RSC Advances

PAPER

View Article Online
View Journal | View Issue

Open Access Article. Published on 28 February 2022. Downloaded on 2/10/2026 1:46:45 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Check for updates

Cite this: *RSC Adv.*, 2022, **12**, 6783

Cheminformatics analysis of molecular datasets of transcription factors associated with quorum sensing in *Pseudomonas aeruginosa*†

Felipe Victoria-Muñoz,[ab] Norberto Sánchez-Cruz,[c] José L. Medina-Franco[c] and Fabian Lopez-Vallejo*[b]

Received 14th November 2021
Accepted 20th February 2022

DOI: 10.1039/d1ra08352j

rsc.li/rsc-advances

Transcription factors associated with quorum sensing in *P. aeruginosa* are promising targets for discovering new adjuvants against infection with this pathogen. Regulation of these transcription factors offers the possibility of controlling multiple virulence factors related to them as biofilm development, proteases, hydrogen cyanide, among others. Numerous molecules have been tested against these targets, however, the keys responsible for antagonistic activity are still unknown. In this work, the structure–activity relationships of active molecules tested against LasR, PqsR, and RhlR transcription factors are analyzed in order to establish the structural characteristics associated. As part of the study, molecular complexity, scaffold, activity cliffs, and chemical space visualization analyses were conducted to find out characteristics associated with biological activity. In this study, several structural features were identified as significant for antagonist activity, highlighting molecular size and hydrogen bond acceptors.

1 Introduction

Infection with *P. aeruginosa* is a challenge in therapeutic management. These bacteria can adapt to a hostile environment, resist the immune system, and block the action of broad-spectrum antibiotics such as third-generation cephalosporin, carbapenems, and others.[1,2] For this reason, the World Health Organization has classified *P. aeruginosa* as a critical priority for the discovery of new drugs.[3]

One of the critical mechanisms related to the augmented antibiotic resistance of *P. aeruginosa* is quorum sensing (QS) which works like a communication system where each bacteria can detect an existent population in a given environment.[4] Small molecules called autoinducers (AIs) regulate the QS through their binding to specific receptors. They activate the genetic machinery to produce multiple virulence factors.[5]

At present, twenty transcription factors (TFs) involved in QS are known; all of them keep multiple interactions with each other.[6] Las, Rhl, PQS, and IQS are related to the activation of QS and the production of multiple virulence factors (elastase, protease, rhamnolipids, pyocyanin, hydrogen cyanide);[7,8] Las and Rhl are systems of the Lux type, while PQS is a quinolone dependent system.

Las and Rhl systems are key to QS activation. These systems have an inductor protein (LasI and RhlI) and a receptor protein (LasR and RhlR); inductors synthesize AIs and receptors recognize them. Lux type systems work with acyl-homoserine lactone molecules like AIs. These molecules are composed of a lactonic ring and a carbonate chain of variable size connected by an amide. *N*-3-Oxo-dodecanoyl homoserine lactone (3OC12-HSL) is the autoinducer of the Las system, while butyryl homoserine lactone (C4-HSL) is an autoinducer of the Rhl system (Fig. 1).[4,9]

2-Heptyl-3-hydroxy-4-quinolone (PQS) is the associated autoinducer to PQS system. PQS is recognized by PqsR and synthesized by multiple proteins such as PqsB, PqsC, PqsD, among others.[10–12] 2-(2-Hydroxyphenyl)-thiazole-4-carbaldehyde (IQS) is an autoinducer for IQS system. Previous studies established the activation of QS by IQS in deficiency of Las system,

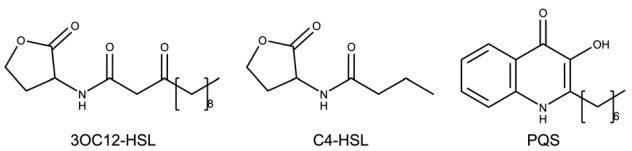

Fig. 1 Molecular structures of autoinducers associated with quorum sensing in *P. aeruginosa*. *N*-3-Oxo-dodecanoyl homoserine lactone (3OC12-HSL), butyryl homoserine lactone (C4-HSL) and 2-heptyl-3-hydroxy-4-quinolone (PQS).[4]

[a]Universidad Nacional de Colombia, Sede Bogotá, Facultad de Ciencias, Departamento de Farmacia, Av. Cra 30 # 45-03, Bogotá D.C., 11001 Colombia
[b]Universidad Nacional de Colombia, Sede Bogotá, Facultad de Ciencias, Departamento de Química, Grupo de Investigación en Productos Naturales Vegetales Bioactivos, Av. Cra 30 # 45-03, Bogotá D.C., 11001 Colombia. E-mail: fhlopezv@unal.edu.co
[c]DIFACQUIM Research Group, Department of Pharmacy, School of Chemistry, Universidad Nacional Autónoma de México, Avenida Universidad 3000, Mexico City, 04510 Mexico

† Electronic supplementary information (ESI) available. See DOI: 10.1039/d1ra08352j

© 2022 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 28 February 2022. Downloaded on 2/10/2026 1:46:45 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

generating interest for their results.[7,13] LasR, RhlR, and PqsR AIs are shown in Fig. 1.

Detecting new antagonists for these TFs gives innovative pathways for the *P. aeruginosa* infection management, nevertheless the structure–activity relationship of these targets is not known yet, highlighting interesting computational studies conducted on this topic.[14–16] For this reason, this paper proposes some relevant keys to contribute to solving this problem by searching for characteristics or structural patterns associated with agonist, antagonist activity, or inactivity. To achieve this goal, a dataset was built of 289 molecules with reported biological activity in major public repositories from LasR, PqsR, and RhlR, and cheminformatics studies were performed.

