



UNIVERSIDAD NACIONAL DEL NORDESTE

Facultad de Cs. Exactas, Naturales y Agrimensura

Doctorado en Química

ENLACES DE HALOGENO EN ENTORNOS BIOLOGICOS

Un nuevo tipo de interacción en complejos ligando-proteína

TESIS DOCTORAL

Bogado María Lucrecia

Director: Dra. Nélida María Peruchena

CoDirector: Dr. Emilio Luis Angelina

Corrientes, Argentina

2023

Prefacio

“Esta tesis se presenta como parte de los requisitos para optar al grado académico de Doctora en Ciencias Químicas de la Universidad Nacional del Nordeste y no ha sido presentada previamente para la obtención de otro título en esta Universidad u otras. La misma contiene los resultados obtenidos en investigaciones llevadas a cabo en el Laboratorio de Estructura Molecular y Propiedades (FACENA-UNNE-CONICET), durante el período comprendido entre Septiembre de 2016 y Diciembre de 2022, bajo la dirección de la Dra. Peruchena, Nélida María y codirección del Dr Emilio Luis Angelina con financiamiento del CONICET”.

Agradecimientos

A la Dra Nélida Peruchena y al Dr Emilio Angelina por haberme dirigido en este trabajo de tesis, por su paciencia y dedicación; y principalmente por haberme dado la posibilidad de llevar a cabo este desafío.

A mis compañeros del Laboratorio de Estructura Molecular y Propiedades (LEMyP) y del IQUIBA-NEA, por su buena predisposición y apoyo.

Agradezco sinceramente a mi familia, mi madre y hermanos, por acompañar y apoyar constantemente mis proyectos.

A mis amigos, “mis llaves”; Ana, Diego, Lore, Marie y Pablo; por haberme acompañado incondicionalmente en los momentos difíciles que me ha tocado atravesar estos últimos años.

Finalmente, aunque ya no nos acompañe mas físicamente, quiero expresar mi agradecimiento a mi padre por haberme enseñado a siempre ser perseverante para alcanzar mis metas.

Trabajos Publicados durante el desarrollo de la tesis

- Parravicini, O., Bogado, M. L., Rojas, S., Angelina, E. L., Andujar, S. A., Gutierrez, L. J., Cabedo, N., Sanz, M. J., López-Gresa, M. P., Cortes, D., & Enriz, R. D. (2017). Tetrahydroisoquinolines functionalized with carbamates as selective ligands of D2 dopamine receptor. *Journal of molecular modeling*, 23(9), 273. <https://doi.org/10.1007/s00894-017-3441-6>.
- Luchi, A., Angelina, E., Bogado, L., Forli, S., Olson, A., & Peruchena, N. (2018). Flap-site Fragment Restores Back Wild-type Behaviour in Resistant Form of HIV Protease. *Molecular informatics*, 37(12), e1800053. <https://doi.org/10.1002/minf.201800053>
- Luchi, A. M., Villafaña, R. N., Gómez Chávez, J. L., Bogado, M. L., Angelina, E. L., & Peruchena, N.M. (2019). Combining Charge Density Analysis with Machine Learning Tools to Investigate the Cruzain Inhibition Mechanism. *ACS Omega*, 4(22), 19582-19594. <https://doi.org/10.1021/acsomega.9b01934>
- Angelina, Emilio & Bustillo, Soledad & Bogado, Lucrecia & Denegri, María & Peruchena, Nélida & Leiva, Laura. (2020). Binding of acidic PLA2 Ba SPII RP4 from Bothrops alternatus snake venom to integrin $\alpha\beta 3$: An in silico study. *Toxicon : official journal of the International Society on Toxinology*. 177 Suppl 1. S52. 10.1016/j.toxicon.2019.12.111.
- Bogado, M. L., Villafaña, R. N., Gómez Chavez, J. L., Angelina, E. L., Sosa, G. L., & Peruchena, N. M. (2022). Targeting Protein Pockets with Halogen Bonds: The Role of the Halogen Environment. *Journal of Chemical Information and Modeling*, 62(24), 6494-6507. https://doi.org/10.1021/ACS.JCIM.2C00475/SUPPL_FILE/CI2C00475_SI_001.PDF

Resumen

Las interacciones no covalentes desempeñan un papel central en la química y la biología. Ellas son las responsables de las propiedades de los polímeros, la estructura terciaria de las macromoléculas, las propiedades moleculares de los sólidos, las estructuras conformacionales preferenciales de las moléculas, la transferencia de energía entre moléculas y restos moleculares, y, en un sentido amplio, son responsables de que ocurran las reacciones químicas. Recientemente, un tipo particular de interacciones moleculares, llamada enlaces de halógeno, EXs, llamados así por su similitud con los enlaces de hidrógeno, EHs, ha cobrado especial interés. El enlace de halógeno, EX, puede definirse como una interacción no covalente del tipo $R-X \cdots Y$, donde X es normalmente yodo, bromo, cloro y en menor medida flúor; Y al que llamaremos *aceptor de halógeno*, es una base de Lewis, normalmente una molécula/átomo con pares electrónicos libres o con una región en la que existe acumulación de densidad electrónica tales como átomos de oxígeno, nitrógeno, azufre o electrones pi; "R" al que denominaremos *dador de halógeno* es un grupo al que el halógeno está unido covalentemente y es por lo general un grupo/átomo aceptor de electrones.

Los halógenos (X), especialmente los más livianos: flúor y cloro, son sustituyentes ampliamente utilizados en química medicinal. Hasta hace poco, fueron meramente percibidos como restos hidrofóbicos y como bases de Lewis debido a su electronegatividad. Muy en contraste con esa percepción, compuestos conteniendo cloro, bromo y yodo también pueden establecer contactos cortos (del tipo $R-X \cdots Y-R'$) entre una región electrofílica en el halógeno X y una región nucleofílica, sobre una base de Lewis, Y. ¿Cómo sucede esto? Con la formación del enlace covalente R-X, la distribución electrónica de X, de simetría esférica en el átomo aislado, se polariza en la dirección del enlace $\sigma(R-X)$, originando en el sentido opuesto a R, una región de densidad electrónica disminuida. Así, el enlace de halógeno, EX, surge de la asimetría de la distribución electrónica del átomo de halógeno puente, es decir la fuerza impulsora de esta interacción $X \cdots Y$ se atribuye a la existencia del denominado por Politzer *agujero- σ* ; este puede visualizarse como una región positiva localizada sobre la superficie del halógeno, en la dirección del enlace R-X y en sentido opuesto a R. Esta deficiencia electrónica que se observa

en la región axial origina un *cinturón rico en electrones* en la región ecuatorial que le permite establecer *interacciones laterales* con electrófilos, al mismo tiempo que establece *interacciones frontales* con nucleófilos, a través del agujero- σ .

En la actualidad un número significativo de drogas y de candidatos a drogas en desarrollo clínico son estructuras halogenadas; la química de los halógenos fue utilizada por los químicos medicinales por más de 70 años, sin embargo hasta no hace mucho tiempo, los halógenos eran considerados meramente como átomos o grupos útiles para la optimización de las llamadas propiedades ADMET (Absorción, Distribución, Metabolismo, Excreción y Toxicidad) ya que mejoran la absorción oral y facilitan la penetración de barreras biológicas, también son útiles para llenar cavidades hidrofóbicas y prolongan la vida media de las drogas en el organismo. Sin embargo, las interacciones directas mediadas por átomos de halógeno han sido anteriormente ignoradas en el contexto del desarrollo de fármacos. Es así que en la actualidad existen numerosos compuestos halogenados de conocida actividad sobre **diferentes blancos moleculares**, para los cuales es necesario re-investigar el rol del átomo de halógeno en la interacción ligando-receptor, en vista de la valoración actual del EX como una interacción no covalente con propiedades única.

Por otro lado, la química computacional es una rama de la química aplicable a numerosos campos. Permite simular numéricamente estructuras, reacciones químicas, interacciones moleculares entre otros, facilitando así el estudio de los fenómenos físico-químicos, y la resolución de problemas que serían más difíciles y costosos de abordar desde el punto de vista experimental. En el diseño de fármacos se utiliza el modelado computacional para predecir cómo la adición o eliminación de un átomo en un nuevo candidato a fármaco aumentará o disminuirá su afinidad al blanco molecular. La mayoría de los átomos se modelan como objetos esféricos, con una distribución isotrópica de carga a través de su superficie. Sin embargo, se ha notado que modelar los átomos de halógeno de esta manera predice incorrectamente los EXs, y puede conducir al fracaso para diseñar compuestos más potentes. Para predecir adecuadamente cómo se comportan los enlaces de halógeno en sistemas bioquímicos primero se debe entender las propiedades geométricas y energéticas de estos enlaces y parametrizarlos adecuadamente en los métodos de Modelado Molecular. En cambio, los métodos mecano-cuánticos describen correctamente los EXs, sin embargo, estos cálculos se vuelven prohibitivos a medida que aumenta el número de átomos en el sistema biológico como también el número de moléculas a ser evaluadas. Con el objeto de realizar una contribución interesante al tema, en esta tesis se elaborará una propuesta que logre describir apropiadamente este tipo de

interacción, utilizando la información estructural disponible para poner a punto las técnicas de modelado molecular, aplicando la aproximación del extra punto (EP) que imita el agujero σ en las técnicas de Docking Molecular y Dinámica Molecular, y así poder aplicarlos al estudio de compuestos halogenados en diversos blancos moleculares, como casos de estudio.

Teniendo en cuenta como los bolsillos de las proteínas que forman un enlace halógeno (EX) con un ligando halogenado constituyen simultáneamente otras interacciones con el átomo de halógeno que puede considerarse como su *entorno de enlace X* (EXent), con el fin de estudiar la implicancia de estas interacciones, como una aporte importante en esta tesis se desarrolla un protocolo para evaluar la viabilidad de los ligandos halogenados para formar un enlace de halógeno en los bolsillos de proteínas, para ello se realiza un análisis de la distribución electrónica y los reordenamientos que presentan principalmente los átomos de hidrógeno, cuando estos forman parte del entorno de un EX, en el contexto de la Teoría Cuántica de Átomos en Moléculas, QTAIM.

La tesis se ha organizado de la siguiente manera:

A modo de introducción en el Capítulo 1 se presentan las consideraciones generales respecto a los EXs, sus propiedades, estructuras y aplicaciones en el diseño de fármacos principalmente. En él se hace una breve reseña acerca de las interacciones no covalentes en la que los átomos de halógenos están presentes, como así también las descripciones teóricas que estos sistemas han tenido a través de la historia y los hallazgos históricos que dieron origen al término enlace de halógeno.

En la sección aplicaciones de los enlaces de halógeno, se explica cómo estas interacciones pueden influir en diferentes áreas. Se brindan algunos ejemplos de cómo los EXs pueden afectar determinadas medidas experimentales, si no se los tienen en cuenta. Sobre el final de este capítulo, se hace una breve reseña sobre casos de enlaces de halógeno en entornos biológicos, y sus características. Finalmente, se definen los objetivos y alcances de la tesis.

En el Capítulo 2 se describen brevemente los fundamentos y la metodología empleada durante el desarrollo de esta tesis, ambos necesarios para la comprensión de los resultados discutidos en capítulos posteriores. Se presenta en primer lugar una breve descripción de los métodos modelado molecular, mecánica molecular y docking molecular. Luego se describen los métodos mecanocuánticos utilizados y la teoría de QTAIM, y finalmente se describe el método de estudio de interacciones a través de “fingerprints” (SIFT).

El Capítulo 3 se divide en dos partes, como se explicó previamente los métodos de modelado molecular no describen correctamente a los EXs, por lo que en la primera de este capítulo tiene como objetivo parametrizar el método de Docking molecular para enlaces de halógeno, para lo que como primer paso se realiza un relevamiento de una base de datos para obtener complejos con ligandos halogenados con EXs, luego siguiendo el modelos de Santos y col. se modifco el campo de fuerzas del software AutoDock, para poder utilizar la aproximación del extra punto (EP) que imita el agujero σ . Se realizan ensayos de Docking sobre todos los complejos halogenados variando el potencial de pozo hasta obtener el valor con el cual los resultados son mas favorables, y luego una vez definido el potencial optimo, se ajusta la direccionalidad del enlace hacia un atomo de oxigeno aceptor.

En la segunda parte de este capitulo se realiza un relevamiento de los complejos ligando-proteína que contienen halógeno, específicamente halobenceno, en entornos biológicos, se estudia el entorno del EXs empleando una herramienta muy rigurosa para describir las interacciones del halógeno, la cual se basa en la distribución de densidad de carga electrónica del sistema (QTAIM). Los resultados de este análisis sugieren que la formación de enlaces X requieren un entorno proteico adecuado para poder establecerse, los átomos que interactúan en mayor proporción son los átomos de hidrogeno indicando que los dadores de enlace H se encuentran preferentemente perpendiculares al eje de enlace covalente $C_{ar}-X$; estas interacciones de enlaces H mejoraría la polarización de la densidad de carga del halógeno fuera de la región axial, dando como resultado un refuerzo del enlace X. Para evaluar la importancia de la distribución de los enlaces de hidrogeno alrededor del halógeno, se procedio a utilizar descriptores basados en densidad de carga electrónica; los resultados obtenidos demuestran que la relacion entre los autovalores de la curvatura en el punto critico de enlace, λ_1 / λ_3 , es el descriptor que mejor representa la preferencia angular y la dependencia con la distancia en EHs alrededor del halógeno.

Finalmente, los datos obtenidos de este relevamiento (parámetros geométricos, densidad electrónica y elementos topológicos) se han usado para generar un sistema de referencia que se utiliza posteriormente para la elaboración de un protocolo para la predicción de EXs en base a su entorno.

En el capítulo 4 consiste en la elaboración de un protocolo que evalúa la fuerza de EXent como una medida de la propensión de un bolsillo de la proteína para albergar un EX. Con este fin se realiza docking parametrizado para EXs a diversos complejos con ligandos halogenados y se

utiliza el sistema de referencia elaborado en el capítulo 3; los descriptores topológicos basados en la densidad de carga en combinación con herramientas de aprendizaje automático se emplean para predecir la formación y la fuerza de las interacciones que conforman el EXent, en función de sus parámetros geométricos. Sobre la base de estos resultados, se propone un protocolo (xenv) que permite predecir en base al EXent la probabilidad de que el bolsillo de una proteína aloje un enlace X. La validación del protocolo se obtuvo realizando redocking de los complejos con EX del sistema de referencia y los entornos de EX predichos en las poses de docking se comparan con los valores reales obtenidos del descriptor λ_1 / λ_3 . Finalmente, con el objetivo de tener una validación externa el rendimiento del protocolo se evaluó más a fondo en 6 complejos unidos a compuestos halogenados que no estaban incluidos en el conjunto de entrenamiento. Las propiedades derivadas de la densidad de carga de los enlaces de hidrógeno H···Cl que conforman los EXent se calcularon con la metodología QTAIM y se predijeron simultáneamente con el protocolo ajustado, obteniendo resultados favorables.

En el capítulo 5 se aplican las técnicas de modelado molecular a uno de los blancos moleculares propuestos en el plan de tesis, los receptores de dopamina D2 (DR2). Como al inicio de este trabajo no se encontraban resueltas las estructuras cristalinas del receptor en forma activa, en una primera etapa se realiza modelado por homología de ambas formas del receptor dentro de una membrana biológica, que luego se someten a simulaciones de Dinámica Molecular. Con estos modelos, y con el fin de profundizar el estudio del blanco molecular en esta tesis se analiza la presencia de los “switches” de activación característicos de los receptores acoplados a la proteína. Posteriormente se analiza el perfil funcional de la dopamina y la m-clorotiramina, aplicando las técnicas de ensemble docking, en este paso para el análisis de las poses de docking se calculan las “fingerprint” basadas en interacciones estructurales (SIFT), siguiendo el protocolo descrito por Deng y col. En una última etapa, se seleccionan dos compuestos clorados, un agonista, m-clorotiramina, y un agonista inverso, clorpromazina. Se realiza Docking Molecular parametrizado para EXs del compuesto agonista con el receptor en la forma activa y del agonista inverso con la estructura del receptor en la forma inactiva, finalmente se aplica el protocolo descripto en el capítulo 4 (xenv) para la predicción de la pose con el entorno más propicio para la formación de un EXs, y además, simultáneamente se calculan con la metodología QTAIM las propiedades derivadas de la densidad de carga de los enlaces H···Cl que conforman los entornos de enlace X para poder comparar los resultados.

En el capítulo 6 se estudian compuestos inhibidores halogenados de la enzima Cruzipain (Cz) del *Trypanosoma cruzi*, a través de simulaciones de Dinámica Molecular ajustada para EXs

siguiendo el protocolo de Ibrahim y col. Se estudian las interacciones establecidas por ligandos halogenados (LX) y sus análogos no halogenados (LH) en el bolsillo de unión de la Cz a través de simulaciones y posterior calculo mecano cuántico de modelos reducidos por QTAIM. Para revelar la importancia de las interacciones en la inhibición de la Cz se compararon los resultados obtenidos computacionalmente con los valores experimentales. Como primer paso se filtran de una base de datos compuestos halogenados inhibidores de la enzima con sus respectivos análogos no halogenados, cada par con una elevada diferencia de actividad; se obtienen del filtrado inhibidores covalentes y no covalentes, aquellos compuestos covalentemente unidos se aplica la técnica de docking covalente para obtener la mejor pose para realizar las simulaciones de dinámica molecular, se parametriza el método para EXs aplicando el protocolo de Ibrahim y col., agregando un EP con carga positiva y sin masa en el campo de fuerza. Para obtener las poses más representativas se realiza un análisis profundo de las dinámicas moleculares, incluyendo el análisis de componentes principales de las trayectorias, obteniendo las estructuras de los máximos y mínimos, sobre las cuales finalmente se realiza el análisis de las interacciones moleculares. Mediante el análisis QTAIM se analizan el enlace de halógeno y las interacciones del entorno del EXs, que demuestran también para este caso de estudio la importancia de los enlaces de hidrogeno en el entorno.

Si bien, al final de cada capítulo se van resumiendo las conclusiones, en el capítulo 7 se detallan las conclusiones generales del trabajo de tesis.

Índice

CAPÍTULO 1: Introducción general y objetivos

1.1. Introducción	2
1.1.1. Interacciones Moleculares.....	2
1.2. Naturaleza del enlace de halógeno	9
1.3. Aplicaciones de los enlaces de halógeno.....	14
1.4. Enlaces de halógeno en entornos biológicos.....	17
1.5. Diseño de Farmacos.....	22
1.6. Planteo del problema y Objetivos.....	26
1.7. Referencias	29

CAPITULO 2: Fundamentos y Metodología

2.1. Modelado por homología.....	40
2.1.1. Modelado con múltiples moldes	43
2.1.2. Servidores web.....	43
2.1.3. Validación del modelo.....	44
2.2. Métodos de Simulación para sistemas moleculares.....	45
2.2.1. Mecánica Molecular.....	46
2.2.1.1Dinamica Molecular	49
2.2.1.2 Cálculos de Energía libre.....	59
2.3. Crivado virtual y Docking Molecular	60
2.3.1 Ensemble Docking.....	66
2.3.2 Docking Covalente.....	67
2.4. Calculos mecano cuanticos.....	68
2.4.1. Propiedades topológicas	70
2.4.2. Campo del vector gradiente de la densidad electrónica.....	72
2.4.3 Propiedades del Laplaciano	74
2.5. Análisis de Interacciones a través de SIFT.....	78
2.6 Aplicaciones	79
2.7 Referencias.....	80

CAPÍTULO 3: Descripción y parametrización de enlaces halógeno en entornos biológicos

3.1. Enlaces de halógeno en métodos de modelado molecular	86
3.2. Procedimiento	90
3.2.1. Parametrización de Docking.....	90

3.2.2. Características de los entornos de los Enlaces de halógeno.....	91
3.3. Resultados.....	92
3.3.1. Parametrización de Docking molecular para compuestos halogenados.....	92
3.3.1.1. Calibración del potencial de pozo.....	92
3.3.1.2. Ajuste de la direccionalidad.....	95
3.3.2. Características de los entornos de los Enlaces de halógeno	97
3.3.2.1. Análisis de densidad electrónica.....	102
3.3.2.2. Fuerza del entorno.....	106
3.4. Conclusiones	111
3.5. Referencias	113

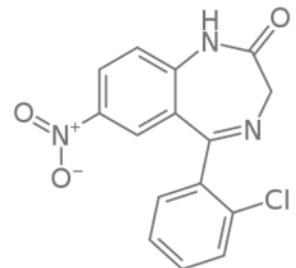
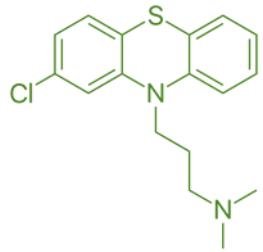
CAPÍTULO 4: Predicción de la formación de Enlaces de halógeno en entornos proteicos

4.1. Enlaces de halógeno en entornos proteicos.....	117
4.2. Procedimiento	122
4.2.1. Entornos de Enlaces de halógeno.....	122
4.2.2. Predicción de la formación de Enlaces de halógeno	122
4.3. Resultados	124
4.3.1. Entornos de Enlaces de halógeno.....	124
4.3.2. Predicción de la formación de Enlaces de halógeno	127
4.3.2.1. Modelo de machine learning.....	127
4.3.2.2. Validación.....	134
4.3.2.3. Entornos reales vs Entornos señuelos.....	135
4.4. Conclusiones	138
4.5. Referencias	140

CAPÍTULO 5: Docking molecular de ligandos halogenados con receptores de dopamina

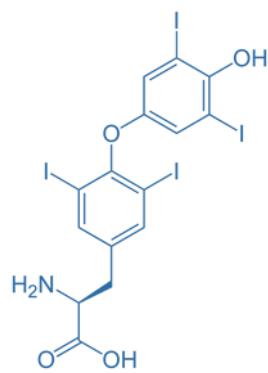
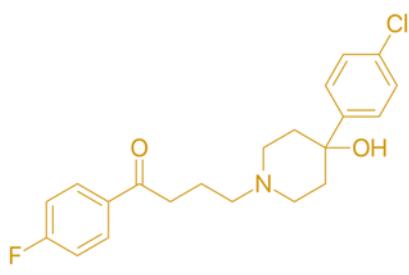
5.1. Introducción.....	145
5.1.1. Receptores acoplados a la proteína G.....	145
5.1.2. Receptores dopaminérgicos.....	150
5.1.3. Rol del halógeno en interacciones ligando-DR2	
5.2. Metodos	159
5.2.1. Construcción de modelos de receptores de dopamina.....	159
5.2.2. Refinamiento de los modelos de homología guiado por el ligando	164
5.2.3. Análisis topológico de la densidad electrónica.....	165
5.2.4. Estructuras cristalinas y comparación con los modelos.....	166
5.2.5. Docking de Dopamina y m-clorotiramina.....	166
5.2.6. Aplicación del protocolo para predecir EXs.....	167

5.3. Resultados.....	167
5.3.1. Analisis del modelado por homología.....	167
5.3.2. Descripción de los switches moleculares.....	170
5.3.3. Activación de D2R por dopamina.....	175
5.3.4. Predicción del perfil funcional de la m-clorotiramina.....	180
5.3.5. Predicción de la formación de EXs basado en entornos proteicos.....	182
5.4. Conclusiones	187
5.5. Referencias	189
CAPÍTULO 6: Estudio de inhibidores halogenados de la enzima Cruzipain y sus interacciones a través de métodos de modelado molecular	
6.1. Introducción.....	196
6.1.1. Enfermedad del Chagas.....	196
6.1.2. Cruzipain.....	201
6.1.3. Inhibidores de la Cruzipain.....	206
6.1.3.1 Enlaces de halógeno en la inhibición de la Cruzipain	
6.2. Metodos	210
6.2.1. Compilación de pares LX/LH con actividad inhibitoria.....	210
6.2.2. Estructuras iniciales de los complejos.....	213
6.2.3 Preparacion de las estructuras.....	213
6.2.4. Simulaciones de Dinámica Molecular parametrizadas para EXs	215
6.2.5. Analisis de las interacciones intermoleculares.....	217
6.3. Resultados	218
6.3.1. Estabilidad de los enlaces S···X / S···H.....	218
6.3.2. Muestreo de la trayectoria.....	219
6.3.2.1. Análisis de Componentes Principales en apo Cz.....	219
6.3.2.2. Análisis de Componentes Principales de Cz/L.....	223
6.3.3. Análisis de densidad electrónica de los pares L/Cz.....	225
6.3.4. Análisis de las interacciones del entorno.....	230
6.4. Conclusiones generales.....	232
6.5. Referencias	234
CAPITULO 7. Conclusiones Generales.....	244
Anexos.....	249



Capítulo 1

Importancia de los Enlaces de Halógeno en el diseño de fármacos



CAPITULO 1

1.1 Introducción

1.1.1 Interacciones moleculares

Hace más de 100 años el conocimiento de los puentes de hidrógeno permitió entender la estructura tridimensional de una molécula fundamental para la vida, como lo es el ADN e increíbles progresos siguieron a este hito científico en biología estructural hasta el conocimiento del genoma humano al inicio de este siglo y mucho más.

Los enlaces de hidrógeno pertenecen al conjunto de interacciones moleculares denominadas *interacciones no covalente*, (NCI), constituyen sin duda la interacción no covalente por excelencia, de mayor abundancia y de relevancia indiscutida en las ciencias de la vida. Como tal, las interacciones intermoleculares (IM) que se establecen entre concentraciones y deplecciones de densidad electrónica (interacciones “hole-lump”) entre dos entidades distintas juegan un rol importantísimo en el reconocimiento molecular L-R (Ligando-Receptor).

A lo largo de la tesis el término interacciones no covalente aparece reservado para interacciones intra e intermoleculares, que no clasifican como típicas interacciones covalentes, pero que son importantes para la estabilización de los sistemas biomoleculares. Así como lo fueron los enlaces de hidrógeno (EH), surge fuertemente hoy en el contexto biológico, el concepto de interacciones mediadas por halógenos, enlaces de halógeno (EX), denominadas también “puentes de halógeno” por su similitud con los puentes de hidrógeno.

El conocimiento de las IM (como los enlaces de halógeno, de hidrógeno, de boro, enlace calcógeno, enlace nicógeno, interacciones $\pi \dots \pi$) aporta una nueva visión en muchas áreas de vanguardia, tanto en el campo del diseño de nuevos materiales funcionales auto-ensamblados, como en el diseño de fármacos asistido por computadora.

En los últimos años, un número significativo de drogas y de candidatos a drogas en desarrollo clínico son estructuras halogenadas; la química de los halógenos fue utilizada por los químicos

medicinales por más de 70 años, sin embargo hasta hace poco tiempo, los halógenos eran considerados meramente como átomos o grupos útiles para la optimización de las llamadas propiedades ADMET (Absorción, Distribución, Metabolismo, Excreción y Toxicidad) ya que mejoran la absorción oral y facilitan la penetración de barreras biológicas; también se ha demostrado su utilidad en el llenado de cavidades hidrofóbicas de los receptores y prolongando la vida media de las drogas en el organismo. Sin embargo, las interacciones directas mediadas por átomos de halógeno han sido ignoradas hasta hace pocos años en el contexto del desarrollo de fármacos. Los átomos de halógeno en compuestos orgánicos halogenados eran considerados típicamente como sitios de alta densidad electrónica dada su alta electronegatividad, debido a este concepto bien establecido se acepta comúnmente que los halógenos puedan establecer interacciones atractivas al funcionar como sitios donantes de electrones, es decir, como nucleófilos. Sin embargo la densidad electrónica en los átomos de halógeno cuando está unido covalentemente a uno o más átomos, por ejemplo a un carbono, se encuentra anisotrópicamente distribuida, es decir hay una región de mayor densidad de electrones, donde el potencial electrostático es negativo, que forma un cinturón ortogonal al enlace covalente C-X y una región de densidad electrónica más baja (el llamado *agujero sigma*) donde el potencial electrostático es frecuentemente positivo, sobre todo si se trata de los halógenos más pesados. Esta región puede formar interacciones atractivas con sitios ricos en electrones, y en general, solo recientemente la capacidad de los átomos de halógenos para interactuar atractivamente con sitios donantes de electrones (nucleófilos) ha sido completamente reconocida y comprendida.

1.1.2 Interacciones no covalentes en entornos biológicos

En los últimos años, se ha extendido el estudio de las interacciones no covalentes a grandes y complejos sistemas biológicos. Una descripción interesante de los eventos de unión (“binding”) puede obtenerse en base al análisis de la distribución de la densidad de carga electrónica en el sitio activo del receptor analizando la red de interacciones intra e intermoleculares que se forman con el ligando, en el bolsillo de unión, considerando no solamente la mejor pose obtenida por docking molecular sino también el conjunto de conformaciones relevantes de los complejos ligando-receptor biológico, obtenidas luego de simulaciones dinámicas (Angelina et al., 2014).

Comenzando con enlaces de hidrógenos fuertes y moderados, como también enlaces de halógenos como se describirá en esta tesis, siguiendo con débiles interacciones polares y

finalizando con interacciones entre anillos aromáticos (“T-shape” y apilamiento o “stacking”) todas ellas pueden investigarse en el contexto de la química cuántica, utilizando la teoría del funcional de la densidad y de la teoría cuántica de átomos en moléculas. Este esquema mecanocuántico de trabajo se ha usado tradicionalmente (desde la década del 90) para estudiar pequeños complejos moleculares, en fase gaseosa, pero aquí se mostrará la aplicación de esta metodología, en el estudio de complejos biomoleculares. Esta se utilizará, a lo largo de la tesis, como una herramienta poderosa, capaz de proveer una descripción muy detallada de los eventos de “binding”.

Uno de los hechos más motivadores para estudiar interacciones NCI radica en sus importantes consecuencias biológicas. Las interacciones moleculares son responsables de la estructura y de la dinámica y consecuentemente también de la función biológica de las macromoléculas. Por ejemplo la interacción de una hormona con su receptor puede producir una cascada de procesos muchos de los cuales dependen o están relacionados con la ruptura o con la formación de NCI. Por lo tanto, entender en profundidad las interacciones no covalentes en complejos biomoleculares constituye un importante desafío. Estas interacciones que van desde interacciones fuertemente electrostáticas a fuertemente dispersivas, han sido extensamente estudiadas por la comunidad científica en sistemas pequeños. En los complejos sistemas biológicos, estas interacciones pueden operar simultáneamente dando origen a interesantes efectos cooperativos, situaciones donde la fuerza colectiva del conjunto de interacciones resulta mayor que la suma de la fortaleza de los enlaces individuales.

Anteriormente, en química medicinal computacional los procesos de reconocimiento molecular se describían con metodología clásica, no cuántica, basadas en el concepto de interacciones de a pares, donde las contribuciones de cada par de átomos se tratan de modo independiente y aditivo. En la actualidad sabemos que todas las interacciones son influenciadas por grupos del entorno, grupos que perturban la distribución de la densidad electrónica y por ello su estudio requiere también de herramientas no clásicas. Los recientes avances en cálculos computacionales hacen posible la apropiada descripción de la red tridimensional de interacciones, enlazantes y no enlazantes, en sistemas biológicos en el contexto de la teoría cuántica de átomos en moléculas.

Las simulaciones de Dinámica Molecular Clásica de diferentes ligandos en el sitio de unión del receptor discriminan bien entre uniones muy fuertes y uniones débiles, pero no permiten distinguir acertadamente en el rango intermedio de fuerza de anclaje. Sin embargo, cuando el sistema que se modela puede reducirse a un modelo 3D contenido solamente los residuos

relevantes del bolsillo de unión, el sistema puede tratarse mecanocuanticamente y la correlación con los valores experimentales de afinidad de unión resulta mayor, para todo el rango de fuerza de anclaje. Además de las mejoras energéticas encontradas, el modelo reducido (RM) permite también representar el grafo molecular de la densidad de carga electrónica, en el sitio de unión.

El trabajo que aquí se presenta pretende contribuir a la comprensión de las NCI (principalmente enlace de hidrógeno y enlace de halógeno) en el contexto de la unión ligando-receptor de varias maneras, por un lado proporcionando una descripción topológica detallada de la red de interacciones del ligando en el bolsillo de unión del receptor, y por otra parte, mostrando la posibilidad de ir *más allá del concepto de interacciones de a pares* para ver los efectos electrónicos dentro del intrincado entorno biológico.

Si bien este tipo de enfoque que involucra la caracterización de las interacciones intermoleculares por medio de QTAIM fue tradicionalmente aplicado al estudio de NCI en complejos de moléculas pequeñas en fase gas, a través de esta tesis se mostrará que esta metodología también es una herramienta muy poderosa para el estudio de complejos biomoleculares, proporcionando una descripción muy detallada de los eventos de unión o de anclaje del ligando con el receptor.

1.1.3 Enlace de halógeno dentro de las NCI

Actualmente es más evidente el hecho de que los halógenos desempeñan un papel directo en la eficacia de ciertos medicamentos, a través de una interacción molecular no covalente, NCI, definida como enlace de halógeno(G. Desiraju et al., 2013) , R-X...Y , donde Y es una base de Lewis o región molecular rica en electrones; es decir, que un halógeno más allá de su conocido comportamiento como base de Lewis puede actuar como ácido de Lewis debido a la distribución asimétrica de su densidad electrónica. El comportamiento electrofílico de los átomos de halógeno es un fenómeno general que afecta a varios procesos, lo que impulsó el interés de la comunidad científica sobre el tema y en los últimos años el número de artículos sobre EX ha crecido muy rápidamente (Figuras 1.1)

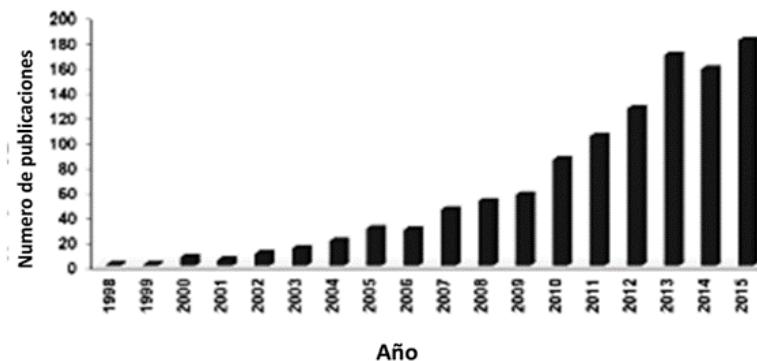


Figura 1.1: Número de publicaciones por año que contienen “halogen bonding” en el título y/o en el resumen.
Tomado de “The Halogen bond” *Chem. Rev.*(Cavallo et al., 2016)

Clark y col. (Clark et al., 2007) estudiaron el origen y la naturaleza del enlace de halógeno en pequeños complejos, atribuyéndola a un “crown” de potencial positivo en la dirección del enlace R-X pero en sentido contrario a R, que hace que el halógeno actúe como un protón (de allí la denominación de enlace de halógeno en similitud con enlace de hidrógeno). En base a potenciales electrostáticos esta interacción fue definida por Politzer como enlace “sigma-hole” descripción que fue adoptada por la comunidad científica(Clark et al., 2005). Sin embargo, desde el Laplaciano de la densidad electrónica, se trata de interacciones hole-lump (Eskandari & Mahmoodabadi, 2013; Zalazar & Peruchena, 2014). Esta característica reconocida hoy en el halógeno en el contexto de la química medicinal, se extiende mucho más allá del enlace de halógeno, por ejemplo, se observa también en enlace calcógeno, enlace tetrel, etc., expandiendo también las posibilidades en la medida que estas interacciones se conozcan más profundamente y puedan explotarse.

Varias interacciones no covalentes pueden clasificarse como interacciones hole-lump. Tomando en consideración la segunda derivada de la densidad electrónica, la función Laplaciana, una interacción hole-lump ocurre cuando una región de depleción de la densidad de carga electrónica (“hole”), en la VSCC (concentración de carga de la capa de valencia) de un átomo (por ejemplo, halógeno, con X = Cl, Br, I), ((Amezaga et al., 2010)(Eskandari & Zariny, 2010). se combina con una región de concentración de carga (“lump”), en el VSCC de un átomo Y (Y= O, N, S, P) para formar un enlace de halógeno (Figura1.2).



Figura 1.2: Interacción *hole-lump* en F-Cl...NH₃, se muestra la envolvente del Laplaciano. Las moléculas están orientadas de modo que el “hole” en la VSCC del cloro está alineado con el “lump” en la VSCC de nitrógeno. Imagen adaptada de Amezaga et al. (Amezaga et al., 2010).

Teniendo en cuenta que en la definición de Laplaciano, (siguiendo el teorema del virial), se utiliza la función energía potencial cuántica completa, el concepto de interacción hole-lump es más útil que el concepto de agujero σ , donde solo se considera la parte electrostática del potencial(Eskandari & Mahmoodabadi, 2013).

Los EXs son interacciones no covalentes muy importantes en los campos de reconocimiento molecular (Corradi et al., 2000; Crihfield et al., 2003; Messina et al., 2001) ingeniería cristalina (Crihfield et al., 2003; Loc Nguyen et al., 2004; Metrangolo et al., 2005; Moorthy et al., 2003), química supramolecular (Lehn, 1988; Politzer, Lane, et al., 2007) y en el diseño de nuevos fármacos (Auffinger et al., 2004; Himmel et al., 2005; Metrangolo et al., 2005), pero además, el hecho de que tanto el átomo de halógeno (X) como la base de Lewis son por lo general átomos con carga parcial negativa, hace que sean sistemas muy interesantes para realizar estudios teóricos (Alkorta et al., 1998, 2008; Lommerse et al., 1996; Metrangolo et al., 2006; Politzer, Murray, et al., 2007; Riley et al., 2009; Wang et al., 2009; Zou et al., 2005)

La idea de que un átomo de halógeno covalentemente unido puede ser atraído (al menos débilmente) por un sitio negativo de otra molécula, puede resultar inverosímil. Sin embargo, este tipo de interacciones se conoce desde hace mucho tiempo. En 1863, Frederick Guthrie dio el primer informe sobre la capacidad de los halógenos para formar aductos bien definidos con especies donadoras de electrones (Guthrie, 1863). En su experimento, el añadía yodo molecular a una solución saturada de nitrato de amonio para formar un complejo. Cuando este complejo era expuesto al aire, se descomponía en forma espontánea en amoníaco y yodo molecular, esto le permitió a Guthrie concluir que el complejo formado era el dímero I₂...NH₃.

La formación de complejos envolviendo EXs con Cloro y Bromo como aceptores de electrones ha sido reportada por primera vez en el siglo 19 por I. Rensen y J.F. Norris, los cuales describieron dímeros formados por Cl_2 y Br_2 con varias aminas (Remsen, 1896). Mientras que el primer compuesto formado por F_2 fue recién obtenido en los años 80 cuando fue posible aislar el anión F^{3-} en condiciones extremas (Ault & Andrews, 1977), y finalmente en los 90 se pudo sintetizar $\text{F}_2\cdots\text{NH}_3$ y $\text{F}_2\cdots\text{OH}_2$ (Legon, 1999).

Los estudios cristalográficos de rayos X realizados por O. Hassel en 1950 fueron cruciales para identificar las características estructurales de la interacción intermolecular que ocurre en varios complejos formados enlaces halógenos y permitieron estudiar muchas de sus propiedades y características. En particular en 1954 mediante técnicas de difracción de rayos X describió la estructura del bromo 1,4-dioxanato(Hassel et al., 1954) como una cadena infinita conteniendo puentes de átomos de halógeno, el experimento reveló que había interacciones intermoleculares presentes entre los átomos de oxígeno del dioxano y los átomos de bromo. La distancia $\text{Br}\cdots\text{O}$ en el cristal era de 2,71 Å, notablemente menor que la suma de los radios de van der Waals del oxígeno y del bromo (3,35 Å), lo que sugería una fuerte interacción entre el átomo de bromo y el átomo de oxígeno. El ángulo $\text{Br}-\text{Br}\cdots\text{O}$ era de aproximadamente 180°. Esto permitió poner en evidencia las características típicas que se encuentran en la formación de enlaces de halógeno, y Hassel llegó a la conclusión de que los átomos de Bromo estaban directamente vinculados a los pares libres del átomo de Oxígeno. Los puentes de halógeno observados en el aducto $\text{Br}_2\cdots\text{O}(\text{CH}_2\text{CH}_2)_2\text{O}$ también fueron encontrado en sistemas similares formados por reacciones de Br_2 y Cl_2 con benceno (Hassel et al., 1959). Estas estructuras eran particularmente interesantes ya que demostró que sistemas π funcionan como donantes de electrones hacia los dihalógenos incluso en estado sólido y sugirió que los aductos con enlaces de halógeno son intermediarios en las reacciones de halogenación de compuestos aromáticos y otros sistemas insaturados. En las sucesivas décadas esta hipótesis fue confirmada (Brown et al., 1990) y J. K. Kochi que mostró que π -donantes forman aductos sólidos con halocarbonos (Rosokha & Kochi, 2008)

Posteriormente, en 1968, el análisis exhaustivo de Bent (Bent, 1968) demostró que en todos los complejos las distancias entre el átomo donador de electrones y el átomo de halógeno eran más cortas que la suma de sus respectivos radios de van der Waals y que los ángulos correspondientes $\text{R}-\text{X}\cdots\text{Y}$ se encontraban cercanos a estaban cerca de 180 °. Estas características geométricas fueron reconocidas 20 años después a través del análisis estadístico

de las estructuras de la Base de datos estructurales de Cambridge, CSD por R. Parthasarathy y G. R. Desiraju (Ramasubbu et al., 1986).

El primer estudio por microondas en fase gaseosa de un complejo que involucra a un átomo de halógeno como aceptor de electrones, fue realizado por Legon y col. en 1975, sobre la estructura del complejo 1:1 formado por el yodotrifluormetano y trimetilamina. Al pasar de los años las geometrías angulares de un gran número de EXs y EHs del tipo B…HCl y B…ClF se han determinado a partir de sus espectros de rotación en fase gaseosa (Legon, 2010).

A pesar de que las interacciones no covalentes que presentan los átomos de halógenos, se conocen desde hace tiempo, el primero en usar el término “halogen bonds” fue Dumas en el contexto de estudios experimentales de los complejos formados por los aceptores: CCl₄, CBr₄, SiCl₄ y SiBr₄ con las bases de Lewis tetrahydrofurano, tetrahydropirano, piridina, anisol y di-n-butileter en solventes orgánicos (Dumas et al., 1983).

1.2 Naturaleza del Enlace de halógeno

En la actualidad nadie desconoce la importancia de los enlaces de hidrógeno en la definición de la estructura y función de las biomoléculas, en cambio, la importancia del enlace de halógeno recién está comenzando a ser reconocida en contextos biológicos. Los halógenos muestran iguales preferencias de interacción con los átomos de nitrógeno y oxígeno de las estructuras proteicas, aunque con diferentes geometrías.

El enlace de halógeno, EX, puede definirse como una interacción de corto alcance del tipo R-X…Y, donde X es normalmente Yodo, Bromo, Cloro y en menor medida Flúor; Y al que llamaremos *aceptor de halógeno*, es una base de Lewis, normalmente una molécula/átomo con pares electrónicos libres o con una región en la que existe acumulación de densidad electrónica tales como átomos de oxígeno, nitrógeno, azufre o electrones π ; “R” al que denominaremos *dador de halógeno* es un grupo al que el halógeno está unido covalentemente y es por lo general un grupo/ átomo aceptor de electrones (Lommers et al., 1996; Metrangolo et al., 2005).

El EX puede considerarse análogo a los enlaces de hidrogeno EH (Legon, 1999; Metrangolo et al., 2005; Odd Hassel, 1970), con la diferencia que en el primero es un halógeno el que actúa como átomo puente entre un dador de halógeno y un aceptor de halógeno, (R-X…Y), mientras que en el segundo el átomo puente es un hidrógeno (R-H…Y). Ambos tipos de interacciones comparten muchas características físicas y químicas, en cuanto a direccionalidad y fortaleza.

En EX como en EH el dador y el aceptor son especies químicas fuertemente electronegativas y con pares electrónicos libres; resulta lógico que los átomos de hidrógeno con cargas parciales positivas interactúen atractivamente con átomos electronegativos o regiones de moléculas con concentración de carga electrónica, sin embargo la idea de que un átomo de halógeno covalentemente unido puede ser atraído por un sitio negativo de otra molécula puede ser poco probable, actualmente se comprobó que esto es posible debido a la formación del enlace covalente R-X, la distribución electrónica de X, de simetría esférica en el átomo aislado, se polariza en la dirección del enlace σ_{R-X} , originando en el sentido opuesto a R, una región de densidad electrónica disminuida. En el marco de la teoría de orbitales moleculares, el origen del agujero sigma se puede explicar por el hecho de que los tres pares de electrones no compartidos del átomo de halógeno, tienen una configuración aproximada $2s^2 2p^2 2p^2$, formando un cinturón de carga negativa en torno a la región ecuatorial al enlace σ_{R-X} , mientras que el tercer orbital p localizado a lo largo del eje R-X es distorsionado hacia el átomo R, para formar el enlace σ_{R-X} , dejando desprotegida electrónicamente la región axial del halógeno, opuesta al enlace σ_{R-X} ; por ejemplo, un átomo de Cloro en una molécula orgánica tiene una configuración de $3s^2 3p^2_X 3p^2_Y 3p^1_Z$, donde el eje z coincide con la dirección del enlace C-Cl. Nota que la hibridación para halógenos y especialmente para halógenos más pesados no es significativa. El electrón en el orbital p_z se localiza principalmente en la región de enlace, lo que resulta en una falta de densidad de electrones en el lóbulo externo (no involucrado). Los pares de electrones en los orbitales p_x y p_y crean una región negativa con una contribución del orbital s, como se muestra en la Figura 1.3

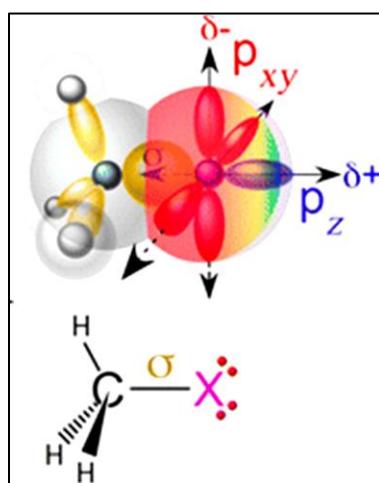


Figura 1.3: Disposición de los orbitales de un átomo de halógeno p_x (blanco), p_y (gris) y p_z (azul). Los electrones están localizados en p_x , p_y y la mitad interna del orbital p_z . La mitad exterior es deficiente en electrones, lo que se

expresa como un agujero. Los dos electrones de valencia restantes están localizados en un orbital s esféricamente simétrico (no dibujado). (Figura adaptada de Mendez, 2017) (Mendez et al., 2017)

Un importante avance en la comprensión de los EXs se dio con la introducción del concepto de agujero sigma; la polarización ocurre porque un electrón de la valencia del halógeno debe participar en el formación del orbital molecular, así, el enlace de halógeno, EX, surge de la asimetría de la distribución electrónica del átomo de halógeno puente, es decir, que la fuerza impulsora de esta interacción X···Y se atribuye a la existencia de un *agujero- σ* (Politzer, Murray, et al., 2007), este puede visualizarse como una región positiva (rojo en la Figura 1.4) localizada sobre el halógeno, en la dirección del enlace R-X y en sentido opuesto a R.

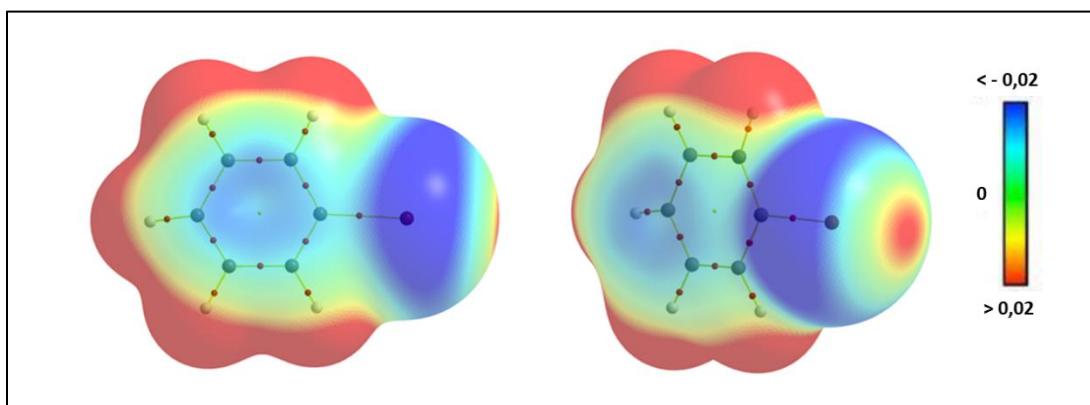


Figura 1.4: Mapa de potencial electrostático molecular montado sobre la densidad electrónica en Bromobenceno. La zona más positiva del potencial de superficie (incluido el orificio σ en el halógeno) está coloreado en rojo, mientras que las regiones más negativas se pueden observar en azul en u.a.

Además, se ha demostrado que los halógenos se polarizan en diferentes grados con una polarización creciente asociada con el aumento de la masa atómica esto significa que el Flúor es el menos polarizable y el Yodo es el más polarizable, en consecuencia, el agujero σ tiende a aumentar y ser más electropositivo en el sentido Cl<Br<I, sin embargo el átomo de Flúor forma enlace de halógeno en algunos casos (Duarte et al., 2015; Metrangolo et al., 2011). Las propiedades geométricas son las que permiten que los halógenos se polaricen, por ejemplo, si una base de Lewis, como el oxígeno, se acerca a un halógeno unido covalentemente, el ángulo de acercamiento más favorable sería lineal con respecto al enlace σ (C — X ... O), porque el oxígeno cargado negativamente interactúa con la zona electrostática positiva del halógeno. El enfoque más desfavorable sería perpendicular al agujero σ , porque el oxígeno cargado negativamente será repelido por el anillo electronegativo alrededor el halógeno en la zona ecuatorial. Además, se ha demostrado que el radio efectivo del halógeno se acorta en un ángulo

de aproximación lineal (Politzer et al., n.d.; Politzer, Murray, et al., 2007) por lo que las distancias de interacción entre halógeno y oxígeno son más cortas de que la suma de las distancias de sus radios.

Podemos decir que, en los EXs, se destacan algunas características sobresalientes que los hacen únicos para determinadas aplicaciones, tales como:

- ⊕ Direccionalidad: presentan una mayor direccionalidad que los EH, debido a que el ángulo formado entre los átomos que intervienen en este tipo de interacciones (R-X---Y) es cercano a 180º. Las estructuras biomoleculares indican una fuerte preferencia por un enfoque casi lineal del átomo aceptor hacia la corona electropositiva del agujero σ , con una caída significativa cuando el aceptor se acerca al punto de cruce entre los potenciales electrostáticos positivos y negativos ($\theta_1 \approx 140^\circ$). La geometría del enlace X en términos del ángulo de aproximación del halógeno hacia el átomo aceptor (θ_2) muestra que, en su mayor parte, el agujero σ es atraído a electrones no enlazantes del aceptor, (Auffinger et al., 2004) con $\theta_2 \approx 120^\circ$. Las excepciones para los enlaces X en entornos biológicos son cuando se encuentran electrones π deslocalizados disponibles, por ejemplo, en la cadena lateral de un aminoácido aromático, allí se encuentran dirigidos perpendicularmente al anillo aromático (Nyburg, 1979).
- ⊕ Fortaleza: Generalmente los EXs son más débiles que los EHs, pero se han encontrado excepciones en las cuales la energía de enlace de los primeros es mayor. Por ejemplo, Regier Voth et al. (A. P. Voth, 2007) han demostrado en una molécula biológica con un enlace de halógeno entre un uracilo bromado y un oxígeno fosfato, cuando una interacción Br---O se dirige en directa competencia con un EH clásico, la primera es aproximadamente entre 2-5 kcal/mol más fuerte que su análogo formado por hidrógeno.
- ⊕ Hidrofobicidad: Los átomos de halógeno son típicamente considerados como residuos hidrofóbicos. De hecho, el sitio de un donante típico de EX (p. ej., un átomo I o Br) es notablemente menos hidrófilo que un sitio típico de donante de EH (por ejemplo, un grupo OH o un grupo NH). Esta diferencia entre los dos sitios donantes produce muchas consecuencias y aplicaciones útiles. Por ejemplo, la absorción de un medicamento y distribución requiere que el compuesto cruce la membrana de las células, los átomos de halógeno se

introducen con frecuencia para aumentar su lipofilia y así atravesar dichas membranas. Por otro lado, la administración oral de una droga a menudo requiere que el compuesto tenga una buena solubilidad en agua, y la presencia de $-OH$ y $-NH_2$ como sustituyentes pueden ser útiles para este propósito.

- Dimensiones del átomo dador: los radios de van der Waals (Bondi, 1964) de los átomos de halógeno (1,47, 1,75, 1,85 y 1,98 Å para F, Cl, Br e I, respectivamente) son más grandes que el radio del hidrógeno (1.20 Å), y esto provoca diferencias en algunas propiedades de las interacciones que involucran estos átomos; EX es, en general, más sensible al impedimento estérico. Por ejemplo, en la cadena infinita formada por 1,4-diyodotetrafluorobenceno con 4,4'- y 2,2'-bipiridina, las distancias C-I ... N son 2.864 y 3.158 Å, respectivamente; cuando la 2,4'-bipiridina forma cristales heteroméricos con el mismo donante EX, solo los 4 nitrógenos se unen con el halógeno y se forman trímeros en donde la distancia C-I ... N es 2.821 Å. (Messina et al., 2001; Syssa-Magale et al., 2005)
- Polarizabilidad: El efecto más directo en la energía, además de la direccionalidad del EX es la polarizabilidad del halógeno, que sigue la serie F < Cl < Br < I, cuanto más polarizable y menos electronegativo es el elemento, más positivos son los agujeros σ . Los electrones de Flúor no forman tan fácilmente el agujero σ ; ya que no hay una desocupación significativa del orbital atómico p_z opuesto y, en consecuencia, F muestra un mínimo agujero σ (Murray et al., 2010; Politzer, Lane, et al., 2007). Por lo tanto, es generalmente considerado como un donante pobre de EX, excepto en los casos en que hay una muy fuerte capacidad de atracción de electrones de parte de la molécula a la que está unida (Duarte et al., 2010; Egner et al., 2005).
- Versatilidad: la fortaleza de la interacción puede modificarse con relativa facilidad sustituyendo el átomo o grupo dador por otro de distinta electronegatividad, o bien sustituyendo grupos en la base de Lewis con la que el halógeno interacciona. Por lo tanto, la acidez del donante y la basicidad del aceptor se combinan para definir en conjunto el potencial estabilizador de los EX. Estudios recientes han demostrado que la sustitución de grupos químicos diferentes en los anillos aromáticos de estos complejos tiene un gran efecto en el comportamiento interactivo del halógeno y por lo tanto podría ser utilizado en el diseño de nuevos productos farmacéuticos (Riley et al., 2009). Sin duda esta gran versatilidad es la característica más distintiva y sumamente útil del EX para su manipulación en el diseño de fármacos, ya que es fácilmente modulable,

“tunneable” por la modificación de su entorno molecular. Además, se debe tener en cuenta que el EX contribuye favorablemente a la estabilidad de los complejos ligando-blanco molecular, participando de *un intrincado conjunto de interacciones* (E. Angelina et al., 2015), debido a su capacidad dual de dador y acceptor de electrones.

1.3 Aplicaciones de los Enlaces de halógeno

Los halógenos se han convertido en una herramienta importante del diseño molecular y su aplicación en los diversos campos de la ciencia ha aumentado drásticamente, principalmente para el desarrollo de una variedad de nuevos materiales. Actualmente, se encuentra en estudio el impacto que los halógenos pueden tener en el ámbito bioquímico y biomédico, específicamente donde se pueden usar para diseñar nuevas terapias contra blancos clínicamente significativos.

A continuación, se describen algunas de sus aplicaciones en diferentes áreas de la química.

- Ingeniería Cristalina: es un área interdisciplinaria cuyo objetivo es el desarrollo de nuevas estructuras cristalinas y ensambles moleculares con propiedades fisicoquímicas específicas. Desiraju definió la Ingeniería Cristalina como *“el entendimiento de interacciones intermoleculares en el contexto del empaquetamiento cristalino y en la utilización de dicho entendimiento en el diseño de nuevos sólidos con propiedades físicas y químicas deseadas”* (G. R. Desiraju, 1996). En un principio se centró la atención en interacciones como enlaces de hidrógeno e interacciones iónicas, al profundizar el conocimiento de las herramientas brindadas por los enlaces halógeno se los comenzó a utilizar en este campo. La estrategia más empleada en la ingeniería cristalina para la obtención de nuevas estructuras, consiste en emplear moléculas (componentes estructurales) con grupos dadores y/o atractores de densidad electrónica, que controlen la organización supramolecular posterior. En este sentido, la capacidad de los halógenos para actuar como especies electrofílicas capaces de interactuar con bases de Lewis, ha despertado un gran interés en este campo. Los halógenos son átomos que se sitúan generalmente en la periferia de las moléculas orgánicas, y por tanto se encuentran en la posición ideal para generar interacciones moleculares, en las que el propio halógeno funciona como una especie aceptora de electrones. Mediante estos enlaces, los componentes estructurales (“building blocks”) quedan unidos a través de

interacciones no covalentes dirigidas linealmente en el espacio y permiten a su vez que la geometría molecular ordenada que poseen, determine la disposición macromolecular posterior. Dichos “building blocks”, han sido empleados para la preparación de cristales líquidos o estructuras microporosas, en electromagnetismo o en óptica no lineal. De sus propiedades la elevada direccionalidad y especificidad son las más importantes para este tipo de aplicaciones, es así como Resnati y Metrangolo las han aprovechado para formar diferentes tipos de estructuras moleculares, como 1,4-diyodotetrafluorobenceno, que ofrece dos sitios colineales, a los átomos de I, que actúan como fuertes dadores de EX.

- ✚ Diseño de nanomateriales: Yamamoto y col. (Yamamoto et al., 2008) han utilizado el enlace de halógeno como una herramienta para formar un nanocable supramolecular con un exterior aislante. La vaina aislante se constituye de enlaces de halógeno formado por aniones haluro como aceptores y moléculas neutras como tetrayodo etileno como dador de halógeno.
- ✚ Utilización en la química de los polímeros: Recientemente Lauher y col. (Lauher et al., 2008) han logrado, la polimerización de simples diyododienos conjugados usando la química del huésped–anfitrión. Los anfitriones oxalamidas alinean a las moléculas de diyododieno, de manera que los predisponen a la polimerización, ya sea espontáneamente o por irradiación. Las moléculas huéspedes se mantienen en su lugar por medio de EHs en algunos casos y por EXs en otros. Liang Luo y col. han logrado valerse de los EXs para sintetizar polímeros determinados. Por ejemplo, han logrado la preparación, el aislamiento, y la caracterización completa del poli(diyododiacetileno) o PIDA. El PIDA, es un polímero conjugado que contiene al poli(diacetileno) PDA como estructura central, con átomos de yodo como sustituyentes. Los monómeros de diyodobutadieno se pueden alinear en el estado sólido con los anfitriones bis(nitrilo)oxalamida por puentes de hidrógeno entre los grupos oxalamida y por medio de EXs entre los nitrilos y yodoalquinos. Como resultado comienzan a formarse cristales de color azul pálido, que se vuelven brillantes y de color cobrizo a medida que la polimerización progresó. El desarrollo de esta metodología de cristalización mejoró enormemente el rendimiento en la obtención del PIDA (alrededor del 50% de rendimiento), permitiendo la caracterización completa del polímero mediante difracción de rayos X, y las espectroscopías de resonancia magnética nuclear del ^{13}C , Raman y electrónica.

Aplicaciones Bioquímicas y Farmacéuticas: se han publicado resultados de investigaciones que vinculan el enlace de halógeno en funciones de importancia biológica y médica tales como el plegado molecular, interacciones ligando receptor, proteína-proteína o proteína-ácidos nucleicos entre otros (Auffinger et al., 2004; A. P. Voth, 2007), debido al contacto de diversos grupos funcionales con halógenos. Dependiendo del entorno estructural y químico, las interacciones mediadas por halógenos pueden ser débil o fuerte. En el caso de algunos complejos ligando-proteína, los enlaces de halógeno son energéticamente y geométricamente comparables a la de los enlaces de hidrógeno. Los organohalógenos que están presentes en el medio ambiente, principalmente como subproductos de la producción comercial de productos químicos, se cree que conducen a efectos biológicos adversos en animales y humanos. Un ejemplo de una proteína transportadora de hormonas cuya capacidad de funcionar adecuadamente se ve obstaculizada por EXs formados por organohalogenados, es la transtiretina, que es la responsable del transporte de las hormonas tiroideas en el plasma sanguíneo (Brouwer et al., 2016; Ghosh et al., 2000). En estudios de unión in vitro con varios grupos de compuestos organohalogenados se ha demostrado que los bifenilos policlorados y bifenilos polibromados son capaces de unirse a la transtiretina, al menos tan fuerte como el ligando natural tiroxina (Lans et al., 1993). La unión de estas moléculas de bifenilo halogenadas es atribuible, al menos en parte, a la presencia de EX (Ghosh et al., 2000).

Como consecuencia de su capacidad para formar EXs, los haloalcanos son también eficaces para romper las interacciones típicas de EHs del tipo O-H…O, N-H…N, N-H…O=C. Este aspecto tiene consecuencias de largo alcance en la actividad biológica de los xenobióticos halogenados. Curiosamente, los anestésicos halofluorocarbonados muestran un paralelismo sorprendente entre la potencia de rompimiento de EHs y la potencia anestésica (Di Paolo & Sandorfy, 1974). Los anestésicos Halofluorocarbonados, que contienen un hidrógeno ácido (halotano, metoxiflurano, enflurano), pueden dar lugar tanto a EHs como a EXs, y la importancia relativa de los dos fenómenos varía de un anestésico a otro.

1.4 Enlaces de halógeno en entornos Biológicos

Ferreira de Freitas y col en un revisión reciente sobre diferentes tipos de interacciones relevantes en proteína-ligando en química medicinal (Ferreira De Freitas & Schapira, 2017) los autores consideran que los enlaces de halógeno se suman al conjunto de interacciones favorables en el reconocimiento molecular y señalan que los EX en algunos casos pueden conducir a significativas ganancias de afinidad, además de facilitar la metabolización de los medicamentos. . Las fortalezas de los enlaces de halógeno se pueden evaluar teóricamente a través de cálculos cuánticos de modelos químicos. Estos estudios también están acompañados de estudios experimentales donde el efecto de la sustitución de hidrógeno por halógeno, sobre la afinidad de unión o binding, se ha evaluado sistemáticamente

Las primeras revisiones sobre contactos de Br--O en sistemas biológicos correspondieron a estructuras del inhibidor IDD594 con la aldolasa reductasa (AR) (Howard et al., 2004; Muzet et al., 2003) donde fue posible observar claramente cómo el Bromo del inhibidor y el Oxígeno del hidroxilo del residuo Thr113 de la proteína formaban un contacto corto con una distancia de 2.97 Å y un ángulo C – Br …O (γ) de 152.8 °, satisfaciendo así sustancialmente los requisitos para un EX, curiosamente, cuando una mutación de un solo átomo era realizada al reemplazar Br con Cl, el valor IC₅₀ aumentó de 500 a 1300 nM, consistente con un papel decisivo de un EX en el evento de reconocimiento y unión asociado con la inhibición de la enzima. Este resultado también refleja el fortalecimiento de la interacción con la polarizabilidad del átomo de halógeno (Br > Cl) (Duarte et al., 2012).

Los requisitos geométricos radiales y angulares en el contexto biomolecular son similares a los que se emplean con frecuencia en otros campos para identificar EX. El requisito radial consiste en la identificación de un contacto corto entre el átomo de halógeno y el átomo base de Lewis asociado, con una distancia interatómica menor que la suma de sus respectivos radios de van der Waals y se considera un ángulo lineal C – X … Y, aunque se ha demostrado que el ángulo X … Y – C (X = donante EX; Y = aceptor EX) esencialmente depende de la distribución de orbitales ricos en electrones en la base de Lewis. Entre los EX encontrados en el PDB (Protein Data Bank), un ángulo X … Y – C cercano a 120 ° es el más comúnmente encontrado, que es consistente con una participación de los orbitales no enlazantes de los grupos C=O. Sin embargo, se produce un máximo local alrededor de 90°, particularmente para los átomos de halógeno más pesados, (Auffinger et al., 2004) que puede estar asociado con interacciones que involucran orbitales π

de la base de Lewis, esta geometría se observa típicamente cuando EX y los donantes EH comparten un aceptor común, lo que resulta en interacciones $X \cdots O \cdots H$ de disposición ortogonal, que es observado en las estructuras secundarias de láminas beta y hélices alfa, la ortogonalidad entre las dos interacciones ha sido demostrada en estudios experimentales (Vasylyeva et al., 2014) y teóricos (A. R. Voth et al., 2009) también en pequeñas moléculas como péptidos, siendo no solo geométrica sino también química, ya que la formación de EX no altera los ensamblajes impulsados por EH.

Debido a la alta polarizabilidad de los halógenos más pesados, la fuerza del EX se puede ajustar introduciendo sustituyentes aceptores de electrones en una posición dada (Riley et al., 2009, 2011). Esta versatilidad de los EX los hace muy útiles para su manipulación en el diseño de fármacos (Wilcken et al., 2013). Más allá del efecto inductivo estándar, Carlsson et al. han demostrado que la adición de un donante de enlace H (incluidos OH, SH o NH₂) junto a un halógeno mejora el potencial del EX a través de una relación sinérgica (Carlsson et al., 2018). En un estudio de ingeniería de proteínas, encontraron que reemplazar una tirosina en la posición 18 (Y18) de la lisozima T4 con m-clorotirosina (mCIY) aumentó la estabilidad térmica de la enzima en 3 kcal/mol. La estructura cristalina de la construcción mCIY18-T4L reveló que el cloro de mCIY forma un enlace X con un oxígeno carbonilo cerca del sitio activo. A pesar de ser una interacción típicamente débil, el enlace X Cl \cdots O se ve significativamente mejorado por la polarización del átomo de Cl a través de un enlace de hidrógeno intramolecular Cl \cdots H con el sustituyente OH adyacente, lo que explica el aumento observado en la proteína estabilidad. En una revisión reciente de la base de datos estructural de Cambridge (CSD) se han encontrado muchos ejemplos de este tipo de EX optimizados con EH (HBeXB) al buscar compuestos aromáticos con Cl, Br o I en posición orto a sustituyentes OH o NH₂ (Riel et al., 2019).

Debido a la prevalencia de los EXs con oxígenos de la cadena principal de las proteínas, aquí se muestran algunos ejemplos visualizando separadamente los contactos con diferentes tipos de cadenas laterales (Figura 1.5). Examinando en el PDB se pueden observar ejemplos de enlaces de halógeno con el oxígeno en serina, treonina o tirosina; el azufre en metionina o cisteína. Algunos estudios han informado contactos favorables de X $\cdots\cdots\pi$. Las interacciones favorables con los sistemas π (fenilalanina, tirosina, triptófano e histidina) son posibles de encontrar en el PDB pero no insignificantes estadísticamente. Se necesitan más conocimientos sobre la dependencia geométrica de esta interacción para facilitar su estudio.

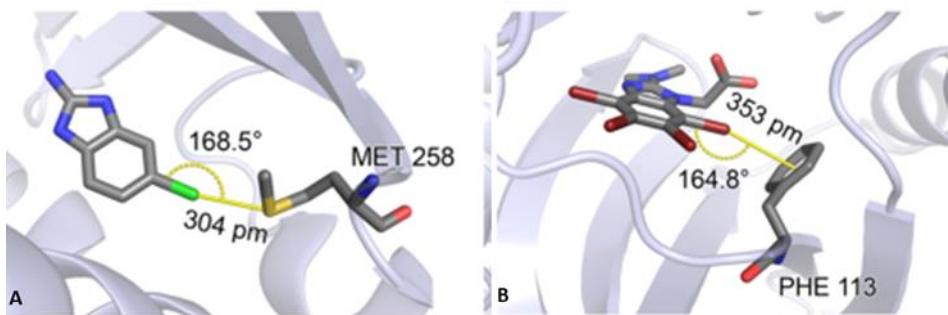


Figura 1.5: Ejemplos seleccionados de EXs en complejos proteína-ligandos A) 3KR1: compuesto formando un contacto Cl...S con el azufre de Met 258 en hPNMT y B)2KXH: compuesto formando contacto Br...π con el anillo de la Phe113 en CK2 quinasa. Tomados de Wickens et al (Wilckens,2012)

Para definir la energía de los EX en biomoléculas se requiere un enfoque computacional más preciso, se debe realizar un alto nivel de cálculos cuánticos (QM) sobre las estructuras de estos complejos; sin embargo, estos son engorrosos y, en la ausencia de estructuras de cristalinas de alta resolución acarrean errores. Lo que plantea la necesidad de un enfoque computacional que incorpore el enlace X en la mecánica molecular actual (MM), y un conjunto de datos experimentales para ayudar a validar tanto los enfoques QM como MM. Debido a este inconveniente hay muy pocas formas directas de medir las energías de los enlaces X, sin embargo con moléculas pequeñas las interacciones entre complejos de geometría conocida se pueden determinar midiendo la termodinámica de fusión del complejo(Corradi et al., 2000), de este modo se ha aplicado un enfoque similar para correlacionar directamente las energías de EX relativos a EH en cristales individuales de ADN, en este sistema, las energías se estimaron inicialmente a través de un ensayo cristalográfico (Figura1.6), en el cual el potencial estabilizador de un EX de bromo compitió directamente contra el de un EH en relaciones 1: 1 y 1: 2 de enlace X a H. En este ensayo, se observaron dos geometrías específicas, con una interacción Br ... O $^{-1/2}$ más corta siendo la energía $\sim 5\text{kcal/mol}$ y para la interacción más larga fue de $\sim 2\text{kcal/mol}$, más estabilizadora que la del EH en este sistema(A. P. Voth, 2007).

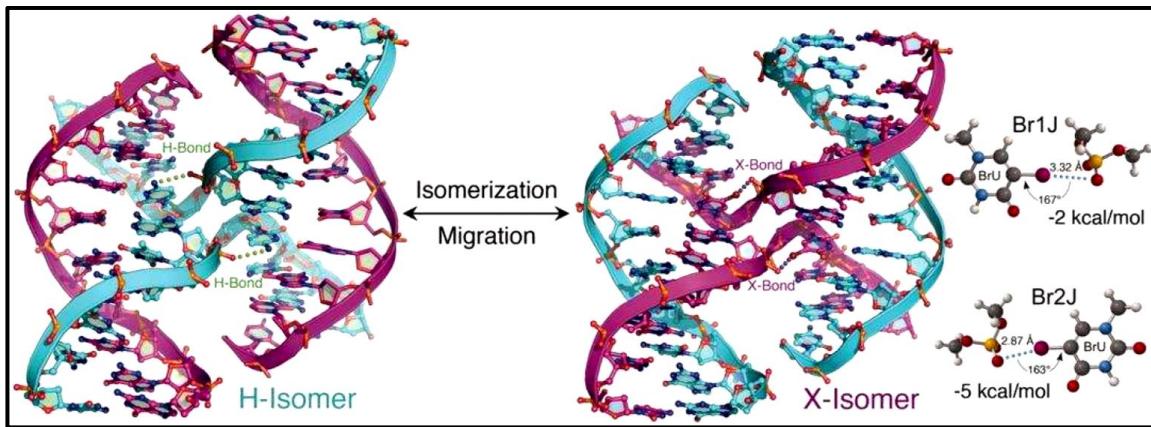


Figura 1.6: Unión de cuatro cadenas de ADN como ensayo competitivo para en las energías de EH frente al EX. Una unión de cuatro hebras puede isomerizarse para colocar bases de citosina para formar enlaces H estabilizadores (hebras celestes) o bases de 5-bromouracilo (BrU) para formar enlaces X estabilizadores (hebras magenta), un giro agudo produce el cruce de la unión, adoptando las formas de isómero H y isómero X, respectivamente. En el isómero X, las dos interacciones posibles $\text{BrU} \cdots \text{OPO}_3^{-1}$ tienen energías que son 2-5 kcal/mol más estabilizadoras que las enlaces H competitivos. (Figura tomada de Scholfield et.al (Scholfield et al, 2013))

Actualmente, en el análisis de estructuras ligando-proteína se utilizan los valores de afinidad, K_i o IC_{50} , que reemplazan indirectamente a los valores de energía, siendo útiles particularmente para comparar las afinidades de los inhibidores halogenado y no halogenado con idénticos dominios proteicos.

En este contexto, debido a sus propiedades los halógenos son incorporados en bibliotecas de cribado utilizadas para buscar compuestos líderes como inhibidores potenciales contra blancos terapéuticos importantes (Lu et al., 2012; Lu, Wang, et al., 2009; A. P. Voth, 2007; Wilcken et al., 2012; Zaldini Hernandes et al., 2010) donde la presencia de un EX puede afectar la afinidad droga-proteína. Por ejemplo, existe un interés particular en las proteínas quinasas por su potencial utilidad en el tratamiento de enfermedades crónicas, enfermedades inflamatorias, cardiovasculares y neurodegenerativas, infecciones virales, y posiblemente incluso en cáncer; varios complejos en donde EX es responsable de la unión de moléculas pequeñas a estas proteínas han sido reportadas (Battistutta et al., 2007; Battistutta et al., 2005; De Moliner et al., 2003; Dobeš et al., 2011; Fedorov et al., 2011; Lu, Shi, et al., 2009; Sarno et al., 2011; A. P. Voth, 2007).

Los átomos de oxígeno del “backbone” de las proteínas pueden mantener el mismo patrón de enlaces de hidrógeno, con los aminoácidos del entorno aún después de la formación de un enlace de halógeno, esta situación se observa en el inhibidor de la transcriptasa inversa del HIV-1, el 3-yodo-4-fenoxipirimidona. La transcriptasa inversa es un enzima clave dado que el virus la utiliza para formar una nueva partícula viral. En el complejo con su inhibidor R221239 (Himmel et al., 2005) se observa un contacto de 3,4 Å entre un átomo de yodo del inhibidor y el átomo de oxígeno del backbone de Tyr 188, con un ángulo C – I … O de 161,6°. La estructura de este co-cristal muestra claramente la formación del enlace de halógeno, donde el oxígeno carbonílico se inclina suavemente hacia el yodo después de la unión con el inhibidor (figura 1.7), pero retiene su estructura nativa de interacciones de EH con Tyr118. Se subraya también que, la sustitución de este yodo con hidrógeno muestra una reducción de la afinidad inhibidor-proteína, de alrededor de 300 veces. Más importante aún, el enlace de halógeno puede ser utilizado para diseñar ligandos que superen la farmacoresistencia. El especial interés en este caso, proviene del hecho de que el contacto con el inhibidor no se forma con el residuo Tyr181, uno de los aminoácidos involucrados en las mutaciones más comunes del virus, que lo vuelven fármaco-resistente.

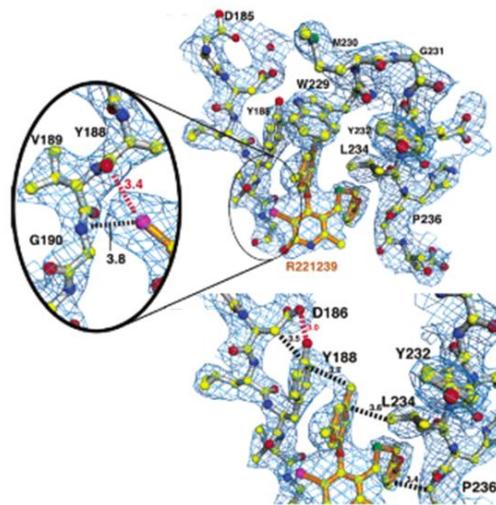


Figura 1.7 : Interacciones clave de R221239 en el bolsillo de NNRTI, el Yodo hace contacto con la cadena principal con de los residuos Tyr188-Gly190, incluida una posible interacción electrostática con el carbonilo de Tyr188 (Figura adaptada de Himmel, 2005 (Himmel et al., 2005))

1.5 Diseño de fármacos

Aunque el descubrimiento y desarrollo de medicamentos se ha hecho durante muchos años usando únicamente métodos experimentales, el aporte beneficioso de los halógenos puede mejorarse más aún con la ayuda de modelos moleculares computacionales utilizables para diseñar racionalmente nuevos medicamentos. Además, en la actualidad se busca con estas técnicas mejorar las propiedades de los medicamentos (como reducir efectos secundarios) o buscar actividades biológicas potenciales de compuestos químicos aislados de fuentes naturales (Medina-Franco, 2013); así como también tienen como objetivo reducir los costos y el tiempo, ya que se estima que el desarrollo de un fármaco tarda aproximadamente entre 10 y 15 años y se invierten millones de dólares, teniendo el mayor costo y tiempo las pruebas clínicas en humanos. El tiempo y costos tan elevados están asociados en gran medida a la gran cantidad de moléculas que fallan una o varias etapas (Figura 1.8)

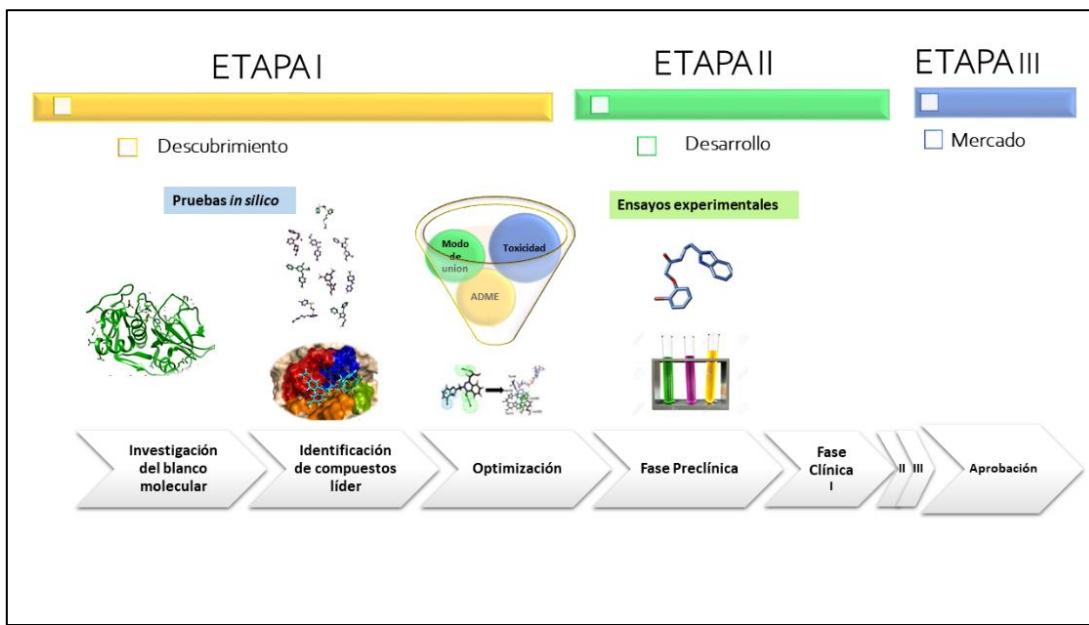


Figura 1.8: Etapas generales del desarrollo de un fármaco.

El desarrollo de fármacos es un proceso complejo el cual inicia con la identificación de compuestos que se unen a un blanco terapéutico o que muestran actividad biológica en un ensayo de tamizaje. Aquellas moléculas que muestran actividad biológica son llamadas *hits*. El siguiente paso es encontrar compuestos que tengan propiedades farmacéuticas atractivas, como baja toxicidad, solubilidad acuosa adecuada para administrarse vía oral, entre otras

propiedades farmacocinéticas. Tales compuestos son llamados *líderes o cabezas de serie*. Típicamente, los *hits* son encontrados por tamizaje de un número vasto de moléculas, mientras que los compuestos *líderes* son desarrollados a partir de los *hits* a través de modificaciones químicas (Saldívar-González et al., 2017). Considerando que el número de moléculas orgánicas que son capaces de ser obtenidas sintéticamente se encuentran entre 10^{20} y 10^{24} (Ertl, 2003), es evidente que su análisis sería muy complejo sin el uso de técnicas computacionales. La aplicación de modelos generados computacionalmente contribuye a entender los mecanismos de acción de los principios activos de los medicamentos o a mejorar las propiedades de los mismos.

El diseño de fármacos se realiza con la ayuda del modelado por computadora para predecir cómo la adición o eliminación de un átomo a un nuevo candidato a fármaco aumentará o disminuirá su afinidad al blanco; la mayoría de los átomos se modelan como objetos esféricos, con una distribución isotrópica de carga a través de su superficie. Sin embargo, se ha notado que modelar los átomos de halógeno de esta manera predice incorrectamente la formación de los EXs, y puede conducir al fracaso para diseñar compuestos más potentes. Surgió entonces la necesidad imperiosa de establecer correcciones que tomen en cuenta de las características de los halógenos. Para predecir adecuadamente cómo se comportan los enlaces de halógeno en sistemas bioquímicos se debe comprender cómo un EX contribuye a la estabilidad de un sistema, primero debemos entender las propiedades geométricas y energéticas de estos enlaces y parametrizarlos adecuadamente en los métodos de Modelado Molecular.

Un trabajo de revisión sobre los medicamentos aprobados recientemente por la Administración de Alimentos y Medicamentos (FDA) que contienen en su estructura, al menos un átomo de halógeno enlazado covalentemente, fue publicado por Tiz et. al (Tiz et al., 2022). Se evidencia que en los últimos años el número de moléculas halogenadas que llegan al mercado se conserva regularmente, y se destaca que 14 de las 50 moléculas aprobadas por la FDA, en ese año, contienen halógenos. Esto resalta el papel emergente de los halógenos y, en particular, del flúor y el cloro en la preparación de fármacos para el tratamiento de varias enfermedades como las infecciones virales, varios tipos de cáncer, enfermedades cardiovasculares, esclerosis múltiple, migraña y enfermedades inflamatorias (como la vasculitis).

Esto demuestra que por un lado, se ha realizado un gran esfuerzo en la búsqueda de nuevas terapias hecho a pesar de la pandemia de COVID-19, y por otro, el uso de halógenos se vuelve regular en la química médica. No solo cabe mencionar los compuestos sintéticos sino también los productos naturales halogenados, ya que muestran una amplia gama de actividades

biológicas (por ejemplo, antibacterianos, antifúngicos y anticancerígenos). Como ejemplo, la vancomicina es un medicamento clínico, antibiótico que contiene cloro obtenido de la bacteria *Streptomyces orientalis*, que se utiliza principalmente para tratar infecciones por *Staphylococcus aureus* resistente a la meticilina (MRSA) (Álvarez-Martínez et al., 2020).

1.5.1 Cloro en química medicinal

Resulta interesante conocer la distribución de halógenos en el campo del desarrollo de fármacos. Sorprendentemente, considerando los cuatro halógenos, el Cloro (Cl) se encuentra con más frecuencia entre las drogas aprobadas incluso con más frecuencia que el Flúor (F) y una de las áreas de más rápido crecimiento en química surge de la utilización de Cloro en química médica o química medicinal. Los elementos de azufre, cloro y flúor se colocan en 5to, 6to y 7mo lugar, respectivamente; después de C, H, O y N. El resto de los diez elementos claves en drogas aprobadas lo constituyen el Fósforo (P), Bromo (Br) y Yodo (I); es decir que compuestos que contienen Cl y F son los halógenos más destacados dentro de las drogas aprobadas (Cl > F > Br > I) (Fang et al., 2019)

Curiosamente, un estudio publicado algunos años atrás por Smith y colaboradores (Smith et al., 2014) sobre 233 drogas aprobadas por la US-FDA, (medicamentos) que contienen cloro, señala que 163 compuestos, aproximadamente el 73% de ellas contiene un solo átomo de cloro, el 23% contiene 2 átomos de cloro; el 2.6% de ellas poseen tres átomos de cloro; el 1.4% de ellas poseen 4 átomos de cloro y el 2,5% poseen 6 átomos de Cl. Sorpresivamente ninguna de las drogas aprobadas tiene 5 átomos de cloro. También entre las drogas analizadas, el 98% fueron monosustituidas (CCl), solamente 4 fueron disustituidas (CCl₂) y ninguna de ellas muestra un grupo tri-sustituido (CCl₃).

Fang y col. listan 139 drogas conteniendo cloro aprobadas por la FDA, que se encuentran en el mercado, desde 1949-2012, para distintas enfermedades y/o dolencias, es decir compuestos de uso clínico para enfermedades cardiovasculares, del sistema nervioso central, del metabolismo alimentario, enfermedades respiratorias, enfermedades infecciosas, enfermedades musculo esqueléticas, genito urinarias, drogas citoestáticas, hormonas, antiretrovirales, entre otras.

1.5.1.1 ¿Cuán importante es el cloro para mejorar las propiedades biológicas?

La presencia del átomo de cloro desempeña un papel fundamental en una serie de productos como los antibióticos clindamicina, vancomicina, cloranfenicol y griseofulvina. Con el transcurso del tiempo se ha encontrado empíricamente que la introducción de un átomo de cloro, en una o más posiciones específicas de una molécula biológicamente activa, puede

mejorar sustancialmente la actividad biológica intrínseca [6]. Las propiedades del enlace carbono-cloro (C-Cl) en los organoclorados han sido analizados por Henschler (Henschler, 1994). Sin embargo, en los productos químicos de bajo peso molecular investigados en ese análisis, la reactividad electrofílica del centro de carbono adyacente al átomo de cloro, que facilita el desplazamiento del cloro por (bio) nucleófilos, determina las propiedades biológicas observadas. El aumento de la lipofilia de toda la molécula por un sustituyente clorado conduce a una mayor partición de un compuesto clorado en la fase lipofílica de una membrana celular o en dominios lipofílicos de una proteína. Esto provoca una mayor concentración local del compuesto próximo o cercano a un sitio de destino biológico, pero, no necesariamente una mayor actividad biológica. El efecto más importante de un átomo de cloro no-reactivo en la actividad biológica de muchos compuestos proviene del cloro como sustituyente en un anillo aromático, en un anillo heteroaromático o en un resto olefínico. Las propiedades mencionadas anteriormente darán lugar a efectos electrónicos y/o estérico de los sustituyentes clorados y conducen a la atracción o la repulsión electrónica y/o interferencia estérica, con cualquier residuo de aminoácido que rodea la posición del átomo de cloro en el bolsillo de unión de la proteína. Esto a su vez puede causar un fortalecimiento o un debilitamiento de las interacciones con los contactos con los aminoácidos cercanos al cloro o en otras partes de la molécula activa. En cualquiera de los dos casos puede afectar la función de la proteína objetivo y causar un aumento o disminución de la actividad biológica. Sin embargo, en otros casos, un cloro sustituyente puede no tener un efecto primario específico sobre las propiedades biológicas de la molécula a la que está unido (Naumann, 2000).

En Thomson Reuters Pharma, el contenido de organohalógenados “para llegar a ser un medicamento” es de 34,6%, 36,4%, 29,8%, 36,4%, 33,1% y 25,5% en etapa de descubrimiento, fase clínica I, fase clínica II, fase clínica fase III, preinscritos y registrados y etapa de lanzamiento al mercado, respectivamente .

Se debe tener en cuenta que las moléculas en etapa de descubrimiento representan el estado actual de las acuerdos e intereses en el descubrimiento de fármacos, y las moléculas en fases clínicas representan las acuerdos e intereses en el descubrimiento de fármacos hace aproximadamente una década, mientras que la fase de lanzamiento recopila las drogas exitosas de las últimas décadas. Resulta valioso mencionar que, en la actualidad, hay más drogas halogenadas (34%) en pruebas clínicas y en etapa de pre-registración o de registro, que las drogas que están en el Mercado (~26%) mostrando que en la actualidad, la halogenación está siendo mucho más apreciada en el diseño de drogas.

Se espera que el estudio llevado a cabo en esta tesis sea útil para inspirar el diseño estructural y el desarrollo de medicamentos, más poderosos y menos tóxicos, basados en halógenos, contra las numerosas enfermedades existentes, muchas de las cuales causan la muerte.

Por otra parte, la unión entre el descubrimiento computacional de fármacos y el de los enlaces de halógeno en sistemas biológicos, ha proporcionado nuevos ímpetus y esperanzas para el descubrimiento de fármacos

1.6 Planteo del Problema y Objetivos

Como se expresó previamente, más del 40% de los principales medicamentos líderes en el mercado son compuestos halogenados, y se ha encontrado que los halógenos persisten durante todo el proceso de desarrollo del fármaco. Actualmente, existe una mayor apreciación de que los halógenos desempeñan un papel directo en la eficacia de ciertos fármacos a través de una interacción molecular entre densidades electrónicas que ahora se define como enlace de halógeno, (interacción hole-lump). Se reconoce en la bibliografía que la sustitución de X por H proporciona un aumento de más de 1.000 veces en la especificidad y afinidad de algunos inhibidores, hacia sus blancos moleculares; sin embargo, estos hechos fueron reconocidos posteriormente, **es decir no fueron tomados en cuenta en el diseño**. Nosotros nos focalizaremos en aplicar el concepto de enlace de halógeno y otras interacciones hole-lump en el descubrimiento y en la optimización de compuestos líder. La incorporación del EX en el diseño de fármacos basado en la estructura requiere modelos computacionales que tengan en cuenta la distribución anisotrópica de la carga electrónica y la forma no esférica de los halógenos covalentemente unidos, características distintivas que conllevan a geometrías de interacción altamente direccionales, y energías estabilizadoras. Este campo de rápido crecimiento impulsará aún más el desarrollo de herramientas computacionales, más precisas y eficientes, para acelerar la explotación de halógenos en química medicinal (Ford,2016).

Los beneficios de los halógenos en estas ciencias se pueden mejorar aún más con la ayuda de modelado por computadora para diseñar racionalmente nuevos medicamentos y materiales. Es evidente que obteniendo más información sobre los halógenos y cómo se comportan en los sistemas bioquímicos, sus aplicaciones serán aún más efectivas e impactantes. A pesar de que los halógenos se usan con frecuencia para desarrollar nuevos materiales, sus propiedades específicas y cómo se comportan aún se no se conocen completamente.

Como se ha descripto previamente el EX comparte con el EH muchas características y propiedades, lo que ha reforzado su valorización en la Química Medicinal principalmente para el diseño de fármacos, pero debido al inconveniente de la existencia del agujero- σ , es decir la distribución asimétrica de la densidad electrónica sobre el átomo de halógeno puente, supone un serio desafío a los actuales enfoques de modelado molecular (Docking, Dinámica Molecular, etc.) dado que estos métodos consideran a los átomos de halógeno como dotados de un potencial electrostático enteramente negativo y por ello fallan en describir correctamente los sistemas enlazados por halógeno, (ej. complejos ligando-proteína). Los métodos mecano-cuánticos en cambio, describen correctamente los EXs, sin embargo, estos cálculos se vuelven prohibitivos a medida que aumenta el número de átomos en el sistema biológico como también el número de moléculas a ser evaluadas, como en el “screening” virtual. Dentro del diseño de fármacos asistido por computadora, la utilización de métodos computacionalmente no-cuánticos “más baratos” resulta inevitable.

1.6.1 Objetivos generales

El objetivo general de este plan de trabajo es contribuir, desde la información provista por la Química Computacional al reconocimiento y factibilidad de las interacciones mediados por halógeno en el diseño de fármacos asistidos por computadora, (Computer Aided Drug Design (CADD)), mejorando además la descripción de los enlaces de halógeno, en los métodos de modelado molecular.

1.6.2 Objetivos particulares

1- Mejorar la descripción de los enlaces de halógeno en los métodos utilizados en el screening virtual de bibliotecas de compuestos frente a los blancos moleculares de interés.

