

Estudio *in silico* de acoplamiento molecular para la búsqueda de potenciales inhibidores para el tratamiento de tuberculosis

Javier Alejandro Daza Galván^{1a}, Laura Alejandra Valencia Torres^{1b}, Karol Dayana Duque Santamaría^{1c}

¹ *Pregrado de Química Farmacéutica, Facultad de Ciencias y Biotecnología, Universidad CES, Cl 10A #22 - 04, Medellín, Antioquia, Colombia.*

² **Correos Electrónicos**

^a *daza.javier@uces.edu.co*

^b *laura.valenciat@uces.edu.co*

^c *duque.karol@uces.edu.co*

RESUMEN

Introducción: Según la Organización Mundial de la Salud (OMS) la tuberculosis (TB), es una enfermedad causada por *Mycobacterium Tuberculosis* (MTB), la cual sigue siendo una de las enfermedades infecciosas de mayor prevalencia a nivel mundial. La tuberculosis es una enfermedad tratable, y en los últimos años se han hecho grandes esfuerzos por reducir la incidencia de dicha enfermedad, sin embargo, factores como la falta de adhesión a los tratamientos por parte de los pacientes ha desencadenado cepas multirresistentes. **Objetivo:** Identificar *in silico* mediante docking molecular potenciales inhibidores para las proteínas de interés que corresponden al dominio uno y dos (D1 y D2) de las proteínas MmpL-3 y MmpL-11. Además de evaluar la capacidad de unión de los ligandos a los dominios y su potencial inhibitorio al bloquear el transporte de sustrato a través de la membrana. **Metodología:** Se construyeron

modelos de los D1 Y D2 de MmpL-3 y MmpL-11, se seleccionaron los mejores modelos, los cuales se utilizaron para el acoplamiento proteína-ligando, se determinaron la interacciones formadas entre los ligandos y los dominios de interés y las afinidades de dicha interacción por medio de base de datos y recursos informáticos disponibles. **Resultados:** Se obtuvo información acerca de los potenciales ligandos, se seleccionaron las tres mejores moléculas: Glecaprevir, digoxina y bisoctrizole, se evaluaron las distintas interacciones entre estos posibles inhibidores y los dominios concluyendo que digoxina fue la molécula más adecuada para convertirse en candidata en el tratamiento de la tuberculosis.

Palabras clave: Acoplamiento molecular, MmpL-3, *M. Tuberculosis*, MmpL-11.

SUMMMARY

Introduction: According to the World Health Organization (WHO) tuberculosis (TB), is a disease caused by Mycobacterium Tuberculosis (MTB), remains one of the most prevalent infectious diseases worldwide. Tuberculosis is a treatable disease, and in recent years great efforts have been made to reduce the incidence of this disease, however, factors such as the lack of adherence to treatment by patients have triggered multi-resistant strains. **Objective:** To identify in silico through molecular coupling possible inhibitors for the proteins of interest that correspond to domain one and two (D1 and D2) of the MmpL-3 and MmpL-11 proteins. In addition to evaluating the binding capacity of the ligands to the domains and their inhibitory potential by blocking the transport of substrate through the membrane. **Methodology:** Models of the D1 and D2 of MmpL-3 and MmpL-11 were built, the best models were selected, which were used for the detection of protein-ligand, the interactions formed between the ligands and the domains of interest were determined and the affinities of said interaction by means of databases and available computing resources. Results: Information about the possible ligands was obtained, the three best

molecules were selected: Glecaprevir, digoxin and bisoctrizole, the different interactions between these possible inhibitors and the domains were evaluated, concluding that digoxin was the most suitable molecule to become a candidate in the tuberculosis treatment.

Keywords: Molecular docking, MmpL-3, *M. Tuberculosis*, MmpL-11.

INTRODUCCIÓN

La tuberculosis (TB) es una patología causada por la bacteria *Mycobacterium tuberculosis* (MTB) que afecta directamente los alvéolos pulmonares, transmitida por contacto persona-persona, a través de bacilos expulsados mediante estornudos, saliva o microgotas generadas por la tos [1]. Actualmente el sistema de farmacovigilancia consiste en cuatro fármacos antituberculosos de primera línea: Isoniazida, rifampicina, etambutol y pirazinamida para tratamientos con regímenes de 6 meses en TB causada por cepas sensibles a fármacos [3]; sin embargo, para tratamientos cuando la enfermedad es causada por cepas multi-farmacorresistentes donde el nivel de complejidad aumenta, existe la necesidad de adicionar medicamentos por periodos de tiempo mayores a los 24 meses [4], como consecuencia se aumenta el riesgo de padecer efectos adversos como neuritis, hepatitis, elevación de transaminasas y bilirrubinas, y nefrotoxicidad [5]. Por otro lado, la problemática en la ineficacia de los fármacos de elección, radica en 2 factores: (1) debido a mutaciones del ADN causados por mecanismos de supervivencia propios de la bacteria, donde se estableció que el tratamiento debe incluir mínimo 4 fármacos potencialmente activos de primera línea, un inyectable (amikacina) y una quinolona; (2) los riesgos de recaída que aumentan en casos de diseminaciones nosocomiales, infecciones por VIH, malnutrición, drogadicción y alcoholismo en un 21%.[3] Según el reporte más reciente de la “Organización Mundial de la Salud” (OMS), se concluye que los objetivos propuestos para poner fin a la tuberculosis durante el período comprendido entre 2015 y 2020 no se cumplieron a cabalidad, pues se esperaba una

reducción del 20% en la incidencia, y una del 35% en la tasa de mortalidad, pero apenas se logró reducir la incidencia en un 11%, y la tasa de mortalidad en un 9.2%. [2]

En los últimos años las investigaciones se han centrado en el estudio de nuevos blancos farmacológicos, en donde se destacan las proteínas transmembranales *Mycobacterial membrane protein large* (MmpL) importantes para la supervivencia de la bacteria, más precisamente MmpL-3 y MmpL-11 las cuales están encargadas de exportar sustratos hidrófobos voluminosos que son esenciales para la estructura única de la envoltura celular y apoyan directamente la capacidad de MTB para infectar y persistir en el huésped [6]. Mediante la técnica DALI (Distance matrix alignment) se han documentado hallazgos de homología estructural entre el dominio D2 de MmpL-11 con otros transportadores conocidos como los RND (Resistance-Nodulation-Cell Division) que participan en el transporte de múltiples fármacos actuando como bombas de eflujo, claves para mecanismos de resistencia bacteriana [7]; por otro lado, no se ha elucidado hasta el momento la dinámica molecular de las interacciones entre los dominios periplásmicos MmpL-3-D1/D2 ni los detalles termodinámicos de las interacciones de los dominios con sus posibles inhibidores. [8]

METODOLOGIA

1. Identificación de las secuencias de aminoácidos

La identificación de las secuencias de aminoácidos para los dominios D1 y D2 tanto para MmpL-3 y MmpL-11 se realizó usando la base de datos de UniProtKB [9]. (Figura 1a y 1b), se consideraron los aminoácidos que están entre las posiciones 35-185 (D1) y 418-562 (D2) de la secuencia de MMPL3_MYCTU (nombre de entrada para MmpL-3 *M.Tuberculosis* en la

plataforma de UniProtKB) obtenida de la cepa Strain ATCC 25618/ H37Rv, y aquellos que se encuentran en las posiciones 34-187 (D1) y 394-526 (D2) de la secuencia de MMPLB_MYCTU (nombre de entrada para MmpL-11 de *M. Tuberculosis* en la plataforma de UniProtKB).

