



OPEN ACCESS

EDITED BY

Alan Talevi,
National University of La Plata, Argentina

REVIEWED BY

Antonio Romo-Mancillas,
Autonomous University of Queretaro, Mexico
Daniela Alejandra Gonzalez,
Northeastern University, United States

*CORRESPONDENCE

Edgar López-López,
✉ elopez.lopez@cinvestav.mx,
✉ edgar.lopez.593@hotmail.com

RECEIVED 13 October 2025

REVISED 21 November 2025

ACCEPTED 25 November 2025

PUBLISHED 10 December 2025

CITATION

López-López E, Pardo-Novoa JC,
Barrientos-Salcedo C, Vichi-Ramírez MM and
Cerdeña-García-Rojas CM (2025) A network-
based protocol to prioritize compounds for
biological testing: discovery of Anti-
Staphylococcus aureus cacalol derivatives.
Front. Drug Discov. 5:1724392.
doi: 10.3389/fddsv.2025.1724392

COPYRIGHT

© 2025 López-López, Pardo-Novoa,
Barrientos-Salcedo, Vichi-Ramírez and Cerdeña-
García-Rojas. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other forums is
permitted, provided the original author(s) and
the copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with these
terms.

A network-based protocol to prioritize compounds for biological testing: discovery of Anti-*Staphylococcus aureus* cacalol derivatives

Edgar López-López^{1*}, Julio C. Pardo-Novoa¹,
Carolina Barrientos-Salcedo², Micheel M. Vichi-Ramírez² and
Carlos M. Cerdeña-García-Rojas¹

¹Department of Chemistry and Graduate Program in Pharmacology, Center for Research and Advanced Studies of the National Polytechnic Institute, Mexico City, Mexico, ²Laboratorio de Química Médica y Quimiogenómica, Facultad de Bioanálisis Campus Veracruz, Universidad Veracruzana, Veracruz, Mexico

The integration of computational tools into early drug discovery has facilitated the rational prioritization of bioactive compounds. In this study, we present a general semi-automated, network-based protocol designed to guide the biological testing of *in-house* compounds by leveraging chemical similarity and reported bioactivity data. As a proof of concept, we constructed a curated database of 127,134 compounds with documented activity against *Staphylococcus aureus* strains. We used this resource to map a biologically relevant chemical space that includes cacalol derivatives, which allowed the identification of two *in-house* cacalol analogues with predicted anti-*S. aureus* activity. These compounds were selected based on pairwise similarity analysis using the ECFP4 fingerprint and the Tanimoto coefficient as the molecular descriptor and similarity metric, respectively. The selected compounds were subsequently validated *in vitro* using disk diffusion and resazurin-based microdilution assays, confirming consistent antibacterial activity (20 mM; 7–10 mm of inhibition zones and 1–2 fold reduction in metabolism activity, respectively). This study aims to bridge the gap between *in silico* and wet-lab approaches, enhancing rapid and intelligent screening to prioritize the biological evaluation of *in-house* compounds.

KEYWORDS

antibacterial activity, chemoinformatics, computer-aided drug design, network-based chemical space, virtual screening, structure-property associations

1 Introduction

Computer-aided drug design (CADD) has become a widely adopted strategy for the rational development of compounds with improved efficiency, safety, and cost-effectiveness. Nowadays, many *in silico* models have emerged as adaptable tools to support the semi-automation of drug design and development campaigns (Talevi and Bellera, 2024; Medina-Franco and López-López, 2024). Chemoinformatics methodologies have proven to be a valuable option for integrating multidisciplinary data to address common research questions, thereby helping to bridge knowledge gaps between different

yet related fields, such as chemistry, biology, and biomedical sciences (e.g., pharmacology and natural products chemistry) (Almeida et al., 2025; Szwarc et al., 2025).

In this regard, network-based drug design is a well-established strategy that leverages available data by using structural similarity to decode and establish complex associations. For example, the prediction of compound bioactivity is based on their structural similarity to previously reported active or inactive compounds (Boezio et al., 2017; Yu et al., 2023). These network-based approaches allow researchers to navigate large chemical spaces and integrate biological data, phenotypic profiles, or disease targets to prioritize candidate molecules with greater biological relevance (Barabási et al., 2011; Lee et al., 2023). For example, chemical space networks are constructed based on structural similarity relationships, where compounds were connected according to their pairwise similarity values, enabling the visualization of clusters of structurally related molecules and the identification of representative analogues with potential biological activity (Vigil-Vásquez and Schüller, 2022).

Staphylococcus aureus, a major cause of hospital- and community-acquired infections, poses a significant global health threat due to the emergence of multidrug-resistant strains (Steinig et al., 2022). Moreover, *S. aureus* infections represent the leading cause of death from bacteremia worldwide, with a case fatality rate ranging from 15% to 30% and an estimated 300,000 deaths annually (Tong et al., 2025). This makes the identification of novel antimicrobial agents with new mechanisms of action an urgent priority.

Natural products, and particularly those derived from plants, have long served as a rich source of structurally diverse and biologically active molecules (Gómez-García et al., 2024; Chandrasekhar et al., 2025). Among these, cacalol, a sesquiterpene alcohol isolated from the Mexican medicinal plants of the genus *Psacalium*, has shown promising antimicrobial properties with reported activities ranging from 0.2–0.7 mg/mL (Robles-Zepeda et al., 2011; García-Pal et al., 2016; Garduño-Ramírez et al., 2001; Soto Hernández, 2002). Additionally, cacalol and its derivatives have been associated with other biological activities, including anticancer, anti-inflammatory, and antimigratory effects, making it an attractive candidate for future medicinal chemistry optimization processes (Liu et al., 2010; Mora-Ramiro et al., 2020; Rostro-Alonso et al., 2024).

In this study, we develop a semi-automated network-based chemoinformatic approach that integrates chemical similarity, network pharmacology, and chemical space analysis to prioritize compounds with potential biological activity. The main objective of this work is to develop a tool for the construction and analysis of network-based chemical spaces, designed to be easily implemented in laboratories with limited expertise in computational methods. We hypothesize that this integrative strategy can effectively guide the identification of bioactive molecules. Our approach was capable of identifying cacalol derivatives with anti-*S. aureus* potential by exploring their connectivity within a biologically informed chemical network. Beyond cacalol, this framework can be readily applied to other scaffolds or compound libraries, providing a scalable and data-driven strategy to accelerate early-stage drug discovery.

2 Materials and methods

This section presents the use of a semi-automated protocol for the generation of network-based chemical spaces, which enabled the identification and prioritization of four *in-house* compounds, i.e., those routinely available in our laboratory. These compounds were subsequently evaluated experimentally against *S. aureus* using locally available biological models. The *in-house* compounds were included in the study due to their accessibility and immediate experimental feasibility, which allowed us to directly link the computational results with biological validation within the context of the present study.

2.1 Data obtaining and pre-processing

From the ChEMBL V. 33 database, 127,128 compounds with reported minimum inhibitory concentration (MIC) activity against *S. aureus* (Target ID: ChEMBL352) were retrieved (Zdrzil et al., 2023). The MIC values were standardized to $\mu\text{g}/\text{mL}$ units. Subsequently, the SMILES codes of all compounds were used to identify and remove duplicate structures, yielding a final dataset of 127,060 unique compounds. Finally, four *in-house* compounds were added, resulting in a total dataset of 127,064 compounds.

