

UNIVERSIDAD AUTÓNOMA DEL ESTADO DE HIDALGO



Escuela Superior de Apan

TESIS DE LICENCIATURA

Estudios *in silico* sobre la proteína Tau en el Tratamiento de la Demencia Fronto-temporal y otras Tauopatías utilizando Compuestos de la Herbolaria Mexicana implementando *Machine Learning*

Para obtener el título de
Licenciada en Ingeniería en Biociencias

P R E S E N T A

Ana Luisa Hernández Cruz

Directores de Tesis

Dr. Aldo Christiaan Jardínez Vera

Dr. Pablo Antonio López Pérez

Apan, Hidalgo., México, Agosto del 2024



Universidad Autónoma del Estado de Hidalgo
Escuela Superior de Apan
Campus Apan

Oficio Núm. ESAp-307/2024
Asunto: Autorización digital de tesis

Mtra. Ojuky del Rocío Islas Maldonado
Directora de Administración Escolar de la UAEH
Presente


Por este conducto le comunico que el jurado que le fue asignado a su trabajo de "TESIS" con el nombre: "Estudios *in silico* sobre la proteína Tau en el Tratamiento de la Demencia Frontotemporal y otras Tauopatías utilizando Compuestos de la Herbolaria Mexicana implementando *Machine Learning*", de la egresada Ana Luisa Hernández Cruz de la Licenciatura en Ingeniería en Biociencias, y que después de revisarlo en reunión de sinodales han decidido autorizar la digitalización del mismo, hechas las correcciones que fueron acordadas.

A continuación, se anotan las firmas de conformidad de los integrantes del jurado:

Presidente	Dr. José Manuel Vásquez Pérez
Secretario	Mtro. Josue Lozada Coronel
Primer vocal	Dr. Pablo Antonio López Pérez
Segundo vocal	Dr. Aldo Christiaan Jardínez Vera
Tercer vocal	Dr. Alberto Navarrete Peón
Primer suplente	Dr. Emmanuel Vallejo Castañeda
Segundo suplente	Dr. José Alfredo Pescador Rojas

Sin otro particular, me es grato reiterar a usted la seguridad de mi atenta consideración y respeto.

ATENTAMENTE
AMOR, ORDEN Y PROGRESO
Apan, Hidalgo; 05 de agosto del 2024


Dr. Jorge Zuno Silva
Director de la Escuela Superior de Apan



Ccp Expediente



Carretera Apan - Calpulapan Km. 8, Colonia Chimalpa Tlalayote, Apan, Hidalgo, México C.P. 43900
Teléfono: 7717172000 Ext. 5800, 5801
esap@uaeh.edu.mx

uaeh.edu.mx

Índice

Resumen.....	11
Abstract.....	12
Abreviaturas y Definiciones	13
Glosario.....	15
Capítulo 1.....	17
Introducción	17
Capítulo 2 Marco Teórico.....	25
2.1 Demencia Frontotemporal (DF): contexto general	26
2.2 Proteína Tau	29
2.2.1 Isoformas de Proteína Tau	29
2.2.2 Modificaciones Post-traduccionales de PTau (fosforilación e hiperfosforilación)	30
2.2.3 Regulación enzimática de la Fosforilación de PTau	31
2.3 Tratamientos actuales de la DF	32
2.4 Herbolaria Mexicana.....	33
2.5 Estudios <i>in silico</i>	38
2.5.1 Enfoques de los estudios <i>in silico</i> : Mecánica Molecular y Mecánica Cuántica.....	40
2.5.2 Dinámicas Moleculares (DM)	41
2.5.3 Construcción de proteínas por Modelado Comparativo.....	41
2.5.4 Acoplamientos Moleculares (AM).....	42
2.6 Inteligencia Artificial (IA)	46
2.6.1 <i>Machine Learning</i> (ML) y Aprendizaje Supervisado.....	47
2.6.2 Algoritmo de Regresión de Bosque Aleatorio (ARBA).....	47
Capítulo 3.....	50
3.1 Planteamiento del problema.....	51
3.2 Justificación.....	52
3.3 Hipótesis.....	53
3.4 Objetivo General	53
3.5 Objetivos Particulares	53
Capítulo 4.....	55
Materiales y Métodos.....	55
4.1 Construcción de la Proteína Tau por Modelado Comparativo.....	56
4.1.1 Selección de secuencia template.....	57
4.1.2 Ensamblaje y selección de la proteína Tau.....	58

4.2	Procesos de Refinamiento y Fosforilación de la proteína Tau	59
4.2.1	Optimización y Fosforilación por <i>CHARMM-GUI</i>	60
4.2.2	Dinámicas Moleculares con <i>Gromacs</i>	64
4.3	Construcción de Base de Datos de los Compuestos de Herbolaria Mexicana (CHM)	67
4.3.1	Selección de Descriptores Moleculares para la base CHM.	70
4.3.2	Optimización de los Compuestos de Herbolaria Mexicana.....	71
4.3.3	Normalización.	72
4.4	Implementación de <i>Machine Learning</i> en la base CHM.....	72
4.4.1	Interpretación de las importancias dadas por ARBA.	75
4.5	Acoplamiento molecular Receptor-Ligantes.....	77
4.5.1	Preparación de la proteína PTau	77
4.5.2	Preparación de los Ligantes	78
4.5.6	Preparación de Acoplamiento Molecular	78
4.5.6.1	Calibración del <i>Gridbox</i>	78
Capítulo 5	80
Resultados y Discusiones	80
5.1	Obtención de Proteína Tau elaborada por Modelado Comparativo	81
5.1.1	Selección del modelo PTau	82
5.1.2	Obtención de PTau	83
5.1.3	Optimización de PTau	84
5.2	Dinámica Molecular para refinamiento de PTau mediante <i>Gromacs</i>	85
5.2.1	Estructura de PTau final	85
5.2.2	Mapa de Potencial Electroestático (MPE).....	86
5.2.3	Raíz de la Desviación Cuadrática Media.....	87
5.2.4	Ramachandran de PTau final.....	88
5.3	Construcción de la Base de datos CHM.....	88
5.3.1	Compuestos optimizados de la Base de datos de CHM	90
5.3.2	Obtención de Descriptores Moleculares de la base de datos de CHM.....	92
5.4	Implementación de <i>Machine Learning (ML)</i>	95
5.4.1	Selección e implementación del Algoritmo de Regresión de Bosque Aleatorio (ARBA)	95
5.4.2	Validación del modelo.....	96
5.4.2	Importancias obtenidas	97
5.4.3	Obtención de <i>Hits</i> y <i>Fails</i>	98
5.5	Acoplamiento Molecular con los <i>Hits</i> y PTau	99

5.5.1 <i>Hits</i> vs Sitio T248.....	101
5.5.2 <i>Hits</i> vs Sitio S411	102
Capítulo 6 Conclusión y perspectivas.....	105
Bibliografía	108
Apéndice	120
Apéndice A “Propiedades ADME y regla de <i>Lipinski</i> ”	121
Apéndice B “Plantillas utilizadas en el Modelado de Tau”	123
Apéndice C “Características del servidor utilizado para los cálculos de DM”	125
Apéndice D “Programación en Python del ARBA”.....	126
Apéndice E “Cargas de <i>Gasteiger</i> aplicadas a los <i>hits</i> ”.....	127
Apéndice F “Error absoluto medio”	128

Lista de Figuras

Figura 1. En la hiperfosforilación de Tau se desestabilizan los microtúbulos alfa y beta, haciendo que se pierda la sinapsis neuronal, este mecanismo es llamado "Hipótesis Tau".	20
Figura 2. Bloque 1, construcción, fosforilación de Tau y selección del mejor modelo.	22
Figura 3. Bloque 2, creación de la base de datos e implementación de Machine Learning.	23
Figura 4. Bloque 3, programación de los Acoplamientos Moleculares en los residuos de interés.	24
Figura 5. Características de la Demencia Frontotemporal.	26
Figura 6. Dominios en Proteína Tau.	29
Figura 7. Mecanismo neurodegenerativo de PTau.	31
Figura 8. Ficha informativa de los compuestos: Acetilcolina, Ácido benzoico, Ácido cafeico y Ácido ferúlico.	34
Figura 9. Ficha informativa de los compuestos: Ácido rosmarínico, Ácido sináptico, Atropina y Convalatoxina.	35
Figura 10. Ficha informativa de los compuestos: Cumarina, Curcumina, Especiofilina y Eucaliptol.	36
Figura 11. Ficha informativa de los compuestos: Plenolina, Santonina y Tujona.	37
Figura 12. Diferencias entre DM y AM.	39
Figura 13. Algoritmo Genético y sus operadores.	43
Figura 14. Posición del ARBA dentro de IA.	48
Figura 15. Aplicación del ARBA.	49
Figura 16. Resumen del diseño de la proteína Tau por Modelado Comparativo en el servidor <i>I-TASSER</i> .	56
Figura 17. Secuencia templete con posibles 40 sitios de hiperfosforilación señalados en color amarillo.	57
Figura 18 <i>Run I-TASSER</i> , se muestra el ingreso de la secuencia <i>templete</i> y la ejecución de <i>I-TASSER</i> sin modificar ningún otro parámetro.	58
Figura 19 Resumen del proceso de obtención y refinamiento de PTau con los servidores <i>CHARMM-GUI</i> y <i>Gromacs</i> .	60
Figura 20. Menú " <i>Solution Builder</i> ".	61
Figura 21. Se muestra el " <i>PDB Manipulation Options</i> " en donde se llevó a cabo la fosforilación de los residuos de acuerdo con la literatura.	62
Figura 22. " <i>Water Size Options</i> " características y tipo de caja para solvatación.	62
Figura 23. Especificaciones para el programa <i>Gromacs</i> .	63
Figura 24. Carpeta GC78.	63
Figura 25. Instrucciones para los cálculos de DM.	64
Figura 26. Comandos de ejecución para la minimización.	65
Figura 27. Comandos de ejecución de la equilibración.	65
Figura 28. Comandos de ejecución de la Dinámicas Moleculares.	66
Figura 29. Archivos obtenidos de la DM.	66
Figura 30. Proceso de minimización de energía de los compuestos recopilados.	71
Figura 31. Cálculo de los Descriptores Moleculares.	71
Figura 32. Serie de comandos para implementar el ARBA.	74
Figura 33. Diagrama "Selección".	76
Figura 34. Parámetros del Algoritmo Genético.	79
Figura 35. Proteínas obtenidas del proceso de Construcción por Modelado Comparativo, y sus valores de <i>TM-score</i> y <i>C-score</i> .	81
Figura 36. Estructura del Modelo 1.	82
Figura 37. PTau a) dentro y b) fuera de la caja de agua, aminoácidos de Serina y Treonina en esferas de color morado.	83
Figura 38. Comparación de Diagramas de Ramachandran.	84
Figura 39. Estructura final de PTau.	86
Figura 40. MPE de PTau final.	86
Figura 41. <i>RMSD</i> de PTau final respecto de PTau inicial.	87
Figura 42. Ramachandran de PTau final.	88
Figura 43. Geometrías optimizadas de la base de datos de CHM.	91
Figura 44. Resultado de las Ponderaciones de las importancias del ARBA sobre la base de datos de los CHM.	97
Figura 45. AM de S411-Fosfato.	100
Figura 46. AM de T248-Fosfato.	100

Figura 47. AM <i>hits</i> vs T248.....	101
Figura 48. AM <i>hits</i> vs S411.....	104

Lista de Tablas

Tabla 1. Principales Tauopatías y sus diferencias.....	20
Tabla 2. Cuadro clínico con los principales rasgos de la DF.....	27
Tabla 3. Variaciones de la Demencia Frontotemporal.....	28
Tabla 4. Isoformas de Tau.....	30
Tabla 5. Enzimas que regulan a P τ en el proceso de Fosforilación.....	32
Tabla 6. Características principales de la Mecánica Molecular y la Mecánica Cuántica.....	40
Tabla 7. Descriptores Moleculares implementados en el diseño de fármacos.....	44
Tabla 8. Compuestos de la base CHM.....	67
Tabla 9. Coordenadas del Gridbox, para su calibración en cajas individuales.....	79
Tabla 10. Ausencia o presencia de propiedades Antinflamatorias y relacionadas con el tratamiento de enfermedades del SNC.....	89
Tabla 11. Descriptores Moleculares calculados para la base de datos de CHM.....	92
Tabla 12. Descriptores Moleculares de THQ4S.....	93
Tabla 13. Normalización de la base de datos de los CHM.....	94
Tabla 14. Selección de compuestos <i>hits</i> y <i>fails</i>	98
Tabla 15. Resultados de AM T248-Fosfato y S411-Fosfato.....	99

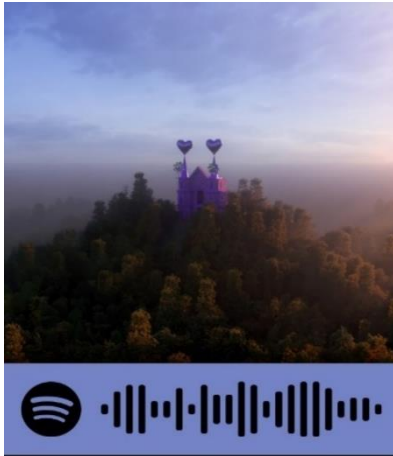
Dedicatorias

A mi padre (†) ...

*“Tantas fotos llenando los marcos, mi propio museo,
no hay muchos trofeos con ustedes tengo.*

*Y aunque te lleve en la sangre me duele sentirte tan lejos,
destellas el cielo y ahora te celebro”*

- Humbe (2023) fantasmas.



A mi madre...

A mis hermanas, sobrinas y cuñados...

A Fátima Lizeth Sandoval y Soraya Gisela Reyes...

“Quiero abrazarte y saber qué se siente renacer en la vida y volverte a amar”

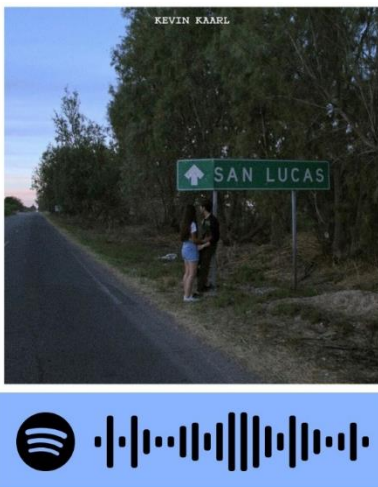
- Siddartha (2022), 00:00.



A mis amigos y compañeros de vida...

“Grita y ponte ya a vivir que todo va a salir mejor”

- Kevin Kaarl (2019), San Lucas.



Agradecimientos

Dr. Aldo Christiaan Jardínez Vera

No tengo palabras suficientes para agradecerle todo lo que ha hecho por mí a lo largo de estos años de conocernos. Gracias por confiar en mí incluso cuando yo no lo hacía, por sus consejos en todos los ámbitos de mi vida, por las horas que dedicó a explicarme y enseñarme cosas nuevas. El impacto que género en mi vida se ve reflejado en todos los logros académicos que hemos conseguido juntos, aún falta mucho camino por recorrer y esto solo es el comienzo.

Dr. Pablo Antonio López Pérez

Gracias por sus consejos, por incentivar en mí el deseo de investigar y salir adelante en la vida, sus enseñanzas, palabras y tiempo invertido en mí hizo esto posible, lo quiero mucho.

Al comité revisor

Dr. José Manuel Vásquez Pérez

Mtro. Josue Lozada Coronel

Dr. Alberto Navarrete Peón

Dr. Emmanuel Vallejo Castañeda

Dr. José Alfredo Pescador Rojas

Gracias por dedicar su tiempo y esfuerzo para transmitirme lo necesario para culminar este proyecto.

A mi Familia

Padre y Madre lo que hicieron y hacen por mí me impulsó a salir adelante para hoy estar cumpliendo este logro, gracias por su crianza y su amor infinito. A mis hermanas Magaly, Liliana y Karla gracias por su apoyo emocional, económico y sobre todo por las palabras de aliento que me hicieron llegar tan lejos, siempre las llevaré en mi corazón. A mis sobrinas Burbuja, Alex y Sari gracias por iluminar mi vida y compartir conmigo su desarrollo como pequeños individuos, las quiero hoy y siempre. A mis cuñados Roberto, Gerardo y Víctor gracias por su apoyo incondicional para llevar a cabo mis sueños. A mi Tobías, gracias por ser mi compañero fiel en estos años, te amo.

A Fátima Lizeth Sandoval Estrada

Creo que mis palabras se quedarían cortas para expresarte lo agradecida que estoy por todo lo que hemos pasado desde hace 5 años, hasta hoy en día. Gracias por abrirme las puertas de tu casa, por cuidar de mí, por ser tan fantástica e inteligente, por compartir conmigo buenos y malos momentos, sobre todo gracias por todas esas noches estudiando juntas hasta lograr entender los temas. Hoy soy la persona que soy por ti, te amo y amaré siempre bolita, estoy muy orgullosa de ti, lo estamos logrando juntas, gracias por cruzarte en mi camino y hacer más amena mi vida.

A Soraya Gisela Reyes Fuentes

Sorayita gracias por ser mi amistad más cercana, compartir contigo estos años ha sido de las mejores experiencias que pude vivir, por ser mi compañera de cuarto y mi cuidadora cuando era necesario, por tus consejos, por las extensas pláticas con Frouz que nos hacían reír horas, por tu existencia, te amo... infinitas gracias.

A la Sra. María Esther Sánchez Garay (†), al Sr. Luis Adrián Sandoval Sánchez y la Sra. Dulce María Ríos García

Gracias por ofrecerme su hogar, como un lugar seguro para poder desarrollarme durante estos años, por compartir conmigo todo su mundo y lo que está a su alcance, por alimentarme, procurarme y aconsejarme. Sra. Mari gracias por sus remedios, su comida y sobre todo por abrigarme con su cariño para hacer más agradable la estancia en Apan, donde quiera que este la llevo en mi corazón y le estaré siempre agradecida.

A mis amigos, compañeros y maestros

A todas las personas que me han acompañado en este camino, gracias totales...

Al grupo de investigación “*Sistemas Energéticos y Materiales Avanzados*” por los recursos computacionales otorgados.

Al “*Laboratorio de Modelado y Simulación Molecular Computacional en Nanomedicina*” por los recursos computacionales otorgados, a mi compañera de laboratorio la *Ingeniera Alejandra Noriega Rubalcaba* por ser mi maestra y amiga durante mis primeros pasos en la química computacional, gracias.

Resumen

En esta investigación se construyó una estructura proteica completa de Tau Fosforilada mediante Modelado Comparativo con técnicas de Dinámica Molecular, dicha estructura contiene los sitios aberrantes donde ocurren las modificaciones post-traduccionales como la hiperfosforilación relacionada a enfermedades pertenecientes al grupo de las Tauopatías entre ellas Alzheimer, Parkinson y la Demencia Frontotemporal (DF). Además, se diseñó una base de Compuestos de Herbolaria Mexicana (CHM) con candidatos inhibidores de la hiperfosforilación de Tau que es la causante de la neurodegeneración en estas Tauopatías, se calcularon Descriptores Moleculares para los CHM evaluados mediante las Reglas de Lipinski y ADME (Absorción, Distribución, Metabolismo y Excreción de fármacos). La base de datos CHM fue sometida a algoritmos de *Machine Learning* (ML) para seleccionar los compuestos ideales (*hits*) en el tratamiento de la DF y descartar los compuestos que no tendrían un buen desempeño (*fails*). Los compuestos *hits* fueron implementados en estudios de Acoplamiento Molecular (AM) rígidos con el fin de analizar su interacción sobre las dianas blanco. De los 40 sitios reportados en la literatura donde se lleva a cabo la hiperfosforilación, nuestra investigación se centró sobre dos residuos específicos, el primero es la serina-411 relacionada con las primeras fases del desarrollo de la DF y el segundo treonina-248 que es fundamental para la unión de estabilización estructural sobre los microtúbulos. Finalmente, los resultados de AM arrojaron una acción competitiva frente a la unión del fosfato mostrando un posible mecanismo que inhiba la hiperfosforilación de Tau.

Abstract

Comparative Modelling with Molecular Dynamics techniques constructed a complete protein structure of phosphorylated Tauopathies in this research. This structure contains the aberrant sites where post-translational modifications occur, such as hyperphosphorylation related to diseases belonging to Tauopathies, including Alzheimer's, Parkinson's and Frontotemporal Dementia (FD). In addition, a database of Mexican Herbal Compounds (CHM) was designed with possible inhibitor candidates of Tau hyperphosphorylation, which is the cause of neurodegeneration in these Tauopathies. Molecular Descriptors were calculated for the CHM evaluated employing Lipinski's Rules and ADME (Absorption, Distribution, Metabolism and Excretion of drugs). The CHM database was subjected to Machine Learning (ML) algorithms to select the ideal compounds (hits) in the treatment of FD and discard compounds that would not perform well (fails). The hits compounds were implemented in rigid Molecular Docking (MD) studies to analyze their interaction with the target determinants. Of the 40 sites reported in the literature where hyperphosphorylation takes place, our research focused on two specific residues, the first one being serine-411, related to the early stages of DF development and the second one threonine-247 which is essential for structural stabilization binding on microtubules. The AM results showed a competitive action against phosphate binding, showing a possible mechanism that inhibits Tau hyperphosphorylation.

Abreviaturas y Definiciones

Acoplamiento Molecular (AM):

Estudio de interacciones moleculares que surgen de la unión entre un receptor y un ligante para la predicción de sus interacciones (ScienceDirect., 2021).

Algoritmo de Regresión de Bosque Aleatorio (ARBA):

Algoritmo de *Machine Learning* de uso común, que mezcla la salida de múltiples árboles de decisión para alcanzar un solo resultado (IBM., 2024).

Aminoácido (a.a):

Unidad base que actúa como estructura de las proteínas, están constituidos por un carbono alfa al cual se unen un grupo funcional amino, uno carboxilo, un hidrógeno y un grupo R o lateral (NIH., 2022).

Compuestos de Herbolaria Mexicana (CHM):

Productos elaborados con material vegetal o algún derivado de la parte aérea o subterránea de una planta. Son extractos, tinturas, jugos, resinas, aceites grasos y esenciales, presentados en forma farmacéutica y cuya eficacia terapéutica ha sido comprobada científicamente (COFEPRIS.,2022).

Demencia Frontotemporal (DF):

Es el término dado para un grupo de trastornos cerebrales neurodegenerativos que afectan los lóbulos frontal y temporal del cerebro, áreas que se asocian con la personalidad, la conducta y el lenguaje (Guimarães *et al.*, 2006).

Dinámica Molecular (DM):

Tipo de simulación molecular computacional que permite analizar el comportamiento y evolución de un sistema a través del tiempo (Meller., 2001).

Inteligencia Artificial (IA):

Es tecnología que permite que las computadoras simulen la inteligencia y las capacidades humanas de resolución de problemas. (IBM., 2024) Rama de la ciencia de la computación empleada en la resolución de problemas cognitivos asociados a la inteligencia humana, como el aprendizaje, creación y reconocimiento de imágenes (NIH., 2022).

Machine Learning (ML):

Rama de la IA que se enfoca en el uso de datos y algoritmos que imitan la forma en que los humanos aprenden, para mejorar gradualmente su precisión (Boden., 2017).

Ovillos neurofibrilares de Tau (NFTs por sus siglas en inglés):

Compilación anormal de proteínas Tau unidas por pequeñas fibrillas dentro de las neuronas (Zhang *et al.*, 2018).

Sistema Nervioso Central (SNC):

Centro de control y procesamiento de información de todas las funciones del cuerpo humano, incluido el pensamiento (Ansari *et al.*, 2020).

Tau Fosforilada (PTau):

Proteína Tau fosforilada de manera normal para vincularse a la unión de los microtúbulos (Dixit *et al.*, 2021).

Tau Hiperfosforilada (HPTau):

Modificación postraducciona aberrante de Tau en sitios específicos que da como resultado agregados anormales causantes de neurodegeneración (Devi., 2023).

Glosario

Aprendizaje Supervisado:

Algoritmo que utiliza un conjunto de datos de entrenamiento para que los algoritmos implementados generen salidas deseadas (Carmona *et al.*, 2015).

Axón:

Conducto de transmisión en donde se propagan impulsos nerviosos (Debanne *et al.*, 2011).

Descriptor Molecular:

Es una propiedad estructural o fisicoquímica de una molécula o alguna de sus partes (Parthasarathi & Dhawan., 2018).

in silico:

Es una expresión que significa “hecho por computador o vía simulación computacional”. Describe al modelado, simulación y visualización de procesos biológicos, físicos y químicos mediante computadoras (Rim., 2020).

Microtúbulo:

Polímeros de tubulina necesarios para la estructura y función celular de la neurona (Devi., 2023).

Neurona:

Célula que forma parte del SNC encargada de transmitir información eléctrica y química (Perea & Araque., 2003).

Tau:

Proteína asociada a microtúbulos que estabiliza los axones neuronales en condiciones fisiológicas normales (Dixit *et al.*, 2021).

Tauopatías:

Enfermedades neurodegenerativas que se caracterizan patológicamente por agregados de Tau anormales en neuronas, responsables de la mayoría de las demencias en todo el mundo (Fedorova., 2019).

Capítulo 1

Introducción

“La ciencia no nos ha enseñado aún si la locura es o no lo más sublime de la inteligencia”.

-Edgar Allan Poe (1809-1849) del cuento "Eleonora" publicado en el libro "Nuevas historias extraordinarias".

“La demencia es en la actualidad la séptima causa de defunción y uno de los principales detonadores de discapacidad y dependencia entre la población de adultos mayores en el mundo”.

-Organización Mundial de la Salud (OMS) (2023).

De acuerdo con la Organización Mundial de la Propiedad Intelectual (OMPI) las Biociencias son las ciencias encargadas del estudio de los seres vivos (OMPI., 2023). Su objetivo principal es innovar o proponer soluciones en los campos de la agricultura, la alimentación, la biología, la medicina y el ambiente.

La era digital ha cambiado los paradigmas clásicos en la investigación de todos los campos de las Biociencias. En específico la industria de la salud no es ajena a la transformación y digitalización de los procesos relacionados con el entendimiento de las enfermedades por medio de la biología computacional (Adler *et al.*, 2011). Una estrategia que ha tomado fuerza durante los últimos años son los estudios *in silico* (Manzoor *et al.*, 2022). El término “*in silico*” hace referencia al diseño de modelos polipeptídicos, celulares o de sistemas biológicos complejos elaborados por simulaciones computacionales (Rim., 2020). Dichos modelos nos permiten explorar y experimentar con sus estructuras tridimensionales, funciones y roles en las que se ven envueltos, hasta poder llegar a predecir respuestas a estímulos farmacológicos (Pelkonen *et al.*, 2011). El desarrollo de proteínas *in silico* relacionadas a padecimientos nos permite analizar su comportamiento molecular frente a candidatos estructurales, que en conjunto con la Inteligencia Artificial (IA), y en específico el *Machine Learning* (ML), brinda la posibilidad de optimizar y acelerar el proceso de selección de

compuestos a aspirantes terapéuticos (Ávila-Tomas *et al.*, 2021). En la actualidad el análisis de los mecanismos relacionados con las enfermedades neurodegenerativas se ve favorecido por las ciencias computacionales, en específico los estudios *in silico* como los Acoplamientos Moleculares (AM) y Dinámicas Moleculares (DM) permiten a los investigadores acelerar el proceso de selección de tratamientos (Makhouri & Ghasemi., 2018).

La proteína Tau pertenece al grupo de polipéptidos encargados de la estructura neuronal y desempeña un papel fundamental en la estabilización de los Microtúbulos, manteniendo la morfología neuronal y contribuyendo al transporte intracelular (Pîrşcoveanu *et al.*, 2017). Las modificaciones postraduccionales ocurridas sobre la proteína Tau, y en específico la hiperfosforilación de dicha proteína, causan el desprendimiento de Tau de los microtúbulos ocasionando daño a la neurona y una acumulación anormal de Tau Hiperfosforilada (HPTau) (Zhang *et al.*, 2018). Este fenómeno biológico cobra relevancia en los últimos años como el responsable del mecanismo neuropatológico de distintos padecimientos neurodegenerativos y se conoce como “Hipótesis Tau”, Figura 1 (Kametani & Hasegawa., 2018), en adición se conocen 40 sitios aberrantes de hiperfosforilación de Tau en donde destacan el relacionado a la unión a microtúbulos “Treonina 248” (T248) y el relacionado a las primeras fases del desarrollo de la DF “Serina-411” (S411) (Reynolds *et al.*, 2000).

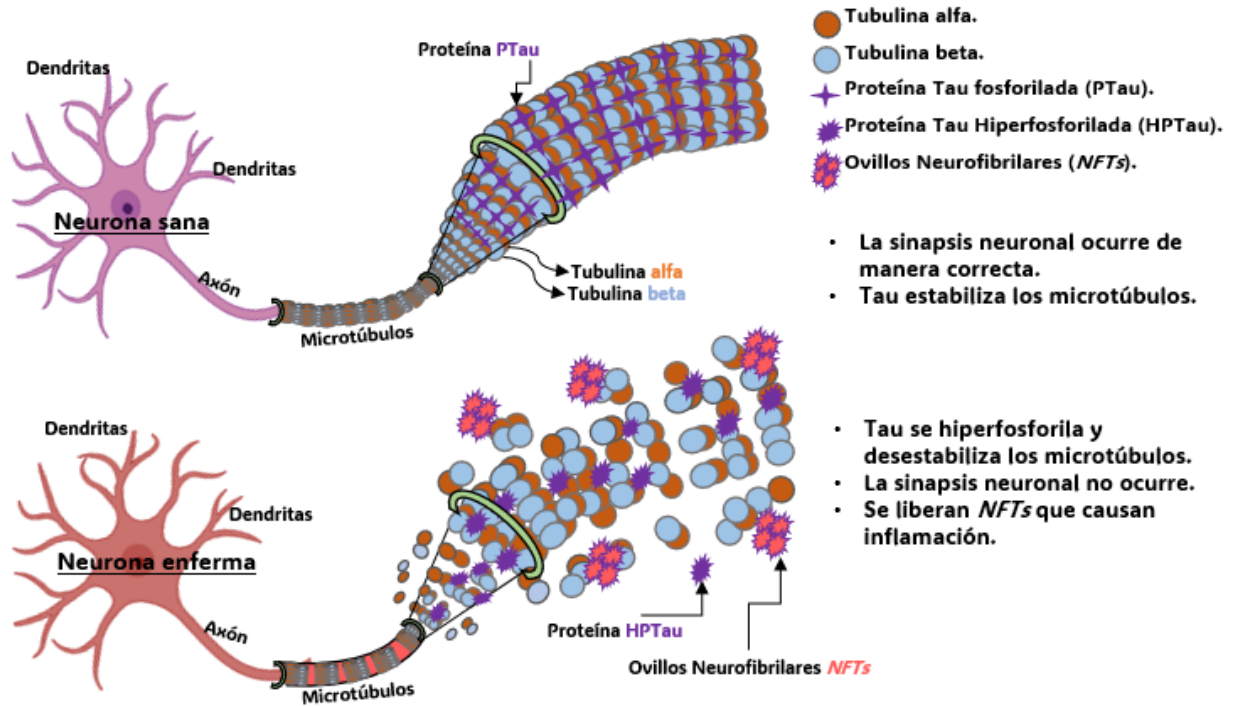


Figura 1. En la hiperfosforilación de Tau se desestabilizan los microtúbulos alfa y beta, haciendo que se pierda la sinapsis neuronal, este mecanismo es llamado "Hipótesis Tau".

Algunos padecimientos como el Alzheimer, el Parkinson y la Demencia Frontotemporal son denominados como Tauopatías (Fedorova., 2019), debido a la presencia de proteínas Tau hiperfosforiladas que ocasionan desajustes en el SNC dañando las estructuras neuronales.

En la Tabla 1 se muestran las principales Tauopatías y sus diferencias, en donde se observa que Tau puede actuar sola o acompañada de otras proteínas para ocasionar el proceso de neurodegeneración (Armstrong., 2016 Bang *et al.*, 2015; Kovacs., 2015; Yoshivama *et al.*, 2012).

Tabla 1. Principales Tauopatías y sus diferencias.

