

# Universidad Católica de Santa María

**Facultad de Ciencias Farmacéuticas, Bioquímicas y  
Biotecnológicas**

**Escuela Profesional de Ingeniería Biotecnológica**



**“APROXIMACIÓN TEÓRICA DEL EFECTO DE  
FLAVONOIDES PROCEDENTES DE *Opuntia ficus-indica*  
(NOPAL) SOBRE LA PROFILINA HUMANA II (PFN2) EN SU  
INTERACCIÓN CON LA ENZIMA HDAC1 Y LA VÍA DE  
SEÑALIZACIÓN TGF-B/SMAD”**

Tesis presentada por el Bachiller:

**Paz Caya, Diego Jeampaul**

para optar el Título Profesional de:

**Ingeniero Biotecnólogo**

Asesor:

Dr. Gómez Valdez, Badhin

AREQUIPA – PERÚ

2018

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

Expediente N°. 20170000031650

N° Trámite en Fac. 354-2017  
Fecha Recep. Fac. 17-07-2017

FORMATO UNICO PARA TRAMITACIÓN DE TÍTULO PROFESIONAL

DE: PAZ CAYA, Diego Jeampaul

PARA OPTAR EL TÍTULO PROFESIONAL DE INGENIERO BIOTECNOLOGO

**"APROXIMACION TEORICA DEL EFECTO DE FLAVONOIDES PROCEDENTES DE *Opuntia ficus-indica* (Nopal) SOBRE LA PROFILINA HUMANA II (PFN2) EN SU INTERACCION CON LA ENZIMA HDAC1 Y LA VIA DE SEÑALIZACION TGF- $\beta$ /Smad"**

DICTAMINADORES: 1) 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

FIRMAS:

(Devolver antes de 8 días hábiles) FECHA 10/08/2017

ASESOR: Dr. Badhin Gómez Valdez

**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 el 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

Atentamente

FIRMA



FECHA 22-10-18

DICTAMINADORES BORRADOR DE TESIS:

- |                                       |                                    |
|---------------------------------------|------------------------------------|
| 1) Dr. Jaime Cárdenas García          | 3) Mgter. Jaime Barreda del Carpio |
| 2) Blgo. Carlos Eitel Iván Paz Aliaga |                                    |

DICTAMEN FINAL:

Señor Decano de la Facultad de Ciencias Farmacéuticas, Bioquímicas y Biotecnológicas, atendiendo a su designación como Dictaminadores del presente Borrador de Tesis y luego de hechas las observaciones y correcciones pertinentes, cumpliendo con las exigencias mínimas establecidas para un trabajo de investigación de Tesis profesional, es que consideramos APTO para continuar con los trámites estipulados en el Reglamento de Grados y Títulos de la Facultad.

Atentamente

FIRMA



(Devolver antes de 15 días hábiles) FECHA 22-11-18

JURADOS: PRESIDENTE DR. JAIME CÁRDENAS GARCÍA  
VOCAL BLGO. CARLOS EITEL IVÁN PAZ ALIAGA  
SECRETARIO MAG. JAIME BARREDA DEL CARPIO

FECHA 3/12/18

HORA 19.00

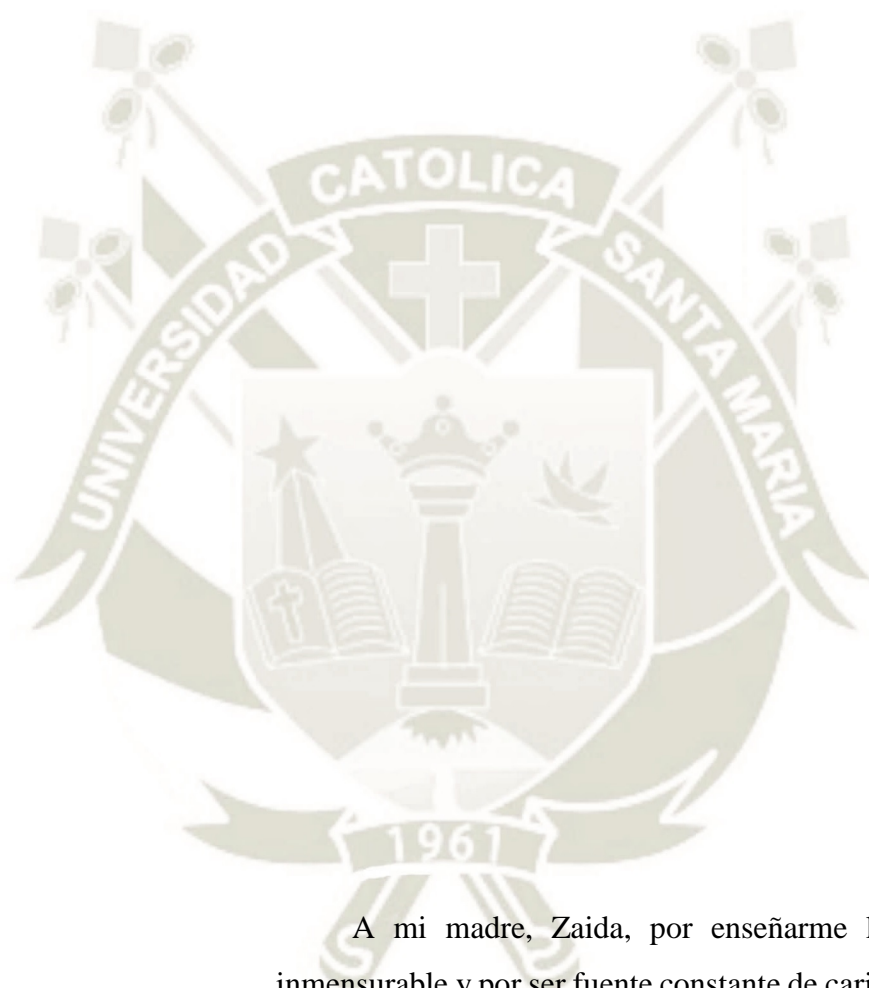
LOCAL C-402

FIRMA DEL DECANO



FECHA 28/11/18

# DEDICATORIA



A mi madre, Zaida, por enseñarme la fortaleza más inmensurable y por ser fuente constante de cariño y motivación; a mi padre, Ricardo, por inculcar en mí raciocinio, por tu guía, soporte y exigencia.

A mi familia, siempre presente y con un apoyo incondicional; a mi compañera, Suzel, por ser la mejor compañía durante esta senda, *conspatior*.

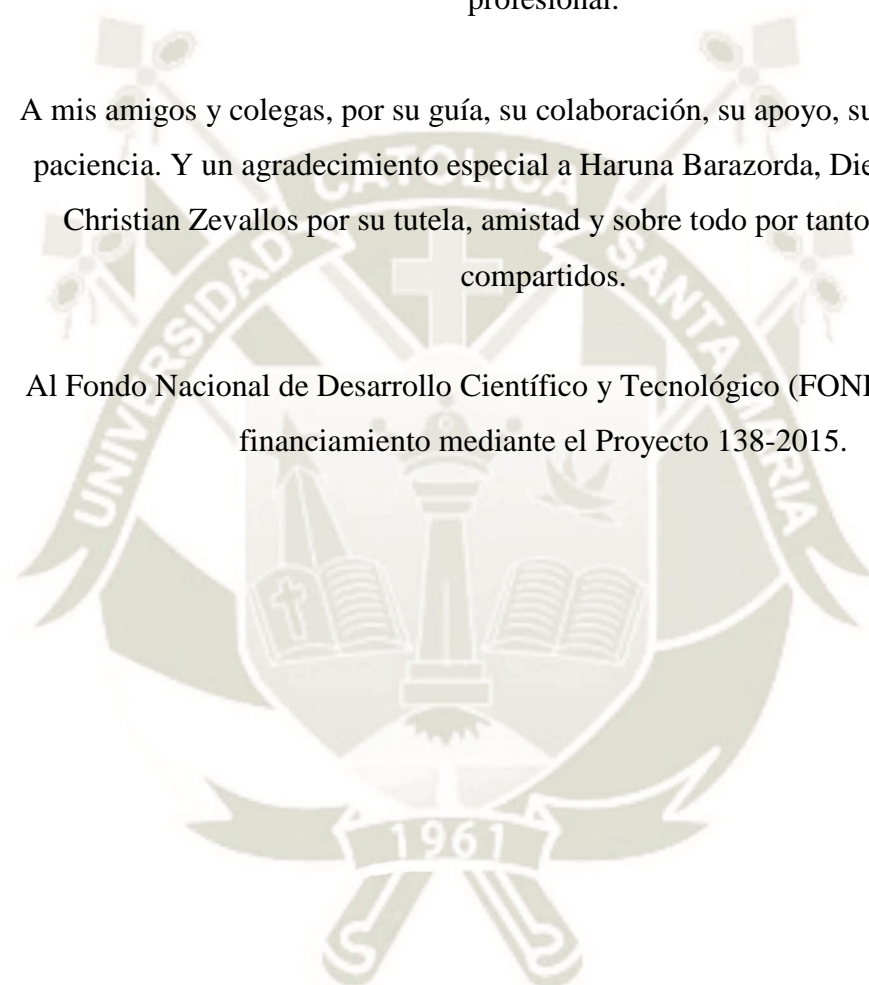


# AGRADECIMIENTOS

Al Dr. Badhin Gómez Valdez, por la oportunidad de ser parte de un increíble grupo de investigación (CIIM), por sus enseñanzas y por su contribución a mi desarrollo profesional.

A mis amigos y colegas, por su guía, su colaboración, su apoyo, sus consejos y su paciencia. Y un agradecimiento especial a Haruna Barazorda, Diego Valencia y Christian Zevallos por su tutela, amistad y sobre todo por tantos momentos compartidos.

Al Fondo Nacional de Desarrollo Científico y Tecnológico (FONDECYT) por el financiamiento mediante el Proyecto 138-2015.





# INTRODUCCIÓN

Los pulmones son órganos vitales para nuestra existencia, en ellos se lleva a cabo el intercambio gaseoso a nivel capilar. Son muchas las afecciones que estos órganos pueden padecer; el cáncer de pulmón es una de estas enfermedades, el cual es capaz de impedir el correcto funcionamiento de estos órganos afectando el tracto respiratorio, el tejido pulmonar y los vasos sanguíneos presentes según su progresión; siendo descrita como una enfermedad de crecimiento excesivo o irregular de las células, lo cual promueve la diseminación e invasión a otros órganos.

Son muchos los factores que promueven el desarrollo de esta enfermedad, uno de los principales es el tabaquismo, asimismo, son varias las rutas metabólicas implicadas en el desarrollo y progresión del cáncer, una de estas rutas vinculada con la progresión de cáncer de pulmón es la ruta de factores de crecimiento transformantes  $\beta$ , la cual presenta una regulación epigenética de la enzima histona desacetilasa a través de su reclutamiento en el citoplasma mediado por la proteína profilina humana II que presenta niveles de sobreexpresión irregulares en estas células neoplásicas; este reclutamiento conlleva a la interacción de ambas proteínas impidiendo el correcto funcionamiento de una de estas a nivel nuclear como regulador de la expresión génica. *Opuntia ficus-indica* es una planta cactácea tradicionalmente usada con fines medicinales que presenta propiedades anticancerígenas y antioxidantes reportadas de manera empírica pero estrechamente relacionadas a su alto contenido de flavonoides que se encuentran en las raíces, pencas, frutos y flores de esta planta; es por ello que se plantea el estudio del efecto de estos compuestos hacia la profilina humana II con gran interés, mediante el uso de herramientas de simulación computacional, una vía de investigación que nos da un acercamiento teórico de los sucesos estudiados así como su viabilidad hacia un desarrollo experimental. Estas herramientas computacionales envuelven grandes campos como la bioinformática y la química computacional y nos permiten evaluar las interacciones de estos compuestos.

## RESUMEN

En el presente estudio se analizó la interacción entre la profilina humana II (PFN2), una proteína de unión a actina, con 9 compuestos flavonoides reportados en flores y frutos de *Opuntia ficus-indica*, y a su vez, el efecto que genera la formación de estos sistemas en la interacción entre PFN2 y la enzima histona desacetilasa 1 (HDAC1) para una regulación epigenética de la ruta del factor de crecimiento transformante beta (TGF- $\beta$ ) implicado en cáncer de pulmón. En primer término, se llevó a cabo la obtención y optimización de las estructuras tridimensionales tanto de las proteínas como de los flavonoides mediante simulaciones de dinámica molecular y mecánica cuántica respectivamente, logrando obtener estructuras en zonas de equilibrio con el uso de un colectivo canónico (NVT) en el software Gromacs y estructuras optimizadas geométricamente en caso de los flavonoides con Gaussian 09; para luego proceder a los acoplamientos moleculares y análisis conformacionales para la validación de estructuras. El acoplamiento entre PFN2 modificada por acción de cada flavonoide y PFN2 nativa con HDAC1 nos permitió dilucidar aquellos compuestos que generan un cambio significativo sobre la PFN2 que altera la interacción de esta con HDAC1, siendo el compuesto quercetina 3-O-glucósido el que presentó resultados más significativos.

### Palabras Clave:

Profilina humana, histona desacetilasa, flavonoides, *in silico*.

## ABSTRACT

In the present study, an analysis was carried on to study the interaction between human profilin II (PFN2), an actin-binding protein, with 9 flavonoid compounds reported in flowers and fruits from *Opuntia ficus-indica*, as well as the effect generated by the formation of these complex in the interaction between PFN2 and the enzyme histone deacetylase 1 (HDAC1) for an epigenetic regulation of the transforming growth factor beta (TGF- $\beta$ ) pathway involved in lung cancer. First of all, three-dimensional structures of both proteins and flavonoids were obtained and then optimized by simulations of molecular dynamics and quantum mechanics, respectively, achieving structures in equilibrium zones by using a canonical assembly (NVT) in the Gromacs software and geometrically optimized structures in the case of flavonoids with Gaussian 09; then proceed to perform molecular couplings and conformational analysis for the structures' validation. The coupling between mutated PFN2 by the action of each flavonoid and native PFN2 with HDAC1 allowed us to elucidate those compounds that generate a significant change on PFN2 in order to alter the interaction of this with HDAC1, being quercetin 3-O-glucoside the one that showed up the most significant results.

### Keywords:

Human profilin, histone deacetylase, flavonoids, *in silico*.



# ÍNDICE GENERAL

DEDICATORIA.....	I
AGRADECIMIENTOS.....	II
INTRODUCCIÓN.....	III
RESUMEN.....	IV
ABSTRACT.....	V
ÍNDICE GENERAL.....	VI
ÍNDICE DE FIGURAS.....	IX
ÍNDICE DE TABLAS.....	XIII
GLOSARIO.....	XIV
1. PLANTEAMIENTO:.....	1
1.1. Objetivos:.....	1
1.1.1. Objetivo General.....	1
1.1.2. Objetivos Específicos.....	1
1.2. Hipótesis:.....	2
2. MARCO TEÓRICO:.....	3
2.1. El Sistema Respiratorio: Los Pulmones:.....	3
2.2. El Cáncer de Pulmón:.....	4
2.2.1. Etiología:.....	5
2.2.2. Clasificación:.....	5
2.2.3. Estadios de NSCLC:.....	6
2.2.4. Síntomas:.....	7
2.2.5. Diagnóstico:.....	8
2.2.6. Tratamiento:.....	8
2.3. Regulación Epigenética Mediada por la Profilina Humana II en Cáncer de Pulmón:.....	9
2.3.1. Vía de Señalización TGF- $\beta$ /Smads:.....	10
2.3.1.1. Factor de Crecimiento Transformante - $\beta$ :.....	11
2.3.1.2. Receptores de Membrana:.....	12
2.3.1.3. Proteínas Smads:.....	12
2.3.2. Enzima Histona Desacetilasa:.....	14
2.3.3. Las Profilinas Humanas:.....	17

2.3.3.1. Rol de las Profilinas Humanas:	18
2.3.3.2. Profilina Humana II (PFN2):	20
2.4. <i>Opuntia ficus-indica</i> :	21
2.4.1. Fisiología:	21
2.4.2. Hábitat:	22
2.4.3. Aplicaciones:	23
2.4.4. Flores y frutos de <i>Opuntia ficus-indica</i> :	23
2.5. Compuestos Flavonoides:	25
2.5.1. Estructura y Clasificación:	25
2.5.2. Propiedades Antioxidantes:	27
2.5.3. Propiedades Anticancerígenas:	27
2.6. Química Computacional y Bioinformática:	28
2.6.1. Mecánica Cuántica:	28
2.6.2. Mecánica Molecular o Mecánica Clásica:	29
2.6.3. Modelamiento de Proteínas:	30
2.6.4. Campos de Fuerza para Modelamiento Molecular:	31
2.6.5. Minimización de Energía:	32
2.6.6. Dinámica Molecular:	33
2.6.8. Acoplamiento Molecular:	35
2.6.9. Validación de Estructuras:	36
3. MATERIALES Y MÉTODOS:	38
3.1. Equipos y Software:	38
3.2. Metodología:	42
3.2.1. Obtención de las Estructuras Tridimensionales de las Proteínas de Interés:	42
3.2.2. Obtención de las Estructuras Tridimensionales y Cargas de Hirshfeld de los Flavonoides de Nopal:	42
3.2.3. Cinco Reglas de Lipinski para Compuestos Flavonoides:	45
3.2.4. Declaración de Nuevos Parámetros en Campos de Fuerza OPLSAA:	45
3.2.4.1. <i>Profilina Humana II</i> :	45
3.2.4.2. <i>Histona Desacetilasa I</i> :	45
3.2.4.3. <i>Compuestos Flavonoides</i> :	46
3.2.5. Minimización de Energías y Dinámica Molecular de las Proteínas:	46
3.2.6. Validación de la Estructura de las Proteínas:	48
3.2.7. Predicción de Bolsillos de Drogabilidad:	48
3.2.8. Docking o Acoplamiento Proteína - Ligando:	49

3.2.9. Dinámica de Complejos Proteína-Ligando: .....	51
3.2.10. Cálculo de Energía Libre con Colectivo NPT: .....	51
3.2.11. Docking o Acoplamiento Proteína - Proteína: .....	52
3.2.11.1. <i>Histona Desacetilasa 1 – Profilina Humana II nativa:</i> .....	52
3.2.11.2. <i>Histona Desacetilasa 1 – Profilina Humana II modificada:</i> .....	52
3.2.12. Refinamiento de los Complejos Proteína - Proteína: .....	53
4. RESULTADOS Y DISCUSIÓN: .....	54
4.1. Obtención de la Estructura Tridimensional de Profilina Humana II (PFN2): ....	54
4.2. Obtención de la Estructura Tridimensional Optimizada de Compuestos Flavonoides: .....	57
4.3. Obtención de la Estructura Tridimensional de Histona Desacetilasa 1: .....	62
4.4. Interacción Proteína - Ligando: .....	68
4.4.1. Acoplamiento PFN2 y quercetina: .....	70
4.4.2. Acoplamiento PFN2 y quercetina 3-O-glucósido: .....	72
4.4.3. Acoplamiento PFN2 y quercetina 3-metil éter: .....	75
4.4.4. Acoplamiento PFN2 y dihidroquercetina: .....	76
4.4.5. Acoplamiento PFN2 e isorhamnetin: .....	79
4.4.6. Acoplamiento PFN2 e isorhamnetin 3-O-galactósido: .....	81
4.4.7. Acoplamiento PFN2 e isorhamnetin 3-O-glucósido: .....	83
4.4.8. Acoplamiento PFN2 y kaempferol: .....	85
4.4.9. Acoplamiento PFN2 y kaempferol 3-metil éter: .....	87
4.5. Cálculo de la Energía Libre de Acoplamiento: .....	89
4.6. Interacción Proteína - Proteína: .....	90
4.6.1. Acoplamiento Molecular: .....	90
4.6.2. Refinamiento de la Interacción del Acoplamiento: .....	94
CONCLUSIONES: .....	96
RECOMENDACIONES: .....	97
REFERENCIAS BIBLIOGRÁFICAS: .....	98



# ÍNDICE DE FIGURAS

<b>Figura 2.1:</b> Gráfica circular de datos los perfiles de mortalidad e incidencia de varios tipos de cáncer que afectan a hombres y mujeres según la OMS (2014) .....	5
<b>Figura 2.2:</b> Cáncer de pulmón en su último estadio y su diseminación a través del cuerpo, Instituto Nacional del Cáncer, U.S. Department of Health and Human Services .....	7
<b>Figura 2.3:</b> Vía de señalización de TGF- $\beta$ , precursores inactivos, receptores de membrana y segundos mensajeros Smads. Nota. Fuente: Meulmeester, E. and P. Ten Dijke (2011). "The dynamic roles of TGF- $\beta$ in cancer." <i>The Journal of pathology</i> 223(2): 206-219. ....	11
<b>Figura 2.4:</b> Funcionamiento regular del proceso de desacetilación (HDAC) y acetilación (HAT). Nota. Adaptado. Fuente: MERCK Presentation .....	15
<b>Figura 2.5:</b> Esquema de los diferentes tipos de HDAC, su clasificación, el largo de las proteínas en base a su número de aminoácidos y en azul su dominio catalítico. Nota. Fuente: De Ruijter, A. J., et al. (2003). "Histone deacetylases (HDACs): characterization of the classical HDAC family." <i>Biochemical Journal</i> 370(Pt 3): 737. ....	16
<b>Figura 2.6:</b> Superposición de estructuras tridimensionales de Profilina Humana I (rojo) y Profilina Humana II (azul), resaltando elementos de su estructura secundaria como son las $\alpha$ -hélices y las láminas $\beta$ . Nota. Fuente: Nodelman, I. M., et al. (1999). "X-ray structure determination of human profilin II: A comparative structural analysis of human profilins." <i>Journal of molecular biology</i> 294(5): 1271-1285. ....	18
<b>Figura 2.7:</b> Rol de las profilinas en la polimerización de actinas. Nota. Fuente: Witke, W. (2004). "The role of profilin complexes in cell motility and other cellular processes." <i>Trends in cell biology</i> 14(8): 461-469. ....	19
<b>Figura 2.8:</b> Opuntia ficus-indica, vista de sus cladodios, espinas, sus frutos y flores.....	22
<b>Figura 2.9:</b> Estructura básica de un compuesto flavonoide o flavano.....	26
<b>Figura 2.10:</b> Estructura secundaria de las proteínas, $\alpha$ -hélices y láminas $\beta$ .....	31
<b>Figura 2.11:</b> Tipos de interacciones descritas en los campos de fuerza, tanto enlazantes como no-enlazantes. Nota. Fuente: Kukol, A. (2008) <i>Molecular Modeling of Proteins</i> , Hatfield, UK .....	32
<b>Figura 2.12:</b> Desacoplamiento de interacciones no enlazantes para el cálculo de la variación de la energía libre de Gibbs haciendo uso del parámetro de desacoplamiento $\lambda$ .....	35
<b>Figura 4.1:</b> Estructura tridimensional del complejo homotetramérico de Profilina Humana II descargada de la base de datos PDB con un ID: 1D1J .....	54
<b>Figura 4.2:</b> Secuencia de aminoácidos de Profilina Humana II en formato FASTA.....	55
<b>Figura 4.3:</b> Estructuras tridimensionales de PFN2 nativa resaltando su estructura secundaria de azul las $\alpha$ -hélices y en verde las láminas $\beta$ . (a) Pre-dinámica (b) Post-dinámica .....	55
<b>Figura 4.4:</b> Ploteo de RMSD de la proteína Profilina Humana II post-dinámica, que alcanza la estabilidad a partir de los 30ns .....	56
<b>Figura 4.5:</b> Diagrama de Ramachandran de la proteína Profilina Humana II, la cual presenta al residuo LYS28 fuera de la zona permitida .....	56
<b>Figura 4.6:</b> Compuesto flavonoide Quercetina y su representación 2D (izquierda) y 3D (derecha).....	57
<b>Figura 4.7:</b> Compuesto flavonoide (+) – Dihidroquercetina y su representación 2D (izquierda) y 3D (derecha).....	57

<b>Figura 4.8:</b> Compuesto flavonoide Quercetina 3-O-rutinósido y su representación 2D (izquierda) y 3D (derecha).....	58
<b>Figura 4.9:</b> Compuesto flavonoide Quercetina 3-O-glucósido y su representación 2D (izquierda) y 3D (derecha).....	58
<b>Figura 4.10:</b> Compuesto flavonoide Quercetina 3-metil éter y su representación 2D (izquierda) y 3D (derecha).....	58
<b>Figura 4.11:</b> Compuesto flavonoide Isorhamnetin y su representación 2D (izquierda) y 3D (derecha).....	59
<b>Figura 4.12:</b> Compuesto flavonoide Isorhamnetin 3-O-galactósido y su representación 2D (izquierda) y 3D (derecha).....	59
<b>Figura 4.13:</b> Compuesto flavonoide Isorhamnetin 3-O-glucósido y su representación 2D (izquierda) y 3D (derecha).....	59
<b>Figura 4.14:</b> Compuesto flavonoide Isorhamnetin 3-O-robinobiósido y su representación 2D (izquierda) y 3D (derecha).....	60
<b>Figura 4.15:</b> Compuesto Kaempferol y su representación 2D (izquierda) y 3D (derecha) ...	60
<b>Figura 4.16:</b> Compuesto Kaempferol 3-metil éter y su representación 2D (izquierda) y 3D (derecha).....	60
<b>Figura 4.17:</b> Estructura tridimensional del complejo heterodimérico de Histona Desacetilasa I con la proteína Asociada a Metástasis (MTA) descargada de la base de datos PDB con un ID: 4BKX.....	63
<b>Figura 4.18:</b> Secuencia de aminoácidos de Histona Desacetilasa 1 en formato FASTA, resaltando la secuencia catalítica que codifica la proteína tridimensionalmente .....	63
<b>Figura 4.19:</b> Estructuras tridimensionales de HDAC1 nativa resaltando su estructura secundaria y el centro catalítico de Zinc. (a) Pre-dinámica (b) Post-dinámica (c) Centro catalítico.....	64
<b>Figura 4.20:</b> Ploteo de RMSD de la enzima Histona Desacetilasa 1 post-dinámica con una zona de estabilidad a partir de los 15ns .....	65
<b>Figura 4.21:</b> Diagrama de Ramachandran de HDAC1, donde se muestra a 10 residuos fuera de la zona permitida .....	66
<b>Figura 4.22:</b> Ploteo de RMSF de la enzima HDAC1 post-dinámica que presenta un pico entre los residuos SER81 y ARG86.....	67
<b>Figura 4.23:</b> Ploteo del radio de giro de la enzima HDAC1 post-dinámica que alcanza estabilidad a partir de los 15ns con un radio de 2.04 aproximadamente .....	67
<b>Figura 4.24:</b> Bolsillos o pockets de la proteína Profilina Humana II calculada en el servidor PockDrug.....	68
<b>Figura 4.25:</b> Acoplamiento de PFN2 con Quercetina en el bolsillo 6, y los residuos que interactúan con este compuesto .....	71
<b>Figura 4.26:</b> Ploteo de RMSD tras una dinámica molecular de la proteína PFN2 acoplada con QCT.....	71
<b>Figura 4.27:</b> Ploteos de RMSF y radio de giro respectivamente, de la proteína PFN2 acoplada a QCT .....	72
<b>Figura 4.28:</b> Acoplamiento de PFN2 con Quercetina 3-O-glucósido en el bolsillo 1, y los residuos que interactúan con este compuesto .....	73
<b>Figura 4.29:</b> Ploteo de RMSD tras una dinámica molecular de la proteína PFN2 acoplada con QGL.....	74



<b>Figura 4.30:</b> Ploteos de RMSF y radio de giro respectivamente, de la proteína PFN2 acoplada a QGL .....	74
<b>Figura 4.31:</b> Acoplamiento de PFN2 con Quercetina 3-metil-éter próximo al bolsillo 6, y los residuos que interactúan con este compuesto en el ploteo .....	75
<b>Figura 4.32:</b> Ploteo de RMSD tras una dinámica molecular de la proteína PFN2 acoplada con QME .....	76
<b>Figura 4.33:</b> Ploteos de RMSF y radio de giro respectivamente, de la proteína PFN2 acoplada a QME .....	76
<b>Figura 4.34:</b> Acoplamiento de PFN2 con Dihidroquercetina en el bolsillo 6, y los residuos que interactúan con este compuesto .....	77
<b>Figura 4.35:</b> Ploteo de RMSD tras una dinámica molecular de la proteína PFN2 acoplada con DHQ.....	78
<b>Figura 4.36:</b> Ploteos de RMSF y radio de giro respectivamente, de la proteína PFN2 acoplada a DHQ .....	78
<b>Figura 4.37:</b> Acoplamiento de PFN2 con Isorhamnetin en el bolsillo 5, y los residuos que interactúan con este compuesto .....	79
<b>Figura 4.38:</b> Ploteo de RMSD tras una dinámica molecular de la proteína PFN2 acoplada con ISO .....	80
<b>Figura 4.39:</b> Ploteos de RMSF y radio de giro respectivamente, de la proteína PFN2 acoplada a ISO.....	80
<b>Figura 4.40:</b> Acoplamiento de PFN2 con Isorhamnetin 3-O-galactósido en el bolsillo 2, y los residuos que interactúan con este compuesto .....	81
<b>Figura 4.41:</b> Ploteo de RMSD tras una dinámica molecular de la proteína PFN2 acoplada con IGA.....	82
<b>Figura 4.42:</b> Ploteos de RMSF y radio de giro respectivamente, de la proteína PFN2 acoplada a IGA .....	82
<b>Figura 4.43:</b> Acoplamiento de PFN2 con Isorhamnetin 3-O-glucósido en el bolsillo 6, y los residuos que interactúan con este compuesto .....	83
<b>Figura 4.44:</b> Ploteo de RMSD tras una dinámica molecular de la proteína PFN2 acoplada con IGL .....	84
<b>Figura 4.45:</b> Ploteos de RMSF y radio de giro respectivamente, de la proteína PFN2 acoplada a IGL.....	84
<b>Figura 4.46:</b> Acoplamiento de PFN2 con Kaempferol en el bolsillo 6, y los residuos que interactúan con este compuesto .....	85
<b>Figura 4.47:</b> Ploteo de RMSD tras una dinámica molecular de la proteína PFN2 acoplada con KMP.....	86
<b>Figura 4.48:</b> Ploteos de RMSF y radio de giro respectivamente, de la proteína PFN2 acoplada a KMP.....	86
<b>Figura 4.49:</b> Acoplamiento de PFN2 con Kaempferol 3-methyl ether en el bolsillo 6, y los residuos que interactúan con este compuesto .....	87
<b>Figura 4.50:</b> Ploteo de RMSD tras una dinámica molecular de la proteína PFN2 acoplada con KME .....	88
<b>Figura 4.51:</b> Ploteos de RMSF y radio de giro respectivamente, de la proteína PFN2 acoplada a KME.....	88
<b>Figura 4.52:</b> Acoplamiento molecular de HDAC1 y PFN2 nativa resaltando sus extremos N- y C-terminales.....	92



<b>Figura 4.53:</b> Acoplamiento molecular de HDAC1 y PFN2 modificada por interacción con el compuesto QGL.....	92
<b>Figura 4.54:</b> Acoplamiento molecular de HDAC1 y PFN2 modificada por interacción con el compuesto QME .....	93
<b>Figura 4.55:</b> Acoplamiento molecular de HDAC1 y PFN2 modificada por interacción con el compuesto DHQ.....	93
<b>Figura 4.56:</b> Acoplamiento molecular de HDAC1 y PFN2 modificada por interacción con el compuesto ISO .....	93
<b>Figura 4.57:</b> Acoplamiento molecular de HDAC1 y PFN2 modificada por interacción con el compuesto IGA<.....	94
<b>Figura A.1:</b> Archivo minima.mdp para la minimización de energías en Gromacs.....	104
<b>Figura A.2:</b> Archivo nvt.md para la dinámica molecular en Gromacs .....	105
<b>Figura A.3:</b> Archivos script inicial, intermedio y final de los 41 scripts para la ejecución de los archivos .mdp para el cálculo NPT de los sistemas profilina humana II - flavonoides..	106
<b>Figura A. 4:</b> Primer archivo para el cálculo de NPT parte I, donde se describen los parámetros.....	107
<b>Figura A.5:</b> Primero archivo para el cálculo de NPT parte 2, donde se describen los estados y distancias que se les va a asignando a las interacciones no enlazantes como las fuerzas de Van der Waals y fuerzas Coulombicas .....	108
<b>Figura A.6:</b> Archivo script para la ejecución de todos los archivos job.sh.....	109
<b>Figura A.7:</b> Cargas de Hirshfeld calculadas para el compuesto QCT .....	109
<b>Figura A.8:</b> Cargas de Hirshfeld calculadas para el compuesto DHQ.....	110
<b>Figura A.9:</b> Cargas de Hirshfeld calculadas para el compuesto QME .....	110
<b>Figura A.10:</b> Cargas de Hirshfeld calculadas para el compuesto QRU .....	111
<b>Figura A.11:</b> Cargas de Hirshfeld calculadas para el compuesto QGL .....	112
<b>Figura A.12:</b> Cargas de Hirshfeld calculadas para el compuesto ISO .....	112
<b>Figura A.13:</b> Cargas de Hirshfeld calculadas para el compuesto IGL.....	113
<b>Figura A.14:</b> Cargas de Hirshfeld calculadas para el compuesto KME .....	113
<b>Figura A.15:</b> Cargas de Hirshfeld calculadas para el compuesto IGA .....	114
<b>Figura A.16:</b> Cargas de Hirshfeld calculadas para el compuesto KMP.....	114
<b>Figura A.17:</b> Cargas de Hirshfeld calculadas para el compuesto IRO .....	115

# ÍNDICE DE TABLAS

<b>Tabla 2.1:</b> Usos Tradicionales de Opuntia ficus-indica según una encuesta popular.....	23
<b>Tabla 2.2:</b> Compuestos flavonoides reportados en frutos y flores de Opuntia ficus-indica .	24
<b>Tabla 2.3:</b> Clasificación de los compuestos flavonoides, la estructura de las moléculas más representativas de cada sub-clase y sus fuentes naturales.....	26
<b>Tabla 4.1.</b> Resultado de los compuestos flavonoides evaluados con las reglas de Lipinski..	61
<b>Tabla 4.2:</b> Score de drogabilidad de los bolsillos o pockets de PFN2 .....	69
<b>Tabla 4.3:</b> Valores de la constante de inhibición y de energía de acoplamiento para cada uno de los sistemas proteína-ligando formados a partir del software AutoDock.....	70
<b>Tabla 4.4:</b> Cálculo de la variación de energía libre de Gibbs ( $\Delta G$ ) en un sistema que va conformando dos estados progresivamente según el módulo BAR.....	89
<b>Tabla 4.5:</b> Valores de coeficientes energéticos según el termino favorecido para el servidor ClusPro .....	91
<b>Tabla 4.6:</b> Valores del acoplamiento HDAC1 y PFN2 nativa y modificadas en servidor ClusPro con los términos de hidrofobicidad favorecidos .....	91
<b>Tabla 4.7:</b> Valores energéticos calculados con el servidor de refinamiento FireDock.....	94

# GLOSARIO

- **HAT:** Histona acetil transferasa
- **HDAC:** Histona desacetilasa
- **PFN1:** Profilina Humana I
- **PFN2:** Profilina Humana II
- **QCT:** Quercetina
- **QME:** Quercetina 3-metil éter
- **QGA:** Quercetina 3-O-galactósido
- **QRU:** Quercetina 3-O-rutinósido
- **QGL:** Quercetina 3-O-glucósido
- **DHQ:** Dihidroquercetina
- **ISO:** Isorhamnetin
- **IGA:** Isorhamnetin 3-O-galactósido
- **IGL:** Isorhamnetin 3-O- glucósido
- **IRO:** Isorhamnetin 3-O-robinobiósido
- **KMP:** Kaempferol
- **KME:** Kaempferol 3-metil éter
- **TGF-B:** Factor de crecimiento transformante  $\beta$
- **TGF- $\beta$ R:** Receptor del factor de crecimiento transformante  $\beta$
- **RMSF:** Fluctuación media cuadrática
- **RMSD:** Desviación media cuadrática
- **QM:** Mecánica cuántica
- **MM:** Mecánica molecular
- **TAC:** Tomografía axial computarizada
- **RMN:** Resonancia magnética nuclear
- **PET:** Tomografía por emisión de positrones
- **SCLC:** Cáncer de pulmón de células pequeñas
- **NSCLC:** Cáncer de pulmón de células no-pequeñas
- **OFI:** *Opuntia ficus-indica*
- **R-SMAD:** Proteína smad receptora



- **I-SMAD:** Proteína smad inhibidora
- **CO-SMAD:** Proteína smad mediadora
- **ATP:** Adenosin trifosfato
- **ADP:** Adenosin difosfato
- **PDB:** Protein data bank
- **MEC:** Matriz extra celular
- **FFT:** Transformada rápida de Fourier



# CAPÍTULO 1

## 1. PLANTEAMIENTO:

### 1.1.Objetivos:

#### 1.1.1. Objetivo General

- Evaluar la inhibición de la capacidad de unión de la profilina humana II (PFN2) a la enzima histona desacetilasa 1 (HDAC1) mediante la interacción de esta primera con flavonoides de la flor y fruto de *Opuntia ficus-indica*.

#### 1.1.2. Objetivos Específicos

- Obtener y optimizar las estructuras de la profilina humana II (PFN2) y de la enzima histona desacetilasa 1 (HDAC1).
- Diseñar y optimizar geoméricamente las estructuras tridimensionales de los compuestos flavonoides reportados en la flor y fruto de *Opuntia ficus-indica*.
- Formar complejos interactuantes entre PFN2 y los compuestos flavonoides, tras lo cual se ha de evaluar sus características geométricas y la energía de unión.
- Evaluar la capacidad de interacción entre el complejo PFN2-flavonoides con la proteína HDAC1.

## 1.2.Hipótesis:

Dado que la sobreexpresión de profilina humana II (PFN2) ha sido asociada al desarrollo de cáncer de pulmón por su interacción con la enzima histona desacetilasa 1 (HDAC1) e impedimento de su funcionamiento a nivel nuclear, es probable que esta proteína, gracias a las regiones hidrofóbicas específicas que presenta y su baja expresión en células de pulmón normales, pueda ser un blanco para moléculas como los flavonoides que han sido descritos como potenciales antioxidantes y relacionados con propiedades anticancerígenas en cáncer de pulmón.





# CAPÍTULO 2

## 2. MARCO TEÓRICO:

### 2.1.El Sistema Respiratorio: Los Pulmones:

Los pulmones son los órganos principales del sistema respiratorio, están ubicados en la caja torácica y a ambos lados del corazón, siendo el pulmón derecho más grande que el izquierdo. Son irrigados por las arterias bronquiales y las arterias pulmonares, siendo estas últimas las encargadas de llevar la sangre para su oxigenación, cumpliendo así su principal función que es el intercambio gaseoso. El flujo de aire ingresa por las fosas nasales y pasa por otros órganos como lo son la faringe, laringe y la tráquea para luego pasar por dos conductos llamados bronquios que son los responsables de comunicar a la tráquea con ambos pulmones, que a su vez, presentan unos pequeños tubos llamados bronquiolos los cuales presentan al final de su estructura unos racimos de diminutos sacos de aire llamados alveolos.

Existen un gran número de afecciones que estos órganos pueden padecer que impiden su normal funcionamiento y están clasificadas en tres tipos:

- Enfermedades de las vías respiratorias: Afectan los conductos encargados del transporte de oxígeno y otros gases al causar un estrechamiento u obstrucción.
- Enfermedades del tejido pulmonar: Afectan la estructura del tejido mediante cicatrización o inflamación, provocando una expansión restrictiva de los pulmones, lo cual minimiza la captación de oxígeno y liberación de dióxido de carbono.
- Enfermedades de la circulación pulmonar: Afectan directamente a los vasos sanguíneos al provocar su coagulación, cicatrización o inflamación. Tiene una fuerte repercusión en la actividad cardiaca en estadios avanzados.

Dentro de esta clasificación podemos someter a todas las enfermedades pulmonares existentes, pudiendo algunas de ellas ser una combinación de los tres tipos según la progresión que presenten. Siendo una de las enfermedades pulmonares más sugerentes el cáncer de pulmón, la cual presenta una elevada tasa de mortalidad a nivel mundial.

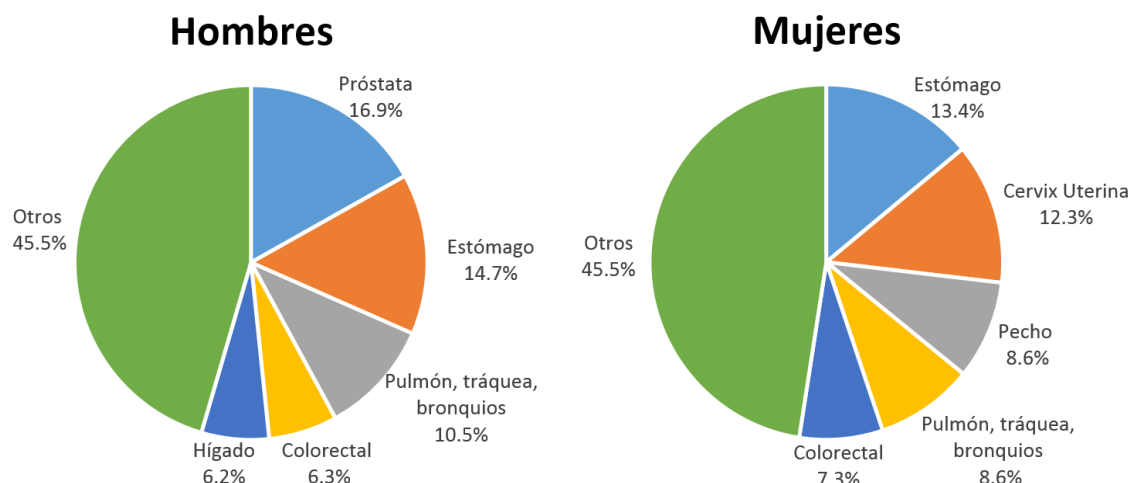
## 2.2. El Cáncer de Pulmón:

Según la Organización Mundial de Salud, el cáncer se define como un gran conjunto de enfermedades que se caracterizan por un crecimiento celular anormal más allá de los límites regulares, por lo cual tiene la capacidad de invadir partes contiguas del cuerpo humano y/o diseminarse a cualquier otro órgano. Esta enfermedad presenta muchos subtipos a nivel anatómico y molecular por lo que requiere de estrategias de manejo o tratamiento específicas.

A su vez, se ha reconocido que el cáncer, en su forma más simple, es una enfermedad genética, o de manera más precisa, es una enfermedad que se debe a una expresión genética anormal<sup>1, 2</sup>. Según una investigación realizada en el instituto suizo para la investigación experimental del cáncer (ISREC), todas las células normales al cambiar progresivamente hacia un estado neoplásico adquieren capacidades propias de toda célula cancerígena. Estas capacidades o sellos distintivos de células tumorales son: capacidad de resistir la muerte celular, señalización proliferativa, inducción de angiogénesis, evasión de supresores de crecimiento, permitir la inmortalidad replicativa, invasión y metástasis<sup>3</sup>.

Por todo ello antes mencionado, el cáncer es la segunda enfermedad con mayor tasa de mortalidad a nivel mundial con un total de 8.8 millones de muertes reportadas en el año 2015. En hombres, los tipos de cáncer más comunes son de pulmón, próstata, colon, estómago e hígado; mientras que en mujeres, los tipos más comunes son de mamas, colon, pulmón, cérvix y estómago.

El cáncer de pulmón es una enfermedad invasiva resultado de una desregulación del crecimiento de células epiteliales en el tracto respiratorio, conllevando a la formación y proliferación de células tumorales. Es uno de los tumores malignos con mayor frecuencia de casos reportados a nivel mundial siendo la mayor incidencia en hombres, y ocupando el cuarto puesto en tipo de cáncer reportado en mujeres. Las personas más propensas a este tipo de cáncer son aquellas fumadoras, en una menor proporción aquellas con antecedentes de enfermedades pulmonares como tuberculosis o fibrosis pulmonar, y fumadores pasivos. El pronóstico de supervivencia de estos pacientes es bastante bajo, llegando a un rango de 5 años solo un 10-15% de pacientes.



