

Steele, K. H., O'Connor, L. H., Burpo, N., Kohler, K., and Johnston, J. W. (2012). Characterization of a ferrous iron-responsive two-component system in nontypeable *Haemophilus influenzae*. *J. Bacteriol.* 194, 6162–6173. doi: 10.1128/JB.01465-1412

Suzuki, I., Los, D. A., Kanesaki, Y., Mikami, K., and Murata, N. (2000). The pathway for perception and transduction of low-temperature signals in *Synechocystis*. *EMBO J.* 19, 1327–1334. doi: 10.1093/emboj/19.6.1327

Trajtenberg, F., Albanesi, D., Ruétalo, N., Botti, H., Mechaly, A. E., Nieves, M., et al. (2014). Allosteric activation of bacterial response regulators: the role of the cognate histidine kinase beyond phosphorylation. *MBio* 5:e02105. doi: 10.1128/mBio.02105-2114

Trajtenberg, F., Graña, M., Ruétalo, N., Botti, H., and Buschiazzo, A. (2010). Structural and enzymatic insights into the ATP binding and autophosphorylation mechanism of a sensor histidine kinase. *J. Biol. Chem.* 285, 24892–24903. doi: 10.1074/jbc.M110.147843

Trajtenberg, F., Imelio, J. A., Machado, M. R., Larrieux, N., Marti, M. A., Obal, G., et al. (2016). Regulation of signaling directionality revealed by 3D snapshots of a kinase:regulator complex in action. *Elife* 5:e21422. doi: 10.7554/eLife.21422

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., et al. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* 46, W296–W303. doi: 10.1093/nar/gky427

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Fernández, Díaz, Ré, Porrini, de Mendoza, Albanesi and Mansilla. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Molecular Biosciences

MINI REVIEW
published: 18 February 2021
doi: 10.3389/fmolb.2021.634479

Metabolic Modeling to Interrogate Microbial Disease: A Tale for Experimentalists

Fabrice Jean-Pierre[1], *Michael A. Henson*[2] *and George A. O'Toole*[1]*

[1]*Department of Microbiology and Immunology, Geisel School of Medicine at Dartmouth, Hanover, NH, United States,* [2]*Department of Chemical Engineering and Institute for Applied Life Sciences, University of Massachusetts, Amherst, MA, United States*

The explosion of microbiome analyses has helped identify individual microorganisms and microbial communities driving human health and disease, but how these communities function is still an open question. For example, the role for the incredibly complex metabolic interactions among microbial species cannot easily be resolved by current experimental approaches such as 16S rRNA gene sequencing, metagenomics and/or metabolomics. Resolving such metabolic interactions is particularly challenging in the context of polymicrobial communities where metabolite exchange has been reported to impact key bacterial traits such as virulence and antibiotic treatment efficacy. As novel approaches are needed to pinpoint microbial determinants responsible for impacting community function in the context of human health and to facilitate the development of novel anti-infective and antimicrobial drugs, here we review, from the viewpoint of experimentalists, the latest advances in metabolic modeling, a computational method capable of predicting metabolic capabilities and interactions from individual microorganisms to complex ecological systems. We use selected examples from the literature to illustrate how metabolic modeling has been utilized, in combination with experiments, to better understand microbial community function. Finally, we propose how such combined, cross-disciplinary efforts can be utilized to drive laboratory work and drug discovery moving forward.

Keywords: metabolic modeling, metabolite cross-feeding, gut microbiome, cystic fibrosis, drug discovery

OPEN ACCESS

Edited by:
Manuel Espinosa, Spanish National Research Council, Spain

Reviewed by:
Fabian Lorenzo-Diaz, University of La Laguna, Spain
Tom, Coenye, Ghent University, Belgium

***Correspondence:**
George A. O'Toole
georgeo@dartmouth.edu

Specialty section:
This article was submitted to Molecular Recognition, a section of the journal Frontiers in Molecular Biosciences

Received: *27 November 2020*
Accepted: *19 January 2021*
Published: *18 February 2021*

Citation:
Jean-Pierre F, Henson MA and O'Toole GA (2021) Metabolic Modeling to Interrogate Microbial Disease: A Tale for Experimentalists. Front. Mol. Biosci. 8:634479. doi: 10.3389/fmolb.2021.634479

INTRODUCTION

Bacteria possess an incredible arsenal of tools allowing them to thrive in diverse environments. Such an adaptive capacity partly stems from their ability to metabolize nutrients that are present in their surroundings thus allowing them to sustain microbial growth in complex environments (Görke and Stülke, 2008). Multiple reports have revealed how metabolic features are important for bacterial persistence, virulence and drug tolerance during the infection process (Eisenreich et al., 2015; Peng et al., 2015; Eisenreich et al., 2017; Adamowicz et al., 2018; Herrero-Fresno and Olsen, 2018; Sprenger et al., 2018; Crabbé et al., 2019; Perinbam et al., 2020). However, obtaining a complete portrait of how metabolism and metabolic interactions can impact these diverse bacterial phenotypes is still a matter of active research (Zuniga et al., 2017).

Observations dating from as early as the 17th century examining dental plaque support the idea that microorganisms often live as polymicrobial, biofilm-like communities and not as single-species

in a planktonic environment (Leewenhoeck, 1684; O'Toole, 2016). Such complex biofilm communities are composed of microbial species that have the potential to interact through exchange of metabolites, ultimately shaping the microbial ecosystem (Zuniga et al., 2017; Antoniewicz, 2020).

Culture-based methods have helped us understand phenotypes associated with specific microorganisms. In contrast, the presence of complex microbial community diversity and spatiotemporal organization existing in some environments, such as in the oral cavity or the human gut, makes it inherently difficult to obtain a clear picture of how metabolic interactions among community members influence, and are being influenced by, their local environment (Peters et al., 2012; Miller et al., 2019). The advent of culture-independent techniques have paved the way to a new golden age in the understanding of microbial biology and ecology, revealing the complexity existing within complex microbial communities and their impact on human health, agriculture and the environment (Kinross et al., 2011; Chaparro et al., 2012). Indeed, it has been shown that diseases such as colorectal cancer and inflammatory bowel disease can be associated with imbalances in the microbiome, also known as dysbiosis (Gilbert et al., 2016). While the use of next-generation sequencing technologies now make it clear that many human niches are colonized by multiple microbial species and taxonomic/functional features can be associated with healthy and disease states, knowing only "who is there" and keeping a "gene catalogue" is not sufficient to probe functional metabolic roles of each member present in a community (Frioux et al., 2020). Furthermore, it is likely that there are instances when communities, classified as "different" based on sequencing data share common metabolic features. For instance, a study published by Jorth and colleagues looking at microbial communities associated to periodontal disease reported that no clear polymicrobial community structure could be associated with disease due to high interpatient variability (Jorth et al., 2014). However, assessing the metabolic capacities of these communities revealed that disease-associated microbial clusters maintained a conserved functional profile (Jorth et al., 2014).

Although metagenomics can alleviate some of the shortcomings of 16S-based rRNA gene amplicon studies by providing an overview of metabolic potential and metabolomics can identify the abundant, accumulated metabolites in a community, these approaches alone are still not sufficient to probe essential metabolic functions and metabolite-based interactions among microbes in a community (Quince et al., 2017; Frioux et al., 2020). The importance of developing novel strategies to understand how metabolic interactions drive community structure and function is further highlighted by recent reports indicating that such microbial or host-driven metabolic interactions can modulate drug resistance and tolerance (Peters et al., 2012; Crabbé et al., 2019; Crabbé et al., 2019; Orazi and O'Toole, 2019).

Therefore, we propose that *metabolic modeling* is one such tool that exploits existing amplicon, metagenomics and metabolomics data to generate hypotheses that can complement and help drive experimental studies, thereby validating the computational interrogations of these existing datasets. Metabolic modeling is an *in silico* predictive mathematical modeling approach that leverages genome-scale metabolic model (GEM) reconstructions of cellular metabolism derived from genomic annotations (Biggs et al., 2015; Zomorrodi and Segre, 2016; Zampieri et al., 2019). Since the *in silico* reconstruction of the first GEM of *Haemophilus influenzae* in 1999 (Edwards and Palsson, 1999) hundreds of novel GEMs have been generated (through automatic, semi-automatic or manual means) as new genomes are being sequenced (Gu et al., 2019). This approach has been used to drive biochemical knowledge of biological systems by 1) translating functional annotation information into metabolic predictions, 2) probing metabolic features involved in metabolite production, and 3) advancing the understanding of metabolic interactions among microbial species with their host (Thiele et al., 2013; Simeonidis and Price, 2015; Kumar et al., 2019). As a growing number of high-quality GEMs are available through several databases (Gu et al., 2019; Noronha et al., 2019; Seaver et al., 2020) and can be used to generate predictions of community biological functions through various modeling tools such as SteadyCOM, CASINO, COMET and BacArena (Harcombe et al., 2014; Shoaie et al., 2015; Bauer et al., 2017; Chan et al., 2017; Altamirano et al., 2020), metabolic modeling represents a powerful approach to interrogate complex community metabolic interactions involved in health and disease (Gu et al., 2019; Kumar et al., 2019). Furthermore, metabolic modeling methods have the capacity to guide laboratory work by integrating and making predictions based on 16S rRNA gene, metagenomics, and metabolomics datasets (Rai and Saito, 2016).

With recent evidence pointing towards the impact of polymicrobial metabolic interactions shaping human health, we review here some of the latest research leveraging metabolic modeling and genome-scale metabolic reconstructions used to interrogate metabolic interactions among the members of complex microbial communities in health and disease. We first start by surveying studies in the context of the human gut microbiome for which metabolic modeling has been mainly applied, and then transition to the case of cystic fibrosis (CF), a genetic disease where persons with CF (pwCF) accumulate thick secretions (mucus) in their airways, which creates an ideal nutrient-rich environment wherein pathogens can thrive, ultimately leading to high morbidity and mortality (Boucher, 2007; O'Sullivan and Freedman, 2009; Surette, 2014). We also present findings indicating how metabolic modeling could be used to develop new antimicrobial drugs. We review this literature through the lens of experimentalists who can integrate such modeling data to focus their efforts on understanding the role of metabolic interactions in driving microbial community structure and function.

Metabolic Modeling to Probe Microbial Interactions and Understand Disease

Human health can be profoundly impacted by the presence of microbial communities, which must include taking into account how the individual microbes in these communities interact with

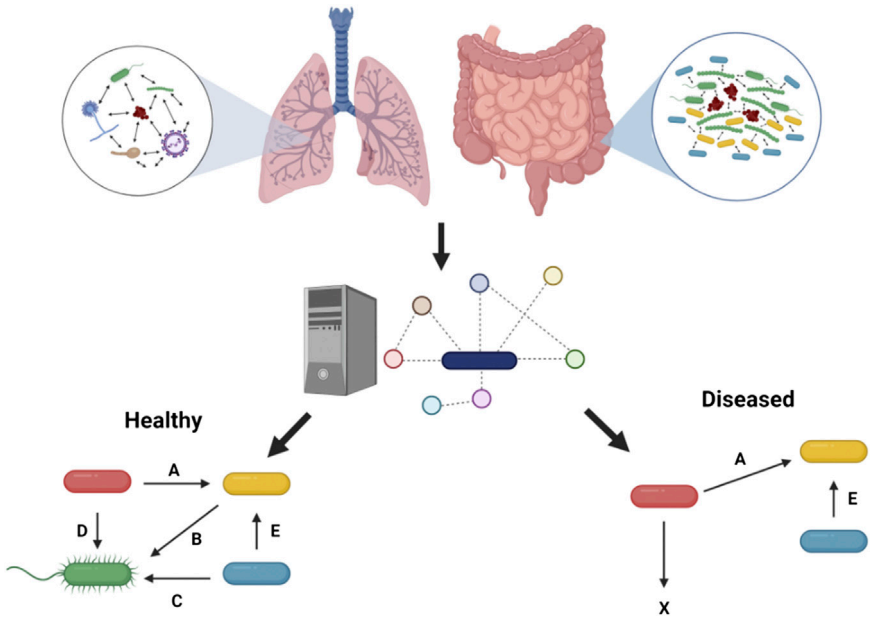


FIGURE 1 | Metabolic modeling to understand health and disease. Microbial interactions observed in the human gut and in the context of chronic lung disease such as in pwCF can be predicted through metabolic modeling to pinpoint metabolic cross-feeding interactions (denoted by letters) driving community structure and associated with healthy or diseased states. These predictions are facilitated by the capacity of this *in silico* approach to integrate *in vivo*-like nutritional and physico-chemical parameters, and ultimately help guide experimentation. Figure designed using BioRender

each other (Peters et al., 2012; Miller et al., 2019; Orazi and O'Toole, 2019). Although some of these metabolic interactions among microbes in a community can be beneficial to maintain health, for example the presence of microbial communities capable of digesting complex molecules in the gut (Backhed et al., 2005), some chronic infectious diseases, such as periodontitis, CF and diabetic foot ulcers, have been shown to be driven by polymicrobial communities (Dowd et al., 2008; Lebeaux et al., 2014; O'Toole, 2018; Miller et al., 2019). We start here with a discussion of findings using metabolic modeling to study microbial community function, largely with a focus on communities relevant to disease, especially in the case of the human gut (**Figure 1**). Many of these studies validate the modeling findings by using *in vitro* experimental models.

As mentioned above, one of the research areas benefiting from metabolic modeling is the study of the human gut microbiome (Shoaie and Nielsen, 2014; Ji and Nielsen, 2015; Kumar et al., 2019; Sen and Oresic, 2019). The human gut is composed of trillions of microbial cells of diverse species that have an important impact on human health as they can perform a number tasks, including but not limited to, metabolizing complex molecules that cannot be degraded by the eukaryotic cells in the intestine to provide energy to colonic cells and playing a role in immune function (Bull and Plummer, 2014; Shoaie and Nielsen, 2014). Short chain fatty acids (SCFAs) produced by the gut microbiota are absorbed by the intestinal epithelial cells and impact several aspect of human health, from energy regulation to the immune programming (Comalada et al., 2006). The concentration of the SCFAs acetate, butyrate and propionate in the human body can vary depending on the diet (den Besten et al., 2013). While acetate is the SCFA mainly found in the blood and contributes to several metabolic functions such as lipogenesis and glyconeogenesis (Comalada et al., 2006), butyrate is preferred by human colonocytes for high energy production (Hamer et al., 2008; Dumas, 2011). Imbalances in the metabolism of SCFAs characterized by low butyrate production have been associated with human diseases such as inflammatory bowel disease, type 2 diabetes, obesity and other pathologies (Kumar et al., 2019).

SCFAs have also been found to be important players in the maintenance of a healthy gut microbiome through microbial interactions (Shoaie and Nielsen, 2014). For example, the presence of enzymatic degradation activity of dietary fibers into SCFAs by microbial communities in the human gut is critical to maintain health (Koh et al., 2016). To better understand how microbial interactions impact human and microbial metabolism, key human gut bacteria, that is, *Bacteroides thetaiotamicron, Eubacterium rectale* and *Methanobrevibacter smithii* (part of the Bacteroides, Firmicute and Euryarchaeota phyla, respectively) previously identified to be SCFA producers (Holmes et al., 2012) have been studied through metabolic modeling. Shoaie and colleagues reconstructed GEMs of these three bacteria by coalescing multiple biochemical reaction databases with manual curation of metabolic pathways (Shoaie and Nielsen, 2014). This work resulted in the creation of three GEMs that were used to 1) predict the metabolic landscape of metabolites that are produced and consumed within that community and 2) infer microbial biomass abundance based on known metabolites present in the environment. This analysis predicted microbe-microbe and host-microbe interactions via metabolites using a

compartmentalization metabolic framework, that is, all the predicted metabolic reactions occurring within an organism were modeled and the produced metabolites allowed to be exchanged freely through a shared compartment (Biggs et al., 2015). The authors observed that acetate was the main metabolite predicted to be cross-fed between *E. rectale* and *B. thetaiotamicron* when modeled *in silico* in the germ-free gut extracellular space. Also, the *B. thetaiotamicron*-derived acetate was predicted to be driving the production of butyrate by *E. rectale*; butyrate was in turn consumed by the gut epithelial cells. These predictions are in agreement with published experimental data whereby germ-free mice co-infected with both species resulted in changing concentrations of these metabolites (Mahowald et al., 2009; Sen and Oresic, 2019). The inclusion of transcriptomic data from infected mice in the model also allowed the investigation of *E. rectale* and *B. thetaiotamicron* gene expression changes in the presence of each other (*vs.* monospecies), and could serve to validate the modeling data. *In silico* modeling indicated that *E. rectale* increases the utilization of the amino acid glutamine, whereas pathways for polysaccharide utilization in *B. thetaiotamicron* were hypothesized to increase when *E. rectale* is present; predictions that are agreement with previously published transcriptomic data (Mahowald et al., 2009). Therefore, in this instance, metabolic modeling represents a useful approach to understand how microbial interactions can impact SCFAs production and ultimately gut health.

The growing number of studies using germ-free mice models in combination with metabolomics and next generation sequencing constitute rich data sets that can feed metabolic predictions and ultimately better probe metabolic features impacting human health (Ji and Nielsen, 2015; Franzosa et al., 2019; Lavelle and Sokol, 2020). Thus, understanding how diet can impact gut health and microbial community structure through modeling approaches constitute a powerful approach that can drive experimentalist to build models to validate testable hypotheses and perhaps better understand chronic gut disease. As more GEMs are constructed and becoming available (Magnúsdóttir et al., 2017) a growing number of studies are building more complex community metabolic models to understand the interactions occurring among microbes and the host in the gut (Kumar et al., 2019).

In one of the most extensive studies modeling multiple microbial interactions in the human intestine, Henson & Phalak leveraged semi-curated genome-scale metabolic reconstructions (described above as GEMs) of 28 bacterial species detected in the human intestine (Henson and Phalak, 2018) available through the Virtual Metabolic Human database (Noronha et al., 2019). This model included a total of 22,203 genes, 26,867 metabolites and 35,031 reactions and revealed that optimal metabolic community growth rate resulted in decreased gut microbial diversity and the enrichment of just a few bacterial genera such as *Escherichia, Enterobacter* and *Citrobacter* (part of the Enterobacteriaceae family), all known to be detected in high numbers in the gut from clinical studies of patients with inflammatory bowel disease (IBD) (Rhodes, 2007; Saleh and Elson, 2011; Kaakoush et al., 2012). Furthermore, optimal community growth rate predictions resulted in different SCFA production levels (acetate overproduction and no butyrate production) compared to published experimental data (den Besten et al., 2013) suggesting that *in vivo* optimal community growth objectives might be kept in check to ensure appropriate SCFA production and healthy gut homeostasis. For example, the authors observed that reduced community growth rate, (i.e. not at its maximal value) increased microbial diversity and produced SCFA synthesis rates more consistent with those observed *in vivo* (den Besten et al., 2013). Interestingly, a recent report examining the changes in microbial community composition in healthy and IBD individuals by 16S rRNA gene amplicon sequencing observed a decrease in microbial diversity and an increase in Enterobacteriaceae abundance in the gut of diseased *vs*. healthy individuals (Alam et al., 2020). This approach of using sub-optimal growth rates in the context of communities is a fundamentally different approach compared to traditional, single-species metabolic modeling approaches that assume maximal growth rates (Biggs et al., 2015); thus, in this case experimental data helped to fine-tune the modeling approach. Moreover, metabolic modeling represents a complementary tool to experimental observations to better understand the metabolic perturbations that might be associated with imbalances in microbial diversity and abundance.

While metabolic modeling can be used to probe interactions hypothesized to drive gut health and microbial community structure, this approach can also be used to understand gut microbial imbalances in the context of infection. The presence of a healthy gut microbiome has been reported to be important for protection against pathogenic species such as the Proteobacterium *Clostridioides difficile* (Lagier, 2016; Leslie et al., 2019). Furthermore, other risk factors such as antimicrobial treatment have been implicated in the emergence of *C. difficile* infections (CDIs) (Lagier, 2016). As CDIs have been reported to cause between 15,000 and 30,000 deaths annually in the US (Lessa et al., 2015) and many healthy individuals are asymptomatic carriers of *C. difficile* (Furuya-Kanamori et al., 2015), it is important to understand how gut metabolic interactions can impact colonization resistance against *C. difficile*. By modeling microbial interactions in multispecies biofilms through the use of GEMs of *B. thetaiotaomicron* (Bacteroidetes), *Faecalibacterium prausnitzii* (Firmicute), *Escherichia coli* (Proteobacteria) and *C. difficile*, Phalak and Henson observed that among the metabolites that were predicted to be produced and cross-fed, that is, acetate, ethanol, formate and succinate, only formate produced by both *F. prausnitzii* and *E. coli* as well as *B. thetaiotaomicron*-derived succinate were hypothesized to be positively driving *C. difficile* abundance. Moreover, *in silico* removal of both organic acids would block this pathogenic microbe's expansion (Phalak and Henson, 2019). Interestingly using a co-colonization gnobiotic mouse model, Ferreyra and colleagues observed that *B. thetaiotaomicron*-derived succinate can be utilized by *C. difficile* to allow this pathogen to thrive in the gut (Ferreyra et al., 2014), a finding consistent with the metabolic modeling data. Phalak & Henson also observed that acetate produced by *C. difficile* was hypothesized to be the only metabolite cross-fed to

F. prausnitzii in a mixed community. However, the hypothesized cross-feeding of metabolites remain to be experimentally validated, perhaps through the development of a four-species community or using a simplified community in a mouse model.

Metabolic modeling has also been used to understand other human diseases such as CF where the lung environment constitutes an ideal locale for disease-causing pathogens to colonize, resulting in chronic infections and promoting negative clinical outcomes (Lipuma, 2010; Surette, 2014). The chronic nature of these infections is due to the presence of polymicrobial biofilm-like communities that exhibit both resistance and tolerance (collectively called recalcitrance) to front-line drugs targeted at these infections (O'Toole, 2018; Orazi and O'Toole, 2019; Vandeplassche et al., 2019). While the development of novel antimicrobials has resulted in the increase in life expectancy for pwCF, these individuals are still burdened with significant clinical symptoms (Limoli et al., 2016; O'Toole, 2018; Bevivino et al., 2019; Limoli and Hoffman, 2019). Several groups have attempted to identify the etiology of chronic CF lung disease, but this remains a topic of active research (O'Toole, 2018). Multiple groups have experimentally reported metabolic interactions among bacterial species that are prevalent and abundant in the CF airway (Crabbé et al., 2014; Orazi and O'Toole, 2017; Tavernier et al., 2017; Tavernier et al., 2018; Crabbé et al., 2019; Orazi and O'Toole, 2019; Orazi et al., 2019; Vandeplassche et al., 2019; Camus et al., 2020; Orazi et al., 2020). Based on these observations, Henson *et al.* tested the hypothesis that it would be possible to infer microbial abundance and map key metabolic interactions that potentially drive CF lung disease by modeling 17 of the most abundant species that encompass most of the 16S rRNA gene reads from three published microbiome studies (Henson et al., 2019). Interestingly, they reported that metabolic simulations performed using different *in silico* lung environments could reproduce experimentally observed 16S rRNA gene normalized reads detected in the CF airway including the infrequent microbial species *Achromobacter, Escherichia* and *Burkholderia*. Metabolic modeling approaches also allowed for the observation of interpatient community composition heterogeneity (β-diversity) that is common in pwCF (Filkins et al., 2012; Price et al., 2013; Carmody et al., 2018). The model accurately predicted that *Prevotella*, *Pseudomonas* and *Streptococcus* would dominate the polymicrobial community, an observation that is strong agreement with multiple experimental studies highlighting the impact of these pathogens on the CF airway (Lipuma, 2010; Scott and O'Toole, 2019; Thornton and Surette, 2020). Furthermore, Henson *et al.*, reported that predicted metabolic interactions in some communities were mainly driven by *Pseudomonas* and/or *Streptococcus*, which are highly abundant microbes capable of metabolizing nutrients such as organic acids, amino acids and secreted alcohols. For instance, the presence of rare pathogens was hypothesized to be driven by 1) acetate, formate and L-lactate produced by *Streptococcus* and consumed by *Escherichia*, 2) acetate, alanine and formate cross-fed from *Pseudomonas* and *Streptococcus* to *Burkholderia*, and to a lesser extent 3) alanine, threonine and L-lactate metabolite consumption by *Achromobacter*.

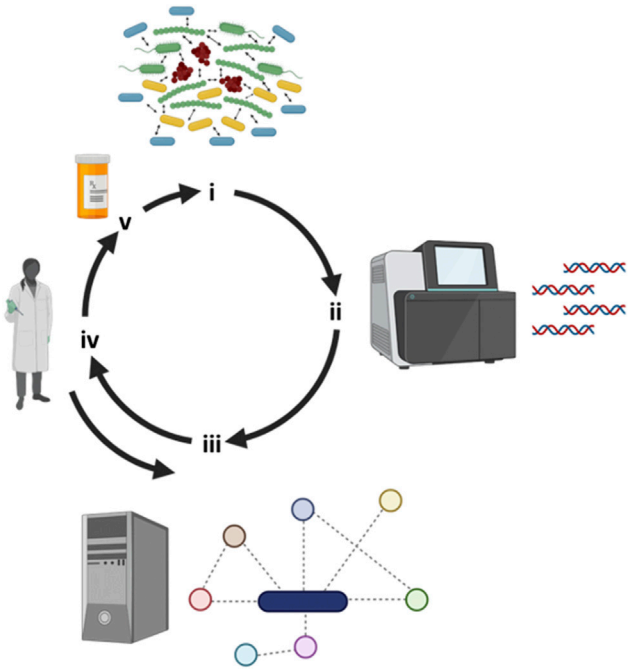


FIGURE 2 | Iterative workflow between clinical observations, metabolic modeling and experimental work. (**i**) Several recalcitrant biofilm-based chronic infectious diseases are polymicrobial in nature and next-generation tools have allowed us to (**ii**) identify "who is there". (**iii**) Metabolic modeling can leverage sequencing information by reconstructing metabolic networks and predicting key metabolites responsible of driving community structure and function. (**iv**) Metabolic predictions can then be experimentally tested and the model fine-tuned with the integration of this new experimental data. (**v**) Through modeling predictions, novel drugs can then be designed or repurposed to negatively impact polymicrobial communities. Figure designed using BioRender.

While some studies have focused on the metabolic interactions of microbes found in the CF airway, most of them have included two microbial species due to complex experimental methods required to culture multiple microbes in a mixed community *in vitro* (Filkins et al., 2015; Flynn et al., 2016; Scott et al., 2019; Camus et al., 2020; Li et al., 2020). However, metabolic modeling represents a useful tool to interrogate metabolic dependencies of more complex multispecies communities detected in pwCF. Furthermore, recent reports indicate that defects in the cystic fibrosis transmembrane regulator impacts mitochondrial metabolism and results in an imbalance of metabolites produced by immune cells, with possible impacts on microbial species such as *P. aeruginosa* and *S. aureus* (Gabryszewski et al., 2019; Riquelme et al., 2019; Riquelme et al., 2020a; Riquelme et al., 2020b).

The findings presented here further support the idea that metabolic modeling represents an effective approach to better understand how metabolic interactions among microorganisms can impact human health and disease such as in the gut and the CF airway. That is, by integrating experimental data from these

studies, metabolic modeling can generate predictions about how polymicrobial communities might interact and a serve as a tool to develop hypotheses to be tested in the laboratory and unlock insight as to which metabolites might drive changes from healthy to diseased states.

Metabolic Modeling as a Tool to Identify Novel Antimicrobial Targets

Targeting pathways implicated in the metabolism of essential nutrients utilized during infection might represent a novel approach for the identification and or development of antimicrobials that can interfere with microbial growth in the context of human disease (**Figure 2**) (Flynn et al., 2016; Flynn et al., 2020). The studies below use a combination of metabolic modeling and experimentation with a focus on individual microbes to better understand responses to antimicrobial therapy. This *in silico* method has also proven to be useful to understand how metabolic reprogramming can impact gene essentiality in microorganisms exposed to various antimicrobials, and thus may also help in the identification of novel drug targets.

Several groups have successfully applied metabolic modeling approaches to understand the global metabolic changes associated with microbial drug exposure, including for *P. aeruginosa* and polymixin (Zhu et al., 2018) and *Acinetobacter baumanii* exposed to colistin (Presta et al., 2017). While remaining to be experimentally validated, these *in silico* predictions revealed that exposure to antimicrobials resulted in known non-essential genes becoming essential in the presence of the drug, a phenomenon sometimes referred to as conditional or synthetic lethality. Although experimental approaches can be used to identify genes and/or synthetic lethal relationships that could represent novel drug targets (Côté et al., 2016), metabolic modeling has the advantage of facilitating the identification of such pathways and/or metabolic interactions by computationally inactivating the pathways, thus potentially helping to guide laboratory work to facilitate antimicrobial development (Raman et al., 2018).

Cesur and colleagues investigated predicted drugs targeted against *Klebsiella pneumoniae*, part of the ***E**nterococcus faecium*, ***S**taphylococcus aureus*, ***K**lebsiella pneumoniae*, ***A**cinetobacter baumannii*, ***P**seudomonas aeruginosa*, and ***E**nterobacter* species (ESKAPE) multidrug resistant group, for which cephalosporin and carbapenem resistance have been identified in part due to widespread acquisition β-lactamase-coding and carbapenemase-coding genes (Cesur et al., 2020; De Oliveira et al., 2020). *Klebsiella* species are capable of causing pneumonia and are associated with increased morbidity and mortality (Bengoechea and Sa Pessoa, 2019). By leveraging GEM of *K. pneumoniae* in host-like conditions in combination with *in silico* structure-based drug screening, these investigators identified several candidate molecules potentially targeting 2-dehydro-3-deoxyphosphooctonate aldolase activity; the gene coding for this enzyme, *kdsA*, is present in 100% of pathogenic *K. pneumoniae* strains and also considered an essential gene (Ramos et al., 2018). However, these predictions remain to be validated experimentally by using *in vitro* and/or *in vivo* models where the essentiality of the *kdsA* gene is tested by inactivating this locus (if possible) and also by verifying the impact of the identified molecules on the viability (or virulence traits) of *K. pneumoniae* species.

Similar approaches have also been used to study fungi. *Candida albicans* is a fungal opportunistic pathogen that is cause of increased mortality in patients suffering of candidiasis (Perlroth et al., 2007). As there are only a few classes of antifungal drugs that are used to treat *Candida* infections and an increase in strains presenting multi-drug resistant phenotypes have been observed (Arendrup and Patterson, 2017), new approaches must be developed to identify compounds that can be used as fungicidal agents *vs. Candida* species. Viana and colleagues reconstructed and validated a GEM of *C. albicans* (Viana et al., 2020) and they tested the *in silico* model for its capacity to predict essential genes in *C. albicans* that had been previously experimentally identified (O'Meara et al., 2015). The model was capable of accurately predicting 78% of the published essential genes and validated the already-known drug target gene ERG11 (Lupetti et al., 2002). Interestingly, the *C. albicans* protein Ura9 (a quinone-containing protein) was identified to be essential through modeling predictions; drugs such as Atovaquone, used to treat malarial infections caused by *Plasmodium falciparum*, could potentially be used to target Ura9. While not experimentally validated and the *in silico* predictions performed using Roswell Park Memorial Institute (RPMI) medium, which likely does not reflect the infection microenvironment in the host, the multiple targets identified through the GEM of *C. albicans* might represent novel targets for drug development or repurposing.

Metabolic modeling has also been used to find potential drug targets for parasites causing diseases in humans (Curran et al., 2020). The parasitic worm *Brugia malayi* is one of the causes of lymphatic filariasis (also known as elephantiasis) and is transmitted through blood-sucking mosquitoes; infection with this parasite can result in the swelling of limbs via chronic leakage of lymph into tissues (Chandy et al., 2011). Current effective anti-parasite drug treatment is limited as these therapies do not target all the life-stages of *B. malayi*, and to be effective, the drug regimen must be performed during the whole lifespan of the adult worm, which has previously been reported to last as long as 15 years (Molyneux et al., 2014). Recent studies have indicated that targeting *Wolbachia*, an endosymbiotic microorganism implicated in *B. malayi* fitness and its reproduction cycle, might represent an new therapeutic avenue (Taylor et al., 2013). In an attempt to identify potential drugs aimed at *Wolbachia* that would result in its elimination of *B. malayi*, Curran and colleagues reconstructed a GEM of the worm and identified 102 essential metabolic reactions predicted to be essential for the survival of the worm (Curran et al., 2020). By filtering out metabolic reactions for which no experimental expression data are available as well as the pathways with homology to human metabolic reactions, *in silico* modeling predicted that the enzymes 1-deoxy-D-xylulose 5-phosphate reductoisomerase, fructose biphosphatase and an adenylate kinase represent potential *Wolbachia* enzymes to be targeted.

Interestingly, the drugs Fosmidomycin and tenofovir (both clinically approved) specifically target those enzymes and were experimentally confirmed to reduce the abundance of *Wolbachia* in *B. malayi*, negatively impacting the worm (Curran et al., 2020).

Taken together, the studies presented in this section support the notion that the integration of experimental expression data in combination with metabolic modeling and *in silico* drug screening might represent a powerful tool to identify drugs that might already be clinically available for other indications and repurposed for new diseases. Future research integrating *in vivo*-like conditions with metabolic modeling will certainly allow for the expansion of the complex metabolic interactions occurring between the host and microbe for the identification of novel antimicrobials with increased activity. Such efforts will surely be facilitated by the construction of integrated microbe-host GEMs to better understand such interactions in the context of human health and disease (Bordbar et al., 2010; Thiele et al., 2020).

CONCLUDING REMARKS

The microbial complexity existing in various ecological niches, including in the context of human infections or microbiomes, necessitates the utilization of novel approaches to probe key metabolic features governing the interactions among microbes and/or the host. We argue that metabolic modeling represents such a tool as it has proven its usefulness in conditions where experimental methods alone were not sufficiently efficient, as for assessing the production of complex mixtures of microbial-derived metabolites (Simeonidis and Price, 2015). Starting with single species and more recently expanding to complex polymicrobial interactions, metabolic modeling has been employed to tackle increasingly complex questions by combining GEMs with nutritional information reflecting the infection environment to pinpoint metabolic pathways driving community structure and metabolite cross-feeding (Bordbar et al., 2010; Thiele et al., 2020).

Perhaps one of the most compelling ecological niches wherein metabolic modeling is currently employed and for which the most data is available is the human gut. Metabolic interactions are likely critical to maintain homeostasis between polymicrobial communities and the host in the gut (Shoaie and Nielsen, 2014; Kumar et al., 2019). Metabolic modeling has also proven to be useful in the context of microbial-based human diseases such as CDI and CF, where key metabolic features have been predicted to drive the abundance of pathogens (Henson et al., 2019; Phalak and Henson, 2019). Obtaining such detailed metabolic understanding of biochemical interactions driving microbial community structure, especially in the context of complex polymicrobial diseases or human microbial communities, could help efforts focused at the identification of new drug targets or probiotic interventions. Furthermore, the development of spatiotemporal metabolic models capable of integrating additional variables such as nutrient and oxygen diffusion/availability (Phalak et al., 2016; Chan et al., 2019; Altamirano et al., 2020) will further improve the computational predictions regarding how microbial community structure and function is impacted in complex polymicrobial diseases.

Given that multiple chronic diseases such as periodontitis, CF and diabetic foot ulcers are driven by the presence of polymicrobial communities that exhibit recalcitrance to antimicrobials (Rams et al., 2014; O'Toole, 2018; Orazi and O'Toole, 2019; Heravi et al., 2020), the field is poised to apply combined modeling and experimental approaches to better understand how to effectively treat such polymicrobial infections. To date however, little or no efforts have been directed towards modeling these complicated, often chronic polymicrobial infections and their predicted responses to antibiotic treatment.

One limitation of the approaches discussed here is the lack of a systematic evaluation of multiple GEMs existing for a same organism. That is, several GEMs with variations in model annotation, reaction stoichiometry, or the presence/absence of cofactors, can exist for a given organism but no unanimously standardized approach for the construction of a GEM has been adopted to date (Lieven et al., 2020). This lack of standardization can in turn lead to inconsistent predictions and a lack of reproducibility among models and laboratories. For instance, in an attempt to identify the best performing GEM for *Mycobacterium tuberculosis* among eight such recently published models, Lopez and colleagues observed that only two of the GEMs were able to generate robust predictions when using parameters such as the number of reaction gaps in the metabolic network, thermodynamically infeasible reactions, dead-end metabolites and gene essentiality predictions using previously published experimental data (Lopez-Agudelo et al., 2020). To remedy such potential shortcomings in the future utilization of GEMs, MEMOTE (MEtabolic MOdel TEsts) an open-source software driven by a community effort, has been published by Lieven and colleagues (Lieven et al., 2020). MEMOTE helps in improving GEM reproducibility and reuse among studies by advocating the utilization of a standardized GEM data format exchange, that is, the Systems Biology Markup Language level 3 flux balance constraints (SBML3FBC) which include defined parameters such as metabolite chemical formulas, charge and annotations, etc. (Lieven et al., 2020). Furthermore, by examining factors including biomass reaction, stoichiometry inconsistencies, annotation validation and basic tests, MEMOTE has the capacity to benchmark GEMs that are currently available throughout several databases, thus helping in the selection and continuous improvement of metabolic models.

Ultimately, we argue that metabolic modeling will prove quite valuable assisting experimentalists to focus their research efforts on a specific sets of questions originating from clinical observations by helping in the identification of key metabolic pathways likely responsible for driving disease. To validate the data from metabolic modeling studies examining complex microbial communities in the context of microbiomes or mixed-species infections, the development of more complex *in vitro/in vivo* communities composed of abundant and prevalent species will be necessary to better probe how

microbe-microbe and polymicrobial community-host metabolic interactions are impacting community structure and function, including the capability of these communities to drive negative clinical outcomes. We also envision that the integration of various *in vivo*-like environmental factors such as oxygen and nutrient diffusion/availability, pH, viscosity, etc. will be critical components to be integrated in experimental models to help in our understanding of how metabolic interactions drive community structure and function. For instance, the development of a an artificial sputum medium developed by the Whiteley group, which was formulated based on nutrients available in the CF airway, (e.g. lactate, amino acids, dextrose, etc.) and mimicking the viscosity observed in the lung environment has allowed for significant breakthroughs in CF research (Palmer et al., 2005; Palmer et al., 2007; Turner et al., 2015; Darch et al., 2018; Cornforth et al., 2020). One could argue that the creation of disease-like growth media for the gut will also result in the discovery of novel findings impacting human health.

In summary, we believe that multi-omics data, (i.e. amplicon, metagenomic, metabolomic, transcriptomic) represent critical information regarding the infection environment that will drive the construction of robust laboratory models, which in combination with metabolic modeling, will trigger an iterative process between computational predictions and laboratory validation to improve our understanding of community structure and function and to facilitate the identification of novel therapeutics.

AUTHOR CONTRIBUTIONS

FJ-P, GAO conceived of the article; FJ-P, MAH, and GO wrote and edited the manuscript.

FUNDING

This work has been supported by R37 AI83256-06 to GO and BBSRC-NSF/BIO Award #2030087 to MH.

REFERENCES

Adamowicz, E. M., Flynn, J., Hunter, R. C., and Harcombe, W. R. (2018). Cross-feeding modulates antibiotic tolerance in bacterial communities. *ISME J.* 12 (11), 2723–2735. doi:10.1038/s41396-018-0212-z

Alam, M. T., Amos, G. C. A., Murphy, A. R. J., Murch, S., Wellington, E. M. H., and Arasaradnam, R. P. (2020). Microbial imbalance in inflammatory bowel disease patients at different taxonomic levels. *Gut Pathog.* 12 (1), 1. doi:10.1186/s13099-019-0341-6

Altamirano, Á., Saa, P. A., and Garrido, D. (2020). Inferring composition and function of the human gut microbiome in time and space: a review of genome-scale metabolic modelling tools. *Comput. Struct. Biotechnol. J.* 18, 3897–3904. doi:10.1016/j.csbj.2020.11.035

Antoniewicz, M. R. (2020). A guide to deciphering microbial interactions and metabolic fluxes in microbiome communities. *Curr. Opin. Biotechnol.* 64, 230–237. doi:10.1016/j.copbio.2020.07.001

Arendrup, M. C., and Patterson, T. F. (2017). Multidrug-resistant *Candida*: epidemiology, molecular mechanisms, and treatment. *J. Infect. Dis.* 216 (3), S445–S451. doi:10.1093/infdis/jix131

Backhed, F., Ley, R. E., Sonnenburg, J. L., Peterson, D. A., and Gordon, J. I. (2005). Host-bacterial mutualism in the human intestine. *Science* 307 (5717), 1915–1920. doi:10.1126/science.1104816

Bauer, E., Zimmermann, J., Baldini, F., Thiele, I., and Kaleta, C. (2017). BacArena: individual-based metabolic modeling of heterogeneous microbes in complex communities. *PLoS Comput. Biol.* 13 (5), e1005544, doi:10.1371/journal.pcbi.1005544

Bengoechea, J. A., and Sa Pessoa, J. (2019). *Klebsiella pneumoniae* infection biology: living to counteract host defences. *FEMS Microbiol. Rev.* 43 (2), 123–144. doi:10.1093/femsre/fuy043

Bevivino, A., Bacci, G., Drevinek, P., Nelson, M. T., Hoffman, L., and Mengoni, A. (2019). Deciphering the ecology of cystic fibrosis bacterial communities: towards systems-level integration. *Trends Mol. Med.* 25 (12), 1110–1122. doi:10.1016/j.molmed.2019.07.008

Biggs, M. B., Medlock, G. L., Kolling, G. L., and Papin, J. A. (2015). Metabolic network modeling of microbial communities. *Wiley Interdiscip Rev Syst Biol Med* 7 (5), 317–334. doi:10.1002/wsbm.1308

Bordbar, A., Lewis, N. E., Schellenberger, J., Palsson, B. Ø., and Jamshidi, N. (2010). Insight into human alveolar macrophage and *M. tuberculosis* interactions via metabolic reconstructions. *Mol. Syst. Biol.* 6 (1), 422. doi:10.1038/msb.2010.68

Boucher, R. C. (2007). Airway surface dehydration in cystic fibrosis: pathogenesis and therapy. *Annu. Rev. Med.* 58, 157–170. doi:10.1146/annurev.med.58.071905.105316

Bull, M. J., and Plummer, N. T. (2014). Part 1: the human gut microbiome in health and disease. *Integr. Med.* 13 (6), 17–22.

