DNA REPLICATION ORIGINS IN MICROBIAL GENOMES, VOLUME 2

EDITED BY: Feng Gao and Alan C. Leonard

frontiers Research Topics

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: researchtopics@frontiersin.org

DNA REPLICATION ORIGINS IN MICROBIAL GENOMES, VOLUME 2

Topic Editors:
Feng Gao, Tianjin University, China
Alan C. Leonard, Florida Institute of Technology, United States
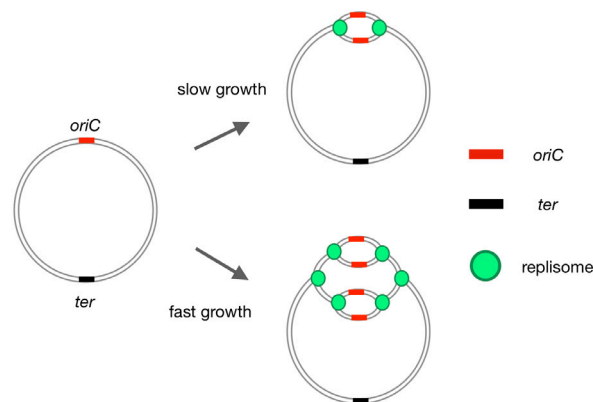
Image: Figure 1 from Trojanowski et al. (2018). Trojanowski D, Hołówka J and Zakrzewska-Czerwińska J (2018) Where and When Bacterial Chromosome Replication Starts: A Single Cell Perspective. *Front. Microbiol*. 9:2819.
doi: 10.3389/fmicb.2018.02819

As guest editor, Prof. Gao has organized the Research Topic "DNA Replication Origins in Microbial Genomes" for Frontiers in Microbiology. Gratifyingly, the papers published in this Research Topic were highly accessed, and well-received by a wide international audience. Given its previous success, we decided to revisit this Research Topic with a second volume. We are pleased that this topic remains one of keen interest, and also surprised by the diversity of the manuscripts submitted for the second volume. The field is certainly moving in interesting new directions. We hope that readers find these articles both informative and entertaining, and we look forward to an exciting future for replication origin research.

Dedicated to the 125th Anniversary of Tianjin University (formerly Peiyang University), the first modern higher education university in China.

Table of Contents

frontiers
in Microbiology

Editorial: DNA Replication Origins in Microbial Genomes, Volume 2

Feng Gao [1,2,3]* and Alan C. Leonard [4]*

[1] *Department of Physics, School of Science, Tianjin University, Tianjin, China,* [2] *Frontier Science Center of Synthetic Biology (MOE), Key Laboratory of Systems Bioengineering (MOE), Tianjin University, Tianjin, China,* [3] *SynBio Research Platform, Collaborative Innovation Center of Chemical Science and Engineering, Tianjin, China,* [4] *Laboratory of Microbial Genetics, Department of Biomedical and Chemical Engineering and Science, Florida Institute of Technology, Melbourne, FL, United States*

Editorial on the Research Topic

DNA Replication Origins in Microbial Genomes, Volume 2

As guest editor, Prof. Gao has organized the Research Topic "DNA Replication Origins in Microbial Genomes" for Frontiers in Microbiology (Gao, 2016). Gratifyingly, the papers published in this volume were highly accessed, and well-received by a wide international audience. Given its previous success we decided to revisit this Research Topic with a second edition in 2017.

We are pleased that this topic remains one of keen interest, and also surprised by the diversity of the manuscripts submitted for the second edition. The field is certainly moving in interesting new directions. We present a total of 11 articles, including 6 original research articles, 4 reviews, and one general commentary, all having undergone rigorous peer review.

Although *Escherichia coli* remains the classic model for studying the mechanisms of DNA replication and regulation in bacteria, there are still uncharted territories and even some surprises. The unique replication origin (*oriC*) encodes instructions for assembly of the initiator protein, DnaA-ATP, into complexes (orisomes) required for the initiation step (Leonard and Méchali, 2013; Wolanski et al., 2015; Katayama et al., 2017), but it remains unclear how orisomes unwind DNA and assist with loading DnaB helicase onto the single-strands. New insights are provided by Sakiyama et al., in this volume, including a model to explain the mechanism of DnaB loading in *E. coli*, and evidence that DnaA AAA+ domain His136 residue directs DnaB to the unwound region. Based on recent studies that show synthetic versions of *oriC* can be activated by the normally inactive DnaA-ADP (Grimwade et al., 2018), Leonard et al. present a new perspective on the requirement for DnaA-ATP in orisome function and timing regulation, and suggest that in *E. coli*, DnaA-ATP is needed for site recognition and occupation instead of mechanical functions. Post-initiation, *E. coli oriC* is sequestered by SeqA protein to prevent re-replication. Surprisingly, Ser36 in the SeqA protein is a target for phosphorylation by the serine-threonine kinase, HipA (Semanjski et al., 2018). However, in this volume, questions about this interesting regulatory feature are raised by Riber et al., who show that mutating the Ser36 residue to alanine (and the loss of phosphorylation) does not affect replication initiation.

Vibrio cholerae has emerged as an important model system due to a genome comprising two chromosomes. Many questions are raised about the regulation of origin licensing and once-per cycle replication, as well as chromosome partitioning in multi-chromosome bacteria. These topics are well-represented in this volume. Fournes et al. review once-per cycle regulation of secondary chromosomes with an insightful perspective based on plasmid systems. One of the key checkpoint regulators of *V. cholerae* chromosome II is a region of chromosome I called *crtS* (Baek and Chattoraj, 2014; Val et al., 2016). Based on an *in vivo* screen, Ciaccia et al. show that a global

transcription factor, Lrp, binds to *crtS* and plays an important role as a licensing factor for chromosome II. In addition to this crosstalk regulation between transcription and chromosome replication, crosstalk must also exist between bacterial chromosome replication and chromosome partitioning (Marczynski et al.; Taylor et al., 2017). Marczynski et al. review replication-partition crosstalk and discuss how *Vibrio cholerae*, has evolved separate and specific replication and partitioning crosstalk systems for its chromosomes. Important for current and future studies are methods to visualize *oriC* regions, chromosome replication, and partitioning in living bacterial cells (for example, see Ginda et al., 2017; Ramachandran et al., 2018). Here, Trojanowski et al. present an in-depth review on single cell imaging methods.

Unexpectedly, *Vibrio cholera* (NSCV1 and NSCV2) strains were found to contain a single chromosome with two replication origins (Xie et al., 2017), adding another level of intrigue to the two chromosome story. In this volume, Bruhn et al. found that both origins can be active (NSCV1) or one origin can be silenced (NSCV2). It is now clear that multi-origin bacterial chromosomes are more prevalent than anticipated (Gao, 2015; Luo and Gao, 2019; Luo et al., 2019), and some thought-provoking issues of regulation raised by this condition are presented here in a commentary by Das and Chattoraj.

It is clear from the two remaining manuscripts in this volume, that the hunt for replication origins on chromosomes remains a worthwhile effort. Jaworski et al. present the novel structure and function of *oriC* in *Campylobacter jejuni*, the bacterium associated with most foodborne infections worldwide. Eukaryotic microbes must also be included, and Wang and Gao present a comprehensive study of *S. cerevisiae* replication origins from genome-wide and population genomics perspectives.

We hope that readers find these articles both informative and entertaining, and we look forward to an exciting future for replication origin research.

AUTHOR CONTRIBUTIONS

FG and AL wrote the manuscript. Both authors read and approved the final manuscript.

FUNDING

REFERENCES

Baek, J. H., and Chattoraj, D. K. (2014). Chromosome I controls chromosome II replication in *Vibrio cholerae*. *PLoS Genet.* 10:e1004184. doi: 10.1371/journal.pgen.1004184

Gao, F. (2015). Bacteria may have multiple replication origins. *Front. Microbiol.* 6:324. doi: 10.3389/fmicb.2015.00324

Gao, F. (2016). Editorial: DNA replication origins in microbial genomes. *Front. Microbiol.* 6:1545. doi: 10.3389/fmicb.2015.01545

Ginda, K., Santi, I., Bousbaine, D., Zakrzewska-Czerwinska, J., Jakimowicz, D., and McKinney, J. (2017). The studies of ParA and ParB dynamics reveal asymmetry of chromosome segregation in *Mycobacteria*. *Mol. Microbiol.* 105, 453–468. doi: 10.1111/mmi.13712

Grimwade, J. E., Rozgaja, T. A., Gupta, R., Dyson, K., Rao, P., and Leonard, A. C. (2018). Origin recognition is the predominant role for DnaA-ATP in initiation of chromosome replication. *Nucleic Acids Res.* 46, 6140–6151. doi: 10.1093/nar/gky457

Katayama, T., Kasho, K., and Kawakami, H. (2017). The DnaA cycle in *Escherichia coli*: activation, function and inactivation of the initiator protein. *Front. Microbiol.* 8:2496. doi: 10.3389/fmicb.2017.02496

Leonard, A. C., and Méchali, M. (2013). DNA replication origins. *Cold Spring Harb. Perspect. Biol.* 5:a010116. doi: 10.1101/cshperspect.a010116

Luo, H., and Gao, F. (2019). DoriC 10.0: an updated database of replication origins in prokaryotic genomes including chromosomes and plasmids. *Nucleic Acids Res.* 47, D74–D77. doi: 10.1093/nar/gky1014

Luo, H., Quan, C. L., Peng, C., and Gao, F. (2019). Recent development of Ori-Finder system and DoriC database for microbial replication origins. *Brief. Bioinform.* 20, 1114–1124. doi: 10.1093/bib/bbx174

Ramachandran, R., Ciaccia, P. N., Filsuf, T. A., Jha, J. K., and Chattoraj, D. K. (2018). Chromosome 1 licenses chromosome 2 replication in *Vibrio cholerae* by doubling the *crtS* gene dosage. *PLoS Genet.* 14:e1007426. doi: 10.1371/journal.pgen.1007426

Semanjski, M., Germain, E., Bratl, K., Kiessling, A., Gerdes, K., and Macek, B. (2018). The kinases HipA and HipA7 phosphorylate different substrate pools in *Escherichia coli* to promote multidrug tolerance. *Sci. Signal.* 11:eat5750. doi: 10.1126/scisignal.aat5750

Taylor, J. A., Panis, G., Viollier, P. H., and Marczynski, G. T. (2017). A novel nucleoid-associated protein coordinates chromosome replication and chromosome partition. *Nucleic Acids Res.* 45, 8916–8929. doi: 10.1093/nar/gkx596

Val, M.-E.,Marbouty, M., de LemosMartins, F., Kennedy, S. P., Kemble, H., Bland, M. J., et al. (2016). A checkpoint control orchestrates the replication of the two chromosomes of *Vibrio cholerae*. *Sci. Adv.* 2:e1501914. doi: 10.1126/sciadv.1501914

Wolanski, M., Donczew, R., Zawilak-Pawlik, A., and Zakrzewska-Czerwinska, J. (2015). *oriC*-encoded instructions for the initiation of bacterial chromosome replication. *Front. Microbiol.* 5:735. doi: 10.3389/fmicb.2014.00735

Xie, G., Johnson, S. L., Davenport, K. W., Rajavel, M., Waldminghaus, T., Detter, J. C., et al. (2017). Exception to the rule: genomic characterization of naturally occurring unusual *Vibrio cholerae* strains with a single chromosome. *Int. J. Genomics* 2017:8724304. doi: 10.1155/2017/8724304

The DnaA AAA+ Domain His136 Residue Directs DnaB Replicative Helicase to the Unwound Region of the Replication Origin, *oriC*

Yukari Sakiyama[†‡], Masahiro Nishimura[†‡], Chihiro Hayashi, Yusuke Akama, Shogo Ozaki and Tsutomu Katayama*

Department of Molecular Biology, Graduate School of Pharmaceutical Sciences, Kyushu University, Fukuoka, Japan

Chromosomal replication initiation requires dynamic mechanisms in higher-order nucleoprotein complexes that are constructed at the origin of replication. In *Escherichia coli*, DnaA molecules construct functional oligomers at the origin *oriC*, enabling localized unwinding of *oriC* and stable binding of DnaB helicases via multiple domain I molecules of *oriC*-bound DnaA. DnaA-bound DnaB helicases are then loaded onto the unwound region of *oriC* for construction of a pair of replisomes for bidirectional replication. However, mechanisms of DnaB loading to the unwound *oriC* remain largely elusive. In this study, we determined that His136 of DnaA domain III has an important role in loading of DnaB helicases onto the unwound *oriC*. DnaA H136A mutant protein was impaired in replication initiation *in vivo*, and in DnaB loading to the unwound *oriC in vitro*, whereas the protein fully sustained activities for *oriC* unwinding and DnaA domain I-dependent stable binding between DnaA and DnaB. Functional and structural analyses supported the idea that transient weak interactions between DnaB helicase and DnaA His136 within specific protomers of DnaA oligomers direct DnaB to a region in close proximity to single stranded DNA at unwound *oriC* bound to DnaA domain III of the DnaA oligomer. The aromatic moiety of His136 is basically conserved at corresponding residues of eubacterial DnaA orthologs, implying that the guidance function of DnaB is common to all eubacterial species.

Keywords: *E. coli, oriC*, DnaA, helicase, AAA+ family, protein–protein interaction

INTRODUCTION

Chromosomal DNA replication is initiated by synergistic mechanisms involving multiple proteins with various functions. The initial steps of replication in *Escherichia coli* occur at the unique replication origin, *oriC*, which has a sophisticated structure that directs unwinding of duplex DNA and loading of replicative helicases (Kaguni, 2011; Leonard and Grimwade, 2015; Wolański et al., 2015; Katayama et al., 2017; **Figure 1A**). In these steps, the initiator DnaA molecules construct specific oligomers with the aid of DiaA (a DnaA-binding protein) and appropriately located DnaA-binding sequences (DnaA boxes) in the *oriC* DnaA-oligomerization region (DOR) (Leonard and Grimwade, 2015; Katayama et al., 2017). ATP-bound DnaA (ATP-DnaA), but not ADP-bound DnaA (ADP-DnaA), efficiently constructs homo-oligomers in a head-to-tail manner, with DnaA–DnaA interactions via the bound ATP and the Arg285 residue of flanking DnaA
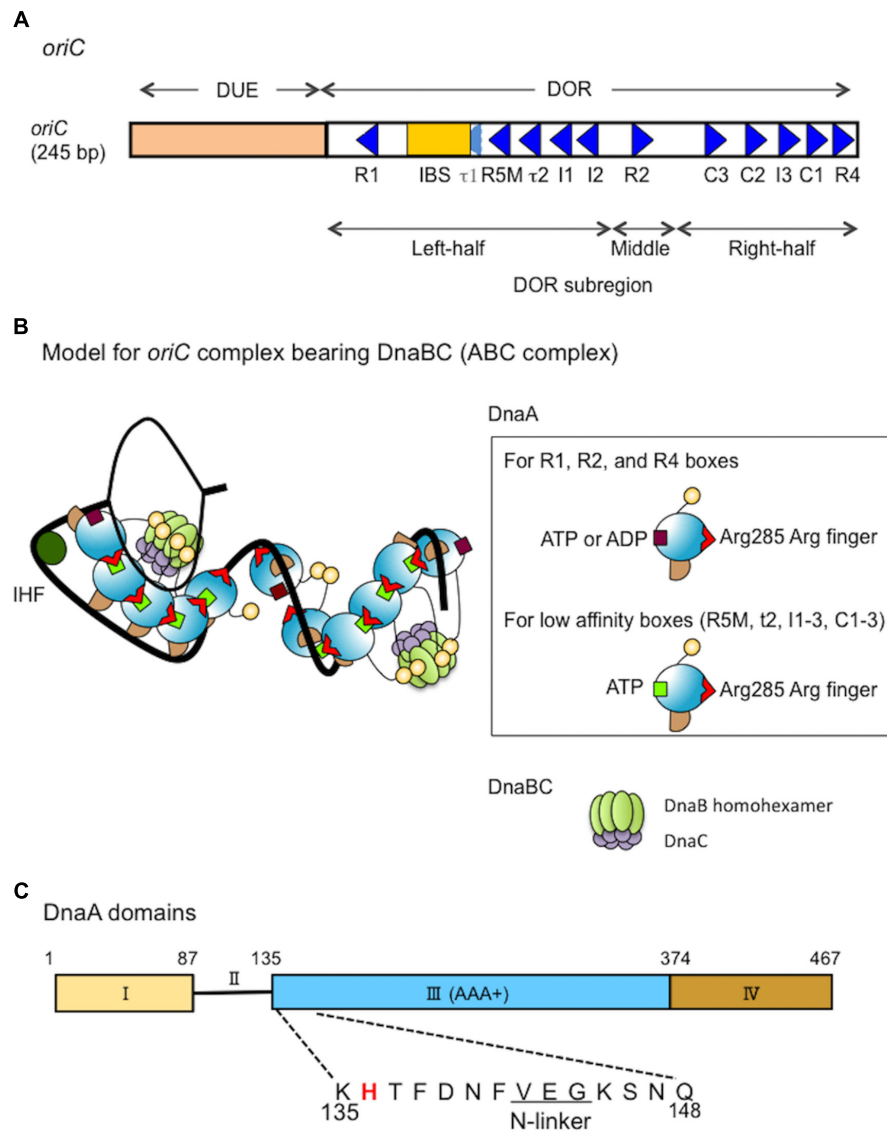
FIGURE 1 | Structures of *oriC*, DnaA and the complex bearing DnaA and DnaBC. **(A)** Structure of *Escherichia coli oriC*. The duplex unwinding element (DUE) and DnaA-oligomerization region (DOR) are indicated. DnaA boxes (consensus sequence TTATNCACA) are shown by triangles, oriented to indicate sequence directionality (McGarry et al., 2004; Kawakami et al., 2005; Rozgaja et al., 2011; Shimizu et al., 2016). The R1 and R4 boxes have high affinity, the R2 box has moderate affinity and the others have low affinity for DnaA. A single integration host factor (IHF)-binding site (IBS) is present. When IHF binds to IBS, DnaA does not bind to the τ1 box (Sakiyama et al., 2017). **(B)** A possible structural model for *oriC*–DnaA complex carrying DnaBC complexes (ABC complex). DnaA domains are colored as in **C**. DnaA oligomers are constructed, and in one DnaA protomer, Arg285 "arginine finger" interacts with ATP bound to the flanking DnaA protomer (Kawakami et al., 2005; Erzberger et al., 2006). In the DnaA oligomers on *oriC* DOR (black line) (Rozgaja et al., 2011), Arg285 faces the middle of the DOR (Noguchi et al., 2015). DnaA domain IV can swivel using a short flexible loop present in its N-terminus, which supports DnaA oligomerization on DOR (Shimizu et al., 2016). In a crystal structure of the DNA-free oligomers of *A. aeolicus* DnaA domain III and domain IV, domain IV of one protomer interacts with domain III of the flanking protomer (Erzberger et al., 2002). By contrast, in molecular dynamics modeling of *E. coli* DnaA complexes at *oriC*, the structural changes induced by DNA binding and swiveling of domain IV prevents domain III-domain IV interaction (Shimizu et al., 2016). We propose that, near the time of initiation, when cellular ATP-DnaA levels peak (Kurokawa et al., 1999; Fujimitsu et al., 2009; Katayama et al., 2010), ATP-DnaA molecules bind to the low affinity DnaA box clusters R5M-τ2-I1-I2 and C3-C2-I3-C1 (Kawakami et al., 2005; Ozaki et al., 2008, 2012a; Keyamura et al., 2009; Sakiyama et al., 2017). We note that other studies (McGarry et al., 2004; Grimwade et al., 2007, 2018) using different methodologies to detect DnaA binding, report that R5M and C1 bind ATP-DnaA and ADP-DnaA with similar affinities. The illustration presents a model at the time of initiation. For stable DUE unwinding, ssDUE binds to DnaA domain III (Ozaki et al., 2008; Duderstadt et al., 2011; Sakiyama et al., 2017). Interactions between domain I of multiple DnaA protomers and each DnaB homohexamer promote stable binding (Abe et al., 2007; Keyamura et al., 2009). Deletion analysis of *oriC* supports the idea that each DnaA subcomplex (one on the left-half DOR and one on the right-half DOR) binds a single DnaB–DnaC complex (Ozaki and Katayama, 2012), although which DnaA protomer binds DnaB is unclear. In addition, the orientation of the DnaA subcomplex at the right-half DOR is important for efficient DnaB loading (Shimizu et al., 2016). When the DUE is unwound (thin black line), DnaB helicases are loaded onto the ssDNA regions and DnaC is released (Kaguni, 2011). For simplicity, DiaA is not shown in this model. **(C)** Domain structure of DnaA. His136 is highlighted in red and the N-linker motif is underlined. See text for details.

molecules (Kawakami et al., 2005; Erzberger et al., 2006; Ozaki et al., 2012a; Noguchi et al., 2015; Sakiyama et al., 2017; **Figures 1B,C**). The left-half DOR adjoins the *oriC* duplex unwinding element (DUE) and includes a specific binding site (IBS) for IHF (Integration host factor), a DNA-bending protein. The DnaA subcomplex including the left-half DOR and IHF causes localized unwinding of DNA in the DUE (Bramhill and Kornberg, 1988; Ozaki and Katayama, 2012; **Figure 1B**). The unwound single-stranded (ss) DUE is stabilized by binding to the DnaA oligomer, which is a prerequisite for DnaB loading (Ozaki et al., 2008; Ozaki and Katayama, 2012; Sakiyama et al., 2017).

DnaB helicase-loading includes critical processes for transition from replication initiation to DNA elongation. DnaB–DnaC complexes bind to DnaA oligomers that are bound to *oriC* (Keyamura et al., 2009; Kaguni, 2011; Ozaki and Katayama, 2012; Soultanas, 2012; Zawilak-Pawlik et al., 2017; **Figures 1B,C**). DnaB is a replicative helicase with a homohexamer structure and ring (or spiral) configuration (Kaguni, 2011; Itsathitphaisarn et al., 2012; Strycharska et al., 2013). DnaC acts in the loading of DnaB by stably binding to DnaB and promoting conformational changes in DnaB hexamer that are required for its loading on the ssDNA (Davey and O'Donnell, 2003; Galletto et al., 2003; Biswas and Biswas-Fiss, 2006; Makowska-Grzyska and Kaguni, 2010; Kaguni, 2011). DnaB C-terminal domain (CTD) is suggested to bind DnaC (Galletto et al., 2003). A pair of DnaB–DnaC complexes is thought to bind to an *oriC*–DnaA complex, resulting in a higher-order complex (Keyamura et al., 2009; Ozaki and Katayama, 2012; Ozaki et al., 2012a; Shimizu et al., 2016; **Figure 1B**): a deletion analysis of *oriC* suggests that each DnaA subcomplex constructed on the left- and right-half subregions of DOR binds a single DnaB helicase (or DnaBC complex) (Ozaki and Katayama, 2012). Also, deletion and insertion analyses suggest that the orientation of DnaA subcomplex constructed on the right-half DOR is optimized for efficient DnaB loading (Shimizu et al., 2016). The two DnaB helicases in the complex are loaded onto the ssDNA region of *oriC* in opposite directions each other, enabling bidirectional migration, leading to loading of one replisome on each strand (Fang et al., 1999; Carr and Kaguni, 2001; Ozaki and Katayama, 2012; Soultanas, 2012; Bell and Kaguni, 2013).

Unlike hyperthermophile bacterium *Aquifex aeolicus* DnaC (Mott et al., 2008), *E. coli* DnaC does not stably bind to DnaA (Keyamura et al., 2009). In *E. coli*, DnaB in DnaBC complex binds to DnaA complexes (Abe et al., 2007; Keyamura et al., 2009) and this mechanism is suggested to be conserved even in *A. aeolicus* (Mott et al., 2008). In addition, we recently demonstrate that YfdR, a protein encoded by an *E. coli* cryptic prophage, binds to DnaA depending on domain I Phe46 (a primary DnaB-binding site; also see below) and competes with the DnaBC complex in DnaA binding (Noguchi and Katayama, 2016). These results also support the notion that, in *E. coli*, DnaC *per se* does not directly bind DnaA, but does so indirectly as part of the DnaB–DnaC complex, in which DnaB binds to DnaA. However, dynamic mechanisms involved in DnaA–DnaB interactions have remained to be further elucidated.

The initiator protein DnaA has four functional domains (Kaguni, 2011; Katayama et al., 2017; **Figure 1C**). Domain I contains a specific site for binding to DnaB and DiaA (Sutton et al., 1998; Seitz et al., 2000; Abe et al., 2007; Keyamura et al., 2009). Domain II is a flexible linker between domains I and III (Nozaki and Ogawa, 2008). Domain III is the AAA+ domain that participates in nucleotide binding, ssDNA recruitment, and DnaA–DnaA interactions (Felczak and Kaguni, 2004; Iyer et al., 2004; Kawakami et al., 2005; Ozaki and Katayama, 2009; Duderstadt et al., 2011; Ozaki et al., 2012a,b). In addition, DnaA domain III N-terminal Val142–Gly144 constitutes a specific N-linker motif, which is thought to structurally interact with the adenine moiety of ATP/ADP (Smith et al., 2004), and our previous results demonstrate that the well-conserved Glu143 residue is specifically important for stable ATP/ADP binding (Ozaki et al., 2012b). DnaA domains II–III are also thought to have a weak binding site for DnaB (Marszalek et al., 1996; Seitz et al., 2000), although the precise location of this site has not been determined. DnaA domain IV binds directly to DnaA boxes, which has the 9-mer consensus sequence of (5′)TTATNCACA(3′) (Sutton and Kaguni, 1997; Fujikawa et al., 2003; Kaguni, 2011). In the N-terminus of domain IV, a short flexible loop enables the swiveling of this domain (Erzberger et al., 2002; Shimizu et al., 2016).

Functional mechanisms of DnaA–DnaB interaction for loading of DnaB on the ssDUE are thought to include multiple steps (Sutton et al., 1998; Seitz et al., 2000; Abe et al., 2007; Keyamura et al., 2009). DnaA domain I is the primary source of weak affinity for the DnaB CTD (Sutton et al., 1998; Seitz et al., 2000). We previously determined that a patch of DnaA that includes the Glu21 and Phe46 residues is exposed on the surface of domain I, binds to DnaB, and supports stable DnaB binding when DnaA oligomers are constructed on *oriC* (Abe et al., 2007; Keyamura et al., 2009). As DnaB is a homohexamer, binding of a single DnaB hexamer to multiple domain I molecules of a DnaA oligomer would effectively increase its affinity for DnaA, stabilizing DnaA–DnaB binding (Abe et al., 2007; Keyamura et al., 2009; Ozaki and Katayama, 2009; Zawilak-Pawlik et al., 2017). Although DnaA domain I is suggested to interact with a site including the N-terminus and its flanking region of DnaB CTD (Seitz et al., 2000), specific amino acids in the region have not been determined.

In addition to domain I, DnaA domains II–III are thought to contain a second site for DnaB binding, which is important for DnaB loading on *oriC* DNA (**Figure 1C**). A previous study assessed specific inhibition with monoclonal anti-DnaA antibodies, and the results suggested that the DnaA Pro111–Gln148 region includes a site for DnaB interaction (Marszalek and Kaguni, 1994; Marszalek et al., 1996; Sutton et al., 1998). Another study assessed functional interactions of various truncated forms of DnaA, and the results suggested that the DnaA Ser130–Gln148 region has a specific interaction site for the DnaB N-terminal domain (NTD) (Seitz et al., 2000). Deletion analysis of domain II has demonstrated that the DnaA Ala99–Val134 region is largely dispensable for DnaA functions in replication initiation at *oriC* (Nozaki and Ogawa, 2008). Taken together, these results suggest that the DnaA domain III

N-terminus spanning Lys135–Gln148 might contain the second essential site for DnaB interaction (**Figure 1C**).

In this study, to determine the role for the second DnaB-binding site of DnaA, we extended functional analysis of the DnaA domain III N-terminus to the region spanning Lys135–Gln148. Alanine-scanning experiments revealed that His136, Phe141, and Val142 are crucial for complementation of *dnaA46* temperature-sensitive mutations. Our previous study indicated that Val142 is an essential constituent of the N-linker (Ozaki et al., 2012b), and Phe141 is its flanking bulky residue. Thus, in this study, we focused our attention on analyses on His136 and found that a substitution of this residue (H136A) impaired specifically DnaA-dependent loading of DnaB on ssDUE, without affecting its assembly at *oriC*, unwinding of DUE, or recruitment of ssDUE. A structural model for the *oriC*–DnaA complex is compatible with a predicted role for His136 in directing DnaB for loading on the ssDUE.

MATERIALS AND METHODS

Nucleic Acids

Plasmid pKA234, a derivative of the pING vector which has an arabinose-inducible promoter, was used for overproduction of wild-type DnaA, and has been described previously (Ozaki et al., 2012b). Derivatives of pKA234 encoding DnaA variants with individual alanine substitutions for each residue from Lys135A to Gln148 were constructed with specific mutagenic primers and QuikChange site-directed mutagenesis protocol (Stratagene [Agilent], Agilent, La Jolla, CA, United States), as previously described (Ozaki et al., 2012b).

The DOR dsDNA fragment (ΔDUE) and 28-mer T-rich ssDUE strand have been described previously (Ozaki and Katayama, 2012). M13KEW101 and pBSoriC are *oriC* plasmids containing intact *oriC*, and pBSoriCΔR4-R2 is a derivative of pBSoriC with only the left half of *oriC*. M13KEW101, pBSoriC, and pBSoriCΔR4-R2 have been described previously (Kawakami et al., 2005; Ozaki and Katayama, 2012). pBSoriCΔDUE is a DUE-deleted derivative of pBSoriC, constructed by outward-directed PCR with pBSoriC as a template and primers ori-1 and MR28r1-r, as described previously (Ozaki et al., 2008; Ozaki and Katayama, 2012); the amplified DNA was digested with *Hinc*II and self-ligated, resulting in pBSoriCΔDUE. Biotinylated *oriC* DNA (bio-*oriC*) has been described previously (Keyamura et al., 2009). M13-A-site ssDNA is a derivative of M13 ssDNA with a hairpin structure that contains a DnaA-box sequence (Masai et al., 1990).

Buffers

Buffer G contained 20 mM HEPES-KOH (pH 7.6), 5 mM magnesium acetate, 1 mM EDTA, 10% glycerol, 0.1% Triton X-100, and 4 mM dithiothreitol (DTT). Buffer F contained 20 mM Tris–HCl (pH 7.5), 8 mM DTT, 10 mM magnesium acetate, 125 mM potassium glutamate, 3 mM ATP, and 0.5 mg/mL bovine serum albumin (BSA). Buffer ABC contained 20 mM Tris–HCl (pH 7.5), 0.1 mg/mL BSA, 8 mM DTT, 8 mM magnesium acetate, 0.01% Brij-58, and 125 mM potassium glutamate.

DnaA Proteins

Wild-type DnaA and DnaA H136A proteins were overproduced in *E. coli* strain KA450 [Δ*oriC1071*::Tn*10 rnhA199*(Am) *dnaA17*(Am)] from pKA234 or pH136A (a derivative of pKA234 encoding DnaA H136A), and purified as previously described (Ozaki et al., 2012b).

Flow-Cytometry Analysis

Flow cytometry was performed as described previously (Noguchi et al., 2015; Inoue et al., 2016; Sakiyama et al., 2017). Briefly, cells were grown in LB medium at 30°C to an absorbance at 660 nm of 0.2, then incubated at 42°C for 150 min. Before and after 42°C incubation, aliquots were withdrawn for analysis of cell mass (or cell volume) with a FACSCalibur flow cytometer (BD Biosciences, Franklin Lakes, NJ, United States). The remaining aliquots of the cell cultures were further incubated in the presence of rifampicin and cephalexin for 4 h, followed by analysis of DNA content by flow cytometry.

oriC Plasmid Replication Assay

The assay was performed essentially as described previously (Keyamura et al., 2009). Briefly, a crude protein extract containing proteins required for *oriC* replication (except for DnaA) was prepared from *E. coli* strain WM433 (*dnaA204*) (Fuller et al., 1981). Reactions (25 µL) were performed with purified DnaA, M13KEW101 *oriC* plasmid (38 fmol as plasmid; 600 pmol nucleotides), and WM433 crude extract (200 µg), as described (Keyamura et al., 2009).

P1 Nuclease Assay for DUE Unwinding

The assay was performed essentially as described previously (Ozaki et al., 2012b). Briefly, ATP-DnaA or ADP-DnaA was incubated for 3 min at 38°C with M13KEW101 *oriC* plasmid (12 fmol as plasmid; 400 ng) and HU protein (16 ng), followed by brief incubation with P1 nuclease. DNA was purified and incubated with restriction enzyme *Bse*RI, followed by analysis by agarose-gel electrophoresis, with ethidium bromide staining.

ssDUE-Recruitment Electrophoretic Mobility Shift Assay (EMSA)

The assay was performed as described previously (Ozaki and Katayama, 2012; Ozaki et al., 2012a; Sakiyama et al., 2017). Briefly, DnaA and the DUE-deleted *oriC* dsDNA fragment, DOR(ΔDUE) (5 nM) were incubated on ice for 5 min, followed by incubation for 5 min at 30°C with ^{32}P-labeled 28-mer T-rich ssDUE strand (2.5 nM) in the presence of λphage DNA (25 ng). The resultant DNA complexes were analyzed by 4% polyacrylamide-gel electrophoresis at 4°C.

Form I* Assay

The assay was performed as previously described (Ozaki and Katayama, 2012; Noguchi et al., 2015; Shimizu et al.,

2016). Briefly, ATP-DnaA and pBSoriC or pBSoriCΔR4-R2 (1.6 nM) were incubated for 15 min at 30°C in buffer (25 μL) containing 760 nM SSB, 76 nM GryA, 100 nM His-GyrB, 42 nM IHF, 100 nM His-DnaB, and 100 nM His-DnaC. Reactions were terminated by addition of 1% SDS, and the DNA samples were purified and analyzed by 0.65% agarose-gel electrophoresis and ethidium bromide staining.

Biotin-Tagged *oriC* Pull-Down Assay

The assay was performed as previously described (Keyamura et al., 2009; Ozaki and Katayama, 2012; Ozaki et al., 2012a; Noguchi and Katayama, 2016). Briefly, bio-*oriC* (419 bp, 100 fmol), including the entire *oriC* and flanking regions (Keyamura et al., 2007), was incubated on ice for 10 min in buffer G (10 μL) containing DnaA and 1 mM ATP. The bio-*oriC* and bound proteins were recovered by pull-down with streptavidin-coated beads (Promega, Madison, WI, United States), washed twice with buffer G (12.5 μL) containing 75 mM KCl, and dissolved in SDS sample buffer. Proteins were quantified using silver staining and quantitative protein standards, as we previously performed (Ozaki and Katayama, 2012; Noguchi and Katayama, 2016). The recovered amounts of bio-*oriC* were deduced by quantifying bio-*oriC* remaining in supernatants. For analysis of DnaB binding to *oriC*–DnaA complexes, bio-*oriC* was coincubated with DnaA and DnaB in the presence of absence of DnaC, followed by the wash step conducted only once with buffer G excluding KCl.

His-DnaB Pull-Down Assay

The assay was performed under similar conditions to the Form I* assay. ATP-DnaA (90 nM) and pBSoriC or pBSoriCΔDUE (1.6 nM) were incubated for 15 min at 30°C in buffer F (10 μL) containing 40 nM IHF, 200 nM His-DnaB K236A, and the native (non-tagged) DnaC (200 nM). After addition of Co^{2+}-conjugated magnetic beads (1 μL bed volume: Dynabeads, Invitrogen, Carlsbad, CA, United States) and incubation for 15 min at 4°C, the beads and bound materials were collected by magnetic pull-down and washed in buffer F containing 100 mM NaCl and excluding BSA. His-DnaB-bound plasmid DNA was eluted in standard SDS sample buffer, and analyzed by 1% agarose-gel electrophoresis and ethidium bromide staining.

ABC Primosome Assay

The assay was performed as described previously (Abe et al., 2007; Keyamura et al., 2009). Briefly, the indicated amounts of DnaA were incubated at 30°C for 15 min in buffer ABC (25 μL) containing M13-A-site ssDNA (1.1 nM as ssDNA; 220 pmol nucleotides), 0.5 μg SSB, 65 ng DnaB, 65 ng DnaC, 72 ng DnaG, 108 ng DNA polymerase III*, 26 ng β-clamp subunit, 1 mM ATP, 0.25 mM each of GTP, CTP, and UTP, and 0.1 mM each of dNTP and [α-^{32}P]dATP. DNA polymerase III* is a subcomplex of DNA polymerase III holoenzyme lacking the β-clamp subunit. Reactions were stopped by addition of 1 mL 10% trichloroacetic acid, and

the amounts of synthesized DNA were measured by liquid scintillation.

RESULTS

DnaA His136 Is Essential for Initiation Activity *in vivo*

A putative DnaB interaction site has been speculated to reside in a DnaA domain III N-terminal region spanning Lys135–Gln148. We conducted an alanine-scanning analysis on all the amino acid residues in this region except for Gly144, which lacks a side chain. Plasmid pKA234 contains the wild-type *dnaA*-coding region downstream of the arabinose-inducible P$_{BAD}$ promoter (Kubota et al., 1997). Derivatives of pKA234 containing *dnaA* mutations encoding alanine substitutions were constructed and used for complementation tests with a temperature (42°C)-sensitive *dnaA46* host strain (KA413). Even in the absence of the inducer arabinose, introduction of pKA234, but not the vector pING1, enabled the KA413 cells to grow at 42°C, presumably because of leaky expression of the pKA234 *dnaA* gene (**Table 1**). In similar experiments that we previously performed, DnaA amounts in cells bearing pKA234 was 1.3- to 2.8-fold higher than those in cells bearing pING1 (Kawakami et al., 2005; Ozaki et al., 2008). Unlike pKA234, plasmids containing *dnaA* alleles encoding H136A, F141A, or V142A substitutions did not complement the *dnaA46* temperature (42°C) sensitive growth (**Table 1**).

Val142 is a component of the N-linker motif (Val142-Glu143-Gly144), a conserved sequence in AAA+ family proteins that is thought to interact with the adenine moiety of ATP (Smith et al., 2004). DnaA V142A has previously been shown to cause overinitiation of replication at 30°C in the absence of wild-type DnaA presumably because its binding to ADP is unstable, resulting in rapid exchange of bound ADP to ATP (Ozaki et al., 2012b), and here consistently, we observed that DnaA V142A resulted in slow colony formation at 30°C and severe inhibition of colony formation at 42°C (**Table 1**). DnaA F141A also resulted in growth inhibition at 42°C, suggesting that substitution of the hydrophobic aromatic side chain of phenylalanine resulted in DnaA structural changes that indirectly inhibited the function of the N-linker motif residues, such as Val142.

We previously indicated that Glu143 within the N-linker motif is important for stable binding of ATP and ADP, and DnaA E143A causes moderate overinitiation of replication at 30°C in the absence of wild-type DnaA, presumably because of rapid exchange of bound ADP to ATP (Ozaki et al., 2012b). When co-expressed with the wild-type DnaA, DnaA E143A was shown not to cause inhibition of cell growth (Ozaki et al., 2012b), which might be a consequence of formation of mixed complexes on *oriC* repressing overinitiation (Ozaki et al., 2012b). These previous observations are consistent with the current data of DnaA E143A at 30°C (**Table 1**). When cells were incubated at 42°C, a temperature at which replication initiation is stimulated even in wild-type cells, a severe inhibitory effect of DnaA E143A for cell growth by

TABLE 1 | Results of plasmid complementation.

Plasmid	*dnaA* allele	Transformation efficiency (cfu/μg DNA)		Ratio (42/30°C)
		30°C	42°C	
pING1(vector)	None	2.8×10^6	$<1.1 \times 10^3$	$<3.9 \times 10^{-4}$
pKA234	Wild-type	2.0×10^5	3.1×10^5	1.6
pK135A	*K135A*	2.1×10^6	1.5×10^6	1.4
pH136A	*H136A*	1.3×10^6	$<1.1 \times 10^3$	$<8.5 \times 10^{-4}$
pT137A	*T137A*	1.4×10^6	8.0×10^5	0.57
pF138A	*F138A*	1.5×10^6	9.0×10^5	0.60
pD139A	*D139A*	1.3×10^6	6.6×10^5	0.51
pN140A	*N140A*	1.5×10^6	9.0×10^5	1.6
pF141A	*F141A*	1.5×10^6	$<2.0 \times 10^3$	$<1.3 \times 10^{-3}$
pV142A	*V142A*	5.6×10^6*	$<1.1 \times 10^3$	$<2.0 \times 10^{-4}$
pE143A	*E143A*	9.6×10^5	2.1×10^6	2.2
pK145A	*K145A*	1.9×10^6	1.2×10^6	1.6
pS146A	*S146A*	1.5×10^6	1.5×10^6	1.0
pN147A	*N147A*	6.5×10^6*	8.9×10^6	1.4
pQ148A	*Q148A*	2.0×10^6	1.3×10^6	1.5

*Plasmids were introduced into Escherichia coli KA413 (dnaA46) cells, which were incubated either at 30°C for 24 h or at 42°C for 12 h on LB-agar plates containing 50 μg/mL thymine and 100 μg/mL ampicillin. The colonies were then counted as described previously (Nishida et al., 2002; Ozaki et al., 2012a). Transformation efficiency [colony forming units (cfu)/μg DNA] at each temperature, and ratios of efficiencies, are shown. *Only tiny colonies were observed by 24 h, so incubation was prolonged to 36 h before colony counting.*

overinitiation might be suppressed. Based on these ideas, we excluded DnaA F141A and V142A for further analyses in this study.

To further elucidate initiation activity *in vivo* of DnaA H136A, we used flow cytometry to determine the chromosomal replication modes of *dnaA46* cells containing derivatives of pKA234. Cells were cultured for exponential growth at 30°C and shifted to 42°C for inactivation of the intrinsic DnaA46. Further incubation with rifampicin and cefalexin enabled run-out replication of the chromosomes. At 30°C, distinct DNA peaks were seen for each strain (**Figure 2**). In *dnaA46* cells with the pING vector, the two-chromosome peak predominated, with minor peaks for one, three, four, and five chromosomes, indicating asynchronous initiations (Skarstad et al., 1986, 1988). Introduction of wild-type *dnaA* (pKA234) into *dnaA46* cells stimulated initiation and resulted in a predominant four-chromosome peak. DnaA F46A protein is impaired in stable DnaB binding and DnaB loading at *oriC* (Abe et al., 2007; Keyamura et al., 2009). Introduction of the *dnaA F46A* allele as a negative control (pF46A) gave a similar profile to that of *dnaA46* cells with the pING vector, whereas moderate stimulation was detected with *dnaA H136A* (pH136A). This stimulation could be a consequence of mixture of DnaA46 and DnaA H136A proteins (see below). At 42°C, the profiles with pING and *dnaA H136A* were fundamentally similar with respect to inhibition of initiation, whereas moderate stimulation occurred with *dnaA F46A*, as previously observed (Abe et al., 2007; Keyamura et al., 2009). These results suggest that DnaA H136A is impaired in the initiation of chromosomal replication *in vivo*. In addition, the smeared peaks of *dnaA H136A* culture at 42°C suggest that progression of replication forks also are moderately inhibited by this mutation. Abnormal interaction between DnaA H136A and DnaB helicases might inhibit fork progression in addition to replication initiation *in vivo*.

Purified DnaA H136A Protein Sustains ATP Binding, but Is Inactive in Initiation *in vitro*

DnaA H136A protein was overproduced and purified, and shown to have high-affinity binding of ATP and ADP at levels similar to wild-type DnaA (**Table 2**), suggesting preservation of the overall protein structure of domain III. By contrast, when replication initiation was assessed with an *oriC* plasmid and a replicative-protein extract, even the ATP form of DnaA H136A was inactive, unlike that of wild-type DnaA (**Figure 3A**).

DnaA H136A Is Active in *oriC* Unwinding and ssDUE Binding

Functions of DnaA His136 were assessed with reconstituted systems for *oriC* unwinding and ssDUE binding. In the P1-nuclease assay, unwinding of *oriC* DUE by initiation complexes produces ssDNA that is sensitive to endonuclease P1. In this assay, the ATP forms of wild-type DnaA and DnaA H136A demonstrated similar activities in specific *oriC* unwinding (**Figure 3B**).

To determine the abilities of these proteins to stabilize the unwound DUE, we analyzed ssDUE-binding activity by electrophoretic mobility shift assay (EMSA) using ssDUE, *oriC* DOR, and DnaA. In this assay, ATP-DnaA molecules can construct homo-oligomers on the DOR and bind ssDUE with high affinity, producing DOR-DnaA-ssDUE ternary complexes (Ozaki et al., 2008; Ozaki and Katayama, 2012; Noguchi et al., 2015; Sakiyama et al., 2017). Construction of the ternary complexes has been shown to occur with ATP-DnaA (but not

FIGURE 2 | Flow-cytometry analysis. KA413 (*dnaA46*) cells containing the indicated plasmids were grown to exponential phase in LB medium containing 50 μg/mL thymine and 50 μg/mL ampicillin at 30°C, and further incubated at 42°C for 150 min. Before and after 42°C incubation, portions of the cultures were withdrawn for analyses of cell size (mass) and DNA content by flow cytometry. Chromosome numbers per cell corresponding to each peak are indicated. Mean cell mass relative to that of KA413 cells containing pING1 and grown at 30°C is indicated at the top right corner of each panel. pKA234 is a DnaA-expressing derivative of pING1. pF46A and pH136A are derivatives of pKA234 encoding variants of DnaA with single amino acid substitutions.

ADP-DnaA) and to require specific DnaA residues involved in construction of *oriC* open complexes (including AAA+ arginine-finger Arg285 and ssDUE-binding H/B motifs Val211 and Arg245) (Ozaki et al., 2008; Ozaki and Katayama, 2012).

TABLE 2 | Binding of ATP and ADP by wild-type (WT) and H136A DnaA.

DnaA	K_D (nM)		Stoichiometry	
	ATP	ADP	ATP	ADP
WT	24	35	0.24	0.17
H136A	21	50	0.10	0.13

DnaA protein (76 nM) was incubated on ice for 15 min with various concentrations of [α-^{32}P]ATP and [2,8-^{3}H]ADP. Nucleotide-bound DnaA was captured on nitrocellulose filters, levels of bound nucleotides were determined by liquid-scintillation counting, and affinities of DnaA proteins for ATP or ADP were deduced from Scatchard plots, as previously described (Ozaki et al., 2012a,b).

Here, the ATP form of DnaA H136A resulted in binding of the ssDUE to the DnaA–DOR complexes at a level similar to that seen with wild-type DnaA (**Figure 3C**). DnaA-ssDUE binding was dependent on DOR as we previously demonstrated. Faint signals in the gel wells were irregular aggregates of DnaA which involved ^{32}P-ssDUE. Those were unstable and slowly degraded during electrophoresis, resulting in faint smeared bands (Ozaki et al., 2008; Ozaki and Katayama, 2012; Sakiyama et al., 2017). These results indicate that DnaA H136A is fully active in the primary reactions required for DUE unwinding and stable binding of ssDUE, which are prerequisites for DnaB loading on ssDUE.

DnaA H136A Is Impaired in DnaB Loading Onto Unwound *oriC*

We conducted the Form I* assay to determine DnaB loading onto the unwound strands of *oriC* (via indirect interactions with DnaC and DnaA) and its subsequent helicase action on DNA strands. Loading on ssDUE activates DnaB helicase, expanding the ssDNA region and introducing positive supercoils. Activation of gyrase then produces highly negatively supercoiled *oriC* plasmid (Form I*), and this topoisomeric form is distinguished from Form I by gel electrophoresis (Baker et al., 1986).

Form I* production was severely impaired for DnaA H136A, even at high levels (**Figure 4A**). Compared with wild-type DnaA, DnaA H136A, and the negative control DnaA F46A (which is inactive in primary DnaB binding) only had low levels of Form I* production, even in the presence of excessive amounts of DnaBC proteins (**Figure 4B**). The left-half *oriC* is a minimal region for DUE unwinding and DnaB loading (Ozaki and Katayama, 2012; Sakiyama et al., 2017). Compared with wild-type DnaA, DnaA H136A, and DnaA F46A also produced only low levels of Form I* even with left-half *oriC* (**Figure 4C**), consistent with the idea that DnaB loading reactions *per se* are inhibited also with the full-length *oriC*, but not with the idea that DnaBC complexes binds simultaneously to left- and right-half DnaA–*oriC* subcomplexes, causing abortive interactions with each other, and inhibiting DnaB loading on ssDUE. Taken together, the results suggest that the DnaA His136 residue has a predominant role in the process of productive loading of DnaB to unwound *oriC*.

A Subgroup of DnaA Molecules in an *oriC* Complex Requires His136 for DnaB Loading

Here, activity of mixtures of the wild-type DnaA and DnaA H136A or DnaA F46A proteins were analyzed using Form I* assay. In these experiments, complexes including both the wild-type DnaA and H136A (or F46A) proteins should be constructed on the same *oriC* molecule. Thus, if a subgroup of DnaA protomers is not used for DnaB loading, the partial inclusion of DnaA H136A (or F46A) which is fully active in DUE unwinding, might retain DUE unwinding and DnaB loading. DnaA F46A, although inactive for DnaB binding by itself, can help achieve optimal Form I* formation when wild-type DnaA is provided suboptimally (**Figure 5A**), as previously demonstrated (Keyamura et al., 2009), suggesting that DnaA

FIGURE 3 | Activities of wild-type (WT) and H136A mutant DnaA in replication initiation and *oriC* unwinding. **(A)** *In vitro oriC* replication assay. Nucleotide-bound WT and H136A DnaA were incubated at 30°C for 20 min with a replicative WM433 (*dnaA204*) crude extract, M13KEW101 *oriC* plasmid (600 pmol nucleotides), and dNTPs containing [$\alpha - ^{32}$P]dATP, as described previously (Nishida et al., 2002). Similar results were obtained in an independent experiment. **(B)** P1-nuclease assay. Nucleotide-bound forms of DnaA were incubated with M13KEW101 *oriC* plasmid and HU protein, followed by incubation with P1 nuclease and then with *Bse*RI. The *Bse*RI restriction site is located distally to *oriC*, so that two DNA fragments (3.5 and 4.4 kb) are produced if *oriC* is unwound and digested by P1 nuclease. Digested samples were analyzed by 1% agarose-gel electrophoresis. *oriC* unwinding (3.5 and 4.4 kb fragments) was quantified. **(C)** Nucleotide-bound forms of DnaA and an *oriC* DOR(ΔDUE) fragment (5 nM) were incubated on ice for 5 min, then for 5 min at 30°C with ^{32}P-labeled 28-mer ssDUE (2.5 nM). Formation of protein–DNA complexes was analyzed by 4% polyacrylamide-gel electrophoresis, and quantified.

F46A can contribute to DnaB loading when it is a part of a complex including the wild-type DnaA at *oriC*. This means that only a subgroup (but not all) of DnaA molecules assembled on *oriC* would require Phe46 for DnaB loading (Keyamura et al., 2009). Here, we obtained similar results with DnaA H136A in the presence of wild-type DnaA (**Figure 5A**).

FIGURE 4 | DnaA-directed loading of DnaB at *oriC*. Representative gel images are shown in black–white inverted mode, and migration positions of negatively supercoiled (Form I) and highly negatively supercoiled (Form I*) plasmid DNA are indicated. Form I* is produced from Form I by DnaB helicase and DNA gyrase activities. Relative amounts of Form I* to input DNA were quantified as "Form I* (%)," and mean values with SD (n = 2) are shown in each graph. WT, wild-type DnaA; F46A, DnaA F46A; H136A, DnaA H136A; None, no DnaA. **(A)** Form I* assay with WT, F46A, and H136A DnaA. Indicated amounts of ATP-DnaA were incubated for 15 min at 30°C in buffer containing DnaB–DnaC (100 nM), SSB, gyrase, IHF, and pBSoriC *oriC* plasmid Form I. The resultant DNA forms were analyzed by agarose-gel electrophoresis. **(B)** Form I* assay with fixed concentration (20 nM) DnaA and various concentrations DnaB–DnaC complexes. Other conditions were as in **(A)**. **(C)** Form I* assay, as in **(A)**, but with pBSoriCΔR4-R2 plasmid.

DnaA H136A and DnaA F46A proteins were mixed at various ratios and assessed by Form I* assay, but did not produce substantial amounts of Form I* (**Figure 5B**). Only weak Form I* formation due to the residual activity of DnaA H136A was observed. These results suggest that, at least in a subgroup of DnaA molecules assembled on *oriC*, both the His136 and Phe46 residues must be present in the same DnaA protomer for DnaB loading.

DnaA H136A Forms Stable *oriC*–DnaA Complexes, but Results in Impaired Loading of DnaB on Unwound *oriC*

DnaA R281A is impaired in stable DnaA assembly on *oriC*, resulting in largely reduced binding of DnaB compared with wild-type DnaA, although DUE unwinding activity is sustained (Felczak and Kaguni, 2004). Arg281 is a constituent of the AAA+ Box VII motif, which is thought to reside at the interface of DnaA oligomers, supporting stable DnaA–DnaA interactions. To assess the activities of DnaA H136A in stable construction of DnaA assembly on *oriC*, we performed a pull-down assay using biotin-tagged *oriC* fragments. DnaA H136A was recovered by *oriC* fragment pull-down at a similar level to wild-type DnaA (**Figure 6A**), indicating that DnaA H136A is competent for DnaA assembly, which is consistent with its activities in DUE unwinding and ssDUE binding. In the presence of DnaA H136A and DnaB, *oriC* pull-down of DnaB was similar to that in the presence of wild-type DnaA and DnaB, indicating stable binding of DnaB helicase by DnaA H136A (**Figure 6B**). Also, even when DnaC was coincubated, DnaA H136A stably bound DnaB at a level similar to wild-type DnaA (**Figure 6C**). As previously shown (Keyamura et al., 2007), recovery of DnaB was increased by the co-incubation of DnaC, suggesting a conformational change of DnaB by binding of DnaC. Recovery of DnaC was moderately less than that of DnaB, which might be caused by moderate dissociation of DnaC during the wash step in this pull-down experiment.

Loading of DnaB was further examined by a novel pull-down assay using His-tagged DnaB, DnaC, DnaA, IHF, and *oriC* plasmid. In this assay, *oriC* is unwound by DnaA complexes, DnaB undergoes DnaC-dependent loading on the ssDUE region, and the resultant DnaB-bound *oriC* plasmids are recovered with Co²⁺-conjugated beads. The use of wild-type DnaA and *oriC* in this assay resulted in a high level of recovery of *oriC* plasmid, which was dependent on the inclusion of DnaB and DnaC (**Figure 6D**), suggesting that stable complexes with DnaB hexameric rings (or spirals) encircling the ssDNA of *oriC* are necessary for *oriC* recovery. A low level of recovery of *oriC* plasmid in the absence of DnaC presumably represented direct binding of DnaB to DnaA oligomers on *oriC* involving DnaA Phe46. These complexes were likely to have been largely dissociated by the use of wash buffer. In agreement with this idea, the use of wild-type DnaA and mutant *oriC* with DUE deletion (ΔDUE), which is inactive for unwinding, only resulted in a low level of recovery of *oriC* plasmid. Moreover, the use of DnaA F46A resulted in no observable recovery of *oriC* plasmid (**Figure 6D**). These
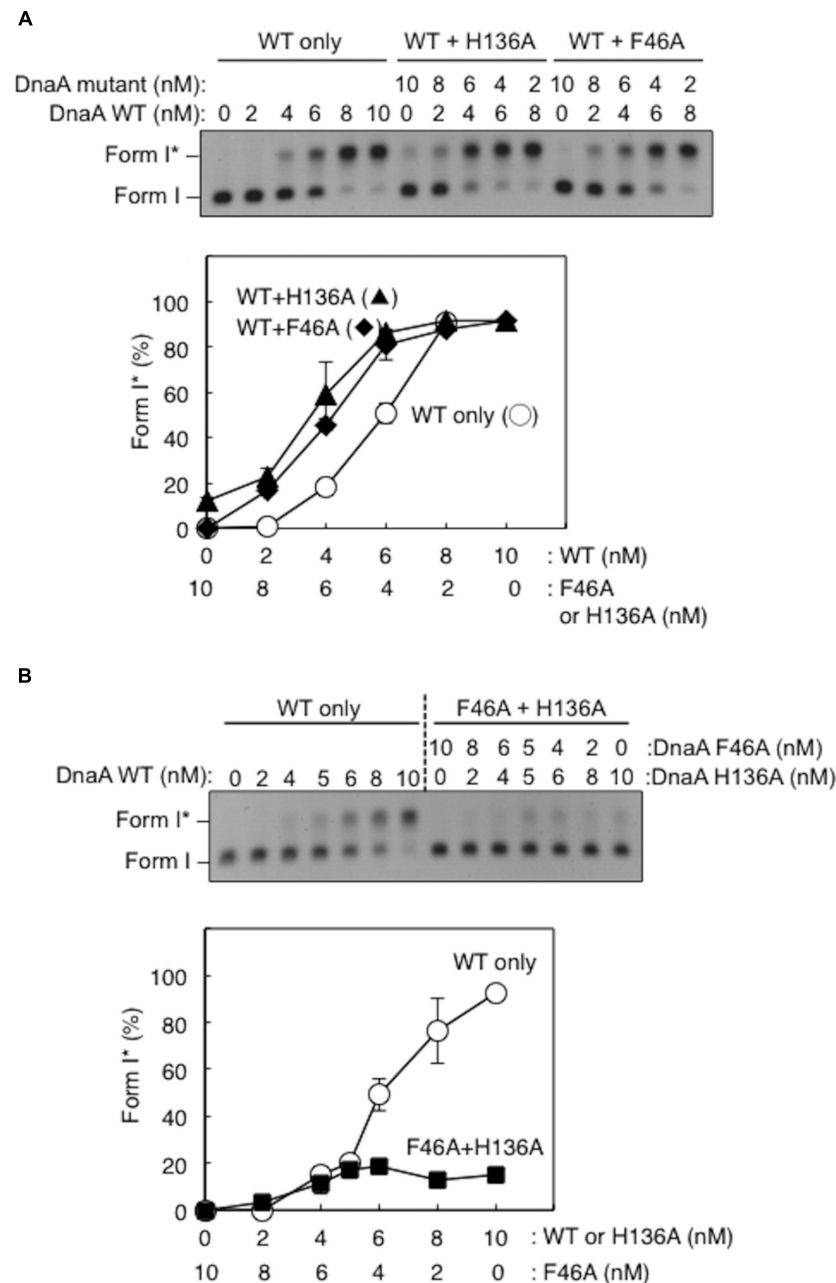
FIGURE 5 | Activities of DnaB loading by mixtures of wild-type (WT) and mutant DnaA. Form I* assays were performed with different combinations of WT, F46A, and H136A DnaA, with pBSoriC *oriC* plasmid. Form I* is highly negatively supercoiled, and is produced from Form I by DnaB helicase and DNA gyrase activities. Reactions were incubated for 15 min at 30°C. Representative gel images are shown in black–white inverted mode. Amounts of Form I* relative to the input DNA were quantified as "Form I* (%)," and mean values with SD (*n* = 2) are shown. **(A)** Mixtures of WT DnaA and DnaA F46A or H136A were used. **(B)** Mixtures of DnaA F46A and DnaA H136A were used.

results were consistent with DnaA Phe46-dependent binding between DnaB and DnaA oligomers constructed on *oriC* being responsible for the basal recovery level, with considerable enhancement of recovery resulting from DnaB loading on ssDNA.

The inclusion of DnaA H136A and wild-type *oriC* in this assay in the presence of DnaB and DnaC resulted in moderate inhibition in *oriC* recovery, compared to the level seen with wild-type DnaA (**Figure 6D**). A low level of recovery (similar to that with wild-type DnaA) was seen with DnaA H136A in the absence of DnaC or in the presence of ΔDUE. These results further support the idea that DnaA H136 residue is specifically important for functional DnaB loading to the unwound site of *oriC* (see also section "Discussion").
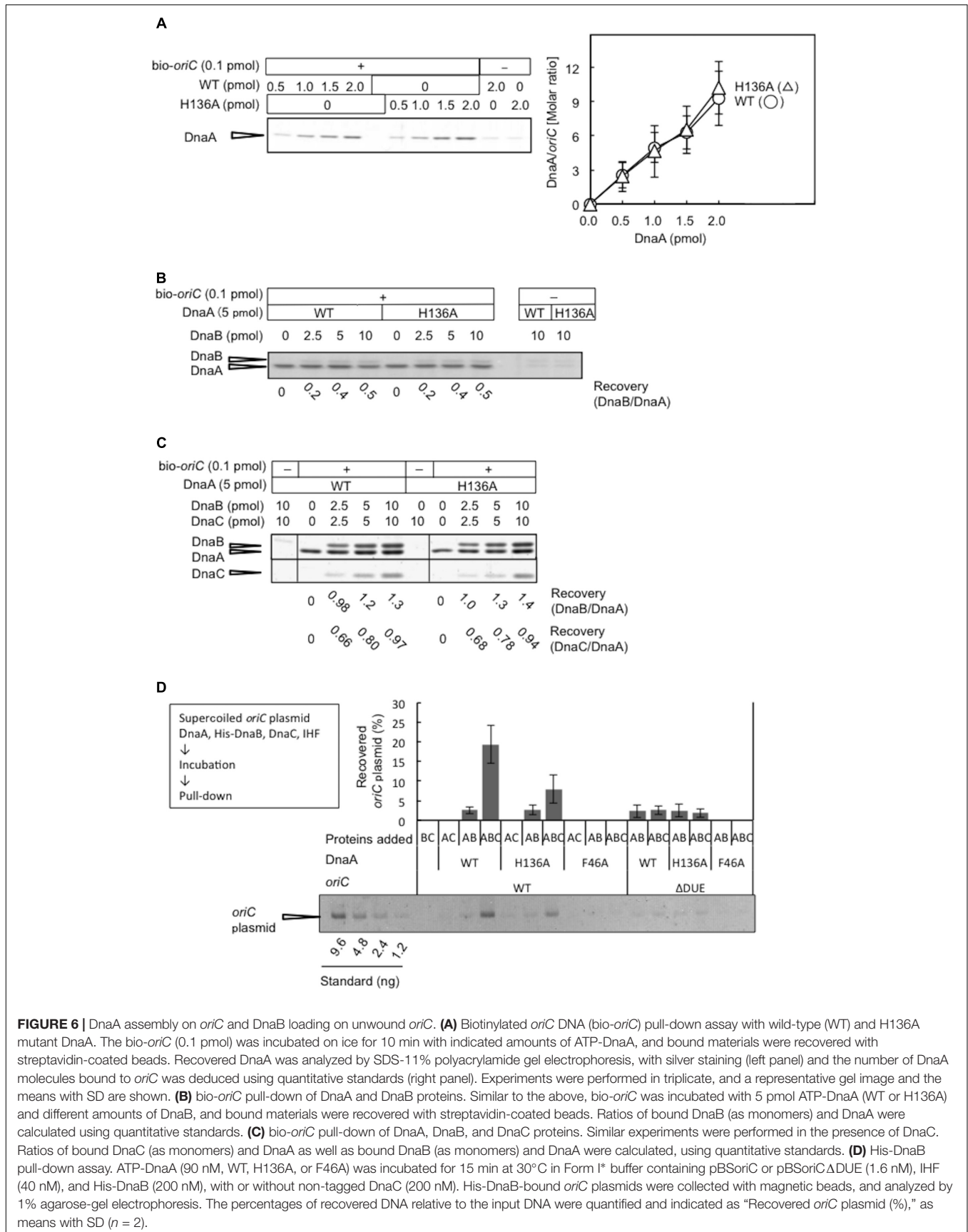
FIGURE 6 | DnaA assembly on *oriC* and DnaB loading on unwound *oriC*. **(A)** Biotinylated *oriC* DNA (bio-*oriC*) pull-down assay with wild-type (WT) and H136A mutant DnaA. The bio-*oriC* (0.1 pmol) was incubated on ice for 10 min with indicated amounts of ATP-DnaA, and bound materials were recovered with streptavidin-coated beads. Recovered DnaA was analyzed by SDS-11% polyacrylamide gel electrophoresis, with silver staining (left panel) and the number of DnaA molecules bound to *oriC* was deduced using quantitative standards (right panel). Experiments were performed in triplicate, and a representative gel image and the means with SD are shown. **(B)** bio-*oriC* pull-down of DnaA and DnaB proteins. Similar to the above, bio-*oriC* was incubated with 5 pmol ATP-DnaA (WT or H136A) and different amounts of DnaB, and bound materials were recovered with streptavidin-coated beads. Ratios of bound DnaB (as monomers) and DnaA were calculated using quantitative standards. **(C)** bio-*oriC* pull-down of DnaA, DnaB, and DnaC proteins. Similar experiments were performed in the presence of DnaC. Ratios of bound DnaC (as monomers) and DnaA as well as bound DnaB (as monomers) and DnaA were calculated, using quantitative standards. **(D)** His-DnaB pull-down assay. ATP-DnaA (90 nM, WT, H136A, or F46A) was incubated for 15 min at 30°C in Form I* buffer containing pBSoriC or pBSoriCΔDUE (1.6 nM), IHF (40 nM), and His-DnaB (200 nM), with or without non-tagged DnaC (200 nM). His-DnaB-bound *oriC* plasmids were collected with magnetic beads, and analyzed by 1% agarose-gel electrophoresis. The percentages of recovered DNA relative to the input DNA were quantified and indicated as "Recovered *oriC* plasmid (%)," as means with SD ($n = 2$).

FIGURE 7 | ABC primosome assay. The indicated concentrations of wild-type (WT) or F46A or H136A mutant DnaA were incubated for 15 min at 30°C in buffer containing DnaB, DnaC, DnaG, SSB, NTPs, dNTPs, and DNA polymerase III holoenzyme, with or without M13 A-site ssDNA (220 pmol nucleotides). DNA replication was quantified by measurement of incorporation of radiolabeled dATP. Similar results were obtained when reactions were incubated at 37 or 42°C.

FIGURE 8 | Bacterial DnaA orthologs possess a conserved aromatic residue at the position corresponding to DnaA His136. Sequence logos of DnaA His136 and its surrounding residues are shown based on the sequence alignment of 581 DnaA homologs from all bacterial phyla (top) and 162 DnaA homologs from γ proteobacteria (bottom). The protein sequences were obtained from the NCBI database and the logos were created using WebLogo (http://weblogo.threeplusone.com).

DnaA H136A Is Active in DnaB Loading in a Simplified System

DnaA–DnaB interactions were also studied using the simplified reconstituted ABC primosome system, which uses M13-phage-derived ssDNA with a hairpin structure containing the DnaA box R1 sequence (A-site ssDNA) (Masai et al., 1990; Masai and Arai, 1995; Carr and Kaguni, 2002; **Figure 7**). DnaA binding to the hairpin structure enables recruitment of DnaBC complex and DnaB loading to the SSB-coated ssDNA, followed by primer and DNA synthesis by DnaG primase and DNA polymerase III holoenzyme. Whereas DnaA F46A was essentially inactive in this system, DnaA H136A was fully active relative to wild-type DnaA (**Figure 7**), indicating that DnaA H136A promoted DnaB loading to ssDNA in this simple system. Notably, unlike the more complex *oriC*, only 2–4 DnaA molecules bind to this M13 hairpin region, with its single DnaA box (Carr and Kaguni, 2002), and except for the hairpin region, the whole template is single-stranded. These specific features may cause the different requirement for His136 between *oriC* and A-site ssDNA (see section "Discussion").

Evolutionary Conservation of DnaA His136

Among bacterial DnaA orthologs, the position corresponding to *E. coli* DnaA His136 is generally occupied by an aromatic residue, such as tyrosine, phenylalanine, or histidine (**Figure 8**). Even in the hyperthermophile *A. aeolicus* DnaA ortholog, the corresponding residue is tyrosine (Erzberger et al., 2002). In γ proteobacteria including *E. coli*, the histidine residue predominates (**Figure 8**). Thus, the aromatic moiety at the position of *E. coli* DnaA His136 appears to have an evolutionarily conserved role in the loading of DnaB helicases at *oriC*.

DISCUSSION

Loading of replicative helicases on unwound origin DNA region is a crucial step in replication initiation of chromosomes. In *E. coli*, this step depends on dynamic interactions between the initiator DnaA and DnaB replicative helicase. Stable binding
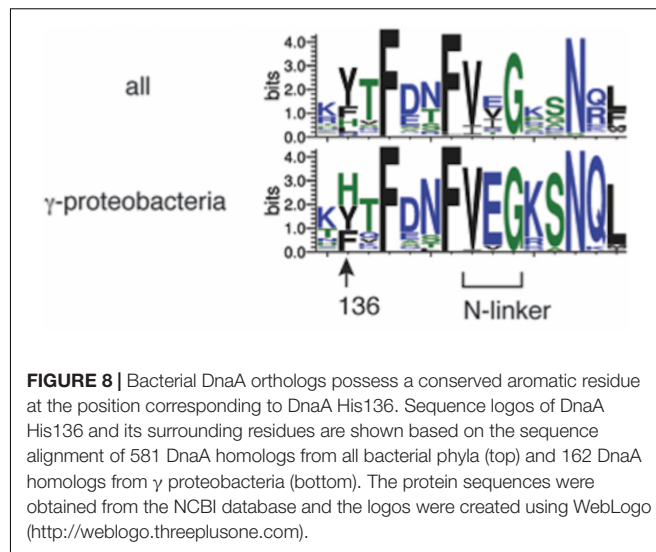
between the two depends on a site including DnaA Glu21 and Phe46 of domain I. This site is suggested to interact with a specific site of DnaB, which resides in a region including the N-terminus and its franking region of DnaB CTD. In addition, the DnaA region Lys135–Gln148 within domain III has been implicated as having weak physical contact with DnaB NTD. However, the biological importance of this interaction has not previously been determined. Here, we used alanine scanning of these DnaA residues to highlight the role of His136 in replication initiation *in vivo*. The plasmid complementation test and flow cytometry analysis demonstrated that the *dnaA H136A* allele has only low *in vivo* initiation activity. In-depth biochemical characterization of DnaA H136A revealed that His136 is indispensable for DnaB loading at *oriC*, but that DUE unwinding, ssDUE binding, and assembly on *oriC* do not depend on this residue. Intriguingly, DnaA H136A was fully active in an *oriC*-independent DnaB loading system for ssDNA replication (i.e., ABC primosome system). These observations suggest that DnaA His136 residue has a distinct role in DnaB loading at *oriC*. Because DnaA H136A-*oriC* complexes maintain the primary contact between DnaA Phe46 and DnaB, we conclude that a secondary, weak contact via His136 might allow DnaA-directed DnaB loading on the ssDUE region. The conservation of His136 suggests that its role is conserved in eubacterial DnaA orthologs.

Loading of DnaB at *oriC* relies on stable formation of DnaA oligomers on *oriC*. DnaA Arg281 residue indirectly facilitates DnaB binding and loading by contributing to stable binding of DnaA molecules to *oriC* (Felczak and Kaguni, 2004). By contrast, the results of the pull-down assay using an *oriC* fragment indicated that the numbers of DnaA and DnaB molecules bound to *oriC* were similar irrespective of whether wild-type DnaA or DnaA H136A was used. Thus, unlike DnaA K281A, DnaA H136A constructs stable complexes at *oriC* which are fully competent in stable DnaB binding. In our highly detailed structural model of DnaA (Shimizu et al., 2016), the location of the His136 residue suggests that it is unlikely to be involved in interactions between DnaA protomers, unlike the

FIGURE 9 | The position of DnaA H136 residue in the left-half DnaA subcomplex. On the basis of a molecular-dynamics-simulation near-atomic model of the *Escherichia coli* initiation complex (Shimizu et al., 2016), the left-half *oriC* complex and important functional residues are shown. In this model, ATP-DnaA domain III–IV molecules are shown complexed with five DnaA boxes (R1, R5M, τ2, I1, and I2) in the left-half DnaA-oligomerization region (DOR). Also, a DNA-bending protein integration host factor (IHF) binds to a unique IHF-binding site. The left-half *oriC* DNA is shown as silver lines. DnaA domain III and IV molecules are shown in violet and light blue. His136 residues are colored pink, and Arg281 residues are colored yellow. ssDUE-binding H/B-motif residues Val211 and Arg245 are shown in orange and dark blue, respectively. IHF is shown in dark green. Two sides of the complexes are shown (upper and lower panels).

Arg281 residue, which resides at the inter-protomer interface (**Figure 9**).

Our previous *in vitro* and *in vivo* studies as well as the present data are in support of the ssDUE-recruitment mechanism in which DnaA oligomers constructed on DOR stably bind ssDUE using DnaA domain III H/B-motifs (**Figure 1B**; Ozaki and Katayama, 2012; Sakiyama et al., 2017). Recent structural analysis also supports this mechanism (Shimizu et al., 2016). In the paper of Duderstadt et al. (2010), DnaA oligomer formation was moderately stimulated by 25-mer ssDNA and largely decreased by the addition of 13-mer dsDNA bearing a single R1 box, resulting in an increase of DnaA monomers. It should be noted in these experiments that glutaraldehyde cross-linking was used because of the instability of ssDNA binding to DnaA and that this cross-linking produced considerable amounts of DnaA oligomers (dimers to tetramers) even in the absence of ssDNA, causing high

background levels. Those results mean that addition of 13-mer dsDNA bearing the R1 box inhibits oligomerization of DnaA and the resultant R1-bound DnaA monomers (but not oligomers) do not bind stably to the ssDNA (and do not mean that dsDNA binding and ssDNA binding of DnaA mutually exclusive). This is consistent with our previous data. We demonstrated that unlike DnaA oligomers constructed on DOR fragments bearing multiple DnaA-binding sites, DnaA bound to a short DNA bearing only a single R1 box is inactive for stable ssDUE binding (Ozaki and Katayama, 2012). In addition, we demonstrated that two T-rich regions in the ssDUE are essential for the binding of ssDUE to DnaA oligomers constructed on the DOR (Ozaki et al., 2008). Consistently, we recently showed that two DnaA molecules bound to the R1 and R5M boxes are crucial for ssDUE binding (Sakiyama et al., 2017). As previously explained (Katayama et al., 2017; Sakiyama et al., 2017), these results concordantly support

our idea that when DnaA molecules form an oligomer on the left-half DOR, stable ssDUE binding occur as a result of the linkage effect that enhances cooperative binding (Stauffer and Chazin, 2004; **Figure 1B**).

Our present results highlight a functional relationship between DnaA His136 and Phe46 (**Figure 5**). We previously reported that the DnaA domain I Phe46 exhibits high-affinity binding to DnaB when DnaA oligomers are constructed on *oriC* (Keyamura et al., 2009). In addition, although DnaA F46A alone has little or no activity for DnaB loading, DnaA F46A contributes to DnaB loading in the presence of wild-type DnaA. Similarly, we have now found that DnaA H136A alone has little or no activity for DnaB loading, but it participates in helicase loading at *oriC* when mixed with wild-type DnaA. This result suggests that His136 and Phe46 are only required in a subset of protomers in DnaA oligomers constructed on *oriC*. Further studies are required to determine which protomers in the DnaA subcomplexes might functionally interact with DnaB during the loading processes.
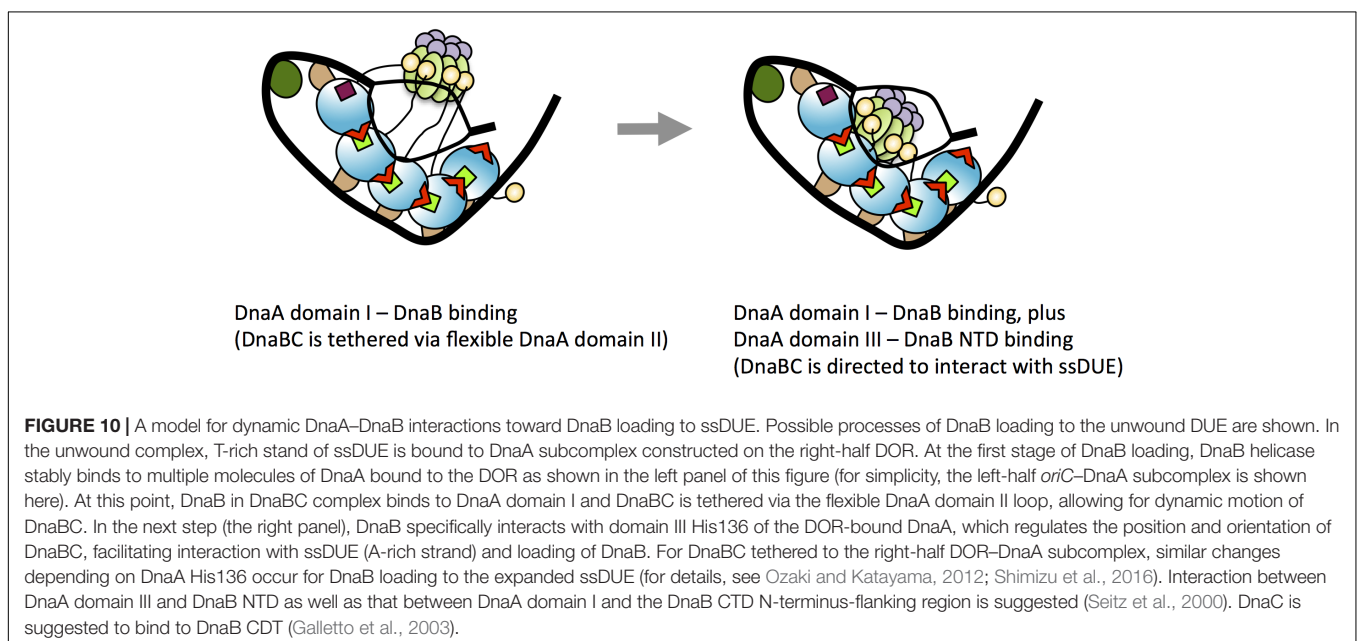
In the *oriC* plasmid-pull down assay using His-DnaB (**Figure 6D**), inhibition by DnaA H136A is moderate compared to the severe inhibition of Form I* production (**Figure 4A**). Given that DnaA Phe46-dependent DnaB interaction is intact even in DnaA H136A, a possible explanation to this is that interaction of DnaB with DnaA H136A at *oriC* results in abortive loading complexes of DnaB that are not competent for helicase activity. Alternatively, loading orientation of DnaB helicases on ssDUE or DUE strand-specific loading of DnaB might be compromised by the DnaA H136A mutation, resulting in abortive complexes.

Two distinct DnaB-binding modes of DnaA seem to be required for the loading of DnaB onto ssDUE. Notably, heterologous complexes formed at *oriC* by a combination of DnaA domain I F46A and domain III H136A are substantially inactive in DnaB loading (**Figure 5B**), suggesting that both

Phe46 and His136 must be present in identical DnaA protomers for helicase loading at *oriC*. Although it has previously been suggested that a weak contact between DnaB and a DnaA region consisting of residues 111–148 (which includes His136) may precede stable binding involving DnaA domain I (Sutton et al., 1998), our results demonstrated that DnaA H136A fully sustained DnaA domain I-dependent DnaB binding. In a previous paper (Sutton et al., 1998) in which Surface Plasmon Resonance analysis was employed, the immobilization of DnaB on the sensor tip may have spatially occluded the DnaA domain I-binding site of DnaB, preventing DnaB interaction with DnaA. Taken together, a likely process is that DnaB binds to DnaA–*oriC* complexes via a primary contact mediated by DnaA domain I, followed by a secondary, weak contact mediated by DnaA His136. Binding of DnaB to DnaA domain I might bring DnaB into proximity with DnaA domain III, enabling the secondary weak contact through DnaA His136 of the same protomer (**Figure 1B**).

We now suggest a model in which the secondary, weak contact via His136 facilitates accessibility of DnaB to the ssDUE (**Figure 10**). Because the T-rich strand of the ssDUE directly binds to the H/B motifs of DnaA domain III (Ozaki et al., 2008; Duderstadt et al., 2011), DnaB bound to DnaA domain I can be brought in close proximity to the ssDUE through the physical contact with DnaA His136 and structural change of the flexible linker domain II (**Figure 10**). Notably, in our high-definition structural model of DnaA–*oriC* complexes, DnaA His136 residues are exposed on the surface of the DnaA protomers, suggesting that they are accessible to DnaB without physically obstructing ssDUE binding to the H/B-motifs (Shimizu et al., 2016; **Figure 9**).

A role for DnaA His136 in regulation of loading of DnaB helicases to the ssDUE might explain why this residue is dispensable for DnaB loading in the ABC primosome system. The ssDNA template in the ABC primosome system most

DnaA domain I – DnaB binding
(DnaBC is tethered via flexible DnaA domain II)

DnaA domain I – DnaB binding, plus
DnaA domain III – DnaB NTD binding
(DnaBC is directed to interact with ssDUE)

FIGURE 10 | A model for dynamic DnaA–DnaB interactions toward DnaB loading to ssDUE. Possible processes of DnaB loading to the unwound DUE are shown. In the unwound complex, T-rich stand of ssDUE is bound to DnaA subcomplex constructed on the right-half DOR. At the first stage of DnaB loading, DnaB helicase stably binds to multiple molecules of DnaA bound to the DOR as shown in the left panel of this figure (for simplicity, the left-half *oriC*–DnaA subcomplex is shown here). At this point, DnaB in DnaBC complex binds to DnaA domain I and DnaBC is tethered via the flexible DnaA domain II loop, allowing for dynamic motion of DnaBC. In the next step (the right panel), DnaB specifically interacts with domain III His136 of the DOR-bound DnaA, which regulates the position and orientation of DnaBC, facilitating interaction with ssDUE (A-rich strand) and loading of DnaB. For DnaBC tethered to the right-half DOR–DnaA subcomplex, similar changes depending on DnaA His136 occur for DnaB loading to the expanded ssDUE (for details, see Ozaki and Katayama, 2012; Shimizu et al., 2016). Interaction between DnaA domain III and DnaB NTD as well as that between DnaA domain I and the DnaB CTD N-terminus-flanking region is suggested (Seitz et al., 2000). DnaC is suggested to bind to DnaB CDT (Galletto et al., 2003).

likely minimizes spatial constraints for DnaB loading, thereby bypassing the strict requirement for H136-mediated positioning of the helicase. In other words, ssDNA replication by the ABC primosome system does not require unwinding of dsDNA, regulated DnaA assembly like that on the left-half DOR–IHF complex or loading of two DnaB helicases in opposite directions. By contrast, at *oriC*, DnaB loading must occur on a short singled-stranded region of the DUE. At the M13-A site, the ssDNA region is much larger and thus DnaB loading is less spatially restricted; i.e., DnaB-loading may not be strictly regulated in the ABC promosome system; this may explain the differences in the roles of His136 at *oriC* and in the ABC promosome. Alternatively, SSB that coats the template ssDNA might have an auxiliary role in directing DnaB helicase for ssDNA loading in the ABC primosome system. Functional interaction between SSB and DnaB is reported to stimulate DnaB helicase activity (Biswas et al., 2002).

The linkage effect means that the presence of multiple contact points can result in a drastic increase in binding avidity (total affinity) even if each individual contact only has weak affinity (Stauffer and Chazin, 2004). A similar mechanism might underlie the DnaB interaction that depends on DnaA H136 residues in the *oriC* system. DnaA H136 residues are predicted to be regularly arrayed on the *oriC*-bound DnaA oligomers (**Figure 9**), which might enable formation of multiple contact points with a single DnaB homohexamer. By contrast, in the ABC primosome system, only a few DnaA protomers are involved in a region bearing only a single DnaA box (Masai et al., 1990; Carr and Kaguni, 2002), suggesting that the specific DnaA oligomer structure causing the His136-dependent linkage effect would not be constructed.

Whereas DnaA His136 was dispensable in the ABC primosome system, DnaA domain III is important. Specifically, two residues (Lys234 and Arg285) exposed on the protein surface that are important for DnaA–DnaA interaction stimulate ssDNA

replication in the ABC primosome system (Kawakami et al., 2005; Ozaki et al., 2008). Binding of a few DnaA molecules to the hairpin structure of the template ssDNA (Carr and Kaguni, 2002) might be stimulated by domain III-dependent DnaA–DnaA interactions.

Dynamic protein complexes are constructed and change structurally at *oriC* for duplex unwinding and helicase loading. This study reveals the essential function for DnaA domain III His136 in the loading of DnaB replicative helicases on the ssDUE. Further analyses are required to dissect the DnaA protomers responsible for the DnaB interaction during DnaB loading on ssDUE and the DnaA-interacting sites on DnaB as well as to uncover the dynamic mechanisms of DnaA–DnaB complexes underlying the loading orientation of DnaB on ssDUE and DUE strand-specific loading of DnaB.

AUTHOR CONTRIBUTIONS

YS, MN, SO, and TK conceived the experiments and wrote the paper. YS, MN, YA, and CH performed the experiments. All authors analyzed the data.

FUNDING

ACKNOWLEDGMENTS

REFERENCES

Abe, Y., Jo, T., Matsuda, Y., Matsunaga, C., Katayama, T., and Ueda, T. (2007). Structure and function of DnaA N-terminal domains: specific sites and mechanisms in inter-DnaA interaction and in DnaB helicase loading on oriC. *J. Biol. Chem.* 282, 17816–17827. doi: 10.1074/jbc.M701841200

Baker, T. A., Sekimizu, K., Funnell, B. E., and Kornberg, A. (1986). Extensive unwinding of the plasmid template during staged enzymatic initiation of DNA replication from the origin of the *Escherichia coli* chromosome. *Cell* 45, 53–64. doi: 10.1016/0092-8674(86)90537-4

Bell, S. P., and Kaguni, J. M. (2013). Helicase loading at chromosomal origins of replication. *Cold Spring Harb. Perspect. Biol.* 5:a010124. doi: 10.1101/cshperspect.a010124

Biswas, E. E., Chen, P. H., and Biswas, S. B. (2002). Modulation of enzymatic activities of *Escherichia coli* DnaB helicase by single-stranded DNA-binding proteins. *Nucleic Acids. Res.* 30, 2809–2816. doi: 10.1093/nar/gkf384

Biswas, S. B., and Biswas-Fiss, E. E. (2006). Quantitative analysis of binding of single-stranded DNA by *Escherichia coli* DnaB helicase and the DnaB-DnaC complex. *Biochemistry* 45, 11505–11513. doi: 10.1021/bi060118d

Bramhill, D., and Kornberg, A. (1988). A model for initiation at origins of DNA replication. *Cell* 54, 915–918. doi: 10.1016/0092-8674(88)90102-X

Carr, K. M., and Kaguni, J. M. (2001). Stoichiometry of DnaA and DnaB protein in initiation at the *Escherichia coli* chromosomal origin. *J. Biol. Chem.* 276, 44919–44925. doi: 10.1074/jbc.M107463200

Carr, K. M., and Kaguni, J. M. (2002). *Escherichia coli* DnaA protein loads a single DnaB helicase at a DnaA box hairpin. *J. Biol. Chem.* 277, 39815–39822. doi: 10.1074/jbc.M205031200

Davey, M. J., and O'Donnell, M. (2003). Replicative helicase loaders: ring dispatch breakers and ring makers. *Curr. Biol.* 13, R594–R596. doi: 10.1016/S0960-9822(03)00523-2

Duderstadt, K. E., Chuang, K., and Berger, J. M. (2011). DNA stretching by bacterial initiators promotes replication origin opening. *Nature* 478, 209–213. doi: 10.1038/nature10455

Duderstadt, K. E., Mott, M. L., Crisona, N. J., Chuang, K., Yang, H., and Berger, J. M. (2010). Origin remodeling and opening in bacteria rely on distinct assembly states of the DnaA initiator. *J. Biol. Chem.* 285, 28229–28239. doi: 10.1074/jbc.M110.147975

Erzberger, J. P., Mott, M. L., and Berger, J. M. (2006). Structural basis for ATP-dependent DnaA assembly and replication-origin remodeling. *Nat. Struct. Mol. Biol.* 13, 676–683. doi: 10.1038/nsmb1115

Erzberger, J. P., Pirruccello, M. M., and Berger, J. M. (2002). The structure of bacterial DnaA: implications for general mechanisms underlying DNA replication initiation. *EMBO J.* 21, 4763–4773. doi: 10.1093/emboj/cdf496

Fang, L., Davey, M. J., and O'Donnell, M. (1999). Replisome assembly at *oriC*, the replication origin of *E. coli*, reveals an explanation for initiation sites outside an origin. *Mol. Cell* 4, 541–553. doi: 10.1016/S1097-2765(00)80205-1

Felczak, M. M., and Kaguni, J. M. (2004). The box VII motif of *Escherichia coli* DnaA protein is required for DnaA oligomerization at the *E. coli* replication origin. *J. Biol. Chem.* 279, 51156–51162. doi: 10.1074/jbc.M409695200

Fujikawa, N., Kurumizaka, H., Nureki, O., Terada, T., Shirouzu, M., Katayama, T., et al. (2003). Structural basis of replication origin recognition by the DnaA protein. *Nucleic Acids Res.* 31, 2077–2086. doi: 10.1093/nar/gkg309

Fujimitsu, K., Senriuchi, T., and Katayama, T. (2009). Specific genomic sequences of *E. coli* promote replicational initiation by directly reactivating ADP-DnaA. *Genes Dev.* 23, 1221–1233. doi: 10.1101/gad.1775809

Fuller, R. S., Kaguni, J. M., and Kornberg, A. (1981). Enzymatic replication of the origin of the *Escherichia coli* chromosome. *Proc. Natl. Acad. Sci. U.S.A.* 78, 7370–7374. doi: 10.1073/pnas.78.12.7370

Galletto, R., Jezewska, M. J., and Bujalowski, W. (2003). Interactions of the *Escherichia coli* DnaB helicase hexamer with the replication factor the DnaC protein: effect of nucleotide cofactors and the ssDNA on protein–protein interactions and the topology of the complex. *J. Mol. Biol.* 329, 441–465. doi: 10.1016/S0022-2836(03)00435-2

Grimwade, J. E., Rozgaja, T. A., Gupta, R., Dyson, K., Rao, P., and Leonard, A. C. (2018). Origin recognition is the predominant role for DnaA-ATP in initiation of chromosome replication. *Nucleic Acids Res.* 46, 6140–6151. doi: 10.1093/nar/gky457

Grimwade, J. E., Torgue, J. J.-C., McGarry, K. C., Rozgaja, T., Enloe, S. T., and Leonard, A. C. (2007). Mutational analysis reveals *Escherichia coli oriC* interacts with both DnaA-ATP and DnaA-ADP during pre-RC assembly. *Mol. Microbiol.* 66, 428–439. doi: 10.1111/j.1365-2958.2007.05930.x

Inoue, Y., Tanaka, H., Kasho, K., Fujimitsu, K., Oshima, T., and Katayama, T. (2016). Chromosomal location of the DnaA-reactivating sequence *DARS2* is important to regulate timely initiation of DNA replication in *Escherichia coli*. *Genes Cells* 21, 1015–1023. doi: 10.1111/gtc.12395

Itsathitphaisarn, O., Wing, R. A., Eliason, W. K., Wang, J., and Steitz, T. A. (2012). The hexameric helicase DnaB adopts a nonplanar conformation during translocation. *Cell* 151, 267–277. doi: 10.1016/j.cell.2012.09.014

Iyer, L. M., Leipe, D. D., Koonin, E. V., and Aravind, L. (2004). Evolutionary history and higher order classification of AAA+ ATPases. *J. Struct. Biol.* 146, 11–31. doi: 10.1016/j.jsb.2003.10.010

Kaguni, J. M. (2011). Replication initiation at the *Escherichia coli* chromosomal origin. *Curr. Opin. Chem. Biol.* 15, 606–613. doi: 10.1016/j.cbpa.2011.07.016

Katayama, T., Kasho, K., and Kawakami, H. (2017). The DnaA cycle in *Escherichia coli*: activation, function and inactivation of the initiator protein. *Front. Microbiol.* 8:2496. doi: 10.3389/fmicb.2017.02496

Katayama, T., Ozaki, S., Keyamura, K., and Fujimitsu, K. (2010). Regulation of the replication cycle: conserved and diverse regulatory systems for DnaA and *oriC*. *Nat. Rev. Microbiol.* 8, 163–170. doi: 10.1038/nrmicro2314

Kawakami, H., Keyamura, K., and Katayama, T. (2005). Formation of an ATP-DnaA-specific initiation complex requires DnaA Arginine 285, a conserved motif in the AAA+ protein family. *J. Biol. Chem.* 280, 27420–27430. doi: 10.1074/jbc.M502764200

Keyamura, K., Abe, Y., Higashi, M., Ueda, T., and Katayama, T. (2009). DiaA dynamics are coupled with changes in initial origin complexes leading to helicase loading. *J. Biol. Chem.* 284, 25038–25050. doi: 10.1074/jbc.M109.002717

Keyamura, K., Fujikawa, N., Ishida, T., Ozaki, S., Su'etsugu, M., Fujimitsu, K., et al. (2007). The interaction of DiaA and DnaA regulates the replication cycle in *E. coli* by directly promoting ATP DnaA-specific initiation complexes. *Genes Dev.* 21, 2083–2099. doi: 10.1101/gad.1561207

Kubota, T., Katayama, T., Ito, Y., Mizushima, T., and Sekimizu, K. (1997). Conformational transition of DnaA protein by ATP: structural analysis of DnaA protein, the initiator of *Escherichia coli* chromosome replication. *Biochem. Biophys. Res. Commun.* 232, 130–135. doi: 10.1006/bbrc.1997.6244

Kurokawa, K., Nishida, S., Emoto, A., Sekimizu, K., and Katayama, T. (1999). Replication cycle-coordinated change of the adenine nucleotide-bound forms of DnaA protein in *Escherichia coli*. *EMBO J.* 18, 6642–6652. doi: 10.1093/emboj/18.23.6642

Leonard, A. C., and Grimwade, J. E. (2015). The orisome: structure and function. *Front. Microbiol.* 6:545. doi: 10.3389/fmicb.2015.00545

Makowska-Grzyska, M., and Kaguni, J. M. (2010). Primase directs the release of DnaC from DnaB. *Mol. Cell* 37, 90–101. doi: 10.1016/j.molcel.2009.12.031

Marszalek, J., and Kaguni, J. M. (1994). DnaA protein directs the binding of DnaB protein in initiation of DNA replication in *Escherichia coli*. *J. Biol. Chem.* 269, 4883–4890.

Marszalek, J., Zhang, W., Hupp, T. R., Margulies, C., Carr, K. M., Cherry, S., et al. (1996). Domains of DnaA protein involved in interaction with DnaB protein, and in unwinding the *Escherichia coli* chromosomal origin. *J. Biol. Chem.* 271, 18535–18542. doi: 10.1074/jbc.271.31.18535

Masai, H., and Arai, K. (1995). DnaA-dependent assembly of the ABC primosome at the A site, a single-stranded DNA hairpin containing a dnaA box. *Eur. J. Biochem.* 230, 384–395. doi: 10.1111/j.1432-1033.1995.tb20573.x

Masai, H., Nomura, N., and Arai, K. (1990). The ABC-primosome. A novel priming system employing dnaA, dnaB, dnaC, and primase on a hairpin containing a dnaA box sequence. *J. Biol. Chem.* 265, 15134–15144.

McGarry, K. C., Ryan, V. T., Grimwade, J. E., and Leonard, A. C. (2004). Two discriminatory binding sites in the *Escherichia coli* replication origin are required for DNA strand opening by initiator DnaA-ATP. *Proc. Natl. Acad. Sci. U.S.A.* 101, 2811–2816. doi: 10.1073/pnas.0400340101

Mott, M. L., Erzberger, J. P., Coons, M. M., and Berger, J. M. (2008). Structural synergy and molecular crosstalk between bacterial helicase loaders and replication initiators. *Cell* 135, 623–634. doi: 10.1016/j.cell.2008.09.058

Nishida, S., Fujimitsu, K., Sekimizu, K., Ohmura, T., Ueda, T., and Katayama, T. (2002). A nucleotide switch in the *Escherichia coli* DnaA protein initiates chromosomal replication: evidence from a mutant DnaA protein defective in regulatory ATP hydrolysis in vitro and in vivo. *J. Biol. Chem.* 277, 14986–14995. doi: 10.1074/jbc.M108303200

Noguchi, Y., and Katayama, T. (2016). The *Escherichia coli* cryptic prophage protein YfdR binds to DnaA and initiation of chromosomal replication is inhibited by overexpression of the gene cluster *yfdQ-yfdR-yfdS-yfdT*. *Front. Microbiol.* 7:239. doi: 10.3389/fmicb.2016.00239

Noguchi, Y., Sakiyama, Y., Kawakami, H., and Katayama, T. (2015). The Arg fingers of key DnaA protomers are oriented inward within the replication origin *oriC* and stimulate DnaA subcomplexes in the initiation complex. *J. Biol. Chem.* 290, 20295–20312. doi: 10.1074/jbc.M115.662601

Nozaki, S., and Ogawa, T. (2008). Determination of the minimum domain II size of *Escherichia coli* DnaA protein essential for cell viability. *Microbiology* 154, 3379–3384. doi: 10.1099/mic.0.2008/019745-0

Ozaki, S., and Katayama, T. (2009). DnaA structure, function, and dynamics in the initiation at the chromosomal origin. *Plasmid* 62, 71–82. doi: 10.1016/j.plasmid.2009.06.003

Ozaki, S., and Katayama, T. (2012). Highly organized DnaA-oriC complexes recruit the single-stranded DNA for replication initiation. *Nucleic Acids Res.* 40, 1648–1665. doi: 10.1093/nar/gkr832

Ozaki, S., Kawakami, H., Nakamura, K., Fujikawa, N., Kagawa, W., Park, S.-Y., et al. (2008). A common mechanism for the ATP-DnaA-dependent formation of open complexes at the replication origin. *J. Biol. Chem.* 283, 8351–8362. doi: 10.1074/jbc.M708684200

Ozaki, S., Noguchi, Y., Hayashi, Y., Miyazaki, E., and Katayama, T. (2012a). Differentiation of the DnaA-oriC subcomplex for DNA unwinding in a replication initiation complex. *J. Biol. Chem.* 287, 37458–37471. doi: 10.1074/jbc.M112.372052

Ozaki, S., Noguchi, Y., Nishimura, M., and Katayama, T. (2012b). Stable nucleotide binding to DnaA requires a specific glutamic acid residue within the AAA+ box II motif. *J. Struct. Biol.* 179, 242–250. doi: 10.1016/j.jsb.2012.05.001

Rozgaja, T. A., Grimwade, J. E., Iqbal, M., Czerwonka, C., Vora, M., and Leonard, A. C. (2011). Two oppositely oriented arrays of low-affinity recognition sites in *oriC* guide progressive binding of DnaA during *Escherichia coli* pre-RC assembly. *Mol. Microbiol.* 82, 475–488. doi: 10.1111/j.1365-2958.2011.07827.x

Sakiyama, Y., Kasho, K., Noguchi, Y., Kawakami, H., and Katayama, T. (2017). Regulatory dynamics in the ternary DnaA complex for initiation of chromosomal replication in *Escherichia coli*. *Nucleic Acids Res.* 45, 12354–12373. doi: 10.1093/nar/gkx914

Seitz, H., Weigel, C., and Messer, W. (2000). The interaction domains of the DnaA and DnaB replication proteins of *Escherichia coli*. *Mol. Microbiol.* 37, 1270–1279. doi: 10.1046/j.1365-2958.2000.02096.x

Shimizu, M., Noguchi, Y., Sakiyama, Y., Kawakami, H., Katayama, T., and Takada, S. (2016). Near-atomic structural model for bacterial DNA replication

initiation complex and its functional insights. *Proc. Natl. Acad. Sci. U.S.A.* 113, E8021–E8030. doi: 10.1073/pnas.1609649113

Skarstad, K., Boye, E., and Steen, H. B. (1986). Timing of initiation of chromosome replication in individual *Escherichia coli* cells. *EMBO J.* 5, 1711–1717.

Skarstad, K., von Meyenburg, K., Hansen, F. G., and Boye, E. (1988). Coordination of chromosome replication initiation in *Escherichia coli*: effects of different dnaA alleles. *J. Bacteriol.* 170, 852–858. doi: 10.1128/jb.170.2.852-858. 1988

Smith, G. R., Contreras-Moreira, B., Zhang, X., and Bates, P. A. (2004). A link between sequence conservation and domain motion within the AAA+ family. *J. Struct. Biol.* 146, 189–204. doi: 10.1016/j.jsb.2003.11.022

Soultanas, P. (2012). Loading mechanisms of ring helicases at replication origins. *Mol. Microbiol.* 84, 6–16. doi: 10.1111/j.1365-2958.2012.08012.x

Stauffer, M. E., and Chazin, W. J. (2004). Structural mechanisms of DNA replication, repair, and recombination. *J. Biol. Chem.* 279, 30915–30918. doi: 10.1074/jbc.R400015200

Strycharska, M. S., Arias-Palomo, E., Lyubimov, A. Y., Erzberger, J. P., O'Shea, V. L., Bustamante, C. J., et al. (2013). Nucleotide and partner-protein control of bacterial replicative helicase structure and function. *Mol. Cell* 52, 844–854. doi: 10.1016/j.molcel.2013.11.016

Sutton, M. D., Carr, K. M., Vicente, M., and Kaguni, J. M. (1998). *Escherichia coli* DnaA protein. The N-terminal domain and loading of DnaB helicase at the *E. coli* chromosomal origin. *J. Biol. Chem.* 273, 34255–34262. doi: 10.1074/jbc. 273.51.34255

Sutton, M. D., and Kaguni, J. M. (1997). Threonine 435 of *Escherichia coli* DnaA protein confers sequence-specific DNA binding activity. *J. Biol. Chem.* 272, 23017–23024. doi: 10.1074/jbc.272.37. 23017

Wolański, M., Donczew, R., Zawilak-Pawlik, A., and Zakrzewska-Czerwińska, J. (2015). *oriC*-encoded instructions for the initiation of bacterial chromosome replication. *Front. Microbiol.* 5:735. doi: 10.3389/fmicb.2014. 00735

Zawilak-Pawlik, A., Nowaczyk, M., and Zakrzewska-Czerwiñska, J. (2017). The role of the N-terminal domains of bacterial initiator DnaA in the assembly and regulation of the bacterial replication initiation complex. *Genes* 8:136. doi: 10.3390/genes8050136

Changing Perspectives on the Role of DnaA-ATP in Orisome Function and Timing Regulation

Alan C. Leonard[1]*, Prassanna Rao[2], Rohit P. Kadam[1] and Julia E. Grimwade[1]

[1] Laboratory of Microbial Genetics, Department of Biomedical and Chemical Engineering and Science, Florida Institute of Technology, Melbourne, FL, United States, [2] Department of Biochemistry, Vanderbilt University School of Medicine, Nashville, TN, United States

Bacteria, like all cells, must precisely duplicate their genomes before they divide. Regulation of this critical process focuses on forming a pre-replicative nucleoprotein complex, termed the orisome. Orisomes perform two essential mechanical tasks that configure the unique chromosomal replication origin, *oriC* to start a new round of chromosome replication: (1) unwinding origin DNA and (2) assisting with loading of the replicative DNA helicase on exposed single strands. In *Escherichia coli*, a necessary orisome component is the ATP-bound form of the bacterial initiator protein, DnaA. DnaA-ATP differs from DnaA-ADP in its ability to oligomerize into helical filaments, and in its ability to access a subset of low affinity recognition sites in the *E. coli* replication origin. The helical filaments have been proposed to play a role in both of the key mechanical tasks, but recent studies raise new questions about whether they are mandatory for orisome activity. It was recently shown that a version of *E. coli oriC* (*oriC^{allADP}*), whose multiple low affinity DnaA recognition sites bind DnaA-ATP and DnaA-ADP similarly, was fully occupied and unwound by DnaA-ADP *in vitro*, and *in vivo* suppressed the lethality of DnaA mutants defective in ATP binding and ATP-specific oligomerization. However, despite their functional equivalency, orisomes assembled on *oriC^{allADP}* were unable to trigger chromosome replication at the correct cell cycle time and displayed a hyper-initiation phenotype. Here we present a new perspective on DnaA-ATP, and suggest that in *E. coli*, DnaA-ATP is not required for mechanical functions, but rather is needed for site recognition and occupation, so that initiation timing is coupled to DnaA-ATP levels. We also discuss how other bacterial types may utilize DnaA-ATP and DnaA-ADP, and whether the high diversity of replication origins in the bacterial world reflects different regulatory strategies for how DnaA-ATP is used to control orisome assembly.