- 1.1 Calibrar las funciones de puntuación y los métodos de docking y de Dinámica Molecular.
- 2- Entender/dilucidar el rol de los enlaces de halógeno en el conjunto de interacciones presentes en el sitio de anclaje del ligando al blanco molecular.
- 3- Analizar comparativamente el rol de los enlaces de halógeno, respecto de los enlaces de hidrógeno, en el sitio de anclaje.

4- Analizar las interacciones intra/intermoleculares mediante el mapeo de las densidades electrónicas (DE) en el sitio de “binding” utilizando un enfoque híbrido QM/MM-QTAIM.

5 -Proponer modificaciones en la estructura del ligando para aumentar el anclaje al blanco molecular.

5.1 Optimizar la estructura del/los compuesto/s líderes aprovechando las características estructurales, electrónicas y energéticas (ductilidad y maleabilidad) de las interacciones mediadas por halógeno, en el proceso de diseño de fármacos, CADD.

5.2 Afinar (ajustar), perfeccionar las estrategias de optimización del ligando basado en análisis de interacciones por QTAIM: QM/MM -QTAIM y DM/QMsemi-QMDFT-QTAIM.

Para resolver los interrogantes planteados se aplicarán estos métodos parametrizados adecuadamente para EXs, en el estudio de posibles inhibidores halogenados en dos blancos moleculares de elevado interés en la Química Medicinal:

i) receptores de Dopamina: D1, D2 y D3 que pertenece a la familia de Receptores Acoplados a Proteína G (GPCR) al momento de iniciar esta tesis no se disponía de una estructura 3D determinada experimentalmente de D2R, se ha construido un modelo por homología de dicho receptor utilizando como molde las estructuras cristalinas del receptor D3 de dopamina (D3R) y receptor β 2 adrenérgico de la forma activa e inactiva. Actualmente existen en el Protein Data Bank 3 estructuras cristalinas de este receptor.

ii) Cruzipain (Cz), la principal cisteína proteasa del *Tripanosoma cruzi*, parásito responsable del mal de Chagas. Actualmente existen depositadas en el Protein Data Bank 25 estructuras de complejos [Ligando-Cruzipaína], obtenidas por difracción de Rayos-X, varias de estas estructuras contienen halógeno.

1.7 Referencias:

- Alkorta, I., Blanco, F., Solimannejad, M., & Elguero, J. (2008). Competition of hydrogen bonds and halogen bonds in complexes of hypohalous acids with nitrogenated bases. *Journal of Physical Chemistry A*, 112(43), 10856–10863. <https://doi.org/10.1021/jp806101t>
- Alkorta, I., Rozas, I., & Elguero, J. (1998). Charge-transfer complexes between dihalogen compounds and electron donors. *Journal of Physical Chemistry A*, 102(46), 9278–9285. <https://doi.org/10.1021/jp982251o>
- Álvarez-Martínez, F. J., Barrajón-Catalán, E., & Micol, V. (2020). Tackling antibiotic resistance with compounds of natural origin: A comprehensive review. *Biomedicines*, 8(10), 1–30. <https://doi.org/10.3390/biomedicines8100405>
- Amezaga, N. J. M., Pamies, S. C., Peruchena, N. M., & Sosa, G. L. (2010). Halogen bonding: A study based on the electronic charge density. *Journal of Physical Chemistry A*, 114(1), 552–562. <https://doi.org/10.1021/jp907550k>
- Angelina, E., Andujar, S., Moreno, L., Garibotto, F., Párraga, J., Peruchena, N., Cabedo, N., Villecco, M., Cortes, D., & Enriz, R. D. (2015). 3-chlorotyramine acting as ligand of the D2 dopamine receptor. Molecular modeling, synthesis and D2 receptor affinity. *Molecular Informatics*, 34(1), 28–43. <https://doi.org/10.1002/minf.201400093>
- Angelina, E. L., Andujar, S. A., Tosso, R. D., Enriz, R. D., & Peruchena, Nlida M. (2014). Non-covalent interactions in receptor-ligand complexes. A study based on the electron charge density. *Journal of Physical Organic Chemistry*, 27(2), 128–134. <https://doi.org/10.1002/poc.3250>
- Auffinger, P., Hays, F. A., Westhoft, E., & Ho, P. S. (2004). Halogen bonds in biological molecules. *PNAS*, 101, 16789–16794. <https://doi.org/10.1073/pnas.0407607101>
- Ault, B. S., & Andrews, L. (1977). Infrared and Raman Spectra of the M+F3-Ion Paris and Their Mixed Chlorine-Fluorine Counterparts in Solid Argon. *Inorganic Chemistry*, 16(8), 2024–2028. <https://doi.org/10.1021/ic50174a040>
- Battistuta, R., Mazzorana, M., Cendron, L., Bortolato, A., Sarno, S., Kazimierczuk, Z., Zanotti, G., Moro, S., & Pinna, L. (2007). *The ATP-Binding Site of Protein kinase CK2 holds a positive electrostatic area and conserved water molecules*.

Battistutta, R., Mazzorana, M., Sarno, S., Kazimierczuk, Z., Zanotti, G., & Pinna, L. A. (2005). Inspecting the structure-activity relationship of protein kinase CK2 inhibitors derived from tetrabromo-benzimidazole. *Chemistry and Biology*, 12(11), 1211-1219. <https://doi.org/10.1016/j.chembiol.2005.08.015>

Bent, H. A. (1968). Structural chemistry of donor-acceptor interactions. *Chemical Reviews*, 68(5), 587-648. https://doi.org/10.1021/CR60255A003/ASSET/CR60255A003.FP.PNG_V03

Bondi, A. (1964). Van der waals volumes and radii. *Journal of Physical Chemistry*, 68(3), 441-451. <https://doi.org/10.1021/J100785A001>

Brouwer, A., Morse, D. C., Lans, M. C., Gerlienke Schuur, A., Murk, A. J., Klasson-Wehler, E., Bergman, Å., & Visser, T. J. (2016). Interactions of Persistent Environmental Organohalogens With the Thyroid Hormone System: Mechanisms and Possible Consequences for Animal and Human Health. <Http://Dx.Doi.Org/10.1177/074823379801400107>, 14(2), 59-84. <https://doi.org/10.1177/074823379801400107>

Brown, R. S., Slebocka-Tilk, H., Bennet, A. J., Bellucci, G., Bianchini, R., & Ambrosetti, R. (1990). Investigation of the Early Stages of Electrophilic Addition of Br₂ to Olefins. Kinetic Evidence for a Reversibly Formed Bromonium Ion in the Reaction of Br₂ with Tetraisobutyl lethylene. *Journal of the American Chemical Society*, 112(17), 6310-6316. https://doi.org/10.1021/JA00173A019/ASSET/JA00173A019.FP.PNG_V03

Carlsson, A. C. C., Scholfield, M. R., Rowe, R. K., Ford, M. C., Alexander, A. T., Mehl, R. A., & Ho, P. S. (2018). Increasing Enzyme Stability and Activity through Hydrogen Bond-Enhanced Halogen Bonds. *Biochemistry*, 57(28), 4135-4147. <https://doi.org/10.1021/acs.biochem.8b00603>

Cavallo, G., Metrangolo, P., Milani, R., Pilati, T., Priimagi, A., Resnati, G., & Terraneo, G. (2016). *The Halogen Bond*. <https://doi.org/10.1021/acs.chemrev.5b00484>

Clark, T., Hennemann, M., Murray, J. S., & Politzer, P. (2007). Halogen bonding: The σ -hole: Proceedings of "Modeling interactions in biomolecules II", Prague, September 5th-9th, 2005. In *Journal of Molecular Modeling* (Vol. 13, Issue 2, pp. 291-296). <https://doi.org/10.1007/s00894-006-0130-2>

Clark, T., Hennermann, M., Murray, J., & Politzer, P. (2005). Halogen bonding:the sigma hole. *J Mol Med*, 13.

Corradi, E., Meille, S. V., Messina, M. T., Metrangolo, P., & Resnati, G. (2000). Halogen bonding

versus hydrogen bonding in driving self-assembly processes. *Angewandte Chemie - International Edition*, 39(10), 1782–1786. [https://doi.org/10.1002/\(SICI\)1521-3773\(20000515\)39:10<1782::AID-ANIE1782>3.0.CO;2-5](https://doi.org/10.1002/(SICI)1521-3773(20000515)39:10<1782::AID-ANIE1782>3.0.CO;2-5)

Crihfield, A., Hartwell, J., Phelps, D., Walsh, R. B., Harris, J. L., Payne, J. F., Pennington, W. T., & Hanks, T. W. (2003). Organic Iodides 2003. *Growth (Lakeland)*.

De Moliner, E., Moro, S., Sarno, S., Zagotto, G., Zanotti, G., Pinna, L. A., & Battistutta, R. (2003). Inhibition of protein kinase CK2 by anthraquinone-related compounds: A structural insight. *Journal of Biological Chemistry*, 278(3), 1831–1836. <https://doi.org/10.1074/jbc.M209367200>

Desiraju, G., Ho, P. S., Kloo, L., Legon, A. C., Marquardt, R., Metrangolo, P., Politzer, P., Resnati, G., & Rissanen, K. (2013). Definition of the halogen bond (IUPAC Recommendations 2013)*. *Pure Appl. Chem.*, 85, 1711–1713. <https://doi.org/10.1351>

Desiraju, G. R. (1996). The supramolecular concept as a bridge between organic, inorganic and organometallic crystal chemistry. *Journal of Molecular Structure*, 374(1-3), 191–198. [https://doi.org/10.1016/0022-2860\(95\)08932-2](https://doi.org/10.1016/0022-2860(95)08932-2)

Di Paolo, T., & Sandorfy, C. (1974). Fluorocarbon anaesthetics break hydrogen bonds. *Nature* 1974 252:5483, 252(5483), 471–472. <https://doi.org/10.1038/252471a0>

Dobeš, P., Řezáč, J., Fanfrlík, J., Otyepka, M., & Hobza, P. (2011). Semiempirical quantum mechanical method PM6-DH2X describes the geometry and energetics of CK2-inhibitor complexes involving halogen bonds well, while the empirical potential fails. *Journal of Physical Chemistry B*, 115(26), 8581–8589. <https://doi.org/10.1021/jp202149z>

Duarte, D. J. R., Angelina, E. L., & Peruchena, N. M. (2012). On the strength of the halogen bonds: Mutual penetration, atomic quadrupole moment and Laplacian distribution of the charge density analyses. *Computational and Theoretical Chemistry*, 998, 164–172. <https://doi.org/10.1016/j.comptc.2012.07.019>

Duarte, D. J. R., De Las Vallejos, M. M., & Peruchena, N. M. (2010). Topological analysis of aromatic halogen/hydrogen bonds by electron charge density and electrostatic potentials. *Journal of Molecular Modeling*, 16(4), 737–748. <https://doi.org/10.1007/s00894-009-0558-2>

Duarte, D. J. R., Peruchena, N. M., & Alkorta, I. (2015). Double hole-lump interaction between halogen atoms. *Journal of Physical Chemistry A*, 119(16), 3746–3752. <https://doi.org/10.1021/jp511118s>

Dumas, J. M., Gomel, L., & Guerin, M. (1983). *Molecular Interactions Involving Organic Halides. The Chemistry of Functional Groups*. 983.

Egner, U., Krätzschmar, J., Kreft, B., Pohlenz, H. D., & Schneider, M. (2005). The target discovery process. *ChemBioChem*, 6(3), 468–479. <https://doi.org/10.1002/cbic.200400158>

Ertl, P. (2003). Cheminformatics Analysis of Organic Substituents: Identification of the Most Common Substituents, Calculation of Substituent Properties, and Automatic Identification of Drug-Like Bioisosteric Groups. *ChemInform*, 34(21), 374–380. <https://doi.org/10.1002/chin.200321198>

Eskandari, K., & Mahmoodabadi, N. (2013). Pnicogen bonds: A theoretical study based on the Laplacian of electron density. *Journal of Physical Chemistry A*, 117(48), 13018–13024. <https://doi.org/10.1021/jp4098974>

Eskandari, K., & Zariny, H. (2010). Halogen bonding: A lump-hole interaction. *Chemical Physics Letters*, 492(1–3), 9–13. <https://doi.org/10.1016/J.CPLETT.2010.04.021>

Fang, W. Y., Ravindar, L., Rakesh, K. P., Manukumar, H. M., Shantharam, C. S., Alharbi, N. S., & Qin, H. L. (2019). Synthetic approaches and pharmaceutical applications of chloro-containing molecules for drug discovery: A critical review. *European Journal of Medicinal Chemistry*, 173, 117–153. <https://doi.org/10.1016/j.ejmech.2019.03.063>

Fedorov, O., Huber, K., Eisenreich, A., Filippakopoulos, P., King, O., Bullock, A. N., Szklarczyk, D., Jensen, L. J., Fabbro, D., Trappe, J., Rauch, U., Bracher, F., & Knapp, S. (2011). Specific CLK inhibitors from a novel chemotype for regulation of alternative splicing. *Chemistry and Biology*, 18(1), 67–76. <https://doi.org/10.1016/j.chembiol.2010.11.009>

Ferreira De Freitas, R., & Schapira, M. (2017). A systematic analysis of atomic protein-ligand interactions in the PDB. *MedChemComm*, 8(10), 1970–1981. <https://doi.org/10.1039/c7md00381a>

Ghosh, M., Meerts, I. A. T. M., Cook, A., Bergman, A., Brouwer, A., & Johnson, L. N. (2000). Structure of human transthyretin complexed with bromophenols: a new mode of binding. *Acta Crystallographica Section D. Biological Crystallography*, 56(9), 1085–1095. <https://doi.org/10.1107/S0907444900008568>

Guthrie, F. (1863). On the iodide of iodammonium. *Journal of the Chemical Society*, 16, 239–244. <https://doi.org/10.1039/JS8631600239>

Hassel, O., Hvoslef, J., Vihovde, E. H., & Sörensen, N. A. (1954). The Structure of Bromine 1,4-Dioxanate. In *Acta Chemica Scandinavica* (Vol. 8, pp. 873–873). <https://doi.org/10.3891/acta.chem.scand.08-0873>

Hassel, O., Strømme, K. O., Hammarsten, E., Hedén, C.-G., Malmgren, B., & Palmstierna, H. (1959). Crystal Structure of the Addition Compound Benzene-Chlorine (1:1). In *Acta Chemica Scandinavica* (Vol. 13, pp. 1781–1786). <https://doi.org/10.3891/acta.chem.scand.13-1781>

Henschler, D. (1994). Toxizität chlororganischer Verbindungen: Einfluß der Einführung von Chlor in organische Moleküle. *Angewandte Chemie*, 106(19), 1997–2012. <https://doi.org/10.1002/ANGE.19941061905>

Himmel, D. M., Das, K., Clark, A. D., Hughes, S. H., Benjahad, A., Oumouch, S., Guillemont, J., Coupa, S., Poncelet, A., Csoka, I., Meyer, C., Andries, K., Nguyen, C. H., Grierson, D. S., & Arnold, E. (2005). Crystal structures for HIV-1 reverse transcriptase in complexes with three pyridinone derivatives: A new class of non-nucleoside inhibitors effective against a broad range of drug-resistant strains. *Journal of Medicinal Chemistry*, 48(24), 7582–7591. <https://doi.org/10.1021/jm0500323>

Howard, E. I., Sanishvili, R., Cachau, R. E., Mitschler, A., Chevrier, B., Barth, P., Lamour, V., Van Zandt, M., Sibley, E., Bon, C., Moras, D., Schneider, T. R., Joachimiak, A., & Podjarny, A. (2004). Ultrahigh resolution drug design I: Details of interactions in human aldose reductase-inhibitor complex at 0.66 Å. *Proteins: Structure, Function and Genetics*, 55(4), 792–804. <https://doi.org/10.1002/prot.20015>

Lans, M. C., Klasson-Wehler, E., Willemse, M., Meussen, E., Safe, S., & Brouwer, A. (1993). Structure-dependent, competitive interaction of hydroxy-polychlorobiphenyls, -dibenz-p-dioxins and -dibenzofurans with human transthyretin. *Chemico-Biological Interactions*, 88(1), 7–21. [https://doi.org/10.1016/0009-2797\(93\)90081-9](https://doi.org/10.1016/0009-2797(93)90081-9)

Lauher, J. W., Fowler, F. W., & Goroff, N. S. (2008). Single-crystal-to-single-crystal topochemical polymerizations by design. *Accounts of Chemical Research*, 41(9), 1215–1229. <https://doi.org/10.1021/ar8001427>

Legon, A. C. (1999). Prereactive complexes of dihalogens XY with Lewis bases B in the gas phase: A systematic case for the halogen analogue B···XY of the hydrogen bond B···HX. *Angewandte Chemie - International Edition*, 38(18), 2686–2714. [https://doi.org/10.1002/\(SICI\)1521-3773\(19990917\)38:18<2686::AID-ANIE2686>3.0.CO;2-6](https://doi.org/10.1002/(SICI)1521-3773(19990917)38:18<2686::AID-ANIE2686>3.0.CO;2-6)

Legon, A. C. (2010). The halogen bond: An interim perspective. *Physical Chemistry Chemical Physics*, 12(28), 7736–7747. <https://doi.org/10.1039/c002129f>

Lehn, J. (1988). *Supramolecular Chemistry-Scope and Perspectives Molecules, Supramolecules and Molecular Devices*.

Loc Nguyen, H., Horton, P. N., Hursthouse, M. B., Legon, A. C., & Bruce, D. W. (2004). Halogen Bonding: A New Interaction for Liquid Crystal Formation. *J. AM. CHEM. SOC*, 126, 16–17. <https://doi.org/10.1021/ja036994l>

Lommers, J. P. M., Stone, A. J., Taylor, R., & Allen, F. H. (1996). The nature and geometry of intermolecular interactions between halogens and oxygen or nitrogen. *Journal of the American Chemical Society*, 118(13), 3108–3116. <https://doi.org/10.1021/ja953281x>

Lu, Y., Liu, Y., Xu, Z., Li, H., Liu, H., & Zhu, W. (2012). Halogen bonding for rational drug design and new drug discovery. *Expert Opinion on Drug Discovery*, 7(5), 375–383. <https://doi.org/10.1517/17460441.2012.678829>

Lu, Y., Shi, T., Wang, Y., Yang, H., Yan, X., Luo, X., Jiang, H., & Zhu, W. (2009). Halogen bonding - A novel interaction for rational drug design? *Journal of Medicinal Chemistry*, 52(9), 2854–2862. <https://doi.org/10.1021/jm9000133>

Lu, Y., Wang, Y., Xu, Z., Yan, X., Luo, X., Jiangs, H., & Zhu, W. (2009). C-X···H contacts in biomolecular systems: How they contribute to protein - ligand binding affinity. *Journal of Physical Chemistry B*, 113(37), 12615–12621. <https://doi.org/10.1021/jp906352e>

Medina-Franco, J. L. (2013). ADVANCES IN COMPUTATIONAL APPROACHES FOR DRUG DISCOVERY BASED ON NATURAL PRODUCTS. *Rev Latinoamericana de Quimica*, 41. <http://pkuxxj.pku.edu.cn/>

Mendez, L., Henriquez, G., Sirimulla, S., & Narayan, M. (2017). Looking back, looking forward at halogen bonding in drug discovery. *Molecules*, 22(9), 22–25. <https://doi.org/10.3390/molecules22091397>

Messina, M. T., Metrangolo, P., Panzeri, W., Pilati, T., & Resnati, G. (2001). Intermolecular recognition between hydrocarbon oxygen-donors and perfluorocarbon iodine-acceptors: The shortest O···I non-covalent bond. *Tetrahedron*, 57(40), 8543–8550. [https://doi.org/10.1016/S0040-4020\(01\)00794-3](https://doi.org/10.1016/S0040-4020(01)00794-3)

Metrangolo, P., Murray, J. S., Pilati, T., Politzer, P., Resnati, G., & Terraneo, G. (2011). Fluorine-centered halogen bonding: A factor in recognition phenomena and reactivity. *Crystal Growth and Design*, 11(9), 4238–4246. <https://doi.org/10.1021/cg200888n>

Metrangolo, P., Neukirch, H., Pilati, T., & Resnati, G. (2005). *Halogen Bonding Based Recognition Processes: A World Parallel to Hydrogen Bonding* †. <https://doi.org/10.1021/ar0400995>

Metrangolo, P., Pilati, T., & Resnati, G. (2006). Halogen bonding and other noncovalent interactions involving halogens: A terminology issue. *CrystEngComm*, 8(12), 946–947. <https://doi.org/10.1039/b610454a>

Moorthy, J. N., Venkatakrishnan, P., Mal, P., Dixit, S., & Venugopalan, P. (2003). Crystal engineering: Identification of a unique supramolecular synthon based on C=O···X interaction in halogen-substituted aromatic carboxaldehydes. *Crystal Growth and Design*, 3(4), 581–585. <https://doi.org/10.1021/cg034001p>

Murray, J. S., Riley, K. E., Politzer, P., & Clark, T. (2010). Directional weak intermolecular interactions: σ -hole bonding. *Australian Journal of Chemistry*, 63(12), 1598–1607. <https://doi.org/10.1071/CH10259>

Muzet, N., Guillot, B., Jelsch, C., Howard, E., & Lecomte, C. (2003). Electrostatic complementarity in an aldose reductase complex from ultra-high-resolution crystallography and first-principles calculations. In *PNAS* (Vol. 100, Issue 15). www.pnas.org/cgi/doi/10.1073/pnas.1432955100

Naumann, K. (2000). Influence of chlorine substituents on biological activity of chemicals: a review. *Pest Management Science*, 56(1). [https://doi.org/10.1002/\(SICI\)1526-4998\(200001\)56:1%3C3::AID-PS107%3E3.0.CO;2-P](https://doi.org/10.1002/(SICI)1526-4998(200001)56:1%3C3::AID-PS107%3E3.0.CO;2-P)

Nyburg, S. C. (1979). 'Polar flattening': non-spherical effective shapes of atoms in crystals. *Acta Crystallographica Section A*, 35(4), 641–645. <https://doi.org/10.1107/S0567739479001492>

Odd Hassel. (1970). Structural Aspects of Interatomic Charge-Transfer Bonding. *Science*, 170(3957), 497–502.

Politzer, P., Lane, P., Concha, M. C., Ma, Y., & Murray, J. S. (n.d.). *An overview of halogen bonding*. <https://doi.org/10.1007/s00894-006-0154-7>

Politzer, P., Lane, P., Concha, M. C., Ma, Y., & Murray, J. S. (2007). An overview of halogen bonding. *Journal of Molecular Modeling*, 13(2), 305–311. <https://doi.org/10.1007/s00894-006-0154-7>

Politzer, P., Murray, J. S., & Concha, M. C. (2007). Halogen bonding and the design of new materials: Organic bromides, chlorides and perhaps even fluorides as donors. *Journal of Molecular Modeling*, 13(6-7), 643–650. <https://doi.org/10.1007/s00894-007-0176-9>

Ramasubbu, N., Parthasarathy, R., & Murray-Rust, P. (1986). Angular Preferences of Intermolecular Forces around Halogens Centers: Preferred Directions of Approach of Electrophiles and Nucleophiles around the Carbon-Halogen Bond. *Journal of the American Chemical Society*, 108(15), 4308–4314. https://doi.org/10.1021/JA00275A012/SUPPL_FILE/JA00275A012_SI_001.PDF

Riel, A. M. S., Rowe, R. K., Ho, E. N., Carlsson, A. C. C., Rappé, A. K., Berryman, O. B., & Ho, P. S. (2019). Hydrogen Bond Enhanced Halogen Bonds: A Synergistic Interaction in Chemistry and Biochemistry. *Accounts of Chemical Research*, 52(10), 2870–2880. <https://doi.org/10.1021/acs.accounts.9b00189>

Riley, K. E., Murray, J. S., Fanfrlík, J., Řezáč, J., Solá, R. J., Concha, M. C., Ramos, F. M., & Politzer, P. (2011). Halogen bond tunability I: The effects of aromatic fluorine substitution on the strengths of halogen-bonding interactions involving chlorine, bromine, and iodine. *Journal of Molecular Modeling*, 17(12), 3309–3318. <https://doi.org/10.1007/s00894-011-1015-6>

Riley, K. E., Murray, J. S., Politzer, P., Concha, M. C., & Hobza, P. (2009). Br···O complexes as probes of factors affecting halogen bonding: Interactions of bromobenzenes and bromopyrimidines with acetone. *Journal of Chemical Theory and Computation*, 5(1), 155–163. <https://doi.org/10.1021/ct8004134>

Rosokha, S., & Kochi, J. (2008). Structures and Electronic Spectra of the π -Halogen Complexes between Halogen Donors and Acceptors with π -Receptors. In *Halogen Bonding*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-540-74330-9>

Saldívar-González, F., Prieto-Martínez, F. D., & Medina-Franco, J. L. (2017). Descubrimiento y desarrollo de fármacos: un enfoque computacional. *Educacion Quimica*, 28(1), 51–58. <https://doi.org/10.1016/j.eq.2016.06.002>

Sarno, S., Papinutto, E., Franchin, C., Bain, J., Elliott, M., Meggio, F., Kazimierczuk, Z., Orzeszko, A., Zanotti, G., Battistutta, R., & A. Pinna, L. (2011). ATP Site-Directed Inhibitors of Protein Kinase CK2: An Update. *Current Topics in Medicinal Chemistry*, 11(11), 1340–1351. <https://doi.org/10.2174/156802611795589638>

Scholfield, M. R., Vander Zanden, C. M., Carter, M., & Ho, P. S. (2013). Halogen bonding (X-bonding): A biological perspective. *Protein Science*, 22(2), 139–152. <https://doi.org/10.1002/pro.2201>

Smith, B. R., Eastman, C. M., & Njardarson, J. T. (2014). Beyond C, H, O, and N! analysis of the elemental composition of U.S. FDA approved drug architectures. *Journal of Medicinal Chemistry*, 57(23), 9764–9773. <https://doi.org/10.1021/jm501105n>

Syssa-Magale, J. L., Boubeker, K., Palvadeau, P., Meerschaut, A., & Schöllhorn, B. (2005). The tailoring of crystal structures via the self-assembly of organic coordination compounds by N···I non-covalent halogen bonds: co-crystals of sterically hindered N-heterocycles and 1,4-diyodo-tetrafluorobenzene. *Undefined*, 7, 302–308. <https://doi.org/10.1039/B500009B>

Tiz, D. , Bagnoli, L., Rosati, O., Marini, F., Sancinetto, L., & Santi, C. (2022). New Halogen-Containing Drugs Approved by FDA in 2021: An Overview on Their Syntheses and Pharmaceutical Use. *Molecules*, 27. <https://doi.org/10.3390/molecules27051643>

Vasylyeva, V., Nayak, S. K., Terraneo, G., Cavallo, G., Metrangolo, P., & Resnati, G. (2014). *Orthogonal halogen and hydrogen bonds involving a peptide bond model Electronic Supplementary Information*. www.ccdc.cam.ac.uk/data_request/cif.

Voth, A. P. (2007). The Role of Halogen Bonding in Inhibitor Recognition and Binding by Protein Kinases. *Current Topics in Medicinal Chemistry*, 7(14), 1336–1348. <https://doi.org/10.2174/156802607781696846>

Voth, A. R., Khuu, P., Oishi, K., & Ho, P. S. (2009). Halogen bonds as orthogonal molecular interactions to hydrogen bonds. *Nature Chemistry*, 1(1), 74–79. <https://doi.org/10.1038/nchem.112>

Wang, W., Ji, B., & Zhang, Y. (2009). Chalcogen bond: A sister noncovalent bond to halogen bond. *Journal of Physical Chemistry A*, 113(28), 8132–8135. https://doi.org/10.1021/JP904128B/ASSET/IMAGES/MEDIUM/JP-2009-04128B_0004.GIF

Wilcken, R., Liu, X., Zimmermann, M. O., Rutherford, T. J., Fersht, A. R., Joerger, A. C., & Boeckler, F. M. (2012). Halogen-enriched fragment libraries as leads for drug rescue of mutant p53. *Journal of the American Chemical Society*, 134(15), 6810–6818. <https://doi.org/10.1021/ja301056a>

Wilcken, R., Zimmermann, M. O., Lange, A., Joerger, A. C., & Boeckler, F. M. (2013). Principles and

applications of halogen bonding in medicinal chemistry and chemical biology. *Journal of Medicinal Chemistry*, 56(4), 1363–1388. <https://doi.org/10.1021/jm3012068>

Yamamoto, H. M., Kosaka, Y., Maeda, R., Yamaura, J.-I., Nakao, A., Nakamura, T., Kato, R., & Riken, †. (2008). *Supramolecular Insulating Networks Sheathing Conducting Nanowires Based on Organic Radical Cations*. <https://doi.org/10.1021/nn700035t>

Zalazar, M. F., & Peruchena, N. M. (2014). Laplacian of the electron density: A hole-lump interaction as a tool to study stereoelectronic control of chemical reactions. *Journal of Physical Organic Chemistry*, 27(4), 327–335. <https://doi.org/10.1002/poc.3260>

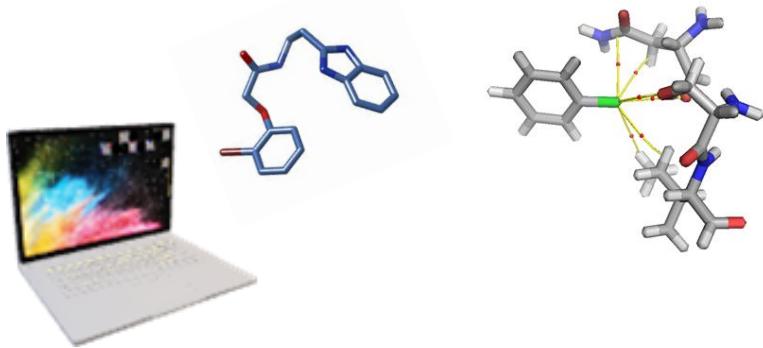
Zaldini Hernandes, M., Melo Cavalcanti, S. T., Rodrigo Moreira, D. M., Filgueira de Azevedo Junior, W., & Cristina Lima Leite, A. (2010). Halogen Atoms in the Modern Medicinal Chemistry: Hints for the Drug Design. *Current Drug Targets*, 11, 303–314.

Zimmermann, M. O., Lange, A., Wilcken, R., Cieslik, M. B., Exner, T. E., Joerger, A. C., Koch, P., & Boeckler, F. M. (2014). Halogen-enriched fragment libraries as chemical probes for harnessing halogen bonding in fragment-based lead discovery. *Future Medicinal Chemistry*, 6(6), 617–639. <https://doi.org/10.4155/fmc.14.20>

Zou, J. W., Jiang, Y. J., Guo, M., Hu, G. X., Zhang, B., Liu, H. C., & Yu, Q. Sen. (2005). Ab initio study of the complexes of halogen-containing molecules RX (X = Cl, Br, and I) and NH₃: Towards understanding the nature of halogen bonding and the electron-accepting propensities of covalently bonded halogen atoms. *Chemistry - A European Journal*, 11(2), 740–751. <https://doi.org/10.1002/chem.200400504>

Capítulo 2

Fundamentos y metodología



CAPITULO 2

El Diseño de fármacos está integrado por diversas áreas de la investigación que incluyen a la quimioinformática, bioinformática, modelado molecular, química teórica y visualización de datos. Cada una comprende métodos que se emplean dependiendo de las características del sistema, información experimental disponible y de los objetivos específicos del proyecto. El modelado computacional comprende a un conjunto de métodos que intentan describir en forma simplificada sistemas y procesos reales aplicando principios matemáticos y físicos. Esto permite interpretar datos experimentales, entender mecanismos interviniéntes en procesos complejos y realizar predicciones. En forma semejante a una prueba experimental, hay que validar los métodos computacionales y encontrar los parámetros óptimos para que los cálculos estén apegados a la realidad.

A continuación, se describen brevemente los métodos computacionales utilizados para el desarrollo de la presente tesis.

2.1 Modelado por homología

El modelado por homología se utiliza para construir un modelo estructural 3D de una proteína (objetivo de modelado) sobre la base de la alineación de su secuencia de aminoácidos con una proteína relacionada de estructura conocida (modelo) (Orry & Abagyan, 2012).

Cualquier enfoque de modelado por homología consta de cuatro pasos (Figura 2.1):

1-Identificación de proteínas similares cuyas estructuras se encuentran disponibles y por lo tanto se pueden utilizar como plantillas o “templates” para modelar

2-Identificar los residuos similares entre la secuencia objetivo de estudio y la estructura de la plantilla, este proceso generalmente se denomina alineamiento.

3-Generar un modelo 3D de una proteína diana en base a la alineación adecuada de la estructura utilizada como “template” y de la secuencia.

4-Estimar la corrección del modelo. Todo el proceso puede ser iterado (reiniciado en cualquiera de los pasos) hasta que se obtenga la calidad satisfactoria o hasta que el modelo ya no puede ser mejorado.

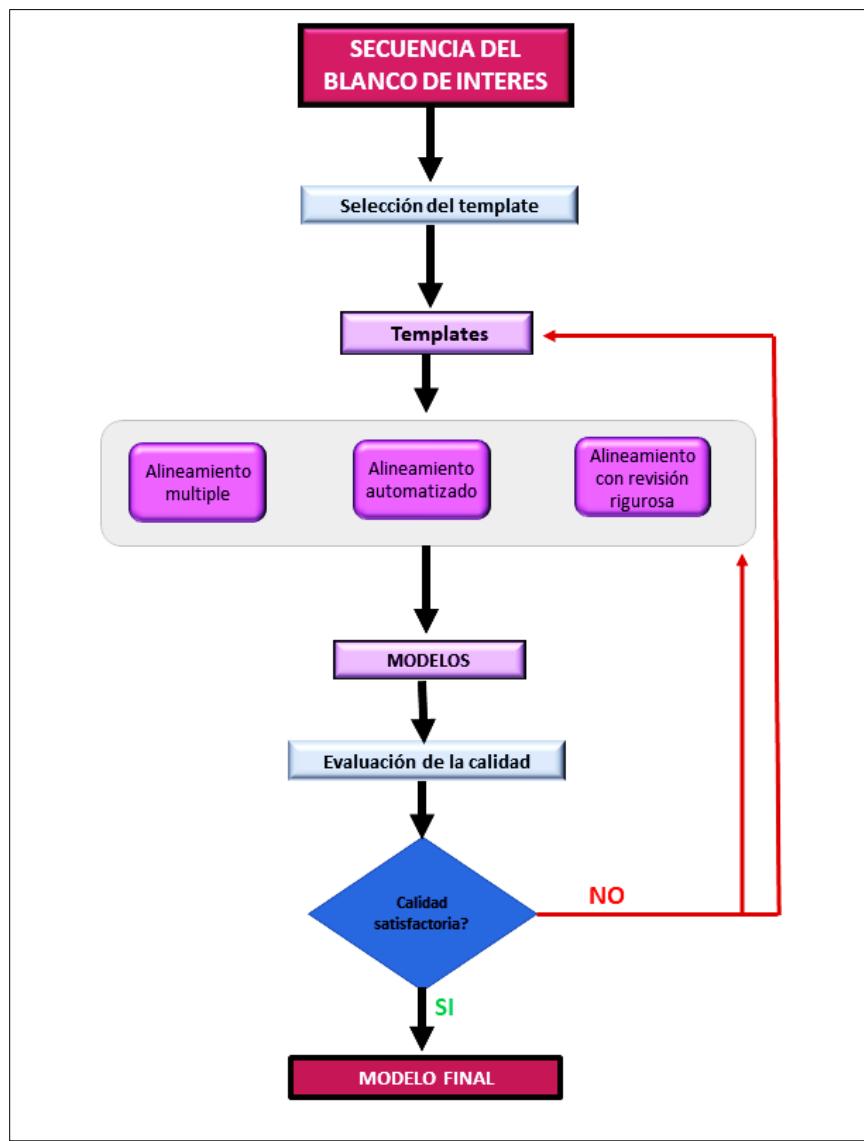


Figura 2.1: Resumen de los pasos que se deben seguir para generar un modelo 3D a través del Modelado por Homología. (Esquema adaptado de Homology Modelling, Methods and Protocols edited by Andrew J. Orry y Ruben)

El paso de alineación de la estructura con la secuencia homóloga busca reproducir la misma con la mayor exactitud posible, pero sin el beneficio de conocer la estructura objetivo del modelado. Obviamente, a menos que el objetivo y la plantilla estén muy estrechamente relacionados, puede haber regiones con diferencias estructurales significativas entre los dos. Estas regiones estructuralmente diferentes, generalmente son resultado de inserciones, delecciones, o cambios extensos en la secuencia de aminoácidos. En otras palabras, una alineación exacta de una determinada secuencia debe incluir todos los pares de residuos equivalentes estructural y evolutivamente, al mismo tiempo, dejando de lado regiones estructuralmente diferentes. Como

el número de estructuras determinadas experimentalmente continúa creciendo constantemente, en muchos casos un objetivo de modelado puede alinearse no sólo con un único “template” sino también con un número (a veces muy grande) de estructuras disponibles como plantillas.

A menudo, una alineación precisa sobre toda la longitud de la secuencia de estudio no puede lograrse con la misma plantilla; en cambio, diferentes dominios pueden alinearse con diferentes plantillas. Esto proporciona una oportunidad para mejorar el modelo, pero introduce una complejidad adicional (Figura 2.2).

Los métodos que detectan homología a través de la alineación de un par de secuencias (alineación “pair-wise”) son simples: utilizan sólo secuencias de aminoácidos de dos proteínas, una tabla de puntuación para sustituciones de residuos y un algoritmo para producir una alineación. Por lo general, los métodos de alineación “pairwise” informan la significancia estadística de las alineaciones resultantes. Indudablemente, la herramienta de búsqueda de base de datos más popular basada en este tipo de alineación es “Basic Local Alignment Search Tool” *BLAST* (Altschul et al., 1990), es muy rápida y tiene una sólida base estadística para la inferencia por homología. La integración del conjunto de programas BLAST junto con bases de datos de secuencias Centro Nacional de Información sobre Biotecnología (NCBI) es otro factor importante que contribuye a la popularidad de BLAST.

Los programas de comparación de pares de secuencias pueden proporcionar una estimación inicial del nivel de dificultad del modelado de homología. Ellos pueden ser adecuadas para detectar proteínas relacionadas que comparten más de 25-30% de residuos idénticos, el rango de similitud entre las secuencias puede denominarse como zona de *luz natural*. Sin embargo, en muchos casos, las alineaciones correspondientes necesitan mejoras. Sólo si las secuencias alineadas son más del 40-50% idénticas entre sí y tienen pocas o no poseen lagunas, puede esperarse que las alineaciones sean precisas en un sentido estructural.

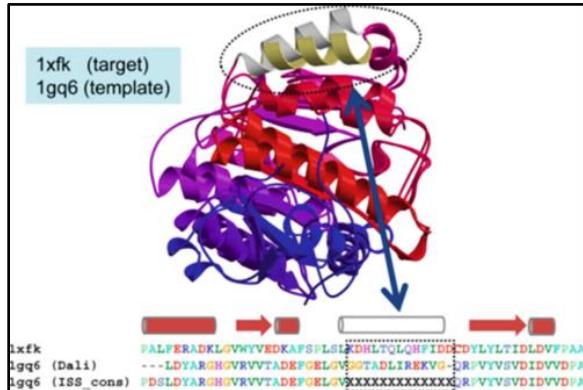


Figura 2.2: Alineamiento donde se indican las de regiones con residuos idénticos en las dos proteínas (modelo y template) (figura tomada de Homology Modelling, Methods and Protocols edited by Andrew J. Orry y Ruben Abagya)

2.1.1 Modelado empleando múltiples moldes (MSA)

Los métodos de alineación de secuencias múltiples representan un caso distinto ya que no están diseñados para detectar secuencias homólogas. En su lugar, alinean un conjunto de secuencias ya identificadas por otros métodos. Principalmente, estos métodos pueden utilizarse para mejorar la calidad de los modelos. Dado un conjunto de secuencias, los métodos MSA pretenden construir una alineación en la que las columnas representan la evolución (estructural) de residuos equivalentes. Aunque en teoría la programación dinámica de algoritmos para la alineación “pairwise” se puede ampliar para una alineación óptima de múltiples secuencias, son computacionalmente exigentes, como resultado, la mayoría de las técnicas actuales utilizan varias aproximaciones heurísticas.

La mayoría de las herramientas modernas utilizan la heurística conocida como alineación progresiva. En esta estrategia, el árbol de guía de alineación se construye primero de acuerdo a las similitudes de secuencia “pairwise”. Usando este árbol guía, las secuencias con mayor similitud se alinean primero. A continuación, son subalineadas entre sí hasta que todas las secuencias se incorporan en el modelo.

2.1.2 Servidores web

El modelado computacional se ha convertido en una valiosa herramienta para complementar la elucidación experimental sobre las estructuras de las proteínas, y para poder hacer la información tridimensional accesible a una amplia comunidad de investigadores biomédicos. Hace casi dos décadas, el primer modelo de servidores web SWISS-MODEL(Ertl, 2003) se puso

a disposición en Internet. Desde entonces, se han desarrollado muchos servidores para modelar las estructuras de las proteínas de una manera automatizada por ejemplo, ModWeb (Pieper et al., 2009), HHpred (Söding et al., 2005), I-TASSER (Yang & Zhang, 2015) y MODELLER (Šali & Blundell, 1993). Recientemente el desarrollo de nuevos métodos tiene por objeto incluir restricciones experimentales adicionales a los procedimientos de modelización y establecer métodos especializados en ciertas familias de proteínas tales como GPCRs o anticuerpos.

Uno de los principales objetivos de automatizar la estructura comparativa de proteínas de modelado con la plantilla seleccionada, alineación de plantillas, construcción de modelos y evaluación de la calidad del modelo es la necesidad de hacer estas tecnologías accesibles a una audiencia de no expertos en bioinformática. Lo que permite el uso de herramientas computacionales que son necesarias en las tareas de modelado.

En segundo lugar, debido al gran número de proteínas cuya estructura aún no ha sido caracterizada experimentalmente, los procedimientos automatizados son esenciales para hacer frente a esta inundación de datos, por ejemplo, para aumentar la cobertura de información estructural para el estudio de proteomas de organismos enteros o familias de proteínas.

2.1.3 Validación del Modelo

Después de la construcción del modelo, es importante comprobar la presencia de errores. Se pueden llevar a cabo dos tipos de validaciones:

- Evaluación interna de la consistencia del modelo, la cual comprueba si el modelo satisface las restricciones usadas para calcularlo.
- Evaluación externa, la cual se refiere a información que no ha sido usada en el cálculo del modelo. A su vez existen dos tipos de evaluaciones externas:
 - a) *Testear si se ha utilizado una plantilla correcta:* se realiza este tipo de evaluación cuando el modelo está basado en menos de un 30% de similitud de la secuencia con respecto de la plantilla. Una manera de predecir si una plantilla es mejor o peor es la comparación en Prosa II del Z-score de las estructuras del modelo y la plantilla. El Z-score de un modelo es una medida de compatibilidad entre su secuencia y su estructura, los cuales deberían ser comparables. Ahora bien, cabe decir que esta evaluación no siempre funciona debido a que la función de potencial utilizada en Prosa II (Wiederstein & Sippl, 2007) no es apropiada para ciertos tipos de plegamientos.

b)-*Reconocer regiones poco fiables en el modelo:* Una vía de aproximación es calculando un perfil de energías con programas como Prosa II. Dicho perfil reporta la energía de cada posición en el modelo. Es posible detectar errores en el modelo porque aparecen como picos de energía positiva en el perfil de energías. Estas regiones del modelo deben ser inspeccionadas cuidadosamente. Otra vía de encontrar regiones poco fiables es evaluar la estereoquímica del modelo (longitud de enlaces, ángulos de enlace, ángulos diedros, solapamiento entre átomos) con programas como Procheck(Laskowski et al., 1993). Finalmente, una herramienta de evaluación importante es el conocimiento experimental acerca de la estructura de la proteína y su función. Un modelo debería ser consistente con las observaciones experimentales como los datos obtenidos por mutagénesis dirigida, datos de crosslinking, unión de ligandos activos, etc.

2.2 Métodos de simulación para sistemas Moleculares

Entre los métodos teóricos utilizados hoy en día para el refinamiento de modelos moleculares para el estudio de interacciones fármaco-receptor se incluyen las minimizaciones de energía (ME) utilizando métodos mecano cuánticos, las simulaciones de Monte Carlo (MC) y las perturbaciones de energía libre (PEL). Todas estas técnicas ofrecen en última instancia la posibilidad de comprender el comportamiento de péptidos, proteínas y ácidos nucleicos y el de sus complejos entre si o con moléculas orgánicas, en términos de fuerzas moleculares fundamentales (Iii & Kollman, 2000).

También se usan para obtener información cinética y termodinámica. Los métodos de simulación de Dinámica Molecular (DM) pueden proporcionar además detalles relativos a las propuestas de resolución de las partículas individuales en función de tiempo. Estos métodos pueden ser utilizados para cuantificar las propiedades de un sistema con una precisión y escala de tiempo que de otro modo sería inaccesible y por lo tanto la simulación es una valiosa herramienta en la comprensión de los sistemas de modelado. Actualmente, la simulación de dinámica molecular (DM), es una de las principales herramientas para el estudio teórico de las moléculas biológicas(Hansson et al., 2002). Este proceso calcula computacionalmente el comportamiento de un sistema molecular con respecto al tiempo y ha proporcionado una gran cantidad de información detallada sobre las variaciones y los cambios conformacionales dentro de las proteínas y los ácidos nucleicos.

2.2.1 Modelado Molecular

La Superficie de Energía Potencial (SEP) es una representación del conjunto de valores de la energía potencial molecular correspondientes a cada una de las disposiciones posibles de los núcleos que conforman el sistema en estudio en función de las variables de las que depende. La energía potencial es la energía que corresponde a la molécula para una configuración fija de núcleos. Para una molécula de N núcleos:

$$V(r^N) = E_{ele}(r^N) + V_{nm}(r^N) \quad (2.1)$$

Donde V_{nm} es la energía de repulsión internuclear y E_{ele} la energía electrónica (obtenida mediante resolución de la ecuación de Schrödinger):

$$H_{ele}\Psi_{ele} = E_{ele}\Psi_{ele} \quad (2.2)$$

La energía potencial nos proporciona las fuerzas que actúan sobre los núcleos en el campo creado por los electrones y por lo tanto nos permite saber que configuraciones son estables, así como las respuestas del sistema ante una perturbación externa.

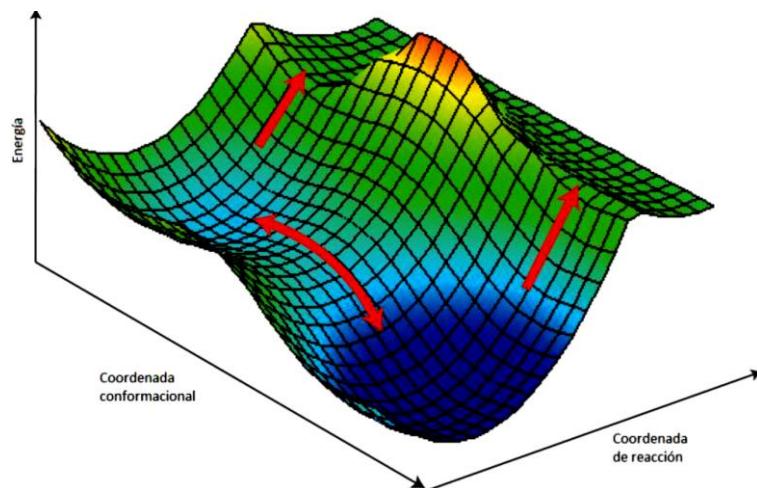


Figura 2.3: Representación de la Superficie de Energía Potencial, conjunto de valores de la energía potencial molecular correspondientes a cada una de las disposiciones posibles de los núcleos que la forman.

Existen dos grandes métodos para calcular la SEP, por cálculos cuánticos y por mecánica molecular. La descripción científicamente correcta de la energía de un sistema molecular está dada por la ecuación de Schrödinger de la mecánica cuántica. Sin embargo, los cálculos requeridos para ello son muy costosos computacionalmente, debido a que es necesario modelar cada electrón en cada orbital, muchos de los cuales se describen por múltiples funciones

(Young, 2009). Por lo tanto, los métodos cuánticos están limitados a sistemas de menos de 100 átomos (ab-initio y DFT) o entre 100-1000 átomos (métodos semi-empíricos). Por esa misma limitación, es difícil realizar un tratamiento riguroso del entorno (disolvente, superficie, proteína) donde tiene lugar el proceso a estudiar. Los cálculos de mecánica molecular (MM) por el contrario, no tratan de manera explícita a los electrones en el sistema. En su lugar realizan cálculos basados en las interacciones entre núcleos. Los efectos de los electrones están incluidos en los campos de fuerza a través de la parametrización del sistema (Foresman & Frisch, 1996). La simulación por mecánica molecular se basa en las leyes de la física clásica para predecir las estructuras y propiedades de las moléculas. La ventaja de estos métodos es que pueden utilizarse en sistemas muy grandes lo que permite incluir el efecto del entorno en los procesos químicos. La principal desventaja es que no es posible tratar procesos que impliquen reorganización electrónica. En MM las moléculas son tratadas como un conjunto de átomos en el espacio, unidas entre sí mediante enlaces y gobernadas por una serie de funciones de potencial mecano-clásicas.

La forma general de la energía potencial V en los métodos MM es:

$$V = \text{Intramolecular} + \text{Intermolecular} \quad (2.3)$$

La contribución intermolecular suele escribirse como la suma de un término electrostático y otro de Van der Waals (*términos no enlazantes*). Por otra parte, la contribución intramolecular o *términos enlazantes* se escriben en función de una serie de términos asociados a las variables que definen la geometría molecular (distancias, ángulos diedros).

En los diferentes métodos de MM todos los parámetros en las ecuaciones son ajustados a los datos experimentales. En conjunto, las ecuaciones y los parámetros asociados se denominan *campos de fuerzas* (Young, 2009). Un campo de fuerza consta de los siguientes componentes:

- Un conjunto de ecuaciones que definen como la energía potencial de la molécula varía con respecto a la distribución de sus respectivos átomos.
- Series de tipos de átomos, que definen las características y comportamiento de los elementos dentro de un contexto químico específico.
- Uno o varios sets de parámetros que ajustan las ecuaciones y el tipo de átomo a los datos experimentales.

Existen dos tipos de campos de fuerza, el de los átomos y el de átomos unidos. En el primero se toman en cuenta todos los átomos que forman parte del sistema a simular, mientras que en el

campo de fuerza de átomos unidos se pueden dejar de tener en cuenta algunos átomos de hidrógeno. En este tipo de campo de fuerza se tienen en cuenta por ejemplo los átomos de hidrógeno polares. Los campos de átomos unidos en ocasiones no funcionan bien, este es el caso de las interacciones aromáticas. La elección del campo de fuerza que uno va a utilizar depende del tipo de sistema biomolecular que necesita simular. La exactitud de los resultados de las simulaciones dependerá de la precisión del método de derivación de los parámetros utilizados y de la forma analítica de representar la energía. Las limitaciones y deficiencias de los campos de fuerza utilizados se ponen de manifiesto cuando se comparan los resultados de las simulaciones con datos experimentales.

Cada método se caracteriza por sus respectivos campos de fuerzas. Las constantes que contribuyen a la energía son lo que diferencia a cada uno de los campos de fuerza que actualmente se conocen y se utilizan (CHARMM, AMBER, GROMOS, OPLS entre otros).

La expresión del campo de fuerza AMBER (Ponder & Case, 2003), el cual fue utilizado para los cálculos realizados en este trabajo, contiene dos clases de contribuciones, las contribuciones enlazantes y las no enlazantes (Ecuación 2.4). Las enlazantes, se computan para átomos formando un enlace, e incluyen penalidades para las desviaciones respecto de su valor de equilibrio de distancias de enlace, ángulos y ángulos diedros. Mientras que las dos primeras se representan a través de un potencial armónico centrado en la distancia de equilibrio, los ángulos diedros o torsiones están descriptos con una función periódica. Las contribuciones no enlazantes, se computan entre átomos que están en diferentes moléculas y se dividen en interacciones electrostáticas basadas en la ley de Coulomb, por un lado, y un potencial de Lennard-Jones, que describe en forma aproximada las interacciones dispersivas y la repulsión originada por el principio de exclusión de Pauli a distancias muy cortas. La expresión de la energía potencial resulta:

$$E(r^N) = \sum_{enlaces} \frac{k_{b,i}}{2} (r_i - r_{i_0})^2 + \sum_{angulos} \frac{k_{\theta,i}}{2} (\theta_i - \theta_{i_0})^2 + \sum_{torsiones} \frac{V_n}{2} [1 + \cos(n\phi - \gamma)] + \sum_{i=1}^{N-1} \sum_{j=i+1}^N \left\{ \epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right\} \quad (2.4)$$

El primer término refiere a la contribución de estiramiento de los enlaces, el cual se describe a través de un potencial armónico. En éste, $k_{b,i}$ corresponde a la constante de fuerza asociada a la unión i y r_{i0} a la distancia de equilibrio de la misma unión. El segundo término corresponde a las contribuciones dadas por las flexiones angulares. Estas también se representan con un potencial armónico $k_{\theta,i}$ y valor de equilibrio θ_i . El término asociado a las torsiones o ángulos diedros presenta tres parámetros: V_n corresponde a la barrera energética asociada a la rotación entre dos mínimos sucesivos, n representa la multiplicidad del potencial, es decir indica el número de mínimos encontrados entre 0 y 360° , y por último γ , que corresponde a la fase de la función sinusoidal. El último término de la ecuación contiene las contribuciones no enlazantes, y depende de la distancia entre pares de átomos, r_{ij} . La primera parte, corresponde al potencial de Van der Waals, y contiene los parámetros ϵ_{ij} asociado a la profundidad del pozo de energía, y σ_{ij} , correspondiente a la distancia entre los átomos correspondiente al mínimo de energía. La parte electrostática contiene como parámetros las cargas sobre los átomos. En el campo de fuerzas de AMBER, la distribución de cargas es representada a través de cargas puntuales q_i ubicadas en las coordenadas de cada átomo. Estas cargas se obtienen de manera de reproducir el potencial electrostático de la molécula, calculado a partir de cálculos cuánticos.

2.2.1.1 Simulación de Dinámica Molecular

En una simulación de dinámica molecular clásica se calcula una evolución temporal de las posiciones y velocidades de los átomos de un sistema a través de la resolución de las ecuaciones de movimiento de Newton. Este método computacional se emplea ya de forma rutinaria para proporcionar información detallada sobre la dinámica (fluctuaciones y cambios conformacionales) y la termodinámica de macromoléculas biológicas como proteínas, polisacáridos y ácidos nucleicos en solución acuosa, así como complejos de asociación proteína-ligando.

Los programas de DM están construidos sobre un algoritmo muy sencillo:

1. Lectura de coordenadas y velocidades iniciales
2. Cálculo del potencial y de la fuerza
3. Integración de las ecuaciones de Newton
4. Cálculo de propiedades

Antes de hablar de la dinámica molecular se debe especificar ciertas cuestiones relacionadas con el sistema que se desea estudiar y de qué manera hacerlo. El objetivo de simular es el estudio de sistemas realistas. Así a la hora de seleccionar un modelo que represente al sistema en estudio se debe prestar atención a tres aspectos:

- El número de moléculas que forman el sistema, limitado por la capacidad de cálculo (tiempo y memoria)
- Número de microestados que se van a generar para obtener las propiedades promedio de nuestro sistema y que vendrá limitado por el tiempo de cálculo.
- El potencial que gobierna el comportamiento del sistema (interacciones intra e intermoleculares).

En química computacional, el problema consiste en evaluar la SEP (por métodos clásicos (MM) o cuánticos (QM)). En la práctica se trata de obtener información sin llegar a explorar toda la SEP. En química cuántica tradicional se realizan optimizaciones de la geometría que corresponden a analizar mínimos de la SEP a temperatura 0°K. Ahora bien, si queremos estudiar un sistema real debemos incluir el efecto de la temperatura. Si se usan métodos cuánticos (QM) la corrección de temperatura se puede obtener a partir de un cálculo de frecuencias en programas de cálculo de estructura electrónica como Gaussian (M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ort, 2016). Si se realizan simulaciones clásicas MM el efecto de la temperatura se puede incluir haciendo uso de la termodinámica estadística. Las leyes mecánicas de los sistemas microscópicos no contienen conceptos tales como la temperatura, el calor o la entropía. La termodinámica estadística muestra como estos conceptos surgen de la incertidumbre natural que existe sobre el estado o configuración del sistema microscópico. De forma general, denominando **A** a las propiedades tales como la presión o la capacidad calorífica dependerán de las posiciones y momentos de las **N** partículas que constituyen el sistema. Estas propiedades macroscópicas del sistema pueden obtenerse como un promedio sobre el conjunto de configuraciones accesibles o colectivo.

Llamando $\rho(\vec{p}^N, \vec{r}^N)$ a la densidad de probabilidad de hallar una configuración con momentos \vec{p}^N y posiciones \vec{r}^N :

$$\langle A \rangle = \int d\vec{p}^N d\vec{r}^N A(\vec{p}^N, \vec{r}^N) \cdot \rho(\vec{p}^N, \vec{r}^N) \quad (2.5)$$

Donde,

$$\rho(\vec{p}^N, \vec{r}^N) = \frac{e^{-E(\vec{p}^N, \vec{r}^N)/kT}}{\int d\vec{p}^N d\vec{r}^N e^{-E(\vec{p}^N, \vec{r}^N)/kT}} \quad (2.6)$$

el denominador de la ecuación anterior es lo que se conoce como función de partición (**Q**) del sistema. Esta función contiene toda la información termodinámica de un sistema de partículas en equilibrio. La función de partición **Q** resulta de explorar todas las conformaciones posibles del sistema. En la práctica obtener **Q** es difícil (solo se hace para gases ideales y suponiendo modelos simplificados de oscilador armónico, etc). Una de las hipótesis básicas de la termodinámica estadística es la hipótesis ergódica, que afirma que el promedio temporal es equivalente a un promedio sobre un colectivo de sistemas idénticos que estando en el mismo estado macroscópico se encuentran en distintos estados microscópicos.

$$\lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau A(\vec{p}^N(t), \vec{r}^N(t)) dt = \int d\vec{p}^N d\vec{r}^N A(\vec{p}^N, \vec{r}^N) \cdot \rho(\vec{p}^N, \vec{r}^N) \quad (2.7)$$

Así, el valor de la propiedad **A** puede en principio también obtenerse promediando la propiedad **A(t)** en función del tiempo y se obtiene cuando el intervalo del promedio temporal tiende al infinito.

$$\bar{A}_\tau = \frac{1}{\tau} \int_{t=0}^\tau A(\vec{p}^N(t), \vec{r}^N(t)) dt \quad (2.8)$$

Dado que obtener **Q** es muy difícil, en las simulaciones de Dinámica Molecular, haciendo uso de la hipótesis ergódica, se realiza un promedio temporal en lugar de configuracional para conocer aproximadamente **Q**. Hay que tener cuidado de que los tiempos de corrida no sean muy cortos para evitar problemas de ergodicidad.

Los sistemas de un tamaño considerable (del orden de las decenas de miles de átomos) difícilmente puedan ser tratados a un nivel mecanocuántico, es decir, considerando explícitamente los electrones y núcleos. Por ello las simulaciones de tipo MD son una herramienta poderosa y nos permite la observación de eventos temporales del orden de pico, nano e incluso microsegundos que puedan suceder en nuestro sistema de interés.

Las simulaciones de tipo MD pueden ser realizadas usando un amplio rango de condiciones que emulan las condiciones experimentales típicas. Las primeras fueron hechas en vacío tomando a las moléculas como entidades aisladas. Luego, se comenzó a desarrollar e incluir al solvente de manera implícita o explícita.

Condiciones periódicas de contorno: este método consiste en rodear el sistema o celda inicial de simulación, por réplicas idénticas. El número de moléculas a describir no crece realmente ya que el movimiento de las réplicas es igual al observado en el sistema inicial. La ventaja es que ahora todas las moléculas del sistema se encuentran rodeadas por más moléculas y por lo tanto desaparecen los efectos de superficie. Si una molécula abandona la celda inicial durante su movimiento, otra entrará desde una celda vecina (Figura 2.5)

Las condiciones periódicas se utilizan de forma muy generalizada en las simulaciones, pero tiene algunos problemas como la imposibilidad de observar fluctuaciones con una longitud de onda mayor que el tamaño de la caja o que su uso impone una periodicidad artificial en el sistema, la cual puede favorecer la aparición de estructuras cristalinas.

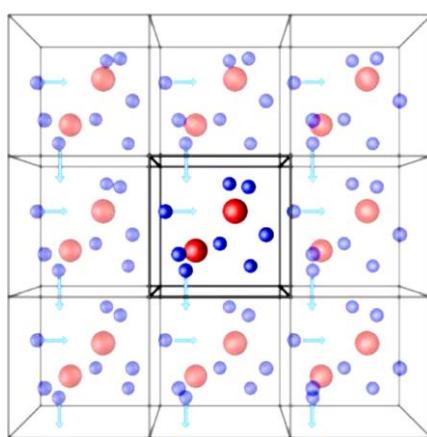


Figura 2.5: Condiciones de borde periódicas en 3D. Nótese el sistema central rodeado por las réplicas.

En la dinámica molecular sucesivas configuraciones del sistema son generadas integrando las leyes de movimiento de Newton. El resultado es una trayectoria que especifica como la posición y la velocidad de las partículas, en un sistema, varían en el tiempo.

La trayectoria se obtiene resolviendo la ecuación de la segunda ley de Newton

$$(F=ma): m_j (d^2r_j/dt^2) = f_j = -\nabla_j V \quad (2.9)$$

Esta ecuación describe el movimiento de una partícula de masa m_j a lo largo de una coordenada (r_j) con una fuerza f_j sobre la partícula (Leach, 2001).

Para sistemas reales, los potenciales de interacción son continuos y la fuerza que actúa sobre cada partícula cambiará cuando la partícula cambie de posición o cuando cualquier otra partícula lo haga. Una vez conocidas las fuerzas que actúan sobre cada partícula puede procederse a integrar las ecuaciones de Newton para conocer las coordenadas y velocidades que puede ir adquiriendo el sistema conforme pasa el tiempo. Los métodos más efectivos para resolver ecuaciones diferenciales como la anterior suelen ser los Métodos de diferencias finitas donde las ecuaciones de movimiento se integran en intervalos de tiempo muy cortos, Δt "time step". Típicamente se dan valores del orden del femtosegundo ($1\text{fs}=10^{-15}\text{s}$). En cada etapa se calcula el vector suma de todas las interacciones que actúan sobre ella. Se calculan las fuerzas sobre todas las partículas y se combinan con la posición y velocidad actuales para generar la posición y velocidad para un tiempo posterior $t+\Delta t$. Los algoritmos para calcular las posiciones, velocidades, aceleraciones, etc., parten de desarrollar en series de Taylor estas magnitudes.

Ensamblés

Un colectivo o ensamble es una colección de microconfiguraciones del sistema en el espacio de fases que satisface las condiciones de un estado termodinámico determinado. Una simulación de dinámica molecular genera una secuencia de configuraciones en el espacio que está concatenada en función del tiempo y recibe el nombre de trayectoria. Cada punto de la trayectoria equivale a una foto de una película que está separada de la anterior y de la siguiente por un intervalo de tiempo dado ("frames"). Todos los puntos de la trayectoria pertenecen al mismo conjunto y corresponden a diferentes configuraciones (posiciones y momentos) del sistema.

Aunque cada estado microscópico sea diferente, todos pertenecen al mismo estado macroscópico termodinámico del sistema. Las simulaciones más ampliamente utilizadas corresponden a los siguientes ensambles:

- ✚ El ensamble canónico (NVT): colección de todos los sistemas cuyo estado termodinámico está caracterizado por un número fijo de átomos (N), volumen fijo (V) y temperatura dada (T).
- ✚ El ensamble isobárico-isoentálpico (NPH).
- ✚ El ensamble isobárico-isotérmico (NPT).

Típicamente, los experimentos se realizan a temperatura y volumen constantes (ensamble canónico) o bien a temperatura y presión constante (ensamble isobárico-isotérmico), por lo tanto es frecuente elegir estas condiciones de simulación para imitar estas condiciones, o porque es esperable encontrarlas en un entorno fisiológico.

Solvatación

Generalmente nuestro sistema de interés es una proteína perteneciente a un organismo vivo, es importante tener en cuenta una representación razonable del solvente con el cual vamos a embeber al sistema. Esto nos permitirá tener una representación más cercana a las condiciones en donde la proteína se encuentra en la naturaleza. La mayor parte de las proteínas se encuentran, al menos parcialmente, embebidas en un entorno acuoso. Por esto, es usual asumir en las simulaciones de estos sistemas, que están completamente solvatadas, ya sea en agua pura o en agua con iones. Sin embargo, una porción considerable del tiempo de cómputo es utilizada en evaluar la interacción solvente-solvente.

Una manera de evitar esto es usar un modelo de solvente implícito, que puede ser adecuado en algunos casos, a pesar de la pérdida del detalle en la descripción. En consecuencia, han sido implementados numerosos modelos de solvente implícito.

- ✚ Solvente implícito: Uno de ellos es el modelo llamado modelo “generalizado de Born” (GB). En general, estos modelos de solvente implícito consideran a todo el solvente como una “jalea” que envuelve a la proteína y está caracterizada por una constante dieléctrica. eligiendo entonces este parámetro se obtienen distintos tipos de solvente implícito (agua, etanol, mezcla de solventes, etc). Los modelos de solvente implícito son incapaces de reproducir fenómenos microscópicos del solvente, por ejemplo, enlaces de hidrógeno, adsorción sobre la proteína, etc. Además, facilita la aparición de

determinadas conformaciones “irreales” de la proteína cuando se las compara con solvente explícito.

- Solvente explícito: El poder de cómputo actual permite usar sin inconvenientes el solvente explícito. Este modelo utiliza una descripción atomística de las moléculas de solvente de forma análoga a la usada para la proteína, es decir asigna carga, enlaces, diedros, parámetros de van der Waals, etc. para cada átomo del solvente. Existen un amplio rango de modelos de agua explícita disponibles. Los más populares son TIP3P, TIP4P, TIP5P, SPC y SCP/E, que tienen la particularidad de tratar a las moléculas de agua como completamente rígidas. Esta última aproximación también ahorra mucho tiempo de cómputo, y se ha demostrado en numerosos trabajos que la mayoría de las propiedades de los sistemas no se ven afectados por dicha simplificación. Finalmente, para lograr el ensamble isobárico-isotérmico comúnmente utilizado necesitamos controlar la presión y temperatura de nuestras simulaciones.

Termostatos

- Termostato de Berendsen (termostato proporcional)

El termostato de Berendsen corrige las desviaciones de la temperatura T respecto de la temperatura de referencia T_0 a tiempo t multiplicando las velocidades de los átomos por un factor λ con el propósito de llevar el sistema global a la temperatura T_0 .

La constante τ_T es la constante de acoplamiento temporal, la cual determina la escala de tiempo para la cual la temperatura deseada es alcanzada. Si se elige un τ_T muy grande el sistema tiende a comportarse de manera aislada.

- Termostato de Langevin (termostato estocástico)

En el caso de los termostatos de tipo estocásticos, todos o un subconjunto de grados de libertad del sistema son sometidos a colisiones con partículas virtuales. Este método está motivado por la ecuación estocástica diferencial de Langevin, la cual describe el movimiento de una partícula dada por la agitación térmica de un baño,

$$ma = -\zeta_v v + f(r) + f' \quad (2.10)$$

donde m es la masa de una partícula, a es su aceleración, $f(r)$ es una fuerza conservadora que actúa sobre la partícula, v es la velocidad de la partícula, ζ es una constante friccional (la fuerza friccional $-\zeta v$ disminuye la energía cinética) y f' es una fuerza aleatoria de tipo gaussiana que

agrega energía cinética a las partículas y su varianza es función de la temperatura y del tiempo. Por consiguiente, la fuerza aleatoria se impone para poder mantener el sistema a T constante, dicho de otra manera, la fuerza friccional y la fuerza aleatoria se compensan tratando de mantener el sistema a la temperatura deseada y permitiendo que evolucione en el tiempo.

Cuanto mayor es el valor de ζ , se incrementan las fluctuaciones térmicas (porque es como considerar más choques virtuales), mientras que para $\zeta=0$ el sistema es aislado (es decir, sería análogo a que τ tienda a infinito para el termostato de Berendsen).

Barostatos

Para mantener constante la presión durante la simulación es necesario permitir que el volumen fluctúe, ajustando para eso las dimensiones de la caja periódica y re-escalando las posiciones de los átomos de acuerdo a esto. Existen numerosos métodos para hacer dinámicas a presión constante. Estos métodos incluyen el algoritmo de sistema extendido (“extended system algorithm”), el de restricción (“constraint algorithm”), un acoplamiento débil con un baño externo, el método híbrido y el método del pistón de Langevin.

Pasos de la Dinámica Molecular:

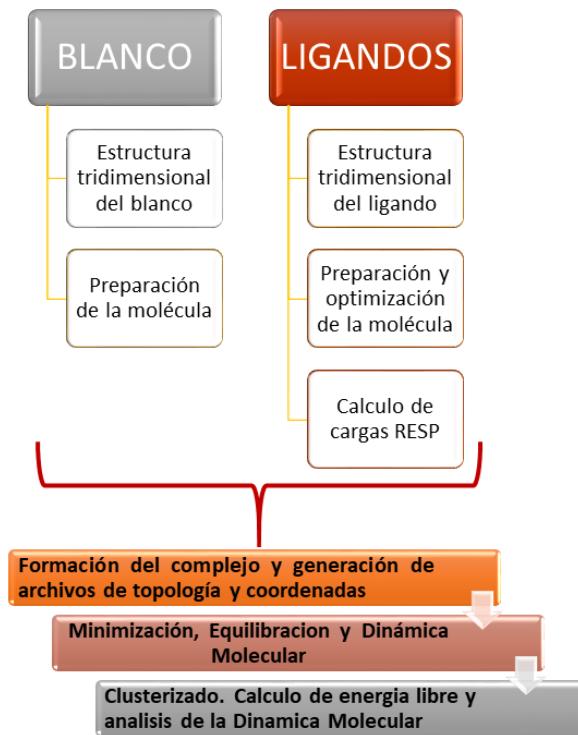


Figura 2.3: Esquema de pasos de la Dinámica Molecular

1- Obtención de la estructura inicial: El primer paso en toda simulación de MD consiste en conseguir una estructura a nivel atómico de la proteína de interés. Descargar de la base de datos Protein Data Bank (<http://www.rcsb.org/>) (Burley et al., 2021) la estructura en formato pdb correspondiente.

2- Preparación de archivos de coordenadas y topología: Una vez obtenida la estructura inicial, se preparará el sistema y se armarán los archivos requeridos para poder realizar las simulaciones de dinámica molecular. Con el archivo pdb correctamente editado, ya podemos construir los dos archivos requeridos por AMBER para realizar cualquier tipo de cálculos: el archivo de coordenadas (llamado archivo ".rst7") y el de parámetros/topología (llamado archivo ".prmtop"). El primero tiene información de las coordenadas (y eventualmente velocidades) de todos los átomos del sistema.

También tiene información sobre los límites de la caja de solvatación. El archivo prmtop contiene los parámetros del campo de fuerzas (parámetros de enlace, ángulos, diédros, etc.) que

se utilizarán y la topología molecular (conectividad de los átomos, tipos de átomos, masas, cargas, etc.). También posee información sobre el tipo de caja de solvente (octahédrica, cúbica, etc.). Para construir estos dos archivos se usa un módulo que se llama “t-Leap” (el cual viene incluido en el paquete de AMBER). Para armar el sistema con sus correspondientes parámetros.

3- Minimización de energía del sistema: Para cristalizar una proteína se la somete a condiciones muy particulares, por ejemplo, baja temperatura, distinta fuerza iónica respecto a su ambiente “natural”, entre otras. Incluso la proteína en el cristal puede adoptar una conformación en donde algunos aminoácidos están demasiado cerca. Todo lo anterior hace que la estructura cristalográfica muchas veces no sea representativa de la estructura a 298 °K, pH fisiológico, ni a fuerza iónica de un entorno biorrellevante, con lo cual muchas veces sucede que la estructura cristalográfica resulta ser tan solo una aproximación o una *foto* de una proteína en su estado natural. Si bien esta estructura puede aportar alguna idea de su comportamiento en condiciones biológicas relevantes, no siempre es concluyente. Es por eso que necesitamos llevar de a poco y gradualmente la estructura del cristal de la proteína a las condiciones más próximas a la *vida real*.

Existen muchas formas de realizar este procedimiento. Generalmente se realiza primero una minimización de energía para “relajar” la estructura de solvatación, solucionando principalmente los problemas electrostáticos dados por la orientación de las moléculas de agua. Se hará entonces una minimización de la energía del sistema restringiendo a los átomos de la proteína (dejándola fija) y se permitirá al solvente de la caja de aguas moverse y “relajarse” alrededor de la proteína. Posteriormente, se hará una segunda minimización sin restricción alguna para obtener la relajación de la geometría de todo el sistema en conjunto (proteína y solvente).

4-Termalización del sistema: En esta etapa se calentará al sistema de 0K a 298 °K de manera gradual, ésta es la fase denominada calentamiento o termalización. Para esto se generará una rampa de temperatura que se irá incrementando hasta alcanzar la esperada.

5- Equilibración del sistema: Habiendo calentado el sistema, ahora se buscará simular durante un tiempo determinado a temperatura y presión constante, de manera que el sistema acomode su volumen para lograr una densidad adecuada. Generalmente, esta fase suele ser de algunos ns de simulación, pero es muy dependiente del tamaño y de las características de cada sistema.

6- Producción de la Dinámica Molecular: Hasta este punto tienen una trayectoria de unos cuantos de nanosegundos para nuestra proteína de interés en solución acuosa. Si hemos realizado todo el protocolo antes mencionado correctamente, podemos estimar que nuestro sistema se encuentra equilibrado a $T=300\text{ }^{\circ}\text{K}$ y a $P=1\text{ atm}$. Se puede asumir que la trayectoria responde bien a un modelo de ensamble a NPT constante. El protocolo empleado permite por lo menos decir que se encuentran en la zona cercana a un mínimo de energía potencial. Con lo cual, pueden intuir que las microconfiguraciones que observan van a ser relevantes en la *vida* de la proteína y por lo tanto también lo serán los eventos que sucedan en la dinámica obtenida.

2.2.1.2 Cálculos de energía libre de unión

En la última década se han realizado numerosos esfuerzos con el propósito de alcanzar mayor precisión en el cálculo de la energía libre de unión, de manera computacionalmente eficiente. Esto ha sido potenciado por el incremento exponencial de recursos computacionales, que se tradujo favorablemente en innovaciones metodológicas. Sin embargo, los cálculos de energía libre y concretamente su aplicación a sistemas macromoleculares sigue siendo problemáticos por distintos factores, entre los cuales se pueden mencionar: el modelo de energía subyacente, el tamaño del sistema de interés y la flexibilidad, ya que se basan en la exactitud de la función de energía potencial del sistema y en una exploración exhaustiva de la hipersuperficie de energía.

Un problema fundamental de los métodos Perturbación de Energía Libre (FEP) e Integración Termodinámica (TI) es su gran demanda computacional, sumado a la dificultad que tienen para alcanzar la convergencia. El poder de cómputo que requieren está relacionado con dos factores principales. El primero es el tratamiento explícito del solvente y el segundo es la exploración exhaustiva de la hipersuperficie de energía potencial que se realiza para determinar la diferencia de energía libre entre dos estados del sistema. Una alternativa atractiva a estos métodos computacionalmente intensivos son los denominados métodos de puntos extremos, que consideran únicamente los estados final e inicial en los cálculos de energía libre. Los métodos “Molecular Mechanics Poisson-Boltzmann Surface Area” (MM-PBSA) y “Molecular Mechanics Generalized Born Surface Area” (MM-GBSA) son los más comúnmente utilizados.

En la práctica, se calcula la energía libre de un complejo y sus componentes, siguiendo la aproximación de trayectoria única. Esta consiste en realizar una única simulación de dinámica molecular y extraer de la misma las conformaciones del complejo proteína-ligando, por lo que

las conformaciones de las partes individuales en el estado libre se obtienen eliminando la proteína o ligando de cada una de las configuraciones, correspondientemente. A pesar de que en este caso no se tiene en cuenta la flexibilidad conformacional que ocurre durante la unión, esta aproximación conduce a resultados precisos siendo más adecuada para sistemas en los que no se esperan encontrar cambios conformacionales significativos entre las estructuras unida y no unida.

El método MM-GBSA utiliza el modelo de solvente basado en la teoría de Born Generalizada, siendo la ecuación:

$$G_s = (1/8\pi) (1/\epsilon_0 \cdot 1/\epsilon) \sum_{A,B} (Q_A Q_B / f_{GB}) \quad (2.11)$$

En esta ecuación, ϵ_0 es la permitividad del vacío, y f_{GB} es una función de la distancia y la interacción soluto-disolvente.

El cálculo de la energía libre total de unión se realizará a partir de las energías libres del complejo proteína-ligando (C) y los reactivos ligando (L) y receptor (R):

$$\Delta G = \langle GC \rangle - \langle GL \rangle - \langle GR \rangle \quad (2.12)$$

La energía libre G de cada uno de los estados se calculará mediante la suma:

$$G = E_{int} + E_{vdW} + E_{el} + E_{polar\ solv} + E_{np\ dissol} - TS \quad (2.13)$$

El primer término, E_{int} , corresponde a la energía interna del estado, formada a su vez por las energías de longitud de enlace, ángulo de enlace y diedros. Los dos términos siguientes, E_{vdW} y E_{el} , determinan las contribuciones de van der Waals y electrostáticas, respectivamente. Los parámetros, $E_{p\ solv}$ y $E_{np\ dissol}$, son las energías de solvatación. La energía de solvatación polar se obtendrá resolviendo la ecuación de GB expresada anteriormente (2.11). El término $E_{np\ dissol}$ está directamente relacionado con el área superficial accesible al disolvente, SASA, y se calcula por medio de una aproximación. Y los dos últimos parámetros, T temperatura, multiplicada por la entropía S.

2.3 Cribado Virtual y Docking molecular

El cribado virtual (CV) es un filtrado computacional (in silico) de moléculas para seleccionar candidatos (hits computacionales) para su evaluación experimental (Saldívar-González et al., 2017; Scior et al., 2012) De esta manera, el cribado virtual reduce significativamente el número

de ensayos biológicos que se harían si no hubiera una selección de compuestos. Sin embargo, es un proceso predictivo que debe integrarse con ensayos experimentales que validen las predicciones de los ensayos *in silico* (Lionta et al., 2014).

Existen diversos filtros que se utilizan para llevar a cabo el cribado virtual los cuales pueden variar según la complejidad de la base de datos y la información experimental de la que se disponga. Por ejemplo, si se conoce la estructura tridimensional (3D) del receptor se sugiere un cribado basado en la estructura (CVBE, Cribado virtual basado en la estructura). Si solo se conocen los compuestos activos, pero no el receptor, entonces la búsqueda se hace basada en el ligando (CVBL, Cribado Virtual Basado en el Ligando). Si se conoce la estructura 3D del receptor y de los compuestos activos se pueden combinar ambos tipos de filtros para facilitar la búsqueda, por ejemplo, se puede realizar un primer tamizaje empleando técnicas de CVBL que son computacionalmente menos demandantes, seguido de un CVBE.

CVBE se apoya en el conocimiento de la estructura del blanco molecular y en la habilidad de los algoritmos de acoplamiento molecular (“docking”) para predecir modo y energías de unión de compuestos en las bases de datos, de tal manera que solo los compuestos mejor *ranqueados* son corroborados experimentalmente (Gil Redondo, 2010).

El acoplamiento molecular es el método más utilizado en CVBE por su balance entre precisión y costo computacional. Sin embargo, puede recurrirse a otros métodos de cribado basado en la estructura dependiendo del balance que se busque entre estas dos variables contrapuestas. Por ejemplo, los modelos farmacofóricos basados en la estructura del blanco molecular permiten un tamizaje rápido de bases de datos de compuestos a expensas de una reducción en la precisión

Screening virtual basado en la estructura, permiten evaluar de una rápida y efectiva las interacciones de una librería de compuestos con blancos moleculares, además permite estudiar el modo de unión de pequeñas moléculas al sitio activo del blanco. “Docking” consiste en la determinación computacional de la afinidad de unión entre una estructura proteica y un ligando. Este método implica un muestreo de todas las poses posibles que puede adoptar el ligando en el bolsillo de unión de la proteína a fin de obtener la geometría de unión óptima, medida por las funciones de puntuación del algoritmo de docking (Gilson & Zhou, 2007). El acoplamiento de moléculas pequeñas se puede realizar generalmente de tres maneras:

- ✚ acoplamiento rígido, en el que el blanco molecular y el ligando se tratan de forma rígido

- ✚ acoplamiento flexible del ligando, en el que el blanco molecular se mantiene rígido;
- ✚ acoplamiento flexible, en el que tanto la proteína como el ligando se consideran flexibles (Mohan et al., 2005).

Existen diversos filtros que se utilizan para llevar a cabo el cribado virtual los cuales pueden variar según la complejidad de la base de datos y la información experimental de la que se disponga. Por ejemplo, si se conoce la estructura tridimensional (3D) del receptor se sugiere un cribado basado en la estructura (acoplamiento molecular). Si solo se conocen los compuestos activos, pero no el receptor, entonces la búsqueda se hace basada en el ligando (similitud molecular). Si se conoce la estructura 3D del receptor y de los compuestos activos se pueden combinar filtros para facilitar la búsqueda, tales como descriptores moleculares y propiedades farmacocinéticas, entre otras. Las técnicas para hacer cribado virtual dependen de la información disponible del sistema.

La combinación de “screening” virtual con el reposicionamiento de fármacos representa hoy en día una metodología eficiente para el desarrollo de nuevos medicamentos. El reposicionamiento de fármacos consiste en encontrar una nueva aplicación terapéutica distinta por la que fue diseñado un fármaco y que no necesariamente tiene que ser la misma dosis. El objetivo de esta estrategia es reducir los gastos y tiempos de investigación (Saldívar-González et al., 2017). A partir de la hipótesis de que *moléculas similares tienen propiedades similares*, se han aplicado estrategias computacionales encaminadas a reposicionar fármacos.

Los protocolos de acoplamiento molecular también se pueden definir como una combinación de un algoritmo de búsqueda y una función de puntuación (Hassan Baig et al., 2016). Ya que tiene dos componentes:

- ✚ “Docking” o proceso de búsqueda de la conformación y orientación de las molécula
- ✚ “Scoring”, que consiste en asignar un valor o puntaje que mida la interacción entre las dos estructuras.

Esta metodología se basa en que el algoritmo de búsqueda proporcione soporte y libertad a la coordinación proteína-ligando para permitir un muestreo preciso y exhaustivo de los posibles modos de unión.

Para describir la metodología del docking es útil dividirla en una serie de etapas (Figura 2.6). La primera de ellas consiste en disponer de la estructura tridimensional de la molécula blanco, la cual debe ser acondicionada para los cálculos subsecuentes, se puede generalizar que las modificaciones más comunes que deben hacerse a estas moléculas son: la adición de los átomos

de hidrógeno, la remoción de las moléculas de agua que están unidas a la superficie de la proteína, y que con gran frecuencia ocupan el sitio de unión; la eliminación de ligandos presentes; y finalmente, el modelado de las regiones que no pudieron ser determinadas experimentalmente o que resultaron ser ambiguas. Una vez preparada la molécula blanco debe identificarse el sitio de unión para lo que existen dos alternativas, la primera es computacional y consiste en revisar la superficie molecular en busca de cavidades con ciertas características de tamaño, hidrofobicidad y grupos funcionales. La segunda es experimental y en ella se requiere determinar los sitios donde se albergan moléculas de solvente orgánico que difunden por el cristal de una proteína (Gao & Serrero, 1999). La ubicación de tales moléculas de solvente se determina por difracción de rayos X y se consideran probables sitios de unión para un ligando (Mattos & Ringe, 1996). Es necesario enfatizar en este momento que la unión de un ligando en una de estas zonas no necesariamente afectará la función de la enzima.

La siguiente etapa requiere poseer un archivo de ligandos potenciales, esto es, moléculas orgánicas con estructuras tridimensionales conocidas y que se encuentran en condiciones adecuadas para simular su asociación al blanco, el acondicionamiento a que deben sujetarse estas moléculas es: a) deben añadirse átomos de hidrógeno, ya que en muchos complejos proteína-ligando están involucradas las interacciones de puentes de hidrógeno; b) debe darse una representación tridimensional a las moléculas, pues en la mayoría de los casos sólo se cuenta con una imagen plana o una fórmula codificada de ellas y; c) debe estimarse la carga parcial de cada uno de los átomos que constituyen al compuesto. En esta última tarea, debe considerarse que las moléculas pequeñas poseen una mayor variedad de tipos de átomo y grupos funcionales que las proteínas conocidas. Por ello, en vez de emplear tablas de valores promedio para la asignación de cargas, se utiliza un algoritmo basado en cálculos de electronegatividad de los átomos involucrados en cada enlace químico (Gasteiger & Marsili, 1980).

La tercera etapa, que es la parte central del método, consiste en un algoritmo que toma cada uno de los ligandos de la base datos y lo coloca dentro del sitio de unión en una gran cantidad de orientaciones (Lengauer & Rarey, 1996). Durante esta etapa es indispensable emplear un criterio numérico para distinguir cuál de todas las disposiciones probadas resulta ser la más adecuada, es decir, la que adoptaría el ligando si tuviera acceso al sitio de unión; a este criterio le llamaremos puntaje. Una vez establecidas las condiciones descritas anteriormente, el proceso para probar orientaciones se repite para cada ligando de la base de datos, ésta es la etapa que consume la mayor cantidad de tiempo de cálculo. Finalmente, el programa de cómputo ordena

los diferentes compuestos probados de acuerdo con el puntaje de su orientación óptima y entonces, el usuario puede analizar estos resultados y planear experimentos para validar los resultados generados en la simulación por computadora.

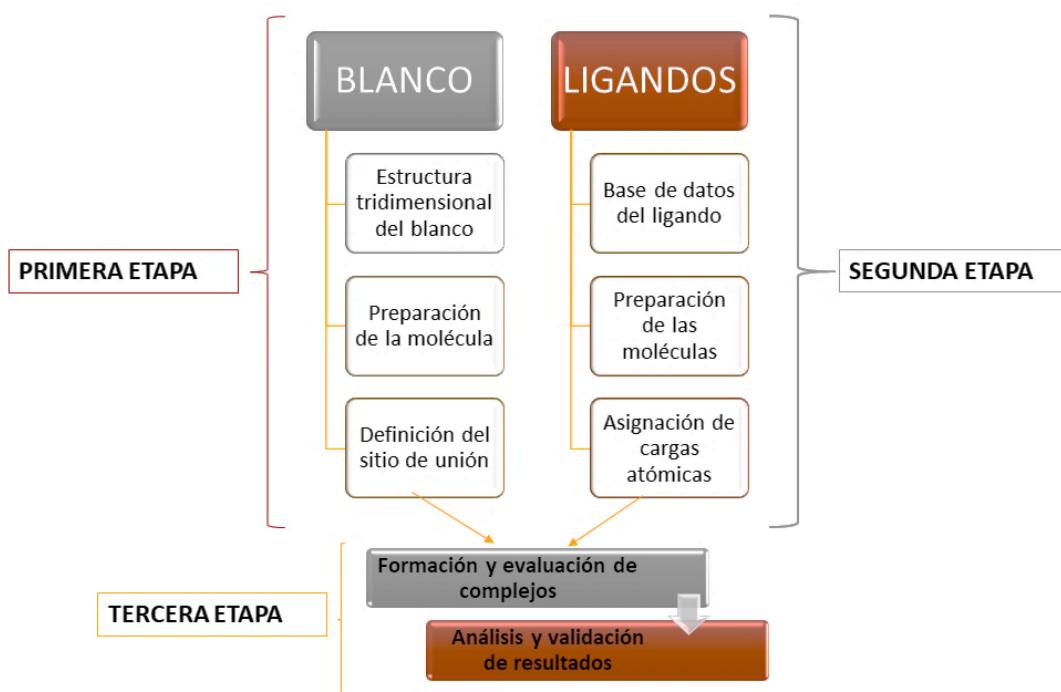


Figura 2.6: Diagrama simplificado de las etapas de simulación del reconocimiento molecular o *docking*.

Resumiendo, el Docking Molecular comprende tres pasos generales (Xu et al., 2013):

- ✚ Generación de un modelo molecular del receptor
- ✚ Pre-tratamiento y muestreo conformacional de los ligandos
- ✚ Obtención de una puntuación o “score” asociado a la energía de interacción de cada potencial ligando de la base de datos al sitio activo del blanco molecular de elección.

El procedimiento de docking es utilizado para explorar y predecir los posibles modos de formación de complejos ligando-blanco molecular (Sousa et al., 2006), para considerar la flexibilidad del ligando generalmente se usan métodos estocásticos como Monte Carlo o algoritmos genéticos, el algoritmo de Monte Carlo opera variando un parámetro a la vez generando una nueva conformación que es aceptada o rechazada de acuerdo a los criterios de Boltzmann. Por otro lado los algoritmos genéticos se basan en la teoría de evolución de Darwin,

la conformación de los ligandos se inicia como una población aleatoria de estados modelada como un conjunto de cromosomas que pueden cruzarse o mutar y producir otras conformaciones, este ciclo se repite hasta que el mínimo local de energía sea alcanzado.

El proceso por el cual un ligando se une a su blanco molecular depende de factores entrópicos y entálpicos. Las funciones de scoring se utilizan para predecir la energía libre de unión del complejo, que es una medida de la potencia de unión de la molécula al blanco biomolecular. Estas funciones se pueden agrupar en tres categorías: puntuación basada en el campo de fuerza, empírica y basada en el conocimiento; las funciones de puntuación basadas en los campos de fuerza se sustentan en las interacciones físicas entre átomos, como las interacciones de VdW, interacciones electrostáticas y fuerzas de extensión, flexión y torsión de los enlaces entre átomos, además de los términos de energía de solvatación, que se calculan con modelos de solvente continuo, como Poisson-Boltzmann (PB) o "Generalized-Born" (GB). Las funciones de campos de fuerza y sus parámetros suelen derivar de datos experimentales y cálculos *ab initio*. En las funciones empíricas, un conjunto de parámetros ponderados, correspondientes a varios tipos de interacciones entre el blanco y el ligando, se calibran con un conjunto de entrenamiento experimental. Los coeficientes de los diferentes términos, que generalmente incluyen interacciones polares tales como enlaces de hidrógeno e interacciones electrostáticas o apolares como las aromáticas e interacciones lipofílicas, se optimizan ajustando las estructuras y afinidades vinculantes del conjunto de entrenamiento en un Modelo de regresión lineal múltiple. Por último, en la puntuación basada en el conocimiento, se extrae información de complejos objetivo-ligando conocidos y se utiliza para derivar potenciales de interacción atómica que describen las interacciones entre los átomos del ligando y el blanco. Debido a que cada método de scoring acarrea errores se ha propuesto el uso de múltiples funciones para mejorar la estimación de la afinidad de unión de los complejos.

Es una técnica computacional que para el caso de complejos proteína-ligando, permite:

- ✚ Identificar las posibles poses de un ligando dado en un sitio de unión de la proteína o receptor;
- ✚ obtener la pose energéticamente más favorable;
- ✚ establecer un ranking de afinidades del sitio para ligandos química o estructuralmente similares.