Côté, J. P., French, S., Gehrke, S. S., MacNair, C. R., Mangat, C. S., Bharat, A., et al. (2016). Author Correction for Côté et al., The Genome-Wide Interaction Network of Nutrient Stress Genes in *Escherichia coli*. *mBio* 7 (6), e01714–01716. doi:10.1128/mBio.01714-16

Camus, L., Briaud, P., Bastien, S., Elsen, S., Doléans-Jordheim, A., Vandenesch, F., et al. (2020). Trophic cooperation promotes bacterial survival of *Staphylococcus aureus* and *Pseudomonas aeruginosa*. *ISME J.* 14 (12), 3093–3105. doi:10.1038/s41396-020-00741-9

Carmody, L. A., Caverly, L. J., Foster, B. K., Rogers, M. A. M., Kalikin, L. M., Simon, R. H., et al. (2018). Fluctuations in airway bacterial communities associated with clinical states and disease stages in cystic fibrosis. *PloS One* 13 (3), e0194060. doi:10.1371/journal.pone.0194060

Cesur, M. F., Siraj, B., Uddin, R., Durmuş, S., and Çakır, T. (2020). Network-based metabolism-centered screening of potential drug targets in *Klebsiella pneumoniae* at genome scale. *Front Cell Infect Microbiol* 9, 447. doi:10.3389/fcimb.2019.00447

Chan, S. H. J., Simons, M. N., and Maranas, C. D. (2017). SteadyCom: predicting microbial abundances while ensuring community stability. *PLoS Comput. Biol.* 13 (5), e1005539: doi:10.1371/journal.pcbi.1005539

Chan, S. H. J., Friedman, E. S., Wu, G. D., and Maranas, C. D. (2019). Predicting the longitudinally and radially varying gut microbiota composition using multi-scale microbial metabolic modeling. *Processes* 7 (7), 394. doi:10.3390/pr7070394

Chandy, A., Thakur, A. S., Singh, M. P., and Manigauha, A. (2011). A review of neglected tropical diseases: filariasis. *Asian Pac J Trop Med* 4 (7), 581–586. doi:10.1016/S1995-7645(11)60150-8

Chaparro, J. M., Sheflin, A. M., Manter, D. K., and Vivanco, J. M. (2012). Manipulating the soil microbiome to increase soil health and plant fertility. *Biol. Fertil. Soils* 48 (5), 489–499. doi:10.1007/s00374-012-0691-4

Comalada, M., Bailón, E., de Haro, O., Lara-Villoslada, F., Xaus, J., Zarzuelo, A., et al. (2006). The effects of short-chain fatty acids on colon epithelial proliferation and survival depend on the cellular phenotype. *J. Canc. Res. Clin. Oncol.* 132 (8), 487–497. doi:10.1007/s00432-006-0092-x

Cornforth, D. M., Diggle, F. L., Melvin, J. A., Bomberger, J. M., and Whiteley, M. (2020). Quantitative framework for model evaluation in microbiology research using *Pseudomonas aeruginosa* and cystic fibrosis infection as a test case. *mBio* 11 (1), e03042–03019. doi:10.1128/mBio.03042-19

Crabbé, A., Jensen, P. Ø., Bjarnsholt, T., and Coenye, T. (2019). Antimicrobial tolerance and metabolic adaptations in microbial biofilms. *Trends Microbiol.* 27 (10), 850–863. doi:10.1016/j.tim.2019.05.003

Crabbé, A., Ledesma, M. A., and Nickerson, C. A. (2014). Mimicking the host and its microenvironment *in vitro* for studying mucosal infections by *Pseudomonas aeruginosa*. *Pathog Dis* 71 (1), 1–19. doi:10.1111/2049-632X.12180

Crabbé, A., Ostyn, L., Staelens, S., Rigauts, C., Risseeuw, M., Dhaenens, M., et al. (2019). Host metabolites stimulate the bacterial proton motive force to enhance the activity of aminoglycoside antibiotics. *PLoS Pathog.* 15 (4), e1007697. doi:10.1371/journal.ppat.1007697

Curran, D. M., Grote, A., Nursimulu, N., Geber, A., Voronin, D., Jones, D. R., et al. (2020). Modeling the metabolic interplay between a parasitic worm and its bacterial endosymbiont allows the identification of novel drug targets. *eLife* 9, e51850. doi:10.7554/eLife.51850

Darch, S. E., Simoska, O., Fitzpatrick, M., Barraza, J. P., Stevenson, K. J., Bonnecaze, R. T., et al. (2018). Spatial determinants of quorum signaling in a *Pseudomonas aeruginosa* infection model. *Proc. Natl. Acad. Sci. U. S. A.* 115 (18), 4779–4784. doi:10.1073/pnas.1719317115

De Oliveira, D. M. P., Forde, B. M., Kidd, T. J., Harris, P. N. A., Schembri, M. A., Beatson, S. A., et al. (2020). Antimicrobial resistance in ESKAPE pathogens. *Clin. Microbiol. Rev.* 33 (3). doi:10.1128/CMR.00181-19

den Besten, G., van Eunen, K., Groen, A. K., Venema, K., Reijngoud, D. J., and Bakker, B. M. (2013). The role of short-chain fatty acids in the interplay between diet, gut microbiota, and host energy metabolism. *J. Lipid Res.* 54 (9), 2325–2340. doi:10.1194/jlr.R036012

Dowd, S. E., Wolcott, R. D., Sun, Y., McKeehan, T., Smith, E., and Rhoads, D. (2008). Polymicrobial nature of chronic diabetic foot ulcer biofilm infections determined using bacterial tag encoded FLX amplicon pyrosequencing (bTEFAP). *PloS One* 3 (10), e3326. doi:10.1371/journal.pone.0003326

Dumas, M. E. (2011). The microbial-mammalian metabolic axis: beyond simple metabolism. *Cell Metabol.* 13 (5), 489–490. doi:10.1016/j.cmet.2011.04.005

Edwards, J. S., and Palsson, B. O. (1999). Systems properties of the Haemophilus influenzae Rd metabolic genotype. *J. Biol. Chem.* 274 (25), 17410–17416. doi:10.1074/jbc.274.25.17410

Eisenreich, W., Heesemann, J., Rudel, T., and Goebel, W. (2015). Metabolic adaptations of intracellular bacterial pathogens and their mammalian host cells during infection ("pathometabolism"). *Microbiol. Spectr.* 3 (3), MBP-0002-2014. doi:10.1128/microbiolspec.MBP-0002-2014

Eisenreich, W., Rudel, T., Heesemann, J., and Goebel, W. (2017). To eat and to be eaten: mutual metabolic adaptations of immune cells and intracellular bacterial pathogens upon infection. *Front Cell Infect Microbiol* 7, 316. doi:10.3389/fcimb.2017.00316

Ferreyra, J. A., Wu, K. J., Hryckowian, A. J., Bouley, D. M., Weimer, B. C., and Sonnenburg, J. L. (2014). Gut microbiota-produced succinate promotes *c. difficile* infection after antibiotic treatment or motility disturbance. *Cell Host Microbe* 16 (6), 770–777. doi:10.1016/j.chom.2014.11.003

Filkins, L. M., Graber, J. A., Olson, D. G., Dolben, E. L., Lynd, L. R., Bhuju, S., et al. (2015). Coculture of *Staphylococcus aureus* with *Pseudomonas aeruginosa* drives *S. aureus* towards fermentative metabolism and reduced viability in a cystic fibrosis model. *J. Bacteriol.* 197 (14), 2252–2264. doi:10.1128/JB.00059-15

Filkins, L. M., Hampton, T. H., Gifford, A. H., Gross, M. J., Hogan, D. A., Sogin, M. L., et al. (2012). Prevalence of streptococci and increased polymicrobial diversity associated with cystic fibrosis patient stability. *J. Bacteriol.* 194 (17), 4709–4717. doi:10.1128/JB.00566-12

Flynn, J. M., Cameron, L. C., Wiggen, T. D., Dunitz, J. M., Harcombe, W. R., and Hunter, R. C. (2020). Disruption of cross-feeding inhibits pathogen growth in the sputa of patients with cystic fibrosis. *mSphere* 5 (2), e00343–20. doi:10.1128/mSphere.00343-20

Flynn, J. M., Niccum, D., Dunitz, J. M., and Hunter, R. C. (2016). Evidence and role for bacterial mucin degradation in cystic fibrosis airway disease. *PLoS Pathog.* 12 (8), e1005846. doi:10.1371/journal.ppat.1005846

Franzosa, E. A., Sirota-Madi, A., Avila-Pacheco, J., Fornelos, N., Haiser, H. J., Reinker, S., et al. (2019). Gut microbiome structure and metabolic activity in inflammatory bowel disease. *Nat Microbiol* 4 (2), 293–305. doi:10.1038/s41564-018-0306-4

Frioux, C., Singh, D., Korcsmaros, T., and Hildebrand, F. (2020). From bag-of-genes to bag-of-genomes: metabolic modelling of communities in the era of metagenome-assembled genomes. *Comput. Struct. Biotechnol. J.* 18, 1722–1734. doi:10.1016/j.csbj.2020.06.028

Furuya-Kanamori, L., Marquess, J., Yakob, L., Riley, T. V., Paterson, D. L., Foster, N. F., et al. (2015). Asymptomatic *Clostridium difficile* colonization: epidemiology and clinical implications. *BMC Infect. Dis.* 15, 516. doi:10.1186/s12879-015-1258-4

Görke, B., and Stülke, J. (2008). Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. *Nat. Rev. Microbiol.* 6 (8), 613–624. doi:10.1038/nrmicro1932

Gabryszewski, S. J., Wong Fok Lung, T., Annavajhala, M. K., Tomlinson, K. L., Riquelme, S. A., Khan, I. N., et al. (2019). Metabolic adaptation in methicillin-resistant *Staphylococcus aureus* pneumonia. *Am. J. Respir. Cell Mol. Biol.* 61 (2), 185–197. doi:10.1165/rcmb.2018-0389OC

Gilbert, J. A., Quinn, R. A., Debelius, J., Xu, Z. Z., Morton, J., Garg, N., et al. (2016). Microbiome-wide association studies link dynamic microbial consortia to disease. *Nature* 535 (7610), 94–103. doi:10.1038/nature18850

Gu, C., Kim, G. B., Kim, W. J., Kim, H. U., and Lee, S. Y. (2019). Current status and applications of genome-scale metabolic models. *Genome Biol.* 20 (1), 121. doi:10.1186/s13059-019-1730-3

Hamer, H. M., Jonkers, D., Venema, K., Vanhoutvin, S., Troost, F. J., and Brummer, R. J. (2008). Review article: the role of butyrate on colonic function. *Aliment. Pharmacol. Ther.* 27 (2), 104–119. doi:10.1111/j.1365-2036.2007.03562.x

Harcombe, W. R., Riehl, W. J., Dukovski, I., Granger, B. R., Betts, A., Lang, A. H., et al. (2014). Metabolic resource allocation in individual microbes determines ecosystem interactions and spatial dynamics. *Cell Rep.* 7 (4), 1104–1115. doi:10.1016/j.celrep.2014.03.070

Henson, M. A., Orazi, G., Phalak, P., and O'Toole, G. A. (2019). Metabolic modeling of cystic fibrosis airway communities predicts mechanisms of pathogen dominance. *mSystems* 4 (2), e00026-00019. doi:10.1128/mSystems.00026-19

Henson, M. A., and Phalak, P. (2018). Suboptimal community growth mediated through metabolite crossfeeding promotes species diversity in the gut microbiota. *PLoS Comput. Biol.* 14 (10), e1006558. doi:10.1371/journal.pcbi.1006558

Heravi, F. S., Zakrzewski, M., Vickery, K., Malone, M., and Hu, H. (2020). Metatranscriptomic analysis reveals active bacterial communities in diabetic foot infections. *Front. Microbiol.* 11, 1688. doi:10.3389/fmicb.2020.01688

Herrero-Fresno, A., and Olsen, J. E. (2018). *Salmonella Typhimurium* metabolism affects virulence in the host - a mini-review. *Food Microbiol.* 71, 98–110. doi:10.1016/j.fm.2017.04.016

Holmes, E., Li, J. V., Marchesi, J. R., and Nicholson, J. K. (2012). Gut microbiota composition and activity in relation to host metabolic phenotype and disease risk. *Cell Metabol.* 16 (5), 559–564. doi:10.1016/j.cmet.2012.10.007

Ji, B., and Nielsen, J. (2015). From next-generation sequencing to systematic modeling of the gut microbiome. *Front. Genet.* 6, 219. doi:10.3389/fgene.2015.00219

Jorth, P., Turner, K. H., Gumus, P., Nizam, N., Buduneli, N., and Whiteley, M. (2014). Metatranscriptomics of the human oral microbiome during health and disease. *mBio* 5 (2), e01012. doi:10.1128/mBio.01012-14

Kaakoush, N. O., Day, A. S., Huinao, K. D., Leach, S. T., Lemberg, D. A., Dowd, S. E., et al. (2012). Microbial dysbiosis in pediatric patients with Crohn's disease. *J. Clin. Microbiol.* 50 (10), 3258–3266. doi:10.1128/JCM.01396-12

Kinross, J. M., Darzi, A. W., and Nicholson, J. K. (2011). Gut microbiome-host interactions in health and disease. *Genome Med.* 3 (3), 14. doi:10.1186/gm228

Koh, A., De Vadder, F., Kovatcheva-Datchary, P., and Bäckhed, F. (2016). From dietary fiber to host physiology: short-chain fatty acids as key bacterial metabolites. *Cell* 165 (6), 1332–1345. doi:10.1016/j.cell.2016.05.041

Kumar, M., Ji, B., Zengler, K., and Nielsen, J. (2019). Modelling approaches for studying the microbiome. *Nat Microbiol* 4 (8), 1253–1267. doi:10.1038/s41564-019-0491-9

Lagier, J.-C. (2016). Gut microbiota and *Clostridium difficile* infections. *Human Microbiome Journal* 2, 10–14. doi:10.1016/j.humic.2016.10.003

Lavelle, A., and Sokol, H. (2020). Gut microbiota-derived metabolites as key actors in inflammatory bowel disease. *Nat. Rev. Gastroenterol. Hepatol.* 17 (4), 223–237. doi:10.1038/s41575-019-0258-z

Lebeaux, D., Ghigo, J. M., and Beloin, C. (2014). Biofilm-related infections: bridging the gap between clinical management and fundamental aspects of

recalcitrance toward antibiotics. *Microbiol. Mol. Biol. Rev.* 78 (3), 510–543. doi:10.1128/MMBR.00013-14

Leewenhoeck, A. (1684). An abstract of a letter from Mr. Anthony Leevvenhoeck at Delft, dated Sep. 17. 1683. Containing some microscopical observations, about animals in the scurf of the teeth, the substance call'd worms in the nose, the cuticula consisting of scales. *Phil. Trans. Roy. Soc. Lond.* 14 (159), 568–574. doi:10.1098/rstl.1684.0030

Leslie, J. L., Vendrov, K. C., Jenior, M. L., and Young, V. B. (2019). The gut microbiota is associated with clearance of *Clostridium difficile* Infection independent of adaptive immunity. *mSphere* 4 (1), e00698-00618. doi:10.1128/mSphereDirect.00698-18

Lessa, F. C., Winston, L. G., and McDonald, L. C. (2015). Burden of *Clostridium difficile* infection in the United States. *N. Engl. J. Med.* 372 (24), 2369–2370. doi:10.1056/NEJMc1505190

Li, K., Gifford, A. H., Hampton, T. H., and O'Toole, G. A. (2020). Availability of Zinc impacts interactions between *Streptococcus sanguinis* and *Pseudomonas aeruginosa* in coculture. *J. Bacteriol.* 202 (2), e00618–00619. doi:10.1128/JB.00618-19

Lieven, C., Beber, M. E., Olivier, B. G., Bergmann, F. T., Ataman, M., Babaei, P., et al. (2020). MEMOTE for standardized genome-scale metabolic model testing. *Nat. Biotechnol.* 38 (3), 272–276. doi:10.1038/s41587-020-0446-y

Limoli, D. H., and Hoffman, L. R. (2019). Help, hinder, hide and harm: what can we learn from the interactions between *Pseudomonas aeruginosa* and *Staphylococcus aureus* during respiratory infections? *Thorax* 74 (7), 684. doi:10.1136/thoraxjnl-2018-212616

Limoli, D. H., Yang, J., Khansaheb, M. K., Helfman, B., Peng, L., Stecenko, A. A., et al. (2016). *Staphylococcus aureus* and *Pseudomonas aeruginosa* co-infection is associated with cystic fibrosis-related diabetes and poor clinical outcomes. *Eur. J. Clin. Microbiol. Infect. Dis.* 35 (6), 947–953. doi:10.1007/s10096-016-2621-0

Lipuma, J. J. (2010). The changing microbial epidemiology in cystic fibrosis. *Clin. Microbiol. Rev.* 23 (2), 299–323. doi:10.1128/CMR.00068-09

Lopez-Agudelo, V. A., Mendum, T. A., Laing, E., Wu, H., Baena, A., Barrera, L. F., et al. (2020). A systematic evaluation of *Mycobacterium tuberculosis* genome-scale metabolic networks. *PLoS Comput. Biol.* 16 (6), e1007533. doi:10.1371/journal.pcbi.1007533

Lupetti, A., Danesi, R., Campa, M., Del Tacca, M., and Kelly, S. (2002). Molecular basis of resistance to azole antifungals. *Trends Mol. Med.* 8 (2), 76–81. doi:10.1016/s1471-4914(02)02280-3

Magnúsdóttir, S., Heinken, A., Kutt, L., Ravcheev, D. A., Bauer, E., Noronha, A., et al. (2017). Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota. *Nat. Biotechnol.* 35 (1), 81–89. doi:10.1038/nbt.3703

Mahowald, M. A., Rey, F. E., Seedorf, H., Turnbaugh, P. J., Fulton, R. S., Wollam, A., et al. (2009). Characterizing a model human gut microbiota composed of members of its two dominant bacterial phyla. *Proc. Natl. Acad. Sci. U.S.A.* 106 (14), 5859–5864. doi:10.1073/pnas.0901529106

Miller, D. P., Fitzsimonds, Z. R., and Lamont, R. J. (2019). Metabolic Signaling and spatial interactions in the oral polymicrobial community. *J. Dent. Res.* 98 (12), 1308–1314. doi:10.1177/0022034519866440

Molyneux, D. H., Hopkins, A., Bradley, M. H., and Kelly-Hope, L. A. (2014). Multidimensional complexities of filariasis control in an era of large-scale mass drug administration programmes: a can of worms. *Parasites Vectors* 7 (1), 363. doi:10.1186/1756-3305-7-363

Noronha, A., Modamio, J., Jarosz, Y., Guerard, E., Sompairac, N., Preciat, G., et al. (2019). The Virtual Metabolic Human database: integrating human and gut microbiome metabolism with nutrition and disease. *Nucleic Acids Res.* 47, D614–D624. doi:10.1093/nar/gky992

O'Meara, T. R., Veri, A. O., Ketela, T., Jiang, B., Roemer, T., and Cowen, L. E. (2015). Global analysis of fungal morphology exposes mechanisms of host cell escape. *Nat. Commun.* 6, 6741. doi:10.1038/ncomms7741

O'Sullivan, B. P., and Freedman, S. D. (2009). Cystic fibrosis. *Lancet* 373 (9678), 1891–1904. doi:10.1016/S0140-6736(09)60327-5

O'Toole, G. A. (2016). Classic Spotlight: before they were biofilms. *J. Bacteriol.* 198 (1), 5. doi:10.1128/JB.00593-15

O'Toole, G. A. (2018). Cystic fibrosis airway microbiome: overturning the old, opening the way for the new. *J. Bacteriol.* 200 (4), e00561–00517. doi:10.1128/JB.00561-17

Orazi, G., Jean-Pierre, F., and O'Toole, G. A. (2020). *Pseudomonas aeruginosa* PA14 enhances the efficacy of norfloxacin against *Staphylococcus aureus* Newman biofilms. *J. Bacteriol.* 202 (18), e00159–00120. doi:10.1128/JB.00159-20

Orazi, G., and O'Toole, G. A. (2019). It takes a village: mechanisms underlying antimicrobial recalcitrance of polymicrobial biofilms. *J. Bacteriol.* 202 (1), e00530–00519. doi:10.1128/JB.00530-19

Orazi, G., and O'Toole, G. A. (2017). *Pseudomonas aeruginosa* alters *Staphylococcus aureus* sensitivity to vancomycin in a biofilm model of cystic fibrosis infection. *mBio* 8 (4), e00873–00817. doi:10.1128/mBio.00873-17

Orazi, G., Ruoff, K. L., and O'Toole, G. A. (2019). *Pseudomonas aeruginosa* increases the sensitivity of biofilm-grown *Staphylococcus aureus* to membrane-targeting antiseptics and antibiotics. *mBio* 10 (4), 1501–1519. doi:10.1128/mBio.01501-19

Palmer, K. L., Aye, L. M., and Whiteley, M. (2007). Nutritional cues control *Pseudomonas aeruginosa* multicellular behavior in cystic fibrosis sputum. *J. Bacteriol.* 189 (22), 8079–8087. doi:10.1128/JB.01138-07

Palmer, K. L., Mashburn, L. M., Singh, P. K., and Whiteley, M. (2005). Cystic fibrosis sputum supports growth and cues key aspects of *Pseudomonas aeruginosa* physiology. *J. Bacteriol.* 187 (15), 5267–5277. doi:10.1128/JB.187.15.5267-5277.2005

Peng, B., Su, Y. B., Li, H., Han, Y., Guo, C., Tian, Y. M., et al. (2015). Exogenous alanine and/or glucose plus kanamycin kills antibiotic-resistant bacteria. *Cell Metabol.* 21 (2), 249–261. doi:10.1016/j.cmet.2015.01.008

Perinbam, K., Chacko, J. V., Kannan, A., Digman, M. A., and Siryaporn, A. (2020). A shift in central metabolism accompanies virulence activation in *Pseudomonas aeruginosa*. *mBio* 11 (2). doi:10.1128/mBio.02730-18

Perlroth, J., Choi, B., and Spellberg, B. (2007). Nosocomial fungal infections: epidemiology, diagnosis, and treatment. *Med. Mycol.* 45 (4), 321–346. doi:10.1080/13693780701218689

Peters, B. M., Jabra-Rizk, M. A., O'May, G. A., Costerton, J. W., and Shirtliff, M. E. (2012). Polymicrobial interactions: impact on pathogenesis and human disease. *Clin. Microbiol. Rev.* 25 (1), 193–213. doi:10.1128/CMR.00013-11

Phalak, P., Chen, J., Carlson, R. P., and Henson, M. A. (2016). Metabolic modeling of a chronic wound biofilm consortium predicts spatial partitioning of bacterial species. *BMC Syst. Biol.* 10 (1), 90. doi:10.1186/s12918-016-0334-8

Phalak, P., and Henson, M. A. (2019). Metabolic modeling of *Clostridium difficile* associated dysbiosis of the gut microbiota. *Processes* 7 (2) 97. doi:10.3390/pr7020097

Presta, L., Bosi, E., Mansouri, L., Dijkshoorn, L., Fani, R., and Fondi, M. (2017). Constraint-based modeling identifies new putative targets to fight colistin-resistant *A. baumannii* infections. *Sci. Rep.* 7 (1), 3706. doi:10.1038/s41598-017-03416-2

Price, K. E., Hampton, T. H., Gifford, A. H., Dolben, E. L., Hogan, D. A., Morrison, H. G., et al. (2013). Unique microbial communities persist in individual cystic fibrosis patients throughout a clinical exacerbation. *Microbiome* 1 (1), 27. doi:10.1186/2049-2618-1-27

Quince, C., Walker, A. W., Simpson, J. T., Loman, N. J., and Segata, N. (2017). Shotgun metagenomics, from sampling to analysis. *Nat. Biotechnol.* 35 (9), 833–844. doi:10.1038/nbt.3935

Rai, A., and Saito, K. (2016). Omics data input for metabolic modeling. *Curr. Opin. Biotechnol.* 37, 127–134. doi:10.1016/j.copbio.2015.10.010

Raman, K., Pratapa, A., Mohite, O., and Balachandran, S. (2018). "Computational prediction of synthetic lethals in genome-scale metabolic models using Fast-SL," in *Metabolic network reconstruction and modeling: methods and protocols*. Editor M. Fondi (New York, NY: Springer), 315–336. doi:10.1007/978-1-4939-7528-0_14

Ramos, P. I. P., Fernández Do Porto, D., Lanzarotti, E., Sosa, E. J., Burguener, G., Pardo, A. M., et al. (2018). An integrative, multi-omics approach towards the prioritization of *Klebsiella pneumoniae* drug targets. *Sci. Rep.* 8 (1), 10755. doi:10.1038/s41598-018-28916-7

Rams, T. E., Degener, J. E., and van Winkelhoff, A. J. (2014). Antibiotic resistance in human chronic periodontitis microbiota. *J. Periodontol.* 85 (1), 160–169. doi:10.1902/jop.2013.130142

Rhodes, J. M. (2007). The role of *Escherichia coli* in inflammatory bowel disease. *Gut* 56 (5), 610–612. doi:10.1136/gut.2006.111872

Riquelme, S. A., Liimatta, K., Wong Fok Lung, T., Fields, B., Ahn, D., Chen, D., et al. (2020a). *Pseudomonas aeruginosa* Utilizes host-derived itaconate to

redirect its metabolism to promote biofilm formation. *Cell Metabol.* 31 (6), 1091–e1096. doi:10.1016/j.cmet.2020.04.017

Riquelme, S. A., Lozano, C., Moustafa, A. M., Liimatta, K., Tomlinson, K. L., Britto, C., et al. (2019). CFTR-PTEN-dependent mitochondrial metabolic dysfunction promotes *Pseudomonas aeruginosa* airway infection. *Sci. Transl. Med.* 11 (499), 4634. doi:10.1126/scitranslmed.aav4634

Riquelme, S. A., Wong Fok Lung, T., and Prince, A. (2020b). Pulmonary pathogens adapt to immune signaling metabolites in the airway. *Front. Immunol.* 11, 385. doi:10.3389/fimmu.2020.00385

Saleh, M., and Elson, C. O. (2011). Experimental inflammatory bowel disease: insights into the host-microbiota dialog. *Immunity* 34 (3), 293–302. doi:10.1016/j.immuni.2011.03.008

Scott, J. E., and O'Toole, G. A. (2019). The Yin and Yang of *Streptococcus* lung infections in cystic fibrosis: a Model for studying polymicrobial interactions. *J. Bacteriol.* 201 (11). doi:10.1128/JB.00115-19

Scott, J. E., Li, K., Filkins, L. M., Zhu, B., Kuchma, S. L., Schwartzman, J. D., et al. (2019). *Pseudomonas aeruginosa* can inhibit growth of streptococcal species *via* siderophore production. *J. Bacteriol.* 201 (8), e00014–00019. doi:10.1128/JB.00014-19

Seaver, S. M. D., Liu, F., Zhang, Q., Jeffryes, J., Faria, J. P., Edirisinghe, J. N., et al. (2020). The ModelSEED Biochemistry Database for the integration of metabolic annotations and the reconstruction, comparison and analysis of metabolic models for plants, fungi and microbes. *Nucleic Acids Res.* 49. 575–588. doi:10.1093/nar/gkaa746

Sen, P., and Orešič, M. (2019). Metabolic modeling of human gut microbiota on a genome scale: an overview. *Metabolites* 9 (2). doi:10.3390/metabo9020022

Shoaie, S., Ghaffari, P., Kovatcheva-Datchary, P., Mardinoglu, A., Sen, P., Pujos-Guillot, E., et al. (2015). Quantifying diet-induced metabolic changes of the human gut microbiome. *Cell Metabol.* 22 (2), 320–331. doi:10.1016/j.cmet.2015.07.001

Shoaie, S., and Nielsen, J. (2014). Elucidating the interactions between the human gut microbiota and its host through metabolic modeling. *Front. Genet.* 5, 86. doi:10.3389/fgene.2014.00086

Simeonidis, E., and Price, N. D. (2015). Genome-scale modeling for metabolic engineering. *J. Ind. Microbiol. Biotechnol.* 42 (3), 327–338. doi:10.1007/s10295-014-1576-3

Sprenger, M., Kasper, L., Hensel, M., and Hube, B. (2018). Metabolic adaptation of intracellular bacteria and fungi to macrophages. *Int J Med Microbiol* 308 (1), 215–227. doi:10.1016/j.ijmm.2017.11.001

Surette, M. G. (2014). The cystic fibrosis lung microbiome. *Ann Am Thorac Soc* 11, S61–S65. doi:10.1513/AnnalsATS.201306-159MG

Tavernier, S., Crabbé, A., Hacioglu, M., Stuer, L., Henry, S., Rigole, P., et al. (2017). Community composition determines activity of antibiotics against multispecies biofilms. *Antimicrob. Agents Chemother.* 61 (9). doi:10.1128/AAC.00302-17

Tavernier, S., Sass, A., De Bruyne, M., Baeke, F., De Rycke, R., Crabbé, A., et al. (2018). Decreased susceptibility of *Streptococcus anginosus* to vancomycin in a multispecies biofilm is due to increased thickness of the cell wall. *J. Antimicrob. Chemother.* 73 (9), 2323–2330. doi:10.1093/jac/dky216

Taylor, M. J., Voronin, D., Johnston, K. L., and Ford, L. (2013). *Wolbachia* filarial interactions. *Cell Microbiol.* 15 (4), 520–526. doi:10.1111/cmi.12084

Thiele, I., Sahoo, S., Heinken, A., Hertel, J., Heirendt, L., Aurich, M. K., et al. (2020). Personalized whole-body models integrate metabolism, physiology, and the gut microbiome. *Mol. Syst. Biol.* 16 (5), e8982. doi:10.15252/msb.20198982

Thiele, I., Swainston, N., Fleming, R. M., Hoppe, A., Sahoo, S., Aurich, M. K., et al. (2013). A community-driven global reconstruction of human metabolism. *Nat. Biotechnol.* 31 (5), 419–425. doi:10.1038/nbt.2488

Thornton, C. S., and Surette, M. G. (2020). Potential contributions of anaerobes in cystic fibrosis airways. *J. of Clinical Microbiology, JCM.* 01813–01819. doi:10.1128/JCM.01813-19

Turner, K. H., Wessel, A. K., Palmer, G. C., Murray, J. L., and Whiteley, M. (2015). Essential genome of *Pseudomonas aeruginosa* in cystic fibrosis sputum. *Proc. Natl. Acad. Sci. U.S.A.* 112 (13), 4110–4115. doi:10.1073/pnas.1419677112

Vandeplassche, E., Tavernier, S., Coenye, T., and Crabbé, A. (2019). Influence of the lung microbiome on antibiotic susceptibility of cystic fibrosis pathogens. *Eur. Respir. Rev.* 28 (152), 190041. doi:10.1183/16000617.0041-2019

Viana, R., Dias, O., Lagoa, D., Galocha, M., Rocha, I., and Teixeira, M. C. (2020). Genome-scale metabolic model of the human pathogen *Candida albicans*: a promising platform for drug target prediction. *Journal of Fungi* 6 (3), 171. doi:10.3390/jof6030171

Zampieri, G., Vijayakumar, S., Yaneske, E., and Angione, C. (2019). Machine and deep learning meet genome-scale metabolic modeling. *PLoS Comput. Biol.* 15 (7), e1007084, doi:10.1371/journal.pcbi.1007084

Zhu, Y., Czauderna, T., Zhao, J., Klapperstueck, M., Maifiah, M. H. M., Han, M. L., et al. (2018). Genome-scale metabolic modeling of responses to polymyxins in *Pseudomonas aeruginosa*. *GigaScience* 7 (4), giy021. doi:10.1093/gigascience/giy021

Zomorrodi, A. R., and Segrè, D. (2016). Synthetic ecology of microbes: mathematical models and applications. *J. Mol. Biol.* 428, 837–861. doi:10.1016/j.jmb.2015.10.019

Zuñiga, C., Zaramela, L., and Zengler, K. (2017). Elucidation of complexity and prediction of interactions in microbial communities. *Microb Biotechnol* 10 (6), 1500–1522. doi:10.1111/1751-7915.12855

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Jean-Pierre, Henson and O'Toole. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 22 February 2021
doi: 10.3389/fmolb.2021.624061

Genome-Wide Analysis of Targets for Post-Transcriptional Regulation by Rsm Proteins in *Pseudomonas putida*

*Óscar Huertas-Rosales[1], Manuel Romero[2], Kok-Gan Chan[3,4], Kar-Wai Hong[3,4], Miguel Cámara[2], Stephan Heeb[2], Laura Barrientos-Moreno[1,2], María Antonia Molina-Henares[1], María L. Travieso[1], María Isabel Ramos-González[1] and Manuel Espinosa-Urgel[1]**

[1]Department of Environmental Protection, Estación Experimental del Zaidín, CSIC, Granada, Spain, [2]National Biofilms Innovation Centre, Biodiscovery Institute and School of Life Sciences, University of Nottingham, Nottingham, United Kingdom, [3]Division of Genetics and Molecular Biology, Institute of Biological Sciences, Faculty of Science, University of Malaya, Kuala Lumpur, Malaysia, [4]International Genome Centre, Jiangsu University, Zhenjiang, China

Post-transcriptional regulation is an important step in the control of bacterial gene expression in response to environmental and cellular signals. *Pseudomonas putida* KT2440 harbors three known members of the CsrA/RsmA family of post-transcriptional regulators: RsmA, RsmE and RsmI. We have carried out a global analysis to identify RNA sequences bound *in vivo* by each of these proteins. Affinity purification and sequencing of RNA molecules associated with Rsm proteins were used to discover direct binding targets, corresponding to 437 unique RNA molecules, 75 of them being common to the three proteins. Relevant targets include genes encoding proteins involved in signal transduction and regulation, metabolism, transport and secretion, stress responses, and the turnover of the intracellular second messenger c-di-GMP. To our knowledge, this is the first combined global analysis in a bacterium harboring three Rsm homologs. It offers a broad overview of the network of processes subjected to this type of regulation and opens the way to define what are the sequence and structure determinants that define common or differential recognition of specific RNA molecules by these proteins.

Keywords: RNA-binding proteins, global regulation, biofilm, rhizosphere, amino acid metabolism, c-di-GMP signaling

OPEN ACCESS

Edited by:
Chew Chieng Yeo, Sultan Zainal Abidin University, Malaysia

Reviewed by:
Frederic Allain, ETH Zürich, Switzerland
Lydia Contreras, University of Texas at Austin, United States

****Correspondence:***
Manuel Espinosa-Urgel
manuel.espinosa@eez.csic.es

Specialty section:
This article was submitted to Molecular Recognition, a section of the journal Frontiers in Molecular Biosciences

Received: *30 October 2020*
Accepted: *21 January 2021*
Published: *22 February 2021*

Citation:
Huertas-Rosales Ó, Romero M, Chan K-G, Hong K-W, Cámara M, Heeb S, Barrientos-Moreno L, Molina-Henares MA, Travieso ML, Ramos-González MI and Espinosa-Urgel M (2021) Genome-Wide Analysis of Targets for Post-Transcriptional Regulation by Rsm Proteins in Pseudomonas putida. *Front. Mol. Biosci. 8:624061. doi: 10.3389/fmolb.2021.624061*

INTRODUCTION

By coordinating the expression of a large number of genes, global regulatory networks allow bacteria to adjust their physiology according to environmental stimuli, changes in their lifestyle, or in nutritional status (Ishihama, 2010; Coggan and Wolfgang, 2012; Shis et al., 2018). Transcriptional regulators and sigma factors influencing the expression of different sets of bacterial genes have been widely studied for decades, starting shortly after the postulates of Jacob and Monod on operons, operators and messenger RNA were put forward (Jacob and Monod, 1961). A second instance of protein expression regulation, namely post-transcriptional modulation affecting mRNA stability, structure or translation, mediated by proteins or small non-coding RNAs, has gained increasing attention in the past 2 decades, but is still less well studied (Romeo et al., 2013; Van Assche et al., 2015; Hör et al., 2018). Among the post-transcriptional regulators identified in prokaryotes, the CsrA/RsmA family of proteins seems to be widely conserved in many bacteria and, in some cases,

more than one member of this family are present in a single species (Reimmann et al., 2005; Morris et al., 2013; Huertas-Rosales et al., 2016; Ferreiro et al., 2018). These small RNA-binding proteins generally function as negative effectors of translation by binding to the 5′untranslated region of target mRNAs, close to or overlapping with the ribosome binding site (Baker et al., 2007; Yakhnin et al., 2007), or by causing premature transcription termination through alterations of the RNA structure that lead to the exposure of a Rho binding site (Figueroa-Bossi et al., 2014). However, they can also influence mRNA stability in a positive way, for example, by masking RNase E cleavage sites (Yakhnin et al., 2013). CsrA/RsmA proteins can interact with two RNA motifs, with a strong preference for 5′-RUACARGGAUGU-3′ consensus sequences located in the loops of short RNA hairpins (Dubey et al., 2005; Duss et al., 2014). Small non-coding RNAs (sRNA) containing multiple binding motifs play an opposing regulatory role by sequestering the CsrA/RsmA proteins, thus limiting their access to target mRNAs (Romeo, 1998; Kay et al., 2005; Sonnleitner and Haas, 2011; Janssen et al., 2018). In *Pseudomonas*, expression of these regulatory RNAs is controlled by the two-component system GacS/GacA (Brencic et al., 2009), in response to as yet not well-defined signal(s).

CsrA/RsmA proteins play a global role in modulating gene expression (Romeo, 1998; Lawhon et al., 2003; Brencic and Lory, 2009; Romeo et al., 2013; Romero et al., 2018). The functions identified as being under this type of regulation in different bacteria include carbohydrate metabolism and storage (Sabnis et al., 1995; Yang et al., 1996; Pannuri et al., 2016), synthesis of flagellar components (Yakhnin et al., 2007, Yakhnin et al., 2013), the production of secondary metabolites (Sonnleitner and Haas, 2011; Morris et al., 2013), quorum sensing signaling (Lenz et al., 2005), or the expression of virulence factors (Heroven et al., 2012; Sterzenbach et al., 2013; Vakulskas et al., 2015; Ferreiro et al., 2018). Global analyses have been done in bacteria harboring one CsrA/RsmA family protein to identify elements in the signaling and regulatory network associated to them (Lawhon et al., 2003; Sowa et al., 2017) or to othologous elements (Romero et al., 2018).

In the plant-root colonizing, beneficial bacterium *Pseudomonas putida* KT2440, three genes have been identified that encode post-transcriptional regulators belonging to the CsrA/RsmA family. These proteins (RsmA, RsmE and RsmI), have opposing effects on surface motility and biofilm formation; deletion of the three genes abolishes swarming motility and stimulates bacterial attachment and biofilm formation, although the biofilms formed by a triple *rsm* mutant are more labile and easily dispersed than wild type biofilms (Huertas-Rosales et al., 2016). These alterations are associated with changes in the expression of some components of the extracellular matrix of biofilms (Huertas-Rosales et al., 2016) and with increased levels of the intracellular second messenger cyclic diguanylate (c-di-GMP) (Huertas-Rosales et al., 2017). The three proteins were found to bind specific motifs in the leader sequence and translation initiation region of the mRNA of *cfcR* (Huertas-Rosales et al., 2017), which encodes a response regulator with diguanylate cyclase activity (Matilla et al., 2011; Ramos-González et al., 2016). Although the binding affinity was different for each Rsm protein, deletion of any single one of the three genes had no significant influence on expression of *cfcR* (Huertas-Rosales et al., 2017), indicating the existence of some functional redundancy between RsmA, RsmE and RsmI. Based on sequence similarity with related strains, the putative antagonistic sRNAs RsmY and RsmZ could also be identified in KT2440 (Huertas-Rosales et al., 2016). Still, little is known about the binding specificities of these proteins in *P. putida*.

To further understand the importance of these proteins in signal transduction and regulation of global gene expression in *P. putida*, we have used a high-throughput approach to identify RNA sequences bound by Rsm proteins. Our data indicate that a significant number of genes are susceptible of being modulated at the post-transcriptional level by these proteins, and support the existence of a certain degree of functional overlap between the three Rsm homologs. This approach has enabled us to gain new insights into the biological function of these post-transcriptional regulators in *P. putida*, including their role in some metabolic processes and bacterial fitness in the plant root environment.

MATERIALS AND METHODS

Bacterial Strains, Culture Media and Growth Conditions

The bacterial strains, plasmids and oligonucleotides used in this study are listed in **Table 1**. *Pseudomonas putida* KT2440 is a plasmid-free derivative of *P. putida* mt-2 (Regenhardt et al., 2002). *Pseudomonas putida* strains were grown at 30°C, in rich LB medium (Lennox, 1955), M9 or M8 defined medium (Sambrook and Russell, 2001) supplemented with 1 mM $MgSO_4$, 6 mg/L ammonium ferric citrate and trace metals as described previously (Yousef-Coronado et al., 2008). Unless otherwise indicated, glucose (20 mM) or sodium citrate (15 mM) were used as carbon sources. *Escherichia coli* strains were grown at 37°C in LB. When appropriate, antibiotics were added to the medium at the following final concentrations (μg/ml): ampicillin (Ap) 100; kanamycin (Km) 25; streptomycin (Sm) 50 (*E. coli*) or 100 (*P. putida*); (Gm) gentamycin 50; tetracycline (Tc) 10 or 20. Cell growth was followed by measuring optical density at 660 nm (OD_{660}).