Keywords: *oriC*, DnaA, DNA replication, replication origin, orisomes, pre-replication complexes, DNA binding proteins, cell cycle

INTRODUCTION

The molecular mechanism responsible for triggering new rounds of chromosome replication in bacteria is precisely regulated. New replication forks are initiated from a fixed chromosomal site (*oriC*) only once during each cell division cycle and at a time that is compatible with the cellular growth rate (Cooper and Helmstetter, 1968; Skarstad et al., 1986; Boye et al., 1996; Boye et al., 2000; Leonard and Méchali, 2013). The molecular machine responsible for unwinding origin DNA

and loading the replicative helicase on exposed single strands (termed the orisome) is assembled at *oriC* and comprises multiple copies of the initiator protein, DnaA (Leonard and Grimwade, 2015), whose activity is regulated by binding to ATP (Sekimizu et al., 1987; Katayama et al., 2017). In *E. coli*, the cellular level of DnaA-ATP fluctuates during the cell cycle (Kurokawa et al., 1999), and the reproducibility of initiation timing from one cell cycle to the next is achieved by coupling orisome assembly to DnaA-ATP levels. This is accomplished via a set of specifically arranged low affinity DnaA-ATP recognition sites in *E. coli oriC* that direct orisome assembly by guiding cooperative binding of the initiator (Zawilak-Pawlik et al., 2005; Rozgaja et al., 2011; described in more detail below).

Following each new round of DNA synthesis, several mechanisms are used by bacteria to restrict inappropriate orisome reassembly, reviewed in Nielsen and Løbner-Olesen (2008), Katayama et al. (2010), and Skarstad and Katayama (2013). The predominant regulatory mechanism used in *E. coli* involves hydrolytic conversion of DnaA-ATP into DnaA-ADP by a replication fork-associated process termed Regulatory Inactivation of DnaA (RIDA) (Katayama and Sekimizu, 1999), which causes rapid hydrolysis of DnaA-ATP shortly after initiation (Kurokawa et al., 1999). The DnaA-ADP that is generated cannot reassemble into active orisomes for two reasons. First, it does not readily interact with all of the low affinity recognition sites in *oriC* (McGarry et al., 2004; Kawakami et al., 2005; Grimwade et al., 2018) (see below). Second, unlike the ATP-bound form, DnaA-ADP is unable to form the oligomeric filaments that are essential for binding to ssDNA, a function that is proposed to mediate both origin unwinding and helicase loading (Erzberger et al., 2006; Duderstadt et al., 2010).
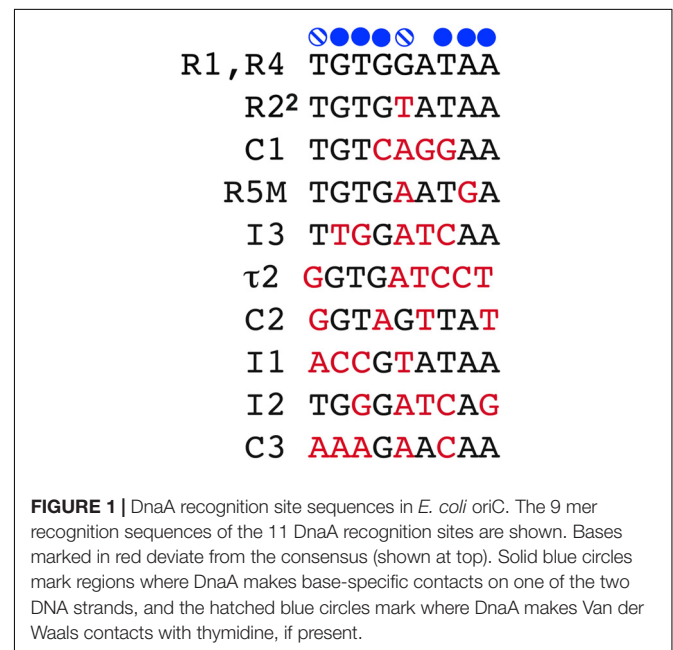
Our main goal for this review is to raise questions about DnaA-ATP's exclusive role as the active initiator form, based on recent findings demonstrating that DnaA-ADP was active in unwinding a synthetic version of *E. coli oriC* (*oriC*^*allADP*) that allows both DnaA-ATP and DnaA-ADP to access all recognition sites (Grimwade et al., 2018). Chromosomal *oriC*^*allADP* was also activated *in vivo* by mutant DnaAs that were defective in adenine nucleotide binding or ATP-dependent oligomerization. However, although functional orisomes were formed on *oriC*^*allADP*, they were unable to trigger properly timed initiation events, revealing that the observed mechanical activity of DnaA-ADP is separate and distinct from the DnaA-ATP-dependent role as a timing regulator. In this review, we discuss the implications of these observations, and discuss how the high level of *oriC* nucleotide sequence diversity among bacterial types may result in orisome assembly pathways that use one or both nucleotide forms for mechanical functions, while reserving the role of DnaA-ATP as a regulator of initiation timing.

ORIGIN RECOGNITION BY DNAA

Almost all bacterial replication origins contain clusters of the 9 bp sequence 5′-TGTGGATAA-3′ (termed the R box) which is the consensus sequence for DnaA recognition. In *E. coli oriC*, there are two R boxes (R1 and R4) that perfectly match the consensus

sequence, and one box (R2) that deviates from consensus by one bp (**Figure 1**); these three sites bind both DnaA-ATP and DnaA-ADP with high affinity (k_d = 4–20 nM) (Sekimizu et al., 1987; Schaper and Messer, 1995). Amino acid residues in the helix-turn-helix motif in DnaA's C-terminal domain (IV) make base-specific hydrogen bonds with nucleotides on one of the two strands at positions 2, 3, 4, 7, 8, and 9 of each R box, as well as Van der Waals contacts with the thymidines that may be present in positions 1 and 6 (Erzberger et al., 2002; Fujikawa et al., 2003) (contacts are summarized at the top of **Figure 1**).

E. coli oriC also contains eight less canonical DnaA binding sites, most of which were identified only after *in vitro* DnaA binding assays (Grimwade et al., 2000; Rozgaja et al., 2011). These cryptic sites deviate from the consensus R box sequence by 2 or more bp (**Figure 1**), which disrupts some base-specific contacts (**Figure 1**). While these sites bind DnaA specifically (McGarry et al., 2004; Rozgaja et al., 2011), their affinity for the initiator is reduced so that dissociation constants for individual sites cannot be measured (Schaper and Messer, 1995). In fact, none of the identified low affinity sites are able to bind DnaA independently; rather, DnaA must be recruited and positioned for them by nearby bound DnaA (Schaper and Messer, 1995; Rozgaja et al., 2011). Six of the lower affinity sites (τ2, I1, I2, I3, C2, and C3) preferentially bind DnaA-ATP (McGarry et al., 2004; Kawakami et al., 2005; Grimwade et al., 2018), and occupation of these sites also requires physiological levels of ATP (0.5–5 mM) (Saxena et al., 2013), as well as interactions between a critical arginine (R285) in DnaA's domain III and the bound ATP of an adjacent DnaA molecule (Kawakami et al., 2005)(discussed further below). While it is not known why these six sites prefer DnaA-ATP, it is probable that conformational differences between DnaA-ATP and DnaA-ADP play a role. The amino acids involved in ATP/ADP binding and hydrolysis are located in a central

FIGURE 1 | DnaA recognition site sequences in *E. coli* oriC. The 9 mer recognition sequences of the 11 DnaA recognition sites are shown. Bases marked in red deviate from the consensus (shown at top). Solid blue circles mark regions where DnaA makes base-specific contacts on one of the two DNA strands, and the hatched blue circles mark where DnaA makes Van der Waals contacts with thymidine, if present.

domain of DnaA (domain III) adjacent to the DNA binding domain (domain IV) (Erzberger et al., 2002; Nishida et al., 2002; Iyer et al., 2004). When bound to ATP, domain IV bends toward domain III, bringing amino acids from both domains into proximity (Erzberger et al., 2006). Physiological levels of ATP are also reported to alter DnaA conformation (Saxena et al., 2015). Conformational changes that alter domain III interactions and allow amino acids outside of domain IV to participate in binding should also increase contacts between DnaA and the low affinity DnaA-ATP sites, thereby compensating for the lack of base-specific DnaA/DNA interactions. Comparing the sequences of the DnaA-ATP sites with the R box sequence (**Figure 1**) suggests that positions 1–4 of the 9 mer binding sites play a greater role in determining preference for DnaA-ATP. It is important to note that not all low affinity sites preferentially bind DnaA-ATP. This is evidenced by the remaining two weak sites in *oriC* (R5M and C1), which were shown by our laboratory to bind both DnaA-ATP and DnaA-ADP (Grimwade et al., 2007, 2018), although there are conflicting reports which show occupation of these sites only by DnaA-ATP (Ozaki et al., 2012). We note that converting the non-discriminatory R5M sequence into the DnaA-ATP-preferring I2 site resulted in delayed initiation *in vivo*, suggesting that R5M is normally occupied by DnaA-ADP (Grimwade et al., 2007).

All of *E. coli*'s 11 DnaA recognition sites lie to the right of the DNA Unwinding Element (DUE) (**Figure 2A**). The three high affinity R boxes are spaced such that R1 is immediately left of the DUE, R2 is central, and R4 is located at the right border of the origin (**Figure 2A**). This widely spaced positioning defines two gap regions where the low affinity sites are located (Rozgaja et al., 2011). Each gap region contains an array of four low affinity sites, each separated from each other by 2 bp (**Figure 2A**). This specific positioning of *oriC* recognition sites facilitates cooperative DnaA binding, and ordered orisome assembly (Rozgaja et al., 2011; described below).
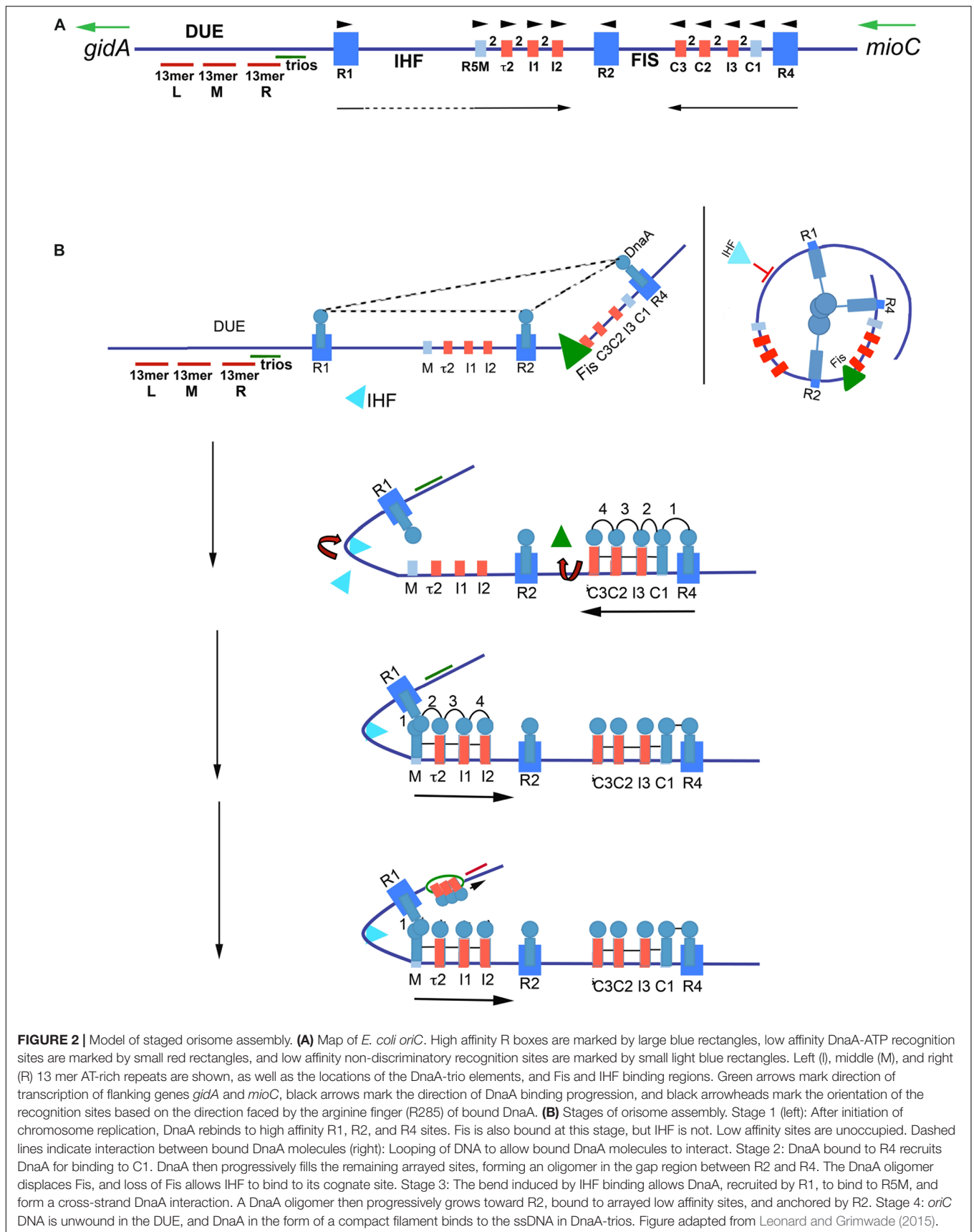
ORDERED ORISOME ASSEMBLY

In *E. coli*, orisome assembly begins when DnaA re-binds to the three high affinity R boxes immediately after the initiation of each round of chromosome replication (Nievera et al., 2006). This tightly bound DnaA plays two important roles. The first is to inhibit unscheduled unwinding of *oriC*, since the DUE is a region of intrinsic helical instability and is subject to spontaneous unwinding when *oriC* is unoccupied (Kowalski and Eddy, 1989). DnaA binding to R1, R2, and R4 constrains *E. coli oriC*, eliminating spontaneous unwinding (Kaur et al., 2014). Although details of the constraint mechanism remain unclear, the most likely scenario involves a trimeric complex formed by interactions among the N-terminal, self-oligomerization domains (domain I) (Simmons et al., 2003) of the bound DnaA molecules (Kaur et al., 2014; **Figure 2B**), perhaps stabilized by the DiaA protein (Ishida et al., 2004). However, domain I-domain I interactions are limited over a distance that is determined by the length of the flexible linker (domain II) that joins each domain I to the rest of the DnaA molecule (Messer et al., 1999;

Nozaki and Ogawa, 2008). Therefore, to make the postulated trimeric complex, *oriC* DNA would need to form loops to place the three bound DnaA molecules close enough to interact, similar to those formed in the nucleosomes of eukaryotes (**Figure 2B**). Alternatively, individual DnaA molecules bound at each R box may be sufficient to clamp the DNA in a way that prevents untwisting without further interactions.

The second role of DnaA binding to R boxes is formation a scaffold that recruits additional DnaA molecules to occupy the adjacent low affinity sites (Miller et al., 2009), and begin the next stage of orisome assembly (**Figure 2B**). Because this role is analogous to that played by the Origin Recognition Complex (ORC) of eukaryotes (Duncker et al., 2009), the structure formed by DnaA binding to the high affinity sites has been termed the bacterial ORC, or bORC (Nievera et al., 2006). DnaA molecules bound to R1 and R4 recruit additional DnaA using their N-terminal domains, and position it for binding to the nearest low affinity site (R1 to R5M and R4 to C1) (Miller et al., 2009; **Figure 2B**). DnaA located at R2 does not normally donate DnaA to either of its nearest sites if R1 and R4 are capable of performing this duty (Rozgaja et al., 2011).

Once DnaA is bound to C1 or R5M, the close positioning of low affinity sites promotes cooperative binding of DnaA-ATP to the remaining sites in the right and left arrays, respectively (**Figure 2B**), progressing from C1 or R5M into the center of *oriC*, toward R2 (Rozgaja et al., 2011). While cooperative binding involves interactions between the domain I regions of donor and recruited DnaA (Rozgaja et al., 2011), domain III regions may also play a role, and the close spacing of the sites is proposed to foster formation of oligomeric DnaA-ATP filaments (Erzberger et al., 2002, 2006; Felczak and Kaguni, 2004; Kawakami et al., 2005). DnaA-ATP oligomers assemble when ATP-associated with DnaA's domain III in one bound molecule interacts with a critical arginine (R285) in the adjacent molecule. R285 comprises DnaA's version of the "arginine finger," a motif that is highly conserved in AAA + (ATPases Associated with various cellular Activities) proteins (Erzberger et al., 2006), with the interaction stabilized by additional amino acid residues (Duderstadt et al., 2010). The orientation of arrayed low affinity binding sites in each half of *oriC* positions bound DnaA-ATP such that their arginine fingers are all facing R2 (Rozgaja et al., 2011; Noguchi et al., 2015). The structures of the two oppositely-oriented DnaA-ATP oligomers have not been solved, but they are presumed to be a more open version of the compact right-handed helical DnaA-ATP filament that has high affinity for single-stranded DNA (Erzberger et al., 2006; Duderstadt et al., 2010).

The 3 bp separation of R4 and C1 allows direct lateral donation of DnaA from a strong to weak site, but the 46 bp distance between R1 and R5M requires DNA bending and cross-strand donation for cooperative binding (Rozgaja et al., 2011). This bend requirement is the basis for a growth rate-regulated switch that ensures synchronous initiations of the multiple copies of *oriC* that obtain during rapid growth conditions (Cooper and Helmstetter, 1968; Roth et al., 1994). During rapid growth, Fis, a growth rate-regulated protein (Nilsson et al., 1992; Mallik et al., 2006), binds to its recognition site between R2 and C3 shortly after the initiation step (Cassler et al., 1995; **Figure 2B**), during the time

FIGURE 2 | Model of staged orisome assembly. **(A)** Map of *E. coli oriC*. High affinity R boxes are marked by large blue rectangles, low affinity DnaA-ATP recognition sites are marked by small red rectangles, and low affinity non-discriminatory recognition sites are marked by small light blue rectangles. Left (l), middle (M), and right (R) 13 mer AT-rich repeats are shown, as well as the locations of the DnaA-trio elements, and Fis and IHF binding regions. Green arrows mark direction of transcription of flanking genes *gidA* and *mioC*, black arrows mark the direction of DnaA binding progression, and black arrowheads mark the orientation of the recognition sites based on the direction faced by the arginine finger (R285) of bound DnaA. **(B)** Stages of orisome assembly. Stage 1 (left): After initiation of chromosome replication, DnaA rebinds to high affinity R1, R2, and R4 sites. Fis is also bound at this stage, but IHF is not. Low affinity sites are unoccupied. Dashed lines indicate interaction between bound DnaA molecules (right): Looping of DNA to allow bound DnaA molecules to interact. Stage 2: DnaA bound to R4 recruits DnaA for binding to C1. DnaA then progressively fills the remaining arrayed sites, forming an oligomer in the gap region between R2 and R4. The DnaA oligomer displaces Fis, and loss of Fis allows IHF to bind to its cognate site. Stage 3: The bend induced by IHF binding allows DnaA, recruited by R1, to bind to R5M, and form a cross-strand DnaA interaction. A DnaA oligomer then progressively grows toward R2, bound to arrayed low affinity sites, and anchored by R2. Stage 4: *oriC* DNA is unwound in the DUE, and DnaA in the form of a compact filament binds to the ssDNA in DnaA-trios. Figure adapted from Leonard and Grimwade (2015).

period that *oriC* is constrained by DnaA occupying the three high affinity sites (Kaur et al., 2014). The Fis-bound bORC prevents IHF from binding and bending at its cognate site between R1 and R5M (Ryan et al., 2004; Kaur et al., 2014), possibly because the constrained bORC does not allow two bends to be simultaneously placed in *oriC*. The inhibition of bending results in a temporary block of DnaA binding in the left half of *oriC*. As DnaA-ATP levels increase during the cell cycle, progressive DnaA occupation of the right array of sites displaces Fis (Ryan et al., 2004), allowing IHF to bind, resulting in a DNA bend that places R1 sufficiently close to R5M to nucleate filling of *oriC*'s left side low affinity sites (Rozgaja et al., 2011). By acting as a temporary partition between the left and right halves of *oriC* (Gille et al., 1991), Fis is able to delay initiation until the total number of DnaA molecules in the cell exceeds that needed for initiation of a single *oriC* copy; thus, when Fis is finally displaced, all origins in the cell can complete orisome assembly and initiate synchronously (Ryan et al., 2004; Rao et al., 2018). In this way, Fis becomes the primary regulator of initiation timing under rapid growth conditions (Flåtten and Skarstad, 2013). In contrast, during slow growth when *E. coli* carries only one *oriC* copy, Fis levels are too low to occupy *oriC* (Nilsson et al., 1992), and IHF is able to bind and bend the DNA between R1 and R5M, promoting low affinity site occupation in the left region of *oriC* independently of the filling of the right region. In this case, orisome completion and initiation timing is dependent only on the cellular levels of DnaA-ATP being high enough to fill the low affinity DnaA-ATP sites (Rao et al., 2018). At all growth rates, DnaA-ATP occupation of the low affinity sites promotes opening of the DNA duplex in the right region of the DUE (Bramhill and Kornberg, 1988; Grimwade et al., 2000; **Figure 2B**). However, there is evidence that not all the low affinity sites in *E. coli oriC* are essential for *in vivo* activity (Stepankiw et al., 2009), and *in vitro*, only R5M needs to be occupied by DnaA for unwinding (Sakiyama et al., 2017).

A variety of models have been proposed to explain the mechanism of unwinding (Speck and Messer, 2001; Erzberger et al., 2006; Ozaki et al., 2008; Duderstadt et al., 2011; Ozaki and Katayama, 2012; Zorman et al., 2012), and both the compact and open versions of DnaA-ATP oligomers are implicated in producing the torsional stress required for DNA unwinding. Proposed mechanisms include: an open DnaA-ATP oligomer bound to double-stranded DNA causing formation of right handed supertwists (Erzberger et al., 2006; Zorman et al., 2012); an open DnaA-ATP oligomer bound to double-stranded DNA in the left array of low affinity sites creating a channel that can engage and unwind DUE DNA (Ozaki et al., 2008, 2012) and a compact DnaA-ATP oligomer stretching and unwinding DUE DNA (Duderstadt et al., 2011; Duderstadt and Berger, 2013).

Once unwound, the single-stranded DNA binds to DnaA-ATP, which stabilizes the open structure (**Figure 2B**) to promote expansion of the initiation bubble and assist with DNA helicase delivery (Yung and Kornberg, 1989; Speck and Messer, 2001). In *Bacillus subtilis*, the additional DnaA-ATP used for this purpose was shown to interact with specialized 3 bp sequence motifs, termed DnaA-trios (Richardson et al., 2016; **Figure 2A**), and it is proposed that the trio elements are a conserved aspect of replication origins. The two end bases of trios can vary, but
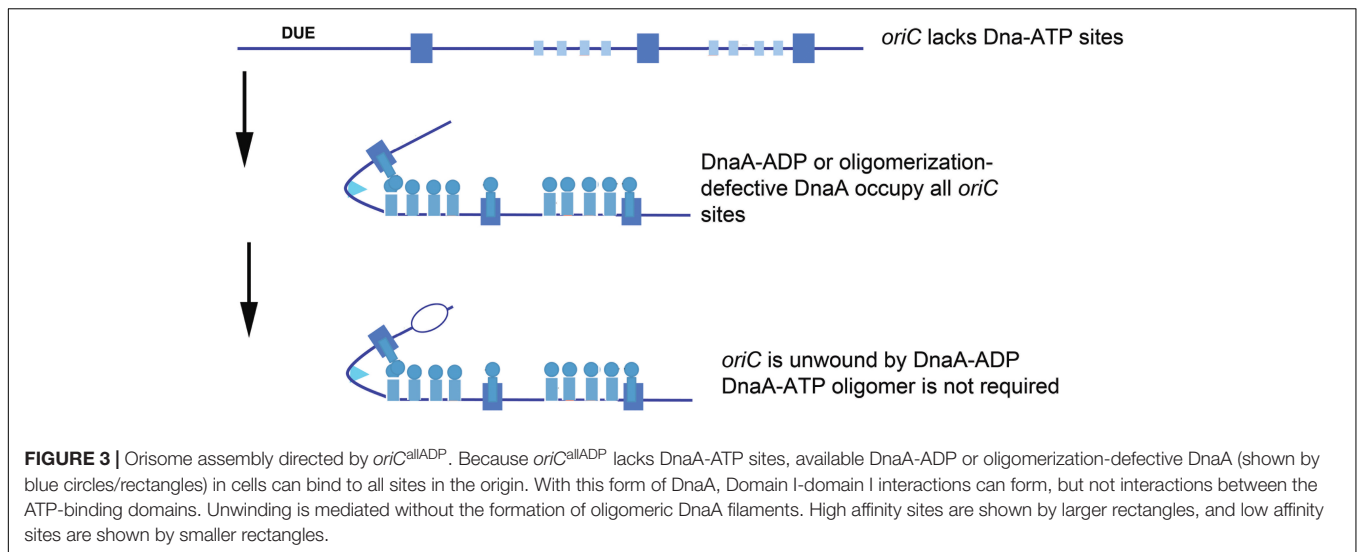
the middle nucleotide must be A (Richardson et al., 2016). In many bacterial types, there are seven to ten direct repeats of DnaA-trios between the DUE and the nearest (3′) high affinity DnaA recognition site (Richardson et al., 2016); *E. coli* has one of the shorter arrays, containing only three trios. In addition to the oligomer formed using trio-elements, the DnaA bound to the right half of *oriC* has also been implicated in DNA helicase loading (Ozaki and Katayama, 2012).

A PREDOMINANT ROLE FOR *E. coli* DNAA-ATP IS IN ORIGIN RECOGNITION AND REGULATION OF INITIATION TIMING

Although DnaA-ATP is required for activation of wild type *E. coli oriC in vitro*, it has been known for several decades that at least some of the DnaA in functional *E. coli* orisomes can be in the ADP-bound form (Yung et al., 1990). The recognition sites occupied by DnaA-ADP in these mixed orisomes was never identified, but all of the R boxes, as well as R5M and C1, are obvious candidates. In support of this idea, a clever heterologous DnaA binding assay was recently used to demonstrate that functional orisomes could be built when either R1 or R4 was occupied by DnaA-ADP (Noguchi et al., 2015).

Regardless of binding locations, the ability to use DnaA-ADP as a component of functional *E. coli* orisomes raises questions about DnaA-ATP as the active form of the initiator. Is DnaA-ATP the active form because it is the only form that can fill all recognition sites, or because it is the only form that can make the higher order oligomeric structures that can perform essential mechanical tasks? To address these issues, a novel version of *oriC* (*oriC*^{allADP}) was constructed that converted every DnaA-ATP recognition site to one that bound either DnaA-ADP or DnaA-ATP with equivalent low affinities (e.g., each low affinity site was made similar to C1 and R5M) (McGarry et al., 2004; Grimwade et al., 2018). By using *oriC*^{allADP}, it was possible to examine the activity of orisomes assembled from only DnaA-ADP. Surprisingly, *in vitro*, *oriC*^{allADP} plasmids were unwound equally by orisomes assembled with either DnaA-ATP and DnaA-ADP. *In vivo*, use of *oriC*^{allADP} as the sole chromosomal replication origin also suppressed the lethality of DnaA mutants with defects in ATP binding and ATP-dependent oligomer formation [DnaA46 and DnaA(R285A), respectively, Grimwade et al., 2018]. Thus, given equal access to *oriC*, both DnaA-ADP and DnaA-ATP are functionally equivalent, with orisomes assembled from either form capable of performing the mechanical actions required to trigger initiation in *E. coli* (**Figure 3**).

These observations lead to the conclusion that the predominant role for DnaA-ATP in activating wild type *E. coli oriC* must be for origin recognition and site occupation. Since it is normally the case that DnaA-ATP preferentially binds most low affinity sites, initiation timing must be coupled to the availability of this form during the cell cycle. Consistent with this idea, cells triggering chromosome replication from *oriC*^{allADP} behaved as if initiation timing was no longer dependent on

FIGURE 3 | Orisome assembly directed by oriC^allADP. Because oriC^allADP lacks DnaA-ATP sites, available DnaA-ADP or oligomerization-defective DnaA (shown by blue circles/rectangles) in cells can bind to all sites in the origin. With this form of DnaA, Domain I-domain I interactions can form, but not interactions between the ATP-binding domains. Unwinding is mediated without the formation of oligomeric DnaA filaments. High affinity sites are shown by larger rectangles, and low affinity sites are shown by smaller rectangles.

DnaA-ATP levels. These cells over-initiated, and consequently showed increased sensitivity to replicative stress (Grimwade et al., 2018). Apparently, since DnaA-ADP is not normally degraded in *E. coli*, it was continuously available at levels sufficient to bind to low affinity sites in *oriC^allADP* and trigger multiple replication rounds. Additional studies, in which only one or two of the DnaA-ATP sites were converted to a version that binds both forms of DnaA equivalently (Rao et al., 2018), revealed that at slow growth rates, each site contributed to the DnaA-ATP regulated initiation timing mechanism. At fast growth rates, Fis, by virtue of its ability to regulate DnaA binding, took over as the major timing regulator, as described above and in Rao et al. (2018). Combined, the data on these synthetic *oriC*s demonstrate that the features of bacterial replication origins involved in mechanical function can be separated from their timing components(s).
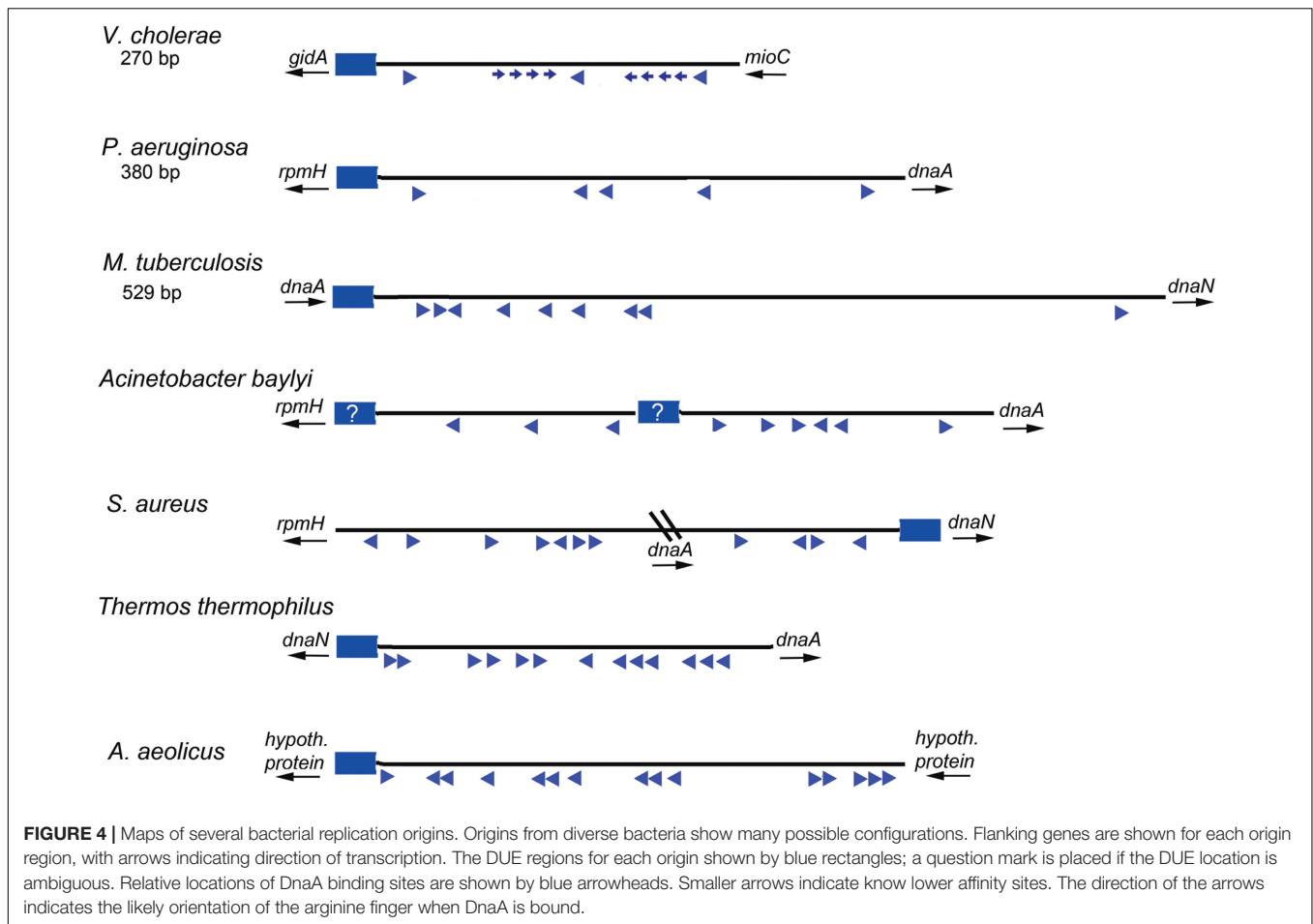
The conclusion that DnaA-ADP can activate *E. coli oriC* does not appear to be compatible with models for *E. coli* origin unwinding that invoke assembly of oligomeric DnaA-ATP filaments (see above), although it has yet to be determined whether orisomes made from only DnaA-ADP or DnaA-ATP function in exactly the same way. It is possible that when DnaA-ADP molecules are aligned by binding to arrayed sites, they are capable of forming an unwinding structure similar to one formed by DnaA-ATP, however, if this is the case, the requirement for DnaA-ATP would still be for binding to arrayed sites, not for a unique ability to oligomerize. Alternatively, unwinding mediated by DnaA-ADP might rely on DnaA's inherent DNA bending activity. DnaA produces a 30–40° bend in DNA when bound to a 9 mer recognition site (Schaper and Messer, 1995). The concerted bending at multiple sites could provide sufficient stress to unwind the DUE. This mechanism could either replace the need for a DnaA-ATP filament, or it could be used by both DnaA-ATP and DnaA-ADP. If the bending model is correct, then DnaA would produce DNA distortions similar to those caused by binding of archaeal and eukaryotic initiator proteins, generating sufficient torsional stress to unwind

the AT-rich DUE (Dueber et al., 2007; Gaudier et al., 2007; Sun et al., 2012).

The observed functionality of DnaA-ADP is also not consistent with mechanisms for unwinding and helicase loading that involve DnaA-ATP filaments associated with DnaA-trios. However, since trio occupation requires DnaA bound to a nearby high affinity R-box (Richardson et al., 2016), and because the trio-proximal R box (R1) is not essential for *E. coli oriC* function (Kaur et al., 2014), it is not known whether DnaA-trios are required in *E. coli*. Thus, *E. coli* may be able to use an alternate mechanism for helicase loading that is not dependent on any unique property of DnaA-ATP.

THOUGHTS ABOUT THE REQUIREMENT FOR DNAA-ATP IN ASSEMBLING ORISIOMES ON DIVERSE REPLICATION ORIGIN TEMPLATES

Based on the studies of *E. coli* orisome assembly, described above, it is clear that the arrangement and nucleotide sequence of DnaA recognition sites in *E. coli oriC* directs ordered orisome assembly, and also couples the cell cycle timing of this process to the availability of DnaA-ATP. Because all other bacterial types must also assemble functional orisomes at the correct cell cycle time, and because DnaA is a highly conserved protein, it is reasonable to expect that the majority of the bacterial *oriC* templates would also be conserved and direct orisome assembly in the same way as *E. coli*. However, this is definitely not the case. A database (DoriC 10.0) containing the nucleotide sequences of thousands of *oriC*s (some putative) reveals enormous diversity among bacterial types, with little overt similarity to most of the features found in *E. coli* other than the presence of multiple R box-type DnaA recognition sites (Luo and Gao, 2019). **Figure 4** depicts a few different *oriC* geographies, showing dramatic differences in the number and relative positions of the R-box-like

FIGURE 4 | Maps of several bacterial replication origins. Origins from diverse bacteria show many possible configurations. Flanking genes are shown for each origin region, with arrows indicating direction of transcription. The DUE regions for each origin shown by blue rectangles; a question mark is placed if the DUE location is ambiguous. Relative locations of DnaA binding sites are shown by blue arrowheads. Smaller arrows indicate know lower affinity sites. The direction of the arrows indicates the likely orientation of the arginine finger when DnaA is bound.

sequences, including both widely separated and closely spaced clusters. However, the variety is far more extensive than can be demonstrated by one figure, and additional details can be found in several papers and reviews (Zawilak-Pawlik et al., 2005; Zakrzewska-Czerwińska et al., 2007; Donczew et al., 2012; Leonard and Méchali, 2013; Wolański et al., 2014; Jaworski et al., 2016). Further, it is likely that cryptic low affinity sites exist in a variety of bacterial origins, but because sequence analysis identifies DnaA binding sites based on their similarity to the consensus R box, DnaA-*oriC* binding assays are required to identify more divergent DnaA recognition sites. Thus, cryptic sites have been mapped in the replication origins of only a few bacterial types other than *E. coli* and its close relatives (Charbon and Lobner-Olesen, 2011; Taylor et al., 2011), and sites similar to the DnaA-ATP sites in *E. coli oriC*'s have not been positively identified in any other bacterial origin.

While we propose that DnaA-ADP might have a greater role than previously believed, it is important to note that a major reason for origin diversity (and the utilization of different forms of DnaA) is that R boxes can be used for functions other than orisome assembly, such as regulating initiation timing (DnaA availability) or for transcriptional regulation. Unlike *E. coli oriC*, which is positioned between the *gidA* and *mioC* genes, many bacterial replication origins are located next to the

dnaA gene (see **Figure 4** for examples). An interesting alternative arrangement in some bacteria places *dnaA* within the interior of *oriC* producing a bi-partite configuration (for examples see *Staphylococcus oriC* in **Figure 4** and the *Helicobacter pylori oriC* (Donczew et al., 2012), such that there are clusters of R boxes on either side of *dnaA*. Since the *dnaA* promoter contains DnaA recognition sites used for autoregulation (Atlung et al., 1985; Braun et al., 1985; Ogura et al., 2001), when *dnaA* and *oriC* are adjacent, it is difficult to distinguish R boxes used to regulate *dnaA* expression (by DnaA-ATP and DnaA-ADP) from those used for orisome assembly. Further, some of the R boxes in certain bacterial origin regions may be used to regulate DnaA availability (and initiation timing) by titration (Moriya et al., 1988). In *E. coli*, sites that can titrate DnaA-ATP or DnaA-ADP (Hansen et al., 1991) are located outside of *oriC*, both as individual DnaA recognition sites distributed around the chromosome as well as within a region with high DnaA capacity termed *datA*, located about 460 kb from *oriC* where bound DnaA-ATP is inactivated (Ogawa et al., 2002; Kasho and Katayama, 2013). For some bacteria, *datA*-like sequences may be found in locations proximal to or within *oriC*. It is also possible that all DnaA binding sites in an origin are simply not necessary for functional orisome assembly. For example, in *E. coli*, low affinity sites between R1 and R2 (left side) are implicated in

origin unwinding, but the right side sites are not (Stepankiw et al., 2009), although they may play a supportive role in helicase loading (Ozaki and Katayama, 2012). Similarly, only a few of the DnaA boxes in the *B. subtilis* origin, near the DUE, are essential for mechanical functions (Richardson et al., 2019). Recognition sites for regulatory proteins could also contribute to origin diversity. Such regulators would include DNA bending proteins (such as analogs of Fis and IHF) (Brassinga et al., 2002), and proteins which block the interaction of DnaA with their respective recognition sites or suppress cooperative DnaA interactions during orisome assembly. Examples of the latter are described below.

Even after considering regulatory and titration sites, the high variability among bacterial origins raises the obvious conclusion that, although the initiator is conserved, and the essential mechanical functions required for initiation are the same in all bacteria, different assembly paths must be used to form the orisomes that ultimately perform these functions (Jakimowicz et al., 2000; Zawilak-Pawlik et al., 2005). The details of these diverse paths, and how they might utilize DnaA-ATP and DnaA-ADP for mechanical and timing functions remain unanswered questions, but we can speculate about several possibilities.

Since many bacteria carry DnaA-trio sequence motifs located between the DUE and its most proximal R box (Richardson et al., 2016), this feature might play a key role in setting the requirement for DnaA-ATP, or even allowing DnaA-ADP to participate in orisome assembly. Although there is insufficient evidence to determine if they are essential for every bacterial origin, in *B. subtilis* and probably other bacteria, DnaA-trios direct the assembly of critical DnaA-ATP oligomers, and could set the amount of DnaA-ATP required for unwinding and the DNA helicase loading steps (Richardson et al., 2019). In some bacteria, only a small amount of DnaA-ATP may be needed to interact at DnaA-trio elements to effect stable strand separation and/or helicase loading, and the rest of the orisome, including a sub-complex that mediates initial unwinding, could be assembled from DnaA-ATP or DnaA-ADP, depending on the specific origin, as described below.

Other than DnaA-trios, some replication origins appear to lack any recognition sites with preference for DnaA-ATP. This seems to be the configuration of the *oriC*s in *B. subtilis*, *C. crescentus*, and *M. tuberculosis*, among others (Leonard and Méchali, 2013; Wolański et al., 2014). For these origins, the most available form of DnaA in the cell would be used to assemble the orisome, but the active form is expected to be tightly regulated at the level of synthesis and during the inter-initiation interval. Some of the different mechanisms that regulate the availability of DnaA-ATP might also apply to the ADP-bound form if it plays a role in orisome assembly or origin activation. For example, in *C. crescentus*, DnaA-ATP is hydrolyzed by RIDA, but the resulting DnaA-ADP is then degraded by Lon protease (Wargachuk and Marczynski, 2015), and in *B. subtilis* and *S. aureus*, DnaA can rapidly exchange the bound ADP for ATP (Kurokawa et al., 2009; Bonilla and Grossman, 2012). Use of inhibitory proteins to block DnaA access to *oriC* binding sites would be equally effective for DnaA-ATP and DnaA-ADP. Known examples

include CtrA in *C. crescentus* (Quon et al., 1998), AdpA in *Streptomyces* (Wolański et al., 2012), MtrA in *Mycobacteria* (Rajagopalan et al., 2010), and HP1021 in *Helicobacter* (Donczew et al., 2015). Topologically-sensitive DnaA binding sites identified in *H. pylori oriC* are an intriguing regulatory feature that would also be compatible with active DnaA-ATP or DnaA-ADP initiator, allowing DnaA to interact at some sites only when binding at other sites changes the origin's superhelical density (Donczew et al., 2014). Anti-cooperativity factors are known to block DnaA-ATP oligomerization at some stage of orisome assembly. Versions include YabA (Merrikh and Grossman, 2011; Scholefield and Murray, 2013), SirA (Rahn-Lee et al., 2011), Soj (Scholefield et al., 2012), DnaD (Bonilla and Grossman, 2012; Scholefield and Murray, 2013), and Spo0A (Boonstra et al., 2013). While not yet identified, it is possible that factors may exist to block cooperative interaction between DnaA-ADP molecules (probably by blocking domain I interactions).

Some bacteria with the ability to assemble orisomes with a mixture of DnaA-ATP and DnaA-ADP may require cooperative binding in order to fill some of the DnaA binding sites, even those that are reported as consensus. For example, in the *Mycobacterium tuberculosis* origin, no individual R box can bind DnaA independently; rather cooperative binding between at least two recognition sites is required for the occupation the *oriC* (Zawilak-Pawlik et al., 2005). (It should be noted that *M. tuberculosis oriC* contains no R boxes with the 5′-TTATCCACA consensus sequence, so it is possible that none of the sites in *oriC^{Mtb}* have high enough affinity for DnaA to bind without cooperative interactions.) Formation of a bORC and progression to complete orisomes in this type of bacteria would require that DnaA recognition sites be closely spaced to allow interactions. This arrangement could be compatible with DnaA-ADP if domain I interactions were sufficient, but not all bacterial DnaAs can associate using domain I (Zawilak-Pawlik et al., 2017), and domain III interactions between DnaA-ATP molecules may be used exclusively. With an expanded view of DnaA-ADP activity, one can also envision origins in which DnaA-ATP is needed for the cooperative interactions used for site filling, but once DnaA-ATP is bound, ATP hydrolysis might provide a conformational change required for origin activation. The hydrolysis step could be intrinsic to the DnaA-ATP complex, or regulated by a factor analogous to the Hda protein in *E coli* (Katayama et al., 2017). Such a mechanism would explain why ATPase activity is required for complete orisome assembly in *M. tuberculosis* (Madiraju et al., 2006). In this scenario, DnaA-ATP would be required for origin recognition, but DnaA-ADP would perform the mechanical functions triggering initiation.

Under certain extreme conditions, such as the high temperatures, DnaA-ATP oligomers may be preferentially used to stabilize the orisome complex. The origins of these bacteria would have to contain closely spaced recognition sites to optimize the interaction between adjacent AAA + domains. Consistent with this idea, R boxes clustered in closely spaced arrays have been observed in the *oriC*s of the thermophilic bacteria *Thermus thermophilus* (**Figure 4**; Schaper et al., 2000) and *Aquifex aeolicus* (Erzberger et al., 2006). For these bacteria, DnaA-ATP oligomerization would be required for initiation, and

it would be unlikely that functional orisomes would assembled using DnaA-ADP, even if that form could bind to the origin.

While the significance of any given arrangement of DnaA recognition sites remains speculative, regardless of the requirements for the ATP or ADP-bound forms, there is ample evidence that the configuration of every bacterial replication origin is optimized for its own DnaA, (Zawilak-Pawlik et al., 2005). For example, the DnaA proteins of both *E. coli* and *B. subtilis* bind with high affinities toward the same DnaA box sequence *in vitro* and create similar multimeric structures when visualized by EM (Krause et al., 1997). However, despite these apparent similarities, neither *E. coli* nor *B. subtilis* DnaA was able to unwind its heterologous partner origin. Similarly, while both *E. coli* and *M. tuberculosis* DnaAs bind well to *S. coelicolor oriC*, neither can bend the origin into the structure formed by the native DnaA protein (Jakimowicz et al., 2000; Zawilak-Pawlik et al., 2005). Further, heterologous *oriCs* replicate autonomously as plasmids or on the chromosome of another bacterial type only when their nucleotide sequences are nearly identical (Takeda et al., 1982; Zyskind et al., 1983; O'Neill and Bender, 1988; Roggenkamp, 2007; Demarre and Chattoraj, 2010). These data, combined with

the many possible versions of *oriC* geography and accompanying regulation, make it difficult to determine whether there are features of orisome assembly widely shared by many bacterial orisomes. It is clear that more extensive analysis of different bacteria, as well as further analysis of synthetic origins, such as *oriC^{allADP}*, will be necessary to reveal common paradigms for bacterial replication initiation and the specific roles of different DnaA forms.

AUTHOR CONTRIBUTIONS

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

FUNDING

REFERENCES

Atlung, T., Clausen, E. S., and Hansen, F. G. (1985). Autoregulation of the dnaA gene of *Escherichia coli* K12. *Mol. Gen. Genet.* 200, 442–450. doi: 10.1007/bf00425729

Bonilla, C. Y., and Grossman, A. D. (2012). The primosomal protein DnaD inhibits cooperative DNA binding by the replication initiator DnaA in Bacillus subtilis. *J. Bacteriol.* 194, 5110–5117. doi: 10.1128/JB.00958-12

Boonstra, M., de Jong, I. G., Scholefield, G., Murray, H., Kuipers, O. P., and Veening, J. W. (2013). Spo0A regulates chromosome copy number during sporulation by directly binding to the origin of replication in *Bacillus subtilis*. *Mol. Microbiol.* 87, 925–938. doi: 10.1111/mmi.12141

Boye, E., Løbner-Olesen, A., and Skarstad, K. (2000). Limiting DNA replication to once and only once. *EMBO Rep.* 1, 479–483. doi: 10.1093/embo-reports/kvd116

Boye, E., Stokke, T., Kleckner, N., and Skarstad, K. (1996). Coordinating DNA replication initiation with cell growth: differential roles for DnaA and SeqA proteins. *Proc. Natl. Acad. Sci. U.S.A.* 93, 12206–12211. doi: 10.1073/pnas.93.22.12206

Bramhill, D., and Kornberg, A. (1988). Duplex opening by dnaA protein at novel sequences in initiation of replication at the origin of the *E. coli* chromosome. *Cell* 52, 743–755. doi: 10.1016/0092-8674(88)90412-6

Brassinga, A. K., Siam, R., McSween, W., Winkler, H., Wood, D., and Marczynski, G. T. (2002). Conserved response regulator CtrA and IHF binding sites in the alpha-proteobacteria *Caulobacter crescentus* and *Rickettsia prowazekii* chromosomal replication origins. *J. Bacteriol.* 184, 5789–5799. doi: 10.1128/jb.184.20.5789-5799.2002

Braun, R. E., O'Day, K., and Wright, A. (1985). Autoregulation of the DNA replication gene dnaA in *E. coli* K-12. *Cell* 40, 159–169. doi: 10.1016/0092-8674(85)90319-8

Cassler, M. R., Grimwade, J. E., and Leonard, A. C. (1995). Cell cycle-specific changes in nucleoprotein complexes at a chromosomal replication origin. *EMBO J.* 14, 5833–5841. doi: 10.1002/j.1460-2075.1995.tb00271.x

Charbon, G., and Lobner-Olesen, A. (2011). A role for the weak DnaA binding sites in bacterial replication origins. *Mol. Microbiol.* 82, 272–274. doi: 10.1111/j.1365-2958.2011.07840.x

Cooper, S., and Helmstetter, C. E. (1968). Chromosome replication and the division cycle of *Escherichia coli* B/r. *J. Mol. Biol.* 31, 519–540. doi: 10.1016/0022-2836(68)90425-7

Demarre, G., and Chattoraj, D. K. (2010). DNA adenine methylation is required to replicate both *Vibrio cholerae* chromosomes once per cell cycle. *PLoS Genet.* 6:e1000939. doi: 10.1371/journal.pgen.1000939

Donczew, R., Makowski, Ł, Jaworski, P., Bezulska, M., Nowaczyk, M., Zakrzewska-Czerwińska, J., et al. (2015). The atypical response regulator HP1021 controls formation of the *Helicobacter* pylori replication initiation complex. *Mol. Microbiol.* 95, 297–312. doi: 10.1111/mmi.12866

Donczew, R., Mielke, T., Jaworski, P., Zakrzewska-Czerwińska, J., and Zawilak-Pawlik, A. (2014). Assembly of *Helicobacter* pylori initiation complex is determined by sequence-specific and topology-sensitive DnaA-oriC interactions. *J. Mol. Biol.* 426, 2769–2782. doi: 10.1016/j.jmb.2014.05.018

Donczew, R., Weigel, C., Lurz, R., Zakrzewska-Czerwinska, J., and Zawilak-Pawlik, A. (2012). *Helicobacter* pylori oriC–the first bipartite origin of chromosome replication in Gram-negative bacteria. *Nucleic Acids Res.* 40, 9647–9660. doi: 10.1093/nar/gks742

Duderstadt, K. E., and Berger, J. M. (2013). A structural framework for replication origin opening by AAA+ initiation factors. *Curr. Opin. Struct. Biol.* 23, 144–153. doi: 10.1016/j.sbi.2012.11.012

Duderstadt, K. E., Chuang, K., and Berger, J. M. (2011). DNA stretching by bacterial initiators promotes replication origin opening. *Nature* 478, 209–213. doi: 10.1038/nature10455

Duderstadt, K. E., Mott, M. L., Crisona, N. J., Chuang, K., Yang, H., and Berger, J. M. (2010). Origin remodeling and opening in bacteria rely on distinct assembly states of the DnaA initiator. *J. Biol. Chem.* 285, 28229–28239. doi: 10.1074/jbc.M110.147975

Dueber, E. L., Corn, J. E., Bell, S. D., and Berger, J. M. (2007). Replication origin recognition and deformation by a heterodimeric archaeal Orc1 complex. *Science* 317, 1210–1213. doi: 10.1126/science.1143690

Duncker, B. P., Chesnokov, I. N., and McConkey, B. J. (2009). The origin recognition complex protein family. *Genome Biol.* 10:214. doi: 10.1186/gb-2009-10-3-214

Erzberger, J. P., Mott, M. L., and Berger, J. M. (2006). Structural basis for ATP-dependent DnaA assembly and replication-origin remodeling. *Nat. Struct. Mol. Biol.* 13, 676–683. doi: 10.1038/nsmb1115

Erzberger, J. P., Pirruccello, M. M., and Berger, J. M. (2002). The structure of bacterial DnaA: implications for general mechanisms underlying DNA replication initiation. *EMBO J.* 21, 4763–4773. doi: 10.1093/emboj/cdf496

Felczak, M. M., and Kaguni, J. M. (2004). The box VII motif of *Escherichia coli* DnaA protein is required for DnaA oligomerization at the *E. coli* replication origin. *J. Biol. Chem.* 279, 51156–51162. doi: 10.1074/jbc.M409695200

Flåtten, I., and Skarstad, K. (2013). The Fis protein has a stimulating role in initiation of replication in *Escherichia coli* in vivo. *PLoS One* 8:e83562. doi: 10.1371/journal.pone.0083562

Fujikawa, N., Kurumizaka, H., Nureki, O., Terada, T., Shirouzu, M., Katayama, T., et al. (2003). Structural basis of replication origin recognition by the DnaA protein. *Nucleic Acids Res.* 31, 2077–2086. doi: 10.1093/nar/gkg309

Gaudier, M., Schuwirth, B. S., Westcott, S. L., and Wigley, D. B. (2007). Structural basis of DNA replication origin recognition by an ORC protein. *Science* 317, 1213–1216. doi: 10.1126/science.1143664

Gille, H., Egan, J. B., Roth, A., and Messer, W. (1991). The FIS protein binds and bends the origin of chromosomal DNA replication, oriC, of *Escherichia coli*. *Nucleic Acids Res.* 19, 4167–4172. doi: 10.1093/nar/19.15.4167

Grimwade, J. E., Rozgaja, T. A., Gupta, R., Dyson, K., Rao, P., and Leonard, A. C. (2018). Origin recognition is the predominant role for DnaA-ATP in initiation of chromosome replication. *Nucleic Acids Res.* 46, 6140–6151. doi: 10.1093/nar/gky457

Grimwade, J. E., Ryan, V. T., and Leonard, A. C. (2000). IHF redistributes bound initiator protein, DnaA, on supercoiled oriC of *Escherichia coli*. *Mol. Microbiol.* 35, 835–844. doi: 10.1046/j.1365-2958.2000.01755.x

Grimwade, J. E., Torgue, J. J., McGarry, K. C., Rozgaja, T., Enloe, S. T., and Leonard, A. C. (2007). Mutational analysis reveals *Escherichia coli* oriC interacts with both DnaA-ATP and DnaA-ADP during pre-RC assembly. *Mol. Microbiol.* 66, 428–439. doi: 10.1111/j.1365-2958.2007.05930.x

Hansen, F. G., Christensen, B. B., and Atlung, T. (1991). The initiator titration model: computer simulation of chromosome and minichromosome control. *Res. Microbiol.* 142, 161–167. doi: 10.1016/0923-2508(91)90025-6

Ishida, T., Akimitsu, N., Kashioka, T., Hatano, M., Kubota, T., Ogata, Y., et al. (2004). DiaA, a novel DnaA-binding protein, ensures the timely initiation of *Escherichia coli* chromosome replication. *J. Biol. Chem.* 279, 45546–45555. doi: 10.1074/jbc.M402762200

Iyer, L. M., Leipe, D. D., Koonin, E. V., and Aravind, L. (2004). Evolutionary history and higher order classification of AAA+ ATPases. *J. Struct. Biol.* 146, 11–31. doi: 10.1016/j.jsb.2003.10.010

Jakimowicz, D., Majkadagger, J., Konopa, G., Wegrzyn, G., Messer, W., Schrempf, H., et al. (2000). Architecture of the *Streptomyces* lividans DnaA protein-replication origin complexes. *J. Mol. Biol.* 298, 351–364. doi: 10.1006/jmbi.2000.3686

Jaworski, P., Donczew, R., Mielke, T., Thiel, M., Oldziej, S., Weigel, C., et al. (2016). Unique and universal features of Epsilonproteobacterial origins of chromosome replication and DnaA-DnaA box interactions. *Front. Microbiol.* 7:1555. doi: 10.3389/fmicb.2016.01555

Kasho, K., and Katayama, T. (2013). DnaA binding locus datA promotes DnaA-ATP hydrolysis to enable cell cycle-coordinated replication initiation. *Proc. Natl. Acad. Sci. U.S.A.* 110, 936–941. doi: 10.1073/pnas.1212070110

Katayama, T., Kasho, K., and Kawakami, H. (2017). The DnaA cycle in *Escherichia coli*: activation, function and inactivation of the initiator protein. *Front. Microbiol.* 8:2496. doi: 10.3389/fmicb.2017.02496

Katayama, T., Ozaki, S., Keyamura, K., and Fujimitsu, K. (2010). Regulation of the replication cycle: conserved and diverse regulatory systems for DnaA and oriC. *Nat. Rev. Microbiol.* 8, 163–170. doi: 10.1038/nrmicro2314

Katayama, T., and Sekimizu, K. (1999). Inactivation of *Escherichia coli* DnaA protein by DNA polymerase III and negative regulations for initiation of chromosomal replication. *Biochimie* 81, 835–840. doi: 10.1016/s0300-9084(00)87167-9

Kaur, G., Vora, M. P., Czerwonka, C. A., Rozgaja, T. A., Grimwade, J. E., and Leonard, A. C. (2014). Building the bacterial orisome: high-affinity DnaA recognition plays a role in setting the conformation of oriC DNA. *Mol. Microbiol.* 91, 1148–1163. doi: 10.1111/mmi.12525

Kawakami, H., Keyamura, K., and Katayama, T. (2005). Formation of an ATP-DnaA-specific initiation complex requires DnaA Arginine 285, a conserved motif in the AAA+ protein family. *J. Biol. Chem.* 280, 27420–27430. doi: 10.1074/jbc.M502764200

Kowalski, D., and Eddy, M. J. (1989). The DNA unwinding element: a novel, cis-acting component that facilitates opening of the *Escherichia coli* replication origin. *EMBO J.* 8, 4335–4344. doi: 10.1002/j.1460-2075.1989.tb08620.x

Krause, M., Rückert, B., Lurz, R., and Messer, W. (1997). Complexes at the replication origin of *Bacillus subtilis* with homologous and heterologous DnaA protein. *J. Mol. Biol.* 274, 365–380. doi: 10.1006/jmbi.1997.1404

Kurokawa, K., Mizumura, H., Takaki, T., Ishii, Y., Ichihashi, N., Lee, B. L., et al. (2009). Rapid exchange of bound ADP on the *Staphylococcus aureus* replication initiation protein DnaA. *J. Biol. Chem.* 284, 34201–34210. doi: 10.1074/jbc.M109.060681

Kurokawa, K., Nishida, S., Emoto, A., Sekimizu, K., and Katayama, T. (1999). Replication cycle-coordinated change of the adenine nucleotide-bound forms of DnaA protein in *Escherichia coli*. *EMBO J.* 18, 6642–6652. doi: 10.1093/emboj/18.23.6642

Leonard, A. C., and Grimwade, J. E. (2015). The orisome: structure and function. *Front. Microbiol.* 6:545. doi: 10.3389/fmicb.2015.00545

Leonard, A. C., and Méchali, M. (2013). DNA replication origins. *Cold Spring Harb. Perspect. Biol.* 5:a010116. doi: 10.1101/cshperspect.a010116

Luo, H., and Gao, F. (2019). DoriC 10.0: an updated database of replication origins in prokaryotic genomes including chromosomes and plasmids. *Nucleic Acids Res.* 47, D74–D77. doi: 10.1093/nar/gky1014

Madiraju, M. V., Moomey, M., Neuenschwander, P. F., Muniruzzaman, S., Yamamoto, K., Grimwade, J. E., et al. (2006). The intrinsic ATPase activity of *Mycobacterium tuberculosis* DnaA promotes rapid oligomerization of DnaA on oriC. *Mol. Microbiol.* 59, 1876–1890. doi: 10.1111/j.1365-2958.2006.05068.x

Mallik, P., Paul, B. J., Rutherford, S. T., Gourse, R. L., and Osuna, R. (2006). DksA is required for growth phase-dependent regulation, growth rate-dependent control, and stringent control of fis expression in *Escherichia coli*. *J. Bacteriol.* 188, 5775–5782. doi: 10.1128/JB.00276-06

McGarry, K. C., Ryan, V. T., Grimwade, J. E., and Leonard, A. C. (2004). Two discriminatory binding sites in the *Escherichia coli* replication origin are required for DNA strand opening by initiator DnaA-ATP. *Proc. Natl. Acad. Sci. U.S.A.* 101, 2811–2816. doi: 10.1073/pnas.0400340101

Merrikh, H., and Grossman, A. D. (2011). Control of the replication initiator DnaA by an anti-cooperativity factor. *Mol. Microbiol.* 82, 434–446. doi: 10.1111/j.1365-2958.2011.07821.x

Messer, W., Blaesing, F., Majka, J., Nardmann, J., Schaper, S., Schmidt, A., et al. (1999). Functional domains of DnaA proteins. *Biochimie* 81, 819–825. doi: 10.1016/s0300-9084(00)87165-5

Miller, D. T., Grimwade, J. E., Betteridge, T., Rozgaja, T., Torgue, J. J., and Leonard, A. C. (2009). Bacterial origin recognition complexes direct assembly of higher-order DnaA oligomeric structures. *Proc. Natl. Acad. Sci. U.S.A.* 106, 18479–18484. doi: 10.1073/pnas.0909472106

Moriya, S., Fukuoka, T., Ogasawara, N., and Yoshikawa, H. (1988). Regulation of initiation of the chromosomal replication by DnaA-boxes in the origin region of the *Bacillus subtilis* chromosome. *EMBO J.* 7, 2911–2917. doi: 10.1002/j.1460-2075.1988.tb03149.x

Nielsen O., and Løbner-Olesen, A. (2008). Once in a lifetime: strategies for preventing re-replication in prokaryotic and eukaryotic cells. *EMBO Rep.* 9, 151–156. doi: 10.1038/sj.embor.2008.2

Nievera, C., Torgue, J. J., Grimwade, J. E., and Leonard, A. C. (2006). SeqA blocking of DnaA-oriC interactions ensures staged assembly of the *E. coli* pre-RC. *Mol. Cell* 24, 581–592. doi: 10.1016/j.molcel.2006.09.016

Nilsson, L., Verbeek, H., Vijgenboom, E., van Drunen, C., Vanet, A., and Bosch, L. (1992). FIS-dependent trans activation of stable RNA operons of *Escherichia coli* under various growth conditions. *J. Bacteriol.* 174, 921–929. doi: 10.1128/jb.174.3.921-929.1992

Nishida, S., Fujimitsu, K., Sekimizu, K., Ohmura, T., Ueda, T., and Katayama, T. (2002). A nucleotide switch in the *Escherichia coli* DnaA protein initiates chromosomal replication: evidnece from a mutant DnaA protein defective in regulatory ATP hydrolysis in vitro and in vivo. *J. Biol. Chem.* 277, 14986–14995. doi: 10.1074/jbc.M108303200

Noguchi, Y., Sakiyama, Y., Kawakami, H., and Katayama, T. (2015). The arg fingers of key DnaA protomers are oriented inward within the replication origin oriC and stimulate DnaA subcomplexes in the initiation complex. *J. Biol. Chem.* 290, 20295–20312. doi: 10.1074/jbc.M115.662601

Nozaki, S., and Ogawa, T. (2008). Determination of the minimum domain II size of *Escherichia coli* DnaA protein essential for cell viability. *Microbiology* 154, 3379–3384. doi: 10.1099/mic.0.2008/019745-0

Ogawa, T., Yamada, Y., Kuroda, T., Kishi, T., and Moriya, S. (2002). The datA locus predominantly contributes to the initiator titration mechanism in the control

of replication initiation in *Escherichia coli. Mol. Microbiol.* 44, 1367–1375. doi: 10.1046/j.1365-2958.2002.02969.x

Ogura, Y., Imai, Y., Ogasawara, N., and Moriya, S. (2001). Autoregulation of the dnaA-dnaN operon and effects of DnaA protein levels on replication initiation in *Bacillus subtilis. J. Bacteriol.* 183, 3833–3841. doi: 10.1128/JB.183.13.3833-3841.2001

O'Neill, E. A., and Bender, R. A. (1988). *Klebsiella pneumoniae* origin of replication (oriC) is not active in *Caulobacter crescentus, Pseudomonas putida,* and *Rhodobacter sphaeroides. J. Bacteriol.* 170, 3774–3777. doi: 10.1128/jb.170.8.3774-3777.1988

Ozaki, S., and Katayama, T. (2012). Highly organized DnaA-oriC complexes recruit the single-stranded DNA for replication initiation. *Nucleic Acids Res.* 40, 1648–1665. doi: 10.1093/nar/gkr832

Ozaki, S., Kawakami, H., Nakamura, K., Fujikawa, N., Kagawa, W., Park, S. Y., et al. (2008). A common mechanism for the ATP-DnaA-dependent formation of open complexes at the replication origin. *J. Biol. Chem.* 283, 8351–8362. doi: 10.1074/jbc.M708684200

Ozaki, S., Noguchi, Y., Hayashi, Y., Miyazaki, E., and Katayama, T. (2012). Differentiation of the DnaA-oriC subcomplex for DNA unwinding in a replication initiation complex. *J. Biol. Chem.* 287, 37458–37471. doi: 10.1074/jbc.M112.372052

Quon, K. C., Yang, B., Domian, I. J., Shapiro, L., and Marczynski, G. T. (1998). Negative control of bacterial DNA replication by a cell cycle regulatory protein that binds at the chromosome origin. *Proc. Natl. Acad. Sci. U.S.A.* 95, 120–125. doi: 10.1073/pnas.95.1.120

Rahn-Lee, L., Merrikh, H., Grossman, A. D., and Losick, R. (2011). The sporulation protein SirA inhibits the binding of DnaA to the origin of replication by contacting a patch of clustered amino acids. *J. Bacteriol.* 193, 1302–1307. doi: 10.1128/JB.01390-10

Rajagopalan, M., Dziedzic, R., Al Zayer, M., Stankowska, D., Ouimet, M. C., Bastedo, D. P., et al. (2010). *Mycobacterium tuberculosis* origin of replication and the promoter for immunodominant secreted antigen 85B are the targets of MtrA, the essential response regulator. *J. Biol. Chem.* 285, 15816–15827. doi: 10.1074/jbc.M109.040097

Rao, P., Rozgaja, T. A., Alqahtani, A., Grimwade, J. E., and Leonard, A. C. (2018). Low affinity DnaA-ATP recognition sites in *E. coli* oriC make non-equivalent and growth rate-dependent contributions to the regulated timing of chromosome replication. *Front. Microbiol.* 9:1673. doi: 10.3389/fmicb.2018.01673

Richardson, T. T., Harran, O., and Murray, H. (2016). The bacterial DnaA-trio replication origin element specifies single-stranded DNA initiator binding. *Nature* 534, 412–416. doi: 10.1038/nature17962

Richardson, T. T., Stevens, D., Pelliciari, S., Harran, O., Sperlea, T., and Murray, H. (2019). Identification of a basal system for unwinding a bacterial chromosome origin. *EMBO J.* 38:e101649. doi: 10.15252/embj.2019101649

Roggenkamp, A. (2007). Phylogenetic analysis of enteric species of the family Enterobacteriaceae using the oriC-locus. *Syst. Appl. Microbiol.* 30, 180–188. doi: 10.1016/j.syapm.2006.06.004

Roth, A., Urmoneit, B., and Messer, W. (1994). Functions of histone-like proteins in the initiation of DNA replication at oriC of *Escherichia coli. Biochimie* 76, 917–923. doi: 10.1016/0300-9084(94)90016-7

Rozgaja, T. A., Grimwade, J. E., Iqbal, M., Czerwonka, C., Vora, M., and Leonard, A. C. (2011). Two oppositely oriented arrays of low-affinity recognition sites in oriC guide progressive binding of DnaA during *Escherichia coli* pre-RC assembly. *Mol. Microbiol.* 82, 475–488. doi: 10.1111/j.1365-2958.2011.07827.x

Ryan, V. T., Grimwade, J. E., Camara, J. E., Crooke, E., and Leonard, A. C. (2004). *Escherichia coli* prereplication complex assembly is regulated by dynamic interplay among Fis, IHF and DnaA. *Mol. Microbiol.* 51, 1347–1359. doi: 10.1046/j.1365-2958.2003.03906.x

Sakiyama, Y., Kasho, K., Noguchi, Y., Kawakami, H., and Katayama, T. (2017). Regulatory dynamics in the ternary DnaA complex for initiation of chromosomal replication in *Escherichia coli. Nucleic Acids Res.* 45, 12354–12373. doi: 10.1093/nar/gkx914

Saxena, R., Fingland, N., Patil, D., Sharma, A. K., and Crooke, E. (2013). Crosstalk between DnaA protein, the initiator of *Escherichia coli* chromosomal replication, and acidic phospholipids present in bacterial membranes. *Int. J. Mol. Sci.* 14, 8517–8537. doi: 10.3390/ijms14048517

Saxena, R., Vasudevan, S., Patil, D., Ashoura, N., Grimwade, J. E., and Crooke, E. (2015). Nucleotide-induced conformational changes in *Escherichia coli* DnaA protein are required for bacterial ORC to Pre-RC conversion at the chromosomal origin. *Int. J. Mol. Sci.* 16, 27897–27911. doi: 10.3390/ijms161126064

Schaper, S., and Messer, W. (1995). Interaction of the initiator protein DnaA of *Escherichia coli* with its DNA target. *J. Biol. Chem.* 270, 17622–17626. doi: 10.1074/jbc.270.29.17622

Schaper, S., Nardmann, J., Lüder, G., Lurz, R., Speck, C., and Messer, W. (2000). Identification of the chromosomal replication origin from *Thermus thermophilus* and its interaction with the replication initiator DnaA. *J. Mol. Biol.* 299, 655–665. doi: 10.1006/jmbi.2000.3764

Scholefield, G., Errington, J., and Murray, H. (2012). Soj/ParA stalls DNA replication by inhibiting helix formation of the initiator protein DnaA. *EMBO J.* 31, 1542–1555. doi: 10.1038/emboj.2012.6

Scholefield, G., and Murray, H. (2013). YabA and DnaD inhibit helix assembly of the DNA replication initiation protein DnaA. *Mol. Microbiol.* 90, 147–159. doi: 10.1111/mmi.12353

Sekimizu, K., Bramhill, D., and Kornberg, A. (1987). ATP activates dnaA protein in initiating replication of plasmids bearing the origin of the *E. coli* chromosome. *Cell* 50, 259–265. doi: 10.1016/0092-8674(87)90221-2

Simmons, L. A., Felczak, M., and Kaguni, J. M. (2003). DnaA Protein of *Escherichia coli*: oligomerization at the *E. coli* chromosomal origin is required for initiation and involves specific N-terminal amino acids. *Mol. Microbiol.* 49, 849–858. doi: 10.1046/j.1365-2958.2003.03603.x

Skarstad, K., Boye, E., and Steen, H. B. (1986). Timing of initiation of chromosome replication in individual *Escherichia coli* cells. *EMBO J.* 5, 1711–1717. doi: 10.1002/j.1460-2075.1986.tb04415.x

Skarstad, K., and Katayama, T. (2013). Regulating DNA replication in bacteria. *Cold Spring Harb. Perspect. Biol.* 5:a012922. doi: 10.1101/cshperspect.a012922

Speck, C., and Messer, W. (2001). Mechanism of origin unwinding: sequential binding of DnaA to double- and single-stranded DNA. *EMBO J.* 20, 1469–1476. doi: 10.1093/emboj/20.6.1469

Stepankiw, N., Kaidow, A., Boye, E., and Bates, D. (2009). The right half of the *Escherichia coli* replication origin is not essential for viability, but facilitates multi-forked replication. *Mol. Microbiol.* 74, 467–479. doi: 10.1111/j.1365-2958.2009.06877.x

Sun, J., Kawakami, H., Zech, J., Speck, C., Stillman, B., and Li, H. (2012). Cdc6-induced conformational changes in ORC bound to origin DNA revealed by cryo-electron microscopy. *Structure* 20, 534–544. doi: 10.1016/j.str.2012.01.011

Takeda, Y., Harding, N. E., Smith, D. W., and Zyskind, J. W. (1982). The chromosomal origin of replication (oriC) of Erwinia carotovora. *Nucleic Acids Res.* 10, 2639–2650. doi: 10.1093/nar/10.8.2639

Taylor, J. A., Ouimet, M. C., Wargachuk, R., and Marczynski, G. T. (2011). The *Caulobacter crescentus* chromosome replication origin evolved two classes of weak DnaA binding sites. *Mol. Microbiol.* 82, 312–326. doi: 10.1111/j.1365-2958.2011.07785.x

Wargachuk, R., and Marczynski, G. T. (2015). The *Caulobacter crescentus* homolog of DnaA (HdaA) also regulates the proteolysis of the replication initiator protein DnaA. *J. Bacteriol.* 197, 3521–3532. doi: 10.1128/JB.00460-15

Wolański, M., Donczew, R., Zawilak-Pawlik, A., and Zakrzewska-Czerwińska, J. (2014). oriC-encoded instructions for the initiation of bacterial chromosome replication. *Front. Microbiol.* 5:735. doi: 10.3389/fmicb.2014.00735

Wolański, M., Jakimowicz, D., and Zakrzewska-Czerwińska, J. (2012). AdpA, key regulator for morphological differentiation regulates bacterial chromosome replication. *Open Biol.* 2:120097. doi: 10.1098/rsob.120097

Yung, B. Y., Crooke, E., and Kornberg, A. (1990). Fate of the DnaA initiator protein in replication at the origin of the *Escherichia coli* chromosome in vitro. *J. Biol. Chem.* 265, 1282–1285.

Yung, B. Y., and Kornberg, A. (1989). The dnaA initiator protein binds separate domains in the replication origin of *Escherichia coli. J. Biol. Chem.* 264, 6146–6150.

Zakrzewska-Czerwińska, J., Jakimowicz, D., Zawilak-Pawlik, A., and Messer, W. (2007). Regulation of the initiation of chromosomal replication in bacteria. *FEMS Microbiol. Rev.* 31, 378–387. doi: 10.1111/j.1574-6976.2007.00070.x

Zawilak-Pawlik, A., Kois, A., Majka, J., Jakimowicz, D., Smulczyk-Krawczyszyn, A., Messer, W., et al. (2005). Architecture of bacterial replication initiation

complexes: orisomes from four unrelated bacteria. *Biochem. J.* 389, 471–481. doi: 10.1042/BJ20050143

Zawilak-Pawlik, A., Nowaczyk, M., and Zakrzewska-Czerwińska, J. (2017). The role of the N-Terminal domains of bacterial initiator DnaA in the assembly and regulation of the bacterial replication initiation complex. *Genes* 8:136. doi: 10.3390/genes8050136

Zorman, S., Seitz, H., Sclavi, B., and Strick, T. R. (2012). Topological characterization of the DnaA-oriC complex using single-molecule nanomanipuation. *Nucleic Acids Res.* 40, 7375–7383. doi: 10.1093/nar/gks371

Zyskind, J. W., Cleary, J. M., Brusilow, W. S., Harding, N. E., and Smith, D. W. (1983). Chromosomal replication origin from the marine bacterium *Vibrio harveyi* functions in *Escherichia coli*: oriC consensus sequence. *Proc. Natl. Acad. Sci. U.S.A.* 80, 1164–1168. doi: 10.1073/pnas.80.5.1164

HipA-Mediated Phosphorylation of SeqA Does not Affect Replication Initiation in *Escherichia coli*

Leise Riber[1]*, Birgit M. Koch[1], Line Riis Kruse[1], Elsa Germain[2] and Anders Løbner-Olesen[1]*

[1] Section for Functional Genomics, Department of Biology, Center for Bacterial Stress Response and Persistence, University of Copenhagen, Copenhagen, Denmark, [2] Laboratoire de Chimie Bactérienne, Université Aix-Marseille, CNRS, Marseille, France

The SeqA protein of *Escherichia coli* is required to prevent immediate re-initiation of chromosome replication from *oriC*. The SeqA protein is phosphorylated at the serine-36 (Ser36) residue by the HipA kinase. The role of phosphorylation was addressed by mutating the Ser36 residue to alanine, which cannot be phosphorylated and to aspartic acid, which mimics a phosphorylated serine residue. Both mutant strains were similar to wild-type with respect to origin concentration and initiation synchrony. The minimal time between successive initiations was also unchanged. We therefore suggest that SeqA phosphorylation at the Ser36 residue is silent, at least with respect to SeqA's role in replication initiation.

Keywords: *E. coli*, SeqA protein, phosphorylation, HipA kinase, initiation synchrony, minimal inter-initiation time

INTRODUCTION

In *Escherichia coli* the DnaA initiator protein binds ATP and ADP with equal affinity (Sekimizu et al., 1987). DnaA binds three high-affinity sites in the origin, *oriC*, throughout the cell cycle irrespective of the bound nucleotide. The relative amounts of DnaA^{ATP} and DnaA^{ADP}, respectively fluctuate during the cell cycle with the DnaA^{ATP}/DnaA^{ADP} ratio peaking at initiation (Kurokawa et al., 1999). This results in binding of a number of additional DnaA binding sites of low affinity and with a preference for DnaA^{ATP} (Skarstad and Katayama, 2013; Leonard and Grimwade, 2015; Katayama et al., 2017). This induces origin opening, allows for helicase loading and replisome assembly.

Immediate re-initiation of new and hemimethylated origins is prevented by SeqA-binding to 11 GATC sites located within the minimal *oriC* (Campbell and Kleckner, 1990; Lu et al., 1994; Boye et al., 2000). The binding of SeqA to the origin prolongs the duration of the DNA hemi-methylated phase; a process called sequestration. Sequestration lasts approximately one-third of a cell cycle where re-initiation is prevented by SeqA denying DnaA^{ATP} access to GATC-containing low affinity DnaA boxes in *oriC* (Nievera et al., 2006). The sequestration period allows the cells to distinguish between "old" and "new" origins, and provides a time window where the DnaA^{ATP} level is lowered by RIDA (Kato and Katayama, 2001) and DDAH (Kasho and Katayama, 2013). Sequestration is finally terminated when GATC sequences within *oriC* become fully methylated by Dam methyltransferase.

In *seqA* mutant cells the sequestration period is shortened or absent (von Freiesleben et al., 2000), re-initiations occur frequently leading to over-initiation, and replication initiation becomes highly asynchronous (Lu et al., 1994). Conversely, excess SeqA protein prolongs the sequestration

period, delays initiation, but does not affect initiation synchrony (Fossum et al., 2003; Charbon et al., 2011).

The SeqA protein contains two functional domains, an N-terminal oligomerization domain (SeqA-N; residues 1–33) and a C-terminal DNA-binding domain (SeqA-C; residues 65–181), which are joined by a flexible linker (residues 34–64; Chung et al., 2009). The interaction of SeqA with DNA occurs mainly in the major groove of the hemimethylated GATC sequences (Guarné et al., 2002), and data have suggested that two adjacent GATC sequences, up to 31 bp apart, interacting with the SeqA dimer are sufficient for strong binding (Guarné et al., 2005).

Recently, a stable isotope labeling by amino acids in cell culture (SILAC)-based quantitative phosphoproteomic approach combined with high-resolution mass spectrometry identified residue serine-36 (Ser36) in SeqA as a direct phosphorylation target for the kinase activity of the high persister protein A, HipA (Semanjski et al., 2018). HipA is an eukaryotic-like serine-threonine protein kinase that induces the stringent response, inhibits cell growth and confers cellular persistence through phosphorylation and inactivation of the glutamyl-tRNA-synthetase, GltX (Germain et al., 2013; Kaspy et al., 2013; Semanjski et al., 2018). The *hipA* gene constitutes a type II TA module with the adjacent upstream *hipB* gene, encoding the HipB antitoxin. HipB interacts directly with HipA to form a protein complex that represses the *hipBA* operon through binding to operators in the *hipBA* promoter region (Black et al., 1994), thereby counteracting the negative effect on cell growth caused by even low amounts of wild-type HipA (Korch and Hill, 2006).

It is not known whether phosphorylation at residue Ser36 of SeqA affects the activity and function of SeqA. Adding a phosphate group with negative charge to a protein, can promote changes in the structural conformation by altering the interactions with nearby amino acids. This might activate or inhibit the activity of the protein (Chao et al., 2014) or result in function modifications (Johnson and Barford, 1993).

Here, we tested the effect of Ser36 phosphorylation of SeqA on chromosome replication initiation. Two variants of SeqA were constructed, in which the Ser36 residue was either mutated to alanine (S36A) or aspartic acid (S36D). The S36A mutation impairs Ser36 phosphorylation, whereas the S36D mutation mimics the conformation of Ser36 phosphorylated SeqA (i.e., phospho-mimetic; (Arany et al., 2013). As both *seqA* mutants were similar to the wild-type with respect to synchrony and length of the sequestration period, our data suggest that HipA-mediated Ser36 phosphorylation of SeqA constitutes a neutral effect on the role of SeqA in *E. coli* replication initiation.

MATERIALS AND METHODS
Media and Growth Conditions
Cells were grown in AB minimal medium (Clark and Maaløe, 1967) supplemented with 1 μg/ml thiamine, 0.2% glucose and 0.5% casamino acids (glucose-CAA medium). When necessary, antibiotic selection was maintained at the following final concentrations: kanamycin, 50 μg/ml; chloramphenicol, 20 μg/ml; tetracycline, 10 μg/ml; ampicillin, 150 μg/ml. All cells were cultured at 37°C, except when otherwise indicated. Cell

growth was monitored by measuring optical density at 450 nm (OD_{450}).

Bacterial Strains
All strains used were derived from *E. coli* K-12 MG1655 (F−, λ−, *rph-1*; Guyer et al., 1981) and are listed in **Table 1**. The Δ*hipBA::frt::kan::frt* (Germain et al., 2013) and *dnaA46 tnaA600::Tn10* (Kogoma and von Meyenburg, 1983) alleles were moved by P1-phage-mediated transduction (Miller, 1972). To construct the chromosomal *seqA* mutant strains ($seqA_{S36A}$ and $seqA_{S36D}$, respectively), base substitutions were made in the codon for Ser36 (5′-TCC-3′ to 5′-**G**CC-3′; Ser(S) to Ala(A), and 5′-TCC-3′ to 5′-**GA**C-3′; Ser(S) to Asp(D), respectively) using splicing by overlap extension (SOEing) polymerase chain reaction (PCR) (Horton et al., 1989). All primers are listed in **Table 2**. For each *seqA* variant two initial PCR products of the MG1655 chromosome were generated. 1) The lower region of the *seqA* gene, spanning residues 27–181, was amplified using primers "SeqA_down_bw_XmaI" and either "SeqA_pos36_SA_fw" or "SeqA_pos36_SD_fw". 2) The upper region of the *seqA* gene was amplified using primers "SeqA_up_fw_SacI" and "SeqA_intern_bw" that generates a fragment with an overlap of 21 bp with the *seqA* downstream PCR product. A secondary amplification was performed using equimolar ratios of the two PCR products as template, and the oligonucleotides, SeqA_down_bw_XmaI, and SeqA_up_fw_SacI, as primers. The resulting PCR fragments were digested with XmaI and SacI, and cloned into the same sites of the 3.9 kb suicide vector, pRUC1437, a derivative of pSW29T (Demarre et al., 2005), carrying the *aph* gene encoding kanaymicn resistance, and the *sacB* gene. The resulting plasmids

TABLE 1 | Bacterial strains.

Strains	Relevant genotype	Plasmid	Reference/Source
MG1655	F−, λ−, *rph-1*	None	(Guyer et al., 1981)
ALO2956	*lacIZYA::cat*[a]	None	This work
ALO3758	Δ*seqA*[b]	None	(Riber et al., 2009)
ALO5105	Δ*hipBA::frt::kan::frt*[b]	None	(Germain et al., 2013)/This work
ALO5695	$seqA_{S36D}$[b]	None	This work
ALO5945	$seqA_{S36A}$[b]	None	This work
ALO6136	*dnaA46 tnaA600*::Tn*10*[b]	None	This work
ALO6138	*dnaA46 tnaA600*::Tn*10* Δ*seqA*[b]	None	This work
ALO6141	*dnaA46 tnaA600*::Tn*10* $seqA_{S36A}$[b]	None	This work
ALO6143	*dnaA46 tnaA600*::Tn*10* $seqA_{S36D}$[b]	None	This work
ALO5090	Δ*seqA*[b]	pFH2102	This work
ALO5093	Δ*seqA*[b]	pMAK7	This work
ALO5095	Δ*seqA*[b]	pLR75	This work
ALO5101	Δ*seqA*[b]	pLR77	This work

[a]*Genotype otherwise as MG1655.*
[b]*Genotype otherwise as ALO2956.*

Name	Sequence
SeqA_down_bw_XmaI	5′-GGCGGCCCCGGGTTTGTCCTTTGTCTGCAACG
SeqA_up_fw_SacI	5′-GGCGGCGAGCTCCAGCTAAGACACTGCACTGG
SeqA_intern_bw	5′-CAACATACGCCGTAAAATGTC
SeqA_pos36_SA_fw	5′-GACATTTTACGGCGTATGTTGAAATTTGCCGCCGCA TCACAGCCTGCTGCTCCG
SeqA_pos36_SD_fw	5′-GACATTTTACGGCGTATGTTGAAATTTGACGCCGCA TCACAGCCTGCTGCTCCG
SeqA_chr_fw	5′-CCATTGTGCCACAGGGCTGCAAC
SeqA_chr_bw	5′-GCACTGCCACGGTGACCGGAAG
SeqA_up_fw_EcoRI	5′-GGCGGCGAATTCCAGCTAAGACACTGCACTGG
SeqA_down_bw_HindIII	5′-GGCGGCAAGCTTTTTGTCCTTTGTCTGCAACG

Plasmid	Relevant genotype	Reference/Source
pBR322	*bla, tet*	(Bolivar et al., 1977)
pFH2102	ori-pBR322, *lacP*$_{A1/04-03}$, *lacI, bla*	(von Freiesleben et al., 2000)
pMAK7	ori-pBR322, *lacP*$_{A1/04-03}$-*seqA*, *lacI, bla*	(von Freiesleben et al., 2000)
pLR75	ori-pBR322, *lacP*$_{A1/04-03}$-*seqA*$_{S36D}$, *lacI, bla*	This work
pLR77	ori-pBR322, *lacP*$_{A1/04-03}$-*seqA*$_{S36A}$, *lacI, bla*	This work

were transformed into strain S17-1 (*recA thi pro hsdR*$^-$*M*$^+$ RP4-2 Tc::Mu-Km::Tn7 λ-*pir* lysogen Tp^R Sm^R; Simon et al., 1983) before being transferred into ALO 2956 cells by conjugation. Selection of exconjugants carrying the chromosomally integrated recombinant suicide plasmids as well as subsequent sucrose-mediated selection for loss of the *sacB* gene (i.e., loss of suicide vector sequences; Donnenberg and Kaper, 1991), leaving either a wild-type or a mutant variant of the *seqA* gene on the MG1655 *lacIZYA::cat* chromosome, was performed as described previously (Riber et al., 2009). Chromosomal *seqA* mutant strains were verified by DNA sequencing of PCR fragments amplified from the *seqA* region using DNA oligonucleotides, "SeqA_chr_fw" and "SeqA_chr_bw," as primers.

Plasmids

All plasmids used are listed in **Table 3**. Plasmids pLR77 and pLR75 were constructed by PCR amplifying the *seqA* variant genes (including the native *seqA* ribosome binding site) from MG1655 *lacIZYA::cat* cells carrying either the *seqA*$_{S36A}$ or *seqA*$_{S36D}$ chromosomal genes (see above), respectively, using DNA oligonucleotides, SeqA_up_fw_EcoRI and SeqA_down_bw_HindIII, as primers. The resultant PCR fragments were digested with EcoRI and HindIII and inserted downstream the IPTG inducible *lacP*$_{A1-04/03}$ promoter (Lanzer and Bujard, 1988) of plasmid pFH2102 (von Freiesleben et al., 2000), cut with the same enzymes. The inserted *seqA* mutant genes were later verified by DNA sequencing.

Flow Cytometry and Cell Cycle Analysis

Exponentially growing cells (OD$_{450}$ = 0.15–0.30) were treated with rifampicin (300 μg/ml; SERVA Electrophoresis GmbH) and cephalexin (36 μg/ml; Sigma-Aldrich) to inhibit initiation of DNA replication and cell division, respectively (Løbner-Olesen et al., 1989). Incubation continued for a minimum of 4 h at 37°C to allow completion of ongoing rounds of replication. Cells were fixed in 70% ethanol and stained with 90 μg/ml mithramycin (SERVA Electrophoresis GmbH) and 20 μg/ml ethidium bromide (Sigma-Aldrich) as described (Løbner-Olesen et al., 1989). Flow cytometry was performed as previously

described (Løbner-Olesen et al., 1989) using an Apogee A10 instrument (Apogee, Inc.). For all samples a minimum of 50.000 cells were analyzed. Numbers of origins per cell and relative cell mass were determined as previously described (Løbner-Olesen et al., 1989).

Immunoblot Procedure

Samples of 2 ml of exponentially growing cells (OD$_{450}$ = 0.3–0.4) were harvested. Proteins were separated by SDS-PAGE and SeqA protein detected by Western blot using rabbit antiserum raised against SeqA protein (Torheim et al., 2000) as previously described (Riber and Lobner-Olesen, 2005). The membrane was scanned using a 230 V GenoView imaging system equipped with a UV transilluminator (VWR). Quantification was done using the ImageJ software.

Multiple Sequence Alignment Analysis

Multiple alignment analysis of SeqA amino acids sequences was performed in the MEGA version 7.0.26 software (Kumar et al., 2016) using the default settings of the integrated ClustalW algorithm (Larkin et al., 2007). Selected species including SeqA protein accession numbers were: *Escherichia coli* K-12 (accession: AAA19855.1), *Vibrio cholerae* (accession: AOY47782.1), *Pasteurella multocida* PM70 (accession: AAK02440.1), *Haemophilus influenzae* Rd (accession: NP_438362.1), *Yersinia enterocolitica* (accession: CNB62546.1), *Serratia marcescens* (accession: KFL03527.1), *Actinobacillus pleuropneumoniae* (accession: SQF64393.1), and *Glaesserella parasuis* (accession: STO80764.1).

RESULTS

Changing the SeqA Ser36 Residue Mainly Affects the Linker Region of SeqA

In order to determine any putative role of SeqA phosphorylation, we generated two mutations at the chromosomal codon 36 of *seqA*. In one strain, the codon for Ser36 was replaced with that of an aspartic acid (*seqA*$_{S36D}$). The SeqA$_{S36D}$ mimics the conformation of Ser36 phosphorylated SeqA (Arany et al., 2013). In a second strain the codon for Ser36 was replaced with that of an alanine (*seqA*$_{S36A}$). The resulting protein, SeqA$_{S36A}$, is phosphorylation impaired at position 36 (Arany et al., 2013).

We used the RaptorX web server (Källberg et al., 2012) to predict the tertiary structures (Peng and Xu, 2011; Ma et al., 2012) of the wild-type SeqA and the SeqA$_{S36A}$ and SeqA$_{S36D}$ proteins. This revealed a significant level of resemblance (**Figure 1A**). By pairwise and multiple structural alignments of the SeqA protein variants, TMScore values above 0.9 were obtained, illustrating a significantly increased likelihood (>90% of chance) that the proteins pairwise and all together share similar folds, RaptorX Structure Alignment Server; (Wang et al., 2011, 2013), with SeqA and SeqA$_{S36D}$ being structurally most alike [TMScore (WT vs. S36D) = 0.96]. The structural differences caused by changing the Ser36 residue seem to affect only the flexible linker region between SeqA-N and SeqA-C (**Figure 1A**; Chung et al., 2009).

Replication Initiation Is Not Affected by *seqA$_{S36A}$* and *seqA$_{S36D}$* Mutations

We used flow-cytometry to determine cell cycle parameters of wild-type and *seqA* mutants. The two *seqA* mutants grew with similar doubling times as wild-type cells in minimal medium supplemented with glucose and casamino acids, whereas cells deficient in SeqA grew with an ∼30% increased doubling time relative to that of wild-type cells (**Figure 1B**). Following treatment with rifampicin and cephalexin, wild-type, *seqA$_{S36A}$* and *seqA$_{S36D}$* cells were similar and contained mainly 2, 4, or 8 fully replicated chromosomes, indicative of initiation synchrony (Skarstad et al., 1986). As the average cell mass and numbers of origins per cell were similar, so was the origin concentration
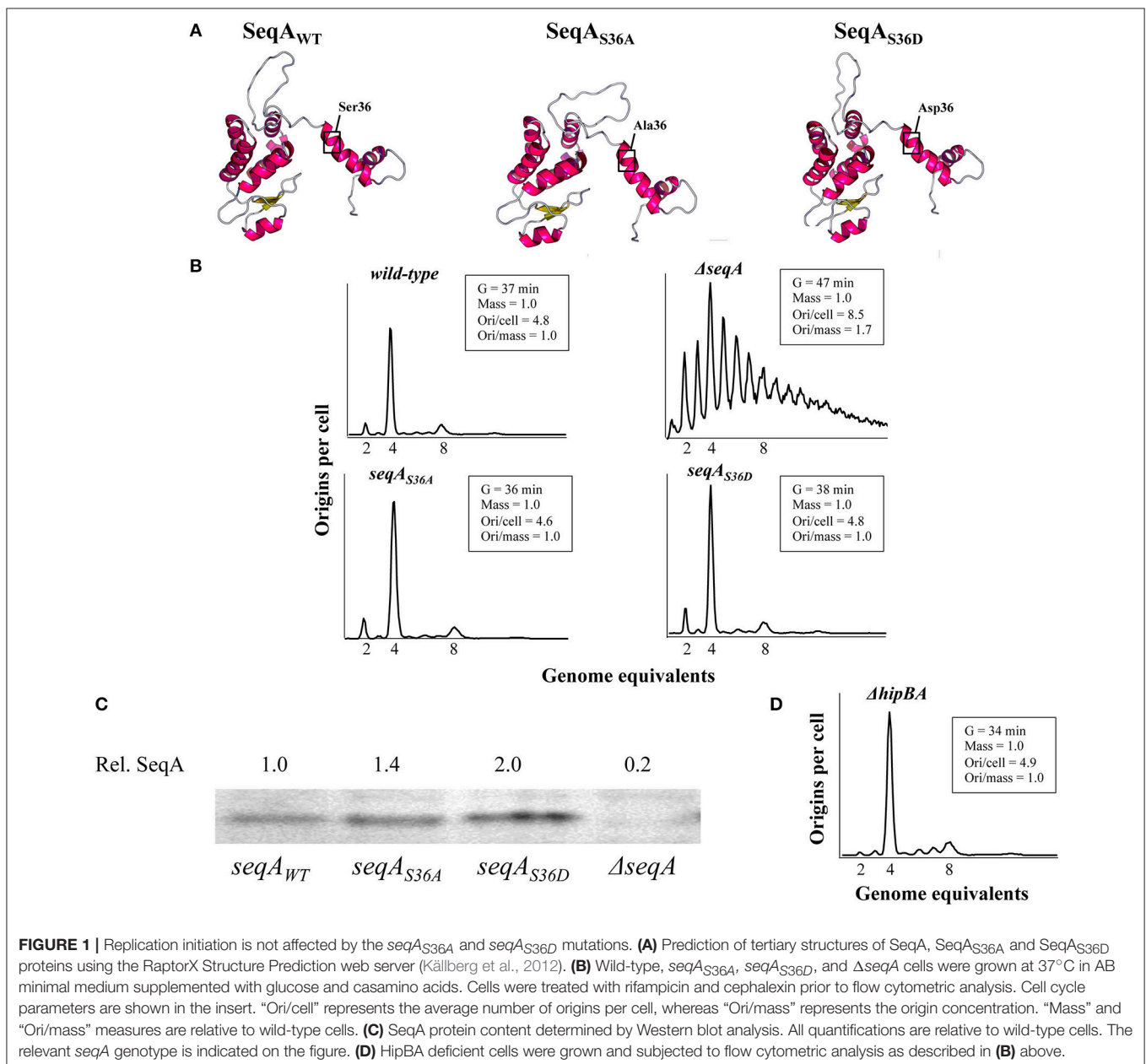
FIGURE 1 | Replication initiation is not affected by the *seqA$_{S36A}$* and *seqA$_{S36D}$* mutations. **(A)** Prediction of tertiary structures of SeqA, SeqA$_{S36A}$ and SeqA$_{S36D}$ proteins using the RaptorX Structure Prediction web server (Källberg et al., 2012). **(B)** Wild-type, *seqA$_{S36A}$*, *seqA$_{S36D}$*, and Δ*seqA* cells were grown at 37°C in AB minimal medium supplemented with glucose and casamino acids. Cells were treated with rifampicin and cephalexin prior to flow cytometric analysis. Cell cycle parameters are shown in the insert. "Ori/cell" represents the average number of origins per cell, whereas "Ori/mass" represents the origin concentration. "Mass" and "Ori/mass" measures are relative to wild-type cells. **(C)** SeqA protein content determined by Western blot analysis. All quantifications are relative to wild-type cells. The relevant *seqA* genotype is indicated on the figure. **(D)** HipBA deficient cells were grown and subjected to flow cytometric analysis as described in **(B)** above.

between these three cell types. SeqA deficient cells showed an asynchronous initiation phenotype with an increased average number of origins, which illustrates a lost ability to negatively regulate replication initiation. The average cell mass was similar to that of wild-type cells resulting in an increased origin concentration (**Figure 1B**).
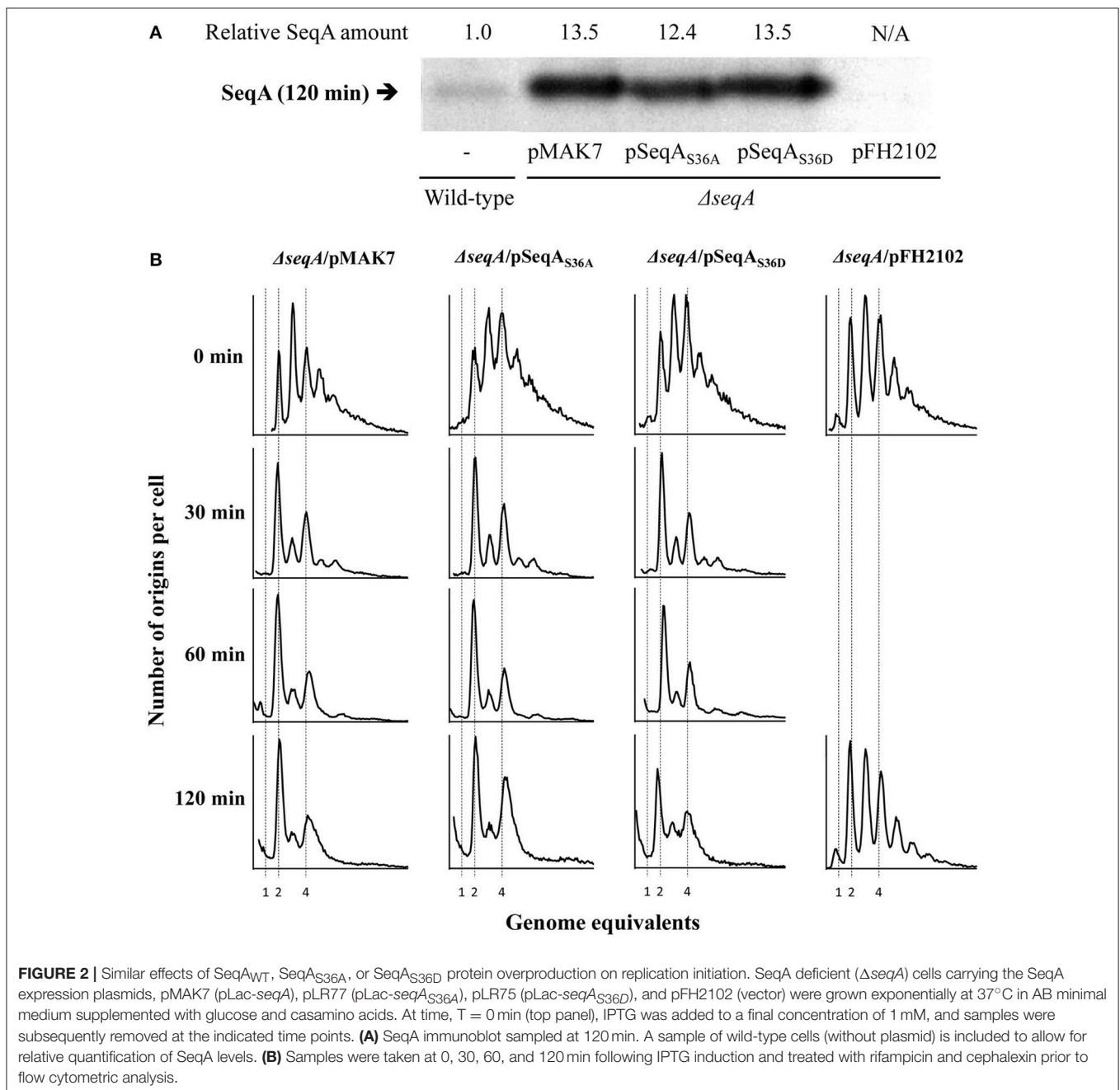
Because the SeqA$_{S36A}$ and SeqA$_{S36D}$ protein levels were comparable or slightly elevated relative to that of wild-type SeqA protein (**Figure 1C**), these data altogether suggest that phosphorylation of SeqA at position 36 has little influence on its activity in replication initiation control. This was further corroborated by analyzing cells deficient in the HipA kinase, i.e., with a knock-out of the *hipA* gene. Here we found that Δ*hipBA::kan* mutant cells displayed similar cell cycle parameters as wild-type cells (**Figure 1D**).