Por ejemplo, la pose de un ligando integra tanto la orientación del ligando respecto del receptor como la conformación intrínseca del ligando. La siguiente figura ilustra ambos conceptos para la molécula de ATP:

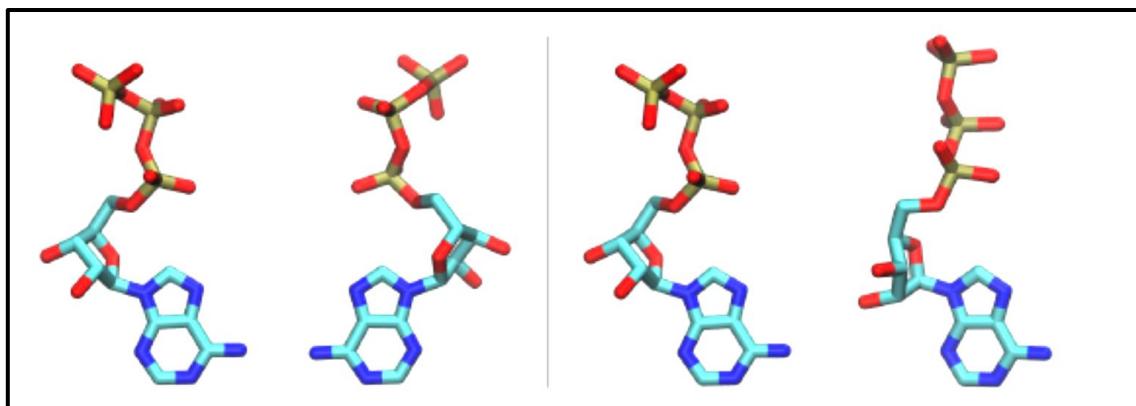


Figura 2.7: Estructura de la molécula de ATP. A la izquierda muestra igual conformación y diferente orientación, a la derecha se muestra diferente conformación con la misma orientación.

La conformación del ligando depende de la capacidad de rotación de sus enlaces. Llamaremos torsiones a los enlaces simples que pueden rotar en un ligando. Cuanto mayor sea el número de torsiones, habrá más posibilidades de generar distintas conformaciones y aumentará el tiempo de cómputo. Cada programa tiene un límite de torsiones pasado el cual la calidad del resultado va perdiendo confiabilidad.

2.3.1 Ensemble docking

El docking molecular puede ser realizado con un único blanco de interés o el docking inverso que se usa para investigar el modo de unión de un fármaco conocido con un panel de blancos terapéuticos, permitiendo así evaluar la selectividad. Una alternativa es obtener un "conjunto" de conformaciones del blanco molecular, mediante el uso de simulaciones de dinámica molecular. La aplicación de este método ha demostrado mostró que los "farmacóforos dinámicos" resultantes eran mejores para predecir la unión de ligandos conocidos que el modelo basado en una única estructura del blanco molecular. Simulaciones de DM se realizan primero para muestrear diferentes conformaciones del receptor diana en forma libre del ligando, con las estructuras obtenidas luego se realiza el docking molecular para identificar posibles ligandos y su modo de unión. El uso de muchas poses de la DM pudo aumentar la flexibilidad del receptor, pero también aumenta el costo del acoplamiento molecular.

2.3.2 Docking Covalente

Los ligandos que se unen a través de un mecanismo covalente no están sujetos a la cinética de equilibrio clásica, ya que su tiempo de residencia en el bolsillo de unión puede durar hasta días. Como consecuencia, la potencia de estos fármacos es capaz de superar los límites teóricos de potencia/eficacia del ligando. En general, un ligando covalente primero requiere la formación de un complejo no covalente inicial con su molécula blanco, seguido por la reacción química entre la *cabeza de guerra* ("warhead") electrofílica del ligando y los residuos nucleofílicos de la proteína. Desde una perspectiva computacional, una vez que se identifican un nucleófilo y una "warhead" apropiadas, se puede usar un enfoque basado en la estructura para seleccionar u optimizar los ligandos para que se ajusten a la sitio de unión, al mismo tiempo que puede colocar la cabeza de guerra en el vecindad del residuo blanco para formar el enlace covalente. En este trabajo usamos la herramienta de Desacoplamiento dinámico (DUck) para la identificación inhibidores covalentes de la enzima Cruzipaina.

2.3.3 Software

AutoDock (Morris et al., 2009) es un software gratuito distribuido bajo licencia GNU y ampliamente utilizado por la comunidad científica. La función de puntaje que utiliza AutoDock sigue un enfoque mixto que combina un potencial con base fisicoquímica similar a un campo de fuerzas de dinámica molecular, cuyos parámetros son luego ajustados a información experimental. Por este motivo, se dice que su función es semiempírica. Para calcular la energía de las distintas poses se utilizan grillas de energía que delimitan una región del espacio (la cual puede abarcar total o parcialmente al receptor) dentro de un cubo o un prisma rectangular que definirá la zona en la que tendrá lugar el cálculo. Luego, previo a la corrida de docking, se evalúa la energía que aportarían los distintos átomos que componen al ligando (carbonos alifáticos, carbonos aromáticos, oxígeno aceptor de enlace de H, etc.) desplazando átomos de prueba a lo largo de los puntos de la grilla y generando así mapas energéticos. Se crea un mapa por cada tipo de átomo y se agrega un mapa electrostático más un mapa de solvatación, los cuales luego se combinan matemáticamente de acuerdo con la ocupación espacial del ligando para determinar el puntaje de la pose dada. La grilla consiste, entonces, en una serie de puntos discretos en el espacio que contienen la información del puntaje energético que resultaría de colocar allí cualquiera de los átomos del ligando.

El campo de fuerzas de AutoDock se basa en la siguiente ecuación:

$$V = W_{vdw} \sum_{i-j} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) + W_{hbound} \sum_{i-j} E(t) \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right) + W_{elec} \sum_{i-j} \frac{q_i q_j}{c(r_{ij}) r_{ij}} + W_{sol} \sum_{i-j} (S_i V_j + S_j V_i) e^{-r_{ij}^2/2\sigma^2} \quad (2.14)$$

Las constantes de ponderación W son las que están optimizadas para calibrar la energía libre empírica. El primer término es un potencial 6/12 para fuerzas de dispersión/repulsión; los parámetros A y B fueron tomados del campo de fuerza de Amber 12. El segundo término es un término de dirección de enlace de H basado en un potencial 10/12. Los parámetros C y D se asignan para dar una profundidad de pozo máxima de 5 kcal/mol a 1,9 Å para OH y NH, y una profundidad de 1 kcal/mol a 2,5 Å para SH. La direccionalidad del enlace de Hidrogeno E(t) depende del angulo t. Las interacciones electrostáticas son consideradas con el Potencial de Coulomb. El término final considera el potencial de desolvatación basado en el volumen (V) de los átomos que rodean un átomo dado, ponderado por un parámetro de solvatación (S) y un término exponencial basado en la distancia.

Los programas de docking se encuentran diseñados para proveer resultados aceptables en una gran gama de compuestos de muy diversas propiedades fisicoquímicas que se ven reflejadas en los conjuntos usados para calibrar la función de puntaje. Esto frecuentemente implica que su performance para con un determinado sistema no sea óptima. Sin embargo, muchas veces uno tiene información sobre el sistema en estudio y puede agregar información a la función de modo de guiar el experimento de docking. Por ejemplo una alternativa es usar la información de los sitios de solvente aplicando la técnica de SSBDM (Arcon et al., 2017), la cual se basa en el hecho de que en las proteínas el solvente tiende a organizarse estructuralmente en torno al sitio de unión imitando las interacciones moleculares de los ligandos cuando se unen al mismo sitio. El método consiste en calcular los sitios de solvente del receptor en el sitio de unión y luego realizar un cálculo de docking modificando los mapas energéticos correspondientes con un sesgo que favorece la ubicación de grupos funcionales/átomos adecuados en las regiones donde se formaron los sitios de solvente.

2.4 Cálculos Mecanoacuánticos

Los métodos de estructura electrónica usan las leyes de la mecánica cuántica para la base de sus cálculos. Establece que la energía y otras propiedades de una molécula pueden ser obtenidas resolviendo la ecuación de Schrödinger:

$$H\Psi = E\Psi \quad (2.15)$$

Sin embargo, aun para sistemas simples las soluciones exactas de la ecuación no son posibles (Hislop et al., 1990). Los métodos utilizados en el cálculo de estructuras electrónicas se caracterizan por las diferentes aproximaciones matemáticas usadas en la resolución de la ecuación. Existen dos clases principales de métodos, los semiempíricos que utilizan parámetros obtenidos experimentalmente para simplificar los cálculos y los métodos ab initio, en los cuales los cálculos se basan solamente en los principios de la mecánica cuántica. Los primeros dependen de la disponibilidad de datos experimentales confiables. Su principal ventaja reside en el tiempo de cálculo requerido, los cuales pueden ser de varios ordenes de magnitud menor que el de los métodos ab initio, los cuales son prohibitivos en problemas que involucran macromoléculas como interacciones receptor-ligando.

La Teoría de Átomos en Moléculas (Atoms in Molecules, AIM) (R. Bader, 1990) se fundamenta en los rigurosos principios de la mecánica cuántica y hereda de ésta, su poder predictivo. Es una teoría en la que no se emplean modelos, ni está sujeta a suposiciones arbitrarias. Proporciona el sustento teórico de la hipótesis de estructura molecular, en la que se reconoce a una molécula como a un conjunto de átomos unidos por una red de enlaces. Así, la teoría AIM se trata de un formalismo que permite definir los conceptos de átomo en una molécula, enlace químico, estructura y estabilidad molecular, cambio estructural y dar, además, descripciones de los diferentes tipos de enlace e interacciones atómicas, en acuerdo con las ideas clásicas de la química descriptiva. La teoría no es dependiente de ningún método de cálculo, ni experimental, y como punto de partida sólo necesita la densidad electrónica del sistema, que puede obtenerse por cálculos químico cuánticos o por técnicas experimentales.

La definición del átomo topológico se basa en la estructura física que exhibe la distribución de carga electrónica, la cual es una manifestación de las fuerzas que actúan dentro del sistema. Las propiedades atómicas se describen por las mismas ecuaciones de movimiento y teoremas asociados a la mecánica cuántica aplicadas al sistema como un todo.

El estudio de las propiedades topológicas de la densidad de carga $\rho(r)$ se ve facilitado a través del estudio del campo del vector gradiente asociado $\nabla\rho(r)$. Las propiedades de este campo proporcionan una definición de los elementos de estructura molecular, estabilidad estructural y permiten delinear los mecanismos de los cambios estructurales, además permite identificar el conjunto de interacciones interatómicas presentes en una molécula (R. F. W. Bader & Nguyen-Dang, 1981). Estas interacciones atómicas pueden caracterizarse mediante el estudio de las

propiedades del campo escalar derivado de la densidad de carga, su distribución Laplaciana $\nabla^2\rho(r)$. El Laplaciano es una propiedad extremadamente importante de un campo escalar, ya que permite determinar en forma directa las regiones en donde el campo está concentrado. El signo del Laplaciano está determinado por el signo y la magnitud relativa de las tres curvaturas locales de $\rho(r)$, autovalores de la matriz Hessiana de $\rho(r)$. Así, existe una íntima unión entre las propiedades topológicas de $\rho(r)$ y su Laplaciano, y a través de las propiedades del Laplaciano se puede establecer una relación entre la forma de la distribución de carga y la mecánica que la gobierna.

2.4.1 Propiedades topológicas de la distribución de la densidad de carga electrónica

La densidad electrónica, $\rho(r)$, es una magnitud física real que tiene asociado un valor en cada punto del espacio y se la puede determinar tanto experimental como teóricamente. El valor local de la distribución de carga, la densidad de carga electrónica $\rho(r)$, para un sistema de N electrones y para una configuración nuclear X fija, se define como:

$$\rho(r, X) = N \sum_{\text{espines}} \int \{ \prod dr \} \psi^*(x, X) \psi(x, X) \quad (2.16)$$

Donde Ψ es una solución de la ecuación de Schrödinger independiente del tiempo, x representa el conjunto de coordenadas electrónicas espaciales y de espín y r las coordenadas espaciales de un electrón.

Se define la densidad electrónica en un punto r del espacio real como la probabilidad por unidad de volumen de encontrar cualquiera de los N electrones del sistema en un elemento diferencial de volumen en torno a la posición r (R. Bader, 1990; P, 2000). $\rho(r)$ se obtiene integrando la función de onda sobre las coordenadas espaciales de todos los electrones excepto uno, y sumando sobre las coordenadas de espín de todos los electrones. Las propiedades topológicas de una distribución de carga electrónica asociada con una configuración nuclear x, se resumen por el número y tipo de puntos críticos de $\rho(r)$, puntos donde el vector gradiente asociado $\nabla\rho(r)$ se anula, $\nabla\rho(r)$ se define como:

$$\nabla\rho(r) = i \frac{\partial\rho(r)}{\partial x} + j \frac{\partial\rho(r)}{\partial y} + k \frac{\partial\rho(r)}{\partial z} \quad (2.17)$$

En la figura 2.7 se muestra el mapa de relieve de la densidad electrónica de la molécula de agua. Se puede apreciar en la figura, que en las posiciones nucleares la densidad electrónica toma los

valores máximos y además, tiende a acumularse en las regiones donde existe un enlace químico. La formación del enlace químico genera una línea de densidad electrónica máxima en la región internuclear, llamada *línea de interacción atómica*, LIA. Se ha demostrado, que la acumulación de carga a lo largo de esta línea de interacción, es una condición necesaria para asegurar la unión entre dos átomos. Cualquier desplazamiento lateral de dicha línea nos lleva a una región con valores menores de la densidad electrónica. A lo largo de esta LIA, los valores de $\rho(r)$ son más pequeños que en los núcleos, lo cual es una consecuencia directa de las capacidades atractoras de los mismos. Además, a lo largo de esta línea, existe un punto en el cual $\rho(r)$ presenta su valor más bajo. Se trata, por lo tanto, de un punto de silla de segundo orden (mínimo en la dirección del enlace y máximo en las otras dos). Ese punto se denomina punto crítico de enlace, PCE (Bader, 1984), y existe entre cada par de átomos que están enlazados químicamente.

Un punto crítico se clasifica en base a su rango y signatura (w, σ) . El rango w de un punto crítico en r_c (donde r_c representa al vector posición del punto crítico) es igual al número de autovalores $\lambda_{1,2,3}$ no nulos de la matriz Hessiana, Λ :

$$\Lambda_{ij} = \left[\frac{\partial^2 \rho}{\partial r_i \partial r_j} \right]_{r=r_c} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \quad (2.18)$$

La signatura σ es la suma algebraica de los signos de los autovalores. Existen cuatro tipos diferentes de puntos críticos, PC, de rango tres: $(3, -3)$, $(3, -1)$, $(3, +3)$ y $(3, +1)$. Los términos autovalores o curvaturas, y autovectores o trayectorias de gradiente, se utilizan indistintamente.

- ⊕ Punto Crítico $(3, -3)$ posee las tres curvaturas negativas y $\rho(r)$ es un máximo local en r_c , coincide habitualmente con las posiciones nucleares, por ello se denomina *atractor nuclear*
- ⊕ Punto Crítico $(3, -1)$ posee dos curvaturas negativas y $\rho(r)$ es un máximo en r_c en el plano definido por los ejes correspondientes. $\rho(r)$ es un mínimo en r_c a lo largo del tercer eje perpendicular a este plano (es mínimo en la dirección de la línea que une los núcleos atómicos entre los que se encuentra y máximo en las dos direcciones perpendiculares a esta). Es un punto de silla y se identifica con un punto crítico de enlace.

- Punto Crítico (3, +1) posee dos curvaturas positivas y $\rho(r)$ es un mínimo en r_c en el plano definido por los ejes correspondientes. $\rho(r)$ es un máximo en r_c a lo largo del tercer eje perpendicular a este plano.
- Punto crítico (3, +3) posee las tres curvaturas positivas y $\rho(r)$ es un mínimo local en r_c , sus autovalores definen un volumen de líneas de gradiente que nacen desde este punto, está asociado a la idea estructural de caja.

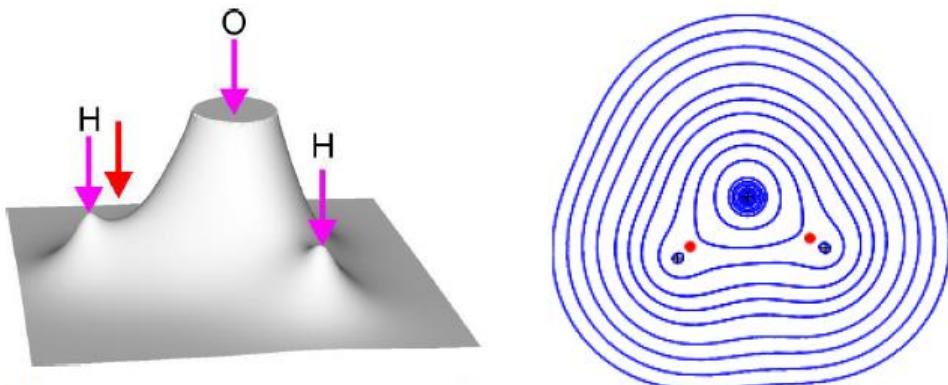


Figura 2.7: A la izquierda mapa de relieve de la densidad electrónica y a la derecha mapa de contorno para la molécula de agua, en el plano que contiene al núcleo de oxígeno y a los núcleos de hidrógeno. Las flechas rosadas señalan los PC (3, -3) correspondiente a los átomos de la molécula de agua, y la flecha roja señala el PC (3, -1). En el mapa de contorno los puntos rojos denotan la posición de los PC (3, -1), y las cruces negras la posición de los PC (3, -3).

2.4.2 Campo del vector gradiente de la densidad electrónica

La función $\nabla\rho(r)$ define un campo de trayectorias o vectores dirigidos en cada punto, a lo largo del gradiente de la densidad de carga. Los vectores o trayectorias de gradiente se originan o terminan en puntos donde $\nabla\rho(r)$ se anula, es decir en puntos críticos de $\rho(r)$. La Figura 2.8 muestra el mapa del campo vector del gradiente de la densidad electrónica de la molécula de agua, donde se observan los elementos de estructura molecular.

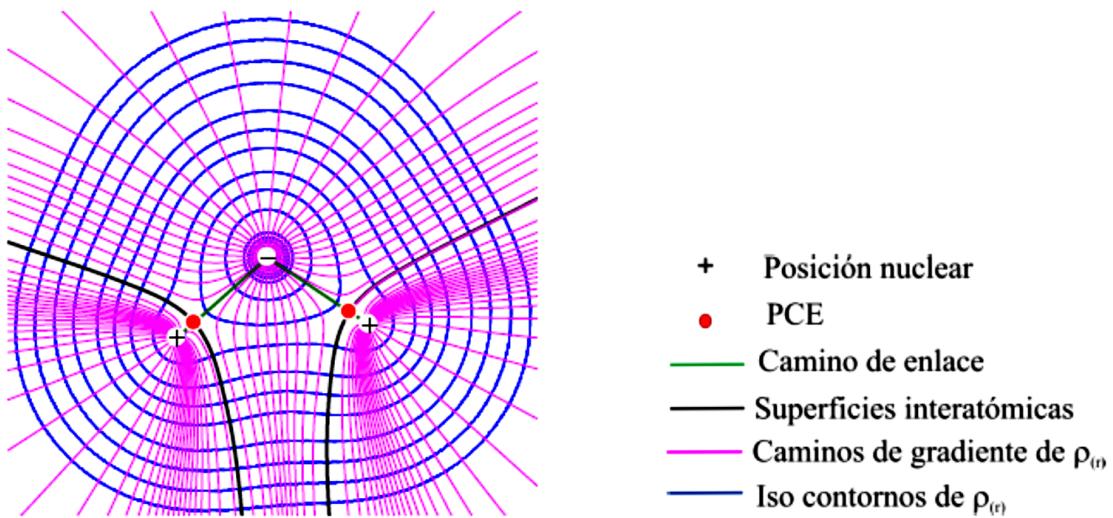


Figura 2.8: Mapa del campo vector del gradiente de la densidad electrónica para el plano que contiene los núcleos de oxígeno e hidrógeno de la molécula de agua, superpuesto con un mapa de contorno de la densidad.

Existe una región abierta del espacio que rodea a un punto crítico $(3, -3)$ tal que todas las trayectorias de gradiente que se originan en este espacio, terminan sobre el punto crítico. Así, este punto $(3, -3)$ se comporta como un atractor del campo del vector gradiente, $\nabla\rho(r)$. Como en general, los máximos de la distribución de carga coinciden con las posiciones nucleares, los núcleos actúan como atractores de la misma. El resultado de esta identificación es tal que el espacio de una distribución de carga molecular es partitionado en regiones disjuntas denominadas cuencas, cada una de las cuales contiene un punto atractor o núcleo. Un átomo (libre ó unido) se define entonces como la unión del atractor y su cuenca asociada.

Alternativamente, un átomo se puede definir en término de sus fronteras o límites. Un átomo en una molécula está separado de sus vecinos por líneas de gradiente que definen superficies interatómicas o IAS por sus siglas en inglés. Las trayectorias de gradiente que definen dicha superficie (en negro) terminan en los puntos críticos de enlace (puntos rojos). Es decir que la existencia de una superficie interatómica indica la presencia de un punto crítico $(3, -1)$ entre dos núcleos vecinos. En la figura 2.8 se muestran solo dos trayectorias de gradiente que terminan en cada punto crítico de enlace, que corresponden a las únicas dos trayectorias que se encuentran en el plano de la molécula. Pero existen infinitas trayectorias de gradiente que inician en el infinito y terminan en cada PCE, que definen las IAS. (Figura 2.9)

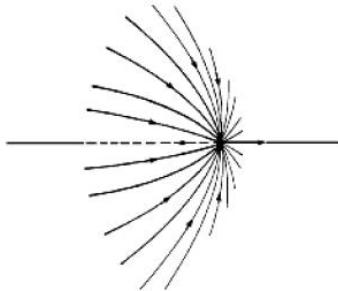


Figura 2.9: Una vista tridimensional del conjunto de trayectorias de $\nabla\rho_r$ que terminan en un punto crítico de enlace y definen una superficie interatómica y del único par de trayectorias que se originan en el mismo punto crítico y definen un camino de enlace (gráfico tomado de Bader, 1990).

Debido a las propiedades topológicas de un punto crítico (3, -1), las superficies interatómicas, $S(r)$, tienen un carácter único: son las únicas superficies continuas y cerradas que satisfacen la condición de flujo cero:

$$\nabla\rho(r) \cdot n(r) = 0 \quad (2.19)$$

Donde $n(r)$ es el vector normal a la superficie S en r . Esta condición equivale a decir que todas las trayectorias generadas en las cercanías de un dado núcleo, terminan sobre este núcleo y ninguna atraviesa la cuenca de un átomo vecino. Finalmente, los pares de caminos de gradiente que se originan en cada punto crítico (3, -1) terminan sobre los núcleos vecinos definiendo una línea a lo largo de la cual, la densidad de carga es máxima con respecto a cualquier línea vecina, esta línea es la llamada *línea de interacción atómica*. La acumulación de carga a lo largo de esta línea de interacción, es una condición necesaria para asegurar la unión entre dos átomos. También es una condición suficiente cuando las fuerzas entre los núcleos están balanceadas y el sistema posee una energía mínima a la separación internuclear de equilibrio. En este caso la línea de máxima densidad de carga entre los núcleos es llamada línea de enlace o camino de enlace, CE, y el punto crítico (3, -1) es llamado punto crítico de enlace, PCE. La red de caminos de enlaces que unen a núcleos vecinos se denomina *grafo molecular*. Se ha observado que los grafos moleculares que se obtienen de esta forma, coinciden con los que se obtendrían uniendo aquellos pares de átomos que en base a argumentos químicos se consideran enlazados.

2.4.3 Propiedades del Laplaciano de la densidad de carga

En una dimensión, la curvatura de $\rho(r)$ es una medida de la diferencia entre su valor promedio en puntos vecinos a r y su valor en r . El análogo tridimensional es el Laplaciano de $\rho(r)$. Un valor negativo para el Laplaciano en algún punto significa que la carga electrónica tiende a

concentrarse en este punto. El Laplaciano a través de las magnitudes relativas de sus tres curvaturas, determina el desequilibrio en las contribuciones de las energías cinética y potencial al virial de un sistema. Esta propiedad de relacionar las regiones de concentración y de reducción de carga con regiones en donde dominan la disminución de energía potencial y aumento de la energía cinética respectivamente constituye la base para la clasificación de las interacciones atómicas.

EL Laplaciano de la densidad electrónica, $\nabla^2\rho$, es utilizado para describir el modelo del agujero σ en EXs, ya que proporciona información sobre las regiones donde la carga electrónica está concentrada, discriminando entre "holes" (cuando $\nabla^2\rho(r) < 0$) o "lumps" (cuando $\nabla^2\rho(r) > 0$). En estrecha analogía este enfoque considera el EX como una interacción entre un *agujero* en un átomo de halógeno y un *bulto* en la base de Lewis base (Eskandari & Zariny, 2010)

Distribución Laplaciana:

El Laplaciano de la densidad de carga ($\nabla^2\rho(r)$) determina las regiones espaciales donde la densidad de carga se concentra o disminuye localmente. La topología de la distribución Laplaciana refleja la estructura de capas de los átomos aislados, existiendo una sub-capa de concentración de carga seguida de otra sub-capa de depleción de carga por cada capa cuántica del átomo. La sub-capa más externa de concentración de carga se denomina capa de concentración de carga de la capa de valencia (VSCC)(R. Bader, 1990). Cuando un átomo está involucrado en un enlace, la simetría esférica de la VSCC se rompe y aparecen regiones de máxima y mínima concentración de carga sobre esta sub-capa. Si bien los máximos en una distribución de carga aparecen sólo en las posiciones nucleares, ello no implica que la contracción de $\rho(r)$ hacia los núcleos sea dominante a todas las distancias del núcleo, ni aún en un átomo aislado.

De forma que además de las regiones en donde la contracción de $\rho(r)$ a lo largo de la línea radial hacia los núcleos es dominante y $\nabla^2\rho(r) > 0$, debe existir una región o regiones donde la magnitud de las curvaturas de $\rho(r)$ perpendiculares a esta línea radial sea dominante, y $\nabla^2\rho(r) < 0$. Así, $\nabla^2\rho(r)$ exhibe nodos esféricos en un átomo y el número de nodos está relacionado con su estructura de capa. Existen pares de regiones, una negativa y otra positiva, para cada capa cuántica principal, siendo las regiones internas, zonas de concentración de carga electrónica

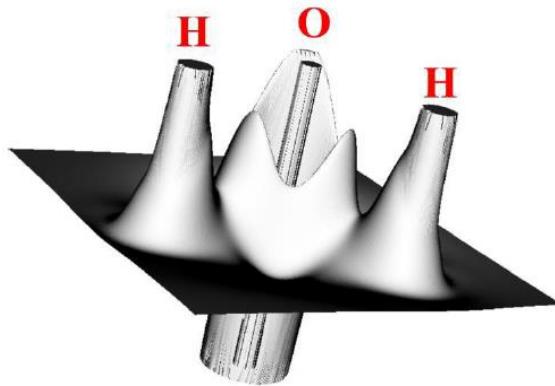


Figura 2.10: Mapa de relieve del Laplaciano de la densidad de carga electrónica de la molécula de agua en el plano que contiene sus átomos. Nótese la VSAC del átomo de oxígeno, con regiones de mayor y menor concentración de carga.

Otra de las propiedades importantes del Laplaciano de la densidad electrónica dentro de la teoría de Átomos en Moléculas es su relación con la densidad de energía. Esta relación sale inmediatamente de la aplicación del teorema del virial local (Bader, 1990; Popelier, 2000), el cual, en unidades atómicas, queda expresado de la siguiente manera:

$$2G(r) + V(r) = \frac{1}{4} \nabla^2 \rho(r) \quad (2.20)$$

Donde $G(r)$, es la densidad de energía cinética y $V(r)$ es la densidad de energía potencial. La ecuación anterior establece que en cada punto del espacio existe un balance entre la densidad de energía cinética y potencial, corregida por el valor del Laplaciano en el punto considerado. Esta densidad de energía es dimensionalmente equivalente a una fuerza por unidad de superficie, esto es, una presión. Así, el Laplaciano puede entenderse como una medida de la presión ejercida sobre la densidad de carga electrónica, relativa al valor nulo que impone la condición de flujo cero en la superficie atómica, que se puede reescribir como una condición de nulidad sobre la integral de volumen del Laplaciano extendida a toda la cuenca atómica. En las regiones en las que el Laplaciano es negativo, la densidad de carga se encuentra fuertemente enlazada y comprimida por encima de su distribución promedio. Sin embargo, en las regiones de Laplaciano positivo la densidad de carga se expande con respecto a su distribución promedio. Esto tiene una influencia decisiva en las propiedades reactivas de la molécula.

Clasificación de las interacciones atómicas

El hecho de que la función $\nabla^2\rho(r)$ es capaz de determinar las regiones de una molécula en donde la carga electrónica está concentrada o disminuida, permite dar una clasificación rigurosa de las interacciones atómicas. Así, si el valor del Laplaciano de ρ en el PCE es negativo, se tiene una interacción de capa compartida (que usualmente se corresponde con los enlaces covalentes o polares de la química tradicional). La distribución de carga en estas interacciones, está dominada por las curvaturas negativas de $\rho(r)$, perpendiculares a la dirección de la interacción interatómica. Es decir, la carga electrónica está concentrada en la región internuclear como resultado de la contracción perpendicular de $\rho(r)$ hacia la línea de interacción o camino de enlace. Esta concentración de carga se refleja en un valor relativamente grande de $\rho(r)$, y en una relación $|\lambda_1|/\lambda_3 > 1$, ambas medidas en el PCE. La región del espacio en el cual $\nabla^2\rho(r) < 0$, contiene al PCE y es contigua sobre las regiones de valencia de ambos átomos. Ellos comparten la densidad electrónica de la capa de valencia conduciendo a una región contigua de energía potencial relativamente baja que se extiende sobre la cuenca de ambos átomos.

El segundo tipo límite de interacción atómica es el que se observa cuando $\nabla^2\rho(r) > 0$, en el PCE, llamado interacciones de capa cerrada (que usualmente se corresponden con los enlaces iónicos, enlaces de hidrógeno e interacciones de van der Walls). Estas interacciones están dominadas por el principio de exclusión de Pauli. $\rho(r)$ toma valores relativamente bajos, la relación $|\lambda_1|/\lambda_3 < 1$, en el PCE. El signo del Laplaciano en estos casos, está determinado por la curvatura positiva de $\rho(r)$ a lo largo de la línea de interacción, mientras que el principio de exclusión conduce a una reducción de carga en la superficie interatómica. Estas interacciones están dominadas por la contracción de carga desde la superficie interatómica hacia cada uno de los núcleos. El Laplaciano es positivo sobre toda la región de interacción y la contribución de la energía cinética en esta región es relativamente elevada.

Por otra parte, se define la densidad de energía total en un punto como:

$$H(\mathbf{r}) = G(\mathbf{r}) + V(\mathbf{r}) \quad (2.21)$$

La densidad de energía electrónica total en un PCE, H , permite caracterizar el tipo de interacción (Arnold & Oldfield, 2000; Jenkins & Morrison, 2000). Un valor negativo de H indica que la energía potencial predomina sobre la cinética, y teniendo en cuenta la ecuación, puede interpretarse como una consecuencia de la acumulación de carga en el PCE. Al existir una

acumulación de carga, interacciones donde $|V| > G$ y $H < 0$, se dice que tienen carácter covalente, mientras que una $H > 0$ siempre es indicativo de interacciones de capa cerrada.

2.5 Análisis de Interacciones a través de SIFT

Consiste en el estudio de interacciones basándose únicamente en distancias interatómicas, siguiendo el enfoque utilizando por Deng (Deng et al., 2004). Esta metodología se basa en la construcción de una huella digital (“fingerprint”) de interacción que traduce la información de unión estructural 3D de un complejo proteína-ligando en una cadena binaria unidimensional 1D (comprendida por 1 denotando presencia y 0 ausencia de una determinada interacción). Este método utiliza el coeficiente de Tanimoto (T_c) como medida cuantitativa de la similitud de la cadena de bits (Willett et al., 1998). El T_c entre dos cadenas de bits A y B se define como:

$$T_c(A, B) = |A \cap B| / |A \cup B| \quad (2.22)$$

donde $|A \cap B|$ representa el número de interacciones o bits encendidos comunes entre los complejos A y B; y $|A \cup B|$ representa el número de interacciones o bits encendidos en uno de los complejos A o B. El valor del coeficiente de Tanimoto oscila entre 0 y 1, donde 0 representa máxima disimilitud (es decir ningún bit o tipo interacción en común) y 1 representa máxima similitud (todos los bits de la huella dactilar en común).

La representación de las interacciones como huellas dactilares permite agrupar, filtrar y perfilar bibliotecas de compuestos utilizando enfoques que están siendo ampliamente empleados en el campo de la diversidad química. SIFT puede servir como organizador molecular y filtro postdocking, ya que permite reorganizar potencialmente millones de poses de docking en función de sus patrones generales de interacción.

2.6 Aplicaciones en el Trabajo de Tesis

Los métodos descriptos de han utilizado en diferentes etapas del desarrollo del presente trabajo, en los capítulos respectivos se detallan específicamente cada una las metodologías y herramientas utilizadas, las cuales se ajustaron en base a los blancos en estudio y a los inconvenientes planteados para la correcta descripción de los Enlaces de halógeno en cada sistema.

El modelado por homología con múltiples “templates” se ha utilizado en el capítulo 5 para construir las estructuras tridimensionales de los Receptores Dopaminérgicos, las cuales no se encontraban cristalizadas, y se lograron modelar, a través de este método, en su forma activa e inactiva, siendo la obtención de estructuras representativas de estos blancos moleculares indispensables para iniciar el trabajo.

Debido a que los métodos de Mecánica Molecular no describen correctamente las propiedades de los Enlaces de Halógeno, previamente hubo que calibrar las técnicas de Dinámica y Docking Molecular para lo que se ajustaron los parámetros de cada método. La parametrización de la técnica de Docking Molecular se detalla en el capítulo 3.

Luego de optimizadas estas metodologías se aplicaron a los blancos moleculares en estudio. Así, las simulaciones de Dinámica Molecular se utilizaron para estudiar los inhibidores de la enzima Cruzipaina en el capítulo 6, mientras que la técnica de Docking Molecular se efectuó para estudiar y caracterizar los fármacos que actúan sobre los receptores dopaminérgicos en el capítulo 5.

La técnica de análisis QTAIM se utilizó a lo largo de todas las etapas de la presente tesis con diferentes objetivos, principalmente se realizó para analizar las características de los EXs en sistemas biológicos como se detalla en el capítulo 3, posteriormente en el caso de la enzima Cruzipaina se aplicó esta metodología para comparar las interacciones en inhibidores halogenados y sus análogos no halogenados, y finalmente en los receptores dopaminérgicos se evaluaron los patrones de las interacciones en compuestos agonistas y agonistas inversos.

2.7 Referencias

- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3), 403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Arcon, J. P., Defelipe, L. A., Modenutti, C. P., López López, E. D., Alvarez-Garcia, D., Barril, X., Adrián, I., Turjanski, A. G., & Martí, M. A. (2017). *Molecular Dynamics in Mixed Solvents Reveals Protein–Ligand Interactions, Improves Docking, and Allows Accurate Binding Free Energy Predictions*. <https://doi.org/10.1021/acs.jcim.6b00678>
- Arnold, W. D., & Oldfield, E. (2000). The Chemical Nature of Hydrogen Bonding in Proteins via NMR: J-Couplings, Chemical Shifts, and AIM Theory. *Journal of the American Chemical Society*, 122(51), 12835–12841. <https://doi.org/10.1021/JA0025705>
- Bader, R. (1990). *Atoms in molecules a quantum theory*. Oxford University Press.
- Bader, R. F. W., & Nguyen-Dang, T. T. (1981). Quantum Theory of Atoms in Molecules–Dalton Revisited. *Advances in Quantum Chemistry*, 14(C), 63–124. [https://doi.org/10.1016/S0065-3276\(08\)60326-3](https://doi.org/10.1016/S0065-3276(08)60326-3)
- Burley, S. K., Bhikadiya, C., Bi, C., Bittrich, S., Chen, L., Crichlow, G. V., Christie, C. H., Dalenberg, K., Di Costanzo, L., Duarte, J. M., Dutta, S., Feng, Z., Ganesan, S., Goodsell, D. S., Ghosh, S., Green, R. K., Guranovic, V., Guzenko, D., Hudson, B. P., ... Zhuravleva, M. (2021). RCSB Protein Data Bank: powerful new tools for exploring 3D structures of biological macromolecules for basic and applied research and education in fundamental biology, biomedicine, biotechnology, bioengineering and energy sciences. *Nucleic Acids Research*, 49(D1), D437–D451. <https://doi.org/10.1093/NAR/GKAA1038>
- Deng, Z., Chuaqui, C., & Singh, J. (2004). Structural Interaction Fingerprint (SIFT): A Novel Method for Analyzing Three-Dimensional Protein-Ligand Binding Interactions. *Journal of Medicinal Chemistry*, 47(2), 337–344. <https://doi.org/10.1021/jm030331x>
- Ertl, P. (2003). Cheminformatics Analysis of Organic Substituents: Identification of the Most Common Substituents, Calculation of Substituent Properties, and Automatic Identification of Drug-Like Bioisosteric Groups. *ChemInform*, 34(21), 374–380. <https://doi.org/10.1002/chin.200321198>

Eskandari, K., & Zariny, H. (2010). Halogen bonding: A lump-hole interaction. *Chemical Physics Letters*, 492(1-3), 9–13. <https://doi.org/10.1016/J.CPLETT.2010.04.021>

Foresman, J., & Frisch, E. (1996). Common ground. In *Exploring Chemistry with Electronic Structure Methods*.

Gao, J., & Serrero, G. (1999). Adipose differentiation related protein (ADRP) expressed in transfected COS-7 cells selectively stimulates long chain fatty acid uptake. *Journal of Biological Chemistry*, 274(24), 16825–16830. <https://doi.org/10.1074/jbc.274.24.16825>

Gasteiger, J., & Marsili, M. (1980). Iterative partial equalization of orbital electronegativity-a rapid access to atomic charges. *Tetrahedron*, 36(22), 3219–3228. [https://doi.org/10.1016/0040-4020\(80\)80168-2](https://doi.org/10.1016/0040-4020(80)80168-2)

Gil Redondo, R. (2010). *Universidad Autónoma de Madrid Desarrollo y Utilización de Métodos Computacionales en la Mejora del Proceso de Obtención de Nuevos Fármacos*.

Gilson, M. K., & Zhou, H. X. (2007). Calculation of protein-ligand binding affinities. *Annual Review of Biophysics and Biomolecular Structure*, 36, 21–42. <https://doi.org/10.1146/annurev.biophys.36.040306.132550>

Hansson, T., Oostenbrink, C., & Van Gunsteren, W. (2002). Molecular dynamics simulations. *Theory and Simulation*.

Hassan Baig, M., Ahmad, K., Roy, S., Mohammad Ashraf, J., Adil, M., Haris Siddiqui, M., Khan, S., Amjad Kamal, M., Provazník, I., & Choi, I. (2016). Computer Aided Drug Design: Success and Limitations. *Current Pharmaceutical Design*, 22(5), 572–581. <https://doi.org/10.2174/1381612822666151125000550>

Hislop, D., Wolfaardt, M. F., & Leach, P. G. L. (1990). The Schrodinger equation for the $f(x)/g(x)$ interaction. *Journal of Physics A: Mathematical and General*, 23(21), L1109–L1112. <https://doi.org/10.1088/0305-4470/23/21/007>

Iii, T. E. C., & Kollman, P. a. (2000). N Ucleic a Cids. *Annual Review of Physical Chemistry*, 51, 435.

Jenkins, S., & Morrison, I. (2000). The chemical character of the intermolecular bonds of seven phases of ice as revealed by ab initio calculation of electron densities. *Chemical Physics Letters*, 317(1-2), 97–102. [https://doi.org/10.1016/S0009-2614\(99\)01306-8](https://doi.org/10.1016/S0009-2614(99)01306-8)

Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993). PROCHECK: a program

to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*, 26(2), 283–291. <https://doi.org/10.1107/s0021889892009944>

Leach, A. R. (2001). *Molecular Modelling Principles*.

Lengauer, T., & Rarey, M. (1996). Computational methods for biomolecular docking. *Current Opinion in Structural Biology*, 6(3), 402–406. [https://doi.org/10.1016/S0959-440X\(96\)80061-3](https://doi.org/10.1016/S0959-440X(96)80061-3)

Lionta, E., Spyrou, G., K. Vassilatis, D., & Cournia, Z. (2014). Structure-Based Virtual Screening for Drug Discovery: Principles, Applications and Recent Advances. *Current Topics in Medicinal Chemistry*, 14.

M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ort, J. B. F. and D. J. F. (2016). *Gaussian 09* (No. A02). Gaussian Inc.

Mattos, C., & Ringe, D. (1996). Locating and Characterizing Binding Sites on Proteins. *Nature Biotechnology*, 14(5), 595–599. <https://doi.org/10.1038/nbt0596-595>

Mohan, V., Gibbs, A., Cummings, M., Jaeger, E., & DesJarlais, R. (2005). Docking: Successes and Challenges. *Current Pharmaceutical Design*, 11(3), 323–333. <https://doi.org/10.2174/1381612053382106>

Morris, G. ., Huey, R., Lindstrom, W., Sanner, M. ., Belew, R., Goodsell, D., & Olson, A. . (2009). AutoDock4 and AutoDockTools4: Automated Dockingwith Selective Receptor Flexibility. *Journal of Computational Chemistry*, 30(16), 2785–2791. <https://doi.org/DOI 10.1002/jcc>

Orry, A., & Abagyan, R. (2012). Homology Modelling. Methods and protocols. In *Humana Press*. <https://doi.org/10.1007/978-1-61779-588-6>

P, P. (2000). Atoms in Molecules. *Chemical Modelling: Applications Ans Theory*, 1(1), 144–193. <https://doi.org/10.1021/ar00109a003>

Pieper, U., Eswar, N., Webb, B. M., Eramian, D., Kelly, L., Barkan, D. T., Carter, H., Mankoo, P., Karchin, R., Marti-Renom, M. A., Davis, F. P., & Sali, A. (2009). MODBASE, a database of annotated comparative protein structure models and associated resources. *Nucleic Acids Research*, 37(SUPPL. 1), 347–354. <https://doi.org/10.1093/nar/gkn791>

Ponder, J. W., & Case, D. A. (2003). Protein Simulations. *Advances in Protein Chemistry*, 66, 27–85. <http://www.sciencedirect.com/science/article/pii/S006532330366002X>

Saldívar-González, F., Prieto-Martínez, F. D., & Medina-Franco, J. L. (2017). Descubrimiento y desarrollo de fármacos: un enfoque computacional. *Educacion Quimica*, 28(1), 51–58. <https://doi.org/10.1016/j.eq.2016.06.002>

Šali, A., & Blundell, T. L. (1993). Comparative protein modelling by satisfaction of spatial restraints. In *Journal of Molecular Biology* (Vol. 234, Issue 3, pp. 779–815). <https://doi.org/10.1006/jmbi.1993.1626>

Scior, T., Bender, A., Tresadern, G., Medina-Franco, L., Martínez-Mayorga, K., Langer, T., Cuanalo-Contreras, K., & Agrafiotis, D. K. (2012). Recognizing Pitfalls in Virtual Screening: A Critical Review. *J. Chem. Inf. Model*, 52, 34. <https://doi.org/10.1021/ci200528d>

Söding, J., Biegert, A., & Lupas, A. N. (2005). The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Research*, 33(SUPPL. 2), 244–248. <https://doi.org/10.1093/nar/gki408>

Sousa, S., Fernandes, P., & Ramos, M. J. (2006). Protein–Ligand Docking: Current Status and Future Challenges. *Proteins: Structure, Function and Genetics*, 65, 15–26. <https://doi.org/DOI 10.1002/prot>

Wiederstein, M., & Sippl, M. J. (2007). ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Research*, 35(SUPPL.2), 407–410. <https://doi.org/10.1093/nar/gkm290>

Willett, P., Barnard, J. M., & Downs, G. M. (1998). Chemical similarity searching. *Journal of Chemical Information and Computer Sciences*, 38(6), 983–996. <https://doi.org/10.1021/ci9800211>

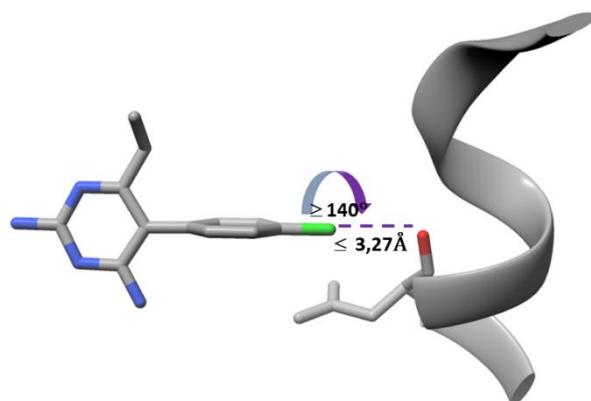
Xu, Y. Y., Cao, Y., Ma, H., Li, H. Q., & Ao, G. Z. (2013). Design, synthesis and molecular docking of α,β -unsaturated cyclohexanone analogous of curcumin as potent EGFR inhibitors with antiproliferative activity. *Bioorganic and Medicinal Chemistry*, 21(2), 388–394. <https://doi.org/10.1016/j.bmc.2012.11.031>

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>

Young, D. C. (2009). Computational Drug Design: A Guide for Computational and Medicinal Chemists. *Computational Drug Design: A Guide for Computational and Medicinal Chemists*, 1-307. <https://doi.org/10.1002/9780470451854>

Capítulo 3

Descripción y parametrización de Enlaces de halógeno en entornos biológicos



CAPITULO 3

La formación de los agujeros σ en los EXs es una consecuencia de la anisotropía de la distribución de densidad de carga alrededor del halógeno lo que supone un desafío a los actuales enfoques del modelado molecular, dado a que estos métodos consideran a los átomos de halógeno como dotados de un potencial electrostático enteramente negativo y fallan al describir correctamente los EXs. Con el fin de mejorar la descripción del EXs en el método Docking Molecular, en la primera parte de este capítulo se buscó parametrizar el agujero σ .

Como estudios previos han demostrado que los residuos de las proteínas que rodean al átomo de halógeno podrían perturbar su distribución electrónica y la fuerza del enlace X. En la segunda parte, se realizó un estudio estructural en bases de datos públicas de complejos proteína-ligando unidos a través de EXs para obtener una descripción del “EX entorno” basado en la densidad de carga, con ayuda de la Teoría Cuántica de los Átomos en Moléculas (QTAIM).

3.1 Enlaces de halógeno en métodos de modelado molecular

Los enlaces halógenos han sido reconocidos en las últimas décadas como una interacción molecular atípica, y se han utilizado en la ciencia de los materiales, así como también para la optimización de compuestos líderes en el diseño de fármacos. Como se explicó previamente, los EXs son interacciones orientadas direccionalmente, que involucran la región electrofílica del halógeno (Zimmermann et al., 2014), y el agujero σ , con carga positiva, debido a la polarización que ocurre cuando un electrón de la capa de valencia del halógeno participa en la formación de orbitales moleculares que se genera al formarse un enlace covalente (Clark et al., 2007).

Comprender que los halógenos unidos covalentemente no son simples átomos ricos en electrones es indispensable para el uso de los EX. La química de los halógenos ha sido explotada debido a que se han considerado útiles para la optimización de las propiedades ADMET ya que mejoran la absorción oral y facilitan el cruce de barreras biológicas; además, son útiles para llenar pequeñas cavidades hidrofóbicas presentes en muchos blancos moleculares como proteínas, prolongando también la vida útil del medicamento.

Un estudio estructural previo de los EX en los complejos de proteína-ligando ha revelado que el oxígeno del carbonilo perteneciente al backbone de la proteína es la base de Lewis

generalmente involucrada en los enlaces X en los sitios de unión a proteínas, probablemente debido a su presencia en cada aminoácido (Wilcken et al., 2013). Con respecto a la estructura donante de enlaces X, los sistemas aromáticos halogenados son los reconocidos como más comunes en la química medicinal.

La existencia del agujero- σ , es decir la distribución asimétrica de la densidad electrónica sobre el átomo de halógeno que interactúa con un donador de electrones supone un serio desafío a los actuales enfoques de modelado molecular en el diseño de fármacos (Docking, Dinámica Molecular, etc.) dado que estos métodos fallan en describir correctamente los sistemas enlazados por halógeno. Entender la asimetría electrónica los EXs ha sido de gran impacto para el modelado computacional de los halógenos para desarrollar nuevos fármacos en los cuales se puede predecir cómo la adición o eliminación de un átomo a un nuevo candidato a fármaco aumentará o disminuirá su afinidad de unión.

La mayoría de los átomos se, en las técnicas de modelado molecular, se tratan como objetos esféricos, con una distribución isotrópica de carga sobre su superficie, pero modelar halógenos de esta manera sería predecir incorrectamente la interacción entre el halógeno y la base de Lewis, y puede conducir al fracaso del desarrollo de nuevos fármacos. Para comprender cómo un EX particular puede contribuir a la estabilidad de un sistema, primero debemos entender la estructura-energía de los enlaces X.

Actualmente, P. S. Ho y sus colegas abordaron el tema derivando un conjunto de ecuaciones para el campo de fuerza MM / MD basadas únicamente en las propiedades físicas de átomos de halógeno unidos covalentemente. Ellos encontraron funciones para los potenciales Coulomb y Lennard-Jones que describen la carga efectiva y la forma de los átomos de Bromo en función del ángulo C – Br … Y (Carter et al., 2012). Este método fue utilizado con éxito para modelar las relaciones estructura-energía en relación con los modelos de unión de ADN estudiados por el mismo grupo.

Los métodos mecano-cuánticos describen correctamente los EXs, pero estos cálculos se vuelven prohibitivos a medida que aumenta el número de átomos en el sistema biológico como también el número de moléculas a ser evaluadas, en un “screening” virtual. Dentro del diseño de fármacos asistido por computadora, la utilización de métodos computacionalmente no-cuánticos más baratos resulta inevitable. Teniendo en cuenta tal deficiencia, en este capítulo de la tesis, se calibrará el método de Docking Molecular cuyo fundamento teórico fue descripto previamente en el capítulo 2, de modo que pueda representar apropiadamente este tipo de interacción.

Para parametrizar este método se utilizará como modelo el trabajo de Santos-Martins (2014) (Santos-Martins et al., 2014) en el cual desarrollaron un modelo direccional e independiente de la carga para que el campo de fuerza describa correctamente las interacciones de coordinación de Zinc en AutoDock4, proporcionando mayor precisión que el campo de fuerza convencional. Debido a que las interacciones se modelan de forma independiente para diferentes tipos de átomos, obteniendo potenciales específicos para cada uno, pudieron parametrizar el campo de fuerza de AutoDock para el Zinc, y así modelaron las interacciones de coordinación de este átomo con los ligandos en este caso. Para ello primero seleccionaron un conjunto de complejos con valores de afinidad experimentales; a través del análisis de este conjunto de datos obtuvieron los parámetros geométricos que fueron modificados y calibrados para ajustar el campo de fuerza y generar el nuevo denominado Auto-Dock4Zn, el cual luego fue validado.

De manera similar, en esta tesis se aprovechará la información estructural disponible para poner a punto las técnicas de modelado para EXs (parámetros del campo de fuerza, mapas-grilla de afinidad pre-calculados, función de scoring, etc.), hasta lograr reproducir las características de unión encontradas experimentalmente. Así, para la parametrización del Docking en una primera etapa ajustaremos el potencial de Lennard-Jones y luego la direccionalidad del EX. El potencial de Lennard-Jones se representa por la ecuación:

$$V_{\text{LJ}}(r) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right]$$

ϵ es la profundidad del potencial o potencial de pozo

σ es la distancia (finita) en la que el potencial entre partículas se anula

r es la distancia entre partículas.

El potencial de Lennard-Jones 12-6 varía en función de la distancia de los átomos interactuantes (Figura 3.1), en la figura se muestran los parámetros asociados al mínimo de este potencial que son la distancia interatómica de equilibrio R_{\min} (en Å) y el potencial del pozo ϵ (en kcal/mol). Estos parámetros pueden ser ajustados para reproducir datos experimentales o pueden ser deducidos de resultados muy precisos de cálculos de química cuántica.

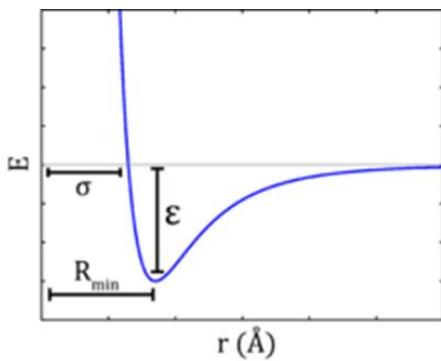


Figura3. 1: Variación del potencial de Lennard-Jones en función de la distancia de los átomos interactuantes

Luego de parametrizada la técnica de Docking Molecular para EXs, se realizará un relevamiento de los complejos ligando-proteína que contienen halógeno, específicamente halobenceno, en entornos biológicos, empleando una herramienta muy rigurosa para describir las interacciones del halógeno, la cual se basa en la distribución de densidad de carga del sistema (QTAIM). A través de esta técnica es posible detectar interacciones no direccionales como, por ejemplo, aquellas que involucran electrones π en anillos aromáticos o enlaces de hidrógeno no convencionales que involucran a donantes C-H entre otros contactos hidrófobos débiles e inusuales, que de otro modo podrían perderse en un análisis meramente geométrico de las interacciones. QTAIM no hace distinciones arbitrarias entre diferentes interacciones no covalentes (es decir, enlace de hidrógeno, contacto hidrófobo, contacto polar, etc.) ya que para todos se aplica la misma condición para la existencia de una interacción, es decir, la presencia de un camino de enlace (BCP) y los puntos críticos (BP) correspondientes que lo conectan a los átomos unidos como se describió en el capítulo 2

Posteriormente, en el capítulo 4, se aplicará este método parametrizado para la elaboración de un protocolo que permita predecir la formación de EXs basado en las interacciones de los entornos proteicos.

3.2 Procedimiento

3.2.1. Parametrización de docking molecular para compuestos halogenados

3.2.1.1 Formación del Enlace de halógeno y calibración del potencial de pozo:

- Para realizar el análisis se utilizó la base de datos Binding MOAD, Mother of All Databases (Hu et al., 2005), que contiene 41409 complejos ligando-proteína, se buscaron ligandos con Cl, Br o I unidos a un carbono aromático aplicando filtros con scripts de Jupyter notebook,, una interfaz web de código abierto que permite la inclusión de texto, video, audio, imágenes así como la ejecución de código a través del navegador en múltiples lenguajes(Kluyver et al., 2016). Posteriormente, a través de herramientas quimioinformáticas se seleccionaron aquellos complejos que forman EXs con un oxígeno carbonílico de la proteína. La distancia se definió a partir de los radios de Van der Waals de cada átomo ($d(I-Y) = 3.50 \text{ \AA}$, $d(Br-Y) = 3.37 \text{ \AA}$, $d(Cl-Y) = 3.27 \text{ \AA}$) y se consideró un ángulo ($C_{ar}-X-Y \geq 140^\circ$). Se seleccionaron aquellos complejos que tenían registrado valores de K_i y luego se los clasificó de acuerdo al Halógeno presente en el ligando (Yodo, Bromo y Cloro). Obteniendo un conjunto de 32 complejos, 3 con ligandos con yodo, 9 con bromo y 20 con ligandos clorados
- Docking Nativo: se realizó Docking con el programa Autodock4(Morris et al., 2009), de cada proteína con el ligando halogenado presente en la estructura nativa. Se calculó los valores de RMSD (desviación cuadrática media) tomando como punto de corte 2 \AA (Figura3.2).
- Parametrización: para ajustar la técnica de Docking Molecular se probó modificar ciertos parámetros, se agregó un pseudo átomo para simular el agujero-sigma, para ello el EP se colocó a lo largo de la línea del enlace covalente entre el halógeno y el carbono aromático (es decir, un ángulo Car-X-EP de 180.0°) a una distancia de 2.00 \AA del núcleo del halógeno, que es aproximadamente el radio de Van der Waals (en promedio) de los halógenos. y se definió un potencial de interacción de Lennard-Jones 12-6 de este pseudo-átomo con átomos de oxígeno para simular el enlace halógeno con el backbone de la proteína, el tipo de EX más común en sistemas biológicos. Para definir este nuevo potencial de interacción se utilizó de referencia el protocolo de trabajo descripto por (Santos-Martins et al., 2014). Se realizaron cálculos de Docking variando

los potenciales de pozo, con el fin de mejorar el rendimiento del algoritmo de docking en la descripción de las poses o modos de unión del ligando en las estructuras cristalina.

3.2.1.2 Ajuste de la direccionalidad del EX:

- ✚ Se partió de los complejos obtenidos del filtrado de la Base de datos Binding MOAD(Hu et al., 2005) utilizando todos los complejos sin la restricción de que tuvieran registrados datos de afinidad, obteniendo así un total de 86 complejos.
- ✚ Se utilizó solamente el anillo halobencenico de cada ligando (es decir, yodobenceno, bromobenceno o clorobenceno). Se agregó un pseudoátomo o extra punto (EP) en el extremo del halógeno para imitar el agujero sigma y además, el oxígeno carbonílico aceptor del enlace X designado se renombro como OX y se introdujo un potencial dentro del campo de fuerza de acoplamiento para impulsar la formación de enlaces X, el EP interactúa con el átomo de oxígeno carbonílico designado a través de un potencial de 3.00 kcal/ mol con una distancia de equilibrio de 3.00 Å (2.00 Å para d (X — EP) + 1.00 Å para d (EP..O))(Santos-Martins et al., 2014) .
- ✚ Se realizó paralelamente con los mismos complejos Docking Molecular sin la parametrización previa, los resultados de ambos se compararon.

3.2.2 Características de los entornos de los Enlaces de halógeno en complejos biológicos

- ✚ Biblioteca de complejos ligando-proteína: se utilizó el mismo conjunto de complejos ligando-proteína del paso anterior (3.2.1.2) Los complejos estructurales en la base de datos Binding MOAD se descargaron en formato PDB, para seleccionar todos los complejos de proteínas cuyos ligandos contienen como subestructura cloro, bromo o yodobenceno se utilizaron herramientas quimioinformáticas, el software Open Babel(O'Boyle et al., 2011) y Jupyter notebook (Kluyver et al., 2016)
- ✚ Preparación de las estructuras PDB: las estructuras seleccionadas se limpiaron eliminando las coordenadas redundantes, si en la estructura había más de una unidad biológica, sólo se conservaba la primera. Para las proteínas con conformaciones alternativas, solo se mantuvo la conformación A. Cuando estaban presentes dos o más cadenas de proteínas, solo se conservó la cadena A. La alineación de los complejos con respecto a las moléculas de halobenceno de referencia se realizó con la función PyMOL “ pair_fit” (Schrodinger, 2010).

La falta de átomos de hidrógeno en las estructuras PDB impone cierta incertidumbre a las ocurrencias de enlaces H en esas estructuras, por eso se empleó el software Reduce con la opción “-BUILD” que agrega hidrógenos y realiza la optimización de grupos ajustables (OH, SH, NH₃₊ , Met-CH₃, y Asn, Gln y la orientación de la cadena lateral His) (Word et al., 1999)

✚ Análisis QTAIM: De los complejos que cumplían los requisitos geométricos se prepararon modelos reducidos, incluyendo el ligando y los residuos del bolsillo de unión, conteniendo solamente los átomos que se encontraban a una distancia ≤ 4 Å del átomo de halógeno con los programas Chimera (Pettersen et al., 2004) y GaussView. Los residuos terminales en los modelos reducidos se completaron manualmente con átomos de hidrógeno. La densidad de carga se calculó mediante la metodología de la teoría del funcional de la densidad con el funcional híbrido corregido por dispersión M06-2x y el conjunto de bases DGDZVP (Siiskonen & Priimaggi, 2017) . Todos los cálculos de la estructura electrónica se realizaron con el paquete Gaussian (M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ort, 2016). El análisis topológico de la densidad de carga se realizó luego con el software Multifuncional Wavefunction Analyzer (Multiwfn) (Lu & Chen, 2012)

3.3 Resultados

3.3.1. Parametrización de Docking molecular para compuestos halogenados

3.3.1.1 Formación del Enlace de halógeno y calibración del potencial de pozo:

Durante una primera etapa se aplicó el método sin calibrar, y posteriormente se parametrizo el agujero σ para lo cual se adaptó la metodología descripta por Santos-Martins (2014), y se realizó Docking al mismo grupo de complejos en estudio para comparar los resultados obtenidos

La formación de EX se logró mediante la adición de un pseudo átomo o punto extra (EP) en el extremo del halógeno (del anillo del ligando de halobenceno) que puede interactuar con el átomo de oxígeno del carbonilo de la cadena principal. El EP imita el agujero sigma de halógeno.

La calibración se realizó variando el potencial de pozo, para obtener el que valor que más se ajusta al modelo y con el cual se obtienen los mejores resultados.

+

Docking Nativo: El rendimiento del programa de docking en cuanto a la predicción de la *pose* correcta puede ser determinada por su habilidad para reproducir los modos de unión en un conjunto de prueba de complejos ligando-proteínas. El indicador de rendimiento más comúnmente usado es el porcentaje de ligandos para el cual la solución mejor ranqueada del programa de docking está dentro de un punto de corte (generalmente 2Å) de desviación cuadrática media (RMSD) en comparación con el modo de unión experimental. Para los resultados del docking nativo, sin parametrizar, se realizó el cálculo de RMSD, observando que en un gran porcentaje de los complejos supera el valor de corte (Figura 3.2). En la figura 3.2 se puede observar que para los complejos con Bromo en la estructura del ligando el 70% obtuvo un RMSD > 2, por lo que sería conveniente la parametrización del método de docking, debido a que los resultados no serían precisos.

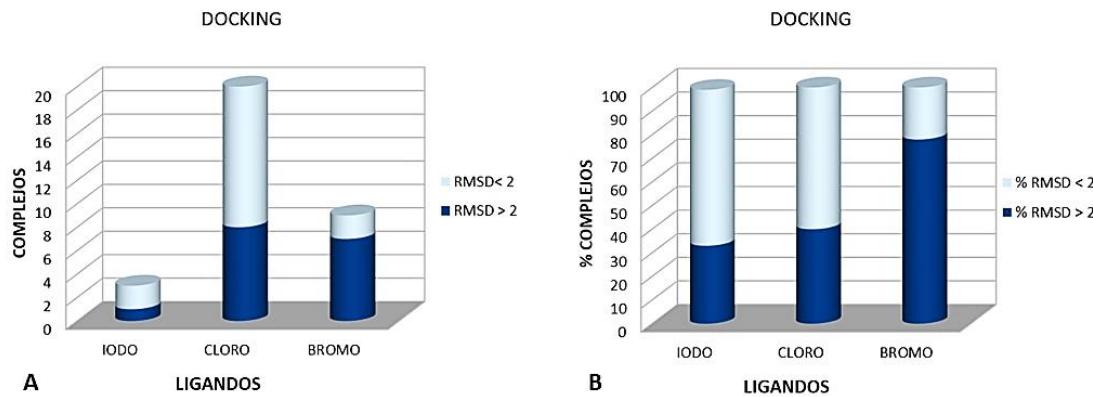


Figura 3.2: Resultados del cálculo RMSD del docking nativo de los complejos de acuerdo al tipo de halógeno del ligando. A) número absoluto de complejos B) número de complejos expresado en forma porcentual.

+

Parametrización: Se fue variando el valor de potencial de pozo en el campo de fuerza de AutoDock4, y se realizó nuevamente el cálculo de docking de los 32 complejos luego de cada ajuste; se compararon las poses predichas por el método respecto a los modos de unión cristalográficos mediante el cálculo de desviación cuadrática media (RMSD) entre las estructuras experimentales y las predichas por el campo de fuerza modificado para cada potencial (Figuras 3.3, 3.4 y 3.5). Los resultados se clasificaron según el tipo de halógeno presente en el ligando.

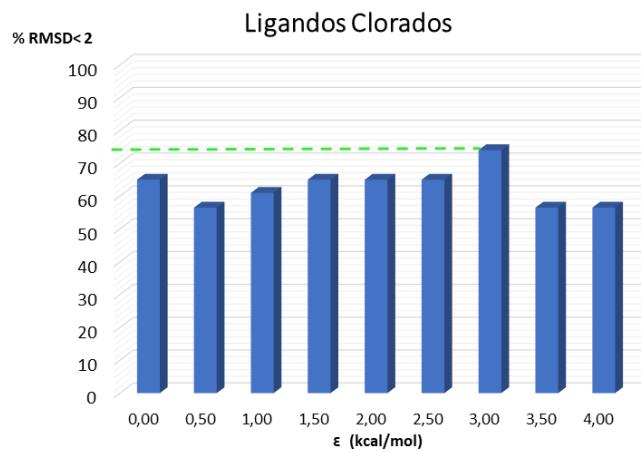


Figura 3.3: Porcentaje de complejos con ligandos clorados con $\text{RMSD} < 2 \text{ \AA}$ en función de la variación de potencial de pozo.



Figura 3.4: Porcentaje de complejos con ligandos bromados con $\text{RMSD} < 2 \text{ \AA}$ en función de la variación de potencial de pozo.

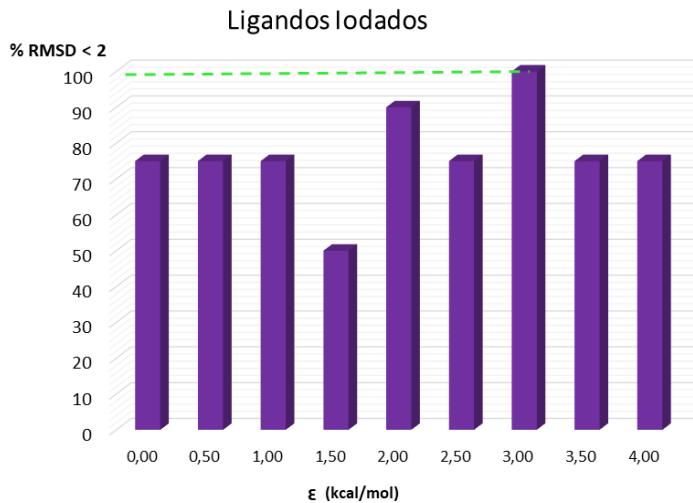


Figura 3.5: Porcentaje de complejos con ligandos iodados con $\text{RMSD} < 2 \text{ \AA}$ en función de la variación de potencial de pozo.

En la figura 3.3 se puede observar en los complejos con ligandos clorados que el mayor porcentaje de complejos con un $\text{RMSD} < 2 \text{ \AA}$ (78%) se obtienen con un potencial de pozo de 3 kcal/mol; en las estructuras que contienen bromo (figura 3.4) se presenta un 88% de complejos con $\text{RMSD} < 2 \text{ \AA}$ con los valores de potencial entre 2-3 kcal /mol. En la figura 3.5 se detalla que en el 98% de los ligandos iodados se obtuvo un valor de $\text{RMSD} < 2 \text{ \AA}$ a un potencial de 3 kcal /mol.

En base a estos resultados, se puede generalizar que hay una mejoría en las poses de docking en mayor número de complejos utilizando un potencial de pozo de 3 kcal/mol en el campo de fuerzas de AutoDock. Por este motivo en las próximas etapas se utilizará este valor.

3.3.1.2 Ajuste de la direccionalidad del EX

Un problema importante que se enfrenta en la parametrización de EX es que cualquier residuo de la proteína con un átomo de oxígeno aceptor (OA) podrá interactuar con el agujero sigma, incluso aquellos residuos de cadenas laterales como los grupos carboxilo de residuos aspártico / glutámico o hidroxilo de serinas / treonina. Esto hace que la tarea de predecir la formación de enlaces sea muy difícil. En consecuencia, se redujo el problema a un bolsillo de unión particular que el usuario está interesado en apuntar con un anillo de halobenceno (por ejemplo,

para mejorar la afinidad del ligando). De ese modo, se obliga a formar el EX deseado (Figura 3.6)

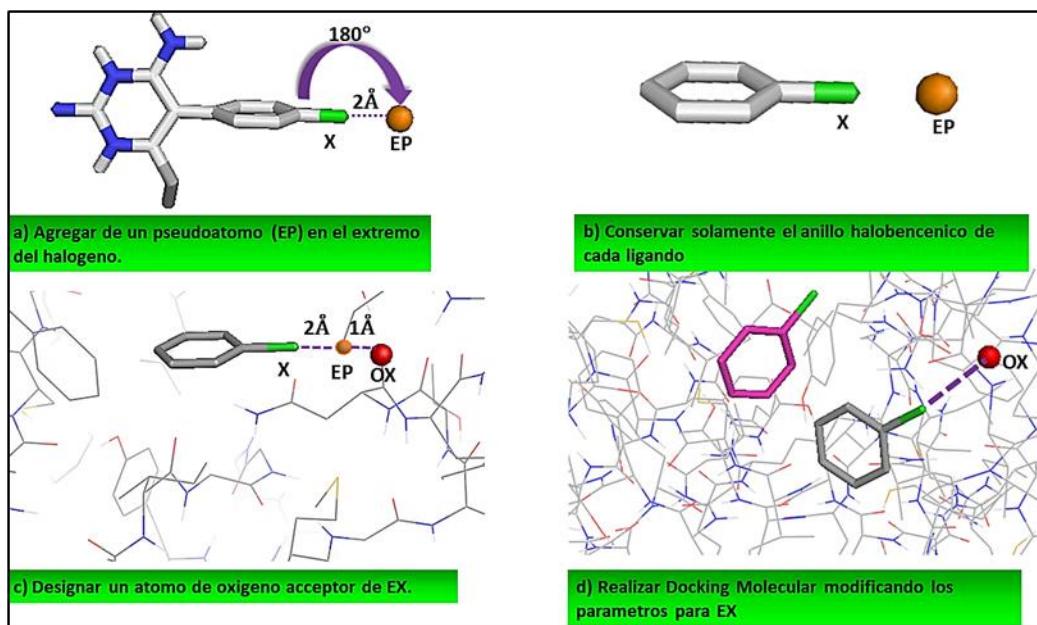


Figura 3.6: Esquema del procedimiento para la calibración de EX en el Docking Molecular, a) Se agrega un pseudoátomo que imita el agujero σ del halógeno a una distancia de 2.00 Å, b) para simplificar, se utiliza solo el anillo de halobenceno del ligando, c) se selecciona qué átomo de oxígeno de carbonílico dentro del bolsillo de la proteína formará el enlace X y se cambia el nombre de O a OX, d) se puede ver el resultado de la pose de Docking Molecular utilizando un potencial de pozo especial que se introduce dentro del campo de fuerza para impulsar la formación de enlaces X (gris) y la pose de Docking realizado sin parametrizar EX (magenta)

Como de la parametrización del potencial del pozo en la etapa anterior se observó que para los tres tipos de halógenos se obtiene un mejor resultado utilizando un potencial de 3 kcal/mol, para los cálculos de docking en este paso se utilizó este valor de potencial.

Comparando con los resultados del Docking Molecular sin parametrizar se pudo observar una mejora en la predicción del enlace X cuando se agrega el potencial y se direcciona el enlace a un oxígeno específico (OX) (Figura 3.7), observando que el EX se mantiene estable, y que las poses obtenidas son similares a las de las estructuras cristalinas.

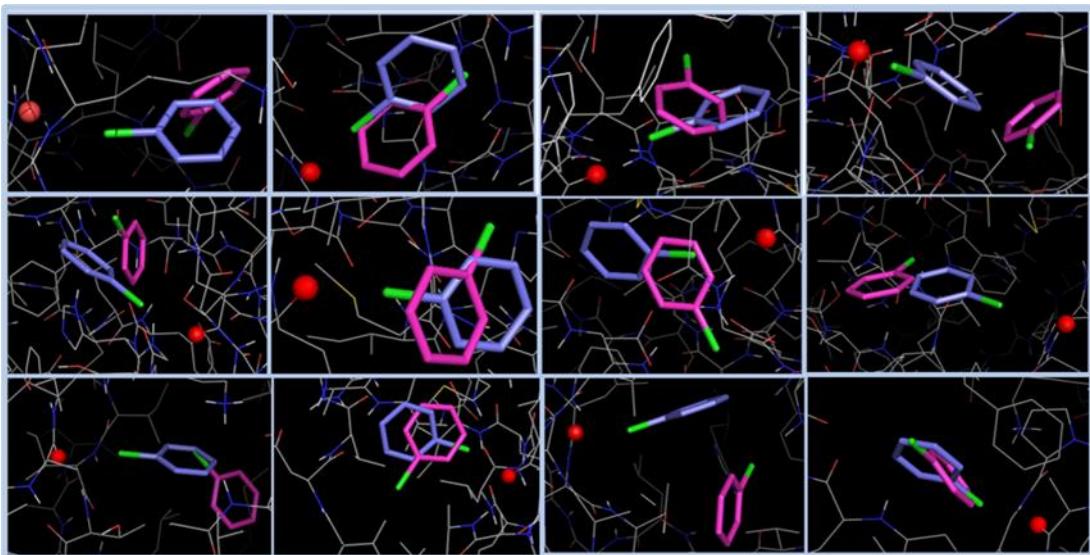


Figura 3.7: Comparación de los resultados obtenidos de Docking Molecular parametrizado para EXs (gris) y las poses de Docking realizado sin parametrizar (magenta). Como una esfera color rojo se puede observar el átomo de oxígeno aceptor del enlace halógeno (OX).

3.3.2 Características de los entornos de los Enlaces de halógeno en complejos biológicos

En una segunda parte de este capítulo, se realizó un relevamiento de las interacciones presentes en entornos biológicos y su influencia en la formación de EXs, basado en el análisis QTAIM, con este fin se siguió el procedimiento detallado en la figura 3.8.

Con este objetivo se realizó el cálculo de densidad electrónica de 86 complejos, 20 con bromobenceno en su estructura, 9 incluyen yodobenceno y 57 clorobenceno. Se obtuvieron también los datos de otros elementos topológicos que se utilizaron para el análisis del entorno (Anexo A)

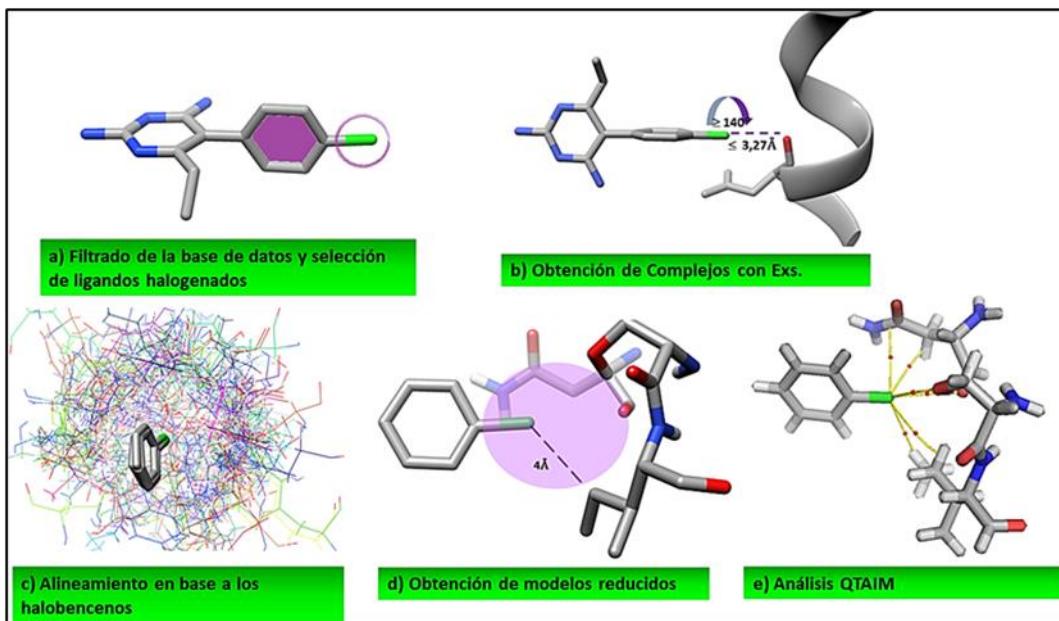


Figura 3.8: Esquema de trabajo, a) se analizó la base de datos Binding Moad (Hu, 2005), se seleccionaron ligandos con un halógeno unido a un carbono aromático, b) Se filtraron los complejos que forman EX al oxígeno de un grupo carbonilo, c) Los complejos fueron alineados por el anillo halobencenico del ligando, d) Se construyeron modelos reducidos de los complejos para calcular la densidad de carga, y e) Análisis de densidad de carga

Las figuras 3.9 y 3.10 muestran una superposición de elementos estructurales y topológicos, respectivamente, para los modelos reducidos alineados. Las esferas que rodean el átomo de halógeno en la figura 3.10 son los puntos críticos de enlace (BCP) de las interacciones de halógeno con los átomos de proteínas. El color de la esfera indica el tipo de átomo que interactúa, mientras que el tamaño de la esfera es proporcional al valor de densidad de carga asociado a la interacción.

En la figura 3.10 se puede observar que los átomos que interactúan en mayor proporción son los átomos de hidrógeno (esferas blancas) formando posiblemente Enlaces de Hidrógeno (EH), el cual se define como una interacción intermolecular de tipo $X-H\cdots Y$ en la cual debe existir evidencia de enlace e involucrar al átomo de H unido covalentemente a X (Pimentel, 1960), en los EHs convencionales X e Y son átomos con electronegatividades mayores al H, que poseen pares electrónicos libres (por ejemplo O, N, F, Cl). El grupo X-H es el *dador de protón* e Y es el *aceptor de protón*.

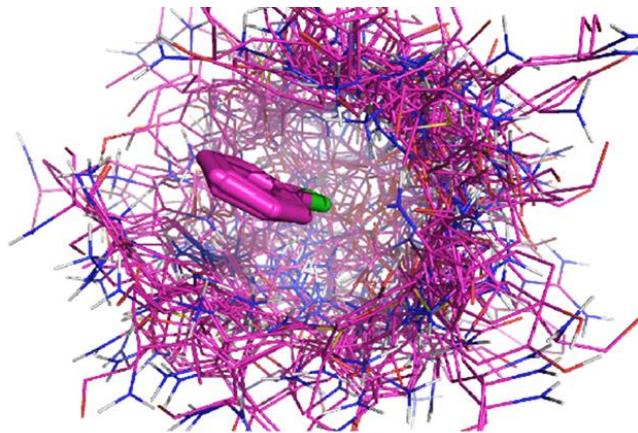


Figura 3.9: Entorno atómico del cloro en los bolsillos de unión a proteínas. La figura muestra la superposición de elementos estructurales de varios complejos de clorobenceno-proteína que previamente estaban alineados por el resto clorobenceno. La imagen es similar a un nido muestra los residuos de proteínas que interactúan con el cloro en los complejos superpuestos.

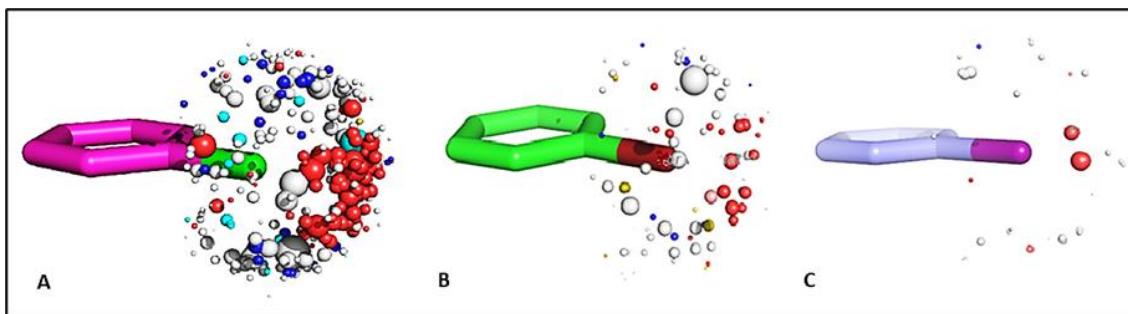


Figura 3.10: Entorno de los halógenos en los bolsillos de unión a proteínas, en la figura A se ve los complejos con Cloro, en B los complejos con ligandos bromados y en C con ligandos iodados. La figura muestra la superposición de los elementos topológicos relacionados con las interacciones de cada halógeno con residuos de las proteínas. La imagen muestra la distribución de los puntos críticos de densidad de carga asociados a las interacciones del halógeno con los átomos de proteínas, estos se representaron con esferas, el color de la esfera indica el tipo de átomo que interactúa con el halógeno (rojo: oxígeno, blanco: hidrógeno, celeste: carbono, azul: nitrógeno, amarillo: azufre) y el tamaño de la esfera es proporcional al valor de densidad de carga asociado a la interacción respectiva.

La figura 3.10 proporciona una descripción general de los entornos de enlace X en proteínas. Las siguientes conclusiones pueden observarse:

- ✚ Los átomos aceptores de enlaces X (es decir, las esferas rojas correspondientes a los átomos de oxígeno de carbonilo) se encuentran cerca del eje formado por el enlace covalente del halógeno ($\theta (C - Cl \cdots O) \sim 180^\circ$).

- Entre los diferentes átomos que rodean al halógeno en entornos proteicos, los hidrógenos son el tipo de átomo más común involucrado en las interacciones con el halógeno (esferas blancas). Los enlaces C-H de cadenas laterales de residuos hidrofóbicos como Leu, Ile y Tyr son los donantes de enlaces H más comunes. Los enlaces C-H del esqueleto de la proteína, especialmente de los residuos de Gly, también son comunes entre los donantes de enlaces H (Figura 3.11).
- Los dadores s de enlace H, a diferencia de los aceptores de enlace X, tienden a ser perpendiculares al eje de enlace covalente de halógeno (θ (C - Cl \cdots H) \sim 90 °).
- Al inspeccionar el tamaño de las esferas (que es proporcional al valor de densidad de carga en el BCP de la interacción), los enlaces X en general son más fuertes que las interacciones de enlaces H.
- Sin embargo, aunque un halógeno puede formar solo una o como máximo dos interacciones de enlaces X en un complejo proteína-ligando, generalmente forma varios enlaces H débiles que juntos realizan una gran contribución a la fuerza de anclaje del ligando.

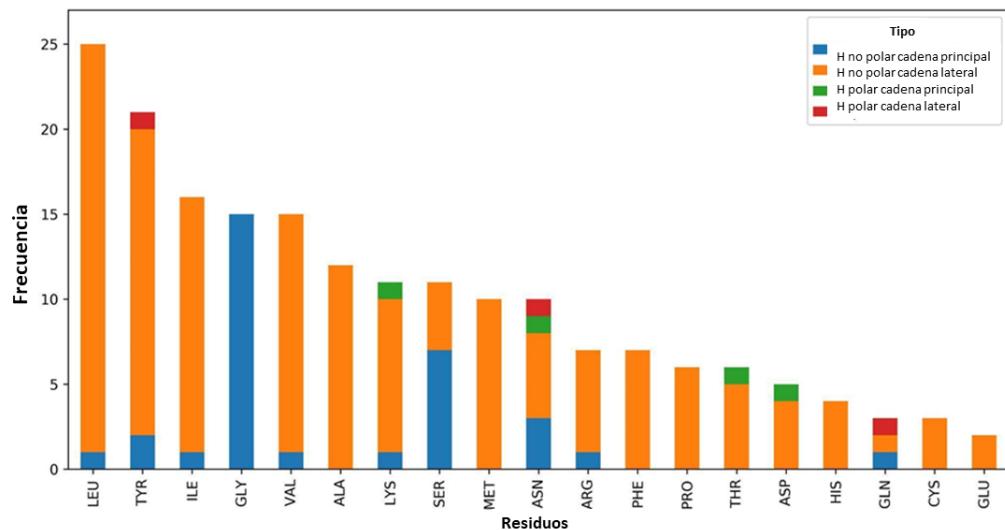


Figura 3.11: Frecuencia de residuos que actúan como donantes de enlaces H contra el Cloro en complejos con enlaces X. Los átomos de hidrógeno se consideran no polares o hidrofóbicos si están unidos al carbono, de lo contrario son polares. Los hidrógenos de la cadena principal de la proteína son aquellos unidos al átomo de N del enlace peptídico (polar) o al C_α (no polar).

En resumen, estos hallazgos sugieren que la formación de enlaces X requieren un entorno proteico adecuado para poder establecerse. Debería haber dadores de enlace H colocados cerca perpendicularmente al eje de enlace covalente C_{ar}-Cl. Debido a que presumiblemente, estas

interacciones de enlaces H mejorarían la polarización de la densidad de carga del halógeno fuera de la región axial, dando como resultado un refuerzo del enlace X.

Si los mismos dadores de enlaces H están fuera de lugar formando un ángulo entre 120° y 140° con respecto al eje de enlace covalente $C_{ar}\text{-}X$, es probable que no puedan polarizar el halógeno y el enlace X no se formaría. Las afirmaciones anteriores están respaldadas por la figura 3.12 que muestra con más detalle la distribución de los donantes de enlaces H y aceptores de enlaces X alrededor del halógeno. Como se puede ver en el gráfico de densidad en la de la figura 3.13, el número de donantes de enlaces H alcanza un máximo en la región de la meseta entre $90^\circ < \theta$ ($C - \text{Cl} \cdots \text{H} < 120^\circ$), pero de repente cae por $\theta (C - \text{Cl} \cdots \text{H}) > 120^\circ$ y $\theta (C - \text{Cl} \cdots \text{H}) < 90^\circ$. Si bien está claro a partir de esta distribución cuáles son las orientaciones preferidas de los donantes de enlaces H alrededor del átomo de halógeno en los complejos proteicos unidos a X, es menor el origen de esa preferencia. Como se argumentó anteriormente, una posibilidad es que el desplazamiento del donante de enlaces H no polarice el halógeno y, por lo tanto, el enlace X formado sea demasiado débil como para mantener una interacción estable. Por otro lado, existe la posibilidad de que algún otro tipo de efecto electrónico pueda estar operando para evitar la formación de enlaces X cuando los donantes de enlaces H se colocan en la región del ángulo intermedio.

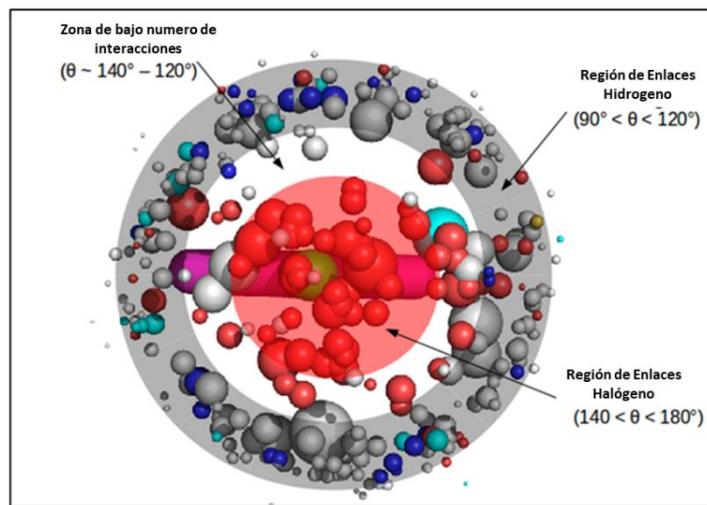


Figura 3.12: Distribución de enlaces H y enlaces X alrededor del átomo de halógeno. Dos regiones angulares alrededor del átomo de cloro están más densamente ocupadas por los donantes de enlace H (gris) y los aceptores de enlace X (rojo), con una región intermedia con un menor número de interacciones.

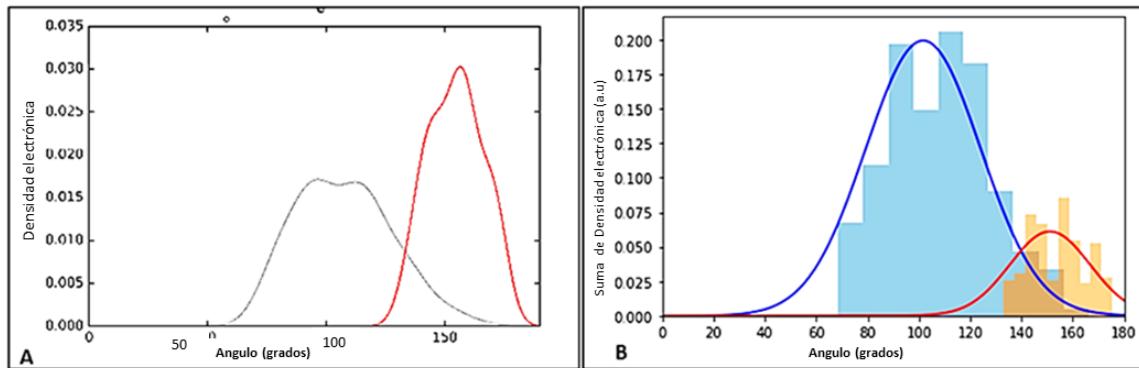


Figura3. 13: A) Gráfico de la densidad de BCP (en a.u.) en función del ángulo EH (gris) y EX (rojo). B) Suma de la densidad de carga del EH (azul) y de EX (rojo) en función del ángulo.

Como señalan Zimmermann (Zimmermann et al., 2015), las geometrías de interacción verdaderamente óptimas se observan muy raramente en el contexto biológico. A menudo, las interacciones pueden competir entre sí, lo que hace que el modo de enlace se desvíe de la geometría óptima. Sin embargo, las figuras 3.12 y 3.13 muestran que los aceptores de EX y los donantes de EH se distribuyen en promedio con el mismo patrón cooperativo que uno esperaría en un entorno menos complejo. En otras palabras, incluso en el entorno de proteínas desordenado, los donantes de EH tienden a unirse perpendicularmente al eje de enlace $C_{ar}\text{-}X$ para contactar el cinturón rico en electrones lateral del halógeno, mejorando así la polarización del halógeno y la fuerza de unión del halógeno. Este es un hallazgo notable que sugiere que un entorno proteico adecuado es una condición necesaria para poder establecer el enlace X.

3.3.2.1 Análisis de densidad electrónica en Entornos EXs: con el fin de evaluar la importancia de la distribución de los Enlaces de Hidrogeno (EH) alrededor del Halógeno, se procedió a utilizar descriptores basados en densidad de carga para analizar la distribución de EH y se realizó el siguiente análisis.

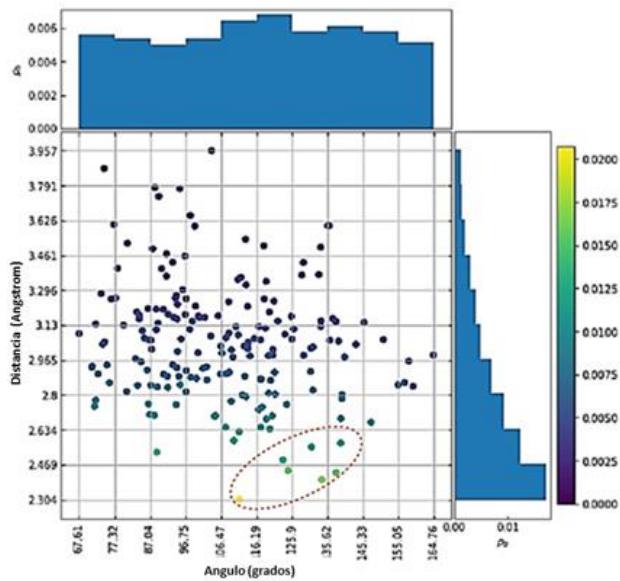


Figura 3.14: Muestra la distribución de densidad electrónica de los donantes de enlaces H alrededor del átomo de halógeno en función de la distancia y el ángulo del enlace H.