<i>Alzheimer</i>	<i>Parkinson</i>	<i>Demencia Fronto-temporal</i>
------------------	------------------	---------------------------------

<i>Proteínas implicadas en el deterioro cognitivo</i>	Beta-amiloide acompañada de Tau	Alfa-sinucleína en conjunto de Tau	Tau
<i>Zona afectada del SNC</i>	Corteza entorrinal e hipocampo, en etapas avanzadas se observa deterioro en la corteza prefrontal del cerebro.	Degeneración de neuronas dopaminérgicas en la materia negra del mesencéfalo.	Lóbulo frontal y temporal del cerebro.
<i>Sintomatología</i>	Deterioro de la memoria acompañado de cambios en el comportamiento y el juicio.	Trastornos motores en conjunto con deterioro cognitivo y problemas del sueño.	Cambios en el comportamiento con alteraciones del lenguaje y la percepción.

La sintomatología de estas enfermedades tiene como rasgo principal la pérdida de funciones cognitivas acompañada de cambios en el comportamiento de quienes las presentan (Kovacs., 2017). Las investigaciones alrededor de los mecanismos para inhibir la hiperfosforilación de Tau aun no son ampliamente explorados, sin embargo, una alternativa para acelerar el proceso de identificación de compuestos que bloqueen la formación de HPTau son los estudios *in silico* (Dixit *et al.*, 2021).

Esta investigación se divide en tres bloques de experimentación computacional. El primer bloque es la construcción de la proteína Tau mediante Modelado Comparativo a través de Dinámicas Moleculares (DM), ver Figura 2, esto como solución a la inexistencia de un modelo experimental

completo de la proteína Tau, en este apartado también se realizó la Fosforilación de Tau (PTau) de acuerdo con la descripción de su estado en el cerebro humano sano (Dixit *et al.*, 2021).

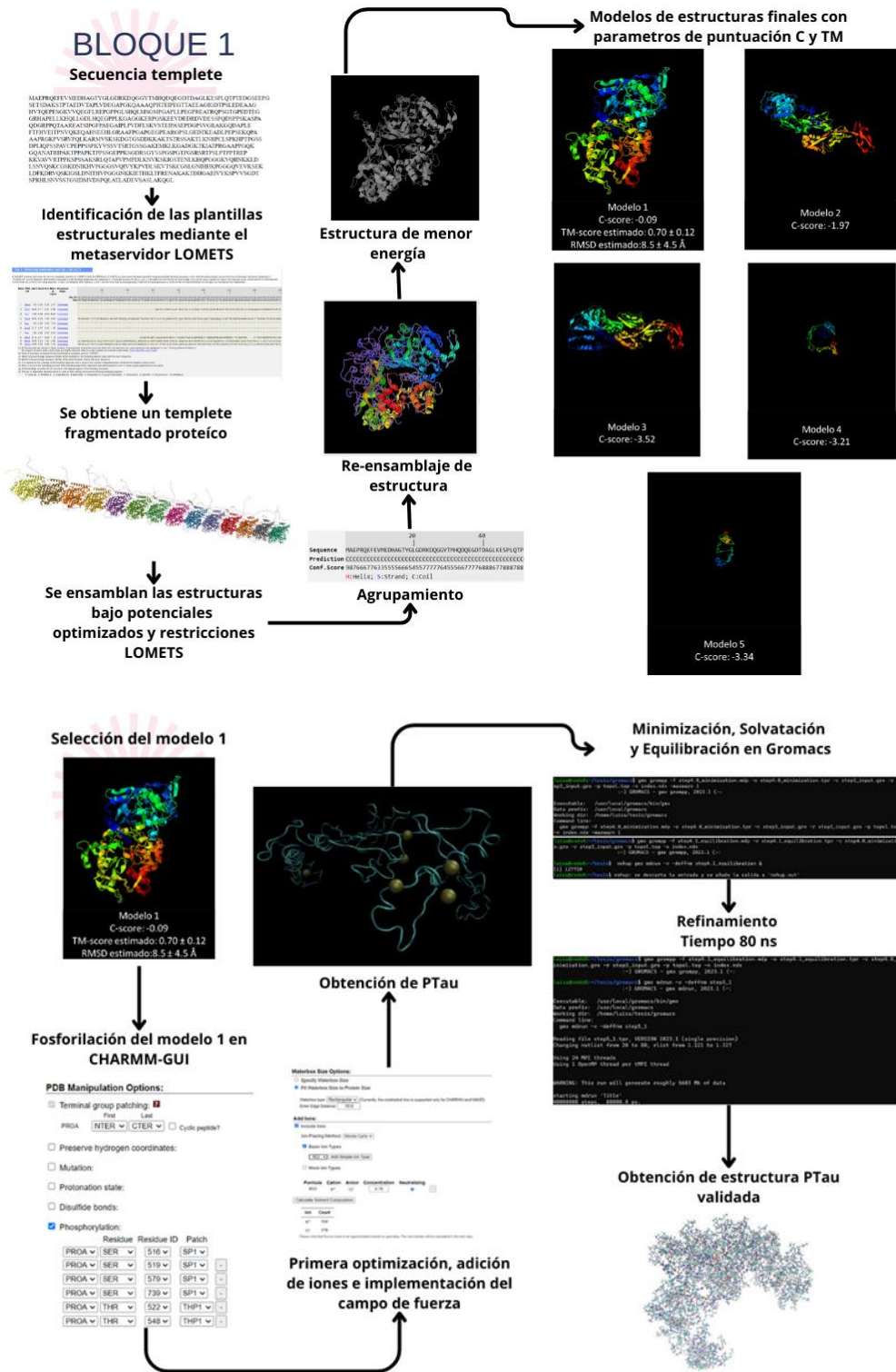


Figura 2. Bloque 1, construcción, fosforilación de Tau y selección del mejor modelo.

En el segundo bloque se realizó la construcción de la base de datos de Compuestos de Herbolaria Mexicana (CHM), ver Figura 3, para su elaboración se tomó como referencia compuestos propuestos en padecimientos del Sistema Nervioso Central (SNC) (Enamorados *et al.*, 2017; Lozoya & Legorreta., 1994), en la selección de moléculas se contempló el uso de propiedades que nos permitieran combatir la inflamación causada por la formación de Ovillos Neurofibrilares de HPTau (*NFTs*) (Ogundele *et al.*, 2014). La base de datos de CHM se analizó a través de algoritmos de *Machine Learning* (ML) correspondientes al grupo de Aprendizaje Supervisado para la selección de candidatos inhibidores de la hiperfosforilación de P τ .

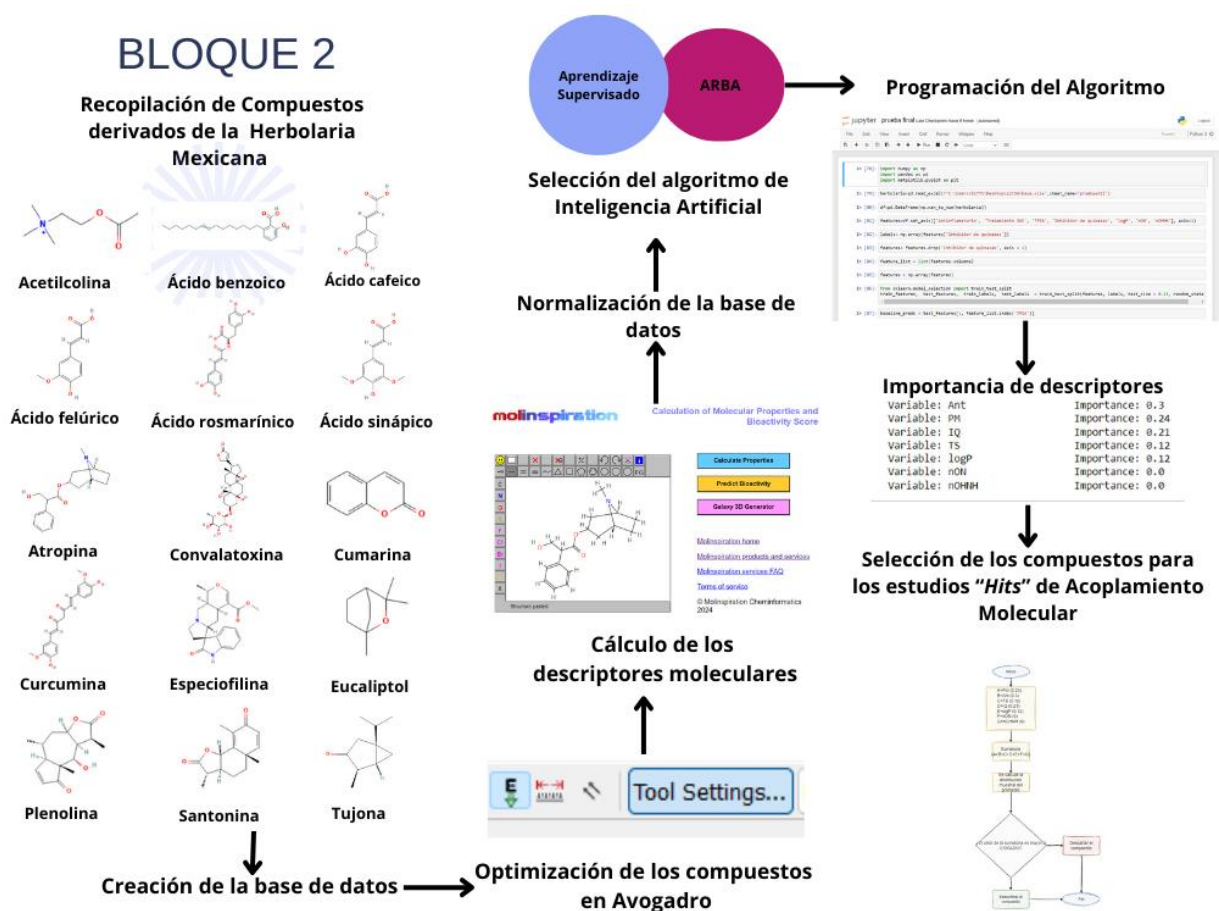


Figura 3. Bloque 2, creación de la base de datos e implementación de Machine Learning.

En la tercera parte de este proyecto se realizaron estudios de Acoplamiento Molecular (AM) con los resultados obtenidos por ML, ver Figura 4, con el fin de proponer nuevos candidatos que inhiban la generación de la HPTau y por consiguiente el deterioro causado por la Demencia Fronto-temporal (DF), dichos AM fueron dirigidos a 2 sitios específicos T248 y S411 que son los relacionados con la unión a los microtúbulos y las primeras fases de la DF (Reynolds *et al.*, 2000).

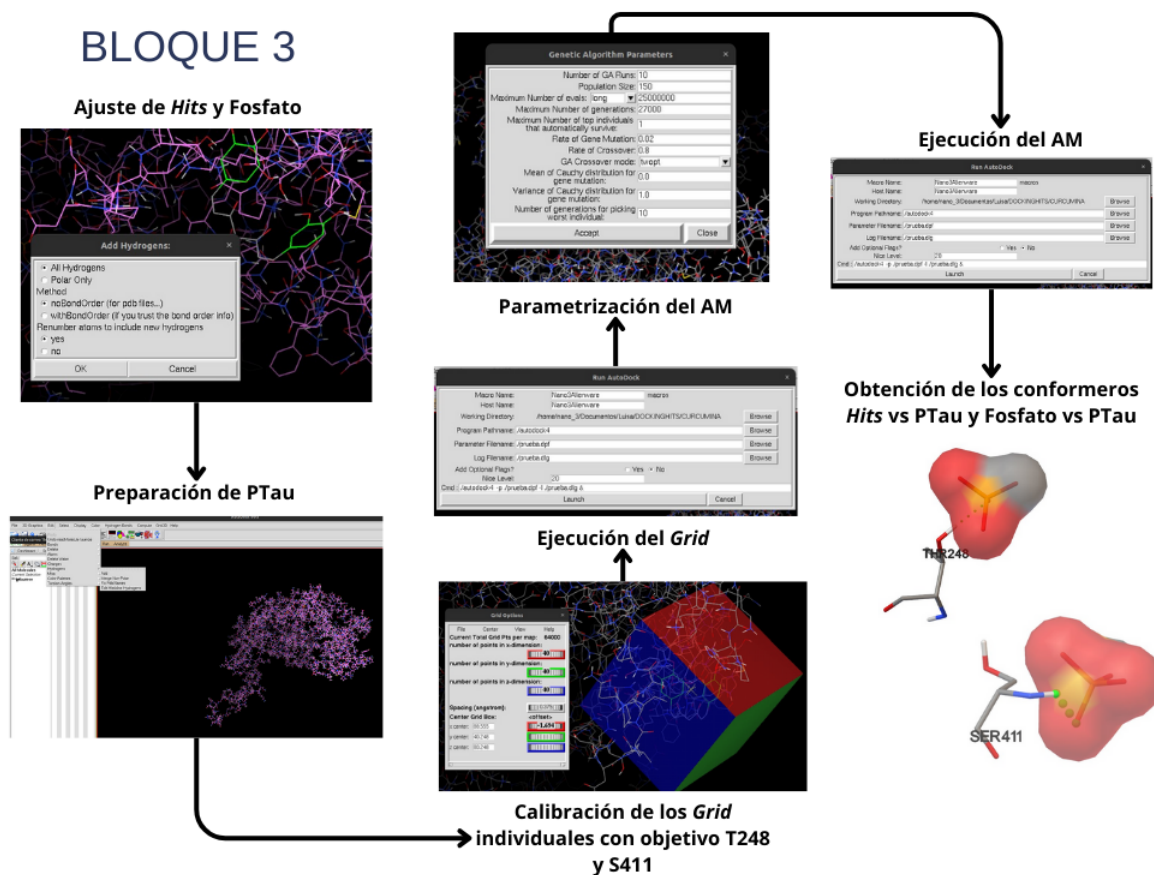


Figura 4. Bloque 3, programación de los Acoplamientos Moleculares en los residuos de interés.

Capítulo 2

Marco Teórico

Demencia Frontotemporal (DF): contexto general

Proteína Tau

Tratamientos actuales de la DF

Herbolaria Mexicana

Estudios in silico

Inteligencia Artificial (IA)

2.1 Demencia Frontotemporal (DF): contexto general

La DF es una proteinopatía denominada Tauopatía neurodegenerativa concentrada en los lóbulos frontal y temporal del cerebro humano. Actualmente se considera la tercera demencia más frecuente alrededor del mundo después de la Enfermedad del Alzheimer y la Demencia de cuerpos de Lewy, sin embargo, dicha enfermedad tiende a aparecer en edades más tempranas que otras demencias, en un rango de 45 a 60 años (Jellinger., 2014). En México no está definida la cifra de personas que presentan DF, debido a la falta de estudios específicos y datos epidemiológicos amplios, se estima que aproximadamente el 14.2% de los adultos mayores de 60 años en adelante presentan algún tipo de demencia (Instituto Mexicano del Seguro Social., 2023). La DF se caracteriza por depósitos interneuronales de proteínas Tau Hiperfosforiladas (HPTau) y acumuladas en forma de Ovillos Neurofibrilares (NFTs), ver Figura 5, que ocasionan un deterioro cognitivo y funcional de quien la padece.

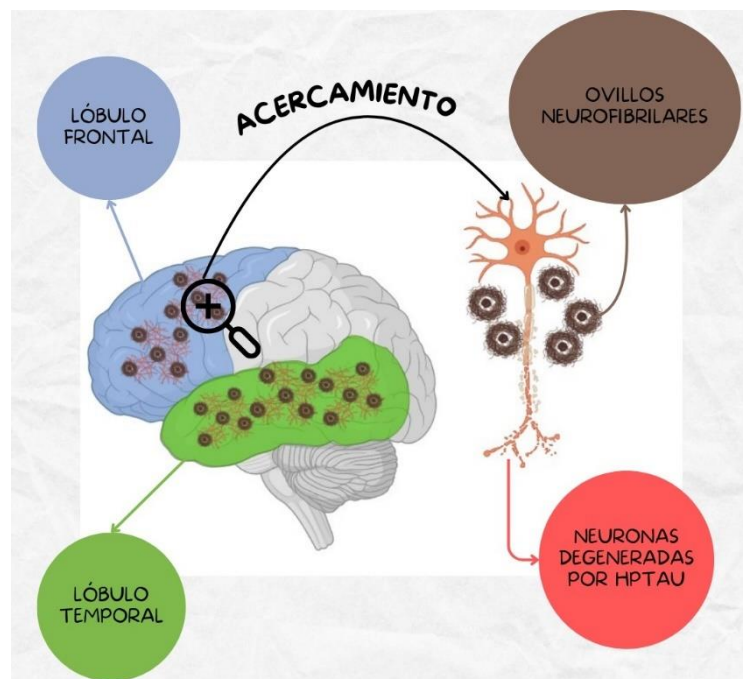


Figura 5. Características de la Demencia Frontotemporal.

Su sintomatología principal son los cambios irracionales en el comportamiento, la disfunción del lenguaje, las dificultades motoras y como síntoma tardío la pérdida de la cognición (NIA., 2024). El cuadro clínico que presentan los pacientes con DF es amplio, Tabla 2, se conoce que existen variaciones de este padecimiento en donde prevalecen un solo tipo de síntomas (Rascovsky et al., 2011).

Tabla 2. Cuadro clínico con los principales rasgos de la DF.

<i>Tipos de síntomas</i>		
<i>Conductuales</i>	<i>Habla y lenguaje</i>	<i>Movimiento</i>
Pérdida de empatía. Conductas antisociales (apatía). Falta de juicio. Ataques de ira. Ingesta de objetos.	Olvidar y errar el significado de palabras. Disociar objetos de su nombre. Imposibilidad de generar ideas claras y completas.	Mala coordinación. Debilidad muscular. Temblores. Rigidez.

La Esclerosis Lateral Amiotrófica (Cervantes *et al.*, 2019), la Afasia Progresiva Primaria (Matías *et al.*, 2013) y la Parálisis Supranuclear Progresiva (Alvarez., 1973), se reportan como variantes de la DF, no obstante, todas comparten la presencia de HPTau como neurodegenerativo en los lóbulos frontal y temporal del cerebro, en la Tabla 3 se especifican las características de estas variantes.

Tabla 3. Variaciones de la Demencia Frontotemporal.

<i>Variación</i>	<i>Breve descripción</i>	<i>Sintomatología principal</i>
<i>Esclerosis Lateral Amiotrófica</i>	Enfermedad producida por la degeneración y muerte de las neuronas motoras, ocasionando la pérdida de conexión con los músculos.	<ol style="list-style-type: none"> 1. Debilidad muscular progresiva. 2. Dificultad para hablar, respirar y comer. 3. Pérdida de la movilidad y atrofia.
<i>Afasia Progresiva Primaria</i>	Síndrome del sistema nervioso que afecta la capacidad para hablar y comunicarse.	<ol style="list-style-type: none"> 1. Dificultad para comprender el lenguaje oral o escrito. 2. Incapacidad para nombrar objetos. 3. Torpeza para formular oraciones.
<i>Parálisis Supranuclear Progresiva</i>	Enfermedad cerebral progresiva que afecta los movimientos oculares y musculares.	<ol style="list-style-type: none"> 1. Pérdida del equilibrio al caminar. 2. Visión borrosa y pérdida del control de la movilidad ocular. 3. Rigidez y dificultad en la marcha.

2.2 Proteína Tau

Tau es la principal proteína asociada a los microtúbulos en las neuronas maduras, su función es, estabilizar la estructura axonal para permitir la conectividad neuronal (Chang *et al.*, 2018). La proteína Tau se expresa en el Sistema Nervioso Central (SNC) y periférico, además de estar presente en riñones, pulmones y testículos (Llorens *et al.*, 2014). Su ubicación se encuentra en el cromosoma 17q21 y es un tipo de Proteína Intrínsecamente Desordenada (PID), se conocen 758 a.a que la integran. Tau tiene ocho dominios distintos: dominio N-terminal (N1, N2), dominio rico en prolina (P1 y P2) y cuatro dominios de unión a microtúbulos.

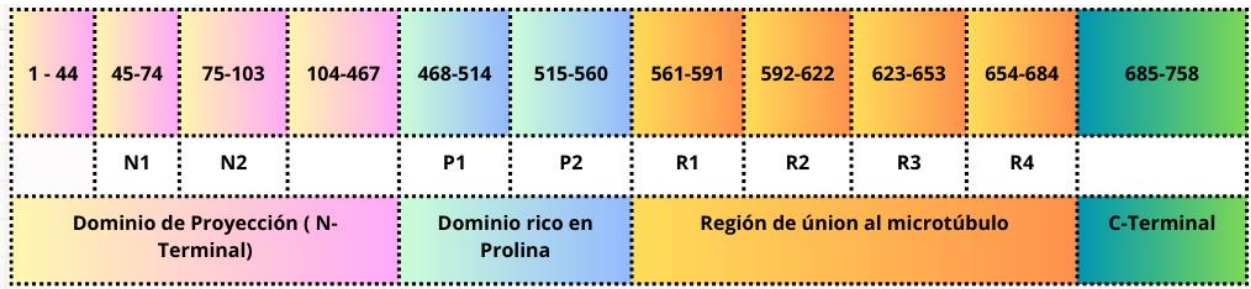


Figura 6. Dominios en Proteína Tau.

2.2.1 Isoformas de Proteína Tau

En la actualidad se conocen 6 isoformas de Tau en el cerebro humano adulto, sus diferencias provienen de la presencia o ausencia de 29 a 58 aminoácidos a causa del proceso de *splicing* alternativo, generando proteínas con 3 y 4 repeticiones en la zona de unión a los microtúbulos (exón 10), en la Tabla 4 se muestran las características de cada isoforma reportada (UniProt., 2024).

Tabla 4. Isoformas de Tau.

<i>Isoforma</i>	<i>Número de a.a</i>	<i>Inclusión del exón 10</i>	<i>Exclusión del exón 10</i>
<i>0N4R</i>	383	Si	No
<i>1N4R</i>	412	Si	No
<i>2N4R</i>	441	Si	No
<i>0N3R</i>	352	No	Si
<i>1N3R</i>	381	No	Si
<i>2N3R</i>	410	No	Si

2.2.2 Modificaciones Post-traduccionales de PTau (fosforilación e hiperfosforilación)

Tau puede sufrir una serie de modificaciones posteriores a su traducción, entre ellas la nitración, ubiquitinación, truncamiento, glicación, oxidación, glicosilación y poliaminación, que influyen en la afinidad que posee PTau con los microtúbulos. En específico se requiere del proceso de Fosforilación de los residuos de Serinas (516,519,579,739) y las Treoninas (522,548) para la unión de Tau a los microtúbulos (Dixit *et al.*, 2021).

Sin embargo, existe una variación en el proceso de fosforilación de PTau, en dicho proceso se lleva a la proteína a un estado de hiperfosforilación en al menos 40 sitios (Dixit *et al.*, 2021) perdiendo afinidad con los microtúbulos, ver Figura 7, y por ende iniciando el proceso de neurodegeneración, este fenómeno cobra relevancia en los últimos años como el causante de un grupo de enfermedades denominadas “Tauopatías” (Fedorova., 2019), que se caracterizan por la presencia de Tau Hiperfosforiladas (HPTau) y marañas compuestas de HPTau (*NFTs*). Se conoce que los residuos afectados en las primeras fases de la DF son los asociados con la unión a los microtúbulos

de los cuales destacan Treonina 248 (T248) y Serina 411 (S411) (Reynolds *et al.*, 2000). Se reporta que para la DF ocurre en al menos 4 de sus isoformas, pero sin descartar que ocurra en todas, siendo un rasgo distintivo frente a otras Tauopatías en donde solo se ve afectada una o dos isoformas de P_{Tau} como la Enfermedad del Alzheimer (EA) (Creekmore *et al.*, 2024).

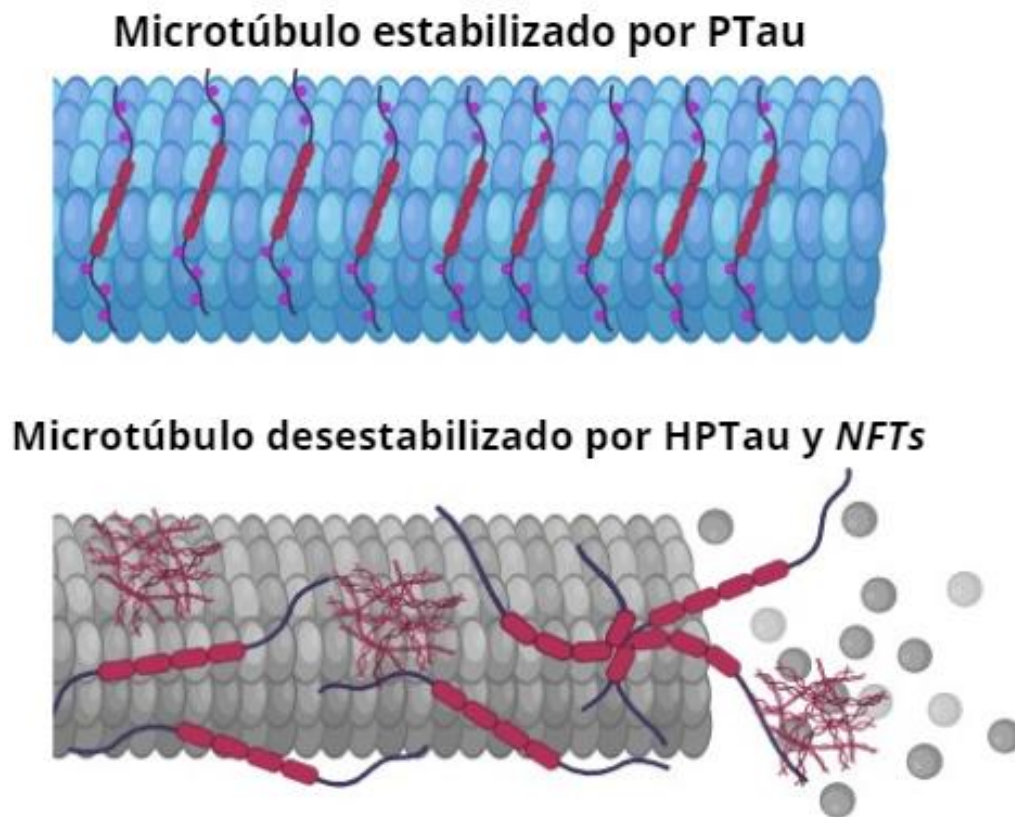


Figura 7. Mecanismo neurodegenerativo de P_{Tau}.

2.2.3 Regulación enzimática de la Fosforilación de P_{Tau}

La interacción entre la Tubulina que compone los microtúbulos y la proteína Tau se regula por la acción de algunas enzimas quinasas que llevan a cabo la Fosforilación y ciertas fosfatasas que eliminan grupos fosfato de la proteína (Avila *et al.*, 2004), ver Tabla 5, ambos procesos ocurren de manera normal en neuronas sanas y en el proceso de formación de HPTau donde aún siguen siendo investigados.

Tabla 5. Enzimas que regulan a P τ en el proceso de Fosforilación.

<i>Quinasas.</i>		Fosfatasas.
Quinasas prolina-dirigidas.	Quinasas no dirigidas por prolina.	PP2A, PP2B y PP2C.
<i>GSK3</i> , <i>cdk5</i> , <i>MAP</i> y <i>JNK</i> .	PKA, PKC y CaMKII.	

2.3 Tratamientos actuales de la DF

Los tratamientos actuales de las Tauopatías son dirigidos a la recaptación o inhibición del neurotransmisor acetilcolinesterasa con el fin de aliviar la sintomatología, pero sin detener la neurodegeneración. En la actualidad no hay resultados efectivos, algunos ejemplos son lecanemab, donepezilo y galantamina (Yiannopoulou & Papageorgiou., 2013). Los tratamientos dirigidos a P τ representan una propuesta prometedora poco explorada, actualmente se desarrollan estudios preclínicos de posibles compuestos que inhiban la formación de HPTau, tal es el caso del derivado de quinolina “THQ4S” (Maccioni *et al.*, 2011) que probó ser efectivo en modelos animales, pero no se reporta éxito del tratamiento en humanos, esto debido a que no se posee una comprensión precisa de la fisiopatología de la DF y otras Tauopatías. Por lo que es necesario buscar estrategias terapéuticas alternativas que cumplan las características de un neuroprotector, es decir, que eviten la progresión y ofrezcan un factor de regeneración o recuperación (Vajda., 2002), en el caso particular de la DF los compuestos seleccionados deberían cumplir los siguientes puntos:

- Aliviar la inflamación causada por la neurodegeneración.
- Inhibir la hiperfosforilación de P τ .
- No causar daño al SNC u otros sistemas del cuerpo humano.

2.4 Herbolaria Mexicana

La Herbolaria se define como la aplicación de plantas o sus derivados para sanar enfermedades (De la Cruz., 1991). El uso de plantas medicinales en México nace junto con su historia, desde la medicina tradicional Tolteca, Azteca y Maya (Lozoya., 1998). Actualmente, la producción de medicamentos sintéticos ha hecho que su eficacia sea menos evidente, sin embargo, sigue siendo indispensable su uso en el desarrollo de nuevos fármacos, los Compuestos de Herbolaria Mexicana como extractos, tinturas, resinas y aceites son implementados en diversas investigaciones en torno a padecimientos del SNC, tal es el caso del Lúpulo y la Valeriana para tratar la ansiedad e insomnio (Jungra *et al.*, 2019). Hace más de 20 años, el 60% de los medicamentos utilizados en México eran a base de plantas medicinales (Lozoya., 1998), en combinación con las tecnologías actuales podría llevar a producir medicamentos herbolarios más seguros, eficaces y económicos (Lozoya & Legorretta., 1994).

Se conoce que el porcentaje de Compuestos de Herbolaria Mexicana estudiados es menor al 10% (Lozoya., 1998). Los compuestos mostrados a continuación son analizados como tratamiento de distintas enfermedades y surgen como propuesta a padecimientos del SNC (Gregory *et al.*, 2021). Sin embargo, en muchos de los casos no se tiene información concreta de sus propiedades individuales, por lo que se tomaron las propiedades generales reportadas para cada planta, como: su información Etnobotánica, su acción farmacológica y sus contraindicaciones por usarlas, ver Figuras 8, 9, 10 y 11.



ACETILCOLINA

COMPUESTO

Planta:
Viscum Album.

Información Etnobotánica:

Se implementan las bayas y hojas frescas para infusiones en el tratamiento de cáncer, diabetes, hipertensión y como neurotransmisor.

Acción Farmacológica:

Estimula la secreción de insulina, induce la apoptosis en los linfocitos y células tumorales.

Contraindicaciones:

En personas con cardiopatías ó insuficiencia renal no se recomienda su uso.

Referencias:

(Urech., 1993)
(Fetrow & Avila., 1999
(Farmacoepa
Homeopática de los
Estados Unidos
Mexicanos., 1998)



ÁCIDO BENZOICO

COMPUESTO

Planta:
Ginkgo biloba.

Información Etnobotánica:

Se implementa en cápsulas con otros compuestos derivados de la planta en el tratamiento de enfermedades cardiovasculares, cáncer, tinnitus y otras afecciones asociadas a la edad.

Acción Farmacológica:

Activador antiplaquetario para enfermedades vasculares e inhibe la agregación del péptido β-amiloide en el SNC.

Contraindicaciones:

Hemorragias intracerebral y trastornos gastrointestinales.

Referencias:

(Smith & Luo., 2004)
(Mahadevan & Park., 2008)



ÁCIDO CAFEICO

COMPUESTO

Planta:
Hypericum perforatum.

Información Etnobotánica:

Se utiliza la planta entera florida y fresca para infusiones o ungüentos para ser aplicados en heridas y quemaduras.

Auxiliar en la prevención de cáncer, aterosclerosis y la Enfermedad de Alzheimer.

Acción Farmacológica:

Actúa como antiinflamatorio, antioxidante y antibacteriano.

Contraindicaciones:

En el embarazo y lactancia no se recomienda su uso.

Referencias:

(Pavlíkova., 2022)
(Czekall *et al.*, 1997)
(Schulz *et al.*, 1998)



ÁCIDO FERÚLICO

COMPUESTO

Planta:
Datura stramonium.

Información Etnobotánica:

Se emplea como analgésico y antiinflamatorio en forma de polvos e infusiones elaboradas de toda la planta. A su vez, se reportan propiedades tranquilizantes y antisépticas.