**Figura 2.1:** Gráfica circular de datos los perfiles de mortalidad e incidencia de varios tipos de cáncer que afectan a hombres y mujeres según la OMS (2014)

### 2.2.1. Etiología:

Existen varios factores de riesgo para el desarrollo de esta enfermedad, el más recurrente es el tabaquismo, otras causas incluyen el consumo por inhalación o ingestión de carcinógenos como el asbesto, el radón y también la influencia de la contaminación urbana y atmosférica<sup>2, 4</sup>. Otros factores endógenos en esta enfermedad son el sexo siendo las mujeres más propensas al desarrollo de cáncer de pulmón ante su exposición al humo de cigarrillo, y el factor genético donde se ha reportado la sobreexpresión de oncogenes Myc, Ras y KRas. Se ha establecido que entre el 80-90% de los casos reportados de cáncer de pulmón se dan en fumadores o personas que abandonaron este vicio recientemente habiendo una relación estrecha entre la probabilidad de adquirir esta enfermedad con la cantidad de cigarrillos por día, la profundidad de inhalación, la edad de inicio del tabaquismo y la duración de este hábito o adicción<sup>5</sup>.

### 2.2.2. Clasificación:

Podemos encontrar diversos tipos de tejidos en nuestros pulmones y por ello diferentes tipos de células, cualquiera de ellas puede iniciar un proceso de desarrollo tumoral, aunque con mayor frecuencia se asocian las células epiteliales que revisten los bronquios, tráquea o alveolos; siendo nombrados como carcinomas<sup>4</sup>. Existen dos



formas de clasificar los carcinomas pulmonares: categorizados por su tamaño seguido de un examen histopatológico bajo microscopio para evaluar la apariencia.

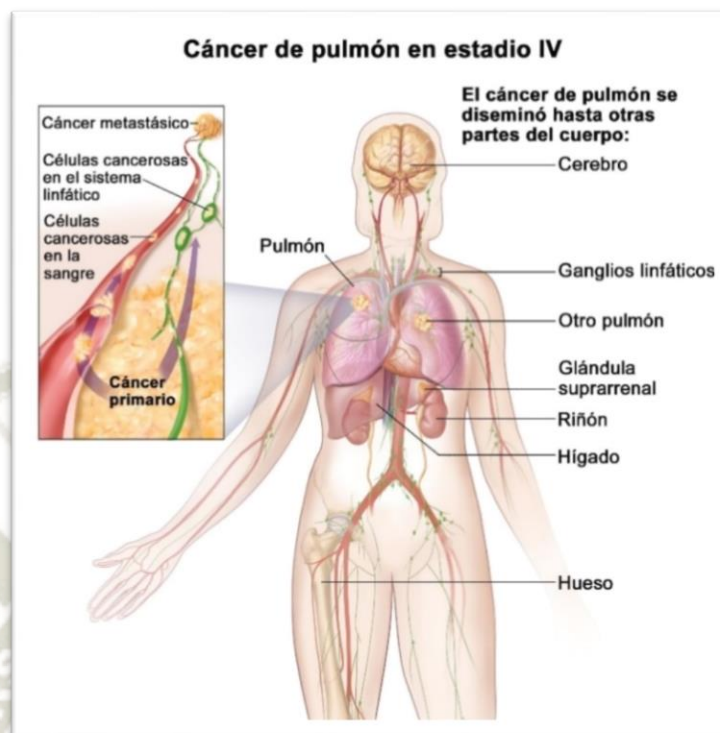
El cáncer de pulmón, desde un punto de vista histológico puede ser clasificado en 4 tipos como lo son: cáncer de células escamosas, adenocarcinomas, cáncer de células grandes y cáncer de células pequeñas<sup>5</sup>; siendo las tres primeras clasificadas como cáncer de pulmón de células no pequeñas y que abarcan un 75-85% de los casos detectados<sup>6</sup>.

Tenemos así el cáncer de pulmón de células pequeñas o microcítico (SCLC) que corresponde a un 15% aproximadamente de los casos reportados, es el más sensible a la quimioterapia en su etapa inicial pero conlleva a un peor pronóstico por su rápida proliferación; por otro lado tenemos el cáncer de pulmón de células no-pequeñas o no-microcítico (NSCLC) que presenta tres subtipos son originarios del tejido pulmonar periférico y tiene mayor probabilidad de diseminación.

### 2.2.3. Estadios de NSCLC:

Esta enfermedad puede diseminarse desde el lugar donde comenzó a diferentes partes del cuerpo por medio de 3 vías: a través del tejido, del sistema linfático y del sistema circulatorio<sup>4</sup>; y se clasifica en los siguientes estadios (Ver Figura 2.2):

- Estadio Oculto: No puede ser detectado mediante pruebas de imágenes o broncoscopia, se requiere pruebas de esputo o lavado bronquial.
- Estadio 0: Se presentan células anormales revistiendo las vías respiratorias.
- Estadio I: Ya hay formación tumoral midiendo 5cm o menos, está presente solo en el pulmón el cual puede presentar una leve inflamación.
- Estadio II: Presenta diseminación hasta los ganglios linfáticos y hasta la capa más interna del pulmón, mide más de 5cm pero no más de 7cm.
- Estadio III: El tumor puede ser de cualquier tamaño y haber invadido el bronquio principal, el diafragma, la traquea, el esófago y la membrana que rodea al corazón.
- Estadio IV: Ambos pulmones presentan tumores y se diseminó a otras partes del cuerpo como el cerebro, los riñones, el hueso y las glándulas suprarrenales.



**Figura 2.2:** Cáncer de pulmón en su último estadio y su diseminación a través del cuerpo, Instituto Nacional del Cáncer, U.S. Department of Health and Human Services

#### 2.2.4. Síntomas:

Esta enfermedad puede presentarse como un caso asintomático el cual requerirá de un hallazgo casual para su detección, y como un caso con múltiples signos dependiendo de la etapa de la enfermedad y su diseminación<sup>4</sup>:

- Síntomas Locorregionales: Dependen directamente del tumor sólido presente en el pulmón y la invasión que este presenta a tejidos adyacentes.
  - Tos(con o sin expectoración)
  - Disnea
  - Dolor torácico
  - Disfonía
  - Disfagia (alimentos sólidos o líquidos)
  - Síndrome de vena cava superior

- Síndromes vinculados al proceso de metástasis: Dependen del órgano al cual migro la enfermedad, entre estos lo más comunes son:
  - Glándulas suprarrenales
  - Hígado
  - Cerebro
  - Huesos
  - Piel
  - Ganglios linfáticos
- Síndromes Paraneoplásicos: Dependen de alteraciones secundarias inducidas por los tumores o un mal funcionamiento del sistema inmune:
  - Hipercalcemia tumoral (niveles elevados de calcio en sangre)
  - Astenia
  - Anorexia y pérdida de peso

#### **2.2.5. Diagnóstico:**

Se ha de formular una historia clínica junto con un examen físico, en caso de sospechar de este tipo de enfermedad se realizan exámenes como radiografías de tórax que confirmen presencia de quistes lo que nos lleva a una toma de muestra mediante biopsia o punción para obtención de líquido pleural; a estas muestras se les ha de realizar un examen anatomopatológico confirmatorio.

Una vez confirmada la enfermedad, se ha de determinar la etapa y extensión que presenta con exámenes como<sup>4</sup>:

- Tomografía Axial Computarizada (TAC)
- Resonancia Magnética Nuclear (RMN)
- Gammagrafía Ósea (requiere de un radiofármaco como el tecnecio)
- Tomografía por Emisión de Positrones (PET)

#### **2.2.6. Tratamiento:**

La mejor forma de combatir el cáncer de pulmón yace en su temprana detección y resección, sin embargo, para aquellos que padecen esta enfermedad la identificación de los síntomas no es posible hasta etapas avanzadas<sup>6</sup>. La etapa o grado de avance del cáncer de pulmón en diversos pacientes es un dato importante para poder realizar un



pronóstico de supervivencia del paciente, así como poder escoger el tratamiento adecuado o más óptimo<sup>7</sup>.

Actualmente, la terapia de primera línea para el tratamiento de células de cáncer de pulmón no pequeñas es la cirugía, sin embargo, tan solo un 20% de pacientes con esta enfermedad puede recurrir a ella considerando las probabilidades que sea una resección curativa exitosa. Aquellos pacientes que no califican como candidatos a cirugía pueden recurrir a otros tratamientos como lo son la quimioterapia, radioterapia y ablación percutánea por radiofrecuencia<sup>8</sup>.

### **2.3.Regulación Epigenética Mediada por la Profilina Humana II en Cáncer de Pulmón:**

Un estudio llevado a cabo en conjunto por un conjunto de laboratorios y centros de investigación en Shangai a cargo del doctor Jian-Guo Song, demostró que existe una vía de regulación epigenética para la señalización de la vía de TGF-  $\beta$  que determina o la sobreexpresión de las proteínas Smad 2 y Smad 3, mecanismo mediado por la profilina humana II (PFN2) en células de cáncer de pulmón específicamente en los adenocarcinomas<sup>9</sup>.

En este tipo de células en las cuales la vía de señalización de factores de crecimiento a través de un mecanismo molecular poco dilucidado ha alterado la expresión de TGF-  $\beta$ ; esta sobreexpresión presenta secuelas que están asociadas directamente a procesos de evasión inmune y metástasis<sup>10</sup>. Además, una sobreexpresión de la proteína profilina humana II (PFN2), una proteína ubicua que presenta niveles muy bajos en este tipo de órganos a diferencia de los niveles que presenta en el cerebro y de la cual se viene clarificando su rol en diferentes funciones celulares como proliferación, motilidad, endocitosis y en la ruta del cáncer, fue considerada como un factor que promovía los niveles de sobreexpresión de la vía de señalización de TGF-  $\beta$  al interactuar con la enzima histona desacetilasa 1 (HDAC1) evitando así su translocación nuclear, siendo el núcleo el lugar donde esta enzima cumple su función de regulador de expresión a nivel del nucleosoma. Estando el nucleosoma expuesto a factores de transcripción, es allí donde el complejo Smad 2 y Smad 3 que cumplen el rol de segundos mensajes, inducen la activación transcripcional de diversos genes según la señal enviada por TGF- $\beta$ , promoviendo de esta manera el crecimiento de células tumorales que

conllevar al proceso de metástasis y un pronóstico desfavorable para pacientes con esta enfermedad.

### **2.3.1. Vía de Señalización TGF- $\beta$ /Smads:**

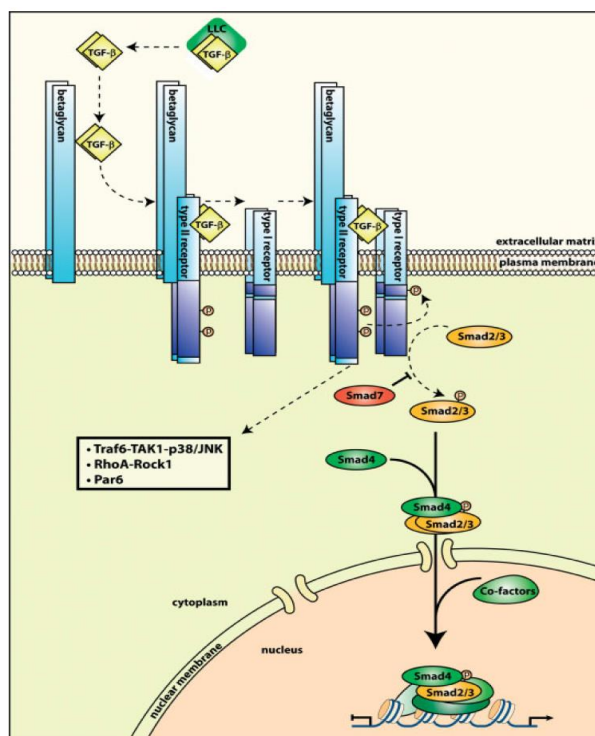
Una de las vías de señalización envueltas en este tipo de cáncer es la del factor de crecimiento transformante  $\beta$  (TGF- $\beta$ ), perteneciente a una familia de alrededor de 35 citocinas, dentro de las cuales se encuentran las inhibinas, activinas, proteína morfogénica de hueso (BMP), hormona anti-muleriana y al factor de crecimiento  $\beta$  propiamente dicho, las cuales controlan el comportamiento celular en tejidos embrionarios y adultos<sup>9-14</sup>.

Estas citocinas comparten una estructura común en la cual está presente una secuencia de residuos de cisteína altamente conservada debido a los enlaces disulfuro intra-moleculares que presenta.

La vía de señalización de TGF- $\beta$  es la encargada de regular el proceso de apoptosis, proliferación celular, diferenciación, migración y también tienen un rol como supresor tumoral mediante delecciones o mutaciones en los componentes principales de su señalización. Está implicada en varias patologías como desordenes autoinmunes y vasculares, enfermedades fibróticas y cáncer<sup>9-11, 15</sup>.

Su función esta intrínsecamente relacionada al contexto involucrado como el tipo de célula o tejido, el tipo de cáncer y la etapa de este. Su mecanismo molecular de señalización involucra receptores de membrana (TGF- $\beta$ RI, TGF- $\beta$ RII y TGF- $\beta$ RIII), proteínas reguladoras como las Smads y factores de transcripción que ayudan a la unión al DNA (Ver Figura 2.3)<sup>10, 13, 16</sup>.

En células tumorales, se presenta una alteración en esta vía de señalización lo que conlleva a una sobreproducción de TGF- $\beta$ , siendo estas cantidades elevadas las que promueven la angiogénesis tumoral, supresión inmune, metástasis y cicatrizaciones fibrosas que al acumularse en órganos como pulmón, hígado o riñón generan un estado patológico de alto riesgo y elevada tasa de mortalidad<sup>11</sup>.



**Figura 2.3:** Vía de señalización de TGF- $\beta$ , precursores inactivos, receptores de membrana y segundos mensajeros Smads. Nota. Fuente: Meulmeester, E. and P. Ten Dijke (2011). "The dynamic roles of TGF- $\beta$  in cancer." *The Journal of pathology* 223(2): 206-219.

### 2.3.1.1. Factor de Crecimiento Transformante $\beta$ :

La razón principal del contundente rol de TGF- $\beta$  en las enfermedades humanas, incluido el cáncer es su función pleiotrópica que cumple teniendo en cuenta los tipos de células sobre los que tiene efecto<sup>11</sup>; dentro de estas funciones están incluidas su implicación en la homeostasis de tejidos, el potencial de estimular la síntesis y depósito de proteínas de matriz extracelular (MEC) especialmente en células endoteliales, favorecer la interacción con estas proteínas de MEC al fomentar la expresión de receptores e integrinas de membrana, además, todos los procesos de inflamación crónicos y procesos de sanación activan esta vía de señalización de citoquinas presentes en la matriz extracelular donde están en cantidades abundantes pero como precursores inactivos<sup>10, 13</sup>.

Se han reportado cinco isoformas de TGF- $\beta$  en diversos organismos; en mamíferos se han identificado a 3 de estas isoformas de TGF- $\beta$  ( $\beta$ 1,  $\beta$ 2 y  $\beta$ 3) que presentan aproximadamente 80% de homología en su secuencia de aminoácidos pese a que su



información reside en diferentes cromosomas (19q13, 1q41 y 14q24 respectivamente)<sup>10, 11</sup>.

TGF- $\beta$  son producidos como precursor inactivo homodímero de 25kDa con una secuencia de 390 aminoácidos que se encuentra anclado a la matriz extracelular. La activación de esta proteína está bajo la influencia de factores como son temperaturas elevadas y/o un pH extremo, y también por rutas como una proteólisis limitada o desglucosilación del péptido asociada a latencia (LAP) o por su unión a la glicoproteína trombospondina 1(TSP-1) presente en la matriz extracelular. Estos procesos generan un corte proteolítico en el aminoácido 278, dejando una cadena donde su extremo C-terminal de 112-114 AA posee una secuencia de 9 residuos de cisteína que representan la región activa de la proteína<sup>10, 11</sup>.

### 2.3.1.2. *Receptores de Membrana:*

En la matriz extracelular también encontramos los receptores de TGF- $\beta$ , se han reportado 3 tipos de receptores (TGF- $\beta$ 1 de 53-65kDa, - $\beta$ 2 de 83-110kDa y - $\beta$ 3 de 250-310kDa) siendo el TGF- $\beta$ RI y TGF- $\beta$ RRI los de mayor relevancia. Ambos receptores presentan algunas diferencias como la conservación del dominio con actividad cinasa y en caso de TGF- $\beta$ RI presenta un dominio serina/treonina anclado a la membrana y es de vital importancia para su activación<sup>9-12, 15</sup>.

Estos receptores forman un complejo heterotretamérico al cual se une el ligando lo que conlleva a la activación de su actividad intrínseca de proteína kinasa; el TGF- $\beta$ RRI es el primer receptor al que se une el ligando y este transfosforila al TGF- $\beta$ RI cuya actividad kinasa es liberada de una inhibición por chaperonas, lo que da paso a la activación por fosforilación de las proteínas Smad 2 y Smad 3<sup>17, 18</sup>. Cabe mencionar que los receptores - $\beta$ RI y - $\beta$ RRI presentan una mayor afinidad por las proteínas TGF- $\beta$ 1 y TGF- $\beta$ 3<sup>10, 12, 14</sup>.

### 2.3.1.3. *Proteínas Smads:*

Smads son una clase de proteínas que tienen la función de transductores de señalización intracelular o segundos mensajeros de la vía de señalización de la familia de proteínas de TGF- $\beta$  hacia genes blanco específicos<sup>9-11, 17, 19</sup>.

Smads reciben su nombre de una contracción de los nombres de *C. elegans* “Sma” (Small body size) y de la proteína de *Drosophila melanogaster* “Mad” (mothers against dpp). Presentan una secuencia de aproximadamente 500 residuos con dos dominios altamente conservados, el MH1 (Mad-homology 1) o dominio-N y el MH2 (Mad-homology 2) o dominio-C, ambos separados por un segmento de unión menos conservado entre los tipos de esta proteína. Se tiene conocimiento de ocho proteínas Smad codificadas en los genomas de humano y de ratón. Solo 5 de estas proteínas Smads actúan como sustratos para los receptores de TGF- $\beta$ . Se subdividen según su función, están aquellos Smad que dependen del receptor (R-Smad) y son las Smad 1, 2, 3, 5 y 8; los Smad con función inhibitoria (Smad 6 y Smad 7) y por ultimo un Smad 4 con función mediadora (Co-Smad)<sup>11, 13, 14, 17, 19, 20</sup>. Los Smads 1, 5 y 8 sirven de sustratos para los receptores de la familia de BMPs mientras que Smad 2 y 3 son sustratos para TGF- $\beta$ , activina y receptores nodales<sup>18, 20</sup>.

La vía de señalización de TGF- $\beta$  fosforila y activa mediante sus receptores a proteínas Smad 2 y Smad 3, estas optan por la formación de un complejo heterotrimérico con la Smad 4 para trasladarse al núcleo donde regula la transcripción de genes específicos<sup>9-12, 17-20</sup>.

Para esto, existe un complejo de factores que regulan la importación nuclear, dentro de estos encontramos a las importinas, proteínas que se encargan del reconocimiento de una secuencia o motivo conocido como señal de localización nuclear (NLS), que al reconocer este motivo promueven la interacción de las proteínas con componentes del poro nuclear o nucleoporinas que permiten finalmente el paso al núcleo. Las proteínas Smads están sometidas a un constante proceso de transporte núcleo-citoplasmático, pero su mecanismo de importación y exportación nuclear puede darse sin la intervención de factores de transporte nuclear debido a que interactúan directamente con las nucleoporinas<sup>20</sup>.

Ya estando en el núcleo, el complejo formado por 2 R-Smads y 1 Co-Smad interactúa con otros factores de transcripción y una doble secuencia promotoras de ADN, una confiere la especificidad de unión a los factores de transcripción asociados a esta transcripción y la segunda secuencia es para la unión directa del complejo Smad, además de conferirle selectividad al complejo con respecto a la primera secuencia. Por tanto, solo un subconjunto de secuencias promotoras que se unen a estos factores de



transcripción asociados y son tomados como blancos para la señalización de Smad, brindándole selectividad a partir de secuencias parcialmente solapantes que permiten su unión<sup>14, 18, 20</sup>.

### **2.3.2. Enzima Histona Desacetilasa:**

Las histonas son proteínas con un alto contenido de lisinas y argininas en su secuencia de aminoácidos, que forman parte del núcleo proteico de un nucleosoma, dotándole de un carácter básico<sup>21</sup>. Han sido identificadas 5 tipos de histonas: H1, H2A, H2B, H3 y H4, de las cuales dos H2A, dos H2B, dos H3 y dos H4 forman el octámero que se presenta en todo nucleosoma<sup>21-25</sup>. Se determinó además que un nucleosoma está compuesto de un fragmento de ADN de 146 pares de bases en promedio, las cuales se enrollan en un sentido anti horario y formando 1.75 giros alrededor de un octámero de histonas<sup>23-25</sup>. Estas histonas nucleosómicas presentan dos dominios, el primer dominio o C-terminal está localizado en el interior del nucleosoma y el segundo dominio o N-terminal está conformado por los residuos de lisina en forma de cadenas laterales que sobresalen fuera del mismo. Estas cadenas laterales del extremo N-terminal pasan a través y alrededor de la doble hélice de ADN, siendo su modificación un suceso determinante para la regulación de la expresión genética. El ADN, las histonas y otras proteínas forman la cromatina, un complejo macromolecular a base de nucleosomas; la cromatina presenta dos formas: heterocromatina, que se caracteriza por ser bastante compacta e inactiva para la transcripción, y la eucromatina, la cual está desplegada y es importante porque promueve los procesos de transcripción y expresión genética<sup>21, 22</sup>.

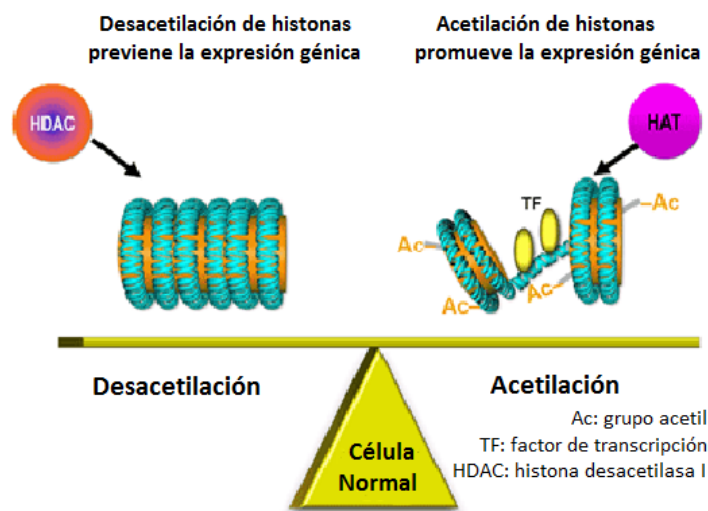
La expresión genética es un proceso que está regulado en parte por el nivel de acetilación de las histonas nucleosomales, tanto para la activación transcripcional (hiperacetilación) o represión de la transcripción (hipoacetilación)<sup>26</sup>. Este fenómeno está estrechamente regulado por los niveles de expresión de dos enzimas: histona acetil transferasa (HAT) y la enzima histona desacetilasa (HDAC)<sup>22, 23, 25</sup>. Los grupos amino de los residuos de lisina que se encuentran en los extremos terminales de las histonas, son los sustratos para estas enzimas. Las HATs transfieren un grupo acetilo a estas cadenas mientras que las HDACs eliminan los grupos acetilos de las ε-N-acetil lisinas. La alteración de los niveles de alguna de estas dos enzimas está estrechamente



relacionado con una expresión génica aberrante y crecimiento tumoral en diversos tipos de cáncer<sup>23, 27-31</sup>.

Dentro de las hipótesis de los mecanismos que describen la desregulación de estas enzimas están<sup>25, 26, 30</sup>:

- Una hiperacetilación puede activar promotores que regularmente están reprimidos, que conlleva a la activación inapropiada de algunas proteínas.
- Una hipoacetilación en regiones promotoras reprime la expresión de genes necesarios para un cierto fenotipo
- El reclutamiento equivoco de HAT o HDAC podría significar un desencadenante patológico (Ver Figura 2.4).



**Figura 2.4:** Funcionamiento regular del proceso de desacetilación (HDAC) y acetilación (HAT). Nota.

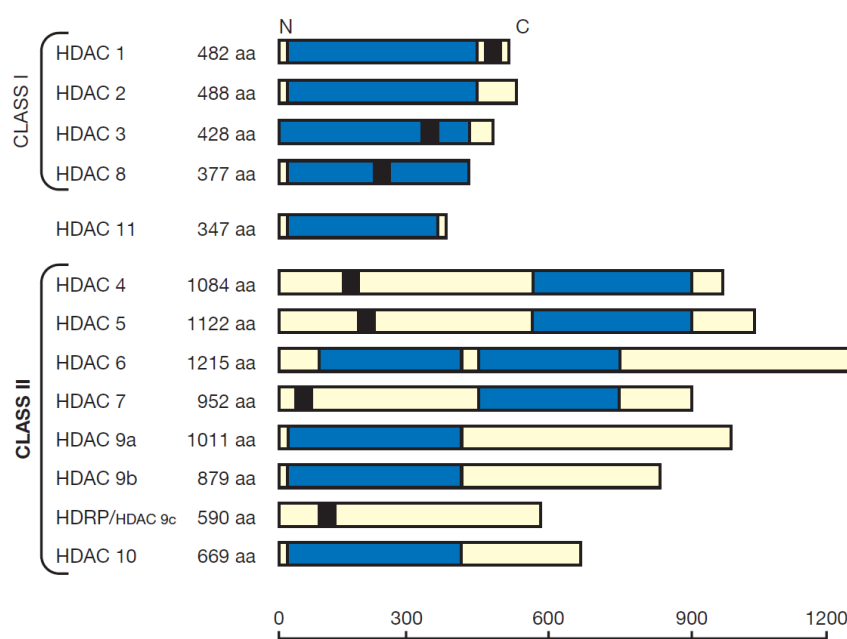
*Adaptado. Fuente: MERCK Presentation*

A pesar de que una alteración directa sobre los genes de estas enzimas no está vinculada con el cáncer, la asociación de sus funciones es clara en caso de genes supresores tumorales.

Las histonas desacetilasas son una súper familia de enzimas (Ver Figura 2.5) que regulan la expresión génica; se han identificado diversas enzimas de este tipo, que a su vez se encuentran clasificadas en tres clases según su homología estructural hacia tres enzimas HDACs de levaduras: Rpd3 (HDACs clase I), Hda1 (HDACs clase II) y Sir2/Hst (HDACs clase III)<sup>21, 28, 32</sup>.

La clase I de HDACs engloba las isoformas HDAC1, HDAC2, HDAC3 y HDAC8; están expresada en diversos tejidos y su localización principal es el núcleo aunque también pueden estar presentes en el citoplasma. La clase II que engloba las isoformas HDAC4, HDAC5, HDAC6, HDAC7, HDAC9 y HDAC10; presentan un mayor tamaño molecular, se encuentran en tejidos específicos y su ubicación varía entre el núcleo y el citoplasma. La clase III engloba una familia de sirtuinas (reguladores de información silenciosos o SIRs) que presentan un mecanismo de acción dependiente del cofactor  $\text{NAD}^{+22, 24, 26, 28, 30, 31, 33}$ .

Ambas clases I/II a diferencia de las sirtuinas, presentan un mecanismo dependiente de zinc, que es el componente principal de su centro catalítico, y están emergiendo como blancos terapéuticos para el tratamiento de cáncer y otras enfermedades. Los inhibidores de HDAC son capaces de reprimir la acción de estas enzimas lo que conlleva a efectos anti-proliferativos in vitro y efectos antitumorales in vivo<sup>22, 24, 34</sup>.



**Figura 2.5:** Esquema de los diferentes tipos de HDAC, su clasificación, el largo de las proteínas en base a su número de aminoácidos y en azul su dominio catalítico. Nota. Fuente: De Ruijter, A. J., et al. (2003). "Histone deacetylases (HDACs): characterization of the classical HDAC family." *Biochemical Journal* 370(Pt 3): 737.

HDAC1 y HDAC 2 son enzimas con un alto porcentaje de similitud alcanzando un 82% aproximadamente, su extremo N-terminal comprende la mayor parte de la enzima así como su dominio catalítico que consiste en un bolsillo tubular hidrofóbico en cuyo extremo se localiza el átomo de zinc ( $Zn^{+2}$ ). Estas enzimas son sintetizadas como precursores inactivos y requieren de cofactores para su activación, también requieren de la unión a complejos proteicos para regular su transcripción genética ya que no pueden unirse directamente al ADN. Algunos de estos complejos multiproteicos son: Sin3, corepresor de receptor nuclear (N-CoR), mediador de silenciamiento del ácido retinoico y del receptor de hormona tiroide (SMRT), activadores y proteínas remodeladoras de cromatina<sup>22, 24, 26, 30, 33</sup>.

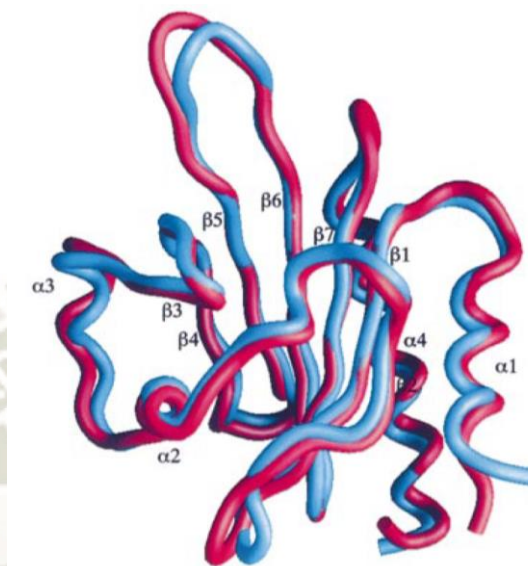
En el presente estudio, se hizo uso de la Histona Desacetilasa 1 (HDAC1) debido a su intervención en la regulación epigenética de la vía de señalización TGF- $\beta$  /Smads<sup>9</sup>. Esta enzima presenta una secuencia de 482 aminoácidos estando como precursor inactivo, pero mediante técnicas de cristalografía de rayos X, se dilucido su estructura activa donde presenta una secuencia de 369 aminoácidos conservando su centro catalítico de Zinc<sup>34</sup>.

### 2.3.3. Las Profilinas Humanas:

Las profilinas (PFNs) conforman una pequeña familia de proteínas con un tamaño entre 12-15 kDa; la primera profilina reportada fue aislada de un tejido de timo de ternero cuyas primeras impresiones fueron el ser un polipéptido que no presentaba homología alguna con otras proteínas, pero de aquí en adelante, esta proteína fue reportada en diversas células y organismos como amebas, levaduras, moscas, plantas y diversos mamíferos<sup>35-42</sup>. Algunas especies expresan homólogos de esta proteínas (Ver Figura 2.6), en el caso de mamíferos, se han descubierto hasta cuatro genes que la codifican (PFN-1, -2, -3 y -4); profilina I (PFN1) es la más representativa de esta familia y es expresada ubicuamente en todo tipo de tejidos a excepción del musculo esquelético; profilina II (PFN2) que esta expresada mayormente en tejidos de cerebro, músculo esquelético e hígado, a su vez esta proteína presenta dos variantes (PFN2a y PFN2b), PFN2a que es la isoforma mas predominante entre ambas y está altamente expresada en células neuronales, y PFN2b cuyo mRNA ha sido detectado en un limitado número de tejidos adultos como en el hígado; PFN3 y PFN4 por el contrario



de las anteriores han sido poco estudiadas debido a que la expresión de estas proteínas se limita a los testículos y en menor proporción en el hígado, estando implicadas en el proceso de espermatogénesis<sup>37-39, 43-48</sup>.



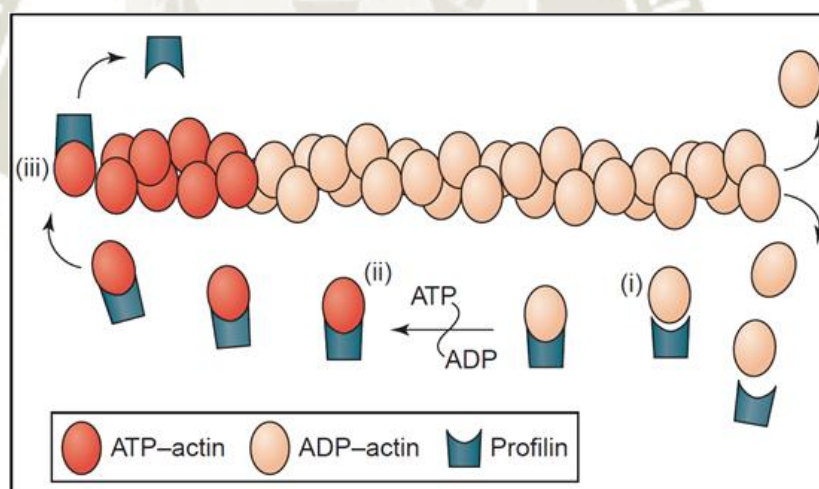
**Figura 2.6:** Superposición de estructuras tridimensionales de Profilina Humana I (rojo) y Profilina Humana II (azul), resaltando elementos de su estructura secundaria como son las  $\alpha$ -hélices y las láminas  $\beta$ . Nota. Fuente: Nodelman, I. M., et al. (1999). "X-ray structure determination of human profilin II: A comparative structural analysis of human profilins." *Journal of molecular biology* 294(5): 1271-1285.

### 2.3.3.1. Rol de las Profilinas Humanas:

Las funciones de las profilinas en tejidos eucariotas aún no han sido totalmente dilucidadas, pese a que varias de estas ya han sido reportadas; estas proteínas presentan un papel fundamental en el funcionamiento celular estando estrechamente relacionadas con su gran capacidad de regulación de la dinámica de polimerización de actina (Ver Figura 2.7), entre las otras funciones de estas proteínas, se encuentran la proliferación celular, migración, endocitosis y transcripción génica<sup>35, 36, 45, 48-50</sup>.

Las actinas son proteínas globulares que forman microfilamentos para el citoesqueleto celular y a su vez hidrolizan moléculas de ATP, estando presente en forma libre como monómero de G-actina o como parte de un polímero lineal (microfilamento) en forma de F-actina<sup>37, 48, 51, 52</sup>. Las profilinas fueron descritas inicialmente como el factor responsable de la despolimerización de filamentos de actina in vitro, ya que al unirse

con monómeros de actina evitaban así su polimerización, años después lograron desmentir esto al notar que las profilinas aceleran el intercambio de nucleótidos ADP unidos a los monómeros de actina por un complejo ATP-actina, incrementando la disponibilidad de esta para una polimerización más veloz al interaccionar con el extremo creciente del microfilamento. Si bien esta formación del complejo actina-profilina presenta gran afinidad e importancia siendo un mecanismo único ya que ninguna otra proteína de unión a actina entre las 70 clases diferentes que existen tiene un efecto similar, la profilina es una proteína multifuncional que puede interactuar con muchas otras proteínas que regulan una compleja red de interacciones moleculares<sup>51, 53</sup>. Dentro de estas interacciones predominan moléculas como fosfatidilinositol 4,5-bisfostato (PIP<sub>2</sub>) y proteínas con una secuencia o dominio poli-L-prolina (PLP) como lo son la fosfoproteína de estimulación de vasodilatación (VASP), Enabled (Ena), la proteína del síndrome Wiskott-Aldrich (WASP), su homólogo neuronal (N-WASP) y mammalian diaphanous (mDia1)<sup>35, 37, 40, 44, 46-48, 53-56</sup>.



**Figura 2.7:** Rol de las profilinas en la polimerización de actinas. Nota. Fuente: Witke, W. (2004). "The role of profilin complexes in cell motility and other cellular processes." *Trends in cell biology* 14(8): 461-469.

A pesar de una baja homología en la secuencia de las profilinas, existe una gran similitud bioquímica y estructural entre estas proteínas, tal es el caso que las estructuras de PFN1 y PFN2 pueden ser superpuestas casi en su totalidad, lo cual sugiere que las secuencias para su unión a ligandos se mantengan conservadas<sup>35, 41</sup>. Sin

embargo, se demostró que PFN1 presenta una afinidad de unión a actina casi 5 veces mayor que PFN2, y también una mayor afinidad de unión a PIP2 mientras que PFN2 presenta mayor afinidad de unión a proteínas con dominios de poli-L-prolina<sup>44, 46</sup>.

Se cree que la mayoría de ligandos han de unirse a la profilina por su dominio o secuencia para poli-L-prolinas, aunque varios de estos ligandos presentan repetición de estos aminoácidos con límite de tres o cuatro consecutivos teniendo a otros aminoácidos como glicina, prolina o alanina en reemplazo formando así secuencias definidas como ZPPX (donde Z suele ser alguno de los aminoácidos mencionados anteriormente y X suele ser un residuo preferentemente hidrofóbico), esta secuencia ZPPX puede ser encontrada en miles de proteínas por lo que la especificidad para unión con profilina está aún bajo estudio, pero se observó que influyen ampliamente sus extremos –N y –C terminales del dominio de unión de esta secuencia. Además, fueron reportados algunos ligandos que difieren de esta presunción e interactúan con un dominio diferente como gephyrin<sup>37, 47</sup>.

### 2.3.3.2. *Profilina Humana II (PFN2):*

La estructura de la profilina humana II fue determinada por técnicas de cristalografía y rayos X a una resolución de 2.2Å, su estructura tridimensional consta de 7 láminas betas ( $\beta$ -sheets) antiparalelas y enlazadas, y de 2 hélices alfa ( $\alpha$ -helix) en los extremos –N y –C terminales y otras dos hélices en el otro extremo<sup>35, 51</sup>. Presenta a diferencia de PFN1 una extensión aromática en el sitio de unión dispuesto para poli-L-prolina que se debe principalmente al cambio en la secuencia de aminoácidos de una serina en la posición 29 por una tirosina (S29Y) y le confiere la capacidad de formar una nueva interacción por apilamiento de bases y la formación de un enlace puente de hidrogeno, destacando así su afinidad por aquellas proteínas con esta secuencia de prolinas<sup>35</sup>.

La expresión de PFN2 en cerebro está limitada a células neuronales (detectada mayormente en el citoplasma) donde es reclutada hacia las espinas dendríticas por medio de unos neurotransmisores AMPA, las espinas dendríticas son las efectoras de la sinapsis neuronal y debido a la alta motilidad que presentan tiene un alto contenido de actina que pasa a ser polimerizado por la PFN2 pero también en menor grado por PFN1, al ser polimerizadas las actinas las espinas dendríticas reducen su motilidad lo que favorece la consolidación de la sinapsis<sup>39</sup>. Además de esta función, PFN2 ha sido



vinculada a múltiples casos de cáncer en diversos tejidos, aunque con un actuar inconsistente, siendo así un regulador negativo del crecimiento tumoral en un caso de carcinoma oral de células escamosas, pero opuesto a esto, eleva la tasa de migración e invasión celular en casos de cáncer colorrectal y cáncer de pulmón<sup>39</sup>. En este último caso de cáncer de pulmón, la profilina humana II (PFN2) ha sido implicada en la regulación epigenética de la vía TGF- $\beta$ /Smad mediante su unión a una enzima desacetilasa de histonas (HDAC)<sup>9</sup>.

## 2.4. *Opuntia ficus-indica*:

Las cactáceas son una familia de plantas que engloban un aproximado de 2000 especies dentro de las cuales resaltan las del genero *Opuntia*, del cual están reportadas 190 especies de amplia distribución en el continente americano. *Opuntia ficus-indica* es la especie más conocida también denominada como tuna, penca, higuera o chumbera que se encuentra distribuida desde el sur de Canadá hasta la Patagonia<sup>57-61</sup>. La Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO) fomenta el cultivo de estas plantas debido a que presenta gran éxito para agricultores que no cuentan con los recursos necesarios y que residen en zonas áridas, sin mencionar las cuantiosas propiedades de la planta y sus frutos<sup>59, 62, 63</sup>.

### 2.4.1. Fisiología:

Es una planta arbustiva con tallos ligeramente espinosos, aplanados y ramificados que se les conoce como cladodios o pencas, que pueden llegar a alcanzar de 20-60 centímetros de altura y 10-40cm de ancho, su color varia de un verde claro en su etapa de juventud a un gris en su etapa más adulta<sup>57</sup>. Estos cladodios son los principales almacenes de agua y la epidermis que los recubre presenta tres funciones<sup>59, 64</sup>:

- Regulación del intercambio gaseoso de dióxido de carbono y oxígeno.
- Retención del agua dentro del cladodio.
- Protección contra hongos, bacterias e insectos.

Sus flores brotan a partir de areolas que se forman en los bordes de los cladodios, se presentan en los meses de marzo, abril y mayo principalmente, se desarrollan hasta 20 flores por cada cladodio que presente la planta en un periodo de 3 a 5 semanas,

alcanzan de 7-10 cm de largo y tienen una coloración amarilla, rojo, naranja o rosa. Sus frutos son de coloración amarilla, roja, naranja o purpura con una forma ovalada de 5-10 cm de largo y 4-9 cm de ancho<sup>57, 59, 65</sup>.

Esta planta presenta un mecanismo fotosintético conocido como Metabolismo del Ácido Crasuláceo (MAC), este fija el carbono por la noche al disminuir las temperaturas y aumentar la humedad relativa del ambiente y manteniendo los estomas cerrados durante el día, promoviendo un uso eficiente del agua.



**Figura 2.8:** *Opuntia ficus-indica*, vista de sus cladodios, espinas, sus frutos y flores

#### 2.4.2. Hábitat:

*Opuntia ficus-indica* o nopal como es su nombre genérico, es un cultivo de origen desconocido pero que presenta gran adaptación a condiciones ambientales extremas como altas temperaturas y radiación UV extrema en zonas áridas o semiáridas, estando presente en varias regiones de nuestro país, pudiendo cultivarse en altitudes desde los 800 hasta los 2500 msnm<sup>58, 62</sup>.

Requieren poca cantidad de agua para su cultivo resistiendo épocas de sequía, pueden desarrollarse en suelos pobres en nutrientes y en el medio en el que se desarrollan presenta temperaturas entre los 10 y 50°C como mínimo y máximo respectivamente, su alta tolerancia a calor está mediada por la acumulación del aminoácido prolina principalmente en sus raíces y también diversos órganos de la planta; sin embargo, muchas de sus características bioquímicas permanecen desdeñadas<sup>58, 59, 62</sup>.

### 2.4.3. Aplicaciones:

Tradicionalmente y primordialmente se ha hecho uso de esta planta con fines medicinales para la cura de: estreñimiento, úlceras estomacales, quemaduras, diarrea, fiebre, dolores estomacales, etc<sup>59</sup>. Es así que, por sus propiedades antiinflamatorias, diuréticas, analgésicas, antioxidantes, neuroprotectoras, hepatoprotectoras y anticancerígenas entre otras reportadas empíricamente, ha sido objeto de estudios científicos que revelan su alto contenido en fibras de lignina, celulosa, pectina, minerales, vitaminas, beta carotenos, flavonoides y diversos metabolitos que podrían confirmar estas propiedades<sup>57, 59, 60, 63, 64, 66-70</sup>.

**Tabla 2.1:** Usos Tradicionales de *Opuntia ficus-indica* según una encuesta popular

Partes de la Planta Usadas	Método de Preparación	Prácticas Tradicionales
Cladodios	El material seco se pulverizó y se disolvió en agua	Para aliviar el dolor estomacal
	Se colectó mucilago y se bebió en mezcla con leche	Para enfermedades del hígado
	Se cortan en rodajas y se colocan en salmuera	Reduce los niveles de colesterol
Flores	Material seco y pulverizado se hirvió con agua	Para enfermedades del hígado
	Material seco y pulverizado se hirvió con agua	Para la enfermedad de próstata
	Material seco y pulverizado se mezcló con miel	Usado para el asma
Semillas	Se secaron y pulverizaron	Para enfermedades de la piel
	Se secaron y pulverizaron	Para dolencias cardiovasculares
Raíces	Se secaron, pulverizaron y mezclaron con leche y miel	Usado para el asma

**Nota.** Fuente: Adaptado de Mouhaddach A, El-hadi A, Taghzouti K, Bendaou M, Hassikou R. Assessment of *Opuntia ficus-indica* in vivo following ethnobotanical survey: Confirmation of its analgesic activity. *Phytothérapie*. 2017:1-6.