Topology

Feature key	Position(s)	Description	Actions	Graphical view	Length
Topological domain ¹	1 – 13	Cytoplasmic By similarity	Add BLAST		13
Transmembrane ¹	14 – 34	Helical Sequence analysis	Add BLAST		21
Topological domain ¹	35 – 185	Periplasmic By similarity	Add BLAST		151
Transmembrane ¹	186 – 206	Helical Sequence analysis	Add BLAST		21
Topological domain ¹	207 – 209	Cytoplasmic By similarity	Add BLAST		3
Transmembrane ¹	210 – 230	Helical Sequence analysis	Add BLAST		21
Topological domain ¹	231 – 235	Periplasmic By similarity	Add BLAST		5
Transmembrane ¹	236 – 256	Helical Sequence analysis	Add BLAST		21
Topological domain ¹	257 – 286	Cytoplasmic By similarity	Add BLAST		30
Transmembrane ¹	287 – 307	Helical Sequence analysis	Add BLAST		21
Topological domain ¹	308 – 314	Periplasmic By similarity	Add BLAST		7
Transmembrane ¹	315 – 335	Helical Sequence analysis	Add BLAST		21
Topological domain ¹	336 – 396	Cytoplasmic By similarity	Add BLAST		61
Transmembrane ¹	397 – 417	Helical Sequence analysis	Add BLAST		21
Topological domain ¹	418 – 562	Periplasmic By similarity	Add BLAST		145
Transmembrane ¹	563 – 583	Helical Sequence analysis	Add BLAST		21
Topological domain ¹	584 – 586	Cytoplasmic By similarity	Add BLAST		3
Transmembrane ¹	587 – 607	Helical Sequence analysis	Add BLAST		21
Topological domain ¹	608 – 616	Periplasmic By similarity	Add BLAST		9
Transmembrane ¹	617 – 637	Helical Sequence analysis	Add BLAST		21
Topological domain ¹	638 – 672	Cytoplasmic By similarity	Add BLAST		35
Transmembrane ¹	673 – 693	Helical Sequence analysis	Add BLAST		21
Topological domain ¹	694 – 698	Periplasmic By similarity	Add BLAST		5
Transmembrane ¹	699 – 719	Helical Sequence analysis	Add BLAST		21
Topological domain ¹	720 – 944	Cytoplasmic By similarity	Add BLAST		225

A

Topology

Feature key	Position(s)	Description	Actions	Graphical view	Length
Transmembrane ¹	13 – 33	Helical Sequence analysis	Add BLAST		21
Transmembrane ¹	188 – 208	Helical Sequence analysis	Add BLAST		21
Transmembrane ¹	214 – 234	Helical Sequence analysis	Add BLAST		21
Transmembrane ¹	235 – 255	Helical Sequence analysis	Add BLAST		21
Transmembrane ¹	279 – 299	Helical Sequence analysis	Add BLAST		21
Transmembrane ¹	311 – 331	Helical Sequence analysis	Add BLAST		21
Transmembrane ¹	373 – 393	Helical Sequence analysis	Add BLAST		21
Transmembrane ¹	527 – 547	Helical Sequence analysis	Add BLAST		21
Transmembrane ¹	557 – 577	Helical Sequence analysis	Add BLAST		21
Transmembrane ¹	595 – 615	Helical Sequence analysis	Add BLAST		21
Transmembrane ¹	646 – 666	Helical Sequence analysis	Add BLAST		21
Transmembrane ¹	668 – 688	Helical Sequence analysis	Add BLAST		21

B

Figura 1a: Imagen tomada del sitio web de UniProtKB que describe la topología en concordancia con la secuencia proteínica de MmpL-3 que corresponden a las distintas estructuras, en donde se destacan las líneas anaranjadas grandes que corresponden a los dominios periplásmicos. Siendo las posiciones 35-185 correspondientes a D1 y 418-562 correspondientes a D2. **Figura 1b:** Imagen tomada del sitio web de UniProtKB que describe la secuencia corresponde a MmpL-11, que basándonos en la similitud con

MmpL-3, se determinó que los espacios señalados en rojo corresponden a los fragmentos de la secuencia de los dominios D1 (posiciones 34-187) y D2 (posiciones 394-526).

2. Desarrollo de los modelos proteínicos 3D y Validación de MmpL-11 en i-TASSER

Para modelar *in silico* la interacción entre los dominios D1 y D2 tanto de MmpL-3 como de MmpL-11 y posteriormente correr el docking de los dominios D1 y D2 de cada proteína individualmente contra una serie de inhibidores descritos en la literatura es ideal tener la estructura cristalina completa o parcial de las proteínas de interés, sin embargo, en el Protein Data Bank (PDB) se han reportado hasta la fecha la estructura cristalina de MmpL-11 D2 de MTB con el código 4Y0L y la estructura cristalina de MmpL-3 de *mycobacterium tuberculosis* con código 7NVH elucidada por el método criomicroscopía electrónica (Cryo-EM) [10] se identificaron los dominios D1 y D2 de la proteína a partir de las posiciones anteriormente descritas dentro de las secuencias correspondientes a cada uno de estos dominios, para así obtener un modelo estructural solamente de los dominios tanto juntos como separados cada una de las proteínas mediante la herramienta software Chimera UCSF [11]. Para modelar la estructura de MmpL-11 se utilizó el software bioinformático Iterative Threading ASSEmbly Refinement (I-TASSER) el cual predice el modelo de estructura tridimensional de proteínas a partir de secuencias de aminoácidos y detecta las plantillas de estructura del PDB mediante una técnica llamada reconocimiento de plegado [12,13,14], se realizó la separación de los dominios D1 y D2 de la proteína mediante Chimera UCSF. Antes de hacer el docking se validó la calidad de los modelos mediante el uso de parámetros energéticos locales y globales con la finalidad de seleccionar el mejor modelo, de este modo se tuvieron en cuenta los parámetros que evalúan la calidad en I-TASSER junto con los modelos escogidos de MmpL-11 para hacer la separación de sus dominios anteriormente mencionados.

Por último, se utilizó el servidor ProSA Web para determinar la validación de los modelos con únicamente los dominios D1 y D2 de ambas proteínas [15, 16].

