

Universidad Católica de Santa María

Facultad de Ciencias Farmacéuticas, Bioquímicas y Biotecnológicas

Escuela profesional de Ingeniería Biotecnológica



EVALUACIÓN DE LAS INTERACCIONES MOLECULARES DEL COMPLEJO PROTEÍNICO Erp, Rv1417 y Rv2617c DEL *Mycobacterium tuberculosis* Y SU POTENCIAL OBJETIVO FRENTE A DERIVADOS DE LA NICOTINA UTILIZANDO TÉCNICAS DE MODELAMIENTO COMPUTACIONAL

Tesis presentada por la Bachiller:

Paco Chipana, Margot Inés

para optar el Título Profesional de:

Ingeniero Biotecnólogo

Asesor: Dr. Gómez Valdez Badhin

**Arequipa – Perú
2019**

UNIVERSIDAD CATOLICA SANTA MARIA
Facultad de Ciencias Farmacéuticas, Bioquímicas
y Biotecnológicas
Escuela Profesional de Ingeniería Biotecnológica

Expediente N°.20190000004203
N° Trámite en Fac. 423-2019
Fecha Recep. Fac. 31-01-2019

14

FORMATO UNICO PARA TRAMITACIÓN DE TÍTULO PROFESIONAL

DE: **PACO CHIPANA, Margot Inés**

PARA OPTAR EL TÍTULO PROFESIONAL DE INGENIERO BIOTECNOLOGO

"EVALUACION DE LAS INTERACCIONES MOLECULARES DEL COMPLEJO PROTEINICO ERP, Rv1417 Y Rv1617c DEL Mycobacterium tuberculosis Y SU POTENCIAL OBJETIVO FRENTE A DERIVADOS DE LA NICOTINA UTILIZANDO TECNICAS DE MODELAMIENTO COMPUTACIONAL"

DICTAMINADORES: **Dr. Jaime Cárdenas García 2) Mgter. Jaime Barreda del Carpio**

DICTAMEN DE PLAN: Señor Decano de la Facultad de Ciencias Farmacéuticas, Bioquímicas y Biotecnológicas, en atención a su designación, el Jurado Dictaminador del Plan de Tesis informa que: hechas las observaciones y subsanadas las correcciones; consideramos se encuentra APTO para continuar con el trámite de acuerdo al Reglamento de Grados y Títulos de la Facultad

Atentamente
FIRMAS:

FECHA 14-03-19

ASESOR: **Dr. Badhin Gómez Valdéz**

DICTAMEN ASESORÍA: Señor Decano de la Facultad de Ciencias Farmacéuticas, Bioquímicas y Biotecnológicas, en atención a su designación como asesor del trabajo de investigación presentado por la recurrente, tengo a bien informar que se ha verificado el cumplimiento de los objetivos y redacción del informe con los resultados, discusión y conclusiones correspondientes por lo cual considero que se encuentra APTO para continuar con los trámites estipulados en el Reglamento de Grados y Títulos de la Facultad

FIRMA

FECHA 09-10-19

DICTAMINADORES BORRADOR DE TESIS:

- 1) **Dr. Jaime Cárdenas García**
- 2) **Dr. Fredy Molina Rodríguez**
- 3) **Mgter. Jaime Barreda del Carpio**

DICTAMEN FINAL: Señor Decano de la Facultad de Ciencias Farmacéuticas, Bioquímicas y Biotecnológicas, atendiendo a su designación como Dictaminadores del presente Borrador de Tesis sugiriendo se cambie el título a: **"EVALUACION DE LAS INTERACCIONES MOLECULARES DEL COMPLEJO PROTEINICO ERP, Rv1417 Y Rv2617c DEL Mycobacterium tuberculosis Y SU POTENCIAL OBJETIVO FRENTE A DERIVADOS DE LA NICOTINA UTILIZANDO TECNICAS DE MODELAMIENTO COMPUTACIONAL"** y luego de hechas las observaciones y correcciones pertinentes, cumpliendo con las exigencias mínimas establecidas para un trabajo de investigación de Tesis profesional, por lo que consideramos APTO para continuar con los trámites estipulados en el Reglamento de Grados y Títulos de la Facultad

Atentamente

FECHA 12-11-19

JURADOS PRESIDENTE **DR. JAIME CARDENAS GARCIA**
VOCAL **DR. FREDY MOLINA RODRIGUEZ**
SECRETARIO **MAG. JAIME BARREDA DEL CARIPO**

FECHA 20/11/19

HORA 19:00

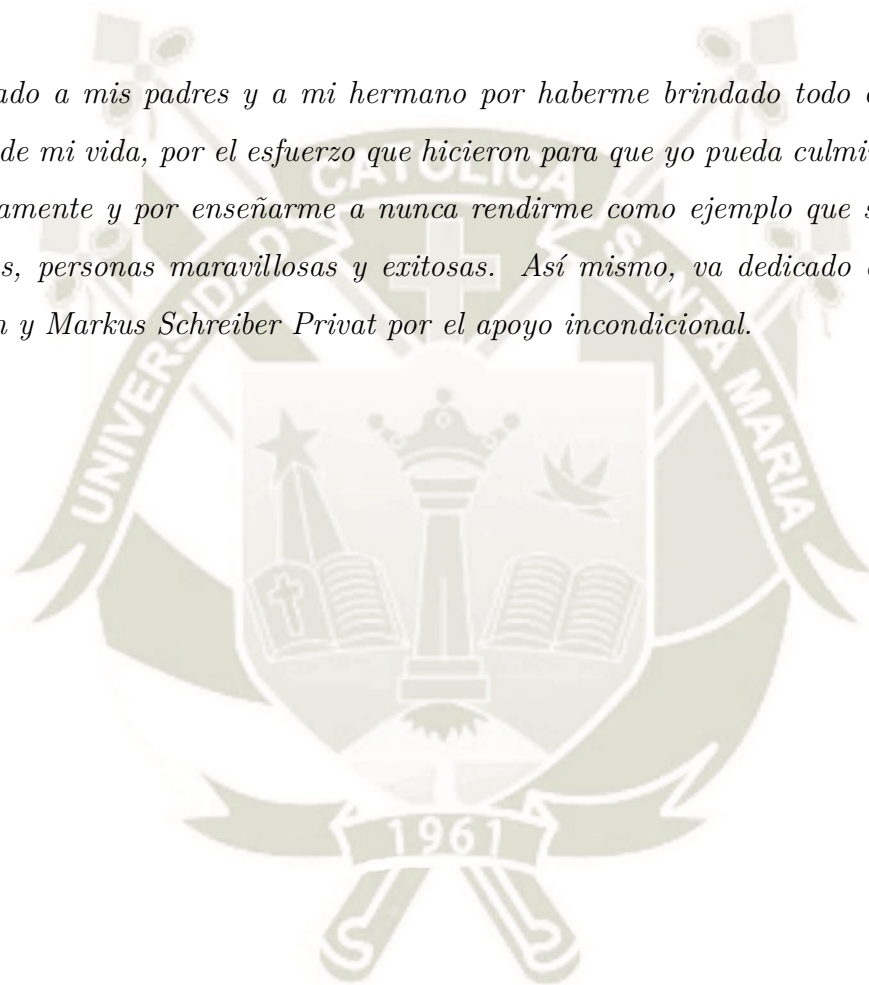
LOCAL C-402

FIRMA DEL DECANO

FECHA

Dedicatoria

Dedicado a mis padres y a mi hermano por haberme brindado todo el apoyo en esta etapa de mi vida, por el esfuerzo que hicieron para que yo pueda culminar mis estudios exitosamente y por enseñarme a nunca rendirme como ejemplo que son cada uno de ustedes, personas maravillosas y exitosas. Así mismo, va dedicado a Marcos Degen Dublin y Markus Schreiber Privat por el apoyo incondicional.



Agradecimientos

Primeramente, un agradecimiento sincero a mi asesor Dr. Badhin Gómez Valdez, director del Centro de Investigación en Ingeniería Molecular (CIIM), por permitirme realizar la tesis en el área de la bioinformática y brindarme el conocimiento necesario para su culminación. De la misma forma agradezco a Haruna Barazorda Cahuana y Diego Valencia Chambi por el apoyo para la realización de este proyecto, por aclararme las dudas puestas en ella y por brindarme su amistad incondicional.

También, un agradecimiento al Vicerrectorado de Investigación por el financiamiento del proyecto con resolución N° 25175-R-2018 y a todos los integrantes del CIIM de esta universidad por el apoyo y los gratos momentos que pasé durante el desarrollo de mi tesis.

Contenido

Contenido	iv
Índice de Figuras	vii
Índice de Tablas	x
Glosario	xii
Resumen	xiv
Abstract	xv
Introducción	xvi
Objetivos	xviii
Hipótesis	xix
1 Marco Teórico	1
1.1 Tuberculosis	1
1.1.1 Fisiopatología	1
1.1.2 Diagnóstico y tratamiento	3
1.2 <i>Mycobacterium tuberculosis</i>	4
1.2.1 Genoma	4

1.2.2	Variación de cepas y resistencia a antibióticos	6
1.2.3	Factores de Virulencia	7
1.3	Derivados de la Nicotina	11
1.4	Bioinformática	13
1.4.1	Modelamiento de Proteínas por Homología	13
1.4.2	Acoplamiento Molecular	15
1.5	Mecánica Molecular	16
1.5.1	Dinámica Molecular	16
1.6	Mecánica Cuántica	17
1.6.1	Indices de Reactividad Local y Global	17
2	Herramientas Computacionales y Metodología	19
2.1	Herramientas Computacionales	19
2.1.1	Equipo Computacional	19
2.1.2	Base de Datos	19
2.1.3	Softwares	20
2.1.4	Servidores	20
2.2	Metodología	22
2.2.1	Preparación de los Derivados de la Nicotina	22
2.2.2	Preparación de las proteínas Erp, Rv1417 y Rv2617c	23
2.2.3	Interacción Molecular	26
2.2.4	Análisis de las Interacciones Moleculares	27
3	Resultados y Discusión	28
3.1	Derivados de la Nicotina	28
3.2	Análisis cuántico de las estructuras moleculares de los Derivados de la Nicotina	28

3.3	Análisis de la Simulación de Dinámica Molecular	37
3.3.1	Validación de las estructuras de las proteínas Erp, Rv1417 y Rv2617c	39
3.3.2	Predicción de los bolsillos	49
3.3.3	Interacción de los Derivados de la Nicotina con las proteínas Erp, Rv1417 y Rv2617c	51
3.3.4	Análisis de dinámica molecular del acoplamiento	52
3.3.5	Análisis de la estructura secundaria de las tres proteínas	55
	Referencias Bibliográficas	68
A	Energías Corregidas para el Cálculo del Potencial de Ionización (PI) y Afinidad Electrónica (AE).	84
B	Predicción de Acoplamiento en PatchDock	86
C	Interacciones Moleculares de las Proteínas y los Derivados de la Nicotina	95

Índice de Figuras

1.1	Respuesta latente y activa del <i>Mycobacterium tuberculosis</i> dentro del macrófago alveolar.	2
1.2	Esquema de la envoltura celular del <i>Mycobacterium tuberculosis</i> y sus componentes.	5
1.3	Mapa circular del <i>Mycobacterium tuberculosis</i> cepa H37Rv.	6
1.4	Repeticiones PGLTS de la secuencia de la Erp.	9
1.5	Organización Genómica de la proteína Rv1417 homóloga en diferentes especies.	12
1.6	Pasos para realizar un modelado de proteína.	14
3.1	Derivados de la nicotina optimizados.	30
3.2	Potencial electrostático de los derivados de la nicotina.	31
3.3	Orbitales moleculares más altos ocupados (HOMO) de los derivados de la nicotina.	32
3.4	Orbitales moleculares más bajos desocupados (LUMO) de los derivados de la nicotina.	33
3.5	Función de Fukui susceptible a un ataque electrofílico de los derivados de la nicotina.	34
3.6	Función de Fukui susceptible a un ataque nucleofílico de los derivados de la nicotina.	35
3.7	Secuencias en formato FASTA de las tres proteínas.	37

3.8	Cambios conformacionales estructurales de las tres proteínas.	38
3.9	Diagrama de Ramachandran de la Erp.	40
3.10	Diagrama de Ramachandran de la proteína Rv1417.	41
3.11	Diagrama de Ramachandran de la proteína Rv2617c.	42
3.12	Análisis de calidad de la estructura tridimensional de Erp.	43
3.13	Análisis de calidad de la estructura tridimensional de la proteína Rv1417.	44
3.14	Análisis de calidad de la estructura tridimensional de la proteína Rv2617c.	45
3.15	RMSD de las proteínas Erp, Rv1417 y Rv2617c.	46
3.16	RMSF de las proteínas Erp, Rv1417 y Rv2617c.	47
3.17	Radio de Giro de las proteínas Erp, Rv1417 y Rv2617c.	48
3.18	Enlaces de Hidrógeno de las proteínas Erp, Rv1417 y Rv2617c.	49
3.19	Bolsillos para las tres proteínas.	50
3.20	RMSD de la interacción de las tres proteínas con cada derivado de la nicotina.	53
3.21	RMSF de la interacción de las tres proteínas con cada derivado de la nicotina.	54
3.22	Aminoácidos interactuantes de la proteína Erp cada derivado.	56
3.23	Aminoácidos interactuantes de la proteína Rv1417 con cada derivado.	57
3.24	Aminoácidos interactuantes de la proteína Rv2617c con cada derivado.	58
3.25	Análisis de la estructura secundaria de las tres proteínas con el derivado A.	59
3.26	Análisis de la estructura secundaria de las tres proteínas con el derivado B.	60
3.27	Análisis de la estructura secundaria de las tres proteínas con el derivado C.	61
3.28	Análisis de la estructura secundaria de las tres proteínas con el derivado D.	62

3.29	Análisis de la estructura secundaria de las tres proteínas con el derivado	
E.	63
C.1	Interacción de la proteína Erp con cada derivado de la nicotina.	96
C.2	Interacción de la proteína Rv1417 con cada derivado de la nicotina. . .	97
C.3	Interacción de la proteína Rv2617c con cada derivado de la nicotina. . .	98



Índice de Tablas

1.1	Proteínas de la Membrana Celular y sus funciones del <i>Mycobacterium tuberculosis</i> estudiado por Chatterjee y colaboradores	11
2.1	Parámetros de trabajo para la simulación con el método "Simulated Annealing".	25
3.1	Estructura y Nomenclatura Química de los Derivados de la Nicotina.	29
3.2	Energías de los orbitales de frontera de los derivados de la nicotina.	35
3.3	Índices de reactividad global de los derivados de la nicotina.	36
3.4	Datos del gráfico de Ramachandran de las tres proteínas estudiadas.	43
3.5	Datos de la predicción de bolsillos de las tres proteínas.	51
3.6	Datos del Acomplamiento Molecular obtenidas en el servidor PATCHDOCK.	52
3.7	RMSD, RMSF, Radio de Giro y Enlaces de hidrógeno de las tres proteínas estudiadas.	55
3.8	Comparación de la estructura secundaria de las proteínas y la interacción entre proteínas y derivados	64
3.9	Aminoácidos interactuantes de cada proteína con los derivados de la nicotina.	65
A.1	Energías corregidas en los estados neutro, catión y anión de los derivados de la nicotina.	85
B.1	Predicción del acoplamiento de la Erp con el derivado A.	87

B.2	Predicción del acoplamiento de la Erp con el derivado B.	87
B.3	Predicción del acoplamiento de la Erp con el derivado C.	88
B.4	Predicción del acoplamiento de la Erp con el derivado D.	88
B.5	Predicción del acoplamiento de la Erp con el derivado E.	89
B.6	Predicción del acoplamiento de la proteína Rv1417 con el derivado A. .	89
B.7	Predicción del acoplamiento de la proteína Rv1417 con el derivado B. .	90
B.8	Predicción del acoplamiento de la proteína Rv1417 con el derivado C. .	90
B.9	Predicción del acoplamiento de la proteína Rv1417 con el derivado D. .	91
B.10	Predicción del acoplamiento de la proteína Rv1417 con el derivado E. .	91
B.11	Predicción del acoplamiento de la proteína Rv2617c con el derivado A.	92
B.12	Predicción del acoplamiento de la proteína Rv2617c con el derivado B.	92
B.13	Predicción del acoplamiento de la proteína Rv2617c con el derivado C.	93
B.14	Predicción del acoplamiento de la proteína Rv2617c con el derivado D.	93
B.15	Predicción del acoplamiento de la proteína Rv2617c con el derivado E. .	94

Glosario

1. OMS = Organización Mundial de la Salud.
2. IFN = Interferón.
3. BAAR = Bacilos Alcohol Ácidos Resistentes.
4. Erp = Export Repetitive Protein.
5. CAM-B3LYP = Coulomb Attenuated Method - B3LYP, hidrido de intercambio y correlación funcional.
6. TZVP = Valence triple-zeta polarization.
7. NCBI = National Center for BIotechnology Information.
8. PDB = Protein Data Bank.
9. DFT = Density Functional Theory.
10. HOMO = Highest Occupied Molecular Orbital.
11. LUMO = Lowest Unoccupied Molecular Orbital.
12. RMSD = Root Mean Square Derivation.
13. RMSF = Root Mean Square Fluctuation.
14. ACE = Atomic Contact Energy.
15. ESP = Eelectrostatic Potential.
16. nm = nano-metro.
17. ns = nano-segundo.

18. ps = pico-segundo.
19. aa = aminoácido.
20. ALA = Alanina.
21. THR = Treonina.
22. PHE = Fenilalanina.
23. LEU = Leucina.
24. SER = Serina.
25. GLY = Glicina.
26. PRO = Prolina.
27. GLN = Glutamina.
28. ASP = Aspartato.
29. ARG = Arginina.
30. TYR = Tirosina.
31. VAL = Valina.
32. ILE = Isoleucina.
33. HIS = Histidina.
34. GLU = GLutamina.
35. TRP = Triptófano.
36. Å = Amstrong.
37. Score = Puntuación.
38. Loops = Regiones que conectan las estructuras secundarias de una proteína.

Resumen

La tuberculosis es una enfermedad alarmante a nivel mundial provocada principalmente por el *Mycobacterium tuberculosis*, motivo por el cual está siendo muy estudiada, descubriéndose cada vez nuevas formas de contrarrestarla. El objetivo del presente trabajo fue evaluar las interacciones moleculares de las proteínas que forman el complejo Erp, Rv1417 y Rv2617c del *Mycobacterium tuberculosis* frente a derivados de la nicotina utilizando técnicas de modelamiento computacional. Para su realización, se modelaron y minimizaron las tres proteínas mediante técnicas de dinámica molecular. Cada proteína se interaccionó con cinco derivados de la nicotina, a estos sistemas se les realizó una pequeña dinámica molecular y se analizaron las estructuras secundarias comparando la estructura de la proteína inicial y después de su interacción, teniendo como resultados cambios en su estructura secundaria de las proteínas con los derivados. Con respecto a la Erp se observa mayores fluctuaciones al interaccionar con el derivado E seguido de los derivados A y B, en cambio con la proteína Rv2617c se observa cambios de fluctuación con el derivado B, desapareciendo las láminas beta. Por otro lado, las láminas beta de la proteína Rv1417 se muestran disminuidas con los derivados B, C, D y E, pero se observa una mayor fluctuación y desestabilización de la proteína con el derivado A. En conclusión las proteínas llegan a alterar su estructura secundaria y a desestabilizarse al interaccionar especialmente con los derivados de la nicotina A, B y E.

Keywords: *Mycobacterium tuberculosis*, factores de virulencia, derivados de la nicotina, dinámica molecular, interacción, desestabilización

Abstract

Tuberculosis is an alarming disease worldwide caused mainly by the *Mycobacterium tuberculosis*, which is why it is being studied, discovering new ways to counteract it. The objective of this work was to evaluate the molecular interactions of the proteins that form the Erp, Rv1417 and Rv2617c complex of *Mycobacterium tuberculosis* against nicotine derivatives using computational modeling techniques. For its realization, the three proteins were modeled and minimized by molecular dynamics techniques. Each protein was interacted with five nicotine derivatives, these systems had a small molecular dynamics and secondary structures were analyzed comparing the structure of the initial protein and after their interaction, resulting in changes in their secondary structure. proteins with derivatives. With respect to the Erp, greater fluctuations are observed when interacting with derivative E followed by derivatives A and B, instead with Rv2617c protein, fluctuation changes are observed with derivative B, disappearing beta sheets. On the other hand, the beta sheets of the Rv1417 protein are shown to be diminished with derivatives B, C, D and E, but there is a greater fluctuation and destabilization of the protein with derivative A. In conclusion, proteins can alter their structure secondary and destabilize by interacting especially with nicotine derivatives A, B and E.

Keywords: *Mycobacterium tuberculosis*, virulence factors, nicotine derivatives, molecular dynamics, interaction, destabilization

Introducción

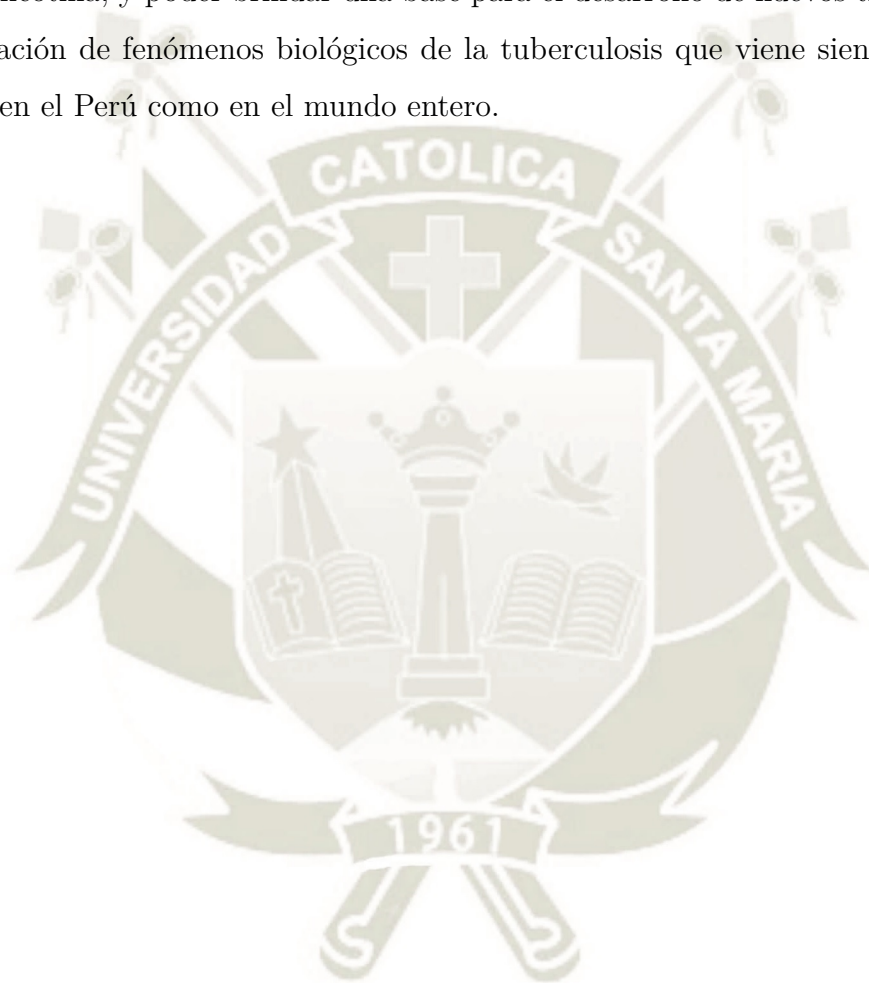
La tuberculosis es una enfermedad infecciosa causada principalmente por el *Mycobacterium tuberculosis* que tiene implicaciones por todo el mundo. Uno de los problemas más grandes que presenta la tuberculosis es la resistencia a los antibióticos, por lo que es necesario la investigación de nuevos blancos terapéuticos y el desarrollo de fármacos que nos permitan atacar el problema desde otra perspectiva.¹

En los últimos años, diversos estudios vienen demostrando la existencia de diferentes genes vinculados a la virulencia de la micobacteria, permitiéndoles sobrevivir y persistir en los hospedadores. La mayoría de estos genes codifican enzimas de la vía lipídica y proteínas reguladoras de los sistemas de transducción de señales. Uno de estos genes de virulencia es la Erp (*Exported Repetitive Protein*), la cual se ha demostrado que tiene una relación directa con la viabilidad de la bacteria, teniendo un rol importante en la síntesis de la pared celular.² En el estudio de Klepp y colaboradores se ha demostrado que, junto a esta proteína, también se encuentran la Rv1417 y la Rv2617c formando un complejo necesario para cumplir como factor de virulencia a través de múltiples interacciones siendo específica en la especie *Mycobacterium tuberculosis*.

La Erp es una proteína que se ha encontrado tanto intracelular como en regiones extracelulares,³⁻⁵ por estas razones no se sabe con exactitud su localización. Mientras que las proteínas Rv1417 y Rv2617c, fueron localizadas en la membrana celular.⁶ A pesar de la importancia de estas proteínas, los reportes científicos sobre ellos son escasos, teniendo la esperanza de encontrar un gran potencial para la identificación de un objetivo terapéutico; por lo que se podrían realizar ensayos con moléculas que han tenido efectividad de forma experimental, como derivados de la nicotina a los que se les ha quitado la capacidad adictiva.

El presente estudio se desarrolla dentro del área de la bioinformática, química,

física y biología computacional, áreas que vienen creciendo a la par con el desarrollo tecnológico en la búsqueda de nuevas estrategias terapéuticas para enfermedades de todo tipo. Por lo que se busca brindar una descripción de las estructuras 3D y un acercamiento al comportamiento de las proteínas Erp, Rv1417 y Rv2617c frente a los derivados de la nicotina sin capacidad adictiva, evaluando de esta manera si estas proteínas pueden ser alteradas, lo que conllevaría a la desestabilización por los derivados de la nicotina, y poder brindar una base para el desarrollo de nuevos tratamientos y la explicación de fenómenos biológicos de la tuberculosis que viene siendo un problema tanto en el Perú como en el mundo entero.



Objetivos

General

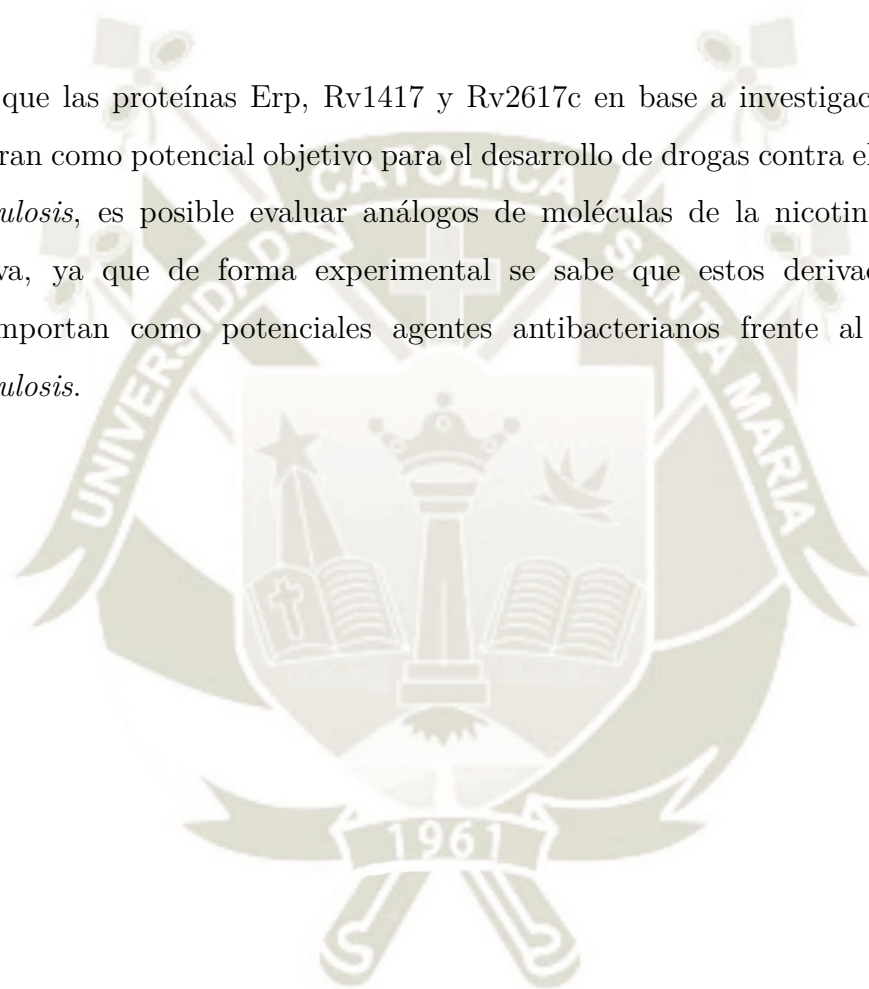
Evaluar las interacciones moleculares de las proteínas que forman el complejo Erp, Rv1417 y Rv2617c del *Mycobacterium tuberculosis* frente a derivados de la nicotina utilizando técnicas de modelamiento computacional para brindar una descripción y un acercamiento al comportamiento de dichas proteínas.

Específicos

- Obtener las estructuras moleculares correspondientes a los derivados de la nicotina y de las proteínas Erp, Rv1417 y Rv2617c del *Mycobacterium tuberculosis* en su estado tridimensional.
- Minimizar y estabilizar mediante técnicas de mecánica molecular y mecánica cuántica las proteínas Erp, Rv1417 y Rv2617c y los derivados de la nicotina respectivamente.
- Realizar la simulación de dinámica molecular, con un ensamble NPT de las proteínas Erp, Rv1417 y Rv2617c a condiciones fisiológicas.
- Realizar las interacciones moleculares entre las moléculas de los derivados de la nicotina y las proteínas Erp, Rv1417 y Rv2617c.

Hipótesis

Dado que las proteínas Erp, Rv1417 y Rv2617c en base a investigaciones previas se muestran como potencial objetivo para el desarrollo de drogas contra el *Mycobacterium tuberculosis*, es posible evaluar análogos de moléculas de la nicotina sin capacidad adictiva, ya que de forma experimental se sabe que estos derivados de nicotina se comportan como potenciales agentes antibacterianos frente al *Mycobacterium tuberculosis*.



Capítulo 1

Marco Teórico

1.1 Tuberculosis

La tuberculosis es una enfermedad causada principalmente por el *Mycobacterium tuberculosis* ocasionando la muerte e infección a un tercio de la población mundial, por el cual, se ha observado en los últimos años un incremento de investigaciones sobre su epidemiología, factores de riesgo, fisiopatología, diagnósticos y nuevos tratamientos para contrarrestar esta enfermedad.⁷ La tuberculosis ha ido prevaleciendo durante muchos años, esto lleva a la OMS poner fin a esta epidemia entre unos 15 años.^{8,9} Una de las propuestas por la OMS es enfocarse en las personas con tuberculosis latente, ya que no presenta síntomas y pueden vivir como parásitos dentro del cuerpo humano durante muchos años, los estudios mencionan que de cada cuatro personas una se encuentra infectada con tuberculosis latente,¹⁰ por lo que es un objetivo clave enfocarse en las personas que desarrollan la infección con tuberculosis latente.

1.1.1 Fisiopatología

Para que se dé una infección por el *Mycobacterium tuberculosis* es necesario una fuente de transmisión, a partir de ésta la micobacteria se propagará a través de gotas de aire ya sea tosiendo, estornudando o hablando,¹¹ por lo tanto, la transmisión es de manera aerógena ya que en la gran mayoría de casos ocurre mediante la inhalación de las secreciones respiratorias. Un pequeño porcentaje del total pasará por los

pulmones llegando hasta los alvéolos, y el resto serán atrapados por las vías respiratorias superiores siendo expulsadas por células de la mucosa. Una vez instaladas en los alvéolos serán fagocitadas por los macrófagos, que serán incapaces de matar y digerir a la micobacteria, dando así dos respuestas durante el primer mes, en la primera las micobacterias podrán atenuarse, evitando así su multiplicación durante un periodo de tiempo muy largo; como segunda respuesta, debido a un sistema inmune deficiente, las micobacterias se seguirán multiplicando hasta volverse activas, ocasionando una lisis celular del macrófago alveolar para continuar con su proliferación (ver figura 1.1).¹²

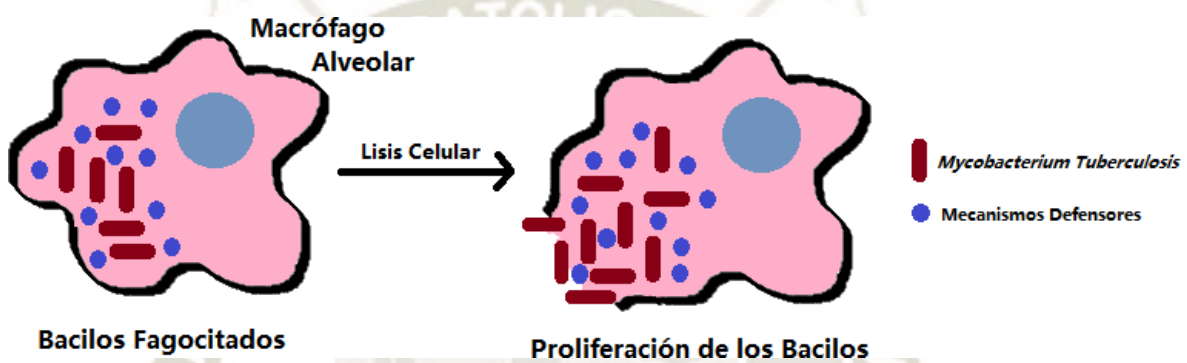


Figura 1.1: Respuesta latente y activa del *Mycobacterium tuberculosis* dentro del macrófago alveolar.

Una vez que las micobacterias se encuentren activas, las siguientes respuestas se caracterizan por la aparición de lesiones granulomatosas llamados tubérculos. Los macrófagos se volverán a reactivar estimulando la producción de linfocitos creando un ambiente necrótico debido a la escasa presencia de oxígeno y un pH bajo, así como también se dará una respuesta del IFN tipo I que como respuesta ante la enfermedad inhibirá a los IFN tipo II,^{13,14} de esa manera tratarán de contrarrestar su proliferación. El *Mycobacterium tuberculosis* también puede proliferar por diferentes partes del cuerpo como: los ganglios linfáticos y/o biliares, vasos linfáticos, el cerebro y riñones ocasionando finalmente la muerte.⁷

La prevalencia del *Mycobacterium tuberculosis* en el hombre se da debido a sus factores de virulencia, hasta el momento se conoce tres principales acciones dentro del macrófago: la inhibición de la acidificación del ambiente, la inhibición de la formación del fagolisosoma (el más conocido y estudiado) y el dominio que tiene el *Mycobacterium*

tuberculosis hacia la respuesta inmune humana mediante la inhibición de la formación del interferón gamma que aumenta la capacidad fagocitaria del macrófago.^{15,16}

1.1.2 Diagnóstico y tratamiento

Cada año se realiza nuevas investigaciones para encontrar métodos de diagnóstico que sean simples y confiables para determinar la enfermedad. Primeramente se empezó con diagnósticos bacteriológicos como la microscopía y el cultivo de esputo,¹⁷ luego aparecieron los diagnósticos imageneológicos como la radiografía convencional (detección a través del tórax), tomografía computarizada para una detección masiva y la resonancia magnética para partes específicas del cuerpo,^{18,19} también se desarrollaron diagnósticos con pruebas serológicas que consiste en la detección de anticuerpos en muestras de sangre, centrándose en respuestas del *Mycobacterium tuberculosis* sobre el huésped que en la detección de la misma.²⁰ Por otro lado, con los estudios del genoma del *Mycobacterium tuberculosis* se demostraron que la inhibición de la actividad del lipoarabinomano (LAM) de la micobacteria disminuye su proliferación, esto llevó a realizar diagnósticos de LAM en la orina acompañado del recuento de CD4, resultando no ser confiable y sensible. También se desarrollaron biomarcadores, pruebas moleculares, pruebas genotípicas basadas en la amplificación del ácido nucleico, ensayos con sondas de línea e identificación de mutaciones en el genoma del *Mycobacterium tuberculosis*,⁷ pero estos ensayos requieren de un buen equipamiento de laboratorio.

El desafío que ha perdurado durante todos estos años ha sido poder diferenciar, a través de las pruebas, a personas con *Mycobacterium tuberculosis* latente y activa. Así también, dentro de su diagnóstico se están evaluando las resistencias que poseen estas micobacterias hacia los medicamentos.²¹ Esta resistencia es causada muchas veces por el tiempo prolongado de tratamiento que se les da; se sabe que los medicamentos de primera línea utilizadas son: la isoniazida, rifampicina, pirazinamida y etambutol, teniendo un tratamiento primero de 2 meses, seguido de los dos primeros medicamentos mencionados durante 4 meses, haciendo un total de 6 meses.^{22,23} Muchas veces no se les hace un seguimiento a las personas con tratamiento por lo que tienden a abandonarlo y de esta manera desarrollan resistencia, llevando al desarrollo de nuevos blancos terapéuticos como son las debaquilinas y delamanidas,⁷ entre otros.