Además, esta planta obtiene gran importancia debido a su uso como materia prima para el cultivo de la cochinilla (*Dactylopius coccus*), un insecto utilizado como tinte orgánico, siendo esta característica su potencial en el mercado mundial y siendo nuestra región Arequipa el líder exportador<sup>59</sup>.

### 2.4.4. Flores y frutos de *Opuntia ficus-indica*:

En cuanto a las flores y frutos de *Opuntia ficus-indica*, estos no presentan tanto protagonismo como los cladodios pese a su amplio uso en medicina tradicional, siendo usados para reducir la inflamación, reducir el dolor, tratar mordidas de serpientes, sus efectos diuréticos entre otros<sup>63, 67</sup>; en las pocas investigaciones acerca de las flores y frutos se reportó un alto contenido en compuestos polifenólicos biofuncionales siendo kaempferol, isorhamnetin y quercetina los más reportado<sup>57, 60, 61, 67, 70, 71</sup>.



Los frutos o tunas, aunque son de amplio consumo se desconoce todas sus propiedades bioquímicas; los reportes según el doctor Silos Espino, indican que tiene en su contenido 35% de glucosa, 2.9% de fructosa, 5.1% de proteínas en la pulpa y 8.3% en la cascara, 0.1% de lípidos, 49mg de calcio, 220mg de potasio y 85mg de magnesio en 100g de materia seca además de la alta cantidad de fibra presente<sup>58</sup>. A partir de un estudio fotoquímico llevado a cabo por Jung Wha Kim, se logró identificar 13 compuestos fenólicos dentro de los cuáles 7 son reportados como flavonoides, mientras que en el estudio llevado a cabo por Hyang Dok-Go se identificó que en frutos y semillas de nopal se encuentran 3 flavonoides (Ver Tabla 2.2)<sup>72, 73</sup>.

Las flores por otra parte, son la materia prima que presenta menor importancia en esta planta, suelen ser objeto de desecho o en algunos casos para infusiones en la medicina tradicional. Investigaciones reportan que estas, presentan un alto contenido de flavonoides (Ver Tabla 2.2) con un valor de 81.75mg por cada gramo de material fresco según el estudio llevado a cabo por M. De Leo donde también reporta presencia de compuestos flavonoides presentes en flores<sup>61, 65</sup>.

**Tabla 2.2:** *Compuestos flavonoides reportados en frutos y flores de Opuntia ficus-indica*

Partes de la Planta	Compuesto Flavonoide	Abreviación
Fruto	Quercetin 3-O-rutinoside	QRU
	Quercetin 3-O-glucoside	QGL
	Isorhamnetin 3-O-robinobioside	IRO
	Isorhamnetin 3-O-galactoside	IGA
	Isorhamnetin 3-O-glucoside	IGL
	(+) - Dihydroquercetin	DHQ
Flor	Quercetin	QCT
	Quercetin 3-methyl ether	QME
	Isorhamnetin	ISO
	Kaempferol	KMP
	Kaempferol 3-methyl ether	KME

## 2.5. Compuestos Flavonoides:

Los compuestos fotoquímicos más relevantes en las vegetales son los compuestos fenólicos, habiendo sido identificados más de 8000 de estos compuestos, de los cuales aproximadamente la mitad son compuestos flavonoides, estando estos presentes en una gran variedad de plantas vasculares inclusive presentando altos niveles, siendo además asociados a la pigmentación de flores y frutos; aunque también están presentes en sus semillas y raíces<sup>68, 71, 74-76</sup>.

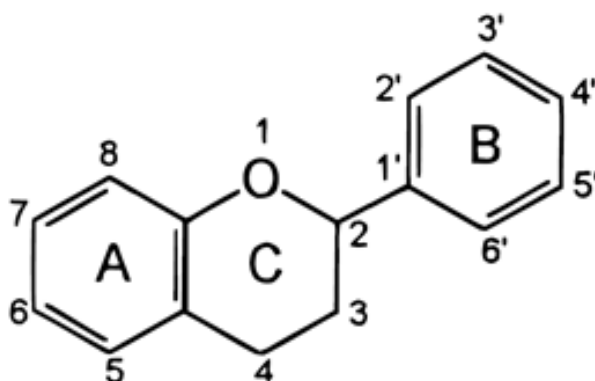
Su rol en los vegetales y otras células eucariotas en los que está presente es diverso, siendo un compuesto antimicrobiano, antioxidante, fotoreceptor y repelente; además, estudios han reportado que estos compuestos presentan actividad biológica teniendo propiedades antialérgicas, antivirales, antiinflamatorias, vasodilatadoras y anticancerígenas, dentro de las cuales resalta su rol anticancerígeno y su actividad antioxidante al reducir la formación de radicales libres y lidiar con aquellos ya presentes<sup>68, 73, 74, 77-79</sup>.

El rango de consumo de flavonoides esta entre 50 y 800mg/día, dependiendo principalmente de la ingesta de frutas y vegetales donde son los flavonoles: quercetina y kaempferol los más abundantes<sup>68</sup>. Todo flavonoide al ser ingerido pasa por un proceso de degradación hacia ácidos fenólicos en el estómago e intestino, y de absorción en este último. En los últimos años, se ha demostrado solo un porcentaje no muy elevado de estos compuestos es absorbido, manteniendo estos flavonoides y sus metabolitos o ácidos fenólicos que permanecen su actividad antioxidante que puede desarrollarse *in vivo*; esto ha sido comprobado al evaluar los niveles de antioxidantes en plasma, los niveles de vitamina E regulados en la membrana de eritrocitos y una baja densidad de las lipoproteínas<sup>68</sup>.

### 2.5.1. Estructura y Clasificación:

Los flavonoides son compuestos de bajo peso molecular que se sintetizan en plantas a partir de aminoácidos aromáticos como lo son la fenilalanina (Phe) y Tirosina (Tyr) junto con el anión malonato, siguiendo la vía metabólica de fenilpropanoide<sup>66, 68, 74</sup>. Presentan una estructura básica conocida como flavano, que está compuesta por 15 átomos de carbono distribuidos en tres anillos, dos anillos benceno unidos por una

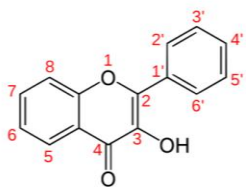
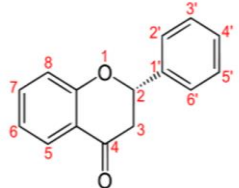
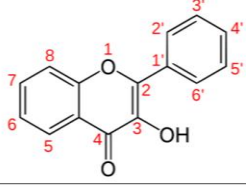
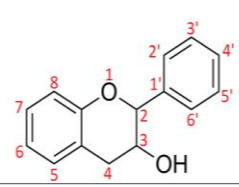
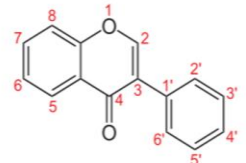
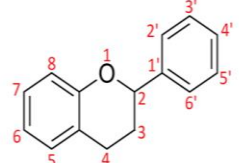
cadena heterocíclica de tres carbonos (C6-C3-C6), cada uno de los anillos fue etiquetado según un consenso (Ver Figura 2.9)<sup>68, 74, 75</sup>.



**Figura 2.9:** Estructura básica de un compuesto flavonoide o flavano

Los flavonoides se clasifican en 6 subclases principales según su estructura molecular y las varias combinaciones de grupos hidroxilo, metoxi y grupos O-glucósidos, siendo los siguientes: flavonoles, flavonas, isoflavonas, flavanonas, flavonoles y antocianidinas (Ver Tabla 2.3)<sup>74, 76, 79-81</sup>.

**Tabla 2.3:** Clasificación de los compuestos flavonoides, la estructura de las moléculas más representativas de cada sub-clase y sus fuentes naturales

Compuestos Flavonoides	Estructura	Representantes	Fuente natural	Compuestos Flavonoides	Estructura	Representantes	Fuente natural
Flavonol		Quercetina Kaempferol Isorhamnetin	Cebolla Brócoli	Flavanona		Hesperetina Naringenina	Frutos Cítricos Tomates
Flavona		Apigenina Luteolin	Perejil Apio	Flavan-3-ol		Catequina Epicatequina	Chocolate Vino Tinto
Isoflavona		Daidzeina Genisteína	Productos de Soya	Antocianidina		Cianidina Malvidina	Vino Tinto Berries



El nivel de oxidación y el patrón de sustitución del anillo C en estos compuestos los diferencia; y, es el patrón de sustitución de los anillos A y B lo que diferencia las diferentes moléculas en estos grupos. Los flavonoides se hidroxilan en los carbonos 2, 3, 5, 7, 3', 4' y 5', además pueden haber especies metil-éter y acetil-éster en estos – OH, así como unión de glicósidos en los carbonos 3 o 7 como D-glucosa, galactosa, arabinosa y otros. Es así que, por ejemplo, los flavonoles se diferencian de las flavononas por un grupo hidroxilo en el carbono 3 y un doble enlace entre este carbono 3 y el carbono 2 ( $C2 = C3$ )<sup>79, 80, 82</sup>.

### **2.5.2. Propiedades Antioxidantes:**

Estas propiedades de los compuestos flavonoides están intrínsecamente ligadas a sus grupos funcionales alrededor de la estructura base. La sustitución en los carbonos y el número total de grupos hidroxilos son los que confieren principalmente esta característica de antioxidante<sup>77</sup>.

El anillo B y sus grupo hidroxilo son los que más determinan la capacidad de erradicar los radicales libres ya que donan un átomo de hidrogeno y un electrón a una especie radical para estabilizarla; siendo los flavonoles y flavonas los de mayor actividad por la presencia de un grupo catecol en su anillo B<sup>76, 82</sup>.

Es así que los flavonoides presentan también capacidad de absorción de rayos UV del sol y eliminación de los compuestos reactivo de oxígeno generados, brindando protección a las plantas en las que están presentes<sup>83</sup>.

### **2.5.3. Propiedades Anticancerígenas:**

Estudios epidemiológicos han podido comprobar que hay una asociación entre el consumo de flavonoides y una protección contra el cáncer, siendo el cáncer de pulmón el único reportado en tener una correlación inversa respecto al consumo de estos compuestos<sup>68, 75</sup>. Los flavonoides tienen la capacidad de interactuar con diversas proteínas, estando varias de ellas relacionadas con vías metabólicas que promueven el desarrollo de cáncer<sup>71, 76, 79, 81</sup>.

## 2.6. Química Computacional y Bioinformática:

La química computacional y la bioinformática son dos líneas de investigación muy relacionadas, ambas hacen uso de herramientas computacionales para resolver problemas químicos mediante simulaciones de sistemas moleculares y para hacer un análisis de datos biológicos respectivamente<sup>84</sup>.

La bioinformática implica diversas disciplinas como matemática aplicada, estadística, informática, química y biología molecular, que conjuntamente brindan herramientas para análisis de datos, simulación de sistemas biológicos a un nivel molecular, siendo varias de las líneas de investigación aquellas encargadas de dilucidar: alineamiento de secuencias (mRNA, cDNA), diseño de primers, predicción de genes codificantes, expresión génica, secuenciamiento de aminoácidos, estructura tridimensional de proteínas y filogenia<sup>85, 86</sup>.

La química computacional por su parte, se enfoca a la resolución de problemas o fenómenos que el método experimental no puede dilucidar ya sea por el alto costo implicado y/o por la carencia de equipamiento tecnológico que brinde un mayor alcance en sus campos de investigación. Es así que, mediante el uso de conceptos y modelos de áreas como la mecánica cuántica, mecánica molecular, física clásica y química teórica incorporados en software para el modelamiento de estructuras, cálculo de energías a nivel atómico o molecular de complejos en estado gaseoso o en solución y capacidad de interacción entre moléculas, se busca complementar la información obtenida mediante experimentación, orientar el diseño experimental en búsqueda del diseño de nuevos fármacos y/o materiales, y finalmente en el poder predecir un fenómeno químico que no presente antecedentes<sup>87</sup>.

### 2.6.1. Mecánica Cuántica:

La mecánica cuántica es una rama de la física que se define como un conjunto de leyes que rigen el comportamiento de partículas como los electrones, núcleos atómicos hasta algunas moléculas. Surgió a partir de la necesidad de explicar ciertos fenómenos, siendo uno de ellos el carácter ondulatorio-corpúscular de la luz; para lo cual se establecieron parámetros de partícula (energía  $E$  y momento  $p$  de un fotón) y parámetros de onda (frecuencia angular  $\omega$ , vector de onda  $k$ , frecuencia  $\nu$  y longitud de onda  $\lambda$ ), siendo más tarde postulada la teoría de la dualidad de la materia de Broglie

que refiere al hecho de que la materia puede ser estudiada tanto como partículas o como ondas dependiendo de la propiedad que se quiera evaluar.

A partir de ello, surge la teoría de la incertidumbre de Heisenberg, que establece que es imposible el poder determinar tanto la posición como el momento o velocidad de una partícula de manera simultánea, un hecho que a macro escala es insignificante pero a escala cuántica la variación de la medida de una magnitud puede hacer variar el estado del sistema.

Finalmente, fue Erwin Schrödinger quien describe que al tener la materia un carácter dual onda-corpúsculo, entonces cada partícula debe estar asociada a una función de onda  $\Psi(r,t)$  que define su evolución temporal, siendo  $r$  las coordenadas tridimensionales y  $t$  el tiempo o momento. Es así que, para describir el estado de un sistema se describe una función de onda  $\Psi$  dependiente del tiempo, y teniendo la primera derivada de esta se puede calcular la función de onda futura en cualquier tiempo; quedando descrita la ecuación de Schrödinger de la siguiente manera<sup>88</sup>:

$$H \Psi(r,t) = E \Psi(r,t) \quad (2.1)$$

Esta ecuación describe lo siguiente: el operador hamiltoniano que define a los electrones y al núcleo de un sistema,  $\Psi$  es la función de onda que contiene la información sobre el estado de los electrones en una posición y momento predeterminado,  $E$  es la energía asociada al estado del electrón ligada a la posición y momento deseado.

### 2.6.2. Mecánica Molecular o Mecánica Clásica:

La mecánica molecular es parte del modelado molecular, haciendo uso de los principios de la física clásica para describir los modelos o sistemas moleculares, percibiendo a los átomos como esferas (los electrones y el núcleo forman un conjunto) que se unen a través de puentes o muelles que representan los enlaces químicos entre átomos pero adicionalmente contemplan las fuerzas no enlazantes en el sistema.

La mecánica molecular ha sido usada desde sus inicios para evaluar sistemas grandes de estudio de miles o millones de átomos tales como proteínas, ácidos nucleicos y ADN debido al elevado costo y capacidad computacional que requerían los cálculos cuánticos, que a la actualidad ya evalúan pequeños sistemas de un par de cientos de átomos; a su vez, se ha desarrollado métodos híbridos (QM/MM) que combinan



ambos, tanto la mecánica clásica como la mecánica cuántica, para el estudio de sistemas grandes y la aplicación de métodos cuánticos a las regiones de interés.

Entre las principales aplicaciones de la mecánica molecular está el cálculo de energía de un sistema molecular a partir de su estructura tridimensional y mediante la cual se puede encaminar el estudio de propiedades asociadas; para este tipo de cálculos se tiene dos grandes herramientas que son la minimización de energías a partir de una optimización geométrica de la estructura en estudio y la simulación de dinámica molecular del sistema a través del tiempo y sometido a una temperatura constante.

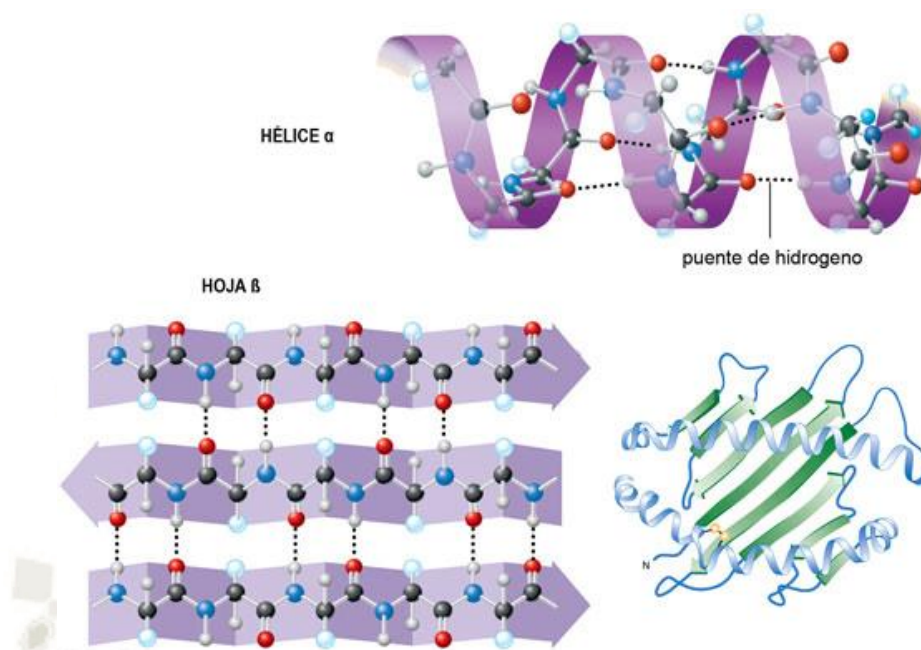
### 2.6.3. Modelamiento de Proteínas:

Modelamiento hace referencia a una predicción de la disposición espacial de los átomos de una proteína intrínsecos a su secuencia de aminoácidos, tiene como parte de sus objetivos el estudio de su función molecular, el diseño de ligandos y el cálculo de sus afinidades. Cada proteína presenta una secuencia particular de aminoácidos o residuos que se disponen en forma de cadena, siendo cada uno de estos diferente al tener ciertas propiedades como polaridad, hidrofobicidad y carga; siendo en total 20 los aminoácidos presentes en las proteínas y siendo estos los que definen sus propiedades físico-químicas y características tridimensionales<sup>85, 89-91</sup>.

Dentro de las características tridimensionales de una proteína, se destaca su estructura secundaria que proviene del plegamiento entre los residuos cercanos de la cadena polipeptídica y se debe principalmente a la formación de enlaces puente de hidrogeno entre el átomo de oxígeno del grupo carbonilo (-CO-) y el átomo de hidrogeno del grupo amino (-NH-) de los carbonos de cada residuo, siendo los principales las  $\alpha$ -hélices y las láminas  $\beta$  (Ver Figura 2.10)<sup>92, 93</sup>.

Las  $\alpha$ -hélices son conformaciones espirales con sentido horario que usualmente involucran a 13 residuos de la cadena, presentando 3.6 residuos por cada vuelta y siendo sus ángulos diedros  $\psi$  (psi) y  $\phi$  (phi)  $-60^\circ$  y  $-45^\circ$  respectivamente<sup>91, 94</sup>.

Las láminas  $\beta$  por su parte se forman a partir del posicionamiento paralelo de dos cadenas de aminoácidos en la misma proteína, en un rango de 3 a 10 residuos y presenta la característica de que los grupos -R de los aminoácidos que la conforman, están posicionados por encima y por debajo de la estructura para no generar un impedimento estérico.



**Figura 2.10:** Estructura secundaria de las proteínas,  $\alpha$ -hélices y láminas  $\beta$

#### 2.6.4. Campos de Fuerza para Modelamiento Molecular:

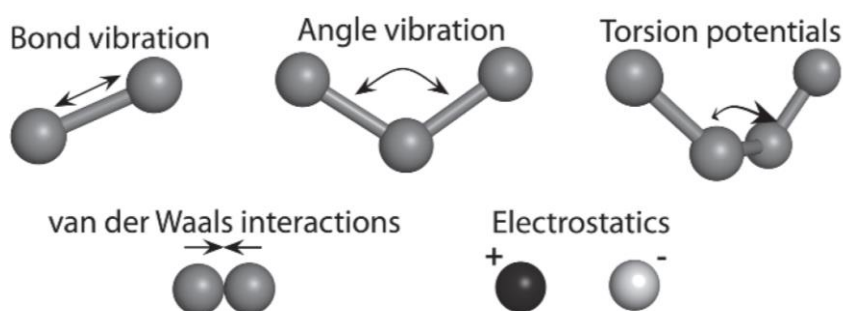
En química, se denomina campo de fuerza al conjunto de parámetros y ecuaciones necesarias para llevar a cabo simulaciones de mecánica molecular. Se basan en el cálculo de la energía potencial como suma de las energías de los enlaces, de los ángulos de enlace, los ángulos de torsión así como de energías no enlazantes de las interacciones de Van der Waals y electrostáticas descritas por la Ley de Coulomb según la ecuación (2.2)<sup>95-97</sup>.

$$E_p = E_{bonds} + E_{angle} + E_{dihedral} + E_{non-bonded} \quad (2.2)$$

Los tipos de interacciones (Ver Figura 2.11) se clasifican según:

- Interacciones enlazantes: presenta enlaces químicos que son intermediarios en la unión de dos o más átomos, engloba los términos de enlaces covalentes, ángulos, diedros o torsiones.
- Interacciones no enlazantes: los átomos no se encuentran unidos por enlaces químicos pero existen energías electrostáticas y de Van der Waals entre estos<sup>97</sup>.





**Figura 2.11:** Tipos de interacciones descritas en los campos de fuerza, tanto enlazantes como no-enlazantes. Nota. Fuente: Kukol, A. (2008) *Molecular Modeling of Proteins*, Hatfield, UK

En la actualidad, se han desarrollado varios campos de fuerza entre los cuales destacan aquellos que describen macromoléculas como OPLS-AA<sup>98, 99</sup>, AMBER<sup>95</sup>, GROMOS<sup>100</sup> y CHARMM<sup>101</sup>; cada uno de ellos contiene 4 archivos principales que describen en su mayoría los parámetros particulares de cada uno según el método de cálculo y precisión usados, siendo estos archivos:

- Atomtypes.atp: Describe las características asignadas a cada átomo así como su nomenclatura en este campo de fuerza, su carga y el tipo de molécula del que forma parte.
- Aminoacids.rtp: Describe la topología de cada residuo, describiendo el nombre asignado, su carga y grupo de carga de cada átomo de cada residuo además de los átomos con los que se enlaza formando enlaces, ángulos y diedros.
- Ffbonded.itp: Describe los diversos tipos de átomos que reconoce el campo de fuerza que forman enlaces, ángulos o diedros entre sí, y a su vez describe sus valores en distancias, grados del ángulo formado.
- FFnonbonded.itp: Describe la nomenclatura asignada a cada átomo, su tipo de enlace, su masa molecular y su carga que en conjunto van a describir las interacciones no enlazantes en el sistema.

### 2.6.5. Minimización de Energía:

El objetivo de esta herramienta es encontrar las coordenadas para una estructura molecular, las cuales han de presentar un mínimo local de energía potencial, ya que en estos estados de menor energía la estructura es más estable, y son estas estructuras las



que son usadas para el estudio de sus funciones en procesos químicos, físicos y biológicos<sup>91</sup>.

Se hace uso de las ecuaciones de un campo de fuerza, el cual describe las propiedades físicas y superficies de energía potencial de la molécula, teniendo métodos derivativos de primer y segundo orden que se basan en una gradiente de energía<sup>98</sup>.

La primera derivada de la energía potencial con respecto a las coordenadas cartesianas busca un valor 0, el cual nos indica un punto de inflexión que se pretende fuese un mínimo global de energía pero estos cálculos tienen usualmente a dirigir al sistema hacia un punto mínimo de energía cercano a las coordenadas de partida o mínimos locales. La derivada de segundo orden, es la que revela si los puntos de inflexión encontrados por la derivada de primer orden son cóncavos o convexos al obtener valores positivos o negativos respectivamente, siendo lo anhelado un valor positivo que indica un mínimo.

Entre los métodos derivativos más conocidos tenemos: steepest descent, gradiente conjugado y Newton-Raphson; siendo el método steepest descent o de fase estacionaria el que ha de usarse en este estudio, siendo un algoritmo que aproxima en primera instancia al sistema hacia un mínimo local cercano pese a estar alejado de este al modificar las coordenadas cartesianas y evaluar los cambios energéticos que están asociados a las perturbaciones de los grados de libertad geométricos, siendo un proceso repetitivo hasta que estas perturbaciones alcancen un valor dentro de los límites permitidos que han sido predefinidos<sup>102</sup>.

#### 2.6.6. Dinámica Molecular:

Es una técnica de simulación molecular que permite la interacción de átomos y moléculas en un periodo determinado de tiempo en la escala de los nanosegundos, se basa en el hecho de que toda la materia está constituida de partículas, las cuáles están en constante movimiento e interacción<sup>91</sup>. Las teorías que fundamentan esta metodología provienen de la física, química y matemática, resaltando la segunda ley del movimiento de Newton (Ver Ecuación 2.3)<sup>103</sup>.

$$F_i = m_i \frac{\partial^2 r_i}{\partial t^2} \quad (2.3)$$

La integración de esta ecuación nos brinda las coordenadas de los átomos de un sistema en el tiempo a la vez que la trayectoria que va desarrollando.

Existen diversos programas para llevar a cabo este tipo de simulaciones como Gromacs, CHARMM, AMBER, NAMD y OpenMM, y a su vez un gran número de campos de fuerza. Dentro de estos software, destaca Gromacs (Gröningen Machine for Chemical Simulations) debido a su alta capacidad de cálculo con alta precisión, por su alta velocidad de cálculo y por ser un programa de licencia libre el cual permite el acceso a estas herramientas de dinámica molecular<sup>104</sup>.

La dinámica molecular a su vez hace uso de colectivos o principios físico, en este estudio se hizo uso del colectivo canónico o NVT<sup>105</sup>, donde se trabaja a un volumen y masa constantes; la temperatura no es constante pero sí su promedio al hacer uso del termostato de Nose-Hoover que es un método determinista para el control de temperatura mas no de energía, dentro de las cuales destacan la energía potencial, energía cinética y las interacciones no enlazantes como las de Van der Waals descritas por el potencial de Lennard Jones y las fuerzas electrostáticas descritas por la ley de Coulomb<sup>97</sup>.

#### **2.6.7. Cálculo de Energía Libre:**

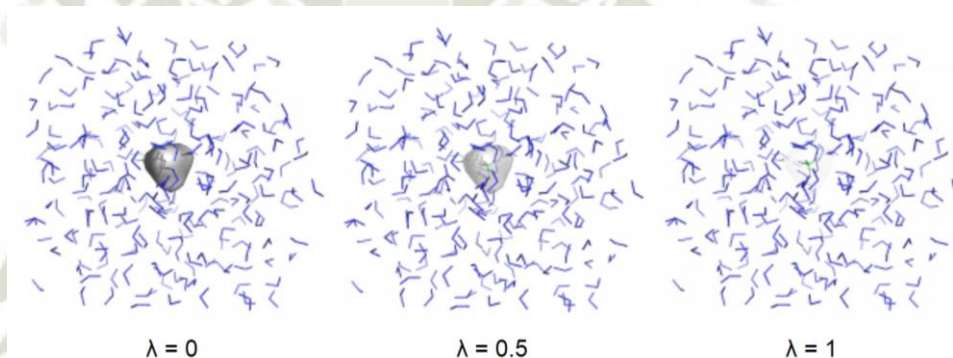
La energía libre es muy importante cuando evaluamos un sistema desde un punto de vista termodinámico, ya que muchas otras propiedades físicas o química dependen de este valor del sistema<sup>106</sup>. Puede ser vista como una propiedad estadística que calcula la probabilidad de encontrar a un sistema en cierto estado, un cálculo de alta precisión para este tipo de energías requiere el evaluar todo el espacio en el que se desarrolla el sistema, siendo un proceso de alto costo y de baja precisión con las herramientas actuales; sin embargo, una variante es el cálculo de la diferencia de energía libre entre dos estados de un mismo sistema<sup>91, 107</sup>.

La energía libre puede ser expresada según dos funciones: función de Helmholtz ( $\Delta F$ ) o función de Gibbs ( $\Delta G$ )<sup>108</sup>, esta última requiere de un colectivo NPT (isobárico – isotérmico) donde el número de partículas sea constante al igual que la presión y temperatura<sup>103</sup>.

Una de las aplicaciones más comunes es el cálculo de la energía libre de unión de pequeñas moléculas a una macromolécula o receptor que usualmente son proteínas, lo cual envuelve el acoplamiento o desacoplamiento de ambas obteniendo como resultado el cambio de energía libre. Para este tipo de cálculos de desacoplamiento,

son las interacciones no enlazantes las que pueden ser modificadas sin afectar el sistema, siendo así restringidas las interacciones de Van der Waals e interacciones Coulombicas.

Gromacs es una paquetería de herramientas que nos ayudan también en este tipo de cálculos, haciendo uso del módulo BAR (Bennet Acceptance Ratio)<sup>109</sup> que calcula las diferencias de energía libre del estado A al estado B ( $\Delta G_{AB}$ ) con un parámetro de desacoplamiento  $\lambda$  (Ver Figura 2.12), el cual indica el nivel de cambio entre ambos estados al restringir las interacciones no enlazantes con una equidistancia del espaciamento  $\lambda$  de 0 (A) a 1 (B) y valores de espaciado intermedios de 0.025 para nuestro estudio.



**Figura 2.12:** Desacoplamiento de interacciones no enlazantes para el cálculo de la variación de la energía libre de Gibbs haciendo uso del parámetro de desacoplamiento  $\lambda$

#### 2.6.8. Acoplamiento Molecular:

El acoplamiento molecular es una herramienta computacional esencial para diversas aplicaciones como cristalografía de rayos X, diseño de fármacos, estudios de mecanismos químicos, etc. Dentro de las técnicas reportadas se distinguen dos categorías: métodos de emparejamiento y simulaciones de acoplamiento.

Los métodos de emparejamiento (Matching Methods) generan sitios activos basándose en sitios donadores de enlaces puente de hidrogeno y sitios sin restricciones estéricas a los cuales se les intenta acoplar, por su forma estructural, una molécula rígida en sus enlaces, ángulos y conformación con una función de interés.

La segunda técnica para el acoplamiento molecular se basa en el uso de un ligando el cual va a ser sometido aleatoriamente a una serie de cambios traslacionales, de



orientación y de conformación hasta encontrar una estructura que presente afinidad por el sitio activo o bolsillo, sin incumplir en normas conformacionales o de impedimento estérico.

En este estudio se usaron dos programas para llevar a cabo el acoplamiento molecular, AutoDock<sup>110, 111</sup> y ClusPro<sup>112, 113</sup>, uno para la formación de complejos proteína – ligando y el otro en formación de complejos proteína – proteína respectivamente. AutoDock está diseñado para predecir la unión de un ligando pequeño con un receptor o proteína que tenga una estructura tridimensional conocida; presenta una gran velocidad de cálculo con alta precisión y provee de una función de score para hallar la energía libre según la afinidad de unión entre proteína y ligando. Por su parte, ClusPro es un servidor cuya función es el acoplamiento o docking de complejos rígidos proteína – proteína requiriendo únicamente sus coordenadas cartesianas basados en una función de correlación denominada transformada rápida de Fourier (FFT)<sup>114-116</sup>, puede desarrollar el acoplamiento basándose en seis funciones de energía diferentes que van a depender de las proteínas en estudio, siendo importante tener en cuenta los objetivos de funcionamiento celular u organización.

Además, para una mejor precisión en el acoplamiento especialmente en un sistema proteína-proteína se hace uso de FireDock<sup>117</sup> un servidor para el refinamiento de estructuras que permite corregir estas estructuras acopladas mediante una reevaluación de sus energías de interacción tomando en cuenta esta vez las cadenas laterales y átomos del *backbone* como moléculas flexibles.

#### 2.6.9. Validación de Estructuras:

Varias son las herramientas desarrolladas para la validación de conformaciones o modelos estructurales mayormente de proteínas en estudio, el principal criterio de evaluación es la estereoquímica de cada modelo que ha de presentar distancias interatómicas asignadas para cada tipo de enlace según los átomos interactuantes, además los ángulos de enlace que han de estar dentro del rango permitido. Dentro de las herramientas disponibles y usadas este estudio están:

- Desviación Media Cuadrática (RMSD):

Evalúa el grado de similitud entre modelos estructurales de dos proteínas entre pares de átomos equivalentes pertenecientes al *backbone* de estas estructuras superpuestas.

El cotejo cuantifica la diferencia entre una proteína nativa o estructura cristalizada y el plegamiento que va asumiendo esta proteína al evolucionar en el tiempo<sup>91</sup>.

- Fluctuación Media Cuadrática (RMSF):

Cuando un sistema al evolucionar por el tiempo define una posición promedio al alcanzar su estabilidad, he aquí donde se evalúa la fluctuación que sus átomos presentan respecto a esta posición local para determinar si las partículas fluctúan debido únicamente a difusión o si hay alguna fuerza que contribuye a ello.

- Radio de Giro (RoG):

Es la distancia radial de un punto donde se asume se encuentra la masa del sistema más concentrada hacia un punto alejado en el sistema que se asume son los átomos de un residuo de una proteína que se alejan o acercan dependiendo del proceso de plegamiento.

- Diagrama de Ramachandran:

Evalúa los ángulos diédricos  $\psi$  (psi) y  $\phi$  (phi) que forman los carbonos alfa (C- $\alpha$ ) de las secuencia de residuos en una proteína. Siendo un ángulo phi aquel que define el giro alrededor de un enlace entre un carbono alfa y un carbono del residuo posterior, mientras que un ángulo psi es aquel que define el giro alrededor de un enlace entre un carbono alfa y un nitrógeno del residuo anterior. Este diagrama define si la conformación tridimensional adoptada por la proteína respecto a estos ángulos es permitida<sup>94</sup>.

# CAPÍTULO 3

## 3. MATERIALES Y MÉTODOS:

### 3.1. Equipos y Software:

La parte computacional requirió de una *workstation* con un *hardware* que se describe a continuación:

- Intel® Core™ i7 5ta Generación Serie X 3.3GHz x 16
- 64GB Memoria RAM
- 11TB HDD
- 2 tarjetas gráficas NVIDIA GTX980 4GB

El entorno virtual tuvo como base el sistema operativo Linux, distribución Debian 8.0 y requirió de los siguientes servidores y programas:

- **Base de Datos Protein Data Bank (PDB):**

Disponible en <http://www.rcsb.org>, Protein Data Bank (PDB) es una base de datos que recopila la estructura tridimensional de estructuras macromoleculares como proteínas y ácidos nucleicos obtenidas mediante técnicas de cristalografía de rayos X y/o resonancia magnética nuclear. Los archivos con extensión “.pdb” presentan la información completa sobre los autores, fuente de la macromolécula en estudio, publicación de referencia, descripción de la estructura incluidos los ligandos usados para su estabilización previo a la cristalización y el método experimental desarrollado.

- **Base de Datos PubChem:**

Disponible en <https://pubchem.ncbi.nlm.nih.gov>, PubChem es una base de datos parte del Instituto Nacional de Salud (NIH), que colecta información de las estructuras químicas, propiedades tanto químicas como físicas, patentes, hojas de seguridad, información de toxicidad, etc.; de moléculas como drogas, nucleótidos, carbohidratos, lípidos y pequeños péptidos. Los principales contribuyentes a esta base de datos son



universidades, compañías farmacéuticas, agencias gubernamentales de investigación y editores de revistas de investigación.

- **Base de Datos NCBI:**

Disponible en <https://www.ncbi.nlm.nih.gov>, el Centro Nacional para la Información Biotecnológica (NCBI) es parte de la Biblioteca Nacional de Medicina de los Estados Unidos de América, cuya misión radica en el desarrollo de sistemas para el almacenamiento, estudio y análisis de información sobre datos de biología molecular, genética, bioquímica; administra bases de datos como GenBank que es referente de secuencias genómicas, un índice de artículos científicos en PubMed aparte de ofrecer herramientas bioinformáticas para el análisis de secuencias de ARN, ADN y proteínas.

- **Gromacs 2016:**

Disponible en <http://www.gromacs.org>, es un paquete computacional para la simulación de dinámica molecular, teniendo como base las ecuaciones Newtonianas de movimiento para sistemas que van de cientos a millones de partículas. Constantemente actualiza su código al introducir algoritmos de optimización haciendo un cálculo de alto rendimiento y de alta velocidad para el cálculo de interacciones enlazantes y no enlazantes, siendo estas últimas las predominantes en sistemas de moléculas como proteínas, lípidos y ácidos nucleicos.

- **Gaussian 09:**

Es un paquete computacional para el desarrollo de amplio uso en estudios de química teórica, se basa en las leyes fundamentales de la mecánica cuántica como lo es la ecuación de Schrödinger basada en la teoría de orbitales moleculares y el cálculo de la función de onda. Predice energías, coordenadas cartesianas, carga y multiplicidad, frecuencias vibracionales y diversas propiedades atómicas y moleculares.

- **GaussView 6:**

Es una interfaz gráfica que complementa al paquete computacional Gaussian 09, permite construir o importar estructuras moleculares de interés; puede lanzar, monitorear y controlar los cálculos hechos en Gaussian y ver los resultados gráficamente.

▪ **Chimera 1.12:**

Disponible en <https://www.cgl.ucsf.edu/chimera>, es un programa para la visualización y análisis de estructuras moleculares que es capaz de generar imágenes de alta calidad y pequeñas películas que describen las trayectorias y ensambles conformacionales de una dinámica molecular; a su vez, incluye mapeo de densidades, alineamiento de secuencias, resultados de acoplamiento molecular

▪ **Autodock 4:**

Disponible en <http://autodock.scripps.edu>, es un conjunto de herramientas para el acoplamiento molecular, está diseñado para predecir la vía por la cual moléculas de tamaño pequeño, como drogas o sustratos, pueden interactuar y/o enlazarse a una molécula de mayor tamaño con una conformación tridimensional conocida como proteínas. Consta de dos programas principales: autogrid que se encarga de calcular mapas de grillas donde describen los valores máximos y mínimos de energía de la grilla; y el segundo programa es autodock que desarrolla con gran efectividad el acoplamiento molecular entre el ligando y el mapa de grillas que describen la proteína objetivo.

▪ **ClusPro 2.0:**

Disponible en <https://cluspro.bu.edu>, es un servidor online que realiza acoplamiento molecular proteína-proteína, siendo necesario archivos que tengan las coordenadas tridimensionales de estas proteínas o sus códigos PDB. Luego, evalúa en etapas aquellos complejos formados con mejor complementariedad de superficies, energías de electrostática y desolvatación para un agrupamiento o clustering final.

▪ **PockDrug Server:**

Disponible en <http://pockdrug.rpbs.univ-paris-diderot.fr>, es un servidor online cuya función principal es la predicción de la capacidad de un bolsillo de proteína para unirse con un fármaco y la afinidad que pueda presentar esta unión.

▪ **MolProbity:**

Disponible en <http://molprobity.biochem.duke.edu>, es un servidor que permite validar estructuras proteicas y de ácidos nucleicos en diagramas de Ramachandran al detectar valores atípicos en este debido usualmente a cadenas laterales de proteínas ramificadas y/o pliegues incorrectos de los azúcares.

▪ **TPPMKTOP:**

Disponible en <http://erg.biophys.msu.ru/wordpress/archives/32>, es un servidor online que permite generar la topología necesaria para una nueva molécula en un campo de fuerza como lo es OPLS-AA. Este servidor contiene los parámetros e información del campo de fuerza antes mencionado con lo cual reconoce los átomos, coordenadas, enlaces y nombre de los elementos de la molécula de interés y les atribuye las características, nombres y descripción propia del campo de fuerza.

▪ **Lipinski Rule of Five:**

Disponible en <http://www.scfbio-iitd.res.in/software/drugdesign/lipinski.jsp>, es un servidor online que permite evaluar la similitud de una molécula a un fármaco basándose en reglas cuyos valores de aceptación en diferentes unidades son múltiplos de 5. Entre sus principios está evaluar el tamaño y lipofilidad de la molécula.

▪ **FireDock:**

Disponible en <http://bioinfo3d.cs.tau.ac.il/FireDock/>, es un servidor online para el refinamiento de complejos rígidos de proteína-proteína o proteína-ligando, brindando a su vez un nuevo score de unión y energías de interacción.

▪ **LigPlot 4.5.3:**

Disponible en <https://www.ebi.ac.uk/thornton-srv/software/LIGPLOT/>, LigPlot es un programa para la generación de diagramas esquemáticos que dilucidan la interacción proteína-ligando mediadas por enlaces puente de hidrogeno y/o por contacto hidrofóbico de un archivo de coordenadas con extensión .pdb.



### 3.2. Metodología:

#### 3.2.1. Obtención de las Estructuras Tridimensionales de las Proteínas de Interés:

En la base de datos del Protein Data Bank (PDB), se realizó una búsqueda por nombre de las proteínas: Profilina Humana II (PFN2) e Histona Desacetilasa 1 (HDAC1), estando ambas presentes bajo dominio público para su uso y manejo libre, y presentan a su vez la metodología que fue empleada para su dilucidación por técnicas de cristalografía de rayos X. Cada proteína presenta un código PDB o PDB ID, siendo “1D1J” el código asignado para la profilina humana II, mientras que “4BKX” es el código asignado para la enzima Histona Desacetilasa 1.

Se descargaron ambos archivos con una extensión .pdb; luego se procedió a su lectura en el software de editor de textos predeterminado de Linux: Gedit, con el cual pudimos observar las coordenadas de cada átomo de toda la secuencia de aminoácidos de cada proteína y los componentes y/o ligandos adjuntos que se necesitaron para su cristalización, siendo estos eliminados al ser de importancia nula para la investigación.

#### 3.2.2. Obtención de las Estructuras Tridimensionales y Cargas de Hirshfeld de los Flavonoides de Nopal:

Los compuestos flavonoides reportados en flores y frutos de *Opuntia ficus-indica* según las investigaciones a cargo de los doctores M. De Leo et. Al<sup>61</sup> y Jung Wha Kim et. Al<sup>72</sup> respectivamente, fueron descargados de la base de datos PubChem, obteniendo archivos con sus coordenadas tridimensionales. Fueron 11 los compuestos descargados tomando como base a Quercetina, Isorhamnetin, Kaempferol y sus derivados glucósidos.

Los archivos descargados presentaron una extensión .sdf; se hizo uso del programa GaussView para poder visualizar estas pequeñas moléculas y poder guardarlas con una extensión .gjf, a partir de la cual pudimos realizar una optimización geométrica y luego realizar un single-point mediante el uso del software Gaussian 09, un paquete computacional basado en leyes de mecánica cuántica, fundamentales para el cálculo de la función de onda y/o densidad electrónica.