Acción Farmacológica:

Antiespasmódica, antiasmática y estimulante del SNC.

Contraindicaciones:

En personas con ansiedad y depresión no se recomienda su uso.

Referencias:

(Fetrow & Avila., 1999)
(Gupta *et al.*, 2021)

Figura 8. Ficha informativa de los compuestos: Acetilcolina, Ácido benzoico, Ácido cafeico y Ácido ferúlico.



Planta:
Rosmarinus officinalis.

Información Etnobotánica:
Se elaboran extractos e infusiones de toda la planta, las infusiones se recomiendan para aliviar procesos alérgicos, así como, para mejorar el rendimiento cognitivo.

Acción Farmacológica:
Astringente, antioxidante, antimutágena, antibacteriana y antiviral.

Contraindicaciones:
No se recomienda su uso en niños.

Referencias:
(Petersen & Simmonds, 2003)
(Hitl *et al.*, 2021)
(Kernou *et al.*, 2023)



Planta:
Saccharum officinarum.

Información Etnobotánica:
Se consume en infusión, tintura o polvo como neuroprotector.

Acción Farmacológica:
Antioxidante, antiinflamatorio, anticancerígeno, hepatoprotector, cardioprotector, renoprotector, neuroprotector, antidiabético y ansiolítico.

Contraindicaciones:
Contraindicado en casos de hipersensibilidad a fármacos antiinflamatorios.

Referencias:
(Pandi & Kalappan., 2021)
(Chen., 2016)



Planta:
Atropa belladonna.

Información Etnobotánica:
Con la planta entera se generan polvos, infusiones y pomadas que se aplican en contusiones y llagas, también se utiliza como auxiliar en cefaleas.

Acción Farmacológica:
Reduce las secreciones glandulares, actúa como sedante y presenta una acción parasimpaticolítica, continua siendo investigada.

Contraindicaciones:
En personas con hipertensión arterial, glaucoma, ansiedad, psicosis y depresión no se recomienda su uso.

Referencias:
(Arteche *et al.*, 1998)
(Montastruc & Tran., 1998)
(Schittkowski & Sturm., 2018)



Planta:
Convallaria majalis.

Información Etnobotánica:
Se utilizan las flores, hojas y raíces para infusiones y tinturas para tratar problemas gastrointestinales y cardíacos.

Acción Farmacológica:
Cardiotónica, diurética y purgante.

Contraindicaciones:
Interacciona con otros medicamentos cardiotónicos.

Referencias:
(NIH., 2024)
(Wildgrube *et al.*, 2011)

Figura 9. Ficha informativa de los compuestos: Ácido rosmarínico, Ácido sinápico, Atropina y Convalatoxina.



Planta:
Citrus paradisi.

Información Etnobotánica:
En forma de extracto o polvo se implementa en el tratamiento de afecciones cardíacas.

Acción Farmacológica:
Anticoagulante y antiinflamatoria.

Contraindicaciones:
En personas con insuficiencia renal o enfermedades hepáticas no se recomienda su uso.

Referencias:
(NIH., 2024)
(Liu., 2011)
(Farinola & Piller., 2005)



Planta:
Curcuma longa.

Información Etnobotánica:
A partir de la raíces se elaboran polvos e infusiones que son utilizadas para aliviar la inflamación, equilibrar los niveles de colesterol, tratar alergias, estimular la digestión y aumentar la inmunidad.

Acción Farmacológica:
Antiinflamatoria, antiséptica, antioxidante y antibacteriana.

Contraindicaciones:
No se recomienda su consumo a personas con enfermedades hepáticas o mujeres embarazadas.

Referencias:
(Voulgaropoulou *et al.*, 2019)
(Sharifi *et al.*, 2020)



Planta:
Uncaria tomentosa.

Información Etnobotánica:
Utilizado en forma de polvo o infusión en padecimientos como la artritis reumatoide y el cáncer.

Acción Farmacológica:
Antiinflamatorio, neuroprotector e inmunomodulador en el envejecimiento celular.

Contraindicaciones:
No se recomienda su uso si consumes medicamentos anticoagulantes.

Referencias:
(Gregory *et al.*, 2021)
(Della *et al.*, 2017)



Planta:
Artemisa vulgaris.

Información Etnobotánica:
Se elaboran aceites e infusiones de las ramas de la planta que son utilizados en dolores agudos.

Acción Farmacológica:
Antiinflamatoria, analgésica, antibacteriana, antiepiléptica y depresora del SNC.

Contraindicaciones:
No se recomienda su consumo a niños pequeños ni mujeres embarazadas debido a su acción abortiva.

Referencias:
(Villavicencio & Pérez., 1998)
(Fetrow & Avila., 1999)
(Yin *et al.*, 2020)

Figura 10. Ficha informativa de los compuestos: Cumarina, Curcumina, Especiofilina y Eucaliptol.



Planta:
Arnica montana.

Información Etnobotánica:
Se utiliza en forma de tinturas, infusiones y compresas, posee propiedades astringentes y cicatrizantes.

Acción Farmacológica:
Acción astringente en las mucosas y antihistamínica en músculo liso.

Contraindicaciones:
En personas embarazadas no se recomienda su uso.

Referencias:
(Klass *et al.*, 2002)
(Arteche *et al.*, 1998)
(Foster., 1999)



Planta:
Artemisa marítima

Información Etnobotánica:
La planta entera es utilizada en forma de infusión como antihelmíntica y hepaprotectora.

Acción Farmacológica:
Estimulante cardíaco, antiséptico y tónico. Modulador de acción rápida del receptor GABA.

Contraindicaciones:
En personas con Parkinson, epilepsia y embarazadas no se recomienda su uso.

Referencias:
(Singh *et al.*, 2019)
(Farmacopea Homeopática de los Estados Unidos Mexicanos., 1998)



Planta:
Artemisa absinthium.

Información Etnobotánica:
La planta entera es utilizada en forma de pasta o tintura principalmente para el dolor de estomago y bilis, también se implementa como auxiliar en trastornos de la conducta alimentaria.

Acción Farmacológica:
Produce un incremento de secreciones gástricas además de poseer mecanismos anticonvulsivos y analgésicos.

Contraindicaciones:
En personas con Parkinson, embarazadas y epilepsia no se recomienda su uso.

Referencias:
(Villavicencio *et al.*, 1998)
(Farmacopea Homeopática de los Estados Unidos Mexicanos.,1998)
(Bielenber., 2007)

Figura 11. Ficha informativa de los compuestos: Plenolina, Santonina y Tujona.

2.5 Estudios *in silico*

La frase "*in silico*" proviene del latín "*en silicio*", que hace referencia a la implementación de ordenadores y simulaciones (Rim., 2020). Los estudios *in silico* son investigaciones realizadas mediante simulaciones por computadora para predecir resultados, analizar datos y simular procesos biológicos, químicos o físicos. Estos estudios son especialmente útiles cuando los experimentos prácticos son costosos, peligrosos o éticamente problemáticos, desempeñando un papel crucial en el descubrimiento de nuevos fármacos por varias razones (Manzoor *et al.*, 2022):

1. Reducción de costos y tiempo:

Permiten realizar una gran cantidad de simulaciones y análisis computacionales en un período relativamente corto de tiempo y a un costo mucho menor que los ensayos experimentales en laboratorio. Esto acelera el proceso de descubrimiento de fármacos al reducir la necesidad de realizar múltiples ensayos físicos.

2. Selección de objetivos terapéuticos:

Pueden ayudar a identificar blancos terapéuticos "Dianas" potenciales al predecir la interacción de compuestos químicos con proteínas, enzimas, receptores u otros biomarcadores relacionados con enfermedades. Esto permite a los investigadores centrarse en las moléculas más prometedoras para el desarrollo de fármacos.

3. Diseño de fármacos:

Ayudan al diseño racional de compuestos químicos con propiedades farmacológicas deseadas, mediante técnicas como el modelado molecular y el cribado virtual, los científicos pueden predecir la actividad biológica de nuevos compuestos y optimizar su estructura para mejorar su eficacia y selectividad.

4. Predicción de propiedades farmacocinéticas y toxicidad:

Antes de llevar a cabo ensayos preclínicos y clínicos, los estudios *in silico* pueden predecir propiedades farmacocinéticas (absorción, distribución, metabolismo, excreción) y toxicidad de los compuestos candidatos. Esto ayuda a identificar posibles problemas de seguridad y lograr optimizar la estructura de los fármacos para mejorar su perfil farmacológico.

5. Optimización de procesos de síntesis:

Los estudios *in silico* pueden ayudar a optimizar los procesos de síntesis de fármacos al predecir la viabilidad y eficiencia de diferentes rutas de síntesis química. Esto puede reducir los costos y los residuos asociados con la producción de compuestos farmacéuticos.

La simulación mediante métodos de Mecánica Molecular como la Dinámica Molecular (DM) y los Acoplamientos Moleculares (AM) hacen uso de algoritmos de búsqueda empleando optimización local o algoritmos genéticos estos son ejemplos de herramientas computacionales para realizar este tipo de estudios (Ardon *et al.*, 1987), ver Figura12.

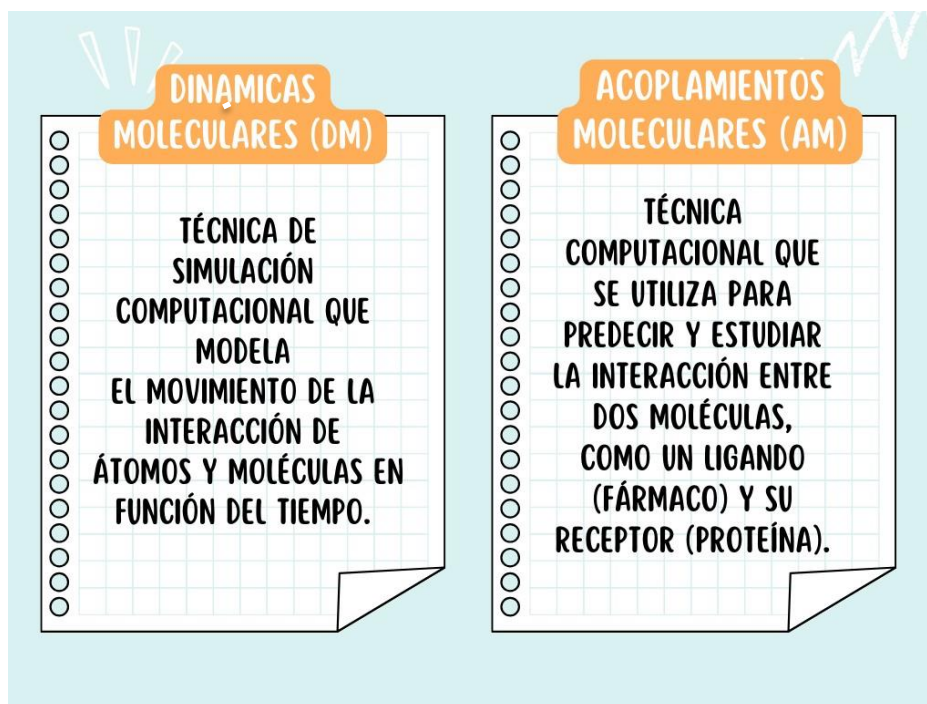


Figura 12. Diferencias entre DM y AM.

2.5.1 Enfoques de los estudios *in silico*: Mecánica Molecular y Mecánica Cuántica

En la búsqueda de fármacos la Mecánica Molecular y la Mecánica Cuántica se utilizan a menudo de manera complementaria, ver Tabla 6, (Boyd & Lipkowitz., 1982). Por ejemplo, la Mecánica Molecular puede proporcionar una aproximación rápida y eficiente para el cribado virtual de compuestos en grandes bases de datos, mientras que la mecánica cuántica puede utilizarse para estudiar detalles más precisos de las interacciones moleculares una vez que se han identificado los candidatos prometedores. En conjunto, estas técnicas ayudan a acelerar el proceso de descubrimiento y diseño de fármacos al proporcionar información detallada sobre la estructura, la actividad y las propiedades de los compuestos candidatos.

Tabla 6. Características principales de la Mecánica Molecular y la Mecánica Cuántica.

<i>Mecánica Molecular</i>	<i>Mecánica Cuántica</i>
<ul style="list-style-type: none"> • La mecánica molecular es una aproximación clásica para describir el movimiento y las interacciones entre átomos y moléculas. • Se basa en los principios de la mecánica clásica. • Las interacciones entre átomos y moléculas se describen mediante potenciales de energía clásicos, como los potenciales de Lennard-Jones para las fuerzas de van der Waals y los potenciales electrostáticos de 	<ul style="list-style-type: none"> • Estudia las características y el comportamiento de las partículas atómicas y subatómicas. • Se basa en la ecuación de <i>Schrödinger</i>. • Se utiliza para estudiar sistemas donde las interacciones electrónicas son críticas, como la estructura electrónica de las moléculas, las reacciones químicas y la predicción de propiedades moleculares.

Coulomb para las interacciones entre cargas eléctricas.

2.5.2 Dinámicas Moleculares (DM)

Tipo de simulación molecular que permite analizar el comportamiento y evolución de un sistema a través del tiempo. Existen dos variantes principales de las DM:

- Mecánica Molecular (MM)

Se basa en la Mecánica Clásica y sus leyes fundamentales, en donde se implementan las leyes de Newton para predecir la trayectoria de todas las partículas en el sistema a lo largo del tiempo y en función de las fuerzas intermoleculares (Ardon *et al.*, 1987; Boyd & Lipkowitz, 1982).

- Química Cuántica (QM)

Combina los principios de la DM de física clásica con la física cuántica para simular sistemas moleculares teniendo en cuenta los efectos cuánticos en la dinámica de los electrones. Un ejemplo de estos métodos es el denominado “*ab initio*” en donde se calculan las propiedades de un sistema molecular directamente desde principios fundamentales de la mecánica cuántica, sin utilizar parámetros empíricos ajustados experimentalmente (Groenhof, 2013).

2.5.3 Construcción de proteínas por Modelado Comparativo

Enfoque utilizado en bioinformática para predecir la composición tridimensional de una proteína utilizando como referencia la estructura de proteínas homólogas ya conocidas. Esta metodología cobra relevancia al permitir el análisis de proteínas que no poseen información cristalográfica para su estudio. Se compone por 5 pasos (Venclovas *et al.*, 2005):

1. Selección de la proteína objetivo.

2. Identificación de proteínas homólogas con estructuras conocidas.
3. Alineación de secuencias.
4. Construcción del modelo.
5. Evaluación y refinamiento del modelo.

2.5.4 Acoplamientos Moleculares (AM)

El Acoplamiento Molecular se puede describir como una herramienta computacional que representa la predicción de la estructura de un complejo intermolecular formado entre dos o más moléculas. Uno de los casos más estudiados de este enfoque son las interacciones proteína-ligando, seguidas de las interacciones proteína-proteína (Meng *et al.*, 2011). El objetivo principal es predecir energías y conformaciones entre ligando receptor (Morris & Lim-Wilby., 2008). Los AM son implementados en medicina debido a que permiten simular y predecir cómo interactúan candidatos a fármacos con sus objetivos biológicos “Dianas”, lo que es significativo para el desarrollo de terapias efectivas y seguras. Existen 2 formas principales de llevar a cabo Acoplamientos Moleculares (Meng *et al.*, 2011):

- Acoplamientos Moleculares rígidos:
 - Se toman a las moléculas como estructuras rígidas.
 - No se permiten cambios en la geometría o la conformación de las moléculas durante la simulación.
- Acoplamientos Moleculares flexibles:
 - Se permite cierto grado de flexibilidad en las partes de una molécula o entre las moléculas.
 - Permite la rotación de enlaces químicos, cambios en la conformación molecular o adaptabilidad estructural en respuesta a las interacciones con otras moléculas.

2.5.2.1 Algoritmos Genéticos (AG)

Son llamados así porque siguen la lógica entorno a la evolución biológica y su base genético/molecular, ver Figura 13, de acuerdo con los operadores (Selección, Cruce, Mutación y Sustitución) se decide cuáles son los individuos más adaptados que sobreviven y cuáles son los menos aptos para ser descartados.

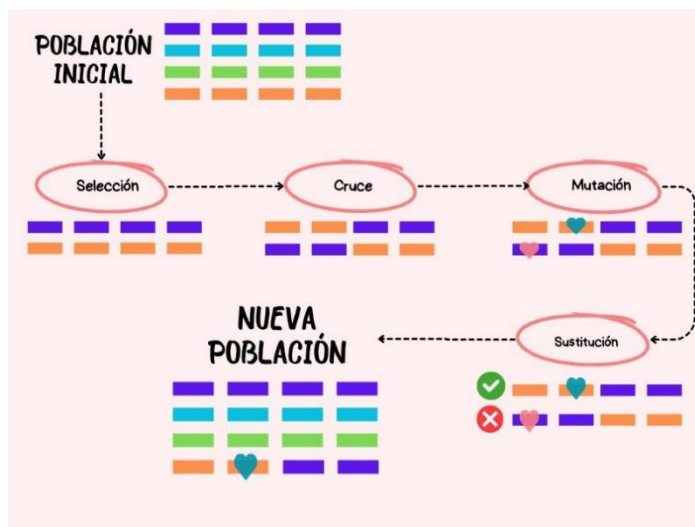


Figura 13. Algoritmo Genético y sus operadores.

Los Algoritmos Genéticos tienen diversas aplicaciones, en el programa *AutoDock* se implementa de la siguiente manera (Taylor *et al.*, 2002):

- Los grados de libertad se codifican en genes o cadenas dobles y al conjunto de genes, o cromosoma, se le asigna un genoma basado en una función de puntuación.
- La codificación de genes se utiliza para representar tanto los diedros giratorios como los enlaces de hidrógeno ligando-receptor.
- Se aplica un operador de movimiento de tipo AG a los cromosomas principales que se eligen aleatoriamente de la población existente con un sesgo hacia los miembros más idóneos.

- Los enlaces de hidrógeno ligando-receptor se combinan con un protocolo de ajuste de mínimos cuadrados para maximizar el número de enlaces de hidrógeno intermoleculares para cada movimiento de los AG.
- Cada estructura se clasifica en función de una función de aptitud más compleja. La función de aptitud (o puntuación) es la suma de un término de enlace de hidrógeno y un potencial de dispersión intermolecular.

2.5.4.1 Descriptores Moleculares

Un descriptor molecular es una característica numérica que se utiliza para describir una molécula en términos de sus propiedades físicas, químicas, estructurales o termodinámicas (Parthasarathi & Dhawan., 2018). En el descubrimiento de fármacos se utilizan para cuantificar y comparar las propiedades de diferentes moléculas, así como para predecir su actividad biológica, toxicidad, solubilidad, entre otras propiedades (Chandrasekaran *et al.*, 2018), esto apoyado de reglas como “*Lipinski*” y las propiedades ADME (la Absorción, la Distribución, la Metabolización y la Excreción), en el Apéndice A se encuentra información adicional. Dentro de la Tabla 7 se encuentran los descriptores moleculares asociados con la búsqueda de nuevos fármacos para inhibir el proceso de hiperfosforilación de Tau.

Tabla 7. Descriptores Moleculares implementados en el diseño de fármacos.

<i>Descriptor</i>	<i>Información</i>	<i>Referencia</i>
<i>Peso Molecular</i> (<i>PM</i>)	<ul style="list-style-type: none"> • Se utiliza para describir el tamaño de una molécula. • Es la suma de las masas atómicas de todos los átomos presentes en la molécula. 	(Atkins & de Paula., 2018)

TPSA	<ul style="list-style-type: none"> • Área de Superficie Polar Total (TPSA por sus siglas en ingles), es un descriptor molecular utilizado en quimioinformática para cuantificar la polaridad de una molécula. (Prasanna & Doerken., 2009) • Se utiliza como un indicador de la capacidad de una molécula para formar enlaces de hidrógeno.
<i>logP</i>	<ul style="list-style-type: none"> • Logaritmo del coeficiente de partición octanol/agua, brinda la medida de la lipofilidad de una molécula. (Vraka <i>et al.</i>, 2017) • Cuantifica la distribución relativa de una molécula entre una fase orgánica (octanol, que es un solvente lipofílico) y una fase acuosa (agua, un solvente hidrofílico).
<i>Inhibidor de Quinasas (IQ)</i>	<ul style="list-style-type: none"> • Describe la capacidad de una molécula para inhibir la actividad de quinasas. (Zhang <i>et al.</i>, 2009) • Se comparan las características estructurales y fisicoquímicas con moléculas con capacidad de inhibir quinasas para encontrar su similitud.
<i>nON</i>	<ul style="list-style-type: none"> • Indica la capacidad de una molécula para aceptar átomos de hidrógeno en una reacción química. (Burden., 2014)

$nOHNH$

- Indica la capacidad de una molécula para donar átomos de hidrógeno en una reacción química. (Burden., 2014)

2.6 Inteligencia Artificial (IA)

Es una rama de la informática que busca desarrollar sistemas capaces de realizar tareas que requieren de inteligencia humana (Rouhiainen., 2018). Estos sistemas utilizan algoritmos y modelos matemáticos para procesar grandes cantidades de datos para obtener predicciones, tomar decisiones o realizar acciones basándose en la información recopilada. La IA posee distintos subcampos dentro de los que se encuentran (Sharifani & Amini., 2023):

- El *Machine Learning (ML)* se centra en el desarrollo de algoritmos con modelos matemáticos que permiten a las computadoras aprender patrones y tomar decisiones basadas en datos.
- El *Deep Learning* se inspira en la estructura y función del cerebro humano, especialmente en las redes neuronales. Utiliza redes neuronales artificiales con múltiples capas para aprender representaciones jerárquicas de datos.

Las aplicaciones de la IA son múltiples, sin embargo, en el campo de la investigación médica destaca su uso por los siguientes puntos (Topol., 2019):

1. Descubrimiento de fármacos: Se pueden analizar grandes bases de datos de compuestos químicos y predecir qué moléculas tienen el potencial de ser fármacos efectivos, optimizando el proceso de selección.
2. Optimización de ensayos clínicos: La IA puede ayudar a diseñar ensayos clínicos más eficientes, identificando criterios de inclusión y exclusión óptimos.

2.6.1 *Machine Learning* (ML) y Aprendizaje Supervisado

En esencia el *Machine Learning*, trata de enseñar a las computadoras a partir de ejemplos y experiencias pasadas para mejorar su rendimiento en tareas específicas. Dentro del *ML* el Aprendizaje Supervisado es un enfoque donde el algoritmo recibe un conjunto de datos de entrenamiento que incluye ejemplos de entrada y las respuestas correctas asociadas con estas. El objetivo es que aprendan los algoritmos una función que mapee las entradas a las respuestas deseadas, basándose en los ejemplos de entrenamiento proporcionados (Bishop., 2006). Los ejemplos comunes de aprendizaje supervisado incluyen los siguientes algoritmos y la combinación de ellos:

- Algoritmos de Clasificación: El objetivo es predecir la clase o categoría de una observación, como "sí" o "no", "gato" o "perro", entre otros ejemplos.
- Algoritmos de Regresión: Donde el objetivo es predecir un valor numérico, como por ejemplo la temperatura o el precio de una casa.
- Algoritmo de Regresión y Clasificación, es la mezcla de los algoritmos mencionados anteriormente en donde el más destacado es el Algoritmo de Regresión de Bosque Aleatorio.

El aprendizaje supervisado es uno de los enfoques más utilizados en el *ML* y se aplica en una amplia variedad de áreas, como reconocimiento de voz, visión por computadora, medicina, finanzas, entre otros.

2.6.2 Algoritmo de Regresión de Bosque Aleatorio (ARBA)

El Algoritmo de Regresión de Bosque Aleatorio (ARBA) es una técnica utilizada en el campo del Aprendizaje Supervisado para resolver problemas de involución, ver Figura 14. Combina dos algoritmos: árboles de decisión y bosques aleatorios. En el desarrollo de este algoritmo se construyen múltiples árboles de decisión independientes y se promedia su salida para obtener una predicción más precisa.

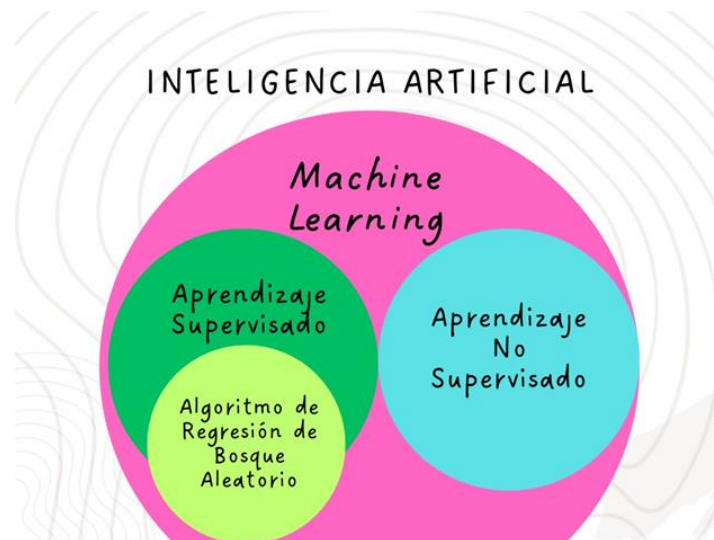


Figura 14. Posición del ARBA dentro de IA.

Una de las ventajas del bosque aleatorio es que tiene la capacidad de manejar grandes conjuntos de datos con una alta dimensionalidad y resistencia a valores atípicos. La metodología ARBA se basa en los siguientes puntos (Liaw & Wiener., 2002), Figura 15:

- Construcción de múltiples árboles de decisión:

Se construye un gran número de árboles de decisiones, cada uno entrenado en una sub-set aleatoria del conjunto de datos de adiestramiento.

- Selección de características aleatorias:

Para cada árbol de decisión, existe una división del árbol, de la cual se selecciona un subconjunto aleatorio de características (predictores) del conjunto completo de propiedades.

- Construcción de árboles de decisión:

Cada árbol de decisión se construye de acuerdo con el algoritmo de clasificación, dividiendo repetidamente los datos en ramas basadas con características previamente seleccionadas, con el objetivo de minimizar la varianza del objetivo en cada nodo.

- Predicción:

Para hacer una conjetura de una nueva instancia, se hace un análisis con cada árbol en el bosque y se promedian los resultados. Esto puede proporcionar una predicción más estable y robusta que la de un solo árbol.

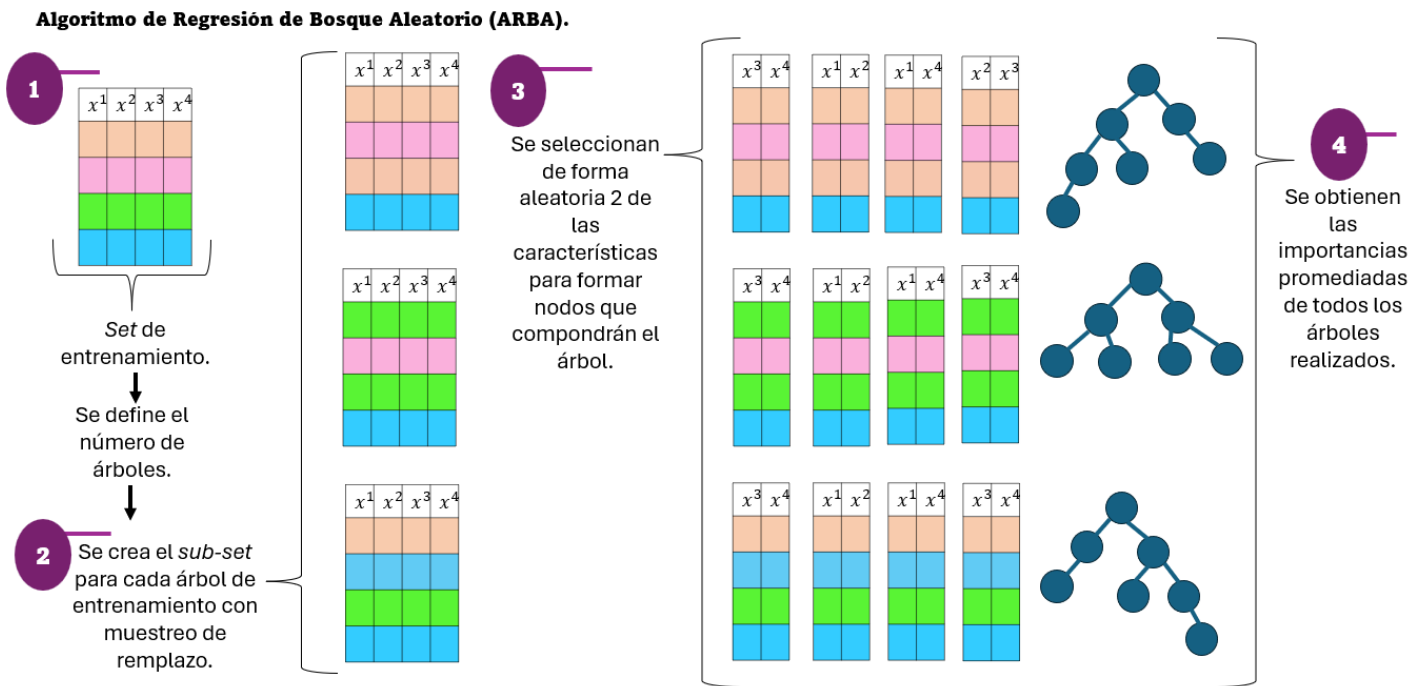


Figura 15. Aplicación del ARBA.

Capítulo 3

Planteamiento del Problema

Justificación

Hipótesis

Objetivo General

Objetivos Particulares

3.1 Planteamiento del problema

La Demencia Frontotemporal (DF) engloba varios síndromes clínicos que implican un cambio progresivo en el comportamiento y/o el lenguaje como una sintomatología común (Perry *et al.*, 2013). Este padecimiento pertenece a la familia de las Tauopatías que son un grupo de enfermedades neurodegenerativas donde la proteína Tau sufre un proceso post-traducciona de Hiperfosforilación (HPTau) (Farwa *et al.*, 2019). Tau está involucrada en el ensamble de monómeros de tubulina dentro de los microtúbulos, que contribuye al mantenimiento de la forma celular y funcionan como vía de transporte dentro de los axones neuronales (Mandelkow *et al.*, 1995). Los cambios patológicos causados por la Hiperfosforilación de Tau conllevan a la separación de la proteína de los microtúbulos y la generación de inclusiones intracelulares tóxicas llamados ovillos neurofibrilares de Tau (*NFTs* por sus siglas en inglés) (Zhang *et al.*, 2018).

Actualmente no existe ningún tratamiento que evite el proceso de daño neurodegenerativo causado por la Hiperfosforilación de Tau. Algunas prescripciones médicas están basadas en fármacos dirigidos a otros padecimientos como la depresión (Kimura & Takamatsu., 2013), alteraciones del sueño (Tisher & Salardini., 2019) y las crisis epilépticas (Rodríguez *et al.*, 2018). Sin embargo, no está comprobada la efectividad de estos medicamentos en la DF. La necesidad de un tratamiento que impida la neurodegeneración causada por la HPTau representa un problema crítico y de salud mundial (Zheng *et al.*, 2021).

La bibliografía muestra poca información sobre los datos experimentales y computacionales acerca de la morfología proteica de Tau (Wang *et al.*, 2021), por ejemplo, no se tiene una estructura cristalográfica experimental completa (Hrushikesh *et al.*, 2020) que nos permita estudiar los sitios de supresión inhibitoria y por ende describir el comportamiento sucesivo a las modificaciones postraduccionales, expresadas en los microtúbulos axonales de las neuronas (Wang *et al.*, 2021).