2.2 Network-based chemical space

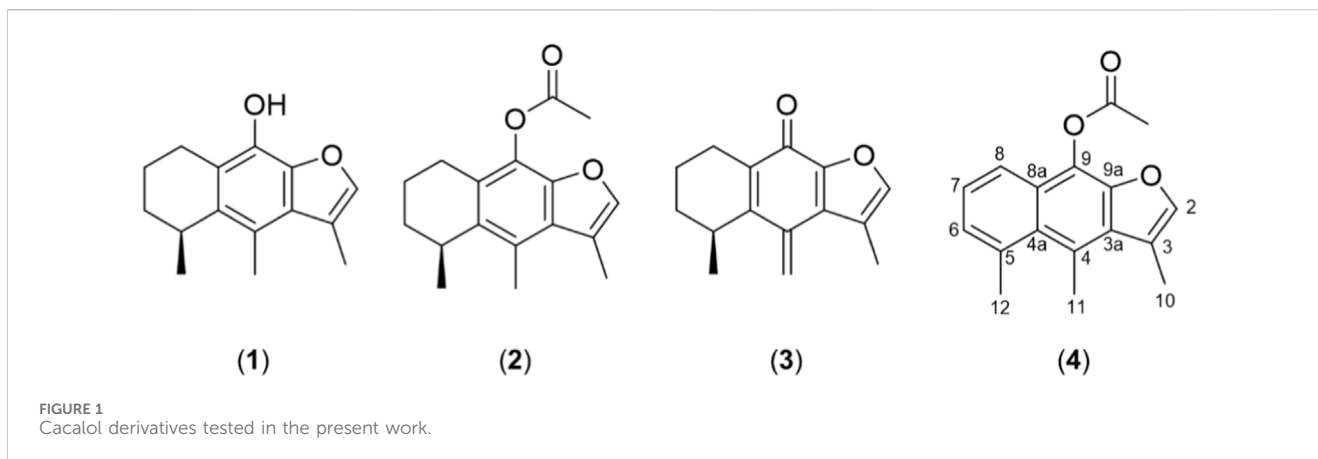
An activity-based network allows the visual representation of many chemical compounds through the paired distance calculation between each pair of compounds in a dataset. In this protocol, compounds with anti-*S. aureus* activity and *in-house* compounds were used to construct it.

2.2.1 Distance-paired calculations

The automation of this step was implemented in KNIME software, using different nodes, including KNIME base nodes, KNIME base chemistry nodes, RDKit nodes, and KNIME distance matrix extension (see in the supplementary material) (KNIME Base nodes, 2025; KNIME Base Chemistry Types and Nodes, 2025; KNIME Distance Matrix, 2025; KNIME RDKit node, 2025). ECFP4 fingerprint (1,024 bits) was calculated for each compound, and the Tanimoto distance was calculated for each pair of compounds (Dunn et al., 2023). The 10 most similar compounds in each case were considered to construct the final chemical space visualization. [Supplementary Figure S1](#) in the supplementary material section illustrates the workflow implemented in KNIME software.

2.2.2 Chemical space visualization

A dimensional reduction method (t-SNE) was used to construct a chemical space visualization, using classical druglikeness descriptors (i.e., total molecular weight, cLogP, number of H acceptors, number of H donors, polar surface area, and number of rotatable bonds) (Lipinski et al., 1997; Veber et al., 2002; Ghose et al., 1999) calculated in DataWarrior software V. 06.01 (Sander et al., 2015). The parameters used in the dimensional reduction



method (*i.e.*, perplexity = 20, source dimensions = 6, and number of interactions = 1,000) were maintained as default values.

2.2.3 Interactive visualization

The activity-based network was constructed using Datawarrior software (Sander et al., 2015; López-López et al., 2019). For this, the t-SNE 1 and t-SNE 2 coordinates were graphed as scatter plots. Each compound (node) was connected according to the 10 most similar compounds previously calculated. The activity values were represented using a color scale from green (most active compound; 0.01 mg/mL to red (most inactive compound; 100,000 mg/mL). Finally, the *in-house* compounds were manually analyzed, while compound pairs exhibiting similarity values greater than 0.5 Tanimoto units were examined following well-established practices in cheminformatics (Maggiore et al., 2014; Bajusz et al., 2015). To guarantee the reproducibility of this protocol, the automated KNIME workflow and the interactive activity-based network generated in DataWarrior software are available in the supplementary material section.

2.3 Chemical experimental procedures

NMR spectra were measured at 300 MHz for ^1H and 75.4 Hz for ^{13}C on a Varian Mercury 300 spectrometer or at 500 MHz for ^1H and 125 MHz for ^{13}C on a Jeol[®] ECA 500 spectrometer from CDCl_3 solutions containing tetramethylsilane as the internal reference. Chemical shift values are reported in ppm and coupling constants (*J*) are in Hz. HRESIMS spectra were measured on an Orbitrap Exploris 120 mass spectrometer. Optical rotations were recorded in CHCl_3 solutions on a PerkinElmer 341 polarimeter. IR spectra were obtained on a BioTools dualPEM ChiralIR FT spectrophotometer in CHCl_3 solutions. UV spectra were measured in EtOH on a PerkinElmer Lambda 12 spectrophotometer. Silica gel 230–400 mesh (Merck) and aluminium oxide 90 active neutral 70–230 mesh (Merck) were used for column chromatography.

2.3.1 Cacalol derivatives

The chemical conditions to obtain and purify compounds 1–3 (Figure 1) were reproduced by previous protocols (López-López et al., 2025b). Their ^1H and ^{13}C NMR data agreed with the reported spectrum data (Liu et al., 2010; del Río et al., 2021; López-López

et al., 2025b). The spectrometric data are available in Supplementary Figures S2–S7 in the supplementary material section.