Para el primer paso descrito anteriormente, haciendo uso de un editor de textos se incluyó en la cabecera del archivo con extensión .gjf los siguientes datos:

```
%chk=quercetin.chk
%Mem=64GB
%Nproc=16
# B3LYP/TZVP Opt=(tight, cartesian) Freq

Quercetina

0 1

C          -2.46800000    0.45700000   -0.19200000
C          -1.61500000   -0.22700000   -1.25900000
C           0.43900000    0.30800000   -0.00400000
C           1.92400000   -0.03900000    0.21200000
C          -0.21900000    0.97700000    0.99100000
```

**Figura 3.1:** Comandos del archivo para optimización geométrica o de frecuencias

En la figura 3.1. se observan los comandos para este cálculo que fueron descritos en la cabecera de nuestros archivos, de donde “%chk=nombredelflavonoide.chk” es un archivo *check-point* o punto de control que va guardando los cálculos que va realizando el programa y que a su vez sirve para un relanzamiento del cálculo en caso de un término imprevisto; los siguientes “%Mem=64GB” y “%Nproc=16” describen la memoria RAM y el número de procesadores pertenecientes al computador en uso; mientras que “# B3LYP/TZVP Opt=Freq” son los comandos que va a ejecutar el software, siendo B3LYP una funcional híbrida entre cálculo de función de onda y la teoría funcional de la densidad (DFT), TZVP representa las frecuencias vibracionales, Opt representa la optimización geométrica y Freq el cálculo de las frecuencias luego de la optimización. El nombre “Quercetina” representa el nombre de la estructura trabajada en este cálculo y finalmente los números “0 1” representan la carga y multiplicidad respectivamente, siendo importante resaltar que lo óptimo es trabajar con moléculas con carga neutra. Este archivo fue guardado con la extensión .gjf para proceder con la optimización cuántica, siendo necesario abrir una ventana del terminal de Linux donde se ejecutó el siguiente comando:

*g09 nombredelflavonoide.gjf > nombre del flavonoide.log &*

Terminado el cálculo, obtuvimos archivos con las extensiones .chk y .log; el archivo con la extensión .chk fue convertido a un archivo con extensión .xyz para lo cual se ejecutó el siguiente comando en la terminal:

```
newzmat -ichk nombredelflavonoide.chk -oxyz nombredelflavonoide.xyz
```

Este archivo .xyz final contiene las coordenadas tridimensionales finales optimizadas de la molécula para proseguir el cálculo del single-point y hallar las cargas asignadas a cada átomo, se procedió a convertir el archivo .xyz a un archivo .gjf en el terminal con el comando:

```
mv nombredelflavonoide.xyz nombredelflavonoide.gjf
```

Al igual que en el caso anterior, se cambiaron los comandos de cabecera en estos archivos, pero con unas pequeñas variaciones como se ve en la figura 3.2, donde “# B3LYP/TZVP Pop=Hirshfeld” es el nuevo comando para el cálculo del single-point donde Pop refiere a Population, término que define a los átomos como una población y al analizarlos se pudo obtener datos de sus orbitales moleculares y asignarles cargas atómicas que en este caso serán las cargas atómicas de Hirshfeld y serán usadas más adelante para la generación de campos de fuerza de los compuestos flavonoides en Gromacs.

```
%chk=quercetin.chk
%Mem=64GB
%Nproc=16
# B3LYP/TZVP Pop=Hirshfeld

Quercetina

0 1

C          -2.46800000    0.45700000   -0.19200000
C          -1.61500000   -0.22700000   -1.25900000
C           0.43900000    0.30800000   -0.00400000
C           1.92400000   -0.03900000    0.21200000
C          -0.21900000    0.97700000    0.99100000
```

**Figura 3.2:** Comandos del archivo para el cálculo de cargas de Hirshfeld

Para realizar este último cálculo se requirió abrir una ventana del terminal de Linux donde se ejecutó el siguiente comando:

```
g09 nombredelflavonoide.gjf > nombre del flavonoide.log &
```

Terminado el cálculo, obtuvimos archivos con las extensiones .chk y .log; el archivo con la extensión .log contiene las cargas asignadas a cada átomo.



### 3.2.3. Cinco Reglas de Lipinski para Compuestos Flavonoides:

Las cinco reglas de Lipinski refieren a la similitud que ha de presentar una molécula o compuesto químico con cierta actividad farmacológica a un medicamento o droga con actividad tras ingesta oral en humanos; basándose en que los medicamentos han de ser relativamente pequeños y con propiedades lipofílicas. Para tal motivo, se requirió de los archivos .pdb de cada compuesto flavonoide ya optimizados, que fueron enviados al servidor que se encargó de procesar estos archivos y como resultado se obtuvo una tabla con los valores que se les asigna a cada uno, en esta tabla se evalúan 4 criterios como son el número de donadores de enlaces puente de hidrogeno, el número de átomos (nitrógeno u oxígeno) receptores de enlaces puente de hidrogeno, la masa molecular y el coeficiente de reparto o coeficiente de partición octanol-agua el cual determina la solubilidad de un compuesto en una mezcla bifásica formada por dos disolventes inmiscibles (octanol y agua), habiéndose reportado los logaritmos del coeficiente calculado para regular el rango de valores.

### 3.2.4. Declaración de Nuevos Parámetros en Campos de Fuerza OPLSAA:

Entre las estructuras que fueron descargadas se encontraron 2 proteínas nativas y 11 compuestos flavonoides los cuales han de someterse a procesos descritos por la mecánica molecular, motivo por el cual se realizaron modificaciones en las bases de datos del campo de fuerza OPLSAA, el cual fue elaborado para moléculas orgánicas y péptidos, y que lo hace un campo de fuerza ideal para este trabajo.

#### 3.2.4.1. *Profilina Humana II:*

La secuencia obtenida de la estructura cristalizada de esta proteína reporta 138 aminoácidos, todos los cuales están parametrizados en el campo de fuerza OPLS-AA, por ello solo fue necesario corroborar su validez en Gromacs 2016.

#### 3.2.4.2. *Histona Desacetilasa 1:*

La secuencia obtenida de la estructura cristalizada de esta proteína reportó 369 aminoácidos y un átomo de zinc que es parte de su centro catalítico, el archivo HDAC1.pdb fue cargado al servidor TPPKMKTOP para obtener los parámetros

necesarios del ion Zinc para el campo de fuerza OPLS-AA que a su vez fue declarado con una carga +2.

### 3.2.4.3. *Compuestos Flavonoides:*

Son moléculas pequeñas que no sobrepasan los 80 átomos compuestas por tan solo 4 elementos químicos como lo son el carbono, hidrógeno, nitrógeno y oxígeno, dispuestos de manera diversa siendo algunas estructuras isómeros.

Fueron optimizados mediante el software Gaussian y convertidos a archivos .xyz, para finalmente convertirlos a archivos .pdb ejecutando el siguiente comando en la terminal:

```
openbabel -ixyz nombredelflavonoide.xyz -opdb nombredelflavonoide.pdb
```

Tras ser creados los archivos pdb, fueron abiertos con un editor de textos como Gedit, donde se identificó una denotación de 3 letras: LIG para todos los átomos lo cual se reemplazó con el nombre o abreviación de 3 letras designadas para cada molécula para su diferenciación; siguiendo con el proceso de generación de topología, los archivos .pdb fueron sometidos al servidor TPPKMKTOP para obtener la topología necesaria para incluir estas moléculas en la raíz del campo de fuerza. Los datos declarados de estas moléculas son los nombres que recibieron los átomos según su conformación, los enlaces, ángulos y diedros que forman entre ellos y las cargas de Hirshfeld calculadas para una carga global neutra mediante herramientas de mecánica cuántica.

### 3.2.5. Minimización de Energías y Dinámica Molecular de las Proteínas:

El software necesario para llevar a cabo este tipo de cálculos fue Gromacs. El primer paso en el proceso de obtención de una molécula en equilibrio, con carga neutra y en fase solvente fue la minimización de energía. Para iniciar, se requirió de los archivos con extensión .pdb, los cuales contienen las coordenadas tridimensionales de los átomos de las proteínas que fueron convertidos a un archivo con extensión .gro con el uso del comando *editconf* que a su vez define una caja que rodea a la proteína limitando así el área dentro de la cual se ha de trabajar. Se procedió a agregar un solvente al medio que tuvo que calzar en la caja creada hasta completar el volumen total de esta, el comando utilizado fue *solvate*. Seguido se ejecutó el comando *pdb2gmx* que generó

archivos con las extensiones .top, .gro, y .itp tomando como base los parámetros del campo de fuerza OPLS-AA y el modelo de aguas SPC, siendo el archivo .top el principal ya que almacena toda la información de la molécula incluida su secuencia de aminoácidos, sus cargas, el número de moléculas de solvente y iones (en caso la carga no sea neutra). El siguiente comando ejecutado fue *grompp*, con el cual se creó un archivo .tpr que presentaba los datos de trayectoria inicial de la molécula, que son necesarios para iniciar el cálculo de minimización, los parámetros fueron provistos por un archivo de entrada minima.mdp (Ver Apéndice 1).

Por último, el comando *mdrun* que es el iniciador del cálculo a partir del desplazamiento e interacción de los átomos conforme a los parámetros establecidos en el archivo minima.mdp que tiene como método principal Steepest Descent, el cual busca que la conformación alcance un mínimo de energía; este comando generó archivos .cpt, .log, .edr, .gro, .xtc y .trr; siendo este último el de mayor tamaño ya que almacena todas las trayectorias finales de cada átomo a través del tiempo de cálculo que va en la escala de los nanosegundos.

Terminado el proceso de minimización, se procedió a realizar la dinámica molecular la cual explora la relación entre estructura, movimiento y función de las proteínas de interés basándose en el hecho que las moléculas no son modelos estáticos sino por el contrario son máquinas en movimiento, teniendo protagonismo los átomos con sus interacciones, torsiones y saltos. Se basan en leyes del movimiento de Newton que son básicas para la explicación del movimiento de los cuerpos según la mecánica clásica. Se procedió de igual manera que en el caso anterior pero tomando como archivo inicial el último archivo .gro generado, el cual se sometió a los comandos *pdb2gmx*, *grompp* y *mdrun*, pero a diferencia del caso anterior, en vez de usar el archivo minima.mdp se utilizó el archivo nvt.mdp (Ver Apéndice 2), que contiene parámetros diferentes que se basan en un colectivo canónico NVT, cuyo algoritmo se basa en mantener un número de partículas invariable, un volumen constante, una temperatura promedio a través del tiempo haciendo uso del termostato de Nosé-Hoover y por último evalúa las energías que generan los átomos de la molécula entre sí y con su medio interactuante.



### 3.2.6. Validación de la Estructura de las Proteínas:

Se hizo uso del servidor MolProbity, teniendo que cargar en este los archivos PFN2.pdb y HDAC1.pdb respectivamente que fueron creados a partir del último archivo .gro con el siguiente comando:

```
gmx editconf -f proteína.gro -o proteína.pdb
```

Terminado el cálculo, el servidor proporcionó un diagrama conocido como diagrama de Ramachandran en un formato .pdf.

Por otra parte, dentro de la paquetería de Gromacs, existen comandos para calcular la desviación media cuadrática (RMSD), la fluctuación (RMSF) y el radio de giro (GYRATE) de las proteínas durante las simulaciones de dinámica molecular; requieren de los archivos de trayectoria .tpr y .trr, y fueron ejecutadas con los siguientes comandos:

```
gmx rms -s proteína.tpr -f proteína.trr -o rmsd.xvg
```

```
gmx rmsf -s proteína.tpr -f proteína.trr -o rmsf.xvg
```

```
gmx gyrate -s proteína.tpr -f proteína.trr -o radiodegiro.xvg
```

### 3.2.7. Predicción de Bolsillos de Drogabilidad:

El servidor online PockDrug fue un valioso recurso para evaluar los bolsillos o *pockets* de la profilina humana II que son sitios de reactividad o drogabilidad que presenta la proteína tomando en cuenta parámetros de hidrofobicidad de su superficie; fue necesario cargar en el servidor el archivo PFN2.pdb que contenía la información de los aminoácidos que componen esta proteína así como la configuración espacial de cada átomo de esta. El resultado fueron archivos .pdb que contienen las coordenadas tridimensionales de los sitios activos o *pockets* de la proteína que pueden ser visualizados con el software UCSF Chimera en forma de átomos o superficies, y además una tabla con los valores de porcentaje de drogabilidad de cada bolsillo.



The screenshot shows the PockDrug website interface. At the top, there is a navigation bar with links: Home, Druggability Prediction, Contacts, Download, and Links. The main header features the PockDrug logo and the text "Pocket Druggability Prediction". Below this, a section titled "Pocket Druggability Prediction" contains the instruction "Please submit your files in order to predict pocket(s) druggability". A form titled "Complete the form:" follows, with two radio button options: "Druggability Prediction using pocket:" and "Druggability Prediction using protein(s):". The second option is selected. Under "Protein(s) information\*", there are three input methods: "PDB ID(4 char) ex: 12AS", "or upload your PDB file" with a "Browse" button, and "or PDB ID list" with a "Browse" button. Each has an "Info" button. Under "Pocket estimation method(s)\*", there are two radio button options: "fpocket" and "prox", both with "Info" buttons. A slider for "Ligand proximity threshold between 4 Å and 12 Å\*" is set to 5.5. A note below the slider states: "Based on the literature, we recommend a ligand proximity threshold of 4 Å or 5.5 Å". At the bottom, there is a text field for "All information will be displayed on the following site. If you desire to receive your results by e-mail and to be notified once your job is finished, please enter your e-mail address: (optional field)". Finally, there are "Submit" and "Reset" buttons.

**Figura 3.3:** Pantalla de inicio del servidor de predicción de bolsillos o pockets: PockDrug

### 3.2.8. Docking o Acoplamiento Proteína - Ligando:

Para realizar el acoplamiento o *docking* de estas moléculas y formar complejos proteína-ligando se hizo uso del software AutoDock 4, el cual usa inicialmente archivos .pdb a partir de los cuales genera archivos .pdbqt que son propios del programa y con los cuales procede a ejecutar los comandos *autogrid* y *autodock* respectivamente. Para este proceso, fueron creadas 11 carpetas, cada una nombrada por cada compuesto flavonoide, dentro de las cuales se añadió el archivo .pdb de cada compuesto y el archivo PFN2.pdb de la proteína. Se procedió a usar AutoDock Tools, que es la interfaz gráfica para realizar los procesos de acoplamiento y análisis de resultados (ver Figura 3.4); en este se cargó el archivo PFN2.pdb como proteína receptora en cada caso y los archivos .pdb de cada flavonoide por separado como ligando, se procedió a guardar cada archivo con extensión .pdbqt y luego se ejecutó el comando *autogrid* que genera grillas imaginarias en la proteína siguiendo un patrón y de las cuales evalúa su energía máxima y mínima generando mapas de grillas. El

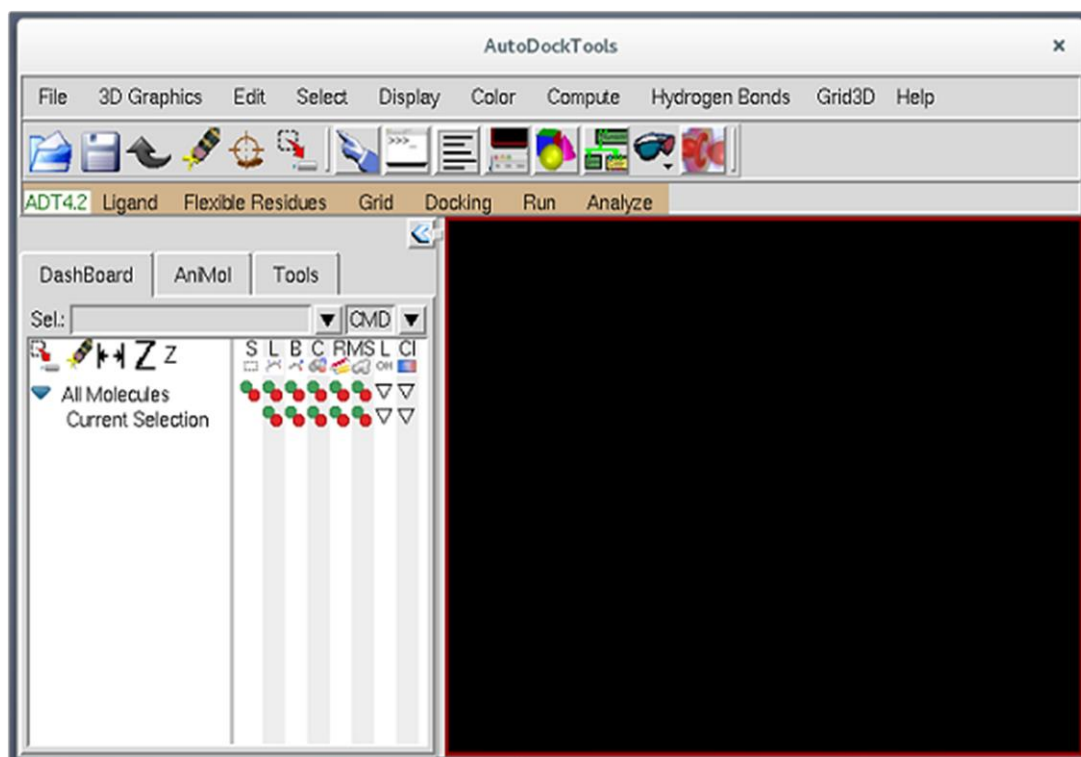
comando necesario para este proceso final fue ejecutado en la terminal de Linux con el siguiente comando:

```
autogrid -p proteina-ligando.gpf -l proteina-ligando.glg
```

Luego siguiendo en la interfaz gráfica, se escogió el número de corridas de acoplamiento que se van a realizar y el algoritmo apropiado. Se generó un archivo final .dpf necesario para ejecutar el comando *autodock* en la terminal:

```
autodock -p proteina-ligando.dpf -l proteina-ligando.dlg
```

El archivo generado .dlg contiene la información de todos los acoplamientos realizados entre la proteína y el ligando usado además de los valores energéticos de unión así como también las coordenadas de cada acoplamiento realizado, siendo el último paso guardar el complejo con mejor valor energético y geométrico como un archivo .pdb.



**Figura 3.4:** Software AutoDock Tools para el acoplamiento proteína – ligando



### 3.2.9. Dinámica de Complejos Proteína-Ligando:

Los complejos formados con el uso del software AutoDock 4 fueron guardados en archivo pdb, dichos complejos fueron sometidos a una minimización de energía, y una dinámica molecular para validar su interacción a través del tiempo con un sistema interactuante como un solvente para simular con mayor precisión un entorno real. Se hizo uso de los pasos ya descritos en el punto 3.2.5 para minimización de energía y dinámica molecular. Los archivos .gro finales se convirtieron a archivos .pdb.

### 3.2.10. Cálculo de Energía Libre con Colectivo NPT:

Los complejos que fueron sometidos a una corrida de dinámica molecular para validar su conformación y estabilidad, procedieron a ser evaluados con un colectivo isobárico-isotérmico o también conocido como colectivo NPT, con este tipo de colectivo se puede calcular la variación de la energía libre de Gibbs ( $\Delta G$ ) en un sistema termodinámico cerrado como este.

Los archivos .pdb finales, fueron modificados en el editor de textos Gedit, en el cual se añadió un terminador de cadena (diferenciando proteína de ligando), luego se ejecutó el comando *pdb2gmx* para generar archivos .gro, .top y más importante, generar archivos .itp específicos para la proteína como para el ligando con denominación Other2. Luego, al llevar a cabo el cálculo de energía libre, se produjo el desacoplamiento de la proteína y el ligando en una simulación, al ir restringiendo la energía de Van der Waals presente en la interacción de ambas, este proceso se lleva a cabo en 41 pasos según el descriptor  $\lambda$ , por ende se requiere de 41 scripts .sh (Ver Apéndice 3) que van a ejecutar específicamente cada uno de los 41 archivos .mdp (Ver Apéndice 4-5) que contengan los parámetros necesarios, adicionalmente se hace uso de un script run.sh (ver Apéndice 6) que ejecuta en orden numérico y serial de cada uno de los scripts anteriores con el siguiente comando en el terminal:

*nohup ./run.sh &*

Cada script creó una carpeta donde se generaron diversos archivos correspondientes a cada cálculo según avanzaba el desacoplamiento y se obtuvieron principalmente archivos con extensión .xvg en cada carpeta, en estas están reportadas las energías de Gibbs calculadas por cada *script* según la restricción de la energía de Van der Waals,

y finalmente con el siguiente comando se obtuvo un promedio que representa el resultado final de variación de la energía libre ( $\Delta G$ ) del sistema.

$$gmx\ bar\ -fmd*.xvg\ -o\ -oi\ -oh$$

### 3.2.11. Docking o Acoplamiento Proteína - Proteína:

Los acoplamientos fueron llevados a cabo en el servidor online ClusPro que es un software para el docking de complejos proteína – proteína (Ver Figura 3.5). Este servidor se encarga de desarrollar el acoplamiento con una serie de tres pasos; primero realiza un acoplamiento de cuerpos rígidos basado en el enfoque de correlación FFT generando billones de conformaciones, luego realiza un análisis de desviación media cuadrática (RMSD) para evaluar la estabilidad de estructuras que presenten las 1000 conformaciones con la menor energía, haciendo uso de una metodología de Clusters o agrupamientos de las conformaciones con una estructura similar generada; por último selecciona los grupos con mayor número de población y con una menor energía reportada para realizar un refinamiento mediante una minimización de energías.

Se evaluaron los complejos bajo dos consideraciones:

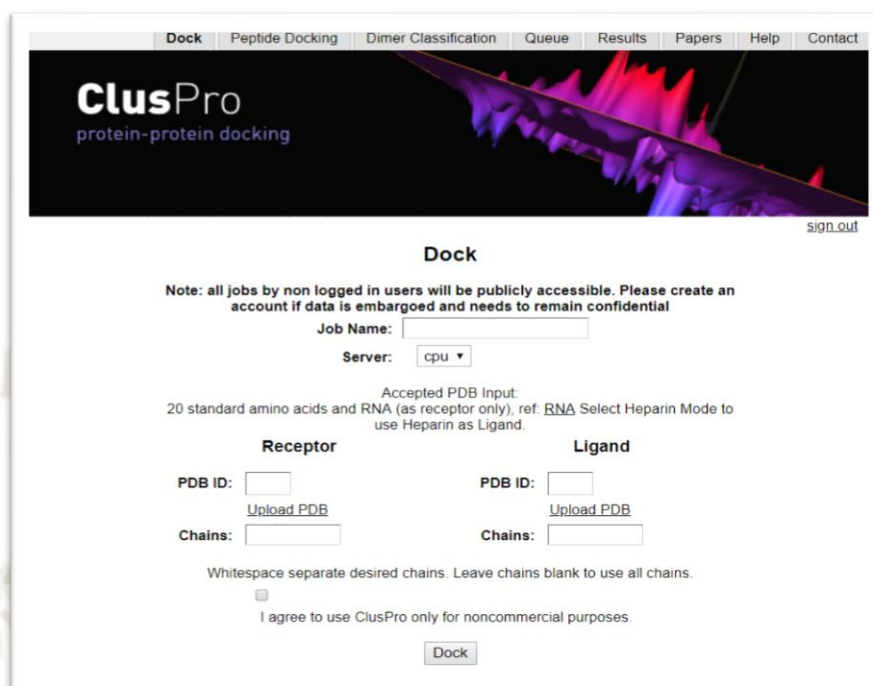
#### 3.2.11.1. *Histona Desacetilasa 1 – Profilina Humana II nativa:*

Se cargaron los archivos .pdb de cada proteína al servidor, siendo el archivo de la proteína profilina humana II aquel que no tuvo ningún acoplamiento previo con ligandos.

#### 3.2.11.2. *Histona Desacetilasa 1 – Profilina Humana II modificada:*

En este caso, se usó aquella profilina humana que ha pasado por un proceso de acoplamiento con los compuestos flavonoides, el archivo .pdb final de esta unión se modificó en el editor de textos Gedit de donde se eliminaron los datos (átomos, enumeración y coordenadas) del ligando, quedando las nuevas coordenadas que adquiere la profilina tras la interacción de estos. Este archivo .pdb se cargó al servidor junto con HDAC1.pdb para el respectivo acoplamiento.

El servidor brindó las estructuras de las proteínas acopladas junto con una tabla de los valores obtenidos en este cálculo, ambas basadas en afinidad hidrofóbica.

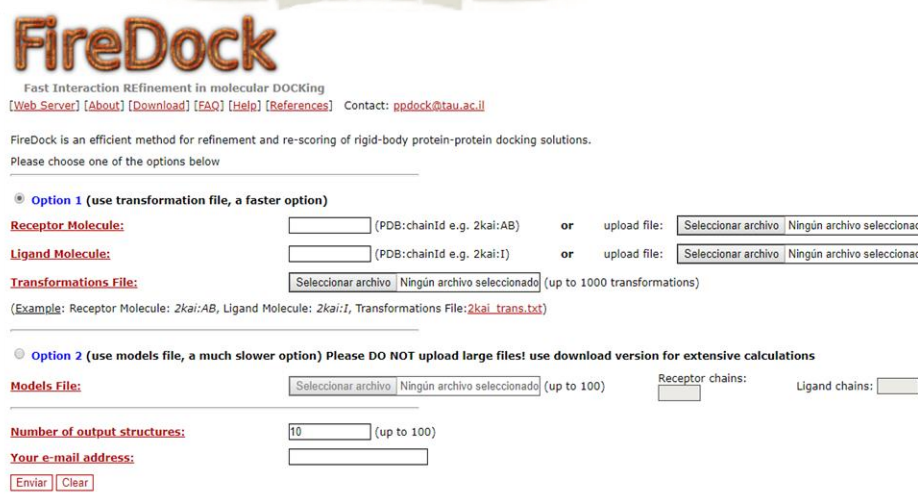


**Figura 3.5:** Servidor ClusPro para acoplamiento proteína – proteína

### 3.2.12. Refinamiento de los Complejos Proteína - Proteína:

Haciendo uso del servidor FireDock, se cargaron los archivos .pdb de los complejos formados por la enzima Histona Desacetilasa 1 y la Profilina Humana II nativa y aquellas modificadas.

Este servidor brinda tabla con valores de energías globales y energías de interacción de los complejos proteína-proteína tras el refinamiento efectuado.



**Figura 3.6:** Pantalla de trabajo del servidor FireDock para refinamiento molecular

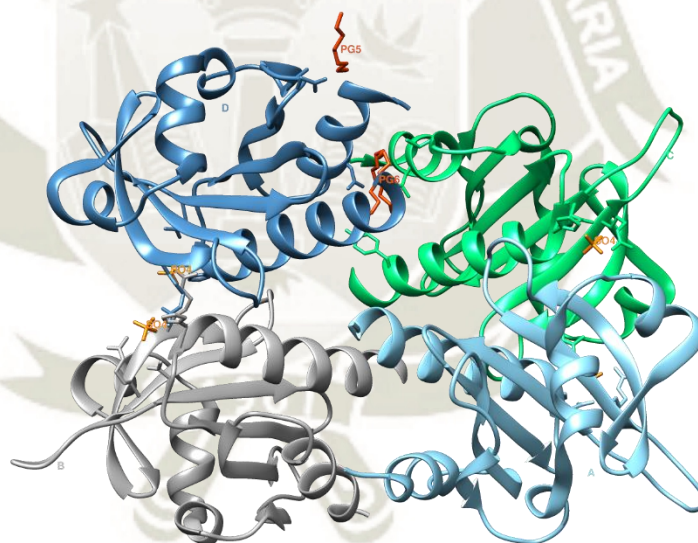


# CAPÍTULO 4

## 4. RESULTADOS Y DISCUSIÓN:

### 4.1. Obtención de la Estructura Tridimensional de Profilina Humana II (PFN2):

La búsqueda fue llevada a cabo en la base de datos del Protein Data Bank (PDB), donde se obtuvo la estructura tridimensional de la proteína cristalizada a una resolución de 2.2 Å con un total de 138 aminoácidos; la estructura está formando un homotetrámero (Ver Figura 4.1), junto con iones sulfato por motivos de estabilidad en el proceso de cristalización, siendo descargada ésta en formato .pdb y haciendo uso del editor de textos Gedit para eliminar las tres cadenas sobrantes.



**Figura 4.1:** Estructura tridimensional del complejo homotetramérico de Profilina Humana II descargada de la base de datos PDB con un ID: 1D1J

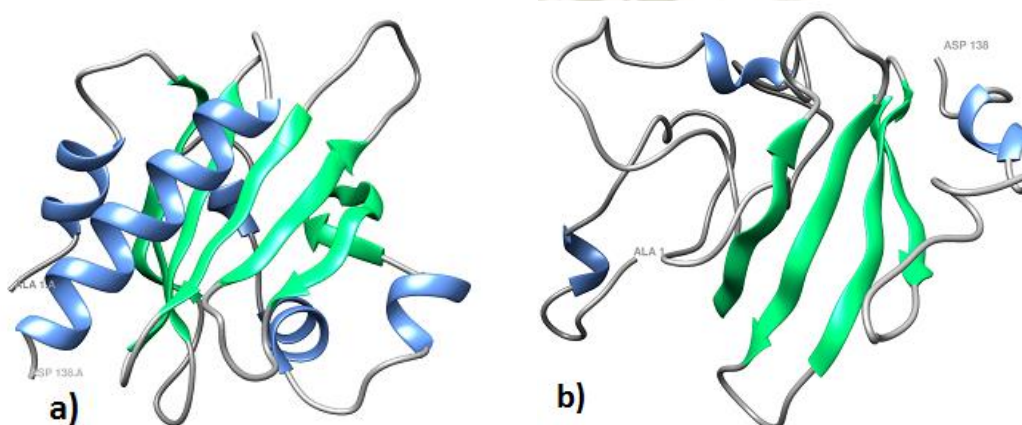
La secuencia de aminoácidos de esta proteína se encuentra consignada en el archivo PFN2.pdb junto a la descripción de sus átomos, sin embargo, se hizo búsqueda de esta misma secuencia en la base de datos NCBI-Protein donde obtuvimos un archivo en formato FASTA donde describe la secuencia de los aminoácidos (Ver Figura 4.2).

>AAH95444.1 Profilin 2 [Homo sapiens]

```
MAGWQSYVDNLMCDGCCQEAAIVGYCDAYVWAATAGGVFQSITPIEIDMIVGKDREGFFTNGLTGAKK
CSVIRDSLYVDGDC TMDIRTKSQGGEPTYNVAVGRAGRALVIVMGKEGVHGGT LNK KAYELALYLRRSDV
```

**Figura 4.2:** Secuencia de aminoácidos de Profilina Humana II en formato FASTA

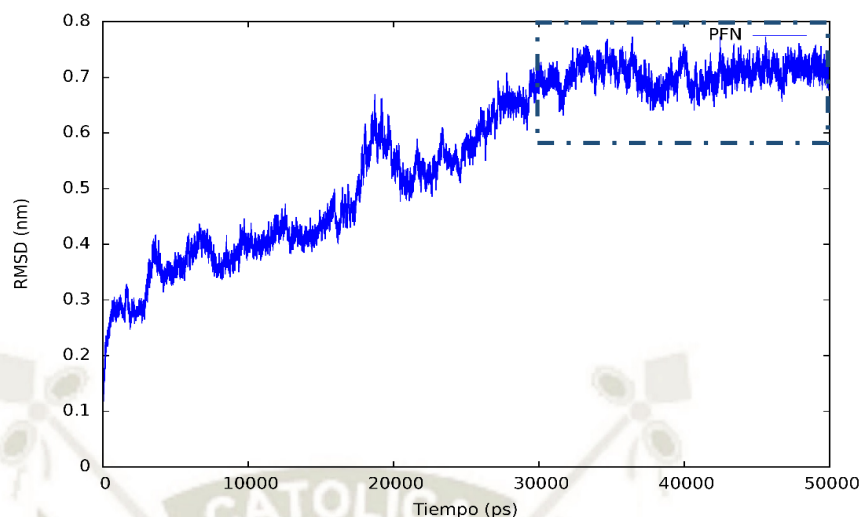
Obtenida la proteína Profilina Humana II (PFN2), se procedió con el trabajo, para lo cual haciendo uso del software Gromacs 2016, se realizó una minimización de energías de esta proteína con un integrador steep a 20000ps con un solvente SPC; luego se procedió a una dinámica molecular con un colectivo canónico NVT, el cual mediante el termostato de Nose-Hoover mantiene la temperatura a 309.65°K del sistema que incluye el solvente SPC y que evoluciona a través del tiempo durante 50000ps o 50ns.



**Figura 4.3:** Estructuras tridimensionales de PFN2 nativa resaltando su estructura secundaria de azul las  $\alpha$ -hélices y en verde las láminas  $\beta$ . (a) Pre-dinámica (b) Post-dinámica

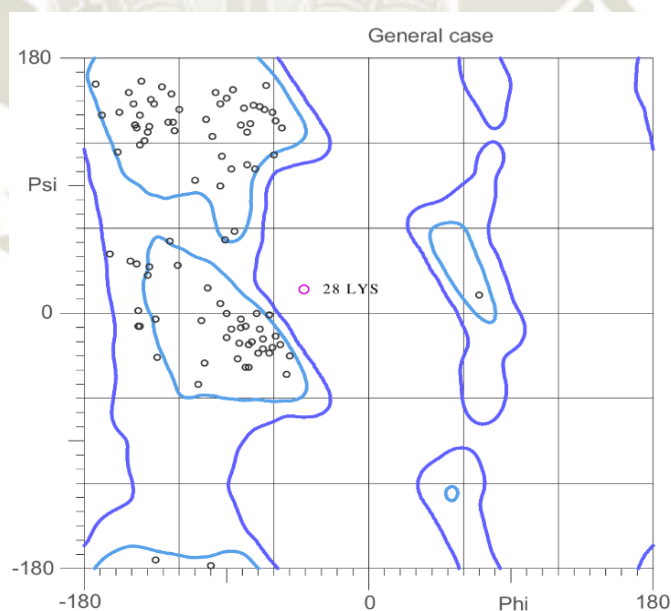
En la figura 4.3, podemos observar como la estructura varía luego de un proceso de minimización y dinámica molecular, habiendo una variación en su estructura secundaria tanto en el número de  $\alpha$ -hélices como el número de láminas  $\beta$ , siendo ambas determinantes en la estabilidad de la proteína, es por ello que se procedió al uso de las herramientas *gmx rms* y diagrama de Ramachandran para evaluar esta estabilidad en la estructura de nuestra proteína.

La herramienta *rms* o *rmsd* calcula la desviación media cuadrática con respecto a la distancia haciendo uso de las trayectorias generadas durante la dinámica molecular obteniendo un gráfico que puede ser visualizado con la herramienta de ploteo de gráficos Grace, en la gráfica se muestra que la proteína alcanza la estabilidad pasado los 30ns (Ver figura 4.4).



**Figura 4.4:** Ploteo de RMSD de la proteína Profilina Humana II post-dinámica, que alcanza la estabilidad a partir de los 30ns

A continuación, se usó el servidor online MolProbity para el ploteo del diagrama de Ramachandran, el cual mostró que el residuo LYS28 está fuera de la zona permitida debido posiblemente a su cadena lateral la cual está presente en un extremo de la proteína lo cual la hace propensa a interacciones; pese a ello, hay un 99.3% de residuos (135/136) que si están presentes en zonas permitidas (Ver Figura 4.5).



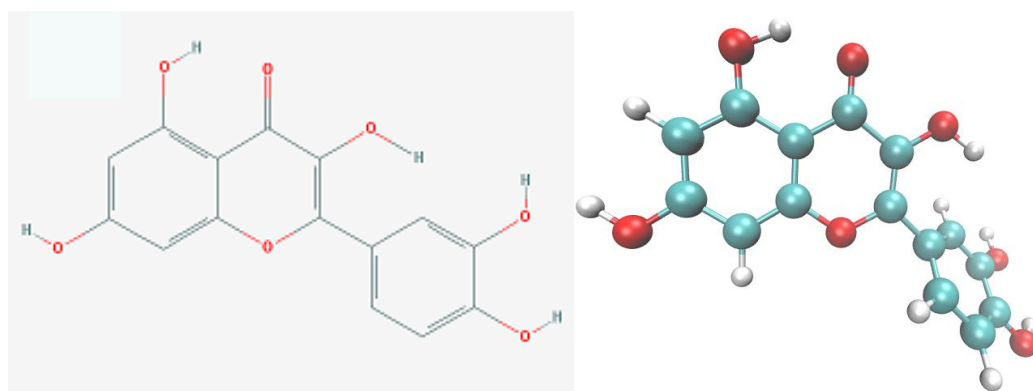
**Figura 4.5:** Diagrama de Ramachandran de la proteína Profilina Humana II, la cual presenta al residuo LYS28 fuera de la zona permitida



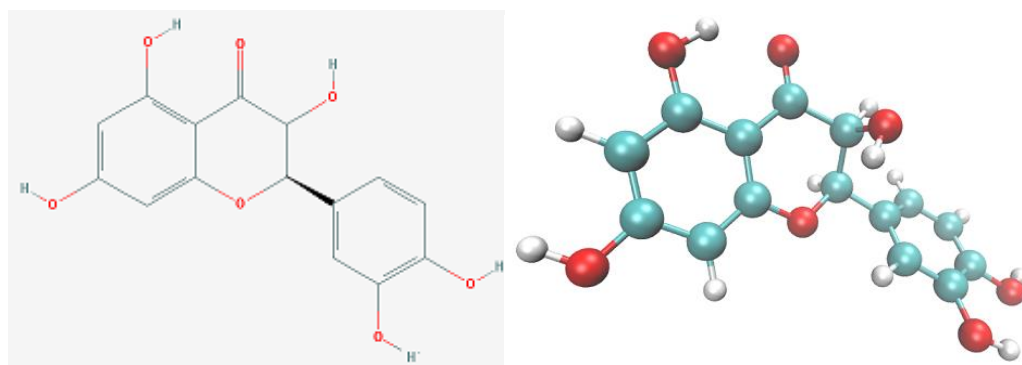
## 4.2. Obtención de la Estructura Tridimensional Optimizada de Compuestos Flavonoides:

La búsqueda de los compuestos flavonoides reportados en flor y fruto de *Opuntia ficus-indica* según la tabla 2.2, fue llevada a cabo en la base de datos de PubChem de donde se descargó 11 archivos conteniendo la estructura de estos compuestos que luego fueron optimizados geoméricamente haciendo uso del software Gaussian 09 y la funcional híbrida CAM-B3LYP (Ver Figuras 4.6 – 4.16).

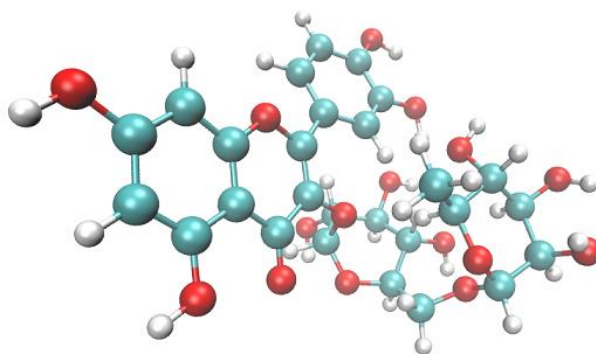
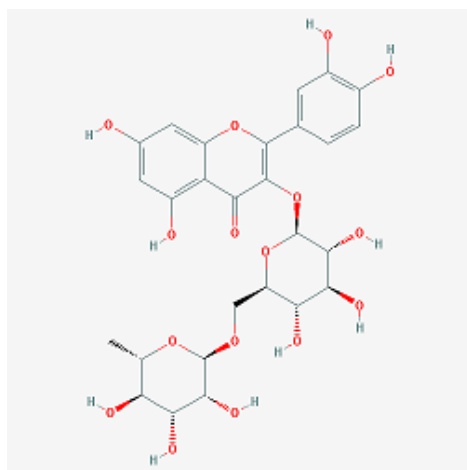
Terminada la optimización geométrica de los nuestros compuestos flavonoides, se realizó un single point para el cálculo de las cargas de Hirshfeld igualmente con el software Gaussian 09, este análisis identifica los orbitales moleculares y asigna las cargas atómicas a cada átomo de la molécula basados en la trayectoria generada a partir de la optimización y se basa en la teoría funcional de la densidad (DFT).