Overproduction of SeqA, SeqA$_{S36A}$, or SeqA$_{S36D}$ Proteins All Restore Initiation Synchrony in Δ*seqA* Mutant Cells

We proceeded to examine whether overexpression of wild-type and mutant SeqA proteins could reveal any difference in activity

FIGURE 2 | Similar effects of SeqA$_{WT}$, SeqA$_{S36A}$, or SeqA$_{S36D}$ protein overproduction on replication initiation. SeqA deficient (Δ*seqA*) cells carrying the SeqA expression plasmids, pMAK7 (pLac-*seqA*), pLR77 (pLac-*seqA*$_{S36A}$), pLR75 (pLac-*seqA*$_{S36D}$), and pFH2102 (vector) were grown exponentially at 37°C in AB minimal medium supplemented with glucose and casamino acids. At time, T = 0 min (top panel), IPTG was added to a final concentration of 1 mM, and samples were subsequently removed at the indicated time points. **(A)** SeqA immunoblot sampled at 120 min. A sample of wild-type cells (without plasmid) is included to allow for relative quantification of SeqA levels. **(B)** Samples were taken at 0, 30, 60, and 120 min following IPTG induction and treated with rifampicin and cephalexin prior to flow cytometric analysis.

among the phospho-impaired (S36A), phospho-mimetic (S36D) and wild-type SeqA proteins.

We expressed the *seqA*, *seqA_{S36A}*, and *seqA_{S36D}* genes from the IPTG-inducible *lacP_{A1/04−03}* promoter in SeqA deficient cells. Exponentially growing cells were induced with 1 mM IPTG at time 0 min (T = 0 min). Immunoblot analysis of cells sampled at 120 min following the addition of IPTG indicated that all SeqA proteins were expressed to comparable levels corresponding to an ∼12- to 14-fold increase in SeqA level relative to wild-type cells (**Figure 2A**). Both wild-type and mutant SeqA proteins complemented Δ*seqA* cells to the same extent when produced from a plasmid. Cells containing mainly two or four origins, indicative of initiation synchrony, dominated the population already after 30 min induction of the *seqA* variant genes (**Figure 2B**). A larger increase in mutant SeqA proteins (T = 120 min) resulted in no significant asynchrony relative to wild-type (**Figure 2B**). This is in agreement with earlier data on SeqA overproduction (Fossum et al., 2003).

The Minimal Time Between Successive Initiations Is Not Altered by *seqA_{S36A}* and *seqA_{S36D}* Mutations

Changes in the duration of sequestration by increasing or decreasing the level of Dam methylase (von Freiesleben et al., 2000) or by increasing the SeqA level (Charbon et al., 2011) were previously found to have relatively modest effects on the cell cycle relative to complete loss of sequestration. We therefore proceeded to determine whether the SeqA mutant proteins affected the length of the sequestration period, defined as the minimal time between successive initiations (von Freiesleben et al., 2000).

We introduced the *dnaA46* allele into *seqA_{S36A}* and *seqA_{S36D}* cells by P1-transduction. The resultant strains are initiation proficient at 30°C (permissive temperature), but not at 42°C (non-permissive temperature) due to a reversible defect in nucleotide binding (Carr and Kaguni, 1996). Wild-type, *seqA_{S36A}* and *seqA_{S36D}* cells carrying the *dnaA46* allele were grown exponentially at 30°C. The average number of origins per cell for all three strains was close to 2 (**Figures 3A–C**) and the SeqA proteins were produced in similar amounts (**Figure 3D**). When cells were shifted to 42°C, initiations ceased whereas cells continued to grow and divide, resulting in most cells ending up having one fully replicated chromosome after 90 min (**Figures 3A–C**). Upon a shift back to 30°C, where the DnaA46 protein was reactivated, all cells initiated replication, i.e., doubled their origin content, within a short period of time. This round of initiation was followed by a period of ∼20 min where all newly formed origins were inert to further initiation, after which replication initiation resumed (**Figures 3A–C**). This 20-min period represents the minimal time between successive initiations (von Freiesleben et al., 2000), and it did not differ between wild-type and *seqA* mutant cells (**Figures 3A–C**). SeqA deficient cells were previously shown to reinitiate frequently without this 20-min delay (von Freiesleben et al., 2000).

Serine36 of SeqA Is Not phylogenetically Conserved

We aligned SeqA amino acid sequences from the *Vibrio cholerae*, *Pasteurella multocida* PM70, *Haemophilus influenzae*
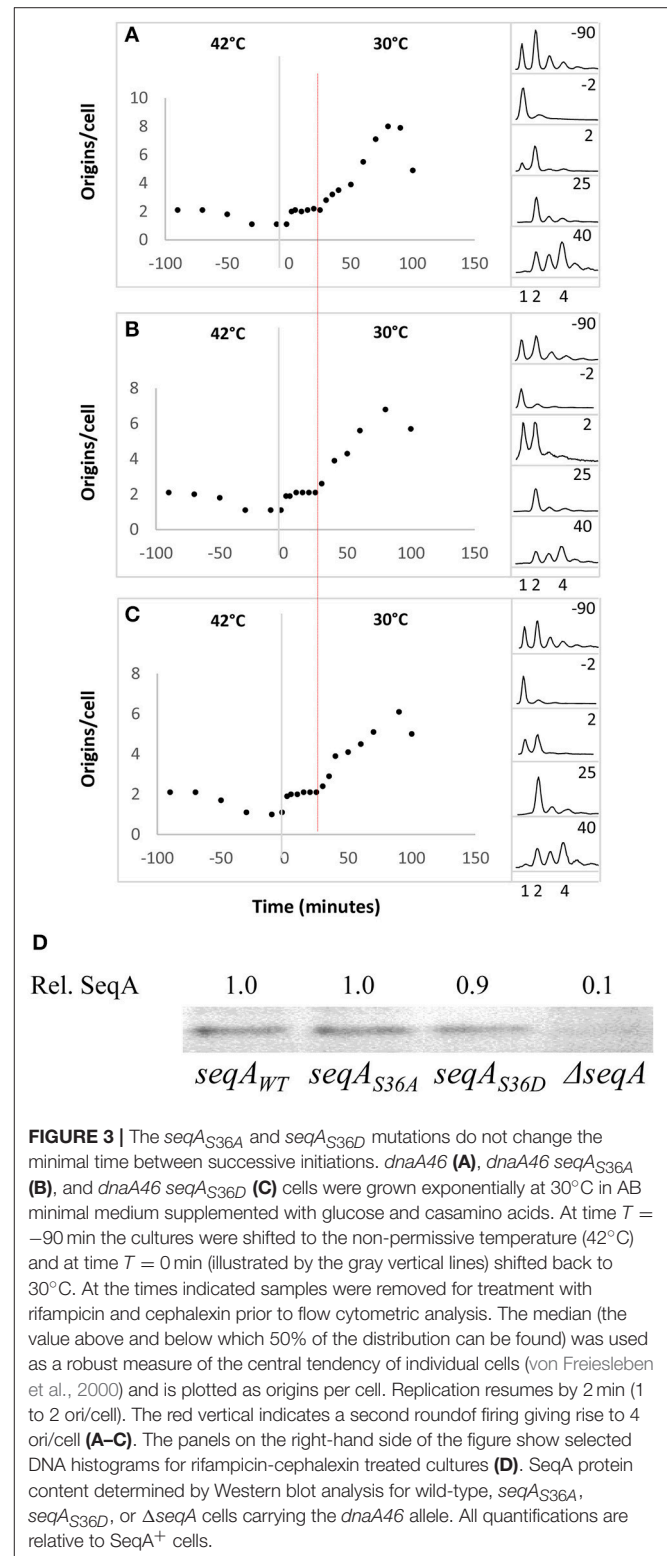

FIGURE 3 | The *seqA_{S36A}* and *seqA_{S36D}* mutations do not change the minimal time between successive initiations. *dnaA46* **(A)**, *dnaA46 seqA_{S36A}* **(B)**, and *dnaA46 seqA_{S36D}* **(C)** cells were grown exponentially at 30°C in AB minimal medium supplemented with glucose and casamino acids. At time *T* = −90 min the cultures were shifted to the non-permissive temperature (42°C) and at time *T* = 0 min (illustrated by the gray vertical lines) shifted back to 30°C. At the times indicated samples were removed for treatment with rifampicin and cephalexin prior to flow cytometric analysis. The median (the value above and below which 50% of the distribution can be found) was used as a robust measure of the central tendency of individual cells (von Freiesleben et al., 2000) and is plotted as origins per cell. Replication resumes by 2 min (1 to 2 ori/cell). The red vertical indicates a second roundof firing giving rise to 4 ori/cell **(A–C)**. The panels on the right-hand side of the figure show selected DNA histograms for rifampicin-cephalexin treated cultures **(D)**. SeqA protein content determined by Western blot analysis for wild-type, *seqA_{S36A}*, *seqA_{S36D}*, or Δ*seqA* cells carrying the *dnaA46* allele. All quantifications are relative to SeqA+ cells.

TABLE 4 | Serine 36 of *E. coli* SeqA is not conserved between bacterial species.

Escherichia coli K12	Thr18	Ile21	Ala25	Phe35	Ser36	Ala37	Arg116	Thr117	Arg118	Asn150	Asn152
Vibrio cholerae	+	+	+	−	−	−	+	−	+	+	+
Pasteurella multocida PM70	+	+	+	+	+	−	+	+	+	+	+
Haemophilus influenzae Rd	+	+	+	−	−	−	+	+	+	+	+
Yersinia enterocolitica	+	+	+	+	−	+	+	+	+	+	+
Serratia marcescens	+	+	+	+	−	+	+	+	+	+	+
Actinobacillus pleuropneumoniae	+	+	+	−	−	−	−	−	+	+	+
Glaesserella parasuis	+	+	+	−	−	+	−	−	+	+	+

The amino acid sequences of the *E. coli* K12, *Vibrio cholera, Pasteurella multocida* PM70, *Haemophilus influenzae* Rd, *Yersinia enterocolitica, Serratia mercescens, Actinobacillus pleuropneumoniae* and *Glaesserella parasuis* SeqA proteins were aligned. + and − indicate presence or absence of the indicated amino acid, respectively.

Rd, *Yersinia enterocolitica, Serratia marcescens, Actinobacillus pleuropneumoniae* and *Glaesserella parasuis* with that of *E. coli* K12. All of these bacteria are known to carry *hipBA* genes. We looked for conservation of Ser36 along with the two flanking amino acids Phe35 and Ala37 (**Table 4**). None of these amino acids were conserved among the species with Ser36 showing the least degree of conservation. On the other hand Thr18, Ile21, and Ala25 which are instrumental in oligomerization of SeqA (Guarné et al., 2005), were completely conserved. For Arg116, Thr117, Arg118, Asn150, and Asn152 that make contact with the GATC sequence in DNA (Fujikawa et al., 2004) we also observed a high degree of conservation between species (**Table 4**). This may indicate a limited role of Ser36 for SeqA function.

DISCUSSION

Recently, it was shown that residue Ser36 in the SeqA protein is a target for phosphorylation by the serine-threonine kinase, HipA (Semanjski et al., 2018). HipA is mostly known for its role in bacterial persister formation through phosphorylation of a conserved serine, Ser239, residue in the GltX aminoacyl-tRNA synthetase, which inactivates the enzyme to arrest cell growth (Germain et al., 2013; Kaspy et al., 2013). Here, we wanted to determine whether Ser36 phosphorylation could alter SeqA activity. It was tempting to speculate that the Ser36 phosphorylation would activate SeqA, thereby enhancing its inhibition of replication initiation, which would contribute to shut down chromosomal replication in persister cells. SeqA was found to be endogenous phosphorylated in wild-type *E. coli* cells, and was revealed as a direct phosphorylation target of HipA *in vitro*. When the *hipA* gene was expressed from a p15A based plasmid, the fraction of wild-type SeqA found to be phosphorylated at residue Ser36 was ∼7% following 95 min induction (Semanjski et al., 2018). It could be argued that this is a relative small fraction of the total SeqA protein. However, one should be aware that the actual phosphorylation status of SeqA may depend on the specific conditions provided. In the Semanjski study HipA expression was countered by the antitoxin HipB produced from the chromosome. The fraction of phosphorylated SeqA may therefore not reflect the fraction of SeqA being phosphorylated

during an actual stress-induced situation where HipA becomes fully induced without HipB-mediated neutralization, and where the overall protein synthesis is affected. Also, it remains unknown whether all SeqA molecules present in the cell are actually available to HipA-mediated phosphorylation. The oligomerization domain of SeqA (residue 1–33) is located close to the HipA phosphorylation domain at residue Ser36 (see below), and hence it is not clear whether SeqA oligomers are available to phosphorylation, or whether only SeqA monomers become phosphorylated.

The Ser36 residue is located in the flexible linker between the N-terminal oligomerization domain and the C-terminal DNA binding domain (Chung et al., 2009). Neither of the phospho-impaired (S36A) nor the phospho-mimetic (S36D) SeqA proteins have any change in linker length nor are they affected in prolin or other amino acid residues suggested as most preferred in linker regions (George and Heringa, 2002), suggesting that changes in flexibility and hydrophobicity are non-significant upon phosphorylation of Ser36. This agrees well with the tertiary structural predictions of the SeqA, SeqA$_{S36A}$, and SeqA$_{S36D}$ proteins that indicated the mutations to cause minor structural changes to the linker region only, leaving the N- and C-terminal domains unaffected. This might explain our observations that function and activity of the SeqA mutant proteins seemed unaffected by the Ser36 mutations with respect to replication initiation control.

Although we have assumed that substituting a serine residue with a negatively charged amino acid, such as aspartic acid, imparts the negative charge associated with serine phosphorylation, caution should be taken as this is not always the case. The phospho-mimetic proteins may fail to recapitulate the true steric and charge-based nature of phosphorylation (Paleologou et al., 2008). Also, the "phosphorylation status" mimicked by phospho-mimetics is non-reversible, and hence cannot reflect the true state of phosphorylation-mediated protein modification. Therefore, the SeqA$_{S36D}$ protein may deviate in activity from the phosphorylated wild-type SeqA protein.

However, because removal of the HipA kinase in wild-type cells revealed no replication phenotype, we find it unlikely that HipA-mediated Ser36 phosphorylation affects the activity of SeqA, at least with respect to its function in replication

initiation, and at least under the conditions provided in this study. SeqA phosphorylation may therefore be an example of a silent phosphorylation. This has previously been observed for pepsin and ovalbumin, where serine phosphorylation did not affect protein activity, and the function of the phosphate group remained unknown (Johnson and Barford, 1993). The proposal that SeqA phosphorylation is silent is reinforced by the low degree of Ser36 conservation between *hipBA* carrying bacterial species compared to highly conserved amino acids crucial for oligomerization and DNA binding activity.

AUTHOR CONTRIBUTIONS

LR, EG, and AL-O planned the experiments. LR, BK, and LK performed the experiments. LR, BK, LK, and AL-O analyzed data. LR, BK, and AL-O wrote the manuscript.

REFERENCES

Arany, I., Clark, J., Reed, D. K., and Juncos, L. A. (2013). Chronic nicotine exposure augments renal oxidative stress and injury through transcriptional activation of p66shc. *Nephrol. Dial Trans.* 28, 1417–1425. doi: 10.1093/ndt/gfs596

Black, D. S., Irwin, B., and Moyed, H. S. (1994). Autoregulation of hip, an operon that affects lethality due to inhibition of peptidoglycan or DNA synthesis. *J. Bacteriol.* 176, 4081–4091.

Bolivar, F., Rodriguez, R. L., Greene, P. J., Betlach, M. C., Heyneker, H. L., Boyer, H. W., et al. (1977). Construction and characterization of new cloning vehicles. II. a multipurpose cloning system. *Gene* 2, 95–113.

Boye, E., Lobner-Olesen, A., and Skarstad, K. (2000). Limiting DNA replication to once and only once. *EMBO Rep.* 1, 479–483. doi: 10.1093/embo-reports/kvd116

Campbell, J. L., and Kleckner, N. (1990). *E. coli* oriC and the dnaA gene promoter are sequestered from dam methyltransferase following the passage of the chromosomal replication fork. *Cell* 62, 967–979.

Carr, K. M., and Kaguni, J. M. (1996). The A184V missense mutation of the dnaA5 and dnaA46 alleles confers a defect in ATP binding and thermolability in initiation of Escherichia coli DNA replication. *Mol. Microbiol.* 20, 1307–1318.

Chao, J. D., Wong, D., and Av-Gay, Y. (2014). Microbial protein-tyrosine kinases. *J. Biol. Chem.* 289, 9463–9472. doi: 10.1074/jbc.R113.520015

Charbon, G., Riber, L., Cohen, M., Skovgaard, O., Fujimitsu, K., Katayama, T., et al. (2011). Suppressors of DnaA(ATP) imposed overinitiation in *Escherichia coli*. *Mol. Microbiol.* 79, 914–928. doi: 10.1111/j.1365-2958.2010.07493.x

Chung, Y. S., Brendler, T., Austin, S., and Guarne, A. (2009). Structural insights into the cooperative binding of SeqA to a tandem GATC repeat. *Nucleic Acids Res.* 37, 3143–3152. doi: 10.1093/nar/gkp151

Clark, D. J., and Maaløe, O. (1967). DNA replication and the division cycle in *Escherichia coli*. *J. Mol. Biol.* 23, 99–112.

Demarre, G., Guerout, A. M., Matsumoto-Mashimo, C., Rowe-Magnus, D. A., Marliere, P., and Mazel, D. (2005). A new family of mobilizable suicide plasmids based on broad host range R388 plasmid (IncW) and RP4 plasmid (IncPalpha) conjugative machineries and their cognate *Escherichia coli* host strains. *Res. Microbiol.* 156, 245–255. doi: 10.1016/j.resmic.2004.09.007

Donnenberg, M. S., and Kaper, J. B. (1991). Construction of an eae deletion mutant of enteropathogenic *Escherichia coli* by using a positive-selection suicide vector. *Infect. Immun.* 59, 4310–4317.

Fossum, S., Søreide, S., and Skarstad, K. (2003). Lack of SeqA focus formation, specific DNA binding and proper protein multimerization in the *Escherichia coli* sequestration mutant seqA2. *Mol. Microbiol.* 47, 619–632. doi: 10.1046/j.1365-2958.2003.t01-1-03329.x

Fujikawa, N., Kurumizaka, H., Nureki, O., Tanaka, Y., Yamazoe, M., Hiraga, S., et al. (2004). Structural and biochemical analyses of hemimethylated DNA binding by the SeqA protein. *Nucleic Acids Res.* 32, 82–92. doi: 10.1093/nar/gkh173

George, R. A., and Heringa, J. (2002). An analysis of protein domain linkers: their classification and role in protein folding. *Protein Eng.* 15, 871–879. doi: 10.1093/protein/15.11.871

Germain, E., Castro-Roa, D., Zenkin, N., and Gerdes, K. (2013). Molecular mechanism of bacterial persistence by HipA. *Mol. Cell* 52, 248–254. doi: 10.1016/j.molcel.2013.08.045

Guarné, A., Brendler, T., Zhao, Q., Ghirlando, R., Austin, S., and Yang, W. (2005). Crystal structure of a SeqA-N filament: implications for DNA replication and chromosome organization. *EMBO J.* 24, 1502–1511. doi: 10.1038/sj.emboj.7600634

Guarné, A., Zhao, Q., Ghirlando, R., and Yang, W. (2002). Insights into negative modulation of *E. coli* replication initiation from the structure of SeqA-hemimethylated DNA complex. *Nat. Struct. Biol.* 9, 839–843. doi: 10.1038/nsb857

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, 135–140.

Horton, R. M., Hunt, H. D., Ho, S. N., Pullen, J. K., and Pease, L. R. (1989). Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene* 77, 61–68.

Johnson, L. N., and Barford, D. (1993). The effects of phosphorylation on the structure and function of proteins. *Annu. Rev. Biophys. Biomol. Struct.* 22, 199–232. doi: 10.1146/annurev.bb.22.060193.001215

Källberg, 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

Kasho, K., and Katayama, T. (2013). DnaA binding locus datA promotes DnaA-ATP hydrolysis to enable cell cycle-coordinated replication initiation. *Proc. Natl. Acad. Sci. U.S.A.* 110, 936–941. doi: 10.1073/pnas.121207011

Kaspy, I., Rotem, E., Weiss, N., Ronin, I., Balaban, N. Q., and Glaser, G. (2013). HipA-mediated antibiotic persistence via phosphorylation of the glutamyl-tRNA-synthetase. *Nat. Commun.* 4:3001. doi: 10.1038/ncomms4001

Katayama, T., Kasho, K., and Kawakami, H. (2017). The DnaA Cycle in *Escherichia coli*: activation, function and inactivation of the initiator protein. *Front. Microbiol.* 8:2496. doi: 10.3389/fmicb.2017.02496

Kato, J., and Katayama, T. (2001). Hda, a novel DnaA-related protein, regulates the replication cycle in *Escherichia coli*. *EMBO J.* 20, 4253–4262. doi: 10.1093/emboj/20.15.4253

Kogoma, T., and von Meyenburg, K. (1983). The origin of replication, oriC, and the dnaA protein are dispensable in stable DNA replication (sdrA) mutants of *Escherichia coli* K-12. *EMBO J.* 2, 463–468.

Korch, S. B., and Hill, T. M. (2006). Ectopic overexpression of wild-type and mutant hipA genes in *Escherichia coli*: effects on macromolecular synthesis and persister formation. *J. Bacteriol.* 188, 3826–3836. doi: 10.1128/JB.01740-05

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054

Kurokawa, K., Nishida, S., Emoto, A., Sekimizu, K., and Katayama, T. (1999). Replication cycle-coordinated change of the adenine nucleotide-bound forms of DnaA protein in *Escherichia coli*. *EMBO J.* 18, 6642–6652.

Lanzer, M., and Bujard, H. (1988). Promoters largely determine the efficiency of repressor action. *Proc. Natl. Acad. Sci. U.S.A.* 85, 8973–8977.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., et al. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947–2948. doi: 10.1093/bioinformatics/btm404

Leonard, A. C., and Grimwade, J. E. (2015). The orisome: structure and function. *Front. Microbiol.* 6:545. doi: 10.3389/fmicb.2015.00545

Løbner-Olesen, A., Skarstad, K., Hansen, F. G., von Meyenburg, K., and Boye, E. (1989). The DnaA protein determines the initiation mass of *Escherichia coli* K-12. *Cell* 57, 881–889.

Lu, M., Campbell, J. L., Boye, E., and Kleckner, N. (1994). SeqA: a negative modulator of replication initiation in E. coli. *Cell* 77, 413–426.

Ma, J., Peng, J., Wang, S., and Xu, J. (2012). A conditional neural fields model for protein threading. *Bioinformatics* 28, i59–66. doi: 10.1093/bioinformatics/bts213.

Miller, J. H. (1972). *Experiments in Molecular Genetics.* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Nievera, C., Torgue, J. J., Grimwade, J. E., and Leonard, A. C. (2006). SeqA blocking of DnaA-oriC interactions ensures staged assembly of the *E. coli* pre-RC. *Mol. Cell* 24, 581–592. doi: 10.1016/j.molcel.2006.09.016

Paleologou, K. E., Schmid, A. W., Rospigliosi, C. C., Kim, H. Y., Lamberto, G. R., Fredenburg, R. A., et al. (2008). Phosphorylation at Ser-129 but not the phosphomimics S129E/D inhibits the fibrillation of alpha-synuclein. *J. Biol. Chem.* 283, 16895–16905. doi: 10.1074/jbc.M800747200

Peng, J., and Xu, J. (2011). RaptorX: exploiting structure information for protein alignment by statistical inference. *Proteins* 10(Suppl. 79), 161–171. doi: 10.1002/prot.23175

Riber, L., Fujimitsu, K., Katayama, T., and Lobner-Olesen, A. (2009). Loss of Hda activity stimulates replication initiation from I-box, but not R4 mutant origins in *Escherichia coli. Mol. Microbiol.* 71, 107–122. doi: 10.1111/j.1365-2958.2008.06516.x

Riber, L., and Lobner-Olesen, A. (2005). Coordinated replication and sequestration of oriC and dnaA are required for maintaining controlled once-per-cell-cycle initiation in *Escherichia coli. J. Bacteriol.* 187, 5605–5613. doi: 10.1128/JB.187.16.5605-5613.2005

Sekimizu, K., Bramhill, D., and Kornberg, A. (1987). ATP activates dnaA protein in initiating replication of plasmids bearing the origin of the *E. coli* chromosome. *Cell* 50, 259–265.

Semanjski, M., Germain, E., Bratl, K., Kiessling, A., Gerdes, K., and Macek, B. (2018). The kinases HipA and HipA7 phosphorylate different substrate pools in *Escherichia coli* to promote multidrug tolerance. *Sci. Signal.* 11:eat5750. doi: 10.1126/scisignal.aat5750

Simon, R., Priefer, U., and Puhler, A. (1983). A broad host range mobilization system for *in vivo* genetic-engineering–transposon mutagenesis in gram-negative bacteria. *Bio-Technology* 1, 784–791. doi: 10.1038/nbt1183-784

Skarstad, K., Boye, E., and Steen, H. B. (1986). Timing of initiation of chromosome replication in individual *Escherichia coli* cells. *EMBO J.* 5, 1711–1717.

Skarstad, K., and Katayama, T. (2013). Regulating DNA replication in bacteria. *Cold Spring Harb.Perspect. Biol.* 5:a012922. doi: 10.1101/cshperspect.a012922

Torheim, N. K., Boye, E., Lobner-Olesen, A., Stokke, T., and Skarstad, K. (2000). The *Escherichia coli* SeqA protein destabilizes mutant DnaA204 protein. *Mol. Microbiol.* 37, 629–638. doi: 10.1046/j.1365-2958.2000.02031.x

von Freiesleben, U., Krekling, M. A., Hansen, F. G., and Lobner-Olesen, A. (2000). The eclipse period of *Escherichia coli. EMBO J.* 19, 6240–6248. doi: 10.1093/emboj/19.22.6240

Wang, S., Ma, J., Peng, J., and Xu, J. (2013). Protein structure alignment beyond spatial proximity. *Sci. Rep.* 3:1448. doi: 10.1038/srep01448

Wang, S., Peng, J., and Xu, J. (2011). Alignment of distantly related protein structures: algorithm, bound and implications to homology modeling. *Bioinformatics* 27, 2537–2545. doi: 10.1093/bioinformatics/btr432

Replicate Once Per Cell Cycle: Replication Control of Secondary Chromosomes

Florian Fournes[1,2], Marie-Eve Val[1,2], Ole Skovgaard[3] and Didier Mazel[1,2]*

[1] Unité Plasticité du Génome Bactérien, Département Génomes et Génétique, Institut Pasteur, Paris, France, [2] UMR3525, Centre National de la Recherche Scientifique, Paris, France, [3] Department of Science and Environment, Roskilde University, Roskilde, Denmark

Faithful vertical transmission of genetic information, especially of essential core genes, is a prerequisite for bacterial survival. Hence, replication of all the replicons is tightly controlled to ensure that all daughter cells get the same genome copy as their mother cell. Essential core genes are very often carried by the main chromosome. However they can occasionally be found on secondary chromosomes, recently renamed chromids. Chromids have evolved from non-essential megaplasmids, and further acquired essential core genes and a genomic signature closed to that of the main chromosome. All chromids carry a plasmidic replication origin, belonging so far to either the iterons or repABC type. Based on these differences, two categories of chromids have been distinguished. In this review, we focus on the replication initiation controls of these two types of chromids. We show that the sophisticated mechanisms controlling their replication evolved from their plasmid counterparts to allow a timely controlled replication, occurring once per cell cycle.

Keywords: megaplasmids, chromids, *repABC*, iterons, replication initiation

INTRODUCTION

The genome of most of bacteria is carried by a single circular chromosome, which is replicated bi-directionally from a single origin in a highly controlled manner. Approximately 10% of the bacterial species have their genome divided in two, or more, large replicative DNA molecules, with a main chromosome, and one or several secondary replicons (second chromosomes and/or megaplasmids) (Harrison et al., 2010; Touchon and Rocha, 2016; diCenzo and Finan, 2017). Several evidences suggest that second chromosomes originate from plasmids that have been domesticated by their ancestral host to become *bona fide* chromosomes (Harrison et al., 2010). Plasmids could represent up to 30% of the bacterial genomes, and in some cases large plasmids were called megaplasmids. One of the founding events of plasmid or megaplasmid domestication involves the transfer of essential core genes from the main chromosome to the plasmid. Certainly because of their plasmid ancestry, all studied secondary chromosomes carry a plasmid-like replication system. In the alpha-proteobacteria *Rhodobacter sphaeroides* the secondary replicon carries a *repABC* replication system (Suwanto and Kaplan, 1989; Cevallos et al., 2008), while, all the species belonging to the *Vibrionaceae* family have a specific iteron plasmid-like replication system dedicated to their second chromosome (Okada et al., 2005). Nonetheless, mechanisms controlling the second chromosomes replication appear to be more sophisticated than that controlling plasmid replication.

Combining essential core genes and plasmid-like replication origin, second chromosomes exhibit features of chromosomes and plasmids, and thus were named chromids. From now on, we will use this terminology for such replicons (Harrison et al., 2010).

Faithful transmission of genetic information from a mother cell to daughter cells requires cell cycle coordinated replication and segregation of the genetic material before cell division. Chromosomal replication has an elaborated control of when to start DNA replication (timing of initiation); an accurate replication-elongation stage and a termination that untangles the replicated chromosomes now ready for partitioning (Schekman et al., 1974; Reyes-Lamothe et al., 2012). Chromosomes differ from plasmids in part by their replication controls, both in terms of initiation process and by their integration to the cell cycle. Chromosome replication generally occurs once per cell cycle and responds to cell growth parameters. On the contrary, plasmids may replicate in a cell cycle independent manner and their replication can be initiated randomly during the cell cycle (Nordström and Dasgupta, 2006). That being said, this last affirmation has been for years subject to debate, as for example, the F and R1 model plasmids supposedly replicate at a particular time during the cell cycle (Zeuthen and Pato, 1971; Pritchard et al., 1975). Replication initiation of almost all replicons starts when the origin-specific replication initiator recognizes and binds motifs located in a well-defined origin region (Wegrzyn et al., 2016). With the exception of certain symbiotic species and few cyanobacteria, chromosomal DNA replication is initiated at a conserved replication origin, oriC, and is orchestrated by DnaA, the "universal" initiator of chromosomal replication in bacteria (Akman et al., 2002; Ohbayashi et al., 2016; Hansen and Atlung, 2018). Plasmid replication can be controlled either by the binding of an initiator to repeated sequences called iterons, or by a small antisense RNA (Chattoraj, 2000; Brantl, 2014; Gaimster and Summers, 2015). Chromids contain a replication origin related to the one of plasmids and thus have retained many of plasmid-like features. Megaplasmids and chromids seem both to share a more tightly controlled replication (Rasmussen et al., 2007; Frage et al., 2016). However, due to their large size, chromids probably necessitated additional mechanisms of initiation control, which permit a well-defined replication initiation mostly integrated to the cell cycle. Two major types of chromids are distinguished based on their replication mechanisms: iteron chromids and repABC chromids. The repABC chromids are exclusively found in the alphaproteobacteria and their replication is dependent on an operon composed of three genes: repA, repB, and repC. Even if, only RepC, the initiator, is essential for DNA replication, all three proteins RepA, RepB, and RepC are required to tightly control replication initiation (Pinto et al., 2012). The iteron chromids are found in the two other classes of proteobacteria (beta and gamma), and their replication origin is mainly composed of short repeated sequences, called iterons, localized near a gene encoding the replication initiator (Heidelberg et al., 2000; Du et al., 2016). Vibrio cholerae has served as the model for investigations of iteron chromids replication and its connection with the cell cycle.
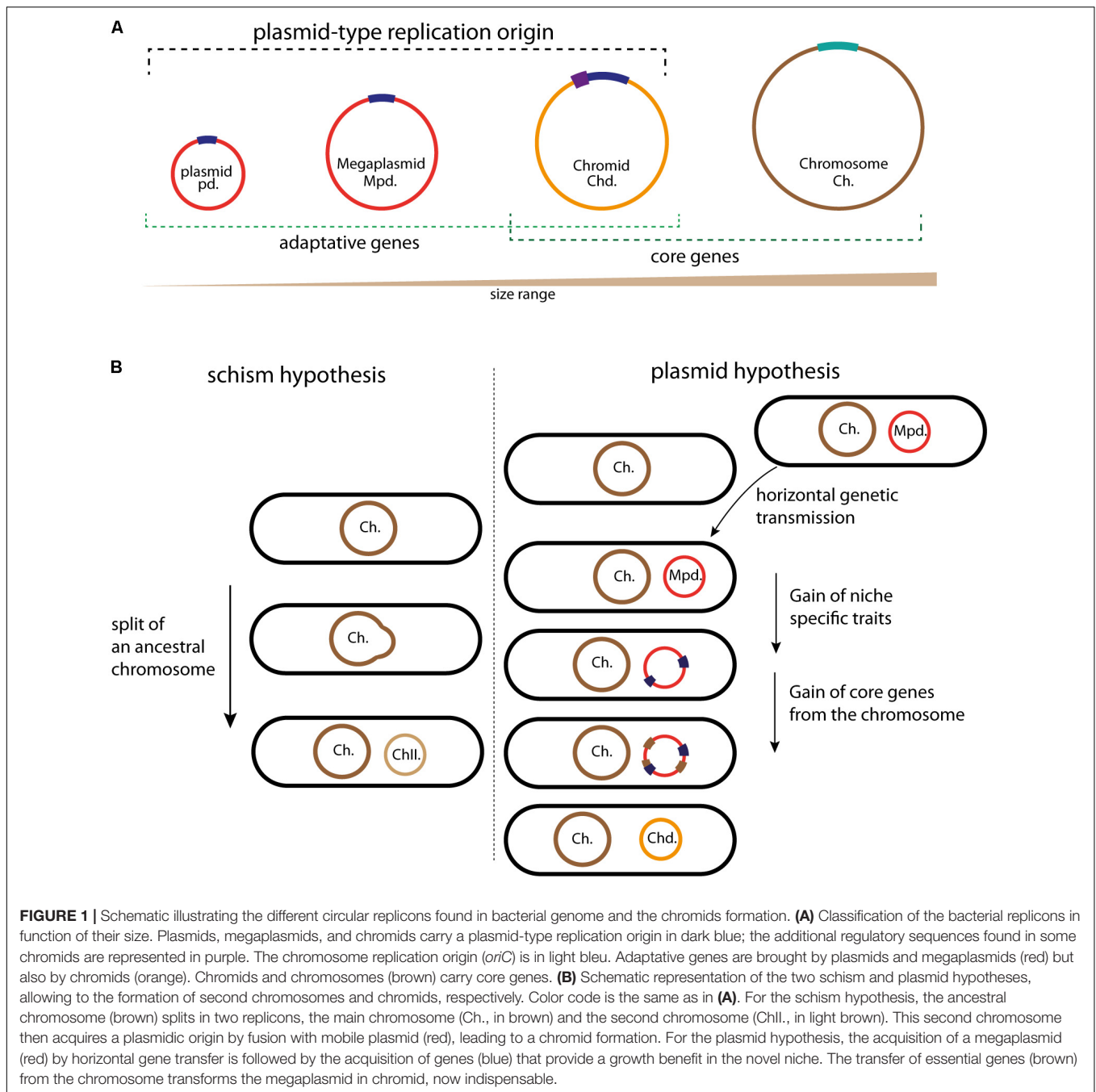
Here we review and discuss the mechanisms controlling the replication initiation of these two types of chromids: iteron and repABC. We highlight the complex levels of control found in chromids, compared to those of their ancestral plasmids, which allow chromids to replicate once, and only once, per cell cycle. We also discuss the timing of replication initiation of the iteron and repABC chromids and their integration to the cell cycle.

FROM MEGAPLASMIDS TO CHROMIDS

Origin of Chromids

Bacterial genomes always include one chromosome and may also include plasmids. Plasmids provide beneficial accessory traits for the organism, for example, antibiotics resistance and anabolic pathways, but do not carry essential genes and thus are dispensable (**Figure 1A**). On the contrary, chromosomes harbor essential genes and are indispensable. This dogma changed, first, with the identification of linear chromosomes and plasmids (Hirochika and Sakaguchi, 1982; Baril et al., 1989), and in 1989, when Suwanto and Kaplin, using a pulsed-field gel electrophoresis, discovered a large second replicon in the alpha-proteobacteria R. sphaeroides (Suwanto and Kaplan, 1989). This replicon carrying essential genes was called "second chromosome". The definition of second replicons as chromosome is mostly based on their essentiality in the bacteria growth and survival. In the 1990s other chromids were identified in Agrobacterium tumefaciens, Brucella melitensis, Leptospira interrogans, and in several Vibrio species (Allardet-Servent et al., 1993; Michaux et al., 1993; Zuerner et al., 1993; Trucksis et al., 1998; Yamaichi et al., 1999). In parallel, large replicons were discovered and called megaplasmids (**Figure 1A**) (Rosenberg et al., 1982). Compared to chromids, megaplasmids are non-essential, they encode their own replication and partition system, and carry adaptative genetic information such as the capacity for Shigella flexneri to invade the eukaryotic cells or, for the Rhizobiaceae to create a symbiosis with legumes (Buchrieser et al., 2000; Marchetti et al., 2010). The difference between plasmids and megaplasmids is currently based on the replicon size, and it will be of great benefit to establish if specific and functional characteristics discriminate plasmids and megaplasmids (**Figure 1A**). Chromids are normally larger than the accompanying plasmids and smaller than the associate chromosome. Comparative analysis of the relative synonymous codon usage of bacterial replicons demonstrates that individual replicons have distinct codon usage characteristics, and that chromids are much closer in codon usage to chromosomes than to plasmids (Harrison et al., 2010; Petersen et al., 2013). This observation implies that chromids have been acquired earlier than plasmids and have spent more time in the same cellular environment as the associated chromosome. Thus, codon usage analysis can be useful to chromids classification, as it was the case with the Rhodobacteraceae (alpha-proteobacteria) genomes analysis (Petersen et al., 2013; Frank et al., 2015a). In addition to that, three main criteria have been proposed to robustly distinguish chromids from chromosomes and plasmids or megaplasmids (**Figure 1A**) (Harrison et al., 2010). Replicons called chromids use a plasmid type maintenance and replication system, harbor a nucleotide composition close to that of the

FIGURE 1 | Schematic illustrating the different circular replicons found in bacterial genome and the chromids formation. **(A)** Classification of the bacterial replicons in function of their size. Plasmids, megaplasmids, and chromids carry a plasmid-type replication origin in dark blue; the additional regulatory sequences found in some chromids are represented in purple. The chromosome replication origin (*oriC*) is in light bleu. Adaptative genes are brought by plasmids and megaplasmids (red) but also by chromids (orange). Chromids and chromosomes (brown) carry core genes. **(B)** Schematic representation of the two schism and plasmid hypotheses, allowing to the formation of second chromosomes and chromids, respectively. Color code is the same as in **(A)**. For the schism hypothesis, the ancestral chromosome (brown) splits in two replicons, the main chromosome (Ch., in brown) and the second chromosome (ChII., in light brown). This second chromosome then acquires a plasmidic origin by fusion with mobile plasmid (red), leading to a chromid formation. For the plasmid hypothesis, the acquisition of a megaplasmid (red) by horizontal gene transfer is followed by the acquisition of genes (blue) that provide a growth benefit in the novel niche. The transfer of essential genes (brown) from the chromosome transforms the megaplasmid in chromid, now indispensable.

chromosome and carry essential core genes that are found on the chromosome of other species (Harrison et al., 2010). Prediction of the essentiality of core genes located within chromids is largely based on automated gene annotations. Experimental validations have in some cases shown that the predicted essential gene actually is dispensable (Cheng et al., 2007; Agnoli et al., 2012). For instance, in the case of the replicon pSymB of *Sinorhizobium meliloti* the *minCDE* genes were predicted to be essential, nonetheless, disruption of the *minE* gene is possible and only provokes a nitrogen fixation defect involved in symbiosis (Cheng et al., 2007). However, pSymB also carries core genes in unique

copy, such as *engA* and $tRNA^{arg}$ and can still be considered as a chromid (diCenzo et al., 2013). Furthermore, chromids can be dispensable under smooth laboratory conditions, but must be required to bacteria survival in the harsh natural environment (Dziewit et al., 2014; Frank et al., 2015b; Soora et al., 2015). Thus, it was proposed to subdivide chromids into two types: "primary" and "secondary" chromids (Dziewit et al., 2014). Primary chromids are indispensable for host viability, while secondary chromids are considered as "facultatively" essential (Dziewit et al., 2014). However, many secondary replicons, such as megaplasmids carrying, for example, antibiotic resistance

genes, which are essential for bacterial growth in presence of theses antibiotics and yet are not considered as chromids. Then, environment-specific beneficial or essential genes are insufficient to associate a replicon with the chromid term (diCenzo and Finan, 2017). Thus, even if this subdivision of chromids would be useful, we should be aware that it has to be carefully used.

The comparison of the available data helps us to determine the extent of megaplasmids and chromids relationship. Two main adaptive traits differentiate megaplasmids and chromids, leading to a stable and cell cycle integrated replicon: the acquisition of genomic signatures similar to those of cognate chromosomes (GC content and codon usage to limit physiological perturbation) and of essential genes. Two hypotheses have been proposed to explain the formation mechanism of an essential secondary replicon (Moreno, 1998; Egan et al., 2005; Prozorov, 2008; Harrison et al., 2010; diCenzo and Finan, 2017) (**Figure 1**). The first, called schism hypothesis, proposes that the formation of second essential replicon is the consequence of a split of an ancestral chromosome into two replicons: main and second chromosomes (**Figure 1B**). The second chromosome could then acquire the plasmid like replication system by fusion with a mobile plasmid, then becoming a chromid (Harrison et al., 2010; diCenzo and Finan, 2017). This was originally proposed to explain the formation of chromids found in *Brucella suis* and *R. sphaeroides*, but it seems to be able to explain solely rare chromids formation (Choudhary et al., 1997; Jumas-Bilak et al., 1998; diCenzo and Finan, 2017). Indeed, in bacteria, there is no evidence for the formation of chromids through the schism hypothesis. However, a recent study in the Archeon *Haloferax volcanii* describes the formation of a prokaryotic multipartite genome in agreement with the schism hypothesis. *H. volcanii* has a multipartite genome, consisting of a main chromosome, three secondary essential replicons and a plasmid, and its main chromosome has three origins, which are already well controlled (Norais et al., 2007; Hartman et al., 2010). In response to an *orc* gene deletion (*orc* encode the replication initiator Orc1), the multi-origin chromosome of *H. volcanii* split by homologous recombination into two elements, thus leading to the creation of a stable second chromosome (Ausiannikava et al., 2018). Contrary to the first hypothetical model, the second, called plasmid hypothesis, states that chromids evolved from megaplasmids (**Figure 1B**). This hypothesis implies that the coevolution of a megaplasmid with a chromosome will result in a transformation of the megaplasmid genomic signatures to that of the chromosome. This transformation is accompanied by the acquisition of essential genes (**Figure 1B**). This is supported by examples belonging to both the *repABC* and iterons chromids, which all carry a plasmid-like replication system and harbor a codon usage similar to that of the chromosome (Harrison et al., 2010; Pinto et al., 2012). Furthermore, the distribution of essential genes and the functional annotation onto the chromids are different compared to those of the chromosomes (Heidelberg et al., 2000; Goodner et al., 2001; Chao et al., 2013). As introduced above, these steps of evolution are the two main adaptive traits of a stable replicon. Strikingly, all observations gathered so far concluded that the plasmid hypothesis could explain the formation of all the studied chromids.

The acquisition of essential genes, prerequisite to the chromid formation, is driven by gene transfers from the chromosome to a megaplasmid (**Figure 1B**). Two possible mechanisms can explain the transfer of essential genes (diCenzo and Finan, 2017). First, inter-replicon genetic transfers could be catalyzed by homologous recombination, for example, by shared insertion sequences (IS), or IS using replicative transposition and resolution by recombination between different IS copies (Lesic et al., 2012). This transfer of genes leads to essential gene deletion from the chromosome. For instance, this is the case for the *engA* and the *tRNA^{arg}* genes in the chromid pSymB, which resulted from the transfer of a 69Kb DNA fragment from the *S. meliloti* chromosome to the pSymB ancestor (diCenzo et al., 2013). On the other hand, the second mechanism takes into account the genetic redundancy due to inter-replicon gene duplication or to the acquisition of an orthologous gene by lateral genetic transfer. Several such examples of redundancy have been pointed in the genome sequences of *V. cholerae*, *R. sphaeroides* and *S. meliloti* (Heidelberg et al., 2000; Bavishi et al., 2010; diCenzo and Finan, 2015). For instance, massive inactivation experiments in *S. meliloti* chromosome has shown that more than 10% of the chromosomal genes have redundant functional copy on the megaplasmid pSymA or on the chromid pSymB, and this is a possible consequence of genes duplication (diCenzo and Finan, 2015).

Where and Why Multipartite Genomes Appeared?

Bacterial genomes carried by more than one large replicon, thus containing megaplasmids and/or chromids, correspond to a divided or a multipartite genome. Increase in genome sequencing over the last years revealed that approximately 10% of the complete bacterial genomes are multipartite (Harrison et al., 2010; Touchon and Rocha, 2016; diCenzo and Finan, 2017). Multipartite genomes are found allover the bacterial kingdom but chromids are mainly found in proteobacteria, including the alpha, beta, and gamma proteobacteria (Harrison et al., 2010). Interestingly, megaplasmids are rarely conserved among genera, but are common in genera containing bacteria involved in symbiotic and pathogenic relationship. Furthermore, they carry genes specific to strains and species. In contrast, chromids are conserved among different genera and carry genus specific characters and genes (Harrison et al., 2010). For instance, pSymA is present only in few closely related *S. meliloti* species, and there is a high genes variation in individual strains (Cevallos et al., 2008; Guo et al., 2009). On the other hand, pSymB is supposed to be an old acquired replicon, sharing common ancestry with *Brucella* chromids, and pSymB chromids belonging to *S. meliloti* genomes show a high synteny between different isolates (Cevallos et al., 2008; Guo et al., 2009; Galardini et al., 2013). Thus, even if it could be difficult to differentiate chromids from megaplasmids with a systematic study of the genome, these observations may be key criteria to distinguish the two replicons. Besides the fact that chromids carry indispensable core genes, the advantages of multipartite genomes are not yet clearly established. Several hypotheses have been proposed. Multipartite genomes could

allow bacteria to have a larger genome, and reduce the complexity of the circular replicons, which permit to correctly manage their heredity (*e.g.*, resolution of chromosome dimers) (Val et al., 2008, 2012). Indeed, the total genome size of the multipartite genomes are on average larger than the non-multipartite genomes, and the differences in genome sizes is correlated to the chromids size and not to the chromosomal size (diCenzo and Finan, 2017). In agreement with the previous hypothesis, the fast growing rhizobia contain a chromid contrary to the slow growing rhizobia (Yamaichi et al., 1999; Pastorino et al., 2003; MacLean et al., 2007). A second hypothesis is that chromids could permit the coordination and regulation of gene expression, contributing to the bacteria adaptation into novel niches. For instance, genes carried by *V. cholerae* chromid are differentially expressed *in vitro* and during the colon colonization. Indeed, during colon infection, *V. cholerae* induces a higher expression of chromid genes (Xu et al., 2003). These genes are involved in response to environmental stresses, allowing intra-intestinal growth and biofilm formation (Xu et al., 2003; Silva and Benitez, 2016).

The previous paragraphs highlighted the prevalence of chromids and their essentiality in the bacterial kingdom. The following sections will present what we know about their maintenance in the cell, focusing on the replication system of the iterons and *repABC* chromids.

ITERON-CHROMIDS AND *Vibrio cholerae* PARADIGM

The genome of *V. cholerae* is divided in two replicons of different sizes: the main chromosome (Chr1) of 3 Mbp and the chromid (Chr2) of 1 Mbp (Trucksis et al., 1998; Yamaichi et al., 1999; Heidelberg et al., 2000). Each replicon encodes a specific partition system, ParAB1 and ParAB2, which recognize different *parS* sites carried on their cognate replicons. Their replication is also differentially regulated (Duigou et al., 2006; Yamaichi et al., 2007). The replication origin of Chr1 is highly related to the chromosomal origin of *Escherichia coli*, and is controlled by the ubiquitous replication initiator DnaA (Duigou et al., 2006). The control of the replication by DnaA is elaborate, and involves, in addition to the regulation of the DnaA concentration in the cell, a balance of the binding affinity of DnaA to multiple sites within or outside the replication origin. The different levels of control of the DnaA replication process have been recently reviewed in (Hansen and Atlung, 2018). The *V. cholerae* main chromosome origin (*ori1*) contains DnaA binding sites, an IHF binding site and several GATC sites for methylation catalyzed by the DNA adenine methyl-transferase (Dam). Dam methylation is not essential to initiate the replication of Chr1, but SeqA, which recognize the hemi-methylated DNA, is required to restrict *ori1* initiation once per cell cycle (Demarre and Chattoraj, 2010). *ori1* can functionally replace the *E. coli*, *oriC*, and sustains chromosome replication (Koch et al., 2010). DnaA can bind ATP or ADP, but only ATP-DnaA can initiate the chromosomal replication initiation (Hase et al., 1998; Kawakami et al., 2005; Katayama et al., 2010; Hansen and Atlung, 2018). The regeneration of the ATP-DnaA, from the ADP-DnaA, is crucial for chromosome

replication control. One of the mechanisms catalyzing this regeneration involves two intergenic regions called DARS1 and DARS2 (DnaA Reactivating Sequence) (Fujimitsu et al., 2009). DARS-like sequences are also found, with the same localization (between *uvrB* and *mutH*), in *V. cholerae* (Fujimitsu et al., 2009). All together, these observations suggest that *V. cholerae* Chr1 and *E. coli* chromosomes share many similar mechanisms to control their initiation.

This, however, does not exclude the involvement of *V. cholerae* species-specific elements to control the DnaA dependent replication. Indeed, the replication regulation of *Bacillus subtilis* and *Caulobacter crescentus*, two other model bacteria, which also use DnaA as initiator, involves additional and specific factors (Murray and Errington, 2008; Scholefield et al., 2012; Duan et al., 2016; Felletti et al., 2018). For example, Soj, an homolog of the partition protein ParA, controls the replication initiation during the *B. subtilis* vegetative growth (Ogura et al., 2003). Soj performs two opposite activities depending on its monomeric or dimeric state. Indeed, Soj monomers inhibit replication by preventing DnaA oligomerization (Murray and Errington, 2008; Scholefield et al., 2012). Conversely, Soj dimers, which require binding to ATP, activate replication by promoting DnaA oligomerization (Murray and Errington, 2008; Hansen and Atlung, 2018). *E. coli* has no *par* genes, but as mentioned above *V. cholerae* has one for each chromosome, and the *V. cholerae parB1* deletion induces Chr1 over-initiation; the same phenomenon is observed with an over-expression of ParA1, suggesting that ParA1 stimulates chromosome replication initiation as Soj does in *B. subtilis* (Kadoya et al., 2011).

Players in the Replication of the *V. cholerae* Chromid: *ori2* and RctB

Vibrio cholerae chromid, Chr2, carries a different replication origin (*ori2*) compared to the origin of the main chromosome (**Figure 2A**). Initiation of the replication at *ori2* is catalyzed by a specific factor named RctB, which is highly conserved within the *Vibrionaceae* family. The ∼900 bp *ori2* has retained many of iteron-plasmid features for replication control. *Ori2* is organized into two functional domains: *ori2-min*, which supports the replication alone and an adjacent sequence, *ori2-inc*, which acts as a negative regulator of replication (**Figure 2A**). Both parts contain a variety of RctB binding sites, which are named based on their length: 11-mers, 12-mers, 29-mer, and 39-mers (**Figure 2A**). The iterons, 11-mers and 12-mers, are closely related, without any similarity with the 29-mer and 39-mers. The 29-mer corresponds to a truncated 39-mer, missing 10 nt in its center (Venkova-Canova et al., 2012). The *ori2-min* harbors an array of six 12-mers oriented in a head-to-tail manner with a regular spacing of 10 or 11 base pairs and each 12-mer contains a GATC Dam methylation site. As *ori1*, *ori2* also contains a DnaA binding site, though a single one, and an IHF binding site (IBS) (**Figure 2A**). Furthermore, *ori2* DnaA binding site is required for the Chr2 replication but DnaA is not limiting to control the timing of replication initiation, suggesting that it must have another function (Duigou et al., 2006). The exact implication of the DnaA binding site and of the IBS in Chr2 initiation
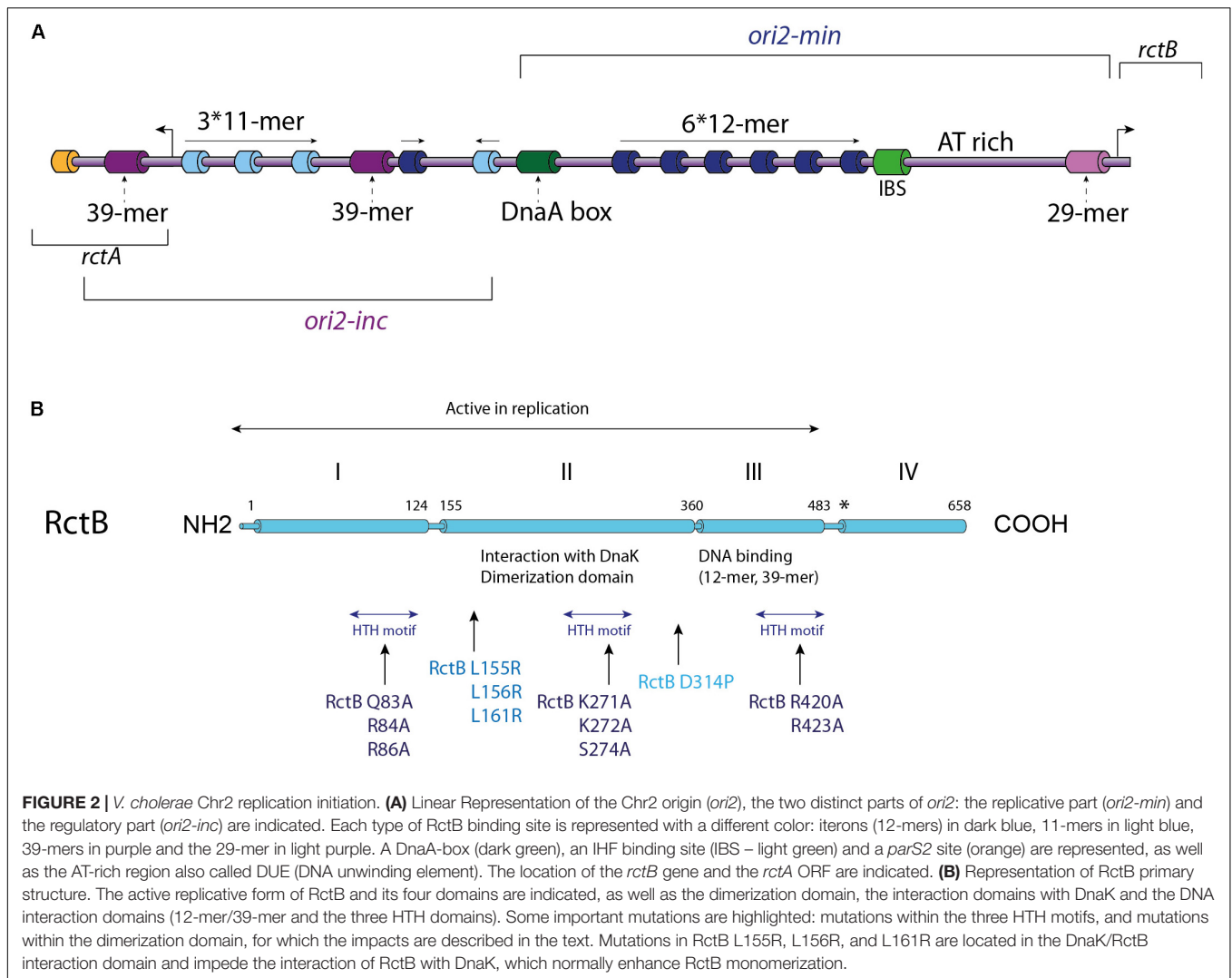
FIGURE 2 | *V. cholerae* Chr2 replication initiation. **(A)** Linear Representation of the Chr2 origin (*ori2*), the two distinct parts of *ori2*: the replicative part (*ori2-min*) and the regulatory part (*ori2-inc*) are indicated. Each type of RctB binding site is represented with a different color: iterons (12-mers) in dark blue, 11-mers in light blue, 39-mers in purple and the 29-mer in light purple. A DnaA-box (dark green), an IHF binding site (IBS – light green) and a *parS2* site (orange) are represented, as well as the AT-rich region also called DUE (DNA unwinding element). The location of the *rctB* gene and the *rctA* ORF are indicated. **(B)** Representation of RctB primary structure. The active replicative form of RctB and its four domains are indicated, as well as the dimerization domain, the interaction domains with DnaK and the DNA interaction domains (12-mer/39-mer and the three HTH domains). Some important mutations are highlighted: mutations within the three HTH motifs, and mutations within the dimerization domain, for which the impacts are described in the text. Mutations in RctB L155R, L156R, and L161R are located in the DnaK/RctB interaction domain and impede the interaction of RctB with DnaK, which normally enhance RctB monomerization.

is still unknown (Gerding et al., 2015; Schallopp et al., 2017). DnaA binding sites have been found in the replication origin of many plasmids (Lu et al., 1998; Wegrzyn et al., 2016), and two hypotheses have been proposed for the possible role of DnaA in plasmid replication. First, it has been suggested that DnaA could help the stabilization of the origin opening catalyzed by the plasmid replication initiators (Rep proteins), and second that DnaA was needed for the helicase loading. Thus, it is tempting to think that DnaA and IHF have conserved the same hypothetic regulatory functions for *V. cholerae* Chr2 replication initiation. Moreover, a recent study showed that DnaA negatively regulates the replication of a mini R1-1 plasmid (Yao et al., 2018). This observation suggests that DnaA, bound to *ori2*, could be also involved in a negative regulation of the *ori2* replication initiation, interacting with RctB. The remaining part of *ori2-min* contains an A-T rich region and a 29-mer RctB binding site overlapping the *rctB* promoter (**Figure 2A**). The regulatory *ori2-inc* part is mainly composed of one 39-mer and of a second 39-mer found at the outskirt, overlapping a transcribed but non-translated ORF *rctA*. 39-mers do not contain Dam methylation site. Four 11-mers

containing GATC sites and one single 12-mer are also located in *ori2-inc* (Venkova-Canova and Chattoraj, 2011) (**Figure 2A**). All these sites are known to play a replication initiation regulatory role, which we will describe below.

RctB is a 658 amino acids protein consists of four domains and its sequence has no detectable homology with other replication initiator (Orlova et al., 2017) (**Figure 2B**). RctB, with a molecular mass of 75.3 kDa is larger than other chromosomal or plasmidic initiator proteins, suggesting that it performs additional functions compared to DnaA and Rep proteins. The first 500 residues, including domains I, II, and III, are sufficient to promote *ori2* replication initiation (Yamaichi et al., 2011; Jha et al., 2012; Koch et al., 2012) (**Figure 2B**). The domain IV is supposed to mediate protein-protein interaction, and thus play a regulatory role in the RctB oligomerization on the origin (Yamaichi et al., 2011; Koch et al., 2012; Orlova et al., 2017). Recent structural and biochemical studies of domains II and III showed that RctB adopts a head-to-head dimeric form in solution (Jha et al., 2017; Orlova et al., 2017) (**Figure 2B**). Interestingly, the structure of these two central domains exhibit significant

similarities with plasmid-type Rep proteins, including π from the R6K plasmid and RepE from the F plasmid (Komori et al., 1999; Swan et al., 2006). Despite the fact that domains III and IV was predicted to be a dimerization interface (Jha et al., 2014), structure of the RctB dimer, restricted to domains II and III, shows that the interaction is mediated by the domain II. Furthermore, substitution of a proline within the beta strand closest to the dimer interface disrupts dimer formation and produces a monomeric mutant in the full length RctB (D314P; **Figure 2B**) (Orlova et al., 2017).

As RctB is the *Vibrio* central player of chromid replication initiation, it should be able to take on different functions. The first of these is the recognition and binding to its target sites. The interaction between RctB and the 12-mer and 11-mer is dependent of the DNA methylation state, while its binding to the 39-mer and the 29-mer is methylation independent (Demarre

and Chattoraj, 2010; Venkova-Canova et al., 2012). DNA/protein interaction experiments, using different RctB mutants, revealed that the domains interacting with the 12-mer and the 39-mer are spatially close and localized in the domain III (**Figure 2B**) (Jha et al., 2014). It was first proposed that RctB binds to the methylated 12-mer both as a monomer and a dimer (Jha et al., 2012) (**Figure 3A**). However, the head to head dimeric form of RctB is incompatible with the head to tail arrangement of 12-mer within *ori2-min* (Orlova et al., 2017) (**Figure 3A**). The crystal structure reveals that RctB contains more DNA binding surface than previously thought, with at least three helix-turn-helix (HTH) motifs identified, each one localized in a given domain (I, II, and III) (**Figure 2B**). Mutations in these three HTH reduce the RctB binding to all its target sites suggesting that all this three HTH are involved in DNA interactions. Furthermore, mutations in the three domains do not exhibit the

FIGURE 3 | Main regulatory mechanisms controlling Chr2 replication initiation. **(A)** Representation of the two different models of RctB binding to the iterons. **(i)** RctB dimers, in blue, are transformed in monomers (light blue) by DnaK/J. Both dimers and monomers are able to binds to the iterons. **(ii)** RctB binding to the iterons is only possible under its monomeric form. The DnaK/J interaction with RctB not only causes its monomerization, but also its oligomerization (dark blue) onto the DNA containing iterons allowing to the origin unwinding. **(B)** Representation of the mechanisms involved in *ori2* replication initiation. RctB binding sites within the *ori2* are indicated and color codes are identical to those of the **(A)**. A black arrow illustrates RctB binding to its binding sites. A positive control is represented by a green arrow associated to (+), and a negative control is represented by flat end red arrow associated to (–). SeqA (orange) impedes the RctB binding to iterons, ParB2 (yellow) and *rctA* transcription (brown arrow) impede the RctB binding to 39-mers (bar black arrows). The handcuffing of the 39-mer with iterons within *ori2-inc* has a positive control on *ori2* replication initiation since it competes with the 39-mer handcuffing with *ori2-min* iterons (bar blue arrow).

same behavior regarding binding activity to the 11–12-mers and to the 29–39-mers. Indeed, all three domain I, II, and III, seem to be involved in the methylation dependent DNA binding (12-mer and 11-mer), while only domain II is involved in the methylation independent binding (29-mer and 39-mer) (Orlova et al., 2017).

In the iteron-plasmids mechanism of replication initiation, DnaK and DnaJ enhance initiator binding to the origin (Wickner et al., 1991). DnaK and DnaJ were first discovered as factors required for the bacteriophage lambda replication and later as enhancers for the replication of plasmids containing iterons within their origin (Friedman et al., 1984; Wickner et al., 1991). Plasmid initiators can dimerize, but in general bind to the origin only as monomers. DnaK/DnaJ system helps to monomerize plasmid initiator and promote the replication initiation. Based on structural data of the plasmid initiators RepA and RepE, it was proposed that monomerization is not sufficient to initiate the replication, and that monomers have to be remodeled, likely to catalyze origin unwinding (Díaz-López et al., 2003; Giraldo et al., 2003; Nakamura et al., 2007). In solution the RctB dimeric form is the most stable, this implies that monomerization of the protein has to be triggered to permit DNA binding (Jha et al., 2017; Orlova et al., 2017). RctB is remodeled from dimer to monomer by the chaperones DnaJ and DnaK *via* an interaction between DnaK and RctB domain II (**Figure 3A**) (*cf.* mutations L155R, L156R, and L161R; **Figure 2B**) (Jha et al., 2014, 2017). For Chr2 replication initiation, DnaK and DnaJ are strictly required to promote *ori2* replication initiation, and were shown to promote RctB binding to both activating and inhibiting sites (12-mers and 39-mers) (Jha et al., 2012). That being said, the elucidation of the precise characteristics of the RctB-DNA interaction needs further structural and biochemical studies, for example, to experimentally show the incapacity of RctB dimer to bind DNA. RctB mutants reducing the dimerization (*e.g.,* F311P) are still DnaKJ dependent to initiate the replication, suggesting that RctB monomers have to be remodeled to correctly work (Jha et al., 2017) (**Figure 3A**). Once bound to the *ori2-min* 12-mer, RctB has to oligomerize to open the adjacent A-T rich region (unwinding activity). The nature of this last process remains obscure. Thus, experimental data determining the role of DnaK and J, the identification of the RctB domain(s) involved in its oligomerization, as well as the precise role of A-T rich sequences needed to stabilize the opening of *ori2* are still missing.

V. cholerae Chromid Controls of Replication Initiation

Vibrio cholerae Chr2 replicate once per cell cycle, pointing to a tight control through the balance between positive and negative effectors (Egan and Waldor, 2003; Egan et al., 2004; Venkova-Canova and Chattoraj, 2011; Baek and Chattoraj, 2014; Val et al., 2016). To summarize, RctB acts on two major types of sites, the 12-mer (iteron) to promote the replication initiation by unwinding the AT-rich region, and the 39-mer to inhibit it (**Figure 3B**). In *E. coli*, a plasmid carrying the entire *ori2* replicates at a copy number equal to that of the *E. coli* chromosome, and a plasmid carrying only *ori2-min* has a copy number increased by about 10 fold. Furthermore, the addition of the 39-mer to

a plasmid containing *ori2-min* drastically reduced the plasmid copy number in the cell (Venkova-Canova and Chattoraj, 2011; Koch et al., 2012; Messerschmidt et al., 2015). The two main mechanisms of inhibition correspond to (1) the RctB titration and (2) the handcuffing between the 39-mer and the *ori2-min* 12-mer mediated by RctB (**Figure 3B**) (Venkova-Canova and Chattoraj, 2011). The inhibitory activity of the 39-mer is central, and the majority of the mechanisms that enhance replication initiation modulate the RctB/39-mer interactions (Pal et al., 2005; Venkova-Canova et al., 2006; Yamaichi et al., 2011).

The regulatory function of the iterons found in the *ori2-inc* region is dual. Indeed, they have a titration activity, similar to the 39-mer, but, additionally, they help to restrain the 39-mer inhibitory activity by enhancing the handcuffing inside the *ori2-inc* region, thus releasing the *ori2-min* 12-mers (Venkova-Canova and Chattoraj, 2011) (**Figure 3B**). Furthermore, the ParB2 protein, which binds Chr2 specific centromeres localized closer to the *ori2-inc*, serves as RctB competitor for the 39-mers binding by two mechanisms: (1) spreading from the *parS2* site closer to the leftmost 39-mer and (2) direct interaction with the central 39-mer (Yamaichi et al., 2011; Venkova-Canova et al., 2013) (**Figure 3B**). In addition, as the leftmost 39-mer is covered by the *rctA* transcript, this also interferes with the RctB binding at this site and thus impede its inhibitory activity (Venkova-Canova et al., 2006) (**Figure 3B**). These mechanisms controlling the 39-mer/RctB interactions release RctB from the inhibitor sites, first decreasing the titration phenomenon and second the handcuffing. Furthermore, as found for DnaA, the concentration of available RctB in the cell controls the Chr2 replication initiation. Thus, RctB gene expression is also tightly controlled. RctB auto-regulates its own expression through binding to the 29-mer located in the *rctB* promoter, where it plays a role of transcriptional repressor and exerts a negative feedback regulation (Pal et al., 2005; Egan et al., 2006) (**Figure 3B**). This 29-mer is also implicated in the *ori2* iterons handcuffing and is able to functionally replace the 39-mer (Venkova-Canova et al., 2012). In addition to this transcriptional regulation, the RctB concentration available to initiate the replication is also significantly controlled by its titration on various regulatory sites. As introduced above, the *ori2-inc* iterons together with the 39-mers and 29-mer can titrate RctB and reduce RctB binding to the *ori2-min* replicative iterons. Chromatin immunoprecipitation (Chip-chip) experiments have revealed that RctB also binds to a number of sites clustered within a 74 Kbp sequence on the Chr2 located 40 Kbp away from the *ori2* (Baek and Chattoraj, 2014). This 74 Kbp sequence contains six RctB binding sites: five iterons and one 39-mer like sequence, which also negatively regulate the *ori2* replication initiation. This locus titrate RctB and inhibit the *ori2* replication initiation, its activity and localisation suggest that it is comparable to the *E. coli datA* titration locus (Kitagawa et al., 1998; Kasho and Katayama, 2013).

The mechanisms of control also involve the methylation state of *ori2*, which prevents the replication restart during the same cell cycle (Demarre and Chattoraj, 2010). Contrary to the Chr1 origin, *ori1*, the Dam methylation of *ori2* is strictly required for its replication initiation (Demarre and Chattoraj,

2010; Val et al., 2014). Indeed, a *dam* mutant of *V. cholerae* can survive only when Chr1 and Chr2 are fused (Val et al., 2014). *ori2* has an overrepresentation of Dam methylation sites and is thus subjected to sequestration by SeqA (**Figure 3B**) (Demarre and Chattoraj, 2010). The SeqA sequestration prevents the immediate re-initiation of the replication, as in the case of Chr1, by temporally inhibiting the full-methylation of the DNA and initiator binding. Thus, the RctB binding to the iterons, which is dependent on the DNA methylation, is integrated to the cell cycle contrary, to its binding to the 39-mers and 29-mer. This methylation binding balance is involved in the cell cycle control of the Chr2 replication initiation.

Integration of Iteron-Chromids Initiation Replication to the Cell Cycle

In *V. cholerae*, Chr2 replication initiation is delayed compared to Chr1 replication initiation. Chr2 replication initiation starts when 2/3 of the replication period is completed. Besides, as Chr2 has a size equal to the 1/3 of Chr1, the replication termination of the two replicons is synchronous (Rasmussen et al., 2007) (**Figure 4A**). Marker frequency analysis (MFA) of a wide selection of Vibrios, with large variations in Chr1 and Chr2 sizes, suggests that there is a selective pressure for a termination synchrony, despite the fact that the control of Chr2 replication is at the initiation level (Kemter et al., 2018). Furthermore, in mutants where Chr2 finishes replicating earlier than Chr1, no impact on fitness was detected (Val et al., 2016). However, in these mutants the Chr2 terminus region (*ter2*) was shown to relocate earlier to mid-cell than in the *wt*, and remained localized at mid-cell until late in the cell cycle (Val et al., 2016). Despite early Chr2 replication termination, *ter2* retention at mid-cell suggests a secondary safeguard. How and why *ter2* segregation is delayed and results in re-synchronization with the Chr1 terminus region (*ter1*) is unknown. The mechanism coordinating the synchronous termination of the two replicons is driven by a locus found on the main chromosome. In *V. cholerae*, this locus, a short non-coding DNA sequence, is bound *in vivo* by RctB (Baek and Chattoraj, 2014). It is localized in the right replichore at around 800 Kbp downstream from *ori1*, and presents no homology with previously described RctB binding sites (*e.g.,* 12-mer and 39-mer) (Baek and Chattoraj, 2014). In *V. cholerae*, the deletion of this locus induces growth defects linked to cell filamentation and Chr2 loss (Val et al., 2016). Interestingly, moving the *V. cholerae crtS* to different location along the main chromosome led to a change of replication initiation timing of the Chr2 (Val et al., 2016). Replication of this Chr1 site triggers the replication of Chr2, which initiate after a short delay corresponding to the time needed for the replication of 200 Kbp. Thus, this checkpoint locus was named *crtS* for "chromosome 2 replication triggering site". (Val et al., 2016) (**Figure 4A**). Besides, by employing chromosome conformation capture (3C) experiments, it has further been demonstrated that *ori2* and *crtS* are in a physical contact. These observations suggest that this *ori2* replication initiation regulatory mechanism could involve a structural interplay between Chr1 and Chr2 (Val et al., 2016). In *E. coli*, the presence of ectopic *V. cholerae*
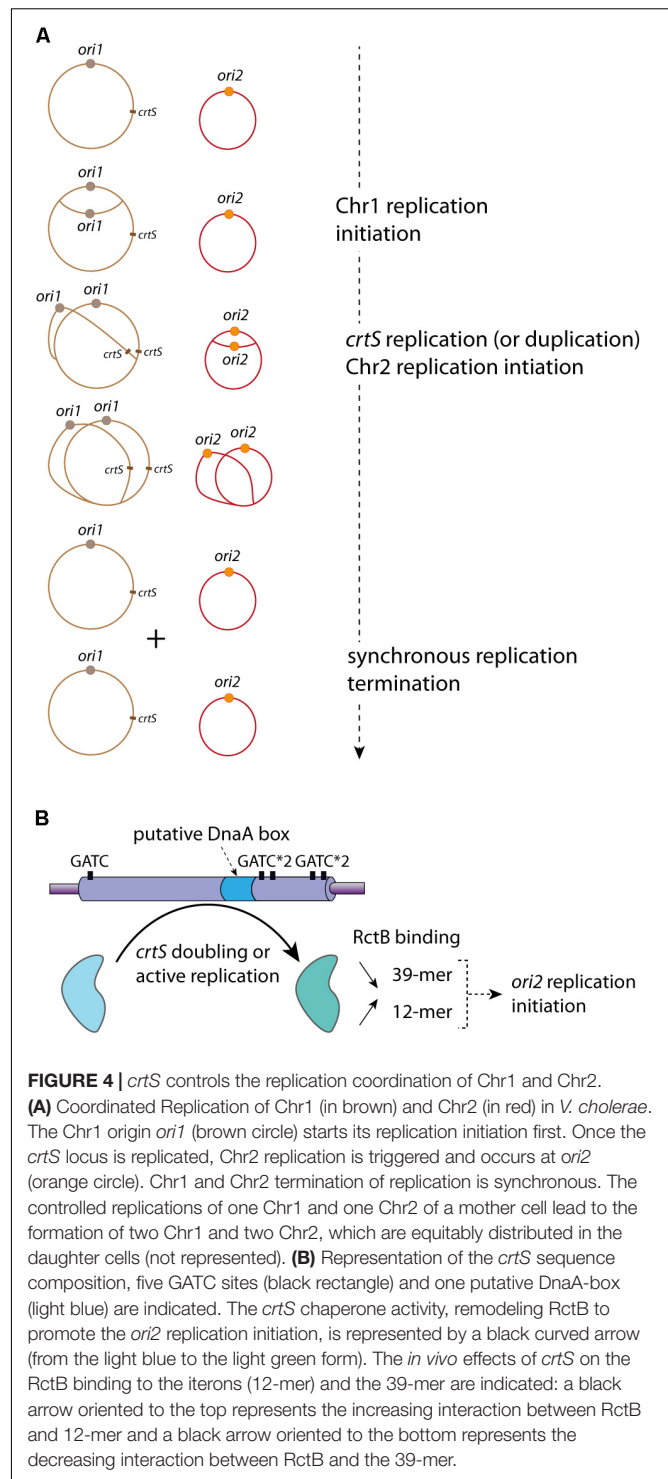

FIGURE 4 | *crtS* controls the replication coordination of Chr1 and Chr2. **(A)** Coordinated Replication of Chr1 (in brown) and Chr2 (in red) in *V. cholerae*. The Chr1 origin *ori1* (brown circle) starts its replication initiation first. Once the *crtS* locus is replicated, Chr2 replication is triggered and occurs at *ori2* (orange circle). Chr1 and Chr2 termination of replication is synchronous. The controlled replications of one Chr1 and one Chr2 of a mother cell lead to the formation of two Chr1 and two Chr2, which are equitably distributed in the daughter cells (not represented). **(B)** Representation of the *crtS* sequence composition, five GATC sites (black rectangle) and one putative DnaA-box (light blue) are indicated. The *crtS* chaperone activity, remodeling RctB to promote the *ori2* replication initiation, is represented by a black curved arrow (from the light blue to the light green form). The *in vivo* effects of *crtS* on the RctB binding to the iterons (12-mer) and the 39-mer are indicated: a black arrow oriented to the top represents the increasing interaction between RctB and 12-mer and a black arrow oriented to the bottom represents the decreasing interaction between RctB and the 39-mer.

or *Vibrio nigripulchritudo crtS* increase the copy number of plasmids carrying different *ori2*, from *Vibrio tubiashi* or *Vibrio furnissi*. However, the copy number of plasmids containing the *ori2* of *Photobacterium profundum*, *Vibrio vulnificus*, or *Vibrio harveyi*, is not increased when *crtS* from other species (*e.g.,* *V. cholerae crtS* and *V. parahaemolyticus crtS*) are provided *in*

trans (Kemter et al., 2018). These discrepancies could be due to the independence of the *P. profundum*, *V. vulnificus*, and *V. harveyi ori2* from *crtS* to regulate their replication, or to a species-specific mechanism. Thus, the *crtS* control activity is conserved, and *crtS* sites of divergent *Vibrio* species seem, to a certain extent, to be interchangeable for triggering the *ori2* replication initiation, showing a loose *crtS* species-specific activity (Kemter et al., 2018).
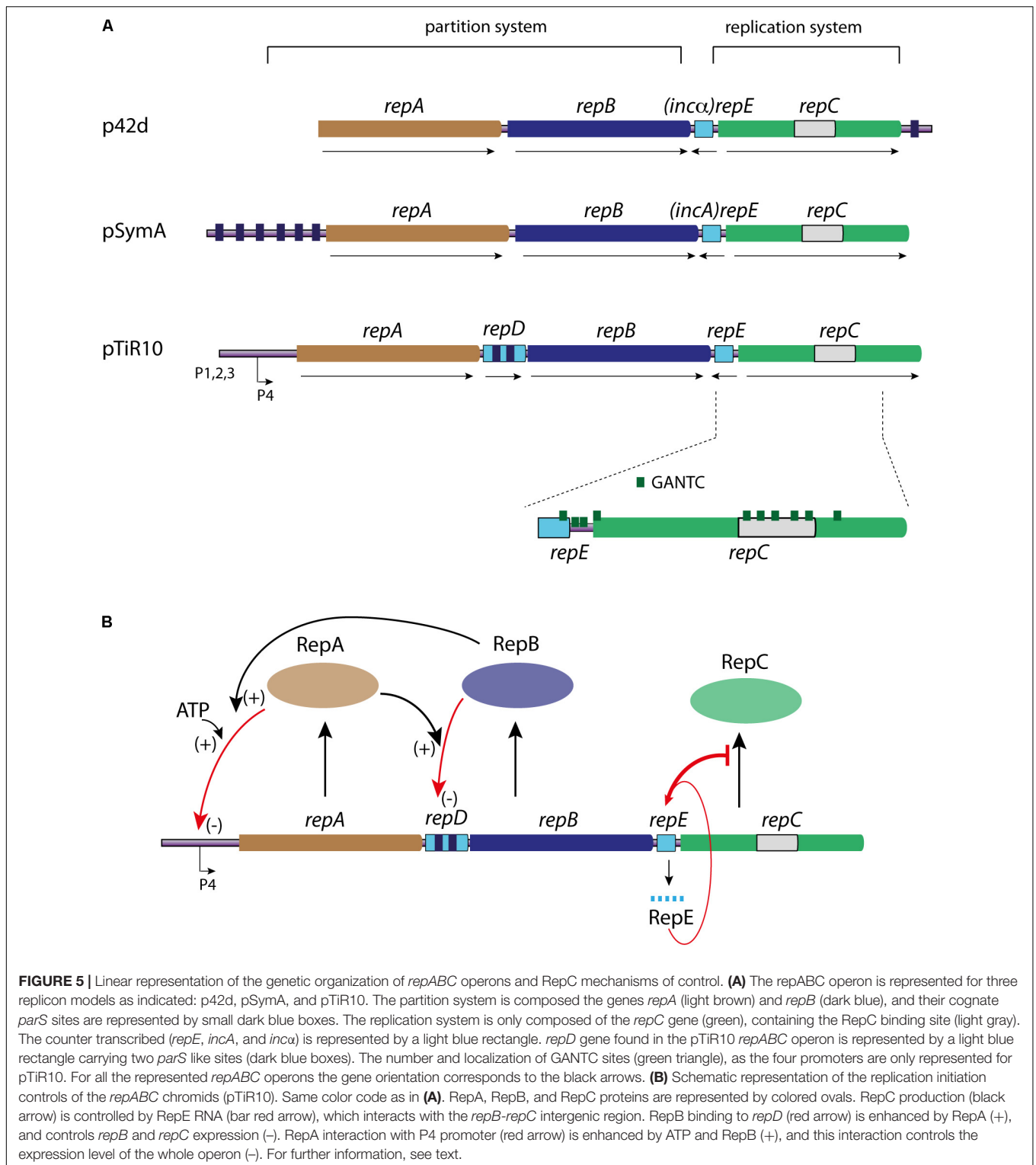
The alignment of different *crtS* sites shows a high sequence conservation among *Vibrionaceae*, including several GATC sites and a putative DnaA binding site (Baek and Chattoraj, 2014; Kemter et al., 2018) (**Figure 4B**). The RctB binding to *crtS* is hardly detected *in vitro* by DnaseI footprint experiments or by electrophoretic mobility shift assay (Baek and Chattoraj, 2014). It was proposed that, in *E. coli*, the *crtS* presence remodel RctB, decreasing its affinity for the 39-mer and conversely increasing it for the 12-mer (Baek and Chattoraj, 2014) (**Figure 4B**). This was drawn from *in vivo* data, but the *in vitro* experiments (electrophoretic mobility shift assay) did not allow obtaining clear results. Indeed, the authors observed only an *in vitro* decrease of RctB affinity to the 39-mer in presence of *crtS*, which could also reflect the competition between two types of RctB binding sites (Baek and Chattoraj, 2014). Thus, from these results it is difficult to differentiate a simple competition from an *in vitro crtS* remodeling activity. Moreover, in *E. coli* the presence of *crtS* makes DnaKJ dispensable for replication of *ori2* based plasmid (Baek and Chattoraj, 2014). This result, in addition to the effect of *crtS* on the RctB/DNA (12-mer and 39-mer) interactions, suggests a *crtS* DNA chaperone activity, which, by remodeling RctB, promotes Chr2 replication initiation. The *crtS* activity triggering *ori2* replication initiation is independent on methylation state of its GATC sites (de Lemos et al., in rev). However, the *crtS* form responsible for the DNA chaperone activity is still unknown. The passage of the replication fork across *crtS* would induce the formation of transient hemimethylated GATC sites, and the hemimethylated *crtS* may impact the RctB binding. Passage of the replication complex also generates single stranded DNA on the template of the lagging strand synthesis and could allow the formation of DNA hairpin. Thus, replication of *crtS* and the supposed DNA modifications it induces may be responsible for the *crtS* DNA chaperone activity. Nevertheless, the replication of *crtS* could simply lead to the duplication of the site, which could change the balance of free active RctB to catalyze the *ori2* opening. When already two copies of *crtS* were inserted on Chr1, Chr2 copy number was doubled suggesting that it is the presence of two *crtS* sites (after replication) that is important (Val et al., 2016). Indeed, a recent paper shows that the *crtS* duplication, without active replication, is sufficient to initiate *ori2* replication initiation (Ramachandran et al., 2018). However, it seems difficult to explain the *crtS* DNA chaperone activity solely from doubling its gene dosage. Further experimental data are needed to understand if either the active replication or the duplication of *crtS* is the signal controlling Chr2 replication initiation.

In conclusion, the molecular mechanisms by which the replication of *crtS* triggers the initiation of Chr2 through RctB are largely unknown. In *E. coli*, several mechanisms are responsible

for the coordinated initiation of multiple origins (DnaA titration, regulatory inactivation of DnaA, origin sequestration and DnaA reactivation sequences) (Hansen and Atlung, 2018). All these mechanisms control the availability of the active form of DnaA in initiating replication from *oriC*. If the control of *ori2* initiation by *crtS* was performed only by controlling the availability of the RctB active form, we would expect a similar synchrony in the firing of multiple *ori2* and this would be observed by cells containing only 2^n *ori2* foci (*e.g.,* two or four). However, using cells with two chromosomal copies of *crtS*, the duplication of one *crtS* triggers the firing of only one *ori2* (Val et al., 2016). This suggests that Chr2 initiation firing may necessitate a contact between *crtS* and *ori2*. The contacts between *ori2* and Chr1, introduced above, may be caused by the simultaneous binding of RctB to *ori2* and *crtS* (Val et al., 2016). The most frequent contacts between *ori2* and Chr1 occur immediately downstream of *crtS*. A possible explanation is that, following the duplication of the *crtS* locus, the replication machineries of Chr1 and Chr2 are maintained in the vicinity of each other until the end of replication of the two chromosomes. Non-replicating cells (*i.e.,* stationary phase) lose the contacts observed between Chr1 and Chr2 replichores during exponential growth, suggesting that replication is indeed responsible for the contacts of the two chromosomes along their chromosomal arms. Overall, the 3C analysis of the *V. cholerae* chromosomes points to a direct interplay between 3D organization and replication regulation. How *trans* topological contacts would drive a functional interaction between the two chromosomes remains unknown.

REPABC CHROMIDS REPLICATION MECHANISMS AND CONTROLS

The genetic information of alpha-proteobacteria is commonly carried by a multipartite genome (Landeta et al., 2011; diCenzo and Finan, 2017). Whatever their nature, megaplasmids or chromids, the replication and segregation of those replicons involve, in most cases, three genes organized in operon: *repA, repB, and repC* (Galibert et al., 2001; Cevallos et al., 2008; García-de Los Santos et al., 2008; Petersen et al., 2013) (**Figure 5A**). The proteins encoded by the *repABC* operon are involved in two distinct mechanisms; RepC is essential for replication, and RepA and RepB are dispensable for replication but required for the partition. The *repA*, *repB,* and *repC* genes are expressed from promoters found upstream of *repA*. Most of our knowledge about the transcriptional regulation of the *repABC* operon comes from the *A. tumefaciens* megaplasmid pTiR10, where data show that *repABC* transcription is regulated by environmental cues (Ramírez-Romero et al., 2001; Pappas and Winans, 2003a,b). Indeed, the pTiR10 *repABC* operon contains four promoters (**Figure 5A**). The promoter P4 ensures the basal expression of the operon, but this promoter can be activated by the regulator VirG once phosphorylated by VirA, in response to plant pheromones (Cho and Winans, 2005). Furthermore, the pTiR10 four promoters are activated by the LuxR-family quorum sensing system (Pappas and Winans, 2003a).

FIGURE 5 | Linear representation of the genetic organization of *repABC* operons and RepC mechanisms of control. **(A)** The repABC operon is represented for three replicon models as indicated: p42d, pSymA, and pTiR10. The partition system is composed the genes *repA* (light brown) and *repB* (dark blue), and their cognate *parS* sites are represented by small dark blue boxes. The replication system is only composed of the *repC* gene (green), containing the RepC binding site (light gray). The counter transcribed (*repE*, *incA*, and *incα*) is represented by a light blue rectangle. *repD* gene found in the pTiR10 *repABC* operon is represented by a light blue rectangle carrying two *parS* like sites (dark blue boxes). The number and localization of GANTC sites (green triangle), as the four promoters are only represented for pTiR10. For all the represented *repABC* operons the gene orientation corresponds to the black arrows. **(B)** Schematic representation of the replication initiation controls of the *repABC* chromids (pTiR10). Same color code as in **(A)**. RepA, RepB, and RepC proteins are represented by colored ovals. RepC production (black arrow) is controlled by RepE RNA (bar red arrow), which interacts with the *repB-repC* intergenic region. RepB binding to *repD* (red arrow) is enhanced by RepA (+), and controls *repB* and *repC* expression (–). RepA interaction with P4 promoter (red arrow) is enhanced by ATP and RepB (+), and this interaction controls the expression level of the whole operon (–). For further information, see text.

The Replication Initiator: RepC

RepC proteins are considered as the initiator protein of the *repABC* replicons and are found only in the alpha-proteobacteria (Palmer et al., 2000; Petersen et al., 2009). The *repC* gene alone is able to replicate a plasmid, showing that the origin is localized inside *repC* (Cevallos et al., 2008; Cervantes-Rivera et al., 2011; Pinto et al., 2011). At the structural level, the origin of the *repABC* replicons are lacking iterons and DnaA-boxes (Cervantes-Rivera et al., 2011; Pinto et al., 2012; Rajewska et al., 2012). The purified pTiR10 RepC binds to a 150 nt region containing an imperfect

dyad near an AT-rich region. This sequence is localized in the middle of the *repC* coding sequence (**Figure 5A**) (Cervantes-Rivera et al., 2011; Pinto et al., 2011). RepC binds it cooperatively with a high specificity. Indeed, overexpression of RepC in *A. tumefaciens* induces an increase in plasmid copy number in *cis*, but does not change copy number of plasmids containing a parental origin in *trans*. Thus, RepC functions only in *cis*. The same phenomenon is observed for the RepC protein of the *R. elti* p42d replicon (Cervantes-Rivera et al., 2011; Pinto et al., 2011). RepC exhibits no homology with other replication initiators. Its predicted secondary structure suggests that RepC is divided in two domains: an amino-terminal (NTD) domain from residues 1 to 265 and a carboxy-terminal (CTD) domain from residues 298 to 439. The two NTD and CTD domains are connected by a linker peptide comprising 30 hydrophilic amino acids peptide (Pinto et al., 2011). The NTD domain of pTir10 RepC is essential for DNA binding but poorly contributes to the binding specificity, *a contrario*, the CTD domain is unable to bind the DNA alone but allows the discrimination between specific and non-specific binding (Pinto et al., 2011). Finally, in the case of the p42d RepC, the last 39 amino acids residues are shown to be involved in the incompatibility phenotype (Cervantes-Rivera et al., 2011). Inside the NTD domain, the region spanning residues 26–158 exhibits a structural similarity with the MarR family of transcription factors and is sufficient to bind the DNA. MarR binds DNA as a dimer, *via* a helix-turn helix (HTH) motif, suggesting that RepC can bind the DNA also as a dimer (Pinto et al., 2011). The supposed dimerization of RepC *via* its CTD domain has been proposed to play a role in the incompatibility between *repABC* replicons (del Solar et al., 1998; Cervantes-Rivera et al., 2011).

Partition System and Replication Regulation

The control of replication initiation catalyzed by RepC is dependent of two major mechanisms, which both act on the *repC* expression level. These mechanisms involve the proteins RepA and RepB on one hand, and an antisense RNA on the other (**Figure 5**). RepA and RepB are members of the ParA and ParB families of partitioning proteins, and follow the same general mechanism of action (Williams and Thomas, 1992; Ramírez-Romero et al., 2000). The position and number of *parS* centromere-like site vary widely in the *repABC* replicons family (**Figure 5A**). These sites are essential for plasmid stability and are involved in the incompatibility mechanism between parental plasmids (MacLellan et al., 2006). Indeed, point mutations in the *parS* sites upstream the *repA2* of pSymA reduce the RepB binding and impede the incompatibility between pSymA parental plasmids. This incompatibility is presumably due to the competition between the two parental plasmids for the same partitioning system. RepA and RepB, together with *parS* sites, also participate to the negative transcriptional regulation of the operon, and thus act on the replication control of *repABC* replicons. Indeed, RepA binds to the *parS* sites and this binding may be enhanced by the presence of RepB and ATP. As an example, the RepA protein of pTiR10 auto-represses the P4 promoter, which is located within a 70 nt region protected

against DnaseI digestion by RepA (Pappas and Winans, 2003b) (**Figure 5B**). Some bacteria belonging to the alpha-proteobacteria may have up to six *repABC* replicons; and the question of the RepA and RepB specific activity at their cognate sites and not at heterologous sites is still open. Two given RepA proteins share no more than 61% of identity and RepB proteins no more than 51%, this may be a key for avoiding cross interactions (incompatibility) (Cevallos et al., 2008; Castillo-Ramírez et al., 2009; Pinto et al., 2012). Thus, the high specific interactions between RepA, RepB and their cognate binding *parS* sites, together with proteins evolution and divergence, likely allow the coexistence of multiple *repABC* replicons in the same bacteria (Żebracki et al., 2015; Koper et al., 2016). In pTiR10-like replicons, a fourth transcribed and translated gene, *repD*, is located between *repA* and *repB* genes and contains two RepB binding sites (*parS*) (Chai and Winans, 2005b) (**Figures 5A,B**). It seems that the RepD protein is not involved in the replication and partition of pTiR10-like replicons (Chai and Winans, 2005b). The RepB binding to *repD* is enhanced by the presence of RepA. *repD* coding sequence is involved in the plasmid partitioning and negatively regulates *repB* and *repC* expression, adding another level of control to replication initiation (**Figure 5B**) (Chai and Winans, 2005b).

In addition to the negative regulation of the operon transcription by RepA and RepB, an antisense RNA also negatively regulates RepC (**Figures 5A,B**). This locus, located between *repB* and *repC*, encodes a 50 nucleotides antisense RNA (ctRNA) (Venkova-Canova et al., 2004; Chai and Winans, 2005a; MacLellan et al., 2005). This ctRNA includes a predicted stem-loop, which can act as a transcription terminator and form a complex with the *repABC* mRNA within the *repB-repC* intergenic region (Chai and Winans, 2005a) (**Figures 5A,B**). This RNA, known as RepE in pTir10, is conserved in most, if not all, replicons belonging to the *repABC* family (Cevallos et al., 2008). The RepE action model, proposed for the *A. tumefaciens* pTiR10 replicon by Chain and Winans, and supported by Cervantes-Rivera and collaborators for the *R. elti* p42d replicon, can be easily applicable to the other *repABC* replicons. In this model, the *repABC* mRNA can adopt two alternative secondary structures in the *repB-repC* intergenic region, depending to the presence or absence of RepE. In the absence of RepE, the intergenic region *repB-repC* is predicted to fold in a large stem-loop, leaving the *repC* Shine-Dalgarno sequence and its initiation codon single stranded, thus permitting the *repC* translation. In presence of RepE, its interaction with the target mRNA induces the re-folding of the sequence downstream of the interaction site, and creates two new stem-loops. One of the new stem-loops forms a Rho-independent termination site upstream of the *repC* ribosome-binding site leading to a premature termination (**Figure 5B**) (Chai and Winans, 2005a; Cervantes-Rivera et al., 2010). The *repB-repC* intergenic region, containing RepE, is also involved in the incompatibility between parental plasmids, and RepE was also named *incA* or *incα* in plasmids pSymA and p42d, respectively (**Figure 5A**) (Ramírez-Romero et al., 2000; Soberón et al., 2004; MacLellan et al., 2005). Mutations reducing the RepE expression or remodeling its structure have been indeed shown to decrease the incompatibility (Chai and Winans, 2005a; Venkova-Canova et al., 2006; Pinto et al., 2012). All together, these mechanisms, *i.e*:

the RepE ctRNA and the RepA/RepB negative regulation bring a fine tuning of the RepC expression level and thus control the replication initiation of the *repABC* replicons.

Integration of *repABC*-Chromids Replication to the Cell Cycle

The replication and segregation of the alpha-proteobacteria multipartite genomes containing a *repABC* chromid is poorly documented. Nevertheless, the comparison of the data obtained for the bacteria *A. tumefaciens*, *S. meliloti*, and *B. abortus*, suggests the existence of a coordination mechanism for their two or three replicons (Kahng and Shapiro, 2003; Deghelt et al., 2014; Frage et al., 2016). The genome of *B. abortus* is divided in two replicons: the 2.1 Mbp chromosome and the 1.2 Mbp *repABC* chromid. The two replicons of *B. abortus* are oriented along the cell length axis, and the chromosome origin displays a bipolar orientation after its replication initiation, contrary to the chromid origin, which drift apart during the cell cycle and displays no sign of polar attachment (Deghelt et al., 2014). This last observation is similar to the results obtained for the *repABC* replicons of *A. tumefaciens* and *S. meliloti* (Kahng and Shapiro, 2003). Furthermore, the origin duplication of the *B. abortus* chromid occurs after the chromosome origin duplication and segregation of the chromid terminal region occurs before cell septation, while chromosome terminal region segregation is observed at the time of cell constriction. In the tripartite genome bacterium, *S. meliloti*, the partitioning of the three replicons (chromosomes, pSymA and pSymB) follows a highly conserved temporal order. The replication of the three replicons occurs once per cell cycle, and the segregation pattern is such that the chromosome segregates first, followed then by pSymA, and then by pSymB (Frage et al., 2016). Interestingly, the pSymA *repABC* region is sufficient to confer the spatiotemporal behavior of this replicon to a small plasmid. Besides, alterations of the DnaA activity, either positively or negatively, only impact the chromosome replication, and have no effect on the secondary replicons replication (Frage et al., 2016). Thus, it is likely that the strict timing of replication and segregation of *repABC* replicons only involve genetic components located within the repABC operon.

Finally, compared to the *V. cholerae* Chr2, there are no direct evidences of a subservient interplay between two replicons in the same cell in the alpha-proteobacteria, and thus no described mechanism. Nonetheless, the origin of replication and the promoters of the counter-transcribed *repE* gene of *repABC* chromids and mega-plasmids are rich in GANTC, which correspond to the Cell cycle-regulated Methylase (CcrM) methylation sites. In the alpha-proteobacteria *C. crescentus*, the A base of GANTC sites is methylated by CcrM (Marczynski and Shapiro, 2002; Wion and Casadesús, 2006). CcrM is functionally related to the *E. coli* methylase Dam, but there are important differences between them. Indeed, compared to Dam, which is active throughout the cell cycle, CcrM is synthesized and active only in predivisional cells. Unlike Dam, CcrM is not required for replication initiation or DNA mismatch repair (Gonzalez et al., 2014). However, CcrM overexpression results

in abnormal chromosomes content per cell in *C. crescentus*. Thus, CcrM is essential for normal chromosomal replication. *C. crescentus* chromosome replicates once per cell cycle, and this seems to be controlled by the CcrM system (Stephens et al., 1996; Marczynski, 1999; Collier, 2012). CcrM is conserved across the alpha-proteobacteria and its orthologs has been studied in *S. meliloti*, *B. abortus*, and *A. tumefaciens* (Wright et al., 1997; Robertson et al., 2000; Kahng and Shapiro, 2001). Interestingly, with the notable exception of *C. crescentus*, the methylation of GANTC sites by CcrM seems to be essential in the other alpha-proteobacteria (Brilli et al., 2010; Fioravanti et al., 2013; Mohapatra et al., 2014). In the alpha-proteobacteria, a conserved master regulator, named CtrA, is involved in the control of the cell division and takes part in the spatio-temporal regulation of the replication initiation linked to the cell cycle (Wolański et al., 2014; Pini et al., 2015; Francis et al., 2017). CtrA is involved in the regulation of *ccrM* expression in both *C. crescentus* and *A. tumefaciens*, and it is likely the case in the other alpha-proteobacteria (Quon et al., 1996; Kahng and Shapiro, 2001). Therefore, the methylation state of the GANTC sites found in the *repABC* operon (*e.g.,* pTiR10) could be timely controlled, impacting the *repC* expression and *repE* transcription, and bringing a cell cycle integrated regulation of the *repABC* replication initiation. However, the *in vitro* binding of RepC to the origin is independent on the DNA methylation state (Pinto et al., 2011), but this does not exclude that other, yet unknown, replication factors might have a binding activity dependent on the GANTC methylation in the origin of *repABC* chromids.

CONCLUSION AND REMARKS

In order to permit a faithful transmission of the genetic information, but also to avoid any problems due to polyploidy, chromids have to be replicated once and only once per cell cycle. In this review, we gave a short overview of chromid domestication history, and further focused our analysis on their replication and how they became integrated in the bacterial cell cycle. Most of our knowledge on chromid replication initiation comes from the *repABC* and iteron models, where controls mainly occur at the initiation step. Both types of chromid present multi-scale mechanisms to timely manage the replication initiation of the replicon, which first involves the recognition of the replication origin by the initiator protein (RepC/RctB). These controls are mostly centered on the initiator proteins both at the gene expression level and through the regulation of their specific activities. This first step is already controlled by diverse and numerous mechanism. Thus, iterons and *repABC* chromids seem to correspond to two different evolutionary ways of achieving a tight replication initiation control.

One of the mechanisms to avoid over-replication of iteron chromids is dependent on the Dam/SeqA couple. There is no SeqA homolog in the alpha-proteobacteria, but a yet unknown protein could play an analogous function of sequestration (Pinto et al., 2012). Besides, usually all the large replicons found in the alpha-proteobacteria

carry a *repABC* operon, while iterons origins are only found in small plasmids in these bacteria (Szymanik et al., 2006). These observations suggest that iteron chromids, which are SeqA dependent, could not allow the tight replication initiation control of the alpha-proteobacteria megaplasmids and chromids.

After the initiator/origin interaction, the following steps, which correspond to the unwinding of the AT-rich region and to the recruitment of the replisome proteins, might also be regulated, but this has not been studied yet. In the case of the iterons plasmids, the recruitment and loading of the helicase DnaB involves a direct interaction of the helicase with DnaA and/or the plasmid initiator (Zhong et al., 2003; Wegrzyn et al., 2016). The interaction between RctB and DnaB has to be shown, as well as the DnaA binding to *ori2*, and its involvement in the DnaB loading. On the contrary, the *repABC* chromid origins do not contain DnaA boxes and thus it is tempting to think that RepC proteins interact directly with the helicase.

An important feature distinguishing the replication control of the two types of chromid is based on the existence (or not) of controls driven by other replicons. Indeed, the *repABC* multi-scale controls seem to be strictly intra-molecular, meaning that all the necessary sequences are carried by the replicon and located within the *repABC* operon (Frage et al., 2016). The results obtained with *B. abortus* and *A. tumefaciens* chromids reveal that these chromids initiate their replication once per cell cycle and after the chromosome (Kahng and Shapiro, 2003; Deghelt et al., 2014). This raises the question of how the *repABC* chromids can be replicated in synchrony with the main chromosome. In contrast, replication control of *Vibrio* iteron chromids involves an inter-molecular interaction (Val et al., 2016). The recent discovery of *crtS* and of the physical contacts between Chr1 and Chr2 reveals a unique checkpoint control of replication in bacteria (Baek and Chattoraj, 2014; Val et al., 2016). The determinants of this contact between Chr1 and Chr2 still have to be identified. Contacts between *crtS* and *ori2* may alter RctB binding and handcuffing activity, or other unknown process involved in Chr2 replication initiation. This new checkpoint implies a transfer of information between the two replicons, which apparently take a time equivalent to the replication of 200 Kbp (Val et al., 2016). This temporal delay

corresponds to the time necessary to deliver the message of the *crtS* replication to the *ori2*, allowing to the remodeling of RctB activities and to the recruitment of the replisome, but the precise events and players involved in it, have yet to be determined.