Algunos compuestos surgen como propuestas tentativas para combatir la Hiperfosforilación de Tau, dentro de estos están los derivados de la quinolina, que son un grupo de compuestos alcaloides que se utilizan para tratar la malaria no complicada por *Plasmodium falciparum*, son antipiréticos y analgésicos (Paintaud *et al.*, 1994). Los estudios alrededor de los derivados de quinolina sobre HPTau han sido dirigidos en tejido cerebral humano post-mortem (Maccioni *et al.*, 2011) y muestran resultados prometedores, sin embargo, el uso prolongado de estos compuestos tiene desventajas como la resistencia a antibióticos (Vandekerckhove & D'hooghe., 2015) y en algunos casos daño hepático (Cano *et al.*, 2017). Una propuesta viable para combatir las enfermedades causadas por HPTau son los compuestos de Herbolaria Mexicana (Enamorados *et al.*, 2017), ya que han probado tener mecanismos de acción eficientes y menos agresivos en el cuerpo humano en el tratamiento de enfermedades como el asma (Cazzola & Matera., 2020), la insuficiencia cardiaca congestiva (Patocka *et al.*, 2020) y como analgésicos en el tratamiento de cáncer (Wiffen *et al.*, 2017) por lo que representan una alternativa sin explorar para la inhibición de la Hiperfosforilación de Tau.

3.2 Justificación

Debido a la poca información relacionada en combatir la problemática mencionada, es necesario un estudio sistemático *in silico* que explique el comportamiento de la estructura y los sitios reactivos que integran la proteína Tau. En este proyecto se pretende utilizar estudios computacionales para determinar la estructura de la proteína Tau en su forma Fosforilada e Hiperfosforilada y analizar el Acoplamiento de los Compuestos de Herbolaria Mexicana seleccionados mediante *Machine Learning* en los sitios de hiperfosforilación de Tau para detener la neurodegeneración ocasionada en la DF.

3.3 Hipótesis

Los compuestos de Herbolaria Mexicana presentarán una mejor actividad *in silico* frente al grupo fosfato en los sitios de hiperfosforilación de Tau, para ser tomados en cuenta como una alternativa de investigación experimental para evitar la neurodegeneración causada por la formación de ovillos neurofibrilares de HPTau y como consecuencia la aparición de la Demencia Frontotemporal.

3.4 Objetivo General

Evaluar el potencial de inhibición mediante estudios *in silico* de Compuestos de Herbolaria Mexicana sobre la proteína Tau, implementando *Machine Learning* para suprimir la desintegración de los microtúbulos y la formación de los ovillos neurofibrilares presentes en la Demencia Frontotemporal y otras Tauopatías.

3.5 Objetivos Particulares

- I. Obtener la secuencia canónica de la proteína Tau para la construcción del modelo polipeptídico.
- II. Construir por Modelado Comparativo la estructura de la proteína Tau con ayuda del servidor *I-TASSER*.
- III. Elegir y calibrar la mejor estructura de Tau para su posterior optimización, con parámetros *C-score* y *TM-score*.
- IV. Preparar la estructura de la proteína Tau resultante del paso anterior para su Fosforilación en los sitios específicos mediante una segunda optimización generando así Tau Fosforilada (PTau) con el servidor *CHARMM-GUI*.

- V. Realizar Dinámicas Moleculares en la proteína PTau con el programa *Gromacs* para obtener una estructura homóloga para los Acoplamiento Moleculares dirigidos.
- VI. Elaborar una base de Compuestos de Herbolaria Mexicana para seleccionar candidatos a estudio de la DF a través de *Machine Learning* (ML).
- VII. Efectuar los Acoplamiento Moleculares de PTau como el receptor y los ligantes seleccionados del filtrado de la base de datos de los CHM con el programa *AutoDock*.
- VIII. Evaluar y analizar los resultados obtenidos de los Acoplamiento Moleculares y compararlos frente a la acción competitiva del grupo fosfato.

Capítulo 4

Materiales y

Métodos

Construcción de la proteína Tau por Modelado Comparativo

Procesos de refinamiento y Fosforilación de la proteína Tau

Construcción de Base de Datos de los Compuestos de Herbolaria Mexicana

Implementación de *Machine Learning* para la selección de ligandos

Acoplamiento Moleculares entre Receptor-Ligantes

4.1 Construcción de la Proteína Tau por Modelado Comparativo

En la literatura reportada se encuentran estructuras cristalinas de Tau construida en fragmentos y con bastantes limitaciones (Dixit *et al.*, 2021), una estrategia para obtener un modelo proteico íntegro es unificar las estructuras fraccionadas y completar las partes faltantes de la secuencia, utilizando un templete de construcción de proteínas por Modelado Comparativo. Esta es una técnica predictiva que nos permite construir una estructura atómica aproximada de una proteína basada en modelos homólogos de referencia experimental (Venclovas *et al.*, 2005). Para este estudio se utilizó el servidor *I-TASSER* por sus siglas en inglés (Refinamiento de Ensamblaje de Subprocesos Iterativos) que nos permite obtener modelos atómicos completos contruidos por



Figura 16. Resumen del diseño de la proteína Tau por Modelado Comparativo en el servidor *I-TASSER*.

simulaciones repetitivas de ensamble en plantillas fragmentadas (Zhou *et al.*, 2022; Zheng *et al.*, 2021; Yang & Zhang., 2015). En la Figura 16, se muestra un resumen del proceso de construcción de la proteína Tau utilizando la técnica de Modelado Comparativo la cual será explicada en una serie de pasos: selección de secuencia templete, ensamblaje de la proteína y obtención de estructuras completas.

4.1.1 Selección de secuencia templete

Se realizó una búsqueda bibliográfica de la proteína Tau a través del servidor *UniProt* (*UniProt Consortium.*, 2023), con el identificador de registro ID: P10636 · *TAU_HUMAN*. Esta exploración recopiló 6 isoformas de la proteína Tau producidas por *splicing* alternativo (Xu *et al.*, 2021), sin embargo, para esta investigación se decidió utilizar la secuencia canónica, la cual contiene 758 aminoácidos (a.a), en la Figura 17 se muestran los a.a que conforman la cadena polipeptídica que se implementó para la construcción del modelo proteico, este *templete* contiene alrededor de 40 sitios de posible hiperfosforilación (Reynolds *et al.*, 2000).

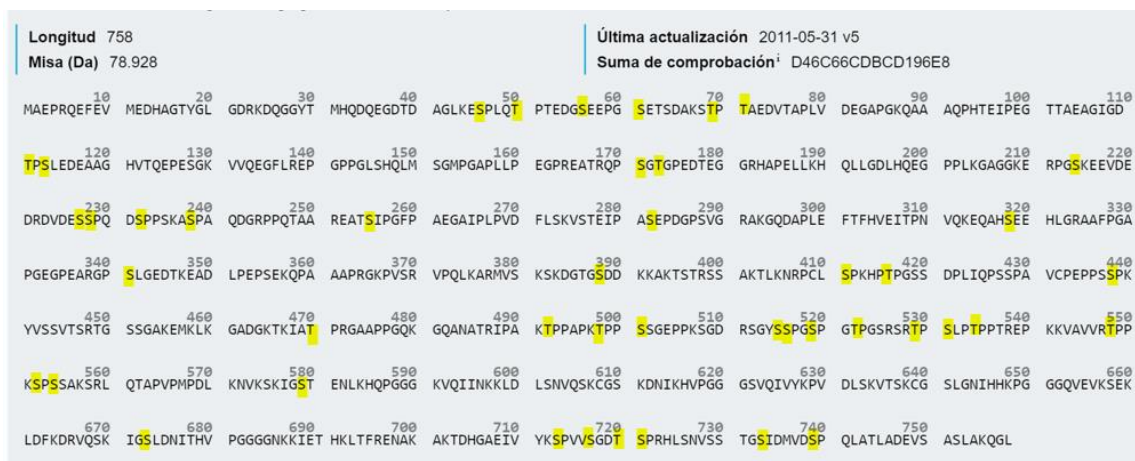


Figura 17. Secuencia templete con posibles 40 sitios de hiperfosforilación señalados en color amarillo.

4.1.2 Ensamblaje y selección de la proteína Tau

Utilizando el servidor *I-TASSER* en el módulo de construcción se introdujo la secuencia *templete* de la proteína Tau (Zhou *et al.*, 2022; Zheng *et al.*, 2021; Yang & Zhang., 2015) y se ejecutó la opción de “*RUN I-TASSER*” sin modificar ningún otro parámetro en el tablero, ver Figura 18.

The screenshot displays the I-TASSER web interface. At the top, there is a navigation bar with links for Home, Research, COVID-19, Services, Publications, People, Teaching, Job Opening, News, Forum, and Lab Only. The main content area features the I-TASSER logo and a brief description of the server's capabilities, including its use of a hierarchical approach and a template-based fragment assembly simulation. Below this, there is a section for the I-TASSER On-line Server, which includes a text input field for a protein sequence in FASTA format. The sequence entered is:

```
SSPQDSPPSKASPAQGRPPQTAAREATSIPGFPAGEATPLPVDFLSKVSTEIPASEPDGSPVGRAGQDAPLEF  
TFHVEITPMVQKEQHSSEHLGRAAFPGAPGEGEARGPSLGEDTKEADLPEPSEKQPAAPRGKPVSRVPLKA  
RNVKSKDGTSSDQKAKTSTRSSAKTLKIRCLSPHPTSSDPLIQPSSPAVCEPPSPKPVYSVSTRGES  
SGAKEPKLKGADGKTLATPRGAAPGQGGQANATRIPAKTPPAKTPPSGEGPPKSGDSEQVSSPQSGPTGSR  
SRTPLPTPTTREPKNVAVRTPPKSPSAKSLQATPVPPDQKIMVSKIGSTESLKHQGGGQVQIINKLIDL  
SNVQKCCGSDIKIHPVGGGQIVYKPVDLKVTSKCGSLGIIHHPGGGQVKESEKLDKDRQVSKTGLSDL  
ITHVPGGGGKIKIETHKLTFRENAKATDHGAEIVYKPVVSGDTPRHLSNVSSTGSDHWDSQPLATLADSV  
ASLAKQGL
```

Below the input field, there are options to upload the sequence from a local computer or to select an archive. The interface also includes a section for additional options, such as assigning restraints, excluding templates, and specifying secondary structure. A checkbox is checked for "Keep my results public". At the bottom, there are buttons for "Run I-TASSER" and "Clear form".

Figura 18 Run *I-TASSER*, se muestra el ingreso de la secuencia *templete* y la ejecución de *I-TASSER* sin modificar ningún otro parámetro.

Posterior al proceso de ingreso de la secuencia *templete*, el servidor recuperó plantillas de la biblioteca *PDB* de proteínas de pliegues similares o estructuras secundarias con aproximación al *templete* de referencia mediante el metaservidor *LOMETS* (*Local Meta-Threading Server*, por sus siglas en inglés) (Zheng *et al.*, 2022; Zheng *et al.*, 2019; Wu *et al.*, 2019) que esta alimentado

por múltiples programas de subprocesos, donde cada programa generó decenas de miles de alineaciones de plantillas. Para finalizar el proceso de selección del servidor *I-TASSER* se utilizó la puntuación *Z*, la cual mide la diferencia entre las puntuaciones brutas y medias de la desviación estándar de las alineaciones de roscado (explicado en el capítulo 2). A valores superiores de $Z > 1$, significa una buena alineación, valores menores o iguales a 1 significa una alineación ineficiente. En el Apéndice B, se muestran las 10 mejores plantillas obtenidas del metaservidor *LO-METS*, con una puntuación *Z* mayor a 1. La topología de la estructura secundaria de los modelos completos se construyó ensamblando los fragmentos alineados de las plantillas mencionadas anteriormente, y las regiones faltantes se crearon desde cero con el método *ab initio* (Zhang., 2008) a través de Simulaciones de Monte Carlo (Zhang *et al.*, 2003) en *SPICKER* (Zhang & Skolnick., 2004) que identificó las conformaciones proteicas de energía libre más baja. Luego *I-TASSER* mediante *FG-MD* (por sus siglas en inglés *Fragment Guide MD Simulation*) (Zhang *et al.*, 2011) y *ModRefiner* (Xu & Zhang., 2011) el cual refinó las estructuras completas mediante simulaciones atómicas. Para finalizar y derivar la función biológica de los modelos *I-TASSER* finales se compararon con las proteínas de la biblioteca *BioLiP* (Yang *et al.*, 2013) para la conservación de características químicas y alineaciones de perfiles de secuencia, obteniendo los modelos finales. El servidor *I-TASSER* arrojó 5 modelos finales, sin embargo, bajo criterio del servidor y parámetros *C-score* y *TM-score* (estos resultados se analizarán en el capítulo posterior), donde el modelo 1 fue el seleccionado para continuar con el proceso de optimización y fosforilación.

4.2 Procesos de Refinamiento y Fosforilación de la proteína Tau

Para optimizar los modelos obtenidos por Modelado Comparativo (MC) y volverlos útiles para análisis bioinformáticos se realizó un proceso de refinamiento estructural (Xiang., 2006), en donde

se vieron involucradas variables como el tiempo, pH y el solvente. Para nuestro estudio fue necesario un paso extra de modificaciones post-traduccionales para lograr obtener PTau. El proceso de refinamiento y Fosforilación se divide en 2 grandes bloques, el realizado con *CHARMM-GUI* (Jo

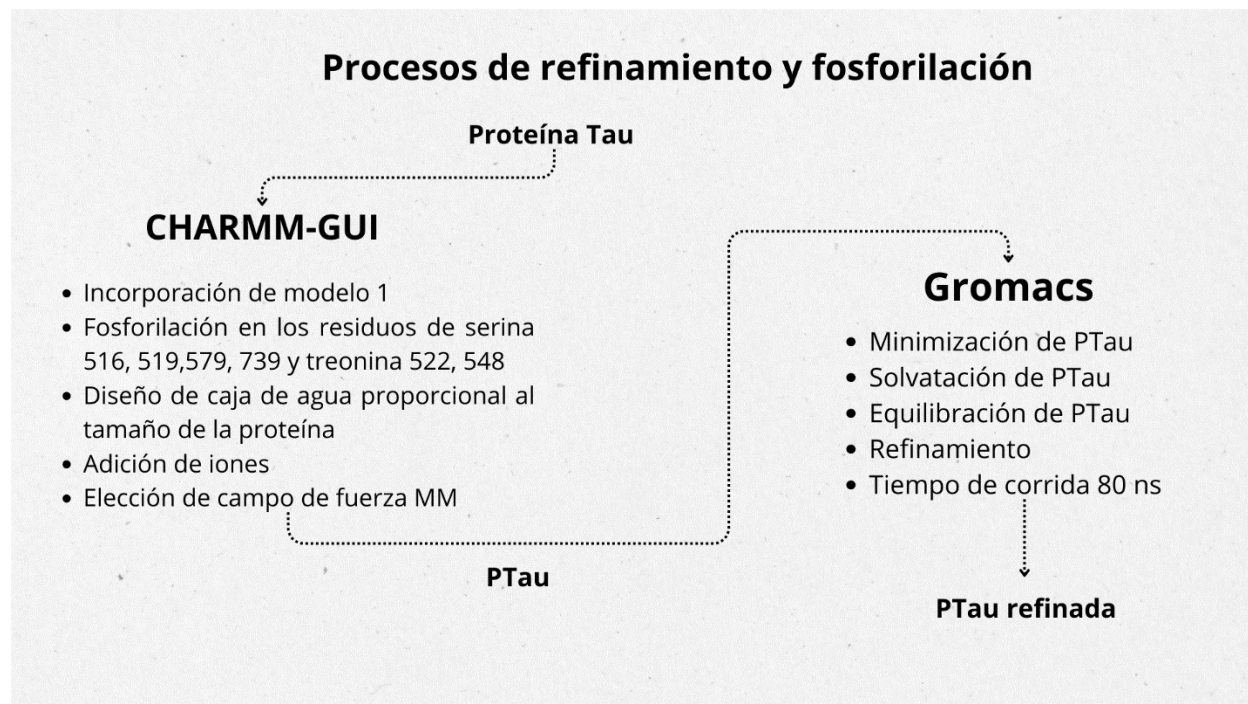


Figura 19 Resumen del proceso de obtención y refinamiento de PTau con los servidores *CHARMM-GUI* y *Gromacs*.

et al., 2008; Brooks *et al.*, 2009; Lee *et al.*, 2016) con el que se llevó a cabo una primera optimización del modelo 1 junto a la Fosforilación de este, obteniendo PTau como resultado, posteriormente en el segundo paso se refinó PTau con Dinámicas Moleculares (DM) en el programa *Gromacs* (Abraham *et al.*, 2015). En la Figura 19 se muestra un resumen de este proceso.

4.2.1 Optimización y Fosforilación por *CHARMM-GUI*

Al acceder al servidor *CHARMM-GUI* se seleccionó la opción “*Solution Builder*” en donde se cargó el modelo 1, automáticamente se desplegó una Tabla como se aprecia en la Figura 20, con las características de la proteína seleccionada. En este apartado se rectificó que todos los datos sean acordes al archivo de entrada, corroborando que no se modificó ningún parámetro.

Free Energy Calculator
LBS Finder & Refiner
Ligand Designer
High-Throughput Simulator
QM/MM Interfacer
PBEQ Solver
Implicit Solvent Modeler
UNICORN Builder
MAP Utilizer
DEER Facilitator
NMR Structure Calculator
Boundary Potential Utilizer
GCMC/BD Ion Simulator

- if you are not familiar with the first PDB reading step, please first watch these [video demos](#).

Reference for Solution Builder:
S. Jo, T. Kim, V.G. Iyer, and W. Im (2008)
CHARMM-GUI: A Web-based Graphical User Interface for CHARMM. [J. Comput. Chem. 29:1859-1865](#)
J. Lee, X. Cheng, J.M. Swails, M.S. Yeom, P.K. Eastman, J.A. Lemkul, S. Wei, J. Buckner, J.C. Jeong, Y. Qi, S. Jo, V.S. Pande, D.A. Case, C.L. Brooks III, A.D. MacKerell Jr, J.B. Klauda, and W. Im (2016)
CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations using the CHARMM36 Additive Force Field. [J. Chem. Theory Comput. 12:405-413](#)
J. Lee, M. Hitznerberger, M. Røger, N.R. Kern, M. Zacharias, and W. Im (2020)
CHARMM-GUI supports the Amber force fields. [J. Chem. Phys. 153:035103](#)

Protein Solution System

Download PDB File: Download Source:

Upload PDB File: model1.pdb

PDB Format: PDB PDB/waterCIF CHARMM


Check/Correct PDB Format

Water Box Only System

Select Water Box Type:

X: Å Y: Å Z: Å

Include Ions

Next Step:  Select Model/Chain

A brief description for each step:

CHARMM-GUI
Effective Simulation Input Generator and More

CHARMM is a versatile program for atomic-level simulation of many-particle systems, particularly macromolecules of biological interest. - M. Karplus

about us :: input generator :: Q&A :: forum :: archive :: lectures :: movie gallery :: video demo :: citations :: update log :: jobs & events :: giving

Some [lectures](#), [job postings](#), and [FAQ](#) are now available. See [update log](#) for update history and [giving](#) for donation. [Contact](#) info is given below.

Input Generator

Job Retriever
Force Field Converter
PDB Reader & Manipulator
Glycan Reader & Modeler
Ligand Reader & Modeler
Glycolipid Modeler
LPS Modeler
Nanomaterial Modeler
Multicomponent Assembler
Solution Builder
Membrane Builder
Martini Maker
PACE CG Builder
Polymer Builder
Drude Prepper

Solution Builder User Profile

PDB Info CHARMM PDB Solvator PBC Setup Input Generator JOB ID: 9750413669

Title				
PDB ID	MODEL1			
Type	Protein			
Experimental Method	Unknown			

Model/Chain Selection Option:

Click on the chains you want to select.

Type	SEGID	PDB ID	Residue ID		Engineered Residues
			First	Last	
<input checked="" type="checkbox"/> Protein	PROA	A	1	758	None

CHARMM-GUI uses internal segid format PRO[A-Z] (protein), DNA[A-Z] (DNA), RNA[A-Z] (RNA), and HET[A-Z] (ligands), instead of PDB chain id.

Figura 20. Menú "Solution Builder"

El siguiente paso se efectuó el llenado del menú "PDB Manipulation Options". En la Figura 21, se muestran los a.a de Serina 516, 519, 579, 739 y Treonina 522, 548 en donde se realizó la Fosforilación del modelo 1 (estos sitios son los reportados en la literatura para la proteína P τ au) (Dixit *et al.*, 2021).

PDB Manipulation Options:

Terminal group patching: ?

First
Last
 PROA Cyclic peptide?

Preserve hydrogen coordinates:

Mutation:

Protonation state:

Disulfide bonds:

Phosphorylation:

Residue	Residue ID	Patch
PROA	SER	516
PROA	SER	519
PROA	SER	579
PROA	SER	739
PROA	THR	522
PROA	THR	548

Figura 21. Se muestra el "PDB Manipulation Options" en donde se llevó a cabo la fosforilación de los residuos de acuerdo con la literatura.

En el apartado "Waterbox Size Options" se muestran opciones predeterminadas de acuerdo con el tamaño del modelo 1 y se seleccionó una caja de tipo rectangular, ver Figura 22, a su vez en la sección de "Add Ions" el servidor también nos arrojó opciones predeterminadas añadiendo los iones mostrados con el método de Monte-Carlo.

Waterbox Size Options:

Specify Waterbox Size

Fit Waterbox Size to Protein Size

Waterbox type: (Currently, the octahedral box is supported only for CHARMM and NAMD)

Enter Edge Distance:

Add Ions:

Include Ions

Ion Placing Method:

Basic Ion Types

More Ion Types

Formula	Cation	Anion	Concentration	Neutralizing
KCl	K ⁺	Cl ⁻	<input type="text" value="0.15"/>	<input checked="" type="checkbox"/>

Ion	Count
K ⁺	194
Cl ⁻	179

Please note that the ion count is an approximation based on geometry. The real number will be calculated in the next step.

Figura 22. "Water Size Options" características y tipo de caja para solvatación.

Para finalizar en el apartado de “*Force Field Options*” se utilizó el campo de energía recomendado por el programa y de acuerdo con el tipo de proteína resultante del modelo 1 se seleccionó “*CHARMM36m*” que se implementa en proteínas intrínsecamente desordenadas; para la opción de “*Input Generation Options*” se escogió la opción de *Gromacs* y en los siguientes apartados no se modificó ningún parámetro, esta programación se muestra en la Figura 23.

Force Field Options:

CHARMM36m ▾

WYF parameter for cation-pi interactions

Hydrogen mass repartitioning Y. Gao, J. Lee, I.P.S. Smith, H. Lee, S. Kim, Y. Qi, J.B. Klauda, G. Widmalm, S. Khalid, and W. Im *Journal of Chemical Information and Modeling* 2021 61 (2), 831-839. |

Multi-site Ca²⁺ Currently only available for CHARMM, NAMD, GROMACS, and OpenMM
Zhang, A., Yu, H., Liu, C. et al. The Ca²⁺ permeation mechanism of the ryanodine receptor revealed by a multi-site ion model. *Nat. Commun.* 11, 922 (2020). [DOI](#).

Input Generation Options:

The input generation scheme has been changed recently. Please check your system, input, and restraint files carefully. Let us know if you see any issue.

NAMD

GROMACS

AMBER

OpenMM

CHARMM/OpenMM

GENESIS

Desmond

LAMMPS

Tinker

Equilibration Input Generation Options:

NVT Ensemble

Dynamics Input Generation Options:

NPT Ensemble

NVT Ensemble

Temperature: K

Figura 23. Especificaciones para el programa *Gromacs*.

El servidor *CHARMM-GUI* dio como resultado la carpeta “*charmm-gui-7892949621*” que denotamos con el nombre “CG78”, para continuar con el refinamiento en el programa *Gromacs* mediante una DM, estos archivos se muestran en la Figura 24.










 index.ndx	Archivo NDX	2,575 KB
 README	Archivo	2 KB
 step3_input.gro	Archivo GRO	8,818 KB
 step3_input.pdb	Archivo PDB	15,480 KB
 step3_input.psf	Archivo PSF	27,582 KB
 step4.0_minimization.mdp	Archivo MDP	1 KB
 step4.1_equilibration.mdp	Archivo MDP	2 KB
 step5_production.mdp	Archivo MDP	2 KB
 topol.top	Archivo TOP	1 KB

Figura 24. Carpeta GC78.

4.2.2 Dinámicas Moleculares con *Gromacs*

Con los resultados de la preoptimización en *CHARMM-GUI* se programa el proceso final de refinamiento. El archivo denominado “*step5_production.mdp*” contiene las instrucciones de configuración de los cálculos necesarios para las iteraciones totales de la DM. Se realizaron las modificaciones de tiempo a partir de 40 millones de pasos obteniendo 80,000 ps equivalentes a 80 ns. En

```
integrator          = md
dt                 = 0.002
nsteps             = 40000000
nstxtcout          = 40000
nstvout            = 40000
nstfout            = 40000
nstcalcenergy      = 400
nstenergy          = 1000
nstlog             = 40000
;
cutoff-scheme      = Verlet
nstlist            = 20
vdwtype            = Cut-off
vdw-modifier        = Force-switch
rvdw_switch        = 1.0
rvdw               = 1.2
rlist              = 1.2
rcoulomb           = 1.2
coulombtype        = PME
;
tcoupl             = Nose-Hoover
tc_grps            = SOLU SOLV
tau_t              = 1.0 1.0
ref_t              = 303.15 303.15
;
pcoupl             = Parrinello-Rahman
pcoupltype         = isotropic
tau_p              = 5.0
compressibility     = 4.5e-5
ref_p              = 1.0
;
constraints        = h-bonds
constraint_algorithm = LINCS
continuation        = yes
;
nstcomm            = 100
comm_mode          = linear
comm_grps          = SOLU SOLV
;
```

Figura 25. Instrucciones para los cálculos de DM.

la literatura se reporta una estabilidad proteica en P τ para este tiempo en DM (Dixit *et al.*, 2021). Los cambios realizados en el archivo de producción se muestran a continuación en la Figura 25. Donde se especifican todos los parámetros integradores de este archivo y la manera de estandarizar el refinamiento molecular. La carpeta “CG78” fue cargada en un servidor con las características

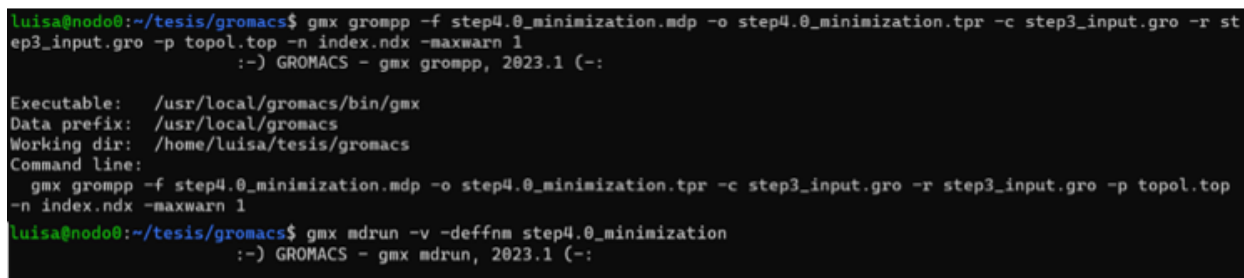
mostradas en el Apéndice C. Los cálculos fueron efectuados bajo condiciones de temperatura y pH fisiológicos.

Para correr el archivo se comenzó llamando al programa *Gromacs*, ver Figura 26, con el atajo preestablecido “*gmx*”. Se continuó con el proceso de minimización al ingresar el comando:

```
“gmx grompp -f step4.0_minimization.mdp -o step4.0_minimization.tpr -c step3_input.gro -r step3_input.gro -p topol.top -n index.ndx -maxwarn 1”
```

Se ejecutó la minimización con el comando:

```
“gmx mdrun -v -deffnm step4.0_minimization”
```



```
luisa@nodo0:~/tesis/gromacs$ gmx grompp -f step4.0_minimization.mdp -o step4.0_minimization.tpr -c step3_input.gro -r step3_input.gro -p topol.top -n index.ndx -maxwarn 1
:~) GROMACS - gmx grompp, 2023.1 (-:
Executable: /usr/local/gromacs/bin/gmx
Data prefix: /usr/local/gromacs
Working dir: /home/luisa/tesis/gromacs
Command line:
gmx grompp -f step4.0_minimization.mdp -o step4.0_minimization.tpr -c step3_input.gro -r step3_input.gro -p topol.top -n index.ndx -maxwarn 1
luisa@nodo0:~/tesis/gromacs$ gmx mdrun -v -deffnm step4.0_minimization
:~) GROMACS - gmx mdrun, 2023.1 (-:
```

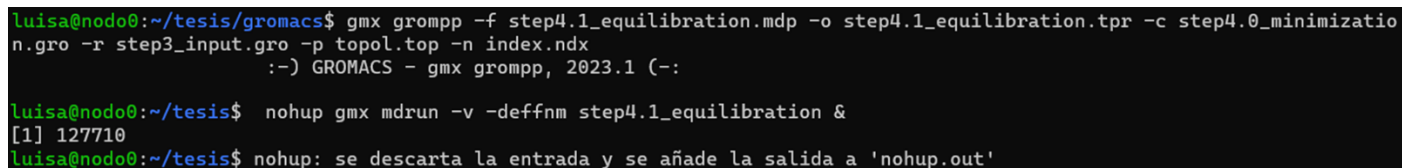
Figura 26. Comandos de ejecución para la minimización.

En el proceso de equilibración se implementaron los siguientes comandos:

```
“gmx grompp -f step4.1_equilibration.mdp -o step4.1_equilibration.tpr -c step4.0_minimization.gro -r step3_input.gro -p topol.top -n index.ndx”
```

La corrida de equilibración se realizó con la instrucción:

```
“gmx mdrun -v -deffnm step4.1_equilibration”
```



```
luisa@nodo0:~/tesis/gromacs$ gmx grompp -f step4.1_equilibration.mdp -o step4.1_equilibration.tpr -c step4.0_minimization.gro -r step3_input.gro -p topol.top -n index.ndx
:~) GROMACS - gmx grompp, 2023.1 (-:
luisa@nodo0:~/tesis$ nohup gmx mdrun -v -deffnm step4.1_equilibration &
[1] 127710
luisa@nodo0:~/tesis$ nohup: se descarta la entrada y se añade la salida a 'nohup.out'
```

Figura 27. Comandos de ejecución de la equilibración.

Mediante el comando siguiente se cargaron los archivos de topología, equilibración, número de moléculas y de producción:

```
“gmx grompp -f step5_production.mdp -o step5_1.tpr -c step4.1_equilibration.gro -p topol.top -n  
index.ndx”
```

El último paso fue la producción final de la dinámica en donde se ingresó el siguiente código:

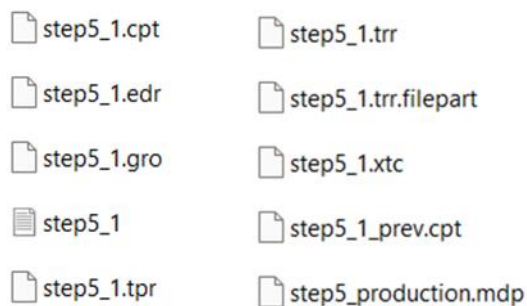
```
“gmx mdrun -v -deffnm step5_1”
```

El tiempo de cálculo total fue de 80 ns, con un recurso computacional de 1,008 horas equivalente a 42 días, ver Figura 28. El estudio final se corroboró llevando a cabo el cálculo por triplicado.

```
luisa@nodo0:~/tesis/gromacs$ gmx grompp -f step4.1_equilibration.mdp -o step4.1_equilibration.tpr -c step4.0_m  
inimization.gro -r step3_input.gro -p topol.top -n index.ndx  
:-) GROMACS - gmx grompp, 2023.1 (-:  
luisa@nodo0:~/tesis/gromacs$ gmx mdrun -v -deffnm step5_1  
:-) GROMACS - gmx mdrun, 2023.1 (-:  
Executable: /usr/local/gromacs/bin/gmx  
Data prefix: /usr/local/gromacs  
Working dir: /home/luisa/tesis/gromacs  
Command line:  
gmx mdrun -v -deffnm step5_1  
Reading file step5_1.tpr, VERSION 2023.1 (single precision)  
Changing nstlist from 20 to 80, rlist from 1.221 to 1.317  
Using 24 MPI threads  
Using 1 OpenMP thread per tMPI thread  
WARNING: This run will generate roughly 5603 Mb of data  
starting mdrun 'Title'  
40000000 steps, 80000.0 ps.
```

Figura 28. Comandos de ejecución de la Dinámicas Moleculares.

Los archivos resultantes de la DM final llevan por nombre step5_1 con las extensiones mostradas en la Figura 29.