DNA Techniques

Digestion with restriction enzymes, dephosphorylation, ligation and electrophoresis were carried out using standard methods (Ausubel et al., 1987; Sambrook and Russell, 2001) and following manufacturers' instructions. Plasmid DNA isolation and recovery of DNA fragments from agarose gels were carried out using QIAGEN miniprep and gel extraction kits, respectively. Competent cells were prepared using calcium chloride, and transformations were performed using standard protocols (Sambrook and Russell, 2001). Electrotransformation of freshly plated *Pseudomonas* cells was performed as previously described (Choi et al., 2006).

Triparental Conjugations

Transfer of plasmids from *E. coli* to *P. putida* strains was performed by triparental matings using as a helper for RP4

TABLE 1 | Bacterial strains, plasmids and oligonucleotides used in this work.

Strains	Genotype/relevant characteristics	References/source
P. putida		
KT2440	Wild-type, pWWO-free derivative of *P. putida* mt-2	PRCC[a]
ΔI	Null *rsmI* derivative of KT2440	10
ΔE	Null *rsmE* derivative of KT2440	10
ΔA	Null *rsmA* derivative of KT2440	10
ΔIE	Double null *rsmI, rsmE* derivative of KT2440	10
ΔIA	Double null *rsmI*, *rsmA* derivative of KT2440	10
ΔEA	Double null *rsmE*, *rsmA* derivative of KT2440	10
ΔIEA	Triple null *rsmI*, *rsmE*, *rsmA* derivative of KT2440	10
KT2440-miniTn7-Km	Km^R, miniTn7Km-tagged derivative of KT2440	This work
KT2440-miniTn7-Sm	Sm^R, miniTn7Sm-tagged derivative of KT2440	This work
ΔA-miniTn7-Sm	Sm^R, miniTn7Sm-tagged derivative of Δ*rsmA*	This work
ΔE-miniTn7-Sm	Sm^R, miniTn7Sm-tagged derivative of Δ*rsmE*	This work
ΔI-miniTn7-Sm	Sm^R, miniTn7Sm-tagged derivative of Δ*rsmI*	This work
ΔEA-miniTn7-Sm	Sm^R, miniTn7Sm-tagged derivative of Δ*rsmEA*	This work
ΔIA-miniTn7-Sm	Sm^R, miniTn7Sm-tagged derivative of Δ*rsmIA*	This work
ΔIE-miniTn7-Sm	Sm^R, miniTn7Sm-tagged derivative of Δ*rsmIE*	This work
ΔIEA-miniTn7-Sm	Sm^R, miniTn7Sm-tagged derivative of Δ*rsmIEA*	This work
E. coli		
AKN63 (pBK-miniTn7-ΩSm1)	Ap^R, Sm^R, strain for delivery of miniTn7Sm	42
Pir1 (pUC18R6KT-mini-Tn7Km)	Ap^R, Km^R, strain for delivery of miniTn7Km	Addgene
SM10 λ*pir* (pUX-BF13)	Ap^R, RP4 transfer functions and miniTn7 transposase helper	42
DH5α (pRK600)	Cm^R, helper for conjugation	PRCC
Plasmids		
pME6032-*rsmA*	Tc^R, derivative of pME6032 for expression of RsmA-His_6	10
pME6032-*rsmE*	Tc^R, derivative of pME6032 for expression of RsmE-His_6	10
pME6032-*rsmI*	Tc^R, derivative of pME6032 for expression of RsmI-His_6	10
Oligonucleotides	Sequence (5′→3′)b	Use
PT7*rpoS*Fw	TTTTCTGCAG**TAATACGACTCACTATAGG**CTCAAGCGCTGCCAGGGA	EMSA, *rpoS* amplification
P*rpoS*FTRv	AAAAAAAACCCCCCCCCTTTACTGAGAGCCATTG	EMSA, *rpoS* amplification
PT7*rsmY*Fw	TTGCGGCCGCTTTTTT**TAATACGACTCACTATAGG**GTTCTAAGATTGGATCCACTG	EMSA, *rsmY* amplification
P*rsmY*FTRv	AAAAGCGGCCGCAAAAAAAACCCCCCCCCGCCGAAGCGGGGTTTTCCAG	EMSA, *rsmY* amplification

[a]*Pseudomonas Reference Culture Collection (http://artemisa.eez.csic.es/prcc/).*

[b]*Restriction sites are underlined, inserted T7 polymerase promoter is indicated in bold and sequences used to hybridize with ATTO700-labelled DNA oligonucleotide are highlighted in gray.*

transfer functions *E. coli* (pRK600) or SM10 λ*pir* (pUX-BF13), the latter when miniTn7 transposase was required for intergenic site-specific insertion of miniTn7 derivatives near *glmS* (Koch et al., 2001), used to tag the wild type and each *rsm* mutant (**Table 1**) for rhizosphere assays (see below). For each strain, cells were collected from 0.5 ml of overnight LB cultures via centrifugation, then washed and suspended in 50 µL of fresh LB, and finally spotted on nitrocellulose filters (0.22 µm pore diameter) on LB-agar plates. After overnight incubation at 30°C, cells were scraped off from the mating filter and suspended in 2 ml of M9, and serial dilutions were plated on selective medium (M9 with citrate and the appropriate antibiotics) to select exconjugants and counter-select donor, helper, and recipient strains.

Purification of Total RNA and RNA from Rsm-RNA Complexes

Previously constructed derivatives of expression vector pME6032 harboring each *rsm* gene (Huertas-Rosales et al., 2016) were used to express His-tagged Rsm proteins in *P. putida* KT2440. Overnight cultures (10 ml) of wild type KT2440 harboring each construct were inoculated in 500 ml of LB medium. Three biological replicates were run in parallel. Cultures were incubated at 30°C under shaking until an OD_{660} of 0.8 was reached. At that point, expression of each His-tagged-Rsm protein was induced by the addition of IPTG to a final concentration of 0.5 mM. Cultures were allowed to grow for six additional hours. Aliquots of 1.5 ml were then harvested by centrifugation, instantly frozen with liquid nitrogen and stored at −80°C for total RNA purification. Cells from the remaining culture volume were also harvested and pellets were stored at −80°C until use. His-tagged Rsm-RNA complexes were isolated using Ni-NTA Fast Start purification kit (Qiagen). Three replicate extractions were done for each culture replica. Elution aliquots were analyzed by SDS-PAGE.

Total RNA and RNA from Rsm-RNA complex was extracted using RNA isolation kit (Macherey-Nagel) following the manufacturer's instructions. RNA samples were subsequently treated with RNase-free DNase I (Turbo DNA-free kit, Ambion) to remove DNA traces, as specified by the supplier. Total RNA quality was assessed using Agilent RNA 6000 Nano Kit (Agilent Technologies) in the Agilent 2100 Bioanalyzer. RNA concentration was measured using Qubit RNA BR assay kit (Life Technologies). 1 µg of RNA was used for rRNA depletion using Ribo-Zero rRNA Removal Kit (Illumina). One of the biological

replicates of RsmA and one of RsmI did not meet the required quality and quantity standards and were not used in further analysis.

Generation of c-DNA Libraries and Sequencing

The generation of cDNA libraries was carried out using NEBNext Ultra Directional RNA Library Prep kit for Illumina (NEB). Dual Index Primers Set one was used to generate bar-coded multiplex libraries (NEB). Library QC was performed using bioanalyser HS kit (Agilent biotechnologies). cDNA libraries were quantified using qPCR (Kapa Biosystems). Libraries were pooled at the desired concentrations, denatured and loaded for sequencing according to manufacturer's instructions. Sequencing was performed on the Illumina MiSeq Benchtop Sequencer to generate 2 × 75 bp reads. The number of reads obtained ensured a minimum of 76× coverage of the whole genome for control RNA and a minimum of 60× coverage for Rsm-bound RNA.

Bioinformatic Analysis

Filtered reads were aligned to reference genome *P. putida* KT2440 (GenBank; RefSeq NC_002947.3) with Bowtie v2 (Langmead et al., 2009). Alignment. sam file was analyzed using MACS v14 to identify and evaluate the significance of reads-enriched regions in the genome, the output being one file containing the peak chromosome coordinates, and one containing the genome coordinates, summit, *p*-value, fold_enrichment and false discovery rate (FDR) of each peak (Zhang et al., 2008). The average number of tags in the control samples after filtering was approximately 2,220,000 (RsmA), 1,420,000 (RsmE), and 1,654,000 (RsmI); the average number of tags in the Rsm-bound samples after filtering was approximately 1,302,000 (RsmA), 550,000 (RsmE), and 1,054,000 (RsmI). Only those peaks present in the three technical replicates were considered. Identity and annotation of the targets above the cut-off values were further validated by individually inspecting the corresponding chromosomal regions in the *Pseudomonas* genome database (www.pseudomonas.com; Winsor et al., 2016).

EMSA Analysis of *in vitro* RNA-Protein Binding

For purification of Rsm proteins, overnight cultures (10 ml) of *P. putida* KT2440 harboring plasmids pME6032-*rsmA*, pME6032-*rsmE*, and pME6032-*rsmI* (Huertas-Rosales et al., 2016) were used to inoculate 490 ml fresh LB medium supplied with Tc. Cultures were grown at 30°C and 200 rpm until reaching an OD_{660} of 0.8. At this point, IPTG (0.5 mM) was added to induce the expression of the proteins. After 6 h of further growth, cells were harvested by centrifugation and pellets subsequently stored at −80°C. Protein purification was carried out using QIAexpress Ni-NTA Fast Start Kit (Qiagen), following the manufacturers' instructions.

Electrophoretic mobility shift assays (EMSA) were carried out following a method described previously (Romero et al., 2018). DNA templates corresponding to the target gene sequences were amplified by PCR using primers that incorporated a T7 promoter at the 5′ end and a 17-nt extension at the 3′ end (**Table 1**). The purified PCR product was used for RNA synthesis *in vitro* using the MAXIscript T7 kit (Life Technologies). The RNA obtained was visualized by hybridization of an ATTO700-labeled DNA primer to the 3′ extension of the RNA. Rsm proteins were incubated with target gene RNA (5 or 10 nM) in 1 × binding buffer [25]. Binding in the absence or presence of unlabeled competitor RNA (100-fold excess) was carried out for 30 min at 30°C. Then Bromophenol Blue was added (0.01%, wt/vol) before electrophoresis on 6% (w/v) non-denaturing polyacrylamide TBE gel at 4°C. Imaging was performed using a 9201 Odyssey Imaging System (LI-COR Biosciences).

Growth with Different Carbon and Nitrogen Sources

BIOLOG plates (Biolog Inc. Hayward, CA, USA) were used for initial assessment of growth of KT2440 and the triple Δ*rsmIEA* mutant with different nitrogen and carbon sources. Cultures grown overnight at 30°C in M9 minimal medium with glucose as carbon source were used for inoculation in the plates at an initial OD_{660} of 0.05. Turbidity was measured at different times over 24 h in a Tecan Sunrise microplate reader. Further experiments were done in 96-well plates using M9 or M8 with the indicated carbon or nitrogen sources, at a final concentration of 5 mM. Growth was followed for 24 h at 3°C with continuous shaking in a Bioscreen apparatus C MBR equipped with a wide band filter (420–580 nm).

Competitive Root Colonization Assays

Surface sterilization, germination of corn seeds, and bacterial inoculation of the seedlings were performed as described previously (Ramos-González et al., 2013). Briefly, at least six two-days old seedlings were inoculated with a 1:1 mix (~5 × 10^5 CFU/strain) of KT2440Tn7-Km, as the wild type, and the wild type or mutant derivatives tagged with miniTn7Sm by triparental conjugation as described above. Inocula sizes were monitored by plating on LB-agar supplied with kanamycin or streptomycin. After 7 days, bacteria were recovered from the rhizosphere or the root tip as specified (Ramos-González et al., 2013) and the number of cells of each strain in the population was estimated by plating on LB-agar supplied with kanamycin or streptomycin. Data are presented as the index of colonization fitness (Ramos-González et al., 2013). SigmaStat software package (Systat software) was used for statistical analysis. The data were compared using Student's *t*-test for independent samples ($p < 0.05$).

Data-Availability

Raw and processed data files have been deposited in the Gene Expression Omnibus Database (www.ncbi.nlm.nih.gov/geo/) and are available under accession number GSE154204.

TABLE 2 | Summary of the high throughput data analysis of Rsm targets in *P. putida* KT2440.

Biological replica	total peaks	CI 95[a] fold enrichment (lower limit)	Cut-off value	CI 95 -10 × log (*P*val) (lower limit)	Cut-off value	peaks above cut-off values	peaks in all biological replicas	unique RNA targets[b]
RsmA1			2.15		130		243	241
RsmA2	595	2.33		135		266		
RsmA3	596	2.33		128		244		
RsmE1	908	4	4	180	170	327	270	261
RsmE2	851	4.22		150		320		
RsmE3	850	4.62		181		402		
RsmI1	557	2.68	2.5	153	145	244	209	206
RsmI2								
RsmI3	596	2.53		137		238		

Shadowed in gray the discarded replicas.

[a]*Confidence intervals of each data distribution, α = 0.05.*

[b]*After discarding redundancies where the analysis identifies more than one peak in a single RNA molecule.*

RESULTS AND DISCUSSION

Identification of RNAs Bound to Rsm Proteins

Affinity purification of RNA-protein complexes followed by sequencing analysis (RAP-Seq), was carried out to identify genes that could potentially be regulated at the post-transcriptional level by RsmA, RsmE or RsmI in *P. putida* KT2440. The methodology was similar to that previously described for genome-wide mapping of targets for RsmN in *Pseudomonas aeruginosa* (Romero et al., 2018) and is summarized in **Supplementary Figure S1**. Each recombinant His-tagged Rsm protein (RsmA-His_6, RsmE-His_6 and RsmI-His_6, previously described; Huertas-Rosales et al., 2016) was over-expressed in KT2440 as described in Materials and Methods. Proteins were purified by affinity chromatography and their associated target RNAs subsequently isolated. As control for transcription levels, total RNA was isolated from each culture in parallel. Total and Rsm-bound RNA were analyzed for purity, quantified and converted to cDNA for Illumina sequencing. One of the three biological replicates of RsmA and one of RsmI were below optimal quality and quantity and were not used in further analysis. Sequence reads from the cDNA libraries were mapped to the genomic sequence of *P. putida* KT2440 and analyzed to identify the regions corresponding to transcripts that were significantly enriched in the Rsm-bound RNA population with respect to the total RNA controls. Rsm-enriched RNAs that were not represented in the three technical replicates from each culture were not included in further analysis.

A first noticeable result was that the number of sequences corresponding to RsmE-bound transcripts overrepresented with respect to total RNA was much higher than those associated to RsmA or RsmI (**Table 2**). Data were grouped in intervals and histograms were built to analyze the distribution of fold-enrichment (FE) values and *p*-value (PV) data–shown as $-10\times\log_{10}(p\text{-value})$ for ease of representation–in each case (**Figure 1**). In all cases, the distribution was similar between biological replicates for each Rsm regulator. The distribution of FE values was similar for RsmA and RsmI, with slightly lower values in the former (**Table 2**). For RsmE, the distribution was different and the average values higher. These different values between Rsm proteins could reflect differences in expression of each construct. In fact, controls for each protein indicate that higher amount of RsmE than of the other two proteins is recovered after purification (**Supplementary Figure S2**).

The analysis of distributions and the confidence intervals calculated for each technical replicate were the basis to establish cut-off values for further analysis of Rsm targets (**Table 3**). Is should be noted that in this analysis we opted for rather strict parameters in order to take into account sequences strongly overrepresented in the Rsm-bound RNA population with respect to the total RNA controls, and also to minimize the number of potential false positives, at the expense of missing some RNA sequences that are actual targets of these proteins. Thus, the following cut-off values were established: PV > 130, FE ≥ 2.15 for RsmA; PV > 170, FE ≥ 4 for RsmE; PV > 145, FE ≥ 2.5 for RsmI. Targets for which either value was below the cut-off in one of the replicates were discarded.

Using these parameters, 241, 261, and 206 RNA sequences were identified as targets for RsmA, RsmE and RsmI, respectively (**Supplementary Table S1**), corresponding to 437 unique transcripts, with 75 targets being shared by all three Rsm proteins and between 36 and 45 shared by two of them (**Figure 2A**). Interestingly, around 40% of the RsmE and RsmA targets are exclusive for each protein, while only 22% of the RsmI targets are unique for this paralog. It should be noted that the cDNA libraries generated in this study were not strand-specific and therefore, did not allow distinguishing the DNA strand to which the transcript corresponds. Consequently, in some cases where divergently transcribed genes are adjacent in the genome, it is not straightforward to discern which of them is the actual target, although the length of overlap between the enriched sequences and each gene, and the location of the summit (i.e. the position of maximal overlap of the reads corresponding to one region) can indicate the most likely target. Despite this limitation, the above results indicate that at least 12% of the transcripts in *P. putida* KT2440 are bound *in vivo* by Rsm proteins under the conditions tested in this study. The data provide a broad overview of the regulon and potential functions of these riboregulators.

TABLE 3 | Annotation and functional classification of common targets for the three Rsm proteins in *P. putida* KT2440.

Locus	Gene	Annotation	Notes
Metabolism			
PP_0053		sulfide:quinone oxidoreductase	Sulfur metabolism
PP_0158	*gcdH*	Glutaryl-CoA dehydrogenase	Operon (PP_0159: family III CoA-transferase)
PP_0292	*hisA*	Phosphoribosylformimino-5-aminoimidazole carboxamide ribonucleotide isomerase	Operon *hisBHAF*; histidine synthesis
PP_0626	*ndh*	NADH dehydrogenase	
PP_0711	*ycaC-I*	Putative hydrolase	Isochorismatase family
PP_1032	*guaA*	Glutamine-hydrolyzing GMP synthase	Purine metabolism
PP_1073	*glpD*	Aerobic glycerol-3-phosphate dehydrogenase	
PP_2080	*gdhB*	NAD-specific glutamate dehydrogenase	
PP_2217		Enoyl-CoA hydratase	
PP_2437		Acyl-CoA dehydrogenase	
PP_2640		GNAT family acetyltransferase	
PP_2681	*pqqD-II*	Pyrroloquinoline quinone biosynthesis chaperone	
PP_4571	*cysK*	Cysteine synthase K	Cysteine biosynthesis
PP_5003	*phaA*	poly (3-hydroxyalkanoate) polymerase	Synthesis of carbon/energy reserve polymers
PP_5079	*aroK*	Shikimate kinase	Biosynthesis of aromatic amino acids
PP_5199	*ubiH*	2-Octaprenyl-6-methoxyphenyl hydroxylase	Ubiquinone biosynthesis
Protein synthesis, degradation and modification			
PP_1434	*era*	GTPase Era	Maturation of 16S rRNA and assembly of the 30S ribosomal subunit
PP_1443	*lon-I*	DNA-binding ATP-dependent protease	Might indirectly regulate the levels and activity of sRNAs through stability of hfq
PP_3620	*ycaO*	Ribosomal protein S12 methylthiotransferase accessory factor	Post-translational peptide modification
PP_4559	*def-II*	Peptide deformylase	Processing of nascent peptides
PP_5364	*clsA*	Cardiolipin synthase	
Transport and secretion			
PP_0907		RND family multidrug transporter	Antimicrobial resistance
PP_1015	*gtsA*	Mannose/glucose ABC transporter substrate-binding protein	
PP_2195		Periplasmic putrescine-binding protein	
PP_3089	*hcp1*	Type VI secretion system effector protein	Part of the K1-T6SS
PP_3099	*tssC1*	Type VI secretion system	Part of the K1-T6SS
PP_3108	*tke2*	Type VI secretion system effector protein	Part of the K1-T6SS
PP_4542		ABC transporter ATP-binding protein/permease	
Stress response			
PP_1210	*dps*	DNA-binding stress protein	
PP_3156		Universal stress protein family	
PP_3234		HSP20 family heat shock protein	Putative chaperone
PP_3312		HSP20 family heat shock protein	Putative chaperone
PP_4541	*ppiA*	Peptidyl-prolyl *cis*-trans isomerase A	Protein folding; stress response and biofilm in different bacteria
Signal transduction and regulation			
PP_0173		Transcriptional factor-like protein	Winged helix DNA binding domain
PP_0546		Sigma-54 dependent transcriptional regulator	Putative acetoin metabolism regulator
PP_0563		Two-component system respose regulator - GGDEF domain	c-di-GMP turnover
PP_1492	*wspE*	Two-component system sensor histidine kinase/response regulator (CheA/WspE)	Part of the *wsp* cluster (biofilm formation)
PP_3761	*cfcA*	Two-component system sensor histidine kinase/response regulator	CfcR phosphorylation and activation
PP_3765	*turB*	H-NS family protein (MvaT homolog)	Repressor of gene expression
PP_3832	*rsmE*	Post-transcriptional regulatory protein RsmE	
PP_4099	*uvrY*	Two-componenent system response regulator	Operon with *uvrC* (nucleotide excision repair)
Non-coding RNAs			
PP_mr05	*rsmY*	Non-coding RNA	
PP_mr44		Non-coding RNA	
PP_mr52		Non-coding RNA	

(Continued on following page)

TABLE 3 | (*Continued*) Annotation and functional classification of common targets for the three Rsm proteins in *P. putida* KT2440.

Locus	Gene	Annotation	Notes
DNA recombination and transposition			
PP_1813		comEA-like protein	DNA binding and recombination domain
PP_1865		Transposase ISPpu8 + intergenic region	
PP_2114		Transposase ISPpu8 + intergenic region	
intergenic		Putative ISPpu8 insertion site	Long palindromic region downstream PP_3547
PP_4318		Transposase ISPpu8 + intergenic region	
Cell envelope and appendages (LPS, EPS, pili)			
PP_0063	(*lpxL*)	Lipid a biosynthesis lauroyl acyltransferase	LPS synthesis
PP_2926	*udg*	UDP-glucose 6-dehydrogenase	may be involved in LPS synthesis
PP_3139		Group 1 family glycosyl transferase	Part of the *pea* operon - EPS synthesis
PP_4795		Lipoprotein	LptE family - LPS assembly
PP_4920		Lipoprotein	PdaC superfamily (polysaccharide deacetylase)
PP_5083	*pilM*	Type IV pili biogenesis protein PilM	Operon 5083-5079 (type IV pili + *aroK*)
Hypothetical/unknown function			
PP_5720		Pseudogene	
PP_0085		Hypothetical protein	YqjD/ElaB family (stress response?)
PP_1499		Hypothetical protein	
PP_1887		Hypothetical protein - YD repeat domain	Repeats may bind carbohydrates
PP_2219		Hypothetical protein	
PP_2345		Hypothetical protein	
PP_2396		Hypothetical protein	Periplasmic protein
PP_3007		Hypothetical protein	
PP_3010		Hypothetical protein	Putative lipoprotein
PP_3130		Hypothetical protein	Glycoside hydrolase family (EPS turnover?)
PP_3580		Hypothetical protein	
PP_3662		Hypothetical protein	Nucleotide 5′-monophosphate nucleosidase YgdH-like superfamily
PP_3901		Hypothetical protein	Predicted phage protein
PP_3909		Hypothetical protein	
PP_3963		Hypothetical protein	Stress-induced protein (KGG repeat) domain
PP_4793		Hypothetical protein	
PP_5099		Hypothetical protein	
PP_5209		Hypothetical protein	FliL-like protein
PP_5232		Hypothetical protein	
PP_5395		Hypothetical protein	Branched-chain polyamine synthase domain

Analysis of RsmA, RsmE and RsmI Targets

The 75 targets common to the three Rsm proteins are compiled in **Table 3**, broadly classified according to their functions. As reflected in **Figure 2B**, about half of the common targets correspond to two categories: metabolism-related and hypothetical proteins. The fact that functions related to central metabolism and carbon storage are among these common targets is not unexpected, since carbon metabolism was at the origins of the identification of the CsrA/Rsm family of proteins (Sabnis et al., 1995). Other expected elements include RNAs corresponding to Rsm proteins themselves: *rsmE* is a target for the three proteins, and *rsmA* is recognized by RsmE, confirming previous expression data that indicated the existence of self- and cross-regulation of these proteins (Huertas-Rosales et al., 2016). Also expected was the small non-coding RNA *rsmY*, known to bind and titrate Rsm proteins (Sonnleitner and Haas, 2011; Janssen et al., 2018). Binding of *rsmY* to the three Rsm proteins could be confirmed *in vitro* by EMSA analysis (**Figure 3**), serving as positive control that the high throughput methodology used was successful, although the affinity appears to be different in each case, being RsmI the protein that required higher concentrations for binding to be detected. A second small RNA previously identified in *P. putida*, *rsmZ* (Huertas-Rosales et al., 2016), involved in titration of Rsm proteins in other bacteria (Janssen et al., 2018), is among the targets common to RsmA and RsmE (**Table 4**, **Supplementary Table S1**). Other non-coding RNAs could also be identified as bound to one, two or the three Rsm proteins (**Table 2**). This might suggest a possible ancillary role in titration of Rsm proteins, which led us to analyze them in some detail. Secondary structure predictions indicate that GGA motifs in short stem-loop structure, typical targets for CsrA/RsmA recognition (Dubey et al., 2005; Duss et al., 2014), are present in at least some of these RNA molecules, namely PP_mr15, PP_mr49 and PP_mr55 (**Supplementary Figure S3**). However, while *rsmY* and *rsmZ* show several of these motifs, only one per molecule was present in the other three. Also, conserved GacA binding sites, involved in transcriptional regulation of Rsm-titrating RNAs in other *Pseudomonas* species (Brencic et al., 2009) could be identified in the regions upstream

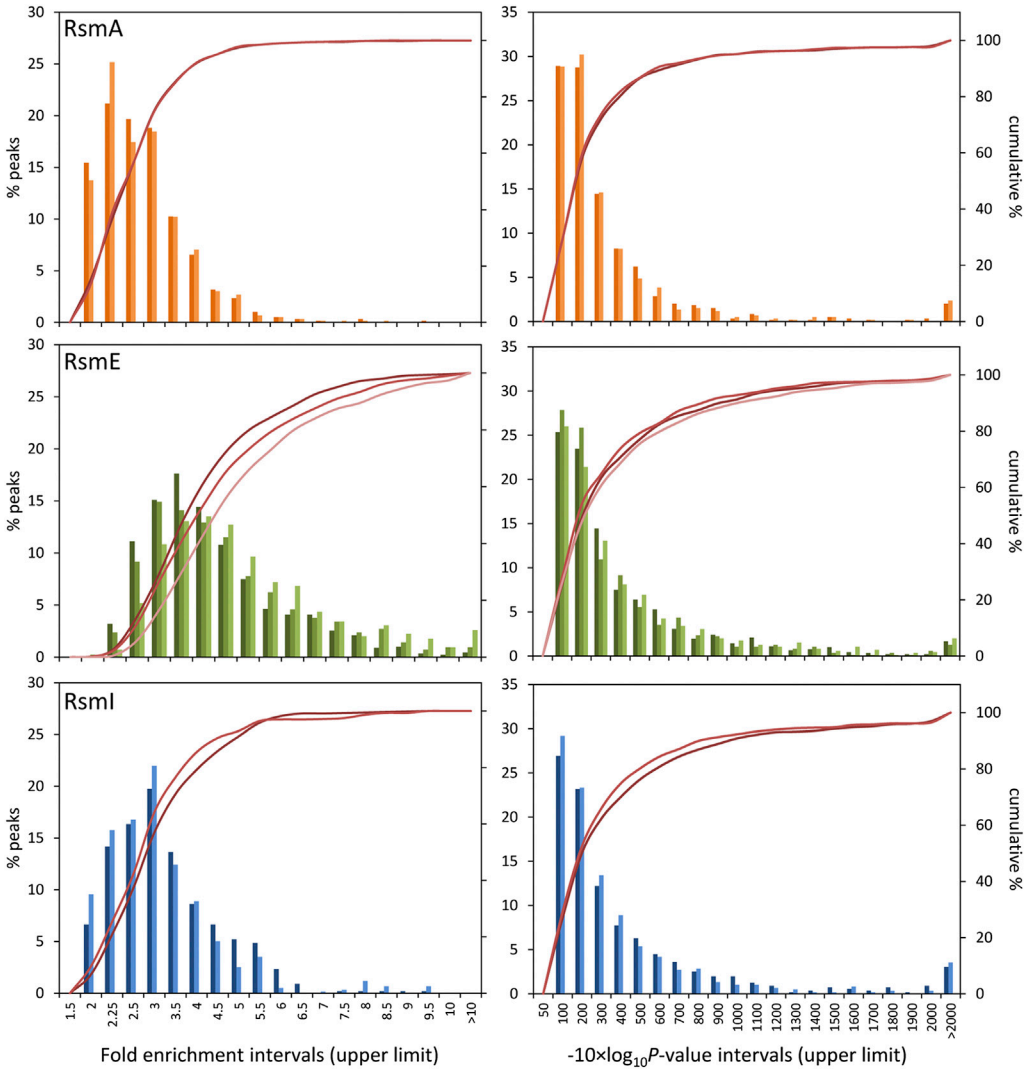


FIGURE 1 | Distribution of fold enrichment **(left)** and $-10 \times \mathrm{Log}_{10}$(*p*-value) scores **(right)** data of overrepresented RNA sequences associated to each Rsm protein (orange, RsmA; green, RsmE; blue, RsmI) with respect to total RNA. Data were grouped in the value intervals indicated in the *X*-axis, where the number corresponds to the upper limit of each interval. The histograms (primary *Y*-axis) show the frequency of targets in each interval. The crimson lines (secondary *Y*-axis) correspond to the cumulative percent values. The different color intensities in the histograms and lines correspond to the different biological replicates for each protein.

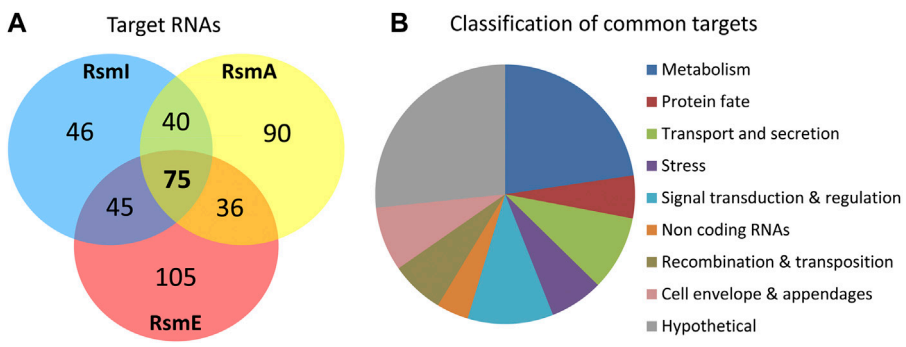


FIGURE 2 | (A) Venn diagram summarizing the number of targets identified for the three Rsm proteins in *P. putida* KT2440, according to the selection parameters detailed in the text. **(B)** Graphical representation of the different functional categories encompassing the 75 common targets.

TABLE 4 | Non-coding RNAs identified as targets for each Rsm protein.

Locus[b]	Length (bp)	Fold-enrichment[a]		
		RsmA	RsmE	RsmI
PP_mr05 (*rsmY*)	127	8.36	16.46	6.88
PP_mr15	209	3.83	5.37	(3.49)
PP_mr22 (*rsmZ*)	134	4.58	11.58	(5.11)
PP_mr44	385	2.87	6.49	3.22
PP_mr49	97	3	—	(3)
PP_mr52	149	2.36	6.96	4.72
PP_mr53	395	3.93	(3.76)	3.53
PP_mr55	82	—	5.20	—
PP_mr57	254	8.17	—	—
PP_mr59	97	2.42	(2.96)	(2.1)

[a]*Values in parentheses indicate targets below the* p*-value and/or fold-enrichment cutoffs. Minus sign indicates the target is not present in one or more of the enriched RNA replicas.*
[b]*Annotated according to 54.*

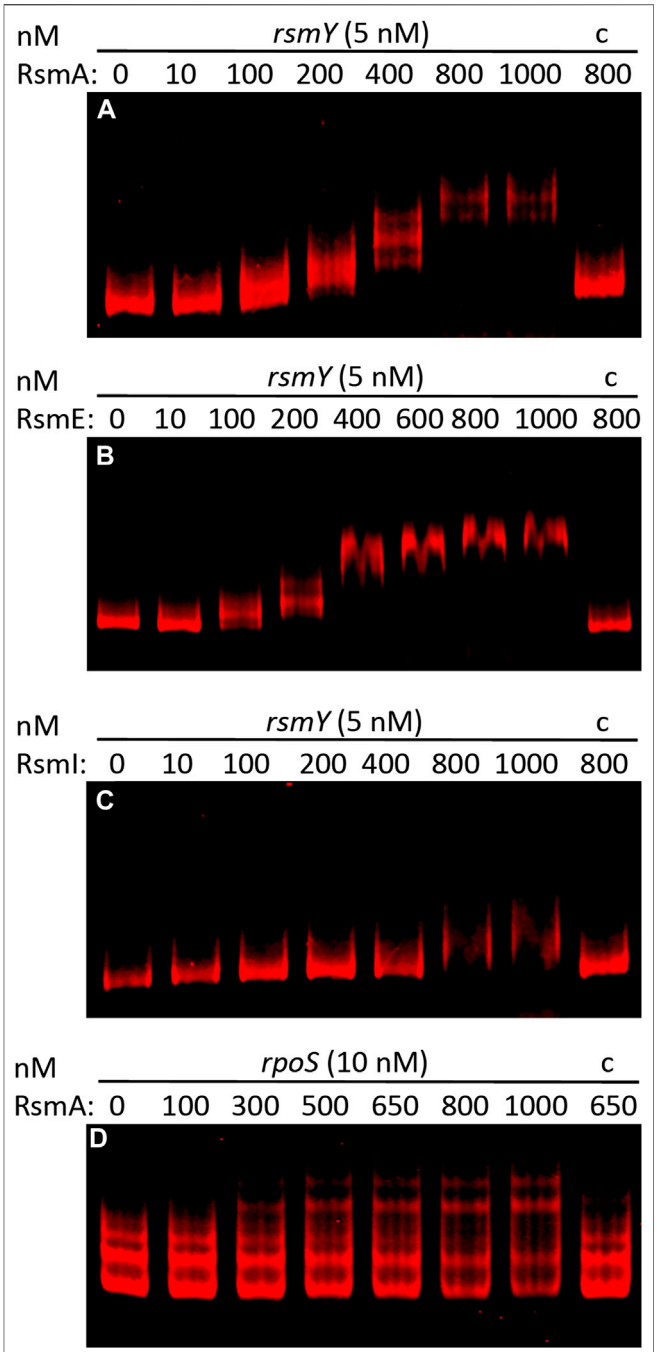


FIGURE 3 | Binding of RsmA-His$_6$ **(A)**, RsmE-His$_6$ **(B)** and RsmI-His$_6$ **(C)** to full-length *rsmY* RNA transcribed *in vitro*, and of RsmA-His$_6$ to a synthetic RNA fragment comprising the untranslated upstream region of *rpoS* **(D)**. EMSAs were carried out using fluorescently labeled RNA (5 or 10 nM) and incubated in the absence or presence of increasing concentrations of purified Rsm proteins, as indicate above each lane, or in the presence of 100-fold excess of unlabelled RNA competitor and the indicated concentration of protein **(C)**.

rsmY and *rsmZ*, being only partially conserved in PP_mr55 (**Supplementary Figure S4**), and absent in the remaining RNAs (not shown). All these data suggest that *rsmY* and *rsmZ* are the main, if not the only, true antagonists of Rsm proteins in *P. putida*. The remaining non-coding RNAs that are targets of these proteins may regulate further downstream elements in the Gac/Rsm cascade.

Additionally, several intergenic sequences were found among the targets for the three proteins. These correspond to regions adjacent to the different copies of the *IS*Ppu8 transposase in the genome of *P. putida* KT2440 or are potential insertion sites (or remnants of previous insertions) of this transposase, based on their sequence identity with *IS*Ppu8 flanking regions. Whether Rsm proteins influence the activity of this mobile genetic element in KT2440 is an interesting issue that deserves further study.

As mentioned above, a significant number of targets seem to be exclusive for one of the three proteins. A few of these correspond to enriched peaks that were below the established cut-off parameters in some samples, and therefore could represent common targets showing different affinities for each protein, with a strong preference for one of them. Such is the case, for example, of PP_0013 (*gyrB*), which is among the above-cut-off targets for RsmA but slightly below the FE and/or *p*-value cut-off for the other two proteins. Other targets are only enriched in association with one of the Rsm proteins and are not present in the RNA population associated to either of the other two, indicating they are truly specific for that Rsm homolog, e.g. PP_1656 (*relA*) for RsmA, PP_0168 (*lapA*) for RsmE, or PP_1803 (*wpbV*) for RsmI. Identifying the molecular basis for such specificity will require detailed bioinformatics analysis combined with *in vitro* and *in vivo* assays.

Supplementary Figure S5 provides data corresponding to β-galactosidase activity of several translational fusions of identified targets to *lacZ* in the wild type KT2440 and a triple *rsmArsmErsmI* deletion mutant, ΔIEA (Huertas-Rosales et al., 2016). In most cases, expression was enhanced in the mutant, indicating that the observed binding to their RNA targets leads to translation repression by these proteins; these results are consistent with previous observations on their influence upon

biofilm-related elements (see below). However, the inverse was true for two of the tested fusions, those corresponding to PP_1088 (*argG*, involved in arginine synthesis; Ramos-González et al., 2016) and PP_4482 (part of the gene cluster encoding the main arginine transporter and its regulator; Barrientos-Moreno et al., 2020). In these cases, expression was approximately halved in the mutant with respect to the wild type.

c-Di-GMP Signaling and Biofilm-Related Targets

As indicated in the Introduction, the three Rsm proteins have been shown to bind a specific site in the mRNA corresponding to the response regulator with diguanylate cyclase activity CfcR, and a triple *rsm* mutant in KT2440 shows increased levels of c-di-GMP and altered biofilm dynamics (Huertas-Rosales et al., 2016; Huertas-Rosales et al., 2017). Intriguingly, *cfcR* mRNA was not among the common targets listed in **Table 3**, being only found as target for RsmE and RsmI. The *cfcR* transcript is actually among the RsmA-bound RNA sequences overrepresented with respect to the control RNA, but the FE values (1.82 and 1.88 in each replicate, respectively) are below the established cut-off. This could indicate that binding of RsmA to the mRNA of *cfcR in vivo* is hampered by competition with the other two Rsm proteins, a possibility that has been previously suggested, based on *in vivo* expression data of *cfcR* compared to the high affinity observed for the RsmA/*cfcR* mRNA interaction *in vitro* (Huertas-Rosales et al., 2017).

Besides *cfcR*, transcripts from four other genes encoding proteins predicted to participate in c-di-GMP turnover and signaling were identified in this analysis, one of them (encoded by locus PP_0563) as target of the three proteins, and the rest (PP_0386, PP_0914, PP_2505), as targets of RsmE. Of these, the proteins encoded by PP_0563 and PP_2505 present GGDEF domains, characteristic of diguanylate cyclases. The first corresponds to GcbA, a diguanylate cyclase conserved in *Pseudomonas*, which has been reported to influence initial attachment and swimming motility (Petrova et al., 2014; Xiao et al., 2016). PP_0914 corresponds to the EAL domain-containing phosphodiesterase BifA, described to regulate biofilm development in *P. putida* (Jiménez-Fernández et al., 2015), and PP_0386 encodes a protein containing both GGDEF and EAL domains. Other relevant biofilm-related transcripts bound by RsmE included: *lapA*, encoding the main adhesin of *P. putida*, essential for initial attachment and biofilm formation (Hinsa et al., 2003; Martínez-Gil et al., 2010); the first gene in the cellulose synthesis operon; and genes in the operon encoding the species-specific EPS Pea

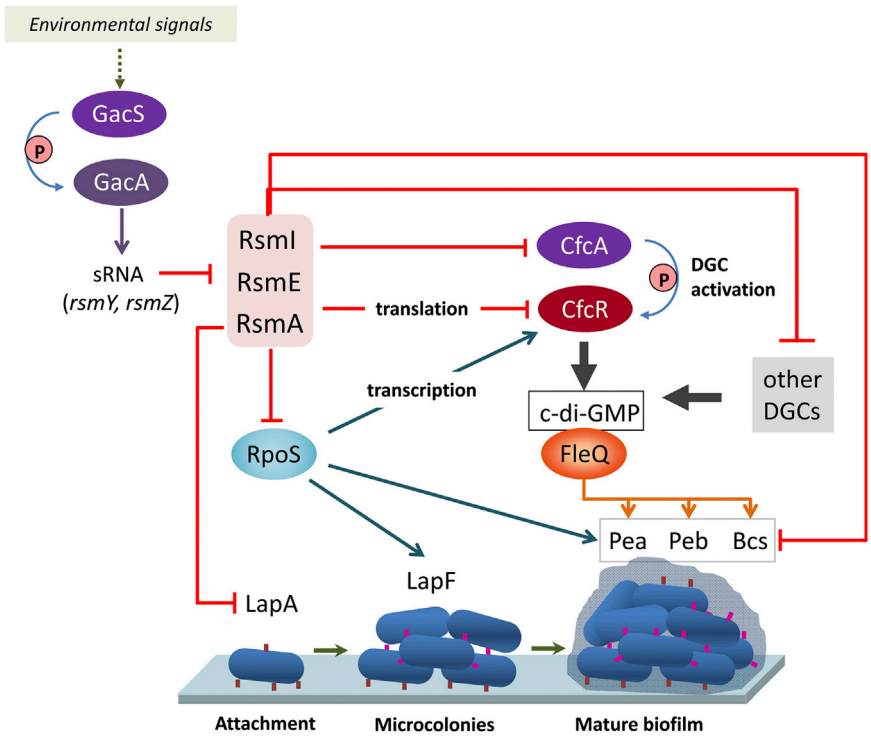


FIGURE 4 | Schematic view of the connection between the Gac/Rsm signaling cascade with biofilm formation through c-di-GMP turnover, adhesins and exopolysaccharides. The Rsm proteins exert a complex regulation on the response regulator with DGC activity CfcR, at the transcriptional (via RpoS), translational (via direct binding), and post-translational (via the histidine kinase CfcA, required for phosphorylation of CfcR) levels. Direct influence of Rsm proteins on structural elements of the biofilm matrix, and indirect influence via RpoS and the c-di-GMP associated regulator FleQ are also depicted. Blue arrows indicate positive regulation and red lines negative regulation.