At the moment, the reasons for the requirement of a replication delay for secondary replicons remains unknown. In *V. cholerae*, initiation of replication of Chr2 is delayed such that replication termination of Chr1 and Chr2 occurs at the same time. This could facilitate the coordination of the final steps of segregation before cell division. The location of *crtS* is highly conserved within the *Vibrio*. The *crtS* position may have been selected throughout evolution by the constraint imposed by this activation delay. The importance of multiple chromosomes to coordinate their replication and the importance for Chr1 and Chr2 to finish replicating at the same time remains in the realm of conjecture.

REFERENCES

Agnoli, K., Schwager, S., Uehlinger, S., Vergunst, A., Viteri, D. F., Nguyen, D. T., et al. (2012). Exposing the third chromosome of *Burkholderia cepacia* complex strains as a virulence plasmid. *Mol. Microbiol.* 83, 362–378. doi: 10.1111/j.1365-2958.2011.07937.x

Akman, L., Yamashita, A., Watanabe, H., Oshima, K., Shiba, T., Hattori, M., et al. (2002). Genome sequence of the endocellular obligate symbiont of tsetse flies, *Wigglesworthia glossinidia*. *Nat. Genet.* 32, 402–407. doi: 10.1038/ng986

Allardet-Servent, A., Michaux-Charachon, S., Jumas-Bilak, E., Karayan, L., and Ramuz, M. (1993). Presence of one linear and one circular chromosome in the *Agrobacterium tumefaciens* C58 genome. *J. Bacteriol.* 175, 7869–7874. doi: 10.1128/jb.175.24.7869-7874.1993

Ausiannikava, D., Mitchell, L., Marriott, H., Smith, V., Hawkins, M., Makarova, K. S., et al. (2018). Evolution of genome architecture in Archaea: spontaneous generation of a new chromosome in *Haloferax volcanii*. *Mol. Biol. Evol.* doi: 10.1093/molbev/msy075 [Epub ahead of print].

Baek, J. H., and Chattoraj, D. K. (2014). Chromosome I controls chromosome II replication in *Vibrio cholerae*. *PLoS Genet.* 10:e1004184. doi: 10.1371/journal.pgen.1004184

Baril, C., Richaud, C., Baranton, G., and Saint Girons, I. S. (1989). Linear chromosome of *Borrelia burgdorferi*. *Res. Microbiol.* 140, 507–516. doi: 10.1016/0923-2508(89)90083-1

Bavishi, A., Lin, L., Schroeder, K., Peters, A., Cho, H., and Choudhary, M. (2010). The prevalence of gene duplications and their ancient origin in *Rhodobacter sphaeroides* 2.4.1. *BMC Microbiol.* 10:331. doi: 10.1186/1471-2180-10-331

Brantl, S. (2014). Plasmid replication control by antisense RNAs. *Microbiol. Spectr.* 2:LAS-0001-2013. doi: 10.1128/microbiolspec.PLAS-0001-2013

Brilli, M., Fondi, M., Fani, R., Mengoni, A., Ferri, L., Bazzicalupo, M., et al. (2010). The diversity and evolution of cell cycle regulation in alpha-proteobacteria: a comparative genomic analysis. *BMC Syst. Biol.* 4:52. doi: 10.1186/1752-0509-4-52

Buchrieser, C., Glaser, P., Rusniok, C., Nedjari, H., D'Hauteville, H., Kunst, F., et al. (2000). The virulence plasmid pWR100 and the repertoire of proteins

secreted by the type III secretion apparatus of *Shigella flexneri*. *Mol. Microbiol.* 38, 760–771. doi: 10.1046/j.1365-2958.2000.02179.x

Castillo-Ramírez, S., Vázquez-Castellanos, J. F., González, V., and Cevallos, M. A. (2009). Horizontal gene transfer and diverse functional constrains within a common replication-partitioning system in *Alphaproteobacteria*: the *repABC* operon. *BMC Genomics* 10:536. doi: 10.1186/1471-2164-10-536

Cervantes-Rivera, R., Pedraza-López, F., Pérez-Segura, G., and Cevallos, M. A. (2011). The replication origin of a repABC plasmid. *BMC Microbiol.* 11:158. doi: 10.1186/1471-2180-11-158

Cervantes-Rivera, R., Romero-López, C., Berzal-Herranz, A., and Cevallos, M. A. (2010). Analysis of the mechanism of action of the antisense RNA that controls the replication of the repABC plasmid p42d. *J. Bacteriol.* 192, 3268–3278. doi: 10.1128/JB.00118-10

Cevallos, M. A., Cervantes-Rivera, R., and Gutiérrez-Ríos, R. M. (2008). The repABC plasmid family. *Plasmid* 60, 19–37. doi: 10.1016/j.plasmid.2008.03.001

Chai, Y., and Winans, S. C. (2005a). A small antisense RNA downregulates expression of an essential replicase protein of an *Agrobacterium tumefaciens* Ti plasmid. *Mol. Microbiol.* 56, 1574–1585. doi: 10.1111/j.1365-2958.2005.04636.x

Chai, Y., and Winans, S. C. (2005b). RepB protein of an *Agrobacterium tumefaciens* Ti plasmid binds to two adjacent sites between repA and repB for plasmid partitioning and autorepression. *Mol. Microbiol.* 58, 1114–1129. doi: 10.1111/j.1365-2958.2005.04886.x

Chao, M. C., Pritchard, J. R., Zhang, Y. J., Rubin, E. J., Livny, J., Davis, B. M., et al. (2013). High-resolution definition of the *Vibrio cholerae* essential gene set with hidden Markov model-based analyses of transposon-insertion sequencing data. *Nucleic Acids Res.* 41, 9033–9048. doi: 10.1093/nar/gkt654

Chattoraj, D. K. (2000). Control of plasmid DNA replication by iterons: no longer paradoxical. *Mol. Microbiol.* 37, 467–476. doi: 10.1046/j.1365-2958.2000.01986.x

Cheng, J., Sibley, C. D., Zaheer, R., and Finan, T. M. (2007). A *Sinorhizobium meliloti minE* mutant has an altered morphology and exhibits defects in legume symbiosis. *Microbiology* 153, 375–387. doi: 10.1099/mic.0.2006/001362-0

Cho, H., and Winans, S. C. (2005). VirA and VirG activate the Ti plasmid repABC operon, elevating plasmid copy number in response to wound-released chemical signals. *Proc. Natl. Acad. Sci. U.S.A.* 102, 14843–14848. doi: 10.1073/pnas.0503458102

Choudhary, M., Mackenzie, C., Nereng, K., Sodergren, E., Weinstock, G. M., and Kaplan, S. (1997). Low-resolution sequencing of *Rhodobacter sphaeroides* 2.4.1T: chromosome II is a true chromosome. *Microbiology* 143(Pt 10), 3085–3099. doi: 10.1099/00221287-143-10-3085

Collier, J. (2012). Regulation of chromosomal replication in *Caulobacter crescentus*. *Plasmid* 67, 76–87. doi: 10.1016/j.plasmid.2011.12.007

Deghelt, M., Mullier, C., Sternon, J.-F., Francis, N., Laloux, G., Dotreppe, D., et al. (2014). G1-arrested newborn cells are the predominant infectious form of the pathogen *Brucella abortus*. *Nat. Commun.* 5:4366. doi: 10.1038/ncomms5366

del Solar, G., Giraldo, R., Ruiz-Echevarría, M. J., Espinosa, M., and Díaz-Orejas, R. (1998). Replication and control of circular bacterial plasmids. *Microbiol. Mol. Biol. Rev.* 62, 434–464.

Demarre, G., and Chattoraj, D. K. (2010). DNA adenine methylation is required to replicate both *Vibrio cholerae* chromosomes once per cell cycle. *PLoS Genet.* 6:e1000939. doi: 10.1371/journal.pgen.1000939

Díaz-López, T., Lages-Gonzalo, M., Serrano-López, A., Alfonso, C., Rivas, G., Díaz-Orejas, R., et al. (2003). Structural changes in RepA, a plasmid replication initiator, upon binding to origin DNA. *J. Biol. Chem.* 278, 18606–18616. doi: 10.1074/jbc.M212024200

diCenzo, G., Milunovic, B., Cheng, J., and Finan, T. M. (2013). The tRNA[arg] gene and engA are essential genes on the 1.7-Mb pSymB megaplasmid of *Sinorhizobium meliloti* and were translocated together from the chromosome in an ancestral strain. *J. Bacteriol.* 195, 202–212. doi: 10.1128/JB.01758-12

diCenzo, G. C., and Finan, T. M. (2015). Genetic redundancy is prevalent within the 6.7 Mb *Sinorhizobium meliloti* genome. *Mol. Genet. Genomics* 290, 1345–1356. doi: 10.1007/s00438-015-0998-6

diCenzo, G. C., and Finan, T. M. (2017). The divided bacterial genome: structure, function, and evolution. *Microbiol. Mol. Biol. Rev.* 81, e00019-17. doi: 10.1128/MMBR.00019-17

Du, W.-L., Dubarry, N., Passot, F. M., Kamgoué, A., Murray, H., Lane, D., et al. (2016). Orderly replication and segregation of the four replicons of *Burkholderia cenocepacia* J2315. *PLoS Genet.* 12:e1006172. doi: 10.1371/journal.pgen.1006172

Duan, Y., Huey, J. D., and Herman, J. K. (2016). The DnaA inhibitor SirA acts in the same pathway as Soj (ParA) to facilitate oriC segregation during *Bacillus subtilis* sporulation. *Mol. Microbiol.* 102, 530–544. doi: 10.1111/mmi.13477

Duigou, S., Knudsen, K. G., Skovgaard, O., Egan, E. S., Løbner-Olesen, A., and Waldor, M. K. (2006). Independent control of replication initiation of the two *Vibrio cholerae* chromosomes by DnaA and RctB. *J. Bacteriol.* 188, 6419–6424. doi: 10.1128/JB.00565-06

Dziewit, L., Czarnecki, J., Wibberg, D., Radlinska, M., Mrozek, P., Szymczak, M., et al. (2014). Architecture and functions of a multipartite genome of the methylotrophic bacterium *Paracoccus aminophilus* JCM 7686, containing primary and secondary chromids. *BMC Genomics* 15:124. doi: 10.1186/1471-2164-15-124

Egan, E. S., Duigou, S., and Waldor, M. K. (2006). Autorepression of RctB, an initiator of *Vibrio cholerae* chromosome II replication. *J. Bacteriol.* 188, 789–793. doi: 10.1128/JB.188.2.789-793.2006

Egan, E. S., Fogel, M. A., and Waldor, M. K. (2005). Divided genomes: negotiating the cell cycle in prokaryotes with multiple chromosomes. *Mol. Microbiol.* 56, 1129–1138. doi: 10.1111/j.1365-2958.2005.04622.x

Egan, E. S., Løbner-Olesen, A., and Waldor, M. K. (2004). Synchronous replication initiation of the two *Vibrio cholerae* chromosomes. *Curr. Biol.* 14, R501–R502. doi: 10.1016/j.cub.2004.06.036

Egan, E. S., and Waldor, M. K. (2003). Distinct replication requirements for the two *Vibrio cholerae* chromosomes. *Cell* 114, 521–530. doi: 10.1016/S0092-8674(03)00611-1

Felletti, M., Omnus, D. J., and Jonas, K. (2018). Regulation of the replication initiator DnaA in *Caulobacter crescentus*. *Biochim. Biophys. Acta* doi: 10.1016/j.bbagrm.2018.01.004 [Epub ahead of print].

Fioravanti, A., Fumeaux, C., Mohapatra, S. S., Bompard, C., Brilli, M., Frandi, A., et al. (2013). DNA binding of the cell cycle transcriptional regulator GcrA depends on N6-adenosine methylation in *Caulobacter crescentus* and other *Alphaproteobacteria*. *PLoS Genet.* 9:e1003541. doi: 10.1371/journal.pgen.1003541

Frage, B., Döhlemann, J., Robledo, M., Lucena, D., Sobetzko, P., Graumann, P. L., et al. (2016). Spatiotemporal choreography of chromosome and megaplasmids in the *Sinorhizobium meliloti* cell cycle. *Mol. Microbiol.* 100, 808–823. doi: 10.1111/mmi.13351

Francis, N., Poncin, K., Fioravanti, A., Vassen, V., Willemart, K., Ong, T. A. P., et al. (2017). CtrA controls cell division and outer membrane composition of the pathogen *Brucella abortus*. *Mol. Microbiol.* 103, 780–797. doi: 10.1111/mmi.13589

Frank, O., Göker, M., Pradella, S., and Petersen, J. (2015a). Ocean's Twelve: flagellar and biofilm chromids in the multipartite genome of *Marinovum algicola* DG898 exemplify functional compartmentalization. *Environ. Microbiol.* 17, 4019–4034. doi: 10.1111/1462-2920.12947

Frank, O., Michael, V., Päuker, O., Boedeker, C., Jogler, C., Rohde, M., et al. (2015b). Plasmid curing and the loss of grip–the 65-kb replicon of *Phaeobacter inhibens* DSM 17395 is required for biofilm formation, motility and the colonization of marine algae. *Syst. Appl. Microbiol.* 38, 120–127. doi: 10.1016/j.syapm.2014.12.001

Friedman, D. I., Olson, E. R., Georgopoulos, C., Tilly, K., Herskowitz, I., and Banuett, F. (1984). Interactions of bacteriophage and host macromolecules in the growth of bacteriophage lambda. *Microbiol. Rev.* 48, 299–325.

Fujimitsu, K., Senriuchi, T., and Katayama, T. (2009). Specific genomic sequences of *E. coli* promote replicational initiation by directly reactivating ADP-DnaA. *Genes Dev.* 23, 1221–1233. doi: 10.1101/gad.1775809

Gaimster, H., and Summers, D. (2015). Plasmids in the driving seat: the regulatory RNA Rcd gives plasmid ColE1 control over division and growth of its *E. coli* host. *Plasmid* 78, 59–64. doi: 10.1016/j.plasmid.2014.11.002

Galardini, M., Pini, F., Bazzicalupo, M., Biondi, E. G., and Mengoni, A. (2013). Replicon-dependent bacterial genome evolution: the case of *Sinorhizobium meliloti*. *Genome Biol. Evol.* 5, 542–558. doi: 10.1093/gbe/evt027

Galibert, F., Finan, T. M., Long, S. R., Puhler, A., Abola, P., Ampe, F., et al. (2001). The composite genome of the legume symbiont *Sinorhizobium meliloti*. *Science* 293, 668–672. doi: 10.1126/science.1060966

García-de Los Santos, A., López, E., Cubillas, C. A., Noel, K. D., Brom, S., et al. (2008). Requirement of a plasmid-encoded catalase for survival of *Rhizobium etli* CFN42 in a polyphenol-rich environment. *Appl. Environ. Microbiol.* 74, 2398–2403. doi: 10.1128/AEM.02457-07

Gerding, M. A., Chao, M. C., Davis, B. M., and Waldor, M. K. (2015). Molecular dissection of the essential features of the origin of replication of the second *Vibrio cholerae* chromosome. *mBio* 6:e00973. doi: 10.1128/mBio.00973-15

Giraldo, R., Fernández-Tornero, C., Evans, P. R., Díaz-Orejas, R., and Romero, A. (2003). A conformational switch between transcriptional repression and replication initiation in the RepA dimerization domain. *Nat. Struct. Biol.* 10, 565–571. doi: 10.1038/nsb937

Gonzalez, D., Kozdon, J. B., McAdams, H. H., Shapiro, L., and Collier, J. (2014). The functions of DNA methylation by CcrM in *Caulobacter crescentus*: a global approach. *Nucleic Acids Res.* 42, 3720–3735. doi: 10.1093/nar/gkt1352

Goodner, B., Hinkle, G., Gattung, S., Miller, N., Blanchard, M., Qurollo, B., et al. (2001). Genome sequence of the plant pathogen and biotechnology agent *Agrobacterium tumefaciens* C58. *Science* 294, 2323–2328. doi: 10.1126/science.1066803

Guo, H., Sun, S., Eardly, B., Finan, T., and Xu, J. (2009). Genome variation in the symbiotic nitrogen-fixing bacterium *Sinorhizobium meliloti*. *Genome* 52, 862–875. doi: 10.1139/g09-060

Hansen, F. G., and Atlung, T. (2018). The DnaA tale. *Front. Microbiol.* 9:319. doi: 10.3389/fmicb.2018.00319

Harrison, P. W., Lower, R. P. J., Kim, N. K. D., and Young, J. P. W. (2010). Introducing the bacterial "chromid": not a chromosome, not a plasmid. *Trends Microbiol.* 18, 141–148. doi: 10.1016/j.tim.2009.12.010

Hartman, A. L., Norais, C., Badger, J. H., Delmas, S., Haldenby, S., Madupu, R., et al. (2010). The complete genome sequence of *Haloferax volcanii* DS2, a model archaeon. *PLoS One* 5:e9605. doi: 10.1371/journal.pone.0009605

Hase, M., Yoshimi, T., Ishikawa, Y., Ohba, A., Guo, L., Mima, S., et al. (1998). Site-directed mutational analysis for the membrane binding of DnaA protein. Identification of amino acids involved in the functional interaction between DnaA protein and acidic phospholipids. *J. Biol. Chem.* 273, 28651–28656. doi: 10.1074/jbc.273.44.28651

Heidelberg, J. F., Eisen, J. A., Nelson, W. C., Clayton, R. A., Gwinn, M. L., Dodson, R. J., et al. (2000). DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature* 406, 477–483. doi: 10.1038/35020000

Hirochika, H., and Sakaguchi, K. (1982). Analysis of linear plasmids isolated from *Streptomyces*: association of protein with the ends of the plasmid DNA. *Plasmid* 7, 59–65. doi: 10.1016/0147-619X(82)90027-0

Jha, J. K., Demarre, G., Venkova-Canova, T., and Chattoraj, D. K. (2012). Replication regulation of *Vibrio cholerae* chromosome II involves initiator binding to the origin both as monomer and as dimer. *Nucleic Acids Res.* 40, 6026–6038. doi: 10.1093/nar/gks260

Jha, J. K., Ghirlando, R., and Chattoraj, D. K. (2014). Initiator protein dimerization plays a key role in replication control of *Vibrio cholerae* chromosome 2. *Nucleic Acids Res.* 42, 10538–10549. doi: 10.1093/nar/gku771

Jha, J. K., Li, M., Ghirlando, R., Miller Jenkins, L. M., Wlodawer, A., and Chattoraj, D. (2017). The DnaK chaperone uses different mechanisms to promote and inhibit replication of *Vibrio cholerae* chromosome 2. *mBio* 8:e000427-17. doi: 10.1128/mBio.00427-17

Jumas-Bilak, E., Michaux-Charachon, S., Bourg, G., O'Callaghan, D., and Ramuz, M. (1998). Differences in chromosome number and genome rearrangements in the genus *Brucella*. *Mol. Microbiol.* 27, 99–106. doi: 10.1046/j.1365-2958.1998.00661.x

Kadoya, R., Baek, J. H., Sarker, A., and Chattoraj, D. K. (2011). Participation of chromosome segregation protein ParAI of *Vibrio cholerae* in chromosome replication. *J. Bacteriol.* 193, 1504–1514. doi: 10.1128/JB.01067-10

Kahng, L. S., and Shapiro, L. (2001). The CcrM DNA methyltransferase of *Agrobacterium tumefaciens* is essential, and its activity is cell cycle regulated. *J. Bacteriol.* 183, 3065–3075. doi: 10.1128/JB.183.10.3065-3075.2001

Kahng, L. S., and Shapiro, L. (2003). Polar localization of replicon origins in the multipartite genomes of *Agrobacterium tumefaciens* and *Sinorhizobium meliloti*. *J. Bacteriol.* 185, 3384–3391. doi: 10.1128/JB.185.11.3384-3391.2003

Kasho, K., and Katayama, T. (2013). DnaA binding locus datA promotes DnaA-ATP hydrolysis to enable cell cycle-coordinated replication initiation. *Proc. Natl. Acad. Scci. U.S.A.* 110, 936–941. doi: 10.1073/pnas.1212070110

Katayama, T., Ozaki, S., Keyamura, K., and Fujimitsu, K. (2010). Regulation of the replication cycle: conserved and diverse regulatory systems for DnaA and oriC. *Nat. Rev. Microbiol.* 8, 163–170. doi: 10.1038/nrmicro2314

Kawakami, H., Keyamura, K., and Katayama, T. (2005). Formation of an ATP-DnaA-specific initiation complex requires DnaA Arginine 285, a conserved motif in the AAA+ protein family. *J. Biol. Chem.* 280, 27420–27430. doi: 10.1074/jbc.M502764200

Kemter, F. S., Messerschmidt, S. J., Schallopp, N., Sobetzko, P., Lang, E., Bunk, B., et al. (2018). Synchronous termination of replication of the two chromosomes is an evolutionary selected feature in Vibrionaceae. *PLoS Genet.* 14:e1007251. doi: 10.1371/journal.pgen.1007251

Kitagawa, R., Ozaki, T., Moriya, S., and Ogawa, T. (1998). Negative control of replication initiation by a novel chromosomal locus exhibiting exceptional affinity for *Escherichia coli* DnaA protein. *Genes Dev.* 12, 3032–3043. doi: 10.1101/gad.12.19.3032

Koch, B., Ma, X., and Løbner-Olesen, A. (2010). Replication of *Vibrio cholerae* chromosome I in *Escherichia coli*: dependence on dam methylation. *J. Bacteriol.* 192, 3903–3914. doi: 10.1128/JB.00311-10

Koch, B., Ma, X., and Løbner-Olesen, A. (2012). rctB mutations that increase copy number of *Vibrio cholerae* oriCII in *Escherichia coli*. *Plasmid* 68, 159–169. doi: 10.1016/j.plasmid.2012.03.003

Komori, H., Matsunaga, F., Higuchi, Y., Ishiai, M., Wada, C., and Miki, K. (1999). Crystal structure of a prokaryotic replication initiator protein bound to DNA at 2.6 A resolution. *EMBO J.* 18, 4597–4607. doi: 10.1093/emboj/18.17.4597

Koper, P., Żebracki, K., Marczak, M., Skorupska, A., and Mazur, A. (2016). RepB proteins of the multipartite *Rhizobium leguminosarum* bv. *trifolii* genome discriminate between centromere-like parS sequences for plasmid segregational stability. *Mol. Microbiol.* 102, 446–466. doi: 10.1111/mmi.13472

Landeta, C., Dávalos, A., Cevallos, M. Á., Geiger, O., Brom, S., and Romero, D. (2011). Plasmids with a chromosome-like role in rhizobia. *J. Bacteriol.* 193, 1317–1326. doi: 10.1128/JB.01184-10

Lesic, B., Zouine, M., Ducos-Galand, M., Huon, C., Rosso, M.-L., Prévost, M.-C., et al. (2012). A natural system of chromosome transfer in *Yersinia pseudotuberculosis*. *PLoS Genet.* 8:e1002529. doi: 10.1371/journal.pgen.1002529

Lu, Y. B., Datta, H. J., and Bastia, D. (1998). Mechanistic studies of initiator-initiator interaction and replication initiation. *EMBO J.* 17, 5192–5200. doi: 10.1093/emboj/17.17.5192

MacLean, A. M., Finan, T. M., and Sadowsky, M. J. (2007). Genomes of the symbiotic nitrogen-fixing bacteria of legumes. *Plant Physiol.* 144, 615–622. doi: 10.1104/pp.107.101634

MacLellan, S. R., Smallbone, L. A., Sibley, C. D., and Finan, T. M. (2005). The expression of a novel antisense gene mediates incompatibility within the large repABC family of alpha-proteobacterial plasmids. *Mol. Microbiol.* 55, 611–623. doi: 10.1111/j.1365-2958.2004.04412.x

MacLellan, S. R., Zaheer, R., Sartor, A. L., MacLean, A. M., and Finan, T. M. (2006). Identification of a megaplasmid centromere reveals genetic structural diversity within the repABC family of basic replicons. *Mol. Microbiol.* 59, 1559–1575. doi: 10.1111/j.1365-2958.2006.05040.x

Marchetti, M., Capela, D., Glew, M., Cruveiller, S., Chane-Woon-Ming, B., Gris, C., et al. (2010). Experimental evolution of a plant pathogen into a legume symbiont. *PLoS Biol.* 8:e1000280. doi: 10.1371/journal.pbio.1000280

Marczynski, G. T. (1999). Chromosome methylation and measurement of faithful, once and only once per cell cycle chromosome replication in *Caulobacter crescentus*. *J. Bacteriol.* 181, 1984–1993.

Marczynski, G. T., and Shapiro, L. (2002). Control of chromosome replication in *Caulobacter crescentus*. *Annu. Rev. Microbiol.* 56, 625–656. doi: 10.1146/annurev.micro.56.012302.161103

Messerschmidt, S. J., Kemter, F. S., Schindler, D., and Waldminghaus, T. (2015). Synthetic secondary chromosomes in *Escherichia coli* based on the replication origin of chromosome II in *Vibrio cholerae*. *Biotechnol. J.* 10, 302–314. doi: 10.1002/biot.201400031

Michaux, S., Paillisson, J., Carles-Nurit, M. J., Bourg, G., Allardet-Servent, A., and Ramuz, M. (1993). Presence of two independent chromosomes in the *Brucella melitensis* 16M genome. *J. Bacteriol.* 175, 701–705. doi: 10.1128/jb.175.3.701-705.1993

Mohapatra, S. S., Fioravanti, A., and Biondi, E. G. (2014). DNA methylation in *Caulobacter* and other *Alphaproteobacteria* during cell cycle progression. *Trends Microbiol.* 22, 528–535. doi: 10.1016/j.tim.2014.05.003

Moreno, E. (1998). Genome evolution within the alpha *Proteobacteria*: why do some bacteria not possess plasmids and others exhibit more than one different chromosome? *FEMS Microbiol. Rev.* 22, 255–275. doi: 10.1111/j.1574-6976.1998.tb00370.x

Murray, H., and Errington, J. (2008). Dynamic control of the DNA replication initiation protein DnaA by Soj/ParA. *Cell* 135, 74–84. doi: 10.1016/j.cell.2008.07.044

Nakamura, A., Wada, C., and Miki, K. (2007). Structural basis for regulation of bifunctional roles in replication initiator protein. *Proc. Natl. Acad. Sci. U.S.A.* 104, 18484–18489. doi: 10.1073/pnas.0705623104

Norais, C., Hawkins, M., Hartman, A. L., Eisen, J. A., Myllykallio, H., and Allers, T. (2007). Genetic and physical mapping of DNA replication origins in *Haloferax volcanii*. *PLoS Genet.* 3:e00077. doi: 10.1371/journal.pgen.0030077

Nordström, K., and Dasgupta, S. (2006). Copy-number control of the *Escherichia coli* chromosome: a plasmidologist's view. *EMBO Rep.* 7, 484–489. doi: 10.1038/sj.embor.7400681

Ogura, Y., Ogasawara, N., Harry, E. J., and Moriya, S. (2003). Increasing the ratio of Soj to Spo0J promotes replication initiation in *Bacillus subtilis*. *J. Bacteriol.* 185, 6316–6324. doi: 10.1128/JB.185.21.6316-6324.2003

Ohbayashi, R., Watanabe, S., Ehira, S., Kanesaki, Y., Chibazakura, T., and Yoshikawa, H. (2016). Diversification of DnaA dependency for DNA replication in cyanobacterial evolution. *ISME J.* 10, 1113–1121. doi: 10.1038/ismej.2015.194

Okada, K., Iida, T., Kita-Tsukamoto, K., and Honda, T. (2005). Vibrios commonly possess two chromosomes. *J. Bacteriol.* 187, 752–757. doi: 10.1128/JB.187.2.752-757.2005

Orlova, N., Gerding, M., Ivashkiv, O., Olinares, P. D. B., Chait, B. T., Waldor, M. K., et al. (2017). The replication initiator of the cholera pathogen's second chromosome shows structural similarity to plasmid initiators. *Nucleic Acids Res.* 45, 3724–3737. doi: 10.1093/nar/gkw1288

Pal, D., Venkova-Canova, T., Srivastava, P., and Chattoraj, D. K. (2005). Multipartite regulation of rctB, the replication initiator gene of *Vibrio cholerae* chromosome II. *J. Bacteriol.* 187, 7167–7175. doi: 10.1128/JB.187.21.7167-7175.2005

Palmer, K. M., Turner, S. L., and Young, J. P. (2000). Sequence diversity of the plasmid replication gene repC in the Rhizobiaceae. *Plasmid* 44, 209–219. doi: 10.1006/plas.2000.1488

Pappas, K. M., and Winans, S. C. (2003a). A LuxR-type regulator from *Agrobacterium tumefaciens* elevates Ti plasmid copy number by activating transcription of plasmid replication genes. *Mol. Microbiol.* 48, 1059–1073.

Pappas, K. M., and Winans, S. C. (2003b). The RepA and RepB autorepressors and TraR play opposing roles in the regulation of a Ti plasmid repABC operon. *Mol. Microbiol.* 49, 441–455.

Pastorino, G. N., Martinez Alcántara, V., and Balatti, P. A. (2003). Identification of fast and slow growing rhizobia nodulating soybean (*Glycine max* [L.] Merr) by a multiplex PCR reaction. *FEMS Microbiol. Lett.* 229, 153–158. doi: 10.1016/S0378-1097(03)00796-1

Petersen, J., Brinkmann, H., and Pradella, S. (2009). Diversity and evolution of repABC type plasmids in *Rhodobacterales*. *Environ. Microbiol.* 11, 2627–2638. doi: 10.1111/j.1462-2920.2009.01987.x

Petersen, J., Frank, O., Göker, M., and Pradella, S. (2013). Extrachromosomal, extraordinary and essential–the plasmids of the *Roseobacter* clade. *Appl. Microbiol. Biotechnol.* 97, 2805–2815. doi: 10.1007/s00253-013-4746-8

Pini, F., De Nisco, N. J., Ferri, L., Penterman, J., Fioravanti, A., Brilli, M., et al. (2015). Cell cycle control by the master regulator CtrA in *Sinorhizobium meliloti*. *PLoS Genet.* 11:e1005232. doi: 10.1371/journal.pgen.1005232

Pinto, U. M., Flores-Mireles, A. L., Costa, E. D., and Winans, S. C. (2011). RepC protein of the octopine-type Ti plasmid binds to the probable origin of replication within repC and functions only in cis. *Mol. Microbiol.* 81, 1593–1606. doi: 10.1111/j.1365-2958.2011.07789.x

Pinto, U. M., Pappas, K. M., and Winans, S. C. (2012). The ABCs of plasmid replication and segregation. *Nat. Rev. Microbiol.* 10, 755–765. doi: 10.1038/nrmicro2882

Pritchard, R. H., Chandler, M. G., and Collins, J. (1975). Independence of F replication and chromosome replication in *Escherichia coli*. *Mol. Gen. Genet.* 138, 143–155. doi: 10.1007/BF02428118

Prozorov, A. A. (2008). Additional chromosomes in bacteria: properties and origin. *Mikrobiologiia* 77, 437–447. doi: 10.1134/S0026261708040012

Quon, K. C., Marczynski, G. T., and Shapiro, L. (1996). Cell cycle control by an essential bacterial two-component signal transduction protein. *Cell* 84, 83–93. doi: 10.1016/S0092-8674(00)80995-2

Rajewska, M., Wegrzyn, K., and Konieczny, I. (2012). AT-rich region and repeated sequences - the essential elements of replication origins of bacterial replicons. *FEMS Microbiol. Rev.* 36, 408–434. doi: 10.1111/j.1574-6976.2011.00300.x

Ramachandran, R., Ciacca, P. N., Filsuf, T. A., Jha, J. K., and Chattoraj, D. K. (2018). Chromosome 1 licenses chromosome 2 replication in *Vibrio cholerae* by doubling the crtS gene dosage. *PLoS Genet.* 14:e1007426. doi: 10.1371/journal.pgen.1007426

Ramírez-Romero, M. A., Soberón, N., Pérez-Oseguera, A., Téllez-Sosa, J., and Cevallos, M. A. (2000). Structural elements required for replication and incompatibility of the *Rhizobium etli* symbiotic plasmid. *J. Bacteriol.* 182, 3117–3124. doi: 10.1128/JB.182.11.3117-3124.2000

Ramírez-Romero, M. A., Téllez-Sosa, J., Barrios, H., Pérez-Oseguera, A., Rosas, V., and Cevallos, M. A. (2001). RepA negatively autoregulates the transcription of the repABC operon of the *Rhizobium etli* symbiotic plasmid basic replicon. *Mol. Microbiol.* 42, 195–204. doi: 10.1046/j.1365-2958.2001.02621.x

Rasmussen, T., Jensen, R. B., and Skovgaard, O. (2007). The two chromosomes of *Vibrio cholerae* are initiated at different time points in the cell cycle. *EMBO J.* 26, 3124–3131. doi: 10.1038/sj.emboj.7601747

Reyes-Lamothe, R., Nicolas, E., and Sherratt, D. J. (2012). Chromosome replication and segregation in bacteria. *Annu. Rev. Genet.* 46, 121–143. doi: 10.1146/annurev-genet-110711-155421

Robertson, G. T., Reisenauer, A., Wright, R., Jensen, R. B., Jensen, A., Shapiro, L., et al. (2000). The *Brucella abortus* CcrM DNA methyltransferase is essential for viability, and its overexpression attenuates intracellular replication in murine macrophages. *J. Bacteriol.* 182, 3482–3489. doi: 10.1128/JB.182.12.3482-3489.2000

Rosenberg, C., Casse-Delbart, F., Dusha, I., David, M., and Boucher, C. (1982). Megaplasmids in the plant-associated bacteria *Rhizobium meliloti* and *Pseudomonas solanacearum*. *J. Bacteriol.* 150, 402–406.

Schallopp, N., Milbredt, S., Sperlea, T., Kemter, F. S., Bruhn, M., Schindler, D., et al. (2017). Establishing a system for testing replication inhibition of the *Vibrio cholerae* secondary chromosome in *Escherichia coli*. *Antibiotics* 7:E3. doi: 10.3390/antibiotics7010003

Schekman, R., Weiner, A., and Kornberg, A. (1974). Multienzyme systems of DNA replication. *Science* 186, 987–993. doi: 10.1126/science.186.4168.987

Scholefield, G., Errington, J., and Murray, H. (2012). Soj/ParA stalls DNA replication by inhibiting helix formation of the initiator protein DnaA. *EMBO J.* 31, 1542–1555. doi: 10.1038/emboj.2012.6

Silva, A. J., and Benitez, J. A. (2016). *Vibrio cholerae* biofilms and cholera pathogenesis. *PLoS Negl. Trop. Dis.* 10:e0004330. doi: 10.1371/journal.pntd.0004330

Soberón, N., Venkova-Canova, T., Ramírez-Romero, M. A., Téllez-Sosa, J., and Cevallos, M. A. (2004). Incompatibility and the partitioning site of the repABC basic replicon of the symbiotic plasmid from *Rhizobium etli*. *Plasmid* 51, 203–216. doi: 10.1016/j.plasmid.2004.01.005

Soora, M., Tomasch, J., Wang, H., Michael, V., Petersen, J., Engelen, B., et al. (2015). Oxidative stress and starvation in *Dinoroseobacter shibae*: the role of extrachromosomal elements. *Front. Microbiol.* 6:233. doi: 10.3389/fmicb.2015.00233

Stephens, C., Reisenauer, A., Wright, R., and Shapiro, L. (1996). A cell cycle-regulated bacterial DNA methyltransferase is essential for viability. *Proc. Natl. Acad. Sci. U.S.A.* 93, 1210–1214. doi: 10.1073/pnas.93.3.1210

Suwanto, A., and Kaplan, S. (1989). Physical and genetic mapping of the *Rhodobacter sphaeroides* 2.4.1 genome: presence of two unique circular chromosomes. *J. Bacteriol.* 171, 5850–5859. doi: 10.1128/jb.171.11.5850-5859.1989

Swan, M. K., Bastia, D., and Davies, C. (2006). Crystal structure of pi initiator protein-iteron complex of plasmid R6K: implications for initiation of plasmid DNA replication. *Proc. Natl. Acad. Sci. U.S.A.* 103, 18481–18486. doi: 10.1073/pnas.0609046103

Szymanik, M., Welc-Faleciak, R., Bartosik, D., and Włodarczyk, M. (2006). Replication system of plasmid pMTH4 of *Paracoccus methylutens* DM12 contains an enhancer. *Pol. J. Microbiol.* 55, 261–270.

Touchon, M., and Rocha, E. P. C. (2016). Coevolution of the organization and structure of prokaryotic genomes. *Cold Spring Harb. Perspect. Biol.* 8:a018168. doi: 10.1101/cshperspect.a018168

Trucksis, M., Michalski, J., Deng, Y. K., and Kaper, J. B. (1998). The *Vibrio cholerae* genome contains two unique circular chromosomes. *Proc. Natl. Acad. Sci. U.S.A.* 95, 14464–14469. doi: 10.1073/pnas.95.24.14464

Val, M.-E., Kennedy, S. P., El Karoui, M., Bonné, L., Chevalier, F., and Barre, F.-X. (2008). FtsK-dependent dimer resolution on multiple chromosomes in the pathogen *Vibrio cholerae*. *PLoS Genet.* 4:e1000201. doi: 10.1371/journal.pgen.1000201

Val, M.-E., Kennedy, S. P., Soler-Bistué, A. J., Barbe, V., Bouchier, C., Ducos-Galand, M., et al. (2014). Fuse or die: how to survive the loss of Dam in *Vibrio cholerae*. *Mol. Microbiol.* 91, 665–678. doi: 10.1111/mmi.12483

Val, M.-E., Marbouty, M., de Lemos Martins, F., Kennedy, S. P., Kemble, H., Bland, M. J., et al. (2016). A checkpoint control orchestrates the replication of the two chromosomes of *Vibrio cholerae*. *Sci. Adv.* 2:e1501914. doi: 10.1126/sciadv.1501914

Val, M.-E., Skovgaard, O., Ducos-Galand, M., Bland, M. J., and Mazel, D. (2012). Genome engineering in *Vibrio cholerae*: a feasible approach to address biological issues. *PLoS Genet.* 8:e1002472. doi: 10.1371/journal.pgen.1002472

Venkova-Canova, T., Baek, J. H., Fitzgerald, P. C., Blokesch, M., and Chattoraj, D. K. (2013). Evidence for two different regulatory mechanisms linking replication and segregation of *Vibrio cholerae* chromosome II. *PLoS Genet.* 9:e1003579. doi: 10.1371/journal.pgen.1003579

Venkova-Canova, T., and Chattoraj, D. K. (2011). Transition from a plasmid to a chromosomal mode of replication entails additional regulators. *Proc. Natl. Acad. Sci. U.S.A.* 108, 6199–6204. doi: 10.1073/pnas.1013244108

Venkova-Canova, T., Saha, A., and Chattoraj, D. K. (2012). A 29-mer site regulates transcription of the initiator gene as well as function of the replication origin of *Vibrio cholerae* chromosome II. *Plasmid* 67, 102–110. doi: 10.1016/j.plasmid.2011.12.009

Venkova-Canova, T., Soberón, N. E., Ramírez-Romero, M. A., and Cevallos, M. A. (2004). Two discrete elements are required for the replication of a repABC plasmid: an antisense RNA and a stem-loop structure. *Mol. Microbiol.* 54, 1431–1444. doi: 10.1111/j.1365-2958.2004.04366.x

Venkova-Canova, T., Srivastava, P., and Chattoraj, D. K. (2006). Transcriptional inactivation of a regulatory site for replication of *Vibrio cholerae* chromosome II. *Proc. Natl. Acad. Sci. U.S.A.* 103, 12051–12056. doi: 10.1073/pnas.0605120103

Wegrzyn, K. E., Gross, M., Uciechowska, U., and Konieczny, I. (2016). Replisome assembly at bacterial chromosomes and iteron plasmids. *Front. Mol. Biosci.* 3:39. doi: 10.3389/fmolb.2016.00039

Wickner, S., Hoskins, J., and McKenney, K. (1991). Monomerization of RepA dimers by heat shock proteins activates binding to DNA replication origin. *Proc. Natl. Acad. Sci. U.S.A.* 88, 7903–7907. doi: 10.1073/pnas.88.18.7903

Williams, D. R., and Thomas, C. M. (1992). Active partitioning of bacterial plasmids. *J. Gen. Microbiol.* 138, 1–16. doi: 10.1099/00221287-138-1-1

Wion, D., and Casadesús, J. (2006). N6-methyl-adenine: an epigenetic signal for DNA-protein interactions. *Nat. Rev. Microbiol.* 4, 183–192. doi: 10.1038/nrmicro1350

Wolański, M., Jakimowicz, D., and Zakrzewska-Czerwińska, J. (2014). Fifty years after the replicon hypothesis: cell-specific master regulators as new players in chromosome replication control. *J. Bacteriol.* 196, 2901–2911. doi: 10.1128/JB.01706-14

Wright, R., Stephens, C., and Shapiro, L. (1997). The CcrM DNA methyltransferase is widespread in the alpha subdivision of *proteobacteria*, and its essential functions are conserved in *Rhizobium meliloti* and *Caulobacter crescentus*. *J. Bacteriol.* 179, 5869–5877. doi: 10.1128/jb.179.18.5869-5877.1997

Xu, Q., Dziejman, M., and Mekalanos, J. J. (2003). Determination of the transcriptome of *Vibrio cholerae* during intraintestinal growth and midexponential phase *in vitro*. *Proc. Natl. Acad. Sci. U.S.A.* 100, 1286–1291. doi: 10.1073/pnas.0337479100

Yamaichi, Y., Fogel, M. A., McLeod, S. M., Hui, M. P., and Waldor, M. K. (2007). Distinct centromere-like parS sites on the two chromosomes of *Vibrio* spp. *J. Bacteriol.* 189, 5314–5324. doi: 10.1128/JB.00416-07

Yamaichi, Y., Gerding, M. A., Davis, B. M., and Waldor, M. K. (2011). Regulatory cross-talk links *Vibrio cholerae* chromosome II replication and segregation. *PLoS Genet.* 7:e1002189. doi: 10.1371/journal.pgen.1002189

Yamaichi, Y., Iida, T., Park, K. S., Yamamoto, K., and Honda, T. (1999). Physical and genetic map of the genome of *Vibrio parahaemolyticus*: presence of two chromosomes in *Vibrio* species. *Mol. Microbiol.* 31, 1513–1521. doi: 10.1046/j.1365-2958.1999.01296.x

Yao, Y., Enkhtsetseg, S., Odsbu, I., Fan, L., and Morigen, M. (2018). Mutations of DnaA-boxes in the oriR region increase replication frequency of the MiniR1-1 plasmid. *BMC Microbiol.* 18:27. doi: 10.1186/s12866-018-1162-3

Żebracki, K., Koper, P., Marczak, M., Skorupska, A., and Mazur, A.(2015). Plasmid-encoded RepA proteins specifically autorepress individual repABC operons in the multipartite *Rhizobium leguminosarum* bv. *trifolii* Genome. *PLoS One* 10:e0131907. doi: 10.1371/journal.pone.0131907

Zeuthen, J., and Pato, M. L. (1971). Replication of the F'lac sex factor in the cell cycle of *Escherichia coli*. *Mol. Gen. Genet.* 111, 242–255. doi: 10.1007/BF00433109

Zhong, Z., Helinski, D., and Toukdarian, A. (2003). A specific region in the N terminus of a replication initiation protein of plasmid RK2 is required for recruitment of *Pseudomonas aeruginosa* DnaB helicase to the plasmid origin. *J. Biol. Chem.* 278, 45305–45310. doi: 10.1074/jbc.M306058200

Zuerner, R. L., Herrmann, J. L., and Saint Girons, I. (1993). Comparison of genetic maps for two *Leptospira interrogans* serovars provides evidence for two chromosomes and intraspecies heterogeneity. *J. Bacteriol.* 175, 5445–5451. doi: 10.1128/jb.175.17.5445-5451.1993

A Requirement for Global Transcription Factor Lrp in Licensing Replication of *Vibrio cholerae* Chromosome 2

Peter N. Ciaccia, Revathy Ramachandran* and Dhruba K. Chattoraj

Laboratory of Biochemistry and Molecular Biology, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD, United States

The human pathogen, *Vibrio cholerae*, belongs to the 10% of bacteria in which the genome is divided. Each of its two chromosomes, like bacterial chromosomes in general, replicates from a unique origin at fixed times in the cell cycle. Chr1 initiates first, and upon duplication of a site in Chr1, *crtS*, Chr2 replication initiates. Recent *in vivo* experiments demonstrate that *crtS* binds the Chr2-specific initiator RctB and promotes its initiator activity by remodeling it. Compared to the well-defined RctB binding sites in the Chr2 origin, *crtS* is an order of magnitude longer, suggesting that other factors can bind to it. We developed an *in vivo* screen to identify additional *crtS*-binding proteins and identified the global transcription factor, Lrp, as one such protein. Studies *in vivo* and *in vitro* indicate that Lrp binds to *crtS* and facilitates RctB binding to *crtS*. Chr2 replication is severely defective in the absence of Lrp, indicative of a critical role of the transcription factor in licensing Chr2 replication. Since Lrp responds to stresses such as nutrient limitation, its interaction with RctB presumably sensitizes Chr2 replication to the physiological state of the cell.

Keywords: *V. cholerae* Chr2 replication, replication licensing, *crtS*, RctB, Lrp, coordination of replication

INTRODUCTION

In bacteria, chromosomes initiate replication at fixed times in the cell cycle that vary depending upon the bacteria and their physiological state. Nearly 10% of bacteria from diverse genera possess divided genomes comprising more than one chromosome (Egan et al., 2005). In such bacteria, timely duplication of all chromosomes prior to cell division is crucial for genome maintenance. *Vibrio cholerae* has emerged as the model organism for studying replication control in multi-chromosome bacteria. It possesses two chromosomes, Chr1, 3 Mb, and Chr2, 1 Mb. Chr1 initiates replication first, and only upon the passage of a Chr1 replication fork across a site, *crtS,* does Chr2 initiate replication (Val et al., 2016). The *crtS* site (Chr2 replication triggering site) is thought to function by interacting with and remodeling the Chr2-specific initiator, RctB (Baek and Chattoraj, 2014). It appears that when duplication of a *crtS* site is prevented within a cell cycle, the site still shows modest activity in licensing Chr2 replication but it is insufficient to do so in a timely fashion (Ramachandran et al., 2018). Duplication of the site as a consequence of a single round of replication increases this activity sufficiently to permit initiation of Chr2 replication in each cell cycle.

The *crtS* site is essential for Chr2 replication in *V. cholerae* (Val et al., 2016). Increasing the copy number of *crtS* increases Chr2 replication in *V. cholerae*, indicating that the activity of the site is limiting for Chr2 replication. The *crtS* site also functions in *Escherichia coli*; the presence of *crtS* in a plasmid increases copy number of plasmids containing the Chr2 origin of replication (p*ori2*) and a source of RctB (Baek and Chattoraj, 2014).

The structure and function of *crtS* are fairly well conserved in the Vibrionaceae family (Kemter et al., 2018). The size of *crtS* (∼153-bp) is rather large for a protein binding site and is much larger than the RctB binding sites in the Chr2 replication origin (12-mers and 39-mers). The region in *crtS* protected by RctB covers only 18 bp (Baek and Chattoraj, 2014). There is plenty of room for other factors to interact with the site. One such factor is RNA polymerase as *crtS* possesses a sigma-70 promoter, called P_{crtS} here, which remains repressed by unknown factors. The repressed promoter allowed us to screen for host genes responsible for that repression and to determine their influence on *crtS* function.

Here, we show that in addition to RctB, *crtS* binds Lrp, a global transcription factor that responds to nutritional status (Calvo and Matthews, 1994; Cho et al., 2008). The protein is largely responsible for keeping P_{crtS} repressed and mediating RctB binding to *crtS*. In the absence of Lrp, Chr2 replication is severely defective. The regulation of Chr2 replication by a global regulator of nutritional status may provide a link between chromosomal replication and the physiological state of the cell.

MATERIALS AND METHODS

Bacterial Strains and Growth Conditions

The bacterial strains and plasmids used in this study are listed in **Supplementary Tables S1,S2**, respectively. All *E. coli* strains are K12 derivatives and all *V. cholerae* strains are El Tor N16961 derivatives, and were maintained in lysogeny broth (LB) at 37°C and 30°C, respectively, unless otherwise specified. When required, media was supplemented with antibiotics at the following concentrations for *E. coli*: 100 µg/ml ampicillin, 25 µg/ml chloramphenicol, 25 µg/ml kanamycin, 40 µg/ml spectinomycin, and 25 µg/ml zeocin. *V. cholerae* strains were maintained with the same antibiotic concentrations as above except for chloramphenicol, which was used at 5 µg/ml.

Microscopy

Single colonies grown overnight in LB with the appropriate antibiotics were used to inoculate 1X M63 medium supplemented with 1 mM $CaCl_2$, 1 mM $MgSO_4$, 0.001% vitamin B1, 0.2% fructose, 0.1% casamino acids, and 100 µM IPTG (to induce GFP-P1ParB). Cultures were grown at 30°C to an OD_{600} of 0.3, added to the center of a glass P35 dish (MatTek corporation, Ashland, MA), and overlaid with 1% agarose prepared with the same medium. Dishes were imaged and analyzed as previously described (Ramachandran et al., 2018).

Natural Transformation of *V. cholerae* to Replace *crtS* With Δ3′crtS, Δ5′crtS, or Δ5′Δ3′crtS

Natural transformations of CVC3058 (HapR$^+$ derivative of El Tor N16961 with P1*parS* cloned at +40 kb on Chr2 for visualizing *ori2* as GFP-P1ParB foci) and CVC3061 (CVC3058 with extra *crtS* site cloned 10 kb upstream of native site) were performed as described (Ramachandran et al., 2018). The native copy of *crtS* was replaced with truncated versions using linear DNA amplified from pPC143, pPC144, and pPC145 containing Δ3′*crtS*, Δ5′*crtS*, and Δ5′Δ3′*crtS*, respectively, flanked by 1 kb of homologous DNA present in plasmid pBJH245. pPC143 was assembled from Δ3′*crtS* DNA amplified from pBJH188 with primers PNC47 and PNC56 and from pBJH245 amplified with primers PNC46 and PNC54. Primers used here are described in **Supplementary Table S3**. pPC144 was assembled from Δ5′*crtS* DNA amplified from pBJH188 with primers PNC51 and PNC55 and from pBJH245 amplified with primers PNC50 and PNC53. pPC145 was assembled from Δ5′Δ3′*crtS* DNA amplified from pBJH188 with primers PNC51 and PNC56 and from pBJH245 amplified with primers PNC50 and PNC54. Plasmids were assembled using the HiFi DNA assembly kit (NEB).

β-Galactosidase Assay

Plasmid Construction: Truncated *crtS* species were transcriptionally fused to *lacZ* in pMLB1109 in order to measure promoter activity. The *crtS* fragments, Δ3′*crtS*, Δ5′*crtS*, and Δ5′Δ3′*crtS* were amplified from pBJH188 using primers PNC15 and PNC17, PNC16 and PNC18, and PNC16 and PNC18, respectively. The fragments were then ligated into pMLB1109 digested with EcoRI and SmaI to produce plasmids pPC066, pPC067, and pPC068, respectively. *E. coli* Δ*lrp* strains was complemented with Lrp using plasmid pPC401. Plasmid pPC401 contained $P_{trc}lrp$ amplified from pJWD-2 using primers PNC139 and PNC140 in a pACYC177 backbone that was amplified using primers PNC141 and PNC142.

Assay Protocol: β-galactosidase assays were performed in 96-well flat bottom plates (Costar 3596) and adapted from Schaefer et al. (2016). Colonies grown overnight on LB plates with appropriate antibiotics were used to inoculate LB. Log-phase culture, 80 µl each, was loaded in duplicates in a 96-well plate to which 120 µl of a custom mix (ONPG + Popculture reagent), prepared as described in (Schaefer et al., 2016), was added. Plates were incubated at 30°C in a Epoch2 plate reader (Biotek, United States) set to double orbital shaking and A_{420} measurements were taken in one- to 5-min intervals. Equivalent Miller Units (MU) were calculated using a Python program that parses OD_{600} and A_{420} values from the plate readers and plots A_{420} and ΔA_{420} values as a function of time, to identify maxima. Plotted β-galactosidase activity in MU represent means from three biological replicates and error bars depict SEM.

Screen of Transposon-Insertion Mutants

The EZ-Tn5 transposome kit (Lucigen, WI) was used to generate random Tn insertions in *E. coli* DH10-β harboring pBJH235. The transformation mixtures were spread first on LB plates

with appropriate antibiotics and incubated overnight at 37°C. The following day colonies were patched on MacConkey agar plates (MacConkey Agar Base [Difco, MD] supplemented with 1% (w/v) lactose and 3 mM 2-phenylethyl β-D-thiogalactoside (PETG, [Biosynth, IL], pH adjusted to 7.1). [PETG, a competitive inhibitor of β-galactosidase, was titrated to 3 mM, the concentration at which colonies with 75 MU appear white and those with 180 MU appear red (**Supplementary Figure S1**)]. Plates were incubated for 16 h at 37°C and colonies were monitored for development of red color. Candidate colonies were grown in LB overnight for genomic DNA isolation. Genomic DNA was extracted using the DNeasy Blood and Tissue Kit (Qiagen, CA, United States), digested with EcoRV-HF (NEB, MA, United States), and ligated overnight at 16°C with T4 DNA ligase (NEB). The ligation product was used to transform DH5α(λpir) cells to recover transposon containing circularized genomic DNA, which were replication competent by virtue of the presence of R6Kγori within the transposon. Plasmid DNA was extracted from individual colonies using the QIAprep Spin Miniprep Kit (Qiagen) and sequenced using the primers supplied in the EZ-Tn5™ transposome kit to identify the locations of Tn insertion.

Deletion of *lrp* in *E. coli*

Δ*lrp-787*::*kan* was transduced from *E. coli* JW0872-2 into *E. coli* DH10-β and BR8706 (constitutive *araE*) using P1*vir* (Miller, 1992). The *kan* cassette was excised by expressing Flp recombinase from pCP20 and subsequently curing the plasmid by overnight growth at 42°C (Datsenko and Wanner, 2000).

Deletion of *lrp* in *V. cholerae*

Deletion of *lrp* from CVC3058 (derivative of El Tor N16961 with P1*parS* cloned at +40 kb on Chr2 for visualizing *ori2* as GFP-P1ParB foci) was performed in the presence of a plasmid carrying *E. coli lrp* (pJWD-2), by natural transformation with linear DNA amplified from pPC352 that contained a *zeocin* cassette flanked by 1 kb upstream and downstream homology sequences. pPC352 was assembled using four DNA fragments: 1 kb upstream homology (amplified from genomic DNA using primers PNC123 and PNC124), 1 kb downstream homology (amplified from genomic DNA using primers PNC121 and PNC122), *zeocin* cassette (amplified from pEM7-Zeo using primers PNC127 and PNC128) and the backbone (amplified from pEM7-Zeo using primers PNC125 and PNC126). Linear DNA used for natural transformation was amplified from pPC352 using primers PNC131 and PNC132. Deletion of *lrp* was confirmed by PCR. The plasmid pJWD-2 was cured by growing overnight in the absence of antibiotic and screening colonies that had lost antibiotic resistance, to generate strain CVC3286. The deletion was verified by whole genome sequencing.

Purification of Lrp and MBP-RctB

Lrp was purified from plasmid pJWD-2 (Ernsting et al., 1993). 5 ml of overnight culture of *E. coli* BL21 containing pJWD-2 was used to inoculate 1 liter of LB supplemented with ampicillin and grown at 37°C. Protein expression was induced at an OD$_{600}$ $_{nm}$ of 0.8 by adding IPTG to the final concentration of 0.5 mM, and growth was allowed to continue for 2.5 h. The pellet was resuspended in PC Buffer [50 mM phosphate buffer (pH 7.4), 100 mM NaCl, 0.1 mM EDTA, 10 mM β-mercaptoethanol and 10% glycerol, (de los Rios and Perona, 2007)] supplemented with 1× protease inhibitor cocktail (Sigma-Aldrich, St. Louis, MO) and lysed by French press. The lysate was clarified by centrifugation for 1 h at 18,000 × g before loading onto a Hitrap SP HP column (GE Healthcare Life sciences, Chicago, IL, United States), pre-equilibrated with PC Buffer. Lrp was eluted using a gradient of PC Buffer + 1 M NaCl. The fractions containing Lrp were purified further by cation exchange on a Mono S column (GE Healthcare Life sciences) equilibrated with Cat2 buffer (50 mM Hepes (pH 8.0), 1 mM EDTA, 0.2% Tween20, 5% Glycerol and 100 mM NaCl. Lrp was eluted using a gradient of Cat2 buffer + 1 M NaCl. MBP-RctB was purified as described previously (Jha et al., 2017).

Electrophoretic Mobility Shift Assay (EMSA)

Interaction of purified Lrp with *crtS* was captured *in vitro* using EMSA. The 153 bp *crtS* was flanked by ~100 bp of lambda DNA and amplified from pBJH170 using FAM-labeled primers RR202 and RR214. Non-specific DNA was amplified from pTVC243 using the same primers as above and contained only the 100-bp flanks. Truncated *crtS* constructs Δ3′*crtS*, Δ5′*crtS*, and Δ5′Δ3′*crtS* were amplified from pPC189, pPC225 and pPC009, respectively, using primers PNC77 and PNC78. Increasing amounts of Lrp protein were added to 20 μl reactions that contained 5 nM each of fluorescent probe and vector DNA, 20 mM Hepes (pH 7.4), 1 mM EDTA, 0.2% Tween20, 5% glycerol, 200 ng poly dI-dC, 1 mM dithiothreitol, 70 mM potassium glutamate and 4 mM magnesium acetate. Leucine was added at 10 mM, when desired. The reaction was incubated at room temperature for 10 min before loading on a 5% native polyacrylamide gel and electrophoresed at 12 V/cm in 0.5 × TBE. The gel was scanned using Typhoon FLA 9500 (GE Healthcare Life Sciences, MA, United States). The image was analyzed, and band intensities quantified using Fiji software (Schindelin et al., 2012). The percent DNA bound was plotted against concentration of protein and K_D values were obtained by performing non-linear regression analysis assuming one site specific binding using GraphPad Prism version 7.0 a (La Jolla, CA, United States). Following EMSA of *crtS* with Lrp and RctB, the super-shifted band was excised from the native polyacrylamide gel and presence of both proteins confirmed by mass spectrometry performed at the Collaborative Protein Technology Resource (CCR, NIH) as previously described (Jha et al., 2017).

Measurement of Plasmid Copy Number

Copy number experiments were performed using either WT *E. coli* (BR8706, constitutive *araE*) or Δ*lrp E. coli* (derivatives of BR8706: CVC3260, Δ*lrp-787*::FRT-*kan*-FRT and CVC3274 Δ*lrp-787*). BR8706 and CVC3260 were transformed with pTVC11

(p*rctB*) and either pTVC243 (vector), pBJH170 (p*crtS*), or pBJH239 (p*crtS-10 m*). CVC3274 was transformed with pTVC11, pPC401 (p*lrp)*, and either pTVC243, pBJH170, or pBJH239. To maintain high levels of RctB, competent cells were grown in 0.2% arabinose before and after transformation with pTVC22 (p*ori2*). Cultures were inoculated at an $OD_{600\ nm}$ of 0.005 and grown at 37°C with shaking to an $OD_{600\ nm}$ of 0.2. Eight OD units were pelleted and used for plasmid isolation. Relative plasmid copy number was measured essentially as described (Das and Chattoraj, 2004) but normalized to pTVC11.

Whole Genome Sequencing

Genomic DNA was extracted from 1 ml of cells grown overnight at 37°C in LB using DNeasy Tissue Kit (Qiagen, Hilden, Germany). DNA was sequenced on the Illumina MiSeq platform at the NCI CCR genomics sequencing core. 1–6 million reads were obtained for each sample, which were trimmed and mapped to a CVC3058 reference genome using the CLC Genomics Workbench (Qiagen). The reference genome for CVC3058 was constructed by *de novo* assembly.

RESULTS

5′ Terminal Sequences of *crtS* Are Important for Licensing Chr2 Replication in *V. cholerae*

In *E. coli*, the presence of a plasmid containing 153 bp of *V. cholerae* Chr1 [coordinates 817947 to 818099 bp of Heidelberg et al. (Heidelberg et al., 2000)] (**Figure 1A**) increases the copy number of *ori2*-containing plasmids (p*ori2*) about threefold in the presence of RctB (Baek and Chattoraj, 2014). The 153 bp sequence was called *crtS* (Val et al., 2016). The central 54 - 123 bp, called △5′△3′*crtS* here, also increases the p*ori2* copy number about twofold in *E. coli* (Baek and Chattoraj, 2014). To test if △5′△3′*crtS* was sufficient to support replication of Chr2 in *V. cholerae* as well, the *crtS* sequence in Chr1 was replaced with △5′△3′*crtS* using natural transformation. Chr2 replication was followed by visualizing GFP-P1ParB bound to the *P1parS* site inserted 40 kb away from *ori2*, as previously described (Ramachandran et al., 2018). We found that △5′△3′*crtS* replacement resulted in a loss of *ori2* foci in 70% of cells (**Figure 1B**, top panel). Truncation of the 5′ and 3′ sequences separately revealed that this functional deficiency is due to the 5′ truncation, as truncation of the 3′ did not significantly alter the foci distribution. This suggests that the 1–123 bp of *crtS* locus spanning chromosomal coordinates 817947 to 818069 is sufficient for licensing Chr2 replication, despite the low conservation of the 5′ bases (**Figure 1A**). Furthermore, the results were indistinguishable in the △5′ constructs, whether or not the 3′ region was present (**Figure 1B**, top row).

Deletion of *crtS* leads to suppressor mutations in *rctB* or fusion of Chr1 and Chr2 (Val et al., 2016). To avoid the selection of suppressors while replacing the native *crtS* locus with truncated species, we repeated the replacements in strains that

also possessed a second functional copy of *crtS* 10 kb upstream of the native locus (Ramachandran et al., 2018). The presence of two full length copies of *crtS* causes over-replication of Chr2 (Val et al., 2016) (**Figure 1B**, bottom row). This over replication was not seen when the native *crtS* locus was replaced with △5′*crtS*. Replacement with △3′*crtS* did not alter *crtS* function, as the distribution of *ori2* foci was similar to that of cells with two intact *crtS* copies. In sum, although the exact bounds of *crtS* remain to be defined, it appears that the 5′ sequence of *crtS* is essential for licensing replication from *ori2*.
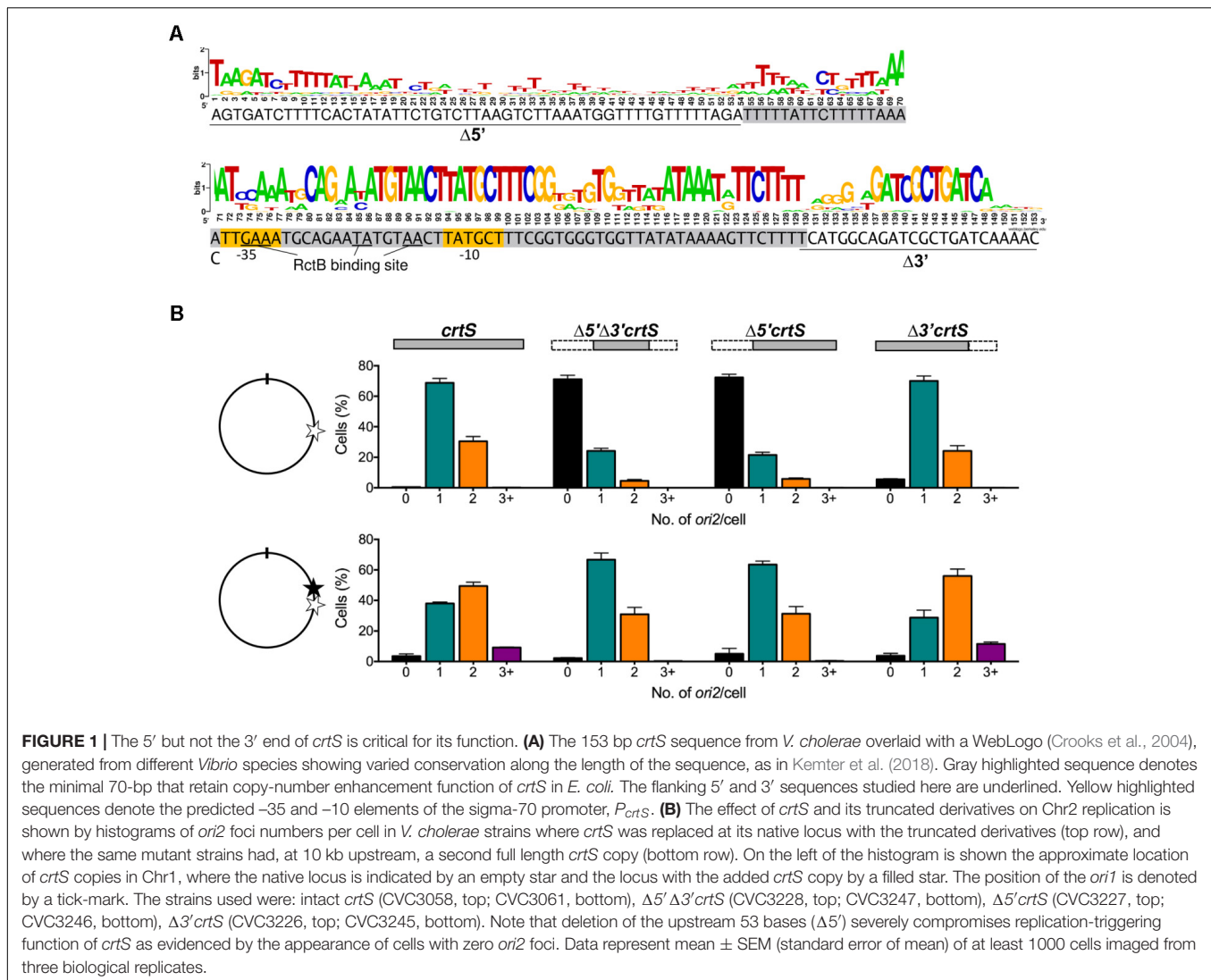
The Promoter Within *crtS* Is Repressed

In spite of the importance of the 5′ terminal sequences of *crtS*, they are not well conserved among the various *Vibrio* species (Kemter et al., 2018). Apart from AT-richness, the region does not have any known sequence features. *crtS*, however, possesses −35 and −10 promoter elements in the more conserved central region (**Figure 1A**). The promoter within *crtS*, called P_{crtS} here, was previously shown to be expressed only from △5′△3′*crtS* but not from full length *crtS* (Baek and Chattoraj, 2014). From these results, it appears that the promoter repression and replication enhancement functions of *crtS* are correlated, and that the promoter repression may be necessary for *crtS* function. To quantify the promoter repression, we fused a promoterless *lacZ* gene to the *crtS* constructs used in **Figure 1B**. In *E. coli*, P_{crtS} activity was as low as in the promoterless vector, but the activity increased fourfold in △5′*crtS* (**Figure 2**, left panel). These results indicate that an *E. coli* factor interacts with the 5′ terminal sequences of *crtS* and represses P_{crtS}. A test of whether RctB, the only protein previously found to bind *crtS* (Baek and Chattoraj, 2014), could also repress the promoter showed that it did, but only partially (black vs. white bars, **Figure 2**). The expression of P_{crtS} is thus controlled by at least two repressors. The deletion of the 3′ 30 bp had only a marginal effect on promoter activity.

In *V. cholerae*, truncation of the 5′ sequences results in only a slight increase in promoter activity (**Supplementary Figure S2**). To test whether the lack of increase could be due to the binding of *crtS* by RctB, the experiments were repeated in a strain of *V. cholerae*, MCH1, that lacks RctB and where Chr2 is maintained by fusion to Chr1 (Val et al., 2012). In MCH1, P_{crtS} was expressed threefold higher in △5′*crtS* and △5′△3′*crtS* than *crtS*, mirroring the *E. coli* results (**Figure 2**, right panel). Addition of RctB caused partial repression of promoter activity in △5′*crtS* and △5′△3′*crtS*, as in *E. coli*. Together, these results strongly suggest that a factor other than RctB, common to both *E. coli* and *V. cholerae*, binds *crtS* and is responsible for the additional repression of the promoter within *crtS*.

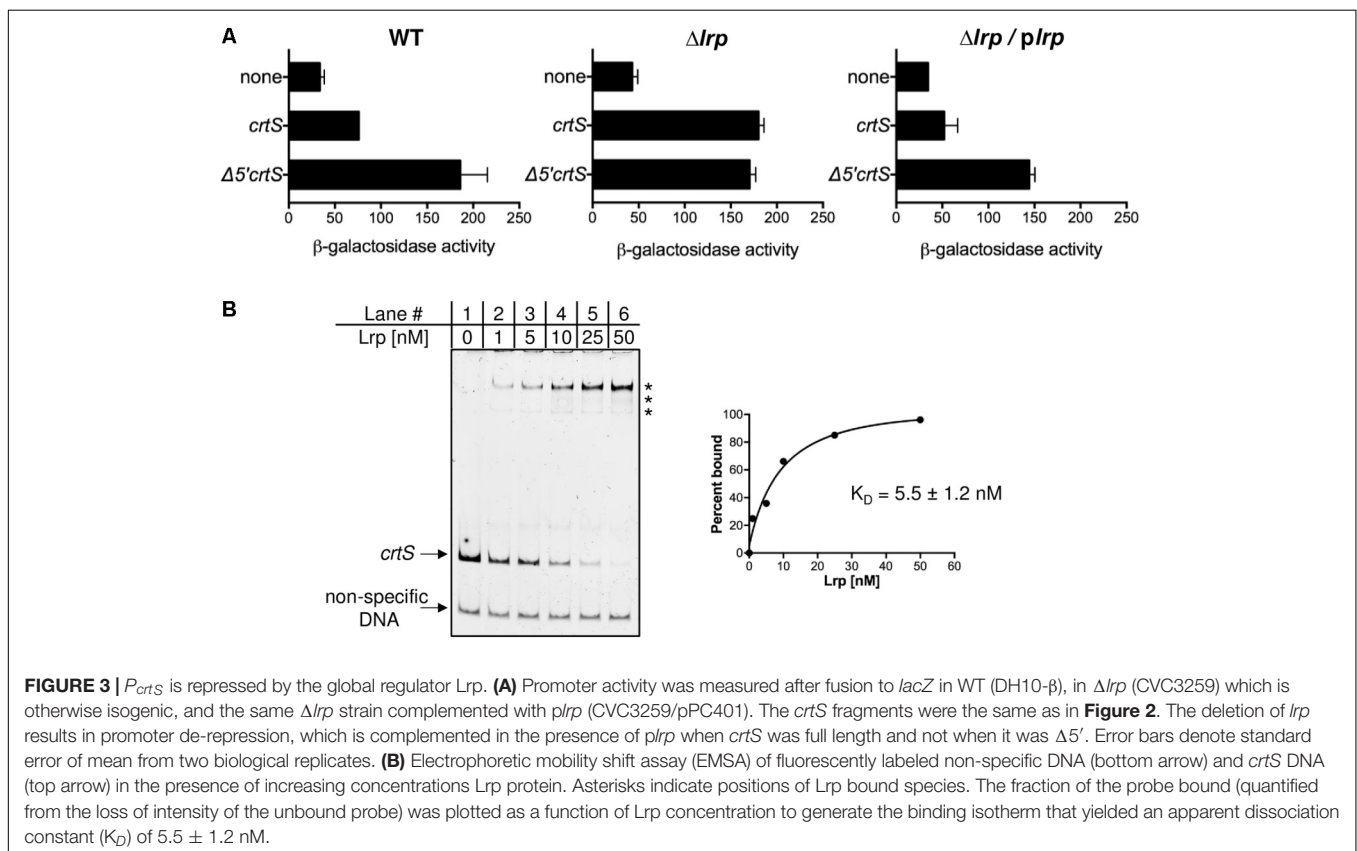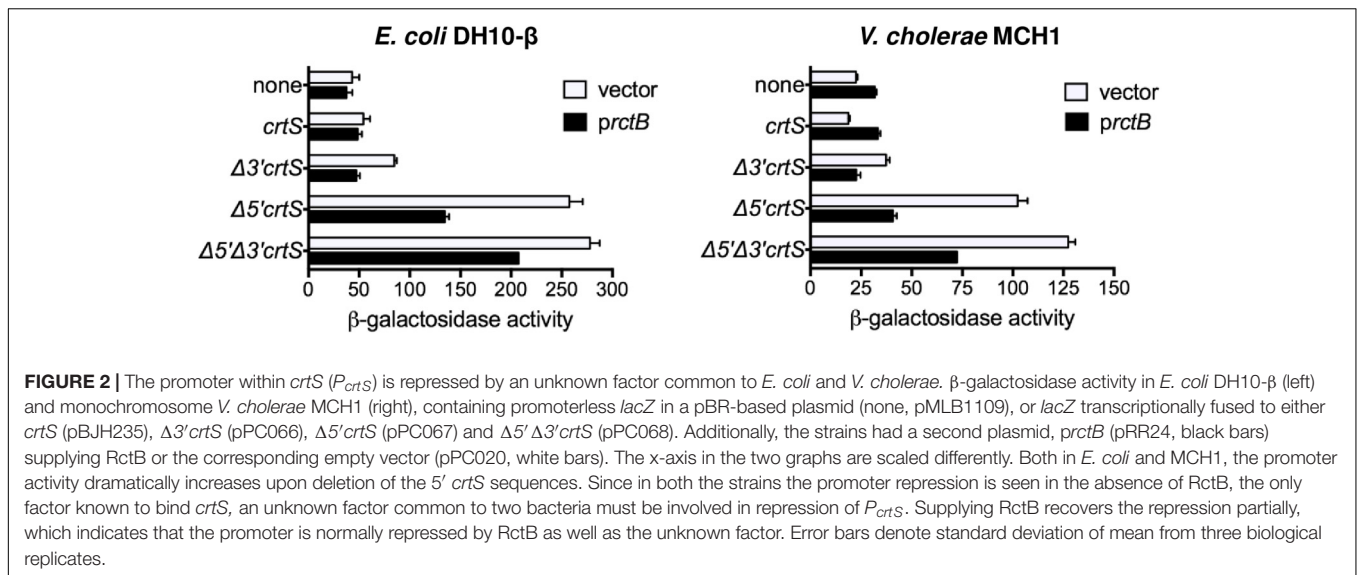### P_{crtS} Is Repressed by the Global Regulator Lrp in *E. coli* and *V. cholerae*

The putative *E. coli* factor responsible for repressing P_{crtS} was identified by performing a transposon (Tn) insertional mutagenesis screen in strains that contained a plasmid with transcriptional-fusion of *crtS* to *lacZ*. Colonies with higher *lacZ* activity were identified by plating on MacConkey agar

FIGURE 1 | The 5′ but not the 3′ end of *crtS* is critical for its function. **(A)** The 153 bp *crtS* sequence from *V. cholerae* overlaid with a WebLogo (Crooks et al., 2004), generated from different *Vibrio* species showing varied conservation along the length of the sequence, as in Kemter et al. (2018). Gray highlighted sequence denotes the minimal 70-bp that retain copy-number enhancement function of *crtS* in *E. coli*. The flanking 5′ and 3′ sequences studied here are underlined. Yellow highlighted sequences denote the predicted −35 and −10 elements of the sigma-70 promoter, P*crtS*. **(B)** The effect of *crtS* and its truncated derivatives on Chr2 replication is shown by histograms of *ori2* foci numbers per cell in *V. cholerae* strains where *crtS* was replaced at its native locus with the truncated derivatives (top row), and where the same mutant strains had, at 10 kb upstream, a second full length *crtS* copy (bottom row). On the left of the histogram is shown the approximate location of *crtS* copies in Chr1, where the native locus is indicated by an empty star and the locus with the added *crtS* copy by a filled star. The position of the *ori1* is denoted by a tick-mark. The strains used were: intact *crtS* (CVC3058, top; CVC3061, bottom), Δ5′Δ3′*crtS* (CVC3228, top; CVC3247, bottom), Δ5′*crtS* (CVC3227, top; CVC3246, bottom), Δ3′*crtS* (CVC3226, top; CVC3245, bottom). Note that deletion of the upstream 53 bases (Δ5′) severely compromises replication-triggering function of *crtS* as evidenced by the appearance of cells with zero *ori2* foci. Data represent mean ± SEM (standard error of mean) of at least 1000 cells imaged from three biological replicates.

supplemented with 3 mM PETG, an inhibitor of β-galactosidase, that allowed clearer distinction between red and white colonies (Golding et al., 1991) (**Supplementary Figure S1**). In most of these colonies, the Tn was found to have inserted into the plasmid expressing *lacZ*. In one colony, the Tn was found to have inserted within the 5′ untranslated region of the *lrp* gene. To determine whether Lrp is responsible for the observed repression of P*crtS*, the promoter activity was measured in a *E. coli* Δ*lrp* strain [from Keio collection, (Baba et al., 2006)] and, although *crtS* was full length, the activity was as high as from Δ5′*crtS* (**Figure 3A**). Complementing the Δ*lrp* strain with Lrp using plasmid pJWD-2 (Ernsting et al., 1993) resulted in repression of P*crtS*, when present in intact *crtS* but not when present in Δ5′*crtS*. These results are fully consistent with Lrp being the factor that, directly or indirectly, keeps P*crtS* repressed.

An *in vitro* experiment was performed to test whether Lrp itself binds to *crtS*. *E. coli* Lrp protein was purified from plasmid pJWD-2 to about 95% purity. The *E. coli* protein is 92% identical 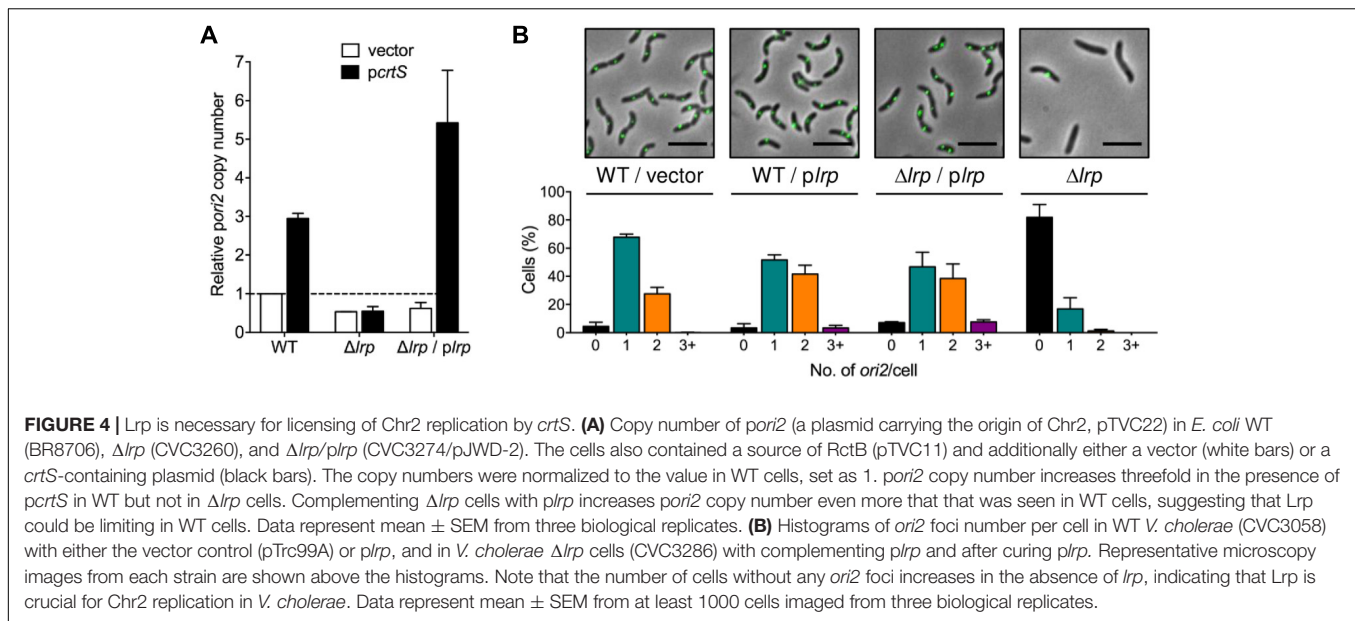to the *V. cholerae* Lrp protein and is completely conserved in the helix-turn-helix motif (Lintner et al., 2008). In EMSA using fluorescently labeled *crtS*, Lrp was seen to bind *crtS* with an approximate K_D of 5.5 ± 1.2 nM (**Figure 3B**). The addition of leucine altered the distribution of the Lrp-shifted species, indicating that *crtS*-binding is responsive to the presence of leucine (**Supplementary Figure S3**). Lrp was seen to bind equally well to Δ3′*crtS* and Δ5′*crtS* and slightly less well to Δ5′Δ3′*crtS*, indicating that it has multiple binding sites within *crtS*, but it appears that the site(s) within the 5′ terminal sequences are required for promoter repression (**Supplementary Figure S4**). A search for Lrp binding site within *crtS* using a SELEX derived consensus sequence (Cui et al., 1995) revealed a putative site with 12/15 matches covering 50 - 64 bp region. The first four bp of this putative Lrp binding site are lost upon truncation of the 5′ sequences, possibly explaining the loss of repression in Δ5′*crtS*. In addition to the 15 bp consensus, the three to five flanking bases also contribute to specific binding by Lrp (Cui et al., 1995), which are also missing in Δ5′*crtS*.

FIGURE 2 | The promoter within *crtS* (P*crtS*) is repressed by an unknown factor common to *E. coli* and *V. cholerae*. β-galactosidase activity in *E. coli* DH10-β (left) and monochromosome *V. cholerae* MCH1 (right), containing promoterless *lacZ* in a pBR-based plasmid (none, pMLB1109), or *lacZ* transcriptionally fused to either *crtS* (pBJH235), Δ*3′crtS* (pPC066), Δ*5′crtS* (pPC067) and Δ*5′ Δ3′crtS* (pPC068). Additionally, the strains had a second plasmid, p*rctB* (pRR24, black bars) supplying RctB or the corresponding empty vector (pPC020, white bars). The x-axis in the two graphs are scaled differently. Both in *E. coli* and MCH1, the promoter activity dramatically increases upon deletion of the 5′ *crtS* sequences. Since in both the strains the promoter repression is seen in the absence of RctB, the only factor known to bind *crtS*, an unknown factor common to two bacteria must be involved in repression of P*crtS*. Supplying RctB recovers the repression partially, which indicates that the promoter is normally repressed by RctB as well as the unknown factor. Error bars denote standard deviation of mean from three biological replicates.

FIGURE 3 | P*crtS* is repressed by the global regulator Lrp. **(A)** Promoter activity was measured after fusion to *lacZ* in WT (DH10-β), in Δ*lrp* (CVC3259) which is otherwise isogenic, and the same Δ*lrp* strain complemented with p*lrp* (CVC3259/pPC401). The *crtS* fragments were the same as in **Figure 2**. The deletion of *lrp* results in promoter de-repression, which is complemented in the presence of p*lrp* when *crtS* was full length and not when it was Δ5′. Error bars denote standard error of mean from two biological replicates. **(B)** Electrophoretic mobility shift assay (EMSA) of fluorescently labeled non-specific DNA (bottom arrow) and *crtS* DNA (top arrow) in the presence of increasing concentrations Lrp protein. Asterisks indicate positions of Lrp bound species. The fraction of the probe bound (quantified from the loss of intensity of the unbound probe) was plotted as a function of Lrp concentration to generate the binding isotherm that yielded an apparent dissociation constant (K_D) of 5.5 ± 1.2 nM.

Lrp Is Required for Chr2 Replication-Licensing by *crtS* in *E. coli* and *V. cholerae*

In order to test the effect of Lrp on the replication enhancement function of *crtS*, the copy number of *ori2*-containing plasmids was measured in Δ*lrp* strains. While in WT *E. coli*, the copy number of p*ori2* increased about threefold in the presence of

p*crtS* compared to the empty vector, no such increase was observed in the Δ*lrp* strain (**Figure 4A**). This indicates that *crtS* fails to function as an enhancer of Chr2 replication in the absence of Lrp. Upon complementing with an Lrp-expressing plasmid, p*lrp*, the copy number of p*ori2* increased about fourfold in the presence of P*crtS*, whereas the vector copy number was unaffected, indicating that Lrp is essential for *crtS* function in *E. coli*. To test whether the Lrp was required solely to repress

FIGURE 4 | Lrp is necessary for licensing of Chr2 replication by *crtS*. **(A)** Copy number of p*ori2* (a plasmid carrying the origin of Chr2, pTVC22) in *E. coli* WT (BR8706), Δ*lrp* (CVC3260), and Δ*lrp/plrp* (CVC3274/pJWD-2). The cells also contained a source of RctB (pTVC11) and additionally either a vector (white bars) or a *crtS*-containing plasmid (black bars). The copy numbers were normalized to the value in WT cells, set as 1. p*ori2* copy number increases threefold in the presence of p*crtS* in WT but not in Δ*lrp* cells. Complementing Δ*lrp* cells with p*lrp* increases p*ori2* copy number even more that that was seen in WT cells, suggesting that Lrp could be limiting in WT cells. Data represent mean ± SEM from three biological replicates. **(B)** Histograms of *ori2* foci number per cell in WT *V. cholerae* (CVC3058) with either the vector control (pTrc99A) or p*lrp*, and in *V. cholerae* Δ*lrp* cells (CVC3286) with complementing p*lrp* and after curing p*lrp*. Representative microscopy images from each strain are shown above the histograms. Note that the number of cells without any *ori2* foci increases in the absence of *lrp*, indicating that Lrp is crucial for Chr2 replication in *V. cholerae*. Data represent mean ± SEM from at least 1000 cells imaged from three biological replicates.

P_{crtS}, a promoter-defective mutant of *crtS* (*crtS-10m*, (Baek and Chattoraj, 2014)) was used in which two bases within the −10 element of the promoter were mutated. The mutation was previously shown to retain the replication enhancement function of *crtS* in WT *E. coli* while possessing low promoter activity. However, p*crtS-10m* failed to increase p*ori2* copy number in the *E. coli* Δ*lrp* strain (**Supplementary Figure S5**). Introduction of the p*lrp* plasmid restored the function of p*crtS-10m.* This indicates that keeping the promoter repressed may not be the sole and perhaps not the primary function of Lrp on *crtS*.

Lrp is not essential for viability of *E. coli* or *V. cholerae* (Lintner et al., 2008; Srivastava et al., 2011; Fu et al., 2013). To test the requirement of Lrp for Chr2 replication in *V. cholerae*, the *lrp* gene was deleted in a strain where fluorescently tagged *ori2* foci could be visualized. The deletion was initially made in the presence of a plasmid supplying Lrp (pJWD-2). Upon deletion of chromosomal *lrp* and curing of p*lrp*, the percentage of cells without an *ori2* focus increased dramatically (from 7 to 80%) (**Figure 4B**). This indicates that although *lrp* gene is not essential, the protein contributes dramatically to Chr2 replication. The contribution seems greater in the defined medium used for microscopy, where the growth was slower than in LB (**Supplementary Figure S6**). In fact, the Δ*lrp* strain never appears to enter logarithmic growth in the microscopy medium. At least in *E. coli*, an Lrp-associated minimal medium growth defect results largely from effects in nitrogen assimilation (Paul et al., 2007; van Heeswijk et al., 2013). The requirement of Lrp in Chr2 replication/cell growth thus exhibits media-dependency.
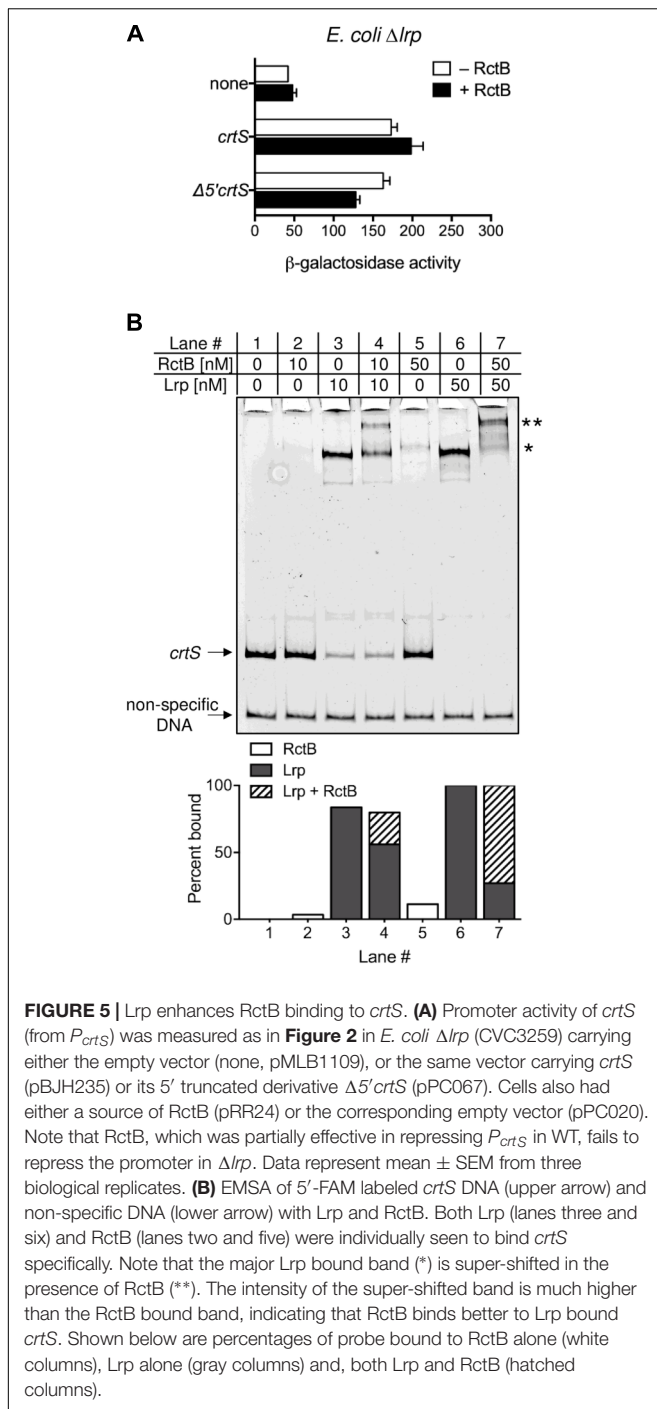
Interestingly, in cells where the complementing p*lrp* plasmid was not cured, Chr2 copy number was higher than when the cells had the empty vector (**Figure 4B**). This outcome was obtained in the WT strain containing p*lrp* as well, where 45% of cells possessed two or more *ori2* foci as compared to 30% when cells contained the empty vector, suggesting that Lrp may normally be limiting for Chr2 replication. A test of whether

Lrp functions solely via *crtS* to increase Chr2 replication was performed using a previously isolated Δ*crtS* strain (Baek and Chattoraj, 2014). This Δ*crtS* mutant possesses a mutation in *rctB,* which makes it a hyper-initiator (in WT *V. cholerae*) and apparently can compensate for the Chr2 replication defect that the absence of *crtS* confers. In this Δ*crtS* strain, p*lrp* failed to increase Chr2 replication (**Supplementary Figure S7**) and the distribution of *ori2* foci in Δ*crtS* with p*lrp,* resembled that of the vector control, indicating that Lrp functions via *crtS* in licensing Chr2 replication.

Lrp Enhances RctB Binding to *crtS*

How could Lrp stimulate *crtS* function? One possibility is that Lrp modulates the interaction of *crtS* with the rate-limiting factor for Chr2 replication, which is known to be RctB (Pal et al., 2005; Duigou et al., 2006). This hypothesis was tested *in vivo* and *in vitro*. In WT *E. coli*, RctB was able to repress the derepressed P_{crtS} activity from Δ5′*crtS* by about 50% (**Figure 2**) but was unable to do so in the isogenic Δ*lrp* strain (**Figure 5A**). These results suggest that RctB binding to *crtS* is Lrp dependent. To test this hypothesis, we performed an EMSA of *crtS* DNA with both RctB and Lrp. RctB was previously shown to bind *crtS* only when the site is supercoiled (Baek and Chattoraj, 2014). However, by changing the buffer condition, it was possible to detect RctB binding to linear *crtS* fragments (**Figure 5B**). Lrp alone also bound to the same fragment but with more affinity. When both Lrp and RctB were present, a super-shifted band was seen, indicating that both proteins are bound to *crtS* simultaneously. Presence of RctB and Lrp in the super-shifted band was confirmed by mass spectrometry.

Leucine does not affect the binding of RctB to Lrp-bound *crtS* significantly (**Supplementary Figure S8**). While the total DNA bound by Lrp alone and by Lrp+RctB remained nearly the same, the intensity of the super-shifted band increased at the expense of the species bound to Lrp alone. Apparently, RctB

FIGURE 5 | Lrp enhances RctB binding to *crtS*. **(A)** Promoter activity of *crtS* (from P*crtS*) was measured as in **Figure 2** in *E. coli* Δ*lrp* (CVC3259) carrying either the empty vector (none, pMLB1109), or the same vector carrying *crtS* (pBJH235) or its 5′ truncated derivative Δ*5′crtS* (pPC067). Cells also had either a source of RctB (pRR24) or the corresponding empty vector (pPC020). Note that RctB, which was partially effective in repressing P*crtS* in WT, fails to repress the promoter in Δ*lrp*. Data represent mean ± SEM from three biological replicates. **(B)** EMSA of 5′-FAM labeled *crtS* DNA (upper arrow) and non-specific DNA (lower arrow) with Lrp and RctB. Both Lrp (lanes three and six) and RctB (lanes two and five) were individually seen to bind *crtS* specifically. Note that the major Lrp bound band (*) is super-shifted in the presence of RctB (**). The intensity of the super-shifted band is much higher than the RctB bound band, indicating that RctB binds better to Lrp bound *crtS*. Shown below are percentages of probe bound to RctB alone (white columns), Lrp alone (gray columns) and, both Lrp and RctB (hatched columns).

binds with higher affinity to Lrp-bound *crtS* than to naked *crtS*. This is quantified by measuring RctB binding to either *crtS* or Lrp-bound *crtS* (**Supplementary Figure S9A**). The affinity of RctB to Lrp-bound *crtS* is nearly 10-fold higher than that to naked *crtS* (**Supplementary Figure S9B**). Lrp could also enhance the binding of RctB to Δ5′Δ3′*crtS* fragment (**Supplementary Figure S10**). This is not surprising, considering that Δ5′Δ3′*crtS* is functional in *E. coli* in multicopy, suggesting that RctB and Lrp can favorably interact on the Δ5′Δ3′*crtS* fragment, where at least

one Lrp binding site also exists (**Supplementary Figure S4**). Lrp thus could enhance Chr2 replication by enhancing RctB binding to *crtS*.

DISCUSSION

Requirement for Lrp in *crtS* Function

The Chr1 encoded *crtS*-mediated licensing of Chr2 replication is so far the only known mechanism by which the replication of one chromosome regulates the timing of the replication of the other. Here we report that the licensing function of *crtS* depends on the global transcription regulator, Lrp. Although many general DNA binding proteins, such as IHF, HU, Fis, and SeqA, are known to participate in DNA replication, this is the first evidence for Lrp participation, a protein sensitive to the environment and, in particular, to the intracellular concentration of leucine and other amino acids (Hart and Blumenthal, 2011). Growth phase control of DNA replication initiation is a little-studied aspect of cell cycle in bacteria, although starvation induced nucleotide alarmone (p)ppGpp has been known to inhibit new rounds of replication initiation for some time (Chiaramello and Zyskind, 1990). In the *E. coli* chromosome, the growth-phase regulated Fis protein signals to *oriC* to turn off DNA replication as the bacteria enter stationary phase (Cassler et al., 1995). So far, no Fis involvement in the origin of Chr2 replication has been detected in *V. cholerae*. The involvement of Lrp in Chr2 replication mirrors the involvement of Fis in sensitizing the chromosome to changes in cell physiology. The involvement of Lrp also makes *crtS* more comparable to DARS2 (DnaA reactivation site) of the *E. coli* chromosome (Kasho et al., 2014). Both *crtS* and DARS2 are involved in initiator remodeling, and both bind the cognate initiator and an additional factor, Lrp and Fis, respectively. The role of Lrp in *crtS* function thus could be analogous to Fis in DARS2 function.

We find that increasing Lrp concentration increases Chr2 replication (**Figure 4B**). Lrp concentration increases in stationary phase and upon other stresses to the cell (Landgraf et al., 1996). This suggests that Lrp could be utilized to promote Chr2 replication preferentially under stressful conditions. It is not possible to specify how the cells benefit from this preferential replication since functions of most of the genes in Chr2 are not known. It is known, however, that many more Chr2 genes are expressed during intestinal growth than during liquid culture, the basis of which is yet to be understood (Xu et al., 2003). One function of Lrp at *crtS* may be to maintain parity of chromosome numbers in stationary phase. In rich medium, Chr1 is maintained at two-fold higher copy number than Chr2 (Srivastava and Chattoraj, 2007; Stokke et al., 2011). When cells reach stationary phase both chromosomes have one copy each. To make this adjustment Chr2 must replicate an additional round after Chr1 replication has ceased. Increased Lrp concentrations during entry to stationary phase could help achieve this parity by stimulating Chr2 replication via *crtS*.

Lrp has been reported to control more genes in *E. coli* than any other global transcriptional regulators (Kroner et al., 2018; Shimada et al., 2018). A deletion of *lrp* in *E. coli*, however, is easily

tolerated. In contrast, a deletion of *lrp* in *V. cholerae* is obtained only in the presence of a complementing plasmid and the deleted strain shows significant growth defect (Srivastava et al., 2011). Tn-seq analysis also showed fewer hits in *lrp* compared to many other targets considered "non-essential" in *V. cholerae* (Fu et al., 2013). The requirement of Lrp in *crtS* function, and hence in Chr2 replication, may explain why in *V. cholerae* Lrp is critical. Whole genome sequencing of Δ*lrp* strains cured of complementing plasmids in this study did not reveal any suppressor mutations in 2/2 cases. At least in our growth conditions (in LB), the Δ*lrp* strains appear to be viable, although slow-growing, whereas growth is more severely affected in the poorer synthetic medium used for microscopy (**Supplementary Figure S6**). *V. cholerae* possesses three hypothetical genes with significant identity (>35%) to Lrp (**Supplementary Table S4**), in addition to the widely distributed local regulator AsnC (Caspi et al., 2016; Unoarumhi et al., 2016). It is possible that in the absence of Lrp, some of its functions could be compensated for by these paralogs. If any paralogs exist in *E. coli*, they do not seem to substitute for Lrp. In *E. coli* the protein seems to be essential for *crtS* function (**Figure 4A**).

The Importance of the Less-Conserved 5′ Region of *crtS*

An intriguing feature of *crtS* is that its 5′ region, although less conserved than the remainder of the site, is crucial for its replication enhancement function. On the other hand, that same function is unaffected by deletion of the downstream sequences, which are better conserved. The conservation of non-essential region suggests that *crtS* serves additional functions that are not yet recognized. Variant forms of Lrp or its orthologs in different species may account for the relatively poor conservation of the 5′ region. If so, this likely involves the differences at the amino termini of the different *Vibrio* Lrp orthologs (Hart et al., 2011; Unoarumhi et al., 2016). Although *crtS* sequences from different *Vibrio* species are able to increase the copy number of orthologous p*ori2*, the failure of certain *crtS* sequences to function with a few other p*ori2*, could be due to differences in their cognate Lrp proteins (Kemter et al., 2018).

The 5′ region of *crtS* provides Lrp binding sites required for promoter repression as well as the enhancement of replication initiation. The nature of the relationship of the two functions to each other remains to be clarified, but they appear to be anti–correlated: truncation that resulted in increased promoter expression reduces the efficiency with which *crtS* can license Chr2 replication. It is possible that occupancy by RNA polymerase interferes with RctB binding to *crtS*. The presence of Lrp could thus aid RctB binding to *crtS* by preventing RNA polymerase from binding to the promoter. Lrp usually forms an octameric ring composed of two tetramers, upon which DNA is wrapped, causing significant bending to the DNA (de los Rios and Perona, 2007). It is possible that the bases on *crtS* preferred by RctB are made more accessible by bound Lrp, or that constructive protein-protein contacts are made between Lrp and RctB.

The low P_{crtS} activity under our laboratory conditions measured with a transcriptional fusion to *lacZ* was also evident from previous RNA-Seq analyses (**Figure 2**) (Baek and Chattoraj, 2014; Papenfort et al., 2015). RNA-Seq reads in *V. cholerae* at low and high cell densities did not reveal any measurable transcripts originating from P_{crtS}. Unless some conditions are found that activate the promoter naturally, the presence of the promoter might well be incidental to the Lrp requirement in *crtS* function (Alice and Crosa, 2012; Lin et al., 2007). Uncoupling the replication enhancement and promoter repressor function of Lrp by mutating the −10 box of P_{crtS} did not relieve the site from Lrp dependence (**Supplementary Figure S5**). This indicates that reduction of the promoter activity cannot be the only role of Lrp. To the extent analyzed, increasing RctB binding appears to be the main function of Lrp. In the immediate future, we seek to delineate the details of interactions among *crtS*, RctB and Lrp with the ultimate aim of understanding how they help RctB to license Chr2 replication and regulate that essential function.

AUTHOR CONTRIBUTIONS

PC, RR, and DC designed the study and wrote the manuscript. PC and RR performed the experiments. All authors read and approved the final version.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2018.02103/full#supplementary-material

REFERENCES

Alice, A. F., and Crosa, J. H. (2012). The TonB3 system in the human pathogen *Vibrio vulnificus* is under the control of the global regulators Lrp and cyclic AMP receptor protein. *J. Bacteriol.* 194, 1897–1911. doi: 10.1128/JB.06614-11

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., et al. (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* 2:20060008. doi: 10.1038/msb4100050

Baek, J. H., and Chattoraj, D. K. (2014). Chromosome I controls chromosome II replication in *Vibrio cholerae*. *PLoS Genet.* 10:e1004184. doi: 10.1371/journal.pgen.1004184

Calvo, J. M., and Matthews, R. G. (1994). The leucine-responsive regulatory protein, a global regulator of metabolism in *Escherichia coli*. *Microbiol. Rev.* 58, 466–490.

Caspi, R., Billington, R., Ferrer, L., Foerster, H., Fulcher, C. A., Keseler, I. M., et al. (2016). The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res.* 44, D471–D480. doi: 10.1093/nar/gkv1164

Cassler, M. R., Grimwade, J. E., and Leonard, A. C. (1995). Cell cycle-specific changes in nucleoprotein complexes at a chromosomal replication origin. *EMBO J.* 14, 5833–5841.