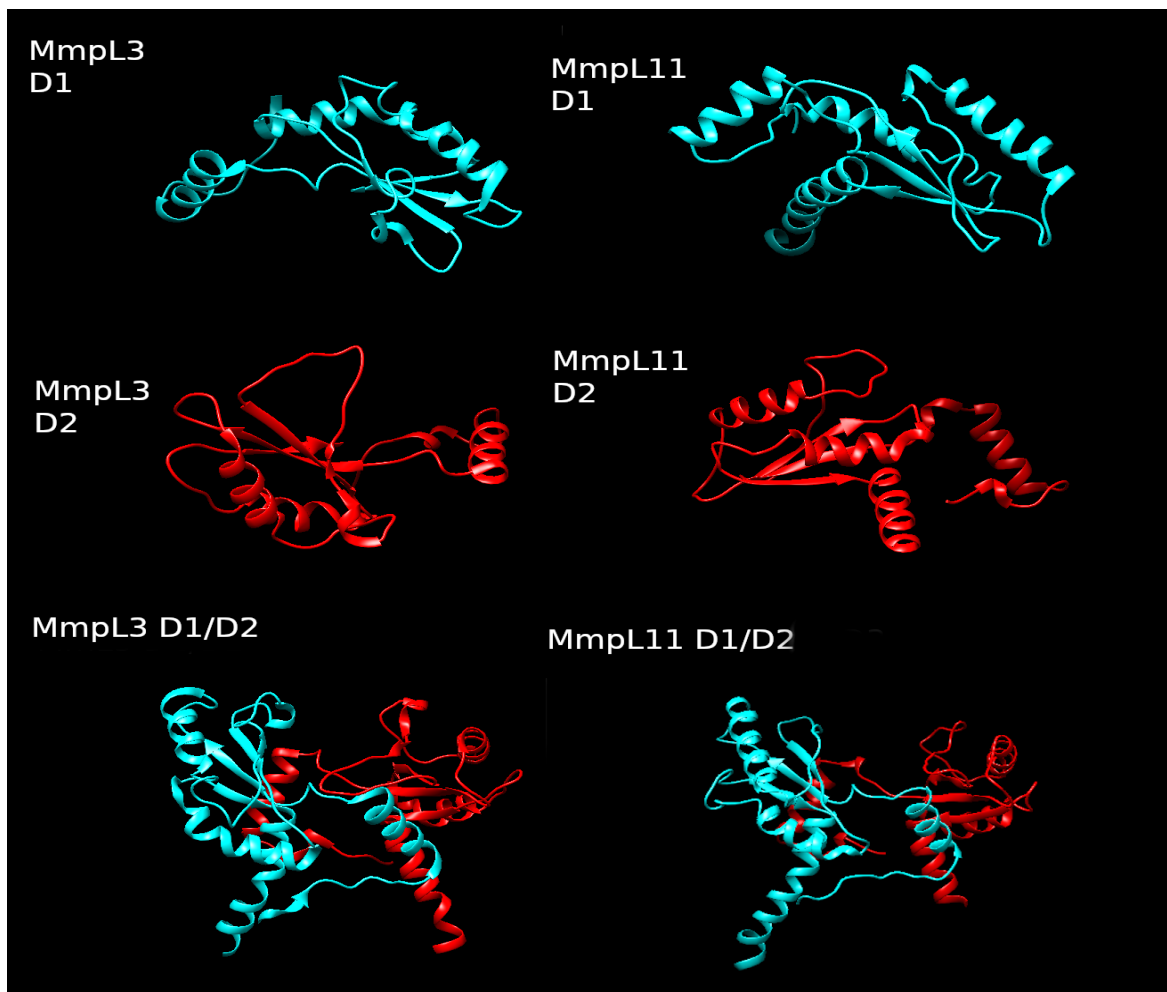


Figura 2 Modelos utilizados para el docking molecular. Los modelos MmpL-3 D1/D2 y MmpL-11 D1/D2 fueron los utilizados para el tamizaje virtual contra la librería de ligandos. Los modelos MmpL-3 D1, MmpL-3 D2, MmpL-11 D1, MmpL-11 D2 fueron los utilizados en la plataforma de ClusPro para determinar cómo se da la interacción y la afinidad de dicha interacción cuando se forma el dímero.

3. Construcción de la librería de ligandos.

Se construyó una librería propia de 132 ligandos usando como base moléculas a partir de los ligandos: SQ109, BM2012 y AU12345, utilizando Swiss Similarity [18] tomando como

referencia las librerías de ZINC Drug like, PDB ligands, ChEMBL y compuestos naturales aprobados para determinar nuevos posibles ligandos o si estos podrían tener un alto potencial para un segundo uso, se escogieron las moléculas con un score de 0.5 de similaridad, para añadir las moléculas a la librería propia de ligandos se transformaron los canonical smiles de cada molécula a formato PDB para poder obtener la estructura tridimensional mediante la herramienta Online SMILES Translator and Structure File Generator CACTUS Web Server del National Cancer Institute.[19]

Los modelos de las estructuras de las proteínas obtenidos anteriormente se prepararon junto con los ligandos en AutodockTools [17] con el fin de que queden en las condiciones adecuadas en el formato de archivo adecuado mediante la adición de cargas gasteiger asignadas automáticamente por el programa y la adición de hidrógenos polares para finalmente ejecutar el docking en Autodock Vina.

3.2 Acoplamiento proteína-ligando

Para el docking proteína-ligando, se realizó un corrido de la librería de ligandos utilizando el software de PyRx [20] que viene con Autodock Tools y AutoDock Vina integrados [21,22], con el fin de facilitar y optimizar la programación del docking. Es de suma importancia identificar los posibles sitios activos de unión de los sustratos putativos, para esto se usó AutoDockTools, para la determinación de la posición tridimensional del ligando se estableció el tamaño y las posiciones X, Y y Z de el centro del grid en el sitio activo el cual abarca toda su totalidad tal como se observa en la tabla 1.

Valores para el Acoplamiento Molecular		
	MmpL-3	MmpL-11
Centro x	136.4487	97.9270
Centro y	128.8425	62.5173
Centro z	167.8876	74.2841
Dimensión x	25 puntos	25 puntos
Dimensión y	25 puntos	25 puntos
Dimensión z	25 puntos	25 puntos
Valor de exhaustividad	8	8

Tabla 1. Valores utilizados en Grid para el acoplamiento molecular. Se determinaron los valores de exhaustividad y dimensiones correspondientes para MmpL-3 y MmpL-11. Esto se hizo para especificar el sitio del docking basado en la posición del ligando reportado en el PDB como AV0 que interactúa con el cristal de MmpL-3 7NVH.

Para la interpretación de los resultados del docking entre las 132 moléculas de la biblioteca con los dominios D1 y D2 de MmpL-3 y MmpL-11, se tomó como ligando de referencia a comparar AU1235, considerando que los ligandos evaluados que posean una energía de afinidad menor a este ligando presentan un alto potencial como inhibidor. Las 3 interacciones con mayor afinidad, es decir, con una menor energía de afinidad necesaria para formar el complejo proteína-ligando resultantes del acoplamiento molecular se analizaron en LIGPLOT+ para identificar las diferentes interacciones electrostáticas (puentes de hidrógeno) e hidrofóbicas que se generaron, para esto, partiendo de los resultados de salida del docking obtenidos en PyRx, se crearon modelos tridimensionales en PyMol [23], en donde se juntaron cada resultado de salida con la proteína MmpL-3 D1/D2 para crear un archivo en formato .pdb para posteriormente transformarlos en

modelos gráficos, es decir, en imágenes que evidencian de forma más sencilla de interpretar cómo se dan las diversas interacciones de cada ligando con MmpL-3 D1/D2.

Es de suma importancia considerar cómo interactúan los dominios D1 y D2 para tener una idea de la unión de estos con el fin de identificar cuál o cuáles moléculas interfieren con esta interacción de forma que evite que se forme el dímero, por ende para el acoplamiento molecular entre los dominios de MmpL-3 y MmpL-11 se utilizó el servidor web de ClusPro [24], este servidor tiene 3 pasos principales para realizar el docking: (1) acoplamiento de cuerpo rígido mediante el muestreo de miles de millones de conformaciones, (2) agrupación basada en la desviación cuadrática media (RMSD) de las 1000 estructuras de energía más baja generadas para encontrar las agrupaciones más grandes que representarán los modelos más probables del complejo, y (3) refinamiento de estructuras seleccionadas usando minimización de energía con el fin de remover los impedimentos estéricos. El paso de acoplamiento de cuerpo rígido utiliza PIPER, un programa de acoplamiento basado en el enfoque de correlación Fast Fourier Transform (FFT), PIPER conserva las 1000 estructuras acopladas de menor energía para su posterior procesamiento y espera que este conjunto incluya al menos algunas que estén cerca de la estructura nativa del complejo. Para interpretar adecuadamente los resultados en ClusPro se seleccionaron los clusters más poblados en estructuras de baja energía en lugar de simplemente las conformaciones de energía más baja como predicciones del complejo proteínico, puesto que el tamaño de cada cluster representa el ancho del pozo de energía correspondiente y, por lo tanto, brinda más información sobre las contribuciones entrópicas a la energía libre. En el servidor de ClusPro se escogió al dominio D1 como blanco proteínico y al D2 como ligando, tomando en cuenta las recomendaciones del tutorial descrito en la página del servidor web, entre las cuales destaca que es recomendable utilizar como ligando al dominio de menor peso molecular. [24]