En la figura 3.14 se puede observar que la densidad electrónica ρ_b muestra una fuerte dependencia con la distancia de interacción, cuanto más corta es la distancia, mayor es el valor de ρ_b . Sin embargo, la densidad de carga en sí misma no es capaz de predecir la preferencia angular por la unión de Hidrógeno alrededor del halógeno.

A diferencia de la densidad electrónica el Laplaciano de la densidad de carga ($\nabla^2\rho(r)$) es una función más sensible pero lamentablemente, $\nabla^2\rho$ también falla al describir la preferencia angular del enlace H alrededor del halógeno, como se evidencia en la Figura 3.15.

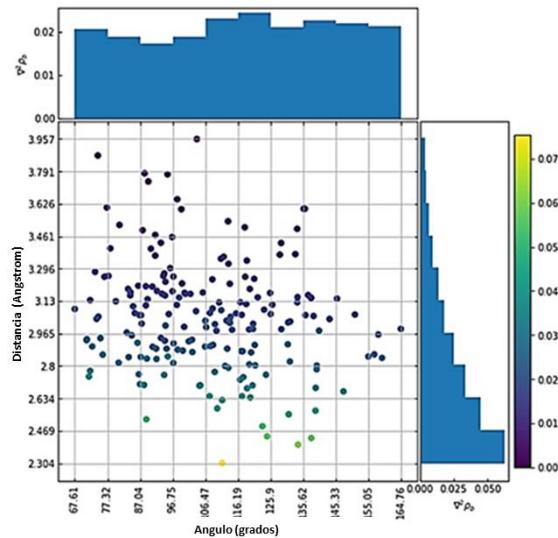


Figura 3.15: Distribución de los donantes de enlaces H alrededor del átomo de halógeno en función del ángulo y la distancia del enlace H. Los puntos de dispersión se colorearon de acuerdo con los valores de $\nabla^2\rho$ (en unidades atómicas).

Como $\nabla^2\rho$ se puede descomponer en las tres curvaturas principales de la densidad con respecto a los tres ejes principales como se observa en la Figura 3.16, λ_3 es el valor propio positivo correspondiente a la curvatura de la densidad de carga a lo largo de la línea de interacción, por lo que podría capturar la dependencia radial de los enlaces H alrededor del halógeno.

Los dos primeros valores propios negativos λ_1 y λ_2 describen las curvaturas de la densidad de carga en las direcciones perpendiculares a la línea de interacción, por lo que podrían capturar la dependencia angular de los enlaces H alrededor del halógeno. El vector propio asociado a λ_2 apunta en la dirección de la disposición de máximos en forma de toroide en el VSCC (concentración de carga de capa de valencia) del halógeno.

Como λ_1 apunta hacia la región axial del agujero σ , indica que las curvaturas de la densidad de carga en los BCP (Puntos críticos de enlace) de los EH que involucran al átomo de halógeno son lo suficientemente sensibles y representa la preferencia angular que nos interesa (Figura 3.17).

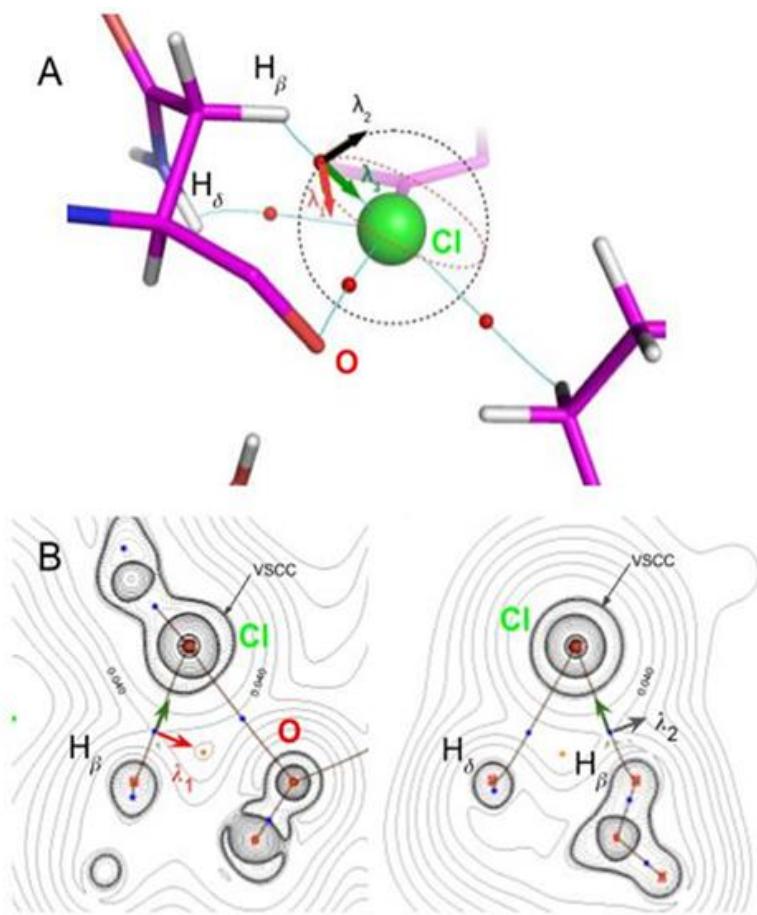


Figura 3.16: A) Curvaturas principales de la densidad de carga (vectores propios de $\nabla^2\rho$) en el BCP de un enlace H-Cl típico (PDB: 2bla). B) Gráficos de contorno de $\nabla^2\rho(r)$ en planos subtendidos por vectores propios λ_1 y λ_3 (izquierda) y vectores propios λ_2 y λ_3 (derecha).

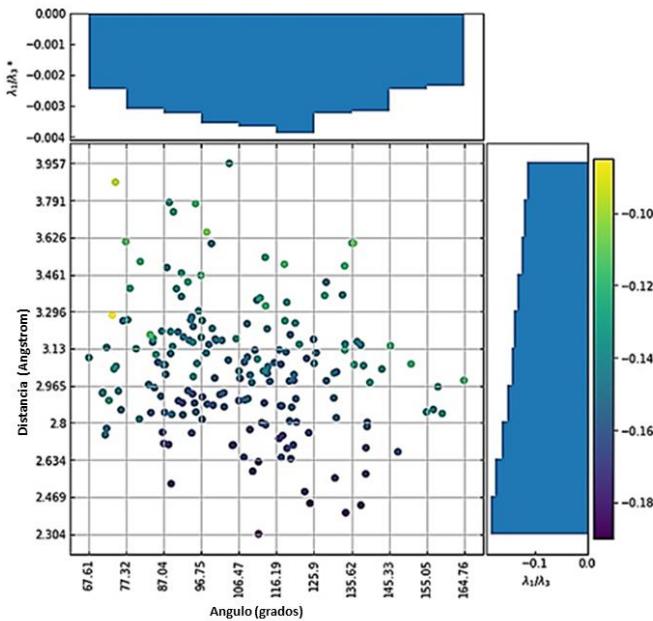


Figura 3.17: Distribución de los donantes de enlaces H alrededor del átomo de halógeno en función del ángulo y la distancia del enlace H. Los puntos de dispersión se colorearon de acuerdo con la relación λ_1 / λ_3 (en unidades atómicas).

La división de λ_1 por λ_3 tiene el efecto de debilitar la dependencia radial de λ_1 , así como de aumentar su dependencia angular. Como se puede ver en el gráfico de dispersión, Figura 3.17, hay un área de valores más negativos de la relación λ_1 / λ_3 en la región angular central entre $\sim 87^\circ$ y $\sim 130^\circ$. Esto se evidencia más claramente en el gráfico de barra marginal superior.

3.3.2.2 Fuerza del entorno de EX: Una de las características más atractivas de las propiedades topológicas locales de $\rho(r)$ es la aditividad. Por ejemplo, la integración de una determinada propiedad topológica local derivada de ρ dentro de la región molecular delimitada por superficies de flujo cero en el campo vectorial de gradiente de $\rho(r)$ da una estimación a nivel de átomo de esa propiedad (es decir, propiedades atómicas). Del mismo modo, la suma de los valores de $\rho(r)$ en todos los BCP intermoleculares ($\sum \rho_b$) se ha relacionado con la fuerza de anclaje de una molécula de ligando en el bolsillo de unión a la proteína (Tosso et al., 2017). Aquí se explotó esa propiedad aditiva de $\rho(r)$ para analizar las fortalezas de los entornos de EX, sumando los valores de las propiedades derivadas de ρ medidas en los BCP del enlace H Cl...H

La Figura 3.18 muestra la contribución relativa de los enlaces H y el enlace X a la fuerza de anclaje general del halógeno al bolsillo de la proteína. Sorprendentemente, el entorno del enlace

X (EXent), es decir, los enlaces H establecidos por el átomo de cloro, proporcionan un anclaje aún más fuerte al bolsillo de la proteína que el EX real, en la mayoría de los casos. Las figuras 3.19 y 3.20 muestran que la observación anterior también es válida para el EXent del bromo y el yodo. Sin embargo, debido a la limitación en el número de complejos con enlace X de estos halógenos, el análisis adicional del EXent y la posterior implementación del protocolo de predicción del enlace X se limitaron al cloro.

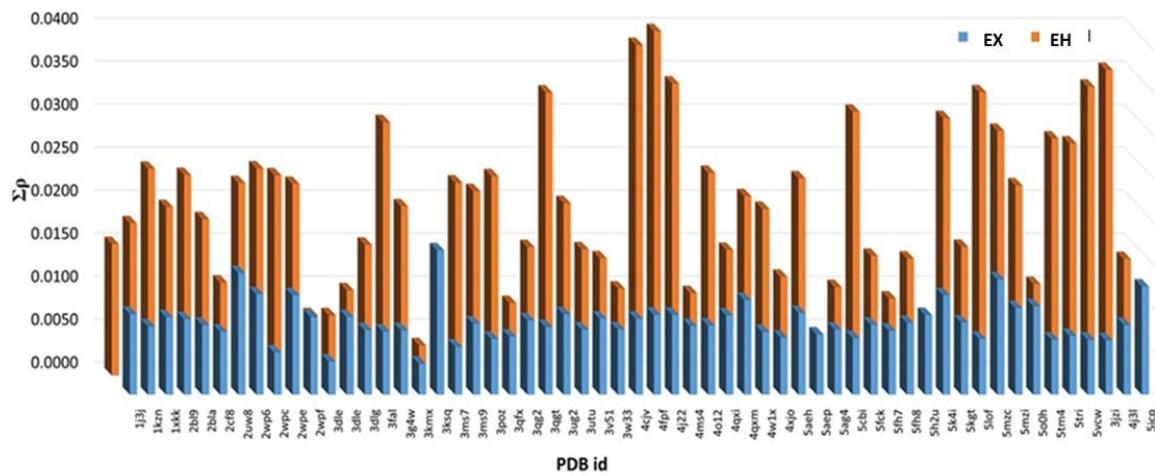


Figura 3.18. Valores de $\sum\rho$ (en unidades atómicas) para las interacciones intermoleculares del átomo de cloro en complejos con enlace X/H

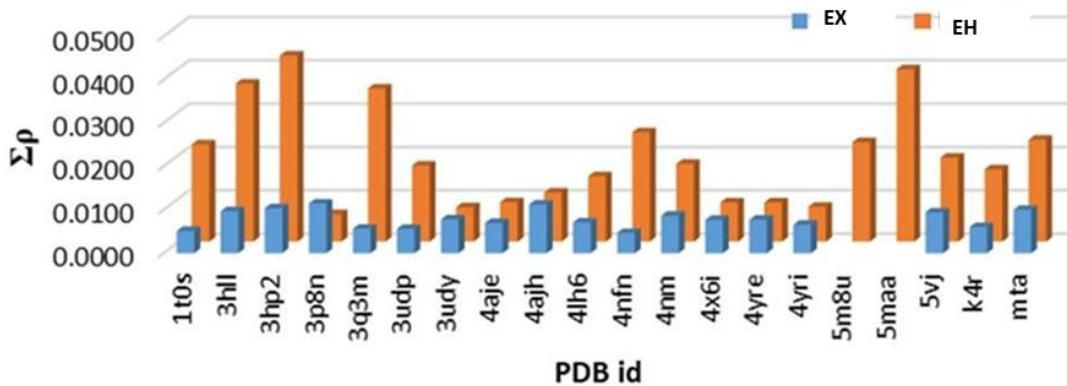


Figura 3.19. Valores de $\sum\rho$ (en unidades atómicas) para las interacciones intermoleculares del átomo de Bromo en complejos con enlace X/H

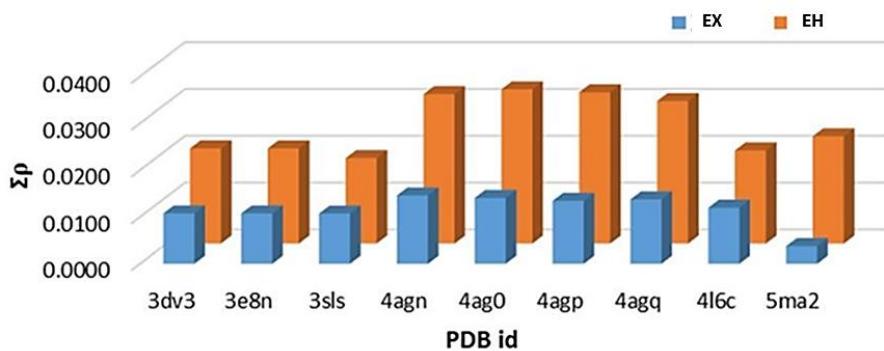


Figura 3.20: Valores de $\sum\rho$ (en unidades atómicas) para las interacciones intermoleculares del átomo de Yodo en complejos con enlace X/H

Trabajos previos han demostrado que un solo EH intramolecular con el cinturón rico en electrones del halógeno puede mejorar el potencial de en EX tanto en receptores de moléculas pequeñas (Riel et al., 2018) como en proteínas modificadas (Carlsson et al., 2018). Por otro lado; ya se ha demostrado anteriormente en el punto 3.3.2 que los átomos de hidrógeno de la proteína se distribuyen alrededor del halógeno de tal manera que interactúan principalmente con el cinturón rico en electrones del halógeno. Por lo tanto, incluso en los entornos de proteínas, uno podría esperar un patrón cooperativo similar entre los EH y los EX como se observa en entornos menos complejos.

Para verificar esa suposición, se calculó las propiedades derivadas de la densidad de carga en el BCP del enlace X, habiendo eliminado previamente los átomos del entorno de los complejos estudiados (es decir, solo el halobenceno y la N-metilacetamida involucrados en el EX se mantuvieron para el cálculo de la densidad de carga).

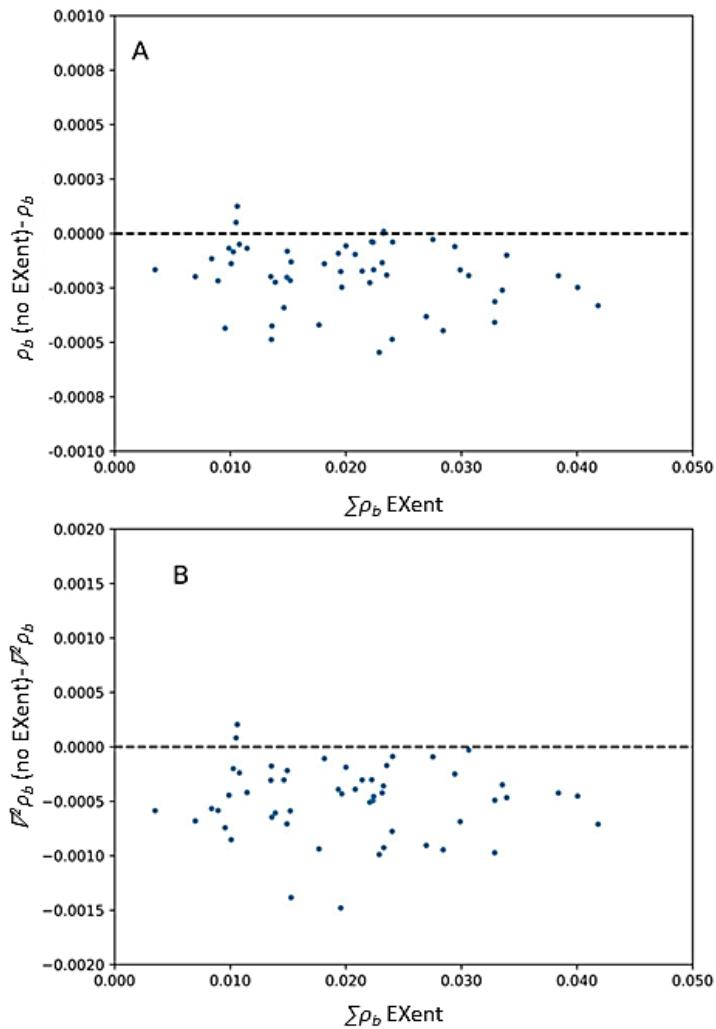


Figura 3.21: Gráficas de diferencias (con y sin considerar el EXent) de ρ_b (A) y $\sqrt{\rho_b}$ (B) en el BCP del enlace X, en función de la fuerza del EXent ($\sum \rho_b$)

La Figura 3.21 muestra el cambio en los valores de ρ_b y $\sqrt{\rho_b}$ en el BCP del enlace X cuando se eliminan las interacciones EXent. A medida que el entorno se *apaga*, se observa una disminución (casi) sistemática en ρ_b (Figura 3.21 A), lo que confirma el papel clave del entorno proteico para la mejora del enlace X. Por otro lado, $\sqrt{\rho_b}$, que es una propiedad más sensible a los entornos químicos, casi duplica la disminución de ρ_b tras la eliminación de EXent (Figura 3.21B). Los valores más pequeños de $\sqrt{\rho_b}$ en los complejos sin EXent probablemente reflejen un agotamiento de la densidad electrónica menos pronunciado en la región del agujero σ , debido a un halógeno menos polarizado.

Por otro lado, a pesar de la clara mejora del EX por EXent, no se puede observar una relación lineal entre esta mejora y $\sum \rho_b$ (eje x). Esto probablemente se deba a algunos enlaces H que no están posicionados correctamente alrededor del halógeno para polarizarlo aún más, pero aun así contribuyen al anclaje del halógeno.

En resumen, se pudo observar que el EXent opera de dos formas para estabilizar el enlace X, mediante la polarización adicional del halógeno (que a su vez aumenta el tamaño del orificio σ y la fuerza del enlace X) y mediante el anclaje adicional del donante del enlace X al bolsillo de proteínas

3. 4 Conclusiones

Como los métodos de modelado molecular no describen correctamente la anisotropía electrónica de los EXs una primera parte de este capítulo fue dedicado a parametrizar el EX para poder aplicar Docking Molecular con compuestos halogenados, con este fin se amplió el campo de fuerzas estándar de AutoDock para incluir un potencial especializado que describe las interacciones de EXs. Este potencial describe tanto la energía como los componentes geométricos de la interacción. Los experimentos de Docking demostraron que dicha modificación proporciona una mejora considerable en el rendimiento en comparación con el campo de fuerzas convencional de AutoDock4. Además, luego se ajustó la direccionalidad del enlace de halógeno, marcando el posible átomo de oxígeno aceptor. Con los parámetros ajustados la descripción del enlace mejora y este se mantiene estable, representando poses similares a la de los complejos en su estructura cristalina.

El modelado de estos parámetros fue posible debido a la naturaleza totalmente configurable de AutoDock4, por lo que el nuevo potencial se pudo implementar modificando las grillas. Además, se puede destacar que el potencial en sí mismo no agrega ninguna sobrecarga al cálculo, por lo tanto, la velocidad de Docking no se ve afectada en absoluto.

En la segunda parte de este capítulo, reconociendo que los átomos de halógeno en el bolsillo de proteína pueden formar otras interacciones no covalentes además del enlace X, se estudió el papel del entorno hidrófobo del halógeno en los complejos ligando-receptor. Es decir, centramos nuestro estudio en lo que denominamos el "*entorno de enlaces de halógeno*" (EXent), los resultados evidenciaron la importancia de los enlaces de Hidrógeno, los cuales representan las interacciones mayoritarias y se ha demostrado su posible implicancia en la fortaleza de la formación de EXs.

Para describir con precisión el entorno de los EXs se realizó un relevamiento sobre 86 complejos ligando-proteína unidos a halógenos inspeccionando no solo el enlace X sino también otros tipos de interacciones no covalentes que forma el halógeno en el contexto de la proteína, se calculó la densidad de carga de los complejos y se utilizó descriptores topológicos QTAIM para medir la fuerza de ese entorno.

De los elementos topológicos analizados se ha demostrado que el descriptor λ_1/λ_3 es el que mejor representa la preferencia angular y la dependencia con la distancia de los enlaces de hidrógeno.

Estos cálculos confirmaron la suposición de que el *entorno de enlace X* hace una gran contribución a la fuerza de anclaje del halógeno, incluso mayor que el enlace X por sí mismo.

3.5 Referencias:

- Carlsson, A. C. C., Scholfield, M. R., Rowe, R. K., Ford, M. C., Alexander, A. T., Mehl, R. A., & Ho, P. S. (2018). Increasing Enzyme Stability and Activity through Hydrogen Bond-Enhanced Halogen Bonds. *Biochemistry*, 57(28), 4135–4147. <https://doi.org/10.1021/acs.biochem.8b00603>
- Carter, M., Rappe, A., & Ho, P. S. (2012). Carter2012.pdf. *Journal of Chemical Theory and Computation*, 2461–2473.
- Clark, T., Hennemann, M., Murray, J. S., & Politzer, P. (2007). Halogen bonding: The σ -hole: Proceedings of “Modeling interactions in biomolecules II”, Prague, September 5th-9th, 2005. In *Journal of Molecular Modeling* (Vol. 13, Issue 2, pp. 291–296). <https://doi.org/10.1007/s00894-006-0130-2>
- Hu, L., Benson, M. L., Smith, R. D., Lerner, M. G., & Carlson, H. A. (2005). Binding MOAD (Mother of All Databases). *Proteins: Structure, Function and Genetics*, 60(3), 333–340. <https://doi.org/10.1002/prot.20512>
- Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B., Bussonnier, M., Frederic, J., Kelley, K., Hamrick, J., Grout, J., Corlay, S., Ivanov, P., Avila, D., Abdalla, S., & Willing, C. (2016). Jupyter Notebooks—a publishing format for reproducible computational workflows. *Positioning and Power in Academic Publishing: Players, Agents and Agendas - Proceedings of the 20th International Conference on Electronic Publishing, ELPUB 2016*, 87–90. <https://doi.org/10.3233/978-1-61499-649-1-87>
- Lu, T., & Chen, F. (2012). Multiwfn: A multifunctional wavefunction analyzer. *Journal of Computational Chemistry*, 33(5), 580–592. <https://doi.org/10.1002/jcc.22885>
- M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ort, J. B. F. and D. J. F. (2016). *Gaussian 09* (No. A02). Gaussian Inc.
- Morris, G. ., Huey, R., Lindstrom, W., Sanner, M. ., Belew, R., Goodsell, D., & Olson, A. . (2009). AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *Journal of Computational Chemistry*, 30(16), 2785–2791. <https://doi.org/10.1002/jcc>
- O’Boyle, N. M., Banck, M., James, C. A., Morley, C., Vandermeersch, T., & Hutchison, G. R. (2011).

Open Babel: An open chemical toolbox. *Journal of Cheminformatics*, 3(33), 33. <https://jcheminf.biomedcentral.com/track/pdf/10.1186/1758-2946-3-33> <http://www.ncbi.nlm.nih.gov/pubmed/21982300> <http://www.ncbi.nlm.nih.gov/articlerender.fcgi?artid=PMC3198950>

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, 25(13), 1605–1612. <https://doi.org/10.1002/jcc.20084>

Riel, A. M. S., Decato, D. A., Sun, J., Massena, C. J., Jessop, M. J., & Berryman, O. B. (2018). The intramolecular hydrogen bonded-halogen bond: A new strategy for preorganization and enhanced binding. *Chemical Science*, 9(26), 5828–5836. <https://doi.org/10.1039/c8sc01973h>

Santos-Martins, D., Forli, S., Ramos, M. J., & Olson, A. J. (2014). AutoDock4Zn: An improved AutoDock force field for small-molecule docking to zinc metalloproteins. *Journal of Chemical Information and Modeling*, 54(8), 2371–2379. <https://doi.org/10.1021/ci500209e>

Schrodinger, L. (2010). *The PyMOL Molecular Graphics System* (1.8).

Siiskonen, A., & Priimaggi, A. (2017). Benchmarking DFT methods with small basis sets for calculation of halogen-bond strengths. *Journal of Molecular Modeling*, 23, 50.

Tosso, R. D., Vettorazzi, M., Andujar, S. A., Gutierrez, L. J., Garro, J. C., Suvire, F. D., Enriz, R. D., Angelina, E., Rodríguez, R., Nogueras, M., & Cobo, J. (2017). The electronic density obtained from a QTAIM analysis used as molecular descriptor. A study performed in a new series of DHFR inhibitors. *Journal of Molecular Structure*, 1134(2017), 464–474. <https://doi.org/10.1016/j.molstruc.2016.12.060>

Wilcken, R., Zimmermann, M. O., Lange, A., Joerger, A. C., & Boeckler, F. M. (2013). Principles and applications of halogen bonding in medicinal chemistry and chemical biology. *Journal of Medicinal Chemistry*, 56(4), 1363–1388. <https://doi.org/10.1021/jm3012068>

Word, J. M., Lovell, S. C., Richardson, J. S., & Richardson, D. C. (1999). Asparagine and glutamine: Using hydrogen atom contacts in the choice of side-chain amide orientation. *Journal of Molecular Biology*, 285(4), 1735–1747. <https://doi.org/10.1006/jmbi.1998.2401>

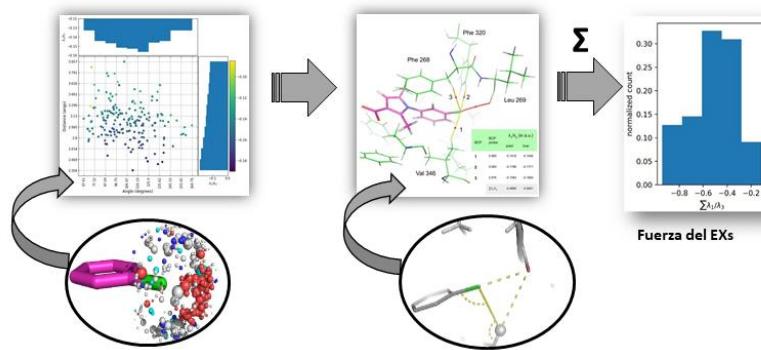
Zimmermann, M. O., Lange, A., & Boeckler, F. M. (2015). Evaluating the potential of halogen bonding in molecular design: Automated scaffold decoration using the new scoring function xbscore. *Journal of Chemical Information and Modeling*, 55(3), 687–699.

<https://doi.org/10.1021/ci5007118>

Zimmermann, M. O., Lange, A., Wilcken, R., Cieslik, M. B., Exner, T. E., Joerger, A. C., Koch, P., & Boeckler, F. M. (2014). Halogen-enriched fragment libraries as chemical probes for harnessing halogen bonding in fragment-based lead discovery. *Future Medicinal Chemistry*, 6(6), 617–639.
<https://doi.org/10.4155/fmc.14.20>

Capítulo 4

Predicción de la formación de Enlaces de Halógeno en entornos proteicos



CAPITULO 4

Los bolsillos de las proteínas que forman un enlace halógeno (EX) con un compuesto halogenado, forman simultáneamente otras interacciones con el átomo de halógeno que puede considerarse como su *entorno de enlace X* (EXent). Como el estudio de las propiedades de este EXent no ha sido profundizado previamente en sistemas biológicos, en este capítulo, se realizó un protocolo que evalúa la fuerza de EXent como una medida de la propensión de un bolsillo de la proteína para albergar un EX. Los descriptores topológicos basados en la densidad de carga en combinación con herramientas de aprendizaje automático se emplearon para predecir la formación y la fuerza de las interacciones que conforman el EXent, en función de sus parámetros geométricos. Sobre la base de estos resultados, se propone un protocolo (xenv) que permite predecir en base al EXent la probabilidad de que el bolsillo de una proteína aloje un enlace X.

4.1 Enlaces de halógeno en entornos proteicos

En las proteínas, la halogenación de aminoácidos (particularmente Tyrosina) se sabe que ocurre naturalmente como resultado del estrés oxidativo, catalizada por enzimas como mieloperoxidases (Domigan et al., 1995). La mieloperoxidasa cataliza la reacción de peróxido de hidrógeno y haluros (Cl-, Br-, I-) o iones de pseudohaluro (SCN-), generando ácidos hipohalosos (HOX), que reaccionan rápidamente con diferentes objetivos en las proteínas incluyendo los átomos de azufre de cisteína o metionina, átomos de nitrógeno de α -aminoácidos o histidina, lisina y cadenas laterales de arginina y los anillos aromáticos de tirosina y triptófano(Davies, 2011). Niveles elevados de residuos de tirosina halogenada han sido detectados en proteínas aisladas de pacientes con aterosclerosis, asma y fibrosis quística (Pattison & Davies, 2004; Wu et al., 2000). Mientras que la existencia de EX que involucran tales variantes de proteínas halogenadas aún no ha sido probada, el hecho de que pueden ser reconocidos por anticuerpos policlonales (Kambayashi et al., 2009) sugiere su existencia.

Estudios recientes sobre fragmentos amiloidogénicos de péptidos de calcitonina humana mostraron una amplificación de su capacidad de fibrilación tras la halogenación de residuos de

Phe, con una eficacia que sigue el orden de la fuerza establecida para EX, Cl < Br < I (Bertolani et al., 2015).

En sistemas biológicos, el aceptor de EX generalmente son macromoléculas, en un trabajo realizado por Xu y colaboradores (Xu et al., 2014) encontraron 778 enlaces que involucran halógenos en estructuras del PDB, y en 211 de ellos sistemas de electrones π deslocalizados participan de la interacción por ejemplo, en residuos Arg (Kortagere et al., 2008; Scholfield et al., 2013) y otros derivados aromáticos (Saraogi et al., 2003). De las restantes interacciones, 567 son contactos C-X…Y y 430 involucran residuos de proteínas, alrededor del 84% de estos involucran átomos de O como aceptores de EX, mientras que los átomos de N están involucrados sólo en el 10,4% de los casos.

Los únicos casos conocidos de EX en los que el aceptor no es parte de una biomacromolécula son cuando moléculas de agua actúan como aceptores de EX, a menudo una sola molécula de agua esté involucrada simultáneamente tanto en EX y en EH; se han realizado cálculos químicos cuánticos que demostraron que los EX con agua son termodinámicamente más estables que otras interacciones que involucran estas moléculas (Zhou, 2010) además en el relevamiento del PDB realizado por Xu y colaboradores se ha descubierto que estas interacciones ocupan alrededor del 16,6% de los EX en sistemas biológicos conocidos (Xu et al., 2014).

Los complejos del fármaco diclofenac con el complejo enzimático citocromo P450936 (Wester et al., 2003) y lactoferrina (Mir et al., 2009) son ejemplos de moléculas de agua que actúan como puentes entre enlaces de hidrógeno y halógeno. En el primer caso, el EX con una molécula de agua posee una distancia Cl…O de 2,85 Å y un ángulo C-Cl…O de 158,1° junto con un EH que involucra la misma molécula de agua y un oxígeno carboxilo de la cadena lateral de Glu297 (Figura 4.1). En el segundo caso se observó una estructura similar, aunque con una distancia Cl…O levemente mayor, con un ángulo algo menos lineal, e involucrando el residuo Thr663 en el EH.

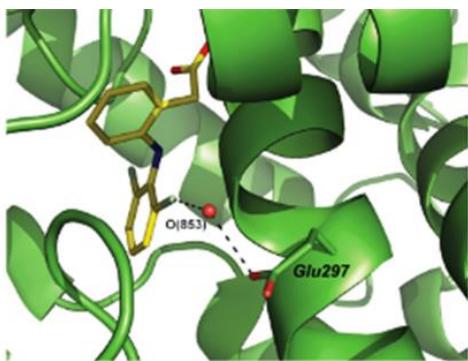


Figura 4.1: Complejo diclofenac-citocromo P450 que muestra que una molécula de agua forma al mismo tiempo un EX con un átomo de cloro y un EH con un oxígeno carboxilo del Glu297. Figura tomada de Parisini y col.(Parisini et al., 2011)

La geometría del enlace X en términos del ángulo de aproximación del halógeno hacia el átomo aceptor muestra que, en su mayor parte, el agujero σ es atraído por electrones no enlazantes del aceptor (Auffinger et al., 2004) con $\Theta_2 \approx 120^\circ$ lo que concuerda con las geometrías detalladas para moléculas pequeñas(Ouvrard et al., 2003). Las excepciones para los enlaces X en sistemas biológicos son cuando están disponibles electrones π deslocalizados, por ejemplo, de la cadena lateral de un aminoácido aromático o del enlace peptídico de un esqueleto proteico. Estos tipos de enlaces π -X se identificaron por primera vez a partir de un relevamiento de estructuras cristalinas de moléculas pequeñas en la base de datos de Cambridge (Prasanna & Guru Row, 2000) demostrando que el halógeno se encuentra perpendicular al anillo aromático. Los primeros enlaces X en biomoléculas fueron observados con el anillo aromático de los residuos de fenilalanina en los complejos de inhibidores de la proteína quinasas CK2 y CDK2 (A. P. Voth, 2007), este estudio se amplió a otros compuestos aromáticos y junto con cálculos MP2 se les pudo atribuir una energía de interacción de 2,5 kcal/mol(Matter et al., 2009).

Los enlaces peptídicos de la mayoría de los aminoácidos participan en los enlaces H que ayudan a definir estructuras secundarias y terciarias de una proteína plegada. En estos casos, los orbitales no enlazantes del oxígeno carbonílico de los enlaces peptídicos está ocupado por enlaces H; por lo tanto, el único potencial electronegativo disponible para el enlace X proviene de los orbitales π (Alkorta et al., 2008). Se demostró que este tipo de interacciones es predominante en complejos proteicos, y se encuentra orientada perpendicularmente a los enlaces H, siendo energéticamente independientes. Estas observaciones conducen a la hipótesis de que los enlaces X biológicos son interacciones ortogonales a los enlaces H cuando ambos comparten el oxígeno carbonílico de un enlace peptídico como el aceptor común, lo que sugiere que un enlace X se puede introducir como una interacción para el reconocimiento sin

alterar la estructura estabilizada por EH de la proteína diana. Este tipo de geometría se observa en cadenas alfa y beta de las proteínas, habiéndose demostrado esta ortogonalidad entre ambas interacciones en péptidos pequeños de forma experimental y teórica (A. R. Voth et al., 2009) pudiendo observar que químicamente no altera la estabilidad de los EHs. (Figura 4.2)

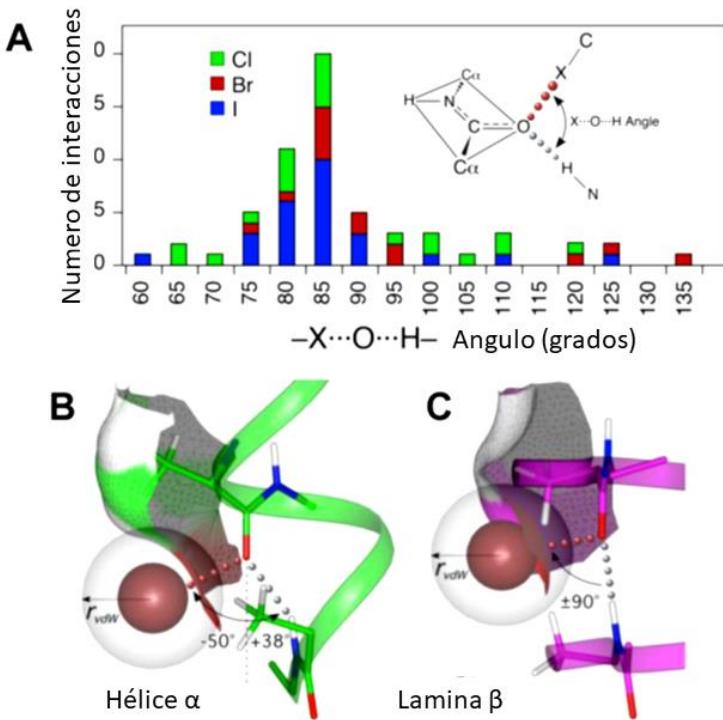


Figura 4.2: A) Distribución relativa de los ángulos de EHs y EXs con una base de Lewis común, subdividida de acuerdo a los halógenos implicados. B) Representación de la ortogonalidad EXs y EHs en láminas beta y C) en alfa hélices. Imagen adaptada de Voth, 2009 (A. R. Voth et al., 2009)

En los últimos años Zimmerman y col. (2015) (Zimmermann et al., 2015) derivaron una función de puntuación empírica derivada de la Mecánica Cuántica para evaluar la interacción entre ligandos halogenados aromáticos y el átomo de oxígeno carbonílico de la cadena principal de las proteínas. Al aplicar esa función, desarrollaron un algoritmo que evalúa el potencial de los ligandos halogenados unidos a proteínas para formar posibles EX. Si bien su función de puntuación del enlace se basa en cálculos precisos de MP2, se parametrizó en un sistema con un modelo reducido compuesto por halobenceno (yodobenceno, bromobenceno o clorobenceno) y N-metilacetamida (que representa la estructura de la proteína). Por lo que, no tendría en cuenta el entorno del EX en los complejos ligando-proteína.

Los EXs en los complejos de proteína-ligando no se encuentran aislados, sino que están rodeados por residuos de las proteínas que conforman el bolsillo donde se une el halógeno, constituyendo el *ambiente* del EX dentro de la estructura de la proteína. Las propiedades de estos entornos al enlace podrían afectar significativamente la formación del enlace. Dado que la formación del agujero σ es una consecuencia de la anisotropía de la distribución de la densidad eletrónica alrededor del halógeno, los residuos de las proteínas que rodean el átomo de halógeno pueden perturbar esa distribución de carga, lo que influye en el tamaño del agujero σ y, en consecuencia, en la propensión o no para la formación del EX. Además, los residuos podrían estar en un rango de distancia adecuado para interactuar directamente con el halógeno.

Además, Wilcken y col. (2013) (Wilcken et al., 2013) han estudiado el efecto de los sustituyentes del anillo de halobenceno en el tamaño del agujero σ como una forma de ajustar la fuerza de los EX en el contexto de los proyectos de optimización de compuestos líderes; pero nuevamente no ha sido considerado el efecto de los residuos de las proteínas que rodean el átomo de halógeno sobre la fuerza del EX.

Por otro lado, actualmente, Shinada et al (2019) (Shinada et al., 2019) han analizado no solo los enlaces halógenos, sino también los entornos interactivos completos de los átomos de halógeno en el contexto proteína-ligando. Aunque realizaron un estudio exhaustivo de las estructuras que contienen halógeno en el PDB (Protein Data Bank), su estudio se basó en solamente en criterios geométricos para evaluar las interacciones de halógenos con los residuos circundantes. La distancia y los rangos angulares están bien definidos para las interacciones convencionales que varían de fuerte a moderada, sin embargo, para enlaces de hidrógeno débiles que involucran grupos hidrófobos, los rangos geométricos no se encuentran tan bien definidos.

En la presente tesis con el fin profundizar el estudio de los entornos biológicos en EXs se desarrolló un protocolo para evaluar la viabilidad de los ligandos halogenados para formar un EX en los bolsillos de proteínas. Esta técnica, podría ser útil en el contexto de proyectos de optimización de compuestos líderes en los que un ligando se modifica estructuralmente para aumentar la afinidad de unión a su objetivo proteico.

Como los bolsillos de las proteínas que forman un enlace halógeno (EX) con un ligando halogenado constituyen simultáneamente otras interacciones con el átomo de halógeno que puede considerarse como su *entorno de enlace X* (EXent), a diferencia de protocolos anteriores

para la predicción del EX en los complejos de proteína-ligando, el realizado aquí (xenv), tiene en cuenta dicho entorno dentro de la proteína, es decir, interacciones distintas al EX en sí mismo, que el átomo de halógeno podría formar con los residuos de las proteínas, fundamentalmente enlaces de hidrógeno. Así, el protocolo obtenido (xenv) permite evaluar la fuerza de EXent como una medida de la propensión de un bolsillo de proteína para albergar un enlace X. Los descriptores topológicos basados en la densidad electrónica en combinación con herramientas de aprendizaje automático se emplearon para predecir la formación y la fuerza de las interacciones que conforman el EXent, en función de sus parámetros geométricos. Sobre la base de estos resultados, propusimos que el EXent se puede usar para evaluar la posibilidad de que un bolsillo de que una proteína forme un EX.

4.2 Procedimiento

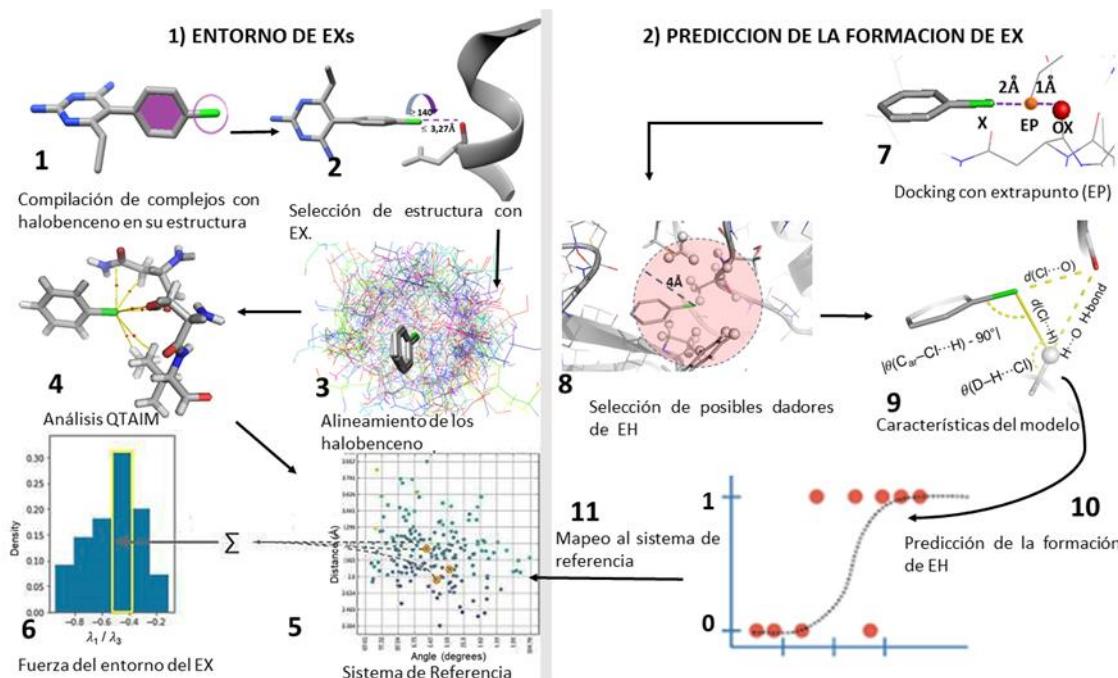


Figura 4.3: Esquema de trabajo de los pasos del protocolo para predecir la capacidad de formación de un EX en base las propiedades del entorno.

4.2.1 Entornos EXs: se utilizo para elaborar un sistema de referencia los datos obtenidos del analisis de densidad electronica y elementos topológicos de los 86 complejos utilizados en el relevamiento en el punto 3.2.2 .

4.2.2 Prediccion de la formacion de EXs

4.2.2.1 Docking Molecular

Los cálculos de docking se realizaron con AutoDock4 (Santos-Martins et al., 2014). Previamente, las estructuras PDB se prepararon de acuerdo con el siguiente procedimiento. Se empleó el software Reduce para agregar y optimizar las posiciones de los átomos de hidrógeno. Luego, el ligando (si está presente) se eliminó de la estructura PDB y la proteína (es decir, el archivo receptor.pdb) se preparó para los cálculos de acoplamiento con los scripts de AutoDock Tools prepare_receptor4.py con la opción -A 'checkhydrogens' para preservar los estados de protonación predicho por Reduce. El archivo receptor.pdbqt resultante tiene cargas atómicas añadidas e hidrógenos no polares fusionados en sus átomos de carbono enlazados.

El docking molecular se realizó con el campo de fuerzas modificado para descubrir la asimetría del EX, como se explicó en la sección 3.2.1, utilizando la aproximación del extra punto (EP) que imita el agujero σ . El aceptor del enlace X designado se renombra de O a OX para ajustar la direccionalidad del enlace hacia ese atomo, e interactúan entre sí a través de un potencial de van der Waals 12-6 de Lennard Jones con una distancia de equilibrio de $\sim 1,00 \text{ \AA}$, de modo que la distancia de equilibrio del enlace X $d(X\cdots OX)$ es de aproximadamente $3,00 \text{ \AA}$ ($2,00 \text{ \AA}$ para $d(X\cdots EP) + 1,00 \text{ \AA}$ para $d(EP\cdots X)$), como se indica en la figura 4.3 punto 7.

La profundidad del pozo de potencial se fijó en $\varepsilon=3,00 \text{ kcal/mol}$, lo que garantiza la formación de enlaces X sin ser tan fuerte como para anular las interacciones del halobenceno con otros átomos del entorno proteico. Se generaron cincuenta poses diferentes con el algoritmo de acoplamiento. Es importante señalar que este algoritmo solo está destinado a generar diferentes poses de ligando con EX. La capacidad del bolsillo de la proteína para albergar un enlace X se evaluará más adelante de acuerdo con el protocolo de predicción propuesto.

4.2.2.2 Detalles del modelo de machine learning

Los datos de densidad de carga electronica calculados en el sistema de referencia se emplearon para construir un modelo de regresión logística simple. Se procesaron los complejos del sistema de referencia como si fueran poses de docking. Todos los átomos de Hidrogeno de la proteina en un rango de distancia de 4,00 Å del halógeno fueron seleccionados y aquellos conectados y no conectados a través de un BCP al cloro fueron marcados como 1 y 0, respectivamente. Mediante este procedimiento se seleccionó un número total de 501 átomos de hidrógeno, de los cuales 189 estaban realmente conectados a un BCP (Figura 4.3 punto 8 y 9).

Con respecto a las propiedades utilizadas de entrada, además de las coordenadas radiales ($d(Cl\cdots H)$) y angulares ($\theta(Car-Cl\cdots H)$) alrededor del halógeno, se incorporaron otros parametros adicionales al modelo de aprendizaje automático. En particular, se consideraron las características del ángulo donante del enlace H ($(D-H\cdots Cl)$, con D = C, N, O o S), así como el enlace H con el aceptor de oxígeno del enlace X. Esta última característica se codificó en una variable que toma valor 1 o 0 dependiendo de si el H está unido por EH al O o no. Los criterios para la formación del EH $H\cdots O$ fueron la distancia $d(H\cdots O) \leq 2,40$ Å y el ángulo $\theta(D-H\cdots O) \geq 100^\circ$. Se incluyó la característica del EH $H\cdots O$ porque se sospechaba un patrón anti-cooperativo entre la formación de $Cl\cdots H$ BCP y el EH con el aceptor de oxígeno del EX. Además, para investigar una posible relación entre la fuerza del enlace X y la formación del enlace H $Cl\cdots H$, la distancia del enlace X $d(Cl\cdots O)$ también se incluyó como una característica del modelo (Figura 4.3, punto número 9).

4.3 Resultados

4.3.1 Entornos de EXs

Para profundizar el conocimiento de los Enlaces de Halógeno se realizó un relevamiento del entorno de Enlaces de Halógeno en los sitios de unión de proteínas en complejos de ligando-proteína obtenidos de la base de datos (Hu et al., 2005); mediante la inspección no solo de los enlaces X sino también de todos los otros tipos de interacciones no covalentes que el halógeno podría formar en el contexto proteico (Wilcken et al., 2013) y posteriormente se aplicó este análisis a las poses obtenidas del Docking Molecular parametrizado para EXs.

Para realizar el protocolo para la predicción de la formación de EXs en este capítulo se utilizaron los datos de densidad electrónica calculados en el capítulo 3, en complejos ligando-proteína unidos a través de EXs (sistema de referencia) para obtener una estimación de las fortalezas de EXent. Esta estimación luego la usaremos para juzgar la predicción de los bolsillos de proteínas para albergar un enlace X.

Realizando un análisis más detallado de los enlaces X del cloro revela que la distribución de los valores de $\sum \rho_b$ para los enlaces H y los enlaces X están bastante separados, y los enlaces H tienen un valor medio de $\sum \rho_b$ más alto que los enlaces X (Figura 4.4).

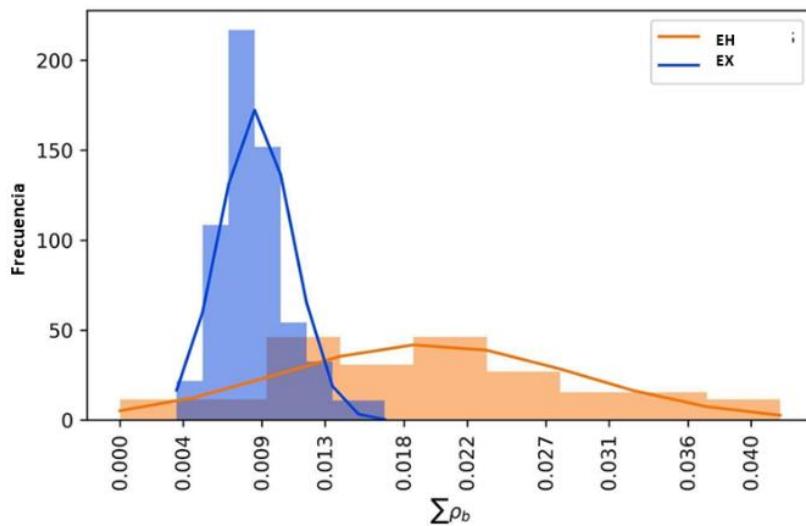


Figura 4.4: Distribución general de los valores de $\sum \rho_b$ de los enlaces X/H (en unidades atómicas) en los complejos con enlaces X. Las distribuciones están bastante separadas, con medias \pm desviaciones estándar de $0,019 \pm 0,010$ au. y $0,009 \pm 0,002$ au, para enlaces H y enlaces X, respectivamente.

Como se explicó en el capítulo 3 como sustituto de $\sum \rho_b$ para explicar mejor las preferencias angulares de los donantes de enlaces H alrededor del átomo de halógeno de ahora en adelante emplearemos $\sum \lambda_1 / \lambda_3$. Cuanto más negativo sea el valor de $\sum \lambda_1 / \lambda_3$, más fuerte será el EXent. De acuerdo con la distribución de los complejos de cloro con enlace X en función del valor $\sum \lambda_1 / \lambda_3$. (Figura 4.5), aproximadamente el 60 % de los complejos estudiados demostraron tener EXent de fuerza intermedia, con valores de $\sum \lambda_1 / \lambda_3$ que oscilan entre -0,28 y -0,60 au. El 30 % y el 10 % restantes de los complejos tienen EXent más fuertes ($\sum \lambda_1 / \lambda_3 < -0,60$ au) y más débiles ($\sum \lambda_1 / \lambda_3 > -0,28$ au), respectivamente.

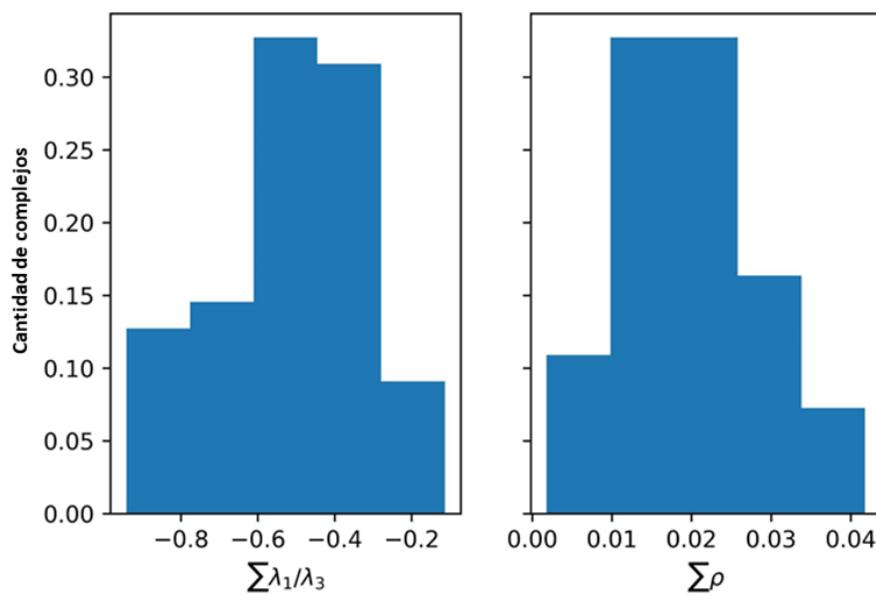


Figura 4.5: Distribución de los complejos con enlaces X en función de la suma de los valores ρ_b y λ_1 / λ_3 asociados a los enlaces H Cl \cdots H (en unidades atómicas).

Un EXent más fuerte podría reflejar una mejor distribución de donantes de enlaces H alrededor del halógeno (según las preferencias de distribución angular y/o radial), pero también puede deberse a un mayor número de donantes de enlaces H. La Figura 4.6 muestra que el número de donantes de enlaces H en los complejos con enlaces X varía de 1 a 6, siendo 3 el número más común de donantes ($\sim 35\%$ de los complejos).

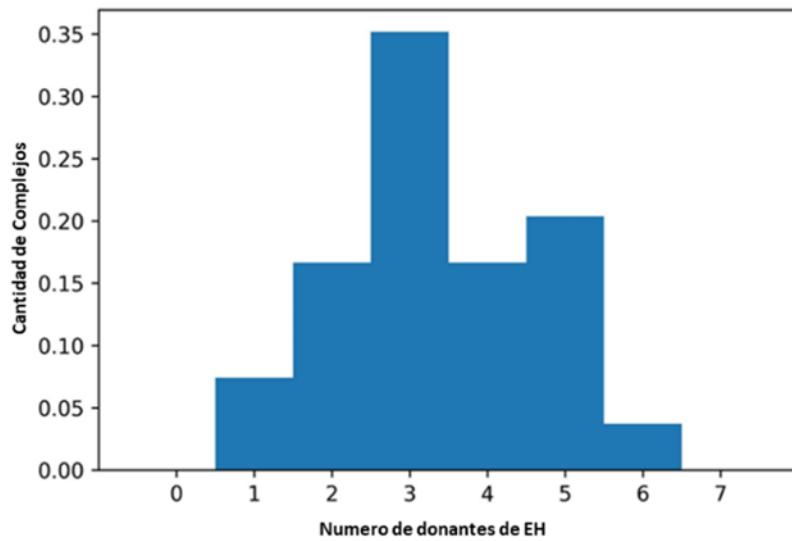


Figura 4.6: Distribución de complejos con enlaces X en función del número de donantes de enlaces H al átomo de cloro.

4.3.2 Predicción de la formación de EXs

La capacidad del bolsillo de la proteína para albergar un EX se evaluó de acuerdo con el protocolo de predicción de la figura 4.3, con este fin de las 50 poses obtenidas del docking molecular con EX parametrizado, como se explicó en la sección 4.2.2.1 para cada complejo ligando-proteína se calculó la distancia desde el átomo de halógeno hasta los átomos de H y O del receptor en un rango de distancia de 4 Å, todos los átomos de H de la proteína que se encuentran a esta distancia del átomo de halógeno ($d(Cl\cdots H) \leq 4,00 \text{ \AA}$) se consideran como posibles interaccionantes.

4.3.2.1 Modelo de Machine learning

En la primera etapa, se entrenó un clasificador de machine learning para predecir la formación de puntos críticos de enlace (BCP) y en la segunda etapa, se realizó el mapeo al sistema de referencia solo para aquellos átomos de H que se predijo la formación de un BCP en la etapa 1. La Tabla 4. 1 muestra la magnitud y el signo de los coeficientes del modelo de regresión logística, como era de esperar, la propiedad geométrica más importante que gobierna la

formación de $\text{Cl}\cdots\text{H}$ BCP (es decir, el coeficiente de mayor magnitud) es la distancia $d(\text{Cl}\cdots\text{H})$ entre ambos átomos que interactúan, como lo indica el valor negativo del coeficiente correspondiente, cuanto más corta es la distancia $d(\text{Cl}\cdots\text{H})$, mayor es la probabilidad de formación de BCP.

Tabla4.1: Coeficientes del modelo de clasificación de machine learning

Propiedad	Coeficiente
$d(\text{Cl}\cdots\text{H})$	-1.77036965
$\text{ABS}(\theta(\text{C}_{\text{ar}}-\text{Cl}\cdots\text{H}) - 90^\circ)$	-0.60877101
$\theta(\text{D}-\text{H}\cdots\text{Cl})$	1.27193649
$\text{H}\cdots\text{O}$ H-bond	-0.39989555
$d(\text{Cl}\cdots\text{O})$	0.09545043

Curiosamente, el ángulo donante del enlace H ($\text{D}-\text{H}\cdots\text{Cl}$), con $\text{D} = \text{C, N, O o S}$, resultó ser más determinante para la formación de BCP que el ángulo del enlace H alrededor del halógeno ($\text{C}_{\text{ar}}-\text{Cl}\cdots\text{H}$), a juzgar por las magnitudes de los respectivos coeficientes. Es decir, la desviación de 180° en el ángulo ($\text{D}-\text{H}\cdots\text{Cl}$) es menos tolerada que la desviación de $\sim 90^\circ$ en el ángulo $\theta(\text{C}_{\text{ar}}-\text{Cl}\cdots\text{H})$. Los signos opuestos de los coeficientes de ambos ángulos tienen sentido ya que la probabilidad de formación de BCP aumenta a medida que aumenta ($\text{D}-\text{H}\cdots\text{Cl}$) (correlación positiva) y $|(C_{\text{ar}}-\text{Cl}\cdots\text{H}) - 90^\circ|$ disminuye (correlación negativa). Además, dado que sospechábamos una competencia entre las formaciones de BCP de $\text{Cl}\cdots\text{H}$ y los enlaces de H con el átomo de O acceptor del enlace X , también se incluyó una característica binaria que explica la formación de enlaces de H $\text{H}\cdots\text{O}$. Como lo demuestra el signo negativo del coeficiente, el EH $\text{H}\cdots\text{O}$ se opone a la formación de $\text{Cl}\cdots\text{H}$ BCP, lo que confirma la anticooperatividad entre ambas interacciones.

La Figura 4.7 muestra el grafo molecular de densidad electrónica para uno de los complejos con enlace X estudiados, una estructura de anillo está subtendida por los tres enlaces H $\text{Cl}\cdots\text{H}$, $\text{H}\cdots\text{O}$

y O…Cl. En la topología de la densidad de carga de electrones, la estructura de anillo se evidencia por la existencia de un Punto Crítico de Anillo (RCP, esfera amarilla). La proximidad del RCP al Cl…H BCP indica que la estructura del anillo está un poco cerca del colapso debido a la debilidad de la interacción Cl…H. Esta última interacción es debilitada por la interacción competitiva H…O.

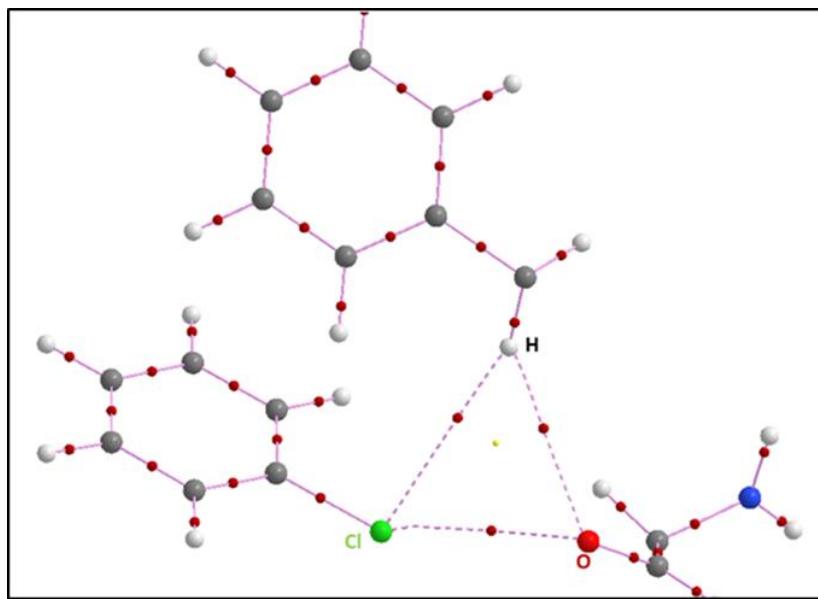


Figura 4.7: Grafo molecular de densidad electrónica (PDB: 5MZI, complejo formado por la enzima quinurenina 3-monooxigenasa de la Pseudomonas fluorescens con un inhibidor clorado), en el que una estructura de anillo está subtendida por las tres interacciones, Cl…H, H…O y O…Cl. En la topología de la densidad de carga, la estructura del anillo se evidencia por la existencia de un Punto Crítico del Anillo (RCP, esfera amarilla).

Finalmente, de acuerdo con hallazgos previos, no parece haber ninguna relación entre la fuerza del enlace X y la formación individual de Cl…H BCP, a juzgar por la pequeña magnitud del coeficiente asociado a $d(Cl\cdots O)$. De hecho, cuando se elimina esta función y se vuelve a entrenar el modelo, el rendimiento del clasificador aumenta un poco. Sin embargo, hemos demostrado anteriormente que cuando el EXent se considera como un todo (es decir, todos los BCP de Cl…H se eliminan a la vez) en realidad hay una mejora del enlace X.

Para la segunda etapa el mapeo se realiza de la siguiente manera: a) las coordenadas de ángulo y distancia en el sistema de referencia se dividen en 10 bins cada una; b) los átomos de H que se prevé que interactúen con el cloro se mapean a través de sus coordenadas polares, c) la propiedad derivada de la densidad electrónica de todos los BCP se promedian juntos; d) el valor medio de la propiedad se le asigna al átomo de H. Después de asignar todos los posibles

donantes de enlaces H a los BCP del sistema de referencia, se suman los valores de densidad electrónica para obtener una estimación global de la fuerza del entorno del enlace X, $\sum\rho_{\text{poses}}$.

La fuerza estimada del entorno del EX $\sum\rho_{\text{poses}}$ se puede contrastar con las de los complejos con EX conocidos en el sistema de referencia $\sum\rho_{\text{ref}}$ para juzgar la capacidad del bolsillo de proteína para albergar un enlace X (Figura 4.5)

4.3.2.2 Validación

Para evaluar el rendimiento del protocolo, realizamos un redocking de los complejos con EX del sistema de referencia y los entornos de EX predichos en las poses de docking se compararon con los valores reales obtenidos de λ_1 / λ_3 . (Figura 4.8). Los puntos naranjas sobre la línea corresponden a los valores reales de la propiedad derivada de la densidad electrónica, mientras que los puntos azules son las predicciones para las poses obtenidas al realizar un redocking de los complejos. La mayoría de los entornos de EX se predijeron correctamente y los que no lo fueron, en la mayoría de los casos tienen errores relativos dentro del 35 % del valor real.

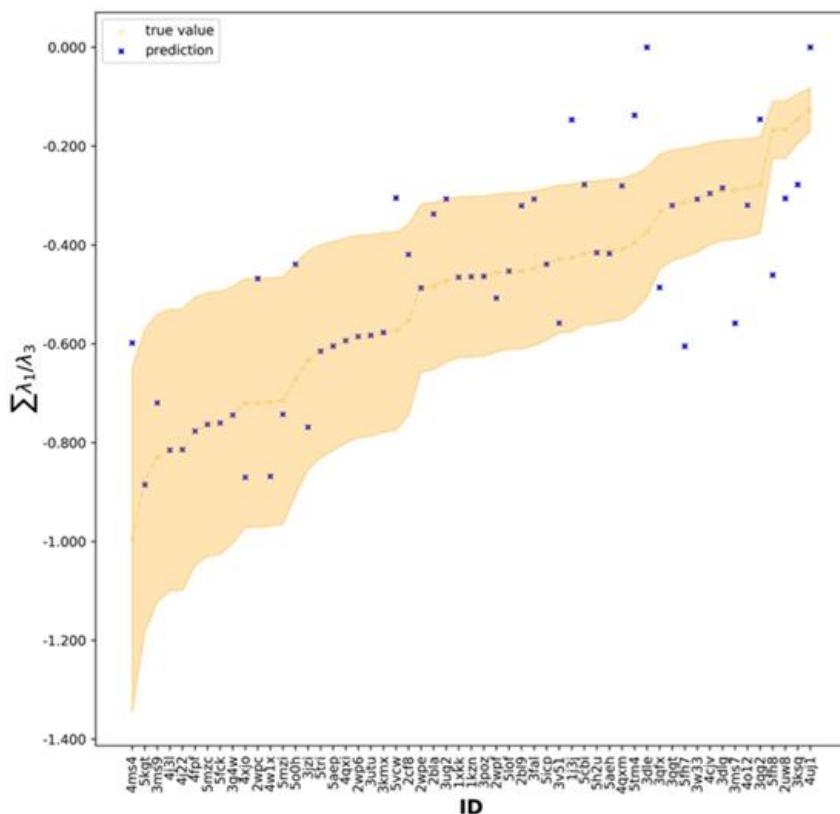


Figura 4.8: Rendimiento del protocolo de predicción del entorno EX. Los valores reales y predichos de la propiedad derivada de la densidad de carga, $\sum \lambda_1 / \lambda_3$, se representaron con puntos naranjas y azules, respectivamente. El área sombreada en naranja abarca errores de predicción relativos de hasta el 35 % de los valores verdaderos correspondientes.

Con el fin de tener una validación externa el rendimiento del protocolo se evaluó más a fondo en 6 complejos unidos a compuestos halogenados que no estaban incluidos en el conjunto de entrenamiento. Las propiedades derivadas de la densidad de carga de los enlaces H H \cdots Cl que conforman los EXent se calcularon con la metodología QTAIM y se predijeron simultáneamente con el protocolo ajustado. Las figuras 4.9 y 4.10 muestran los grafos moleculares de densidad de carga de cada complejo ligando-proteína con EX junto con la probabilidad estimada de formación del enlace de hidrógeno (EH(BCP)), así como los valores verdadero y predicho de la relación λ_1 / λ_3 para cada H \cdots Cl (BCP).

El protocolo predijo correctamente la formación de las interacciones de los enlaces de Hidrógeno H \cdots Cl que están realmente presentes en los grafos moleculares calculados y con valores estimados de la relación λ_1 / λ_3 que están muy cerca de sus valores reales. Además, según el histograma de la Figura 4.5, en general los complejos tienen un entorno de enlace X con una fuerza moderada.

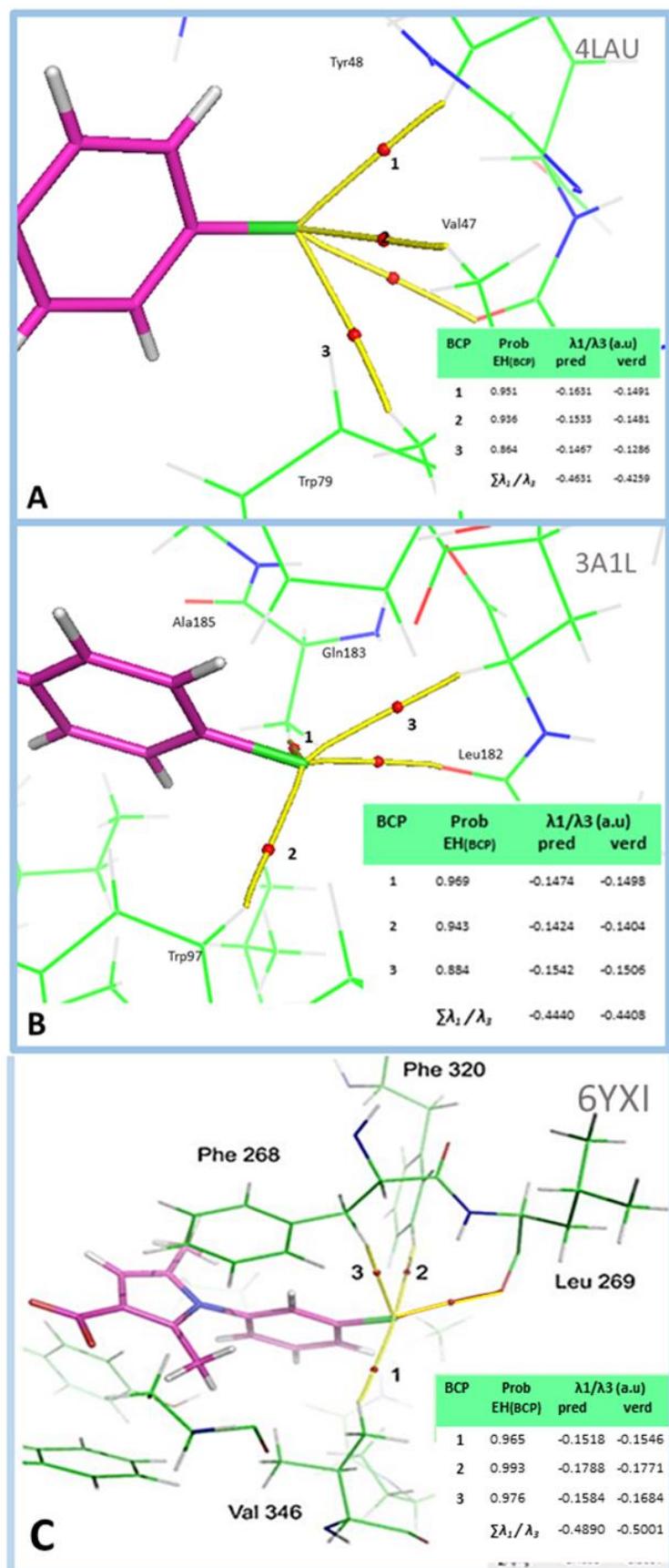


Figura 4.9: Grafos de densidad electrónica de los complejos con EX, código de PDB 4LAU (A), 3A1L (B) y código PDB 6YXI (C). Los caminos de enlace (BP) que conectan el cloro con el átomo de O (enlace X) y los átomos de H (enlaces H) se representan en amarillo. Los círculos rojos en la parte superior de los BP son los puntos críticos de enlace (BCP). Los BCP para los enlaces de hidrógeno H···Cl están numerados y sus propiedades correspondientes se enumeran en la tabla insertada (probabilidad estimada de formación del EH(BCP) y valor predicho/verdadero de la relación λ_1 / λ_3).

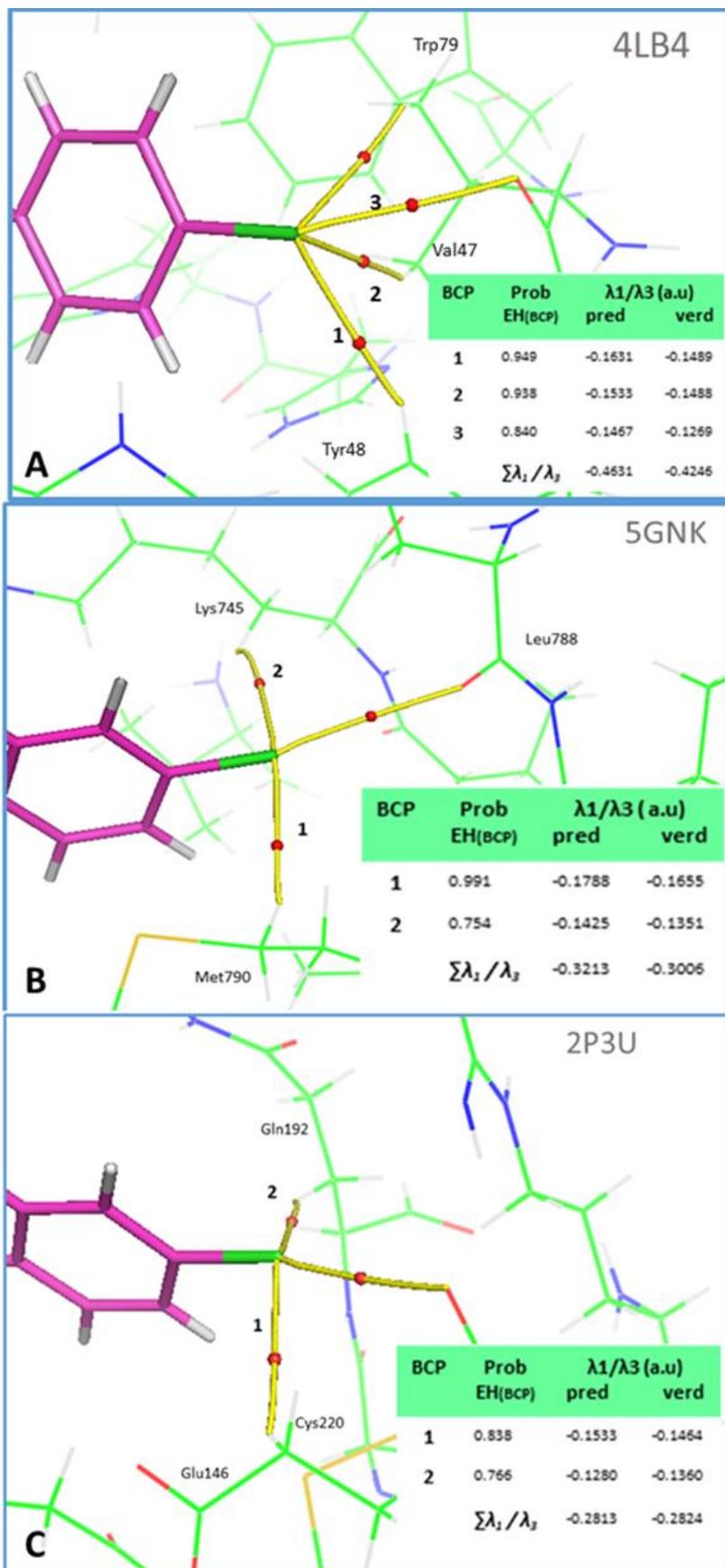


Figura 4.10: Grafos de densidad electrónica de los complejos con EX, código PDB 4LB4(A), 5GNK (B) y 2P3U (C). Los caminos de enlace (BP) que conectan el cloro con el átomo de O (enlace X) y los átomos de H (enlaces H) se representan en amarillo. Los círculos rojos en la parte superior de los BP son los puntos críticos de enlace (BCP). Los BCP para los enlaces de hidrógeno H···Cl están numerados y sus propiedades correspondientes se enumeran en la tabla insertada (probabilidad estimada de formación del EH(BCP) y valor predicho/verdadero de la relación λ_1 / λ_3)

4.3.2.3 Entorno de EX reales versus EXent señuelos

Habiendo validado el protocolo, podría, en principio, aplicarse de forma segura en un entorno prospectivo. Sin embargo, en la realización del protocolo de predicción del EX basado en el EXent, hemos asumido implícitamente que los subbolsillos de proteínas capaces de albergar un enlace X tienen un EXent más fuerte que los que no lo son. Necesitamos verificar que esta suposición realmente se cumple en la práctica.

Dado que el EXent hace una gran contribución a la fuerza de anclaje del halógeno en los complejos con enlaces X, los subbolsillos de proteínas capaces de albergar un EX deberían tener un EXent más fuerte. Para probar esta hipótesis, aplicando la parametrización de EX detallada en el punto 4.2.1, generamos cincuenta poses de docking en las que la estructura de clorobenceno se ve obligada a formar: (a) el EX conocido (real) y (b) un enlace X diferente en un subbolsillo de proteína cercano sin evidencia estructural sobre su existencia. Llamamos a este último enlace X artificial, como un *EXseñuelo*.

Los EXent basados en la densidad electrónica para las poses de docking se estimaron mapeando el entorno de la proteína al cloro en el sistema de referencia, siguiendo el protocolo explicado en la Figura 4.3. Se encontraron EX señuelos para solo 19 del total de estructuras que conforman el sistema de referencia. La Figura 4.11 muestra las predicciones de EXent para las poses de acoplamiento que forman el enlace X conocido (barras rojas) o un enlace X señuelo (barras amarillas), en esos 19 complejos. Como puede verse en la figura 4.11, solo en 5 de los 19 complejos, el EXent es más fuerte para el enlace X señuelo que para el enlace X conocido, lo que respalda parcialmente nuestra hipótesis.

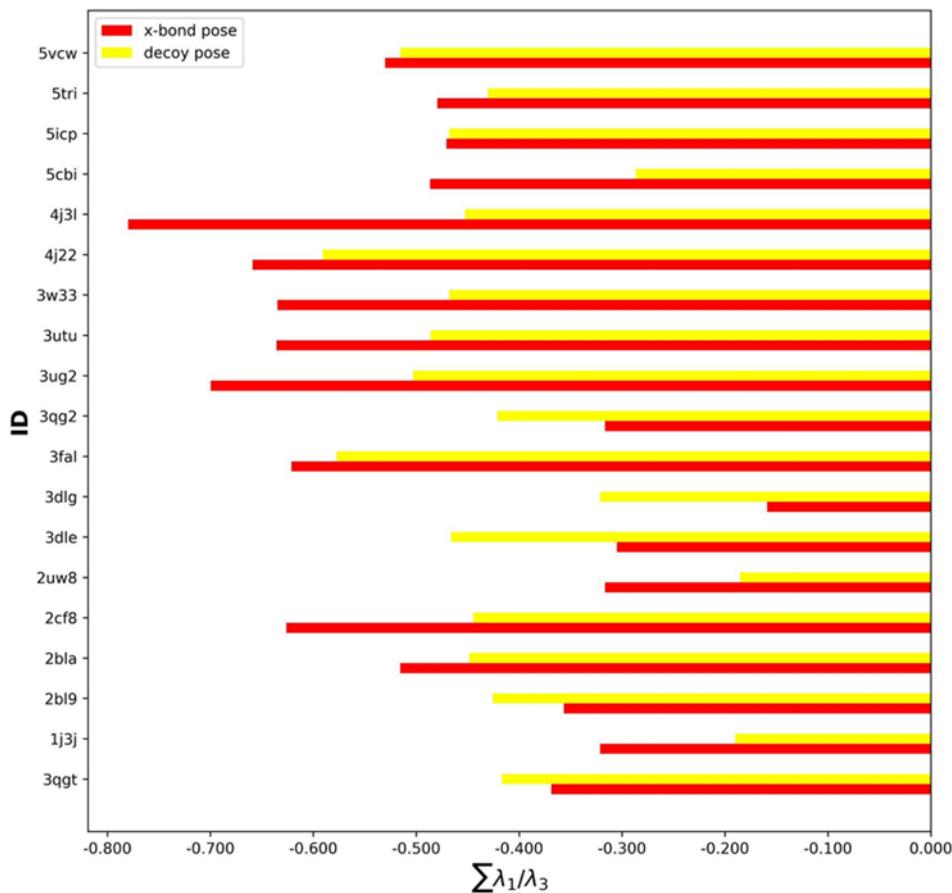


Figura 4.11: Fortaleza del EXent para las predicciones de las poses de docking que forman EX real (rojo) y el señuelo (amarillo).

La figura 4.12 muestra las poses de los enlaces X conocidos y señuelos de las estructuras seleccionadas, a modo de comparación.

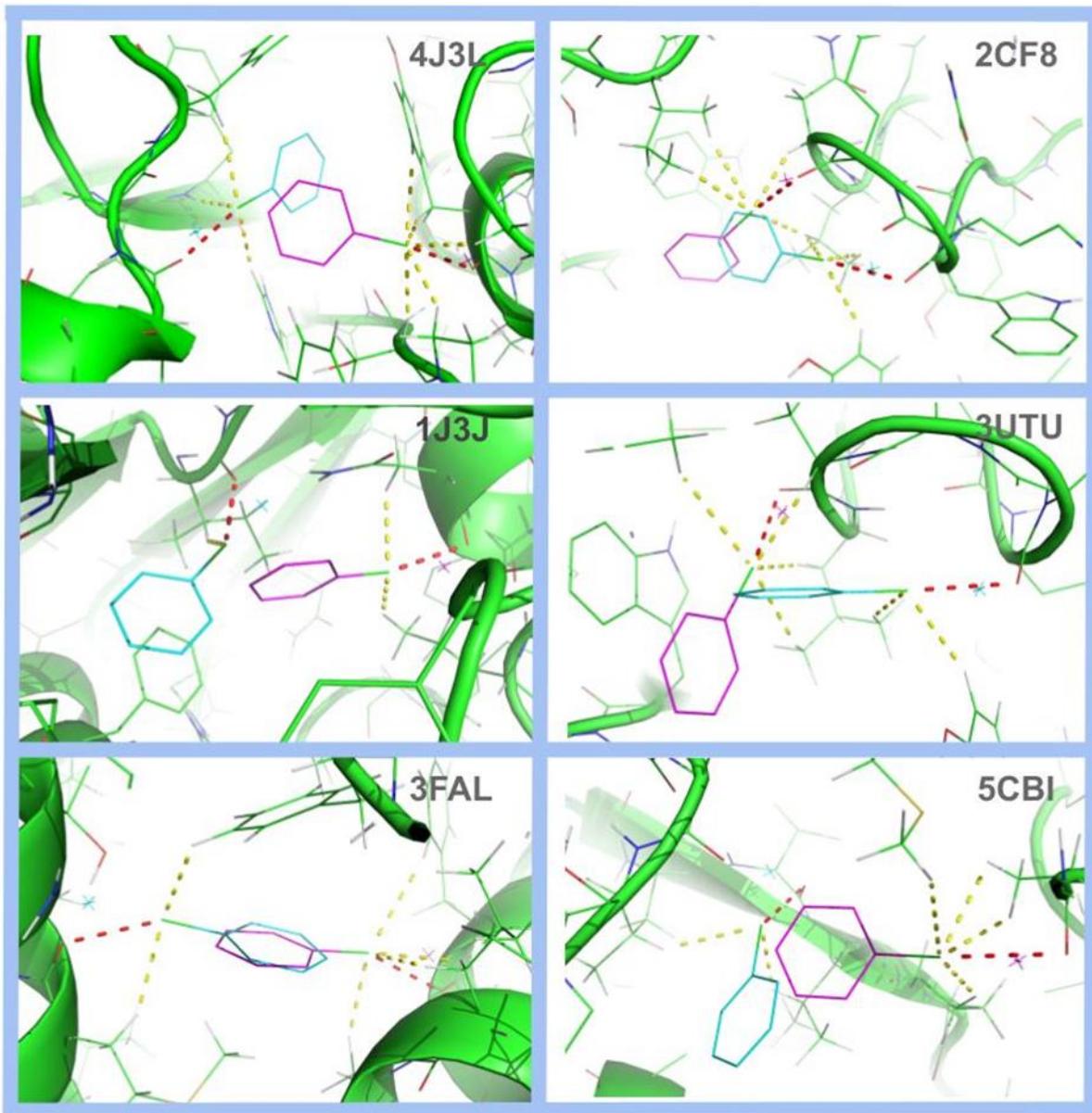


Figura 4.12: Poses de docking de halobenceno que forman un enlace X conocido (magenta) y señuelo (celeste) en las estructuras proteicas seleccionadas. Las interacciones de enlace H y enlace X se representaron con líneas discontinuas amarillas y rojas, respectivamente.

Los subbolsillos de las proteínas que albergan enlaces X reales forman una red más compleja de enlaces H con cloro que los con enlaces X señuelo (Figura 4.11). Una inspección más cuidadosa revela que las poses señuelo tienden a ser impulsadas por interacciones del anillo aromático, como por ejemplo interacciones de “stacking” $\text{C-H}\cdots\pi$, en lugar de interacciones con el halógeno. Además, en varios casos, la pose señuelo ha sido clasificada mejor por la función

de puntuación de docking de la pose experimental real que forma el verdadero enlace X (datos no mostrados). Este hallazgo indica que la función de puntaje de docking está subestimando las interacciones del entorno del enlace X.

4.4 Conclusiones:

Reconociendo que los átomos de halógeno pueden formar otras interacciones no covalentes además del enlace X en el bolsillo de una proteína, se estudió el papel del entorno hidrófobo del halógeno en los complejos ligando-receptor con EX. Es decir, se centró el estudio en lo que se denominó el "entorno de EX".

Se utilizaron los datos de densidad de carga electrónica calculados previamente en el capítulo 3 en los complejos ligando-proteína unidos a través de EX conocidos, para construir un modelo que puede predecir la formación y la fuerza de las interacciones del entorno del enlace X, basándose únicamente en sus datos geométricos. De esta forma, el modelo se puede aplicar directamente a las posiciones de acoplamiento para predecir la propensión de entornos de enlaces de halógeno sin la necesidad de calcular la densidad electrónica del sistema.

Para probar la hipótesis de que el enlace X en un bolsillo de proteína necesita un entorno adecuado para establecerse, se aplicó el protocolo de predicción (xenv) a enlaces X "reales" y "señuelos", así se confirmó que los primeros tienen un EXent más fuerte que los últimos. Además, se realizó una validación externa aplicando el procedimiento en estructuras de complejos con ligandos halogenados que no se encontraban en el conjunto de entrenamiento, y realizando simultáneamente los respectivos cálculos cuánticos, los datos obtenidos se compararon obteniendo resultados favorables.

Dado que la formación de enlaces X en los bolsillos de proteínas ciertamente requiere un entorno adecuado, podemos aplicar con confianza el protocolo de predicción del entorno de enlaces X para evaluar la propensión de los bolsillos de proteínas a albergar un enlace X.

Esta metodología podría ser de interés en el contexto de proyectos de optimización de compuestos líderes donde un ligando se modifica estructuralmente para aumentar la afinidad de unión por su proteína diana, lo que facilita la aplicabilidad de enlaces halógenos en el diseño molecular.

Finalmente, se destaca como un aporte importante de este trabajo que el código y los datos del sistema de referencia que se utilizaron en la elaboración del protocolo, para la predicción de la presencia de enlaces de halógeno en base al entorno (xenv) descripto en este capítulo, fueron publicados y se encuentra disponibles en el repositorio de Github (<https://github.com/emilioangelina/xbenv>)

4.5 Referencias:

- Alkorta, I., Blanco, F., Solimannejad, M., & Elguero, J. (2008). Competition of hydrogen bonds and halogen bonds in complexes of hypohalous acids with nitrogenated bases. *Journal of Physical Chemistry A*, 112(43), 10856–10863. <https://doi.org/10.1021/jp806101t>
- Auffinger, P., Hays, F. A., Westhoft, E., & Ho, P. S. (2004). Halogen bonds in biological molecules. *PNAS*, 101, 16789–16794. <https://doi.org/10.1073/pnas.0407607101>
- Bertolani, A., Pirrie, L., Stefan, L., Houbenov, N., Haataja, J. S., Catalano, L., Terraneo, G., Giancane, G., Valli, L., Milani, R., Ikkala, O., Resnati, G., & Metrangolo, P. (2015). Supramolecular amplification of amyloid self-assembly by iodination. *Nature Communications*, 6(May), 1–9. <https://doi.org/10.1038/ncomms8574>
- Davies, M. J. (2011). Myeloperoxidase-derived oxidation: mechanisms of biological damage and its prevention Overview of the Action of Myeloperoxidase and Other Heme Peroxidases. *J. Clin. Biochem. Nutr*, 48(1), 8–19. <https://doi.org/10.3164/jcbn.11>
- Domigan, N. M., Charlton, T. S., Duncan, M. W., Winterbourn, C. C., & Kettle, A. J. (1995). Chlorination of tyrosyl residues in peptides by myeloperoxidase and human neutrophils. *Journal of Biological Chemistry*, 270(28), 16542–16548. <https://doi.org/10.1074/jbc.270.28.16542>
- Hu, L., Benson, M. L., Smith, R. D., Lerner, M. G., & Carlson, H. A. (2005). Binding MOAD (Mother of All Databases). *Proteins: Structure, Function and Genetics*, 60(3), 333–340. <https://doi.org/10.1002/prot.20512>
- Kambayashi, Y., Ogino, K., Takemoto, K., Imagama, T., Takigawa, T., Kimura, S., Hibino, Y., Hitomi, Y., & Nakamura, H. (2009). Preparation and characterization of a polyclonal antibody against brominated protein. *Journal of Clinical Biochemistry and Nutrition*, 44(1), 95–103. <https://doi.org/10.3164/JCBN.08-196>
- Kortagere, S., Ekins, S., & Welsh, W. J. (2008). Halogenated ligands and their interactions with amino acids: Implications for structure-activity and structure-toxicity relationships. *Journal of Molecular Graphics and Modelling*, 27(2), 170–177. <https://doi.org/10.1016/j.jmgm.2008.04.001>
- Matter, H., Nazar, M., Güssregen, S., Will, D. W., Schreuder, H., Bauer, A., Urmann, M., Ritter, K.,

Wagner, M., & Wehner, V. (2009). Evidence for C-Cl/C-Br...π interactions as an important contribution to protein-ligand binding affinity. *Angewandte Chemie - International Edition*, 48(16), 2911–2916. <https://doi.org/10.1002/anie.200806219>

Mir, R., Singh, N., Vikram, G., Kumar, R. P., Sinha, M., Bhushan, A., Kaur, P., Srinivasan, A., Sharma, S., & Singh, T. P. (2009). The structural basis for the prevention of nonsteroidal antiinflammatory drug-induced gastrointestinal tract damage by the C-lobe of bovine colostrum lactoferrin. *Biophysical Journal*, 97(12), 3178–3186. <https://doi.org/10.1016/j.bpj.2009.09.030>

Ouvrard, C., Le Questel, J. Y., Berthelot, M., & Laurence, C. (2003). Halogen-bond geometry: a crystallographic database investigation of dihalogen complexes. *Urn:ISSN:0108-7681*, 59(4), 512–526. <https://doi.org/10.1107/S0108768103011510>

Parisini, E., Metrangolo, P., Pilati, T., Resnati, G., & Terraneo, G. (2011). Halogen bonding in halocarbon–protein complexes: A structural survey. *Chemical Society Reviews*, 40(5), 2267–2278. <https://doi.org/10.1039/c0cs00177e>

Pattison, D. I., & Davies, M. J. (2004). Kinetic Analysis of the Reactions of Hypobromous Acid with Protein Components: Implications for Cellular Damage and Use of 3-Bromotyrosine as a Marker of Oxidative Stress. *Biochemistry*, 43(16), 4799–4809. <https://doi.org/10.1021/bi035946a>

Prasanna, M. D., & Guru Row, T. N. (2000). C-halogen...π interactions and their influence on molecular conformation and crystal packing: a database study. *Crystal Engineering*, 3(2), 135–154. [https://doi.org/10.1016/S1463-0184\(00\)00035-6](https://doi.org/10.1016/S1463-0184(00)00035-6)

Santos-Martins, D., Forli, S., Ramos, M. J., & Olson, A. J. (2014). AutoDock4Zn: An improved AutoDock force field for small-molecule docking to zinc metalloproteins. *Journal of Chemical Information and Modeling*, 54(8), 2371–2379. <https://doi.org/10.1021/ci500209e>

Saraogi, I., Vijay, V. G., Das, S., Sekar, K., & Guru Row, T. N. (2003). C-halogen...π interactions in proteins: A database study. *Crystal Engineering*, 6(2), 69–77. [https://doi.org/10.1016/S1463-0184\(03\)00068-6](https://doi.org/10.1016/S1463-0184(03)00068-6)

Scholfield, M. R., Vander Zanden, C. M., Carter, M., & Ho, P. S. (2013). Halogen bonding (X-bonding): A biological perspective. *Protein Science*, 22(2), 139–152. <https://doi.org/10.1002/pro.2201>

Shinada, N. K., De Brevern, A. G., & Schmidtke, P. (2019). Halogens in Protein-Ligand Binding Mechanism: A Structural Perspective. *Journal of Medicinal Chemistry*, 62(21), 9341–9356. <https://doi.org/10.1021/acs.jmedchem.8b01453>

Voth, A. P. (2007). The Role of Halogen Bonding in Inhibitor Recognition and Binding by Protein Kinases. *Current Topics in Medicinal Chemistry*, 7(14), 1336–1348. <https://doi.org/10.2174/156802607781696846>

Voth, A. R., Khuu, P., Oishi, K., & Ho, P. S. (2009). Halogen bonds as orthogonal molecular interactions to hydrogen bonds. *Nature Chemistry*, 1(1), 74–79. <https://doi.org/10.1038/nchem.112>

Wester, M. R., Johnson, E. F., Marques-Soares, C., Dijols, S., Dansette, P. M., Mansuy, D., & Stout, C. D. (2003). Structure of Mammalian Cytochrome P450 2C5 Complexed with Diclofenac at 2.1 Å Resolution: Evidence for an Induced Fit Model of Substrate Binding†,‡. *Biochemistry*, 42(31), 9335–9345. <https://doi.org/10.1021/BI034556L>

Wilcken, R., Zimmermann, M. O., Lange, A., Joerger, A. C., & Boeckler, F. M. (2013). Principles and applications of halogen bonding in medicinal chemistry and chemical biology. *Journal of Medicinal Chemistry*, 56(4), 1363–1388. <https://doi.org/10.1021/jm3012068>

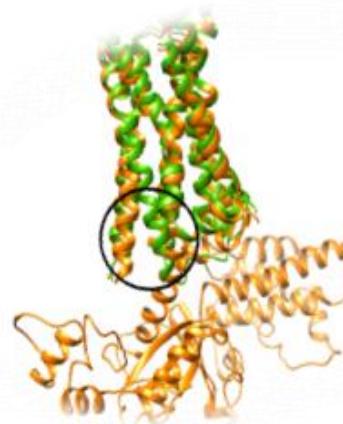
Wu, W., Samoszuk, M. K., Comhair, S. A. A., Thomassen, M. J., Farver, C. F., Dweik, R. A., Kavuru, M. S., Erzurum, S. C., & Hazen, S. L. (2000). Eosinophils generate brominating oxidants in allergen-induced asthma. *Journal of Clinical Investigation*, 105(10), 1455–1463. <https://doi.org/10.1172/JCI9702>

Xu, Z., Yang, Z., Liu, Y., Lu, Y., Chen, K., & Zhu, W. (2014). Halogen bond: Its role beyond drug-target binding affinity for drug discovery and development. *Journal of Chemical Information and Modeling*, 54(1), 69–78. <https://doi.org/10.1021/ci400539q>

Zimmermann, M. O., Lange, A., & Boeckler, F. M. (2015). Evaluating the potential of halogen bonding in molecular design: Automated scaffold decoration using the new scoring function xbscore. *Journal of Chemical Information and Modeling*, 55(3), 687–699. <https://doi.org/10.1021/ci5007118>

Capítulo 5

Docking molecular de ligandos
halogenados con receptores de
dopamina



CAPÍTULO 5

En este capítulo se estudiaron las interacciones de EX formadas por ligandos halogenados unidos al receptor D2 de dopamina (D2R). Por su extensión y para facilitar su comprensión, se presenta subdividido en tres partes. La primera está enfocada en el receptor; debido a la falta de disponibilidad de estructuras cristalinas de los receptores de dopamina al momento del inicio de esta tesis, la primer parte consistió en la generación de modelos por homología del receptor de dopamina D2 (D2R) en su forma inactiva (D2Ri) y activa (D2Ra), esta última acoplada a la proteína Gi. Posteriormente se estudiaron los cambios conformacionales de cada receptor al unirse a un compuesto agonista, en el caso del receptor activo, y a un agonista inverso en el caso del receptor inactivo. Durante el desarrollo de esta tesis se cristalizaron las estructuras del receptor D2 (Wang et al., 2018; Yin et al., 2020) esto despertó el interés en estudiar también los cambios conformacionales involucrados en la activación del receptor por el ligando, la evaluación de la calidad de los modelos obtenidos por homología en la primera parte, además de los interruptores moleculares. La segunda parte se centró en el docking molecular de los ligandos dopaminérgicos halogenados, LDX, tomando en consideración **los parámetros para halógenos** incluidos en la función del campo de fuerzas, como se describe en el capítulo 3. La tercera parte consistió en la aplicación del Protocolo predictivo de formación de EX basado en el entorno (Xenv), para predecir **la pose con el entorno más propicio del LDX en el sitio ortoestérico del receptor**, como se describe en el capítulo 4. Finalmente los resultados predichos se confirmaron mediante cálculos mecanocuánticos de la densidad electrónica, utilizando QTAIM.