Chiaramello, A. E., and Zyskind, J. W. (1990). Coupling of DNA replication to growth rate in *Escherichia coli*: a possible role for guanosine tetraphosphate. *J. Bacteriol.* 172, 2013–2019. doi: 10.1128/jb.172.4.2013-2019.1990

Cho, B. K., Barrett, C. L., Knight, E. M., Park, Y. S., and Palsson, B. O. (2008). Genome-scale reconstruction of the Lrp regulatory network in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 105, 19462–19467. doi: 10.1073/pnas.0807227105

Crooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004). WebLogo: a sequence logo generator. *Genome Res.* 14, 1188–1190. doi: 10.1101/gr.849004

Cui, Y., Wang, Q., Stormo, G. D., and Calvo, J. M. (1995). A consensus sequence for binding of Lrp to DNA. *J. Bacteriol.* 177, 4872–4880. doi: 10.1128/jb.177.17.4872-4880.1995

Das, N., and Chattoraj, D. K. (2004). Origin pairing ('handcuffing') and unpairing in the control of P1 plasmid replication. *Mol. Microbiol.* 54, 836–849. doi: 10.1111/j.1365-2958.2004.04322.x

Datsenko, K. A., and Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U.S.A.* 97, 6640–6645. doi: 10.1073/pnas.120163297

de los Rios, S., and Perona, J. J. (2007). Structure of the *Escherichia coli* leucine-responsive regulatory protein Lrp reveals a novel octameric assembly. *J. Mol. Biol.* 366, 1589–1602. doi: 10.1016/j.jmb.2006.12.032

Duigou, S., Knudsen, K. G., Skovgaard, O., Egan, E. S., Lobner-Olesen, A., and Waldor, M. K. (2006). Independent control of replication initiation of the two *Vibrio cholerae* chromosomes by DnaA and RctB. *J. Bacteriol.* 188, 6419–6424. doi: 10.1128/JB.00565-06

Egan, E. S., Fogel, M. A., and Waldor, M. K. (2005). Divided genomes: negotiating the cell cycle in prokaryotes with multiple chromosomes. *Mol. Microbiol.* 56, 1129–1138. doi: 10.1111/j.1365-2958.2005.04622.x

Ernsting, B. R., Denninger, J. W., Blumenthal, R. M., and Matthews, R. G. (1993). Regulation of the gltBDF operon of *Escherichia coli*: how is a leucine-insensitive operon regulated by the leucine-responsive regulatory protein? *J. Bacteriol.* 175, 7160–7169. doi: 10.1128/jb.175.22.7160-7169.1993

Fu, Y., Waldor, M. K., and Mekalanos, J. J. (2013). Tn-Seq analysis of *Vibrio cholerae* intestinal colonization reveals a role for T6SS-mediated antibacterial activity in the host. *Cell Host Microbe* 14, 652–663. doi: 10.1016/j.chom.2013.11.001

Golding, A., Weickert, M. J., Tokeson, J. P., Garges, S., and Adhya, S. (1991). A mutation defining ultrainduction of the *Escherichia coli* gal operon. *J. Bacteriol.* 173, 6294–6296. doi: 10.1128/jb.173.19.6294-6296.1991

Hart, B. R., and Blumenthal, R. M. (2011). Unexpected coregulator range for the global regulator Lrp of *Escherichia coli* and *Proteus mirabilis*. *J. Bacteriol.* 193, 1054–1064. doi: 10.1128/JB.01183-10

Hart, B. R., Mishra, P. K., Lintner, R. E., Hinerman, J. M., Herr, A. B., and Blumenthal, R. M. (2011). Recognition of DNA by the helix-turn-helix global regulatory protein Lrp is modulated by the amino terminus. *J. Bacteriol.* 193, 3794–3803. doi: 10.1128/JB.00191-11

Heidelberg, J. F., Eisen, J. A., Nelson, W. C., Clayton, R. A., Gwinn, M. L., Dodson, R. J., et al. (2000). DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature* 406, 477–483. doi: 10.1038/35020000

Jha, J. K., Li, M., Ghirlando, R., Miller Jenkins, L. M., Wlodawer, A., and Chattoraj, D. (2017). The DnaK chaperone uses different mechanisms to promote and inhibit replication of *Vibrio cholerae* chromosome 2. *mBio* 8:e00427-17. doi: 10.1128/mBio.00427-17

Kasho, K., Fujimitsu, K., Matoba, T., Oshima, T., and Katayama, T. (2014). Timely binding of IHF and Fis to DARS2 regulates ATP-DnaA production

and replication initiation. *Nucleic Acids Res.* 42, 13134–13149. doi: 10.1093/nar/gku1051

Kemter, F. S., Messerschmidt, S. J., Schallopp, N., Sobetzko, P., Lang, E., Bunk, B., et al. (2018). Synchronous termination of replication of the two chromosomes is an evolutionary selected feature in Vibrionaceae. *PLoS Genet.* 14:e1007251. doi: 10.1371/journal.pgen.1007251

Kroner, G. M., Wolfe, M. B., and Freddolino, P. (2018). *Escherichia coli* Lrp regulates one-third of the genome via direct, cooperative, and indirect routes. *bioRxiv* [Preprint]. doi: 10.1101/276808

Landgraf, J. R., Wu, J., and Calvo, J. M. (1996). Effects of nutrition and growth rate on Lrp levels in *Escherichia coli*. *J. Bacteriol.* 178, 6930–6936. doi: 10.1128/jb.178.23.6930-6936.1996

Lin, W., Kovacikova, G., and Skorupski, K. (2007). The quorum sensing regulator HapR downregulates the expression of the virulence gene transcription factor AphA in *Vibrio cholerae* by antagonizing Lrp- and VpsR-mediated activation. *Mol. Microbiol.* 64, 953–967. doi: 10.1111/j.1365-2958.2007.05693.x

Lintner, R. E., Mishra, P. K., Srivastava, P., Martinez-Vaz, B. M., Khodursky, A. B., and Blumenthal, R. M. (2008). Limited functional conservation of a global regulator among related bacterial genera: Lrp in *Escherichia*, *Proteus* and *Vibrio*. *BMC Microbiol.* 8:60. doi: 10.1186/1471-2180-8-60

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

Pal, D., Venkova-Canova, T., Srivastava, P., and Chattoraj, D. K. (2005). Multipartite regulation of rctB, the replication initiator gene of *Vibrio cholerae* chromosome II. *J. Bacteriol.* 187, 7167–7175. doi: 10.1128/JB.187.21.7167-7175.2005

Papenfort, K., Forstner, K. U., Cong, J. P., Sharma, C. M., and Bassler, B. L. (2015). Differential RNA-seq of *Vibrio cholerae* identifies the VqmR small RNA as a regulator of biofilm formation. *Proc. Natl. Acad. Sci. U.S.A.* 112, E766–E775. doi: 10.1073/pnas.1500203112

Paul, L., Mishra, P. K., Blumenthal, R. M., and Matthews, R. G. (2007). Integration of regulatory signals through involvement of multiple global regulators: control of the *Escherichia coli* gltBDF operon by Lrp, IHF, Crp, and ArgR. *BMC Microbiol.* 7:2. doi: 10.1186/1471-2180-7-2

Ramachandran, R., Ciaccia, P. N., Filsuf, T. A., Jha, J. K., and Chattoraj, D. K. (2018). Chromosome 1 licenses chromosome 2 replication in *Vibrio cholerae* by doubling the crtS gene dosage. *PLoS Genet.* 14:e1007426. doi: 10.1371/journal.pgen.1007426

Schaefer, J., Jovanovic, G., Kotta-Loizou, I., and Buck, M. (2016). Single-step method for beta-galactosidase assays in *Escherichia coli* using a 96-well microplate reader. *Anal. Biochem.* 503, 56–57. doi: 10.1016/j.ab.2016.03.017

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat. Methods* 9, 676–682. doi: 10.1038/nmeth.2019

Shimada, T., Ogasawara, H., and Ishihama, A. (2018). Single-target regulators form a minor group of transcription factors in *Escherichia coli* K-12. *Nucleic Acids Res.* 46, 3921–3936. doi: 10.1093/nar/gky138

Srivastava, D., Harris, R. C., and Waters, C. M. (2011). Integration of cyclic di-GMP and quorum sensing in the control of vpsT and aphA in *Vibrio cholerae*. *J. Bacteriol.* 193, 6331–6341. doi: 10.1128/JB.05167-11

Srivastava, P., and Chattoraj, D. K. (2007). Selective chromosome amplification in *Vibrio cholerae*. *Mol. Microbiol.* 66, 1016–1028. doi: 10.1111/j.1365-2958.2007.05973.x

Stokke, C., Waldminghaus, T., and Skarstad, K. (2011). Replication patterns and organization of replication forks in *Vibrio cholerae*. *Microbiology* 157, 695–708. doi: 10.1099/mic.0.045112-0

Unoarumhi, Y., Blumenthal, R. M., and Matson, J. S. (2016). Evolution of a global regulator: Lrp in four orders of gamma-*Proteobacteria*. *BMC Evol. Biol.* 16:111. doi: 10.1186/s12862-016-0685-1

Val, M. E., Marbouty, M., de Lemos Martins, F., Kennedy, S. P., Kemble, H., Bland, M. J., et al. (2016). A checkpoint control orchestrates the replication of the two chromosomes of *Vibrio cholerae*. *Sci. Adv.* 2:e1501914. doi: 10.1126/sciadv.1501914

Val, M. E., Skovgaard, O., Ducos-Galand, M., Bland, M. J., and Mazel, D. (2012). Genome engineering in *Vibrio cholerae*: a feasible approach to address biological issues. *PLoS Genet.* 8:e1002472. doi: 10.1371/journal.pgen.1002472

van Heeswijk, W. C., Westerhoff, H. V., and Boogerd, F. C. (2013). Nitrogen assimilation in *Escherichia coli*: putting molecular data into a systems perspective. *Microbiol. Mol. Biol. Rev.* 77, 628–695. doi: 10.1128/MMBR.00 025-13

Xu, Q., Dziejman, M., and Mekalanos, J. J. (2003). Determination of the transcriptome of *Vibrio cholerae* during intraintestinal growth and midexponential phase in vitro. *Proc. Natl. Acad. Sci. U.S.A.* 100, 1286–1291. doi: 10.1073/pnas.0337479100

Crosstalk Regulation Between Bacterial Chromosome Replication and Chromosome Partitioning

Gregory T. Marczynski*, Kenny Petit† and Priya Patel†

Department of Microbiology and Immunology, McGill University, Montreal, QC, Canada

Despite much effort, the bacterial cell cycle has proved difficult to study and understand. Bacteria do not conform to the standard eukaryotic model of sequential cell-cycle phases. Instead, for example, bacteria overlap their phases of chromosome replication and chromosome partitioning. In "eukaryotic terms," bacteria simultaneously perform "S-phase" and "mitosis" whose coordination is absolutely required for rapid growth and survival. In this review, we focus on the signaling "crosstalk," meaning the signaling mechanisms that advantageously commit bacteria to start both chromosome replication and chromosome partitioning. After briefly reviewing the molecular mechanisms of replication and partitioning, we highlight the crosstalk research from *Bacillus subtilis*, *Vibrio cholerae,* and *Caulobacter crescentus*. As the initiator of chromosome replication, DnaA also mediates crosstalk in each of these model bacteria but not always in the same way. We next focus on the *C. crescentus* cell cycle and describe how it is revealing novel crosstalk mechanisms. Recent experiments show that the novel nucleoid associated protein GapR has a special role(s) in starting and separating the replicating chromosomes, so that upon asymmetric cell division, the new chromosomes acquire different fates in *C. crescentus*'s distinct replicating and non-replicating cell types. The *C. crescentus* PopZ protein forms a special cell-pole organizing matrix that anchors the chromosomes through their centromere-like DNA sequences near the origin of replication. We also describe how PopZ anchors and interacts with several key cell-cycle regulators, thereby providing an organized subcellular environment for more novel crosstalk mechanisms.

Keywords: DnaA, GapR, PopZ, chromosome replication, partitioning, cell cycle

INTRODUCTION: BACTERIAL CELL CYCLES REQUIRE CROSSTALK AND COORDINATION

To ensure their survival and proliferation, bacteria overlap and compress cell-cycle processes that are complex and time consuming. This overlap in bacteria contrasts with eukaryotes, which have sequential and non-overlapping phases for chromosome replication (S-phase), partition/segregation (mitosis), and cell-division/cytokinesis. Each eukaryotic phase of the cell cycle takes time, and while their sequential ordering enables accurate checkpoint controls, this system also prolongs the cell cycle and consequently limits the growth rates. Bacteria overcome this limitation and increase their growth rates by overlapping the phases of chromosome

replication, partition/segregation, and cell wall growth/cell division (Helmstetter et al., 1968). The initiation of chromosome replication immediately precedes the initiation of chromosome partitioning and chromosome movement into separate cell spaces that will eventually become the daughter cells at cell division (Toro and Shapiro, 2010). This close temporal link suggests that it would be especially advantageous to co-regulate replication and partitioning. In a previous review article from our lab, we argued that eubacteria use one origin of replication (*ori*) per chromosome not because they are simpler organisms, but because a single *ori* allows for a more rapid and efficient control of replication (Marczynski et al., 2015). Bacterial chromosomes with one *ori* can more easily respond to many inputs both from outside and from inside the cell. We will argue that input/signals from inside the cell and crosstalk/ signals with chromosome partitioning (*par*) systems are especially important. While most studies illustrate *par* components signaling replication, we will also highlight recent studies of crosstalk in the reverse direction. However, before presenting concrete examples of crosstalk, we will first outline the basic features of both *ori* and *par* systems and emphasize their potential for regulation.

ORIGINS OF REPLICATION RECEIVE SIGNALS AND DYNAMIC PROTEIN ASSEMBLIES

Like transcription promoters, bacterial *oris* are platforms for assembling replication proteins and their regulators (Kornberg and Baker, 1992). The *Escherichia coli oriC* and DnaA model for initiating chromosome replication has revealed the most detailed molecular mechanisms that operate inside *oris* (Kaguni, 2011; Skarstad and Katayama, 2013; Kaur et al., 2014). In broad outline, a bacterial *ori* is a specific place where the DnaA protein binds an array of DnaA boxes to self-assemble and then to promote the assembly of the downstream replication proteins (Wolanski et al., 2014a,b).

In *E. coli*, chromosome replication starts from one "oriC" when a threshold level of activated DnaA (ATP bound ATP-DnaA) is reached (Katayama et al., 2010; Skarstad and Katayama, 2013). Both forms of DnaA, ATP-DnaA and ADP-DnaA, bind to the strong/high affinity DnaA boxes in *oriC*, but only the activated ATP-DnaA proteins will bind to weak DnaA box motifs and oligomerize on *oriC* through neighboring AAA+ domains (McGarry et al., 2004; Kawakami et al., 2005; Erzberger and Berger, 2006; Grimwade et al., 2018). Such DnaA self-assembly starts from strategically placed "anchor" DnaA boxes (Rozgaja et al., 2011), and the resulting protein-DNA structure (and possibly a helix) causes DNA unwinding and a further altered structure with new protein surfaces that recruit downstream replication proteins. More specifically, *oriC* DNA unwinding allows DnaA to recruit DnaB (the replicative DNA helicase) bound to DnaC, the helicase escort/loader, on to the single-stranded DNA of the AT-rich region (Mott and Berger, 2007). It is likely that two types of DnaA protein-DNA structures form on *oriC*; one

that unwinds and keeps the AT-rich region open and single stranded and another DnaA-DNA structure that recruits and loads two DnaB hexamers around the single-stranded DNA. Once loaded, the two DnaB hexamers move apart, expanding the single-stranded DNA region, thereby permitting the recruitment of primase DnaG. Next, the DNA polymerase III holoenzyme composed of the Pol III and the beta-clamp (DnaN) is recruited, and together with the clamp-loading proteins, these form the "replisome" that synthesizes the complementary DNA strands (Kaguni, 2011; Skarstad and Katayama, 2013; Katayama, 2017).

Since most eubacteria use the DnaA protein to initiate chromosome replication (Wolanski et al., 2014a,b), DnaA and the assembly reactions at *oriC* are major targets for the regulators of chromosome replication (Wolanski et al., 2014a,b). Recent reviews have described many proposed and established regulators of replication, and an especially good review with fine graphic summaries was provided by Katayama et al. (2010).

Most importantly for our topic, DnaA assembly at *oriC* is dynamic, and *in vivo* there is probably both back and forth assembly and dis-assembly of DnaA until the critical amount of DnaA oligomerization and active structure formation is reached (Leonard and Grimwade, 2011; Kaur et al., 2014). This dynamic feature of *E. coli* replication initiation implies that there are many ways to shift the assembly versus dis-assembly of DnaA and DnaB. This process has the potential to integrate many signals that can be constantly added or subtracted in real time before the final commitment to replication is made. We will describe below how this view of dynamic *oriC*/DnaA assemblies helps us to understand the regulatory crosstalk with chromosome partitioning.

DNA PARTITIONING SYSTEMS

Many bacteria use systems often called "*parABS*" for mitotic-like chromosome separation and partitioning into cell compartments, and their proximity to origins of replication (*oris*) suggests functional linkages (Livny et al., 2007). These partitioning systems were originally studied on large low-copy plasmids, and they account for faithful and consistent plasmid distribution to both progeny cells (Austin and Abeles, 1983; Ogura and Hiraga, 1983; Gerdes et al., 1985). Despite much effort, exactly how the *parABS* systems work to move and to position plasmid and chromosome DNAs remains incompletely understood and in parts controversial (Gerdes et al., 2010). Here we want to present the basic information and sketch what appear to us the most relevant models for our topic. Knowledge of the detailed mechanisms is required not just to understand how *parABS* systems work to partition DNA but also to understand and speculate how evolution has harnessed these systems for other functions and particularly for crosstalk with chromosome replication. With respect to deep evolutionary potentials, *parABS* systems have also been harnessed for protein positioning and localization, as, for example, organizing chemotaxis proteins and other large protein assemblies (Vecchiarelli et al., 2012).

In bare outline, the three-component *parABS* system works as follows: The *parS* DNA acts as a "centromere-like" locus with specific DNA sequences that bind and hold ParB proteins. ParA protein binds and hydrolyses ATP, and it somehow imparts motion to the ParB-*parS* complex through interactions with ParB. These basic functions need to be controlled and organized by regulators and structures that change during the cell cycle. As a main topic, we will address some key regulators and structures below, including, for example, the cell-pole proteins that anchor the chromosome ParB-*parS* complexes.

There are several significant variations to the above bare outline of *parABS* system. For example, some bacteria apparently use several *parS* loci, while others appear to use just one. *Bacillus subtilis* probably uses 10 *parS* loci and 8/10 loci cluster toward the *oriC* side of the chromosome (Breier and Grossman, 2007). *Myxococcus xanthus* may use as many as 22 *parS* loci, likewise near its *oriC*, for partitioning its exceptionally large 9.1 Mb chromosome (Iniesta, 2014). In contrast, the *Caulobacter crescentus* (Mohl and Gober, 1997) and the *Vibrio cholerae* (Espinosa et al., 2017) chromosomes appear to use just one *parS* per chromosome, and these single *parS* loci are likewise closely linked to their corresponding *oriCs*. Why does one bacterium need one *parS* and another several? There is no good answer yet, but this distinction may be too simplistic. For example, a recent study showed that *C. crescentus* has several yet substantially weaker ParB-binding sites (Tran et al., 2018), and it may be more correct to speak of a "*parS* region" surrounding the origin of replication as described further below.

There are also significant variations in how ParB binds DNA to create a "centromere-like" locus. ParB binds specifically to *parS* DNA and less specifically to other parts of the chromosome. First, ParB binds specifically to an inverted DNA repeat that is typical of many standard dimeric helix-turn-helix DNA-binding proteins, and these sites are easily found and used to identify *parS* sites in most bacterial genomes (Livny et al., 2007; Iniesta, 2014). However, ParB is reported to have additional modes of DNA binding. *In vivo* cross-linking and transcription reporter experiments imply that ParB binds to *parS* sites and then spreads to adjacent DNA as if forming a filament across the DNA to distant sites. It is not likely that "spreading" is an experimental artifact because spreading is required for partitioning. ParB mutants that do not spread do not partition DNA (Rodionov et al., 1999; Graham et al., 2014).

The exact DNA/protein structure(s) of these "spreading" ParB molecules is not known, but interactions can be inferred from crystal structures (Chen et al., 2015). ParB can bind other ParB molecules through lateral contacts that reach adjacent DNA and through bridging contacts that bring distant DNAs together with loops. This capacity for non-specific DNA binding suggests that ParB can be classified as one among many nucleoid-associated proteins (NAPs) that compact and organize bacterial chromosomes. Recently, a ParB "caging model" has been proposed whereby *parS* organizes a large chromosome subdomain through dynamic ParB-ParB and ParB-DNA interactions (Funnell, 2016). This model is further supported by *in vitro* experiments with magnetic tweezers, which suggest

that the overall ParB-DNA complex is not well ordered and vaguely resembles a phase separation from the rest of the nucleoid (Taylor et al., 2015). In summary, considering the proximity of *parS* to *oriC*, ParB protein certainly has the potential to influence chromosome replication, and we will describe specific examples below.

PARTITION PROTEIN PARA CAN BE A MOTOR AND A REGULATOR

The preceding observations argue that ParA imparts motion not just to a small ParB-*parS* locus but also to a large ParB-DNA subdomain of the chromosome. Exactly how ParA drives ParB-DNA motion also remains controversial. However, ParA has several established and speculative properties that enable it to serve both as a propeller and as a regulator. We will focus below on two properties required for regulation: First, we explain that ParA can act like a "molecular switch" and second, we explain that ParA (like ParB) can bind and influence large domains of DNA.

ParA "switches" within a biochemical cycle: ParA monomers bind ATP, the ParA-ATP dimerizes, and this form binds DNA non-specifically. ATP hydrolysis creates ParA-ADP molecules, which disassociate from the DNA as monomers. When ParA binds ParB, specific protein-protein contacts stimulate ATP hydrolysis, thereby resetting the ParA-ATP/DNA binding versus ParA-ADP/DNA release cycle (Vecchiarelli et al., 2010). A protein contact switch seems ideal for regulation, and as an interesting example, we will describe below how *Bacillus subtilis* has harnessed ParA to also regulate chromosome replication through direct contacts with DnaA.

Exactly how this ParA cycle drives ParB-DNA motion remains controversial. It is also not clear if propulsion and switching/regulation are separable functions. Here we can only superficially comment on this literature, and we will focus on how ParA binds to the nucleoid. For example, it has been proposed that ParA binds ParB and then retracts to pull the ParB-DNA along its path. This could be an active process where ParA imparts the force of motion or it could be a more passive mechanism, for example, a "catch and release" mechanism whereby ParA guides and biases a random "DNA flapping" motion. ParA may be organized as "microtubule-like" or as "cloud-like" structures that move forward and recede by assembly and dis-assembly. The literature is not consistent. However, there are credible reports that during partition, ParA forms dynamic cloud-like patterns on the surface of the nucleoid, and this pattern is interpreted as a gradient that recedes and seems to draw the ParB bound to *parS* (Hatano and Niki, 2010; Ah-Seng et al., 2013). Nucleoid patterning by ParA proteins resembles membrane patterning by the *E. coli* MinCDE system (Vecchiarelli et al., 2012), which imparts positional information for cell division, so that the septum forms at mid-cell (Lutkenhaus, 2007). Furthermore, the ParA and Min proteins belong to the same class of ATPases, and their mechanisms for molecular positioning may be fundamentally similar (Vecchiarelli et al., 2012).

Ietswaart et al. have presented an important synthesis between what seemed at first to be distinct and contradictory par mechanisms (Ietswaart et al., 2014). They demonstrate that the *par* system stimulates plasmid DNA motion above the random Brownian motion kinetics, thereby demonstrating that the *par* system can impart an active force and does not simply bias a random motion. Also, very importantly, Ietswaart et al. have argued that nucleoid structure plays an essential role in ordering the bound ParA-ATP structures. For example, helical nucleoid folds might provide grooves for channeling ParA-ATP aggregates into filaments or elongated clouds. Their model requires linear arrays of DNA-bound ParA-ATP and not necessarily that they be microtubule-like filaments. In other words, ParA-ATP linearity imparts the directionality to DNA motion and short disjoint filaments or elongated clouds (where individual ParA-ATP dimers bound to the nucleoid need not touch) will equally satisfy their model. In summary, the *par* literature argues that both ParA and ParB shape and respond to the structure of the nucleoid. Consequently, NAPs should significantly impact both chromosome replication and its partitioning. We will therefore discuss NAPs as regulators further below.

ESTABLISHED EXAMPLES OF CROSSTALK: THE *BACILLUS SUBTILIS* SYSTEM

Bacillus subtilis provides clear examples of crosstalk and a series of papers provide the best and earliest evidence. For example, early studies showed that *B. subtilis*, Spo0J(ParB), is required for the normal positioning of the *oriC* region and for restricting its replication. Wild type cells prior to replication place their *oriC* regions at the lateral mid-cell position and when they duplicate their *oriC* regions, they position them around the cell quarter-length positions. However, in *spo0J(parB)*-null strains, the duplicated *oriC* regions are positioned significantly closer together and toward the mid-cell. Interestingly, these *spo0J(parB)*-null strains had more *oriC* DNA per cell, as determined by flow cytometry. Apparently, *spo0J(parB)*-null cells had increased chromosome content from an excessive and/or an asynchronous initiation of DNA replication from *oriC* (Lee et al., 2003).

One general question is whether asynchronous firing of *B. subtilis oriC* was caused indirectly by *oriC* mislocalization or whether the ParAB system directly interacts with the replication system. Later studies showed that the *B. subtilis* ParAB proteins directly target DnaA (Murray and Errington, 2008). Using fluorescence-tagged ParA and ParB proteins, Murray and Errington showed that these proteins dynamically localize as specific foci (spots) near *B. subtilis* cell poles and nucleoids and that ParA can both inhibit and activate DnaA to alter chromosome replication. The inferred cytogenetic interactions between ParA and DnaA were supported by direct *in vivo* crosslinking and two-hybrid assays. In addition to this direct mechanistic link, this article also made several other interesting

observations: For example, *parA*-null mutants behave like wild-type cells arguing for redundant or multiple regulatory inputs. Revealing the hidden cell-cycle interactions required assaying mutant protein forms. For example, revealing DnaA-dependent ParA foci at *oriC* required expressing a fluorescent ParA protein that bound ATP but did not bind DNA. Presumably, the weaker binding of ParA to DnaA protein at *oriC* would be otherwise obscured by its stronger binding to the larger/bulkier chromosome DNA. Similarly, revealing ParB-dependent ParA foci required fluorescent ParA that was deficient for ATPase and therefore apparently remained bound for longer times to the DNA.

Furthermore, the cell-cycle roles of ParAB were originally hidden because *parAB* mutants were first classified as sporulation genes. ParB was called Spo0J because null alleles were blocked in the earliest 0-stage of sporulation. ParA was called Soj, "suppressor of spo gene J," because its null alleles allowed sporulation of spo0J null strains (Ireton et al., 1994; Quisel and Grossman, 2000). We now know that sporulation is inhibited by ParA (Soj), which requires ParA-ATP dimerization and that ParB (Spo0J) counteracts ParA (Soj) by stimulating ParA-ATP hydrolysis. Murray and Errington also showed that ParA (Soj) acts through the Sda-dependent DNA replication checkpoint (Murray and Errington, 2008). Sporulation is not just a simple response to starvation. Sporulation also requires passing several checkpoints and conditions that perturb chromosome replication block sporulation by expressing a sporulation inhibitor, Sda (Ruvolo et al., 2006). Most interestingly, the transcription promoter of *sda* has many DnaA boxes, and like *oriC*, it essentially acts as a sensor for DnaA activity. In other words, one had to look through one layer of regulation (Sda check point regulation) to see the other layer of *oriC*/DnaA regulation. Note also that both sporulation and chromosome replication are long processes that require a "full commitment" following a "deliberation process" with multiple inputs, and that evolution has recruited DnaA in both cases as an integrating component.

Subsequent studies showed how ParA changes DnaA oligomerization at the *B. subtilis oriC*. For example, Scholefield et al. showed that the initiation of chromosome replication is inhibited by monomeric ParA-ADP (Soj) and conversely activated by dimeric ParA-ATP (Scholefield et al., 2011). This study also identified specific amino-acid contacts on coregulator ParB (Spo0J) that touch ParA and "flip the switch" to its inactive form. Next, in their following paper, Scholefield et al. demonstrated specific amino-acid contacts between ParA and DnaA with both molecular-genetic and biochemical (e.g. SPR sensorgram and crosslinking) experiments. Most impressively, this study showed that monomeric ParA represses *oriC* replication by depolymerizing DnaA (Scholefield et al., 2012). These experiments used a functional double-cysteine version of DnaA (DnaA-CC) that allowed stable crosslinking of the DnaA-CC oligomers during *in vitro* and *in vivo* experiments. These oligomers presumably reflect the assembly of the DnaA-*oriC* DNA complexes, and their summary model implies that monomer ParA acts as a negative input during the dynamic assembly and dis-assembly process that tips *oriC* either toward or away from replication.

Recent microscopic studies have more directly confirmed this rapid assembly and dis-assembly model of DnaA at *B. subtilis oriC* and the proposed regulatory roles of ParA (Soj) in this dynamic process (Schenk et al., 2017). More specifically, "FRAP" fluorescence recovery and photobleaching analysis of a functional fluorescent YFP-DnaA protein showed that DnaA is bound to *oriC* with a short half-time of only 2.5 s. As predicted, a genetic deletion of *parA* (*soj*) increased the DnaA residence time at *oriC* and this in turn caused over-replication of the chromosome, presumably by shifting the equilibrium more frequently toward DnaA-*oriC* DNA complex formation. Furthermore, single-molecule YFP-DnaA microscopy showed that DnaA oscillates between polar-oriented *oriC* foci with a very short ~2 s periodicity. This last observation unexpectedly shows that DnaA can behave more like the *par* and *min* (cell division) system proteins than previously suspected (Schenk et al., 2017).

The overall view that emerges from these studies is that ParA (Soj) is an important *oriC*/DnaA regulator or more accurately, a key regulatory input. This regulation is not essential but instead seems to fine tune the cell cycle in growing cells and their timely exit into sporulation. ParA (Soj) can either delay or advance the start of *oriC* replication depending on its monomer versus dimer states and its contacts with ParB (Spo0J). However, exactly how these factors link *oriC*/DnaA regulation to chromosome movements and perhaps to other cell-cycle processes remains vague and speculative.

ESTABLISHED EXAMPLES OF CROSSTALK: THE *VIBRIO CHOLERAE* SYSTEM

Vibrio cholerae presents another interesting, evolutionary very divergent and well-studied system for addressing chromosome replication and partitioning. This topic has recently been well reviewed (Espinosa et al., 2017). *V. cholerae* is closely related to *E. coli*, and while these bacteria have expected similarities, they also have some very surprising differences. For example, the *V. cholerae oriC* and the *E. coli oriC* seem to function and use DnaA very similarly. However, unlike *E. coli*, *V. cholerae* has two chromosomes, one replicated by an *E. coli*-like *oriC* (Chrom I) and the other by a distinct plasmid-like origin of replication (Chrom II). The *V. cholerae* Chrom I and *E. coli oriC*s have identical DnaA box distributions, and as expected, DnaA is the primary initiator (Egan and Waldor, 2003). In contrast, the *V. cholerae* Chrom II *ori* has only one DnaA box, and it instead uses an "iteron" organization, i.e., a long array of binding sites for the initiator protein RctB (Gerding et al., 2015). Yet, despite such major differences both Chrom I and II are well integrated into the *V. cholerae* cell cycle, and their replication is strictly timed (Espinosa et al., 2017).

V. cholerae Chrom I and II have evolved separate and specific replication and partitioning crosstalk systems. For example, the control of chromosome replication through ParA and ParB, seen above in *B. subtilis*, also seems to apply to the large

chromosome (Chrom I) of *V. cholerae* (Kadoya et al., 2011). Interestingly, each Chrom I and II has its own chromosome-specific *parABS* system. Accordingly, Chrom I has corresponding *parA1*, *parB1*, and *parS1* linked to its *E. coli*-like *oriC*. Prior to the start of replication, this *V. cholerae par/oriC* DNA region is positioned at the cell pole. Deletion of either *parA1* or *parS1* caused delocalization away from the cell pole. Deletion of *parB1* caused a similar delocalization as expected, plus an increased *oriC* copy number indicating that lack of ParB1 causes over-replication. Therefore, as in *B. subtilis*, ParB1 limits ParA1 activity, which then presumably targets *oriC* through DnaA. This view is supported, by, for example, double *parB1* and *parA1* deletions, which reduce and restore approximately normal levels of *oriC* replication presumably by eliminating the stimulus of ParA1-ATP dimers. Unfortunately, direct evidence for ParA1 and *V. cholerae* DnaA interactions is lacking. It is tempting to speculate that like *B. subtilis* ParA (Soj), the *V. cholerae* ParA1 also directly contacts the AAA+ domain of DnaA and more specifically that it too both stabilizes and destabilizes the DnaA structure on *oriC*. However, there are many ways to regulate DnaA activity, and considering the evolutionary distance between Gram (+) and Gram (−) bacteria, other mechanisms are likely, and the details of this broad outline need to be investigated.

The *V. cholerae* (*Vc*) Chrom II system is significantly different from Chrom I: Its *ori* is flanked by two genetic loci *rctA* and *rctB* (Egan and Waldor, 2003). While *rctB* simply encodes the DNA-binding initiator protein, the *rctA* locus seems to be a complex regulatory system with the *Vc parS2* "centromere" embedded among its regulatory elements (Gerding et al., 2015). Also, Chrom II seems to have an interesting parallel regulation with that of the *Caulobacter crescentus* (*Ccr*) chromosome, which will be described further below: As with most *parABS* systems, the *Vc parS* centromere locus in *rctA* binds *Vc* ParB2 and the *Ccr parS* binds *Cr* ParB. However, very interestingly, both centromere loci also bind their main replication initiator proteins, *Vc* RctB (Gerding et al., 2015) and *Ccr* DnaA, respectively (Mera et al., 2014). This is probably an example of convergent functional evolution, since *Vc* RctB and *Ccr* DnaA are otherwise unrelated.

Despite these two clear examples of crosstalk, the details of their mechanisms, as far as they are known, appear to be very different. The details of *Ccr parS* and *Ccr* DnaA interactions will be described further below in the context of cell-cycle control. Here we will note some mechanistic similarities and differences. For example, the *rctA/parS2* locus of *Vc* Chrom II binds RctB protein and seems to repress replication by titrating RctB away from the nearby origin of replication (Yamaichi et al., 2011). This is clearly different than *Ccr* DnaA protein that binds *parS* to apparently trigger DNA movement. Also, RctB has at least two separate DNA-binding domains (Yamaichi et al., 2011), one to bind *rctA* DNA and the other to bind the iteron motifs inside the adjacent Chrom II *ori*. In contrast, DnaA uses its single domain IV to bind DnaA boxes in both *parS* and *ori* DNA (Mera et al., 2014). Moreover, *rctA/parS2* seems to be a more complex locus. Its small ORF does not seem to encode a functional protein, and instead, it seems to function by providing

an RNA molecule, or as a platform for transcription activity (perhaps to alter DNA topology), and as a platform for binding proteins, including ParB2 (at the main *parS2* sequences) and RctB. Both ParB2 and RctB can bind and simultaneously occupy *rctA* DNA in what appears to be adjacent binding zones (Yamaichi et al., 2011). ParB2 binding to *rctA* DNA counteracts *rctA* repression of replication, yet ParB2 protein does not seem to displace the bound RctB protein. This last observation argues that simple RctB protein titration away from the *ori* does not obviously explain how the *rctA* locus acts through RctB protein to repress replication or how ParB2 binding counteracts this effect. A fuller explanation is needed, and it may need to invoke altered protein and DNA structures.

Separate studies confirm the preceding antagonistic relationships among *rctA/parS2*, ParB2, and RctB, but the inferred mechanism does not involve RctB titration (Venkova-Canova et al., 2013). Instead, it was argued that RctB binds short 12-mer DNA sequences to activate replication and to longer 39-mer DNA sequences to repress replication. Apparently, ParB2 has two ways to relieve this repression. In the first way, ParB2 binds at *rctA/parS2* and spreads laterally across the DNA into a nearby 39-mer, thereby displacing RctB and relieving its repression. In the second way, ParB2 has a secondary intrinsic affinity for the 39-mer DNA, and so ParB2 competes for RctB repressor binding at a distant 39-mer without the spreading mechanism from *parS2*.

Furthermore, RctB and ParB2 provide a second level of crosstalk since they control transcription of the downstream *parAB2* operon. As observed in similar *par* systems, ParB2 binds *parS2/rctA* and auto-represses the *parAB* operon. However, RctB binding stimulates transcription, thereby increasing ParB2. Therefore, RctB and ParB2 have mutually antagonistic effects on both *parAB2* operon transaction and on Chrom II replication (Yamaichi et al., 2011; Gerding et al., 2015). Overall, these observations suggest a dynamic back and forth switch between *par* and *ori* control that is yet to be fully understood.

In summary, the *V. cholerae* two chromosome system provides interesting examples of *ori* and *par* crosstalk. At Chrom I, evolution has apparently conserved the ParB1, ParA1, and DnaA signaling pathway between *parS1* and the origin of replication. However, at Chrom II, evolution has modified the paralogous ParB2 protein to interact more directly with a very different type of origin of replication through direct contact or through competition with its iteron-binding protein RctB.

THE *CAULOBACTER CRESCENTUS* CELL-CYCLE MODEL FOR CROSSTALK

Caulobacter crescentus provides further evidence of *ori* and *par* crosstalk. As a chief advantage, this bacterium allows crosstalk studies in the context of a synchronized and well-studied cell cycle (**Figure 1A**). This is a "di-morphic" cell cycle where the transition from the "swarmer cell" to the "stalked cell" also marks the key steps of replication and chromosome partitioning. Conceptually, the cell cycle starts with the motile and non-replicating

swarmer cell. Its chromosome replication is blocked by the CtrA regulator with five-binding sites inside the *C. crescentus* origin of replication (*Cori*) (Siam and Marczynski, 2000). The *C. crescentus parS* is only about 8 kb from *Cori* (Mohl and Gober, 1997), and this whole region of the chromosome is polarized and held near the flagellated cell pole by *parS*-binding ParB, which in turn is bound to a polar matrix protein called "PopZ" (Bowman et al., 2008). We will describe PopZ further below and argue that it can serve as a "hub" for many regulatory interactions, but the most conspicuous role for PopZ is to serve as the substrate that binds ParB, which thereby anchors the *parS* and *Cori* region in the swarmer cell (Bowman et al., 2008).

The cell-cycle transition from swarmer cell to stalked cell coincides with many molecular events that suggest *Cori* and *parS* crosstalk (Marczynski et al., 2015). While the swarmer cell ejects its flagellum and starts to grow its stalk (a tubular cell wall outgrowth), the CtrA protein is inactivated (dephosphorylated) and degraded, the *parS-Cori* region detaches from the cell pole, and the chromosome initiates replication from *Cori* (Toro and Shapiro, 2010). The initiation of chromosome partitioning is practically simultaneous with the initiation of chromosome replication, and both processes continue through most of the cell division cycle. Note especially that the dividing cell poles are different, one pole has a stalk, while the other is building a new flagellum, so this is an "asymmetric" cell division cycle (**Figure 1**). Eventually, one whole chromosome is placed in the nascent swarmer cell compartment, while the other chromosome is placed in the stalked cell compartment. In other words, with respect to chromosome replication, one chromosome will be placed into an inactive swarmer cell, and the other chromosome will be placed into an active stalked cell.

Such asymmetric cell division implies that the initiation of chromosome replication and partitioning coincide with the critical chromosome symmetry-splitting step of the cell cycle (**Figure 1A**). Time-lapse fluorescence microscopy showed that *parS-Cori* region DNA partitioning (visualized with fluorescent ParB) is a complex process involving at least the following steps: First *parS-Cori* separation, then *parS-Cori* discrimination, such that one *parS-Cori* region seems to be chosen for reattachment to the stalked pole. Then, the other (apparently unattached) *parS-Cori* region moves slowly away from the stalked pole to approximately the quarter-cell length position before moving more rapidly to the new swarmer pole (Shebelut et al., 2010). Further analysis showed that only the last fast-movement phase requires ParA ATPase activity and that the early slow movement of *parS-Cori* to the quarter-cell length position occurs faithfully when a dominant-negative ParA allele is expressed (Shebelut et al., 2010). Since fluorescently labeled ParB is bound to *parS* during this early slow-movement phase, then how does *parS*-ParB move without ParA?

More importantly, how do these early partitioning steps faithfully split chromosome symmetry to channel them toward two different cell fates (**Figure 1A**)? What are the regulators and the motors during the early partitioning steps? How do they communicate with chromosome replication? These questions are starting to be addressed in the following paragraphs.
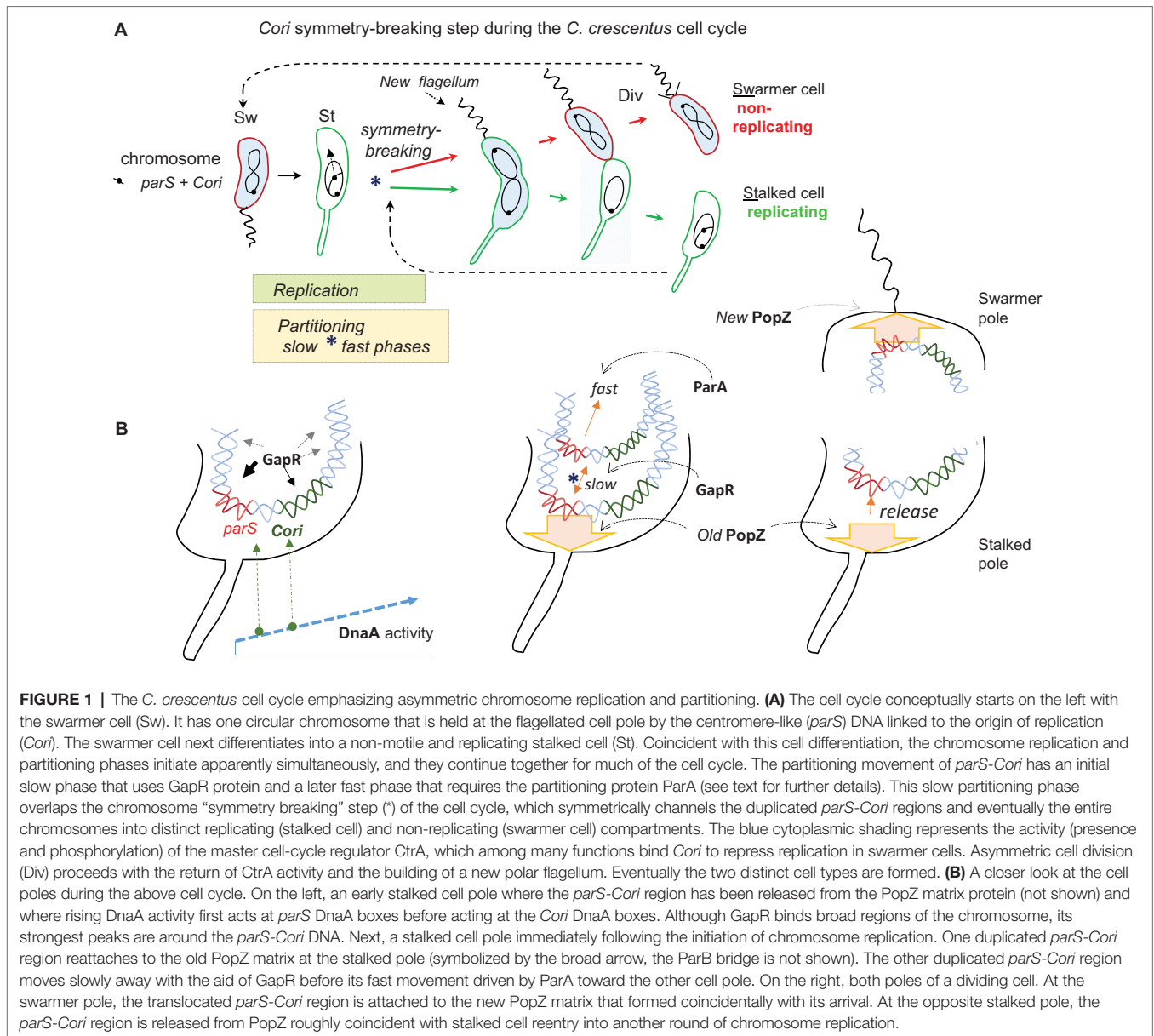
FIGURE 1 | The *C. crescentus* cell cycle emphasizing asymmetric chromosome replication and partitioning. **(A)** The cell cycle conceptually starts on the left with the swarmer cell (Sw). It has one circular chromosome that is held at the flagellated cell pole by the centromere-like (*parS*) DNA linked to the origin of replication (*Cori*). The swarmer cell next differentiates into a non-motile and replicating stalked cell (St). Coincident with this cell differentiation, the chromosome replication and partitioning phases initiate apparently simultaneously, and they continue together for much of the cell cycle. The partitioning movement of *parS-Cori* has an initial slow phase that uses GapR protein and a later fast phase that requires the partitioning protein ParA (see text for further details). This slow partitioning phase overlaps the chromosome "symmetry breaking" step (*) of the cell cycle, which symmetrically channels the duplicated *parS-Cori* regions and eventually the entire chromosomes into distinct replicating (stalked cell) and non-replicating (swarmer cell) compartments. The blue cytoplasmic shading represents the activity (presence and phosphorylation) of the master cell-cycle regulator CtrA, which among many functions bind *Cori* to repress replication in swarmer cells. Asymmetric cell division (Div) proceeds with the return of CtrA activity and the building of a new polar flagellum. Eventually the two distinct cell types are formed. **(B)** A closer look at the cell poles during the above cell cycle. On the left, an early stalked cell pole where the *parS-Cori* region has been released from the PopZ matrix protein (not shown) and where rising DnaA activity first acts at *parS* DnaA boxes before acting at the *Cori* DnaA boxes. Although GapR binds broad regions of the chromosome, its strongest peaks are around the *parS-Cori* DNA. Next, a stalked cell pole immediately following the initiation of chromosome replication. One duplicated *parS-Cori* region reattaches to the old PopZ matrix at the stalked pole (symbolized by the broad arrow, the ParB bridge is not shown). The other duplicated *parS-Cori* region moves slowly away with the aid of GapR before its fast movement driven by ParA toward the other cell pole. On the right, both poles of a dividing cell. At the swarmer pole, the translocated *parS-Cori* region is attached to the new PopZ matrix that formed coincidentally with its arrival. At the opposite stalked pole, the *parS-Cori* region is released from PopZ roughly coincident with stalked cell reentry into another round of chromosome replication.

C. CRESCENTUS DNAA ALSO SIGNALS CHROMOSOME PARTITIONING

A study by Mera et al. implicated *C. crescentus* DnaA in chromosome partitioning (Mera et al., 2014). A conditional DnaA expression strain failed to initiate chromosome replication when DnaA was shut-off, as expected (Gorbatyuk and Marczynski, 2001), and as expected kept a single fluorescent ParB-*parS* centromere complex at the old stalked cell pole while the cell attempted to grow and divide. However, and very surprisingly, DnaA expression at low levels that could not initiate chromosome replication could still initiate and complete *parS-Cori* partitioning. Under these low DnaA conditions, many cells that had only a single, i.e. an un-replicated ParB-*parS* centromere complex

could still move it completely from the old stalked pole to the new swarmer cell pole. Mera et al. clearly showed that DnaA binds the *parS* and that DnaA-ATP is required for this partitioning since a DnaA allele that does not bind ATP does not support partitioning. The view suggested by these results is that as DnaA activity rises (as both protein abundance and DnaA-ATP) during the swarmer cell to stalked cell transition, DnaA first acts at *parS* to perhaps commit the chromosome to partitioning before acting at *Cori* to commit it to chromosome replication (**Figure 1B**). This view is attractive considering that DnaA often acts as a global regulator of cell-cycle gene expression (Hottes et al., 2005) and chromosome replication (Gorbatyuk and Marczynski, 2001) and now apparently chromosome partitioning as well.

CONTROL BY NUCLEOID-ASSOCIATED PROTEINS

Nucleoid organization can theoretically impact both chromosome replication and partitioning (Badrinarayanan et al., 2015). Unlike eukaryotic cells, bacteria do not possess histones, and instead, several small proteins called nucleoid-associated proteins (NAPs) compact and organize their genomes (Luijsterburg et al., 2006; Stavans and Oppenheim, 2006). Bacterial NAPs are not always conserved, but they share many features such as a small size, a high expression level, and a tight DNA binding (Krogh et al., 2018). NAPs impact DNA topology, which must be regulated for efficient transcription and replication (Donczew et al., 2014; Dorman and Dorman, 2016). For example, negatively supercoiled genes are more efficiently transcribed than positively supercoiled genes suggesting transcriptional control by NAPs (Sobetzko et al., 2012). The most investigated and one of the most conserved NAPs is the HU protein of *E. coli* (Ali Azam et al., 1999). HU exists as a homo- or hetero-dimer of α and β chains depending on the growth phase. DNA-binding affinity is different for each dimer, leading to differential nucleoid compaction and differential transcription between the growth phases. HU also stabilizes the pre-replication complex essential for the initiation of *E. coli oriC* replication (Chodavarapu et al., 2008). Other NAPs also influence the initiation of DNA replication. In *E. coli*, NAPs "FIS" and "IHF" repress and stimulate the initiation of DNA replication (Ryan et al., 2004; Wolanski et al., 2014a,b). In *B. subtilis*, the NAP "ROK" recruits and interacts with the bacterial replication initiator DnaA. ROK thereby directs DnaA to repress transcription and to help shape the nucleoid (Seid et al., 2017).

C. CRESCENTUS GAPR IS A NOVEL NAP THAT AIDS CHROMOSOME REPLICATION AND PARTITIONING

In *C. crescentus*, the recently identified and now best characterized NAP "GapR" is implicated in cell-cycle control including chromosome replication and partitioning. GapR is an essential 89 amino-acid protein exclusively found in the alpha-proteobacteria, which are also known for their asymmetric cell division (Brilli et al., 2010). It is therefore tempting to speculate that GapR contributes to the chromosome asymmetry of *C. crescentus* (**Figure 1A**). Recent papers report that GapR has several relevant properties. For example, GapR binds DNA in AT-rich regulatory regions and next to highly expressed genes. Interestingly, the bulk distribution of GapR on the chromosome forms a gradient that decreases from the *parS-Cori* region to the terminus region (Arias-Cartin et al., 2017). Recently, another NAP "HupB" in *M. smegmatis* (Holowka et al., 2017) was shown to have a similar chromosome-wide gradient distribution. In the absence of GapR, both DNA replication and cell division are impaired (Arias-Cartin et al., 2017; Taylor et al., 2017). However, depletion of GapR only

slightly affects global gene expression and most of the genes that are overexpressed belong to the DNA damage stress response and could be induced by indirect DNA damage. These observations argue that GapR is not primarily a transcription regulator. ChIP-seq analysis and fluorescence microscopy have shown that binding of GapR on the chromosome is dynamic and changes throughout the cell cycle. The strongest GapR peaks accumulate near *Cori* and downstream of *parB* near *parS* during the initiation of DNA replication (Taylor et al., 2017). Through this binding GapR somehow enhances the early slow phase of chromosome partitioning (**Figure 1B**), because without GapR, the *parS-Cori* region duplicates and then collapses into one focus before repeating the separation/partitioning process. This is the critical time when separation of the two chromosomes directs them to their alternative fates (Taylor et al., 2017). Subsequently, as the chromosome replicates and partitions, GapR localization correlates with the moving replisome and the replication fork seems to displace the protein from the DNA (Arias-Cartin et al., 2017). Consistent with these observations, X-ray protein crystallography has shown that two GapR dimers assemble to encircle DNA that must be overly twisted to fit inside the hole (Guo et al., 2018). Such overly twisted DNA is either found in front of the replication forks or downstream highly transcribed genes. Although the molecular details have still to be explored, it was proposed that once bound to the overly twisted DNA, GapR enhances or recruits the gyrase activity to dissipate (+) supercoiled DNA produced by replication forks and by RNA polymerase (Guo et al., 2018). Therefore, unlike most NAPs that primarily compact the nucleoid, GapR seems to primary facilitate nucleoid replication and partitioning perhaps at least in part by strategically directing DNA gyrase and perhaps other "molecular machines" including RNA and DNA polymerases.

C. CRESCENTUS PROTEIN POPZ IS A POLAR ORGANIZING "HUB"

Multiple cell-cycle regulators act through the cell poles, and PopZ is their polar "hub protein" acting at the heart of chromosome replication and partitioning (Bowman et al., 2010). PopZ is an intrinsically disordered network protein that fills and forms special apical zones in the cytoplasm. Molecular recognition features "MoRFs" (Holmes et al., 2016) allow PopZ to engage and to localize many cell-cycle proteins. PopZ is initially found at the cell poles, where it binds ParB to anchor *parS* (Bowman et al., 2008; Ebersbach et al., 2008). In addition to this key function, PopZ serves as a platform for other cell-cycle regulators. For example, CtrA and its kinases CckA regulate chromosome replication. CtrA binds *Cori* and both CtrA and CckA are recruited to the stalked cell pole in a PopZ-dependent manner (Bowman et al., 2010; Holmes et al., 2016). Moreover, PopZ sequesters and restrains the CtrA-targeting protease ClpXP (Joshi et al., 2018). In the absence of PopZ, ClpXP exhibits unprecedently high CtrA degradation rates. Under normal conditions, the

PopZ-recruited adaptor protein CdpR modulates ClpXP activity also by CckA-mediated phosphorylation. When PopZ is lost, CckA localization is hindered, and CdpR remains in its "active" dephosphorylated state. Consequently, overly active CdpR recruits more ClpXP to accelerate the proteolysis of CtrA. Interestingly, over-expression of PopZ also stimulates the proteolysis of CtrA but by a different mechanism. Under these abnormal conditions, CtrA and ClpXP are thought to concentrate at the cell pole and directly interact without using the CdpR adaptor (Joshi et al., 2018).

While CtrA inactivation is required for the initiation of chromosome replication in stalked cells (**Figure 1A**), its re-accumulation and phosphorylation in late S-phase are also required for cell-cycle transcription control and to prevent premature replication in the new swarmer cell compartment (Sanselicio et al., 2015). Accordingly, MopJ (motility PAS domain associated with DivJ) emerged as an important enhancing factor for CtrA accumulation (Sanselicio et al., 2015). At the cell poles, MopJ attenuates the DivJ-DivK-DivL kinase pathway that is also involved in the downregulation of CtrA through ClpXP. Once again, PopZ lies at the heart of this molecular interaction because the PopZ polar matrix localizes DivJ to the stalked pole, which in turn drives the polarization of DivK, DivL, and MopJ (Ebersbach et al., 2008).

During chromosome replication, the role of PopZ in partitioning switches from passive anchoring to an active participation in the movement of *parS-Cori*. Co-Immunoprecipitation experiments revealed that PopZ interacts directly with ParB, and a PopZ-ParB-*parS* complex presumably accounts for the initial polar anchoring/tethering at the early stalked pole (Bowman et al., 2008). Somehow the *parS-Cori* region is released from PopZ, and upon replication initiation, the duplicated DNA regions are separated such that one region seems to reattach, while the other moves slowly toward the quarter cell-length position. This corresponds to the slow phase of chromosome partitioning (**Figure 1**) that, as we described above, requires GapR but not ParA (Taylor et al., 2017). This is also the symmetry splitting point in the cell cycle that determines the subsequent fates of the chromosomes. Once this step is reached, the subsequent fast phase of partitioning uses ParA-ATPase activity. As the ParB-*parS* chromosome complex contacts DNA-bound ParA-ATP, the stimulated ATP hydrolysis causes subsequent ParA release. Such repeated interactions of binding and unbinding presumably cause the movement toward the new pole (Laloux and Jacobs-Wagner, 2013; Ptacin et al., 2014). Interestingly, the PopZ matrix directly sequesters the DNA-released ParA subunits at the new pole and then revives their ATP-bound state and their affinity for nucleoid DNA (Ptacin et al., 2014). This "recycling" or "rejuvenating" function of PopZ presumably enhances partitioning, since by concentrating and reactivating ParA-ATP dimers, PopZ will create a sharper ParA gradient that leads to the new cell pole. Interestingly, another cell pole "landmark" protein "TipN" shares functional redundancy with PopZ as it also recruits ParA to prevent reversal of the segregating ParB-*parS* complex (Ptacin et al., 2014). Accordingly, the ΔtipNΔpopZ double mutation is synthetically lethal (Schofield et al., 2010), and TipN polar localization is disrupted in the absence of PopZ (Ebersbach et al., 2008).

Further studies suggest an added layer of communication between ParA and PopZ. The redistribution of PopZ to the new swarmer pole (**Figure 1**) is coordinated with the arrival of the second ParB-*parS* focus at the new pole (Bowman et al., 2008; Ebersbach et al., 2008; Laloux and Jacobs-Wagner, 2013). Therefore, the ParA-dependent partitioning process somehow also drives the bi-polar organization of PopZ. In support of this notion, delayed partitioning caused by TipN depletion postponed PopZ accumulation at the new pole (Laloux and Jacobs-Wagner, 2013). ParA participates in the formation of the new PopZ matrix, as its loss disrupts PopZ bi-polarity. While other means of PopZ localization have been suggested such as self-organization by nucleoid occlusion (Ebersbach et al., 2008; Saberi and Emberly, 2010), these are clearly not enough, and a ParA-mediated PopZ-localization mechanism is required. If basal levels of ParA initiate PopZ recruitment, this may trigger a positive-feedback loop where ParA and PopZ will accumulate together through mutual support (Laloux and Jacobs-Wagner, 2013). As mentioned above, TipN also recruits ParA, and therefore, this polar landmark protein may also start or contribute to the growth of the PopZ matrix.

PopZ interactions are certainly complex yet robust, and however, this happens in wild-type *C. crescentus* cells, a new PopZ matrix always forms in time to meet and anchor the ParB-*parS* complex arriving at the new swarmer pole (**Figure 1B**). Interestingly, this cell-cycle pattern is very similar to that of the *V. cholerae* Chrom I, which is anchored through *parS1*-ParB1 to a polar PopZ-like protein called "HubP" (Yamaichi et al., 2012). Yet despite such a striking functional correspondence, HubP and PopZ are otherwise evolutionarily unrelated proteins.

The cell-cycle regulated zinc-finger protein ZitP offers yet another mechanism to control PopZ, independent of the *parABS* system (Berge et al., 2016). When ZitP is removed in a strain expressing a variant of PopZ that cannot bind ParB, bi-polar ParB fluorescent foci are rarely seen. However, the resupply of ZitP restores ParB foci at both cell poles, which implies the restoration of localized PopZ anchors (Berge et al., 2016). In this situation, the chromosome anchoring function may rely solely on ZitP since the PopZ variant is unable to bind ParB, but in wild-type cells, the role of ZitP in anchoring would be considered supportive. Normally, PopZ-bound ZitP indirectly binds to *parS*-flanking sites, where it functions to enhance ParB nucleation on the *parS* DNA. This assembly of ZitP-PopZ-ParB on the chromosome effectively restrains segregation (Berge et al., 2016).

It seems that the common theme for this multifaceted PopZ protein is its capacity for two-way interactions with many regulating and cell organizing proteins. For example, ZitP also relies on PopZ to recruit and position pilus biogenesis and swarming motility systems (Mignolet et al., 2016). In summary, PopZ is certainly a "hub" for cell-cycle communication that is yet to be fully explored as a mediator of crosstalk. Future studies promise new insights and new mechanisms of crosstalk between chromosome replication, partitioning, and probably the other landmarks of the cell cycle.

AUTHOR CONTRIBUTIONS

FUNDING

REFERENCES

Ah-Seng, Y., Rech, J., Lane, D., and Bouet, J. Y. (2013). Defining the role of ATP hydrolysis in mitotic segregation of bacterial plasmids. *PLoS Genet.* 9:e1003956. doi: 10.1371/journal.pgen.1003956

Ali Azam, T., Iwata, A., Nishimura, A., Ueda, S., and Ishihama, A. (1999). Growth phase-dependent variation in protein composition of the *Escherichia coli* nucleoid. *J. Bacteriol.* 181, 6361–6370.

Arias-Cartin, R., Dobihal, G. S., Campos, M., Surovtsev, I. V., Parry, B., and Jacobs-Wagner, C. (2017). Replication fork passage drives asymmetric dynamics of a critical nucleoid-associated protein in Caulobacter. *EMBO J.* 36, 301–318. doi: 10.15252/embj.201695513

Austin, S., and Abeles, A. (1983). Partition of unit-copy miniplasmids to daughter cells. I. P1 and F miniplasmids contain discrete, interchangeable sequences sufficient to promote equipartition. *J. Mol. Biol.* 169, 353–372. doi: 10.1016/S0022-2836(83)80055-2

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

Berge, M., Campagne, S., Mignolet, J., Holden, S., Theraulaz, L., Manley, S., et al. (2016). Modularity and determinants of a (bi-)polarization control system from free-living and obligate intracellular bacteria. *eLife* 5, pii: e20640–30. doi: 10.7554/eLife.20640

Bowman, G. R., Comolli, L. R., Gaietta, G. M., Fero, M., Hong, S. H., Jones, Y., et al. (2010). Caulobacter PopZ forms a polar subdomain dictating sequential changes in pole composition and function. *Mol. Microbiol.* 76, 173–189. doi: 10.1111/j.1365-2958.2010.07088.x

Bowman, G. R., Comolli, L. R., Zhu, J., Eckart, M., Koenig, M., Downing, K. H., et al. (2008). A polymeric protein anchors the chromosomal origin/ParB complex at a bacterial cell pole. *Cell* 134, 945–955. doi: 10.1016/j.cell.2008.07.015

Breier, A. M., and Grossman, A. D. (2007). Whole-genome analysis of the chromosome partitioning and sporulation protein Spo0J (ParB) reveals spreading and origin-distal sites on the *Bacillus subtilis* chromosome. *Mol. Microbiol.* 64, 703–718. doi: 10.1111/j.1365-2958.2007.05690.x

Brilli, M., Fondi, M., Fani, R., Mengoni, A., Ferri, L., Bazzicalupo, M., et al. (2010). The diversity and evolution of cell cycle regulation in alpha-proteobacteria: a comparative genomic analysis. *BMC Syst. Biol.* 4, 52. doi: 10.1186/1752-0509-4-52

Chen, B. W., Lin, M. H., Chu, C. H., Hsu, C. E., and Sun, Y. J. (2015). Insights into ParB spreading from the complex structure of Spo0J and parS. *Proc. Natl. Acad. Sci. USA* 112, 6613–6618. doi: 10.1073/pnas.1421927112

Chodavarapu, S., Felczak, M. M., Yaniv, J. R., and Kaguni, J. M. (2008). *Escherichia coli* DnaA interacts with HU in initiation at the *E. coli* replication origin. *Mol. Microbiol.* 67, 781–792. doi: 10.1111/j.1365-2958.2007.06094.x

Donczew, R., Zakrzewska-Czerwinska, J., and Zawilak-Pawlik, A. (2014). Beyond DnaA: the role of DNA topology and DNA methylation in bacterial replication initiation. *J. Mol. Biol.* 426, 2269–2282. doi: 10.1016/j.jmb.2014.04.009

Dorman, C. J., and Dorman, M. J. (2016). DNA supercoiling is a fundamental regulatory principle in the control of bacterial gene expression. *Biophys. Rev.* 8, 209–220. doi: 10.1007/s12551-016-0205-y

Ebersbach, G., Briegel, A., Jensen, G. J., and Jacobs-Wagner, C. (2008). A self-associating protein critical for chromosome attachment, division, and polar organization in Caulobacter. *Cell* 134, 956–968. doi: 10.1016/j.cell.2008.07.016

Egan, E. S., and Waldor, M. K. (2003). Distinct replication requirements for the two *Vibrio cholerae* chromosomes. *Cell* 114, 521–530. doi: 10.1016/S0092-8674(03)00611-1

Erzberger, J. P., and Berger, J. M. (2006). Evolutionary relationships and structural mechanisms of AAA+ proteins. *Annu. Rev. Biophys. Biomol. Struct.* 35, 93–114. doi: 10.1146/annurev.biophys.35.040405.101933

Espinosa, E., Barre, F. X., and Galli, E. (2017). Coordination between replication, segregation and cell division in multi-chromosomal bacteria: lessons from *Vibrio cholerae. Int. Microbiol.* 20, 121–129. doi: 10.2436/20.1501.01.293

Funnell, B. E. (2016). ParB partition proteins: complex formation and spreading at bacterial and plasmid centromeres. *Front. Mol. Biosci.* 3:44. doi: 10.3389/fmolb.2016.00044

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

Gerdes, K., Larsen, J. E., and Molin, S. (1985). Stable inheritance of plasmid R1 requires two different loci. *J. Bacteriol.* 161, 292–298.

Gerding, M. A., Chao, M. C., Davis, B. M., and Waldor, M. K. (2015). Molecular dissection of the essential features of the origin of replication of the second *Vibrio cholerae* chromosome. *MBio* 6:e00973. doi: 10.1128/mBio.00973-15

Gorbatyuk, B., and Marczynski, G. T. (2001). Physiological consequences of blocked *Caulobacter crescentus* dnaA expression, an essential DNA replication gene. *Mol. Microbiol.* 40, 485–497. doi: 10.1046/j.1365-2958.2001.02404.x

Graham, T. G., Wang, X., Song, D., Etson, C. M., van Oijen, A. M., Rudner, D. Z., et al. (2014). ParB spreading requires DNA bridging. *Genes Dev.* 28, 1228–1238. doi: 10.1101/gad.242206.114

Grimwade, J. E., Rozgaja, T. A., Gupta, R., Dyson, K., Rao, P., and Leonard, A. C. (2018). Origin recognition is the predominant role for DnaA-ATP in initiation of chromosome replication. *Nucleic Acids Res.* 46, 6140–6151. doi: 10.1093/nar/gky457

Guo, M. S., Haakonsen, D. L., Zeng, W., Schumacher, M. A., and Laub, M. T. (2018). A bacterial chromosome structuring protein binds overtwisted DNA to stimulate type II topoisomerases and enable DNA replication. *Cell.* 175, 583–597.e23. doi: 10.1016/j.cell.2018.08.029

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

Helmstetter, C., Cooper, S., Pierucci, O., and Revelas, E. (1968). On the bacterial life sequence. *Cold Spring Harb. Symp. Quant. Biol.* 33, 809–822. doi: 10.1101/SQB.1968.033.01.093

Holmes, J. A., Follett, S. E., Wang, H., Meadows, C. P., Varga, K., and Bowman, G. R. (2016). Caulobacter PopZ forms an intrinsically disordered hub in organizing bacterial cell poles. *Proc. Natl. Acad. Sci. USA* 113, 12490–12495. doi: 10.1073/pnas.1602380113

Holowka, J., Trojanowski, D., Ginda, K., Wojtas, B., Gielniewski, B., Jakimowicz, D., et al. (2017). HupB is a bacterial nucleoid-associated protein with an indispensable eukaryotic-like tail. *MBio* 8, pii: e01272–17. doi: 10.1128/mBio.01272-17

Hottes, A. K., Shapiro, L., and McAdams, H. H. (2005). DnaA coordinates replication initiation and cell cycle transcription in *Caulobacter crescentus. Mol. Microbiol.* 58, 1340–1353. doi: 10.1111/j.1365-2958.2005.04912.x

Ietswaart, R., Szardenings, F., Gerdes, K., and Howard, M. (2014). Competing ParA structures space bacterial plasmids equally over the nucleoid. *PLoS Comput. Biol.* 10:e1004009. doi: 10.1371/journal.pcbi.1004009

Iniesta, A. A. (2014). ParABS system in chromosome partitioning in the bacterium *Myxococcus xanthus. PLoS One* 9:e86897. doi: 10.1371/journal.pone.0086897

Ireton, K., Gunther, N. W. t., and Grossman, A. D. (1994). spo0J is required for normal chromosome segregation as well as the initiation of sporulation in *Bacillus subtilis. J. Bacteriol.* 176, 5320–5329. doi: 10.1128/jb.176.17.5320-5329.1994

Joshi, K. K., Battle, C. M., and Chien, P. (2018). Polar localization hub protein PopZ restrains adaptor-dependent ClpXP proteolysis in *Caulobacter crescentus. J. Bacteriol.* 200, pii: e00221–18. doi: 10.1128/JB.00221-18

Kadoya, R., Baek, J. H., Sarker, A., and Chattoraj, D. K. (2011). Participation of chromosome segregation protein ParAI of *Vibrio cholerae* in chromosome replication. *J. Bacteriol.* 193, 1504–1514. doi: 10.1128/JB.01067-10

Kaguni, J. M. (2011). Replication initiation at the *Escherichia coli* chromosomal origin. *Curr. Opin. Chem. Biol.* 15, 606–613. doi: 10.1016/j.cbpa.2011.07.016

Katayama, T. (2017). Initiation of DNA replication at the chromosomal origin of *E. coli*, oriC. *Adv. Exp. Med. Biol.* 1042, 79–98. doi: 10.1007/978-981-10-6955-0_4

Katayama, T., Ozaki, S., Keyamura, K., and Fujimitsu, K. (2010). Regulation of the replication cycle: conserved and diverse regulatory systems for DnaA and oriC. *Nat. Rev. Microbiol.* 8, 163–170. doi: 10.1038/nrmicro2314

Kaur, G., Vora, M. P., Czerwonka, C. A., Rozgaja, T. A., Grimwade, J. E., and Leonard, A. C. (2014). Building the bacterial orisome: high-affinity DnaA recognition plays a role in setting the conformation of oriC DNA. *Mol. Microbiol.* 91, 1148–1163. doi: 10.1111/mmi.12525

Kawakami, H., Keyamura, K., and Katayama, T. (2005). Formation of an ATP-DnaA-specific initiation complex requires DnaA Arginine 285, a conserved motif in the AAA+ protein family. *J. Biol. Chem.* 280, 27420–27430. doi: 10.1074/jbc.M502764200

Kornberg, A., and Baker, T. A. (1992). *DNA replication. 2nd Edn.* New York: W.H. Freeman.

Krogh, T. J., Moller-Jensen, J., and Kaleta, C. (2018). Impact of chromosomal architecture on the function and evolution of bacterial genomes. *Front. Microbiol.* 9:2019. doi: 10.3389/fmicb.2018.02019

Laloux, G., and Jacobs-Wagner, C. (2013). Spatiotemporal control of PopZ localization through cell cycle-coupled multimerization. *J. Cell Biol.* 201, 827–841. doi: 10.1083/jcb.201303036

Lee, P. S., Lin, D. C., Moriya, S., and Grossman, A. D. (2003). Effects of the chromosome partitioning protein Spo0J (ParB) on oriC positioning and replication initiation in *Bacillus subtilis*. *J. Bacteriol.* 185, 1326–1337. doi: 10.1128/JB.185.4.1326-1337.2003

Leonard, A. C., and Grimwade, J. E. (2011). Regulation of DnaA assembly and activity: taking directions from the genome. *Annu. Rev. Microbiol.* 65, 19–35. doi: 10.1146/annurev-micro-090110-102934

Livny, J., Yamaichi, Y., and Waldor, M. K. (2007). Distribution of centromere-like parS sites in bacteria: insights from comparative genomics. *J. Bacteriol.* 189, 8693–8703. doi: 10.1128/JB.01239-07

Luijsterburg, M. S., Noom, M. C., Wuite, G. J., and Dame, R. T. (2006). The architectural role of nucleoid-associated proteins in the organization of bacterial chromatin: a molecular perspective. *J. Struct. Biol.* 156, 262–272. doi: 10.1016/j.jsb.2006.05.006

Lutkenhaus, J. (2007). Assembly dynamics of the bacterial MinCDE system and spatial regulation of the Z ring. *Annu. Rev. Biochem.* 76, 539–562. doi: 10.1146/annurev.biochem.75.103004.142652

Marczynski, G. T., Rolain, T., and Taylor, J. A. (2015). Redefining bacterial origins of replication as centralized information processors. *Front. Microbiol.* 6:610. doi: 10.3389/fmicb.2015.00610

McGarry, K. C., Ryan, V. T., Grimwade, J. E., and Leonard, A. C. (2004). Two discriminatory binding sites in the *Escherichia coli* replication origin are required for DNA strand opening by initiator DnaA-ATP. *Proc. Natl. Acad. Sci. USA* 101, 2811–2816. doi: 10.1073/pnas.0400340101

Mera, P. E., Kalogeraki, V. S., and Shapiro, L. (2014). Replication initiator DnaA binds at the Caulobacter centromere and enables chromosome segregation. *Proc. Natl. Acad. Sci. USA* 111, 16100–16105. doi: 10.1073/pnas.1418989111

Mignolet, J., Holden, S., Berge, M., Panis, G., Eroglu, E., Theraulaz, L., et al. (2016). Functional dichotomy and distinct nanoscale assemblies of a cell cycle-controlled bipolar zinc-finger regulator. *eLife* 5. doi: 10.7554/eLife.18647

Mohl, D. A., and Gober, J. W. (1997). Cell cycle-dependent polar localization of chromosome partitioning proteins in *Caulobacter crescentus*. *Cell* 88, 675–684. doi: 10.1016/S0092-8674(00)81910-8

Mott, M. L., and Berger, J. M. (2007). DNA replication initiation: mechanisms and regulation in bacteria. *Nat. Rev. Microbiol.* 5, 343–354. doi: 10.1038/nrmicro1640

Murray, H., and Errington, J. (2008). Dynamic control of the DNA replication initiation protein DnaA by Soj/ParA. *Cell* 135, 74–84. doi: 10.1016/j.cell.2008.07.044

Ogura, T., and Hiraga, S. (1983). Partition mechanism of F plasmid: two plasmid gene-encoded products and a cis-acting region are involved in partition. *Cell* 32, 351–360. doi: 10.1016/0092-8674(83)90454-3

Ptacin, J. L., Gahlmann, A., Bowman, G. R., Perez, A. M., von Diezmann, A. R., Eckart, M. R., et al. (2014). Bacterial scaffold directs pole-specific centromere segregation. *Proc. Natl. Acad. Sci. USA* 111, E2046–E2055. doi: 10.1073/pnas.1405188111

Quisel, J. D., and Grossman, A. D. (2000). Control of sporulation gene expression in *Bacillus subtilis* by the chromosome partitioning proteins Soj (ParA) and Spo0J (ParB). *J. Bacteriol.* 182, 3446–3451. doi: 10.1128/JB.182.12.3446-3451.2000

Rodionov, O., Lobocka, M., and Yarmolinsky, M. (1999). Silencing of genes flanking the P1 plasmid centromere. *Science* 283, 546–549. doi: 10.1126/science.283.5401.546

Rozgaja, T. A., Grimwade, J. E., Iqbal, M., Czerwonka, C., Vora, M., and Leonard, A. C. (2011). Two oppositely oriented arrays of low-affinity recognition sites in oriC guide progressive binding of DnaA during *Escherichia coli* pre-RC assembly. *Mol. Microbiol.* 82, 475–488. doi: 10.1111/j.1365-2958.2011.07827.x

Ruvolo, M. V., Mach, K. E., and Burkholder, W. F. (2006). Proteolysis of the replication checkpoint protein Sda is necessary for the efficient initiation of sporulation after transient replication stress in *Bacillus subtilis*. *Mol. Microbiol.* 60, 1490–1508. doi: 10.1111/j.1365-2958.2006.05167.x

Ryan, V. T., Grimwade, J. E., Camara, J. E., Crooke, E., and Leonard, A. C. (2004). *Escherichia coli* prereplication complex assembly is regulated by dynamic interplay among Fis, IHF and DnaA. *Mol. Microbiol.* 51, 1347–1359. doi: 10.1046/j.1365-2958.2003.03906.x

Saberi, S., and Emberly, E. (2010). Chromosome driven spatial patterning of proteins in bacteria. *PLoS Comput. Biol.* 6:e1000986. doi: 10.1371/journal.pcbi.1000986

Sanselicio, S., Berge, M., Theraulaz, L., Radhakrishnan, S. K., and Viollier, P. H. (2015). Topological control of the Caulobacter cell cycle circuitry by a polarized single-domain PAS protein. *Nat. Commun.* 6, 7005. doi: 10.1038/ncomms8005

Schenk, K., Hervas, A. B., Rosch, T. C., Eisemann, M., Schmitt, B. A., Dahlke, S., et al. (2017). Rapid turnover of DnaA at replication origin regions contributes to initiation control of DNA replication. *PLoS Genet.* 13:e1006561. doi: 10.1371/journal.pgen.1006561

Schofield, W. B., Lim, H. C., and Jacobs-Wagner, C. (2010). Cell cycle coordination and regulation of bacterial chromosome segregation dynamics by polarly localized proteins. *EMBO J.* 29, 3068–3081. doi: 10.1038/emboj.2010.207

Scholefield, G., Errington, J., and Murray, H. (2012). Soj/ParA stalls DNA replication by inhibiting helix formation of the initiator protein DnaA. *EMBO J.* 31, 1542–1555. doi: 10.1038/emboj.2012.6

Scholefield, G., Whiting, R., Errington, J., and Murray, H. (2011). Spo0J regulates the oligomeric state of Soj to trigger its switch from an activator to an inhibitor of DNA replication initiation. *Mol. Microbiol.* 79, 1089–1100. doi: 10.1111/j.1365-2958.2010.07507.x

Seid, C. A., Smith, J. L., and Grossman, A. D. (2017). Genetic and biochemical interactions between the bacterial replication initiator DnaA and the nucleoid-associated protein Rok in *Bacillus subtilis*. *Mol. Microbiol.* 103, 798–817. doi: 10.1111/mmi.13590

Shebelut, C. W., Guberman, J. M., van Teeffelen, S., Yakhnina, A. A., and Gitai, Z. (2010). Caulobacter chromosome segregation is an ordered multistep process. *Proc. Natl. Acad. Sci. USA* 107, 14194–14198. doi: 10.1073/pnas.1005274107

Siam, R., and Marczynski, G. T. (2000). Cell cycle regulator phosphorylation stimulates two distinct modes of binding at a chromosome replication origin. *EMBO J.* 19, 1138–1147. doi: 10.1093/emboj/19.5.1138

Skarstad, K., and Katayama, T. (2013). Regulating DNA replication in bacteria. *Cold Spring Harb. Perspect. Biol.* 5, a012922. doi: 10.1101/cshperspect.a012922

Sobetzko, P., Travers, A., and Muskhelishvili, G. (2012). Gene order and chromosome dynamics coordinate spatiotemporal gene expression during the bacterial growth cycle. *Proc. Natl. Acad. Sci. USA* 109, E42–E50. doi: 10.1073/pnas.1108229109

Stavans, J., and Oppenheim, A. (2006). DNA-protein interactions and bacterial chromosome architecture. *Phys. Biol.* 3, R1–10. doi: 10.1088/1478-3975/3/4/R01

Taylor, J. A., Panis, G., Viollier, P. H., and Marczynski, G. T. (2017). A novel nucleoid-associated protein coordinates chromosome replication and chromosome partition. *Nucleic Acids Res.* 45, 8916–8929. doi: 10.1093/nar/gkx596

Taylor, J. A., Pastrana, C. L., Butterer, A., Pernstich, C., Gwynn, E. J., Sobott, F., et al. (2015). Specific and non-specific interactions of ParB with DNA:

implications for chromosome segregation. *Nucleic Acids Res.* 43, 719–731. doi: 10.1093/nar/gku1295

Toro, E., and Shapiro, L. (2010). Bacterial chromosome organization and segregation. *Cold Spring Harb. Perspect. Biol.* 2, a000349. doi: 10.1101/cshperspect.a000349

Tran, N. T., Stevenson, C. E., Som, N. F., Thanapipatsiri, A., Jalal, A. S. B., and Le, T. B. K. (2018). Permissive zones for the centromere-binding protein ParB on the *Caulobacter crescentus* chromosome. *Nucleic Acids Res.* 46, 1196–1209. doi: 10.1093/nar/gkx1192

Vecchiarelli, A. G., Han, Y. W., Tan, X., Mizuuchi, M., Ghirlando, R., Biertumpfel, 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., Mizuuchi, K., and Funnell, B. E. (2012). Surfing biological surfaces: exploiting the nucleoid for partition and transport in bacteria. *Mol. Microbiol.* 86, 513–523. doi: 10.1111/mmi.12017

Venkova-Canova, T., Baek, J. H., Fitzgerald, P. C., Blokesch, M., and Chattoraj, D. K. (2013). Evidence for two different regulatory mechanisms linking replication and segregation of vibrio cholerae chromosome II. *PLoS Genet.* 9:e1003579. doi: 10.1371/journal.pgen.1003579

Wolanski, M., Donczew, R., Zawilak-Pawlik, A., and Zakrzewska-Czerwinska, J. (2014a). oriC-encoded instructions for the initiation of bacterial chromosome replication. *Front. Microbiol.* 5, 735. doi: 10.3389/fmicb.2014.00735

Wolanski, M., Jakimowicz, D., and Zakrzewska-Czerwinska, J. (2014b). Fifty years after the replicon hypothesis: cell-specific master regulators as new players in chromosome replication control. *J. Bacteriol.* 196, 2901–2911. doi: 10.1128/JB.01706-14

Yamaichi, Y., Bruckner, R., Ringgaard, S., Moll, A., Cameron, D. E., Briegel, A., et al. (2012). A multidomain hub anchors the chromosome segregation and chemotactic machinery to the bacterial pole. *Genes Dev.* 26, 2348–2360. doi: 10.1101/gad.199869.112

Yamaichi, Y., Gerding, M. A., Davis, B. M., and Waldor, M. K. (2011). Regulatory cross-talk links *Vibrio cholerae* chromosome II replication and segregation. *PLoS Genet.* 7:e1002189. doi: 10.1371/journal.pgen.1002189

Where and When Bacterial Chromosome Replication Starts: A Single Cell Perspective

Damian Trojanowski†, Joanna Hołówka† and Jolanta Zakrzewska-Czerwińska*†

Department of Molecular Microbiology, Faculty of Biotechnology, University of Wrocław, Wrocław, Poland

Bacterial chromosomes have a single, unique replication origin (named *oriC*), from which DNA synthesis starts. This study describes methods of visualizing *oriC* regions and the chromosome replication in single living bacterial cells in real-time. This review also discusses the impact of live cell imaging techniques on understanding of chromosome replication dynamics, particularly at the initiation step, in different species of bacteria.

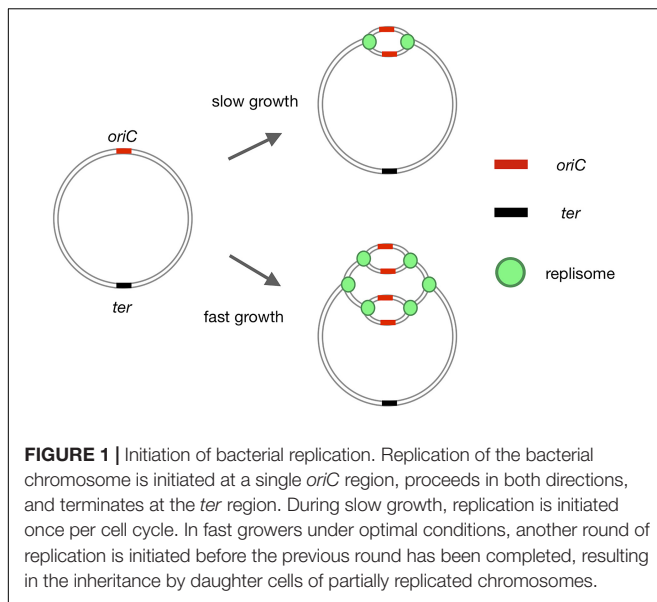Keywords: replication initiation, *oriC*, replisome, single-cell, bacterial chromosome

INTRODUCTION

DNA replication is an enormously intricate process, in which a few dozen enzymes catalyze a series of reactions, including DNA unwinding and the synthesis of sister DNA strands. This process must be highly precise and accurately timed to prevent any unnecessary loss of energy and to ensure that DNA is faithfully and completely replicated only once per cell-division cycle (Leonard and Grimwade, 2015). In all three domains of life, chromosomal replication is mainly regulated at the initiation step (Nielsen and Løbner-Olesen, 2008; Aves, 2009; Skarstad and Katayama, 2013), an important cell cycle checkpoint guaranteeing that DNA replication begins at the right place and time.

Most bacterial genomes consist of one covalently closed chromosome (**Figure 1**). In a few bacteria, however, the genetic information is distributed on two [e.g., *Vibrio cholerae* (Trucksis et al., 1998)] or even more [e.g., *Paracoccus denitrificans* (Winterstein and Ludwig, 1998)] chromosomes. Interestingly, some bacteria possess linear chromosomes [e.g., *Streptomyces* (Lin et al., 1993)].

In contrast to eukaryotes, bacterial chromosomes have a single, unique origin of replication (*oriC*) (Bird et al., 1972; Kaguni and Kornberg, 1984; Gao and Zhang, 2008; Masai et al., 2010; Méchali, 2010; Katayama, 2017). DNA synthesis is initiated at this unique *oriC*, generating a single replication eye per chromosome (**Figure 1**). Cooperative binding of the initiator protein, DnaA, to multiple DnaA-recognition sites (DnaA boxes) within the *oriC* region triggers separation of the DNA strands at the DNA unwinding element (DUE), providing an entry site for the machinery of replication (replisome, **Figures 1**, **2A**; Skarstad et al., 1986, 1990; Bach et al., 2008; Leonard and Grimwade, 2011; Wolański et al., 2014; Richardson et al., 2016).

Enormous progress has been made in recent years toward understanding the mechanisms of replication initiation, particularly the organization and function of *oriC* regions in different bacteria (Donczew et al., 2012; Makowski et al., 2016; Jaworski et al., 2018; Midgley-Smith et al., 2018; Samadpour and Merrikh, 2018). Less is known, however, about the subcellular localization of replication processes during the cell cycle in various bacterial species. The development of sophisticated cell biology techniques has allowed examination of when and where the replication machinery is assembled within the bacterial cells, and how the initiation of replication is coordinated with the cell cycle (Donczew et al., 2012; Harms et al., 2013; Santi and McKinney, 2015;

FIGURE 1 | Initiation of bacterial replication. Replication of the bacterial chromosome is initiated at a single *oriC* region, proceeds in both directions, and terminates at the *ter* region. During slow growth, replication is initiated once per cell cycle. In fast growers under optimal conditions, another round of replication is initiated before the previous round has been completed, resulting in the inheritance by daughter cells of partially replicated chromosomes.

Trojanowski et al., 2015; Böhm et al., 2017). This process is particularly interesting in bacteria with two chromosomes (*V. cholerae*) (Demarre et al., 2014; Ramachandran et al., 2018) and in those that undergo complex cell differentiation (*Caulobacter crescentus*) (Jensen et al., 2001; Toro et al., 2008) and/or exhibit complicated life cycles, e.g., *Myxococcus xanthus* (Harms et al., 2013; Lin et al., 2017) and *Streptomyces* species (Kois-Ostrowska et al., 2016). In these bacteria, the regulatory networks that control replication initiation are likely to be intricate and require specific mechanisms that can synchronize the initiation of chromosomal replication with developmental processes.
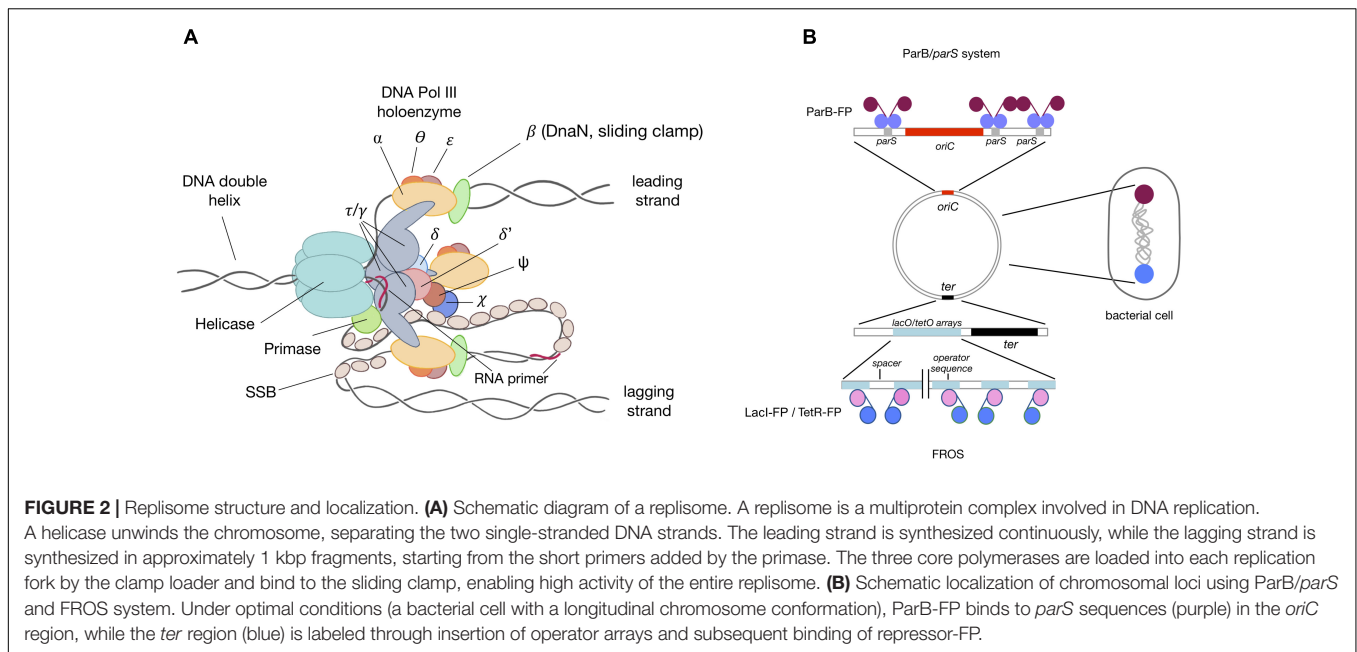
The main goal of this review is to highlight imaging techniques that allow the determination of the subcellular location of *oriC* regions and the initiation of chromosome replication (i.e., assembly of the replication machinery) in single living bacterial cells in real time. This review also discusses the impact of real-time single-cell imaging on understanding of chromosome replication dynamics, particularly at the initiation step, in different bacteria.

VISUALIZATION OF REPLICATION INITIATION AND REPLISOME DYNAMICS IN LIVE CELLS

The development of live cell imaging techniques has allowed the visualization of replisomes (**Figure 2A**; Jensen et al., 2001; Reyes-Lamothe et al., 2008; Wang and Sherratt, 2010; Harms et al., 2013; Santi and McKinney, 2015; Trojanowski et al., 2015; Mangiameli et al., 2017) in live cells and the study of DNA replication dynamics, including the timing and localization of replication initiation, in real time at the single-cell level. Microscopic analysis of live cells has several advantages over analysis of fixed samples. Fixing the cells, a process that involves dehydration

and/or intracellular cross-linking, may influence the localization of proteins or subcellular structures of interest. Moreover, some fusions with fluorescent proteins (FP) are sensitive to the harsh conditions used during fixation. For example, different sample preparation of *Mycobacterium smegmatis* cells results in ParA-EGFP localizing either apically or as a cloud arising from the new cell pole (Ginda et al., 2013, 2017). Furthermore, permeabilization of the bacterial cell wall during immunostaining may contribute to a loss of cytoplasmic content or, due to cellular crowding, may generate high background noise or alter the localization of large immunocomplexes, particularly when using secondary antibodies for signal amplification. Although several high quality studies of fixed samples have provided invaluable data, the conditions found in cells fixed on a coverslip only approximate the conditions found in live cells.

Replication is visualized primarily by the fusion of different replisome (DNA polymerase III) subunits (**Figure 2A**) to a variety of FP. The choice of subunit to create the fusion protein should be guided by the specific application and the specific type of bacterium. *Escherichia coli* is the best characterized bacterial model for tracking live replication (Kongsuwan et al., 2002; Bates and Kleckner, 2005; Fossum et al., 2007; Reyes-Lamothe et al., 2008, 2010; Su'etsugu and Errington, 2011; Wang et al., 2011; Moolman et al., 2014; Beattie et al., 2017). However, several reports have tracked replication in other organisms, including *Bacillus subtilis* (Lemon and Grossman, 1998; Migocki et al., 2004; Berkmen and Grossman, 2006; Mangiameli et al., 2017; Li et al., 2018), *C. crescentus* (Jensen et al., 2001; Fernandez-Fernandez et al., 2013; Arias-Cartin et al., 2017), *V. cholerae* (Srivastava and Chattoraj, 2007; Stokke et al., 2011), *M. smegmatis* (Santi et al., 2013; Santi and McKinney, 2015; Trojanowski et al., 2015, 2017), *Streptomyces coelicolor* (Ruban-Ośmiałowska et al., 2006; Wolański et al., 2011), *Corynebacterium glutamicum* (Böhm et al., 2017), *Pseudomonas aeruginosa* (Vallet-Gely and Boccard, 2013), *M. xanthus* (Harms et al., 2013), and *Streptococcus pneumoniae* (Raaphorst et al., 2017). Findings of these studies may help in the construction of fluorescent fusions of replisome components in other bacteria. It is also important to consider alternative N- and C-terminal fusion, as one, or sometimes both, ends of target proteins may be implicated in inter- or intra-molecular interactions. The sliding clamp (**Figure 2A**) is the protein of choice in most studies and both N- and C-terminal fusions proved to be functional in a range of species (Kongsuwan et al., 2002; Reyes-Lamothe et al., 2010; Su'etsugu and Errington, 2011; Moolman et al., 2014; Santi and McKinney, 2015; Trojanowski et al., 2015; Arias-Cartin et al., 2017; Böhm et al., 2017; Mangiameli et al., 2017; Hołówka et al., 2018). However, the sliding clamp also participates in processes other than DNA replication, including recombination and DNA repair, possibly altering the distribution of DnaN-FP (or FP-DnaN) foci in these cells. This is not usually a concern in wild-type-like fluorescent reporter strains, under both optimal and minimal conditions, but may be of concern in knock-out/overproducing mutant strains, involving, for example, genes engaged in DNA repair, or when studying replication dynamics under stress-inducing conditions such as in the presence of antibiotics, mutagenic compounds like

FIGURE 2 | Replisome structure and localization. **(A)** Schematic diagram of a replisome. A replisome is a multiprotein complex involved in DNA replication. A helicase unwinds the chromosome, separating the two single-stranded DNA strands. The leading strand is synthesized continuously, while the lagging strand is synthesized in approximately 1 kbp fragments, starting from the short primers added by the primase. The three core polymerases are loaded into each replication fork by the clamp loader and bind to the sliding clamp, enabling high activity of the entire replisome. **(B)** Schematic localization of chromosomal loci using ParB/*parS* and FROS system. Under optimal conditions (a bacterial cell with a longitudinal chromosome conformation), ParB-FP binds to *parS* sequences (purple) in the *oriC* region, while the *ter* region (blue) is labeled through insertion of operator arrays and subsequent binding of repressor-FP.

mitomycin, and replication inhibitors. In these experiments, choosing another replisome component may be advisable. Beside the siding clamp, DnaX (Lemon and Grossman, 2000; Bates and Kleckner, 2005; Berkmen and Grossman, 2006; Vallet-Gely and Boccard, 2013; Raaphorst et al., 2017) (particularly its C-terminal fusion) is frequently used as a replisome localization marker. The *dnaX* gene encodes two alternative proteins, τ – the full-length protein encoded by the *dnaX* gene, and γ, which originates from ribosome switching during translation, resulting in premature termination of translation and generating a truncated protein. Single-stranded DNA binding protein (SSB) (**Figure 2A**) has also been tested in several studies (Reyes-Lamothe et al., 2008, 2010; Harms et al., 2013; Sukumar et al., 2014; Santi and McKinney, 2015; Mangiameli et al., 2017; Raaphorst et al., 2017). Monitoring replisome dynamics in strains expressing fusion proteins encoded on an episomal plasmid is not recommended, as plasmid replication is triggered mainly by the same protein components that trigger chromosomal replication. Fusion with catalytic core subunits (Lemon and Grossman, 1998; Migocki et al., 2004; Trojanowski et al., 2017) is also possible, although additional cargo attached to core Pol-DNA III may affect nucleotide incorporation rates and influence the kinetic parameters of the entire replication complex. This was shown for *M. smegmatis*, where the C-terminal fusion of a catalytic alpha subunit to EYFP prolonged the C-period (Trojanowski et al., 2017). Thus proteins other than the catalytic core complex may be a better choice for studies of replisome dynamics. Other fusions successfully used for replisome tracking include DnaB (DNA helicase) (Jensen et al., 2001; Beattie et al., 2017), DnaQ (Reyes-Lamothe et al., 2008, 2010; Wallden et al., 2016; Mangiameli et al., 2017), and χ and δ' subunits (Jensen et al., 2001; Reyes-Lamothe et al., 2008). When designing a fluorescent fusion for replisome visualization, additional features should be taken into account, especially oligomerization status, fluorescence yield and

spectral properties. FP (especially GFP derivatives) are likely to form low-affinity oligomers (Costantini et al., 2012), which may influence the dynamics of the studied protein complex, especially when the fusion protein is produced at a high level. Thus, choosing a fluorescent variant with a lower tendency to undergo oligomerization (e.g., mCherry, mCherry2, mCitrine, and mScarlett) is recommended. Spectral characteristics and brightness are essential, especially when replisomes are localized together with other cellular components (e.g., chromosome and membrane) (Shaner et al., 2005). Importantly, FP are sensitive to pH and cannot be utilized to analyze anaerobic bacteria, as maturation of the chromophore requires oxygen molecules (Shaner et al., 2005; Landete et al., 2015). Fluorescent fusion proteins are suitable for both qualitative long-term live cell imaging and quantitative analysis. For example, Y-Pet fusion with a variety of replisome subunits was used to quantify the numbers of copies of particular proteins within a replication eye *in vivo* (Reyes-Lamothe et al., 2010). However, most of these variants lacked the properties required for super-resolution imaging. In the latter case, proteins of interest should be fused with photoactivated or photoconvertible proteins. Recently published studies may provide hints regarding single-molecule resolution microscopy of replication complexes (Georgescu et al., 2012; Stracy et al., 2014; Liao et al., 2016; Lewis et al., 2017). The fusion of replisome subunits with HaloTag may be an alternative to FP. The size of HaloTag is similar to that of FP, but the ligands that bind to HaloTag have better fluorescence yield, resulting in a higher signal compared with standard FPs (HaloTag® Protein Purification System, 2018). The advantage of using direct fluorescent ligands (e.g., dTMR and dR110) is that they do not need to be washed out before acquisition. Halo ligands are also suitable for high-resolution microscopy.

Replication tracking (particularly initiation of replication) is often accompanied by localization of nascent *oriCs* (**Figure 2B**).

The fluorescence repressor operator system (FROS) or ParB/*parS* is frequently used for live cell tracking (Lau et al., 2003). The FROS system (**Figure 2B**) consists of two components: operator sequences (usually *lacO* or *tetO* arrays repeated up to several hundred times in tandem and interspersed by oligonucleotide spacers) and an FP-tagged repressor protein (LacI-FP or TetR-FP), which binds to the operator sequences. FROS was efficiently used to localize chromosomal loci, including *oriC*, terminus and other specific loci on both replichores in a variety of species (Viollier et al., 2004; Fogel and Waldor, 2005; Frunzke et al., 2008; Liu et al., 2010; Vallet-Gely and Boccard, 2013; Wang et al., 2014; Santi and McKinney, 2015). However, it is often difficult to insert the large operator arrays into the chromosome, particularly in highly transcribed regions such as *oriC* (Le and Laub, 2014). Moreover, overexpression of repressor may result in replication/transcription hold-up or alteration in segregation of replicated regions (Possoz et al., 2006; Mettrick and Grainge, 2016). Thus, low levels of repressor should be produced, usually by using inducible promoters. Additionally, tracking *oriCs* together with replisomes requires delivery of the repressor-FP fusion protein from the chromosomal locus, either as a part of an operator array construct or inserted into an attachment site. Although FROS may provide invaluable data, its instability is a major drawback.
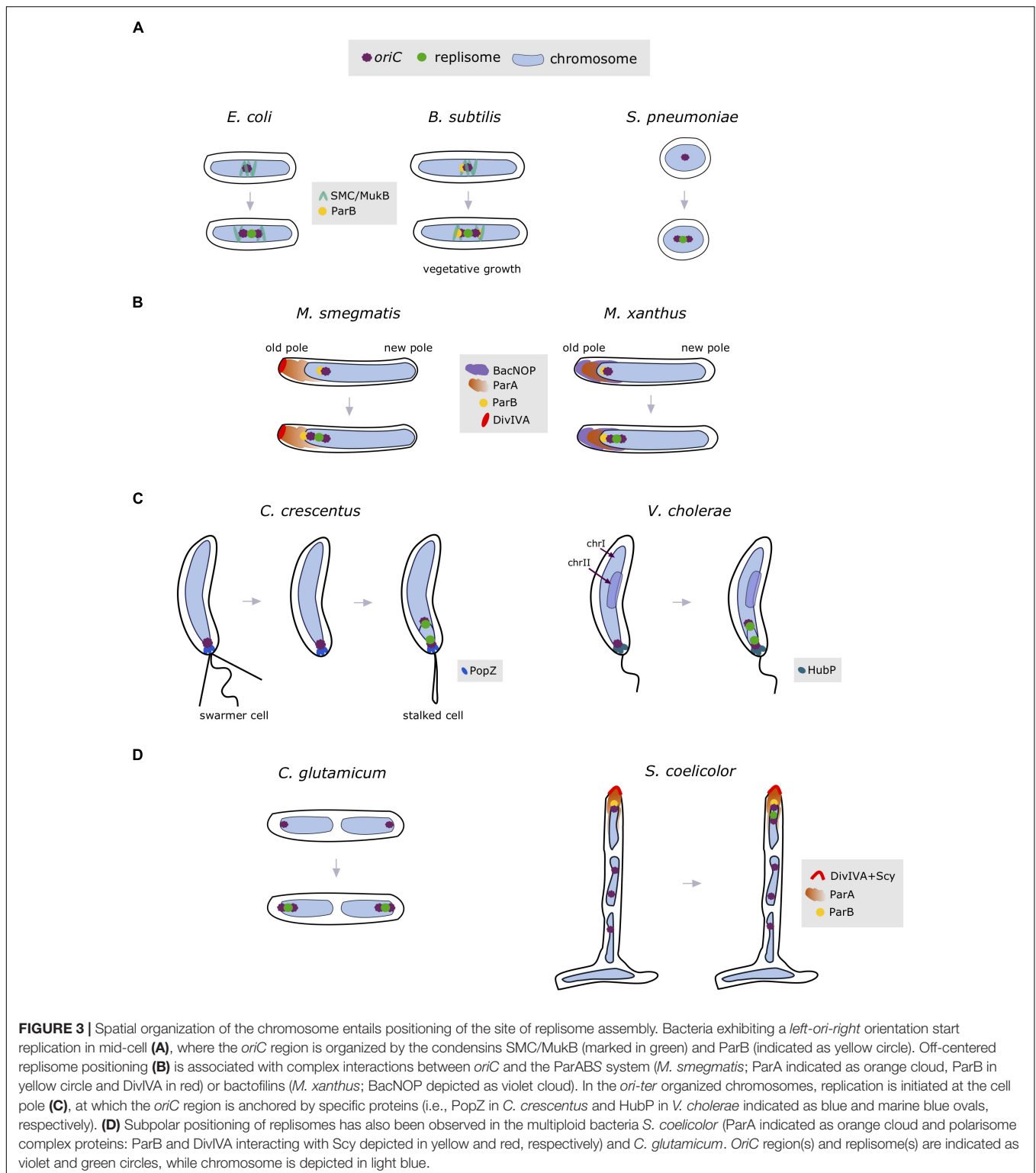
The ParB-FP/*parS* system (which originated from naturally existing chromosome and/or plasmid partitioning strategies) (**Figure 2B**) represents an easier alternative to FROS. This system uses an intrinsic feature of ParB, its binding to centromere-like *parS* sequences (Wang et al., 2011; Reyes-Lamothe et al., 2012; Badrinarayanan et al., 2015). Most bacterial species possess the ParABS chromosome segregation system, except for several well-studied Gammaproteobacteria, including *E. coli*. Because most chromosomal *parS* sites are localized proximal to the *oriC*-proximal regions (Livny et al., 2007), introduction of fluorescent ParB, which oligomerizes within *parS* sequences, addresses all of the system requirements for successful *oriC* labeling. This approach has been shown effective in a number of bacteria, including *Mycobacterium*, *M. xanthus* (Harms et al., 2013), *Streptomyces* (Donczew et al., 2016; Kois-Ostrowska et al., 2016), *C. crescentus* (Laloux and Jacobs-Wagner, 2013), and *C. glutamicum* (Donovan et al., 2010; Böhm et al., 2017). In bacteria lacking a chromosomal ParABS system (e.g., *E. coli*), plasmid-derived partitioning components (phage P1 or *Yersinia pestis* MT1ParB/*parS* systems) are frequently used (Youngren et al., 2000; Li et al., 2002; Nielsen et al., 2006, 2007). The use of plasmid-derived *parS*/ParB is also beneficial, as it does not interfere with the endogenous chromosomal ParABS system or another plasmid-derived *parS*/ParB system (P1/MT1), allowing the simultaneous localization of multiple chromosomal loci. Its major advantage compared with FROS is that insertion of only a few copies of *parS* is sufficient for strong fluorescent signals after ParB-FP binding.

