



UNIVERSIDAD
NACIONAL
DE COLOMBIA

Contribución a la identificación de potenciales antagonistas asociados al *quorum sensing* en *Pseudomonas aeruginosa* a través de métodos quimioinformáticos

Daniel Felipe Victoria Muñoz

Universidad Nacional de Colombia
Facultad de Ciencias, Departamento de Farmacia
Bogotá D.C., Colombia

2021

Contribución a la identificación de potenciales antagonistas asociados al *quorum sensing* en *Pseudomonas aeruginosa* a través de métodos quimioinformáticos

Daniel Felipe Victoria Muñoz

Tesis presentada como requisito parcial para optar al título de:

Magister en Ciencias Farmacéuticas

Director (a):

Fabian López Vallejo

Línea de Investigación:

Química medicinal y modelamiento molecular

Grupo de Investigación:

Química de Productos Naturales Vegetales Bioactivos (QuiProNaB)

Molecular Design and Modelling Laboratory (MDML)

Universidad Nacional de Colombia

Facultad de Ciencias, Departamento de Farmacia

Bogotá D.C., Colombia

2021

Contribution to the identification of potential antagonists associated with quorum sensing in *Pseudomonas aeruginosa* through chemoinformatic methods

Daniel Felipe Victoria Muñoz

A thesis presented as partial fulfillment of the requirement for the degree of:
Master of Science in Pharmaceutical Sciences

Advisor:

PhD. Fabian López Vallejo

Research line:

Medicinal Chemistry and Molecular Modelling

Research groups:

Química de Productos Naturales Vegetales Bioactivos (QuiProNaB)

Molecular Design and Modelling Laboratory (MDML)

Universidad Nacional de Colombia

Facultad de Ciencias, Departamento de Farmacia

Bogotá D.C., Colombia

2021

It always seems impossible until it's done

Nelson Mandela

Agradecimientos

A la Universidad Nacional de Colombia, de la cual he recibido toda mi formación profesional en estos últimos diez años. Especialmente a los departamentos de Farmacia y Química por ser los lugares donde descubrí mi amor por la ciencia, el trabajo y la investigación.

A la Facultad de Ciencias de la Universidad Nacional, por la Beca Auxiliar Docente durante todo mi proceso de formación en la maestría.

A la fundación para la promoción de la investigación y la tecnología (FPIT) por la financiación del proyecto 4.553.

A el profesor José Luis Medina-Franco de la Universidad Nacional Autónoma de México (UNAM) en Ciudad de México, México, quien es uno de los mayores exponentes del diseño de fármacos asistido por computadora, al cual tuve la fortuna de conocer. De su mano entendí el mundo de la investigación en esta área y me ayudo a plantear mejor mis proyectos de investigación, así mismo la estancia realizada en su grupo de investigación DIFACQUIM me permitió conocer personas maravillosas que comparten mi gran pasión por la investigación y las computadoras, en donde se destaca el Dr. Norberto Sánchez-Cruz, mi gran amigo y maestro en ramas de programación relacionadas con ciencias químicas y biológicas, además de ser uno de los mejores guías de la CDMX.

A mi madre Edilmita, a mi hermano el enano y a la que siempre me espero en casa Lulu, por su apoyo incondicional siempre, por permitirme soñar y desarrollar mis locos sueños de la academia e investigación; sin este apoyo muy probablemente no habría podido llegar a este punto. Ellos saben que gran parte de esta tesis es de ellos y que los retos siempre serán más fáciles si estamos juntos.

Al profesor Luis Enrique Cuca, por la oportunidad de trabajar en su grupo de investigación; donde encontré personas maravillosas que me permitieron entrar a sus vidas y encontrar un gran apoyo en la investigación y en la rumba.

A los miembros del grupo QuiProNaB, que nunca permitieron un día de investigación aburrido, entre los que recuerdo con mucho aprecio: a Daniela Torres 'Alias la jefe', a Juan Cardozo, a Alberto Osorio,

a Laura López, a Jeisson Martínez, a Luis Fernando y a Erika Plazas, doctora y gran investigadora de la cual pude aprender mucho.

Al profesor Fabian López, por su confianza, orientación y ayuda; la cual me permitió lograr todos los objetivos de esta maestría; con su apoyo pude llegar a conocer otros países, grupos y formas de investigar. Adicionalmente a lo mencionado me brindo toda la libertad creativa para encontrar la mejor manera de resolver los problemas a los cuales me enfrente con esta investigación y consciente de mis falencias o errores siempre busco la mejor manera para tratar de subsanarlos. Espero siempre poder contar con su amistad y experiencia; aunque él sabe que siempre contara con la mía.

A Andrés, Diego, Miguel, Yesid compañeros de colegio y de vida. A mis grandes amigos Daniel, Iván, Sebastián y Jessi; con los que pude emprender este camino de ser QF y a los cuales siempre espero tener a mi lado.

En pocas palabras gracias a todos los que intervinieron en el desarrollo de esta investigación.

Resumen

De acuerdo con el centro de control de enfermedades de los Estados Unidos (CDC's) en su reporte de resistencia a antibióticos, las infecciones causadas por *P. aeruginosa* multi-fármaco resistente comprometieron alrededor de 32,600 pacientes y generaron 2,700 muertes durante el 2017. En Colombia el Sistema Nacional de Vigilancia de Resistencia Antimicrobiana (IAAS) en su informe del 2018 mostró que la *P. aeruginosa* es el patógeno con mayor recurrencia de la metalo β -lactamasa VIM en el territorio colombiano, alcanzando 840 casos contabilizados y dificultando el tratamiento de las infecciones de este patógeno.

Como parte de la búsqueda de nuevas soluciones contra este patógeno, el mecanismo de comunicación intercelular denominado *quorum sensing* y los factores de transcripción que intervienen en él resultan ser blancos promisorios como adyuvantes en el tratamiento de infecciones por este patógeno. Por esta razón, en la presente investigación se seleccionaron los factores de transcripción LasR, PqsR y RhlR como blancos de estudio en el control de esta vía de comunicación.

A través de la construcción de bases de datos con moléculas a las que se les ha demostrado su actividad biológica y su análisis mediante diversas herramientas del diseño de fármacos basado en el ligando, se encontraron múltiples características estructurales involucradas con la actividad contra estos factores de transcripción.

Para LasR, anillos de 5 y 6 miembros (lactónicos y fenílicos), cadenas carbonadas de 10 o más carbonos, dos aceptores y un donador de puente de hidrogeno y el tamaño molecular superior a los 15 átomos fueron las características estructurales mínimas asociadas a moléculas con actividad sobre este factor. Adicionalmente se identificaron las interacciones con los residuos Tyr 47, ala 105, Thr 80, Leu 36, Ile 36, Ile 52 y Tyr 56 como posibles responsables de la actividad antagonista utilizando herramientas como *fingerprints* de interacción proteína-ligando y simulación de dinámicas molecular; siendo estos asociados al diseño de fármacos basado en la estructura.

Los compuestos con actividad sobre PqsR exhibieron las siguientes características estructurales: bicíclos nitrogenados de 6 miembros (quinoleína y quinazolina), cadenas alifáticas de un tamaño

mínimo de 10 carbonos, mínimo un aceptor y donador de puente de hidrogeno y se debe igualar o superar a los 15 átomos; las interacciones con los aminoácidos que podrían estar involucrados en la actividad antagonista sobre este factor son: Tyr 258, Phe 221, Ala 102, Leu 207, Ile 149, siendo estas caracterizadas con los instrumentos del diseño e fármacos basado en la estructura.

A partir de la base de datos construida para RhIR y los análisis quimioinformáticos se detectó que las moléculas con un anillo de 5 o 6 miembros (lactonas y fenilos), cadenas carbonadas de 3 a 6 carbonos; al menos 2 aceptores de puente de hidrogeno, 1 donador de puente de hidrogeno y que el conteo de átomos oscile entre 8 y 12. Como en las bases de datos públicas al momento no se encuentra elucidada la estructura terciaria de este factor de transcripción, en este trabajo se planteó y validó un modelo de estructura: con la anterior se procedió a hacer los análisis de *fingerprints* interacción proteína-ligando y la simulación de dinámica molecular para encontrar los residuos involucrados con la actividad antagonista, encontrado a las interacciones con la Val 60, Leu 69, Tyr 64, Tyr 72 y Ala 83 como las posibles responsables.

Con el fin de complementar las características estructurales encontradas anteriormente, se crearon los farmacóforos a través del alineamiento estructural de las conformaciones obtenidas con los acoplamientos de las 10 moléculas con mejor actividad antagonista sobre cada uno de los factores de transcripción. Adicionalmente, se construyeron modelos de aprendizaje de maquina supervisados, donde los compuestos de la base de datos fueron clasificados como activos e inactivos y con el *molecular fingerprint* ECFP4 se caracterizó cada una de las moléculas para el entrenamiento posterior de 14 modelos; los cuales fueron clasificados según 6 métricas diferentes. De lo anterior se encontró que los mejores clasificadores fueron *extreme gradient boosting* y *k-nearest neighbors classifier* para todos los factores de transcripción y se propone el uso conjunto para la clasificación de los compuestos con posible actividad.

Para integrar la información anterior, se realizó un cribado virtual sobre una base de datos molecular de productos naturales; encontrando a los compuestos: litseanolido C2, 1-hidroxi-3-metoxi-N-metil acridona y hematoma de metilo como los mejores clasificados para ser antagonistas contra LasR, PqsR y RhIR; además se identificó a las familias Rutaceae, Lauraceae y Piperaceae como una fuente potencial de compuestos con actividad contra estos factores de transcripción

Palabras clave: *Quorum sensing*, *P. aeruginosa*, factores de transcripción, LasR / PqsR / RhIR, quimioinformatica, *protein-ligand interaction fingerprints*, simulaciones de dinámica molecular, *machine learning*, diseño de fármacos asistido por computadora

Abstract

According to the United States Center for Disease Control (CDC's) in its report of antibiotic resistance, infections caused by multi-drug resistant *P. aeruginosa* affected around 32,600 patients and generated 2,700 deaths during 2017. In Colombia the National Antimicrobial Resistance Surveillance System (IAAS) in its 2018 report showed that *P. aeruginosa* is the pathogen with the highest recurrence of Metallo β -lactamase VIM in the Colombian territory, reaching 840 counted cases and making it difficult to treat the infections of this pathogen.

As part of the search for new solutions against this pathogen, the intercellular communication mechanism called quorum sensing and the transcription factors involved in it turn out to be promising targets as adjuvants in the treatment of infections caused by this pathogen. For this reason, in this research, the transcription factors LasR, PqsR, and RhIR were selected as study targets in the control of this communication pathway.

Through the construction of databases with active molecules. In this thesis using various tools of drug design based on the ligand, multiple structural characteristics involved with the activity against these transcription factors were found.

For LasR, 5 and 6-membered rings (lactonic and phenyl), carbon chains of 10 or more carbons, two acceptors and a hydrogen bridge donor, and molecular size greater than 15 atoms were the minimum structural characteristics associated with molecules with activity on this factor. Additionally, interactions with residues Tyr 47, ala 105, Thr 80, Leu 36, Ile 36, Ile 52, and Tyr 56 were identified as possibly responsible for the antagonist activity using tools such as protein-ligand interaction fingerprints and molecular dynamics simulation; these being associated with structure-based drug design.

Compounds with activity on PqsR exhibited the following structural characteristics: 6-membered nitrogenous bicycles (quinoline and quinazoline), aliphatic chains of a minimum size of 10 carbons, at least one hydrogen bridge acceptor and donor, and should equal or exceed those 15 atoms; the interactions with the amino acids that could be involved in the antagonist activity on this factor are:

Tyr 258, Phe 221, Ala 102, Leu 207, Ile 149, being these characterized with the instruments of design and drugs based on the structure.

From the database built for RhIR and the chemoinformatics analyzes, it was detected that molecules with a 5 or 6-membered ring (lactones and phenyl), carbon chains of 3 to 6 carbons; at least 2 hydrogen bridge acceptors, 1 hydrogen bridge donor, and that the atom count ranges between 8 and 12. As in the public databases at the moment the tertiary structure of this transcription factor is not elucidated. This work proposed and validated a structure model: with the previous one, we proceeded to do the analysis of protein-ligand interaction fingerprints and the simulation of molecular dynamics to find the residues involved with the antagonist activity, found in the interactions with Val 60, Leu 69, Tyr 64, Tyr 72 and Ala 83 as the possible culprits.

In order to complement the structural characteristics, found previously, the pharmacophores were created through the structural alignment of the conformations obtained with the couplings of the 10 molecules with the best antagonist activity on each of the transcription factors. Additionally, supervised machine learning models were built, where the compounds in the database were classified as active and inactive, and with the molecular fingerprint ECFP4, each of the molecules was characterized for the subsequent training of 14 models; which were classified according to 6 different metrics. From the above, it was found that the best classifiers were extreme gradient boosting and k-nearest **neighbor's** classifier for all transcription factors, and joint use is proposed for the classification of compounds with possible activity.

To integrate the above information, virtual screening was performed on a molecular database of natural products; finding the compounds: C2 litseanolide, 1-hydroxy-3-methoxy-N-methyl acridone, and methyl hematamate as the best classified to be antagonists against LasR, PqsR, and RhIR; Furthermore, the Rutaceae, Lauraceae, and Piperaceae families were identified as a potential source of compounds with activity against these transcription factors.

Keywords: Quorum sensing, *P. aeruginosa*, transcriptional factors, LasR / PqsR / RhIR, cheminformatics, protein-ligand interaction fingerprints, molecular dynamics simulation, machine learning, computer-aided drug design.

Contenido

	Pág.
Resumen.....	IX
Lista de figuras.....	XV
Lista de tablas.....	XIX
Lista de símbolos y abreviaturas	XX
Introducción	1
Hipótesis.....	3
Objetivos.....	4
1 Estado del arte	5
1.1 Infecciones causadas por <i>P. aeruginosa</i>	5
1.1.1 Prevalencia de las infecciones por <i>P. aeruginosa</i>	5
1.1.2 Resistencia a antibióticos por <i>P. aeruginosa</i>	7
1.1.3 Tratamiento de infecciones causadas por <i>P. aeruginosa</i>	10
1.2 <i>Quorum sensing</i> y factores de transcripción.....	12
1.2.1 Sistemas <i>Las</i> y <i>Rhl</i>	17
1.2.2 Sistema <i>Pqs</i>	17
1.2.3 Sistema <i>Iqs</i>	18
1.3 Diseño de fármacos asistido por computadora (DIFAC)	19
1.3.1 Diseño de fármacos basado en la estructura (DFBE)	20
1.3.2 Diseño de fármacos basado en el ligando (DFBL)	22
2 Metodología	25
2.1 Diseño basado en el ligando.....	25
2.1.1 Construcción y preparación de las bases de datos	25
2.1.2 Complejidad molecular	26
2.1.3 Análisis de núcleos base.....	26
2.1.4 Análisis de <i>activity cliffs</i>	27
2.1.5 Visualización del espacio químico.....	27
2.2 Diseño basado en la estructura.....	28
2.2.1 Alineamientos de las estructuras cristalográficas de los FT 's.....	28
2.2.2 Construcción del modelo de RHR.....	28
2.2.3 Acoplamiento moleculares automatizados.....	28
2.2.4 Simulaciones de dinámica molecular	30
2.3 Modelado farmacológico	31
2.4 Clasificadores en base a <i>machine learning</i>	31
2.5 Cribado virtual	33
3 Resultados.....	34

3.1	Diseño de fármacos basado en el ligando (DFBL)	34
3.1.1	Construcción de las bases de datos.....	34
3.1.2	Complejidad molecular	34
3.1.3	Análisis de núcleos base	36
3.1.4	Análisis de <i>activity cliffs</i>	39
3.1.5	Visualización del espacio químico	41
3.2	Diseño de fármacos basado en la estructura (DFBE)	44
3.2.1	Alineamientos de las estructuras cristalográficas de los FT 's.....	44
3.2.2	Construcción del modelo de RhIR	46
3.2.3	Acoplamiento molecular automatizado	48
3.2.4	Simulaciones de dinámica molecular	53
3.3	Modelado farmacofórico	64
3.4	Clasificadores de <i>machine learning</i>	65
3.5	Cribado virtual	73
4	Conclusiones y recomendaciones.....	79
	Bibliografía	131

Lista de figuras

Pág.

Figura 1. 1: Prevalencia de las neumonías causadas por <i>P. aeruginosa</i> en el estudio realizado por Aliberti et-al, a) neumonía adquirida en la comunidad asociada a <i>P. aeruginosa</i> , b) neumonía asociada a <i>P. aeruginosa</i> resistentes a antibióticos. Tomada de Aliberti et-al ⁴⁴	6
Figura 1. 2: Prevalencia de los brotes de infecciones por microorganismo durante el 2019 en Colombia. Adaptada del boletín epidemiológico semanal 07 del 2020, Instituto Nacional de Salud ⁴⁵	7
Figura 1. 3: Mecanismos de resistencia a antibióticos en <i>P. aeruginosa</i> : intrínseco, adquirido y adaptativo. Tomada de Rehm et-al ³⁶	9
Figura 1. 4: Distribución geográfica de casos de carbapenemasas en Colombia, septiembre 2012-2018. Adaptada del informe de resultados de vigilancia por laboratorio de resistencia antimicrobiana en infecciones asociadas a la atención en salud (IAAS) 2018 ⁴⁵	10
Figura 1. 5: Potenciales nuevos antibióticos posiblemente aprobados por la US-FDA, clasificados según su tasa de innovación, siendo community old: núcleos base usualmente usado en terapéutica, community new: núcleos base nuevos, NBTI's: nuevos inhibidores de topoisomerasa tipo II , BLI: inhibidores de β-lactamasa , FabI: inhibidores de FabI , específicos para estafilococos. Tomado de Theuretzbacher et al ⁵⁸	12
Figura 1. 6: Representación esquemática funcionamiento de los sistemas quorum sensing, adaptada de O'Reilly & Blackwell⁶².....	13
Figura 1. 7: Interconexión entre los sistemas del QS. Tomada de Greenberg & Schuster ²¹	14
Figura 1. 8: Autoinductores tipo acil-homoserin-lactona, a) AI del sistema Lux: N-(3-oxohexanoil)- homoserin-lactona (OHHL), b) AI del sistema Las: N-(3-oxododecanoil)-homoserin-lactona (OdDHL), c) AI del sistema Rhl: N-butiril-homoserin-lactona (BHL).....	17
Figura 1. 9: Estructuras químicas asociadas a la activación natural de PqsR, a) 3,4-dihidroxy-2-heptilquinolona (POS), b) 2-heptil-4-hidroxiquinolona (HHQ).....	18
Figura 1. 10: Representación esquemática de Iqs y su relación con otros FT's. a) Factores de estrés ambiental y la jerarquía de señalización , b) estructura química de IQS. Adaptado de Lee et al ⁶⁴	19
Figura 1. 11: Matriz de trabajo de DIFAC según la información que se tiene disponible. Tomada de Lopez-Vallejo et al ⁸¹	20
Figura 1. 12: Avance en la predicción estructural en los modelos por homología realizados en el CASP13 según el grado de dificultad. Tomada de Croll et al ⁹⁵	22

Figura 1. 13: Representación esquemática de los molecular fingerprints, a) fingerprint basado en características, b) fingerprint topológico. Adaptado de Bajusz et al ¹⁰⁰.23

Figura 2.1: Diagrama general de la metodología aplicada en el estudio.25

Figura 2.2: Metodología general en la aplicación de los acoplamientos moleculares automatizados.29

Figura 2.3: Diagrama ilustrativo de la metodología implementada en el análisis y la clasificación de un residuo importante para la actividad agonista o antagonista, en base a la diferencia del acumulado de las interacciones por tipo de actividad30

Figura 3.1: Distribución de los valores de actividad biológica para las tres categorías de complejidad en las BD's, donde a) representa los compuestos presentes en LasR, b) PqsR y c) RhIR. En el eje-x se encuentran las cuatro métricas y en el eje-y la distribución de los valores pIC_{50} , la línea horizontal en los gráficos representa la mediana del valor de complejidad.35

Figura 3.2: Núcleos base tipo Bemis-Murcko más recurrentes en las bases de datos, a la izquierda se muestra la frecuencia de estos y a la derecha la estructura. A) representa la BD de LasR, b) la BD de PqsR y c) la BD de RhIR; el tipo de actividad es mostrado según el color de las barras, azul – inactivo, verde – agonista y amarillo – antagonista.38

Figura 3.3: Ejemplos representativos del análisis de máxima común subestructura. N representa el número de moléculas por cada scaffold.38

Figura 3.4: activity cliffs hallados en las BD's, aquí los pares moleculares asociados a la actividad agonista se muestran a la izquierda, mientras que los relacionados con la actividad antagonista se muestran a la derecha41

Figura 3.5: representación visual del espacio químico de las BD's obtenida por t-distributed stochastic neighbor embedding (t-SNE) en base al fingerprint ECFP4. Todos los compuestos son clasificados de acuerdo con la actividad, inactivos-gris, LasR agonistas (verde claro), LasR antagonistas (verde oscuro), PqsR agonistas (rosado), PqsR antagonistas (rojo), RhIR agonistas (azul claro), RhIR antagonistas (azul oscuro).42

Figura 3.6: gráfico de constelación de la base de datos LasR. Aquí los núcleos más representativos de dicha BD son mostrados, teniendo este los mismos parámetros de la figura 3. 5 (t-SNE, ECFP4, métrica de Tanimoto), cada punto **representa un núcleo, el tamaño de los círculos indica el “n” numero de los compuestos** relacionados, adicionalmente las conexiones entre los círculos indican que se comparten los núcleos, los rectángulos representan las series de análogos y el color muestra el promedio de pIC_{50}43

Figura 3.7: alineamiento de las estructuras cristalográficas almacenadas en el PDB; a) complejos del FT LasR, b) complejos del FT PqsR. Los colores de los complejos muestran el tipo de molécula con la cual son acomplejados: azul – agonista, naranja – antagonista, gris – inactivo y beige – AI, el círculo rojo encierra la principal diferencia entre las estructuras de los **FT's**.45

Figura 3.8 Modelo de RhIR obtenido de la plataforma SWISS-MODEL, en azul se encuentra el sitio de unión del FT y en naranja el sitio de unión al ADN.48

Figura 3.9 resultados de los re-acoplamiento de los AI's frente a los FT's, los numerales a), b) y c) representan los alineamientos de las conformaciones de las estructuras cristalográficas en morado y las conformaciones de los re-

acoplamientos en verde, representando LasR-AI, PqsR-AI y RhIR-AI respectivamente; los numerales d). e), f) muestran las interacciones establecidas entre estas moléculas y los residuos de los sitios de unión obtenidas de los re-acoplamientos, representando LasR-AI, PqsR-AI y RhIR-AI respectivamente.	50
Figura 3.10 fingerprints de interacción ligando-proteína (PLIF's) obtenidos para los acoplamientos realizados entre las moléculas de las bases de datos y los FT's estudiados. En el eje "Y" se observa la prevalencia del tipo de interacción, siendo positivo para las interacciones preferentes para los agonistas y negativo para las interacciones antagonistas; mientras que el eje "X" muestra el residuo aminoacídico con el cual sucede la interacción. Los PLIF's asociados a LasR son mostrados por a), a PqsR por b) y a RhIR por c); el color de las barras indica el tipo de interacción relacionada con cada residuo.	53
Figura 3.11 mapa de calor y dendograma de similitud entre las diferentes trayectorias de las dinámicas realizadas por FT a través de la raíz de la desviación cuadrática media (RMSD); donde se representan las dinámicas de a) LasR, b) PqsR y c) RhIR; al lado izquierdo se observa el dendograma que agrupa según su similitud a las 4 trayectorias, mientras al lado derecho se representa a través de una escala de color la diferencia entre ellas. Significando AI: auto-inductor, ANT: antagonista, AGO: agonista y INA: inactiva.	55
Figura 3.12 cambios de conformación entre las dinámicas de los complejos LasR-antagonista (naranja) y LasR-agonista (azul). A la derecha se observa el mapa de calor que compara la diferencia entre los RMSD de las dinámicas y a la izquierda se muestra el alineamiento de los complejos cuando presentan la mayor variación en estos RMSD. El recuadro rojo indica los cuadros (frames) con mayores diferencias de RMSD entre los dos complejos.	55
Figura 3.13 cambios de conformación entre las dinámicas de los complejos PqsR-antagonista (naranja) y PqsR-agonista (azul). A la derecha se observa el mapa de calor que compara la diferencia entre los RMSD de las dinámicas y a la izquierda se muestra el alineamiento de los complejos cuando presentan la mayor variación en estos RMSD. El recuadro rojo indica los cuadros (frames) con mayores diferencias de RMSD entre los dos complejos.	56
Figura 3. 14 cambios de conformación entre las dinámicas de los complejos RhIR-antagonista (naranja) y RhIR-agonista (azul). A la derecha se observa el mapa de calor que compara la diferencia entre los RMSD de las dinámicas y a la izquierda se muestra el alineamiento de los complejos cuando presentan la mayor variación en estos RMSD. El recuadro rojo indica los cuadros (frames) con mayores diferencias de RMSD entre los dos complejos.	57
Figura 3.15 RMSF por residuo en las dinámicas realizadas; se representan las dinámicas: a) LasR, b) PqsR y c) RhIR. El color de cada gráfica representa el tipo de actividad que tiene la molécula correspondiente, en donde los antagonistas son naranja, los agonistas son azul y los inactivos gris. Adicionalmente los recuadros rojos muestran los cambios más significativos. .	59
Figura 3.16 RMSF mostrado para cada uno de los átomos de las moléculas con las cuales se realizaron las dinámicas moleculares. Los compuestos asociados y la gráfica de variación de RMSF con respecto al FT LasR se muestra en a); con respecto a PqsR en b) y a RhIR en c). El color de cada gráfica hace referencia al tipo de actividad que posee la molécula, naranja los antagonistas, azul los agonistas y gris los inactivos.	60
Figura 3.17 interacciones presentadas por los complejos de compuestos agonista y antagonista con el FT LasR. En a) se representan las interacciones con el antagonista y en b) se representan las interacciones con el agonista.	61
Figura 3.18 interacciones presentadas por los complejos de compuestos agonista y antagonista con el FT PqsR. En a) se representan las interacciones con el antagonista y en b) se representan las interacciones con el agonista.	62
Figura 3.19 interacciones presentadas por los complejos de compuestos agonista y antagonista con el FT RhIR. En a) se representan las interacciones con el antagonista y en b) se representan las interacciones con el agonista.	63

Figura 3.20 modelos farmacofóricos desarrollados para la actividad antagonista en los FT's. A la izquierda se representa el alineamiento de las diez moléculas más activas y a la derecha los puntos farmacofóricos encontrados con sus respectivas distancias, estando representado el modelo farmacofórico de LasR en a), de PqsR en b) y de RhIR c).	65
Figura 3.21 exactitud y error para los modelos de clasificación k-nearest neighbors (k-NN Model) y Extreme gradient boosting (Xgb Model) frente a posibles moléculas activas e inactivas en los FT's a) LasR, b) PqsR y c) RhIR.	68
Figura 3.22: bits o características más relevantes para la clasificación de moléculas activas frente a LasR, según el modelo Extreme gradient boosting (Xgb). En la parte superior e inferior de la figura se representan las partes estructurales asociadas a las 2 actividades, siendo las superiores (naranja) activas y las inferiores (gris) inactivas; en el medio se observa la influencia de cada uno de los bits y su importancia para la clasificación.	70
Figura 3.23: bits o características más relevantes para la clasificación de moléculas activas frente a PqsR, según el modelo Extreme gradient boosting (Xgb). En la parte superior e inferior de la figura se representan las partes estructurales asociadas a las 2 actividades, siendo las superiores (naranja) activas y las inferiores (gris) inactivas; en el medio se observa la influencia de cada uno de los bits y su importancia para la clasificación.	71
Figura 3.24: bits o características más relevantes para la clasificación de moléculas activas frente a RhIR, según el modelo Extreme gradient boosting (Xgb). En la parte superior e inferior de la figura se representan las partes estructurales asociadas a las 2 actividades, siendo las superiores (naranja) activas y las inferiores (gris) inactivas; en el medio se observa la influencia de cada uno de los bits y su importancia para la clasificación.	72
Figura 3.25 compuestos escogidos por el cribado como los más probables de tener actividad antagonista sobre los FT's. A la izquierda de la figura se observa el número de interacciones con los aminoácidos agonistas (verde) y los antagonistas (amarillo) y a la derecha las estructuras de las moléculas resultantes.	77
Figura 3.26 Gráfico de radar con las familias de las especies de donde fueron extraídos los compuestos con mayor potencial antagonista sobre los FT's. A la izquierda se observa el mapa de radar, donde los ejes representan el número de compuestos por Familias que podrían tener actividad antagonista contra LasR (azul), PqsR (naranja) y RhIR (verde); a la derecha un ejemplo de compuestos con el mejor puntaje por cada FT en la familia con mayor prevalencia, en azul el ejemplo de LasR, en naranja PqsR y en verde RhIR.	78

Lista de tablas

	Pág.
Tabla 1. 1: Antibióticos de uso predilecto en infecciones generadas por <i>P. aeruginosa</i> . Adaptado de la guía sensibilización a antibióticos: mensajes claves, Instituto de salud pública de Inglaterra ⁵⁷	10
Tabla 1. 2: Factores de virulencia dependientes del QS producidos por <i>P. aeruginosa</i> . Adaptada de Rhem et al ³⁶	15
Tabla 2. 1: Métricas usadas para evaluar el desempeño de los modelos de machine learning	32
Tabla 3. 1: Moléculas presentes en las BD	34
Tabla 3. 2 Características estructurales asociadas con la actividad frente a los FT. *Las interacciones hidrofóbicas y los apilamientos de anillos no se pudieron cuantificar.	44
Tabla 3. 3 resultados del alineamiento local realizado sobre el FT RhIR con la plataforma BLAST	46
Tabla 3. 4 métricas de validación para los modelos construidos del FT RhIR.....	47
Tabla 3. 5 parámetros usados en los acoplamientos frente a los FT's; PH: puente de hidrogeno.....	49
Tabla 3. 6 aminoácidos relevantes por cada tipo de actividad para todos los FT's estudiados.	63
Tabla 3. 7 modelos de machine learning y métricas de desempeño utilizadas para la evaluación de la clasificación de moléculas activas e inactivas frente al FT LasR.....	66
Tabla 3. 8 mejores resultados del cribado virtual. Dónde: ML significa machine learning, PLIF's significa Protein ligand Interaction fingerprint, AA agonistas son los aminoácidos tipo agonista con los que interacciona la molécula y AA antagonistas son los son los aminoácidos tipo antagonista con los que interacciona la molécula.....	74

Lista de símbolos y abreviaturas

Símbolos con letras latinas

Símbolo	Término	Unidad SI	Definición
μM	Micro molar	M	Unidad de concentración

Abreviaturas

Abreviatura	Término
<i>P. aeruginosa</i>	<i>Pseudomonas aeruginosa</i>
3-oxo-C12-HSL	3-oxododecanoil-L-homoserin-lactona (Autoinductor <i>Ias</i>)
ada	<i>Ada boost classifier</i>
AHL	Acil-homoserin-lactona
AI	Autoinductor
AUC	Área bajo la curva
BD	Base de datos
C4-HSL	N-Butanoil-homoserin-lactona (Autoinductor <i>Rhl</i>)
CASP	<i>Critical assessment of techniques for protein structure prediction</i>
catboost	<i>CatBoost classifier</i>
CRPA	<i>Carbapenem resistant Pseudomonas aeruginosa</i>
DFBE	Diseño de fármacos basado en la estructura
DFBL	Diseño de fármacos basado en el ligando
DIFAC	Diseño de fármacos asistido por computadora
DM	Simulaciones de dinámica molecular
dt	<i>Decision tree classifier</i>

Abreviatura	Término
EC ₅₀	Concentración efectiva cincuenta
ECFP2	<i>Extended connectivity fingerprint radius 2</i>
ECFP4	<i>Extended connectivity fingerprint radius 4</i>
et	<i>Extra trees classifier</i>
FN	Falsos negativos
FP	Falsos positivos
FT	Factor de transcripción
gbc	<i>Gradient boosting classifier</i>
HHQ	2-heptil-4-hidroxiquinolona
I&D	Investigación y desarrollo
IC ₅₀	Concentración inhibitoria cincuenta
lqs	<i>Integrated quorum sensing</i>
IQS	2-(2-hidroxi-fenil)-tiazolo-4-carbaldehido (Autoinductor lqs)
kappa	Coefficiente kappa de Cohen
k-NN	<i>k-nearest neighbors classifier</i>
lda	<i>Linear discriminant analysis</i>
lightgbm	<i>Light gradient boosting machine</i>
lr	<i>Logistic regression classifier</i>
MACCS keys	<i>Molecular access system keys</i>
MCC	Coefficiente de correlación de Mathews
MCS	<i>Maximum common substructure analyze</i>
MHFP	<i>MiniHash fingerprint</i>
nb	<i>Native bayes</i>
NCBI	<i>National Center of Biotechnology Information</i>
PCA	<i>Principal Component Analyze</i>
PDB	<i>Protein data bank</i>
PLIF's	<i>Protein-ligand interaction fingerprints</i>

Abreviatura	Término
PME	<i>Partice mesh Ewald</i>
Pqs	<i>Pseudomonas quinolone signal</i>
POS	2-heptil-3-hidroxi-4-quinolona (Autoinductor Pqs)
qda	<i>Quadratic discriminant analysis</i>
QS	<i>Quorum Sensing</i>
QSAR	<i>Quantity structure activity-relationship</i>
REA	Relación estructura actividad
rf	<i>Random forest classifier</i>
ridge	<i>Ridge Classifier</i>
RMSD	<i>Root mean square deviation of atomic positions</i>
RMSF	<i>Root mean square fluctuation of atomic positions</i>
SECFP	<i>SMILES extended connectivity fingerprint</i>
SMILES	<i>Simplified molecular input line entry specification</i>
SMOTE	<i>Synthetic minority over-sampling technique</i>
SVM	<i>Support vector machine</i>
svm	<i>Supported vector machine</i>
t-SNE	<i>t-distributed stochastic neighbor embedding</i>
Vina	AutoDock Vina 1.1.2.
VMD	<i>Visual Molecular Dynamics</i>
VN	Verdaderos negativos
VP	Verdaderos positivos
xgboost	<i>Extreme gradient boosting</i>

Introducción

Varios estudios sugieren que la primera forma de vida en el planeta se puede asociar con organismos unicelulares, que pueblan la tierra entre 3 y 4 billones de años atrás^{1,2}. Al sobrevivir tanto tiempo dichos microorganismos han desarrollado diferentes formas de adaptarse para poder mantener su descendencia sobre el planeta tierra, uno de estos mecanismos resulta en la invasión de un organismo multicelular de mayor tamaño, el cual tiene las condiciones óptimas para su supervivencia. Si se asocia la presencia del microorganismo con una enfermedad en específico a través de los postulados modificados de Koch, se puede afirmar que ese microorganismo produce una infección^{3,4}.

La investigación en nuevos blancos moleculares, para el desarrollo de moléculas antimicrobianas, ha cobrado especial relevancia en los últimos años debido al aumento del reporte de bacterias que presentan resistencia a los antibióticos que están disponibles en el mercado. La pérdida de efectividad por parte de estos medicamentos tiene origen en factores como: la baja tasa de nuevos compuestos antimicrobianos en el mercado que diversifiquen las opciones terapéuticas⁵, el uso indiscriminado de los mismos por parte diferentes actores, entre los que se encuentran pacientes, personal de salud, industria y demás^{6,7}.

La infección por *Pseudomonas aeruginosa* constituye uno de los grandes retos a nivel de infecciones hospitalarias. Esto a causa del aumento de la resistencia que presenta este microorganismo y a su prevalencia sobre pacientes inmunocomprometidos^{8,9}. Según la lista “*WHO priority pathogens list for R&D of new antibiotics*” que clasifica la prioridad en el desarrollo de antibióticos según el patógeno, se clasificó a la *P. aeruginosa* resistente a carbapenem (CRPA por sus siglas en inglés) con prioridad crítica para la investigación y el desarrollo de nuevas moléculas contra este microorganismo¹⁰.

En los últimos dos años la Administración de Medicamentos y Alimentos de los Estados Unidos (US-FDA) ha aprobado seis nuevos medicamentos con indicación de antibióticos, dos de ellos tetraciclinas^{5,11}, un carbapenem con inhibidor de β -lactamasa¹², aminoglicosido¹², una cefalosporina^{13,14} y un carbapenem más inhibidor de dehidropeptidasa y un inhibidor de β -lactamasa^{13,14}; ninguno de los anteriores fue clasificado para el uso contra CRPA, evidenciando la falta de innovación en nuevos blancos terapéuticos.

Entre los blancos de estudio para combatir la *P. aeruginosa* se destaca el sistema de comunicación denominado *quorum sensing* (QS), este sistema permite a la bacteria reconocer la concentración bacteriana en un ambiente determinado, a través de la identificación de una molécula específica denominadas autoinductor (AI); una vez alcanzada una población determinada, el receptor que reconoce a los AI, conocidos como factores de transcripción (FT), activa la maquinaria genética para la producción de diferentes mecanismos de virulencia^{15,16}.

La red QS se basa en una relación de proteínas que se encargan tanto de la síntesis de los AI's, como del reconocimiento de los mismos¹⁷. Las principales relaciones proteicas conocidas a la fecha son los sistemas LasI/LasR – RhII/RhIR¹⁸, PqsBCDE/PqsR¹⁹, y AmbBCDE/IqsR²⁰. Los AI's de estos sistemas son las siguientes moléculas respectivamente: 3-oxododecanoil-L-homoserin-lactona (3-oxo-C12-HSL), N-butanoil-homoserin-lactona (C4-HSL), 2-heptil-3-hidroxi-4-quinolona (POS) y 2-(2-hidroxi-fenil)-tiazolo-4-carbaldehído (IQS). Aunque existe una gran cantidad de sistemas asociados al QS, los mencionados anteriormente están en la cima de esta compleja red jerárquica²¹.

El diseño de fármacos asistido por computadora (DIFAC) es una herramienta muy útil, usada en el proceso de I&D en el descubrimiento de fármacos, este disminuye el número de ensayos biológicos al mejorar las probabilidades de encontrar moléculas con la actividad deseada^{22,23}, generando una optimización de los recursos, tiempo y dinero invertido en esta etapa específica²⁴. Varios autores clasifican en dos ramas las metodologías usadas en el DIFAC: diseño de fármacos basado en el ligando (DFBL) y diseño de fármacos basado en la estructura (DFBE)²⁵⁻²⁷.

En el DIFAC se tiene que partir de información experimental, para esto el “big data” asociado al descubrimiento de fármacos (grandes cantidades de información biológica, química y farmacológica) cobra relevancia^{28,29}; esta información suele estar almacenada en grandes bases de datos (BD's), algunas públicas como: ChEMBL³⁰, PubChem BioAssay³¹, DrugBank³², PDB³³ y otras privadas que suelen pertenecer a compañías farmacéuticas. De las herramientas computacionales usadas para el análisis de big data destacan: análisis de espacio químico, complejidad molecular, estudios de similitud y *activity cliffs*, análisis de *scaffolds* (SFA por sus siglas en inglés) y análisis de pares moleculares emparejados (MMPA por sus siglas en inglés)²⁸.

En el presente trabajo se utilizó la información biológica disponible de los factores LasR, PqsR y RhIR asociados al QS con el objetivo de encontrar las características estructurales relevantes que se pueden asociar a la actividad antagonista, para esto se aplicó el diseño DIFAC en sus dos modalidades: DFBE y DFBL, contribuyendo así con la búsqueda y la optimización de los compuestos con actividad sobre estos blancos.

Hipótesis

La actividad antagonista de moléculas pequeñas sobre los factores de transcripción LasR, PqsR y RhlR de *P. aeruginosa* está determinada por interacciones intermoleculares con algunos residuos claves que no son importantes para la actividad agonista.

Objetivos

Objetivo general

Contribuir a la identificación de características estructurales relacionadas con la actividad antagonista de moléculas pequeñas sobre los factores de transcripción LasR, PqsR y RhlR en *P. aeruginosa*.

Objetivos específicos

- Caracterizar las posibles relaciones de estructura-actividad de antagonistas sobre los blancos de estudio utilizando herramientas quimioinformáticas basadas en el ligando.
- Identificar, a través de metodologías del diseño de fármacos basado en la estructura y el ligando, compuestos con potencial afinidad por los blancos de estudio.
- Aportar al establecimiento de las características estructurales o farmacóforos relacionados a la actividad antagonista de los blancos de estudio.

1 Estado del arte

1.1 Infecciones causadas por *P. aeruginosa*

La *P. aeruginosa* es una bacteria ubicua gran negativa presente en múltiples ambientes, de los cuales el ambiente hospitalario es donde presenta mayor índice de resistencia, adicionalmente es causante de fibrosis quística^{34,35}; este patógeno suele habitar en superficies como suelo, agua, implementos quirúrgicos y paredes, lo que incrementa exponencialmente la posibilidad de infectar a un paciente. Además de lo anterior tiene una alta adaptabilidad y una alta resistencia a los antibióticos, generando uno de los mayores retos en las infecciones oportunistas nosocomiales³⁶.

Parte de la adaptabilidad que posee la *P. aeruginosa* se debe a sus grandes cadenas genómicas (~5-7Mbp), éstas pueden producir una gran cantidad de metabolitos secundarios, polímeros y estructuras carbonadas capaces de ser aceptoras de electrones; constituyéndose en la defensa primaria contra cualquier tipo de molécula que pueda causarle algún daño³⁷.

La capacidad de producir múltiples factores de virulencia lo vuelven uno de los patógenos más peligrosos en pacientes inmunocomprometidos, siendo éstos encontrados usualmente en ambientes hospitalarios³⁸. Estos factores de virulencia son diversos y cumplen múltiples funciones, entre los cuales tenemos: piocianina, pioverdina, elastasas, alcalin proteasas, lectinas, fosfolipasas, ramnolipidos, exotoxinas y ácido cianhídrico; siendo causantes de la agravación de las infecciones producidas y de algunas de la resistencias intrínsecas que posee este patógeno³⁹.

Otro factor que complementa la infectividad de este patógeno son los múltiples tipos de resistencia, entre las que se tienen: resistencia intrínseca, adquirida y adaptativa⁴⁰. Todo lo anterior se une para generar la alta peligrosidad de este microorganismo, llevando a que hoy en día poseamos pocas defensas contra el mismo⁴¹ y desencadenando la clasificación dada por la OMS como prioridad crítica para el desarrollo de antibióticos contra este patógeno¹⁰.

1.1.1 Prevalencia de las infecciones por *P. aeruginosa*

De acuerdo con el Centro de Control y Prevención de Enfermedades de los Estados Unidos (CDC), un estimado de 32.600 infecciones asociadas a *P. aeruginosa* con cuidado de personal sanitario ocurrieron en los Estados Unidos durante el 2017, de los cuales 2.700 pacientes fallecieron a causa de la infección, gastando más de 767 millones de dólares en su cuidado y prevención⁴². Por otra parte, en el estudio

realizado por Suetens *et-al* en el 2018, en donde se reporta la prevalencia de las infecciones asociadas a *P. aeruginosa* en la Unión Europea en el periodo 2016-2017, este patógeno generó 16.869 infecciones, de las cuales 9.569 fueron adquiridas intrahospitalariamente, siendo el cuarto microorganismo más común en producir infecciones resistentes a carbapenem⁴³.

En la investigación realizada por Aliberti *et-al*, donde se establece la prevalencia de las infecciones producidas por *P. aeruginosa* en una muestra de 3193 pacientes localizados en 54 países del mundo, se evidencia que la región de Sudamérica tiene la tercera mayor tasa de infecciones de neumonía adquirida en comunidad (4,9 %); mientras que en la tasa de las infecciones resistentes a antibióticos por este patógeno, Sudamérica ocupa el segundo lugar (3,0%), como se evidencia en la figura 1-1⁴⁴.

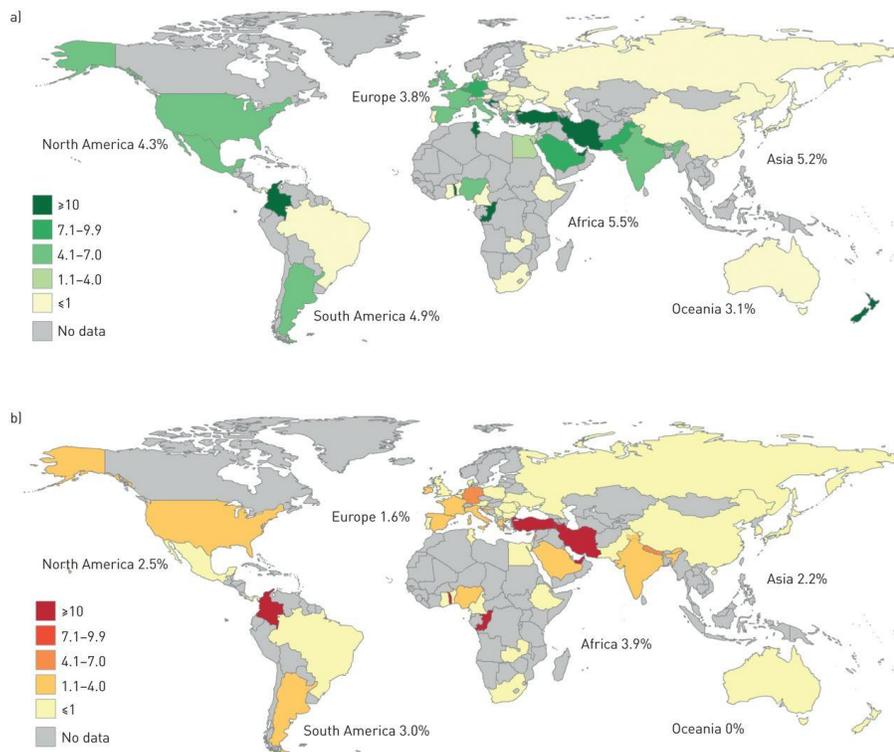


Figura 1. 1: Prevalencia de las neumonías causadas por *P. aeruginosa* en el estudio realizado por Aliberti *et-al*, a) neumonía adquirida en la comunidad asociada a *P. aeruginosa*, b) neumonía asociada a *P. aeruginosa* resistentes a antibióticos. Tomada de Aliberti *et-al* ⁴⁴.

Cabe resaltar de la figura 1-1 que el país de la región con mayor proporción de neumonías adquiridas en comunidad y el de neumonías con resistencia a antibióticos es Colombia, con más del 10 %. De acuerdo con los resultados del programa de vigilancia de brotes de infecciones asociadas a la atención en salud (IAAS) del 2019, el 27 % de los casos reportados corresponden a infecciones ocasionadas por

P. aeruginosa (92 casos), siendo el microorganismo con mayor índice de infección en ese año⁴⁵, como se observa en la figura 1-2

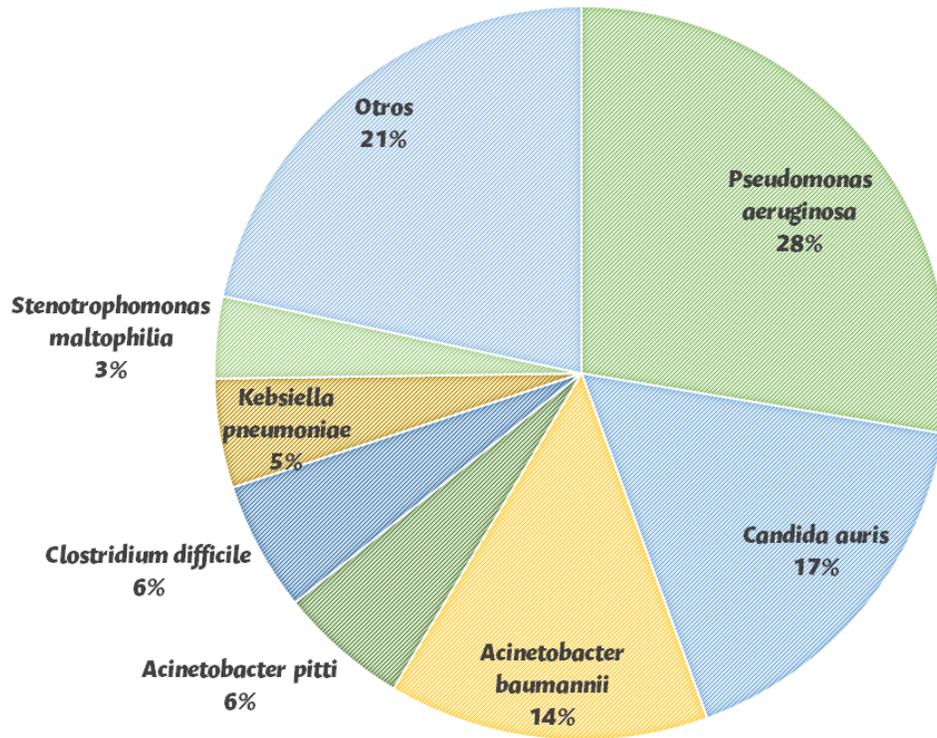


Figura 1. 2: Prevalencia de los brotes de infecciones por microorganismo durante el 2019 en Colombia. Adaptada del boletín epidemiológico semanal 07 del 2020, Instituto Nacional de Salud⁴⁵.

Al encontrarse ampliamente distribuidas las infecciones causadas por *P. aeruginosa* en todo el mundo y especialmente en el territorio colombiano, urge encontrar otro tipo de soluciones para atender esta amenaza pública.

1.1.2 Resistencia a antibióticos por *P. aeruginosa*

En efecto, un microorganismo que presenta resistencia a antibióticos multiplica las posibilidades de producir una infección que desencadene una muerte. Como se mencionó con anterioridad, a *P. aeruginosa* se le han reportado múltiples mecanismos de resistencia; siendo estos clasificados por Rehm et-al en resistencias intrínsecas, adquiridas y adaptativas.

Una resistencia intrínseca se puede definir como la producida por genes que codifican proteínas propias de una bacteria cambiando su estructura o composición celular, sin que ningún medio externo intervenga, disminuyendo la probabilidad de que una molécula exógena pueda causarle daño. La supresión de canales de membrana o porinas es una de las resistencias más representativas de este

grupo, aquí se reemplazan estas porinas por canales selectivos evitando la entrada de compuestos tóxicos⁴⁶. Por este mecanismo la *P. aeruginosa* adquiere resistencia a antibióticos de amplio espectro como los carbapenémicos o las cefalosporinas⁴⁷.

Algunas de las bombas de eflujo también entran en esta clasificación, entre las que tenemos: MexAB-OprM, MexXY/OprM(OprA), MexCD-OprJ, y MexEF-OprN, éstas se encargan de la eliminación de xenobióticos tóxicos que logran ingresar a la célula⁴⁸. El último mecanismo intrínseco que se conoce es la producción de la β -lactamasa AmpC, encargada de la hidrólisis de las moléculas β -lactámicas⁴⁹.

A diferencia de la resistencia intrínseca, la adquirida necesita de procesos como mutaciones o transferencias de plásmidos que tienen un mayor efecto que a las intrínsecas; estas suelen presentarse en ambientes hospitalarios, lo que incrementa las poblaciones resistentes. Aquí los plásmidos juegan un rol clave al ser uno de los mejores vehículos para la transmisión de estas resistencias; se ha comprobado que por esta vía se pueden transmitir genes que codifican β -lactamasas de espectro extendido, carbapenemasas, enzimas modificantes de aminoglucósidos y 16S rRNA metilasas, consiguiendo así la ineffectividad de antibióticos de primera línea, como β -lactámicos, fluoroquinolonas y aminoglucósidos⁵⁰.

La resistencia adaptativa aún no se ha comprendido del todo, pero se puede definir como la resistencia que solo se presenta cuando existe la presencia del xenobiótico, como los antibióticos, desinfectantes u otros mecanismos de estrés ambiental; estos compuestos inducen alteraciones en la expresión de algunos genes que a su vez modifican la producción de proteínas, sin embargo si la molécula causante es retirada del medio se puede regresar al efecto basal, generando de nuevo susceptibilidad a la misma⁵¹.

Las β -lactamas, aminoglucósidos, polimixinas y fluoroquinolonas se suelen ver afectadas por la resistencia adaptativa, además de algunos desinfectantes como la clorhexidina⁵². Pero la colistina suele ser el antibiótico con mayor afectación por esta vía, generando su desuso a nivel hospitalario, molecularmente se ha observado el cambio en algunos aminoácidos importantes para la unión de la colistina con su blanco molecular⁵³. En la figura 1-3 se pueden observar todas las formas de resistencias clasificadas por su mecanismo; aunque suelen compartirse, se diferencian en su modo de activación.

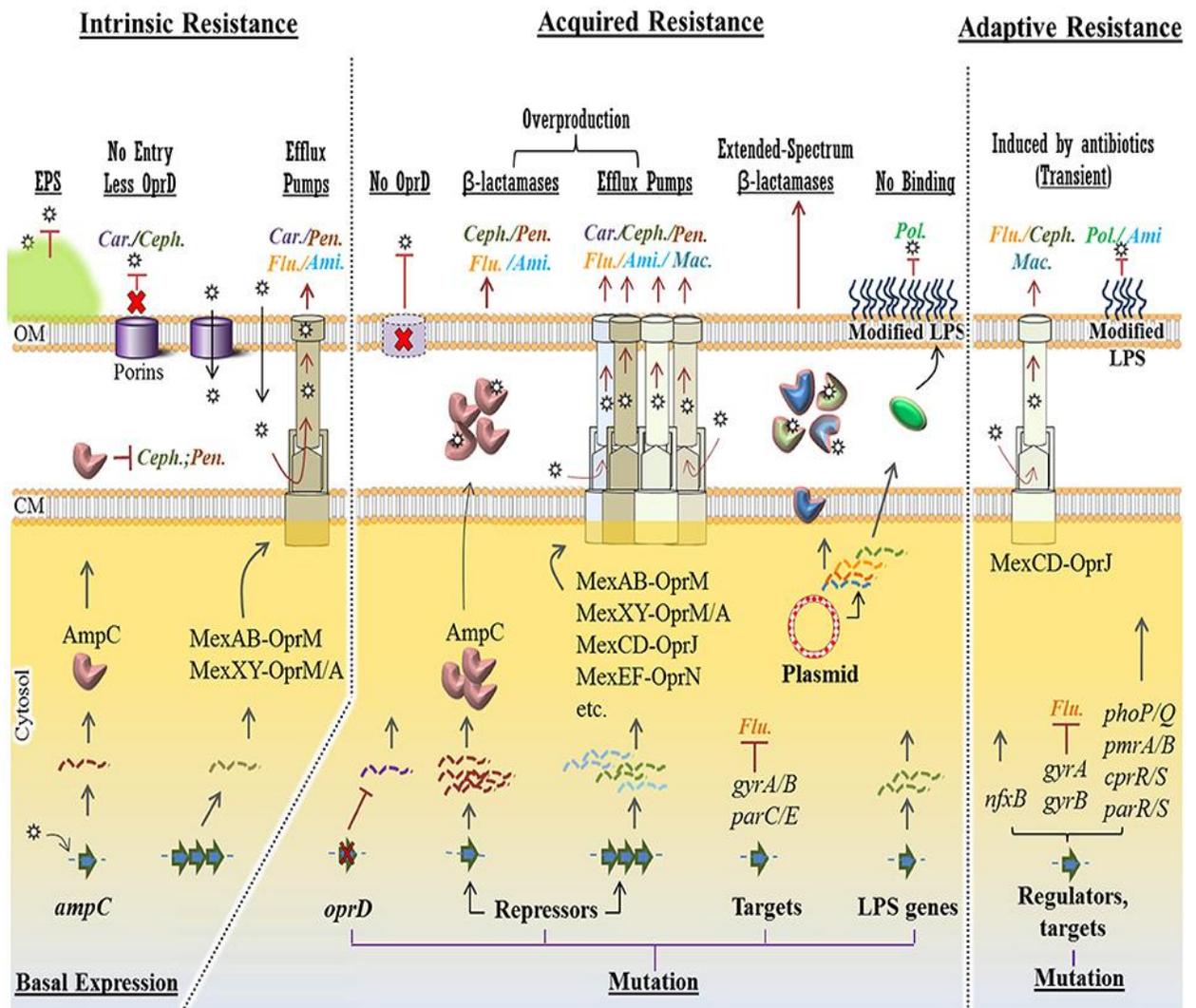


Figura 1. 3: Mecanismos de resistencia a antibióticos en *P. aeruginosa*: intrínseco, adquirido y adaptativo. Tomada de Rehm et-al³⁶.

Como se menciona anteriormente, la OMS clasificó a la *P. aeruginosa* resistente a carbapenémicos CRPA como prioridad crítica para el desarrollo de antibióticos ¹⁰, en este aspecto el Sistema Nacional de Vigilancia de Resistencia Antimicrobiana (IAAS) en su informe del 2018 mostró que las carbapenemasas tienen una alta incidencia en las infecciones producidas por *P. aeruginosa* a lo largo de la geografía colombiana, como se evidencia en la figura 1-4, dificultando el tratamiento de las infecciones producidas por este patógeno.

10 Contribución a la identificación de potenciales antagonistas asociados al quorum sensing en *Pseudomonas aeruginosa* a través de métodos quimioinformáticos

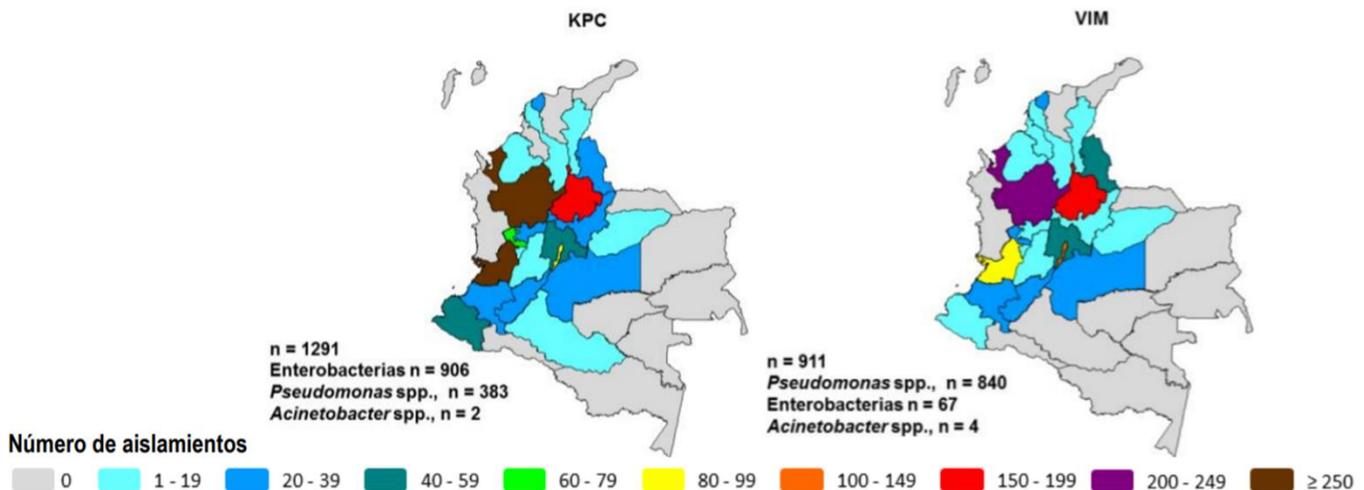


Figura 1. 4: Distribución geográfica de casos de carbapenemasas en Colombia, septiembre 2012-2018. Adaptada del informe de resultados de vigilancia por laboratorio de resistencia antimicrobiana en infecciones asociadas a la atención en salud (IAAS) 2018⁴⁵.