Un estudio reciente del 2018 menciona que el Perú tiende a tener retrasos para el diagnóstico de la tuberculosis conduciendo al prolongamiento de la infección e incrementando la carga de la enfermedad, dándose principalmente en pobladores de zonas rurales, especialmente mujeres.²²

1.2 *Mycobacterium tuberculosis*

El *Mycobacterium tuberculosis* es una bacteria aerobia estricta, y por este motivo le gusta depositarse en los alveólos pulmonares ya que estas se encuentran oxigenadas;²⁴ otra característica principal es su envoltura celular, conformada por una amplia gama de ácidos grasos, empezando por una membrana celular, una capa de peptidoglicanos, seguida de arabinogalactanos y ácidos micólicos, ésta última se considera importante para la supervivencia del *Mycobacterium tuberculosis* ya que posee cadenas largas de 60 a 90 átomos de carbono que le brinda protección.²⁵ La envoltura celular a su vez tiene componentes como el fosfatidilinositol manósido (PIMs), phthiocerol dimycocerosato (PDIM), diaciltrehalosa (DAT), poliaciltrehalosa (PAT), sulfolípidos (SL-1), lipomanano (LM), dimicolato de trehalosa (TDM), monomicolato de trehalosa (TMM) y lipoarabinomanano (LAM) tal como se observan en la figura 1.2. El LAM también se considera importante por comportarse como un factor de virulencia. La gran cantidad de componentes que presenta la envoltura celular, hace que el *Mycobacterium tuberculosis* se encuentre considerada dentro del grupo de los bacilos alcohol ácidos resistentes (BAAR) y para su identificación microscópica debe realizarse una tinción de Zielh-Nielsen.

1.2.1 Genoma

El primer genoma estudiado fue la cepa H37Rv que se usa comúnmente como control para la identificación del *Mycobacterium tuberculosis*.²⁶⁻²⁸ Tiene un tamaño de 4 millones de pares de bases con 3959 genes.²⁷ Actualmente el genoma del *Mycobacterium tuberculosis* aún sigue en estudio ya que aproximadamente el 60% de los genes no tienen caracterizada su función, teniendo una posible función para la mitad de estos genes. En la figura 1.3 podemos observar el genoma del *Mycobacterium tuberculosis* que se

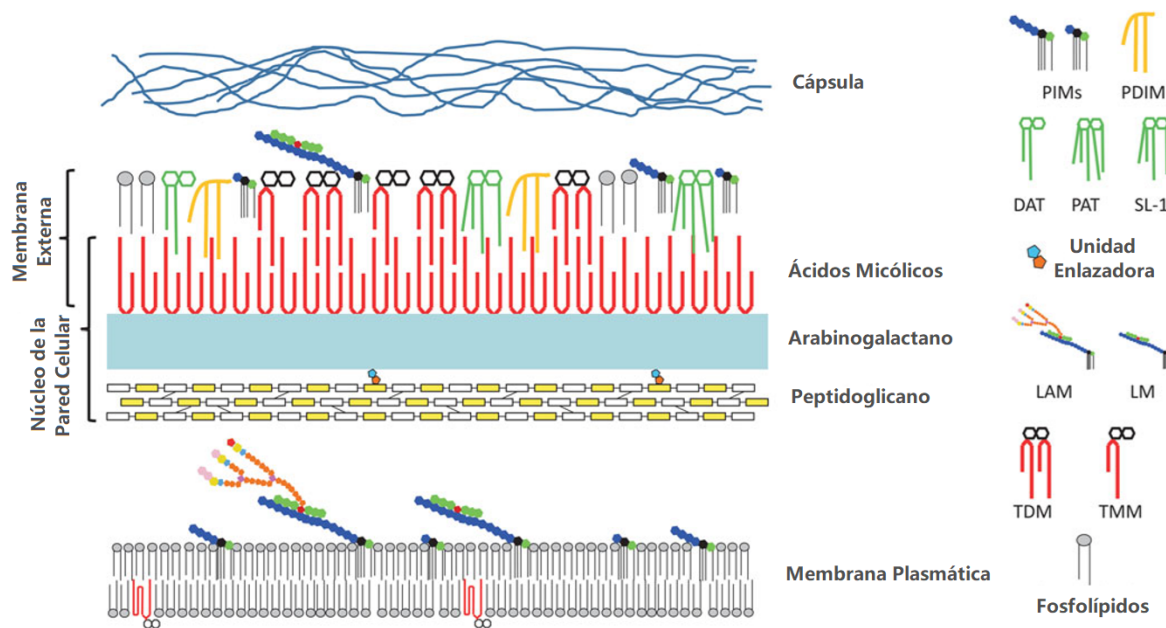


Figura 1.2: Esquema de la envoltura celular del *Mycobacterium tuberculosis* y sus componentes.²⁵

caracteriza por tener un cromosoma circular de 4.411 kpb, en el anillo central se observa un histograma de color amarillo y rojo representado el contenido de G+C, siendo <65% de color amarillo y >65% de color rojo; seguidamente se tiene las posiciones de secuencias repetitivas polimórficas (PGR) de color rojo oscuro, luego se encuentra la familia de proteínas PE (Prolina-ácido glutámico) de color púrpura seguida de la familia de proteínas PPE (prolina-prolina-ácido glutámico) de color verde. Siguiendo el orden observamos el ADN repetitivo de color naranja y la familia de las Erp de colores rosa, azul y prófago. La secuencia de codificación por hebra se muestra de color verde oscuro y claro. Finalmente se observan las posiciones de los genes de ARN estables, siendo de color azul para el ARNt y la región de repetición directa de color rosado en forma de cubo. La escala del *Mycobacterium tuberculosis* se observa en el círculo exterior, siendo “0” el origen de replicación.²⁹ En su genoma también se encuentran 4019 y 4300 marcos de lectura abiertos, así como 48 y 104 espaciadores intergénicos.³⁰

Los estudios de ingeniería genética se encuentran limitadas por la presencia de la pared gruesa que posee el *Mycobacterium tuberculosis*. Esta envoltura celular hace difícil una captación eficiente de ADN ya que al cultivarlas crecen como una película densa. Además, la exposición del laboratorista para la manipulación del *Mycobacterium tuberculosis* pone en riesgo su vida con este patógeno mortal. Esto llevó a cabo la

construcción de vectores de ADN para su estudio.³¹

Por estos motivos tanto la tecnología de secuenciación, la bioinformática y los trabajos *in silico* se ve incrementada con diferentes desarrollos de softwares e investigaciones que ayuda al análisis genómico y a entender el comportamiento de sus genes y/o proteínas presentes en el *Mycobacterium tuberculosis* en un corto periodo de tiempo y a un bajo costo.³²

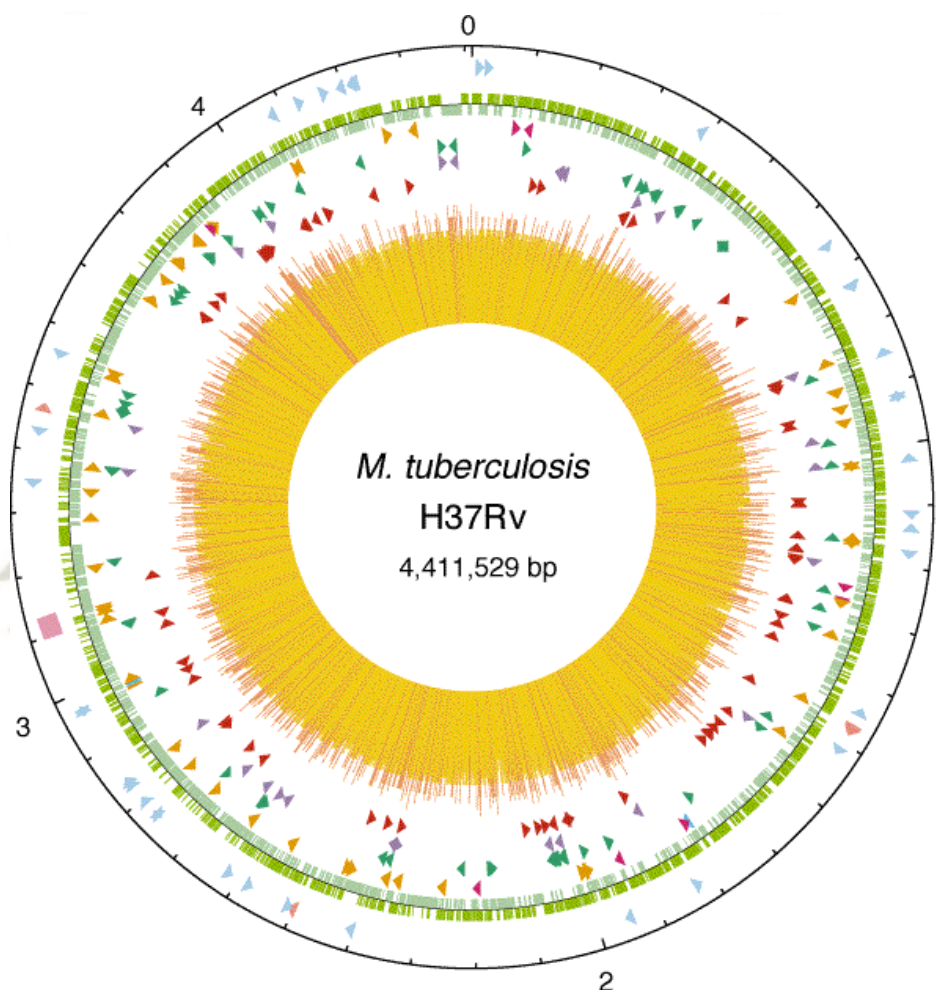


Figura 1.3: Mapa circular del *Mycobacterium tuberculosis* cepa H37Rv.²⁹

1.2.2 Variación de cepas y resistencia a antibióticos

Existe una diversidad de cepas de *Mycobacterium tuberculosis*, con una identidad muy relacionada entre sí.³³ Hasta la actualidad se ha realizado la secuenciación completa del genoma de 8 cepas: CDC1551, H37Ra, F11, KZN1435, KZN4207, CCDC5180, CCDC5079 Y CTRI-2³⁴ que revelaron una gran cantidad de datos, especialmente para

el entendimiento de sus genes de virulencia y sobre la resistencia a los antibióticos.³⁵⁻³⁹ Por otro lado, la tipificación de cepas ha tomado importancia ya que brinda evidencia ya sea a favor o en contra de la transmisión del *Mycobacterium tuberculosis*.⁴⁰

Un tratamiento estándar para la tuberculosis tiene una duración de 6 a 8 meses con antibióticos. Los medicamentos de primera línea más utilizados son: la isoniazida, rifampicina, pirazinamida y etambutol.⁴¹ Sin embargo, el prolongado tratamiento que reciben los pacientes ha llevado a problemas de adherencia y altos niveles de cepas resistentes hacia los medicamentos de primera línea, en especial a la rifampicina e isoniazida; así como también a los de segunda línea como el amikacin, kanamicina, capreomusina; y finalmente a los fluoroquinolonas y moxifloxacinas.^{42,43} Los estudios en busca de una solución demostraron que al realizar una terapia de combinación con fármacos selectivos como la sulfametoxazol, sulfadiazina, clofazimina, linezolid, minociclina, amoxicilina/ácido clavulónico y meropenem brindan una alternativa para contrarrestar al *Mycobacterium tuberculosis* resistente.⁴⁴

Lo que actualmente se está buscando es que los medicamentos actúen sobre la pared celular, proteínas de membrana y ácidos nucleicos, así como también se busca productos genéticos que ataquen directamente a los responsables de la síntesis de los genes de virulencia y reduciendo el tiempo de su tratamiento para evitar la resistencia de éstos medicamentos.^{6,45-47} Por otro lado, se busca desarrollar nuevos candidatos de vacunas como RUTI y ID93/GLA-SE para disminuir y prevenir la infección por el *Mycobacterium tuberculosis*.⁴⁸

1.2.3 Factores de Virulencia

El cuerpo humano normalmente desarrolla mecanismos de defensa para eliminar agentes extraños de nuestro organismo, pero esto no sucede cuando somos infectados por el *Mycobacterium tuberculosis*, ésta micobacteria puede modular nuestro sistema inmune, y a su vez los macrófagos no llegan a madurar el fagolisosoma, por lo tanto, no se llega a formar un ambiente ácido. Por otro lado, el *Mycobacterium tuberculosis* tiene un sistema de formación de poros que lo utiliza para romper el fagosoma liberándose al citosol, una vez liberada la micobacteria continúa con su proliferación.⁴⁹ Para poder entender el comportamiento del *Mycobacterium tuberculosis* se han estudiado a sus

factores de virulencia, tomando como centro de atención a los genes involucrados en los componentes de la envoltura celular que les da resistencia y protección; uno de éstos genes es la Erp, demostrando en diversos estudios que tras su inhibición tiende a disminuir la proliferación del *Mycobacterium tuberculosis*, de la misma forma también se consideran importantes los componentes de la membrana celular como factores de virulencia que actualmente están siendo muy estudiadas.⁴

1.2.3.1 Proteínas de la Envoltura Celular

La envoltura celular del *Mycobacterium tuberculosis* está conformada por lípidos y proteínas complejas, si nos fijamos en la figura 1.2 su núcleo se compone de peptidoglicanos unidos covalentemente a un galactofurano lineal, a su vez unido a varias cadenas de arabinofuranos altamente ramificados unidos a ácidos micólicos que son ácidos grasos α -alquil- β -hidroxi de cadena muy larga. Dentro de su núcleo se encuentran los PDIM, TDM, SL y PIM. Así como también, los fosfatidilo-miinositol manósidos, lopomanano (LM) y lipoarabinomanano (LAM) anclándose a la membrana celular y extendiéndose hacia el exterior de la pared celular.^{24,49} Las proteínas integradas en la envoltura celular están relacionadas con la síntesis y el mantenimiento de la pared celular que su vez son responsables de su adhesión, infección, transporte de solutos y supervivencia.⁵⁰⁻⁵² Todas las bacterias cuentan con una vía para secretar proteínas (Sec), este sistema se encuentra altamente conservado entre las especies de *Mycobacterium*, constituida por SecY, SecE, SecG, SecD, SecF y SecATPasa, carentes de un SecB, la conformación de este sistema de secreción es diferente a otras bacterias, y se desconoce cómo es que atraviesan las capas de la membrana celular.⁵³ Al realizar comparaciones genómicas entre las especies de *Mycobacterium* se observó que el *Mycobacterium tuberculosis* muestra regiones de diferencia (RD);⁵⁰ en los análisis de secuenciación del ADN se identificaron 14 regiones (RD1-14). La región RD1 es responsable de la virulencia que contribuye a la secreción de citosinas y también puede causar daño celular inespecífico, ésta región se encuentra ausente en algunas especies de *Mycobacterium* que resultan no ser patógenas.⁵⁴ La presencia de estas 14 regiones varía entre las diferentes especies de *Mycobacterium*.

Diversos estudios demostraron que los ácidos micólicos también son importantes

como factores de virulencia, ya que éstas poseen proteínas que en conjunto son denominadas como "Cord Factor" haciendo que el *Mycobacterium tuberculosis* crezca dentro del macrófago en forma de serpentin. Los estudios crearon cepas mutantes defectuosos en la síntesis de los ácidos micólicos como es la eliminación del *pcaA*, catalizador de la ciclopropanación proximal de los ácidos micólicos, así como también la modificación del gen *kasB*, observando una alteración morfológica del *Mycobacterium tuberculosis*; estos experimentos fueron probados en ratas infectadas dando resultados positivos, quedando la infección atenuada por la no proliferación del *Mycobacterium tuberculosis*.^{55,56}

1.2.3.2 Proteína Repetitiva Exportada (Erp)

La Erp es una proteína muy estudiada, ya que ésta ha sido de interés mucho antes de hacerse completado la secuenciación del genoma, por lo que ha sido generalmente encontrado en los sobrenadantes de los cultivos de *Mycobacterium tuberculosis*.⁴⁶ Conocida también como P36 encontrada en diferentes especies del género *Mycobacterium*. Esta proteína pesa aproximadamente 36 kDa, se caracteriza por tener dominios de repeticiones de prolina-glicina-leucina-treonina-serina (PGLTS) en su secuencia como la podemos ver en la figura 1.4 que se encuentra enmarcada de color rojo, esta proteína está comprendida por 284 aminoácidos con dominios N y C terminal que se encuentran conservados.²

GenBank: TFB22176.1

[GenPept](#) [Identical Proteins](#) [Graphics](#)

```
>TFB22176.1 exported repetitive protein [Mycobacterium tuberculosis]
MPNRRRRLSTAMSAVAALAVASPCAYFLVYESTETTERPEHHEFKQAAVLTDLPGLMSALSQGLSQFG
INIPPVPSLTGSGDASTGLTGPGLTSPGLTSPGLTSPGLTDPALTSPGLTPTLPGLSLAAPGTTLAPTPGV
GANPALTNPALTSPTGATPGLTSPGLTDPALGGANEIPITTPVGLDPGADGTYPILGDPTLGTIPSSPAT
TSTGGGGLVNDVMQVANELGASQAIIDLLKGVLMPSIMQAVQNGGAAAPAASPPVPIAAAAVPPTDPIIT
VPVA
```

Figura 1.4: Repeticiones PGLTS de la secuencia de la Erp. Secuencia disponible en: <https://www.ncbi.nlm.nih.gov/protein/TFB22176.1?report=fasta>.

Hasta el momento no se sabe con exactitud la función del Erp y se cree que las repeticiones de PGLTS se relacionan con la capacidad de virulencia del *Mycobacterium*

tuberculosis. En el estudio de Klepp se planteó la idea de que ésta región repetidora participa en la interacción con otras proteínas por lo que puede tender a formar complejos de proteínas.⁶ Por otro lado, se menciona que la región carboxi terminal se asocia a la pared celular pero no es esencial para la virulencia, demostrado en los estudios de Kocincová y colaboradores.⁴⁶

1.2.3.3 Proteínas de la Membrana Celular

Del total de proteínas presentes en el banco de datos *Protein Data Bank* (PDB) solo el 0.2% son proteínas de membrana del *Mycobacterium tuberculosis* que aún no han sido estudiadas,⁵⁷ pero se sabe que carecen de un péptido señal,⁵⁸ de los cuales sólo existen 10 estructuras cristalizadas en este banco de datos, ya que la determinación experimental de estructuras tridimensionales de proteínas de membrana es complicado de realizarlas. Por estos motivos la bioinformática y/o la biología computacional está tomando importancia para ayudar a la comprensión experimental del comportamiento de éstas proteínas, así como las interacciones proteína-ligando.⁵⁹

En la tabla 1.1 podemos observar información de 14 proteínas más estudiadas de la membrana celular del *Mycobacterium tuberculosis* con sus respectivas funciones. La mayoría de estas 14 proteínas de membrana cumplen funciones actuando como factores de virulencia, por lo que, habiendo proteínas aún no estudiadas, es posible que cumplan funciones importantes para que esta micobacteria pueda seguir proliferándose y desarrollando resistencia a los antibióticos. Esta información fue recopilada en el estudio de Chatterje y colaboradores.⁶⁰

En la membrana celular también se encuentran las proteínas quinasas, que ayudan a regular las señales de transducción, a la diferencia y división celular. El *Mycobacterium tuberculosis* codifica 11 proteínas quinasas, PknA a PknL, excepto PknC.⁵⁴ También se menciona que la PknG está involucrada en la inhibición de la formación del fagolisosoma dentro del macrófago.

Tabla 1.1: Proteínas de la Membrana Celular y sus funciones del *Mycobacterium tuberculosis* estudiado por Chatterjee y colaboradores

Proteína	Función
CpnT	Absorción de nutrientes y secreción de toxinas.
Rv0888	Potencia la multiplicación del <i>Mycobacterium tuberculosis</i> dentro del macrófago.
Rv0747	Metabolismo de los lípidos.
Rv0278c	Propaga la virulencia, actúa en la desintoxicación y ayuda a la adaptación.
Rv0916c	Ayuda a la respiración y metabolismo.
Rv1168c	Ayuda a la respiración y metabolismo.
Pe/PPE	Rol vital en la virulencia.
PPE62	Unión a pequeñas moléculas involucradas en el acceso de nutrientes.
PPE36	Utilización del grupo Hemo.
Rv2623	Homóloga de la proteína del estrés universal (USP).
Rv1747	Transportador esencial para la proliferación.
Mpt64	Proteínas de filtrado de cultivo.
CpsA	Alineación de la respuesta inmune del huésped.
Rv1738	Desactivación de la síntesis de proteínas ribosómicas.

Por otro lado, tenemos el estudio mencionado anteriormente del cual se obtiene información de un complejo formado por las proteínas Erp, Rv1417 y Rv2617c.⁶ La evaluación de la interacción de estas proteínas demostró que la proteína Rv2617c es específica para la especie *Mycobacterium tuberculosis*, mientras que la proteína Rv1417 se encuentra conservada en diferentes especies de *Mycobacterium*, encontrando que están flanqueadas por genes involucradas en la síntesis de rivo flavina, proponiendo así que ésta proteína es parte del operón de la rivo flavina (ribA2) tal como se observa en la figura 1.5.

1.3 Derivados de la Nicotina

La nicotina y sus derivados han sido utilizados para enfermedades como el alzheimer, esquizofrenia, parkinson, depresión, trastorno por déficit de atención

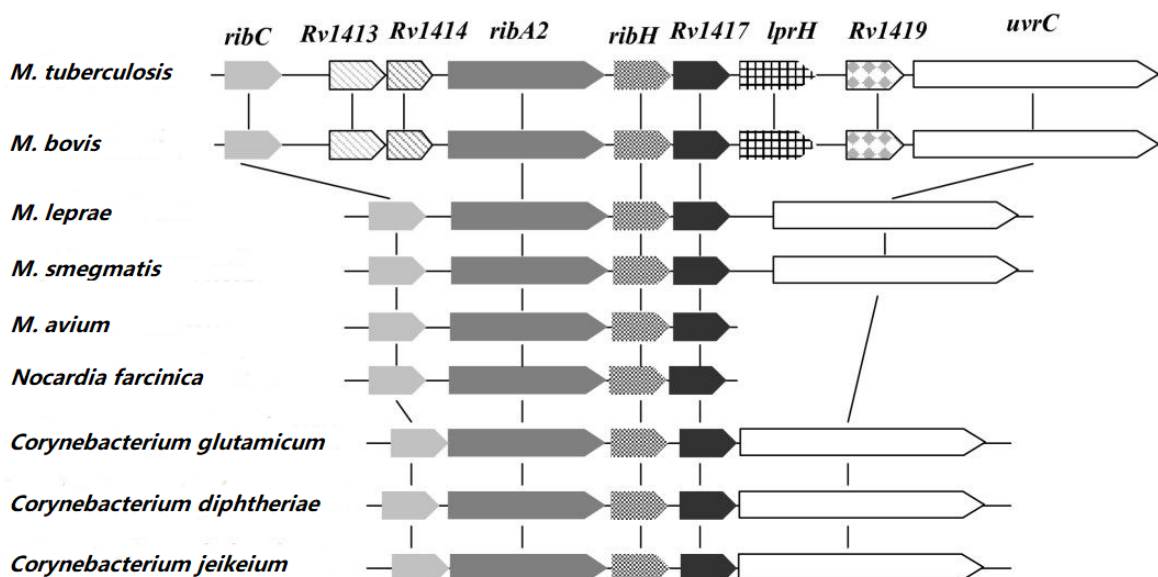


Figura 1.5: Organización Genómica de la proteína Rv1417 homóloga en diferentes especies.⁶

con hiperactividad, y recientemente ha tomado interés contra las infecciones micobacterianas.⁶¹ La piridina presente en la estructura de la nicotina ha demostrado tener actividades antioxidantes, antibacterianas, anticancerígenas y antiinflamatorias;⁶² por otro lado, la imidazol piridina carboxamidas inhibe la dihidrofolato reductasa (DHRF), que es crucial para la supervivencia del *Mycobacterium tuberculosis*.⁶³ El fármaco de primera línea utilizada para la enfermedad de la tuberculosis conocida como isoniazida es un análogo del ácido isonicotínico, isómero de ácido nicotínico, siendo así los derivados de nicotina un punto de partida para el desarrollo de nuevos fármacos.⁶⁴ Asu vez, en un estudio demostraron que los derivados de nicotina tienen un efecto inhibitorio contra el *Mycobacterium tuberculosis* añadiéndole a estos análogos la propiedad de no ser adictivos siendo así un potencial terapéutico.⁶⁵ Hoy en día se consideran importantes los estudios *in silico* de acoplamiento y dinámica molecular, ya que nos ayudan a entender el comportamiento de los fármacos, así como su rápido descubrimiento utilizando bajos recursos.⁶⁶

1.4 Bioinformática

La bioinformática está descrita como una ciencia multidisciplinaria utilizada en investigaciones biológicas mediante el uso de computadoras y fuentes masivas de almacenamiento. Actualmente la bioinformática incluye algoritmos, representación de secuencias, modelado de Markov, redes neuronales para predecir estructuras de proteínas y métodos de modelado computacional.⁶⁷ Este campo tienen sus inicios en el proyecto genoma humano donde se observa un incremento del secuenciamiento del genoma de otros seres vivos, a su vez las secuencias de proteínas, interacciones entre ellas y el descubrimiento de nuevos genes ha creado una demanda de la elaboración de base de datos para un manejo eficiente de toda ésta información.⁶⁸ La bioinformática ha ido evolucionando llegando al punto de responder preguntas biológicas que a través del análisis de un conjunto de datos complejos se pueda llegar a conclusiones biológicas válidas que puedan revelar datos desconocidos.⁶⁹ Hoy en día se buscan desarrollar nuevos métodos para un mejor manejo en el alineamiento de secuencias múltiples, la predicción y modelado de proteínas y el reconocimiento de genes eucariotas para que de esta manera se pueda responder a muchos problemas que aún están por resolverse.⁷⁰

1.4.1 Modelamiento de Proteínas por Homología

El método más utilizado para el modelamiento de proteínas es el modelamiento por homología, que consiste inicialmente en hacer un alineamiento de dos secuencias (secuencia objetivo y secuencia molde), los porcentajes aceptables de homología son desde un 50% a partir de 100 aminoácidos a más.⁷¹ A partir de la secuencias molde se puede predecir la estructura de la proteína en 3D; en los estudios de Fiser y colaboradores se mencionan 5 pasos para realizar un modelado de proteína:⁷²

1. Buscar estructuras de proteínas relacionadas entre sí.
2. Selección de uno o más moldes.
3. Alineamiento de la secuencia objetivo y de la secuencia molde.
4. Construcción de las estructuras tridimensionales de la proteína.

5. Evaluación de las estructuras tridimensionales de la proteína.

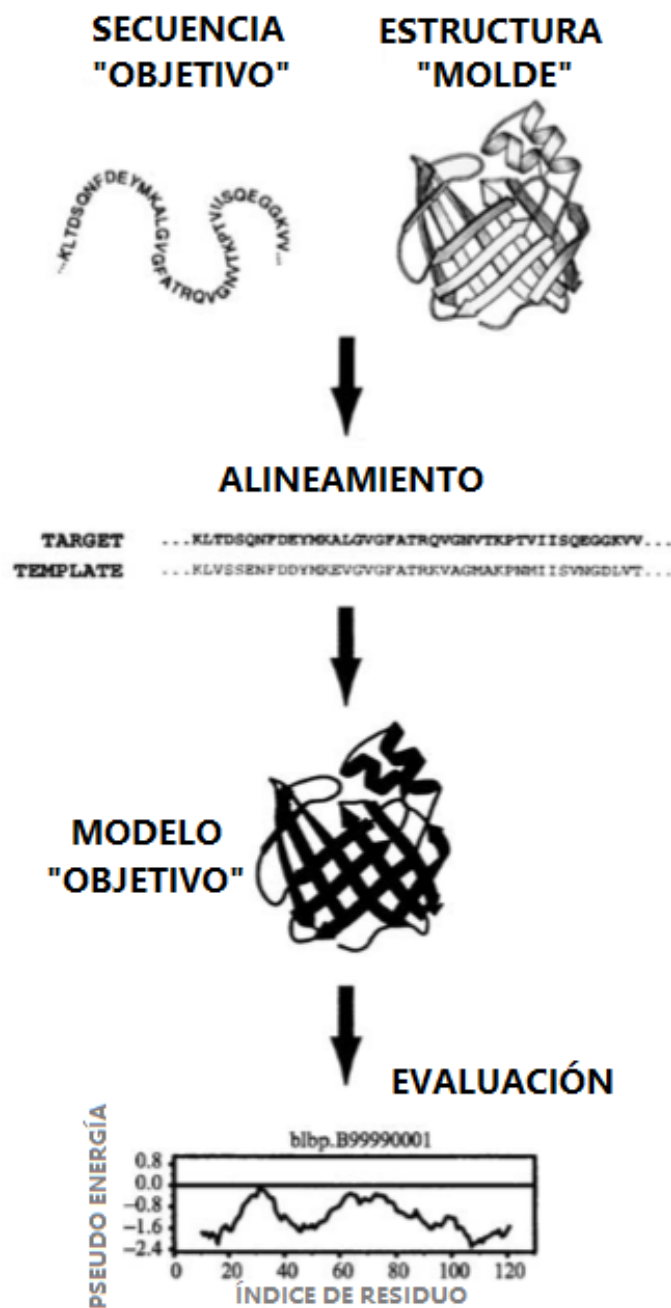


Figura 1.6: Pasos para realizar un modelado de proteína.⁷²

Por otro lado, cuando el porcentaje de homología no supera los valores aceptables es posible modelar la proteína por partes, método conocido como "Threading".⁷³ Este método consiste en tomar varios moldes que a través de su alineamiento de aminoácidos se pueda conocer la estructura secundaria, tomando cada parte coincidente de los moldes para formar finalmente la estructura tridimensional.⁷⁴

Cuando se obtiene la predicción de la estructura de la proteína en 3D se recomienda utilizar un refinamiento a la estructura, así como también se recomienda que la estructura no se encuentre atrapado en mínimos energéticos locales.⁷⁵ La metodología de "Simulated Annealing" ha sido aplicado en simulaciones de dinámica molecular para el mejoramiento de las estructuras proteínicas, éste método consiste en realizar cambios de temperatura; generalmente las simulaciones empiezan con una temperatura elevada seguido de un enfriamiento, éste procedimiento consigue que los átomos de la proteínas puedan reorganizarse encontrando así una conformación de energía más baja. En la etapa de enfriamiento es recomendable trabajar de manera gradual y lenta para que la proteína pueda tener el tiempo necesario de acomodarse.^{76,77}

1.4.2 Acoplamiento Molecular

El acoplamiento molecular, conocido como "Docking" en inglés, es una herramienta de la bioinformática utilizado en el diseño de fármacos, ya sea a través de una interacción ligando-receptor o para realizar un "virtual screening" (análisis de una biblioteca virtual).⁷⁸ Existen diferentes algoritmos que pueden predecir los acoplamientos de un complejo que pueda estimar la afinidad de unión, generalmente se toma en cuenta la baja energía de unión, ya sea de un complejo proteína-ligando (receptor-ligando) o proteína-proteína, ya que de este modo tendría mayor tiempo de contacto y estabilidad.^{79,80} Según los estudios de García-Nieto y colaboradores, existen 3 componentes que describen la interacción mediante una función de energía,⁸¹ que son:

- La traducción de los valores de los ejes "x", "y" y "z" en el espacio de coordenadas cartesianas del ligando.
- La orientación del ligando, modelada con cuatro variables (incluye la pendiente del ángulo).
- Las flexibilidades del ligando (ángulos diedros) y receptor (cadenas laterales de los aminoácidos).

Las flexibilidades tanto del ligando como del receptor son de importancia ya que de ésta forma se puede reproducir de forma nativa su geometría a su vez se explica los cambios conformacionales inducidos por la unión del ligando al receptor.⁸²

1.5 Mecánica Molecular

Existen diferentes métodos computacionales para el estudio de moléculas, una de ellas es la mecánica molecular, cuyo método es generalmente utilizado para trabajar con proteínas, DNA y RNA. Este método utiliza una base de datos para la parametrización de cualquier molécula en estudio,⁸³ estos parámetros están conformados por enlaces, estiramiento, flexión de ángulos, ángulos diédricos e interacciones no enlazantes, diseñando así los denominados "campos de fuerza".⁸⁴

Matemáticamente, los cálculos en mecánica molecular se dan a través de la energía de tensión total (gobernado por una función de energía potencial):⁸⁵

$$E_{strain} = E(\tau) + E(\theta) + E(\varphi) + E(\delta) \quad (1.1)$$

Donde $E(\tau)$ es la energía de enlace, $E(\theta)$ es la energía de la distorsión de ángulo, $E(\varphi)$ es la energía de la tensión torsional y $E(\delta)$ es la energía de las interacciones no enlazantes. A partir de los cálculos de mecánica molecular se busca dar una aproximación al comportamiento experimental de un sistema,⁸⁶ desarrollándose así a través de las simulaciones, la dinámica molecular.

1.5.1 Dinámica Molecular

La dinámica se encarga del movimiento y/o la evolución de un sistema a través del tiempo. En ésta área la dinámica molecular es un herramienta muy popular que estudia el movimiento, deformación e interacción de las moléculas a través del tiempo, siendo útil para el estudio de sistemas biológicos.⁸⁷ Las ecuaciones diferenciales de segundo orden representadas por la segunda ley de Newton hacen posible las simulaciones de dinámica molecular:

$$f_i(t) = m_i a_i(t) \quad (1.2)$$

$$f_i(t) = -\frac{\partial V(x(t))}{\partial x_i(t)} \quad (1.3)$$

Donde $f_i(t)$ es la fuerza neta que actúa sobre el átomo (i) en un punto dado en el tiempo (t), m_i es la masa y a_i es la aceleración (ecuación 1.2). Esta ecuación está

representada por el vector $x(t)$ describiendo la posición de los átomos en el espacio cartesiano x_i (ecuación 1.3).⁸⁸

Las simulaciones de dinámica molecular utilizan conjuntos termodinámicos, siendo el más práctico las simulaciones a temperatura y presión constantes utilizando en ensamble canónico NVT e isobárico-isotérmico NPT,⁸⁹ siendo N el número de partículas y T la temperatura para ambos casos, V como el volumen y P como la presión; éstos ensambles permiten que el sistema se encuentre en equilibrio.

1.6 Mecánica Cuántica

1.6.1 Índices de Reactividad Local y Global

Los índices de Reactividad se basan en la Teoría Funcional de la Densidad (*Density Functional Theory*, DFT por sus siglas en inglés) desarrollado por Hohenberg y Kohn, donde se observa que este lenguaje funcional de densidad tiene la capacidad de definir y aclarar conceptos de reactividad de estructuras moleculares, así nace el "*DFT concetual*" desarrollado por Parr y colaboradores el cual permite comprender y predecir el comportamiento químico de una molécula,⁹⁰ definiendo de esta manera los índices de reactividad local y global.

Los índices de reactividad local son: el orbital molecular más alto ocupado (*Highest Occupied Molecular Orbital*, HOMO por sus siglas en inglés), el orbital molecular más bajo desocupado (*Lowest Unoccupied Molecular Orbital*, LUMO por sus siglas en inglés) y las funciones que Fukui nucleofílica y electrofílica, que ayudan a predecir si los átomos de la molécula tienen la capacidad de ser donadores o aceptores de una cantidad de densidad electrónica.⁹¹ Para el cálculo de los índices de reactividad global por el método de aproximaciones de diferencias finitas es necesario el cálculo del potencial de ionización (PI) y la afinidad electrónica (AE).⁹² Los índices de reactividad global calculados en este trabajo de investigación fueron los siguientes:

- Electronegatividad (χ) y potencial químico (μ), estos índices de reactividad se encuentran relacionadas entre sí, ya que el potencial químico DFT mide la tendencia del movimiento de los electrones desde regiones con mayor potencial

hacia regiones de menor potencial hasta alcanzar una uniformidad. ⁹³ Siendo la electronegatividad la capacidad de un átomo de atraer electrones, el potencial químico viene a ser el equivalente al negativo de la electronegatividad, calculadas mediante la siguiente fórmula:

$$\mu = -\frac{1}{2}(PI + AE) = -\chi \quad (1.4)$$

- Dureza (η) y Blandura (s), estos índices de reactividad ayuda a explicar la estabilidad de una molécula, introducida por Pearson basándose en conceptos de la "Dureza" como estados pequeños, cargados y con baja polarizabilidad, en cambio la "Blandura" son estados grandes, poca carga y alta polarizabilidad, ^{93,94} calculadas mediante las siguientes fórmulas:

$$\eta = \frac{1}{2}(PI - AE) \quad (1.5)$$

$$s = \frac{1}{2\eta} \quad (1.6)$$

- Electrofilicidad (ω), este índice de reactividad combina la electronegatividad y la dureza, el cual mide la estabilidad de una interacción covalente por acción de un electrófilo, que puede ser importante para estudios de la unión de ligandos con sistemas bioquímicos. ⁹² calculada mediante la siguiente fórmula:

$$\omega = \frac{\mu^2}{2\eta} \quad (1.7)$$

Capítulo 2

Herramientas Computacionales y Metodología

2.1 Herramientas Computacionales

2.1.1 Equipo Computacional

El estudio se realizó en una estación computacional de trabajo de 64 GB de memoria RAM y 2 tarjetas gráficas NVIDIA GeForce GTX 1080 integradas.