2.3.2 Compound 4: 9-*O*-acetyloxy-3,4,5-trimethylnaphtho[2,3-*b*]furan

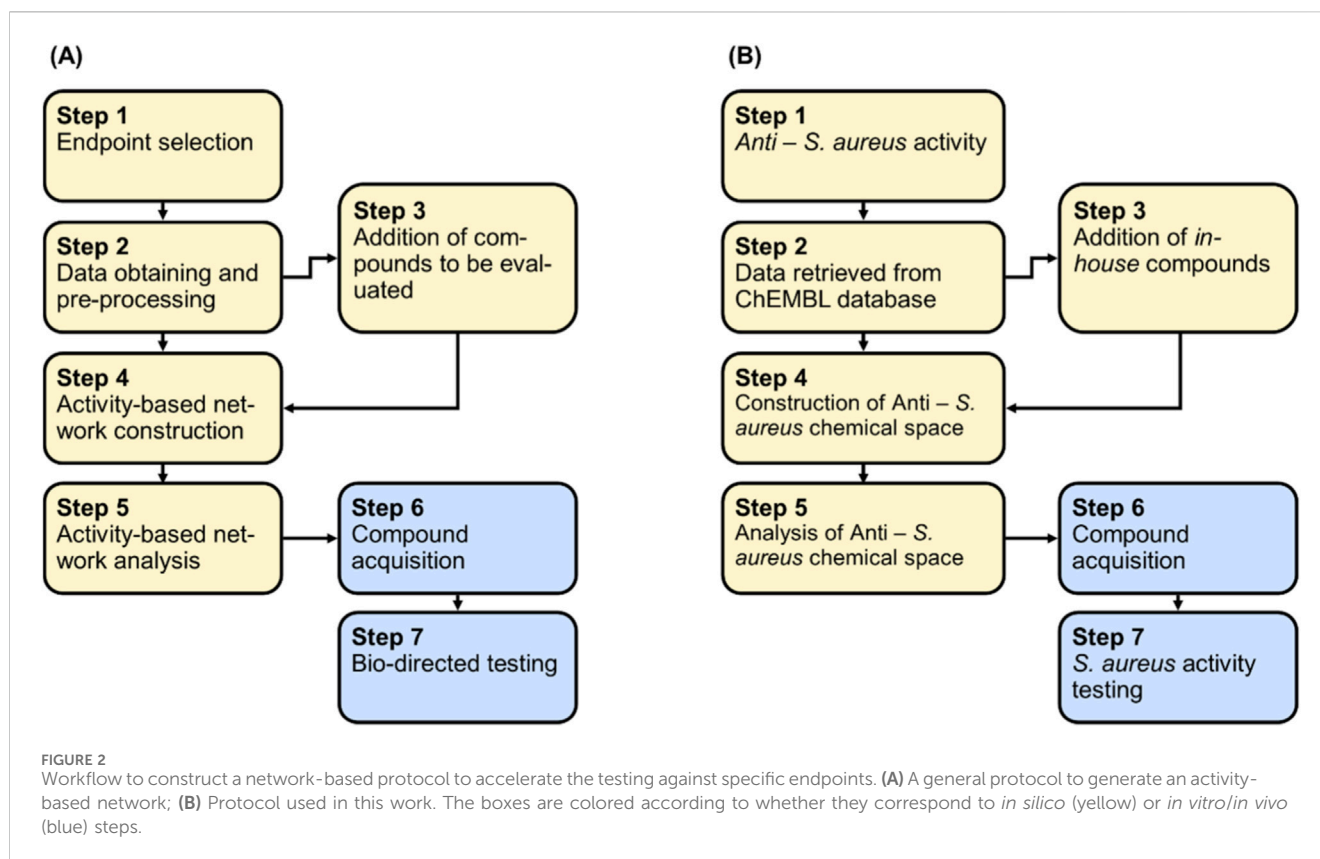
A sample of cacalol acetate (2) (200 mg) in *o*-xylene (15 mL) was treated with *p*-chloranil (700 mg). The mixture was refluxed for 20 h, poured over ice– H_2O , and evaporated to dryness. The residue was purified by column chromatography on alumina using mixtures of hexanes–EtOAc as the eluent, affording 9-*O*-acetyloxy-3,4,5-trimethylnaphtho[2,3-*b*]furan (4) (8 mg, 4%) in fractions 12–17 (hexanes–EtOAc, 99:1). White solid; mp 154 °C–156 °C; lit. mp 154 °C–155 °C. IR, UV, and ^1H NMR data were in agreement with literature values (Romo and Joseph-Nathan, 1964; Joseph-Nathan et al., 1966); ^{13}C NMR δ 168.7 (CO, Ac), 144.1 (CH, C-2), 143.9 (C, C-9a), 135.7 (C, C-5), 130.7 (C, C-4a), 129.9 (C, C-3a), 128.2 (CH, C-6), 127.1 (2C, C-4 or C-9), 126.4 (C, C-8a), 124.7 (CH, C-7), 119.3 (CH, C-8), 116.8 (C, C-3), 26.8 (CH_3 , C-11), 20.7 (CH_3 , Ac), 19.7 (CH_3 , C-12), 12.3 (CH_3 , C-10). The spectrometric data are available in Supplementary Figures S8–S13 in the supplementary material section.

2.4 Disk diffusion testing

The antibacterial susceptibility of *S. aureus* (ATCC 29213) was evaluated using the disk diffusion method following the guidelines outlined in the CLSI document M02-A14 (M02, 2025). Bacterial suspensions were prepared from broth cultures and adjusted to a 0.5 McFarland standard ($\approx 1.5 \times 10^8$ CFU/mL). Mueller-Hinton agar plates (90 mm in diameter) were uniformly inoculated with the suspension to create a bacterial lawn. Subsequently, disks containing the tested compounds (0.5 mM, 1 mM, or 20 mM) were placed on the agar surface and incubated at 37 °C for 16 h. After that, inhibition zones were measured in millimeters, and results were interpreted based on CLSI M100-S35 performance standards (M100, 2025; Skov et al., 2006).

2.5 Resazurin-based microdilution assay

A 0.01% (w/v) resazurin solution was prepared by dissolving resazurin sodium salt in sterile phosphate-buffered saline (PBS) at



pH 7.2. The solution was sterilized by filtration through a 0.22 μm membrane filter and stored at 4 $^{\circ}\text{C}$, protected from light, for up to 3 days to maintain stability and prevent degradation. As a sensitivity control, penicillin 10 U/mL (6.027 $\mu\text{g/mL}$) solutions were prepared following CLSI guidelines (Elshikh et al., 2016). Bacterial inoculum was prepared by suspending colonies from overnight cultures in sterile saline solution and adjusted to match the turbidity of a 0.5 McFarland standard ($\approx 1.5 \times 10^8$ CFU/mL).

In a sterile 96-well microtiter plate, 100 μL of the standardized bacterial suspension was dispensed into each well. For positive control wells, 50 μL of the control antibiotic (≤ 2.0 $\mu\text{g/mL}$ oxacillin) or tested compounds were added. For negative control wells, a volume of DMSO was added, while for the experimental wells, 50 μL of each test compound preparation in DMSO was added to 0.5 mM. All assays were performed in triplicate. The plates were incubated at 37 $^{\circ}\text{C}$ for 60 min. Following incubation, 50 μL of the sterile 0.01% resazurin solution was added to each well, and the plates were reincubated for 4 h at 37 $^{\circ}\text{C}$. Optical density readings at 560 nm were obtained using a multiscan spectrophotometer (ThermoScientific). Background absorbance was corrected by subtracting the readings from wells containing only resazurin and a standardized bacterial suspension (Bobenchik et al., 2015).