(Nilsson et al., 2011), which is also a target for RsmA. Neither *lapF*, encoding the second relevant adhesin present in KT2440 (Martínez-Gil et al., 2010), nor the other two EPS operons described in this strain (Nilsson et al., 2011) were identified in this analysis.

Since altered expression of translational fusions corresponding to some of the genes indicated aboved has been previously observed in a triple *rsmAEI* mutant (Huertas-Rosales et al., 2016), it is likely that the observed expression changes are due to an indirect effect of Rsm proteins through other regulators. One such regulator is the stationary phase sigma factor RpoS, which controls expression of *lapF* (Martínez-Gil et al., 2010). RpoS was found to be negatively regulated by RsmA In *P. protegens* CHA0 (Heeb et al., 2005), and the mRNA corresponding to *rpoS* (PP_1623) is among the targets for RsmA in KT2440 (**Supplementary Table S1**). Moreover, binding of RsmA to an *in vitro* transcribed RNA fragment containing the ribosome binding site and start codon of *rpoS* was confirmed via EMSA (**Figure 3**). Given that expression of *cfcR* is also regulated by RpoS (Matilla et al., 2011), these results support the previously proposed model whereby Rsm proteins exert a dual control, direct and indirect, on *cfcR* (Huertas-Rosales et al., 2017). Remarkably, the gene *cfcA*, encoding a sensor histidine kinase essential for activation of CfcR (Ramos-González et al., 2016), is among the targets for the three proteins, indicating that c-di-GMP signaling through CfcR is tightly regulated by the Gac/Rsm network. A schematic view of this signaling cascade connecting external stimuli with biofilm formation through Rsm elements is depicted in **Figure 4**.

Influence of Rsm Proteins on Nutrient Utilization and Rhizosphere Fitness of *P. putida* KT2440

Among the shared targets for the three riboregulators, about a third of the transcripts with known or predicted functions correspond to genes with metabolism-related roles in KT2440 (**Figure 2B** and **Table 4**). This was expected since previous findings have established direct and indirect connections between the CsrA/Rsm system and metabolism as well as carbon storage functions (Sabnis et al., 1995; Yang et al., 1996; Pannuri et al., 2016; Romero et al., 2018). Furthermore, a significant number of targets identified for one or more of the Rsm proteins included transcripts from genes related to transport of nutrients (**Figure 2B** and **Table 3**). All this prompted us to carry out a preliminary high throughput study comparing the growth of KT2440 wild type strain and a triple *rsmAEI* mutant derivative (ΔIEA; Huertas-Rosales et al., 2016) in BIOLOG plates using 192 and 96 compounds as sole carbon or nitrogen source, respectively. A sample of some of the obtained data is shown in **Supplementary Figure S5**. Growth differences between the two strains were observed for several compounds, particularly certain amino acids and their derivatives. Further detailed evaluation of growth in some of these compounds confirmed the existence of differences between KT2440 and the triple *rsm* mutant. In particular, a prolonged

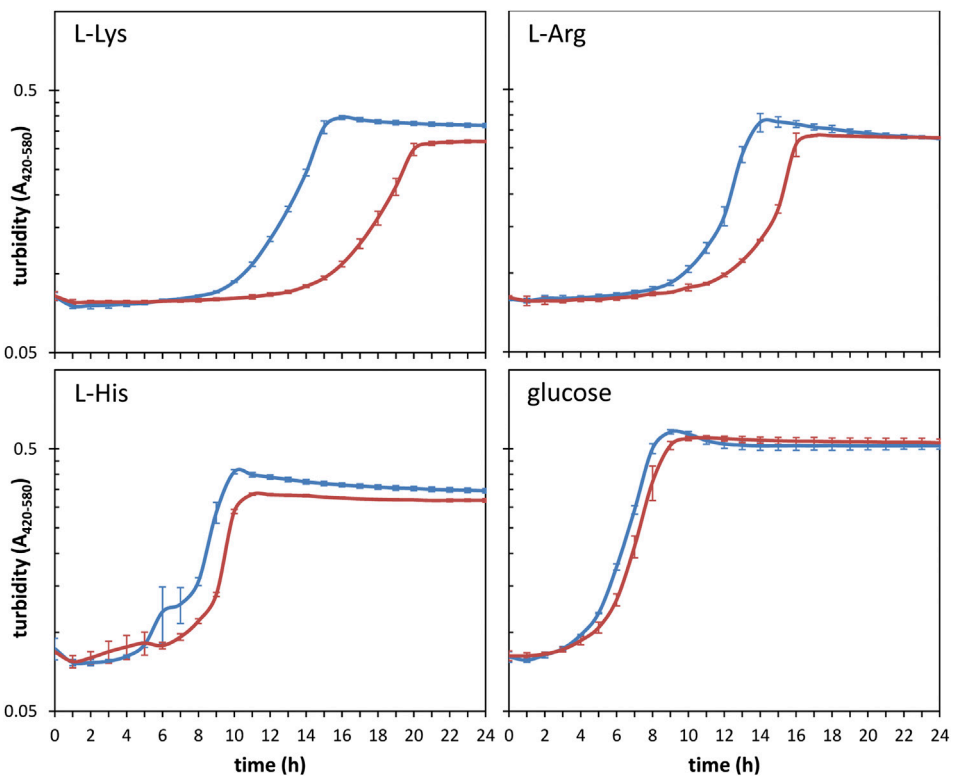


FIGURE 5 | Growth of KT2440 (blue) and the triple *rsm* mutant (red) in M9 minimal medium with 5 mM L-lysine, L-arginine, L-histidine or glucose as carbon source. Cultures were grown in 96-well plates at 30°C and $OD_{420–580}$ was read at 1-h intervals in an automated BioScreen equipped with a wide band filter (420–580 nm) to minimize potential alterations due to changes in medium color. Data are averages and standard deviations of experiments done in duplicate, with three technical replicas each.

lag phase was observed in the mutant with L-lysine as carbon source, and to a lesser extent with L-arginine, although the final turbidity reached by both strains with this last amino acid was similar (**Figure 5**). In contrast, growth differences were less evident with the other basic amino acid L-histidine (**Figure 5**).

These results evidenced that the Rsm system may play a regulatory role in metabolism and/or uptake of nutrients, particularly of basic amino acids, in KT2440. Hence, potential Rsm-enriched targets explaining these divergences were explored. A survey of the identified RNA molecules related to these processes indicated that those corresponding to *arcD* (PP_1002; arginine-ornithine antiporter) and *hisP* (PP_4483; ATP-binding subunit of a histidine/lysine/arginine/ornithine transporter) are targets for RsmE; *artJ* (PP_0282; L-arginine ABC transporter substrate-binding subunit) and *amaD* (PP_3596; D-lysine oxidase) are targets for RsmA, and *amaB* (PP_5258; L-piperidine-6-carboxylate dehydrogenase) is a target for both RsmE and RsmI (**Supplementary Table S1**). Based on data from the Kyoto Encyclopedia of Genes and Genomes (www.genome.jp/kegg/), AmaB is involved in catabolic pathways for L-lysine and L-arginine in *P. putida*, participating in the conversion of aminobutanal, N4-acetyl-aminobutanal, and 4-trimethyl-ammoniobutanal into their corresponding butanoate derivatives.

Besides glucose metabolism, a connection has been established in *E. coli* between CsrA and the stringent response (Edwards et al., 2011), a regulatory network triggered by amino acid limitation and controlled by the RelA and SpoT proteins. It should be noted that besides the targets indicated above in relation with arginine and lysine utilization, the *relA* (PP_1656) mRNA is also a target for RsmA (**Supplementary Table S1**).

Previous reports have established that *hisP* and genes involved in lysine catabolism are preferentially expressed in KT2440 in the rhizosphere of corn plants (Espinosa-Urgel and Ramos, 2001; Matilla et al., 2007). In addition, a connection beween arginine transport and c-di-GMP signaling in this strain has been recently reported (Barrientos-Moreno et al., 2020). These facts, and the influence of Rsm proteins on biofilm formation and surface motility, along with the identification of c-di-GMP turnover elements and other biofilm-related genes as targets of these proteins, made us analyze if mutants in *rsm* genes showed altered fitness in the rhizosphere. Competitive root colonization assays were done in corn (*Zea mays* L.) plants with the wild type and single (ΔI, ΔE, ΔA), double (ΔIE, ΔIA, ΔEA) and the triple (ΔIEA) *rsm* mutants (Huertas-Rosales et al., 2016). For that purpose, germinated corn seeds were inoculated with miniTn7Km-tagged KT2440 and each mutant tagged with miniTn7Sm, in a 1:1 proportion. Plants were sown in sterile sand and the population of each strain was evaluated in the whole root and in the root tip after 7 days. Results are shown in **Figure 6**. The mutation in *rsmA* caused a slight reduction in competitive colonization of the whole root and a much larger effect when root tip colonization was evaluated. This phenotype was not observed in the other single mutants, nor in the ΔIE double mutant. However, while the results with the ΔIA double mutant were similar to those obtained with the ΔA single mutant, a cumulative effect could be observed in the ΔEA strain, which showed a significant decrease in colonization of both the whole root and the root tip. This phenotype was very similar to that observed in the triple mutant. In this set of experiments the Sm resistance marker seemed to confer a slight advantage over the Km resistance marker, according to the results obtained with the wild type derivatives tagged with each mini-Tn7 (**Figure 6**). This could suggest that the actual influence of the *rsm* mutations might be larger than observed here.

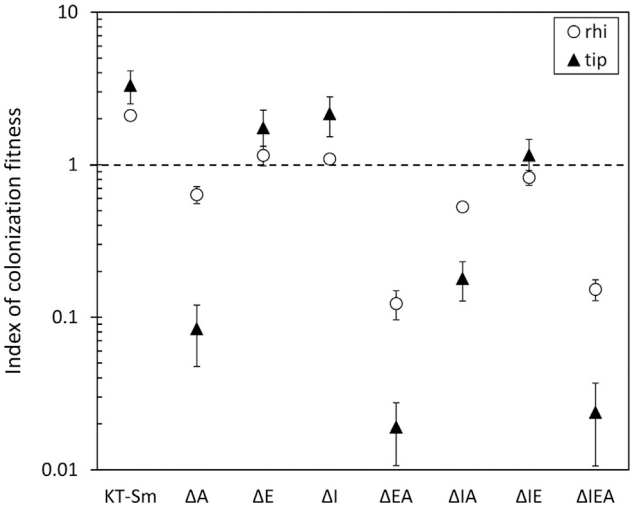


FIGURE 6 | Colonization capacity of mutants in *rsm* genes in competition with the wild type in the rhizosphere (circles) and the root tip (triangles) of corn plants. KT2440 tagged with miniTn7-Km was co-inoculated in a 1:1 proportion with each of the mutants, which were tagged with miniTn7-Sm. To check the potential effect of the antibiotic resistance markers, a competitive colonization assay co-inoculating KT2440 miniTn7-Km and KT2440 miniTn7-Sm (KT-Sm) was also included as control. The index of colonization fitness was measured as 1/[(percentage of recovered wild type vs. each mutant per plant 7 days post-inoculation)/(percentage wt vs. mutant in the initial inocula)]. A value of 1 (broken line) indicates equal colonization capacity. Data are the averages and standard errors for six plants.

The loss of fitness in the rhizosphere of the ΔEA and ΔIEA mutants may result from the overall influence of Rsm proteins on different metabolic processes, including amino acid transport, while the reduction in root tip colonzation could correlate with the previously observed decrease in swimming motility in the ΔEA and ΔIEA strains and the lack of swarming motility of all the rsm mutants, with the exception of ΔI, which still retained some surface motility (Huertas-Rosales et al., 2016). It will also be of interest to explore whether the type VI secretion system K1, some of whose genes were identified in our study (**Table 3**), may contribute to the fitness of KT2440 in the rhizosphere. It has been reported that this system can provide a competitive advantage to this strain in the presence of phytopathogenic bacteria (Bernal et al., 2017).

CONCLUDING REMARKS

This work represents the first effort to define the global regulatory network commanded by Rsm proteins in *P. putida*. Besides exposing its complexity and the vast influence that post-transcriptional regulation is bound to have in this bacterium, ranging from amino acid metabolism to potential transposon-mediated DNA rearrangements, the information obtained leads to a relevant question to be analyzed in detail, i.e. the characteristics by which

an RNA molecule constitutes a shared target for the three Rsm proteins or is selectively bound by only one or two of them.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

OH-R, MR, K-WH, LB-M, MAM-H, MT performed the experimental work; OH-R, MR, MC, SH, MIR-G, ME-U designed the work; OH-R, MR, K-GC, MC, SH, MIR-G, ME-U analyzed data; MC, SH, K-GC, MIR-G, MEU obtained funding; OH-R, ME-U wrote the paper; All authors have revised the manuscript.

ACKNOWLEDGMENTS

This work was supported by grants BFU2013-43469-P, BFU2016-80122-P and PID2019-109372GB-I00 from the Plan Estatal de I+D+I (Agencia Estatal de Investigación, Spanish Ministry of Science and Innovation and FEDER funds). Funding from the Biotechnology and Biological Sciences Research Council, United Kingdom (BB/R012415/1), and the University of Malaya (FRGS grant FP022-2018A and HIR grant H-50001-00-A000027) are also gratefully acknowledged.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2021.624061/full#supplementary-material.

REFERENCES

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., et al. (1987). *Current protocols in molecular biology*. New York, NY: Wiley.

Baker, C. S., Eöry, L. A., Yakhnin, H., Mercante, J., Romeo, T., and Babitzke, P. (2007). CsrA inhibits translation initiation of *Escherichia coli hfq* by binding to a single site overlapping the Shine-Dalgarno sequence. *J. Bacteriol.* 189, 5472–5481. doi:10.1128/JB.00529-07

Barrientos-Moreno, L., Molina-Henares, M. A., Ramos-González, M. I., and Espinosa-Urgel, M. (2020). Arginine as an environmental and metabolic cue for cyclic diguanylate signalling and biofilm formation in *Pseudomonas putida*. *Sci. Rep.* 10 (1), 13623. doi:10.1038/s41598-020-70675-x

Bernal, P., Allsopp, L. P., Filloux, A., and Llamas, M. A. (2017). The *Pseudomonas putida* T6SS is a plant warden against phytopathogens. *ISME J.* 11, 972–987. doi:10.1038/ismej.2016.169

Brencic, A., and Lory, S. (2009). Determination of the regulon and identification of novel mRNA targets of *Pseudomonas aeruginosa* RsmA. *Mol. Microbiol.* 72, 612–632. doi:10.1111/j.1365-2958.2009.06670.x

Brencic, A., McFarland, K. A., McManus, H. R., Castang, S., Mogno, I., Dove, S. L., et al. (2009). The GacS/GacA signal transduction system of *Pseudomonas aeruginosa* acts exclusively through its control over the transcription of the RsmY and RsmZ regulatory small RNAs. *Mol. Microbiol.* 73, 434–445. doi:10.1111/j.1365-2958.2009.06782.x

Choi, K. H., Kumar, A., and Schweizer, H. P. (2006). A 10-min method for preparation of highly electrocompetent *Pseudomonas aeruginosa* cells: application for DNA fragment transfer between chromosomes and plasmid transformation. *J. Microbiol. Methods* 64, 391–397. doi:10.1016/j.mimet.2005.06.001

Coggan, K. A., and Wolfgang, M. C. (2012). Global regulatory pathways and cross-talk control *Pseudomonas aeruginosa* environmental lifestyle and virulence phenotype. *Curr. Issues Mol. Biol.* 14, 47–70.

Dubey, A. K., Baker, C. S., Romeo, T., and Babitzke, P. (2005). RNA sequence and secondary structure participate in high-affinity CsrA-RNA interaction. *RNA* 11, 1579–1587. doi:10.1261/rna.2990205

Duss, O., Michel, E., Diarra dit Konté, N., Schubert, M., and Allain, F. H. (2014). Molecular basis for the wide range of affinity found in Csr/Rsm protein-RNA recognition. *Nucleic Acids Res.* 42, 5332–5346. doi:10.1093/nar/gku141

Edwards, A. N., Patterson-Fortin, L. M., Vakulskas, C. A., Mercante, J. W., Potrykus, K., Vinella, D., et al. (2011). Circuitry linking the Csr and stringent response global regulatory systems. *Mol. Microbiol.* 80, 1561–1580. doi:10.1111/j.1365-2958.2011.07663.x

Espinosa-Urgel, M., and Ramos, J. L. (2001). Expression of a *Pseudomonas putida* aminotransferase involved in lysine catabolism is induced in the rhizosphere. *Appl. Environ. Microbiol.* 67, 5219–5224. doi:10.1128/AEM.67.11.5219-5224.2001

Ferreiro, M. D., Nogales, J., Farias, G. A., Olmedilla, A., Sanjuán, J., and Gallegos, M. T. (2018). Multiple CsrA proteins control key virulence traits in *Pseudomonas syringae* pv. tomato DC3000. *Mol. Plant Microbe Interact.* 31, 525–536. doi:10.1094/MPMI-09-17-0232-R

Figueroa-Bossi, N., Schwartz, A., Guillemardet, B., D'Heygère, F., Bossi, L., and Boudvillain, M. (2014). RNA remodeling by bacterial global regulator CsrA promotes Rho-dependent transcription termination. *Genes Dev.* 28, 1239–1251. doi:10.1101/gad.240192.114

Heeb, S., Valverde, C., Gigot-Bonnefoy, C., and Haas, D. (2005). Role of the stress sigma factor RpoS in GacA/RsmA-controlled secondary metabolism and resistance to oxidative stress in *Pseudomonas fluorescens* CHA0. *FEMS Microbiol. Lett.* 243, 251–258. doi:10.1016/j.femsle.2004.12.008

Heroven, A. K., Böhme, K., and Dersch, P. (2012). The Csr/Rsm system of *Yersinia* and related pathogens: a post-transcriptional strategy for managing virulence. *RNA Biol.* 9, 379–391. doi:10.4161/rna.19333

Hinsa, S. M., Espinosa-Urgel, M., Ramos, J. L., and O'Toole, G. A. (2003). Transition from reversible to irreversible attachment during biofilm formation by *Pseudomonas fluorescens* WCS365 requires an ABC transporter and a large secreted protein. *Mol. Microbiol.* 49, 905–918. doi:10.1046/j.1365-2958.2003.03615.x

Hör, J., Gorski, S. A., and Vogel, J. (2018). Bacterial RNA biology on a genome scale. *Mol. Cell* 70, 785–799. doi:10.1016/j.molcel.2017.12.023

Huertas-Rosales, Ó., Ramos-González, M. I., and Espinosa-Urgel, M. (2016). Self-regulation and interplay of Rsm family proteins modulate the lifestyle of *Pseudomonas putida*. *Appl. Environ. Microbiol.* 82, 5673–5686. doi:10.1128/AEM.01724-16

Huertas-Rosales, Ó., Romero, M., Heeb, S., Espinosa-Urgel, M., Cámara, M., and Ramos-González, M. I. (2017). The *Pseudomonas putida* CsrA/RsmA homologues negatively affect c-di-GMP pools and biofilm formation through the GGDEF/EAL response regulator CfcR. *Environ. Microbiol.* 19, 3551–3566. doi:10.1111/1462-2920.13848

Ishihama, A. (2010). Prokaryotic genome regulation: multifactor promoters, multitarget regulators and hierarchic networks. *FEMS Microbiol. Rev.* 34, 628–645. doi:10.1111/j.1574-6976.2010.00227.x

Jacob, F., and Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* 3, 318–356. doi:10.1016/s0022-2836(61)80072-7

Janssen, K. H., Diaz, M. R., Golden, M., Graham, J. W., Sanders, W., Wolfgang, M. C., et al. (2018). Functional analyses of the RsmY and RsmZ small noncoding regulatory RNAs in *Pseudomonas aeruginosa*. *J. Bacteriol.* 200, e00736–17. doi:10.1128/JB.00736-17

Jiménez-Fernández, A., López-Sánchez, A., Calero, P., and Govantes, F. (2015). The c-di-GMP phosphodiesterase BifA regulates biofilm development in *Pseudomonas putida*. *Environ Microbiol Rep* 7, 78–84. doi:10.1111/1758-2229.12153

Kay, E., Dubuis, C., and Haas, D. (2005). Three small RNAs jointly ensure secondary metabolism and biocontrol in *Pseudomonas fluorescens* CHA0. *Proc. Natl. Acad. Sci. U.S.A.* 102, 17136–17141. doi:10.1073/pnas.0505673102

Koch, B., Jensen, L. E., and Nybroe, O. (2001). A panel of Tn7-based vectors for insertion of the *gfp* marker gene or for delivery of cloned DNA into Gram-negative bacteria at a neutral chromosomal site. *J. Microbiol. Methods* 45, 187–195. doi:10.1016/s0167-7012(01)00246-9

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10, R25. doi:10.1186/gb-2009-10-3-r25

Lawhon, S. D., Frye, J. G., Suyemoto, M., Porwollik, S., McClelland, M., and Altier, C. (2003). Global regulation by CsrA in *Salmonella typhimurium*. *Mol. Microbiol.* 48, 1633–1645. doi:10.1046/j.1365-2958.2003.03535.x

Lennox, E. S. (1955). Transduction of linked genetic characters of the host by bacteriophage P1. *Virology* 1, 190–206. doi:10.1016/0042-6822(55)90016-7

Lenz, D. H., Miller, M. B., Zhu, J., Kulkarni, R. V., and Bassler, B. L. (2005). CsrA and three redundant small RNAs regulate quorum sensing in *Vibrio cholerae*. *Mol. Microbiol.* 58, 1186–1202. doi:10.1111/j.1365-2958.2005.04902.x

Martínez-Gil, M., Yousef-Coronado, F., and Espinosa-Urgel, M. (2010). LapF, the second largest *Pseudomonas putida* protein, contributes to plant root colonization and determines biofilm architecture. *Mol. Microbiol.* 77, 549–561. doi:10.1111/j.1365-2958.2010.07249.x

Matilla, M. A., Espinosa-Urgel, M., Rodríguez-Herva, J. J., Ramos, J. L., and Ramos-González, M. I. (2007). Genomic analysis reveals the major driving forces of bacterial life in the rhizosphere. *Genome Biol.* 8, R179. doi:10.1186/gb-2007-8-9-r179

Matilla, M. A., Travieso, M. L., Ramos, J. L., and Ramos-González, M. I. (2011). Cyclic diguanylate turnover mediated by the sole GGDEF/EAL response regulator in *Pseudomonas putida*: its role in the rhizosphere and an analysis of its target processes. *Environ. Microbiol.* 13, 1745. doi:10.1111/j.1462-2920.2011.02499.x

Morris, E. R., Hall, G., Li, C., Heeb, S., Kulkarni, R. V., Lovelock, L., et al. (2013). Structural rearrangement in an RsmA/CsrA ortholog of *Pseudomonas aeruginosa* creates a dimeric RNA-binding protein, RsmN. *Structure* 21, 1659–1671. doi:10.1016/j.str.2013.07.007

Nilsson, M., Chiang, W. C., Fazli, M., Gjermansen, M., Givskov, M., and Tolker-Nielsen, T. (2011). Influence of putative exopolysaccharide genes on *Pseudomonas putida* KT2440 biofilm stability. *Environ. Microbiol.* 13, 1357–1369. doi:10.1111/j.1462-2920.2011.02447.x

Pannuri, A., Vakulskas, C. A., Zere, T., McGibbon, L. C., Edwards, A. N., Georgellis, D., et al. (2016). Circuitry linking the catabolite repression and Csr global regulatory systems of *Escherichia coli*. *J. Bacteriol.* 198, 3000–3015. doi:10.1128/JB.00454-16

Petrova, O. E., Cherny, K. E., and Sauer, K. (2014). The *Pseudomonas aeruginosa* diguanylate cyclase GcbA, a homolog of *P. fluorescens* GcbA, promotes initial attachment to surfaces, but not biofilm formation, via regulation of motility. *J. Bacteriol.* 196, 2827–2841. doi:10.1128/JB.01628-14

Ramos-González, M. I., MA, Matilla., Quesada, J. M., Ramos, J. L., and Espinosa-Urgel, M. (2013). "Using genomics to unveil bacterial determinants of rhizosphere life style," in *Molecular microbial ecology of the rhizosphere*. Editor F. J. de Bruijn (Hoboken, NJ: John Wiley & Sons), Vol. 1, 7–16.

Ramos-González, M. I., Travieso, M. L., Soriano, M. I., Matilla, M. A., Huertas-Rosales, Ó., Barrientos-Moreno, L., et al. (2016). Genetic dissection of the regulatory network associated with high c-di-GMP levels in *Pseudomonas putida* KT2440. *Front. Microbiol.* 7, 1093. doi:10.3389/fmicb.2016.01093

Regenhardt, D., Heuer, H., Heim, S., Fernández, D. U., Strömpl, C., Moore, E. R., et al. (2002). Pedigree and taxonomic credentials of *Pseudomonas putida* strain KT2440. *Environ. Microbiol.* 4, 912–915. doi:10.1046/j.1462-2920.2002.00368.x

Reimmann, C., Valverde, C., Kay, E., and Haas, D. (2005). Posttranscriptional repression of GacS/GacA-controlled genes by the RNA-binding protein RsmE acting together with RsmA in the biocontrol strain *Pseudomonas fluorescens* CHA0. *J. Bacteriol.* 187, 276–285. doi:10.1128/JB.187.1.276-285.2005

Romeo, T. (1998). Global regulation by the small RNA-binding protein CsrA and the non-coding RNA molecule CsrB. *Mol. Microbiol.* 29, 1321–1330. doi:10.1046/j.1365-2958.1998.01021.x

Romeo, T., Vakulskas, C. A., and Babitzke, P. (2013). Post-transcriptional regulation on a global scale: form and function of Csr/Rsm systems. *Environ. Microbiol.* 15, 313–324. doi:10.1111/j.1462-2920.2012.02794.x

Romero, M., Silistre, H., Lovelock, L., Wright, V. J., Chan, K. G., Hong, K. W., et al. (2018). Genome-wide mapping of the RNA targets of the *Pseudomonas aeruginosa* riboregulatory protein RsmN. *Nucleic Acids Res.* 46, 6823–6840. doi:10.1093/nar/gky324

Sabnis, N. A., Yang, H., and Romeo, T. (1995). Pleiotropic regulation of central carbohydrate metabolism in *Escherichia coli* via the gene *csrA*. *J. Biol. Chem.* 270, 29096–29104. doi:10.1074/jbc.270.49.29096

Sambrook, J., and Russell, D. W. (2001). *Molecular cloning: a laboratory manual*. New York, NY: Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Shis, D. L., Bennett, M. R., and Igoshin, O. A. (2018). Dynamics of bacterial gene regulatory networks. *Annu. Rev. Biophys.* 47, 447–467. doi:10.1146/annurev-biophys-070317-032947

Sonnleitner, E., and Haas, D. (2011). Small RNAs as regulators of primary and secondary metabolism in *Pseudomonas* species. *Appl. Microbiol. Biotechnol.* 91, 63–79. doi:10.1007/s00253-011-3332-1

Sowa, S. W., Gelderman, G., Leistra, A. N., Buvanendiran, A., Lipp, S., Pitaktong, A., et al. (2017). Integrative FourD omics approach profiles the target network of the carbon storage regulatory system. *Nucleic Acids Res.* 45, 1673–1686. doi:10.1093/nar/gkx048

Sterzenbach, T., Nguyen, K. T., Nuccio, S. P., Winter, M. G., Vakulskas, C. A., Clegg, S., et al. (2013). A novel CsrA titration mechanism regulates fimbrial gene expression in *Salmonella typhimurium*. *EMBO J.* 32, 2872–2883. doi:10.1038/emboj.2013.206

Vakulskas, C. A., Potts, A. H., Babitzke, P., Ahmer, B. M., and Romeo, T. (2015). Regulation of bacterial virulence by Csr (Rsm) systems. *Microbiol. Mol. Biol. Rev.* 79, 193–224. doi:10.1128/MMBR.00052-14

Van Assche, E., Van Puyvelde, S., Vanderleyden, J., and Steenackers, H. P. (2015). RNA-binding proteins involved in post-transcriptional regulation in bacteria. *Front. Microbiol.* 6, 141. doi:10.3389/fmicb.2015.00141

Winsor, G. L., Griffiths, E. J., Lo, R., Dhillon, B. K., Shay, J. A., and Brinkman, F. S. (2016). Enhanced annotations and features for comparing thousands of *Pseudomonas* genomes in the *Pseudomonas* genome database. *Nucleic Acids Res.* 44 (D1), D646–D653. doi:10.1093/nar/gkv1227

Xiao, Y., Nie, H., Liu, H., Chen, W., and Huang, Q. (2016). Expression of the diguanylate cyclase GcbA is regulated by FleQ in response to cyclic di-GMP in *Pseudomonas putida* KT2440. *Environ Microbiol Rep* 8, 993–1002. doi:10.1111/1758-2229.12478

Yakhnin, A. V., Baker, C. S., Vakulskas, C. A., Yakhnin, H., Berezin, I., Romeo, T., et al. (2013). CsrA activates flhDC expression by protecting flhDC mRNA from RNase E-mediated cleavage. *Mol. Microbiol.* 87, 851–866. doi:10.1111/mmi.12136

Yakhnin, H., Pandit, P., Petty, T. J., Baker, C. S., Romeo, T., and Babitzke, P. (2007). CsrA of *Bacillus subtilis* regulates translation initiation of the gene encoding the flagellin protein (*hag*) by blocking ribosome binding. *Mol. Microbiol.* 64, 1605–1620. doi:10.1111/j.1365-2958.2007.05765.x

Yang, H., Liu, M. Y., and Romeo, T. (1996). Coordinate genetic regulation of glycogen catabolism and biosynthesis in *Escherichia coli* via the *csrA* gene product. *J. Bacteriol.* 178, 1012–1017. doi:10.1128/jb.178.4.1012-1017.1996

Yousef-Coronado, F., Travieso, M. L., and Espinosa-Urgel, M. (2008). Different, overlapping mechanisms for colonization of abiotic and plant surfaces by *Pseudomonas putida*. *FEMS Microbiol. Lett.* 288, 118–124. doi:10.1111/j.1574-6968.2008.01339.x

Zhang, Y., Liu, T., Meyer, C. A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., et al. (2008). Model-based analysis of ChIP-seq (MACS). *Genome Biol.* 9, R137. doi:10.1186/gb-2008-9-9-r137

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Huertas-Rosales, Romero, Chan, Hong, Cámara, Heeb, Barrientos-Moreno, Molina-Henares, Travieso, Ramos-González and Espinosa-Urgel. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 05 March 2021
doi: 10.3389/fmolb.2021.634461

Acidic pH Decreases the Endonuclease Activity of Initiator RepB and Increases the Stability of the Covalent RepB-DNA Intermediate while Has Only a Limited Effect on the Replication of Plasmid pMV158 in *Lactococcus lactis*

Rafael Valdelvira, Lorena Bordanaba-Ruiseco, Cristina Martín-Huestamendía, José Angel Ruiz-Masó* and Gloria del Solar*

Department of Microbial and Plant Biotechnology, Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas, Madrid, Spain

OPEN ACCESS

Edited by:
Tatiana Venkova,
Fox Chase Cancer Center,
United States

Reviewed by:
Christian Magni,
CONICET Instituto de Biología Molecular y Celular de Rosario (IBR), Argentina
Arijit Dutta,
The University of Texas Health Science Center at San Antonio, United States
Itziar Alkorta,
University of the Basque Country, Spain

****Correspondence:***
José Ángel Ruiz-Masó
jaruizmaso@cib.csic.es
Gloria del Solar
gdelsolar@cib.csic.es

Specialty section:
This article was submitted to Molecular Recognition, a section of the journal Frontiers in Molecular Biosciences

Received: *27 November 2020*
Accepted: *27 January 2021*
Published: *05 March 2021*

Citation:
Valdelvira R, Bordanaba-Ruiseco L, Martín-Huestamendía C, Ruiz-Masó JA and del Solar G (2021) Acidic pH Decreases the Endonuclease Activity of Initiator RepB and Increases the Stability of the Covalent RepB-DNA Intermediate while Has Only a Limited Effect on the Replication of Plasmid pMV158 in Lactococcus lactis. *Front. Mol. Biosci. 8:634461. doi: 10.3389/fmolb.2021.634461*

Plasmid vectors constitute a valuable tool for homologous and heterologous gene expression, for characterization of promoter and regulatory regions, and for genetic manipulation and labeling of bacteria. During the last years, a series of vectors based on promiscuous replicons of the pMV158 family have been developed for their employment in a variety of Gram-positive bacteria and proved to be useful for all above applications in lactic acid bacteria. A proper use of the plasmid vectors requires detailed knowledge of their main replicative features under the changing growth conditions of the studied bacteria, such as the acidification of the culture medium by lactic acid production. Initiation of pMV158 rolling-circle replication is catalyzed by the plasmid-encoded RepB protein, which performs a sequence-specific cleavage on one of the parental DNA strands and, as demonstrated in this work, establishes a covalent bond with the 5′-P end generated in the DNA. This covalent adduct must last until the leading-strand termination stage, where a new cleavage on the regenerated nick site and a subsequent strand-transfer reaction result in rejoining of the ends of the cleaved parental strand, whereas hydrolysis of the newly-generated adduct would release the protein from a nicked double-stranded DNA plasmid form. We have analyzed here the effect of pH on the different *in vitro* reactions catalyzed by RepB and on the *in vivo* replication ability of plasmid pMV158. We show that acidic pH greatly impairs the catalytic activity of the protein and reduces hydrolysis of the covalent RepB-DNA adduct, as expected for the nucleophilic nature of these reactions. Conversely, the ability of pMV158 to replicate *in vivo*, as monitored by the copy number and segregational stability of the plasmid in *Lactococcus lactis*, remains almost intact at extracellular pHs ranging from 7.0 to 5.0, and a significant reduction (by ~50%) in the plasmid copy number per chromosome equivalent is only observed at pH 4.5. Moreover, the RepB to pMV158 molar ratio is

increased at pH 4.5, suggesting the existence of compensatory mechanisms that operate *in vivo* to allow pMV158 replication at pH values that severely disturb the catalytic activity of the initiator protein.

Keywords: rolling circle replication (RCR), promiscuous plasmid pMV158, nucleophilic attack, endonuclease activity, strand-transfer activity, covalent RepB-DNA adduct

INTRODUCTION

Rolling-circle replication (RCR) was first described for single-stranded DNA (ssDNA) coliphages about half a century ago (Gilbert and Dressler, 1968; Dressler, 1970; Baas, 1985). Later, the RCR mechanism was reported for various small multicopy plasmids isolated from Gram-positive (Koepsel et al., 1985; te Riele et al., 1986; del Solar et al., 1987) and Gram-negative (Yasukawa et al., 1991; Solar et al., 1993) bacteria, as well as from archaea (Marsin and Forterre, 1999) and mitochondria (Backert et al., 1997; see also reviews by Khan, 2005; Ruiz-Masó et al., 2015). Recently, evidence of RCR has been obtained for some excised integrative and conjugative elements (ICE) (Burrus, 2017), and this replicative mechanism may also account for the maintenance of certain integrative and mobilizable elements (IME) while they are transiently excised from the chromosome (Guédon et al., 2017). Also ssDNA viruses of plants and animals are known to replicate by a RCR-like mechanism (Dolja and Koonin, 2011; Chandler et al., 2013). Basically, RCR initiation implies the sequence-specific cleavage of the non-template strand of a double-stranded DNA (dsDNA) form of the genetic element to be replicated. The cleavage reaction, which is catalyzed by a Rep protein encoded by the genetic element, generates a 5′-phosphate end covalently attached to the Rep protein (the so-called Rep-DNA adduct), and a free 3′-OH end that serves as a primer for the synthesis of the DNA using host cell enzymes. Hence, a singularity of the RCR is that the newly-synthesized DNA remains covalently linked to the cleaved parental strand during the entire elongation step. In all reported cases, a Tyr residue located in a conserved motif of the Rep protein is used as the nucleophile for the initial endonucleolytic cleavage of the parental DNA (Koonin and Ilyina, 1993; Chandler et al., 2013).

RCR in plasmids is subjected to an strict control that allows them to maintain their characteristic intracellular concentration (i.e., copy number) in a given host under defined growth conditions, thus reducing plasmid loss and metabolic burdening of the cell. Control of plasmid RCR is exerted by 1) self-encoded elements that inhibit the synthesis of the Rep initiator protein at the transcriptional and/or translational level, and 2) complementary mechanisms that avoid recycling and ensure inactivation of the Rep protein, so that each initiator molecule drives a single round of replication (reviewed in Ruiz-Masó et al., 2015).

Catalytic residues involved in the RCR initiation and termination of the leading strand have been best characterized in Rep proteins of the pT181 and pC194 plasmid families. The initiation cleavage at the specific nick site of one of the parental DNA strands is catalyzed by a Tyr residue located either in one of the subunits of a dimeric Rep protein (pT181 plasmid family) or in a monomeric Rep protein (pC194 plasmid family). In the members of the pT181 family, once the synthesis of the leading-strand has been completed, the catalytic Tyr of the second Rep subunit mediates a cleavage reaction on the reconstituted nick sequence and becomes covalently attached to the 5′-end of the newly-synthesized strand. On its turn, the 3′-OH end generated in this reaction acts as a nucleophile that attacks the phosphotyrosine bond of the first Rep subunit and thus leads to the generation of a circular ssDNA intermediate corresponding to the parental strand cleaved at the replication initiation step, which will serve as the template for the lagging-strand synthesis. The released catalytic Tyr of the first Rep subunit still mediates one more cleavage reaction on a second nick sequence that has been reconstituted by extension of the nascent leading strand by ~10 additional nucleotides beyond the nick site, and becomes covalently attached to this short oligonucleotide (Khan, 2005). This termination mechanism has two important implications for the regulation of the number of replication rounds driven by the same Rep dimer: 1) it prevents recycling (i.e., continuous rounds of leading-strand synthesis driven by the same Rep molecule), and 2) the heterodimeric Rep/Rep-DNA protein is inactive in replication initiation (Novick, 1998; Ruiz-Masó et al., 2015). In contrast, in plasmids of the pC194 family, a Glu residue of the monomeric Rep protein mediates the hydrolysis of the DNA at the specific site of the reconstituted nick sequence during the replication termination step. Hence, covalent linkage of Rep to the 5′-phosphate end of the newly-synthesized leading strand and subsequent recycling of the protein are prevented (Noirot-Gros et al., 1994). The 3′-OH end generated in the cleavage of the reconstituted nick sequence next attacks the adduct, thus releasing the free protein along with the circular ssDNA intermediate.

RCR plasmids of the pMV158 family are singular in that their Rep protein rather resembles the Rep initiators of small eukaryotic viruses: all of them form hexameric rings with a central electropositive channel that can lodge a ssDNA but not a dsDNA fragment (Enemark and Joshua-Tor, 2006; Boer et al., 2009, 2016). The atomic structure of the full-length RepB initiator encoded by pMV158 has been solved and remains to date the only published example among the Rep proteins from RCR plasmids and bacteriophages (Boer et al., 2009). The crystalographic analysis of RepB shows a hexameric molecule whose protomers consist of two domains, namely the N-terminal origin binding domain (OBD) and the C-terminal oligomerization domain (OD), conected by a short hinge region that allows multiple orientations of the OBDs relative to the rigid cylindrical scaffold constituted by the ODs. The OBD accounts for both the specific dsDNA binding ability and the

catalytic activities of the protein, whereas the OD is responsible for hexamerization (Boer et al., 2009). The OBD and OD domains of RepB have been purified separately as monomeric and hexameric proteins, respectively, and the OBD was shown to retain the DNA binding and catalytic activities of the protein (Boer et al., 2016). The RepB OBD domain belongs to the Rep class of the HUH endonuclease superfamily, which is featured by a His-hydrophobic-His (HUH) sequence motif involved in the coordination of the divalent metal ion of the active site (Chandler et al., 2013; Boer et al., 2016).

Although the mechanistic details of the reactions that take place in initiation and termination of the leading-strand synthesis of pMV158 (and related plasmids) remain still unveiled, they have to involve the initial cleavage of the parental strand at the nick site and the formation of a covalent adduct, most likely mediated by the Tyr99 residue located at the active center of the RepB protein (Boer et al., 2009). This adduct must last until the termination step, despite its *in vitro* identification has remained elusive for a long time (Moscoso et al., 1995; Moscoso et al., 1997). On the other hand, termination must involve at least a new cleavage reaction, catalyzed either by Tyr99 of a different protein subunit or by a second catalytic residue of the same subunit, as well as a strand-transfer reaction that seals the cleaved parental strand leading to the generation of the ssDNA replicative intermediate (del Solar et al., 1987). Both cleavage and strand-transfer reactions imply nucleophilic attacks, and hence could be expected to be sensitive to acidity, although very few studies have been performed about the effect of the pH on these reactions (Lovendahl et al., 2017). Moreover, despite the lack of information about the stability of Rep-DNA covalent adducts involving RCR initiators, the phosphoryl bond of phosphotyrosine residues of various proteins has been shown to be quite stable to basic conditions that give complete hydrolysis of peptide bonds (Martensen, 1982), and the phosphotyrosine linkage between *Escherichia coli* or *Micrococcus luteus* DNA topoisomerase I is stable in aqueous buffers ranging from pH 1 to 13 (Tse et al., 1980).

In this work, we have taken advantage of the great promiscuity of plasmid pMV158, which has been shown to replicate in the γ-proteobacterium *E. coli*, in several actinobacteria, and in many species from the Firmicutes phylum, including lactic acid bacteria (LAB) that can grow under very low pH conditions. Whereas the intracellular pH (pH_i) of neutralophilic hosts (like *E. coli*) is maintained close to neutral when the medium is acidified, many LAB have been reported to reduce their pH_i as the extracellular pH (pH_{ex}) decreases, thus generating smaller proton gradients across the membrane (Siegumfeldt et al., 2000). Studies on the homeostasis against acid stress in *Lactococcus lactis* subsp. *cremoris* MG1363 have revealed the relationship between the pH of the culture medium and the actual value of the pH_i in this model LAB strain (Mercade et al., 2000). This information has allowed us to test the effect of the pH not only on the *in vitro* reactions of cleavage, strand transfer and hydrolysis of the adduct, but also on the *in vivo* replication of pMV158 by checking its copy number and segregational stability in the lactococcal host.