2 Methods

2.1 Construction and preparation of compound datasets

Compound datasets (DS) were built with ChEMBL version 26, PubChem, and scientific articles published in peer-reviewed journals. For ChEMBL, the following identifiers were used: LasR – CHEMBL1075207, PqsR – CHEMBL2424500, and RhlR – CHEMBL3112386. In PubChem bioassay, the search was done directly with AIs names. Additionally, for the previous search, only molecules with IC_{50} and EC_{50} values were taken into account. Keywords 'LasR', 'PqsR', 'RhlR' and 'quorum sensing' were used to search for molecules in scientific articles at SciFinder. Afterwards, all the results were filtered taking into account their biological activity values 'IC_{50}' or 'EC_{50}' according to all the articles published between 1998 and 2020.

All the molecules in DS were classified as "active" or "inactive" using a cutoff value of 10 μM for LasR and PqsR, while 12 μM was used for RhlR. These cutoff values were chosen based on the mean minus one standard deviation of biological activity. For a better understanding, some analysis required the pIC_{50} calculation ($-\log IC_{50}$). Then, the following steps were applied to prepare the molecules: (1) the largest fragment in each molecule was selected; (2) non-organic molecules were removed; (3) charges were neutralized by adding and/or removing hydrogens where possible; (4) it was ensured that the most potent acid groups were first ionized in partially ionized molecules; (5) stereochemistry was removed; (6) the most populated tautomer was selected; (7) duplicate molecules were discarded. This curation was done with RDKit and MolVS libraries for Python.[17,18]

2.2 Molecular complexity

Molecular complexity has been associated with target selectivity and success in clinical stages.[19–21] Here the relation between active compounds and molecular complexity was examined. For this study, complexity metrics based on topological descriptors like the fractions of sp^3 atoms (Fsp^3) and chiral centers (FCC) were used.[22] Likewise, the Bertz and DataWarrior complexity were calculated.[23,24] All the metrics were computed with RDKit and DataWarrior 5.0.0.[17,24]

2.3 Scaffolds analysis

Scaffolds can be defined as the core of a molecule used in drug design.[25–27] Scaffolds are intuitive and informative when exploring SAR and Structure-Multiple Activity Relationships (SmART) of compounds datasets as a way of identifying supports to drug design.[28] This work explores the relationships between scaffolds and the biological activity for LasR, PqsR, and RhlR datasets. For scaffold generation, in accordance with Bemis and Murcko as implemented in RDKit,[17,29] a scaffold is defined as the core structure of a compound consisting of all of its rings and connecting linkers. A maximum common substructure analysis was done on the most frequent scaffold in each DS.[30,31] The data were collected using RDKit.[17]

2.4 Activity cliffs analysis

Activity cliffs are small changes in the structure associated with unexpected large changes in biological activity.[32–34] Identification of activity cliffs is a relevant and informative analysis of the SAR because they point to specific and minor structural modifications that strongly influence biological activity. It is important to note, however, that there might be "artificial" cliffs (or artifacts) in compound data sets, for instance, when the biological activity is not measured correctly.[35] In this work, to identify the possible presence of activity cliffs in the DS, consensus activity cliffs were calculated between multiple 2D fingerprints (MACCS Keys 166-bits, ECFP4, ECFP3, ECFP2 and path length) as is proposed by Medina-Franco and coworkers,[36] finding the activity cliff common in all the fingerprints with Tanimoto coefficient. RDKit and MayaChemTools[44] were employed for this analysis.[17,37]

2.5 Chemical space visualization

The representation of a set of molecules in a two- or three-dimensional plot based on a set of pre-defined descriptors is usually known as visualization of the chemical space. In this work, the statistical analysis called t-distributed Stochastic Neighbor Embedding (t-SNE) was done so as to represent chemical space occupied by datasets, with a perplexity equal to 10, a learning rate equal to 100, using the Jaccard index as distance metric and ECFP4 to represent the compounds. Additionally, constellation plots recently developed were built for the datasets.[38] These plots integrate a methodology of analog series with the visualization of the chemical space, associating the molecular analog to several molecules and their biological activity.

3 Results and discussion

Table 1 summarizes the unique compounds found in ChEMBL, PubChem Bioassays and the scientific papers. In more detail, for ChEMBL can be found molecules with LasR, PqsR, and RhlR activities. For instance, for LasR there are 173 molecules reported, but only 82 have IC_{50} and EC_{50} values; 29 molecules are reported for PqsR, of which 8 have EC_{50} values; while from 31 compounds found for RhlR, only 4 of them have IC_{50} values. In comparison, PubChem Bioassay has 87, 10, and 17 compounds

 © 2022 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 28 February 2022. Downloaded on 2/10/2026 1:46:45 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Table 1 TFs datasets with unique compounds

		Agonist		Antagonist		Inactive	
DS	Total	*n*	%	*n*	%	*n*	%
LasR	188	52	27.7	41	21.8	95	50.5
PqsR	54	10	18.5	15	27.8	29	53.7
RhlR	47	12	25.6	7	14.9	28	59.5

with activity for LasR, PqsR, and RhlR, respectively. The remainder of the molecules was found in scientific articles. More than 70 percent of molecules are shared with ChEMBL. LasR is the TF with the highest number of molecules. Furthermore, the LasR dataset has the highest proportion of active molecules among the three datasets. These datasets can be found in the ESI.†

3.1 Molecular complexity

Target selectivity is a major issue in drug discovery. The relationship between higher molecular complexity and less target promiscuity had already been suggested.[19,20,22,39,40] In the present study, an association between DataWarrior complexity and biological activity was found as shown in Fig. 2.