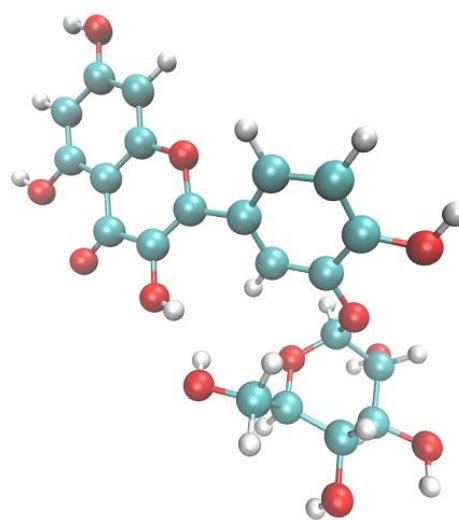
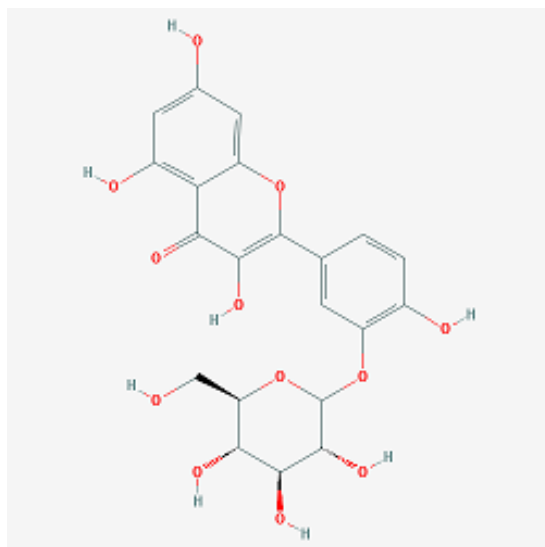
**Figura 4.6:** Compuesto flavonoide Quercetina y su representación 2D (izquierda) y 3D (derecha)



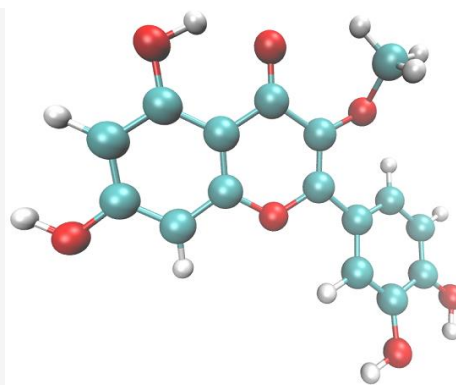
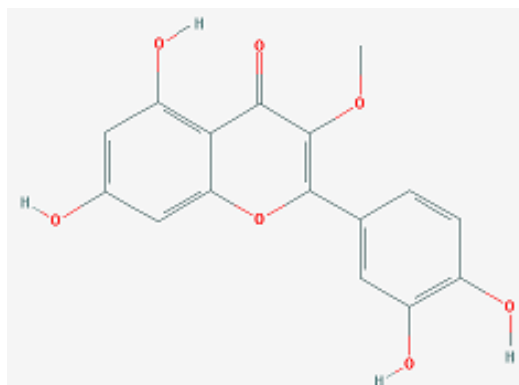
**Figura 4.7:** Compuesto flavonoide (+) – Dihidroquercetina y su representación 2D (izquierda) y 3D (derecha)



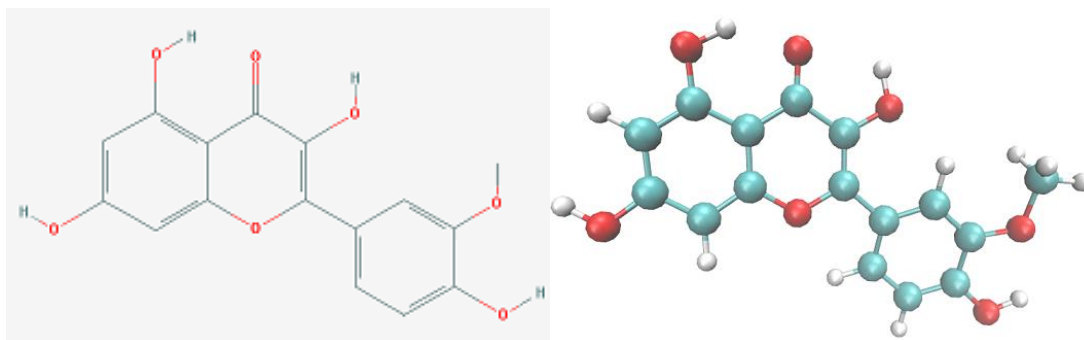
**Figura 4.8:** Compuesto flavonoide Quercetina 3-O-rutinósido y su representación 2D (izquierda) y 3D (derecha)



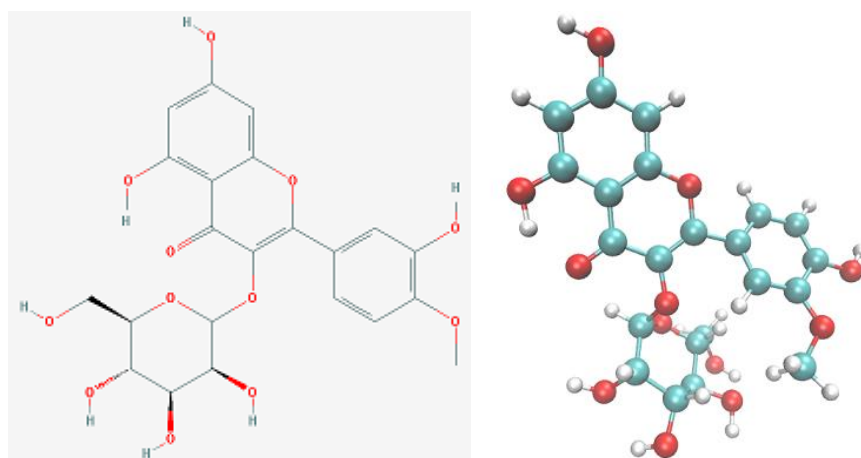
**Figura 4.9:** Compuesto flavonoide Quercetina 3-O-glucósido y su representación 2D (izquierda) y 3D (derecha)



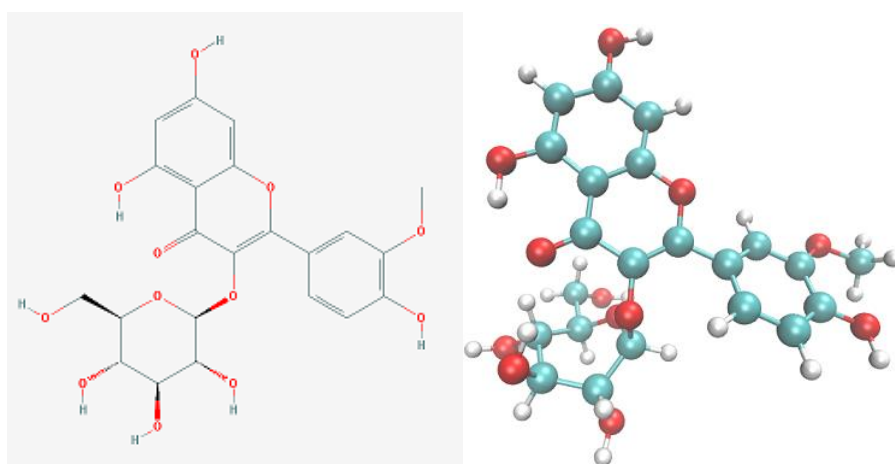
**Figura 4.10:** Compuesto flavonoide Quercetina 3-metil éter y su representación 2D (izquierda) y 3D (derecha)



**Figura 4.11:** Compuesto flavonoide Isorhamnetin y su representación 2D (izquierda) y 3D (derecha)

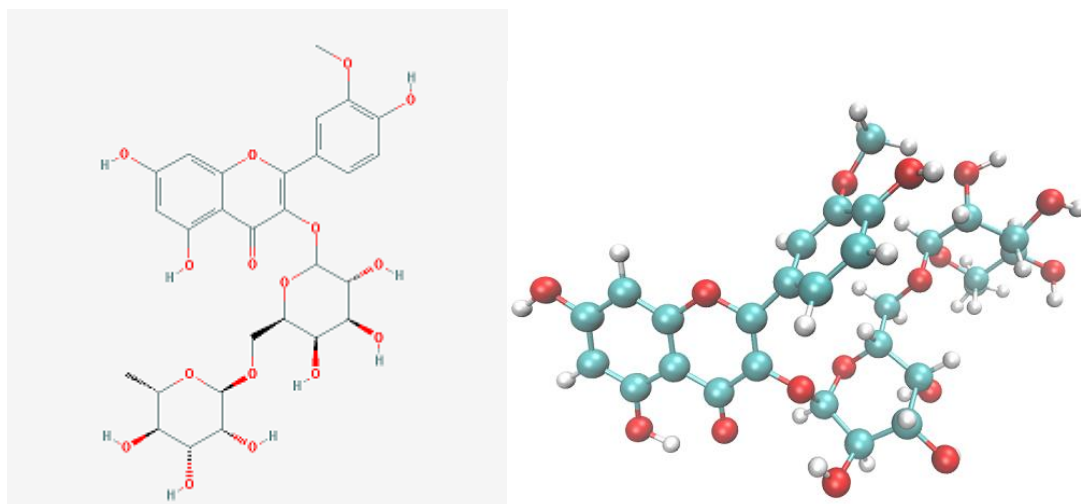


**Figura 4.12:** Compuesto flavonoide Isorhamnetin 3-O-galactósido y su representación 2D (izquierda) y 3D (derecha)

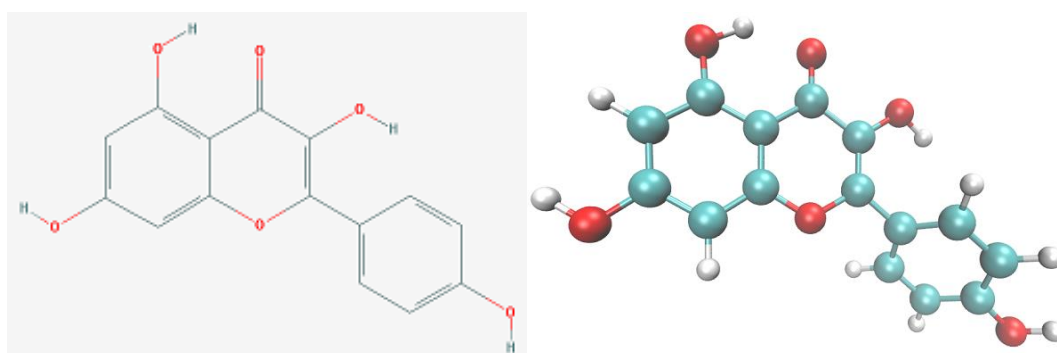


**Figura 4.13:** Compuesto flavonoide Isorhamnetin 3-O-glucósido y su representación 2D (izquierda) y 3D (derecha)

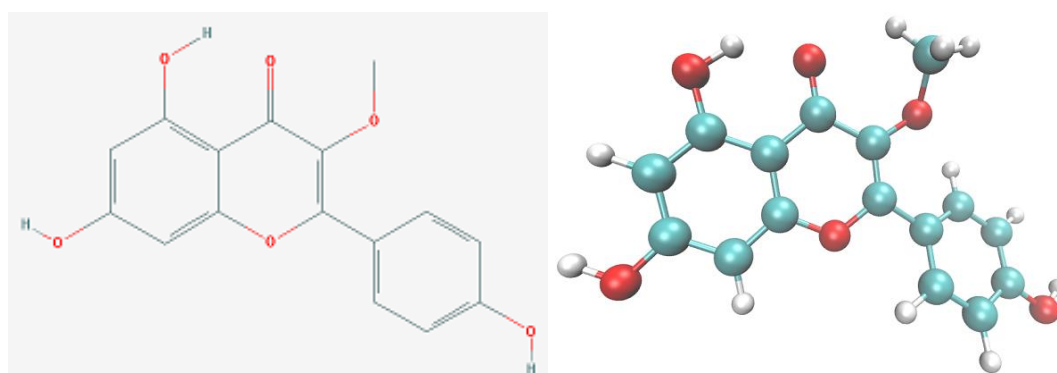




**Figura 4.14:** Compuesto flavonoide Isorhamnetin 3-O-robinobióside y su representación 2D (izquierda) y 3D (derecha)



**Figura 4.15:** Compuesto Kaempferol y su representación 2D (izquierda) y 3D (derecha)



**Figura 4.16:** Compuesto Kaempferol 3-metil éter y su representación 2D (izquierda) y 3D (derecha)

Así, habiendo obtenido cada compuesto optimizado y sus respectivas cargas, se procedió a generar la topología necesaria de estos en el servidor TPPKMKTOP que asignó los parámetros del campo de fuerza OPLS-AA a cada átomo perteneciente a los diferentes compuestos flavonoides, que se consignan en la raíz del software Gromacs 2016. Para este último paso, se requirió el acceso al sistema como super usuario a través de la terminal de Linux, accediendo hasta la raíz de Gromacs a la carpeta del campo de fuerza OPLS-AA ubicado en /usr/local/Gromacs/share/Gromacs/top/OPLS-AA.ff, siendo de especial interés modificar el archivo *aminoacids.rtp* en el cual se describe la denominación de cada átomo, el tipo de átomo que es según la estructura y función y finalmente las cargas que adquieren según lo descrito en otro archivo *atomtypes.atp* pero en nuestro caso se asignó las cargas de Hirshfeld obtenidas por cálculos de mecánica cuántica (Ver Apéndice 7-17).

Con las moléculas ya optimizadas y parametrizadas, se hizo un análisis previo a emplearlas en acoplamiento molecular, con un servidor que hace uso de las cinco reglas de Lipinski para identificar la similitud de una molécula con un fármaco y su viabilidad basada en 4 criterios que en sus rangos tienen valores que son múltiplos de 5, he de allí el nombre. Los resultados de este análisis están presentes en la tabla 4.1.

**Tabla 4.1.** Resultado de los compuestos flavonoides evaluados con las reglas de Lipinski

COMPUESTO	MASS (<500Da)	H-BOND DONOR ( $\leq 5$ )	H-BOND ACCEPTOR ( $\leq 10$ )	Log P (<5)
QME	322	5	7	1.0056
DHQ	304	5	7	1.1863
ISO	320	5	7	2.06179
KMP	288	5	6	2.5815
KME	306	4	6	2.2798
QCT	306	6	7	1.20259
IGL	486	8	12	-1.411001
IGA	480	8	12	-0.1859
QGL	468	9	12	-2.608911
QRU	614	11	16	-1.690923
IRO	626	10	16	-1.334102

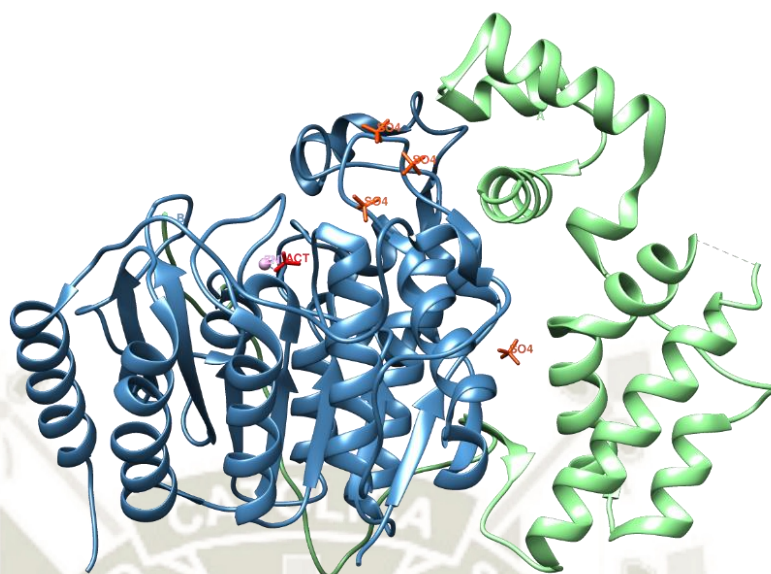
Según los criterios de las cinco reglas de Lipinski, una molécula debe de tener una masa molecular menor a 500 Dalton, no tener más de 5 átomos donantes de enlaces de puente de hidrogeno, no tener más de 10 átomos aceptores de enlaces puente de hidrogeno y por último un logaritmo de coeficiente de partición octanol-agua no mayor de 5; teniendo que cumplir con al menos 3 de las 4 reglas para cumplir con el objetivo aunque existen excepciones especialmente en moléculas de origen natural. Aunque adicionalmente existen variantes para mejorar estas reglas como el rango del logaritmo del coeficiente de partición entre -0.4 a 5.6, cálculo de área de superficie polar no mayor a  $140\text{\AA}^2$ , no más de 10 enlaces rotables entre otros.

En la tabla 4.1, podemos observar que los compuestos QME, DHQ, ISO, KMP, KME cumplieron en su totalidad con las reglas, mientras que QCT incumplió con solo una de estas, y los compuestos IGL, IGA, QGL incumplieron con dos de las reglas pudiendo estar en el límite de aceptación según el rango de tolerancia para estos compuestos de origen natural, mientras que los compuestos quercetina 3-O-rutinósido (QRU) e isorhamnetin 3-O-robinobiósido (IRO) incumplieron con la mayoría, por ende fueron descartados del estudio.

### **4.3. Obtención de la Estructura Tridimensional de Histona Desacetilasa 1:**

La búsqueda fue llevada a cabo en la base de datos del Protein Data Bank (PDB), donde se obtuvo la estructura tridimensional de la proteína cristalizada a una resolución de  $3\text{\AA}$  con una secuencia de 369 aminoácidos; la estructura está formando un heterodímero junto con la proteína MTA1 de 176 aminoácidos y además iones potasio, iones sulfato, iones acetato y un ión zinc que es centro catalítico de la HDAC1 (Ver Figura 4.17). La estructura se descargó en formato .pdb y en el editor de textos Gedit se eliminó la proteína MTA1 y todos los iones a excepción del ion zinc ( $\text{Zn}^{+2}$ ).





**Figura 4.17:** Estructura tridimensional del complejo heterodimérico de Histona Desacetilasa I con la proteína Asociada a Metástasis (MTA) descargada de la base de datos PDB con un ID: 4BKX

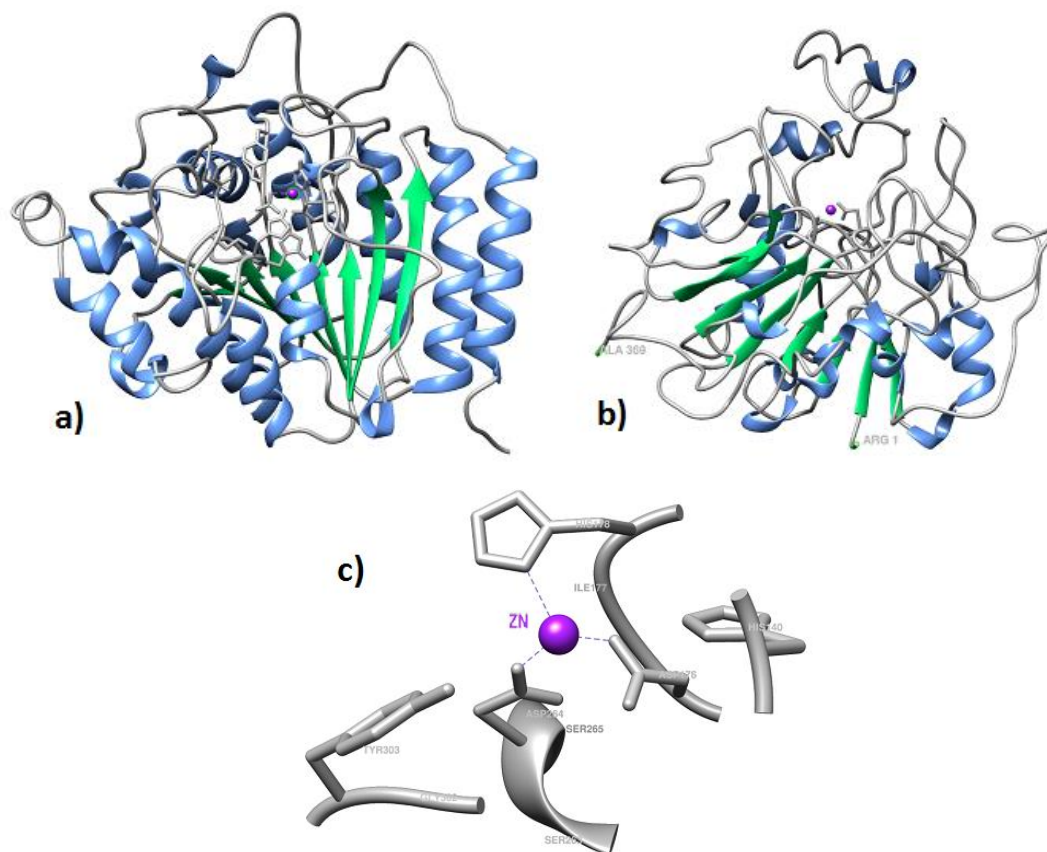
Al igual que en el caso anterior, se realizó una búsqueda de la secuencia de aminoácidos en la base de datos NCBI-Protein donde obtuvimos un archivo en formato FASTA donde describe la secuencia de los aminoácidos (Ver Figura 4.18), aquí podemos ver una diferencia notoria ya que la secuencia de aminoácidos de toda la proteína en el formato FASTA presenta 482 aminoácidos, mientras que el archivo .pdb presenta tan solo 369, esta secuencia excedente de aminoácidos convierte a la proteína en un precursor inactivo, es por ello que en el proceso de aislamiento y cristalización, perdió esta parte de su cadena de aminoácidos.

```
>CAG46518.1 HDAC1 [Homo sapiens]
MAQTQGTTRRKVCYYYDGDVGNYYYGGHPMKPHRIRMTHNLLLNYGLYRKMEIYRPHKANAEEMTKYHSD
DYIKFLRSIRPDNMSEYSKQMQRFNVGEDCPVFDGLFEFCQLSTGGSVASAVKLNKQQTDAIVNWAGGLH
HAKKSEASGFCYVNDIVLAILELLKXHQVLYIDIDIHHGDGVVEEAFYTTDRVMTVSVFKHYGEYFPGTGD
LRDIGAGKGKYYAVNYPLRDGIDDESIEAIFKPVMSKVMEMFQPSAVVLQCGSDSLSGDRLGCFNLTIKG
HAKCVEFVKSFNLPMLMLGGGGYTIRNVARCWYETAVALDTEIPNELPYNDYFEYFGPDFKLHISPSNM
TNQNTNEYLEKIKQRLFENLRMLPHAPGVQMQAIPEDAIPEESGDEDEDDPKRISICSSDKRIACEEEF
SDSEEEGEGGRKNSSNFKAKRVKTEDEKEKDPEEKKEVTEEEKTKEEKPEAKGVKEEVKLA
```

**Figura 4.18:** Secuencia de aminoácidos de Histona Desacetilasa 1 en formato FASTA, resaltando la secuencia catalítica que codifica la proteína tridimensionalmente

La proteína HDAC1 junto con su centro catalítico de Zn, fue sometida a un proceso de minimización de energías, haciendo uso de un integrador steep a 20000ps con un

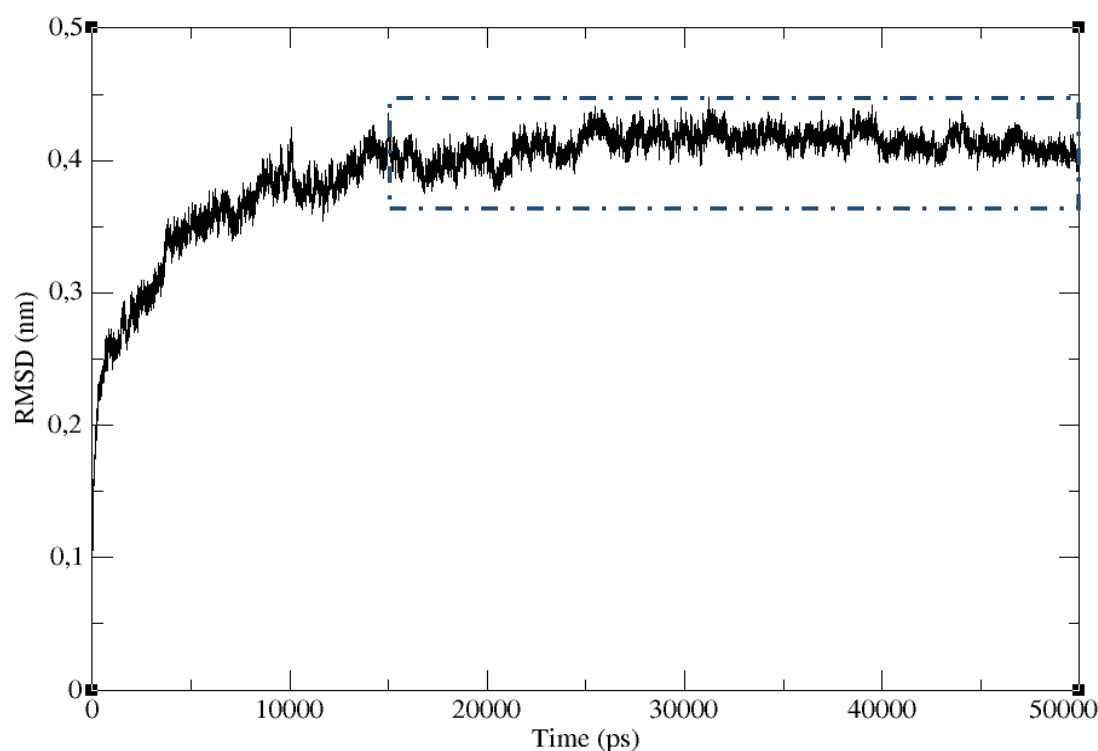
solvente SPC; luego se procedió a una dinámica molecular con un colectivo canónico NVT, el cual mediante el termostato de Nose-Hoover mantiene la temperatura a 309.65°K del sistema que incluye el solvente SPC y que evoluciona a través del tiempo durante 50000ps o 50ns.



**Figura 4.19:** Estructuras tridimensionales de HDAC1 nativa resaltando su estructura secundaria y el centro catalítico de Zinc. (a) Pre-dinámica (b) Post-dinámica (c) Centro catalítico

En la figura 4.19, se logra observar como la estructura secundaria de la proteína varió luego de un proceso de minimización y dinámica molecular, reduciéndose el número de  $\alpha$ -hélices como el número de láminas  $\beta$ , siendo ambas determinantes en la estabilidad de la proteína, además se observa el centro de catalítico de zinc que presenta esta enzima y los residuos que interactuaron con este. Se procedió al uso de las herramientas *gmx rms*, *gmx rmsf*, *gmx gyrate* y diagrama de Ramachandran para evaluar esta estabilidad en la estructura de nuestra proteína.

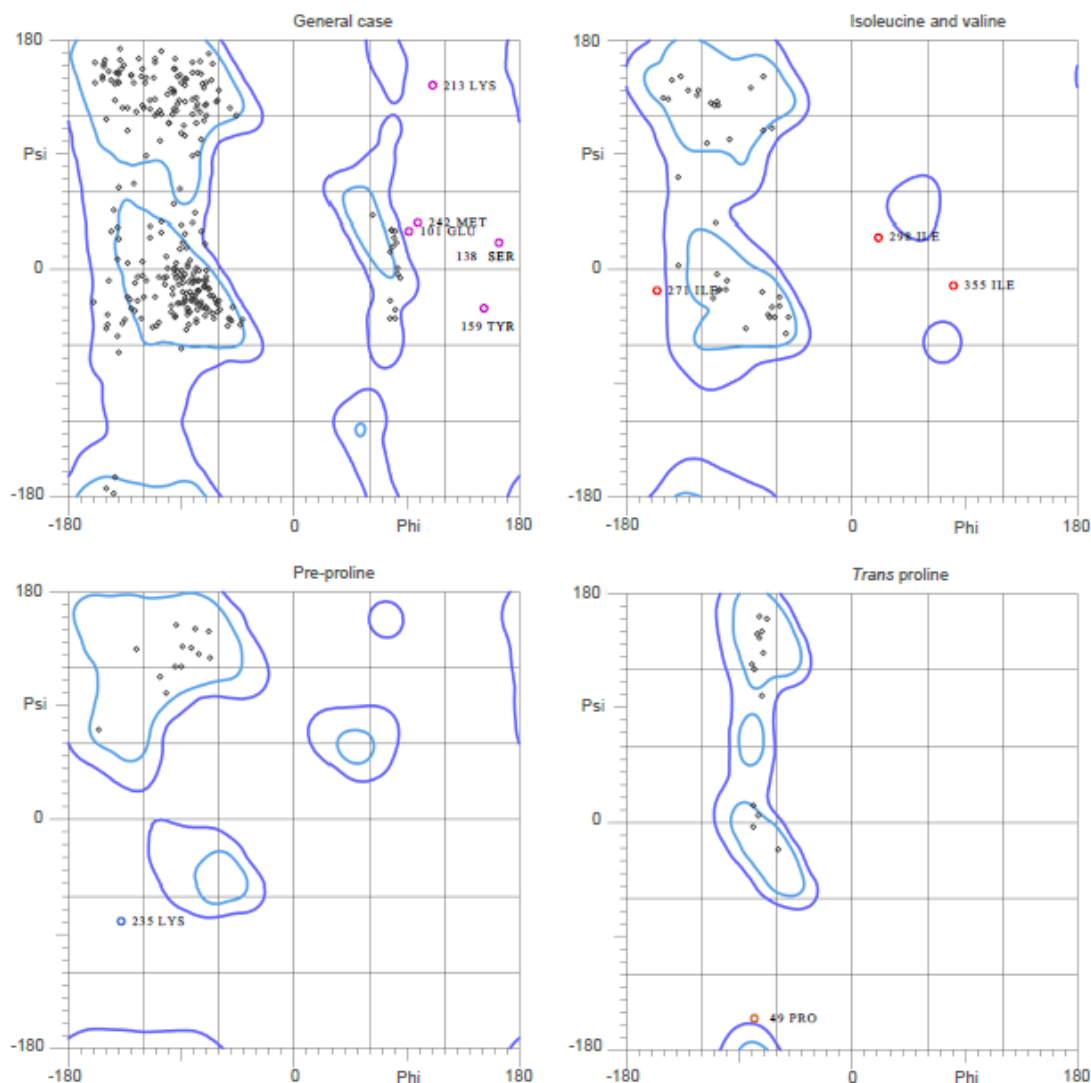
La herramienta *rmsd* para el cálculo de la desviación media cuadrática con respecto a la distancia, mostró un gráfico generado por la herramienta de gráficos Grace, en el que la proteína alcanzó tempranamente la estabilidad a los 15ns (Ver figura 4.20).



**Figura 4.20:** Ploteo de RMSD de la enzima Histona Desacetilasa 1 post-dinámica con una zona de estabilidad a partir de los 15ns

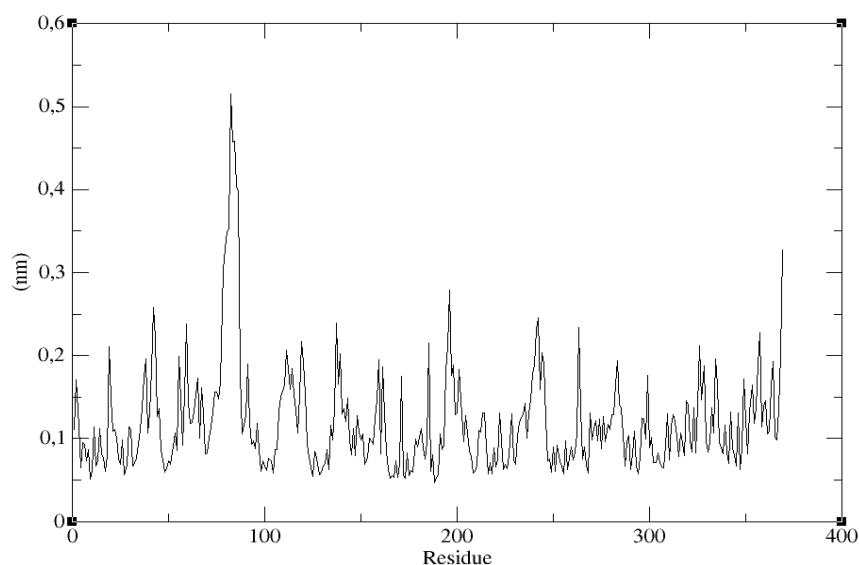
A continuación, con el uso del servidor online MolProbity se plotó el diagrama de Ramachandran, el cual nos muestra que varios residuos están fuera de la zona permitida siendo estos: PRO49, GLU101, SER138, TYR159, LYS213, LYS235, MET242, ILE271, ILE298 y ILE355; pese a ser 10 aminoácidos que están en zonas no permitidas, los demás residuos que si están presentes en zonas permitidas (357/367) representan un 97.3% (Ver Figura 4.21).





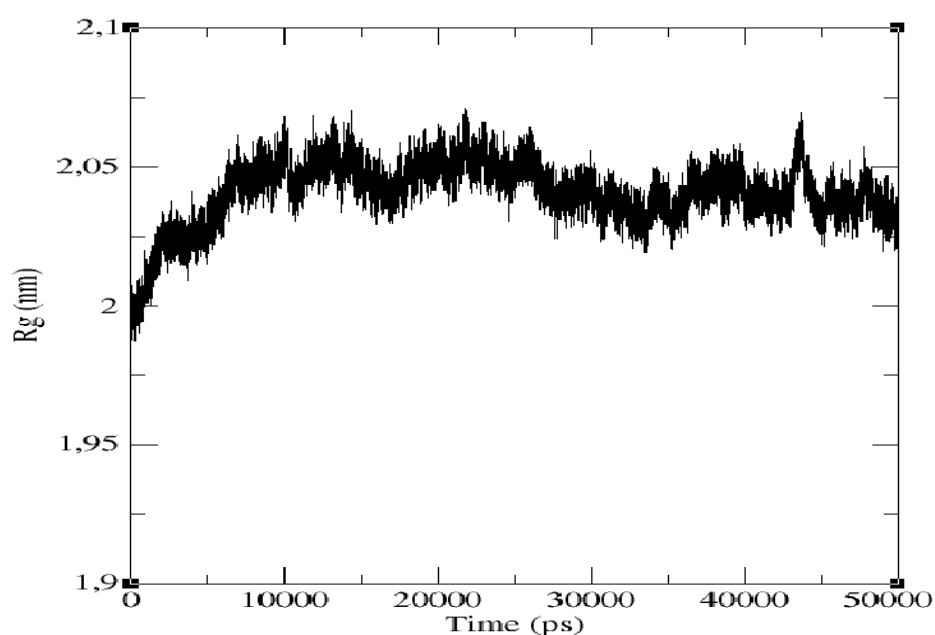
**Figura 4.21:** Diagrama de Ramachandran de HDAC1, donde se muestra a 10 residuos fuera de la zona permitida

Debido a la presencia de estos residuos fuera de la zona permitida, se hizo uso de dos herramientas complementarias para la evaluación de esta enzima. La primera fue *rmsf*, que se utilizó para calcular la fluctuación de la desviación media cuadrática por residuos, pudiendo observar que los residuos LYS82, GLN83, MET84, GLN85 y ARG86 presentaron un pico por su alta fluctuación y que esta notablemente vinculada con el área de la proteína en la que están presentes, al estar sus cadenas laterales orientadas hacia el exterior en un extremo de la proteína (Ver Figura 4.22).



**Figura 4.22:** Ploteo de RMSF de la enzima HDAC1 post-dinámica que presenta un pico entre los residuos SER81 y ARG86

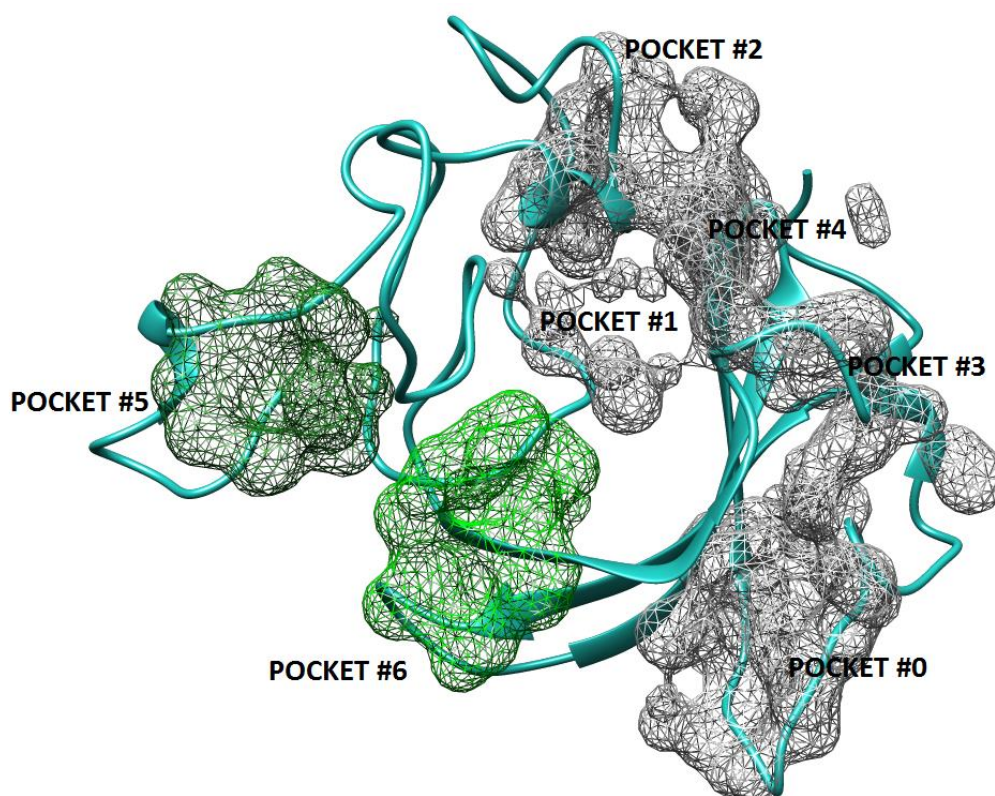
La última herramienta utilizada fue el cálculo del radio de giro o *gyrate*, esta herramienta genera un centro de masa de la proteína y de aquí evalúa la variación distal entre este centro y el centro de masa de aquel residuo que este más alejado a través del tiempo de evaluación de 50ns, evaluando a su vez el plegamiento o compactación. Pudiendo ver en la figura 4.23 que la distancia hacia el extremo más lejano es en promedio 2,04nm y que se estabiliza rápidamente.



**Figura 4.23:** Ploteo del radio de giro de la enzima HDAC1 post-dinámica que alcanza estabilidad a partir de los 15ns con un radio de 2.04 aproximadamente

#### 4.4. Interacción Proteína - Ligando:

Teniendo nuestra proteína Profilina Humana II (PFN2) ya optimizada por dinámica molecular, con un archivo final .gro, se procedió a convertirlo en un archivo con extensión .pdb. Este archivo se usó para el análisis de la estructura y la predicción de bolsillos o lugares de acoplamiento con el servidor PockDrug, el cual hace uso de algoritmos que identifican las zonas hidrofóbicas y geométricamente aceptables de una proteína para el acoplamiento de una molécula de interés. Este servidor tuvo 7 archivos finales con extensión .pdb, cada uno de los cuales representaba un bolsillo en zonas diversas de la proteína (Ver Figura 4.24).



**Figura 4.24:** Bolsillos o pockets de la proteína Profilina Humana II calculada en el servidor PockDrug

A su vez, el servidor brindó un score de probabilidad de unión o porcentaje de drogabilidad de cada uno de estos bolsillos, los cuales se describen en la tabla 4.2.



**Tabla 4.2:** Score de drogabilidad de los bolsillos o pockets de PFN2

Número de Bolsillo	Probabilidad de Drogabilidad
#0	0.61
#1	0.19
#2	0.72
#3	0.05
#4	0.28
#5	0.98
#6	0.65

Ya con los lugares de acoplamiento dilucidados, se procedió a realizar el acoplamiento molecular de esta proteína con todos los compuestos flavonoides que cumplieran con las cinco reglas de Lipinski, requiriendo de los archivos con extensión .pdb tanto de la proteína como de los flavonoides.

El acoplamiento molecular se llevó a cabo con el software AutoDock, el primero paso fue el uso de la herramienta AutoGrid que reconoce las coordenadas tridimensionales de la proteína, genera mapas de estos y calcula su potencial electrostático y su potencial de desolvatación, generando archivos con extensión .glg, .map y .pdbqt; a partir de las grillas y los valores generados por esta herramienta se procedió con el uso de AutoDock el cual acopló ambas estructuras generando archivos con extensión .dlg y .pdb con las coordenadas finales del cálculo que contemplaba 50 corridas o evaluaciones de las cuales se generó un ranking según la constante de inhibición calculada y la energía de unión o acoplamiento (Ver Tabla 4.3), estando la energía de acoplamiento basada en la siguiente ecuación:

$$\text{Energía de Acoplamiento Estimada (Kcal/mol)} = [1] + [2] + [3] - [4] \quad (4.1)$$

De donde [1] refiere a las energías de Van der Waals, puentes de hidrógeno y energía de desolvatación en kcal/mol; [2] refiere a la energía interna total en kcal/mol, [3] refiere a la energía libre de torsión en kcal/mol y por último, [4] refiere a la energía del sistema desacoplado que es igual a la energía interna total.

**Tabla 4.3:** Valores de la constante de inhibición y de energía de acoplamiento para cada uno de los sistemas proteína-ligando formados a partir del software AutoDock

Compuesto	Abreviación	AUTODOCK		
		Constante de Inhibición ( $\mu\text{M}$ )	Energía de Unión (Kcal/mol)	Energía de Unión (KJ/mol)
1 Quercetin	QCT	3.69	-7.41	-31.00
2 Quercetin 3-O-glucoside	QGL	14.35	-6.61	-27.65
3 Quercetin 3-methyl ether	QME	8.97	-7.33	-30.67
4 (+) - Dihydroquercetin	DHQ	6.12	-7.11	-29.75
5 Isorhamnetin	ISO	3.86	-7.39	-30.92
6 Isorhamnetin 3-O-Galactoside	IGA	14.29	-6.61	-27.66
7 Isorhamnetin 3-O-Glucoside	IGL	9.61	-6.84	-28.62
8 Kaempferol	KMP	3.49	-7.45	-31.17
9 Kaempferol 3-methyl ether	KME	6.77	-7.05	-29.50

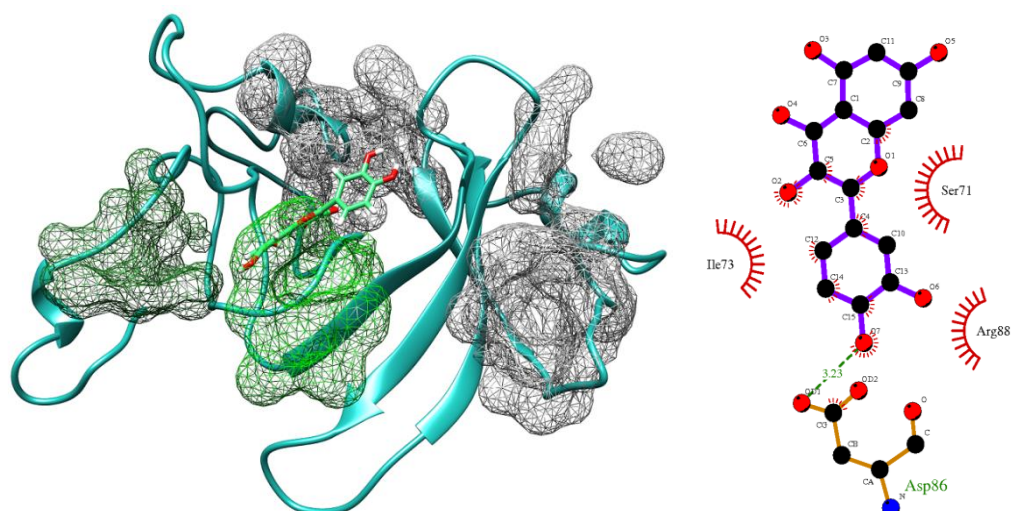
De esta tabla pudimos inferir que los valores de la constante de inhibición son menores en aquellos compuestos flavonoides que no presentan una cadena glucosídica como es el caso de quercetina, isorhamnetin y kaempferol; mientras que estas mismas presentan los valores más elevados de energía de acoplamiento en Kcal/mol pero sin una diferencia muy notoria por sobre los demás compuestos.

El software AutoDock generó un archivo final .pdb en el cual se encuentran las coordenadas tridimensionales del complejo proteína-ligando, se abrió este archivo con un editor de textos para modificarlo primeramente al agregar los términos TER después del último átomo del ultimo residuo de la proteína y después del último átomo del flavonoide para que se diferencien ambas cadenas; luego se modificó la denominación LIG por los códigos de tres letras asignados a cada flavonoide en estudio para que sea reconocido nuevamente por el campo de fuerza OPLS-AA. Para poder evaluar la interacción de este sistema y si el acoplamiento se mantiene a través del tiempo, se realizó una minimización de energía haciendo uso del archivo final .pdb y luego se procedió con la dinámica molecular haciendo uso de un colectivo canónico NVT y uso del termostato de Nose-Hoover a una temperatura de 309.65°K del sistema que incluye el solvente SPC y que evoluciona a través del tiempo durante 50000ps.

#### 4.4.1. Acoplamiento PFN2 y quercetina:

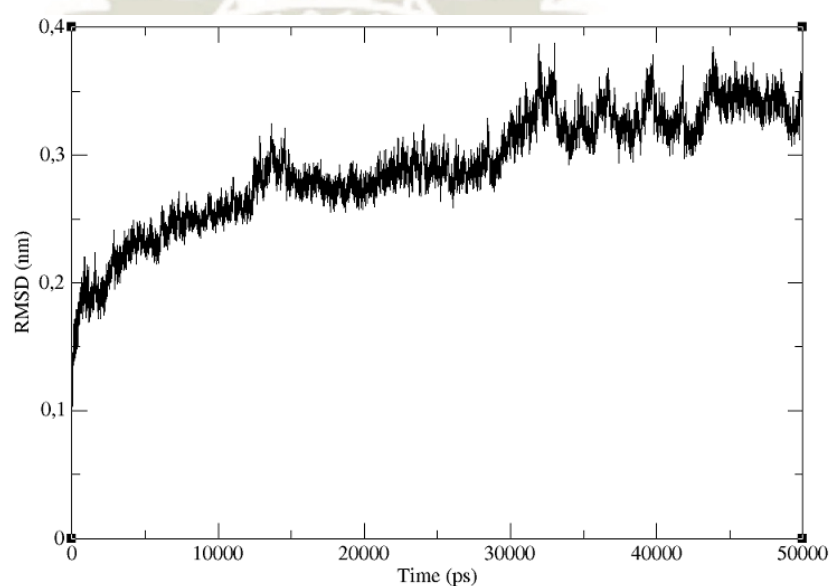
Con el visualizador Chimera se observó la estructura generada por AutoDock resaltando el bolsillo en el que se acopló, y con la herramienta LigPlot se visualizó los residuos de la proteína que interactuaron con el flavonoide tras la dinámica molecular (Ver Figura 4.25).





**Figura 4.25:** Acoplamiento de PFN2 con Quercetina en el bolsillo 6, y los residuos que interactúan con este compuesto

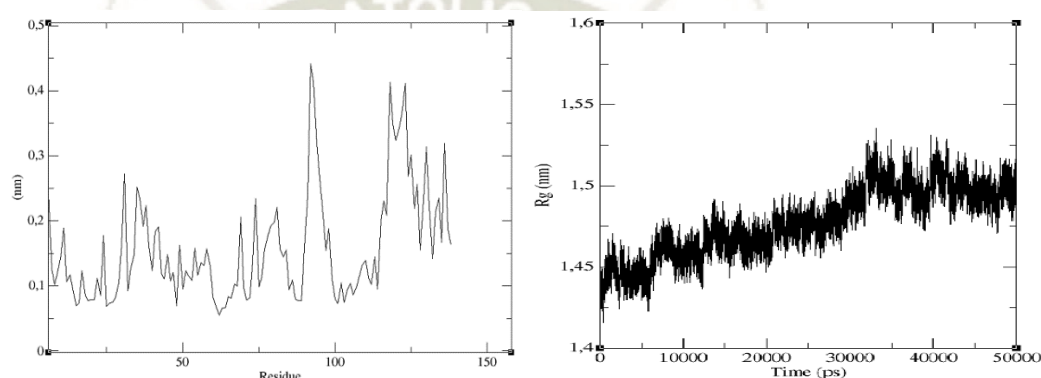
Se pudo observar de la figura anterior que QCT se acopló inicialmente al bolsillo #6 resaltado en verde claro, y tras la dinámica molecular interactuó con los residuos SER71, ILE73, ARG88 y formó un puente de hidrogeno con el residuo ASP86, aunque estos residuos no forman parte de un *pocket*; luego para evaluar el sistema, se hizo uso de las herramientas *rmsd*, *rmsf* y radio de giro para evaluar esta estabilidad en la estructura de nuestra proteína. La herramienta *rmsd* calcula la desviación media cuadrática obteniendo un gráfico con la herramienta de ploteo de gráficos Grace, mostrando en la gráfica que la proteína alcanzó la estabilidad cerca a los 35ns (Ver Figura 4.26).