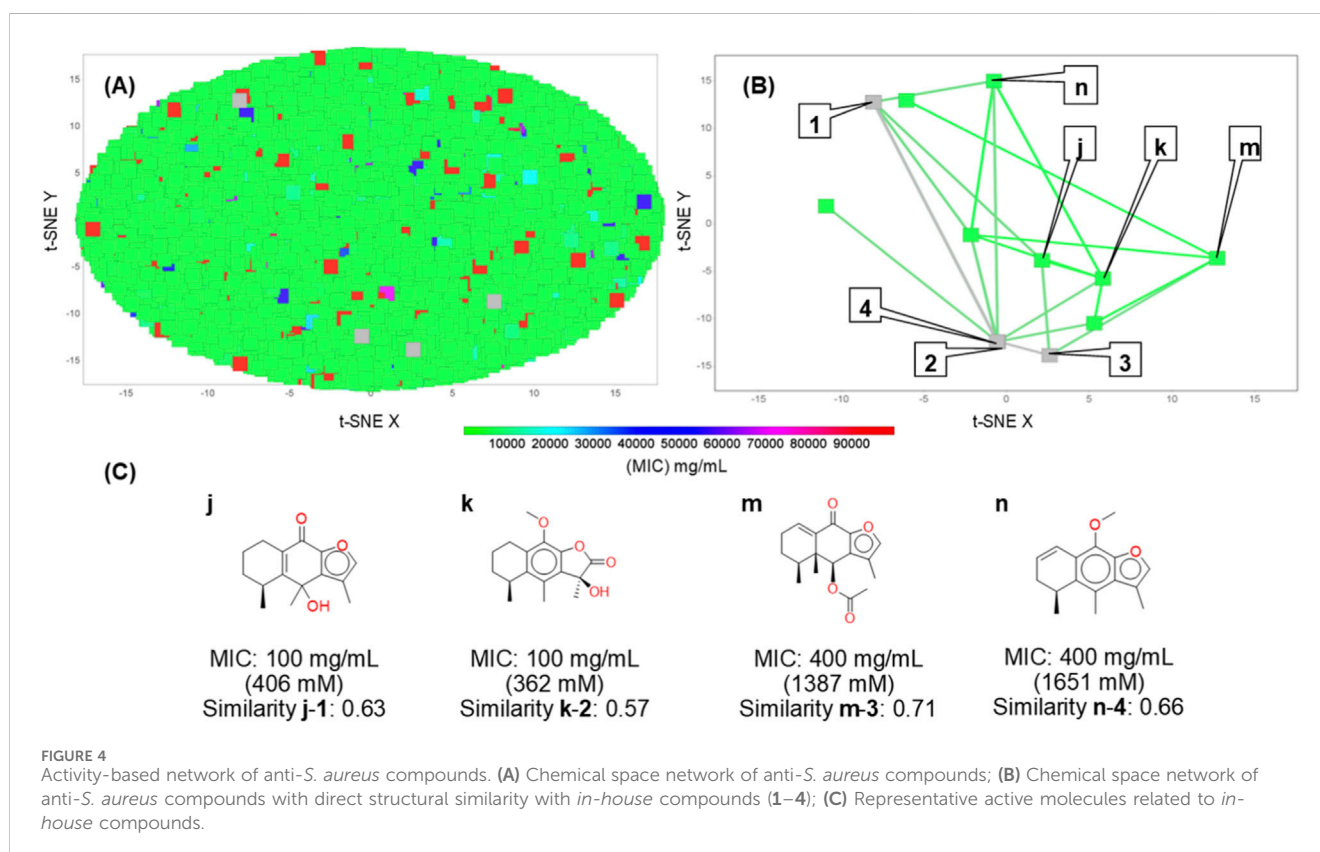
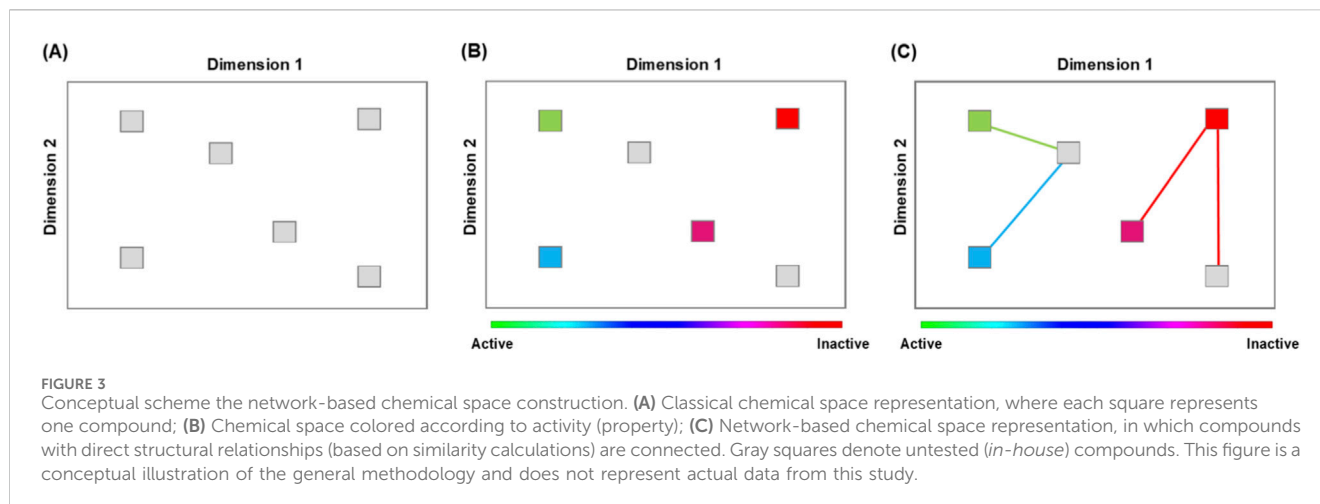
3 Results

In this work, a general network-based protocol was developed (Figure 2A) and validated through *in vitro* assays to demonstrate its

utility as a rapid and convenient strategy for prioritizing compound evaluations against defined biological endpoints. Subsequently, this protocol was specifically applied to construct a network-based chemical space of compounds with reported activity against *S. aureus* (Figure 2B), which guided the biological testing of *in-house* compounds (1–4).

Figure 2 illustrates the general protocol for constructing network-based chemical spaces, divided into seven steps: (1) Endpoint selection: choosing a relevant *in vitro* or *in vivo* model guided by prior experimental evidence, literature, or model availability; (2) Data acquisition: collecting information from public or proprietary databases containing biological data (e.g., activity against specific targets, cell line assays, side effects, *in vivo* outcomes, etc.); (3) Compound addition: incorporating compounds to be tested; (4) Network-based chemical space construction: generating a network from molecular descriptors and similarity calculations, where nodes represent compounds and edges connect structurally related neighbors; (5) Network analysis: identifying clusters, key nodes, and structure–property associations, focusing on compounds related to active ones; (6) Compound acquisition: selecting candidate molecules with structurally related neighbors exhibiting the desired property or activity; and (7) Experimental validation: testing the selected compounds to confirm the associations predicted by the network-based approach, thereby linking computational prediction to experimental confirmation.

Figure 3 illustrates the process of constructing a network-based chemical space. Starting from a traditional chemical space, an additional dimension is projected through the use of color coding



(Figures 3A,B), leveraging known property information for each compound in the initial dataset while also highlighting *in-house* compounds lacking prior evaluation of that property. The protocol further enables the identification of close structural neighbors (similar compounds; Figure 3C) within the dataset, capturing associations both between compounds with known values of the property of interest and those without prior experimental information, thus facilitating the prioritization of untested molecules with high structural resemblance to tested (active) ones.

3.1 Network-based chemical space

Figure 4 illustrates the network-based chemical space constructed in this work, which enables the identification of neighbors around the *in-house* compounds (Figures 4A,B). In this network, *in-house* compounds (1–4) exhibited similarity values above 0.57 with respect to confirmed active compounds (j–n). The active compounds display MIC values ranging from 100 to 400 mg/mL (406 to 1,651 mM; Figure 4C), suggesting that structurally related compounds may exhibit a similar activity range.