Determination of the specific point (and subcellular localization) at which replication is initiated requires long-term imaging of living cells (from several minutes to hours, depending on the bacterial growth rate and the conditions being tested, e.g., rich versus minimal medium). The simplest way to analyze replication at the single-cell level is to spread the cells of the reporter strain on the agar pad (a thin agar layer between the microscope slide and the cover glass) or on the bottom of solidified medium inside culture dishes (Joyce et al., 2011; Dhar and Manina, 2015). Although simple and low-cost, this approach is not always applicable (e.g., labeling and medium changing). Microfluidic flow chambers are used for the latter purposes, as well as for rapidly changing culture conditions (e.g., applying stress). Various microfluidic chips and plates are commercially available from an increasing number of companies, whereas custom made (usually PDMS) chips are a cost-reducing alternative and also allow for more personalized applications (Wang et al., 2010; Cattoni et al., 2013; Dhar and Manina, 2015; Trojanowski et al., 2015; Wallden et al., 2016). The architecture of microfluidic chips and plates varies among studies and choosing the right one should be dictated by the specific study purpose and the availability of additional equipment, e.g., peristaltic/syringe/pressure pumps, flow controllers, or automation.

SPATIOTEMPORAL LOCALIZATION OF THE REPLISOME DURING REPLICATION INITIATION

Localization of the replication machinery at the beginning of DNA synthesis is dependent on *oriC* position, and is therefore connected with the spatial arrangement of the chromosome. In bacteria having *oriC* and *ter* regions positioned at the mid-cell, the intervening chromosomal regions (i.e., the left and right chromosomal arms) are stretched out toward opposite cell poles, creating a *left-ori-right* pattern, whereas cells having *oriC* and *ter* regions localized to opposite poles show an *ori-ter* chromosomal arrangement (Wang and Rudner, 2014). Replisomes in the cells exhibiting a *left-ori-right* configuration are assembled in the mid-cell region of the chromosome. This pattern has been observed in *E. coli* cells (Postow et al., 2004; Valens et al., 2004; Boccard et al., 2005) and during the vegetative growth of *B. subtilis* (the chromosome in *B. subtilis* is oscillating between *left-ori-right* and *ori-ter* configuration) (Wang et al., 2014; **Figure 3A**). During sporulation, however, the *B. subtilis* chromosome adopts an *ori-ter* orientation to segregate an entire copy of the chromosome within each spore. Positioning of the *oriC* at the mid-cell of *B. subtilis* and *E. coli* is maintained by the condensins SMC and MukB (a structural homolog of SMC), respectively (Niki et al., 1992; Danilova et al., 2007; Sullivan et al., 2009). SMC can compact large chromosomal regions, and, by interacting with ParB protein, organizes the *oriC*-proximal regions in *B. subtilis*, with ParB binding to *parS* sequences located near *oriC* (Gruber and Errington, 2009). The interaction of MukB with the nucleoid associated protein HU ensures proper *oriC* positioning in *E. coli* cells (Lioy et al., 2018). After initiation, *E. coli* replisomes oscillate near the cell center, while newly replicated *oriCs* are segregated toward the cell poles (Reyes-Lamothe et al., 2008). In comparison, *B. subtilis* replisomes colocalize throughout replication (Migocki et al., 2004), and are therefore visible as a single fluorescent focus. Replisome positioning in the cell center can be also found in oval-shaped *S. pneumoniae* (Kjos and Veening, 2014; van

FIGURE 3 | Spatial organization of the chromosome entails positioning of the site of replisome assembly. Bacteria exhibiting a *left-ori-right* orientation start replication in mid-cell **(A)**, where the *oriC* region is organized by the condensins SMC/MukB (marked in green) and ParB (indicated as yellow circle). Off-centered replisome positioning **(B)** is associated with complex interactions between *oriC* and the ParABS system (*M. smegmatis*; ParA indicated as orange cloud, ParB in yellow circle and DivIVA in red) or bactofilins (*M. xanthus*; BacNOP depicted as violet cloud). In the *ori-ter* organized chromosomes, replication is initiated at the cell pole **(C)**, at which the *oriC* region is anchored by specific proteins (i.e., PopZ in *C. crescentus* and HubP in *V. cholerae* indicated as blue and marine blue ovals, respectively). **(D)** Subpolar positioning of replisomes has also been observed in the multiploid bacteria *S. coelicolor* (ParA indicated as orange cloud and polarisome complex proteins: ParB and DivIVA interacting with Scy depicted in yellow and red, respectively) and *C. glutamicum*. OriC region(s) and replisome(s) are indicated as violet and green circles, while chromosome is depicted in light blue.

Raaphorst et al., 2017), which, similar to many other bacteria including *B. subtilis*, encodes an SMC homolog.

Some bacteria, such as *M. smegmatis* (Santi and McKinney, 2015; Trojanowski et al., 2015) and *M. xanthus* (Harms et al., 2013), exhibit off-center replisome localization during the

initiation of replication (see **Figure 3B**). In *M. smegmatis*, segregation of the newly replicated *oriC*s starts immediately after initiation of replication, with one *oriC* remaining near the old cell pole and the other traveling toward the opposite pole (Ginda et al., 2017; Hołówka et al., 2018). Replisomes oscillate in the

old-pole-proximal cell half during most of the replication process, but localize closer to the new cell pole prior to termination (Trojanowski et al., 2015). A slight asymmetry in mycobacterial replisome positioning is associated with the apical growth mode of these bacteria. Positioning of *oriC* region(s) in *Mycobacterium* depends on the interaction of ParB with ParA protein, which in turn interacts with the polar growth determinant, DivIVA protein (Ginda et al., 2013).

As a result of the asymmetric location of *oriC*, *M. xanthus* replisomes are positioned at the subpolar regions (**Figure 3B**; Harms et al., 2013). Although *M. xanthus* contains a DivIVA homolog, suggesting analogous interactions at the pole as described for *Mycobacterium*, deletion of this homolog does not affect cell division or chromosome segregation. Rather, localization of the ParA and ParB-*parS* complexes (and thus the *oriC* region) in *M. xanthus* is controlled by the bactofilins BacNOP, through the direct interactions of ParA and ParB with the scaffold created by BacNOP (Lin et al., 2017).

Bacteria exhibiting complex life cycles often show an *ori-ter* chromosome orientation (**Figure 3C**). In *C. crescentus* stalked cells, chromosome replication starts at the old cell pole (Jensen et al., 2001). The anchorage of the chromosome at the old cell pole is maintained by the protein PopZ (Bowman et al., 2008). Similarly, in *V. cholerae*, the origin (*ori*I) of one of the two chromosomes, chrI, is attached to the old pole by HubP protein (Yamaichi et al., 2012), thereby setting the subcellular position for assembly of the replication machinery. In contrast, the origin (*ori*II) of the second, smaller chromosome (chrII) is located at mid-cell. Replication of *V. cholerae* chrII starts later than that of chrI to synchronize the termination of replication of both chromosomes (Demarre et al., 2014; Ramachandran et al., 2018). As a result of the subpolar localization of *C. crescentus* and *V. cholerae* (chrI) replisomes near the old cell pole, one of the newly replicated *oriC* regions travels across the chromosome to the opposite cell pole with the assistance of the ParAB*S* system (Toro et al., 2008; Ramachandran et al., 2014). Interestingly, in *P. aeruginosa* exhibiting *ori-ter* orientation, the chromosome is apparently not anchored to the cell pole, as shown by the cytoplasmic gap between *oriC* and the cell pole (Vallet-Gely and Boccard, 2013).

The multiploid and apically growing bacterial species *S. coelicolor*, exhibits another mode of spatiotemporal replisome localization, in which replication is initiated during vegetative growth (**Figure 3D**; Kois-Ostrowska et al., 2016). Replication of multiple copies of the *S. coelicolor* chromosome starts asynchronously, and newly replicated sister chromosomes follow the extending hyphal tip. Similar to *Mycobacterium*, positioning of the tip-proximal *oriC* (and hence the replisomes) is maintained through ParA interactions with the polarisome complex, which includes the proteins ParB, DivIVA, and Scy (Flärdh et al., 2012; Ditkowski et al., 2013). In the closely related and diploid species *C. glutamicum*, replisomes are assembled on each chromosome asymmetrically, in proximity to the cell poles (**Figure 3D**; Böhm et al., 2017). Fluorescently tagged ParB attaches to the cell poles, suggesting an *ori-ter-ter-ori* spatial orientation of *C. glutamicum* chromosomes.

Described differences among bacteria in the positioning of *oriC* regions during the replication initiation reflect the different modes of chromosome segregation. Mid-cell replisomes location results in symmetric segregation of *oriC*s toward the opposite cell poles, while polar and off-center replisome positioning imply asymmetric segregation of the newly replicated *oriC* regions. Furthermore, polar localization requires the complex system to either anchor *oriC* directly at the pole (e.g., PopZ and HubP proteins) or to maintain the subpolar position by protein complexes (e.g., the interaction of ParAB*S* system with the DivIVA or the BacNOP). Such variety in the composition of multiprotein complexes involved in *oriC*(s) positioning provides an opportunity for the discovery of novel genus/species-specific drug targets.

CONCLUSION

Single-cell fluorescence imaging and fluorescence tagging techniques allow researchers to precisely visualize proteins and their complexes inside living bacterial cells in real time. These techniques revealed that many proteins are targeted to distinct subcellular positions, where they participate in various cellular processes including chromosome replication. Recent studies using advanced live-cell imaging demonstrated that chromosome replication is coordinated with other key steps of the cell cycle, such as chromosome segregation and cell division. Proteins (or protein complexes) involved in condensation (i.e., SMC/MukB), chromosome segregation (i.e., ParAB in Gram-negative and Gram-positive bacteria) and/or cell division (DivIVA in Gram-positive bacteria) take part directly or indirectly in *oriC* positioning, thus indicating the site of replisome assembly. Additionally, other proteins guiding the *oriC* region have been recently identified. Interestingly, they vary significantly among different bacteria, e.g., PopZ (*C. crescents*), HubP (*V. cholerae*, chromosome I), and bactofilins (*M. xanthus*). The diversity and complexity of the systems involved in *oriC* (and thus replisome) subcellular positioning suggest the possibility of developing new antimicrobial therapies and/or altering existing treatments (Kaguni, 2018).

AUTHOR CONTRIBUTIONS

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

REFERENCES

Arias-Cartin, R., Dobihal, G. S., Campos, M., Surovtsev, I. V., Parry, B., and Jacobs-Wagner, C. (2017). Replication fork passage drives asymmetric dynamics of a critical nucleoid-associated protein in *Caulobacter*. *EMBO J.* 36, 301–318. doi: 10.15252/embj.201695513

Aves, S. J. (2009). DNA replication initiation. *Methods Mol. Biol. Clifton NJ* 521, 3–17. doi: 10.1007/978-1-60327-815-7_1

Bach, T., Morigen, and Skarstad, K. (2008). The initiator protein DnaA contributes to keeping new origins inactivated by promoting the presence of hemimethylated DNA. *J. Mol. Biol.* 384, 1076–1085. doi: 10.1016/j.jmb.2008.09.042

Badrinarayanan, A., Le, T. B. K., 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

Bates, D., and Kleckner, N. (2005). Chromosome and replisome dynamics in *E. coli*: loss of sister cohesion triggers global chromosome movement and mediates chromosome segregation. *Cell* 121, 899–911. doi: 10.1016/j.cell.2005.04.013

Beattie, T. R., Kapadia, N., Nicolas, E., Uphoff, S., Wollman, A. J., Leake, M. C., et al. (2017). Frequent exchange of the DNA polymerase during bacterial chromosome replication. *eLife* 6:e21763. doi: 10.7554/eLife.21763

Berkmen, M. B., and Grossman, A. D. (2006). Spatial and temporal organization of the *Bacillus subtilis* replication cycle. *Mol. Microbiol.* 62, 57–71. doi: 10.1111/j.1365-2958.2006.05356.x

Bird, R. E., Louarn, J., Martuscelli, J., and Caro, L. (1972). Origin and sequence of chromosome replication in *Escherichia coli*. *J. Mol. Biol.* 70, 549–566. doi: 10.1016/0022-2836(72)90559-1

Boccard, F., Esnault, E., and Valens, M. (2005). Spatial arrangement and macrodomain organization of bacterial chromosomes. *Mol. Microbiol.* 57, 9–16. doi: 10.1111/j.1365-2958.2005.04651.x

Böhm, K., Meyer, F., Rhomberg, A., Kalinowski, J., Donovan, C., and Bramkamp, M. (2017). Novel chromosome organization pattern in actinomycetales—overlapping replication cycles combined with diploidy. *mBio* 8:e00511-17. doi: 10.1128/mBio.00511-17

Bowman, G. R., Comolli, L. R., Zhu, J., Eckart, M., Koenig, M., Downing, K. H., et al. (2008). A polymeric protein anchors the chromosomal origin/ParB complex at a bacterial cell pole. *Cell* 134, 945–955. doi: 10.1016/j.cell.2008.07.015

Cattoni, D. I., Fiche, J.-B., Valeri, A., Mignot, T., and Nöllmann, M. (2013). Super-resolution imaging of bacteria in a microfluidics device. *PLoS One* 8:e76268. doi: 10.1371/journal.pone.0076268

Costantini, L. M., Fossati, M., Francolini, M., and Snapp, E. L. (2012). Assessing the tendency of fluorescent proteins to oligomerize under physiologic conditions. *Traffic Cph. Den.* 13, 643–649. doi: 10.1111/j.1600-0854.2012.01336.x

Danilova, O., Reyes-Lamothe, R., Pinskaya, M., Sherratt, D., and Possoz, C. (2007). MukB colocalizes with the oriC region and is required for organization of the two *Escherichia coli* chromosome arms into separate cell halves. *Mol. Microbiol.* 65, 1485–1492. doi: 10.1111/j.1365-2958.2007.05881.x

Demarre, G., Galli, E., Muresan, L., Paly, E., David, A., Possoz, C., et al. (2014). Differential management of the replication terminus regions of the two *Vibrio cholerae* chromosomes during cell division. *PLoS Genet.* 10:e1004557. doi: 10.1371/journal.pgen.1004557

Dhar, N., and Manina, G. (2015). Single-cell analysis of mycobacteria using microfluidics and time-lapse microscopy. *Methods Mol. Biol. Clifton NJ* 1285, 241–256. doi: 10.1007/978-1-4939-2450-9_14

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

Donczew, M., Mackiewicz, P., Wróbel, A., Flärdh, K., Zakrzewska-Czerwińska, J., and Jakimowicz, D. (2016). ParA and ParB coordinate chromosome segregation with cell elongation and division during *Streptomyces sporulation*. *Open Biol.* 6:150263. doi: 10.1098/rsob.150263

Donczew, R., Weigel, C., Lurz, R., Zakrzewska-Czerwińska, J., and Zawilak-Pawlik, A. (2012). *Helicobacter* pylori oriC—the first bipartite origin of chromosome replication in gram-negative bacteria. *Nucleic Acids Res.* 40, 9647–9660. doi: 10.1093/nar/gks742

Donovan, C., Schwaiger, A., Krämer, R., and Bramkamp, M. (2010). Subcellular localization and characterization of the ParAB system from *Corynebacterium glutamicum*. *J. Bacteriol.* 192, 3441–3451. doi: 10.1128/JB.00214-10

Fernandez-Fernandez, C., Grosse, K., Sourjik, V., and Collier, J. (2013). The β-sliding clamp directs the localization of HdaA to the replisome in *Caulobacter crescentus*. *Microbiology* 159, 2237–2248. doi: 10.1099/mic.0.068577-0

Flärdh, K., Richards, D. M., Hempel, A. M., Howard, M., and Buttner, M. J. (2012). Regulation of apical growth and hyphal branching in *Streptomyces*. *Curr. Opin. Microbiol.* 15, 737–743. doi: 10.1016/j.mib.2012.10.012

Fogel, M. A., and Waldor, M. K. (2005). Distinct segregation dynamics of the two *Vibrio cholerae* chromosomes. *Mol. Microbiol.* 55, 125–136. doi: 10.1111/j.1365-2958.2004.04379.x

Fossum, S., Crooke, E., and Skarstad, K. (2007). Organization of sister origins and replisomes during multifork DNA replication in *Escherichia coli*. *EMBO J.* 26, 4514–4522. doi: 10.1038/sj.emboj.7601871

Frunzke, J., Bramkamp, M., Schweitzer, J.-E., and Bott, M. (2008). Population heterogeneity in *Corynebacterium glutamicum* ATCC 13032 caused by prophage CGP3. *J. Bacteriol.* 190, 5111–5119. doi: 10.1128/JB.00310-08

Gao, F., and Zhang, C.-T. (2008). Ori-Finder: a web-based system for finding oriCs in unannotated bacterial genomes. *BMC Bioinformatics* 9:79. doi: 10.1186/1471-2105-9-79

Georgescu, R. E., Kurth, I., and O'Donnell, M. E. (2012). Single-molecule studies reveal the function of a third polymerase in the replisome. *Nat. Struct. Mol. Biol.* 19, 113–116. doi: 10.1038/nsmb.2179

Ginda, K., Bezulska, M., Ziółkiewicz, M., Dziadek, J., Zakrzewska-Czerwińska, J., and Jakimowicz, D. (2013). ParA of *Mycobacterium smegmatis* co-ordinates chromosome segregation with the cell cycle and interacts with the polar growth determinant DivIVA. *Mol. Microbiol.* 87, 998–1012. doi: 10.1111/mmi.12146

Ginda, K., Santi, I., Bousbaine, D., Zakrzewska-Czerwińska, J., Jakimowicz, D., and McKinney, J. (2017). The studies of ParA and ParB dynamics reveal asymmetry of chromosome segregation in Mycobacteria. *Mol. Microbiol.* 105, 453–468. doi: 10.1111/mmi.13712

Gruber, S., and Errington, J. (2009). Recruitment of condensin to replication origin regions by ParB/SpoOJ promotes chromosome segregation in *B. subtilis*. *Cell* 137, 685–696. doi: 10.1016/j.cell.2009.02.035

HaloTag® Protein Purification System (2018). *HaloTag Protein Purification System*. Available at: https://pl.promega.com/products/protein-purification/protein-purification-kits/halotag-protein-purification-system/ [Accessed September 28, 2018].

Harms, A., Treuner-Lange, A., Schumacher, D., and Søgaard-Andersen, L. (2013). Tracking of chromosome and replisome dynamics in *Myxococcus xanthus* reveals a novel chromosome arrangement. *PLoS Genet.* 9:e1003802. doi: 10.1371/journal.pgen.1003802

Hołówka, J., Trojanowski, D., Janczak, M., Jakimowicz, D., and Zakrzewska-Czerwińska, J. (2018). The origin of chromosomal replication is asymmetrically positioned on the mycobacterial nucleoid, and the timing of its firing depends on HupB. *J. Bacteriol.* 200:e00044-18. doi: 10.1128/JB.00044-18

Jaworski, P., Donczew, R., Mielke, T., Weigel, C., Stingl, K., and Zawilak-Pawlik, A. (2018). Structure and function of the *Campylobacter jejuni* chromosome replication origin. *Front. Microbiol.* 9:1533. doi: 10.3389/fmicb.2018.01533

Jensen, R. B., Wang, S. C., and Shapiro, L. (2001). A moving DNA replication factory in *Caulobacter crescentus*. *EMBO J.* 20, 4952–4963. doi: 10.1093/emboj/20.17.4952

Joyce, G., Robertson, B. D., and Williams, K. J. (2011). A modified agar pad method for mycobacterial live-cell imaging. *BMC Res. Notes* 4:73. doi: 10.1186/1756-0500-4-73

Kaguni, J. M. (2018). The macromolecular machines that duplicate the *Escherichia coli* chromosome as targets for drug discovery. *Antibiot. Basel Switz.* 7:E23. doi: 10.3390/antibiotics7010023

Kaguni, J. M., and Kornberg, A. (1984). Replication initiated at the origin (oriC) of the *E. coli* chromosome reconstituted with purified enzymes. *Cell* 38, 183–190. doi: 10.1016/0092-8674(84)90539-7

Katayama, T. (2017). Initiation of DNA replication at the chromosomal origin of *E. coli*, oriC. *Adv. Exp. Med. Biol.* 1042, 79–98. doi: 10.1007/978-981-10-6955-0_4

Kjos, M., and Veening, J.-W. (2014). Tracking of chromosome dynamics in live *Streptococcus pneumoniae* reveals that transcription promotes chromosome segregation. *Mol. Microbiol.* 91, 1088–1105. doi: 10.1111/mmi.12517

Kois-Ostrowska, A., Strzałka, A., Lipietta, N., Tilley, E., Zakrzewska-Czerwińska, J., Herron, P., et al. (2016). Unique function of the bacterial chromosome segregation machinery in apically growing streptomyces – Targeting the chromosome to new hyphal tubes and its anchorage at the tips. *PLoS Genet.* 12:e1006488. doi: 10.1371/journal.pgen.1006488

Kongsuwan, K., Dalrymple, B. P., Wijffels, G., and Jennings, P. A. (2002). Cellular localisation of the clamp protein during DNA replication. *FEMS Microbiol. Lett.* 216, 255–262. doi: 10.1111/j.1574-6968.2002.tb11444.x

Laloux, G., and Jacobs-Wagner, C. (2013). Spatiotemporal control of PopZ localization through cell cycle–coupled multimerization. *J. Cell Biol.* 201, 827–841. doi: 10.1083/jcb.201303036

Landete, J. M., Langa, S., Revilla, C., Margolles, A., Medina, M., and Arqués, J. L. (2015). Use of anaerobic green fluorescent protein versus green fluorescent protein as reporter in lactic acid bacteria. *Appl. Microbiol. Biotechnol.* 99, 6865–6877. doi: 10.1007/s00253-015-6770-3

Lau, I. F., Filipe, S. R., Søballe, B., Økstad, O.-A., Barre, F.-X., and Sherratt, D. J. (2003). Spatial and temporal organization of replicating *Escherichia coli* chromosomes. *Mol. Microbiol.* 49, 731–743. doi: 10.1046/j.1365-2958.2003.03640.x

Le, T. B., and Laub, M. T. (2014). New approaches to understanding the spatial organization of bacterial genomes. *Curr. Opin. Microbiol.* 22, 15–21. doi: 10.1016/j.mib.2014.09.014

Lemon, K. P., and Grossman, A. D. (1998). Localization of bacterial DNA polymerase: evidence for a factory model of replication. *Science* 282, 1516–1519. doi: 10.1126/science.282.5393.1516

Lemon, K. P., and Grossman, A. D. (2000). Movement of replicating DNA through a stationary replisome. *Mol. Cell* 6, 1321–1330. doi: 10.1016/S1097-2765(00)00130-1

Leonard, A. C., and Grimwade, J. E. (2011). Regulation of DnaA assembly and activity: taking directions from the genome. *Annu. Rev. Microbiol.* 65, 19–35. doi: 10.1146/annurev-micro-090110-102934

Leonard, A. C., and Grimwade, J. E. (2015). The orisome: structure and function. *Front. Microbiol.* 6:545. doi: 10.3389/fmicb.2015.00545

Lewis, J. S., Spenkelink, L. M., Jergic, S., Wood, E. A., Monachino, E., Horan, N. P., et al. (2017). Single-molecule visualization of fast polymerase turnover in the bacterial replisome. *eLife* 6:e23932. doi: 10.7554/eLife.23932

Li, Y., Schroeder, J. W., Liao, Y., Chen, Z., Simmons, L. A., and Biteen, J. S. (2018). Super-resolution imaging of bacterial replisome dynamics in live *Bacillus subtilis*. *Biophys. J.* 114:539a. doi: 10.1016/j.bpj.2017.11.2942

Li, Y., Sergueev, K., and Austin, S. (2002). The segregation of the *Escherichia coli* origin and terminus of replication. *Mol. Microbiol.* 46, 985–996. doi: 10.1046/j.1365-2958.2002.03234.x

Liao, Y., Li, Y., Schroeder, J. W., Simmons, L. A., and Biteen, J. S. (2016). Single-molecule DNA polymerase dynamics at a bacterial replisome in live cells. *Biophys. J.* 111, 2562–2569. doi: 10.1016/j.bpj.2016.11.006

Lin, L., Osorio Valeriano, M., Harms, A., Søgaard-Andersen, L., and Thanbichler, M. (2017). Bactofilin-mediated organization of the ParABS chromosome segregation system in *Myxococcus xanthus*. *Nat. Commun.* 8:1817. doi: 10.1038/s41467-017-02015-z

Lin, Y.-S., Kieser, H. M., Hopwood, D. A., and Chen, C. W. (1993). The chromosomal DNA of *Streptomyces lividans* 66 is linear. *Mol. Microbiol.* 10, 923–933. doi: 10.1111/j.1365-2958.1993.tb00964.x

Lioy, V. S., Cournac, A., Marbouty, M., Duigou, S., Mozziconacci, J., Espéli, O., et al. (2018). Multiscale structuring of the *E. coli* chromosome by nucleoid-associated and condensin proteins. *Cell* 172, 771.e18–783.e18. doi: 10.1016/j.cell.2017.12.027

Liu, X., Wang, X., Reyes-Lamothe, R., and Sherratt, D. (2010). Replication-directed sister chromosome alignment in *Escherichia coli*. *Mol. Microbiol.* 75, 1090–1097. doi: 10.1111/j.1365-2958.2009.06791.x

Livny, J., Yamaichi, Y., and Waldor, M. K. (2007). Distribution of centromere-like parS sites in bacteria: insights from comparative genomics. *J. Bacteriol.* 189, 8693–8703. doi: 10.1128/JB.01239-07

Makowski, Ł., Donczew, R., Weigel, C., Zawilak-Pawlik, A., and Zakrzewska-Czerwińska, J. (2016). Initiation of chromosomal replication in predatory bacterium *Bdellovibrio* bacteriovorus. *Front. Microbiol.* 7:1898. doi: 10.3389/fmicb.2016.01898

Mangiameli, S. M., Veit, B. T., Merrikh, H., and Wiggins, P. A. (2017). The replisomes remain spatially proximal throughout the cell cycle in bacteria. *PLoS Genet.* 13:e1006582. doi: 10.1371/journal.pgen.1006582

Masai, H., Matsumoto, S., You, Z., Yoshizawa-Sugata, N., and Oda, M. (2010). Eukaryotic chromosome DNA replication: where, when, and how? *Annu. Rev. Biochem.* 79, 89–130. doi: 10.1146/annurev.biochem.052308.103205

Méchali, M. (2010). Eukaryotic DNA replication origins: many choices for appropriate answers. *Nat. Rev. Mol. Cell Biol.* 11, 728–738. doi: 10.1038/nrm2976

Mettrick, K. A., and Grainge, I. (2016). Stability of blocked replication forks in vivo. *Nucleic Acids Res.* 44, 657–668. doi: 10.1093/nar/gkv1079

Midgley-Smith, S. L., Dimude, J. U., Taylor, T., Forrester, N. M., Upton, A. L., Lloyd, R. G., et al. (2018). Chromosomal over-replication in *Escherichia coli* recG cells is triggered by replication fork fusion and amplified if replichore symmetry is disturbed. *Nucleic Acids Res.* 46, 7701–7715. doi: 10.1093/nar/gky566

Migocki, M. D., Lewis, P. J., Wake, R. G., and Harry, E. J. (2004). The midcell replication factory in *Bacillus subtilis* is highly mobile: implications for coordinating chromosome replication with other cell cycle events. *Mol. Microbiol.* 54, 452–463. doi: 10.1111/j.1365-2958.2004.04267.x

Moolman, M. C., Krishnan, S. T., Kerssemakers, J. W. J., van den Berg, A., Tulinski, P., Depken, M., et al. (2014). Slow unloading leads to DNA-bound β2-sliding clamp accumulation in live *Escherichia coli* cells. *Nat. Commun.* 5:5820. doi: 10.1038/ncomms6820

Nielsen, H. J., Ottesen, J. R., Youngren, B., Austin, S. J., and Hansen, F. G. (2006). The *Escherichia coli* chromosome is organized with the left and right chromosome arms in separate cell halves. *Mol. Microbiol.* 62, 331–338. doi: 10.1111/j.1365-2958.2006.05346.x

Nielsen, H. J., Youngren, B., Hansen, F. G., and Austin, S. (2007). Dynamics of *Escherichia coli* chromosome segregation during multifork replication. *J. Bacteriol.* 189, 8660–8666. doi: 10.1128/JB.01212-07

Nielsen, O., and Løbner-Olesen, A. (2008). Once in a lifetime: strategies for preventing re-replication in prokaryotic and eukaryotic cells. *EMBO Rep.* 9, 151–156. doi: 10.1038/sj.embor.2008.2

Niki, H., Imamura, R., Kitaoka, M., Yamanaka, K., Ogura, T., and Hiraga, S. (1992). *E. coli* MukB protein involved in chromosome partition forms a homodimer with a rod-and-hinge structure having DNA binding and ATP/GTP binding activities. *EMBO J.* 11, 5101–5109. doi: 10.1002/j.1460-2075.1992.tb05617.x

Possoz, C., Filipe, S. R., Grainge, I., and Sherratt, D. J. (2006). Tracking of controlled *Escherichia coli* replication fork stalling and restart at repressor-bound DNA in vivo. *EMBO J.* 25, 2596–2604. doi: 10.1038/sj.emboj.7601155

Postow, L., Hardy, C. D., Arsuaga, J., and Cozzarelli, N. R. (2004). Topological domain structure of the *Escherichia coli* chromosome. *Genes Dev.* 18, 1766–1779. doi: 10.1101/gad.1207504

Raaphorst, R., van Kjos, M., and Veening, J.-W. (2017). Chromosome segregation drives division site selection in *Streptococcus pneumoniae*. *Proc. Natl. Acad. Sci. U.S.A.* 114, E5959–E5968. doi: 10.1073/pnas.1620608114

Ramachandran, R., Ciaccia, P. N., Filsuf, T. A., Jha, J. K., and Chattoraj, D. K. (2018). Chromosome 1 licenses chromosome 2 replication in *Vibrio cholerae* by doubling the crtS gene dosage. *PLoS Genet.* 14:e1007426. doi: 10.1371/journal.pgen.1007426

Ramachandran, R., Jha, J., and Chattoraj, D. (2014). Chromosome segregation in *Vibrio cholerae*. *J. Mol. Microbiol. Biotechnol.* 24, 360–370. doi: 10.1159/000368853

Reyes-Lamothe, R., Nicolas, E., and Sherratt, D. J. (2012). Chromosome replication and segregation in bacteria. *Annu. Rev. Genet.* 46, 121–143. doi: 10.1146/annurev-genet-110711-155421

Reyes-Lamothe, R., Possoz, C., Danilova, O., and Sherratt, D. J. (2008). Independent positioning and action of *Escherichia coli* replisomes in live cells. *Cell* 133, 90–102. doi: 10.1016/j.cell.2008.01.044

Reyes-Lamothe, R., Sherratt, D. J., and Leake, M. C. (2010). Stoichiometry and architecture of active DNA replication machinery in *Escherichia coli*. *Science* 328, 498–501. doi: 10.1126/science.1185757

Richardson, T. T., Harran, O., and Murray, H. (2016). The bacterial DnaA-trio replication origin element specifies single-stranded DNA initiator binding. *Nature* 534, 412–416. doi: 10.1038/nature17962

Ruban-Ośmiałowska, B., Jakimowicz, D., Smulczyk-Krawczyszyn, A., Chater, K. F., and Zakrzewska-Czerwińska, J. (2006). Replisome localization in vegetative and aerial hyphae of *Streptomyces coelicolor*. *J. Bacteriol.* 188, 7311–7316. doi: 10.1128/JB.00940-06

Samadpour, A. N., and Merrikh, H. (2018). DNA gyrase activity regulates DnaA-dependent replication initiation in *Bacillus subtilis*. *Mol. Microbiol.* 108, 115–127. doi: 10.1111/mmi.13920

Santi, I., Dhar, N., Bousbaine, D., Wakamoto, Y., and McKinney, J. D. (2013). Single-cell dynamics of the chromosome replication and cell division cycles in mycobacteria. *Nat. Commun.* 4:2470. doi: 10.1038/ncomms3470

Santi, I., and McKinney, J. D. (2015). Chromosome organization and replisome dynamics in *Mycobacterium smegmatis*. *mBio* 6:e01999-14. doi: 10.1128/mBio.01999-14

Shaner, N. C., Steinbach, P. A., and Tsien, R. Y. (2005). A guide to choosing fluorescent proteins. *Nat. Methods* 2, 905–909. doi: 10.1038/nmeth819

Skarstad, K., Baker, T. A., and Kornberg, A. (1990). Strand separation required for initiation of replication at the chromosomal origin of *E. coli* is facilitated by a distant RNA–DNA hybrid. *EMBO J.* 9, 2341–2348. doi: 10.1002/j.1460-2075.1990.tb07406.x

Skarstad, K., Boye, E., and Steen, H. B. (1986). Timing of initiation of chromosome replication in individual *Escherichia coli* cells. *EMBO J.* 5, 1711–1717. doi: 10.1002/j.1460-2075.1986.tb04415.x

Skarstad, K., and Katayama, T. (2013). Regulating DNA replication in bacteria. *Cold Spring Harb. Perspect. Biol.* 5:a012922. doi: 10.1101/cshperspect.a012922

Srivastava, P., and Chattoraj, D. K. (2007). Selective chromosome amplification in *Vibrio cholerae*. *Mol. Microbiol.* 66, 1016–1028. doi: 10.1111/j.1365-2958.2007.05973.x

Stokke, C., Waldminghaus, T., and Skarstad, K. (2011). Replication patterns and organization of replication forks in *Vibrio cholerae*. *Microbiology* 157, 695–708. doi: 10.1099/mic.0.045112-0

Stracy, M., Uphoff, S., Garza de Leon, F., and Kapanidis, A. N. (2014). In vivo single-molecule imaging of bacterial DNA replication, transcription, and repair. *FEBS Lett.* 588, 3585–3594. doi: 10.1016/j.febslet.2014.05.026

Su'etsugu, M., and Errington, J. (2011). The replicase sliding clamp dynamically accumulates behind progressing replication forks in *Bacillus subtilis* cells. *Mol. Cell* 41, 720–732. doi: 10.1016/j.molcel.2011.02.024

Sukumar, N., Tan, S., Aldridge, B. B., and Russell, D. G. (2014). Exploitation of *Mycobacterium tuberculosis* reporter strains to probe the impact of vaccination at sites of infection. *PLoS Pathog.* 10:e1004394. doi: 10.1371/journal.ppat.1004394

Sullivan, N. L., Marquis, K. A., and Rudner, D. Z. (2009). Recruitment of SMC by ParB-parS organizes the origin region and promotes efficient chromosome segregation. *Cell* 137, 697–707. doi: 10.1016/j.cell.2009.04.044

Toro, E., Hong, S.-H., McAdams, H. H., and Shapiro, L. (2008). Caulobacter requires a dedicated mechanism to initiate chromosome segregation. *Proc. Natl. Acad. Sci. U.S.A.* 105, 15435–15440. doi: 10.1073/pnas.0807448105

Trojanowski, D., Ginda, K., Pióro, M., Hołówka, J., Skut, P., Jakimowicz, D., et al. (2015). Choreography of the *Mycobacterium* replication machinery during the cell cycle. *mBio* 6:e02125-14. doi: 10.1128/mBio.02125-14

Trojanowski, D., Hołówka, J., Ginda, K., Jakimowicz, D., and Zakrzewska-Czerwińska, J. (2017). Multifork chromosome replication in slow-growing bacteria. *Sci. Rep.* 7:43836. doi: 10.1038/srep43836

Trucksis, M., Michalski, J., Deng, Y. K., and Kaper, J. B. (1998). The *Vibrio cholerae* genome contains two unique circular chromosomes. *Proc. Natl. Acad. Sci. U.S.A.* 95, 14464–14469. doi: 10.1073/pnas.95.24.14464

Valens, M., Penaud, S., Rossignol, M., Cornet, F., and Boccard, F. (2004). Macrodomain organization of the *Escherichia coli* chromosome. *EMBO J.* 23, 4330–4341. doi: 10.1038/sj.emboj.7600434

Vallet-Gely, I., and Boccard, F. (2013). Chromosomal organization and segregation in *Pseudomonas aeruginosa*. *PLoS Genet.* 9:e1003492. doi: 10.1371/journal.pgen.1003492

van Raaphorst, R., Kjos, M., and Veening, J.-W. (2017). Chromosome segregation drives division site selection in *Streptococcus pneumoniae*. *Proc. Natl. Acad. Sci. U.S.A.* 114, E5959–E5968. doi: 10.1073/pnas.1620608114

Viollier, P. H., Thanbichler, M., McGrath, P. T., West, L., Meewan, M., McAdams, H. H., et al. (2004). Rapid and sequential movement of individual chromosomal loci to specific subcellular locations during bacterial DNA replication. *Proc. Natl. Acad. Sci. U.S.A.* 101, 9257–9262. doi: 10.1073/pnas.0402606101

Wallden, M., Fange, D., Lundius, E. G., Baltekin, Ö., and Elf, J. (2016). The synchronization of replication and division cycles in individual *E. coli* cells. *Cell* 166, 729–739. doi: 10.1016/j.cell.2016.06.052

Wang, P., Robert, L., Pelletier, J., Dang, W. L., Taddei, F., Wright, A., et al. (2010). Robust growth of *Escherichia coli*. *Curr. Biol. CB* 20, 1099–1103. doi: 10.1016/j.cub.2010.04.045

Wang, X., Lesterlin, C., Reyes-Lamothe, R., Ball, G., and Sherratt, D. J. (2011). Replication and segregation of an *Escherichia coli* chromosome with two replication origins. *Proc. Natl. Acad. Sci. U.S.A.* 108, E243–E250. doi: 10.1073/pnas.1100874108

Wang, X., Montero Llopis, P., and Rudner, D. Z. (2014). *Bacillus subtilis* chromosome organization oscillates between two distinct patterns. *Proc. Natl. Acad. Sci. U.S.A.* 111, 12877–12882. doi: 10.1073/pnas.1407461111

Wang, X., and Rudner, D. Z. (2014). Spatial organization of bacterial chromosomes. *Curr. Opin. Microbiol.* 22, 66–72. doi: 10.1016/j.mib.2014.09.016

Wang, X., and Sherratt, D. J. (2010). Independent segregation of the two arms of the *Escherichia coli* ori region requires neither RNA synthesis nor MreB dynamics. *J. Bacteriol.* 192, 6143–6153. doi: 10.1128/JB.00861-10

Winterstein, C., and Ludwig, B. (1998). Genes coding for respiratory complexes map on all three chromosomes of the *Paracoccus denitrificans* genome. *Arch. Microbiol.* 169, 275–281. doi: 10.1007/s002030050572

Wolański, M., Donczew, R., Zawilak-Pawlik, A., and Zakrzewska-Czerwińska, J. (2014). oriC-encoded instructions for the initiation of bacterial chromosome replication. *Front. Microbiol.* 5:735. doi: 10.3389/fmicb.2014.00735

Wolański, M., Wali, R., Tilley, E., Jakimowicz, D., Zakrzewska-Czerwińska, J., and Herron, P. (2011). Replisome trafficking in growing vegetative hyphae of *Streptomyces coelicolor* A3(2). *J. Bacteriol.* 193, 1273–1275. doi: 10.1128/JB.01326-10

Yamaichi, Y., Bruckner, R., Ringgaard, S., Möll, A., Cameron, D. E., Briegel, A., et al. (2012). A multidomain hub anchors the chromosome segregation and chemotactic machinery to the bacterial pole. *Genes Dev.* 26, 2348–2360. doi: 10.1101/gad.199869.112

Youngren, B., Radnedge, L., Hu, P., Garcia, E., and Austin, S. (2000). A plasmid partition system of the P1-P7par family from the pMT1 virulence plasmid of *Yersinia pestis*. *J. Bacteriol.* 182, 3924–3928. doi: 10.1128/JB.182.14.3924-3928.2000

Conflict of Interest Statement: 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.

Check for
updates

Functionality of Two Origins of Replication in *Vibrio cholerae* Strains With a Single Chromosome

Matthias Bruhn[1], Daniel Schindler[2], Franziska S. Kemter[1], Michael R. Wiley[3],
Kitty Chase[3], Galina I. Koroleva[3], Gustavo Palacios[3], Shanmuga Sozhamannan[4,5]* and
Torsten Waldminghaus[1]*

[1] LOEWE Centre for Synthetic Microbiology-SYNMIKRO, Philipps-Universität Marburg, Marburg, Germany, [2] Manchester
Institute of Biotechnology, The University of Manchester, Manchester, United Kingdom, [3] United States Army Medical
Research Institute of Infectious Diseases, Frederick, MD, United States, [4] Defense Biological Product Assurance Office,
Frederick, MD, United States, [5] The Tauri Group, LLC, Alexandria, VA, United States

Chromosomal inheritance in bacteria usually entails bidirectional replication of a single
chromosome from a single origin into two copies and subsequent partitioning of one
copy each into daughter cells upon cell division. However, the human pathogen *Vibrio
cholerae* and other *Vibrionaceae* harbor two chromosomes, a large Chr1 and a small
Chr2. Chr1 and Chr2 have different origins, an *oriC*-type origin and a P1 plasmid-type
origin, respectively, driving the replication of respective chromosomes. Recently, we
described naturally occurring exceptions to the two-chromosome rule of *Vibrionaceae*:
i.e., Chr1 and Chr2 fused single chromosome *V. cholerae* strains, NSCV1 and NSCV2,
in which both origins of replication are present. Using NSCV1 and NSCV2, here we
tested whether two types of origins of replication can function simultaneously on the
same chromosome or one or the other origin is silenced. We found that in NSCV1, both
origins are active whereas in NSCV2 *ori2* is silenced despite the fact that it is functional
in an isolated context. The *ori2* activity appears to be primarily determined by the copy
number of the triggering site, *crtS,* which in turn is determined by its location with respect
to *ori1* and *ori2* on the fused chromosome.

**Keywords: DNA replication, secondary chromosome, plasmid, multipartite genome, replication initiation,
pathogens, cholera**

INTRODUCTION

The generally accepted paradigm of chromosome replication in bacteria is elucidated in *Escherichia
coli*. Replication is initiated at a unique singular sequence, the origin of replication (*oriC*) by DnaA,
proceeds bidirectionally along the chromosome and ends at the terminus diametrically opposite
to *oriC* on the circular chromosome. In *E. coli* and related bacteria, immediate re-initiation of
chromosome replication is hindered due to the hemi-methylated status of the sister chromosomes
and sequestration of *oriC* by SeqA which has a high binding affinity to hemimethylated *ori*
sequences (Lu et al., 1994; Slater et al., 1995; Waldminghaus and Skarstad, 2009). Most bacteria
have single chromosomes and follow this general replication paradigm. However, about 10% of
bacterial species have more than one chromosome and exhibit some deviation from this norm
(Fournes et al., 2018). Among these, *Vibrio cholerae* with chromosome 1 (Chr1, ~3 Mbps) and

chromosome 2(Chr2, ∼1 Mbps) has served as a model system for studies pertaining to multi-chromosome replication mechanisms, and in recent years, an extensive body of information has been accumulated on various aspects of Chr1 and Chr2 replication (Egan et al., 2005; Jha et al., 2012; Val et al., 2014b; Espinosa et al., 2017; Ramachandran et al., 2017).

Chr1 in *V. cholerae* is similar to the *E. coli* chromosome in that the replication follows the same pattern: replication origin, *ori1*, contains multiple DnaA boxes, which are bound by DnaA that unwinds the DNA and initiate replication (Duigou et al., 2006). The similarity is so striking that *V. cholerae ori1* can functionally substitute the *E. coli* replication origin *oriC* (Egan and Waldor, 2003; Demarre and Chattoraj, 2010; Koch et al., 2010; Kamp et al., 2013).

In contrast, the *V. cholerae* Chr2 appears to have an origin that resembles those of low copy number plasmids such as P1 and F (Fournes et al., 2018). The *ori2* contains an array of repeats (iterons) where the Chr2 specific initiator protein, RctB, binds and unwinds the DNA for *ori2* firing (Egan and Waldor, 2003; Duigou et al., 2008) but also exerts a form of negative regulation, termed 'handcuffing,' originally discovered in plasmids (Venkova-Canova and Chattoraj, 2011). Although *ori2* has plasmid-like features, Chr2 resembles typical chromosomes in some respects: (1) Participation of SeqA and Dam in regulation of *ori2* (Saint-Dic et al., 2008; Demarre and Chattoraj, 2010; Koch et al., 2010; Stokke et al., 2011). (2) Indispensability of Chr2, unlike plasmids, for cell survival because it harbors essential genes (Heidelberg et al., 2000; Kamp et al., 2013). (3) High level of coordination of replication between Chr1 and Chr2 in order to prevent over replication of Chr2 and ensure a guaranteed inheritance of a single copy of both chromosomes (Baek and Chattoraj, 2014; Val et al., 2016; Ramachandran et al., 2018). This raises the question on how coordination between Chr1 and Chr2 with respect to their timing of replication initiation is achieved given the disparity in their sizes and mechanisms of replication.

Chr1 replication is initiated at the onset of the replication period while initiation of Chr2 is delayed and occurs only when 2/3rd of Chr1 replication has been completed. Since Chr2 is 1/3rd the size of Chr1, both chromosomes consequently terminate their replication roughly at the same time (Rasmussen et al., 2007; Stokke et al., 2011). This termination synchrony appears not to be accidental but is selected for during evolution and is conserved within *Vibrionaceae* despite differing ratios of chromosome sizes (Kemter et al., 2018). This synchrony occurs through the *crtS* (*Chr2 replication triggering Site*) present on Chr1 that positively regulates *ori2* initiation (Val et al., 2016). Translocation of the *crtS* locus on Chr1, either closer to *ori1* or farther away, resulted in a corresponding shift in Chr2 initiation time as revealed by marker frequency analysis (Val et al., 2016), indicating that the native position of *crtS* sets the timing of Chr2 replication initiation such that its replication terminates synchronously with Chr1 (Kemter et al., 2018). The exact mechanism of *crtS* action remains to be elucidated but may include physical contacts between *crtS* and *ori2* as well as sequestration of Chr2 replication initiator protein, RctB (Baek and Chattoraj, 2014; Val et al., 2016). Recently, the global transcription factor Lrp was shown to bind to the *crtS* site and to facilitate RctB binding (Ciaccia et al., 2018).

Heterologous *E. coli* systems have been established based on *ori2* mini-chromosomes demonstrating that *crtS* provided *in trans* increases mini chromosome copy number indicating a positive role played by *crtS* in *ori2* firing (Baek and Chattoraj, 2014; Schallopp et al., 2017; de Lemos Martins et al., 2018). Recently, it was demonstrated that the copy number of *crtS* rather than the act of replicating the *crtS* is critical in the triggering of *ori2* firing (de Lemos Martins et al., 2018; Ramachandran et al., 2018).

In order to assess the differential genetic requirements of Chr1 and Chr2 replication, an artificial single chromosome *V. cholerae* strain has been created by genetic engineering in which *ori1* drives the replication of the fused chromosome (Val et al., 2012). In this strain, designated MCH1, the sequences to the left and right of *ori2* were fused to the terminus of Chr1. In this arrangement the direction in which chromosome arms are replicated is conserved to minimize conflicts between DNA replication and transcription. The MCH1 strain was instrumental in establishing the essentiality of Dam methyltransferase in *V. cholerae* because of its role in *ori2* function which was first shown by Demarre and Chattoraj (2010). This conclusion was further supported by the finding that depletion of Dam leads to spontaneous chromosomal fusion (Val et al., 2014a). In this case, the entire genome is replicated from *ori1* which can tolerate the absence of Dam (Val et al., 2014a).
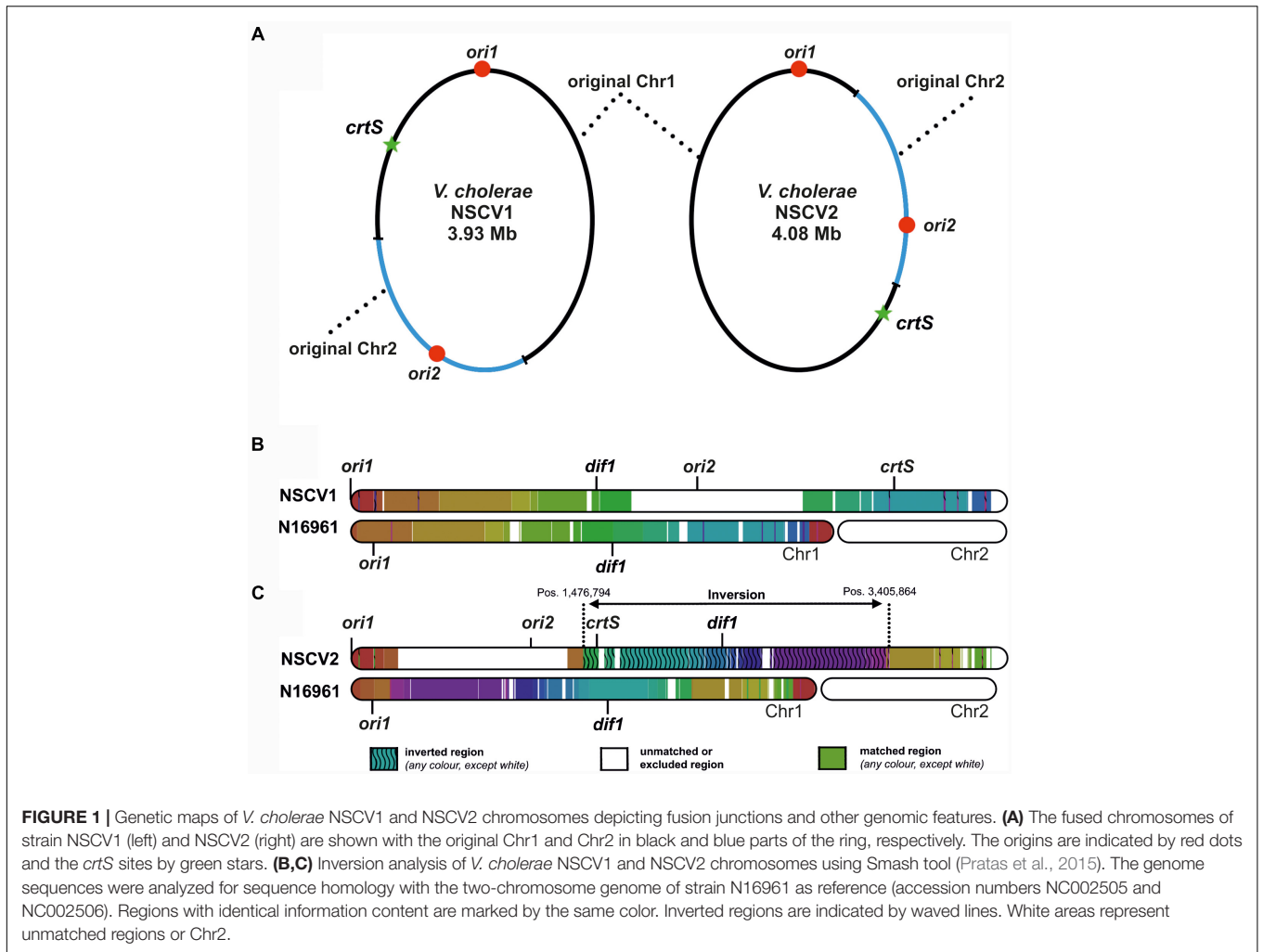
Recently, we described two naturally occurring *V. cholerae* strains (NSCV1 and NSCV2) containing both *ori1* and *ori2* on the same chromosome (Chapman et al., 2015; Xie et al., 2017). In these strains, Chr1 and Chr2 are fused at two different locations (**Figure 1**). The locations of relevant features such as *ori1*, *ori2*, and *crtS* sites are indicated in **Figure 1A**. In NSCV1, *crtS* is located about 670 kbs away from *ori1* (**Figure 1**). It is similar to the standard two-chromosome reference strain N16961 with respect to distance, where *crtS* is located 695 kbs away from *ori1*. In NSCV2, the distance between *ori1* and *crtS* is 1,566 kbs due to a large inversion that has occurred around the terminus region of the chromosome (**Figure 1C**).

This genomic architecture raised interesting questions about whether both origins are functional in the same cell or one or the other *ori* is silenced since in principle a single origin should suffice to replicate the fused chromosome. The strains also allowed us to ask if the chromosomal fusions are maintained without resorting to genome splitting which is the predominant genome configuration in *Vibrionaceae*. We found that in NSCV1 both origins are active in the same cell whereas in NSCV2 *ori2* appears to be silent. Further, these chromosomes appear to be in a locked configuration since even after prolonged continuous growth they remain fused without splitting into two.
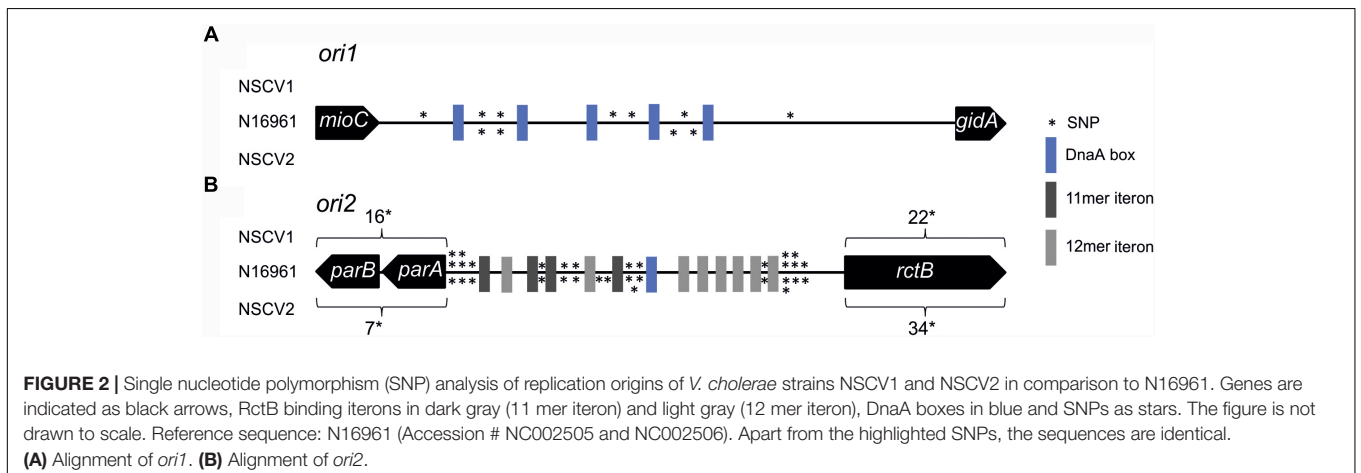
RESULTS

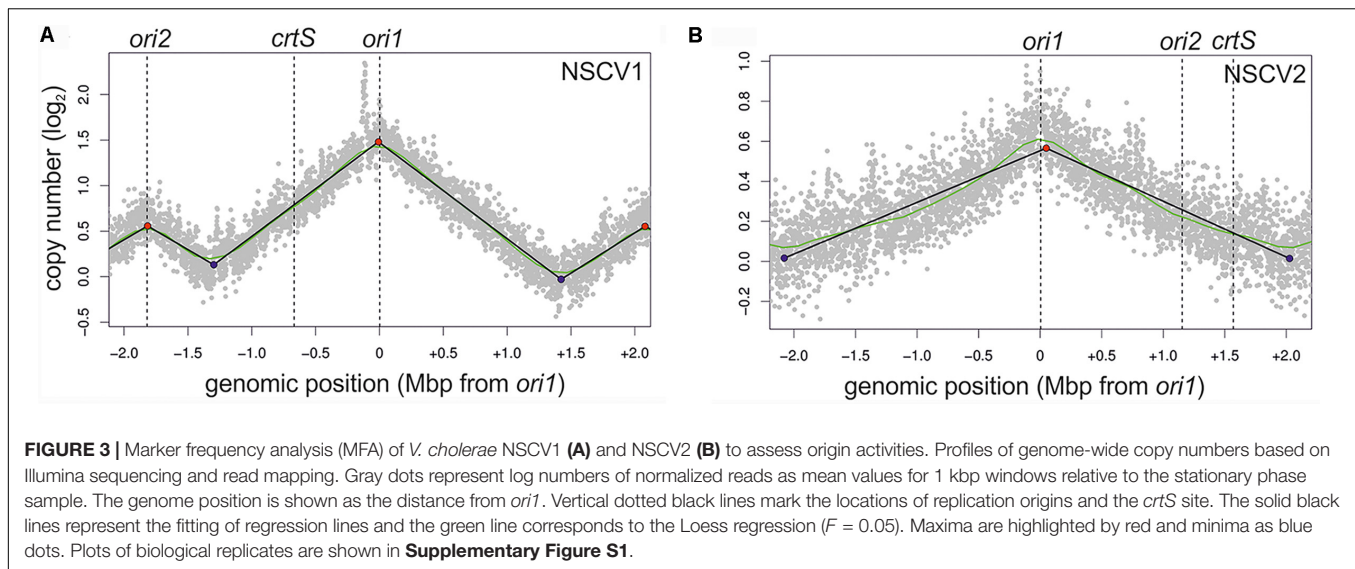## Activity of *ori1* and *ori2* in *V. cholerae* NSCV Strains

In general, the replication in bacteria relies on one origin of replication for one replicon. Fusion of two replicons, such as seen in NSCV1 and NSCV2, would initially give rise to a chromosome with two functional replication origins where one could be

FIGURE 1 | Genetic maps of *V. cholerae* NSCV1 and NSCV2 chromosomes depicting fusion junctions and other genomic features. **(A)** The fused chromosomes of strain NSCV1 (left) and NSCV2 (right) are shown with the original Chr1 and Chr2 in black and blue parts of the ring, respectively. The origins are indicated by red dots and the *crtS* sites by green stars. **(B,C)** Inversion analysis of *V. cholerae* NSCV1 and NSCV2 chromosomes using Smash tool (Pratas et al., 2015). The genome sequences were analyzed for sequence homology with the two-chromosome genome of strain N16961 as reference (accession numbers NC002505 and NC002506). Regions with identical information content are marked by the same color. Inverted regions are indicated by waved lines. White areas represent unmatched regions or Chr2.

superfluous. This raises the question of whether both, *ori1* and *ori2* are active in the fused chromosomes of NSCV1 and NSCV2, or only one of the origins is active and the other is silent. Inspection of the sequence of the origins and the replication initiator genes revealed no obvious mutational changes that could indicate non-functionality of one or the other of the origins in the two NSCV strains (**Figure 2**) (Xie et al., 2017). Compared to strain N16961, both strains possess complete *ori1* sequences,

FIGURE 2 | Single nucleotide polymorphism (SNP) analysis of replication origins of *V. cholerae* strains NSCV1 and NSCV2 in comparison to N16961. Genes are indicated as black arrows, RctB binding iterons in dark gray (11 mer iteron) and light gray (12 mer iteron), DnaA boxes in blue and SNPs as stars. The figure is not drawn to scale. Reference sequence: N16961 (Accession # NC002505 and NC002506). Apart from the highlighted SNPs, the sequences are identical. **(A)** Alignment of *ori1*. **(B)** Alignment of *ori2*.

FIGURE 3 | Marker frequency analysis (MFA) of *V. cholerae* NSCV1 **(A)** and NSCV2 **(B)** to assess origin activities. Profiles of genome-wide copy numbers based on Illumina sequencing and read mapping. Gray dots represent log numbers of normalized reads as mean values for 1 kbp windows relative to the stationary phase sample. The genome position is shown as the distance from *ori1*. Vertical dotted black lines mark the locations of replication origins and the *crtS* site. The solid black lines represent the fitting of regression lines and the green line corresponds to the Loess regression (*F* = 0.05). Maxima are highlighted by red and minima as blue dots. Plots of biological replicates are shown in **Supplementary Figure S1**.

with seven SNPs (NSCV1) and four SNPs (NSCV2) spanning the 474 bps long *gidA-mioC* intergenic region and none of the DnaA boxes were affected by mutations. The 5,656 bps long *ori2* regions (including genes *parB*, *parA,* and *rctB*) are also intact, with 54 SNPs in NSCV1 and 57 SNPs in NSCV2. Notably, the RctB-binding iteron sequences are not affected by mutations.

To test the activity of replication origins experimentally, we carried out marker frequency analyses (MFA). It is known that actively growing cells have a higher copy number of *ori* proximal sequences compared to *ori* distal/*ter* proximal sequences. We employed next generation sequencing technology to obtain whole genome sequences of DNAs isolated from logarithmic and stationary phase cultures of NSCV1 and NSCV2 and analyzed the sequence data for marker frequency. Read data from stationary phase DNAs were used for normalization and read mapping plots were created with *ori1* repositioned at the center of the plot to represent bidirectional replication as well as for easy visualization of *ori* activity (**Figure 3**). Both NSCV1 and NSCV2 exhibited a maximum copy number of reads close to *ori1* and a decreasing gradient on either side of *ori1* moving toward the terminus creating a tent-shape (Skovgaard et al., 2011), indicating an active *ori1* and bidirectional replication in both strains. In strain NSCV1, a local higher marker frequency was observed at the *ori2* position, indicating that *ori2* also is active in this strain. In contrast, no such local peak in marker frequency was found at the *ori2* position in NSCV2 consistent with a silent *ori2*. In addition, an almost 3X lower *ori1/ter* ratio was observed in NSCV2 compared to NSCV1 indicating less or no overlap of replication cycles. The rationale for this interpretation is as follows: On a replicon with overlapping replication, the *ori/ter* ratio would be four since replication is initiated twice before termination can occur. A replicon that initiates only once would consequently have an *ori/ter* ratio of two. Considering that the culture represents a mixed population of cells before and after termination an *ori/ter* ratio of higher than two, as in the case of NSCV1, indicates overlapping replication cycles while values less than two indicate no overlap of replication cycles.

NSCV1 and NSCV2 Carry a Functional *ori2*

The apparent inactivity of *ori2* in strain NSCV2 raises the question of whether this origin of replication is functional but silenced or non-functional. To answer this question, we cloned the origins into a mini replicon and assessed independent replication in a plasmid backbone. Plasmid pMA135 carries *oriR6K* and can replicate conditionally in *E. coli* strains that provide *in trans*, the replication initiator protein, Pir, from a lambda prophage. In addition pMA135 can be transferred by conjugation from a donor to a recipient (Messerschmidt et al., 2015). In the absence of λ*pir*, no exconjugants are obtained unless replication is driven by another fully functional origin of replication. The *ori2* fragments of NSCV1 and NSCV2, including the core *ori2* plus the genes *rctB* and *parAB2*, were cloned into pMA135 independently and the number of exconjugants was enumerated in an *E. coli* strain that does not contain λ*pir*. The *ori2* minichromosome constructs from all three strains (N16961, NSCV1, and NSCV2) yielded exconjugants at a frequency of about 1% (10^{-2}) of the recipients, as did the positive control F plasmid *ori* (**Table 1**). Minichromosomes based on *oriII* of strain N16961 have been shown not to integrate into the *E. coli* chromosome (Messerschmidt et al.,

TABLE 1 | Conjugation efficiencies of replicons with NSCV *ori2* replication origins.

Replicon[†]	Origin of replication	Conjugation[‡] efficiency
pMA568	N16961*ori2*	5.5×10^{-3}
pMA739	NSCV1 *ori2*	1.7×10^{-2}
pMA755	NSCV2 *ori2*	4.3×10^{-2}
pMA899	F plasmid origin	3.2×10^{-2}
pMA135	*oriR6K* only	0

[†]*All replicons possess an oriR6K and conjugated from E. coli strain WM3064 to MG1655.* [‡]*Efficiency of one representative experiment is given as ratio of conjugant CFU over total recipient CFU.*

2016). To test if the replicons based on NSCV *oriII* are also replicating autonomously, we performed a plasmid isolation procedure for five individual clones for each of the two tested NSCV replicons. In all cases, we were able to isolate the corresponding minichromosomes as evidenced by agarose gel electrophoresis (data not shown), verifying autonomous replication without integration into the primary chromosome. The *oriR6K* replicon by itself did not yield any exconjugant. We conclude that both NSCV1 and NSCV2 carry a functional *ori2*.

Genetic Stability of Chromosome Fusions in NSCV1 and NSCV2

An active *ori2* in strain NSCV1 could potentially allow the chromosome fusion to be reversed. Similarly, the silent *ori2* in NSCV2 might be inactive only in the context of a fused chromosome; in either case, splitting of Chr1 and Chr2 is conceivable. It was observed that *V. cholerae* with an artificially fused chromosome grew slower compared to the two-chromosome parental strain indicating a negative fitness burden on the bacterium (Val et al., 2012). Similarly, we observed an increased doubling time of strains NSCV1 (20 ± 0.5 min) and NSCV2 (29 ± 1.3 min) compared to the two-chromosome strain N16961 (16 ± 0.2 min) (**Supplementary Table S4**). In addition, microscopic examination showed that strain NSCV2 cells exhibited a distinct phenotype. While the two-chromosome reference strain N16961 and NSCV1 cells are comma-shaped as typical for *V. cholerae*, NSCV2 cells are much more curled and occasionally S-shaped (**Figure 4**). It is unknown whether this phenotype is related to the chromosome fusion.
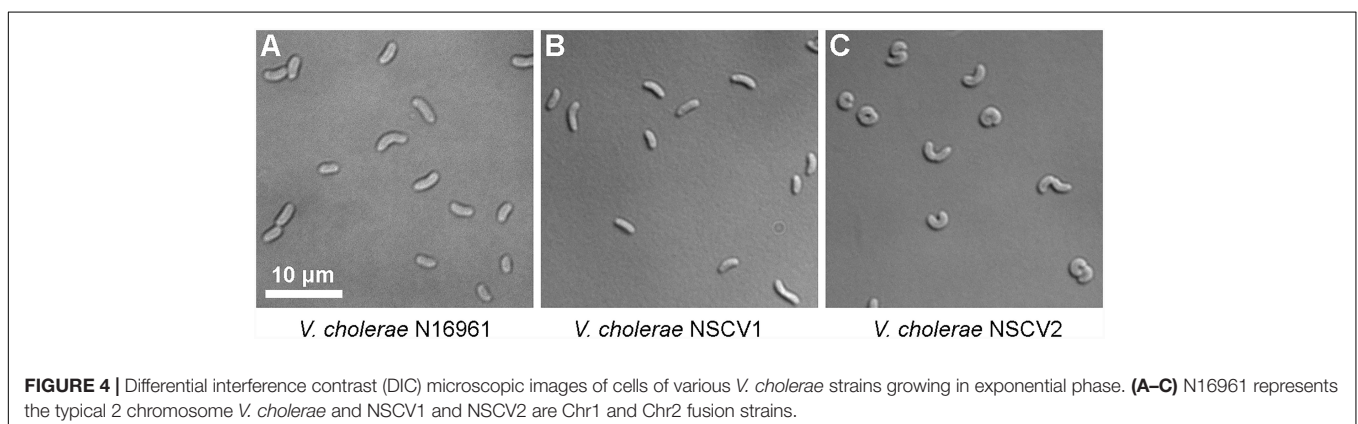
It is conceivable that the slower growth of the fused chromosome strains may have negative fitness value leading to instability, thus promoting genome splitting. On the other hand, if the fused chromosome were under a positive selection pressure to remain fused even at a greater cost in terms of slower growth rate or if they are locked in the fusion configuration due to genetic defects that prevent splitting, then the fusion would be stable even after long term continuous culturing. This led us to test whether the two NSCV strains potentially could revert back to a two-chromosome arrangement

upon prolonged continuous culturing. If splitting of the fused chromosomes occurs and if this splitting leads to a fitness advantage due to increased growth rates one would expect two-chromosome clones to appear in a population of NSCV1 and NSCV2 and these clones could potentially replace the fused chromosome cells after prolonged growth by clonal expansion. To test this hypothesis, we cultured the two NSCV strains for 16 days in 100 ml of liquid medium with replenishment of fresh medium every 24 h resulting in approximately 160 generations of growth in total. To examine if the NSCV strains reverted back to a two-chromosome arrangement we isolated DNA from long-term grown cells and performed long-read DNA sequencing using the PacBio technology to be able to detect long range chromosomal rearrangements such as the chromosomal fusion junctions/or junctions (∼24–51 kbs) of the split chromosomes. A *de novo* genome assembly led to one single contig for both NSCV1 and NSCV2 reflecting the original one-chromosome configuration. In conclusion, the chromosome fusions in NSCV1 and NSCV2 appear to be stable and chromosome splitting is not a frequent event or the fused state is probably under positive selection pressure.

Replication Origins Are Active in an Engineered System Resembling the NSCV Arrangement

Why is *ori2* of strain NSCV2 silenced while it is active in NSCV1? One obvious difference between the two strains is the differential positioning of the two replication origins *ori1* and *ori2* to one another. While the distance between *ori1* and *ori2* is 1.828 Mbps in NSCV1 it is only 1.118 Mbps in the genome of NSCV2.

To test experimentally, if the *ori* positioning of NSCV1 and NSCV2 influences the initiation outcome, we re-constructed the *ori1* to *ori2* arrangement in a genetically accessible system since NSCV1 and NSCV2 are recalcitrant for genetic manipulations. To accomplish this, we transferred a functional *hapR* to *V. cholerae* strain MCH1 to render it naturally competent (Lo Scrudato and Blokesch, 2012, 2013). Strain MCH1 was derived from the prototype *V. cholerae* strain N16961 by fusion of the two chromosomes with a deletion of *ori2* (Val et al., 2012). We inserted a copy of *ori2* including the flanking genes
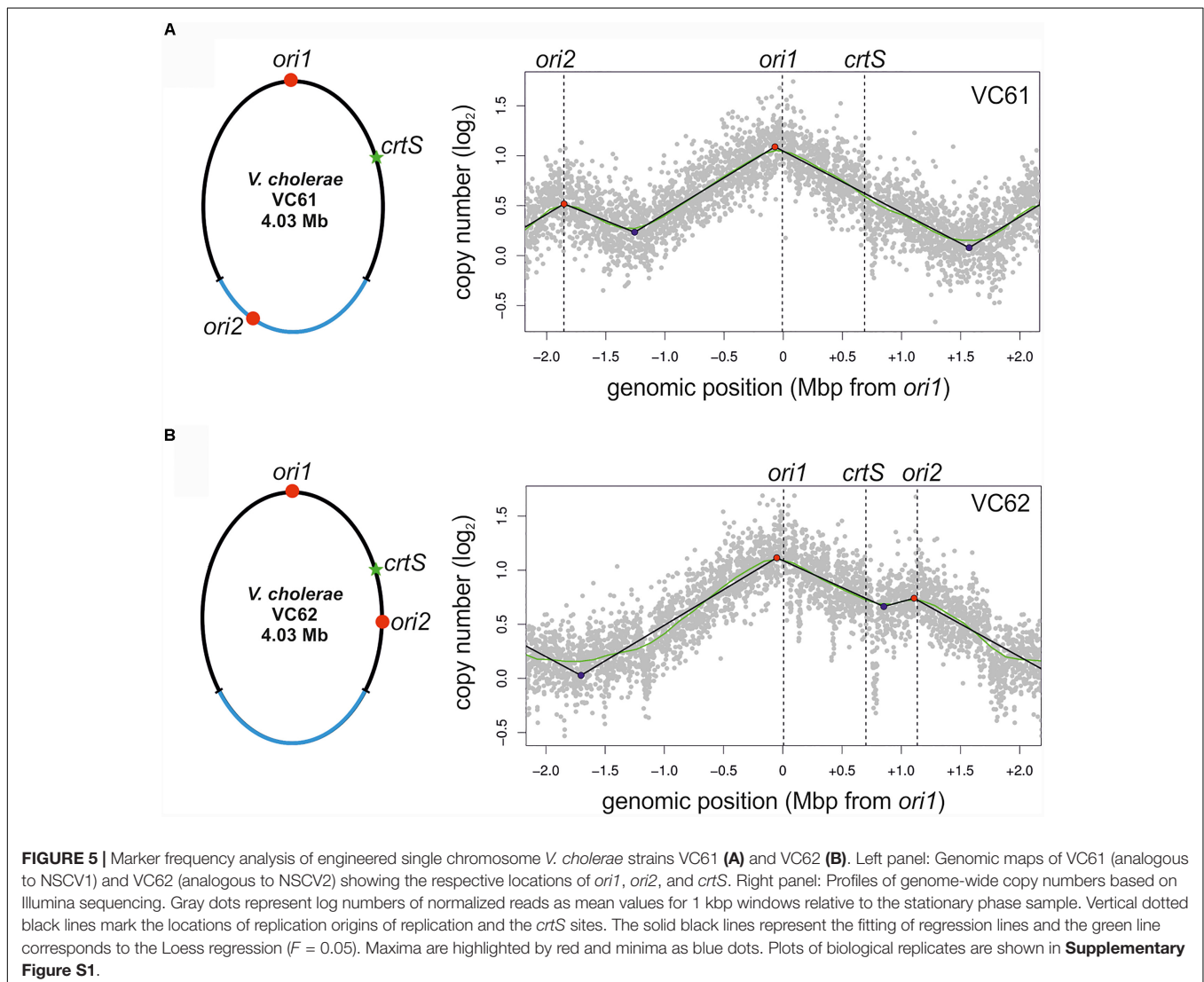

FIGURE 4 | Differential interference contrast (DIC) microscopic images of cells of various *V. cholerae* strains growing in exponential phase. **(A–C)** N16961 represents the typical 2 chromosome *V. cholerae* and NSCV1 and NSCV2 are Chr1 and Chr2 fusion strains.

parAB and *rctB* into MCH1 at positions analogous to NSCV1 and NSCV2 with respect to the distance from *ori1* giving rise to strains VC61 and VC62, respectively, and performed marker frequency analysis of exponentially grown cultures (**Figure 5**).

The MFA plot of strain VC61 resembled the pattern seen in NSCV1 (**Figure 5A**, compared to **Figure 3A**). In contrast, a clear local peak of marker frequency was seen at the *ori2* position of strain VC62 (**Figure 5B**). Thus, even though the origin arrangement is similar in NSCV2 and VC62, the *ori2* copy appears to be active only in the engineered strain VC62 and not in NSCV2. To test if the *ori2* positioning at the NSCV2 position has more negative effect on growth compared to the *ori2* insertion at the NSCV1 position we measured the doubling times of strains VC61 (26 ± 0.1 min) and VC62 (27 ± 0.5 min) (**Supplementary Table S4**). The difference in doubling time was only marginally affected suggesting that the differences in DNA replication between the two strains do not result in severe impairment in growth. We conclude that an *ori2* insertion at positions analogous to those in the NSCV strains relative to *ori1* can be active as determined by MFA.

Dam Methyltransferase Is Functional in NSCV Strains

We have shown above that *ori2* in strain NSCV2 is functional but appears to be not active. One potential cause of *ori2* silencing could be an inactive Dam methylation system. Methylation of the adenine within the Dam recognition site 'GATC' present at *ori2* locus is a prerequisite for *ori2* activity (Demarre and Chattoraj, 2010). Sequence analyses showed intact *dam* genes in both strains (Xie et al., 2017). The methylation status of GATC sites can be analyzed by the differential sensitivities of the genomic DNA to *Dpn*I (cleaves only methylated/hemimethylated GATC sites), *Dpn*II (cleaves only unmethylated GATC sites) and *Sau*3A1 (cleaves both unmethylated and methylated GATC sites) restriction enzymes.

FIGURE 5 | Marker frequency analysis of engineered single chromosome *V. cholerae* strains VC61 **(A)** and VC62 **(B)**. Left panel: Genomic maps of VC61 (analogous to NSCV1) and VC62 (analogous to NSCV2) showing the respective locations of *ori1*, *ori2*, and *crtS*. Right panel: Profiles of genome-wide copy numbers based on Illumina sequencing. Gray dots represent log numbers of normalized reads as mean values for 1 kbp windows relative to the stationary phase sample. Vertical dotted black lines mark the locations of replication origins of replication and the *crtS* sites. The solid black lines represent the fitting of regression lines and the green line corresponds to the Loess regression (F = 0.05). Maxima are highlighted by red and minima as blue dots. Plots of biological replicates are shown in **Supplementary Figure S1**.

Genomic DNAs of NSCV1 and NSCV2 were sensitive to *Dpn*I and *Sau*3A1 and resistant to *Dpn*II, indicative of full methylation at GATC sites (**Figure 6**). This result was further confirmed by the PacBio sequence data. In PacBio SMRT (Single Molecule Real Time) sequencing, presence of modified base (A in GATC) in the DNA template results in a delayed incorporation of the corresponding T nucleotide, i.e., longer inter-pulse duration (IPD) compared to template lacking the modification (Flusberg et al., 2010). These kinetic measurements create specific signatures for different types of base modifications. Analyses of the PacBio sequence data for modified bases indicated that both NSCV1 (38571/38572 sites) and NSCV2 (37573/37590) have fully methylated (>99.99%) GATC sites. We conclude that Dam is functional in both NSCV strains and a lack of methylation cannot explain the inactivity of *ori2* in NSCV2.

crtS Sites of NSCV Strains Are Functional

Another possibility for *ori2* silencing in NSCV2 is through alterations of *crtS* activity. This short DNA sequence is found on Chr1 in all available whole genome sequences of two-chromosome strains of *Vibrionaceae* and its replication triggers *ori2* firing on Chr2 (Baek and Chattoraj, 2014; Val et al., 2016; Kemter et al., 2018). Consequently, a non-functional *crtS* could lead to an inactive *ori2*. To decipher the functionality of *crtS* sites in NSCV1 and NSCV2

FIGURE 6 | Testing Dam methylation sensitivities of genomic DNAs of strains *V. cholerae* NSCV1 and NSCV2. Restriction digestion of 1 μg of genomic DNA using the indicated enzymes was carried out and the various strains are as follows: 1. *E. coli* MG1655; 2. *E. coli* MG1655Δ*dam*; 3 and 4. *V. cholerae* NSCV1; 5 and 6. *V. cholerae* NSCV2. *Dpn*I cleaves methylated GATC sequences. *Dpn*II cleaves only unmethylated GATC sequences. *Sau*3AI cleaves GATCs independent of the methylation state. DNA cleavage is evident from the disappearance of high molecular weight band and appearance of low molecular weight streak of DNA.

in silico, we extracted the respective sequences and aligned them to the consensus sequence we established recently (**Figure 7A**) (Kemter et al., 2018). Notably, all highly conserved parts of the *crtS* sequence are also conserved in the *crtS* sequences of NSCV1 and NSCV2 (**Figure 7A**). To test the functionality of *crtS* in the context of a fused chromosome experimentally, we deleted the *crtS* in strain VC62 (resulting in strain VC71) and performed marker frequency analysis (**Figure 7B**).

As expected, no local peak of copy number increase was seen at the *ori2* position of the Δ*crtS* strain VC71 in contrast to the MFA of the parental strain VC62 (compare **Figures 5B,7B**) confirming the necessity of a functional *crtS* site for *ori2* firing. Incidentally, the doubling time of this strain is not much different (25 ± 0.6 min) from that of VC61 and VC62 (**Supplementary Table S4**). To analyze the functionality of the *crtS* sites from NSCV1 and NSCV2, we inserted these sequences into the genome of *E. coli* and transformed the corresponding strains with the *ori2*-based minichromosome synVicII. The rationale behind this approach is the observation of a copy number increase of an *ori2*-based minichromosome in *E. coli* strains carrying a functional *crtS* site (Baek and Chattoraj, 2014; Kemter et al., 2018). The copy number of *ori2* minichromosomes was measured in *E. coli* strains either carrying the *crtS* sites of strain N16961, NSCV1, NSCV2 or no *crtS* (**Figure 7C**). In this case, we tested the resistance level of the respective strains to ampicillin as an indirect measure of the minichromosome copy number as described previously (Schallopp et al., 2017). The *crtS* sites of both NSCV strains increased the copy number of the *ori2* minichromosome compared to the strain lacking any *crtS* site and to a similar extent as the *crtS* site from the two-chromosome *V. cholerae* strain N16961 used as positive control (**Figure 7C**). We conclude that *crtS* sites in NSCV1 and NSCV2 are functional in a heterologous system and a defective *crtS* might not explain the silent *ori2* in strain NSCV2.

Relative ori1, ori2, and crtS Locations Determine ori2 Activity in NSCV Strains

In a canonical two-chromosome *V. cholerae*, replication of the *crtS* on Chr1 triggers the initiation at *ori2* on Chr2. This scenario can in principle be analogous in a fused chromosome such as NSCV1 in which *ori1*, *ori2* and *crtS* are present on the same molecule. Here the *crtS* is located about 677 kbps away from *ori1* and is replicated first with replication forks originating at *ori1* which subsequently triggers *ori2* firing, which is further downstream of *crtS*. The organization is different in strain NSCV2 where the *crtS* lies about 416 kbps downstream of *ori2*. Thus, in NSCV2, *ori2* precedes *crtS*. This arrangement is the consequence of a large inversion of the Chr1 part of the genome (**Figure 1**). The replicative outcome of this arrangement is difficult to predict considering the known triggering role of *crtS*; i.e., *ori2* being replicated first by replication forks originating at *ori1*, subsequently upon duplication of the *crtS* site. In other words, *crtS* can also retrospectively trigger replication of already replicated *ori2* that *crtS* replication is supposed to trigger in the
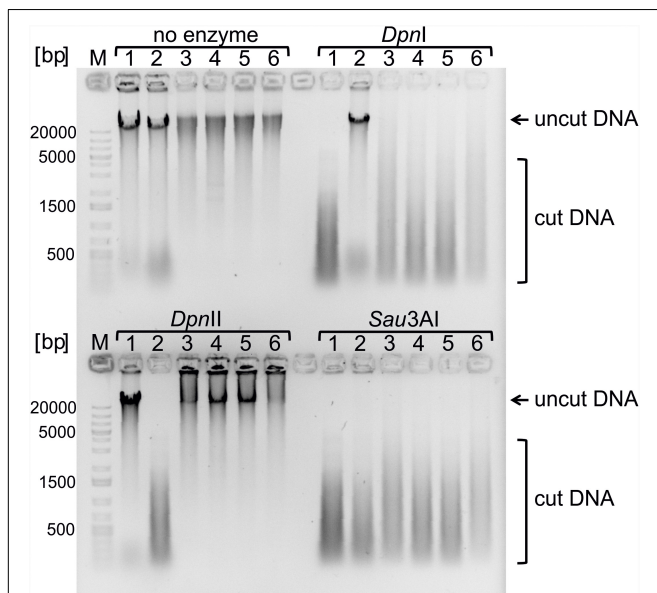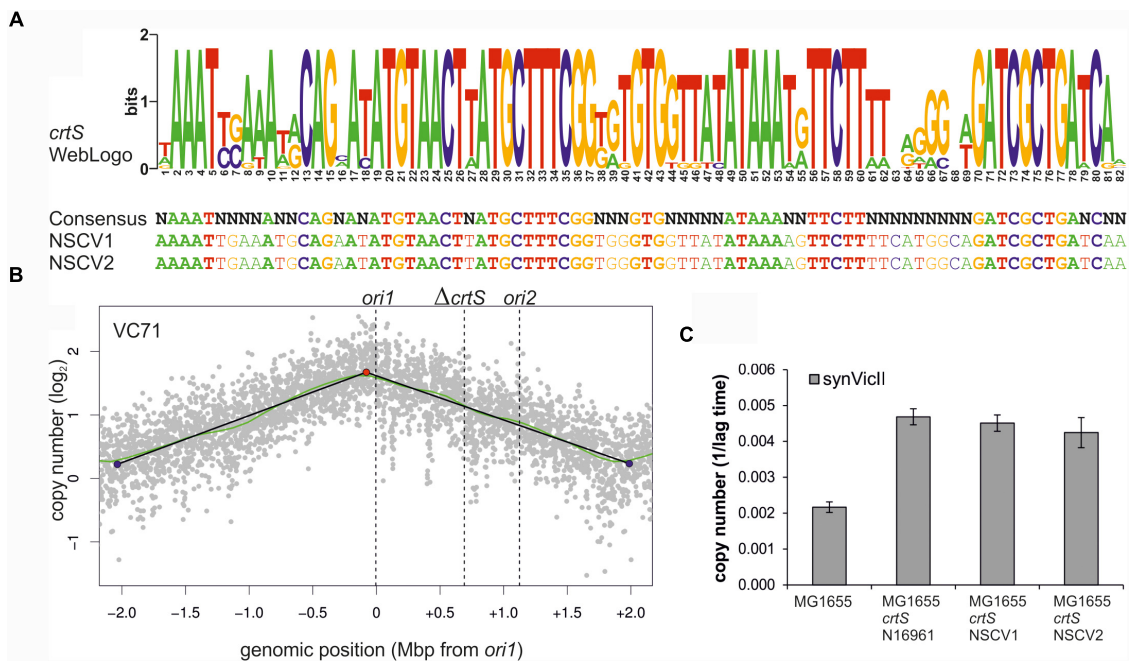
FIGURE 7 | *crtS* sites of NSCV strains are functional. **(A)** An alignment of *crtS* sites from 13 sequenced *Vibrio* genomes was used to calculate a WebLogo (top panel). The height of the nucleotides in each position represents the measure of conservation. All positions with 100% sequence conservation are conserved in the *crtS* sequences of NSCV1 and NSCV2 as well (indicated in bold, lower panel) (Kemter et al., 2018). **(B)** Marker frequency analysis of *V. cholerae* strain VC71 lacking a functional *crtS* site. Gray dots represent log numbers of normalized reads as mean values for 1 kbp windows relative to the stationary phase sample. The genome position is shown as the distance from *ori1*. Vertical dotted black lines mark the locations of replication origins of replication and the *crtS* sites. The solid black lines represent the fitting of regression lines and the green line corresponds to the Loess regression ($F = 0.05$). Maxima are highlighted by red and minima as blue dots. Plots of biological replicates are shown in **Supplementary Figure S1**. **(C)** *crtS* sites of NSCV1 and NSCV2 increase the copy number of *ori2*-based mini-chromosomes in *E. coli*. *E. coli* strains harboring the *ori2*-based mini-chromosome synVicII and chromosomal insertions of *crtS* from different *V. cholerae* strains as indicated were grown in LB medium with 500 µg/ml ampicillin in a 96-well plate at 37°C. As strains with a lower replicon copy number have a longer lag period to initiate growth, the value 1 divided by the time to reach an $OD_{600} \geq 0.1$ was used as measure for the replicon copy number. Values are the mean of three biological replicates with indicated standard deviation. Growing the strains in standard concentrations of ampicillin did not show any difference between *wild type* and *crtS* carrying strains as expected (data not shown).

first place. To test this scenario, we moved the *crtS* site in strain VC62 to the position analogous to strain NSCV2 (resulting in strain VC73). VC73 did not exhibit any significant difference in doubling time compared to other strains (26 ± 0.2 min) (**Supplementary Table S4**). Respective MFA analysis revealed an active *ori2,* suggesting that an *ori2* copy on a fused chromosome can be triggered by a *crtS* site located downstream relative to *ori2*, contrary to what was observed in NSCV2 (compare **Figures 3B– 8**). The peak at the *ori2* position is not very strong but it is important to note that the fitting of regression lines is automated based on the maxima found within the mapped reads (Kemter et al., 2018). There is no pre-selection of *ori* positions and the fact that the peak is detected at the *ori2* position by the computational fitting model implies that it is an active origin as also seen in the biological replicate (**Supplementary Figure S1F**).

DISCUSSION

Vibrio cholerae has partitioned its genome between a true bacterial chromosome and a "domesticated" plasmid replicon

(Heidelberg et al., 2000; Venkova-Canova and Chattoraj, 2011; Val et al., 2014b). Unlike most bacteria, this two-replicon arrangement is conserved within the family of *Vibrionaceae* (Okada et al., 2005; Val et al., 2014b; Ramachandran et al., 2017). It was postulated that the bipartite genome in *Vibrio* species enables varying the copy numbers of both chromosomes in a niche-specific manner under certain environmental conditions as an adaptation strategy (Heidelberg et al., 2000; Schoolnik and Yildiz, 2000; Srivastava and Chattoraj, 2007). Alternatively, the two-chromosome setup in *Vibrio* species can be considered as an adaptive feature that enables rapid genome duplication if multiple chromosomes are replicated simultaneously. The latter assumption fits with the observation that *Vibrio* species is one of the fastest growing bacteria known, a feature that has led to a recent proposal for using the non-pathogenic *V. natriegens* as the new workhorse for bioengineering (Weinstock et al., 2016; Dalia et al., 2017; Hoffart et al., 2017). However, it brings up the question of how the two replicons are coordinating their replication and regulate over or under replication in order to ensure inheritance of the just one copy of the full complement of the genome into daughter cells upon cell division. Furthermore, the evolutionary driving

force (or forces) that led to the two-chromosome setup in *Vibrionaceae* remain speculative. One way to better understand the evolutionary significance of this unique feature of *Vibrio* species is to study naturally occurring exceptions to the two chromosome rule: strains which have evolved into a single chromosome by chromosomal fusion. We have previously identified two Natural Single Chromosome *Vibrio* (NSCV) *cholerae* strains (Chapman et al., 2015; Johnson et al., 2015; Xie et al., 2017) and in this study we addressed the functionality and activity of the two origins of replication on the same chromosome.

Fortuitously, the two NSCV strains have different fusion junctions and other genetic rearrangements that provided an opportunity to compare and contrast their *ori* functionality. We used marker frequency analysis as an indirect measure of *ori* activity and found *ori1* to be active in both strains. In strain NSCV1, we detected an additional, replication activity peak at the *ori2* locus indicating that this second origin is active as well. Since the MFA experiments are population based, we cannot exclude that some cells use *ori1* only and other cells *ori2* only to replicate the fused chromosome. However, we consider it more likely that both origins are firing within the same cell. Interestingly, the copy number ratio of *ori2/ter* was lower than for *ori1/ter* consistent with a time delayed initiation of *ori2* on the fused chromosome. Such an origin initiation differential has also been observed in conventional two-chromosome *Vibrio* strains resulting in synchronous termination of replication of the two chromosomes despite their different sizes (Rasmussen et al., 2007; Stokke et al., 2011; Kemter et al., 2018). Conservation of the orchestrated replication timing of the two origins in NSCV1 indicates that the single-chromosome setup does not lead to any origin interference. In two-chromosome *Vibrio* species, replication of the *crtS* site located on Chr1 has been shown to trigger *ori2* firing (Baek and Chattoraj, 2014; Val et al., 2016; Ramachandran et al., 2018). In *V. cholerae* N16961, the time between *crtS* replication and *ori2* initiation corresponds to the time the replication fork needs to replicate about 200 kbps of DNA (Val et al., 2016). Considering a similar delay in strain NSCV1, *ori2* would have fired long before replication fork originating at *ori1* reaches it since the distance between *crtS* to *ori2* is more than 1 Mbps. As a consequence, replication forks originating at *ori1* and *ori2* will meet at some position of the fused chromosome, the timing of which will dependent on *ori1* location and its regulation just as it occurs in two-chromosome *V. cholerae*.

The scenario is much different in NSCV2. Here, *ori2* is replicated before *crtS* by replication forks originating at *ori1* (**Figure 1A**). Intuitively, this would lead to a chaotic replication perturbance because *ori2* firing precedes *crtS* duplication. In the two-chromosome context, *crtS* duplication occurs first which then triggers *ori2* firing. Hence, if replication of the *crtS* triggers *ori2* initiation also in this genomic arrangement it would happen on two copies of the already replicated *ori2*. In addition, the replication forks coming from this newly initiated *ori2* copies have the potential to replicate the *crtS* in just a few minutes because of their close proximity and could potentially lead to additional rounds of *ori2* initiation

and thus an uncontrolled *ori2* firing. If this were to happen, not only replication control at *ori2* would be severely perturbed but also *ori1* functioning can be interfered because *ori1* might be replicated passively from replication forks, coming from *ori2*. Interestingly, what we observed in strain NSCV2 is not a replication out of control but instead a simple silencing of *ori2* activity. Our data clearly demonstrate that *ori2* of NSCV2 is functional in an isolated context as shown by its ability to drive replication of a mini-chromosome in *E. coli*. In addition, critical factors involved in regulation of *ori2* appear to be fully functional in NSCV2, namely the Dam methylation system and the *crtS* site. Paradoxically, *ori2* is not used to initiate DNA replication in the fused chromosome strain, NSCV2.

A recent study offers a simple explanation for our observation on *ori2* silencing in NSCV2 (Ramachandran et al., 2018). These authors showed that it is the doubling of the *crtS* dosage rather than the process of replicating the *crtS* that triggers *ori2* initiation which results in an even number of *crtS* and *ori2* copies. This tendency of the regulatory system to produce similar copy numbers of *crtS* and *ori2* have also been found in engineered systems with multiple *crtS* sites (de Lemos Martins et al., 2018). If the *crtS* site is duplicated before *ori2* as it is the case in two-chromosome *V. cholerae* strains, the *ori2* copy number/ter will be lower compared to *crtS* and consequently the initiation at *ori2* will be triggered to restore *crtS/ori2* copy number balance. If the *crtS* site is replicated after *ori2* has been copied by replication forks coming from *ori1* as in the case of NSCV2, there is no need to initiate at *ori2* because the copy number of *crtS* and *ori2* are in balance already. The lack of initiation at *ori2* in NSCV2 is therefore fully consistant with and confirmatory to the findings of the aforementioned studies (de Lemos Martins et al., 2018; Ramachandran et al., 2018). However, in contrast to the expectation of the *crtS*-to-*ori2* copy control model, we observed that *ori2* is active in the engineered strain VC73 in which the *crtS* lies downstream of *ori2* analogous to the arrangement in NSCV2 (**Figure 8**). However, it is important to note that although the positioning of *ori1*, *ori2* and *crtS* are similar in strains NSCV2 and VC73, the genomic contexts of *ori2* and *crtS* are entirely artificial compared to their native positions and furthermore, VC73 does not share the large chromosomal inversion found in NSCV2. More precisely, in VC73, *ori2* lies in a region of the original Chr1 and *crtS* within a region of the original Chr2 (**Figure 8**, left panel). The discrepancy in *ori2* firing between the naturally occurring NSCV2 and the engineered strain VC73 may therefore be explained by their different genomic context.

One interesting question that arises is if the two *V. cholerae* strains with fused chromosomes are evolutionarily stable. It could also be that the fusion occurred only transiently as an artifact of lab cultivation or during the original strain isolation from a patient sample. In fact, chromosome fusions in *V. cholerae* have been observed previously to occur frequently within a population but the two chromosome configuration seems to provide a selective advantage leading to rapid elimination of the fused chromosomes cells from the population (Val et al., 2014a). Based on this observation, we grew the NSCV strains for about 160 generations and expected a splitting of the

FIGURE 8 | Marker frequency analysis of *V. cholerae* strain VC73. Left panel: Genomic map of VC73 showing the locations of *ori1*, *ori2*, and *crtS*. Right panel: Profile of genome-wide copy numbers. Gray dots represent log numbers of normalized reads as mean values for 1 kbp windows relative to the stationary phase sample. The genome position is shown as the distance from *ori1*. Vertical dotted black lines mark the locations of replication origins of replication and the *crtS* sites. The solid black lines represent the fitting of regression lines and the green line corresponds to the Loess regression ($F = 0.05$). Maxima are highlighted by red and minima as blue dots. Plots of biological replicates are shown in **Supplementary Figure S1**.

fused chromosome that may provide a selective advantage thereby eliminating the cells with single chromosomes from the population. Contrary to the expectation, we found that cells retained the fused chromosomes indicating that they are in fact, locked in this configuration and not easily revertible. An alternative explanation would be that the fused state is under positive selection pressure. It remains to be seen what the significance of the fused chromosome status in NSCV1 and NSCV2 is, in contrast to vast majority of other *V. cholerae* strains where two chromosome status is the norm. In any case, the single-chromosome *V. cholerae* appears to be more frequent than expected as yet another NSCV strain was discovered recently (Yamamoto et al., 2018). We expect future studies of NSCV strains which are a deviation from the norm, would to lead to a better understanding of why most *V. cholerae* carry their genome split into two as the norm.

MATERIALS AND METHODS

Strains, Plasmids, Oligonucleotides, and Growth Conditions

All strains, plasmids and oligonucleotides used in this study are listed in the **Supplementary Tables S1–S3**. Unless indicated otherwise, the bacterial cells were grown in LB medium at a temperature of 37°C. Antibiotic selection was performed at the following concentrations, if not indicated differently: Ampicillin 100 μg/ml, Kanamycin 35 μg/ml for *E. coli* and 70 μg/ml for *V. cholerae*, Spectinomycin 100 μg/ml, Gentamicin 20 μg/ml, Chloramphenicol 35 μg/ml. Where needed, diaminopimelic acid (DAP) was added to the medium at a concentration of 300 μM. To determine doubling times, cells were grown in LB medium in a 96-well plate at 37°C in a microplate reader (Infinite M200 pro multimode microplate reader, Tecan). OD_{600} was measured every 5 min for 18 h. Doubling times were calculated in exponential phase for OD_{600} values between 0.01

and 0.1. For continues cultivation of NSCV strains to investigate potential chromosome splitting strains were inoculated in the morning and grown in 100 ml liquid medium for 24 h. The next morning, new flasks were inoculated 1:1,000 from the previous cultures. After 4 working days (on day 5), 1 ml samples of the cultures were frozen at −80°C as glycerin stocks. Two days later, new flasks were inoculated from these glycerin stocks and the procedure was continued for a total of 16 days. Before sequencing, these were re-streaked on TCBS medium to verify that it is *V. cholerae* and to obtain single colonies. For each strain one single colony was used to inoculate a culture for DNA isolation and sequencing.

Construction of Replicons and Strains

NSCV minichromosomes were constructed by PCR amplifying the respective origins of replication using the primers 1227/1228 on *V. cholerae* NSCV1 genomic DNA for pMA739 or *V. cholerae* NSCV2 genomic DNA for pMA755. The vector pMA135 was digested with *Asc*I and the PCR products were integrated in the linearized vector using Gibson Assembly (Gibson et al., 2009) used to transform *E. coli* WM3064. For construction of chromosomal integrations and deletions, integration cassettes were assembled by using the MoClo system as described previously (Weber et al., 2011; Milbredt et al., 2016; Schindler et al., 2016). The *ori2* insertion cassettes on pMA735 and pMA736, as well as the *crtS* deletion and insertion cassettes on pMA748 and pMA749 were assembled by MoClo reactions using the plasmids indicated in **Supplementary Table S2**, which themselves were assembled by MoClo reactions of the respective PCR products into the respective backbones (primers, templates, and backbones indicated in **Supplementary Table S2**). The linear cassettes were released by restriction enzyme digestions with *Bsa*I. Triparental mating was performed to deliver the plasmids pUXBF13 and pGP704-mTn7-hapR_ATN from *E. coli* S17-1 λ*pir* to *V. cholerae* MCH1 (Bao et al., 1991; Meibom et al., 2005; Val et al., 2012). The created strain *V. cholerae* VC49 was naturally

competent and by transforming it with the *ori2* insertion cassettes released from pMA735 or pMA736, respectively, followed by transformation with pBR-*flp* and flippase reaction strains VC61 and VC62 were created (De Souza Silva and Blokesch, 2010). *V. cholerae* strain VC71 was constructed by deletion of the *crtS* sequence by transforming *V. cholerae* VC49 with the *crtS* deletion cassette from pMA748 and a flippase reaction to excise the resistance marker. Subsequently, strain VC71 was transformed with the *crtS* insertion cassette from pMA749 and the flippase reaction was performed to create strain VC73. To construct plasmid pMA449, *crtS* was amplified with primers 1439/1440 from gDNA of *V. cholerae* NSCV1 and for pMA450 with primers 1439/1440 from gDNA of *V. cholerae* NSCV2. PCR products were assembled in pMA349 by MoClo assembly as described (Schindler et al., 2016) and used to transform *E. coli* TOP10 cells. Details of further MoClo assemblies are provided in **Supplementary Table S2**. Assemblies were used to transform *E. coli* DH5α λpir. The derived integration cassettes were cut out with BsaI, integrated into the chromosome of *E. coli* AB330 and transferred to *E. coli* MG1655 per P1-transduction. FRT recombination was used to remove the resistance marker.

Sequence Comparison

Sequences from *V. cholerae* N16961, NSCV1 and NSCV2 were compared by multiple sequence alignment using Clustal Omega (Goujon et al., 2010; Sievers et al., 2011) to find single nucleotide polymorphisms. To analyze the whole genomes of the *V. cholerae* strains NSCV1 and NSCV2 for sequence homology, the alignment free chromosome comparison tool SMASH (Pratas et al., 2015) was used with a minimum block size setting of 5,000 bp.