1.1.3 Tratamiento de infecciones causadas por *P. aeruginosa*

La *P. aeruginosa* es capaz de producir diversas infecciones entre las que se tienen: infecciones del tracto urinario, neumonía adquirida en comunidad, neumonía asociada a ventilador, endocarditis, meningitis, infecciones oculares, infecciones de oído (pericondritis de oído, otitis externa, otitis externa maligna) e infecciones de piel y tejidos blandos (foliculitis, osteomielitis, infección de pies, ectima gangrenosa, fascitis necrotizante y quemaduras)⁵⁴, de las cuales las neumonías resultan de los mayores retos en terapéutica especialmente las asociadas a fibrosis quística.

Para el tratamiento de estas infecciones se tienen disponibles múltiples clases de antibióticos como se muestra en la tabla 1-1, los cuales pueden atacar a esta bacteria mediante distintos blancos terapéuticos; sin embargo, debido a la resistencia bacteriana y a la toxicidad asociada, muchas de estas opciones no suelen estar disponibles para su uso. Adicionalmente, algunos de estos antibióticos no se encuentran aprobados por el Instituto Nacional de Vigilancia de Medicamentos y Alimentos (INVIMA), como: ticarcilina, mezlocilina y azlocilina, según las normas farmacológicas del 2020; complicando así el tratamiento de *P. aeruginosa* resistente⁵⁵. Entre los antibióticos aprobados en los últimos cinco años por la US-FDA, en Colombia no se encuentra aprobado ninguno por el INVIMA, siendo estos: cefiderocol, ceftolozan-tazobactam, ceftazidime-avibactam, meropenem-vaborbactam, imipienem-relebactam y plazomicina; los cuales se suelen usar frecuentemente en infecciones provocadas por CRPA⁵⁶.

Tabla 1. 1: Antibióticos de uso predilecto en infecciones generadas por *P. aeruginosa*. Adaptado de la guía sensibilización a antibióticos: mensajes claves, Instituto de salud pública de Inglaterra⁵⁷.

Clase de antibiótico	Blanco molecular	Antibiótico
Aminoglucósidos	Inhibición de la subunidad 30S del ribosoma	Gentamicina Amikacina Tobramicina
Quinolonas	1ra y 2da generación: inhibición del dominio ligasa de la topoisomerasa II. 3ra y 4ta generación: inhibición selectiva por el dominio ligasa de la topoisomerasa IV.	Ciprofloxacina - 2da generación. Levofloxacina - 3ra generación
Polimixinas	Anfipáticas con actividad superficial, interactuando de manera potente con los fosfolípidos, rompiendo así la membrana celular	Polimixina B Colistina
Cefalosporinas	Inhibición de la síntesis del peptidoglicano	Ceftazidima Cefepime Cefoperazona Cefpiroma Ceftobiprole
Carboxipenicilinas	Inhibición de la síntesis del peptidoglicano	Carbencilina Ticarcilina
Ureidopenicilinas	Inhibición de la síntesis del peptidoglicano	Mezlocilina Azlocilina Piperacilina
Carbapenemas	Inhibición de la síntesis del peptidoglicano	Meropenem Imipenem Doripenem

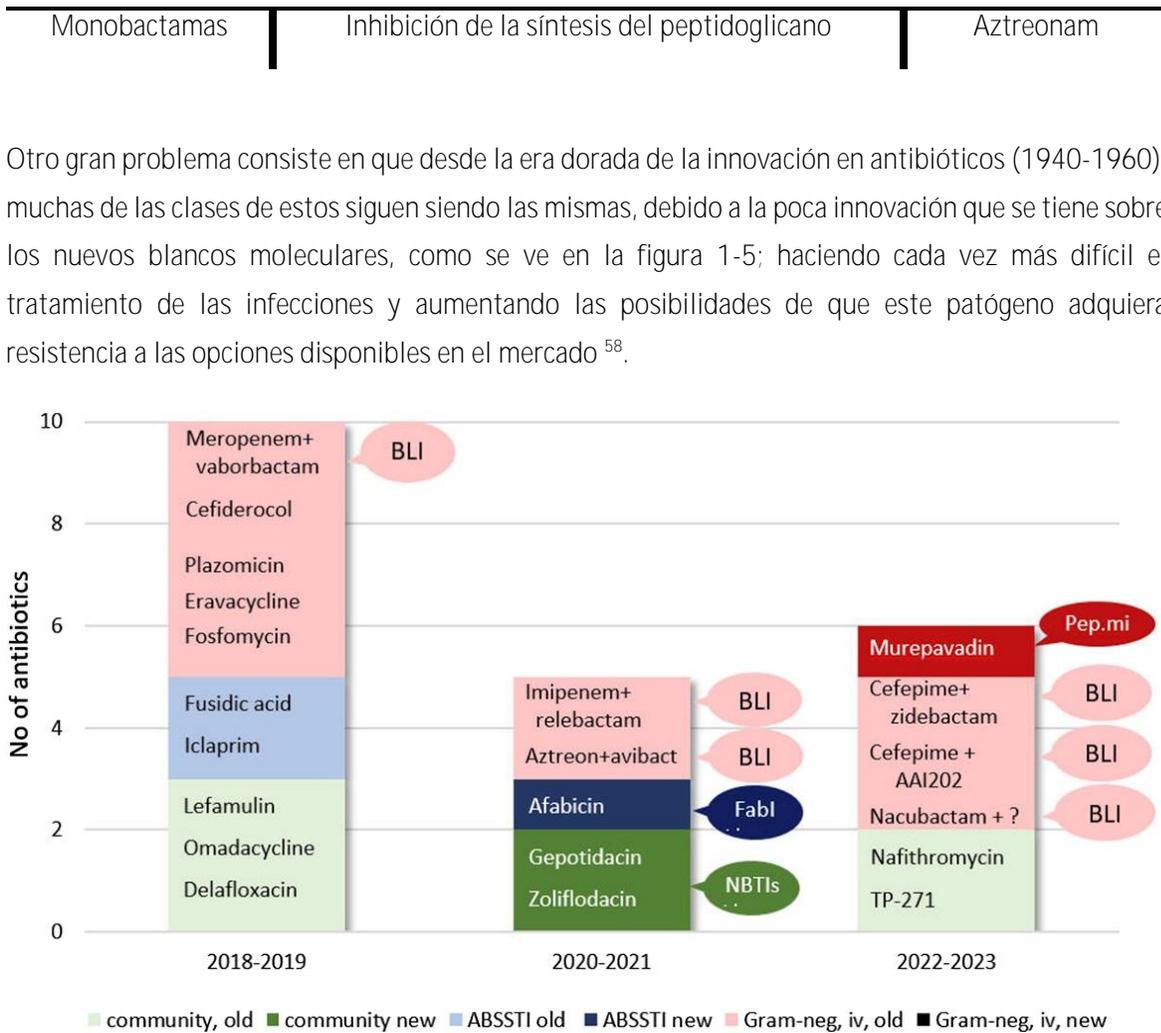


Figura 1. 5: Potenciales nuevos antibióticos posiblemente aprobados por la US-FDA, clasificados según su tasa de innovación, siendo *community old*: núcleos base usualmente usado en terapéutica, *community new*: núcleos base nuevos, NBTI's: nuevos inhibidores de topoisomerasa tipo II, BLI: inhibidores de β -lactamasa, FabI: inhibidores de FabI, específicos para estafilococos. Tomado de Theuretzbacher et al⁵⁸.

1.2 Quorum sensing y factores de transcripción

El censo de grupo o *quorum sensing* (QS) se puede definir como un mecanismo que le permite a una bacteria conocer la población en su ambiente circundante, generando que estos microorganismos puedan controlar la expresión de múltiples factores de virulencia^{36,39,59}. Este censo posibilita la dedicación de un mayor esfuerzo a la colonización por parte de las bacterias, aumentando así sus índices de supervivencia. Los sistemas que controlan el QS en bacterias usualmente se componen de un equilibrio entre dos proteínas, una sintasa o inductora encargada de la producción del AI y otra

receptora que reconoce al AI y activa la maquinaria genética generando la producción de algunos factores de virulencia asociados^{15,16}.

Como se muestra en la figura 1-6 la concentración de los AI dependen directamente de la población bacteriana circundante, generando que estos mecanismos solo se activen, cuando se sobrepasa un umbral determinado de población en el ambiente^{60,61}.

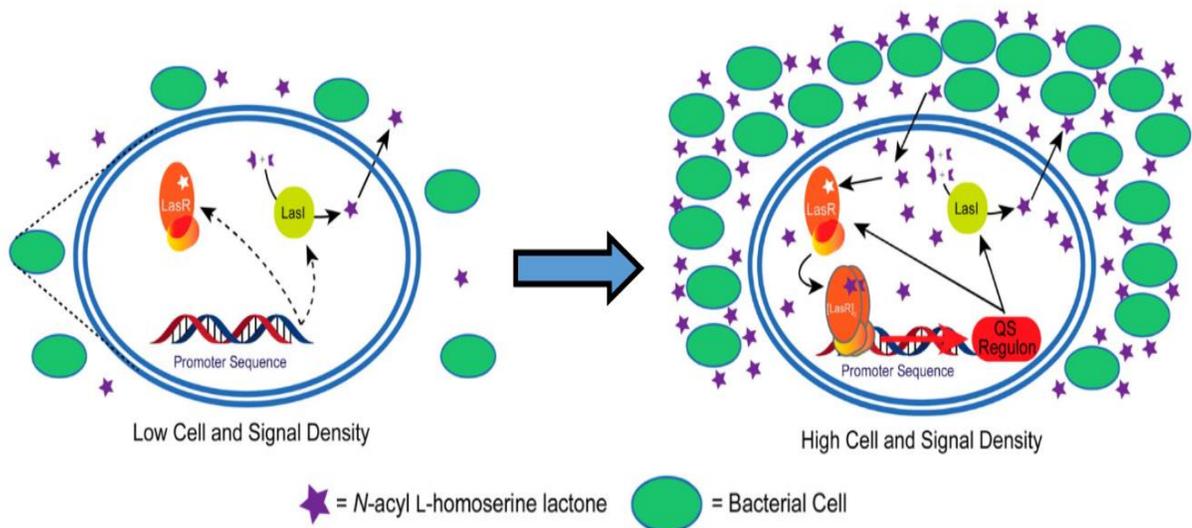


Figura 1. 6: Representación esquemática funcionamiento de los sistemas *quorum sensing*, adaptada de O'Reilly & Blackwell⁶².

Este mecanismo ha sido ampliamente estudiado en los últimos años y se han identificado, tanto los factores de virulencia producidos por el QS, como los sistemas de inductores/receptores que los producen. Existe una compleja red jerárquica entre los sistemas asociados al QS, a la fecha se conocen los siguientes factores de transcripción: LasR, PqsR, RhlR, IqsR, VqsM, AlgR, CdpR, GacA, ExsA, MexT, QscR, SoxR, PhoR, GdbR, PchR, SphR, RsaL y FleQ⁶³, siendo ésta una compleja red en donde todos se interconectan entre sí, activándose o inactivándose dependiendo de las condiciones ambientales^{21,64}.

La producción de estos factores de virulencia es una de las estrategias más exitosas de la *P. aeruginosa* contra la defensa del sistema inmune y el avance patogénico de la infección; sin embargo, esto implica un gran esfuerzo metabólico por parte de este patógeno, por este motivo la bacteria solo lo puede activar cuando existe determinada población garantizando así su efectividad⁶⁵. Estos factores se suelen asociar con resistencias mecánicas, químicas y enzimáticas que aumentan la virulencia y la citotoxicidad de este microorganismo.

En la tabla 1-2 se mencionan los factores de virulencia asociados al QS y sus principales características asociadas a la infectividad desarrollada por la *P. aeruginosa* ³⁶. Recientes hallazgos sugieren una correlación entre las concentraciones de algunos AI's con difíciles estatus en infecciones pulmonares ⁶⁶.

Esta directa relación que se muestra entre el QS y los factores de virulencia hacen de este blanco uno de los más interesantes para la búsqueda de nuevos fármacos, ya que al solo controlar esta vía podríamos detener las complicaciones clínicas producidas por este patógeno. Sin embargo, como se mencionaba antes, este mecanismo presenta múltiples factores de transcripción y una jerarquía asociada a estos.

Las recientes investigaciones sobre la jerarquía presente en el QS muestran que todos estos sistemas están interconectados entre sí, como se observa en la figura 1.7, pero ahora cuatro de estos sistemas se encuentran encima de esta jerarquía, debido a que sin ellos no se puede activar el QS, siendo estos: *Las*, *Rhl*, *Pqs* e *Iqs*.

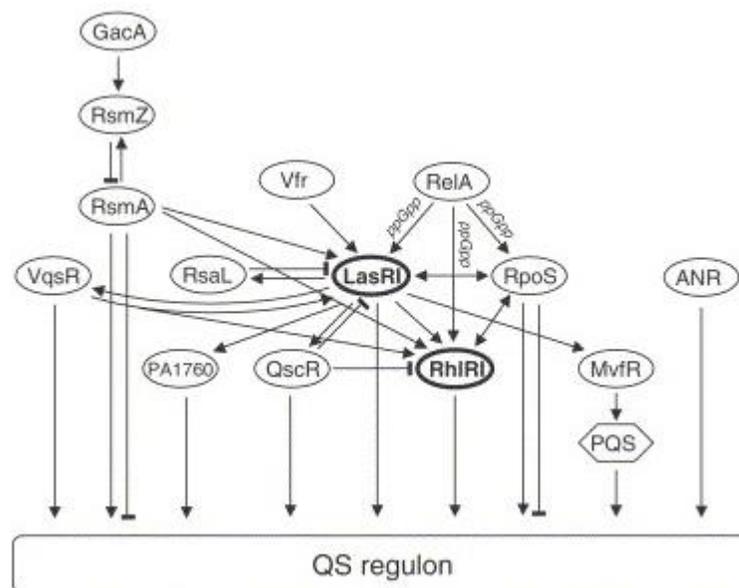


Figura 1.7: Interconexión entre los sistemas del QS. Tomada de Greenberg & Schuster ²¹.

En los últimos años se han reportado múltiples estudios de inhibidores del QS, entre las aproximaciones exploradas se encuentran péptidos ⁶⁷, moléculas y extractos de productos naturales ^{68,69}, compuestos de síntesis ^{70,71}, moléculas y extractos de productos marinos ^{72,73}, venenos animales ⁷⁴ y nanopartículas metálicas ⁷⁵, entre otras opciones.

Tabla 1.2: Factores de virulencia dependientes del QS producidos por *P. aeruginosa*. Adaptada de Rhem et al³⁶

Factor de Virulencia	Clase/Química	Síntesis	Secreción	Propiedad	Rol en la patogénesis
Piocianina	Metabolito secundario / fenacina tricíclica	Operones <i>phzA1-G1</i> y <i>phzA2-G2</i>	T2SS	Redox activo, zwitteriónico	Citotóxica / Daño a células hospederas, tisulares y del sistema inmune / inductor de apoptosis y causante de estrés oxidativo.
Pioverdina	Cromóforo unido a péptido tipo dihidroquinolina	Sintetasas peptídicas no ribosomales	Bombas de eflujo PvdRT-opmQ y MexAB-OprM	Alta afinidad por Fe(III)/ quelante férrico y fluorescente	Quelante del hierro y otros metales / crucial para la infección y el desarrollo de la biopelícula.
Elastasa-LasA	Estafilocolítica/ serin proteasa	<i>lasA</i>	La vía Sec y T2SS	Actividad proteasa y elastolítica (preferente de proteínas contenedoras de glicina)	Actividad estafilocolítica / potenciación de LasB y de las proteasas elastolíticas del hospedero.
Elastasa-LasB	Familia M4 termolisina peptidasa/ metaloproteasa	<i>lasB</i>	La vía Sec y T2SS	Actividad proteasa y elastolítica	Degradante de proteínas del hospedero (elastina, colágeno y fibrina) / crucial para la invasión de tejidos.
Alcalin proteasa (aeruginolisina)	Familia M10 peptidasa/Zinc-metaloendopeptidasa	<i>aprA</i>	T1SS	Amplia actividad proteasa	Degradante de proteínas tisulares como la laminina / causante de la necrosis hemorrágica tisular / inactivante clave de los componentes del sistema inmune.
Lectina A	Proteína tetramérica	<i>lecA</i> (o <i>pa1L</i>)	Intracelular, una pequeña fracción en la superficie celular	Galactofílica / adhesiva	Citotóxico / deterioro respiratorio de las células epiteliales / importante para el anclaje celular.
PlcB	Fosfolipasas C	<i>plcB</i>	La vía Sec y T2SS	Hidrolizante fosfatidil-colina y fosfatidiletíl-anolamina	Actividad citolítica / importante para la destrucción de membranas celulares e invasión de tejidos
Rhamnolipidos	Compuestos glicolipídicos que contienen Ramnosa.	Operón <i>rhlAB</i> y <i>rhlC</i>	-	Biosurfactante / estructura tipo detergente / actividad hemolítica	Citotóxico / Causante de la invasión y el daño tisular / importante para el mantenimiento de la arquitectura de la biopelícula y motilidad bacteriana
Exotoxina A (ToxA)	Pertenece a la familia de toxinas AB de dos componentes.	<i>toxA</i>	T2SS	Modificación de la elongación del factor-2 en células eucariotas	Uno de los factores de virulencia más tóxicos / inhibición de la síntesis de proteínas del hospedero.
Ácido cianhídrico (HCN)	Metabolito secundario.	Operón <i>hcnABC</i>	Difusible	Altamente tóxico / inhibidor potente de la Citocromo C Oxidasa y otras metaloenzimas.	Citotóxico / supresor de la respiración aeróbica por su rápida difusión en tejidos

1.2.1 Sistemas Las y Rhl

Los sistemas *Las* y *Rhl* están clasificados entre los sistemas tipo *Lux*, debido a que estos son activados por moléculas tipo acil-homoserin-lactona (AHL). Las AHL poseen un anillo lactónico conectado a una cadena lateral carbonada de tamaño variable a través de un grupo amida (Figura 1.8) ⁶⁰. El AI del sistema *Las* es N-(3-oxododecanoil)-homoserin-lactona (OdDHL), mientras que la N-butiril-homoserin-lactona (BHL) está encargada de la activación del sistema *Rhl*, pero estas no son las únicas tipo *Lux*, en realidad en la naturaleza más de 20 bacterias poseen sistemas de QS, regulados por AHL ⁷⁶.

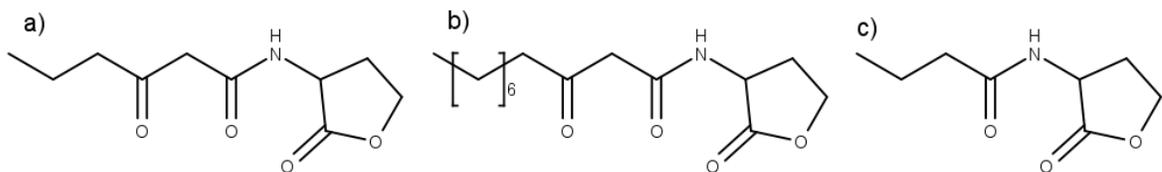


Figura 1.8: Autoinductores tipo acil-homoserin-lactona, a) AI del sistema *Lux*: N-(3-oxohexanoil)-homoserin-lactona (OHHL), b) AI del sistema *Las*: N-(3-oxododecanoil)-homoserin-lactona (OdDHL), c) AI del sistema *Rhl*: N-butiril-homoserin-lactona (BHL).

Las proteínas inductoras LasI y RhlI producen los AI's que son reconocidos por sus receptores. El sitio de unión de la lactona en las proteínas LasR y RhlR es altamente conservado, mientras que el bolsillo hidrofóbico de las cadenas acílicas varía según el tamaño de los grupos acil ⁷⁶. Una vez se realiza el acoplamiento de OdDHL con LasR y BHL con RhlR, estos complejos se dimerizan con el ADN produciendo así los factores de virulencia asociados. Los factores tipo *Lux* se encuentran en la cima de la jerarquía del QS en *P. aeruginosa* ya que son los responsables de la activación de los otros sistemas ⁶⁵. A la fecha ninguna base de datos pública posee la estructura terciaria de RhlR, sin embargo, se puede encontrar su secuencia de aminoácidos en National Center for Biotechnology Information (NCBI) con código AAC44036.1.

1.2.2 Sistema Pqs

A diferencia de los sistemas anteriormente descritos, *Pqs* posee un autoinductor quinolónico, el 3,4-dihidroxi-2-heptilquinolona (PQS), como también el precursor metabólico 2-heptil-4-hidroxiquinolona (HHQ), donde HHQ es transformado en PQS a través de la enzima PqsH. Aunque ambos son capaces de unirse al receptor PqsR, PQS es 100 veces más potente que HHQ ⁷⁷, encontrando la primera característica relacionada con la actividad sobre este FT. Ambas moléculas se muestran en la figura 1.9.

Se ha encontrado que PqsR es capaz de unirse a varios locus en el genoma de *P. aeruginosa*, incluyendo a los reguladores de *Las* y *Rhl* como a otros genes asociados al estrés oxidativo ⁷⁸. En un estudio

realizado por Starkey *et al*, en donde realizan ensayos de alta eficiencia sobre este patógeno con análogos de PQS y HHQ encuentran que estos pueden disminuir de una manera eficiente los factores de virulencia, no solo los asociados a PqsR si no también los asociados a LasR y RhlR, mostrando que esta vía puede ser efectiva para el control del QS ⁷⁹.

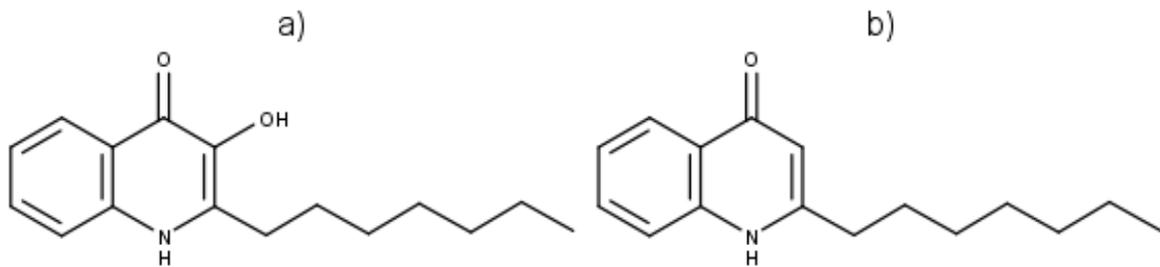


Figura 1.9: Estructuras químicas asociadas a la activación natural de PqsR, a) 3,4-dihidroxy-2-heptilquinolona (PQS), b) 2-heptil-4-hidroxiquinolona (HHQ).

1.2.3 Sistema *lqs*

En los últimos años se ha reportado una cuarta señal que genera la integración del estrés ambiental sobre la bacteria con la red QS, llamada *lqs* (*integrated quorum sensing*), que tiene como AI al 2-(2-hidroxifenil)-tiazol-4-carbaldehído (IQS) ²⁰. Según Lee *et al*, cuando existe ausencia de PQS y BHL en el medio y no se producen los factores de virulencia como piocianinas y ramnolipidos, una adición de 10 nM de IQS es capaz de restaurar la producción de estos factores de virulencia ⁶⁴. La mayoría de los factores responden a diferentes condiciones ambientales como se muestran en la figura 1.10, sin embargo, la jerarquía al desencadenar los factores de virulencia se mantiene.

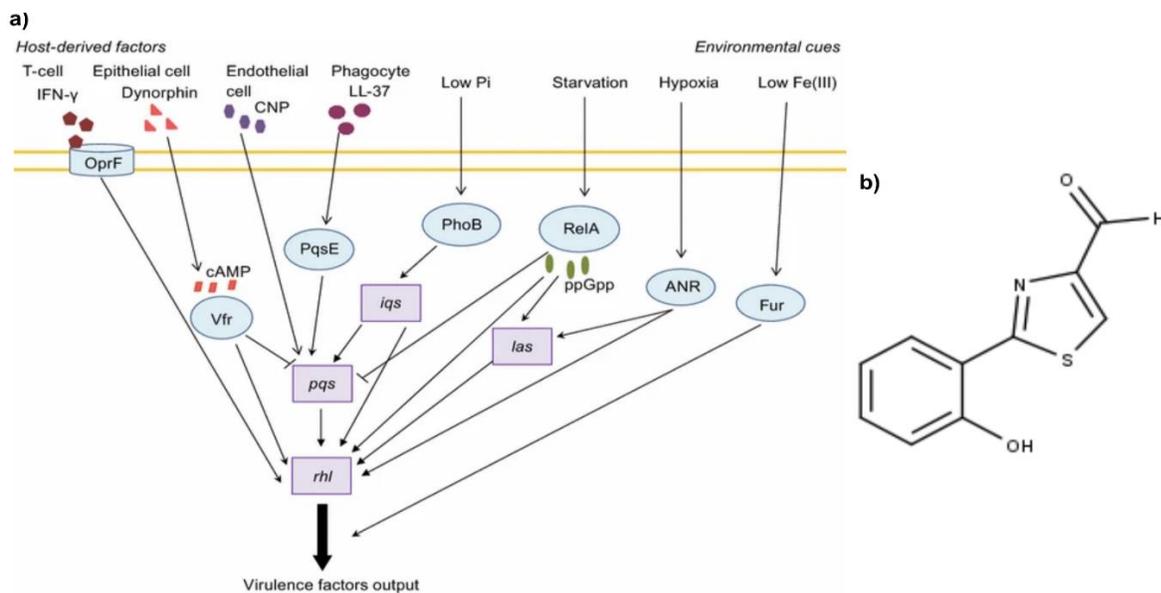


Figura 1.10: Representación esquemática de Iqs y su relación con otros FT's. a) Factores de estrés ambiental y la jerarquía de señalización, b) estructura química de IQS. Adaptado de Lee et al ⁶⁴.

En el 2019 Wang et al mostraron la relación entre IQS y la inducción de la apoptosis celular de las células del hospedero, a través del disparo de la vía de señalización POT-p53 involucrada en la ataxia telangiectasia ⁸⁰. A la fecha de escritura de esta tesis, ninguna base de datos pública posee la secuencia aminoacídica del receptor IqsR, ni la estructura terciaria del mismo, dificultando así la búsqueda de nuevas formas de bloqueo de esta vía.

El antagonismo del QS, desde los FT's *Las*, *Rhl*, *Pqs* y *Iqs*, se ha concentrado en bloquear los receptores, específicamente *LasR*, *RhIR*, *PqsR* y *IqsR*; ya que suelen haber múltiples inductores para cada sistema.

1.3 Diseño de fármacos asistido por computadora (DIFAC)

El diseño de fármacos asistido por computadora (DIFAC) tiene como objetivo poder establecer o comprender las relaciones estructura-actividad (REA), que puede ser biológica o farmacológica producidas por una molécula en específico ⁸¹. Los objetivos del DIFAC⁸² son:

- Diseñar e identificar nuevos compuestos: Obtener a través del diseño nuevas moléculas que tengan un efecto biológico concreto; aquí también se incluye la identificación de estos compuestos de las grandes BD's disponibles.
- Seleccionar candidatos: Con diferentes técnicas computacionales se pretende escoger las moléculas que tengan mayor probabilidad de éxito en la actividad biológica deseada, ahorrando recursos materiales y monetarios. Aquí es importante aclarar que un método computacional no determina su actividad biológica, para esto necesitamos ensayos experimentales.
- Optimizar moléculas líderes: Aquí se parte de un compuesto activo y se busca mejorar sus propiedades.

Aunque se tiende a pensar que el DIFAC es una metodología ya establecida, realmente este es una colección de diferentes métodos que pueden variar y ser usados según el objetivo a alcanzar; estos métodos incluyen quimioinformática, bioinformática, modelado molecular, química teórica y visualización de datos ²⁴. Esta colección de métodos puede ser dividida en dos grandes grupos, el diseño de fármacos basado en estructura (DFBE) y el diseño de fármacos basado en el ligando (DFBL).

1.3.1 Diseño de fármacos basado en la estructura (DFBE)

El DFBE se basa en el conocimiento del blanco molecular; las diferentes estrategias utilizan las estructuras terciarias de los blancos para la búsqueda de nuevas moléculas con las que puedan interactuar y generar el efecto biológico deseado, además de generar el conocimiento necesario para la optimización de moléculas a las cuales ya se les conoce su actividad biológica ⁸¹.

		¿Receptor conocido?	
		sí	no
¿Ligandos conocidos?	sí	Combinación de métodos	Diseño basado en el ligando
	no	Diseño basado en el receptor	Información experimental

Figura 1.11: Matriz de trabajo de DIFAC según la información que se tiene disponible. Tomada de Lopez-Vallejo et al ⁸¹.

Las estructuras terciarias elucidadas experimentalmente suelen obtenerse de experimentos como: difracción de rayos X, crio-microscopía electrónica y resonancia magnética nuclear; cabe resaltar que al 2020 existen más de 169963 estructuras disponibles en la base de datos pública del PDB ⁸³.

Este diseño comprende tres tipos de metodologías: acoplamiento molecular, dinámica molecular y modelado por homología. Los acoplamientos moleculares tratan de predecir la interacción que puede haber entre dos moléculas, generalmente entre una molécula pequeña y una proteína, obteniendo de este los modos de unión de la molécula frente a la proteína⁴⁰. Existen dos componentes básicos en los softwares de acoplamiento molecular: el algoritmo de muestreo y la función de puntaje. El algoritmo de muestreo es el encargado de buscar las conformaciones y orientaciones de las moléculas pequeñas, mientras que la función de puntaje trata de clasificar las conformaciones asociadas y predecir cuál es la biológicamente correcta⁸⁴. Los algoritmos de muestreo más comunes son: *Matching algorithm* (SANDOCK), *Incremental construction* (DOCK 4.0), *Monte Carlo* (AutoDock), algoritmo genético (AutoDock 4.0) y métodos jerárquicos (GLIDE)⁸⁴. En cuanto a las funciones de puntaje, éstas se basan en campos de fuerza que se componen de múltiples términos como: radios atómicos, cargas atómicas,

interacciones van der Waals, constantes dieléctricas, fuerzas de enlace (lineal, angular, tetraédrica), fuerza de desolvatación, interacciones lipofílicas y de rotabilidad ⁸⁵.

En los últimos años los estudios de acoplamiento molecular han sido una de las herramientas computacionales más usadas en el descubrimiento de fármacos; según Prieto-Martínez et al del 2000 al 2017 las publicaciones que incluyen acoplamientos moleculares han crecido en más de un 870 por ciento ⁸⁶. La mayoría de estos son usados en cribados virtuales para seleccionar candidatos a probar biológicamente.

El incremento de las capacidades de cálculo provistas por un acceso a hardware cada vez más económico y robusto ha incrementado dramáticamente el uso de la dinámica molecular (DM) en diversos campos del descubrimiento de fármacos. La DM es capaz de predecir cómo cada átomo perteneciente a una molécula (generalmente macromoléculas como proteínas o ADN) puede moverse en función del tiempo, usualmente basándose en modelos que expresan las interacciones interatómicas ⁸⁷. La función de esta dinámica se mueve en un amplio espectro del conocimiento de los procesos biomoleculares, entre los que se destacan: cambios conformacionales, unión de ligandos y plegamiento de proteínas.

Según Hollingsworth & Dror, en el 2017 aproximadamente 1000 artículos publicados en los 250 *journals* con mejor factor de impacto en el mundo incluyeron el término “*molecular dynamics*”, esto representa un enorme crecimiento si se coincidiera que 20 años antes se publicaron menos de 100. Parte de los estudios conformacionales producidos por diferentes moléculas han permitido encontrar las diferencias de modos de unión entre agonistas y antagonistas para diferentes receptores como: los androgénicos⁸⁸, κ -opioides⁸⁹ y μ -opioides^{90,91}.

El modelado por homología es uno de los métodos de predicción de estructura más usados, este permite encontrar el plegamiento de la proteína adecuado desde la secuencia de aminoácidos basado en una plantilla ya elucidada experimentalmente ⁹². El proceso se puede dividir en tres grandes etapas: identificación y selección de la plantilla, construcción del modelo y validación de este.

En la identificación de la plantilla se realiza un alineamiento múltiple de secuencias sobre la secuencia problema hasta encontrar la secuencia con estructura 3D conocida experimentalmente homóloga con respecto a la secuencia problema. Para la generación de los modelos se pueden usar varias metodologías, entre las que se clasifican en métodos de ensamblaje de cuerpo rígido, métodos de unión de segmentos, métodos de restricción espacial y métodos de evolución espacial ⁹³. Para la validación se usan múltiples softwares que muestran las interacciones espaciales de las cadenas laterales, la

longitud de los enlaces, los ángulos de torsión, los ángulos de rotación y los gráficos de Ramachandran, encontrando que estos estén acordes a lo conocido experimentalmente ⁹⁴.

El Critical Assessment of Techniques for Protein Structure Prediction (CASP) es un evento donde múltiples laboratorios de elucidación experimental de estructuras y otros dedicados a la construcción computacional, tratan de elucidar cada dos años un set de secuencias aminoacídicas, permitiendo evaluar objetivamente los servidores de modelación estructural ⁹⁵, en la edición CASP13 llevada a cabo en el 2018 se produjeron 57249 modelos siendo 17753 escogidos como los mejores por su similitud con las estructuras elucidadas experimentalmente, mostrando así la mejora de estos servidores y softwares como se ve en la figura 1-12 ⁹⁶.

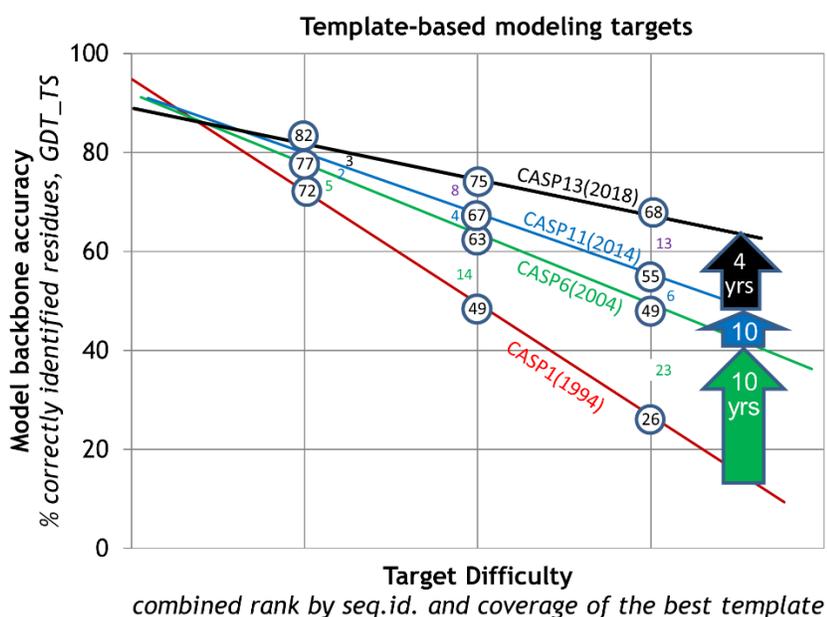


Figura 1. 12: Avance en la predicción estructural en los modelos por homología realizados en el CASP13 según el grado de dificultad. Tomada de Croll et al ⁹⁵.

1.3.2 Diseño de fármacos basado en el ligando (DFBL)

Todas las estructuras que poseen un tipo de actividad comprobadas experimentalmente son un recurso clave al realizar estudios de REA, además de permitir la identificación de otras posibles moléculas con actividad en el blanco en cuestión. Parte del atractivo para ejecutar esta estrategia es el aumento en la cantidad de información en BD's públicas, como PubChem ³¹, ChEMBL ³⁰ o Binding Database ⁹⁷. Las técnicas asociadas a este tipo de diseño son: análisis quimiinformáticos, búsqueda por similitud estructural, modelado farmacofórico, QSAR y técnicas de machine learning ²⁷.

La quimioinformática, o ciencia de la información química, es una de las más recientes disciplinas de la química teórica usada en el descubrimiento de fármacos; esta disciplina se basa en una serie de herramientas informáticas que permiten manipular una gran cantidad de moléculas a través de descriptores estructurales, topológicos o fisicoquímicos⁹⁸. Los descriptores estructurales transforman la información de la estructura química en un código numérico denominado en la literatura como huella dactilar molecular o *molecular fingerprint*, ya sea en dos o tres dimensiones.

Los *molecular fingerprints* pueden ser clasificados en tres categorías: basados en características, topológicos e híbridos. Los *fingerprints* basados en características arman su código en base a una lista de subestructuras presentes en la molécula, el tamaño del código depende de la cantidad de características presentes como se ve en la figura 1-13a⁹⁹; uno de los problemas de este tipo de *fingerprint* es que tiene dificultades para distinguir moléculas con los mismos grupos funcionales, en diferente posición¹⁰⁰. Los *fingerprints* topológicos (Figura 1-13b) generan un bit por cada subestructura presente en la molécula, determinada por un radio específico sobre cada átomo; esto genera que cada molécula tenga una mayor cantidad de bits y que ellos puedan distinguir entre moléculas similares debido a que tienen en cuenta la conectividad de la misma¹⁰⁰.

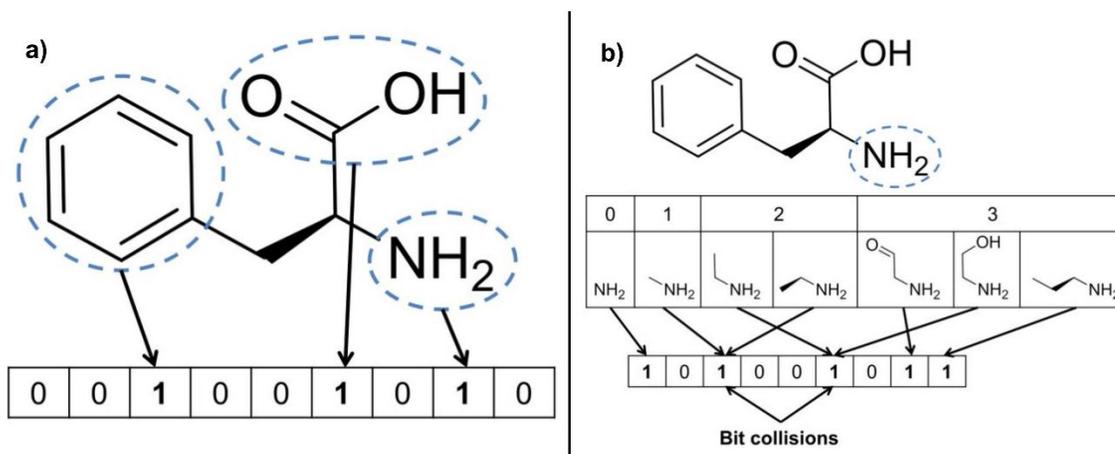


Figura 1. 13: Representación esquemática de los *molecular fingerprints*, a) *fingerprint* basado en características, b) *fingerprint* topológico. Adaptado de Bajusz et al¹⁰⁰.

Al poder representar las moléculas como códigos binarios, la búsqueda por similitud estructural en grandes BD's resulta ser mucho más rápida y útil, a partir de esta metodología Raymond et al pudieron establecer una base de datos con los compuestos químicos presentes en el universo, a partir de 17 átomos con base de C,N,O,S y halógenos, encontrando la sorprendente cantidad de 166 mil millones de compuestos aproximadamente, cabe aclarar que esta construcción se hizo bajo reglas retrosintéticas dándole un sentido lógico a los compuestos¹⁰¹.

En 1909 Paul Ehrlich introduce la primera definición del término farmacóforo, combinando las palabras en latín “*phoros*” que significa que lleva y “*pharmakon*” que significa droga o medicamento; actualmente podemos definir al farmacóforo como un marco molecular que posee las características cruciales para que una molécula produzca una actividad biológica¹⁰². De acuerdo con la anterior definición se tienen que las características típicas que posee un farmacóforo son: donadores y aceptores de puentes de hidrogeno, cargas positivas y negativas, características hidrofóbicas y anillos aromáticos. Hoy en día podemos producir modelos farmacofóricos a partir de un grupo de ligandos, alineándolos y tomando los puntos comunes o basados en la interacción de los mismos con el sitio de unión¹⁰³.

Estos modelos pueden ser usados en cribados virtuales, donde se realizan alineamientos múltiples con el fin de encontrar moléculas que posean las características farmacofóricas en un espacio de tres dimensiones, como lo realizado por Liu *et al*, donde generan un modelo farmacofórico a partir de amino alcoholes y encuentran 2 moléculas biológicamente activas contra *echinococcus*¹⁰⁴.

El *quantitative structure-activity relationships* (QSAR) utiliza modelos estadísticos para establecer una ecuación matemática que correlaciona una estructura o una familia de compuestos con su actividad biológica²⁷. Por otro lado, el *machine learning* o aprendizaje de maquina es definido por Ethem Alpaydin como un programa computacional que optimiza la resolución de un problema usando información de ejemplo o experiencia pasada¹⁰⁵; en el descubrimiento de fármacos podemos aprovechar esta poderosa herramienta para la clasificación de un compuesto determinado según una propiedad deseada (actividad biológica, toxicidad, farmacocinética, entre otras).

Este aprendizaje se clasifica en dos grandes grupos, aprendizaje supervisado o no supervisado, en el aprendizaje de maquina supervisado se tienen modelos como análisis de regresión lineal, *k*-nearest neighbor (*k*-NN), modelos bayesianos, máquinas de soporte vectorial (SVM), bosques aleatorios y redes neuronales. Mayr *et al* en el 2018 realizaron una comparación de diferentes modelos de predicción de actividad biológica sobre la BD publica de ChEMBL, aquí encontraron que los métodos de aprendizaje profundo (redes neuronales) poseen una precisión superior sobre a otros métodos¹⁰⁶.

En el aprendizaje no supervisado tenemos modelos como el análisis de componentes principales (PCA), la incrustación de vecinos estocásticos distribuidos en *t* (*t*-SNE), entre otros. Éstos nos permiten clasificar grandes grupos de moléculas en *clusters* según su similitud estructural, siendo realmente útiles para la visualización del espacio químico. En los últimos años la exploración de este espacio ha tenido un gran impacto en el descubrimiento de fármacos, destacando aquí los gráficos de constelación creados por Naveja *et al* en donde se asocia el espacio químico con fragmentos retrosintéticos y una propiedad determinada, lo que permite encontrar fácilmente núcleos con actividad¹⁰⁷.

2 Metodología

En la figura 2.1 se muestra la metodología aplicada en esta tesis.

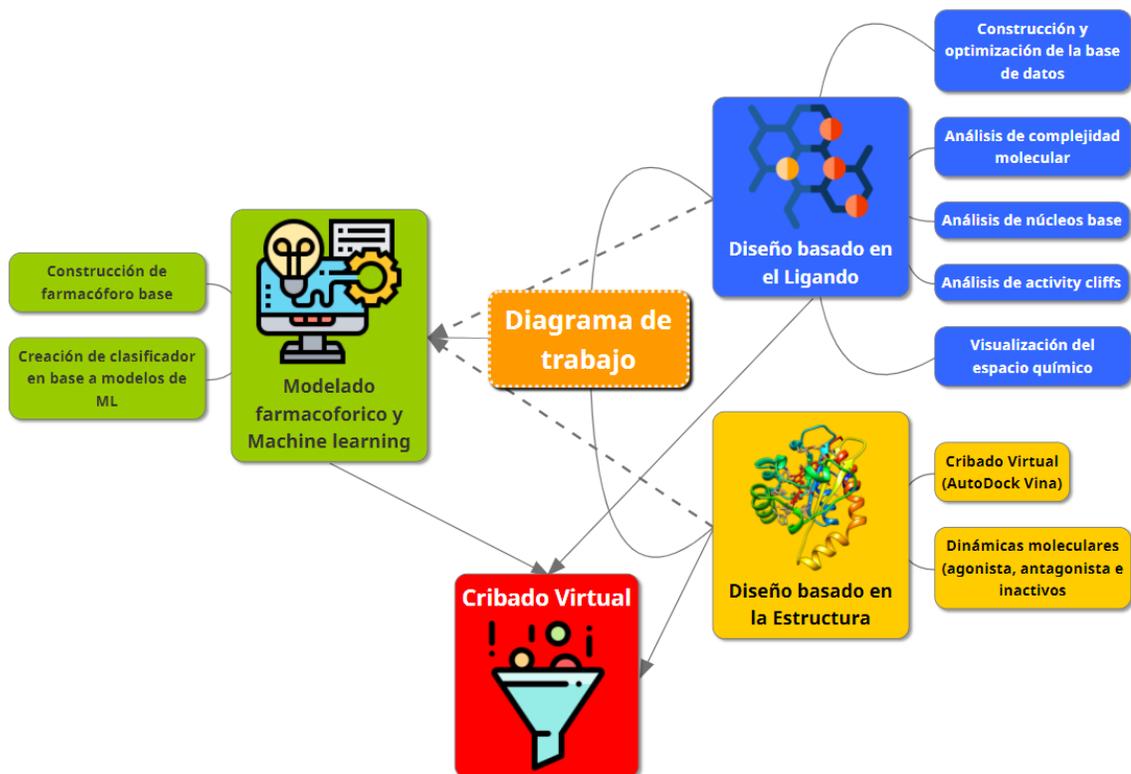


Figura 2.1: Diagrama general de la metodología aplicada en el estudio.

2.1 Diseño basado en el ligando

2.1.1 Construcción y preparación de las bases de datos

Las BD's fueron construidas con información del ChEMBL - versión 26, PubChem y artículos científicos; para ChEMBL se usaron los siguientes identificadores: LasR - CHEMBL1075207, PqsR - CHEMBL2424500 y RhIR - CHEMBL3112386, en PubChem Bioassay la búsqueda fue hecha directamente con los AI's. Las palabras clave "LasR", "PqsR", "RhIR" y "quorum sensing" fueron usadas para la búsqueda de moléculas en los artículos científicos bajo el metabuscador de SciFinder; todos los artículos se filtraron por los reportes de actividad biológica teniendo en cuenta solo "EC₅₀" e "IC₅₀", siendo seleccionados los artículos publicados entre 1998 y 2019. Esta búsqueda fue llevada a cabo entre enero y Julio del 2019.

Las moléculas obtenidas en las BD's fueron clasificadas como "activas" e "inactivas" usando como valor de corte 10 μM para LasR y PqsR, mientras 12 μM fue usado para RhlR; estos valores se escogieron con base al valor resultante entre la resta de la media y la desviación estándar (SD) de las actividades biológicas encontradas. Luego, los siguientes pasos fueron aplicados para la preparación de los compuestos o curado de las BD's: 1) el fragmento más grande en cada molécula fue seleccionado, 2) los compuestos no orgánicos se removieron, 3) las cargas fueron neutralizadas por adición y/o eliminación de hidrógenos donde fue posible, 4) se aseguró que los grupos ácidos fueran ionizados primero en las moléculas parcialmente cargadas, 5) se eliminó la estereoquímica, 6) el tautómero más frecuente fue seleccionado, 7) las moléculas duplicadas fueron descartadas. Esta curación fue realizada con las librerías de Python RDKit 2019.09.1¹⁰⁸ y MolVS 0.1.1¹⁰⁹.

2.1.2 Complejidad molecular

La complejidad molecular ha sido asociada con la selectividad por el blanco y el éxito en diferentes etapas clínicas¹¹⁰⁻¹¹². Para este estudio, la fracción de átomos sp^3 (F_{sp^3}) y la fracción centros quirales (FCC) fueron usados como métricas de complejidad topológica¹¹³, además la complejidad de Bertz¹¹⁴ y DataWarrior¹¹⁵ basada en métodos de grafos-teóricos que también fueron aplicadas. Todas las moléculas se clasificaron en tres grupos usando el método propuesto por Méndez-Lucio et al¹¹³, siendo simple si su valor de complejidad es \leq a la media - 1 desviación estándar (SD), intermedia si la media - 1 SD es $<$ que el valor de complejidad y $>$ a la media + 1 SD o compleja si el valor de complejidad es \geq media + 1 SD. Las métricas fueron calculadas con RDKit 2019.09.1¹⁰⁸ y DataWarrior 5.0.0¹¹⁵, mientras la librería de Python Pandas 1.1.1¹¹⁶ fue empleada para la clasificación de las BD's según su complejidad.

2.1.3 Análisis de núcleos base

Los núcleos base o *scaffolds* pueden ser definidos como el soporte de una molécula bioactiva¹¹⁷⁻¹¹⁹, por esta razón los núcleos base son una herramienta intuitiva e informativa para explorar las REA y las relaciones de estructura-actividad múltiple (SmART)¹²⁰ de las BD's moleculares. En este trabajo, se exploró la conexión existente entre los *scaffolds* y la actividad biológica sobre las BD's de los FT's. Para la generación de estos, se usó la definición propuesta por Bemis & Murcko¹²¹, donde los núcleos base de un compuesto son todos los anillos que lo componen y sus respectivos conectores moleculares. La herramienta implementada para llevar este análisis a cabo fue la librería de Python RDKit 2019.09.1

Todos los núcleos fueron clasificados de acuerdo con el tipo de actividad (agonista, antagonista o inactivo). Con el *scaffold* más frecuente en cada BD se realizó el análisis de máxima común subestructura (MCS) ^{122,123}, permitiendo encontrar la mayor cantidad de relaciones entre los núcleos y la actividad biológica para los FT's. Toda esta información fue colectada a través de la herramienta de Python RDKit 2019.09.1 ¹⁰⁸.

2.1.4 Análisis de *activity cliffs*

Los *activity cliffs* o acantilados de actividad son pequeños cambios en la estructura que están asociados con grandes cambios en la actividad biológica ¹²⁴⁻¹²⁷. La identificación de estos es un análisis relevante e informativo de las REA, porque permite especificar que modificaciones tienen una fuerte influencia sobre la actividad biológica. En este trabajo se trata de hallar la posible presencia de *activity cliffs* en las BD's, para esto se calcularon los 2D *molecular fingerprints* (MACCS Keys 166-Bits, ECFP2, ECFP3, ECFP4 y *Path Length*), luego se determinó la similitud molecular usando la métrica del coeficiente de Tanimoto, para finalmente caracterizar los *activity cliffs* a través del consenso (*activity cliffs* comunes en todos los *fingerprints*), método propuesto por Medina-Franco et al ¹²⁸. Este análisis se realiza bajo el uso de la librería de Python RDKit 2019.09.1 ¹⁰⁸ y el software MayaChemTools ¹²⁹.

2.1.5 Visualización del espacio químico

La representación de un conjunto de moléculas en un gráfico de dos o tres dimensiones basados en un grupo de descriptores predefinidos se conoce como visualización del espacio químico ¹³⁰⁻¹³³. Para este **trabajo el análisis estadístico denominado “t-distributed stochastic neighbor embedding (t-SNE)”** es usado para la representación del espacio químico ocupado por las BD's, éste se realizó bajo las siguientes condiciones: perplejidad igual a 10, tasa de aprendizaje de 100, métrica de Jaccard y el descriptor *molecular fingerprint* ECFP4. El análisis se realizó con las librerías de Python RDKit 2019.09.1 y SciKit Learn 0.23.1.

Adicionalmente se elaboraron los gráficos de constelación basados en el trabajo de Naveja & Medina-Franco ¹⁰⁷ para todas las BD's, estos gráficos integran una metodología de series de análogos con la visualización del espacio químico, asociando así los análogos moleculares a compuestos determinados y sus actividades biológicas.

2.2 Diseño basado en la estructura

2.2.1 Alineamientos de las estructuras cristalográficas de los FT's

Con las estructuras cristalográficas de los FT's almacenadas en el *Protein Data Bank* se realizó un alineamiento estructural. **En esta base de datos se encuentran complejos de los FT's con moléculas que** generan diversas respuestas biológicas frente a los mismos; dicho análisis permite caracterizar las partes de la estructura proteica que cambian según la molécula con la cual forma el complejo. Los alineamientos y las imágenes respectivas fueron realizadas con el software libre UCSF Chimera ¹³⁴.

2.2.2 Construcción del modelo de RhIR

Por la ausencia de una estructura elucidada experimentalmente para el FT RhIR, se realizó la construcción de su modelo de estructura terciaria mediante modelado por homología, para ello se utilizó la herramienta *basic local alignment search tool* (BLAST) ¹³⁵, que permite encontrar a través del porcentaje de identidad proteínas similares al objetivo. Luego, al utilizar como plantilla base la estructura terciaria con mayor porcentaje de identidad se generó diferentes modelos con los servidores I-TASSER ¹³⁶, ROBBETA ¹³⁷, MODELLER ¹³⁸, PHYRE2 ¹³⁹ y SWISS-MODEL ¹⁴⁰.

Estos modelos fueron evaluados con las siguientes métricas: RMSD con respecto a la referencia, QMEANDisCo, LGscore, MaxSub y porcentaje de acomodación en el gráfico de Ramachandran. De los anteriores el modelo con las mejores características fue seleccionado para continuar con los análisis posteriores.

2.2.3 Acoplamiento moleculares automatizados

Para los acoplamiento moleculares se utilizaron las estructuras cristalográficas con código PDB: 3IX3 para LasR, código PDB: 4JVD para PqsR y el modelo de homología construido para RhIR. Las estructuras proteicas se prepararon bajo el siguiente procedimiento: 1) se seleccionó una única cadena, 2) se removió el ligando cristalográfico, el solvente y los iones no acoplados, 3) se cambiaron las selenometioninas presentes por metioninas, 4) se adicionaron las cadenas incompletas usando la librería de rotámeros propuesta por Dunbrack & Shapovalov ¹⁴¹, 5) se adicionaron los hidrógenos faltantes teniendo en cuenta las interacciones de la cadena base y la cadena secundaria, además de los factores estéricos, 6) se agregaron los estados de protonación correspondientes a los aminoácidos cargados usando los modelos de Gasteiger ¹⁴², como se observa en la figura 2.2. Este procedimiento fue realizado con el módulo DockPrep del software UCSF Chimera ¹³⁴.

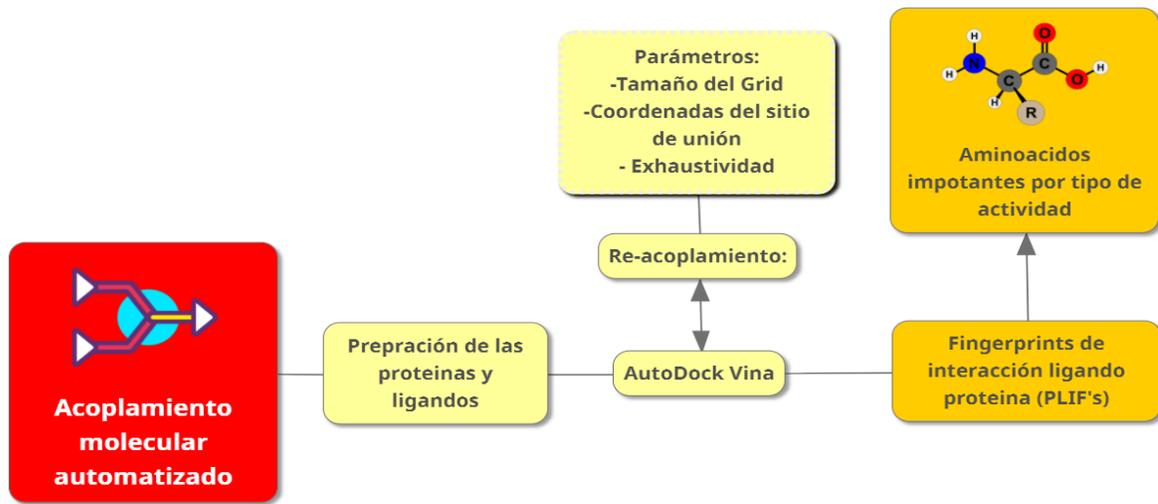


Figura 2.2: Metodología general en la aplicación de los acoplamientos moleculares automatizados.

A los ligandos cristalográficos se les realizó una minimización energética utilizando el campo de fuerza MMFF¹⁴³, luego con las coordenadas que poseían éstos en los archivos PDB, se realiza el re-acoplamiento con una caja de 15 Å³, obteniendo así los parámetros iniciales para realizar los acoplamientos moleculares posteriores. Con las moléculas de la base de datos construida para cada FT se realizó el acoplamiento sobre su respectivo blanco, buscando poder establecer relaciones estructura actividad (REA) frente a las interacciones que se obtienen por complejo. Para este procedimiento se utilizó el software AutoDock Vina 1.1.2 (Vina)¹⁴⁴.

Para analizar las interacciones de los acoplamientos entre las moléculas de las BD's, se utilizó la metodología de *protein-ligand interaction fingerprints (PLIF's)* como se puede observar en la figura 2.3, que evalúa la presencia de contactos entre los residuos proteicos y las moléculas, siendo clasificados en los siguientes tipos: donador y receptor de puente de hidrogeno, cargados positiva o negativamente, hidrofóbicos, aromáticos o polares.

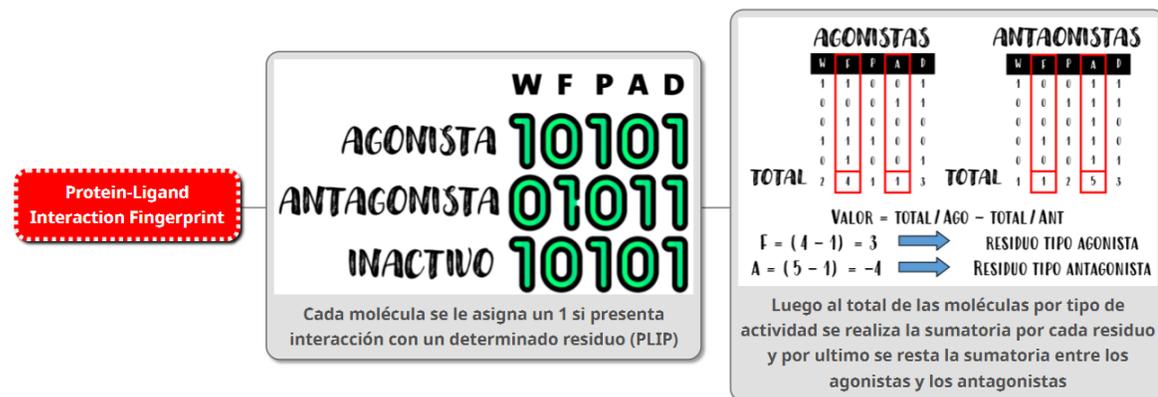


Figura 2.3: Diagrama ilustrativo de la metodología implementada en el análisis y la clasificación de un residuo importante para la actividad agonista o antagonista, en base a la diferencia del acumulado de las interacciones por tipo de actividad

Para este estudio solo se tuvieron en cuenta las interacciones con una distancia menor a 4.5 Å, siendo obtenidas con la herramienta *protein-ligand interaction profiler* 2.1.6¹⁴⁵. Una vez se consiguieron, se procedió a revisar cuales interacciones son predominantes según el tipo de actividad (Agonista - Antagonista) utilizando la siguiente expresión matemática:

$$TA = \frac{\Sigma_{iag}}{n_{ag}} - \frac{\Sigma_{ian}}{n_{an}} \quad (2.1)$$

El tipo de interacción (TA) está dado por la diferencia entre la fracción de la sumatoria de las interacciones que tuvo x residuo con las moléculas agonistas (Σ_{iag}) y el número de compuestos agonistas (n_{ag}) contra la fracción de la sumatoria de las interacciones que tuvo el mismo x residuo con las moléculas antagonistas (Σ_{ian}) y el número de compuestos antagonistas (n_{an}). Si el valor de TA es positivo nos indica que ese residuo tiende a tener un mayor número de interacciones con las moléculas agonistas mientras que si es negativo esa interacción se encuentra con mayor frecuencia en los complejos formados con los compuestos antagonistas. Este procesamiento se realizó con las librerías de Python RDKit 2019.09.1¹⁰⁸ y Pandas 1.1.1¹¹⁶, mientras que las gráficas resultantes se obtuvieron con la librería Matplotlib 3.3.3¹⁴⁶.