2.1.2 Base de Datos

- Base de datos NCBI (*National Center for Biotechnology Information* por sus siglas en inglés), brinda información sobre la biotecnología, biología molecular, bioquímica y genética; también brinda y realiza investigaciones sobre métodos bioinformáticos para analizar estructuras, secuencias y funciones biológicamente importantes. Disponible en: <https://www.ncbi.nlm.nih.gov/>.
- Base de datos RCSB PDB (*Protein Data Bank* por sus siglas en inglés), especializada en biología estructural, proporciona acceso a datos obtenidos de forma experimental de estructuras 3D especialmente de proteínas, así como también DNA y RNA. Es una base de datos que tiene como visión "*Permitir el acceso abierto al conocimiento acumulado de la estructura, función y evolución de*

las macromoléculas en 3D, expandiendo las fronteras de la biología fundamental, la biomedicina y la biotecnología". Disponible en: <https://www.rcsb.org/>.

2.1.3 Softwares

- Gaussian 16, a través de estructuras electrónicas gaussianas sirve para modelar estructuras utilizando métodos de mecánica cuántica, prediciendo energía, frecuencias vibracionales, propiedades moleculares y reacciones químicas. Las estructuras se pueden visualizar y construir conjuntamente con el software Gaussview 5.0.9 desarrollado por su grupo.⁹⁵ Disponible en: <https://gaussian.com/gaussian16/>.
- GROMACS 2018.5, (*GRO*ningen *M*achine for *C*hemical *S*imulations por sus siglas en inglés), sirve para realizar simulaciones de dinámica molecular principalmente para proteínas, lípidos y ácidos nucleicos a través del uso de campos de fuerza. GROMACS es un software libre, integrando cada vez nuevos algoritmos, así como compiladores, funcionando de forma paralela aumentando su eficiencia y rendimiento.⁹⁶ Disponible en: <http://www.gromacs.org/>.
- Chimera UCSF 1.13.1, utilizada para el análisis y visualización interactiva de estructuras moleculares. Entre los análisis podemos observar mapas de densidad electrónica, acoplamientos, trayectorias, etc.⁹⁷ Disponible en: <https://www.cgl.ucsf.edu/chimera/download.html>.
- VMD 1.9.3. (*V*isual *M*olecular *D*ynamics por sus siglas en inglés), al igual que Chimera UCSF sirve para la visualización y análisis de estructuras moleculares. La mayor diferencia entre estos dos es que en VMD se puede trabajar de manera eficiente con sistemas que presentan grandes estructuras.⁹⁸ Disponible en: <https://www.ks.uiuc.edu/Research/vmd/>.

2.1.4 Servidores

- I-Tasser (*I*terative *T*hreading *A*SSEmbly *R*efinement por sus siglas en inglés), utilizada para el modelamiento de estructuras de proteínas tridimensionales a

través de los métodos "Threading" y "Ab initio". Luego se fueron incorporando la predicción de funciones de las estructuras modeladas. Este servidor, identifica la plantilla, busca plantillas homólogas, realiza el ensamble a través del montaje de fragmentos continuos alineados generando varias proteínas de modelado, finalmente se hace la selección del modelo y un refinamiento.⁹⁹ Disponible en: <https://zhanglab.ccmb.med.umich.edu/I-TASSER/>.

- ModRefiner (*High-resolution Protein Structure Refinement*), a través del desarrollo de su algoritmo sirve para el refinamiento realizadas a un nivel atómico de estructuras proteicas. Generalmente es utilizada para proteínas que han sido modeladas por homología y/o a partir de una estructura primaria (secuencias de aminoácidos) siendo mejoradas estructuralmente.¹⁰⁰ Disponible en: <https://zhanglab.ccmb.med.umich.edu/ModRefiner/>.
- RAMPAGE (*Rammachandran Plot Assessment*), analiza y genera un diagrama de Rammachandran de una proteína de entrada en formato PDB, cuyo resultado muestra regiones permitidas, favorecidas y no permitidas de cada residuo a través de la visualización de sus ángulos diédricos (ψ y ϕ).¹⁰¹ Disponible en: <http://mordred.bioc.cam.ac.uk/rapper/rampage.php>.
- ProSA-Web (*Protein Structure Analysis* por sus siglas en inglés), utilizada para la validación de estructuras de proteínas 3D buscando posibles errores. Este servidor reconoce los errores haciendo una comparación con estructuras experimentales disponibles, así como también muestra una gráfica de las energías que muestran los residuos.¹⁰² Disponible en: <https://prosa.services.came.sbg.ac.at/prosa.php>.
- PockDrug (*Pock Druggability Prediction* por sus siglas en inglés), sirve para la predicción de bolsillos ("*Pockets*") de proteínas que tengan alta capacidad de drugabilidad. El servidor te muestra un ranking de posibles bolsillos con archivos en formato PDB de salida y una tabla de datos.¹⁰³ Disponible en: <http://pockdrug.rpbs.univ-paris-diderot.fr/cgi-bin/index.py?page=Druggability>.
- PATCHDOCK (*Molecular Docking Algorithm BAse don Shape Complementarity Principles*), algoritmo utilizado para el acoplamiento molecular de dos sistemas

(ligando-receptor). Consiste en la entrada del receptor y del ligando en formatos PDB generando un ranking de 10 a más interacciones y una tabla de datos de todas las interacciones encontradas con sus respectivos archivos de salida en PDB.^{104,105} Disponible en: <https://bioinfo3d.cs.tau.ac.il/PatchDock/index.html>.

- PDBsum, muestra muchas aplicaciones. Para este estudio se utilizó el servidor para analizar la estructura secundaria de la proteína después de su acoplamiento e interacción, así como también el análisis de los aminoácidos interactuantes y la formación de enlaces de la proteína con el ligando a través del diagrama LIGPLOT.¹⁰⁶ Disponible en: <http://www.ebi.ac.uk/thornton-srv/databases/pdbsum/Generate.html>.

2.2 Metodología

2.2.1 Preparación de los Derivados de la Nicotina

2.2.1.1 Optimización de las estructuras

Primeramente, las estructuras de los derivados de la Nicotina fueron obtenidas del trabajo de Gandhi y colaboradores,⁶¹ donde estas estructuras demostraron poseer actividad antimicrobiana. Se dibujaron 5 derivados de la Nicotina en el software Gaussview 5.0.9 seguidamente fueron optimizadas en el software Gaussian 16 utilizando la funcional CAM-B3LYP con una base TZVP en fase solvente-etanol, dándole de esta manera un ambiente hidrofóbico. Para la optimización de las estructuras también se realizaron cálculos de las frecuencias vibracionales para cerciorarnos que las estructuras se encuentren en un mínimo global de energía.

2.2.1.2 Cálculo de los índices de Reactividad Local y Global

Usando el modelo DFT se calcularon los índices de reactividad local (orbitales moleculares de frontera y las funciones de Fukui nucleofílica y electrofílica) y global (electronegatividad, potencial químico, dureza, blandura y electrofilicidad) con el software Gaussian 16 a través del método de aproximaciones de diferencias finitas.⁹²

Primeramente, se calcularon las energías corregidas de cada estructura en un estado de catión, anión y neutro ($N0$), a través de estas energías se calculó el potencial de ionización (PI) y la afinidad electrónica (AE) de cada estructura:

$$PI = E(N0 - 1) - E(N0) \quad (2.1)$$

$$AE = E(N0) - E(N0 + 1) \quad (2.2)$$

A través del potencial de ionización (PI) y la afinidad electrónica (AE) se empezó a calcular los siguientes índices de reactividad global:

- Electronegatividad, (χ).
- Potencial Químico, (μ).
- Dureza, (η).
- Blandura, (s).
- Electrofilicidad, (ω).

También se hicieron cálculos para los índices de reactividad local como los orbitales moleculares HOMO y LUMO, seguido de sus funciones de Fukui nucleofílica y electrofílica.

2.2.2 Preparación de las proteínas Erp, Rv1417 y Rv2617c

2.2.2.1 Modelamiento por homología

El modelamiento por homología se realizó utilizando el método "Threading" en el servidor I-Tasser. A través de los estudios de Klepp y colaboradores,⁶ se tomaron atención a tres proteínas del *Mycobacterium tuberculosis* (Erp, Rv1417 y Rv2617c). Mediante la base de datos NCBI se obtuvieron las secuencias FASTA de aminoácidos de las tres proteínas para su modelamiento.

Las secuencias FASTA se subieron al servidor I-Tasser para obtener estructuras tridimensionales de las proteínas. Para la proteína Erp se eligió el primer modelo con un C-score de 3.20 y un RMSD de 13.8 ± 3.9 Å, para la proteína Rv1417 se eligió el

segundo modelo con un C-score de -4.24 y un RMSD de $15.0 \pm 3.5 \text{ \AA}$, finalmente, para la proteína Rv2617c se eligió el segundo modelo con un C-score de -4.17 y un RMSD de $12.0 \pm 4.4 \text{ \AA}$.^{107,108} Una vez obtenidas las estructuras tridimensionales se procedió a realizar un refinamiento de las estructuras mediante el servidor ModRefiner.

2.2.2.2 Simulación de Minimización de Energía

La simulación de Minimización de energía de las tres proteínas se realizó en el software GROMACS versión 2018.5. Cada proteína se centró en una caja periódica con un tamaño de 1.5 \AA en medio solvente-etanol, simulando un ambiente hidrofóbico. Seguidamente se obtuvieron las topologías utilizando el campo de fuerza OPLS-AA para poder finalmente minimizar las estructuras con el fin de relajar la geometría de cada proteína.

2.2.2.3 Predicción de bolsillos

Con el fin de comprender la drugabilidad de las proteínas, fue el análisis de las superficies que son capaces de albergar un fármaco. Estas zonas se denominan bolsillos o pockets. Estos bolsillos fueron analizados a través del servidor PockDrug, el cual utiliza descriptores fisicoquímicos y geométricos del bolsillo. En esta investigación se eligieron las 6 primeras regiones de mayor drugabilidad para la proteína Rv2617c y las 5 primeras regiones para las proteínas Erp y Rv1417.

2.2.2.4 Simulated Annealing

A las estructuras minimizadas, se les aplicó el método "*Simulated Annealing*", utilizado para proteínas modeladas por homología a partir de una secuencia, sometiénolas a un calentamiento y enfriamiento, para que de ésta manera sea posible encontrar los mínimos globales de las proteínas.¹⁰⁹ Si bien, se ha observado reportes con mejores resultados a temperaturas de aproximadamente 300 K dándole restricciones a la proteína.^{110,111} Sin embargo, este método es específico para cada tipo de estudio, por lo que los parámetros de temperatura, tiempo, etc. pueden cambiar dependiendo del tipo de proteína.⁷⁶ En la tabla 2.1 se muestran los parámetros de trabajo utilizados

para cada proteína.

Tabla 2.1: Parámetros de trabajo para la simulación con el método "Simulated Annealing".

Proteína	Temperatura de "Annealing"	Tiempo de "Annealing"	Restricción	Barostato
Erp	310	0	NO	Berendsen
	350	30		
	325	60		
	310	300		
	100	500		
Rv1417	330	0	SI	Parrinello–Rahman
	360	100		
	330	300		
	320	500		
	100	600		
Rv2617c	310	0	SI	Parrinello–Rahman
	350	100		
	325	400		
	310	500		
	100	600		

La temperatura se encuentra reportada en Kelvin y el tiempo en pico–segundos.

La simulación se realizó periódicamente a 10 ns para las proteínas Rv1417 y Rv2617c, en cambio para la Erp se realizó 50 ns. Finalmente, se analizaron las estructuras mediante el ploteo de Ramachandran y la visualización con el software VMD versión 1.9.3.

2.2.2.5 Simulación de Dinámica Molecular

Las tres proteínas fueron sometidas a un equilibrio de ensamble canónico (NVT) a 50 ns con restricciones utilizando un termostato de tipo "*v-rescale*", y a un equilibrio de ensamble isotérmico-isobárico (NPT) a 200 ns sin restricciones utilizando un barostato de tipo "Parrinello-Rahman", para ambos ensambles se trabajó con medio solvente etanol a condiciones fisiológicas ($T = 309.65K$ y $P = 1atm$). Previamente las cargas fueron neutralizadas reemplazando moléculas del solvente por iones de Cl^- o Na^+ . El ensamble canónico nos ayudó a darle un equilibrio con enfoque al solvente y finalmente se equilibró a todo el sistema bajo condiciones de presión (P), temperatura (T) y número de partículas (N) con el ensamble NPT.^{112,113} Uno de los archivos de salida obtenidos más importante es el archivo de trayectoria, ésta indica como es que todo el sistema se comportó a través del tiempo; a partir del procesamiento de éste archivo obtuvimos las coordenadas y la información de la proteína y no de todo el sistema (proteína, solvente e iones).

Las proteínas modeladas fueron validadas con el gráfico de Ramachandran a través del servidor RAMPAGE, también se analizó la calidad de las estructuras a través del servidor ProSA-Web y finalmente, se analizaron los RMSD, RMSF, el radio de giro y el número de los enlaces de hidrógeno a través de los cálculos en GROMACS versión 2018.5.

2.2.3 Interacción Molecular

Las proteínas fueron acopladas con los derivados de la Nicotina en el servidor PATCHDOCK obteniendo como archivos de salida un ranking de 10 posibles sitios de interacción mediante un acoplamiento rígido-rígido.

Con ayuda del tamaño de los bolsillos y la alta drugabilidad se eligió la mejor interacción teniendo en cuenta la menor energía de contacto atómico (ACE) con un volumen grande.¹¹⁴ Seguidamente, a cada interacción se le realizó una dinámica molecular modificando primeramente los campos de fuerza OPLS-AA agregando información de los derivados de la nicotina a la base de datos. La simulación de dinámica molecular consistió en una minimización sometiéndolas en medio

solvente-etanol, seguido de un equilibrio de ensamble canónico a 2.5 ns y equilibrio de ensamble isotérmica-isobárica a 10 ns con un termostato tipo "*v-rescale*" y un barostato tipo "*Berendsen*" respectivamente. De la misma forma los archivos de trayectoria fueron procesados para obtener información del complejo proteína-ligando.

2.2.4 Análisis de las Interacciones Moleculares

Se compararon los análisis de los últimos 2ns del RMSD, RMSF, radio de giro y los enlaces de hidrógeno del complejo proteína-ligando y la proteína sin ligando, obtenidas a través del software GROMACS versión 2018.5. También se analizó la formación de enlaces fuertes, como los enlaces de hidrógeno de los aminoácidos interactuantes con los derivados de la nicotina, obtenidas a través del servidor PDBsum. Finalmente se comparó y analizó la estructura secundaria del complejo proteína-ligando y la proteína sin ligando, también obtenidas en el servidor PDBsum, teniendo en cuenta las fluctuaciones de cada residuo proporcionadas por el RMSF.

Cada paso de la metodología se realizó individualmente para cada proteína.

Capítulo 3

Resultados y Discusión

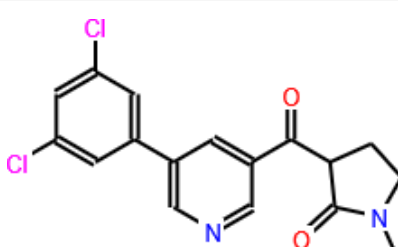
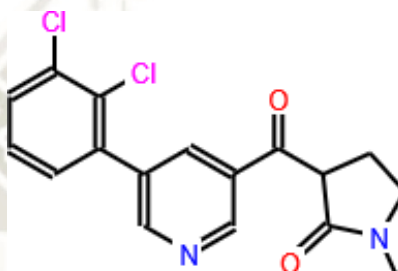
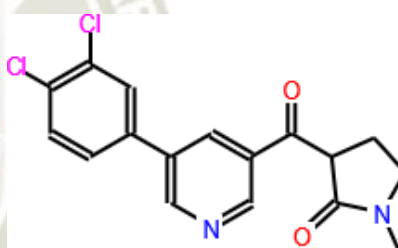
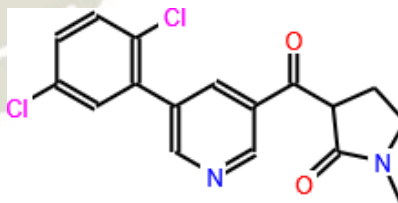
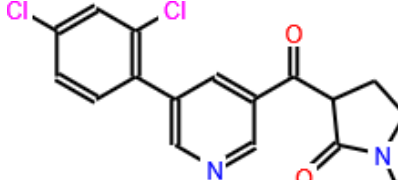
3.1 Derivados de la Nicotina

En la tabla 3.1 se muestran las estructuras con sus respectivas nomenclaturas químicas de los cinco derivados de la nicotina con las que se han trabajado, obtenidas del estudio de Gandhi y colaboradores,⁶¹ de las cuales se nombraron con letras desde la A hasta la E. Las diferencias entre los derivados de la nicotina son las posiciones de los átomos de Cloro, teniendo para los derivados "A" y "E" una sustitución "meta" en posiciones 3,5 y 2,4 en el diclorofenol respectivamente ; para los derivados "B" y "C" una sustitución "orto" en posiciones 2,3 y 3,4 en el diclorofenil respectivamente, y para el derivado "D" una sustitución "para" en la posición 2,5 en el diclorofenil.

3.2 Análisis cuántico de las estructuras moleculares de los Derivados de la Nicotina

Las estructuras de los derivados de la nicotina lograron ser optimizadas llegando a los mínimos correspondientes de energía que fueron corroboradas a través del cálculo de sus frecuencias, encontrando como resultados, ninguna frecuencia vibracional imaginaria para todas las estructuras, confirmando que estamos en los estados fundamentales de cada derivado de nicotina.

Tabla 3.1: Estructura y Nomenclatura Química de los Derivados de la Nicotina.

Derivado de Nicotina	Nomenclatura Química	Estructura Química
A	3-(5-(3,5 diclorofenil) nicotinoil)-1-metilpirrolidin-2-ona	
B	3-(3-(5-(2,3 diclorofenil) nicotinoil)-1-metilpirrolidin-2-ona	
C	3-(5-(3,4 diclorofenil) nicotinoil)-1-metilpirrolidin-2-ona	
D	3-(5-(2,5 diclorofenil) nicotinoil)-1-metilpirrolidin-2-ona	
E	3-(5-(2,4 diclorofenil) nicotinoil)-1-metilpirrolidin-2-ona	

En la figura 3.1 se encuentran representadas las estructuras optimizadas de los derivados de la nicotina, observando de color verde a los átomos de cloro, de color rojo a los átomos de oxígeno, de color azul a los átomos de nitrógeno, de color gris a los átomos de carbono y de color blanco a los átomos de hidrógeno. También se observó

que la posición del diclorofenil en el espacio tridimensional cambia para cada derivado dependiendo de la posición y sustitución en el que se encuentran los átomos de cloro en el anillo, siendo más notoria en el derivado C.

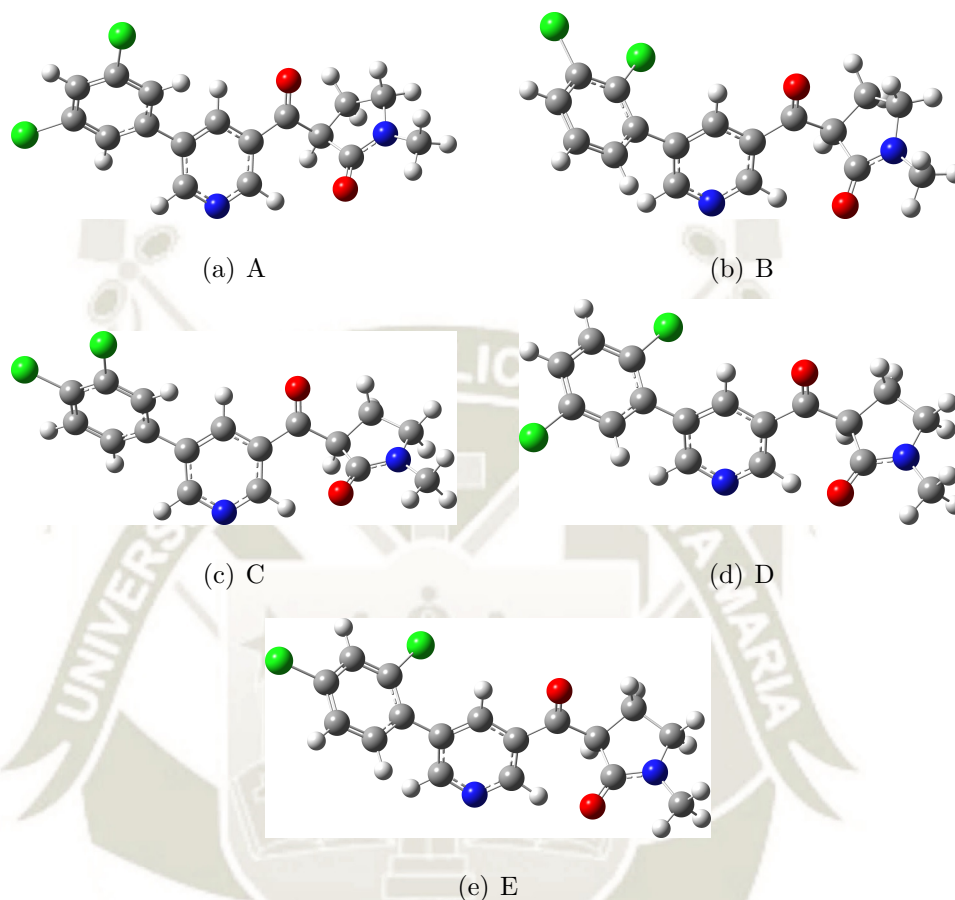


Figura 3.1: Derivados de la nicotina optimizados.

Una vez optimizadas se analizaron de los potenciales electrostáticos (conocida como ESP por sus siglas en inglés) en el que fueron graficados con un corte de isovalor de 0.04 para cada estructura, siendo las superficies de color rojo las zonas que tienen mayor concentración de electrones.

En la figura 3.2 podemos observar los gráficos de los potenciales electrostáticos notando que todas las estructuras muestran un comportamiento electrostático similar, teniendo un mayor potencial negativo en los átomos de oxígeno y nitrógeno del nicotinoil, y el oxígeno de la metilpirrolidona, teniendo cargas puntuales, esto se debe a que los oxígenos presentan dos pares de electrones libres y un par de electrones libres para el caso del nitrógeno, teniendo mayor contribución para los derivados B, D y E.

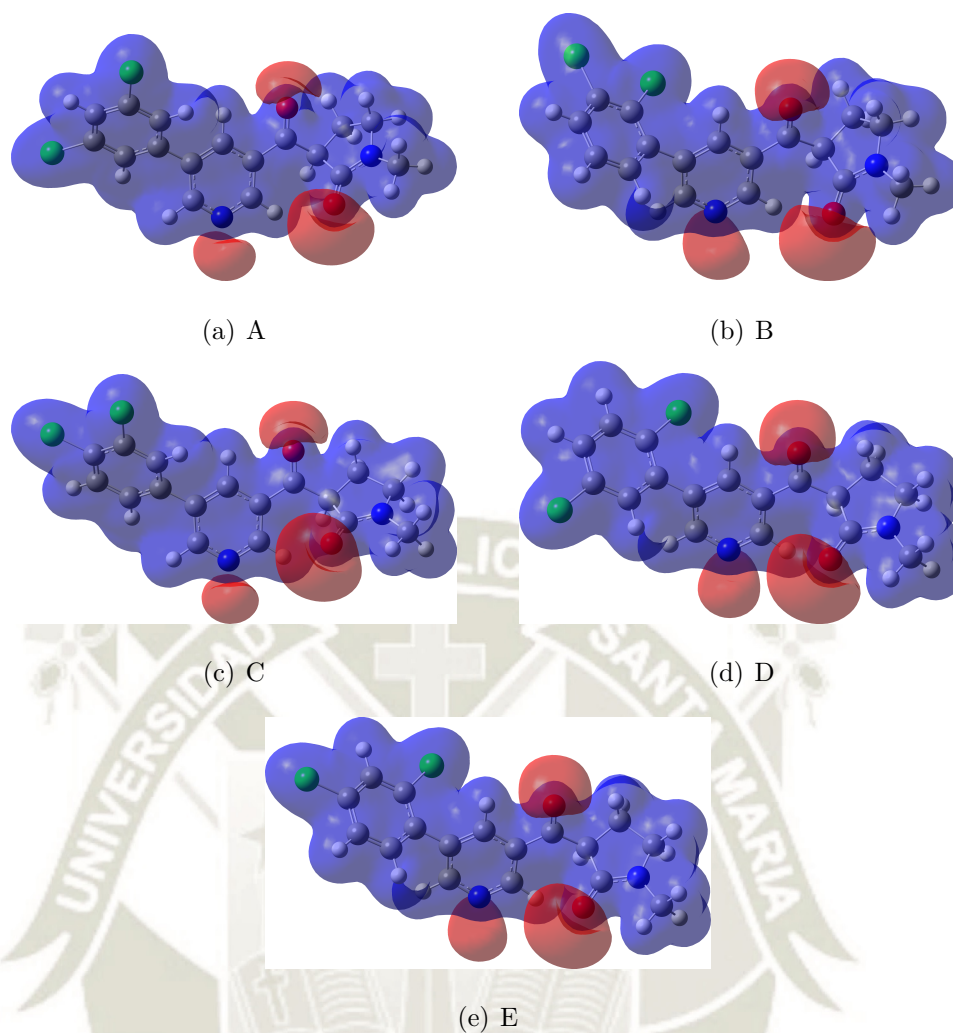


Figura 3.2: Potencial electrostático de los derivados de la nicotina.

Con los orbitales moleculares de frontera podemos obtener un adecuado y concreto análisis de la reactividad local de las estructuras. En las figura 3.3 observamos que en todos los casos el comportamiento de la reactividad se centra en los dobles enlaces del nicotinoil y diclorofenil. Para los derivados "A" y "B" se observa que existe una pequeña contribución en el oxígeno y nitrógeno de la metilpirrolidona y el oxígeno de la nicotinoil. La contribución del grupo de nicotinoil puede ser atribuida a un proceso de resonancia sobre el diclorofenil, a su vez, éste grupo muestra mayor contribución, siendo más favorable la sustitución "para" del cloro observada en el derivado D.

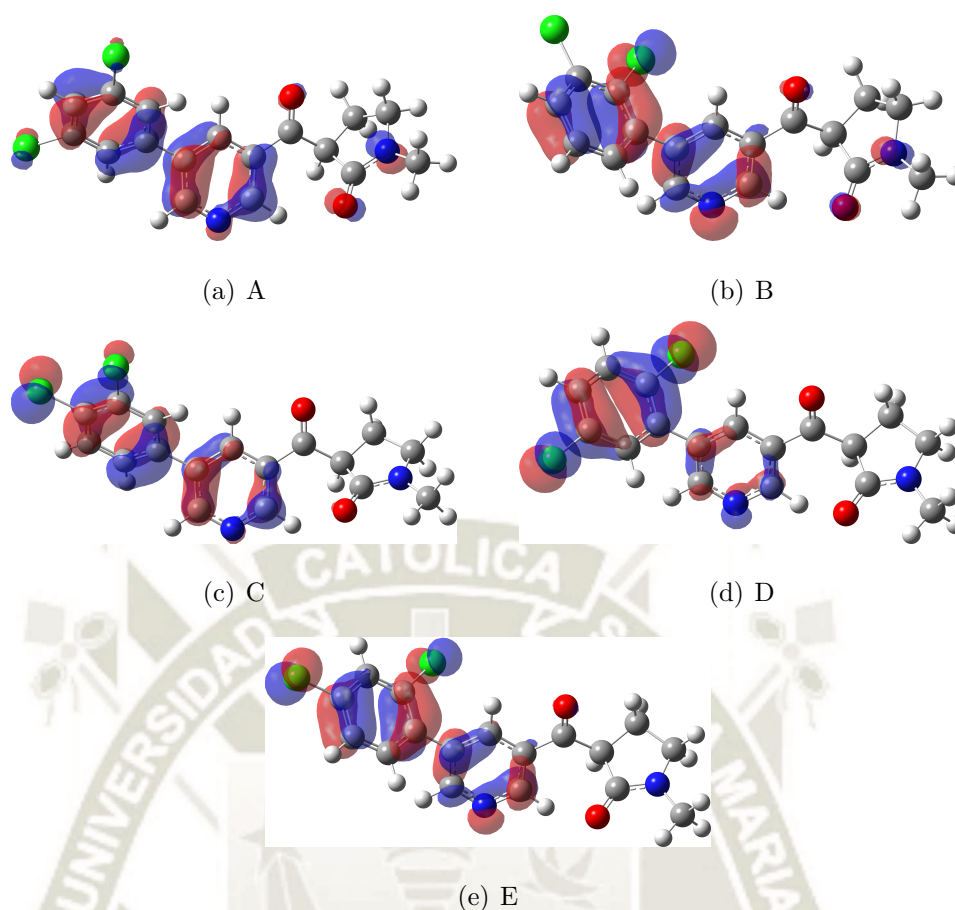


Figura 3.3: Orbitales moleculares más altos ocupados (HOMO) de los derivados de la nicotina.

En general, estos resultados nos indican que, si algún agente electrofílico desea interactuar con los derivados, lo harían principalmente con el diclorofenil y parte del nicotinoil, siendo el de mayor contribución con los derivados "C", "D" y "E" por la posición de los átomos de Cloro, que es el que los diferencia, tal como se observa en la figura 3.3, siendo de color rojo orbitales con funciones de onda positivas y de color azul funciones de onda negativas.

En el análisis de los orbitales moleculares desocupados de energía más baja (LUMO) se observa que se hallan centrados en los dobles enlaces del nicotinoil para todos los derivados de la nicotina (ver figura 3.4). También se observa una pequeña contribución en el carbono del grupo cetona del metilpirrolidona. Tanto el HOMO como el LUMO fueron graficados con un corte de isovalor de 0.04 para todas las estructuras.

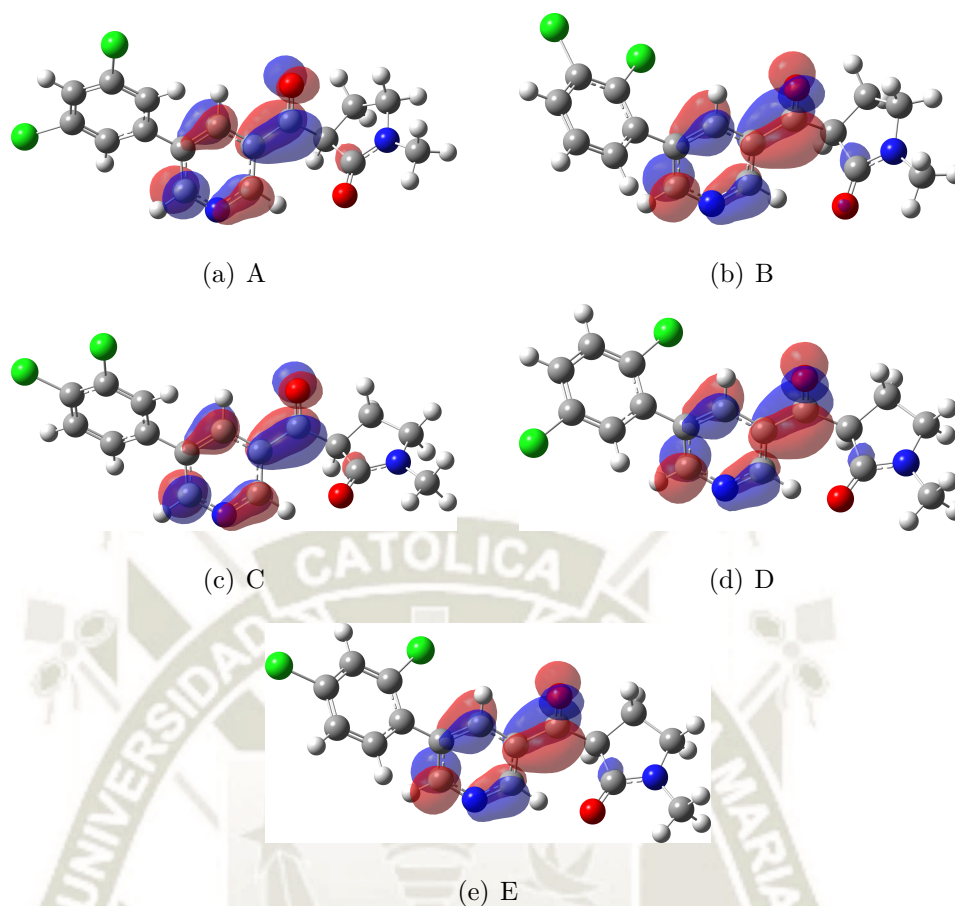


Figura 3.4: Orbitales moleculares más bajos desocupados (LUMO) de los derivados de la nicotina.

La función electrónica de Fukui nos ayuda a identificar de forma precisa el centro involucrado en potenciales mecanismos de reacción. En la figura 3.5 podemos analizar la función de Fukui para un ataque electrofílico centrándose tanto en el nicotinoil como en el diclorofenil. En los derivados "C", "D" y "E" se muestran que los cloros están predispuestos para un ataque electrofílico, así como también podemos contrastarlas con los resultados del análisis del HOMO con algunos cambios significativos. Sin embargo con los derivados "A" y "B" sólo contribuyen para los anillos del diclorofenil y el nicotinoil pero se observa una pequeña contibución con el nitrógeno y el oxígeno de la metilpirrolidona, siendo mayor para el nitrógeno. Asimismo observamos que, el diclorofenil del derivado "D" está completamente susceptible a un ataque electrofílico ya que, para la piridina del nicotinoil se va perdiendo el ataque electrofílico.

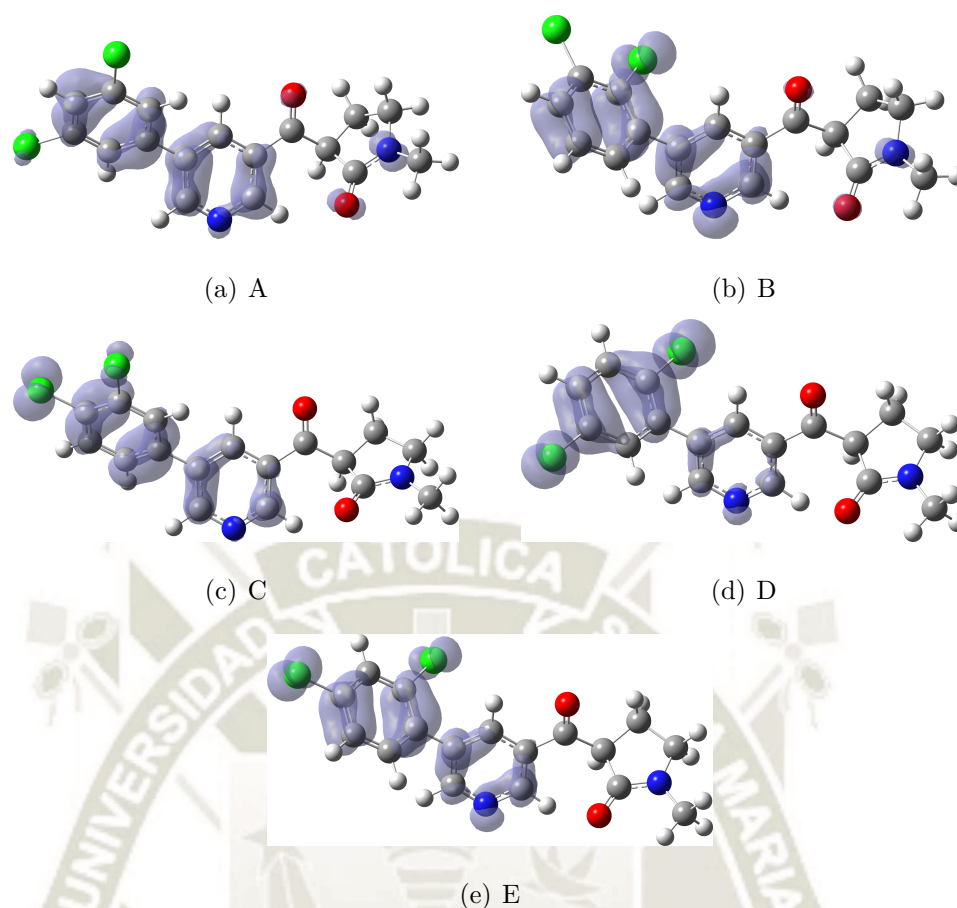


Figura 3.5: Función de Fukui susceptible a un ataque electrofílico de los derivados de la nicotina.