The majority of PqsR active molecules have the highest DataWarrior complexity and some of them have pIC_{50} higher than 7. In contrast, RhlR compounds are towards the left bottom corner of Fig. 2, where molecules with low activity and poor complexity are located. Finally, LasR active compounds could not be associated with specific complexity values because this dataset has a high molecular diversity as it is evidenced in analysis further on. Antagonist molecules in all the datasets do not have a clear behaviour associated with complexity for the metrics studied. These results are available in the ESI.†

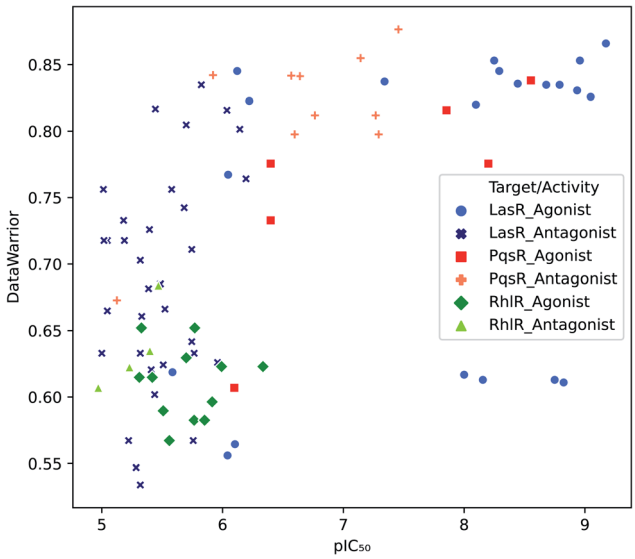

Fig. 2 Distribution of biological activity values against DataWarrior complexity. Molecules with activity against LasR, PqsR and RhlR are in blue, red and green respectively.

3.2 Scaffold analysis

Fig. 3 shows scaffold frequency of the DS on the left side; scaffold structures are displayed on the right. The frequency of scaffolds found for DS varies according to DS size (Fig. 3). Nevertheless, the five most frequent scaffolds represent around 20% of the total, highlighting a variety of types of biological activity in these scaffolds. In LasR, it is found that those first five scaffolds have molecules with activity, being the most frequent the agonist activity, which is shown in Fig. 3a. Rings of six members connected by a short amide chain are the most relevant ones for this DS.

Three phenyl system is present in the second and fourth most populated scaffolds in Fig. 3a but the position of amide in chain allows antagonistic activity for the fourth scaffold. Even though SC1 and SC2 scaffolds are based on the common lactone ring in autoinducers, they show all the three activities, therefore these scaffolds are unsuitable for similarity search. SC3 and SC4 could be useful.

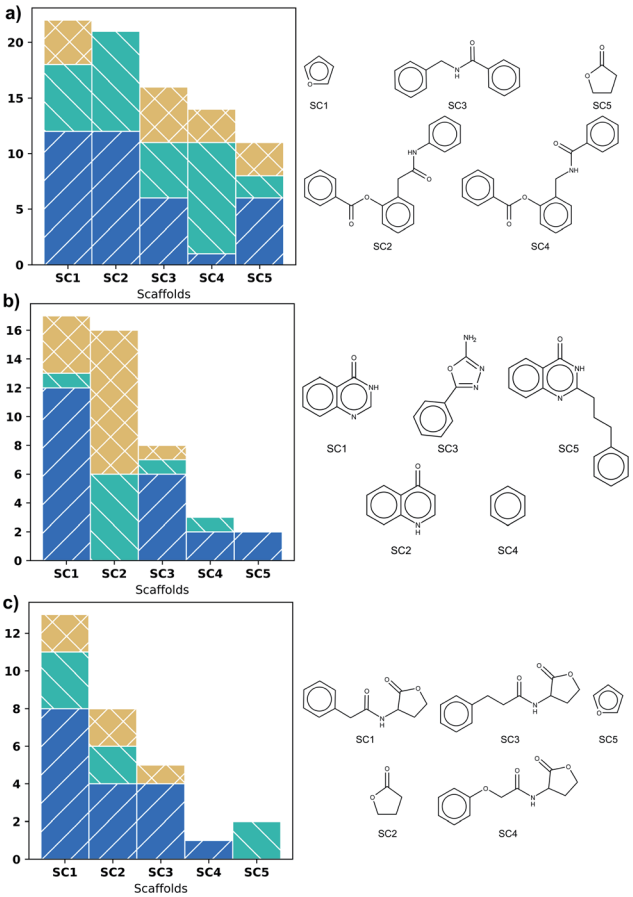

Fig. 3 First five most frequent Bemis-Murcko scaffolds in: (a) LasR dataset, (b) PqsR dataset and (c) RhlR dataset. Activity type is represented by colours: inactive (blue), agonist (green), and antagonist molecules (orange).

Open Access Article. Published on 28 February 2022. Downloaded on 2/10/2026 1:46:45 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Scaffolds of PqsR DS are compounds with six-membered rings fused, a common feature of the natural autoinducer. All the molecules with quinoline ring (SC2) have an activity lower than 10 μM (Fig. 3b), making it a suitable candidate for similarity search of new molecules with potential biological activity. It is important to mention that the first two scaffolds (SC1 and SC2) represent more than fifty percent of the molecules, revealing a low scaffold diversity in this DS.

The biological activity of the first five scaffolds found in RhlR DS is higher than the activity of scaffolds in the rest of DS. Here all the scaffolds have a lactone ring, but scaffold (SC4), the one with the largest percent of antagonist molecules, also presents an amide chain with an ether and a phenyl moiety.

Analysis of the maximum common substructure (MCS) allows to detect small structural differences between molecules with different types of activity. All the molecules shown in Fig. 4 for LasR DS were reported by O'Reilly and collaborators,[41] they have a triphenyl scaffold showing the activity related to a specific ring substitution. The inactive molecule is *ortho* and *meta* disubstituted by fluorine atom in the first ring, the agonist and antagonist are monosubstituted in the *ortho* position by chlorine atom in the first ring, but the antagonist is additionally monosubstituted in the *para* position in the third ring. These substitutions can change the inductive effect generating in this way different interactions in the binding site.[42]

Ilangovan *et al.* reported several compounds associated with PqsR activity.[43] Those compounds differ from each other only in the substitution of nitrogen in the quinazoline ring: the inactive compound has a hydroxyl group, the agonist has ethyl, and the antagonist has an amine, finding the association of biological activity related to hydrogen bonds acceptors and donors of this position (Fig. 4). For RhlR DS, it is possible to establish a steric effect in the binding site with the type of activity, as it is shown in Fig. 4. This analysis allows to establish a mechanism of action cliffs (MOA-cliffs) for these scaffolds.[44] All of these scaffolds and its frequency are available in the ESI.†

3.3 Activity cliffs

Fig. 5 shows the structure of representative activity cliffs of the agonists and antagonist for all the TFs. It is evident the relationship of agonist and antagonist molecules, supported by our previous analysis. Regarding LasR agonists, a change of the EC_{50} values can be noted by a variation in one carbon on the side chain: BO1 has a side chain of 11 carbons, while CH202 has 10 carbons, causing a 1-fold decrease in biological activity. Furthermore, natural autoinducer has a 12-carbons side-chain. All this allows us to understand the importance of the size of side-chain in agonist activity. The molecular pair found for the antagonist activity of LasR, MA3 and CH184, differ in the halogen atom at the end of the side-chain: MA3 has chlorine while CH184 has bromine. Thus, MA3 is about 10 times more active than CH184 relating the possibility to make a halogen bond, which could be a possible cause of antagonist activity.

LU3-1 and IL4 is the molecular pair found for agonist activity of PqsR, LU3-1 has a 7-carbon side-chain, while IL4 has a 9-carbon side-chain, agonist activity of LU3-1 is about two orders of magnitude greater than IL4, once again confirming the importance of side-chain length in agonist activity. On the other hand, the pair of compounds LU2 and LU10 were activity cliffs related to antagonist activity; both have a 6-carbons side-chain, but LU2 has an oxygen atom at position four, while in LU10 they all are carbons. This change is associated with a potency difference in two orders of magnitude. The potency difference could be explained by the lack of a hydrogen bond acceptor in LU10.

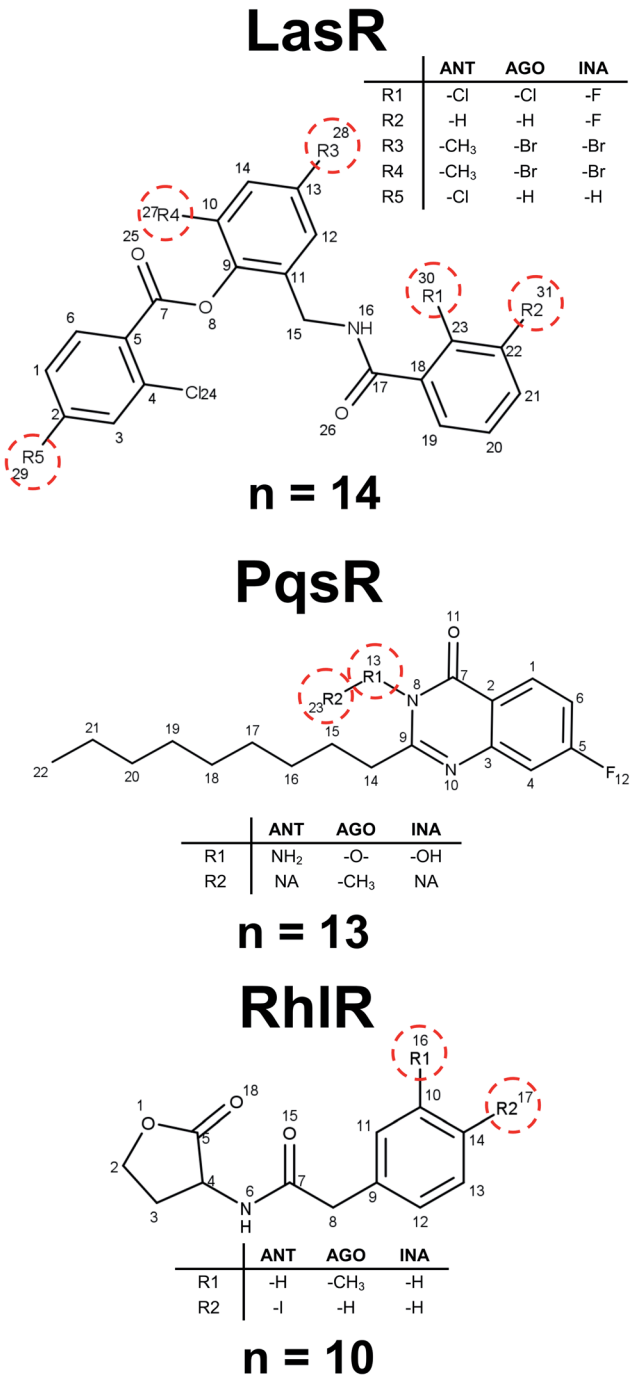

Fig. 4 Representative examples of maximum common substructures. *n* represents the number of molecules with this scaffold.

 © 2022 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 28 February 2022. Downloaded on 2/10/2026 1:46:45 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

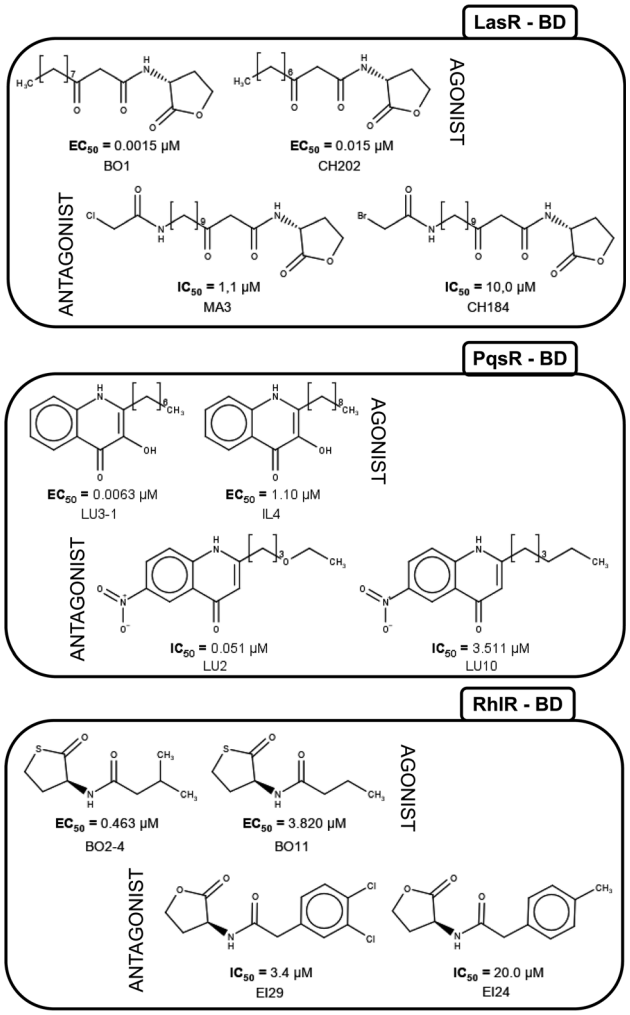

Fig. 5 Activity cliffs found in datasets. The agonist is shown to the left and the antagonist to the right.

Activity cliffs for RhlR agonists are molecules with a thiolactone ring and a side-chain, here BO11 side chain changes a hydrogen atom in the fourth position to a methyl group in compound BO2–4; activity for BO2–4 is about eight times greater than BO11.

It is notable that the natural autoinducer of RhlR has an EC_{50} (8.95 μM) lower than these compounds, and its difference is a lactonic ring *versus* thiolactone ring as reported by Boursier and co-workers (Fig. 6).[45] The substitution on the phenyl ring in molecular pairs found for antagonist activity generates a change of this activity, while EI29 is two chlorine-substituted, EI24 is monomethyl substituted, linking the size and inductive effects with antagonist activity.[46]

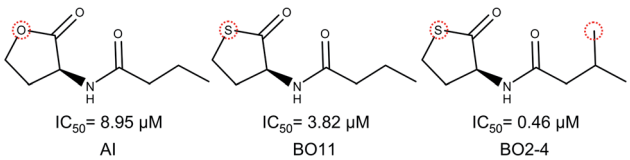

Fig. 6 Structural comparison of autoinducer (AI) and agonist activity cliffs of RhlR dataset (BO11 and BO2–4).[40]

3.4 Chemical space

Visualization of the Euclidean distance between molecules by using structural parameters as a classification method provides a comprehensive view of the molecular data sets. Therefore, a representation was done of a chemical space of DS related to a type of biological activity, which generates a general overview of the structural diversity of the molecules (Fig. 7). The chemical space visualization, occupied by compounds in DS and based on ECFP4 fingerprints, shows that the agonist and antagonist molecules are close in space due to the fact that the type of compounds are similar to each other, according to our previous results. In the lower right corner of the plot, molecules with five-membered rings are displayed, usually lactone and thiolactone rings. In this region of the chemical space, compounds of LasR and RhlR DS can be found. LasR agonists and antagonists are closer to each other than molecules of RhlR since the spatial location is given by the structure.

In the upper right corner of Fig. 7, molecules with triphenyl scaffold are found, a common characteristic of LasR DS. However, in order to find activity cliffs, a group of antagonist molecules is displayed separated from the agonist compounds; this might be a critical starting point for optimizing the searching of antagonist activity or similarity in big databases. Compounds of PqsR DS are displayed in the center of the plot; quinoline and quinazoline rings are usual scaffolds with PqsR activity.

In the visualization of the chemical space in Fig. 7, all the DS are close in the space, especially LasR and RhlR DS. This is an indication that it is possible to find a molecule with antagonist activity against both TFs; this is attributable to its high similarity. The proximity between agonist and antagonist compounds shows the challenge of finding specific structural characteristics associated with antagonist activity against all the TFs.

3.5 Constellation plots

A constellation plot is a hybrid approach of visualization of chemical space and substructure analysis. As described in detail elsewhere,[38] the main analog series of a compound data set were generated. The structural similarity of the analog series is displayed in a two-dimensional representation of chemical space such as t-SNE. Constellation plots represent a method to visualize and analyse StARs (Structure–Activity Relationships) in chemical space.[47,48]

For this paper, constellation plots were built following the methodology proposed by Naveja and Medina-Franco.[38] These analyses provide with the relation of standard cores in the DS with biological activity and the number of compounds that they have. Constellation plot for LasR DS (Fig. 8) shows nineteen cores grouped in ten analog series; they contained one hundred and six compounds (around half of the dataset).

The first three analog series display small structural changes in diphenyl rings, both the first and third series show diphenyl

Open Access Article. Published on 28 February 2022. Downloaded on 2/10/2026 1:46:45 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

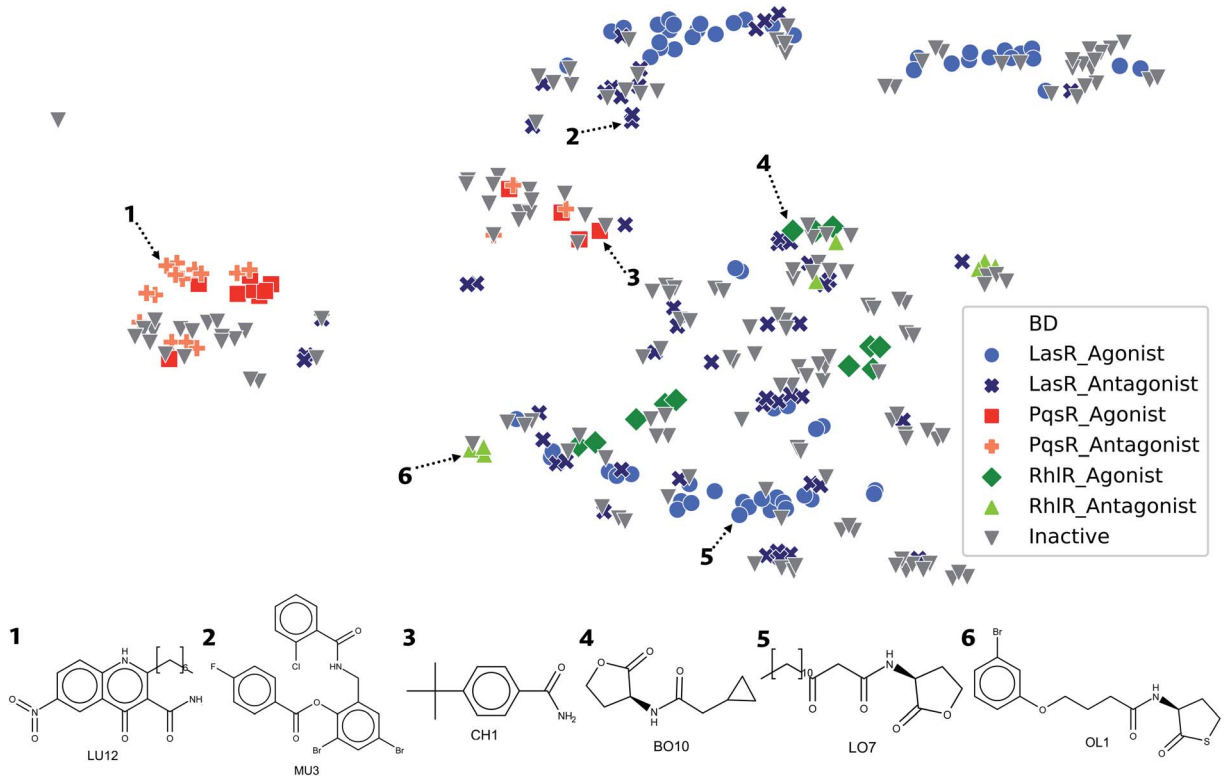

Fig. 7 Visual representation of the structural chemical space obtained by t-distributed stochastic neighbor embedding (t-SNE). Molecules are classified according to their activity: inactive (grey), LasR agonist (light blue), LasR antagonist (dark blue), PqsR agonist (red), PqsR antagonist (pink), RhlR agonist (light green), and RhlR antagonist (dark green).

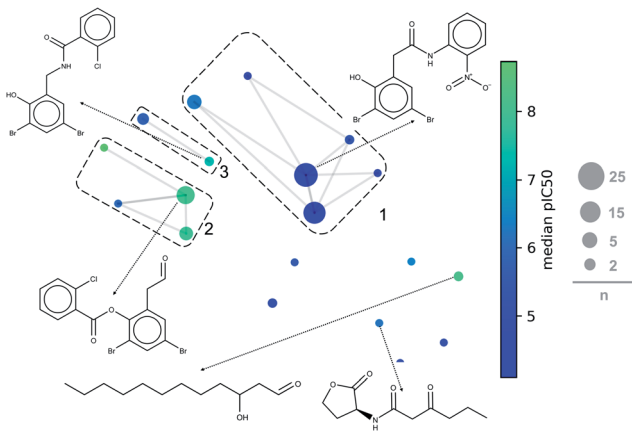

Fig. 8 General constellation plot for a dataset of LasR. The most representative cores are shown. This plot has the same parameters of Fig. 7 (t-SNE, ECFP4, Tanimoto Matrix). Each point represents a core, the size of circles indicates the number "*n*" of compounds annotated; connected circles are cores sharing compounds; grids represent analog series, and colour shows the average of pIC_{50}.

rings connected by a chain of three members with an amide group in the middle of them. The second series has diphenyl rings connected by a chain of two members with an ester group in the middle of them, that series has a better activity profile which sets an excellent base for searching other compounds or a useful scaffold for optimizing biological activity.

Other essential cores for this activity are a carbonate chain of twelve members and a lactone ring. The carbon chain compounds have an excellent mean activity for this target, but the lactone ring does not; this is because the lactone ring has multiple compounds with a harmful activity while carbon chain has a significant number of compounds with functional activity. PqsR and RhlR DS could not be analysed with this method because the rules of fragmentation used does not involve carbon chains or rings.[43]

4 Conclusions

PqsR active molecules showed a high DataWarrior complexity. In RhlR DS, all the molecules were correlated with low DataWarrior complexity; in contrast, LasR active compounds were associated with both high and low complexity values, making impossible a specific association with DataWarrior complexity.

For LasR antagonist compounds, it was identified that the side-chain size in homoserine lactone scaffolds should range from 10 to 12 atoms. In contrast, the rings substitution in triphenyl scaffolds generates a specific type of activity. In molecules active against PqsR, the size of the side-chain quinazoline rings and substitution on the nitrogen result in a certain type of activity. Active homoserine lactone scaffolds for RhlR show that phenyl or ethyl substitution on the side-chain does give also a type of activity.

 © 2022 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 28 February 2022. Downloaded on 2/10/2026 1:46:45 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

The maximum common substructure analysis and chemical space visualization have indicated that there is no scaffold associated with a unique type of activity in any TF. Yet, the fact that most of them are active should be considered as an encouraging compound to optimize. The LasR constellation plot allowed to determine that the diphenyl scaffold is one of the most promising for optimizing molecules since it is present in more than 25 molecules with LasR biological activity.

Nevertheless, it is important to highlight the ambiguity of these characteristics as in some cases it could generate a change in the mode of action, as described in MOA-cliffs, making compound optimization difficult. To establish the structural keys with a type of activity, it is advisable to use molecular docking and dynamics simulations for a deeper understanding at the molecular level.

Author contributions

Felipe Victoria-Muñoz: conceptualization, data curation, formal analysis, methodology, writing – original draft. Norberto Sanchez-Cruz: conceptualization, formal analysis, methodology, data curation. Jose L Medina-Franco: conceptualization, methodology, formal analysis, review & editing, supervision. Fabian Lopez-Vallejo: conceptualization, methodology, formal analysis, review & editing, project administration, supervision, funding acquisition.

Conflicts of interest

The authors declare that there is no financial or commercial conflict of interest.

Acknowledgements

We thank Dr Jesus Naveja for sharing the code to build the Constellation Plots. This research was supported by the Ministerio de Ciencia, Tecnología, e Innovación (Minciencias) (contract 835-2017/grant number 110177758105) with the project entitled "Identificación de moléculas líderes de origen natural con acción multidiana como inhibidores de quorum sensing en *Pseudomonas aeruginosa* multirresistente" and Fundación para la Promoción de la Investigación y la Tecnología (FPIT) (grant number 4.553) with the project entitled "Análisis quimioinformático de moléculas activas contra factores de transcripción asociados a quorum sensing en *Pseudomonas aeruginosa*".

References

1 C. Peña, C. Suarez, F. Tubau, A. Dominguez, M. Sora, M. Pujol, F. Gudiol and J. Ariza, *Eur. J. Clin. Microbiol. Infect. Dis.*, 2009, **28**, 519–522.
2 A. Potron, L. Poirel and P. Nordmann, *Int. J. Antimicrob. Agents*, 2015, **45**, 568–585.
3 E. Tacconelli, E. Carrara, A. Savoldi, S. Harbarth, M. Mendelson, D. L. Monnet, C. Pulcini, G. Kahlmeter, J. Kluytmans, Y. Carmeli, M. Ouellette, K. Outterson, J. Patel, M. Cavaleri, E. M. Cox, C. R. Houchens, M. L. Grayson, P. Hansen, N. Singh, U. Theuretzbacher, N. Magrini, A. O. Aboderin, S. S. Al-Abri, N. Awang Jalil, N. Benzonana, S. Bhattacharya, A. J. Brink, F. R. Burkert, O. Cars, G. Cornaglia, O. J. Dyar, A. W. Friedrich, A. C. Gales, S. Gandra, C. G. Giske, D. A. Goff, H. Goossens, T. Gottlieb, M. Guzman Blanco, W. Hryniewicz, D. Kattula, T. Jinks, S. S. Kanj, L. Kerr, M. P. Kieny, Y. S. Kim, R. S. Kozlov, J. Labarca, R. Laxminarayan, K. Leder, L. Leibovici, G. Levy-Hara, J. Littman, S. Malhotra-Kumar, V. Manchanda, L. Moja, B. Ndoye, A. Pan, D. L. Paterson, M. Paul, H. Qiu, P. Ramon-Pardo, J. Rodríguez-Baño, M. Sanguinetti, S. Sengupta, M. Sharland, M. Si-Mehand, L. L. Silver, W. Song, M. Steinbakk, J. Thomsen, G. E. Thwaites, J. W. van der Meer, N. Van Kinh, S. Vega, M. V. Villegas, A. Wechsler-Fördös, H. F. L. Wertheim, E. Wesangula, N. Woodford, F. O. Yilmaz and A. Zorzet, *Lancet Infect. Dis.*, 2018, **18**, 318–327.
4 V. C. Kalia, *Quorum Sensing vs. Quorum Quenching: A Battle with No End in Sight*, Springer India, New Delhi, 2015.
5 G. A. March Rosselló and J. M. Eiros Bouza, *Med. Clin.*, 2013, **141**, 353–357.
6 H. Huang, X. Shao, Y. Xie, T. Wang, Y. Zhang, X. Wang and X. Deng, *Nat. Commun.*, 2019, **10**, 1–13.
7 J. Lee and L. Zhang, *Protein Cell*, 2015, **6**, 26–41.
8 M. Schuster and E. Peter Greenberg, *Int. J. Med. Microbiol.*, 2006, **296**, 73–81.
9 M. B. Miller and B. L. Bassler, *Annu. Rev. Microbiol.*, 2001, **55**, 165–199.
10 E. C. Pesci, J. B. J. Milbank, J. P. Pearson, S. Mcknight, A. S. Kende, E. P. Greenberg and B. H. Iglewski, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 11229–11234.
11 L. A. Gallagher, S. L. McKnight, M. S. Kuznetsova, E. C. Pesci and C. Manoil, *J. Bacteriol.*, 2002, **184**, 6472–6480.
12 E. Déziel, F. Lépine, S. Milot, J. He, M. N. Mindrinos, R. G. Tompkins and L. G. Rahme, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 1339–1344.
13 J. Lee, J. Wu, Y. Deng, J. Wang, C. Wang, J. Wang, C. Chang, Y. Dong, P. Williams and L.-H. Zhang, *Nat. Chem. Biol.*, 2013, **9**, 339–343.
14 O. Kindler, O. Pulkkinen, A. G. Cherstvy and R. Metzler, *Sci. Rep.*, 2019, **9**, 1–19.
15 A. A. M. Kamal, L. Petrera, J. Eberhard and R. W. Hartmann, *Org. Biomol. Chem.*, 2017, **15**, 4620–4630.
16 J. E. Paczkowski, A. R. McCready, J.-P. Cong, Z. Li, P. D. Jeffrey, C. D. Smith, B. R. Henke, F. M. Hughson and B. L. Bassler, *ACS Chem. Biol.*, 2019, **14**, 378–389.
17 RDKit, *Open-source cheminformatics*, http://www.rdkit.org.
18 MolVS, *Molecule Validation and Standardization*, https://molvs.readthedocs.io/en/latest/index.html.
19 F. Lovering, J. Bikker and C. Humblet, *J. Med. Chem.*, 2009, **52**, 6752–6756.
20 F. Lovering, *MedChemComm*, 2013, **4**, 515–519.
21 P. A. Clemons, N. E. Bodycombe, H. A. Carrinski, J. A. Wilson, A. F. Shamji, B. K. Wagner, A. N. Koehler and

Open Access Article. Published on 28 February 2022. Downloaded on 2/10/2026 1:46:45 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

S. L. Schreiber, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 18787–18792.
22 O. Méndez-Lucio and J. L. Medina-Franco, *Drug Discovery Today*, 2017, **22**, 120–126.
23 S. Bertz, *J. Am. Chem. Soc.*, 1981, **103**, 3599–3601.
24 T. Sander, J. Freyss, M. Von Korff and C. Rufener, *J. Chem. Inf. Model.*, 2015, **55**, 460–473.
25 Y. Hu, D. Stumpfe and J. Bajorath, *J. Chem. Inf. Model.*, 2011, **51**, 1742–1753.
26 D. Stumpfe, D. Dimova and J. Bajorath, *J. Med. Chem.*, 2016, **59**, 7667–7676.
27 B. Zdrazil and R. Guha, *J. Med. Chem.*, 2018, **61**, 4688–4703.
28 F. I. Saldívar-González, J. J. Naveja, O. Palomino-Hernández and J. L. Medina-Franco, *RSC Adv.*, 2017, **7**, 632–641.
29 G. W. Bemis and M. A. Murcko, *J. Med. Chem.*, 1996, **39**, 2887–2893.
30 A. Golbamaki, A. M. Franchi and G. Gini, in *Advances in QSAR Modeling. Challenges and Advances in Computational Chemistry and Physics*, 24th edn, 2017, pp. 149–165.
31 Y. J. Xu and M. Johnson, *J. Chem. Inf. Comput. Sci.*, 2002, **42**, 912–926.
32 G. M. Maggiora, *J. Chem. Inf. Model.*, 2006, **46**, 1535.
33 M. Cruz-Monteagudo, J. L. Medina-Franco, Y. Pérez-Castillo, O. Nicolotti, M. N. D. S. Cordeiro and F. Borges, *Drug Discovery Today*, 2014, **19**, 1069–1080.
34 Y. Hu and J. Bajorath, *Eur. J. Med. Chem.*, 2013, **69**, 393–398.
35 J. L. Medina-Franco, *Chem. Biol. Drug Des.*, 2013, **81**, 553–556.
36 M. González-Medina, F. D. Prieto-Martínez, J. R. Owen and J. L. Medina-Franco, *J. Cheminf.*, 2016, **8**, 1–11.
37 M. Sud, *J. Chem. Inf. Model.*, 2016, **56**, 2292–2297.
38 J. J. Naveja and J. L. Medina-Franco, *Front. Chem.*, 2019, **7**, 1–10.
39 Y. Hu, D. Gupta-Ostermann and J. Bajorath, *Comput. Struct. Biotechnol. J.*, 2014, **9**, e201401003.
40 S. Monteleone, J. E. Fuchs and K. R. Liedl, *Front. Pharmacol.*, 2017, **8**, 1–9.
41 M. C. O'Reilly, S. H. Dong, F. M. Rossi, K. M. Karlen, R. S. Kumar, S. K. Nair and H. E. Blackwell, *Cell Chem. Biol.*, 2018, **25**, 1128–1139.
42 *The Practice of Medicinal Chemistry*, ed. W. Camille Georges, Academic Press, London, 2011.
43 A. Ilangovan, M. Fletcher, G. Rampioni, C. Pustelny, K. Rumbaugh, S. Heeb, M. Cámara, A. Truman, S. R. Chhabra, J. Emsley and P. Williams, *PLoS Pathog.*, 2013, **9**, 1–17.
44 M. Hao, S. H. Bryant and Y. Wang, *J. Cheminf.*, 2016, **8**, 37.
45 M. E. Boursier, J. B. Combs and H. E. Blackwell, *ACS Chem. Biol.*, 2019, **14**, 186–191.
46 N. R. Eibergen, J. D. Moore, M. E. Mattmann and H. E. Blackwell, *Chembiochem*, 2015, **16**, 2348–2356.
47 J. L. Medina-Franco, J. J. Naveja and E. López-López, *Drug Discovery Today*, 2019, **24**, 2162–2169.
48 J. L. Medina-Franco, N. Sánchez-Cruz, E. López-López and B. I. Díaz-Eufracio, *J. Comput.-Aided Mol. Des.*, 2022, DOI: 10.1007/s10822-021-00399-1.

 © 2022 The Author(s). Published by the Royal Society of Chemistry