2.2.4 Simulaciones de dinámica molecular

Para realizar las simulaciones de dinámica molecular (DM) se seleccionaron el grupo de moléculas encontradas en el análisis de MCS el cual provee una molécula por cada tipo de actividad (agonista, antagonista e inactivo) **sobre los FT's** además del AI, llevando a cabo 12 dinámicas moleculares, 4 por **cada FT's**.

Estas DMs se llevaron a cabo bajo el software DESMOND¹⁴⁷, los complejos obtenidos de los estudios de acoplamiento molecular fueron solvatados en una caja ortorrómbica con el modelo de agua TIP3P, teniendo una distancia margen de 8.0 Å entre la esquina de la caja y el complejo; los términos electrostáticos fueron definidos con el algoritmo *particle mesh Ewald* (PME) usando 12.0 Å como valor de corte, adicionalmente iones sodio y cloruro se colocaron aleatoriamente hasta llegar a la neutralidad del sistema.

Todas las DMs siguieron el mismo procedimiento, donde se ejecutaron una etapa de minimización, calentamiento, equilibrio con un tiempo de 2 ns mientras que la producción duro 50 ns, estas se

Llevaron a cabo bajo condiciones isotérmicas e isobáricas a 300 K y 1 atm de presión, el control de estas variables se hizo bajo el uso de los algoritmos de Langevin y Berendsen respectivamente. La visualización se realizó con VMD ¹⁴⁸ y USCF Chimera ¹³⁴, mientras el análisis de trayectorias fue hecho con la librería de Python MDAnalysis ¹⁴⁹ y Maestro 11.4 ¹⁴⁷. La similitud de las trayectorias se calculó bajo la métrica de Hausdorff ¹⁵⁰, que es definida como la distancia entre dos puntos espaciales en el mismo intervalo de tiempo.

2.3 Modelado farmacofórico

Para las 10 moléculas con los mejores valores de EC₅₀ y IC₅₀ **en cada uno de los FT's se construyó un** modelo de farmacóforo, usando las conformaciones de las moléculas después de ser acopladas frente a sus blancos respectivos; de aquí se extraen las siguientes propiedades: aceptores y donadores de puente de hidrogeno, cargas positivas y negativas, características hidrofóbicas, aromáticas e interacciones con ligando metálico; estas características debían estar presentes en más del 50 % de las moléculas acopladas y a una distancia de no más de 4.5 Å. Los modelos fueron construidos con el programa MOE 2018 ¹⁵¹.

2.4 Clasificadores en base a *machine learning*

La construcción de los clasificadores fue realizada con las moléculas presentes en las BD's **de los FT's**, dichos compuestos se clasificaron como activos o inactivos, luego se representaron con 6 tipos de *molecular fingerprints* (ECFP2, ECFP4, RDKit, MACCS keys, MHFP y SECFP) que permitieron desglosar **las características estructurales en '1' y '0' mostrando la presencia o ausencia** de estos descriptores estructurales. Con dichos *fingerprints* y sus respectivas etiquetas (activo o inactivo) se entrenaron los siguientes modelos de aprendizaje supervisado: *K-nearest Neighbors Classifier* (k-NN), *Random Forest Classifier* (rf), *Logistic Regression* (lr), *Extra Trees Classifier* (et), *Light Gradient Boosting Machine* (lightgbm), *Extreme Gradient Boosting* (xgboost), *CatBoost Classifier* (catboost), *Ridge Classifier* (ridge), *Gradient Boosting Classifier* (gbc), *Decision Tree Classifier* (dt), *Supported Vector Machine* (svm), *Ada Boost Classifier* (ada), *Linear Discriminant Analysis* (lda), *Native Bayes* (nb), *Quadratic Discriminant Analysis* (qda).

Con cada uno de los *fingerprints* se evaluaron todos los modelos anteriores, dividiendo los datos iniciales entre el conjunto de entrenamiento (70 %) y el conjunto de prueba (30 %); el conjunto de entrenamiento a su vez fue dividido en 10 partes iguales usando la metodología *StratifiedKFold* para realizar una validación cruzada utilizando 10 iteraciones diferentes. El desempeño de los modelos fue evaluado con las métricas: exactitud, área bajo la curva (AUC), *recall*, precisión, F1, coeficiente Kappa

de Cohen (kappa) y el coeficiente de correlación de Mathews (MCC), cuyas expresiones matemáticas se muestran la tabla 2.1.

Tabla 2.1: Métricas usadas para evaluar el desempeño de los modelos de *machine learning*

Métrica	Ecuación
Exactitud	$\frac{VP + VN}{N}$
Área bajo la curva (AUC)	$VP \sum_s VN$
Recall	$\frac{VP}{VP + FN}$
Precisión	$\frac{VP}{VP + FP}$
F1	$2 * \frac{\text{precisión} * \text{recall}}{\text{precisión} + \text{recall}}$
Kappa	$1 - \frac{1 - P_0}{1 - P_e}$ $P_e = \frac{1}{N} \sum_k n_{k1} n_{k2}$
MCC	$\frac{VP * VN - FP * FN}{\sqrt{(VP + FP)(VP + FN)(VN + FP)(VN + FN)}}$

Donde VP son los verdaderos positivos predichos por el modelo, VN los verdaderos negativos, FP los falsos positivos, FN los falsos negativos, N el total de moléculas, P_0 la ratio de los clasificados correctamente, P_e la probabilidad hipotética de clasificar todo correctamente, K el número de categorías, n_{ki} el número de veces de i predicho en la categoría k . Luego de evaluar estas métricas a los mejores modelos se les aumentó el número de iteraciones (50, 150, 300 y 500) con el fin de optimizar los hiperparámetros, hasta llegar al mejor resultado. Todo lo anterior se realizó con las librerías de Python RDKit 2019.09.1¹⁰⁸ y PyCaret 1.0.0¹⁵².

2.5 Cribado virtual

Para el cribado virtual se usó la librería de moléculas QuiProNaB, en ella se depositan compuestos extraídos de fuentes vegetales. Esta contiene estructura molecular (SMILES), especie y familia de la fuente de origen además de la información bibliográfica correspondiente. Dicha BD se preparó de igual manera que las BD's de los FT's.

Luego esta base de datos fue cribada usando los dos mejores clasificadores obtenidos de los modelos de *machine learning*. Por último, A partir de los resultados previos se puntuaron los compuestos obtenidos bajo la fórmula 2.2, donde al calcular los *fingerprints* de interacción proteína-ligando para las moléculas resultantes del paso anterior; se restaron las interacciones con los aminoácidos asociados a la actividad antagonista menos los residuos encontrados claves para la actividad agonista, con el fin de dejar las moléculas que tengan la mayor probabilidad de ser antagonistas. Todo lo anterior fue realizado con las librerías de Python RDKit 2019.09.1 ¹⁰⁸, PyCaret 1.0.0 ¹⁵² y Pandas 1.1.1 ¹¹⁶ además de los softwares: *protein-ligand interaction profiler* ¹⁴⁵ y AutoDock Vina ¹⁴⁴.

$$P = Aa_{ant} - Aa_{ago} \quad (2.2)$$

3 Resultados

3.1 Diseño de fármacos basado en el ligando (DFBL)

3.1.1 Construcción de las bases de datos

En la tabla 3.1 se muestran la proporción compuestos encontrados para los FT's en las diferentes fuentes de búsqueda. En ChEMBL 173 moléculas poseían actividad sobre LasR, de las cuales 82 tienen valores de IC_{50} y EC_{50} ; para PqsR se encontraron 29 moléculas, 8 estaban ligados con la actividad EC_{50} ; mientras que para RhIR se hallaron 56 moléculas y de estas 4 tenían actividad IC_{50} . PubChem Bioassay almacenaba 87, 10 y 17 compuestos con actividad sobre LasR, PqsR y RhIR respectivamente, el resto de las moléculas se obtuvieron de artículos consultados. Más del 70 % son compartidas entre todas las fuentes de consulta; adicionalmente de las BD's generadas, LasR tiene el mayor número de moléculas y el mayor número de moléculas activas.

Tabla 3.1: Moléculas presentes en las BD's

	Total	Agonista		Antagonista		Inactivo	
		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
RhIR	47	12	25,6	7	14,9	28	59,5

El valor de actividad de todos los compuestos se expresará en pIC_{50} o pEC_{50} , que comprende el menos logaritmo en base diez de la actividad reportada en una concentración molar.

3.1.2 Complejidad molecular

La selectividad es uno de los mayores desafíos en el descubrimiento de fármacos, en este sentido diversos autores proponen la asociación entre una alta complejidad molecular y una baja promiscuidad por los blancos. Por eso en este trabajo al clasificar las moléculas según su complejidad se intenta mostrar la selectividad que podrían tener estas **por los FT's**.

El análisis de complejidad molecular fue realizado mediante el grafico de cajas mostrados en la figura 3.1, aquí se relacionan las cuatro métricas de complejidad por FT, para mostrar su comportamiento múltiple.

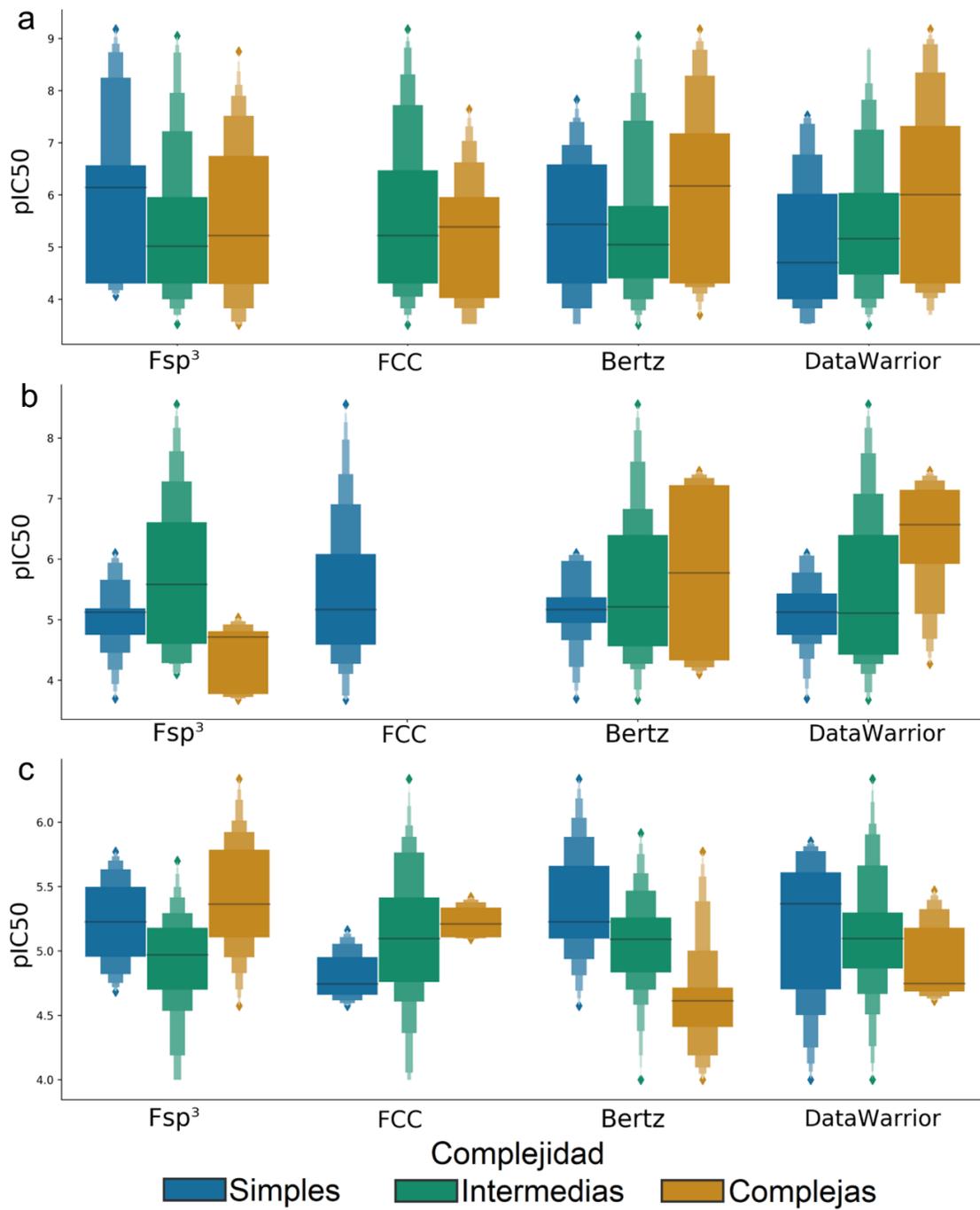


Figura 3.1: Distribución de los valores de actividad biológica para las tres categorías de complejidad en las BD's, donde a) representa los compuestos presentes en LasR, b) PqsR y c) RhIR. En el eje-x se encuentran las cuatro métricas y en el eje-y la distribución de los valores pIC₅₀, la línea horizontal en los gráficos representa la mediana del valor de complejidad.

Para LasR (Figura 3.1a), una parte importante de la base de datos fue clasificada como compleja por todas las métricas, sugiriendo una alta selectividad por el blanco, sin embargo, como algunos compuestos están clasificados en la categoría de simples e intermedios es recomendable llevar a cabo los ensayos biológicos de selectividad que permitan corroborar esta teoría.

En la figura 3.1b se observa la complejidad molecular encontrada para los compuestos presentes en la BD de PqsR, en esta se clasifican a las moléculas en la categoría de simples, relacionando la falta de selectividad de estas moléculas por el blanco. Los compuestos presentes en RhIR tienden a ser clasificados con complejidad intermedia mostrando así la posible promiscuidad sobre diferentes blancos por las moléculas presentes en esta BD (Figura 3.1c).

3.1.3 Análisis de núcleos base

La frecuencia de los 5 núcleos base más recurrentes en las BD's se muestra a la izquierda de la figura 3.2, mientras su estructura se encuentra al lado derecho de la misma figura. Esta recurrencia varía acorde al tamaño de cada BD, sin embargo, estos cinco *scaffolds* representan alrededor del 75 % del **total de los compuestos en todas las BD's, resaltando aquí que un único núcleo no puede ser asociado** a un tipo de actividad biológica (agonista, antagonista o inactivo). Los 5 *scaffolds* más recurrentes en la BD de LasR se caracterizan por estar formados de anillos de 5 o 6 miembros, además estos suelen poseer algún tipo de actividad siendo la agonista la más frecuente.

Los sistemas trifenílicos presentes en el segundo y cuarto núcleo de la figura 3.2 a) cambian solo en la posición de la amida dentro de la cadena que conecta a los anillos, más sin embargo este cambio genera que el cuarto *scaffold* posea moléculas con actividad antagonista. Adicionalmente, cabe resaltar que el primer y quinto núcleo basados en el anillo lactónico y común en el AI, presentan múltiples tipos de actividad, lo que impide asociar este *scaffold* con el antagonismo sobre este FT. Los núcleos base más frecuentes de la BD PqsR están compuestos de dos anillos de seis miembros fusionados, característicos del AI natural. Todas las moléculas con el anillo quinoleínico (SC2) tienen actividad menor a 10 μM (Figura 3.2b), haciendo de este un buen candidato para la búsqueda de nuevas moléculas con actividad; cabe mencionar que los dos *scaffolds* más frecuentes de esta BD representan a más del 50 % de las moléculas, evidenciando la poca diversidad estructural de estos compuestos.

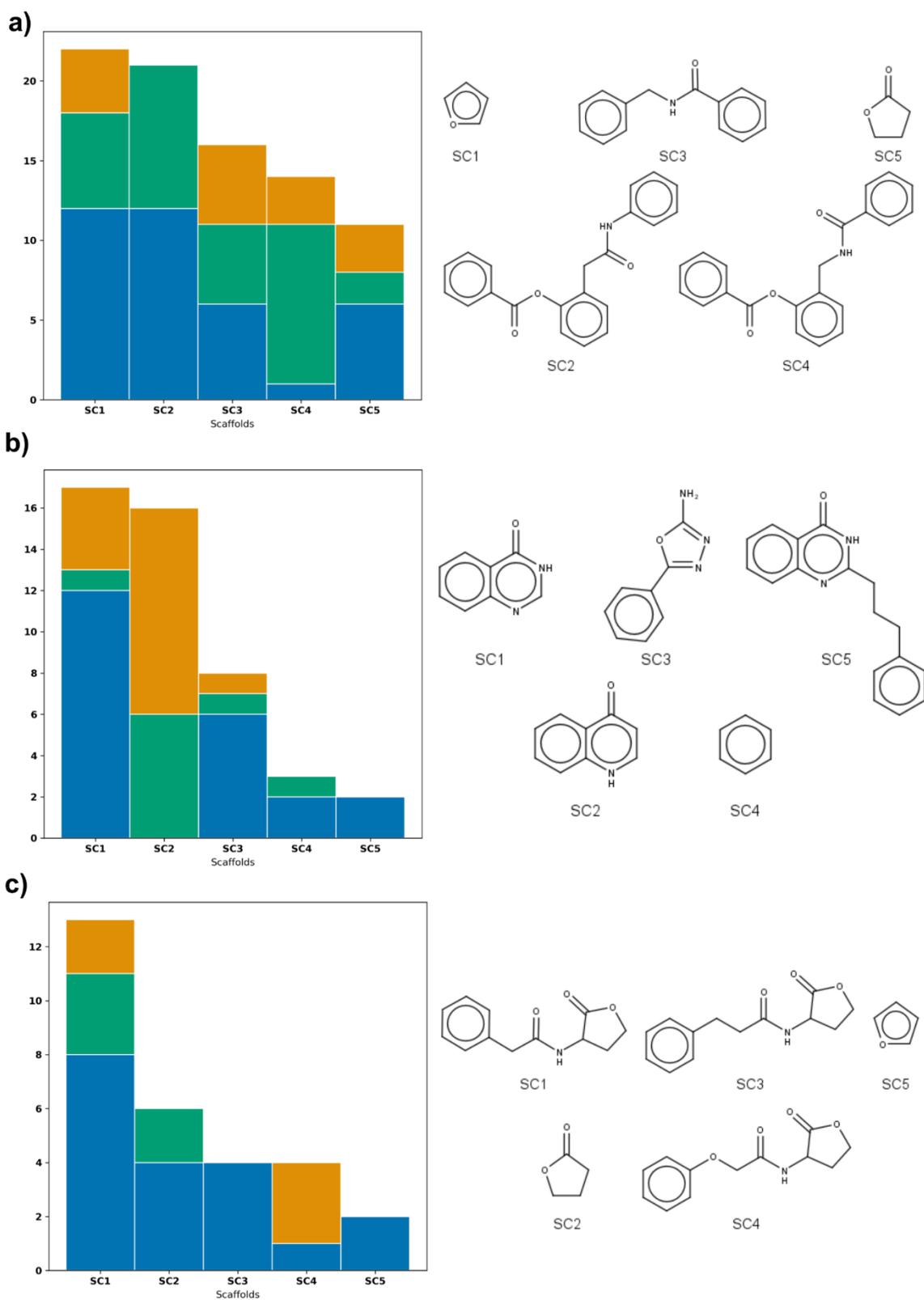


Figura 3.2: Núcleos base tipo Bemis-Murcko más recurrentes en las bases de datos, a la izquierda se muestra la frecuencia de estos y a la derecha la estructura. A) representa la BD de LasR, b) la BD de PqsR y c) la BD de RhIR; el tipo de actividad es mostrado según el color de las barras, azul – inactivo, verde – agonista y amarillo – antagonista.

La actividad biológica de las moléculas asociadas a los primeros cinco núcleos encontrados en la BD de RhIR es mucho menor que la presentada por compuestos en las **otras BD's**, aquí destaca que, aunque el SC4 y el SC2 son similares entre sí, el cambio de un C por O en la cadena que conecta al fenilo provoca que compuestos con este núcleo presenten actividad antagonista (Figura 3.2c).

El análisis de máxima común subestructura (MCS) permitió encontrar pequeños cambios estructurales entre moléculas con diferentes tipos de actividad, como se puede observar en la figura 3.3.

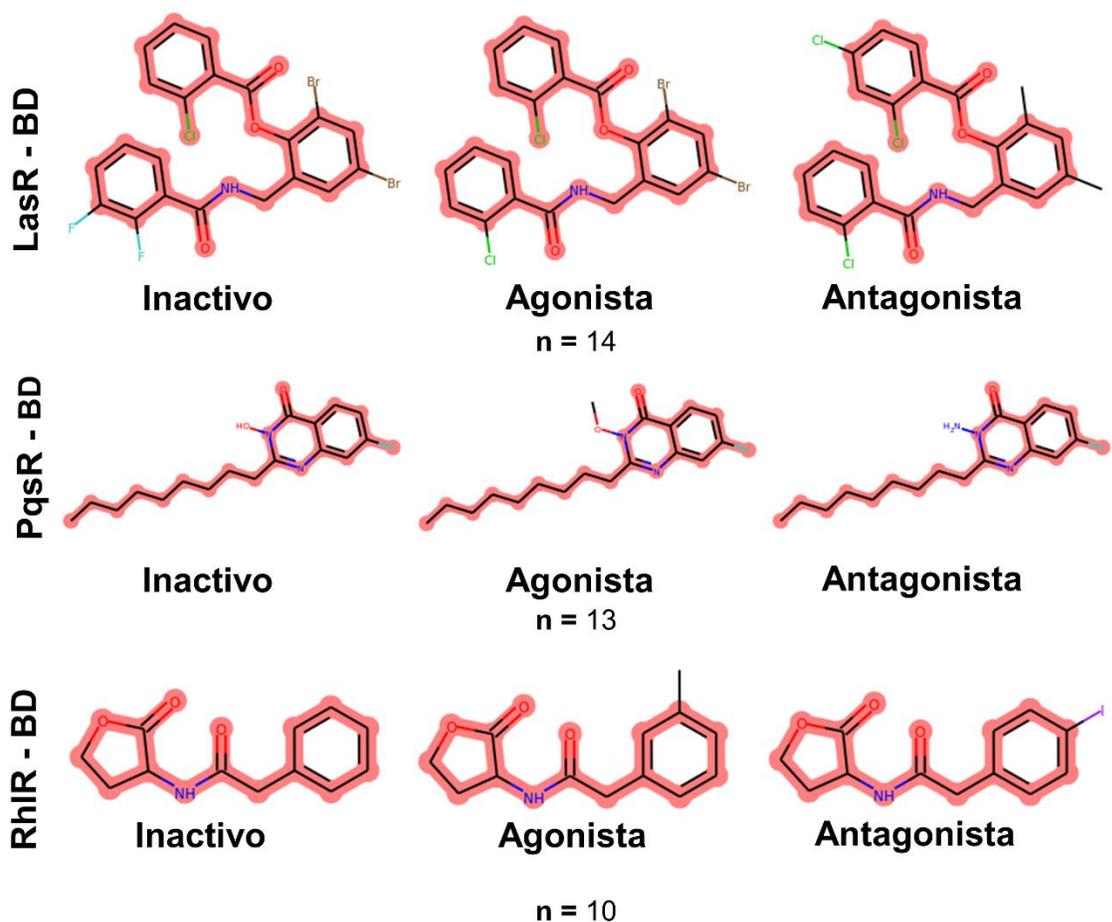


Figura 3.3: Ejemplos representativos del análisis de máxima común subestructura. N representa el número de moléculas por cada *scaffold*.

Todas las moléculas mostradas en la figura 3.3 que tienen actividad sobre LasR fueron reportadas por **O'Reilly y sus colaboradores**¹⁵³; éstas poseen un núcleo base trifenílico y cada actividad se relaciona con una sustitución específica en los anillos fenílicos, donde la molécula antagonista se caracterizó por estar sustituida por metilos en el anillo intermedio y cloros en los anillos de los extremos. La modificación de actividades debido a estas sustituciones puede deberse a un cambio en los efectos inductivos sobre estos anillos o un puente halógeno que pueda generarse en el sitio activo¹⁵⁴.

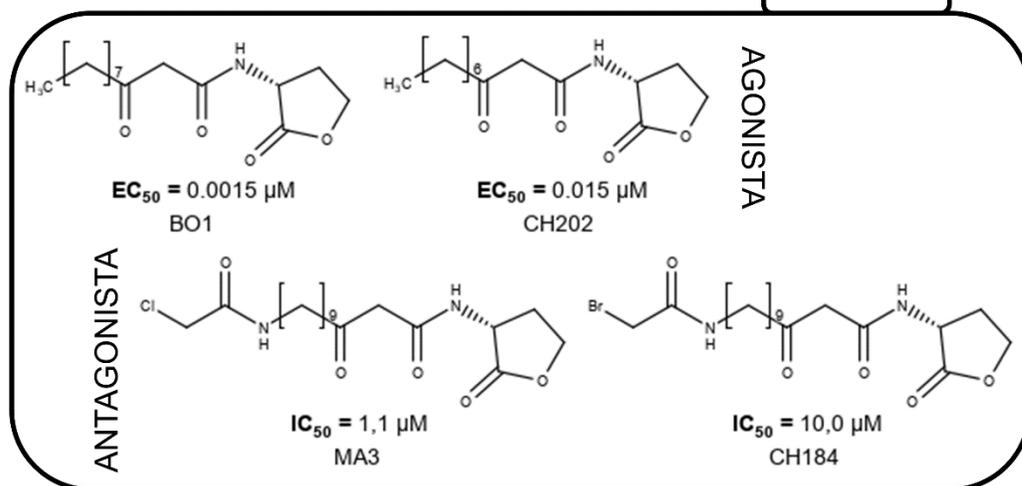
Ilangovan *et al.* reportaron los compuestos encontrados en el análisis MCS para PqsR, estos difieren unos de otros solo en la sustitución del nitrógeno en el anillo quinazolínico; el compuesto inactivo tiene un hidroxilo, el agonista un etilo y el antagonista una amida, mostrando que un posible puente de hidrogeno con el ángulo correcto, nos puede llevar a la actividad antagonista sobre este FT. Los compuestos del análisis de MCS relacionados con el FT RhIR evidencian que las interacciones como los puentes halógenos u algún efecto estérico pueden asociarse a la actividad antagonista sobre este FT, como se muestra en la figura 3.3.

Este análisis nos permite afirmar la presencia de *mechanism of actions cliffs* (MOA-cliffs), ya que al compartir los mismos *scaffolds* pero al generar diversas actividades según las sustituciones presentes se muestra una disrupción en las asociaciones estructura actividad¹⁵⁵.

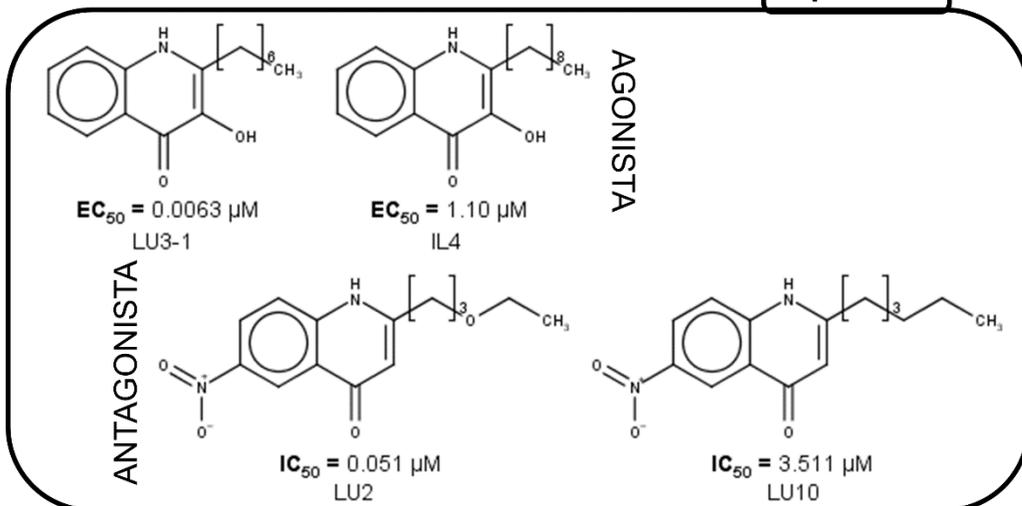
3.1.4 Análisis de *activity cliffs*

Los *activity cliffs* representativos en las **BD's para todos los FT's son mostrados** en la figura 3.4; siendo evidente la relación que poseen las moléculas agonistas con las antagonistas. Para el par mostrado con actividad agonista sobre LasR se evidencia que la variación de un solo carbono en la cadena lateral de las moléculas tipo AHL puede disminuir en diez ordenes de magnitud la potencia de estos compuestos; mientras BO1 tiene 9 carbonos, CH202 tiene 10 carbonos; evidenciando así la importancia del tamaño de esta cadena para la actividad en este tipo de moléculas. En el par encontrado para la actividad antagonista de LasR (MA3 y CH184) se observa el cambio en el halógeno que se encuentra al final de la cadena, mientras MA3 posee un Cloro, CH184 tiene un Bromo; pero la actividad de CH184 es solo el diez por ciento con respecto a MA3.

LasR - BD



PqsR - BD



RhIR - BD

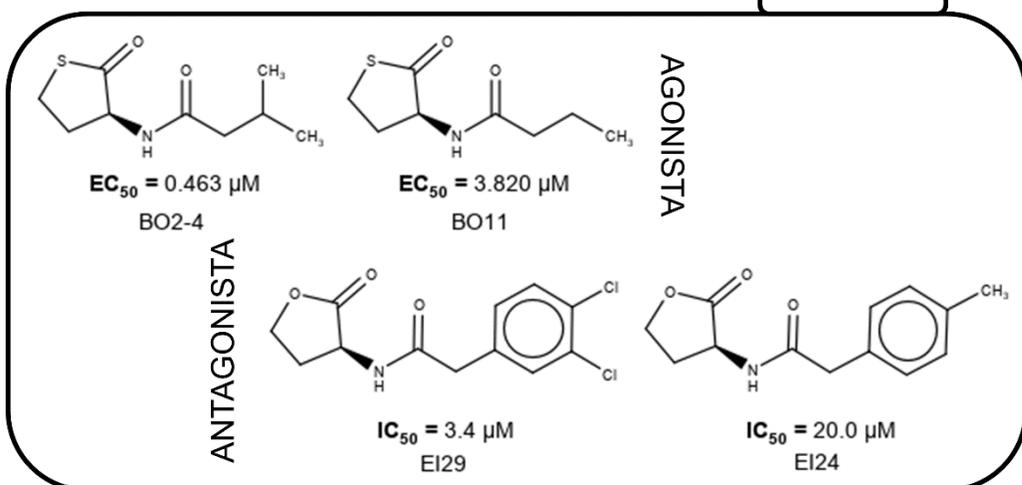


Figura 3.4: *activity cliffs* hallados en las BD's, aquí los pares moleculares asociados a la actividad agonista se muestran a la izquierda, mientras que los relacionados con la actividad antagonista se muestran a la derecha

Esto muestra la importancia de la longitud de la cadena, ya que los que generan algún tipo de interacción y tienen un menor tamaño permiten aumentar la actividad antagonista.

LU3-1 y IL4 son los pares moleculares hallados para la actividad agonista en PqsR, LU3-1 posee una cadena de siete carbonos mientras IL4 tiene la misma cadena con nueve carbonos; sin embargo, la actividad es alrededor de cien veces más grande en LU3-1, mostrando la relación entre el tamaño de las cadenas y la actividad agonista sobre PqsR. El par molecular encontrado para la actividad antagonista de PqsR posee el mismo tamaño de cadena, sin embargo, LU2 tiene un oxígeno en la cuarta posición mientras que en LU10 esta cadena solo se encuentran carbonos; este cambio genera una pérdida de actividad en LU10 de seis ordenes de magnitud, lo que puede ser explicado por la pérdida del donador de puente de hidrogeno, o la generación del ángulo adecuado para ocupar un sitio hidrofóbico.

Los *activity cliffs* relacionados con la actividad agonista sobre RhIR son compuestos con anillos tiolactónicos y una cadena lateral que es conectada a través de una amida, diferenciándose entre sí, en un grupo metilo en la cuarta posición para BO2-4, mientras BO11 no está sustituido en esta posición; este cambio genera que BO2-4 sea ocho veces más potente que BO11, mostrando la importancia de los sustituyentes en esta cadena ya que al poseer un átomo adicional posiblemente se puede ocupar un espacio hidrofóbico en el sitio de acción, tal como lo reportan Boursier y sus colaboradores ¹⁵⁶.

La sustitución del anillo fenílico en el par molecular encontrado para la actividad antagonista en RhIR, genera una pérdida de potencia de seis ordenes de magnitud aproximadamente; mientras EI29 es disustituido por dos cloros, EI24 es monosustituido con un grupo metilo; permitiendo asociar los efectos inductivos y estéricos con la actividad antagonista ¹⁵⁷.

3.1.5 Visualización del espacio químico

En la esquina superior derecha de la visualización del espacio químico mostrada en la figura 3.5, se encuentran ubicadas las moléculas con núcleos base trifenílicos, característicos de la BD LasR; aquí se observan los *activity cliffs* presentes, sin embargo, un grupo de moléculas antagonistas está separado del *cluster* principal, siendo estas un buen punto de partida para la optimización de la actividad antagonista y la búsqueda por similitud en grandes bases de datos. Los compuestos de la BD PqsR se muestran en el centro del gráfico, teniendo estas moléculas con núcleos quinoleínicos y quinazolínicos.

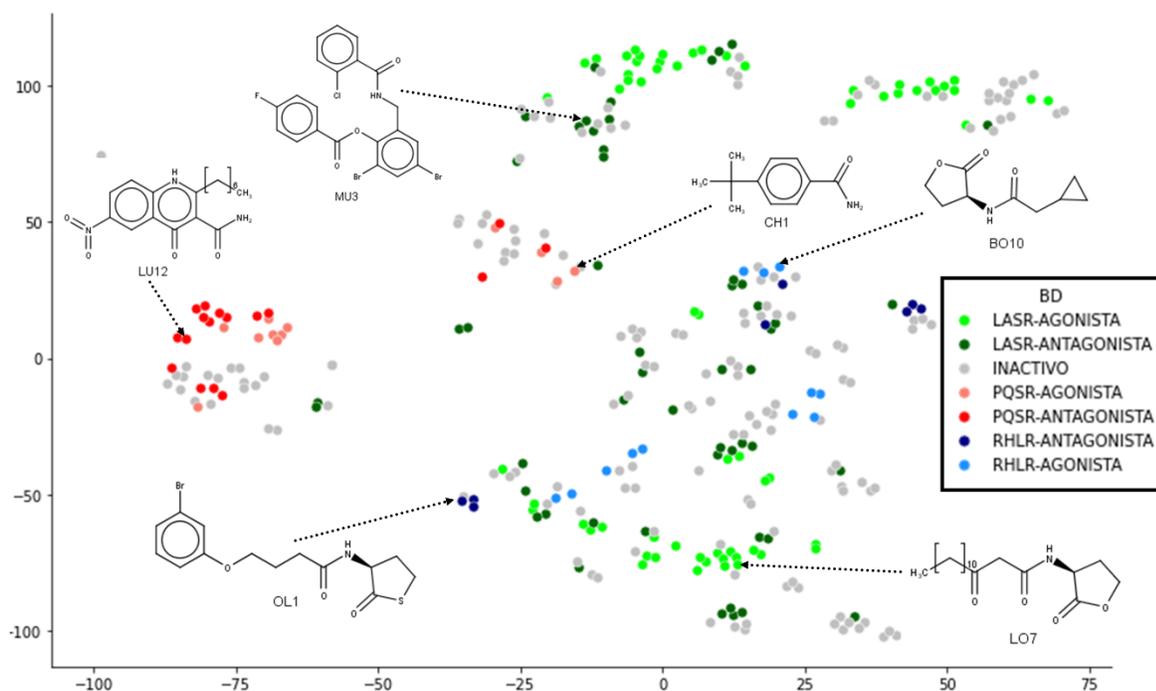


Figura 3.5: representación visual del espacio químico de las BD's obtenida por *t-distributed stochastic neighbor embedding* (t-SNE) en base al *fingerprint* ECFP4. Todos los compuestos son clasificados de acuerdo con la actividad, inactivos-gris, LasR agonistas (verde claro), LasR antagonistas (verde oscuro), PqsR agonistas (rosado), PqsR antagonistas (rojo), RhIR agonistas (azul claro), RhIR antagonistas (azul oscuro).

Todas las moléculas representadas en esta visualización están cercanas en el espacio, especialmente LasR y RhIR; lo que evidencia la relación estructural entre los compuestos que tienen actividad sobre estos FT's. Adicionalmente a lo anteriormente dicho, este espacio químico da de manifiesto la dificultad para separar las *scaffolds* o moléculas con actividad antagonista específica.

Los gráficos de constelación nos dan una **relación entre los núcleos en las BD's con actividad biológica** y el número de compuestos que incluyen los mismos. La figura 3.6 muestra el gráfico de constelación obtenido para la BD de LasR. Aquí son representados 19 núcleos agrupados en 10 series de análogos que contienen 106 compuestos (alrededor de la mitad de esta BD). Las primeras tres series de análogos muestran los cambios estructurales en los anillos difenílicos, la primera y tercera serie evidencian los difenilos conectados por una cadena de tres miembros con un grupo amida en medio de ellos.

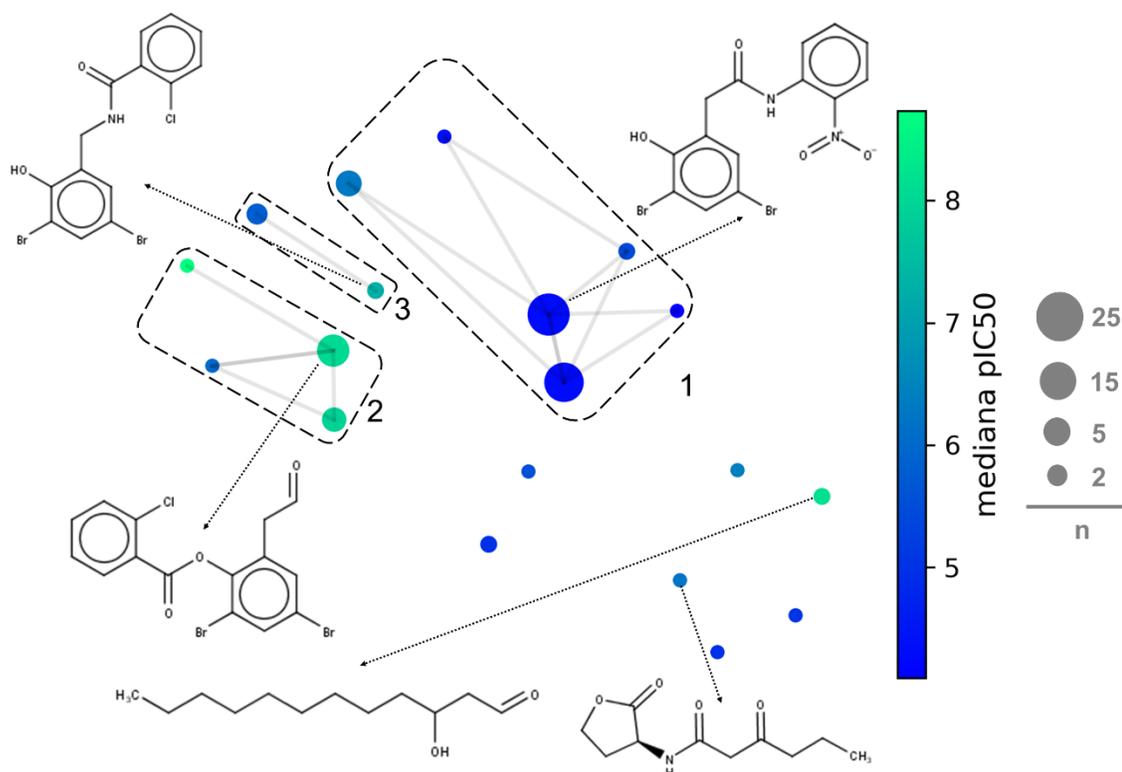


Figura 3.6: gráfico de constelación de la base de datos LasR. Aquí los núcleos más representativos de dicha BD son mostrados, teniendo este los mismos parámetros de la figura 3. 5 (t-SNE, ECFP4, métrica de Tanimoto), cada punto representa un núcleo, el tamaño de los círculos indica el “n” número de los compuestos relacionados, adicionalmente las conexiones entre los círculos indican que se comparten los núcleos, los rectángulos representan las series de análogos y el color muestra el promedio de pIC₅₀.

La segunda serie tiene los anillos difenólicos conectados a través de una cadena de dos miembros con un grupo éster en medio de ella, mostrando un perfil de actividad superior a las otras series, por lo tanto, estos núcleos son ideales para la búsqueda de nuevas moléculas con actividad sobre este blanco o la optimización de las ya existentes. Adicionalmente la cadena de 12 átomos de carbono tiene un excelente promedio de actividad sobre esta diana; en cambio, el núcleo lactónico no, evidenciando que el tamaño molecular y el número de enlaces rotables son características relevantes para este FT. Las **BD's de PqsR** y **RhlR** no pudieron ser analizadas por este método, debido a que las reglas de fragmentación no involucran la ruptura de anillos o cadenas carbonadas ¹⁰⁷.

En la tabla 3.2 se resumen las características estructurales de las moléculas **activas frente a los FT's** encontradas en el DFBL; según lo visto con los análisis previos no fue posible identificar con estos análisis características propias de las moléculas antagonistas exclusivamente; debido a la fuerte relación estructural con las agonistas e inactivas.

Tabla 3.2 Características estructurales asociadas con la actividad frente a los FT's. *Las interacciones hidrofóbicas y los apilamientos de anillos no se pudieron cuantificar.

	Núcleos base	Interacciones clave	Tamaño molecular
LasR	Anillos de 5 miembros (lactona)	2 aceptores de puentes de hidrogeno	Superior a 15 átomos
	3 anillos de 6 miembros interconectados (trifenilo)	1 donador de puente de hidrogeno	
	-Cadenas alifáticas (más de 10 carbonos)	Fuertes interacciones hidrofóbicas y apilamientos de anillos*	
PqsR	Biciclos nitrogenados (quinoleína y quinazolina)	1 aceptor de puentes de hidrogeno 1 donador de puente de hidrogeno	Superior a 15 átomos
	Cadenas alifáticas (más de 10 carbonos)	Fuertes interacciones hidrofóbicas y apilamientos de anillos*	
RhIR	Anillos de 5 miembros (Lactonas)	2 aceptores de puentes de hidrogeno	Entre 8 y 12 átomos
	Cadenas alifáticas (entre 3 y 6 carbonos)	1 donador de puente de hidrogeno	
	Anillos de 6 miembros (fenilos)	Interacción hidrofóbica y apilamiento de anillos	

Dentro de las características comunes, resaltan la presencia de interacciones hidrofóbicas que según los análisis anteriores podrían ser los responsables del cambio del tipo de actividad de un núcleo base.

3.2 Diseño de fármacos basado en la estructura (DFBE)

3.2.1 Alineamientos de las estructuras cristalográficas de los FT's

Al alinear estructuralmente los diferentes complejos de las estructuras cristalográficas almacenadas en el PDB, se pretendía encontrar posibles cambios estructurales relevantes para el tipo de actividad que puede tener una molécula. En la figura 3.7a se muestra el alineamiento de las estructuras de LasR,

donde los complejos con la molécula agonista (código PDB: 6D6B), la inactiva (código PDB: 6D6P) y el AI (código PDB: 2UVO) son representados con los colores azul, gris y beige respectivamente.

El principal cambio sucede en el bucle formado entre la Pro 41 hasta la Phe 51; aquí resalta la interacción con el residuo Tyr 47, el AI interactúa hidrofóbicamente mientras que el agonista interactúa mediante un apilamiento de anillos π - π , esta interacción no está presente en el complejo con la molécula inactiva, lo que induce al cambio de orientación de dicho bucle. Dicha interacción con la Tyr 47 fue descrita previamente como importante para la actividad por Vivas-Reyes et-al ¹⁵⁸.

En cuanto el alineamiento de los complejos de PqsR mostrado en la figura 3.7 b), la estructura del FT y la molécula agonista (código PDB: 4JVD) y la del AI (código PDB: 6Q7U) tienen el bucle formado por los residuos comprendidos entre la Leu 257 y Ilu 263 con una orientación diferente a la formada por el complejo con el compuesto antagonista (código PDB: 4JVI), siendo esta una de las partes claves para la actividad biológica.

En los complejos de las moléculas agonista y antagonista la interacción hidrofóbica con la Tyr 258 es la única que se presenta con los aminoácidos del bucle anteriormente mencionado; mostrando así la posible importancia de este residuo en la actividad biológica de PqsR.

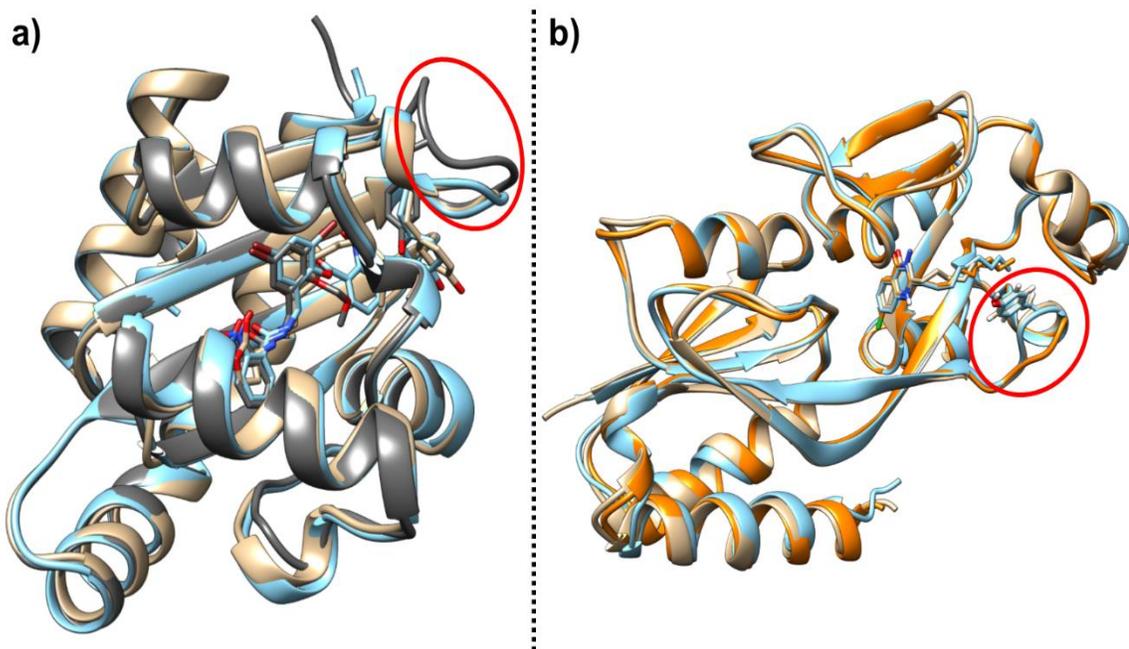


Figura 3.7: alineamiento de las estructuras cristalográficas almacenadas en el PDB; a) complejos del FT LasR, b) complejos del FT PqsR. Los colores de los complejos muestran el tipo de molécula con la cual

son acomplejados: azul – agonista, naranja – antagonista, gris – inactivo y beige – AI, el círculo rojo encierra la principal diferencia entre las estructuras de los FT's.

3.2.2 Construcción del modelo de RhIR

Para la construcción de este modelo se realizó el alineamiento de la secuencia proteica de RhIR obtenido de UniProt (código P54292), usando la herramienta BLAST^{135,159}, con la matriz de sustitución BLOSUM 62. Aquí se encuentra que las 2 primeras secuencias que tienen estructura cristalográfica conocida poseen un porcentaje de identidad mayor al 40 %, considerado adecuado para realizar el modelado por homología^{92,93,160}. De las anteriores, la identificada con el código PDB: 4LFU presento el mayor porcentaje de identidad y el menor e-value como se observa en la tabla 3.3, considerados críticos a la hora de evaluar los alineamientos de secuencias.

Tabla 3.3 resultados del alineamiento local realizado sobre el FT RhIR con la plataforma BLAST

Código PDB	Nombre y origen	Puntaje frente a RhIR			Resolución (Å)	Referencia
		Porcentaje de identidad	Score	e-value		
4LFU	SdIA – proteína regulatoria <i>E. coli</i>	40.76	190	4 x10 ⁻⁶⁰	2.26	Kim et al ¹⁶¹
4Y13	ftsQAZ – regulador transcripcional <i>E. coli</i>	40.34	182	8 x10 ⁻⁵⁷	3.10	Nguyen et al ¹⁶²
3SZT	OscR – represor quorum sensing <i>P. aeruginosa</i>	32.47	139	4 x10 ⁻⁴⁰	2.55	Lintz et al ¹⁶³
6MWH	LasR – FT quorum sensing <i>P. aeruginosa</i>	33.79	117	1 x10 ⁻³¹	2.20	Paczkowski et al ¹⁶⁴
6MWZ	LasR – FT quorum sensing <i>P. aeruginosa</i>	34.25	117	1 x10 ⁻³¹	1.66	Paczkowski et al ¹⁶⁴

Las cinco primeras proteínas halladas en el alineamiento local guardan relación con RhIR, encontrando las primeras tres ortólogas en referencia a RhIR, mientras que LasR resulta ser paróloga con respecto a RhIR; por lo tanto, al tener una relación biológica y al todas cumplir con funciones de regulación, cualquiera de ellas sería una buena plantilla para la construcción del modelo desde un punto de vista

biológico. Siendo escogido 4LFU como plantilla base para el modelado, por ser la secuencia que tiene la menor resolución con más del 40% de porcentaje de identidad.

Con las plataformas I-TASSER ¹³⁶, ROBBETA ¹³⁷, MODELLER ¹³⁸, PHYRE2 ¹³⁹, SWISS-MODEL ¹⁴⁰ se construyeron los posibles modelos a usar y se evaluaron usando las métricas mostradas en la tabla 3.4. El modelo de I-TASSER es descartado por tener 2 violaciones con respecto al gráfico de Ramachandran, reflejando ángulos incorrectos que desencadenan en problemas estéricos entre los aminoácidos ¹⁶⁵.

Entre los demás modelos se buscaba encontrar el balance entre los parámetros QMEANDisCo, RMSD, Lgscore y MaxSub; el primero evalúa las consistencias entre $C\alpha-C\alpha$ por pares del modelo con las restricciones extraídas de las estructuras homologas, además de tener en cuenta su potencial de solvatación y sus ángulos de rotación; para este parámetro todos los modelos tuvieron puntajes parecidos y superiores al 60 % ¹⁶⁶. Aunque se buscaba que no hubiera una diferencia significativa con el modelo de partida y para esto se evaluó el RMSD contra la estructura de referencia (4LFU), ninguno de los modelos supero los 1.5 Å, por lo cual este parámetro no fue relevante para la selección de la estructura adecuada.

Tabla 3.4 métricas de validación para los modelos construidos del FT RhIR

Modelo	Parámetro de Validación				
	RMSD	QMEANDisCo (± 0.05)	Lgscore	MaxSub	Ramachandran (%)
4FLU	0	0.89	4.589	0.236	99.57
I-TASSER	0.467	0.69	4.017	0.212	89.54
SWISS-MODEL	0.475	0.70	4.506	0.419	96.58
MODELLER	0.230	0.68	3.914	0.358	97.01
ROBETTA	0.960	0.69	5.047	0.270	99.58
PHYRE2	0.499	0.68	4.150	0.388	96.65

Lgscore busca los segmentos de la proteína no continuos correctamente conformados y MaxSub es calculado a partir de la mayor cantidad de residuos que se pueden encontrar en una distancia corta de 3.5 Å entre el modelo y una estructura comparada; siendo buenos modelos los valores mayores a 3 y 0.4 respectivamente ¹⁶⁷. Basado en las métricas anteriores el modelo con el mejor balance resulta ser el creado por SWISS-MODEL, siendo éste seleccionado para seguir con los estudios posteriores, su estructura se muestra en la figura 3.8.

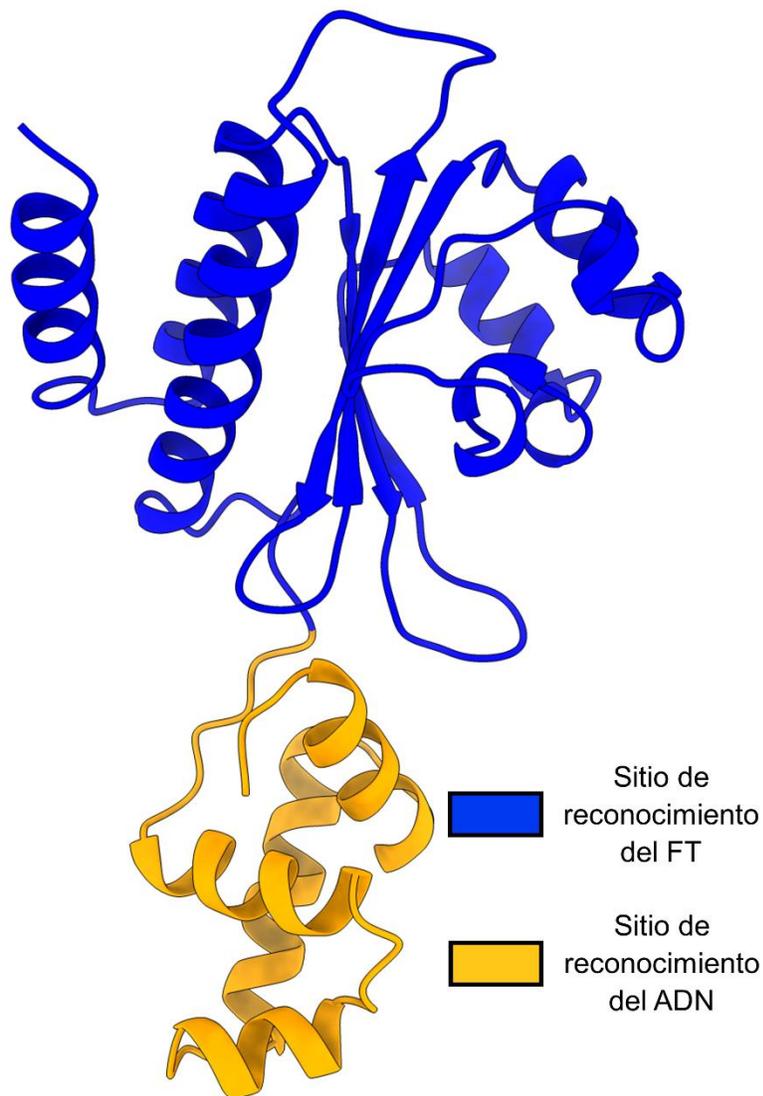


Figura 3.8 Modelo de RhIR obtenido de la plataforma SWISS-MODEL, en azul se encuentra el sitio de unión del FT y en naranja el sitio de unión al ADN.

Dicho modelo, a diferencia de las estructuras cristalográficas de LasR y PqsR, tiene el sitio de unión a ADN, siendo este un cambio importante a la hora de evaluar las posibles modificaciones que desencadenen diferentes actividades biológicas, como es observado en la figura 3.8.

3.2.3 Acoplamiento molecular automatizado

La preparación de las estructuras cristalográficas de LasR y PqsR se realizó con el software UCSF-Chimera como fue planteado en la metodología, obteniendo los blancos listos para realizar el re-

acoplamiento y encontrando así los parámetros a utilizar en los *dockings* posteriores como se observa en la tabla 3.5. El RMSD y las interacciones encontradas entre los AI's y los FT's, fueron los criterios guía para replicar las conformaciones más parecidas a lo experimentalmente observado. En todos los re-acoplamientos se lograron emular las interacciones de las estructuras cristalográficas, como se ve reflejado en los alineamientos de la figura 3.9.

Tabla 3.5 parámetros usados en los acoplamientos frente a los FT's; PH: puente de hidrogeno

Característica	LasR	PqsR	RhIR
Estructura			
(Código PDB)	3IX3	4JVD	N.A. / SWISS-MODEL
Resolución (Å)	0.64	2.95	N.A.
Grid	X = 10.68	X = -35.33	X = -23.38
	Y = 4.47	Y = 57.40	Y = 11.95
	Z = 20.95	Z = 9.17	Z = 15.79
	Tamaño = 15.0 Å	Tamaño = 15.0 Å	Tamaño = 15.0 Å
RMSD (Å)	0.72	1.47	N. A.
Enlaces de hidrogeno	PH = Trp 60		PH = Trp 68
	PH = Asp 73	No encontradas	PH = Asp 81
	PH = Ser 126		PH = Tyr 64
	PH = Tyr 56		

En el re-acoplamiento realizado para LasR se identifican interacciones importantes entre el AI y el sitio de unión, destacando los siguientes: puentes de hidrogeno entre el carbonilo del anillo lactónico con el Trp 60, nitrógeno de la amida con el Asp 73, carbonilo de la amida con la Tyr 56 y la Ser 129; estas son mostradas en la figura 3.9 d).

En PqsR no se logra identificar ninguna interacción relevante entre el AI y el sitio de unión, mientras que para RhIR los puentes de hidrogeno entre: carbonilo del anillo lactónico y el Trp 68, el nitrógeno de la amida con el Asp 81 y el carbonilo de la amida con la Tyr 64 parecen ser relevantes en la unión a este FT. Las interacciones relacionadas anteriormente podrían cumplir la función de anclaje de las moléculas con el sitio de unión respectivo.