Para los análisis de la función de Fukui susceptible a un ataque nucleofílico en la figura 3.6 se observa el mismo comportamiento para todos los derivados, teniendo una mayor contribución con el derivado B y E seguido del D y A, siendo menor con el derivado C, donde el centro de mayor reactividad está dado por el nicotinoil especialmente en el doble enlace del oxígeno. También podemos observar una pequeña contribución en el carbono del grupo cetona de la metilpirrolidona así como se observa en los resultados del LUMO. Tanto la función de Fukui susceptible a un ataque nucleofílico como electrofílico se graficaron usando un corte de isovalor de 0.002.

Para poder calcular los índices de reactividad global primeramente se obtuvieron las energías del potencial de ionización (PI), obteniendo para los derivados A, B, C, D y E los valores de 159.16, 156.78, 156.34, 159.09 y 159.44 Kcal/mol respectivamente, con respecto a las energías de afinidad electrónica (AE), para los derivados A, B, C, D y E se obtuvieron valores de -57.03, -56.45, -56.18, -56.09 y -56.12 Kcal/mol respectivamente.

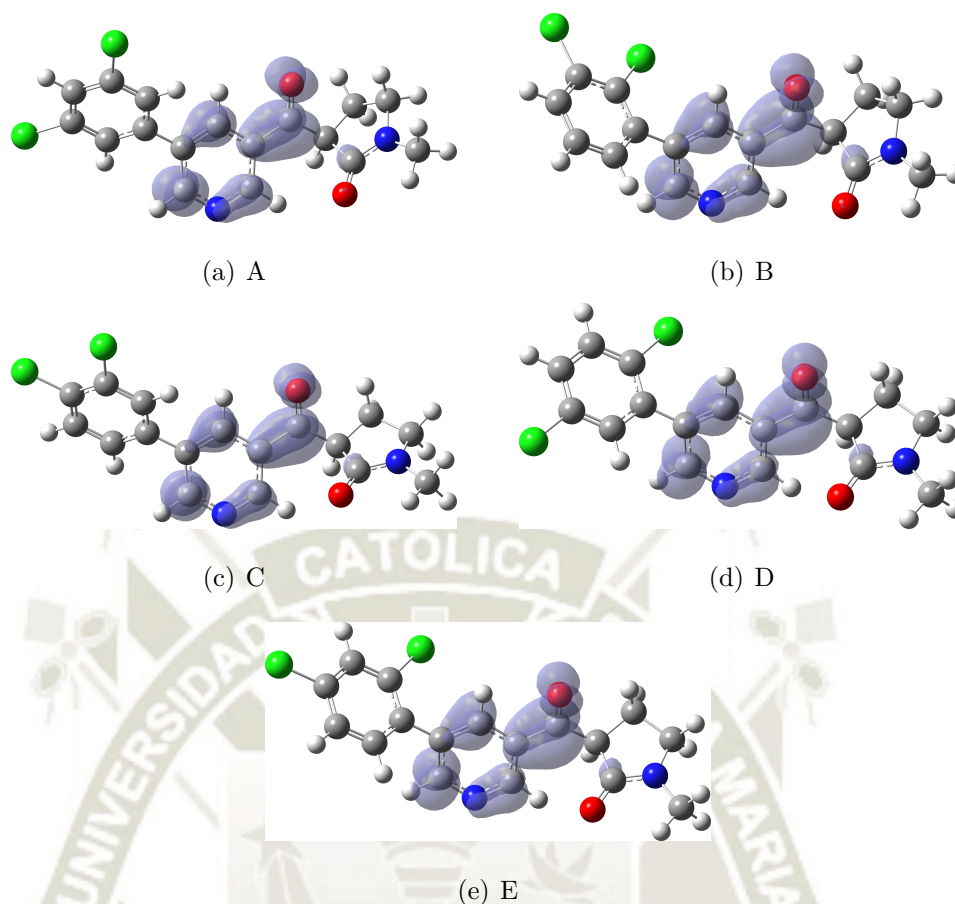


Figura 3.6: Función de Fukui susceptible a un ataque nucleofílico de los derivados de la nicotina.

En la tabla 3.2 podemos observar las energías de los orbitales de frontera que, nos ayudaron a comprender la reactividad química de los derivados de la nicotina.

Tabla 3.2: Energías de los orbitales de frontera de los derivados de la nicotina.

Derivado	A	B	C	D	E
E_{HOMO}	-191.58	-194.77	-186.83	-190.25	-190.64
E_{LUMO}	-21.86	-21.53	-21.06	-21.17	-21.09
GAP	169.73	173.24	165.77	169.08	169.55

*Energías reportadas en Kcal/mol.

Por otro lado, cuando la brecha de energía HOMO y LUMO es menor, será más reactivo químicamente.¹¹⁵ Observamos que en la energía del LUMO no hay mucha

variación entre los cinco derivados, sin embargo, la energía del HOMO disminuye para el derivado C y aumenta para el derivado B. Con respecto a los valores de su brecha energética (GAP), el derivado C es el más reactivo seguidos del D, E y A, siendo el menos reactivo el derivado B.

Tabla 3.3: Índices de reactividad global de los derivados de la nicotina.

Derivado	A	B	C	D	E
η	51.06	50.17	50.08	51.50	51.66
s	0.02	0.02	0.02	0.02	0.02
ω	114.41	113.29	112.73	112.38	112.42
μ	-108.10	-106.62	-106.26	-107.59	-107.78
χ	108.10	106.62	106.26	107.59	107.78

η =dureza; s =blandura; ω =electrofilicidad, μ =potencial químico, χ =electronegatividad.

*Energías reportadas en Kcal/mol.

Asímismo, en la tabla 3.3 podemos observar los índices de reactividad global mostrando que para la dureza el más estable viene a ser el derivado E seguido del D y los menos estables son el derivado C seguido del B, con respecto al derivado A, ésta se encuentra en medio de los 5 derivados; por otro lado, con respecto a la blandura, se observó el mismo comportamiento para todos los derivados de la nicotina teniendo un valor de 0.02, demostrando así, que las estructuras no sufren muchos cambios en su densidad electrónica. También se calculó el índice de electrofilicidad, que nos brinda información sobre la afinidad de electrones, es decir, el flujo de electrones entre el donante y el receptor en una interacción molecular; ¹¹⁶ con respecto a nuestros resultados, el derivado A muestra mayor afinidad por los electrones seguido del derivado B, siendo menor para los derivados C, D y E respectivamente. Finalmente, el potencial químico y la electronegatividad al tener los mismos valores sin tomar en cuenta los signos por la parte electronegativa, muestra ser mejor con el derivado A, seguidos por el E y D, siendo el derivado A el mejor atractor de electrones, quedando como últimos, los derivados B y C.

3.3 Análisis de la Simulación de Dinámica Molecular

Las proteínas fueron obtenidas de la base de datos NCBI encontradas con los códigos *TFB22176.1* (Erp), *NP_217133.1* (Rv2617c) y *NP_215933.1* (Rv1417), estas secuencias se observan en la figura 3.7. Las secuencias fueron obtenidas de la cepa H37Rv, que durante mucho tiempo ha sido utilizada como modelo de estudio.

A

```
>TFB22176.1 exported repetitive protein [Mycobacterium tuberculosis]
MPNRRRRKLTAMSAVAALAVASPCAYFLVYESTETTERPEHHEFKQAAVLTDLPGELMSALSQGLSQFG
INIPPVPSLTGSGDASTGLTGPGLTSPGLTSPGLTSPGLTDPALTSPLTPTLPGSLAAPGTTLAPTPGVGAN
PALTNPALTSPTGATPGLTSPTGLDPALGGANEIPITTPVGLDPGADGTYPILGDPTLTGIPSSPATTSTGGG
GLVNDVMQVANELGASQAIDLLKGVLMPSIMQAVQNGGAAAPAASPPVPIAAAAPPTDPITVPVA
```

B

```
>NP_217133.1 transmembrane protein [Mycobacterium tuberculosis H37Rv]
MSIRPTTSPALADQLKDPAYSAYVLLRTLFTVPILFGLDKFFNLLTHPQHWNMYLAGWINDLVPGTADQ
CMYLVGAIEIVAGVLVAVAPRIGAWVVAWLAVLAGIILNLVTGPGFYDIALRDFGLLVGAIALARLAQGVHS
GGIGRP
```

C

```
>NP_215933.1 membrane protein [Mycobacterium tuberculosis H37Rv]
MTAAPNDWDVLRPHWTPLFAYAAAFLIAVAHVAGGLLLKVGSSGVVFQTADQVAMGALGLVLAGAV
LLFARPRLRVGSAGLSVRNLLGDRIVGWSEVIGVSFPGGSRWARIDLADDEYIPVMAIQAVDKDRAVAA
MDTVRSLRLARYRPDLCAR
```

Figura 3.7: Secuencias en formato FASTA de las tres proteínas. A. Erp B. Rv2617c y C. Rv1417.

Los modelos tridimensionales de las proteínas se optimizaron encontrando el mínimo global de energía mediante el método de *Simulated Annealing*. Este método ha sido utilizado durante mucho tiempo para proteínas que han sido simuladas partiendo desde una secuencia o estructura primaria.¹¹⁷ En la figura 3.8 podemos observar una comparación de cómo cada proteína ha ido cambiando su conformación empezando por la estructural inicial, seguido por la aplicación del *Simulated Annealing* y de un equilibrio y dinámica molecular. Se observan mejoras relevantes después de haber realizado la simulación mediante un *Annealing* obteniendo mejores resultados en la validación y su análisis de los diagramas de Ramachandran. En la proteína Erp se observan drásticos cambios conformacionales, sin embargo, en la proteína Rv1417 se observa la aparición de 2 láminas beta adicionales a comparación de la estructura inicial y en la proteína Rv2617c sólo se observa un ligero cambio en la zona terminal.

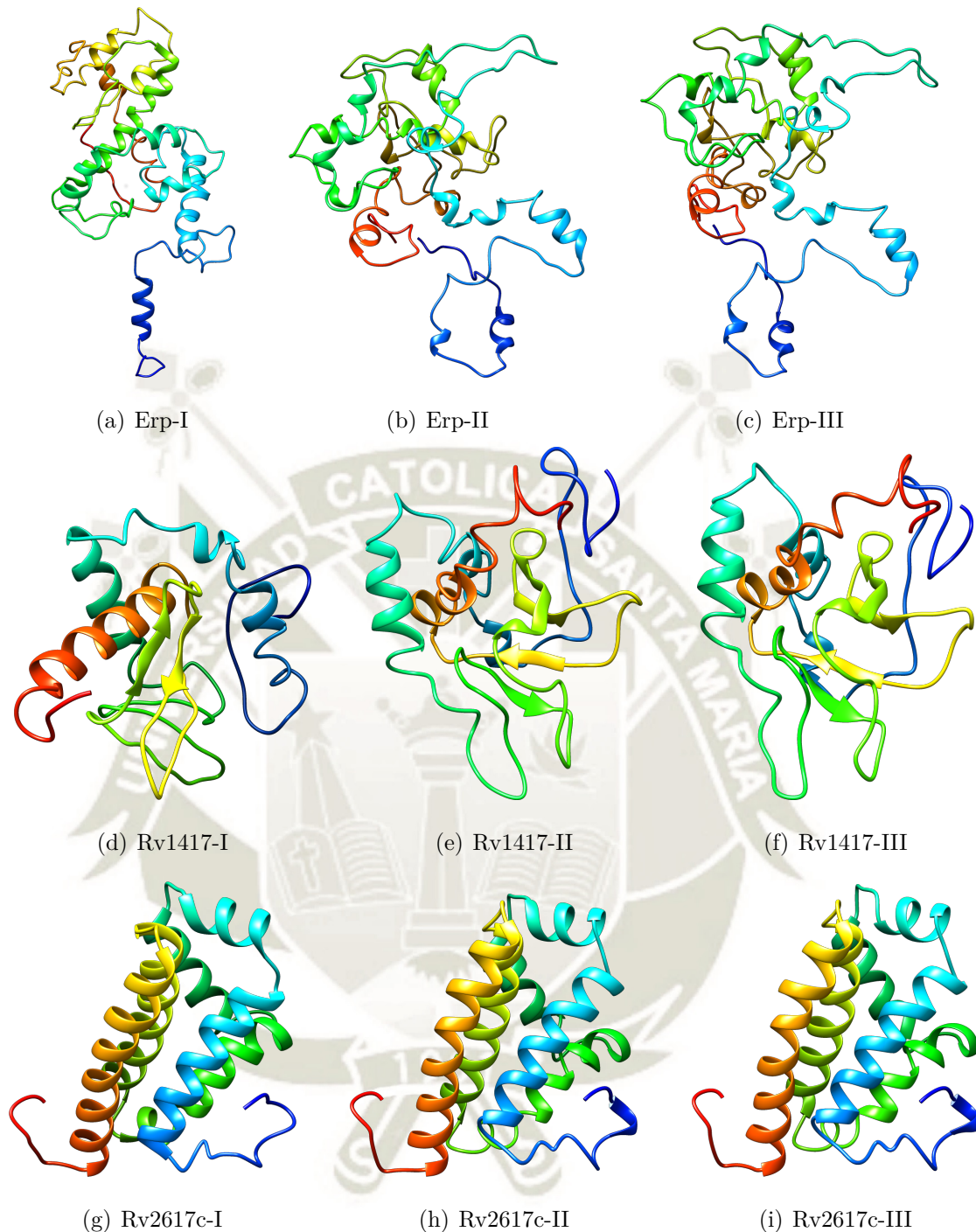


Figura 3.8: Cambios conformacionales estructurales de las tres proteínas. I. Estructura tridimensional inicial, II. Simulated Annealing y III. Dinámica Molecular.

Es necesario recalcar que en el *Simulated Annealing* se simuló a diferentes rangos de temperatura, tiempo y otros especificados, dándonos cuenta que cada proteína tiene una manera diferente de trabajo. La Erp dio mejores resultados trabajando sin restricciones ya sea de forma lineal o periódica con respecto a los valores de

temperatura, en cambio para las proteínas Rv1417 y Rv2617c fue necesario trabajar restringiendo las zonas conformadas por estructuras secundarias y de forma periódica. Las tres proteínas tuvieron mejores resultados cuando las temperaturas no llegaban a temperaturas extremas, ya sea en el calentamiento o enfriamiento, trabajando en un intervalo comprendido entre 100 y 360 K dándole mayor tiempo de cálculo al momento de su enfriamiento.

3.3.1 Validación de las estructuras de las proteínas Erp, Rv1417 y Rv2617c

Para la validación de las estructuras tridimensionales se analizaron los diagramas de Ramachandran en el cual podemos observar las regiones favorecidas donde los aminoácidos son aceptables por la formación de sus ángulos diédricos al igual que las regiones permitidas, también nos muestra regiones no favorecidas en el que los aminoácidos no forman los ángulos diédricos adecuados. Por otro lado, también se analizaron gráficos obtenidos en el servidor ProSA-Web en el que nos muestra la calidad de las proteínas modeladas a través de una comparación de su estructura con proteínas cristalizadas y su energía. También se calcularon los valores de RMSD, RMSF, el radio de giro y los enlaces de hidrógeno.

En la figura 3.9 podemos observar el diagrama de Ramachandran de la Erp, siendo de color celeste y naranja las regiones favorecidas, mostrándose puntos de color negro los residuos pertenecientes en la región, obteniendo un porcentaje del 81.5%; de color celeste y naranja claro se encuentran las regiones permitidas, mostrándose puntos de color naranja los residuos pertenecientes en la región, teniendo un porcentaje del 15.3% y finalmente, de color blanco se encuentran las regiones no favorecidas teniendo un porcentaje del 3.2% encontrándose 9 residuos: 277ASP, 261SER, 224GLN, 215GLY, 120THR, 116SER, 80THR, 31TYR y 20ALA marcadas de color rojo, del cual se observó que se encuentran ubicadas a través de toda la proteína y no en un lugar específico.

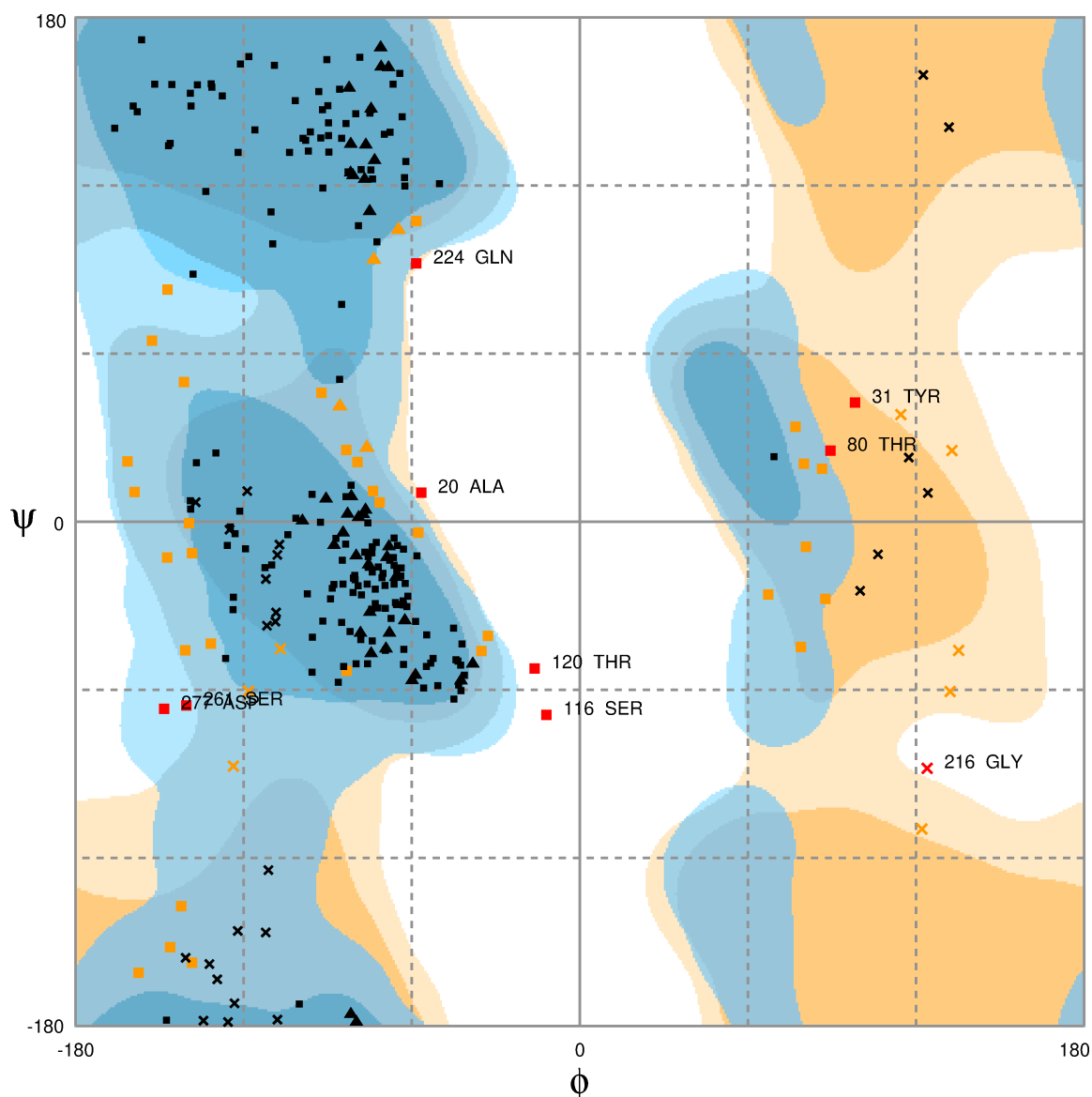


Figura 3.9: Diagrama de Ramachandran de la Erp.

Para el caso de la proteínas Rv1417 en la figura 3.10 se observa un porcentaje del 74.8% en la región favorecida (marcadas de color negro dentro de las zonas celeste y naranja), un 19.9% para la región permitida (marcadas de color naranja en las zonas celeste y naranja claro) y para las regiones no favorecidas un porcentaje del 5.3% (marcadas de color rojo en la zona blanca) observando 8 residuos: 148ARG, 118GLU, 108ARG, 85ARG, 84VAL, 63LEU, 36GLY y 4ALA, observando que éstos aminoácidos se encuentran en toda la proteína, siendo la zona entre las alfa hélice y las lámina beta la de menor conformación (ver figura 3.8).

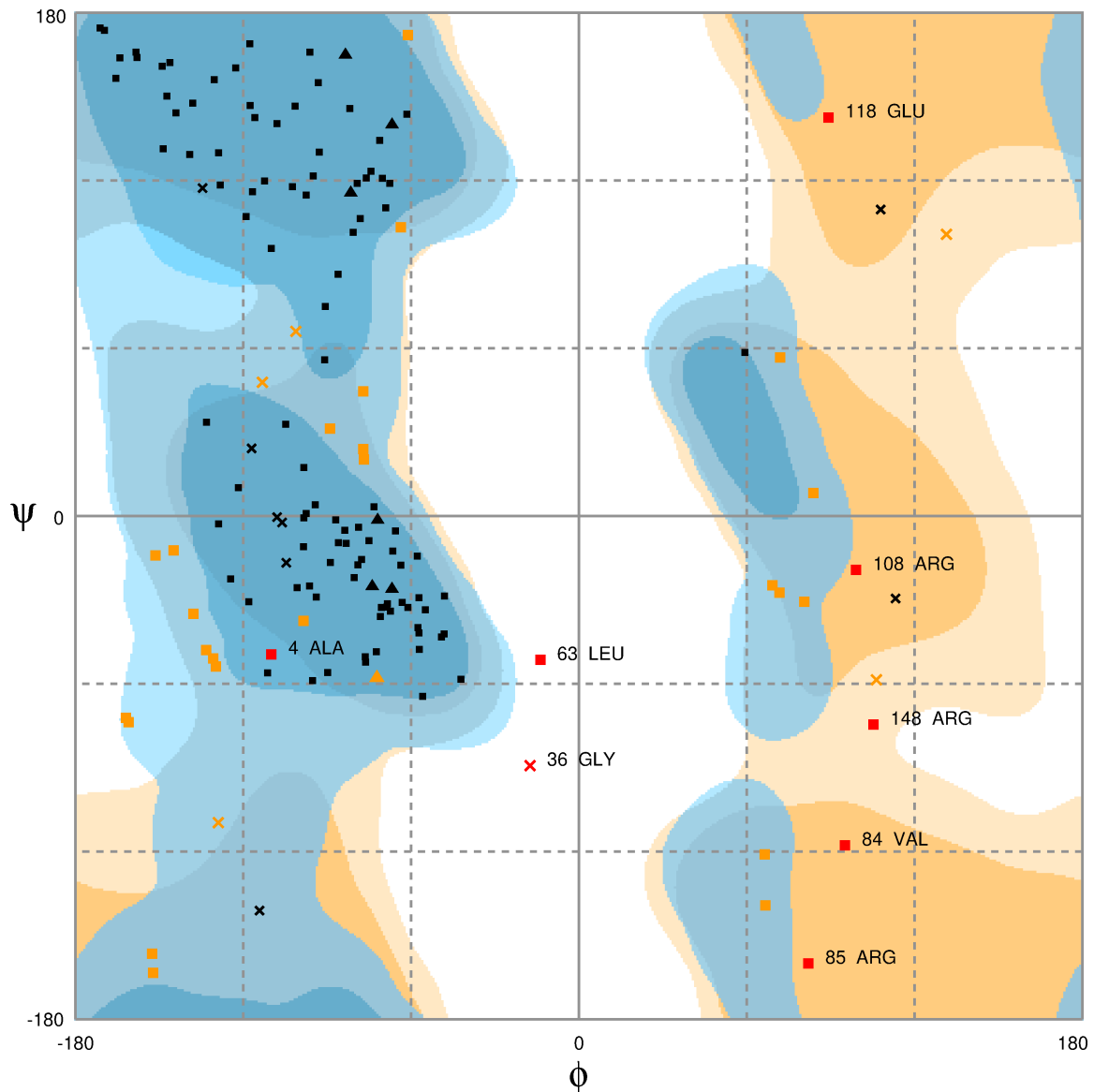


Figura 3.10: Diagrama de Ramachandran de la proteína Rv1417.

Finalmente, con respecto a la proteína Rv2617c, la región favorecida nos da un porcentaje del 85.3% (marcadas de color negro dentro de las zonas celeste y naranja), para las regiones permitidas se obtiene un porcentaje del 9.8% (marcadas de color naranja en las zonas celeste y naranja claro) y para las regiones no favorecidas un porcentaje del 4.9% (marcadas de color rojo en la zona blanca) cuyos aminoácidos no favorecidos son: 111GLY, 114PHE, 39LEU, 38GLY, 3ILE, 14GLN y 10ALA, 2 de éstos aminoácidos se encuentran casi al final de toda la proteína y el resto se encuentran al comienzo de la proteína, observando que ambas zonas no se encuentran conformadas. (ver figuras 3.8 y 3.11).

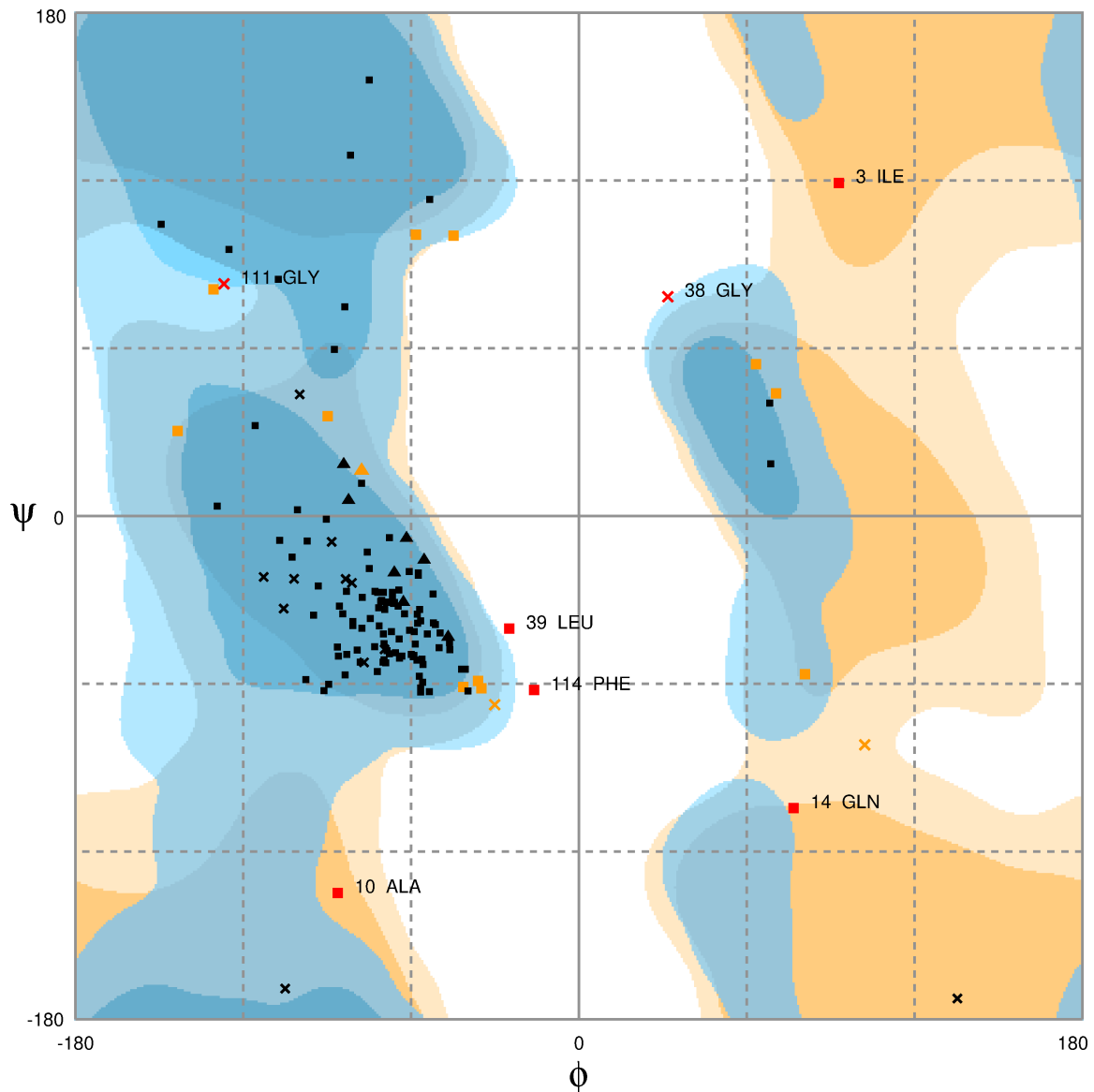


Figura 3.11: Diagrama de Ramachandran de la proteína Rv2617c.

Complementariamente en la tabla 3.4 se observan los datos tanto en número de residuos como en porcentaje de los valores de las regiones favorecidas, permitidas y no favorecidas o atípicas. La proteína con mejores resultados fue la Erp con un total del 96.8% y 272 residuos, seguido de la Rv2617c con un 95.1% y 136 residuos y por último la Rv1417 con un 94.7% y 143 residuos entre la región favorecida y permitida, con respecto a las regiones atípicas entre las tres proteínas se observan que de 7 a 9 residuos no han llegado a formar los ángulos ψ y ϕ .

Tabla 3.4: Datos del gráfico de Ramachandran de las tres proteínas estudiadas.

Proteínas	Región Favorecida		Región Permitida		Región Atípica	
	N° Residuos	%	N° Residuos	%	N° Residuos	%
Erp	229	81.5	43	15.3	9	3.2
Rv2617c	122	85.3	14	9.8	7	4.9
Rv1417	113	74.8	30	19.9	8	5.3

Si bien, los valores aceptables para la evaluación con el gráfico de Ramachandran son del 98%, en los trabajos de investigación en la que se trabajaron con proteínas modeladas a partir de una secuencia, reportan ser aceptables los resultados de los gráficos de Ramachandran cuyos valores oscilen entre los 80 y 90% entre la región favorecida y permitida.^{118–121}

Por otro lado, los resultados de los gráficos obtenidos en el servidor ProSA-Web, nos muestran gráficas de la calidad general del modelo y gráficas de las energías de los residuos en función a su posición.

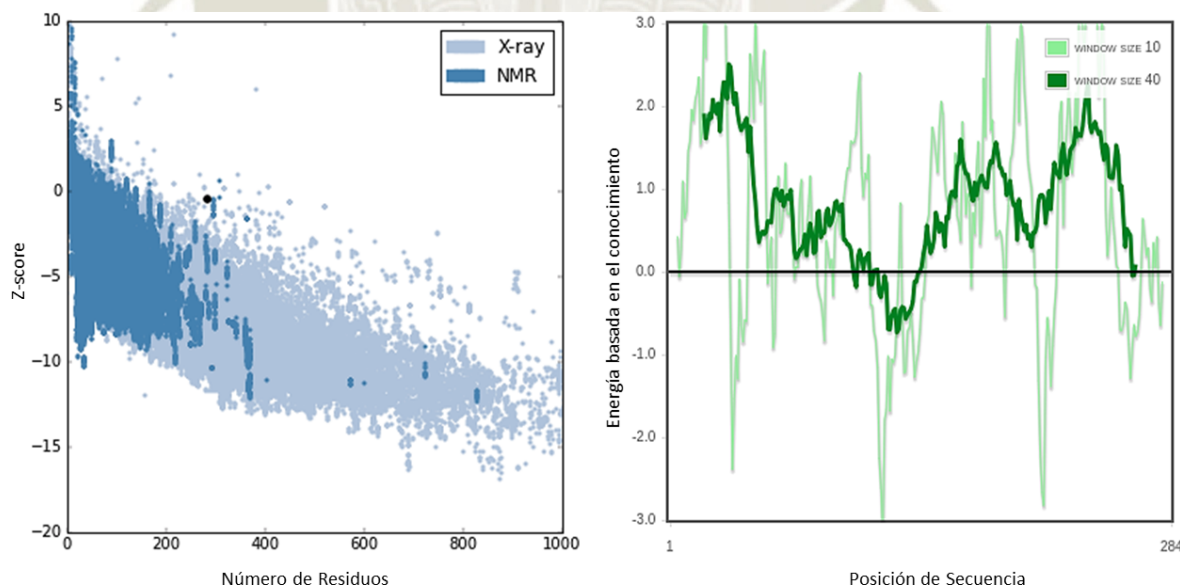


Figura 3.12: Análisis de calidad de la estructura tridimensional de Erp.

En la figura 3.12 podemos observar a la Erp obteniendo un Z-score de -0.43 en la gráfica de calidad general, estando cerca a las estructuras cristalizadas por NMR y x-Ray encontradas en la base de datos del servidor, con respecto a la segunda

gráfica observamos que la proteína muestra fluctuaciones (color verde oscuro), pero con tendencia a disminuir ya que parte de su estructura al encontrarse por debajo de "cero" tiende a estar relajada.

Con respecto a la proteína Rv1417 podemos observar que se encuentra dentro de la zona de las proteínas cristalizadas por x-Ray acercándose a la zona del NMR, teniendo un Z-score de -1.59, el cual es mejor que la Erp. Con respecto a la gráfica de las energías (color verde oscuro) de los residuos se observó que parte de su estructura tiende a bajar su energía cercandose a "cero" y la otra parte de su estructura se encuentra por debajo y tomando un valor "cero" estando relajada. Tanto para la proteína Rv1417 como para la Erp, requieren más tiempo de cálculo al momento de realizar la simulación de dinámica molecular para que puedan encontrar un estado más relajado.

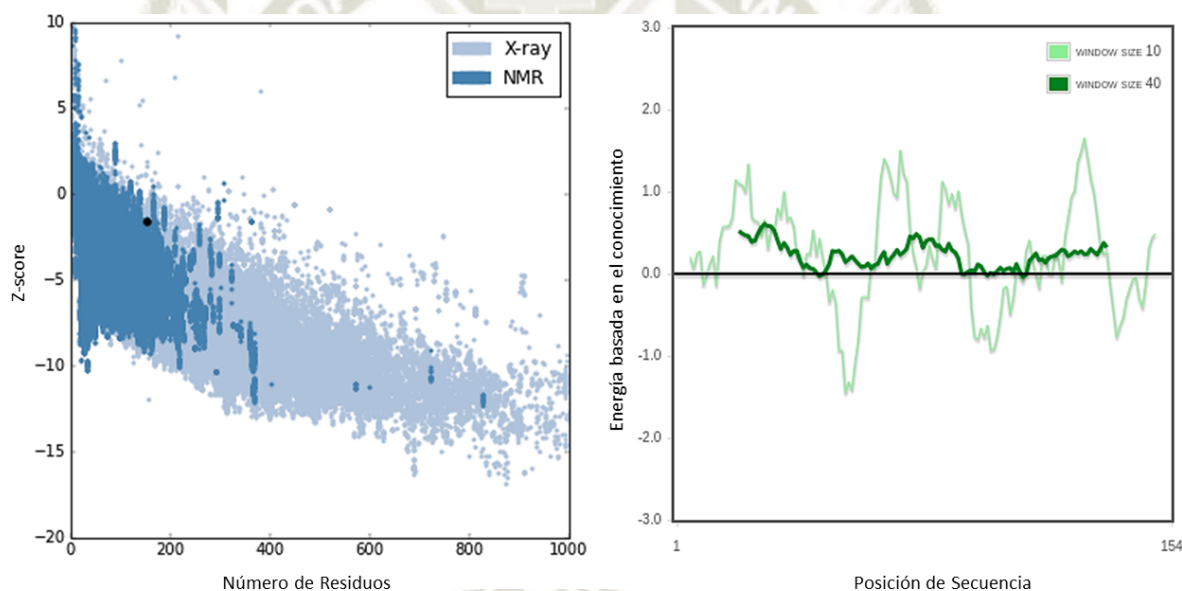


Figura 3.13: Análisis de calidad de la estructura tridimensional de la proteína Rv1417.

En la figura 3.14 podemos observar a la proteína Rv2617c, encontrándose dentro de la base de datos de las estructuras cristalizadas por NMR, siendo ésta la mejor, teniendo un z-score de -4.01, y para la gráfica de energías de los residuos (color verde oscuro) se observó que la proteína se encuentra por debajo de "cero" encontrándose relajada, concluyendo que esta estructura tridimensional presenta una buena calidad de modelado, observando que gran mayoría de sus aminoácidos encontraron una conformación.