5.1. Introducción

5.1.1 Receptores acoplados a la proteína G (GPCR)

Los receptores de Dopamina pertenecen a una gran familia de receptores denominados Receptores acoplados a la proteína G que se caracterizan por poseer siete segmentos transmembranas (7TM) relativamente hidrofóbicos los cuales están ubicados en la membrana celular. Estos segmentos acoplados a la proteína G (Šukalović et al., 2005) interactúan con moléculas efectoras citoplasmáticas que regulan funciones neuronales, de este modo median la mayoría de las respuestas celulares a las hormonas y a los neurotransmisores, como los sentidos de la vista, el olfato y el gusto y por lo tanto están asociados con una multitud de enfermedades que hacen de los miembros de esta familia, importantes blancos farmacológicos. Se estima que los GPCRs son blancos de alrededor del 50% de los fármacos que se encuentran actualmente en el mercado (Cheng et al., 2010). El nombre GPCR se refiere a la forma de señalización de estos receptores a través de proteínas de unión a GTP (guanosina trifosfato) que se encuentran en el interior celular (Figura 5.1).

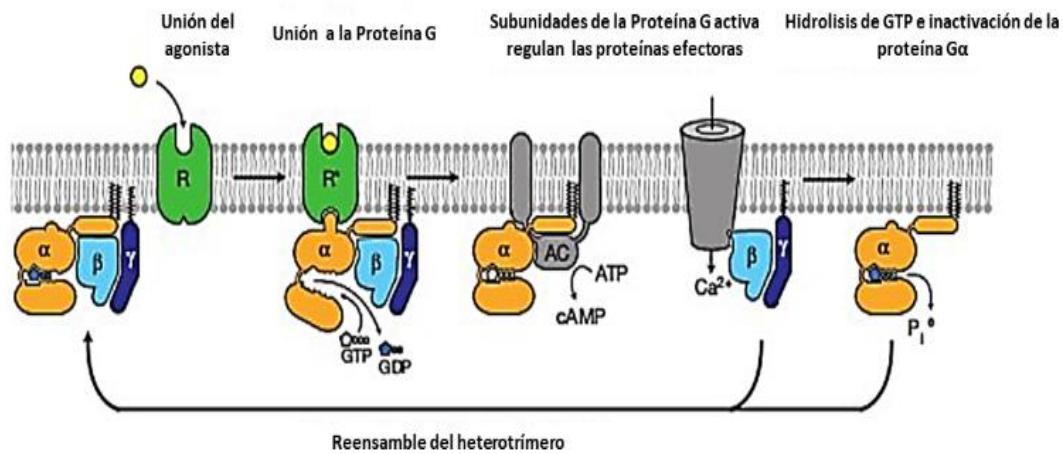


Figura 5.1: Ciclo de la proteína G. La unión del agonista extracelular al receptor R provoca el reordenamiento de los segmentos transmembrana en sus extremos citoplasmáticos, lo cual permite que el heterotrimero de proteína G (α , β y γ) se una al receptor. Luego de la formación del complejo R-G, el nucleótido GDP se libera de la subunidad G α . El GTP se une a G α resultando en la disociación de las subunidades α y β del receptor. Las subunidades regulan sus

respectivas proteínas efectoras, adenil ciclase (AC) y canales de Ca^{2+} . Posteriormente, la actividad GTPasa intrínseca de $\text{G}\alpha$ lleva a la hidrólisis de GTP en GDP y el reensamblaje del heterotrímero a partir de las subunidades α y β .

Los GPCRs son proteínas dinámicas que pueden adoptar un número de conformaciones en la membrana, dos de ellas son las más relevantes llamadas conformación activa e inactiva, siendo solo la forma activa capaz de transmitir “la señal” al interior celular. Los ligandos se clasifican de acuerdo con el efecto que tienen sobre el equilibrio de conformaciones del receptor. Los agonistas inversos se unen y estabilizan la forma inactivada del receptor con lo cual reducen la actividad basal o constitutiva del receptor.

Los antagonistas neutros no tienen efecto sobre la actividad de señalización, pero pueden evitar que otros ligandos se unan al receptor. Por el contrario, la unión de un agonista aumenta la afinidad por la proteína G en el interior de la célula y favorece la conformación del receptor activado lo cual se traduce en un aumento de la respuesta biológica por encima del nivel basal (Figura 5.2).

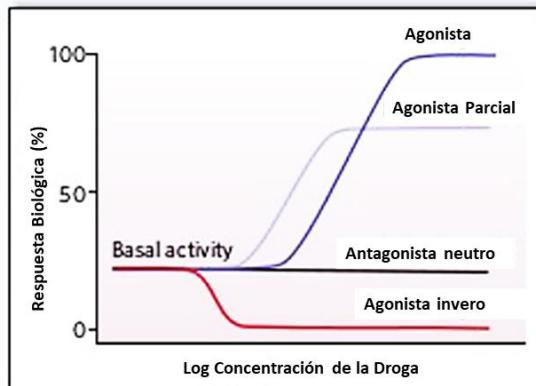


Figura 5.2: Efecto de los diferentes tipos de ligandos en la actividad del receptor

El estado activo de un GPCR puede definirse como la conformación que acopla y estabiliza la proteína G libre de nucleótido (Rasmussen et al., 2011a). El mecanismo general de activación de los GPCRs se debe a formación del “complejo ternario” formado por el agonista extracelular, el GPCR transmembrana y la proteína G intracelular. (De Lean et al., 1980) . En este complejo ternario, la proteína G tiene mayor afinidad por GTP que por GDP (ver figura 5.1) y el GPCR tiene varias veces mayor afinidad por el agonista que el receptor solo. La mayoría de los GPCRs resueltos proporcionan una visión del receptor en su forma inactivada. Esto se debe a la inestabilidad inherente del estado activado del GPCR en ausencia de proteína G (solo unido a

agonista), lo cual hace difícil capturar dicho estado en una estructura cristalina (Rasmussen et al., 2011).

Debido a los desafíos técnicos de trabajar con complejos de señalización nativos que incluyan a la proteína G, se han desarrollado varios sustitutos que imitan la proteína G para simplificar la cristalización de los GPCR en su estado activo (Carpenter & Tate, 2017). La primera estructura resuelta de un GPCR en estado activado fue la opsina (Scheerer et al., 2008). En 2011, la estructura del receptor adrenérgico $\beta 2$ ($\beta 2$ -AR) se determinó en complejo con un nanoanticuerpo (Rasmussen et al., 2011) y con Gs heterotrimérica (Rasmussen et al., 2011). Desde entonces se han determinado tres GPCRs adicionales en su estado activo, el receptor muscarínico de acetilcolina M2 (Kruse et al., 2013), el receptor μ -opioide (Huang & Thathiah, 2015) (Carpenter et al., 2016) y recientemente, durante el desarrollo de esta tesis, el receptor D2 de dopamina (Yin et al., 2020), el cual se encuentra cristalizado formando un complejo con el compuesto agonista bromocriptina.

La cuestión fundamental acerca del mecanismo de activación de GPCRs sigue siendo entender ¿cómo la unión de un agonista y los cambios resultantes a nivel de las interacciones del ligando en el bolsillo de unión conducen a cambios conformacionales que se propagan desde la porción extracelular del receptor hacia la superficie citoplasmática involucrada en la unión a proteína G?

Los métodos bioquímicos y cristalográficos junto con las simulaciones de dinámica molecular y otras técnicas teóricas proporcionaron modelos de activación del receptor basados en la acción de los llamados "interruptores moleculares" que están "escondidos" en la estructura del receptor (Trzaskowski et al., 2012).

El análisis topológico de la densidad electrónica a partir de la teoría de Bader de átomos en moléculas, puede ser determinante en este estudio. La información crucial que brinda predice cómo el efecto de la interacción intermolecular redistribuye la densidad de carga electrónica del ligando en el sitio activo de la proteína. La caracterización topológica de la densidad electrónica del sitio activo con el ligando, a su vez define el enlace químico entre ellos; es decir, el análisis topológico de interacciones intermoleculares confirma la existencia de tales interacciones. La densidad electrónica de las interacciones intermoleculares se correlaciona en gran medida con la energía de disociación de los grupos que interactúan. Además, el análisis de densidad electrónica del ligando y del bolsillo de unión de la proteína (complejo ligando-proteína) permite también comprender las propiedades electrostáticas del ligando en el sitio

activo y la superficie de potencial electrostático de la molécula proporciona la forma o figura de localización de los sitios reactivos.

Los interruptores moleculares o switchs descriptos hasta el momento incluyen: el bloqueo iónico (ionic lock switch), el interruptor de transmisión (transmission switch o rotamer toggle switch), el bloqueo 3-7 (3-7 lock switch) y el interruptor de tirosina (tyrosine toggle switch) (Figura 5.3).

Dado que los interruptores moleculares están usualmente altamente conservados dentro de los GPCRs (particularmente en los GPCR de tipo A, de la familia de la Rodopsina), se piensa que éstos serían parte de una vía o mecanismo común de activación de los GPCRs. Sin embargo, varios experimentos posteriores mostraron que probablemente no exista tal vía común para todos los miembros de la familia de GPCRs(Trzaskowski et al., 2012) .

A continuación, se describen los interruptores más importantes:

- *Interruptor de bloqueo iónico:* En la superficie citoplásica, una diferencia estructural importante entre los GPCRs activados por ligando y GPCRs en estado inactivado reside en el “bloqueo iónico”, una fuerte interacción intramolecular entre residuos E3.49 / R3.50 del motivo (e/d)Ry en TM3 y los residuos Glu6.30 / T6.34 en TM6 que mantiene juntas ambas hélices transmembrana estabilizando así al receptor en su estado inactivado (Figura 5.3). La ruptura de esta interacción provoca el desplazamiento de TM6 hacia afuera del ovillo de TMs (alrededor de 14Å en el complejo ternario agonista β 2-AR-Gs (Rasmussen et al., 2011) y de 6Å en una forma activada de rodopsina(Park et al., 2008); TM5 también sufre un desplazamiento menor. El resultado de los cambios conformacionales genera una cavidad entre TM3, TM5 y TM6 en la cual la proteína G puede anclarse.
- *Interruptor de transmisión:* Este interruptor que involucra la secuencia conservada cwxP, describe los cambios conformacionales inducidos por el agonista en las interfaces inter-hélice que dan como resultado la rotación de TM6 y el rearreglo de TM5 y TM7 (Figura 5.3). El movimiento de Trp6.48 conduce a la reubicación de una **red de enlaces de hidrógeno mediada por agua**, que en estado activo abarca desde el sitio del ligando hacia el sitio de unión a proteína G, seguido del movimiento de W6.48 en rodopsina el residuo F6.44 situado en un giro de hélice hacia el lado citoplásico de TM6, rota hacia Leu5.51 ya que todo el TM6 está girando horizontalmente, mientras que I3.40 se aleja de P5.50 iniciando así un cambio estructural en TM5 inducido por la prolina. Tales

translocaciones forman parte del "interruptor de transmisión" descripto por Deupi y Standfuss (Deupi & Standfuss, 2011) (previamente conocido como "rotamer toggle switch").

- ✚ **Interruptor de bloqueo 3-7:** Este interruptor se denomina bloqueo 3-7 porque involucra a los segmentos TM3 y TM7. En los receptores aminérgicos (ej. de dopamina, histamina H1, etc), el interruptor se compone generalmente de los residuos Y7.43 (TM7) y D3.32 (TM3) que forman un puente entre ambos segmentos en la forma inactivada del receptor que tras la activación se rompe.
- ✚ **Interruptor de tirosina y barrera hidrofóbica:** Para explicar el **reordenamiento de enlaces de hidrógeno mediado por moléculas de agua** durante la activación de GPCRs, se ha propuesto un interruptor denominado "tyrosine toggle switch" que involucra la secuencia conservada nPxxy. En la forma inactivada de GPCRs una región llamada *barrera hidrofóbica* separa la red de enlaces de hidrógeno mediados por moléculas de agua del motivo (d/e)Ry que es crítico para la activación de los GPCRs.

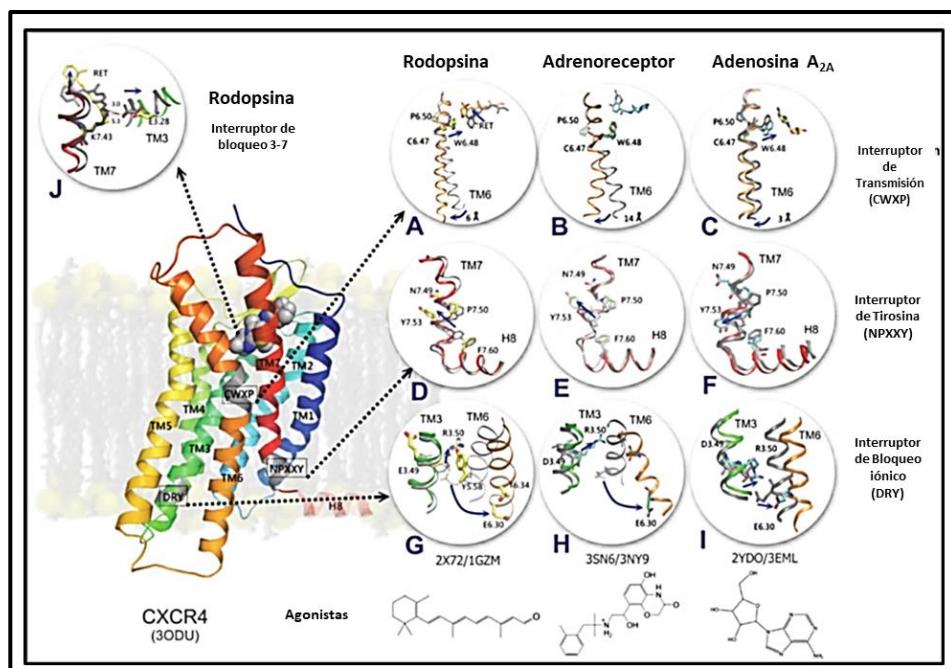


Figura 5.3: Acción de interruptores moleculares en GPCRs. Se muestran cuatro interruptores: interruptor de transmisión, interruptor de palanca de tirosina, bloqueo iónico y bloqueo 3-7. Se muestran basados en las estructuras cristalinas de la rodopsina, Adrenoreceptor y receptor de Adenosina A2 con agonistas, antagonistas, agonistas inversos. Se proporcionan sus números de Id del Protein Data Bank: primer número para inactivo y segundo para receptor activo. Además, se muestran las fórmulas estructurales de los agonistas de las estructuras

cristalinas de los receptores activos. El esquema general de la estructura de GPCR se muestra en base a la estructura cristalina del receptor de quimiocinas CXCR4 con un ligando pequeño. Las flechas azules en paneles circulares indican movimientos de la estructura del receptor durante la acción de un interruptor en particular. La estructura inactiva del receptor se muestra en gris mientras la activa se encuentra en color. Los residuos están numerados de acuerdo con el esquema de numeración de Ballesteros-Weinstein. (Figura reproducida de Trzaskowski et al, 2012). Los interruptores moleculares junto con la red de enlaces de hidrógeno entre residuos conservados, motivos y moléculas de agua estructurales constituyen una interfaz extendida entre diferentes áreas de los GPCRs que facilitan los movimientos concertados que conectan la unión del ligando con la señalización celular.

5.1.2 Receptores Dopaminérgicos

La Dopamina o 3,4-dihidroxifeniletilamina (Figura 5.4), es una de las catecolaminas naturales que se biosintetizan por acciones enzimáticas en las neuronas terminales dopaminérgicas usando como precursor L-Tirosina proveniente del aminoácido esencial Tirosina. La Dopamina se sintetiza en neuronas específicas, localizadas fundamentalmente en el cuerpo estriado y en el sistema límbico, almacenándose en las vesículas sinápticas. Este neurotransmisor catecolaminérgico precursor de la síntesis de noradrenalina y neuromediador del sistema nervioso central y periférico, se libera a la hendidura sináptica por influencia de estímulos nerviosos.

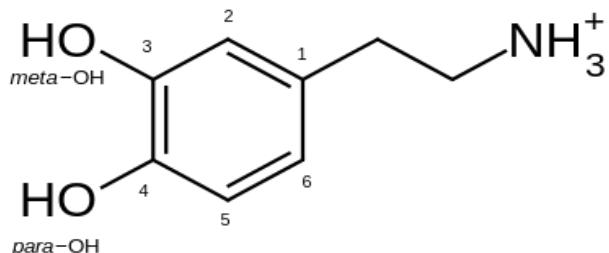


Figura 5.4: Estructura de la dopamina.

La neurotransmisión dopaminérgica juega un papel muy importante en los desórdenes que involucran el sistema nervioso central (SNC) tales como la esquizofrenia, la enfermedad de Parkinson, el síndrome de Tourette y las adicciones psicoestimulantes. Las cuatro vías dopaminérgicas principales son la vía nigroestriada, vía mesolímbica, vía mesocortical y vía tuberoinfundibular (Schertler, 1999; Teeter et al., 1994), por ejemplo la Esquizofrenia ocurre por un aumento de los niveles de dopamina en las vías del sistema mesolímbico, mientras que

en la enfermedad de Parkinson existe una disminución de dopamina en las neuronas nigroestriadas.

La clasificación farmacológica más utilizada de los receptores dopaminérgicos es la de Kebabian y Calne(Andujar et al., 2008) que los subdivide en dos subtipos D1 y D2, con diferentes propiedades bioquímicas y farmacológicas y con distintas funciones fisiológicas. En la actualidad, la aplicación de técnicas de clonación ha permitido identificar 5 subtipos de receptores clasificados en dos subfamilias, tipo D1 (D1 y D5) y tipo D2 (D2, D3 y D4) (Lan et al., 2006). Los receptores similares a D1 se acoplan a la subunidad alfa estimuladora de la Proteína G (Gs), activando la adenilciclasa; mientras que los receptores similares a D2 se acoplan a la subunidad alfa inhibidora de la proteína G (Gi), inhibiendo de este modo la adenilciclasa (Chien et al., 2010).

Los receptores tipo D1 presentan 80 % de similitud entre si y propiedades fisiológicas similares. Poseen alta afinidad por los ligandos benzacepínicos, los cuales son antagonistas selectivos para dichos subtipos. Presentan además una moderada afinidad por los agonistas dopaminérgicos típicos como la apomorfina. Existen pequeñas diferencias de afinidad de algunos compuestos para los receptores D1 y D5 y estas se han observado especialmente en algunos agonistas. Los receptores D1 son postsinápticos y se distribuyen en casi todas las áreas cerebrales que poseen sinapsis dopaminérgica. Su estimulación permite la activación de la adenilciclasa a través de una proteína Gs produciéndose por lo tanto un aumento del AMPc intracelular en la neurona postsináptica. Los receptores D5 se diferencian de los D1 en que poseen una mayor afinidad por la dopamina y se localizan únicamente en el cortex frontal, cuerpo estriado, hipotálamo e hipocampo (Yashar et al., n.d.). Las funciones de los receptores D5 no son bien conocidas, sin embargo, se sabe que los receptores D1 están implicadas en importantes acciones de la dopamina relacionadas con el control de movimiento, así como funciones cognitivas y cardiovasculares.

En cuanto a los receptores tipo D2, se conoce que D2 y D3 muestran un 75% de similitud, mientras que entre los receptores D2 y D4 solo hay un 53% de similitud, pero todos ellos presentan propiedades farmacológicas similares. Las butirofenonas son antagonistas que poseen alta afinidad por dichos receptores al igual que las fenotiazinas y los tioxantenos. Sin embargo, existen algunas diferencias entre las afinidades de estos tres receptores (D2, D3 y D4) con algunos ligandos específicos. La benzamida raclopride por ejemplo, presenta una alta afinidad por los receptores D2 y D3 y baja por los receptores D4. En general los receptores D2 presentan una fuerte afinidad frente a los neurolépticos, así como la dopamina y la mayoría de

los agonistas dopaminérgicos. Los receptores tipo D2 se encuentran distribuidos mayoritariamente en los sistemas nigroestriado y mesolímbico, se localizan tanto a nivel presináptico como postsináptico y su estimulación inhibe la adenilciclasa a través de una proteína G, disminuyendo el contenido de AMPc intracelular. Poseen una función importante en el control de movimientos, en ciertos aspectos del comportamiento cerebral y en la secreción de prolactina. Las funciones de los receptores D3 y D4 son desconocidas, aunque su localización en el área límbica del cerebro les sugiere papeles cognitivos, emocionales y de comportamiento. Los receptores tipo D2 presentan una gran afinidad por los fármacos utilizados para tratar la esquizofrenia y Parkinson.

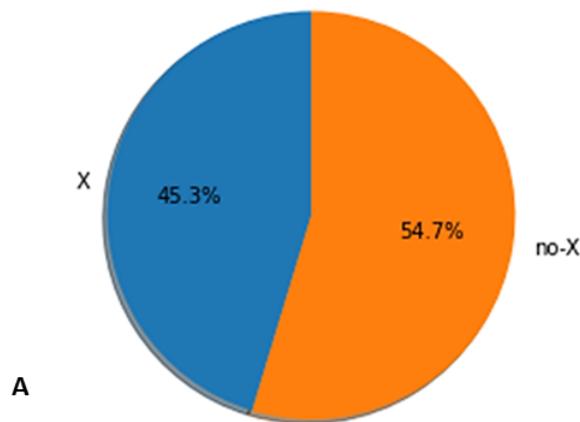
Actualmente, el receptor D2 (D2R) es de gran interés terapéutico, incluido sus principales sitios de acción para la mayoría de los nuevos fármacos que se están desarrollando con actividad antiparkisoniana y antipsicóticos. Se ha creado con gran éxito un gran número de agonistas de los receptores D2, los cuales reducen los síntomas característicos de la enfermedad de Parkinson como son la bradiquinesia, la rigidez y los temblores, como así también otros síntomas menos comunes como son los trastornos de movimiento incluyendo el síndrome de Ekbon (Cox et al., 1992; Mansour et al., 1992). Para el tratamiento de la esquizofrenia todavía se utilizan neurolépticos tradicionales, flufenazina y haloperidol y otros más modernos antipsicóticos de segunda generación, los cuales son menos potentes como agentes neurolépticos, pero son fuertes antagonistas de los receptores de serotonina.

5.1.3 Relevamiento de ligandos dopaminérgicos halogenados

Un relevamiento de la base de datos ChEMBL (Gaulton et al., 2012) muestra que de los 11733 compuestos con anotaciones de afinidad frente a D2R (con valores de K_i), 5317 de ellos contienen, al menos, un átomo de halógeno en su estructura (Figura 5.5A).

Por otra parte, de los 57 fármacos y candidatos clínicos actualmente disponibles, incluyendo agonistas, agonistas parciales y antagonistas de D2R, 33 de ellos contienen, por lo menos, un átomo de halógeno en su estructura (Figura 5.5B)

Compuestos con datos de afinidad a D2R (ki)



Drogas seleccionadas para ensayos clínicos

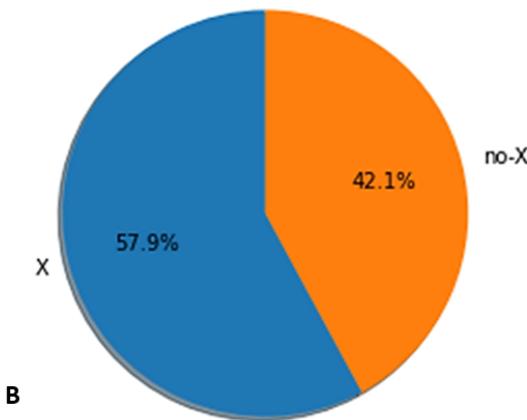


Figura 5.5: Relevamiento de la base de datos ChEMBL de compuestos ligandos de D2R, halogenados (X) y no-halogenados (no-X).

La elevada frecuencia de átomos de halógeno en los ligandos dopaminérgicos obedece principalmente a que su incorporación en la etapa de optimización estructural del líder permite mejorar sus propiedades ADMET, facilitando la absorción y el cruce de las barreras biológicas, y prolongando, además, su vida útil.

Sin embargo, como se expresó previamente a lo largo de esta tesis, el átomo de halógeno puede interaccionar de manera no covalente (a través de enlaces de hidrógeno y de halógeno) con los átomos de la proteína blanco, influyendo también sobre la afinidad de la unión ligando-receptor. En otras palabras, la incorporación de un átomo de halógeno en la estructura de un ligando dopaminérgico influye no solo sobre su farmacocinética sino también sobre su farmacodinamia. Más aún, como influye la halogenación sobre la afinidad de unión del ligando al blanco terapéutico, no ha sido todavía, suficientemente estudiado.

Surge entonces en esta tesis, el particular interés de abordar esta cuestión. Más específicamente se pretende saber si ¿es factible la formación de una interacción de enlace de halógeno en el sitio ortostérico de D2R? y ¿qué implicancias puede tener su formación desde el punto de vista biológico?

Para ello, una primera cuestión que se debemos conocer es ¿cuántos de los 5571 ligandos dopaminérgicos halogenados podrían potencialmente formar un enlace de halógeno?

Entre los halógenos, el átomo de flúor reúne características especiales: es el más electronegativo de la tabla periódica, es el más compacto y el menos polarizable de ellos; lo que determina que solo en determinadas situaciones se ha demostrado su capacidad formadora de EX, por lo que no será considerado en este estudio.

Por otra parte, es también importante el efecto atractor de electrones del átomo unido covalentemente al halógeno (dador de halógeno), cuanto mayor su electronegatividad mayor será el tamaño del agujero- σ . Entre todos los compuestos, los sistemas halogenados aromáticos, representan los andamiajes más comunes en química medicinal, forman enlaces de halógeno relativamente estables, debido a la mayor electronegatividad del carbono aromático (Car) en relación a uno alifático (Cal). Aplicando los filtros previos, es decir considerando únicamente los ligandos que contienen la siguiente subestructura: Car-X donde X = Cl, Br o I, el número de dadores de halógeno potenciales se reduce a 3199, que representa alrededor del 60% del total de ligandos dopaminérgicos halogenados (Figura 5.6)

Ligandos dopaminergicos potenciales dadores de halógeno

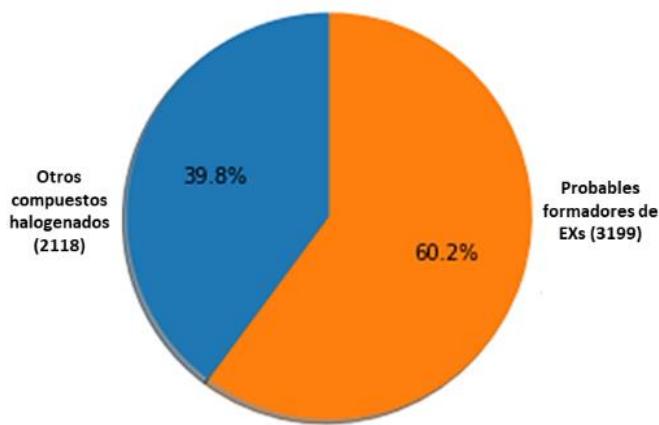


Figura 5.6: Distribución de los Ligandos dopaminergicos halogenados según su potencialidad de formar EX.

Para conocer si estos potenciales dadores de halógeno pueden efectivamente establecer un enlace de halógeno, se debería modelar explícitamente al ligando en el sitio orto estérico del receptor D2R y considerar explícitamente sus interacciones intermoleculares (Figura 5.7)

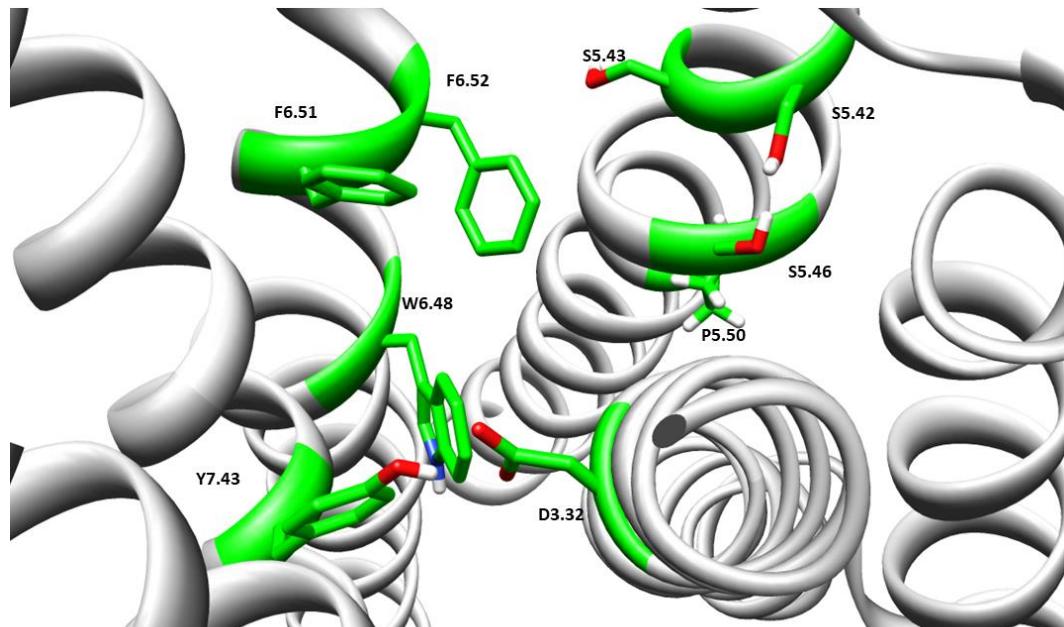


Figura 5.7: Sitio de unión de ligandos del D2R, se detallan los residuos más relevantes de acuerdo con el esquema de numeración de Ballesteros-Weinstein.

En la siguiente sección se describen las interacciones de algunos potenciales dadores de halógeno en el sitio ortostérico de D2R, que fueron estudiados por el grupo de investigación, en trabajos previos. (Angelina et al.). Para comprender mejor el rol que jugaba el átomo de halógeno en la unión ligando-receptor, la estrategia adoptada en los trabajos mencionados fue evaluar el efecto de la sustitución del átomo de halógeno por un grupo OH, analizando a la vez cómo se modificaban las afinidades de unión con dicha sustitución.

5.1.4. Rol del halógeno en la interacción ligando-D₂R

Angelina et al. (Angelina et al., 2015) estudiaron el efecto del reemplazo de los grupos OH de la dopamina por un átomo de halógeno y encontraron que al reemplazar el OH en posición *meta* por Cloro (*m*-clorotiramina) la afinidad del ligando por D2R se conservaba, sin embargo cuando se reemplaza el OH en posición *para* (*p*-clorotiramina), ésta se reducía notablemente.

El análisis basado en la densidad electrónica del entorno del átomo de halógeno muestra que *p*-Cl está anclado en un sub bolsillo altamente hidrofóbico del D2R, donde el halógeno forma solo enlaces de hidrógeno del tipo C-H...Cl. Por el contrario, *m*-Cl está alojado en un sub bolsillo más polar en donde forma interacciones hidrofóbicas, pero también un enlace de halógeno O...Cl con Ser5.42 del segmento transmembrana 5 (TM5).

Si bien en el trabajo descrito se logró explicar las diferencias en afinidades de unión en término de interacciones moleculares, es importante aclarar que en estos estudios anteriores, se utilizó el campo de fuerzas de AMBER sin corregir o sin modificar, en las simulaciones de dinámica molecular, es decir, se lo consideró como un átomo con distribución de carga esférica.

Esta tesis aporta (en el capítulo 3) una mejora en la herramienta de estudio al tomar en cuenta el carácter anisotrópico de la distribución de carga electrónica del átomo de halógeno incorporando la aproximación del **extra punto positivo** (PEP) de Ibrahim. .

Empleando el protocolo de Ibrahim (Ibrahim, 2012) en las simulaciones de dinámica molecular, Luchi et al. (Luchi et al., 2016) estudiaron la formación de enlaces de halógeno en el receptor D2R ,modelado en base al D3R,(Chien et al., 2010).

Para poner en evidencia el rol del enlace de halógeno en la estabilización de los complejos ligando-receptor, se realizó una búsqueda en bases de datos públicas de pares de ligandos dopaminérgicos que difieran únicamente en la presencia de un halógeno (LX) o un hidroxilo (LOH), en la misma posición. En particular, se seleccionaron pares LX/LOH en los cuales el

compuesto halogenado tenía mayor afinidad de unión por D2R (medida experimentalmente) que su contraparte hidroxilado (Figura 5.8)

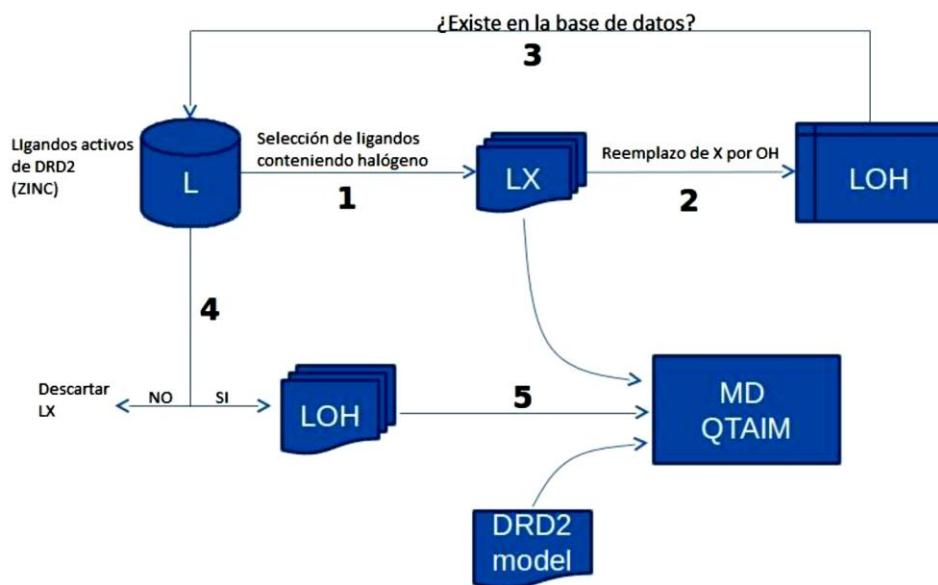


Figura 5.8: Diagrama de flujo empleado para la compilación de los pares de ligandos LX/LOH tomados de la base de datos ZINC (figura tomada de Luchi et al., 2016)

El análisis de las trayectorias de dinámica molecular mostró que los compuestos halogenados parametrizados con un extra-punto forman un enlace de halógeno X...O con el oxígeno del backbone del residuo Ser5.42 mientras que su contraparte hidroxilada no tiene la misma tendencia a formar el correspondiente enlace de hidrógeno OH...O.

A su vez, las simulaciones muestran que este enlace de halógeno atenúa o reduce la inherente propensión a desplegarse/desenrollarse de la parte extracelular del segmento de transmembrana 5 (TM5) de D2R. *Estos resultados sugieren un posible rol del enlace de halógeno como modulador de la estructura secundaria de las proteínas, como consecuencia de la capacidad del halógeno para interactuar con el backbone de las proteínas.*

La incorporación de un extra-punto en los campos de fuerza de los algoritmos de docking y en las simulaciones de dinámica molecular constituye ciertamente una mejora en el tratamiento del enlace de halógeno, con respecto a los campos de fuerza de potencial esférico clásicos. Sin embargo, se hace mucho énfasis en esta tesis en que el entorno del enlace de halógeno es tanto o más importante que el enlace de halógeno en sí mismo. En este capítulo se aplica el protocolo de detección de enlace de halógeno basado en el entorno (EXent) para predecir el modo de unión de ligandos halogenados sobre D2R.

En los últimos años, la irrupción de la tecnología de criomicroscopía electrónica (cryo-EM) ha acelerado la resolución de la estructura tridimensional de las proteínas, a tal punto que actualmente se encuentran disponibles las estructuras del D2R tanto en su forma inactiva (D2Ri) como su forma activada acoplada a la proteína Gi (D2Ra-Gi). Esta situación o permitiría predecir con mayor exactitud el modo de unión y las interacciones de ligandos halogenados con D2R, respecto a los estudios previos realizados sobre receptores modelados por homología.

Más aún, la disponibilidad de ambas formas del receptor, activado (7jvr) e inactivado (6cm4), proporciona una oportunidad para intentar predecir no sólo la afinidad de unión de los ligandos halogenados sino también su perfil funcional, es decir si actuarían como agonistas o antagonistas del D2R.

En este capítulo se simularon por Dinámica Molecular ambas estructuras del receptor D2R inserto en una membrana fosfolipídica y se analizaron sus diferencias estructurales en término de los *interruptores moleculares*.

Luego se modelaron las interacciones moleculares de la dopamina sobre diferentes conformaciones de ambas formas del receptor, verificando que ésta se une con mayor afinidad a la forma activada, como era de esperarse, por tratarse del agonista endógeno del D2R.

Posteriormente se realizó el docking del análogo halogenado de dopamina sustituido por cloro en posición *meta* (***m*-clorotiramina**) para investigar a qué forma del receptor se uniría más fuertemente, o con mayor afinidad con el fin de predecir si tras la halogenación, se conservaría su función como agonista, o si por el contrario se convertiría en un agonista inverso/antagonista.

De esta manera se constató que el docking sobre un “ensemble” de conformaciones activas e inactivas del receptor fue capaz de predecir no solo la afinidad de los ligandos dopaminérgicos sino también su perfil funcional.

Finalmente, en la última parte se aplicó el **protocolo de detección de enlace de halógeno basado en el entorno** para conocer cuál sería la orientación más probable del halógeno de *m*-Cl dentro del sitio ortostérico de D2R del receptor activado. En el caso del receptor inactivado se aplicó el protocolo para el agonista inverso, **clorpromazina**. Los resultados predictivos se confirmaron mediante cálculos mecano cuánticos de la densidad electrónica, con QTAIM

5.2 Métodos

5.2.1 Construcción de modelos de receptores de dopamina por modelado por homología

Al inicio del desarrollo de esta tesis no se encontraba resuelta la estructura del receptor D2R, por lo cual se debió construir un modelo por homología del receptor en las formas inactiva y activa acoplada a la proteína Gi para poder aplicar el estudio de compuestos halogenados. La figura 5.9 resume los pasos del procedimiento.

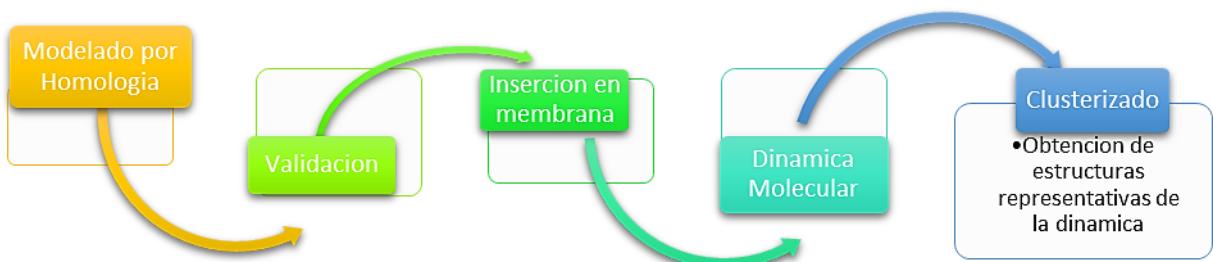


Figura 5.9: Esquema de los pasos de trabajo para obtener los modelos por homología

5.2.1.1 Modelado por homología

Se conocen al menos dos formas del receptor D2, D2A o forma corta y D2B o forma larga, la diferencia entre estas formas se basan en la constitución del loop intracelular, hasta el momento la forma larga del receptor se encuentra más implicada en patologías nerviosas (Soriano, 2011). En los últimos años se ha construido un modelo por homología de dicho receptor utilizando como molde las estructuras cristalinas del receptor D3 de dopamina (D3R), receptor $\beta 2$ adrenérgico y receptor A2 α de adenosina; posteriormente Ralf C. Kling describió el primer modelo de homología del D2R unido a la proteína G basándose en la estructura cristalina del receptor $\beta 2$ adrenérgico en complejo con la proteína Gs (Kling et al., 2013) y durante la realización de esta tesis se ha cristalizado el D2R en su forma inactiva y activa (Wang et al., 2018; Yin et al., 2020).

El modelado por homología se emplea para encontrar la conformación más exacta de una proteína (Orry & Abagyan, 2012). Entre los receptores dopaminérgicos el único cuya estructura

había sido resuelta hasta el momento en que se inició esta tesis era el D3R (Chien et al., 2010), por ese motivo y debido a la elevada homología de la secuencia entre los diferentes subtipos de receptores de dopamina, la estructura de D3R constituye una buena plantilla para modelar por homología la estructura del D2R. Sin embargo, D3R fue capturado en una forma inactiva con lo cual los modelos resultantes también estarán en dicha forma, por lo que, a diferencia de los trabajos citados previamente, la estrategia utilizada en este trabajo consistió en emplear más de una plantilla para poder modelar el D2R en su forma activa. Esta estrategia de alineamiento con plantillas múltiples implica la escisión del receptor en varios dominios y la posterior selección de la plantilla más adecuada para cada uno de estos dominios, aumentando así la precisión de los modelos (Orry & Abagyan, 2012). Aplicando esta estrategia se ha construido un modelo del receptor D2 acoplado a la proteína Gi, tomando como modelo las estructuras cristalizadas del DRD3 (PDB:3PBL),y la estructura del receptor $\beta 2$ adrenérgico acoplado a la proteína Gs (PDB:3SN6). A través del modelado por homología basado en múltiples plantillas se lograron obtener dos modelos del D2R, en su forma activada e inactivado, se realizó posteriormente un refinamiento manual del alineamiento, en el cual se corroboró la presencia de las secuencias conservadas correspondientes a cada hélice transmembrana. (Figura5.8). Para el modelo del D2R acoplado a la proteína Gi, la plantilla de $\beta 2$ -Gs permitió modelar el acoplamiento de D2R con la proteína G en la región intracelular más conservada entre los GPCRs mientras que la plantilla de D3R permitió modelar con más precisión la región extracelular de unión al ligando de D2R, la cual está menos conservada entre los GPCRs (Figura5.10)

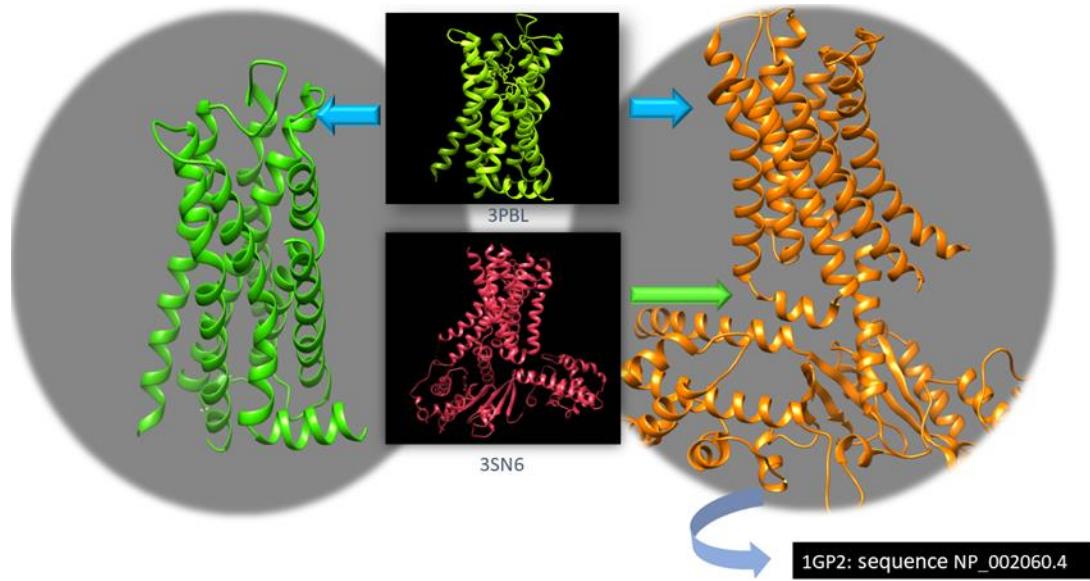


Figura 5.10: Moldes utilizados para modelado por homología para cada parte del receptor (3PBL,3SN6 y 1GP2), a la izquierda el modelo obtenido del receptor D2 en su forma inactiva y a la derecha el modelo obtenido en su forma activada acoplado a la proteína Gi.

En resumen, para construir el modelo del receptor activado se utilizaron dos plantillas, el receptor de dopamina D3 (PDB-ID: 3PBL), y el receptor B2 adrenergico (PDB-ID: 3SN6), acoplado a la proteína Gs, ambas estructuras cristalinas fueron obtenidas del Protein Data Bank y se les removió el loop intracelular 3, ya que se ha demostrado que su ausencia no interfiere en la unión del receptor a la Proteína G (Rubenstein et al., 1987).

Para la construcción de los modelos se utilizó Modeller v.9.15 (Šali & Blundell, 1993) se usó la secuencia de la isoforma larga del receptor D2 (NP_000786.1) la cual fue obtenida de la base de datos BLAST. Para el modelado de la proteína G se utilizó solamente la subunidad alfa del trímero de dicha proteína debido a que se ha demostrado que incluir solo esta región es suficiente para que se produzca la interacción entre la proteína G y el receptor (Herrmann et al., 2006)

Se alinearon las secuencias del receptor D2 y de la subunidad alfa inhibitoria (Gi) de la proteína G con las dos plantillas. A diferencia del receptor $\beta 2$ adrenérgico que se acopla a Gs, D2R se une a la proteína Gi, sin embargo ambas isoformas presentan alta homología se secuencia por lo tanto la proteína Gs acoplada a $\beta 2$ en la estructura 3SN6 puede utilizarse

para modelar la estructura de Gi acoplada a D2R. En el caso del modelo del receptor inactivado no se consideró la proteína G en el modelado, utilizando como única plantilla la estructura del D3R.

Los alineamientos obtenidos fueron sometidos a una inspección visual y a un refinamiento manual, teniendo en cuenta las secuencias conservadas en cada dominio de transmembrana (Figura 5.11).

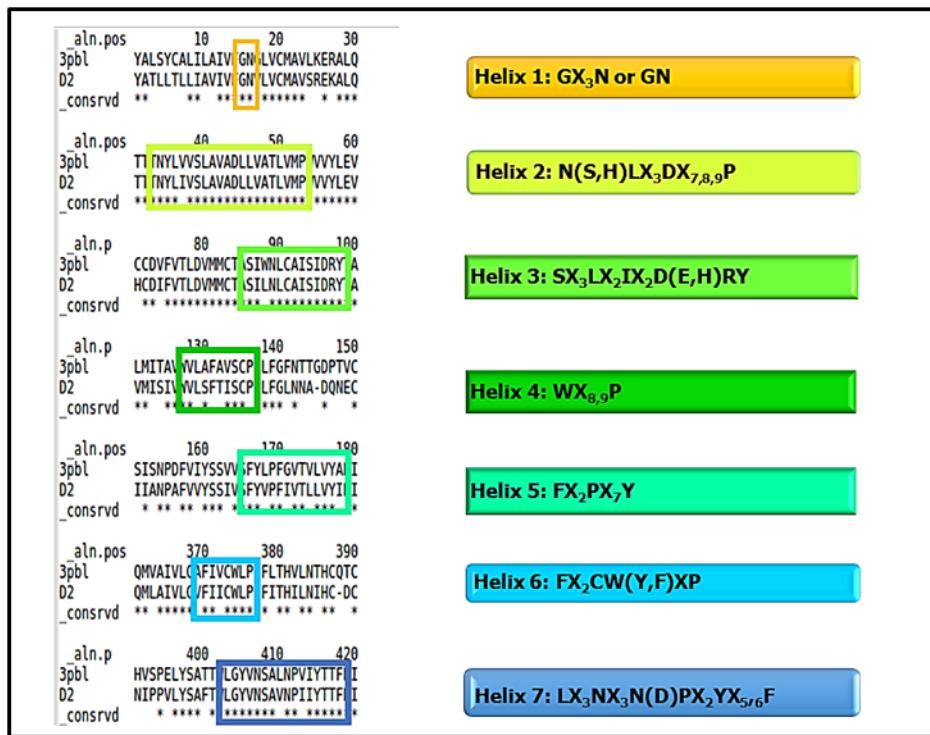


Figura 5.11: Secuencias conservadas de los receptores transmembrana presentes en el alineamiento del modelo del DR2.

5.2.1.2 Validación del modelo

Para analizar estas características, en primer lugar se hizo una selección de los mejores 5 modelos obtenidos en función de su puntuación DOPE (Discrete Optimized Protein Energy, de sus siglas en inglés), provista por el software Modeller (Šali & Blundell, 1993), la cual está diseñada para la selección de la mejor estructura de una colección de modelos construidos por este software; se seleccionó el modelo de valor más bajo de energía de cada uno y este fue sometido a la validación de sus parámetros energéticos y estereoquímicos con PROCHECK (Laskowski et al., 1993) (ANEXO B). Además,

posteriormente se comparó el modelo obtenido del D2R con la estructura recientemente cristalizada.

5.2.1.3 Inserción del receptor en la membrana

Posteriormente a haber obtenido ambos modelos del receptor de dopamina D2 acoplado a la proteína Gi y en su forma inactiva, cada uno fue insertado en una membrana biológica, con este fin sus coordenadas fueron preorientadas con respecto a las de una membrana normal con el servidor OPM (Database of orientation of proteins in membranes) (Lomize et al., 2006).

Cada receptor fue colocado dentro de una membrana heterogénea compuesta por CHL: POPC: POPE (25:50:50, replacement method) con el servidor CHARMM-GUI (Mahmood et al., 2013), el sistema fue solvatado con agua; se agregaron iones de K y Cl hasta obtener una concentración 0.15M.

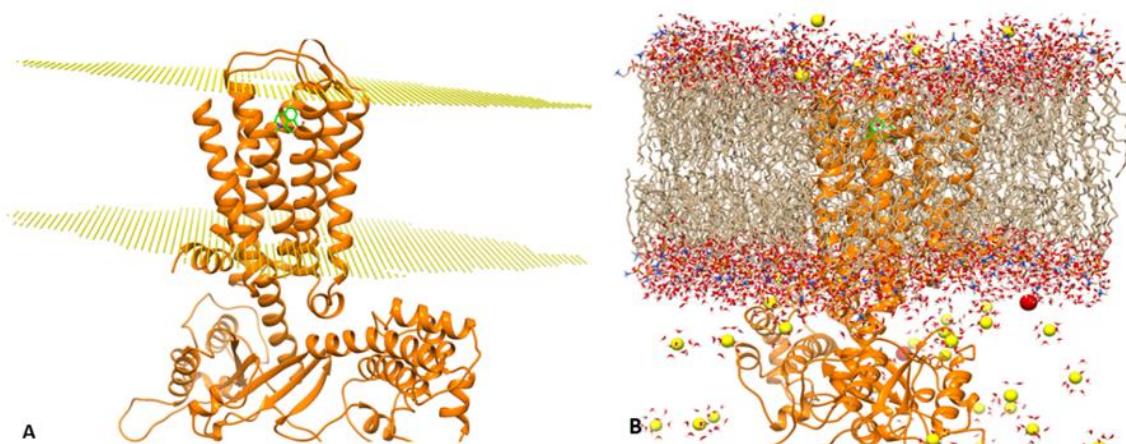


Figura 5.12: Modelo por homología del D2R orientado por OPM (A) y luego insertado en la membrana lipídica (B).

5.2.1.4 Dinámica Molecular

Con el fin de refinar los modelos de homología se realizaron simulaciones de dinámica molecular de los modelos embebidos en membrana. La intención es que las simulaciones permiten relajar los sistemas moleculares para permitirles que alcancen (o al menos se aproximen) a la estructura que estas presentan en su estado nativo.

Los sistemas para la simulación fueron preparados con ayuda del servidor CHARMM-GUI, se utilizó el campo de fuerza CHARMM 36 para proteínas, lípidos, agua e iones.

Las simulaciones fueron realizadas con AMBER 14, luego de la minimización el sistema fue sometido a seis pasos de equilibración seguido por el paso de producción de 100 ns. Los resultados de la trayectoria fueron analizados con CPPTRAJ, módulo de AMBER 14; para la visualización se utilizaron los programas Chimera (Pettersen et al., 2004) y Pymol (Schrodinger, 2010).

Con el módulo CPPTRAJ de AMBER 14 se realizó un clusterizado de las simulaciones de dinámica molecular de cada receptor, así se obtuvieron 10 estructuras representativas de la trayectoria de cada uno.

5.2.2 Refinamiento de los modelos de homología guiado por el ligando

En un último paso de refinamiento de los modelos de homología se realizaron simulaciones de un agonista y un agonista inverso en el sitio ortostérico de las formas activada e inactivada del receptor, respectivamente, con la intención de hacer más evidentes los cambios que ocurren durante la activación en el lado extracelular del receptor, donde se une el ligando. El siguiente esquema resume el procedimiento:

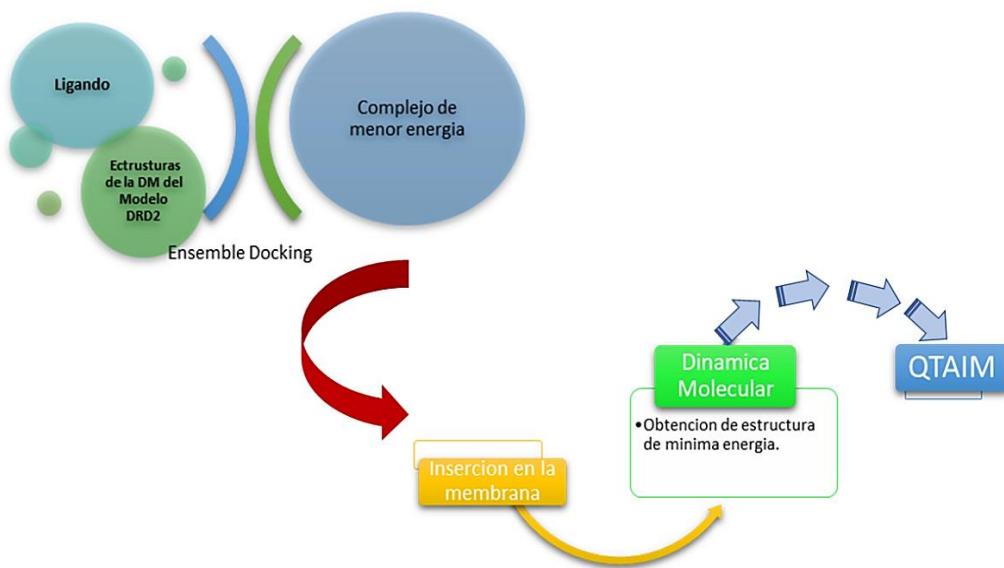


Figura 5.13: Esquema de trabajo para el refinamiento de los modelos por homología. Con flechas se indican los pasos subsiguientes, hasta realizar el estudio de interacciones intermoleculares

Para definir las modos de unión iniciales de los ligandos se realizó Docking Molecular de las estructuras resultantes del clusterizado de la dinámica molecular de los modelos 3D de los receptores, utilizando como ligandos un agonista como la apomorfina (Trujillo et al., 2000) para el caso del receptor en su forma activada, acoplado a la proteína Gi, y un agonista inverso, spiperona (Lan et al., 2006), en el caso del receptor de D2R en su estado inactivado.

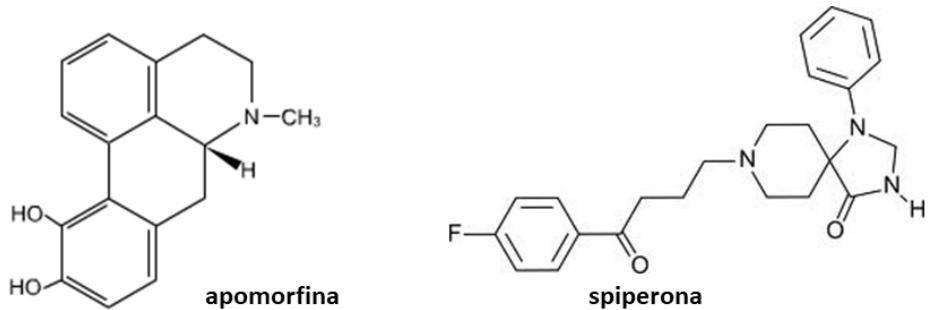


Figura 5.14: Estructura de los ligandos agonista (apomorfina) y agonista inverso (spiperona).

Los ligandos y los receptores fueron convertidos de pdb a formato pdbqt con MGLTools, los cálculos de docking fueron realizados utilizando AutoDock4 (Morris et al., 2009). De los diferentes complejos obtenidos del Docking de cada receptor, la conformación de menor energía de cada uno fue considerada como la orientación más favorable y fue usada para la simulación de Dinámica molecular del complejo.

Ambos complejos, del D2R acoplado a la proteína Gi y en su forma inactivada con sus respectivos ligandos, fueron insertados en una membrana biológica y se realizaron simulaciones de Dinámica Molecular siguiendo los pasos explicados previamente en el punto anterior.

5.2.3 Análisis topológico de la distribución de la densidad electrónica

Se construyeron sistemas reducidos de los complejos de menor energía de las trayectorias de DM, del D2R inactivado y activado, conservando en cada receptor solamente los residuos que interaccionan directamente con los ligandos. La densidad electrónica de estos sistemas se obtuvo a través de cálculos DFT con funcional híbrido PBE y conjunto de bases 6-31G(d), utilizando Gaussian 09 (M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ort, 2016). El análisis

topológico de la densidad de carga se realizó en base a la Teoría cuántica de Atomos (QTAIM) con ayuda del software Multiwfn.

5.2.4 Estructuras cristalinas y comparación con los modelos de homología.

Como se dijo previamente durante el desarrollo de esta tesis se cristalizaron las estructuras del receptor D2 tanto en su forma inactiva (PDB: 6CM4)(Wang et al., 2018) unida a las droga antipsicótica Risperidona como en la forma activa formando un complejo ternario con la proteína Gi y el compuesto agonista Bromocriptina (PDB: 7JVR)(Zhuang et al., 2021).

Esta situación proporcionó la posibilidad de evaluar la calidad de los modelos de homología obtenidos previamente, tanto de forma cuantitativa (ej. mediante superposición de los modelos de homología con las correspondientes estructuras resueltas y cálculo de la desviación cuadrática media de los $C\alpha$ (RMSD)) como de forma cualitativa (ej. mediante análisis comparativo de los “switches” moleculares presentes). Los resultados de este análisis comparativo se muestran en el ANEXO B

Asimismo, los estudios y análisis subsiguientes fueron realizados empleando las estructuras resueltas experimentalmente.

5.2.5 Docking de Dopamina y m-clorotiramina

Para lograr un mejor muestreo de las conformaciones de dopamina y m-clorotiramina en el sitio ortostérico D2R, se profundizo más allá de la simplificación del modelo de dos estados de activación del receptor y se realizó el docking sobre un conjunto de varias conformaciones de ambos estados del receptor.

Para generar las diferentes conformaciones del receptor se realizaron previamente simulaciones de dinámica molecular de 100 ns de las estructuras cristalinas de ambas formas insertadas en una bicapa fosfolipídica, siguiendo un procedimiento similar al descrito para los modelos de homología.

Luego se realizó el docking de ambos ligandos sobre las conformaciones pre-generadas del receptor. Para el análisis de las poses de docking se calcularon las “fingerprint” basadas en interacciones estructurales (SIFT), siguiendo el protocolo descrito por Deng et al. (Deng et al., 2004). Brevemente, para cada pose se tomaron 23 residuos correspondientes al bolsillo de

unión del receptor junto con el ligando y se calcularon 7 tipos de interacciones intermoleculares para cada residuo: (i) Interacciones débiles (del tipo dispersión de London); (ii) Nube π (cara a cara); (iii) Nube π (borde a cara); (iv) Puente de hidrógeno (proteína dador/ ligando aceptor); (v) Puente de hidrógeno (proteína aceptor/ ligando dador); (vi) Puente salino (proteína +/ lignado -); y (vii) Puente salino (proteína -/ ligando +). Dichas interacciones fueron codificadas en una “fingerprint” de 7 bits binarios para cada residuo (7×23 bits) donde cada bit está “encendido” (ej. toma el valor 1) cuando está presente la interacción o “apagado” (ej. toma el valor 0) en otro caso. Mediante la comparación de las “fingerprint”s calculadas para cada pose, empleando el coeficiente de Tanimoto, es posible agruparlas por similaridad en base a las interacciones intermoleculares que se establecen en cada caso.

Empleando este protocolo se seleccionaron poses representativas de los ligandos en ambos receptores y se calcularon las energías libres de unión relativas $\Delta(\Delta G)$ siguiendo protocolo MM-GBSA incluido en AMBER.

5.2.6 Aplicación del protocolo para predecir EXs basado en las interacciones de los entornos proteicos

Se seleccionaron dos compuestos clorados, un agonista, m-clorotiramina, y un agonista inverso, clorpromazina. Las estructuras 3D de estos compuestos se obtuvieron mediante la utilización de herramientas quimioinformáticas incluidas en OpenBabel(O’Boyle et al., 2011).

Luego se realizó Docking Molecular parametrizado para EXs del compuesto agonista con el receptor en la forma activada y del agonista inverso con la estructura del receptor en la forma inactiva, finalmente se aplicó el protocolo descripto en el capítulo 4 (xenv) para la predicción de la pose con el entorno más propicio para la formación de un EXs, además simultáneamente se calcularon con la metodología QTAIM las propiedades derivadas de la densidad de carga de los enlaces H…Cl que conforman los entornos de enlace X para poder comparar los resultados.

5.3 Resultados

5.3.1 Análisis de los modelos de homología

La Figura 5.15 muestra la superposición de los modelos de homología de D2R activado (D2Ra) e inactivado (D2Ri) luego del refinamiento de los modelos mediante simulaciones guiadas por ligando. Al pasar de la forma inactivada a la activada de D2R se observa un marcado

desplazamiento del segmento intracelular de TM6 hacia fuera del ovillo de segmentos transmembrana lo que permite el acoplamiento con la proteína Gi.

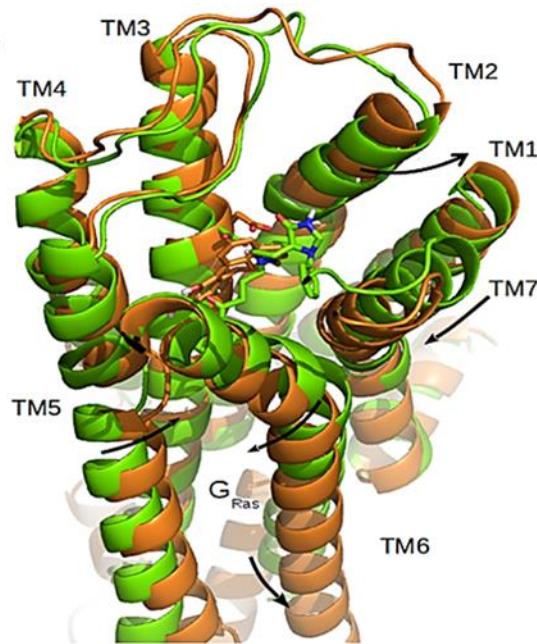


Figura 5.15: Superposición de las estructuras de las formas activa e inactiva del DR2 formando complejo con el ligando.

La Figura 5.16 muestra con más detalle las interacciones de los ligandos en el sitio ortostérico. Tanto en el caso de la apomorfina con el receptor activado, como en la spiperona con el receptor inactivado, las orientaciones de los ligandos coinciden con los reportados previamente (Lan et al., 2006). Ambos se anclan al receptor mediante el característico puente salino entre el grupo amino protonado del ligando y D3.32 (Figura 5.16). Esto se evidencia con los caminos de enlace que conectan el protón del grupo amonio del ligando con uno de los átomos de oxígeno del carboxilato de la cadena lateral de D3.32. Estudios previos de mutagénesis han revelado que esta interacción es clave para el evento de reconocimiento molecular (Mansour et al., 1992). Otra interacción clave establecida por la apomorfina a través de sus grupos OH catecólicos son los enlaces de hidrógeno con S5.46 y S5.42. La capacidad de la apomorfina para formar interacciones con el grupo de residuos de serina de TM5 es una característica común de los agonistas de la dopamina; es bien conocido que estas interacciones están involucradas en la activación de D2R.

Por el contrario, el agonista inverso no forma interacciones con las serinas de TM5; el anillo fluorobencénico de la spiperona se aloja más abajo en un bolsillo hidrofóbico formado por W6.48, F6.44, F6.52, F5.47 e I3.40.

La formación del bolsillo hidrofóbico es impulsada por el agonista inverso, por desplazamiento de la cadena lateral de W6.48 (de lo contrario este chocaría con el anillo fluorobencénico de la spiperona) que es acompañado por el desplazamiento del segmento TM6 en su parte central hacia TM7. Dicho desplazamiento se ve facilitado por la presencia de un residuo de prolina P6.50 altamente conservado, el cual, al carecer de un grupo amino libre no puede participar en la red de enlaces de hidrógeno del backbone de las α hélices y por lo tanto constituye un punto de fácil perturbación de la hélice.

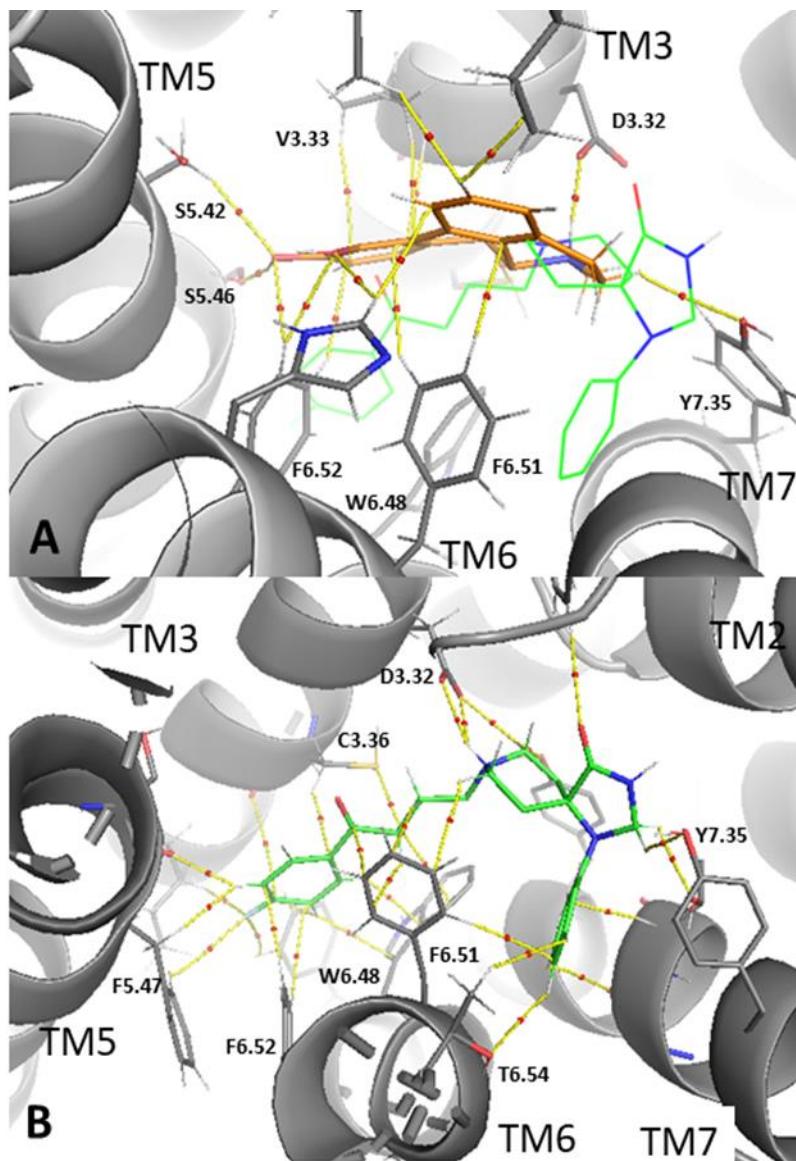


Figura 5.16: Bolsillo ortostérico de D2R en sus estados activo (A) e inactivo (B) unido al agonista apomorfina (naranja) y al agonista inverso spiperona (verde), respectivamente. También se representan los elementos topológicos de la densidad de carga para las interacciones ligando/receptor. Los caminos de enlace (BP) que conectan los átomos del ligando con los átomos del receptor se representan con líneas amarillas y los puntos críticos de enlace (BCP) se representan con pequeñas esferas rojas sobre los BP. En A, se superpuso la estructura del agonista inverso (verde) tal como está unido a D2Ri para hacer evidentes las diferencias en el modo de unión de los distintos ligandos. Se utilizó la convención de Ballesteros-Weinstein para la numeración de los residuos de GPCR (Ballesteros et. al., 1995)

Por otra parte, mediante el análisis de densidad electrónica se observó que los valores de densidad electrónica para cada residuo son mucho mayores en el D2Ra, principalmente D3.32 (ANEXO B)

5.3.2 Descripción de los “switches” moleculares

Como se describió previamente la activación de los receptores acoplados a la proteína G ocurre a través de una serie de cambios conformacionales denominados *switches* moleculares, los cuales se pueden observar comparando la estructura del mismo receptor unida a un agonista o un agonista inverso. De este modo a través del análisis de las estructuras de menor energía de cada dinámica se pudieron observar dichos cambios en las dos formas del receptor (Trzaskowski et al., 2012).

La superposición de los modelos activado e inactivado en la Figura 5.17 permite visualizar los “switches” moleculares, particularmente en el lado extracelular del receptor.

Tras la activación de D2R, el residuo W6.48 (y con él la región central de TM6) se desplaza para llenar el bolsillo hidrofóbico que es ocupado usualmente por los agonistas inversos como la espiperona que se unen a la forma inactiva del receptor. Estos cambios alrededor de W6.48 forman parte del interruptor de transmisión observado previamente en rodopsina, entre otros GPCRs. Similarmente a D2R, en la rodopsina, el anillo de beta-ionona del 11-cis-retinal es el que ocupa el bolsillo hidrofóbico que luego será ocupado por W6.48 tras la isomerización de retinal.

Otros de los residuos de TM6 que muestran un marcado movimiento al comparar las formas activada e inactivada de D2R son I3.40 y F6.44. Al pasar de la forma inactivada a la activa se produce un desplazamiento de F6.44 hacia TM5 lo que posibilita la formación de interacciones de tipo stacking con F5.51 y F5.47 (Figura 5.17B). Estas interacciones podrían explicar en parte

el acercamiento de TM5 hacia el centro del ovillo de TMs y la menor separación TM5-TM6 que se observa en la parte extracelular de D2R tras la activación del receptor (Figura 5.17A). El desplazamiento de F6.44 también es acompañado por la translocación de la cadena lateral de I3.40, de lo contrario se produciría un choque estérico entre ambos residuos. En la forma inactivada la cadena lateral de Ile3.40 se encuentra en proximidad a P5.50 bloqueando el movimiento de TM5 hacia el interior del ovillo de TMs. Tras la activación, la cadena lateral de I3.40 se aleja de P5.50 y se desbloquea el desplazamiento de TM5 hacia el centro de los TMs en la parte extracelular de D2R (Figura 5.17B). Estas translocaciones de los residuos de F6.44 e I3.40 en D2R forman parte del interruptor de transmisión y son notablemente similares a las descriptas en rodopsina y otros GPCRs.

Finalmente, las interacciones del agonista con las serinas de TM5 también facilitarían (junto con las translocaciones mencionadas) el acercamiento de este segmento transmembrana hacia el interior del ovillo de TMs que se observa tras la activación de D2R.

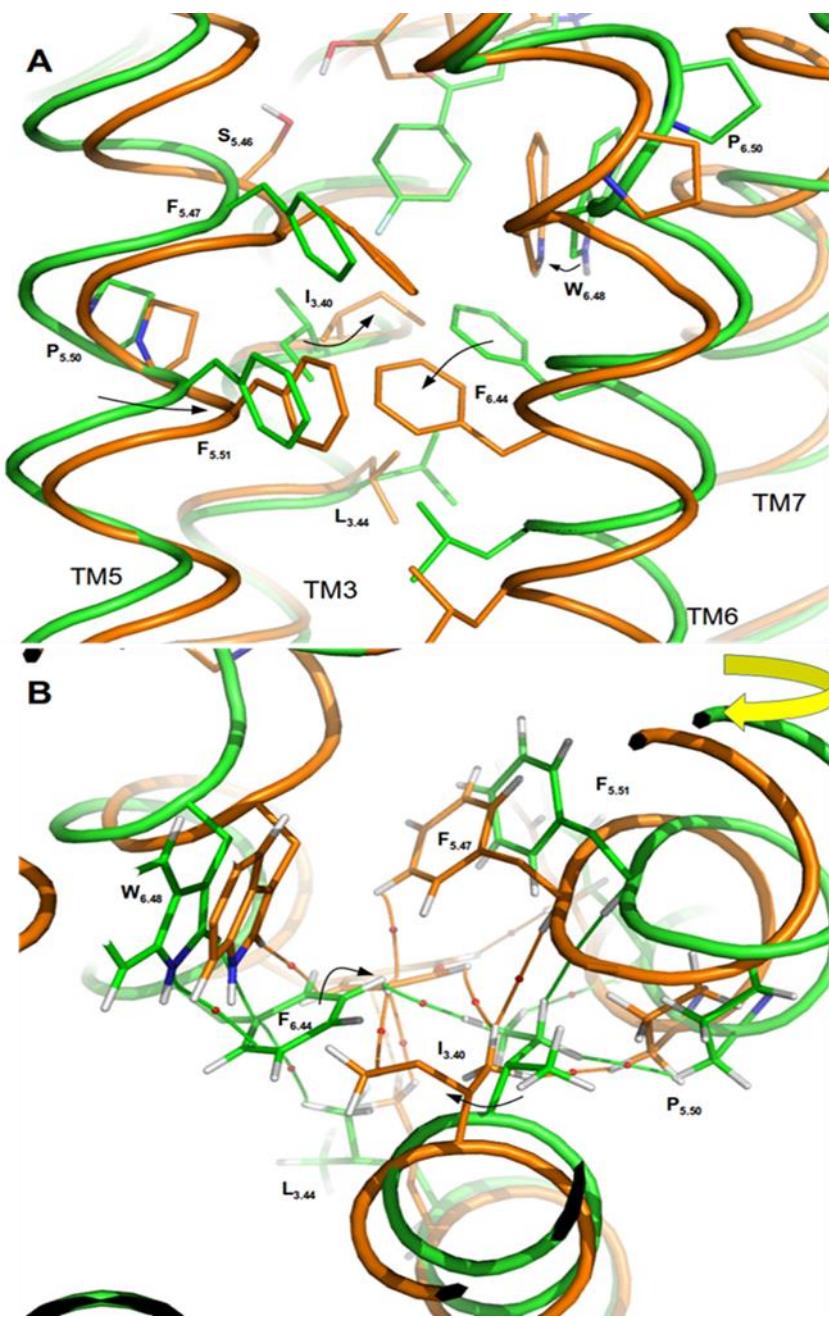


Figura 5.17. Superposición estructural de los estados D2Ra-Gi y D2Ri unidos al agonista apomorfina (naranja) y al agonista inverso spiperona (verde), respectivamente. Se muestran los residuos implicados en el cambio de transmisión en el lado extracelular de las hélices transmembrana. Las flechas indican los desplazamientos de átomos/hélices provocados por la activación del receptor. En A se muestra una vista lateral de los elementos estructurales que intervienen en el interruptor de transmisión. En B, se representa una vista superior/girada de los mismos residuos junto con los elementos topológicos de la densidad de carga correspondientes a las interacciones de los residuos clave F6.44 e I3.40.

La Figura 5.18 muestra la superposición de las formas activada e inactivada de D2R en el lado citoplasmático del receptor. La diferencia más sobresaliente entre ambas formas del receptor en el lado citoplasmático es el desplazamiento hacia afuera del TM6 para dar lugar a la subunidad alfa de la proteína G en el estado activado. Este desplazamiento de TM6 implica la ruptura de interacciones de este segmento con residuos de TM3 que mantienen ambos segmentos juntos en el estado inactivado del receptor. En D2R se produce la ruptura del puente entre los residuos R3.50 y T6.34, dicha interacción forma parte del interruptor de bloqueo iónico previamente descripto en rodopsina y otros GPCRs.

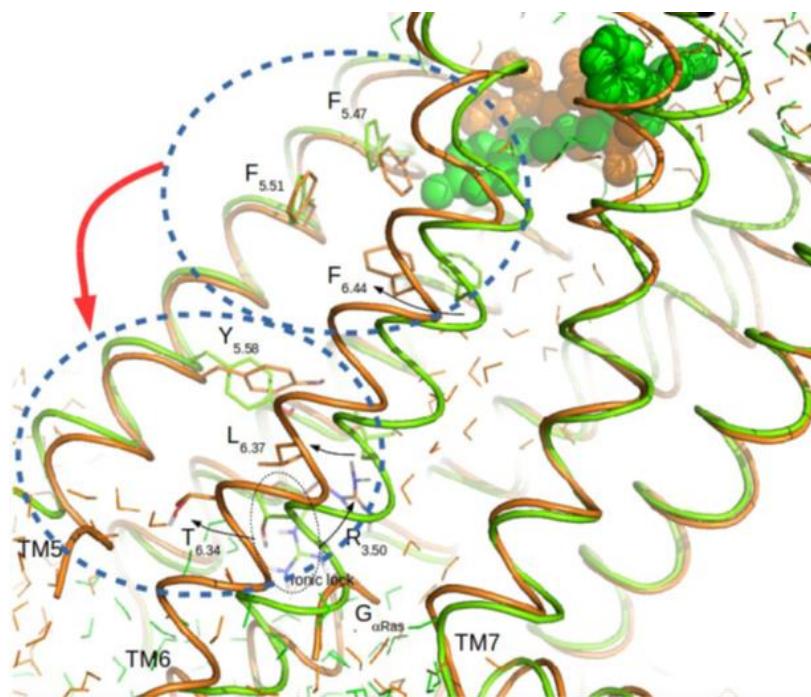


Figura 5.18: Superposición del lado citoplasmático de D2R en sus formas inactivada (verde) y activada acoplada a proteína G (naranja). Se indican con flechas los movimientos de residuos al pasar de la forma inactivada a la activada del receptor. Con líneas verdes y anaranjadas se muestran las moléculas de agua de la forma inactivada y activada, respectivamente. Nótese que solo en la forma activada las moléculas de agua pueden atravesar el canal central del ovillo de TMs desde el lado extracelular hasta el citoplasmático.

El desplazamiento hacia afuera de TM6 en su extremo citoplasmático al activarse el receptor hace que éste segmento se separe de TM3 y TM7 y al mismo tiempo se acerque a TM5 con lo cual se refuerza la interacción TM5-TM6. Dado que el efecto de la unión del agonista en el lado extracelular del receptor es el acercamiento y aumento de la interacción entre TM5 y TM6, es lógico pensar que estos efectos se puedan transmitir hacia el lado citoplasmático del receptor.

Es decir que el desplazamiento/rotación de TM6 hacia afuera del ovillo de TMs en el lado intracelular podría ser (al menos en parte) una consecuencia del refuerzo de la interacción TM5-TM6 impulsada por el agonista en el lado extracelular del receptor, proporcionando así un mecanismo de transmisión de la señal desde el lado extracelular al lado intracelular del receptor.

La Figura 5.19 representa los elementos topológicos de densidad electrónica asociados con las interacciones no covalentes en la interfaz TM5-TM6 para ambas formas de receptor. Queda claro a partir de esta figura que la asociación entre ambos segmentos TM se fortalece con la activación del receptor en los extremos extracelular e intracelular del receptor. Mientras que en la región media de la interfaz TM5-TM6, el cambio más notable en la red de enlaces H del residuo Y5.58 que, a medida que se acerca a TM5, se apila encima de L6.37, como lo demuestran los caminos de enlace que conectan ambos residuos.

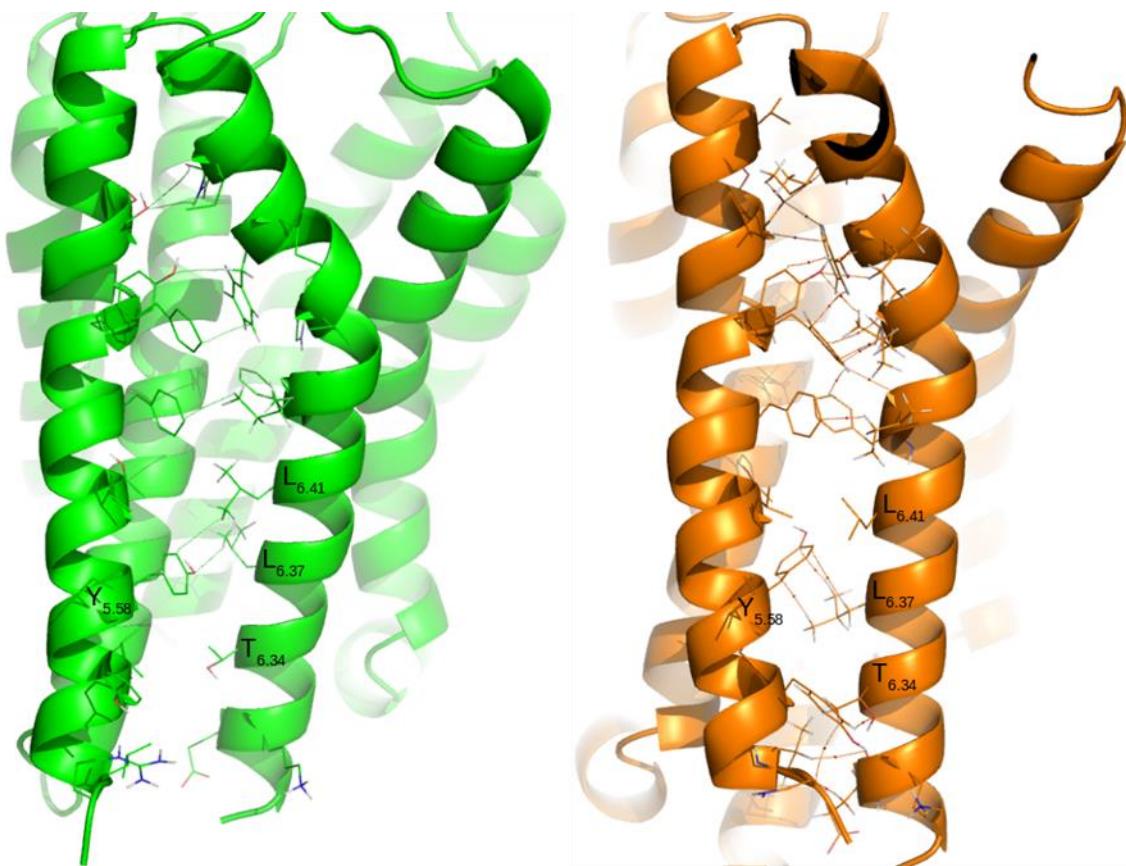


Figura 5.19. Interacciones entre residuos en la interfaz TM5-TM6 para formas inactivas (verde) y activas (naranja) de D2R. Las interacciones se evidencian representando los elementos topológicos de la densidad de carga: puntos críticos de enlace (BCP) y los caminos de enlace (BP) correspondientes que conectan los núcleos enlazados.

Por otra parte en la bibliografía se encuentra que las moléculas de agua también han sido señaladas como involucradas en la activación de los GPCRs (Trzaskowski et al., 2012). Nótese en la Figura 5.18 que solo en la forma activada las moléculas de agua pueden atravesar el canal central del ovillo de TMs desde el lado extracelular hasta el citoplasmático (moléculas de agua en anaranjado). Así, podría pensarse que la ruptura del interruptor de bloqueo iónico se vería también facilitada por el ingreso de moléculas de agua en el canal central desde el lado extracelular.

5.3.3 Activación de D₂R por Dopamina

Más allá de la descripción de los “switches” moleculares, una cuestión aún no resuelta es entender la serie de eventos moleculares que desencadenarían el *encendido* de los “switches” y en última instancia la activación del D₂R. Desde esa perspectiva, resulta interesante indagar cómo la unión del ligando endógeno, la dopamina, estaría involucrada en dichos eventos de activación. De acuerdo con un modelo de activación por ajuste inducido, la dopamina al unirse a la forma inactiva del receptor induciría de alguna manera al encendido de los interruptores.

Por otra parte, de acuerdo a la teoría de selección conformacional, los agonistas deberían estabilizar primordialmente la forma activada del receptor mientras que los agonistas inversos deberían unirse con mayor afinidad a la forma inactivada.

La disponibilidad de estructuras de D₂R en ambas formas, activada e inactivada, brinda la posibilidad de analizar estos aspectos relativos a la activación del receptor por dopamina.

Siguiendo el protocolo descrito en métodos, se obtuvieron las poses de docking de dopamina en diferentes conformaciones de ambas formas del receptor.

Las huellas dactilares (“fingerprint”s) de interacción estructurales (SIFT) se calcularon para las poses de docking y se realizó un clusterizado jerárquico aglomerativo de las mismas (dendograma). Para obtener los clusters se realizó corte del dendograma a una altura $h=5$, revelando dos modos principales de unión C1 y C2 (Figura 5.20). Los dos grupos muestran la interacción característica del puente salino con D3.32, pero solo en C1 la dopamina está bien posicionada dentro del sitio ortostérico para interactuar con los residuos de TM5 y TM6. Luego,

se tomó el grupo C1 y se fragmentó en grupos más pequeños con un agrupamiento con altura $h=1$, obteniendo 10 grupos diferentes.

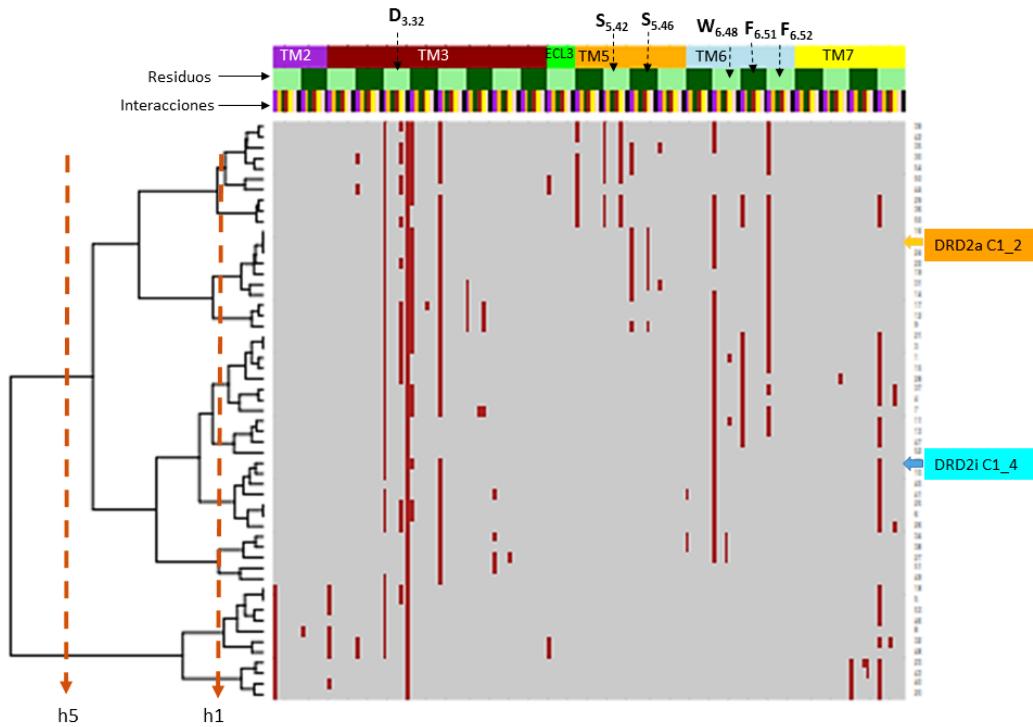


Figura 5.20: Agrupación jerárquica de los “fingerprint” de interacción estructural (SIFT) de las poses de acoplamiento de dopamina en conformaciones D2R activas e inactivas. Las filas en el mapa binario representan los SIFts (es decir, una pose de acoplamiento). Los bits (interacciones) que están “encendidos” se muestran en rojo, en las columnas del mapa binario. El dendrograma de la izquierda representa los resultados del agrupamiento jerárquico en las huellas dactilares. Los SIFts (filas) en el mapa binario se reorganizan de acuerdo con el orden dado por la agrupación jerárquica. El agrupamiento se realizó realizando el corte del árbol dendrograma a dos alturas diferentes, $h = 5$ y $h = 1$. Marcadas a la derecha en color naranja y celeste, las poses seleccionadas con el D2Ra y D2Ri respectivamente.