MATERIALS AND METHODS

Bacterial Strains and Growth Conditions

L. lactis subsp. *cremoris*, strain MG1363 (Gasson, 1983) harboring plasmid pMV158 (*L. lactis* MG1363/pMV158) was cultivated routinely in M17 (Difco™) medium supplemented with yeast extract to a final concentration of 0.5% and with 0.5% glucose (M17′; Sánchez et al., 2008) at 30°C without aeration. When needed, tetracycline (Tc) was added at a concentration of 1 μg/ml. Growth was monitored with the aid of a Spectronic 20D+ equipment by checking the optical density (OD) at 660 nm.

For growing *L. lactis* MG1363/pMV158 at different pH conditions, the pH of the M17′ medium was adjusted with HCl to the values analyzed in this study: 7.0, 6.5, 6.0, 5.5, 5.0, and 4.5. A bacterial inoculum was prepared in M17 supplemented with 0.5% glucose and 1 μg/ml Tc, and grown statically at 30°C to an OD_{660} of 0.4. Several aliquots of this starter culture were taken, and the cells washed and resuspended in Tc-free M17′ adjusted to different pHs, to obtain cultures at an initial OD_{660} of 0.1. These cultures were incubated under the same conditions until the OD_{660} reached again 0.4, which corresponds to two generations of cellular growth. This procedure of cell collection, washing, dilution in antibiotic-free M17′ adjusted to the corresponding pH and incubation until OD_{660} of 0.4 was repeated for 20 generations of cellular growth, thus avoiding an excessive culture medium acidification due to the production of lactic acid by the bacteria. In fact, by following this procedure we only observed decreases of less than 0.3 units of pH in the cultures at OD_{660} of 0.4 respect to the initial pH value of the medium. Aliquots corresponding to 10 and 20 generations of cell growth at the different pHs were taken with two purposes: 1) extraction of genomic DNA (gDNA) for plasmid copy number (PCN) determination, and 2) analysis of the segregational stability of the plasmid by determination of the percentage of colony forming units (cfu) resistant to Tc (Tc^R). To estimate the percentage of Tc^R cfu, appropriate dilutions of culture samples were spread on M17-agar plates supplemented with glucose 0.5% and containing or not Tc.

Genomic DNA Preparations

Lactococcal gDNA, used as template for real-time quantitative PCR (qPCR), was isolated as described in (Garay-Novillo et al., 2019). Purified gDNA was digested with EcoRI, which linearizes the pMV158 plasmid DNA but leaves intact the template DNA segments containing the plasmidic and chromosomal amplicons that will be amplified in the qPCR assays. This method was developed to obtain accurate qPCR-based copy number results for plasmids (Providenti et al., 2006). Concentration of the gDNA was determined with a Qubit fluorometer by using the Qubit HS dsDNA Assay Kit (Molecular Probes). DNA integrity was examined by 0.8% agarose gel electrophoresis and DNA staining with GelRed (Biotium). Irrespective of the pH of the medium, the gDNA yield was ~10 μg from 1 ml of culture at OD_{660} of 0.8. Hence, similar intracellular concentrations of chromosomal DNA were inferred for lactococcal cells grown at the different pHs.

Determination of PCN in *L. lactis* by qPCR and Agarose Gel Quantification

For the determination of the PCN per chromosomal equivalent by qPCR, two primer pair sets specific to the PcrA helicase single-copy reference gene (*pcrA*) of *L. lactis* MG1363 (Wegmann et al., 2007) and to the RepB replication protein gene (*repB*) of pMV158 were used (Garay-Novillo et al., 2019). Criteria used during primer design were that primers had a predicted Tm of ~59°C and that they generated amplicons of ~140 bp in length. qPCRs were conducted as described (Ruiz-Masó et al., 2017). Diluted preparations of EcoRI (Nzytech)-digested gDNA (5, 1.7, and 0.17 ng per reaction) were analyzed using 0.5 μM of the specific forward and reverse primers of the primer pair used. Three independent qPCR trials were conducted for each template source. In each trial, triplicate samples of the three different amounts of template were analyzed. PCN of pMV158 at the different pH conditions studied was calculated as described (Garay-Novillo et al., 2019).

For PCN determination by agarose gel quantification, gDNAs samples were loaded into 0.8% agarose gels, run for 60 min at 6.5 V/cm and dyed with GelRed for 10 min. Gels were visualized with the aid of a Gel Doc (Bio-Rad). Non-saturated images of the gels were analyzed with the Quantity One v4.5.2 software (Bio-Rad). Plasmid bands corresponding to the covalently-closed circular (CCC) and open circular (OC) forms and chromosomal DNA band in each lane were identified and their densities analyzed. For estimation of the pMV158 PCN in cells grown at different pHs, the ratio of the signal of plasmid (CCC × 1.36 + OC; Projan et al., 1983) to the signal of chromosome was quantitatively compared with that of *L. lactis*/pMV158 cells grown at pH 7.0, whose value was taken as a reference to which the PCN determined by qPCR was assigned.

Protein Purification

Hexameric full-length RepB ($RepB_6$) and His-tagged OBD (hereafter termed OBD) were purified as described previously (Ruiz-Masó et al., 2004; Boer et al., 2009). Protein concentration was measured by two methods: 1) spectrophotometrically using the theoretical molar absorption coefficient, and 2) SDS-PAGE and quantitative densitometry. In the latter method, gels were stained with Coomassie Brilliant blue and the densitometric value of the specific protein band was quantitatively compared with that of a $RepB_6$ or OBD sample of known concentration.

Identification of the RepB-DNA Covalent Adduct

The formation of a covalent complex between RepB and the 5′-phosphate end of the cleaved DNA substrate was analyzed by using $RepB_6$ and His-tagged OBD proteins. For the detection assays, 5 pmol of the 27-mer oligo substrate (5′-TGCTTCCGT ACTACG/ACCCCCCATTAA-3′; where "/" indicates the RepB nick site) fluorescently labeled with Cy5 at its 3′ end was mixed with 7.5 pmol of the protein. The reaction was performed in 20 μL of reaction buffer 8 (20 mM Tris-HCl pH 8.0) supplemented with 300 mM NaCl and 2 mM $MnCl_2$. The mixture was incubated for 15 s or 5 min at 37°C, and SDS was subsequently added to a final concentration of 0.2%. Prior to electrophoresis, the samples were mixed with the appropriate volume of 10× DNA loading buffer (60% glycerol and 10 mM EDTA) and heated at 95°C for 3 min. The products were separated on 10% PAA (19:1 acrylamide:bis-acrylamide), 0.3% SDS denaturing gels. After electrophoresis, the gels were analyzed by using a FLA-3000 (FUJIFILM) imaging system (λ_{ex} 633 nm; λ_{em} 675 nm), and the fluorescence intensity of each band was quantified with the Quantity One software (Bio-Rad). After visualization of the fluorescent reaction products, the same gel was used for western blotting analysis with RepB antiserum. For that purpose, a Trans-Blot® Turbo™ system was employed and a conventional semi-dry blotting protocol was followed for western blotting. Prior to transfer, the gel and the PVDF membrane were equilibrated in Towbin buffer (25 mM Tris, 192 mM glycine pH 8.3, 20% methanol) for 10 min. Once the blotting sandwich was assembled, the transfer was carried out by applying 25 V constant voltage for 30 min. Then, the membrane was washed with Tris-buffered saline (50 mM Tris-HCl pH 7.5, 150 mM NaCl) containing 0.1% Tween-20 (TTBS) and incubated with the appropriate dilution of the primary RepB antiserum in TTBS supplemented with 0.2% casein. After at least 2 h of incubation with gentle agitation, the membrane was washed with TTBS and further incubated for 2 h with the appropriate dilution of the peroxidase-conjugated anti-rabbit IgG secondary antibody. Finally, the membrane was washed with TTBS to remove the remaining secondary antibody, and incubated with an ECL western susbtrate for chemiluminescence detection. The membrane was visualized with the aid of a chemiluminescent detector Fuji LAS-3000.

Analysis of the Nicking and Strand-Transfer Activities on Single-Stranded Oligonucleotides.

For the cleavage assays, 5 pmol of the 27-mer oligo substrate fluorescently labeled with Cy5 at its 3′ end was mixed with 150 pmol of an unlabeled 30-mer oligo (5′-TACTGCGGAATT CTGCTTCCATCTACTACG-3′) that provides the 3′-OH substrate for strand transfer. The mixture was incubated for 1 min at 25°C with OBD (0.1 pmol) in 20 μL of the appropriate reaction buffer. For reactions at pH 8.0 and 7.0, reaction buffers 8 and 7 (20 mM Tris-HCl adjusted to pH 8.0 and 7.0, respectively) supplemented with 300 mM NaCl, 2 mM $MnCl_2$ and 0.2 mg/ml BSA, were used. For reactions at pH 6.5, 6.0, 5.5, 5.0, and 4.5, reaction buffers 6.5, 6, 5.5, 5, and 4.5 contained 20 mM MES adjusted to the indicated pH value and supplemented with 300 mM NaCl, 2 mM $MnCl_2$ and 0.2 mg/ml BSA. Protein samples were previously diluted in 20 mM Tris-HCl buffer (pH 8.0) supplemented with 430 mM NaCl and 0.2 mg/ml BSA. The presence in the reaction mixtures of a 30-fold molar excess of the 30-mer oligo relative to the 27-mer oligo makes rejoining of the cleaved 27-mer oligo highly unlikely. After incubation for 1 min at 25°C, the reaction mixtures were

treated with proteinase K (0.1 mg/ml) and 0.2% SDS for 15 min at 37°C.

To evaluate the strand-transfer activity of OBD at different pHs, the protein and the 27-mer oligo substrate were incubated at a 1:1 molar ratio in reaction buffer 7 at 25°C for 1 min. Then, the reaction mixture was diluted 10 times in buffers 7, 5, and 4.5, and an aliquot of this dilution was taken as a reference. Next, the unlabeled 30-mer oligo was added at a 50-fold molar excess relative to the initial 27-mer oligo and the strand-transfer reaction proceeded for 5 min at 25°C. After incubation, an aliquot of the three reactions mixtures was taken and, together with the reference aliquot, treated with SDS and proteinase K as stated above. Prior to electrophoresis, the samples were mixed with an appropriate volume of 10× DNA loading buffer and denatured by heating at 95°C for 3 min. The products were separated on 20% PAA (19:1 acrylamide:bis-acrylamide), 8 M urea denaturing gels. After electrophoresis, the fluorescent reaction products were visualized as depicted in **Identification of the RepB-DNA Covalent Adduct**.

Analysis of the Effect of Temperature, Depletion of Divalent Cations and pH on the Stability of the OBD-DNA Covalent Adduct.

The OBD-DNA adduct was formed by mixing 5 pmol of the fluorescently labeled 27-mer oligo substrate with 7.5 pmol of OBD in 10 μL of buffer 7. The mixture was incubated for 2 min at 25°C and subsequently diluted 30 times with the same reaction buffer equilibrated to the reaction temperature. To analyze the effect of the temperature, diluted reactions were incubated at 0°C, 22°C or 37°C for up to 1 h, and aliquots (10 μL) were taken at intervals. The effect of the depletion of Mn^{2+} ions was analyzed by treating the diluted reaction with 20 mM EDTA and incubating the mixture at 22°C, with 10-μL aliquots being withdrawn at intervals over a 1-h period. The stability of the covalent adduct at different pHs was analyzed by a similar procedure. The covalent complex was formed in buffer 7 as described above. After being treated with 10 mM EDTA for 1 min at 25°C, the reaction mixture was diluted 30 times in the corresponding reaction buffer adjusted to the pH of analysis (buffers 7, 6, 5.5, 5 or 4.5) and incubated 5 min at 4°C. Next, the mixture was incubated at 25°C and 10-μL aliquots were withdrawn at intervals over a 1-h period. In all cases, an initial $t = 0$ aliquot was also taken. All aliquots were treated with proteinase K (0.1 mg/ml) and 0.2% SDS for 15 min at 37°C. Samples were analyzed by electrophoresis and the gels visualized as depicted in **Identification of the RepB-DNA Covalent Adduct**.

The half-life time ($t_{1/2}$) of the covalent complex, defined as the time it takes for a 50% reduction of the adduct fraction, was estimated according to the equation:

$$[adduct_t] \Big/ [adduct_{t_0}] = 1 \Big/ 2^{t/t_{1/2}} \quad (1)$$

where $[adduct_t]$ and $[adduct_{t_0}]$ represent the fraction of the nicking products in the form of adduct at time t and $t = 0$, respectively.

This equation can be converted into a linear function by taking logarithms:

$$\log([adduct_t]/[adduct_{t_0}]) = -tlog2/t_{1/2} \quad (2)$$

The $\log([adduct_t]/[\text{adduct}_{t0}])$ was plotted against time (t), and the experimental data were fitted to **Eq. 2** by linear regression. The term $-log2/t_{1/2}$ represents the slope (m) of the fit, so that the half-life time ($t_{1/2}$) can be calculated from the expression:

$$t_{1/2} = -log2/m \quad (3)$$

Preparation of Total Protein Extracts and Detection of RepB by Western Blotting

L. lactis MG1363/pMV158 was grown for 10 and 20 generations in Tc-free M17′ adjusted to pH 7.0, 5.0 or 4.5, and aliquots (4 ml) of the cultures were taken when they reached an OD_{660} of 0.4. Cells were harvested by centrifugation and resuspended in 300 μL of lysis solution (50 mM Tris-HCl pH 7.0, 0.3% SDS, 200 mM DTT, 3 mM $MgCl_2$, 16 μg/ml of protease-free RNase I and 60 μg/ml of protease-free DNase I). Cell suspension was transferred to 2-mL tubes containing 0.1 mm silica spheres lysing matrix (FastPrep®) and the cells were lysed by beating the tubes at 6 m/s for 45 s at 25°C with the aid of a Hybaid Ribolyser homogenizer. The cell debris was removed by centrifugation at 14,000 × g for 20 min at 4°C. Total protein concentration of the extracts was determined with a Qubit fluorometer by using the Qubit Protein dsDNA Assay kit (Molecular Probes). Similar total protein concentrations (maximum difference of 11%) were obtained across the extracts. Samples were analyzed by electrophoresis on 12% SDS-glycine-PAA gels and transferred to a PVDF membrane for western blotting analysis with RepB antiserum as described in **Identification of the RepB-DNA Covalent Adduct**. Images of the membrane were analyzed with the Quantity One (Bio-Rad) software and the luminescent bands corresponding to the RepB protein were quantified.

The total protein amount loaded for western blotting was controlled in two different ways: 1) by adjusting the volume of protein extracts loaded in the same gel in order to keep constant the amount of total protein in the different samples analyzed; and 2) by normalization to an isolated non-specific band (internal control) that corresponds to a chromosome-encoded protein whose concentration remained rather constant across the different extracts. Similar results were obtained when comparing the RepB levels between protein extracts from cells grown at different pHs, irrespective of whether normalization to the internal control band was made or not.

Statistical Analysis

Statistical analyses were carried out using SigmaPlot software (v12.5, Systat software, San Jose, CA, United States). One-way ANOVA tests were conducted to analyze the effect of pH on the OBD endonuclease activity, the PCN and the bacterial growth rate, as well as the effect of pH and temperature on the adduct half-life time. Student t-tests were used to compare the relative

RepB level and the RepB to pMV158 ratio in *L. lactis* MG1363/pMV158 cells grown at different conditions with respect to those of the reference condition (10 generations at pH 7.0). The level of significance in both statistical analyses was set at $p < 0.05$.

RESULTS

RepB-mediated DNA cleavage at the nick site renders a transient covalent nucleoprotein complex.

Direct verification of the existence of a RepB-DNA covalent intermediate in DNA cleavage and rejoining has largely remained elusive and only indirect indications could be reported for a long time (Moscoso et al., 1995; Moscoso et al., 1997). Recently, we succeeded in detecting a product arising from the RepB endonuclease activity at the nick site that corresponded to the 3′-fragment of a 3′-end fluorescently labeled substrate oligonucleotide and exhibited a decreased electrophoretic mobility relative to the corresponding protein-free DNA product on 20% PAA-urea gels. This slowly migrating product could only be observed when the temperature and time of incubation of the substrate DNA with either $RepB_6$ or OBD were reduced with respect to previously-reported conditions (from 60°C to ≤37°C and from 30 min to ≤15 min, respectively), and after treatment of the reaction products with 0.2% SDS and 0.1 mg/ml proteinase K (Ruiz-Masó et al., 2016). The anomalous migration of the labeled reaction product (which actually appears as a set of 2–3 bands) was hypothesized to arise from a small protein polypeptide that remained bound to the 3′-fragment generated upon cleavage of the substrate oligonucleotide by the $RepB_6$ or OBD protein. The discrete bands assigned to this putative covalent complex were shown to result, at least in part, from partial digestion of the initiator protein with proteinase K (not shown). Moreover, we observed that once the substrate utilization had gone to completion, further incubation of the reaction mixture resulted in the decay of the slowly migrating DNA product and the concomitant rise of its protein-free counterpart (see below). All these observations suggested that, upon RepB-mediated specific cleavage at the nick site, a covalent link between the protein and the 5′-phosphate end of the cleaved DNA was formed that dissociated rapidly.

To ascertain the nature of the transient nucleoprotein complex resulting from the RepB-mediated cleavage of the 27-mer oligo substrate (which contains the nick sequence and is fluorescently labeled at its 3′-end), the reaction products were subjected to two different analyses. The reactions were performed by incubating the 27-mer oligo DNA with a 1.5-fold molar excess of either $RepB_6$ (consisting of six 24.12 kDa polypeptide chains) or OBD (a single 16.4 kDa polypeptide chain) at 37°C for the indicated times. After stopping the reactions with 0.2% SDS, the mixtures were processed as described (**Identification of the RepB-DNA Covalent Adduct** in **Materials and Methods**). Treatment with SDS denatures proteins, dissociates the oligomeric structures and confers uniform negative charge densities to the polypeptides, thus allowing the electrophoretic separation of covalent nucleoprotein complexes according to their molecular sizes. The fluorescent DNA species were visualized with the Phosphorimager FLA-3000 system (**Figure 1A**). In addition to the 27-mer oligo substrate and to the 12-mer oligo (the protein-free fluorescently-labeled final product of the $RepB_6$/OBD-mediated cleavage), slower-migrating fluorescent DNA species were observed whose electrophoretic mobilities were inversely related to the polypeptide chain length of the protein employed ($RepB_6$ or OBD; **Figure 1A**). We speculated that these retarded bands could correspond to covalent complexes between the target RepB or OBD protein and the fluorescent 12-mer DNA generated by cleavage of the 27-mer substrate oligonucleotide. To confirm this hypothesis, once the fluorescent bands corresponding to DNA and nucleoprotein complexes were imaged (**Figure 1A**), free RepB/OBD proteins and RepB/OBD-DNA complexes in the gel were detected by chemiluminescent western blotting employing anti-RepB polyclonal serum as the primary antibody and peroxidase-conjugated anti-rabbit IgG as secondary antibody (**Figure 1B**). Chemiluminescent signals appeared linked to the fluorescent DNA species that exhibited retarded mobility, which were thus confirmed to correspond to covalent (SDS-resistant) complexes between the 12-mer oligo and either RepB or OBD (see reaction scheme in **Figure 1C**). Apart from these nucleoprotein complexes, chemiluminescent bands that also displayed electrophoretic mobilities inversely related to the length of the target protein polypeptide chain were observed in samples containing $RepB_6$ or OBD, irrespective of the presence or absence of DNA (either the 27-mer substrate or the 12-mer product oligo), and were ascribed to the DNA-free polypeptides of the corresponding protein (**Figure 1B**).

Given that the OBD domain retains the specific endonuclease and strand-transfer activities of the RepB protein, we chose to employ OBD, which is produced more easily due to the presence of the His-tag, in all subsequent experiments throughout this work.

Low pH Impairs the Endonuclease Activity of OBD

The DNA cleavage activity of OBD was assayed in the presence of a 50:1 molar ratio of the 27-mer oligo substrate to the enzyme. To avoid regeneration of the substrate oligo by the strand-transfer activity of the protein, which would entail an underestimation of its cleavage activity, a 30-fold molar excess of the unlabeled 30-mer oligo (carrying the nick sequence 5′ to the nick site) relative to the fluorescently-labeled 27-mer oligo was added to the reaction mixture. In this way, virtually the only strand-transfer product should correspond to the 42-mer oligo that arises from joining of the 30-mer oligo with the 3′-fluorescently labeled 12-mer cleavage product and is thus distinguishable from the substrate 27-mer oligo (**Figures 2A,C**).

In previous works, all assays of the catalytic activity of the RepB protein or its OBD domain had been carried out at a pH value of 8.0. This pH was chosen based on early reported conditions used for analyzing the activity of the plasmid pT181 Rep protein (Koepsel et al., 1985), which, like RepB, catalyzes DNA cleavage and rejoining reactions mediated by nucleophilic attacks. In the present work, we aimed to analyze the pMV158 replication capacity and the RepB catalytic activities

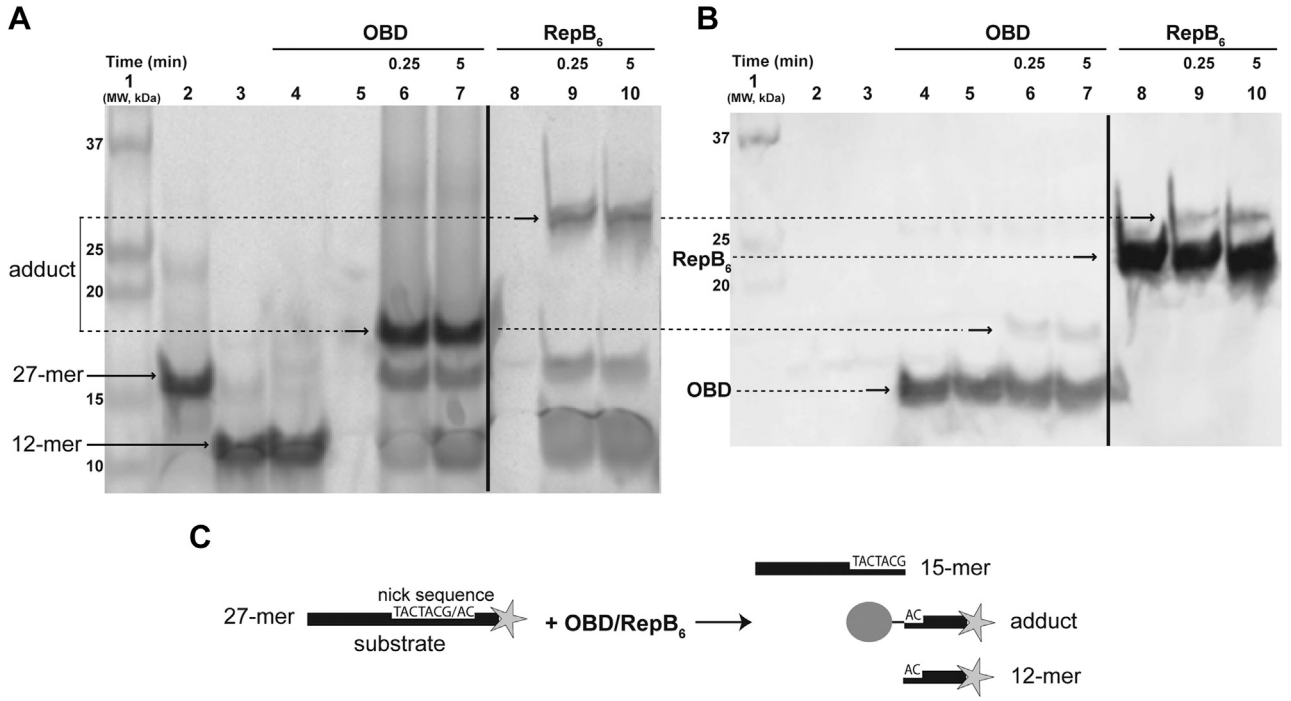


FIGURE 1 | Detection of the RepB-DNA covalent complex. **(A)** Direct visualization of the fluorescent DNA species that arise from the endonuclease activity of $RepB_6$ and OBD on the ssDNA 27-mer oligo substrate. A prestained and chemiluminescent protein ladder was loaded in lane 1. Lanes 2 and 3 show the migration of fluorescent 27-mer substrate and chemically-synthesized 12-mer oligo, respectively. In lane 4, a mixture of OBD and chemically-synthesized fluorescent 12-mer oligo (which is not a substrate of the protein) was loaded. OBD (lanes 6 and 7) and $RepB_6$ (lanes 9 and 10) were incubated for the indicated times with the 27-mer substrate. The slower migrating bands corresponding to covalent complexes between $RepB_6$ or OBD and the fluorescent 12-mer DNA generated by cleavage of the 27-mer substrate are indicated. The dividing line indicates grouping of different parts of the same gel. **(B)** Detection of $RepB_6$ and OBD in the gel shown in panel A by chemiluminescent western blotting employing anti-RepB polyclonal serum. Lanes 5 and 8 contain OBD and $RepB_6$ proteins, respectively. The position of the slower migrating chemiluminescent bands that correspond to covalent complexes between $RepB_6$ or OBD and the 12-mer oligo is indicated. The presence of DNA-free protein subunits was also detected and their position indicated. The dividing line indicates grouping of images of the same gel as indicated in panel **(A)**. **(C)** Schematic description of the endonuclease reaction mediated by $RepB_6$ or OBD.

in a range of pHs, varying from neutral (7.0) to acidic (4.5) values, at which LAB usually grow (Schillinger et al., 2006). Thus, we first compared the endonuclease activity of OBD at pHs 8.0 and 7.0, and did not found significant differences between both conditions (not shown). All subsequent *in vivo* and *in vitro* experiments were performed at pH values of 7.0, 6.5, 6.0, 5.5, 5.0, and 4.5, and the data obtained were normalized with respect to those obtained at pH 7.0, to which 100% activity was assigned.

The endonuclease activity of OBD clearly decreased with decreasing pH values, the activity at pH 4.5 becoming ~16% of that observed at pH 7.0. Statistically significant decreases were found between all pH values, except for pHs 6.5 and 6.0 (**Figure 2B**).

The Strand-Transfer Activity of OBD is Not Significantly Affected Over a Wide pH Range.

To test the strand-transfer activity of the RepB OBD domain, the protein and the 3′-fluorescently labeled 27-mer oligo substrate were incubated at a 1:1 molar ratio, in reaction buffer 7, at 25°C, until reaction products (mainly consisting of the covalent adduct) reached 80–90% of total DNA. After diluting 10 times the reaction mixture in buffers at each of the pHs studied (7.0, 5.0, and 4.5), the unlabeled 30-mer oligo was added at a 50-fold molar excess relative to the initial fluorescently-labeled 27-mer substrate and the strand-transfer reaction was allowed to proceed for 5 min at 25°C. Similar ratios of the 42-mer product to the OBD-DNA adduct were obtained irrespective of the pH, showing that the strand-transfer activity of the protein is not significantly affected by the acidic conditions of the reaction mixture (**Figure 3**).

Preservation of the strand-transfer activity of the RepB OBD domain even at the most acidic conditions analyzed suggests that the active center of the protein shows structural stability in the pH range from 8.0 to 4.5.

Low pH Improves the Stability of the OBD-DNA Covalent Adduct.

To accurately determine the influence of the pH in the stability of the covalent link between OBD and the 5′-phosphoryl group generated upon cleavage of the oligo substrate, we first tested the optimal temperature conditions for these experiments. To this end, reaction mixtures that contained a 1.5:1 molar ratio of OBD

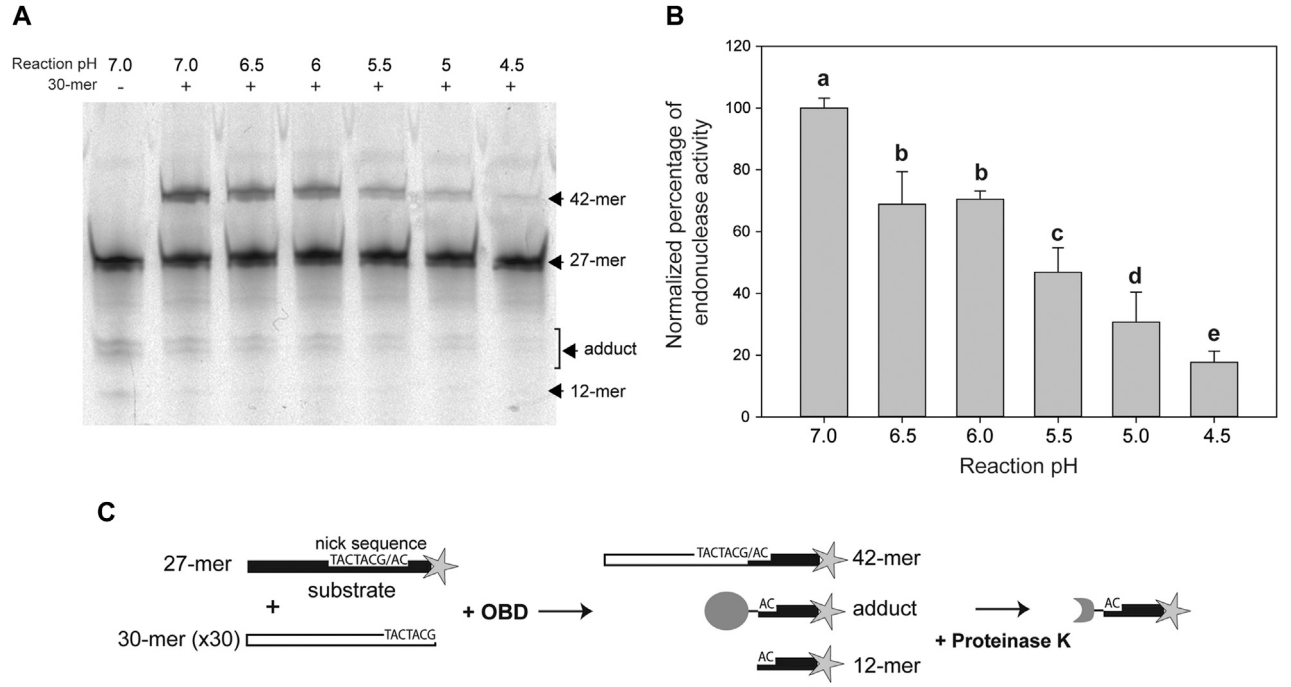


FIGURE 2 | Endonuclease activity of OBD at different pHs. **(A)** Reaction product pattern generated by the nicking and strand-transfer activities of OBD on ssDNA oligos at the indicated reaction pHs. The 27-mer oligo substrate (500 nM), labeled at its 3′-end with the fluorescent dye Cy5 (indicated by a star in panel **(C)**), and a 30-fold molar excess of the unlabeled 30-mer oligo were incubated with the protein (10 nM) at 25°C for 2 min. The resultant fluorescent oligos were analyzed by electrophoresis in 20% PAA-urea gels, and visualized with the aid of a FLA 3000 (FUJIFILM) imaging system. A representative gel from one of four independent experiments is displayed. The bands corresponding to the 12-mer DNA generated by cleavage of the 27-mer substrate and to the 42-mer DNA strand-transfer product are indicated. The covalent adduct is represented by a set of 2–3 bands that arise from partial digestion of OBD with proteinase K. **(B)** Vertical bar graph showing the percentage of reaction products rendered by OBD at the indicated pHs. The sum of all reaction products was considered for catalytic activity quantification. The percentage of activity obtained at pH 7.0 was used as a reference for normalization. Vertical bars represent the average value of four independent experiments. Error bars represent standard deviations. Average values with different lowercase letters were significantly ($p < 0.05$) different. **(C)** Schematic description of the different reaction products.

to the 3′-fluorescently labeled 27-mer oligo were prepared in buffer 7 and incubated at 25°C for 2 min, so that the oligo substrate was almost completely (>80%) exhausted. Then, the mixtures were diluted 30 times with the same buffer 7 and further incubated at either 0°C, 22°C or 37°C to analyze the change in the adduct fraction over time. A sample of the diluted mixture was brought to 20 mM EDTA and incubated at 22°C in order to study whether depletion of Mn^{2+} affected the stability of the OBD-DNA covalent linkage. Moreover, the effect of denaturing the protein on the stability of the adduct was analyzed by adding 0.2% SDS to a sample of the freshly-diluted mixture and incubating at 22°C for 24 h. The half-life time of the adduct, which measures its stability, was calculated according to **Eq. 3**. Hydrolysis of the adduct was found to be prevented in the presence of SDS (**Figure 4A**), indicating that this process is catalyzed by the protein and requires the native folding of the OBD. The stability of the covalent adduct clearly decreased with increasing temperatures, the half-life at 0°C (~220 min) being about 2.5- and 10-fold longer than those determined at 22°C (~84 min) and 37°C (~25 min), respectively (**Supplementary Figure S1**). On the other hand, chelation of Mn^{2+} did not affected significantly the stability of the adduct (**Supplementary Figure S1**). We finally chose to investigate the effect of the pH on the stability of the OBD-DNA adduct at a temperature of 25°C, which can be easily controlled in a heating block and was expected to confer a medium stability to the covalent nucleoprotein complex.

To test the effect of the pH on the stability of the OBD-DNA adduct, reaction mixtures containing a 1.5:1 molar ratio of OBD to the 3′-fluorescently labeled 27-mer oligo were prepared in buffer 7, incubated at 25°C for 2 min and treated with 10 mM EDTA before being diluted 30 times in buffer 7, 6, 5.5, 5 or 4.5 for further incubation at the same temperature (**Figures 4A,B**). A remarkable and statistically significant stabilization of the OBD-DNA adduct was observed as the pH was decreased, with the half-life of the covalent complex increasing from ~1 h (at pH 7.0) to more than 6 h (at pH 4.5) (**Figure 4**). Hence, it seems that acidity greatly impairs hydrolysis of the adduct catalyzed by the protein.

The PCN is Inversely Related to the Growth Rate of *L. lactis* MG1363/pMV158 Cells That Grow in Tc-Containing Media Adjusted to Different pHs.

Plasmid pMV158 has been previously shown to be stably inherited and not to cause any statistically significant decrease

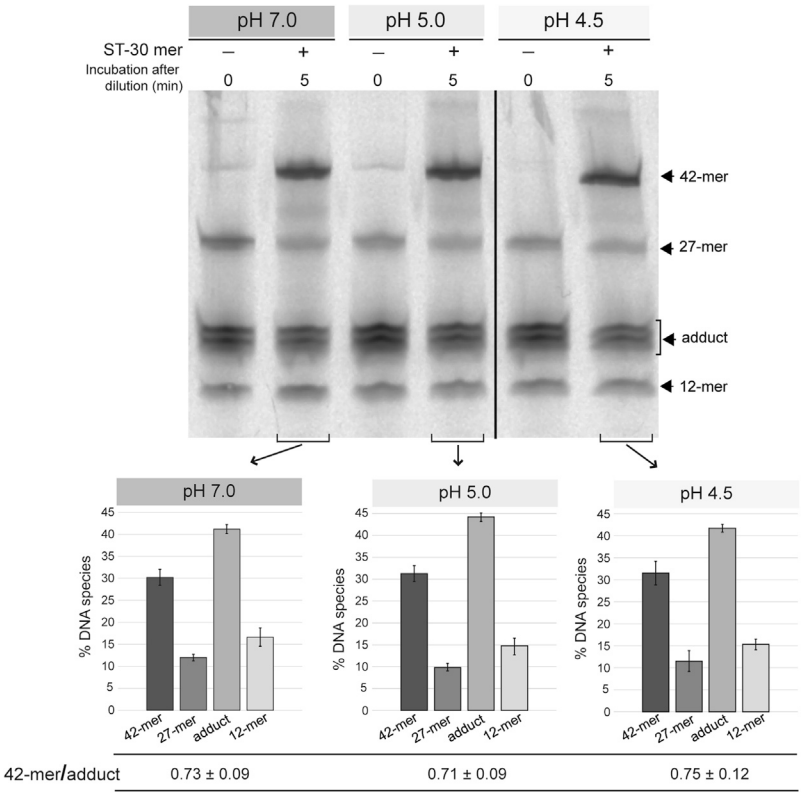
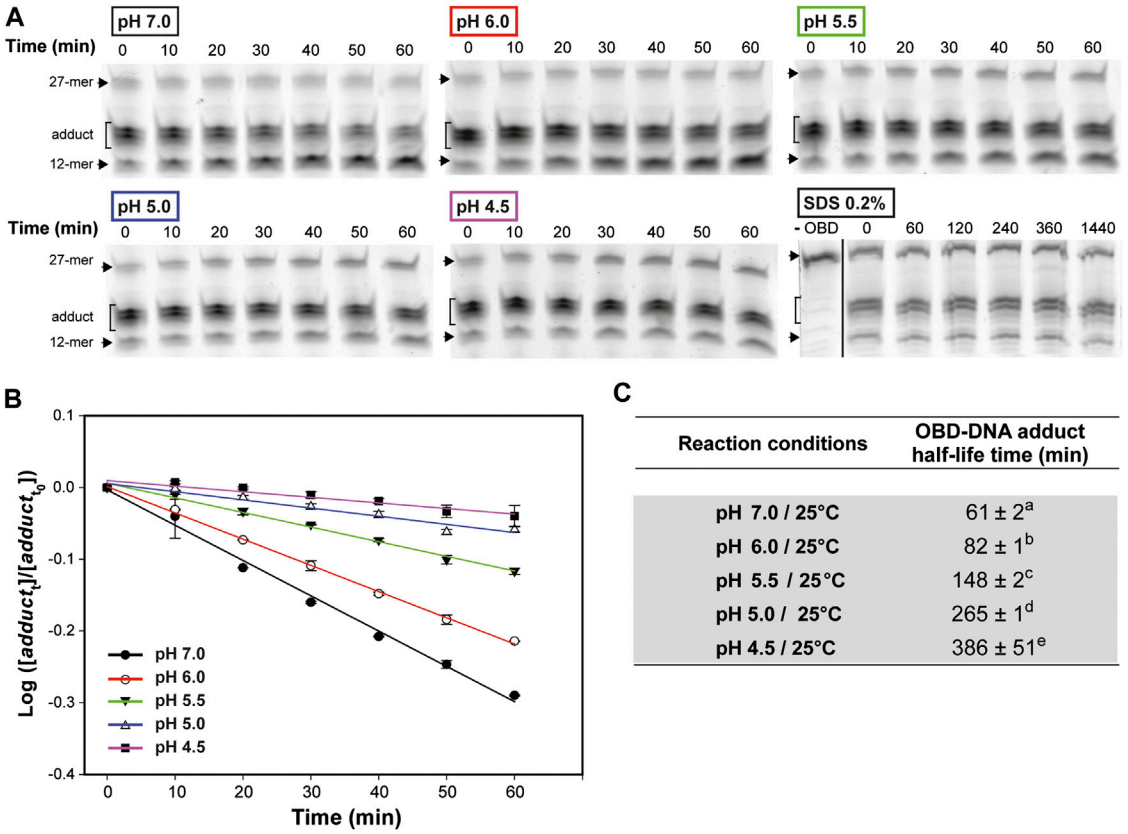


FIGURE 3 | Strand-transfer activity of OBD at different pHs. Gel showing the reaction products that arise from the strand-transfer activity of OBD at different pHs. After setting up a nicking reaction (OBD:27-mer molar ratio of 1:1) in reaction buffer 7 at 25°C for 1 min, the mixture was diluted 10 times in buffers 7, 5.0, and 4.5. Next, the unlabeled 30-mer oligo was added to a 50-fold molar excess relative to the initial 27-mer oligo and the strand-transfer reaction was allowed to proceed for 5 min under the same conditions used for the nicking reaction. Aliquots of the three reactions were taken and analyzed. Aliquots of the diluted reactions before adding the 30-mer oligo were also analyzed and used as reference. The bands corresponding to the 12-mer DNA generated by cleavage of the 27-mer substrate and the 42-mer strand-transfer product are indicated (see reaction scheme in **Figure 2C**). The covalent adduct is represented by a set of 2–3 bands that arise from partial digestion of OBD with proteinase K. A representative gel of three different experiments is displayed. Below the gel image, three bar graphs represent the quantification of the different DNA species generated at the three reaction pHs analyzed. Vertical bars represent the average value of three different experiments and error bars represent standard deviations. The ratio (mean ± SD) between the percentages of the 42-mer product and the adduct for each quantification is also shown.

in the growth rate of *L. lactis* MG1363 host cells growing in M17 medium supplemented with 4% glucose. Conversely, when the lactococcal host cells were grown in M17 supplemented with 4% glucose and 1 μg/ml Tc (the resistance to which is encoded by pMV158), the bacterial growth rate underwent a statistically significant 25% reduction (Garay-Novillo et al., 2019). Since expression of the plasmid-encoded *tetL* gene has been reported to be constitutive (Lacks et al., 1986), the decrease in the growth rate of the bacterial host was attributed directly to the presence of the antibiotic, which was thought to cause some noxious effect before being expelled from the cell. In the lactococcal host cells grown under these latter conditions (i.e., in M17 plus glucose and Tc), pMV158 exhibited ~18 copies per chromosome equivalent (Garay-Novillo et al., 2019). The copy number of pMV158 may vary, however, between host cells growing in different media, as it has been reported that the PCN shows a strong dependence on the bacterial growth rate and that, in general, the plasmid concentration decreases with increasing growth rates (Summers, 1996; Klumpp, 2011; Ruiz-Masó et al., 2017). For this reason, we first determined the PCN and the bacterial growth rate in *L. lactis* MG1363/pMV158 cells cultured in M17′ adjusted to different pHs and supplemented with 1 μg/ml Tc.