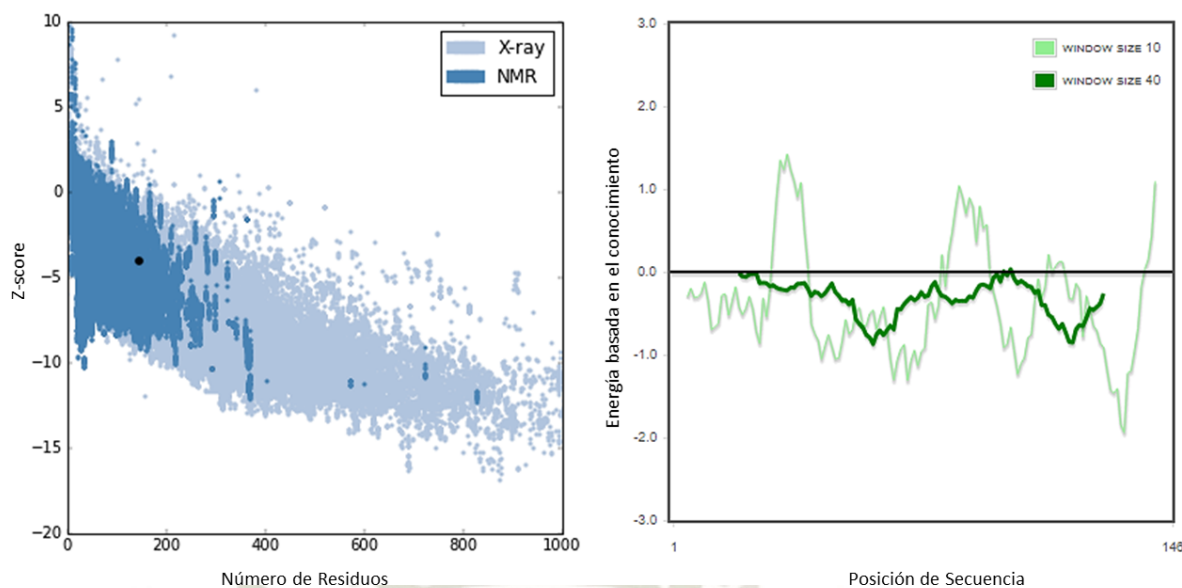


Figura 3.14: Análisis de calidad de la estructura tridimensional de la proteína Rv2617c.

Los gráficos de RMSD, RMSF, el radio de giro y los enlaces de hidrógeno nos muestran una mejor percepción de cómo la molécula ha ido comportándose a través del tiempo. En la figura 3.15 observamos que la proteína Rv2617c ha sido estable a partir de los 20 ns, llegando a subir su RMSD a partir de los 100 ns e incrementándose a partir de los 190 ns aproximadamente, mostrando un rango de 0.2 nm entre estos intervalos de tiempo, mientras que las demás proteínas han ido encontrando estabilidad a partir de los 60 ns, observando que la Erp es la más inestable a comparación de las 3 proteínas, tendiendo a subir y bajar con respecto al valor de su RMSD, mostrando un rango de 0.5 nm a partir de los 120 ns. Sin embargo la proteína Rv1417 se muestra estable con un rango de 0.3 nm a partir de los 60 ns hasta los 200 ns, pero a su vez también muestra estabilidad entre los 5 y 50 ns.

Los rangos de RMSD en los que encuentran estabilidad las proteínas son 2Å, 3Å y 5Å para la Rv2617c, Rv1417 y Erp respectivamente. Los rangos normales para proteínas conocidas modeladas por homología son aceptados hasta valores de <3Å, en cambio para proteínas modeladas desde su primera estructura (secuencias de aminoácidos) se considera un éxito valores de 6Å, incluso se encuentran valores de 8 y 10Å.^{122,123}

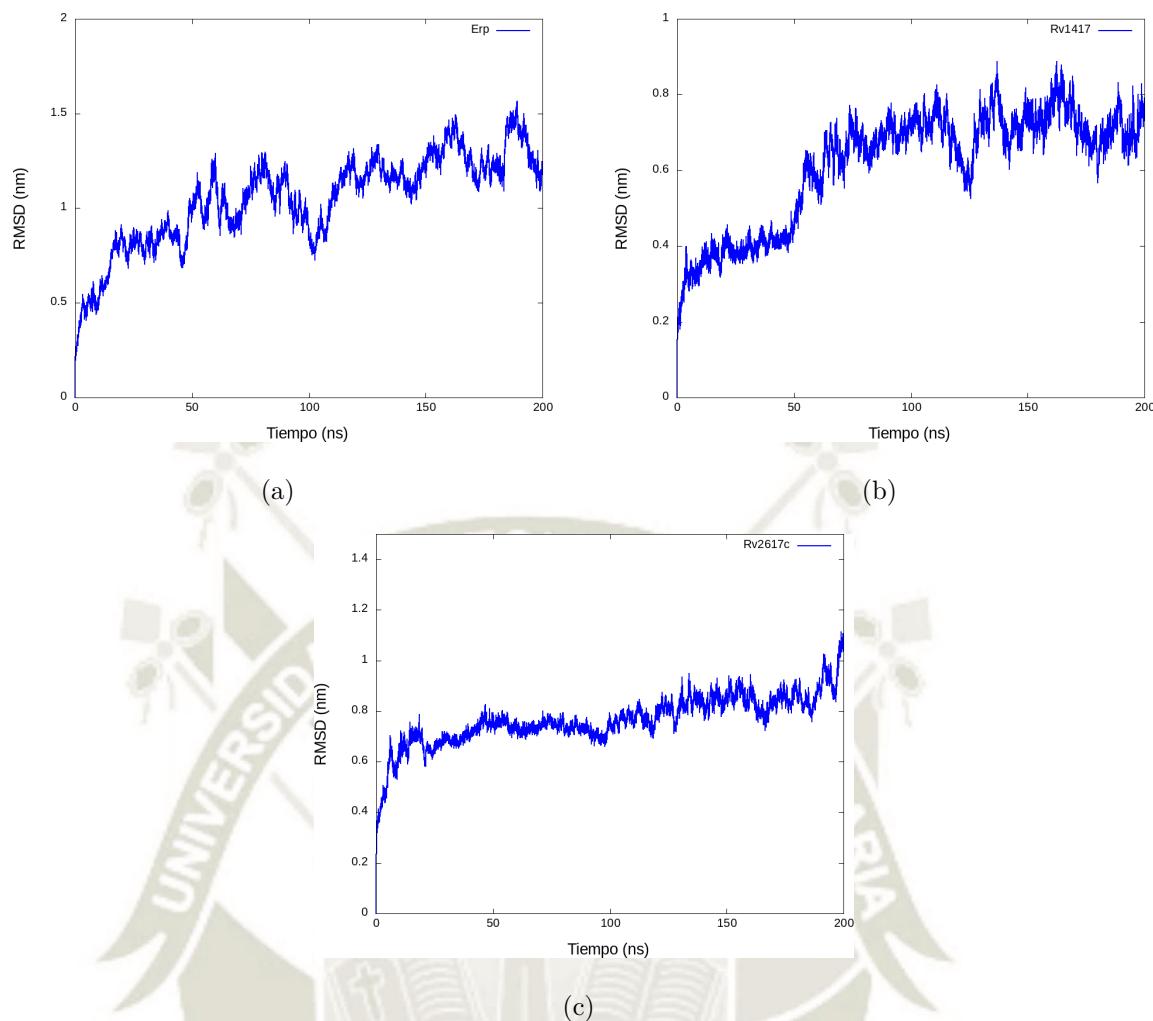


Figura 3.15: RMSD de las proteínas. (a) Erp, (b) Rv1417 y (c) Rv2617c .

Las gráficas de RMSF nos muestran las fluctuaciones de los residuos de cada proteína de forma local a través del tiempo y el Radio de Giro nos da a conocer la distancia de los átomos desde el centro de masa de la proteína. ¹²⁴ Para el RMSF, se tomaron en cuenta los 200 ns de cálculo para observar todo su comportamiento (ver figura 3.16). Para el caso de la proteína Erp observamos que la zona más fluctuante se da entre los residuos 180(THR) y 284(ALA). La proteína Rv1417 muestra una mayor fluctuación entre los residuos 20(PHE) y 70(PHE), a partir de ésta la proteína se muestra estable. Finalmente, la proteína Rv2617c muestra fluctuación por regiones de aminoácidos, observando una elevada fluctuación con respecto a su último residuo. De todo éste análisis podemos comprobar que este tipo de estructuras son altamente fluctuantes, particularmente en las zonas que no han podido ser reestructuradas.

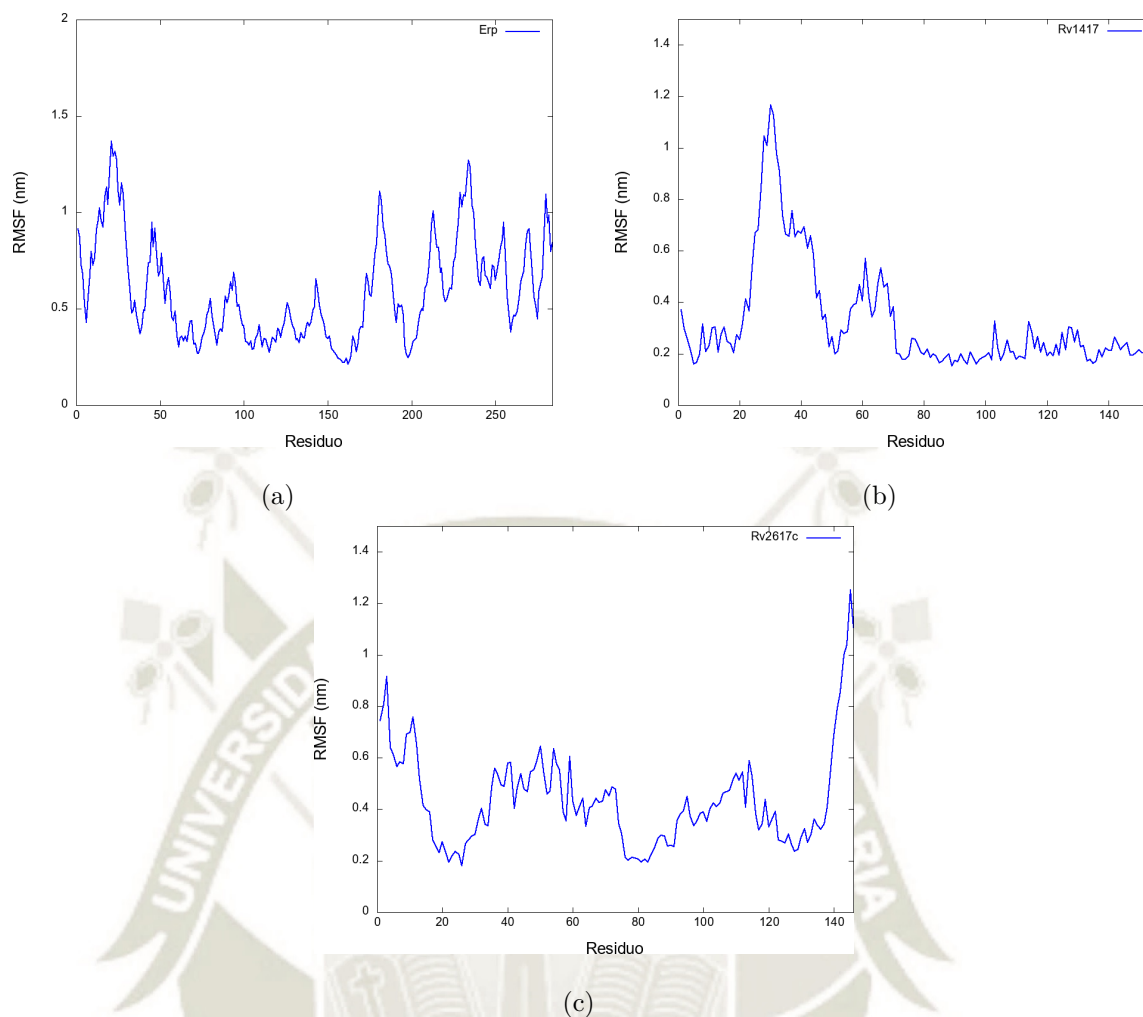


Figura 3.16: RMSF de las proteínas. (a) Erp (b) Rv1417 y (c) Rv2617c.

En la figura 3.17 observamos las gráficas del Radio de Giro de cada proteína en un tiempo de 200 ns, que podemos enriquecerlas con las gráficas de RMSF, ya que a mayor distancia de los átomos con respecto a un radio de giro, mayor será la fluctuación. El radio de giro de la Erp se muestra mayor al comienzo de la simulación entre los 50 y 100 ns, tratando de mantenerse estable los últimos 10 ns, sin embargo, la proteína Rv1417 incrementa su radio de giro entre los 50 y 60 ns, llegando a mantener estable el valor de su radio de giro así como sus fluctuaciones hasta los 200 ns. La proteína Rv2617c muestra estabilidad entre los 20 y 100 ns, llegando a incrementar su radio de giro hasta los 200 ns, tal como se observa una mayor fluctuación en la gráfica del RMSF.

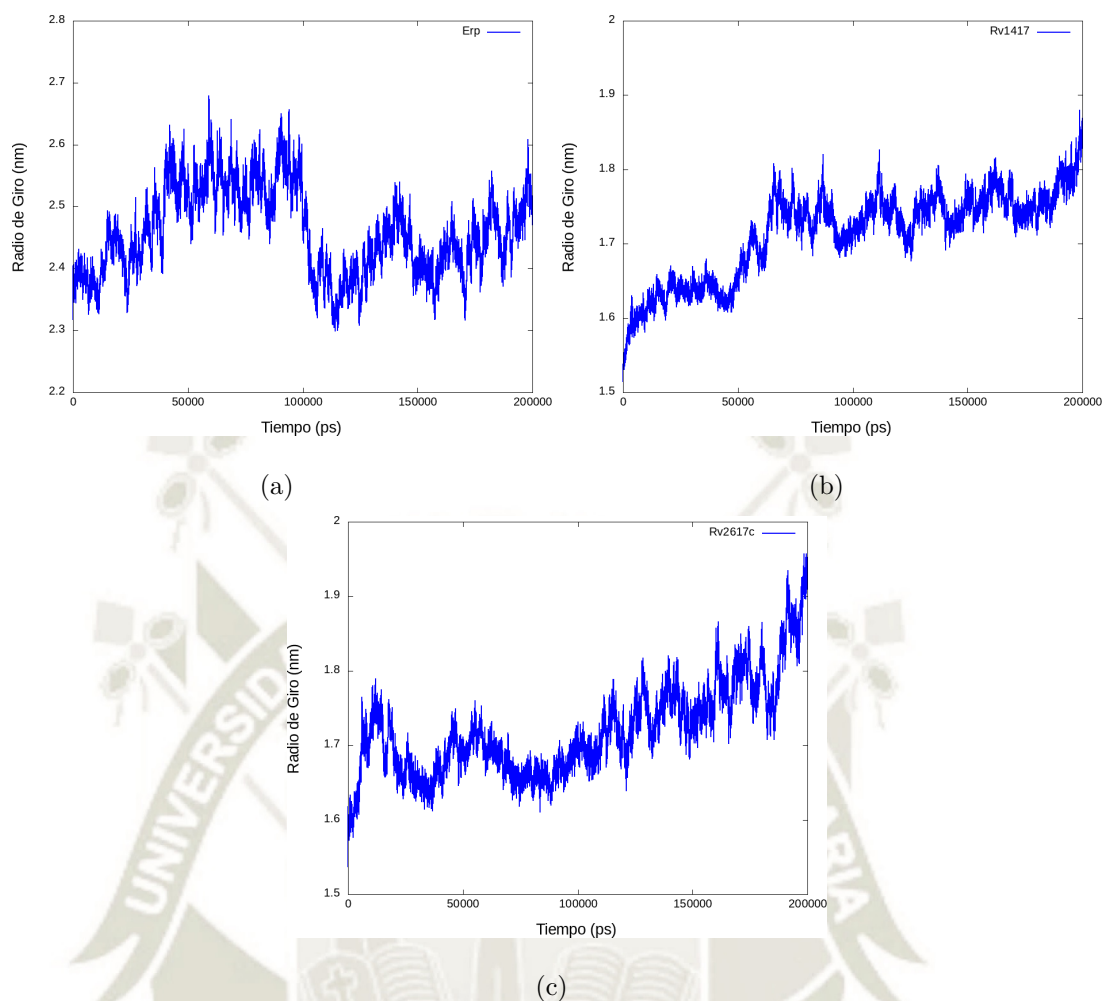


Figura 3.17: Radio de Giro de las proteínas. (a) Erp, (b) Rv1417 y (c) Rv2617c.

Las formaciones de los enlaces de hidrógeno se observan en la figura 3.18. La Erp mantiene la formación de enlaces de hidrógeno desde los 50 ns hasta los 170 ns, tendiendo a disminuir para los últimos 20 ns, sin embargo, la proteína Rv1417 en un comienzo muestra inestabilidad en la formación de enlaces hasta los 100 ns, pasado este tiempo la formación de los enlaces tiende a bajar pero se muestra más estable, esto se debe a la presencia y formación de láminas beta. Finalmente, la proteína Rv2617c muestra estabilidad en la formación de los enlaces de hidrógeno hasta los 100 ns, pasado éste tiempo los enlaces de hidrógenos tienden a disminuir. Por lo general, la variación de enlaces de hidrógeno que existe a través del tiempo en las tres proteínas es mínima.

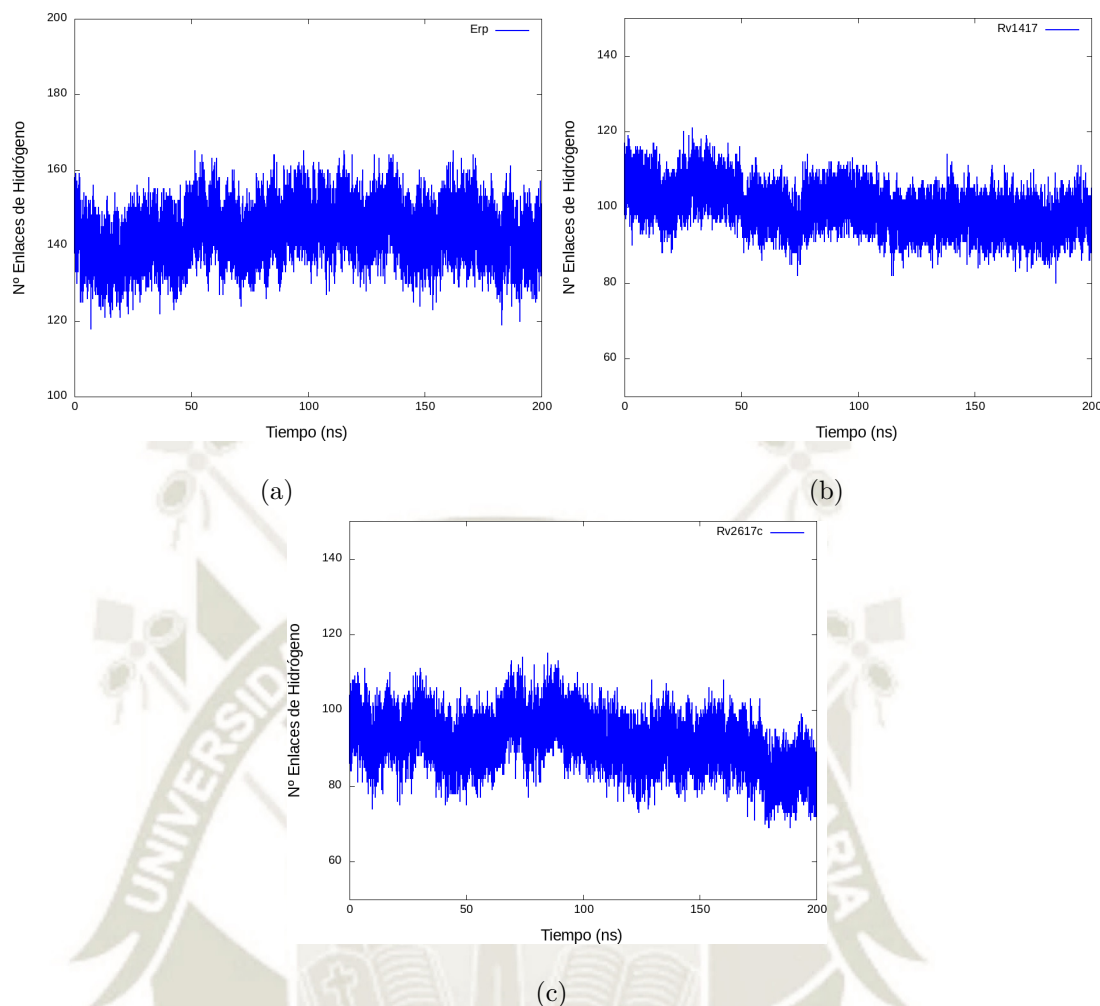


Figura 3.18: Enlaces de Hidrógeno de las proteínas. (a) Erp, (b) Rv1417 y (c) Rv2617c.

3.3.2 Predicción de los bolsillos

Para poder realizar el acoplamiento molecular primeramente se realizaron la predicción de posibles bolsillos para cada proteína, teniendo como resultados finales 2 bolsillos para las proteínas Erp (bolsillo "c" de color naranja y "b" de color turquesa) y Rv2617c (bolsillo "d" de color turquesa y "f" de color naranja) y un solo bolsillo para la proteína Rv1417 (bolsillo "a" de color turquesa) como se observa en la figura 3.19. A través del servidor PockDrug se obtuvo un ranking de 6 bolsillos para la Rv2617c y 5 bolsillos para las proteínas Erp y Rv1417, escogiendo las mejores.

La probabilidad de drugabilidad de los dos bolsillos de la Erp ("b" y "c") según la tabla 3.5 son de 0.95 y 0.81, comparando con el bolsillo "a" son menores, pero también es necesario tener en cuenta el tamaño del bolsillo para obtener una mejor interacción

con el ligando. Para el caso de la proteína Rv1417 el único bolsillo aceptable es el "a" por tener un volumen grande donde los derivados pueden acoplarse completamente, a su vez el bolsillo presenta un valor alto en residuos polares y aromáticos. Finalmente, la proteína Rv2617c muestra altas probabilidades de drugabilidad, siendo el máximo valor para los bolsillos "a" y "b", sin embargo, los derivados tienen preferencia por el bolsillo "d" teniendo valores muy cercanos a los primeros en el ranking, pero superan los valores en los residuos polares y aromáticos. Por otro lado, al no tener preferencia el derivado D por los bolsillos mencionados se eligió otro bolsillo evaluando la superficie de hidrofobicidad de la proteína, por el mismo hecho de que las drogas se favorecen en este ambiente.¹²⁵

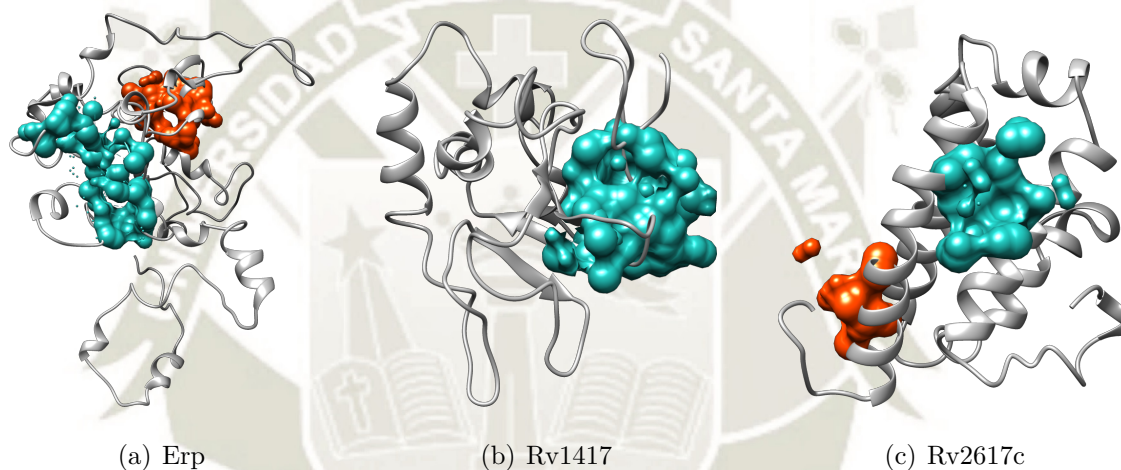


Figura 3.19: Bolsillos para las tres proteínas. (a) Erp, (b) Rv1417 y (c) Rv2617c.

Los residuos aromáticos demostraron estar relacionados con las cadenas laterales hidrófobas alifáticas en presencia de residuos polares para el reconocimiento proteína-ligando en el diseño de fármacos formando enlaces CH- π y catión- π fuertes.¹²⁶ También se encuentran estudios donde se trabajan con compuestos de nicotina que fueron reconocidos por grupos de residuos aromáticos interactuando con el nitrógeno de la pirrolidina formando enlaces catión- π , de la misma forma encontrándose sustituyentes de cloro hidrofóbico.^{127,128} De acuerdo a los resultados podemos observar que los derivados de la nicotina trabajados en este estudio tienden a tener el mismo comportamiento, escogiendo bolsillos con mayor cantidad de residuos aromáticos.

Tabla 3.5: Datos de la predicción de bolsillos de las tres proteínas.

Bolsillo	Volumen de Casco	Kyte Hidrofób.*	Residuos Polares	Residuos Aromát.	Átomos Otyr*	Residuos del Bolsillo	Prob. de Drugab.	Desv. Estand.
Erp								
a	1381.92	1.33	0.22	0.00	0.00	18.0	0.98	0.00
b	3926.00	0.71	0.28	0.03	0.00	40.0	0.96	0.01
c	1533.96	0.20	0.35	0.04	0.00	23.0	0.81	0.03
d	5962.62	-0.21	0.45	0.02	0.01	47.0	0.75	0.06
e	1274.89	-0.03	0.39	0.00	0.00	18.0	0.67	0.06
Rv1417								
a	1916.44	0.33	0.40	0.16	0.00	25.0	0.92	0.02
b	384.30	2.12	0.20	0.20	0.04	10.0	1.00	0.00
c	328.13	0.75	0.36	0.18	0.00	11.0	0.96	0.01
d	430.37	0.61	0.38	0.25	0.00	8.0	0.96	0.02
e	272.32	0.93	0.25	0.00	0.00	8.0	0.93	0.02
Rv2617c								
a	447.63	1.57	0.30	0.20	0.00	10.0	1.00	0.00
b	320.24	2.65	0.09	0.27	0.00	11.0	1.00	0.00
c	499.15	0.70	0.40	0.40	0.03	10.0	0.99	0.01
d	465.71	0.69	0.44	0.56	0.03	9.0	0.99	0.01
e	310.03	1.90	0.10	0.00	0.00	10.0	0.99	0.00
f	280.77	0.80	0.25	0.38	0.04	8.0	0.99	0.00

*Kyte hidrofóbico = residuos estudiados por Kyte. ¹²⁹

*Átomos Otyr = frecuencia de átomos de Otyr (TyrOH) en el bolsillo. ¹³⁰

3.3.3 Interacción de los Derivados de la Nicotina con las proteínas Erp, Rv1417 y Rv2617c

El acoplamiento proteína-ligando para cada sistema se realizó en el servidor PATCHDOCK obteniendo un ranking de 10 interacciones. En la tabla 3.6 se observan las mejores interacciones para cada proteína con cada derivado. Para la proteína Erp con el derivado A se encuentra en el bolsillo "c" teniendo la mejor energía de contacto atómico (ACE) y para el resto de los derivados se encuentran en el bolsillo "b" teniendo valores similares de "score", área y ACE.

La energía de contacto atómico es la energía de desolvatación, reemplazando un contacto proteína-átomo de agua con un contacto proteína-ligando a nivel atómico.

Cuando los valores de ACE son positivos tiende a disminuir la unión de una proteína y/o ligando, en cambio, un valor negativo es favorable para un acoplamiento.^{114,131} Todas las interacciones observadas en la tabla 3.6 muestran un buen "score" y valores negativos de ACE que nos proporciona buenas interacciones y que además es factible un acoplamiento espontáneo en estos bolsillos.

Tabla 3.6: Datos del Acoplamiento Molecular obtenidas en el servidor PATCHDOCK.

Proteína	Derivado	Score	Área	ACE
Erp	A	4898	562.20	-189.76
	B	4642	521.40	-287.07
	C	4976	539.50	-221.08
	D	4666	519.40	-268.48
	E	4738	533.40	-271.71
Rv1417	A	4508	596.10	-322.54
	B	4658	521.40	-222.73
	C	4344	541.20	-274.46
	D	4552	566.60	-271.22
	E	4856	616.60	-325.44
Rv2617c	A	3962	499.70	-205.65
	B	4028	500.00	-189.37
	C	4098	498.30	-208.75
	D	4272	529.40	-291.30
	E	4326	525.70	-209.75

3.3.4 Análisis de dinámica molecular del acoplamiento

Los RMSD de la proteína Erp después del acoplamiento muestra mayores cambios en la unión con el derivado E, seguido del derivado C, D, A y B; para la proteína Rv1417 muestran cambios notorios con el derivado A seguido del B, y de los derivados C, D y E que muestran comportamientos similares. Finalmente, la proteína Rv2617c al igual que la proteína Rv1417, los cambios que se muestran son muy notorios empezando con

el derivado B, seguido de los derivados C y E que también muestran comportamientos similares (ver figura 3.20).

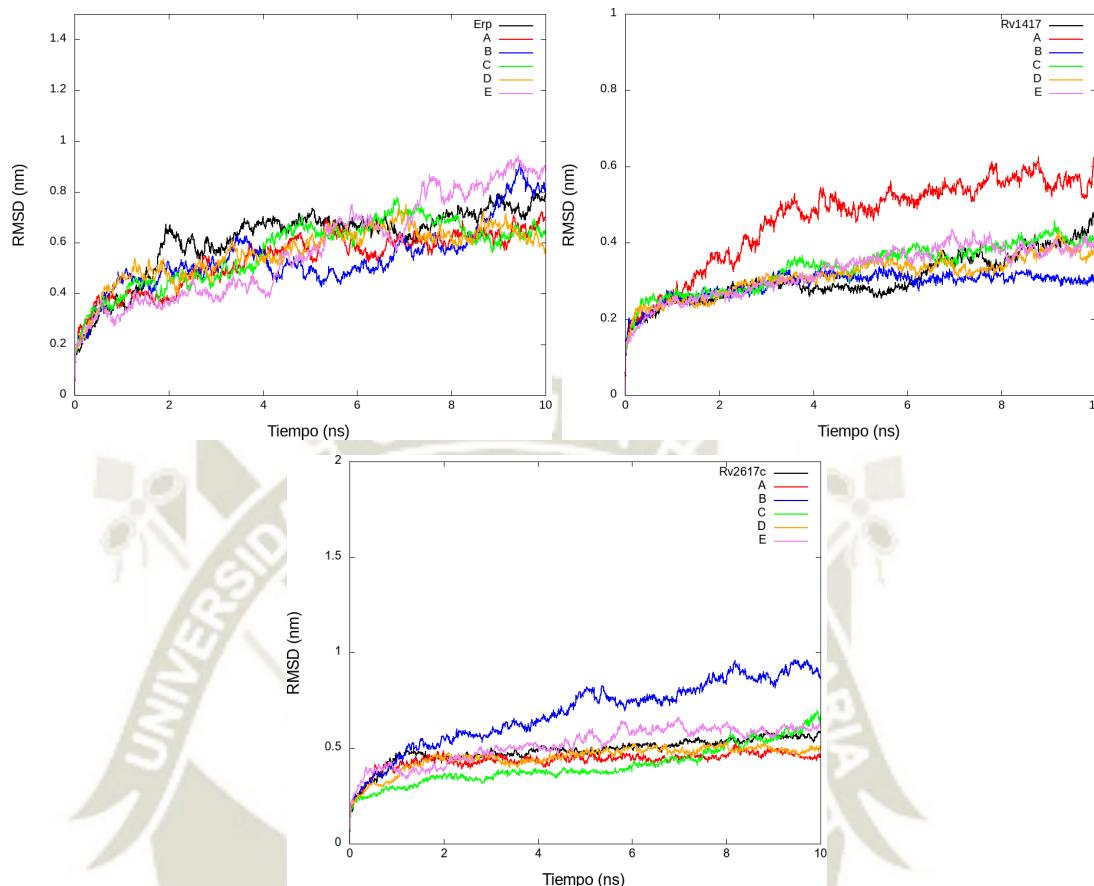


Figura 3.20: RMSD de la interacción de las tres proteínas con cada derivado la de nicotina. (a) Erp, (b) Rv1417 y (c) Rv2617c.

Con respecto a un enfoque local, en la figura 3.21 podemos observar los RMSF de cada proteína. En caso de la Erp con el derivado B se observan muchos picos de fluctuación, luego se encuentran los derivados E, D y A que muestran comportamientos similares, por último, tenemos al derivado C. Mayormente, las fluctuaciones se dan en los residuos 94(LEU), 213(THR), 232(SER). La proteína Rv1417 tiene diferentes comportamientos tendiendo a tener picos muy elevados con los primeros residuos, en cambio con el derivado B los picos de los primeros residuos tienden a bajar y a ser más estable, pero con los derivados C y D que tienen comportamientos similares seguido del derivado E se observan pocos cambios; y para el derivado A se observa el mismo comportamiento que la proteína inicial. Finalmente, la proteína Rv2617c muestra mayores fluctuaciones en la zona C-terminal, siendo drásticos para el derivado B, en

cambio con el derivado D los picos tienden a disminuir a comparación de la proteína sola; los derivados C y E tienden a tener comportamientos similares entre sí, seguido del derivado A que se asemeja a las fluctuaciones iniciales de la proteína.

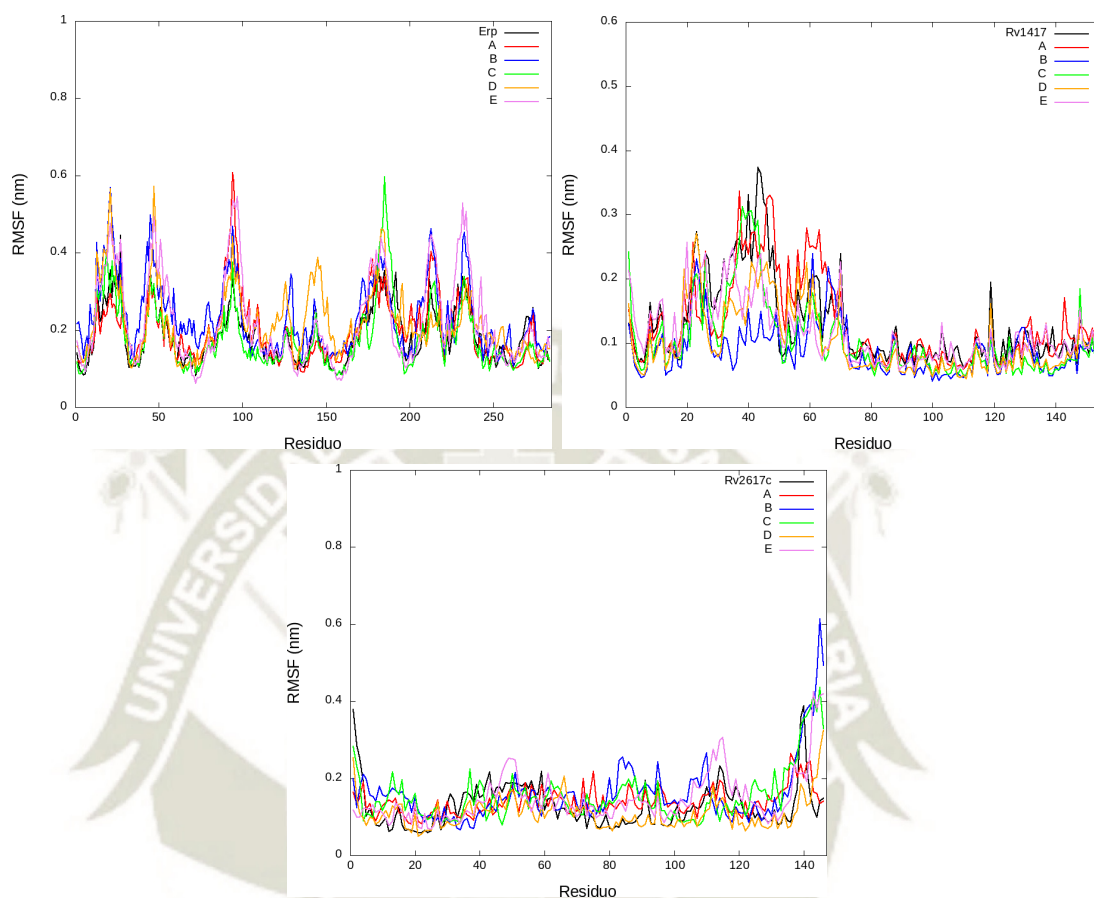


Figura 3.21: RMSF de la interacción de las tres proteínas con cada derivado de la nicotina. (a) Erp, (b) Rv1417 y (c) Rv2617c

En la tabla 3.7 se muestra valores promedio del RMSD, RMSF, Radio de Giro y los enlaces de Hidrógeno de los últimos 2ns de simulación donde, la proteína Erp con el derivado E muestra mayores cambios con respecto a su estabilidad como su conformación, con respecto a las fluctuaciones se observan elevadas tanto con el derivado B y el E, con respecto al radio de giro los derivados A y E incrementan su valor, finalmente, los enlaces de hidrógeno aumentan con los derivados A y B, sin embargo, disminuyen con el derivado E. Si bien, los enlaces de hidrógeno tienden a estabilizar las estructuras secundarias de la proteína ^{132,133}, así como también al presenciar interacciones favorables tienden a formarse enlaces de hidrógeno. La proteína Rv1417 con el derivado A muestra cambios drásticos, pero localmente se dan en pocos residuos;

con el derivado B también se observan cambios y elevadas fluctuaciones, y los enlaces de hidrógeno tienden a aumentar al igual que la proteína Rv2617c. Finalmente, la proteína Rv2617c con el derivado B la proteína muestra mayores cambios y fluctuaciones, pero los enlaces de hidrógeno tienden a aumentar, esto puede deberse a la fuerte unión entre el derivado y la proteína (ver tabla 3.7).