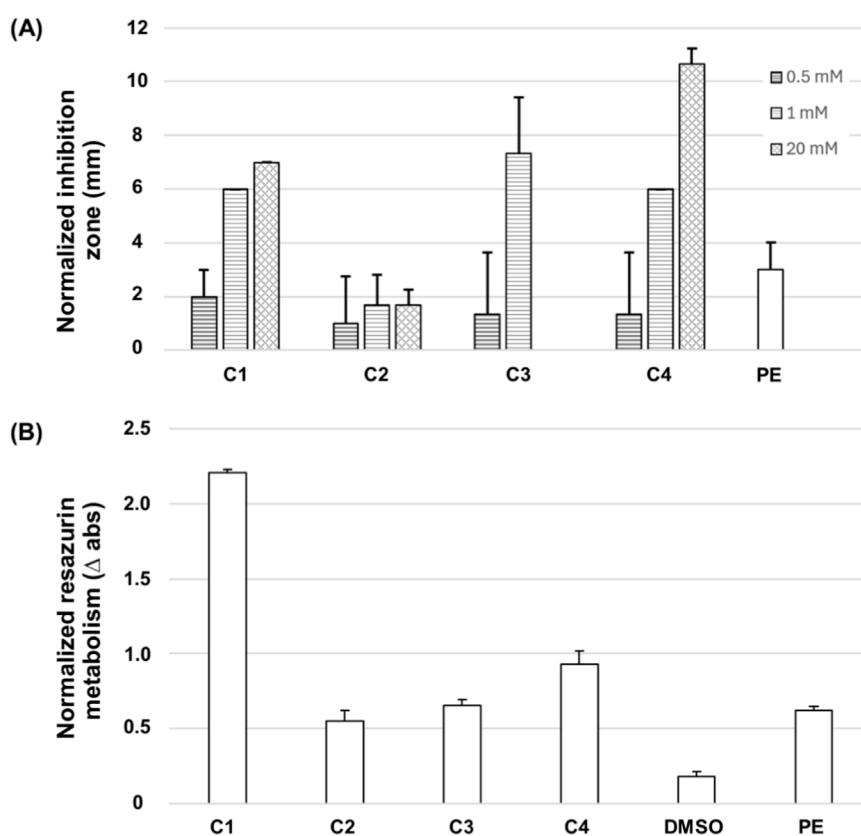


FIGURE 5
In vitro S. aureus testing. **(A)** Normalized inhibition zone of cacalol derivatives. The vehicle (DMSO, 4%) did not exhibit a halo of inhibition (0 mm; not shown). The data were normalized using the size of the inhibition disc. Evaluation of compound **3** at 20 mM was not possible due to low solubility. **(B)** Normalized resazurin metabolism activity of cacalol derivatives. Each compound was evaluated at a concentration of 0.5 mM; PE = Penicillin (6.027 μ g/mL; 0.01692 mM). Each experimental condition was analyzed in triplicate.

3.2 Anti-*S. aureus* activity

The antibacterial susceptibility of *S. aureus* (ATCC 29213) to cacalol derivatives (**1–4**) was assessed through two complementary methods: the disk diffusion assay and the resazurin-based microdilution assay (Figures 5A,B). The disk diffusion method offered a qualitative evaluation of antibacterial activity by measuring the ability of each compound to inhibit bacterial growth on solid media; larger inhibition zones correspond to stronger antibacterial effects (Saar et al., 2025). Conversely, the resazurin-based microdilution assay provided a quantitative measure of inhibitory concentration, where higher absorbance (inhibition of resazurin metabolism) indicates higher antibacterial potency (Foerster et al., 2017). Notably, compounds **1** and **4** demonstrated modest anti-*S. aureus* activity across both assays, which is consistent with the structure-activity associations suggested by our network-based approaches.

4 Discussion

Network-based chemical space is an emergent approach used to accelerate the discovery and development of new drug candidates, which allows for simplifying multidimensional variables in two or

three dimensions, improving the explainability and analysis of complex data (Talevi, A., 2024; Medina-Franco et al., 2021). Nowadays, network-based chemical space representations are among the most widely used chemical space visualizations, offering multiple utilities in molecular design and discovery. For example, regarding the decoding of structure-property (activity) and polypharmacological relationships data (Maggiore and Bajorath, 2014; Wen et al., 2024; Kumari and Subbarao, 2023; Medina-Franco et al., 2022), on the prediction of ADMET properties, side and off-target effects (Yi et al., 2023; Shabani-Mashcool et al., 2020), on the identification of putative targets (You et al., 2022), and to prioritize the selection of potentially active compounds through the use of virtual screening campaigns (Goel et al., 2022; Wu et al., 2022).

However, there are some issues to resolve before exploiting the use of this kind of approach in drug design areas, which include (1) the creation of more efficient protocols, to reduce the high computational cost which limits their exponential implementation for large chemical libraries; (2) the semi-automatization of this kind of protocols; and (3) the creation of protocols that do not require extensive programming knowledge, which would accelerate its adoption in medicinal chemistry, pharmacology, and biology laboratories (Bajorath et al., 2022). For this reason, in this work, we propose a general protocol

implemented in KNIME, which is a free and easy-to-use software that could be easily adapted for the hardware possibilities of each kind of laboratory, which could democratize the use of this novel strategy in underdeveloped countries, and not only in specialized *in silico* laboratories.

Our results demonstrate the utility of the semi-automated network-based chemical space protocol for the efficient discovery of bioactive molecules, through the establishment of structure-activity associations between active, inactive, and not-tested compounds (*in-house*). In this regard, the derivatives of cacalol were associated with an anti-*S. aureus* activity, which has been experimentally validated. Interestingly, previous studies have confirmed the anti-*S. aureus* activity of the plant extract, which contains cacalol derivatives (Robles-Zepeda et al., 2011; Garcia-Pal et al., 2016), opening new perspectives for the use of the presented protocol to decode the activity of natural product extracts, which contain a vast list of compounds. This is particularly useful when their activity cannot be associated directly with experimental data. Additionally, other possible applications of the network-based protocols are the use of multiparametric data (*i.e.*, chemical, biological, toxicological, environmental, or clinical data) to establish multiple-structure activity associations of novel compounds, reducing the possibility of solubility issues, side effects, or bioavailability limitations (Medina-Franco et al., 2024; López-López et al., 2022). For example, our case study highlights that other factors influencing compound behavior must still be taken into account by the users. Indeed, it is common in virtual screening protocols for compounds to appear inactive due to physicochemical limitations—such as poor solubility—rather than true lack of activity (Chand Dakal et al., 2025). For instance, compound 3 exhibited solubility issues during disk diffusion assays, which limit the complete characterization of its activity using aqueous *in vitro* models. This is a well-documented limitation in virtual screening workflows, where compounds that fail physical property filters can generate misleading outcomes. Therefore, we recommend adapting this methodology by incorporating classical virtual screening approaches focused on ADMET-related criteria before proceeding to *in vitro* evaluations (Corrêa Veríssimo et al., 2024).

In addition, it is important to acknowledge that the evaluated cacalol derivatives (1–4) exhibited only modest antibacterial activity, falling within the high-micromolar to low-millimolar range (0.5–20 mM). This limited potency is consistent with the activity of the most similar known active compounds in the dataset: the reference molecules j–n, which share similarity values above 0.57 with the *in-house* compounds, display MIC values between 100 and 400 mg/mL (406–1,651 mM; Figure 5C). These observations suggest that the chemical neighborhood explored here is intrinsically associated with weak antibacterial activity—an aspect that likely reflects intrinsic limitations of the underlying scaffold rather than shortcomings of the computational workflow itself. From a conventional hit-discovery perspective, compounds with low-micromolar activity are generally preferred as starting points for optimization, whereas millimolar-range activity is typically considered too weak to constitute a viable hit (Zhu et al., 2013). Accordingly, the cacalol derivatives identified here should be regarded as preliminary hits that offer structural insight rather than as fully validated antibacterial leads.

Nevertheless, the primary goal of this study was not to deliver optimized hit-level compounds but to evaluate the ability of the network-based protocol to prioritize chemical candidates, reveal structure–activity associations, and support early decision-making. Even weakly active compounds can be mechanistically informative when they form coherent clusters, align with independent ethnopharmacological evidence, or highlight conserved chemical features. Looking forward, several strategies could further enhance the likelihood of identifying more potent candidates, including expanding the chemical space to incorporate additional bioactivity annotations, integrating physicochemical or ADMET-guided filters, applying alternative similarity metrics or network embeddings, and exploring scaffold hopping to reach adjacent, more bioactive chemical neighborhoods. Such refinements, together with targeted experimental validation, will be essential for advancing the most promising candidates identified through the proposed workflow.