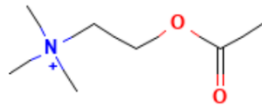
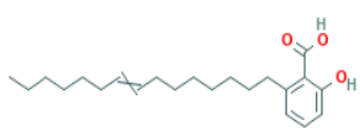
step5_1.cpt	step5_1.trr
step5_1.edr	step5_1.trr.filepart
step5_1.gro	step5_1.xtc
step5_1	step5_1_prev.cpt
step5_1.tpr	step5_production.mdp

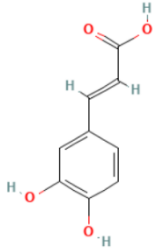
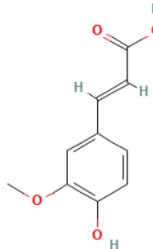
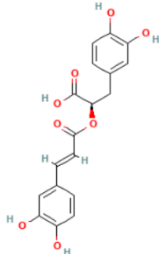
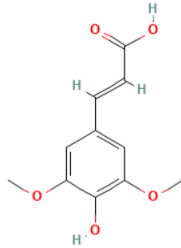
Figura 29. Archivos obtenidos de la DM.

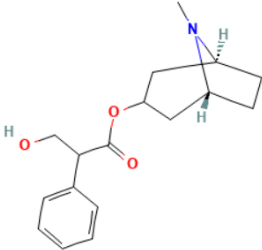
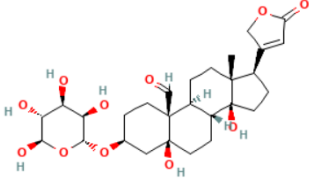
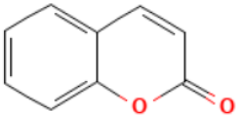
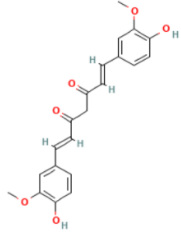
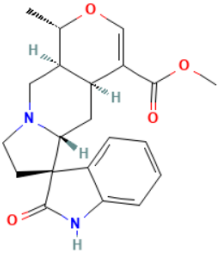
4.3 Construcción de Base de Datos de los Compuestos de Herbolaria Mexicana (CHM)


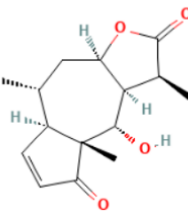
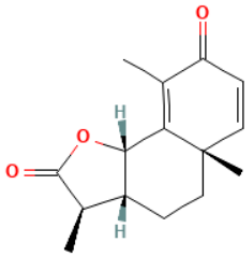
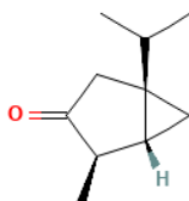
En este proyecto se construyó una base de Compuestos de Herbolaria Mexicana (CHM) que se correlacionaron con parámetros de utilidad en la selección de posibles candidatos en el tratamiento de la DF mediante el uso de *Machine Learning* para obtener los compuestos mejor valuados “*hits moleculares*” que sirvieron para los estudios de Acoplamiento Molecular en el sitio de hiperfosforilación en el modelo P τ au. La obtención de estos compuestos se llevó a cabo en las bases moleculares *PubMed* (Sayers *et al.*, 2022) y *PubChem* (Kim *et al.*, 2023), tomando en consideración sus propiedades reportadas como antiinflamatorias y su uso en el tratamiento de padecimientos en el Sistema Nervioso Central (SNC), la Tabla 8 muestra los compuestos que integraron la base final CHM.

Tabla 8. Compuestos de la base CHM.

Compuesto	Planta Herbolaria Mexicana	Propiedades antiinflamatorias y tratamiento en SNC	CID
Acetilcolina	<i>Viscum album L.</i> , “Muerdago” (Fetrow., 1999)	(Wagner <i>et al.</i> , 1986)	187 
Ácido benzoico	<i>Ginkgo biloba L.</i> , “Ginkgo biloba” (Da Rocha <i>et al.</i> , 2011)	(Gregory <i>et al.</i> , 2021)	368156 

Ácido cafeico	<i>Hypericum perforatum</i> L, “Hierba de San Juan” (Lozoya., 1998)	(Pavlíková N., 2022)	689043	
Ácido ferúlico	<i>Datura stramonium</i> L, “Estramonio” (OMS., 2001)	(Malik <i>et al.</i> , 2019)	445858	
Ácido rosmarínico	<i>Rosmarinus officinalis</i> L, “Romero” (Karadağ <i>et al.</i> , 2019)	(Khaw <i>et a.</i> , 2017)	5281792	
Ácido sinápico	<i>Saccharum officinarum</i> L, “Caña de azúcar” (Vilela <i>et al.</i> , 2017)	(Yoon <i>et al.</i> , 2007)	637775	

Atropina	<p><i>Atropa belladonna</i> L, “Belladona” (Dhuley., 2000)</p>	(Ellinwood <i>et al.</i> , 1990)	174174	
Convalatoxina	<p><i>Convallaria ma-</i> <i>jalis</i> L, “Convalla-</p>	(Liu <i>et al.</i> , 2022)	86574223	
Cumarina	<p><i>Citrus paradisi</i> <i>Macfad</i>, “To-</p>	(De la Rosa <i>et al.</i> , 2016)	323	
Curcumina	<p><i>Curcuma longa</i> L, “Curcuma” (Gre-</p>	(Gregory <i>et al.</i> , 2021)	969516	
Especiofilina	<p><i>Uncaria tomen-</i> <i>tosa</i>, “Uña de</p>	(Baitha <i>et al.</i> , 2020)	168985	

Eucaliptol	<i>Artemisa vulgaris</i>		2758
	<i>L.</i> , “Artemisa” (Ceremuga <i>et al.</i> , 2017) (Stermitz <i>et al.</i> , 2002)		
Plenolina	<i>Arnica montana</i>		442294
	<i>L.</i> , “Árnica” (Lyss <i>et al.</i> , 1997) (Klaas <i>et al.</i> , 2002)		
Santonina	<i>Artemisa marítima</i> <i>L.</i> , “Ajenjo marino” (Wang <i>et al.</i> , 2019)	(Zhang <i>et al.</i> , 2019)	7566380
			
Tujona	<i>Artemisa absinthium</i> <i>L.</i> , “Ajenjo común” (Bielenberg, 2007)	(Bielenberg, 2007)	261491
			

4.3.1 Selección de Descriptores Moleculares para la base CHM.

Para esta investigación se seleccionaron dos Descriptores Moleculares de importancia terapéutica denominados como “Ant” para las propiedades antiinflamatorias y “TS” para el uso en padecimientos del Sistema Nervioso Central (SNC). Se clasificó con 1 si presentaban estas características o con 0 en ausencia de ellas. Posteriormente se seleccionaron los Descriptores Moleculares de bioactividad TPSA, Inhibidor de Quinasas (IQ), logP, nON y nOHNH.

4.3.2 Optimización de los Compuestos de Herbolaria Mexicana.

Los compuestos seleccionados formaron la base de datos y se sometieron a un proceso de minimización de energía en el programa *Avogadro*, este paso se describe en la Figura 30, el campo de fuerza implementado fue MMFF94 debido a su precisión y fiabilidad en la predicción de conformaciones de moléculas orgánicas y biomoléculas.

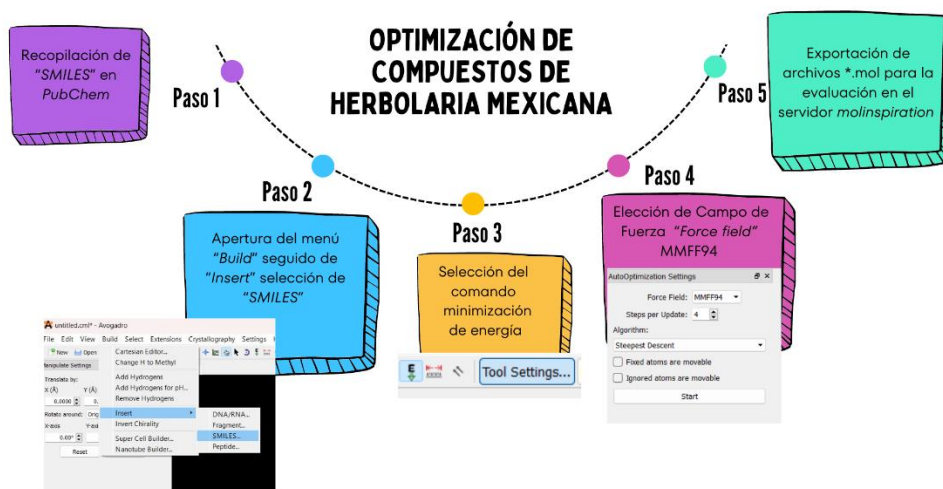


Figura 30. Proceso de minimización de energía de los compuestos recopilados.

En la Figura 31 se muestra la programación del servidor *molinspiration* (Molinspiration., 2023) para la obtención de los descriptores. El cálculo se realizó en el apartado "Calculation of Molecular Properties and Bioactivity Score".

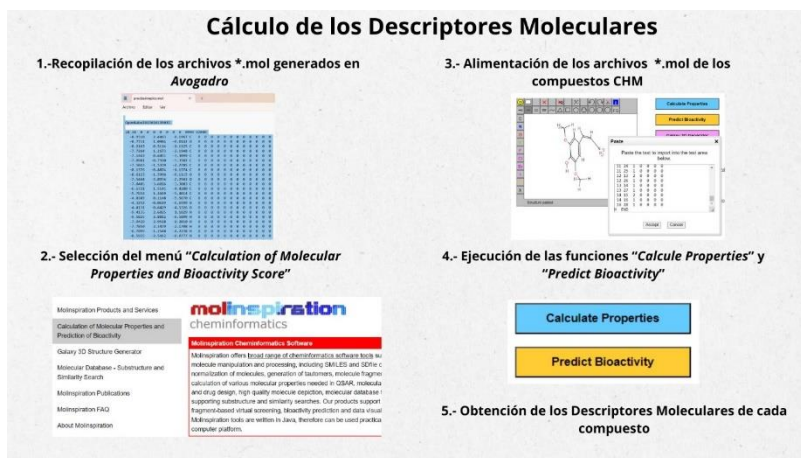


Figura 31. Cálculo de los Descriptores Moleculares.

4.3.3 Normalización.

Los descriptores moleculares utilizados PM, logP, nON y nOHNH se normalizaron aplicando la regla de 5 de “*Lipinski*” (Lipinski *et al.*, 1997), la cual menciona que las moléculas similares a los fármacos con biodisponibilidad tienen que cumplir los siguientes criterios:

- Peso Molecular (PM) menor o igual a 500 g/mol.
- logP menor o igual a 5.
- El número de aceptores de enlaces de hidrógeno (nON) debe ser menor o igual a 10.
- El número de donantes de enlaces de hidrogeno (nOHNH) debe ser menor o igual a 5.

Para el caso del descriptor IQ se evaluó con el rango proporcionado por el servidor *molinspiration*, donde la capacidad de inhibir quinasas se ve reflejada en los valores de -0.5 a 1.5. Para el TPSA no se normalizaron los datos, debido a que todas las moléculas se encuentran dentro del rango deseado para penetrar la barrera hematoencefálica (Hassan & George., 2005), si cumplían los criterios de evaluación por presencia se le colocaba un valor 1 y por ausencia un valor 0.

Estos resultados sirvieron para la construcción de la base denominada CHM, misma que fue programada para la selección de compuestos mediante *Machine Learning*

4.4 Implementación de *Machine Learning* en la base CHM.

Se realizó una búsqueda de Algoritmos de Aprendizaje Supervisado que se clasificaran los compuestos de la base CHM en dos grupos: los candidatos mejor puntuados para Acoplamiento Molecular (que denominamos como: *Hits*) y los compuestos con menor puntuación (etiquetados como: *Fails*).

La implementación propuesta de ML fue el Algoritmo de Regresión de Bosque Aleatorio (ARBA) (Lozada *et al.*, 2022), el cual asignó un porcentaje a las variables basado en su impacto para la

toma de decisiones, esta herramienta evaluó si el compuesto “es” o “no” viable como tratamiento para la DF mediante la aplicación de la regla de Lipinski y los criterios ADME.

La ejecución del ARBA se llevó a cabo en el servidor *jupyter* (Kluyver *et al.*, 2016), mediante un *script* con lenguaje de programación *Python 3*, la rutina completa se encuentra dentro del Apéndice D se siguieron los comandos que se muestran en la Figura 32, realizando la función del Algoritmo basado en el principio de correlación de las variables bidimensionales en un enfoque aleatorio.



IMPLEMENTACIÓN DEL ARBA

IMPORTACIÓN DE LIBRERÍAS

```
In [8]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
```

LLAMADA DE LA BASE CHM

```
In [9]: herbolaria=pd.read_excel(r'C:\Users\52771\Desktop\LUISA\base.xlsx',sheet_name='nuevabase')
```

TRANSFORMACIÓN A PANEL BIDIMENSIONAL

```
In [10]: df=pd.DataFrame(np.nan_to_num(herbolaria))
```

RENOMBRAMIENTO DE LAS COLUMNAS

```
In [11]: features=df.set_axis(['PM','Ant','TS','TPSA','IQ','logP','nON','nOHNH'], axis=1)
```

SELECCIÓN DEL DESCRIPTOR A DESCARTAR "TPSA"

```
In [12]: labels= np.array(features['TPSA'])
In [13]: features= features.drop('TPSA', axis = 1)
```



Figura 32. Serie de comandos para implementar el ARBA.

4.4.1 Interpretación de las importancias dadas por ARBA.

Con las importancias obtenidas del proceso del ARBA, se realizó una multiplicación del número contenido en la base de datos por el cociente de importancia, dando origen a las siguientes variables:

$$A_1 = PM(0.24),$$

$$A_2 = Ant(0.30),$$

$$A_3 = TS(0.12),$$

$$A_4 = IQ((0.21),$$

$$A_5 = \log P(0.12),$$

$$A_6 = nON(0),$$

$$A_7 = nOHNH(0)$$

Posteriormente se realizó la sumatoria por filas (k) de las variables, a este parámetro se le denominó porcentaje (P_k) dentro de la base CHM:

$$\sum_{n=1}^7 A_{kn} = \text{porcentaje}(P_k), \quad k = 1, 2, \dots, 14$$

Con los porcentajes obtenidos se estimó la media poblacional ($\hat{\mu}$) de los compuestos:

$$\sum_{k=1}^{14} \text{porcentaje}(P_k) =$$

$$\begin{aligned} &P_{\text{Acetilcolina}} + P_{\text{Ác.benzoico}} + P_{\text{Ác.cafeico}} + P_{\text{Ác.ferúlico}} + P_{\text{Ác.rosmarínico}} \\ &+ P_{\text{Ác.sinápico}} + P_{\text{Atropina}} + P_{\text{Convalatoxina}} + P_{\text{Cumarina}} + P_{\text{Curcumina}} \\ &+ P_{\text{Especiofilina}} + P_{\text{Eucaliptol}} + P_{\text{Plenolina}} + P_{\text{Santonina}} + P_{\text{Tujona}} \end{aligned}$$

$$\hat{\mu} = \frac{\sum_{k=1}^{14} \text{porcentaje}(P_k)}{\text{número total de compuestos}} = \frac{\sum_{k=1}^{14} \text{porcentaje}(P_k)}{14}$$

La media estimada estableció el valor para la selección de los compuestos, bajo las siguientes condiciones:

$$\begin{cases} \text{Si } P_{k \text{ compuesto}} > \hat{\mu} \Rightarrow \text{se selecciona} \\ \text{Si } P_{k \text{ compuesto}} < \hat{\mu} \Rightarrow \text{se descarta} \end{cases}$$

4.4.2.1 Diagrama de flujo del proceso de selección

Con el fin de simplificar el proceso de selección de los compuestos de la base CHM, se construyó el Diagrama “Selección”, que incluye los resultados de las condiciones mencionadas en la sección anterior, con este paso se clasificó en dos grupos a los compuestos:

- *Hits*: Compuestos mejor puntuados y seleccionados para los estudios de Acoplamiento Molecular.
- *Fails*: Compuestos con menor puntuación y descartados para los estudios de Acoplamiento Molecular.

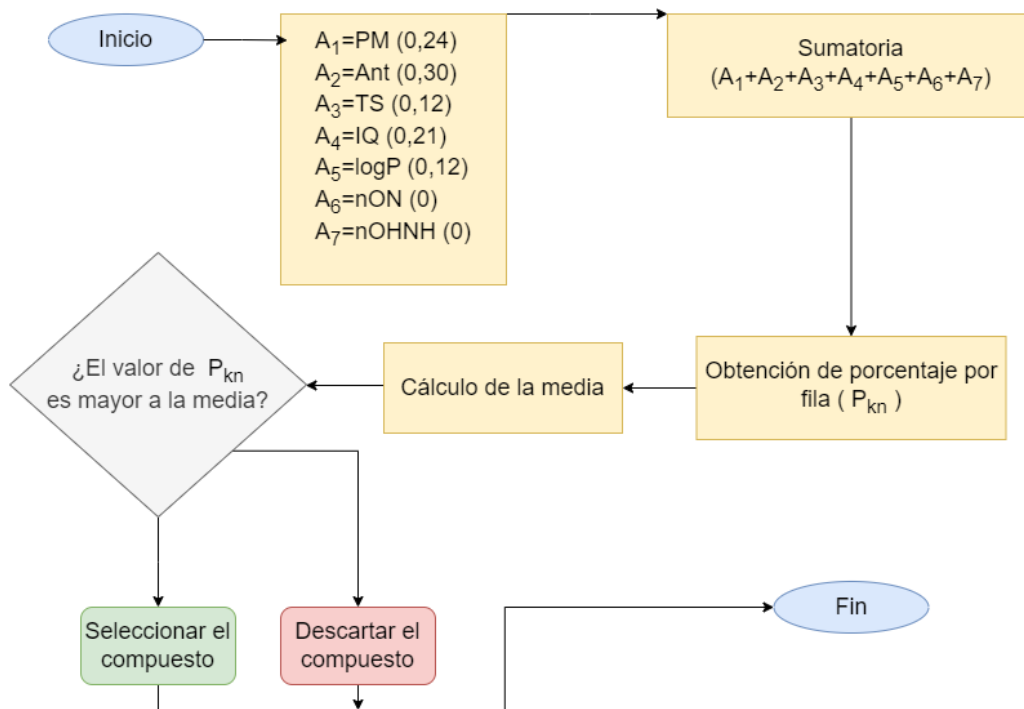


Figura 33. Diagrama "Selección".

4.5 Acoplamiento molecular Receptor-Ligantes

Los estudios de Acoplamiento Molecular entre Receptor y los Ligantes se desarrollaron en el software *AutoDock* (Garrett., 2010), la optimización y construcción de los ligandos se llevó a cabo en *Avogadro* (Marcus *et al.*, 2012). Los compuestos *hits* obtenidos de la sección anterior son los implementados en estos estudios. El tipo de Acoplamiento Molecular desarrollado en esta investigación corresponde a ligando flexible con receptor rígido.

La proteína Tau se fosforila de forma habitual en 6 sitios y se hiperfosforila de manera aberrante en 40 sitios (Reynolds *et al.*, 2000). Dentro de este último grupo existen residuos asociados a las primeras fases de la DF, tal es el caso de Serina 411 (Jicha *et al.*, 1999). A su vez existen aminoácidos que forman parte del sitio a unión de los microtúbulos y en específico Treonina 248 afecta la unión entre P τ y los microtúbulos (Reynolds *et al.*, 2000), ambos aminoácidos mencionados anteriormente representan nuevas dianas de investigación en la inhibición de hiperfosforilación de Tau, por lo que se decidió tomarlas como blanco para este estudio.

4.5.1 Preparación de la proteína P τ

Con los archivos generados en la sección de optimización (*step5_1.gro* y *step5_1.xtc*) se creó el archivo **P τ .pdb** en *VMD* (Humphrey *et al.*, 1996) seleccionando únicamente la proteína y suprimiendo la caja de agua. Posteriormente se cargó el archivo **P τ .pdb** al programa *AutoDock* para su preparación siguiendo los pasos enlistados a continuación:

- 1) Adición de hidrógenos no polares.
- 2) Implementación de la carga de *Kollman* (-51.68)
- 3) Creación del archivo **P τ .pdbqt**

4.5.2 Preparación de los Ligantes

La preparación para los acoplamientos consistió en los siguientes pasos:

- 1) Adición de hidrógenos
- 2) Implementación de la carga de *Gasteiger* (la información acerca de las cargas proporcionadas se encuentra en la Tabla “Cargas de *Gasteiger*” en la sección de Apéndice E)
- 3) Creación de los archivos en formato. pdbqt

4.5.2.1 Construcción y optimización del grupo Fosfato

La evaluación de la competitividad entre la unión de los *hits* y la proteína fue medida por la unión que ocurre entre el grupo Fosfato y PTau en los sitios aberrantes, por lo tanto, fue necesario la construcción de la molécula Fosfato y su optimización siguiendo los pasos mencionados en la sección de “Optimización de los Compuestos de Herbolaria Mexicana” aplicados al Fosfato.

4.5.6 Preparación de Acoplamientos Moleculares

En este estudio se realizaron 7 Acoplamientos Moleculares por cada sitio (Serina 411 y Treonina 248), 6 correspondientes a los *hits* y 1 a la unión PTau-Fosfato dando un total de 14 Acoplamientos con características especiales para cada caso.

4.5.6.1 Calibración del *Gridbox*

Los Acoplamientos Moleculares fueron realizados para la PTau de forma rígida, estableciendo la búsqueda entorno al ligando, seleccionando los criterios (*set map types, choose ligands*) y las especificaciones generales del *Gridbox*: 40 \AA_x , 40 \AA_y , 40 \AA_z y un *Spacing grid*: 0.375. En la Tabla 9 se muestran las coordenadas (*Center Gridbox*) implementadas para cada *hit* en presencia de los a.a.

Tabla 9. Coordenadas del Gridbox, para su calibración en cajas individuales.

Ligantes (<i>hits</i> y grupo Fosfato)	Coordenadas	
	Center Gridbox en T248	Center Gridbox en S411
Fosfato	57.893, 93.410, 60.099	90.142, 39.172, 77.744
Ácido cafeico	56.842, 93.381, 60.881	90.251, 40.248, 80.740
Ácido sinápico	57.115, 93.704, 59.292	92.645, 40.248, 84.907
Curcumina	59.015, 91.716, 57.189	90.251, 40.248, 81.020
Especiofilina	55.843, 88.465, 59.183	90.251, 40.248, 82.769
Eucaliptol	60.374, 91.993, 60.381	90.300, 40.300, 83.010

4.5.6.2 Parametrización del Acoplamiento Molecular

En el apartado de parámetros de Algoritmos genéticos que se utilizó para la parametrización del Acoplamiento Molecular, se indicó a PTau como receptor rígido, los valores de búsqueda fueron los que se muestran en la Figura 34 y se implementó el Algoritmo *Lamarckian*.

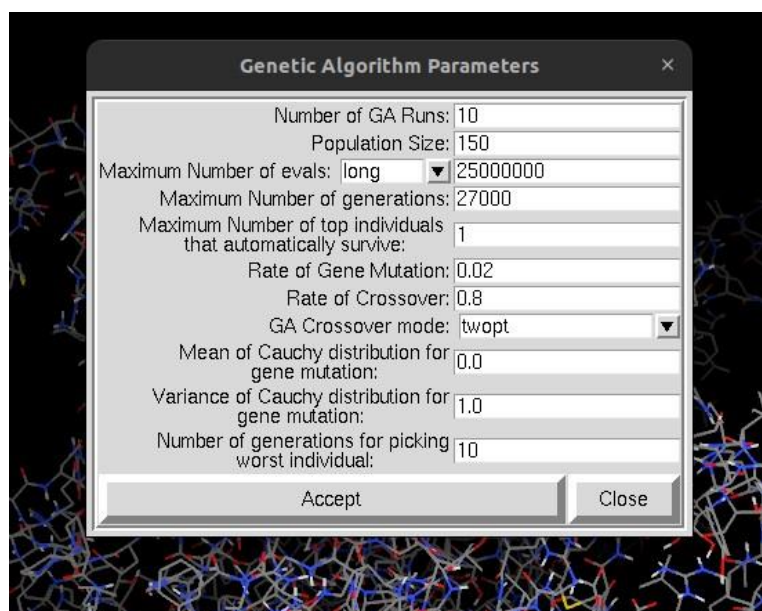


Figura 34. Parámetros del Algoritmo Genético.

Capítulo 5

Resultados y

Discusiones

Obtención de Proteína Tau elaborada por Modelado Comparativo

Dinámica Molecular para refinamiento de PTau mediante *Gromacs*

Construcción de la base de datos CHM

***Machine Learning* sobre CHM**

Acoplamientos Moleculares entre PTau y los *hits* de CHM

5.1 Obtención de Proteína Tau elaborada por Modelado Comparativo

Por medio del programa “*I-TASSER*” se obtuvieron un total de 5 modelos proteicos de Tau, ver Figura 35. A los cuales se les aplicaron cálculos de valoración *C-score* y *TM-score* para evaluar su construcción:

- El *C-score* es implementado por el servidor “*I-TASSER*” para estimar la calidad del modelo predicho, basado en la alineación de la plantilla y los parámetros de convergencia de la estructura, en donde el valor más alto de “5” indica un modelo de mayor confianza.
- Cuando el *TM-score* toma un valor de “1” indica una coincidencia perfecta entre las estructuras homólogas, sin embargo, cuando toma un valor de “0” nos dice que las proteínas no están relacionadas, esto nos permite analizar si la construcción se llevó a cabo de manera correcta.

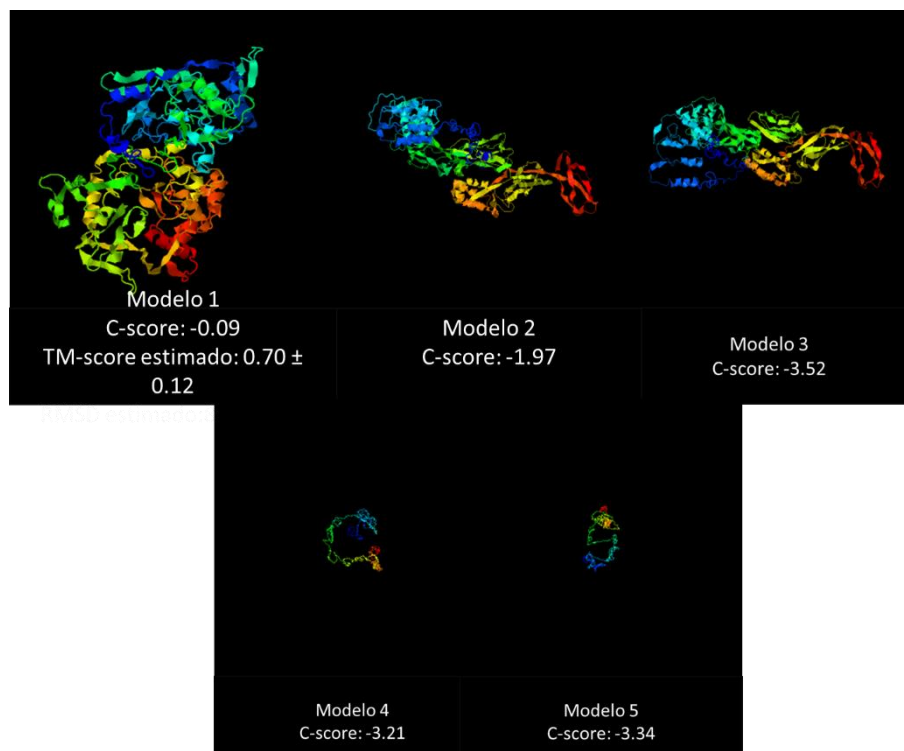


Figura 35. Proteínas obtenidas del proceso de Construcción por Modelado Comparativo, y sus valores de *TM-score* y *C-score*.

En la Figura 36 podemos observar que el cálculo de *TM-score* solo fue aplicado al Modelo 1, debido al descarte de la viabilidad por el servidor de los modelos restantes. Estas estructuras proteicas concuerdan con lo esperado a las proteínas intrínsecamente desordenadas en donde no se posee una estructura bien definida, sin embargo, en los últimos años se ha reportado que esto no afecta su funcionalidad en condiciones fisiológicas, aunque no se descarta que al unirse con otras macromoléculas defina su estructura.

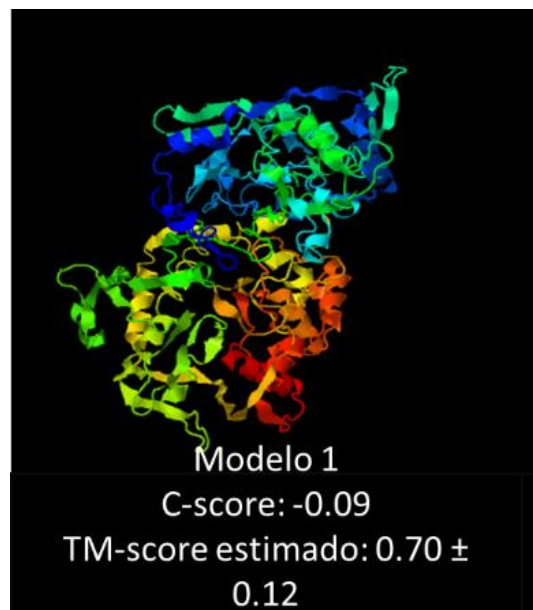


Figura 36. Estructura del Modelo 1.

5.1.1 Selección del modelo P τ

De acuerdo con lo mencionado en la sección anterior el Modelo 1, se seleccionó para continuar con el proceso de fosforilación y refinamiento. Los criterios para tomar en cuenta fueron los siguientes:

C-score: -0.09

- De las 5 proteínas obtenidas, el Modelo 1 posee el valor de *C-score* más cercano a 0 y por lo tanto el valor numérico más grande, interpretándolo como el modelo mejor predicho.

TM-score: 0.70 ± 0.12

- Lo que indica que su porcentaje de coincidencia con proteínas homólogas reportadas cristalográficamente es alto en un rango de (0.58 a 0.82).

5.1.2 Obtención de PTau

Al finalizar el proceso de *CHARMM-GUI* se obtuvo la estructura proteica de PTau dentro de una caja de agua, ver Figura 37 a), y los residuos Serina 516, 519, 579, 739, así como los de Treonina 522, 548 fueron fosforilados y se muestran como esferas moradas en la Figura 37 b).

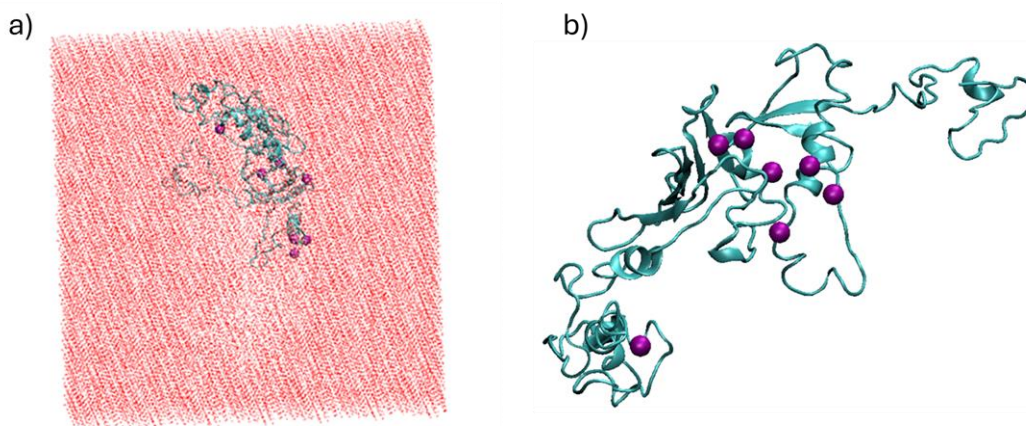


Figura 37. PTau a) dentro y b) fuera de la caja de agua, aminoácidos de Serina y Treonina en esferas de color morado.

El campo de fuerza implementado correspondió a “*CHARMM36m*” del nivel de teoría de Mecánica Molecular. Este campo de fuerza nos permitió obtener una mayor precisión en el proceso de fosforilación sin afectar la flexibilidad de la molécula y permitiendo el manejo estable de las hélices alfa presentes en Tau, obteniendo una PTau completa sin alteraciones en la cadena central.