Tabla 3.7: RMSD, RMSF, Radio de Giro y Enlaces de hidrógeno de las tres proteínas estudiadas.

Sistema	RMSD	RMSF	Radio de G.	Enlaces de H.
Erp	0.737 ± 0.029	0.185 ± 0.072	2.389 ± 0.024	142.060 ± 6.072
A	0.642 ± 0.030	0.200 ± 0.083	2.445 ± 0.015	144.129 ± 5.470
B	0.733 ± 0.102	0.232 ± 0.092	2.428 ± 0.021	144.677 ± 5.896
Erp	0.634 ± 0.031	0.178 ± 0.085	2.418 ± 0.014	135.622 ± 5.966
C	0.634 ± 0.031	0.178 ± 0.085	2.418 ± 0.014	135.622 ± 5.966
D	0.635 ± 0.034	0.219 ± 0.088	2.427 ± 0.027	142.119 ± 5.789
E	0.856 ± 0.043	0.223 ± 0.117	2.497 ± 0.021	137.343 ± 5.136
Rv1417	0.398 ± 0.029	0.129 ± 0.067	1.616 ± 0.010	102.657 ± 4.152
A	0.567 ± 0.023	0.131 ± 0.067	1.667 ± 0.015	108.463 ± 4.882
B	0.307 ± 0.009	0.089 ± 0.040	1.534 ± 0.007	111.662 ± 4.785
Rv1417	0.408 ± 0.014	0.105 ± 0.059	1.565 ± 0.009	105.701 ± 4.825
C	0.408 ± 0.014	0.105 ± 0.059	1.565 ± 0.009	105.701 ± 4.825
D	0.374 ± 0.020	0.103 ± 0.053	1.595 ± 0.007	103.731 ± 4.504
E	0.383 ± 0.016	0.115 ± 0.048	1.599 ± 0.008	109.149 ± 4.350
Rv2617c	0.557 ± 0.015	0.129 ± 0.058	1.626 ± 0.017	90.975 ± 4.603
A	0.473 ± 0.017	0.139 ± 0.035	1.593 ± 0.012	87.279 ± 4.923
B	0.897 ± 0.034	0.160 ± 0.079	1.846 ± 0.015	93.811 ± 5.301
Rv2617c	0.574 ± 0.046	0.153 ± 0.067	1.717 ± 0.018	93.801 ± 4.313
C	0.574 ± 0.046	0.153 ± 0.067	1.717 ± 0.018	93.801 ± 4.313
D	0.493 ± 0.013	0.109 ± 0.042	1.592 ± 0.011	95.134 ± 4.104
E	0.594 ± 0.020	0.147 ± 0.066	1.721 ± 0.014	91.622 ± 4.994

RMSD, RMSF y Radio de Giro reportadas en nm.

3.3.5 Análisis de la estructura secundaria de las tres proteínas

En las siguientes figuras podemos observar la localización del acoplamiento de los derivados de la nicotina y cómo éstas interactúan con las proteínas, los enlaces de

formación y los aminoácidos involucrados en el complejo proteína–ligando. Entre las tres proteínas podemos observar que la proteína Rv1417 tiene fuertes interacciones y mayor formación de enlaces de hidrógeno con los derivados.

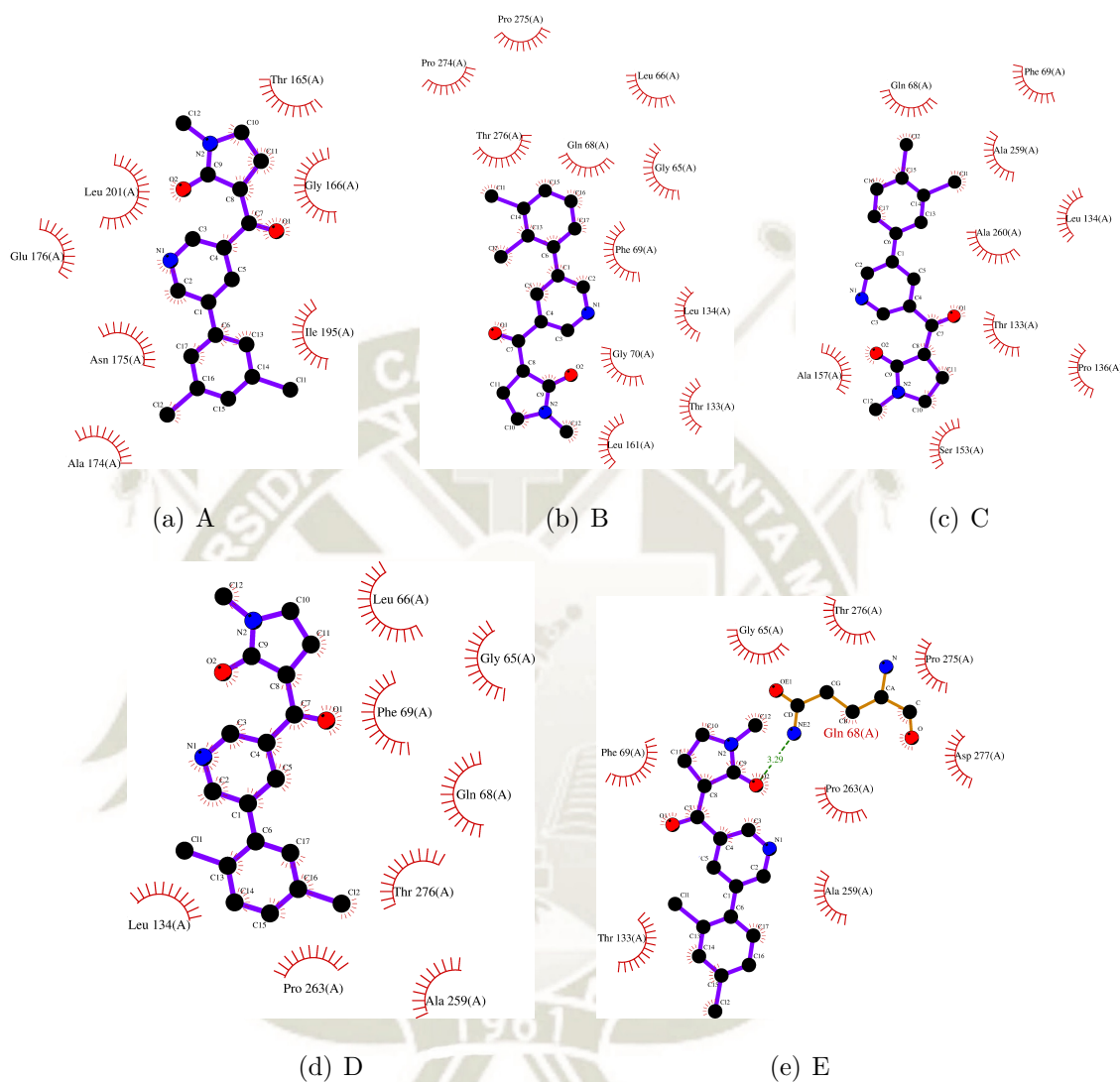


Figura 3.22: Aminoácidos interactuantes de la proteína Erp cada derivado.

Los aminoácidos involucrados con el Erp y los átomos de cloro son generalmente la alanina, treonina y prolina, en cambio con la proteína Rv2617c los aminoácidos involucrados son la fenilalanina seguido de la leucina. Para la proteína Rv1417, con mayor frecuencia se encuentra la prolina seguido de la Isoleucina. Como podemos observar, todos los aminoácidos involucrados son hidrofóbicos a excepción de la treonina que es un aminoácido polar sin carga.

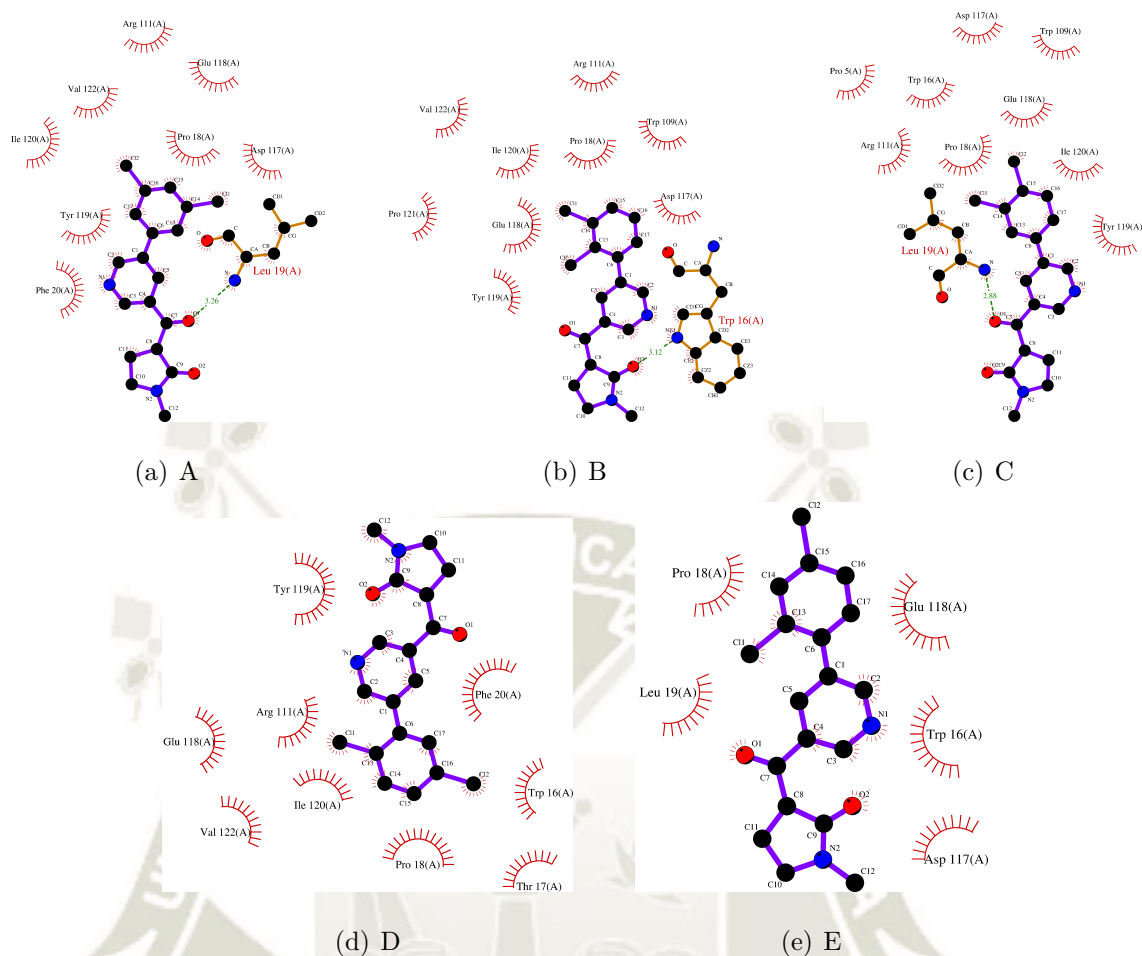


Figura 3.23: Aminoácidos interactuantes de la proteína Rv1417 con cada derivado.

Por otro lado, si nos fijamos en el derivado E del Erp y el derivado B de las proteínas Rv2617c y Rv1417 tiende a formar un enlace de hidrógeno entre el oxígeno del anillo de la pirrolidona y los aminoácidos glutamina, tirosina y el triptófano respectivamente, siendo los últimos dos aminoácidos aromáticos tal como se discutió en la predicción de los bolsillos, estos derivados de nicotina tienden a unirse a los aminoácidos aromáticos y cloros hidrofóbicos, con la excepción de que la glutamina del Erp es un aminoácido polar sin carga. La proteína Rv1417 también forma enlaces de hidrógeno entre el oxígeno del nicotinoil y la leucina, para los derivados A y C, que viene a ser un aminoácido hidrofóbico.

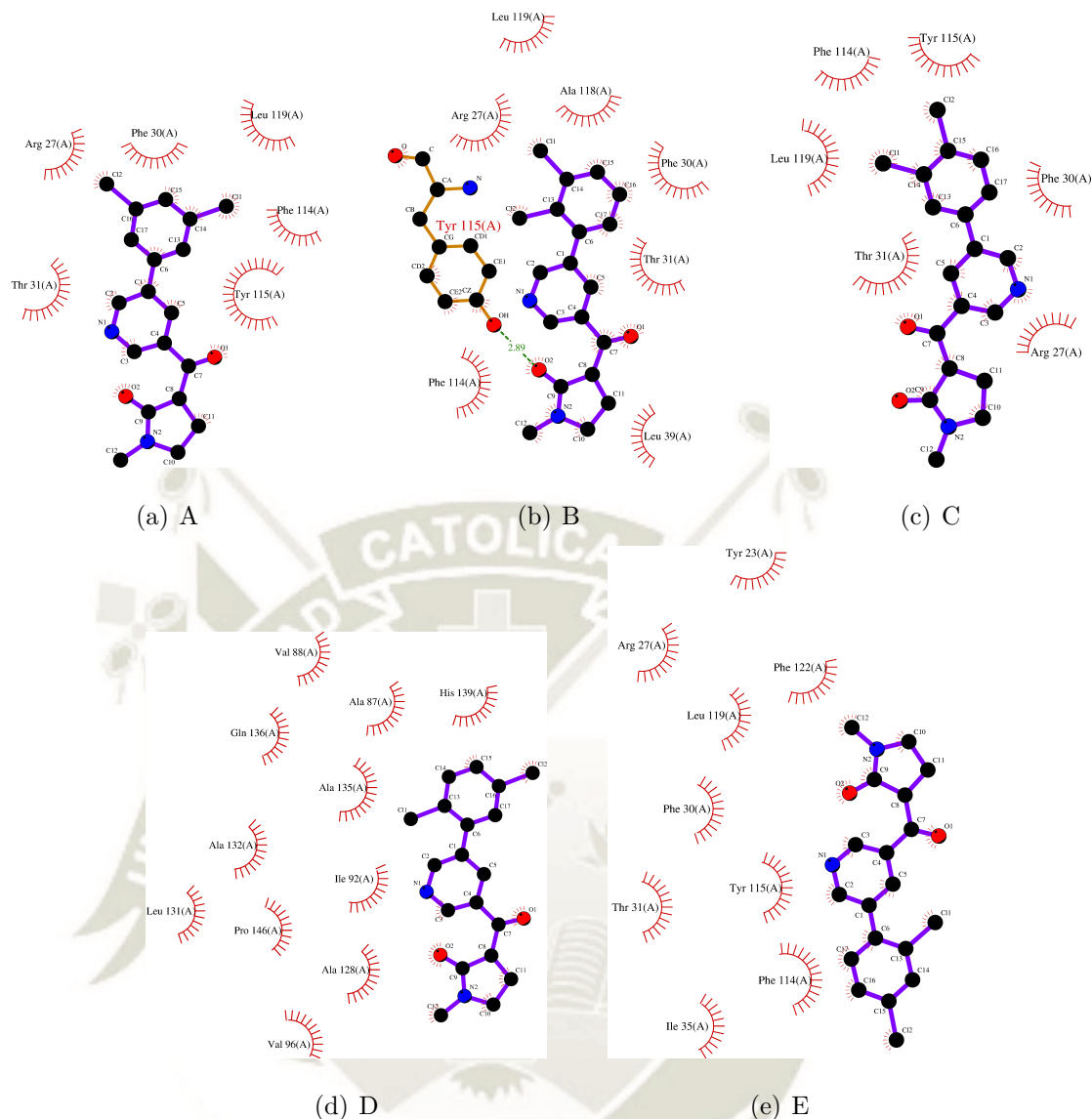


Figura 3.24: Aminoácidos interactuantes de la proteína Rv2617c con cada derivado.

A continuación, podemos observar el cambio de la estructura secundaria con los resultados del RMSF. Observamos que la Erp con el derivado A presenta cambios entre los residuos 60–65 moviéndose la posición de las estructuras hélice, y entre los residuos 150–160 observando la desaparición de estructuras hélice, por otro lado, entre los residuos 80–100 no tienen conformación alguna, formando "loops" y por lo tanto observaremos mayores fluctuaciones. Sin embargo, para la proteína Rv1417 se observa la desaparición de dos estructuras hélice entre los residuos 40–43 y 51–53, a su vez se observan fluctuaciones en regiones conformadas por "loops". Finalmente, la proteína Rv2617c sólo muestra la formación y desaparición de estructuras hélice entre los residuos 74–77 y 121–125 respectivamente. El derivado A tiende a fluctuar más con

la proteína Rv1417. (ver figura 3.25)

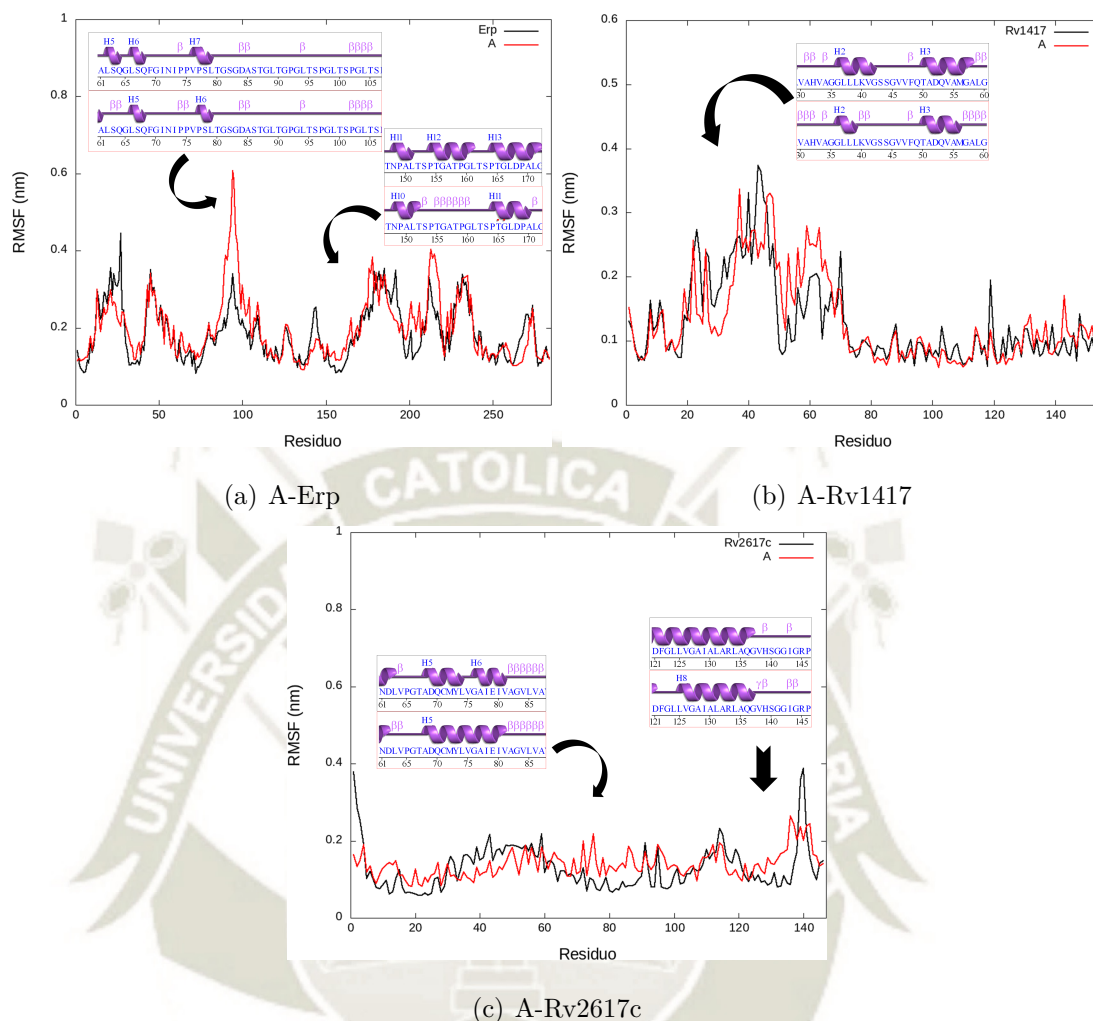


Figura 3.25: Análisis de la estructura secundaria de las tres proteínas con el derivado A.

La Erp con el derivado B muestra varias fluctuaciones que como consecuencia se observa la desaparición de estructuras hélice entre los residuos 17–23 y 200–204; en los residuos <120 se observan la formación de una alfa-hélice. En caso de la proteína Rv1417 se observa la desaparición de una alfa hélice pero en ésta zona las fluctuaciones tienden a bajar. Para la proteína Rv2617c se observa la desaparición de la lámina beta observando fluctuaciones en esa zona, también se observa la desaparición de una alfa hélice y elevadas fluctuaciones en los residuos finales de la proteína observando que éstas zonas están conformadas por "loops". Con respecto al derivado B, se observa que desestabiliza más a la proteína Rv2617c seguido de la Erp (ver figura 3.26).

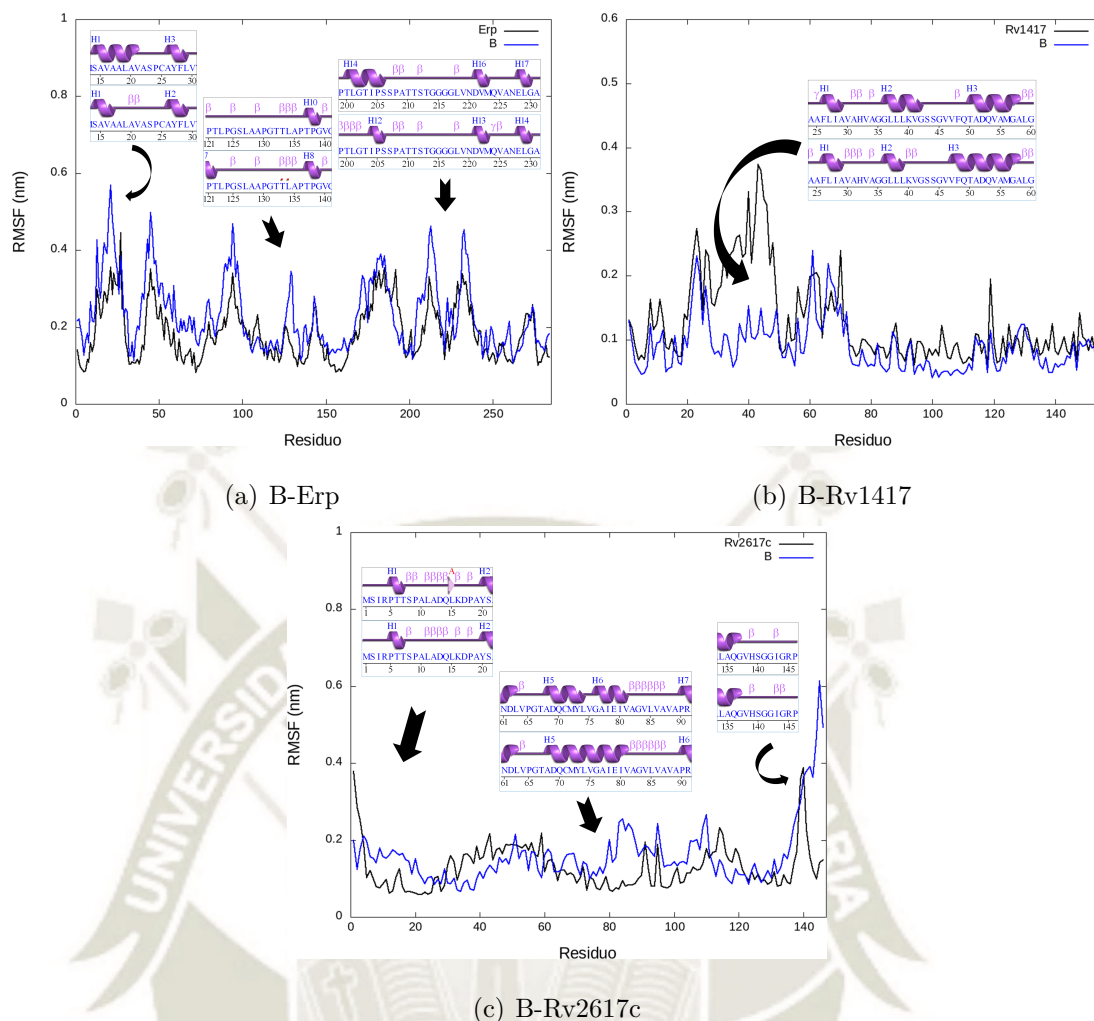


Figura 3.26: Análisis de la estructura secundaria de las tres proteínas con el derivado B.

Para el caso del derivado C con la Erp se observa una fuerte fluctuación entre los residuos 155–205 desapareciendo varias estructuras hélice. Sin embargo, para la proteína Rv1417 no se observan muchas diferencias ya que las fluctuaciones son muy parecidas a la proteína inicial. Con respecto a la proteína Rv2617c se observa la desaparición de una lámina beta y entre los residuos 79–85 se observa la desaparición de una alfa hélice. Para éste caso el derivado C las fluctuaciones con respecto a las tres proteínas no varían mucho pero actúa mejor con la Erp (ver figura 3.27).

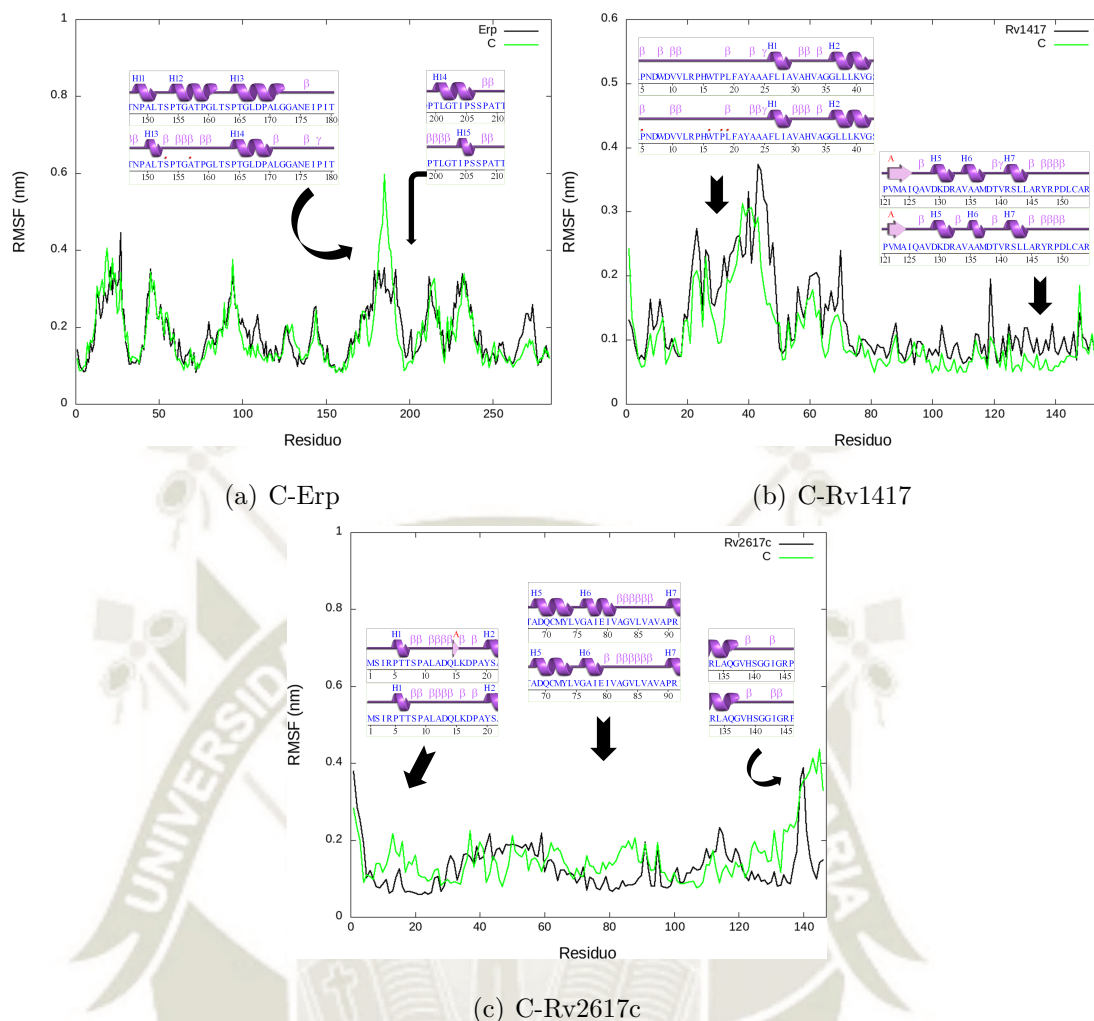


Figura 3.27: Análisis de la estructura secundaria de las tres proteínas con el derivado C.

La Erp con el derivado D tiende a presentar cambios, formando y desapareciendo estructuras hélice en los residuos 59–65, 154–161, 170–173 y 204–206. Con la proteína Rv1417 se observa la desaparición de una estructura hélice entre los residuos 30–50 y 130–145. Finalmente, para la proteína Rv2617c se observa leves fluctuaciones, teniendo la desaparición de dos estructuras hélice entre los residuos 5–8 y 125–128 (ver figura 3.28). Al igual que el derivado C, la Erp tiende a desestabilizarse más en comparación a las demás proteínas.

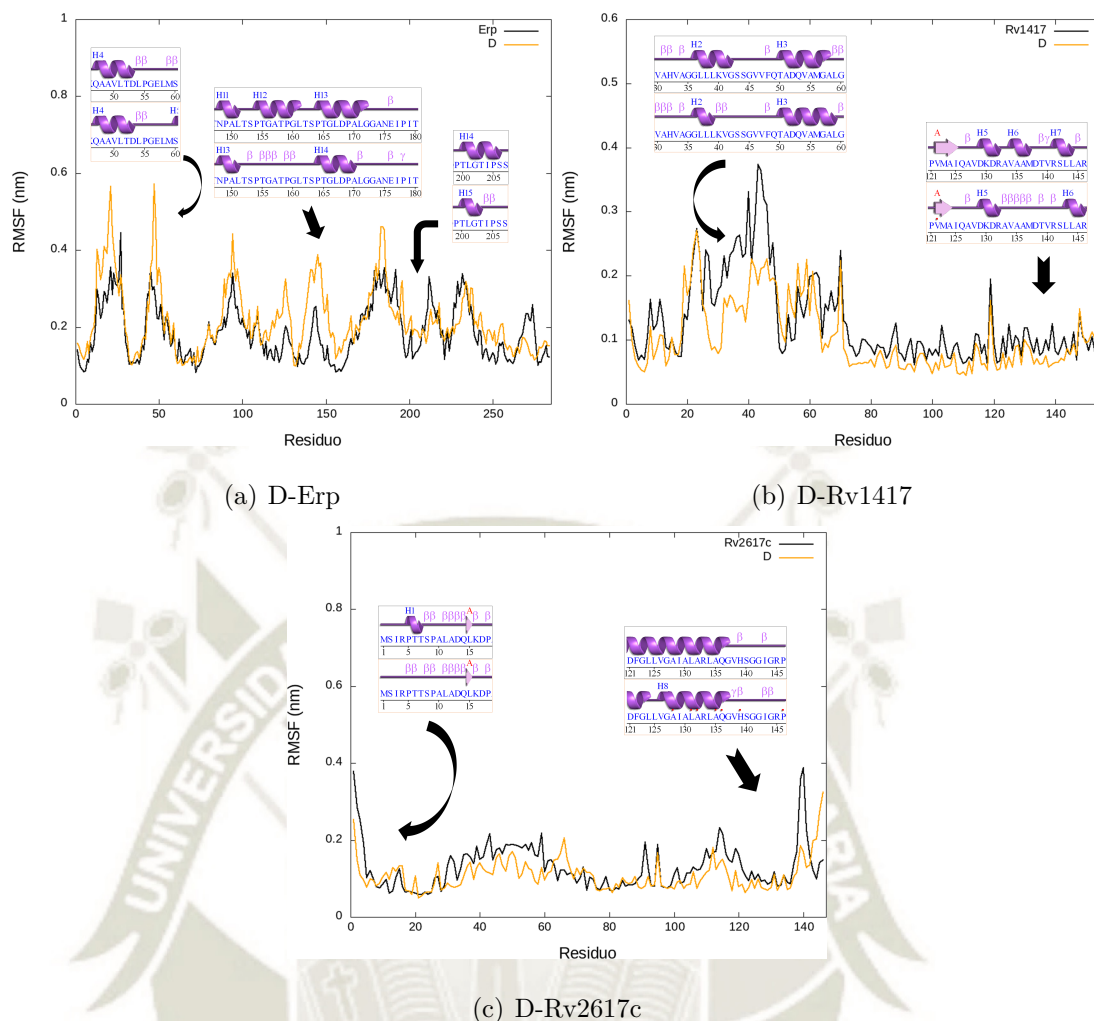


Figura 3.28: Análisis de la estructura secundaria de las tres proteínas con el derivado D.

En la figura 3.29 el derivado E con la Erp se observan mayores fluctuaciones tanto en las estructuras conformadas, como en los "loops", llegando a desaparecer las estructuras hélices. Para la proteína Rv1417 se observa la desaparición de una estructura hélice entre los residuos 130–145, con respecto a las zonas terminales se observan mayores fluctuaciones siendo zonas conformadas por "loops". Lo interesante de la proteína Rv1417 es que no desaparece por completo las lámina beta de su estructura con respecto a todos los derivados. Finalmente, la proteína Rv2617c presenta pequeñas fluctuaciones llegando a la desaparición de una lámina beta y afa hélice en la región de los residuos 10–15 y 80–83 respectivamente. Para el caso del derivado E, muestra mayor contribución con las proteínas Erp y Rv1417.