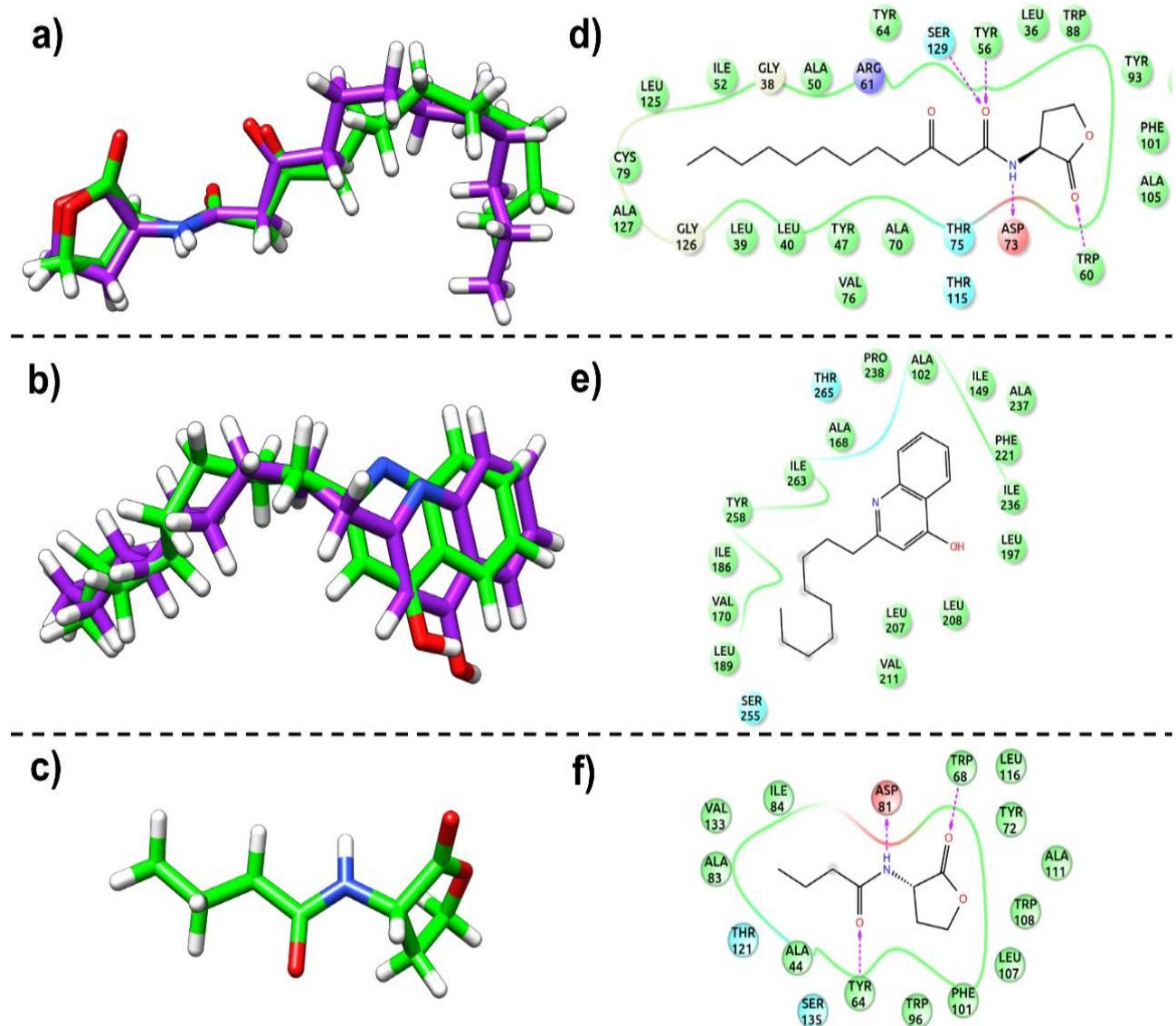


Figura 3.9 resultados de los re-acoplamiento de los AI's frente a los FT's, los numerales a), b) y c) representan los alineamientos de las conformaciones de las estructuras cristalográficas en morado y las conformaciones de los re-acoplamiento en verde, representando LasR-AI, PqsR-AI y RhIR-AI respectivamente; los numerales d), e), f) muestran las interacciones establecidas entre estas moléculas

y los residuos de los sitios de unión obtenidas de los re-acoplamiento, representando LasR-AI, PqsR-AI y RhIR-AI respectivamente.

Una vez se identificaron las interacciones relevantes **entre los AI's y los sitios** de unión, se procedió a **acoplar las moléculas de las BD's contra los FT's y a través de la metodología de *protein-ligand interaction fingerprint* (PLIF)** se busca encontrar cuales son las interacciones relevantes para determinado tipo de actividad según la ecuación 2.1 mostrada en la sección de metodología, que muestra las interacciones importantes para los agonistas con valores positivos y a las importantes para los antagonistas con valores negativos.

En la figura 3.10a se puede observar que las moléculas que poseen actividad sobre LasR presentan múltiples interacciones con los residuos que rodean el sitio activo, superando estas las 20 interacciones. Las interacciones con la Thr 115 y la Leu 40 se producen con mayor frecuencia en los compuestos con actividad agonista, siendo estas de tipo polar e hidrofóbica respectivamente; mientras que para la actividad antagonista resalta la interacción con la Ala 105, aunque esta no supero el 10 % de frecuencia por moléculas antagonistas.

Para la actividad agonista contra PqsR en este análisis se resaltaron las interacciones con la Leu 197 y la Ser 196, la primera de tipo donador de puente de hidrogeno y la segunda de tipo aceptor de puente de hidrogeno; sin embargo, ninguna de estas supero el 10 % de frecuencia. Los residuos Ala 102, Phe 221, Pro 238 y Ile 236 superaron el 40 % de frecuencia de interacción por compuestos antagonistas mostrando la gran relevancia de estos por este tipo de actividad, la mayoría de estas interacciones fueron hidrofóbicas solo la interacción de la Phe 221 fue de tipo aromática como se muestra en la figura 3.10b.

Las interacciones de las moléculas activas frente a los residuos Ile 84, Leu 107, Val 133, Leu 69 y Val 60 de RhIR superaron el 20 % de frecuencia mostrando la importancia de estos frente la actividad contra este FT; los tres primeros residuos fueron importantes para la actividad agonista, mientras que los dos últimos resultaron relevantes para la actividad antagonista. Cabe resaltar que todas las interacciones anteriormente mencionadas fueron de tipo hidrofóbico (Figura 3.10c).

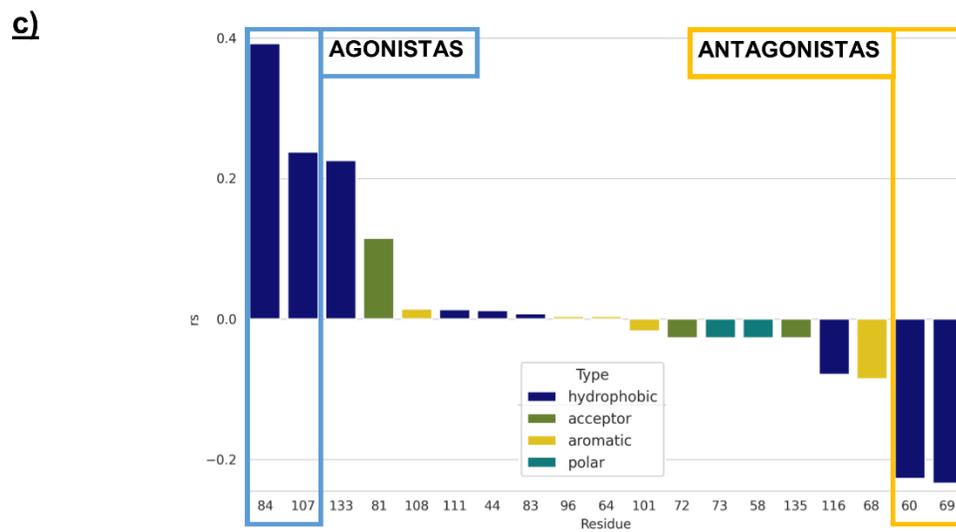
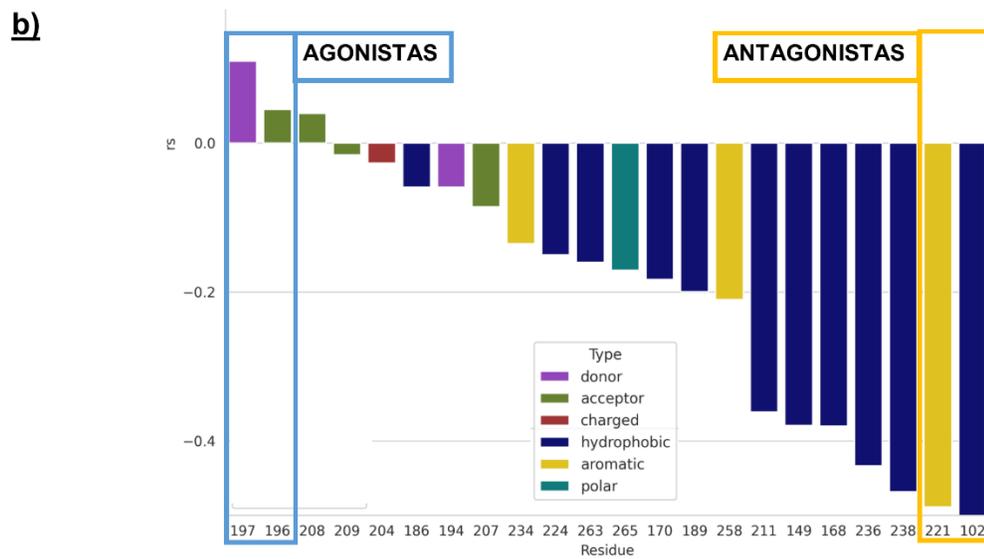
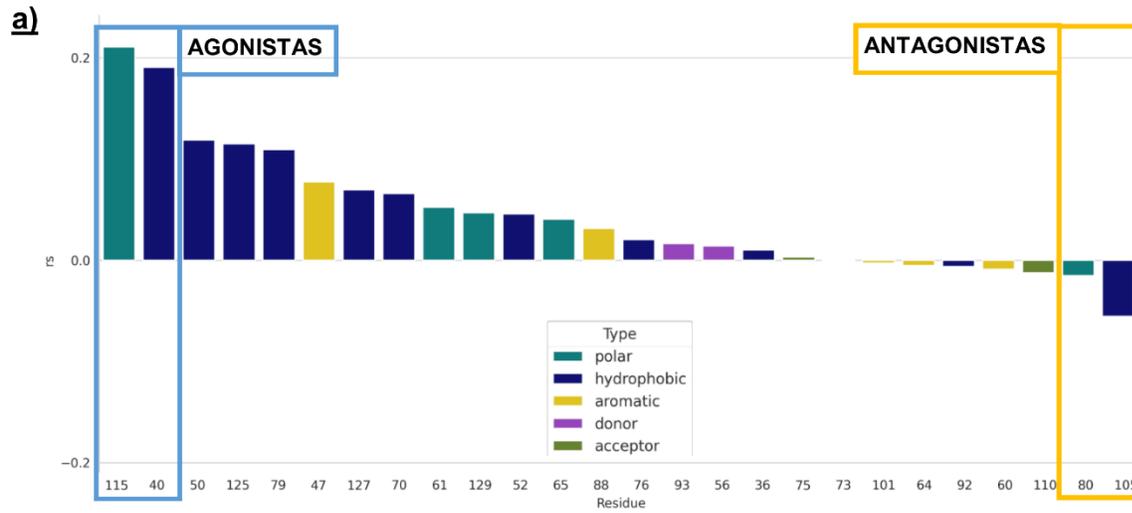


Figura 3.10 *fingerprints* de interacción ligando-proteína (PLIF's) obtenidos para los acoplamientos realizados entre las moléculas de las bases de datos y los FT's estudiados. En el eje "Y" se observa la prevalencia del tipo de interacción, siendo positivo para las interacciones preferentes para los agonistas y negativo para las interacciones antagonistas; mientras que el eje "X" muestra el residuo aminoacídico con el cual sucede la interacción. Los PLIF's asociados a LasR son mostrados por a), a PqsR por b) y a RhIR por c); el color de las barras indica el tipo de interacción relacionada con cada residuo.

3.2.4 Simulaciones de dinámica molecular

Con el fin de evaluar posibles diferencias de conformaciones presentadas entre los complejos de los FT's y las moléculas según su tipo de actividad, se realizaron simulaciones de dinámica molecular; éstas fueron llevadas a cabo con los compuestos hallados en el análisis de máxima común subestructura, que son aquellos que poseen una gran similitud entre sí, pero cambian su tipo de actividad con un pequeño cambio estructural (ver figura 3.3). Para demostrar que existen diferencias de movimiento entre los complejos formados por las moléculas según su tipo de actividad, se calcula la similitud de las trayectorias usando la métrica de Hausdorff¹⁵⁰ y luego se gráfica en un mapa de calor y un dendograma como se puede observar en la figura 3.11.

En los gráficos mostrados en la figura 3.11 se observa que las trayectorias de los complejos de compuestos agonistas y antagonistas difieren entre ellos, mostrando que un posible cambio de conformación generara un tipo de actividad diferente. Para LasR las trayectorias de los complejos tienen una diferencia de 1.48 Å en promedio; mostrando la rigidez de estos complejos (Figura 3.11a).

En PqsR se encuentra que las trayectorias de los complejos del FT con los compuestos agonista e inactivo son las que presentan la mayor diferencia entre sí, alcanzado los 2.42 Å en promedio. Mientras que la diferencia entre el complejo agonista y el antagonista supera los 2.5 Å, esto debido al cambio de un hidroxilo por un amino que sustituye el nitrógeno de las quinazolininas en las moléculas con las cuales se realizaron las dinámicas, como se puede observar en la figura 3.3b.

Las trayectorias de los complejos RhIR son las que presentan las mayores distancias entre sí, alcanzado un valor de 8.8 Å; en donde la divergencia entre las trayectorias de los complejos agonista y antagonista tiene un valor de 6.0 Å, mostrando un posible cambio conformacional que explicaría las dos actividades. Cabe mencionar que las menores similitudes entre las trayectorias se hallaron para las dinámicas de RhIR, debido a que el sitio de unión a ADN se encuentra conectado por un bucle que le permite una amplia movilidad, siendo esto previamente descrito por Chowdhury et al¹⁶⁸ y mostrándose en la figura 3.8.

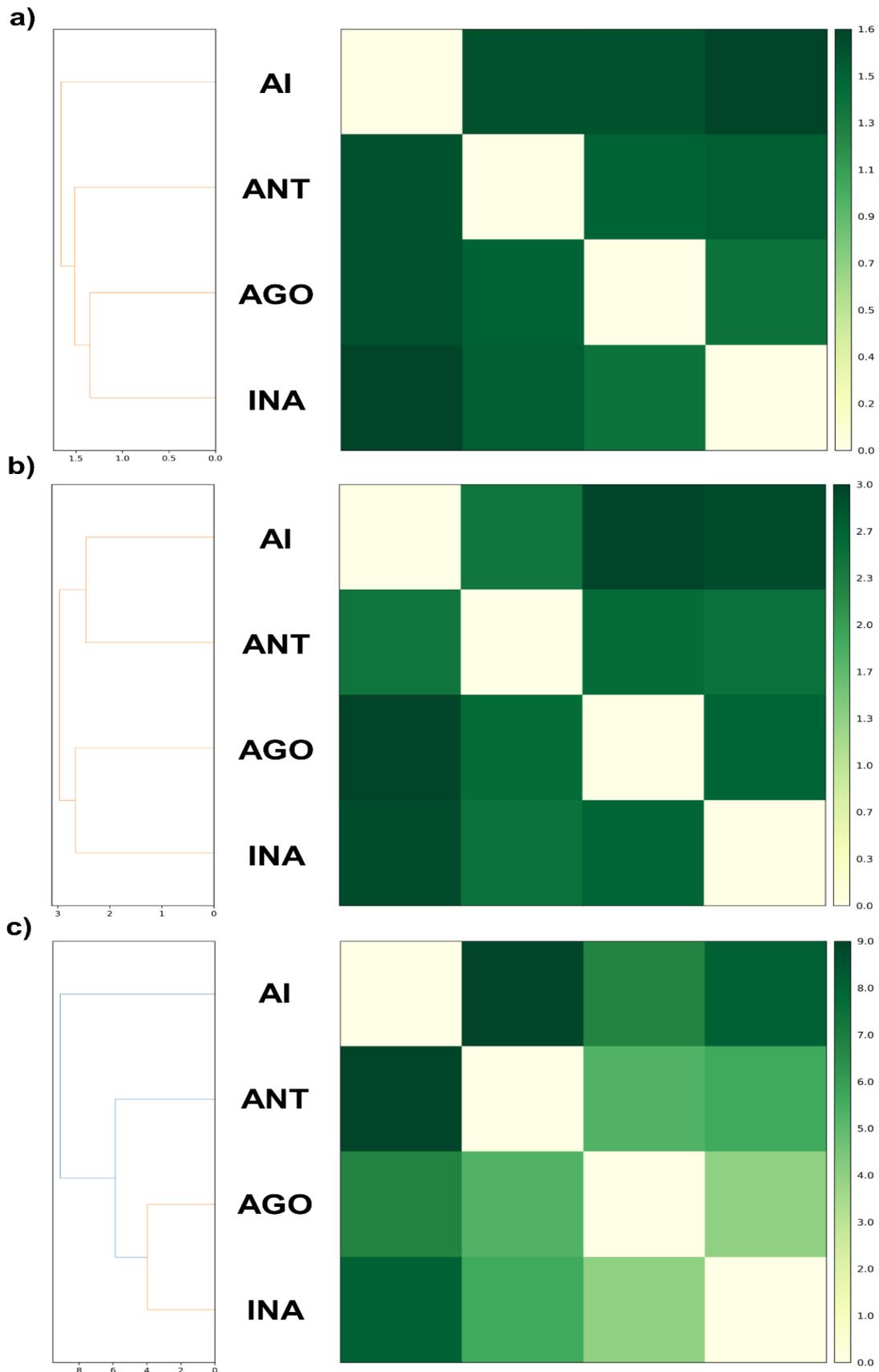


Figura 3.11 mapa de calor y dendograma de similitud entre las diferentes trayectorias de las dinámicas realizadas por FT a través de la raíz de la desviación cuadrática media (RMSD); donde se representan las dinámicas de a) LasR, b) PqsR y c) RhlR; al lado izquierdo se observa el dendograma que agrupa según su similitud a las 4 trayectorias, mientras al lado derecho se representa a través de una escala de color la diferencia entre ellas. Significando AI: auto-inductor, ANT: antagonista, AGO: agonista y INA: inactiva.

Para encontrar los mayores cambios entre las conformaciones de los complejos agonista y antagonista se decide comparar los RMSD de estas dinámicas, dichas comparaciones son mostradas en las figuras 3.12, 3.13 y 3.14. En LasR se encontró que la mayor divergencia estaba en el cuadro (*frame*) 8144 de los complejos formados con los compuestos agonista y antagonista, luego al alinear las estructuras de este cuadro (ver figura 3.12) se encontraron las interacciones que pueden explicar estas diferencias. En el caso del complejo LasR-agonista se presentan 2 puentes de hidrogeno con la Thr 75 y la Ala 127, mientras que en el complejo formado con el antagonista las interacciones hidrofóbicas con la Tyr 56, la Leu 36 y la Ile 52 generando el cambio en la hélice que se encuentra en el sitio de unión del FT, que a su vez lleva a una conformación diferente en el sitio de unión del ADN y su posterior actividad antagonista.

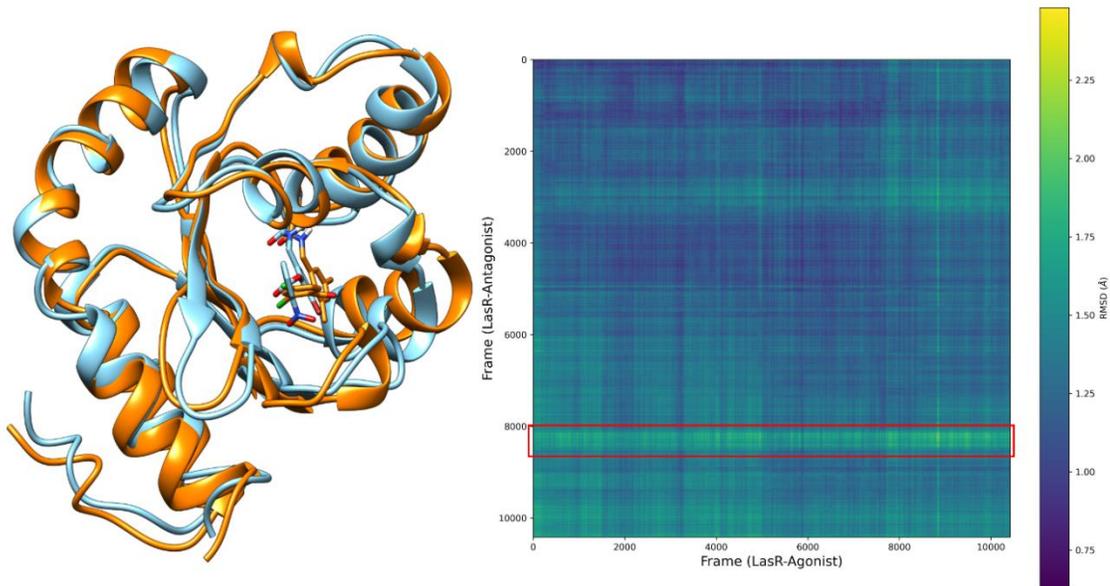


Figura 3.12 cambios de conformación entre las dinámicas de los complejos LasR-antagonista (naranja) y LasR-agonista (azul). A la derecha se observa el mapa de calor que compara la diferencia entre los RMSD de las dinámicas y a la izquierda se muestra el alineamiento de los complejos cuando presentan la mayor variación en estos RMSD. El recuadro rojo indica los cuadros (*frames*) con mayores diferencias de RMSD entre los dos complejos.

Al comparar los RMSD de las dinámicas de los complejos formados entre PqsR y las moléculas agonista y antagonista en el mapa de calor mostrado en la figura 3.13 se observa que las diferencias de estas superan los 4.0 Å durante toda la dinámica, siendo la conformación mostrada en el cuadro 1623 la que presenta la mayor divergencia. Aquí el complejo de PqsR y el compuesto agonista forman un puente de hidrogeno con la Gln 194, además de una interacción hidrofóbica con la Tyr 258; mientras que el complejo formado con la molécula antagonista genera un puente halógeno con la Ala 102, un puente de hidrogeno con la Thr 256, explicando las diferentes conformaciones adoptadas por estos complejos.

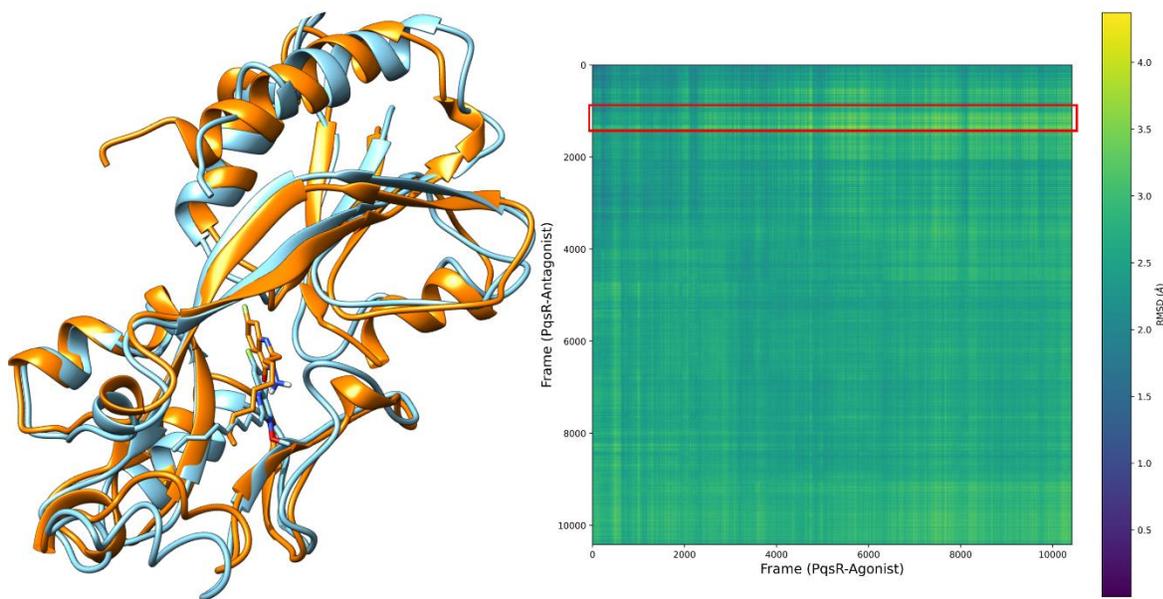


Figura 3.13 cambios de conformación entre las dinámicas de los complejos PqsR-antagonista (naranja) y PqsR-agonista (azul). A la derecha se observa el mapa de calor que compara la diferencia entre los RMSD de las dinámicas y a la izquierda se muestra el alineamiento de los complejos cuando presentan la mayor variación en estos RMSD. El recuadro rojo indica los cuadros (frames) con mayores diferencias de RMSD entre los dos complejos.

En el complejo RhlR-agonista se encontró solo la interacción hidrofóbica con la Val 133, mientras que para el complejo antagonista hallaron múltiples interacciones, entre las que destacan el apilamiento de anillos con la Tyr 72, los puentes de hidrogeno con la Tyr 64, Tyr 72 y la Ser 132, además de la interacción hidrofóbica con el Trp 96, siendo estos los posibles causantes de la modificación de estas conformaciones y la explicación de las diferencias con la actividad biológica. Adicionalmente, el alineamiento permitió ver un cambio de conformación en el dominio de unión a ADN entre los complejos de agonista y antagonista, siendo este una posible explicación de la pérdida de la actividad agonista.

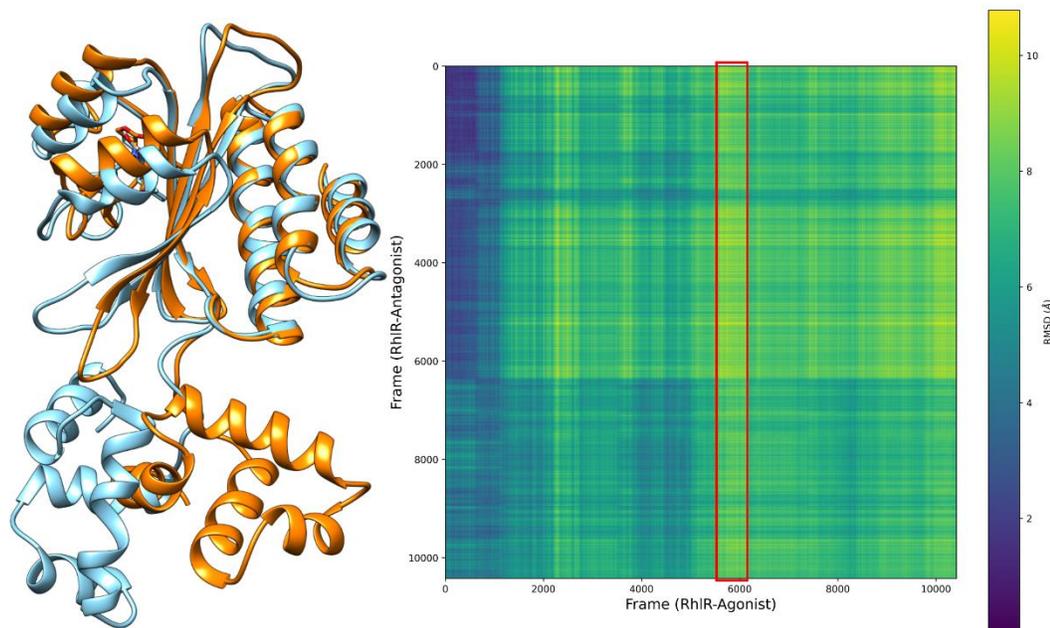


Figura 3. 14 cambios de conformación entre las dinámicas de los complejos RhIR-antagonista (naranja) y RhIR-agonista (azul). A la derecha se observa el mapa de calor que compara la diferencia entre los RMSD de las dinámicas y a la izquierda se muestra el alineamiento de los complejos cuando presentan la mayor variación en estos RMSD. El recuadro rojo indica los cuadros (frames) con mayores diferencias de RMSD entre los dos complejos.

Para el análisis del movimiento de cada aminoácido se utiliza la *root mean square fluctiation* (RMSF) de las posiciones atómicas, que mide la tasa de cambio de la posición de un residuo en Å. Permitiendo encontrar cuales aminoácidos cambian por la acción de las fuerzas presentes en el sistema 169,170. En la figura 3.15 se observan las gráficas de las fluctuaciones medias para cada residuo en las diferentes dinámicas realizadas, los rectángulos rojos en cada grafica señalan las regiones de mayor movimiento **en las estructuras de los FT's con los compuestos agonistas, lo que implicaría que interacciones** con los antagonistas permiten una mayor rigidez de la estructura proteica impidiendo la conformación adecuada para la unión con el ADN. En el caso de LasR se señalan los residuos: Lys 42, Asp 43 y Gln 45, que pertenecen al bucle señalado en el alineamiento de las estructuras cristalográficas como posibles responsables de la actividad antagonista, además la Gln 94 cercana al sitio de unión también presentó un mayor movimiento en la dinámica con el agonista (Figura 3.15a).

En la figura 3.15b se observa el RMSF de las dinámicas de los complejos PqsR; los aminoácidos Leu 115, Leu 116, Asp 118 y Met 121 que se muestran en el primer recuadro rojo, presentan un mayor movimiento en la dinámica realizada con la molécula agonista pudiendo ser esta la causa del cambio de conformación y la posible activación del FT; también la Gly 202 y la His 204 presentan un valor alto de RMSF.

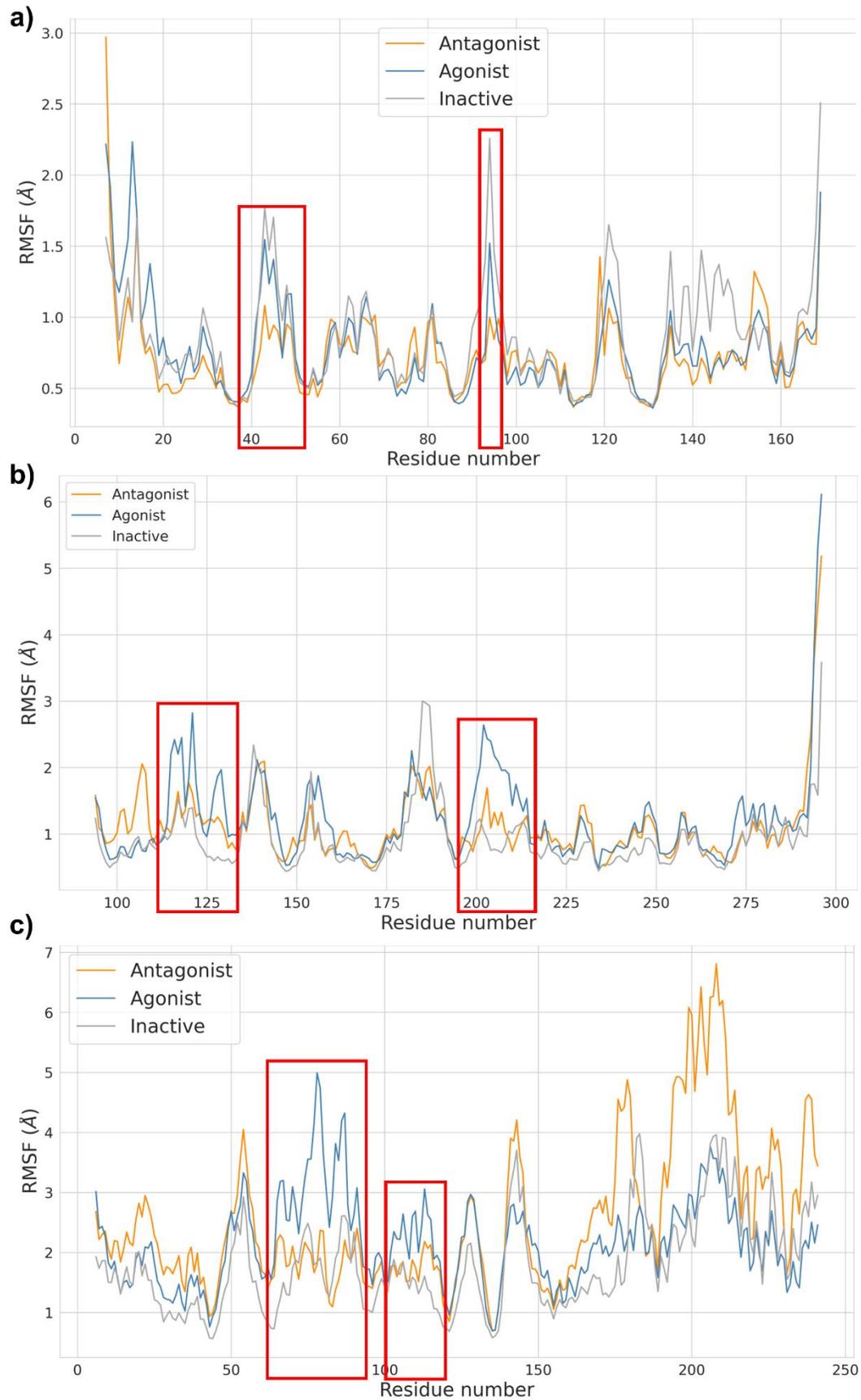


Figura 3.15 RMSF por residuo en las dinámicas realizadas; se representan las dinámicas: a) LasR, b) PqsR y c) RhIR. El color de cada grafica representa el tipo de actividad que tiene la molécula correspondiente, en donde los antagonistas son naranja, los agonistas son azul y los inactivos gris. Adicionalmente los recuadros rojos muestran los cambios más significativos.

En cuanto a las fluctuaciones presentadas por los aminoácidos en las dinámicas realizadas con el FT RhIR (Figura 3.15c), los residuos 77 al 87 y 106 al 113 presentan un movimiento mayor en el complejo formado con el compuesto agonista; en cambio de los residuos 178 al 239 asociados a la región de unión a ADN tuvieron un mayor movimiento en el complejo con la molécula antagonista, lo que evidencia que si se tiene una mayor rigidez en el sitio de unión al AI, se podría producir un mayor movimiento en la región de unión al ADN y el FT no podrá unirse de manera adecuada, impidiendo así su función biológica.

Contrario a lo esperado, la fluctuación presentada por los átomos del compuesto antagonista en la dinámica con el FT LasR fue mucho mayor que los de los compuestos agonistas o inactivos, mostrando que este complejo podría llegar a ser más inestable en el tiempo. En la gráfica mostrada en la figura 3.16a se observa que las tres moléculas tienen la menor fluctuación en el anillo que está conectado por la amida representado por los átomos 16 al 23, cabe también mencionar que el átomo 24, un cloro y los átomos 1 al 6 y un anillo bencénico unido por el éster presentaron los mayores valores de RMSF en la dinámica del complejo con el antagonista, mostrando que este podría ser clave para la actividad agonista.

Los valores de RMSF de los átomos asociados a las moléculas acomplejadas con el FT PqsR fueron relativamente similares, sin embargo, los carbonos de la cadena lateral 14, 15, 16, 17 y 18 permanecieron mucho más estables en la dinámica con el compuesto antagonista, que con el agonista mostrando que esta parte de la molécula al establecer alguna interacción podría inducir a la actividad antagonista; como se observa en la figura 3.16 b.

En cuanto los átomos 12 y 13, un flúor y R1, de los compuestos utilizados en la dinámica con PqsR, se observó que en las moléculas activas (agonista y antagonista), el flúor posee una menor fluctuación de movimiento, generando que el átomo 13 (N en el antagonista y O en el agonista) tenga un mayor cambio en el espacio lo que posiblemente ocasione la pérdida de la interacción polar (Figura 3.16b).

Con respecto a los cambios presentados por los átomos de las moléculas en las dinámicas de RhIR, se observa que, la que presentó una menor variación fue la molécula inactiva; no superando un valor de 1.5 Å en el RMSF por ninguno de sus átomos, mientras que para los compuestos activos se hallaron átomos con RMSF cercanos o superiores a los 2.0 Å (ver figura 3.16c). Sin embargo, esta gráfica mostró

que en términos generales el compuesto antagonista es mucho más estable que el agonista, generando estas mayores interacciones y a su vez mas rigidez en la estructura.

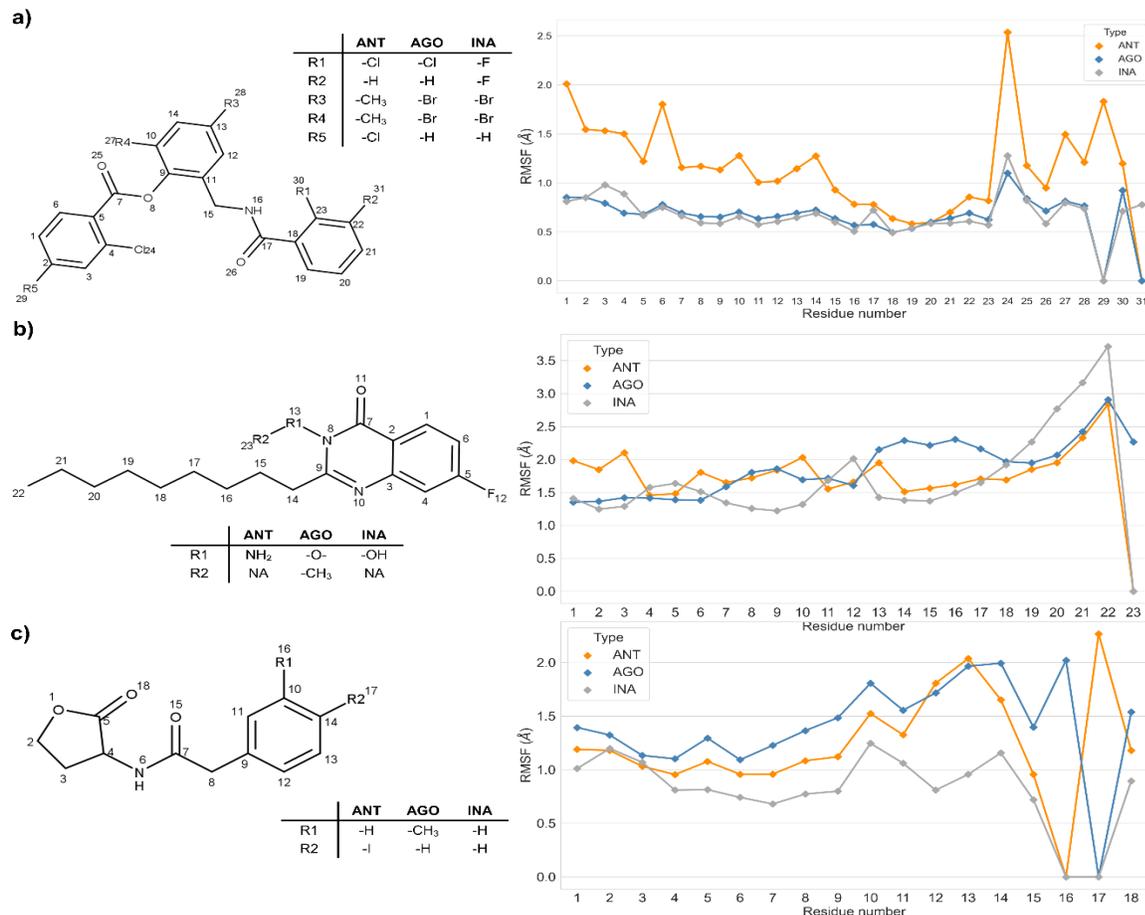


Figura 3.16 RMSF mostrado para cada uno de los átomos de las moléculas con las cuales se realizaron las dinámicas moleculares. Los compuestos asociados y la gráfica de variación de RMSF con respecto al FT LasR se muestra en a); con respecto a PqsR en b) y a RhIR en c). El color de cada gráfica hace referencia al tipo de actividad que posee la molécula, naranja los antagonistas, azul los agonistas y gris los inactivos.

Por último, con el fin de determinar que interacciones fueron relevantes durante la dinámica, se revisó la fracción de interacciones presentadas entre la molécula con los aminoácidos que circundan el sitio de unión. En la figura 3.17 se observan las interacciones realizadas por los compuestos antagonista y agonista con el FT LasR, aquí los residuos Leu 36, Tyr 47, Ile 52 y Tyr 56 tienen interacciones más prolongadas con el compuesto antagonista que agonista, todos interaccionaron hidrofóbicamente y adicionalmente las Tyr 47 y 56 generaron un puente de agua y uno de hidrogeno respectivamente. Las

interacciones con la Tyr 64, el Asp 73, el Trp 88, la Phe 101 y la Ser 129 fueron comunes para las dos moléculas y adicionalmente tuvieron una duración similar.

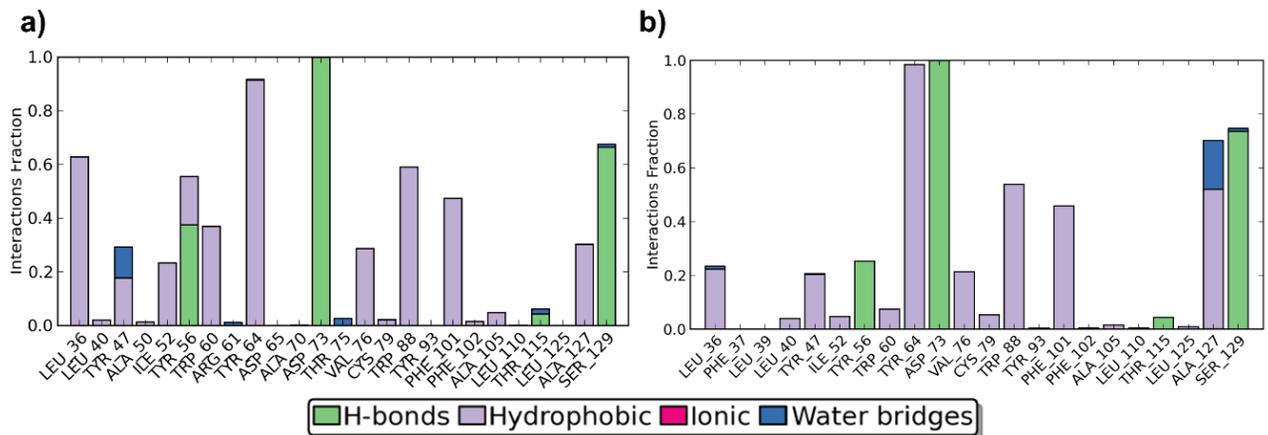


Figura 3.17 interacciones presentadas por los complejos de compuestos agonista y antagonista con el FT LasR. En a) se representan las interacciones con el antagonista y en b) se representan las interacciones con el agonista.

En cuanto las interacciones con mayor duración en el compuesto agonista solo se encontraron la realizada con la Ala 127, esta se mantuvo aproximadamente el doble del tiempo y se caracterizó por tener una parte hidrofóbica y otra de puente de hidrogeno, lo que le da mayor estabilidad. Por último, se resalta la interacción de la Tyr 47 con el compuesto antagonista, ya que esta también estuvo presente en el alineamiento de estructuras cristalográficas y fue descrita previamente como importante por Vivas-Reyes et al ¹⁵⁸.

Las interacciones importantes con mayor porcentaje de duración para la molécula antagonista contra PqsR fueron las realizadas con los residuos Ile 149, Leu 207 y Pro 238, la primera y la tercera fueron de tipo hidrofóbica, mientras que la realizada con la Leu 207 fue predominantemente un puente de hidrogeno, que a su vez en algunos momentos generaba puentes de agua e interacción hidrofóbica como se observa en la figura 3.18. La interacción con la Ile 236 resultó tener un comportamiento similar entre las dinámicas realizadas de las moléculas agonista y antagonista, aunque la interacción con el antagonista fue predominantemente por un puente de agua, al contrario del agonista que resulto ser hidrofóbica.

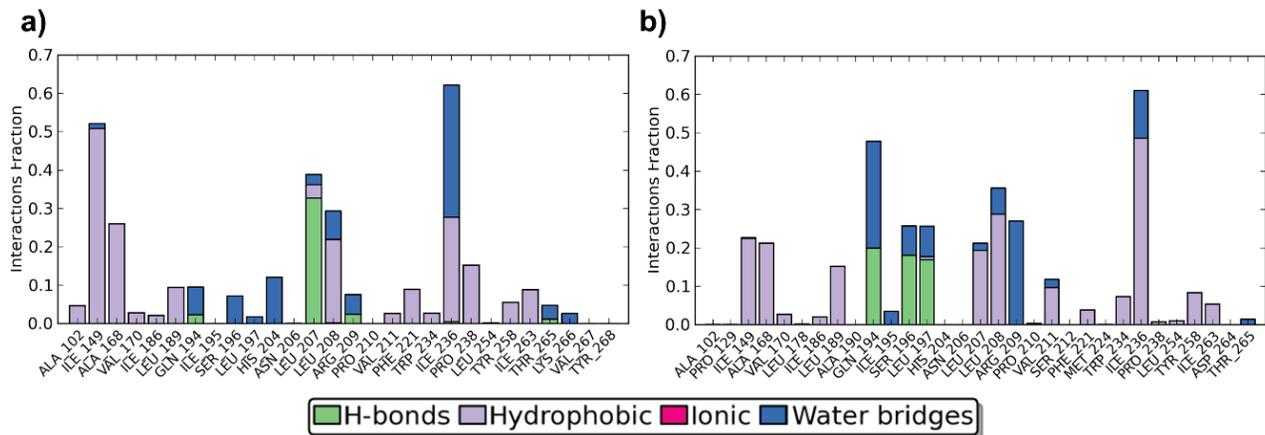


Figura 3.18 interacciones presentadas por los complejos de compuestos agonista y antagonista con el FT PqsR. En a) se representan las interacciones con el antagonista y en b) se representan las interacciones con el agonista.

La dinámica con el compuesto agonista de PqsR, mostró tener un mayor tiempo de interacción con los residuos Gln 194, Leu 197, Ser 196 y Arg 209; siendo todos de tipo polar, donde los primeros tres establecen puentes de hidrogeno entre la molécula y los residuos; mientras que todos forman puentes de agua con estos aminoácidos. Cabe resaltar que las interacciones hechas con la Leu 197 y la Ser 196 también fueron descritas en el análisis de *protein ligand interaction fingerprints*, como importantes para esta actividad.

En la gráfica 3.19 se muestran las interacciones realizadas entre los compuestos antagonista y agonista y el FT RhIR durante la duración de las dinámicas. Aunque no se observan interacciones únicas con la molécula antagonista, si se resaltan las realizadas con los residuos Tyr 64, Trp 68 y Tyr 72, ya que tuvieron un mayor tiempo de contacto. La dinámica con el compuesto agonista tuvo más contactos, entre los que se resaltan con Leu 69, Phe 101, Ala 111, Thr 121, Val 133 y Ser 135; la mayoría de éstos establecen interacciones de forma hidrofóbica, solo los dos últimos residuos (Val 133 y Ser 135) generan puentes de hidrógenos con esta molécula.

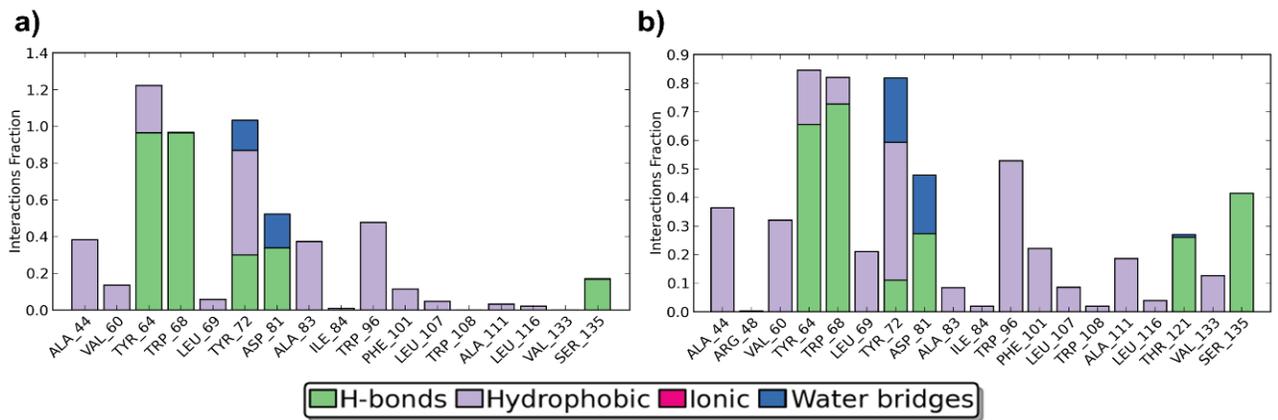


Figura 3.19 interacciones presentadas por los complejos de compuestos agonista y antagonista con el FT RhIR. En a) se representan las interacciones con el antagonista y en b) se representan las interacciones con el agonista.

Este estudio de diseño de fármacos basado en la estructura permitió hallar, como claves, las interacciones realizadas con los residuos que aparecen en la tabla 3.6, la mayoría de estos interactúan con las moléculas de forma hidrofóbica y polar. Aquí se resaltan los resultados hallados para el FT PqsR, este fue el que presentó mayor relación entre los estudios conducidos. Estas interacciones establecidas como importantes fueron las determinantes en el cribado virtual realizado posteriormente

Tabla 3.6 aminoácidos relevantes por cada tipo de actividad para todos los FT's estudiados por PLIP y dinámica molecular

	LasR		PqsR		RhIR	
Tipo de actividad	Aminoácido	Estudio	Aminoácido	Estudio	Aminoácido	Estudio
Agonista	Thr 115	PLIF	Ser 196	PLIF – Dinámica	Ile 84	PLIF
	Leu 40	PLIF	Leu 197	PLIF – Dinámica	Leu 107	PLIF
	Thr 75	Dinámica	Gln 194	Dinámica	Leu 69	Dinámica
	Ala 127	Dinámica	Arg 209	Dinámica	Phe 101	Dinámica
					Thr 121	Dinámica
Antagonista	Tyr 47	Alineamiento – Dinámica	Tyr 258	Alineamiento	Val 133	Dinámica
	Ala 105	PLIF	Phe 221	PLIF	Ser 135	Dinámica
				Val 60	PLIF	
				Leu 69	PLIF	

Thr 80	PLIF	Ala 102	PLIF – Dinámica	Tyr 64	Dinámica
Leu 36	Dinámica	Leu 207	Dinámica	Tyr 72	Dinámica
Ile 52	Dinámica	Ile 149	Dinámica	Ala 83	Dinámica
Tyr 56	Dinámica				

3.3 Modelado farmacofórico

En la figura 3.20 se observan los modelos farmacofóricos encontrados para los diez antagonistas más activos frente a cada uno de **los FT's**, estos fueron realizados usando alineamientos rígidos con las conformaciones obtenidas a partir del acoplamiento molecular. Para el FT LasR se encontró que 4 de las 5 características son de tipo hidrofóbico y aromático, algo que es soportado en los análisis previos; sin embargo, la parte final de las cadenas laterales de estas moléculas no se acercaron lo suficiente para representar el punto hidrofóbico que se esperaría que estuviera a una distancia mayor de los 12 Å, siendo esta una característica que se deber tener en cuenta en la optimización este modelo.

En la figura 3.20b se observa el modelo hallado para el FT PqsR, este al igual que en LasR representa en su mayoría a las características hidrofóbicas y aromáticas encontradas en los anillos quinoleínicos y quinazolínicos, sin embargo, las cadenas laterales no se superpusieron de manera correcta, impidiendo la generación de la característica correspondiente.

El modelo farmacofórico hallado para la actividad antagonista frente a RhIR tiene una cobertura total frente a las diez moléculas halladas, esto debido a la gran similitud estructural. Aquí, al igual que en modelos hallados con anterioridad; los puntos hidrofóbicos y aromáticos resultaron tener la mayor relevancia, sin descartar las características de aceptores de puente de hidrogeno que se muestran en la figura 3.20c.

En general, estos modelos farmacofóricos nos permitieron encontrar las similitudes estéricas y electrónicas que tienen los **antagonistas de los FT's entre sí**. Siendo **los de mayor tamaño LasR y PqsR, debido a que el sitio activo de estos FT's es más grande que en RhIR**, lo que genera que dichas características estén a una menor distancia en este FT. Adicionalmente, todas poseen un aceptor como punto de anclaje al sitio de unión.

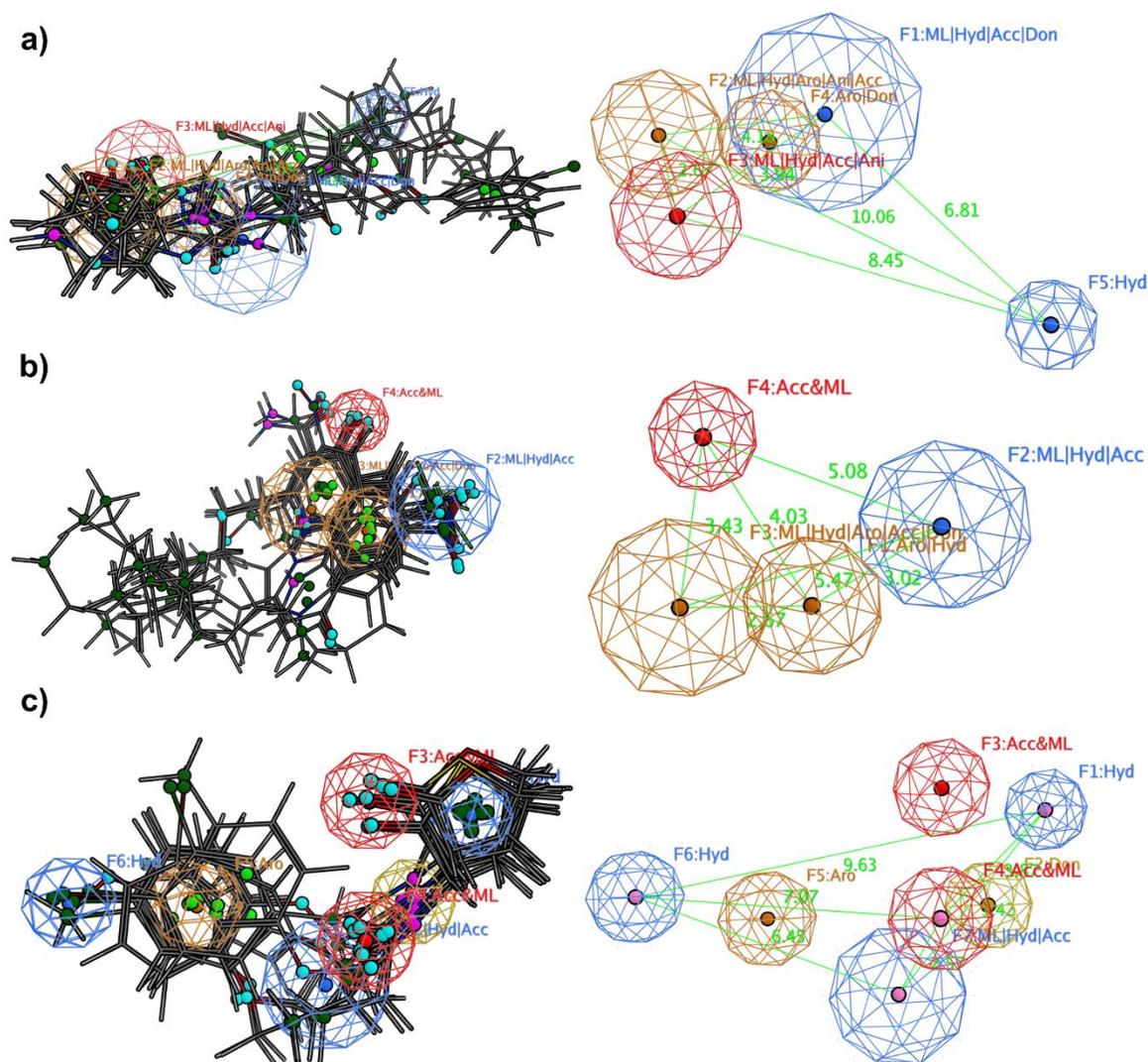


Figura 3.20 modelos farmacofóricos desarrollados para la actividad antagonista en los FT's. A la izquierda se representa el alineamiento de las diez moléculas más activas y a la derecha los puntos farmacofóricos encontrados con sus respectivas distancias, estando representado el modelo farmacofórico de LasR en a), de PqsR en b) y de RhlR c).

3.4 Clasificadores de machine learning

Al realizar las primeras pruebas de los diferentes algoritmos para la clasificación de las moléculas entre los tres tipos de actividad relacionados en esta tesis (agonistas, antagonistas e inactivos), estos clasificadores no lograron etiquetar a más del 50% de las moléculas en el grupo agonista y en el antagonista, probablemente por la alta similitud estructural presente en estos grupos, por esta razón se decidió agrupar a los agonistas y antagonistas dentro de una sola categoría, denominada activos.

Con las categorías de activos e inactivos se construyeron los modelos para cada FT comparándolos contra las métricas propuestas en la sección de metodología. En la tabla 3.7 se puede observar el comportamiento de cada modelo de clasificación para el FT LasR frente a cada métrica. Según lo reportado por autores como Kim et. Al ¹⁶⁹, Mayr et. Al ¹⁰⁶ y Yang et. Al ¹⁷⁰. Al usar *molecular fingerprints*, la mejor métrica para encontrar el mejor clasificador sería el AUC, que representaría que tan bien están clasificando los modelos a cada punto dentro de la categoría que les corresponde. Sin embargo, debido al desbalance entre activos e inactivos en todas las bases de datos (LasR = inactivos: 311, activos: 116; PqsR = inactivos: 88, activos: 30, RHIR = inactivos: 45, activos: 38), al optimizar los modelos bajo esta métrica se tiende a que la clasificación de estas moléculas sea siempre como inactivas; debido a que así se reduciría el error y se aumentarían los aciertos ^{171,172}.

Tabla 3.7 modelos de *machine learning* y métricas de desempeño utilizadas para la evaluación de la clasificación de moléculas activas e inactivas frente al FT LasR

Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC
Random Forest Classifier	0.7780	0.8287	0.5208	0.6402	0.5617	0.4184	0.4301
Extreme Gradient Boosting	0.7716	0.8199	0.4750	0.6124	0.5200	0.3816	0.3918
Light Gradient Boosting Machine	0.7683	0.8180	0.4931	0.5959	0.5330	0.3837	0.3897
Extra Trees Classifier	0.7748	0.7999	0.5208	0.6307	0.5578	0.4118	0.4231
CatBoost Classifier	0.7616	0.7997	0.3597	0.6354	0.4487	0.3154	0.3405
K-nearest Neighbors Classifier	0.7447	0.7843	0.5569	0.5433	0.5392	0.3661	0.3739
Logistic Regression	0.7452	0.7770	0.4569	0.5556	0.4942	0.3284	0.3347
Gradient Boosting Classifier	0.7349	0.7642	0.4097	0.5516	0.4639	0.2940	0.3037
Decision Tree Classifier	0.7380	0.6945	0.5708	0.5585	0.5497	0.3697	0.3796
Ada Boost Classifier	0.6814	0.6701	0.3972	0.4339	0.4076	0.1944	0.1970
Quadratic Discriminant Analysis	0.5202	0.6302	0.8792	0.3580	0.5080	0.1833	0.2474
Linear Discriminant Analysis	0.6208	0.6300	0.5333	0.3709	0.4344	0.1637	0.1731
Naive Bayes	0.6278	0.6183	0.5944	0.3861	0.4635	0.1976	0.2149
SVM - Linear Kernel	0.6710	0.000	0.6083	0.4984	0.4902	0.2681	0.3044
Ridge Classifier	0.7117	0.000	0.5097	0.4811	0.4862	0.2919	0.2952

Para solucionar el desbalance entre las bases de datos, se utiliza la metodología *Synthetic Minority Over-sampling Technique* (SMOTE) propuesta por Chawla et al ¹⁷³, donde se genera un sobre muestreo con muestras sintéticas del grupo minoritario; permitiendo un mejor balance entre los dos grupos y mejorando la clasificación de las moléculas activas. Por otro lado, se decide utilizar la métrica F1 como

parámetro de optimización, la cual al relacionar la precisión (verdaderos positivos sobre el total verdadero) y el *recall* (verdaderos positivos sobre el total identificado) podrá generar los mejores hiperparámetros para llegar al modelo con el mayor número de aciertos y el menor número de errores.