Preparation of Genomic DNA From Bacteria

The desired amount of culture (between 0.1 ml and 5 ml) was mixed 1:1 with ice-cold killing buffer and centrifuged at maximum speed and 4°C for 3 min. The pellet was resuspended in 300 µl TE, 40 µl SDS and 3 µl 0.5 M EDTA. After 5 min incubation at 65°C, 750 µl Isopropanol were added and the sample was centrifuged 5 min at maximum speed. The pellet was resuspended in 500 µl TE and 2 µl RNAse were added. After 30 min at 37°C, 2 µl of Proteinase K were added and incubated for another 15 min. The sample was purified via twofold Phenol-Chloroform extraction and precipitated over night at −20°C by mixing it with 1 ml pure ethanol and 40 µl 3 M sodium acetate.

The next day, DNA was spinned down at maximum speed for 10 min, the pellet was washed with 70% ethanol and centrifuged another 10 min. The pellet was dried and resuspended in 50 µl pure water.

Short Read Sequencing (Illumina MiSeq)

V. cholerae genomic DNA was isolated as described (Kemter et al., 2018) from various strains grown under different conditions (log phase vs. stationary) were quantitated using the Qubit fluorimeter (Thermo Fisher) and adjusted to 0.2 ng/µL with nuclease-free water. Sequencing libraries were prepared using

the Illumina Nextera XT kit, processed and pooled according to manufacturer's instructions (Illumina). The final, pooled sample was paired-end sequenced (2×300 bp) using the Illumina MiSeq with a v3 chemistry, 600 cycle kit. Post sequencing processing was performed the systems software packages and the final demultiplexed fastq reads produced by the instrument were used for MFA against the reference genome. Raw sequencing data are available on request.

Long Read Sequencing (PacBio)

Whole genome sequencing was performed on a Pacific Biosciences RSII platform. The sequencing library was prepared using the SMRTbell[TM] Template Prep Kit (Pacific Biosciences, Menlo Park, CA, United States) following manufacturer's protocol. 5 µg of DNA was fragmented using gTUBE (Covaris Inc., Woburn, MA, United States) to ∼20 kb. After DNA damage repair and ends repair, blunt hairpin adapters were ligated to the template. Non-ligated products were digested by ExoIII and ExoVII exonucleases. Resulting SMRTbell template were purified with AMPure PB beads and size selected on BluePippin system (Sage Science, Beverly, MA, United States), using 0.75% dye-free agarose cassette, with 4–10 kb Hi-Pass protocol and lower cut set on 4 kb. Size selected purified libraries were quantified by Qubit dsDNA High Sensitivity assay. After primer annealing, and P6 polymerase binding, templates were bound to MagBeads for loading. Each sample was sequenced on two SMRT cells, using C4 sequencing kit and 360-min movies per SMRT cell. Presence of unidentified contaminant in two libraries (NSV2 16 days) inhibited sequencing reactions, which manifested in extremely low P1 (2%) and super short reads. Both libraries were subjected to the cleanup procedure that involves binding annealed SMRTbell libraries to magnetic beads, washing the bound annealed DNA SMRTbell templates to remove potential contaminants, and eluting the purified, annealed DNA SMRTbell templates from the magnetic beads. The purified SMRTbell templates were then re-quantified by Qubit and prepared for sequencing on the PacBio RSII according to the Binding Calculator. After cleanup procedure, PacBio RS II instrument sequencing yields were comparable to the other samples. Raw sequencing data are available on request.

Marker Frequency Analysis

Sequencing reads from a NGS were mapped to the respective genome using the program Geneious (Biomatters Ltd.; Kearse et al., 2012). Read densities were extracted and plotted using custom R scripts as described previously (Kemter et al., 2018).

Semiquantitative Conjugation

All used replicons possess an *oriR6K* and were conjugated from *E. coli* strain WM3064 to MG1655. For overnight cultures of donor and recipient strains the OD600 was determined and the amount of cells corresponding to 1 ml of OD600 = 1 was centrifuged 1 min at 13,000 × *g*. The cells were washed twice in TBS and resuspended in 100 µl TBS. From each donor strain, 50 µl were mixed with 50 µl of the recipient strain and dropped on LB agar including DAP. After 6 h cells were scraped off the plate, washed twice in TBS. The total CFU of recipient cells was

determined by plating dilutions on LB, while the CFU of plasmid bearing recipients was determined by plating the same dilutions on selective media. The selective CFU was then normalized to total CFU.

Microscopy

Differential interference contrast (DIC) microscopy was performed on 1% (w/v) agarose pads in PBS buffer using the Nikon Ti fluorescence microscope (100× objective, NA 1.45).

Quantification of Replicon Copy Number via Antibiotic Sensitivity

The copy-up effect of *crtS* was measured as described (Messerschmidt et al., 2016). Cells were grown in LB medium with either 100 or 500 μg/ml ampicillin at 37°C in 96-well plates in a microplate reader (Infinite M200 pro multimode microplate reader, Tecan). The main culture (150 μL) was inoculated 1:1,000 and growth curves recorded for 15 h. For better visualization, 1 divided by the time needed to reach an OD_{600} of 0.1 was defined as measure of the copy number.

AUTHOR CONTRIBUTIONS

MB, DS, SS, and TW contributed to the conceptualization of the study. MB, DS, FK, MW, KC, GK, and GP contributed to the methodology. MB, DS, KC, and GK contributed to the software. MB, DS, FK, MW, KC, GK, GP, SS, and TW contributed to the formal analysis of the study. MB, DS, FK, MW, KC, GK, and GP executed the investigation for the study. MB, SS, and TW contributed to the writing of the original draft. SS and TW

reviewed and edited the manuscript. DS, SS, and TW provided the supervision. SS, GP, and TW contributed to the funding acquisition.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2018.02932/full#supplementary-material

REFERENCES

Baek, J. H., and Chattoraj, D. K. (2014). Chromosome I controls chromosome II replication in *Vibrio cholerae. PLoS Genet.* 10:e1004184. doi: 10.1371/journal.pgen.1004184

Bao, Y., Lies, D. P., Fu, H., and Roberts, G. P. (1991). An improved Tn7-based system for the single-copy insertion of cloned genes into chromosomes of gram-negative bacteria. *Gene* 109, 167–168. doi: 10.1016/0378-1119(91)90604-A

Chapman, C., Henry, M., Bishop-Lilly, K. A., Awosika, J., Briska, A., Ptashkin, R. N., et al. (2015). Scanning the landscape of genome architecture of non-O1 and non-O139 *Vibrio cholerae* by whole genome mapping reveals extensive population genetic diversity. *PLoS One* 10:e0120311. doi: 10.1371/journal.pone.0120311

Ciaccia, P. N., Ramachandran, R., and Chattoraj, D. K. (2018). A requirement for global transcription factor Lrp in licensing replication of *Vibrio cholerae* chromosome 2. *Front. Microbiol.* 9:2103. doi: 10.3389/fmicb.2018.02103

Dalia, T. N., Hayes, C. A., Stolyar, S., Marx, C. J., McKinlay, J. B., and Dalia, A. B. (2017). Multiplex genome editing by natural transformation (MuGENT) for synthetic biology in *Vibrio natriegens. ACS Synth. Biol.* 6, 1650–1655. doi: 10.1021/acssynbio.7b00116

de Lemos Martins, F., Fournes, F., Mazzuoli, M. V., Mazel, D., and Val, M. E. (2018). *Vibrio cholerae* chromosome 2 copy number is controlled by the methylation-independent binding of its monomeric initiator to the chromosome 1 *crtS* site. *Nucleic Acids Res.* 46, 10145–10156. doi: 10.1093/nar/gky790

De Souza Silva, O., and Blokesch, M. (2010). Genetic manipulation of *Vibrio cholerae* by combining natural transformation with FLP recombination. *Plasmid* 64, 186–195. doi: 10.1016/j.plasmid.2010.08.001

Demarre, G., and Chattoraj, D. K. (2010). DNA adenine methylation is required to replicate both *Vibrio cholerae* chromosomes once per cell cycle. *PLoS Genet.* 6:e1000939. doi: 10.1371/journal.pgen.1000939

Duigou, S., Knudsen, K. G., Skovgaard, O., Egan, E. S., Lobner-Olesen, A., and Waldor, M. K. (2006). Independent control of replication initiation of the two *Vibrio cholerae* chromosomes by DnaA and RctB. *J. Bacteriol.* 188, 6419–6424. doi: 10.1128/JB.00565-06

Duigou, S., Yamaichi, Y., and Waldor, M. K. (2008). ATP negatively regulates the initiator protein of *Vibrio cholerae* chromosome II replication. *Proc. Natl. Acad. Sci. U.S.A.* 105, 10577–10582. doi: 10.1073/pnas.0803904105

Egan, E. S., Fogel, M. A., and Waldor, M. K. (2005). Divided genomes: negotiating the cell cycle in prokaryotes with multiple chromosomes. *Mol. Microbiol.* 56, 1129–1138. doi: 10.1111/j.1365-2958.2005.04622.x

Egan, E. S., and Waldor, M. K. (2003). Distinct replication requirements for the two *Vibrio cholerae* chromosomes. *Cell* 114, 521–530. doi: 10.1016/S0092-8674(03)00611-1

Espinosa, E., Barre, F. X., and Galli, E. (2017). Coordination between replication, segregation and cell division in multi-chromosomal bacteria: lessons from *Vibrio cholerae. Int. Microbiol.* 20, 121–129. doi: 10.2436/20.1501.01.293

Flusberg, B. A., Webster, D. R., Lee, J. H., Travers, K. J., Olivares, E. C., Clark, T. A., et al. (2010). Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nat. Methods* 7, 461–465. doi: 10.1038/nmeth.1459

Fournes, F., Val, M. E., Skovgaard, O., and Mazel, D. (2018). Replicate once per cell cycle: replication control of secondary chromosomes. *Front. Microbiol.* 9:1833. doi: 10.3389/fmicb.2018.01833

Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A., and Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* 6, 343–345. doi: 10.1038/nmeth.1318

Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J., et al. (2010). A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res.* 38, W695–W699. doi: 10.1093/nar/gkq313

Heidelberg, J. F., Eisen, J. A., Nelson, W. C., Clayton, R. A., Gwinn, M. L., Dodson, R. J., et al. (2000). DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature* 406, 477–483. doi: 10.1038/35020000

Hoffart, E., Grenz, S., Lange, J., Nitschel, R., Muller, F., Schwentner, A., et al. (2017). High substrate uptake rates empower *Vibrio natriegens* as production host for industrial biotechnology. *Appl. Environ. Microbiol.* 83:e01614-17. doi: 10.1128/AEM.01614-17

Jha, J. K., Baek, J. H., Venkova-Canova, T., and Chattoraj, D. K. (2012). Chromosome dynamics in multichromosome bacteria. *Biochim. Biophys. Acta* 1819, 826–829. doi: 10.1016/j.bbagrm.2012.01.012

Johnson, S. L., Khiani, A., Bishop-Lilly, K. A., Chapman, C., Patel, M., Verratti, K., et al. (2015). Complete genome assemblies for two single-chromosome *Vibrio cholerae* Isolates, strains 1154-74 (Serogroup O49) and 10432-62 (Serogroup O27). *Genome Announc.* 3:e00462-15. doi: 10.1128/genomeA.00462-15

Kamp, H. D., Patimalla-Dipali, B., Lazinski, D. W., Wallace-Gadsden, F., and Camilli, A. (2013). Gene fitness landscapes of *Vibrio cholerae* at important stages of its life cycle. *PLoS Pathog.* 9:e1003800. doi: 10.1371/journal.ppat.1003800

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649. doi: 10.1093/bioinformatics/bts199

Kemter, F. S., Messerschmidt, S. J., Schallopp, N., Sobetzko, P., Lang, E., Bunk, B., et al. (2018). Synchronous termination of replication of the two chromosomes is an evolutionary selected feature in Vibrionaceae. *PLoS Genet.* 14:e1007251. doi: 10.1371/journal.pgen.1007251

Koch, B., Ma, X., and Lobner-Olesen, A. (2010). Replication of *Vibrio cholerae* chromosome I in *Escherichia coli*: dependence on dam methylation. *J. Bacteriol.* 192, 3903–3914. doi: 10.1128/JB.00311-10

Lo Scrudato, M., and Blokesch, M. (2012). The regulatory network of natural competence and transformation of *Vibrio cholerae*. *PLoS Genet.* 8:e1002778. doi: 10.1371/journal.pgen.1002778

Lo Scrudato, M., and Blokesch, M. (2013). A transcriptional regulator linking quorum sensing and chitin induction to render *Vibrio cholerae* naturally transformable. *Nucleic Acids Res.* 41, 3644–3658. doi: 10.1093/nar/gkt041

Lu, M., Campbell, J. L., Boye, E., and Kleckner, N. (1994). SeqA: a negative modulator of replication initiation in *E. coli*. *Cell* 77, 413–426. doi: 10.1016/0092-8674(94)90156-2

Meibom, K. L., Blokesch, M., Dolganov, N. A., Wu, C. Y., and Schoolnik, G. K. (2005). Chitin induces natural competence in *Vibrio cholerae*. *Science* 310, 1824–1827. doi: 10.1126/science.1120096

Messerschmidt, S. J., Kemter, F. S., Schindler, D., and Waldminghaus, T. (2015). Synthetic secondary chromosomes in *Escherichia coli* based on the replication origin of chromosome II in *Vibrio cholerae*. *Biotechnol. J.* 10, 302–314. doi: 10.1002/biot.201400031

Messerschmidt, S. J., Schindler, D., Zumkeller, C. M., Kemter, F. S., Schallopp, N., and Waldminghaus, T. (2016). Optimization and characterization of the synthetic secondary chromosome *synVicII* in *Escherichia coli*. *Front. Bioeng. Biotechnol.* 4:96. doi: 10.3389/fbioe.2016.00096

Milbredt, S., Farmani, N., Sobetzko, P., and Waldminghaus, T. (2016). DNA replication in engineered *Escherichia coli* genomes with extra replication origins. *ACS Synth. Biol.* 5, 1167–1176. doi: 10.1021/acssynbio.6b00064

Okada, K., Iida, T., Kita-Tsukamoto, K., and Honda, T. (2005). Vibrios commonly possess two chromosomes. *J. Bacteriol.* 187, 752–757. doi: 10.1128/JB.187.2.752-757.2005

Pratas, D., Silva, R. M., Pinho, A. J., and Ferreira, P. J. (2015). An alignment-free method to find and visualise rearrangements between pairs of DNA sequences. *Sci. Rep.* 5:10203. doi: 10.1038/srep10203

Ramachandran, R., Ciaccia, P. N., Filsuf, T. A., Jha, J. K., and Chattoraj, D. K. (2018). Chromosome 1 licenses chromosome 2 replication in *Vibrio cholerae* by doubling the *crtS* gene dosage. *PLoS Genet.* 14:e1007426. doi: 10.1371/journal.pgen.1007426

Ramachandran, R., Jha, J., Paulsson, J., and Chattoraj, D. (2017). Random versus cell cycle-regulated replication initiation in bacteria: insights from studying *Vibrio cholerae* chromosome 2. *Microbiol. Mol. Biol. Rev.* 81, e00033-16. doi: 10.1128/MMBR.00033-16

Rasmussen, T., Jensen, R. B., and Skovgaard, O. (2007). The two chromosomes of *Vibrio cholerae* are initiated at different time points in the cell cycle. *EMBO J.* 26, 3124–3131. doi: 10.1038/sj.emboj.7601747

Saint-Dic, D., Kehrl, J., Frushour, B., and Kahng, L. S. (2008). Excess SeqA leads to replication arrest and a cell division defect in *Vibrio cholerae*. *J. Bacteriol.* 190, 5870–5878. doi: 10.1128/JB.00479-08

Schallopp, N., Milbredt, S., Sperlea, T., Kemter, F. S., Bruhn, M., Schindler, D., et al. (2017). Establishing a system for testing replication inhibition of the *Vibrio cholerae* secondary chromosome in *Escherichia coli*. *Antibiotics* 7:E3. doi: 10.3390/antibiotics7010003

Schindler, D., Milbredt, S., Sperlea, T., and Waldminghaus, T. (2016). Design and assembly of DNA sequence libraries for chromosomal insertion in bacteria based on a set of modified MoClo vectors. *ACS Synth. Biol.* 5, 1362–1368. doi: 10.1021/acssynbio.6b00089

Schoolnik, G. K., and Yildiz, F. H. (2000). The complete genome sequence of *Vibrio cholerae*: a tale of two chromosomes and of two lifestyles. *Genome Biol.* 1, 1016.1–1016.3. doi: 10.1186/gb-2000-1-3-reviews1016

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 7:539. doi: 10.1038/msb.2011.75

Skovgaard, O., Bak, M., Lobner-Olesen, A., and Tommerup, N. (2011). Genome-wide detection of chromosomal rearrangements, indels, and mutations in circular chromosomes by short read sequencing. *Genome Res.* 21, 1388–1393. doi: 10.1101/gr.117416.110

Slater, S., Wold, S., Lu, M., Boye, E., Skarstad, K., and Kleckner, N. (1995). *E. coli* SeqA protein binds oriC in two different methyl-modulated reactions appropriate to its roles in DNA replication initiation and origin sequestration. *Cell* 82, 927–936. doi: 10.1016/0092-8674(95)90272-4

Srivastava, P., and Chattoraj, D. K. (2007). Selective chromosome amplification in *Vibrio cholerae*. *Mol. Microbiol.* 66, 1016–1028. doi: 10.1111/j.1365-2958.2007.05973.x

Stokke, C., Waldminghaus, T., and Skarstad, K. (2011). Replication patterns and organization of replication forks in *Vibrio cholerae*. *Microbiology* 157, 695–708. doi: 10.1099/mic.0.045112-0

Val, M. E., Kennedy, S. P., Soler-Bistue, A. J., Barbe, V., Bouchier, C., Ducos-Galand, M., et al. (2014a). Fuse or die: how to survive the loss of Dam in *Vibrio cholerae*. *Mol. Microbiol.* 91, 665–678. doi: 10.1111/mmi.12483

Val, M. E., Soler-Bistue, A., Bland, M. J., and Mazel, D. (2014b). Management of multipartite genomes: the *Vibrio cholerae* model. *Curr. Opin. Microbiol.* 22, 120–126. doi: 10.1016/j.mib.2014.10.003

Val, M. E., Marbouty, M., de Lemos Martins, F., Kennedy, S. P., Kemble, H., Bland, M. J., et al. (2016). A checkpoint control orchestrates the replication of the two chromosomes of *Vibrio cholerae*. *Sci. Adv.* 2:e1501914. doi: 10.1126/sciadv.1501914

Val, M. E., Skovgaard, O., Ducos-Galand, M., Bland, M. J., and Mazel, D. (2012). Genome engineering in *Vibrio cholerae*: a feasible approach to address biological issues. *PLoS Genet.* 8:e1002472. doi: 10.1371/journal.pgen.1002472

Venkova-Canova, T., and Chattoraj, D. K. (2011). Transition from a plasmid to a chromosomal mode of replication entails additional regulators. *Proc. Natl. Acad. Sci. U.S.A.* 108, 6199–6204. doi: 10.1073/pnas.1013244108

Waldminghaus, T., and Skarstad, K. (2009). The *Escherichia coli* SeqA protein. *Plasmid* 61, 141–150. doi: 10.1016/j.plasmid.2009.02.004

Weber, E., Engler, C., Gruetzner, R., Werner, S., and Marillonnet, S. (2011). A modular cloning system for standardized assembly of multigene constructs. *PLoS One* 6:e16765. doi: 10.1371/journal.pone.0016765

Weinstock, M. T., Hesek, E. D., Wilson, C. M., and Gibson, D. G. (2016). *Vibrio natriegens* as a fast-growing host for molecular biology. *Nat. Methods* 13, 849–851. doi: 10.1038/nmeth.3970

Xie, G., Johnson, S. L., Davenport, K. W., Rajavel, M., Waldminghaus, T., Detter, J. C., et al. (2017). Exception to the rule: genomic characterization of naturally occurring unusual *Vibrio cholerae* strains with a single chromosome. *Int. J. Genomics* 2017:8724304. doi: 10.1155/2017/8724304

Yamamoto, S., Lee, K. I., Morita, M., Arakawa, E., Izumiya, H., and Ohnishi, M. (2018). Single circular chromosome identified from the genome sequence of the *Vibrio cholerae* O1 bv. El Tor Ogawa strain V060002. *Genome Announc.* 6:e00564-18. doi: 10.1128/genomeA.00564-18

Commentary: Functionality of Two Origins of Replication in *Vibrio cholerae* Strains With a Single Chromosome

Bhabatosh Das[1] and Dhruba K. Chattoraj[2]*

[1] *Translational Health Science and Technology Institute, Faridabad, India,* [2] *Center of Cancer Research (CCR), National Cancer Institute (NCI) and National Institute Health (NIH), Bethesda, MD, United States*

A Commentary on

Functionality of Two Origins of Replication in *Vibrio cholerae* Strains With a Single Chromosome

by Bruhn, M., Schindler, D., Kemter, F. S., Wiley, M. R., Chase, K., Koroleva, G. I., et al. (2018). Front. Microbiol. 9:2932. doi: 10.3389/fmicb.2018.02932

This paper is about divided genomes in bacteria. In the era of genomics, it has become clear that about 10% of bacteria have multiple chromosomes, as is the norm in eukaryotes. This observation raises the questions as to how they originated and how they are maintained, in particular whether their replication and segregation are independently or coordinately controlled, and what evolutionary advantage the divided genome might have to discourage reversion to the single-chromosome state, the norm in bacteria.

The prevailing view is that multi-chromosome bacteria have originated from single-chromosome bacteria by transferring some essential genes from the chromosome to plasmids, thus making the plasmid an indispensable component of the genome or in other words, another chromosome (Fournes et al., 2018). The best evidence for this view comes from studies of *Vibrio cholerae* (Vc), which has one main chromosome (Chr1), analogous to the paradigmatic *Escherichia coli* chromosome, carrying most of the housekeeping genes, and a second chromosome (Chr2) with distinct hallmarks of certain low-copy number *E. coli* plasmids, such as P1 and F, but carrying some essential genes not present in Chr1.

Genomes of many naturally occurring *Vibrio* strains have been analyzed, and in the *Vibrionaceae* family that includes Vc, the two-chromosome genome has been the rule. However, in a recent analysis of 91 *Vibrio* strains from the Sakazaki collection, two strains were found with a single chromosome that resulted from fusion of Chr1 and Chr2 (Chapman et al., 2015; Xie et al., 2017). This is the first report of a naturally occurring single-chromosome *Vibrio* (NSCV), although forced fusions in the laboratory were achieved earlier (Val et al., 2012, 2014, 2016). Since then, another *Vibrio* with single chromosome has been reported (Yamamoto et al., 2018). Note that in all the laboratory-achieved fusions, the Chr2 replicon was inactive, and the strains survived because Chr2 could be passively maintained as an integral part of Chr1. In contrast, both Chr1 and Chr2 origins (*ori1* and *ori2*) were active in the strain NSCV1 of Xie et al. (**Figure 1**) (Bruhn et al., 2018). The commentary is based on this exceptional finding.

The basic claim that both the origins can function in a fused chromosome is reasonable. Particularly, the authors verified that the two special features of Chr2 replication, dependence on Dam methylation and on two copies of a replication enhancer site *crtS*, are retained after the fusion.
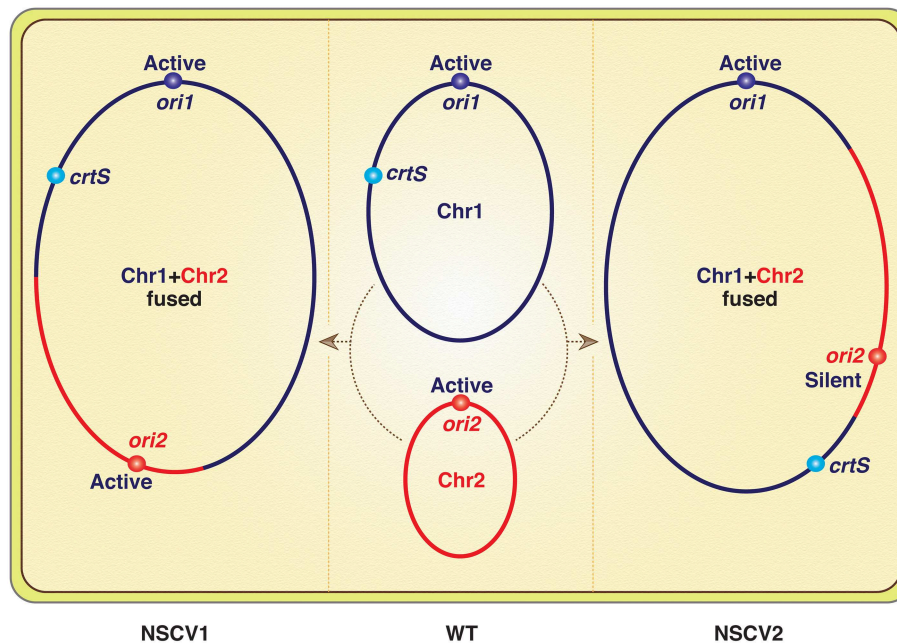
FIGURE 1 | Schematics of *V. cholerae* chromosomes (Chr1 and Chr2) before (in WT) and after fusions (in NSCV1 and NSCV2). Note that a replication fork emanating from *ori1* will encounter *crtS* before *ori2* in NSCV1 whereas the opposite will be the case in NSCV2. The inactivity of *ori2* in NSCV2 is expected since *crtS* duplication is a prerequisite for *ori2* firing, although the authors have reasons to believe that this may not be the real explanation.

In the other fused chromosome strain (NSCV2), the fusion junctions were different, and *ori2* was silent. The authors attributed this to an altered genomic context of the regulatory sites (*ori1, ori2,* and *crtS,* **Figure 1**) and not on their relative positions, which is currently believed to be important for *ori2* function. Although how the context matters was not elaborated on, a new perspective on Chr2 replication was provided to explain the results. The idea is that the regulation of Chr2 replication is such that it maintains the parity of *crtS* to *ori2* copy numbers (de Lemos Martins et al., 2018; Ramachandran et al., 2018). The *crtS* site normally resides in Chr1 and when the site number doubles upon passage of the Chr1 replication fork, Chr2 replication initiates and restores the *crtS*/*ori2* ratio.

A bacterial chromosome with two functional origins is unprecedented and is unexpected. A reason for why bacterial chromosomes have one origin whereas eukaryotic chromosomes have multiple origins, has been proposed (Kuzminov, 2014). In eukaryotes, chromosomes segregate at the end of replication, and the entire chromosome segregates as a unit, whereas in bacteria the two arms of a replication bubble start segregating away from each other soon after their synthesis. In other words, segregation proceeds much before the completion of replication. If there are two replication bubbles on the same chromosome from two differently located origins, then productive segregation of the replicated arms would require that the parental Watson strand of both the bubbles go in the same direction, and the parental Crick strand of both the bubbles go in the opposite direction. No mechanism for such non-random segregation is known. It might well be that to avoid random segregation of locally replicated

arms, which can potentially entangle rather than segregate the replicated arms, bacteria with a single origin might have enjoyed a significant selective advantage.

The two single-chromosome strains, however, were stable when grown over 160 generations. How? Fusion junctions indicate that complex genetic rearrangements accompanied the joining of the two chromosomes, which would prevent the simple reversal of the integration event. As argued above, the stability of an irreversibly fused chromosome can be improved by silencing one of the origins, which is the case in NSCV2. In NSCV1, it is still possible that only one of the functional origins fires in any one cell cycle. Even if both the origins fire in the same cell cycle, then silencing or overriding of one of the two segregation systems of Vc would avoid the mess that random segregation of replicated arms might cause. Also, Chr1 initiates replication first and its segregation system is set in motion well-before the onset of the Chr2 segregation. In the fused chromosome, the Chr1 system most likely dominates, which is a testable prediction.

The finding that chromosomes can fuse and that the fused chromosome can be stably maintained with two functional origins raises the question: What keeps the chromosomes from fusing in the vast majority of cases? This is even more surprising because the chromosomes share plenty of regions for homologous recombination (Heidelberg et al., 2000). Fused-chromosome strains should be viewed as an exception and, moving forward, the emphasis should be on understanding the selective advantages of maintaining the divided state. Since the majority of bacteria have one chromosome, the selection of the divided state must have some species-specific basis. For

example, *Vibrios* are one of the fastest growing bacteria (Lee et al., 2019). The high growth rate entails multi-fork replication and dividing the genome lessens the demand for more forks (Srivastava and Chattoraj, 2007). A chromosome with fewer forks should be less vulnerable to damage, which would be worth exploring.

AUTHOR CONTRIBUTIONS

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

REFERENCES

Bruhn, M., Schindler, D., Kemter, F. S., Wiley, M. R., Chase, K., Koroleva, G. I., et al. (2018). Functionality of two origins of replication in *Vibrio cholerae* strains with a single chromosome. *Front. Microbiol.* 9:2932. doi: 10.3389/fmicb.2018.02932

Chapman, C., Henry, M., Bishop-Lilly, K. A., Awosika, J., Briska, A., Ptashkin, R. N., et al. (2015). Scanning the landscape of genome architecture of non-O1 and non-O139 *Vibrio cholerae* by whole genome mapping reveals extensive population genetic diversity. *PLoS ONE* 10:e0120311. doi: 10.1371/journal.pone.0120311

de Lemos Martins, F., Fournes, F., Mazzuoli, M. V., Mazel, D., and Val, M. E. (2018). *Vibrio cholerae* chromosome 2 copy number is controlled by the methylation-independent binding of its monomeric initiator to the chromosome 1 crtS site. *Nucleic Acids Res.* 46, 10145–10156. doi: 10.1093/nar/gky790

Fournes, F., Val, M. E., Skovgaard, O., and Mazel, D. (2018). Replicate once per cell cycle: replication control of secondary chromosomes. *Front. Microbiol.* 9:1833. doi: 10.3389/fmicb.2018.01833

Heidelberg, J. F., Eisen, J. A., Nelson, W. C., Clayton, R. A., Gwinn, M. L., Dodson, R. J., et al. (2000). DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae. *Nature* 406, 477–483. doi: 10.1038/35020000

Kuzminov, A. (2014). The precarious prokaryotic chromosome. *J. Bacteriol.* 196, 1793–1806. doi: 10.1128/JB.00022-14

Lee, H. H., Ostrov, N., Wong, B. G., Gold, M. A., Khalil, A. S., and Church, G. M. (2019). Functional genomics of the rapidly replicating bacterium *Vibrio natriegens* by CRISPRi. *Nat. Microbiol.* doi: 10.1038/s41564-019-0423-8

Ramachandran, R., Ciaccia, P. N., Filsuf, T. A., Jha, J. K., and Chattoraj, D. K. (2018). Chromosome 1 licenses chromosome 2 replication in *Vibrio cholerae* by doubling the crtS gene dosage. *PLoS Genet.* 14:e1007426. doi: 10.1371/journal.pgen.1007426

Srivastava, P., and Chattoraj, D. K. (2007). Selective chromosome amplification in *Vibrio cholerae*. *Mol. Microbiol.* 66, 1016–1028. doi: 10.1111/j.1365-2958.2007.05973.x

Val, M. E., Kennedy, S. P., Soler-Bistue, A. J., Barbe, V., Bouchier, C., Ducos-Galand, M., et al. (2014). Fuse or die: how to survive the loss of Dam in *Vibrio cholerae*. *Mol. Microbiol.* 91, 665–678. doi: 10.1111/mmi.12483

Val, M. E., Marbouty, M., de Lemos Martins, F., Kennedy, S. P., Kemble, H., Bland, M. J., et al. (2016). A checkpoint control orchestrates the replication of the two chromosomes of *Vibrio cholerae*. *Sci. Adv.* 2:e1501914. doi: 10.1126/sciadv.1501914

Val, M. E., Skovgaard, O., Ducos-Galand, M., Bland, M. J., and Mazel, D. (2012). Genome engineering in *Vibrio cholerae*: a feasible approach to address biological issues. *PLoS Genet.* 8:e1002472. doi: 10.1371/journal.pgen.1002472

Xie, G., Johnson, S. L., Davenport, K. W., Rajavel, M., Waldminghaus, T., Detter, J. C., et al. (2017). Exception to the rule: genomic characterization of naturally occurring unusual *Vibrio cholerae* strains with a single chromosome. *Int. J. Genomics* 2017:8724304. doi: 10.1155/2017/8724304

Yamamoto, S., Lee, K. I., Morita, M., Arakawa, E., Izumiya, H., and Ohnishi, M. (2018). Single circular chromosome identified from the genome sequence of the *Vibrio cholerae* O1 bv. El Tor Ogawa Strain V060002. *Genome Announc.* 6:e00564–18. doi: 10.1128/genomeA.00564-18

Structure and Function of the *Campylobacter jejuni* Chromosome Replication Origin

Pawel Jaworski[1], Rafal Donczew[1†], Thorsten Mielke[2], Christoph Weigel[3], Kerstin Stingl[4] and Anna Zawilak-Pawlik[1*]

[1] Department of Microbiology, Ludwik Hirszfeld Institute of Immunology and Experimental Therapy, Polish Academy of Sciences, Wrocław, Poland, [2] Max Planck Institute for Molecular Genetics, Berlin, Germany, [3] Department of Life Science Engineering, Fachbereich 2, HTW Berlin, Berlin, Germany, [4] National Reference Laboratory for Campylobacter, Department of Biological Safety, Federal Institute for Risk Assessment, Berlin, Germany

Campylobacter jejuni is the leading bacterial cause of foodborne infections worldwide. However, our understanding of its cell cycle is poor. We identified the probable *C. jejuni* origin of replication (*oriC*) – a key element for initiation of chromosome replication, which is also important for chromosome structure, maintenance and dynamics. The herein characterized *C. jejuni oriC* is monopartite and contains (i) the DnaA box cluster, (ii) the DnaA-dependent DNA unwinding element (DUE) and (iii) binding sites for regulatory proteins. The cluster of five DnaA boxes and the DUE were found in the *dnaA-dnaN* intergenic region. Binding of DnaA to this cluster of DnaA-boxes enabled unwinding of the DUE *in vitro*. However, it was not sufficient to sustain replication of minichromosomes, unless the cluster was extended by additional DnaA boxes located in the 3′ end of *dnaA*. This suggests, that *C. jejuni oriC* requires these boxes to initiate or to regulate replication of its chromosome. However, further detailed mutagenesis is required to confirm the role of these two boxes in initiation of *C. jejuni* chromosome replication and thus to confirm partial localization of *C. jejuni oriC* within a coding region, which has not been reported thus far for any bacterial *oriC*. *In vitro* DUE unwinding by DnaA was inhibited by Cj1509, an orphan response regulator and a homolog of HP1021, that has been previously shown to inhibit replication in *Helicobacter pylori*. Thus, Cj1509 might play a similar role as a regulator of *C. jejuni* chromosome replication. This is the first systematic analysis of chromosome replication initiation in *C. jejuni*, and we expect that these studies will provide a basis for future research examining the structure and dynamics of the *C. jejuni* chromosome, which will be crucial for understanding the pathogens' life cycle and virulence.

Keywords: Epsilonproteobacteria, *Campylobacter jejuni,* initiation of chromosome replication, *oriC*, DnaA, DnaA box, orisome

INTRODUCTION

Campylobacter jejuni is a Gram-negative, microaerophilic bacterium that belongs to the Epsilon class of Proteobacteria, which has recently been proposed to constitute a separate phylum, the Epsilonbacteryota (Parkhill et al., 2000; Eppinger et al., 2004; Waite et al., 2017). *C. jejuni* colonizes the intestine of diverse animal species in a commensal manner. However, in humans, *C. jejuni*

often invades intestinal epithelial cells and causes acute bacterial gastroenteritis (O Cróinín and Backert, 2012). The infection is generally self-limiting, but complications can arise and may include autoimmune sequelae like reactive arthritis, irritable bowel syndrome and Guillain–Barré syndrome (Kaakoush et al., 2015). *C. jejuni* is isolated from environmental samples; however, the main source of *C. jejuni* infections is the handling and consumption of contaminated poultry meat. Survival in or colonization of diverse niches indicates that *C. jejuni*, although quite stress-sensitive, can resist varying environmental conditions such as low temperature or atmospheric oxygen concentrations (Murphy et al., 2006). Chromosome replication is one of the most vulnerable processes, which has to be highly regulated, because unexpected interruption of replication (e.g., under stress conditions) may be fatal for the bacterium. Hence, a molecular understanding of chromosome replication in *C. jejuni* and other pathogens may provide new reliable ways to combat infections.

To precisely and efficiently regulate chromosome replication, bacteria control the process at the very first step – the initiation. Generally, bacteria initiate replication at *oriC*, a single, unique site on the chromosome. Bacterial *oriCs* consists of one or more clusters of DnaA binding sites (DnaA boxes), the DNA unwinding element (DUE) and the binding sites for regulatory proteins called oriBPs – origin binding proteins (Wolański et al., 2014; Marczynski et al., 2015). Upon initiation, the initiator protein DnaA binds to DnaA boxes at *oriC* and assembles into a filament that can distort double-stranded (ds) DNA at the DUE (Duderstadt and Berger, 2013; Wolański et al., 2014; Leonard and Grimwade, 2015; Katayama et al., 2017). Subsequently, the open complex serves as a platform for the assembly of a multiprotein replication machinery, called the replisome, which will synthesize the nascent chromosome. *oriC* can be mono- or bipartite, i.e., DnaA boxes can be gathered into a single (*Escherichia coli, Mycobacterium tuberculosis*) or two clusters (*Bacillus subtilis, Helicobacter pylori*) located in intergenic regions, usually in the vicinity of the *dnaA – dnaN* locus (Briggs et al., 2012; Wolański et al., 2014). Typical DnaA boxes are non-palindromic, oriented (i.e., 3′–5′ directed) 9-mers with sequences similar to the "perfect" high-affinity R-type *E. coli* DnaA box 5′-TTWTNCACA-3′ (Schaper and Messer, 1995; Leonard and Grimwade, 2011; Wolański et al., 2014). However, in *E. coli* and *Caulobacter cresentus* low-affinity DnaA boxes were also identified, which differ by 3–4 nucleotides from R-type DnaA boxes (McGarry et al., 2004; Kawakami et al., 2005; Taylor et al., 2011). High-affinity DnaA boxes are bound by ATP- and ADP-DnaA, while low-affinity DnaA boxes are exclusively bound by ATP-DnaA. The DUE is located outside of the cluster, adjacent to the last DnaA box in the scaffold. The DUE region has usually a size of around 50 bps and is rich in thymines and adenines (an AT-rich region) (Rajewska et al., 2012), which lower the thermodynamic stability of the DUE and makes it helically unstable, even in the absence of DnaA (Kowalski and Eddy, 1989). It should be noted that, despite minor differences, the *oriC* region unwound either due to intrinsic instability of the helix or in the presence of DnaA was shown to be similar [e.g., *E. coli oriC* (Kowalski and Eddy, 1989; Hwang and Kornberg, 1992a) or

B. subtilis oriC (Krause et al., 1997)]. It has recently been shown that in many bacteria the DUE region proximal to the DnaA-box cluster encodes a 5′-TAG-3′ motif, named DnaA-trio. This motif is required by DnaA to open DNA and to assemble on ssDNA (Richardson et al., 2016). The primary role of the last module, oriBPs binding site, which binds to proteins that control *oriC* activity, is to efficiently transmit feedback information from the cell and/or environment to the *oriC* to rapidly adjust the replication rate (Wolański et al., 2014; Marczynski et al., 2015). oriBPs binding sequences can overlap with DnaA boxes or be located within the DUE or elsewhere within the *oriC*. They bind different classes of proteins, such as nucleoid-associated proteins (NAPs, e.g., *E. coli* IHF, Fis, SeqA) [(Waldminghaus and Skarstad, 2009; Wolański et al., 2014; Leonard and Grimwade, 2015) and references herein] or response regulators of two component systems (e.g., *E. coli* ArcA, *B. subtilis* Spo0A, *H. pylori* HP1021, *M. tuberculosis* MtrA) (Lee et al., 2001; Castilla-Llorente et al., 2006; Donczew et al., 2015; Purushotham et al., 2015). Thus, the oriBPs binding modules are highly diverse, both in structure and species specificity.

oriCs of four Epsilonproteobacteria have been identified to date (Donczew et al., 2012; Jaworski et al., 2016). The Epsilonproteobacterial origins typically co-localize with *ruvC-dnaA-dnaN*, with the exception of Helicobacteraceae species, in which this gene order is not conserved (e.g., *H. pylori dnaA* is located between *punB* and *comH*). They likely constitute bipartite origins, with clusters of DnaA boxes localized upstream (*oriC1*) and downstream (*oriC2*) of *dnaA*; DNA unwinding was shown to occur in *oriC2* (Donczew et al., 2012; Jaworski et al., 2016). The typical Epsilonproteobacterial 9-mer DnaA box consists of the core nucleotide sequence 5′-TTCAC-3′ (4–8 nt of a 9-mer), with the 5th residue strictly conserved. This specific DnaA box sequence, together with the significant changes in the DNA-binding motif of corresponding DnaAs, determines the unique molecular mechanism of the DnaA-DNA interaction (Jaworski et al., 2016). There are two known regulators of *H. pylori* chromosome replication: HobA and HP1021. HobA, a homolog of *E. coli* DiaA (Keyamura et al., 2007; Natrajan et al., 2007), binds to DnaA and controls its oligomerization upon *oriC* binding (Zawilak-Pawlik et al., 2007, 2011), while HP1021 binds to *H. pylori oriC* to preclude DnaA binding to DnaA boxes and inhibit DNA unwinding at *oriC* (Donczew et al., 2015). Homologs of HP1021 and HobA are found in other Epsilonproteobacteria, including *C. jejuni* (Schär et al., 2005; Zawilak-Pawlik et al., 2011).

In this work, we identified and characterized the probable *C. jejuni oriC*. The *C. jejuni oriC* is most likely monopartite with the initial unwinding region located between *dnaA* and *dnaN*. We call this region DnaA-dependent DNA unwinding element (DUE), although, unlike in *E. coli* or *B. subtilis* DUE (Kowalski and Eddy, 1989; Krause et al., 1997), formal proof of protein-independent instability was not conducted. However, this region is AT rich, it is predicted to be helically unstable and it is unwound by DnaA. The *dnaA–dnaN* intergenic region contains a cluster of five DnaA binding sites which enable DnaA to build up a complex capable of DUE unwinding *in vitro*. However, for self-replication of minichromosomes in *C. jejuni* an additional 3′ end of *dnaA* comprising further DnaA binding sites are essential.

Thus, *C. jejuni oriC* might be the first example of a bacterial origin that is partially located within a coding region, however, further studies are required to confirm the essentiality of these two DnaA boxes for initiation of *C. jejuni* chromosome replication. There are numerous DnaA binding sites located in the vicinity of *oriC*, mainly within the *dnaA* gene, which might play regulatory roles in controlling the initiation of complex assembly. We also identified Cj1509 (*C. jejuni* 81116 or Cj1608 in *C. jejuni* 11168), a homolog of *H. pylori* HP1021 (Schär et al., 2005). Here, we show that Cj1509 binds to *C. jejuni oriC* and inhibits *oriC* unwinding at the DUE. We speculate that these orphan regulators are regulating chromosome replication in Epsilonproteobacteria in response to as yet unknown signaling pathways.

MATERIALS AND METHODS

Materials, Strains and Culture Conditions

The plasmids, proteins and bacterial strains used in this work are listed in Supplementary Table S1. The oligonucleotide sequences are presented in Supplementary Table S2. *C. jejuni* 81116 genomic DNA was used as a template to amplify DNA fragments for cloning. *E. coli* was grown at 37°C on solid or liquid Luria-Bertani medium supplemented with 100 μg ml^{-1} ampicillin or 25 μg ml^{-1} kanamycin where necessary. *C. jejuni* was cultivated at 37°C or 42°C under microaerophilic conditions on Columbia Blood (CB) Agar (CM0331, Oxoid) or Brain Heart Infusion (BHI) Broth (CM1135, Oxoid) supplemented with trimethoprim and polymyxin B at final concentrations of 10 μg ml^{-1} and 2.5 U ml^{-1}, respectively; colistin (10 μg ml^{-1}) and kanamycin (25 μg ml^{-1}) were added when necessary.

Minichromosome Maintenance

pRY107d was prepared by HindIII digestion of pRY107 and religation. The IGR regions were amplified by PCR using specific primer pairs (Supplementary Figure S1 and Supplementary Table S2) and inserted between the EcoRI and PstI restriction sites of pRY107d to generate pRY_X (X-respective IGR region, see **Figure 1B**, Supplementary Table S1 and Supplementary Figure S2A). The pRY plasmids were introduced into *C. jejuni* 81–176 using a conjugation protocol that has been described previously (Van Vliet et al., 1998; Zeng et al., 2015), with minor modifications. *C. jejuni* recipient cells were grown overnight on CB agar plates at 37°C under microaerophilic conditions and subsequently harvested using an inoculating loop and 2 ml of BHI broth pre-warmed to 50°C, diluted to OD600 = 1 and incubated at 50°C for 30 min. Washed *E. coli* S17-1 donor cells were resuspended in 0.5 ml of *C. jejuni* cells. The mixture was concentrated by centrifugation to 100 μl and placed on a CB agar plate without antibiotics. After a 5-h incubation at 42°C under microaerophilic conditions, the cells were harvested with 1 ml of BHI, centrifuged, resuspended in 100 μl of BHI and spread on CB plates supplemented with trimethoprim, polymyxin B, colistin and kanamycin. The plates were incubated at 42°C under microaerophilic conditions for 2–4 days. Four colonies of each conjugation were streaked on CB agar plates supplemented with selective antibiotics, incubated for 2 days and harvested. Genomic

DNA was purified and used as a template for PCR with the following primer pairs: B4-B5, F1-F2, M13-rM13, and F3-F4 or in Southern blot analysis. Southern blot was performed as described (Sambrook and Russell, 2001). Briefly 10–20 μg of *C. jejuni* genomic DNA and 10 ng of a control plasmid DNA isolated from *E. coli*, undigested or digested with appropriate restriction enzymes, were resolved in 1% agarose gel. DNA was transferred onto a nylon membrane and incubated at 58°C with digoxigenin-labeled DNA probe (314 bp DNA, amplified with primers A3–A4). Southern blot was developed by colorimetric reaction using anti-digoxigenin antibody (Anti-Digoxigenin-AP, Fab fragments, Roche). The presence of the self-replicating pRY4_6 plasmid in *C. jejuni* was additionally analyzed by transformation of *E. coli* TOP10 competent cells with genomic DNA isolated from *C. jejuni* pRY4_6 conjugants. Plasmid DNA was purified from *E. coli* colonies and analyzed by digestion with EcoRI and PstI.

Protein Expression and Purification

The *dnaA* gene was amplified by PCR using primer pairs B1–B2 and B1–B3 and inserted between the BamHI-XhoI restriction sites of pET28a(+) and pT21Strep, to generate pET28CjDnaA and pET21CjDnaA, respectively. pET21Strep is pET21b(+) derivative modified by removal of T7tag sequence and insertion of the Strep-tag sequence (for details see Supplementary Table S1 and Supplementary Figure S4A). 6HisCjDnaA (N-terminal His-tagged) and StrepCjDnaA (C-terminal Strep-tagged) (Supplementary Figure S4B) were expressed and purified as described previously (Zawilak-Pawlik et al., 2006) or according to Strep-Tactin manufacturer's protocol (IBA Lifesciences). The activities of recombinant His-tagged and Strep-tagged proteins were similar to other Epsilonproteobacterial DnaAs, which was confirmed by using similar experimental conditions as previously published. In particular, in the P1 nuclease or the DMS footprinting assays CjDnaA was used at 20:1 to 160:1 DnaA:*oriC* ratios while *H. pylori* DnaA was used at up to 80:1 DnaA:*oriC* ratio in P1 (Donczew et al., 2012) and DMS assays (Donczew et al., 2014). We did not observe significant differences in concentration of HisCjDnaA and StrepCjDnaA required for specific unwinding of *C. jejuni oriC*, thus it was concluded that the tags did not interfere with *C. jejuni* DnaA unwinding or DNA binding activities. StrepCjDnaA was preferentially used in analyses that could require a free N-terminus [e.g., long-range interactions in electron microscopy (EM) or putative interactions with Cj1509]. In all analyses, DnaA was supplemented with 3 mM ATP (EM) or 5 mM ATP (footprinting and P1 nuclease assay).

Plasmid pET28Cj1509 was constructed by ligating the PCR-amplified gene using primer pair B4-B5 into BamHI/SalI-digested pET28a(+). Cj1509 protein expression was induced in *E. coli* BL21 by the addition of IPTG to a final concentration of 0.05 mM, followed by an additional incubation for 3 h at 30°C. Cells were centrifuged, resuspended in His-A buffer (50 mM NaH$_2$PO$_4$, 300 mM NaCl, pH 8.0), sonicated and centrifuged (30 min, 15,000 × *g*, 4°C). The supernatant was incubated for 1 h at 4°C with HIS-Select Nickel Affinity Gel (Sigma-Aldrich), washed twice with His-A buffer supplemented with 10 mM imidazole and eluted with His-A buffer with increasing concentration of imidazole (20–200 mM).
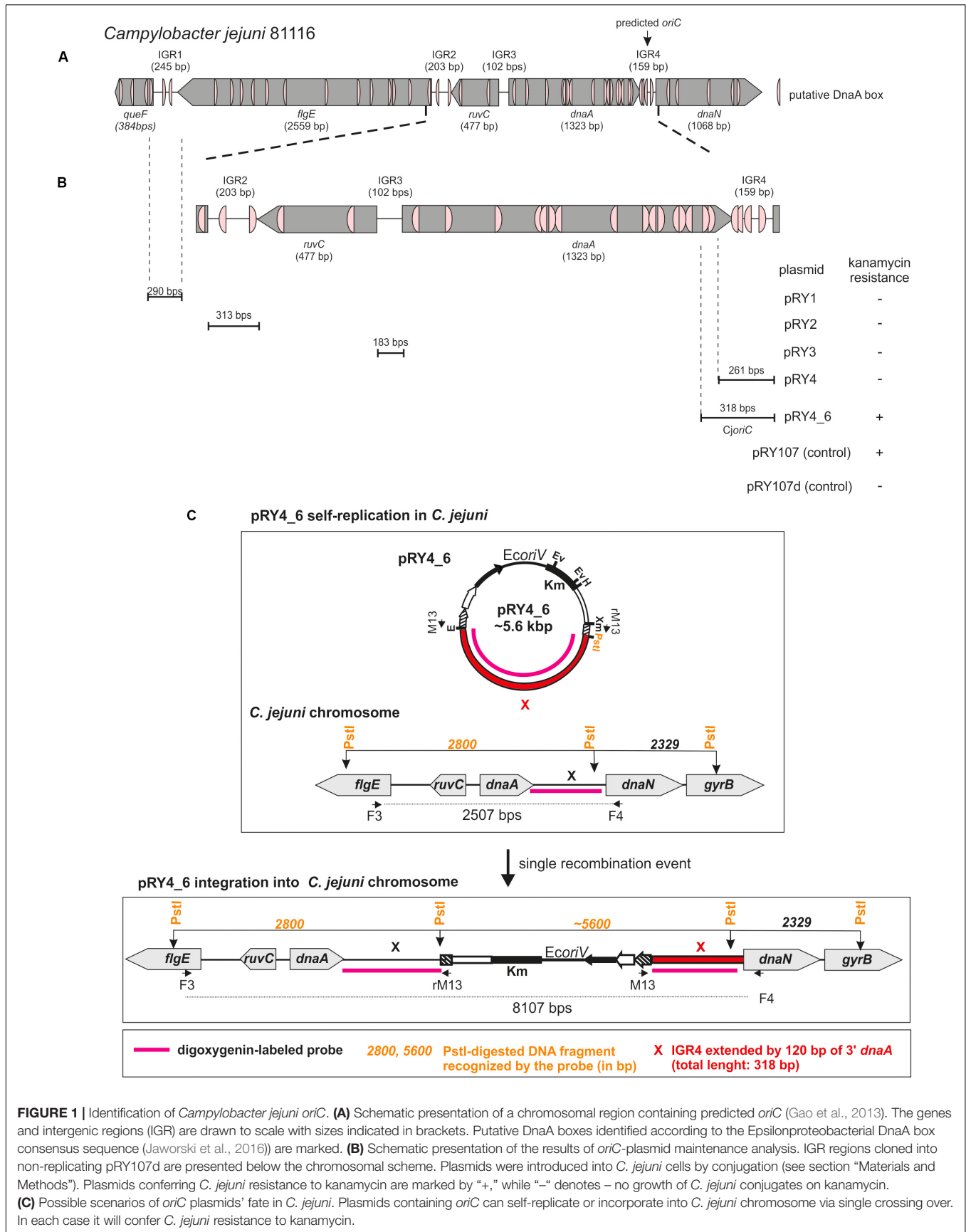
FIGURE 1 | Identification of *Campylobacter jejuni oriC*. **(A)** Schematic presentation of a chromosomal region containing predicted *oriC* (Gao et al., 2013). The genes and intergenic regions (IGR) are drawn to scale with sizes indicated in brackets. Putative DnaA boxes identified according to the Epsilonproteobacterial DnaA box consensus sequence (Jaworski et al., 2016)) are marked. **(B)** Schematic presentation of the results of *oriC*-plasmid maintenance analysis. IGR regions cloned into non-replicating pRY107d are presented below the chromosomal scheme. Plasmids were introduced into *C. jejuni* cells by conjugation (see section "Materials and Methods"). Plasmids conferring *C. jejuni* resistance to kanamycin are marked by "+," while "−" denotes – no growth of *C. jejuni* conjugates on kanamycin. **(C)** Possible scenarios of *oriC* plasmids' fate in *C. jejuni*. Plasmids containing *oriC* can self-replicate or incorporate into *C. jejuni* chromosome via single crossing over. In each case it will confer *C. jejuni* resistance to kanamycin.

DMS Footprinting

In vitro DNA modification was performed as described previously (Cassler et al., 1995; Donczew et al., 2014) in a concentration range of 0.4–1.6 μM or 0.5–2.1 μM for 6HisCjDnaA and 6HisCj1509, respectively. Methylated pOC_24 and pOC_IGR4 were used as a template for primer extension (PE) reactions with appropriate primers (see Supplementary Table S2).

P1 Nuclease Assay

The P1 nuclease assay was conducted as previously described (Donczew et al., 2012). Reaction mixtures contained 300 ng of pOC_IGR plasmid DNA (approximately 10 nM) and 6HisCjDnaA protein (up to 1.6 μM), or a mixture of StrepCjDnaA (up to 0.8 μM) and 6HisCj1509 (up to 1.6 μM), in a total volume of 15 μl. P1 activity was analyzed by SspI restriction enzyme digestion and 1% agarose gel separation or PE analysis. The gels were scanned with a GelDoc Xr+ Imaging System (Bio-Rad).

Primer Extension (PE) Reactions

The modification sites introduced either by DMS or P1 nuclease were monitored by PE analysis. Reaction conditions, mixture separation and product visualization were conducted as described previously (Jaworski et al., 2016).

Electrophoretic Mobility Shift Assay (EMSA)

Electrophoretic mobility shift assay was conducted as described previously (Jaworski et al., 2016). The IGR regions were amplified by PCR using IRD800 labeled E1-E2 or FAM-labeled E5-E6 primer pairs (Supplementary Table S2) and specific template pOC_X for IGRX to give IRD800-IGRX or FAM-IGRX (X-respective IGR region, pOC plasmids are described in Supplementary Table S1). The NC control region was amplified by PCR using IRD800 labeled E3–E4 primer pair and specific template pTZ_NC (Supplementary Tables S1, S2) to give IRD800-NC. Probes representing Cj1509 boxes were designed as described previously (Jullien and Herman, 2011). Oligonucleotides (Supplementary Table S2), suspended in H₂O, were mixed in equimolar concentration (33.3 μM each, in threes: E7–E8–E9 for Cj1509 box 1, E7–E10–E11 Cj1509 box 2 and E7–E12–E13 Cj1509 box 3) and hybridized. 0.4 μM of the complete annealed probe was used in gel shift analyses. IRD800- or FAM-labeled DNA fragments were incubated with 6HisCjDnaA protein (up to 30 nM) or 6HisCj1509 (up to 400 nM) at room temperature for 15 min. Bound complexes were resolved on 4% polyacrylamide gels (1x TBE at 7.5 V/cm) and visualized on an Odyssey CLx Infrared Imaging System (Li-Cor) or Typhoon FLA9500 Variable Mode Imager (GE Healthcare).

Electron Microscopy (EM)

Electron microscopy was performed as described previously (Spiess and Lurz, 1988; Donczew et al., 2012, 2014) with few modifications. 90 ng (approximately 110 nM) of StrepCjDnaA protein was incubated with 60 ng (approximately 1.4 nM) of pOC_24 plasmid DNA. Complex localization was measured using ImageJ 1.46v software (Schneider et al., 2012). To calculate the binding and distribution of protein, approximately 300 DNA molecules were analyzed.

In Silico Analysis

The prediction of *oriC*-type replication origins in the *C. jejuni* 81116 chromosome was performed using a stepwise procedure, as described previously (Donczew et al., 2012; Jaworski et al., 2016). The DnaA box consensus sequence was generated by WebLogo3 (Schneider and Stephens, 1990; Crooks et al., 2004). The DnaA and Cj1509 box search was performed by Pattern Locator (Mrázek and Xie, 2006). DnaA alignment was prepared by Praline (Bawono and Heringa, 2014).

RESULTS

We identified the probable *C. jejuni* origin of chromosome replication and characterized its three basic modules: DUE, DnaA boxes and OriBP binding sites in a three-step approach: *in silico* analysis of putative origins, analyses of mini-chromosome replication *in vivo* and DnaA-DNA and Cj1509-DNA interactions *in vitro*.

C. jejuni oriC Is Located Downstream of *dnaA*

DoriC predicted *C. jejuni* 81116 *oriC* for pos. 1324–1482 [ORI92240122] (i.e., *dnaA–dnaN* intergenic region) (Gao et al., 2013). According to this database, the predicted *oriC* contains two DnaA boxes highly similar to *E. coli* perfect DnaA box (no more than one mismatch from 5′-TTATCCACA-3′). However, Epsilonproteobacterial DnaA boxes differ from the perfect *E. coli* DnaA box (Jaworski et al., 2016). Thus, we searched for *E. coli* consensus DnaA box sequence (5′-TTWTNCACA-3′) allowing for two mismatches (Schaper and Messer, 1995). The search also filtered for the presence of a 5′-TCAC-3′ (5–8 nt) sequence to meet more stringent Epsilonproteobacterial criteria (Jaworski et al., 2016). We found a cluster of four DnaA boxes at predicted *oriC* and the putative DUE sequence downstream of the last DnaA box (**Figure 1A** and Supplementary Figure S1). Significantly, a degenerated 'DnaA trio' motif (5′-TAG-3′) (Richardson et al., 2016) was found between the DnaA box cluster and the putative DUE. Due to these structural similarities to other *oriC*s of Epsilonproteobacteria, we assumed that replication started at *dnaA–dnaN* intergenic region. However, numerous DnaA boxes were also found in the intergenic regions upstream of *dnaA* (intergenic regions were denoted IGR1–IGR3, *dnaA–dnaN* was denoted IGR4 for clarity of description, **Figure 1** and Supplementary Figure S1). Since the known Epsilonproteobacterial *oriC* regions are bipartite (Jaworski et al., 2016), IGR4-proximate IGR2 or less likely IGR4-distal IGR1, which contain predicted DnaA boxes, may be involved in the initiation of *C. jejuni* replication similarly to Epsilonproteobacterial *oriC1* (Jaworski et al., 2016).

Thus, we analyzed the functionality of all four selected IGRs as putative *oriC* (sub-)regions in *C. jejuni*. We used

a minichromosome approach, which has been successfully applied to identify or to characterize chromosomal origins of replication of several bacterial species, for example, *E. coli*, *Streptomyces coelicolor*, and *B. subtilis* (Yasuda and Hirota, 1977; Moriya et al., 1992; Zakrzewska-Czerwińska et al., 1995). The minichromosome is a plasmid that contains a chromosomal *oriC* as sequence that supports the plasmid's autonomous replication in a cell of a given species (Dasgupta and Løbner-Olesen, 2004). In addition, minichromosomes contain selection markers and may contain sequences for propagation of a plasmid (plasmid *oriV*) in a heterologous host strain, which is usually *E. coli*. To study the *C. jejuni oriC*, we used a derivative of a shuttle *E. coli–C. jejuni* plasmid pRY107 as a cloning vector (Yao et al., 1993). We removed Cj*oriV* supporting pRY107 replication in *C. jejuni* and obtained pRY107_d, which was incapable of replicating in *C. jejuni* but contained E*coriV* and, thus, could still replicate in *E. coli* (**Figure 1B** and Supplementary Figure S2). pRY107_d was further used for cloning of the IGRs previously determined by *in silico* analyses as putative *C. jejuni oriC* regions (see section "Materials and Methods"). A series of plasmids was obtained containing the intergenic regions IGR1–IGR4 (pRY1, pRY2, pRY3, pRY4, respectively) (**Figure 1B**, Supplementary Table S1 and Supplementary Figure S2). The plasmids were introduced into *C. jejuni* 81–176 via conjugation (see section "Materials and Methods"); pRY107 was introduced in a parallel conjugation and served as a positive control for conjugation, replication and selection in *C. jejuni*. None of the cloned IGR regions supported the replication of the plasmids in *C. jejuni* because no colonies were obtained after conjugation. Conversely, the control pRY107 plasmid replicated in *C. jejuni* because kanamycin-resistant *C. jejuni* conjugant colonies grew on selective plates. The *in silico* analysis indicated that there were numerous DnaA boxes within the *dnaA* gene and other IGR regions, which could be necessary to support IGR4 activity (**Figure 1A**). Therefore, we prepared a series of plasmids, all of which contained IGR4 but differed in the length of DNA extending upstream of IGR4, up to IGR2 (pRY4_6 and data not shown, Supplementary Table S1 and **Figures 1B,C**). Plasmid pRY4_6, which contained IGR4 extended by 120 bp of 3′ of *dnaA* encompassing two additional putative DnaA boxes, was the shortest construct, which, when introduced into *C. jejuni* by conjugation, successfully supported *C. jejuni* growth on selective kanamycin plates. We confirmed by PCR that the genomic DNA isolated from *C. jejuni* pRY4_6 conjugants contained the pRY4_6 plasmid (Supplementary Figure S2B). We also excluded by PCR the possibility that residual *E. coli* DNA contaminated isolated *C. jejuni* genomic DNA (Supplementary Figure S2B). Thus, the pRY4_6 plasmid detected by PCR was carried by *C. jejuni* conjugants. pRY4_6 could be carried by *C. jejuni* as a self-replicating plasmid or could have been integrated into *C. jejuni* chromosome (**Figure 1C**). To confirm that pRY4_6 self-replicated in *C. jejuni,* we isolated pRY4_6 plasmid from *C. jejuni*, transformed *E. coli* with this plasmid, re-isolated it from transformants and digested the extracted plasmid by EcoRI and PstI. The restriction pattern confirmed the authenticity of pRY4_6 (**Figure 2A**). Finally, we performed Southern blot to confirm the presence of self-replicating pRY4_6 plasmid in *C. jejuni*. We resolved undigested and PstI-digested

C. jejuni genomic DNA and the pRY4_6 plasmid in 1% agarose gel (Supplementary Figure S3) and transferred DNA onto a nylon membrane. The membrane was then incubated with the digoxigenin-labeled, 314 bp-DNA probe, corresponding to the sequence of IGR4 extended by approximately 120 bps of the 3′ *dnaA* sequence (**Figure 1C**). In undigested genomic DNA the probe detected IGR4 within high-molecular weight *C. jejuni* chromosomal DNA, both in *C. jejuni* wild type and *C. jejuni* pRY4_6 conjugant (**Figures 2B,C**). In pRY4_6 conjugant strain, but not in the wild type strain, the additional single band corresponding to intact, self-replicating plasmid DNA was detected. The intensities of bands representing the self-replicating pY4_6 plasmid were low when compared to the signal detected within undigested *C. jejuni* chromosomal DNA, which suggested that only a fraction of cells maintained the self-replicating plasmid, while in majority of *C. jejuni* cells the plasmid recombined with the chromosome. In PstI-digested genomic DNA, in *C. jejuni* wild type strain, the probe detected a single DNA band corresponding to wild type IGR4 genomic locus. In *C. jejuni* pRY4_6 conjugant, the probe detected two bands corresponding to wild type IGR4 genomic locus and pRY4_6 plasmid which integrated into the *C. jejuni* chromosome. Molecular weight and approximately 1:1 stoichiometry of detected bands suggested that recombination occurred via single crossing over (**Figure 1C**). The PCR reaction, performed using F3–F4 primer pair (**Figure 1C**), confirmed the presence of both intact and recombined *oriC* loci in DNA isolated from *C. jejuni* pRY4_6 conjugants (Supplementary Figure S2C), which further confirmed the presence of two populations of *C. jejuni* cells. Altogether, the results showed that *C. jejuni* pRY4_6 is unstable and tends to integrate into the *C. jejuni* chromosome, which is a known and common phenomenon observed for minichromosomes in diverse bacterial species (see section "Discussion"). Nonetheless, since pRY4_6 was re-isolated from *C. jejuni* and was detected as a self-replicating plasmid in southern blot, we conclude, that the a 318 bp fragment containing the intergenic region between *dnaA* and *dnaN* (IGR4) and two predicted DnaA boxes located in the 3′ region of the *dnaA* gene is sufficient to maintain replication of mini-chromosomes. We named this region probable *C. jejuni oriC* (Cj*oriC*) (**Figure 2D**).

C. jejuni DnaA Unwinds Cj*oriC* in Vitro

In the next step, we assessed whether the identified Cj*oriC* region is unwound by DnaA *in vitro* since the presence of the region that undergoes specific DnaA-dependent unwinding is the most unequivocal *in vitro* indication of *oriC* functionality. To experimentally validate DNA unwinding and to identify the DnaA-dependent DUE position in predicted Cj*oriC*, P1 nuclease assay was applied (Sekimizu et al., 1987). A plasmid containing the IGR4 region was constructed (pOC_IGR4), and the *C. jejuni* DnaA protein was purified (see section "Materials and Methods," **Figure 3A**, Supplementary Table S1 and Supplementary Figure S4). As a control, we constructed a pOC_IGR2 plasmid containing IGR2 (Supplementary Table S1 and **Figure 3A**), which was predicted to exhibit significant helical instability (data not shown), however, it was unable to support minichromosomal replication in *C. jejuni* (**Figure 1B**). The supercoiled plasmids
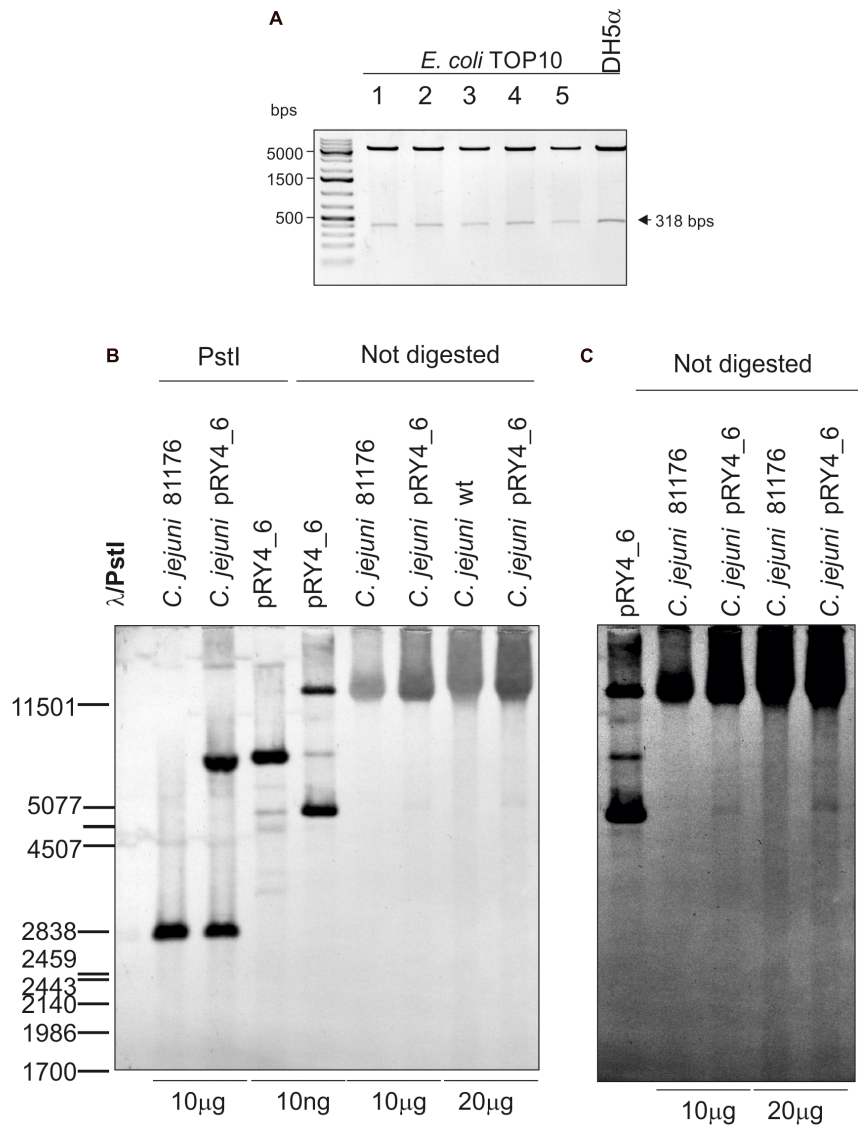
FIGURE 2 | Analysis of pRY4_6 self-replication in *C. jejuni*. **(A)** Gel analysis of plasmids isolated from *E. coli* TOP10 cells transformed with genomic DNA of *C. jejuni* pRY4_6 conjugant strain. Plasmids were digested by EcoRI and PstI. The plasmid propagated in *E. coli* DH5a is shown as a control. **(B)** Southern blot analysis of *C. jejuni* pRY4_6 conjugant strain. *C. jejuni* 81176 wild type and *C. jejuni* pRY4_6 conjugant strains genomic DNA and the pRY4_6 plasmid, isolated from *E. coli*, undigested or digested with PstI, were resolved in 1% agarose gel (Supplementary Figure S3), transferred onto a nylon membrane and probed with the digoxigenin-labeled DNA probe (**Figure 1C**). Southern blot was developed by anti-digoxigenin antibody followed by colorimetric reaction. **(C)** The fragment of the blot, presenting probe hybridisation to not digested DNA, was digitally manipulated to increase intensity of bands corresponding to the self-replicating pRY4_6 and to reduce the background. Adjustment was applied to the whole image. **(D)** Nucleotide sequence of *C. jejuni* DNA, which enabled self-replication of the pRY4_6 minichromosome. The sequence in bold presents DNA region cloned into pRY4, which did not replicate in *C. jejuni*.

FIGURE 3 | Identification of the DUE in *C. jejuni oriC*. **(A)** Supercoiled plasmids pOC_IGR2 and pOC_IGR4 were incubated with the indicated amounts of 6HisDnaA and successively digested by P1 nuclease and SspI. Digested plasmids were resolved in 1% agarose gel, and the fragments resulting from DnaA-dependent unwinding are indicated by red arrowheads next to the gel, while the DNA fragments resulting from DnaA independent unwinding within the plasmid pBR322 *ori* site are indicated by green arrows. A map of the pOC_IGR plasmids used in the P1 nuclease assay is presented below the gel. The IGR regions, pBR322 plasmid origin of replication and positions of the most important restriction sites are marked. P1-sensitive DnaA-dependent unwinding is distinguished from DnaA-independent by red and green lines, respectively. **(B)** The supercoiled plasmid pOC_IGR4 was incubated with the indicated amounts of 6HisDnaA, digested by P1 nuclease and used as a template for PE reactions with ^{32}P-labeled primers C1 and C2. Dashed lines on the right of the PE gel indicate the nucleotides susceptible for the P1 nuclease treatment, while the boundaries of the DUE are marked with continuous-line arrows next to the presented sequences. The sequence of the DUE (dashed line, yellow shadowing), the degenerated DnaA-trio motif (solid line) and predicted DUE-proximal DnaA box (light-red) are presented below the PE gel. Yellow shadowing is used to mark DUE in subsequent figures.

were incubated with increasing amounts of DnaA protein and digested with P1 nuclease to cleave the resulting single-stranded DNA regions. Subsequently, site-specific digestion by SspI excised the DNA fragment from the plasmid. The size of the fragment allowed us to estimate the position of a region unwound by DnaA. In the case of pOC_IGR4, DNA fragments of approximately 1000 and 1400 bp were excised by P1/SspI, indicating specific formation of a single-stranded DNA within the *CjoriC* region (**Figure 3A**). pOC_IGR2 was only linearized by SspI. Thus, DnaA-dependent unwinding occurred in IGR4, but not in IGR2 (**Figure 3A**). IGR4 represents Cj*oriC* lacking the two DnaA boxes distal to DUE (**Figure 2D**), but it was previously shown that not all DnaA boxes are required for DUE unwinding

in the *oriC* cloned into a plasmid (Ozaki and Katayama, 2012; Donczew et al., 2014; Sakiyama et al., 2017). Regardless of the DnaA presence or concentration, all lanes contained additional DNA fragments of 400 and 2000 bp because the plasmids were also unwound at a site corresponding to the plasmid origin of replication (Kowalski et al., 1988). In addition, a fraction of pOC_IGR4 molecules was unwound simultaneously at IGR4 and plasmid origin, which resulted in three DNA fragments: one 400 bp and two 1000 bp. Therefore, the overall intensity of the 1000 bps DNA band on the gel was higher than that of 1400 bp band (**Figure 3A**).

To precisely determine the unwound regions, primer extension (PE) reactions with ^{32}P-labeled primers were

performed using the P1-digested pOC_IGR4 plasmid template (**Figure 3B**; the primers are specified in Supplementary Table S2). The primers hybridized to the template DNA approximately 40–80 bp away from the *in silico*-predicted DUE region within IGR4, which was extended by Taq polymerase until the P1 nuclease digestion site was encountered. The detailed PE analysis confirmed that the IGR4 region underwent DnaA-dependent unwinding, and thus it contained the DUE sequence (**Figure 3B**). The DUE encompasses approximately 35 bps and contains 19% GC residues (overall chromosomal GC content is 30.5%). The main part of the identified DUE region is an AT-rich region, which is a typical feature of bacterial origins. Analysis of the DUE sequence did not reveal any repeats similar to 13-mer *E. coli* L, M, R repeats. We detected a degenerate DnaA-trio (5′-TAG-3′), but no GC-rich region was found in the *C. jejuni oriC* region (Richardson et al., 2016).

Taken together, the above *in vivo* (minichromosomes) and *in vitro* (P1 plasmid unwinding) data indicate that IGR4 extended by approximately 120 bps of a 3′ region of *dnaA* is probably the functional *C. jejuni oriC.*

DnaA Boxes Are Present in *oriC* and in the *oriC*-Vicinal Region

The DNA is unwound by the DnaA protein bound to DnaA boxes at *oriC*. The *in silico* analysis predicted numerous DnaA binding sites in the vicinity of Cj*oriC* (**Figure 1A** and Supplementary Figure S1). Accordingly, the gel shift results indicated that IRD800-labeled DNA fragments comprising IGR1, IGR2, and IGR4, for which the DnaA boxes were predicted, were bound by DnaA, while no binding was observed for a fragment comprising IGR3, which contained only 1 predicted, apparently non-optimal DnaA box sequence (Supplementary Figure S5). The number of distinct nucleoprotein complexes that formed between DnaA and IRD800-IGR4 was higher than between DnaA and IRD800-IGR2 or IRD800-IGR1, indicating a higher number of DnaA binding sites at IRD800-IGR4 than at IRD800-IGR2 and IRD800-IGR1.

Therefore, the next step was to precisely identify DnaA boxes at *C. jejuni oriC* by dimethyl sulfate (DMS) footprinting (Sasse-Dwight and Gralla, 1991; Cassler et al., 1995; Shaw and Stewart, 2009) and to establish the *C. jejuni* DnaA box consensus sequence. Briefly, DMS methylates guanine or adenine residues. DMS footprinting allows the detection of DNA sequences that are bound by a protein and thus protected by the protein against guanine or adenine methylation. These protected residues are distinguishable as DNA bands with a decreased intensity on a footprinting gel. The pOC_24 plasmid (Supplementary Table S1) containing the IGR2-*ruvC*-IGR3-*dnaA*-IGR4 region was incubated with increasing concentrations of the *C. jejuni* DnaA protein and methylated by DMS. To determine protein binding sites, sets of primers (Supplementary Table S2), that were complementary to the upstream regions of predicted DnaA boxes were used in PE reactions. We detected multiple G residues that were protected by DnaA in IGR4 (**Figures 4A–C** and Supplementary Figure S6). The subsequent comparison of the DNA sequences in the vicinity of protected G residues allowed to classify the identified DnaA binding sites as DnaA

boxes, because they resembled DnaA boxes characterized in other bacterial species (Wolański et al., 2014) (Supplementary Figure S7) (see also section "Discussion"). In total we identified five DnaA boxes in the IGR4 intergenic region and two DnaA boxes enclosed within the 3′ end of the *dnaA* gene that were essential to support minichromosome replication in *C. jejuni* as discussed above (**Figure 2D**). Thus, the *C. jejuni oriC* contained seven DnaA boxes bound by DnaA *in vitro,* from which one DnaA box (5′-AATTTCAAC-3′, DnaA box 1) was not predicted *in silico*, because of absence of the Epsilonproteobacterial core sequence (Supplementary Figure S6B). The *oriC* region preserved the general features of a typical bacterial origin of replication, namely, the distance (approximately 1–2 helical turns) and spatial orientation between the DUE and CjDnaA box 1 (**Figure 4G**) and the fact that CjDnaA box 1 is accompanied by a second DnaA box 2 in head-to-tail orientation. This array has been proposed to be essential for the formation of a functional orisome and DUE unwinding in a few bacterial species including *H. pylori* (Donczew et al., 2014). The unique feature of *C. jejuni oriC* is the possible requirement for DnaA boxes located in the *dnaA* gene for the activity of *oriC in vivo* (see section "Discussion").

We additionally analyzed DnaA binding to two other regions outside *oriC* at which DnaA boxes were predicted: IGR2, which is supposed to contain two DnaA boxes and the middle region of *dnaA* in which four DnaA boxes were predicted (Supplementary Figures S1, S6B). The DMS footprinting analysis confirmed DnaA binding to four of those predicted DnaA boxes (DnaA binding sites 9–11 and 13), while it allowed to identify two new DnaA binding sites: 8 (5′-AACTGCACA-3′) and 12 (5′-TATTACACA-3′), which were not predicted due to deviation from the core of the typical Epsilonproteobacterial DnaA box (**Figures 4D–F** and Supplementary Figure S6). Please note, that our analysis do not preclude further binding of DnaA to other oppositely oriented or non-clustered DnaA boxes, not detected by using by using a single PE primer. Nonetheless DMS footprinting and *in silico* analyses indicated a high number of DnaA binding sites in the *oriC*-proximal region (**Figure 4G**); however, they were more scattered in regions outside of *oriC* than those enclosed within *C. jejuni oriC,* with the exception of DnaA boxes in the middle and in the 3′ region of *dnaA* (see also below). The sequences of 13 *in vitro*-determined DnaA binding sites were assembled to generate a logo of the *C. jejuni* DnaA box sequence, 5′-NHHWDCAMH-3′ (Supplementary Figure S7), with the majority of boxes at *oriC* following the more stringent Epsilonproteobacterial DnaA box core consensus 5′-WWHTTCACW-3′ sequence (**Figure 4H** and Supplementary Figure S7) (see also section "Discussion").

The DnaA-*oriC* Nucleoprotein Complex Engages *oriC*-Proximal DNA

C. jejuni oriC is most likely monopartite. However, the results presented thus far indicated the presence of numerous DnaA binding sites in *oriC*-vicinal regions, such as *ruvC, flgE,* and *dnaA* genes or the IGR1, IGR2, and IGR4 intergenic regions (**Figures 1A**, **4G**). The binding of DnaA to IGR1 and IGR2 intergenic regions was confirmed by gel shift and footprinting
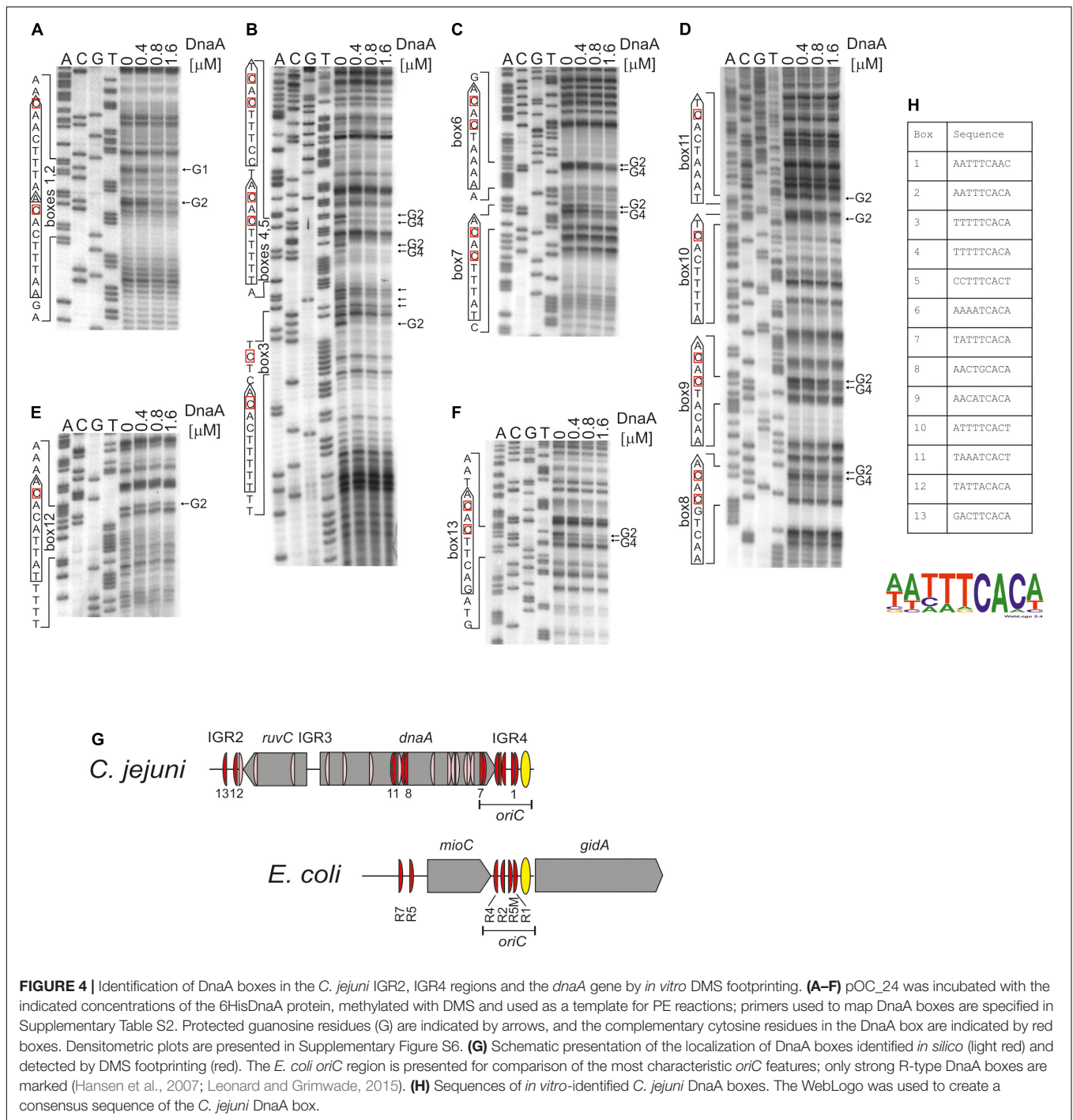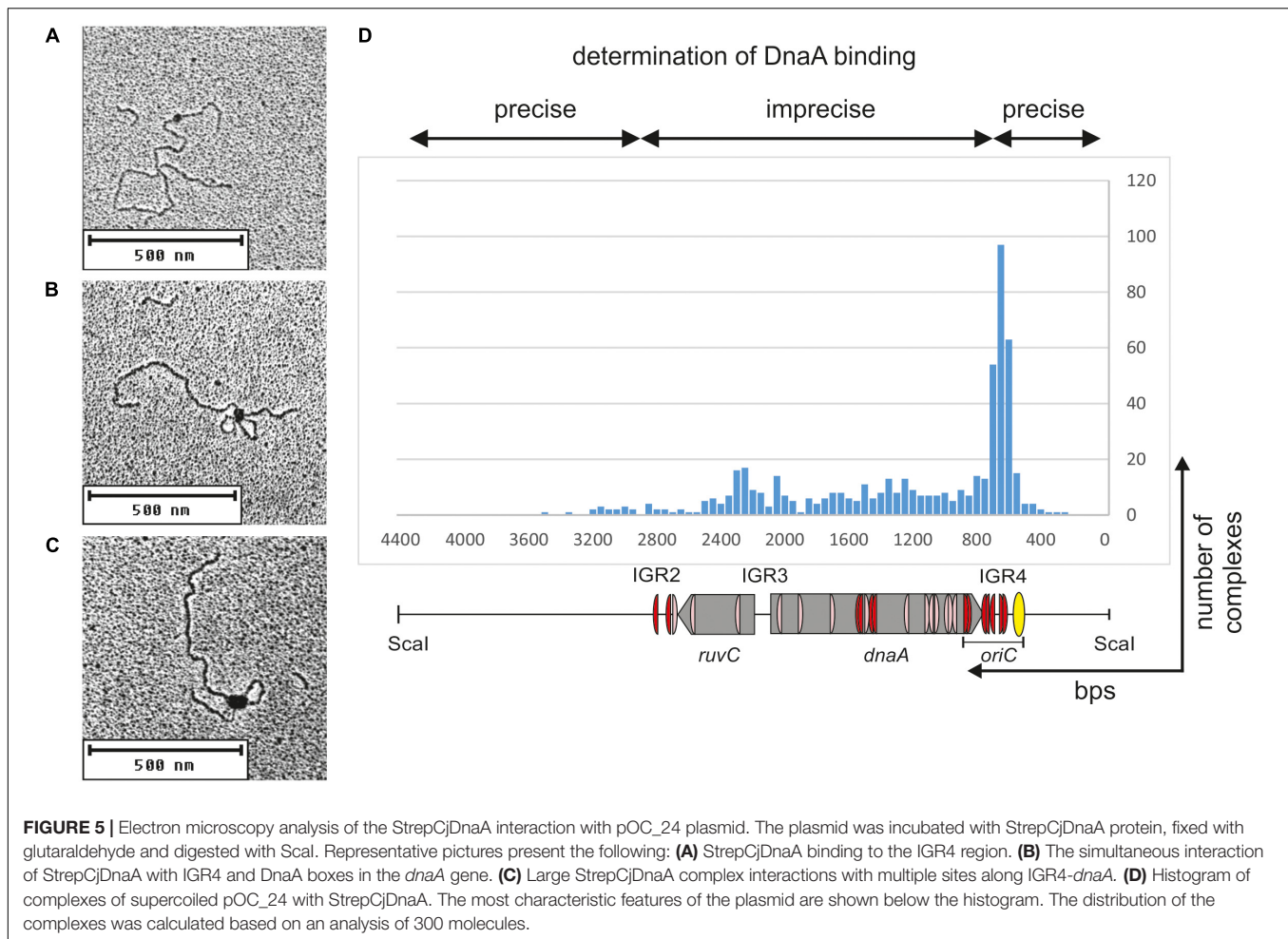
FIGURE 4 | Identification of DnaA boxes in the *C. jejuni* IGR2, IGR4 regions and the *dnaA* gene by *in vitro* DMS footprinting. **(A–F)** pOC_24 was incubated with the indicated concentrations of the 6HisDnaA protein, methylated with DMS and used as a template for PE reactions; primers used to map DnaA boxes are specified in Supplementary Table S2. Protected guanosine residues (G) are indicated by arrows, and the complementary cytosine residues in the DnaA box are indicated by red boxes. Densitometric plots are presented in Supplementary Figure S6. **(G)** Schematic presentation of the localization of DnaA boxes identified *in silico* (light red) and detected by DMS footprinting (red). The *E. coli oriC* region is presented for comparison of the most characteristic *oriC* features; only strong R-type DnaA boxes are marked (Hansen et al., 2007; Leonard and Grimwade, 2015). **(H)** Sequences of *in vitro*-identified *C. jejuni* DnaA boxes. The WebLogo was used to create a consensus sequence of the *C. jejuni* DnaA box.

assays (Supplementary Figure S5 and **Figure 4**). The results suggested that the nucleoprotein complex formed by *C. jejuni* DnaA might extend beyond the Cj*oriC* region required to support replication of the minichromosome (pRY4_6, **Figure 1B**). Such auxiliary binding sites may play regulatory roles in the initiation of *C. jejuni* chromosome replication or indicate further functions of DnaA, for example, in chromosome maintenance or structure. Thus, to better characterize DnaA binding to *oriC* proximal regions and, especially, to monitor intermolecular interactions

between DnaA, we further analyzed the binding of DnaA to a plasmid that contained the entire region between *flgE* and *dnaN* by electron microscopy (EM) (**Figure 5**). The supercoiled pOC_24 plasmid that contained IGR2-*ruvC*-IGR3-*dnaA*-IGR4 (Supplementary Table S1) was incubated with the StrepCjDnaA protein. The StrepCjDnaA protein contained a C-terminal Strep-tag and a native N-terminus, which might be essential for long-range protein–protein interactions (Zawilak-Pawlik et al., 2017). The nucleoprotein complexes were subsequently stabilized by

FIGURE 5 | Electron microscopy analysis of the StrepCjDnaA interaction with pOC_24 plasmid. The plasmid was incubated with StrepCjDnaA protein, fixed with glutaraldehyde and digested with ScaI. Representative pictures present the following: **(A)** StrepCjDnaA binding to the IGR4 region. **(B)** The simultaneous interaction of StrepCjDnaA with IGR4 and DnaA boxes in the *dnaA* gene. **(C)** Large StrepCjDnaA complex interactions with multiple sites along IGR4-*dnaA*. **(D)** Histogram of complexes of supercoiled pOC_24 with StrepCjDnaA. The most characteristic features of the plasmid are shown below the histogram. The distribution of the complexes was calculated based on an analysis of 300 molecules.
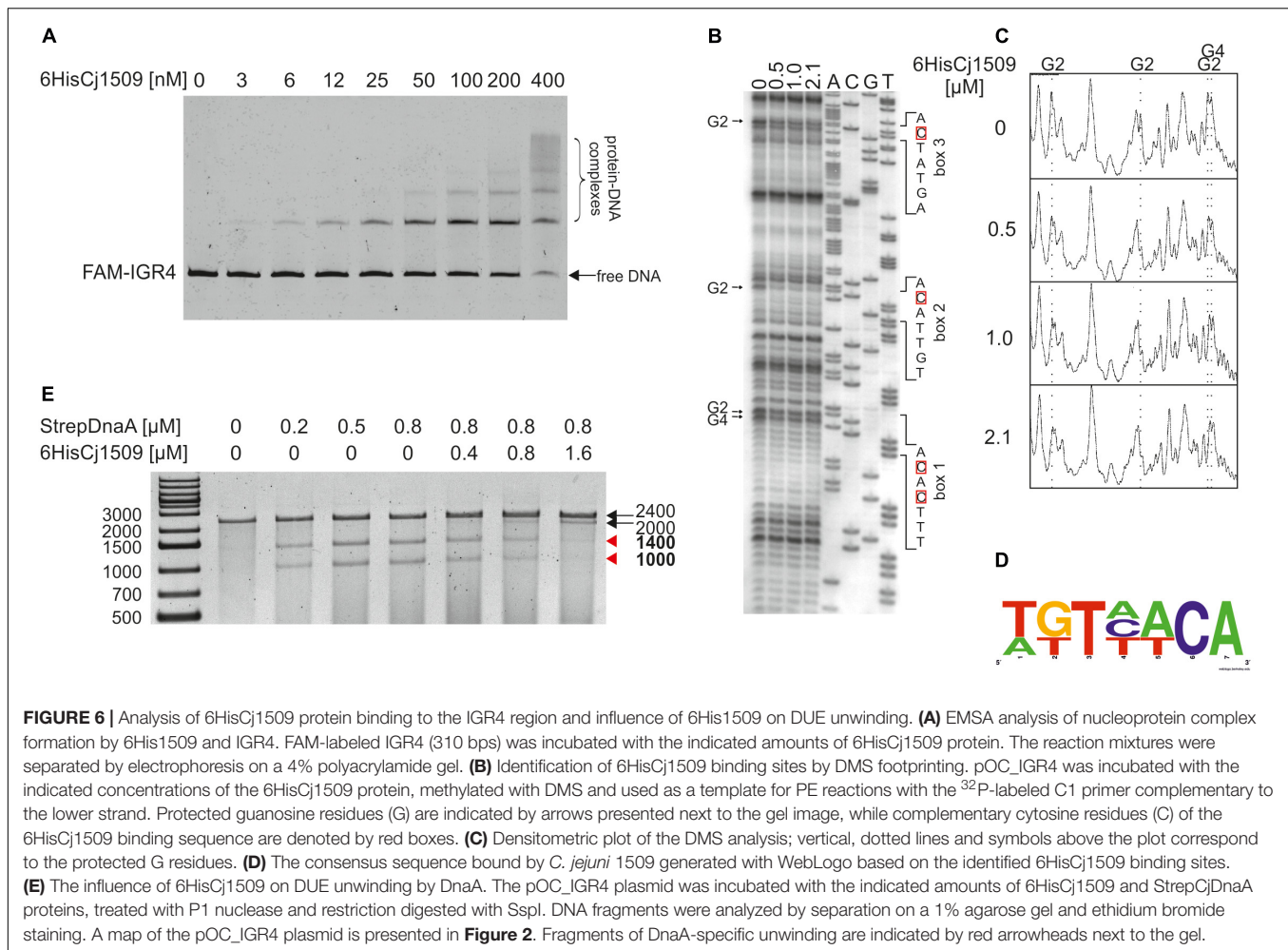
glutaraldehyde crosslinking and digested by ScaI to linearize the plasmid molecules. EM analysis revealed that the majority (94%) of the analyzed plasmid molecules were bound by DnaA (**Figure 5**). Two predominant kinds of nucleoprotein complexes were formed: (i) plasmid molecules with a single protein complex bound to a single plasmid region (type 1 complexes, **Figure 5A**) and (ii) looped DNA structures with a relatively large protein complex (or complexes) bound to at least two separated DNA regions (type 2 complexes, **Figures 5B,C**).

The type 1 complexes constituted 37% of all bound molecules. The distance measurements confirmed the binding of DnaA to the *oriC* region in 52% of the single nucleoprotein complexes (19% of all complexes), while 15% of the type 1 protein complexes were formed by DnaA binding between *oriC* and IGR2 (5.6% of all complexes, **Figure 5D**). Additionally, 33% of the type 1 complexes (12% of all complexes) were formed outside of *C. jejuni* DNA (i.e., on a DNA of the vector pOC), and thus they should be considered non-specific.

The type 2 complexes constituted 63% of all bound molecules. The distance measurements between the plasmid ends and the protein core confirmed the simultaneous binding of DnaA to *oriC* and to the region located between IGR2 and *oriC* in 66% of the complexes (42% of all bound plasmid molecules, **Figures 5B,C**).

However, in the type 2 complexes, DnaA bound to and condensed a large portion of *ruvC*-IGR3-*dnaA*-*oriC* DNA, with only a short unbound DNA stretch being looped out into 1 or 2 short loops. Thus, only the borders of DnaA binding (i.e., the length of the plasmid between ScaI-IGR2 and *oriC*-ScaI) could be relatively precisely measured along a plasmid molecule. Approximately 30% of the type 2 complexes (19% of all complexes) were bound at IGR2-*ruvC*-IGR3-*dnaA* without binding to *oriC*, and 4% of the type 2 complexes (2% of all complexes) were formed in non-specific regions (**Figure 5D**).

Taken together, the EM analysis indicated that in two types of complexes, type 1 and type 2, DnaA exhibited a higher affinity or more stable binding toward *oriC* than toward other regions because 60% of the complexes involved *oriC* (**Figure 5D**). Nonetheless, the majority of the complexes also included other regions of *C. jejuni* DNA (36 and 63% of types 1 and 2 complexes, respectively). This result indicated that (i) *C. jejuni* DnaA bound to multiple sites along DNA and (ii) DnaA exhibited high dimerization or oligomerization potential, and probably the ability to establish long-range interactions. These two findings suggest DnaA activity in processes beyond initiation of *C. jejuni* chromosome replication (see section "Discussion").

FIGURE 6 | Analysis of 6HisCj1509 protein binding to the IGR4 region and influence of 6His1509 on DUE unwinding. **(A)** EMSA analysis of nucleoprotein complex formation by 6His1509 and IGR4. FAM-labeled IGR4 (310 bps) was incubated with the indicated amounts of 6HisCj1509 protein. The reaction mixtures were separated by electrophoresis on a 4% polyacrylamide gel. **(B)** Identification of 6HisCj1509 binding sites by DMS footprinting. pOC_IGR4 was incubated with the indicated concentrations of the 6HisCj1509 protein, methylated with DMS and used as a template for PE reactions with the ^{32}P-labeled C1 primer complementary to the lower strand. Protected guanosine residues (G) are indicated by arrows presented next to the gel image, while complementary cytosine residues (C) of the 6HisCj1509 binding sequence are denoted by red boxes. **(C)** Densitometric plot of the DMS analysis; vertical, dotted lines and symbols above the plot correspond to the protected G residues. **(D)** The consensus sequence bound by *C. jejuni* 1509 generated with WebLogo based on the identified 6HisCj1509 binding sites. **(E)** The influence of 6HisCj1509 on DUE unwinding by DnaA. The pOC_IGR4 plasmid was incubated with the indicated amounts of 6HisCj1509 and StrepCjDnaA proteins, treated with P1 nuclease and restriction digested with SspI. DNA fragments were analyzed by separation on a 1% agarose gel and ethidium bromide staining. A map of the pOC_IGR4 plasmid is presented in **Figure 2**. Fragments of DnaA-specific unwinding are indicated by red arrowheads next to the gel.
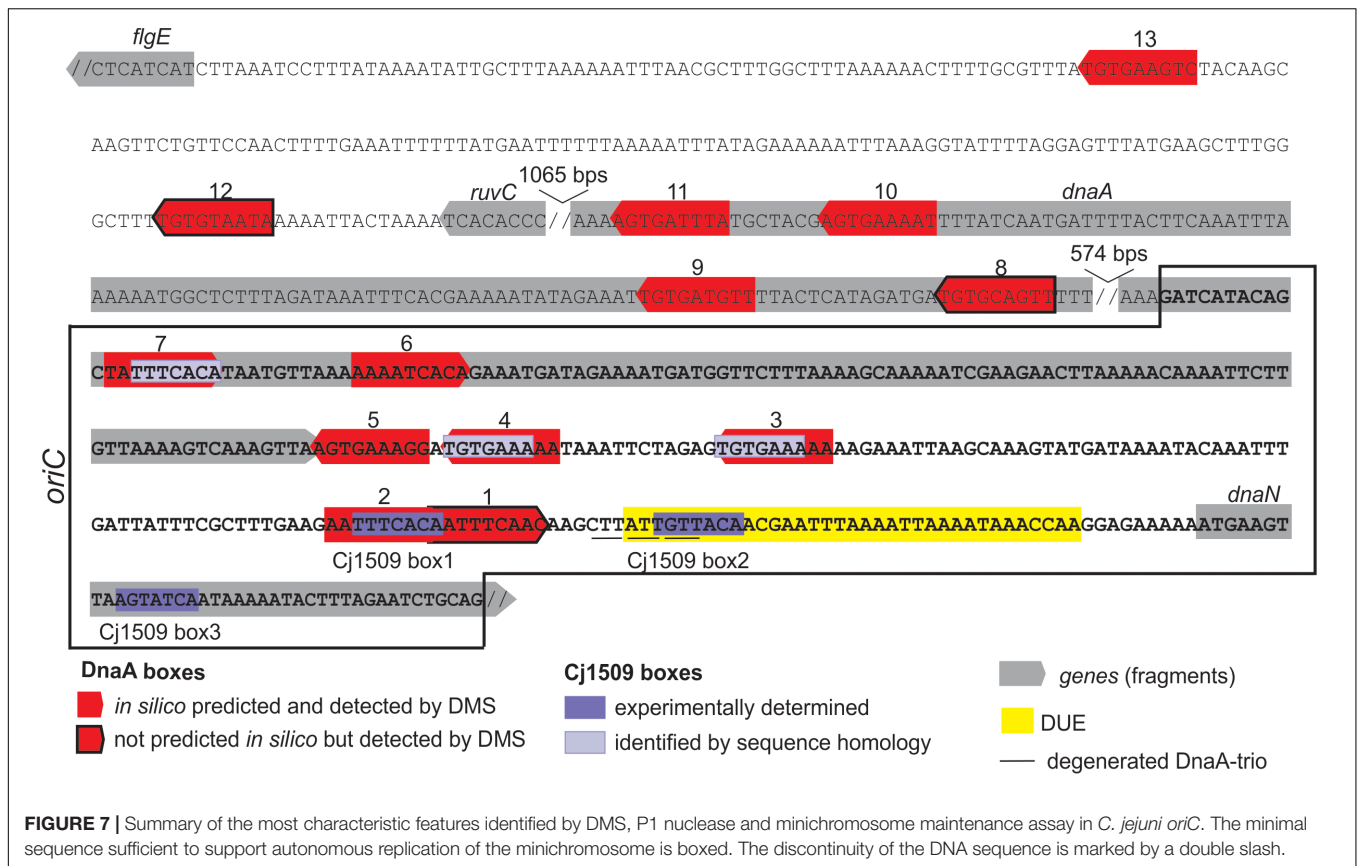
The Cj1509 Orphan Response Regulator Interacts With *CjoriC* and Inhibits DUE Unwinding

The orphan response regulator HP1021 was recently identified in *H. pylori* as a negative replication initiation regulator (Donczew et al., 2015). Since the two bacteria are related and the HP1021 protein is conserved and unique in Epsilonproteobacteria, we decided to analyze the role of the *C. jejuni* homolog Cj1509 in the initiation of *C. jejuni* chromosome replication.