4. Validación De Los Modelos de Proteínas

Los resultados se obtuvieron con el fin de definir un modelo adecuado de MmpL-11 para posteriormente separar los dominios D1/D2 y validar dicho modelo para el docking molecular mediante ProSA web. Con base a los parámetros de validación utilizados en los modelos en la plataforma de i-TASSER, el modelo 1 de la proteína MmpL-11 fue el más adecuado entre los 5 mejores. En un conjunto de pruebas de referencia de 500 proteínas no homólogas realizadas por i-TASSER, se encontró que el C-Score está altamente correlacionado con la puntuación TM y RMSD. El coeficiente de correlación del C-Score del primer modelo con la puntuación TM con la estructura nativa es 0.91, mientras que el coeficiente del C-Score con el RMSD de la estructura nativa es 0.75. Estos datos realmente sientan la base para la predicción confiable de la puntuación TM y RMSD utilizando la puntuación C. Los valores informados en las Columnas 3 y 4 (Tabla 2) son los valores estimados de la puntuación TM y RMSD en función de su correlación con la puntuación C.

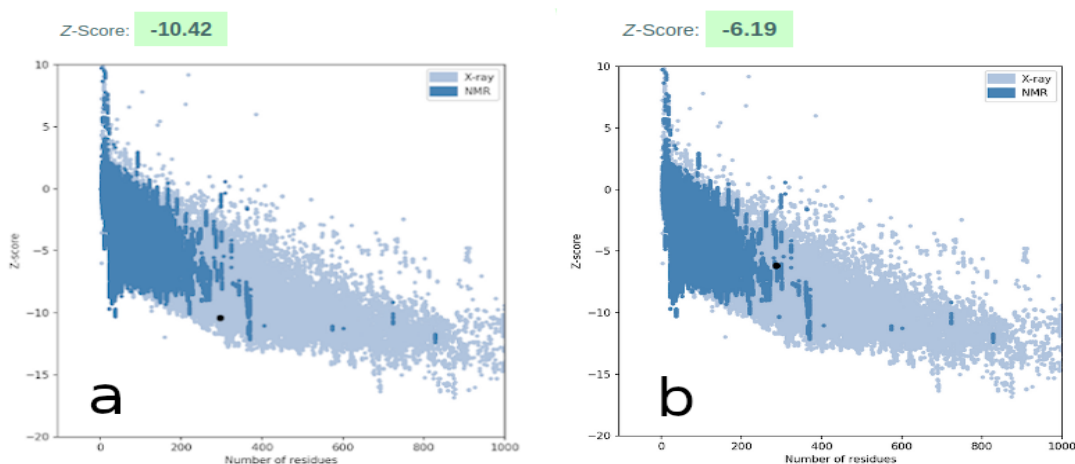
Modelo de MmpL-11	C.score	Exp.TM-Score	Exp.RMSD
Modelo 1	-0.49	0.65+-0.13	10.0+-4.6
Modelo 2	-1.36	N/A	N/A
Modelo 3	-1.70	N/A	N/A
Modelo 4	-3.19	N/A	N/A
Modelo 5	-3.33	N/A	N/A

Tabla 2. Resultados de los parámetros de validación de los 5 mejores modelos de MmpL-11. El C-score es una puntuación de confianza para estimar la calidad de los modelos predichos por I-TASSER. El C-score está típicamente en el rango de [-5,2], donde un C-score de valor más alto significa un modelo con una alta confianza, que en este caso el modelo de mayor confianza sería el modelo 1. El TM-score y el RMSD son estándares conocidos para medir la similitud estructural entre dos estructuras que generalmente se usan para medir la precisión del modelado de estructuras cuando se conoce la estructura nativa. En este caso, como no se conoce la estructura nativa, es necesario predecir la calidad de la predicción del modelado, para esto i-TASSER intentó estimar la puntuación TM y la RMSD de los modelos predichos en relación con las estructuras nativas en función del C-score Solo se reportó la predicción de calidad (TM-Score y RMSD) para el primer modelo, porque de acuerdo con i-TASSER se encontró que la correlación entre el C-Score y TM-Score es débil para los modelos de rango más bajo. Sin embargo, se enumeraron los C-score de todos los modelos solo como referencia. Un TM-score >0.5 indica un modelo de mayor exactitud topológica, hecho que se evidencia para el modelo 1.

La validación de los modelos de MmpL-3 D1/D2 y MmpL-11 D1/D2 mediante ProSA Web dieron resultados satisfactorios que evidencian alta confianza y calidad en ambos modelos (figura 3). Su valor se evidencia en una gráfica que contiene los z-score (es un indicador de la calidad general para cada modelo [14,16] de todas las cadenas de proteínas determinadas experimentalmente en el PDB (Figura 3a y 3b)). En este gráfico se distingue por colores los grupos de estructuras de diferentes fuentes (rayos X, RMN), y se pudo comprobar que el z-score de ambos modelos se encuentra dentro del rango de puntajes en donde normalmente se hallan las proteínas nativas de tamaño similar. Los gráficos de puntuaciones de los residuos (Figura 3c y 3d) evidencian la energía promedio sobre cada ventana de fragmentos de 40 y 10 residuos de aminoácidos, en donde se determinó que el modelo con mayor partes problemáticas corresponde

a MmpL 11 D1/D2 para una ventana de 10 aminoácidos, aun así para una ventana de 40 residuos el modelo muestra una alta confiabilidad; para el modelo MmpL-3 D1/D2 no se evidenciaron partes de la proteína para una ventana de 40 residuos que fuese problemática, en cambio, los 10 residuos de aminoácidos evidenciaron pocas partes de la proteína que fuesen problemáticas o erróneas.

Modelos de calidad general



Modelos de calidad local

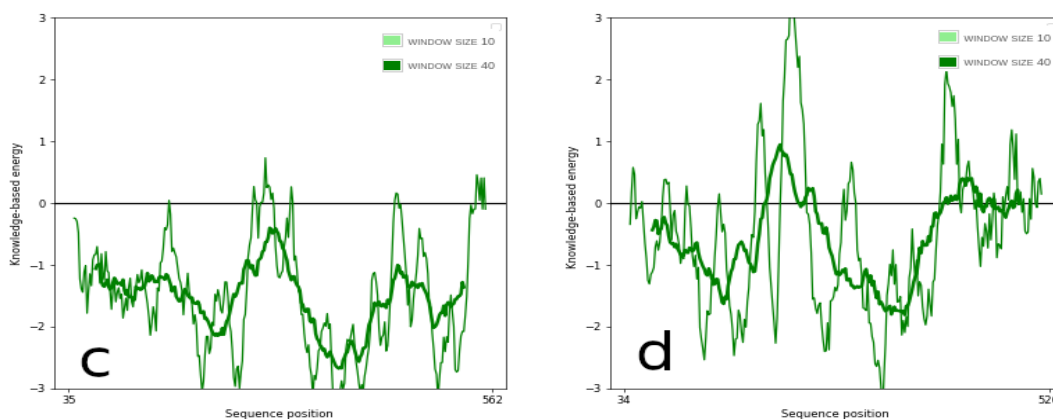


Figura 3: Calidad general para cada modelo y gráficos de puntuaciones de los residuos. Las gráficas de calidad general para cada modelo (figura 3a y 3b) permiten determinar si el modelo se comporta como

una proteína nativa del PDB. Los gráficos de puntuaciones de residuos evidencian la energía promedio sobre cada ventana de fragmentos de 40 y 10 residuos de aminoácidos. Estos gráficos muestran la calidad local de los modelos representando las energías en función de la posición de la secuencia de aminoácidos.