De los *molecular fingerprints* evaluados los que mostraron tener los valores más altos en F1 fueron ECFP2 y ECFP4, que son clasificados dentro del grupo de la conectividad. Para mejorar la clasificación dada por estos *fingerprints*, se decidió probar los valores de 512 y 1024 como longitud de cada cadena de bits, además de eliminar las columnas que no poseían algún valor dentro de ellas; logrando encontrar los mejores clasificadores dentro de los *fingerprints* ECFP4 de longitud de 512 bits. Con este *fingerprint* fueron optimizados todos los modelos de clasificación hasta encontrar los valores F1 más altos, la exactitud mayor y el error de clasificación más bajo; siendo seleccionados los modelos *K-Nearest Neighbors Classifier* (k-NN) y *Extreme Gradient Boosting* (Xgb).

Al ser optimizados estos se obtienen los valores reportados en la figura 3.21; aquí para el caso del calificador LasR se encuentra que la mayor exactitud está en el modelo Xgb (79.3 %) para los compuestos activos, sin embargo, dicho modelo presenta un error de clasificación superior al 50 % en este mismo grupo, generando una gran cantidad de falsos positivos dentro de las moléculas clasificadas. En comparación el modelo de clasificación k-NN presenta un error de clasificación y una exactitud más bajas que el Xgb (33.3 %, 62.1 % respectivamente); generando en este la reducción de los falsos positivos y de los verdaderos positivos, por lo que se propone el consenso entre los dos modelos como el método adecuado para encontrar moléculas con posible actividad (Figura 3.21a).

En el caso de los clasificadores de moléculas con actividad sobre PqsR, se encontraron que estos poseen exactitudes y errores de clasificación similares en ambos modelos (k-NN y Xgb); llegando ambos a una exactitud aproximada del 75 % y un error del 50 %, siendo evidente la fuerte cantidad de falsos positivos que podrían llegar a clasificar ambos (Figura 3.21b).

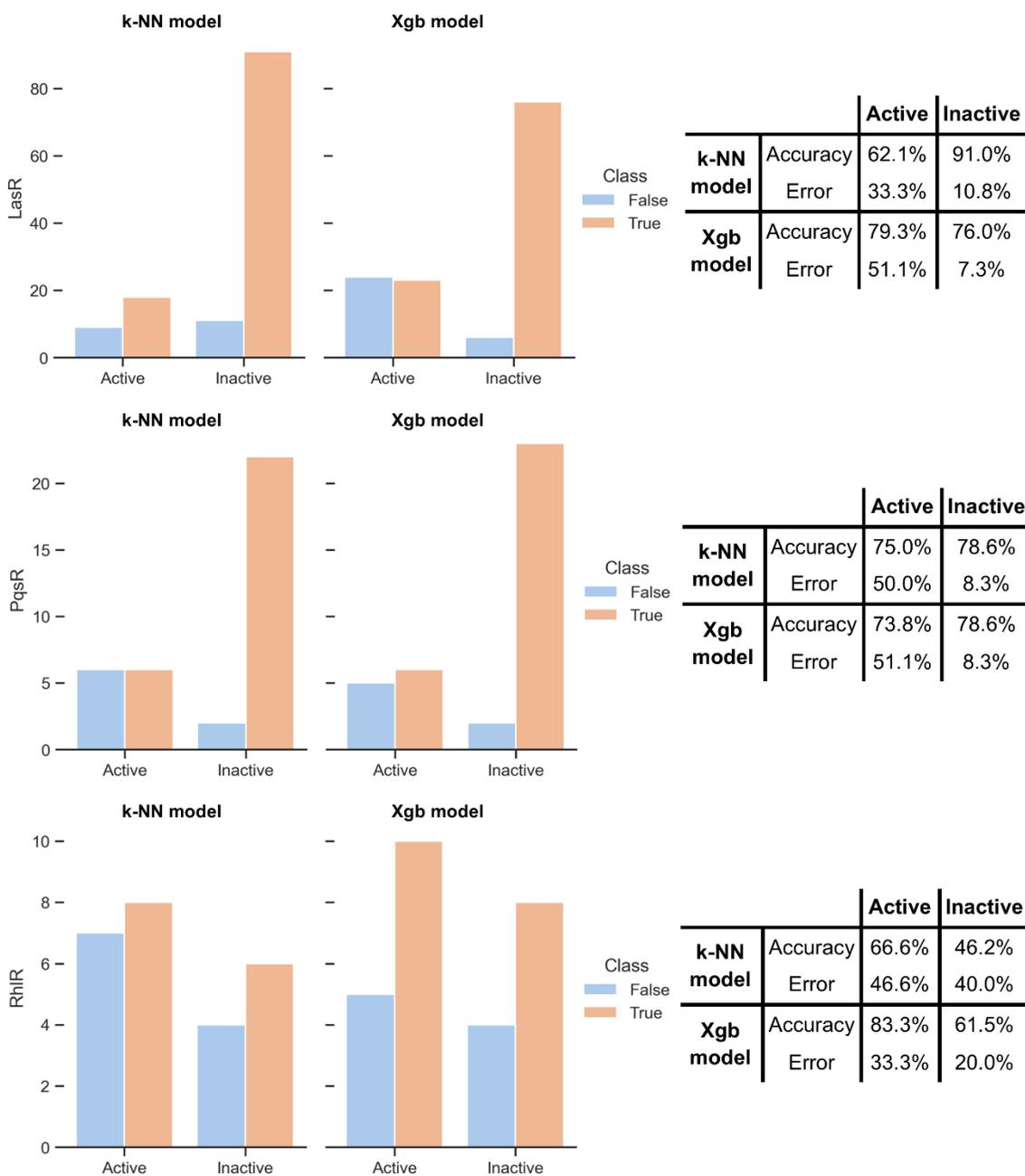


Figura 3.21 exactitud y error para los modelos de clasificación *k-nearest neighbors* (k-NN Model) y *Extreme gradient boosting* (Xgb Model) frente a posibles moléculas activas e inactivas en los FT's a) LasR, b) PqsR y c) RhIR.

En cuanto a los clasificadores de compuestos activos frente a RhIR, se observó que el que presentaba las mejores métricas era Xgb, sin embargo, como se evidencia en la figura 3.21c el número de moléculas

con las que fueron entrenados estos modelos no superaron las 10 en ambas categorías siendo estas insuficientes para generar modelos con una calidad adecuada, problema que también sucedía con PqsR, ya que éste no superó las 20 moléculas por categoría; por lo que se requieren un mayor número de puntos para entrenar adecuadamente estos modelos predictivos.

Debido a que el clasificador *Extreme Gradient Boosting* (Xgb) trabaja bajo la metodología de árbol de decisión; donde el modelo realiza una serie de múltiples preguntas hasta llegar a la clasificación del problema que se le presenta^{174,175}; se puede obtener de estos, el peso de cada una de las características con las que es alimentado el modelo para la clasificación dentro de cada uno de los grupos (activos e inactivos), ya que se estos modelos son dependientes del *molecular fingerprint* ECFP4 y que éste representa a las moléculas como bits asociados a una parte de los compuestos, fue posible extraer que parte estructural que tenía una mayor relevancia en la clasificación de la actividad frente a cada uno de los FT's.

Para el caso del clasificador Xgb hecho para predecir la actividad frente a LasR, se encontraron 5 características relevantes para la clasificación como activas, conforme a lo mostrado en la figura 3.22. Aquí se evidencia, que no necesariamente las características estructurales más complejas son indicativas de actividad sobre este FT, siendo los bits de radio 1 como el 295 y el 483 (carbono sp² e hidroxilo) los más relevantes en la clasificación de las moléculas activas. En cuanto a la clasificación de las moléculas inactivas, todos los bits importantes fueron de radio 3 y las fracciones estructurales resultaron ser más complejas, pasando por carbonos aromáticos con condiciones específicas (bits: 99, 135) hasta carbonos de cadenas laterales cercanos a oxígeno y nitrógeno (bits: 510, 377).

En las características estructurales que tuvieron mayor peso dentro de las moléculas activas en el clasificador Xgb para el FT PqsR, se destaca el flúor representado por el bit 429 de radio 1 que se encuentra unido a un carbono sp³; los otros bits son de un radio mucho mayor en donde el centro del 393 de radio 3 resultó ser un nitrógeno perteneciente a un anillo y el 63 también de radio 3 donde se representa el carbono que une los dos anillos aromáticos de la quinoleína; siendo esta última parte estructural similar con la de los inactivos, en donde la variación es el grado de saturación de estos biciclos como se observa en la figura 3.23.

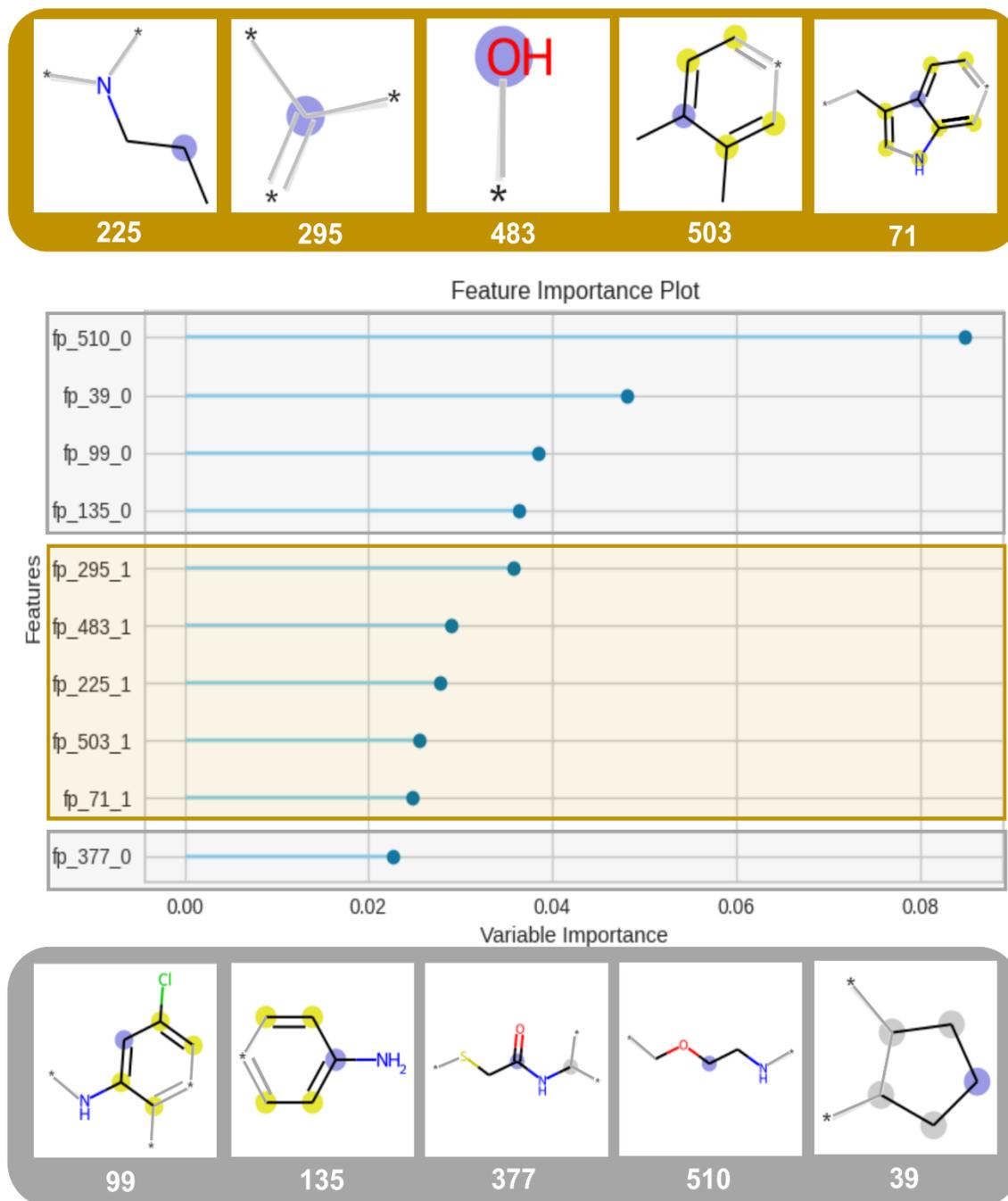


Figura 3.22: bits o características más relevantes para la clasificación de moléculas activas frente a LasR, según el modelo *Extreme gradient boosting* (Xgb). En la parte superior e inferior de la figura se representan las partes estructurales asociadas a las 2 actividades, siendo las superiores (naranja) activas y las inferiores (gris) inactivas; en el medio se observa la influencia de cada uno de los bits y su importancia para la clasificación.

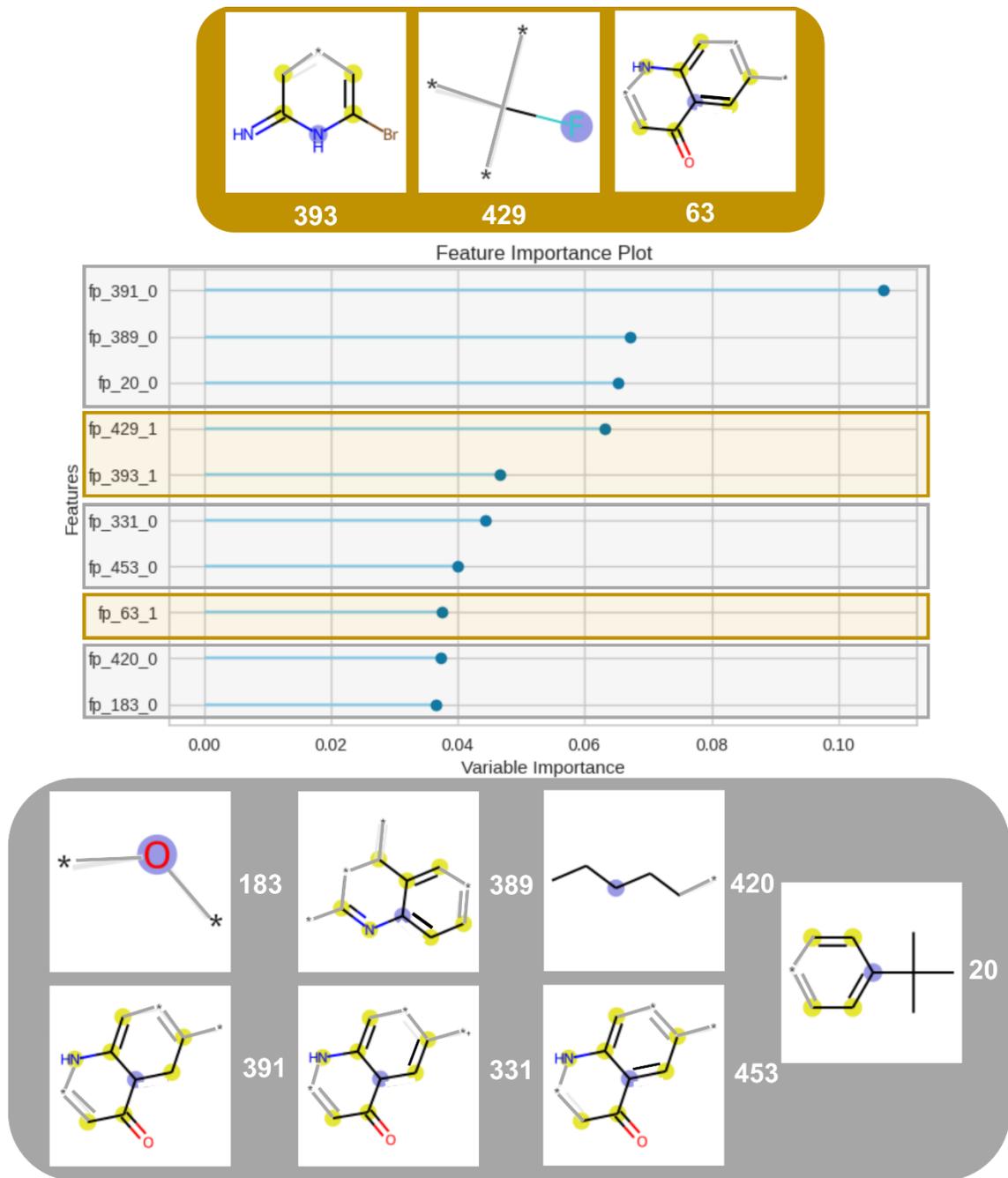


Figura 3.23: bits o características más relevantes para la clasificación de moléculas activas frente a PqsR, según el modelo *Extreme gradient boosting* (Xgb). En la parte superior e inferior de la figura se representan las partes estructurales asociadas a las 2 actividades, siendo las superiores (naranja) activas y las inferiores (gris) inactivas; en el medio se observa la influencia de cada uno de los bits y su importancia para la clasificación.

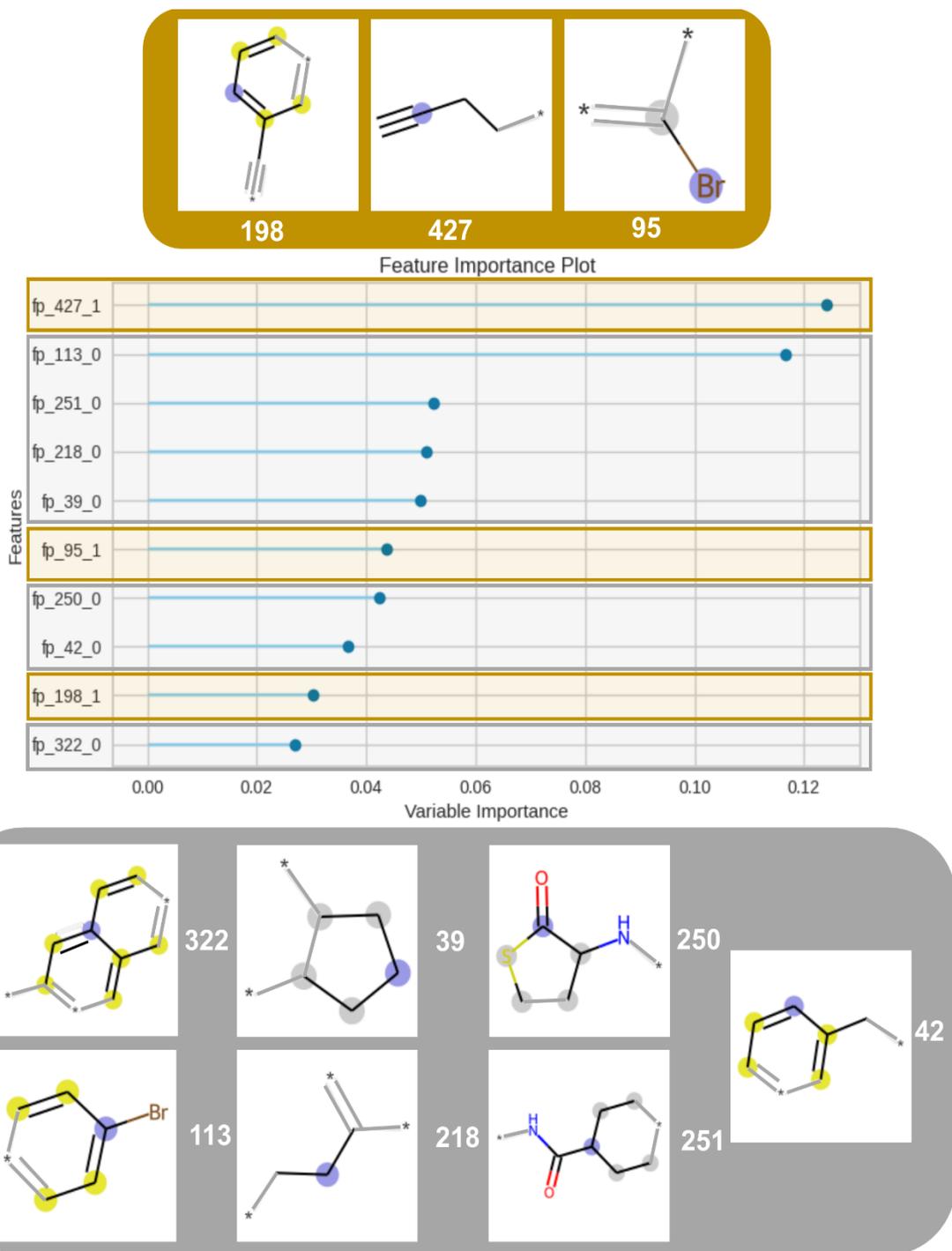


Figura 3.24: bits o características más relevantes para la clasificación de moléculas activas frente a RhIR, según el modelo *Extreme gradient boosting* (Xgb). En la parte superior e inferior de la figura se representan las partes estructurales asociadas a las 2 actividades, siendo las superiores (naranja) activas y las inferiores (gris) inactivas; en el medio se observa la influencia de cada uno de los bits y su importancia para la clasificación.

Otros bits que estos modelos encontraron como relevantes en las moléculas inactivas fueron los 183, 420, 389 y 20, representado estos un: oxígeno conectado a carbonos sp^3 , un carbono entre una cadena lateral, un carbono entre el biciclo de la quinoleína cercano al nitrógeno y el carbono de un anillo aromático conectado a un isobutano respectivamente. En este clasificador se evidencia los problemas que tiene estos *fingerprints* para distinguir estos compuestos, que poseen grandes cadenas laterales; debido a la similitud entre los bits encontrados de los activos e inactivos.

En los bits usados por el clasificador Xgb para encontrar moléculas activas en el FT RhIR, se resaltan el 427 y el 198 que poseen en su cercanía un grupo alquino, como se observa en la figura 3.24; esto debido a que uno de los compuestos usados para construir este clasificador poseía este grupo funcional; y al partir de tan pocos puntos, éste determinó que este grupo era relevante para la actividad. Aunque el clasificador lo tome como un punto importante para la actividad, es necesario mencionar que este no siempre será indicativo de actividad. En cuanto a las inactivas el modelo ponderó a 7 bits de diferentes longitudes como relevantes para la clasificación de esta actividad, teniendo el mayor peso el bit 113 de radio 3, donde se representa a un carbono sp^2 de un anillo aromático unido a bromo.

De este último análisis cabe mencionar que, aunque estos clasificadores dan mayor importancia a unos bits determinados, no significa que estos necesariamente al estar presentes en una molécula generen algún tipo de actividad; si no que son características relevantes para cada una de los compuestos dentro de las categorías asignadas.

3.5 Cribado virtual

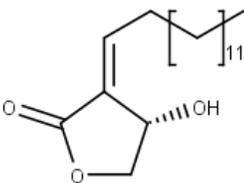
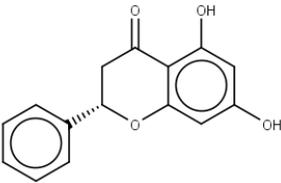
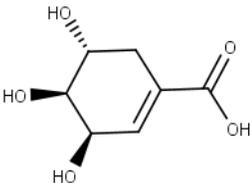
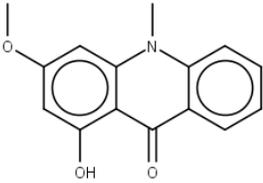
Según los estudios previamente realizados por Yang et al ¹⁷⁰ y Mayr et al ¹⁰⁶ donde se propone el uso de modelos de *machine learning* para la búsqueda de nuevas moléculas con actividad biológica y la aplicación de los PLIF propuestas por Vass et al ¹⁷⁶ y Velásquez-Libera et al ¹⁷⁷; en este trabajo se propone una combinación de ambos métodos, uno basado en los ligandos (*machine learning*) y el otro en la estructura (PLIF) con el fin de aumentar la tasa de éxito del cribado de la BD QuiProNab.

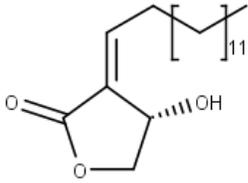
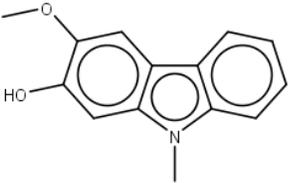
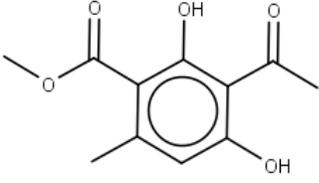
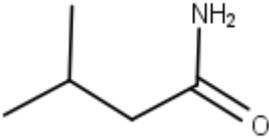
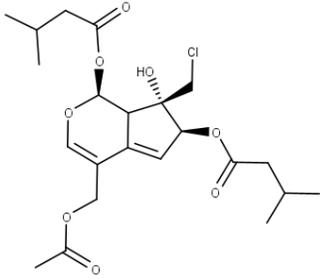
Para el cribado virtual, a las moléculas ya curadas de la BD QuiProNaB (propia del grupo de investigación) se les calculó el *molecular fingerprint* ECFP4 de 512 bits, este código fue filtrado con los modelos k-NN y Xgb de cada FT; encontrando que 10, 3 y 185 compuestos fueron clasificados dentro del grupo de activos para LasR, PqsR y RhIR respectivamente en ambos modelos, como se observa en la tabla 3.8.

Según los estos resultados, los clasificadores de PqsR tienden al sobreajuste; debido a que estos seleccionaron a muy pocas moléculas dentro del grupo de activas, mientras que los de RhIR por el

contrario fallaron al clasificar a los compuestos; todo lo anterior se debe a que se contaba con un bajo número de compuestos para el entrenamiento de estos modelos y la alta similitud estructural entre las moléculas activas e inactivas. Los clasificadores asociados a la actividad sobre LasR presentaron un mejor balance, sin embargo, es importante destacar que estos también tienden al sobreajuste.

Tabla 3.8 mejores resultados del cribado virtual. Dónde: ML significa *machine learning*, PLIF significa *Protein Ligand Interaction fingerprint*, AA agonistas son los aminoácidos tipo agonista con los que interacciona la molécula y AA antagonistas son los son los aminoácidos tipo antagonista con los que interacciona la molécula.

	Filtro - ML	Filtro PLIF's	Ejemplos	AA. agonistas	AA. antagonistas	Puntaje
LasR	10	1	 <p>Litseanolido C2</p>	Thr 115 Thr 75 Ala 127	Tyr 47 Ala 105 Thr 80 Leu 36 Ile 52 Tyr 56	3
			 <p>5, 7-dihydroxiflavanona</p>	Leu 40 Ala 127	Tyr 47 Leu 36 Ile 52 Tyr 46	2
			 <p>(-) ácido shikimico</p>	Thr 115 Thr 75 Ala 127	Tyr 47 Thr 80 Leu 36 Ile 52 Tyr 56	2
PqSR	3	1	 <p>1-hidroxi-3-metoxi-N-metil acridona</p>	Leu 197 Arg 209	Ala 102 Ile 149 Leu 207 Phe 221 Tyr 258	3

RhIR	185	4		<p>Gln 194 Ser 196 Leu 197 Arg 209</p>	<p>Ala 102 Ile 149 Leu 207 Phe 221</p>	0	
			Litseanolido C2				
				<p>Gln 194 Ser 196 Leu 197 Arg 209</p>	<p>Ala 102 Ile 149 Leu 207 Phe 221</p>	0	
			3-metoxi-9-metil-9H-carbazo-2-ol				
				<p>Ser 135</p>	<p>Val 60 Tyr 64 Leu 69 Tyr 72 Ala 83</p>	4	
			Hematomoato de metilo				
				<p>Ser 135</p>	<p>Val 60 Tyr 64 Leu 69 Tyr 72 Ala 83</p>	4	
			Isovaleramida				
				<p>Ser 135</p>	<p>Val 60 Tyr 64 Leu 69 Tyr 72 Ala 83</p>	4	
			Clorhidrina valtrato				

A las moléculas resultantes activas de los clasificadores basados en *machine learning*, se les determina el fingerprint de interacción ligando-proteína a partir de su acoplamiento molecular, para luego realizar la resta de las interacciones con los aminoácidos asociados a la actividad antagonista contra los asociados a la actividad agonista como se muestra en ecuación 2.2; dando lugar al puntaje mostrado en la tabla 3.8. En esta tabla se evidencia que en las moléculas seleccionadas para LasR y RhIR, las interacciones con los aminoácidos claves en la actividad agonista fueron bajas en comparación con los aminoácidos clave en la actividad antagonista, generando puntajes altos. En el caso de PqsR los compuestos seleccionados como activos por los clasificadores de *machine learning* tuvieron muchas más interacciones con los residuos asociados a la actividad agonista, lo que muestra la dificultad para la selección de los compuestos antagonistas para este FT.

En la figura 3.25 se pueden observar los compuestos con mejor puntaje dentro del cribado y el balance de sus interacciones entre los aminoácidos clave para la actividad antagonista y agonista. En el caso de LasR, el litseanólido C2 coincide con las características tabuladas en la tabla 3.2; este posee un anillo de 5 miembros con 3 aceptores de puente de hidrogeno y un donador de puente hidrogeno; además de tener una cadena lateral de 14 átomos de carbono, sin embargo, debido a su similitud con el AI se debería verificar experimentalmente la hipótesis de que este compuesto posee actividad antagonista.

El compuesto 1-hidroxi-3-metoxi-N-metil acridona fue seleccionado por el cribado, como el de mayor posibilidad de ser un compuesto antagonista de PqsR (ver figura 3.25), este posee un núcleo de acridina; el cual tiene una alta similitud con el *scaffold* de quinoleina hallado relevante para la actividad sobre este FT. Adicional a lo anterior este compuesto posee los donadores y aceptores de puente de hidrogeno identificados como relevantes para la actividad deseada; sin embargo, la falta de las interacciones con la región hidrofóbica del sitio de unión del FT, por la ausencia de una cadena carbonada genera la necesidad de verificar su actividad a través de ensayos biológicos.

El ultimo compuesto mostrado en la figura 3.25, el hematomato de metilo, es un potencial antagonista del FT RhIR, tiene el anillo aromático de 6 miembros como núcleo base, sin embargo, esa molécula no cumpliría con el número de átomos mínimos para que pueda generar actividad sobre RhIR, tal como se ve en la tabla 3.2 y en la figura del farmacóforo de la actividad antagonista sobre RhIR mostrado en la figura 3.20, por lo cual es recomendable confirmar la actividad antagonista experimentalmente.

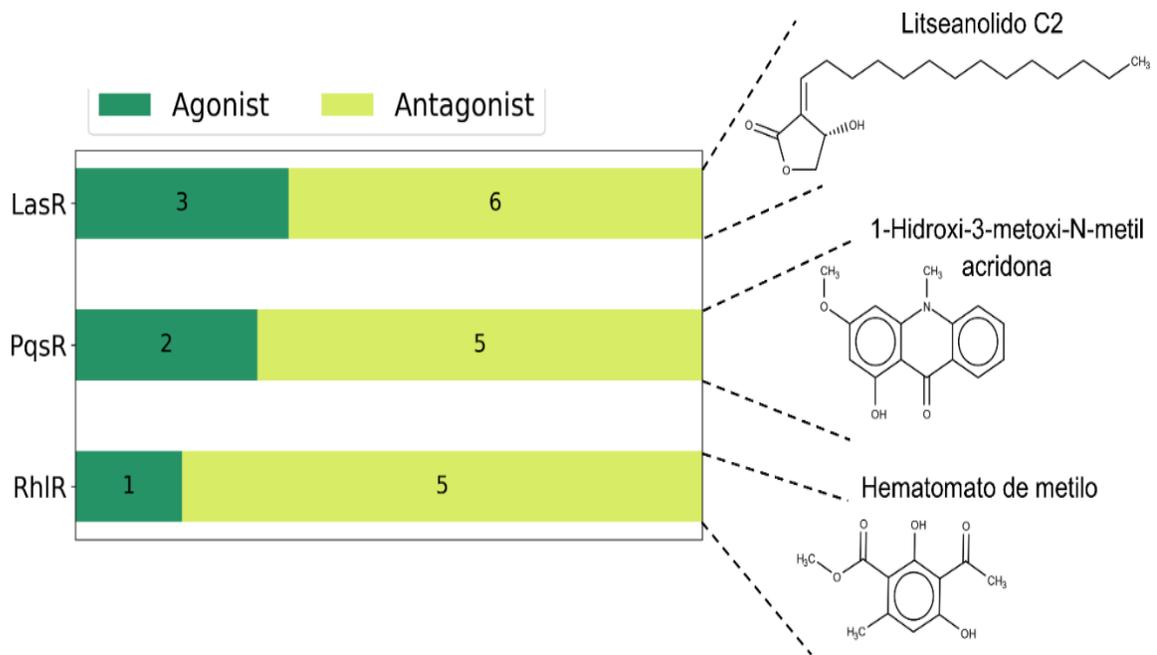


Figura 3.25 compuestos escogidos por el cribado como los más probables de tener actividad antagonista sobre los FT's. A la izquierda de la figura se observa el número de interacciones con los aminoácidos agonistas (verde) y los antagonistas (amarillo) y a la derecha las estructuras de las moléculas resultantes.

De los compuestos encontrados se revisaron las familias de las especies de donde fueron extraídos, representando el resultado de este análisis en la figura 3.26; aquí se puede observar que para los tres factores de transcripción los compuestos presentes en las familias Rutaceae, Piperaceae y Lauraceae serían una gran fuente de compuestos con la actividad deseada. Adicionalmente, este gráfico nos permite pensar en que un extracto de una de las especies del género *Zanthoxylum* podría llegar a tener actividad múltiple sobre los tres FT's, potenciando la actividad contra *P. aeruginosa* reportada previamente por diversos autores¹⁷⁸⁻¹⁸⁰.

Es importante tener en cuenta que las especies con mayor prevalencia (Rutaceae, Piperaceae y Lauraceae) de posibles compuestos también tienen asociados un mayor número de moléculas en esta BD; donde la familia Lauraceae posee el 29.3 %, la Rutaceae el 26.6 % y la Piperaceae el 16.3 %, generando así un posible sesgo en este resultado.

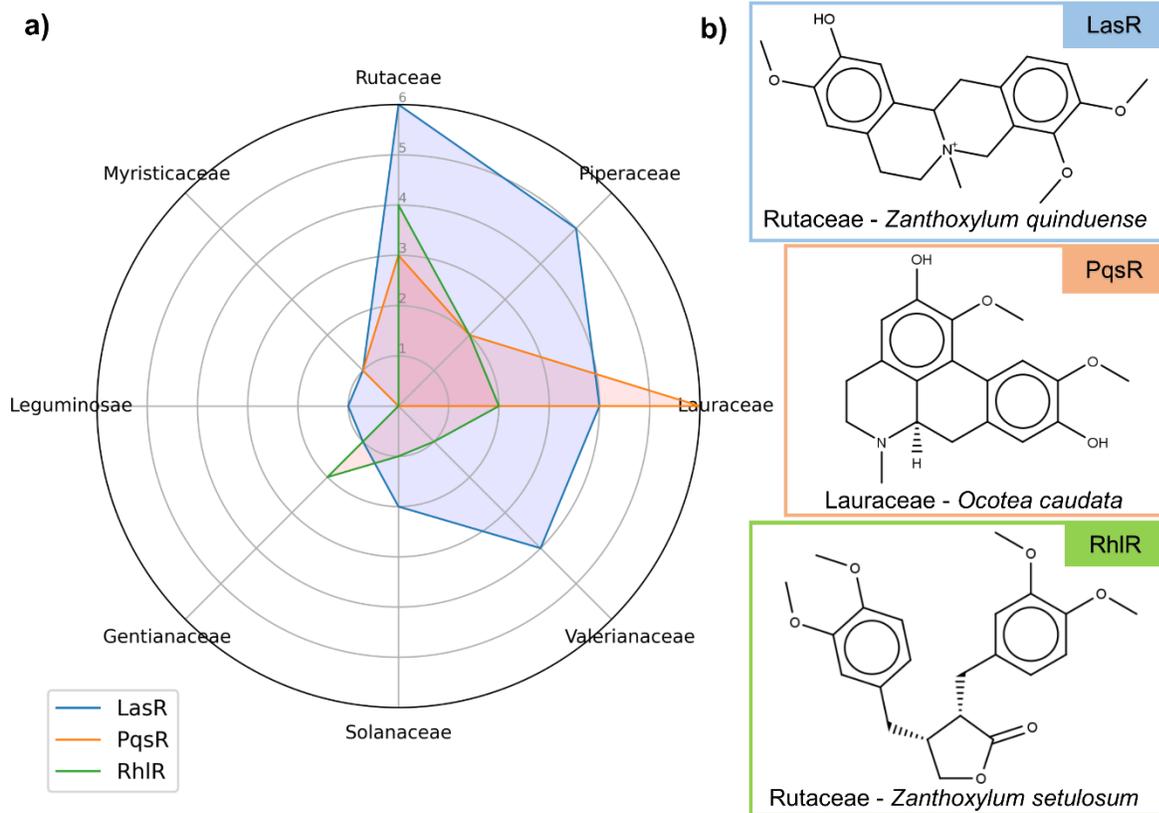


Figura 3.26 Gráfico de radar con las familias de las especies de donde fueron extraídos los compuestos con mayor potencial antagonista sobre los FT's. A la izquierda se observa el mapa de radar, donde los ejes representan el número de compuestos por Familias que podrían tener actividad antagonista contra LasR (azul), PqsR (naranja) y RhIR (verde); a la derecha un ejemplo de compuestos con el mejor puntaje por cada FT en la familia con mayor prevalencia, en azul el ejemplo de LasR, en naranja PqsR y en verde RhIR.

4 Conclusiones y recomendaciones

Se llevaron a cabo múltiples análisis basados en el DiFAC, éstos permitieron obtener características **estructurales asociadas a la actividad antagonista de las moléculas sobre los FT's LasR, PqsR y RhIR** asociados al *quorum sensing* en *P. aeruginosa*; además de la caracterización de los posibles aminoácidos involucrados en este tipo de actividad frente a estos blancos.

A partir de la construcción de BD's con moléculas que clasificadas con el tipo de actividad (agonista, antagonista e inactivo) por cada uno de los **FT's**, se concluye que no existen núcleos base únicos asociados con la actividad antagonista y que los compuestos a los cuales se les conoce la actividad **dentro de estas BD's pueden llegar a ser MOA-cliffs y activity-cliffs**, en donde un pequeño cambio estructural puede llevar a un viraje del tipo de actividad o la modificación en los valores de actividad.

Para LasR se asociaron los núcleos base tipo lactónico, trifenílicos y cadenas carbonadas iguales o superiores a los 10 carbonos con potencial actividad sobre este blanco. Adicionalmente, se encontró que estas moléculas debían tener por lo menos dos aceptores de puente de hidrogeno, un donador de puente de hidrogeno, además de una parte que interaccione hidrofóticamente con los residuos del sitio de unión. Por último, se evidenció la necesidad de moléculas con gran tamaño, superior a los 15 átomos.

En PqsR los bicíclicos nitrogenados como quinoleínas y quinazolininas resultaron ser los núcleos base más comunes dentro de los compuestos activos, además de la cadena carbonada con 10 o más carbonos; para este FT se caracterizaron como necesarios para la actividad un aceptor de puente de hidrogeno, un donador de puente de hidrogeno y múltiples interacciones de tipo hidrofóbico. Adicional a lo anterior, las moléculas con actividad sobre este FT suelen tener más de 15 átomos.

Los ciclos de cinco miembros, las cadenas alifáticas y los anillos de 6 miembros fueron los núcleos base que prevalecieron en las moléculas que poseían actividad sobre el FT RhIR; aquí dichas moléculas suelen poseer dos aceptores de puente de hidrogeno, un donador de puente de hidrogeno, interacciones π - π e hidrofóbicas. En cuanto al tamaño molecular se encontraron que estas moléculas suelen estar entre los 8 a 12 átomos.

El análisis de DFBE permitió establecer una posible estructura para el FT RhIR, además de encontrar posibles aminoácidos con relevancia sobre la actividad agonista y antagonista para cada uno de los **FT's**; **aquí las** herramientas de los PLIF y las dinámicas moleculares fueron claves para la selección de los residuos según el tipo de actividad y ver los posibles cambios conformacionales que ocurren en los complejos con moléculas según su tipo de actividad. Sin embargo, se recomienda realizar las dinámicas

con un tiempo mayor para corroborar estos resultados, así como los posibles ensayos experimentales (Cristalización de los complejos y difracción de rayos X, RMN o crio microscopía electrónica).

Para LasR los aminoácidos Thr 115, Leu 40, Thr 75 y Ala 127 se asociaron con la actividad agonista; mientras Tyr 47, Ala 105, Thr 80, Leu 36, Ile 56 y Tyr 56 fueron relevantes para la actividad antagonista. En cuanto a PqsR la Ser 196, Leu 197, Gln 194 y Arg 209 tuvieron importancia en la actividad agonista, adicional a lo anterior la actividad antagonista fue relacionada con los residuos Phe 221, Ala 102, Leu 207, Ile 149 y Tyr 258. Por último, a la actividad agonista del FT RhIR se le asoció los aminoácidos: Leu 107, Leu 69, Phe 101, Thr 121, Val 133 y Ser 135; y a la antagonista los residuos: Val 60, Leu 69, Tyr 64, Tyr 72 y Ala 83.

Los modelos farmacofóricos propuestos en este trabajo fueron construidos con las moléculas que presentaron mejor actividad de las bases de datos, resumiendo las características estructurales y electrónicas necesarias para esta actividad, incluyendo las posibles distancias intramoleculares que deben **ser tenidas en cuenta para esta actividad sobre los FT's. Sin embargo, se recomienda validar** estos modelos y buscar el método que permita la inclusión de las cadenas laterales dentro de ellos, **sobre todo para los FT's LasR y PqsR.**

Se construyeron modelos de clasificación basados en *machine learning* que permiten la distinción de moléculas activas e inactivas frente a **los FT's; estos tuvieron desempeños acordes a la información de** partida, constituyéndose como uno de los grandes problemas la falta de una mayor cantidad de moléculas con actividad conocida experimentalmente. A pesar de lo anterior se lograron obtener modelos con valores de desempeño similares a los encontrados en la literatura, pero se recomienda buscar nuevas formas de representación molecular y entrenar modelos no supervisados más poderosos, como las redes neuronales profundas a partir de estas bases de datos, todo esto con el fin de mejorar el clasificador de RhIR y eliminar el sobreajuste en el clasificador de PqsR.

Por último, se diseñó una metodología para la búsqueda de posibles moléculas antagonistas contra **estos FT's, integrando la información** recolectada sobre el comportamiento de los compuestos antagonistas ya conocidos y de los posibles cambios estructurales que inducen sobre estas proteínas. Aquí se encontraron características estructurales como los anillos de 5 y 6 miembros para los antagonistas de LasR y RhIR; mientras que los bicíclicos de 6 miembros fueron importantes para esta **actividad sobre PqsR, la única característica en común para los tres FT's fue cadenas carbonadas de más de 5 carbonos.**

Todos estos compuestos con actividad antagonista presentaron mínimo un aceptor y un donador para **todos los FT's, mientras que para LasR y RhIR se presentó un aceptor adicional en todos estos compuestos.** Por último, **se encontró que todas las moléculas con dicha actividad sobre estos FT's** poseen mínimo 8 átomos y algunas de ellas alcanzaban hasta los 15 átomos. Para la actividad antagonista mediante los estudios basados en la estructura pudimos caracterizar que interacciones con **tirosinas, alaninas y leucinas en todos los sitios de unión de los FT's;** además de otros aminoácidos específicos por cada FT.

El análisis de la BD QuiProNaB arrojó como posibles antagonistas al litseanolido C2, al 1-hidroxi-3-metoxi-N-metil acridona y al hematoma de metilo de los FTs: LasR, PqsR y RhIR respectivamente; a los cuales se les recomienda llevar los ensayos experimentales correspondientes para la validación de su actividad biológica y del modelo de cribado. Adicionalmente el estudio de esta BD permite recomendar la búsqueda de moléculas sobre especies de las familias Rutaceae, Lauraceae y Piperaceae como posibles antagonistas de los FT's y a los extractos de las especies pertenecientes al género *Zanthoxylum* como posibles antagonistas múltiples de los tres factores de transcripción.

A. Anexo: Base de datos molecular LasR

ID	SMILES	Actividad	Valor (nM)	Tipo de actividad	pIC50	Referencia
AM10	<chem>S(CCCCCCCCCC(=O)C C(=O)N[C@H]1CCOC1=O)(=O)(=O)C</chem>	INACTIVO	1500	EC50	5.82	1
AM2	<chem>S=C=NC[C@H](F)CCCCC CCC(=O)CC(=O)N[C@H]1CCOC1=O</chem>	INACTIVO	154000	IC50	3.81	1
AM3	<chem>Br[C@H](CCCCCCCC(=O))CC(=O)N[C@H]1CCOC1=O)CN=C=S</chem>	INACTIVO	95000	IC50	4.02	1
AM4	<chem>Cl[C@H](CCCCCCCC(=O))CC(=O)N[C@H]1CCOC1=O)CN=C=S</chem>	INACTIVO	143000	IC50	3.84	1
AM5	<chem>SCCCCCCCCCC(=O)CC(=O)N[C@H]1CCOC1=O</chem>	INACTIVO	115000	IC50	3.94	1
AM7	<chem>O1CC[C@H](NC(=O)CC(=O)CCCCCCC[C@H]2OC2)C1=O</chem>	AGONISTA	23	EC50	7.64	1
AM8	<chem>S(CCCCCCCCCC(=O)CC(=O)N[C@H]1CCOC1=O)(=O)(=O)C</chem>	INACTIVO	10000	EC50	5	1
AM9	<chem>S(CCCCCCCCCC(=O)CC(=O)N[C@H]1CCOC1=O)(=O)(=O)C</chem>	INACTIVO	6000	EC50	5.22	1
BO1	<chem>O1CC[C@H](NC(=O)CC(=O)CCCCCCCC)C1=O</chem>	AGONISTA	1.5	EC50	8.82	2
BO11	<chem>O1CC[C@H](OC(=O)CC(=O)CCCCCCCCNC(=O)C C(=O)CCCCCCCC)C1=O</chem>	INACTIVO	250000	IC50	3.6	2
BO3	<chem>S1CC[C@@H](NC(=O)C C(=O)CCCCCCCC)C1=O</chem>	AGONISTA	1.5	EC50	8.82	2
BO4	<chem>O=C1NCC[C@H]1NC(=O))CC(=O)CCCCCCCC</chem>	AGONISTA	30	EC50	7.52	2
BO5	<chem>O=C1CCC[C@H]1NC(=O))CC(=O)CCCCCCCC</chem>	AGONISTA	15	EC50	7.82	2

BO6	<chem>O=C(CCCCCCCC)CC(=O)NC1CCCC1</chem>	AGONISTA	160	EC50	6.8	2
BO7	<chem>O1C[C@H](NC(=O)CC(=O)CCCCCCCC)CC1</chem>	AGONISTA	910	EC50	6.04	2
BO8	<chem>O[C@H]1CCC[C@H]1NC(=O)CC(=O)CCCCCCCC</chem>	INACTIVO	1900	EC50	5.72	2
BO9	<chem>O1CCC[C@H]1CNC(=O)CC(=O)CCCCCCCC</chem>	AGONISTA	260	EC50	6.59	2
BR5	<chem>O1CC[C@@H](n2nnc(c2)CCCCCCCC)C1=O</chem>	INACTIVO	307700	IC50	3.51	3
CA1	<chem>Brc1cc(Br)cc(CC(=O)Nc2cccc2[N+](=O)[O-])c1OC(=O)c1cccc1Cl</chem>	AGONISTA	372	EC50	6.43	4
CA10	<chem>Brc1cc(Br)cc(CC(=O)Nc2cccc2[N+](=O)[O-])c1OC(=O)c1cccc1C</chem>	AGONISTA	757	EC50	6.12	4
CA11	<chem>Brc1cc(Br)cc(CC(=O)Nc2cccc2[N+](=O)[O-])c1OC(=O)c1cc(ccc1)C</chem>	INACTIVO	50000	EC50	4.3	4
CA12	<chem>Brc1cc(Br)cc(CC(=O)Nc2cccc2[N+](=O)[O-])c1OC(=O)c1ccc(cc1)C</chem>	INACTIVO	50000	EC50	4.3	4
CA13	<chem>Brc1cc(Br)cc(CC(=O)Nc2cccc2[N+](=O)[O-])c1OC(=O)c1cccc1OC</chem>	AGONISTA	246	EC50	6.61	4
CA14	<chem>Brc1cc(Br)cc(CC(=O)Nc2cccc2[N+](=O)[O-])c1OC(=O)c1cccc1[N+](=O)[O-]</chem>	AGONISTA	489	EC50	6.31	4
CA15	<chem>Brc1cc(Br)cc(CC(=O)Nc2cccc2[N+](=O)[O-])c1OC(=O)c1ccc(F)cc1Br</chem>	AGONISTA	436	EC50	6.36	4
CA16	<chem>Brc1cc(Br)cc(CC(=O)Nc2cccc2[N+](=O)[O-])c1OC(=O)c1c(Cl)cccc1Cl</chem>	AGONISTA	313	EC50	6.5	4
CA17	<chem>Brc1cc(Br)cc(CC(=O)Nc2cccc2[N+](=O)[O-])c1OCc1c(Cl)cccc1Cl</chem>	AGONISTA	161	EC50	6.79	4
CA18	<chem>Brc1cc(Br)cc(CC(=O)Nc2cccc2[N+](=O)[O-])c1OC(=O)c1cccc1</chem>	INACTIVO	50000	EC50	4.3	4

84 Contribución a la identificación de potenciales antagonistas asociados al quorum sensing en *Pseudomonas aeruginosa* a través de métodos quimioinformáticos.

CA2	<chem>Brc1cc(Br)cc(CC(=O)Nc2cccc2[N+](=O)[O-])c1OC(=O)c1cc(Cl)ccc1</chem>	INACTIVO	50000	EC50	4.3	4
CA21	<chem>Brc1cc(Br)cc(CC(=O)Nc2cc([N+](=O)[O-])ccc2)c1OC(=O)c1cccc1Cl</chem>	INACTIVO	50000	EC50	4.3	4
CA22	<chem>Brc1cc(Br)cc(CC(=O)Nc2ccc([N+](=O)[O-])cc2)c1OC(=O)c1cccc1Cl</chem>	INACTIVO	50000	EC50	4.3	4
CA25	<chem>Brc1cc(Br)cc(CC(=O)Nc2cc(F)c(F)cc2[N+](=O)[O-])c1OC(=O)c1cccc1Cl</chem>	AGONISTA	334	EC50	6.48	4
CA26	<chem>Brc1cc(Br)cc(CC(=O)Nc2ncncc2)c1OC(=O)c1cccc1Cl</chem>	AGONISTA	601	EC50	6.22	4
CA27	<chem>Brc1cc(Br)cc(CC(=O)Nc2cc[nH+]cc2[N+](=O)[O-])c1OC(=O)c1cccc1Cl</chem>	AGONISTA	446	EC50	6.35	4
CA28	<chem>Brc1cc(Br)cc(CC(=O)N[C@H]2CCOC2=O)c1OC(=O)c1cccc1Cl</chem>	AGONISTA	0.896	EC50	9.05	4
CA29	<chem>Brc1cc(Br)cc(CC(=O)Nc2ncccc2)c1OC(=O)c1cccc1Cl</chem>	INACTIVO	42000	IC50	4.38	4
CA3	<chem>Brc1cc(Br)cc(CC(=O)Nc2cccc2[N+](=O)[O-])c1OC(=O)c1ccc(Cl)cc1</chem>	INACTIVO	50000	EC50	4.3	4
CA30	<chem>Brc1cc(Br)cc(CNC(=O)c2ncccc2)c1OC(=O)c1cccc1Cl</chem>	INACTIVO	79300	IC50	4.1	4
CA31	<chem>Brc1cc(Br)cc(CC(=O)Nc2ccc(F)c2F)c1OC(=O)c1cccc1Cl</chem>	AGONISTA	45.6	EC50	7.34	4
CA32	<chem>Brc1cc(Br)cc(CNC(=O)c2ccc(F)c2F)c1OC(=O)c1cccc1Cl</chem>	INACTIVO	64500	IC50	4.19	4
CA34	<chem>Brc1cc(Br)cc(CNC(=O)c2ncncc2)c1OC(=O)c1cccc1Cl</chem>	INACTIVO	1630	EC50	5.79	4
CA35	<chem>Brc1cc(Br)cc(CNC(=O)c2cccnc2)c1OC(=O)c1cccc1Cl</chem>	AGONISTA	488	EC50	6.31	4
CA36	<chem>Brc1cc(Br)cc(CNC(=O)c2ncccc2)c1OC(=O)CCCN1C(=O)C=CC1=O</chem>	INACTIVO	76800	IC50	4.11	4

CA37	<chem>Brc1cc(Br)cc(CC(=O)N[C@@H]2CCOC2=O)c1OC(=O)CCCN1C(=O)C=CC1=O</chem>	INACTIVO	100000	IC50	4	4
CA38	<chem>Brc1cc(Br)cc(CC(=O)N[C@@H]2CCOC2=O)c1OC(=O)CCN1C(=O)C=CC1=O</chem>	INACTIVO	23800	IC50	4.62	4
CA4	<chem>Brc1cc(Br)cc(CC(=O)Nc2cccc2[N+](=O)[O-])c1OC(=O)c1cccc1Br</chem>	AGONISTA	271	EC50	6.57	4
CA5	<chem>Brc1cc(Br)cc(CC(=O)Nc2cccc2[N+](=O)[O-])c1OC(=O)c1cc(Br)ccc1</chem>	INACTIVO	50000	EC50	4.3	4
CA6	<chem>Brc1cc(Br)cc(CC(=O)Nc2cccc2[N+](=O)[O-])c1OC(=O)c1ccc(Br)cc1</chem>	INACTIVO	50000	EC50	4.3	4
CA7	<chem>Brc1cc(Br)cc(CC(=O)Nc2cccc2[N+](=O)[O-])c1OC(=O)c1cccc1F</chem>	INACTIVO	1290	EC50	5.89	4
CA8	<chem>Brc1cc(Br)cc(CC(=O)Nc2cccc2[N+](=O)[O-])c1OC(=O)c1ccc(F)cc1</chem>	INACTIVO	50000	EC50	4.3	4
CA9	<chem>lc1cccc1C(Oc1c(cc(Br)c1Br)CC(=O)Nc1cccc1[N+](=O)[O-])=O</chem>	INACTIVO	1580	EC50	5.8	4
CH165	<chem>O1CC[C@H](NC(=O)CCC(CCC)C1=O</chem>	INACTIVO	51500	IC50	4.29	5
CH167	<chem>O1CC[C@H](NC(=O)C[C@H]2CCC=C2)C1=O</chem>	INACTIVO	33100	IC50	4.48	5
CH174	<chem>O=C1CCCC[C@H]1NC(=O)CC(=O)CCCCCCCC</chem>	ANTAGONISTA	10000	IC50	5	5
CH175	<chem>O=C([O-])Cc1nn(nn1)CCCCCCCC</chem>	ANTAGONISTA	30	IC50	7.52	5
CH176	<chem>S1CC[C@H](NC(=O)CCC(CCCCCC)C1=O</chem>	ANTAGONISTA	40	IC50	7.4	5
CH177	<chem>Brc1ccc(cc1)CCC(=O)N[C@H]1CCOC1=O</chem>	ANTAGONISTA	340	IC50	6.47	5
CH184	<chem>BrCC(=O)NCCCCCCCCC(=O)CC(=O)N[C@H]1CCOC1=O</chem>	INACTIVO	100000	IC50	4	5
CH189	<chem>O1CC[C@H](NC(=O)CCC(CCCC)C1=O</chem>	ANTAGONISTA	1750	IC50	5.76	5

CH190	<chem>O1CC[C@H](NC(=O)CCCCCCCC)C1=O</chem>	ANTAGONISTA	250	IC50	6.6	5
CH191	<chem>O1CC[C@H](NC(=O)CC(=O)CCCC)C1=O</chem>	ANTAGONISTA	110	IC50	6.96	5
CH193	<chem>lc1ccc(cc1)CC(=O)N[C@H]1CCOC1=O</chem>	ANTAGONISTA	1720	IC50	5.76	5
CH202	<chem>O1CC[C@H](NC(=O)CC(=O)CCCCCCC)C1=O</chem>	AGONISTA	15	EC50	7.82	5
CH203	<chem>O1CC[C@H](NC(=O)C[C@H](O)CCCCCCCC)C1=O</chem>	AGONISTA	54	EC50	7.27	5
CH208	<chem>O1CC[C@H](NC(=O)CC(=O)CCCCCCCCCCCC)C1=O</chem>	AGONISTA	10	EC50	8	5
CH209	<chem>O1CC[C@H](NC(=O)CC(=O)Cc2ccccc2)C1=O</chem>	AGONISTA	540	EC50	6.27	5
CH210	<chem>Brc1cc(Br)cc(CC(=O)Nc2ccccc2[N+](=O)[O-])c1OC(=O)CCN1C(=O)CC1=O</chem>	AGONISTA	3.6	EC50	8.44	5
CH211	<chem>Brc1cc(Br)cc(CNC(=O)c2ccccc2[N+](=O)[O-])c1OC(=O)CCCN1C(=O)CCC1=O</chem>	AGONISTA	119	EC50	6.92	5
CH213	<chem>Brc1cc(Br)cc(CNC(=O)c2ccccc2[N+](=O)[O-])c1OC(=O)c1ccccc1Cl</chem>	AGONISTA	28	EC50	7.55	5
GE12	<chem>O1CC[C@H](NC(=O)CCCc2c3c([nH]c2)cccc3)C1=O</chem>	INACTIVO	14800	IC50	4.83	6
GE7	<chem>Oc1ccccc1NC(=O)CC(=O)CCCC</chem>	INACTIVO	37300	IC50	4.43	6
GE2	<chem>O=C(CCCCCCCC)CC(=O)Nc1ccccc1</chem>	ANTAGONISTA	4800	IC50	5.32	6
GE3	<chem>O1CC[C@H](NC(=O)Cc2cc([N+](=O)[O-])ccc2)C1=O</chem>	ANTAGONISTA	510	IC50	6.29	6
HA1	<chem>S(Cc1nnn(c1)-c1ccc(cc1)C(=O)N[C@H]1CCOC1=O)c1ccccc1</chem>	ANTAGONISTA	2640	IC50	5.58	7
LO13	<chem>FC(F)(F)Oc1ccc(OCC(=O)N[C@H]2CCOC2=O)cc1</chem>	ANTAGONISTA	4670	IC50	5.33	8
LO2	<chem>O1CC[C@H](NC(=O)CCCCCCCC)C1=O</chem>	AGONISTA	33	EC50	7.48	8