To determine the affinity of Cj1509 for DNA, a gel-shift assay was conducted. For this purpose, a recombinant 6HisCj1509 protein was purified (see section "Materials and Methods," Supplementary Figure S4). A PCR-amplified *CjoriC* region was incubated with purified 6HisCj1509, and the complexes were resolved in a polyacrylamide gel (**Figure 6A**). The Cj1509 protein bound DNA *in vitro* and three nucleoprotein complexes were formed, suggesting either multiple Cj1509 binding sites or cooperative binding of Cj1509 with *oriC*. Therefore, we decided to precisely determine the sequences bound by Cj1509 at *oriC* by DMS footprinting. The pOC_IGR4 plasmid (Supplementary Table S1) was incubated with increasing amounts of

6HisCj1509 and methylated by DMS. Nucleotides protected by the Cj1509 protein were detected by subsequent PE using the C1 primer (Supplementary Table S2) complementary to the upstream region of the DUE (**Figure 6B**). The densitometric analysis revealed four G residues that were specifically protected by Cj1509 already at the lowest tested Cj1509 concentration (0,5 μM, **Figure 6C**); further increase in Cj1509 concentration did not greatly increase the protection. The DNA sequences in the vicinity of protected residues were assembled to generate a consensus sequence of the Cj1509 box 5′-WKTHWCA-3′ (**Figure 6D**), which is less stringent than that of the HP1021 box (5′-TGTTWCW-3′). To confirm binding of Cj1509 to each of the identified boxes gel shift assay was performed. Cj1509 was incubated with FAM-labeled probes representing each of the Cj1509 boxes and the complexes were resolved in a polyacrylamide gel (Supplementary Figure S8). The lower intensity of the complexes formed by Cj1509 with boxes 1 and 3 than that formed with the box 2 suggested that Cj1509 exhibits higher affinity toward the Cj1509 box 2 than toward boxes 1 or 3. Notably, the Cj1509 box 2 (5′-TGTTACA-3′) has the highest similarity (no mismatches) to the HP1021 box consensus sequence compared to Cj1509 box 1 (5′-TTTCACA-3′) or

flgE

//CTCATCATCTTAAATCCTTTATAAAATATTGCTTTAAAAAATTTAACGCTTTGGCTTTAAAAAACTTTTGCGTTTATGTGAAGTCTACAAGC

AAGTTCTGTTCCAACTTTTGAAATTTTTTATGAATTTTTTAAAAATTTATAGAAAAAATTTAAAGGTATTTTAGGAGTTTATGAAGCTTTGG

12   *ruvC*   1065 bps   **11**   **10**   *dnaA*

GCTTTTGTGTAATAAAAATTACTAAAATCACACCC//AAAAGTGATTTATGCTACGAGTGAAAATTTTATCAATGATTTTACTTCAAATTTA

9   **8**   574 bps

AAAAATGGCTCTTTAGATAAATTTCACGAAAAATATAGAAATTGTGATGTTTTACTCATAGATGATGTGCAGTTTTT//AAA**GATCATACAG**

7   **6**

CTATTTCACATAATGTTAAAAAAATCACAGAAATGATAGAAAATGATGGTTCTTTAAAAGCAAAAATCGAAGAACTTAAAAACAAAATTCTT

5   **4**   **3**

GTTAAAAGTCAAAGTTAAGTGAAAGGA**TGTGAAAAA**TAAATTCTAGAGT**GTGAAA**AAAAGAAATTAAGCAAAGTATGATAAAATACAAATTT

2   **1**   *dnaN*

GATTATTTCGCTTTGAAGATTTCACA**ATTTCAAC**AGCTTAT**TGTTACA**ACGAATTTAAAATTAAAATAAACCAAGGAGAAAAAATGAAGT

    Cj1509 box1   Cj1509 box2

oriC

TA**AGTATCA**ATAAAAATACTTTAGAATCTGCAG//

Cj1509 box3

DnaA boxes
- 🟥 *in silico* predicted and detected by DMS
- 🟥 (black border) not predicted *in silico* but detected by DMS

Cj1509 boxes
- 🟪 experimentally determined
- 🟦 identified by sequence homology

- ⬜ *genes* (fragments)
- 🟨 DUE
- — degenerated DnaA-trio

FIGURE 7 | Summary of the most characteristic features identified by DMS, P1 nuclease and minichromosome maintenance assay in *C. jejuni oriC*. The minimal sequence sufficient to support autonomous replication of the minichromosome is boxed. The discontinuity of the DNA sequence is marked by a double slash.

Cj1509 box 3 (5′-AGTATCA-3′) (each of which has two mismatches).

Two out of three Cj1509 boxes were located within sequences important for *oriC* activity: DnaA box 2 and DUE (**Figure 7**). The DnaA box 2 is probably important for assembly of a DnaA oligomer capable of DNA unwinding. Thus, overlapping DnaA and Cj1509 interaction sites suggested that these two proteins compete for binding to DNA. The 3′ DUE sequence is probably crucial for the formation of a complex between the DnaA filament and ssDNA. To analyze the influence of Cj1509 on DnaA-dependent DNA unwinding, we performed a P1 nuclease test. DnaA was incubated with pOC_4 plasmid, which led to significant unwinding at the DUE (**Figure 6E**). The addition of 6HisCj1509 at an equimolar or higher concentration than DnaA inhibited DUE unwinding, which indicated that the binding of Cj1509 to *C. jejuni oriC* inhibited DnaA-dependent DNA unwinding, as previously shown for HP1021 in *H. pylori*.

DISCUSSION

Correct timing and synchronization of chromosome replication with other processes are vital for bacterial fitness. Bacterial *oriCs* have been proposed to act as centralized information processors that receive and transmit information reflecting the current state of the bacterial cell (Marczynski et al., 2015). Thus, the identification of *oriC* and characterization of the mechanism of initiation of chromosome replication are crucial for studies on the bacterial cell cycle, as well as adaptation to the environment and host-pathogen interactions. In this work, we present the first data pinpointing the *oriC* of the *C. jejuni* chromosome and identifying possible regulators of replication initiation in this bacterium.

Three Modules of *C. jejuni oriC*

We found that *C. jejuni oriC* is composed of the three standard replication origin modules: DnaA-box cluster, DUE and OriBP binding site, all of which were located in the *dnaA – dnaN* locus (**Figure 7**). This localization of *C. jejuni oriC* is typical for many bacteria. Unlike other known Epsilonproteobacterial origins, which are most likely bipartite, we showed herein that *C. jejuni oriC* is probably monopartite because the single intergenic region between *dnaA* and *dnaN*, accompanied by 2 DnaA boxes at the 3′ end of *dnaA*, was sufficient to sustain self-replication of a minichromosome (see also below). In addition, the looped DNA structures typical for bipartite origins (Krause et al., 1997; Donczew et al., 2012; Jaworski et al., 2016) were not observed in EM upon *C. jejuni* orisome formation. The reason for the presence of a mono- or bipartite origin is unknown. It has been proposed that although the *E. coli* DnaA oligomer is assembled on a monopartite origin, it is structurally divided into sub-oligomers of different functions (DNA opening, DnaB loading) (Rozgaja et al., 2011;

Ozaki et al., 2012; Shimizu et al., 2016). Thus, in bipartite origins, the DnaA sub-oligomers, which are additionally spatially divided, may provide another level of control during orisome formation.

There are at least 7 DnaA boxes within *C. jejuni oriC* that are bound by DnaA *in vitro* (see also Discussion below). They do not form any clear-cut oligomerization pattern, similar to oppositely oriented R1-I2 and C3-R4 DnaA box arrays in *E. coli* (Rozgaja et al., 2011; Ozaki et al., 2012; Shimizu et al., 2016). Rather, the boxes are grouped in three divergently oriented sub-arrays (DnaA boxes 1–2, 3–4–5, and 7–8), providing a scaffold for a DnaA oligomer of unknown composition and structure. Bacterial origins identified to date are highly variable in number, spacing and orientation of DnaA boxes, and there is still no general understanding of the DnaA filament assembly with respect to the array of DnaA boxes. Interestingly, DnaA boxes 6–7, which are located within the *dnaA* gene, were shown to sustain replication of *C. jejuni oriC* as minichromosome, while the cluster of five DnaA boxes of the intergenic region IGR4 was not. It has been shown previously that in *Bdellovibrio bacteriovorus*, in which *oriC* is also located in the *dnaA-dnaN* intergenic region, the two DnaA boxes in the 3′ region of *dnaA* are bound by DnaA *in vitro*; however, it is not known whether these boxes are required for the initiation of *B. bacteriovorus* chromosome replication (Makowski et al., 2016). It should be noted that DnaA boxes required to support *oriC* activity on a minichromosome or on a chromosome might be different (Bates et al., 1995; Weigel et al., 2001; Riber et al., 2009). Studies on *E. coli oriC* have shown that mutations of some boxes, including the DUE-distal R4 DnaA box corresponding to DUE distal *C. jejuni* DnaA boxes 6–7 (**Figures 4, 7**), are not tolerated on minichromosomes while they can be mutagenized on chromosomal *oriC*. However, such mutations, although tolerated, often trigger perturbations in the initiation of chromosome replication and bacterial fitness, and thus they are conditionally required (Bates et al., 1995; Stepankiw et al., 2009). Moreover, minichromosomes are often kept at low copy number, they are unstable and tend to integrate into chromosomes (Zakrzewska-Czerwińska and Schrempf, 1992; Duret et al., 1999; Weigel et al., 2001; Cordova et al., 2002; Lartigue et al., 2003). *E. coli* minichromosomes are exceptional when compared to minichromosomes of other species because they self-replicate at relatively high-copy number [(Løbner-Olesen et al., 1987; Løbner-Olesen and von Freiesleben, 1996) and references herein]. Thus further studies and detailed mutagenesis of DnaA boxes directly within *C. jejuni* chromosomal *oriC* are required to analyze the role and essentiality of DnaA binding sites for *C. jejuni oriC* activity upon initiation of chromosome replication *in vivo*.

Five DnaA boxes at *C. jejuni oriC* resemble the typical Epsilonproteobacterial DnaA box consensus sequence, in which a core sequence 5′-TTCAC-3′ (4–8 nt) is strictly conserved; the DnaA box 1 differs at the 8th position (C→A), while the DnaA box 6 differs at the 4th position (T→A). However, *C. jejuni* DnaA, unlike other Epsilonproteobacterial DnaAs, also recognizes DnaA boxes in which the thymine residue at the 5th position is substituted by another nucleotide (guanine

in DnaA box 8 or adenine in DnaA box 12) (see also below).

The DUE (ca. 35 bps, 19% GC) is preceded by twin DnaA boxes 1–2, which overlap by 1 bp. The boxes are followed by ca. 5 bps that probably remains double-stranded upon DUE unwinding. The GC-rich sequence between DnaA boxes and the DUE, which occurs in some origins (Richardson et al., 2016), was not present in *C. jejuni*. Similarly, the DnaA trio consensus motif, which was previously reported to be important for ssDNA binding by DnaA upon DNA unwinding (Richardson et al., 2016), was detected in *C. jejuni oriC*, however, its sequence was degenerated when compared to the consensus 5′-TAG-3′ sequence. It should be noted that the DnaA trio motif is very short, and it is difficult to predict the mismatches that preclude its activity. In addition, it is not known why these motifs are conserved in many but not in all bacterial origins and how ssDNA is bound by bacteria that lack the obvious DnaA-trio motif (e.g., *E. coli*).

We also identified OriBP binding sites, i.e., sequences that are bound by the orphan response regulator Cj1509. There are two experimentally determined Cj1509 binding sites at Cj*oriC*, Cj1509 boxes 1 and 2; the Cj1509 box 3 is located within the *dnaN* gene (**Figure 7**). However, the sequence of Cj1509 box 1 (5′-TTTCACA-3′) was found at three other locations in *oriC* (**Figure 7**). The position of Cj1509 boxes did not seem to be distributed randomly because they overlap with the DnaA boxes (2 and possibly 3, 4, and 7) and the DUE. By competition with the binding sites, Cj1509 potentially interfered with DnaA-*oriC* interactions and precluded DNA unwinding (**Figure 6B**). A similar mechanism was observed in *H. pylori*, in which HP1021, a homolog of Cj1509, bound to DnaA boxes and the DUE, leading to the inhibition of DUE unwinding. Thus, HP1021 has been proposed to act as a repressor of chromosome replication (Donczew et al., 2015). A similar mechanism of inhibition has been previously proposed for *E. coli* and *M. tuberculosis*. It is postulated that under anaerobic conditions, *E. coli* ArcA∼P binds to the AT-rich region of *oriC* and inhibits DNA unwinding (Hwang and Kornberg, 1992b; Lee et al., 2001). *M. tuberculosis* MtrA∼P binds to *oriC* and also the promoter of *dnaA* and inhibits the initiation of *M. tuberculosis* chromosome replication (Fol et al., 2006; Purushotham et al., 2015). The signal that triggers MtrA phosphorylation is still unknown, but it is linked to the infection of macrophages (Fol et al., 2006). Thus, *E. coli* and *M. tuberculosis* proteins control chromosome replication and cell cycle progression in response to environmental conditions, including the stage of host infection. By analogy, Cj1509 and HP1021 might be activated by external stimuli, and these factors might regulate the initiation of chromosome replication in response to host–pathogen interactions. No changes in the level of Cj1509 expression were observed upon infection (de Vries et al., 2017). However, it is likely that, similarly to HP1021 (Müller et al., 2007), post-translational modification of Cj1509 rather than a change in expression level is important for Cj1509 function. Therefore, further studies are required to

identify the mechanism of Cj1509 activation and signal transduction.

The Roles of DnaA Boxes and Cj1509 Binding Sequences Beyond the Initiation of Chromosome Replication

DnaA boxes are also found outside *oriC*, which suggests that DnaA plays a role in processes other than initiation complex formation and DNA unwinding. The *C. jejuni* 5′-NHHWDCAMH-3′ DnaA box consensus sequence derived from 13 experimentally determined DnaA binding sites is relatively relaxed when compared to other Epsilonproteobacterial or *E. coli* DnaA boxes (Jaworski et al., 2016), including the 5th position of the DnaA box previously shown to be strictly conserved in Epsilonproteobacteria (Jaworski et al., 2016). Hence, *C. jejuni* DnaA is exceptional among the Epsilonproteobacteria studied to date. However, the *C. jejuni* DnaA amino acid sequence of the DNA binding helix-turn-helix motif within domain IV is highly similar to that of other Epsilonproteobacteria (Supplementary Figure S9) (Fujikawa et al., 2003; Tsodikov and Biswas, 2011; Jaworski et al., 2016). It should be noted that our studies did not allow to distinguish between low- and high-affinity DnaA binding sites, which differ in nucleotide sequence. For example, *E. coli* low affinity I-DnaA binding sites usually differ by 3–4 bases from high affinity R-type DnaA boxes (McGarry et al., 2004) while 6-mer τ DnaA binding sites are shorter than 9-mer R-type DnaA boxes (Kawakami et al., 2005). Thus, herein determined *C. jejuni* DnaA box consensus sequence may be relaxed because it includes high- and low- affinity DnaA binding sites. On the other hand, *B. bacteriovorus* DnaA box sequence was shown to be conserved within seven nucleotides, while the two nucleotides at the 5′ sequence of DnaA box were not conserved (Makowski et al., 2016). Therefore it is possible that some bacterial DnaAs specifically recognize sequences shorter than typical 9-mer DnaA box sequences. Thus, further studies are required to explain the loose *C. jejuni* DnaA specificity for DnaA consensus sequences.

Twenty DnaA boxes were predicted *in silico* in the vicinity of *C. jejuni oriC* (IGR2-*ruvC*-IGR3-*dnaA*). The binding of DnaA to the *oriC* proximal region was also observed by electron microscopy, and we confirmed the binding of CjDnaA to four randomly chosen DnaA boxes by DMS footprinting (DnaA boxes 9–11 in *dnaA* and 13 in IGR2); additionally we identified 2 DnaA boxes that were not predicted (DnaA box 8 in *dnaA* and 12 in IGR2) (Supplementary Figure S6). DnaA boxes 12 and 13 were separated by 104 bps, which likely excluded cooperativity in DnaA-DNA interactions and indicated that *C. jejuni* DnaA could efficiently recognize single DnaA boxes. There are numerous putative DnaA boxes on the *C. jejuni* 81116 chromosome. For example, there are 2659 DnaA boxes that follow the stringent *C. jejuni* DnaA box consensus sequence 5′-WWHWTCACW-3′ (Mrázek and Xie, 2006). Excluding the *oriC* region, the DnaA boxes were evenly distributed along the chromosome. The role of DnaA boxes scattered along the chromosome is not known. As observed by EM, DnaA bound to *oriC* and proximal DnaA boxes formed a large nucleoprotein complex that might affect

the activity of the initiation complex. Thus, the *oriC* proximal DnaA boxes (**Figures 4**, **5**) might play a regulatory function in the control of *C. jejuni* chromosome replication. The genome-wide scattered boxes located in the intergenic regions (100 DnaA boxes, 3.7% of all the boxes) might contribute to regulation of gene expression, similarly to *E. coli* or *B. subtilis* (Messer and Weigel, 2003; Washington et al., 2017). *C. jejuni* DnaA was shown to organize DNA into higher order structures such as loops and wraps (**Figure 5**). Thus, we speculate that DnaA participates in the global control of the nucleoid structure, especially since *C. jejuni* chromosome lacks many nucleoid associated proteins such as H-NS, IHF, and Fis (Shortt et al., 2016), which help to maintain the chromosome structure in other bacteria (Browning et al., 2010; Badrinarayanan et al., 2015; Dame and Tark-Dame, 2016). However, further experimental studies are required to determine the actual DnaA binding sites in the *C. jejuni* chromosome, especially since *in silico* predictions of DnaA boxes often overestimate the number of actual DnaA binding sites. Moreover, binding of DnaA to DNA might differ *in vivo* and *in vitro* (Smith and Grossman, 2015).

The actual binding sites of Cj1509 on the *C. jejuni* chromosome are difficult to predict because the Cj1509 consensus 5′-WKTWWCA-3′ sequence is too relaxed to provide reliable data for *in silico* analysis. Preliminary results suggested that Cj1509 exhibits the highest affinity toward Cj1509 box 2 (5′-TGTTACA-3′). However, further studies are required to better characterize Cj1509 affinity to individual Cj1509 boxes and, since there are multiple Cj1509 binding sites at *C. jejuni oriC*, to determine whether Cj1509-DNA interactions might exhibit cooperativity. Nonetheless, Cj1509 could control the expression of selected *C. jejuni* genes in response to unknown stimuli; however, Cj1509 has been reported to be non-essential in some strains (de Vries et al., 2017). Moreover, in different laboratories, Cj1509 has been shown to be an essential and non-essential gene in the same strain, e.g., *C. jejuni* 11168 (Metris et al., 2011; de Vries et al., 2017; Mandal et al., 2017). Thus, it is possible that strain-specific genome content and growth conditions determine the actual role of Cj1509 in *C. jejuni* physiology.

In summary, we identified and characterized the probable *oriC* region of *C. jejuni*. We anticipate that these studies will initiate further research on the structure and dynamics of the *C. jejuni* chromosome, which in turn will facilitate studies on the *C. jejuni* life cycle in the context of its biology and pathogenicity. The results are also important for further comparative investigations of the initiation of chromosome replication and other cellular processes throughout the whole class of Epsilonproteobacteria, which include established and emerging pathogens associated with gastrointestinal diseases and/or reproductive disorders in animals, as well as non-pathogenic symbiotic or free-living species (Eppinger et al., 2004; Gupta, 2006; Nakagawa and Takaki, 2009; Waite et al., 2017). Such diverse life styles of Epsilonproteobacteria might be reflected by the diversity of the initiation or regulatory factors involved in the initiation of chromosome replication of species inhabiting various ecological niches.

AUTHOR CONTRIBUTIONS

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2018.01533/full#supplementary-material

REFERENCES

Badrinarayanan, A., Le, T. B. K., 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

Bates, D. B., Asai, T., Cao, Y., Chambers, M. W., Cadwell, G. W., Boye, E., et al. (1995). The DnaA box R4 in the minimal oriC is dispensable for initiation of *Escherichia coli* chromosome replication. *Nucleic Acids Res.* 23, 3119–3125.

Bawono, P., and Heringa, J. (2014). PRALINE: a versatile multiple sequence alignment toolkit. *Methods Mol. Biol.* 1079, 245–262. doi: 10.1007/978-1-62703-646-7_16

Briggs, G. S., Smits, W. K., and Soultanas, P. (2012). Chromosomal replication initiation machinery of low-g + c-content firmicutes. *J. Bacteriol.* 194, 5162–5170. doi: 10.1128/JB.00865-12

Browning, D. F., Grainger, D. C., and Busby, S. J. (2010). Effects of nucleoid-associated proteins on bacterial chromosome structure and gene expression. *Curr. Opin. Microbiol.* 13, 773–780. doi: 10.1016/j.mib.2010.09.013

Cassler, M. R., Grimwade, J. E., and Leonard, A. C. (1995). Cell cycle-specific changes in nucleoprotein complexes at a chromosomal replication origin. *EMBO J.* 14, 5833–5841.

Castilla-Llorente, V., Muñoz-Espín, D., Villar, L., Salas, M., and Meijer, W. J. J. (2006). Spo0A, the key transcriptional regulator for entrance into sporulation, is an inhibitor of DNA replication. *EMBO J.* 25, 3890–3899. doi: 10.1038/sj.emboj.7601266

Cordova, C. M. M., Lartigue, C., Sirand-Pugnet, P., Renaudin, J., Cunha, R. A. F., and Blanchard, A. (2002). Identification of the origin of replication of the *Mycoplasma pulmonis* chromosome and its use in oriC replicative plasmids. *J. Bacteriol.* 184, 5426–5435.

Cróinín, T., and Backert, S. (2012). Host epithelial cell invasion by *Campylobacter jejuni*: trigger or zipper mechanism? *Front. Cell. Infect. Microbiol.* 2:25. doi: 10.3389/fcimb.2012.00025

Crooks, G. E., Hon, G., Chandonia, J.-M., and Brenner, S. E. (2004). WebLogo: a sequence logo generator. *Genome Res.* 14, 1188–1190. doi: 10.1101/gr.849004

Dame, R. T., and Tark-Dame, M. (2016). Bacterial chromatin: converging views at different scales. *Curr. Opin. Cell Biol.* 40, 60–65. doi: 10.1016/j.ceb.2016.02.015

Dasgupta, S., and Løbner-Olesen, A. (2004). Host controlled plasmid replication: *Escherichia coli* minichromosomes. *Plasmid* 52, 151–168. doi: 10.1016/j.plasmid.2004.08.001

de Vries, S. P., Gupta, S., Baig, A., Wright, E., Wedley, A., Jensen, A. N., et al. (2017). Genome-wide fitness analyses of the foodborne pathogen *Campylobacter jejuni* in in vitro and in vivo models. *Sci. Rep.* 7:1251. doi: 10.1038/s41598-017-01133-4

Donczew, R., Makowski, Ł., Jaworski, P., Bezulska, M., Nowaczyk, M., Zakrzewska-Czerwińska, J., et al. (2015). The atypical response regulator HP1021 controls formation of the *Helicobacter pylori* replication initiation complex. *Mol. Microbiol.* 95, 297–312. doi: 10.1111/mmi.12866

Donczew, R., Mielke, T., Jaworski, P., Zakrzewska-Czerwińska, J., and Zawilak-Pawlik, A. (2014). Assembly of *Helicobacter pylori* initiation complex is determined by sequence-specific and topology-sensitive DnaA-oriC interactions. *J. Mol. Biol.* 426, 2769–2782. doi: 10.1016/j.jmb.2014.05.018

Donczew, R., Weigel, C., Lurz, R., Zakrzewska-Czerwińska, J., and Zawilak-Pawlik, A. (2012). *Helicobacter pylori* oriC —the first bipartite origin of chromosome replication in Gram-negative bacteria. *Nucleic Acids Res.* 40, 9647–9660. doi: 10.1093/nar/gks742

Duderstadt, K. E., and Berger, J. M. (2013). A structural framework for replication origin opening by AAA+ initiation factors. *Curr. Opin. Struct. Biol.* 23, 144–153. doi: 10.1016/j.sbi.2012.11.012

Duret, S., Danet, J. L., Garnier, M., and Renaudin, J. (1999). Gene disruption through homologous recombination in *Spiroplasma citri*: an scm1-disrupted motility mutant is pathogenic. *J. Bacteriol.* 181, 7449–7456.

Eppinger, M., Baar, C., Raddatz, G., Huson, D. H., and Schuster, S. C. (2004). Comparative analysis of four Campylobacterales. *Nat. Rev. Microbiol.* 2, 872–885. doi: 10.1038/nrmicro1024

Fol, M., Chauhan, A., Nair, N. K., Maloney, E., Moomey, M., Jagannath, C., et al. (2006). Modulation of *Mycobacterium tuberculosis* proliferation by MtrA, an essential two-component response regulator. *Mol. Microbiol.* 60, 643–657. doi: 10.1111/j.1365-2958.2006.05137.x

Fujikawa, N., Kurumizaka, H., Nureki, O., Terada, T., Shirouzu, M., Katayama, T., et al. (2003). Structural basis of replication origin recognition by the DnaA protein. *Nucleic Acids Res.* 31, 2077–2086.

Gao, F., Luo, H., and Zhang, C.-T. (2013). DoriC 5.0: an updated database of oriC regions in both bacterial and archaeal genomes. *Nucleic Acids Res.* 41, D90–D93. doi: 10.1093/nar/gks990

Gupta, R. S. (2006). Molecular signatures (unique proteins and conserved indels) that are specific for the epsilon proteobacteria (*Campylobacterales*). *BMC Genomics* 7:167. doi: 10.1186/1471-2164-7-167

Hansen, F. G., Christensen, B. B., and Atlung, T. (2007). Sequence characteristics required for cooperative binding and efficient in vivo titration of the replication initiator protein DnaA in *E. coli*. *J. Mol. Biol.* 367, 942–952. doi: 10.1016/j.jmb.2007.01.056

Hwang, D. S., and Kornberg, A. (1992a). Opening of the replication origin of *Escherichia coli* by DnaA protein with protein HU or IHF. *J. Biol. Chem.* 267, 23083–23086.

Hwang, D. S., and Kornberg, A. (1992b). Opposed actions of regulatory proteins, DnaA and IciA, in opening the replication origin of *Escherichia coli*. *J. Biol. Chem.* 267, 23087–23091.

Jaworski, P., Donczew, R., Mielke, T., Thiel, M., Oldziej, S., Weigel, C., et al. (2016). Unique and universal features of Epsilonproteobacterial origins of chromosome replication and DnaA-DnaA box interactions. *Evol. Genomic Microbiol.* 7:1555. doi: 10.3389/fmicb.2016.01555

Jullien, N., and Herman, J.-P. (2011). LUEGO: a cost and time saving gel shift procedure. *Biotechniques* 51, 267–269. doi: 10.2144/000113751

Kaakoush, N. O., Castaño-Rodríguez, N., Mitchell, H. M., and Man, S. M. (2015). Global epidemiology of *Campylobacter* infection. *Clin. Microbiol. Rev.* 28, 687–720. doi: 10.1128/CMR.00006-15

Katayama, T., Kasho, K., and Kawakami, H. (2017). The DnaA cycle in *Escherichia coli*: activation, function and inactivation of the initiator protein. *Front. Microbiol.* 8:2496. doi: 10.3389/fmicb.2017.02496

Kawakami, H., Keyamura, K., and Katayama, T. (2005). Formation of an ATP-DnaA-specific initiation complex requires DnaA Arginine 285, a conserved motif in the AAA+ protein family. *J. Biol. Chem.* 280, 27420–27430. doi: 10.1074/jbc.M502764200

Keyamura, K., Fujikawa, N., Ishida, T., Ozaki, S., Su'etsugu, M., Fujimitsu, K., et al. (2007). The interaction of DiaA and DnaA regulates the replication cycle in *E. coli* by directly promoting ATP DnaA-specific initiation complexes. *Genes Dev.* 21, 2083–2099. doi: 10.1101/gad.1561207

Kowalski, D., and Eddy, M. J. (1989). The DNA unwinding element: a novel, cis-acting component that facilitates opening of the *Escherichia coli* replication origin. *EMBO J.* 8, 4335–4344.

Kowalski, D., Natale, D. A., and Eddy, M. J. (1988). Stable DNA unwinding, not "breathing," accounts for single-strand-specific nuclease hypersensitivity of specific A+T-rich sequences. *Proc. Natl. Acad. Sci. U.S.A.* 85, 9464–9468.

Krause, M., Rückert, B., Lurz, R., and Messer, W. (1997). Complexes at the replication origin of *Bacillus subtilis* with homologous and heterologous DnaA protein. *J. Mol. Biol.* 274, 365–380. doi: 10.1006/jmbi.1997.1404

Lartigue, C., Blanchard, A., Renaudin, J., Thiaucourt, F., and Sirand-Pugnet, P. (2003). Host specificity of mollicutes oriC plasmids: functional analysis of replication origin. *Nucleic Acids Res.* 31, 6610–6618.

Lee, Y. S., Han, J. S., Jeon, Y., and Hwang, D. S. (2001). The arc two-component signal transduction system inhibits in vitro *Escherichia coli* chromosomal initiation. *J. Biol. Chem.* 276, 9917–9923. doi: 10.1074/jbc.M008629200

Leonard, A. C., and Grimwade, J. E. (2011). Regulation of DnaA assembly and activity: taking directions from the genome. *Annu. Rev. Microbiol.* 65, 19–35. doi: 10.1146/annurev-micro-090110-102934

Leonard, A. C., and Grimwade, J. E. (2015). The orisome: structure and function. *Front. Microbiol.* 6:545. doi: 10.3389/fmicb.2015.00545

Løbner-Olesen, A., Atlung, T., and Rasmussen, K. V. (1987). Stability and replication control of *Escherichia coli* minichromosomes. *J. Bacteriol.* 169, 2835–2842.

Løbner-Olesen, A., and von Freiesleben, U. (1996). Chromosomal replication incompatibility in Dam methyltransferase deficient *Escherichia coli* cells. *EMBO J.* 15, 5999–6008.

Makowski, Ł., Donczew, R., Weigel, C., Zawilak-Pawlik, A., and Zakrzewska-Czerwińska, J. (2016). Initiation of chromosomal replication in predatory bacterium *Bdellovibrio bacteriovorus*. *Front. Microbiol.* 7:1898. doi: 10.3389/fmicb.2016.01898

Mandal, R. K., Jiang, T., and Kwon, Y. M. (2017). Essential genome of *Campylobacter jejuni*. *BMC Genomics* 18:616. doi: 10.1186/s12864-017-4032-8

Marczynski, G. T., Rolain, T., and Taylor, J. A. (2015). Redefining bacterial origins of replication as centralized information processors. *Front. Microbiol.* 6:610. doi: 10.3389/fmicb.2015.00610

McGarry, K. C., Ryan, V. T., Grimwade, J. E., and Leonard, A. C. (2004). Two discriminatory binding sites in the *Escherichia coli* replication origin are required for DNA strand opening by initiator DnaA-ATP. *Proc. Natl. Acad. Sci. U.S.A.* 101, 2811–2816. doi: 10.1073/pnas.0400340101

Messer, W., and Weigel, C. (2003). DnaA as a transcription regulator. *Methods Enzymol.* 370, 338–349.

Metris, A., Reuter, M., Gaskin, D. J., Baranyi, J., and van Vliet, A. H. (2011). In vivo and in silico determination of essential genes of *Campylobacter jejuni*. *BMC Genomics* 12:535. doi: 10.1186/1471-2164-12-535

Moriya, S., Atlung, T., Hansen, F. G., Yoshikawa, H., and Ogasawara, N. (1992). Cloning of an autonomously replicating sequence (ars) from the *Bacillus subtilis* chromosome. *Mol. Microbiol.* 6, 309–315.

Mrázek, J., and Xie, S. (2006). Pattern locator: a new tool for finding local sequence patterns in genomic DNA sequences. *Bioinformatics* 22, 3099–3100. doi: 10.1093/bioinformatics/btl551

Müller, S., Pflock, M., Schär, J., Kennard, S., and Beier, D. (2007). Regulation of expression of atypical orphan response regulators of *Helicobacter pylori*. *Microbiol. Res.* 162, 1–14. doi: 10.1016/j.micres.2006.01.003

Murphy, C., Carroll, C., and Jordan, K. N. (2006). Environmental survival mechanisms of the foodborne pathogen *Campylobacter jejuni*. *J. Appl. Microbiol.* 100, 623–632. doi: 10.1111/j.1365-2672.2006.02903.x

Nakagawa, S., and Takaki, Y. (2009). "Nonpathogenic Epsilonproteobacteria," in *Encyclopedia of Life Sciences* (Chichester: John Wiley & Sons, Ltd.). Available at: https://onlinelibrary.wiley.com/

Natrajan, G., Hall, D. R., Thompson, A. C., Gutsche, I., and Terradot, L. (2007). Structural similarity between the DnaA-binding proteins HobA (HP1230) from *Helicobacter pylori* and DiaA from *Escherichia coli*. *Mol. Microbiol.* 65, 995–1005. doi: 10.1111/j.1365-2958.2007.05843.x

Ozaki, S., and Katayama, T. (2012). Highly organized DnaA-oriC complexes recruit the single-stranded DNA for replication initiation. *Nucleic Acids Res.* 40, 1648–1665. doi: 10.1093/nar/gkr832

Ozaki, S., Noguchi, Y., Hayashi, Y., Miyazaki, E., and Katayama, T. (2012). Differentiation of the DnaA-oriC subcomplex for DNA unwinding in a replication initiation complex. *J. Biol. Chem.* 287, 37458–37471. doi: 10.1074/jbc.M112.372052

Parkhill, J., Wren, B. W., Mungall, K., Ketley, J. M., Churcher, C., Basham, D., et al. (2000). The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. *Nature* 403, 665–668. doi: 10.1038/35001088

Purushotham, G., Sarva, K. B., Blaszczyk, E., Rajagopalan, M., and Madiraju, M. V. (2015). *Mycobacterium tuberculosis* oriC sequestration by MtrA response regulator. *Mol. Microbiol.* 98, 586–604. doi: 10.1111/mmi.13144

Rajewska, M., Wegrzyn, K., and Konieczny, I. (2012). AT-rich region and repeated sequences - the essential elements of replication origins of bacterial replicons. *FEMS Microbiol. Rev.* 36, 408–434. doi: 10.1111/j.1574-6976.2011.00300.x

Riber, L., Fujimitsu, K., Katayama, T., and Løbner-Olesen, A. (2009). Loss of Hda activity stimulates replication initiation from I-box, but not R4 mutant origins in *Escherichia coli*. *Mol. Microbiol.* 71, 107–122. doi: 10.1111/j.1365-2958.2008.06516.x

Richardson, T. T., Harran, O., and Murray, H. (2016). The bacterial DnaA-trio replication origin element specifies single-stranded DNA initiator binding. *Nature* 534, 412–416. doi: 10.1038/nature17962

Rozgaja, T. A., Grimwade, J. E., Iqbal, M., Czerwonka, C., Vora, M., and Leonard, A. C. (2011). Two oppositely oriented arrays of low-affinity recognition sites in oriC guide progressive binding of DnaA during *Escherichia coli* pre-RC assembly. *Mol. Microbiol.* 82, 475–488. doi: 10.1111/j.1365-2958.2011.07827.x

Sakiyama, Y., Kasho, K., Noguchi, Y., Kawakami, H., and Katayama, T. (2017). Regulatory dynamics in the ternary DnaA complex for initiation of chromosomal replication in *Escherichia coli*. *Nucleic Acids Res.* 45, 12354–12373. doi: 10.1093/nar/gkx914

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

Sasse-Dwight, S., and Gralla, J. D. (1991). Footprinting protein-DNA complexes in vivo. *Methods Enzymol.* 208, 146–168.

Schaper, S., and Messer, W. (1995). Interaction of the initiator protein DnaA of *Escherichia coli* with its DNA target. *J. Biol. Chem.* 270, 17622–17626.

Schär, J., Sickmann, A., and Beier, D. (2005). Phosphorylation-independent activity of atypical response regulators of *Helicobacter pylori*. *J. Bacteriol.* 187, 3100–3109. doi: 10.1128/JB.187.9.3100-3109.2005

Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* 9, 671–675.

Schneider, T. D., and Stephens, R. M. (1990). Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res.* 18, 6097–6100.

Sekimizu, K., Bramhill, D., and Kornberg, A. (1987). ATP activates dnaA protein in initiating replication of plasmids bearing the origin of the *E. coli* chromosome. *Cell* 50, 259–265.

Shaw, P. E., and Stewart, A. F. (2009). Identification of protein/DNA contacts with dimethyl sulfate: methylation protection and methylation interference. *Methods Mol. Biol.* 543, 97–104. doi: 10.1007/978-1-60327-015-1_8

Shimizu, M., Noguchi, Y., Sakiyama, Y., Kawakami, H., Katayama, T., and Takada, S. (2016). Near-atomic structural model for bacterial DNA replication initiation complex and its functional insights. *Proc. Natl. Acad. Sci. U.S.A.* 113, E8021–E8030. doi: 10.1073/pnas.1609649113

Shortt, C., Scanlan, E., Hilliard, A., Cotroneo, C. E., Bourke, B., and Ó Cróinín, T. (2016). DNA supercoiling regulates the motility of *Campylobacter jejuni* and

is altered by growth in the presence of chicken mucus. *mBio* 7:e01227-16. doi: 10.1128/mBio.01227-16

Smith, J. L., and Grossman, A. D. (2015). In Vitro whole genome DNA binding analysis of the bacterial replication initiator and transcription factor DnaA. *PLoS Genet.* 11:e1005258. doi: 10.1371/journal.pgen.100 5258

Spiess, E., and Lurz, R. (1988). *13 Electron Microscopic Analysis of Nucleic Acids and Nucleic Acid-Protein Complexes.* Cambridge, MA: Academic Press, 293–323.

Stepankiw, N., Kaidow, A., Boye, E., and Bates, D. (2009). The right half of the *Escherichia coli* replication origin is not essential for viability, but facilitates multi-forked replication. *Mol. Microbiol.* 74, 467–479. doi: 10.1111/j.1365-2958. 2009.06877.x

Taylor, J. A., Ouimet, M.-C., Wargachuk, R., and Marczynski, G. T. (2011). The *Caulobacter crescentus* chromosome replication origin evolved two classes of weak DnaA binding sites. *Mol. Microbiol.* 82, 312–326. doi: 10.1111/j.1365-2958.2011.07785.x

Tsodikov, O. V., and Biswas, T. (2011). Structural and thermodynamic signatures of DNA recognition by *Mycobacterium tuberculosis* DnaA. *J. Mol. Biol.* 410, 461–476. doi: 10.1016/j.jmb.2011.05.007

Van Vliet, A. H. M., Wood, A. C., Wooldridge, K., Ketley, J. M., and Henderson, J. (1998). "7.7 genetic manipulation of enteric *Campylobacter* Species," in *Methods in Microbiology Bacterial Pathogenesis*, eds P. Williams, J. Ketley, and G. Salmond (Cambridge, MA: Academic Press), 407–419. doi: 10.1016/S0580-9517(08)70301-5

Waite, D. W., Vanwonterghem, I., Rinke, C., Parks, D. H., Zhang, Y., Takai, K., et al. (2017). Comparative genomic analysis of the class Epsilonproteobacteria and proposed reclassification to epsilonbacteraeota (phyl. nov.). *Front. Microbiol.* 8:682. doi: 10.3389/fmicb.2017.00682

Waldminghaus, T., and Skarstad, K. (2009). The *Escherichia coli* SeqA protein. *Plasmid* 61, 141–150. doi: 10.1016/j.plasmid.2009.02.004

Washington, T. A., Smith, J. L., and Grossman, A. D. (2017). Genetic networks controlled by the bacterial replication initiator and transcription factor DnaA in *Bacillus subtilis. Mol. Microbiol.* 106, 109–128. doi: 10.1111/mmi.13755

Weigel, C., Messer, W., Preiss, S., Welzeck, M., Morigen, and Boye, E. (2001). The sequence requirements for a functional *Escherichia coli* replication origin are different for the chromosome and a minichromosome. *Mol. Microbiol.* 40, 498–507.

Wolański, M., Donczew, R., Zawilak-Pawlik, A., and Zakrzewska-Czerwińska, J. (2014). oriC-encoded instructions for the initiation of bacterial chromosome replication. *Front. Microbiol.* 5:735. doi: 10.3389/fmicb.2014.00735

Yao, R., Alm, R. A., Trust, T. J., and Guerry, P. (1993). Construction of new *Campylobacter* cloning vectors and a new mutational cat cassette. *Gene* 130, 127–130.

Yasuda, S., and Hirota, Y. (1977). Cloning and mapping of the replication origin of *Escherichia coli. Proc. Natl. Acad. Sci. U.S.A.* 74, 5458–5462.

Zakrzewska-Czerwińska, J., Majka, J., and Schrempf, H. (1995). Minimal requirements of the *Streptomyces lividans* 66 oriC region and its transcriptional and translational activities. *J. Bacteriol.* 177, 4765–4771.

Zakrzewska-Czerwińska, J., and Schrempf, H. (1992). Characterization of an autonomously replicating region from the *Streptomyces lividans* chromosome. *J. Bacteriol.* 174, 2688–2693.

Zawilak-Pawlik, A., Donczew, R., Szafrański, S., Mackiewicz, P., Terradot, L., and Zakrzewska-Czerwińska, J. (2011). DiaA/HobA and DnaA: a pair of proteins co-evolved to cooperate during bacterial orisome assembly. *J. Mol. Biol.* 408, 238–251. doi: 10.1016/j.jmb.2011.02.045

Zawilak-Pawlik, A., Kois, A., Stingl, K., Boneca, I. G., Skrobuk, P., Piotr, J., et al. (2007). HobA–a novel protein involved in initiation of chromosomal replication in *Helicobacter pylori. Mol. Microbiol.* 65, 979–994. doi: 10.1111/j.1365-2958.2007.05853.x

Zawilak-Pawlik, A., Nowaczyk, M., and Zakrzewska-Czerwińska, J. (2017). The role of the N-terminal domains of bacterial initiator DnaA in the assembly and regulation of the bacterial replication initiation complex. *Genes* 8:136. doi: 10.3390/genes8050136

Zawilak-Pawlik, A. M., Kois, A., and Zakrzewska-Czerwinska, J. (2006). A simplified method for purification of recombinant soluble DnaA proteins. *Protein Expr. Purif.* 48, 126–133. doi: 10.1016/j.pep.2006.01.010

Zeng, X., Ardeshna, D., and Lin, J. (2015). Heat shock-enhanced conjugation efficiency in standard *Campylobacter jejuni* strains. *Appl. Environ. Microbiol.* 81, 4546–4552. doi: 10.1128/AEM.00346-15

Comprehensive Analysis of Replication Origins in *Saccharomyces cerevisiae* Genomes

Dan Wang[1] and Feng Gao[1,2,3]*

[1] *Department of Physics, School of Science, Tianjin University, Tianjin, China,* [2] *Key Laboratory of Systems Bioengineering, Ministry of Education, Tianjin University, Tianjin, China,* [3] *SynBio Research Platform, Collaborative Innovation Center of Chemical Science and Engineering, Tianjin, China*

DNA replication initiates from multiple replication origins (ORIs) in eukaryotes. Discovery and characterization of replication origins are essential for a better understanding of the molecular mechanism of DNA replication. In this study, the features of autonomously replicating sequences (ARSs) in *Saccharomyces cerevisiae* have been comprehensively analyzed as follows. Firstly, we carried out the analysis of the ARSs available in *S. cerevisiae* S288C. By evaluating the sequence similarity of experimentally established ARSs, we found that 94.32% of ARSs are unique across the whole genome of *S. cerevisiae* S288C and those with high sequence similarity are prone to locate in subtelomeres. Subsequently, we built a non-redundant dataset with a total of 520 ARSs, which are based on ARSs annotation of *S. cerevisiae* S288C from SGD and then supplemented with those from OriDB and DeOri databases. We conducted a large-scale comparison of ORIs among the diverse budding yeast strains from a population genomics perspective. We found that 82.7% of ARSs are not only conserved in genomic sequence but also relatively conserved in chromosomal position. The non-conserved ARSs tend to distribute in the subtelomeric regions. We also conducted a pan-genome analysis of ARSs among the *S. cerevisiae* strains, and a total of 183 core ARSs existing in all yeast strains were determined. We extracted the genes adjacent to replication origins among the 104 yeast strains to examine whether there are differences in their gene functions. The result showed that the genes involved in the initiation of DNA replication, such as *orc3*, *mcm2*, *mcm4*, *mcm6,* and *cdc45*, are conservatively located adjacent to the replication origins. Furthermore, we found the genes adjacent to conserved ARSs are significantly enriched in DNA binding, enzyme activity, transportation, and energy, whereas for the genes adjacent to non-conserved ARSs are significantly enriched in response to environmental stress, metabolites biosynthetic process and biosynthesis of antibiotics. In general, we characterized the replication origins from the genome-wide and population genomics perspectives, which would provide new insights into the replication mechanism of *S. cerevisiae* and facilitate the design of algorithms to identify genome-wide replication origins in yeast.

Keywords: replication origin, DNA replication, *Saccharomyces cerevisiae*, genome-wide analysis, autonomously replicating sequence

INTRODUCTION

DNA replication is a highly orchestrated process, which is tightly controlled to duplicate the genetic materials into both daughter cells (Bell and Labib, 2016). The specific sites where DNA replication initiates and double-stranded DNA starts unwinding are termed replication origins (ORI) (Jacob et al., 1963; Gilbert, 2001). The identification of ORIs has long been a critical issue, which is helpful to elucidate the molecular mechanism of DNA replication.

The base composition asymmetry widely exists in bacterial genomes (Lobry, 1996; Rocha et al., 1999; Zhang and Gao, 2017; Quan and Gao, 2019). Based on this phenomenon, some strategies to predict replication origin of chromosomes (*oriC*s) have been developed, for instance, GC skew (Lobry, 1996), cumulative GC skew (Grigoriev, 1998), skewed oligomers (Salzberg et al., 1998) and Z-curve (Zhang and Zhang, 2014). Considering the distributions of DnaA boxes and the conserved *oriC*s-adjacent genes in different phyla (Mackiewicz et al., 2004; Luo et al., 2018), the web server Ori-Finder (Gao and Zhang, 2008; Luo et al., 2014) has been developed based on the Z-curve method to predict *oriC*s in bacteria.

For eukaryotes, due to the long linear chromosomes, initiation of DNA replication occurs at multiple discrete sites and activates following the specific timing of DNA replication during the S phase (Taylor, 1960; Hand, 1978; Friedman et al., 1997). The characteristics of eukaryotic replication origins are best understood in the budding yeast *Saccharomyces cerevisiae*. Sequences conferring the ability of the autonomous replication on circular plasmid molecule are termed autonomously replicating sequences (ARSs) (Stinchcomb et al., 1981) that are regarded as ORIs in yeast chromosomes (Brewer and Fangman, 1987). Taking ARS1 as an example, it consists of the A element (ARS consensus sequence, ACS) (Marahrens and Stillman, 1992) where the ATP-dependent origin recognition complex (ORC) specifically recognizes and binds (Bell and Stillman, 1992; Li et al., 2018), the B1 element partially involved in ORC-DNA interaction (Duderstadt and Berger, 2008; Li et al., 2018), and the B2 element associated with mini-chromosome maintenance (MCM) proteins (Wilmes and Bell, 2002). ARS1 also contains the binding site for site-specific DNA-binding protein ABF1 (ARS binding factor I) (Diffley and Stillman, 1988), whereas ABF1 is not a universal ARS-binding factor. The experimental methods for identifying ARSs in yeast such as the two-dimensional (2D) gel analysis (Brewer and Fangman, 1987; Newlon et al., 1993), microarray-based approaches (Lee et al., 2007), chromatin immunoprecipitation (ChIP) including microarray (ChIP-chip) (Wyrick et al., 2001) and sequencing (ChIP-seq) (Eaton et al., 2010) as well as deep sequencing approaches (Müller et al., 2013) have provided plentiful accurate and reliable results. However, they are costly and time-consuming.

With the accumulation of experimental data and sequencing genomes, the available databases related to replication origins in yeast such as SGD (Cherry et al., 2012), OriDB (Nieduszynski et al., 2006a), DeOri (Gao et al., 2012) and DNA replication (Cotterill and Kearsey, 2008) have been established and updated, which brings new opportunities to study ORIs in yeast genome

via more efficient and faster bioinformatic methods. For example, Breier et al. (2004) developed an Oriscan algorithm to predict ORIs in the *S. cerevisiae* genome utilizing both the ACS motif and its flanking AT-rich region. Consequently, 84% of the top 100 Oriscan predictions matched known ARSs or replication protein binding sites, whereas with the accumulation of predictions, only 56% of the top 350 Oriscan predictions were matched. The result indicated that the algorithm using the similarity to 26 featured origins may limit the discovery of new potential ARSs. The machine learning-based techniques for predicting ORIs in yeast genome have been developed in recent years. Both iRO-3wPseKNC (Liu et al., 2018) and PseKNC2.0 (Dao et al., 2019) web-servers generated the sample formulation based on the mode of PseKNC (pseudo K-tuple nucleotide composition) for describing nucleotide sequences, and utilized the machine learning methods of random forest (RF) and support vector machine (SVM), respectively. Both the web-servers are user-friendly and efficiently performing, whereas direct extraction of the overall ORI sequence information without highlighting the characteristic conservative motifs will undoubtedly dilute the specific features of ORIs, resulting in lowering the prediction accuracy and increasing the false positives. Nieduszynski et al. (2006b) combined the results of ACS motif searches, phylogenetic conservation and microarray data, which enabled the prediction of essential ORIs throughout the *S. cerevisiae* genome. The result of phylogenetic conservation of replication origin sequences among closely related *Saccharomyces* species evidently improves the determination of the genome-wide location of replication origins, which suggested that multi-aspect analysis of replication origin sequences will facilitate the performance of prediction models. Although ORIs are essential for the maintenance of *S. cerevisiae* genome, the yeast chromosome harboring multiple origin deletions has been reported to replicate relatively normally (Dershowitz et al., 2007; Bogenschutz et al., 2014), that is to say, for an individual ORI, it is optional or redundant, which reflects the unexpected flexibility of DNA replication and also implies that there are still a number of potential replication origins to be discovered. Research on replication origins of only one or several strains may provide limited information. In recent years, the accumulation of published *S. cerevisiae* whole genome sequences (Strope et al., 2015; Zhu et al., 2016; Peter et al., 2018) is unprecedented, hence we can not only comprehensively analyze the ORI features at a genome-wide level, but also have the chance to compare the similarities and differences of replication origin sequences among diverse budding yeast strains from a population genomics respective. However, there are no such reports on the analysis of replication origins based on large-scale genomic data so far.

In this study, we firstly summarized and analyzed the characteristics of replication origin sequences in the reference genome of *S. cerevisiae* S288C, including classification of ARSs and the specific features in different types. Then we retrieved 104 genome sequences of budding yeasts with high genome integrity, and built a non-redundant dataset based on the published ARSs of *S. cerevisiae* S288C and supplemented with the confirmed ARSs from OriDB and DeOri databases, which makes it possible for us to conduct a large-scale comparison of replication

origin sequences among various yeast strains from a population genomics perspective. By a pan-genome analysis of ARSs among the *S. cerevisiae* strains, we determined the core ARSs existing in all yeast strains. We also analyzed the distribution bias of various ARSs with different conservation along the chromosomes. To explore whether the ARS-adjacent genes are conserved or not, we extracted genes adjacent to replication origins among the 104 yeast strains. Subsequently, we compared the enriched function of genes adjacent to various ARSs with different conservation and attempted to explain the relationships between replication origins and their adjacent genes.

MATERIALS AND METHODS

Strains and Datasets

We retrieved the reference genome sequence of *S. cerevisiae* S288C (version: R64-2-1) as well as 103 well-annotated budding yeast genome sequences with high genome integrity (>95%) from the NCBI FTP site[1]. The annotation of *S. cerevisiae* S288C was downloaded from the SGD FTP site[2]. From literature sources, the information of ecological origins and geographical origins of above *S. cerevisiae* strains were obtained (Strope et al., 2015; Peter et al., 2018) and the datasets of temporal replication of yeast chromosomes were collected (Raghuraman et al., 2001). We also acquired the available ACS sequences from YeastMine database[3] populated by SGD.

Evaluating the Sequence Similarity of ARSs

Sequences similarity analysis of ARSs in *S. cerevisiae* S288C reference genome was conducted by local BLAST 2.7.1+ (Camacho et al., 2009) (cutoff: E-value $\geq$ 5e-10, identity $\geq$90%, coverage $\geq$90%). Subsequently, the ARS sequences that have multiple alignment results were visualized by ClicO FS (Cheong et al., 2015).

Scan of Homologous ARSs in 104 *S. cerevisiae* Strains

In this study, a non-redundant dataset consisting of 520 ARSs was constructed based on the available ARSs of *S. cerevisiae* S288C from SGD database (Cherry et al., 2012) and supplemented with the confirmed ARSs from OriDB database (Nieduszynski et al., 2006a)[4] and ARSs of *S. cerevisiae* from DeOri 6.0 database (Gao et al., 2012)[5]. For each ARS in this dataset, we performed a BLAST against 104 *S. cerevisiae* genomes (cutoff: E-value $\geq$ 5e-10, identity $\geq$90%, coverage $\geq$90%). Then, a custom Python script was used to extract the aligned information, which was converted into a GFF3 format file. The conservative profile of ARSs was statistically analyzed according to the frequency of homologous ARSs among 104 *S. cerevisiae* strains. Finally, we divided the

annotation file of sorted homologous ARSs into 104 separate annotation files based on strain names.

Extraction of the Protein-Coding Genes Adjacent to Replication Origins

Firstly, we merged the ARS annotation files with the CDS annotation files of the corresponding 104 *S. cerevisiae* strains. Then, a custom Python script was used to extract the adjacent genes on both sides of ARSs, and the interval or intersected distance between ARSs and their adjacent genes were calculated. Subsequently, the sequences of these genes adjacent to ARSs among 104 *S. cerevisiae* strains were collected to BLAST against the dataset obtained from the latest version (February 2019) of UniProtKB[6] by the "blastp" program with the e-value cutoff of 1e-5.

Functional Enrichment Analysis

Lists of genes adjacent to conserved ARSs and non-conserved ARSs were prepared through the above steps. Then the gene lists were submitted to DAVID website (Huang et al., 2008) to perform the enrichment analysis of the Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways by Fisher's exact test. False discovery rate (fdr) was used to filter out the false-positive results with the cutoff of 0.05 for statistical significance.

RESULTS AND DISCUSSION

Characteristic Analysis of ARSs in *S. cerevisiae* Genomes

Overview of ARSs in *S. cerevisiae* S288C Reference Genome

There are a total of 352 ARSs in *S. cerevisiae* S288C reference genome available in the SGD database (**Supplementary Table 1**). The length of these experimentally verified ARSs ranged from 51 to 1324 bp and mainly (63.63%) concentrated in 70–250 bp (**Supplementary Figure 1A**), and the median length of ARSs in each chromosome is mainly around 240 bp (**Supplementary Figure 1B**). By linear regression analysis, the fitting line suggested that the count of ARSs was positively correlated to the length of chromosomes with the correlation coefficient of 0.7758 (**Supplementary Figure 1C**).

The *S. cerevisiae* S288C genome sequence has an average GC content of 38.38%, while that of ORI sequences is only 29.65%. With the accumulation of experimental data, there are 196 published ACS sequences in the YeastMine database populated by SGD. About 87.43% of ARSs that contain the ACS element are within the length of 300 bp. Here, we took the ACS element as the center to explore the base distribution of ARSs (**Supplementary Figure 2**) by WebLogo plot (Crooks et al., 2004). The ACS is visible as the high central peak with a high proportion of T residues. With the increasing number of ACS subjects, the

degenerate ACS is slightly changed from 5′-WTTTATRTTTW-3′ (Broach et al., 1983) to 5′-WTTTAYRTTTW-3′ (Marahrens and Stillman, 1992). In this study, we collected those published 196 ACS sequences from YeastMine and generated the matrix profile of ACS motif (**Supplementary File 1**) using "Bio.motifs" package included in Biopython for subsequent prediction of ACS in the candidate sequence. A broad region located directly 3′ to the ACS termed B elements (Lucas and Raghuraman, 2003), which showed low sequence similarity among various ARSs. And only minimal conservations located 3′ to the ACS were detected, for example the conserved 5′-TT-3′ of B1 element and a relatively high frequency of adenine residues start from around 30 to 110 bp, which are consistent with the previous research (Broach et al., 1983; Huang and Kowalski, 1996; Lucas and Raghuraman, 2003; Breier et al., 2004).

In order to investigate the uniqueness of the ARS sequences in *S. cerevisiae* S288C yeast genome, we conducted a sequence similarity analysis of all these 352 annotated ARSs. The majority of ARS sequences (94.32%) are unique, while only 20 ARSs have multiple alignment results (**Supplementary File 2**) distributed in intra-chromosomes and inter-chromosomes (**Figure 1**). Interestingly, all the similar ARS pairs distributed in inter-chromosomes are biased to locate in subtelomeric regions generally within 20 kb of both ends of yeast chromosome (Brown et al., 2010). Yue et al. (2017) found that subtelomeres possess a higher level of copy number variants (CNV) accumulation than those from the internal chromosomal cores and non-reciprocal exchanges and duplications among subtelomeric regions appear to be widespread among eukaryotes (Eichler and Sankoff, 2003), which supported our findings. We subsequently scanned the homologous ARSs of these 20 ARSs among 104 *S. cerevisiae* strains (**Supplementary Table 2**), and we found that the most conserved pairs among 104 strains are two pairs internal ARSs compared with those of subtelomeric ARSs. One is the pair of ARS 810 and ARS 811 closed to the tandem array of CUP1 that are associated with resistance to the toxicity of copper (Fogel and Welch, 1982). The other is the pair of ARS1200-1 and ARS 1200-2 known as rARSs (Miller and Kowalski, 1993) that are associated with yeast life span (Kwan et al., 2013). These results suggested that compared to similar ARS pairs distributed in inter-chromosomes, similar ARS pairs located in intra-chromosomes prefer to be shared between strains.
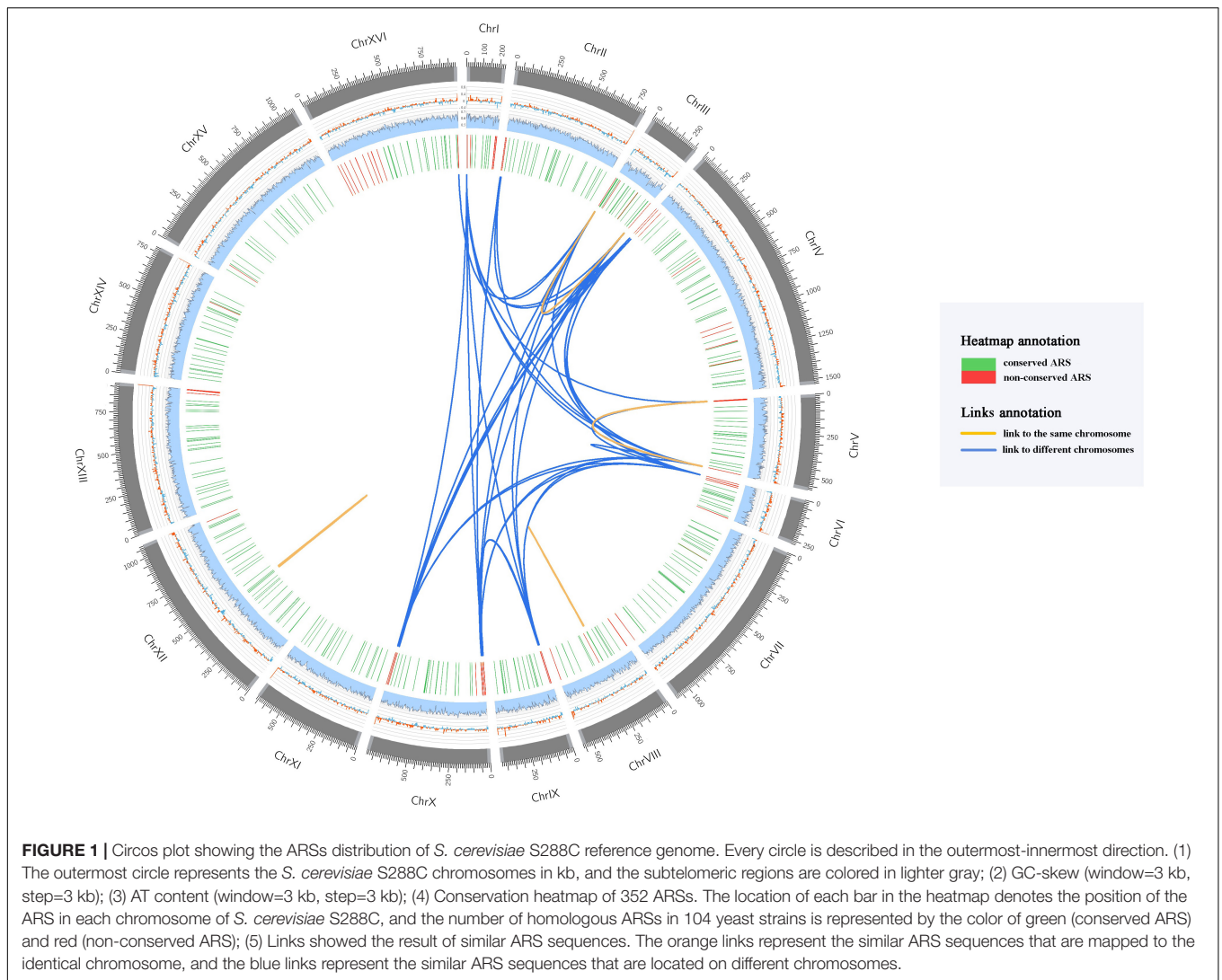
Population Genomic Analysis of ARSs Among 104 Yeast Genomes

The 104 *S. cerevisiae* strains that we focused on in this study showed broad genotypic and phenotypic diversity (Strope et al., 2015; Peter et al., 2018). What is the extent of distribution of replication origins in these *S. cerevisiae* strains with various phylogenetic distance? Therefore, we built a non-redundant dataset with a total of 520 ARSs based on ARSs annotation of *S. cerevisiae* S288C from SGD and supplemented with the confirmed ARSs from OriDB and DeOri database (**Supplementary Table 4**).

A marked similarity between two nucleotide sequences may reflect the fact that they come from the same ancestral

sequence driven by evolution (Petsko and Ringe, 2004). By mapping ARS sequences from the non-redundant dataset to 104 *S. cerevisiae* genomes, we determined the homologous ARSs in each chromosome among these yeast strains. Even though the strict alignment conditions had been set, plentiful homologous ARS sequences were found among various yeast genomes and the majority of the homologous ARSs were mapped to the chromosome of the corresponding ARS of the non-redundant dataset. We measured the proportion of homologous ARSs in all strains within the species of *S. cerevisiae* and displayed it in heat map plot (**Supplementary Figure 3**), illustrating that the conservation profile of ARSs is not evenly distributed along the chromosomes. It should be noted that only the homologous ARSs located in the corresponding chromosome of the 520 ARS dataset were collected. As for the non-unique ARSs, their chromosomal regions were also taken into consideration. According to the number of homologous ARSs in 104 budding yeast strains, we defined those ARSs existing in more than 90% of the yeast strains as conserved ARSs, and the rest as non-conserved ARSs. We found 430 conserved ARSs accounting for 82.7% of the ARSs from the dataset, and these ARSs are not only conserved in sequence but also relatively conserved in the chromosomal position among various *S. cerevisiae* genomes (**Supplementary Figure 4A**), which likely served as the organizational framework for *S. cerevisiae* genomes. Although large-scale structural variants might exist in the chromosome XII of different yeast strains, the relative position between ARSs in the fragments with structural variants are conserved. Interestingly, about 80% of ARSs located in the subtelomeric regions are non-conserved ARSs. The number of homologous ARSs from subtelomeres is less than those from internal chromosomal regions (one-side Mann–Whitney *U* test, *p*-value < 0.01). Subtelomeric regions profoundly contribute to genetic and phenotypic diversity, which recognized as peculiarly dynamic regions of chromosomal evolution (Eichler and Sankoff, 2003; Dujon, 2010; Yue et al., 2017). These areas with rampant genomic rearrangement are the hotspots of reciprocal translocations (Eichler and Sankoff, 2003), which could interpret that ARSs located in subtelomeric regions possess lower conservation than those located in internal chromosomal regions. Subtelomeric regions showed a strong relevance of rapid adaptation to novel niches (Brown et al., 2010; Bergström et al., 2014; Yue et al., 2017), and were known as the hot spots of genetic variation (Peter et al., 2018), which helps accelerate genome evolution and divergence (Cohn et al., 2006). The distribution of non-conserved ARSs in subtelomeres of different *S. cerevisiae* strains may reflect the strain specificity during genome evolution.

We also performed the pan-genome analysis of ARSs (pan-ARSs) among 104 *S. cerevisiae* strains by PanGP (Zhao et al., 2014) software. The pan-ARSs size curve illustrated closed pan-ARSs (**Supplementary Figure 5**). Since we adopted limited ARSs pool to map to yeast strains, it seems that the additional strains could not provide new ARSs to *S. cerevisiae* pan-ARSs. The result showed that a small number of yeast strains are sufficient to cover the majority of *S. cerevisiae* pan-ARSs. Core ARSs represent the ARSs exist in all strains of *S. cerevisiae*. The core ARSs curve showed that the size of core ARS

FIGURE 1 | Circos plot showing the ARSs distribution of *S. cerevisiae* S288C reference genome. Every circle is described in the outermost-innermost direction. (1) The outermost circle represents the *S. cerevisiae* S288C chromosomes in kb, and the subtelomeric regions are colored in lighter gray; (2) GC-skew (window=3 kb, step=3 kb); (3) AT content (window=3 kb, step=3 kb); (4) Conservation heatmap of 352 ARSs. The location of each bar in the heatmap denotes the position of the ARS in each chromosome of *S. cerevisiae* S288C, and the number of homologous ARSs in 104 yeast strains is represented by the color of green (conserved ARS) and red (non-conserved ARS); (5) Links showed the result of similar ARS sequences. The orange links represent the similar ARS sequences that are mapped to the identical chromosome, and the blue links represent the similar ARS sequences that are located on different chromosomes.

approached to a constant value, suggesting that these 183 core ARSs might serve as the organizational framework for the *S. cerevisiae* genome.

The total number of homologous ARSs corresponding to the chromosome of ARS from non-redundant dataset for each strain was calculated, and the amount of data we obtained was sufficient for subsequent analysis. Based on the geographic and environmental origins of yeast strains (**Supplementary Table 5**), we classified the strains into several subsets. By illustrating the distribution of homologous ARSs in each subset (**Supplementary Figure 6**), we found that the average number of homologous ARSs in each category are relatively similar, however, due to distinct ecological niche and various degree of human association of the isolated strains, the data fluctuation range are various from each class, which may underline a key role of human-driven activities in shaping the distribution of ARSs in *S. cerevisiae* (Strope et al., 2015; Peter et al., 2018).

In the process of DNA replication, the ORC is recruited to replication origins, followed by the binding of CDC6 (cell division cycle 6) and CDT1 (Cdc10-dependent transcript 1)

as well as loading of the MCM helicase complex, which formed the pre-RC (pre-replication complex) proteins (Fragkos et al., 2015). The pre-RCs would bind to all potential origins, however, potential replication origins are in excess and only a small fraction of assembled pre-RCs will be activated at each cell cycle. In addtion, the activation of pre-RCs does not occur simultaneously. Some are fired in the early S phase, and others are activated in the mid or late S phase (Méchali, 2010). The number of corresponding homologous ARSs found in different yeast strains showed the conservation of ARS within the species of *S. cerevisiae*. To analyze the correlation coefficient between the ARS conservation and the replication fire time, the data of the replication time (Raghuraman et al., 2001) together with the number of homologous ARSs among *S. cerevisiae* strains were adopted. The result showed that the conservation of ARSs was non-randomly associated with replication time, the Pearson correlation between the conservation of ARSs and replication time shows the value of -0.484 (p-value < 0.01). By comparing the different replication time between the conserved ARSs and non-conserved ARSs (**Supplementary Figure 7**), we found that

the higher conservation of ARSs, the earlier it might initiate (the pairwise Wilcox.test, *p*-value < 0.001). Combining with the previous findings, we could conclude that in the species of budding yeast, the ARSs biased toward the subtelomeric regions tends to possess weaker conservation and later replicated fire time, which was consistent with the previous conclusion that subtelomeric regions generally possess late DNA replication and low levels of transcription (Barton et al., 2003; Yamazaki et al., 2013). We could also infer that those conservative and earlier replicated replication origins may possess more vital missions than others in chromosomes, for instance their neighboring genes have the priority to early replicate to maintain the growth of yeast strains.

Functional Analysis of Genes Adjacent to Replication Origins in *S. cerevisiae* Genomes

Genes Adjacent to ARSs in *S. cerevisiae* S288C Reference Genome

In bacteria, the distribution of *oriC*s and its corresponding adjacent replication-related genes such as *dnaA*, *dnaN* or *gidA* are highly conserved among different phyla and around 43% of the *oriC*s are biased close to *dnaA* among a total of 2740 bacterial chromosomes distributed in various phyla (Luo et al., 2018). The relationship between the *oriC* and adjacent replication-related genes has been successfully applied to predict the location of *oriC*s in bacterial chromosomes (Gao and Zhang, 2008). In archaea, replication origins are found to locate next to *cdc6/cdc1* (Norais et al., 2007). It is worth surveying the distribution profile of eukaryotic replication origins and their corresponding adjacent genes.

It is generally accepted that the locations of replication origins are exclusively restricted to intergenic regions in eukaryotes (Brewer, 1994; Gilbert, 2001). Here we extracted protein-coding genes adjacent to replication origins of the well-annotated reference genome of *S. cerevisiae* S288C. In recent years, studies on minimal ARS (miniARS) in yeast have been reported (Liachko et al., 2013; Tsai et al., 2014). However, research on systematically and accurately identifying the precise boundaries of minimal functional replication regions have not been performed due to a large number of replication origins in yeast chromosomes. Please note that the boundaries of the ARSs used in this study are all collected from the original literature, which may be not confined to the minimum essential regions.

According to *S. cerevisiae* reference genome annotation, we classified the replication origins based on the positional relationships between ARSs and their adjacent genes in the chromosomes. It is defined as the intergenic ORI if there is no intersection between replication origin and its corresponding adjacent protein-coding genes, otherwise as the intersected ORI that the replication origin sequence partially or completely overlaps the adjacent protein-coding genes (**Supplementary Figure 8**). The result showed that the intergenic ORIs account for 68.18% of known replication origins of *S. cerevisiae* S288C. Distance distribution among ARSs and their adjacent protein-coding genes showed that their interval distances are mainly

less than 1000 bp (**Supplementary Figure 9**). Although there are 112 intersected ORIs with the average length of 395 bp, about 55.35% of their overlapped segments are less than 30% of their own lengths. Since the ACS element is an essential and conservative element in ARSs, we subsequently scanned the overlapped segments between the intersected ORI and its overlapping protein-coding genes using the matrix profile of ACS motif (**Supplementary File 1**). The result showed that there are 40 overlapped segments contain ACS motif, suggesting the overlapped segments may be important for these ORIs. We also identified the repeats in ARSs executed by REPuter (Kurtz et al., 2001) program (options:./repfind -c -f -p -r -l 8 -best 50 -h 0 –s) with the *e*-value cutoff of 5e-2. We found that the majority (92.90%) of ARSs contain repeats (**Supplementary Table 3**). For intergenic ORIs, repeats (average AT content of 91.10%) are generally characterized by continuous A base, continuous T base, or alternating repeats of A and T base (**Supplementary Figure 10**). However, for the overlapped segment with ACS motif between the intersected ORI and its overlapping protein-coding gene, repeats in these segments (average AT content of 83.52%) are biased to possess higher GC content (**Supplementary Table 3**), which suggest that the sequence composition of the overlapped segment between the intersected ORI and its overlapping gene may be constrained by the gene composition.

For *S. cerevisiae*, the effectiveness of ARSs is various, and restraining the initiation of certain ARSs could affect the expression of neighboring genes. Histone gene pairs (HTA1-HTB1, HHT1-HHF1) are closely positioned to replication origins (ARS428, ARS209) in *S. cerevisiae* S288C genome. Inactivation of ARSs that are proximal to HTA1–HTB1 gene pairs significantly delayed replication of HTA1 and HTB1, resulting in halving the expression of histone genes (Muller and Nieduszynski, 2017). The delay in replication of centromeric regions (including ARS919 and ARS920) contributes to chromosome instability (Natsume et al., 2013). Nevertheless, *S. cerevisiae* with multiple origin deletions (ARS600, ARS601/2, ARS603, ARS603.5, ARS604, ARS506, and ARS606) in chromosome VI can replicate relatively normally without detectable growth defects (Dershowitz et al., 2007). Essential genes are those indispensable for the survival of an organism (Giaever et al., 2002), and we found 47 intersected ORIs overlapped with essential genes (data from DEG database[7]) (Gao et al., 2015; **Supplementary Table 1**). Any genetic variation occurred in the overlapped region may cause changes in both replication origin and the essential gene, which may disturb the stability and integrity of genomes and even threaten the viability of yeast cells.

Subsequently, in order to assess which factors could determine the difference between the intergenic ORIs and the intersected ORIs, principal component analysis (**Supplementary Figure 11**) was conducted by integrating the comprehensive features of ARSs in *S. cerevisiae* S288C including length, GC content, the positional relationship with the adjacent gene, relative chromosomal position, the number of homologous ARSs we obtained among *S. cerevisiae* species as well as replication time [data from Raghuraman et al. (2001)] and gene expression
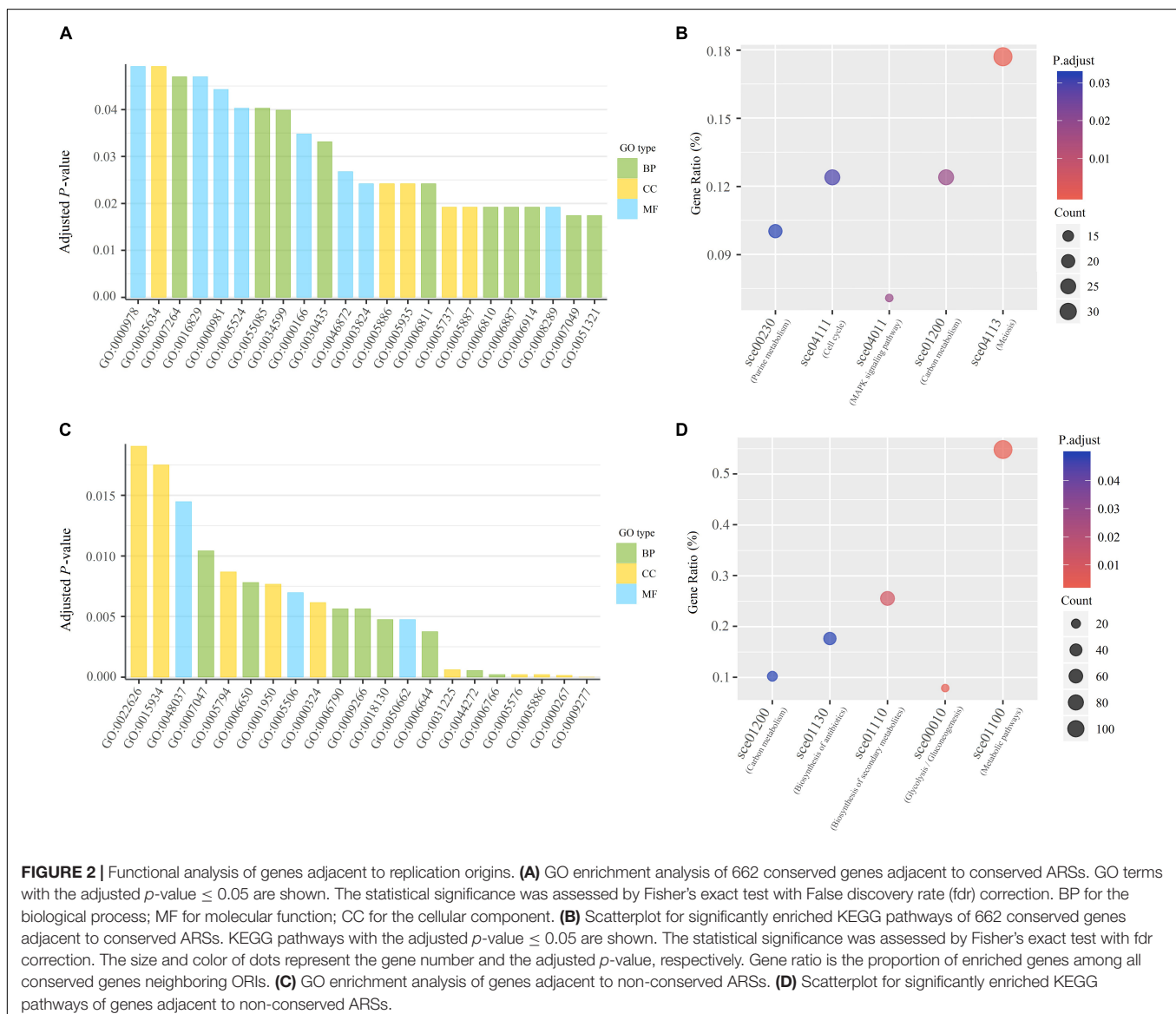
---

[7]http://tubic.org/deg/

profiles of *S. cerevisiae* [data from Arava et al. (2003)]. The intergenic ORIs and intersected ORIs showed relatively visible distinction in the PCA score plot. The average expression of genes adjacent to intersected ORI was significantly lower than that of genes adjacent to intergenic ORI (one-side Mann–Whitney *U* test, *p*-value < 0.05). We guessed that replication-related proteins that bind or be recruited at replication origins may interfere with the expression of overlapping genes, and the influence of expression of overlapping genes has to be considered in space as well as time, because only a subset of origins are activated during every cell cycle and the activation of replication origins may vary according to the cell fate or environmental conditions (Méchali, 2010; Fragkos et al., 2015). However, the comparison was only based on the inference of statistical results, and the specific relationships between replication origins and their adjacent genes require more detailed experimental studies. We found that most of the above mentioned factors were linearly uncorrelated to the

ORIs that possessed the various positional relationship with the adjacent gene, but the intergenic ORIs and the intersected ORIs could be roughly distinguished through the PCA analysis if these factors were comprehensively considered.

Genes Adjacent to Conserved ORIs Among 104 *S. cerevisiae* Genomes

In the study of population genomic analysis of ORIs among 104 yeast genomes, we identified 430 conserved ARSs from the ARSs dataset. Based on the homologous ARSs of various yeast strains, we extracted their adjacent genes from the corresponding yeast genomes. According to the number of genes located next to each of the corresponding conserved ORIs in 104 budding yeast strains, we defined the genes existing in more than 90% of the yeast strains as conserved adjacent genes, and the rest as non-conserved adjacent genes. As a result, a total of 662 conserved adjacent genes were collected and 27.64% of them

FIGURE 2 | Functional analysis of genes adjacent to replication origins. **(A)** GO enrichment analysis of 662 conserved genes adjacent to conserved ARSs. GO terms with the adjusted *p*-value ≤ 0.05 are shown. The statistical significance was assessed by Fisher's exact test with False discovery rate (fdr) correction. BP for the biological process; MF for molecular function; CC for the cellular component. **(B)** Scatterplot for significantly enriched KEGG pathways of 662 conserved genes adjacent to conserved ARSs. KEGG pathways with the adjusted *p*-value ≤ 0.05 are shown. The statistical significance was assessed by Fisher's exact test with fdr correction. The size and color of dots represent the gene number and the adjusted *p*-value, respectively. Gene ratio is the proportion of enriched genes among all conserved genes neighboring ORIs. **(C)** GO enrichment analysis of genes adjacent to non-conserved ARSs. **(D)** Scatterplot for significantly enriched KEGG pathways of genes adjacent to non-conserved ARSs.

belong to essential genes based on DEG database (Gao et al., 2015; **Supplementary Table 4**). We also found that the conserved adjacent genes relatively conserved in both chromosomal position and orientation among the various *S. cerevisiae* genomes (**Supplementary Figure 4B**), which suggested that the adjacent relationship between ARSs and their corresponding genes are conserved in chromosomes among 104 yeast strains.

In bacteria, the replication-related genes such as *dnaA* and *dnaN*, are highly conservatively close to *oriC*s (Luo et al., 2018). Likewise, genes involved in the initiation of DNA replication in *S. cerevisiae* are found to conservatively locate next to the replication origins, such as *orc3*, *mcm2*, *mcm4*, *mcm6*, and *cdc45* (**Supplementary Table 4**). During G1 phase, ORC complex recognizes and bind sequence-specifically to ACS in the presence of ATP (Bell and Stillman, 1992; Watson et al., 2013). And helicase-loading proteins, CDC6 and CDT1, are recruited to load MCM2–7 complexes onto the replication origin (Bell and Labib, 2016). During S phase, loaded helicases are activated by CDK (cyclin dependent kinase) and DDK (Dbf4-dependent kinase). Those two factors, CDC45 (cell division cycle 45) and GINS (Go, Ichi, Ni, and San) complex are tightly associated with MCM2-7 at replication forks to form the activated helicase called the CMG complex (CDC45–MCM–GINS) (Moyer et al., 2006; Pacek et al., 2006). Then the CMG complex would be well assembled and activated to unwind the double-stranded DNA and start to initiate DNA synthesis (Fragkos et al., 2015).

We found that the conserved adjacent genes were significantly enriched in twenty-three GO terms and five KEGG pathways related to DNA binding, enzyme activity, transportation and energy, including sequence-specific DNA binding, nucleotide binding, lyase activity, catalytic activity, ion transport, transmembrane transport and ATP binding (**Figures 2A,B** and **Supplementary Table 6**). Obviously, these GO terms and KEGG pathways are strongly correlated with the process of DNA replication. Due to the advantage in chromosomal position, the genes adjacent to replication origins are preferentially replicated after the double-stranded DNA start unwinding. We could infer that replication origin neighboring genes involved in DNA binding, enzyme activity, transportation, and energy might have a higher priority to replicate. It is likely that the aggregation of ORI and its conserved adjacent genes enable the gene-encoded products to be more effectively involved in the DNA replication initiation in the localized cellular space. It is possible that these preferential replicated genes enable normal and efficient DNA replication and possibly make it orderly organized in protein-DNA and protein-protein interactions, which may guarantee stable operations in the yeast cell cycle. With regard to non-conserved ARSs, also considered as strain-specific ARSs, we found that these genes are significantly enriched in response to environmental stress (such as temperature and drug), metabolites biosynthetic process and biosynthesis of antibiotics (**Figures 2C,D** and **Supplementary Table 7**). We speculated that the preferential replication related to these adjacent genes is likely to provide raw materials for the active metabolism of yeast strains, and may enhance the adaptability of the strains to survive in the environmental stress.

CONCLUSION

In this study, we comprehensively analyzed features of replication origin sequences of *S. cerevisiae* from genome-wide and population genomics perspectives. We conducted the data-analytic work for investigating the similarities and genomic positions of the ARS sequences among the diverse budding yeast strains obtained from various ecological and geographical backgrounds. We also performed a characterization of the genes that are adjacent to the conserved and non-conserved ARSs among the 104 yeast strains. These results presented here may provide insights into the replication mechanism of *S. cerevisiae* and facilitate the development of algorithms for further prediction of replication origins in budding yeast genomes. For examples, the conserved ARS-adjacent genes should be taken into consideration in the design of prediction algorithms, just like Ori-Finder considering the conserved *oriC*-adjacent genes (such as *dnaA*, *dnaN*, and *gidA*), which would make the prediction more robust and reliable. In addition, as modular parts, the core ARSs and their conserved adjacent genes might provide a useful reference for the rational design of replication origins for the synthetic *S. cerevisiae* genome. However, the conserved ARSs and their corresponding adjacent genes are obtained based on sequence alignment and statistical results, whereas the biological significance of the positional conservation of the replication origins and their adjacent genes requires more detailed experimental proof. Since DNA replication is one of the highly conserved processes of eukaryotic cell, and almost all the proteins related to the DNA replication in yeast correspond to a single ortholog in humans and other eukaryotic species (Bell and Labib, 2016), the features and rules of DNA replication initiation found in *S. cerevisiae* genomes may be extended to higher eukaryotes.

AUTHOR CONTRIBUTIONS

DW conducted the data analysis and drafted the manuscript. FG supervised the study and revised the manuscript. Both authors read and approved the final manuscript.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2019.02122/full#supplementary-material

REFERENCES

Arava, Y., Wang, Y., Storey, J. D., Liu, C. L., Brown, P. O., and Herschlag, D. (2003). Genome-wide analysis of mRNA translation profiles in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U.S.A.* 100, 3889–3894. doi: 10.1073/pnas. 0635171100

Barton, A. B., Su, Y., Lamb, J., Barber, D., and Kaback, D. B. (2003). A function for subtelomeric DNA in *Saccharomyces cerevisiae*. *Genetics* 165, 929–934.

Bell, S. P., and Labib, K. (2016). Chromosome duplication in *Saccharomyces cerevisiae*. *Genetics* 203, 1027–1067. doi: 10.1534/genetics.115.186452

Bell, S. P., and Stillman, B. (1992). ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature* 357, 128–134. doi: 10.1038/357128a0

Bergström, A., Simpson, J. T., Salinas, F., Barré, B., Parts, L., Zia, A., et al. (2014). A high-definition view of functional genetic variation from natural yeast genomes. *Mol. Biol. Evol.* 31, 872–888. doi: 10.1093/molbev/msu037

Bogenschutz, N. L., Rodriguez, J., and Tsukiyama, T. (2014). Initiation of DNA replication from non-canonical sites on an origin-depleted chromosome. *PLoS One* 9:e114545. doi: 10.1371/journal.pone.0114545

Breier, A. M., Chatterji, S., and Cozzarelli, N. R. (2004). Prediction of *Saccharomyces cerevisiae* replication origins. *Genome Biol.* 5:R22.

Brewer, B. J. (1994). Intergenic DNA and the sequence requirements for replication initiation in eukaryotes. *Curr. Opin. Genet. Dev.* 4, 196–202. doi: 10.1016/ s0959-437x(05)80045-0

Brewer, B. J., and Fangman, W. L. (1987). The localization of replication origins on ARS plasmids in S. cerevisiae. *Cell* 51, 463–471. doi: 10.1016/0092-8674(87) 90642-8

Broach, J. R., Li, Y. Y., Feldman, J., Jayaram, M., Abraham, J., Nasmyth, K. A., et al. (1983). Localization and sequence analysis of yeast origins of DNA replication. *Cold Spring Harb. Symp. Quant. Biol.* 47(Pt 2), 1165–1173. doi: 10.1101/sqb. 1983.047.01.132

Brown, C. A., Murray, A. W., and Verstrepen, K. J. (2010). Rapid expansion and functional divergence of subtelomeric gene families in yeasts. *Curr. Biol.* 20, 895–903. doi: 10.1016/j.cub.2010.04.027

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinformatics* 10:421. doi: 10.1186/1471-2105-10-421

Cheong, W.-H., Tan, Y.-C., Yap, S.-J., and Ng, K.-P. (2015). ClicO FS: an interactive web-based service of Circos. *Bioinformatics* 31, 3685–3687. doi: 10.1093/ bioinformatics/btv433

Cherry, J. M., Hong, E. L., Amundsen, C., Balakrishnan, R., Binkley, G., Chan, E. T., et al. (2012). Saccharomyces genome database: the genomics resource of budding yeast. *Nucleic Acids Res.* 40, D700–D705. doi: 10.1093/nar/gkr1029

Cohn, M., Liti, G., and Barton, D. B. (2006). "Telomeres in fungi," in *Comparative Genomics*, ed. Pa. P. J Sunnerhagen, (Heidelberg: Springer), 101–130. doi: 10. 1007/4735_108

Cotterill, S., and Kearsey, S. E. (2008). DNAReplication: a database of information and resources for the eukaryotic DNA replication community. *Nucleic Acids Res.* 37, D837–D839. doi: 10.1093/nar/gkn726

Crooks, G. E., Hon, G., Chandonia, J.-M., and Brenner, S. E. (2004). WebLogo: a sequence logo generator. *Genome Res.* 14, 1188–1190. doi: 10.1101/gr.849004

Dao, F. Y., Lv, H., Wang, F., Feng, C. Q., Ding, H., Chen, W., et al. (2019). Identify origin of replication in *Saccharomyces cerevisiae* using two-step feature selection technique. *Bioinformatics* 35, 2075–2083. doi: 10.1093/bioinformatics/bty943

Dershowitz, A., Snyder, M., Sbia, M., Skurnick, J. H., Ong, L. Y., and Newlon, C. S. (2007). Linear derivatives of *Saccharomyces cerevisiae chromosome* III can be maintained in the absence of autonomously replicating sequence elements. *Mol. Cell. Biol.* 27, 4652–4663. doi: 10.1128/mcb.01246-06

Diffley, J. F., and Stillman, B. (1988). Purification of a yeast protein that binds to origins of DNA replication and a transcriptional silencer. *Proc. Natl. Acad. Sci. U.S.A.* 85, 2120–2124. doi: 10.1073/pnas.85.7.2120

Duderstadt, K. E., and Berger, J. M. (2008). AAA+ ATPases in the initiation of DNA replication. *Crit. Rev. Biochem. Mol. Biol.* 43, 163–187. doi: 10.1080/ 10409230802058296

Dujon, B. (2010). Yeast evolutionary genomics. *Nat. Rev. Genet.* 11, 512–524. doi: 10.1038/nrg2811

Eaton, M. L., Galani, K., Kang, S., Bell, S. P., and Macalpine, D. M. (2010). Conserved nucleosome positioning defines replication origins. *Genes Dev.* 24, 748–753. doi: 10.1101/gad.1913210

Eichler, E. E., and Sankoff, D. (2003). Structural dynamics of eukaryotic chromosome evolution. *Science* 301, 793–797. doi: 10.1126/science.1086132

Fogel, S., and Welch, J. W. (1982). Tandem gene amplification mediates copper resistance in yeast. *Proc. Natl. Acad. Sci. U.S.A.* 79, 5342–5346. doi: 10.1073/ pnas.79.17.5342

Fragkos, M., Ganier, O., Coulombe, P., and Mechali, M. (2015). DNA replication origin activation in space and time. *Nat. Rev. Mol. Cell Biol.* 16, 360–374. doi: 10.1038/nrm4002

Friedman, K. L., Brewer, B. J., and Fangman, W. L. (1997). Replication profile of *Saccharomyces cerevisiae* chromosome VI. *Genes Cells* 2, 667–678.

Gao, F., Luo, H., and Zhang, C. T. (2012). DeOri: a database of eukaryotic DNA replication origins. *Bioinformatics* 28, 1551–1552. doi: 10.1093/bioinformatics/ bts151

Gao, F., Luo, H., Zhang, C.-T., and Zhang, R. (2015). Gene essentiality analysis based on DEG 10, an updated database of essential genes. *Methods Mol. Biol.* 1279, 219–233. doi: 10.1007/978-1-4939-2398-4_14

Gao, F., and Zhang, C.-T. (2008). Ori-Finder: a web-based system for finding oriC s in unannotated bacterial genomes. *BMC Bioinformatics* 9:79. doi: 10.1186/1471-2105-9-79

Giaever, G., Chu, A. M., Ni, L., Connelly, C., Riles, L., Veronneau, S., et al. (2002). Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* 418, 387–391.

Gilbert, D. M. (2001). Making sense of eukaryotic DNA replication origins. *Science* 294, 96–100. doi: 10.1126/science.1061724

Grigoriev, A. (1998). Analyzing genomes with cumulative skew diagrams. *Nucleic Acids Res.* 26, 2286–2290. doi: 10.1093/nar/26.10.2286

Hand, R. (1978). Eucaryotic DNA: organization of the genome for replication. *Cell* 15, 317–325. doi: 10.1016/0092-8674(78)90001-6

Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2008). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57. doi: 10.1038/nprot.2008.211

Huang, R. Y., and Kowalski, D. (1996). Multiple DNA elements in ARS305 determine replication origin activity in a yeast chromosome. *Nucleic Acids Res.* 24, 816–823. doi: 10.1093/nar/24.5.816

Jacob, F., Brenner, S., and Cuzin, F. (1963). On the regulation of DNA replication in bacteria. *Cold Spring Harb. Symp. Quant. Biol.* 28, 329–348.

Kurtz, S., Choudhuri, J. V., Ohlebusch, E., Schleiermacher, C., Stoye, J., and Giegerich, R. (2001). REPuter: the manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Res.* 29, 4633–4642. doi: 10.1093/nar/29.22. 4633

Kwan, E. X., Foss, E. J., Tsuchiyama, S., Alvino, G. M., Kruglyak, L., Kaeberlein, M., et al. (2013). A natural polymorphism in rDNA replication origins links origin activation with calorie restriction and lifespan. *PLoS Genet.* 9:e1003329. doi: 10.1371/journal.pgen.1003329

Lee, W., Tillo, D., Bray, N., Morse, R. H., Davis, R. W., Hughes, T. R., et al. (2007). A high-resolution atlas of nucleosome occupancy in yeast. *Nat. Genet.* 39, 1235–1244. doi: 10.1038/ng2117

Li, N., Lam, W. H., Zhai, Y., Cheng, J., Cheng, E., Zhao, Y., et al. (2018). Structure of the origin recognition complex bound to DNA replication origin. *Nature* 559, 217–222. doi: 10.1038/s41586-018-0293-x

Liachko, I., Youngblood, R. A., Keich, U., and Dunham, M. J. (2013). High-resolution mapping, characterization, and optimization of autonomously replicating sequences in yeast. *Genome Res.* 23, 698–704. doi: 10.1101/gr. 144659.112

Liu, B., Weng, F., Huang, D.-S., and Chou, K.-C. (2018). iRO-3wPseKNC: identify DNA replication origins by three-window-based PseKNC. *Bioinformatics* 34, 3086–3093. doi: 10.1093/bioinformatics/bty312

Lobry, J. (1996). Asymmetric substitution patterns in the two DNA strands of bacteria. *Mol. Biol. Evol.* 13, 660–665. doi: 10.1093/oxfordjournals.molbev. a025626

Lucas, I. A., and Raghuraman, M. K. (2003). The dynamics of chromosome replication in yeast. *Curr. Top. Dev. Biol.* 55, 1–73. doi: 10.1016/s0070-2153(03) 01001-9

Luo, H., Quan, C. L., Peng, C., and Gao, F. (2018). Recent development of Ori-Finder system and DoriC database for microbial replication origins. *Brief. Bioinformatics* doi: 10.1093/bib/bbx174 [Epub ahead of print].

Luo, H., Zhang, C. T., and Gao, F. (2014). Ori-Finder 2, an integrated tool to predict replication origins in the archaeal genomes. *Front. Microbiol.* 5:482. doi: 10.3389/fmicb.2014.00482

Mackiewicz, P., Zakrzewska-Czerwiñska, J., Zawilak, A., Dudek, M. R., and Cebrat, S. (2004). Where does bacterial replication start? Rules for predicting the *oriC* region. *Nucleic Acids Res.* 32, 3781–3791. doi: 10.1093/nar/gkh699

Marahrens, Y., and Stillman, B. (1992). A yeast chromosomal origin of DNA replication defined by multiple functional elements. *Science* 255, 817–823. doi: 10.1126/science.1536007

Méchali, M. (2010). Eukaryotic DNA replication origins: many choices for appropriate answers. *Nat. Rev. Mol. Cell Biol.* 11, 728–738. doi: 10.1038/nrm2976

Miller, C. A., and Kowalski, D. (1993). cis-acting components in the replication origin from ribosomal DNA of *Saccharomyces cerevisiae*. *Mol. Cell Biol.* 13, 5360–5369. doi: 10.1128/mcb.13.9.5360

Moyer, S. E., Lewis, P. W., and Botchan, M. R. (2006). Isolation of the Cdc45/Mcm2-7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase. *Proc. Natl. Acad. Sci. U.S.A.* 103, 10236–10241. doi: 10.1073/pnas.0602400103

Müller, C. A., Hawkins, M., Retkute, R., Malla, S., Wilson, R., Blythe, M. J., et al. (2013). The dynamics of genome replication using deep sequencing. *Nucleic Acids Res.* 42:e3. doi: 10.1093/nar/gkt878

Muller, C. A., and Nieduszynski, C. A. (2017). DNA replication timing influences gene expression level. *J. Cell Biol.* 216, 1907–1914. doi: 10.1083/jcb.201701061

Natsume, T., Muller, C. A., Katou, Y., Retkute, R., Gierlinski, M., Araki, H., et al. (2013). Kinetochores coordinate pericentromeric cohesion and early DNA replication by Cdc7-Dbf4 kinase recruitment. *Mol. Cell* 50, 661–674. doi: 10.1016/j.molcel.2013.05.011

Newlon, C. S., Collins, I., Dershowitz, A., Deshpande, A. M., Greenfeder, S. A., Ong, L. Y., et al. (1993). Analysis of replication origin function on chromosome III of *Saccharomyces cerevisiae*. *Cold Spring Harb. Symp. Quant. Biol.* 58, 415–423. doi: 10.1101/sqb.1993.058.01.048

Nieduszynski, C. A., Hiraga, S.-I., Ak, P., Benham, C. J., and Donaldson, A. D. (2006a). OriDB: a DNA replication origin database. *Nucleic Acids Res.* 35, D40–D46.

Nieduszynski, C. A., Knox, Y., and Donaldson, A. D. (2006b). Genome-wide identification of replication origins in yeast by comparative genomics. *Genes Dev.* 20, 1874–1879. doi: 10.1101/gad.385306

Norais, C., Hawkins, M., Hartman, A. L., Eisen, J. A., Myllykallio, H., and Allers, T. (2007). Genetic and physical mapping of DNA replication origins in Haloferax volcanii. *PLoS Genet.* 3:e77. doi: 10.1371/journal.pgen.0030077

Pacek, M., Tutter, A. V., Kubota, Y., Takisawa, H., and Walter, J. C. (2006). Localization of MCM2-7, Cdc45, and GINS to the site of DNA unwinding during eukaryotic DNA replication. *Mol. Cell* 21, 581–587. doi: 10.1016/j.molcel.2006.01.030

Peter, J., De Chiara, M., Friedrich, A., Yue, J.-X., Pflieger, D., Bergström, A., et al. (2018). Genome evolution across 1,011 *Saccharomyces cerevisiae* isolates. *Nature* 556, 339–344. doi: 10.1038/s41586-018-0030-5

Petsko, G. A., and Ringe, D. (2004). *Protein Structure and Function. Chapter 4: From Sequence to Function.* London: New Science Press.

Quan, C. L., and Gao, F. (2019). Quantitative analysis and assessment of base composition asymmetry and gene orientation bias in bacterial genomes. *FEBS Lett.* 593, 918–925. doi: 10.1002/1873-3468.13374

Raghuraman, M., Winzeler, E. A., Collingwood, D., Hunt, S., Wodicka, L., Conway, A., et al. (2001). Replication dynamics of the yeast genome. *Science* 294, 115–121. doi: 10.1126/science.294.5540.115

Rocha, E. P., Danchin, A., and Viari, A. (1999). Universal replication biases in bacteria. *Mol. Microbiol.* 32, 11–16. doi: 10.1046/j.1365-2958.1999.01334.x

Salzberg, S. L., Salzberg, A. J., Kerlavage, A. R., and Tomb, J.-F. (1998). Skewed oligomers and origins of replication. *Gene* 217, 57–67. doi: 10.1016/s0378-1119(98)00374-6

Stinchcomb, D. T., Mann, C., Selker, E., and Davis, R. W. (1981). "DNA sequences that allow the replication and segregation of yeast chromosomes," in *The Initiation of DNA Replication*, ed. D. S. Ray, (New York, NY: Academic Press), 473–488. doi: 10.1016/b978-0-12-583580-0.50036-4

Strope, P. K., Skelly, D. A., Kozmin, S. G., Mahadevan, G., Stone, E. A., Magwene, P. M., et al. (2015). The 100-genomes strains, an *S. cerevisiae* resource that illuminates its natural phenotypic and genotypic variation and emergence as an opportunistic pathogen. *Genome Res.* 25, 762–774. doi: 10.1101/gr.185538.114

Taylor, J. H. (1960). Asynchronous duplication of chromosomes in cultured cells of Chinese hamster. *J. Biophys. Biochem. Cytol.* 7, 455–464.

Tsai, H.-J., Baller, J. A., Liachko, I., Koren, A., Burrack, L. S., Hickman, M. A., et al. (2014). Origin replication complex binding, nucleosome depletion patterns, and a primary sequence motif can predict origins of replication in a genome with epigenetic centromeres. *mBio* 5:e01703-14. doi: 10.1128/mBio.01703-14

Watson, J. D., Baker, T. A., Bell, S. P., Gann, A., Levine, M., and Losick, R. (2013). "The replication of DNA," in *Molecular Biology of the Gene*, ed. J. D. Watson, (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), 257–312.

Wilmes, G. M., and Bell, S. P. (2002). The B2 element of the *Saccharomyces cerevisiae* ARS1 origin of replication requires specific sequences to facilitate pre-RC formation. *Proc. Natl. Acad. Sci. U.S.A.* 99, 101–106. doi: 10.1073/pnas.012578499

Wyrick, J. J., Aparicio, J. G., Chen, T., Barnett, J. D., Jennings, E. G., Young, R. A., et al. (2001). Genome-wide distribution of ORC and MCM proteins in *S. cerevisiae*: high-resolution mapping of replication origins. *Science* 294, 2357–2360. doi: 10.1126/science.1066101

Yamazaki, S., Hayano, M., and Masai, H. (2013). Replication timing regulation of eukaryotic replicons: Rif1 as a global regulator of replication timing. *Trends Genet.* 29, 449–460. doi: 10.1016/j.tig.2013.05.001

Yue, J.-X., Li, J., Aigrain, L., Hallin, J., Persson, K., Oliver, K., et al. (2017). Contrasting evolutionary genome dynamics between domesticated and wild yeasts. *Nat. Genet.* 49, 913–924. doi: 10.1038/ng.3847

Zhang, G., and Gao, F. (2017). Quantitative analysis of correlation between AT and GC biases among bacterial genomes. *PLoS One* 12:e0171408. doi: 10.1371/journal.pone.0171408

Zhang, R., and Zhang, C.-T. (2014). A brief review: the z-curve theory and its application in genome analysis. *Curr. Genomics* 15, 78–94. doi: 10.2174/1389202915999140328162433

Zhao, Y., Jia, X., Yang, J., Ling, Y., Zhang, Z., Yu, J., et al. (2014). PanGP: a tool for quickly analyzing bacterial pan-genome profile. *Bioinformatics* 30, 1297–1299. doi: 10.1093/bioinformatics/btu017

Zhu, Y. O., Sherlock, G., and Petrov, D. A. (2016). Whole genome analysis of 132 clinical *Saccharomyces cerevisiae* strains reveals extensive ploidy variation. *G3* 6, 2421–2434. doi: 10.1534/g3.116.029397

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: info@frontiersin.org | +41 21 510 17 00

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