Finally, although this work was primarily focused on exploring structure–activity associations (López-López et al., 2025a), the concept developed here can be extended to the study of other properties not related to the biological or pharmacological domains. In particular, the same framework can be adapted for the decoding of properties in diverse fields such as food chemistry (Foodinformatics, 2014), materials science (Yosipof et al., 2016), and polymer chemistry (Bărbulescu and Barbeș, 2025), among others, thereby broadening its applicability beyond the scope of the present study. Finally, this study has some limitations. The dimensionality reduction approach may oversimplify structural relationships, potentially leading to an overestimation of the predicted activity of *in-house* compounds (Jia et al., 2022; Orlov et al., 2025). For this reason, we suggest applying the proposed methodology in combination with classical virtual screening approaches. Nonetheless, the network-based framework proved effective as a rapid and informative strategy to prioritize compounds and guide future optimization efforts.

5 Conclusion

Network-based chemical space analysis offers a powerful approach for decoding structure–activity associations (or relationships) and prioritizing candidate molecules in large and diverse chemical libraries. Here, we implemented a semi-automated protocol that integrates chemical similarity and chemical space concepts to guide the biological evaluation of *in-house* compounds. Using a curated database of compounds with documented activity against *S. aureus*, we identified cacalol derivatives as promising candidates and experimentally validated analogues with consistent antibacterial activity.

This proof of concept illustrates how network-based workflows can bridge computational associations and experimental validation, accelerating early-stage drug discovery while lowering technical and computational barriers. The methodology can be extended to multiparametric datasets that integrate chemical, biological, ADMET-related criteria, clinical, or environmental data, improving the selection of compounds with balanced activity and favorable properties. By making such workflows accessible to laboratories with varying resource levels, this approach

contributes to the democratization of advanced chemoinformatics tools and supports the discovery of new therapeutic candidates against pressing health threats.

Data availability statement

The KNIME workflows and datasets associated with this study are available at: <https://doi.org/10.5281/zenodo.16814759>.

Author contributions

EL-L: Funding acquisition, Resources, Visualization, Writing – original draft, Formal Analysis, Validation, Data curation, Project administration, Investigation, Conceptualization, Supervision, Writing – review and editing, Methodology, Software. JP-N: Methodology, Writing – review and editing. CB-S: Writing – review and editing, Project administration, Methodology, Resources. MV-R: Writing – review and editing, Methodology. CC-G-R: Methodology, Writing – review and editing, Resources.

Funding

The authors declare that no financial support was received for the research and/or publication of this article.

Acknowledgements

E.L.-L. is grateful to Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI), Mexico, for the Ph.D. scholarship number 762342 (No. CVU: 894234). We are thankful to Q. F. B. Verónica Reyes Olivares and Q. F. B. Elvia Celina Álvarez

References

- Almeida, R. L., Campera, G. M., Pöhner, I., and Maltarollo, V. G. (2025). Artificial intelligence in drug design: why a ‘one-size-fits-all’ approach remains out of reach. *Exp. Op. Drug Discov.* 20 (10), 1239–1250. doi:10.1080/17460441.2025.2543802
- Bajorath, J., Chávez-Hernández, A. L., Duran-Frigola, M., Fernández-de Gortari, E., Gasteiger, J., López-López, E., et al. (2022). Chemoinformatics and artificial intelligence colloquium: progress and challenges in developing bioactive compounds. *J. Cheminformatics* 14 (1), 82. doi:10.1186/s13321-022-00661-0
- Bajusz, D., Rácz, A., and Héberger, K. (2015). Why is tanimoto index an appropriate choice for fingerprint-based similarity calculations? *J. Cheminform.* 7 (1), 20. doi:10.1186/s13321-015-0069-3
- Barabási, A. L., Gulbahce, N., and Loscalzo, J. (2011). Network medicine: a network-based approach to human disease. *Nat. Rev. Genet.* 12, 56–68. doi:10.1038/nrg2918
- Bărbulescu, A., and Barbeș, L. (2025). Cheminformatics approaches to the analysis of additives for sustainable polymeric materials. *Polymers* 17 (11), 1522. doi:10.3390/polym17111522
- Bobenchik, A. M., Deak, E., Hindler, J. A., Charlton, C. L., and Humphries, R. M. (2015). Performance of Vitek 2 for Antimicrobial Susceptibility testing of enterobacteriaceae with vitek 2 (2009 FDA) and 2014 CLSI breakpoints. *J. Clin. Microbiol.* 53 (3), 816–823. doi:10.1128/jcm.02697-14
- Boezio, B., Audouze, K., Ducrot, P., and Taboureau, O. (2017). Network-based approaches in pharmacology. *Mol. Inf.* 36 (10), 1700048. doi:10.1002/minf.201700048
- Chand Dakal, T., Joseph, J. K., Maurya, P. K., Sharma, N. K., and Rajagopal, S. (2025). Comprehensive report on biochemical, pharmacological, and pharmacokinetic

Cisneros, Department of Chemistry, Cinvestav-IPN, for specialized instrumental determinations and NMR measurements, respectively.

Conflict of interest

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

Generative AI statement

The authors declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fddsv.2025.1724392/full#supplementary-material>

properties of tool compounds relevant to human pathologies. *RSC Pharm.* 2, 480–489. doi:10.1039/d4pm00331d

Chandrasekhar, V., Rajan, K., Kanakam, S. R. S., Sharma, N., Weissenborn, V., Schaub, J., et al. (2025). COCONUT 2.0: a comprehensive overhaul and curation of the collection of open natural products database. *Nucleic Acids Res.* 53 (D1), D634–D643. doi:10.1093/nar/gkae1063

Corrêa Verissimo, G., Salgado Ferreira, R., and Gonçalves Maltarollo, V. (2024). Ultra-large virtual screening: definition, recent advances, and challenges in drug design. *Mol. Inf.* 44 (1), e202400305. doi:10.1002/minf.202400305

del Río, R. E., Pardo-Novoa, J. C., Cerda-García-Rojas, C. M., and Joseph-Nathan, P. (2021). Vibrational circular dichroism behavior of quinaldine derivatives from Psacalium aff. sinuatum. *J. Mol. Struct.* 1224 (2020), 128987. doi:10.1016/j.molstruc.2020.128987

Dunn, T. B., López-López, E., Kim, T. D., Medina-Franco, J. L., and Miranda-Quintana, R. A. (2023). Exploring activity landscapes with extended similarity: is Tanimoto enough? *Mol. Inf.* 42 (7), 2300056. doi:10.1002/minf.202300056

Elsikh, M., Ahmed, S., Funston, S., Dunlop, P., McGaw, M., Marchant, R., et al. (2016). Resazurin-based 96-well plate microdilution method for the determination of minimum inhibitory concentration of biosurfactants. *Biotech. Lett.* 38 (6), 1015–1019. doi:10.1007/s10529-016-2079-2

Foerster, S., Desilvestro, V., Hathaway, L. J., Althaus, C. L., and Unemo, M. (2017). A new rapid resazurin-based microdilution assay for antimicrobial susceptibility testing of *Neisseria gonorrhoeae*. *J. Antimicrob. Chemother.* 72 (7), 1961–1968. doi:10.1093/jac/dkx113