The growth rate of the lactococcal host decreased abruptly as the acidity of the culture medium was increased, and approached to 0 at pH 4.5 ($\mu = 0.05 \pm 0.02$; **Table 1**). This growth slowdown at the lowest pHs was much more pronounced than that previously reported for the plasmid-free *L. lactis* MG1363 cells, whose growth rate at pH 4.5 was only decreased by ~50% with respect to pH 6.5 (Mercade et al., 2000). Since pMV158 does not seem to burden significantly the lactococcal host by itself (Ruiz-Masó et al., 2017), we speculate that the strong pH-dependence of the growth rate of the *L. lactis* MG1363/pMV158 cells growing in the presence of Tc arises from the antibiotic efflux mechanism mediated by the plasmid-encoded TetL protein, which exchanges Tc-divalent metal complexes for protons (Krulwich et al., 2001). Hence, the activity of the pMV158-encoded Tc^R determinant could further increase the intracellular acidity of the lactococcal host grown at low pHs,



C

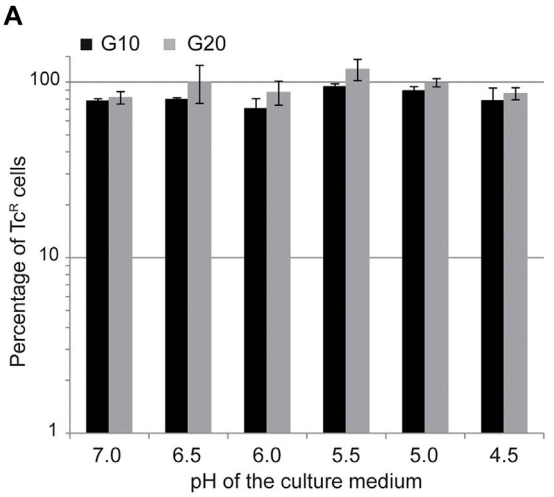
Reaction conditions	OBD-DNA adduct half-life time (min)
pH 7.0 / 25°C	61 ± 2[a]
pH 6.0 / 25°C	82 ± 1[b]
pH 5.5 / 25°C	148 ± 2[c]
pH 5.0 / 25°C	265 ± 1[d]
pH 4.5 / 25°C	386 ± 51[e]

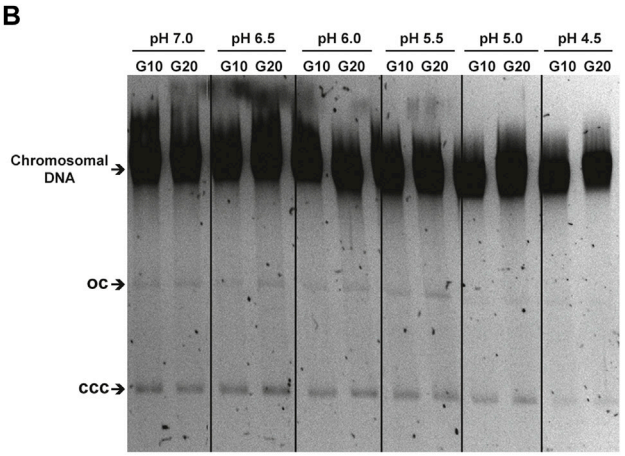
FIGURE 4 | OBD-DNA adduct stability at different pHs. **(A)** Gels showing the effect of pH on the time course of the adduct-hydrolysis reaction. The reaction mixtures contained a 1.5:1 molar ratio of OBD to the 3′-fluorescently labeled 27-mer oligo substrate and were prepared in buffer 7. After 2 min of incubation at 25°C, the endonuclease reaction was stopped with 10 mM EDTA before diluting the mixtures in buffer 7, 6, 5.5, 5 or 4.5 for 1-h incubation at the same temperature. Several aliquots were taken every 10 min and analyzed in 20% PAA-urea gels. A gel showing the effect of denaturing the protein on the stability of the adduct is also shown. In this case, the endonuclease reaction was performed in buffer 7 for 2 min at 22°C, and then the mixture was diluted in buffer 7, treated with 0.2% SDS and incubated at the same temperature for 24 h. Several aliquots were taken at the indicated times and analyzed in 20% PAA-urea gels. A control lane (-OBD) showing the migration of the fluorescent 27-mer substrate is included. The dividing line indicates grouping of different parts of the same gel. **(B)** Kinetics of the hydrolysis of the OBD-DNA adduct at different pHs. The logarithm of the ratio between the fraction of adduct at each of the time intervals analyzed ($adduct_t$) and the initial fraction ($adduct_{t0}$) was plotted against time, and the experimental data were fitted to **Eq. 2** by linear regression. The OBD-DNA adduct half-life time was calculated from the slope of the regression curve according to **Eq. 3**. Plots display the adduct decay curves obtained from at least five independent experiments for each pH, with the symbols and vertical bars representing the mean and standard deviation, respectively. **(C)** Table showing the adduct half-life time calculated from at least five independent experiments performed under the indicated reaction conditions. The values are expressed as mean ± SD. Values with different superscript letters were significantly ($p < 0.05$) different.

TABLE 1 | pMV158 PCN and growth rate of *L. lactis*/pMV158 cells grown in Tc-containing M17' adjusted to different pH values.

M17′ pH	μ (h^{-1}) M17′ + Tc (mean ± SD)	pMV158 copy number (mean ± SD)	
		Gel quantification	qPCR quantification
7.0	0.58 ± 0.06$^{\alpha}$	7.2^{A}	7.2 ± 2.0^{a}
6.5	0.34 ± 0.02$^{\beta}$	9.6 ± 0.8^{A}	9.5 ± 1.0^{a}
6.0	0.30 ± 0.03$^{\beta,\gamma}$	10.5 ± 1.7^{A}	11.8 ± 1.2^{a}
5.5	0.24 ± 0.03$^{\gamma}$	23.4 ± 3.0^{B}	23.5 ± 0.8^{b}
5.0	0.15 ± 0.03$^{\delta}$	41.3 ± 4.4^{C}	43.1 ± 2.3^{c}
4.5	0.05 ± 0.02$^{\theta}$	47.9 ± 4.2^{C}	52.9 ± 4.3^{d}

The number of pMV158 plasmid molecules per chromosome equivalent was determined by qPCR or by densitometric analysis of agarose gels. In this latter method, the plasmidic to chromosomal DNA ratios were compared with that obtained in the lactococcal host cells grown at pH 7.0 in the presence of Tc, to which the corresponding qPCR-determined pMV158 PCN value (red color number) was assigned. The copy number and growth rate values given in the table are expressed as mean ± standard deviation of three independent experiments (biological samples). All the values with the same superscript letter were not found to have statistically significant differences.





C

M17' pH	μ (h^{-1}) M17' -Tc (mean ± SD)	pMV158 copy number at G10 (mean ± SD)		pMV158 copy number at G20 (mean ± SD)	
		Gel quantification	qPCR quantification	Gel quantification	qPCR quantification
7.0	1.21 ± 0.17^{a}	7.5	7.5 ± 1.1^{A}	7.2 ± 0.9	7.8 ± 0.9^{A}
6.5	1.23 ± 0.15^{a}	6.2 ± 1.3	7.1 ± 0.7^{A}	7.0 ± 1.1	7.5 ± 0.5^{A}
6.0	1.13 ± 0.16^{a,b}	6.4 ± 1.1	7.1 ± 0.9^{A}	7.4 ± 0.7	7.2 ± 0.5^{A}
5.5	0.87 ± 0.10^{b,c}	6.1 ± 1.2	6.6 ± 0.9^{A}	6.9 ± 0.8	6.4 ± 0.5^{A}
5.0	0.66 ± 0.08^{c}	6.5 ± 1.3	6.8 ± 0.5^{A}	6.0 ± 2.0	6.4 ± 1.1^{A}
4.5	0.30 ± 0.05^{d}	3.6 ± 1.1	3.7 ± 0.5^{B}	4.5 ± 1.4	4.1 ± 0.2^{B}

FIGURE 5 | Segregational stability and PCN of pMV158 in *L. lactis* grown in M17′ adjusted to different pHs. **(A)** Analysis of the segregational stability of plasmid pMV158 in *L. lactis* MG1363. The stability of inheritance of pMV158 was analyzed after growing the lactoccocal cells for 10 and 20 generations in M17′ adjusted to different pH values in the absence of selective pressure. The vertical bars represent an estimation of the percentage of Tc resistant (Tc^{R}) cells after 10 (black bars) and 20 (gray bars) generations at the indicated pHs. Bar height represents the average value of three different analysis and error bars are standard deviations. **(B)** The gel shows the total DNA content of the *L. lactis* MG1363/pMV158 strain grown at the indicated pHs for 10 and 20 generations in the absence of Tc. Supercoiled monomeric (CCC) and open circle (OC) forms of pMV158 are indicated in the gel. Dividing lines indicate grouping of different parts of the same gel. This image is representative of at least three agarose gels used in the analysis and quantification of the DNA bands for PCN determination. **(C)** The table shows the pMV158 PCN in the lactococcal cells grown for 10 and 20 generations in Tc-free M17′ adjusted to different pHs. The pMV158 PCN per chromosome equivalent was determined by qPCR or by densitometric analysis of stained agarose gels. In this latter method, the plasmidic to chromosomal DNA ratios were compared with the one obtained in the lactococcal host cells grown at pH 7.0 for 10 generations, to which the corresponding qPCR-determined PCN value (red color number) was assigned. The table also includes the growth rates of *L. lactis*/pMV158 in Tc-free M17′ media adjusted to the different pH values. The copy number and growth rate values given in the table are expressed as mean ± SD of three independent experiments (biological samples). All the values with the same superscript letter were not found to have statistically significant differences. The statistical analysis of the PCN was only performed for the results obtained by qPCR, for which a lower standard deviation was found.

causing acid stress damage of the DNA (Jeong et al., 2008). This hypothesis is supported by the observations that: 1) a significant fraction of the plasmidic DNA in the lactococcal cells grown at pH 5.0 and especially at pH 4.5 in the presence of Tc, but not in the absence of the antibiotic, is found as open circular forms; and 2) a smear clearly appears in the gDNA extracted from the lactococcal cells grown at pH 4.5 in the presence, but not in the absence, of Tc (not shown).

The pMV158 PCN per chromosome equivalent in *L. lactis* MG1363/pMV158 cells grown at different pHs in the presence of Tc was inversely correlated with the bacterial growth rate. In fact, the PCN in cells of the lactococcal host grown in Tc-containing M17′ at pH 4.5 was ~7.5-fold higher than that in the host cells grown at pH 7.0 (**Table 1**). Increase in the PCN as the bacterial growth slows down has been reported for theta-replicating plasmid R1, which also encodes a Rep initiator protein (Engberg and Nordström, 1975). Also, a 3.5-fold increase in the PCN was observed for RCR lactococcal plasmid pLd3 at near-zero growth rates of the host cells (van Mastrigt et al., 2018).

Plasmid pMV158 is Stably Inherited by the Lactococcal Host Cells Irrespective of the pH of the Culture Medium.

A first evaluation of the capacity of pMV158 to replicate in the lactococcal host cells growing in media with different pHs was carried out by analyzing its segregational stability over a period of 10 and 20 generations in the absence of selection for the plasmid-encoded antibiotic resistance marker (Tc^R). The plasmid was maintained stably, irrespective of the pH of the culture medium, indicating that regulated replication was feasible in all cases (**Figures 5A,B**).

The Copy Number of pMV158 Decreases in Lactococcal Cells Grown at the Lowest pH.

A second approach to evaluate the capacity of pMV158 to replicate in lactococcal host cells that grow at different pHs consisted in determining the copy number of the plasmid after 10 and 20 generations in the absence of Tc. The growth rate of *L. lactis* MG1363/pMV158 in the different media was also determined to find out whether the foreseeable slowdown at the lowest pHs resulted in an increase of the PCN, as was the case with the host cells that grew in the presence of Tc (see **The PCN is Inversely Related to the Growth Rate of *L. lactis* MG1363 /pMV158 Cells That Grow in Tc-Containing Media Adjusted to Different pHs**).

Compared with the medium containing Tc, the growth of *L. lactis* MG1363/pMV158 in antibiotic-free medium showed a weaker dependence on the pH, with growth rates ranging from ~1.1-1.2 (at pHs 7.0, 6.5, and 6.0) to 0.3 (at pH 4.5) (**Figure 5C**).

The pMV158 copy number per chromosome equivalent showed only a rather small but statistically significant decrease (by ~50%) at the lowest pH (**Figure 5C**). The PCN reduction in the lactococcal cells grown in acidic media was observed irrespective of whether the gDNA was analyzed by qPCR or by densitometric analysis of plasmidic and chromosomal DNA bands on stained agarose gels (**Figure 5B**).

We have observed that the overall yield of gDNA from cells grown at different pHs to the same OD is quite similar. Moreover, since the cell density is kept constant, the mass and volume of the bacterial cells have the same dependence on the OD of the culture (Klumpp, 2011). Hence, it can be inferred that lactococcal cells grown at different pH conditions have a similar intracellular concentration of the chromosome, which is used as a reference for the determination of the PCN. Taking this into account, we can conclude that the estimated PCN correlates directly with the concentration of plasmid molecules within the lactococcal host cells.

The Molar Ratio of RepB to Plasmid DNA is Significantly Increased at a pH Value of 4.5.

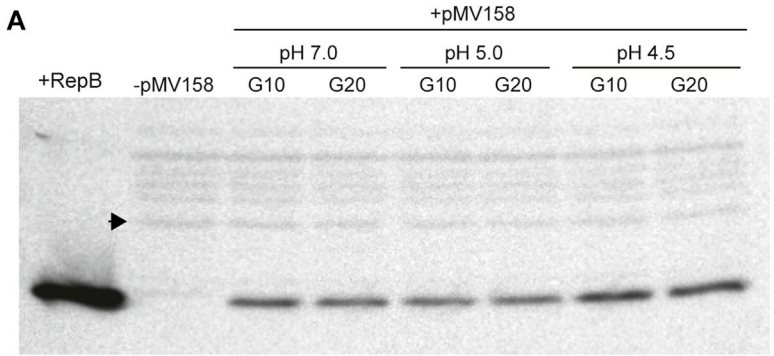
The intracellular concentration of the initiator RepB protein in the lactococcal host grown in media adjusted to pH 5.0 and 4.5 relative to that at pH 7.0, which was taken as the standard condition, was analyzed by western blot of protein extracts obtained from bacterial cultures at each condition (**Figure 6A**). The intracellular level of RepB (normalized to the internal control) was virtually the same in cells grown for either 10 or 20 generations at pH 7.0, and that estimated for 10 generations was arbitrarily taken as reference and assigned the value 1 (**Figure 6B**). The RepB level in the host cells grown at pH 5.0 also did not differ significantly from that observed in the cells that grew at pH 7.0, whereas the relative protein concentration within the cells grown for 10 or 20 generations at pH 4.5 was ~20% higher than the reference value and these differences were found to be statistically significant (**Figure 6B**). The RepB to pMV158 molar ratio relative to that within the host cells grown for 10 generations at pH 7.0 was next estimated for the lactococcal cells that had been grown for up to 20 generations at pH values of 7.0, 5.0, and 4.5. After 20 generations of growth at pH 7.0, the host cells presented the same RepB to plasmid molar ratio as the cells grown for 10 generations at neutral pH (reference value). Also, the protein to plasmid molar ratio in cells grown for 10 or 20 generations at pH 5.0 was close to that in cells that grew at pH 7.0. Remarkably however, the decrease of the pMV158 copy number in cells that grew for 10 or 20 generations at pH 4.5 occurred despite the increase in RepB concentration, so that the protein to plasmid molar ratio became ~2.4-fold higher than the ratio estimated for the cells grown at the reference pH value (**Figure 6B**).

DISCUSSION

So far, there have been only a very few reports about the effect of the pH on the plasmid RCR, a process that is initiated and terminated by Rep-mediated nucleophilic reactions. In fact, there is a lack of information on the efficiency of the nucleophilic DNA cleavage catalyzed by Rep proteins of RCR plasmids, bacteriophages and eukaryotic viruses under acidic conditions. The susceptibility to low pH of the strand-transfer reaction, as well as the stability of the Rep-DNA adduct against hydrolysis, have not yet been tested either.

In the present work, we have undertaken a global analysis to determine whether the acidity affected both the different *in vitro* reactions catalyzed by the pMV158-encoded RepB initiator protein and the parameters (PCN and stability) that prove the *in vivo* replication ability of the plasmid. In this sense, the pMV158 system provides the advantage of its singular promiscuity, which enables the plasmid to replicate in a wide variety of hosts, many of them able to grow at quite low pH conditions (Schillinger et al., 2006; Ruiz-Masó et al., 2017). Importantly, these studies may help to optimize the use of pMV158-based vectors in LAB, which constitute a large and heterogeneous group of bacteria widely employed, among other things, for the manufacturing of fermented and/or functional foods, for the production of probiotics and for the design and development of mucosal vaccines (Juarez del Valle et al., 2014; Mohedano et al., 2014; Russo et al., 2014).

The endonuclease activity of RepB on the nick sequence, which takes place at both the replication initiation and



B

	pH 7.0 G10	pH 7.0 G20	pH 5.0 G10	pH 5.0 G20	pH 4.5 G10	pH 4.5 G20
Relative normalized level of RepB	1.00 ± 0.08	1.02 ± 0.14	0.93 ± 0.10	0.96 ± 0.03	1.19 ± 0.04*	1.20 ± 0.06*
Relative RepB/PCN ratio	1.00 ± 0.14	1.00 ± 0.15	1.08 ± 0.13	1.15 ± 0.20	2.46 ± 0.34*	2.23 ± 0.15*

FIGURE 6 | Estimation of the relative RepB concentration and RepB to pMV158 molar ratio in lactococcal cells grown at different pHs. **(A)** Western blot analysis with RepB antiserum of protein extracts from *L. lactis*/pMV158 cells grown for 10 and 20 generations in Tc-free M17′ adjusted to pH 7.0, 5.0, and 4.5. Pure RepB protein and a protein extract from a pMV158-free *L. lactis* culture (–) were used as controls. The sample volume of the protein extracts was adjusted in order to keep constant the amount of total protein loaded in the gel. A black arrowhead indicates the band used as an internal loading control. **(B)** The normalized RepB level (9.91 ± 0.77) and RepB/PCN ratio (1.32 ± 0.18) were taken as reference and assigned the value 1 (in red color) for comparison with the different conditions. Data are presented as mean ± SD of the analysis of four different western blot assays. Values with an asterisk were found to have statistically significant differences with the corresponding reference value.

termination steps, is severely impaired at low pHs (**Figure 2**). Conversely, the strand-transfer activity, which is essential for the generation of the circular ssDNA intermediate that is used as the template for lagging-strand synthesis, does not seem to be affected by the pH (**Figure 3**). Finally, hydrolysis of the RepB-DNA adduct, which is shown here to require preservation of the native structure of the protein and hence may be mediated by the catalytic center of RepB, decreases with decreasing pH (**Figure 4**). The lability of the covalent linkage between RepB and the DNA has been hypothesized to play a critical role in releasing the inactive hexameric protein at the termination step, thus preventing recycling of RepB and continuous rounds of replication of the plasmid leading strand (Ruiz-Masó et al., 2015; Bordanaba-Ruiseco, 2017). According to this model, stabilization of the RepB-DNA adduct could lead to recycling of the protein and subsequent uncontrolled synthesis of the leading strand, a process that might counteract partially the impairment of the endonuclease activity at low pH. Based on the results of PCN in the host cells grown at the lowest pHs (**Figures 5B,C**), it does not seem however that RepB recycling, if occurs, has significant outcomes for the plasmid replication. Persistence of the covalent adduct might be a required but not sufficient condition for continuous replication using the same plasmid DNA template molecule.

Impaired RepB-mediated cleavage could result in a decreased frequency of the *in vivo* replicative events per RepB molecule, which would in turn entail a reduction in the intracellular concentration of pMV158 along with an increase in the protein to plasmid molar ratio. This is, in fact, what is observed in the lactococcal host cells grown in the absence of Tc at the lowest pH. Namely, it can be seen that at pH 4.5: 1) the PCN per chromosome equivalent (and hence the plasmid concentration) decreases by ~50% as compared with that at pH 7.0 (**Figure 5C**), and 2) the ratio between the concentration of RepB and the PCN is ~2.4-fold higher than that at pH 7.0 (**Figure 6B**). The decrease in the pMV158 copy number at low pH occurs even though a significant host growth slowdown, such as that observed in the most acidic medium, usually entails an increase in plasmid concentration or PCN per chromosome equivalent (Engberg and Nordström, 1975; Klumpp, 2011; van Mastrigt et al., 2018). It thus seems that the diminished pMV158 copy number observed within the host cells grown in M17' medium adjusted to pH 4.5 is directly accounted for by the low intracellular pH. With regard to the concentration of Rep initiator molecules relative to the plasmid concentration, it has been previously reported for staphylococcal RCR plasmid pT181 (wt and various copy-number mutants) that there is approximately one RepC dimer per plasmid copy (Jin et al., 1997). We have also estimated before that in the streptococcal host (*S. pneumoniae* 708) one RepB hexamer exists per pMV158 molecule, indicating that each hexameric RepB drives one plasmid replication event (Ruiz-Masó et al., to be published

elsewhere). We assume that such an efficient utilization of RepB would also occur in the lactococcal host cells grown at pH 7.0, which was taken as the reference for estimating the RepB to plasmid molar ratio at pH values of 5.0 and 4.5 (**Figure 6B**). This ratio is only significantly increased at pH 4.5, suggesting a less efficient utilization of RepB, with each replication event potentially requiring 2 or 3 RepB molecules in the lactococcal host cells grown at the lowest pH. However, the plasmid segregational stability points to the occurrence of regulated replication even at the lowest pH (**Figure 5A**).

It should be taken into account that acid-base homeostasis implies a pH gradient between the cytoplasm and the culture medium; the more extreme the pH, the greater the gradient. The pH homeostasis of *L. lactis* MG1363 in acid stress conditions has been studied and shown to entail a pH gradient value ranging from 0.7 (at outer pH 6.6) to 1.2 (at outer pH 4.4), so that the internal pH of the lactococcal cells grown in M17' adjusted to pH 4.5 would be ~5.6, which is the minimal internal pH value that enables the bacterial growth (Mercade et al., 2000). Hence, the pH homeostasis most likely accounts for the rather moderate reduction in PCN and plasmid concentration observed in the lactococcal host cells grown in medium at pH 4.5, which does not result in plasmid loss for at least 20 generations (**Figure 5A**).

The present work constitutes the first *in vitro* and *in vivo* global analysis of the potential effects of the acidity on the replication of a RCR plasmid. Previously, the effect of the culture medium pH on the PCN of RCR lactococcal plasmid pLd3, which belongs to a plasmid replicon family different from pMV158, had been investigated. No statistically significant decrease of the pLd3 PCN was found in the lactococcal cells grown at pH 5.5 (the only acidic condition investigated) compared with those grown at pH 7.0 (van Mastrigt et al., 2018). It should be highlighted nonetheless that the reduction in the pMV158 PCN in the lactococcal host is only observed at external pH values lower than 5.5, to which a cytoplasmic pH of 6.2–6.3 would be expected (**Figure 5B**; van Mastrigt et al., 2018)**.**

This study is also pioneer in the analysis of the pH-dependence of the different reactions, all of them involving nucleophilic attacks, catalyzed by members of the HUH endonuclease superfamily (Chandler et al., 2013). The results obtained in the present work show that the RepB-mediated DNA cleavage and the adduct hydrolysis, but not the strand-transfer reaction, are pH-dependent and greatly impaired in acidic conditions (**Figures 2–4**). The lability of the protein-DNA covalent linkage seems to be anyhow a singularity of the pMV158-encoded RepB initiator protein, as it has not been reported for the covalent adducts generated upon DNA cleavage by other HUH endonucleases. The lability of the protein-DNA covalent linkage has also been reported for the filamentous phage gpII RCR initiator protein, which, like the Rep proteins encoded by the pT181-family plasmids, lacks the HUH motif involved in binding of the divalent metal cation at the active site and hence does not belongs to the HUH endonuclease superfamily (Ilyina and Koonin, 1992). In the case of the pMV158 replication system, the instability of the RepB-DNA adduct may have an important physiological role in the termination step by precluding recycling and allowing the release of the hexameric inactive form of the initiator protein. In this way, the lability of the RepB-DNA covalent linkage might make it feasible the effective control of the replication, and hence of the copy number, of pMV158 by the plasmid-encoded elements that regulate the synthesis of the initiator protein, namely transcriptional repressor CopG and antisense RNAII (del Solar et al., 1995; Ruiz-Masó et al., 2015).

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

RV performed the determination of the pMV158 PCN in *L. lactis*, analyzed its segregational stability, and contributed to the OBD endonuclease and strand-transfer activity assays. LB-R performed the identification of the covalent adduct and contributed to the adduct stability assays. CM-H contributed to the OBD endonuclease activity assays. JAR-M performed the determination of the intracellular concentration of RepB and contributed to the OBD strand transfer assays and to the adduct stability assays. RV, JAR-M, and GS prepared the figures. GS and JAR-M contributed to the conception and design of the study. GS wrote the manuscript. All authors contributed to the manuscript revision, read, and approved the submitted version.

FUNDING

The Spanish Ministry of Science, Innovation and Universities (Grant RTI2018-097114-B-100 to GS) supported this work.

ACKNOWLEDGMENTS

We thank Ivan Belinchón-Esteban for critical reading of the manuscript. We thank Fernando Corrales, head of the CNB Proteomics Facility (CSIC), and Sergio Ciordia for access to the Typhoon biomolecular imager. We acknowledge support of the publication fee by the CSIC Open Access Publication Support Initiative through its Unit of Information Resources for Research (URICI).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2021.634461/full#supplementary-material.

REFERENCES

Baas, P. D. (1985). DNA replication of single-stranded *Escherichia coli* DNA phages. *BBA - Gene Struct. Expr.* 825, 111–139. doi:10.1016/0167-4781(85)90096-X

Backert, S., Meißner, K., and Börner, T. (1997). Unique features of the mitochondrial rolling circle-plasmid mp1 from the higher plant Chenopodium album (L.). *Nucleic Acids Res.* 25, 582–589. doi:10.1093/nar/25.3.582

Boer, D. R., Ruiz-Masó, J. A., López-Blanco, J. R., Blanco, A. G., Vives-Llàcer, M., Chacón, P., et al. (2009). Plasmid replication initiator RepB forms a hexamer reminiscent of ring helicases and has mobile nuclease domains. *EMBO J.* 28, 1666–1678. doi:10.1038/emboj.2009.125

Boer, D. R., Ruiz-Masó, J. A., Rueda, M., Petoukhov, M. V., Machón, C., Svergun, D. I., et al. (2016). Conformational plasticity of RepB, the replication initiator protein of promiscuous streptococcal plasmid pMV158. *Sci. Rep.* 6, 20915. doi:10.1038/srep20915

Bordanaba-Ruiseco, L. (2017). Caracterización de las actividades de RepB y su implicación en el proceso replicativo del plásmido promiscuo pMV158. Doctoral Thesis. Complutense University of Madrid. Available at: https://www.educacion.gob.es/teseo/mostrarRef.do?ref=433542

Burrus, V. (2017). Mechanisms of stabilization of integrative and conjugative elements. *Curr. Opin. Microbiol.* 38, 44–50. doi:10.1016/j.mib.2017.03.014

Chandler, M., de la Cruz, F., Dyda, F., Hickman, A. B., Moncalian, G., and Ton-Hoang, B. (2013). Breaking and joining single-stranded DNA: the HUH endonuclease superfamily. *Nat. Rev. Microbiol.* 11, 525–538. doi:10.1038/nrmicro3067

del Solar, G., Acebo, P., and Espinosa, M. (1995). Replication control of plasmid pLS1: efficient regulation of plasmid copy number is exerted by the combined action of two plasmid components, CopG and RNA II. *Mol. Microbiol.* 18, 913–924. doi:10.1111/j.1365-2958.1995.18050913.x

del Solar, G. H., Puyet, A., and Espinosa, M. (1987). Initiation signals for the conversion of single stranded to double stranded DNA forms in the streptococcal plasmid pLS1. *Nucleic Acids Res.* 15, 5561–5580. doi:10.1093/nar/15.14.5561

Dolja, V. V., and Koonin, E. V. (2011). Common origins and host-dependent diversity of plant and animal viromes. *Curr. Opin. Virol.* 1, 322–331. doi:10.1016/j.coviro.2011.09.007

Dressler, D. (1970). The rolling circle for X DNA replication, II. Synthesis of single-stranded circles. *Proc. Natl. Acad. Sci.* 67, 1934–1942. doi:10.1073/pnas.67.4.1934

Enemark, E. J., and Joshua-Tor, L. (2006). Mechanism of DNA translocation in a replicative hexameric helicase. *Nature* 442, 270–275. doi:10.1038/nature04943

Engberg, B., and Nordström, K. (1975). Replication of R-factor R1 in *Scherichia coli* K-12 at different growth rates. *J. Bacteriol.* 123, 179–186. doi:10.1128/JB.123.1.179-186.1975

Garay-Novillo, J. N., García-Morena, D., Ruiz-Masó, J. A., Barra, J. L., and del Solar, G. (2019). Combining modules for versatile and optimal labeling of lactic acid bacteria: two pMV158-family promiscuous replicons, a pneumococcal system for constitutive or inducible gene expression, and two fluorescent proteins. *Front. Microbiol.* 10, 1431. doi:10.3389/fmicb.2019.01431

Gasson, M. J. (1983). Plasmid complements of *Streptococcus lactis* NCDO 712 and other lactic streptococci after protoplast-induced curing. *J. Bacteriol.* 154, 1–9. doi:10.1128/JB.154.1.1-9.1983

Gilbert, W., and Dressler, D. (1968). DNA replication: the rolling circle model. *Cold Spring Harb. Symp. Quant. Biol.* 33, 473–484. doi:10.1101/SQB.1968.033.01.055

Guédon, G., Libante, V., Coluzzi, C., Payot, S., and Leblond-Bourget, N. (2017). The obscure world of integrative and mobilizable elements, highly widespread elements that pirate bacterial conjugative systems. *Genes* 8, 337. doi:10.3390/genes8110337

Ilyina, T. V., and Koonin, E. V. (1992). Conserved sequence motifs in the initiator proteins for rolling circle DNA replication encoded by diverse replicons from eubacteria, eucaryotes and archaebacteria. *Nucleic Acids Res.* 20, 3279–3285. doi:10.1093/nar/20.13.3279

Jeong, K. C., Hung, K. F., Baumler, D. J., Byrd, J. J., and Kaspar, C. W. (2008). Acid stress damage of DNA is prevented by Dps binding in *Escherichia coli* O157:H7. *BMC Microbiol.* 8, 181. doi:10.1186/1471-2180-8-181

Jin, R., Rasooly, A., and Novick, R. P. (1997). *In vitro* inhibitory activity of RepC/C*, the inactivated form of the pT181 plasmid initiation protein, RepC. *J. Bacteriol.* 179, 141–147. doi:10.1128/JB.179.1.141-147.1997

Juarez del Valle, M., Laiño, J. E., Savoy de Giori, G., and LeBlanc, J. G. (2014). Riboflavin producing lactic acid bacteria as a biotechnological strategy to obtain bio-enriched soymilk. *Food Res. Int.* 62, 1015–1019. doi:10.1016/j.foodres.2014.05.029

Khan, S. A. (2005). Plasmid rolling-circle replication: highlights of two decades of research. *Plasmid* 53, 126–136. doi:10.1016/j.plasmid.2004.12.008

Klumpp, S. (2011). Growth-rate dependence reveals design principles of plasmid copy number control. *PLoS One* 6, e20403. doi:10.1371/journal.pone.0020403

Koepsel, R. R., Murray, R. W., Rosenblum, W. D., and Khan, S. A. (1985). The replication initiator protein of plasmid pT181 has sequence-specific endonuclease and topoisomerase-like activities. *Proc. Natl. Acad. Sci. U.S.A.* 82, 6845–6849. doi:10.1073/pnas.82.20.6845

Koonin, E. V., and Ilyina, T. V. (1993). Computer-assisted dissection of rolling circle DNA replication. *Biosystems* 30, 241–268. doi:10.1016/0303-2647(93)90074-M

Krulwich, T. A., Jin, J., Guffanti, A. A., and Bechhofer, H. (2001). Functions of tetracycline efflux proteins that do not involve tetracycline. *J. Mol. Microbiol. Biotechnol.* 3, 237–246.

Lacks, S. A., Lopez, P., Greenberg, B., and Espinosa, M. (1986). Identification and analysis of genes for tetracycline resistance and replication functions in the broad-host-range plasmid pLS1. *J. Mol. Biol.* 192, 753–765. doi:10.1016/0022-2836(86)90026-4

Lovendahl, K. N., Hayward, A. N., and Gordon, W. R. (2017). Sequence-directed covalent protein–DNA linkages in a single step using HUH-tags. *J. Am. Chem. Soc.* 139, 7030–7035. doi:10.1021/jacs.7b02572

Marsin, S., and Forterre, P. (1999). The active site of the rolling circle replication protein Rep75 is involved in site-specific nuclease, ligase and nucleotidyl transferase activities. *Mol. Microbiol.* 33, 537–545. doi:10.1046/j.1365-2958.1999.01498.x

Martensen, T. M. (1982). Phosphotyrosine in proteins. Stability and quantification. *J. Biol. Chem.* 257, 9648–9652. doi:10.1016/S0021-9258(18)34121-8

Mercade, M., Lindley., N., and Loubière, P. (2000). Metabolism of *Lactococcus lactis* subsp. *cremoris* MG 1363 in acid stress conditions. *Int. J. Food Microbiol.* 55, 161–165. doi:10.1016/S0168-1605(00)00190-2

Mohedano, M. L., García-Cayuela, T., Pérez-Ramos, A., Gaiser, R. A., Requena, T., and López, P. (2014). Construction and validation of a mCherry protein vector for promoter analysis in *Lactobacillus acidophilus*. *J. Ind. Microbiol. Biotechnol.* 42, 247–253. doi:10.1007/s10295-014-1567-4

Moscoso, M., Eritja, R., and Espinosa, M. (1997). Initiation of replication of plasmid pMV158: mechanisms of DNA strand-transfer reactions mediated by the initiator RepB protein. *J. Mol. Biol.* 268, 840–856. doi:10.1006/jmbi.1997.1012

Moscoso, M., Solar, G., deland Espinosa, M. (1995). Specific nicking-closing activity of the initiator of replication protein RepB of plasmid pMV158 on supercoiled or single-stranded DNA. *J. Biol. Chem.* 270, 3772–3779. doi:10.1074/jbc.270.8.3772

Noirot-Gros, M. F., Bidnenko, V., and Ehrlich, S. D. (1994). Active site of the replication protein of the rolling circle plasmid pC194. *EMBO J.* 13, 4412–4420. doi:10.1002/j.1460-2075.1994.tb06761.x

Novick, R. P. (1998). Contrasting lifestyles of rolling-circle phages and plasmids. *Trends Biochem. Sci.* 23, 434–438. doi:10.1016/S0968-0004(98)01302-4

Projan, S. J., Carleton, S., and Novick, R. P. (1983). Determination of plasmid copy number by fluorescence densitometry. *Plasmid* 9, 182–190. doi:10.1016/0147-619X(83)90019-7

Providenti, M. A., O'Brien, J. M., Ewing, R. J., Paterson, E. S., and Smith, M. L. (2006). The copy-number of plasmids and other genetic elements can be determined by SYBR-Green-based quantitative real-time PCR. *J. Microbiol. Methods* 65, 476–487. doi:10.1016/j.mimet.2005.09.007

Ruiz-Masó, J. A., Bordanaba-Ruiseco, L., Sanz, M., Menéndez, M., and del Solar, G. (2016). Metal-induced stabilization and activation of plasmid replication initiator RepB. *Front. Mol. Biosci.* 3, 56. doi:10.3389/fmolb.2016.00056

Ruiz-Masó, J. A., López-Zumel, C., Menéndez, M., Espinosa, M., and del Solar, G. (2004). Structural features of the initiator of replication protein

RepB encoded by the promiscuous plasmid pMV158. *Biochim. Biophys. Acta – Proteins Proteomics* 1696, 113–119. doi:10.1016/j.bbapap.2003.09.010

Ruiz-Masó, J. A., Luengo, L. M., Moreno-Córdoba, I., Díaz-Orejas, R., and del Solar, G. (2017). Successful establishment of plasmids R1 and pMV158 in a new host requires the relief of the transcriptional repression of their essential *rep* genes. *Front. Microbiol.* 8, 2367. doi:10.3389/fmicb.2017.02367

Ruiz-Masó, J. A., Machón, C., Bordanaba-Ruiseco, L., Espinosa, M., Coll, M., and Del Solar, G. (2015). Plasmid rolling-circle replication. *Microbiol. Spectr.* 3, PLAS-0035-2014. doi:10.1128/microbiolspec.PLAS-0035-2014

Russo, P., Capozzi, V., Arena, M. P., Spadaccino, G., Dueñas, M. T., López, P., et al. (2014). Riboflavin-overproducing strains of *Lactobacillus fermentum* for riboflavin-enriched bread. *Appl. Microbiol. Biotechnol.* 98, 3691–3700. doi:10.1007/s00253-013-5484-7

Sánchez, C., Neves, A. R., Cavalheiro, J., dos Santos, M. M., García-Quintáns, N., López, P., et al. (2008). Contribution of citrate metabolism to the growth of *Lactococcus lactis* CRL264 at low pH. *Appl. Environ. Microbiol.* 74, 1136–1144. doi:10.1128/AEM.01061-07

Schillinger, U., Holzapfel, W. H., and Björkroth, K. J. (2006). "Lactic acid bacteria," in *Food spoilage microorganisms* (Elsevier), 541–578.

Siegumfeldt, H., Rechinger, K. B., and Jakobsen, M. (2000). Dynamic changes of intracellular pH in individual lactic acid bacterium cells in response to a rapid drop in extracellular pH. *Appl. Environ. Microbiol.* 66, 2330–2335. doi:10.1128/AEM.66.6.2330–2335.2000

Solar, G., Moscoso, M., and Espinosa, M. (1993). Rolling circle-replicating plasmids from Gram-positive and Gram-negative bacteria: a wall falls. *Mol. Microbiol.* 8, 789–796. doi:10.1111/j.1365-2958.1993.tb01625.x

Summers, D. K. (1996). *The biology of plasmids*. Oxford, UK: Blackwell Publishing Ltd.

te Riele, H., Michel, B., and Ehrlich, S. D. (1986). Are single-stranded circles intermediates in plasmid DNA replication?. *EMBO J.* 5, 631–637. doi:10.1002/j.1460-2075.1986.tb04257.x

Tse, Y. C., Kirkegaard, K., and Wang, J. C. (1980). Covalent bonds between protein and DNA. Formation of phosphotyrosine linkage between certain DNA topoisomerases and DNA. *J. Biol. Chem.* 255, 5560–5565. doi:10.1016/s0021-9258(19)70666-8

van Mastrigt, O., Lommers, M. M. A. N., de Vries, Y. C., Abee, T., and Smid, E. J. (2018). Dynamics in copy numbers of five plasmids of a dairy *Lactococcus lactis* strain under dairy-related conditions including near-zero growth rates. *Appl. Environ. Microbiol.* 84, e00314-18. doi:10.1128/AEM.00314-18

Wegmann, U., O'Connell-Motherway, M., Zomer, A., Buist, G., Shearman, C., Canchaya, C., et al. (2007). Complete genome sequence of the prototype lactic acid bacterium *Lactococcus lactis* subsp. *cremoris* MG1363. *J. Bacteriol.* 189, 3256–3270. doi:10.1128/JB.01768-06

Yasukawa, H., Hase, T., Sakai, A., and Masamune, Y. (1991). Rolling-circle replication of the plasmid pKYM isolated from a gram-negative bacterium. *Proc. Natl. Acad. Sci. U.S.A.* 88, 10282–10286. doi:10.1073/pnas.88.22.10282

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Valdelvira, Bordanaba-Ruiseco, Martín-Huestamendía, Ruiz-Masó and del Solar. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Molecular Biosciences

REVIEW
published: 18 March 2021
doi: 10.3389/fmolb.2021.648468

Multiple Layered Control of the Conjugation Process of the *Bacillus subtilis* Plasmid pLS20

Wilfried J. J. Meijer[1], D. Roeland Boer[2], Saúl Ares[3], Carlos Alfonso[4], Fernando Rojo[5], Juan R. Luque-Ortega[6] and Ling Juan Wu[7]**

[1]Laboratory 402, Centro de Biología Molecular "Severo Ochoa" (CSIC-UAM), Universidad Autónoma, Canto Blanco, Madrid, Spain, [2]ALBA Synchrotron Light Source, Barcelona, Spain, [3]Laboratory 35, C. Grupo Interdisciplinar de Sistemas Complejos and Departamento de Biología de Sistemas, Centro Nacional de Biotecnología, CSIC, Madrid, Spain, [4]Laboratory B08, Systems Biochemistry of Bacterial Division Lab, Centro de Investigaciones Biológicas Margarita Salas (CSIC), Madrid, Spain, [5]Laboratory 216, Department of Microbial Biotechnology, Centro Nacional de Biotecnología, CSIC, Madrid, Spain, [6]Laboratory S07, Molecular Interactions Facility, Centro de Investigaciones Biológicas Margarita Salas (CSIC), Madrid, Spain, [7]Centre for Bacterial Cell Biology, Biosciences Institute, Newcastle University, Newcastle Upon Tyne, United Kingdom

OPEN ACCESS

Edited by:
Chew Chieng Yeo, Sultan Zainal Abidin University, Malaysia

Reviewed by:
Günther Koraimann, University of Graz, Graz, Austria
Imrich Barak, Institute of Molecular Biology (SAS), Slovakia

***Correspondence:**
Wilfried J. J. Meijer
wmeijer@cbm.csic.es
Ling Juan Wu
l.j.wu@newcastle.ac.uk

Specialty section:
This article was submitted to Molecular Recognition, a section of the journal Frontiers in Molecular Biosciences

Received: *31 December 2020*
Accepted: *08 February 2021*
Published: *18 March 2021*

Citation:
Meijer WJ J, Boer DR, Ares S, Alfonso C, Rojo F, Luque-Ortega JR and Wu LJ (2021) Multiple Layered Control of the Conjugation Process of the Bacillus subtilis *Plasmid pLS20. Front. Mol. Biosci. 8:648468. doi: 10.3389/fmolb.2021.648468*

Bacterial conjugation is the main horizontal gene transfer route responsible for the spread of antibiotic resistance, virulence and toxin genes. During conjugation, DNA is transferred from a donor to a recipient cell via a sophisticated channel connecting the two cells. Conjugation not only affects many different aspects of the plasmid and the host, ranging from the properties of the membrane and the cell surface of the donor, to other developmental processes such as competence, it probably also poses a burden on the donor cell due to the expression of the large number of genes involved in the conjugation process. Therefore, expression of the conjugation genes must be strictly controlled. Over the past decade, the regulation of the conjugation genes present on the conjugative *Bacillus subtilis* plasmid pLS20 has been studied using a variety of methods including genetic, biochemical, biophysical and structural approaches. This review focuses on the interplay between Rco_{pLS20}, Rap_{pLS20} and Phr^*_{pLS20}, the proteins that control the activity of the main conjugation promoter P_c located upstream of the conjugation operon. Proper expression of the conjugation genes requires the following two fundamental elements. First, conjugation is repressed by default and an intercellular quorum-signaling system is used to sense conditions favorable for conjugation. Second, different layers of regulation act together to repress the P_c promoter in a strict manner but allowing rapid activation. During conjugation, ssDNA is exported from the cell by a membrane-embedded DNA translocation machine. Another membrane-embedded DNA translocation machine imports ssDNA in competent cells. Evidences are reviewed indicating that conjugation and competence are probably mutually exclusive processes. Some of the questions that remain unanswered are discussed.