RESULTADOS

1. Resultados de los ensayos *in silico* de Acoplamiento molecular.

Los acoplamientos moleculares resultantes al correr la librería de ligandos para identificar potenciales inhibidores de MmpL-3 D1/D2 en PyRx revelaron 15 ligandos claves relacionados, directamente con la energía de unión de cada interacción con posible función inhibitoria, en donde se evidencian 6 moléculas ya aprobadas para su uso en el ámbito clínico, de acuerdo con la información obtenida en el DrugBank [25] (Rifamicina S (Rifamycin S), Digoxina (Digoxin), Glecaprevir, Bisotriczole, Ácido Glicirricinico, Cardiogreen (indocianina verde)), de los demás ligandos no se tienen como tal avances o estudios que den indicios que puedan ser usados en el tratamiento para TB, no obstante, sería necesario evaluar en el futuro si estos ligandos cumplen con las condiciones de las reglas de Lipinski para determinar si estos son adecuados como posibles inhibidores que puedan estudiarse en ensayos *in vitro*, preclínicos y clínicos. Partiendo de los resultados del listado, se escogieron tres ligandos cuya energía de afinidad fuese bastante menor con respecto al inhibidor de referencia, en este caso se hizo uso del inhibidor AU1235, que presentó un energía de afinidad de -7.6 kcal/mol al correr el docking con MmpL-3 D1/D2. Los 3 ligandos que formaron los complejos más estables con MmpL-3 D1/D2 fueron: Rifamicina S (Rifamycin S), Glecaprevir y Bisotriczole, con valores de energía afinidad de -10.8 kcal/mol, -10.3 kcal/mol y -10.2 kcal/mol respectivamente, por ende, en base a estos resultados se puede definir a Rifamicina S como el ligando que forma el complejo más estable con MmpL-3 D1/D2;

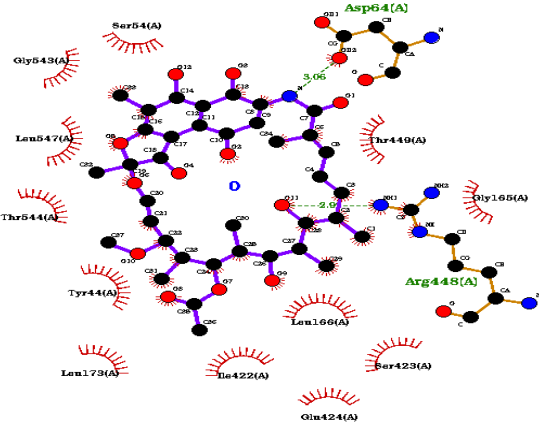
por otra parte se deben tener en cuenta las interacciones relacionadas con estas energías de afinidad obtenidas partiendo de los resultados de LIGPLOT+ en las interacción de cada uno de estos ligandos con MmpL-3 D1/D2 (Tabla 3). Se puede concluir que el ligando que presenta el complejo con el mayor número de interacciones de puente de hidrógeno (total de 2 puentes de hidrógeno) formados entre este ligando y los aminoácidos sea la Rifamicina S, que de hecho, es una molécula perteneciente de la familia de las rifamicinas en donde se destacan varios fármacos ya utilizados para tratar la tuberculosis, no obstante, al ser esta una molécula base para la formación de otras moléculas de uso clínico como la Rifampicina [26], se eligió a Glecaprevir como el ligando de interés.

Los 15 ligandos con mayor afinidad para MmpL3 D1/D2				
Número	Ligando	Energía de afinidad (kcal/mol)	RMSD/lb	RMSD/ub
1	Rifamycin S	-10.8	0	0
2	Glecaprevir	-10.3	0	0
3	Bisectriazole	-10.2	0	0
4	Glycyrrhizic acid	-10.2	0	0
5	Digoxin	-10.2	0	0
6	CHEMBL1088314	-10	0	0
7	CHEMBL1080747	-9.4	0	0
8	CHEMBL1081302	-9.4	0	0
9	CHEMBL1086754	-9.4	0	0
10	Cardiogreen	-9.4	0	0
11	CHEMBL482221	-9.4	0.893	1.357
12	CHEMBL1087773	-9.2	0	0
13	CHEMBL1088312	-9.2	0	0
14	ZINC33139598	-9.1	0	0
15	CHEMBL1080671	-9.1	0	0
	AU1235	-7.6	0	0

Los 15 ligandos con mayor afinidad para MmpL11 D1/D2				
Número	Ligando	Energía de afinidad (kcal/mol)	RMSD/lb	RMSD/ub
1	Glecaprevir	-11.4	0	0
2	Digoxin	-10.3	0	0
3	Glycyrrhizic acid	-10.1	0	0
4	Bisectriazole	-9.4	0	0
5	CHEMBL1080671	-9.4	0	0
6	Cardiogreen	-9.4	0	0
7	Rifamycin S	9.2	0	0
8	CHEMBL1214177	-8.9	0	0
9	CHEMBL1086754	-8.7	0	0
11	CHEMBL1087773	-8.7	0	0
12	CHEMBL258392	-8.7	0	0
13	CHEMBL1818393	-8.7	0	0
14	CHEMBL482221	-8.5	0	0
15	CHEMBL1088314	-8.5	0	0
	AU1235	-7.9	0	0

Tabla 3 (Superior) y Tabla 4 (Inferior). Lista de Resultados Obtenidos de los 15 mejores ligandos para la proteína MmpL-3 D1/D2 y MmpL-11 D1/D2 mediante Docking Molecular Masivo. **RMSD/lb:** Root-Mean-Square Deviation/ lower bound. **RMSD/ub:** Root-Mean-Square Deviation/ upper bound.

Ejecutando una serie de acoplamientos moleculares de características similares con el que se realizó la identificación de potenciales inhibidores para MmpL-3 D1/D2, tal como se observa en la Tabla 4, se determinaron también 15 ligandos claves, entre los cuales se destacan los que presentaron las tres primeras interacciones con la mayor afinidad hacia MmpL-11 D1/D2: Ácido Glicirricínico (Glycyrrhizic Acid) con una energía de afinidad de -10.1 kcal/mol, Digoxina (Digoxin) con -10.3 kcal/mol, y por último, Glecaprevir con -11.4 kcal/mol, siendo este el ligando que forma el complejo con mayor afinidad hacia MmpL-11 D1/D2; sin embargo, se optó por el ligando Digoxina como complejo de mayor afinidad debido a que presenta mayor número de interacciones por puentes de hidrógeno. Hay que destacar que los resultados que se tuvieron en cuenta fueron los resultados de salida de los acoplamientos moleculares cuyo Root Mean Square Deviation (Desviación estándar de la media al cuadrado) tanto para los enlaces inferiores como superiores (RMSD/lb y RMSD/up respectivamente) presenten un valor no mayor a 1.5 unidades.