5.1.3 Optimización de PTau

Para analizar el cambio estructural entre el Modelo 1 y PTau se realizaron los cálculos de Diagrama de Ramachandran para evaluar y visualizar los ángulos de torsión Φ (φ) y Ψ (ψ) de los enlaces peptídicos en la cadena principal, ver la Figura 38. De acuerdo a la literatura (Dixit *et al.*, 2021) la

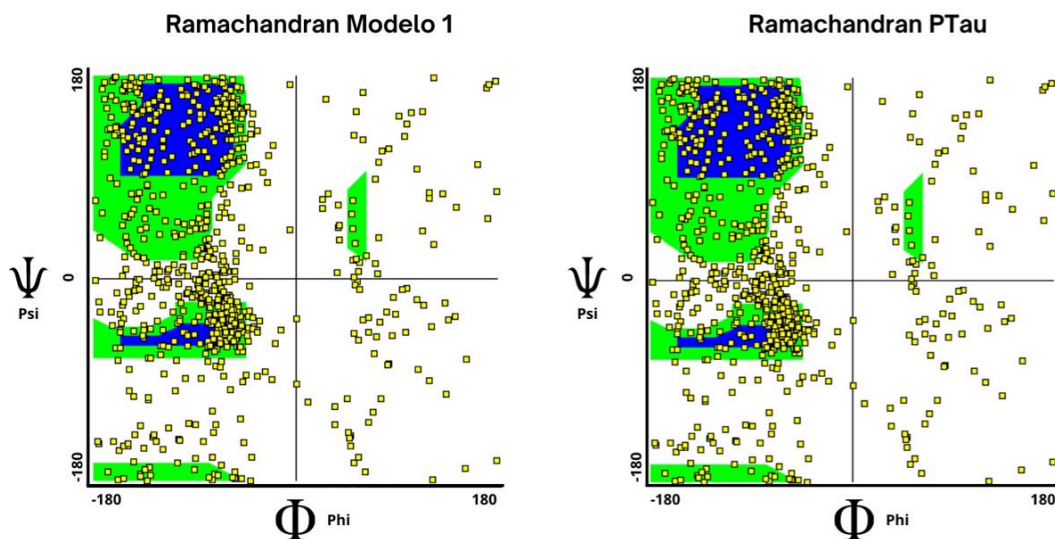


Figura 38. Comparación de Diagramas de Ramachandran.

estructura proteica de Tau fosforilada (PTau) no percibe alteraciones en su cadena principal en el proceso de fosforilación normal, por lo que se esperaba un Rammachandran idéntico al de Tau sin fosforilar “Modelo1”, lo que nos permitió visualizar que nuestra metodología de modificación fue eficiente y no causo alteraciones en la cadena central, en adición se percibió que los ángulos de torsión se encuentran en su mayoría en regiones comúnmente observadas en proteínas naturales (cuadrantes del centro hacia la izquierda) obteniendo una conformación estructural con la proporción de residuos fuera de los rangos viables correspondientes al 14% del total.

5.2 Dinámica Molecular para refinamiento de PTau mediante *Gromacs*

El tiempo de cálculo implementado para el refinamiento de PTau correspondió a los 80 ns, nuestro punto de partida se tomó como referencia un estudio similar en donde se estabilizó a PTau en el tiempo antes mencionado (Dixit *et al.*, 2021). No obstante, en nuestro estudio se apreció un equilibrio estructural de PTau a los 50 ns. Durante el proceso de DM la proteína se mantuvo como una sola cadena sin romper sus enlaces, se percibió una relajación estructural sin perder su forma intrínsecamente desordenada como lo reporta la literatura (Chang *et al.*, 2018).

5.2.1 Estructura de PTau final

Al finalizar los cálculos se realizó un conteo de los residuos presentes para asegurarnos de que no hubiera una delección o desnaturalización de la cadena durante el refinamiento, el conteo final arrojó un total de 758 aminoácidos, que corresponde al número inicial de residuos presentes en la cadena, asimismo se visualizó que durante el proceso de Dinámica la cadena no se desnaturalizara. Como se mencionó en el capítulo de antecedentes no existe reporte de una estructura proteica experimental completa de PTau por lo que en este apartado se adjunta información sobre los elementos de la estructura que presenta nuestro modelo final, ver Figura 39:

- Estructuras secundarias:

Hélices α

Laminas β (paralelas y antiparalelas)

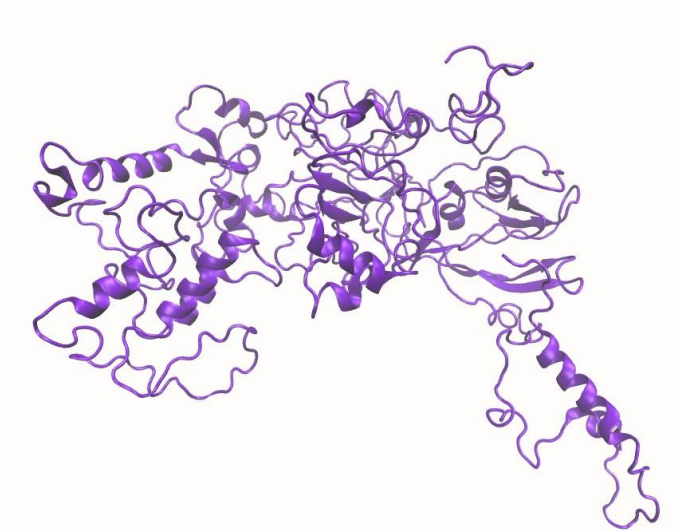


Figura 39. Estructura final de PTAu.

5.2.2 Mapa de Potencial Electroestático (MPE)

Se realizó la construcción del MPE mediante el programa *Pymol* (Schrödinger & DeLano., 2020), se implementó las cargas que contenía la proteína al momento de finalizar la DM, lo que nos permitió hacer una aproximación de la densidad electrónica presente en PTAu final. En la Figura 40, se pueden observar las regiones de mayor potencial electrostático (ataque nucleofílico) en color rojo y las de menor potencial electrostático (zonas de ataque electrofílico) se muestran de color azul, respectivamente. Las zonas blancas, grises y verdes indican un potencial electrostático

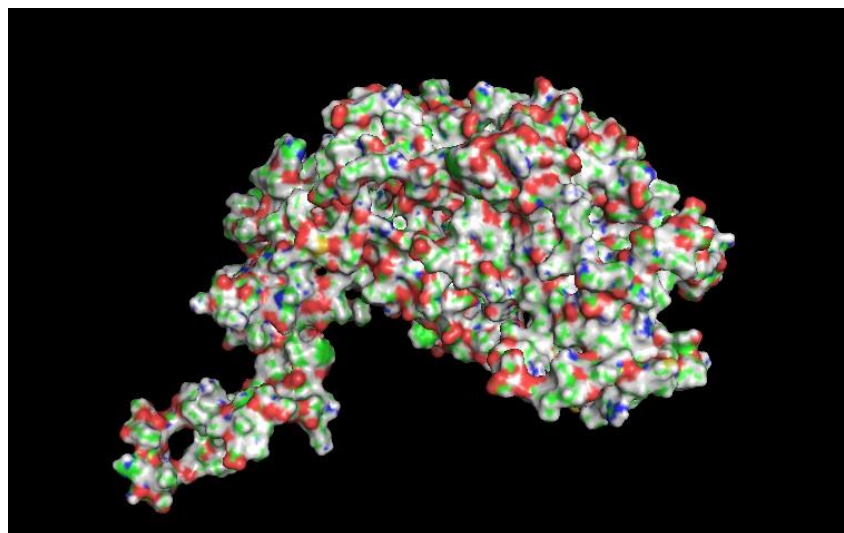


Figura 40. MPE de PTAu final.

intermedio. Al ser una estructura proteica se esperaba que los colores indicativos se distribuyeran de forma uniforme denotando que cada a.a presente tiene zonas con alta electronegatividad, lo que favorece la unión de otras moléculas a los residuos de PTau.

5.2.3 Raíz de la Desviación Cuadrática Media

En la Figura 39, se puede apreciar la gráfica de la Raíz de la Desviación Cuadrática Media por sus siglas en inglés (*RMSD*) que se construyó a manera de analizar el movimiento de PTau final respecto de PTau inicial, el tiempo de la DM del modelo fue de 80 ns, sin embargo, se obtuvo un equilibrio estructural a partir de los 50 ns, en donde la pendiente de la gráfica comienza a mantenerse dentro de un rango de valores en el “eje Y”, que no tienden a superar las 2 unidades lo que se interpreta como cambios leves a moderados en las posiciones de los residuos proteicos. Las variaciones del *RMSD* dentro de los primeros 45 ns se deben al movimiento de un grupo de aminoácidos que estiran sus enlaces dentro de la caja de agua.

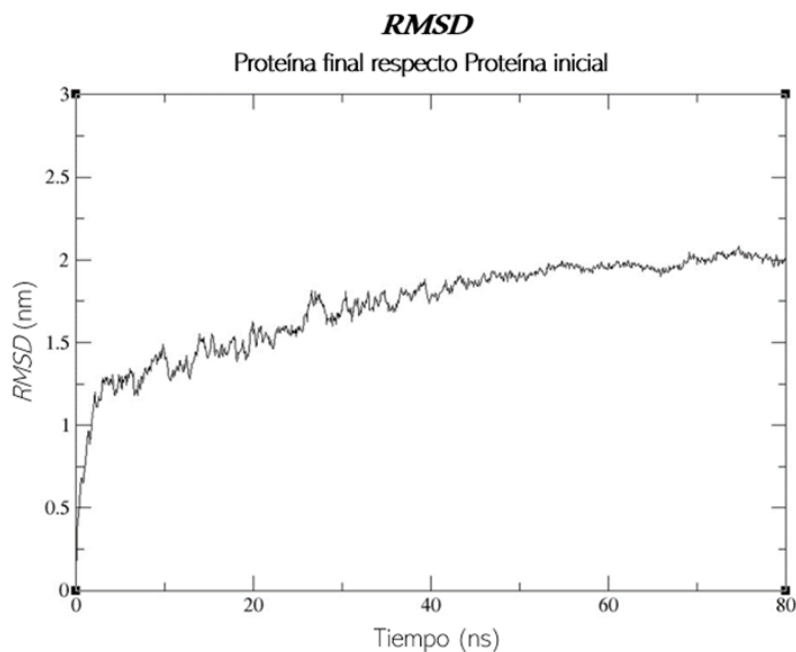


Figura 41. *RMSD* de PTau final respecto de PTau inicial.

5.2.4 Ramachandran de PTau final

Los cambios estructurales mostrados en la Figura 42 respecto a los Ramachandran de la sección

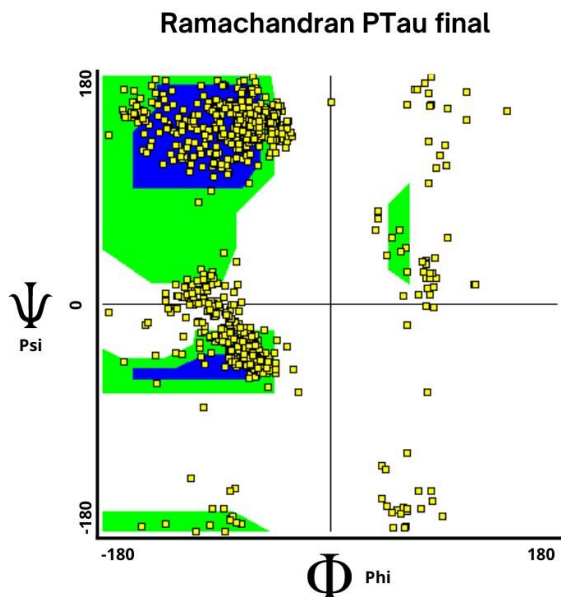


Figura 42. Ramachandran de PTau final.

“5.1.3 Optimización de PTau” nos permiten visualizar una mejoría de los ángulos de torsión colocando mayor cantidad de residuos en la zona favorable (cuadrantes izquierdos), el porcentaje de residuos fuera de la zona favorable corresponde al 4%, esto es un 10% de diferencia de lo obtenido en los diagramas anteriores, esta disminución del porcentaje se atribuye al proceso de refinamiento en donde se alcanzó el equilibrio estructural y por lo tanto las torsiones viables de los a.a de PTau.

5.3 Construcción de la Base de datos CHM

Aplicando criterios de evaluación relacionados con propuestas a tratamientos de la Demencia Frontotemporal se seleccionaron un total de 15 Compuestos de Herbolaria Mexicana.

El punto de selección para estos compuestos fue que inicialmente contaran con el dato de Peso Molecular (PM) y Formula química reportada, posteriormente se realizó la recopilación bibliográfica de las siguientes propiedades:

- Antinflamatoria general en humanos (Ant, por su sigla como identificador de Descriptor Molecular)
- La propiedad identificada como coadyuvante en alguna enfermedad neurodegenerativa priorizando su uso en alguna Tauopatía (TS, por su sigla como identificador de Descriptor Molecular)

Los valores obtenidos se reportan en la Tabla 10, en esta se etiquetaron aquellos compuestos que presentan las propiedades mencionadas con el valor de 1, y para aquellas moléculas que carecen de esta actividad se les colocó el número 0, evaluando presencia y ausencia correspondientemente. Los compuestos que fueron evaluados con 0 no se descartaron debido a que existen estudios en donde aún se sigue explorando si poseen mecanismos que ayuden en el tratamiento de ciertas Tauopatías.

Tabla 10. Ausencia o presencia de propiedades Antinflamatorias y relacionadas con el tratamiento de enfermedades del SNC.

<i>Compuesto</i>	<i>PM g/mol</i>	<i>Formula Molecular</i>	<i>Ant</i>	<i>TS</i>
Acetilcolina	146.21	$C_7H_{16}NO_2$	0	1
Ácido benzoico	346.50	$C_{22}H_{34}O_3$	0	1
Ácido cafeico	180.16	$C_9H_8O_4$	1	1
Ácido ferúlico	194.18	$C_{10}H_{10}O_4$	1	0
Ácido rosmarínico	360.30	$C_{18}H_{16}O_8$	0	1
Ácido sinápico	224.21	$C_{11}H_{12}O_5$	1	1
Atropina	694.80	$C_{34}H_{50}N_2O_1$	0	0
Convalatoxina	552.60	$C_{28}H_{40}O_{11}$	0	0
Cumarina	146.14	$C_9H_6O_2$	1	0
Curcumina	368.40	$C_{21}H_{20}O_6$	1	1
Especiofilina	368.40	$C_{21}H_{24}N_2O_4$	1	1
Eucaliptol	154.25	$C_{10}H_{18}O$	1	1

Plenolina	264.32	$C_{15}H_{20}O_4$	1	0
Santonina	246.30	$C_{15}H_{18}O_3$	0	0
Tujona	152.23	$C_{10}H_{16}O$	0	1

En la Tabla 10, se puede observar como el promedio del Peso Molecular de los CHM es de 314.21 g/mol. No obstante, el valor máximo reportado corresponde al compuesto Atropina con 694.80 g/mol, que es más del doble del promedio por lo que se estima que este compuesto podría quedar fuera de los seleccionados para AM. Dentro de el mismo Descriptor “PM” existen dos CHM con el mismo valor: Curcumina y Especiofilina con un peso de 368.40 g/mol, pero ambos difieren en el número de átomos y composición química por lo que no son conformeros y no están relacionados entre sí. Todas las moléculas presentes en la Base CHM se encuentran dentro de los compuestos orgánicos teniendo como elementos principales en su composición al carbono, al hidrógeno y al oxígeno, y para este caso, los compuestos Acetilcolina, Atropina y Especiofilina cuentan con átomos de Nitrógeno que podrían favorecer la interacción específica con PTau a través de la formación de puentes de hidrógeno y otros tipos de interacciones no covalentes.

5.3.1 Compuestos optimizados de la Base de datos de CHM

Para el cálculo de los Descriptores Moleculares fue necesario someter a los compuestos de la base de datos de CHM a una optimización a través del programa *Avogadro* (Marcus *et al.*, 2012) con el fin de obtener las geometrías ideales para su descripción, ver Figura 43. El campo de fuerza implementado en estos cálculos correspondió a el “*MMFF94*” por sus siglas en inglés (*Merck Molecular Force Field 94*) que pertenece al nivel de teoría de Mecánica Molecular, y es aplicado para predecir la estructura y propiedades de moléculas orgánicas con interés farmacéutico (Halgren., 1996).

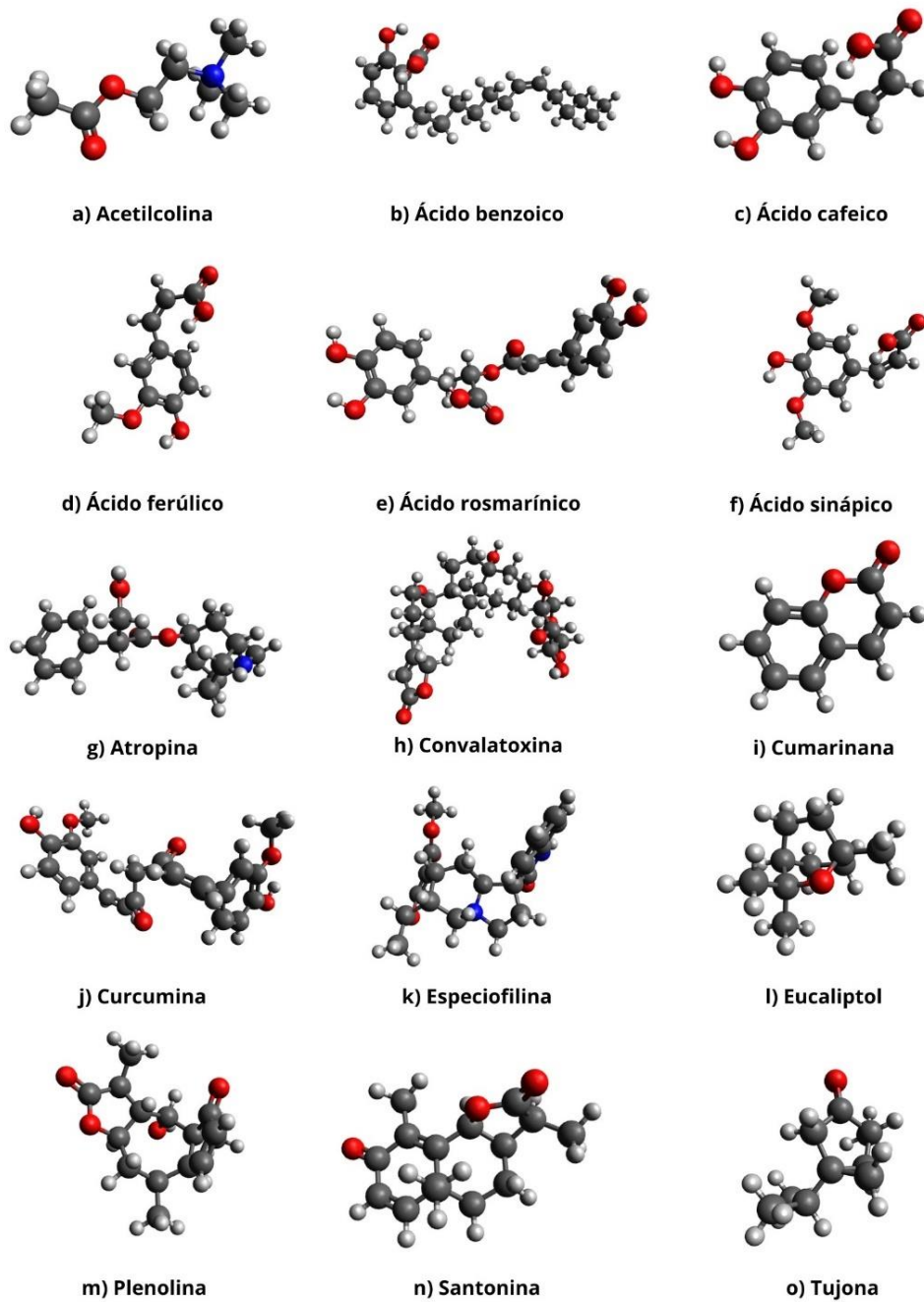


Figura 43. Geometrías optimizadas de la base de datos de CHM.

5.3.2 Obtención de Descriptores Moleculares de la base de datos de CHM

Los Descriptores Moleculares TPSA, IQ, logP, nON y nOHNH fueron calculados para obtener una aproximación de sus interacciones como fármacos. En la Tabla 11, se encuentran los valores obtenidos de los cálculos.

Tabla 11. Descriptores Moleculares calculados para la base de datos de CHM.

<i>Compuesto</i>	<i>TPSA</i>	<i>IQ</i>	<i>logP</i>	<i>nON</i>	<i>nOHNH</i>
Acetilcolina	26.30	-1.54	-3.56	3	0
Ácido benzoico	57.53	-0.07	7.33	3	2
Ácido cafeico	77.75	-0.81	0.94	4	3
Ácido ferúlico	66.76	-0.72	1.25	4	2
Ácido rosmarínico	144.52	-0.18	1.63	8	5
Ácido sinápico	76.00	-0.47	1.26	5	2
Atropina	49.77	-0.09	1.77	4	1
Convalatoxina	183.21	-0.39	0.11	11	6
Cumarina	30.21	-1.57	2.01	2	0
Curcumina	93.07	-0.26	2.30	6	2
Especiofilina	67.88	-0.34	2.18	6	1
Eucaliptol	9.23	-1.6	2.72	1	0
Plenolina	63.60	-1.09	0.51	4	1
Santonina	43.38	-1.06	1.31	3	0
Tujona	17.07	-1.63	2.16	1	0

Con la Tabla anterior se estimaron los valores promedios siguientes para cada columna:

- TPSA

Promedio: 71.87, no tiene moda, valor máximo:183.21 y valor mínimo :9.23

- IQ

Promedio: -11.82, no tiene moda, valor máximo: -1.63 y valor mínimo: -0.07

- LogP

Promedio: 1.70, no tiene moda, valor máximo :7.33 y valor mínimo: -3.56

- nON

Promedio: 4.64, moda: 4, valor máximo:11 y valor mínimo: 1

- nOHNH

Promedio: 1.78, moda:0, valor máximo:6 y valor mínimo:0

Con los siguientes parámetros se puede estimar que los Descriptores Moleculares nON y nOHNH al tener más de un dato repetido “moda” podrían quedar fuera de la selección de importancias del ARBA al no generar alto impacto en su delección, sin embargo, Descriptores Moleculares como TPSA, IQ y logP podrían ser un preámbulo en la decisión de selección de los *hits* al no tener valores repetidos. El CHM “Acetilcolina” destaca por tener un valor de logP negativo lo que quiere decir que la solubilidad en agua que posee es alta y por lo tanto su capacidad para atravesar membranas celulares es baja, lo que compromete su absorción y distribución en el cuerpo humano. Como se mencionó en el capítulo de antecedentes una propuesta en la inhibición de la hiperfosforilación de Tau es el derivado de quinolina “THQ4S” (Maccioni *et al.*, 2011), que reporta los siguientes valores:

Tabla 12. Descriptores Moleculares de THQ4S.

<i>Compuesto</i>	TPSA	IQ	logP	nON	nOHNH
THQ4S	12.82	-1	2.66	1	0

Existen parámetros aproximados al compuesto THQ4S dentro de la base CHM. Sin embargo, no existe dentro de nuestra base un compuesto que asimile todos los Descriptores de la misma forma que THQ4S. Con el objetivo de implementar un algoritmo de *ML* a nuestra base de CHM se aplicaron los criterios ADME y la regla de 5 de *Lipinski* para obtener una base normalizada en parámetros de cumplimiento o infracción de los puntos antes mencionados. En el caso del Descriptor

TPSA no se encontró ningún dato que reportará un rango de valores para su evaluación de impacto en nuevos fármacos, es por esto por lo que se decidió tomarlo como punto de referencia al momento de aplicar el ARBA. Para los Descriptores nON y nOHNH se obtuvo un 92.85% de cumplimiento en las reglas, teniendo como único infractor al compuesto Convalatoxina, se prevé que ambos descriptores podrían tener poco o nulo valor en el proceso de selección de los *hits*. Descriptores Moleculares con mayor varianza en los valores como Ant y TS podrían ser los que obtengan un mayor peso al momento de la selección de compuestos para AM.

Tabla 13. Normalización de la base de datos de los CHM, los números en rojo indican la única variación de los descriptores nON y nOHNH.

<i>Compuesto</i>	<i>PM g/mol</i>	<i>Formula</i>	<i>Ant</i>	<i>TS</i>	<i>TPSA</i>	<i>IQ</i>	<i>logP</i>	<i>nON</i>	<i>nOHNH</i>
Acetilcolina	1	$C_7H_{16}NO_2$	0	1	26.30	0	1	1	1
Ácido benzoico	1	$C_{22}H_{34}O_3$	0	1	57.53	1	0	1	1
Ácido cafeico	1	$C_9H_8O_4$	1	1	77.75	0	1	1	1
Ácido ferúlico	1	$C_{10}H_{10}O_4$	1	0	66.76	0	1	1	1
Ácido rosmarínico	1	$C_{18}H_{16}O_8$	0	1	144.52	1	1	1	1
Ácido sinápico	1	$C_{11}H_{12}O_5$	1	1	76.00	1	1	1	1
Atropina	0	$C_{34}H_{50}N_2O_1$	0	0	49.77	1	1	1	1
Convalatoxina	0	$C_{28}H_{40}O_{11}$	0	0	162.98	1	1	0	0
Cumarina	1	$C_9H_6O_2$	1	0	30.21	0	1	1	1
Curcumina	1	$C_{21}H_{20}O_6$	1	1	93.07	1	1	1	1
Especiofilina	1	$C_{21}H_{24}N_2O_4$	1	1	67.88	1	1	1	1
Eucaliptol	1	$C_{10}H_{18}O$	1	1	9.23	0	1	1	1
Plenolina	1	$C_{15}H_{20}O_4$	1	0	63.60	0	1	1	1

Santonina	1	$C_{15}H_{18}O_3$	0	0	43.38	0	1	1	1
Tujona	1	$C_{10}H_{16}O$	0	1	17.07	0	1	1	1

5.4 Implementación de *Machine Learning* (ML)

La aplicación de *ML* a la base de datos de CHM nos permitió realizar una metodología para obtener los compuestos que serían probados en Acoplamiento Molecular. El proceso de selección del algoritmo se fundamentó principalmente en los siguientes criterios:

- Debería ser un algoritmo que tuviera la capacidad de ser implementado a un volumen grande de datos, aunque en este punto nuestra base es pequeña se pretende que en un futuro sea aplicado a bases de datos más grandes para obtener un filtrado de otros posibles candidatos a fármacos inhibidores de la hiperfosforilación de P τ au.
- El mecanismo de selección tendría que estar basado en las importancias de cada descriptor, asignándoles un valor numérico para facilitar el proceso de obtención de los *hits*.
- El Algoritmo debería tener alta precisión para reducir el riesgo de un sobreajuste o una deficiencia en su ejecución.

5.4.1 Selección e implementación del Algoritmo de Regresión de Bosque Aleatorio (ARBA)

Tomando en cuenta los puntos mencionados anteriormente la selección del modelo nos condujo al grupo de algoritmos de Aprendizaje Supervisado y en específico al ARBA (Algoritmo de Regresión de Bosque Aleatorio), que nos proporcionó la capacidad para manejar bases de datos robustas sin perder la precisión y asignando importancias de selección a cada Descriptor (Lozada *et al.*, 2022).

Para comenzar la programación del algoritmo se descartó el Descriptor “TPSA” al cual no se le dio un peso en la toma de decisión de las importancias debido a que no se encontró una manera de normalizar sus valores y ponerlos en criterios de cumplimiento o infracción a alguna regla relacionada con su farmacodinámica, además de que el rango que muestra es grande, teniendo datos de 9.23 como mínimo y 162.98 como máximo, esto indica que el algoritmo es susceptible a subestimación. En la etapa de prueba del algoritmo se realizó la partición del total de los datos de la base CHM en forma aleatoria (15 con 8 descriptores) para tomar el 25% (4 compuestos) para este paso y dejando el 75% (11 compuestos) para ejecutar la prueba del ARBA, el número total de árboles implementados correspondió a 1000 regresiones.

5.4.2 Validación del modelo

La validación es un paso esencial en la implementación de algoritmos de IA, garantizando que sean precisos, fiables y capaces de proporcionar su aplicación en problemas del mundo real (Lozada *et al.*, 2021). Para conocer si la aplicación del ARBA a la base de datos de los CHM se realizó de manera óptima, se analizó las diferencias entre el “Error absoluto medio” estimado en la etapa de entrenamiento (e_2) y el “Error absoluto medio” de la prueba con respecto al PM (e_1), dicho parámetro nos permite obtener la diferencia en valor absoluto de la predicción y los valores reales. A pesar de que el Descriptor “TPSA” se descartó en la programación fue aplicado para el cálculo del Error absoluto medio, esto debido a que es la variable que posee la mayor varianza, Apéndice F.

$$e_1 = 42$$

$$e_2 = 25.87$$

La diferencia de valor en error absoluto entre el entrenamiento y la prueba fue de 16.13, lo que corrobora que hubo una mejora en el proceso de ponderación de las importancias, además muestra

que el algoritmo no está sobre ajustado ni sesgado al percibir una mejora entre los dos errores medios.

5.4.2 Importancias obtenidas

Como resultado final de la aplicación del ARBA se obtuvieron las importancias mostradas en la Figura 44. El Descriptor Molecular “Ant” obtuvo la menor varianza del error total promediado de cada uno de los árboles en que se implementó como selector, su valor de ponderación es el más alto con 0.3, seguido de PM con 0.24, IQ con 0.21, TS y logP con 0.12, teniendo un total de 0.99. Los Descriptores Moleculares nON y nOHNH se ponderaron con un valor de 0, como se había predicho en el apartado de “Obtención de Descriptores Moleculares” la repetición de los datos

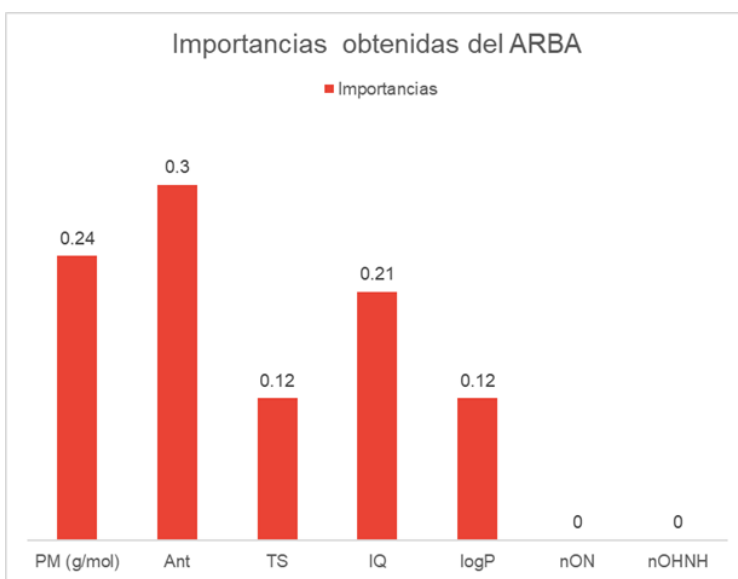


Figura 44. Resultado de las Ponderaciones de las importancias del ARBA sobre la base de datos de los CHM.

entre compuestos impedía que fuera un parámetro de decisión. Se estimaba que las importancias sumarian 1, sin embargo, se obtuvo un error de 0.01 atribuido al redondeo de las ponderaciones de los últimos Descriptores.

5.4.3 Obtención de *Hits* y *Fails*

Con la metodología menciona en la sección “Interpretación de las importancias dadas por ARBA” se obtuvo el siguiente criterio:

Para $\hat{\mu} = 0.70$:

$$\begin{cases} P_k \text{ compuesto} > \hat{\mu} \Rightarrow \text{Hit} \\ P_k \text{ compuesto} < \hat{\mu} \Rightarrow \text{Fail} \end{cases}$$

En la Tabla 14, se muestran los compuestos “*hits*” en filas color azul y los compuestos que no superaron la media poblacional ($\hat{\mu}$) denotados como “*fails*” en filas de color blanco. Se obtuvo un total de 5 *hits* y 10 *fails*. Los CHM que obtuvieron el total del puntaje permitido por el ARBA (0.99) son los siguientes: Ácido sinápico, Curcumina y Especiofilina. El puntaje obtenido por los *hits* restantes fue de 0.78 en donde clasificaron el Ácido cafeico y el Eucaliptol.