Keywords: quorum sensing, *Bacillus subtilis*, Gram-positive bacteria, antibiotic resistance, horizontal gene transfer, DNA looping, transcriptional regulation, conjugation

INTRODUCTION

Horizontal Gene Transfer (HGT) refers to the process by which a DNA element/region is transferred from one cell to another. HGT occurs at large scale in bacteria, and a single HGT event can result in the acquisition of several to many genes by a cell and therefore plays a major role in bacterial evolution (Soucy et al., 2015). Unfortunately, HGT also contributes importantly to the dissemination of antibiotic resistance, virulence and toxin genes. The spread of these genes has now become a major global problem, with serious consequences on both human lives and economy (Partridge et al., 2018; CDC, 2019; Lerminiaux and Cameron, 2019). A better understanding of the HGT processes is needed to design drugs or strategies to impede the spread of antibiotic resistance and other pernicious genes. There are three main mechanisms responsible for HGT: transformation, transduction and conjugation (Ochman et al., 2000; Frost et al., 2005; Thomas and Nielsen, 2005; Novick et al., 2010). Of these, conjugation is the mechanism that is majorly responsible for the spread of antibiotic resistance, toxin and virulence genes (Mazel and Davies, 1999; Davies and Davies, 2010).

During conjugation, a DNA element is transferred from a donor cell to a recipient cell; hence, it requires direct contact between the two cells. Conjugative elements are often present on plasmids, but they can also be integrated in a bacterial chromosome. These latter elements are named integrative and conjugative elements (ICE). Conjugation occurs in both Gram-positive (G+) and Gram-negative (G−) bacteria, and although there are important differences (see below), the basics of the conjugation process are conserved in G+ and G− bacteria. Conjugation can be divided into four discernable stages, which have been described in detail previously (Chen et al., 2005; Arutyunov and Frost, 2013; Cabezon et al., 2014; Goessweiner-Mohr et al., 2014; Koraimann, 2018; Li et al., 2019; Waksman, 2019). Below we briefly summarize each stage, to illustrate the large number of proteins involved in conjugation, and the complexity of the process in which the many proteins act together in a temporally orchestrated manner. During stage 1) transfer competent donor cells select and attach to a suitable recipient cell (also known as mating pair formation). This process involves surface-located proteins, some of which are adhesion proteins that establish contact with the recipient cell. The adhesion proteins encoded by conjugative elements present in G− bacteria are generally located at the tip of a pilus structure. Conjugative elements of G+ bacteria do not form so-called sex-pili but probably encode adhesions to fulfill this function. While some proteins are involved in establishing contact, other surface-located proteins serve to inhibit conjugation between two donor cells, in a process named exclusion. In stage 2) the conjugation element is processed to generate the DNA that is transferred to the recipient cell, which in most cases is single-stranded DNA (ssDNA). The DNA processing reaction requires a relaxase, which forms a nucleoprotein complex called the relaxosome that introduces a site- and strand-specific nick within the origin of transfer (*oriT*). Host and/or plasmid-encoded proteins, named auxiliary proteins, often assist in the DNA processing reaction. During stage 3) a sophisticated membrane-embedded translocation machinery, also called transferosome, is synthesized through which the DNA is transferred. The intercellular transferosome, which is a form of type IV secretion system, is composed of at least eight but often more different proteins (Christie et al., 2014; Christie, 2016; Li et al., 2019; Waksman, 2019). Moreover, most of these proteins are present in multiple copies in the transferosome. The plasmid is recruited to the transferosome through interactions of an ATPase, named coupling protein (CP) and which is located at the cytoplasmic entry side of the transferosome, and the relaxosome proteins. Next, the relaxase pilots transfer of the ssDNA into the recipient cell in an energy consuming process that involves, besides the CP, one or two additional ATPases. Stage 4) is the establishment of the DNA once it has entered the recipient cell. Once transferred, the ssDNA has to be converted into double-stranded DNA (dsDNA). Many conjugative elements encode a primase that is involved in the conversion of ss to dsDNA. Finally, as most bacteria have evolved defense systems that inactivate incoming foreign DNA, including conjugative elements, conjugative elements encode proteins that transiently inactivate these defense mechanisms. So far, little is known about this step and the proteins involved.

Expression of the conjugation genes of most systems are strictly controlled because of at least two reasons. First, expression of the many proteins involved in the conjugation process imposes an energetic burden on the cell. Second, activation of the conjugation process has major consequences for both the plasmid and the host. Conjugative plasmids replicate by the theta type of replication mechanism during vegetative growth of the host. However, replication switches to the rolling-circle mode during conjugation in order to generate the ssDNA molecule that is transferred to the recipient cell (Smillie et al., 2010; Cabezon et al., 2014). Moreover, probably all conjugative plasmids possess a partitioning system to ensure that each daughter cell receives a copy of the plasmid upon cell division (Baxter and Funnell, 2014; Brooks and Hwang, 2017; Bouet and Funnell, 2019). During conjugation, however, the plasmid is recruited to the entry site of the transferosome. The conjugation process also has consequences for the host cell, as it alters characteristics of the membrane and the cell surface.

Studies on conjugative plasmids replicating in G- bacteria have shown that the conjugation genes are strictly controlled by a combination of plasmid and host-encoded factors. In particular, the mechanism controlling conjugation genes of the IncF incompatibility group has been studied in depth (Frost and Koraimann, 2010; Koraimann, 2018). Compared to G− bacteria, conjugation in G+ bacteria is less understood, and little is known about the regulation of conjugation genes (Singh and Meijer, 2014; Kohler et al., 2019). Exceptions are the two related enterococcal plasmids pCF10 and pAD1. Conjugation of these plasmids is induced by pheromones that are produced by plasmid-free recipient cells (for review see, Clewell, 2011; Dunny and Johnson, 2011; Chatterjee et al., 2013; Dunny and Berntsson, 2016).

The 65 kb plasmid pLS20 was originally identified in the *Bacillus subtilis natto* strain IFO3335 (Tanaka and Koshikawa,

1977) and is known to be conjugative in liquid as well as on solid media (Koehler and Thorne, 1987; Meijer et al., 1998; Itaya et al., 2006). Although very little was known about conjugation systems present on *B. subtilis* plasmids, as a host *B. subtilis* is one of the best studied G+ bacteria (Sonenshein et al., 1993; sonenshein et al., 2001). There are also other reasons why *B. subtilis* is an interesting host to study conjugation. *B. subtilis* is closely related to fastidious and pathogenic bacilli, and more distantly related to the Gram+ pathogen *Listeria*. *B. subtilis* is a soil-dwelling bacterium that is found all over the world; therefore pLS20 may be a representative of a group of globally distributed conjugative plasmids. Finally, *B. subtilis* strains are also gut commensals in animals and humans (Cutting, 2011), so it is possible that *B. subtilis* plasmids play a role in HGT in the gut.

Earlier studies on pLS20 have determined the replication region of the plasmid (Meijer et al., 1995), and shown that faithful segregation of the plasmid during cell division is ensured by the actin-like Alp7A protein (Derman et al., 2009). However, little was known about the conjugation process and how it was regulated. In this review we describe the advances that have been made in recent years in understanding how expression of the pLS20 conjugation genes is regulated, and identify the similarities and differences with two other conjugative systems of Gram+ origin. We also discuss future directions to unanswered questions. The most important gaps of knowledge concern structural insights in 1) how repressor molecules generate DNA looping, 2) how the antirepressor interacts with the repressor, and 3) if and how genes involved in different steps of the conjugation process–all located in a single large operon-are differentially expressed. More general reviews of regulation of conjugation in G+ and G– can be found elsewhere (Grohmann et al., 2003; Frost and Koraimann, 2010; Koraimann and Wagner, 2014; Singh and Meijer, 2014; Kohler et al., 2019).

EVIDENCE THAT pLS20 CONJUGATION IS STRICTLY CONTROLLED AND NOT ACTIVATED BY PHEROMONES

Determination of conjugation efficiencies as a function of growth revealed that maximum conjugation efficiencies, in the range of 1×10^{-3}, could be obtained only during a short window of time near the end of the exponential growth phase. Conjugation efficiency decreases sharply when donor cells enter the stationary phase, and eventually declines below the detection level ($<1 \times 10^{-8}$) (Singh et al., 2013). These results indicate that the conjugation process is strictly regulated. Conjugation efficiencies, which are also low at early time points, increase during the exponential growth phase. However, this increase in efficiency was not due to recipient-produced pheromones though as occurs for the enterococcal pheromone-responsive conjugative plasmids (Dunny, 2013), because similar results were obtained when the growth medium was replaced by fresh medium before mating (Singh et al., 2013). This demonstrated that regulation of pLS20 conjugation is fundamentally different from the enterococcal pheromone-responsive plasmids. The narrow time window of efficient conjugation does not depend on the growth phase of the recipient cells. Thus, pLS20 conjugation is strictly repressed except for a small window of time near the end of the exponential growth phase.

pLS20 CONTAINS A LARGE CONJUGATION OPERON THAT IS PRECEDED BY A DIVERGENTLY ORIENTED GENE, rco_{pLS20}, ENCODING THE MASTER REGULATOR OF CONJUGATION

Our annotation of the plasmid revealed a large putative operon spanning genes 28-74 in which some genes share similarity with conjugation genes (Singh et al., 2013; see **Figure 1A** for a schematic view of plasmid pLS20). This indicated that genes 28-74 constitute a large conjugation operon, which was later confirmed (see below). A single divergently oriented gene, 27c, is located immediately upstream of the conjugation operon. Its product, protein p27c, is a putative Xre-type transcriptional regulator predicted to contain a typical DNA binding Helix-Turn-Helix (HTH) domain near its

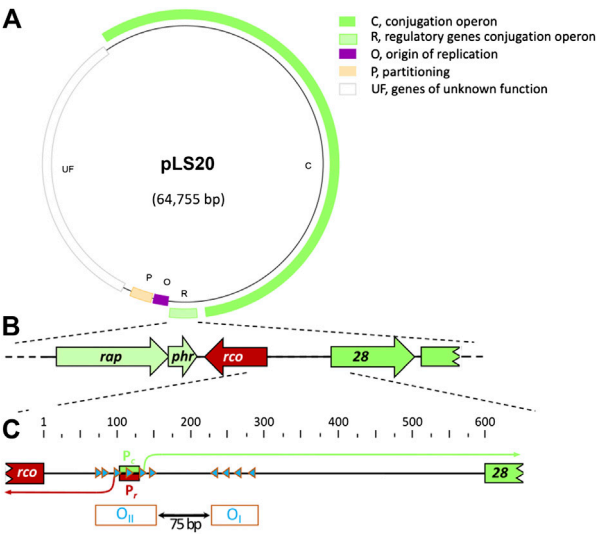


FIGURE 1 | General characteristics of pLS20 and genetic organization of the pLS20 region involved in regulation of the conjugation genes. **(A).** Map of pLS20 (accession number AB615352.1). Different modules are indicated with colors. Light green box, region containing the three genes regulating the activity of the main conjugation promoter P_c (indicated with "R"); dark green box, conjugation operon (indicated with "C"). Purple box, origin of replication region (indicated with "O"); orange box, partitioning genes (indicated with "P"); and white box, module with unknown genes (indicated with "UF"). **(B).** Blow-up of the region encoding the genes rap_{pLS20} (*rap*), phr_{pLS20} (*phr*) and rco_{pLS20} (*rco*) involved in regulating the expression of the conjugation genes, together with the 5' region of the conjugation operon. **(C).** Blow-up of the gene *28-*rco_{pLS20} intergenic region and its features involved in regulating expression of the conjugation genes. The position of the divergently oriented and overlapping promoters P_c and P_r are indicated with green and red boxes, respectively. Transcription initiation sites are indicated with bent arrows and the Rco_{pLS20} binding sites with blue triangles. The numbers at the top correspond to bp of the almost 600 bp intergenic region. The operators O_I and O_{II}, which are separated by a 75 bp spacer, are indicated at the bottom.

N-terminus. Protein p27c has been shown to be the master regulator of pLS20 conjugation and was named Rco_{pLS20} (regulator of conjugation of pLS20, see **Figure 1B**) (Singh et al., 2013). Ectopic overexpression of rco_{pLS20} (= gene *27c*) from a chromosomal location blocked conjugation, and conjugation levels were constitutively high in a derivative of pLS20 lacking the gene (pLS20rco). Plasmid pLS20rco was very unstable in wild type *B. subtilis* cells and its integrity required the ectopic expression of gene rco_{pLS20} in the host, suggesting that constitutive expression of the conjugation operon is generating high levels of stress to the host. Evidence that Rco_{pLS20} is a transcriptional regulator was obtained by RNAseq analyses (Singh et al., 2013). Thus, genes within the conjugation operon were expressed among the highest level of all plasmid-encoded genes when samples were taken from donor cell cultures at late exponential growth phase, but their levels were >16 fold lower for samples of cells in which rco_{pLS20} was ectopically overexpressed.

PROMOTERS DRIVING EXPRESSION OF rco_{pLS20} AND THE CONJUGATION OPERON

Several analyses, including RNAseq, deletion analyses and primer extension experiments revealed that the promoters driving expression of the divergently oriented genes rco_{pLS20} and gene *28*, the first gene in the conjugation operon, are located in the 595 bps intergenic region (Ramachandran et al., 2014; see also **Figure 1**). The promoter upstream of rco_{pLS20} is named P_r, and the one upstream of the conjugation operon P_c. The P_c promoter is located 462 bp upstream of gene *28* and, interestingly, overlaps with the divergently oriented P_r promoter. The strength of the two promoters has been determined using transcriptional *lacZ* fusions (Ramachandran et al., 2014). These analyses showed that the conjugation promoter P_c is a strong promoter, and promoter P_r, driving expression of rco_{pLS20}, is a very weak promoter whose activity was barely detectable. Surprisingly, analyses of additional strains containing, besides a transcriptional *lacZ* fusion, either pLS20 or a copy of rco_{pLS20} under the control of an inducible promoter, revealed that Rco_{pLS20} does not just repress the strong promoter P_c, but also regulates its own promoter, P_r. At low expression levels Rco_{pLS20} stimulates the activity of P_r but at high expression level it represses its own promoter (Ramachandran et al., 2014). It is important to mention that the maximum strength of the P_r promoter is about 50–75-fold lower than that of the P_c promoter (Singh et al., 2020). In summary, Rco_{pLS20} represses the strong promoter P_c and simultaneously regulates the activity of its own promoter P_r (Ramachandran et al., 2014).

REGULATION OF P_c AND P_r BY THE TETRAMERIC Rco_{pLS20} INVOLVES DNA LOOPING

The binding regions of Rco_{pLS20} were initially determined by analyzing the response of P_r and P_c promoters to ectopically expressed rco_{pLS20}, using strains containing *lacZ* fused to promoter-containing fragments of different lengths (Ramachandran et al., 2014). This approach revealed that proper regulation required the presence of two DNA regions. These two regions contain 10 octamer boxes whose sequences are (nearly) identical to CAGTGAAA (see **Figure 1C** for a schematic view). In the first region a cluster of six of these boxes, all in the same orientation, are located near, and overlapping with the P_r and P_c promoters. The second region contains four clustered boxes whose orientation is opposite to those present in the first region. The two clusters of boxes are separated by 75 bp. Binding of Rco_{pLS20} to these boxes was demonstrated by DNase I footprinting assays, and also by gel retardation experiments using fragments with/out mutation in the octamers (Ramachandran et al., 2014). The regions containing the four and six octamers are named operator O_I and O_{II}, respectively (see **Figure 1C**).

The presence of two convergently oriented Rco_{pLS20} operators raised the possibility that binding of Rco_{pLS20} to these operators could result in DNA looping. However, the spacer region between the two operators is only 75 bp. Due to intrinsic stiffness of DNA spacer regions in DNA-loops are generally larger than 90 bp. DNA loops containing smaller spacer lengths are only possible when the DNA region is intrinsically bent. Circular permutation assays showed that the region containing the Rco_{pLS20} operators is indeed intrinsically bent and that the center of the bent is located in approximately the middle of the 75 bp spacer region between operators O_I and O_{II} (Ramachandran et al., 2014). Several lines of evidence supported that binding of Rco_{pLS20} to operators O_I and O_{II} results in DNA looping. First, ultracentrifugation analysis showed that Rco_{pLS20} forms tetramers in solution (Ramachandran et al., 2014), which is in line with the observation that proteins generating DNA looping commonly form multimers in solution. Second, particularly for short spacer regions, DNA looping requires a certain phasing between the two operators so that operator-bound proteins are in the required configuration to generate the DNA loop. Enlarging the spacer region between the two operators with a half helical turn (i.e., 5 bp) disrupted Rco_{pLS20} mediated regulation of the P_c promoter *in vivo* (Ramachandran et al., 2014). Finally, Rco_{pLS20}-mediated DNA looping was further supported by gel retardation and ultracentrifugation experiments, which also provided important insights into the mode of DNA binding (Singh et al., 2020). DNA looping may be generated as a consequence of one of the two fundamentally different modes by which tetrameric Rco_{pLS20} units bind DNA, as schematically shown in **Figure 2A**. The difference between these two modes of DNA binding can be distinguished by gel retardation experiments using DNA fragments containing only one operator, as the second binding mode generates a maximum of two retarded species, while the first binding mode would allow the generation of more than two retarded species. Retardations using DNA fragments containing either operator O_I or O_{II} both generated a maximum of two retarded species (Ramachandran et al., 2014), indicating that DNA looping occurs through interaction between two Rco_{pLS20} tetramers each bound to an operator. This conclusion was supported by additional gel retardation experiments using two DNA fragments, each containing one operator but having different lengths and distinct flanking

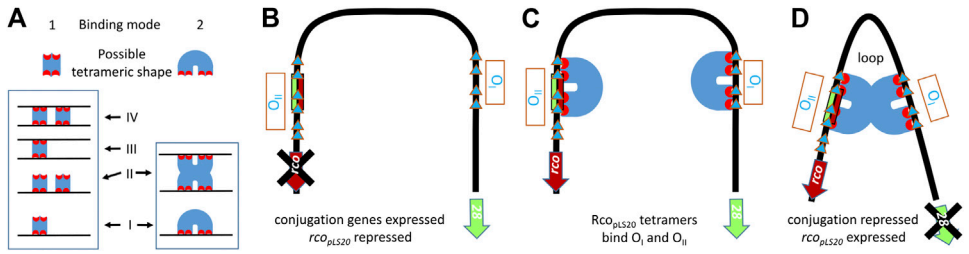


FIGURE 2 | Probable mode of DNA binding by Rco_{pLS20} to the intrinsically bent DNA region encompassing Rco_{pLS20} operators O_I and O_{II}. **(A).** Schematic representation of two different binding modes by which tetrameric DNA binding proteins can bind a DNA fragment containing one operator. Binding modes 1 and 2 can generate up to four and two different protein-bound DNA configurations, respectively. **(B–D)**. Schematic view of conjugation unrepressed state and repressed state involving DNA looping due to binding of Rco_{pLS20} to operators O_I and O_{II} according to binding mode 2 explained in panel **A**. **(B).** Unrepressed state. The conjugation operon starting with gene *28* is derepressed when the operators O_I and O_{II} are not bound by Rco_{pLS20} tetramers. **(C).** Rco_{pLS20} tetramers can bind to operators O_I and O_{II}, resulting in a local high concentration and probably a specific configuration stimulating interaction between the two tetramers. **(D).** Repressed state. Interaction between the two Rco_{pLS20} tetramers bound to operators O_I and O_{II} results in repression of the conjugation genes and simultaneous activation of rco_{pLS20}. Blue triangles: Rco_{pLS20} binding sites. Blue U shapes: Rco_{pLS20} tetramers.

sequences, allowing the fragments to be distinguished by Southern blotting. Gel retardation using a mixture of the small and large DNA fragment resulted in the appearance of an additional retarded species composed of a small and a large DNA fragment (Singh et al., 2020). Furthermore, SAXS analysis showed that Rco_{pLS20} forms complexes of increased size at concentrations above 200 μM, indicating that Rco_{pLS20} tetramers can interact with each other (Crespo et al., 2020), which would be expected if DNA binding occurs via the second mode. The combined evidences strongly indicate that efficient repression of the P_c promoter and simultaneous activation of the P_r promoter by Rco_{pLS20} is achieved through DNA looping as a consequence of interaction between the Rco_{pLS20} tetramers bound to operators O_I and O_{II}. We cannot fully excluded the possibility that, in addition to DNA looping, binding of Rco_{pLS20} also introduces other effects on the DNA configuration or orientation.

rco_{pLS20} is under the control of a very weak promoter (see above). Nonetheless, despite the low levels of Rco_{pLS20} present in the cell, the conjugation process is strictly controlled as the number of transconjugants is below the detection level under conjugation-unfavorable conditions. Probably, this is achieved through DNA looping because it causes a high local concentration of the transcriptional regulator at the right location, which has been shown to be able to increase its specificity and affinity, and concomitantly controls the stochasticity of cellular processes (Vilar and Saiz, 2005; Oehler and Muller-Hill, 2010). Thus, DNA looping allows strict regulation at low protein concentration, while remaining sensitive for accurate activation of conjugation when appropriate conditions –as detailed in the following section- occur. Examples of DNA looping playing a crucial role in transcriptional regulation have been described before. One of the best-studied examples is DNA looping exerted by the CI protein and which is crucial for the genetic switch between the lysogenic and lytic state of the *Escherichia coli* phage λ (for review see, Dodd et al., 2005; Oppenheim et al., 2005). There are clear differences between the DNA looping mechanisms of λ and pLS20. Whereas Rco_{pLS20} forms tetramers in solution, CI forms dimers. In addition, the region separating the operators is about 2.3 kb in phage λ, and 75 bp in pLS20.

Rco_{pLS20}-MEDIATED REPRESSION IS RELIEVED BY THE QUORUM SENSING-RESPONSIVE ANTI-REPRESSOR Rap_{pLS20} BELONGING TO THE RRNPP FAMILY OF PROTEINS

Expression of the conjugation genes requires suspension of Rco_{pLS20}-mediated repression of the conjugation promoter P_c. In case of pLS20, the repressor Rco_{pLS20} is inactivated by an antirepressor, Rap_{pLS20}. Annotation of pLS20cat revealed that a two-gene *rap-phr* cassette is located downstream of rco_{pLS20}, with direction of transcription convergent to that of rco_{pLS20} (see **Figure 1B**). Rap proteins belong to a large family of signal-peptide receptor proteins of the so-called RRNPP family, which is named after its prototypical members Rap, Rgg, NprR, PlcR and PrgX (for review see, Declerck et al., 2007; Rocha-Estrada et al., 2010; Dunny and Berntsson, 2016; Neiditch et al., 2017). RRNPP proteins, which are encoded by many bacteria belonging to the phylum of Firmicutes, are generally co-transcribed with a *phr* gene encoding the cognate signaling peptide. The Phr signaling peptides are synthesized as a pre-proprotein. After being secreted they are cleaved again to generate the mature peptide, which in most cases correspond to the C-terminal region of the pre-proprotein, and which can be imported into the cell by the oligopeptide permease system (Pottathil and Lazazzera, 2003; Monnet et al., 2016). The chromosome of *B. subtilis* contains eight *rap-phr* cassettes, and some *Bacillus* plasmids also contain such cassettes. Many of these *rap-phr* cassettes, including those on plasmids, function to interfere with developmental processes such as sporulation, competence and degradative enzyme production (Perego, 2001; Bongiorni et al., 2006; Smits et al., 2007; Parashar et al., 2013). Ectopically overexpressing the *rap-phr* cassette of pLS20 in *B. subtilis* did not seem to affect developmental processes of the host. However, it strongly

affected the conjugation process, resulting in 1) increased maximum conjugation levels, 2) overexpression of genes located in the conjugation operon, and 3) allowing conjugation to occur during a much broader time window. Moreover, conjugation levels were severely affected for a pLS20 derivative lacking *rap-phr* (Singh et al., 2013). Addition of a synthetic peptide corresponding to the five C-terminal residues of the unprocessed Phr peptide also inhibited conjugation, showing that the activity of Rap is regulated by the mature signaling peptide (Singh et al., 2013). Based on these results the genes were named rap_{pLS20} and phr_{pLS20} and the mature Phr peptide was referred to as Phr^{*}_{pLS20}.

MINIMAL *IN VIVO* REGULATORY CIRCUIT OF pLS20 CONJUGATION

The results outlined above suggest that Rco_{pLS20}, Rap_{pLS20} and Phr^{*}_{pLS20} are the only pLS20-encoded proteins involved in regulation of the conjugation genes. This was confirmed using an *in vivo* system in which the regulatory components were uncoupled from their native setting in a *B. subtilis* strain where the rco_{pLS20} and rap_{pLS20} genes are placed under different inducible promoters, and a *lacZ* reporter gene fused to the main conjugation promoter P_c. The main conjugation promoter P_c was active in the absence of both rco_{pLS20} and rap_{pLS20} expression. P_c was repressed when rco_{pLS20} alone was induced, strongly supporting the view that Rco_{pLS20} is directly responsible for repressing the P_c promoter, which is in line with *in vitro* results (see above). The P_c promoter became active when both rco_{pLS20} and rap_{pLS20} were induced, showing that Rap_{pLS20} alone is sufficient to relieve Rco_{pLS20}-mediated repression of the P_c promoter. Finally, when both rco_{pLS20} and rap_{pLS20} were induced in the presence of the mature peptide Phr^{*}_{pLS20} (added to the growth medium), the P_c promoter remained inactive, demonstrating that Phr^{*}_{pLS20} is required and sufficient to inactivate Rap_{pLS20} (Singh et al., 2020).

RELATIVE PROMOTER STRENGTHS

The interplay between Rco_{pLS20}, Rap_{pLS20} and Phr^{*}_{pLS20} requires an understanding of the relative strengths of their promoters, as well as the activity of promoter P_c, which these proteins regulate. The strengths of P_c, P_r and the promoter driving expression of rap_{pLS20} and phr_{pLS20}, named P_{rap}, were determined using strains containing *gfp* reporter cassettes in which either of these promoters was fused to the *gfp* gene (Singh et al., 2020). The results obtained with the *gfp*-fusion strains confirmed that P_c and P_r are a strong and a very weak promoter, respectively. P_{rap} was found to be a weak promoter, whose strength was similar to the P_r promoter in cells containing pLS20, and about 50 to 75-fold lower than that of P_c (see **Figure 3**).

P_{rap} drives the expression of rap_{pLS20} and phr_{pLS20}. However, the concentration of Phr_{pLS20} will drop once it is exported from the cell and it is plausible that proper signaling of Phr^{*}_{pLS20} requires higher expression levels than that of Rap_{pLS20}. Analysis of an additional strain containing *gfp* fused to the upstream region of phr_{pLS20} revealed that the expression of phr_{pLS20} indeed is controlled by a second promoter, named P_{phr}, whose activity is about two-fold higher than that of P_{rap}. Unlike the P_c and P_r promoters, the activities of P_{rap} and P_{phr} were not affected by the presence of pLS20 in the cell, indicating that Rco_{pLS20} does not regulate these promoters. The absence of Rco_{pLS20} binding motifs in the vicinity of the P_{rap} and P_{phr} promoters supports this conclusion (Singh et al., 2020). In summary, 1) the conjugation genes are under the control of the strong P_c promoter, 2) the rco_{pLS20} and rap_{pLS20} genes, driving expression of the proteins regulating the activity of the P_c promoter, are preceded by very weak promoters, 3) expression of *phr* is controlled by P_{rap} and additionally by the two-fold stronger promoter P_{phr}, presumably to accommodate for the dilution effect upon secretion, and 4) Rco_{pLS20} represses P_c and regulates the P_r promoter, but not promoters P_{rap} and P_{phr}.

PROMOTER P_c IS HOMOGENEOUSLY EXPRESSED

The multi-layered regulation of the P_c promoter in combination with the Rap/Phr-based quorum sensing mechanism might be expected to cause heterogeneous activation of the P_c promoter in a population of donor cells. This seemed not to be the case though, because flow cytometry analyses of samples of *B. subtilis* cells having a chromosomal cassette containing the P_c-*gfp* fusion, or those harboring in addition pLS20, both showed homogeneous activity of the P_c promoter (Singh et al., 2020). In this set up, the P_c promoter was present on the bacterial chromosome whereas proteins regulating its activity were encoded by the resident plasmid containing its own P_c promoter. A homogeneous pattern of P_c promoter activity was also found for cells containing a derivative of pLS20cat in which a copy of the *gfp* gene was placed on pLS20 behind gene *28*, the first gene of the conjugation operon. Together, these results strongly indicate that the multi-levels of regulation results in a sensitive genetic switch that activates the P_c promoter in a coordinated manner in most or all pLS20-containing cells (Singh et al., 2020). This does not automatically imply that activation of the P_c promoter guarantees successful plasmid transfer or even the initiation of the conjugation process in all cells. In fact, in most if not all conjugation systems studied, the full conjugation process appears to be activated only in a subpopulation of the donor cells (reviewed in, Koraimann and Wagner, 2014; Stingl and Koraimann, 2017). In the case of pLS20, checkpoints may be present downstream the P_c promoter. Moreover, successful transfer may not occur even when all conjugation genes are expressed due to, for example, unsuccessful mating pair formation or failure of the plasmid to establish itself in the new host. Moreover, it should be noted that the laboratory conditions under which conjugation is studied are very different from natural conditions in which environmental fluctuations at macro and microscale are likely to affect individual cells or subpopulations.

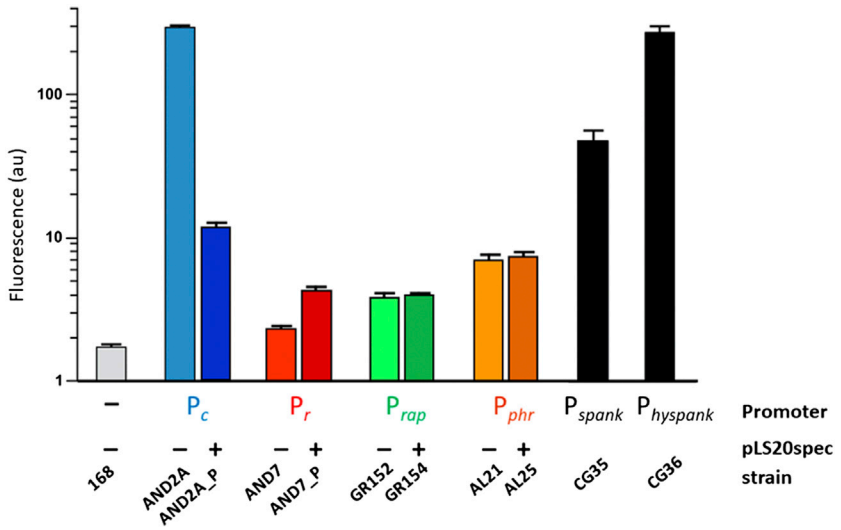


FIGURE 3 | Relative strength of promoters P_c, P_r, P_{rap} and P_{phr} determined by FACS analysis of strains containing transcriptional *gfp* fusions. Relative promoter strengths determined by cytometry using strains containing promoter P_c, P_r, P_{rap} or P_{phr} transcriptionally fused to *gfp*. A negative control strain (-) and two positive control strains that contain *gfp* fused to the IPTG-inducible promoter P_{spank} or $P_{hyspank}$ were included. Values correspond to late exponentially growing cultures (OD_{600} between 0.8-1). Light and dark colored bars reflect strains lacking or containing pLS20spec, respectively. The figure is adapted from (Singh et al., 2020).

Rap_{pLS20} ACTIVATES CONJUGATION BY DETACHING Rco_{pLS20} FROM ITS OPERATORS AND FORMING HETEROCOMPLEXES WITH Rco_{pLS20}

All members of the RRNPP family of proteins analyzed so far, are all-helical proteins composed of a large C-terminal domain (CTD) and a much smaller N-terminal domain (NTD). Whereas the large CTD interacts with its cognate signaling peptide resulting in conformational changes that affects the activity of the protein, the small NTD constitutes the effector domain that interacts with a target molecule (D'Andrea and Regan, 2003; Rocha-Estrada et al., 2010; Neiditch et al., 2017). Although the effector domains of all RRNPP proteins are composed of three α-helices, they exert their regulatory functions in three fundamentally different ways. In the first subset of RRNPP proteins these three α-helices have evolved into Helix-Turn-Helix (HTH) DNA binding motifs, which allow them to negatively regulate protein expression by binding to DNA. Examples of such RRNPP proteins are PlcR from *Bacillus thuringiensis* (Declerck et al., 2007) and PrgX from *Enterococcus faecalis* (Shi et al., 2005; Dunny and Berntsson, 2016). Particularly the PrgX protein is interesting in the context of this review because it regulates the conjugation genes present on the plasmid pCF10 (see below). In the second subset of RRNPP proteins, the effector domain has phosphatase activity. In *B. subtilis*, the developmental process of sporulation is induced by the phosphorylated form of the master regulator Spo0A, Spo0A~P. Spo0A phosphorylation is controlled by a phosphorelay consisting of several kinases and two phosphorelay proteins, Spo0F and Spo0B (Hoch, 1993; Grossman, 1995; Molle et al., 2003; Chastanet and Losick, 2011). Several Rap proteins, including RapA, RapB, RapE, RapH and RapJ, can interrupt the phosphorylation of Spo0A by dephosphorylating Spo0F~P. RapA was the first Rap member discovered to have this activity and was named Rap, which stands for Response regulator Aspartate Phosphatase (Perego, 2001). Finally, in the third subset of RRNPP proteins, the effector domain binds to its cognate effector protein, which impedes the function of its cognate effector protein (in)directly by modulating the expression of the differentiation pathway involved. Examples of these are RapC, RapF, RapG, RapH and RapK (note that not all Rap proteins have phosphatase activity). It is also worth mentioning that the effector domain of some RRNPP proteins, for instance RapH and NprR, exhibit both activities.

Based on the features of the known RRNPP proteins, one possibility of how Rap_{pLS20} might activate the pLS20 conjugation genes is by competing with Rco_{pLS20} for binding to DNA near the P_c promoter. However, both gel retardation and analytical ultracentrifugation (AUC) analyses showed that Rap_{pLS20} has no DNA binding activity, and the NTD of Rap_{pLS20} does not adopt a HTH configuration (Crespo et al., 2020; Singh et al., 2020). Instead, Rap_{pLS20} seems to exert its antirepressor activity by binding directly to Rco_{pLS20}, as bacterial two-hybrid analysis showed that Rap_{pLS20} and Rco_{pLS20} interact *in vivo* (Singh et al., 2020). *In vitro* evidence that Rap_{pLS20} interacts directly with Rco_{pLS20} has also been obtained by SAXS and analytical size exclusion chromatography (Crespo et al., 2020). Furthermore, AUC analyses provided conclusive evidence that Rap_{pLS20} forms predominantly dimers in solution. Sedimentation velocity experiments of mixtures of Rap_{pLS20} and Rco_{pLS20} resulted in the appearance of a novel species whose deduced molecular weight was compatible with a complex of one Rap_{pLS20} dimer and one Rco_{pLS20} tetramer (Singh et al., 2020). This conclusion

was reinforced by multi-signal sedimentation velocity (MSSV) analysis from which it was determined that the heterocomplex was composed of a stoichiometry of 2.1 mol of Rco_{pLS20} per mol of Rap_{pLS20}. More importantly, besides demonstrating that Rap_{pLS20} interacts with free Rco_{pLS20}, AUC analyses also showed that Rap_{pLS20} is able to interact with DNA-bound Rco_{pLS20}, and that this results in the release of Rco_{pLS20} from the DNA and the concomitant formation of Rap_{pLS20}/Rco_{pLS20} complexes. Complementary gel retardation experiments provided evidence that Rap_{pLS20} preferentially interacts with Rco_{pLS20} tetramers involved in DNA looping (Singh et al., 2020). In summary, Rap_{pLS20} appears to activate the conjugation genes by preferentially interacting with Rco_{pLS20} involved in DNA-looping, resulting in the detachment of Rco_{pLS20} from its operators, and forming a heterocomplex in which one Rap_{pLS20} dimer binds one Rco_{pLS20} tetramer. A schematic view of how Rap_{pLS20} activates the conjugation genes is shown in **Figure 4**.

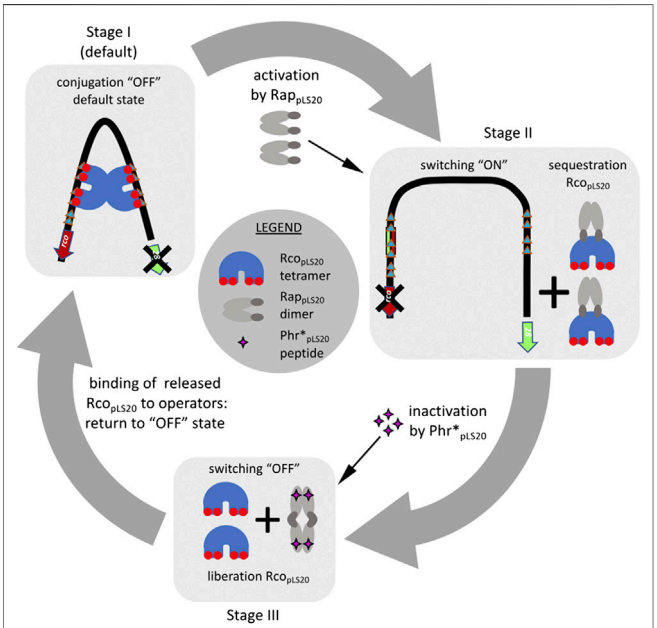


FIGURE 4 | Overview of the current understanding of how the pLS20 conjugation genes are regulated by Rco_{pLS20}, Rap_{pLS20} and Phr^*_{pLS20}. Stage I, as default, the two Rco_{pLS20} tetramers bound to operators O_I and O_{II} interact with each other resulting in DNA-looping. This configuration causes a steady state in which the strong conjugation promoter P_c is repressed and the weak P_r promoter driving expression of *rco*$_{pLS20}$ is simultaneously stimulated. Stage II, the conjugation genes are activated when two apo-form of Rap_{pLS20} dimers detach two Rco_{pLS20} tetramers from the DNA through direct interaction and generating Rap_{pLS20}:Rco_{pLS20} heterocomplexes in a stoichiometry of 1:2. Stage III, binding of the Phr^*_{pLS20} signaling peptide provokes a conformational change of Rap_{pLS20} that stimulates tetramerization of Rap_{pLS20} via the so-called foot-2-foot configuration in which the N-terminal effector domains interact with each other and are, presumably, no longer available to interact with Rco_{pLS20}. The liberated Rco_{pLS20} tetramers are then free to bind its operators thereby returning the system to the default conjugation repressed "OFF" state (stage I).

THE ACTIVATION STATE OF THE CONJUGATION GENES IS ULTIMATELY DETERMINED BY THE SIGNALING PEPTIDE Phr^*_{pLS20}

The presence of the mature signaling peptide Phr^*_{pLS20} suppresses the Rap_{pLS20}–mediated activation of the conjugation genes (see above). For other RRNPP members it has been shown that signaling peptides interact directly with the CTDs of their cognate RRNPP proteins, causing inactivation through peptide-induced conformational changes of the RRNPP protein. As outlined below, this is also the case for Rap_{pLS20} and Phr^*_{pLS20}. Phr^*_{pLS20} binds directly to Rap_{pLS20} (see below) and the dissociation constant, K_d, for Phr^*_{pLS20} was determined to be 2.1 and 7.4 μM by ultracentrifugation and fluorescence anisotropy approaches, respectively (Crespo et al., 2020; Singh et al., 2020). Although AUC analysis showed that Rap_{pLS20} forms predominantly dimers in solution (Crespo et al., 2020, see also above), it seems to have an intrinsic ability to form tetramers at higher concentrations, as 1.1% of Rap_{pLS20} was detected as tetramers in solution at 4.5 μM, and this percentage increased to 3.6% at 25 μM. Interestingly, AUC analysis showed that in the presence of Phr^*_{pLS20}, Rap_{pLS20} forms almost exclusively tetramers in solution (Crespo et al., 2020). The structures of the apo and peptide-bound forms of Rap_{pLS20} have been determined recently, which show that Rap_{pLS20} forms the typical structure of RRNPP family proteins (Crespo et al., 2020). Rap_{pLS20} is an all α-helical protein (17 antiparallel helices) that is composed of a large 14-helix C-terminal domain (CTD) that can bind the signal peptide, and a much smaller 3-helix N-terminal domain (NTD). The three helices of the NTD do not form a HTH topology, but have an approximate parallel configuration. The N- and CTDs are separated by a 13-residue long flexible loop that connects α-helices three and four. Rap_{pLS20} dimerizes through interactions between the CTDs, and these interactions involve helices H5-H7, H16-H17 and the C-terminal residues.