- Foodinformatics (2014). Cham: Springer International Publishing. doi:10.1007/978-3-319-10226-9
- García-Pal, P., Garza-Ramo, M. A. de la, Galindo-Ro, S. A., Oranday-Ca, A., Leos-Rivas, C., Verde-Star, M. J., et al. (2016). Antimicrobial effect of the methanolic extract *Pscacalium decompositum* on periodontopathogenic bacteria. *Intern. J. Pharmacol.* 12 (7), 737–742. doi:10.3923/ijp.2016.737.742
- Garduño-Ramírez, M. L., Trejo, A., Navarro, V., Bye, R., Linares, E., and Delgado, G. (2001). New modified eremophilanes from the roots of *Pscacalium radulifolium*. *J. Nat. Prod.* 64 (4), 432–435. doi:10.1021/np000385z
- Ghose, A. K., Viswanadhan, V. N., and Wendoloski, J. J. (1999). A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. I. A qualitative and quantitative characterization of known drug databases. *J. Comb. Chem.* 1 (1), 55–68. doi:10.1021/cc9800071
- Goel, M., Aggarwal, R., Sridharan, B., Pal, P. K., and Priyakumar, U. D. (2022). Efficient and enhanced sampling of drug-like chemical space for virtual screening and molecular design using modern machine learning methods. *WIREs Comput. Mol. Sci.* 13 (2), e1637. doi:10.1002/wcms.1637
- Gómez-García, A., Acuña Jiménez, D. A., Zamora, W. J., Barazorda-Cahuana, H. L., Chávez-Fumagalli, M. Á., Valli, M., et al. (2024). Latin American natural product database (LANaPDB): an update. *J. Chem. Inf. Model.* 64 (22), 8495–8509. doi:10.1021/acs.jcim.4c01560
- Jia, W., Sun, M., Lian, J., and Hou, S. (2022). Feature dimensionality reduction: a review. *Complex & Intel. Syst.* 8 (3), 2663–2693. doi:10.1007/s40747-021-00637-x
- Joseph-Nathan, P., Morales, J. J., and Romo, J. (1966). Contribution to the chemistry of cacalol. *Tetrahedron* 22 (1), 301–307. doi:10.1016/0040-4020(66)80131-x
- KNIME Base Chemistry Types & Nodes (2025). KNIME community hub. Available online at: <https://hub.knime.com/knime/extensions/org.knime.features.chem.types/latest/> (Accessed September 5, 2025).
- KNIME Base nodes (2025). KNIME community hub. Available online at: <https://hub.knime.com/knime/extensions/org.knime.features.base/latest/> (Accessed September 5, 2025).
- KNIME Distance Matrix (2025). KNIME community hub. Available online at: <https://hub.knime.com/knime/extensions/org.knime.features.distmatrix/latest/> (Accessed September 5, 2025).
- KNIME RDKit node (2025). KNIME community hub. Available online at: <https://www.knime.com/rdkit> (Accessed September 5, 2025).
- Kumari, M., and Subbarao, N. (2023). Convolutional neural network-based quantitative structure–activity relationship and fingerprint analysis against inhibitors of anthrax lethal factor. *Fut. Med. Chem.* 15 (10), 853–866. doi:10.4155/fmc-2023-0093
- Lee, C. W., Kim, S. M., Sa, S., Hong, M., Nam, S. M., and Han, H. W. (2023). Relationship between drug targets and drug-signature networks: a network-based genome-wide landscape. *BMC Med. Gen.* 16 (1), 17. doi:10.1186/s12920-023-01444-8
- Lipinski, C. A., Lombardo, F., Dominy, B. W., and Feeney, P. J. (1997). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* 23 (1–3), 3–25. doi:10.1016/S0169-409X(96)00423-1
- Liu, W., Furuta, E., Shindo, K., Watabe, M., Xing, F., Pandey, P. R., et al. (2010). Cacalol, a natural sesquiterpene, induces apoptosis in breast cancer cells by modulating Akt-SREBP-FAS signaling pathway. *Breast Cancer Res. Treat.* 128 (1), 57–68. doi:10.1007/s10549-010-1076-8
- López-López, E., Naveja, J. J., and Medina-Franco, J. L. (2019). DataWarrior: an evaluation of the open-source drug discovery tool. *Exp. Op. Drug Discov.* 14 (4), 335–341. doi:10.1080/17460441.2019.1581170
- López-López, E., Fernández-de Gortari, E., and Medina-Franco, J. L. (2022). Yes SIR! on the structure–inactivity relationships in drug discovery. *Drug Discov. Today* 27 (8), 2353–2362. doi:10.1016/j.drudis.2022.05.005
- López-López, E., Espinoza-Castañeda, J. I., Martínez-Mayorga, J. L., and Medina-Franco, A. C. (2025a). Structure–property associations: breaking paradigms for linking chemical structures and biological properties in drug discovery. *Chem. Med. Chem.* 25, e202500847. doi:10.1002/cmdc.202500847
- López-López, E., Medina-Franco, J. L., Salinas-Arellano, E., Ardila-Fierro, K. J., Pardo-Novoa, J. C., del Río, R. E., et al. (2025b). A covalent and modulable inhibitor of the tubulin-microtubule system: insights into the mechanism of cacalol. *Chem. Biol. Drug Des.* 106 (1–12), e70165. doi:10.1111/cbdd.70165
- M02 (2025). Clinical and Laboratory Standards Institute. Malvern, PA. Available online at: <https://clsi.org/shop/standards/m02/> (Accessed September 5, 2025).
- M100 (2025). Clinical and Laboratory Standards Institute. Malvern, PA. Available online at: <https://clsi.org/shop/standards/m100/> (Accessed September 5, 2025).
- Maggiore, G. M., and Bajorath, J. (2014). Chemical space networks: a powerful new paradigm for the description of chemical space. *J. Comp.-Aided Mol. Des.* 28 (8), 795–802. doi:10.1007/s10822-014-9760-0
- Maggiore, G., Vogt, M., Stumpfe, D., and Bajorath, J. (2014). Molecular similarity in medicinal chemistry: miniperspective. *J. Med. Chem.* 57 (8), 3186–3204. doi:10.1021/jm401411z
- Medina-Franco, J. L., and López-López, E. (2024). What is the plausibility that all drugs will be designed by computers by the end of the decade? *Exp. Op. Drug Discov.* 19 (5), 507–510. doi:10.1080/17460441.2024.2331734
- Medina-Franco, J. L., Sánchez-Cruz, N., López-López, E., and Díaz-Eufracio, B. I. (2021). Progress on open chemoinformatic tools for expanding and exploring the chemical space. *J. Comp.-Aided Mol. Des.* 36 (5), 341–354. doi:10.1007/s10822-021-00399-1
- Medina-Franco, J. L., Rodríguez-Pérez, J. R., Cortés-Hernández, H. F., and López-López, E. (2024). Rethinking the “best method” paradigm: the effectiveness of hybrid and multidisciplinary approaches in chemoinformatics. *Artif. Intel. Life Sci.* 6 (2024), 100117. doi:10.1016/j.ailesci.2024.100117
- Medina-Franco, J. L., Chávez-Hernández, A. L., López-López, E., and Saldívar-González, F. I. (2022). Chemical multiverse: an expanded view of chemical space. *Mol. Inf.* 41 (11), 2200116. doi:10.1002/minf.202200116
- Mora-Ramiro, B., Jiménez-Estrada, M., Zentella-Dehesa, A., Ventura-Gallegos, J. L., Gomez-Quiroz, L. E., Rosiles-Alanis, W., et al. (2020). Cacalol acetate, a sesquiterpene from *Pscacalium decompositum*, exerts an anti-inflammatory effect through LPS/NF- κ B signaling in raw 264.7 macrophages. *J. Nat. Prod.* 83 (8), 2447–2455. doi:10.1021/acs.jnatprod.0c00300
- Orlov, A. A., Akhmetshin, T. N., Horvath, D., Marcou, G., and Varnek, A. (2025). From high dimensions to human insight: exploring dimensionality reduction for chemical space visualization. *Mol. Inf.* 44 (1), e202400265. doi:10.1002/minf.202400265
- Robles-Zepeda, R. E., Velázquez-Contreras, C. A., Garibay-Escobar, A., Gálvez-Ruiz, J. C., and Ruiz-Bustos, E. (2011). Antimicrobial activity of Northwestern Mexican plants against *Helicobacter pylori*. *J. Med. Food* 14 (10), 1280–1283. doi:10.1089/jmf.2010.0263
- Romo, J., and Joseph-Nathan, P. (1964). The constituents of *Cacalea decomposita* A. Gray. *Tetrahedron* 20 (10), 2331–2337. doi:10.1016/s0040-4020(01)97622-7
- Rostro-Alonso, G. O., Castillo-Montoya, A. I., García-Acosta, J. C., Aguilar-Llanos, E. F., Quintas-Granados, L. I., Villegas-Vazquez, E. Y., et al. (2024). Cacalol acetate as anticancer agent: antiproliferative, Pro-apoptotic, cytostatic, and anti-migratory effects. *Curr. Issues Mol. Biol.* 46 (9), 9298–9311. doi:10.3390/cimb46090550
- Saar, M., Wawrzyk, A., Pastuszek-Lewandoska, D., and Bielec, F. (2025). Cefiderocol antimicrobial susceptibility testing by disk diffusion: influence of agar media and inhibition zone morphology in *K. pneumoniae* Metallo- β -lactamase. *Antibiotics* 14 (5), 527. doi:10.3390/antibiotics14050527
- Sander, T., Freyss, J., von Korff, M., and Rufener, C. (2015). DataWarrior: an open-source program for chemistry aware data visualization and analysis. *J. Chem. Inf. Model.* 55 (2), 460–473. doi:10.1021/ci500588j
- Shabani-Mashcool, S., Marashi, S.-A., and Gharaghani, S. (2020). NDDSA: a network-and domain-based method for predicting drug-side effect associations. *Inf. Proces. & Manag.* 57 (6), 102357. doi:10.1016/j.ipm.2020.102357
- Skov, R., Smyth, R., Larsen, A. R., Bolmström, A., Karlsson, A., Mills, K., et al. (2006). Phenotypic detection of Methicillin resistance in *Staphylococcus aureus* by disk diffusion testing and estest on mueller-hinton agar. *J. Clin. Microbiol.* 44 (12), 4395–4399. doi:10.1128/jcm.01411-06
- Soto Hernández, A. (2002). *Evaluación de la actividad antimicrobiana del cacalol y sus derivados*. [dissertation/bachelor's thesis]. [Mexico (CDMX)]. Mexico City, Mexico: Universidad Nacional Autónoma de Mexico City, Mexico.
- Steinig, E., Aglua, I., Duchene, S., Meehan, M. T., Yoannes, M., Firth, C., et al. (2022). Phylogenetic signatures in the emergence of community-associated MRSA. *Proceed. Natl. Acad. Sci.* 119 (45), e2204993119. doi:10.1073/pnas.2204993119
- Szwarc, S., Rutz, A., Lee, K., Meiri, Y., Bonnet, O., Hazni, H., et al. (2025). Translating community-wide spectral library into actionable chemical knowledge: a proof of concept with monoterpene indole alkaloids. *J. Cheminf.* 17 (1), 62–15. doi:10.1186/s13321-025-01009-0
- Talevi, A. (2024). “Network-pharmacology and central nervous system drug development.” in *AAPS introductions in the pharmaceutical sciences*, (Cham: Springer Nature Switzerland), 213–227. doi:10.1007/978-3-031-62604-3_9
- Talevi, A., and Bellera, C. (2024). An update on the novel methods for the discovery of antiseizure and antiepileptogenic medications: where are we in 2024? *Exp. Op. Drug Discov.* 19 (8), 975–990. doi:10.1080/17460441.2024.2373165
- Tong, S. Y. C., Fowler, V. G., Skalla, L., and Holland, T. L. (2025). Management of *Staphylococcus aureus* bacteremia: a review. *JAMA.* 334 (9), 798–808. doi:10.1001/jama.2025.4288
- Veber, D. F., Johnson, S. R., Cheng, H. Y., Smith, B. R., Ward, K. W., and Kopple, K. D. (2002). Molecular properties that influence the oral bioavailability of drug candidates. *J. Med. Chem.* 45 (12), 2615–2623. doi:10.1021/jm020017n
- Vigil-Vásquez, C., and Schüller, A. (2022). *De novo* prediction of drug targets and candidates by chemical similarity-guided network-based inference. *Int. J. Mol. Sci.* 23 (17), 9666. doi:10.3390/ijms23179666
- Wen, H., Nan, S., Zhang, J., Lei, Z., and Shen, W. (2024). Chemical space deconstruction-based dynamic model ensemble architecture for molecular property prediction. *Chem. Eng. Sci.* 295, 120118. doi:10.1016/j.ces.2024.120118