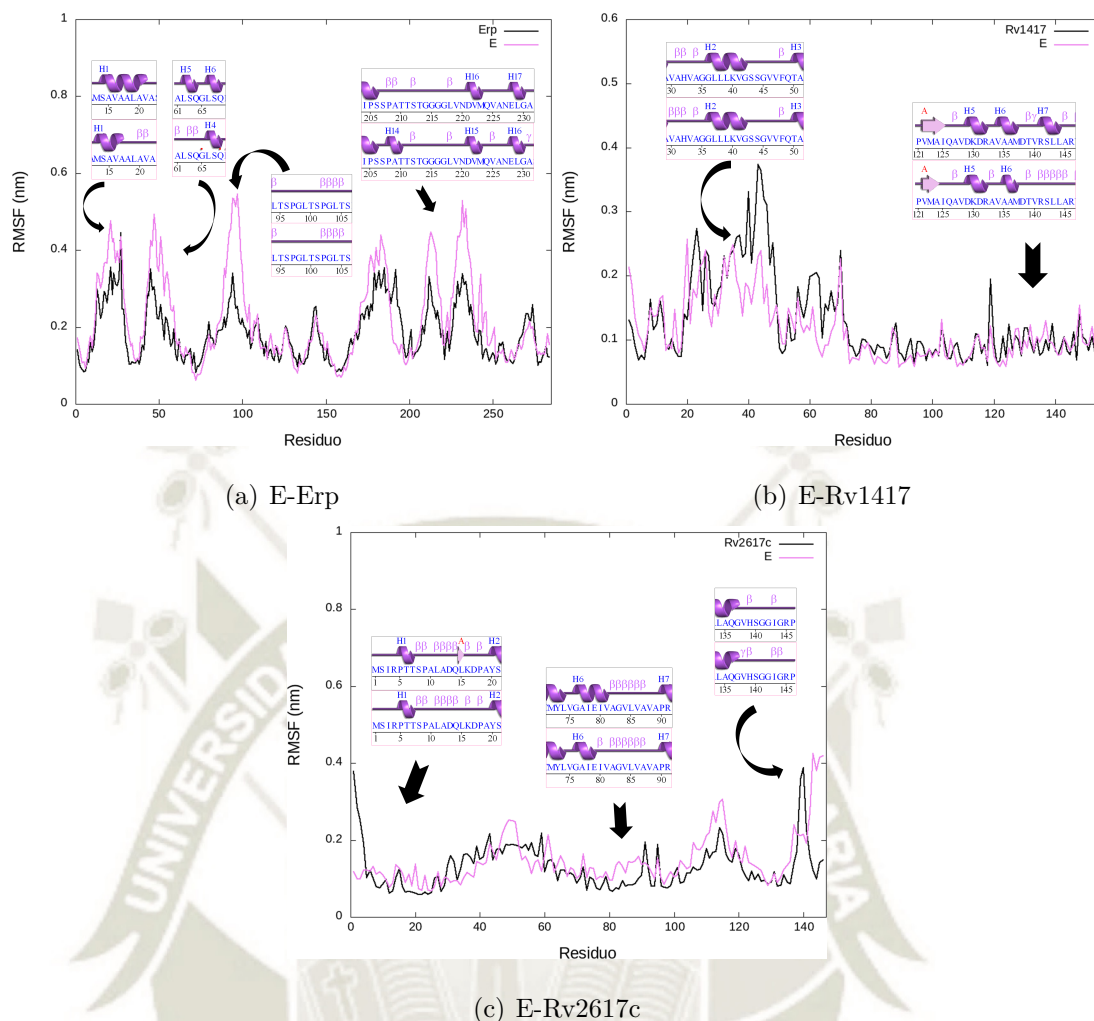


Figura 3.29: Análisis de la estructura secundaria de las tres proteínas con el derivado E.

Por otro lado, en la tabla 3.8 se muestra información de las estructuras secundarias para cada proteína después de su interacción con los derivados de la nicotina y la proteína inicial. La proteína Erp con el derivados A, C y D muestra mayor disminución de estructuras hélices, en cambio con el derivado B tiende a tener el mismo comportamiento manteniendo los alfa-hélices y disminuyendo los hélices de tipo 3_10 , a excepción del derivado E que tiende a aumentar una estructura alfa-hélice. En caso de la proteína Rv1417 con los derivados B, C, D y E se observa la disminución de cadenas beta, manteniéndose las estructuras hélices para los derivados B y C; con respecto al derivado A se observa una mayor disminución de estructuras alfa-hélices, seguido de los derivados E y D. Finalmente, la proteína Rv2617c con los derivados B, C y E se observa la desaparición de cadenas beta, aumentando una estructura alfa-hélice para

el derivado B y disminuyendo éste tipo de estructuras para los derivados C y E.

Tabla 3.8: Comparación de la estructura secundaria de las proteínas y la interacción entre proteínas y derivados

Sistema	Cadenas beta		Alfa hélice		3,10 Hélices		Otros	
	%	N°	%	N°	%	N°	%	N°
Erp	0.0	0	12.7	36	14.1	40	73.2	208
A	0.0	0	10.9	31	10.9	31	78.2	222
B	0.0	0	12.7	36	11.3	32	76.1	216
Erp	0.0	0	10.2	29	14.8	42	75.0	213
D	0.0	0	10.6	30	14.8	42	74.6	212
E	0.0	0	13.0	37	11.3	32	75.7	215
Rv1417	9.1	14	18.2	28	3.9	6	68.8	106
A	9.1	14	14.9	23	3.9	6	72.1	111
B	7.8	12	18.2	28	3.2	5	70.8	109
Rv1417	7.8	12	18.2	28	3.2	5	70.8	109
D	7.8	12	16.9	26	1.3	2	74.0	114
E	7.8	12	15.6	24	3.2	5	73.4	113
Rv2617c	2.7	4	63.7	93	2.1	3	31.5	46
A	2.7	4	61.6	90	2.1	3	33.6	49
B	0.0	0	64.4	94	2.1	3	33.6	49
Rv2617c	0.0	0	63.0	92	2.1	3	34.9	51
D	2.7	4	63.0	92	0.0	0	34.2	50
E	0.0	0	62.3	91	2.1	3	35.6	52

Teniendo en cuenta los cambios en la estructura secundaria y los valores de RMSF de las fluctuaciones de aminoácidos en las proteínas se puede observar que la Erp muestra mayores cambios al interactuar con los derivados A, B y E, la proteína Rv1417 con el derivado A y finalmente la proteína Rv2617c con el derivado B. En la tabla 3.9 podemos observar que para los 3 derivados con sus respectivas proteínas presentan aminoácidos interactuantes de tipo hidrofóbicos y aminoácidos polares sin carga a excepción de

las dos últimas proteínas que tienen presencia de un aminoácido con carga positiva que es la Arginina. La presencia de éste aminoácido interactuante podría mejorar la capacidad antimicrobiana de los derivados de la nicotina.¹³⁴ En la tabla 3.9 también podemos observar que la proteína Rv1417 tiende a formar enlaces de hidrógeno con tres derivados, esto indica un mejor acoplamiento hacia los derivados de la nicotina, asimismo ésta proteína tiende a encontrar más estabilidad con el derivado B, cuyo comportamiento nos dá una posibilidad de poder entender su comportamiento con estudios posteriores. Por otro lado, también se observa que la Erp tiende a tener un comportamiento diferente a comparación de las otras dos proteínas, esto puede deberse a los aminoácidos presentes en la proteína y a su localización dentro de la micobacteria.

Tabla 3.9: Aminoácidos interactuantes de cada proteína con los derivados de la nicotina.

Proteína	Deriv.	N° Aa.	N° E.H.	Aminoácidos
Erp	A	7	0	THR165, GLY166, ALA 174, ASN175, GLU176, ILE195, LEU201.
	B	11	0	GLY65, LEU66, GLN68, PHE69, GLY70, THR133, LEU134, LEU161, PRO274, PRO275, THR276.
	C	9	0	GLN68, PHE69, THR133, LEU134, PRO136, SER153, ALA257, ALA259, ALA260.
	D	8	0	GLY65, LEU66, GLN68, PHE69, LEU134, ALA259, PRO263, THR276.
	E	9	1	GLY65, GLN68, PHE69, THR133, ALA259, PRO263, PRO275, THR276, ASP277.
Rv1417	A	9	1	PRO18, LEU19, PHE20, ARG111, ASP117, GLU118, TYR119, ILE120, VAL122.
	B	10	0	TRP16, PRO18, TRP109, ARG111, ASP117, GLU118, TYR119, ILE120, PRO121, VAL122.
	C	10	0	PRO5, TRP16, PRO18, LEU19, TRP109, ARG111, ASP117, GLU118, TYR119, ILE120.
	D	9	0	TRP16, THR17, PRO18, PHE20, ARG111, GLU118, TYR119, ILE120, VAL122.
	E	5	1	TRP12, PRO18, LEU19, ASP117, GLU118.
Rv2617c	A	6	0	ARG27, PHE30, THR31, PHE114, TYR115, LEU119.
	B	8	1	ARG27, PHE30, THR31, LEU39, PHE114, TYR115, ALA118, LEU119.
	C	6	0	ARG27, PHE30, THR31, PHE114, TYR115, LEU119.
	D	11	0	ALA87, VAL88, ILE92, VAL96, ALA128, LEU131, ALA132, ALA135, GLN136, HIS139, PRO146.
	E	9	1	TYR23, ARG27, PHE30, THR31, ILE35, PHE114, TYR115, LEU119, PHE122.

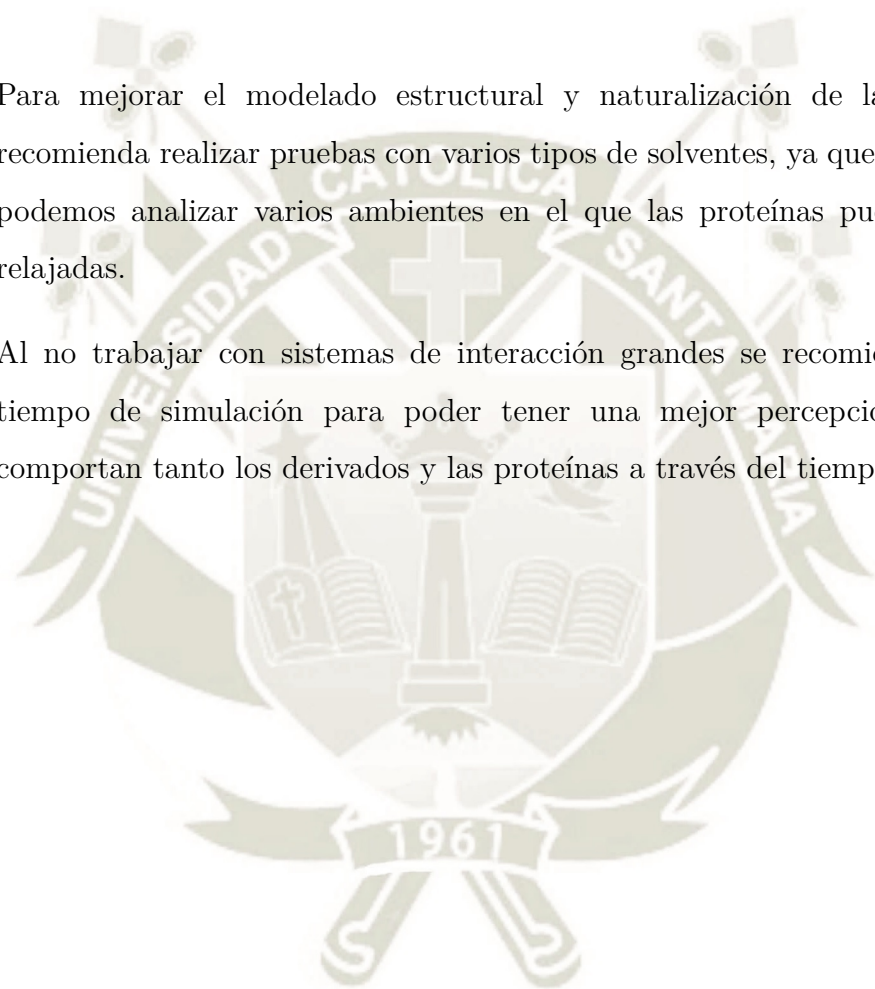
Deriv. = Derivados, Aa = Aminoácidos, E.H. = Enlaces de Hidrógeno.

Conclusiones

1. Se obtuvieron las estructuras moleculares de los cinco derivados de la nicotina y se modelaron las proteínas Erp, Rv1417 y Rv2617c del *Mycobacterium tuberculosis* en su estado tridimensional.
2. A partir de un buen modelado estructural se minimizaron y estabilizaron las tres proteínas a condiciones fisiológicas mediante técnicas de mecánica molecular llegando a un mínimo global.
3. Se realizó una simulación de dinámica molecular mediante un ensamble NPT a 200 ns para las tres proteínas a condiciones fisiológicas, teniendo mejores resultados con la proteína Erp al momento de realizar su validación en el diagrama de Ramchandran y calidad. Sin embargo, en los análisis del RMSD, RMSF y Radio de Giro dieron mejores resultados para las proteínas Rv1417 y Rv2617c.
4. Se realizaron las interacciones moleculares entre las moléculas de los derivados de la nicotina y las proteínas Erp, Rv1417 y Rv2617c, obteniendo altos valores de ACE y drugabilidad que favorecieron dichas interacciones.
5. Finalmente, se evaluaron las interacciones moleculares entre los cinco derivados de la nicotina y las proteínas Erp, Rv1417 y Rv2617c. Se llegó a obtener la desestabilización y pérdida de estructuras secundarias de las proteínas al interactuar con los derivados A, B y E concluyendo que al no tener una conformación estructural pierdan su actividad funcional dentro del *Mycobacterium tuberculosis*, que implica la vulnerabilidad de supervivencia de la micobacteria, brindando así, una base para el desarrollo de nuevos blancos terapéuticos.

Recomendaciones

- Para mejorar el modelado estructural y naturalización de las proteínas se recomienda realizar pruebas con varios tipos de solventes, ya que de esta manera podemos analizar varios ambientes en el que las proteínas puedan estar más relajadas.
- Al no trabajar con sistemas de interacción grandes se recomienda darle más tiempo de simulación para poder tener una mejor percepción de como se comportan tanto los derivados y las proteínas a través del tiempo.



Referencias Bibliográficas

- [1] Jason D Simmons, Catherine M Stein, Chetan Seshadri, Monica Campo, Galit Alter, Sarah Fortune, Erwin Schurr, Robert S Wallis, Gavin Churchyard, Harriet Mayanja-Kizza, et al. Immunological mechanisms of human resistance to persistent mycobacterium tuberculosis infection. *Nature Reviews Immunology*, 18(9):575, 2018.
- [2] Leila de Mendonça-Lima, Mathieu Picardeau, Catherine Raynaud, Jean Rauzier, Yves-Olivier Goguet de la Salmonière, Lucia Barker, Fabiana Bigi, Angel Cataldi, Brigitte Gicquel, and Jean-Marc Reyrat. Erp, an extracellular protein family specific to mycobacteria. *Microbiology*, 147(8):2315–2320, 2001.
- [3] Christine L Cosma, Kathryn Klein, Rosa Kim, Dana Beery, and Lalita Ramakrishnan. Mycobacterium marinum erp is a virulence determinant required for cell wall integrity and intracellular survival. *Infection and immunity*, 74(6):3125–3133, 2006.
- [4] François-Xavier Berthet, Jean Rauzier, Eng Mong Lim, Wolfgang Philipp, Brigitte Gicquel, and Denis Portnoi. Characterization of the mycobacterium tuberculosis erp gene encoding a potential cell surface protein with repetitive structures. *Microbiology*, 141(9):2123–2130, 1995.
- [5] François-Xavier Berthet, Micheline Lagranderie, Pierre Gounon, Christine Laurent-Winter, Danielle Ensergueix, Pierre Chavarot, Françoise Thouron, Eddie Maranghi, Vladimir Pelicic, Denis Portnoi, et al. Attenuation of virulence by disruption of the mycobacterium tuberculosis erp gene. *Science*, 282(5389):759–762, 1998.

- [6] Laura I Klepp, Marcelo Soria, Federico C Blanco, María V Bianco, María P Santangelo, Angel A Cataldi, and Fabiana Bigi. Identification of two proteins that interact with the erp virulence factor from mycobacterium tuberculosis by using the bacterial two-hybrid system. *BMC molecular biology*, 10(1):3, 2009.
- [7] Jennifer Furin, Helen Cox, and Madhukar Pai. Tuberculosis. *THE LANCET*, 393(10181):1642–1656, 2019.
- [8] Olivia Oxlade, Anete Trajman, Andrea Benedetti, Mênonli Adjobimey, Victoria J Cook, Dina Fisher, Gregory James Fox, Federica Fregonese, Panji Hadisoemarto, Philip C Hill, et al. Enhancing the public health impact of latent tuberculosis infection diagnosis and treatment (act4): protocol for a cluster randomised trial. *BMJ open*, 9(3):e025831, 2019.
- [9] Adrián Sánchez-Montalvá, Fernando Salvador, Daniel Molina-Morant, and Israel Molina. Tuberculosis e inmigración. *Enfermedades infecciosas y microbiología clinica*, 36(7):446–455, 2018.
- [10] Richard Douglas Turner. Cough in pulmonary tuberculosis: Existing knowledge and general insights. *Pulmonary pharmacology & therapeutics*, 2019.
- [11] Joseph Mendy, Sheikh Jarju, Rhiannon Heslop, Adama L Bojang, Beate Kampmann, and Jayne S Sutherland. Changes in mycobacterium tuberculosis-specific immunity with influenza co-infection at time of tb diagnosis. *Frontiers in immunology*, 9, 2018.
- [12] Gerhard Walzl, Katharina Ronacher, Willem Hanekom, Thomas J Scriba, and Alimuddin Zumla. Immunological biomarkers of tuberculosis. *Nature Reviews Immunology*, 11(5):343, 2011.
- [13] Thomas J Braciale, Jie Sun, and Taeg S Kim. Regulating the adaptive immune response to respiratory virus infection. *Nature Reviews Immunology*, 12(4):295, 2012.
- [14] Paul S Redford, Katrin D Mayer-Barber, Finlay W McNab, Evangelos Stavropoulos, Andreas Wack, Alan Sher, and Anne O’Garra. Influenza a virus impairs control of mycobacterium tuberculosis coinfection through a type

- i interferon receptor–dependent pathway. *The Journal of infectious diseases*, 209(2):270–274, 2013.
- [15] Lewis K Schragar, Angelo Izzo, and Kamalakannan Velmurugan. Immunopathogenesis of tuberculosis and novel mechanisms of vaccine activity. *Tuberculosis*, 99:S3–S7, 2016.
- [16] Phillip P Salvatore and Ying Zhang. Tuberculosis: molecular basis of pathogenesis. 2017.
- [17] Kamran Siddiqi, Marie-Laurence Lambert, and John Walley. Clinical diagnosis of smear-negative pulmonary tuberculosis in low-income countries: the current evidence. *The Lancet infectious diseases*, 3(5):288–296, 2003.
- [18] Evangelia Skoura, Alimuddin Zumla, and Jamshed Bomanji. Imaging in tuberculosis. *International Journal of Infectious Diseases*, 32:87–93, 2015.
- [19] Jamshed B Bomanji, Narainder Gupta, Parveen Gulati, and Chandan J Das. Imaging in tuberculosis. *Cold Spring Harbor perspectives in medicine*, 5(6):a017814, 2015.
- [20] Bareng AS Nonyane, Mark P Nicol, Nicholas J Andreas, Stefanie Rimmele, Nicole Schneiderhan-Marra, Lesley J Workman, Mark D Perkins, Thomas Joos, Tobias Broger, Jerrold J Ellner, et al. Serologic responses in childhood pulmonary tuberculosis. *The Pediatric infectious disease journal*, 37(1):1, 2018.
- [21] Jennifer Furin. Advances in the diagnosis, treatment, and prevention of tuberculosis in children. *Expert review of respiratory medicine*, 13(3):301–311, 2019.
- [22] Lily V Bonadonna, Matthew J Saunders, Heinner Guio, Roberto Zegarra, and Carlton A Evans. Socioeconomic and behavioral factors associated with tuberculosis diagnostic delay in lima, peru. *The American journal of tropical medicine and hygiene*, 98(6):1614–1623, 2018.
- [23] Hannah Alsdurf, Philip C Hill, Alberto Matteelli, Haileyesus Getahun, and Dick Menzies. The cascade of care in diagnosis and treatment of latent tuberculosis

- infection: a systematic review and meta-analysis. *The Lancet Infectious Diseases*, 16(11):1269–1278, 2016.
- [24] Sadaf Kalsum. *Characterizing phenotypes of Mycobacterium tuberculosis and exploring anti-mycobacterial compounds through high content screening*, volume 1657. Linköping University Electronic Press, 2019.
- [25] Shiva Kumar Angala, Juan Manuel Belardinelli, Emilie Huc-Claustre, William H Wheat, and Mary Jackson. The cell envelope glycoconjugates of mycobacterium tuberculosis. *Critical reviews in biochemistry and molecular biology*, 49(5):361–399, 2014.
- [26] P Bifani, S Moghazeh, B Shopsin, J Driscoll, A Ravikovitch, and BN Kreiswirth. Molecular characterization of mycobacterium tuberculosis h37rv/ra variants: distinguishing the mycobacterial laboratory strain. *Journal of clinical microbiology*, 38(9):3200–3204, 2000.
- [27] ST Cole and BG Barrell. Analysis of the genome of mycobacterium tuberculosis h37rv. In *Genetics and Tuberculosis: Novartis Foundation Symposium 217*, pages 160–177. Wiley Online Library, 1998.
- [28] Chunyan Li, Qiming Li, Yuan Zhang, Zhen Gong, Sai Ren, Ping Li, and Jianping Xie. Characterization and function of mycobacterium tuberculosis h37rv lipase rv1076 (lipu). *Microbiological research*, 196:7–16, 2017.
- [29] STea Cole, R Brosch, J Parkhill, T Garnier, C Churcher, D Harris, SV Gordon, K Eiglmeier, S Gas, CE Barry Iii, et al. Deciphering the biology of mycobacterium tuberculosis from the complete genome sequence. *Nature*, 393(6685):537, 1998.
- [30] Otmame Lamrabet and Michel Drancourt. Genetic engineering of mycobacterium tuberculosis: A review. *Tuberculosis*, 92(5):365–376, 2012.
- [31] M Daffe and G Etienne. The capsule of mycobacterium tuberculosis and its implications for pathogenicity. *Tubercle and lung disease*, 79(3):153–169, 1999.
- [32] Gloria P Monterrubio-López, Rosa María Ribas-Aparicio, et al. Identification of novel potential vaccine candidates against tuberculosis based on reverse vaccinology. *BioMed research international*, 2015, 2015.

- [33] Sebastien Gagneux. Ecology and evolution of mycobacterium tuberculosis. *Nature Reviews Microbiology*, 16(4):202, 2018.
- [34] Zoheira Djelouadji, Didier Raoult, and Michel Drancourt. Palaeogenomics of mycobacterium tuberculosis: epidemic bursts with a degrading genome. *The Lancet infectious diseases*, 11(8):641–650, 2011.
- [35] Tanya Parish, Gretta Roberts, Francoise Laval, Merrill Schaeffer, Mamadou Daffé, and Ken Duncan. Functional complementation of the essential gene *fabg1* of mycobacterium tuberculosis by mycobacterium smegmatis *fabg* but not *escherichia coli* *fabg*. *Journal of bacteriology*, 189(10):3721–3728, 2007.
- [36] Tanya Parish, Merrill Schaeffer, Gretta Roberts, and Ken Duncan. Hemz is essential for heme biosynthesis in mycobacterium tuberculosis. *Tuberculosis*, 85(3):197–204, 2005.
- [37] Jennifer C Chang, Nada S Harik, Reiling P Liao, and David R Sherman. Identification of mycobacterial genes that alter growth and pathology in macrophages and in mice. *The Journal of infectious diseases*, 196(5):788–795, 2007.
- [38] Courtney E Maus, Bonnie B Plikaytis, and Thomas M Shinnick. Mutation of *tlyA* confers capreomycin resistance in mycobacterium tuberculosis. *Antimicrobial agents and chemotherapy*, 49(2):571–577, 2005.
- [39] Kevin Pethe, Dana L Swenson, Sylvie Alonso, Jennifer Anderson, Carren Wang, and David G Russell. Isolation of mycobacterium tuberculosis mutants defective in the arrest of phagosome maturation. *Proceedings of the National Academy of Sciences*, 101(37):13642–13647, 2004.
- [40] Jeanett Bauer, Vibeke O Thomsen, Sigrid Poulsen, and Ase B Andersen. False-positive results from cultures of mycobacterium tuberculosis due to laboratory cross-contamination confirmed by restriction fragment length polymorphism. *Journal of clinical microbiology*, 35(4):988–991, 1997.
- [41] Jarmila Vinsova and Martin Kratky. *Development of New MDR-Tuberculosis Drugs*. Nova Science Publisher/Nova Biomedical, 2010.

- [42] Tianao Yuan and Nicole S Sampson. Hit generation in tb drug discovery: from genome to granuloma. *Chemical reviews*, 118(4):1887–1916, 2018.
- [43] Michael L Chen, Akshith Doddi, Jimmy Royer, Luca Freschi, Marco Schito, Matthew Ezewudo, Isaac S Kohane, Andrew Beam, and Maha Farhat. Beyond multidrug resistance: Leveraging rare variants with machine and statistical learning models in mycobacterium tuberculosis resistance prediction. *EBioMedicine*, 2019.
- [44] Kamal Dua, Vamshi Krishna Rapalli, Shakti Dhar Shukla, Gautam Singhvi, Madhur D Shastri, Dinesh Kumar Chellappan, Saurabh Satija, Meenu Mehta, Monica Gulati, Terezinha De Jesus Andreoli Pinto, et al. Multi-drug resistant mycobacterium tuberculosis & oxidative stress complexity: Emerging need for novel drug delivery approaches. *Biomedicine & Pharmacotherapy*, 107:1218–1229, 2018.
- [45] Arsheed Ahmad Ganaie, Garima Trivedi, Amanpreet Kaur, Sidharth Shankar Jha, Shashi Anand, Vibhuti Rana, Amit Singh, Shekhar Kumar, and Charu Sharma. Interaction of erp protein of mycobacterium tuberculosis with rv2212 enhances intracellular survival of mycobacterium smegmatis. *Journal of bacteriology*, 198(20):2841–2852, 2016.
- [46] Dana Kocíncová, Berit Sondén, Leila de Mendonça-Lima, Brigitte Gicquel, and Jean-Marc Reyrat. The erp protein is anchored at the surface by a carboxy-terminal hydrophobic domain and is important for cell-wall structure in mycobacterium smegmatis. *FEMS microbiology letters*, 231(2):191–196, 2004.
- [47] Luis Solans, Santiago Uranga, Nacho Aguilo, Carmen Arnal, Ana Belen Gomez, Marta Monzon, Juan Jose Badiola, Brigitte Gicquel, and Carlos Martin. Hyper-attenuated mtbvac erp mutant protects against tuberculosis in mice. *Vaccine*, 32(40):5192–5197, 2014.
- [48] Arshid Yousefi-Avarvand, Mohsen Tafaghodi, Saman Soleimanpour, and Farzad Khademi. Hspx protein as a candidate vaccine against mycobacterium tuberculosis: an overview. *Frontiers in Biology*, 13(4):293–296, 2018.

- [49] Niels Zondervan, Jesse van Dam, Peter Schaap, Vitor Martins dos Santos, and Maria Suarez-Diez. Regulation of three virulence strategies of mycobacterium tuberculosis: a success story. *International journal of molecular sciences*, 19(2):347, 2018.
- [50] Harsha P Gunawardena, Meghan E Feltcher, John A Wrobel, Sheng Gu, Miriam Braunstein, and Xian Chen. Comparison of the membrane proteome of virulent mycobacterium tuberculosis and the attenuated mycobacterium bovis bcg vaccine strain by label-free quantitative proteomics. *Journal of proteome research*, 12(12):5463–5474, 2013.
- [51] Clemens Hermann, Alexander D Giddey, Andrew JM Nel, Nelson C Soares, and Jonathan M Blackburn. Cell wall enrichment unveils proteomic changes in the cell wall during treatment of mycobacterium smegmatis with sub-lethal concentrations of rifampicin. *Journal of proteomics*, 191:166–179, 2019.
- [52] Richard M Armstrong, Dominique C Carter, Samantha N Atkinson, Scott S Terhune, and Thomas C Zahrt. Association of mycobacterium proteins with lipid droplets. *Journal of bacteriology*, 200(16):e00240–18, 2018.
- [53] Patricia A DiGiuseppe Champion and Jeffery S Cox. Protein secretion systems in mycobacteria. *Cellular microbiology*, 9(6):1376–1384, 2007.
- [54] Marina A Forrellad, Laura I Klepp, Andrea Gioffré, Julia Sabio y Garcia, Hector R Morbidoni, María De La Paz Santangelo, Angel A Cataldi, and Fabiana Bigi. Virulence factors of the mycobacterium tuberculosis complex. *Virulence*, 4(1):3–66, 2013.
- [55] Apoorva Bhatt, Nagatoshi Fujiwara, Kiranmai Bhatt, Sudagar S Gurcha, Laurent Kremer, Bing Chen, John Chan, Steven A Porcelli, Kazuo Kobayashi, Gurdyal S Besra, et al. Deletion of kasB in mycobacterium tuberculosis causes loss of acid-fastness and subclinical latent tuberculosis in immunocompetent mice. *Proceedings of the National Academy of Sciences*, 104(12):5157–5162, 2007.
- [56] Michael S Glickman, Jeffery S Cox, and William R Jacobs Jr. A novel mycolic acid cyclopropane synthetase is required for cording, persistence, and virulence of mycobacterium tuberculosis. *Molecular cell*, 5(4):717–727, 2000.

- [57] Lintao Bu and Charles L Brooks Iii. De novo prediction of the structures of m. tuberculosis membrane proteins. *Journal of the American Chemical Society*, 130(16):5384–5385, 2008.
- [58] Houhui Song, Reatha Sandie, Ying Wang, Miguel A Andrade-Navarro, and Michael Niederweis. Identification of outer membrane proteins of mycobacterium tuberculosis. *Tuberculosis*, 88(6):526–544, 2008.
- [59] Dominique Sydow, Lindsey Burggraaff, Angelika Szengel, Herman WT van Vlijmen, Adriaan P IJzerman, Gerard JP van Westen, and Andrea Volkamer. Advances and challenges in computational target prediction. *Journal of chemical information and modeling*, 59(5):1728–1742, 2019.
- [60] Nina Chatterjee, Rupal Ojha, Nazia Khatoon, and Vijay Kumar Prajapati. Scrutinizing mycobacterium tuberculosis membrane and secretory proteins to formulate multiepitope subunit vaccine against pulmonary tuberculosis by utilizing immunoinformatic approaches. *International journal of biological macromolecules*, 118:180–188, 2018.
- [61] Paresh T Gandhi, Thimmasandra Narayanappa Athmaram, and Gundaiah Ramesh Arunkumar. Novel nicotine analogues with potential anti-mycobacterial activity. *Bioorganic & medicinal chemistry*, 24(8):1637–1647, 2016.
- [62] Tomohiro Kinjo, Yuji Koseki, Maiko Kobayashi, Atsumi Yamada, Koji Morita, Kento Yamaguchi, Ryoya Tsurusawa, Gulcin Gulden, Hideyuki Komatsu, Hiroshi Sakamoto, et al. Identification of compounds with potential antibacterial activity against mycobacterium through structure-based drug screening. *Journal of chemical information and modeling*, 53(5):1200–1212, 2013.
- [63] Wei Hong, Yu Wang, Zhe Chang, Yanhui Yang, Jing Pu, Tao Sun, Sargit Kaur, James C Sacchettini, Hunmin Jung, Wee Lin Wong, et al. The identification of novel mycobacterium tuberculosis dhfr inhibitors and the investigation of their binding preferences by using molecular modelling. *Scientific reports*, 5:15328, 2015.

- [64] Maria CS Lourenco, Marcus VN de Souza, Alessandra C Pinheiro, Marcelle de L Ferreira, Raoni SB Gonçalves, Thais Cristina M Nogueira, and Monica A Peralta. Evaluation of anti-tubercular activity of nicotinic and isoniazid analogues. *Arxivoc*, 15:181–191, 2007.
- [65] PT Gandhi and TN Athmaram. Evaluation of novel nicotine analogues for their anti-bacterial and anti-fungal activity. *J Microbiol Exp*, 3(1):00079, 2016.
- [66] Rakesh Narang, Balasubramanian Narasimhan, Sunil Sharma, Dharmarajan Sriram, Perumal Yogeewari, Erik De Clercq, Christophe Pannecouque, and Jan Balzarini. Synthesis, antimycobacterial, antiviral, antimicrobial activities, and qsar studies of nicotinic acid benzylidene hydrazide derivatives. *Medicinal Chemistry Research*, 21(8):1557–1576, 2012.
- [67] Kal Renganathan Sharma. *Bioinformatics: sequence alignment and Markov models*. McGraw Hill Professional, 2008.
- [68] Venkatarajan Mathura and Pandjassarame Kanguane. *Bioinformatics: a concept-based introduction*. Springer Science & Business Media, 2008.
- [69] Andreas D Baxevanis and BF Francis Ouellette. *Bioinformatics: a practical guide to the analysis of genes and proteins*, volume 43. John Wiley & Sons, 2004.
- [70] Shiyi Shen. *Theory and Mathematical methods in Bioinformatics*. Springer Science & Business Media, 2008.
- [71] William R Pearson. An introduction to sequence similarity (“homology”) searching. *Current protocols in bioinformatics*, 42(1):3–1, 2013.
- [72] András Fiser and Andrej Šali. Modeller: generation and refinement of homology-based protein structure models. In *Methods in enzymology*, volume 374, pages 461–491. Elsevier, 2003.
- [73] Jianyi Yang and Yang Zhang. I-tasser server: new development for protein structure and function predictions. *Nucleic acids research*, 43(W1):W174–W181, 2015.

- [74] Jian Peng and Jinbo Xu. Low-homology protein threading. *Bioinformatics*, 26(12):i294–i300, 2010.
- [75] Juan Lin, Yiwen Zhong, Ena Li, Xiaoyu Lin, and Hui Zhang. Multi-agent simulated annealing algorithm with parallel adaptive multiple sampling for protein structure prediction in ab off-lattice model. *Applied Soft Computing*, 62:491–503, 2018.
- [76] Ozan S Sariyer and Can Güven. Sequence alignment using simulated annealing. *Physica A: Statistical Mechanics and its Applications*, 389(15):3007–3012, 2010.
- [77] Srinivasaraghavan Kannan and Martin Zacharias. Simulated annealing coupled replica exchange molecular dynamics—an efficient conformational sampling method. *Journal of structural biology*, 166(3):288–294, 2009.
- [78] Carlos Pérez and Angel R Ortiz. Evaluation of docking functions for protein-ligand docking. *Journal of medicinal chemistry*, 44(23):3768–3785, 2001.
- [79] Meenakshi Gupta, Ruchika Sharma, and Anoop Kumar. Docking techniques in pharmacology: How much promising? *Computational biology and chemistry*, 76:210–217, 2018.
- [80] Yi Fu, Zhiguo Chen, and Jun Sun. Random drift particle swarm optimisation algorithm for highly flexible protein-ligand docking. *Journal of theoretical biology*, 457:180–189, 2018.
- [81] José García-Nieto, Esteban López-Camacho, María Jesús García-Godoy, Antonio J Nebro, and José F Aldana-Montes. Multi-objective ligand-protein docking with particle swarm optimizers. *Swarm and evolutionary computation*, 44:439–452, 2019.
- [82] Giovanni Bottegoni, Irina Kufareva, Maxim Totrov, and Ruben Abagyan. Four-dimensional docking: a fast and accurate account of discrete receptor flexibility in ligand docking. *Journal of medicinal chemistry*, 52(2):397–406, 2008.
- [83] KI Ramachandran, Gopakumar Deepa, and Krishnan Namboori. *Computational chemistry and molecular modeling: principles and applications*. Springer Science & Business Media, 2008.

- [84] E Lewars et al. Introduction to the theory and applications of molecular and quantum mechanics. *J. Computational chemistry, Ontario Canada*, 2003.
- [85] Francis A Carey and Richard J Sundberg. *Advanced organic chemistry: part A: structure and mechanisms*. Springer Science & Business Media, 2007.
- [86] Thomas T Joseph and Jérôme Hémin. Molecular mechanics parameterization of anesthetic molecules. In *Methods in enzymology*, volume 602, pages 61–76. Elsevier, 2018.
- [87] Ben Leimkuhler and Charles Matthews. *Molecular Dynamics*. Springer, 2016.
- [88] Marco De Vivo, Matteo Masetti, Giovanni Bottegoni, and Andrea Cavalli. Role of molecular dynamics and related methods in drug discovery. *Journal of medicinal chemistry*, 59(9):4035–4061, 2016.
- [89] Jaroslav M Ilnytskyi and Mark R Wilson. A domain decomposition molecular dynamics program for the simulation of flexible molecules of spherically-symmetrical and nonspherical sites. ii. extension to nvt and npt ensembles. *Computer physics communications*, 148(1):43–58, 2002.
- [90] Luis Domingo, Mar Ríos-Gutiérrez, and Patricia Pérez. Applications of the conceptual density functional theory indices to organic chemistry reactivity. *Molecules*, 21(6):748, 2016.
- [91] Luis R Domingo, Patricia Pérez, and José A Sáez. Understanding the local reactivity in polar organic reactions through electrophilic and nucleophilic parr functions. *RSC Advances*, 3(5):1486–1494, 2013.
- [92] Paul Geerlings, F De Proft, and W Langenaeker. Conceptual density functional theory. *Chemical reviews*, 103(5):1793–1874, 2003.
- [93] Shu-Bin Liu. Conceptual density functional theory and some recent developments. *Acta Physico-Chimica Sinica*, 25(3):590–600, 2009.
- [94] Pratim K Chattaraj and Debesh R Roy. Conceptual density functional theory of chemical reactivity. In *Advances in Mathematical Chemistry and Applications*, pages 196–221. Elsevier, 2015.