	<p>2 Puentes de hidrógeno formado por las interacciones entre los aminoácidos Asp 64 y Arg 448 con Rifamicina S.</p>	<p>12 Interacciones hidrofóbicas de los aminoácidos de MmpL-3 D1/D2 con el Rifamycin S: Tyr 44 Ser 54, Gly 165 Leu 166, Leu 173, Ile 422, Ser 423, Glu 424, The 449, Thr 544, Leu 547, Gly 543.</p>
-------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

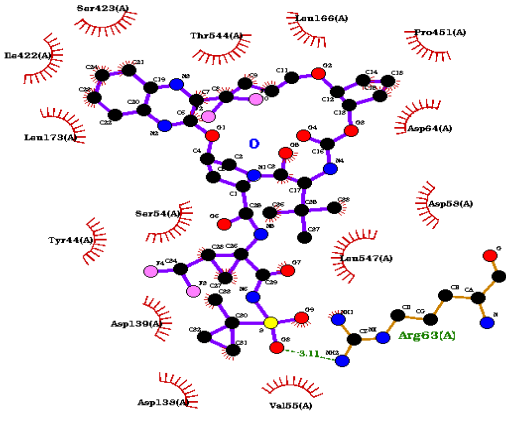
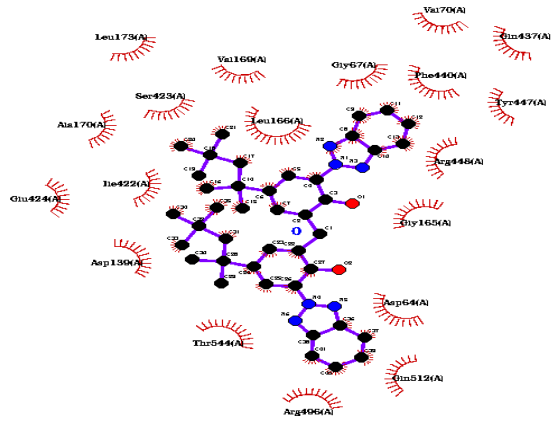
<p>b)</p> 	<p>Se presenta una interacción de puente de hidrógeno entre Glecaprevir con el aminoácido MmpL-3 D1/D2.</p>	<p>El total de las 14 interacciones hidrofóbicas corresponden a los siguientes aminoácidos: Tyr 44, Ser 54, Val 55, Asp 59, Asp 64, Asp 138, Asp 139, Leu 166, Leu 179, Ile 422, Ser 423, Pro 451, Thr 544 y Leu 547.</p>
<p>c)</p> 	<p>No se presentan interacciones por puente de hidrógeno entre Bisectrizole y MmpL-3 D1/D2</p>	<p>19 Interacciones hidrofóbicas de los aminoácidos de MmpL-3 D1/D2 y Bisectrizole: Asp 64, Gly 67, Val 70, Asp 139, Gly 165, Leu 166, Val 169, Ala 170, Leu 173, Ile 422, Ser 423, Glu 424, Gln 437, Phe 440, Tyr 447, Arg 448, Arg 496, Gln 512, Thr 544.</p>

Tabla 5. Análisis del acoplamiento molecular en LIGPLOT+ de los diferentes complejos MmpL-3 D1/D2-ligando. Los enlaces de hidrógeno se muestran como líneas discontinuas verdes y las interacciones hidrofóbicas se muestran como arcos rojos. En la figura 5a) se evidencian las diferentes interacciones del complejo MmpL-3 D1/D2- Rifamicina S, figura 5b) interacciones del complejo MmpL-3 D1/D2 - Glecaprevir y por último figura 5c) interacciones del complejo MmpL-3 D1/D2.

<p>a)</p>	<p>Se observa únicamente una interacción por puente de hidrógeno entre el aminoácido Ala 421 de MmpL-11 D1/D2 con Glecaprevir.</p>	<p>Se observa un total de 18 interacciones hidrofóbicas de Glecaprevir con los siguientes aminoácidos: Pro 66, Ala 69, Arg 109, Ile 161, Gly 162, Gln 163, Asn 394, Ser 395, Leu 396, Leu 397, Ala 410, Gly 420, Leu 422, Gly 423, Asp 482, Asp 485, Pro 513, Thr 514.</p>
<p>b)</p>	<p>Se evidencian en total 6 interacciones de puente de hidrógeno entre Digoxina (Digoxin) y los siguientes aminoácidos de MmpL-11 D1/D2: Gly 39, Ala 128, Ser 131, Gln 163, Gly 423, Asp 482.</p>	<p>Se observan 10 interacciones hidrofóbicas entre Digoxina (Digoxin) y los siguientes aminoácidos de MmpL-11 D1/D2: Thr 130, Leu 166, Ser 167, Asn 394, Ala 421, Leu 422, Asp 485, Val 491, Pro 513, Thr 514,</p>

<p>c)</p>	<p>Se evidencian 7 interacciones por puente de hidrógeno entre el Ácido Glicirricínico (Glycyrrhizic acid) y los siguientes aminoácidos de MmpL-11: Glu 43, Asp 57, Asp 62, Asn 127, Ser 131, Ser 167, Asn 394.</p>	<p>Se observan 10 interacciones hidrofóbicas entre el Ácido Glicirricínico (Glycyrrhizic acid) y los aminoácidos siguientes de MmpL-11: Glu 43, Ala 128, Thr 130, Gln 163, Leu 166, Ser 395, Leu 396, Leu 422, Gly 423, Pro 513.</p>
-----------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Tabla 6. Análisis del acoplamiento molecular en LIGPLOT+ de los diferentes complejos MmpL-11 D1/D2-ligando. En la figura 6a) se evidencian las diferentes interacciones del complejo MmpL-3 D1/D2-Glecaprevir, figura 6b) interacciones del complejo MmpL-3 D1/D2 - Digoxina y por último figura 6c) interacciones del complejo MmpL-3 D1/D2- Ácido Glicirricínico.

2. Resultados del docking proteína-proteína realizado en ClusPro.

Para conocer cómo puede influir en la inhibición de la formación del dímero D1-D2 tanto para MmpL-3 como MmpL-11 los ligandos anteriormente considerados, es de suma importancia saber cómo interaccionan estos dominios. Según los resultados obtenidos del acoplamiento molecular proteína-proteína entre los dominios D1/D2 tanto para MmpL-3 y MmpL-11 (Tablas 7 y 8). Se determinó que el modelo cluster 0 en ambos casos es el más adecuado teniendo en cuenta que en ellos se encuentra el mayor número de miembros en el cluster, 97 miembros para el modelo cluster 0 D1-D2 en MmpL-3, y 126 miembros para el modelo cluster 0 D1-D2 en MmpL-11.

Puntuación de los 10 mejores modelos del docking D1-D2 para MmpL3			
Cluster	Miembros	Representante	Puntuación ponderada
0	97	Centro	-939.8
		Menor energía	-1639
1	88	Centro	-896.9
		Menor energía	-1139.1
2	67	Centro	-961.2
		Menor energía	-1175.6
3	65	Centro	-1056.4
		Menor energía	-1073.4
4	37	Centro	-917.2
		Menor energía	-1223.4
5	36	Centro	-877.1
		Menor energía	-991.6
6	35	Centro	-873.9
		Menor energía	-993.4
7	32	Centro	-1021.5
		Menor energía	-1069.2
8	32	Centro	-856.2
		Menor energía	-991.2
9	29	Centro	-901
		Menor energía	-932.3

Puntuación de los 10 mejores modelos del docking D1-D2 para MmpL11			
Cluster	Miembros	Representante	Puntuación Ponderada
0	126	Centro	-1120.3
		Menor energía	-1330.5
1	59	Centro	-931.9
		Menor energía	-1150.1
2	45	Centro	-941.2
		Menor energía	-1024.9
3	41	Centro	-908.3
		Menor energía	-1154.6
4	41	Centro	-923.8
		Menor energía	-1037.5
5	36	Centro	-954.8
		Menor energía	-1069.5
6	31	Centro	-950
		Menor energía	-1019.8
7	31	Centro	-1308.1
		Menor energía	-1340.2
8	30	Centro	-932.5
		Menor energía	-1106.9
9	28	Centro	-899.2
		Menor energía	-1056.3

Tablas 7 (Izquierda) y 8 (Derecha). Puntuaciones de los 10 mejores modelos del docking D1-D2 tanto para MmpL-3 como MmpL-11.