LO6	<chem>S1CC[C@@H](NC(=O)CC(=O)CCCCCCCC)C1=O</chem>	AGONISTA	7	EC50	8.15	8
LO8	<chem>O1CC[C@H](NC(=O)CC(=O)CC\C\CCCC)C1=O</chem>	AGONISTA	10	EC50	8	8
MA3	<chem>C1CC(=O)NCCCCCCCC(=O)CC(=O)N[C@H]1CCOC1=O</chem>	ANTAGONISTA	1100	IC50	5.96	9
MA5	<chem>O=C([O-])Cc1nn(n1)CCCCCCC</chem>	INACTIVO	30000	IC50	4.52	9
MC2-1	<chem>S1CC[C@H](NC(=O)CC(=O)CCC)C1=O</chem>	INACTIVO	13000	IC50	4.89	10
MC2-10	<chem>BrC1ccc(cc1)CC(=O)N[C@H]1CCSC1=O</chem>	AGONISTA	400	EC50	6.4	10
MC2-11	<chem>S1CC[C@H](NC(=O)C2ccc(cc2)-c2cccc2)C1=O</chem>	INACTIVO	2900	EC50	5.54	10
MC2-12	<chem>S1CC[C@H](NC(=O)CC2[nH]c3c(c2)cccc3)C1=O</chem>	ANTAGONISTA	1800	IC50	5.74	10
MC2-3	<chem>S1CC[C@H](NC(=O)CCC)C1=O</chem>	INACTIVO	1100	EC50	5.96	10
MC2-4	<chem>S1CC[C@H](NC(=O)CCC)C1=O</chem>	AGONISTA	790	EC50	6.1	10
MC2-5	<chem>S1CC[C@H](NC(=O)CCC)C1=O</chem>	AGONISTA	140	EC50	6.85	10
MC2-7	<chem>S1CC[C@H](NC(=O)C2cccc2)C1=O</chem>	INACTIVO	2500	EC50	5.6	10
MC2-8	<chem>S1CC[C@H](NC(=O)C2cc([N+](=O)[O-])ccc2)C1=O</chem>	ANTAGONISTA	4100	IC50	5.39	10
MC2-9	<chem>S1CC[C@H](NC(=O)C2ccc(OC)cc2)C1=O</chem>	INACTIVO	7200	EC50	5.14	10
ME1	<chem>BrC1ccc(cc1)CC(=O)N[C@H]1CCOC1=O</chem>	ANTAGONISTA	4800	IC50	5.32	11
ME3	<chem>C1CC(=O)NCCCCCCCC(=O)CC(=O)N[C@H]1CCOC1=O</chem>	ANTAGONISTA	3100	IC50	5.51	11
ME4	<chem>BrCC(=O)NCCCCCCCC(=O)CC(=O)N[C@H]1CCOC1=O</chem>	INACTIVO	26800	IC50	4.57	11
ME5	<chem>S=C=NCCCCCCCC(=O)CC(=O)N[C@H]1CCOC1=O</chem>	INACTIVO	39100	IC50	4.41	11

ME7	<chem>S=C=NCCCCCCCCC(=O)CC(=O)N[C@H]1CCOC1=O</chem>	ANTAGONISTA	3900	IC50	5.41	11
MO10	<chem>S=C=NCCCCCCCCC(=O)CC(=O)N[C@H]1CCOC1=O</chem>	AGONISTA	2600	EC50	5.59	12
MO12	<chem>O(C)c1cc(NC(=O)CC(=O)CCCCCCCC)ccc1</chem>	INACTIVO	200000	IC50	3.7	12
MO14	<chem>Brc1cc(OCCCC(=O)N[C@H]2CCSC2=O)ccc1</chem>	INACTIVO	4200	EC50	5.38	12
MO15	<chem>O=C(NC1CCCC1)CCCCC</chem>	INACTIVO	50000	IC50	4.3	12
MO16	<chem>O=C(CC(=O)NCCCCC)CCc1cccc1</chem>	ANTAGONISTA	5200	IC50	5.28	12
MO18	<chem>Clc1cc(ccc1)C(=O)Nc1ccc1NC(=O)c1cc(Cl)ccc1</chem>	INACTIVO	69000	IC50	4.16	12
MO2	<chem>O1CC[C@H](NC(=O)CC(=O)CCC)C1=O</chem>	INACTIVO	40000	IC50	4.4	12
MO21	<chem>Oc1cccc1NC(=O)CC(=O)CCCCCCCC</chem>	INACTIVO	17000	EC50	4.77	12
MO5	<chem>O1CC[C@H](NC(=O)CCCc2cccc2)C1=O</chem>	INACTIVO	175000	IC50	3.76	12
MO8	<chem>FC(F)(F)c1ccc(cc1)CCC(=O)N[C@H]1CCOC1=O</chem>	ANTAGONISTA	3000	IC50	5.52	12
MO9	<chem>Clc1ccc(OCCCC(=O)N[C@H]2CCOC2=O)cc1</chem>	INACTIVO	21000	IC50	4.68	12
MU3	<chem>Brc1cc(Br)cc(CNC(=O)c2cccc2Cl)c1OC(=O)c1ccc(F)cc1</chem>	ANTAGONISTA	240	IC50	6.62	13
MU4	<chem>Brc1cc(Br)cc(CNC(=O)c2cccc2Cl)c1OC(=O)c1ccc1C</chem>	ANTAGONISTA	54	IC50	7.27	13
MU5	<chem>Clc1cc(Cl)ccc1C(Oc1c(cc(cc1C)C)C)CNC(=O)c1cccc1Cl)=O</chem>	ANTAGONISTA	920	IC50	6.04	13
OB10	<chem>Brc1cc(Br)cc(CC(=O)Nc2cccc2[N+](=O)[O-])c1OC(=O)CNC(=O)CCl</chem>	INACTIVO	99500	IC50	4	14
OB11	<chem>Brc1cc(Br)cc(CC(=O)Nc2cccc2[N+](=O)[O-])c1OC(=O)N1C(=O)C=C1=O</chem>	INACTIVO	86800	IC50	4.06	14
OB12	<chem>Brc1cc(Br)cc(CC(=O)Nc2cccc2[N+](=O)[O-])c1OC(=O)CC</chem>	ANTAGONISTA	3600	IC50	5.44	14

OB13	<chem>Brc1cc(Br)cc(CC(=O)Nc2cccc2[N+](=O)[O-])c1OC(=O)CCN1C(=O)C=CC1=O</chem>	INACTIVO	78100	IC50	4.11	14
OB14	<chem>Brc1cc(Br)cc(CNC(=O)c2cccc2[N+](=O)[O-])c1OC(=O)CN1C(=O)C=CC1=O</chem>	INACTIVO	51900	IC50	4.28	14
OB15	<chem>Brc1cc(Br)cc(CNC(=O)c2cccc2[N+](=O)[O-])c1OC(=O)CCN1C(=O)C=CC1=O</chem>	ANTAGONISTA	1500	IC50	5.82	14
OB3	<chem>O1CC[C@H](NC(=O)CC(=O)CCCCCCCCC)C1=O</chem>	INACTIVO	30000	IC50	4.52	14
OB5	<chem>[Br+](c1cc(CC(=O)Nc2ccc2[N+](=O)[O-])c(OC(=O)CCCN=C=S)c(Br)c1)[CH2-]</chem>	INACTIVO	148000	IC50	3.83	14
OB6	<chem>[Br+](c1cc(CNC(=O)c2ccc2[N+](=O)[O-])c(OC(=O)CCCN=C=S)c(Br)c1)[CH2-]</chem>	INACTIVO	200000	IC50	3.7	14
OB7	<chem>Brc1cc(Br)cc(CC(=O)Nc2cccc2[N+](=O)[O-])c1OCCCN=C=S</chem>	INACTIVO	200000	IC50	3.7	14
OB8	<chem>Brc1cc(Br)cc(CC(=O)Nc2cccc2[N+](=O)[O-])c1OCCCN=C=S</chem>	INACTIVO	145000	IC50	3.84	14
OB9	<chem>Brc1cc(Br)cc(CC(=O)Nc2cccc2[N+](=O)[O-])c1OC(=O)[N-]C(=O)CCl</chem>	INACTIVO	73800	IC50	4.13	14
OR1	<chem>O1CC[C@H](NC(=O)CC(=O)CCCCCCCC)C1=O</chem>	AGONISTA	1.78	EC50	8.75	15
OR10	<chem>Brc1cc(Br)cc(CNC(=O)c2cccc2[N+](=O)[O-])c1OC(=O)c1ccc(Br)cc1</chem>	AGONISTA	1.62	EC50	8.79	15
OR11	<chem>Brc1cc(Br)cc(CNC(=O)c2cccc2[N+](=O)[O-])c1OC(=O)c1ccc(OC)cc1</chem>	AGONISTA	9.2	EC50	8.04	15
OR12	<chem>Brc1cc(Br)cc(CNC(=O)c2cccc2[N+](=O)[O-])c1OC(=O)CCCCCCC</chem>	AGONISTA	12.2	EC50	7.91	15
OR13	<chem>Brc1cc(Br)cc(CNC(=O)c2cccc2[N+](=O)[O-])c1OC(=O)C</chem>	AGONISTA	20.5	EC50	7.69	15

OR14	<chem>Brc1cc(Br)cc(CNC(=O)c2cccc2[N+](=O)[O-])c1OCOC</chem>	AGONISTA	37.6	EC50	7.42	15
OR15	<chem>Brc1cc(Br)cc(CNC(=O)c2cccc2[N+](=O)[O-])c1OC</chem>	AGONISTA	60	EC50	7.22	15
OR16	<chem>Brc1cc(Br)cc(CNC(=O)c2cccc2[N+](=O)[O-])c1O</chem>	AGONISTA	900	EC50	6.05	15
OR17	<chem>Brc1cc(Br)cc(CNC(=O)c2cccc2Cl)c1OC(=O)c1ccc1[N+](=O)[O-]</chem>	AGONISTA	5.12	EC50	8.29	15
OR18	<chem>Brc1cc(Br)cc(CNC(=O)c2cccc2Cl)c1OC(=O)c1ccc1Cl</chem>	AGONISTA	7.99	EC50	8.1	15
OR2-1	<chem>Fc1cc(ccc1)CNC(=O)c1cccc1[N+](=O)[O-]</chem>	INACTIVO	21000	IC50	4.68	16
OR2-10	<chem>FC(F)(F)c1ccc(cc1)CNC(=O)c1cccc1[N+](=O)[O-]</chem>	INACTIVO	35000	IC50	4.46	16
OR2-11	<chem>O=C(NC[C@H](C)c1ccc1)c1cccc1[N+](=O)[O-]</chem>	INACTIVO	37000	IC50	4.43	16
OR2-12	<chem>O=C(NCCc1cccc1)c1cccc1[N+](=O)[O-]</chem>	INACTIVO	21000	IC50	4.68	16
OR2-13	<chem>O=C(NCCCCC)c1cccc1[N+](=O)[O-]</chem>	INACTIVO	40000	IC50	4.4	16
OR2-14	<chem>O=C(NCCC=1CCCC=1)c1cccc1[N+](=O)[O-]</chem>	INACTIVO	52000	IC50	4.28	16
OR2-15	<chem>O=C(NCc1cc([N+](=O)[O-])ccc1)c1cccc1[N+](=O)[O-]</chem>	INACTIVO	25000	EC50	4.6	16
OR2-16	<chem>O=C(NCCCCCCCCC)c1cccc1[N+](=O)[O-]</chem>	AGONISTA	358	EC50	6.45	16
OR2-2	<chem>Fc1cc(cc(F)c1)CNC(=O)c1cccc1[N+](=O)[O-]</chem>	ANTAGONISTA	6600	IC50	5.18	16
OR2-3	<chem>Fc1cc(ccc1F)CNC(=O)c1cccc1[N+](=O)[O-]</chem>	INACTIVO	16000	IC50	4.8	16
OR2-4	<chem>Fc1cc(F)ccc1CNC(=O)c1cccc1[N+](=O)[O-]</chem>	INACTIVO	45000	IC50	4.35	16
OR2-5	<chem>Clc1cc(ccc1)CNC(=O)c1cccc1[N+](=O)[O-]</chem>	ANTAGONISTA	9000	IC50	5.05	16
OR2-6	<chem>Brc1cc(ccc1)CNC(=O)c1cccc1[N+](=O)[O-]</chem>	ANTAGONISTA	9600	IC50	5.02	16

OR2-7	<chem>lc1cc(ccc1)CNC(=O)c1ccc1[N+](=O)[O-]</chem>	ANTAGONISTA	6500	IC50	5.19	16
OR2-8	<chem>Brc1ccc(cc1)CCNC(=O)c1cccc1[N+](=O)[O-]</chem>	ANTAGONISTA	4800	IC50	5.32	16
OR2-9	<chem>O=C(NCc1cccc1[N+](=O)[O-])c1cccc1[N+](=O)[O-]</chem>	ANTAGONISTA	9700	IC50	5.01	16
OR5	<chem>Brc1cc(Br)cc(CNC(=O)c2cccc2[N+](=O)[O-])c1OC(=O)c1cccc1</chem>	AGONISTA	1.16	EC50	8.94	16
OR6	<chem>Brc1cc(Br)cc(CNC(=O)c2cccc2[N+](=O)[O-])c1OC(=O)c1cccc1[N+](=O)[O-]</chem>	AGONISTA	0.67	EC50	9.17	16
OR7	<chem>Brc1cc(Br)cc(CNC(=O)c2cccc2[N+](=O)[O-])c1OC(=O)c1cccc1OC</chem>	AGONISTA	1.1	EC50	8.96	16
OR8	<chem>Brc1cc(Br)cc(CNC(=O)c2cccc2[N+](=O)[O-])c1OC(=O)c1cccc1C#N</chem>	AGONISTA	5.64	EC50	8.25	16
OR9	<chem>Brc1cc(Br)cc(CNC(=O)c2cccc2[N+](=O)[O-])c1OC(=O)c1ccc(Cl)cc1</chem>	AGONISTA	2.09	EC50	8.68	16
PE1	<chem>BrC/1=CC(O\C1=C/Br)=O</chem>	ANTAGONISTA	2000	IC50	5.7	17
PE12	<chem>S(Cc1cccc1)CC(=O)N[C@@H]1CCOC1=O</chem>	INACTIVO	100000	IC50	4	17
PE13	<chem>S(Cc1cccc1)CC(=O)N[C@H]1CCOC1=O</chem>	INACTIVO	100000	IC50	4	17
PE14	<chem>S(CC(=O)N[C@@H]1CCOC1=O)c1cc2c(cc1)cccc2</chem>	INACTIVO	150000	IC50	3.82	17
PE15	<chem>S(CCCCCC)CC(=O)NC1CCCC1</chem>	INACTIVO	150000	IC50	3.82	17
PE16	<chem>S(CCCCCC)CC(=O)NC1CCCC1</chem>	INACTIVO	75000	IC50	4.12	17
PE17	<chem>S(CCCCCC)CC(=O)N[C@@H]1CCC[C@H]1O</chem>	INACTIVO	100000	IC50	4	17
PE18	<chem>S(CCCCCC)CC(=O)N[C@@H]1CCCC1=O</chem>	INACTIVO	150000	IC50	3.82	17
PE19	<chem>S(CCCCCC)CC(=O)N[C@@H]1CCCC1=O</chem>	INACTIVO	150000	IC50	3.82	17
PE2	<chem>S(CCCCC)CC(=O)N[C@@H]1CCOC1=O</chem>	INACTIVO	300000	IC50	3.52	17

PE21	<chem>S(=O)(CCCCC)CC(=O)N[C@H]1CCOC1=O</chem>	INACTIVO	300000	IC50	3.52	17
PE22	<chem>S(=O)(CCCCCCC)CC(=O)N[C@H]1CCOC1=O</chem>	INACTIVO	50000	IC50	4.3	17
PE4	<chem>S(CCCCCC)CC(=O)N[C@@H]1CCOC1=O</chem>	ANTAGONISTA	6000	IC50	5.22	17
PE5	<chem>S(CCCCCCCC)CC(=O)N[C@@H]1CCOC1=O</chem>	INACTIVO	50000	IC50	4.3	17
PE6	<chem>S(CC(=O)N[C@@H]1CCOC1=O)c1cccc1</chem>	INACTIVO	300000	IC50	3.52	17
PE7	<chem>S(CC(=O)N[C@@H]1CCOC1=O)c1ccc(cc1)C</chem>	INACTIVO	50000	IC50	4.3	17
PE8	<chem>S(CC(=O)N[C@@H]1CCOC1=O)c1ccc(O)cc1</chem>	INACTIVO	100000	IC50	4	17
SK1	<chem>S1CC[C@@H](NC(=O)Cn2c3c(nc2)N(C)C(=O)N(C)C3=O)C1=O</chem>	INACTIVO	200000	EC50	3.7	18
SK2	<chem>O(CC(=O)NN1C(=Nc2c(c3cc2)C1=O)C)c1cccc1</chem>	INACTIVO	120000	EC50	3.92	18
SK3	<chem>O=C(Nc1ncccc1C)c1n(nc(c1)C(=O)Nc1ncccc1C)C</chem>	INACTIVO	35000	EC50	4.46	18
SK4	<chem>s1ccnc1NS(=O)(=O)c1cc(c(NS(=O)(=O)c2cccc2[N+](=O)[O-]))cc1</chem>	INACTIVO	60000	IC50	4.22	18
SK5	<chem>s1ccnc1NS(=O)(=O)c1cc(c(N)cc1</chem>	ANTAGONISTA	9000	IC50	5.05	18
ST1	<chem>O1CC[C@H](NC(=O)Cn2nnc(c2)CCCC)C1=O</chem>	INACTIVO	17600	IC50	4.75	19
ST2	<chem>O1CC[C@H](NC(=O)Cn2nnc(c2)CCCC)C1=O</chem>	ANTAGONISTA	3270	IC50	5.49	19
ST3	<chem>O1CC[C@H](NC(=O)Cn2nnc(c2)C2CCCC2)C1=O</chem>	INACTIVO	16300	IC50	4.79	19
ST4	<chem>O1CC[C@H](NC(=O)Cn2nnc(c2)C2CCCC2)C1=O</chem>	INACTIVO	12400	IC50	4.91	19
ST5	<chem>O1CC[C@H](NC(=O)Cn2nnc(c2)-c2cccc2)C1=O</chem>	INACTIVO	15900	IC50	4.8	19
ST6	<chem>Fc1cc(cc(F)c1)-c1nnn(c1)CC(=O)N[C@H]1CCOC1=O</chem>	INACTIVO	13700	IC50	4.86	19
ST7	<chem>O1CC[C@H](NC(=O)CCc2nnn(c2)-c2ccc(OC)cc2)C1=O</chem>	INACTIVO	16500	IC50	4.78	19

ST8	<chem>O1CC[C@H](NC(=O)CCc2nnn(c2)-c2ccc(cc2)C)C1=O</chem>	INACTIVO	10900	IC50	4.96	19
ST9	<chem>O1CC[C@H](NC(=O)CCc2nnn(c2)-c2cc(N)ccc2)C1=O</chem>	ANTAGONISTA	4030	IC50	5.39	19
YA1	<chem>O=C1NC=Nc2[nH]nnc12</chem>	ANTAGONISTA	720	IC50	6.14	20
YA2	<chem>Fc1cc(ccc1)C[C@@H]([NH3+])C(=O)[O-]</chem>	ANTAGONISTA	1800	IC50	5.74	20
YA3	<chem>O=C1NC(Nc2[nH]nnc12)=N</chem>	ANTAGONISTA	640	IC50	6.19	20
YA4	<chem>O=C([O-])[C@@H]([NH3+])[C@H](O)c1cccc1</chem>	ANTAGONISTA	3650	IC50	5.44	20
YA5	<chem>O=C([O-])c1c2c([nH]c1)cccc2</chem>	ANTAGONISTA	2080	IC50	5.68	20

ID: Identificador usado en esta tesis.

Actividad: Tipo de actividad biológica frente a LasR, según la clasificación realizada en esta tesis.

Referencia: Origen bibliográfico de las moléculas, a continuación, se listan las referencias:

1. Ahumado, M., Díaz, A. & Vivas-Reyes, R. Theoretical and structural analysis of the active site of the transcriptional regulators LasR and TraR, using molecular docking methodology for identifying potential analogues of acyl homoserine lactones (AHLs) with anti-quorum sensing activity. *Eur. J. Med. Chem.* 45, 608–615 (2010).
2. Boursier, M. E., Manson, D. E., Combs, J. B. & Blackwell, H. E. A comparative study of non-native N-acyl L-homoserine lactone analogs in two *Pseudomonas aeruginosa* quorum sensing receptors that share a common native ligand yet inversely regulate virulence. *Bioorg. Med. Chem.* 26, 5336–5342 (2018).
3. Brackman, G. et al. Synthesis and evaluation of the quorum sensing inhibitory effect of substituted triazolylidihydrofuranones. *Bioorg. Med. Chem.* 20, 4737–4743 (2012).
4. Capilato, J. N. et al. Development of a novel series of non-natural triaryl agonists and antagonists of the *Pseudomonas aeruginosa* LasR quorum sensing receptor. *Bioorg. Med. Chem.* 25, 153–165 (2017).
5. Mendez, D. et al. ChEMBL: Towards direct deposition of bioassay data. *Nucleic Acids Res.* 47, D930–D940 (2019).
6. Gerdt, J. P., McInnis, C. E., Schell, T. L., Rossi, F. M. & Blackwell, H. E. Mutational Analysis of the Quorum-Sensing Receptor LasR Reveals Interactions that Govern Activation and Inhibition by Nonlactone Ligands. *Chem. Biol.* 21, 1361–1369 (2014).
7. Hansen, M. R. et al. Triazole-containing N-acyl homoserine lactones targeting the quorum sensing system in *Pseudomonas aeruginosa*. *Bioorg. Med. Chem.* 23, 1638–1650 (2015).

8. Lowery, C. A., Salzameda, N. T., Sawada, D., Kaufmann, G. F. & Janda, K. D. Medicinal Chemistry as a Conduit for the Modulation of Quorum Sensing. *J. Med. Chem.* 53, 7467–7489 (2010).
9. Malladi, V. L. A. et al. Substituted lactam and cyclic azahemiacetals modulate *Pseudomonas aeruginosa* quorum sensing. *Bioorg. Med. Chem.* 19, 5500–6 (2011).
10. McInnis, C. E. & Blackwell, H. E. Thiolactone modulators of quorum sensing revealed through library design and screening. *Bioorg. Med. Chem.* 19, 4820–4828 (2011).
11. Amara, N. et al. Fine-Tuning Covalent Inhibition of Bacterial Quorum Sensing. *ChemBioChem* 17, 825–835 (2016).
12. Welsh, M. A., Eibergen, N. R., Moore, J. D. & Blackwell, H. E. Small Molecule Disruption of Quorum Sensing Cross-Regulation in *Pseudomonas aeruginosa* Causes Major and Unexpected Alterations to Virulence Phenotypes. (2015).
13. Müh, U. et al. Novel *Pseudomonas aeruginosa* quorum-sensing inhibitors identified in an ultra-high-throughput screen. *Antimicrob. Agents Chemother.* 50, 3674–9 (2006).
14. O'Brien, K. T., Noto, J. G., Nichols-O'Neill, L. & Perez, L. J. **Potent Irreversible Inhibitors of LasR** Quorum Sensing in *Pseudomonas aeruginosa*. *ACS Med. Chem. Lett.* 6, 162–167 (2015).
15. O'Reilly, M. C. et al. **Structural and Biochemical Studies of Non-native Agonists of the LasR Quorum-Sensing Receptor Reveal an L3 Loop "Out" Conformation for LasR.** *Cell Chem. Biol.* 25, 1128-1139.e3 (2018).
16. O'Reilly, M. C. & Blackwell, H. E. **Structure-Based Design and Biological Evaluation of Triphenyl Scaffold-Based Hybrid Compounds as Hydrolytically Stable Modulators of a LuxR-Type Quorum Sensing Receptor.** *ACS Infect. Dis.* 2, 32–38 (2016).
17. Pesci, E. C., Pearson, J. P., Seed, P. C. & Iglewski, B. H. Regulation of *las* and *rhl* quorum sensing in *Pseudomonas aeruginosa*. *J. Bacteriol.* 179, 3127–32 (1997).
18. Skovstrup, S. et al. Identification of LasR Ligands through a Virtual Screening Approach. *ChemMedChem* 8, 157–163 (2013).
19. Stacy, D. M. et al. Synthesis and biological evaluation of triazole-containing N-acyl homoserine lactones as quorum sensing modulators. *Org. Biomol. Chem.* 11, 938–954 (2013).
20. Yang, S. et al. Bicyclic brominated furanones: a new class of quorum sensing modulators that inhibit bacterial biofilm formation. *Bioorg. Med. Chem.* 22, 1313–7 (2014).

B. Anexo: Base de datos molecular PqsR

ID	SMILES	Actividad	Valor (nM)	Tipo de actividad	pIC50	Referencia
CH1	<chem>O=C(N)c1ccc(cc1)C(C)(C)C</chem>	AGONISTA	900	Kd	6.05	1
CH17	<chem>o1c(nnc1N)-c1cc(ccc1)C</chem>	INACTIVO	12400	Kd	4.91	1
CH21	<chem>FC(F)(F)c1cc(ccc1)-c1oc(nn1)C</chem>	INACTIVO	12900	Kd	4.89	1
CH25	<chem>FC(F)(F)c1cc(ccc1)-c1ocnn1</chem>	INACTIVO	6500	Kd	5.19	1
CH29	<chem>s1cccc1-c1[nH]nc(c1)C(F)(F)F</chem>	INACTIVO	8500	Kd	5.07	1
CH3	<chem>o1c(nnc1N)-c1cc(ccc1)C(C)(C)C</chem>	INACTIVO	3100	Kd	5.51	1
CH35	<chem>O=C(N)c1cc2c(cc1)cccc2</chem>	AGONISTA	800	Kd	6.1	1
CH41	<chem>Brc1cc(ccc1)C(=O)N</chem>	INACTIVO	6800	Kd	5.17	1
CH47	<chem>Brc1nc(N)ccc1</chem>	INACTIVO	6800	Kd	5.17	1
CH50	<chem>Clc1ccc(cc1)-c1oc(nn1)N</chem>	INACTIVO	200000	Kd	3.7	1
CH54	<chem>Clc1ccc(cc1)-c1sc(nn1)N</chem>	INACTIVO	3700	Kd	5.43	1
CH63	<chem>o1c(nnc1N)-c1cc(ccc1)C#N</chem>	INACTIVO	17800	EC50	4.75	1
CH72	<chem>FC(F)(F)c1cc(ccc1)-c1oc(nn1)N</chem>	AGONISTA	1300	EC50	5.89	1
CH79	<chem>Clc1cc(ccc1)-c1oc(nn1)N</chem>	INACTIVO	26700	EC50	4.57	1
IL10	<chem>Fc1cc2N=C(N([O-])C(=O)c2cc1)CCCCCCCC</chem>	INACTIVO	21000	EC50	4.68	2

96 Contribución a la identificación de potenciales antagonistas asociados al quorum sensing en *Pseudomonas aeruginosa* a través de métodos quimioinformáticos.

IL11	<chem>C1c1cc2N=C(N(OC)C(=O)c2cc1)CCCCCCCC</chem>	INACTIVO	37300	EC50	4.43	2
IL12	<chem>Fc1cc2N=C(N(OC)C(=O)c2cc1)CCCCCCCC</chem>	AGONISTA	2200	EC50	5.66	2
IL13	<chem>Fc1cc2c(N=C(N(OC)C2=O)CCCCCCCC)cc1F</chem>	INACTIVO	53600	EC50	4.27	2
IL14	<chem>O=C1N(N)C(=Nc2c1cccc2)CCCCC</chem>	INACTIVO	54000	IC50	4.27	2
IL15	<chem>O=C1N(N)C(=Nc2c1cccc2)CCCCCCCC</chem>	INACTIVO	77200	IC50	4.11	2
IL16	<chem>O=C1N(N)C(=Nc2c1cccc2)CCCCCCCCC</chem>	INACTIVO	210000	IC50	3.68	2
IL17	<chem>C1c1cc2N=C(N(N)C(=O)c2cc1)CCCCCCCC</chem>	ANTAGONISTA	5000	IC50	5.3	2
IL18	<chem>Fc1cc2N=C(N(N)C(=O)c2cc1)CCCCCCCC</chem>	ANTAGONISTA	3900	IC50	5.41	2
IL19	<chem>Fc1cc2c(N=C(N(N)C2=O)CCCCCCCC)cc1F</chem>	ANTAGONISTA	1200	IC50	5.92	2
IL20	<chem>O=C1N(CC[NH3+])C(=Nc2c1cccc2)CCCCCCCC</chem>	INACTIVO	169000	IC50	3.77	2
IL21	<chem>C1c1cc2c(N=C(N(CC[NH3+])C2=O)CCCCCCCC)cc1</chem>	ANTAGONISTA	9300	IC50	5.03	2
IL22	<chem>C1c1cc2N=C(N(CC[NH3+])C(=O)c2cc1)CCCCCCCC</chem>	INACTIVO	19300	IC50	4.71	2
IL23	<chem>C1c1cc2N=C(N(CCC[NH3+])C(=O)c2cc1)CCCCCCC</chem>	INACTIVO	15500	IC50	4.81	2
IL24	<chem>C1c1cc2N=C(N(CC[NH3+])C(=O)c2cc1)CCc1cccc1</chem>	INACTIVO	79200	IC50	4.1	2

IL25	<chem>C1c1cc2N=C(N(N)C(=O)c2cc1)CCCC1cccc1</chem>	INACTIVO	39600	IC50	4.4	2
IL3	<chem>O=C1C=C(Nc2c1cccc2)CCCCCCC</chem>	AGONISTA	400	EC50	6.4	2
IL4	<chem>OC=1C(=O)c2c(NC=1CCCCCCCC)cccc2</chem>	AGONISTA	1100	EC50	5.96	2
IL5	<chem>C1c1cc2NC(CCCCCC)=C(O)C(=O)c2cc1</chem>	AGONISTA	14	EC50	7.85	2
IL6	<chem>O=C1c2c(NC(CCCCCC)=C1N)cccc2</chem>	AGONISTA	400	EC50	6.4	2
IL7	<chem>O=C1N([O-])C(=Nc2c1cccc2)CCCCC</chem>	INACTIVO	54000	EC50	4.27	2
IL8	<chem>O=C1N([O-])C(=Nc2c1cccc2)CCCCC</chem>	INACTIVO	30000	EC50	4.52	2
IL9	<chem>C1c1cc2N=C(N([O-])C(=O)c2cc1)CCCCCCC</chem>	INACTIVO	12700	EC50	4.9	2
LU1	<chem>FC(F)(F)c1cc2c(NC(=CC2=O)CCCCC)cc1</chem>	ANTAGONISTA	54	IC50	7.27	3
LU10	<chem>O=C1C=C(Nc2c1cc([N+](=O)[O-])cc2)CCCCOCC</chem>	ANTAGONISTA	3511	IC50	5.45	3
LU11	<chem>O=C1c2cc([N+](=O)[O-])ccc2NC(CCCCCC)=C1C</chem>	ANTAGONISTA	72	IC50	7.14	3
LU12	<chem>O=C1c2cc([N+](=O)[O-])ccc2NC(CCCCCC)=C1C(=O)N</chem>	ANTAGONISTA	35	IC50	7.46	3
LU2	<chem>O=C1C=C(Nc2c1cc([N+](=O)[O-])cc2)CCCCC</chem>	ANTAGONISTA	51	IC50	7.29	3
LU3-1	<chem>OC=1C(=O)c2c(NC=1CCCCC)cccc2</chem>	AGONISTA	6.3	EC50	8.2	3
LU3-3	<chem>OC=1C(=O)c2cc([N+](=O)[O-])ccc2NC=1CCCCC</chem>	AGONISTA	2.8	EC50	8.55	3
LU5	<chem>FC(F)(F)c1cc2c(N(C)C(=C2=O)CCCCC)cc1</chem>	ANTAGONISTA	270	IC50	6.57	3

LU6	<chem>FC(F)(F)c1cc2c(N(C)C(=CC2=O)CCCCC)cc1</chem>	ANTAGONISTA	229	IC50	6.64	3
LU7	<chem>FC(F)(F)c1cc2c(NC(=CC2=O)COCCCC)cc1</chem>	ANTAGONISTA	172	IC50	6.76	3
LU8	<chem>O=C1C=C(Nc2c1cc([N+](=O)[O-])cc2)COCCCC</chem>	ANTAGONISTA	253	IC50	6.6	3
LU9	<chem>FC(F)(F)c1cc2c(NC(=CC2=O)CCCCOCC)cc1</chem>	ANTAGONISTA	1712	IC50	5.77	3
SO5	<chem>O=C(N[O-])c1ccc(cc1)C(C)(C)C</chem>	INACTIVO	23600	IC50	4.63	4
SO6	<chem>S(=O)(=O)(C)c1nc(-n2nnnc2)ccn1</chem>	INACTIVO	15000	IC50	4.82	4
SO7	<chem>S(=O)(=O)(CCCCC)c1nc(-n2nncc2)ccn1</chem>	INACTIVO	50000	IC50	4.3	4
ZE12	<chem>Brc1cc(ccc1)-c1oc(nn1)N</chem>	ANTAGONISTA	7500	IC50	5.12	5
ZE15	<chem>FC(F)(F)Oc1cc(ccc1)-c1oc(nn1)N</chem>	INACTIVO	46500	IC50	4.33	5

ID: Identificador usado en esta tesis.

Actividad: Tipo de actividad biológica frente a PqsR, según la clasificación realizada en esta tesis.

Referencia: Origen bibliográfico de las moléculas, a continuación, se listan las referencias:

1. Mendez, D. et al. ChEMBL: Towards direct deposition of bioassay data. *Nucleic Acids Res.* 47, D930–D940 (2019).
2. Ilangovan, A. et al. Structural Basis for Native Agonist and Synthetic Inhibitor Recognition by the *Pseudomonas aeruginosa* Quorum Sensing Regulator PqsR (MvfR). *PLoS Pathog.* (2013).
3. Lu, C. et al. Discovery of Antagonists of PqsR, a Key Player in 2-Alkyl-4-quinolone-Dependent Quorum Sensing in *Pseudomonas aeruginosa*. *Chem. Biol.* 19, 381–390 (2012).
4. Soheili, V., Tajani, A. S., Ghodsi, R. & Bazzaz, B. S. F. Anti-PqsR compounds as next-generation antibacterial agents against *Pseudomonas aeruginosa*: A review. *Eur. J. Med. Chem.* 172, 26–35 (2019).

5. Zender, M. et al. Discovery and Biophysical Characterization of 2-Amino-oxadiazoles as Novel Antagonists of PqsR, an Important Regulator of *Pseudomonas aeruginosa* Virulence. *J. Med. Chem.* 56, 6761–6774 (2013).

C. Anexo: Base de datos molecular RhIR

ID	SMILES	Actividad	Valor (nM)	Tipo de actividad	pIC50	Referencia
BO1	<chem>O1CC[C@H](NC(=O)CCC)C1=O</chem>	INACTIVO	8080	EC50	5.09	1
BO10	<chem>O1CC[C@H](NC(=O)CC2CC2)C1=O</chem>	AGONISTA	2760	EC50	5.56	1
BO11	<chem>S1CC[C@H](NC(=O)CCC)C1=O</chem>	AGONISTA	3820	EC50	5.42	1
BO12	<chem>O=C1CCC[C@@H]1NC(=O)CCC</chem>	INACTIVO	14300	EC50	4.84	1
BO2	<chem>O1CC[C@H](NC(=O)CCC=C)C1=O</chem>	INACTIVO	7930	EC50	5.1	1
BO2-1	<chem>O=C1CCC[C@@H]1NC(=O)C1CCC1</chem>	INACTIVO	5940	EC50	5.23	1
BO2-10	<chem>Ic1ccc(OCC(=O)NC[C@H]2OCCC2)cc1</chem>	INACTIVO	100000	IC50	4	1
BO2-11	<chem>Ic1ccc(OCC(=O)N[C@H]2CCSC2=O)cc1</chem>	INACTIVO	31400	IC50	4.5	1
BO2-2	<chem>S1CC[C@H](NC(=O)C2CC2)C1=O</chem>	AGONISTA	1720	EC50	5.76	1
BO2-3	<chem>O=C1CCC[C@@H]1NC(=O)CC(C)C</chem>	INACTIVO	7580	EC50	5.12	1
BO2-4	<chem>S1CC[C@H](NC(=O)CC(C)C)C1=O</chem>	AGONISTA	463	EC50	6.33	1
BO2-6	<chem>O1CC[C@H](NC(=O)CC(C)C)C1=O</chem>	AGONISTA	1020	EC50	5.99	1
BO2-9	<chem>O=C(NC1CCCC1)C1CCC1</chem>	INACTIVO	26700	IC50	4.57	1
BO3	<chem>O1CC[C@H](NC(=O)C2CC2)C1=O</chem>	AGONISTA	1220	EC50	5.91	1
BO4	<chem>O1CC[C@H](NC(=O)CCC)C1=O</chem>	INACTIVO	10800	EC50	4.97	1

BO5	<chem>O1CC[C@H](NC(=O)C(C)C)C1=O</chem>	AGONISTA	4890	EC50	5.31	1
BO7	<chem>O1CC[C@H](NC(=O)[C@H](CC)C)C1=O</chem>	INACTIVO	7770	EC50	5.11	1
BO8	<chem>O1CC[C@H](NC(=O)C\C=C\C)C1=O</chem>	INACTIVO	6930	EC50	5.16	1
BO9	<chem>O1CC[C@H](NC(=O)C2CC2)C1=O</chem>	AGONISTA	1410	EC50	5.85	1
EI12	<chem>O1CC[C@H](NC(=O)Cc2cc(ccc2)C)C1=O</chem>	AGONISTA	2000	EC50	5.7	2
EI13	<chem>O1CC[C@H](NC(=O)Cc2cc(OC)ccc2)C1=O</chem>	AGONISTA	4700	EC50	5.33	2
EI14	<chem>O1CC[C@H](NC(=O)Cc2cc(ccc2)C#N)C1=O</chem>	AGONISTA	1700	EC50	5.77	2
EI15	<chem>S(C)c1cc(ccc1)CC(=O)N[C@H]1CCOC1=O</chem>	INACTIVO	6600	EC50	5.18	2
EI16	<chem>Clc1ccc(cc1)CCC(=O)N[C@H]1CCOC1=O</chem>	INACTIVO	6600	EC50	5.18	2
EI17	<chem>Brc1cc(ccc1)CCC(=O)N[C@H]1CCOC1=O</chem>	INACTIVO	11100	EC50	4.95	2
EI18	<chem>O1CC[C@H](NC(=O)CCc2ccc([N+](=O)[O-])cc2)C1=O</chem>	INACTIVO	27100	EC50	4.57	2
EI2	<chem>O1CC[C@H](NC(=O)CC2CCC2)C1=O</chem>	AGONISTA	3100	EC50	5.51	2
EI21	<chem>Ic1ccc(cc1)CC(=O)N[C@H]1CCOC1=O</chem>	ANTAGONISTA	8100	IC50	5.09	2
EI22	<chem>O1CC[C@H](NC(=O)Cc2ccc([N+](=O)[O-])cc2)C1=O</chem>	INACTIVO	17900	IC50	4.75	2

102 Contribución a la identificación de potenciales antagonistas asociados al quorum sensing en *Pseudomonas aeruginosa* a través de métodos quimioinformáticos.

EI24	<chem>O1CC[C@H](NC(=O)Cc2c cc(cc2)C)C1=O</chem>	INACTIVO	20000	IC50	4.7	2
EI25	<chem>FC(F)(F)c1ccc(cc1)CC(=O) N[C@H]1CCOC1=O</chem>	INACTIVO	24400	IC50	4.61	2
EI29	<chem>Clc1cc(ccc1Cl)CC(=O)N[C@ H]1CCOC1=O</chem>	ANTAGONISTA	3400	IC50	5.47	2
EI3	<chem>O1CC[C@H](NC(=O)Cc2c cccc2)C1=O</chem>	INACTIVO	14700	EC50	4.83	2
EI30	<chem>O1CC[C@H](NC(=O)Cc2c c3c(cc2)cccc3)C1=O</chem>	INACTIVO	20700	IC50	4.68	2
EI31	<chem>O1CC[C@H](NC(=O)COc2 ccc(cc2)C)C1=O</chem>	ANTAGONISTA	10700	IC50	4.97	2
EI32	<chem>O1CC[C@H](NC(=O)COc2 ccc(OC)cc2)C1=O</chem>	ANTAGONISTA	12000	IC50	4.92	2
EI33	<chem>Brc1ccc(OCC(=O)N[C@H] 2CCOC2=O)cc1</chem>	ANTAGONISTA	5900	IC50	5.23	2
EI34	<chem>lc1ccc(OCC(=O)N[C@H]2 CCOC2=O)cc1</chem>	INACTIVO	17300	IC50	4.76	2
EI36	<chem>S(C)c1ccc(cc1)CCC(=O)N[C@H]1CCOC1=O</chem>	INACTIVO	21800	IC50	4.66	2
EI4	<chem>Fc1ccc(cc1)CC(=O)N[C@ H]1CCOC1=O</chem>	INACTIVO	9000	EC50	5.05	2
EI7	<chem>Clc1cc(ccc1)CC(=O)N[C@ H]1CCOC1=O</chem>	INACTIVO	5500	EC50	5.26	2
EI9	<chem>lc1cc(ccc1)CC(=O)N[C@H]1CCOC1=O</chem>	INACTIVO	5800	EC50	5.24	2
OL1	<chem>Brc1cc(OCCCC(=O)N[C@ H]2CCSC2=O)ccc1</chem>	ANTAGONISTA	4000	IC50	5.4	3
OL2	<chem>[Br+](c1cc(OCCCC(=O)N[C@@H]2CCSC2=O)ccc1)[CH2-]</chem>	INACTIVO	100000	IC50	4	3
OL4	<chem>Clc1cc(OCCCC(=O)N[C@ H]2CCSC2=O)ccc1</chem>	ANTAGONISTA	9000	IC50	5.05	3

OL5	<chem>S=C=NCCCCCCCCC(=O)CC(=O)N[C@H]1CCOC1=O</chem>	INACTIVO	56000	IC50	4.25	3
OL6	<chem>O=C(CC(=O)NCCCCCCCC)C1CCCCC1</chem>	INACTIVO	18000	IC50	4.74	3

ID: Identificador usado en esta tesis.

Actividad: Tipo de actividad biológica frente a PqsR, según la clasificación realizada en esta tesis.

Referencia: Origen bibliográfico de las moléculas, a continuación, se listan las referencias:

1. Blackwell, H. E., Boursier, M. E. & Moore, J. Synthetic ligands that modulate the activity of the rhIR quorum sensing receptor. (2017).
2. Eibergen, N. R., Moore, J. D., Mattmann, M. E. & Blackwell, H. E. Potent and Selective Modulation of the RhIR Quorum Sensing Receptor by Using Non-native Ligands: An Emerging Target for Virulence Control in *Pseudomonas aeruginosa*. *ChemBioChem* 16, 2348–2356 (2015).
3. **O’Loughlin, C. T. et al. A quorum-sensing inhibitor blocks *Pseudomonas aeruginosa* virulence and biofilm formation. *Proc. Natl. Acad. Sci.* 110, 17981–17986 (2013).**

D. Anexo: Base de datos molecular QuiProNaB

ID	SMILES	Familia	Referencia
QPN0001	<chem>O(C(C(O)c1cc(OC)c(O)cc1)C)c1c(OC)cc(/C=C/C)cc1</chem>	Myristicaceae	Aislamiento de Metabolitos Secundarios de Hojas de <i>Iryanthera ulei</i> evaluación de su Actividad Antibacteriana y Síntesis de un Neolignano
QPN0002	<chem>O=C(O[C@@H]1C(OC(=O)C)C(OC(=O)C)[C@H](C)OC1OC1C(=O)c2c(O)cc(OC(=O)C)cc2OC1c1cc(O)C(=O)C)c(OC(=O)C)cc1)C</chem>	Myristicaceae	
QPN0003	<chem>O=C1C(OC2[C@H](O)C(O)C(O)[C@H](C)O2)C(c2ccc(O)cc2)Oc2c1c(O)cc(O)c2</chem>	Myristicaceae	
QPN0004	<chem>O=C1OC2(C)C(C(C)(C)CC(O)C2)=C1</chem>	Myristicaceae	
QPN0005	<chem>O=Cc1cc(OC)c(OC(C(O)c2cc(OC)c(O)cc2)C)cc1</chem>	Myristicaceae	
QPN0006	<chem>OC1(C)C2C(C(C(C)C)CC1)C=C(C)CC2</chem>	Myristicaceae	
QPN0007	<chem>C1C[C@@]1(O)[C@@H](OC(=O)CC(C)C)C=C2C(C)OC(=O)C=CO[C@@H](OC(=O)CC(C)C)C12</chem>	Valerianaceae	Aislamiento e identificación de metabolitos activos sobre el sistema nervioso central obtenidos de <i>Valeriana pavonii</i>
QPN0008	<chem>O=C(N)CC(C)C</chem>	Solanaceae	
QPN0009	<chem>O=C(OCC1=CO[C@@H](O)C2[C@@](O)(CO)[C@@H](OC(=O)C)CC12)CC(C)C</chem>	Valerianaceae	
QPN0010	<chem>O=C(OCC1=CO[C@@H](OC(=O)CC(C)C)C2[C@@](O)(COC(=O)C)[C@@H](O)CC12)CC(C)C</chem>	Valerianaceae	
QPN0011	<chem>O=C(OC[C@@]1(O)[C@@H](O)C=C2C(COC(=O)C)=CO[C@@H](OC(=O)CC(C)C)C12)CC(C)C</chem>	Valerianaceae	
QPN0012	<chem>O=C(OC[C@@]1(O)[C@@H](OC(=O)CC(C)C)C=C2C(COC(=O)C)=CO[C@@H](OC(=O)CC(C)C)C12)CC(C)C</chem>	Valerianaceae	
QPN0013	<chem>O=C(O[C@@H]1[C@](O)(COC(=O)C)C2[C@H](OC(=O)CC(C)C)OC=C(COC(=O)C)C2=C1)CC(C)C</chem>	Valerianaceae	
QPN0014	<chem>O=C(O[C@@H]1[C@](O)(COC)C2[C@H](OC(=O)CC(C)C)OC=C(COC(=O)C)C2=C1)CC(C)C</chem>	Valerianaceae	
QPN0015	<chem>O=C(O[C@@H]1[C@](O)(COCC)C2[C@H](OC(=O)CC(C)C)OC=C(COC(=O)C)C2=C1)CC(C)C</chem>	Valerianaceae	

QPN0016	<chem>O(C)c1c(C[C@@H](O)C(O)(C)C)cc2c(OC(=O)C=C2)c1</chem>	Rutaceae	Aislamiento y caracterización de alcaloides presentes en dos especies del genero <i>Zanthoxylum</i> (rutaceae) síntesis de análogos benzofenantridínicos y evaluación de actividad antifúngica y antibacteriana
QPN0017	<chem>O(C)c1c(O)c(OC)cc([C@H]2OC[C@@H]3C(c4cc(O)C)c(O)c(OC)c4)OC[C@H]23)c1</chem>	Rutaceae	
QPN0018	<chem>O(C)c1c(O)cc2C(=O)NCCc2c1</chem>	Rutaceae	
QPN0019	<chem>O(C)c1c(O)ccc2c1cnc1c3c(cc4OCOc4c3)ccc21</chem>	Rutaceae	
QPN0020	<chem>O(C)c1c(OC)cc2c(c3c(nc2)c2c(cc4OCOc4c2)cc3)c1</chem>	Rutaceae	
QPN0021	<chem>O(C)c1c(OC)ccc2c1C[N+]1(C)CCc3c(cc(O)c(OC)c3)[C-]1C2</chem>	Rutaceae	
QPN0022	<chem>O(C)c1c(OC)ccc2c1C[N+]1(C)CCc3c(cc(OC)c(OC)c3)[C-]1C2</chem>	Rutaceae	
QPN0023	<chem>O(C)c1c(OC)ccc2c1cnc1c3c(cc4OCOc4c3)ccc21</chem>	Rutaceae	
QPN0024	<chem>O(CC=C(C)C)c1ccc(/C=C/CO)cc1</chem>	Rutaceae	
QPN0025	<chem>O([C@@H]1CC=2C(C)(C3C(C4C(C)(C(C/C=C/C(C)C)CC)C)CC4)CC3)CC=2)CC1)C1[C@@H](O)[C@H](O)[C@@H](O)[C@H](CO)O1</chem>	Rutaceae	
QPN0026	<chem>O=C(C[C@H]1N(C)c2c3c(cc4OCOc4c3)ccc2-c2c1c(OC)c(OC)cc2)C</chem>	Rutaceae	
QPN0027	<chem>O=C(O)C[C@H]1N(C)c2c3c(cc4OCOc4c3)ccc2-c2c1c(OC)c(OC)cc2</chem>	Rutaceae	
QPN0028	<chem>O=C(O)[C@@H]1N(Cc2c(O)c(C(=O)C)c(OC)cc2OC)CCC1</chem>	Rutaceae	
QPN0029	<chem>O=C(O)c1cc(OC)c(O)cc1</chem>	Rutaceae	
QPN0030	<chem>O=C1Oc2c(cc3c(O[C@H](C(O)(C)C)C3)c2)C=C1</chem>	Rutaceae	
QPN0031	<chem>O=C1Oc2c(ccc3O[C@H](C(O)(C)C)Cc23)C=C1</chem>	Rutaceae	
QPN0032	<chem>O=C1[C@@H](Cc2cc3OCOc3cc2)[C@H](Cc2cc3OCOc3cc2)CO1</chem>	Rutaceae	
QPN0033	<chem>O=CNc1c(-c2c(O)c(OC)c(O)c(OC)c2)ccc2c1cc1OCOc1c2</chem>	Rutaceae	
QPN0034	<chem>O=Cc1c(OC)cc(OC)cc1O</chem>	Rutaceae	
QPN0035	<chem>O=Cc1ccc(O)cc1</chem>	Rutaceae	

QPN0036	<chem>OC1CC=2[C@@](C)(C3C(C4[C@@](C)(C5C(C)C6N(C5C4)CC(C)C4C6OC(C)C4)CC3)CC=2)CC1</chem>	Solanaceae	Aislamiento y elucidación estructural de algunos metabolitos secundarios del extracto etanolico de la especie <i>Solanum cornifolium</i> . Sección geminata
QPN0037	<chem>O(C)c1c(O)cc2c(OC(=O)C=C2)c1</chem>	Solanaceae	Aislamiento y estructura de algunos metabolitos secundarios de <i>Solanum validinervium</i> (Solanaceae) Sección Geminata
QPN0038	<chem>O(CCO)[C@@H]1[C@@H](O)C(O)[C@H](Oc2c(O)cc3c(OC(=O)C=C3)c2)OC1CO</chem>	Solanaceae	
QPN0039	<chem>O=C1Oc2c(cc(O)c(O)c2)C=C1</chem>	Solanaceae	
QPN0040	<chem>OCCc1ccc(O)cc1</chem>	Solanaceae	
QPN0041	<chem>O=C(O)c1c(C)cc(O)c2c1oc1c2c2c(c(O)c1)C(=O)OC2</chem>	Stereocaulaceae	Antioxidantes de los líquenes <i>Stereocaulon strictum</i> (Stereocaulaceae) y <i>Lobariella pallida</i> (lobariaceae) y determinación de su potencial citotoxicidad
QPN0042	<chem>O=C(OC)c1c(O)c(C(=O)C)c(O)c(C)c1C</chem>	lobariaceae	
QPN0043	<chem>O=C(OC)c1c(O)c(C(=O)C)c(O)cc1C</chem>	Stereocaulaceae	
QPN0044	<chem>O=C(OC)c1c(O)c(C)c(O)cc1C</chem>	Stereocaulaceae	
QPN0045	<chem>O=C(OC)c1c(O)c(C)c(OC(=O)c2c(O)c(C=O)c(O)cc2C)cc1C</chem>	Stereocaulaceae	
QPN0046	<chem>O=C(OC)c1c(O)cc(O)cc1C</chem>	lobariaceae	
QPN0047	<chem>O=Cc1c(OC)cc(C)c(C)c1</chem>	Stereocaulaceae	
QPN0048	<chem>OC(C(O)CO)C(O)CO</chem>	lobariaceae	
QPN0049	<chem>O(C)c1c(O)c2-c3c(cc(O)c(OC)c3)CC3N(C)CCc(c23)c1</chem>	Lauraceae	Bioprospección de la especie <i>Ocotea caudata</i> nees (Lauraceae)
QPN0050	<chem>O(C)c1c(O)c2-c3c(cc4OCOc4c3)CC3NCCc(c23)c1</chem>	Lauraceae	
QPN0051	<chem>O(C)c1c(O)cc(C=2Oc3c(C(=O)C=2)ccc(OC)c3)cc1</chem>	Lauraceae	
QPN0052	<chem>O(C)c1c(O)cc(CCN(C)C)c2c1c1c(cc3OCOc3c1)cc2</chem>	Lauraceae	
QPN0053	<chem>O(C)c1c(O)cc(CCN(C/C=C/c2cccc2)C)c2c1c1c(cc(O)c(OC)c1)cc2</chem>	Lauraceae	

QPN0054	<chem>O(C)c1c(O)cc(CCN(CC#C)C)c2c1c1c(cc(O)c(OC)c1)cc2</chem>	Lauraceae	Búsqueda de sustancias antifungicas para el control de hongos fitopatogenos apartir de constituyentes químicos presentes en Piper Pesaresanum.
QPN0055	<chem>O(C)c1c(O)cc(CCN(CC=C(C)C)C)c2c1c1c(cc(O)c(O)C)c1)cc2</chem>	Lauraceae	
QPN0056	<chem>O(C)c1c(O)cc2C(Cc3cc(O)c(OC)cc3)N(C)CCc2c1</chem>	Lauraceae	
QPN0057	<chem>O(C)c1c(O)cc2c(c1)C13C(=CC(=O)C(OC)=C1)C(N(C)CC3)C2</chem>	Lauraceae	
QPN0058	<chem>O(C)c1c(O)cc2c3c1-c1c(cc(O)c(OC)c1)CC3N(C)CC2</chem>	Lauraceae	
QPN0059	<chem>O(C)c1c(OC)c2-c3c(cccc3)CC3NCCc(c1O)c23</chem>	Lauraceae	
QPN0060	<chem>O(C)c1c(OC)c2c3c(c1O)-c1c(cc(O)c(OC)c1)CC3NCC2</chem>	Lauraceae	
QPN0061	<chem>O(C)c1c(OC)cc(CCN(C)C)c2c1c1c(cc3OCOc3c1)cc2</chem>	Lauraceae	
QPN0062	<chem>O(CC=C(C)C)c1c(OC)cc2c3c(OC)c(O)cc(CCN(CC=C(C)C)C)c3ccc2c1</chem>	Lauraceae	
QPN0063	<chem>O=C(O)C(=C)C1C[C@H]2C(=C)CCCC2(C)CC1</chem>	Lauraceae	
QPN0064	<chem>O=C(O)C1N(C)CC(O)C1</chem>	Lauraceae	
QPN0065	<chem>O=C(O)CCCCCCCCCCCCCCC</chem>	Lauraceae	
QPN0066	<chem>O=C(O)Cc1ccc(O)cc1</chem>	Lauraceae	
QPN0067	<chem>O=C(OC)C(=C)C1CC2C(=C)CCCC2(C)CC1</chem>	Lauraceae	
QPN0068	<chem>O=C(OCC)Cc1ccc(O)cc1</chem>	Lauraceae	
QPN0069	<chem>O=C1C(OC2C(O)C(O)C(O)C(CO)O2)=C(c2cc(O)c(O)cc2)Oc2c1c(O)cc(O)c2</chem>	Lauraceae	
QPN0070	<chem>O=C1C(OC2C(O)C(O)C(O)C(CO)O2)=C(c2ccc(O)c2)Oc2c1c(O)cc(O)c2</chem>	Lauraceae	
QPN0071	<chem>O=C1c2c(O)cc(O)cc2OC(c2cc(O)c(O)cc2)=C1</chem>	Lauraceae	
QPN0072	<chem>OC(=C)CCCCCCC/C=C\C/C=C\C\CCCC</chem>	Lauraceae	
QPN0073	<chem>OCCCCCCCCCCCCCCCCCCCCCCCCCCCC</chem>	Lauraceae	
QPN0074	<chem>O(C)c1cc(O)c(CCCc2ccccc2)c(O)c1</chem>	Piperaceae	
QPN0075	<chem>O=C(O)c1cc(C(=O)/C=C(\C)/C)c(O)c(C/C=C(\C)/C)c1</chem>	Piperaceae	
QPN0076	<chem>O=C(O)c1cc(C/C=C(\C)/C)c(O)c(C/C=C(\C)/C)c1</chem>	Piperaceae	
QPN0077	<chem>O=C(O)c1cc(C/C=C(\C)/C)c(OC)c(C/C=C(\C)/C)c1</chem>	Piperaceae	
QPN0078	<chem>O=C(O)c1cc(CCC(C)C)c(OC)c(CCC(C)C)c1</chem>	Piperaceae	
QPN0079	<chem>O=C(OC)c1cc(C/C=C(\C)/C)c(OC)c(C/C=C(\C)/C)c1</chem>	Piperaceae	

QPN0080	<chem>O=C(OCC)c1cc(C/C=C(\C)/C)c(OC)c(C/C=C(\C)/C)c1</chem>	Piperaceae	
QPN0081	<chem>C(=C/CCC(C=C)(C)c1cc2OCOc2cc1)\CC/C=C(\C)/C</chem>	Piperaceae	Constituyentes químicos provenientes de <i>Piper peltatum</i> y evaluación de su potencial para el control de <i>Moniliophthora roreri</i> y <i>M. perniciosa</i>
QPN0082	<chem>C(c1cc2OCOc2cc1)C1C(Cc2cc3OCOc3cc2)COC1</chem>	Piperaceae	
QPN0083	<chem>Oc1c(O)ccc(C(C=C)(CC/C=C(/CC/C=C(\C)/C)\C)C)c1</chem>	Piperaceae	
QPN0084	<chem>Oc1c(O)ccc(C(CCCC(CCCC(C)C)C)(CC)C)c1</chem>	Piperaceae	
QPN0085	<chem>[C@H]1(c2cc3OCOc3cc2)OC[C@H]2[C@@H](c3cc4OCOc4cc3)OC[C@@H]12</chem>	Piperaceae	
QPN0086	<chem>O(C)c1c(O)cc2C(c3cc4OCOc4cc3)C(C)C(C)C(O)c2c1</chem>	Myristicaceae	Contribución al estudio fitoquímico de hojas de <i>Osteophloeum sulcatum</i>
QPN0087	<chem>OC(C(C(Cc1cc2OCOc2cc1)C)C)c1cc2OCOc2cc1</chem>	Myristicaceae	
QPN0088	<chem>OC1(c2cc3OCOc3cc2)C(C)C(C)Cc2c1c1OCOc1cc2</chem>	Myristicaceae	
QPN0089	<chem>O(C)c1c(O)ccc([C@@H]2[C@@H](C)[C@H](C)[C@H](c3cc(OC)c(O)cc3)O2)c1</chem>	Myristicaceae	Contribución al estudio fitoquímico del fruto (mesocarpo) de <i>Osteophloeum sulcatum</i>
QPN0090	<chem>O(C)C1=C(CC=C(C)C)C(=O)N(C)c2c1cccc2</chem>	Rutaceae	Determinación de terpenos cumarinas y alcaloides del extracto etanólico de la madera de la especie <i>Esenbeckia runyonii</i> (Rutaceae)
QPN0091	<chem>O(C)c1c(OC)cc2N=C3OC=CC3C(OC)c2c1</chem>	Rutaceae	
QPN0092	<chem>O(C)c1c(C/C=C(\C)/C)c2C(=O)c3c(O)c(C/C=C(\C)/C)c(O)cc3Oc2cc1</chem>	Clusiaceae	Diabetes y obesidad actividad frente a lipasa pancreática alfa amilasa alfa glucosidasa
QPN0093	<chem>O(C)c1c(C/C=C(\C)/C)c2C(=O)c3c(O)c(C/C=C(\C)/C)c(O)cc3Oc2cc1O</chem>	Clusiaceae	
QPN0094	<chem>O(C)c1cc(C/C=C(\C)/C)c2Oc3c(c(O)c(C/C=C(\C)/C)c(O)c3)C(=O)c2c1</chem>	Clusiaceae	
QPN0095	<chem>O=C1c2c(O)c(C/C=C(\C)/C)c(O)c(C/C=C(\C)/C)c2Oc2c1c(O)c(C/C=C(\C)/C)cc2</chem>	Clusiaceae	
QPN0096	<chem>O=C1c2c(O)c(C/C=C(\C)/C)c(O)cc2Oc2c(C/C=C(\C)/C)cccc12</chem>	Clusiaceae	

QPN0097	<chem>O(C)c1c(O)cc2C(=O)Nc3c2c1c1c(c3)cccc1</chem>	Piperaceae	Estudio fitoquímico de Algunos Metabolitos Secundarios Presentes en el Extracto Etanolico de Madera de la Especie <i>Piper hispidum</i>
QPN0098	<chem>O=C1C=2C(C)(C3C(C4C(C)(C(C(/C=C/C(C(C)C)CC)C)CC4)CC3)C1)CCC(=O)C=2</chem>	Piperaceae	
QPN0099	<chem>O=C1C=2C(C)(C3C(C4C(C)(C(C(CCC(C(C)C)CC)C)CC4)CC3)C1)CCC(=O)C=2</chem>	Piperaceae	
QPN0100	<chem>O(C)c1c(O)c2c(-c3c4OCOc4c(OC)c4c3[C@@H](N(C)CC4)C2)cc1OC</chem>	Lauraceae	Estudio fitoquímico de Compuestos Fijo y Aceite Esencial de hojas de la especie <i>Rhodostemonoda phne</i> sp.
QPN0101	<chem>O(C)c1c(O)c2c(-c3c4OCOc4cc4c3[C@@H](N(C)CC4)C2)cc1OC</chem>	Lauraceae	
QPN0102	<chem>O(C)c1c2OCOc2c-2c3[C@@H](N(C)CCc13)Cc1c-2cc(OC)c(OC)c1</chem>	Lauraceae	
QPN0103	<chem>O(C)c1c(OC)cc2c(-c3c4OCOc4cc4c3[C@@H](N(C)CC4)C2)c1</chem>	Lauraceae	
QPN0104	<chem>O(C)c1c(OC)cc2c(OC(=O)C(C(C=C)(C)C)=C2)c1</chem>	Rutaceae	
QPN0105	<chem>O=C(NCC(C)C)/C=C/C=C/CCCC</chem>	Rutaceae	Estudio fitoquímico de los Extractos Etanolicos de Hojas Corteza y Madera de <i>Esenbeckia cf alata</i> (Rutaceae) Evaluacion de Actividades biológicas y Etnobotánica de Cinco Familias de Montes de María
QPN0106	<chem>O=C1C(C(C=C)(C)C)=Cc2c(O1)cc1OC(C(=C)C)Cc1c2</chem>	Rutaceae	
QPN0107	<chem>O=C1Oc2c(c(O)c3c(OC(C(=C)C)C3)c2)C=C1</chem>	Rutaceae	
QPN0108	<chem>O=C1Oc2c(cc3c(OC(C(=C)C)C3)c2)C=C1</chem>	Rutaceae	
QPN0109	<chem>O=C1c2c(O)cccc2OC(C)C1</chem>	Rutaceae	
QPN0110	<chem>C(=C)Cc1cc2OCOc2cc1</chem>	Lauraceae	Estudio fitoquímico de los Extractos Etanolicos de madera corteza y hojas de <i>Beilschmiedia tovarensis</i> (Meins)Sa. Nishida (Lauraceae)
QPN0111	<chem>C(=C/c1cc2c(C)c(-c3cc4OCOc4cc3)oc2cc1)\C</chem>	Lauraceae	
QPN0112	<chem>O(C)c1c(O)cc2c(c1)C1=C3OCOC3C(OC)C3=C1C(NCC3)C2</chem>	Lauraceae	
QPN0113	<chem>O(C)c1c(O)cc2c(c1)C1=C3OCOC3CC3=C1C(NCC3)C2</chem>	Lauraceae	
QPN0114	<chem>O(C)c1c(O)ccc(-c2c(C)c3c(o2)ccc(/C=C/C)C3)c1</chem>	Lauraceae	
QPN0115	<chem>O(C)c1c(O)ccc(-c2c(C)c3c(o2)ccc(C(O)C(O)C)C3)c1</chem>	Lauraceae	
QPN0116	<chem>O(C)c1c(O)ccc(/C=C/CO)c1</chem>	Lauraceae	
QPN0117	<chem>O(C)c1c(OC)c(OC)c(/C=C/Cc2cccc2)c(O)c1OC</chem>	Lauraceae	

110 Contribución a la identificación de potenciales antagonistas asociados al quorum sensing en *Pseudomonas aeruginosa* a través de métodos quimioinformáticos.