Tabla 14. Selección de compuestos *hits* y *fails*.

Compuesto	A_1	Formula	A_2	A_3	TPSA	A_4	A_5	A_6	A_7	P_k	Mayor a .70
Acetilcolina	0.24	$C_7H_{16}NO_2$	0.00	0.12	26.30	0.00	0.12	0.00	0.00	0.48	Fail
Ácido benzoico	0.24	$C_{22}H_{34}O_3$	0.00	0.12	57.53	0.21	0.00	0.00	0.00	0.57	Fail
Ácido cafeico	0.24	$C_9H_8O_4$	0.30	0.12	77.75	0.00	0.12	0.00	0.00	0.78	Hit
Ácido ferúlico	0.24	$C_{10}H_{10}O_4$	0.30	0.00	66.76	0.00	0.12	0.00	0.00	0.66	Fail
Ácido rosmarínico	0.24	$C_{18}H_{16}O_8$	0.00	0.12	144.52	0.21	0.12	0.00	0.00	0.69	Fail
Ácido sinápico	0.24	$C_{11}H_{12}O_5$	0.30	0.12	76.00	0.21	0.12	0.00	0.00	0.99	Hit
Atropina	0.00	$C_{34}H_{50}N_2O_1$	0.00	0.00	49.77	0.21	0.12	0.00	0.00	0.33	Fail
Convalatoxina	0.00	$C_{28}H_{40}O_{11}$	0.00	0.00	162.98	0.21	0.12	0.00	0.00	0.33	Fail
Cumarina	0.24	$C_9H_6O_2$	0.30	0.00	30.21	0.00	0.12	0.00	0.00	0.66	Fail
Curcumina	0.24	$C_{21}H_{20}O_6$	0.30	0.12	93.07	0.21	0.12	0.00	0.00	0.99	Hit

Especiofilina	0.24	$C_{21}H_{24}N_2O_4$	0.30	0.12	67.88	0.21	0.12	0.00	0.00	0.99	<i>Hit</i>
Eucaliptol	0.24	$C_{10}H_{18}O$	0.30	0.12	9.23	0.00	0.12	0.00	0.00	0.78	<i>Hit</i>
Plenolina	0.24	$C_{15}H_{20}O_4$	0.30	0.00	63.60	0.00	0.12	0.00	0.00	0.66	<i>Fail</i>
Santonina	0.24	$C_{15}H_{18}O_3$	0.00	0.00	43.38	0.00	0.12	0.00	0.00	0.36	<i>Fail</i>
Tujona	0.24	$C_{10}H_{16}O$	0.00	0.12	17.07	0.00	0.12	0.00	0.00	0.48	<i>Fail</i>

5.5 Acoplamiento Molecular con los *Hits* y PTau

Los estudios de AM se realizaron mediante la metodología rígida. Como resultado se obtuvieron 12 conformeros, 2 conformaciones de referencia al grupo Fosfato y 10 estructuras de tipo PTau-*hits*. Se sometió a PTau en AM contra el grupo Fosfato para medir la energía de enlace presente en los sitios aberrantes de hiperfosforilación, para su estudio, se analizaron 2 tipos de interacciones:

- Puente de Hidrógeno
- *Van der Waals*

En la Tabla 15 se resume la información correspondiente a las uniones entre los sitios T248, S411 y el grupo Fosfato.

Tabla 15. Resultados de AM T248-Fosfato y S411-Fosfato.

Sitio	Tipo de interacción	
	Puente de Hidrógeno	Interacción <i>Van der Waals</i>
T248	- 0.240 <i>kcal/mol</i>	-0.58 <i>kcal/mol</i>
S411	+3.008 <i>kcal/mol</i>	+4.32 <i>kcal/mol</i>

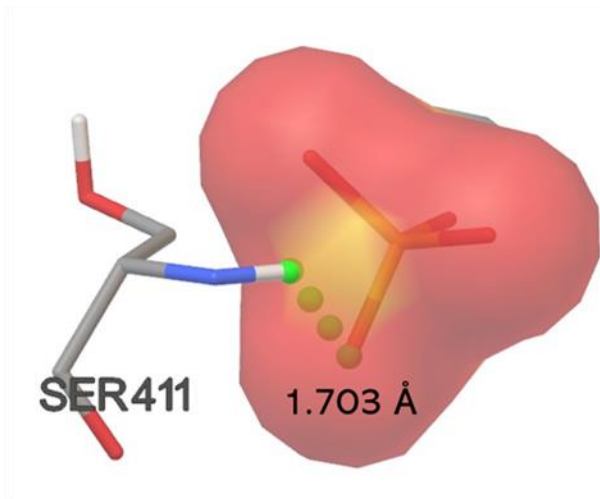


Figura 45. AM de S411-Fosfato.

El AM S411-Fosfato, en la Figura 45, presentó una distancia de interacción por Puente de Hidrógeno de 1.703 Å y una energía de unión de -0.240 kcal/mol , se teoriza que este comportamiento se da en la naturaleza del cerebro humano. Sin embargo, no existen cristalográficas reportadas de esta Modificación Post-traducciona. La energía del Puente de Hidrógeno formado por el confórmero se encuentra

cercana a 0 y por lo tanto se considera como una interacción débil (Desiraju., 1996).

El AM T248-Fosfato, ver Figura 46, dio como resultado un Puente de Hidrógeno con una distancia de interacción de 1.88 Å y una energía de $+3.008 \text{ kcal/mol}$, como se observa tenemos la presencia de una energía positiva, lo que indica que la unión del confórmero no es la óptima y por lo tanto no es termodinámicamente favorable. Sin embargo, al igual que el caso anterior se estima que esta conformación existe en la naturaleza.

Para ambos casos los AM fueron realizados de forma rígida y no se descartó que los AM flexibles puedan tener mayor precisión biológica y por lo tanto permitir conformaciones estables.

En ambos sitios seleccionados T248 y S411 se realizaron estudios de AM rígidos, con el fin de observar si poseen o no la capacidad de inhibir la hiperfosforilación de P_Tau.

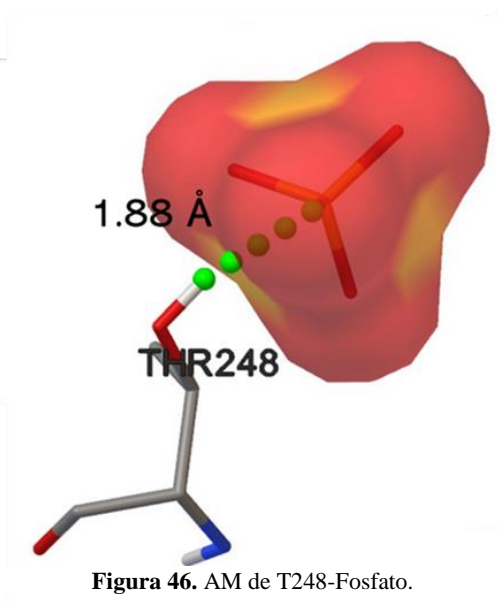


Figura 46. AM de T248-Fosfato.

5.5.1 Hits vs Sitio T248

Las conformaciones resultantes de los AM con el sitio T248 aparecen en la Figura 47, en donde se obtuvieron interacciones de *Van der Waals* y Puentes de Hidrógeno, las distancias entre enlaces se muestran con números negros al costado de las esferas, para este caso se encontraron energías de enlace positivas y negativas como se muestra a continuación:

$$E_{de\ enlace\ para\ \acute{A}cido\ cafeico} = 0.83\ kcal/mol$$

$$E_{de\ enlace\ para\ \acute{A}cido\ sinapico} = -3.03\ kcal/mol$$

$$E_{de\ enlace\ para\ Curcumina} = -3.82\ kcal/mol$$

$$E_{de\ enlace\ para\ Especiofilina} = 26.28\ kcal/mol$$

$$E_{de\ enlace\ para\ Eucaliptol} = -2.44\ kcal/mol$$

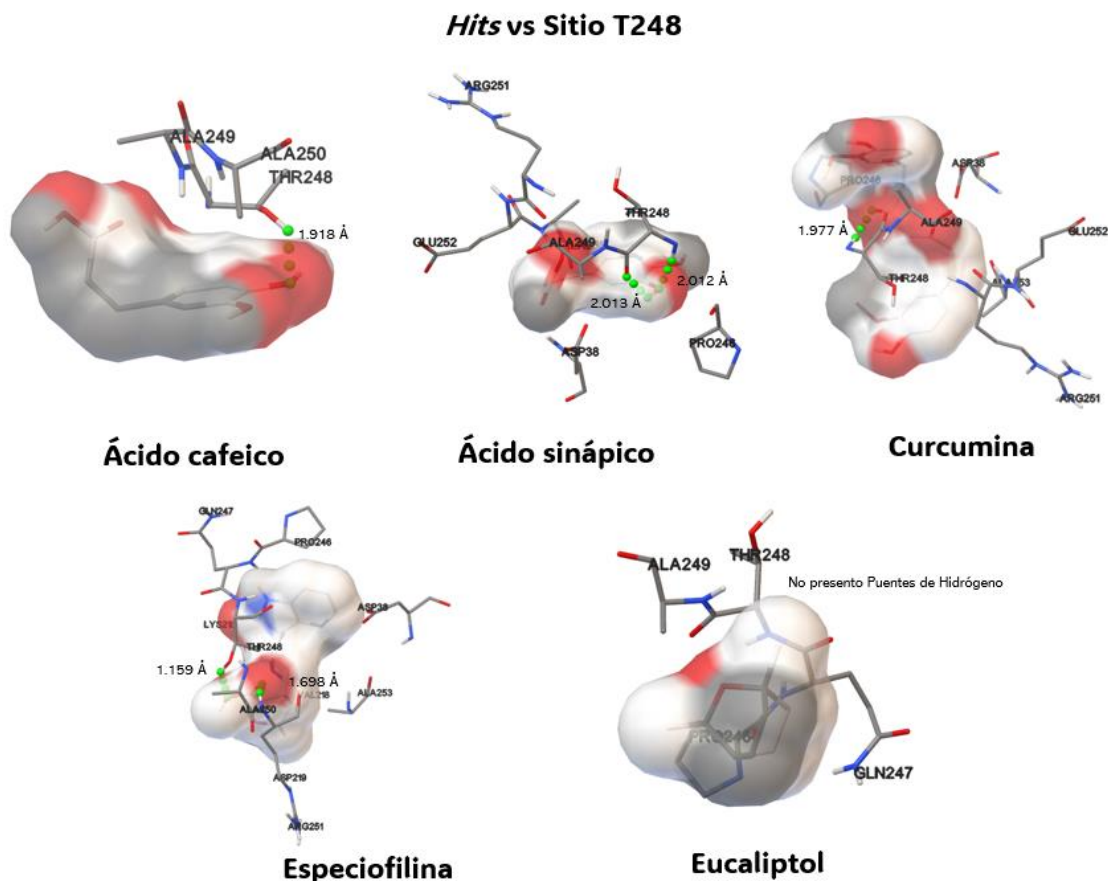


Figura 47. AM hits vs T248.

Exceptuando a la Especiofilina, identificada como *hit* molecular, nuestros compuestos obtuvieron uniones favorables respecto a las presentadas con el fosfato en el sitio T248. Los mejores candidatos para inhibir la hiperfosforilación de P τ en este blanco corresponden a Curcumina y Ácido sinápico presentando las energías con más afinidad electrónica y por lo tanto las más negativas. En el caso del Ácido sinápico percibimos una doble interacción de Puente de Hidrógeno con el residuo T248 y clasifica como la segunda interacción mejor calificada. El promedio de las distancias de interacción es de 1.796 Å, no obstante, para el *hit* Eucaliptol no hubo presencia de interacción de Puente de Hidrógeno. Los AM arrojaron en un caso otras interacciones de PH con otros residuos, tal es el caso de Especiofilina con ALA250 obteniendo una distancia de enlace de 1.698 Å.

5.5.2 Hits vs Sitio S411

En la Figura 48 se encuentran los AM correspondientes al sitio S411, en todos los casos se obtuvieron interacciones de *Van der Waals* y Puentes de Hidrógeno, las distancias entre enlaces se muestran con números negros al costado y con esferas verdes, las energías de enlace calculadas son las siguientes:

$$E_{de\ enlace\ para\ \acute{a}cido\ cafeico} = -5.1\ kcal/mol$$

$$E_{de\ enlace\ para\ \acute{a}cido\ sinapico} = -4.6\ kcal/mol$$

$$E_{de\ enlace\ para\ Curcumina} = -6.24\ kcal/mol$$

$$E_{de\ enlace\ para\ Especiofilina} = -5.34\ kcal/mol$$

$$E_{de\ enlace\ para\ Eucaliptol} = -2.74\ kcal/mol$$

Para estos resultados todas las energías fueron negativas lo que indica una unión favorable contra el sitio S411. En comparación con la interacción que tuvo el grupo fosfato, los *hits* se ven

favorecidos en las energías de unión, brindando así la posibilidad de inhibir el proceso de hiperfosforilación, el compuesto mejor puntuado fue Curcumina al poseer la interacción más fuerte con P τ au. En el caso del Ácido cafeico percibimos una doble interacción de Puente de Hidrógeno con el residuo S411 y clasifica como la tercera interacción mejor puntuada, lo que sugiere, que, al aumentar el número de interacciones, presenta mayor fuerza y estabilidad molecular. Dentro de la Figura 48, se pueden observar las distancias obtenidas de las interacciones, el promedio calculado de las distancias es de 1.931 Å para S411, no obstante, los AM arrojaron en algunos casos interacciones de PH con otros residuos y en la literatura no se tiene evidencia de que la unión con estos residuos afecte la estructura proteica de P τ au (Dixit *et al.*, 2021):

- Ácido cafeico con LYS375.
- Ácido sinápico y Especiofilina con LYS413, aunque en este último caso la interacción fue débil.
- Curcumina con CYS409.

Hits vs Sitio S411

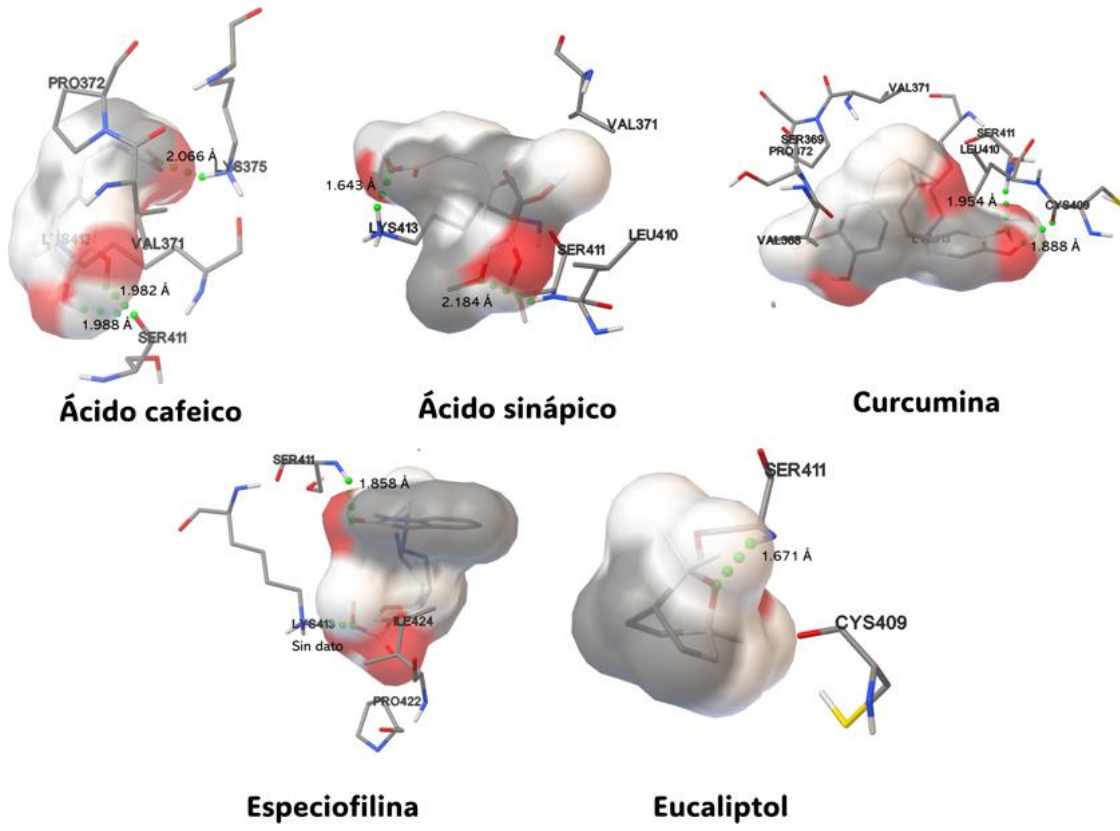


Figura 48. AM hits vs S411.

Capítulo 6

Conclusión y

perspectivas

Los compuestos de Herbolaria Mexicana presentaron una mejor actividad *in silico* frente al grupo fosfato en los sitios de hiperfosforilación de Tau, destacando al CHM “curcumina” que puede ser considerado como una alternativa de investigación experimental para evitar la neurodegeneración causada en la Demencia Frontotemporal.

Con esta investigación se generó un modelo proteico completo de Tau Fosforilada, que nos permitió realizar estudios entorno a la inhibición de la hiperfosforilación de P_{Tau}, se implementó la metodología “Construcción por Modelado Comparativo”, obteniendo la estructura “P_{Tau} Final” en un tiempo de 50 ns, pero completando la corrida a los 80 ns, lo que representa una mejora respecto de experimentos similares en donde se aplicó una metodología de construcción distinta, además la cadena principal de P_{Tau} no sufrió modificaciones en su secuencia, lo que permitió el análisis completo de la predicción de la densidad electrónica y la presencia del 96% de residuos en zonas favorables de predicción estructural dentro del gráfico Ramachandran.

En la búsqueda de compuestos capaces de competir con el grupo fosfato en sitios aberrantes de fosforilación, se construyó la base de datos de CHM que fue integrada por 15 Compuestos de Herbolaria Mexicana con posibles aplicaciones Antinflamatorias (Ant) y el TS como coadyuvantes en el tratamiento de enfermedades relacionadas con el Sistema Nervioso Central. Los Descriptores Moleculares PM, TPSA, logP, nON y nOHNH permitieron ponderar las infracciones o cumplimientos de parámetros ideales para fármacos pertenecientes a la Regla de 5 de *Lipinski* y las propiedades ADME, ofreciendo un acercamiento a sus interacciones dentro del cuerpo humano. Con la perspectiva de mejorar la selección de los compuestos con mejores propiedades farmacodinámicas, se implementó el Algoritmo de Regresión de Bosque Aleatorio (ARBA), el cual predijo que los descriptores con mayor impacto en la inhibición de la formación de HPTau son Ant, PM

y IQ, formando así un criterio de decisión que nos condujo a la selección de los “5 *hits*”, dichos compuestos fueron probados en Acoplamiento Molecular dirigidos a 2 sitios (T248 y S411). Los conformeros obtenidos de las regiones mostraron en su mayoría una actividad favorable de unión frente al grupo Fosfato, el mayor éxito de ambos grupos de estudio es el Compuesto de Herbolaria Mexicana “Curcumina” que presentó una energía de enlace superior en ambos casos. Finalmente, las perspectivas futuras para el proyecto es complementar el modelo teórico con la construcción de un microtúbulo rodeado de PTau para aproximar el comportamiento de los *hits* en un axón neuronal con la finalidad de acortar el tiempo de prueba en un proyecto experimental que permitirá estar más cerca de un tratamiento de la Demencia Frontotemporal y otras Tauopatías.

Bibliografía

Abraham, M.J., Murtola, T., Schulz, R., Páll, S., Smith, J.C., Hess, B., *et al.* (2015). GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1(2), 19-25.

Adler, S., Basketter, D., Creton, S., Pelkonen, O., van Benthem, J., Zuang, V., Andersen, K. E., Angers-Loustau, A., Aptula, A., Bal-Price, A., Benfenati, E., Bernauer, U., Bessems, J., Bois, F. Y., Boobis, A., Brandon, E., Bremer, S., Broschard, T., Casati, S., Coecke, S., Zaldivar, J. M. (2011). Alternative (non-animal) methods for cosmetics testing: current status and future prospects-2010. *Archives of toxicology*, 85(5), 367–485. <https://doi.org/10.1007/s00204-011-0693-2>.

Alvarez G. (1973). Parálisis supranuclear progresiva [Progressive supranuclear paralysis]. *Revista médica de Chile*, 101(9), 718–719.

Ansari, F., Pourjafar, H., Tabrizi, A., & Homayouni, A. (2020). The Effects of Probiotics and Prebiotics on Mental Disorders: A Review on Depression, Anxiety, Alzheimer, and Autism Spectrum Disorders. *Current pharmaceutical biotechnology*, 21(7), 555–565. <https://doi.org/10.2174/1389201021666200107113812>.

Armstrong, R. A. (2016). The molecular biology of tauopathies. *Journal of Neurogenetics*, 30(1-2), 1-17. <https://doi.org/10.1080/01677063.2016.1201213>

Atkins, P., & de Paula, J. (2018). *Physical Chemistry*. Oxford University Press.

Ávila, J., Mayer, M., & Quesada, V. (2021). La inteligencia artificial y sus aplicaciones en medicina II: importancia actual y aplicaciones prácticas [Artificial intelligence and its applications in medicine II: Current importance and practical applications]. *Atencion primaria*, 53(1), 81–88. <https://doi.org/10.1016/j.aprim.2020.04.014>.

Batiha, G., Magdy, A., Wasef, L., Elewa, Y., Abd El-Hack, M., Taha, A., Al-Sagheer, A., Devkota, H., Tufarelli, V. (2015). A Review on Chemical Constituents and Biological Activities. *Appl. Sci.* 2020, 10, 2668. <https://doi.org/10.3390/app10082668>.

Benjamin, C.C., Ryohei, W & Edward, B.L. (2024). Tauopatías de enfermedades neurodegenerativa. *Revisión Anual de Patología: Mecanismos de enfermedad*, 19, 345-370. <https://doi.org/10.1146/annurev-pathmechdis-051222-120750>.

Bielenberg J. (2007). Zentralnervöse Effekte durch Thujon [Thujone]. *Medizinische Monatsschrift für Pharmazeuten*, 30(9), 322–326.

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer. Boden, M. A. (2017). *Inteligencia artificial*. Turner.

Boyd, D. B., & Lipkowitz, K. B. (1982). Molecular mechanics: The method and its underlying philosophy. *Journal of Chemical Education*, 59(4), 269–274. <https://doi.org/10.1021/ed059p269>.

Brooks, R., Brooks, L., Mackerell, A., Nilsson, L., Petrella, R., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., *et al.* (2009). CHARMM: the biomolecular simulation program. *Journal of computational chemistry*, 30(10), 1545–1614. <https://doi.org/10.1002/jcc.21287>.

Burden, F. R. (2014). *Molecular structure and properties: From quantum chemical calculations to phenomenological relationships*. Elsevier.

Cano, A., Posada, L., & Amariles, P. (2017). Toxicidad hepática causada por medicamentos: revisión estructurada. *Revista Colombiana de Gastroenterología*, 32(4), 337–348. <https://doi.org/10.22516/25007440.177>.

Carmona, C.J., Pulgar-Rubio, F.J., García-Vico, A.M., González, P., & Jesus, M.J. (2015). Análisis descriptivo mediante aprendizaje supervisado basado en patrones emergentes. *Proceedings of the VII Simposio Teoría y Aplicaciones de Minería de Datos*, 685–694.

Cazzola, M., & Matera, M. G. (2020). The effect of doxofylline in asthma and COPD. *Respiratory medicine*, 164, 105904. <https://doi.org/10.1016/j.rmed.2020.105904>.

Ceremuga, T., McClellan, C., Green, X., Heber, B., Jolly, M., Malone, T., Schaaf, J., & Isaacs, A. (2017). Investigation of the Anxiolytic and Antidepressant Effects of Eucalyptol (1,8-Cineole), a Compound From Eucalyptus, in the Adult Male Sprague-Dawley Rat. *AANA journal*, 85(4), 277–284.

Cervantes-Aragón, I., Ramírez-García, S. A., Baltazar-Rodríguez, L. M., García-Cruz, D., & Castañeda-Cisneros, G. (2019). Aproximación genética en la esclerosis lateral amiotrófica. *Gaceta médica de México*, 155(5), 513–521. <https://doi.org/10.24875/GMM.19004927>.

Chen C. (2016). Sinaptic Acid and Its Derivatives as Medicine in Oxidative Stress-Induced Diseases and Aging. *Oxidative medicine and cellular longevity*, 2016, 3571614. <https://doi.org/10.1155/2016/3571614>.

Cruz, M. (2016). Variedades de toronja para suelos arcillosos limosos de la Huasteca Potosina. México. *Revista Mexicana de Ciencias Farmacéuticas* 47(2), 22-35. <http://biblioteca.inifap.gob.mx:8080/xmlui/bitstream/handle/123456789/737/176.pdf?sequenc>.

Da Rocha, M., Viegas, F., Campos, H., Nicastro, P., Fossaluzza, P., Fraga, C., Barreiro, E., Viegas, C., (2011). The role of natural products in the discovery of new drug candidates for the treatment of neurodegenerative disorders II: Alzheimer’s disease. *CNS Neurol. Disord. Drug Targets* 2011, 10, 251–270. DOI: 10.2174/187152711794480429.

De la Cruz, M.: “*Libellus de Medicinalibus Indorum herbis*”, México, Fondo de Cultura Económica. 2ª. Edición. 1991.

De la Rosa, M., Wong, J., Muñiz, D., Carrillo, M., & Sánchez, J. (2016). Compuestos fenólicos bioactivos de la toronja (*Citrus paradisi*) y su importancia en la industria farmacéutica y alimentaria. *Revista Mexicana de Ciencias Farmacéuticas*, 47 (2), 22-35. ISSN: 1870-0195.

Debanne, D., Campanac, E., Bialowas, A., Carlier, E., & Alcaraz, G. (2011). Axon physiology. *Physiological reviews*, 91(2), 555–602. <https://doi.org/10.1152/physrev.00048.2009>.

Della, VV. (2017). Uncaria tomentosa. *Giornale italiano di dermatologia e venereologia : organo ufficiale, Societa italiana di dermatologia e sifilografia*, 152(6), 651–657. <https://doi.org/10.23736/S0392-0488.17.05712-1>.

Desiraju, G. R. (1996). The C—H···O hydrogen bond in crystals: What is it? *Accounts of Chemical Research*, 29(9), 441–449. <https://doi.org/10.1021/ar950174c>.

Devi, G. (2023). The tauopathies. *Handbook of clinical neurology*, 196, 251–265.

Dixit, H., Kumar C., Chaudhary, R., Thaker, D., Gadewal, N., & Dasgupta, D. (2021). Role of Phosphorylation and Hyperphosphorylation of Tau in Its Interaction with $\beta\alpha$ Dimeric Tubulin Studied from a Bioinformatics Perspective. *Avicenna journal of medical biotechnology*, 13(1), 24–34. <https://doi.org/10.18502/ajmb.v13i1.4579>.

Ellinwood, E., Nikaido, A., Gupta, S., Heatherly, D., & Nishita, J. (1990). Comparison of central nervous system and peripheral pharmacodynamics to atropine pharmacokinetics. *The Journal of pharmacology and experimental therapeutics*, 255(3), 1133–1139.

Enamorados, Y., Ilyina, A., Silva, Y., Vargas, A., Martínez, J., & Segura, E. (2017). Uso de plantas mexicanas con efecto inhibidor sobre la enzima acetilcolinesterasa como un posible tratamiento para la enfermedad de Alzheimer. *Revista Mexicana de Ciencias Farmacéuticas*, 48, (4), 7-16. https://www.redalyc.org/pdf/579/Resumenes/Resumen_57956617002_1.pdf.

Farinola, N & Piller, N. (2005). Pharmacogenomics: Its Role in Re-establishing Coumarin as Treatment for Lymphedema. *Lymphatic Research and Biology*. 3(2), 81–86.

Fedorova Y. (2019). Sovremennye podkhody k terapii lobno-visochnykh dementsii [Current approaches to the treatment of fronto-temporal dementia]. *Zhurnal nevrologii i psikiatrii imeni S.S. Korsakova*, 119(9. Vyp. 2), 18–24. doi: 10.17116/jnevro201911909218.

Gregory, AJ., Charles, W, Eric, L., Cintia, V., Yvonne, K., Julia, R & Peter, D. (1999). cAMP-Dependent Protein Kinase Phosphorylations on Tau in Alzheimer's Disease. *The Journal of Neuroscience*. 19 (17), 7486-7494. <https://doi.org/10.1523/JNEUROSCI.19-17-07486.1999>.

Gregory, J., Vengalasetti, Y. V., Bredesen, D. E., & Rao, R. V. (2021). Neuroprotective Herbs for the Management of Alzheimer's Disease. *Biomolecules*, 11(4), 543. <https://doi.org/10.3390/biom11040543>.

Groenhof, G. (2013). Introduction to QM/MM simulations. *Methods in Molecular Biology*, 924, 43–66. https://doi.org/10.1007/978-1-62703-17-5_3.

Guimarães, J., Fonseca, R., & Garrett, C. (2006). Demencia frontotemporal: ¿Que entidades? [Frontotemporal dementia]. *Acta medica portuguesa*, 19(4), 319–324.

Gupta, A., Singh, A. K., Loka, M., Pandey, A. K., & Bishayee, A. (2021). Ferulic acid-mediated modulation of apoptotic signaling pathways in cancer. *Advances in protein chemistry and structural biology*, 125, 215–257. <https://doi.org/10.1016/bs.apcsb.2020.12.005>.

Halgren, T. A. (1996). Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *Journal of Computational Chemistry*, 17(5-6), 490-519. [https://doi.org/10.1002/\(SICI\)1096-987X\(199604\)17:5/6<490::AID-JCC1>3.0.CO;2-P](https://doi.org/10.1002/(SICI)1096-987X(199604)17:5/6<490::AID-JCC1>3.0.CO;2-P).

Heitzman, M., Neto, C., Winiarz, E., Vaisberg, A., Hammond, G. (2005). Ethnobotany, phytochemistry and pharmacology of *Uncaria* (Rubiaceae). *Phytochemistry* 2005, 66, 5–29. DOI: 10.1016/j.phytochem.2004.10.022.

Hitl, M., Kladar, N., Gavarić, N., & Božin, B. (2021). Rosmarinic Acid-Human Pharmacokinetics and Health Benefits. *Planta medica*, 87(4), 273–282. <https://doi.org/10.1055/a-1301-8648>.

Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual Molecular Dynamics. *Journal of Molecular Graphics*, 14, 33-38, 0263-7855. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5).

Instituto Mexicano del Seguro Social. (2023). *Demencia: un reto a la salud de importancia creciente*. Gobierno de México. <https://www.gob.mx/imss/articulos/demencia-un-reto-a-la-salud-de-importancia-creciente>

Jellinger K. Neuropathology of Dementia Disorders. *JAlzheimers Dis Parkinsonism* 2014;4:1-17.

Jo, S., Kim, T., Iyer, V. G., & Im, W. (2008). CHARMM-GUI: a web-based graphical user interface for CHARMM. *Journal of computational chemistry*, 29(11), 1859–1865. <https://doi.org/10.1002/jcc.20945>.

Jugran, A. K., Rawat, S., Bhatt, I. D., & Rawal, R. S. (2019). Valeriana jatamansi: An herbaceous plant with multiple medicinal uses. *Phytotherapy research : PTR*, 33(3), 482–503. <https://doi.org/10.1002/ptr.6245>.

Kametani, F., & Hasegawa, M. (2018). Reconsideration of Amyloid Hypothesis and Tau Hypothesis in Alzheimer's Disease. *Frontiers in neuroscience*, 12, 25. <https://doi.org/10.3389/fnins.2018.00025>.