Las interacciones entre los dominios generan una variada estequiometría cuando se trata de la formación de los dímeros, ya que a pesar de ser dominios de proteínas homólogas las interacciones que se presentan son diferentes. La obtención de los modelos tridimensionales de ClusPro es clave para determinar cómo se da la disposición espacial que adquieren los aminoácidos mediante las distintas interacciones moleculares involucradas en la unión de los dominios D1 y D2 de MmpL-3 y MmpL-11; de este modo se observó que modelo cluster 0 de cada acoplamiento molecular predice con mayor exactitud la orientación que se origina entre ambos dominios para formar el dímero MmpL-3 D1/D2 y el dímero MmpL-11 D1/D2 (Tabla 9).

Modelos	Cluster	Miembros	Representante Centro	Representante Menor Energía
---------	---------	----------	----------------------	-----------------------------

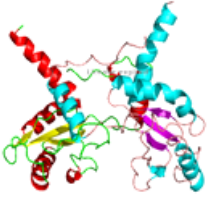
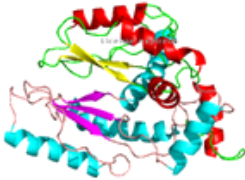
	0	97	-939.8	-1639
	0	126	-1120.3	-1330.5

Tabla 9. Los mejores modelos en ClusPro para cada proteína describen la disposición más acertada en el espacio de los dominios D1 y D2 para formar el dímero.

Por otro lado, la determinación de las distintas interacciones proteína-proteína entre los dominios D1 y D2 para MmpL-3 se evidencian en la Figura 5. Se observan como los aminoácidos Tyr 44, Ser, 50, Asp 58, Tyr 61, Gly 62, Asp 64, Arg 65, Ser 66, Lys 162 que forman parte del dominio D1 interactúan mediante puentes de hidrógenos con los siguientes aminoácidos del dominio D2 respectivamente: Ile 422, Glu 439, Tyr 447, Glu 495, Arg 496, Gly 501, Arg 509, Gln 512, Ala 546. Al comparar todo estas interacciones entre aminoácidos involucrados en la estabilidad del dímero MmpL-3 D1/D2 que interactúan con el ligando glecaprevir, se evidencian que como tal estos aminoácidos involucrados Tyr 44, Ser 54, Arg 63, Asp 64, Leu 547 y Ille 422. En el caso de las interacciones intramoleculares detalladas en la Figura 6 entre los dominios D1 y D2 de MmpL-11 se observan como los aminoácidos Gly 40, Phe 42, Glu 43, Val 44, Asp 57, Arg 63, Arg 109, Ala 128, Arg 115, Tyr 159, Ser 167 que forman parte del dominio D1, interactúan mediante puentes

de hidrógenos con los siguientes aminoácidos del dominio D2 respectivamente: Leu 422, Ala 421, Asn 394, Ser 395, Arg 398, Ala 433, Gly 434, Asp 469 Asp 485, Asp 509, Leu 516. Al igual que en el caso de la interacciones intermoleculares entre los dominios D1 y D2 para MmpL-3, la gran mayoría de interacciones son hidrofóbicas. Al realizarse la misma comparación con el fin de identificar a los aminoácidos involucrados en la estabilidad del dímero MmpL-11 D1/D2 con los que interactúan con Digoxina, se obtuvieron siguientes aminoácidos: Gly 39, Ala 128, Gln 163, Leu 166, Ser 167, Asn 394, Ala 421, Leu 422, Asp 485, Pro 513.

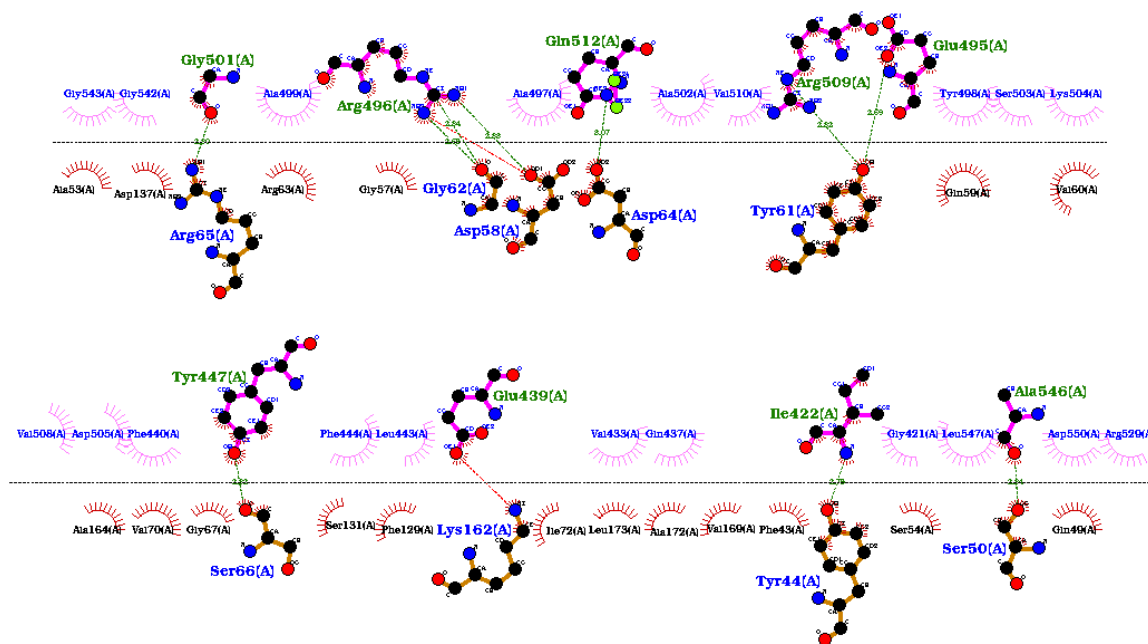


Figura 5. Interacciones intermoleculares entre los aminoácidos de los dominios D1-D2 de MmpL-3 a partir de Cluspro y LIGPLOT+.

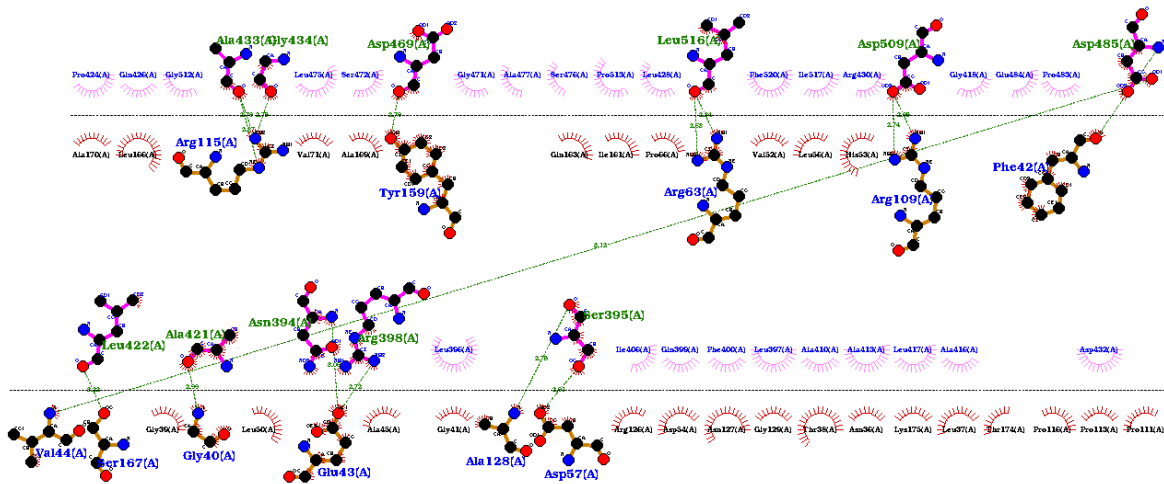


Figura 6. Interacciones intermoleculares entre los aminoácidos de los dominios D1-D2 de MmpL-11 a partir de Cluspro y LIGPLOT.