QPN0118	<chem>O(C)c1c(OC)cc(OC2C(O)C(O)C(O)C(COC3C(O)C(O)C(O)C(C)O3)O2)cc1OC</chem>	Lauraceae
QPN0119	<chem>O(C)c1c(OC)cc([C@@H]2OCC3C(c4cc(OC)c(OC)c(OC)c4)OCC23)cc1OC</chem>	Lauraceae
QPN0120	<chem>O(C)c1c(OC)cc([C@@H]2OCC3C(c4cc(OC)c5OCO c5c4)OCC23)cc1OC</chem>	Lauraceae
QPN0121	<chem>O(C)c1c(OC)ccc(/C=C/CO)c1</chem>	Lauraceae
QPN0122	<chem>O(C)c1c(OC)ccc(CC=C)c1</chem>	Lauraceae
QPN0123	<chem>O(C)c1c2OCOc2cc(C2C3(CC=C)C(=O)C=C4OCOC 4(C2C)C3)c1</chem>	Lauraceae
QPN0124	<chem>O(C)c1c2OCOc2cc(C2C3(CC=C)C(=O)CC4OCOC4(C2C)C3)c1</chem>	Lauraceae
QPN0125	<chem>O(C)c1c2OCOc2cc(C2C3(CC=C)C(O)CC4(OCOC4) C(C2C)C3)c1</chem>	Lauraceae
QPN0126	<chem>O(C)c1c2oc(c(C)c2cc(/C=C/C)c1)-c1cc2OCOc2cc1</chem>	Lauraceae
QPN0127	<chem>O(Cc1cc(OC)c(OC)cc1)[C@@H]1[C@@H](C)[C@] 2(CC=C)C(O)(C(=O)C=CC2)O1</chem>	Lauraceae
QPN0128	<chem>O(Cc1cc(OC)c(OC)cc1)[C@H]1[C@@H](C)[C@]2(CC=C)C(OC)(C(=O)C=CC2)O1</chem>	Lauraceae
QPN0129	<chem>O(Cc1cc2OCOc2cc1)[C@@H]1[C@@H](C)[C@]2(CC=C)C(OC)(C(=O)C=CC2)O1</chem>	Lauraceae
QPN0130	<chem>O(Cc1cc2OCOc2cc1)[C@H]1[C@@H](C)[C@]2(C C=C)C(O)(C(=O)C=CC2)O1</chem>	Lauraceae
QPN0131	<chem>O([C@@H](C[C@@H](O)CC1OC(=O)C=CC1)CCC Cc1cccc1)C(=O)C</chem>	Lauraceae
QPN0132	<chem>O([C@H](C[C@H](O)CCCCc1cccc1)CC1OC(=O)C =CC1)C(=O)C</chem>	Lauraceae
QPN0133	<chem>O=C(/C=C/CCC/C=C\C/C=C\C\CCCC)CC(O)(O)COC(=O)C</chem>	Lauraceae
QPN0134	<chem>O=C(/C=C/c1ccc(O)cc1)CNCc1c2c([nH]c1)cccc2</chem>	Lauraceae
QPN0135	<chem>O=C(C(OC1ccc(/C=C/C)cc1)C)c1ccc(O)cc1</chem>	Lauraceae
QPN0136	<chem>O=C(C)c1c(OC)cc2c(- c3c4OCOc4cc4c3C(N(C)CC4)C2)c1</chem>	Lauraceae
QPN0137	<chem>O=C(CCC1cccc1)c1c(OC)c(O)c(OC)c(O)c1OC</chem>	Lauraceae
QPN0138	<chem>O=C(C[C@@H](O)COC(=O)C)CCCC/C=C\C/C=C\C CCCC</chem>	Lauraceae

QPN0139	<chem>O=C(O)[C@@]12[C@@H]([C@@](O)(C)C(C)CC1)C=1C(C)([C@@]3(C)C([C@]4(C)C(C)(C)[C@@H](O)C(O)C4)CC3)CC=1)CC2</chem>	Lauraceae	
QPN0140	<chem>O=C(OC)/C/C(O)C(=O)C=C\CCCCCCCCCCCC</chem>	Lauraceae	
QPN0141	<chem>O=C(OC)N1C2C=3C(c4c(cc(O)c(OC)c4)C2)=C2OCOC2C(OC)C=3CC1</chem>	Lauraceae	
QPN0142	<chem>O=C(OC)N1C2c3c(c4OCOc4cc3CC1)-c1c(cc(O)c(OC)c1)C2</chem>	Lauraceae	
QPN0143	<chem>O=C(OC/C=C/c1cc2OCOc2cc1)C</chem>	Lauraceae	
QPN0144	<chem>O=C(OC1C2=C(C)C(O)OC2(O)C[C@]2(C)C1C(=C)C1C2C1)C</chem>	Lauraceae	
QPN0145	<chem>O=C/C=C/CCCCCCCC/C=C/1\[C@H](O)COC\1=O</chem>	Lauraceae	
QPN0146	<chem>O=C/C=C/c1cc(OC)c(OC)cc1</chem>	Lauraceae	
QPN0147	<chem>O=C/C=C/c1cc2OCOc2cc1</chem>	Lauraceae	
QPN0148	<chem>O=C/C=C/c1cc2c(C)c(-c3cc4OCOc4cc3)oc2cc1</chem>	Lauraceae	
QPN0149	<chem>O=C1/C(=C/CCCCCCCC/C=C/C=C)/[C@H](O)CO1</chem>	Lauraceae	
QPN0150	<chem>O=C1/C(=C/CCCCCCCC#C)/[C@H](O)CO1</chem>	Lauraceae	
QPN0151	<chem>O=C1/C(=C/CCCCCCCCCCCC)/[C@H](O)CO1</chem>	Lauraceae	
QPN0152	<chem>O=C1C(/C=C/CCCCCCCC/C=C/C=C)=CC(C)O1</chem>	Lauraceae	
QPN0153	<chem>O=C1C(/C=C/CCCCCCCC#C)=CC(C)O1</chem>	Lauraceae	
QPN0154	<chem>O=C1C(/C=C/CCCCCCCCCCCC)=CC(C)O1</chem>	Lauraceae	
QPN0155	<chem>OC(C(O)c1cc2c(C)c(-c3ccc(O)cc3)oc2cc1)C</chem>	Lauraceae	
QPN0156	<chem>OC(CO)CC(O)CCCCCCCCCCCC#C</chem>	Lauraceae	
QPN0157	<chem>OC(CO)CC(O)CCCCCCCCCCCC=C</chem>	Lauraceae	
QPN0158	<chem>OC(CO)CC(O)CCCCCCCCCCCC</chem>	Lauraceae	
QPN0159	<chem>OC/C=C/c1cc2OCOc2cc1</chem>	Lauraceae	
QPN0160	<chem>OC1C(C)=C2[C@@](C3[C@@]4(C)C=CCC(=C)C4C4=C(C)C(O)OC34)(O1)C[C@@]1(C)C=CCC(=C)C1C2</chem>	Lauraceae	
QPN0161	<chem>O[C@@H]1[C@@H]2[C@H](C(C)CCCC3C(C)(C)C23)CC1</chem>	Lauraceae	
QPN0162	<chem>O[C@@]12OCC(C)=C1CC1C(=C)CC=C[C@]1(O)C2</chem>	Lauraceae	
QPN0163	<chem>Oc1ccc(-c2c(C)c3c(o2)ccc(/C=C/C)c3)cc1</chem>	Lauraceae	
QPN0164	<chem>O(C)c1c(O)ccc([C@@H]2[C@@H](O)Cc3c(O2)cc(O)cc3)c1</chem>	Myristicaceae	Estudio fitoquímico y Toxicidad de los Extractos
QPN0165	<chem>O=C(CCCCCCCCCc1cccc1)C=1C(=O)C(O)CCC=1O</chem>	Myristicaceae	

112 Contribución a la identificación de potenciales antagonistas asociados al quorum sensing en *Pseudomonas aeruginosa* a través de métodos quimioinformáticos.

QPN0166	<chem>O=C(CCCCCCCCCc1ccccc1)C=1C(=O)[C@@H](O)CCC=1O</chem>	Myristicaceae	Etanolicos de <i>Virola venosa</i> (Myristicaceae).
QPN0167	<chem>O[C@@H]1[C@@H](c2cc(O)c(O)cc2)Oc2c(ccc(O)c2)C1</chem>	Myristicaceae	
QPN0168	<chem>O(C)c1c2OCOc2cc([C@@H]2[C@@H](C)[C@]3(O)C[C@H](O)[C@@]2(OC)C(=O)C(CC=C)=C3)c1</chem>	Lauraceae	Estudio fitoquímico y citotóxico del extracto etanolico del fruto de la <i>Ocotea heterochroma</i>
QPN0169	<chem>O=C(O[C@@H]1[C@]2(OC)[C@@H]([C@@H](C)[C@@]1(OC)C=C(CC=C)C2=O)c1cc(OC)c2OCOc2c1)C</chem>	Lauraceae	
QPN0170	<chem>O(C)c1c(OC)c2c3c(c1OC)C(=O)Nc3cc1c2cc(O)cc1</chem>	Anonaceae	Estudio fitoquímico y de actividad biológica de los extractos etanolicos de hojas, corteza y madera de la especie <i>Rollinia membranacea</i> (Anonaceae)
QPN0171	<chem>O=C1C(CCCC)=CC(C)O1</chem>	Anonaceae	
QPN0172	<chem>O=C1c2c(OC(c3cc4OCOc4cc3)=C1)cc(O)cc2</chem>	Anonaceae	
QPN0173	<chem>C=C1C2C(C(C)(C)C2)CC=C(C)CCC1</chem>	Myristicaceae	Estudio Parcial del Extracto bencenico de las hojas de <i>Virola elongata</i>
QPN0174	<chem>C=C1C2C(C(C)(C)C2)CCC(C)=CCC1</chem>	Myristicaceae	
QPN0175	<chem>O=C(CC(C)C)CC1(C)OC(c2cocc2)CC1</chem>	Myristicaceae	
QPN0176	<chem>O[C@@H]1CC=2[C@@](C)([C@@H]3[C@H]([C@H]4[C@@](C)(C([C@@H](CCCC(C)C)CC4)CC3)CC=2)CC1</chem>	Myristicaceae	
QPN0177	<chem>O=C(OC1OC(=O)C2=CO(OC3C(O)C(O)C(O)C(CO)O3)C(C=C)C2C1)C(C)C</chem>	Gentianaceae	Estudio de hojas de <i>Chelonantus alatus</i>
QPN0178	<chem>O=C(OC1OC(=O)C2=CO(OC3C(O)C(O)C(O)C(CO)O3)C(C=C)C2C1)C(CC)C</chem>	Gentianaceae	
QPN0179	<chem>O=C1OC(OC)CC2C(C=C)C(OC3C(O)C(O)C(O)C(CO)O3)OC=C12</chem>	Gentianaceae	
QPN0180	<chem>O=C1OCCC2C(C=C)C(OC3C(O)C(O)C(O)C(CO)O3)OC=C12</chem>	Gentianaceae	
QPN0181	<chem>C(CCCC(CCCCCCCCC)C)(CCCC(CCCCCCCCC)C)C</chem>	Solanaceae	Estudio de metabolitos fijos y volátiles en tres morfotipos de cocona <i>Solanum sessiliflorum</i> Dunal
QPN0182	<chem>C(CCCCCCCCC(C)C)(CCCCCCCCCCCCCCCC)C</chem>	Solanaceae	
QPN0183	<chem>C(CCCCCCCCCCCCC(C)C)(CCCCCCCCCCCC)C</chem>	Solanaceae	
QPN0184	<chem>C(CCCCCCCCCCCCCCCCC)C</chem>	Solanaceae	
QPN0185	<chem>O=C(O)/C=C/c1ccc(O)cc1</chem>	Solanaceae	

QPN0186	<chem>O=C(O)CCC/C=C\CCCCCCCCCO(=O)C</chem>	Solanaceae	procedentes del Guaviare
QPN0187	<chem>O=C(O)CCCCCCCCCCCCCCC</chem>	Solanaceae	
QPN0188	<chem>O=C(O)CCc1ccc(O)cc1</chem>	Solanaceae	
QPN0189	<chem>O=C(OCC)CCCCCCC/C=C/C/C=C/CCCCC</chem>	Solanaceae	
QPN0190	<chem>O=C(OCC)CCCCCCC/C=C/CCCCCCC</chem>	Solanaceae	
QPN0191	<chem>O=C(OCC)CCCCCCCCCCCCCCC</chem>	Solanaceae	
QPN0192	<chem>O=C(OCC)CCC1C(C)C=2N=C1C=1C(C(=O)OC)C(=O)c3c(C)[cH0]([nH][cH0]=13)=CC=1C(CC)=C(C)C(N=1)=Cc1c(C=C)c(C)c([nH]1)C=2</chem>	Rubiaceae	Estudio fitoquímico de algunos metabolitos secundarios presentes en el extracto de éter de petróleo de hojas de <i>Uncaria guianensis</i>
QPN0193	<chem>Clc1c(/C=C/C(=O)c2ccc(N)cc2)cccc1</chem>	Piperaceae	Estudio fitoquímico de dos especies del género <i>Piper</i> (<i>Piper submentosum</i> , <i>Piper septuplinervium</i>) Piperaceae y determinación de actividad insecticida sobre <i>Spodoptera frugiperda</i> (Lepidoptera)
QPN0194	<chem>Clc1c(/C=C/C(=O)c2ccc(NC(=O)Nc3ccc(F)cc3)cc2)cccc1</chem>	Piperaceae	
QPN0195	<chem>Fc1ccc(NC(=O)Nc2ccc(C(=O)/C=C/c3cc([N+](=O)[O-])ccc3)cc2)cc1</chem>	Piperaceae	
QPN0196	<chem>O(C)c1c(CCCc2c(O)cc(O)cc2)cc2OCOc2c1</chem>	Piperaceae	
QPN0197	<chem>O(C)c1c(O)cc(C2=C(OC)C(=O)c3c(O)cc(OC)cc3O2)cc1</chem>	Piperaceae	
QPN0198	<chem>O(C)c1cc(O)c2C(=O)CC(c3ccc(OC)cc3)Oc2c1</chem>	Piperaceae	
QPN0199	<chem>O(C)c1cc(O)c2C(=O)C[C@@H](c3cccc3)Oc2c1</chem>	Piperaceae	
QPN0200	<chem>O(C1C(O)C(O)C(O)C(CO)O1)C1CC=2C(C)(C3C(C4C(C)(C(C(CCC(C(C)C)CC)C)CC4)CC3)CC=2)CC1</chem>	Piperaceae	
QPN0201	<chem>O=C(/C=C/c1ccc(C)cc1)c1ccc(N)cc1</chem>	Piperaceae	
QPN0202	<chem>O=C(/C=C/c1ccc(N(C)C)cc1)c1ccc(N)cc1</chem>	Piperaceae	
QPN0203	<chem>O=C(/C=C/c1ccc(N(C)C)cc1)c1cccc1</chem>	Piperaceae	
QPN0204	<chem>O=C(/C=C/c1cccc1)c1cccc1</chem>	Piperaceae	
QPN0205	<chem>O=C(CCc1cccc1)c1c(OC)cc(O)cc1O</chem>	Piperaceae	
QPN0206	<chem>O=C(NCCc1ccc(O)cc1)/C=C/c1ccc(O)cc1</chem>	Piperaceae	
QPN0207	<chem>O=C(O)c1cc(O)c(O)cc1</chem>	Piperaceae	
QPN0208	<chem>O=C(OC1=C(c2cccc2)Oc2c(c(O)cc(O)c2)C1=O)C</chem>	Piperaceae	
QPN0209	<chem>O=C(OC1C(O)C(OC(=O)/C=C/c2ccc(O)cc2)C(C)OC1OC1=C(c2ccc(O)cc2)Oc2c(c(O)cc(O)c2)C1=O)/C=C/c1ccc(O)cc1</chem>	Piperaceae	

QPN0210	<chem>O=C(Oc1c(C(=O)CCc2ccccc2)c(OC)cc(OC(=O)C)c1)C</chem>	Piperaceae	Estudio fitoquímico de las hojas de la especie vegetal <i>Croton schiedeanus</i>
QPN0211	<chem>O=C(Oc1c2C(=O)C(OC(=O)C)=C(c3ccccc3)Oc2cc(OC(=O)C)c1)C</chem>	Piperaceae	
QPN0212	<chem>O=C(Oc1cc(O)c2C(=O)C(OC(=O)C)=C(c3ccccc3)Oc2c1)C</chem>	Piperaceae	
QPN0213	<chem>O=C(Oc1cc(OC)c(C(=O)CCc2ccccc2)c(O)c1)C</chem>	Piperaceae	
QPN0214	<chem>O=C1C(O)=C(c2ccccc2)Oc2c1c(O)cc(O)c2</chem>	Piperaceae	
QPN0215	<chem>O=C1Oc2c([C@@H](c3cc(O)c(O)cc3)C1)c(O)cc1OC(c3cc(O)c(O)cc3)C(O)Cc21</chem>	Piperaceae	
QPN0216	<chem>O=C1c2c(O)cc(O)cc2OC(c2ccccc2)=C1</chem>	Piperaceae	
QPN0217	<chem>O=[N+](O)c1cc(/C=C/C(=O)c2ccc(N)cc2)ccc1</chem>	Piperaceae	
QPN0218	<chem>Oc1c(CCCc2cc3OCOc3cc2)ccc(O)c1</chem>	Piperaceae	
QPN0219	<chem>C(=C)(C)C1C(C=C)(C)CCC(C(=O)C)C1</chem>	Euphorbiaceae	
QPN0220	<chem>C(C)(C)C1/C=C/C(=O)CC/C=C(\C)/CC1</chem>	Euphorbiaceae	
QPN0221	<chem>C(C=C)(=O)CCC=C(C)C</chem>	Euphorbiaceae	
QPN0222	<chem>C=C1C2C(C)(C)C(C2)CC1</chem>	Euphorbiaceae	
QPN0223	<chem>CC1(C)C2C(C)=CCC1C2</chem>	Euphorbiaceae	
QPN0224	<chem>O(C)c1c(O)ccc(C2=C(OC)C(=O)c3c(O)cc(OC)cc3O2)c1</chem>	Euphorbiaceae	
QPN0225	<chem>O(C)c1c(OC)ccc(C2=C(OC)C(=O)c3c(O)cc(OC)cc3O2)c1</chem>	Euphorbiaceae	
QPN0226	<chem>O(C)c1cc(O)c2C(=O)C(O)=C(c3cc(O)c(O)cc3)Oc2c1</chem>	Euphorbiaceae	
QPN0227	<chem>O(C)c1cc(O)c2C(=O)C(OC)=C(c3cc(O)c(O)cc3)Oc2c1</chem>	Euphorbiaceae	
QPN0228	<chem>O=C(Oc1c(OC(=O)C)ccc(C2=C(OC)C(=O)c3c(O)cc(OC)cc3O2)c1)C</chem>	Euphorbiaceae	
QPN0229	<chem>O=C(Oc1c(OC(=O)C)ccc(C2=C(OC)C(=O)c3c(OC(=O)C)cc(OC)cc3O2)c1)C</chem>	Euphorbiaceae	
QPN0230	<chem>O=C(Oc1c(OC)ccc(C2=C(OC)C(=O)c3c(O)cc(OC)c3O2)c1)C</chem>	Euphorbiaceae	
QPN0231	<chem>O(C)c1c(O)cc2c(c3c(nc2)c2c(cc4OCOc4c2)cc3)c1</chem>	Rutaceae	Estudio fitoquímico parcial de corteza de <i>zanthoxylum quinduense</i>

QPN0232	<chem>C(=C)CCCCCCCCCCCCCCC</chem>	Leguminosae	Estudio fitoquímico parcial del extracto isopropanolico de corteza de <i>Clitoria</i> sp
QPN0233	<chem>NCCCCC1c(/C=C/CCC)cccc1</chem>	Leguminosae	
QPN0234	<chem>O=C(O)CCCCCCCCCCCCC</chem>	Leguminosae	
QPN0235	<chem>O=C(O)CCCCCCCCCCCCC</chem>	Leguminosae	
QPN0236	<chem>O=C(OCCCC)CCC</chem>	Leguminosae	
QPN0237	<chem>O=C(OCCCC)CCCC</chem>	Leguminosae	
QPN0238	<chem>O=C(OCCCC)CCCCC</chem>	Leguminosae	
QPN0239	<chem>OCCCCCCCCCCCCCCC</chem>	Leguminosae	
QPN0240	<chem>O(C)c1cc(O)c2C(=O)C[C@@H](c3ccc(O)cc3)Oc2c1</chem>	Piperaceae	Estudio fitoquímico y evaluación de actividad Leshmanicida de la especie <i>Piper hispidum</i>
QPN0241	<chem>O(C)c1ccc(C2Oc3c(c(O)cc(O)c3)C(=O)C2)cc1</chem>	Piperaceae	
QPN0242	<chem>O=C(C)c1ccc(C2Oc3c(c(O)cc(OC)c3)C(=O)C2)cc1</chem>	Piperaceae	
QPN0243	<chem>O=C(Cc1cccc1)c1c(O)cc(O)cc1O</chem>	Piperaceae	
QPN0244	<chem>O=C(Cc1cccc1)c1c(O)cc(OC)cc1O</chem>	Piperaceae	
QPN0245	<chem>O=C(Cc1cccc1)c1c(OC)cc(O)cc1OC</chem>	Piperaceae	
QPN0246	<chem>O=C(Cc1cccc1)c1c(OC)cc(OC)cc1OC</chem>	Piperaceae	
QPN0247	<chem>O=C(Oc1c2C(=O)CC(c3ccc(OC)cc3)Oc2cc(OC)c1)C</chem>	Piperaceae	
QPN0248	<chem>O=C(Oc1c2C(=O)CC(c3cccc3)Oc2cc(OC)c1)C</chem>	Piperaceae	
QPN0249	<chem>O=C1c2c(O)cc(O)cc2OC(c2ccc(O)cc2)C1</chem>	Piperaceae	
QPN0250	<chem>O=C1c2c(O)cc(O)cc2O[C@H](c2cccc2)C1</chem>	Piperaceae	
QPN0251	<chem>Oc1c(O)ccc(O)c1</chem>	Piperaceae	
QPN0252	<chem>O=C1C=CC(O)(CC/C=C\CCCCC)C=C1</chem>	Piperaceae	
QPN0253	<chem>Oc1c(O)ccc(CC/C=C/CCCCC)c1</chem>	Piperaceae	
QPN0254	<chem>Oc1ccc(CC/C=C/CCCCC)cc1</chem>	Piperaceae	
QPN0255	<chem>C(C)(C)C1C2C(C)(C3OC2C(=C)CC3)CC1</chem>	Lauraceae	Estudio fitoquímico bioguiado de los extractos de corteza de las especies <i>Endlicheria oreocola</i> y <i>Endlicheria arenosa</i> (Lauraceae) en busca de fuentes para el control del hongo
QPN0256	<chem>O(C)c1c(OC)c2c3c(C(=O)c4c(-c3c1OC)cccc4)ncc2</chem>	Lauraceae	
QPN0257	<chem>O=C1/C(=C/CCCCCCCCC)/[C@H](O)[C@H](C)O1</chem>	Lauraceae	
QPN0258	<chem>O=C1/C(=C\CCCCCCCCC)/C(O)C(C)O1</chem>	Lauraceae	
QPN0259	<chem>O=C1C(C(O)CCCCCCCC)=C(C)C(O)(C)O1</chem>	Lauraceae	
QPN0260	<chem>OC1C(C(C)C)CCC2(C)C(O)CCC(=C)C12</chem>	Lauraceae	
QPN0261	<chem>O[C@@H]1CC=2[C@@](C)([C@@H]3[C@H]([C@H]4[C@@](C)([C@@H]([C@@H]1/C=C/[C@H](C)C)CC)C)CC4)CC3)CC=2)CC1</chem>	Lauraceae	

QPN0262	<chem>O[C@@H]1CC=2[C@@](C)([C@@H]3[C@H]([C@H]4[C@@](C)([C@@H]([C@@H](CC[C@H](C(C)C)C)CC4)CC3)CC=2)CC1</chem>	Lauraceae	fitopatogeno <i>Colletotrichum tamarilloi</i>
QPN0263	<chem>O[C@@H]1CC=2[C@@](C)([C@@H]3[C@H]([C@H]4[C@@](C)([C@@H]([C@@H](CC[C@H](C(C)C)C)CC4)CC3)CC=2)CC1</chem>	Lauraceae	
QPN0264	<chem>O(C)c1c(OC)cc2-c3c4c(cc5OCoc35)ccnc4C(=O)c2c1</chem>	Magnoliaceae	Estudio fitoquímico de las hojas y corteza de <i>Talauma arcabucoana</i> (MAGNOLIACEAE) y evaluación de su toxicidad y actividad antimicrobiana
QPN0265	<chem>O(C)c1c(OC)cc2c(-c3c4OCoc4cc4c3C(NCC4)C2)c1</chem>	Magnoliaceae	
QPN0266	<chem>O=C(O)C1=C[C@@H](O)[C@@H](O)[C@H](O)C1</chem>	Magnoliaceae	
QPN0267	<chem>O=C1C(O)(C)C(O)CO1</chem>	Magnoliaceae	
QPN0268	<chem>O=CN1C2c3c(c4OCoc4cc3CC1)-c1c(cc(OC)c(OC)c1)C2</chem>	Magnoliaceae	
QPN0269	<chem>O(C)c1c(OC)ccc(CC2C(=O)OCC2Cc2cc(OC)c(OC)c2)c1</chem>	Rutaceae	Estudio fitoquímico de los extractos etanolicos de hojas corteza y madera de <i>Zanthoxylum setulosum</i> (RUTACEAE)
QPN0270	<chem>O(C)c1c(OC)ccc(C[C@@H]2[C@H](Cc3cc4OCoc4cc3)C(=O)OC2)c1</chem>	Rutaceae	
QPN0271	<chem>O=C(C(C)C1C2(C)C(C3C(C4(C)C(=CC3)C[C@@H](O)CC4)CC2)CC1)CC(C(C)C)C</chem>	Rutaceae	
QPN0272	<chem>O=C(C)c1cc(O)c(O)cc1</chem>	Rutaceae	
QPN0273	<chem>O=C(O)CCCCCCCCCCCCCCCCCCCCCCCCC</chem>	Rutaceae	
QPN0274	<chem>O=C(Oc1cc(O)c2C(=O)CC(c3ccc(O)cc3)Oc2c1)C</chem>	Rutaceae	
QPN0275	<chem>O=C(Oc1cc(O)c2C(=O)CC(c3ccc(OC(=O)C)cc3)Oc2c1)C</chem>	Rutaceae	
QPN0276	<chem>O=C(Oc1ccc(C2Oc3c(c(O)cc(O)c3)C(=O)C2)cc1)C</chem>	Rutaceae	
QPN0277	<chem>O=C(CCc1cc(OC)c(OC)cc1)c1c(OC)cc(O)cc1O</chem>	Myristicaceae	Estudio fitoquímico de una parte del extracto bencenico de corteza <i>Iryanthera laevis</i>
QPN0278	<chem>CC1=C(c2cc3OCoc3cc2)c2c3OCoc3ccc2CC1C</chem>	Myristicaceae	Estudio fitoquímico del fruto de <i>Osteophoeum sulcatum</i>
QPN0279	<chem>CC1C(c2cc3OCoc3cc2)c2c3OCoc3ccc2CC1C</chem>	Myristicaceae	
QPN0280	<chem>O(C)c1ccc(C=2C(=O)c3c(OC=2)cc(O)cc3)cc1</chem>	Myristicaceae	

QPN0281	<chem>O=C(CC1C(CCCCC)OC2OC(CO)C(OC(=O)c3ccc(O)cc3)C(O)C2O1)C</chem>	Myristicaceae	Estudio fitoquímico de <i>Compsonera capitellata</i> (Myristicaceae), <i>Zanthoxylum rigidum</i> (Rutaceae) y <i>Ocotea longifolia</i> (Lauraceae) y evaluación de su posible aplicación como biocontroladores de <i>Sitophilus</i> sp
QPN0282	<chem>O=C(Cc1cc(O)c(OC)cc1)c1c(OC)cc(O)cc1O</chem>	Myristicaceae	
QPN0283	<chem>O=C(Cc1cc2OCOc2cc1)c1c(OC)cc(O)cc1O</chem>	Myristicaceae	
QPN0284	<chem>O=C(O)/C=C/c1cc(O)c(OC2[C@H](O)CC(O)(C(=O)O)C[C@@H]2O)cc1</chem>	Rutaceae	
QPN0285	<chem>O=C(O)C(C)C1CC2C(=C)CCCC2(C)CC1</chem>	Lauraceae	
QPN0286	<chem>O=C(O)C1C2C(C(C)C)CCC(O)(C)[C@@H]2CC1</chem>	Myristicaceae	
QPN0287	<chem>O=C(OCC(O)CO)CCCCCCCCCCCC</chem>	Rutaceae	
QPN0288	<chem>O=C(OCC)/C=C/c1cc(O)c(O)cc1</chem>	Rutaceae	
QPN0289	<chem>O=C1/C(=C\CCCCCCCCC)/[C@@H](O)C(=C)O1</chem>	Lauraceae	
QPN0290	<chem>O=C1/C(=C\c2cc3OCOc3cc2)/C(Cc2cc3OCOc3cc2)CO1</chem>	Rutaceae	
QPN0291	<chem>O=C1C(C)(C)C=2C(C)(C3C(C)(C4C(C)CC5C(C(C=C)C)CC5)C4CC3)CC=2)CC1</chem>	Rutaceae	
QPN0292	<chem>O=C1C(C)=C2C(C)(O1)CC1(C)C(C(=C)CCC1)C2</chem>	Lauraceae	
QPN0293	<chem>O=C1C(C)=C2C(O)(O1)CC1(C)C(C(=C)CCC1)C2C</chem>	Lauraceae	
QPN0294	<chem>O=C1C(C)=C2C(O)C3C(=C)CCCC3(C)C=C2O1</chem>	Lauraceae	
QPN0295	<chem>O=C1C(CCCCCCCCCCCCCC/C=C/CCCC)C(O)C(C)O1</chem>	Myristicaceae	
QPN0296	<chem>O=C1C(Cc2cc3OCOc3cc2)=C(Cc2cc3OCOc3cc2)CO1</chem>	Rutaceae	
QPN0297	<chem>O=Cc1cc(O)c(O)cc1</chem>	Myristicaceae	
QPN0298	<chem>OC(CO)(C)C1CC2C(=C)CCCC2(C)CC1</chem>	Lauraceae	
QPN0299	<chem>OC1C(C)(C)C=2C(C)(C3C(C)(C4(C)C(C5C(C)(CC4)CCC5(=C)C)CC3)CC=2)CC1</chem>	Rutaceae	
QPN0300	<chem>OCC12C(C(C(=C)C)CC1)C1C(C3(C)C(C4(C)C(C(C)(C)C(O)CC4)=CC3)CC1)C(C)C2</chem>	Rutaceae	
QPN0301	<chem>O[C@]1(C)[C@H]2C(C)(CCC(C(=C)C)C2)CCC1</chem>	Myristicaceae	
QPN0302	<chem>O(C)c1c(O)ccc(C[C@@H]([C@@H](Cc2cc(OC)c(O)cc2)C)C)c1</chem>	Magnoliaceae	Estudio fitoquímico y búsqueda de compuestos de interés terapéutico de <i>Dugandiodendron argyrotrichum</i>
QPN0303	<chem>O(C)c1c(O)ccc(C[C@@H]([C@@H](Cc2cc3OCOc3cc2)C)C)c1</chem>	Magnoliaceae	
QPN0304	<chem>O=C(C)N1C2c3c(c4OCOc4cc3CC1)-c1c(ccc1)C2</chem>	Magnoliaceae	
QPN0305	<chem>O=C1C(=C)[C@@H]2[C@@H](O1)[C@@H]1[C@](C)(O1)CC/C=C/C\CC2</chem>	Magnoliaceae	

QPN0306	<chem>O=C(Cc1ccc(OC)cc1)c1c(OC)cc(O)cc1O</chem>	Myristicaceae	Estudio parcial del extracto tolueno de la corteza de <i>Virola sebifera</i>
QPN0307	<chem>O(C)c1c(O)ccc(C2Oc3c(ccc(OC)c3)CC2)c1</chem>	Myristicaceae	Estudio químico de especies de la familia Myristicaceae <i>Virola calophylloidea</i>
QPN0308	<chem>O(C)c1c(O)ccc(CCCc2c(O)cc(OC)cc2)c1</chem>	Myristicaceae	
QPN0309	<chem>O(C)c1c2OCOc2c2C(c3cc4OCOc4cc3)=C(C)[C@H](C)Cc2c1</chem>	Myristicaceae	
QPN0310	<chem>O(C)c1c2OCOc2c2[C@](O)([C@H](C)[C@H](C)Cc2c1)c1cc2OCOc2cc1</chem>	Myristicaceae	
QPN0311	<chem>O=C(C(O)Cc1ccc(OC)cc1)c1c(O)cc(OC)cc1</chem>	Myristicaceae	
QPN0312	<chem>O(C)c1c(O)c(-c2c(O)c(OC)cc(CC=C)c2)cc(CC=C)c1</chem>	Myristicaceae	Estudio químico del extracto bencénico de las hojas de <i>Virola carinata</i>
QPN0313	<chem>O(C)c1c(O)ccc(C2[C@@H](CO)c3c(c(OC)cc(CC=C)c3)O2)c1</chem>	Myristicaceae	
QPN0314	<chem>O(C)c1c(OC)ccc([C@@H]2[C@@H](CO)c3c(c(OC)cc(CC=C)c3)O2)c1</chem>	Myristicaceae	
QPN0315	<chem>O(C)c1c2OC([C@@H](CO)c2cc(CC=C)c1)c1cc2OCOc2cc1</chem>	Myristicaceae	
QPN0316	<chem>O(C)c1c(C(O)C)c(O)c2C(=O)C=C(OC)C(=O)c2c1</chem>	Lauraceae	
QPN0317	<chem>O(C)c1c(O)c(OC)cc(C2C(CO)C(CO)C(c3cc(OC)c(O)c(OC)c3)O2)c1</chem>	Lauraceae	
QPN0318	<chem>O(C)c1c(O)c(OC)cc2c1OC(=O)C=C2</chem>	Lauraceae	
QPN0319	<chem>O(C)c1c(O)ccc(C2C(O)Cc3c(OC)cc(OC)cc3O2)c1</chem>	Lauraceae	
QPN0320	<chem>O(C)c1c(OC)cc(O[C@@H]2C(O)C(O)[C@H](O)C(CO)O2)cc1OC</chem>	Lauraceae	
QPN0321	<chem>O(C)c1c2c(OC(c3cc4OCOc4cc3)C(O)C2)cc(OC)c1</chem>	Lauraceae	
QPN0322	<chem>O(C)c1ccc(C2=C(OC3C(O)C(O)C(C)O3)C(=O)c3c(O)cc(O)cc3O2)cc1</chem>	Lauraceae	
QPN0323	<chem>O([C@H]1C(O)C(O)[C@@H](O)C(C)O1)C1=C(c2c(O)c(O)cc2)Oc2c(c(O)cc(O)c2)C1=O</chem>	Lauraceae	
QPN0324	<chem>O([C@H]1C(O)C(O)[C@@H](O)C(C)O1)C1=C(c2c(O)cc(O)cc2)Oc2c(c(O)cc(O)c2)C1=O</chem>	Lauraceae	
QPN0325	<chem>O([C@H]1C(O)C(O)[C@H](CO)O1)C1=C(c2cc(O)c(O)cc2)Oc2c(c(O)cc(O)c2)C1=O</chem>	Lauraceae	
QPN0326	<chem>O([C@H]1C(O)C(O)[C@H](CO)O1)C1=C(c2ccc(O)cc2)Oc2c(c(O)cc(O)c2)C1=O</chem>	Lauraceae	

QPN0327	<chem>O=C(O[C@@H]1C(O)C(OC(=O)/C=C/c2ccc(O)cc2)[C@H](OC2=C(c3ccc(OC)cc3)Oc3c(c(O)cc(O)c3)C2=O)CC1C)/C=C\c1ccc(O)cc1</chem>	Lauraceae	
QPN0328	<chem>O=C(O[C@@H]1C(OC(=O)/C=C/c2ccc(O)cc2)C(O)[C@H](OC2=C(c3ccc(OC)cc3)Oc3c(c(O)cc(O)c3)C2=O)CC1C)/C=C/c1ccc(O)cc1</chem>	Lauraceae	
QPN0329	<chem>O=C1C(O)=C(c2cc(O)c(O)cc2)Oc2c1c(O)cc(O)c2</chem>	Lauraceae	
QPN0330	<chem>O([C@@H]1C(O)C(O)[C@H](OC2[C@H](O)OC(CO)[C@@H](O)C2O)OC1CO)c1c(O)ccc(C=2Oc3c(O)c(O)cc(O)c3C(=O)C=2)c1</chem>	Cannaceae	Estudio químico de algunos metabolitos secundarios de la especie <i>Canna edulis</i>
QPN0331	<chem>O(C)c1c(O)ccc(CCCc2c(OC)cc(O)cc2)c1</chem>	Myristicaceae	Estudio químico de los extractos bencénicos de madera y hojas de <i>Iryanthera triconis</i> Ducke
QPN0332	<chem>O=C(CCc1ccc(OC)cc1)c1c(O)c(C(C(Cc2ccc(O)cc2)C)C)c2ccc(O)cc2)c(O)cc1OC</chem>	Myristicaceae	
QPN0333	<chem>O=C(CCc1ccc(OC)cc1)c1c(O)c(C(C=C)c2ccc(O)cc2)c(O)cc1OC</chem>	Myristicaceae	
QPN0334	<chem>O=C(CCc1ccc(OC)cc1)c1c(O)cc(O)cc1O</chem>	Myristicaceae	
QPN0335	<chem>C(C)(C)[C@H]1C2[C@](C)([C@H]3[C@@H]2C(=C)CC3)CC1</chem>	Lauraceae	Estudio químico del tallo y volátiles de hojas de <i>Ocotea macrophylla</i> (Lauraceae) y evaluación de actividad biológica
QPN0336	<chem>C(C)(C)[C@H]1[C@@]2(C)[C@@H](C(C)CC1)CCC(C)=C2</chem>	Lauraceae	
QPN0337	<chem>CC1(C)C2C=C(C)/C=C\C=C(\C)/CCC12</chem>	Lauraceae	
QPN0338	<chem>O(C)c1c(OC)c2c3[C@@H]([N+](=C/O)/CC2)Cc2c(-c3c1O)cc1OCOc1c2</chem>	Lauraceae	
QPN0339	<chem>O(C)c1c(OC)c2c3[C@@H](NCC2)Cc2c(-c3c1O)cc1OCOc1c2</chem>	Lauraceae	
QPN0340	<chem>O=C(OC)/C=C(\C=C\C1(O)C2(C)OCC1(C)CC(O)C2)/C</chem>	Lauraceae	
QPN0341	<chem>O=C(OC)CCCCCCCCCCCCC</chem>	Lauraceae	
QPN0342	<chem>O=C(OCC)N1[C@@H]2c3c(c(O)c(OC)c(OC)c3CC1)-c1c(cc3OCOc3c1)C2</chem>	Lauraceae	
QPN0343	<chem>OC/C=C(/CCC=C(C)C)\C</chem>	Lauraceae	
QPN0344	<chem>OCCC(CCC=C(C)C)C</chem>	Lauraceae	
QPN0345	<chem>O[C@]1(C)C2[C@@H]3C(C)(C)[C@@H]3CCC(=C)C2CC1</chem>	Lauraceae	
QPN0346	<chem>O=C1N2C(=CC=C1)C1CN(C)CC(C2)C1</chem>	Fabaceae	Estudio químico y evaluación de la
QPN0347	<chem>O=C1N2C(=CC=C1)C1CN3C(C(C2)C1)CCCC3</chem>	Fabaceae	

QPN0348	<chem>O=C1N2C(=CC=C1)C1CNCC(C2)C1</chem>	Fabaceae	actividad antifúngica de alcaloides quinolozidínicos de <i>Genista monspessulana</i> (Fabaceae)
QPN0349	<chem>O=C1N2C(C=CC1)C1C3N(CCC1)CCCC3C2</chem>	Fabaceae	
QPN0350	<chem>O=C1N2C([C@H]3CN4C([C@H]1C3)CCCC4)CCCC2</chem>	Fabaceae	
QPN0351	<chem>O=CN1CC2C=3N(C(=O)C=CC=3)CC(C1)C2</chem>	Fabaceae	
QPN0352	<chem>O=C(OCCCCCCCCCCCCC)/C=C/c1cc(OC)c(O)cc1</chem>	Solanaceae	
QPN0353	<chem>C(Cc1cc2OCoc2cc1)C</chem>	Piperaceae	Fenilpropanoides de origen natural con actividad antialimentaria y repelente para el control de <i>T. castaneum</i> .
QPN0354	<chem>O(C)c1c(O)ccc(CCC)c1</chem>	Piperaceae	
QPN0355	<chem>O(C)c1c(OC)c2OCoc2cc1CC=C</chem>	Piperaceae	
QPN0356	<chem>O(C)c1c(OC)c2OCoc2cc1CCC</chem>	Piperaceae	
QPN0357	<chem>O(C)c1c(OC)ccc(CCC)c1</chem>	Piperaceae	
QPN0358	<chem>O(C)c1c2OCoc2cc(CC=C)c1</chem>	Piperaceae	
QPN0359	<chem>O(C)c1ccc(CCC)cc1</chem>	Piperaceae	
QPN0360	<chem>O=C(Oc1c(OC)cc(/C=C/C=O)cc1)C</chem>	Piperaceae	
QPN0361	<chem>O=C(Oc1c(OC)cc(CC=C)cc1)C</chem>	Piperaceae	
QPN0362	<chem>O=C(Oc1c(OC)cc(CCC)cc1)C</chem>	Piperaceae	
QPN0363	<chem>O=C/C=C/c1c(OC)c(OC)c2OCoc2c1</chem>	Piperaceae	
QPN0364	<chem>O=C/C=C/c1cc(OC)c2OCoc2c1</chem>	Piperaceae	
QPN0365	<chem>O=C/C=C/c1ccc(OC)cc1</chem>	Piperaceae	
QPN0366	<chem>O=C(/C=C(/CC/C=C(\C)/C)\C)c1c(O)ccc(O)c1</chem>	Piperaceae	Macroproyecto: bioprospección de especies del género piper para el control de agentes
QPN0367	<chem>O=C(CC(O)(CC/C=C(\C)/C)C)c1c(O)ccc(O)c1</chem>	Piperaceae	
QPN0368	<chem>O(C)c1c(C)cc(C2=C(OC3C(O)C(O)C(O)C(C)O3)C(=O)c3c(O)cc(O)cc3O2)cc1</chem>	Lauraceae	Metabolitos secundarios aislados de hojas de <i>Ocotea heterochroma</i> (Lauraceae)
QPN0369	<chem>O(C)c1c(OC)cc([C@@H]2[C@@H](C)[C@@]3(OC)C(=C(O)C(=O)[C@H](CC=C)C3)O2)cc1OC</chem>	Lauraceae	
QPN0370	<chem>O(C)c1c2OCoc2cc(C2[C@@H](C)[C@@]3(OC)C(O2)=CC(=O)C(CC=C)=C3)c1</chem>	Lauraceae	
QPN0371	<chem>O(C)c1c2c(cc(C3[C@@H](C)[C@]4(OC)C(O)[C@@]3(O)C(=O)C(CC=C)=C4)c1)CCC2</chem>	Lauraceae	
QPN0372	<chem>O=C(OC1C(O)C(O)C(CO)OC1OC1=C(C2ccc(O)cc2)Oc2c(c(O)cc(O)c2)C1=O)/C=C/c1ccc(O)cc1</chem>	Lauraceae	
QPN0373	<chem>O=C(OCC)CCc1ccc(O)cc1</chem>	Lauraceae	
QPN0374	<chem>O(C)C=1C(=O)C=C2[C@@](CC=C)([C@H](C)[C@H](c3cc4OCoc4cc3)O2)C=1</chem>	Lauraceae	
QPN0375	<chem>O(C)C=1C(=O)[C@H]2[C@@H]([C@@H](C)[C@](CC=C)(C2=O)C=1)c1cc2OCoc2cc1</chem>	Lauraceae	

QPN0376	<chem>O(C)[C@@]1(CC=C)C(=O)C=C2O[C@H]([C@@H](C)C2=C1)c1cc2OCOc2cc1</chem>	Lauraceae	corteza de <i>Mezilaurus itauba</i> (Meissa.) Taubert ex Mez (Lauraceae)
QPN0377	<chem>O(C)c1c(O)c(CC=C)c2O[C@@H]([C@@H](C)c2c1)c1cc2OCOc2cc1</chem>	Lauraceae	
QPN0378	<chem>O(C)c1c(O)c(CC=C)c2oc(c(C)c2c1)-c1cc2OCOc2cc1</chem>	Lauraceae	
QPN0379	<chem>O(C)c1c(O)ccc([C@@H]2[C@@H](C)[C@]3(CC=C)C(=O)[C@H]2C(=O)C(OC)=C3)c1</chem>	Lauraceae	
QPN0380	<chem>O(CC=C)c1c(OC)cc2[C@H](C)[C@@H](c3cc4OCOc4cc3)Oc2c1</chem>	Lauraceae	
QPN0381	<chem>O=C1[C@@H](O)[C@@H]2[C@H]([C@H](C)[C@@](CC=C)(C2O)C1)c1cc2OCOc2cc1</chem>	Lauraceae	
QPN0382	<chem>O=C(c1ccccc1)C1C(=O)OCCC1</chem>	Rutaceae	Nueva benzoilactona y otros constituyentes de <i>Pilocarpus alvaradonii</i>
QPN0383	<chem>O=C1Oc2c(cc3c(occ3)c2)C=C1</chem>	Rutaceae	
QPN0384	<chem>OCC12C(C(C(=C)C)CC1)C1C(C)(C3(C)C(C4(C)C(C(C)(C)[C@H](O)CC4)CC3)CC1)CC2</chem>	Rutaceae	
QPN0385	<chem>O(C)C1=CC2OC(c3cc(O)c(O)cc3)=CC(=O)C2C=C1</chem>	Myristicaceae	Perfilado metabólico de plantas del genero <i>Virola</i> Spp (Myristicaceae) provenientes de la Amazonia y Orinoquia Colombiana y evaluacion de la actividad antifungica contra <i>Fusarium oxysporum</i>
QPN0386	<chem>O(C)c1c(O)cc(C=2Oc3c(C(=C)C=2)ccc(OC)c3)cc1</chem>	Myristicaceae	
QPN0387	<chem>O(C)c1c(O)ccc(C2CC(=O)c3c(cc(OC)cc3)C2)c1</chem>	Myristicaceae	
QPN0388	<chem>O(C)c1cc2OC(c3cc4OCOc4cc3)=CC(=O)c2cc1</chem>	Myristicaceae	
QPN0389	<chem>O=C(O)c1cc(C/C=C(/CC/C=C(\C)/C)\C)c(OC)cc1</chem>	Piperaceae	
QPN0390	<chem>O(C)c1c2c(c(OC)ccc2)nc2occc12</chem>	Rutaceae	Aislamiento y caracterización de alcaloides flavonoides y terpenos del extracto etanolico de la corteza de <i>Esenbeckia runyonii</i>
QPN0391	<chem>O(C)c1c2c(nc3occc13)cc(OC)c(OC)c2</chem>	Rutaceae	
QPN0392	<chem>O(C)c1cc(O)c2C(=O)c3c(N(C)c2c1)cccc3</chem>	Rutaceae	
QPN0393	<chem>C=C1C2C(C(C)(C)C2)CCC2(C)OC2CC1</chem>	Piperaceae	Contribución al estudio fitoquímico de la parte aérea de <i>Piper</i> cf. <i>Cumanense</i> Kunth
QPN0394	<chem>O=C(O)c1cc(C(=O)/C=C(/CCC=C(C)C)\C)c(O)c(CC=C(C)C)c1</chem>	Piperaceae	
QPN0395	<chem>O=C(O)c1cc(C/C=C(/CCC=C(C)C)\C)c2OC(C)(C)C=Cc2c1</chem>	Piperaceae	

122 Contribución a la identificación de potenciales antagonistas asociados al quorum sensing en *Pseudomonas aeruginosa* a través de métodos quimioinformáticos.