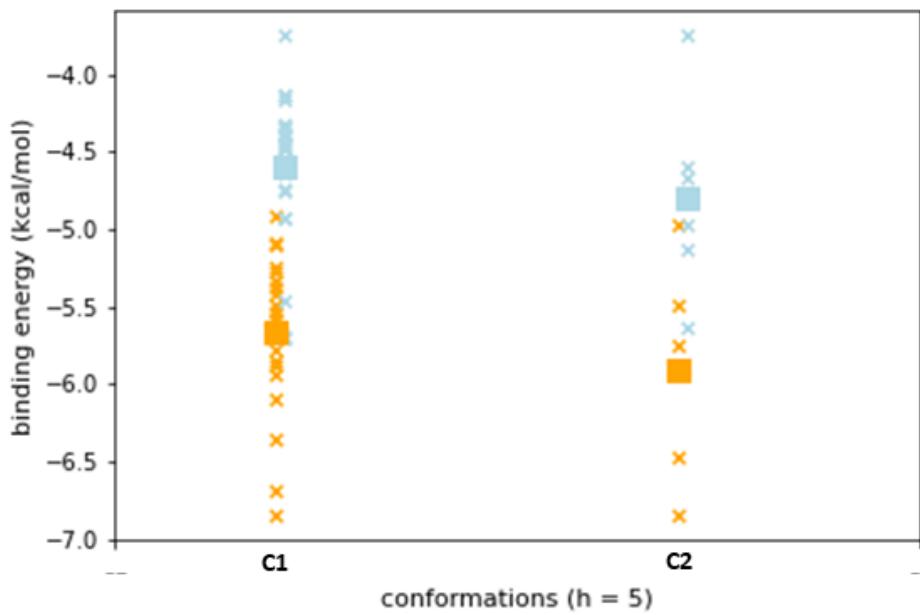


Figura 5.21: Grafico de energía de las diferentes poses de acuerdo al agrupamiento que se realizó a la altura $h=5$. Los grupos naranjas corresponden a la forma activa del receptor y los grupos celestes a la forma inactiva.

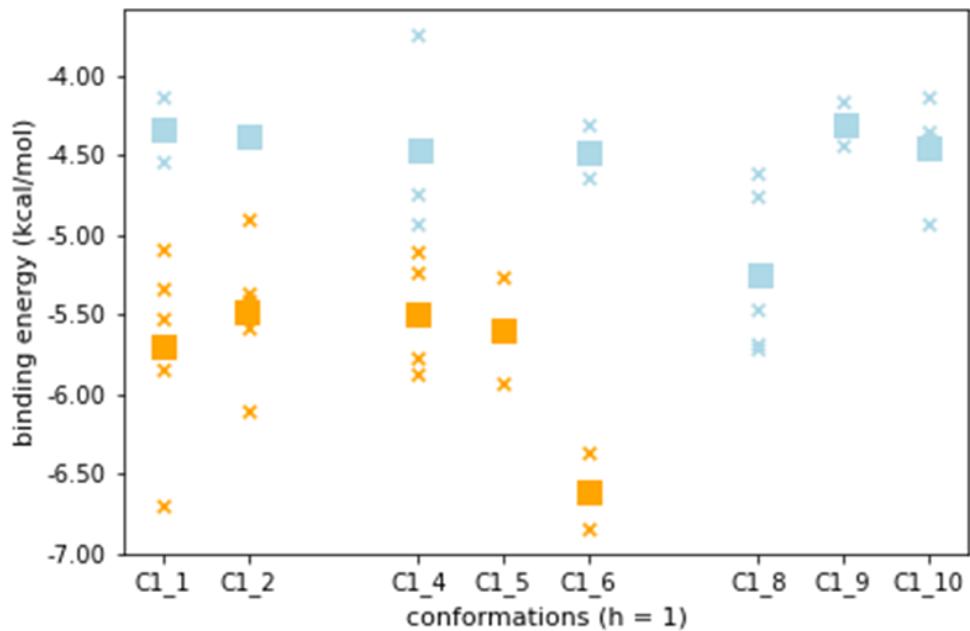


Figura 5.22: Grafico de energía de las diferentes poses de acuerdo al agrupamiento que se realizó a la altura, $h = 1$. Los grupos naranjas corresponden a la forma activa del receptor y los grupos celestes a la forma inactiva.

La comparación de la estabilidad general de los diferentes modos de unión de dopamina (Figuras 5.21 y 5.22) reveló que las poses de docking en las conformaciones de receptores activos son, en promedio, más estables que las de las conformaciones de receptores inactivos.

Para corroborar estos datos, se seleccionaron dos poses, una con el receptor en la forma inactiva (C1_4, Figura 5.20), en la cual el anillo de catecol de dopamina se une al bolsillo hidrofóbico en la parte inferior del sitio ortostérico del receptor, adoptando así un patrón de unión similar al de los agonistas inversos, y otra del receptor en la forma activa en la cual la dopamina adopta un modo de unión similar al agonista e interactúa con los residuos de serina relevantes de TM5, como se esperaba (C1_2, Figura 5.20). Se realizaron simulaciones de DM de 100ns.

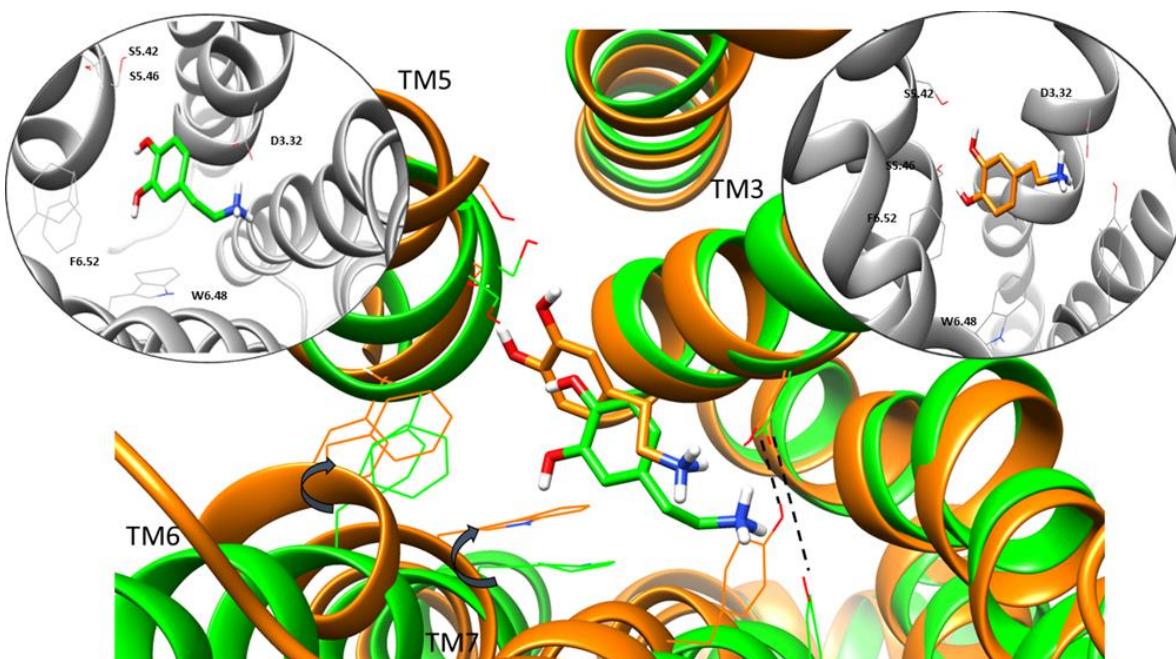


Figura 5.23: Poses de docking seleccionadas a través del análisis de SiFt. En color naranja la pose con el receptor en la forma activa (grupo C1_2) y en verde la pose seleccionada de la dopamina con el receptor en la forma inactiva (grupo C1_4).

Al analizar las poses seleccionadas (Figura 5.23), se puede observar los cambios en el sitio activo del receptor en sus diferentes formas, lo que lleva a que la dopamina, aunque sea un ligando pequeño, adopte diferentes poses dependiendo de la conformación del receptor. En otras palabras, el modo de unión de dopamina viene determinada por la distribución espacial de los residuos proteicos, particularmente de aquellos involucrados en los “switches” moleculares. Por ejemplo, el interruptor de tirosina, el cual se encuentra “apagado” en este caso

en la forma inactiva permite que dopamina adopte un modo de unión en el cual interacciona con los residuos de TM6 y TM7, pero se encuentra alejado para formar puentes de hidrógeno con el cluster de serinas de TM5. En cambio, en la forma activada, la formación de la interacción entre Y7.43 y D3.32 disminuye las posibilidades conformacionales de dopamina dentro del sitio ortostérico, con lo cual ésta termina necesariamente interaccionando con los residuos de TM5.

Asimismo, el switch de transmisión que involucra la translocación del W6.48, también determina la conformación que adopta la dopamina, en la forma activada este residuo se encuentra desplazado hacia el centro del ovillo de hélices transmembrana, impulsando la translocación de dopamina hacia TM5.

De las simulaciones de DM de cada complejo se realizó el cálculo de energía libre relativa de unión ($\Delta\Delta G$) calculada a partir de las trayectorias de DM empleando el protocolo MM-GBSA (Tabla 4.1), comprobando que las simulaciones con el receptor en la forma activa son más estables que con el receptor en la forma inactiva, dato compatible con los de valores de energía obtenidos del análisis de Docking Molecular y SiFt. Estos resultados son congruentes con el hecho de que la dopamina es un compuesto agonista por lo cual se une de una forma más estable al receptor en la forma activa, lo cual esta de acuerdo con el modelo de selección conformacional.

Tabla 5.1. Energías libres de unión relativas ($\Delta\Delta G$) de los complejos ligando-receptor

Complejo L-R	$\Delta\Delta G$ (kcal/mol)
D2Ra C1_2	-19,12
D2Ri C1_4	-11,84

Por otra parte, si consideramos un modelo de ajuste inducido, la dopamina debería unirse primero a la forma inactiva e inducir el encendido de los “switches” moleculares. Nótese que en la forma inactiva el grupo amonio de la dopamina se ubica entre los residuos del interruptor de tirosina Y7.43 y D3.32, formando enlaces de hidrógeno simultáneos con ambos residuos. Así, uno podría especular que la unión de dopamina contribuye de alguna manera a acercar ambos residuos, facilitando el encendido del interruptor de tirosina. En efecto, los enlaces de hidrógeno simultáneos, en sus diferentes configuraciones, han sido previamente implicados en este tipo de transiciones conformacionales de proteínas (Jeffrey & Saenger, n.d.).

5.3.4 Predicción del perfil funcional de la m-Clorotiramina

Con el fin de estudiar el posible rol del enlace de halógeno en la estabilización de complejos ligando-D2R, se realizó docking de la m-clorotiramina (m-Cl), donde el OH en meta- de dopamina se reemplazó por un átomo de Cloro. En estudios previos este compuesto mostró una afinidad significativa por el D2R indicando un valor K_i de 0,146 mM (Angelina, 2014), muy similar a la reportada para dopamina ($K_i=0.52$ mM) (Kozell & Neve, 1997). Así también, los estudios de modelado molecular previos han demostrado un modo de unión similar al de la dopamina, en el cual el átomo de cloro “mimetiza” las interacciones que establece el OH en meta- de la dopamina, debido a la capacidad dual del halógeno de actuar como dador y aceptor (de halógeno), en analogía con la posibilidad del OH de actuar como dador y aceptor de hidrógeno.

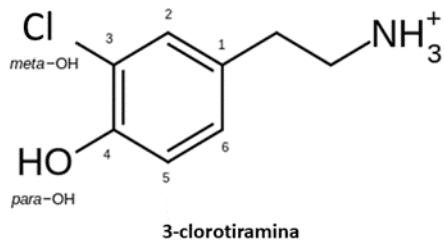


Figura 5.24: Estructura de 3-clorotiramina

Se aplicó el mismo protocolo de docking sobre ensembles de conformaciones de la DM empleado para dopamina. Para realizar el clusterizado basado en “fingerprint”s de interacciones se siguieron los pasos detallados en la sección de métodos, las “fingerprint”s de interacción estructural (SIFT) se calcularon para las poses de docking y se agruparon en consecuencia, con altura $h=1$, revelando dos modos principales de unión C1 y C2 (Figura 5.25). Los dos grupos muestran la interacción característica del puente salino con D3.32, pero en C1 se ubica en la zona del TM5 dentro del sitio ortostérico donde podría interaccionar con las serinas como un compuesto agonista.

La comparación de la estabilidad general de los diferentes modos de unión de clorotiramina (Figura 5.26) reveló que las poses del docking en las conformaciones de receptores activos son en general, más estables que las de las conformaciones de receptores inactivos como sería de

esperar para un compuesto agonista. No obstante, estas predicciones deberían confirmarse con los correspondientes ensayos funcionales.

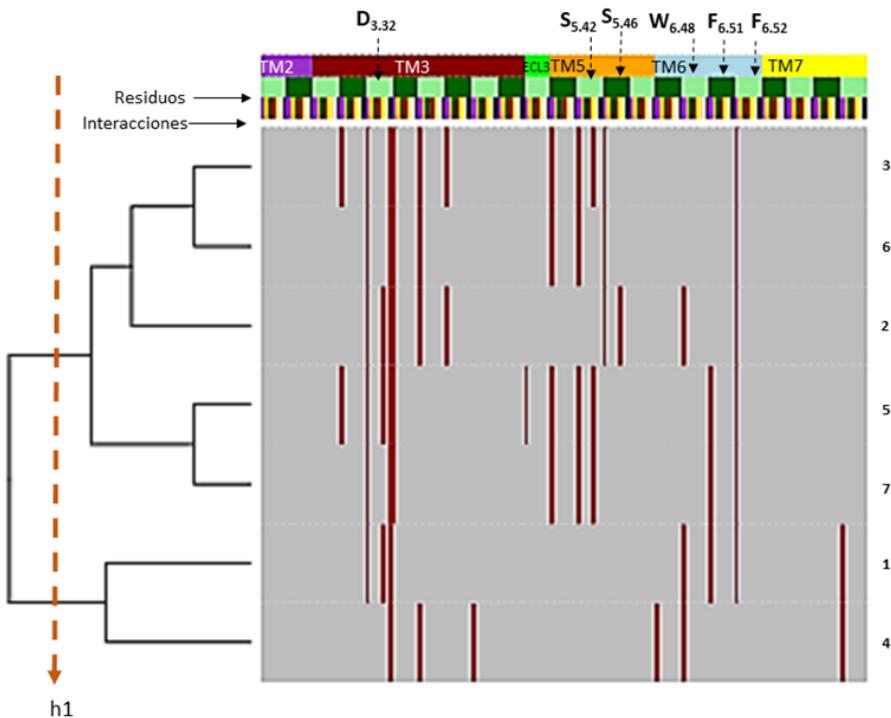


Figura 5.25: Agrupación jerárquica de los “fingerprint” de interacción estructural (SIFT) de las poses de acoplamiento de 3-clorotiramina en conformaciones D2R activas e inactivas. Las filas en el mapa binario representan los SIFts (es decir, una pose de acoplamiento). Los bits (interacciones) que están *encendidos* se muestran en rojo, en las columnas del mapa binario. El dendograma de la izquierda representa los resultados del agrupamiento jerárquico en las huellas dactilares. Los SIFts (filas) en el mapa binario se reorganizan de acuerdo con el orden dado por la agrupación jerárquica. El agrupamiento se realizó realizando el corte del árbol dendrograma a altura $h = 1$.

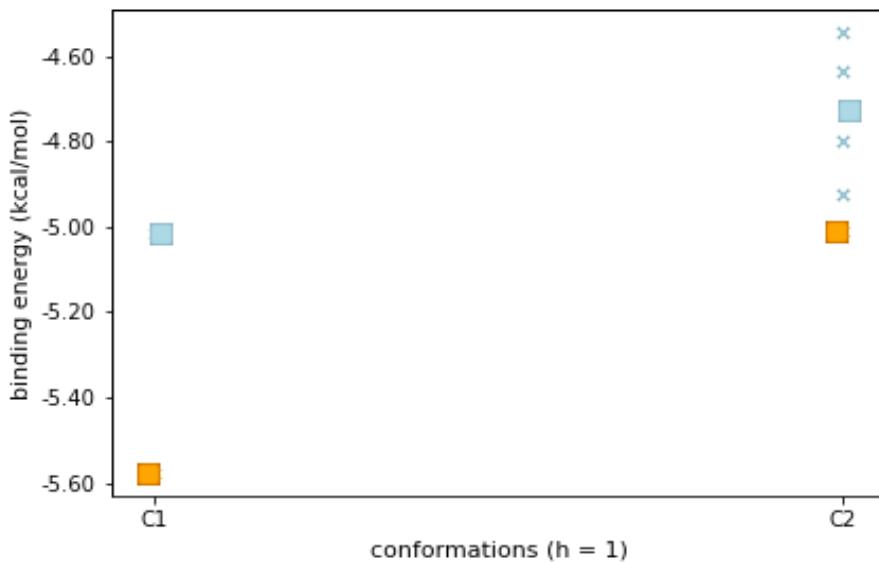


Figura 5.26: Gráfico de energía de las diferentes poses de acuerdo al agrupamiento que se realizó a la altura, $h = 1$. Los grupos naranjas corresponden a la forma activa del receptor y los grupos celestes a la forma inactiva.

5.3.5 Predicción de EXs basado en entornos proteicos

En la sección previa se analizó la interacción de 3-clorotiramina con el receptor y se identificó un grupo de poses en las cuales ésta forma interacciones con los residuos de serina TM5. Asumiendo que la 3-clorotiramina actuaría como una agonista del D2R, éste sería su modo de unión más probable dentro del sitio ortostérico del receptor.

En esta sección queremos indagar más en profundidad cuáles serían las interacciones específicas que establecería la 3-clorotiramina con las serinas de TM5. En estudio previos de ligandos dopaminérgicos halogenados, el enlace de halógeno más frecuentemente observado durante las simulaciones era el que se formaba con el oxígeno carbonílico del backbone de Ser5.42 (O@Ser5.42) como aceptor. No obstante, enlaces de halógeno con el oxígeno hidroxílico de Ser5.42 (OH@Ser5.42) y con el oxígeno del backbone de Ser5.46 (O@Ser5.46) también se formaban ocasionalmente (Luchi et al., 2016).

Para conocer cuál sería la orientación más probable del halógeno de la 3-clorotiramina dentro del sitio ortostérico de D2R del receptor activado, se aplicó el protocolo de detección de enlace de halógeno basado en el entorno implementado en los capítulos previos.

Se consideraron los 3 aceptores de halógeno (OA) observados en las simulaciones previas de LDX, O@Ser5.42, OH@Ser5.42 y O@Ser5.46; se realizó docking molecular parametrizado para EXs y calculó la fortaleza del entorno del átomo de halógeno en cada caso como lo indica el protocolo del capítulo 4. Para confirmar las predicciones se realizó el cálculo de la distribución de la densidad electrónica y se obtuvieron los grafos moleculares (red de caminos de enlaces) de las poses seleccionadas.

Las figuras 5.27, 5.28 y 5.29 muestran los grafos moleculares de densidad de carga de los complejos 3-clorotiramina/D2Ra enlazados por EX. Para cada enlace de hidrógeno del entorno del halógeno H···Cl, se muestra el valor verdadero de la relación λ_1/λ_3 . De los tres entornos posibles se obtuvo mejores resultados con el oxígeno carbonílico de S5.42 ya que el protocolo predijo correctamente las interacciones H···Cl, además dio un valor de $\sum \lambda_1/\lambda_3$ mayor a -0,60 au. lo que indica un entorno de EX de fuerza elevada, a diferencia de los otros dos casos donde el protocolo predijo solamente la formación de una interacción (Figura 5.28) y los valores de $\sum \lambda_1/\lambda_3$ son más bajos, indicando un entorno de EX más débil.

Mediante este análisis pudimos aplicar el protocolo para obtener la forma más probable de unión de la 3-clorotiramina en el sitio activo del receptor, basándonos en el entorno y la fuerza del mismo para formar EX.

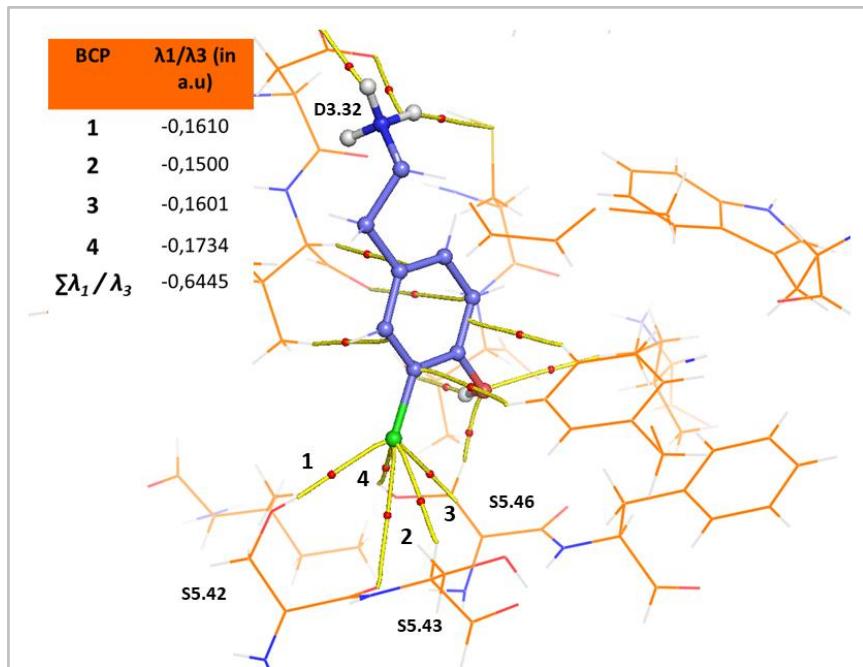


Figura 5.27. Grafos de densidad electrónica del complejo clorotiramina/D2Ra con el oxígeno carbonílico de la Ser5.42 como OA. Los caminos de enlace (BP) que conectan el cloro con el átomo de O (enlace X) y los átomos de H

(enlace H) se representan amarillo. Los círculos rojos en la parte superior de los BP son los puntos críticos de enlace (BCP). Los BCP para los enlaces de hidrógeno H···Cl están numerados y sus propiedades correspondientes se enumeran en la tabla insertada (valor de la relación λ_1 / λ_3).

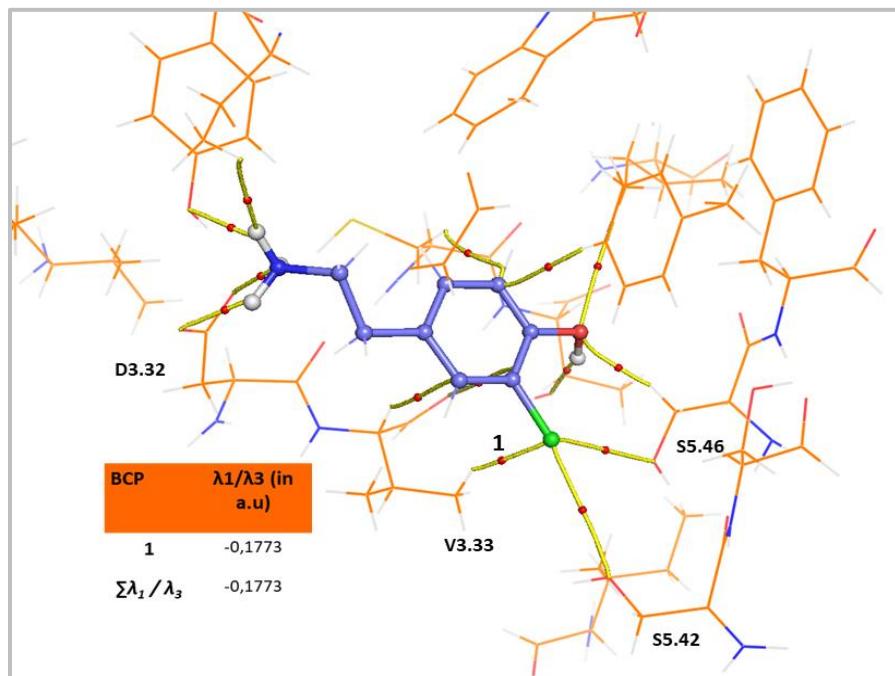


Figura 5.28. Grafo de densidad electrónica del complejo clorotiramina/D2Ra con el oxígeno del hidroxilo de la Ser5.42 como OA. Los caminos de enlace (BP) que conectan el cloro con el átomo de O (enlace X) y los átomos de H (enlace H) se representan en amarillo. Los círculos rojos en la parte superior de los BP son los puntos críticos de enlace (BCP). Los BCP para los enlaces de hidrógeno H···Cl están numerados y sus propiedades correspondientes se enumeran en la tabla insertada (valor verdadero de la relación λ_1 / λ_3).

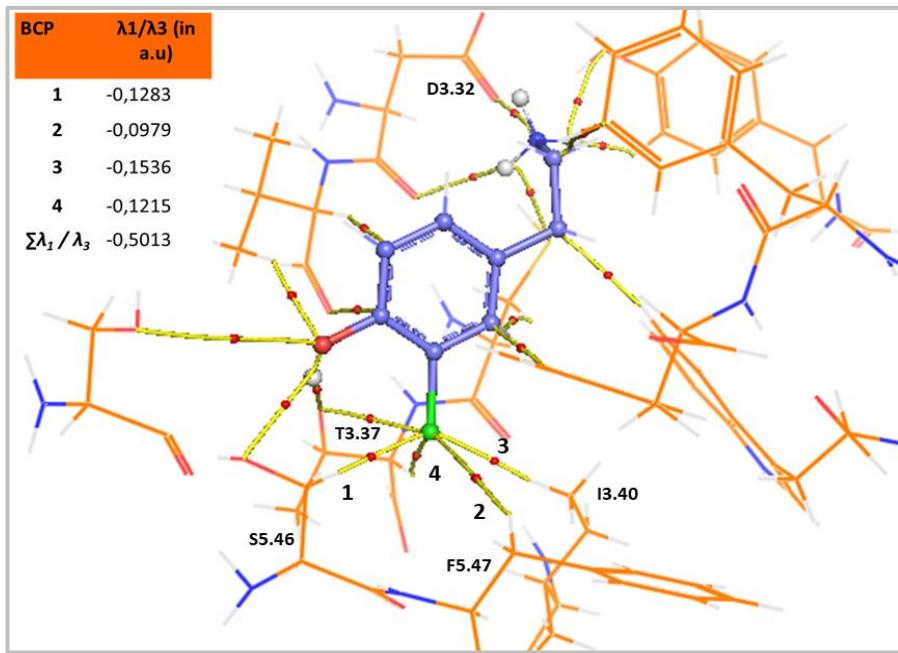


Figura 5.29. Grafo de densidad electrónica del complejo clorotiramina/D2Ra con el oxígeno carbonílico de la Ser5.46 como OA. Los caminos de enlace (BP) que conectan el cloro con el átomo de O (enlace X) y los átomos de H (enlace H) se representan en amarillo. Los círculos rojos en la parte superior de los BP son los puntos críticos de enlace (BCP). Los BCP para los enlaces de hidrógeno H···Cl están numerados y sus propiedades correspondientes se enumeran en la tabla insertada (valor verdadero de la relación λ_1/λ_3).

Con el fin de presentar un caso de estudio con el receptor en la forma inactiva, se utilizó el agonista inverso halogenado, clorpromazina (Figura 5.30). El protocolo predijo correctamente la formación del EX y de dos interacciones H···Cl que están realmente presentes en el grafo molecular calculado y con valores estimados de la relación λ_1/λ_3 que están muy cerca de sus valores reales. Además, de acuerdo al histograma de la Figura del capítulo 4, este complejo estudiado tiene un entorno de enlace X de fuerza intermedia, con valores de $\sum \lambda_1 / \lambda_3$ que entre -0,28 a -0,60 au.

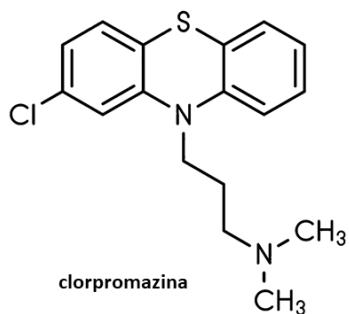


Figura 5.30: Estructura del agonista inverso clorpromazina

La Figura 5.31 clorpromazina/D2Ri respectivamente con EX junto con la probabilidad estimada de formación de EH (BCP) (prob EH), así como los valores verdadero y predicho de la relación λ_1/λ_3 para cada H···Cl. BCP. En la pose seleccionada por el protocolo, el Cloro se encuentra formando un EX con el Oxígeno del residuo D3.32.

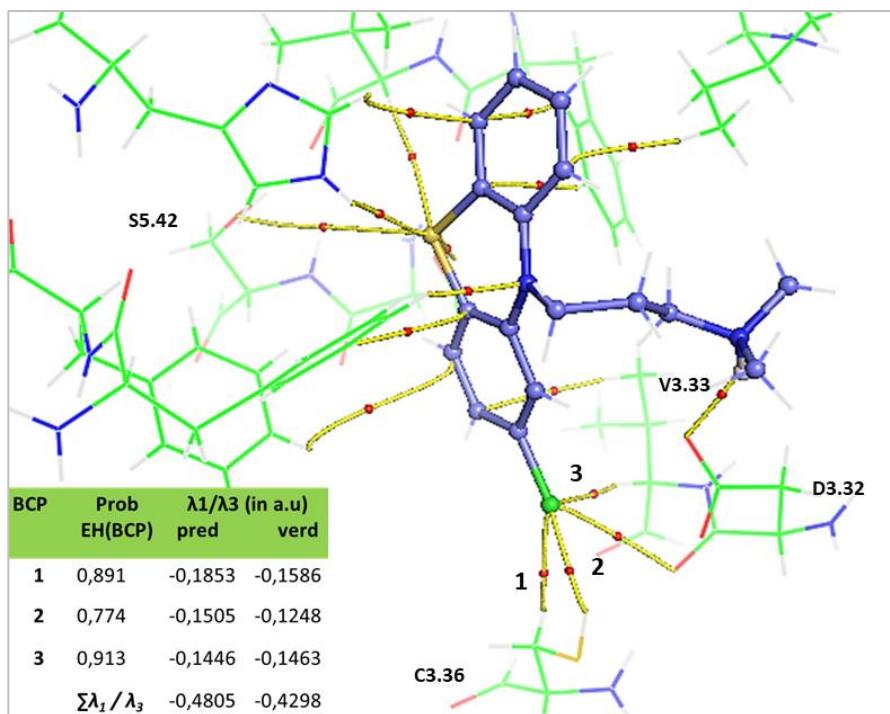


Figura 5.32. Grafo de densidad electrónica del complejo clorpromazina/D2Ra. Los caminos de enlace (BP) que conectan el cloro con el átomo de O (enlace X) y los átomos de H (enlaces H) se representan en amarillo. Los círculos rojos en la parte superior de los BP son los puntos críticos de enlace (BCP). Los BCP para los enlaces de hidrógeno H···Cl están numerados y sus propiedades correspondientes se enumeran en la tabla insertada (probabilidad estimada de formación de EH (BCP) y valor predicho/verdadero de la relación λ_1/λ_3)

5.4 Conclusiones

En este capítulo se modelaron las estructuras 3D del D2R en estados activo e inactivo en función de los datos cristalográficos existentes. Luego, los modelos predichos se validaron utilizando diferentes métodos. Para obtener un modelo de homología del receptor en su forma activada acoplado a la proteína Gi se propuso el uso de múltiples plantillas combinadas con el refinamiento de dichas estructuras por simulaciones de Dinámica Molecular embebidas en un modelo de membrana. Los modelos obtenidos se usaron para estudiar los cambios conformacionales que se producen en el D2R durante su activación a nivel molecular y en las interacciones entre átomos, así se pudo analizar la presencia de los “switches” moleculares de activación estudiados previamente en otros receptores acoplados a la proteína G. Se validaron los modelos estructurales de formas activas e inactivas de D2R generadas mediante simulaciones de dinámica molecular y “docking” molecular. El estudio detallado de las interacciones moleculares en el sitio de unión se realizó mediante análisis QTAIM.

Una vez obtenidas las estructuras cristalinas de ambas formas del receptor, en una segunda etapa, se procedió a la comparación con los modelos obtenidos durante la tesis. Se trabajó con estas estructuras con el fin de explorar el comportamiento conformacional y el modo de unión en el sitio activo de los agonistas dopamina y 3-clorotiramina, con este fin se realizó un estudio de las interacciones sobre ambas formas del receptor. Los resultados obtenidos demostraron que estos compuestos presentan un modo de unión más estable con las formas activa del receptor, lo que se comprobó a través de los perfiles de energía del Docking Molecular corroborando así el hecho de que agonistas actúan preferentemente con el receptor activado.

Como un resultado importante del trabajo realizado en este capítulo se considera que este tipo de análisis podría ser de utilidad en la clasificación de ligandos agonistas y agonistas inversos de receptores dopaminérgicos.

Finalmente, se pudo aplicar el protocolo para predecir la formación de EXs (xenv) en entornos proteicos, detallado en el capítulo 4, a ambas formas del receptor D2Ra y D2Ri, unido al compuesto agonista 3-clorotiramina y al agonista inverso, clorpromazina, respectivamente, y se compararon los resultados de la predicción con los datos obtenidos del análisis QTAIM, comprobando que concuerdan correctamente.

5.5 Referencias:

- Andujar, S. A., de Angel, B. M., Charris, J. E., Israel, A., Suárez-Roca, H., López, S. E., Garrido, M. R., Cabrera, E. V., Visbal, G., Rosales, C., Suvire, F. D., Enriz, R. D., & Angel-Guío, J. E. (2008). Synthesis, dopaminergic profile, and molecular dynamics calculations of N-aralkyl substituted 2-aminoindans. *Bioorganic & Medicinal Chemistry*, 16(6), 3233–3244. <https://doi.org/10.1016/J.BMC.2007.12.027>
- Angelina, E., Andujar, S., Moreno, L., Garibotto, F., Párraga, J., Peruchena, N., Cabedo, N., Villecco, M., Cortes, D., & Enriz, R. D. (2015). 3-chlorotyramine acting as ligand of the D2 dopamine receptor. Molecular modeling, synthesis and D2 receptor affinity. *Molecular Informatics*, 34(1), 28–43. <https://doi.org/10.1002/minf.201400093>
- Carpenter, B., Nehmé, R., Warne, T., Leslie, A. G. W., & Tate, C. G. (2016). Structure of the adenosine A2A receptor bound to an engineered G protein. *Nature*, 536(7614), 104–107. <https://doi.org/10.1038/nature18966>
- Carpenter, B., & Tate, C. G. (2017). Active state structures of G protein-coupled receptors highlight the similarities and differences in the G protein and arrestin coupling interfaces. *Current Opinion in Structural Biology*, 45(Figure 1), 124–132. <https://doi.org/10.1016/j.sbi.2017.04.010>
- Cheng, Z., Garvin, D., Paguio, A., Stecha, P., Wood, K., & Fan, F. (2010). Luciferase Reporter Assay System for Deciphering GPCR Pathways. *Current Chemical Genomics*, 4, 84–91. <https://doi.org/10.2174/1875397301004010084>
- Chien, E. Y. T., Liu, W., Zhao, Q., Katritch, V., Han, G. W., Hanson, M. A., Shi, L., Newman, A. H., Javitch, J. A., Cherezov, V., & Stevens, R. C. (2010). Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist. *Science*, 330(6007), 1091–1095. <https://doi.org/10.1126/science.1197410>
- Cox, B. A., Henningsen, R. A., Spanoyannis, A., Neve, R. L., & Neve, K. A. (1992). Contributions of Conserved Serine Residues to the Interactions of Ligands with Dopamine D2 Receptors. *Journal of Neurochemistry*, 59(2), 627–635. <https://doi.org/10.1111/j.1471-4159.1992.tb09416.x>
- De Lean, A., Stadel, J. M., & Lefkowitz, R. J. (1980). A ternary complex model explains the agonist-

specific binding properties of the adenylyl cyclase-coupled beta-adrenergic receptor. *Journal of Biological Chemistry*, 255(15), 7108–7117. [https://doi.org/10.1016/S0021-9258\(20\)79672-9](https://doi.org/10.1016/S0021-9258(20)79672-9)

Deng, Z., Chuaqui, C., & Singh, J. (2004). Structural Interaction “fingerprint” (SIFT): A Novel Method for Analyzing Three-Dimensional Protein-Ligand Binding Interactions. *Journal of Medicinal Chemistry*, 47(2), 337–344. <https://doi.org/10.1021/jm030331x>

Deupi, X., & Standfuss, J. (2011). Structural insights into agonist-induced activation of G-protein-coupled receptors. *Current Opinion in Structural Biology*, 21(4), 541–551. <https://doi.org/10.1016/j.sbi.2011.06.002>

Gaulton, A., Bellis, L. J., Bento, A. P., Chambers, J., Davies, M., Hersey, A., Light, Y., McGlinchey, S., Michalovich, D., Al-Lazikani, B., & Overington, J. P. (2012). ChEMBL: A large-scale bioactivity database for drug discovery. *Nucleic Acids Research*, 40(D1), 1100–1107. <https://doi.org/10.1093/nar/gkr777>

Herrmann, R., Heck, M., Henklein, P., Hofmann, K. P., & Ernst, O. P. (2006). Signal Transfer from GPCRs to G Proteins: ROLE OF THE G α N-TERMINAL REGION IN RHODOPSIN-TRANSDUCIN COUPLING. *Journal of Biological Chemistry*, 281(40), 30234–30241. <https://doi.org/10.1074/JBC.M600797200>

Huang, Y., & Thathiah, A. (2015). Regulation of neuronal communication by G protein-coupled receptors. *FEBS Letters*, 589(14), 1607–1619. <https://doi.org/10.1016/j.febslet.2015.05.007>

Ibrahim, M. A. A. (2012). AMBER empirical potential describes the geometry and energy of noncovalent halogen interactions better than advanced semiempirical quantum mechanical method PM6-DH2X. *Journal of Physical Chemistry B*, 116(11), 3659–3669. <https://doi.org/10.1021/jp3003905>

Jeffrey, G. A., & Saenger, W. (n.d.). *Hydrogen Bonding in Biological Structures*.

Kling, R. C., Lanig, H., Clark, T., & Gmeiner, P. (2013). Active-State Models of Ternary GPCR Complexes: Determinants of Selective Receptor-G-Protein Coupling. *PLoS ONE*, 8(6). <https://doi.org/10.1371/journal.pone.0067244>

Kozell, L. B., & Neve, K. A. (1997). Constitutive Activity of a Chimeric D2/D1Dopamine Receptor. *Molecular Pharmacology*, 52(6), 1137–1149. <https://doi.org/10.1124/MOL.52.6.1137>

Kruse, A. C., Ring, A. M., Manglik, A., Hu, J., Hu, K., Eitel, K., Hübner, H., Pardon, E., Valant, C., Sexton, P. M., Christopoulos, A., Felder, C. C., Gmeiner, P., Steyaert, J., Weis, W. I., Garcia, K. C., Wess, J., & Kobilka, B. K. (2013). Activation and allosteric modulation of a muscarinic acetylcholine receptor. *Nature*, 504(7478), 101–106. <https://doi.org/10.1038/nature12735>

Lan, H., DuRand, C. J., Teeter, M. M., & Neve, K. A. (2006). Structural determinants of pharmacological specificity between D1 and D2 dopamine receptors. *Molecular Pharmacology*, 69(1), 185–194. <https://doi.org/10.1124/mol.105.017244>

Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*, 26(2), 283–291. <https://doi.org/10.1107/s0021889892009944>

Lomize, M. A., Lomize, A. L., Pogozheva, I. D., & Mosberg, H. I. (2006). OPM: Orientations of proteins in membranes database. *Bioinformatics*, 22(5), 623–625. <https://doi.org/10.1093/bioinformatics/btk023>

Luchi, A. M., Angelina, E. L., Andujar, S. A., Enriz, R. D., & Peruchena, N. M. (2016). Halogen bonding in biological context: a computational study of D2 dopamine receptor. *Journal of Physical Organic Chemistry*, 29(11), 645–655. <https://doi.org/10.1002/poc.3586>

M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ort, J. B. F. and D. J. F. (2016). *Gaussian 09* (No. A02). Gaussian Inc.

Mahmood, M. I., Liu, X., Neya, S., & Hoshino, T. (2013). Influence of lipid composition on the structural stability of G-protein coupled receptor. *Chemical and Pharmaceutical Bulletin*, 61(4), 426–437. <https://doi.org/10.1248/cpb.c12-01059>

Mansour, A., Meng, F., Meador-Woodruff, J. H., Taylor, L. P., Civelli, O., & Akil, H. (1992). Site-directed mutagenesis of the human dopamine D2 receptor. *European Journal of Pharmacology: Molecular Pharmacology*, 227(2), 205–214. [https://doi.org/10.1016/0922-4106\(92\)90129-J](https://doi.org/10.1016/0922-4106(92)90129-J)

Morris, G. ., Huey, R., Lindstrom, W., Sanner, M. ., Belew, R., Goodsell, D., & Olson, A. . (2009). AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility.

Journal of Computational Chemistry, 30(16), 2785–2791. <https://doi.org/DOI 10.1002/jcc>

O'Boyle, N. M., Banck, M., James, C. A., Morley, C., Vandermeersch, T., & Hutchison, G. R. (2011).

Open Babel: An open chemical toolbox. *Journal of Cheminformatics*, 3(33), 33.

<https://jcheminf.biomedcentral.com/track/pdf/10.1186/1758-2946-3-33>

[33%0Ahttp://www.ncbi.nlm.nih.gov/pubmed/21982300%0Ahttp://www.ncbi.nlm.nih.gov/articlerender.fcgi?artid=PMC3198950](http://www.ncbi.nlm.nih.gov/pubmed/21982300)

Orry, A., & Abagyan, R. (2012). Homology Modelling. Methods and protocols. In *Humana Press*.

<https://doi.org/10.1007/978-1-61779-588-6>

Park, J. H., Scheerer, P., Hofmann, K. P., Choe, H. W., & Ernst, O. P. (2008). Crystal structure of the

ligand-free G-protein-coupled receptor opsin. *Nature*, 454(7201), 183–187.

<https://doi.org/10.1038/nature07063>

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera - A visualization system for exploratory research and analysis.

Journal of Computational Chemistry, 25(13), 1605–1612.

<https://doi.org/10.1002/jcc.20084>

Rasmussen, S. G. F., Choi, H., Fung, J. J., Pardon, E., Chae, P. S., Devree, B. T., Rosenbaum, D. M.,

Thian, F. S., Kobilka, T. S., Schnapp, A., Konetzki, I., Sunahara, R. K., Samuel, H., Pautsch, A.,

Steyaert, J., Weis, W. I., & Kobilka, B. K. (2011). Structure of a nanobody-stabilized active

state of the $\beta 2$ adrenoceptor. *Nature*, 469(7329), 175–180.

<https://doi.org/10.1038/nature09648.Structure>

Rubenstein, R. C., Wong, S. K. F., & Ross, E. M. (1987). The hydrophobic tryptic core of the β -

adrenergic receptor retains G(s) regulatory activity in response to agonists and thiols.

Journal of Biological Chemistry, 262(34), 16655–16662. [https://doi.org/10.1016/s0021-9258\(18\)49305-2](https://doi.org/10.1016/s0021-9258(18)49305-2)

Šali, A., & Blundell, T. L. (1993). Comparative protein modelling by satisfaction of spatial

restraints. In *Journal of Molecular Biology* (Vol. 234, Issue 3, pp. 779–815).

<https://doi.org/10.1006/jmbi.1993.1626>

Scheerer, P., Park, J. H., Hildebrand, P. W., Kim, Y. J., Krauß, N., Choe, H. W., Hofmann, K. P., &

Ernst, O. P. (2008). Crystal structure of opsin in its G-protein-interacting conformation.

Nature, 455(7212), 497–502. <https://doi.org/10.1038/nature07330>

Schertler, G. F. X. (1999). Structure of rhodopsin. *Novartis Foundation Symposium*, 224, 54–69.
<https://doi.org/10.1002/9780470515693.ch5>

Schrodinger, L. (2010). *The PyMOL Molecular Graphics System* (1.8).

Šukalović, V., Andrić, D., Roglić, G., Kostić-Rajačić, S., Schrattenholz, A., & Šoškić, V. (2005). Synthesis, dopamine D 2 receptor binding studies and docking analysis of 5-[3-(4-arylpiperazin-1-yl)propyl]-1H-benzimidazole, 5-[2-(4-arylpiperazin-1-yl)ethoxy]-1H-benzimidazole and their analogs. *European Journal of Medicinal Chemistry*, 40(5), 481–493. <https://doi.org/10.1016/j.ejmech.2004.10.006>

Teeter, M. M., Froimowitz, M., Stec, B., & DuRand, C. J. (1994). Homology Modeling of the Dopamine D2 Receptor and Its Testing by Docking of Agonists and Tricyclic Antagonists. *Journal of Medicinal Chemistry*, 37(18), 2874–2888.
<https://doi.org/10.1021/jm00044a008>

Trujillo, R., Flores, G., & Arias-Montaño, J. (2000). Dopamina : síntesis , liberación y receptores en el Sistema. *Rev Ista Biomedica*, 11(1), 39–60.

Trzaskowski, B., Latek, D., Yuan, S., Ghoshdastider, U., Debinski, A., & Filipek, S. (2012). Action of Molecular “switches” in GPCRs - Theoretical and Experimental Studies. *Current Medicinal Chemistry*, 19(8), 1090–1109. <https://doi.org/10.2174/092986712799320556>

Wang, S., Che, T., Levit, A., Schochet, B. K., Wacker, D., & Roth, B. L. (2018). STRUCTURE OF THE D2 DOPAMINE RECEPTOR BOUND TO THE ATYPICAL ANTISSCHIZOPHRENIC DRUG RISPERIDONE. *Nature*, 555(7695), 269–273. <https://doi.org/10.1038/nature25758>

Yashar, M., Kalani, S., Vaidehi, N., Hall, S. E., Trabanino, R. J., Freddolino, P. L., Kalani, M. A., Floriano, W. B., Wai, V., Kam, T., & Goddard Iii, W. A. (n.d.). *The predicted 3D structure of the human D2 dopamine receptor and the binding site and binding affinities for agonists and antagonists.* Retrieved September 21, 2017, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC374327/pdf/1013815.pdf>

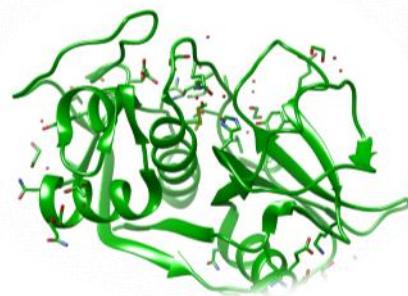
Yin, J., Chen, K. Y. M., Clark, M. J., Hijazi, M., Kumari, P., Bai, X. chen, Sunahara, R. K., Barth, P., & Rosenbaum, D. M. (2020). Structure of a D2 dopamine receptor–G-protein complex in a lipid membrane. *Nature*, 584(7819), 125–129. <https://doi.org/10.1038/s41586-020-2379-5>

Zhuang, Y., Xu, P., Mao, C., Wang, L., Krumm, B., Zhou, X. E., Huang, S., Liu, H., Cheng, X., Huang, X.

P., Shen, D. D., Xu, T., Liu, Y. F., Wang, Y., Guo, J., Jiang, Y., Jiang, H., Melcher, K., Roth, B. L., ...
Xu, H. E. (2021). Structural insights into the human D1 and D2 dopamine receptor
signaling complexes. *Cell*, 184(4), 931-942.e18.
<https://doi.org/10.1016/j.cell.2021.01.027>

Capítulo 6

Estudio de inhibidores de la enzima
Cruzipaina y sus interacciones
a través de métodos de modelado
molecular



CAPITULO 6

El presente capítulo de esta tesis tiene como objetivo estudiar la importancia de los EX en la estabilidad y afinidad de inhibidores halogenados de la enzima Cruzipaina, a través de la Mecánica Cuántica/Mecánica Molecular (QM/MM) aplicando el análisis de densidad electrónica y simulaciones de dinámica molecular. Para ello se trabajó con ligandos covalentes y no covalentes de la enzima, comparando los resultados de los ligandos halogenados con sus respectivos análogos no halogenados.

6.1 Introducción

6.1.1 Enfermedad del Chagas

La enfermedad de Chagas, causada por el parásito protozoario *Trypanosoma cruzi*, es endémica en América Central y Sudamérica y representa la miocarditis más frecuente a nivel mundial. El establecimiento de la infección crónica conduce a una patología cardíaca debilitante que puede llevar a la muerte. El parásito *T. cruzi* es transmitido por vectores hematófagos pertenecientes a la subfamilia Triatominae, familia Reduviidae, orden Hemiptera, los cuales depositan sus heces infectadas sobre la piel al picar. *Triatoma infestans*, *Rhodnius prolixus* y *Panstrongylus megistus* son los vectores más importantes (Ibañez De Barrett et al., 2004). Actualmente, existen 8-9 millones de personas infectadas por *T. cruzi* y otros 40 millones con riesgo de adquirir la enfermedad. Investigaciones recientes indican que cada año ocurren 14.000 muertes asociadas con esta condición ubicándose a nivel mundial en tercer lugar entre las afecciones parasitarias (<https://dndial.org/es/>). (Figura 6.1)

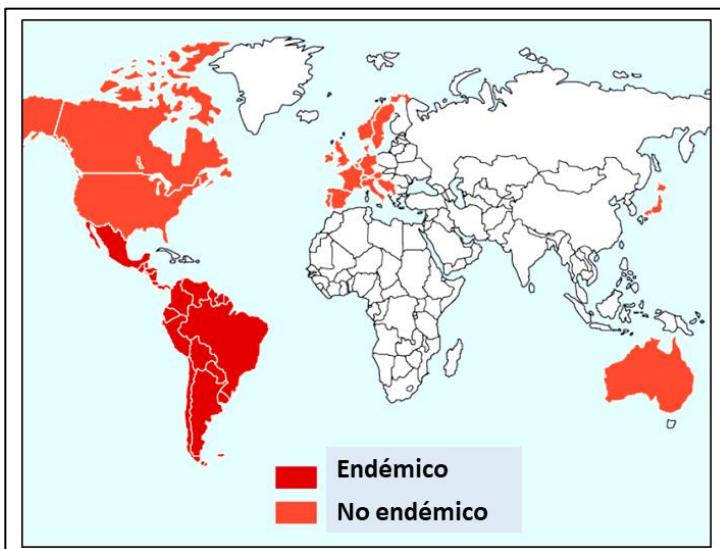


Figura 6.1: Distribución geográfica de la enfermedad de Chagas. Adaptado de www.dndi.org

El principal factor de transmisión del *Trypanosoma cruzi* es la picadura de la vinchuca, *Triatoma infestans*. En cada ingesta de sangre la vinchuca deposita sus heces sobre la piel del mamífero y el Trypanosoma penetra por las microescoriaciones que la víctima genera en su piel través del rascado o por casualidad, al poner en contacto las heces del insecto con mucosas o conjuntivas. El segundo mecanismo de transmisión de la enfermedad es la transfusión de sangre infectada, que ocurre con frecuencia y es la responsable de miles de casos nuevos cada año. La incidencia está estrechamente relacionada con los programas de salud en vigencia en las zonas donde se efectúan las transfusiones. El tercer factor en importancia es la transmisión congénita, que genera hasta 8000 casos nuevos por año en las regiones endémicas (Gloss et al., 1990).

La enfermedad de Chagas tiene 3 etapas: una aguda, una indeterminada y una crónica (Carlomagno M, Cura E, 1989). La fase aguda se presenta inmediatamente después de la infección, suele producir manifestaciones específicas, como la oclusión de un ojo o *signo de Romagna* y va acompañada de una fuerte parasitemia, inrnunosupresión y activación policlonal. Una vez concluida la fase aguda, los infectados pasan por un largo período intermedio en el cual no hay manifestaciones clínicas de la infección. En un tercio de la población infectada se desarrollan los síntomas propios de la etapa crónica, como la miocardiopatía o las patologías del aparato digestivo con diversos grados de severidad. La infección en los adultos suele pasar desapercibida. En los niños, en cambio, es alta la probabilidad de desarrollar los síntomas de la etapa aguda, en la cual los parásitos pueden ser detectados en la sangre. Si bien en algunos casos la fase aguda resulta fatal, en la mayoría de los

enfermos la respuesta inmune logra reducir el nivel de parasitemia a niveles indetectables, y el hospedador entra en la etapa crónica de la enfermedad.

El *T. cruzi* presenta un ciclo de vida complejo, con estadios especializados para la supervivencia tanto en el vector como en el mamífero (Figura 6.2). Es un parásito con un ciclo cerrado que involucra vertebrados e invertebrados y que asume tres distintas morfologías y características bioquímicas en cada estadio (Bern, 2015; Carr et al., 2021).

Las dos formas replicativas, epimastigote y amastigote, están presentes en el intestino medio del insecto y en el citoplasma de una célula de mamífero, respectivamente. Los trypomastigotes, una forma no replicativa que puede invadir activamente las células de mamíferos, está presente en el torrente sanguíneo o en el intestino posterior del insecto (trypomastigotes metacíclicos) (Brener, 2003).

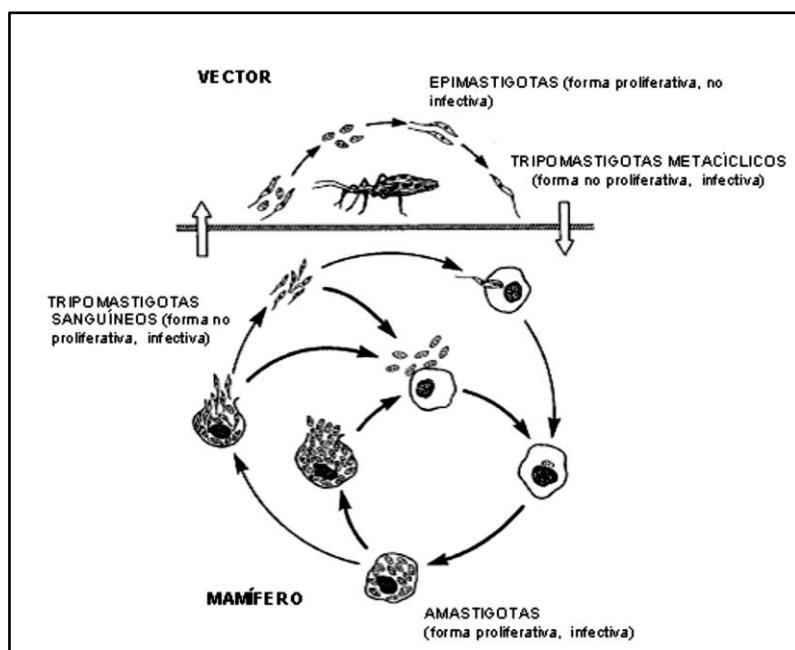


Figura 6.2: Ciclo de vida de *T. cruzi*. Adaptado de Urbina, 1999.

Amastigote: esférico u ovalado, es la forma reproductiva en el interior de las células mamíferas.

Epimastigote: alargado y con el cinetoplasto localizado con anterioridad al núcleo, es la forma reproductiva en el tracto digestivo de los invertebrados y en medios de cultivo.

Trypomastigote: también alargado, pero con el cinetoplasto localizado posteriormente al núcleo. Se encuentra en la sangre de los mamíferos (en donde es el estadio infectante), así como también en las heces de los insectos. Esta forma no se divide.

Después de ser ingeridos por el insecto, los trypomastigotes sanguíneos se transforman en epimastigotes, los cuales proliferan en la luz del tracto gastrointestinal. Los epimastigotes se adhieren a la superficie epitelial en la porción terminal del tubo digestivo (recto) y se diferencian a la forma infectiva trypomastigote metacíclica. Estas formas son eliminadas en las heces cuando el insecto pica a un mamífero e ingresan a éste a través de la piel o mucosas. Estas formas no proliferativas evaden el sistema inmune y pueden invadir diversos tipos de células, incluyendo células musculares y nerviosas del corazón y el tracto gastrointestinal, así como también células del sistema retículo endotelial e incluso neuronas (Tyler & Engman, 2001; Urbina, 1999).

La membrana plasmática de *Trypanosoma cruzi* tiene moléculas (péptidos y carbohidratos) que participan en distintas etapas de la invasión a la célula hospedadora. Algunas intervienen en la adhesión del parásito, otras en su internalización y una vez en el interior de la célula, algunas moléculas participan en el escape de la vacuola fagocítica. *Trypanosoma cruzi* tiene la capacidad de invadir y replícarse en células fagocíticas o no fagocíticas del hospedador mamífero. Producida la penetración, el trypanosoma queda encerrado en una vacuola fagocítica de la cual escapa rápidamente hacia el citoplasma. Los receptores de *T. cruzi* presentes en la superficie de las células de mamíferos están sialilados y se ha visto que su remoción afecta a la invasión. Por otro lado, existen evidencias de que la presencia de ácido siálico en la superficie del trypanosoma beneficia dicho proceso (Schenkman & Eichinger, 1993).

La invasión por *T. cruzi* es un proceso activo y de eventos simultáneos. En primer lugar, el parásito deberá adherirse a moléculas ricas en ácido siálico, proteoglicanos o moléculas en la matriz extracelular, como el colágeno y la fibronectina. En el proceso de penetración parecen sucederse reacciones de sialidación y trans-sialidación de las moléculas del parásito y de la célula huésped (Schenkman & Eichinger, 1993). Luego de atravesar la membrana plasmática, el trypomastigote se encontrará en el fagolisosoma o vacuola parasitófora, que está adosada a la membrana plasmática y tiene un pH ácido. La trans-sialidasa está involucrada en el escape de la vacuola parasitófora a través de la remoción del ácido siálico de las glicoproteínas lisosomales. Además, *Trypanosoma cruzi* posee una proteína llamada TcTOX, relacionada inmunológicamente con el factor C9 del sistema de complemento. Es una toxina acídica formadora de poros transmembranares, también conocida como penetrina (Andrews et al., 1990).

Después de la invasión, los parásitos abandonan los fagolisosomas y se diferencian a la forma amastigota, la cual prolifera por fisión binaria y eventualmente se diferencia a la forma

trypomastigote sanguínea a través de formas intermedias, algunas de las cuales presentan una morfología epimastigote (Almeida et al., 1999; Tyler & Engman, 2001). Una vez formados, los trypomastigotes sanguíneos destruyen la célula hospedera pasando a la sangre pudiendo invadir otras células o ser ingeridos por un vector apropiado, completándose así el ciclo (Urbina, 1999).

Los amastigotes alteran la fisiología de las células influenciando directamente la homeostasis microvascular. Un ejemplo de ello es que la adherencia plaquetaria a las células endoteliales infectadas es mayor que la observada en células normales. Estos estudios han generado una hipótesis alternativa para la cardiopatía chagásica: las lesiones en la microcirculación pueden ser la base de la disfunción y la patología cardiaca. Los factores que contribuyen a la disfunción endotelial son varios y están directamente asociados con los amastigotes, así como lo están indirectamente con la respuesta inmune.

Las proteasas de *Trypanosoma cruzi* participan en la nutrición del parásito a expensas del hospedador y en otros aspectos de dicha relación (Cazzulo, 1994); hay algunas evidencias que sugieren un rol en la penetración del trypomastigote en la célula hospedadora. Dicho evento es sensible a algunos inhibidores de proteasas, tal como demostraron Piras et al. (Piras et al., 1985). Las proteasas también podrían participar en la evasión de la respuesta inmune mediante la digestión intracelular de inmunoglobulinas G humanas endocitadas una vez reconocido el antígeno (Santana et al., 1992).

Dentro de las proteasas presentes en el parásito se destaca la Cruzipaína (Cz) (Burleigh et al., 2002; Dos Reis et al., 2006). La cual fue purificada a través de electroforesis por Bontempi et al., (Bontempi et al., 1984) a partir de epimastigotes de la cepa Tu12 de *T. cruzi*. La enzima resultó ser monomérica y de 60 kDa de peso molecular aparente. Entre sus características principales se destacó la utilización de azocaseína, caseína, seroalbúmina bovina y hemoglobina como sustratos. La máxima actividad se obtuvo sobre seroalbúmina bovina a pH 4 y también se observó actividad sobre las proteínas solubles del parásito, con un pH óptimo 3.

Estudios posteriores demostraron que la Cruzipaína tiene la capacidad de degradar IgGs humanas (Bontempi & Cazzulo, 1990). El fragmento Fab de tales inmunoglobulinas fue degradado levemente, no así el fragmento Fc, que fue extensamente hidrolizado a péptidos menores. La cisteína-proteasa principal de *T. cruzi* es también capaz de degradar las cadenas A y B de la insulina bovina oxidada (Raimondi et al., 1991). Ambas resultan extensamente hidrolizadas, con 4 sitios de corte en la cadena A y 6 en la cadena B, varios de los cuales pertenecen a residuos de tipo hidrofóbico.

Al no existir vacunas, la quimioterapia sigue siendo el único método para tratar individuos afectados o para disminuir la trasmisión poblacional. El tratamiento de la enfermedad de Chagas involucra los medicamentos de primera línea nifurtimox y benznidazol. Estos dos medicamentos antiguos son efectivos para curar la infección principalmente en la fase aguda, con una cura exitosa de hasta el 80%, pero muestran escasa efectividad en pacientes con infección crónica. Además, debido a sus efectos colaterales, nifurtimox ya no está disponible en la mayoría de los países latinoamericanos. Así también, se ha observado resistencia a los medicamentos en algunas cepas de trypanosomas. Por lo tanto, es necesario el descubrimiento de medicamentos nuevos, más seguros y más efectivos para tratar el Chagas. Dadas las limitaciones del restringido arsenal terapéutico disponible actualmente, se han señalado como necesidades de investigación prioritarias la investigación de nuevas dianas moleculares, la exploración/screening de bibliotecas químicas públicas y privadas, la explotación de la biodiversidad de las Américas en la búsqueda de nuevas terapias contra la enfermedad de Chagas y el ensayo preclínico y clínico de nuevos candidatos (WHO, 2007).

Algunos procesos y/o enzimas que se utilizan actualmente como blancos terapéuticos son: tripanotión reductasa; biosíntesis de esteroles de membrana; metabolismo del pirofosfato; enzimas glicolíticas; trans-sialidasas y cisteín-proteasas, entre otras (V. Duschak & Couto, 2007). El papel fundamental que juegan las cisteín proteasas en las infecciones parasitarias sumado a las sutiles diferencias que presentan con respecto a las correspondientes enzimas mamíferas, han validado esta familia de proteasas como blancos atractivos para el desarrollo de nuevos agentes quimioterápicos (Engel et al., 1999). Dentro de estas enzimas, Cruzipaína, la principal cisteín proteasa presente en *T. cruzi* ha sido una de las más estudiadas.

6.1.2 Cruzipaína

La Cruzipaína (Cz), es un miembro de la superfamilia de las papaínas, se expresa como una mezcla compleja de isoformas en los diferentes estadíos de desarrollo del parásito. Esta glicoproteína participa en la internalización del *T. cruzi* en las células mamíferas, lo que ha sido demostrado con inhibidores específicos de la enzima que interfieren en la invasión celular y la replicación del parásito (Engel et al., 1999). Además, esta enzima genera una fuerte respuesta immune en individuos infectados. Estas características hacen de la Cruzipaína un potencial blanco de drogas terapéuticas.

La Cz es una cisteína-proteasa análoga a la catepsina L (también conocida como gp51/57 o cruzaína, en su forma recombinante), responsable de la mayor parte de la actividad proteolítica de *T. cruzi*, en todos los estadios de su ciclo de vida (Caffrey et al., 2000; Inst et al., 2009), es codificada por diversos genes colocados en serie y separados por espacios intragénicos de algo más de 400 pares de bases (Campetella et al., 1992). Esta variabilidad génica resulta en la expresión de diversas isoformas de la enzima, algunas de las cuales presentan variaciones no conservativas en la secuencia de aminoácidos que influyen en la especificidad de la misma (V. G. Duschak et al., 2003; Lima et al., 2001). Se ha demostrado que la composición de isoformas varía entre cepas y en diferentes estadios del ciclo de vida del parásito lo cual podría explicar las variaciones inter-cepa observadas en la susceptibilidad a distintos inhibidores (Gomes et al., n.d.).

Diferentes estudios han demostrado que Cz se expresa de manera diferencial en los cuatro estadios principales del ciclo de vida de *T. cruzi* (Inst et al., 2009) (McKerrow et al., 2009), encontrándose en mayores niveles en la forma epimastigotes localizándose en organelas análogas a los lisosomas que almacenan proteínas que son digeridas por el parásito durante la etapa de diferenciación a la forma tripomastigote metacíclico (Souto-Padron et al., 1990); también se localiza en menor proporción en la membrana plasmática del parásito en todas sus formas, pero en mayores niveles en la forma amastigote (Nascimento and de Souza, 1996).

Según la convención adoptada por Schechter y Berger (Schechter & Berger, 1967) para identificar los residuos del sustrato y los subsitutos de la enzima, **P** designa a residuos presentes en el sustrato y **S** refiere a los correspondientes subsitutos de la enzima. Con una prima (**P'**) se identifica a los residuos ubicados hacia el extremo C-terminal del enlace peptídico hidrolizado, para diferenciarlos de los que se ubican en el extremo N-terminal.

Estructuralmente, consta de una única cadena polipeptídica que posee un dominio catalítico de 215 aminoácidos y un dominio C-terminal de 130 aminoácidos, que es característico de las cisteína proteasas de tripanosomatídeos (Figura 6.3). La estructura cristalográfica muestra que la cadena polipeptídica de 215 aminoácidos se pliega en dos dominios definidos. El dominio izquierdo presenta tres regiones de hélices alfa, una de las cuales contiene el residuo Cys25 en su extremo N-terminal. El dominio derecho presenta principalmente una disposición de hojas beta antiparalelas. La tríada catalítica, formada por los residuos de Cys25, His162 y Asn182, y los distintos subsitutos (S1-S4 y S1'-S3') de unión al sustrato se encuentran en una hendidura entre los dos dominios (Huang et al., 2003; McGrath et al., 1995) (Huang et al., 2003; McGrath et al., 1995).

Con respecto al domino C-terminal, el mismo es característico de las cisteín proteasas de Tipo I de los trypanosomátidos y otras proteasas de parásitos (Eakinsb et al., 1992). Este dominio se elimina de forma autocatalítica al madurar la enzima (Hellman et al., 1991) y se ha demostrado, en pruebas de inhibición por sustrato, que este fragmento de la enzima no interfiere en la misma (Stoka et al., 1998) .

Como se detalló previamente los distintos subsitios (S1- S4 y S1'-S3') de unión al sustrato se encuentran en una hendidura entre los dos dominios (McGrath et al., 1995). La especificidad de los distintos subsitios de Cz ha sido ampliamente estudiada y se ha observado que S1' presenta una notoria especificidad por agrupamientos aromáticos en P1', definida principalmente por la presencia del residuo de Trp184 (Brinen et al., 2000). Los sitios S1 y S3 presentan especificidad por agrupamientos básicos. La especificidad del sitio S2' no es clara, pero el residuo de Gln19 presente en el mismo juega un rol fundamental en la estabilización del oxianión formado durante la catálisis (Leung et al., 2000). Se ha notado que el sitio S2 es el principal determinante de la especificidad de esta enzima. Este subsitio es capaz de unir tanto residuos hidrofóbicos como básicos debido al comportamiento del residuo de Glu205. Este residuo, cargado negativamente a pH fisiológico, es capaz de adoptar una conformación orientada hacia el sustrato cuando el mismo presenta *agrupamientos básicos* en P2, con los que establece interacciones electrostáticas. Cuando P2 sustenta *agrupamientos hidrofóbicos*, Glu205 adopta una conformación orientada hacia el disolvente. Este comportamiento también explica la mayor selectividad de la enzima por sustratos hidrofóbicos a pH ácido debido a la protonación de Glu205 (Gillmor et al., 1997).

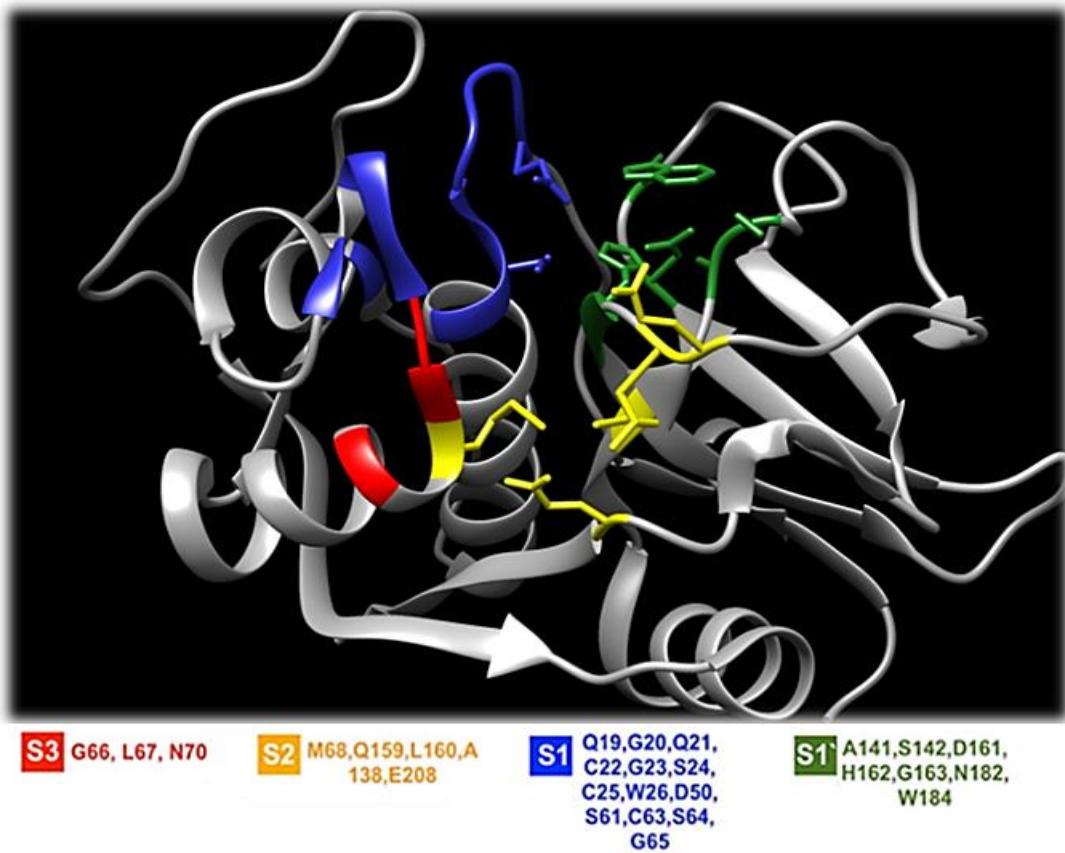


Figura 6.3: Estructura de Cz determinada por difracción de rayos X a una resolución de 1.28 Å, (PDB:3KKU). El inhibidor fue retirado del sitio activo para mayor claridad. En colores se muestran los principales residuos que delimitan la hendidura del sitio activo: azul S1; amarillo S2; rojo S3; verde S1'.

En relación al mecanismo de acción propuesto (Leung,2000) a través del cual ocurre la hidrólisis, se postula que la Cys25 y la His162 forman parte del sitio activo de la enzima, pero se ha cuestionado si alcanzan esos dos residuos, únicamente, para una actividad catalítica completa. El grupo imidazol de la histidina polariza el tiol de la cisteína permitiendo la desprotonación aún en condiciones neutras o débilmente ácidas. El par iónico S-imidazolium que se produce es altamente nucleófilico. Las cisteína proteasas hidrolizan el enlace amida de manera similar a las serin proteasas. Luego de la unión al sustrato se forma un complejo no covalente de Michaelis. El anión tiolato ataca al grupo carbonilo del enlace carbonilo que se romperá (Figura 6.4a). Luego se forma un complejo tetraédrico el cual es estabilizado por el hueco del oxianión (Figura 6.4b). A continuación, el grupo tiol de la enzima es acilado y el primer producto es liberado. La hidrólisis del complejo aciloenzima da como resultado la

formación del segundo intermediario tetraédrico (Figura 6.4c). Luego el intermediario colapsa y el ácido se libera regenerando la enzima (Figura 6.4d) (Leung et al., 2000).

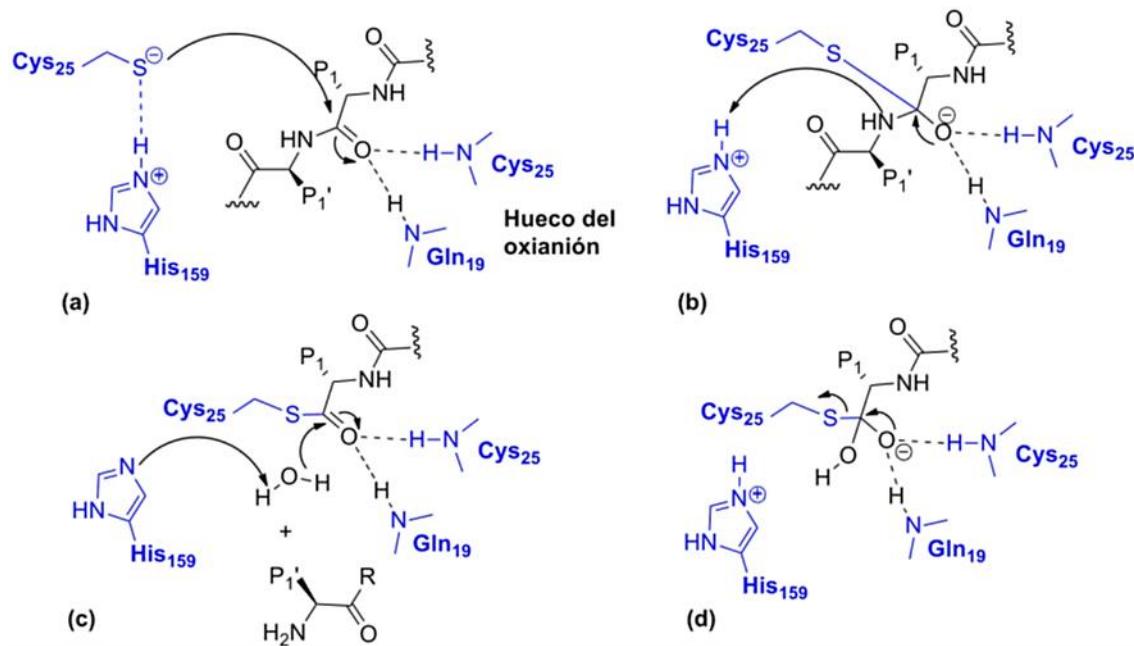


Figura 6.4: Mecanismo de acción propuesto para cisteína-proteasas (Leung, 2000)

La relevancia funcional del blanco molecular Cz es evidente, la misma juega un rol importante en diversas funciones del parásito tales como la invasión de las células huésped, degradando tejidos del huésped y facilitando la penetración del parásito en las células (Aparicio et al., 2004), y además participa en la evasión de la respuesta inmune por diferentes mecanismos (Benítez-Hernández et al., 2010). La participación de la enzima en la proliferación de epimastigotes y amastigotes y en la metaciclogénesis también ha sido demostrada (Santos et al., 2005; Tomás & Kelly, 1996); por otro lado, dada su función autocatalítica, la inhibición de Cz conduce a la acumulación de un precursor en el aparato de Golgi, provocando importantes alteraciones morfológicas, dando lugar eventualmente a un shock osmótico y a la muerte de *T. cruzi*; estos cambios morfológicos en el aparato de Golgi se han observado en parásitos tratados con diferentes inhibidores, lo cual sugiere que independientemente del mecanismo de inhibición, todos los inhibidores de Cz producen el mismo efecto (Capaci Rodrigues et al., 2011; Inst et al., 2009).

La enzima es estable en un rango de pH entre 4.0 y 9.0 (Cazzulo et al., 1990; V. G. Duschak et al., 2003) siendo el pH óptimo para su actividad dependiente del sustrato. La hidrólisis de seroalbúmina bovina (BSA) y hemoglobina es óptima a pH 4.0, mientras que para la azocaseína

la máxima actividad se observa a pH 5.0 (V. G. Duschak et al., 2003). El rango óptimo de pH para distintos sustratos cromogénicos varía entre 7.0 y 9.0 mientras que para sustratos fluorogénicos es de 6.3 a 7.2 ((Cazzulo et al., 1990; Stoka et al., 1995).

En los últimos quince años diversos grupos de investigación han enfocado sus proyectos en desarrollar inhibidores de Cz (Capaci Rodrigues et al., 2011; V. Duschak & Couto, 2009; Inst et al., 2009; *World Health Organization (WHO) Special Programme for Research and Training in Tropical Diseases (TDR) Neglected Tropical Diseases (NTD) - Global Health Impact Group*, n.d.) La mayoría de los de potenciales antichagásicos registradas entre 2000 y 2006 corresponden a inhibidores de cisteína proteasas (V. Duschak & Couto, 2007).

Las investigaciones relacionadas a inhibidores de Cz se enfocaban primeramente en *inhibidores irreversibles*, generalmente peptidomiméticos, pero se encontraron dificultades principalmente vinculadas a la carencia de selectividad, alta reactividad, toxicidad y baja biodisponibilidad oral de los mismos (Capaci Rodrigues et al., 2011; Jacobsen et al., 2000; Mallari et al., 2008; Powers et al., 2002).

Por otra parte, los pocos inhibidores reversibles que se han sintetizados muestran una muy buena actividad en Cz *in vitro*, pero cuando se evaluaron *in vivo*, resultaron inactivos debido a problemas de biodisponibilidad.

6.1.3 Inhibidores de la Cruzipain

Existen varios estudios enfocados a identificar inhibidores irreversibles que se unan covalentemente a la proteasa. Estos inhibidores contienen grupos electrofílicos entre los cuales se encuentran las peptidildiazometilcetonas, peptidilfluorometilcetonas y peptidil vinil sulfonas. Estos inhibidores son capaces de bloquear la diferenciación del parásito en las distintas fases de su ciclo de vida. El derivado vinil sulfónico, la N-piperazina-Fenil-hFenil-vinil sulfona fenil resultó ser un excelente inhibidor de la Cruzipain, provocando la muerte de los parásitos mediante la inducción de la acumulación de la proteasa no procesada en el aparato de Golgi. Los inhibidores covalentes irreversibles de la enzima reaccionan después de unirse a ella y, por lo tanto, la inactivación de la enzima no es un proceso de equilibrio como con la inhibición reversible, y requiere la reexpresión de la enzima para revertir la acción del fármaco, que puede ocurrir después de la eliminación del mismo del organismo. En el desarrollo de fármacos se observa una tendencia a evitar los inhibidores irreversibles y covalentes para evitar el riesgo de: 1) efectos secundarios impredecibles como la generación de proteínas alergénicas modificadas (haptenos); 2) modificación irreversible no específica de proteínas

fuerza del objetivo; y 3) la dificultad en el seguimiento de los metabolitos cuando se une covalentemente a las proteínas. Sin embargo, se ha aprovechado el advenimiento de los inhibidores covalentes en la terapia del cáncer para abrir las compuertas a dichos medicamentos.

Las cisteína proteasas son inhibidas eficazmente por varias clases de inhibidores (tabla 6.1) incluidos aquellos basados en estados de transición reversibles e irreversibles (Powers et al., 2002). Ejemplos de inhibidores reversibles son aldehídos peptídicos, α -dicetonas, α -cetoésteres, α -cetoamidas y α -cetoácidos (Götz et al., 2004). También se inhiben irreversiblemente por peptidildiazometil cetonas, fluorometil cetonas, péptidoepóxidos y vinilsulfonas (Scheidt et al., 1998). Los inhibidores con estructuras peptídicas tienen el problema adicional de ser inestables en plasma, aunque en los últimos tiempos se ha mejorado mucho con la utilización aminoácidos no naturales (McKerrow et al., 2006).

Aunque la relación entre la inhibición reversible e irreversible no se ha explorado completamente, en general los inhibidores irreversibles de la Cruzipaína han curado con éxito la infección parasitaria, lo que implica que la unión fuerte a la enzima puede ser esencial.

Actualmente, 27 estructuras químicas están asociadas a este blanco molecular en el *Protein Data Bank* (rcsb.org)(Burley et al., 2021) donde la Cz se encuentra cristalizada con inhibidores reversibles e irreversibles. Entre los inhibidores, los que contienen una vinilsulfona en su estructura presentan buena selectividad y un favorable desarrollo prospectivo a pesar de la naturaleza irreversible. En este contexto, Jaishankar (Jaishankar et al., 2008) sintetizó y determinó las constantes de inhibición de los ligandos frente a Cz de una serie de vinilsulfonas análogas al inhibidor K-777.

Tomando como esqueleto los inhibidores reversibles no básicos de las catepsinas, se han desarrollado potentes inhibidores reversibles para la Cruzipaína. Se demostró que la presencia de un motivo parecido a la fenilalanina era muy importante y el mejor sustituyente fue el 4-ciano-2fluorofenilalanina, que presentó un incremento de ocho veces en potencia con respecto a la fenilalanina no sustituida. El mejor sustituyente en la posición P3 fue el 4-metanosulfonil bifenilo. Estudios previos habían demostrado que los grupos funcionales posicionados en P2 determinaban la selectividad entre la Cruzipaína y las catepsinas (Gauthier et al., 2008).

También, se han descubierto inhibidores no covalentes a través de plataformas de detección de alto rendimiento (HTS) y, a pesar de su menor potencia relativa a compuestos covalentes reportados previamente, ellos representan importantes avances en el desarrollo de

compuestos no peptídicos con características similares a las drogas (Ferreira et al., 2010; Schirmeister & Kaeplker, 2005)

Tabla 6.1: Estructura de los principales inhibidores de la enzima Cruzipain

Familia	Estructura	Inhibición
Vinilsulfonas		Irreversible $IC_{50}=4\text{ nM}$ (Cheng,2008)
Aciloxi metil cetonas		Irreversible $k_2/K_i = 1.6 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$ (Brak,2008)
Alpha-dicarbonilos		Análogo del Estado de Transición $IC_{50}=0,64\text{ nM}$ (Choe,2005)
Cetonas cíclicas		Reversible competitivo $K_i=16\text{ nM}$ (Huang,2003)
Nitrilos		Reversible $IC_{50}=0.2\text{ nM}$ (Mott,2010)
Hidroximetilcetonas		Reversible competitivo $K_i=65\text{ nM}$ (Huang,2003)
Tiosemicarbazonas		Irreversible $IC_{50}=17\text{nM}$ (Fuji,2005; Siles,2006)

6.1.3.1 Enlace de Halógeno en la inhibición de Cruzipain

Un relevamiento en bases de datos públicas reveló que varios inhibidores de Cruzipain (Cz) que contenían un átomo de halógeno en su estructura (X = Cloro, Bromo o Yodo) eran, en algunos casos, muchas veces más activos que sus análogos no halogenados.

Adicionalmente, la evidencia estructural revela que los compuestos halogenados forman un Enlace de Halógeno (EX) X…S con la metionina 68 de Cz (Ferreira et.al ,2010).

La estructura de alta resolución del inhibidor B95, (N-[2-(1H-benzimidazol-2-yl)ethyl]-2-(2-bromophenoxy)acetamide) anclado en la hendidura de unión de la Cruzipain se muestra en la Figura 6.5 (código PDB: 3KKU).

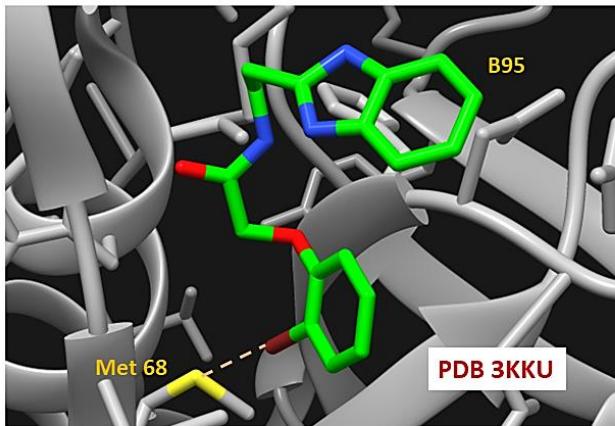


Figura 6.5: Sitio de unión de la enzima Cruzipain en la estructura del PDB: 3KKU obtenida del Protein Data Bank.

En dicha estructura, el átomo de bromo del resto bromofenóxido del inhibidor B95 se encuentra próximo al residuo Met 68 que limita, por debajo, el sub-bolsillo hidrofóbico S2 de la enzima. La distancia S…Br de 3,31 Å y el ángulo C_{ar}-Br-S 172,45° se encuentran dentro de los límites geométricos de un enlace de halógeno entre los átomos involucrados. La metionina generalmente es considerado como un compañero de interacción altamente hidrofóbico, pero debido a la anisotropía de la distribución de electrones del azufre y su polarizabilidad, es un ejemplo interesante de un aceptor de enlace de halógeno(Wilcken et al., 2013).

Sobre la base de estos antecedentes, en este capítulo se estudió la importancia del EX (Met 68)S…X (con X=Cl, Br) en la inhibición de Cruzipain, bajo la hipótesis de que esta interacción estaría detrás del aumento observado en la actividad inhibitoria de los ligandos halogenados, en comparación con sus análogos no halogenados.

La estrategia consistió en relevar bases de datos públicas de compuestos con actividad biológica, identificando inhibidores de cruzipain en los cuales la incorporación de un átomo de halógeno aumentaba su potencia inhibitoria. En particular, se prestó especial atención a aquellos casos en los cuales la actividad aumentaba en más de 100 veces, lo que se conoce como un “activity cliff” o “acantilado de actividad”. Los acantilados de actividad generalmente se definen como pares o grupos de compuestos estructuralmente similares que son activos contra el mismo objetivo, pero tienen grandes diferencias en potencia (Stumpfe et. al, 2012). Los “activity cliffs” capturan modificaciones químicas que influyen fuertemente en la actividad biológica. En nuestro caso particular, el requerimiento de un activity cliff resulta una estrategia promisoria para poder asociar fehacientemente el aumento de actividad con la presencia del halógeno y en particular, con las interacciones que este establece.

Algunas preguntas que queremos responder a través de este estudio son: ¿Cómo puede una única interacción no covalente, como un enlace de halógeno, aumentar la potencia de inhibición contra Cz tan drásticamente?, ¿Hay algún otro fenómeno involucrado en la generación de los “acantilados de actividad”?

6.2. Métodos

6.2.1 Compilación de pares LX/LH con actividad inhibitoria conocida sobre Cz

De la base de datos CHEMBL(Gaulton et al., 2012) (ebi.ac.uk/chembl/) se obtuvieron ligandos halogenados (LX, con X = Cl, Br o I) junto con sus correspondientes anotaciones de actividad contra Cz. En la misma base de datos se buscó, para cada compuesto halogenado, su correspondiente análogo no halogenado (LH/LOH) empleando herramientas quiminformáticas (OpenBabel(O’Boyle et al., 2011)). Más concretamente, se utilizó la representación de SMILES del compuesto halogenado (excluyendo el halógeno) para realizar el filtrado de la base de datos, seleccionado de esta manera pares que se diferenciaban únicamente en la posición de sustitución del halógeno (Figura 6.6).

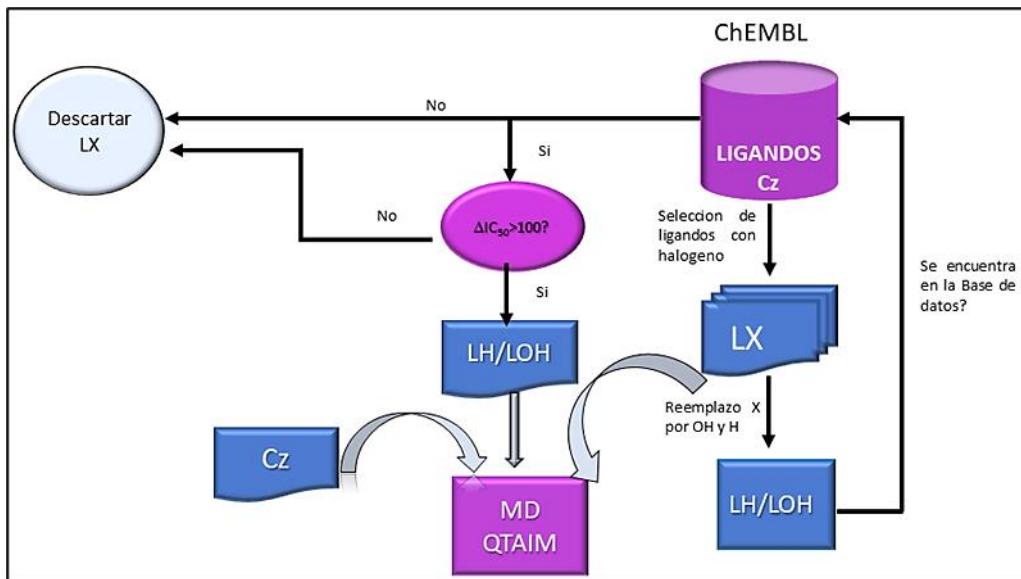


Figura6.6: Procedimiento para la selección de pares de ligandos LX/LH y LX/LOH

De los pares encontrados, se seleccionaron finalmente 5 de ellos, los cuales cumplían con la condición de que el análogo halogenado fuese al menos 100 veces más activo que su contraparte no halogenada (“activity Cliff”). En estos pares el átomo de halógeno X estaba sustituido por hidrógeno (LH) u oxhidrilo (LOH).

Los 5 grupos seleccionados se encuentran dentro de la familia de las tiosemicarbazonas (Figura 6.7A) las cuales se unen covalentemente al residuo de Cisteína 25 del sitio activo de la enzima y han demostrado ser inhibidores irreversibles de la Cruzipain (Du et al., 2002; Siles et al., 2006).