**Figura 4.26:** Ploteo de RMSD tras una dinámica molecular de la proteína PFN2 acoplada con QCT



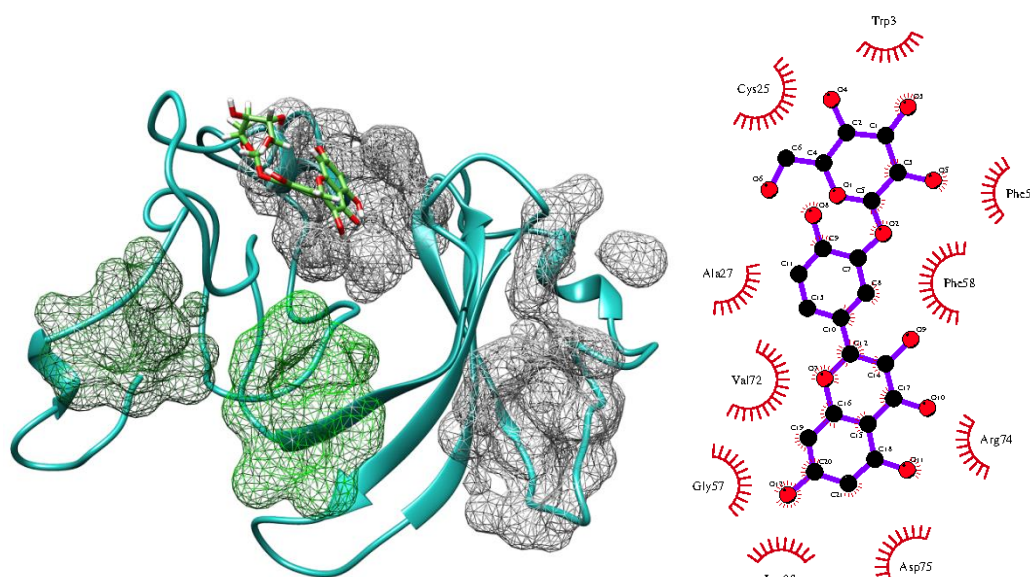
Luego, se hizo uso de dos herramientas complementarias para la evaluación de la proteína en este sistema. La primera fue *rmsf*, que se utilizó para calcular la fluctuación de la desviación media cuadrática por residuos, pudiendo observar que el cálculo presentó dos picos que involucran a los residuos GLN92, GLY93, VAL118 y LEU123 respectivamente; mientras que la segunda herramienta utilizada fue el cálculo del radio de giro o *gyrate*, para ver el plegamiento o estiramiento de la proteína tras la interacción. Pudiendo ver que la distancia hacia el extremo más lejano es de aproximadamente 1,5nm y que se termina de estabilizar pasado los 30ns o 30000ps (Ver Figura 4.27).



**Figura 4.27:** Ploteos de RMSF y radio de giro respectivamente, de la proteína PFN2 acoplada a QCT

#### 4.4.2. Acoplamiento PFN2 y quercetina 3-O-glucósido:

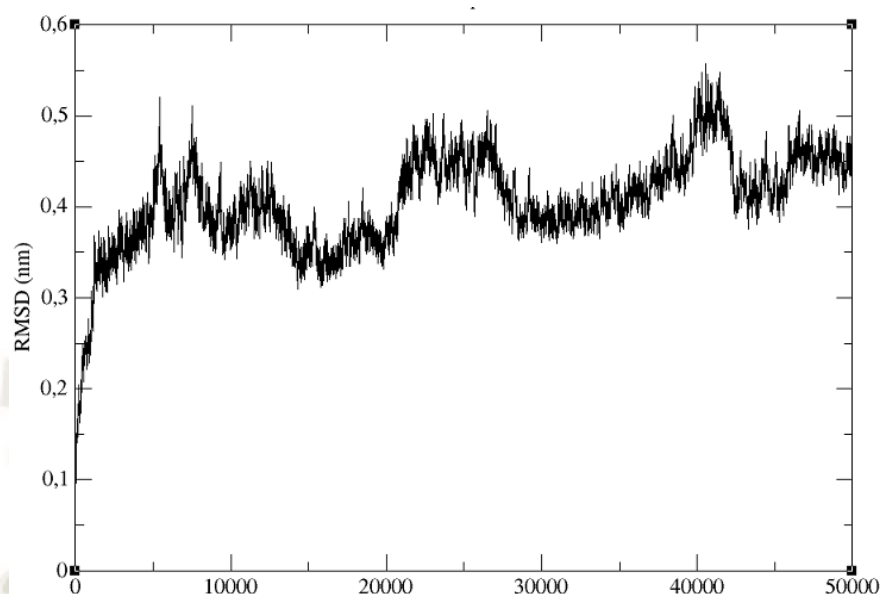
Con el visualizador Chimera se observó la estructura generada por AutoDock y con la herramienta LigPlot se visualizó los residuos de la proteína que interactúan con el flavonoide (Ver Figura 4.28).



**Figura 4.28:** Acoplamiento de PFN2 con Quercetina 3-O-glucósido en el bolsillo 1, y los residuos que interactúan con este compuesto

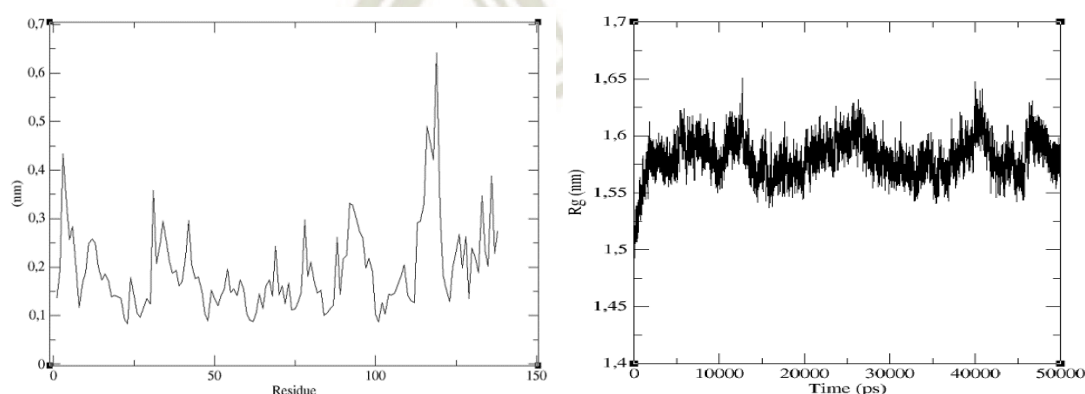
Se pudo observar de la figura anterior que QGL se acopló inicialmente al bolsillo #1 que es uno de los menos favorecidos para unión y por ende de baja drogabilidad, sin embargo, tras la dinámica molecular, los residuos que interaccionaron en este acoplamiento fueron TPR3, CYS25, ALA27, LYS28, GLY 57, PHE58, PHE59, VAL72, ARG74 y ASP75 sin formar algún puente de hidrogeno, aunque siendo algunos de estos residuos los mismos que interaccionan con ISO en el bolsillo #5.

Luego para evaluar el sistema, se hizo uso de las herramientas *rmsd*, *rmsf* y radio de giro para evaluar esta estabilidad en la estructura de nuestra proteína. La herramienta *rms* o *rmsd* calcula la desviación media cuadrática obteniendo un gráfico con la herramienta de ploteo de gráficos Grace, en la gráfica se muestra que la proteína presenta varios picos y se estabiliza en ciertas regiones pero no es consistente en ello, llegando a los 50ns con una estructura ligeramente estable (Ver figura 4.29).



**Figura 4.29:** Ploteo de RMSD tras una dinámica molecular de la proteína PFN2 acoplada con QGL

Luego, se hizo uso de dos herramientas complementarias para la evaluación de la proteína en este sistema. La primera fue *rmsf*, que se utilizó para calcular la fluctuación de la desviación media cuadrática por residuos, pudiendo observar que el cálculo presento un pico que involucran a los residuos GLU116, GLY117, VAL118 y HIS119; mientras que la segunda herramienta utilizada fue el cálculo del radio de giro o *gyrate*, para ver el plegamiento o estiramiento de la proteína tras la interacción. Pudiendo ver que la distancia hacia el extremo más lejano es de aproximadamente 1,57nm y que se estabilizo casi desde el inicio (Ver Figura 4.30).

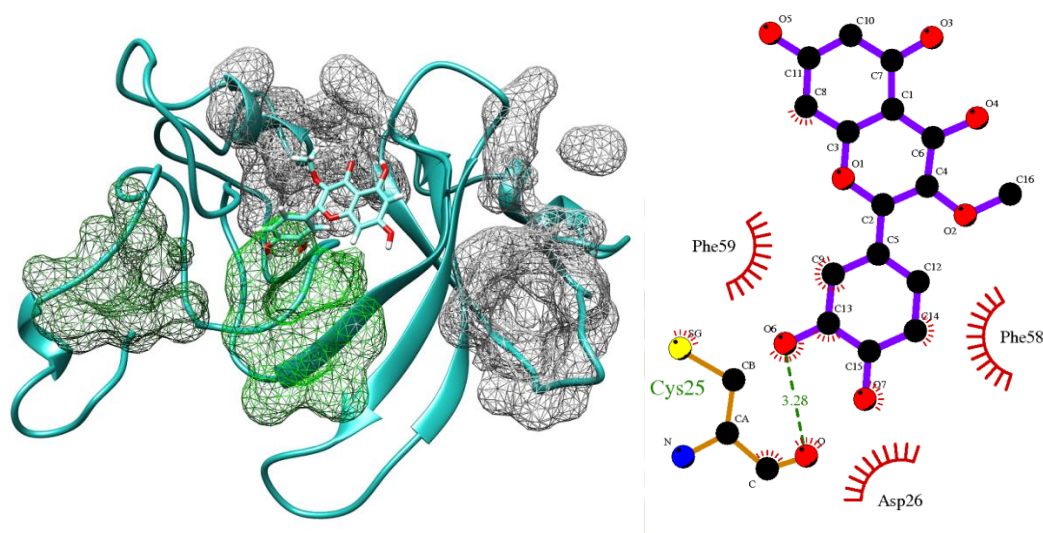


**Figura 4.30:** Ploteos de RMSF y radio de giro respectivamente, de la proteína PFN2 acoplada a QGL



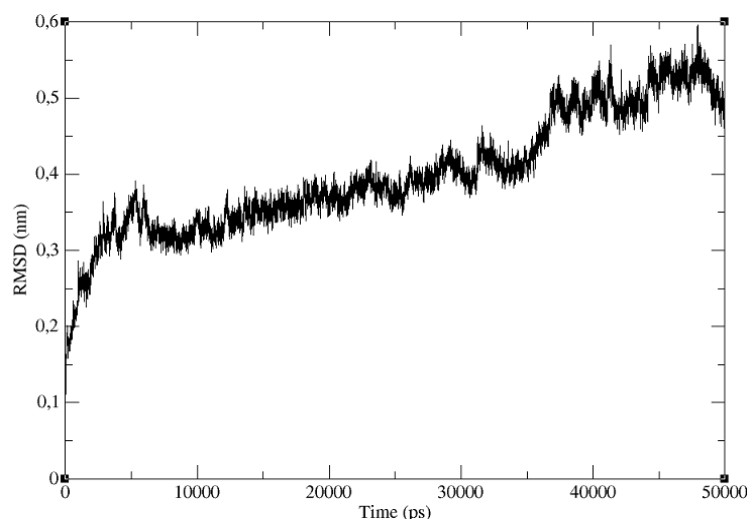
#### 4.4.3. Acoplamiento PFN2 y quercetina 3-metil éter:

Con el visualizador Chimera se observó la estructura generada por AutoDock y con la herramienta LigPlot se visualizó los residuos de la proteína que interactúan con el flavonoide tras la dinámica molecular (Ver Figura 4.31).



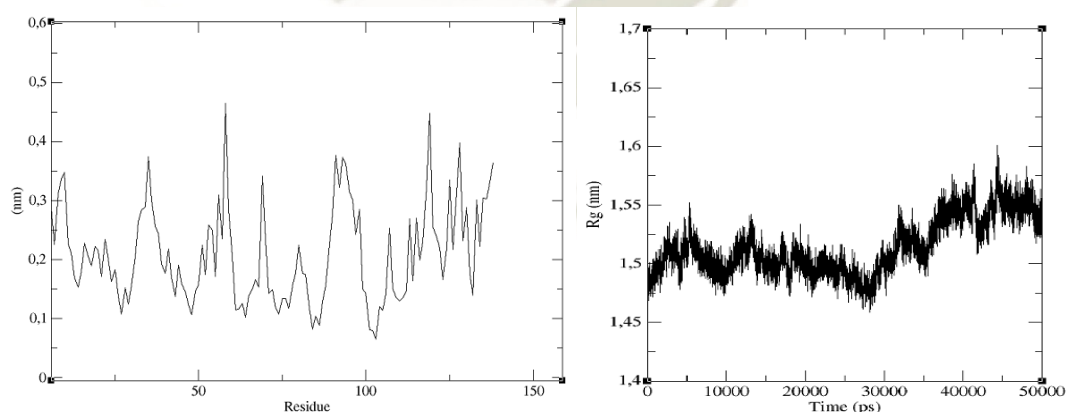
**Figura 4.31:** Acoplamiento de PFN2 con Quercetina 3-metil-éter próximo al bolsillo 6, y los residuos que interactúan con este compuesto en el ploteo

Se puede observar de la figura anterior que QME se acopló inicialmente en el bolsillo #6 resaltado en verde claro, y tras la dinámica molecular interacciona con los residuos ASP26, PHE58, PHE59 y forma un puente de hidrogeno con el residuo CYS25, conservando su lugar inicial de acoplamiento. Luego para evaluar el sistema, se hizo uso de las herramientas *rmsd*, *rmsf* y radio de giro para evaluar esta estabilidad en la estructura de nuestra proteína. La herramienta *rms* o *rmsd* calcula la desviación media cuadrática obteniendo un gráfico con la herramienta de ploteo de gráficos Grace, en la gráfica se muestra que la proteína alcanzó la estabilidad cerca a los 40ns (Ver Figura 4.32).



**Figura 4.32:** Ploteo de RMSD tras una dinámica molecular de la proteína PFN2 acoplada con QME

Luego, se hizo uso de dos herramientas complementarias para la evaluación de la proteína en este sistema. La primera fue *rmsf*, que se utilizó para calcular la fluctuación de la desviación media cuadrática por residuos, pudiendo observar que el cálculo presento varios picos de fluctuación pero ninguno que marcara una diferencia significativa; mientras que la segunda herramienta utilizada fue el cálculo del radio de giro o *gyrate*, para ver el plegamiento o estiramiento de la proteína tras la interacción. Pudiendo ver que la distancia hacia el extremo más lejano es de aproximadamente 1,5nm y que se termina de estabilizar aproximadamente a los 40ns (Ver Figura 4.33).



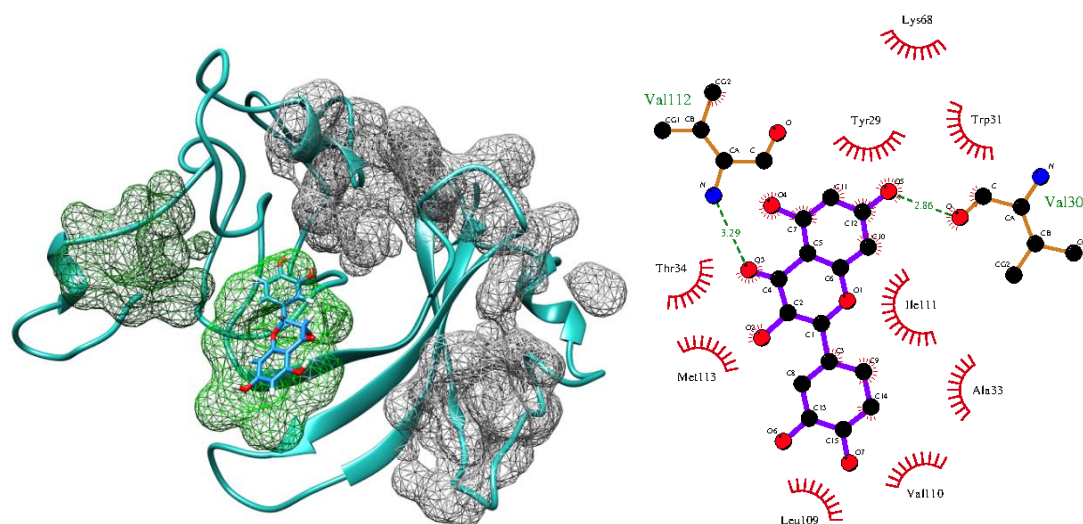
**Figura 4.33:** Ploteos de RMSF y radio de giro respectivamente, de la proteína PFN2 acoplada a QME

#### 4.4.4. Acoplamiento PFN2 y dihidroquercetina:

Con el visualizador Chimera se observó la estructura generada por AutoDock mientras que con la herramienta LigPlot se visualizaron los residuos de la proteína que

interaccionaron con el flavonoide dihidroquercetina tras la dinámica molecular (Ver Figura 4.34).

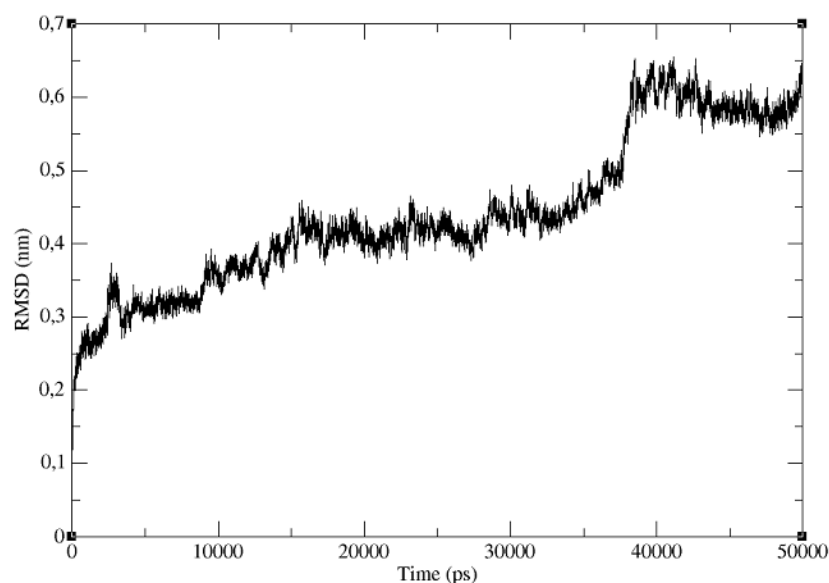
De la figura pudo observarse que DHQ está acoplada en el bolsillo #6 resaltado en verde claro tras el acoplamiento, e interacciona con los residuos TYR29, TRP31, ALA33, THR34, LYS68, LEU109, VAL110, ILE111, MET113 y forma enlace puente de hidrogeno con los residuos VAL30 y VAL112 tras la dinámica molecular, variando la posición de esta molécula y siendo los residuos con los que interactúa pertenecientes al bolsillo #2.



**Figura 4.34:** Acoplamiento de PFN2 con Dihidroquercetina en el bolsillo 6, y los residuos que interactúan con este compuesto

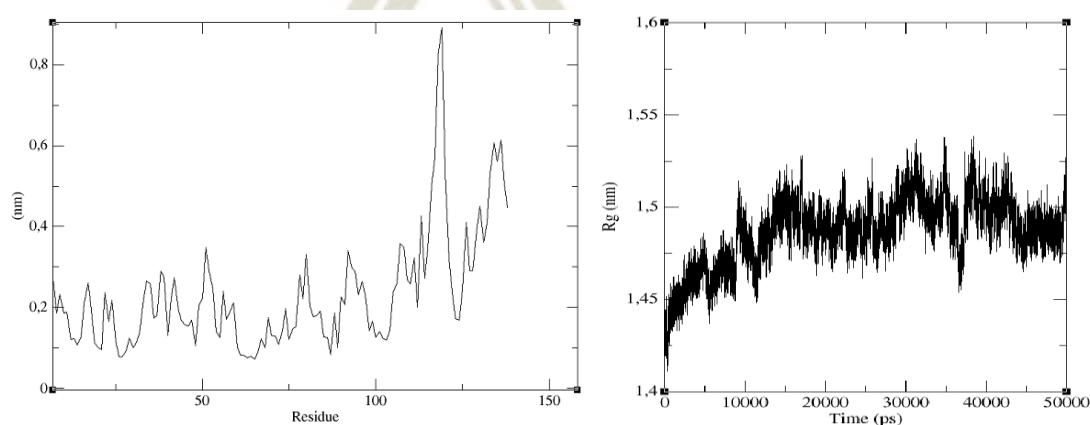
Luego para evaluar el sistema, se hizo uso de las herramientas *rmsd*, *rmsf* y radio de giro para evaluar esta estabilidad en la estructura de nuestra proteína. La herramienta *rms* o *rmsd* calcula la desviación media cuadrática obteniendo un gráfico con la herramienta de ploteo de gráficos Grace, en la gráfica se muestra que la proteína alcanza la estabilidad aproximadamente a los 40ns (Ver Figura 4.35).





**Figura 4.35:** Ploteo de RMSD tras una dinámica molecular de la proteína PFN2 acoplada con DHQ

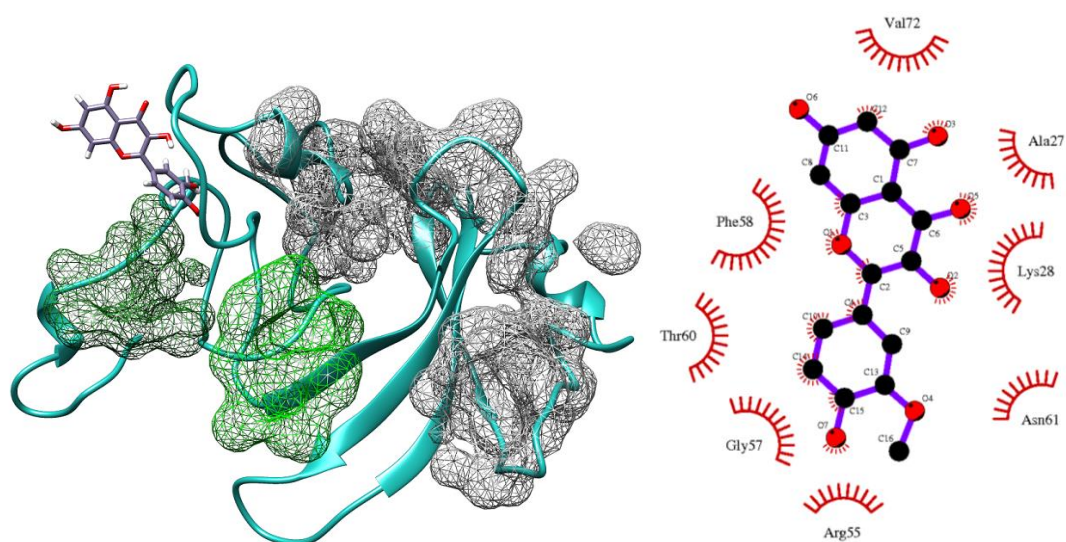
Luego, se hizo uso de dos herramientas complementarias para la evaluación de la proteína en este sistema. La primera fue *rmsf*, que se utilizó para calcular la fluctuación de la desviación media cuadrática por residuos, pudiendo observar que el cálculo presento dos picos que involucran a los residuos GLU116 - GLY120 y TYR133 - SER137 respectivamente; mientras que la segunda herramienta utilizada fue el cálculo del radio de giro o *gyrate*, para ver el plegamiento o estiramiento de la proteína tras la interacción. Pudiendo ver que la distancia hacia el extremo más lejano es de aproximadamente 1,5nm y que se estabilizó pasado los 15ns o 15000ps (Ver Figura 4.36).



**Figura 4.36:** Ploteos de RMSF y radio de giro respectivamente, de la proteína PFN2 acoplada a DHQ

#### 4.4.5. Acoplamiento PFN2 e isorhamnetin:

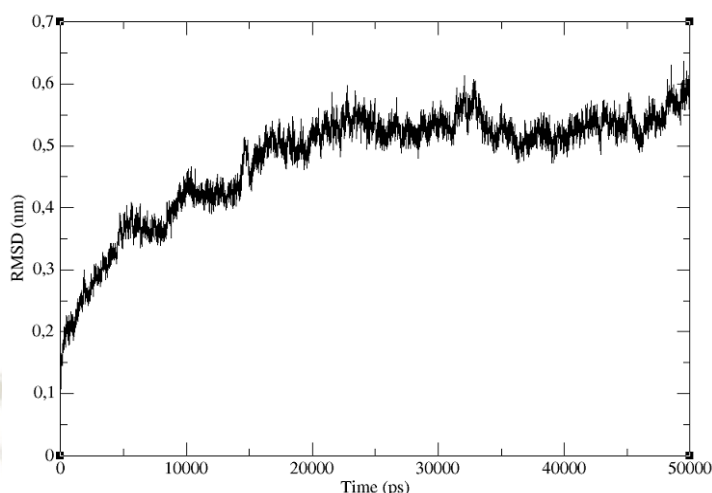
Con el visualizador Chimera se observó la estructura generada por AutoDock y con la herramienta LigPlot se visualizó los residuos de la proteína que interactúan con el flavonoide tras la dinámica molecular (Ver Figura 4.37).



**Figura 4.37:** Acoplamiento de PFN2 con Isorhamnetin en el bolsillo 5, y los residuos que interactúan con este compuesto

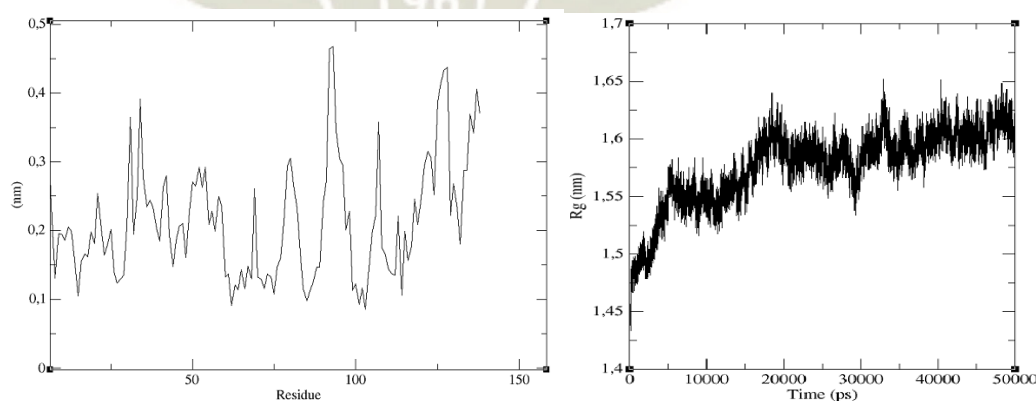
Se pudo observar de la figura anterior que ISO se acopla levemente al bolsillo #5 resaltado en verde oscuro, y tras la dinámica molecular interacciona con los residuos ALA27, LYS28, ARG55, GLY57, PHE58, THR 60, ASN61 y VAL72 sin formar ningún enlace puente de hidrogeno, resaltando su unión al residuo THR60.

Luego para evaluar el sistema, se hizo uso de las herramientas *rmsd*, *rmsf* y radio de giro para evaluar la estabilidad en la estructura de nuestra proteína. La herramienta *rms* o *rmsd* calcula la desviación media cuadrática obteniendo un gráfico con la herramienta de ploteo de gráficos Grace, en la gráfica se muestra que la proteína alcanza la estabilidad aproximadamente a los 20ns (Ver figura 4.38).



**Figura 4.38:** Ploteo de RMSD tras una dinámica molecular de la proteína PFN2 acoplada con ISO

Seguidamente, se utilizaron dos herramientas complementarias para la evaluación de la proteína en este sistema. La primera fue *rmsf*, que se utilizó para calcular la fluctuación de la desviación media cuadrática por residuos, pudiendo observar que el cálculo presento dos picos aunque no muy resaltantes pero se analizaron y se vio que involucran a los residuos GLN92, GLY93 y ALA127, TYR128 respectivamente; mientras que la segunda herramienta utilizada fue el cálculo del radio de giro o *gyrate*, para ver el plegamiento o estiramiento de la proteína tras la interacción. Pudiendo ver que la distancia hacia el extremo más lejano es de aproximadamente 1,6nm y que se termina de estabilizar pasado los 35ns o 35000ps (Ver Figura 4.39).

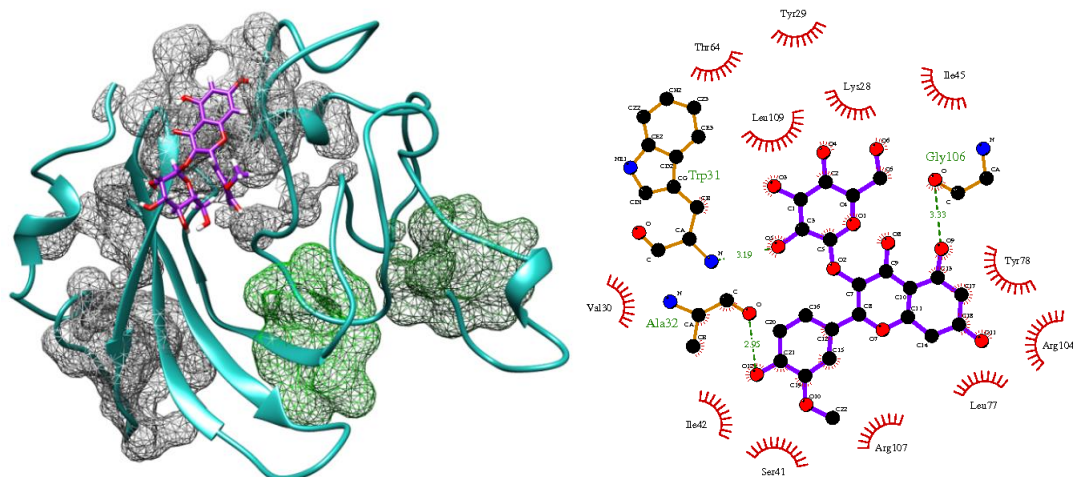


**Figura 4.39:** Ploteos de RMSF y radio de giro respectivamente, de la proteína PFN2 acoplada a ISO



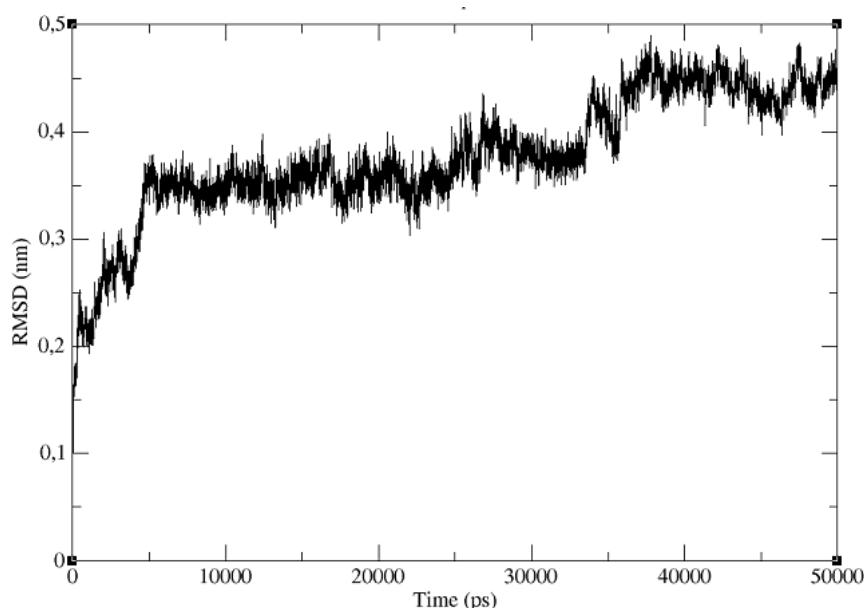
#### 4.4.6. Acoplamiento PFN2 e isorhamnetin 3-O-galactósido:

Con Chimera se observó la estructura generada por AutoDock y con la herramienta LigPlot se visualizó los residuos de la proteína que interactúan con el flavonoide tras la dinámica molecular (Ver Figura 4.40).



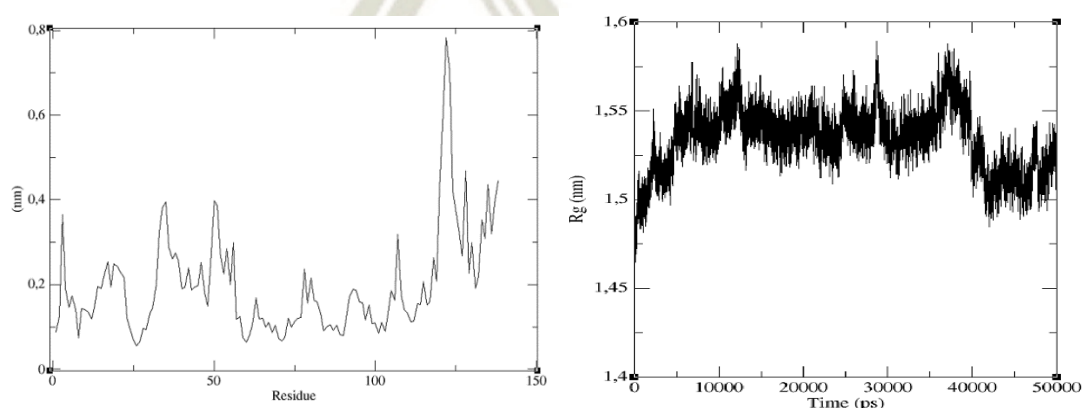
**Figura 4.40:** Acoplamiento de PFN2 con Isorhamnetin 3-O-galactósido en el bolsillo 2, y los residuos que interactúan con este compuesto

Se pudo observar de la figura anterior que IGA está acoplada al bolsillo #2 que aunque no esté resaltado, presentó una probabilidad de drogabilidad de 0.72 siendo bastante significativo, e interacciona con los residuos LYS28, TYR29, VAL30, SER41, ILE42, ILE45, THR64, LEU77, TYR78, ARG104, ARG107, LEU109 y formó 3 enlaces puente de hidrogeno con los residuos TRP31, ALA32 y GLY106 respectivamente. Luego para evaluar el sistema, se hizo uso de las herramientas *rmsd*, *rmsf* y radio de giro para evaluar esta estabilidad en la estructura de nuestra proteína. La herramienta *rms* o *rmsd* calcula la desviación media cuadrática obteniendo un gráfico con la herramienta de ploteo de gráficos Grace, en la gráfica se muestra que la proteína alcanza la estabilidad aproximadamente a los 36ns (Ver Figura 4.41).



**Figura 4.41:** Ploteo de RMSD tras una dinámica molecular de la proteína PFN2 acoplada con IGA

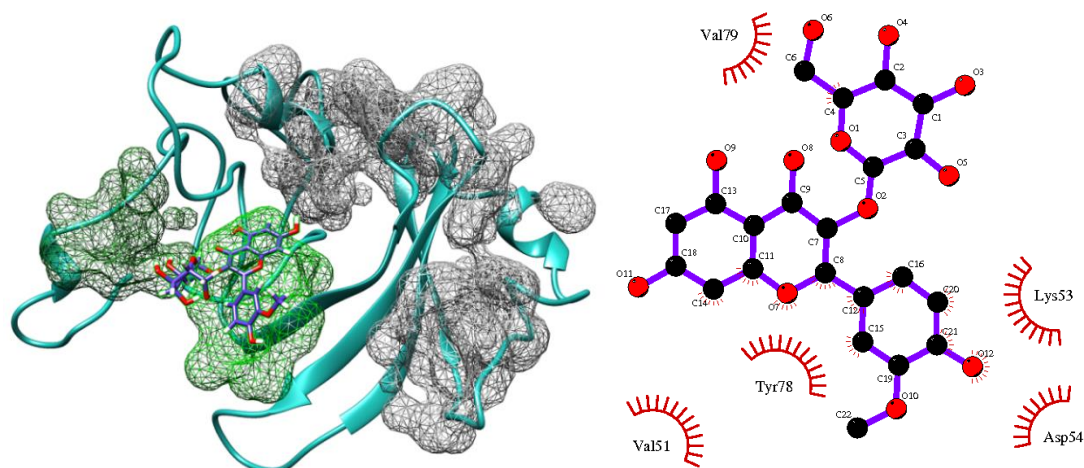
Seguidamente, se hizo uso de dos herramientas complementarias para la evaluación de la proteína en este sistema. La primera fue *rmsf*, que se utilizó para calcular la fluctuación de la desviación media cuadrática por residuos, pudiendo observar que el cálculo presento un pico significativo que involucra a los residuos GLY121, THR122 y LEU123; mientras que la segunda herramienta utilizada fue el cálculo del radio de giro o *gyrate*, para ver el plegamiento o estiramiento de la proteína tras la interacción. Pudiendo ver que la distancia hacia el extremo más lejano es de aproximadamente 1,5nm y que pese al estar estable desde los 10ns hasta los 39ns, muestra una variación hacia los últimos 10ns que es donde se estabiliza finalmente (Ver Figura 4.42).



**Figura 4.42:** Ploteos de RMSF y radio de giro respectivamente, de la proteína PFN2 acoplada a IGA

#### 4.4.7. Acoplamiento PFN2 e isorhamnetin 3-O-glucósido:

Con Chimera se observó la estructura generada por AutoDock y con la herramienta LigPlot se visualizó los residuos de la proteína que interactúan con el flavonoide tras la dinámica molecular (Ver Figura 4.43).

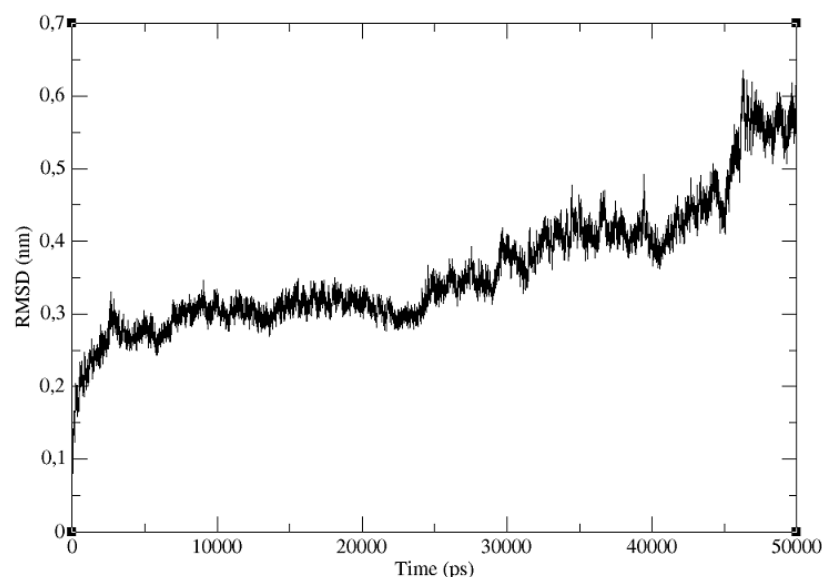


**Figura 4.43:** Acoplamiento de PFN2 con Isorhamnetin 3-O-glucósido en el bolsillo 6, y los residuos que interactúan con este compuesto

Se pudo observar de la figura anterior que IGL está acoplada al bolsillo #6 resaltado en verde claro, pero tras la dinámica molecular interacciona con los residuos VAL51, LYS53, ASP54, TYR78 y VAL79 sin formar ningún enlace puente de hidrogeno y estando estos residuos fuera de cualquier *pocket*.

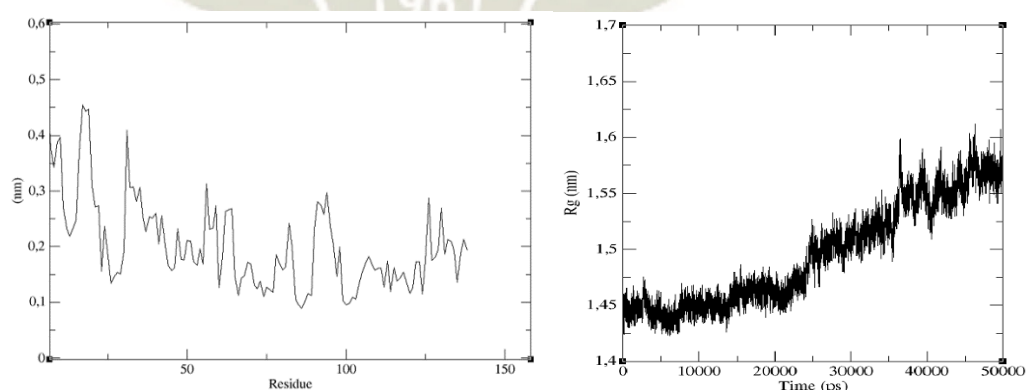
Luego para evaluar el sistema, se hizo uso de las herramientas *rmsd*, *rmsf* y radio de giro para evaluar esta estabilidad en la estructura de nuestra proteína. La herramienta *rms* o *rmsd* calcula la desviación media cuadrática obteniendo un gráfico con la herramienta de ploteo de gráficos Grace, en la gráfica se muestra que la proteína no converge y no halla la estabilidad deseada hasta que llego a los 45ns (Ver Figura 4.44).





**Figura 4.44:** Ploteo de RMSD tras una dinámica molecular de la proteína PFN2 acoplada con IGL

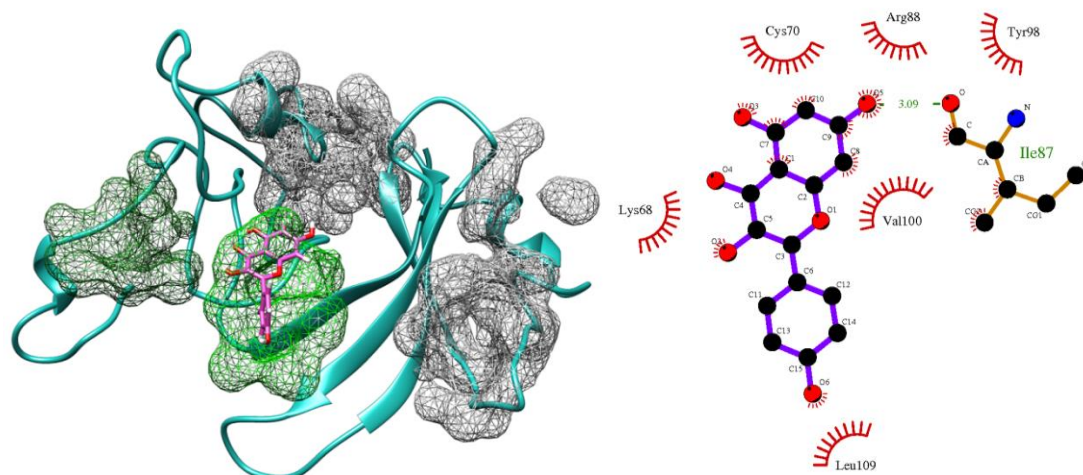
Seguidamente, se hizo uso de dos herramientas complementarias para la evaluación de la proteína en este sistema. La primera fue *rmsf*, que se utilizó para calcular la fluctuación de la desviación media cuadrática por residuos, pudiendo observar que el cálculo presentó un pico que aunque poco significativo involucra a los residuos GLN17, GLU18 y ALA19 estando la fluctuación regularizada; mientras que la segunda herramienta utilizada fue el cálculo del radio de giro o *gyrate*, para ver el plegamiento o estiramiento de la proteína tras la interacción. Pudiendo ver que la distancia hacia el extremo más lejano es de aproximadamente 1,55nm y que se termina de estabilizar pasado los 40ns (Ver Figura 4.45).



**Figura 4.45:** Ploteos de RMSF y radio de giro respectivamente, de la proteína PFN2 acoplada a IGL

#### 4.4.8. Acoplamiento PFN2 y kaempferol:

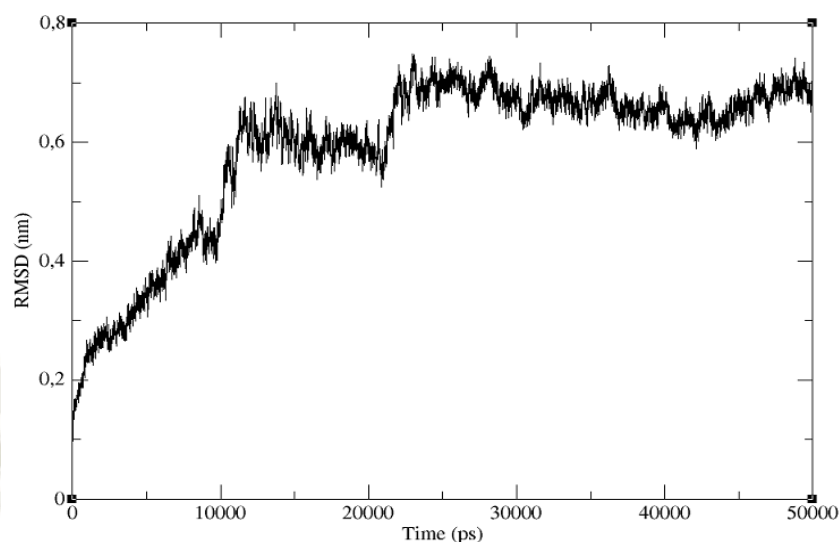
Con el visualizador Chimera se observó la estructura generada por AutoDock y con la herramienta LigPlot se visualizó los residuos de la proteína que interactúan con el flavonoide tras la dinámica molecular (Ver Figura 4.46).



**Figura 4.46:** Acoplamiento de PFN2 con Kaempferol en el bolsillo 6, y los residuos que interactúan con este compuesto

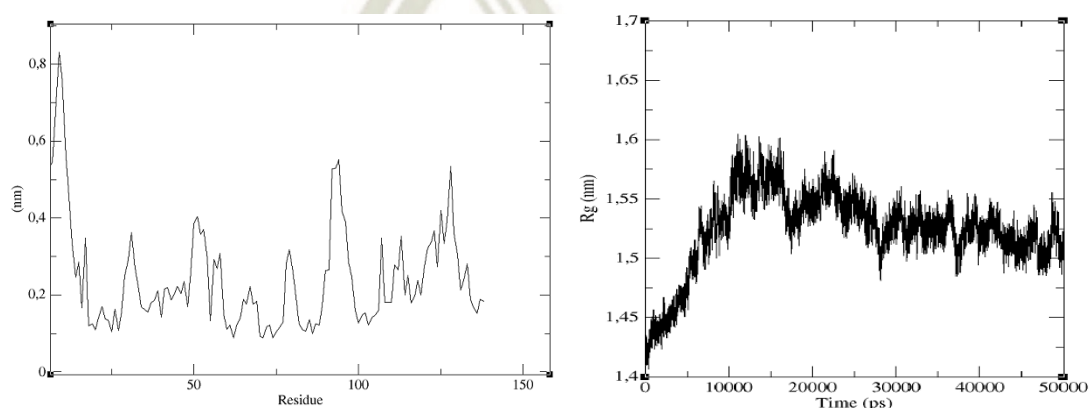
Se pudo observar de la figura anterior que KMP está acoplada al bolsillo #6, y tras la dinámica interacciona con los residuos LYS68, CYS70, ARG88, TYR98, VAL100, LEU109 y formando un enlace puente de hidrogeno con el residuo ILE87, sin embargo, estos residuos no forman parte de ningún *pocket*.

Es así que para evaluar el sistema, se hizo uso de las herramientas *rmsd*, *rmsf* y radio de giro para evaluar esta estabilidad en la estructura de nuestra proteína. La herramienta *rms* o *rmsd* calcula la desviación media cuadrática obteniendo un gráfico con la herramienta de ploteo de gráficos Grace, en la gráfica se muestra que la proteína alcanza la estabilidad aproximadamente a los 23ns (Ver Figura 4.47).



**Figura 4.47:** Ploteo de RMSD tras una dinámica molecular de la proteína PFN2 acoplada con KMP

Luego, se hizo uso de dos herramientas complementarias para la evaluación de la proteína en este sistema. La primera fue *rmsf*, que se utilizó para calcular la fluctuación de la desviación media cuadrática por residuos, pudiendo observar que el cálculo presento tres picos que involucran a los residuos ALA1 –MET11, GLN92 – GLY94 y TYR128 respectivamente; mientras que la segunda herramienta utilizada fue el cálculo del radio de giro o *gyrate*, para ver el plegamiento o estiramiento de la proteína tras la interacción. Pudiendo ver que la distancia hacia el extremo más lejano es de aproximadamente 1,52nm y que se termina de estabilizar pasado los 30ns (Ver Figura 4.48).

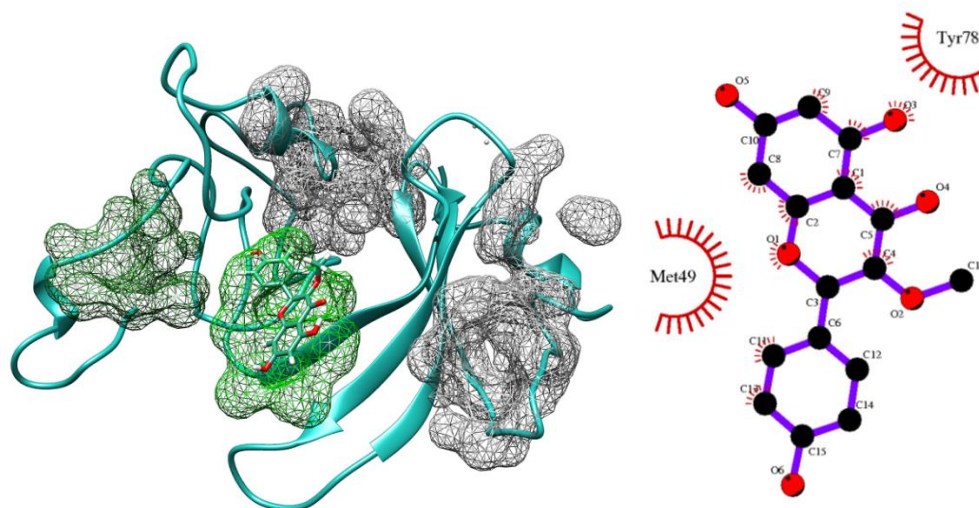


**Figura 4.48:** Ploteos de RMSF y radio de giro respectivamente, de la proteína PFN2 acoplada a KMP



#### 4.4.9. Acoplamiento PFN2 y kaempferol 3-metil éter:

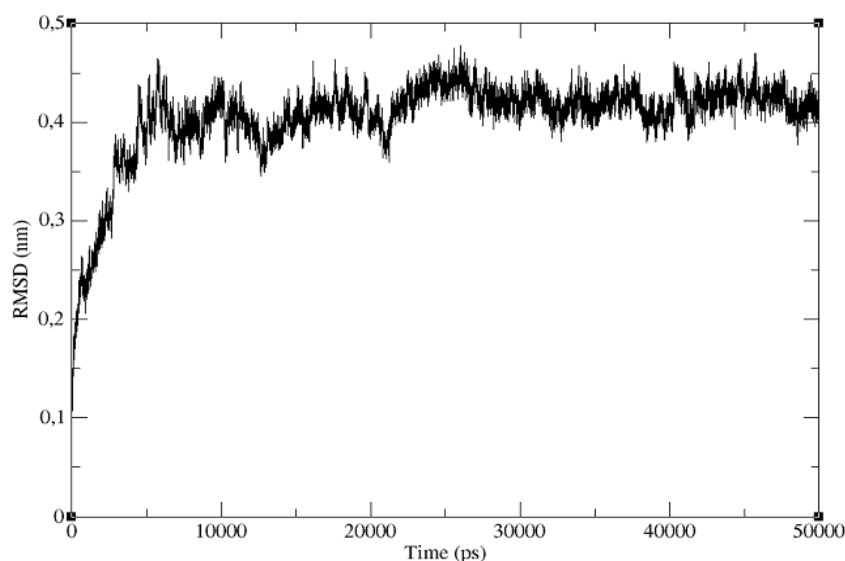
Con el visualizador Chimera se observó la estructura generada por AutoDock y con la herramienta LigPlot se visualizó los residuos de la proteína que interactúan con el flavonoide tras la dinámica (Ver Figura 4.49).



**Figura 4.49:** Acoplamiento de PFN2 con Kaempferol 3-methyl ether en el bolsillo 6, y los residuos que interactúan con este compuesto

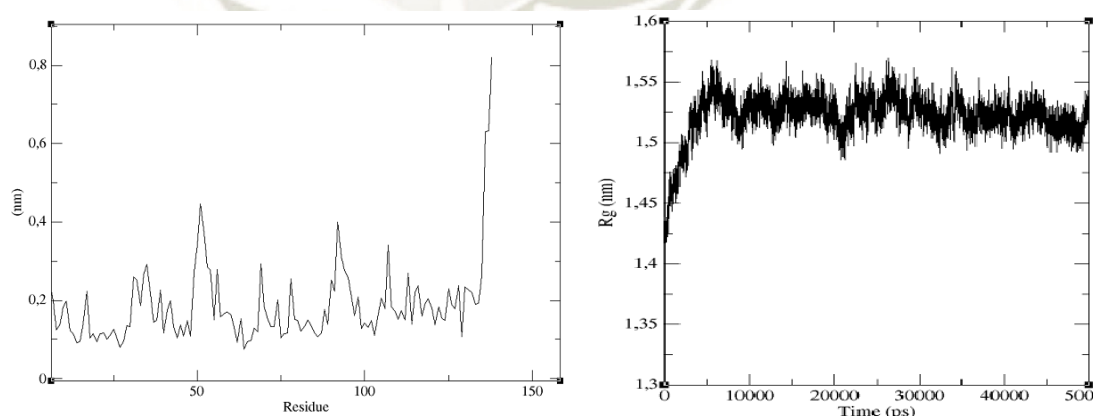
Se pudo observar de la figura anterior que KME está acoplada al bolsillo #6, y tras la dinámica molecular interacciona con los residuos MET49 y TYR78 sin formar ningún enlace puente de hidrógeno y sin ser parte de un *pocket*.

Para evaluar el sistema, se hizo uso de las herramientas *rmsd*, *rmsf* y radio de giro para evaluar esta estabilidad en la estructura de nuestra proteína. La herramienta *rms* o *rmsd* calcula la desviación media cuadrática obteniendo un gráfico con la herramienta de ploteo de gráficos Grace, en la gráfica se muestra que la proteína alcanza la estabilidad aproximadamente a los 5ns (Ver figura 4.50).



**Figura 4.50:** Ploteo de RMSD tras una dinámica molecular de la proteína PFN2 acoplada con KME

Luego, se hizo uso de dos herramientas complementarias para la evaluación de la proteína en este sistema. La primera fue *rmsf*, que se utilizó para calcular la fluctuación de la desviación media cuadrática por residuos, pudiendo observar que el cálculo presento un pico bastante prominente que involucra a los residuos ARG136 – ASP138; mientras que la segunda herramienta utilizada fue el cálculo del radio de giro o *gyrate*, para ver el plegamiento o estiramiento de la proteína tras la interacción. Pudiendo ver que la distancia hacia el extremo más lejano es de aproximadamente 1,52nm y que se termina de estabilizar casi desde un inicio a los 5ns (Ver Figura 4.51).



**Figura 4.51:** Ploteos de RMSF y radio de giro respectivamente, de la proteína PFN2 acoplada a KME

#### 4.5. Cálculo de la Energía Libre de Acoplamiento:

Teniendo ya sistemas conformados tanto por la profilina humana II como cada uno de los compuestos flavonoides y que han pasado por un proceso de dinámica molecular en un tiempo de 50ns en fase solvente con un modelo de aguas SPC, se pudo evaluar cómo evoluciona el sistema a través del tiempo, es decir, se evaluó la interacción entre la proteína y los respectivos ligando en un ambiente más cercano al *in vivo* y se pudo evaluar si el compuesto flavonoide permanecía en el lugar de acoplamiento original, si este interactuaba con otros residuos o si finalmente no tenía interacción alguna.

Terminado este proceso, todas las moléculas mostraron poder mantener una conformación que alcanzaba la estabilidad con la proteína según se pudo visualizar estos sistemas con el software Chimera, pese a ello, no todas las moléculas permanecieron en los lugares originales de acoplamiento y sus interacciones con los residuos descritos fueron evaluadas de manera más precisa mediante el cálculo de la variación de energía libre de estos sistemas bajo un colectivo isobárico-isotérmico en el cual gradualmente se restringe las interacciones no enlazantes como las de Van der Waals y Coulómbicas. Este cálculo de la energía libre de Gibbs ( $\Delta G$ ) se llevó a cabo con el módulo BAR (Bennet Acceptance Ratio) que calcula las diferencias de energía libre entre estado A y estado B ( $\Delta G_{AB}$ ) de y sistema con un parámetro de desacoplamiento  $\lambda$  con valores de espaciado intermedios de 0.025 para nuestro estudio que restringió gradualmente las interacciones no enlazantes ya mencionadas, llegando a obtener los valores reportados en la tabla 4.4.

**Tabla 4.4:** Cálculo de la variación de energía libre de Gibbs ( $\Delta G$ ) en un sistema que va conformando dos estados progresivamente según el módulo BAR.

Compuesto	Abreviación	NPT	
		Delta G (Kcal/mol)	Delta G (KJ/mol)
1 PFN2 - Quercetin	PQCT	-9.467 +/- 0.177	-39.61 +/- 0.74
2 PFN2 - Quercetin 3-O-glucoside	PQGL	-17.22 +/- 0.88	-72.05 +/- 3.69
3 PFN2 - Quercetin 3-methyl ether	PQME	-6.582 +/- 0.17	-27.54 +/- 0.72
4 PFN2 - (+) - Dihydroquercetin	PDHQ	-16.89 +/- 1.097	-70.67 +/- 4.59
5 PFN2 - Isorhamnetin	PISO	-12.19 +/- 0.75	-51.04 +/- 3.12
6 PFN2 - Isorhamnetin 3-O-Galactoside	PIGA	-18.40 +/- 0.87	-77.00 +/- 3.64
7 PFN2 - Isorhamnetin 3-O-Glucoside	PIGL	-10.327 +/- 0.547	-43.21 +/- 2.29
8 PFN2 - Kaempferol	PKMP	-8.045 +/- 0.31	-33.66 +/- 1.30
9 PFN2 - Kaempferol 3-methyl ether	PKME	-5.559 +/- 0.38	-23.26 +/- 1.58



De la tabla 4.4, resaltaron los valores negativos más elevados que a su vez pertenecen a los complejos de PFN2 con los flavonoides IGA, QGL, DHQ e ISO, siendo estos aquellos cuyos residuos de interacción pertenecían a los bolsillos con mejor score de drogabilidad en la proteína; además, se incluyó al compuesto QME que pese a tener un valor negativo más bajo también interactuaba con estos residuos.

#### 4.6. Interacción Proteína - Proteína:

Ambas proteínas, tanto profilina humana II (PFN2) e histona desacetilasa 1 (HDAC1) pasaron por un proceso de optimización mediante dinámica molecular presentando tanto estabilidad energética como geométricamente teniendo archivos finales .gro en cada caso que han de convertirse a archivos .pdb. Adicionalmente, la profilina humana II fue sometida a un proceso de acoplamiento con 9 diferentes compuestos flavonoides y una posterior dinámica molecular para evaluar su interacción en un medio solvatado tras el cual se evalúa la capacidad de unión de estos compuestos hacia la proteína y el efecto que tienen sobre su conformación estructural y energética; es a partir de estos cambios que se evalúa si la profilina humana II modificada aún mantiene su capacidad de unión con histona desacetilasa 1 (HDAC1) tal cual la profilina humana II nativa mediante una interacción proteína-proteína.

##### 4.6.1. Acoplamiento Molecular:

El acoplamiento molecular proteína-proteína se llevó a cabo con el servidor ClusPro, el cual requiere las coordenadas tridimensionales de las proteínas en archivos con extensión .pdb para llevar a cabo los tres pasos necesarios para completar el acoplamiento y el análisis energético a partir de las mejores conformaciones agrupadas a las cuales les sigue un proceso de minimización, obteniendo resultados y estructuras; además, este servidor proporciona diferentes valores según el termino de energía se desee favorecer según la siguiente ecuación:

$$E = w_1 E_{rep} + w_2 E_{attr} + w_3 E_{elec} + w_4 E_{DARS}$$

De donde  $E_{rep}$  y  $E_{attr}$  refiere a las energías de Van der Waals de repulsión y de atracción respectivamente, mientras que  $E_{elec}$  es un término electrostático y finalmente  $E_{DARS}$  (decoys as the reference state o DARS) es una aproximación para la construcción de estructuras en base a los potenciales intermoleculares y que en este caso está

representado por el cambio en la energía libre a partir de la desolvatación o remoción de las moléculas del solvente. Por otra parte, los coeficientes  $w_1$ ,  $w_2$ ,  $w_3$  y  $w_4$  corresponden a los valores que están asociados a cada término de energía según el que se vaya a favorecer dependiendo del tipo de sistema en estudio (Ver Tabla 4.5).

**Tabla 4.5:** Valores de coeficientes energéticos según el termino favorecido para el servidor ClusPro

Términos Energéticos	Valores de los Coeficientes Energéticos			
	$E_{rep}$	$E_{attr}$	$E_{elec}$	$E_{DARS}$
Balanced	0.40	-0.40	600.0	1.0
Electrostatic-favored	0.40	-0.40	1200.0	1.0
Hydrophobic-favored	0.40	-0.40	600.0	2.0
Van der Waals + Electrostatics	0.40	-0.40	600.0	0.0

Los coeficientes  $w_1$ ,  $w_2$ ,  $w_3$  y  $w_4$  asumen según la tabla los valores 0.40, -0.40, 600 y 2 respectivamente favoreciendo una interacción hidrofóbica, este término energético refleja el hecho de que en un sistema, mientras mayor sea la interfaz agua-soluto más inestable es el sistema, por ende, favorece a una reducción de esta interfaz dirigiendo el sistema hacia el equilibrio o en otras palabras favorece la interfaz soluto-soluto que serían las macromoléculas en estudio HDAC1 con PFN2 nativa o modificada. De allí, los resultados obtenidos del acoplamiento de ambas proteínas (Ver Tabla 4.6).

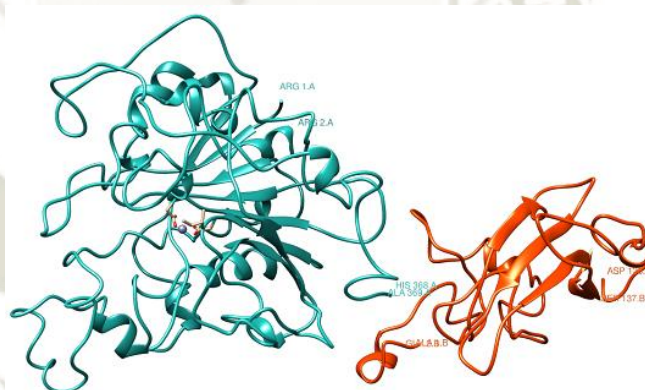
**Tabla 4.6:** Valores del acoplamiento HDAC1 y PFN2 nativa y modificadas en servidor ClusPro con los términos de hidrofobicidad favorecidos

Hydrophobic-Favored		
Compuestos Acoplados	Representación	Puntaje Ponderado
HDAC-PFN	Centro	-990.3
	Menor Energía	-1097
HDAC-PQGL	Centro	-1013.8
	Menor Energía	-1036.1
HDAC-PQME	Centro	-989.8
	Menor Energía	-1219.1
HDAC-PDHQ	Centro	-952.8
	Menor Energía	-1169.9
HDAC-PISO	Centro	-1187.4
	Menor Energía	-1187.4
HDAC-PIGA	Centro	-1050.1
	Menor Energía	-1255.4

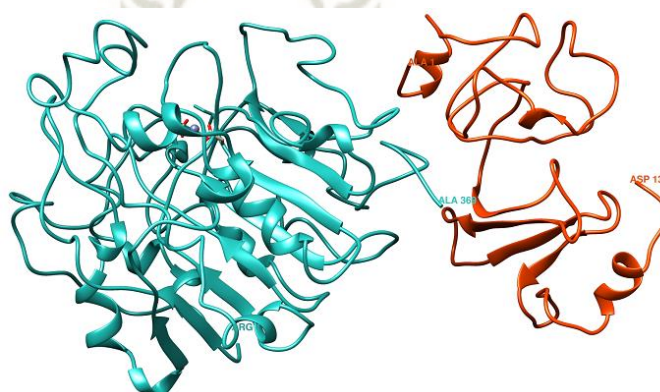


De la tabla podemos inferir que existen dos tipos de representación: centro y menor energía, estas corresponden a los agrupamientos o *clusters*, siendo centro referido al *cluster* o grupo con mayor número de conformaciones que son estructuralmente similares entre sí; mientras que menor energía refiere a la conformación con la menor energía encontrada; siendo estadísticamente más relevante los valores del término centro por el número de conformaciones pertenecientes al *cluster*. A su vez, del mismo servidor ClusPro, se pudo descargar las estructuras tridimensionales de estos complejos proteína-proteína formados (Ver Figuras 4.52-4.57).

Además, de esta tabla encontramos que los valores energéticos en el primer compuesto refieren a la interacción entre HDAC1 con PFN2 nativa mientras los demás presentan la PFN2 modificada; aquellos como es el caso de PQME y PDHQ presentan valores negativos menores en comparación con aquel reportado por la profilina nativa, mientras que PIGA, PISO y PQGL un valor negativo por encima del ya mencionado.

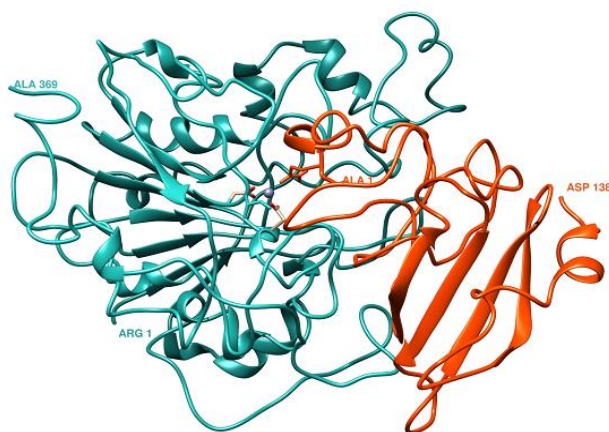


**Figura 4.52:** Acoplamiento molecular de HDAC1 y PFN2 nativa resaltando sus extremos N- y C-terminales



**Figura 4.53:** Acoplamiento molecular de HDAC1 y PFN2 modificada por interacción con el compuesto QGL

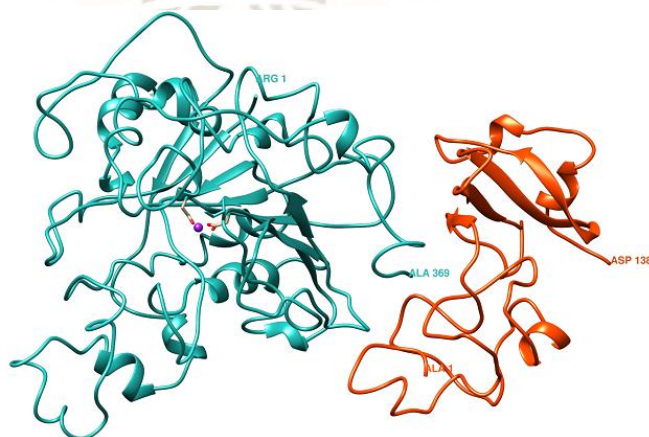




**Figura 4.54:** Acoplamiento molecular de HDAC1 y PFN2 modificada por interacción con el compuesto QME



**Figura 4.55:** Acoplamiento molecular de HDAC1 y PFN2 modificada por interacción con el compuesto DHQ



**Figura 4.56:** Acoplamiento molecular de HDAC1 y PFN2 modificada por interacción con el compuesto ISO



**Figura 4.57:** Acoplamiento molecular de HDAC1 y PFN2 modificada por interacción con el compuesto IGA

#### 4.6.2. Refinamiento de la Interacción del Acoplamiento:

Las estructuras acopladas fueron sometidas al servidor FireDock para un refinamiento energético del acoplamiento que se realiza en dos etapas de cálculo rápido: la primera etapa es un ajuste de las cadenas laterales de los residuos de la proteína buscando una geometría óptima, luego un ajuste de la orientación de las proteínas en el complejo que previamente restringe las nuevas posiciones de las cadenas laterales y genera ligeras rotaciones o movimientos en el *backbone* de la proteína; finalmente este servidor brinda energías de acoplamiento con alto porcentaje de fiabilidad y una categorización de las conformaciones según estas energías (Ver Tabla 4.7).

**Tabla 4.7:** Valores energéticos calculados con el servidor de refinamiento FireDock

Compuestos	FireDock			
	GE or BE	aVdW	rVdW	ACE
HPFN	-64.04	-40.24	10.83	-14.03
HPQGL	-27.59	-33.56	36.55	-10.42
HPQME	-114.02	-51.56	7.57	-21.83
HPDHQ	-85.47	-62.63	67.46	-13.01
HPISO	-80.68	-46.19	15.24	-13.51
HPIGA	-116.21	-52.81	4.64	-15.04



En la tabla podemos visualizar los términos de evaluación como: energía global o energía de acoplamiento (GE or BE), energía de Van der Waals de atracción (AVdW), energía de Van der Waals de repulsión (rVdW) y energía de contacto atómico (ACE), siendo más significativos los valores de GE y ACE. Es así que, se puede inferir que el compuesto HPQGL presentó el menor valor negativo en comparación con las otras conformaciones incluida la conformación que incluye a la profilina humana II nativa. Por ello podemos afirmar que la profilina humana II modificada por efecto de la interacción con el compuesto flavonoide QGL requiere de un mayor gasto energético para su interacción y formación de un complejo con la enzima histona desacetilasa 1, lo que es un indicativo de que este compuesto flavonoide QGL desestabiliza estructural y/o energéticamente de manera significativa a la profilina humana II, disminuyendo su capacidad de unión a histona desacetilasa 1. Adicionalmente los compuestos DHQ e ISO acoplados a PFN2, también presentaron valores de energía de contacto atómico (ACE) significativos para este estudio, en su unión a HDAC1.



## CONCLUSIONES:

1. La estructura cristalizada de PFN2 fue optimizada teniendo un monómero de 138 residuos; mientras que la estructura cristalizada de HDAC1 fue optimizada junto con una molécula de Zinc +2 que representa su sitio activo.
2. Los 11 compuestos flavonoides reportados fueron dibujados y optimizados geométricamente mediante técnicas de mecánica cuántica, a estos también se les generó su correspondiente topología para el campo de fuerza OPLS-AA. Aunque los compuestos QRU e IRO no cumplieron con las cinco reglas de Lipinski por lo cual fueron descartadas del estudio.
3. Los compuestos flavonoides QGL, QME, DHQ, ISO e IGA presentaron una interacción con la profilina humana II en los *pockets* con mayor porcentaje de drogabilidad tras los procesos de acoplamiento, dinámica molecular y cálculo de la energía libre de interacción; además de interactuar con residuos pertenecientes a las zonas activas de esta proteína.
4. Los compuestos flavonoides quercetina 3-O-glucósido o QGL, dihidroquercetina o DHQ, e isorhamnentin o ISO son aquellos que reportan un cambio significativo sobre la profilina humana II, que acarrea un mayor requerimiento energético para el acoplamiento e interacción de ésta con la enzima histona desacetilasa 1.

## RECOMENDACIONES:

1. El tiempo de dinámica molecular ejecutado para las proteínas PFN2 y HDAC1 puede ser extendido a 100ns para confirmar la convergencia de estos sistemas al equilibrio.
2. La estructura cristalizada de profilina humana II carece de dos residuos en comparación con su secuencia genómica; ambos se encuentran en cada extremo de la cadena, y son los aminoácidos metionina y valina que no son significativos pero pueden ser completados mediante homología de secuencias.
3. Se recomienda el diseño de un nanotransportador que tenga como característica la especificidad por estos compuestos flavonoides y por PFN2.

## REFERENCIAS BIBLIOGRÁFICAS:

1. Hayashi E, Kuramitsu Y, Okada F, Fujimoto M, Zhang X, Kobayashi M, et al. Proteomic profiling for cancer progression: Differential display analysis for the expression of intracellular proteins between regressive and progressive cancer cell lines. *Proteomics*. 2005;5(4):1024-32.
2. Coleman WB, Tsongalis GJ. *The molecular basis of human cancer*: Springer Science & Business Media; 2001.
3. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *cell*. 2011;144(5):646-74.
4. Rodríguez JC, Aranda ICB, Torres JMS. *Cáncer de pulmón*.
5. Kenfield SA, Wei EK, Stampfer MJ, Rosner BA, Colditz GA. Comparison of aspects of smoking among the four histological types of lung cancer. *Tobacco control*. 2008;17(3):198-204.
6. Wang G, Reed E, Li QQ. Molecular basis of cellular response to cisplatin chemotherapy in non-small cell lung cancer. *Oncology reports*. 2004;12(5):955-65.
7. Toloza EM, Harpole L, McCrory DC. Noninvasive staging of non-small cell lung cancer: a review of the current evidence. *CHEST Journal*. 2003;123(1\_suppl):137S-46S.
8. Beland MD, Wasser EJ, Mayo-Smith WW, Dupuy DE. Primary non-small cell lung cancer: review of frequency, location, and time of recurrence after radiofrequency ablation. *Radiology*. 2009;254(1):301-7.
9. Tang Y-N, Ding W-Q, Guo X-J, Yuan X-W, Wang D-M, Song J-G. Epigenetic regulation of Smad2 and Smad3 by profilin-2 promotes lung cancer growth and metastasis. *Nature communications*. 2015;6.
10. Meulmeester E, Ten Dijke P. The dynamic roles of TGF- $\beta$  in cancer. *The Journal of pathology*. 2011;223(2):206-19.
11. Gálvez-Gastélum FJ, Sandoval-Rodríguez AS, Armendáriz-Borunda J. El factor de crecimiento transformante beta como blanco terapéutico. *salud pública de méxico*. 2004;46(4):341-50.
12. Lönn P, Morén A, Raja E, Dahl M, Moustakas A. Regulating the stability of TGF $\beta$  receptors and Smads. *Cell research*. 2009;19(1):21-35.
13. Izzi L, Attisano L. Regulation of the TGF $\beta$  signalling pathway by ubiquitin-mediated degradation. *Oncogene*. 2004;23(11):2071-8.
14. Heldin C-H, Miyazono K, Ten Dijke P. TGF- $\beta$  signalling from cell membrane to nucleus through SMAD proteins. *Nature*. 1997;390(6659):465-71.
15. Bellomo C, Caja L, Moustakas A. Transforming growth factor  $\beta$  as regulator of cancer stemness and metastasis. *British journal of cancer*. 2016.
16. Miyazono K, Kamiya Y, Miyazawa K. SUMO amplifies TGF- $\beta$  signalling. *nature cell biology*. 2008;10(6):635.
17. Cerutti J, Ebina K, Matsuo S, Martins L, Maciel R, Kimura E. Expression of Smad4 and Smad7 in human thyroid follicular carcinoma cell lines. *Journal of endocrinological investigation*. 2003;26(6):516-21.
18. Derynck R, Zhang Y, Feng X-H. Transcriptional activators of TGF- $\beta$  responses: Smads. *Cell*. 1998;95(6):737-40.
19. Shi Y, Hata A, Lo RS, Massague J, Pavletich NP. A structural basis for mutational inactivation of the tumour suppressor Smad4. *Nature*. 1997;388(6637):87-93.
20. Massagué J, Seoane J, Wotton D. Smad transcription factors. *Genes & development*. 2005;19(23):2783-810.



21. Acharya MR, Sparreboom A, Venitz J, Figg WD. Rational development of histone deacetylase inhibitors as anti-cancer agents: a review. *Molecular pharmacology*. 2005.
22. Tabraue Chávez M. Diseño y síntesis de derivados del ácido hidroxámico como nuevos inhibidores de las histonas desacetilasas con actividad antitumoral. 2006.
23. Marks PA, Richon VM, Rifkind RA. Histone deacetylase inhibitors: inducers of differentiation or apoptosis of transformed cells. *Journal of the National Cancer Institute*. 2000;92(15):1210-6.
24. Dokmanovic M, Clarke C, Marks PA. Histone deacetylase inhibitors: overview and perspectives. *Molecular cancer research*. 2007;5(10):981-9.
25. Vigushin D, Coombes R. Targeted histone deacetylase inhibition for cancer therapy. *Current cancer drug targets*. 2004;4(2):205-18.
26. Glaser KB, Staver MJ, Waring JF, Stender J, Ulrich RG, Davidsen SK. Gene expression profiling of multiple histone deacetylase (HDAC) inhibitors: defining a common gene set produced by HDAC inhibition in T24 and MDA carcinoma cell lines. *Molecular cancer therapeutics*. 2003;2(2):151-63.
27. Bahr JC, Robey RW, Luchenko V, Basseville A, Chakraborty AR, Kozlowski H, et al. Blocking downstream signaling pathways in the context of HDAC inhibition promotes apoptosis preferentially in cells harboring mutant Ras. *Oncotarget*. 2016;7(43):69804.
28. Drummond DC, Noble CO, Kirpotin DB, Guo Z, Scott GK, Benz CC. Clinical development of histone deacetylase inhibitors as anticancer agents. *Annu Rev Pharmacol Toxicol*. 2005;45:495-528.
29. Shieh JM, Tang YA, Hu FH, Huang WJ, Wang YJ, Jen J, et al. A histone deacetylase inhibitor enhances expression of genes inhibiting Wnt pathway and augments activity of DNA demethylation reagent against non-small-cell lung cancer. *International Journal of Cancer*. 2017;140(10):2375-86.
30. Chueh AC, Tse JW, Tögel L, Mariadason JM. Mechanisms of histone deacetylase inhibitor-regulated gene expression in cancer cells. *Antioxidants & redox signaling*. 2015;23(1):66-84.
31. Adeegbe D, Liu Y, Lizotte PH, Kamihara Y, Aref AR, Almonte C, et al. Synergistic Immunostimulatory Effects and Therapeutic Benefit of Combined Histone Deacetylase and Bromodomain Inhibition in Non-small Cell Lung Cancer. *Cancer Discovery*. 2017:CD-16-1020.
32. Miller TA, Witter DJ, Belvedere S. Histone deacetylase inhibitors. *Journal of medicinal chemistry*. 2003;46(24):5097-116.
33. De Ruijter AJ, Van Gennip AH, Caron HN, Kemp S, Van Kuilenburg AB. Histone deacetylases (HDACs): characterization of the classical HDAC family. *Biochemical Journal*. 2003;370(Pt 3):737.
34. Millard CJ, Watson PJ, Celardo I, Gordiyenko Y, Cowley SM, Robinson CV, et al. Class I HDACs share a common mechanism of regulation by inositol phosphates. *Molecular cell*. 2013;51(1):57-67.
35. Nodelman IM, Bowman GD, Lindberg U, Schutt CE. X-ray structure determination of human profilin II: A comparative structural analysis of human profilins. *Journal of molecular biology*. 1999;294(5):1271-85.
36. Cheng Y-J, Zhu Z-X, Zhou J-S, Hu Z-Q, Zhang J-P, Cai Q-P, et al. Silencing profilin-1 inhibits gastric cancer progression via integrin  $\beta 1$ /focal adhesion kinase pathway modulation. *World Journal of Gastroenterology: WJG*. 2015;21(8):2323.
37. Witke W. The role of profilin complexes in cell motility and other cellular processes. *Trends in cell biology*. 2004;14(8):461-9.
38. Coumans JV, Gau D, Poljak A, Wasinger V, Roy P, Moens PD. Profilin-1 overexpression in MDA-MB-231 breast cancer cells is associated with alterations in proteomics biomarkers

of cell proliferation, survival, and motility as revealed by global proteomics analyses. *Omics: a journal of integrative biology*. 2014;18(12):778-91.

39. Cui X-b, Zhang S-m, Xu Y-x, Dang H-w, Liu C-x, Wang L-h, et al. PFN2, a novel marker of unfavorable prognosis, is a potential therapeutic target involved in esophageal squamous cell carcinoma. *Journal of translational medicine*. 2016;14(1):137.

40. Sohn RH, Chen J, Koblan KS, Bray PF, Goldschmidt-Clermont PJ. Localization of a binding site for phosphatidylinositol 4, 5-bisphosphate on human profilin. *Journal of Biological Chemistry*. 1995;270(36):21114-20.

41. Kwiatkowski DJ, Bruns G. Human profilin. Molecular cloning, sequence comparison, and chromosomal analysis. *Journal of Biological Chemistry*. 1988;263(12):5910-5.

42. Gieselmann R, Kwiatkowski DJ, Janmey PA, Witke W. Distinct biochemical characteristics of the two human profilin isoforms. *The FEBS Journal*. 1995;229(3):621-8.

43. Da Silva JS, Medina M, Zuliani C, Di Nardo A, Witke W, Dotti CG. RhoA/ROCK regulation of neuritogenesis via profilin IIa-mediated control of actin stability. *The Journal of cell biology*. 2003;162(7):1267-79.

44. Lambrechts A, Braun A, Jonckheere V, Aszodi A, Lanier LM, Robbens J, et al. Profilin II is alternatively spliced, resulting in profilin isoforms that are differentially expressed and have distinct biochemical properties. *Molecular and cellular biology*. 2000;20(21):8209-19.

45. Polet D, Lambrechts A, Vandepoele K, Vandekerckhove J, Ampe C. On the origin and evolution of vertebrate and viral profilins. *FEBS letters*. 2007;581(2):211-7.

46. Ding Z, Bae YH, Roy P. Molecular insights on context-specific role of profilin-1 in cell migration. *Cell adhesion & migration*. 2012;6(5):442-534.

47. Witke W, Podtelejnikov AV, Di Nardo A, Sutherland JD, Gurniak CB, Dotti C, et al. In mouse brain profilin I and profilin II associate with regulators of the endocytic pathway and actin assembly. *The EMBO journal*. 1998;17(4):967-76.

48. Yarmola EG, Bubba MR. How depolymerization can promote polymerization: the case of actin and profilin. *BioEssays*. 2009;31(11):1150-60.

49. Haarer B, Brown SS. Structure and function of profilin. *Cytoskeleton*. 1990;17(2):71-4.

50. Ding Z, Roy P. Profilin-1 versus profilin-2: two faces of the same coin? *Breast Cancer Research*. 2013;15(3):311.

51. Rennella E, Sekhar A, Kay LE. Self-Assembly of Human Profilin-1 Detected by Carr-Purcell-Meiboom-Gill Nuclear Magnetic Resonance (CPMG NMR) Spectroscopy. *Biochemistry*. 2017;56(5):692-703.

52. Shao J, Welch WJ, DiProspero NA, Diamond MI. Phosphorylation of profilin by ROCK1 regulates polyglutamine aggregation. *Molecular and cellular biology*. 2008;28(17):5196-208.

53. Suetsugu S, Miki H, Takenawa T. The essential role of profilin in the assembly of actin for microspike formation. *The EMBO journal*. 1998;17(22):6516-26.

54. Hansson A, Skoglund G, Lassing I, Lindberg U, Ingelman-Sundberg M. Protein kinase C-dependent phosphorylation of profilin is specifically stimulated by phosphatidylinositol bisphosphate (PIP2). *Biochemical and biophysical research communications*. 1988;150(2):526-31.

55. Harbeck B, Hüttelmaier S, Schlüter K, Jockusch BM, Illenberger S. Phosphorylation of the vasodilator-stimulated phosphoprotein regulates its interaction with actin. *Journal of Biological Chemistry*. 2000;275(40):30817-25.

56. Ackermann M, Matus A. Activity-induced targeting of profilin and stabilization of dendritic spine morphology. *Nature neuroscience*. 2003;6(11):1194-200.

57. Mouhaddach A, El-hadi A, Taghzouti K, Bendaou M, Hassikou R. Assessment of *Opuntia ficus-indica* in vivo following ethnobotanical survey: Confirmation of its analgesic activity. *Phytothérapie*. 2017:1-6.



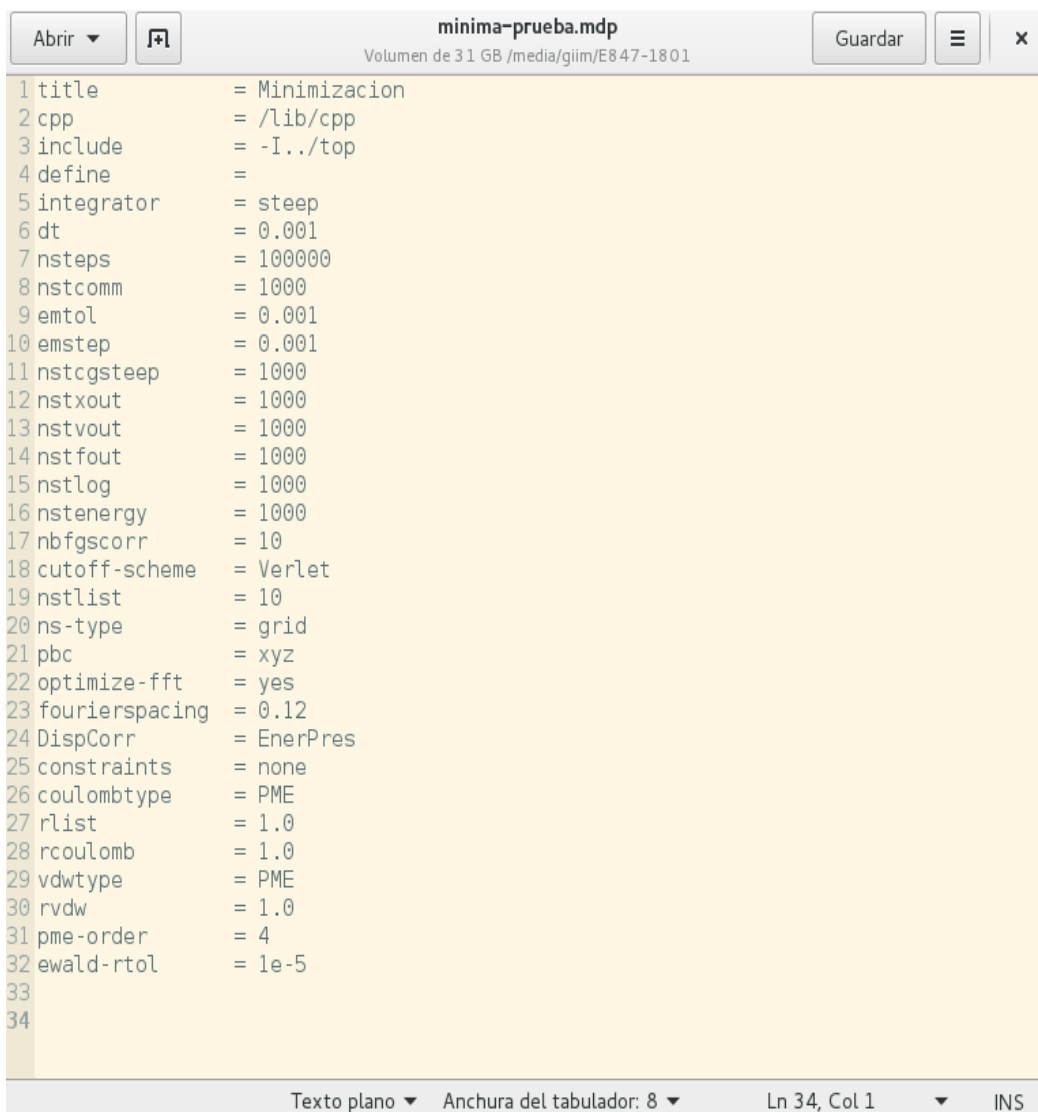
58. Silos Espino H. Estudio molecular del nopal: 1. calidad y proceso de maduración del fruto: 2. transformación genética por medio del sistema natural de agrobacterium tumefaciens. Cuaderno de trabajo: sistema de investigación Miguel Hidalgo. 2002.
59. Escobar O, Breshkovskaya T. Identificación de enzimas que participan en la ruta de biosíntesis de folatos, en cladodios de nopal (*Opuntia ficus indica*) 2008.
60. Msaddak L, Abdelhedi O, Kridene A, Rateb M, Belbahri L, Ammar E, et al. *Opuntia ficus-indica* cladodes as a functional ingredient: bioactive compounds profile and their effect on antioxidant quality of bread. *Lipids in health and disease*. 2017;16(1):32.
61. De Leo M, De Abreu MB, Pawlowska A, Cioni P, Braca A. Profiling the chemical content of *Opuntia ficus-indica* flowers by HPLC–PDA–ESI–MS and GC/EIMS analyses. *Phytochemistry Letters*. 2010;3(1):48-52.
62. Bárcenas Argüello ML. Acumulación de prolina en raíces de dos cultivares de nopal *Opuntia* spp. bajo tensión por calor. 1999.
63. Galati E, Tripodo M, Trovato A, Miceli N, Monforte M. Biological effect of *Opuntia ficus indica* (L.) Mill.(Cactaceae) waste matter: Note I: diuretic activity. *Journal of Ethnopharmacology*. 2002;79(1):17-21.
64. Galati EM, Mondello MR, Monforte MT, Galluzzo M, Miceli N, Tripodo MM. Effect of *Opuntia ficus-indica* (L.) Mill. cladodes in the wound-healing process. *Journal of the Professional Association for Cactus Development*. 2003;5:1-16.
65. Rabhi A, Limam F, Falleh H, Ksouri R, Abdely C, Raies A. *Opuntia ficus indica* flower: a prominent source of phenolic compounds and unsaturated fatty acid. *International J Adv Res*. 2015;3:41-55.
66. Camarena-Rangel NG, Barba-De la Rosa AP, Herrera-Corredor JA, Santos-Díaz MdS. Enhanced production of metabolites by elicitation in *Opuntia ficus-indica*, *Opuntia megacantha*, and *Opuntia streptacantha* callus. *Plant cell, tissue and organ culture*. 2017;129(2):289-98.
67. Antunes-Ricardo M, Gutiérrez-Urbe JA, Guajardo-Flores D. Extraction of isorhamnetin conjugates from *Opuntia ficus-indica* (L.) Mill using supercritical fluids. *The Journal of Supercritical Fluids*. 2017;119:58-63.
68. Pietta P-G. Flavonoids as antioxidants. *Journal of natural products*. 2000;63(7):1035-42.
69. Stintzing FC, Schieber A, Carle R. Identification of betalains from yellow beet (*Beta vulgaris* L.) and cactus pear [*Opuntia ficus-indica* (L.) Mill.] by high-performance liquid chromatography– electrospray ionization mass spectrometry. *Journal of Agricultural and Food Chemistry*. 2002;50(8):2302-7.
70. Dhaouadi K, Raboudi F, Funez-Gomez L, Pamies D, Estevan C, Hamdaoui M, et al. Polyphenolic extract of Barbary-Fig (*Opuntia ficus-indica*) syrup: RP–HPLC–ESI–MS analysis and determination of antioxidant, antimicrobial and Cancer-Cells cytotoxic potentials. *Food Analytical Methods*. 2013:1-9.
71. Wenzel U, Herzog A, Kuntz S, Daniel H. Protein expression profiling identifies molecular targets of quercetin as a major dietary flavonoid in human colon cancer cells. *Proteomics*. 2004;4(7):2160-74.
72. Kim JW, Kim TB, Yang H, Sung SH. Phenolic Compounds Isolated from *Opuntia ficus-indica* Fruits. *Natural Product Sciences*. 2016;22(2):117-21.
73. Dok-Go H, Lee KH, Kim HJ, Lee EH, Lee J, Song YS, et al. Neuroprotective effects of antioxidative flavonoids, quercetin,(+)-dihydroquercetin and quercetin 3-methyl ether, isolated from *Opuntia ficus-indica* var. saboten. *Brain research*. 2003;965(1):130-6.
74. Middleton E, Kandaswami C, Theoharides TC. The effects of plant flavonoids on mammalian cells: implications for inflammation, heart disease, and cancer. *Pharmacological reviews*. 2000;52(4):673-751.