The structure of the peptide-bound form of Rap_{pLS20} showed that Phr^*_{pLS20} binds in a cavity of the CTD. Binding of the Phr^*_{pLS20} peptide causes the N-terminal domains to swing outwards thereby favoring Rap_{pLS20} to form tetramers (Crespo et al., 2020). Interestingly, the dimer-dimer interface is principally formed between the NTDs of each dimer, resulting in a so-called foot-2-foot interaction (see stage III in **Figure 4** for a schematic view). The small NTDs of RRNPP proteins form the effector domain and this is most likely also the case for Rap_{pLS20}, implying that the NTD of Rap_{pLS20} interacts with Rco_{pLS20}. Due to the particular foot-2-foot configuration induced upon peptide binding, the N-terminal effector domain is no longer capable of interacting with Rco_{pLS20} to relieve repression, providing the molecular mechanism by which Phr^*_{pLS20} inactivates the antirepressor activity of Rap_{pLS20}. Phr^*_{pLS20} can also bind to Rap_{pLS20} complexed with Rco_{pLS20}. This conclusion is based on the results of size exclusion chromatography and SAXS experiments, which showed that the addition of Phr^*_{pLS20} peptide to preformed Rap_{pLS20}/Rco_{pLS20} complexes resulted in

disruption of these complexes and the appearance of elution and SAXS patterns that were similar to those of Rco_{pLS20} alone and Rap_{pLS20}-$Phr^{\star}_{pLS20}$ complexes (i.e., Rap_{pLS20} tetramers) (Crespo et al., 2020).

RECIPROCAL INHIBITION OF CONJUGATION AND COMPETENCE PATHWAYS

Under certain growth conditions, *B. subtilis* can develop the state of natural competence in which cells can bind and adsorb extracellular DNA into the cytosol after which it can be integrated into the bacterial genome via homologous recombination (Dubnau, 1991; Hamoen et al., 2003). The processes of conjugation and competence have in common that they transport DNA via a sophisticated membrane-embedded DNA translocation machinery, but in opposite directions: DNA is exported in the case of conjugation, but imported during competence development. In both cases, only one of the two DNA strands (i.e., an ssDNA strand) is transported, and various other similarities exist between competence and conjugation related ssDNA transfer machines (Chen et al., 2005). However, simultaneous import and export of large ssDNA regions by the same cell may be incompatible for several reasons. For instance, simultaneous production and assembly of the ssDNA translocation machineries for competence and conjugation may interfere with each other and/or they may compete for the same cellular positions. In addition, the recombination enzymes synthesized during competence may act on ssDNA of the conjugative element. The following evidence indeed supports that conjugation and competence are incompatible processes and that at least two reciprocal mechanisms have evolved to avoid simultaneous activation of the competence and conjugation pathways in the same cell. Competence development is strongly suppressed in cells containing pLS20 by a plasmid-encoded protein called Rok_{pLS20}. Rok_{pLS20} shares similarity with *B. subtilis* Rok, which is a repressor of *comK* that encodes the transcriptional activator of the recombination and the structural competence genes. As for Rok, Rok_{pLS20} suppresses competence by repressing *comK* (Singh et al., 2012).

Like conjugation, competence is also induced by quorum sensing signals. The two-component ComP-ComA system plays a crucial role in the induction of competence: upon sensing the proper signals, the membrane-embedded kinase ComP activates ComA through phosphorylation which then activates competence genes, leading ultimately to the activation of the master regulator gene *comK* (Hamoen et al., 2003). RapF, encoded by the *B. subtilis* chromosome, inhibits stimulation of the competence pathway by interacting with ComA thereby preventing ComA from activating gene expression. *RapF* is cotranscribed with *phrF* that encodes the pre-proprotein which, after being processed, generates the signaling peptide PhrF*. High levels of PhrF* stimulate competence by inactivating RapF. Intriguingly, $Phr^{\star}_{pLS20}$ and PhrF* are very similar: residues at positions 1, 3 and 4 are identical; position 2 concerns a conserved substitution of Lysine to Arginine, and position 5 a change from Tyrosine to Isoleucine. PhrF* has been shown to cross talk with Rap_{pLS20} (Singh et al., 2020). Thus, PhrF* can interact with Rap_{pLS20} with a K_d that is only 2.5-fold higher than that of $Phr^{\star}_{pLS20}$ as determined by AUC (5.3 and 2.1 μM, respectively), and interaction with PhrF* results in tetramerization of Rap_{pLS20}. More importantly, PhrF* is able to inactivate Rap_{pLS20} *in vivo* when tested using the minimal regulatory system of pLS20 conjugation (see above).

CURRENT VIEW OF THE SIGNAL-PEPTIDE MEDIATED REGULATION OF pLS20 CONJUGATION

All the evidences obtained so far show that the conjugation genes of pLS20 are regulated in a highly sophisticated manner ensuring that conjugation is only activated under conditions that favor a successful conjugation event, i.e. when potential recipient cells are present. A schematic view of the current knowledge of how the conjugation genes are regulated is presented in **Figure 4**. The main and strong conjugation promoter P_c located upstream of the conjugation operon is repressed by Rco_{pLS20}, which also controls the activity of its own weak promoter P_r to ensure that only low levels of Rco_{pLS20} are produced. Despite these low levels, the tetrameric Rco_{pLS20} efficiently represses the P_c promoter. This is most likely achieved through DNA looping that ensures high local concentration of Rco_{pLS20} and simultaneous binding to multiple binding sites in the two operators. The low levels of Rco_{pLS20} probably are also crucial to allow for a sensitive and rapid switch to activate the conjugation genes when favorable conditions occur. Activation of the conjugation genes occurs by the action of the dimeric form of Rap_{pLS20} that relieves Rco_{pLS20}-mediated repression of the P_c promoter. Binding of Rap_{pLS20} to Rco_{pLS20} results in detachment of Rco_{pLS20} from the DNA, and hence it is the apo form of Rap_{pLS20} that activates conjugation. Accumulating levels of $Phr^{\star}_{pLS20}$ cause inactivation of Rap_{pLS20} by inducing a conformational change of Rap_{pLS20} that favors the formation of tetramers. Intriguingly, tetramerization occurs through interactions between the N-terminal effector domains with the consequence that these are presumably no longer available to interact with Rco_{pLS20}. Consequently, Rap_{pLS20} tetramerization results in the liberation of Rco_{pLS20} from the Rap_{pLS20}/Rco_{pLS20} complexes, which are then available to bind its operators and thereby returning the system to the default "OFF" state. Therefore, as schematically shown in **Figure 5**, the activation state of the pLS20 conjugation genes is ultimately regulated by the concentration of the $Phr^{\star}_{pLS20}$ signaling peptide. When all or a majority of the cells within a population contain the plasmid, $Phr^{\star}_{pLS20}$ concentrations will be high and conjugation will be inhibited. The conjugation genes will become activated when $Phr^{\star}_{pLS20}$ concentrations are low, which occurs when donor cells form a minority of the population and are surrounded by many potential recipient cells. In fact, the presence of recipient cells has a dual effect on lowering the $Phr^{\star}_{pLS20}$ concentration: 1) the relative number of donor cells,

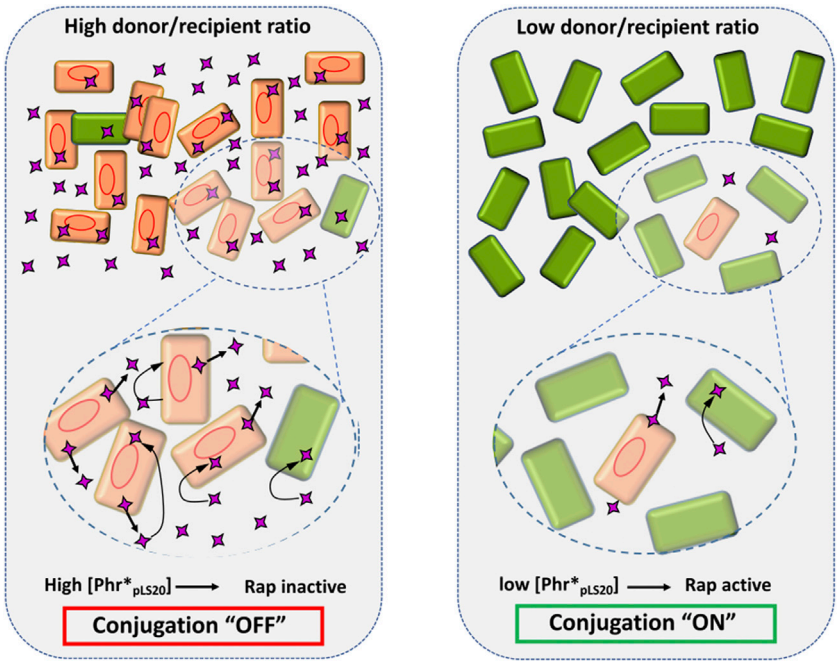


FIGURE 5 | The concentration of the signaling peptide Phr*$_{pLS20}$ determines the activation state of the conjugation genes. Schematic view of how a high and low donor/recipient ratio (left and right panel, respectively) affects the Phr*$_{pLS20}$ concentration, and thereby the activation state of the conjugation process. Donor and recipient cells are indicated with orange and green rectangles, respectively. Red oval lines, pLS20; purple four-pointed star, signaling peptide. The amplified regions in the lower part illustrate secretion of the signaling peptide by donor cells, and its adsorption by donor and recipient cells as indicated by straight and bent arrows, respectively. Note that for simplicity only the mature form of the signaling peptide is shown. See text for details.

which produce Phr*$_{pLS20}$, will be lower in populations containing recipient cells, and 2) recipient cells can actively adsorb the signaling peptide. Finally, it is not known if the Rap$_{pLS20}$-Phr*$_{pLS20}$ complexes are destined for rapid degradation, or whether there exists a mechanism to recycle Rap$_{pLS20}$ by removing Phr*$_{pLS20}$ from the complex, thereby stimulating the activation of the conjugation pathway. It should be noted, however, that the promoter P$_{rap}$ driving expression of *rap$_{pLS20}$* and *phr$_{pLS20}$* has been shown to be constitutive, thus resulting in *de novo* synthesis of cytosolic Rap$_{pLS20}$, and Phr$_{pLS20}$ that is destined for secretion (Singh et al., 2020).

COMPARISON OF REGULATION OF CONJUGATION GENES OF pLS20 AND OTHER CONJUGATIVE ELEMENTS REGULATED BY RRNPP PROTEINS

Signaling peptide receptor proteins belonging to the RRNPP family all have a similar structure that is composed of a small 3-helix N-terminal effector domain and a much larger CTD that, upon interaction with its cognate signal peptide, undergoes a conformational change affecting the functionality of the protein. Despite these conserved features, evolutionary adaptations have led to major effects on the mechanistic action of these proteins. This is exemplified here by comparing the three RRNPP members that regulate the conjugation genes of the conjugative plasmids pLS20 from *B. subtilis*, pCF10 from *Enterococcus faecalis*, and those of the ICE*Bs1* element that is present in many *B. subtilis* strains. The N-terminal effector domain of the pCF10-encoded RRNPP protein PrgX has evolved into a HTH DNA-binding domain. Remarkably, PrgX can interact with two signal peptides that cause opposing effects on its DNA binding. One of these signal peptides, iCF10, is encoded by the plasmid, and the other, cCF10 (the pheromone) by recipient cells. PrgX bound to the plasmid-encoded iCF10 adopts a conformation in which PrgX binds DNA resulting in repression of the conjugation genes. However, when PrgX binds to cCF10, DNA is released resulting in activation of the conjugation genes (for recent review see, Dunny and Berntsson, 2016). The regulatory system of pCF10 differs therefore in two important aspects from that of pLS20. First, in the case of pCF10 recipient cells are directly responsible for activation of the conjugation genes, and second, regulation of the conjugation genes is controlled by one instead of two regulatory proteins (PrgX vs. Rap$_{pLS20}$ and Rco$_{pLS20}$).

The conjugation genes of ICE*Bs*1 are repressed by a regulator named ImmR; an analogue of Rco$_{pLS20}$. Relieve of ImmR-mediated repression can occur in two ways. First, as a consequence of RecA-dependent SOS response to DNA damage, or second, by a mechanism that involves the ICE-encoded RRNPP protein RapI (Auchtung et al., 2007; Bose et al., 2008). However, activation of the conjugation genes by RapI occurs in a different way to Rap$_{pLS20}$. RapI is required but not sufficient to activate the conjugation genes; a protease encoded by gene *immA* that degrades ImmR in a RapI-

dependent manner is also needed. The exact mechanism of how RapI stimulates degradation of ImmR by ImmA is not yet known (Bose et al., 2008).

The presence and absence of a protease dedicated to degrade the repressor is likely to have important effects on reversibility of switching on and off of the conjugation pathway. Similar to Rco_{pLS20}, ImmR represses the conjugation promoter and simultaneously activates its own promoter (Auchtung et al., 2007). Degradation of ImmR will therefore result in activation of the conjugation genes and concomitantly inhibits *de novo* ImmR synthesis, causing conjugation of the ICE*Bs1* to be an irreversible process. In the case of pLS20, conjugation may be a reversible process, or at least it may be more flexible than the ICE*Bs1* system, because, rather than being degraded, Rco_{pLS20} becomes temporally sequestered by Rap_{pLS20}. Inactivation of Rap_{pLS20} by Phr_{pLS20} results in the release of Rco_{pLS20}, which can then bind to DNA and again repress the conjugation genes. Interaction of Rap_{pLS20} with Rco_{pLS20} might inactivate Rco_{pLS20} by altering its oligomerization state, in a similar way that RapF causes dissociation of its dimeric interacting partner ComA upon binding (Griffith and Grossman, 2008; Baker and Neiditch, 2011; Wolf et al., 2016). However, results of different analyses compellingly show that Rco_{pLS20} is released as functional tetramers from the Rap_{pLS20}-Rco_{pLS20} complex upon addition of Phr^*_{pLS20}, lending further support for the view that activation of the pLS20 conjugation genes can be a reversible process.

UNANSWERED QUESTIONS AND FUTURE PERSPECTIVES

Studies of the past years have resulted in major advances in understanding how the conjugation genes of pLS20 are regulated. Yet, multiple questions remain unanswered. Several of these concern multimerization of Rco_{pLS20} and how Rco_{pLS20} interacts with DNA and with Rap_{pLS20}. *In silico* analysis indicates that Rco_{pLS20} contains a HTH motif in its N-terminal region, which probably will be responsible for DNA binding. Therefore, the CTD of Rco_{pLS20} is presumably involved in tetramerization, but how the Rco_{pLS20} monomers interact to form a tetramer is unknown. Currently, it is also not exactly known how two tetramers interact when they are bound to DNA and induce DNA looping. Binding to DNA may generate a conformational change in the tetrameric structure of Rco_{pLS20} that favors octamerization. Alternatively, octamerization may be favored by a local high concentration as the consequence of Rco_{pLS20} binding to the intrinsically bent region containing operators O_I and O_{II}. The interacting surface between Rco_{pLS20} and Rap_{pLS20} is also not known. Rap_{pLS20} may bind the DNA binding surface of Rco_{pLS20} in an analogous way as RapF interacts with the DNA binding domain of ComA (Baker and Neiditch, 2011). However, the observation that Rap_{pLS20} appears to bind preferentially to Rco_{pLS20} tetramers that are involved in DNA looping (see above) suggests that Rap_{pLS20} binds preferentially, but not exclusively, to the Rco_{pLS20} surface generated upon dimerization of two Rco_{pLS20} tetramers; this interface most likely does not involve the DNA binding domains. Various experimental approaches have shown that Rco_{pLS20} does not bind DNA when it forms a complex with Rap_{pLS20} (Crespo et al., 2020; Singh et al., 2020). This may be because Rap_{pLS20} binds the DNA binding domain of Rco_{pLS20}. Alternatively, the formation of an Rco_{pLS20}/Rap_{pLS20} complex may induce a conformational change in Rco_{pLS20} affecting its DNA binding ability. Answers to several or all of these questions may be obtained by unraveling the apo structure of Rco_{pLS20} and the structures of Rco_{pLS20} complexed with DNA and with Rap_{pLS20}.

The activation of several developmental processes in *B. subtilis* such as sporulation, competence and motility depends on stochastic variability in expression of a master regulator that results in heterogeneic behavior of genetically identical cells within a population (Korobkova et al., 2004; Dubnau and Losick, 2006; Veening et al., 2008). A benefit of heterogeneity can be division of labor, where a subpopulation of cells produces products for the benefit of the entire community. Heterogeneity may also lead to the so-called bet-hedging strategies in which a subpopulation of differentiated cells is generated even in the absence of conditions favoring the differentiation process. Like division of labor, this strategy is beneficial for the community at the population level because the differentiated cells are adapted to adverse conditions that may arise suddenly. However, the process of conjugation is an energy consuming process and has major impacts on cell surface and membrane components, requiring tight repression at times when conditions for successful DNA transfer are not apt. Instead of a heterogeneic behavior resulting in a bet-hedging strategy, the conjugation process must be strictly controlled, allowing its activation to occur only under conditions that are favourable for successful transfer events, and this requires an efficient and sensitive genetic switch from an "OFF" to "ON" state only under conditions that favor successful plasmid transfer. Several features described above contribute to this efficient switch, such as overlapping divergent promoters, low levels of Rco_{pLS20} and Rap_{pLS20}, and DNA looping. We consider it possible that the multimerization state of Rco_{pLS20}, and particularly the apo and peptide bound forms of Rap_{pLS20} may also contribute to the sensitivity and efficiency of the switch.

Conjugation and competence appear to be incompatible processes. Although no evidence is currently available, one or two-way interference may also exist between conjugative plasmids and phage infection. Some *Bacillus* phages use a quorum sensing mechanism that plays a crucial role in entering the lytic or the lysogenic cycle (Erez et al., 2017). Perhaps, cross-talk exists between this quorum sensing system and the one regulating conjugation.

The work summarized in this review shows that conjugation is exquisitely controlled at the level of the P_c promoter. However, the conjugation operon is very large and contains many genes that are involved in different stages of the conjugation process, and it is very well possible that subsets of conjugation genes need to be expressed differentially or temporally. This may be achieved, at least in part, by different translation initiation signals, and

protein/RNA stability, but we have recently obtained evidence that at least one other layer of expression is involved. pLS20cat gene *29* encodes a surface-located protein that exerts exclusion; i.e., it inhibits futile transfer of the plasmid between donor cells, thereby contributing to efficient transfer to plasmid-free recipient cells (Gago-Cordoba et al., 2019). Proper functioning of exclusion systems require that the gene(s) involved are expressed in all donor cells, but gene *29* is located inside the conjugation operon. This apparent paradox was solved by the identification of a weak promoter located upstream of gene *29*, and a similar strategy is probably used in other conjugative systems as well (Gago-Cordoba et al., 2019). It is possible that the conjugation operon contains additional constitutive or regulated promoters that may modulate the expression of subsets of genes within the operon.

Contrary to most other conjugative systems, pLS20 conjugates efficiently in liquid medium and therefore it will have an important role in gene transfer under natural conditions. However, the predominant lifestyle of bacteria, including *B. subtilis*, is forming colonies or biofilms (Donlan, 2002). pLS20 also conjugates efficiently on solid medium (Koehler and Thorne, 1987; Meijer et al., 1998). As described above, the activation state of the conjugation process is ultimately determined by the concentration of the Phr^{*}_{pLS20} signaling peptide. The results summarized in this review are based on conjugation experiments carried out in liquid medium, and particularly using shaking cultures. Under these conditions, the concentration of the exported signal peptide will be instantly homogenized, which is very different from the situation in which donor cells form part of a colony or biofilm where secretion of the signal peptide will lead to a concentration gradient. It will be interesting to see how the conjugation genes are regulated under such conditions.

AUTHOR CONTRIBUTIONS

WM and LW wrote the paper. DB, SA, FR, CA, and JRL-O edited the paper, WM and SA made Figures.

FUNDING

Ministry of Science and Innovation of the Spanish Government (PID2019 108778GB C21 (AEI/FEDER, EU)) to WM, (BIO2016-77883-C2-2-P and FIS2015-72574-EXP (AEI/FEDER, EU)), to DB, (PID2019-104544GBI00 (AEI/FEDER/EU)) to CA, (PID2019-109320GB-100/AEI/10.13039/501100011033) to SA, (PGC2018-094143-B-I00, (AEI/FEDER, EU)) to FR, and (FIS2016-78313-P) to SA. The Ministry of Science and Innovation of the Spanish Government also funded the Centro Nacional de Biotecnología (CNB) by means of the "Severo Ochoa" Center of Excellence grant SEV 2017-0712. A Wellcome Investigator grant (209500) to Jeff Errington supported LW. Institutional grants from the 'Fundación Ramón Areces' and 'Banco de Santander' supported the Centro de Biología Molecular 'Severo Ochoa'.

ACKNOWLEDGMENTS

We thank members of the different labs for useful discussions, and Daniel Zeigler of the Bacillus Genetic Stock Center for providing plasmids and strains essential for many experiments described in this review. We acknowledge the publication fee support of the CSIC Open Access Publication Support Initiative through its 'Unit of Information Resources for Research' (URICI).

REFERENCES

Arutyunov, D., and Frost, L. S. (2013). F conjugation: back to the beginning. *Plasmid* 70 (1), 18–32. doi:10.1016/j.plasmid.2013.03.010

Auchtung, J. M., Lee, C. A., Garrison, K. L., and Grossman, A. D. (2007). Identification and characterization of the immunity repressor (ImmR) that controls the mobile genetic element ICEBs1 of *Bacillus subtilis*. *Mol. Microbiol.* 64 (6), 1515–1528. doi:10.1111/j.1365-2958.2007.05748.x

Baker, M. D., and Neiditch, M. B. (2011). Structural basis of response regulator inhibition by a bacterial anti-activator protein. *PLoS Biol.* 9 (12), e1001226. doi:10.1371/journal.pbio.1001226

Baxter, J. C., and Funnell, B. E. (2014). Plasmid partition mechanisms. *Microbiol. Spectr.* 2 (6) 1–20. doi:10.1128/microbiolspec.PLAS-0023-2014

Bongiorni, C., Stoessel, R., Shoemaker, D., and Perego, M. (2006). Rap phosphatase of virulence plasmid pXO1 inhibits *Bacillus anthracis* sporulation. *J. Bacteriol.* 188 (2), 487–498. doi:10.1128/JB.188.2.487-498.2006

Bose, B., Auchtung, J. M., Lee, C. A., and Grossman, A. D. (2008). A conserved anti-repressor controls horizontal gene transfer by proteolysis. *Mol. Microbiol.* 70 (3), 570–582. doi:10.1111/j.1365-2958.2008.06414.x

Bouet, J. Y., and Funnell, B. E. (2019). Plasmid localization and partition in Enterobacteriaceae. *EcoSal Plus* 8 (2) 1–23. doi:10.1128/ecosalplus.ESP-0003-2019

Brooks, A. C., and Hwang, L. C. (2017). Reconstitutions of plasmid partition systems and their mechanisms. *Plasmid* 91, 37–41. doi:10.1016/j.plasmid.2017.03.004

Cabezón, E., Ripoll-Rozada, J., Peña, A., De la Cruz, F., and Arechaga, I. (2014). Towards an integrated model of bacterial conjugation. *FEMS Microbiol. Rev.* 39, 81–95. doi:10.1111/1574-6976.12085

CDC (2019). *Antibiotic resistance threats in the United States*. Atlanta, GA: U.S. Department of health and human services.

Chastanet, A., and Losick, R. (2011). Just-in-time control of Spo0A synthesis in *Bacillus subtilis* by multiple regulatory mechanisms. *J. Bacteriol.* 193 (22), 6366–6374. doi:10.1128/JB.06057-11

Chatterjee, A., Cook, L. C., Shu, C. C., Chen, Y., Manias, D. A., Ramkrishna, D., et al. (2013). Antagonistic self-sensing and mate-sensing signaling controls antibiotic-resistance transfer. *Proc. Natl. Acad. Sci. U.S.A.* 110 (17), 7086–7090. doi:10.1073/pnas.1212256110

Chen, I., Christie, P. J., and Dubnau, D. (2005). The ins and outs of DNA transfer in bacteria. *Science* 310 (5753), 1456–1460. doi:10.1126/science.114021

Christie, P. J. (2016). The mosaic type IV secretion systems. *EcoSal Plus* 7 (1), 10. doi:10.1128/ecosalplus.ESP-0020-2015

Christie, P. J., Whitaker, N., and González-Rivera, C. (2014). Mechanism and structure of the bacterial type IV secretion systems. *Biochim. Biophys. Acta* 1843 (8), 1578–1591. doi:10.1016/j.bbamcr.2013.12.019

Clewell, D. B. (2011). Tales of conjugation and sex pheromones: a plasmid and enterococcal odyssey. *Mob. Genet. Elem.* 1 (1), 38–54. doi:10.4161/mge.1.1.15409

Crespo, I., Bernardo, N., Miguel-Arribas, A., Singh, P. K., Luque-Ortega, J. R., Alfonso, C., et al. (2020). Inactivation of the dimeric RappLS20 anti-repressor of the conjugation operon is mediated by peptide-induced tetramerization. *Nucleic Acids Res.* 48, 8113. doi:10.1093/nar/gkaa540

Cutting, S. M. (2011). Bacillus probiotics. *Food Microbiol.* 28 (2), 214–220. doi:10.1016/j.fm.2010.03.007

D'Andrea, L. D., and Regan, L. (2003). TPR proteins: the versatile helix. *Trends Biochem. Sci.* 28 (12), 655–662. doi:10.1016/j.tibs.2003.10.007

Davies, J., and Davies, D. (2010). Origins and evolution of antibiotic resistance. *Microbiol. Mol. Biol. Rev.* 74 (3), 417–433. doi:10.1128/MMBR.00016-10

Declerck, N., Bouillaut, L., Chaix, D., Rugani, N., Slamti, L., Hoh, F., et al. (2007). Structure of PlcR: insights into virulence regulation and evolution of quorum sensing in gram-positive bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 104 (47), 18490–18495. doi:10.1073/pnas.0704501104

Derman, A. I., Becker, E. C., Truong, B. D., Fujioka, A., Tucey, T. M., Erb, M. L., et al. (2009). Phylogenetic analysis identifies many uncharacterized actin-like proteins (Alps) in bacteria: regulated polymerization, dynamic instability and treadmilling in Alp7A. *Mol. Microbiol.* 73 (4), 534–552. doi:10.1111/j.1365-2958.2009.06771.x

Dodd, I. B., Shearwin, K. E., and Egan, J. B. (2005). Revisited gene regulation in bacteriophage lambda. *Curr. Opin. Genet. Dev.* 15 (2), 145–152. doi:10.1016/j.gde.2005.02.001

Donlan, R. M. (2002). Biofilms: microbial life on surfaces. *Emer. Infect. Dis.* 8 (9), 881–890. doi:10.3201/eid0809.020063

Dubnau, D. (1991). Genetic competence in *Bacillus subtilis*. *Microbiol. Rev.* 55, 395–424. doi:10.1128/mmbr.55.3.395-424.1991

Dubnau, D., and Losick, R. (2006). Bistability in bacteria. *Mol. Microbiol.* 61 (3), 564–572. doi:10.1111/j.1365-2958.2006.05249.x

Dunny, G. M. (2013). Enterococcal sex pheromones: signaling, social behavior, and evolution. *Annu. Rev. Genet.* 47, 457–482. doi:10.1146/annurev-genet-111212-133449

Dunny, G. M., and Berntsson, R. P. (2016). Enterococcal sex pheromones: evolutionary pathways to complex, two-signal systems. *J. Bacteriol.* 198 (11), 1556–1562. doi:10.1128/JB.00128-16

Dunny, G. M., and Johnson, C. M. (2011). Regulatory circuits controlling enterococcal conjugation: lessons for functional genomics. *Curr. Opin. Microbiol.* 14 (2), 174–180. doi:10.1016/j.mib.2011.01.008

Erez, Z., Steinberger-Levy, I., Shamir, M., Doron, S., Stokar-Avihail, A., Peleg, Y., et al. (2017). Communication between viruses guides lysis-lysogeny decisions. *Nature* 541 (7638), 488–493. doi:10.1038/nature21049

Frost, L. S., and Koraimann, G. (2010). Regulation of bacterial conjugation: balancing opportunity with adversity. *Future Microbiol.* 5 (7), 1057–1071. doi:10.2217/fmb.10.70

Frost, L. S., Leplae, R., Summers, A. O., and Toussaint, A. (2005). Mobile genetic elements: the agents of open source evolution. *Nat. Rev. Microbiol.* 3, 722–732. doi:10.1038/nrmicro1235

Gago-Córdoba, C., Val-Calvo, J., Miguel-Arribas, A., Serrano, E., Singh, P. K., Abia, D., et al. (2019). Surface exclusion revisited: function related to differential expression of the surface exclusion system of *Bacillus subtilis* plasmid pLS20. *Front. Microbiol.* 10, 1502. doi:10.3389/fmicb.2019.01502

Goessweiner-Mohr, N., Arends, K., Keller, W., and Grohmann, E. (2014). Conjugation in gram-positive bacteria. *Microbiol. Spectr.* 2 (4) 1–19. doi:10.1128/microbiolspec.PLAS-0004-2013

Griffith, K. L., and Grossman, A. D. (2008). A degenerate tripartite DNA-binding site required for activation of ComA-dependent quorum response gene expression in *Bacillus subtilis*. *J. Mol. Biol.* 381 (2), 261–275. doi:10.1016/j.jmb.2008.06.035

Grohmann, E., Muth, G., and Espinosa, M. (2003). Conjugative plasmid transfer in gram-positive bacteria. *Microbiol. Mol. Biol. Rev.* 67 (2), 277–contents. doi:10.1128/mmbr.67.2.277-301.2003

Grossman, A. D. (1995). Genetic networks controlling the initiation of sporulation and the development of genetic competence in *Bacillus subtilis*. *Annu. Rev. Genet.* 29, 477–508. doi:10.1146/annurev.ge.29.120195.002401

Hamoen, L. W., Venema, G., and Kuipers, O. P. (2003). Controlling competence in *Bacillus subtilis*: shared use of regulators. *Microbiology (Reading, Engl)* 149, 9–17. doi:10.1099/mic.0.26003-0

Hoch, J. A. (1993). "*spo0* genes, the phosphorelay, and the initiation of sporulation," in Bacillus subtilis *and other gram-positive bacteria: biochemistry, physiology, and molecular genetics*. Editors A. L. Sonenshein, J. A. Hoch, and R. Losick (Washington, DC: American Society for Microbiology (ASM)), 747–755.

Itaya, M., Sakaya, N., Matsunaga, S., Fujita, K., and Kaneko, S. (2006). Conjugational transfer kinetics of pLS20 between *Bacillus subtilis* in liquid medium. *Biosci. Biotechnol. Biochem.* 70, 740–742. doi:10.1271/bbb.70.740

Koehler, T. M., and Thorne, C. B. (1987). *Bacillus subtilis (natto)* plasmid pLS20 mediates interspecies plasmid transfer. *J. Bacteriol.* 169 (11), 5271–5278. doi:10.1128/jb.169.11.5271-5278.1987

Kohler, V., Keller, W., and Grohmann, E. (2019). Regulation of gram-positive conjugation. *Front. Microbiol.* 10, 1134. doi:10.3389/fmicb.2019.01134

Koraimann, G. (2018). Spread and persistence of virulence and antibiotic resistance genes: a ride on the F plasmid conjugation module. *EcoSal Plus* 8 (1) 1–23. doi:10.1128/ecosalplus.ESP-0003-2018

Koraimann, G., and Wagner, M. A. (2014). Social behavior and decision making in bacterial conjugation. *Front. Cell Infect. Microbiol.* 4, 54. doi:10.3389/fcimb.2014.00054

Korobkova, E., Emonet, T., Vilar, J. M., Shimizu, T. S., and Cluzel, P. (2004). From molecular noise to behavioural variability in a single bacterium. *Nature* 428 (6982), 574–578. doi:10.1038/nature02404

Lerminiaux, N. A., and Cameron, A. D. S. (2019). Horizontal transfer of antibiotic resistance genes in clinical environments. *Can. J. Microbiol.* 65 (1), 34–44. doi:10.1139/cjm-2018-0275

Li, Y. G., Hu, B., and Christie, P. J. (2019). Biological and structural diversity of type IV secretion systems. *Microbiol. Spectr.* 7 (2) 1–10. doi:10.1128/microbiolspec.PSIB-0012-2018

Mazel, D., and Davies, J. (1999). Antibiotic resistance in microbes. *Cell Mol. Life Sci.* 56 (9–10), 742–754. doi:10.1007/s000180050021

Meijer, W. J., de Boer, A., van Tongeren, S., Venema, G., and Bron, S. (1995). Characterization of the replication region of the *Bacillus subtilis* plasmid pLS20: a novel type of replicon. *Nucleic Acids Res.* 23, 3214–3223. doi:10.1093/nar/23.16.3214

Meijer, W. J., Wisman, G. B., Terpstra, P., Thorsted, P. B., Thomas, C. M., Holsappel, S., et al. (1998). Rolling-circle plasmids from *Bacillus subtilis*: complete nucleotide sequences and analyses of genes of pTA1015, pTA1040, pTA1050 and pTA1060, and comparisons with related plasmids from gram-positive bacteria. *FEMS Microbiol. Rev.* 21 (4), 337–368. doi:10.1111/j.1574-6976.1998.tb00357.x

Molle, V., Fujita, M., Jensen, S. T., Eichenberger, P., González-Pastor, J. E., Liu, J. S., et al. (2003). The Spo0A regulon of *Bacillus subtilis*. *Mol. Microbiol.* 50 (5), 1683–1701. doi:10.1046/j.1365-2958.2003.03818.x

Monnet, V., Juillard, V., and Gardan, R. (2016). Peptide conversations in gram-positive bacteria. *Crit. Rev. Microbiol.* 42 (3), 339–351. doi:10.3109/1040841X.2014.948804

Neiditch, M. B., Capodagli, G. C., Prehna, G., and Federle, M. J. (2017). Genetic and structural analyses of RRNPP intercellular peptide signaling of gram-positive bacteria. *Annu. Rev. Genet.* 51, 311–333. doi:10.1146/annurev-genet-120116-023507

Novick, R. P., Christie, G. E., and Penadés, J. R. (2010). The phage-related chromosomal islands of gram-positive bacteria. *Nat. Rev. Microbiol.* 8 (8), 541–551. doi:10.1038/nrmicro2393

Ochman, H., Lawrence, J. G., and Groisman, E. A. (2000). Lateral gene transfer and the nature of bacterial innovation. *Nature* 405 (6784), 299–304. doi:10.1038/35012500

Oehler, S., and Müller-Hill, B. (2010). High local concentration: a fundamental strategy of life. *J. Mol. Biol.* 395 (2), 242–253. doi:10.1016/j.jmb.2009.10.056

Oppenheim, A. B., Kobiler, O., Stavans, J., Court, D. L., and Adhya, S. (2005). Switches in bacteriophage lambda development. *Annu. Rev. Genet.* 39, 409–429. doi:10.1146/annurev.genet.39.073003.113656

Parashar, V., Konkol, M. A., Kearns, D. B., and Neiditch, M. B. (2013). A plasmid-encoded phosphatase regulates *Bacillus subtilis* biofilm architecture, sporulation, and genetic competence. *J. Bacteriol.* 195 (10), 2437–2448. doi:10.1128/JB.02030-12

Partridge, S. R., Kwong, S. M., Firth, N., and Jensen, S. O. (2018). Mobile genetic elements associated with antimicrobial resistance. *Clin. Microbiol. Rev.* 31 (4), e00088-17. doi:10.1128/CMR.00088-17

Perego, M. (2001). A new family of aspartyl phosphate phosphatases targeting the sporulation transcription factor Spo0A of *Bacillus subtilis*. *Mol. Microbiol.* 42 (1), 133–143. doi:10.1046/j.1365-2958.2001.02611.x

Pottathil, M., and Lazazzera, B. A. (2003). The extracellular Phr peptide-Rap phosphatase signaling circuit of *Bacillus subtilis*. *Front. Biosci.* 8, 32–45. doi:10.2741/913

Ramachandran, G., Singh, P. K., Luque-Ortega, J. R., Yuste, L., Alfonso, C., Rojo, F., et al. (2014). A complex genetic switch involving overlapping divergent promoters and DNA looping regulates expression of conjugation genes of a gram-positive plasmid. *PLoS Genet.* 10 (10), e1004733. doi:10.1371/journal.pgen.1004733

Rocha-Estrada, J., Aceves-Diez, A. E., Guarneros, G., and de la Torre, M. (2010). The RNPP family of quorum-sensing proteins in gram-positive bacteria. *Appl. Microbiol. Biotechnol.* 87 (3), 913–923. doi:10.1007/s00253-010-2651-y

Shi, K., Brown, C. K., Gu, Z. Y., Kozlowicz, B. K., Dunny, G. M., Ohlendorf, D. H., et al. (2005). Structure of peptide sex pheromone receptor PrgX and PrgX/pheromone complexes and regulation of conjugation in *Enterococcus faecalis*. *Proc. Natl. Acad. Sci. U.S.A.* 102 (51), 18596–18601. doi:10.1073/pnas10.1073/pnas.0506163102

Singh, P. K., and Meijer, W. J. (2014). Diverse regulatory circuits for transfer of conjugative elements. *FEMS Microbiol. Lett.* 358 (2), 119–128. doi:10.1111/1574-6968.12526

Singh, P. K., Ramachandran, G., Durán-Alcalde, L., Alonso, C., Wu, L. J., and Meijer, W. J. (2012). Inhibition of *Bacillus subtilis* natural competence by a native, conjugative plasmid-encoded comK repressor protein. *Environ. Microbiol.* 14 (10), 2812–2825. doi:10.1111/j.1462-2920.2012.02819.x

Singh, P. K., Ramachandran, G., Ramos-Ruiz, R., Peiró-Pastor, R., Abia, D., Wu, L. J., et al. (2013). Mobility of the native *Bacillus subtilis* conjugative plasmid pLS20 is regulated by intercellular signaling. *PLoS Genet.* 9 (10), e1003892. doi:10.1371/journal.pgen.1003892

Singh, P. K., Serrano, E., Ramachandran, G., Miguel-Arribas, A., Gago-Cordoba, C., Val-Calvo, J., et al. (2020). Reversible regulation of conjugation of *Bacillus subtilis* plasmid pLS20 by the quorum sensing peptide responsive anti-repressor RappLS20. *Nucleic Acids Res.* 48 (19), 10785–10801. doi:10.1093/nar/gkaa797

Smillie, C., Garcillán-Barcia, M. P., Francia, M. V., Rocha, E. P., and De la Cruz, F. (2010). Mobility of plasmids. *Microbiol. Mol. Biol. Rev.* 74 (3), 434–452. doi:10.1128/MMBR.00020-10

Smits, W. K., Bongiorni, C., Veening, J. W., Hamoen, L. W., Kuipers, O. P., and Perego, M. (2007). Temporal separation of distinct differentiation pathways by a dual specificity Rap-Phr system in *Bacillus subtilis*. *Mol. Microbiol.* 65 (1), 103–120. doi:10.1111/j.1365-2958.2007.05776.x

Sonenshein, A. L., Hoch, J. A., and Losick, R. (1993). *Bacillus subtilis and other Gram-positive bacteria; Biochemistry, physiology, and molecular genetics*. Washington, DC: American Society for Microbiology (ASM).

Sonenshein, A. L., Hoch, J. A., and Losick, R. (2001). *Bacillus subtilis and its closest relatives: from genes to cells*. Washington, DC: ASM Press.

Soucy, S. M., Huang, J., and Gogarten, J. P. (2015). Horizontal gene transfer: building the web of life. *Nat. Rev. Genet.* 16 (8), 472–482. doi:10.1038/nrg3962

Stingl, K., and Koraimann, G. (2017). Prokaryotic information games: how and when to take up and secrete DNA. *Curr. Top. Microbiol. Immunol.* 413, 61–92. doi:10.1007/978-3-319-75241-9_3

Tanaka, T., and Koshikawa, T. (1977). Isolation and characterization of four types of plasmids from *Bacillus subtilis (natto)*. *J. Bacteriol.* 131, 699–701. doi:10.1128/JB.131.2.699-701.1977

Thomas, C. M., and Nielsen, K. M. (2005). Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nat. Rev. Microbiol.* 3 (9), 711–721. doi:10.1038/nrmicro1234

Veening, J. W., Smits, W. K., and Kuipers, O. P. (2008). Bistability, epigenetics, and bet-hedging in bacteria. *Annu. Rev. Microbiol.* 62, 193–210. doi:10.1146/annurev.micro.62.081307.163002

Vilar, J. M., and Saiz, L. (2005). DNA looping in gene regulation: from the assembly of macromolecular complexes to the control of transcriptional noise. *Curr. Opin. Genet. Dev.* 15 (2), 136–144. doi:10.1016/j.gde.2005.02.005

Waksman, G. (2019). From conjugation to T4S systems in gram-negative bacteria: a mechanistic biology perspective. *EMBO Rep.* 20 (2), e47012. doi:10.15252/embr.201847012

Wolf, D., Rippa, V., Mobarec, J. C., Sauer, P., Adlung, L., Kolb, P., et al. (2016). The quorum-sensing regulator ComA from *Bacillus subtilis* activates transcription using topologically distinct DNA motifs. *Nucleic Acids Res.* 44 (5), 2160–2172. doi:10.1093/nar/gkv1242

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Meijer, Boer, Ares, Alfonso, Luque-Ortega and Wu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership