RESEARCH ARTICLE

Computational study of the interactions among structural analogues of acyl homoserine lactones (AHLs) and the *Agrobacterium tumefaciens* TraR binding site [version 1; peer review: 2 approved with reservations]

Daniela Pérez¹, Maicol Ahumado^{1,2}, Eileen Herrera¹, Catalina Vivas-Gomez¹, Ricardo Vivas-Reyes ^{1,2}

¹Grupo de Química Cuántica y Teórica, Programa de Química, Facultad de Ciencias Exactas y Naturales, universidad de Cartagena, Cra 50 # 29-11, Cartagena, Colombia

²Grupo de Investigación CIPTEC, Programa de Tecnología en Producción Industrial, universidad de Cartagena, Cra 44 D # 30a-91, Cartagena, Colombia

V1 First published: 05 Dec 2019, 8:2062
<https://doi.org/10.12688/f1000research.20793.1>
 Latest published: 05 Dec 2019, 8:2062
<https://doi.org/10.12688/f1000research.20793.1>

Abstract

Background: In the present investigation, relationships between a set of 34 analogues of N-acyl-L-homoserine lactones (AHL) and the TraR receptor were studied. The aim was to use molecular modeling as a strategy for elucidating important aspects of the mechanism of chemical signaling in the Gram-negative bacteria *Agrobacterium tumefaciens*, with the idea of identifying some of analogues' structural characteristics and molecular interactions with the active site of the TraR receptor.

Methods: For this purpose, we combine two molecular modeling strategies: molecular docking and three-dimensional quantitative structure-activity relationship (3D-QSAR). First, the molecular docking methodology was applied to a series of 34 analogues of AHL on the TraR transcriptional receptor to simulate the binding of analogues at the active TraR site. Secondly, 3D-QSAR models were generated to describe the correlation with the experimental biological activity using partial least squares (PLS) calculations and steric and electrostatic properties, which theoretically predict the activity of the 34 AHL analogues through statistical parameters and evaluate the prediction of the models obtained. Two alignment models were constructed; one using the optimized structures of the 34 analogues (ligand-based model) and another using the conformations of the best poses generated in the docking with TraR (receptor-based model).

Results: The outcomes obtained for each protein-ligand complex showed that the Aspartic acid 70 and Threonine 129 residues are residues that participate in the formation of hydrogen bonds, while

Open Peer Review

Approval Status ? ?

	1	2
version 1
05 Dec 2019	view	view
1. Nestor Cubillan , University of Atlántico, Barranquilla, Colombia		
2. Leonardo C. Pacheco-Londoño , University of Puerto Rico, Mayaguez, Puerto Rico		
Universidad Simón Bolívar, Barranquilla, Colombia		

Any reports and responses or comments on the article can be found at the end of the article.

residues Alanine 38, Leucine, 40, Tyrosine 53, Glutamine 58, Tyrosine 61, Phenylalanine 62 and Valine 72 form hydrophobic interactions. These interactions are important in determining the antagonistic activity of the analogues under study against TraR.

Conclusions: The ligand-based model produces better statistical results expressed in terms of several rigorous evaluation criteria, such as Q^2 and R^2 for the data sets than those of the receptor-based model.

Keywords

Active TraR site, Quorum Sensing, Molecular Docking, 3D-QSAR

This article is included in the **Chemical Information Science** gateway.

Corresponding author: Ricardo Vivas-Reyes (rvivasr@unicartagena.edu.co)

Author roles: **Pérez D:** Formal Analysis, Investigation, Methodology, Writing – Original Draft Preparation; **Ahumedo M:** Data Curation, Formal Analysis, Investigation, Methodology, Software, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Herrera E:** Data Curation, Methodology, Writing – Original Draft Preparation; **Vivas-Gomez C:** Conceptualization, Data Curation, Investigation, Writing – Original Draft Preparation, Writing – Review & Editing; **Vivas-Reyes R:** Conceptualization, Data Curation, Investigation, Methodology, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: Financial support for this work was provided by the Fundación Universitaria Tecnológico de Comfenalco [008430]. *The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*

Copyright: © 2019 Pérez D *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Pérez D, Ahumedo M, Herrera E *et al.* **Computational study of the interactions among structural analogues of acyl homoserine lactones (AHLs) and the *Agrobacterium tumefaciens* TraR binding site** [version 1; peer review: 2 approved with reservations] F1000Research 2019, **8**:2062 <https://doi.org/10.12688/f1000research.20793.1>

First published: 05 Dec 2019, **8**:2062 <https://doi.org/10.12688/f1000research.20793.1>

Introduction

The intracellular communication mechanism known as quorum sensing (QS) makes it easier for bacteria to develop cooperative behaviors, which coordinate their activities in order to function as a multicellular unit¹. QS is a process based on the synthesis, diffusion, and perception of small signal molecules called autoinducers (AIs), through which a bacterial population can monitor their cell density and regulate specific gene expression in response to local changes in its population density^{2,3}. Bacterial behaviors regulated by QS include the secretion of virulence factors, biofilm formation, bioluminescence, sporulation and conjugation⁴.

QS in Gram-negative bacteria is mediated by N-acyl homoserine lactones (AHLs). In this AHL-dependent system, the QS signal is detected by a cytosolic transcription factor (R). AHL synthase (Protein I) produces AHL molecules that can diffuse through the cell membrane and as they reach a critical extracellular concentration, they enter the cell again and bind to the R protein, forming the R-AHL complex, which activates the expression of some specific genes^{5,6}.

The Gram-negative bacterial species *Agrobacterium tumefaciens* regulates the expression of the genes necessary for the replication and conjugation of the tumor-inducing plasmid in dicotyledonous Ti plants through the QS system, causing the disease known as 'crown gall' that occurs in greenhouse and fruit crops⁷. Its QS system is homologous to LuxI /LuxR by *Vibrio fischeri* and is made up of the autoinducer synthase, the TraI protein, the N-(3-oxo-octanoyl)-L-homoserine lactone (3OC8HSL) autoinducer and the transcriptional receptor, the TraR protein⁵. This communication system is activated when the lactones synthesized by TraI at high population densities accumulate until reaching a certain level and when they reach that concentration threshold, they are detected by TraR. The TraR-3OC₈HSL complexes activate the transcription of the target genes in the Ti plasmid, including those required for conjugation⁸. In effect, the TraR protein is a transcriptional regulator in the QS mechanism of *A. tumefaciens*.

Since QS signaling is vital for the regulation of bacterial gene expression that has an impact on numerous microbial physiological functions, including those related to pathogenicity in humans, plants and animals, there is enormous interest in the design and implementation of QS system inactivation strategies⁹. One of the strategies for interruption of the QS is the blockade of the transduction of the QS signal, which can be achieved by the presence of an AHL antagonist capable of competing or interfering with the binding of the signal molecule to its LuxR receptor¹⁰.

For this purpose, we combined two molecular modeling strategies: molecular docking and three-dimensional quantitative structure-activity relationship (3D-QSAR). First, the molecular docking methodology was applied to a series of 34 analogues of AHL on the TraR transcriptional receptor to simulate the binding of analogues at the active TraR site. Secondly, analysis of the structural characteristics and interactions of the ligands

was carried out in order to evaluate how the inhibitory activity is affected by its properties. Finally, 3D-QSAR models were generated to describe the correlation between the experimental biological activity using partial least squares (PLS) calculations and steric and electrostatic properties, which theoretically predict the activity of the 34 analogues of AHL through statistical parameters and evaluate the prediction of the models obtained.

Methods

Dataset

The chemical structures of 34 AHL analogues were taken together with their half maximal inhibitory concentration (IC₅₀) biological activity from Geske *et al.* (Table 1)^{11,12}. These structures were optimized using the theory of functional density with the functional X3LYP, together with the DGDZVP calculation base^{13,14}, using the Gaussian09 program¹⁵ (as an open-source alternative, GAMESS could be used¹⁶).

Molecular docking

Molecular docking calculations were performed to predict the binding modes among the 34 analogues of AHL and the TraR receptor. The 3D structure of the TraR receptor and its native co-crystallized ligand 3OC8HSL were downloaded from the Protein Data Bank, with the identifier 1L3L. The complex in PDB format was visualized and prepared using the structure preparation tool available in the Sybyl X 2.1.1 package¹⁷ and the Amber force field was used to assign the partial atomic charges of the protein. As an alternative, the open-source program, Avogadro could be used¹⁸. The native ligand and all water molecules were removed from the complex, hydrogen atoms were added and side chain amides and imidazoles were fixed (protonated), assuming a physiological pH of 7. The structure of extracted native ligand was optimized with the same conditions as the 34 analogues, mentioned above.

Once the structure of the native ligand and the TraR receptor was obtained, the docking protocol was initiated in the molecular simulation program AutoDock 4.2¹⁹. The preparation of both the protein and ligand was performed using the graphical user interface of AutoDock, called AutoDockTools²⁰. This preparation involved the addition of hydrogen and the assignment of Gasteiger charges²¹. After this, the potential maps for each type of atom in the ligand were calculated, executing the grid parameter using the Autogrid program incorporated in AutoDock. In this study, the grid box was delimited, centered on the TraR binding site, with dimensions 50x50x50 Å and spacing between points of 0.375 Å. Finally, for the conformational search, the Lamarckian genetic algorithm was chosen and for the validation 500 conformations were registered, with the remaining parameters used as per the default setting. With these parameters, docking validation was carried out by re-coupling the native ligand to the binding site of the TraR receptor.

Once the docking protocol was validated, and its ability to reproduce the interactions that promote protein-ligand binding confirmed for this particular case, TraR-3OC₈HSL, the docking calculations for the 34 AHL analogues and the receptor in

Table 1. Structures and IC₅₀ values of AHL analogues.

Group	2D Structure	ID	Substituent (X)	IC ₅₀ (μM)
AHL		4	-	0.69
	A3	-	0.77
	A4	-	1.05
Other analogues		5	-	0.83
	6	-	1.12
	A9 n=1	-	0.61
		A11 n=3	-	0.83
		A12 n=4	-	3.49
	D6	-	0.57
	D17	-	1.40
PHL		E12		4.2
		8	4-Br	4.73
		C10	4-I	1.25
		C13	3-NO ₃	2.25
		C20	4-CF ₃	0.81
		E6	3-SCH ₃	8.4
		E8	3-SCF ₃	4.7
		E11	3,4-Cl	2.4
POHL		D15	4-OCF ₃	0.46
		E15	4-CH ₃	1.8
		E16	3-CH ₃	1.1
		E17	4-OCH ₃	4.3
		E19	4-F	2.7
		E20	4-Cl	0.62
		E21	4-Br	0.51
		E22	4-I	0.44
		E23	4-CF ₃	1.6
		E24	4-NO ₂	0.29
		B7	4-Br	0.92
PPHL		E30	4-Cl	3.3
		E32	4-I	2.1
		E34	4-SCH ₃	1.1
		E35	4-CF ₃	0.21
		E37	4-NO ₂	1.6

AHL, acyl homoserine lactone; PHL, phenylacetyl homoserine lactone; POHL, phenoxyacetyl homoserine lactone; PPHL, phenylpropionyl homoserine lactone. Structures and IC₅₀ values for each analogue were obtained from Geske *et al.* 2007¹¹.

question were continued following the previous protocol, with the only difference being that 200 conformations were recorded for each analogue, of which the best were chosen according to their binding energy values and their position within the binding site of the TraR protein. The interactions of the selected protein-ligand complexes were analyzed on the Protein-Ligand Interaction Profiler webserver²².

Construction of 3D-QSAR models

Two 3D-QSAR models were constructed from structural alignments. The first model was constructed by taking the structure of the ligand, with the best inhibitory activity, reported by Geske *et al.* 2007¹¹, as a template. The other model was constructed based on the 34 best-selected poses of each analogue-TraR molecular docking²³, which generated an automatic alignment²⁴. The comparative molecular field analysis (CoMFA)²⁵ descriptor was obtained by using the QSAR tool implemented in Sybyl X 2.1.1. Steric and electrostatic field energies were calculated by using the Lennard-Jones and Coulomb potentials, respectively, with a dielectric constant 1/r based on the distance of all interactions. To calculate the steric and electrostatic energies, we used the force field of Sybyl X.2.1.1, Tripos²⁶. As an alternative, the open-source program, Open3DQSAR could be used²⁷.

For the evaluation of the 3D-QSAR models obtained, the cross-validation technique was used with the 'leave-one-out' method. With this validation, the predictive power of the models was determined and, in this way, the best model was determined. An exploratory analysis of the correlation between molecular fields and biological activity was also performed using PLS regression; thus, the correlation coefficients (R^2) of each model were obtained using the SAMPLS algorithm of Sybyl X 2.1.1. Finally, the cross-validation was done, giving the values of Q^2 and R^2_{cv} . According to these values both models were compared, and it was deduced which had the best predictability.

Results and discussion

Docking validation

In order to guarantee the reliability of the protocol used for molecular docking among analogues of AHL and the TraR receptor, validation of the search algorithm and scoring function combination was performed by re-docking the native ligand to the binding site of the TraR receptor using the molecular docking approach. From the large number of conformations generated, the conformations that best reproduced the position and interactions of the native ligand within the binding site were chosen. This conformation was compared with the co-crystallized structure of the same native ligand, as is depicted in Figure 1. In the alignment of the conformations of the native ligand in (green color) and the resulting docking (pink color), it can be seen that the coupled structure located at the TraR binding site adopts a conformation almost identical to the co-crystallized structure of the native ligand and in this way, both the lactone ring and the acyl side chain is located suitably at the receptor binding site.

The mean quadratic root deviation (RMSD) value calculated by AutoDock was used to confirm whether the chosen conformation was correct or not. The RMSD value of the chosen conformation is 0.56 Å; generally, an RMSD value of less than 1.5 or 2 Å is considered to be optimal²⁸. In Figure 2, the interactions established in the conformation resulting from the molecular docking are visualized. In this ligand-receptor complex, the majority of hydrogen bond interactions that contribute to the binding of the native ligand and TraR are reproduced, as well as the correct folding of the TraR receptor²⁹. These interactions include hydrogen bonding between the Trp 57 residue and the carbonyl group of the lactone ring, hydrogen bonding between Asp 70 and the amino group and a bond between the second carbonyl group and Thr 129, with the exception of the bond between the Tyr 53 residue and the first carbonyl group located in the side acyl chain. Hydrophobic

Figure 1. Alignment between the ligand present in the co-crystallized complex in green and the coupled ligand in pink, using the Autodock 4.2 program.

Figure 2. Interactions present in the protein-native ligand complex.

residues also contribute to the stable binding of the native ligand with TraR by hydrophobic interactions with residues Leu 40, Tyr 53, Tyr 61, Val 72, Trp 85. The interactions with Thr 51, Phe 62 and Ile 102 do not occur in said conformation.

Modes of binding among analogues and TraR

The structural interactions of this group of molecules were classified into five groups according to the differences present in the acyl chain reported by Geske *et al.*^{11,12}. This classification was made taking into account the importance of the influence of the variation of the acyl chain on the interactions that form the 34 analogues of the AHL within the TraR binding site. In addition to the classification by groups, the individual structure of each analogue was divided into two regions, A and B, as is shown in Figure 3, the first consisting of the lactone ring and the carbonyl group and the second region comprising the acyl chain, which is the one that varies between the different analogues. In this way, the analysis of the interactions is facilitated since each analog tends to form mostly the same interactions in the region A, unlike those formed in region B, which tend to be different³⁰.

For each analogue-TraR docking, two or more clusters were generated. In these clusters, the different ways of binding can be identified. From the analysis of the clusters of each analogue-TraR docking, the best 34 conformations were chosen using the orientations that reported the lowest binding energy as the criterion of choice. The analogue-TraR complexes chosen, in comparison to 3OC₈HSL, are, in general, similarly oriented at the TraR receptor binding site. The structure of these analogues was constructed based on the structure of homoserine lactone, with variations in the acyl chain of the native ligand. Table 2 shows the IC₅₀ values and the free energies calculated using AutoDock for all 34 ligands.

Ligands 4, D6, C10, E22 and E35 have the highest values of inhibitory activity (IC₅₀), so the analysis of their structure-activity relationship for each group was taken as precedents. Analogues 4, C10, E22 and E35 form two hydrogen bonds with the TraR receptor; one between the carbonyl group of the lactone ring and the amino group in the Trp 57 residue and the second between the carbonyl group of the first carbon of the acyl chain and the residue Thr 129. Ligand D6 also forms these two hydrogen bonds, but also forms another with the Asp 70 residue and the carbonyl group of the acyl chain. The native ligand forms the same three hydrogen bonds mentioned that contribute to the bond at the TraR protein binding site. Similarly, the remaining ligands form these bonds but with the absence of one or two of them and differences in bond distances. The interactions of ligands D6 and E35 are detailed in Figure 4.

On the other hand, the possible hydrophobic interactions formed by these ligands in the hydrophobic pocket of TraR involve both the lactone ring and the acyl chain. Structurally, ligand 4 is classified within the group of non-native AHLs with an acyl chain of seven carbon atoms; one less than the native ligand 3OC₈HSL. This side chain allows the formation of eight possible hydrophobic interactions with Ala38, Leu40, Gln58, Phe62 and two interactions with each residue of Tyr61 and 63. In addition, in the lactone ring are the interactions with Val72 and Trp85, for a total of ten interactions. The D6 ligand has in its structure a side chain a carbonyl group on carbon 4 and a toluyl group. This structural modification places it within the group of other analogues, and between this chain and the protein it is possible to form 11 hydrophobic interactions; eight with the same residues as ligand 4 and also those formed with Ala49, Thr51 and Ala168. The C10 ligand has in its side chain a phenylacetanoyl group and E22 a phenoxyacetanoyl group, which places them in the phenylacetyl homoserine lactones and

Figure 3. Sketch of the 34 acyl homoserine lactones analogues. Region A is formed by the lactone ring and the amide function and region B has variable side chains.

Table 2. Half maximal inhibitory concentration (IC_{50}) values and the free energy of docking of the analogues.

Analogue	IC_{50} (mM)	Energy AutoDock (kcal/mol)	Analogue	IC_{50} (mM)	Energy AutoDock (kcal/mol)
4	0.69	-8.87	E8	4.70	-10.13
5	0.83	-9.73	E11	2.40	-9.25
6	1.12	-9.98	E12	4.20	-10.04
8	4.73	-8.52	E15	1.80	-8.81
A3	0.77	-9.41	E16	1.10	-8.96
A4	1.05	-10.65	E17	4.30	-8.63
A9	0.61	-8.54	E19	2.70	-8.27
A11	0.83	-9.60	E20	0.62	-7.73
A12	3.49	-10.21	E21	0.51	-8.82
B7	0.92	-10.34	E22	0.44	-9.50
C10	1.25	-9.26	E23	1.60	-9.79
C13	2.25	-10.11	E24	0.29	-9.44
C20	0.81	-10.41	E30	3.30	-9.14
D6	0.57	-10.37	E32	2.10	-10.41
D15	0.46	-9.17	E34	1.10	-9.93
D17	1.40	-10.28	E35	0.21	-10.68
E6	8.4	-8.94	E37	1.60	-10.09

Figure 4. Interactions among ligands D6, E35 and TraR. Hydrogen bonds are shown in blue, hydrophobic interactions in gray, π - π stacking in green and halogen bonds in cyan.

phenoxyacetyl homoserine lactones groups, respectively. Both form six hydrophobic interactions; four in the side chain with Leu40 and Tyr53 with different distances of connection, and two in the lactone ring with Val 72 and Trp85. Finally, of the 34 analogues, ligand E35 reported the highest inhibitory activity against TraR. It is classified in the phenylpropionyl homoserine lactones group since it has the phenylpropionyl group in its side chain and its possible hydrophobic interactions are the same as those mentioned for C10 and E22.

The differences among the possible hydrophobic interactions mentioned above for the different ligands representing each group and the interactions of the native ligand in the hydrophobic pocket of TraR are notable. However, these are not the only differences; some analogue-TraR complexes reported other possible interactions, such as π - π stacking, which are favored in ligands that in their structure possess an aromatic group, as can be seen in the complexes of ligands D6, C10, E22 and E35. The formation of this interaction occurs between the benzene rings of these molecules and the Tyr 61 residue. This type of interaction is also reported in other ligands, but with the amino acid Tyr53.

On the other hand, analogues that have halogen atoms in their side chains can form other types of interactions called halogen bonds. In this study, ligands C10, D15, E21, E23, B7, and E35 have halogen bond interactions with the Gln 58 residue and the E23 ligand with Trp 57. Studies of the characteristics of this type of halogen bond establish that they are strongly directional and the distance between the halogen and the nearby electronegative atom should ideally be less than the sum of its Van der Waals radii. Moreover, although the fluorine atom is a halogen, it does not have the capacity to form these interactions due to its great electronegativity³¹. According to the above, the halogen bonds of ligands D15, E23 and E35 are discarded as they are not possible since the interaction occurs between a fluorine atom and an oxygen atom. The sum of the van der

Waals radii for the remaining ligands C10, E21 and B7 are calculated as follows: in the C10 ligand where its iodine atom and the Gln 58 oxygen interact, this value is 3.50 Å; in the E21 ligand where its bromine atom and the Gln 58 oxygen interact, the sum of the radii is 3.37 Å; and in ligand B7 where its bromine atom and oxygen of Gln 58 interact, the sum of the radii is 3.37 Å. The respective distances between the atoms bound for each ligand-receptor complex are 3.90 Å, 2.14 Å and 2.06 Å. Comparing these values, it follows that the only halogen bonds that are shorter than the value of the sum of the radii of the bound atoms are those of ligands E21 and B7 with the amino acid Gln 58.

It is quite clear that the addition of certain groups such as phenyl and halogenated substituents can generate other interactions, which could confer changes in the increase or decrease in the inhibitory activity of these analogues against the TraR receptor. These extra interactions (halogen bond, π - π stacking), formed by the analogues and the TraR receptor, are not present when TraR is in complex with the native ligand.

Analysis of the 3D-QSAR structure-activity quantitative relationship

For ligand-based alignment, ligand E35 was used as a template since it was the ligand with the highest inhibitory activity against TraR and for the receptor-based alignment, the best poses generated in the molecular docking, that is, those that reported the lowest binding energy and the correct orientation in the active site of the receptor, were used. Figure 5A and 5B show the superposition of all ligands for the ligand-based model (Model 2) and receptor-based model (Model 1), respectively.

The correlation analysis between CoMFA values and *in vitro* biological activities of the 34 analogues was performed by PLS and the evaluation of the predictive capacity of the models generated through the leave-one-out cross-validation method. Q^2 values of 0.457 and R^2 are obtained with cross validation of

0.856 for the Model 2, and a Q^2 of 0.697 and R^2 with cross validation of 0.781 for the Model 1. In effect, Model 1 shows a lower correlation so that their statistical confidence is lower than in model 2. Table 3 shows the other results of both PLS and cross-validation analyzes.

Figure 6 shows correlation of models 1 and 2, it can be seen that in model 2 there is a better pattern of adjustment of the points along the trend line, which translates into the prediction of biological activity values closer to those reported experimentally, unlike the graph for Model 2 where the residual dispersion is greater. From these comparisons, the Model 2 is statistically profiled as the one with the best correlation and the greatest predictive power. Even so, the Model 1 should not be discarded altogether since it also takes into account the structural

similarity between the analogues and the native ligand; the position and orientation taken by the ligands at the receptor binding site according to the conformational freedom allowed by the amino acid residues that comprise it.

Through the interpretation of the contour maps obtained by CoMFA, it is possible to know which structural characteristics of the ligands correlate with the changes in their biological activity values. These contours are represented by colors; steric contour maps are represented in green (region favorable for bulky groups) and yellow (region not favorable for bulky groups), while those of red (region favorable for electronegative substituents) and blue (region favorable for electropositive substituents) correspond to electrostatic contours³².

Figure 5. Structural alignment. **A)** Model 1 and **B)** Model 2.

Table 3. Statistical data for comparative molecular field analysis (CoMFA) models.

Statistical Results	CoMFA receptor-based model	CoMFA ligand-based model
Q^2	0.697	0.457
R^2	0.781	0.856
SEE	0.194	0.158
Probability of $R^2=0$	0	0
F	16.387	26.854
Components	6	6
Steric fraction (%)	51.700	68.300
Electrostatic fraction (%)	48.300	31.700

SEE, standard error of estimate.

In **Figure 7** the contour maps of the CoMFA models calculated are shown, taking as an illustrative example those generated for ligand D15, which records one of the highest values of inhibitory activity against TraR in the phenoxyacetyl homoserine lactones group. In these contours, it can be observed that the sterically favored regions are in positions close to the fluorine atoms found in the benzene ring substituent in the ligand structure in both **Figure 7A and 7B**. This is because this region has high electronegativity, which favors interactions with residues that have partially positive charges in their structures, while the sterically unfavorable regions, in yellow, spread around the phenyl group. However, on most of the hydrogen atoms in the phenyl ring, the positively charged regions that favor the electrostatic field are identified, as can be seen in more detail in **Figure 7B**. In the same way, carbonyl and ether groups with a negative charge are seen as electrostatically favorable regions due to the negative potential of oxygen atoms. In addition, the

existence of halogens promotes the formation of halogen bonds with residues containing oxygen atoms in their structure. All these groups mentioned are identified as regions that would contribute to the antagonistic activity of ligand D15 against TraR.

Conclusions

In most of the 34 ligand-receptor complexes obtained by molecular docking, the formation of hydrogen bonds and similar hydrophobic interactions with the same amino acid residues were observed. The most common hydrogen bonds are established with Trp 57, Thr 129 and Asp 70. The most recurrent hydrophobic interactions are formed with Ala 38, Leu 40, Tyr 53, Tyr 61, Val 72, Trp 85. Certain structural modifications such as the presence of aromatic rings and halogen atoms allow the formation of other interactions such as molecular type π - π stacking and halogen bonds. From the above, it is possible to affirm that there is a marked relationship between the type of interactions and the

Figure 6. Predicted vs experimental activity scatter plot obtained by partial least squares analysis using comparative molecular field analysis (CoMFA).

Figure 7. 3D comparative molecular field analysis (CoMFA) contour maps for ligand 5. **A)** receptor-based model and **B)** ligand-based model. Steric field: green: sterically favorable; yellow: sterically unfavorable. Electrostatic field: red: negative potential favorable; blue: negative potential unfavorable.

activity of the ligands, where the differences with those formed by the native ligand may improve the inhibitory activity of the analogues against TraR.

The PLS statistical analysis of each model and subsequent internal validation allowed comparison of the predictive capacity under parameters such as R^2 and Q^2 , analyzing how the predicted inhibitory activity values vary when the experimental biological activities are correlated with the properties of the optimized structures of the ligands (model based on the ligand) and, for the model based on the receptor, when the same experimental activities are correlated with the properties of the poses that the ligands adopt in the molecular docking. This resulted in the best model in statistical terms being the model based on the ligand. Even so, the receptor-based model, since it is based on the simulations of the analogue receptor binding, provides information based on the position and orientation that the analogues can adopt according to the chemical nature of the TraR receptor's active site, unlike the model based on the ligand, which is based solely on optimized structures.

Regarding the study of the relationship of the structural characteristics that influence the values of biological activity, the 3D-QSAR-CoMFA models suggest that it is possible to improve the inhibitory activity of new ligands by adding groups with

positive and negative charges of favorable electrostatic potentials such as carbonyl groups, ether groups, halogens, and in some cases the nitro substituents and the phenyl group, which favors the formation of hydrophobic interactions. Bulky groups attached to the phenyl ring and other groups such as sulfur are identified as unfavorable regions for the design of new analogues with greater inhibitory activity against TraR.

Data availability

Underlying data

Figshare: Receptor_Alum_2019_Pose_E_D_1.mdb (1).zip. <https://doi.org/10.6084/m9.figshare.10255739.v1>²³.

This project contains the following underlying data:

- .mol2 files (templates used for alignment in MOL2 format)

Data are available under the terms of the [Creative Commons Attribution 4.0 International license](#) (CC-BY 4.0).

Acknowledgements

MAM and RVR thank the Universidad de Cartagena (Cartagena, Colombia) and Fundación Universitaria Tecnológico de Comfenalco for continuous support to the group.

References

- Kuttler C, Hense BA: **Interplay of two quorum sensing regulation systems of *Vibrio fischeri***. *J Theor Biol*. 2008; **251**(1): 167–180.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Barreto AC: **Quorum Sensing: Sistemas de comunicación bacteriana**. *Ciencia Actual*. 2013; 5–6.
- Lang J, Faure J: **Functions and regulation of quorum-sensing in *Agrobacterium tumefaciens***. *Front Plant Sci*. 2014; **5**: 14.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Mattmann ME, Blackwell HE: **Small molecules that modulate quorum sensing and control virulence in *Pseudomonas aeruginosa***. *J Org Chem*. 2010; **75**(20): 6737–6746.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Waters CM, Bassler BL: **Quorum sensing: cell-to-cell communication in bacteria**. *Annu Rev Cell Dev Biol*. 2005; **21**: 319–346.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Pedroza CJ: **Caracterización de la enzima acil homoserina lactonasa de una cepa de *Bacillus thuringiensis***. Medellín: Universidad nacional de Colombia. 2010.
[Reference Source](#)
- Alippi A, Lopez A, Balatti P: **Métodos para la detección de *Agrobacterium* a partir de muestras de material vegetal, suelo y agua**. *Revista Argentina de Microbiología*. 2011; **43**(4): 278–286.
[Reference Source](#)
- Lang J, Faure D: **Functions and regulation of quorum-sensing in *Agrobacterium tumefaciens***. *Front Plant Sci*. 2014; **5**: 14.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Dong YH, Wang LH, Zhang LH: **Quorum-quenching microbial infections: mechanisms and implications**. *Philos Trans R Soc Lond B Biol Sci*. 2007; **362**(1483): 1201–1211.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Ortero MR: **Revista real academia Galega de ciencias**. 2010; **XXIX**: 6–7.
[Reference Source](#)
- Geske GD, O'Neill JC, Miller DM, et al.: **Modulation of Bacterial Quorum Sensing with Synthetic Ligands: Systematic Evaluation of *N*-Acyated Homoserine Lactones in Multiple Species and New Insights into Their Mechanisms of Action**. *J Am Chem Soc*. 2007; **129**(44): 13613–13625.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Geske GD, Mattmann ME, Blackwell HE: **Evaluation of a focused library of *N*-aryl L-homoserine lactones reveals a new set of potent quorum sensing modulators**. *Bioorg Med Chem Lett*. 2008; **18**(22): 5978–5981.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Godbout N, Salahub DR, Andzelm J, et al.: **Optimization of Gaussian-type basis sets for local spin density functional calculations. Part I. Boron through neon, optimization technique and validation**. *Can J Chem*. 1992; **70**(2): 560–571.
[Publisher Full Text](#)
- Xu X, Zhang Q, Muller RP, et al.: **An extended hybrid density functional (X3LYP) with improved descriptions of nonbond interactions and thermodynamic properties of molecular systems**. *J Chem Phys*. 2005; **122**(1): 14105.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Frisch MJ, Trucks GW, Schlegel HB, et al.: **Gaussian 09**. Wallingford CT: Gaussian, Inc. 2009.
[Reference Source](#)
- Gordon MS, Schmidt MW: **Advances in electronic structure theory: GAMESS a decade later**. In *Theory and Applications of Computational Chemistry: the first forty years*. Dykstra CE, Frenking G, Kim KS, Scuseria GE, (editors), Elsevier, Amsterdam, 2005; 1167–1189.
[Publisher Full Text](#)
- TRIPOS Associates, Inc: **Sybyl-X Molecular Modeling Software Packages, Version 2.0**. St. Louis, MO USA. 2012.
[Reference Source](#)
- Hanwell MD, Curtis DE, Lonie DC, et al.: **Avogadro: an advanced semantic chemical editor, visualization, and analysis platform**. *J Cheminform*. 2012; **4**(1): 17.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Morris GM, Goodsell DS, Pique ME, et al.: **AutoDock Version 4.2**. Citeseer: 2012; 1–6.
[Reference Source](#)
- Huey R, Morris G: **Using AutoDock 4 with AutoDockTools: A Tutorial**. The Scripps Research Institute, USA: 2008; 1–56.
- Morris GM, Huey R, Olson AJ: **Using AutoDock for ligand-receptor docking**. *Curr Protoc Bioinformatics*. 2008; **24**: Chapter 8:Unit 8.14.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Salentin S, Schreiber S, Haupt VJ, et al.: **PLIP: fully automated protein-ligand**

- interaction profiler. *Nucleic Acids Res.* 2015; **43**(W1): W443–W447.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
23. Vivas-Reyes R: **Receptor_Alum_2019_Pose_E_D_1.mdb (1).zip**. *figshare*. 2019.
<http://www.doi.org/10.6084/m9.figshare.10255739.v1>
 24. Sippl W: **Development of biologically active compounds by combining 3D QSAR and structure-based design methods**. *J Comput Aided Mol Des.* 2002; **16**(11): 825–830.
[PubMed Abstract](#) | [Publisher Full Text](#)
 25. Cramer RD, Patterson DE, Bunce JD: **Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins**. *J Am Chem Soc.* 1988; **110**(18): 5959–5967.
[PubMed Abstract](#) | [Publisher Full Text](#)
 26. Clark M, Cramer RD, Van Opdenbosch N: **Validation of the general purpose tripos 5.2 force field**. *J Comput Chem.* 1989; **10**(8): 982–1012.
[Publisher Full Text](#)
 27. Tosco P, Balle T: **Open3DQSAR: a new open-source software aimed at high-throughput chemometric analysis of molecular interaction fields**. *J Mol Model.* 2011; **17**(1): 201–208.
[PubMed Abstract](#) | [Publisher Full Text](#)
 28. Hevener KE, Zhao W, Ball DM, *et al.*: **Validation of molecular docking programs for virtual screening against dihydropteroate synthase**. *J Chem Inf Model.* 2009; **49**(2): 444–460.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 29. Zhang RG, Pappas KM, Brace JL, *et al.*: **Structure of a bacterial quorum-sensing transcription factor complexed with pheromone and DNA**. *Nature.* 2002; **417**(6892): 971–974.
[PubMed Abstract](#) | [Publisher Full Text](#)
 30. Ahumado M, Drosos JC, Vivas-Reyes R: **Application of molecular docking and ONIOM methods for the description of interactions between anti-quorum sensing active (AHL) analogues and the *Pseudomonas aeruginosa* LasR binding site**. *Mol Biosyst.* 2014; **10**(5): 1162–1171.
[PubMed Abstract](#) | [Publisher Full Text](#)
 31. Saavedra LH: **Entendiendo los puentes de halógeno desde un punto de vista teórico**. Santiago de Chile: Universidad de Chile. 2013.
[Reference Source](#)
 32. Nair PC, Sobhia ME: **CoMFA based *de novo* design of pyridazine analogs as PTP1B inhibitors**. *J Mol Graph Model.* 2007; **26**(1): 117–123.
[PubMed Abstract](#) | [Publisher Full Text](#)

Open Peer Review

Current Peer Review Status: ? ?

Version 1

Reviewer Report 10 March 2020

<https://doi.org/10.5256/f1000research.22871.r58098>

© 2020 Pacheco-Londoño L. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The author(s) is/are employees of the US Government and therefore domestic copyright protection in USA does not apply to this work. The work may be protected under the copyright laws of other jurisdictions when used in those jurisdictions.

Leonardo C. Pacheco-Londoño

¹ R3-C Research and Education Component of ALERT DHS Center of Excellence for Explosives Research, Department of Chemistry, University of Puerto Rico, Mayaguez, PR, Puerto Rico

² School of Basic and Biomedical Science, Universidad Simón Bolívar, Barranquilla, Colombia

The authors try to explain the interactions among structural analogues of acyl homoserine lactones (AHLs) using molecular modeling and support this by a statistical correlation of biological activity with molecular descriptors. The problem of the article is the type of linear model used for the statistical correlation. PLS is a type of correlation and this generates multiple linear correlations. These are known as latent variables and these are the variables that are not directly observed but are inferred (through a mathematical model) from other observed variables or (molecular descriptors). In this case, this type of model is viable when you have many samples (molecules) and many independent variables (molecular descriptors), and this is not the case. In my opinion for the article to be published the authors must include the following:

1. You must indicate how many descriptors are used in the models.
2. They use 34 molecules; they should remember that the number of degrees of freedom (g) for PLS models would be $34 - \# \text{ LV} - \# \text{ descriptors} - 1$.
The authors do not report the number of latent variables (LV). If the g is minimal, the PLS model must be discarded.
3. The authors must realize an analysis of the LV. This should be showing the weights or loading for each LV, and thus indicate which descriptors are contributing to each LV and be able to explain: (1) why the correlations should be independent based on the weights of the descriptors in the LVs and (2) why these descriptors do not work well in a single multiple linear equations and it is necessary to generate several linear combinations (LV) to find a good correlation.
4. If the model has only one descriptor or two, the PLS process of adding LV to improve the

correlation generates what is known as overfitting though that a leave-one-out cross-validation is being used. They should better perform a multiple regression for this case.

5. Most likely, the authors have already tried multiple regression and found no correlation. If this is the case, there is a high probability that there is no true correlation between the data of the descriptors obtained and the biological activity. I invite you to improve the calculation of these descriptors.
6. The authors must include an external validation in addition to the cross-validation

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Partly

If applicable, is the statistical analysis and its interpretation appropriate?

Partly

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Chemometrics, intelligence artificial, sensor and biosensor design, material science, vibrational spectroscopy, and ab-initio calculation

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 24 January 2020

<https://doi.org/10.5256/f1000research.22871.r58629>

© 2020 Cubillan N. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Nestor Cubillan

Chemistry Program, Faculty of Basic Sciences, University of Atlántico, Barranquilla, Colombia

In this work, the authors proposed a dataset for studying interaction of a set of 34 molecules with the *Agrobacterium tumefaciens* TraR binding site through Molecular Docking and 3D QSAR. The work is well written and readable. I have an unique concern about the ligand-based ComFA model results in table 3. The model 1 presents an R2 of 0.781 while Q2 is 0.697 revealing good robustness of this Model. However, Model 2 has an R2 of 0.856 and Q2 of 0.457. This result is evidence of the presence of data leveraging the model. Can the authors interpret this behaviour?

An explanation of this result can be related to noisy variables in predictors overfitting the model. Is this a possible explanation of these results?

I considered this work suitable for publication after making these minimum corrections.

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Partly

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Computational Chemistry, Mathematical Chemistry, Chemometrics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 07 Feb 2020

Ricardo Vivas-Reyes

Dear Sir,

This is the answer to the reviewer, where we give the reason for the two models used in this study. Also waiting for the referee to agree with our response. Thank you in advance for your time and dedication.

The idea of the two models was realized to compare the predictive capacity of both models. So, a good value for Q^2 is a value close to R^2 . That means that your PLS model works independently of the specific data that was used to train the PLS model, which is the case of model 1, while inconsistencies occur in the model.

Thanks in advance

Best Regards,

Ricardo

Competing Interests: Not competing interest

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research