Con base a los resultados obtenidos, se puede concluir que la estabilidad de las interacciones entre los dominios D1 y D2 de las proteínas transmembranales MmpL-3 y MmpL-11 se verían afectadas por los ligandos Glecaprevir y Digoxina, respectivamente, lo cual da indicio de que estos ligandos a pesar de que ya tienen un uso farmacológico (Glecaprevir es un antiviral utilizado para el tratamiento de Hepatitis C [27], por otra parte, Digoxina es un antiarrítmico utilizado para el tratamiento de insuficiencia cardiaca [28]), pueden ser objeto de estudio con el fin de determinar si son adecuados como alternativa a los tratamientos actuales de tuberculosis. Considerando lo anterior, estos resultados generan nuevas expectativas y oportunidades en la mejora de tratamientos de primera línea para todos los tipos de estadios de tuberculosis. Se espera que con este hallazgo, podamos contribuir al desarrollo de nuevas ideas de investigación

AGRADECIMIENTOS

Agradecimientos a la Facultad de Ciencias y Biotecnología de la Universidad CES, y a nuestra tutora Karol Dayana Duque Santamaría por guiarnos y brindarnos los conocimientos bases para el desarrollo de este proyecto. También queremos brindar un agradecimiento especial a nuestras respectivas familias y amigos por el apoyo, la paciencia y la motivación.

CONFLICTOS DE INTERESES

Los autores declaran no tener conflictos de intereses.

REFERENCIAS

1. Centers for Disease Control and Prevention CDC. Preguntas y respuestas sobre la tuberculosis. <https://www.cdc.gov/tb/esp/publications/faqs/TBQASp.pdf>, consultado en marzo de 2021
2. World Health Organization. Global tuberculosis report 2021., Geneva, 2021. License: CC BY-NC-SA 3.0 IGO, p.13-14.
3. Navas Elorza E., Moreno Guillén S. Tuberculosis multirresistente y extremadamente resistente. *Rev. esp. sanid. penit.* **12**(3): 91-98 (2010).
4. Korycka-Machała M, Viljoen A, Pawełczyk J, Borówka P, Dziadek B, Gobis K, et al. 1H-Benzo[d]imidazole Derivatives Affect MmpL3 in Mycobacterium tuberculosis. *Antimicrobial Agents and Chemotherapy*. **63**(10), e00441-19 (2019).
5. Boldú J., Cebollero P., Abu J., Prado A. de. Tratamiento de la tuberculosis pulmonar. *Anales Sis San Navarra.* **30**(Suppl 2), 99-115 (2007).
6. Melly Geoff., Purdy Georgiana E. MmpL proteins in Physiology and Pathogenesis of M. tuberculosis. *Microorganisms* **7**(3):70 (2019)
7. Holm L, Rosenström P. Dali server: Conservation mapping in 3D. *Nucleic Acids Res.* *Nucleic Acids Research*, Volume 38, Issue suppl_2, 1 July 2010, Pages W545–W549 (2010)
8. Chim Nicholas, Torres Rodrigo, Liu Yuqi, Capri Joe, Batott Gaëlle, P. Whitelegge Julian, Gouldin Celia W. The Structure and Interactions of Periplasmic Domains of Crucial

- MmpL Membrane Proteins from *Mycobacterium tuberculosis*. *Chem Biol* **22**(8):1098-107. (2015)
9. The UniProt Consortium. UniProt: the universal protein knowledgebase in 2021, *Nucleic Acids Research*, Volume **49**, Issue D1, Pages D480–D489. (2021).
 10. Adams O, Deme JC, Parker JL, Fowler PW, Lea SM, Newstead S. Cryo-EM structure and resistance landscape of *M. tuberculosis* MmpL3: An emergent therapeutic target. *Structure*. **29**(10):1182-1191.e4. (2021)
 11. UCSF Chimera--a visualization system for exploratory research and analysis. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. *J Comput Chem*. **25**(13):1605-12. (2004)
 12. J Yang, R Yan, A Roy, D Xu, J Poisson, Y Zhang. The I-TASSER Suite: Protein structure and function prediction. *Nature Methods*, **12**: 7-8 (2015).
 13. A Roy, A Kucukural, Y Zhang. I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protocols*, **5**: 725-738 (2010)
 14. Y Zhang. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*, vol. **9**, no. 1, p. 40. (2008)
 15. Wiederstein & Sippl. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Research* **35**:407-410. (2007)
 16. Sippl, M.J. Recognition of Errors in Three-Dimensional Structures of Proteins. *Proteins* **17**, 355-362. (1993)
 17. Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S. and Olson, A. J. (2009) Autodock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J. Computational Chemistry* **30**(16), pp. 2785-91. (2009).
 18. Zoete, V., Daina, A., Bovigny, C., & Michielin, O. SwissSimilarity: A Web Tool for Low to Ultra High Throughput Ligand-Based Virtual Screening., *J. Chem. Inf. Model.* **56**(8), pp.1399-1404. (2016)
 19. National Cancer Institute. Online SMILES Translator.
<https://cactus.nci.nih.gov/translate/index.html#Form> Consultado en abril de 2022
 20. Dallakyan S, Olson AJ. Small-molecule library screening by docking with PyRx. En: “Chemical Biology Methods and Protocols”. Ed. por Jonathan E. Hempel, Charles H.

- Williams ,Charles C. Hong Springer Science+Business Media, New York 2015. Vol **1263**, pp. 243-50.
21. Eberhardt, J., Santos-Martins, D., Tillack, A.F., Forli, S. AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings. *JJ. Chem. Inf. Model.* **61** (8), pp. 3891–3898. (2021)
 22. The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.AxPyMOL.
 23. Trott, O., & Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry*, **31**(2), 455-461. (2010)
 24. Kozakov D, Hall DR, Xia B, Porter KA, Padhorny D, Yueh C, Beglov D, Vajda S. The ClusPro web server for protein-protein docking. *Nat Protoc.* **12**(2):255-278. (2017)
 25. Wishart DS, Knox C, Guo AC, Shrivastava S, Hassanali M, Stothard P, Chang Z, Woolsey J. Drugbank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res.* **1**;34 (Database issue):D668-72. (2006)
 26. Giovanni Battista Migliori, Alimuddin Zumla. 148 - Antituberculosis Agents. En: “ Infectious Diseases (Fourth Edition)s”. Ed. por Jonathan Cohen, William G. Powderly, Steven M. Opal. Elsevier, New York 2017. Volume **2**, Pages 1264-1276.e2
 27. Kosloski Matthew, Wang Haoyu, Pugatch David, Mensa Federico, Gane Edward, Lawitz Eric, Marbury Thomas, Preston Richard, Kort, Jens, Liu Wei. Pharmacokinetics and safety of glecaprevir and pibrentasvir in HCV-negative subjects with hepatic impairment. *European Journal of Clinical Pharmacology.* **75**(2).(2019).
 28. Gordon A. Ewy. Digoxin: The Art and Science. *The American Journal of Medicine.* **128**, 1272-1274. (2015)