- [95] MJ Frisch, GW Trucks, HB Schlegel, GE Scuseria, MA Robb, JR Cheeseman, G Scalmani, V Barone, GA Petersson, H Nakatsuji, et al. Gaussian 16. *Revision A*, 3, 2016.
- [96] Mark James Abraham, Teemu Murtola, Roland Schulz, Szilárd Páll, Jeremy C Smith, Berk Hess, and Erik Lindahl. Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1:19–25, 2015.
- [97] Eric F Pettersen, Thomas D Goddard, Conrad C Huang, Gregory S Couch, Daniel M Greenblatt, Elaine C Meng, and Thomas E Ferrin. Ucsf chimera—a visualization system for exploratory research and analysis. *Journal of computational chemistry*, 25(13):1605–1612, 2004.
- [98] John E Stone, Ryan McGreevy, Barry Isralewitz, and Klaus Schulten. Gpu-accelerated analysis and visualization of large structures solved by molecular dynamics flexible fitting. *Faraday discussions*, 169:265–283, 2014.
- [99] Ambrish Roy, Alper Kucukural, and Yang Zhang. I-tasser: a unified platform for automated protein structure and function prediction. *Nature protocols*, 5(4):725, 2010.
- [100] Dong Xu and Yang Zhang. Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization. *Biophysical journal*, 101(10):2525–2534, 2011.
- [101] Francisco Carrascoza, Snezana Zaric, and Radu Silaghi-Dumitrescu. Computational study of protein secondary structure elements: Ramachandran plots revisited. *Journal of Molecular Graphics and Modelling*, 50:125–133, 2014.
- [102] Markus Wiederstein and Manfred J Sippl. Prosa-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic acids research*, 35(suppl_2):W407–W410, 2007.
- [103] Alexandre Borrel, Leslie Regad, Henri Xhaard, Michel Petitjean, and Anne-Claude Camproux. Pockdrug: A model for predicting pocket druggability

- that overcomes pocket estimation uncertainties. *Journal of chemical information and modeling*, 55(4):882–895, 2015.
- [104] Dina Duhovny, Ruth Nussinov, and Haim J Wolfson. Efficient unbound docking of rigid molecules. In *International workshop on algorithms in bioinformatics*, pages 185–200. Springer, 2002.
- [105] Dina Schneidman-Duhovny, Yuval Inbar, Ruth Nussinov, and Haim J Wolfson. Patchdock and symmdock: servers for rigid and symmetric docking. *Nucleic acids research*, 33(suppl_2):W363–W367, 2005.
- [106] Roman A Laskowski, Jagoda Jabłońska, Lukáš Pravda, Radka Svobodová Vařeková, and Janet M Thornton. Pdbsum: Structural summaries of pdb entries. *Protein Science*, 27(1):129–134, 2018.
- [107] R Parthasarathi, V Subramanian, and PK Chattaraj. Effect of electric field on the global and local reactivity indices. *Chemical physics letters*, 382(1-2):48–56, 2003.
- [108] Eram Khan, Poonam Tandon, Rakesh Maurya, Padam Kumar, et al. A theoretical study on molecular structure, chemical reactivity and molecular docking studies on dalbergin and methyl dalbergin. *Journal of Molecular Structure*, 1183:100–106, 2019.
- [109] Haicheng Li, Roberto Tejero, Daniel Monleon, Donna Bassolino-Klimas, Cory Abate-Shen, Robert E Bruccoleri, and Gaetano T Montelione. Homology modeling using simulated annealing of restrained molecular dynamics and conformational search calculations with congen: Application in predicting the three-dimensional structure of murine homeodomain msx-1. *Protein Science*, 6(5):956–970, 1997.
- [110] Ge-Fei Hao, Wei-Fang Xu, Sheng-Gang Yang, and Guang-Fu Yang. Multiple simulated annealing-molecular dynamics (msa-md) for conformational space search of peptide and miniprotein. *Scientific reports*, 5:15568, 2015.
- [111] Michael Nilges, Angela M Gronenborn, Axel T Brünger, and G Marius Clore. Determination of three-dimensional structures of proteins by simulated

- annealing with interproton distance restraints. application to crambin, potato carboxypeptidase inhibitor and barley serine proteinase inhibitor 2. *Protein Engineering, Design and Selection*, 2(1):27–38, 1988.
- [112] SH [Ubar] ICHI NOSÉ. A molecular dynamics method for simulations in the canonical ensemble. *Molecular Physics*, 100(1):191–198, 2002.
- [113] Tang-Qing Yu, José Alejandre, Roberto López-Rendón, Glenn J Martyna, and Mark E Tuckerman. Measure-preserving integrators for molecular dynamics in the isothermal–isobaric ensemble derived from the liouville operator. *Chemical Physics*, 370(1-3):294–305, 2010.
- [114] Chao Zhang, George Vasmatazis, James L Cornette, and Charles DeLisi. Determination of atomic desolvation energies from the structures of crystallized proteins. *Journal of molecular biology*, 267(3):707–726, 1997.
- [115] Md Junaid, Md Jahangir Alam, Md Kamal Hossain, Mohammad A Halim, and M Obayed Ullah. Molecular docking and dynamics of nickel-schiff base complexes for inhibiting β -lactamase of mycobacterium tuberculosis. *In silico pharmacology*, 6(1):6, 2018.
- [116] Faizul Azam, Nada Hussin Alabdullah, Hadeel Mohammed Ehmedat, Abdullah Ramadan Abulifa, Ismail Taban, and Sreedevi Upadhyayula. Nsaids as potential treatment option for preventing amyloid β toxicity in alzheimer’s disease: an investigation by docking, molecular dynamics, and dft studies. *Journal of Biomolecular Structure and Dynamics*, 36(8):2099–2117, 2018.
- [117] Marco Wiltgen. Algorithms for structure comparison and analysis: Homology modelling of proteins. *Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics*, page 38, 2018.
- [118] B Lakshmi, C Ramakrishnan, G Archunan, R Sowdhamini, and N Srinivasan. Investigations of ramachandran disallowed conformations in protein domain families. *International journal of biological macromolecules*, 63:119–125, 2014.
- [119] Bruno Borguesan, Mariel Barbachan e Silva, Bruno Grisci, Mario Inostroza-Ponta, and Márcio Dorn. Apl: An angle probability list to improve

- knowledge-based metaheuristics for the three-dimensional protein structure prediction. *Computational biology and chemistry*, 59:142–157, 2015.
- [120] A Gopalan, G Deka, M Prabhavathi, HS Savithri, MRN Murthy, and A Raja. Structural and biophysical characterization of rv3716c, a hypothetical protein from mycobacterium tuberculosis. *Biochemical and biophysical research communications*, 495(1):982–987, 2018.
- [121] S Vijayakumar, P Manogar, S Prabhu, M Pugazhenthii, and PK Praseetha. A pharmacoinformatic approach on cannabinoid receptor 2 (cb2) and different small molecules: Homology modelling, molecular docking, md simulations, drug designing and adme analysis. *Computational biology and chemistry*, 78:95–107, 2019.
- [122] Boris A Reva, Alexei V Finkelstein, and Jeffrey Skolnick. What is the probability of a chance prediction of a protein structure with an rmsd of 6 Å? *Folding and Design*, 3(2):141–147, 1998.
- [123] Jennifer Ferina and Valerie Daggett. Visualizing protein folding and unfolding. *Journal of molecular biology*, 431(8):1540–1564, 2019.
- [124] Haruna L Barazorda-Ccahuana, Diego Ernesto Valencia, Jorge Alberto Aguilar-Pineda, and Badhin Gómez. Art v 4 protein structure as a representative template for allergen profilins: Homology modeling and molecular dynamics. *ACS Omega*, 3(12):17254–17260, 2018.
- [125] Jim Vrbanac and Richard Slauter. Adme in drug discovery. In *A Comprehensive Guide to Toxicology in Nonclinical Drug Development*, pages 39–67. Elsevier, 2017.
- [126] Kieran L Hudson, Gail J Bartlett, Roger C Diehl, Jon Agirre, Timothy Gallagher, Laura L Kiessling, and Derek N Woolfson. Carbohydrate–aromatic interactions in proteins. *Journal of the American Chemical Society*, 137(48):15152–15160, 2015.
- [127] Kerem Terah. An evaluation of neonicotinoids’ potential to inhibit human

- cholinesterases: Protein–ligand docking and interaction profiling studies. *Journal of Molecular Graphics and Modelling*, 84:54–63, 2018.
- [128] Kaspars Tars, Svetlana Kotelovica, Gerd Lipowsky, Monika Bauer, Roger R Beerli, Martin F Bachmann, and Patrik Maurer. Different binding modes of free and carrier-protein-coupled nicotine in a human monoclonal antibody. *Journal of molecular biology*, 415(1):118–127, 2012.
- [129] Jack Kyte and Russell F Doolittle. A simple method for displaying the hydrophobic character of a protein. *Journal of molecular biology*, 157(1):105–132, 1982.
- [130] Francesca Milletti and Anna Vulpetti. Predicting polypharmacology by binding site similarity: from kinases to the protein universe. *Journal of chemical information and modeling*, 50(8):1418–1431, 2010.
- [131] Fei Guo, Shuai Cheng Li, Lusheng Wang, and Daming Zhu. Protein-protein binding site identification by enumerating the configurations. *BMC bioinformatics*, 13(1):158, 2012.
- [132] Tap Ha-Duong. Protein backbone dynamics simulations using coarse-grained bonded potentials and simplified hydrogen bonds. *Journal of chemical theory and computation*, 6(3):761–773, 2010.
- [133] Xia Sheng, Xiaotong Jiang, Hailiang Zhao, Dongjin Wan, Yongde Liu, Cleopatra Ashley Ngwenya, and Lin Du. Ftir study of hydrogen bonding interaction between fluorinated alcohol and unsaturated esters. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 198:239–247, 2018.
- [134] Kara J Cutrona, Bethany A Kaufman, Dania M Figueroa, and Donald E Elmore. Role of arginine and lysine in the antimicrobial mechanism of histone-derived antimicrobial peptides. *FEBS letters*, 589(24):3915–3920, 2015.

Anexos A

Energías Corregidas para el Cálculo
del Potencial de Ionización (PI) y
Afinidad Electrónica (AE).

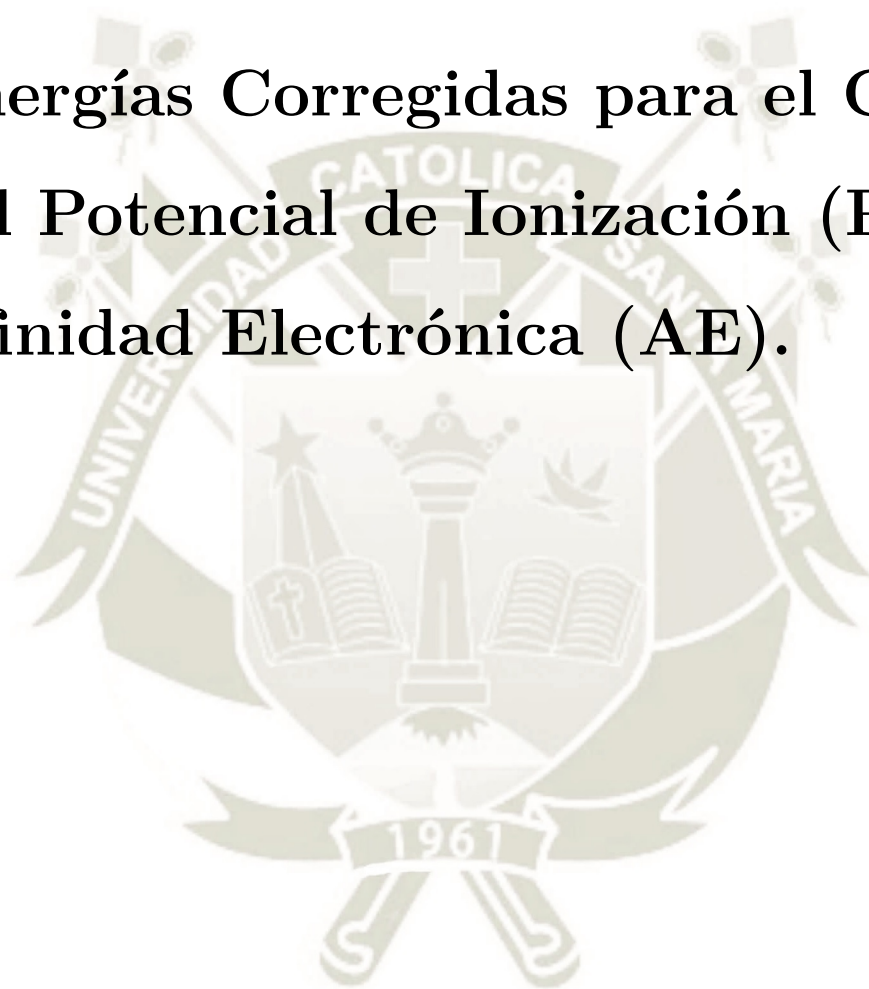


Tabla A.1: Energías corregidas en los estados neutro, catión y anión de los derivados de la nicotina.

Derivado	Estados	Energía	Potencial de Ionización	Afinidad Electrónica
A	A-anión	-1152314.58		
	A-catión	-1152098.39	159.16	57.03
	A-neutro	-1152257.55		
B	B-anión	-1152309.03		
	B-catión	-1152095.79	156.79	56.45
	B-neutro	-1152252.58		
C	C-anión	-1152311.98		
	C-catión	-1152099.47	156.34	56.18
	C-neutro	-1152255.80		
D	D-anión	-1152311.78		
	D-catión	-1152096.61	156.09	56.09
	D-neutro	-1152255.69		
E	E-anión	-1152311.69		
	E-catión	-1152096.14	159.44	56.12
	E-neutro	-1152255.57		

*Todas las energías están reportadas en kcal/mol.

Anexos B

Predicción de Acoplamiento en PatchDock



Tabla B.1: Predicción del acoplamiento de la Erp con el derivado A.

N°	score	Área	ACE
1	4940	575.40	-278.28
2	4898	562.20	-189.76
3	4876	592.10	-238.22
4	4832	592.70	-188.61
5	4818	590.50	-201.80
6	4790	630.70	-357.27
7	4750	606.60	-155.53
8	4734	572.80	-262.29
9	4700	506.10	-270.53
10	4692	581.70	-297.76

Tabla B.2: Predicción del acoplamiento de la Erp con el derivado B.

N°	score	Área	ACE
1	5000	652.40	-335.84
2	4824	583.30	-178.37
3	4664	591.30	-294.11
4	4642	521.40	-287.07
5	4628	574.50	-235.52
6	4624	589.30	-206.92
7	4622	531.50	-233.80
8	4612	507.60	-234.04
9	4610	569.20	-204.74
10	4572	568.80	-255.23

Tabla B.3: Predicción del acoplamiento de la Erp con el derivado C.

N°	score	Área	ACE
1	4984	585.90	-214.13
2	4976	539.50	-221.08
3	4832	606.70	-216.55
4	4770	631.00	-329.56
5	4734	523.60	-214.16
6	4732	593.90	-271.86
7	4702	521.80	-261.32
8	4696	585.90	-348.03
9	4686	549.30	-197.62
10	4666	610.10	-324.29

Tabla B.4: Predicción del acoplamiento de la Erp con el derivado D.

N°	score	Área	ACE
1	5016	604.10	-194.04
2	4814	585.90	-180.31
3	4800	584.30	-213.00
4	4744	638.50	-328.92
5	4690	576.80	-344.78
6	4666	585.60	-203.36
7	4666	519.40	-268.48
8	4652	522.40	-237.76
9	4628	506.10	-201.69
10	4622	542.30	-290.27

Tabla B.5: Predicción del acoplamiento de la Erp con el derivado E.

N°	score	Área	ACE
1	4872	569.10	-248.62
2	4834	595.10	-195.98
3	4738	533.40	-271.71
4	4680	591.80	-193.21
5	4664	538.20	-174.98
6	4632	574.40	-247.14
7	4616	526.20	-263.67
8	4614	635.70	-326.07
9	4610	527.60	-236.81
10	4598	584.40	-235.56

Tabla B.6: Predicción del acoplamiento de la proteína Rv1417 con el derivado A.

N°	score	Área	ACE
1	4928	613.30	-302.89
2	4906	613.40	-261.27
3	4774	602.50	-295.05
4	4508	596.10	-322.54
5	4506	560.50	-284.71
6	4468	536.50	-278.60
7	4456	533.90	-235.11
8	4386	527.30	-212.56
9	4350	516.10	-174.92
10	4318	486.70	-134.81

Tabla B.7: Predicción del acoplamiento de la proteína Rv1417 con el derivado B.

N°	score	Área	ACE
1	4658	521.40	-222.73
2	4574	594.90	-242.28
3	4508	536.60	-241.40
4	4398	541.90	-248.28
5	4364	524.60	-210.99
6	4216	517.70	-286.70
7	4208	499.60	-213.82
8	4208	556.60	-287.96
9	4194	491.80	-237.50
10	4108	552.40	-301.46

Tabla B.8: Predicción del acoplamiento de la proteína Rv1417 con el derivado C.

N°	score	Área	ACE
1	5066	618.50	-287.22
2	4894	592.90	-233.55
3	4742	531.60	-203.94
4	4640	594.60	-252.95
5	4376	547.20	-312.02
6	4344	541.20	-274.46
7	4288	472.40	-140.64
8	4272	562.20	-255.70
9	4260	474.70	-194.06
10	4210	506.60	-246.33

Tabla B.9: Predicción del acoplamiento de la proteína Rv1417 con el derivado D.

N°	score	Área	ACE
1	4912	634.00	-321.15
2	4908	552.80	-240.49
3	4650	564.60	-252.70
4	4608	535.90	-279.62
5	4552	566.60	-271.22
6	4420	506.80	-217.07
7	4302	543.50	-279.74
8	4274	559.00	-180.03
9	4152	524.60	-262.07
10	4104	498.10	-267.32

Tabla B.10: Predicción del acoplamiento de la proteína Rv1417 con el derivado E.

N°	score	Área	ACE
1	4856	616.60	-325.44
2	4612	584.80	-282.07
3	4454	582.50	-228.03
4	4396	510.00	-184.10
5	4324	525.70	-282.52
6	4272	571.60	-219.86
7	4238	476.70	-204.99
8	4216	493.10	-229.39
9	4182	589.80	-285.61
10	4174	506.80	-247.84

Tabla B.11: Predicción del acoplamiento de la proteína Rv2617c con el derivado A.

N°	score	Área	ACE
1	4332	516.60	-195.51
2	4294	535.60	-199.60
3	4186	496.20	-223.85
4	4056	514.00	-266.68
5	4012	490.60	-279.81
6	3990	463.10	-236.47
7	3962	499.70	-205.65
8	3940	465.90	-159.53
9	3918	432.60	-159.89
10	3906	434.90	-188.49

Tabla B.12: Predicción del acoplamiento de la proteína Rv2617c con el derivado B.

N°	score	Área	ACE
1	4296	484.70	-220.37
2	4170	468.80	-232.53
3	4068	493.00	-272.97
4	4028	500.00	-189.37
5	4026	477.30	-207.19
6	3912	435.80	-239.21
7	3900	529.40	-216.57
8	3852	536.60	-295.94
9	3780	484.70	-241.27
10	3778	417.40	-191.20

Tabla B.13: Predicción del acoplamiento de la proteína Rv2617c con el derivado C.

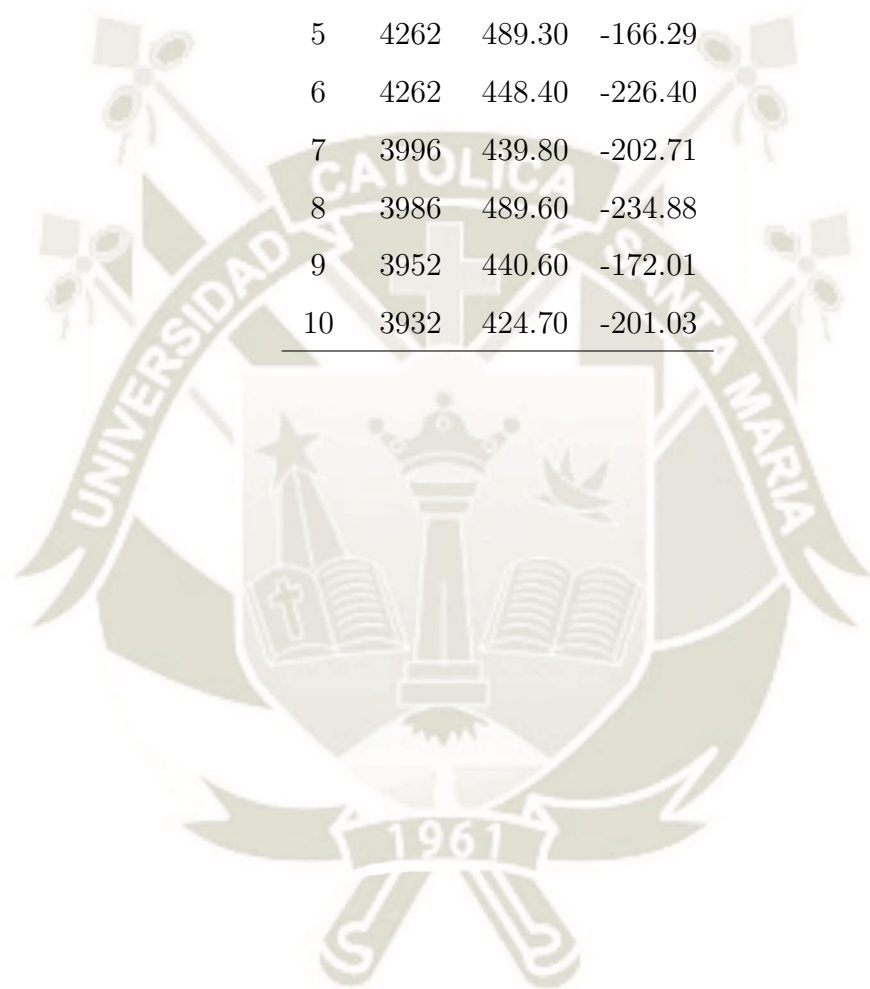
N°	score	Área	ACE
1	4394	516.20	-273.89
2	4194	530.60	-228.62
3	4098	498.30	-208.75
4	4080	447.70	-215.61
5	4008	498.40	-274.65
6	3998	438.50	-207.73
7	3994	465.80	-243.02
8	3982	536.10	-267.19
9	3940	507.90	-141.25
10	3922	463.80	-203.04

Tabla B.14: Predicción del acoplamiento de la proteína Rv2617c con el derivado D.

N°	score	Área	ACE
1	4306	514.90	-194.46
2	4272	529.40	-291.30
3	4090	466.80	-173.70
4	4048	461.80	-254.63
5	4026	465.80	-230.59
6	3972	475.20	-238.82
7	3964	476.10	-174.25
8	3948	428.80	-183.85
9	3938	503.70	-264.49
10	3872	566.20	-281.28

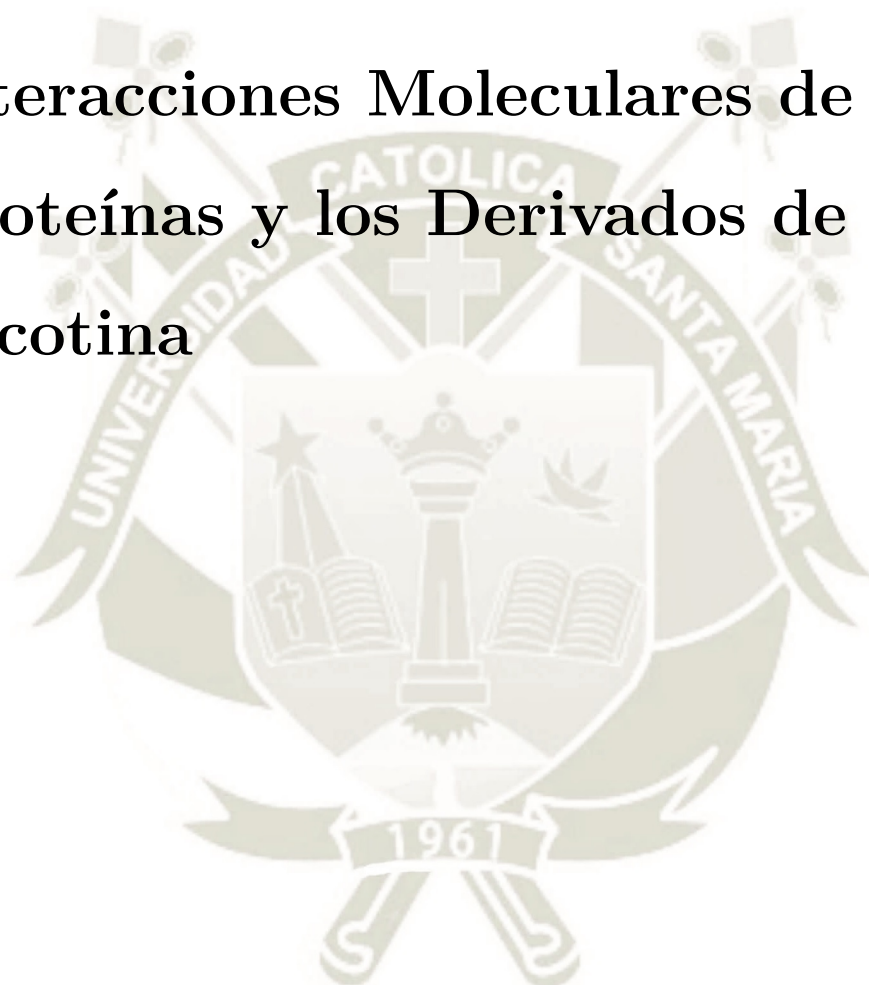
Tabla B.15: Predicción del acoplamiento de la proteína Rv2617c con el derivado E.

N°	score	Área	ACE
1	4400	489.60	-252.61
2	4326	525.70	-209.75
3	4318	501.00	-248.49
4	4274	522.40	-222.76
5	4262	489.30	-166.29
6	4262	448.40	-226.40
7	3996	439.80	-202.71
8	3986	489.60	-234.88
9	3952	440.60	-172.01
10	3932	424.70	-201.03



Anexos C

Interacciones Moleculares de las Proteínas y los Derivados de la Nicotina



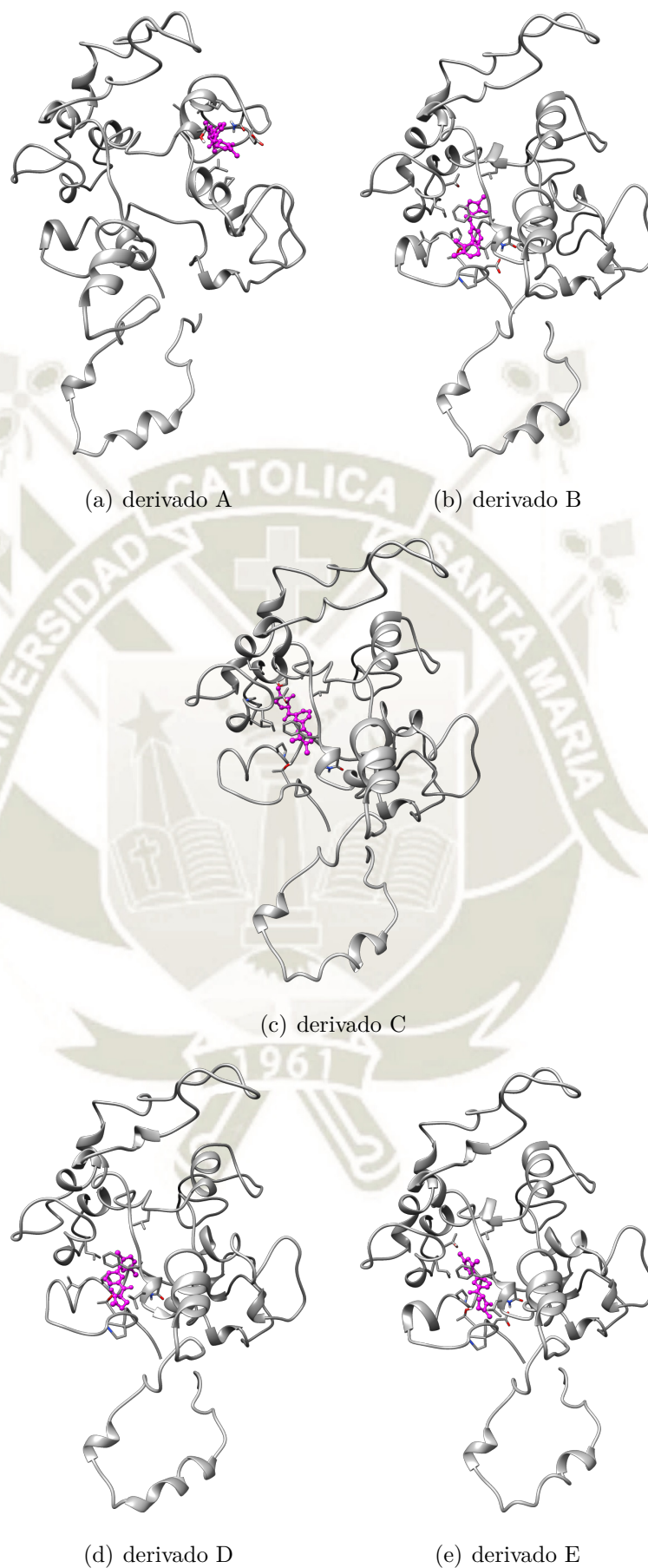


Figura C.1: Interacción de la proteína Erp con cada derivado de la nicotina.

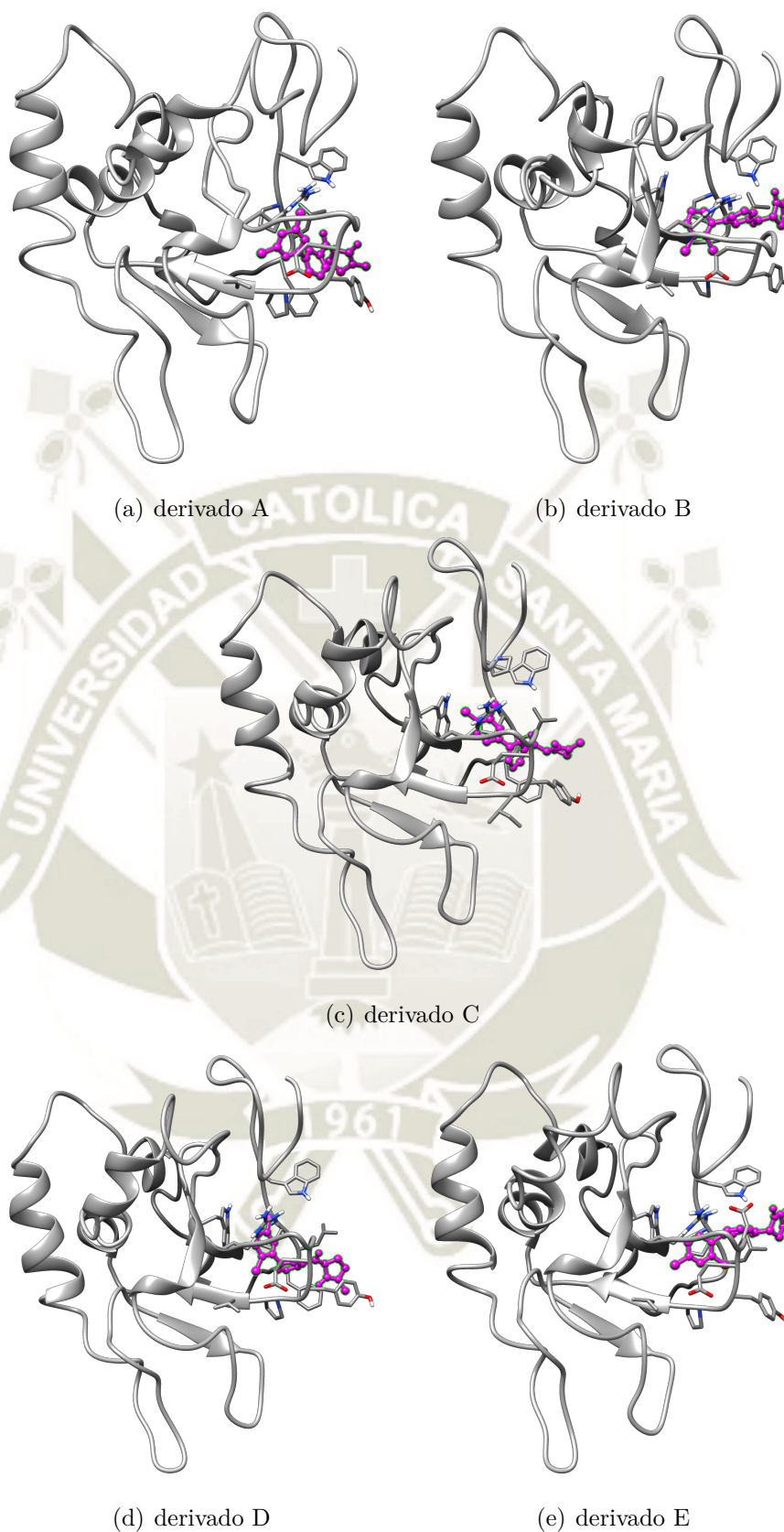


Figura C.2: Interacción de la proteína Rv1417 con cada derivado de la nicotina.

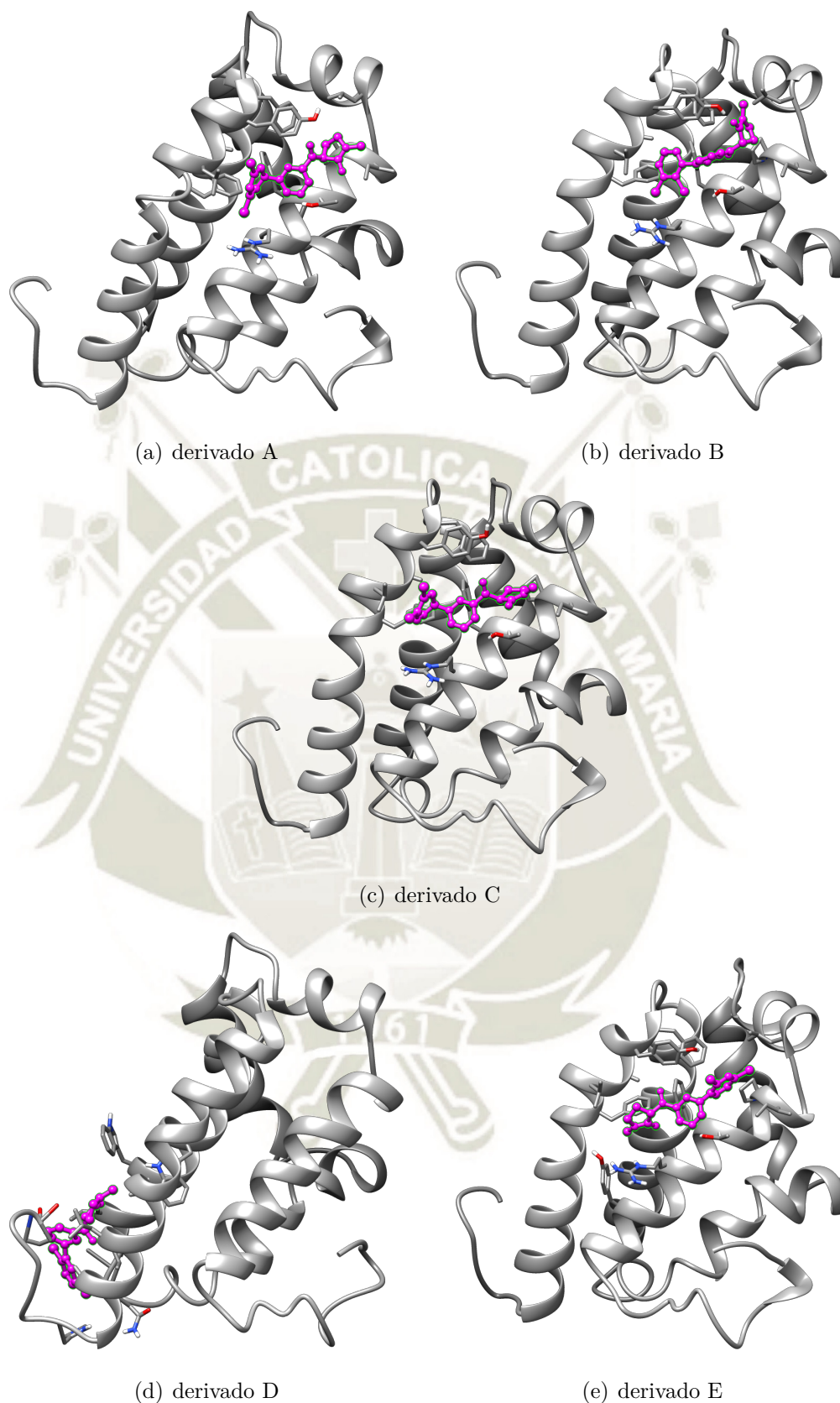


Figura C.3: Interacción de la proteína Rv2617c con cada derivado de la nicotina.