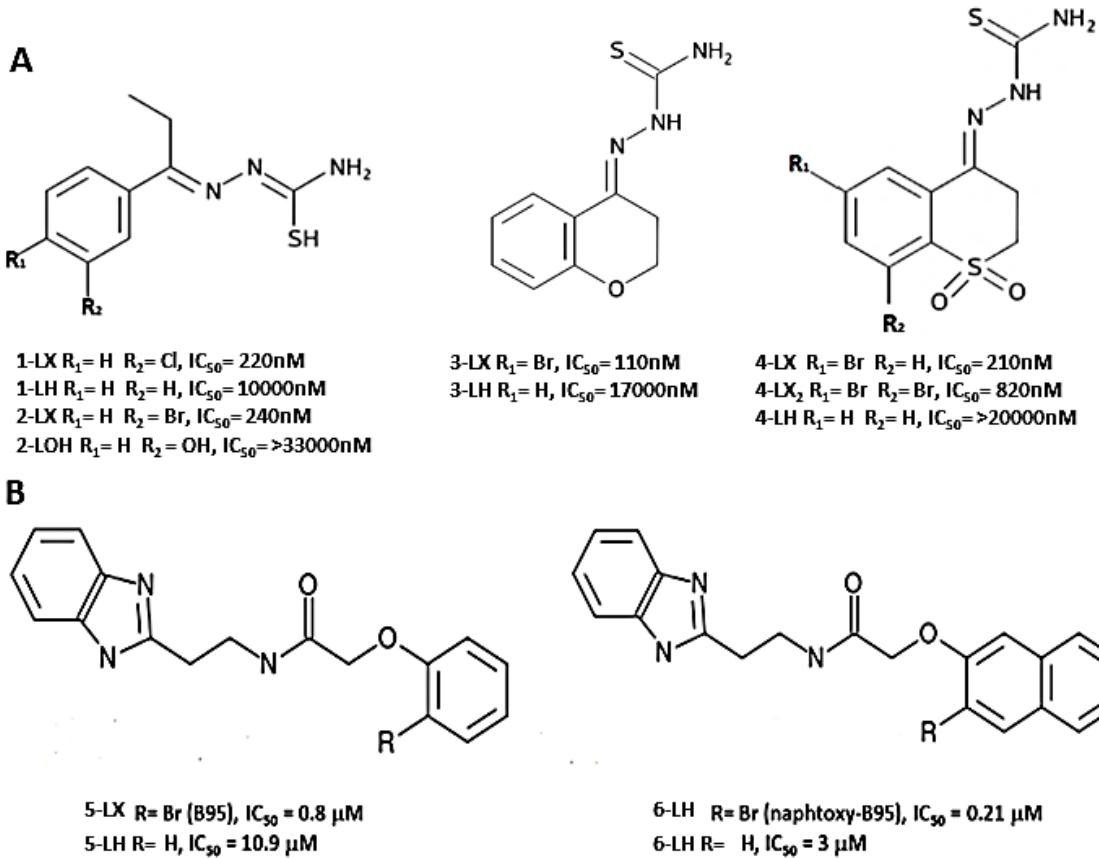


Figura 6.7: Estructura de los inhibidores covalentes (A) y no covalentes (B) de Cruzipaina seleccionados. Para cada andamiaje estructural se indica el sitio y tipo de sustitución para formar los pares LX/LH o LX/LOH junto con las correspondientes anotaciones de actividad (IC₅₀). En paréntesis se indica la numeración asignada a cada par, la cual será utilizada de aquí en adelante para referenciarlos.

Para estudiar también el efecto de la halogenación sobre inhibidores reversibles de Cz, se consideraron dos pares LX/LH adicionales de inhibidores de Cz (Figura 6.7B), que se unen no covalentemente a la enzima y que presentan un núcleo benzimidazol como característica estructural en común (Ferreira et al., 2014). Si bien el análogo halogenado es más activo en ambos casos, es importante aclarar que la diferencia de actividad no es tan significativa como para ser considerado un “activity cliff”, como era el caso de los derivados de tiosemicarbazonas. En el caso del inhibidor B95 (5-LX) previamente co-cristalizado con Cz, éste es 14 veces más activo que su análogo no halogenado 5-LH mientras que 6-LX es 15 veces más activo que 6-LH.

6.2.2 Estructuras iniciales de los complejos

Las estructuras de partida para las simulaciones de las tiosemicarbazonas ancladas en la hendidura de unión de Cz, fueron obtenidas mediante docking covalente. Para ello se utilizó la funcionalidad sdtether de rDock (Rachman et al., 2019) que permite restringir el movimiento de la cabeza de guerra (warhead) cerca del residuo reactivo Cys 25 para simular la unión covalente.

En el caso de los inhibidores no covalentes derivados de benzimidazol, uno de ellos (B95) ya fue co-cristalizado con Cz (codigo PDB: 3KKU), por lo tanto se consideró ésta estructura como referencia para construir a partir de ellas las estructuras de los restantes complejos.

6.2.3 Preparación de las estructuras iniciales para las simulaciones de Dinámica Molecular

a) Tratamiento del enlace covalente: Empleando las funcionalidades incluidas en el paquete AmberTools (antechamber, etc) se construyó un residuo no estándar que incluía a Cys25 unida covalentemente a la tiosemicarbazona. Las cargas atómicas del residuo no estándar fueron calculadas usando el método del potencial electrostático restringido (RESP).

b) Tratamiento del enlace de halógeno: Para simular el agujero- σ sobre el átomo de halógeno, se introdujo un extra-punto (EP) con carga positiva y sin masa en el campo de fuerza de Amber (Ibrahim, 2011). La parametrización del EP se llevó a cabo siguiendo el procedimiento del mismo autor, donde la posición óptima del EP en el átomo de halógeno se exploró estudiando la correlación entre la distancia X - EP y la longitud del enlace halógeno X...Y, y se investigó el efecto de la inclusión de EP en la precisión de las cargas derivadas de potencial electrostático. Se ha demostrado que sin la introducción del EP la distancia del átomo de halógeno excede a la suma de los radios de Van der Waals del halógeno y aceptor, valor límite para la formación del enlace. Este fenómeno se explica por la repulsión entre los átomos de halógeno y el aceptor, ambos con una carga negativa. La parametrización del EP se llevó aplicando el procedimiento descripto en el capítulo 3 (ver Figura 6.8). Brevemente, el procedimiento consiste en los siguientes pasos:

- ✚ Se construyeron o se obtuvieron de bases de datos las estructuras 3D de los ligandos, a las cuales se las debe protonar y optimizar.
- ✚ Posteriormente, un átomo sin masa EP se incluyó en el campo de fuerza del ligando. Se estableció el ángulo C(ar)-X-EP en 180° y la distancia X-EP se igualó al radio atómico del halógeno, es decir, 2.22 Å para el Bromo 2,35 Å para Yodo y 1.95 Å para el Cloro.

- El potencial electrostático para las moléculas halogenadas en estudio se generó al nivel HF/6-31G(d). Las cargas parciales atómicas (incluyendo las del EP) fueron asignadas mediante el método del potencial electrostático restringido (RESP).
- La estructuras cristalinas de los blanco molecular, Cruzipaína (PDBid:3KKU), también debe ser preparada y se protona adecuadamente.
- Una vez listas las estructuras del ligando y la proteína se procesan con el software tleap de AmberTools, para formar el complejo y obtener los archivos de topología y coordenadas, con los cuales se realizarán las simulaciones de Dinámica Molecular.
- Finalmente, las simulaciones se realizaron con el software AMBER, siguiendo los pasos habituales de minimización, equilibración y producción de la dinámica.

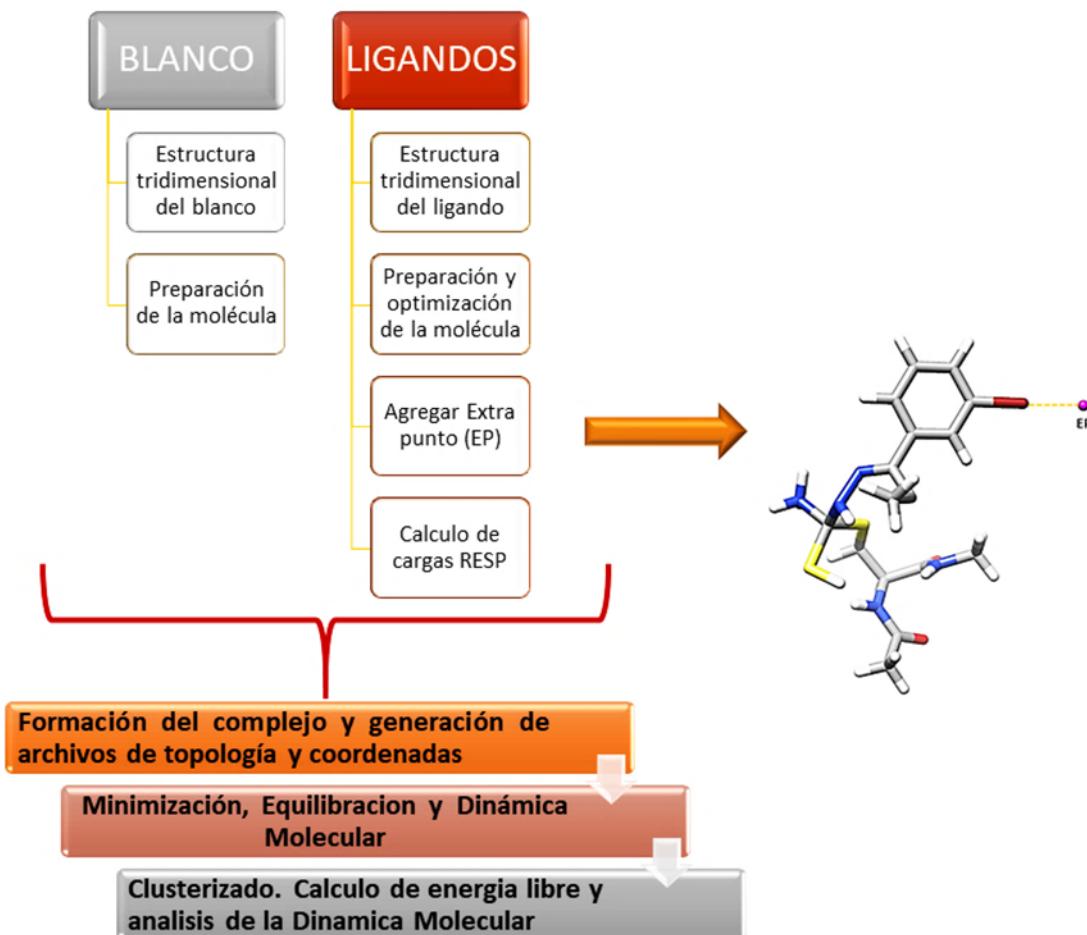


Figura 6.8: Esquema de los pasos seguidos para la parametrización del EX en las simulaciones de Dinámica Molecular. A la derecha se puede ver un ejemplo de la estructura de un ligando al cual se le ha agregado el EP que simula el agujero sigma.

6.2.4 Simulaciones de Dinámica Molecular parametrizadas para EXs

Las simulaciones de DM de los complejos previamente preparados se realizaron utilizando el paquete de software de AMBER (Case et al., 2005) a 300K de temperatura y extendida a 100 nanosegundos de tiempo global de simulación.

Adicionalmente se realizaron simulaciones largas de 1 microsegundo y por triplicado de la forma libre (apo) de la enzima.

6.2.4.1 Análisis de las trayectorias de DM

Para el análisis de las trayectorias se utilizó el módulo CPPTRAJ del paquete AmberTools. Se calculó la desviación cuadrática media (RMSD) para verificar que se alcanzó un estado de equilibrio del sistema molecular. Las interacciones moleculares se analizaron mediante cálculos de distancia y tiempo de vida media de los enlaces de hidrógeno a lo largo de las simulaciones.

Normalización de las distancias de interacción

Como se mencionó previamente los enlaces de halógeno comparten características con los puentes de hidrógeno, como que ser fuertemente direccionales, con un ángulo A—XY óptimo de 180°, y una distancia X—Y idealmente menor que la suma de sus radios van der Waals (Tabla 6.2) . Así, en este caso el ángulo formado por el carbono del anillo aromático, el halógeno y el S de la Met68 debería cumplir este requisito para establecer la existencia de un puente por halógeno. También se considera importante la distancia entre el halógeno y el azufre, un parámetro útil para esta evaluación es el radio de van der Waals que indica la distancia mínima a la cual se puede acercar un átomo que no forma un enlace con otro. Si sumamos los radios de van der Waals del átomo sustituyente y del O se obtiene la distancia máxima a la que se podría identificar una interacción relevante. De esta forma, si las sumas de los radios van der Waals de dos átomos es mayor que la distancia calculada entre ambos, podría deducirse que existe una interacción entre ellos.

Tabla 6.2: Valores de radios van der Waals (Å) para los átomos involucrados en este estudio (Bondi, 1964)

Cl	Br	H	S
1,75	1,87	1,06	1,80

$$\Sigma r \text{ vdw Cl-S} = 3,55$$

$$\Sigma r \text{ vdw Br-S} = 3,67$$

$$\Sigma r \text{ vdw H-S} = 2,86$$

Para poder hacer una comparación correcta entre las distancias de enlace se necesita normalizar los valores entre las moléculas estudiadas, ya que a medida que disminuye la densidad electrónica del sustituyente es necesaria una menor distancia de enlace en comparación con la que necesita un átomo de alta densidad electrónica como el yodo. De esta forma, una distancia que implica una interacción fuerte para un átomo como yodo puede significar una nula interacción para otro de menor o nulo agujero como el flúor. Un parámetro que sirve para normalizar los resultados y hacerlos comparables entre si es el radio de van der Waals. Se divide cada valor de distancia por el valor de la suma de radios de van der Waals de los átomos involucrados. Así, el límite de interacción sería 1. Cuando el factor obtenido es mayor que 1 esto implica que hay una débil o nula interacción, puesto que la distancia calculada es mayor que el radio de van der Waals. Si por el contrario este valor es menor que 1 esto indica que hay una interacción considerable entre los átomos involucrados.

$$\text{Distancia normalizada} = \text{Distancia} (\text{\AA}) / \Sigma r \text{ vdw} (\text{\AA})$$

6.2.4.2 Muestreo de estructuras a partir de las trayectorias

La Cruzipaína es un blanco molecular muy esquivo y difícil de entender, por lo tanto, la construcción de un modelo estructural que permita encontrar relaciones de estructura-actividad significativas, resulta una tarea difícil. Un modelo sencillo basado en una única estructura de la enzima tomada de las trayectorias de DM, posiblemente no logre recuperar todos los aspectos estructurales que impulsan el acantilado de actividad en los pares LX/LH y LX/LOH.

- ⊕ Por lo tanto, se muestrearon y analizaron diferentes estructuras a partir de las simulaciones MD largas, como ser:
 - ⊕ Estructura del mínimo de energía potencial. Se calculó la energía potencial del sistema a lo largo de la trayectoria y de todos los "frames" se seleccionó la estructura con mínima energía.
 - ⊕ Estructuras representativas del clusterizado de la trayectoria. Se realizó un análisis de agrupamiento de las trayectorias según el RMSD del backbone de la proteína. Se empleó el algoritmo K-means (K=10) implementado en CPPTRAJ. Se seleccionó una estructura representativa para cada uno de los 10 clusters para su posterior análisis.

- Proyecciones extremas a lo largo de la Componente Principal 1 (PC1). Se realizó un análisis de componentes principales de las simulaciones y se proyectó la trayectoria a lo largo de la primera componente PC1, seleccionando las estructuras correspondientes a los valores extremos de dicha componente (máximo y mínimo).

6.2.5 Análisis de interacciones intermoleculares

Se siguieron los pasos a continuación:

1. Selección de estructuras representativas: El estudio de las interacciones se realizó a partir de la estructura del mínimo de energía potencial de trayectoria de DM.
2. Se construyeron modelos reducidos, conservando sólo aquellos residuos de la Cruzipain que interactúan directamente con el ligando dentro de un rango de 5 Å. Incluyendo los residuos Gln-19, Cys-22, Gly-23, Ser-24, Cys-25, Trp-26, Ala- 27, Phe-28, Ser-29, Ala-30, Gly-65, Gly-66, Leu- 67, Met-68, Asn-69, Ala-136, Val-137, Ala-138,-138, Val-139, Leu-160, Asp-161, His-162, Gly-163, Val-164, Ser-183, Trp-184.
3. Análisis QTAIM: La densidad de carga electrónica (DE o ρ) de los modelos reducidos fue calculada con el funcional M06-2x y el conjunto de bases 6-31G (d), incluido en el paquete Gaussian 09. Se aplicó la Teoría Cuántica de Átomos en Moléculas QTAIM (Bader, 1990) para obtener los grafos moleculares de la DE, sobre los cuales se realizó el análisis en detalle de las interacciones moleculares. En breve, el procedimiento consiste en “mapear” el campo del vector gradiente ∇ , sobre la DE previamente calculada ($\nabla\rho$), para así exponer los elementos topológicos de la DE, los puntos críticos de enlace (PCE) y caminos de enlace (CE) asociados a las interacciones interatómicas. El análisis topológico de densidad de carga se realizó con el software Multiwfn(Lu & Chen, 2012).

6.3 Resultados:

6.3.1 Estabilidad de los enlaces S...X / S...H

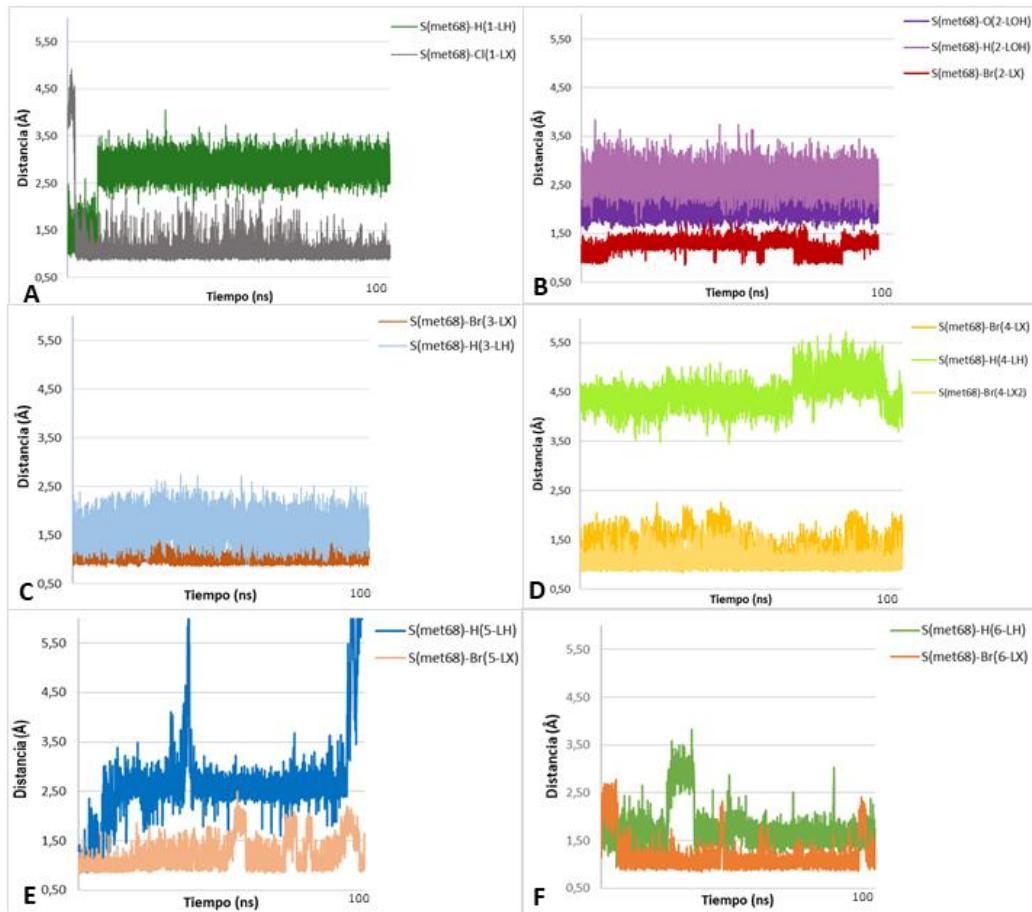


Figura 6.9: Gráficos de distancias normalizadas del S(met68) de la enzima al halógeno o hidrógeno de los pares de ligandos halogenados y sus análogos respectivamente, en función de la trayectoria. A) 1-LX/LH, B)2-LX/LOH, C)3-LX/LH, D)4-LX/LX₂/LH, E)5-LX/LH, F)6-LX/LH

La Figura 6.9 muestra las distancias normalizadas entre el átomo de S de Metionina 68 y el halógeno (o hidrógeno del correspondiente análogo no halogenado), a lo largo de las simulaciones de los pares LX/LH(LOH). Se puede observar que en los complejos halogenados los valores se mantienen próximos a 1 mientras que en los análogos no halogenados este valor fluctúa por encima de 1; demostrando que las interacciones de estos son más débiles, los compuestos halogenados interaccionan permanentemente con la Met68 a diferencia de su contraparte LH, el cual se separa del bolsillo de unión de la Cz.

6.3.2 Muestreo de la trayectoria

Diversos estudios se han enfocado en tratar de definir el bolsillo de unión de Cruzipaina y los diferentes sub-bolsillos que conforman el sitio catalítico (Bryant et al., 2009; Durrant et al., 2010; Turk et al., 1998), lográndose caracterizar 4 sub-bolsillos (S1-S4) bien definidos, a los que se agregan otros (S1', - S3') dependiendo del ajuste conformacional adoptado por Cz inducido por diferentes ligandos (Luchi et al., 2019; Sartori et al., 2019). Debido a que la densidad de carga, medida en el punto crítico de interacción, es una propiedad topológica local, se puede calcular la contribución de un subconjunto molecular, formado por uno o varios residuos, a la fuerza de anclaje total del inhibidor. De esta manera, se podría conocer cual/es subbolsillo/s de la enzima se revelan como sitios adecuados para mejorar las interacciones con el inhibidor. La evolución de la distancia S···X/S···H a lo largo de las simulaciones muestra una posible correlación entre la actividad inhibitoria y la estabilidad del anclaje del halógeno / hidrógeno en el sub-bolsillo S2. Para poder cuantificar dicho anclaje, se necesita estimar la fortaleza de las interacciones moleculares individuales que establece el halógeno con los distintos átomos del entorno.

Desafortunadamente, considerando el tamaño de los sistemas biológicos y el estado de arte de la química computacional, no es posible calcular la densidad electrónica del sistema completo. Resulta necesario entonces modelar el sistema con un sistema modelo reducido que incluya al inhibidor y a los residuos proteicos circundantes, construido sobre una única (o unas pocas) estructuras (frames) muestreadas a partir de las simulaciones.

Por lo tanto, la selección de los residuos a incluir en el modelo reducido y la “muestra” de estructuras tomadas a partir de las simulaciones deben ser lo suficientemente representativas del sistema real, para así lograr capturar los procesos fisicoquímicos claves, detrás de la respuesta biológica observada.

Para “muestrear” las trayectorias de manera más informada, es importante primero tener una perspectiva de la variabilidad conformacional total de la Cruzipaina. Para ello investigamos su espacio conformacional usando análisis de componentes principales (PCA).

6.3.2.1 Análisis de Componentes Principales en apo Cz

Análisis de componentes principales (PCA) sobre la trayectoria de dinámica molecular de la forma apo (libre) de la enzima. La trayectoria se extendió hasta alcanzar los 500 ns de simulación, con la intención de capturar los cambios conformacionales relevantes de la enzima.

Los resultados del PCA muestran que 29% de la variabilidad conformacional total está contenida en las primeras 3 componentes principales (PC1, PC2 y PC3), de las cuales PC1 da cuenta de más de la mitad (15%) mientras que el resto se distribuye entre PC2 (7%) y PC3 (6%). La contribución de las restantes componentes a la varianza es aún menor y se distribuye de manera casi equitativa entre ellas, con lo cual su aporte individual a la variabilidad conformacional resulta despreciable.

El histograma de la figura 6.10 muestra la proyección de la trayectoria sobre las componentes PC1, PC2 y PC3.

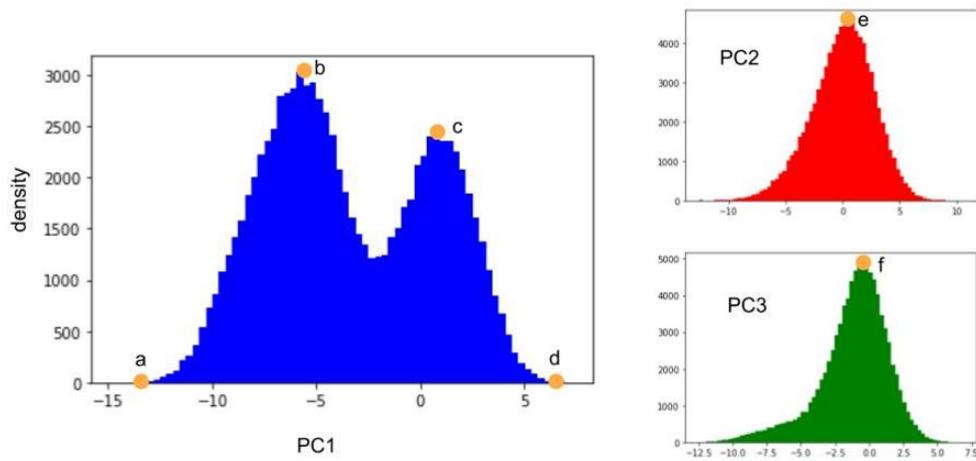


Figura 6.10: Proyección de la trayectoria de la forma apo (libre) de Cruzipaina sobre las componentes principales 1, 2 y 3.

PC1 muestra una distribución bimodal, con un máximo global b y un máximo local c, mientras que PC2 y PC3 presentan un único máximo.

Las estructuras correspondientes a los valores extremos en dichas distribuciones pueden dar una idea de la variabilidad conformacional total de la enzima. La figura 6.10 muestra ambas estructuras extremas en la distribución de PC1 superpuestas (estructuras en celeste y gris, correspondientes a los puntos extremos a y d en la Figura 6.11, respectivamente).

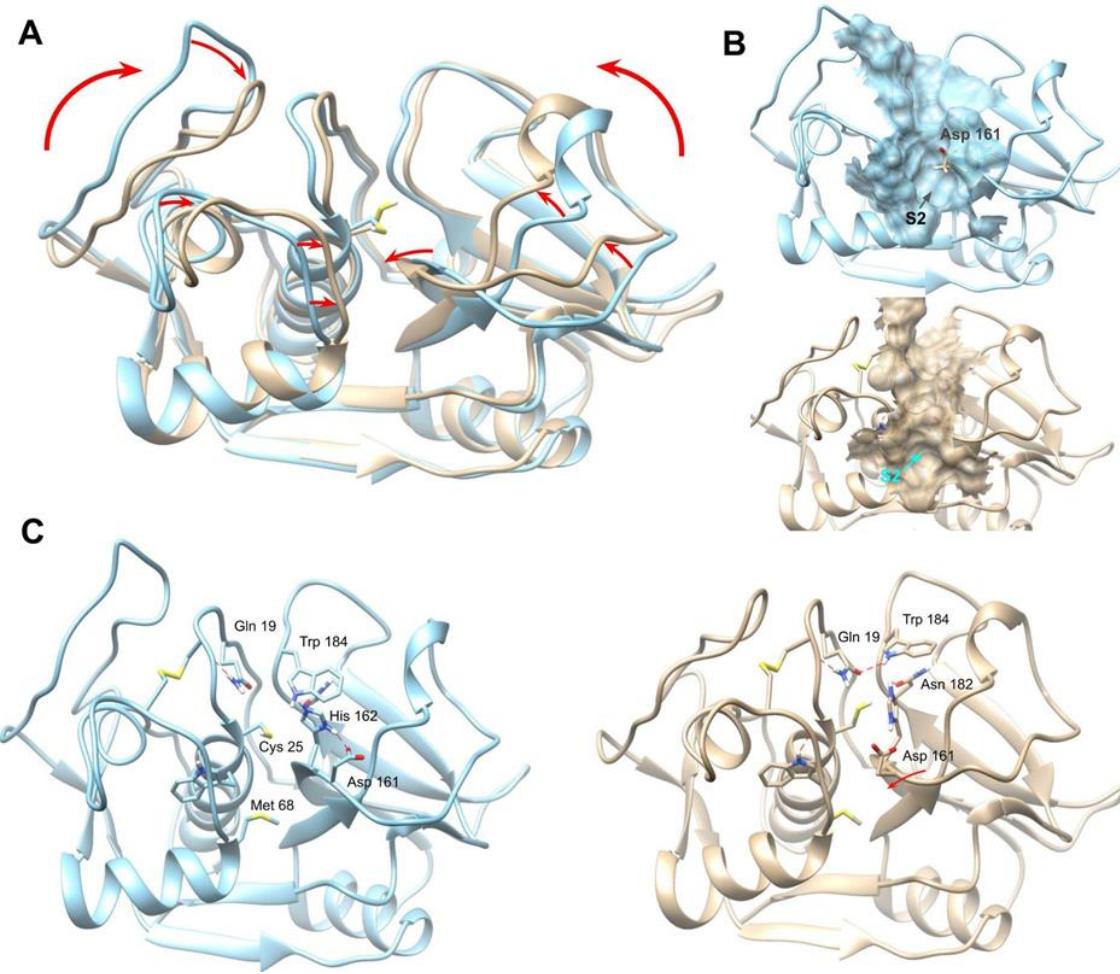


Figura 6.11: Conformaciones extremas de apo Cz en la distribución de PC1. En celeste y gris se muestran las estructuras correspondientes a los puntos extremos a y d en la figura respectivamente.

Las estructuras superpuestas muestran que ambos dominios de la proteína, el dominio de hélices α (izquierdo) y el de láminas β (derecho) experimentan movimientos concertados en dirección opuesta que tienden a cerrar/estrechar la hendidura de unión del sustrato que se encuentra en el medio, entre ambos dominios. Así, podemos asociar a las estructuras en celeste y gris con las conformaciones abierta y cerrada de la enzima, respectivamente.

La superficie accesible de la hendidura de unión correspondiente a ambas conformaciones extremas se muestra en la figura 6.11B. Se observa que en la conformación abierta la hendidura presenta una superficie más amplia y más profunda que en la conformación cerrada, en particular sobre el sub-bolsillo S2. Al superponer ambas estructuras se observa cómo el residuo Asp 161 de la conformación cerrada sobresale dentro del surco de unión de la conformación abierta (Figura 6.11B), lo que explica en parte el estrechamiento de la hendidura en la primera.

La Figura 6.11C muestra algunos residuos del surco de unión que experimentan cambios conformacionales al pasar de la forma abierta a la cerrada. El residuo His 162 parece jugar un papel importante en estos cambios. En la forma abierta, el residuo de His 162 no está formando el característico puente de hidrógeno con Asn 182 que facilita la abstracción del protón de Cys 25 a través de la formación de un relé de carga, (Asn 182)C=O···H-N(His 162)N:···H-S(Cys 25). Esto se debe a que en esta conformación, Trp 184 se interpone entre ambos residuos evitando la formación del relé. En cambio, en la conformación abierta His 162 forma una interacción fuerte con el residuo vecino Asp 161. Al pasar a la conformación cerrada, Trp 184 se desplaza y permite la formación del relé entre Asn 182 e His 162. Al mismo tiempo, se debilita la interacción de His 162 con Asp 161 y la cadena β que contiene ambos residuos se tuerce hacia la hendidura de unión, contribuyendo a su estrechamiento.

Como se dijo más arriba, las estructuras extremas a lo largo de la PC1 dan una idea de la variabilidad conformacional global de la proteína. Sin embargo, éstas no son representativas de la trayectoria dado que presentan una densidad poblacional muy baja (puntos de mínimo a y d en PC1, Figura 6.10).

Por otro lado, los máximos en la distribución de PC1, PC2 y PC3 (puntos b, c, e y f en la Figura 6.10) corresponden a los valores de dichas componentes que son más frecuentemente visitados por Cz a lo largo de la simulación. Por lo tanto, el muestreo de la trayectoria en la proximidad de estos máximos posiblemente recupere conformaciones de la enzima que sean más representativas de la trayectoria.

El análisis de las estructuras de apo Cz muestreadas en la proximidad de los máximos global y local en la distribución de PC1, reveló que estos comparten características comunes con las estructuras abierta y cerrada de la enzima, respectivamente (no mostrado). Si bien las diferencias no son tan marcadas como entre las estructuras correspondientes a los extremos, podemos decir que las conformaciones abierta y cerrada son estructuras representativas de apo Cz. Por otra parte, la densidad relativa de ambas conformaciones de apo Cz indica que la forma abierta sería la conformación más probable de la enzima libre, por tratarse del máximo global en la distribución de PC1. En consecuencia, podemos especular que la conformación abierta de la enzima sería la forma más probable de interacción inicial con el sustrato. Sin embargo, como esta forma no es catalíticamente competente, debido a que no forma el relé de carga, debe producirse el cambio conformacional a la forma cerrada para que ocurra finalmente la proteólisis.

6.3.2.2 Análisis de Componentes Principales de los pares Cz/LX y Cz/LH(LOH)

En un trabajo previo, realizado en este mismo grupo de investigación, se encontró que inhibidores de Cz del tipo vinil sulfona “seleccionaban” una de dos posibles conformaciones alternativas de la enzima (Luchi et al., 2019). Los inhibidores más potentes ($K_i < 170$ nM) estabilizaban una estructura que se parecía a la conformación cerrada o catalíticamente competente, previamente descripta para apo Cz. En cambio los inhibidores menos potentes estabilizaban una estructura más similar a la forma abierta de apo Cz. Estos hallazgos sugieren que los inhibidores, al unirse a Cz, actuarían modificando el equilibrio de conformaciones pre-existentes de la enzima libre. Para investigar esta hipótesis, se procedió a proyectar las trayectorias de los complejos Cz/inhibidor sobre el subespacio de componentes principales de la forma apo de la enzima y a observar hacia donde se desplaza este equilibrio.

La figura 6.12 muestra la distribución de las trayectorias de los complejos Cz/2-LX y Cz/2-LOH, proyectadas sobre la PC1 de apo Cz. Ambas se encuentran desplazadas hacia el máximo local en la PC1 de apo Cz, indicando que la estructura adoptada por la proteína en ambos complejos compartiría características con la forma cerrada de apo Cz.

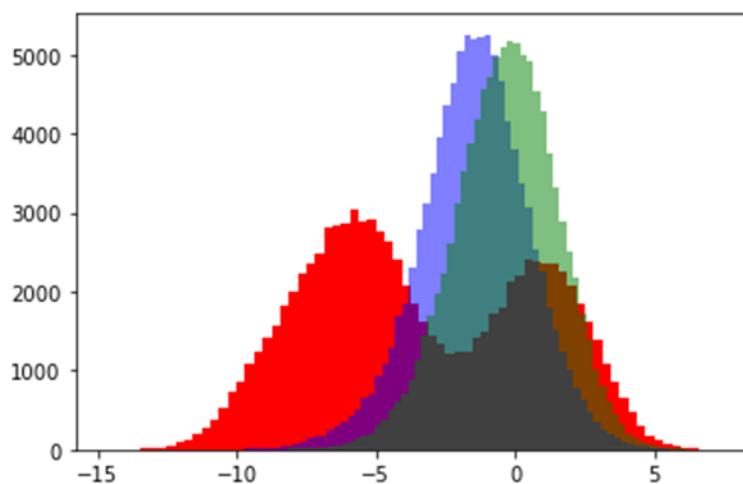


Figura 6.12: Proyección de las trayectorias de los complejos de Cz con 2-LX (azul) y 2-LOH (verde) sobre la PC1 de apo Cz (rojo).

En la Fig 6.13 se muestran las estructuras superpuestas correspondientes a los complejos Cz/2-LX y Cz/2-LOH seleccionados alrededor del máximo en la distribución de PC1 (Figura 6.12), y simultáneamente alrededor del máximo en las distribuciones de PC2 y PC3.

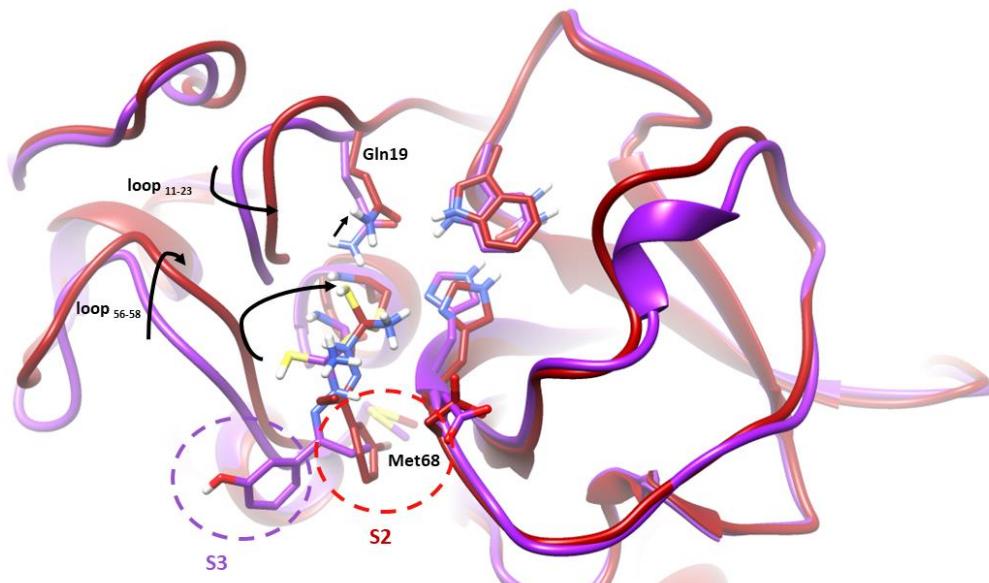


Figura 6.13: Superposición de estructuras representativas de los complejos Cz/2-LX (rojo oscuro) y Cz/2-LOH (violeta). Las estructuras fueron obtenidas mediante muestreo de las trayectorias (previamente proyectadas en el subespacio de componentes principales de apo Cz) alrededor de los valores máximos en las proyecciones de las primeras 3 componentes principales.

La inspección de las estructuras superpuestas reveló que, en efecto, estas presentan algunas características estructurales que recuerdan a la conformación cerrada de apo Cz, en particular la formación del relé de carga entre His 162 y Asn 182 en ambos complejos.

Las mayores diferencias conformacionales entre Cz/2-LX y Cz/2-LOH se observan en el dominio de hélices α , más precisamente en los loops 11-23 y 56-68, mientras que el dominio de láminas β prácticamente no experimenta cambios en el backbone (Figura 6.13).

Al pasar del complejo Cz/2-LOH al Cz/2-LX, el loop 56-68 se desplaza hacia la derecha/atrás, estrechando la hendidura de unión (Figura 6.13). A su vez, este desplazamiento del loop 56-68 es acompañado por el movimiento del loop 11-23 en el mismo sentido. El resultado es el desplazamiento *hacia atrás* de la cadena lateral de Gln 19 que a su vez impulsa el reordenamiento del inhibidor para llenar el espacio que dejó vacante Gln 19 en el sub-bolsillo S1'.

Finalmente, para encontrar el origen de las diferencias conformacionales descritas entre los complejos Cz/2-LOH y Cz/2-LX, resulta necesario analizar la interacción enzima/inhibidor, en las inmediaciones al sitio de sustitución. Nótese en la Figura 6.13 que, mientras el anillo hidroxi fenóxido de 2-LOH se ancla en el sub-bolsillo S3 más expuesto al solvente, el bromo fenóxido de 2-LX se *entierra* en el fondo del sub-bolsillo S2 más profundo y más hidrofóbico que el anterior. Nuestra hipótesis es que el anclaje del resto hidroxi fenóxido en S3 bloquearía el desplazamiento del loop 56-68 y los concomitantes cambios conformacionales que impulsan la retracción de Gln 19 y el reacomodamiento del inhibidor. Por el contrario, debido a la naturaleza más hidrofóbica del anillo de bromo fenóxido, éste tiende a *esconderse* en el sub-bolsillo S2, dejando liberado al loop 56-68.

Los cambios conformacionales que llevan a la retracción de Gln 19 y al consiguiente reordenamiento del inhibidor ya fueron descritos previamente en complejos de Cz con inhibidores de tipo vinil sulfonas (Luchi et al., 2019). En dicho estudio se encontró que solo las vinil sulfonas más potentes eran capaces de inducir estos cambios sobre la enzima. Es lógico suponer entonces, teniendo en cuenta este antecedente, que la capacidad del análogo halogenado 2-LX de inducir un “activity cliff” tiene que ver con su habilidad para impulsar estos cambios conformacionales críticos.

6.3.3 Análisis basado en la densidad electrónica de los pares Cz/LX y Cz/LH(LOH)

A partir de las estructuras representativas seleccionadas de las simulaciones de los complejos Cz/2-LX y Cz/2-LOH se realizó un análisis de la distribución de la densidad electrónica empleando la teoría QTAIM.

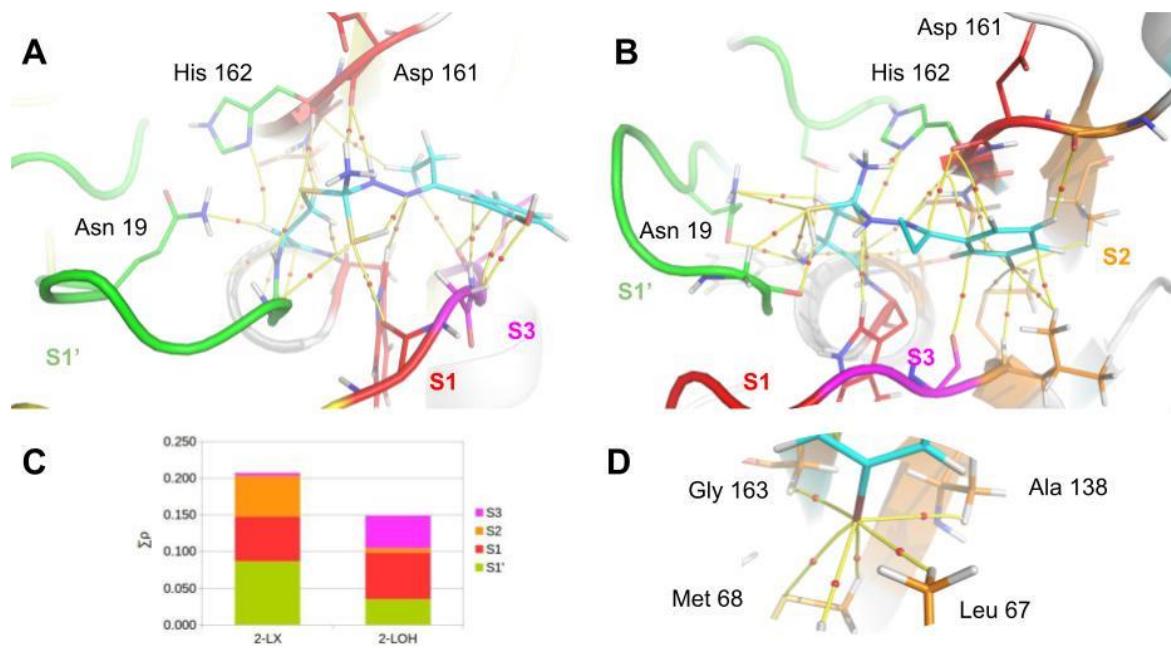


Figura 6.14. Análisis basado en la densidad electrónica de los complejos Cz/2-LOH y Cz/2-LX. En A y B se muestran los grafos moleculares de Cz/2-LOH y Cz/2-LX, respectivamente. Los residuos correspondientes a los sub-bolsillos S1', S1, S2 y S3 se muestran en verde, rojo, naranja y magenta, respectivamente. El gráfico de barras apiladas C muestra la contribución de cada sub-bolsillo al anclaje total del inhibidor. En D se muestran más claramente las interacciones del átomo de bromo.

Los grafos moleculares de los complejos Cz/2-LOH (Figura 6.14 A) y Cz/2-LX (Figura 6.14B) muestran los caminos de enlace (en amarillo) que anclan el inhibidor al surco de unión de la enzima. Cada una de esas interacciones presentan un punto crítico enlace (PCE) asociado (esfera roja) y el valor de la densidad electrónica (ρ) medido en dicho punto da cuenta de la fortaleza de la interacción.

De la misma manera, el valor de la sumatoria de ρ en todos los PCE intermoleculares ($\sum\rho$) puede considerarse como una medida del anclaje global del inhibidor. La Figura 6.14C muestra que el anclaje de 2-LX es alrededor de un 25% más fuerte que el de 2-LOH.

Al analizar la contribución de cada sub-bolsillo, la mayor diferencia entre ambos radica en los sub-bolsillos S2 y S3. El anillo de bromo fenóxido de 2-LX forma interacciones con las cadenas laterales de los residuos del sub-bolsillo S2 (Figura 6.14B), mientras que el anillo hidroxí fenóxido de 2-LOH está conectado por caminos de enlace al backbone del loop 56-68, en el sub-bolsillo S3 (Figura 6.14A). La fortaleza de unión de 2-LX en S2 es prácticamente compensada por el anclaje de 2-LOH en S3 (Figura 6.14C). Es decir que no puede atribuirse la existencia del “activity cliff” a la formación de un enlace de halógeno excepcionalmente fuerte. Por lo tanto,

para poder explicar el mayor anclaje global del análogo halogenado se debe analizar el aporte de los restantes sub-bolsillos. En particular, nótese que la contribución de S1' al anclaje global es casi 2 veces mayor en el análogo halogenado. Como se comentó en la sección previa, los cambios conformacionales que llevan a la retracción de Gln 19 impulsan el desplazamiento del análogo halogenado hacia S1' con la consiguiente formación de nuevas interacciones estabilizantes. En particular destaca la formación de un enlace de hidrógeno fuerte de tipo $\text{NH}_3^+ \cdots \text{N}$ con el residuo His 162 (Figura 6.17B), con un valor de $\rho_b = 0.0257$ ua., que explica la mayor parte del aumento del anclaje de 2-LX en el sub-bolsillo S1'.

Si bien el análisis se realizó para el par 2-LX/2-LOH, estos resultados pueden extrapolarse a los restantes pares de tiosemicarbazonas. La Figura 6.15 muestra que la contribución del halógeno al anclaje total en los análogos halogenados de tiosemicarbazonas no es despreciable, pero tampoco es tan significativa como para poder explicar el “activity Cliff”. Nuevamente, la fuerza impulsora del acantilado de actividad no es el enlace de halógeno en sí mismo, sino a los cambios conformacionales que se producen a consecuencia del enlace de halógeno y que dan lugar a otras interacciones electrostáticas, que permiten un anclaje global más fuerte del inhibidor a la Cz.

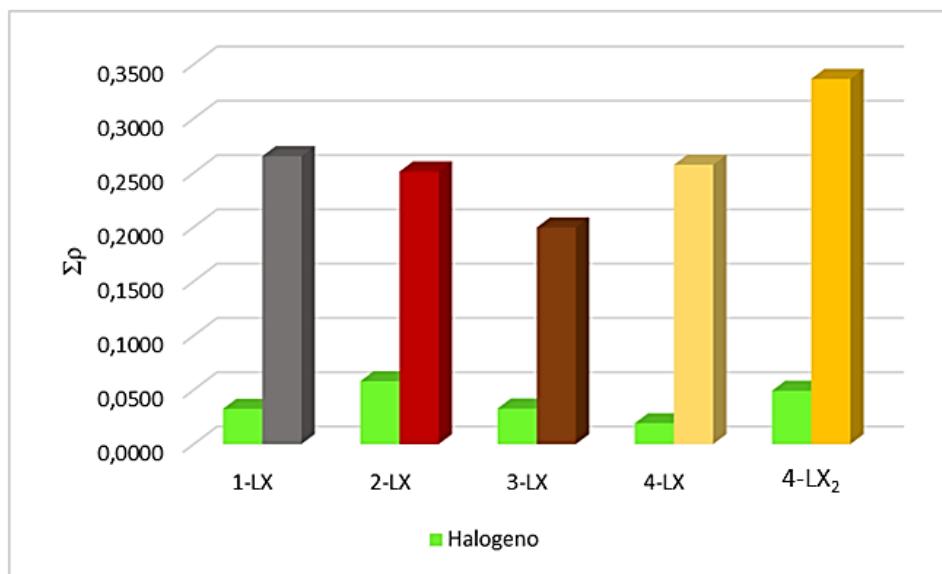


Figura 6.15: Valores de densidad electrónica total (ρ) (u.a.) de los puntos críticos de enlace (BCP) de los ligandos halogenados, en verde se puede observar la contribución del átomo de halógeno (X).

Resultados similares se observan al analizar los pares de inhibidores no covalentes 5-LX/LH y 6-LX/LH. Como se puede observar en la Figura 6.16 el valor de densidad electrónica total en los complejos con compuestos halogenados es mayor que el de su respectivo análogo, aunque la diferencia es menor que en los inhibidores de tiosemicarbazonas, lo cual tiene sentido porque la diferencia de potencia es mucho menor en los primeros (ver Figura 6.7).

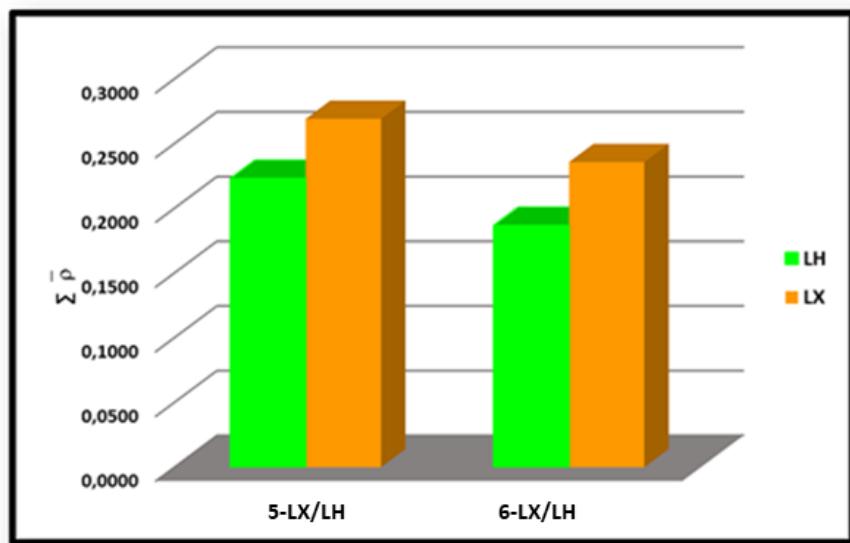


Figura 6.16: suma de los valores de densidad electrónica total de cada complejo

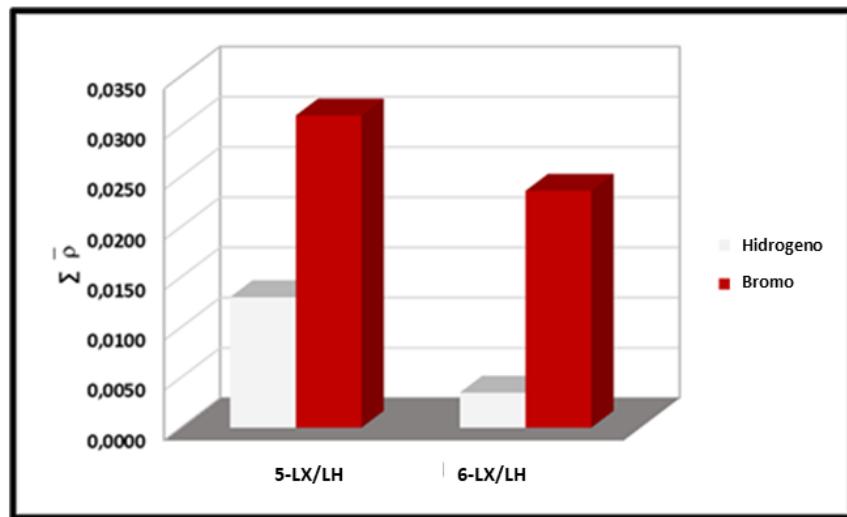


Figura 6.17: suma de los valores de densidad electrónica correspondiente a las interacciones de los átomos sustituidos.

Al analizar la contribución del halógeno (Figura 6.17) se observa que éste contribuye a aumentar el anclaje del inhibidor, en comparación con su análogo no halogenado, y en este caso, el aporte del halógeno podría explicar por sí mismo la diferencia de afinidad.

En la figura 6.18 se representan los grafos moleculares de los complejos 5-LX/5-LH respectivamente, se puede observar la elevada cantidad de caminos de enlace que forma el bromo en comparación al hidrógeno en el compuesto no halogenado, así también se puede corroborar que solo los compuestos halogenados interaccionan con la Met 68. Pero más allá de estas diferencias locales en el sitio de sustitución, no se observan cambios importantes en el modo de unión global, tanto el anillo de Br-fenóxido, en 5-LX, como el H-fenóxido, en 5-LH, se anclan en el sub-bolsillo S2.

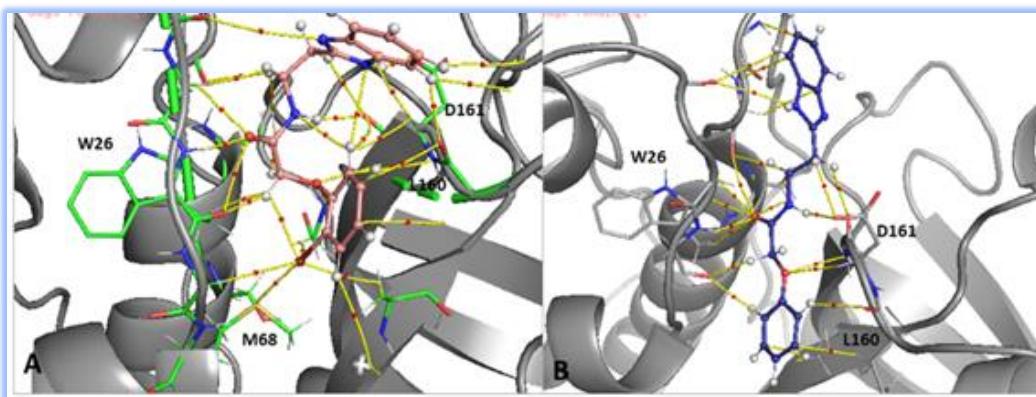


Figura 6.18: Grafos moleculares de las estructuras representativas de la DM de los complejos 5-LX(A) y LH (B). En líneas amarillas se observan los caminos de enlace y las esferas rojas representan los puntos críticos de enlace.

6.3.3.1 Rol del enlace de halógeno en la inducción del “activity cliff”

La figura 6.14 muestra en mayor detalle las interacciones del átomo de bromo de 2-LX con los residuos del sub-bolsillo S2 de Cz. En particular, nótese el camino de enlace que une el Br con el átomo de S de Met 68 que indica la formación de un enlace de halógeno putativo. Sin embargo, aunque el ángulo de interacción ($C-Br\cdots S$) = 155.8° se encuentra dentro de los límites de un enlace de halógeno, la distancia $d(Br\cdots S)$ = 4.00 \AA es superior a la suma de los radios de van der Walls de Br y S, aunque esta distancia desciende por debajo de este límite, en ocasiones, a lo largo de las simulaciones. Además del enlace de halógeno putativo, el Br forma otras 5 interacciones de enlace de hidrógeno $C-H\cdots Br$ con cadenas laterales hidrofóbicas del sub-bolsillo S2 (Fig x5D). Estos enlaces de hidrógeno sumados contribuyen al anclaje del bromo en $\sum \rho_b(C-H\cdots Br) = 0.0406 \text{ ua.}$, mientras que el enlace de halógeno putativo contribuye en $\rho_b(C-Br\cdots S) = 0.0035 \text{ ua.}$, lo que representa menos del 10% de la contribución de los 5 enlaces de

hidrógeno, que juntos constituyen lo que previamente llamamos “el *entorno del enlace de halógeno*” (Bogado et al., 2022). Es evidente que en este caso, la fuerza impulsora para el anclaje del resto de bromo fenóxido en el sub-bolsillo S2, no es el enlace de halógeno putativo sino las interacciones del entorno hidrofóbico. En realidad, las interacciones del entorno tampoco logran explicar el activity cliff por sí solas, es por lo tanto una combinación de factores, que involucra las interacciones del entorno y otros factores que no se pueden observar en las estructuras estáticas en las que se realizó el análisis QTAIM.

6.3.4 Análisis de las interacciones del entorno X..S

En el capítulo 4 de esta tesis se estudiaron las interacciones del entorno del EX, cuando este se forma con el oxígeno carbonílico de la cadena principal de una proteína, y se elaboró un protocolo (XBenv), capaz de predecir las condiciones favorables o el entorno favorable para albergar un enlace de halógeno. En este caso, de inhibidores halogenados de cruzipaina, dicho protocolo no puede utilizarse porque el enlace de halógeno se establece entre el átomo de S de un aminoácido que pertenece de la cadena lateral y el halógeno del ligando halogenado. No obstante, se considera muy interesante realizar un estudio en estos entornos específicos y analizar las interacciones de EH, para estos casos de estudio.

Para analizar más profundamente los enlaces de hidrógeno alrededor del EXs, y ver su implicancia, se estudiaron los datos de densidad electrónica obtenidos del análisis QTAIM; como puede verse en la Figura 6.19, se encuentran representados los valores de densidad electrónica del **total de las interacciones que forma el halógeno en cada complejo**, en comparación con el valor de densidad electrónica solo considerando las interacciones del halógeno a través de puentes de hidrógeno (EXent), se puede observar que en la mayoría de los casos estas interacciones contribuyen en más del 50% al total del valor de densidad electrónica, indicando que podrían ser relevantes para la formación de los EXs y su estabilidad.

En la Figura 6.20 se detallan los grafos moleculares de los complejos con ligandos halogenados, donde se pueden observar específicamente las interacciones de EH que se forman en cada uno de los complejos en el sitio activo alrededor del EX que se forma al interaccionar el halógeno con el azufre de la metionina 68.

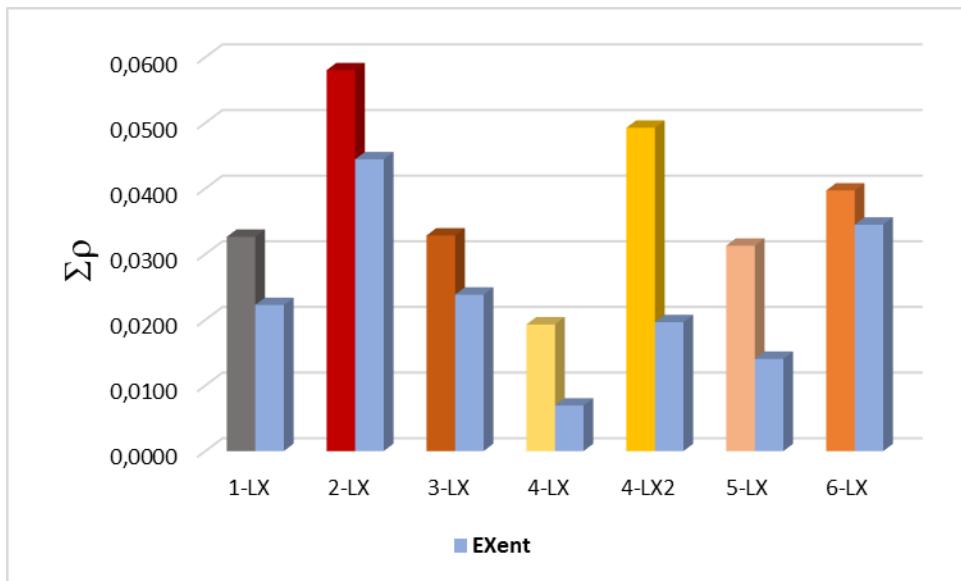


Figura 6.19: Valores de densidad electrónica total (ρ) (u.a.) de los puntos críticos de enlace (BCP) de los ligandos halogenados, en celeste se puede observar la contribución de las interacciones del entorno (EXent).

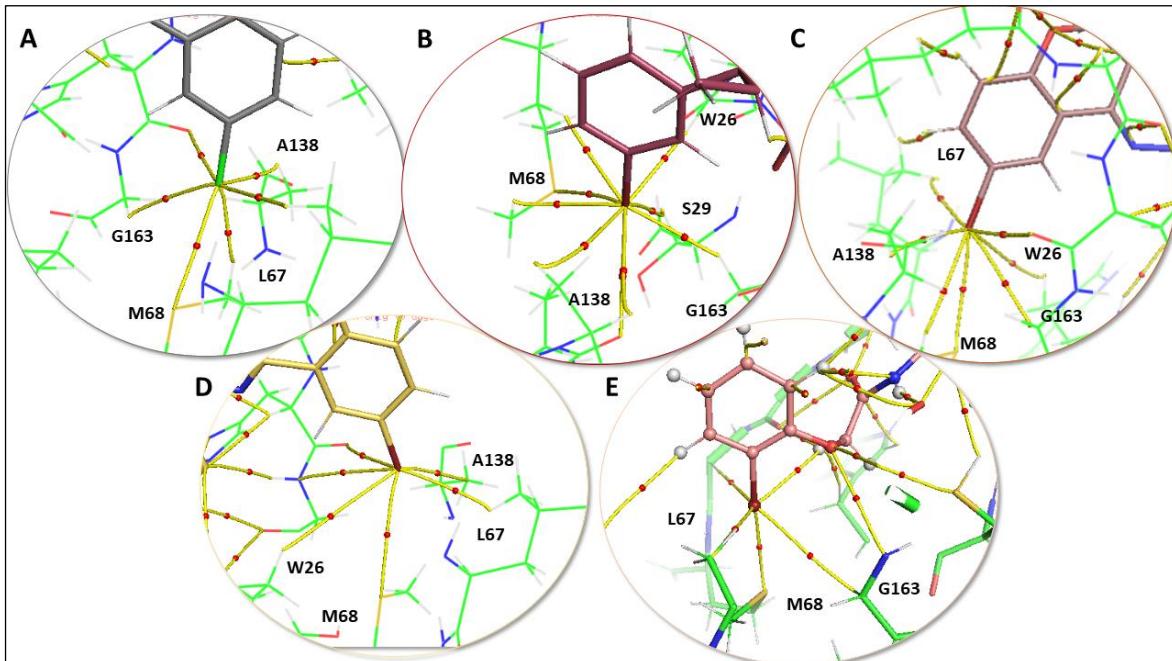


Figura 6.20 Grafos moleculares de las estructuras representativas de la DM de los complejos con ligandos halogenados, donde se detallan los residuos con los que el halógeno forma EXs y EHs. A)1-LX, B)2-LX, C)3-LX, D)4-LX y E) 5-LX. En líneas amarillas se observan los caminos de enlace y las esferas rojas representan los puntos críticos de enlace

6.4 Conclusiones

Los resultados de las simulaciones de Dinámica Molecular parametrizadas para Enlaces de Halógeno reflejan con una buena aproximación las distancias interatómicas y los ángulos involucrados en los enlaces de halógeno que se observan en estructuras cristalinas.

El desarrollo de este trabajo permite comprobar que la representación del *agujero σ* mediante un punto extra de carga permite acercarnos a una mejor comprensión de los enlaces de halógeno.

La aplicación del método de dinámica molecular parametrizado para EX permite estudiar ligandos halogenados desde un punto de vista nuevo, donde la comprensión de los enlaces de halógeno aporta significativamente al entendimiento de las interacciones que afectan positiva o negativamente la actividad biológica.

Los resultados de este capítulo muestran que la sustitución de un hidrógeno por halógeno no cambia drásticamente el modo de unión global del ligando en el receptor. Sin embargo, una mirada más detallada de las interacciones moleculares reveló diferencias en los patrones de interacción.

Los datos obtenidos indicaron que mientras LX mantiene la interacción con Metionina68, su contraparte LH rápidamente se separa del sitio de unión de la Cz. Así se pudo comprender la forma en que inhibidores potencialmente activos interaccionan con el sitio de unión de la Cz, produciendo cambios conformacionales en la enzima que podrían ser considerados responsables de la diferencia de actividad, demostrando que el valor total de densidad electrónica ($\Sigma\rho$), correspondiente a las interacciones establecidas por un grupo particular de átomos del ligando, es una medida de la fuerza de anclaje de ese grupo al bolsillo de unión. El estudio comparativo de pares de ligandos halogenados/no-halogenados puede resultar útil para delinear pautas generales sobre cómo utilizar los halógenos de manera racional para mejorar la afinidad de un compuesto líder por su blanco molecular.

El análisis de las interacciones del entorno, principalmente de los EHs en los casos de estudio permitió destacar su relevancia en la formación y estabilidad de EX no convencionales como los que conforman los ligandos halogenados con el azufre de la Metionina 68.

A través del análisis de las interacciones se ha comprobado en el caso del complejo 2-LX que la capacidad de inducir un “activity cliff” tiene que ver con su habilidad para impulsar estos cambios conformacionales críticos. Siendo la fuerza impulsora el enlace de halógeno en sí mismo, sino a los cambios que se producen a consecuencia del enlace de halógeno y que dan

lugar a otras interacciones electrostáticas, que permiten un anclaje global más fuerte del inhibidor a la Cz.

6.7 Referencias

- Almeida, P. C., Nantes, I. L., Rizzi, C. C. A., Júdice, W. A. S., Chagas, J. R., Juliano, L., Nader, H. B., & Tersariol, I. L. S. (1999). Cysteine Proteinase Activity Regulation. *Journal of Biological Chemistry*, 274(43), 30433–30438. <https://doi.org/10.1074/jbc.274.43.30433>
- Andrews, N. W., Abrams, C. K., Slatin, S. L., & Griffiths, G. (1990). A *T. cruzi*-secreted protein immunologically related to the complement component C9: Evidence for membrane pore-forming activity at low pH. *Cell*, 61(7), 1277–1287. [https://doi.org/10.1016/0092-8674\(90\)90692-8](https://doi.org/10.1016/0092-8674(90)90692-8)
- Aparicio, I. M., Scharfstein, J., & Lima, A. P. C. A. (2004). A new cruzipain-mediated pathway of human cell invasion by *Trypanosoma cruzi* requires trypomastigote membranes. *Infection and Immunity*, 72(10), 5892–5902. <https://doi.org/10.1128/IAI.72.10.5892-5902.2004>
- Bader, R. (1990). *Atoms in molecules a quantum theory*. Oxford University Press.
- Benítez-Hernández, I., Méndez-Enríquez, E., Ostoa, P., Fortoul, T., Ramírez, J. A., Stempin, C., Cerbán, F., Soldevila, G., & Garća-Zepeda, E. A. (2010). Proteolytic cleavage of chemokines by *Trypanosoma cruzi*'s cruzipain inhibits chemokine functions by promoting the generation of antagonists. *Immunobiology*, 215(5), 413–426. <https://doi.org/10.1016/J.IMBIO.2009.06.001>
- Bern, C. (2015). Chagas' Disease. <Https://Doi.Org/10.1056/NEJMra1410150>, 373(5), 456–466. <https://doi.org/10.1056/NEJMRA1410150>
- Bogado, M. L., Villafaña, R. N., Gómez Chavez, J. L., Angelina, E. L., Sosa, G. L., & Peruchena, N. M. (2022). Targeting Protein Pockets with Halogen Bonds: The Role of the Halogen Environment. *Journal of Chemical Information and Modeling*, 62(24), 6494–6507. https://doi.org/10.1021/ACS.JCIM.2C00475/SUPPL_FILE/CI2C00475_SI_001.PDF
- Bondi, A. (1964). Van der waals volumes and radii. *Journal of Physical Chemistry*, 68(3), 441–451. <https://doi.org/10.1021/J100785A001>
- Bontempi, E., & Cazzulo, J. J. (1990). Digestion of human immunoglobulin G by the major cysteine proteinase (cruzipain) from *Trypanosoma cruzi*. *FEMS Microbiology Letters*, 70(3), 337–341. <https://doi.org/10.1111/J.1574-6968.1990.TB14000.X>

Bontempi, E., Franke de Cazzulo, B. M., Ruiz, A. M., & Cazzulo, J. J. (1984). Purification and some properties of an acidic protease from epimastigotes of *Trypanosoma cruzi*. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry*, 77(3), 599–604. [https://doi.org/10.1016/0305-0491\(84\)90282-7](https://doi.org/10.1016/0305-0491(84)90282-7)

Brener, Z. (2003). Biology of *Trypanosoma cruzi*. *Annals of the New York Academy of Sciences*, 1001, 27, 347–382. [https://doi.org/10.1146/ANNUREV.MI.27.100173.002023](https://doi.org/10.1146/annurev.mi.27.100173.002023)

Brinen, L. S., Hansell, E., Cheng, J., Roush, W. R., McKerrow, J. H., & Fletterick, R. J. (2000). A target within the target: Probing cruzain's P1' site to define structural determinants for the Chagas' disease protease. *Structure*, 8(8), 831–840. [https://doi.org/10.1016/S0969-2126\(00\)00173-8](https://doi.org/10.1016/S0969-2126(00)00173-8)

Bryant, C., Kerr, I. D., Debnath, M., Ang, K. K. H., Ratnam, J., Ferreira, R. S., Jaishankar, P., Zhao, D. M., Arkin, M. R., McKerrow, J. H., Brinen, L. S., & Renslo, A. R. (2009). Novel non-peptidic vinylsulfones targeting the S2 and S3 subsites of parasite cysteine proteases. *Bioorganic & Medicinal Chemistry Letters*, 19(21), 6218–6221. <https://doi.org/10.1016/J.BMCL.2009.08.098>

Burleigh, A., Woolsey, A. M., Burleigh, B. A., & Woolsey, A. M. (2002). Cell signalling and *Trypanosoma cruzi* invasion. *Cellular Microbiology*, 4(11), 701–711. <https://doi.org/10.1046/J.1462-5822.2002.00226.X>

Burley, S. K., Bhikadiya, C., Bi, C., Bittrich, S., Chen, L., Crichlow, G. V., Christie, C. H., Dalenberg, K., Di Costanzo, L., Duarte, J. M., Dutta, S., Feng, Z., Ganesan, S., Goodsell, D. S., Ghosh, S., Green, R. K., Guranovic, V., Guzenko, D., Hudson, B. P., ... Zhuravleva, M. (2021). RCSB Protein Data Bank: powerful new tools for exploring 3D structures of biological macromolecules for basic and applied research and education in fundamental biology, biomedicine, biotechnology, bioengineering and energy sciences. *Nucleic Acids Research*, 49(D1), D437–D451. <https://doi.org/10.1093/NAR/GKAA1038>

Caffrey, C., Scory, S., & Steverding, D. (2000). Cysteine proteinases of trypanosome parasites: novel targets for chemotherapy. *Current Drug Targets*, 1(2), 155–162. <https://doi.org/10.2174/1389450003349290>

Campetella, O., Henriksson, J., Åslund, U., Frasch, A. C. C., Pettersson, U., & Cazzulo, J. J. (1992). The major cysteine proteinase (cruzipain) from *Trypanosoma cruzi* is encoded by multiple

polymorphic tandemly organized genes located on different chromosomes. *Molecular and Biochemical Parasitology*, 50(2), 225–234. [https://doi.org/10.1016/0166-6851\(92\)90219-A](https://doi.org/10.1016/0166-6851(92)90219-A)

Capaci Rodrigues, G., Palermo Aguiar, A., Lidio da Silva Goncalves Vianez, J., Macrae, A., Cristina Nogueira de Melo, A., & Beatriz Vermelho, A. (2011). Peptidase Inhibitors as a Possible Therapeutic Strategy for Chagas Disease. *Current Enzyme Inhibition*, 6(4), 183–194. <https://doi.org/10.2174/157340810794578506>

Carlomagno M, Cura E, P. A. y S. E. (1989). *Informe sobre chagas*. 1(2), 2–7. www.produccion-animal.com.ar

Carr, J. A., Pingue-Filho, P., Donizette Malvezi, A., Simão Machado, F., Villar, S. R., Silva Oliveira Barbosa, da E., da Silva Oliveira Barbosa, E., Roggero, E. A., González, F. B., del Valle Fernández, R., Frias Carvalho, V., Bottasso, O. A., & Pérez, A. R. (2021). Evidence in Favor of an Alternative Glucocorticoid Synthesis Pathway During Acute Experimental Chagas Disease. *Frontiers in Endocrinology*, 10, 866. <https://doi.org/10.3389/FENDO.2019.00866>

Case, D., Cheatahm, T., Darden, T., Gohlke, H., Luo, R., Merz, K., Onufriev, A., Simmerling, C., Wang, B., & Woods, R. (2005). *The Amber Biomolecular Simulation Programs*.

Cazzulo, J. J. (1994). Intermediate metabolism in *Trypanosoma cruzi*. *Journal of Bioenergetics and Biomembranes*, 26(2), 157–165. <https://doi.org/10.1007/BF00763064>

Cazzulo, J. J., Cazzulo Franke, M. C., Martínez, J., & Franke de Cazzulo, B. M. (1990). Some kinetic properties of a cysteine proteinase (cruzipain) from *Trypanosoma cruzi*. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology*, 1037(2), 186–191. [https://doi.org/10.1016/0167-4838\(90\)90166-D](https://doi.org/10.1016/0167-4838(90)90166-D)

Dos Reis, F. C. G., Júdice, W. A. S., Juliano, M. A., Juliano, L., Scharfstein, J., & Ana Paula, A. P. C. (2006). The substrate specificity of cruzipain 2, a cysteine protease isoform from *Trypanosoma cruzi*. *FEMS Microbiology Letters*, 259(2), 215–220. <https://doi.org/10.1111/J.1574-6968.2006.00267.X>

Du, X., Guo, C., Hansell, E., Doyle, P. S., Caffrey, C. R., Holler, T. P., McKerrow, J. H., & Cohen, F. E. (2002). Synthesis and structure-activity relationship study of potent trypanocidal thio semicarbazone inhibitors of the trypanosomal cysteine protease cruzain. *Journal of Medicinal Chemistry*, 45(13), 2695–2707. <https://doi.org/10.1021/jm010459j>

- Durrant, J. D., Keränen, H., Wilson, B. A., & Mccammon, J. A. (2010). Computational Identification of Uncharacterized Cruzain Binding Sites. *PLoS Negl Trop Dis*, 4(5), 676. <https://doi.org/10.1371/journal.pntd.0000676>
- Duschak, V., & Couto, A. (2007). An insight on targets and patented drugs for chemotherapy of Chagas disease. *Recent Patents on Anti-Infective Drug Discovery*, 2(1), 19–51. <https://doi.org/10.2174/157489107779561625>
- Duschak, V., & Couto, A. (2009). Cruzipain, the Major Cysteine Protease of *Trypanosoma cruzi*: A Sulfated Glycoprotein Antigen as Relevant Candidate for Vaccine Development and Drug Target. A Review. *Current Medicinal Chemistry*, 16(24), 3174–3202. <https://doi.org/10.2174/092986709788802971>
- Duschak, V. G., Barboza, M., & Couto, A. S. (2003). *Trypanosoma cruzi*: partial characterization of minor cruzipain isoforms non-adsorbed to Concanavalin A-Sepharose. *Experimental Parasitology*, 104(3–4), 122–130. <https://doi.org/10.1016/J.EXPPARA.2003.08.001>
- Eakinsb, A. E., Millslii, A. A., Harthn, G., Mckerrow, J. H., & Craiks, C. S. (1992). THE JOURNAL OF BIOLOGICAL CHEMISTRY The Sequence, Organization, and Expression of the Major Cysteine Protease (Cruzain) from *Trypanosoma cruzi**. *Journal of Biological Chemistry*, 267(11), 7411–7420. [https://doi.org/10.1016/S0021-9258\(18\)42533-1](https://doi.org/10.1016/S0021-9258(18)42533-1)
- Engel, J. C., Doyle, P. S., & McKerrow, J. H. (1999). Efecto antiparasitario de inhibidores de cistein proteasas in vitro e in vivo en la enfermedad de Chagas experimental. *Medicina (B.Aires)*, 171–175.
- Ferreira, R. S., Desso, M. A., Pauli, I., Souza, M. L., Krogh, R., Sales, A. I. L., Oliva, G., Dias, L. C., & Andricopulo, A. D. (2014). Synthesis, biological evaluation, and structure-activity relationships of potent noncovalent and nonpeptidic cruzain inhibitors as anti-*Trypanosoma cruzi* agents. *Journal of Medicinal Chemistry*, 57(6), 2380–2392. <https://doi.org/10.1021/JM401709B>
- Ferreira, R. S., Simeonov, A., Jadhav, A., Eidam, O., Mott, B. T., Keiser, M. J., McKerrow, J. H., Maloney, D. J., Irwin, J. J., & Shoichet, B. K. (2010). Complementarity between a docking and a high-throughput screen in discovering new cruzain inhibitors. *Journal of Medicinal Chemistry*, 53(13), 4891–4905. <https://doi.org/10.1021/jm100488w>
- Gaulton, A., Bellis, L. J., Bento, A. P., Chambers, J., Davies, M., Hersey, A., Light, Y., McGlinchey, S.,

- Michalovich, D., Al-Lazikani, B., & Overington, J. P. (2012). ChEMBL: A large-scale bioactivity database for drug discovery. *Nucleic Acids Research*, 40(D1). <https://doi.org/10.1093/NAR/GKR777>
- Gauthier, J. Y., Chauret, N., Cromlish, W., Desmarais, S., Duong, L. T., Falgueyret, J. P., Kimmel, D. B., Lamontagne, S., Léger, S., LeRiche, T., Li, C. S., Massé, F., McKay, D. J., Nicoll-Griffith, D. A., Oballa, R. M., Palmer, J. T., Percival, M. D., Riendeau, D., Robichaud, J., ... Black, W. C. (2008). The discovery of odanacatib (MK-0822), a selective inhibitor of cathepsin K. *Bioorganic & Medicinal Chemistry Letters*, 18(3), 923–928. <https://doi.org/10.1016/J.BMCL.2007.12.047>
- Gillmor, S. A., Craik, C. S., & Fletterick, R. J. (1997). Structural determinants of specificity in the cysteine protease cruzain. *Protein Science*, 6(8), 1603–1611. <https://doi.org/10.1002/PRO.5560060801>
- Gloss, G., Barrera, M. R., Monteón, V. M., & Reyes, P. A. (1990). [American trypanosomiasis and chronic Chagas cardiopathy at the "Ignacio Chaves" National Institute of Cardiology]. *Archivos Del Instituto de Cardiología de Mexico*, 60(3), 261–266. <https://europepmc.org/article/med/2241399>
- Gomes, S. A. O., Misael, D., Silva, B. A., Feder, D., Silva, C. S., Gonçalves, T. C. M., Santos, A. L. S., & Santos-Mallet, J. R. (n.d.). *Major cysteine protease (cruzipain) in Z3 sylvatic isolates of Trypanosoma cruzi from Rio de Janeiro, Brazil*. <https://doi.org/10.1007/s00436-009-1446-5>
- Götz, M. G., Caffrey, C. R., Hansell, E., McKerrow, J. H., & Powers, J. C. (2004). Peptidyl allyl sulfones: A new class of inhibitors for clan CA cysteine proteases. *Bioorganic and Medicinal Chemistry*, 12(19), 5203–5211. <https://doi.org/10.1016/j.bmc.2004.07.016>
- Hellman, U., Wernstedt, C., & Cazzulo, J. J. (1991). Self-proteolysis of the cysteine proteinase, cruzipain, from *Trypanosoma cruzi* gives a major fragment corresponding to its carboxy-terminal domain. *Molecular and Biochemical Parasitology*, 44(1), 15–21. [https://doi.org/10.1016/0166-6851\(91\)90216-S](https://doi.org/10.1016/0166-6851(91)90216-S)
- Huang, L., Brinen, L. S., & Ellman, J. A. (2003). Crystal structures of reversible ketone-Based inhibitors of the cysteine protease cruzain. *Bioorganic and Medicinal Chemistry*, 11(1), 21–29. [https://doi.org/10.1016/S0968-0896\(02\)00427-3](https://doi.org/10.1016/S0968-0896(02)00427-3)

- Ibañez De Barrett, C. I., Bozzini, J. P., & De Bozzini, M. M. (2004). Differentiation and morphogenesis of *Triatoma infestans* (Klug 1834) female gonads. I - Post embryonic development. *Biocell*, 28(3), 259–269.
- Ibrahim, M. A. A. (2011). Molecular Mechanical Study of Halogen Bonding in Drug Discovery. *Journal of Computational Chemistry*, 32(12), 2564–2574. <https://doi.org/10.1002/jcc>
- Inst, M., Cruz, O., Mckerrow, J. H., Doyle, P. S., Engel, J. C., Podust, L. M., Robertson, S. A., Ferreira, R., Saxton, T., Arkin, M., Kerr, I. D., Brinen, L. S., & Craik, C. S. (2009). *Two approaches to discovering and developing new drugs for Chagas disease*. 104, 263–269.
- Jacobsen, W., Christians, U., & Benet, L. Z. (2000). In Vitro Evaluation of the Disposition of A Novel Cysteine Protease Inhibitor. *Drug Metabolism and Disposition*, 28(11).
- Jaishankar, P., Hansell, E., Zhao, D. M., Doyle, P. S., McKerrow, J. H., & Renslo, A. R. (2008). Potency and selectivity of P2/P3-modified inhibitors of cysteine proteases from trypanosomes. *Bioorganic & Medicinal Chemistry Letters*, 18(2), 624–628. <https://doi.org/10.1016/J.BMCL.2007.11.070>
- Leung, D., Abbenante, G., & Fairlie, D. P. (2000). *Perspective Protease Inhibitors: Current Status and Future Prospects*. <https://doi.org/10.1021/jm990412m>
- Lima, A. P. C. A., Dos Reis, F. C. G., Serveau, C., Lalmanach, G., Juliano, L., Ménard, R., Vernet, T., Thomas, D. Y., Storer, A. C., & Scharfstein, J. (2001). Cysteine protease isoforms from *Trypanosoma cruzi*, cruzipain 2 and cruzain, present different substrate preference and susceptibility to inhibitors. *Molecular and Biochemical Parasitology*, 114(1), 41–52. [https://doi.org/10.1016/S0166-6851\(01\)00236-5](https://doi.org/10.1016/S0166-6851(01)00236-5)
- Lu, T., & Chen, F. (2012). Multiwfn: A multifunctional wavefunction analyzer. *Journal of Computational Chemistry*, 33(5), 580–592. <https://doi.org/10.1002/jcc.22885>
- Luchi, A. M., Villafaña, R. N., Gómez Chávez, J. L., Bogado, M. L., Angelina, E. L., & Peruchena, N. M. (2019). Combining Charge Density Analysis with Machine Learning Tools to Investigate the Cruzain Inhibition Mechanism. *ACS Omega*, 4(22), 19582–19594. <https://doi.org/10.1021/acsomega.9b01934>
- M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ort, J. B. F. and D. J. F. (2016).

- Mallari, J. P., Shelat, A. A., Obrien, T., Caffrey, C. R., Kosinski, A., Connelly, M., Harbut, M., Greenbaum, D., McKerrow, J. H., & Guy, R. K. (2008). Development of potent purine-derived nitrile inhibitors of the trypanosomal protease TbcatB. *Journal of Medicinal Chemistry*, 51(3), 545–552. https://doi.org/10.1021/JM070760L/SUPPL_FILE/JM070760L-FILE002.PDF
- McGrath, M. E., Eakin, A. E., Engel, J. C., McKerrow, J. H., Craik, C. S., & Fletterick, R. J. (1995). The Crystal Structure of Cruzain: A Therapeutic Target for Chagas' Disease. *Journal of Molecular Biology*, 247(2), 251–259. <https://doi.org/10.1006/JMBI.1994.0137>
- McKerrow, J. H., Caffrey, C., Kelly, B., Loke, P., & Sajid, M. (2006). PROTEASES IN PARASITIC DISEASES. [Https://Doi.Org/10.1146/Annurev.Pathol.1.110304.100151](https://doi.org/10.1146/Annurev.Pathol.1.110304.100151), 1, 497–536. <https://doi.org/10.1146/ANNUREV.PATHOL.1.110304.100151>
- O'Boyle, N. M., Banck, M., James, C. A., Morley, C., Vandermeersch, T., & Hutchison, G. R. (2011). Open Babel: An open chemical toolbox. *Journal of Cheminformatics*, 3(33), 33. <https://jcheminf.biomedcentral.com/track/pdf/10.1186/1758-2946-3-33> [http://www.ncbi.nlm.nih.gov/pubmed/21982300](https://www.ncbi.nlm.nih.gov/pubmed/21982300) [http://www.ncbi.nlm.nih.gov/articlerender.fcgi?artid=PMC3198950](https://www.ncbi.nlm.nih.gov/articlerender.fcgi?artid=PMC3198950)
- Piras, M. M., Henriquez, D., & Piras, R. (1985). The effect of proteolytic enzymes and protease inhibitors on the interaction Trypanosoma cruzi — fibroblasts. *Molecular and Biochemical Parasitology*, 14(2), 151–163. [https://doi.org/10.1016/0166-6851\(85\)90034-9](https://doi.org/10.1016/0166-6851(85)90034-9)
- Powers, J. C., Asgian, J. L., Ekici, Ö. D., & James, K. E. (2002). Irreversible inhibitors of serine, cysteine, and threonine proteases. *Chemical Reviews*, 102(12), 4639–4750. https://doi.org/10.1021/CR010182V/ASSET/CR010182V.FP.PNG_V03
- Rachman, M., Scarpino, A., Bajusz, D., Pálfy, G., Vida, I., Perczel, A., Barril, X., & Keserű, G. M. (2019). DUckCov: a Dynamic Undocking-Based Virtual Screening Protocol for Covalent Binders. *ChemMedChem*, 14(10), 1011–1021. <https://doi.org/10.1002/CMDC.201900078>
- Raimondi, A., Wernstedt, C., Hellman, U., & Cazzulo, J. J. (1991). Degradation of oxidised insulin A and B chains by the major cysteine proteinase (cruzipain) from Trypanosoma cruzi epimastigotes. *Molecular and Biochemical Parasitology*, 49(2), 341–344. [https://doi.org/10.1016/0166-6851\(91\)90080-P](https://doi.org/10.1016/0166-6851(91)90080-P)

Santana, J. M., Grellier, P., Rodier, M. H., Schrevel, J., & Teixeira, A. (1992). Purification and characterization of a new 120 kDa alkaline proteinase of *Trypanosoma cruzi*. *Biochemical and Biophysical Research Communications*, 187(3), 1466–1473. [https://doi.org/10.1016/0006-291X\(92\)90467-Y](https://doi.org/10.1016/0006-291X(92)90467-Y)

Santos, C. C., Sant'Anna, C., Terres, A., Cunha-e-Silva, N. L., Scharfstein, J., & Lima, A. P. C. d. A. (2005). Chagasin, the endogenous cysteine-protease inhibitor of *Trypanosoma cruzi*, modulates parasite differentiation and invasion of mammalian cells. *Journal of Cell Science*, 118(Pt 5), 901–915. <https://doi.org/10.1242/JCS.01677>

Sartori, G. R., Leitão, A., Montanari, C. A., & Laughton, C. A. (2019). Ligand-induced conformational selection predicts the selectivity of cysteine protease inhibitors. *PLOS ONE*, 14(12), e0222055. <https://doi.org/10.1371/JOURNAL.PONE.0222055>

Schechter, I., & Berger, A. (1967). On the size of the active site in proteases. I. Papain. *Biochemical and Biophysical Research Communications*, 27(2), 157–162. [https://doi.org/10.1016/S0006-291X\(67\)80055-X](https://doi.org/10.1016/S0006-291X(67)80055-X)

Scheidt, K. A., Roush, W. R., McKerrow, J. H., Selzer, P. M., Hansell, E., & Rosenthal, P. J. (1998). Structure-based design, synthesis and evaluation of conformationally constrained cysteine protease inhibitors. *Bioorganic & Medicinal Chemistry*, 6(12), 2477–2494. [https://doi.org/10.1016/S0968-0896\(98\)80022-9](https://doi.org/10.1016/S0968-0896(98)80022-9)

Schenkman, S., & Eichinger, D. (1993). *Trypanosoma cruzi* trans-sialidase and cell invasion. *Parasitology Today*, 9(6), 218–222. [https://doi.org/10.1016/0169-4758\(93\)90017-A](https://doi.org/10.1016/0169-4758(93)90017-A)

Schirmeister, T., & Kaepller, U. (2005). Non-Peptidic Inhibitors of Cysteine Proteases. *Mini-Reviews in Medicinal Chemistry*, 3(4), 361–373. <https://doi.org/10.2174/1389557033488079>

Siles, R., Chen, S. E., Zhou, M., Pinney, K. G., & Trawick, M. L. (2006). Design, synthesis, and biochemical evaluation of novel cruzain inhibitors with potential application in the treatment of Chagas' disease. *Bioorganic and Medicinal Chemistry Letters*, 16(16), 4405–4409. <https://doi.org/10.1016/j.bmcl.2006.05.041>

Stoka, V., McKerrow, J. H., Cazzulo, J. J., & Turk, V. (1998). Substrate inhibition of cruzipain is not affected by the C-terminal domain. *FEBS Letters*, 429(2), 129–133. [https://doi.org/10.1016/S0014-5793\(98\)00532-8](https://doi.org/10.1016/S0014-5793(98)00532-8)

Stoka, V., Nycander, M., Lenarčič, B., Labriola, C., Cazzulo, J. J., Björk, I., & Turk, V. (1995). Inhibition of cruzipain, the major cysteine proteinase of the protozoan parasite, *Trypanosoma cruzi*, by proteinase inhibitors of the cystatin superfamily. *FEBS Letters*, 370(1-2), 101–104. [https://doi.org/10.1016/0014-5793\(95\)00798-E](https://doi.org/10.1016/0014-5793(95)00798-E)

Stumpfe, D.; Bajorath, J. (2012). Exploring Activity Cliffs in Medicinal Chemistry. *J. Med. Chem.*, 55, 2932– 2942, DOI: 10.1021/jm201706b

Tomás, A. M., & Kelly, J. M. (1996). Stage-regulated expression of cruzipain, the major cysteine protease of *Trypanosoma cruzi* is independent of the level of RNA. *Molecular and Biochemical Parasitology*, 76(1-2), 91–103. [https://doi.org/10.1016/0166-6851\(95\)02545-6](https://doi.org/10.1016/0166-6851(95)02545-6)

Turk, D., Gunčar, G., Podobnik, M., & Turk, B. (1998). Revised definition of substrate binding sites of papain-like cysteine proteases. *Biological Chemistry*, 379(2), 137–147. [https://doi.org/10.1515/BCHM.1998.379.2.137/MACHINEREADABLECITATION/RIS](https://doi.org/10.1515/BCHM.1998.379.2.137)

Tyler, K. M., & Engman, D. M. (2001). The life cycle of *Trypanosoma cruzi* revisited. *International Journal for Parasitology*, 31(5-6), 472–481. [https://doi.org/10.1016/S0020-7519\(01\)00153-9](https://doi.org/10.1016/S0020-7519(01)00153-9)

Urbina, J. A. (1999). Chemotherapy of Chagas' disease: the how and the why. *J Mol Med*, 77, 332–338.

Wilcken, R., Zimmermann, M. O., Lange, A., Joerger, A. C., & Boeckler, F. M. (2013). Principles and applications of halogen bonding in medicinal chemistry and chemical biology. *Journal of Medicinal Chemistry*, 56(4), 1363–1388. <https://doi.org/10.1021/jm3012068>

World Health Organization (WHO) Special Programme for Research and Training in Tropical Diseases (TDR) Neglected Tropical Diseases (NTD) - Global Health Impact Group. (n.d.). Retrieved February 15, 2023, from <https://www.globalhealthig.com/world-health-organization-who-special-programme-for-research-and-training-in-tropical-diseases-tdr-neglected-tropical-diseases-ntd/>

Capítulo 7

Conclusiones Generales

Los enlaces de halógenos, EXs, juegan un importante rol en muchas áreas de la química moderna, y la comprensión de los aspectos fundamentales de estas interacciones es esencial para el diseño racional de fármacos, la inserción de átomos de halógeno en compuestos se realiza predominantemente para explotar sus efectos estéricos, a través de la capacidad de estos átomos para ocupar el sitio de unión en blanco moleculares; a lo largo de este trabajo se ha detallado algunas de las consideraciones necesarias para comprender cómo los enlaces X se pueden utilizar en macromoléculas biológicas.