Wu, Z., Ma, H., Liu, Z., Zheng, L., Yu, Z., Cao, S., et al. (2022). wSDTNBI: a novel network-based inference method for virtual screening. *Chem. Sci.* 13 (4), 1060–1079. doi:10.1039/d1sc05613a

Yi, J., Lee, S., Lim, S., Cho, C., Piao, Y., Yeo, M., et al. (2023). Exploring chemical space for lead identification by propagating on chemical similarity network. *Comp. Struct. Biotech. J.* 21 (2023), 4187–4195. doi:10.1016/j.csbj.2023.08.016

Yosipof, A., Shimanovich, K., and Senderowitz, H. (2016). Materials informatics: statistical modeling in material science. *Mol. Inf.* 35 (11–12), 568–579. doi:10.1002/minf.201600047

You, Y., Lai, X., Pan, Y., Zheng, H., Vera, J., Liu, S., et al. (2022). Artificial intelligence in cancer target identification and drug discovery. *Signal Transduct. Target Ther.* 7 (1), 156. doi:10.1038/s41392-022-00994-0

Yu, Z., Wu, Z., Wang, Z., Wang, Y., Zhou, M., Li, W., et al. (2023). Network-based methods and their applications in drug discovery. *J. Chem. Inf. Model.* 64 (1), 57–75. doi:10.1021/acs.jcim.3c01613

Zdrzil, B., Felix, E., Hunter, F., Manners, E. J., Blackshaw, J., Corbett, S., et al. (2023). The ChEMBL database in 2023: a drug discovery platform spanning multiple bioactivity data types and time periods. *Nucleic Acids Res.* 52 (D1), D1180–D1192. doi:10.1093/nar/gkad1004

Zhu, T., Cao, S., Su, P. C., Patel, R., Shah, D., Chokshi, H. B., et al. (2013). Hit identification and optimization in virtual screening: practical recommendations based on a critical literature analysis. *J. Med. Chem.* 56 (17), 6560–6572. doi:10.1021/jm301916b