QPN0396	<chem>O=C(NCC(C)C)/C=C/C=C/CCC</chem>	Piperaceae	Contribución al estudio químico de la madera de la <i>Piper yanaconasense</i>
QPN0397	<chem>O=C1C(C(=C)C)C=CC(C)=C1</chem>	Piperaceae	
QPN0398	<chem>O(C)C=1C(=O)Oc2c3c(occ3)ccc2C=1</chem>	Piperaceae	Cumarinas aisladas de <i>Pilocarpus</i> Sp. Rutaceae
QPN0399	<chem>O(C)c1c2c(OC(=O)C=C2)c2c(occ2)c1</chem>	Piperaceae	
QPN0400	<chem>O=C1Oc2c3c(occ3)cc(C)c2C=C1</chem>	Piperaceae	
QPN0401	<chem>O(C)c1c2occc2c2[nH]c3c(c2c1)cccc3</chem>	Rutaceae	Estudio de alcaloides lignanos y terpenos de los extractos etanolicos de corteza y madera de <i>Zanthoxylum caribaeum</i>
QPN0402	<chem>o1c2c(c3[nH]c4c(c3cc2)cccc4)cc1</chem>	Rutaceae	
QPN0403	<chem>O(C)C=1c2c(N(C)C(=O)C=1)cccc2</chem>	Rutaceae	Estudio de metabolitos secundarios (alcaloides y cumarinas) en hojas y corteza de dos especies de la familia rutaceae <i>Esenbeckia alata</i> y <i>Raputia heptaphylla</i> síntesis de análogos de alguno de los alcaloides aislados.
QPN0404	<chem>O(C)c1c2N(C)C(=O)C3=C(OC(C)(C)C=C3)c2ccc1</chem>	Rutaceae	
QPN0405	<chem>O(C)c1c2c(n(C)c1)cccc2</chem>	Rutaceae	
QPN0406	<chem>O(C)c1c2c(occ2)cc2OC(=O)C=Cc12</chem>	Rutaceae	
QPN0407	<chem>O(C)c1c2occc2cc2c1OC(=O)C=C2</chem>	Rutaceae	
QPN0408	<chem>O(c1cccc1)C=1N(C)c2c(C(=O)C=1)cccc2</chem>	Rutaceae	
QPN0409	<chem>O=C(NCCc1cccc1)/C=C/c1cc2OCOc2cc1</chem>	Rutaceae	
QPN0410	<chem>O=C(O)C12C(C(C(=C)C)CC1)C1C(C)(C3(C)C(C4(C)C(C(C)(C)C(O)CC4)=CC3)CC1)CC2</chem>	Rutaceae	
QPN0411	<chem>O=C(OC)/C=C\c1c(C(=O)CC)nc2occc2c1</chem>	Rutaceae	
QPN0412	<chem>O=C(OC1[C@@]2(C)C34OC3C(=O)O[C@@H](c3cocc3)C4C[C@@H](O)C2C2([C@H](O)OC(=O)CC2O)C(C(O)(C)C)C1)C</chem>	Rutaceae	
QPN0413	<chem>O=C1OC(c2occc2)C2(C)C3(C4(C)C(=O)CC(C(O)(C)C)C5(C(O)CC(=O)OC5)C4CC2)OC13</chem>	Rutaceae	
QPN0414	<chem>O=C1Oc2c(cc3c(OC(C)(C)C=C3)c2)C=C1</chem>	Rutaceae	
QPN0415	<chem>O=C1Oc2c(ccc(O)c2)C=C1</chem>	Rutaceae	
QPN0416	<chem>O=C1Oc2c(cccc2)C=C1</chem>	Rutaceae	
QPN0417	<chem>O=C1c2c3c(oc2N(C)c2c1cccc2)cccc3</chem>	Rutaceae	
QPN0418	<chem>O[C@@H]1C(C)(C)[C@H]2[C@@](C)([C@@H]3[C@](C)([C@@]4(C)[C@@H]([C@H]5[C@H](C(=C)C)CC[C@]5(C)CC4)CC3)CC2)CC1</chem>	Rutaceae	
QPN0419	<chem>O(C)C=1C(=O)Oc2c(C=1CC=C(C)C)cccc2</chem>	Rutaceae	Estudio de terpenos
QPN0420	<chem>O(C)c1c(O)cc2n(C)c3c(c2c1)cccc3</chem>	Rutaceae	

QPN0421	<chem>O(C)c1c2c(occ2)nc2c1cccc2</chem>	Rutaceae	cumarinas y lignanos del extracto etanolico de la madera de <i>Zanthoxylum rhoifolium</i>
QPN0422	<chem>O=C1N(C)c2c(cccc2)C=2OC(C)(C)C=CC1=2</chem>	Rutaceae	
QPN0423	<chem>O(C)c1c(OC)cc2c(OC(=O)C=C2)c1</chem>	Rutaceae	Estudio fitoquímico de los extractos etanolicos de cuatro especies de la familia rutaceae (<i>Hortia colombiana</i> , <i>Zanthoxylum culantrillo</i> , <i>Zanthoxylum monophyllum</i> y <i>Peltostigma guatemalensis</i>)
QPN0424	<chem>O(C)c1c(OC)ccc([C@@H]2OCC3[C@@H](c4cc5OCC5cc4)OCC23)c1</chem>	Rutaceae	
QPN0425	<chem>O(C)c1c(OC)ccc([C@H]2[C@H](CO)[C@H](Cc3cc4OCCc4cc3)CO2)c1</chem>	Rutaceae	
QPN0426	<chem>O(C)c1c2c(OC(C)(C)C=C2)c2c(OC(=O)C=C2)c1</chem>	Rutaceae	
QPN0427	<chem>O(C)c1cc2c([nH]c3c2CCN2C(=O)c4c(N=C32)cccc4)cc1</chem>	Rutaceae	
QPN0428	<chem>O(C)c1nc2c(c(OC)c1)cccc2</chem>	Rutaceae	
QPN0429	<chem>O(CC(O)C(O)(C)C)c1c(OC)c2nc3occc3c(OC)c2cc1</chem>	Rutaceae	
QPN0430	<chem>O(CC1C(C)(C)O1)c1c(OC)c2nc3occc3c(OC)c2cc1</chem>	Rutaceae	
QPN0431	<chem>O(CC=C(C)C)c1c(OC)c2nc3occc3c(OC)c2cc1</chem>	Rutaceae	
QPN0432	<chem>O=C(C)c1c(OC)cc(O)cc1O</chem>	Rutaceae	
QPN0433	<chem>O=C(C)c1c(OC)cc(OC)cc1O</chem>	Rutaceae	
QPN0434	<chem>O=C(NCCc1ccc(O)cc1)/C(=C/C)/C</chem>	Rutaceae	
QPN0435	<chem>O=C(NCCc1ccc(O)cc1)CCCCCCCC=CCCCCCCC</chem>	Rutaceae	
QPN0436	<chem>O=C(NCCc1ccc(OCC(O)C(O)(C)C)cc1)/C(=C/C)/C</chem>	Rutaceae	
QPN0437	<chem>O=C(O)CCc1c(OC)c2OC(C)(C)C=Cc2c(OC)c1</chem>	Rutaceae	
QPN0438	<chem>O=C(O)CCc1c(OC)c2c(OC(C)(C)C=C2)cc1OC</chem>	Rutaceae	
QPN0439	<chem>O=C(O)Cc1cccc1</chem>	Rutaceae	
QPN0440	<chem>O=C(OC)CC1C(C)(C)C(=O)C2OC32C1(C)C=1C(C)(O3)C=2C(C)(C(c3occc3)OC(=O)C=2)CC=1</chem>	Rutaceae	
QPN0441	<chem>O=C(OC)CC1C(C)(C)C(=O)C=C2C1(C)C1C(C)(O2)C=2C(C)(C(c3occc3)OC(=O)C=2)CC1</chem>	Rutaceae	
QPN0442	<chem>O=C(OC)CCc1c(OC)c2OC(C)(C)C=Cc2c(OC)c1</chem>	Rutaceae	
QPN0443	<chem>O=C(OC)CCc1c(OC)c2c(OC(C)(C)C=C2)cc1</chem>	Rutaceae	
QPN0444	<chem>O=C(OC)CCc1c(OC)c2c(OC(C)(C)C=C2)cc1OC</chem>	Rutaceae	
QPN0445	<chem>O=C(OC)CCc1c(OC)cc2occc2c1</chem>	Rutaceae	
QPN0446	<chem>O=C(OC)c1ccc(O)cc1</chem>	Rutaceae	
QPN0447	<chem>O=C1N2C(=Nc3c1cccc3)c1[nH]c3c(c1CC2)cccc3</chem>	Rutaceae	
QPN0448	<chem>O=C1Nc2c(cccc2)C=2OC(C)(C)C=CC1=2</chem>	Rutaceae	

QPN0449	<chem>O=C1Oc2c(C/C=C(\CO)/C)c(O)ccc2C=C1</chem>	Rutaceae	
QPN0450	<chem>O=C1Oc2c(CC=C(C)C)c(O)ccc2C=C1</chem>	Rutaceae	
QPN0451	<chem>OC(C=C)(CC/C=C(/CCC1C(C)(C)O1)\C)C</chem>	Rutaceae	
QPN0452	<chem>OC(C=C)(CC/C=C(/CCC=C(C)C)\C)C</chem>	Rutaceae	
QPN0453	<chem>OC1[C@@H](Cc2cc3OCOc3cc2)[C@H](Cc2cc3OC Oc3cc2)CO1</chem>	Rutaceae	
QPN0454	<chem>O(C)c1cc(O)c(C=2C(=O)c3c(O)cc(O)cc3OC=2)cc1</chem>	Myristicaceae	Estudio fitoquímico del extracto etanólico de la madera de <i>Virola carinata</i>
QPN0455	<chem>O(C)c1ccc(C=2C(=O)c3c(O)cc(O)cc3OC=2)cc1</chem>	Myristicaceae	
QPN0456	<chem>O(C)c1c2c(OC(=O)C=C2)c2c(OC(C)(C)C=C2)c1</chem>	Rutaceae	Estudio fitoquímico parcial del extracto clorofórmico de la corteza de <i>Peltostigma guatemalense</i>
QPN0457	<chem>O=C(NCCc1ccc(OCC=C(C)C)cc1)/C(=C/C)/C</chem>	Rutaceae	
QPN0458	<chem>O=C(O)CCc1c(OC)c2c(OC(C)(C)C=C2)cc1</chem>	Rutaceae	
QPN0459	<chem>O(C)c1c(O)cc2OC(=O)C=Cc2c1</chem>	Rutaceae	Estudio fitoquímico y de actividad biológica (antimalarica y antimicrobiana) de los extractos etanólicos de hoja corteza y madera de <i>Esenbeckia littoralis</i>
QPN0460	<chem>O(C)c1c(OC)c(O)c2C(=O)C=C(c3ccc(OC)cc3)Oc2c1OC</chem>	Rutaceae	
QPN0461	<chem>O(C)c1c2c(c(OC)c3OCOc3c2)nc2occc12</chem>	Rutaceae	
QPN0462	<chem>O(C)c1c2c(c(OC)cc(OC)c2)nc2occc12</chem>	Rutaceae	
QPN0463	<chem>O(C)c1c2c(nc3occc13)cc1OCOc1c2</chem>	Rutaceae	
QPN0464	<chem>O(C)c1c(OC)ccc2c(OC)c3c(occ3)nc12</chem>	Rutaceae	Estudio químico de algunos de los metabolitos secundarios presentes en el extracto etanólico de madera de <i>Raputia heptaphylla</i>
QPN0465	<chem>O=Cc1cc(OC)c(O)c(OC)c1</chem>	Rutaceae	

E.Anexo: Resultados del cribado virtual QuiProNaB

Mejores resultados por el modelo de cribado con posible actividad sobre LasR

ID	k-NN	Xgb	Aminoácidos										Puntaje
			115	40	75	127	47	105	80	36	52	56	
QPN0162	Activo	Activo	-1	0	-1	-1	1	1	1	1	1	1	3
QPN0264	Activo	Activo	0	-1	0	-1	1	0	0	1	1	1	2
QPN0282	Activo	Activo	-1	0	-1	-1	1	0	1	1	1	1	2
QPN0319	Activo	Activo	0	-1	-1	-1	1	1	0	1	1	1	2
QPN0056	Activo	Activo	-1	0	-1	-1	1	1	0	1	0	1	1
QPN0468	Activo	Activo	-1	-1	-1	-1	1	1	0	1	1	1	1
QPN0205	Activo	Activo	0	-1	-1	-1	0	1	0	1	0	1	0
QPN0465	Activo	Activo	-1	0	-1	-1	0	1	0	1	0	1	0
QPN0216	Activo	Activo	-1	-1	-1	-1	0	1	0	0	1	1	-1
QPN0483	Activo	Activo	-1	-1	-1	-1	1	0	0	1	0	1	-1

Mejores resultados por el modelo de cribado con posible actividad sobre PqsR

ID	k-NN	Xgb	Aminoácidos									Puntaje
			196	197	194	209	258	221	102	207	149	
QPN0458	Activo	Activo	0	-1	0	-1	1	1	1	1	1	3
QPN0162	Activo	Activo	-1	-1	-1	-1	1	1	0	1	1	0
QPN0489	Activo	Activo	-1	-1	-1	-1	0	1	0	1	1	-1

Mejores resultados por el modelo de cribado con posible actividad sobre PqsR

ID	k-NN	Xgb	Aminoácidos											Puntaje
			84	107	101	121	133	135	64	60	69	72	83	
QPN0007	1	1	0	0	0	0	0	-1	1	1	1	1	1	4
QPN0008	1	1	0	0	0	0	0	-1	1	1	1	1	1	4
QPN0051	1	1	0	0	0	0	0	-1	1	1	1	1	1	4
QPN0192	1	1	-1	0	0	0	0	0	1	1	1	1	1	4
QPN0072	1	1	0	0	0	0	-1	-1	1	1	1	1	1	3
QPN0462	1	1	-1	0	-1	0	0	0	1	1	1	1	1	3
QPN0381	1	1	0	0	-1	0	0	-1	1	1	1	1	1	3
QPN0457	1	1	0	0	-1	0	0	-1	1	1	1	1	1	3
QPN0473	1	1	0	-1	0	0	0	-1	1	1	1	1	1	3
QPN0386	1	1	0	0	-1	0	0	-1	1	1	1	1	1	3
QPN0458	1	1	0	0	-1	0	0	-1	1	1	1	1	1	3
QPN0174	1	1	0	0	0	0	-1	-1	1	1	1	1	1	3
QPN0181	1	1	-1	0	-1	0	0	0	1	1	1	1	1	3
QPN0314	1	1	-1	0	0	0	0	-1	1	1	1	1	1	3
QPN0394	1	1	0	0	0	0	0	-1	1	1	0	1	1	3
QPN0342	1	1	-1	0	0	0	0	-1	1	1	1	1	1	3
QPN0486	1	1	-1	0	0	0	0	0	1	1	1	1	0	3
QPN0014	1	1	-1	0	0	0	0	-1	1	1	1	1	1	3
QPN0094	1	1	-1	0	0	0	0	0	1	1	1	1	0	3
QPN0284	1	1	-1	0	0	0	0	-1	1	1	1	1	1	3
QPN0010	1	1	0	0	0	0	0	-1	1	1	0	1	1	3
QPN0481	1	1	-1	0	0	0	0	0	1	1	1	1	0	3
QPN0112	1	1	-1	0	0	0	0	0	1	1	1	1	0	3
QPN0504	1	1	-1	0	-1	0	0	0	1	1	0	1	1	2
QPN0115	1	1	0	-1	-1	0	0	-1	1	1	1	1	1	2
QPN0505	1	1	0	-1	-1	0	0	-1	1	1	1	1	1	2
QPN0502	1	1	0	0	0	0	-1	-1	1	1	1	1	0	2
QPN0135	1	1	-1	0	-1	0	0	-1	1	1	1	1	1	2
QPN0136	1	1	-1	0	-1	0	0	-1	1	1	1	1	1	2
QPN0137	1	1	-1	0	-1	0	0	-1	1	1	1	1	1	2
QPN0109	1	1	0	-1	0	-1	0	-1	1	1	1	1	1	2
QPN0168	1	1	-1	0	0	0	-1	-1	1	1	1	1	1	2
QPN0150	1	1	-1	0	-1	0	0	-1	1	1	1	1	1	2
QPN0151	1	1	0	-1	-1	0	0	-1	1	1	1	1	1	2
QPN0152	1	1	0	-1	-1	0	0	-1	1	1	1	1	1	2
QPN0500	1	1	0	0	-1	0	0	0	1	0	0	1	1	2
QPN0498	1	1	0	0	-1	0	0	0	1	0	0	1	1	2
QPN0169	1	1	-1	0	-1	0	0	-1	1	1	1	1	1	2
QPN0180	1	1	0	0	-1	0	-1	-1	1	1	1	1	1	2
QPN0188	1	1	-1	0	0	0	-1	0	1	1	1	1	0	2

QPN0190	1	1	-1	0	0	0	-1	-1	1	1	1	1	1	2
QPN0490	1	1	0	-1	-1	0	0	-1	1	1	1	1	1	2
QPN0307	1	1	0	-1	-1	0	0	-1	1	1	1	1	1	2
QPN0245	1	1	0	-1	0	0	0	-1	1	1	1	1	0	2
QPN0283	1	1	-1	0	0	0	-1	-1	1	1	1	1	1	2
QPN0111	1	1	0	-1	-1	0	0	-1	1	1	1	1	1	2
QPN0282	1	1	-1	0	-1	0	0	-1	1	1	1	1	1	2
QPN0379	1	1	0	-1	-1	0	0	-1	1	1	1	1	1	2
QPN0456	1	1	0	-1	-1	0	0	-1	1	1	1	1	1	2
QPN0535	1	1	-1	0	0	0	0	-1	1	1	0	1	1	2
QPN0027	1	1	-1	0	0	0	-1	-1	1	1	1	1	1	2
QPN0522	1	1	-1	0	0	0	-1	-1	1	1	1	1	1	2
QPN0025	1	1	-1	0	0	0	-1	-1	1	1	1	1	1	2
QPN0383	1	1	0	-1	-1	0	0	-1	1	1	1	1	1	2
QPN0050	1	1	-1	0	0	0	-1	-1	1	1	1	1	1	2
QPN0515	1	1	-1	0	0	-1	0	-1	1	1	1	1	1	2
QPN0399	1	1	-1	0	0	0	-1	-1	1	1	1	1	1	2
QPN0015	1	1	-1	0	0	0	-1	-1	1	1	1	1	1	2
QPN0053	1	1	-1	0	-1	0	0	-1	1	1	1	1	1	2
QPN0055	1	1	-1	0	-1	0	0	-1	1	1	1	1	1	2
QPN0056	1	1	-1	0	-1	0	0	-1	1	1	1	1	1	2
QPN0453	1	1	0	-1	-1	0	0	-1	1	1	1	1	1	2
QPN0539	1	1	0	0	-1	0	0	-1	1	0	1	1	1	2
QPN0059	1	1	0	-1	-1	0	0	-1	1	1	1	1	1	2
QPN0066	1	1	0	-1	-1	0	0	-1	1	1	1	1	1	2
QPN0067	1	1	-1	0	-1	0	0	-1	1	1	1	1	1	2
QPN0540	1	1	-1	0	-1	0	0	-1	1	1	1	1	1	2
QPN0513	1	1	0	0	0	0	0	-1	1	0	1	1	0	2
QPN0451	1	1	-1	0	0	0	-1	-1	1	1	1	1	1	2
QPN0514	1	1	0	-1	-1	0	0	-1	1	1	1	1	1	2
QPN0085	1	1	-1	0	0	0	0	-1	1	1	0	1	1	2
QPN0086	1	1	0	0	-1	0	-1	-1	1	1	1	1	1	2
QPN0087	1	1	-1	-1	0	0	0	-1	1	1	1	1	1	2
QPN0406	1	1	-1	0	-1	0	-1	-1	1	1	1	1	1	1
QPN0397	1	1	-1	0	0	0	-1	-1	1	1	0	1	1	1
QPN0469	1	1	-1	0	-1	0	-1	-1	1	1	1	1	1	1
QPN0404	1	1	-1	-1	-1	0	0	-1	1	1	1	1	1	1
QPN0299	1	1	-1	0	-1	0	-1	-1	1	1	1	1	1	1
QPN0287	1	1	-1	0	0	-1	-1	-1	1	1	1	1	1	1
QPN0368	1	1	-1	0	-1	0	-1	-1	1	1	1	1	1	1
QPN0329	1	1	-1	0	-1	0	-1	-1	1	1	1	1	1	1
QPN0321	1	1	-1	0	-1	0	-1	-1	1	1	1	1	1	1
QPN0390	1	1	0	-1	-1	0	-1	-1	1	1	1	1	1	1
QPN0322	1	1	-1	0	-1	0	-1	-1	1	1	1	1	1	1
QPN0389	1	1	0	-1	-1	-1	0	-1	1	1	1	1	1	1

QPN0367	1	1	-1	0	0	-1	-1	-1	1	1	1	1	1	1
QPN0332	1	1	-1	0	-1	0	-1	-1	1	1	1	1	1	1
QPN0466	1	1	0	0	-1	0	0	-1	1	0	0	1	1	1
QPN0336	1	1	-1	0	-1	0	-1	-1	1	1	1	1	1	1
QPN0346	1	1	-1	0	0	0	-1	-1	1	0	1	1	1	1
QPN0370	1	1	0	-1	-1	0	0	-1	1	1	0	1	1	1
QPN0353	1	1	-1	0	0	0	-1	-1	1	0	1	1	1	1
QPN0356	1	1	-1	0	-1	0	-1	-1	1	1	1	1	1	1
QPN0369	1	1	0	-1	-1	0	0	-1	1	1	0	1	1	1
QPN0004	1	1	-1	-1	-1	0	0	-1	1	1	1	1	1	1
QPN0156	1	1	-1	-1	-1	0	0	-1	1	1	1	1	1	1
QPN0474	1	1	-1	0	0	0	0	-1	1	1	0	1	0	1
QPN0516	1	1	0	-1	-1	0	0	-1	1	1	0	1	1	1
QPN0510	1	1	-1	0	-1	0	-1	-1	1	1	1	1	1	1
QPN0512	1	1	-1	-1	-1	0	0	-1	1	1	1	1	1	1
QPN0068	1	1	-1	0	-1	0	-1	-1	1	1	1	1	1	1
QPN0065	1	1	-1	0	-1	0	-1	-1	1	1	1	1	1	1
QPN0058	1	1	0	-1	0	-1	-1	-1	1	1	1	1	1	1
QPN0057	1	1	-1	0	-1	0	-1	-1	1	1	1	1	1	1
QPN0054	1	1	-1	0	0	0	-1	-1	1	1	0	1	1	1
QPN0518	1	1	0	-1	-1	0	0	-1	1	1	0	1	1	1
QPN0503	1	1	-1	0	-1	0	0	0	1	0	0	1	1	1
QPN0044	1	1	-1	-1	-1	0	0	-1	1	1	1	1	1	1
QPN0530	1	1	-1	0	-1	0	-1	-1	1	1	1	1	1	1
QPN0026	1	1	-1	-1	-1	0	0	0	1	1	1	1	0	1
QPN0023	1	1	-1	-1	-1	0	0	-1	1	1	1	1	1	1
QPN0013	1	1	0	0	-1	0	0	0	1	0	0	0	1	1
QPN0012	1	1	0	0	-1	0	0	0	1	0	0	0	1	1
QPN0009	1	1	-1	0	0	0	-1	-1	1	1	1	1	0	1
QPN0110	1	1	0	-1	-1	-1	0	-1	1	1	1	1	1	1
QPN0448	1	1	0	-1	-1	0	-1	-1	1	1	1	1	1	1
QPN0493	1	1	-1	0	-1	0	0	-1	1	1	0	1	1	1
QPN0233	1	1	0	-1	-1	-1	0	-1	1	1	1	1	1	1
QPN0134	1	1	-1	0	-1	0	-1	-1	1	1	1	1	1	1
QPN0146	1	1	-1	0	-1	0	-1	-1	1	1	1	1	1	1
QPN0147	1	1	-1	0	-1	0	-1	-1	1	1	1	1	1	1
QPN0153	1	1	0	-1	-1	-1	0	-1	1	1	1	1	1	1
QPN0155	1	1	0	-1	-1	-1	0	-1	1	1	1	1	1	1
QPN0450	1	1	0	0	0	-1	0	-1	1	0	0	1	1	1
QPN0161	1	1	-1	-1	-1	0	0	-1	1	1	1	1	1	1
QPN0494	1	1	0	0	-1	0	0	-1	1	0	0	1	1	1
QPN0484	1	1	-1	-1	-1	0	0	-1	1	1	1	1	1	1
QPN0186	1	1	0	-1	-1	0	-1	-1	1	1	1	1	1	1
QPN0244	1	1	0	-1	-1	-1	0	-1	1	1	1	1	1	1

QPN0235	1	1	0	-1	-1	-1	0	-1	1	1	1	1	1	1
QPN0477	1	1	-1	-1	-1	0	-1	-1	1	1	1	1	1	0
QPN0452	1	1	-1	0	0	0	-1	-1	1	1	0	0	1	0
QPN0501	1	1	0	-1	-1	0	0	-1	1	0	0	1	1	0
QPN0538	1	1	0	-1	-1	0	0	-1	1	0	0	1	1	0
QPN0536	1	1	-1	-1	-1	0	0	-1	1	1	0	1	1	0
QPN0489	1	1	0	-1	-1	0	0	-1	1	0	0	1	1	0
QPN0475	1	1	-1	-1	-1	0	-1	-1	1	1	1	1	1	0
QPN0459	1	1	0	-1	-1	0	-1	-1	1	1	0	1	1	0
QPN0509	1	1	-1	0	-1	-1	-1	-1	1	1	1	1	1	0
QPN0541	1	1	-1	-1	-1	0	0	-1	1	1	0	1	1	0
QPN0449	1	1	-1	-1	-1	0	-1	-1	1	1	1	1	1	0
QPN0213	1	1	-1	0	0	0	-1	-1	0	1	0	1	1	0
QPN0403	1	1	-1	0	-1	0	-1	-1	1	1	0	1	1	0
QPN0130	1	1	-1	-1	-1	0	0	-1	1	1	0	1	1	0
QPN0129	1	1	-1	0	-1	0	-1	-1	1	1	0	1	1	0
QPN0122	1	1	-1	-1	-1	0	0	-1	1	1	0	1	1	0
QPN0096	1	1	-1	-1	-1	0	-1	-1	1	1	1	1	1	0
QPN0062	1	1	-1	-1	-1	0	-1	-1	1	1	1	1	1	0
QPN0052	1	1	-1	0	-1	-1	0	-1	1	1	0	1	1	0
QPN0049	1	1	-1	-1	-1	0	0	-1	1	1	0	1	1	0
QPN0045	1	1	-1	0	-1	0	-1	-1	1	1	0	1	1	0
QPN0037	1	1	-1	0	-1	0	-1	-1	1	1	0	1	1	0
QPN0036	1	1	-1	0	-1	0	-1	-1	1	1	0	1	1	0
QPN0033	1	1	-1	-1	-1	0	-1	-1	1	1	1	1	1	0
QPN0031	1	1	-1	-1	-1	0	-1	-1	1	1	1	1	1	0
QPN0028	1	1	-1	-1	-1	0	-1	-1	1	1	1	1	1	0
QPN0006	1	1	-1	-1	-1	0	-1	-1	1	1	1	1	1	0
QPN0176	1	1	0	-1	-1	0	-1	-1	1	1	0	1	1	0
QPN0162	1	1	-1	-1	-1	-1	0	-1	1	1	1	1	1	0
QPN0232	1	1	-1	0	-1	0	-1	-1	1	1	1	1	0	0
QPN0281	1	1	-1	-1	-1	0	-1	-1	1	1	1	1	1	0
QPN0373	1	1	-1	-1	-1	0	-1	-1	1	1	1	1	1	0
QPN0371	1	1	-1	-1	-1	0	0	-1	1	1	0	1	1	0
QPN0339	1	1	-1	-1	-1	0	-1	-1	1	1	1	1	1	0
QPN0338	1	1	-1	-1	-1	0	-1	-1	1	1	1	1	1	0
QPN0335	1	1	0	-1	-1	-1	-1	-1	1	1	1	1	1	0
QPN0312	1	1	-1	-1	-1	0	-1	-1	1	1	1	1	1	0
QPN0278	1	1	0	-1	-1	0	0	-1	1	1	0	1	0	0
QPN0005	1	1	-1	-1	-1	0	-1	-1	1	1	1	1	1	0
QPN0236	1	1	-1	-1	-1	0	-1	-1	1	1	1	1	1	0
QPN0277	1	1	0	0	-1	0	-1	-1	1	1	0	1	0	0
QPN0042	1	1	-1	0	-1	-1	-1	-1	1	1	0	1	1	-1
QPN0398	1	1	-1	-1	-1	-1	-1	-1	1	1	1	1	1	-1
QPN0392	1	1	-1	-1	-1	0	-1	-1	1	1	0	1	1	-1

QPN0344	1	1	-1	-1	-1	0	-1	-1	1	1	0	1	1	-1
QPN0160	1	1	-1	-1	-1	-1	-1	-1	1	1	1	1	1	-1
QPN0114	1	1	-1	-1	-1	0	-1	-1	1	1	0	1	1	-1
QPN0113	1	1	0	-1	-1	0	-1	-1	1	0	0	1	1	-1
QPN0276	1	1	0	-1	-1	0	-1	-1	1	0	0	1	1	-1
QPN0123	1	1	-1	-1	-1	0	0	-1	1	1	0	1	0	-1
QPN0183	1	1	-1	-1	-1	-1	-1	-1	1	1	1	1	1	-1
QPN0102	1	1	-1	-1	-1	0	-1	-1	1	1	0	1	1	-1
QPN0103	1	1	-1	-1	-1	0	-1	-1	1	1	0	1	1	-1
QPN0275	1	1	0	-1	-1	0	-1	-1	1	0	0	1	1	-1
QPN0249	1	1	-1	-1	-1	0	-1	-1	1	1	0	0	1	-2

Bibliografía

1. Schopf, J. W. Fossil evidence of Archaean life. *Philos. Trans. R. Soc. B Biol. Sci.* 361, 869–885 (2006).
2. Altermann, W. & Kazmierczak, J. Archean microfossils: A reappraisal of early life on Earth. *Res. Microbiol.* 154, 611–617 (2003).
3. Koch, R. Die Ätiologie der Milzbrand-Krankheit, begründet auf die Entwicklungsgeschichte des Bacillus Anthracis. *Dtsch. Medizinische Wochenschrift* 8, 283 (1882).
4. Fredricks, D. N. & Relman, D. A. Sequence-based identification of microbial pathogens: A **reconsideration of Koch's postulates**. *Clin. Microbiol. Rev.* 9, 18–33 (1996).
5. Andrei, S., Valeanu, L., Chirvasuta, R. & Stefan, M.-G. New FDA approved antibacterial drugs: 2015-2017. *FDA Approv. Antibact. Drugs 2015–2017*.
6. Aminov, R. I. A brief history of the antibiotic era: Lessons learned and challenges for the future. *Front. Microbiol.* 1, (2010).
7. Ventola, C. L. [Review] The antibiotic resistance crisis: part 1: causes and threats. *Pharm. Ther.* 40, 277–83 (2015).
8. Peña, C. et al. Carbapenem-resistant *Pseudomonas aeruginosa*: Factors influencing multidrug-resistant acquisition in non-critically ill patients. *Eur. J. Clin. Microbiol. Infect. Dis.* 28, 519–522 (2009).
9. El Zowalaty, M. E. et al. *Pseudomonas aeruginosa*: Arsenal of resistance mechanisms, decades of changing resistance profiles, and future antimicrobial therapies. *Fut. Microbiol.* vol. 10 1683–1706 (2015).
10. Tacconelli, E. et al. Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. *Lancet Infect. Dis.* 18, 318–327 (2018).

11. Drugs Development & Approval Process. Novel Drug Approvals for 2017. *U.S. Food Drug Adm.* 1–36 (2017).
12. Drugs Development & Approval Process. Novel Drug Approvals for 2018. *U.S. Food Drug Adm.* 1–36 (2018).
13. Drugs Development & Approval Process. Through Innovation New Drug Therapy Approvals 2019. *U.S. Food Drug Adm.* 1–44 (2020).
14. Mullard, A. 2019 FDA drug approvals. *Nat. Rev. Drug Discov.* 19, 79–84 (2020).
15. Nealson, K. H., Platt, T. & Hastings, J. W. Cellular control of the synthesis and activity of the bacterial luminescent system. *J. Bacteriol.* 104, 313–22 (1970).
16. Kalia, V. C. *Quorum Sensing vs Quorum Quenching: A Battle with No End in Sight*. (Springer India, 2015).
17. Pérez-Pérez, M., Jorge, P., Pérez Rodríguez, G., Pereira, M. O. & Lourenço, A. Quorum sensing inhibition in *Pseudomonas aeruginosa* biofilms: new insights through network mining. *Biofouling* 33, 128–142 (2017).
18. Pesci, E. C., Pearson, J. P., Seed, P. C. & Iglewski, B. H. Regulation of las and rhl quorum sensing in *Pseudomonas aeruginosa*. *J. Bacteriol.* 179, 3127–32 (1997).
19. Dubern, J. F. & Diggle, S. P. Quorum sensing by 2-alkyl-4-quinolones in *Pseudomonas aeruginosa* and other bacterial species. *Mol. Biosyst.* 4, 882–888 (2008).
20. Lee, J. et al. A cell-cell communication signal integrates quorum sensing and stress response. *Nat. Chem. Biol.* 9, 339–343 (2013).
21. Schuster, M. & Peter Greenberg, E. A network of networks: Quorum-sensing gene regulation in *Pseudomonas aeruginosa*. *Int. J. Med. Microbiol.* 296, 73–81 (2006).
22. Lagarde, N., Zagury, J. F. & Montes, M. Benchmarking Data Sets for the Evaluation of Virtual Ligand Screening Methods: Review and Perspectives. *Journal of Chemical Information and Modeling* vol. 55 1297–1307 (2015).

-
23. Sliwoski, G. R., Meiler, J. & Lowe, E. W. Computational Methods in Drug Discovery Prediction of protein structure and ensembles from limited experimental data View project Antibody modeling, Antibody design and Antigen-Antibody interactions View project. *Comput. Methods Drug Discov.* 66, 334–95 (2014).
 24. Macalino, S. J. Y., Gosu, V., Hong, S. & Choi, S. Role of computer-aided drug design in modern drug discovery. *Arc. of Pharmacal. Res.* vol. 38 1686–1701 (2015).
 25. Bajorath, J. Computer-aided drug discovery [version 1 ; referees : 3 approved] Referee Status : 4, 1–8 (2016).
 26. Jorgensen, W. L. The Many Roles of Computation in Drug Discovery. *Science* vol. 303 1813–1818 (2004).
 27. Aguayo-Ortiz, R. & Fernández-de Gortari, E. Overview of Computer-Aided Drug Design for Epigenetic Targets. in *Epi-Informatics* 21–52 (Academic Press, 2016).
 28. Brown, N. et al. Big Data in Drug Discovery. in *Progress in Medicinal Chemistry* vol. 57 277–356 (Elsevier B.V., 2018).
 29. Polanski, J. Big Data in Structure-Property Studies—From Definitions to Models. in 529–552 (2017).
 30. Mendez, D. et al. ChEMBL: Towards direct deposition of bioassay data. *Nucleic Acids Res.* 47, D930–D940 (2019).
 31. Kim, S. et al. PubChem 2019 update: Improved access to chemical data. *Nucleic Acids Res.* 47, D1102–D1109 (2019).
 32. Wishart, D. S. et al. DrugBank 5.0: A major update to the DrugBank database for 2018. *Nucleic Acids Res.* 46, D1074–D1082 (2018).
 33. Gutmanas, A. et al. PDBe: Protein data bank in Europe. *Nucleic Acids Res.* 42, (2014).
 34. Treepong, P. et al. Global emergence of the widespread *Pseudomonas aeruginosa* ST235 clone. *Clin. Microbiol. Infect.* 24, 258–266 (2018).

35. Gellatly, S. L. & Hancock, R. E. W. *Pseudomonas aeruginosa*: New insights into pathogenesis and host defenses. *Pathog. Dis.* 67, 159–173 (2013).
36. Moradali, M. F., Ghods, S. & Rehm, B. H. A. *Pseudomonas aeruginosa* lifestyle: A paradigm for adaptation, survival, and persistence. *Front. Cell. Infect. Microbiol.* 7, (2017).
37. Frimmersdorf, E., Horatzek, S., Pelnikevich, A., Wiehlmann, L. & Schomburg, D. How *Pseudomonas aeruginosa* adapts to various environments: A metabolomic approach. *Environ. Microbiol.* 12, 1734–1747 (2010).
38. Migiyama, Y. et al. *Pseudomonas aeruginosa* bacteremia among immunocompetent and immunocompromised patients: Relation to initial antibiotic therapy and survival. *Jpn. J. Infect. Dis.* 69, 91–96 (2016).
39. Hauser, A. R. So Many Virulence Factors, So Little Time. *Crit Care Med.* 39, 2193–2194 (2012).
40. Olusegun, A. et al. **We are IntechOpen , the world ' s leading publisher of Open Access books** Built by scientists , for scientists TOP 1 %. *Intech i*, 38 (2012).
41. Tacconelli, E., Carrara, E., Savoldi, A., Kattula, D. & Burkert, F. GLOBAL PRIORITY LIST OF ANTIBIOTIC-RESISTANT BACTERIA TO GUIDE RESEARCH, DISCOVERY, AND DEVELOPMENT OF NEW ANTIBIOTICS. <http://www.cdc.gov/drugresistance/threat-report-2013/>.
42. Centers for Disease Control and Prevention. *Multi-drug resistant Pseudomonas aeruginosa infection.* (2017).
43. Suetens, C. et al. Prevalence of healthcare-associated infections, estimated incidence and composite antimicrobial resistance index in acute care hospitals and long-term care facilities: Results from two european point prevalence surveys, 2016 to 2017. *Eurosurveillance* 23, 1–18 (2018).
44. Restrepo, M. I. et al. Burden and risk factors for *Pseudomonas aeruginosa* community-acquired pneumonia: A multinational point prevalence study of hospitalised patients. *Eur. Respir. J.* 52, (2018).

-
45. Instituto Nacional de Salud. *Boletín epidemiológico semanal 07 de 2020. Boletín epidemiológico semanal* (2020).
 46. Tamber, S. & Hancock, R. E. W. *Frontiers in Bioscience* 8, s472-483, May 1, 2003] ON THE MECHANISM OF SOLUTE UPTAKE IN PSEUDOMONAS. 472-483 (2003).
 47. Baumgart, A. M. K., Molinari, M. A. & de Silveira, A. C. O. Prevalence of carbapenem resistant pseudomonas aeruginosa and acinetobacter baumannii in high complexity hospital. *Brazilian J. Infect. Dis.* 14, 433-436 (2010).
 48. Colclough, A. L. et al. RND efflux pumps in Gram-negative bacteria; Regulation, structure and role in antibiotic resistance. *Future Microbiol.* 15, 143-157 (2020).
 49. Oliver, A., Mulet, X., López-Causapé, C. & Juan, C. The increasing threat of Pseudomonas aeruginosa high-risk clones. *Drug Resist. Updat.* 21-22, 41-59 (2015).
 50. Poole, K. Stress responses as determinants of antimicrobial resistance in Gram-negative bacteria. *Trends Microbiol.* 20, 227-234 (2012).
 51. Fernández, L., Breidenstein, E. B. M. & Hancock, R. E. W. Creeping baselines and adaptive resistance to antibiotics. *Drug Resist. Updat.* 14, 1-21 (2011).
 52. Fraud, S., Campigotto, A. J., Chen, Z. & Poole, K. MexCD-OprJ multidrug efflux system of Pseudomonas aeruginosa: Involvement in chlorhexidine resistance and induction by membrane-damaging agents dependent upon the AlgU stress response sigma factor. *Antimicrob. Agents Chemother.* 52, 4478-4482 (2008).
 53. Lee, J. Y., Park, Y. K., Chung, E. S., Na, I. Y. & Ko, K. S. Evolved resistance to colistin and its loss due to genetic reversion in Pseudomonas aeruginosa. *Sci. Rep.* 6, 1-13 (2016).
 54. Fujitani, S., Moffett, K. & Yu, V. Pseudomonas aeruginosa. *Infect. Dis. Antimicrob. agents* (2018).
 55. Instituto Nacional de Vigilancia de Medicamentos y Alimentos INVIMA. *NORMAS FARMACOLÓGICAS.* (2020).
 56. Bassetti, M., Vena, A., Croxatto, A., Righi, E. & Guery, B. How to manage Pseudomonas

- aeruginosa infections. *Drugs Context* 7, 1–18 (2018).
57. Public Health England. Antibiotic Awareness: Key Messages 2017 World Antibiotic Awareness Week European Antibiotic Awareness Day Antibiotic Guardian. www.facebook.com/PublicHealthEngland (2017).
 58. Theuretzbacher, U. Antibiotic innovation for future public health needs. *Clin. Microbiol. Infect.* 23, 713–717 (2017).
 59. Holden, M. T., Diggle, S. P. & Williams, P. Quorum Sensing. in *Encyclopedia of Life Sciences* (John Wiley & Sons, Ltd, 2007).
 60. Miller, M. B. & Bassler, B. L. Quorum Sensing in Bacteria. *Annu. Rev. Microbiol.* 55, 165–199 (2001).
 61. Bai Aswathanarayan, J. & Ravishankar Rai, V. Quorum-Sensing Systems in *Pseudomonas*. (2015).
 62. **O'Reilly, M. C. & Blackwell, H. E. Structure-Based Design and Biological Evaluation of Triphenyl Scaffold-Based Hybrid Compounds as Hydrolytically Stable Modulators of a LuxR-Type Quorum Sensing Receptor.** *ACS Infect. Dis.* 2, 32–38 (2016).
 63. Huang, H. et al. An integrated genomic regulatory network of virulence-related transcriptional factors in *Pseudomonas aeruginosa*. *Nat. Commun.* 10, (2019).
 64. Lee, J. & Zhang, L. The hierarchy quorum sensing network in *Pseudomonas aeruginosa*. *Protein Cell* 6, 26–41 (2015).
 65. Garcia-Contreras, R. Is quorum sensing interference a viable alternative to treat *Pseudomonas aeruginosa* infections? *Front. Microbiol.* 7, 1–7 (2016).
 66. Barr, H. L. et al. *Pseudomonas aeruginosa* quorum sensing molecules correlate with clinical status in cystic fibrosis. *Eur. Respir. J.* 46, 1046–1054 (2015).
 67. Taha, M. N., Saafan, A. E., Ahmedy, A., El Gebaly, E. & Khairalla, A. S. Two novel synthetic peptides inhibit quorum sensing-dependent biofilm formation and some virulence factors in

- Pseudomonas aeruginosa* PAO1. *J. Microbiol.* 57, 618–625 (2019).
68. Abinaya, M. & Gayathri, M. Inhibition of biofilm formation, quorum sensing activity and molecular docking study of isolated 3, 5, 7-Trihydroxyflavone from *Alstonia scholaris* leaf against *P.aeruginosa*. *Bioorg. Chem.* 87, 291–301 (2019).
 69. Ćirić, A. D. et al. Natural products as biofilm formation antagonists and regulators of quorum sensing functions: A comprehensive review update and future trends. *South African Journal of Botany* vol. 120 65–80 (2019).
 70. Soukarieh, F. et al. In silico and in vitro-guided identification of inhibitors of alkylquinolone-dependent quorum sensing in *pseudomonas aeruginosa*. *Molecules* 23, (2018).
 71. Schütz, C. & Empting, M. Targeting the *Pseudomonas* quinolone signal quorum sensing system for the discovery of novel anti-infective pathoblockers. *Beilstein Journal of Organic Chemistry* vol. 14 2627–2645 (2018).
 72. Sampathkumar, S. J., Srivastava, P., Ramachandran, S., Sivashanmugam, K. & Gothandam, K. M. Lutein: A potential antibiofilm and antiquorum sensing molecule from green microalga *Chlorella pyrenoidosa*. *Microb. Pathog.* 135, 103658 (2019).
 73. Reina, J. C., Pérez-Victoria, I., Martín, J. & Llamas, I. A Quorum-Sensing Inhibitor Strain of *Vibrio alginolyticus* Blocks Qs-Controlled Phenotypes in *Chromobacterium violaceum* and *Pseudomonas aeruginosa*. *Mar. Drugs* 17, 494 (2019).
 74. Teerapo, K., Roytrakul, S., Sistayanarain, A. & Kunthalert, D. A scorpion venom peptide derivative BmKn-22 with potent antibiofilm activity against *Pseudomonas aeruginosa*. *PLoS One* 14, (2019).
 75. Gómez-Gómez, B. et al. Selenium and tellurium-based nanoparticles as interfering factors in quorum sensing-regulated processes: Violacein production and bacterial biofilm formation. *Metallomics* 11, 1104–1114 (2019).
 76. Rutherford, S. T. & Bassler, B. L. Bacterial quorum sensing: Its role in virulence and possibilities for its control. *Cold Spring Harb. Perspect. Med.* 2, 1–25 (2012).

77. Soheili, V., Tajani, A. S., Ghodsi, R. & Bazzaz, B. S. F. Anti-PqsR compounds as next-generation antibacterial agents against *Pseudomonas aeruginosa*: A review. *Eur. J. Med. Chem.* 172, 26–35 (2019).
78. Maura, D., Hazan, R., Kitao, T., Ballok, A. E. & Rahme, L. G. Evidence for direct control of virulence and defense gene circuits by the *pseudomonas aeruginosa* quorum sensing regulator, MvFR. *Sci. Rep.* 6, 1–14 (2016).
79. Starkey, M. et al. Identification of Anti-virulence Compounds That Disrupt Quorum-Sensing Regulated Acute and Persistent Pathogenicity. *PLoS Pathog.* 10, (2014).
80. Wang, J. et al. Bacterial quorum-sensing signal IQS induces host cell apoptosis by targeting POT1-p53 signalling pathway. *Cell. Microbiol.* 21, e13076 (2019).
81. López Vallejo, F., Medina Franco, J. L. & Castillo, R. Diseño de fármacos asistido por computadora. *Educ. Química* 17, 452 (2006).
82. Saldívar-González, F., Prieto-Martínez, F. D. & Medina-Franco, J. L. Descubrimiento y desarrollo de fármacos: un enfoque computacional. *Educ. Quim.* 28, 51–58 (2017).
83. PROTEIN DATA BANK. PDB Statistics: PDB Data Distribution by Residue Count. <https://www.rcsb.org/stats/distribution-residue-count> (2020).
84. Meza Menchaca, T., Juárez-Portilla, C. & C. Zepeda, R. Past, Present, and Future of Molecular Docking. in *Drug Discovery and Development - New Advances* (IntechOpen, 2020).
85. Berry, M., Fielding, B. & Gamielien, J. *Practical Considerations in Virtual Screening and Molecular Docking. Emerging Trends in Computational Biology, Bioinformatics, and Systems Biology: Algorithms and Software Tools* (Elsevier Inc., 2015).
86. Prieto-Martínez, F. D., Arciniega, M. & Medina-Franco, J. L. Acoplamiento Molecular: Avances Recientes y Retos. *TIP Rev. Espec. en Ciencias Químico-Biológicas* 21, 65–87 (2018).
87. Hollingsworth, S. A. & Dror, R. O. Molecular Dynamics Simulation for All. *Neuron* 99, 1129–1143 (2018).

-
88. Azhagiya Singam, E. R., Tachachartvanich, P., La Merrill, M. A., Smith, M. T. & Durkin, K. A. Structural Dynamics of Agonist and Antagonist Binding to the Androgen Receptor. *J. Phys. Chem. B* 123, 7657–7666 (2019).
89. An, X. et al. How Does Agonist and Antagonist Binding Lead to Different Conformational Ensemble Equilibria of the κ -Opioid Receptor: Insight from Long-Time Gaussian Accelerated Molecular Dynamics Simulation. *ACS Chem. Neurosci.* 10, 1575–1584 (2019).
90. Wang, Y. T. & Chan, Y. H. Understanding the molecular basis of agonist/antagonist mechanism of human mu opioid receptor through gaussian accelerated molecular dynamics method. *Sci. Rep.* 7, 1–11 (2017).
91. Kolinski, M. & Filipek, S. Molecular Dynamics of μ Opioid Receptor Complexes with Agonists and Antagonists. *Open Struct. Biol. J.* 2, 8–20 (2008).
92. Wiltgen, M. & Tilz, G. P. Homology modelling: Eine übersicht über die methode am beispiel der strukturbestimmung vom diabetes antigen GAD 65. *Wiener Medizinische Wochenschrift* 159, 112–125 (2009).
93. Muhammed, M. T. & Aki-Yalcin, E. Homology modeling in drug discovery: Overview, current applications, and future perspectives. *Chem. Biol. Drug Des.* 93, 12–20 (2019).
94. Rodriguez, R., Chinea, G., Lopez, N., Pons, T. & Vriend, G. Homology modeling, model and software evaluation: Three related resources. *Bioinformatics* 14, 523–528 (1998).
95. Croll, T. I., Sammito, M. D., Kryshtafovych, A. & Read, R. J. Evaluation of template-based modeling in CASP13. *Proteins Struct. Funct. Bioinforma.* 87, 1113–1127 (2019).
96. Senior, A. W. et al. Protein structure prediction using multiple deep neural networks in the 13th Critical Assessment of Protein Structure Prediction (CASP13). *Proteins Struct. Funct. Bioinforma.* 87, 1141–1148 (2019).
97. Liu, T., Lin, Y., Wen, X., Jorissen, R. N. & Gilson, M. K. BindingDB: A web-accessible database of experimentally determined protein-ligand binding affinities. *Nucleic Acids Res.* 35, 198–201 (2007).

98. Todeschini, R. & Consonni, V. *Handbook of Molecular Descriptors*. (John Wiley & Sons, 2008).
99. Sastry, M., Lowrie, J. F., Dixon, S. L. & Sherman, W. Large-scale systematic analysis of 2D fingerprint methods and parameters to improve virtual screening enrichments. *J. Chem. Inf. Model.* 50, 771–784 (2010).
100. Bajusz, D., Rácz, A. & Héberger, K. Chemical Data Formats, Fingerprints, and Other Molecular Descriptions for Database Analysis and Searching. in *Comprehensive Medicinal Chemistry III* (2017).
101. Ruddigkeit, L., Van Deursen, R., Blum, L. C. & Reymond, J. L. Enumeration of 166 billion organic small molecules in the chemical universe database GDB-17. *J. Chem. Inf. Model.* 52, 2864–2875 (2012).
102. Choudhury, C. & Narahari Sastry, G. Pharmacophore Modelling and Screening: Concepts, Recent Developments and Applications in Rational Drug Design. *Challenges Adv. Comput. Chem. Phys.* 27, 25–53 (2019).
103. Qing, X. et al. Pharmacophore modeling: Advances, Limitations, And current utility in drug discovery. *Journal of Receptor, Ligand and Channel Research* vol. 7 81–92 (2014).
104. Liu, C. et al. Pharmacophore-Based Virtual Screening Toward the Discovery of Novel Anti-echinococcal Compounds. *Front. Cell. Infect. Microbiol.* 10, 1–12 (2020).
105. Alpaydin, E. *Machine Learning: The New AI*. (The MIT Press, 2016).
106. Mayr, A. et al. Large-scale comparison of machine learning methods for drug target prediction on ChEMBL. *Chem. Sci.* 9, 5441–5451 (2018).
107. Naveja, J. J. & Medina-Franco, J. L. Finding Constellations in Chemical Space Through Core Analysis. *Front. Chem.* 7, 1–10 (2019).
108. RDKit: Open-source cheminformatics. <http://www.rdkit.org> (2013).
109. MolVS: Molecule Validation and Standardization. <https://molvs.readthedocs.io/en/latest/index.html> (2016).

-
110. Lovering, F., Bikker, J. & Humblet, C. Escape from flatland: Increasing saturation as an approach to improving clinical success. *J. Med. Chem.* 52, 6752–6756 (2009).
 111. Lovering, F. Escape from Flatland 2: Complexity and promiscuity. *Medchemcomm* 4, 515–519 (2013).
 112. Clemons, P. A. et al. Small molecules of different origins have distinct distributions of structural complexity that correlate with protein-binding profiles. *Proc. Natl. Acad. Sci. U. S. A.* 107, 18787–18792 (2010).
 113. Méndez-Lucio, O. & Medina-Franco, J. L. The many roles of molecular complexity in drug discovery. *Drug Discovery Today* (2017).
 114. Bertz H., S. The First General Index of Molecular Complexity. *J. Am. Chem. Soc.* 103, 3599–3601 (1981).
 115. Sander, T., Freyss, J., Von Korff, M. & Rufener, C. DataWarrior: An open-source program for chemistry aware data visualization and analysis. *J. Chem. Inf. Model.* 55, 460–473 (2015).
 116. McKinney, W. Data Structures for Statistical Computing in Python. *PROC. 9th PYTHON Sci. CONF. (SCIPY 2010)* 51–56 (2010).
 117. Hu, Y., Stumpfe, D. & Bajorath, J. Computational Exploration of Molecular Scaffolds in Medicinal Chemistry. *Journal of Medicinal Chemistry* (2016).
 118. Stumpfe, D., Dimova, D. & Bajorath, J. Computational method for the systematic identification of analog series and key compounds representing series and their biological activity profiles. *J. Med. Chem.* 59, 7667–7676 (2016).
 119. Zdrzil, B. & Guha, R. The Rise and Fall of a Scaffold: A Trend Analysis of Scaffolds in the Medicinal Chemistry Literature. *J. Med. Chem.* 61, 4688–4703 (2018).
 120. Saldivar-González, F. I., Naveja, J. J., Palomino-Hernández, O. & Medina-Franco, J. L. Getting SMART in drug discovery: Chemoinformatics approaches for mining structure-multiple activity relationships. *RSC Advances* vol. 7 632–641 (2017).

121. Bemis, G. W. & Murcko, M. A. The properties of known drugs. 1. Molecular frameworks. *J. Med. Chem.* 39, 2887–2893 (1996).
122. Golbamaki, A., Franchi, A. M. & Gini, G. The Maximum Common Substructure (MCS) Search as a New Tool for SAR and QSAR. in *Advances in QSAR Modeling. Challenges and Advances in Computational Chemistry and Physics* 149–165 (2017).
123. Xu, Y. J. & Johnson, M. Using molecular equivalence numbers to visually explore structural features that distinguish chemical libraries. *J. Chem. Inf. Comput. Sci.* 42, 912–926 (2002).
124. Maggiora, G. M. On outliers and activity cliffs - Why QSAR often disappoints. *Journal of Chemical Information and Modeling* vol. 46 1535 (2006).
125. Cruz-Monteaugudo, M. et al. Activity cliffs in drug discovery: Dr Jekyll or Mr Hyde? *Drug Discov. Today* 19, 1069–1080 (2014).
126. Hu, Y. & Bajorath, J. Activity profile relationships between structurally similar promiscuous compounds. *Eur. J. Med. Chem.* 69, 393–398 (2013).
127. Medina-Franco, J. L. Activity cliffs: Facts or artifacts? *Chem. Biol. Drug Des.* 81, 553–556 (2013).
128. Medina-Franco, J. L. et al. Characterization of activity landscapes using 2D and 3D similarity methods: Consensus activity cliffs. *J. Chem. Inf. Model.* (2009).
129. Sud, M. MayaChemTools: An Open Source Package for Computational Drug Discovery. *J. Chem. Inf. Model.* 56, 2292–2297 (2016).
130. Saldívar-González, F. I., Lenci, E., Trabocchi, A. & Medina-Franco, J. L. Exploring the chemical space and the bioactivity profile of lactams: a chemoinformatic study. *RSC Adv.* 9, 27105–27116 (2019).
131. Firdaus Begam, B., Begam, B. F. & Kumar, J. S. Visualization of Chemical Space Using Principal Component Analysis. *World Appl. Sci. J.* 29, 53–59 (2014).
132. Karlov, D. S., Sosnin, S., Tetko, I. V. & Fedorov, M. V. Chemical space exploration guided by deep neural networks. *RSC Adv.* 9, 5151–5157 (2019).

-
133. De La Vega De León, A. & Bajorath, J. Chemical space visualization: Transforming multidimensional chemical spaces into similarity-based molecular networks. *Future Med. Chem.* 8, (2016).
 134. Pettersen, E. F. et al. UCSF Chimera?A visualization system for exploratory research and analysis. *J. Comput. Chem.* 25, 1605–1612 (2004).
 135. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410 (1990).
 136. Wu, S., Skolnick, J. & Zhang, Y. Ab initio modeling of small proteins by iterative TASSER simulations. *BMC Biol.* 5, 1–10 (2007).
 137. Ovchinnikov, S., Park, H., Kim, D. E., DiMaio, F. & Baker, D. Protein structure prediction using Rosetta in CASP12. *Proteins Struct. Funct. Bioinforma.* 86, 113–121 (2018).
 138. Meier, A. & Söding, J. Automatic Prediction of Protein 3D Structures by Probabilistic Multi-template Homology Modeling. *PLoS Comput. Biol.* 11, (2015).
 139. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. The Phyre2 web portal for protein modeling, prediction and analysis Lawrence. *Nat. Protoc.* 10, 845–858 (2016).
 140. Waterhouse, A. et al. SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* 46, W296–W303 (2018).
 141. Shapovalov, M. V. & Dunbrack Jr., R. L. A smoothed backbone-dependent rotamer library for proteins derived from adaptive kernel density estimates and regressions Maxim. 19, 844–858 (2012).
 142. Gasteiger, J. & Marsili, M. A new model for calculating atomic charges in molecules. *Tetrahedron Lett.* 19, 3181–3184 (1978).
 143. Halgren, T. a. Merck Molecular Force Field. *J. Comput. Chem.* 17, 490–519 (1996).
 144. Trott, O. & Olson, A. J. Software news and update AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J.*

- Comput. Chem.* 31, 455–461 (2010).
145. Salentin, S., Schreiber, S., Haupt, V. J., Adasme, M. F. & Schroeder, M. PLIP: Fully automated protein-ligand interaction profiler. *Nucleic Acids Res.* 43, W443–W447 (2015).
146. Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* 9, 90–95 (2007).
147. Bergdorf, M. et al. Desmond / GPU Performance as of November 2016. 2015, 1–11 (2016).
148. Humphrey, W., Dalke, A. & Schulten, K. VMD -- Visual Molecular Dynamics. *J. Mol. Graph.* 14, 33–38 (1996).
149. Gowers, R. et al. MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations. *Proc. 15th Python Sci. Conf.* 98–105 (2016).
150. Seyler, S. L., Kumar, A., Thorpe, M. F. & Beckstein, O. Path Similarity Analysis: A Method for Quantifying Macromolecular Pathways. *PLoS Comput. Biol.* 11, 1–36 (2015).
151. Chemical Computing Group ULC. Molecular Operating Environment (MOE). (2020).
152. Ali, M. PyCaret: An open source, low-code machine learning library in Python. (2020).
153. O'Reilly, M. C. et al. Structural and Biochemical Studies of Non-native Agonists of the LasR Quorum-Sensing Receptor Reveal an L3 Loop “Out” Conformation for LasR. *Cell Chem. Biol.* 25, 1128–1139.e3 (2018).
154. Wermuth, C. G. *The Practice of Medicinal Chemistry*, 3rd Ed. 982 (2011).
155. Hao, M., Bryant, S. H. & Wang, Y. Cheminformatics analysis of the AR agonist and antagonist datasets in PubChem. *J. Cheminform.* 8, 37 (2016).
156. Boursier, M. E., Combs, J. B. & Blackwell, H. E. N -Acyl I -Homocysteine Thiolactones Are Potent and Stable Synthetic Modulators of the RhIR Quorum Sensing Receptor in *Pseudomonas aeruginosa*. *ACS Chem. Biol.* 14, 186–191 (2019).
157. Eibergen, N. R., Moore, J. D., Mattmann, M. E. & Blackwell, H. E. Potent and Selective Modulation of the RhIR Quorum Sensing Receptor by Using Non-native Ligands: An Emerging

-
- Target for Virulence Control in *Pseudomonas aeruginosa*. *Chembiochem* 16, 2348–56 (2015).
158. Ahumedo Monterrosa, M., Galindo, J. F., Vergara Lorduy, J., Alf-Torres, J. & Vivas-Reyes, R. The role of LasR active site amino acids in the interaction with the Acyl Homoserine Lactones (AHLs) analogues: A computational study. *J. Mol. Graph. Model.* 86, 113–124 (2019).
159. Bateman, A. UniProt: A worldwide hub of protein knowledge. *Nucleic Acids Res.* 47, D506–D515 (2019).
160. Cavasotto, C. N. & Phatak, S. S. Homology modeling in drug discovery: current trends and applications. *Drug Discovery Today* vol. 14 676–683 (2009).
161. Kim, T. et al. Structural insights into the molecular mechanism of *Escherichia coli* SdiA, a quorum-sensing receptor. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 70, 694–707 (2014).
162. Nguyen, Y. et al. Structural and mechanistic roles of novel chemical ligands on the SdiA quorum-sensing transcription regulator. *MBio* 6, 1–10 (2015).
163. Lintz, M. J., Oinuma, K. I., Wysoczynski, C. L., Greenberg, E. P. & Churchill, M. E. A. Crystal structure of OscR, a *Pseudomonas aeruginosa* quorum sensing signal receptor. *Proc. Natl. Acad. Sci. U. S. A.* 108, 15763–15768 (2011).
164. Paczkowski, J. E. et al. An Autoinducer Analogue Reveals an Alternative Mode of Ligand Binding for the LasR Quorum-Sensing Receptor. *ACS Chem. Biol.* 14, 378–389 (2019).
165. Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.* 7, 95–99 (1963).
166. Studer, G. et al. OMEANDisCo—distance constraints applied on model quality estimation. *Bioinformatics* 36, 1765–1771 (2020).
167. Wallner, B. & Elofsson, A. Can correct protein models be identified? *Protein Sci.* 12, 1073–1086 (2003).
168. Chowdhury, N. & Bagchi, A. Identification of ligand binding activity and DNA recognition by RhlR protein from opportunistic pathogen *Pseudomonas aeruginosa*—a molecular dynamic

- simulation approach. *J. Mol. Recognit.* 31, 1–7 (2018).
169. Kim, H. et al. Artificial Intelligence in Drug Discovery: A Comprehensive Review of Data-driven and Machine Learning Approaches. *Biotechnol. Bioprocess Eng.* 25, 895–930 (2020).
170. Yang, M. et al. Machine Learning Models Based on Molecular Fingerprints and an Extreme Gradient Boosting Method Lead to the Discovery of JAK2 Inhibitors. *J. Chem. Inf. Model.* 59, 5002–5012 (2019).
171. Fernández, A. et al. *Learning from Imbalanced Data Sets*. Learning from Imbalanced Data Sets (Springer International Publishing, 2018).
172. Chawla, N. V. Data Mining for Imbalanced Datasets: An Overview. in *Data Mining and Knowledge Discovery Handbook* 853–867 (Springer-Verlag, 2006).
173. Chawla, N. V., Bowyer, K. W., Hall, L. O. & Kegelmeyer, W. P. SMOTE: Synthetic minority over-sampling technique. *J. Artif. Intell. Res.* 16, 321–357 (2002).
174. Ji, X., Tong, W., Liu, Z. & Shi, T. Five-Feature Model for Developing the Classifier for Synergistic vs. Antagonistic Drug Combinations Built by XGBoost. *Front. Genet.* 10, 600 (2019).
175. Chen, T. & Guestrin, C. XGBoost: A scalable tree boosting system. in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* vols 13-17-August-2016 785–794 (Association for Computing Machinery, 2016).
176. Vass, M. et al. Molecular interaction fingerprint approaches for GPCR drug discovery. *Curr. Opin. Pharmacol.* 30, 59–68 (2016).
177. Velázquez-Libera, J. L., Rossino, G., Navarro-Retamal, C., Collina, S. & Caballero, J. Docking, Interaction Fingerprint, and Three-Dimensional Quantitative Structure–Activity Relationship (3D-QSAR) of Sigma1 Receptor Ligands, Analogs of the Neuroprotective Agent RC-33. *Front. Chem.* 7, (2019).
178. Khan, M. F., Tang, H., Lyles, J. T. & Pineau, R. Antibacterial Properties of Medicinal Plants From Pakistan Against Multidrug-Resistant ESKAPE. 9, 1–17 (2018).

179. Atef, N. M., Shanab, S. M., Negm, S. I. & Abbas, Y. A. Evaluation of antimicrobial activity of some plant extracts against antibiotic susceptible and resistant bacterial strains causing wound infection. 9, (2019).
180. Aparicio, R. & Velasco, J. Composición química y actividad antibacteriana del aceite esencial de los frutos de una nueva especie del género. 158-174 (2017).