75. Saralamma VVG, Kim EH, Lee HJ, Raha S, Lee WS, Heo JD, et al. Flavonoids: A new generation molecule to stimulate programmed cell deaths in cancer cells. *Journal of Biomedical and Translational Research*. 2017;18(1):30-7.
76. Martínez-Flórez S, González-Gallego J, Culebras JM, Tuñón MJ. Los flavonoides: propiedades y acciones antioxidantes. *Nutr Hosp*. 2002;17(6):271-8.
77. Agati G, Azzarello E, Pollastri S, Tattini M. Flavonoids as antioxidants in plants: location and functional significance. *Plant Science*. 2012;196:67-76.
78. Raj U, Varadwaj PK. Flavonoids as multi-target inhibitors for proteins associated with Ebola virus: In silico discovery using virtual screening and molecular docking studies. *Interdisciplinary Sciences: Computational Life Sciences*. 2016;8(2):132-41.
79. Lago JHG, Toledo-Arruda AC, Mernak M, Barrosa KH, Martins MA, Tibério IF, et al. Structure-activity association of flavonoids in lung diseases. *Molecules*. 2014;19(3):3570-95.
80. Baptista FI, Henriques AG, Silva AM, Wiltfang J, da Cruz e Silva OA. Flavonoids as therapeutic compounds targeting key proteins involved in Alzheimer's disease. *ACS chemical neuroscience*. 2014;5(2):83-92.
81. Jakobek L. Interactions of polyphenols with carbohydrates, lipids and proteins. *Food Chemistry*. 2015;175:556-67.
82. Novza Y, Popova E. Flavonoids: chemistry and biological activities. *Problems of Environmental Biotechnology*. 2016(1).
83. Brunetti C, Di Ferdinando M, Fini A, Pollastri S, Tattini M. Flavonoids as antioxidants and developmental regulators: relative significance in plants and humans. *International journal of molecular sciences*. 2013;14(2):3540-55.
84. Bustamante J, Dumas V, Gauto D. DINÁMICA MOLECULAR DE PROTEÍNAS.
85. Anfinsen CB. The formation and stabilization of protein structure. *Biochemical Journal*. 1972;128(4):737.
86. Gu J, Bourne PE. *Structural bioinformatics*: John Wiley & Sons; 2009.
87. Gago F. Métodos computacionales de modelado molecular y diseño de fármacos. *Monografías de la Real Academia Nacional de Farmacia*. 1994.
88. Gómez Serván C. *Introducción a la mecánica cuántica*. 2017.
89. Rost B, Sander C. Prediction of protein secondary structure at better than 70% accuracy. *Journal of molecular biology*. 1993;232(2):584-99.
90. Rhee YM, Sorin EJ, Jayachandran G, Lindahl E, Pande VS. Simulations of the role of water in the protein-folding mechanism. *Proceedings of the National Academy of Sciences of the United States of America*. 2004;101(17):6456-61.
91. Rojo RT. *Modeling biomolecules: interactions, forces and free energies*: Universidad de Zaragoza; 2016.
92. Chou PY, Fasman GD. Prediction of protein conformation. *Biochemistry*. 1974;13(2):222-45.
93. Qin Z, Buehler MJ. Molecular dynamics simulation of the  $\alpha$ -helix to  $\beta$ -sheet transition in coiled protein filaments: Evidence for a critical filament length scale. *Physical Review Letters*. 2010;104(19):198304.
94. Hooft RW, Sander C, Vriend G. Objectively judging the quality of a protein structure from a Ramachandran plot. *Bioinformatics*. 1997;13(4):425-30.
95. Weiner SJ, Kollman PA, Nguyen DT, Case DA. An all atom force field for simulations of proteins and nucleic acids. *Journal of computational chemistry*. 1986;7(2):230-52.
96. Allinger NL, Yuh YH, Lii JH. Molecular mechanics. The MM3 force field for hydrocarbons. 1. *Journal of the American Chemical Society*. 1989;111(23):8551-66.
97. Lennard-Jones J. On the forces between atoms and ions. *Proc R Soc Lond A*. 1925;109(752):584-97.

98. Kaminski GA, Friesner RA, Tirado-Rives J, Jorgensen WL. Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *The Journal of Physical Chemistry B*. 2001;105(28):6474-87.
99. Jorgensen WL, Tirado-Rives J. The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *Journal of the American Chemical Society*. 1988;110(6):1657-66.
100. Oostenbrink C, Villa A, Mark AE, Van Gunsteren WF. A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *Journal of computational chemistry*. 2004;25(13):1656-76.
101. Mackerell Jr AD, Bashford D, Bellott M, Dunbrack Jr RL, Evanseck JD, Field MJ, et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *The journal of physical chemistry B*. 1998;102(18):3586-616.
102. Jensen F. *Introduction to computational chemistry*: John Wiley & sons; 2017.
103. Kukol A. *Molecular modeling of proteins*: Springer; 2008.
104. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*. 2015;1:19-25.
105. Zhao J, Wang Q, Liang G, Zheng J. Molecular dynamics simulations of low-ordered Alzheimer  $\beta$ -amyloid oligomers from dimer to hexamer on self-assembled monolayers. *Langmuir*. 2011;27(24):14876-87.
106. Shirts MR, Pitner JW, Swope WC, Pande VS. Extremely precise free energy calculations of amino acid side chain analogs: Comparison of common molecular mechanics force fields for proteins. *The Journal of chemical physics*. 2003;119(11):5740-61.
107. Schapira M, Totrov M, Abagyan R. Prediction of the binding energy for small molecules, peptides and proteins. *Journal of Molecular Recognition*. 1999;12(3):177-90.
108. Knegtel RM, Grootenhuys PD. Binding affinities and non-bonded interaction energies. *Perspectives in drug discovery and design*. 1998;9:99-114.
109. Bennett CH. Efficient estimation of free energy differences from Monte Carlo data. *Journal of Computational Physics*. 1976;22(2):245-68.
110. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, et al. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of computational chemistry*. 2009;30(16):2785-91.
111. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, et al. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *Journal of computational chemistry*. 1998;19(14):1639-62.
112. Kozakov D, Hall DR, Xia B, Porter KA, Padhorny D, Yueh C, et al. The ClusPro web server for protein-protein docking. *Nature protocols*. 2017;12(2):255.
113. Kozakov D, Beglov D, Bohnuud T, Mottarella SE, Xia B, Hall DR, et al. How good is automated protein docking? *Proteins: Structure, Function, and Bioinformatics*. 2013;81(12):2159-66.
114. Ritchie DW, Kemp GJ. Protein docking using spherical polar Fourier correlations. *Proteins: Structure, Function, and Bioinformatics*. 2000;39(2):178-94.
115. Duhovny D, Nussinov R, Wolfson HJ, editors. *Efficient unbound docking of rigid molecules*. International workshop on algorithms in bioinformatics; 2002: Springer.
116. Schneidman-Duhovny D, Inbar Y, Polak V, Shatsky M, Halperin I, Benyamini H, et al. Taking geometry to its edge: fast unbound rigid (and hinge-bent) docking. *Proteins: Structure, Function, and Bioinformatics*. 2003;52(1):107-12.
117. Andrusier N, Nussinov R, Wolfson HJ. FireDock: fast interaction refinement in molecular docking. *Proteins: Structure, Function, and Bioinformatics*. 2007;69(1):139-59.

# APÉNDICES



```

1 title           = Minimizacion
2 cpp             = /lib/cpp
3 include         = -I../top
4 define          =
5 integrator      = steep
6 dt              = 0.001
7 nsteps          = 100000
8 nstcomm         = 1000
9 emtol           = 0.001
10 emstep         = 0.001
11 nstcgsteep     = 1000
12 nstxout        = 1000
13 nstvout        = 1000
14 nstfout        = 1000
15 nstlog         = 1000
16 nstenergy      = 1000
17 nbfgsccorr     = 10
18 cutoff-scheme  = Verlet
19 nstlist        = 10
20 ns-type        = grid
21 pbc            = xyz
22 optimize-fft   = yes
23 fourierspacing = 0.12
24 DispCorr       = EnerPres
25 constraints     = none
26 coulombtype    = PME
27 rlist          = 1.0
28 rcoulomb       = 1.0
29 vdwtype        = PME
30 rvdw           = 1.0
31 pme-order      = 4
32 ewald-rtol     = 1e-5
33
34

```

**Figura A.1:** Archivo *minima.mdp* para la minimización de energías en Gromacs



```

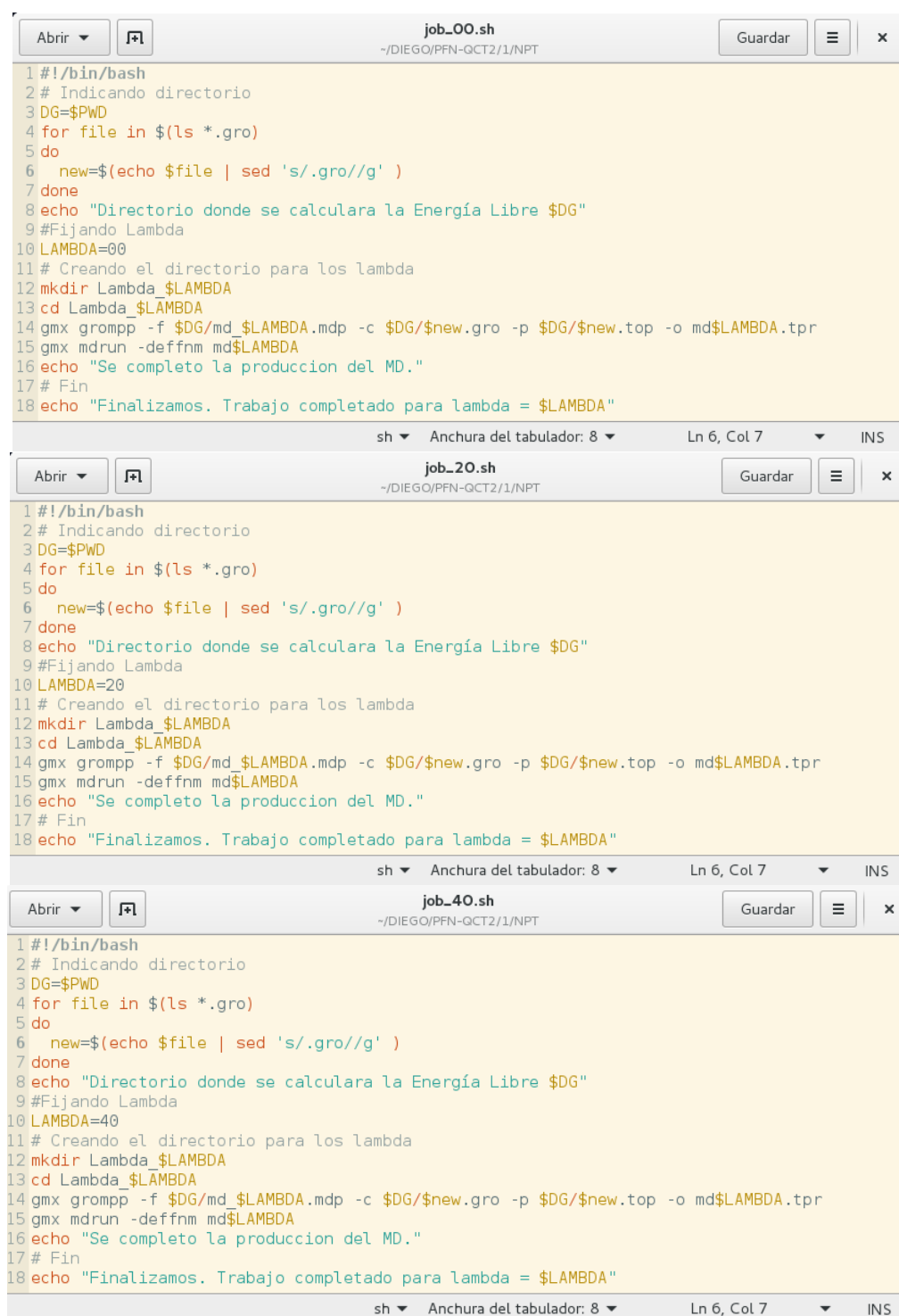
nvt-so.mdp
Volumen de 3.1 GB /media/giim/E847-1801

1 title           = Dinamica Molecular NVT
2 cpp             = /lib/cpp
3 include         = -I../top
4 integrator      = md
5 dt              = 0.001
6 nsteps          = 50000000
7 nstxout         = 1000
8 nstfout         = 1000
9 nstlog          = 1000
10 nstcalcenergy  = -1
11 nstenergy       = 200
12 nstxout-compressed= 1000
13 xtc-grps       =
14 energygrps     =
15 nstlist        = 20
16 ns-type        = grid
17 fourierspacing = 0.39
18 pme-order      = 4
19 ewald-rtol     = 1e-5
20 pbc            = xyz
21 periodic-molecules = yes
22 tcoupl         = v-rescale
23 tc-grps        = system
24 tau-t          = 0.5
25 ref-t          = 309.65
26 gen-vel        = yes
27 gen-seed       = 173529
28 constraints    = none
29 cutoff-scheme  = Verlet
30 coulombtype    = PME
31 rlist          = 0.9
32 rcoulomb       = 0.9
33 vdwttype       = Cut-off
34 rvdw           = 0.9
35
36

```

Texto plano ▼ Anchura del tabulador: 8 ▼ Ln 1, Col 1 ▼ INS

**Figura A.2:** Archivo nvt.md para la dinámica molecular en Gromacs



```

job_00.sh
~/DIEGO/PFN-QCT2/1/NPT

1 #!/bin/bash
2 # Indicando directorio
3 DG=$PWD
4 for file in $(ls *.gro)
5 do
6   new=$(echo $file | sed 's/./gro/g' )
7 done
8 echo "Directorio donde se calculara la Energía Libre $DG"
9 #Fijando Lambda
10 LAMBDA=00
11 # Creando el directorio para los lambda
12 mkdir Lambda_$LAMBDA
13 cd Lambda_$LAMBDA
14 gmx grompp -f $DG/md_$LAMBDA.mdp -c $DG/$new.gro -p $DG/$new.top -o md$LAMBDA.tpr
15 gmx mdrun -deffnm md$LAMBDA
16 echo "Se completo la produccion del MD."
17 # Fin
18 echo "Finalizamos. Trabajo completado para lambda = $LAMBDA"

sh Anchura del tabulador: 8 Ln 6, Col 7 INS

job_20.sh
~/DIEGO/PFN-QCT2/1/NPT

1 #!/bin/bash
2 # Indicando directorio
3 DG=$PWD
4 for file in $(ls *.gro)
5 do
6   new=$(echo $file | sed 's/./gro/g' )
7 done
8 echo "Directorio donde se calculara la Energía Libre $DG"
9 #Fijando Lambda
10 LAMBDA=20
11 # Creando el directorio para los lambda
12 mkdir Lambda_$LAMBDA
13 cd Lambda_$LAMBDA
14 gmx grompp -f $DG/md_$LAMBDA.mdp -c $DG/$new.gro -p $DG/$new.top -o md$LAMBDA.tpr
15 gmx mdrun -deffnm md$LAMBDA
16 echo "Se completo la produccion del MD."
17 # Fin
18 echo "Finalizamos. Trabajo completado para lambda = $LAMBDA"

sh Anchura del tabulador: 8 Ln 6, Col 7 INS

job_40.sh
~/DIEGO/PFN-QCT2/1/NPT

1 #!/bin/bash
2 # Indicando directorio
3 DG=$PWD
4 for file in $(ls *.gro)
5 do
6   new=$(echo $file | sed 's/./gro/g' )
7 done
8 echo "Directorio donde se calculara la Energía Libre $DG"
9 #Fijando Lambda
10 LAMBDA=40
11 # Creando el directorio para los lambda
12 mkdir Lambda_$LAMBDA
13 cd Lambda_$LAMBDA
14 gmx grompp -f $DG/md_$LAMBDA.mdp -c $DG/$new.gro -p $DG/$new.top -o md$LAMBDA.tpr
15 gmx mdrun -deffnm md$LAMBDA
16 echo "Se completo la produccion del MD."
17 # Fin
18 echo "Finalizamos. Trabajo completado para lambda = $LAMBDA"

sh Anchura del tabulador: 8 Ln 6, Col 7 INS
  
```

**Figura A.3:** Archivos script inicial, intermedio y final de los 41 scripts para la ejecución de los archivos .mdp para el cálculo NPT de los sistemas profilina humana II - flavonoides

```

md_00.mdp
~/DIEGO/PFN-QCT2/1/NPT

1; Run control
2 integrator          = sd          ; Langevin dynamics
3 tinit              = 0
4 dt                 = 0.002
5 nsteps             = 500000      ; 1 ns
6 nstcomm            = 100
7 ; Output control
8 nstxout             = 500
9 nstfout             = 500
10 nstfout            = 0
11 nstlog             = 500
12 nstenergy          = 500
13 nstxout-compressed = 0
14 ; Neighborsearching and short-range nonbonded interactions
15 cutoff-scheme      = verlet
16 nstlist            = 20
17 ns_type            = grid
18 pbc                 = xyz
19 rlist              = 1.2
20 ; Electrostatics
21 coulombtype         = PME
22 rcoulomb            = 1.2
23 ; van der Waals
24 vdwttype           = cutoff
25 vdw-modifier        = potential-switch
26 rvdw-switch         = 1.0
27 rvdw                = 1.2
28 ; Apply long range dispersion corrections for Energy and Pressure
29 DispCorr            = EnerPres
30 ; Spacing for the PME/PPPM FFT grid
31 fourierspacing      = 0.12
32 ; EWALD/PME/PPPM parameters
33 pme_order           = 6
34 ewald_rtol          = 1e-06
35 epsilon_surface     = 0
36 ; Temperature coupling
37 ; tcoupl is implicitly handled by the sd integrator
38 tcoupl              = nose-hoover
39 tc_grps             = system
40 tau_t               = 1.0
41 ref_t               = 309.65
42 ; Pressure coupling is on for NPT
43 Pcoupl              = Parrinello-Rahman
44 tau_p               = 1.0
45 compressibility      = 4.5e-05
46 ref_p               = 1.0

```

**Figura A. 4:** Primer archivo para el cálculo de NPT parte I, donde se describen los parámetros




```

md_00.mdp
~/DIEGO/PFN-QCT2/1/NPT

47 ; Free energy control stuff
48 free_energy = yes
49 init_lambda_state = 0
50 delta_lambda = 0
51 calc_lambda_neighbors = 1 ; only immediate neighboring windows
52 ; Vectors of lambda specified here
53 ; Each combination is an index that is retrieved from init_lambda_state for each
simulation
54 ; init_lambda_state 0 1 2 3 4 5 6 7 8 9 10 11
12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29
30 31 32 33 34 35 36 37 38 39 40
55 vdw_lambdas = 0.00 0.05 0.10 0.15 0.20 0.25 0.30 0.35 0.40 0.45 0.50 0.55
0.60 0.65 0.70 0.75 0.80 0.85 0.90 0.95 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
56 coul_lambdas = 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.05 0.10 0.15 0.20 0.25 0.30 0.35 0.40 0.45
0.50 0.55 0.60 0.65 0.70 0.75 0.80 0.85 0.90 0.95 1.00
57 ; We are not transforming any bonded or restrained interactions
58 bonded_lambdas = 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00
0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00
0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00
59 restraint_lambdas = 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00
0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00
0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00
60 ; Masses are not changing (particle identities are the same at lambda = 0 and lambda = 1)
61 mass_lambdas = 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00
0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00
0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00
62 ; Not doing simulated tempering here
63 temperature_lambdas = 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00
0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00
0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00
64 ; Options for the decoupling
65 sc-alpha = 0.5
66 sc-coul = no ; linear interpolation of Coulomb (none in this case)
67 sc-power = 1
68 sc-sigma = 0.3
69 couple-moltype = Other2 ; name of molecule type to decouple
70 couple-intramol = no
71 couple-lambda0 = none ; only van der Waals interactions
72 couple-lambda1 = vdw-q ; turn off everything, in this case only vdw
73 nstdhdl = 10
74 ; Do not generate velocities
75 gen_vel = no
76 ; options for bonds
77 constraints = h-bonds ; we only have C-H bonds here
78 ; Type of constraint algorithm
79 constraint-algorithm = lincs
80 ; Constrain the starting configuration
81 ; since we are continuing from NPT
82 continuation = no
83 ; Highest order in the expansion of the constraint coupling matrix
84 lincs-order = 12

```

**Figura A.5:** Primero archivo para el cálculo de NPT parte 2, donde se describen los estados y distancias que se les va a asignando a las interacciones no enlazantes como las fuerzas de Van der Waals y fuerzas Coulombicas



```

Abrir  run.sh  Guardar
Volumen de 16 GB /media/diego/A0443DD5443DAF42/...

#!/bin/bash
for file in $(ls *.sh)
do
    ./ $file
done

sh  Anchura del tabulador: 8  Ln 1, Col 1  INS
    
```

**Figura A.6:** Archivo script para la ejecución de todos los archivos job.sh

[	QCT	]
[	atoms	]
O1	opls_180	-0.107816 1
O2	opls_167	-0.207040 2
O3	opls_167	-0.217411 3
O4	opls_180	-0.266758 4
O5	opls_167	-0.209607 5
O6	opls_167	-0.209252 6
O7	opls_167	-0.209672 7
C1	opls_147	-0.067253 8
C2	opls_147	0.089296 9
C3	opls_1458	0.062677 10
C4	opls_1458	-0.025770 11
C5	opls_166	0.042162 12
C6	opls_147	0.139913 13
C7	opls_166	0.105952 14
C8	opls_145	-0.096595 15
C9	opls_166	0.093954 16
C10	opls_145	-0.069181 17
C11	opls_145	-0.101710 18
C12	opls_145	-0.045525 19
C13	opls_166	0.057486 20
C14	opls_145	-0.051745 21
C15	opls_166	0.074573 22
H1	opls_146	0.062625 15
H2	opls_146	0.056392 17
H3	opls_146	0.055450 18
H4	opls_146	0.057332 19
H5	opls_146	0.065026 21
H6	opls_168	0.175959 23
H7	opls_168	0.162351 24
H8	opls_168	0.195190 25
H9	opls_168	0.197878 26
H10	opls_168	0.185179 27

**Figura A.7:** Cargas de Hirshfeld calculadas para el compuesto QCT

```
[ DHQ ]
[ atoms ]
01  opls_179  -0.134204  1
02  opls_154  -0.234575  2
03  opls_281  -0.256732  3
04  opls_167  -0.212149  4
05  opls_167  -0.199719  5
06  opls_167  -0.215262  6
07  opls_167  -0.213954  7
C1  opls_149  0.056467  8
C2  opls_158  0.043412  9
C3  opls_147  -0.018325  10
C4  opls_280  0.146669  11
C5  opls_147  -0.081102  12
C6  opls_199  0.088719  13
C7  opls_166  0.109180  14
C8  opls_145  -0.059168  15
C9  opls_145  -0.063201  16
C10 opls_145  -0.096638  17
C11 opls_145  -0.105960  18
C12 opls_166  0.099211  19
C13 opls_166  0.069272  20
C14 opls_145  -0.069916  21
C15 opls_166  0.054836  22
H1  opls_140  0.054494  8
H2  opls_282  0.064007  9
H3  opls_146  0.053637  15
H4  opls_146  0.048622  16
H5  opls_146  0.063401  17
H6  opls_155  0.166253  23
H7  opls_146  0.054938  18
H8  opls_146  0.052203  21
H9  opls_168  0.162890  24
H10 opls_168  0.197034  25
H11 opls_168  0.181662  26
H12 opls_168  0.193998  27
```

**Figura A.8:** Cargas de Hirshfeld calculadas para el compuesto DHQ

```
[ QME ]
[ atoms ]
01  opls_180  -0.099631  1
02  opls_179  -0.145183  2
03  opls_167  -0.217153  3
04  opls_180  -0.257058  4
05  opls_167  -0.205938  5
06  opls_167  -0.211915  6
07  opls_167  -0.211883  7
C1  opls_147  -0.066048  8
C2  opls_145  0.088270  9
C3  opls_147  0.090558  10
C4  opls_199  0.037416  11
C5  opls_145  -0.024700  12
C6  opls_147  0.138629  13
C7  opls_166  0.108083  14
C8  opls_145  -0.097924  15
C9  opls_145  -0.067460  16
C10 opls_145  -0.101849  17
C11 opls_166  0.090737  18
C12 opls_145  -0.045502  19
C13 opls_166  0.055695  20
C14 opls_145  -0.059489  21
C15 opls_166  0.074792  22
C16 opls_181  -0.015535  23
H1  opls_146  0.060362  15
H2  opls_146  0.049619  16
H3  opls_146  0.055071  17
H4  opls_146  0.046586  19
H5  opls_146  0.061128  21
H6  opls_168  0.161667  24
H7  opls_140  0.043548  23
H8  opls_140  0.039792  23
H9  opls_140  0.051820  23
H10 opls_168  0.194999  25
H11 opls_168  0.195119  26
H12 opls_168  0.182377  27
```

**Figura A.9:** Cargas de Hirshfeld calculadas para el compuesto QME



```
[ QRU ]
[ atoms ]
01  op1s_186  -0.182833  1
02  op1s_186  -0.181792  2
03  op1s_186  -0.185387  3
04  op1s_171  -0.230191  4
05  op1s_171  -0.209339  5
06  op1s_179  -0.142827  6
07  op1s_171  -0.253880  7
08  op1s_171  -0.245561  8
09  op1s_171  -0.268267  9
010  op1s_171  -0.208203  10
011  op1s_180  -0.104668  11
012  op1s_180  -0.281606  12
013  op1s_167  -0.184221  13
014  op1s_167  -0.205714  14
015  op1s_167  -0.290653  15
016  op1s_167  -0.216777  16
C1  op1s_183  0.049357  17
C2  op1s_174  0.039572  18
C3  op1s_174  0.037061  19
C4  op1s_174  0.039872  20
C5  op1s_174  0.038484  21
C6  op1s_193  0.116720  22
C7  op1s_174  0.027650  23
C8  op1s_174  0.035654  24
C9  op1s_193  0.122907  25
C10  op1s_182  0.015685  26
C11  op1s_183  0.047424  27
C12  op1s_135  -0.107432  28
C13  op1s_199  0.027569  29
C14  op1s_145B  0.090332  30
C15  op1s_147  0.132698  31
C16  op1s_147  -0.057213  32
C17  op1s_145B  -0.030615  33
C18  op1s_147  0.092280  34
C19  op1s_166  0.102024  35
C20  op1s_145  -0.090008  36
C21  op1s_145  -0.058642  37
C22  op1s_145  -0.052299  38
C23  op1s_145  -0.111174  39
C24  op1s_166  0.091468  40
C25  op1s_166  0.057017  41
C26  op1s_145  -0.063029  42
C27  op1s_166  0.074488  43
H1  op1s_185  0.041771  17
H2  op1s_176  0.039990  18
H3  op1s_176  0.044245  19
H4  op1s_176  0.057331  20
H5  op1s_176  0.058120  21
H6  op1s_194  0.052663  22
H7  op1s_176  0.042465  23
H8  op1s_176  0.041940  24
H9  op1s_194  0.060846  25
H10  op1s_185  0.053974  26
H11  op1s_185  0.044256  26
H12  op1s_185  0.031971  27
H13  op1s_140  0.042887  28
H14  op1s_140  0.042393  28
H15  op1s_140  0.038304  28
H16  op1s_172  0.168610  44
H17  op1s_172  0.172544  45
H18  op1s_172  0.164871  46
H19  op1s_172  0.163339  47
H20  op1s_172  0.152085  48
H21  op1s_172  0.170026  49
H22  op1s_146  0.066734  36
H23  op1s_146  0.036864  37
H24  op1s_146  0.047948  38
H25  op1s_146  0.047875  39
H26  op1s_146  0.059171  42
H27  op1s_168  0.196106  50
H28  op1s_168  0.193760  51
H29  op1s_168  0.153193  52
H30  op1s_168  0.183847  53
```

**Figura A.10:** Cargas de Hirshfeld calculadas para el compuesto QRU

```
[ QGL ]
[ atoms ]
01  opls_186  -0.185600  1
02  opls_179  -0.149949  2
03  opls_171  -0.248356  3
04  opls_171  -0.237947  4
05  opls_171  -0.248702  5
06  opls_154  -0.233151  6
07  opls_180  -0.111727  7
08  opls_167  -0.203417  8
09  opls_167  -0.194326  9
010  opls_180  -0.279584  10
011  opls_167  -0.180867  11
012  opls_167  -0.205716  12
C1  opls_174  0.033552  13
C2  opls_174  0.039183  14
C3  opls_174  0.045403  15
C4  opls_183  0.041328  16
C5  opls_193  0.124668  17
C6  opls_157  0.004234  18
C7  opls_199  0.063171  19
C8  opls_145  -0.060075  20
C9  opls_166  0.073652  21
C10  opls_1458  -0.024411  22
C11  opls_145  -0.068948  23
C12  opls_1458  0.053891  24
C13  opls_145  -0.045601  25
C14  opls_166  0.039290  26
C15  opls_147  -0.058938  27
C16  opls_147  0.091750  28
C17  opls_147  0.135587  29
C18  opls_166  0.102368  30
C19  opls_145  -0.094660  31
C20  opls_166  0.089584  32
C21  opls_145  -0.112916  33
H1  opls_176  0.038399  13
H2  opls_176  0.057237  14
H3  opls_176  0.066447  15
H4  opls_185  0.048437  16
H5  opls_194  0.058281  17
H6  opls_140  0.031003  18
H7  opls_140  0.052077  18
H8  opls_172  0.161648  34
H9  opls_172  0.170895  35
H10  opls_172  0.166749  36
H11  opls_155  0.154850  37
H12  opls_146  0.047264  20
H13  opls_146  0.053803  23
H14  opls_146  0.054608  25
H15  opls_168  0.192714  38
H16  opls_146  0.062221  31
H17  opls_146  0.046341  33
H18  opls_168  0.157864  39
H19  opls_168  0.194049  40
H20  opls_168  0.192343  41
```

**Figura A.11:** Cargas de Hirshfeld calculadas para el compuesto QGL

```
[ ISO ]
[ atoms ]
01  opls_180  -0.096327  1
02  opls_167  -0.198148  2
03  opls_167  -0.214173  3
04  opls_179  -0.157801  4
05  opls_180  -0.284204  5
06  opls_167  -0.201682  6
07  opls_167  -0.213572  7
C1  opls_147  -0.064542  8
C2  opls_1458  0.082820  9
C3  opls_147  0.094724  10
C4  opls_1458  -0.026096  11
C5  opls_166  0.037146  12
C6  opls_147  0.129857  13
C7  opls_166  0.106266  14
C8  opls_145  -0.093879  15
C9  opls_145  -0.064690  16
C10  opls_145  -0.049188  17
C11  opls_166  0.096455  18
C12  opls_145  -0.100333  19
C13  opls_199  0.054848  20
C14  opls_145  -0.060334  21
C15  opls_166  0.074977  22
C16  opls_181  -0.004322  23
H1  opls_146  0.064026  15
H2  opls_146  0.041714  16
H3  opls_146  0.048718  17
H4  opls_146  0.056431  19
H5  opls_146  0.060397  21
H6  opls_168  0.180960  24
H7  opls_168  0.163177  25
H8  opls_168  0.196178  26
H9  opls_168  0.179404  27
H10  opls_140  0.051382  23
H11  opls_140  0.052539  23
H12  opls_140  0.057272  23
```

**Figura A.12:** Cargas de Hirshfeld calculadas para el compuesto ISO

```
[ IGL ]
[ atoms ]
01  opl_186  -0.178983  1
02  opl_179  -0.162937  2
03  opl_171  -0.277278  3
04  opl_171  -0.245786  4
05  opl_171  -0.247614  5
06  opl_154  -0.262185  6
07  opl_180  -0.095870  7
08  opl_180  -0.238759  8
09  opl_167  -0.212651  9
010  opl_179  -0.157850  10
011  opl_167  -0.199221  11
012  opl_167  -0.188572  12
C1  opl_174  0.034135  13
C2  opl_174  0.031413  14
C3  opl_174  0.037928  15
C4  opl_183  0.043988  16
C5  opl_193  0.126436  17
C6  opl_157  0.011148  18
C7  opl_199  0.035555  19
C8  opl_145B  0.099648  20
C9  opl_147  0.135483  21
C10  opl_147  -0.061047  22
C11  opl_147  0.093576  23
C12  opl_145B  -0.021108  24
C13  opl_166  0.110638  25
C14  opl_145  -0.104397  26
C15  opl_145  -0.049419  27
C16  opl_145  -0.049170  28
C17  opl_145  -0.085439  29
C18  opl_166  0.099230  30
C19  opl_199  0.072025  31
C20  opl_145  -0.069639  32
C21  opl_166  0.058487  33
C22  opl_181  -0.016391  34
H1  opl_176  0.047241  13
H2  opl_176  0.031448  14
H3  opl_176  0.048442  15
H4  opl_185  0.048202  16
H5  opl_194  0.044121  17
H6  opl_140  0.036980  18
H7  opl_140  0.050988  18
H8  opl_172  0.133543  35
H9  opl_172  0.173955  36
H10  opl_172  0.175717  37
H11  opl_155  0.162740  38
H12  opl_146  0.056478  26
H13  opl_146  0.060239  27
H14  opl_146  0.050556  28
H15  opl_146  0.066702  29
H16  opl_146  0.054539  32
H17  opl_168  0.159284  39
H18  opl_168  0.196254  40
H19  opl_168  0.193719  41
H20  opl_140  0.041394  34
H21  opl_140  0.042217  34
H22  opl_140  0.049867  34
```

**Figura A.13:** Cargas de Hirshfeld calculadas para el compuesto IGL

```
[ KME ]
[ atoms ]
01  opl_180  -0.099894  1
02  opl_179  -0.146949  2
03  opl_167  -0.220585  3
04  opl_180  -0.258301  4
05  opl_167  -0.203857  5
06  opl_167  -0.212503  6
C1  opl_147  -0.065887  7
C2  opl_147  0.090894  8
C3  opl_145B  0.088229  9
C4  opl_199  0.036144  10
C5  opl_147  0.137030  11
C6  opl_145B  -0.026926  12
C7  opl_166  0.105727  13
C8  opl_145  -0.095702  14
C9  opl_145  -0.101518  15
C10  opl_166  0.094314  16
C11  opl_145  -0.032623  17
C12  opl_145  -0.033199  18
C13  opl_145  -0.074518  19
C14  opl_145  -0.060233  20
C15  opl_166  0.086145  21
C16  opl_181  -0.016871  22
H1  opl_146  0.062825  14
H2  opl_146  0.055451  15
H3  opl_146  0.049435  17
H4  opl_146  0.053320  18
H5  opl_146  0.050977  19
H6  opl_146  0.060443  20
H7  opl_168  0.159161  23
H8  opl_140  0.044900  22
H9  opl_140  0.040254  22
H10  opl_140  0.049156  22
H11  opl_168  0.195167  24
H12  opl_168  0.189994  25
```

**Figura A.14:** Cargas de Hirshfeld calculadas para el compuesto KME



[ IGA ]		
[ atoms ]		
01	opls_186	-0.179685 1
02	opls_179	-0.153210 2
03	opls_171	-0.262370 3
04	opls_171	-0.200796 4
05	opls_171	-0.228253 5
06	opls_154	-0.267163 6
07	opls_180	-0.097024 7
08	opls_180	-0.271584 8
09	opls_167	-0.217267 9
010	opls_179	-0.153442 10
011	opls_167	-0.201469 11
012	opls_167	-0.201461 12
C1	opls_174	0.038018 13
C2	opls_174	0.036278 14
C3	opls_174	0.039763 15
C4	opls_183	0.041431 16
C5	opls_193	0.117488 17
C6	opls_157	0.013790 18
C7	opls_199	0.034090 19
C8	opls_145B	0.095313 20
C9	opls_147	0.134908 21
C10	opls_147	-0.064748 22
C11	opls_147	0.092690 23
C12	opls_145B	-0.024170 24
C13	opls_166	0.106729 25
C14	opls_145	-0.093643 26
C15	opls_145	-0.062139 27
C16	opls_145	-0.049432 28
C17	opls_145	-0.099762 29
C18	opls_166	0.096580 30
C19	opls_199	0.069520 31
C20	opls_145	-0.072715 32
C21	opls_166	0.076979 33
C22	opls_181	-0.016703 34
H1	opls_176	0.050363 13
H2	opls_176	0.047987 14
H3	opls_176	0.050801 15
H4	opls_185	0.032371 16
H5	opls_194	0.049830 17
H6	opls_140	0.050801 18
H7	opls_140	0.037268 18
H8	opls_172	0.154043 35
H9	opls_172	0.178212 36
H10	opls_172	0.183091 37
H11	opls_155	0.143846 38
H12	opls_146	0.064119 26
H13	opls_146	0.044895 27
H14	opls_146	0.050621 28
H15	opls_146	0.056580 29
H16	opls_146	0.052545 32
H17	opls_168	0.159655 39
H18	opls_168	0.196474 40
H19	opls_168	0.191383 41
H20	opls_140	0.055801 34
H21	opls_140	0.030382 34
H22	opls_140	0.042481 34

**Figura A.15:** Cargas de Hirshfeld calculadas para el compuesto IGA

[ KMP ]		
[ atoms ]		
01	opls_180	-0.109165 1
02	opls_167	-0.206797 2
03	opls_167	-0.217937 3
04	opls_180	-0.265859 4
05	opls_167	-0.204355 5
06	opls_167	-0.210159 6
C1	opls_147	-0.067059 7
C2	opls_147	0.089128 8
C3	opls_145B	0.062374 9
C4	opls_147	0.140280 10
C5	opls_166	0.042957 11
C6	opls_145B	-0.028090 12
C7	opls_166	0.105839 13
C8	opls_145	-0.096935 14
C9	opls_166	0.093639 15
C10	opls_145	-0.101846 16
C11	opls_145	-0.038624 17
C12	opls_145	-0.030226 18
C13	opls_145	-0.071400 19
C14	opls_145	-0.053538 20
C15	opls_166	0.087121 21
H1	opls_146	0.062271 14
H2	opls_146	0.055312 16
H3	opls_146	0.058010 17
H4	opls_146	0.060370 18
H5	opls_146	0.054043 19
H6	opls_146	0.063853 20
H7	opls_168	0.176991 22
H8	opls_168	0.162034 23
H9	opls_168	0.194807 24
H10	opls_168	0.191961 25

**Figura A.16:** Cargas de Hirshfeld calculadas para el compuesto KMP

```
[ IRO ]
[ atoms ]
01  opls_186 -0.187678 1
02  opls_186 -0.185406 2
03  opls_186 -0.180976 3
04  opls_171 -0.237436 4
05  opls_171 -0.262993 5
06  opls_179 -0.149084 6
07  opls_171 -0.254143 7
08  opls_171 -0.237052 8
09  opls_171 -0.235319 9
010  opls_171 -0.243895 10
011  opls_180 -0.103169 11
012  opls_180 -0.258786 12
013  opls_167 -0.215251 13
014  opls_167 -0.201172 14
015  opls_179 -0.150182 15
016  opls_167 -0.170981 16
C1  opls_183 0.035943 17
C2  opls_174 0.034750 18
C3  opls_174 0.032281 19
C4  opls_174 0.034644 20
C5  opls_174 0.036091 21
C6  opls_193 0.118493 22
C7  opls_174 0.035592 23
C8  opls_174 0.037844 24
C9  opls_193 0.118396 25
C10  opls_182 0.010590 26
C11  opls_183 0.046203 27
C12  opls_135 -0.107443 28
C13  opls_199 0.035805 29
C14  opls_1458 0.095237 30
C15  opls_147 0.138260 31
C16  opls_147 -0.064849 32
C17  opls_1458 -0.017265 33
C18  opls_147 0.090394 34
C19  opls_166 0.107586 35
C20  opls_145 -0.092355 36
C21  opls_145 -0.044930 37
C22  opls_145 -0.047823 38
C23  opls_145 -0.098672 39
C24  opls_166 0.097020 40
C25  opls_199 0.076622 41
C26  opls_145 -0.064960 42
C27  opls_166 0.073177 43
C28  opls_181 -0.023296 44
H1  opls_185 0.032428 17
H2  opls_176 0.048614 18
H3  opls_176 0.049799 19
H4  opls_176 0.052672 20
H5  opls_176 0.028854 21
H6  opls_194 0.049994 22
H7  opls_176 0.042846 23
H8  opls_176 0.054427 24
H9  opls_194 0.053004 25
H10  opls_185 0.044126 26
H11  opls_185 0.037546 26
H12  opls_185 0.038693 27
H13  opls_140 0.045737 28
H14  opls_140 0.041807 28
H15  opls_140 0.039198 28
H16  opls_172 0.159598 45
H17  opls_172 0.150894 46
H18  opls_172 0.153713 47
H19  opls_172 0.149375 48
H20  opls_172 0.169213 49
H21  opls_172 0.171043 50
H22  opls_146 0.065640 36
H23  opls_146 0.064044 37
H24  opls_146 0.054263 38
H25  opls_146 0.057192 39
H26  opls_146 0.054849 42
H27  opls_168 0.163001 51
H28  opls_168 0.196790 52
H29  opls_168 0.194403 53
H30  opls_140 0.027960 44
H31  opls_140 0.049033 44
H32  opls_140 0.038892 44
```

**Figura A.17:** Cargas de Hirshfeld calculadas para el compuesto IRO