Karadağ, A., Demirci, B., Çaşkurlu, A., Demirci, F., Okur, M., Orak, D., Sipahi, H. & Başer, K. (2019). *In vitro* antibacterial, antioxidant, anti-inflammatory and analgesic evaluation of *Rosmarinus officinalis* L. flower extract fractions. *South African Journal of Botany*, 125, 214-220. <https://doi.org/10.1016/j.sajb.2019.07.039>.

Kernou, O. N., Azzouz, Z., Madani, K., & Rijo, P. (2023). Application of Rosmarinic Acid with Its Derivatives in the Treatment of Microbial Pathogens. *Molecules (Basel, Switzerland)*, 28(10), 4243. <https://doi.org/10.3390/molecules28104243>.

Khaw, K., Parat, M., Shaw, P. & Falconer, J. (2017). Solvent Supercritical Fluid Technologies to Extract Bioactive Compounds from Natural Sources: A Review. *Molecules*, 12, 1-22. DOI: 10.3390/molecules22071186.

Kim, S., Chen, J., Cheng, T., Gindulyte, A., He, J., He, S., Li, Q., Shoemaker, BA, Thiessen, PA, Yu, B., Zaslavsky, L., Zhang, J. & Bolton, EE (2023). Actualización de PubChem 2023. *Ácidos nucleicos.*, 51 (D1), D1373–D1380. <https://doi.org/10.1093/nar/gkac956>.

- Kimura, T., & Takamatsu, J. (2013). Pilot study of pharmacological treatment for frontotemporal dementia: risk of donepezil treatment for behavioral and psychological symptoms. *Geriatrics & gerontology international*, *13*(2), 506–507. <https://doi.org/10.1111/j.1447-0594.2012.00956>.
- Klaas, C., Wagner, G., Laufer, S., Sosa, S., Della Loggia, R., Bomme, U., Pahl, H., & Merfort, I. (2002). Studies on the anti-inflammatory activity of phytopharmaceuticals prepared from Arnica flowers. *Planta medica*, *68*(5), 385–391. <https://doi.org/10.1055/s-2002-32067>.
- Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B.E., Bussonnier, M., Frederic, J., *et al.* (2016). Jupyter Notebooks - a publishing format for reproducible computational workflows. International Conference on Electronic Publishing, 87-90. 10.3233/978-1-61499-649-1-87.
- Kovacs G. (2017). Tauopathies. *Handbook of clinical neurology*, *145*, 355–368. <https://doi.org/10.1016/B978-0-12-802395-2.00025-0>.
- Lee, J., Cheng, X., Swails, J., Yeom, M., Eastman, P., Lemkul, J., Wei, S., Buckner, J., Jeong, J., Qi, Y., Jo, S., Pande, V., Case, D., Brooks, C. L., MacKerell, A., Jr, Klauda, J., & Im, W. (2016). CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *Journal of chemical theory and computation*, *12*(1), 405–413. <https://doi.org/10.1021/acs.jctc.5b00935>.
- Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R News*, *2*(3), 18-22.
- Lipinski, F., Lombardo, B., Dominy, P., Feeney, P. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug. Delivery Rev.* *23*, 4-25 (1997). DOI: 10.1016/s0169-409x(00)00129-0.
- Liu, X., Geng, Z., Ding, X., Lou, Y., & Zhang, X. (2022). Convallatoxin suppresses osteosarcoma cell proliferation, migration, invasion, and enhances osteogenic differentiation by downregulating parathyroid hormone receptor 1 (PTHr1) expression and inactivating Wnt/ β -catenin pathway. *Bioengineered*, *13*(5), 13280–13292. <https://doi.org/10.1080/21655979.2022.2080363>.
- Llorens, M., Jurado, J., Hernández, F. & Ávila, J. (2014). GSK-3 β , a pivotal kinase in Alzheimer disease. *Frontiers in molecular neuroscience*, *7*, 1-11.
- Lozada, J., Torres, A., Reguera, E. & Aguirre, C. Database-Machine Learning Analysis of Cu-Exchanged Zeolites for Methane-Methanol Conversion at Mild Conditions. Available at <http://dx.doi.org/10.2139/ssrn.4429053>.
- Lozoya, L. L., & Legorreta, X. L. (1994). *Plantas, medicina y poder: breve historia de la herbolaria mexicana* (Vol. 5). Editorial Pax México.
- Lyss, G., Schmidt, T., Merfort, I., & Pahl, H. (1997). Helenalin, an anti-inflammatory sesquiterpene lactone from Arnica, selectively inhibits transcription factor NF-kappaB. *Biological chemistry*, *378*(9), 951–961. <https://doi.org/10.1515/bchm.1997.378.9.951>.
- Maccioni, R., Navarrete, L., & San Martín, A. (2011). Método para inhibir la agregación de proteína tau y tratamiento de la enfermedad de alzheimer con un compuesto derivado de la quinolina

(Chile Patent no.pct/CL2011/000028). *Organización Mundial de la Propiedad Intelectual*. <https://patentscope.wipo.int/search/es/detail.jsf?docId=WO2011134098>.

Mahadevan, S., & Park, Y. (2008). Multifaceted therapeutic benefits of Ginkgo biloba L.: chemistry, efficacy, safety, and uses. *Journal of food science*, 73(1), R14–R19. <https://doi.org/10.1111/j.1750-3841.2007.00597.x>.

Makhouri, F., & Ghasemi, J. (2018). *In Silico* Studies in Drug Research Against Neurodegenerative Diseases. *Current neuropharmacology*, 16(6), 664–725. <https://doi.org/10.2174/1570159X15666170823095628>.

Malik, S., de Mesquita, L., Silva, C., de Mesquita, J., de Sá Rocha, E., Bose, J., Abiri, R., de Maria Silva Figueiredo, P., & Costa-Júnior, L. (2019). Chemical Profile and Biological Activities of Essential Oil from *Artemisia vulgaris* L. Cultivated in Brazil. *Pharmaceuticals (Basel, Switzerland)*, 12(2), 49. <https://doi.org/10.3390/ph12020049>.

Mandelkow, E. M., Biernat, J., Drewes, G., Gustke, N., Trinczek, B., & Mandelkow, E. (1995). Tau domains, phosphorylation, and interactions with microtubules. *Neurobiology of aging*, 16(3), 355–363. [https://doi.org/10.1016/0197-4580\(95\)00025-a](https://doi.org/10.1016/0197-4580(95)00025-a).

Manzoor, M., Singh, J., & Gani, A. (2022). Exploration of bioactive peptides from various origin as promising nutraceutical treasures: *In vitro*, *in silico* and *in vivo* studies. *Food chemistry*, 373(Pt A), 131395. <https://doi.org/10.1016/j.foodchem.2021.131395>.

Marcus, DH., Donald, EC., David, CL., Tim, V., Eva, Z & Geoffrey, RH. (2012). Avogadro: An advanced semantic chemical editor, visualization, and analysis platform. *Journal of Cheminformatics*, 4, 17. <https://doi.org/10.1186/1758-2946-4-17>.

Matías-Guiu, J. A., & García-Ramos, R. (2013). Primary progressive aphasia: from syndrome to disease. *Neurologia (Barcelona, Spain)*, 28(6), 366–374. <https://doi.org/10.1016/j.nrl.2012.04.003>.

Meller, J. (2001). Molecular Dynamics. *Encyclopedia of life sciences*, 18.

Meng, XY., Zhang, HX., Mezei, M., & Cui, M. (2011). Molecular docking: a powerful approach for structure-based drug discovery. *Current computer-aided drug design*, 7(2), 146–157. <https://doi.org/10.2174/157340911795677602>.

Mollinsparation. (2023). Tools and Resources. <https://www.mollinsparation.com/tools-and-resources>.

Montastruc, J. L., & Tran, M. A. (1998). Atropine [Atropine]. *La Revue du praticien*, 48(1), 85–87.

Morris, GM., & Lim, WM. (2008). Molecular docking. *Methods in molecular biology (Clifton, N.J.)*, 443, 365–382. https://doi.org/10.1007/978-1-59745-177-2_19.

Ogundele, O., Omoaghe, A., Ajonijebu, D., Ojo, A., Fabiyi, T., Olajide, O., Falode, D., & Adeniyi, P. (2014). Glia activation and its role in oxidative stress. *Metabolic brain disease*, 29(2), 483–493. <https://doi.org/10.1007/s11011-013-9446-7>.

Paintaud, G., Alván, G., Berninger, E., Gustafsson, L. L., Idrizbegovic, E., Karlsson, K. K., & Wakelkamp, M. (1994). The concentration-effect relationship of quinine-induced hearing impairment. *Clinical pharmacology and therapeutics*, 55(3), 317–323. <https://doi.org/10.1038/clpt.1994.32>.

Pandi, A., & Kalappan, V. M. (2021). Pharmacological and therapeutic applications of Sinaptic acid-an updated review. *Molecular biology reports*, 48(4), 3733–3745. <https://doi.org/10.1007/s11033-021-06367-0>.

Parthasarathi, R., & Dhawan, A. (2018). In Silico Approaches for Predictive Toxicology. In *Vitro Toxicology*. Elsevier Inc. <https://doi.org/10.1016/B978-0-12-804667-8.00005-5>.

Patocka, J., Nepovimova, E., Wu, W., & Kuca, K. (2020). Digoxin: Pharmacology and toxicology-A review. *Environmental toxicology and pharmacology*, 79, 103400. <https://doi.org/10.1016/j.etap.2020.103400>.

Pavlíková N. (2022). Caffeic Acid and Diseases-Mechanisms of Action. *International journal of molecular sciences*, 24(1), 588. <https://doi.org/10.3390/ijms24010588>.

Pelkonen, O., Turpeinen, M., & Raunio, H. (2011). *In vivo-in vitro-in silico* pharmacokinetic modelling in drug development: current status and future directions. *Clinical pharmacokinetics*, 50(8), 483–491. <https://doi.org/10.2165/11592400-000000000-00000>.

Perea, G., & Araque, A. (2003). Nuevas vías de información en el sistema nervioso: comunicación entre astrocitos y neuronas [New information pathways in the nervous system: communication between astrocytes and neurones]. *Revista de neurologia*, 36(2), 137–144.

Petersen, M., & Simmonds, M. S. (2003). Rosmarinic acid. *Phytochemistry*, 62(2), 121–125. [https://doi.org/10.1016/s0031-9422\(02\)00513-7](https://doi.org/10.1016/s0031-9422(02)00513-7)

Pîrșcoveanu, D., Pirici, I., Tudorică, V., Bălșeanu, T., Albu, V., Bondari, S., Bumbea, A., & Pîrșcoveanu, M. (2017). Tau protein in neurodegenerative diseases - a review. *Romanian journal of morphology and embryology = Revue roumaine de morphologie et embryologie*, 58(4), 1141–1150.

Prasanna, S., & Doerksen, R. J. (2009). Topological polar surface area: a useful descriptor in 2D-QSAR. *Current medicinal chemistry*, 16(1), 21–41. <https://doi.org/10.2174/092986709787002817>.

Ramakrishnan, P & Alok, D. (2018). In Silico Approaches for Predictive Toxicology. Alok Dhawan, Seok Kwon, In *Vitro Toxicology*, Academic Press, 91-109. <https://doi.org/10.1016/B978-0-12-804667-8.00005-5>.

- Rascovsky, K., Hodges, J. R., Knopman, D., Mendez, M. F., Kramer, J. H., Neuhaus, J., & Miller, B. L. (2011). Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain*, 134(9), 2456-2477.
- Reynolds, C. H., Betts, J. C., Blackstock, W. P., Nebreda, A. R., & Anderton, B. H. (2000). Phosphorylation sites on tau identified by nanoelectrospray mass spectrometry: differences in vitro between the mitogen-activated protein kinases ERK2, c-Jun N-terminal kinase and P38, and glycogen synthase kinase-3beta. *Journal of neurochemistry*, 74(4), 1587–1595. <https://doi.org/10.1046/j.1471-4159.2000.0741587.x>.
- Rim K. T. (2020). *In silico* prediction of toxicity and its applications for chemicals at work. *Toxicology and environmental health sciences*, 12(3), 191–202. <https://doi.org/10.1007/s13530-020-00056-4>.
- Rouhiainen, L. (2018). Inteligencia artificial. *Madrid: Alienta Editorial*, 20-21.
- Sayers, E., Bolton, E., Brister, J., Canese, K., Chan, J., Comeau, D., Connor, R., Funk, K., Kelly, C., Kim, S., Madej, T., Marchler-Bauer, A., Lanczycki, C., Lathrop, S., Lu, Z., Thibaud-Nissen, F., Murphy, T., Phan, L., Skripchenko, Y., Tse, T., ... Sherry, S. T. (2022). Database resources of the national center for biotechnology information. *Nucleic acids research*, 50(D1), D20–D26. <https://doi.org/10.1093/nar/gkab1112>.
- Schittkowski, MP., & Sturm, V. (2017). Atropine for the Prevention of Progression in Myopia - Data, Side Effects, Practical Guidelines. *Klin Monbl Augenheilkd*, 235(4), 385-39.
- Schrödinger, L., & DeLano, W. (2020). *PyMOL*. Retrieved from <http://www.pymol.org/pymol>
- Sharifani, K. y Amini, M. (2023). Aprendizaje automático y aprendizaje profundo: una revisión de métodos y aplicaciones. *Revista mundial de ingeniería y tecnología de la información*, 10 (07), 3897-3904.
- Sharifi, RJ., Rayess, YE., Rizk, AA., Sadaka, C., Zgheib, R., Zam, W., *et al* (2020). Turmeric and Its Major Compound Curcumin on Health: Bioactive Effects and Safety Profiles for Food, Pharmaceutical, Biotechnological and Medicinal Applications. *Frontiers in pharmacology*, 11. <https://doi.org/10.3389/fphar.2020.01021>.
- Singh, R., Mandrah, K., Asati, A., Patel, D. K., Goel, B., Vishwakarma., *et al.* (2019). Transformation of Santonin to a Naproxen Analogue with Anti-Inflammatory Activity. *Journal of natural products*, 82(6), 1710–1713. <https://doi.org/10.1021/acs.jnatprod.8b00318>.
- Smith, J. V., & Luo, Y. (2004). Studies on molecular mechanisms of Ginkgo biloba extract. *Applied microbiology and biotechnology*, 64(4), 465–472. <https://doi.org/10.1007/s00253-003-1527-9>.
- Stermitz, F., Scriven, L., Tegos, G., & Lewis, K. (2002). Two flavonols from *Artemisa annua* which potentiate the activity of berberine and norfloxacin against a resistant strain of *Staphylococcus aureus*. *Planta medica*, 68(12), 1140–1141. <https://doi.org/10.1055/s-2002-36347>.

The effect of curcumin on cognition in Alzheimer's disease and healthy aging: A systematic review of pre-clinical and clinical studies, *Brain Research*, Volume 1725, 2019, 146476, ISSN 0006-8993 <https://doi.org/10.1016/j.brainres.2019.146476>.

Tisher, A., & Salardini, A. (2019). A Comprehensive Update on Treatment of Dementia. *Seminars in neurology*, 39(2), 167–178. <https://doi.org/10.1055/s-0039-1683408>.

Topol, E. J. (2019). *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again*. Basic Books.

UniProt Consortium (2023). UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic acids research*, 51(D1), D523–D531. <https://doi.org/10.1093/nar/gkac1052>.

Vajda F. J. (2002). Neuroprotection and neurodegenerative disease. *Journal of clinical neuroscience : official journal of the Neurosurgical Society of Australasia*, 9(1), 4–8. <https://doi.org/10.1054/jocn.2001.1027>.

Vandekerckhove, S., & D'hooghe, M. (2015). Quinoline-based antimalarial hybrid compounds. *Bioorganic & medicinal chemistry*, 23(16), 5098–5119. <https://doi.org/10.1016/j.bmc.2014.12.018>.

Venclovas, Č., & Margelevičius, M. (2005). Comparative modeling in CASP6 using consensus approach to template selection, sequence-structure alignment, and structure assessment. *Proteins*, 61 Suppl 7, 99–105. <https://doi.org/10.1002/prot.20725>.

Vilela, M., Del Bem, L., Van Sluys, M., de Setta, N., Kitajima, J., Cruz, G., Sforça, D., de Souza, A., Ferreira, P., Grativol, C., Cardoso-Silva, C., Vicentini, R., & Vincentz, M. (2017). Analysis of Three Sugarcane Homo/Homeologous Regions Suggests Independent Polyploidization Events of *Saccharum officinarum* and *Saccharum spontaneum*. *Genome biology and evolution*, 9(2), 266–278. <https://doi.org/10.1093/gbe/evw293>.

Voulgaropoulou, S. D., van Amelsvoort, T. A. M. J., Prickaerts, J., & Vingerhoets, C. (2019). The effect of curcumin on cognition in Alzheimer's disease and healthy aging: A systematic review of pre-clinical and clinical studies. *Brain research*, 1725, 146476. <https://doi.org/10.1016/j.brainres.2019.146476>.

Wang, D., Huang, X., Yan, L., Zhou, L., Yan, C., Wu, J., Su, Z., & Huang, Y. (2021). The Structure Biology of Tau and Clue for Aggregation Inhibitor Design. *The protein journal*, 40(5), 656–668. <https://doi.org/10.1007/s10930-021-10017-6>.

Wang, J., Su, S., Zhang, S., Zhai, S., Sheng, R., Wu, W., & Guo, R. (2019). Structure-activity relationship and synthetic methodologies of α -santonin derivatives with diverse bioactivities: A mini-review. *European journal of medicinal chemistry*, 175, 215–233. <https://doi.org/10.1016/j.ejmech.2019.04.066>.

Wiffen, P. J., Derry, S., & Moore, R. A. (2017). Tramadol with or without paracetamol (acetaminophen) for cancer pain. *The Cochrane database of systematic reviews*, 5(5), CD012508. <https://doi.org/10.1002/14651858.CD012508.pub2>.

- Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, *30*(1), 79-82.
- Willow, JHL (2011). Traditional Herbal Medicine Research Methods: Identification, Analysis, Bioassay, and Pharmaceutical and Clinical Studies. Traditional Herbal Medicine Research Methods, 1-26. 10.1002/9780470921340.
- Xiang Z. (2006). Advances in homology protein structure modeling. *Current protein & peptide science*, *7*(3), 217–227. <https://doi.org/10.2174/138920306777452312>.
- Xu, D., & Zhang, Y. (2011). Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization. *Biophysical journal*, *101*(10), 2525–2534. <https://doi.org/10.1016/j.bpj.2011.10.024>.
- Xu, M., Lu, M., Zhang, W., Jin, Q., & Chen, Y. (2021). Simultaneous Detection of Six Isoforms of Tau Protein in Human Cerebrospinal Fluid by Multidimensional Mass Spectrometry-Based Targeted Proteomics. *Journal of proteome research*, *20*(5), 2299–2307. <https://doi.org/10.1021/acs.jproteome.0c00826>.
- Yang, J., & Zhang, Y. (2015). I-TASSER server: new development for protein structure and function predictions. *Nucleic acids research*, *43*(W1), W174–W181. <https://doi.org/10.1093/nar/gkv342>.
- Yang, J., Roy, A., & Zhang, Y. (2013). BioLiP: a semi-manually curated database for biologically relevant ligand-protein interactions. *Nucleic acids research*, *41*, D1096–D1103. <https://doi.org/10.1093/nar/gks966>.
- Yiannopoulou KG, Papageorgiou SG. Current and future treatments for Alzheimer's disease. *Ther Adv Neurol Disord*. 2013 Jan;*6*(1):19-33. doi: 10.1177/1756285612461679. PMID: 23277790; PMCID: PMC3526946.
- Yin, C., Liu, B., Wang, P., Li, X., Li, Y., Zheng, X., Tai, Y., Wang, C., & Liu, B. (2020). Eucalyptol alleviates inflammation and pain responses in a mouse model of gout arthritis. *British journal of pharmacology*, *177*(9), 2042–2057. <https://doi.org/10.1111/bph.14967>.
- Yoon, B., Jung, J., Lee, J., Cho, Y., Jang, C., Jin, C., T. H., & Ryu, J. (2007). Anxiolytic-like effects of sinapic acid in mice. *Life sciences*, *81*(3), 234–240. <https://doi.org/10.1016/j.lfs.2007.05.007>.
- Yoshiyama, Y., Higuchi, M., & Zhang, B. (2012). Tau pathology in neurodegenerative diseases: The importance of amyloid β precursor protein processing. *Journal of Alzheimer's Disease*, *32*(4), 823-842. <https://doi.org/10.3233/JAD-2012-120634>
- Zhang Y. (2008). Progress and challenges in protein structure prediction. *Current opinion in structural biology*, *18*(3), 342–348. <https://doi.org/10.1016/j.sbi.2008.02.004>.
- Zhang, H., Liu, C., & Zheng, Q. (2019). Development and application of anthelmintic drugs in China. *Acta tropica*, *200*, 105181. <https://doi.org/10.1016/j.actatropica.2019.105181>.

Zhang, J., Yang, P. & Gray, N. Targeting cancer with small molecule kinase inhibitors. *Nat Rev Cancer* 9, 28–39 (2009). <https://doi.org/10.1038/nrc2559>.

Zhang, X., Gao, F., Wang, D., Li, C., Fu, Y., He, W., & Zhang, J. (2018). Tau Pathology in Parkinson's Disease. *Frontiers in neurology*, 9, 809. <https://doi.org/10.3389/fneur.2018.00809>.

Zhang, Y., & Skolnick, J. (2004). SPICKER: a clustering approach to identify near-native protein folds. *Journal of computational chemistry*, 25(6), 865–871. <https://doi.org/10.1002/jcc.20011>.

Zhang, Y., Kolinski, A., & Skolnick, J. (2003). TOUCHSTONE II: a new approach to ab initio protein structure prediction. *Biophysical journal*, 85(2), 1145–1164. [https://doi.org/10.1016/S0006-3495\(03\)74551-2](https://doi.org/10.1016/S0006-3495(03)74551-2).

Zheng, W., Wuyun, Q., Zhou, X., Li, Y., Freddolino, P. L., & Zhang, Y. (2022). LOMETS3: integrating deep learning and profile alignment for advanced protein template recognition and function annotation. *Nucleic acids research*, 50(W1), W454–W464. <https://doi.org/10.1093/nar/gkac248>.

Zheng, W., Zhang, C., Li, Y., Pearce, R., Bell, E. W., & Zhang, Y. (2021). Folding non-homologous proteins by coupling deep-learning contact maps with I-TASSER assembly simulations. *Cell reports methods*, 1(3), 100014. <https://doi.org/10.1016/j.crmeth.2021.100014>.

Zhou, X., Zheng, W., Li, Y. *et al.* I-TASSER-MTD: a deep-learning-based platform for multi-domain protein structure and function prediction. *Nat Protoc* 17, 2326–2353 (2022). <https://doi.org/10.1038/s41596-022-00728-0>.

Apéndice

Apéndice A “Propiedades ADME y regla de *Lipinski*”

Apéndice B “Plantillas utilizadas en el Modelado de Tau”

Apéndice C “Características del servidor utilizado para los cálculos de DM”

Apéndice D “Programación en Python del ARBA”

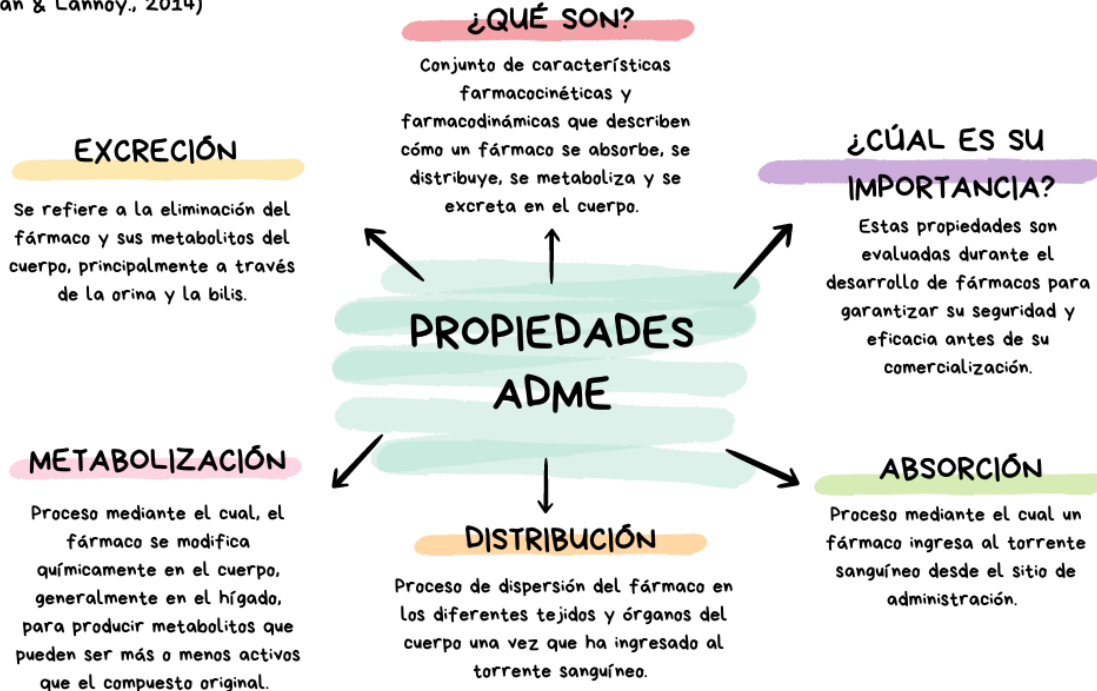
Apéndice E “Cargas de *Gasteiger* aplicadas a los *hits*”

Apéndice F “Error absoluto medio”


Apéndice A “Propiedades ADME y regla de *Lipinski*”

Mapa conceptual propiedades ADME.

Referencias:
(Alqahtani., 2017)
(Fan & Lannoy., 2014)



REGLA DE 5 DE “LIPINSKI”



Christopher A. Lipinski, año 1997.

Criterios utilizados en el diseño de fármacos para evaluar la probabilidad de que un compuesto químico se absorba adecuadamente en el torrente sanguíneo.

- 1** **Peso o masa molecular no superior a 500 Daltons (g/mol).**

Sugiere que los compuestos con una masa molecular mayor a 500 g/mol pueden tener dificultades para atravesar las membranas biológicas, lo que podría limitar su absorción.
- 2** **LogP (Coeficiente de partición octanol/agua) no superior a 5.**

Un valor de LogP no mayor a 5 sugiere una adecuada solubilidad en agua, lo que es importante para la distribución del fármaco en el organismo.
- 3** **Número de aceptores de enlaces de hidrógeno (nON) menor o igual a 10.**




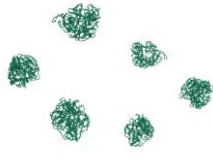

La presencia del aumento de grupos de hidrógeno incrementa la susceptibilidad del fármaco a la metabolización, lo que puede afectar su vida media y su eficacia terapéutica.
- 4** **Número de donadores de enlaces de hidrógeno (nOHNH) menor o igual a 5.**

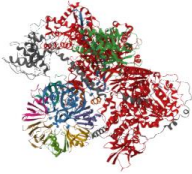


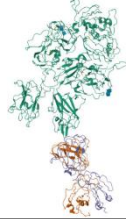

Demasiados grupos donantes de hidrógeno pueden aumentar la polaridad de la molécula y afectar su capacidad para disolverse en los medios biológicos adecuados. Esto podría comprometer la absorción y distribución del fármaco en el cuerpo.

Referencia:
(Lipinski *et al.*, 1997)

Apéndice B “Plantillas utilizadas en el Modelado de Tau”

Esta Tabla indica los valores porcentuales de identidad (Porcentaje de identidad 1 y 2) y la cobertura de alineación de subprocesos respecto a la longitud de la proteína (*Cov*).

Número	PDB	Porcentaje de identidad 1.	Porcentaje de identidad 2.	Cov	Puntuación Z	Imagen
1	7pqcA	1.00	0.26	0.26	2.07	
2	7e2cI	0.06	0.17	0.91	2.96	
3	7pqc	1.00	0.26	0.26	6.20	
4	3ja4A	0.05	0.23	1.00	5.97	
5	7pqc	1.00	0.26	0.26	3.55	

6	5gaoE	0.11	0.07	0 · 4 5	1.46	
7	7pqe	1.00	0.26	0 · 2 5	5.52	
8	2nbiA	0.13	0.17	0 · 6 5	1.74	
9	8g6iA	0.08	0.23	1 · 0 0	2.96	
1 0	7n95A	0.08	0.29	0 · 9 8	2.30	

Apéndice C “Características del servidor utilizado para los cálculos de DM”

Propiedades del servidor	Especificaciones
Memoria	94.4 GB
Procesador	Intel® Xenon(R)CPU X5650 @ 2.67 GHz x 24
Gráficos	llvmpipe (LLVM 11.01.128 bits)
Capacidad del disco	2 TB
Nombre del SO	Debian GNU/Linux 11 (bullseye)
Tipo de SO	64 bits
Versión de GNOME	3.38.5
Sistema de ventanas	Wayland

Apéndice D “Programación en Python del ARBA”

Obtenido de (Lozada *et al.*,2022) modificada para la implementación a la base CHM.

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt

herbolaria=pd.read_excel(r'C:\ruta')

df=pd.DataFrame(np.nan_to_num(herbolaria))

features=df.set_axis(['PM','Ant', 'TS', 'TPSA', 'IQ', 'logP', 'nON', 'nOHNH'], axis=1)

labels= np.array(features['TPSA'])

features= features.drop('TPSA', axis = 1)

feature_list = list(features.columns)

features = np.array(features)

from sklearn.model_selection import train_test_split
train_features, test_features, train_labels, test_labels = train_test_split(features, labels,
test_size = 0.25, random_state =42)

seline_preds = test_features[:, feature_list.index('PM')]

baseline_errors = abs(baseline_preds - test_labels)
print ('Average baseline error:',round(np.mean(baseline_errors),2))

from sklearn.ensemble import RandomForestRegressor
rf= RandomForestRegressor(n_estimators= 1000, random_state =42)
rf.fit(train_features, train_labels);

predictions = rf.predict(test_features)
errors= abs(predictions - test_labels)
print('Mean Absolute Error:', round(np.mean(errors), 2),'degrees.')
```

```
importances = list(rf.feature_importances_)
feature_importances = [(feature, round(importance., 2005)).
```

Apéndice E “Cargas de *Gasteiger* aplicadas a los *hits*”

Ligantes (<i>Hits</i> y Fosfato)	Carga de <i>Gasteiger</i> implementada
Fosfato	-15.047
Ácido cafeico	-16.987
Ácido sinápico	-16.675
Curcumina	-16.044
Especiofilina	-14.987
Eucaliptol	-15.067

Apéndice F “Error absoluto medio”

La fórmula matemática del Error absoluto medio es la siguiente:

$$\text{Error absoluto medio} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

En donde:

n es el número total de observaciones,

y_i es el valor real para la i – ésima observación,

\hat{y}_i es la predicción del modelo para la i – ésima observación,

$|y_i - \hat{y}_i|$ es la diferencia absoluta entre el valor real y la predicción.

En resumen, la forma de calcular este parámetro consiste en que, para cada observación se estima la diferencia absoluta entre el valor real y la predicción, para posteriormente dividir entre el total de observaciones.

En la implementación del ARBA la metodología de cálculo fue la siguiente:

Durante la etapa de prueba el algoritmo se seleccionó aleatoriamente 4 compuestos (25% de los datos totales de la base de datos) y se calculó las diferencias entre su PM y el TPSA en valor absoluto y se definió como e_1 , de la siguiente forma:

$$e_1 = \frac{1}{4} \sum_{i=1}^4 |PM - TPSA|$$

De donde:

$$e_1 = 42$$

Posteriormente al aplicar el algoritmo de bosque aleatorio a los datos de prueba se obtuvo el

error absoluto medio del descriptor TPSA y se denotó por e_2 :

$$e_2 = \frac{1}{4} \sum_{i=1}^4 |TPSA - \widehat{TPSA}|$$

$$e_2 = 25.87$$

$$\therefore e_2 < e_1$$

Lo anterior significa que el modelo ajusta bien a los datos de prueba, sin importar como fueron seleccionados los 4 datos, puesto que siempre se va a cumplir que $e_2 < e_1$.