Durante desarrollo de esta tesis se han utilizado distintas metodologías propias de la química computacional, desde la mecánica clásica en las simulaciones de dinámica molecular a la mecánica cuántica, aplicando en los diversos pasos herramientas quimioinformáticas. El trabajo realizado se puede dividir en tres líneas generales: la parametrización de métodos de modelado molecular para que describan correctamente los enlaces de halógeno, el estudio de las interacciones del entorno de los enlaces de halógeno en los bolsillos de las proteínas, y el modelado de EXs en blancos de estudio de interés en la Química medicinal. Los resultados obtenidos se pueden condensar en las siguientes conclusiones generales:

- En este trabajo se ha demostrado que es necesario poner a punto los métodos no cuánticos de modelado molecular para que representen adecuadamente las características propias de un EX. Específicamente, en esta tesis se propone un procedimiento eficaz para describir apropiadamente este tipo de interacción en el método de Docking Molecular, los resultados obtenidos demuestran que modificando el potencial de pozo en los parámetros del campo de fuerza de Autodock y ajustando la direccionalidad del EX hacia un átomo aceptor el enlace permanece estable y los resultados del Docking Molecular representan poses similares a la de los complejos en su estructura cristalina, a diferencia de las poses obtenidas con el docking sin parametrizar. Para los tres tipos de halógenos (Bromo, Cloro y Yodo) el mayor porcentaje de estructuras que presentaron una mínima desviación cuadrática media (RMSD) entre las estructuras experimentales y las predichas por el campo de fuerza

modificado, fue con un valor de potencial de pozo de 3 kcal/mol, por lo cual fue considerado el potencial óptimo para este tipo de enlaces.

Para la parametrización de la técnica de Dinámica Molecular se ha usado el protocolo propuesto por Ibrahim y col., para simular el agujero- σ sobre el átomo de halógeno, se introdujo un extra-punto (EP) con carga positiva y sin masa en el campo de fuerza de Amber, se estableció el ángulo C(ar)-X-EP en 180° y la distancia X-EP se igualó al radio atómico del halógeno. Los resultados de las simulaciones de Dinámica Molecular parametrizadas para Enlaces de Halógeno realizadas en esta tesis reflejan con una buena aproximación las distancias interatómicas y los ángulos involucrados en los enlaces de halógeno que se observan en estructuras cristalinas. El desarrollo de este trabajo permitió comprobar que la representación del *agujero* σ mediante un extra punto permite acercarnos a una mejor comprensión de los enlaces de halógeno.

- ✚ Los átomos de halógeno en el bolsillo de proteína pueden formar otras interacciones no covalentes además del enlace X, lo que se denomina el *entorno de enlaces de halógeno* (EXent). En esta tesis se presenta por primera vez, un estudio exhaustivo de las interacciones del entorno de los EXs en base al análisis topológico de la densidad de carga electrónica (en el contexto de la teoría AIM), se realizó un relevamiento sobre 86 complejos ligando-proteína unidos a halógenos inspeccionando no solo el enlace X sino también otros tipos de interacciones no covalentes que forma el halógeno, se calculó la densidad de carga electrónica de los complejos y se utilizó descriptores topológicos QTAIM para medir la fuerza de ese entorno. Principalmente en el entorno se destacó el rol importante de los enlaces de Hidrógeno, los cuales representan las interacciones mayoritarias y se ha demostrado su posible implicancia en la fortaleza de la formación de EXs.

Los resultados confirmaron la suposición de que el entorno de enlace X hace una gran contribución a la fuerza de anclaje del halógeno, incluso mayor que el propio enlace X. Como un aporte destacado de esta tesis se ha logrado elaborar un protocolo para la predicción de la presencia de enlaces de halógeno en base al entorno (xenv), que se encuentra disponible en el repositorio de Github.

Los blancos moleculares sobre los que se trabajó en esta tesis fueron el receptor D2 de dopamina (DR2), que pertenece a la familia de Receptores Acoplados a Proteína G (GPCR) y la enzima Cruzipaina, la principal cisteína proteasa del *Tripanosoma cruzi*, parásito responsable del mal de Chagas. En el primer caso de estudio, DR2, en una primera etapa se aplicó las técnicas de modelado por homología y dinámica molecular para obtener dos modelos del receptor en la forma activa e inactiva, los cuales fueron validados y luego comparados con las estructuras cristalinas, obteniendo resultados favorables. Posteriormente se estudiaron los cambios conformacionales que se producen durante la activación del receptor al unirse a un compuesto agonista, en el caso del receptor activo, y a un agonista inverso en el caso del receptor inactivo a través de los *interruptores moleculares*. En una segunda etapa se aplicó la técnica docking molecular parametrizado para enlaces de halógeno este blanco molecular con ligandos halogenados, analizando luego las poses obtenidas a través del método de “fingerprints” de interacción estructural (SIFT), los resultados permitieron corroborar el hecho de que compuestos agonistas actúan preferentemente con el receptor activado. Además, en una última etapa se pudo aplicar el protocolo para predecir la formación de EXs en entornos proteicos (xenv) a ambas formas del receptor D2Ra y D2Ri, y se compararon los resultados de la predicción con los datos obtenidos mediante análisis QTAIM, comprobando que concuerdan correctamente.

En el blanco molecular Cruzipaina se utilizó la técnica de Dinámica Molecular parametrizada para enlaces de halógeno para el estudio de inhibidores halogenados de la enzima. El análisis del comportamiento de los enlaces de halógeno a través de las simulaciones fue un aporte significativo al entendimiento de las interacciones que afectan positiva o negativamente la actividad biológica. Los resultados demostraron que la sustitución de un hidrógeno por halógeno no cambia drásticamente el modo de unión global del ligando en el receptor. Sin embargo, una mirada más detallada de las interacciones moleculares a través del análisis QTAIM reveló diferencias en los patrones de interacción. Finalmente, el análisis de las interacciones del entorno, principalmente de los EHs en los casos de estudio permitió destacar su relevancia en la formación y estabilidad de EX no convencionales como son los que conforman los ligandos halogenados con el azufre en este blanco molecular.

En el desarrollo de esta tesis se pudo comprender la implicancia de las NCI (principalmente enlace de hidrógeno y enlace de halógeno) en el contexto de la unión

ligando-proteína en los casos de estudio, por un lado a través de descripción topológica detallada de la red de interacciones del ligando en el bolsillo de unión de la proteína, y por otra parte, mostrando la posibilidad de ir más allá del concepto de interacciones de pares para ver los efectos electrónicos dentro del complejo entorno biológico, aplicando técnicas de modelado molecular.

ANEXOS

ANEXO A

Descripción y parametrización de Enlaces de halógeno en entornos biológicos

Tabla A1: Resultados de las propiedades topológicas expresadas en unidades atómicas de los átomos de oxígeno de las estructuras utilizadas como sistema de referencia.

PDB id	Atomo 1	Residuo 1	Atomo 2	Residuo 2	Densidad electrónica	Laplaciano	L1	L2	L3
3qfx	CL1	CP6	O	THR	0,0067	0,0278	-0,0046	-0,0043	0,0367
3poz	CL	O3P	O	LEU	0,0065	0,0260	-0,0039	-0,0038	0,0337
3ms9	CL1	MS9	O	LEU	0,0082	0,0332	-0,0060	-0,0057	0,0449
3ms7	CL1	22S	O	ALA	0,0056	0,0208	-0,0038	-0,0037	0,0282
3fal	CL10	LO2	O	PHE	0,0073	0,0296	-0,0051	-0,0048	0,0395
3dlg	CL1	GWE	O	TYR	0,0075	0,0288	-0,0049	-0,0047	0,0384
3ksq	CL	Z96	O	LYS	0,0168	0,0766	-0,0135	-0,0123	0,1024
3g4w	CL6	8CL	O	ASN	0,0075	0,0305	-0,0049	-0,0046	0,0400
3jzi	CL28	JZL	O	ILE	0,0063	0,0249	-0,0043	-0,0036	0,0329
2cf8	CL30	ESH	O	GLU	0,0073	0,0281	-0,0051	-0,0048	0,0381
3dle	CL1	GFA	O	TYR	0,0090	0,0359	-0,0064	-0,0059	0,0482
2wpf	CL1	WPF	O	GLY	0,0090	0,0378	-0,0067	-0,0064	0,0509
2bla	CL1	CP6	O	ASN	0,0081	0,0338	-0,0054	-0,0052	0,0444
2wpe	CL	WPE	O	GLY	0,0115	0,0503	-0,0090	-0,0085	0,0678
1kzn	CL1	CBN	O	GLY	0,0079	0,0299	-0,0057	-0,0049	0,0405
1xkk	CL3	FMM	O	LEU	0,0090	0,0358	-0,0061	-0,0057	0,0476
2wp6	CL	WP6	O	GLY	0,0117	0,0504	-0,0090	-0,0087	0,0681
2wpc	CLA	WP7	O	GLY	0,0049	0,0191	-0,0032	-0,0030	0,0253
2bl9	CL1	CP6	O	SER	0,0088	0,0368	-0,0061	-0,0058	0,0487
2uw8	CL1	GVQ	O	GLY	0,0141	0,0644	-0,0105	-0,0060	0,0809
1j3j	CL1	CP6	O	ASN	0,0094	0,0395	-0,0065	-0,0061	0,0522
5tri	CL1	7HM	O	MET	0,0068	0,0268	-0,0049	-0,0044	0,0361
5vcw	CL1	93J	O	LEU	0,0064	0,0268	-0,0040	-0,0035	0,0343
5tm4	CL1	7E3	O	GLU	0,0064	0,0258	-0,0044	-0,0043	0,0345
5o0h	CL	9FN	O	ALA	0,0103	0,0416	-0,0074	-0,0068	0,0558
5h2u	CL	1N1	O	ALA	0,0092	0,0361	-0,0061	-0,0055	0,0477

5icp	CL1	69Z	O	VAL	0,0126	0,0522	-0,0098	-0,0088	0,0709
5kgt	CL	6SQ	O	GLY	0,0084	0,0353	-0,0059	-0,0056	0,0468
5mzi	CL	FYK	O	LEU	0,0100	0,0406	-0,0075	-0,0072	0,0552
5mzc	CL	8EQ	O	LEU	0,0134	0,0563	-0,0109	-0,0105	0,0778
5l0f	CL1	70R	O	ALA	0,0065	0,0251	-0,0040	-0,0037	0,0329
5ag4	CL	N8N	O	ASP	0,0075	0,0322	-0,0054	-0,0052	0,0428
5cbi	CL06	4ZC	O	ALA	0,0066	0,0253	-0,0041	-0,0040	0,0334
5k4i	CL1	6QB	O	GLY	0,0115	0,0479	-0,0087	-0,0025	0,0591
5fh8	CL	5XK	O	MET	0,0083	0,0353	-0,0059	-0,0056	0,0468
5aeh	CL	8IR	O	GLY	0,0095	0,0438	-0,0070	-0,0063	0,0571
5fck	CL	5WC	O	TRP	0,0081	0,0322	-0,0054	-0,0051	0,0427
5fh7	CL	5XL	O	MET	0,0074	0,0308	-0,0050	-0,0049	0,0407
4qxi	CL4	I98	O	VAL	0,0092	0,0359	-0,0064	-0,0061	0,0484
4qxm	CL1	713	O	MET	0,0110	0,0470	-0,0088	-0,0078	0,0636
5aep	CL	QUP	O	GLY	0,0070	0,0259	-0,0040	-0,0036	0,0335
4j3l	CL	AJ5	O	TYR	0,0081	0,0327	-0,0055	-0,0052	0,0433
4j22	CL	AJ7	O	TYR	0,0093	0,0384	-0,0065	-0,0062	0,0511
4xjo	CL9	41O	O	GLY	0,0066	0,0268	-0,0045	-0,0043	0,0356
4w1x	CL1	3G9	O	GLY	0,0073	0,0295	-0,0050	-0,0049	0,0394
4o12	CL1	2QG	O	GLU	0,0080	0,0337	-0,0054	-0,0049	0,0440
4ms4	CL	2C0	O	PHE	0,0079	0,0314	-0,0056	-0,0048	0,0417
3qgt	CL1	CP6	O	SER	0,0078	0,0324	-0,0052	-0,0050	0,0426
3qg2	CL1	CP6	O	ASN	0,0086	0,0365	-0,0059	-0,0057	0,0481
4fpf	CL	JKK	O	SER	0,0093	0,0390	-0,0069	-0,0067	0,0526
4cjv	CL	IV2	O	GLY	0,0088	0,0360	-0,0060	-0,0056	0,0477
3w33	CL1	W19	O	LEU	0,0076	0,0287	-0,0047	-0,0047	0,0381
3v51	CLD	I76	O	GLY	0,0088	0,0342	-0,0059	-0,0041	0,0442
3utu	CL	1TS	O	ASN	0,0076	0,0286	-0,0046	-0,0036	0,0369
3ug2	CL	IRE	O	LEU	0,0093	0,0404	-0,0063	-0,0059	0,0526

Tabla A2: Datos geométricos de los átomos de oxígeno de las estructuras utilizadas como sistema de referencia

PDB id	Atomo 1	Residuo 1	Atomo 2	Residuo 2	Distancia (Å)	Angulo (°)	Coordenadas
3qfx	CL1	CP6	O	THR	3,208526	147,000875	[14,898 31,075 46,992]
3poz	CL	O3P	O	LEU	3,220511	159,441326	[14,008 30,114 46,222]
3ms9	CL1	MS9	O	LEU	3,140528	151,621385	[13,53 32,05 47,014]
3ms7	CL1	22S	O	ALA	3,356383	152,508696	[13,534 29,573 45,379]
3fal	CL10	LO2	O	PHE	3,190997	154,155463	[14,316 30,039 46,1]
3dlg	CL1	GWE	O	TYR	3,211051	146,128444	[14,146 32,012 47,26]
3ksq	CL	Z96	O	LYS	2,743015	171,713143	[13,35 31,29 46,155]
3g4w	CL6	8CL	O	ASN	3,154867	166,92138	[12,967 30,378 45,661]
3jzi	CL28	JZL	O	ILE	3,254684	155,846826	[12,654 31,991 46,584]
2cf8	CL30	ESH	O	GLU	3,2217	148,147187	[12,527 30,194 44,697]
3dle	CL1	GFA	O	TYR	3,110551	148,728419	[13,766 32,186 47,04]
2wpf	CL1	WPF	O	GLY	3,069741	161,535565	[13,962 30,297 46,241]
2bla	CL1	CP6	O	ASN	3,111679	169,672155	[13,027 30,46 45,64]
2wpe	CL	WPE	O	GLY	2,93501	169,99542	[13,569 30,548 46,053]
1kzn	CL1	CBN	O	GLY	3,19911	143,533392	[13,486 29,559 44,651]
1xkk	CL3	FMM	O	LEU	3,099998	159,589948	[14,083 30,256 46,133]
2wp6	CL	WP6	O	GLY	2,938655	169,201972	[12,782 31,201 45,929]
2wpc	CLA	WP7	O	GLY	3,364828	167,257871	[12,51 31,245 46,268]
2bl9	CL1	CP6	O	SER	3,073163	171,13529	[12,988 31,238 46,251]
2uw8	CL1	GVQ	O	GLY	2,902728	147,769719	[12,684 30,522 44,601]
1j3j	CL1	CP6	O	ASN	3,040706	174,701851	[13,492 30,958 46,108]
5tri	CL1	7HM	O	MET	3,236953	140,662313	[15,223 30,501 46,658]
5vcw	CL1	93J	O	LEU	3,212729	162,913375	[12,436 30,767 45,515]
5tm4	CL1	7E3	O	GLU	3,258795	140,899361	[12,122 30,6 44,323]
5o0h	CL	9FN	O	ALA	3,043878	146,644576	[13,536 32,431 46,94]
5h2u	CL	1N1	O	ALA	3,11422	152,548719	[14,435 30,119 46,105]
5icp	CL1	69Z	O	VAL	2,948502	143,380049	[14,216 32,12 47,023]
5kgf	CL	6SQ	O	GLY	3,121222	144,879572	[14,509 29,69 45,608]
5mzi	CL	FYK	O	LEU	3,055912	154,449606	[14,038 29,952 45,674]
5mzc	CL	8EQ	O	LEU	2,915796	146,862175	[14,17 29,863 45,238]
5lof	CL1	70R	O	ALA	3,254014	160,909077	[13,414 29,935 45,655]
5ag4	CL	N8N	O	ASP	3,127026	165,419932	[12,851 30,438 45,562]
5cbi	CL06	4ZC	O	ALA	3,245994	158,300151	[13,778 31,469 47,101]
5k4i	CL1	6QB	O	GLY	3,054749	151,154993	[12,52 30,537 44,815]
5fh8	CL	5XK	O	MET	3,091943	163,698221	[12,451 30,901 45,476]
5aeh	CL	8IR	O	GLY	3,005843	155,009891	[13,074 32,075 46,664]
5fck	CL	5WC	O	TRP	3,133748	173,391598	[13,267 30,576 46,158]
5fh7	CL	5XL	O	MET	3,157344	165,212178	[12,455 30,955 45,649]
4qxi	CL4	I98	O	VAL	3,111158	141,602711	[15,133 30,688 46,669]
4qxm	CL1	713	O	MET	2,99579	131,757628	[12,865 33,34 46,298]
5aep	CL	QUP	O	GLY	3,254813	168,299913	[12,514 30,828 45,909]
4j3l	CL	AJ5	O	TYR	3,138081	173,703991	[13,295 30,451 46,098]
4j22	CL	AJ7	O	TYR	3,0675	172,21679	[12,963 31,241 46,317]

4xjo	CL9	41O	O	GLY	3,237358	147,827635	[14,195 29,593 45,477]
4w1x	CL1	3G9	O	GLY	3,19621	147,146426	[14,212 29,605 45,436]
4o12	CL1	2QG	O	GLU	3,117695	166,302104	[13,781 30,769 46,551]
4ms4	CL	2C0	O	PHE	3,153944	161,734756	[12,382 31,49 46,009]
3qgt	CL1	CP6	O	SER	3,129594	170,702502	[13,521 30,564 46,201]
3qg2	CL1	CP6	O	ASN	3,071253	174,758323	[13,44 31,114 46,502]
4fpf	CL	JKK	O	SER	3,057797	156,756741	[14,1 30,072 45,879]
4cjv	CL	IV2	O	GLY	3,094704	166,415156	[12,923 30,513 45,609]
3w33	CL1	W19	O	LEU	3,202787	156,602876	[12,274 31,753 46,026]
3v51	CLD	I76	O	GLY	3,170516	143,391355	[14,745 31,535 47,159]
3utu	CL	1TS	O	ASN	3,243358	143,702258	[14,093 32,242 47,296]
3ug2	CL	IRE	O	LEU	3,027465	161,999376	[12,403 31,018 45,471]

Tabla A3: Resultados de las propiedades topológicas expresadas en unidades atómicas de los átomos de hidrógeno de las estructuras utilizadas como sistema de referencia.

PDBid	Atomo 1	Residuo 1	Atomo 2	Residuo 2	Densidad electrónica	Laplaciano	L1	L2	L3
3qfx	CL1	CP6	HG	LEU	0,0155	0,0568	-0,0158	-0,0146	0,0871
3qfx	CL1	CP6	HG21	THR	0,0077	0,0300	-0,0059	-0,0022	0,0381
3poz	CL	03P	HB3	LEU	0,0050	0,0172	-0,0035	-0,0031	0,0238
3poz	CL	03P	HB3	LYS	0,0071	0,0271	-0,0047	-0,0028	0,0346
3poz	CL	03P	HG22	THR	0,0093	0,0349	-0,0079	-0,0066	0,0494
3ms9	CL1	MS9	HB	ILE	0,0047	0,0172	-0,0031	-0,0027	0,0231
3ms9	CL1	MS9	HB2	ALA	0,0014	0,0041	-0,0007	-0,0004	0,0052
3ms9	CL1	MS9	HG22	ILE	0,0048	0,0172	-0,0030	-0,0022	0,0224
3ms9	CL1	MS9	HD21	LEU	0,0036	0,0121	-0,0024	-0,0023	0,0168
3ms9	CL1	MS9	HB2	ALA	0,0062	0,0216	-0,0048	-0,0045	0,0309
3ms9	CL1	MS9	HB2	LEU	0,0018	0,0056	-0,0007	-0,0002	0,0066
3ms7	CL1	22S	HB3	ALA	0,0047	0,0183	-0,0034	-0,0033	0,0250
3ms7	CL1	22S	HA	ASN	0,0058	0,0198	-0,0044	-0,0038	0,0280
3kmx	CL1	G00	HG3	GLN	0,0042	0,0143	-0,0029	-0,0026	0,0198
3kmx	CL1	G00	HA2	GLY	0,0054	0,0214	-0,0043	-0,0042	0,0299
3kmx	CL1	G00	HE22	GLN	0,0047	0,0171	-0,0030	-0,0020	0,0221
3kmx	CL1	G00	HD1	PHE	0,0053	0,0187	-0,0042	-0,0039	0,0268
3fal	CL10	LO2	HB2	LEU	0,0064	0,0240	-0,0043	-0,0021	0,0304
3fal	CL10	LO2	HB2	ALA	0,0064	0,0226	-0,0053	-0,0046	0,0325
3fal	CL10	LO2	HD1	PHE	0,0024	0,0080	-0,0015	-0,0014	0,0109
3dlg	CL1	GWE	HG23	VAL	0,0066	0,0237	-0,0053	-0,0052	0,0342
3dlg	CL1	GWE	HB2	TYR	0,0033	0,0103	-0,0018	-0,0009	0,0130
3ksq	CL	Z96	HD1	TYR	0,0035	0,0121	-0,0025	-0,0020	0,0167
3g4w	CL6	8CL	HB3	LEU	0,0093	0,0332	-0,0084	-0,0077	0,0493
3g4w	CL6	8CL	HG13	VAL	0,0037	0,0119	-0,0025	-0,0022	0,0166
3g4w	CL6	8CL	HD21	LEU	0,0053	0,0182	-0,0039	-0,0036	0,0257
3g4w	CL6	8CL	HB2	ASN	0,0065	0,0230	-0,0054	-0,0050	0,0334
3g4w	CL6	8CL	HE1	MET	0,0047	0,0166	-0,0031	-0,0023	0,0219
3jzi	CL28	JZL	HD3	LYS	0,0099	0,0369	-0,0089	-0,0078	0,0536
3jzi	CL28	JZL	HB2	GLU	0,0087	0,0327	-0,0079	-0,0073	0,0479
3jzi	CL28	JZL	HG2	LYS	0,0080	0,0293	-0,0060	-0,0048	0,0400
3jzi	CL28	JZL	HH	TYR	0,0070	0,0280	-0,0053	-0,0038	0,0371
2cf8	CL30	ESH	HA	ASN	0,0055	0,0238	-0,0043	-0,0038	0,0319
2cf8	CL30	ESH	HG	LEU	0,0031	0,0107	-0,0021	-0,0017	0,0145
2cf8	CL30	ESH	HG22	ILE	0,0022	0,0077	-0,0014	-0,0003	0,0093
2cf8	CL30	ESH	HD13	ILE	0,0073	0,0285	-0,0065	-0,0061	0,0411

3dle	CL1	GFA	HG12	VAL	0,0021	0,0067	-0,0011	-0,0010	0,0088
3dle	CL1	GFA	HB	VAL	0,0018	0,0056	-0,0007	-0,0006	0,0070
3dle	CL1	GFA	HB2	TYR	0,0031	0,0097	-0,0016	-0,0010	0,0122
2wpf	CL1	WPF	HG22	VAL	0,0024	0,0081	-0,0012	-0,0006	0,0099
2wpf	CL1	WPF	HD12	LEU	0,0137	0,0489	-0,0133	-0,0123	0,0745
2wpf	CL1	WPF	HA	SER	0,0061	0,0204	-0,0049	-0,0048	0,0301
2bla	CL1	CP6	HG13	ILE	0,0040	0,0143	-0,0025	-0,0018	0,0186
2bla	CL1	CP6	HD22	ASN	0,0077	0,0283	-0,0055	-0,0041	0,0380
2bla	CL1	CP6	HB2	ASN	0,0116	0,0415	-0,0110	-0,0092	0,0617
2wpe	CL	WPE	HD12	LEU	0,0095	0,0325	-0,0084	-0,0080	0,0490
2wpe	CL	WPE	HE2	TYR	0,0077	0,0273	-0,0063	-0,0061	0,0397
2wpe	CL	WPE	HA	SER	0,0060	0,0202	-0,0047	-0,0044	0,0294
1kzn	CL1	CBN	HA	ILE	0,0053	0,0182	-0,0037	-0,0026	0,0245
1kzn	CL1	CBN	HD3	PRO	0,0096	0,0324	-0,0087	-0,0081	0,0491
1kzn	CL1	CBN	HB3	ARG	0,0028	0,0088	-0,0018	-0,0016	0,0122
1xkk	CL3	FMM	HB3	LYS	0,0096	0,0375	-0,0074	-0,0054	0,0503
1xkk	CL3	FMM	HB3	LEU	0,0067	0,0238	-0,0052	-0,0046	0,0336
1xkk	CL3	FMM	HG22	THR	0,0076	0,0280	-0,0062	-0,0053	0,0396
2wp6	CL	WP6	HA	SER	0,0040	0,0132	-0,0027	-0,0021	0,0180
2wp6	CL	WP6	HA3	GLY	0,0040	0,0144	-0,0026	-0,0013	0,0183
2wp6	CL	WP6	HE2	TYR	0,0071	0,0252	-0,0055	-0,0047	0,0353
2wp6	CL	WP6	HD12	LEU	0,0072	0,0246	-0,0056	-0,0056	0,0358
2wpc	CLA	WP7	HG22	VAL	0,0021	0,0072	-0,0011	-0,0002	0,0085
2wpc	CLA	WP7	HA	SER	0,0048	0,0157	-0,0035	-0,0025	0,0217
2wpc	CLA	WP7	HA3	GLY	0,0028	0,0091	-0,0014	-0,0008	0,0114
2wpc	CLA	WP7	HE2	TYR	0,0057	0,0201	-0,0042	-0,0039	0,0282
2wpc	CLA	WP7	HD12	LEU	0,0087	0,0293	-0,0074	-0,0072	0,0439
2bl9	CL1	CP6	HG12	ILE	0,0060	0,0238	-0,0041	-0,0031	0,0309
2bl9	CL1	CP6	HD12	ILE	0,0051	0,0184	-0,0038	-0,0028	0,0250
2bl9	CL1	CP6	HB2	SER	0,0084	0,0300	-0,0076	-0,0068	0,0443
2uw8	CL1	GVQ	HA	ARG	0,0108	0,0409	-0,0090	-0,0070	0,0569
1j3j	CL1	CP6	HB2	ASN	0,0061	0,0226	-0,0037	-0,0014	0,0277
1j3j	CL1	CP6	HD12	ILE	0,0042	0,0149	-0,0030	-0,0021	0,0200
1j3j	CL1	CP6	HG12	ILE	0,0049	0,0190	-0,0029	-0,0017	0,0236
5tri	CL1	7HM	HB2	MET	0,0026	0,0085	-0,0016	-0,0013	0,0115
5tri	CL1	7HM	HB3	LEU	0,0083	0,0300	-0,0070	-0,0063	0,0433
5tri	CL1	7HM	HB2	PRO	0,0119	0,0421	-0,0116	-0,0109	0,0646
5tri	CL1	7HM	HD12	LEU	0,0047	0,0162	-0,0023	-0,0018	0,0203
5vcw	CL1	93J	HB2	LEU	0,0023	0,0073	-0,0010	-0,0007	0,0090

5vcw	CL1	93J	HG21	THR	0,0109	0,0401	-0,0098	-0,0093	0,0592
5vcw	CL1	93J	H	LYS	0,0072	0,0272	-0,0054	-0,0031	0,0358
5vcw	CL1	93J	HB3	LYS	0,0067	0,0242	-0,0046	-0,0038	0,0326
5tm4	CL1	7E3	HB2	HIS	0,0071	0,0282	-0,0047	-0,0031	0,0360
5tm4	CL1	7E3	HG12	VAL	0,0013	0,0041	-0,0007	-0,0002	0,0049
5tm4	CL1	7E3	HB3	MET	0,0022	0,0072	-0,0011	-0,0005	0,0087
5o0h	CL	9FN	HA2	GLY	0,0009	0,0027	-0,0005	-0,0004	0,0036
5o0h	CL	9FN	HA	LEU	0,0062	0,0242	-0,0042	-0,0025	0,0310
5o0h	CL	9FN	HB3	ALA	0,0068	0,0263	-0,0053	-0,0036	0,0352
5o0h	CL	9FN	HB3	LEU	0,0066	0,0245	-0,0045	-0,0025	0,0315
5o0h	CL	9FN	HB3	PRO	0,0016	0,0049	-0,0005	0,0000	0,0054
5h2u	CL	1N1	H	THR	0,0047	0,0177	-0,0032	-0,0015	0,0223
5h2u	CL	1N1	HB1	ALA	0,0046	0,0170	-0,0030	-0,0020	0,0220
5h2u	CL	1N1	HG2	LYS	0,0042	0,0147	-0,0028	-0,0017	0,0191
5icp	CL1	69Z	HD2	ARG	0,0049	0,0159	-0,0037	-0,0031	0,0228
5icp	CL1	69Z	HA2	GLY	0,0040	0,0142	-0,0026	-0,0016	0,0183
5icp	CL1	69Z	HG23	VAL	0,0047	0,0160	-0,0032	-0,0029	0,0221
5kgt	CL	6SQ	H	ASP	0,0043	0,0160	-0,0026	-0,0013	0,0198
5kgt	CL	6SQ	HB3	TYR	0,0036	0,0118	-0,0023	-0,0021	0,0162
5kgt	CL	6SQ	HB2	CYS	0,0043	0,0141	-0,0030	-0,0027	0,0198
5kgt	CL	6SQ	HB2	ASP	0,0076	0,0280	-0,0063	-0,0058	0,0402
5kgt	CL	6SQ	HA3	GLY	0,0044	0,0157	-0,0033	-0,0028	0,0217
5kgt	CL	6SQ	HA2	GLY	0,0057	0,0201	-0,0042	-0,0040	0,0282
5mzi	CL	FYK	HB2	PHE	0,0039	0,0128	-0,0025	-0,0023	0,0175
5mzi	CL	FYK	HB2	ALA	0,0168	0,0613	-0,0174	-0,0162	0,0949
5mzi	CL	FYK	HD21	LEU	0,0015	0,0043	-0,0006	-0,0005	0,0054
5mzi	CL	FYK	HG11	VAL	0,0038	0,0124	-0,0027	-0,0024	0,0174
5mzi	CL	FYK	HG21	VAL	0,0025	0,0077	-0,0012	-0,0009	0,0098
5mzc	CL	8EQ	HB2	PHE	0,0048	0,0160	-0,0033	-0,0032	0,0226
5mzc	CL	8EQ	HB3	LEU	0,0055	0,0191	-0,0043	-0,0038	0,0272
5mzc	CL	8EQ	HB2	ALA	0,0156	0,0558	-0,0163	-0,0151	0,0872
5mzc	CL	8EQ	HG11	VAL	0,0048	0,0165	-0,0036	-0,0034	0,0235
5mzc	CL	8EQ	HA	VAL	0,0022	0,0067	-0,0012	-0,0007	0,0087
5lof	CL1	70R	HB3	MET	0,0047	0,0166	-0,0035	-0,0028	0,0229
5lof	CL1	70R	HB3	ALA	0,0050	0,0178	-0,0037	-0,0032	0,0247
5lof	CL1	70R	HD1	PHE	0,0053	0,0191	-0,0041	-0,0036	0,0268
5cbi	CL06	4ZC	HG2	PRO	0,0035	0,0117	-0,0023	-0,0021	0,0162
5cbi	CL06	4ZC	HB1	ALA	0,0033	0,0113	-0,0018	-0,0011	0,0142
5cbi	CL06	4ZC	HE1	MET	0,0034	0,0115	-0,0021	-0,0017	0,0153

5fh8	CL	5XK	HB2	ALA	0,0089	0,0317	-0,0081	-0,0074	0,0472
5aeh	CL	8IR	HE1	HIS	0,0034	0,0117	-0,0021	-0,0014	0,0151
5aeh	CL	8IR	HB2	SER	0,0041	0,0161	-0,0029	-0,0021	0,0211
5aeh	CL	8IR	HB2	TYR	0,0040	0,0145	-0,0026	-0,0022	0,0193
5fck	CL	5WC	HA3	GLY	0,0084	0,0298	-0,0072	-0,0066	0,0435
5fck	CL	5WC	HA2	GLY	0,0067	0,0225	-0,0055	-0,0049	0,0329
5fck	CL	5WC	HG3	ARG	0,0045	0,0148	-0,0026	-0,0023	0,0196
5fck	CL	5WC	HD2	HIS	0,0060	0,0219	-0,0048	-0,0048	0,0315
5fck	CL	5WC	HD3	ARG	0,0050	0,0189	-0,0032	-0,0022	0,0244
5fh7	CL	5XL	HG2	MET	0,0051	0,0184	-0,0035	-0,0012	0,0231
5fh7	CL	5XL	HB2	ALA	0,0088	0,0311	-0,0079	-0,0072	0,0461
4qxi	CL4	I98	HA	TYR	0,0041	0,0152	-0,0022	-0,0012	0,0186
4qxi	CL4	I98	HD1	TYR	0,0078	0,0270	-0,0065	-0,0061	0,0397
4qxi	CL4	I98	HG11	VAL	0,0069	0,0233	-0,0055	-0,0050	0,0338
4qxi	CL4	I98	HG21	VAL	0,0047	0,0166	-0,0031	-0,0029	0,0227
4qxm	CL1	713	HB2	MET	0,0043	0,0181	-0,0027	-0,0026	0,0233
4qxm	CL1	713	HG3	MET	0,0038	0,0149	-0,0023	-0,0020	0,0192
4qxm	CL1	713	HA	GLN	0,0065	0,0226	-0,0054	-0,0052	0,0333
5aep	CL	QUP	HA	SER	0,0095	0,0356	-0,0081	-0,0072	0,0509
5aep	CL	QUP	HG23	VAL	0,0060	0,0200	-0,0043	-0,0042	0,0286
5aep	CL	QUP	HA	LYS	0,0058	0,0216	-0,0038	-0,0027	0,0281
5aep	CL	QUP	HE2	LYS	0,0017	0,0050	-0,0010	-0,0010	0,0070
4j3l	CL	AJ5	HA	SER	0,0046	0,0157	-0,0032	-0,0009	0,0198
4j3l	CL	AJ5	HD3	PRO	0,0106	0,0398	-0,0092	-0,0091	0,0580
4j3l	CL	AJ5	HA3	GLY	0,0071	0,0249	-0,0058	-0,0048	0,0354
4j3l	CL	AJ5	HB2	TYR	0,0115	0,0423	-0,0109	-0,0102	0,0635
4j3l	CL	AJ5	HD1	TYR	0,0063	0,0221	-0,0050	-0,0045	0,0316
4j22	CL	AJ7	HD1	TYR	0,0068	0,0240	-0,0056	-0,0050	0,0346
4j22	CL	AJ7	HB2	TYR	0,0119	0,0440	-0,0113	-0,0106	0,0659
4j22	CL	AJ7	HA3	GLY	0,0072	0,0252	-0,0060	-0,0049	0,0361
4j22	CL	AJ7	HA	SER	0,0053	0,0184	-0,0041	-0,0030	0,0255
4j22	CL	AJ7	HD3	PRO	0,0106	0,0389	-0,0097	-0,0092	0,0578
4xjo	CL9	41O	HB3	TYR	0,0040	0,0132	-0,0028	-0,0024	0,0183
4xjo	CL9	41O	HB2	CYS	0,0032	0,0103	-0,0021	-0,0019	0,0143
4xjo	CL9	41O	HB2	ASP	0,0037	0,0130	-0,0026	-0,0024	0,0180
4xjo	CL9	41O	HA3	GLY	0,0052	0,0189	-0,0041	-0,0036	0,0266
4xjo	CL9	41O	HA2	GLY	0,0032	0,0110	-0,0022	-0,0018	0,0151
4uj1	CL	NVX	HG2	LYS	0,0018	0,0056	-0,0008	-0,0007	0,0071
4w1x	CL1	3G9	HB3	TYR	0,0040	0,0132	-0,0028	-0,0023	0,0183

4w1x	CL1	3G9	HB2	CYS	0,0037	0,0120	-0,0025	-0,0023	0,0168
4w1x	CL1	3G9	HB2	ASP	0,0037	0,0127	-0,0025	-0,0023	0,0175
4w1x	CL1	3G9	HA3	GLY	0,0056	0,0200	-0,0044	-0,0039	0,0284
4w1x	CL1	3G9	HA2	GLY	0,0039	0,0136	-0,0028	-0,0024	0,0188
4o12	CL1	2QG	HD3	ARG	0,0066	0,0247	-0,0054	-0,0048	0,0350
4o12	CL1	2QG	HB2	ARG	0,0029	0,0094	-0,0015	-0,0008	0,0116
4ms4	CL	2C0	HB	ILE	0,0036	0,0127	-0,0021	-0,0014	0,0162
4ms4	CL	2C0	HG23	ILE	0,0054	0,0195	-0,0035	-0,0029	0,0259
4ms4	CL	2C0	HE1	TYR	0,0070	0,0259	-0,0060	-0,0055	0,0375
4ms4	CL	2C0	HE1	PHE	0,0037	0,0130	-0,0024	-0,0020	0,0175
4ms4	CL	2C0	HA	TYR	0,0076	0,0262	-0,0062	-0,0061	0,0385
4ms4	CL	2C0	HZ	PHE	0,0024	0,0083	-0,0014	-0,0007	0,0104
4ms4	CL	2C0	HD1	TYR	0,0042	0,0142	-0,0014	-0,0010	0,0166
3qgt	CL1	CP6	HB2	SER	0,0098	0,0339	-0,0090	-0,0085	0,0515
3qgt	CL1	CP6	HD12	ILE	0,0051	0,0183	-0,0037	-0,0025	0,0245
3qg2	CL1	CP6	HD12	ILE	0,0037	0,0129	-0,0025	-0,0015	0,0170
3qg2	CL1	CP6	HG12	ILE	0,0047	0,0181	-0,0028	-0,0016	0,0225
4fpf	CL	JKK	HB2	ASP	0,0064	0,0223	-0,0053	-0,0049	0,0326
4fpf	CL	JKK	HB3	SER	0,0039	0,0128	-0,0026	-0,0016	0,0170
4fpf	CL	JKK	H	ASN	0,0135	0,0523	-0,0134	-0,0102	0,0759
4fpf	CL	JKK	HE1	MET	0,0095	0,0351	-0,0070	-0,0043	0,0464
4fpf	CL	JKK	HB2	ASN	0,0051	0,0190	-0,0032	-0,0006	0,0228
4cjv	CL	IV2	HB2	ASN	0,0054	0,0192	-0,0040	-0,0037	0,0269
4cjv	CL	IV2	HE2	TYR	0,0047	0,0160	-0,0034	-0,0030	0,0225
3w33	CL1	W19	HB3	LEU	0,0057	0,0199	-0,0041	-0,0032	0,0273
3w33	CL1	W19	HG22	THR	0,0079	0,0293	-0,0064	-0,0054	0,0411
3v51	CLD	I76	HD2	HIS	0,0056	0,0245	-0,0043	-0,0032	0,0320
3v51	CLD	I76	HG3	GLU	0,0044	0,0150	-0,0031	-0,0029	0,0210
3v51	CLD	I76	HD13	LEU	0,0046	0,0166	-0,0030	-0,0022	0,0218
3utu	CL	1TS	HG23	ILE	0,0046	0,0157	-0,0035	-0,0032	0,0223
3utu	CL	1TS	HA	ASN	0,0076	0,0270	-0,0062	-0,0053	0,0385
3utu	CL	1TS	HB3	LEU	0,0044	0,0147	-0,0030	-0,0018	0,0195
3utu	CL	1TS	HD21	LEU	0,0034	0,0112	-0,0021	-0,0017	0,0150
3ug2	CL	IRE	HB3	LEU	0,0033	0,0110	-0,0017	-0,0012	0,0139
3ug2	CL	IRE	HG2	MET	0,0207	0,0754	-0,0229	-0,0221	0,1204
3ug2	CL	IRE	HB3	LYS	0,0088	0,0323	-0,0068	-0,0058	0,0449

Tabla A4: Datos geométricos de los átomos de hidrógeno de las estructuras utilizadas como sistema de referencia.

PDBid	Atomo 1	Residuo 1	Atomo 2	Residuo 2	Distancia (Å)	Angulo (°)	Coordenadas
3qfx	CL1	CP6	HG	LEU	2,431611	138,712688	[13,232 30,693 44,28]
3qfx	CL1	CP6	HG21	THR	2,917601	86,132441	[17,309 32,505 45,637]
3poz	CL	O3P	HB3	LEU	3,064441	121,166055	[15,711 32,118 47,046]
3poz	CL	O3P	HB3	LYS	2,93569	69,676008	[17,41 31,297 43,783]
3poz	CL	O3P	HG22	THR	2,753528	95,414647	[12,39 32,559 42,738]
3ms9	CL1	MS9	HB	ILE	3,05434	137,501776	[15,209 30,084 45,989]
3ms9	CL1	MS9	HB2	ALA	3,740311	91,856489	[14,504 29,6 41,789]
3ms9	CL1	MS9	HG22	ILE	3,158615	92,046822	[17,352 31,06 45,257]
3ms9	CL1	MS9	HD21	LEU	3,218418	116,623335	[12,107 31,04 43,039]
3ms9	CL1	MS9	HB2	ALA	2,932935	98,031119	[14,725 34,439 46,343]
3ms9	CL1	MS9	HB2	LEU	3,600981	137,309919	[11,299 32,105 45,086]
3ms7	CL1	22S	HB3	ALA	2,959698	154,872819	[12,893 32,092 46,458]
3ms7	CL1	22S	HA	ASN	2,997342	103,116256	[13,184 34,554 45,659]
3kmx	CL1	G00	HG3	GLN	3,124125	115,696648	[16,351 30,231 45,624]
3kmx	CL1	G00	HA2	GLY	2,858289	159,00804	[12,684 30,86 45,26]
3kmx	CL1	G00	HE22	GLN	3,05114	74,108322	[17,113 30,728 43,352]
3kmx	CL1	G00	HD1	PHE	2,99884	99,357331	[12,102 33,857 43,924]
3fal	CL10	LO2	HB2	LEU	3,029992	109,764543	[12,853 30,673 42,671]
3fal	CL10	LO2	HB2	ALA	2,914168	96,508826	[16,918 30,649 44,809]
3fal	CL10	LO2	HD1	PHE	3,369655	131,64396	[12,157 33,303 46,091]
3dlg	CL1	GWE	HG23	VAL	2,875143	119,271755	[13,211 30,22 43,274]
3dlg	CL1	GWE	HB2	TYR	3,397303	87,081564	[16,067 34,546 46,503]
3ksq	CL	Z96	HD1	TYR	3,228809	91,656224	[17,434 31,544 45,674]
3g4w	CL6	8CL	HB3	LEU	2,700165	120,524219	[15,659 30,079 44,974]
3g4w	CL6	8CL	HG13	VAL	3,204121	93,249994	[13,319 30,749 41,87]
3g4w	CL6	8CL	HD21	LEU	3,015539	84,95827	[17,448 31,847 45,351]
3g4w	CL6	8CL	HB2	ASN	2,874575	118,632135	[11,891 32,475 43,951]
3g4w	CL6	8CL	HE1	MET	3,139721	86,317171	[13,649 35,23 45,192]
3jzi	CL28	JZL	HD3	LYS	2,685352	119,693269	[15,391 29,905 44,564]
3jzi	CL28	JZL	HB2	GLU	2,687509	134,896433	[15,201 31,538 46,623]
3jzi	CL28	JZL	HG2	LYS	2,856936	77,615968	[15,634 30,499 42,285]
3jzi	CL28	JZL	HH	TYR	2,815856	82,932636	[13,106 32,018 41,91]
2cf8	CL30	ESH	HA	ASN	2,847522	153,892754	[14,429 30,445 46,141]
2cf8	CL30	ESH	HG	LEU	3,254311	75,494225	[17,377 30,657 43,555]
2cf8	CL30	ESH	HG22	ILE	3,426501	124,087854	[14,592 33,169 47,515]
2cf8	CL30	ESH	HD13	ILE	2,802585	117,901684	[12,996 33,843 45,871]
3dle	CL1	GFA	HG12	VAL	3,53728	115,723064	[14,758 29,022 43,604]
3dle	CL1	GFA	HB	VAL	3,649609	101,18584	[13,225 30,024 41,969]
3dle	CL1	GFA	HB2	TYR	3,427774	90,343173	[15,536 34,81 46,562]
2wpf	CL1	WPF	HG22	VAL	3,498785	136,462115	[11,616 30,828 44,221]
2wpf	CL1	WPF	HD12	LEU	2,527407	87,535462	[16,35 30,556 43,763]
2wpf	CL1	WPF	HA	SER	2,927787	88,446147	[16,483 33,433 46,297]
2bla	CL1	CP6	HG13	ILE	3,243874	116,535511	[16,076 32,042 47,037]
2bla	CL1	CP6	HD22	ASN	2,836681	94,429749	[13,594 30,928 42,069]
2bla	CL1	CP6	HB2	ASN	2,625188	114,651045	[12,195 32,545 43,677]

2wpe	CL	WPE	HD12	LEU	2,706815	87,071482	[15,916 30,27 43,117]
2wpe	CL	WPE	HE2	TYR	2,814509	98,654753	[13,181 30,974 42,37]
2wpe	CL	WPE	HA	SER	2,942166	88,422658	[16,296 33,757 46,313]
1kzn	CL1	CBN	HA	ILE	3,014167	119,150291	[16,031 30,087 45,485]
1kzn	CL1	CBN	HD3	PRO	2,702361	88,287034	[16,277 30,335 43,503]
1kzn	CL1	CBN	HB3	ARG	3,346438	113,071069	[11,484 32,323 43,381]
1xkk	CL3	FMM	HB3	LYS	2,745581	71,68339	[17,22 31,3 43,767]
1xkk	CL3	FMM	HB3	LEU	2,906275	118,170643	[15,879 32,22 46,804]
1xkk	CL3	FMM	HG22	THR	2,845246	88,921264	[12,452 32,915 42,548]
2wp6	CL	WP6	HA	SER	3,159107	86,387309	[13,339 31,159 41,654]
2wp6	CL	WP6	HA3	GLY	3,150511	129,126751	[13,766 33,269 47,042]
2wp6	CL	WP6	HE2	TYR	2,900589	103,099282	[16,154 32,938 46,651]
2wp6	CL	WP6	HD12	LEU	2,84454	80,332261	[13,797 35,076 44,804]
2wpc	CLA	WP7	HG22	VAL	3,60135	131,36612	[15,484 31,43 47,476]
2wpc	CLA	WP7	HA	SER	3,069429	89,396493	[13,103 31,402 41,775]
2wpc	CLA	WP7	HA3	GLY	3,366969	124,71837	[13,185 33,736 46,808]
2wpc	CLA	WP7	HE2	TYR	2,978977	104,732098	[15,623 33,464 46,794]
2wpc	CLA	WP7	HD12	LEU	2,75697	84,330806	[13,632 34,912 44,784]
2bl9	CL1	CP6	HG12	ILE	2,979343	138,187159	[15,292 30,758 46,465]
2bl9	CL1	CP6	HD12	ILE	3,069725	103,447471	[16,21 32,995 46,791]
2bl9	CL1	CP6	HB2	SER	2,768995	108,172941	[12,566 33,98 44,879]
2uw8	CL1	GVQ	HA	ARG	2,700661	106,425538	[12,19 32,584 43,329]
1j3j	CL1	CP6	HB2	ASN	3,083859	114,571303	[16,381 30,176 45,13]
1j3j	CL1	CP6	HD12	ILE	3,16715	106,783874	[13,112 30,241 42,311]
1j3j	CL1	CP6	HG12	ILE	3,123977	139,502247	[11,975 31,197 44,264]
5tri	CL1	7HM	HB2	MET	3,362212	90,777058	[17,482 30,631 44,821]
5tri	CL1	7HM	HB3	LEU	2,790444	113,163746	[12,035 33,077 44,151]
5tri	CL1	7HM	HB2	PRO	2,582881	105,433584	[14,95 33,778 46,377]
5tri	CL1	7HM	HD12	LEU	3,192298	81,3538	[13,022 35,164 44,243]
5vcw	CL1	93J	HB2	LEU	3,506473	122,167195	[13,519 29,309 43,331]
5vcw	CL1	93J	HG21	THR	2,645226	107,622507	[16,222 30,376 44,765]
5vcw	CL1	93J	H	LYS	2,821848	132,220039	[12,048 32,545 44,885]
5vcw	CL1	93J	HB3	LYS	2,941074	76,114934	[12,709 33,862 42,532]
5tm4	CL1	7E3	HB2	HIS	2,899092	71,688649	[17,51 31,556 44,161]
5tm4	CL1	7E3	HG12	VAL	3,781918	90,892913	[12,081 31,829 41,554]
5tm4	CL1	7E3	HB3	MET	3,518478	81,095268	[12,157 34,482 42,861]
5o0h	CL	9FN	HA2	GLY	3,956694	105,988283	[13,397 29,324 42,225]
5o0h	CL	9FN	HA	LEU	3,005191	93,773206	[16,846 30,338 44,375]
5o0h	CL	9FN	HB3	ALA	2,909819	132,67782	[11,982 31,653 44,27]
5o0h	CL	9FN	HB3	LEU	2,930789	68,651003	[17,615 32,563 44,836]
5o0h	CL	9FN	HB3	PRO	3,872736	72,279085	[14,228 36,251 44,932]
5h2u	CL	1N1	H	THR	3,030631	112,489478	[16,332 30,129 45,19]
5h2u	CL	1N1	HB1	ALA	3,149252	117,780531	[15,648 32,608 47,148]
5h2u	CL	1N1	HG2	LYS	3,135288	72,673924	[12,762 34,417 42,672]
5icp	CL1	69Z	HD2	ARG	3,060216	119,058058	[11,758 32,062 43,675]
5icp	CL1	69Z	HA2	GLY	3,205807	84,226653	[16,655 33,87 46,319]
5icp	CL1	69Z	HG23	VAL	3,110602	94,608162	[13,377 34,941 45,456]
5kgf	CL	6SQ	H	ASP	3,058484	139,639766	[13,093 29,873 44,289]
5kgf	CL	6SQ	HB3	TYR	3,219969	96,684404	[17,234 30,884 45,453]

5kgt	CL	6SQ	HB2	CYS	3,108156	90,55944	[14,229 30,269 42,022]
5kgt	CL	6SQ	HB2	ASP	2,782292	139,617385	[12,113 32,17 45,076]
5kgt	CL	6SQ	HA3	GLY	3,12298	104,750073	[15,388 33,72 46,953]
5kgt	CL	6SQ	HA2	GLY	2,969888	80,228325	[13,076 34,92 44,117]
5mzi	CL	FYK	HB2	PHE	3,180226	97,691005	[16,055 29,659 43,513]
5mzi	CL	FYK	HB2	ALA	2,39805	129,883349	[15,007 32,357 46,53]
5mzi	CL	FYK	HD21	LEU	3,776966	97,077834	[11,699 31,952 42,015]
5mzi	CL	FYK	HG11	VAL	3,171587	123,405065	[12,58 33,753 46,07]
5mzi	CL	FYK	HG21	VAL	3,457002	93,083436	[14,717 35,072 46,454]
5mzc	CL	8EQ	HB2	PHE	3,057974	88,064069	[16,237 30,007 43,156]
5mzc	CL	8EQ	HB3	LEU	2,985429	121,857955	[12,107 31,324 43,486]
5mzc	CL	8EQ	HB2	ALA	2,440383	120,98959	[15,532 32,281 46,456]
5mzc	CL	8EQ	HG11	VAL	3,026168	129,305059	[12,849 33,458 46,281]
5mzc	CL	8EQ	HA	VAL	3,491966	85,734299	[13,33 35,548 45,059]
5lof	CL1	70R	HB3	MET	3,092539	108,580123	[16,322 29,886 44,751]
5lof	CL1	70R	HB3	ALA	3,05003	116,668059	[12,227 31,135 43,112]
5lof	CL1	70R	HD1	PHE	2,986723	125,967984	[12,814 33,559 46,142]
5cbi	CL06	4ZC	HG2	PRO	3,255595	95,250611	[17,33 30,835 45,289]
5cbi	CL06	4ZC	HB1	ALA	3,356403	108,817152	[16,489 32,426 47,048]
5cbi	CL06	4ZC	HE1	MET	3,298134	93,838516	[12,68 34,958 44,854]
5fh8	CL	5XK	HB2	ALA	2,729055	113,606954	[13,228 33,997 45,804]
5aeh	CL	8IR	HE1	HIS	3,336823	119,808954	[14,568 29,143 43,899]
5aeh	CL	8IR	HB2	SER	3,038746	145,256315	[12,151 30,954 44,718]
5aeh	CL	8IR	HB2	TYR	3,174131	102,012311	[13,258 34,737 45,839]
5fck	CL	5WC	HA3	GLY	2,754692	122,714039	[15,822 30,365 45,569]
5fck	CL	5WC	HA2	GLY	2,888042	88,878308	[17,22 31,993 45,607]
5fck	CL	5WC	HG3	ARG	3,175705	85,012383	[13,159 31,363 41,68]
5fck	CL	5WC	HD2	HIS	2,906781	105,633535	[14,761 34,13 46,581]
5fck	CL	5WC	HD3	ARG	3,103837	85,185507	[12,261 33,483 42,627]
5fh7	CL	5XL	HG2	MET	3,148492	102,867552	[13,299 30,413 42,252]
5fh7	CL	5XL	HB2	ALA	2,742349	114,520822	[13,15 33,956 45,784]
4qxi	CL4	I98	HA	TYR	3,061563	154,196859	[13,286 29,987 45,14]
4qxi	CL4	I98	HD1	TYR	2,799275	114,354764	[14,37 29,809 43,437]
4qxi	CL4	I98	HG11	VAL	2,887276	94,774343	[16,502 30,189 44,035]
4qxi	CL4	I98	HG21	VAL	3,06407	92,754109	[17,247 31,881 45,838]
4qxm	CL1	713	HB2	MET	2,988324	167,843784	[13,035 30,598 45,65]
4qxm	CL1	713	HG3	MET	3,14258	143,114711	[14,939 31,25 47,022]
4qxm	CL1	713	HA	GLN	2,877817	99,775616	[12,155 33,623 43,734]
5aep	CL	QUP	HA	SER	2,669742	141,823441	[14,03 32,393 46,746]
5aep	CL	QUP	HG23	VAL	2,955458	80,100431	[17,253 32,786 45,678]
5aep	CL	QUP	HA	LYS	3,021229	128,324213	[12,351 33,315 45,821]
5aep	CL	QUP	HE2	LYS	3,599157	98,817247	[16,797 29,51 44,241]
4j3l	CL	AJ5	HA	SER	3,086185	122,844237	[15,863 29,935 45,534]
4j3l	CL	AJ5	HD3	PRO	2,645077	120,866889	[14,667 29,805 43,9]
4j3l	CL	AJ5	HA3	GLY	2,898106	110,917019	[12,697 30,932 42,714]
4j3l	CL	AJ5	HB2	TYR	2,571179	130,82986	[14,162 32,778 46,621]
4j3l	CL	AJ5	HD1	TYR	2,925008	96,660953	[13,198 34,634 45,322]
4j22	CL	AJ7	HD1	TYR	2,88286	98,219425	[15,71 29,871 43,478]
4j22	CL	AJ7	HB2	TYR	2,552893	132,306222	[13,382 30,364 44,027]

4j22	CL	AJ7	HA3	GLY	2,887745	110,233978	[15,912 32,901 46,802]
4j22	CL	AJ7	HA	SER	2,987848	119,568743	[12,003 33,358 44,927]
4j22	CL	AJ7	HD3	PRO	2,639001	117,228516	[13,638 33,861 46,142]
4xjo	CL9	41O	HB3	TYR	3,167444	99,120308	[17,105 30,695 45,293]
4xjo	CL9	41O	HB2	CYS	3,258265	96,878289	[13,725 30,214 42,049]
4xjo	CL9	41O	HB2	ASP	3,146166	138,657605	[11,783 32,176 45,107]
4xjo	CL9	41O	HA3	GLY	3,018638	109,084302	[15,344 33,495 46,926]
4xjo	CL9	41O	HA2	GLY	3,271553	92,476491	[12,555 34,867 44,556]
4uj1	CL	NVX	HG2	LYS	3,607349	77,895763	[18,143 31,175 44,407]
4w1x	CL1	3G9	HB3	TYR	3,172421	98,555989	[17,149 30,808 45,387]
4w1x	CL1	3G9	HB2	CYS	3,182262	94,543461	[13,937 30,232 42,05]
4w1x	CL1	3G9	HB2	ASP	3,160964	137,731677	[11,733 32,109 45,002]
4w1x	CL1	3G9	HA3	GLY	2,983828	109,582006	[15,335 33,448 46,911]
4w1x	CL1	3G9	HA2	GLY	3,168135	91,588476	[12,579 34,766 44,433]
4o12	CL1	2QG	HD3	ARG	2,868624	121,956724	[14,97 29,629 44,272]
4o12	CL1	2QG	HB2	ARG	3,399131	80,283583	[15,546 30,052 41,903]
4ms4	CL	2C0	HB	ILE	3,254273	118,437054	[15,54 29,347 44,63]
4ms4	CL	2C0	HG23	ILE	3,039648	74,226195	[16,404 30,449 42,606]
4ms4	CL	2C0	HE1	TYR	2,801356	134,034699	[14,976 31,948 46,9]
4ms4	CL	2C0	HE1	PHE	3,208178	88,791136	[12,615 31,868 41,83]
4ms4	CL	2C0	HA	TYR	2,80731	115,965377	[13,164 33,941 45,951]
4ms4	CL	2C0	HZ	PHE	3,467619	92,191491	[11,757 33,836 43,042]
4ms4	CL	2C0	HD1	TYR	3,279411	70,556635	[14,259 35,663 44,706]
3qgt	CL1	CP6	HB2	SER	2,69869	104,813227	[16,355 30,408 44,707]
3qgt	CL1	CP6	HD12	ILE	3,078843	104,644417	[12,526 31,18 42,594]
3qg2	CL1	CP6	HD12	ILE	3,241158	108,603798	[13,173 30,128 42,734]
3qg2	CL1	CP6	HG12	ILE	3,148627	137,966023	[11,953 31,262 44,528]
4fpf	CL	JKK	HB2	ASP	2,896957	92,851319	[15,915 30,004 43,187]
4fpf	CL	JKK	HB3	SER	3,192161	115,356506	[16,225 31,841 46,892]
4fpf	CL	JKK	H	ASN	2,492021	122,616212	[12,5 33,165 44,922]
4fpf	CL	JKK	HE1	MET	2,847709	90,347845	[14,885 34,596 46,042]
4fpf	CL	JKK	HB2	ASN	3,089876	66,341487	[13,423 35,173 43,387]
4cjv	CL	IV2	HB2	ASN	3,012272	108,942372	[16,487 30,229 45,118]
4cjv	CL	IV2	HE2	TYR	3,05893	89,070392	[14,732 30,169 42,166]
3w33	CL1	W19	HB3	LEU	3,031385	116,375903	[12,057 31,477 43,203]
3w33	CL1	W19	HG22	THR	2,834817	88,573771	[17,119 32,194 45,683]
3v51	CLD	I76	HD2	HIS	2,840529	161,796538	[13,618 30,316 45,537]
3v51	CLD	I76	HG3	GLU	3,111751	128,311299	[11,756 31,744 44,072]
3v51	CLD	I76	HD13	LEU	3,129613	78,996832	[12,858 34,938 43,695]
3utu	CL	1TS	HG23	ILE	3,067169	116,121163	[15,753 29,644 44,657]
3utu	CL	1TS	HA	ASN	2,870332	99,9031	[16,701 32,37 46,26]
3utu	CL	1TS	HB3	LEU	3,115318	116,26172	[13,267 34,146 46,312]
3utu	CL	1TS	HD21	LEU	3,258533	76,408139	[13,386 35,45 44,235]
3ug2	CL	IRE	HB3	LEU	3,320651	118,21335	[12,539 30,281 42,869]
3ug2	CL	IRE	HG2	MET	2,303762	110,007032	[15,935 30,587 44,782]
3ug2	CL	IRE	HB3	LYS	2,774462	72,812948	[12,934 33,655 42,402]

ANEXO B

Docking molecular de ligandos halogenados con receptores de dopamina

Construcción de los modelos por homología del D₂R

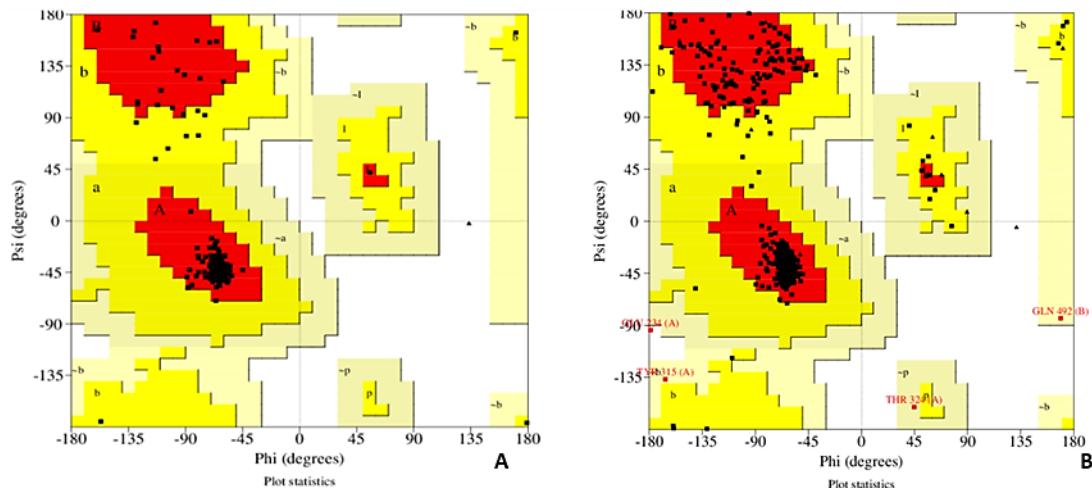


Figura B1: Gráficos de Ramachandran de la validación con PROCHECK de los modelos por homología del receptor D₂. A) corresponde al modelo inactivado y B) al modelo del receptor activado.

Ánálisis de los modelos de homología

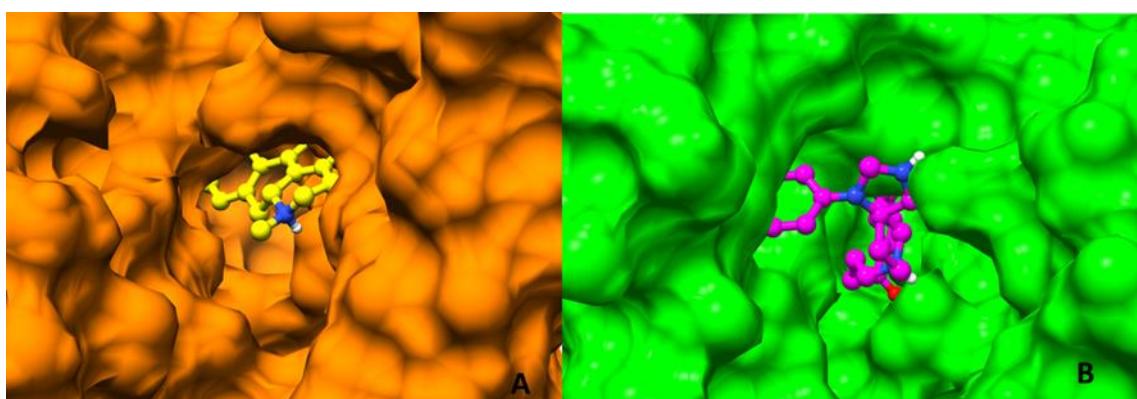


Figura B2: A- superficie del modelo de D₂R en su forma activa. B- superficie del receptor D₂R en su forma inactiva

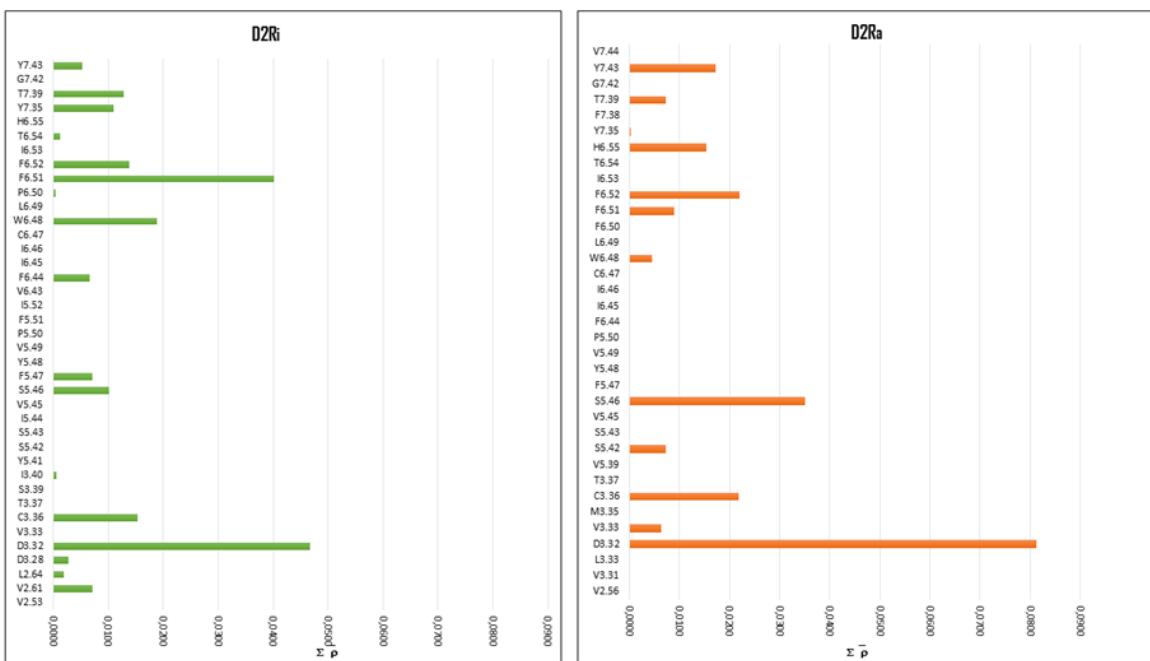


Figure B3: Suma de los valores de densidad electrónica (en unidades atómicas) para las interacciones entre los residuos del receptor y el ligando. (A) D2Ra y apomorfina y (B) D2Ri y spiperona.

A partir nuestro estudio in silico actual del sitio de unión de D2R, se han determinado las interacciones particulares vinculadas a la actividad de los compuestos (Tabla B1); por ejemplo los átomos de nitrógeno protonados están involucrados en interacciones electrostáticas fuertes con un residuo de ácido aspártico conservado en TM3 (D3. 32), el grupo carboxilo del ácido aspártico forma un estrecho puente salino con el grupo amino primario de todos los ligandos, agonistas y agonistas inversos. Si bien en ambas formas del receptor los residuos que están involucrados en las interacciones más importantes son los mismos, se puede observar que en la forma inactiva prevalece el tipo de interacciones hidrofóbicas.

Tabla B 1: Interacciones ligando-receptor de D2Ra con apomorfina y D2Ri con spiperona.

D2Ra		D2Ri	
Enlaces de Hidrogeno	Interacciones hidrofóbicas	Enlaces de Hidrogeno	Interacciones hidrofóbicas
D3.32, S5.46, C3.36, H6.55, F6.52, F5.42, F6.52	Y7.43, F6.52, C3.36, V3.33, H6.55, F6.51, T7.39, W6.48, V2.56	D3.32, S5.46, C3.36, T7.39, Y7.35, F3.28, L2.64, T6.54	F6.51, F6.52, W6.48, C3.36, V2.61, F7.47, F6.44, T7.39, Y7.35, F3.28, L2.64, T6.54, P6.50

Las moléculas de agua también han sido implicadas en la activación de los GPCRs (Trzaskowski et al, 2012). En la forma inactivada de D2R, no se observan moléculas de agua en el lado

citoplasmático del canal central del receptor (Figura B4) debido a la existencia de una barrera hidrofóbica que impide el paso de moléculas de agua desde el lado extracelular hacia esa región. La barrera hidrofóbica en D2R está formada por seis residuos de las hélices TM2, TM3 y TM6: L2.43, L2.46, I3.46, L3.43, V6.40, L6.37 y M6.36. En el receptor activado, TM6 se desplaza hacia afuera en el lado citoplasmático y la barrera hidrofóbica se abre permitiendo el paso de moléculas de agua hacia la parte intracelular del canal central de D2R (Figura B4).

Para explicar cómo ocurriría la apertura de la barrera hidrofóbica se ha propuesto un mecanismo que involucra a las secuencias conservadas cwxP y nPxxy (Standfuss et al, 2011); en D2R no se observan los cambios en la red de enlaces de hidrógeno mediados por moléculas de agua descriptos por Standfuss et al en rodopsina. Sin embargo, se observan cambios conformacionales en los segmentos TM2 y TM7 que juntos actuarían como una “compuerta” regulando el ingreso de moléculas de agua al canal central del ovillo de TMs.

Como se discutió previamente, TM7 muestra un marcado desplazamiento hacia TM6 en su mitad extracelular al pasar de la forma inactivada a la activada del receptor. Este desplazamiento de TM7 ocasiona la ruptura de un puente entre Ser7.36 y Asp2.65 que mantiene proximos a los segmentos TM7 y TM2, tras lo cual TM2 se desplaza en sentido contrario a TM7. Estos cambios se muestran en la Figura B3A.

Además de la ruptura del puente Ser7.36 y Asp2.65 otras interacciones se ven afectadas por el desplazamiento de TM2 y TM7. El desplazamiento de TM7 causa la ruptura del puente de hidrógeno mediado por agua entre Tyr7.43 y Asp3.32 (Figura B4 A,B) abriendo de esta manera el canal central del receptor para el ingreso de moléculas de agua (Figura B4 C). Es importante destacar que la ruptura de esta interacción se verifica también en la activación de otros GPCRs y forma parte del denominado interruptor de bloqueo 3-7.

Asimismo, el desplazamiento de TM2 aumenta el diámetro del canal central del receptor facilitando el ingreso de moléculas de agua y ademas permite la rotación de la cadena lateral de Met3.35 que en la forma inactivada se orienta hacia el interior bloqueando la luz del canal central pero que en la forma activada se dirige hacia afuera entre los segmentos TM3 y TM2 debido a la mayor separación entre ambos segmentos (Figura B4 A, C).

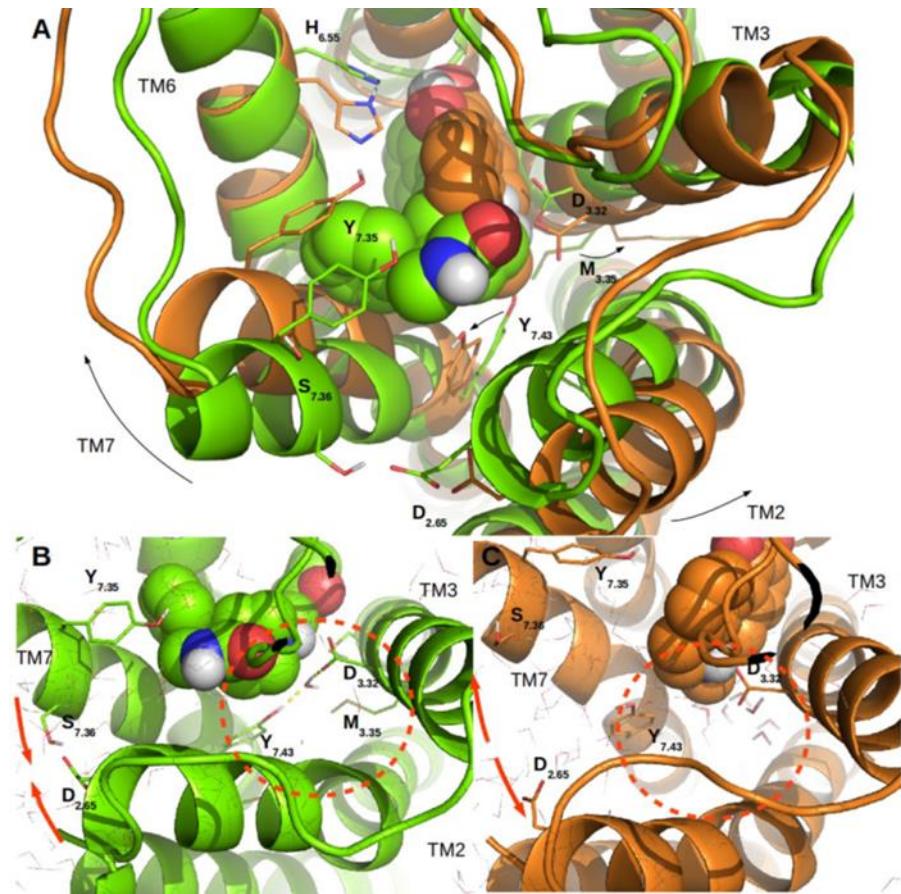


Figura B4: Cambios conformativos en el lado extracelular de los segmentos TM2 y TM7. Se muestran las formas inactiva (verde) y activa (naranja) de D2R superpuestas (A) y por separado (B,C). El desplazamiento de TM7 y TM2 en sentido opuesto abre el canal central de D2R permitiendo el paso de moléculas de agua.

Refinamiento de los modelos de homología guiado por el ligando

Para corroborar los datos obtenidos a través del docking con un agonista y un agonista inverso, y evaluar la calidad y poder predictivo de cada modelo, se obtuvo las estructuras de varios ligandos agonistas y agonistas inversos del D2R de la base de datos ChEMBL, y se realizó docking de los mismos con ambas formas del receptor, activa e inactiva, para poder evaluar posteriormente la energía y el modo de unión con los correspondientes receptores.

Los ligandos y los receptores fueron convertidos de pdb a formato pdbqt con MGLTools, las simulaciones fueron realizadas utilizando AutoDock Vina. De los diferentes complejos obtenidos del Docking de cada receptor, la conformación de menor energía de cada uno fue considerada como la orientación más favorable.

De la base de datos ChEMBL se obtuvo la estructura de compuestos agonistas y agonistas inversos.

Tabla B2 : Lista de ligando de D2R utilizados para realizar el Docking

Agonistas Inversos	Agonistas
Pimozide	Bromocriptina
Thiothixene	Quinpirole
Risperidone	Cabergoline
Spiperone	PPHT
Tefludazine	Rotigotine
Raclopride	Ropinirole
Butaclamol	Apomorphine
Clozapine	Dopamina
Haloperidol	Pergolide
Clorpromazine	Pramipexole

A través del docking molecular realizado con diferentes ligandos, agonistas y agonistas inversos, se pudo observar las diferentes poses de unión con los dos tipos de receptores, activo e inactivo, comprobando que los agonistas conservan el modo de unión convencional al unirse al receptor activado solamente.

En la figura B5 se pueden ver las poses de menor energía resultado del Docking de los diferentes agonistas con el receptor activo e inactivo superpuestas, además se observan los residuos más relevantes del sitio de unión del D2R activado (D3.32, S5.43, S5.42 y S5.46). Como se puede observar en las imágenes los ligandos agonistas unidos al receptor activado mantienen las interacciones con estos residuos, características de los receptores dopaminérgicos, el puente iónico entre el nitrógeno protonado y el residuo D3.32, y las interacciones con los grupos OH de las serinas. Sin embargo, en las poses del docking realizado con el receptor inactivado se puede ver que las orientaciones de los mismos son poco precisas, sin conservar estas proximidades con los residuos del sitio de unión.

En la figura B6 se observan las poses más estables del docking con los ligandos agonistas inversos con ambos receptores, superpuestas. En este caso se observa que en las poses del docking efectuado con el D2R inactivo los ligandos tienden a ubicarse siempre próximos al segmento TM6, en un bolsillo conformado por los residuos W6.48, F6.44 y F6.52; mientras que

con el D2R activado las orientaciones de los agonistas inversos no son estables, sin permanecer en un único sitio de unión.

Analizando las poses obtenidas de los ligandos se puede decir que los agonistas se ubican apropiadamente en el sitio de unión al receptor activado, mientras que los ligandos agonistas inversos se unen de una manera más consistente estando el receptor en su forma inactiva

Las figuras B7 y B8 muestran las diferentes interacciones de los ligandos spiperona y apomorfina en las poses obtenidas del docking con las diferentes formas del receptor, D2Ri y D2Ra.

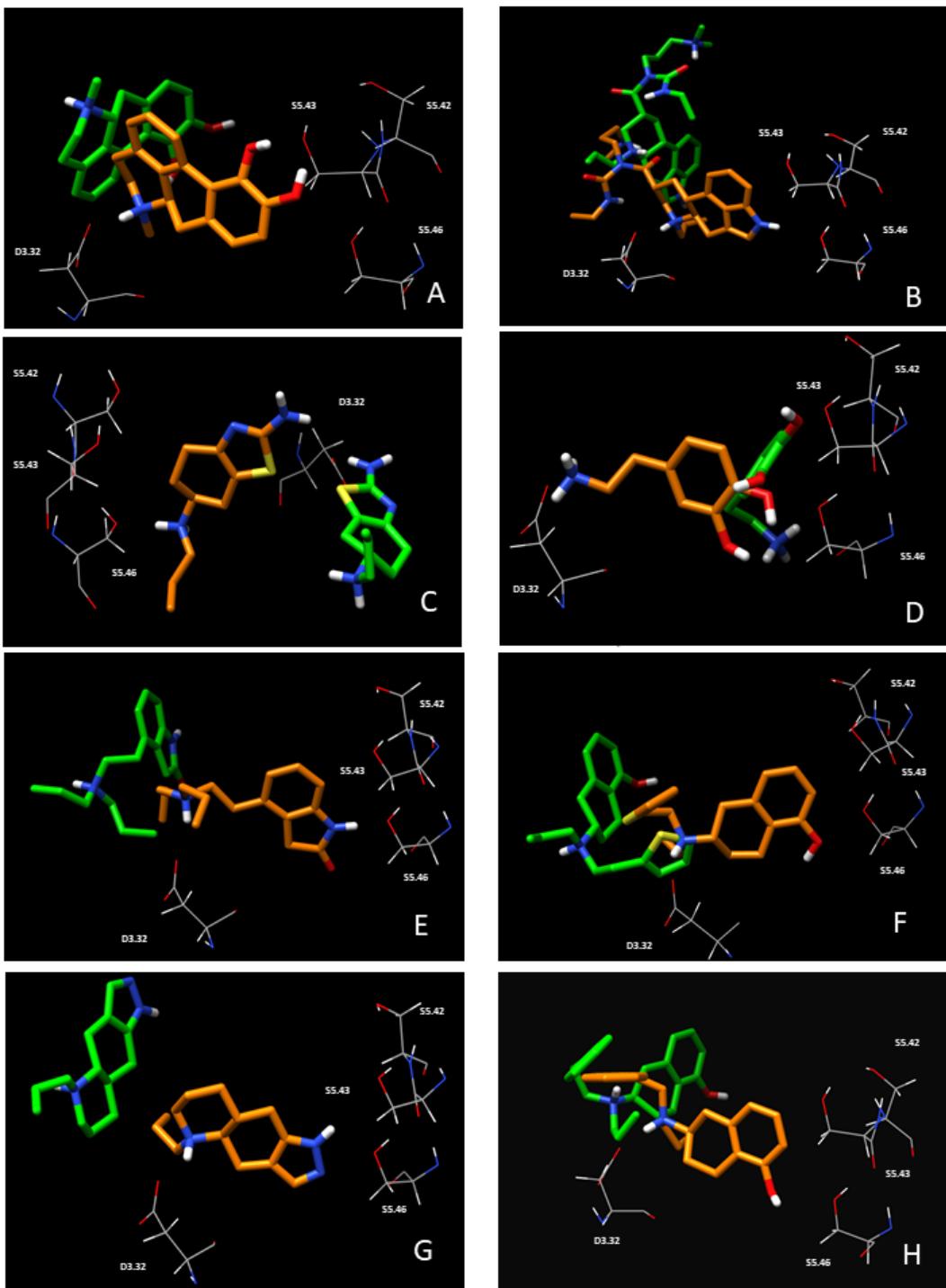


Figura B5: Poses de los ligandos agonistas unidos al receptor activado en color naranja y en color verde unidos al receptor inactivado. A) Apomorfina B) Cabergolina C) Dopamina D) Pramipexole E) Ropinirole F) Rotigotine G)Quinpirole H)PPTH

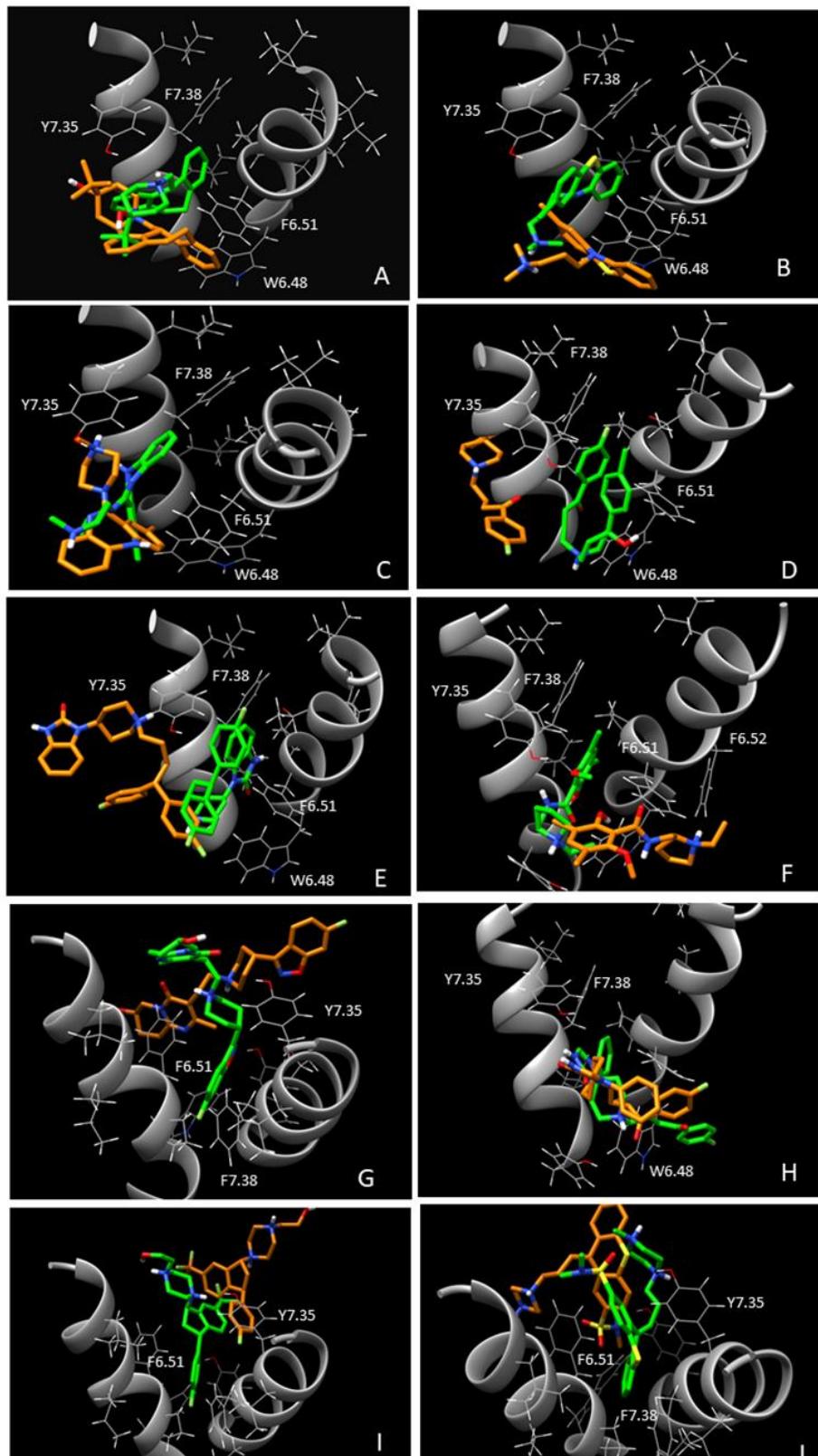


Figura B6: Poses de los ligandos agonistas inversos unidos al receptor activado en color naranja y en color verde unidos al receptor inactivado. A) Butaclamol B) Clorpromazina C) Clozapina D) Haloperidol E) Pimozide F) Raclopride G) Risperidona H) Spiperona I) Tefludazine J) Thiothixene

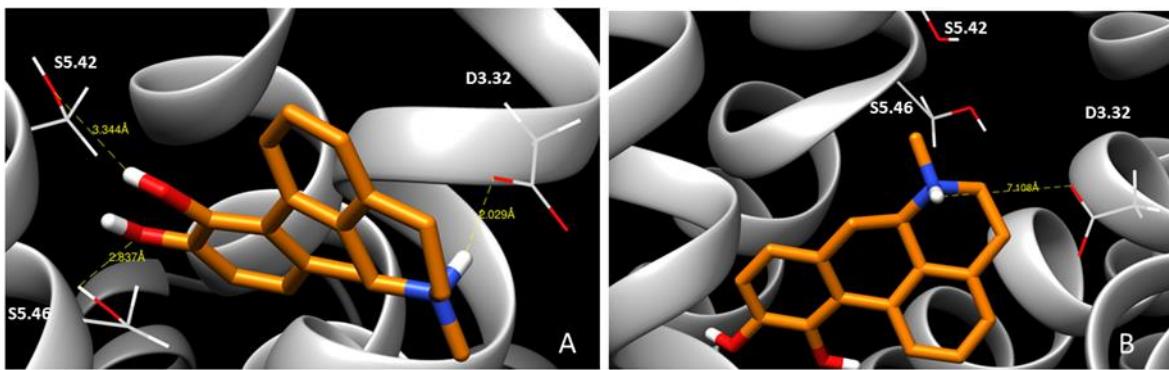


Figure B7: poses de la apomorfina unida al receptor en la forma activada del receptor (A) y a la forma inactiva (B). Se detallan los residuos que forman interacciones más importantes

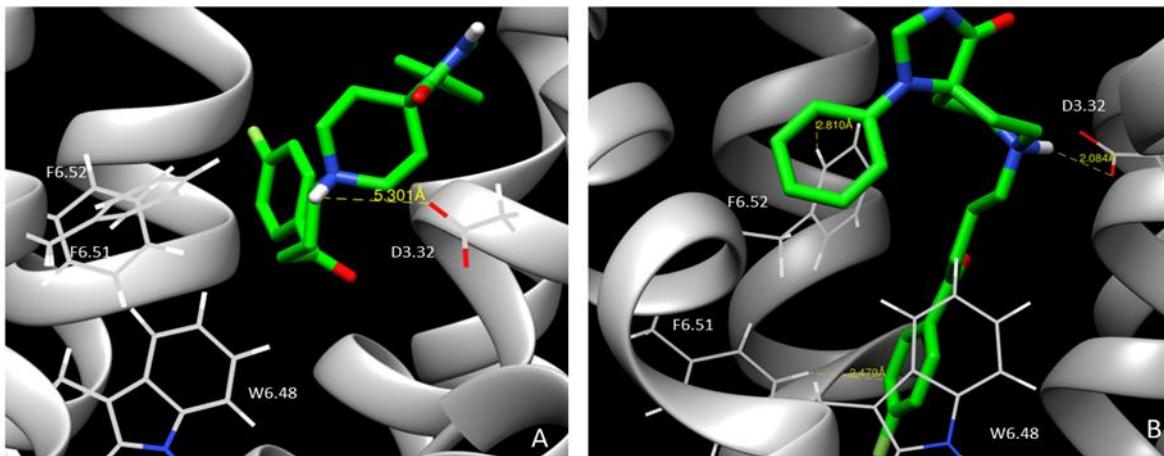


Figure B8: poses de la spiperona unida al receptor en la forma activa del receptor (A) y a la forma inactiva (B). Se detallan los residuos que forman interacciones más importantes.

En la tabla B3 se muestran los valores de energía de unión resultado del Docking para cada receptor, se puede observar que todos los agonistas inversos presentan valores más bajos de energía con el D2R en su forma inactiva, siendo este resultado favorable, ya que indicaría que estos ligandos se unen de una manera más estable al D2Ri que al D2R a. En el caso de los agonistas el 80% presenta valores de energías inferiores con el D2R activado, indicando una tendencia a que este tipo de ligandos se une preferentemente al receptor en su forma activa.

Tabla B3: Resultados obtenidos del Docking con los valores de Energía de unión para cada receptor

	LIGANDOS	Energía de unión a D2Ract (kcal/mol)	Energía de unión a D2Rinact (kcal/mol)	
AGONISTAS	Bromocriptine	-9,1	-9,7	✗
	Apomorphine	-9,2	-8,8	✓
	Cabergoline	-8,9	-8,4	✓
	Dopamina	-5,6	-5,5	✓
	Pergolide	-6,8	-7,8	✗
	Pramipexole	-6,3	-6,2	✓
	Ropinirole	-7,2	-7,1	✓
	Rotigotine	-7,9	-7,6	✓
	Quinpirole	-7,5	-7,3	✓
	PPHT	-8,7	-7,6	✓
AGONISTAS INVERSOS	Thiothixene	-7,7	-8,9	✓
	Pimozide	-8,7	-9,6	✓
	Tefludazine	-8	-8,2	✓
	Risperidone	-8,9	-9,1	✓
	Spiperone	-8,2	-10,1	✓
	Butaclamol	-8,4	-9	✓
	Clozapine	-7,6	-8,1	✓
	Haloperidol	-7,6	-8,7	✓
	Clorpromazine	-7,3	-7,7	✓
	Raclopride	-6,9	-7,1	✓

Estructuras cristalinas y comparación con los modelos de homología

Como se dijo previamente durante el desarrollo de esta tesis se cristalizaron las estructuras del receptor D2. Por lo que pudimos comparar el modelo D2Ri unido al agonista inverso con la estructura cristalina de D2R en su forma inactiva (PDB: 6CM4) y el modelo D2Ra-Gi unido a la apomorfina con la estructura resuelta de D2R (PDB: 6VMS).

Después de la alineación de las estructuras en la forma inactiva, la desviación cuadrática media general (RMSD) entre los átomos del backbone resultó ser de 0,946 Å, lo que indica una gran similitud estructural de nuestro modelo D2R en estado inactivo con respecto a la estructura experimental. La semejanza entre ambas estructuras se puede inspeccionar visualmente en la superposición de la Figura B9. La diferencia más destacada entre ambas estructuras está en la conformación del bucle extracelular 2. Sin embargo, nuestro modelo es capaz de acomodar el agonista inverso en un modo de unión bastante similar al de la estructura resuelta de D2R, como se evidencia en la superposición de la Figura B9. En la estructura cristalina de D2R, la fracción de benzisoxazol de la risperidona se extiende hacia un bolsillo de unión profundo definido por las cadenas laterales de TM3, TM5 y TM6, interactuando con C3.36, T3.37, S5.46, F5.47, F6.44, F6.52 y W6.48 que forman un subbolsillo debajo del sitio ortostérico (Wang et al. al., 2018). De manera similar, el resto fluorobenceno de la spiperona en el modelo D2Ri también penetra en el bolsillo hidrofóbico donde forma enlaces C-H…F H con los residuos I3.40, F5.47 y F6.44 y también interactúa con W6.48 a través de interacciones de stacking.

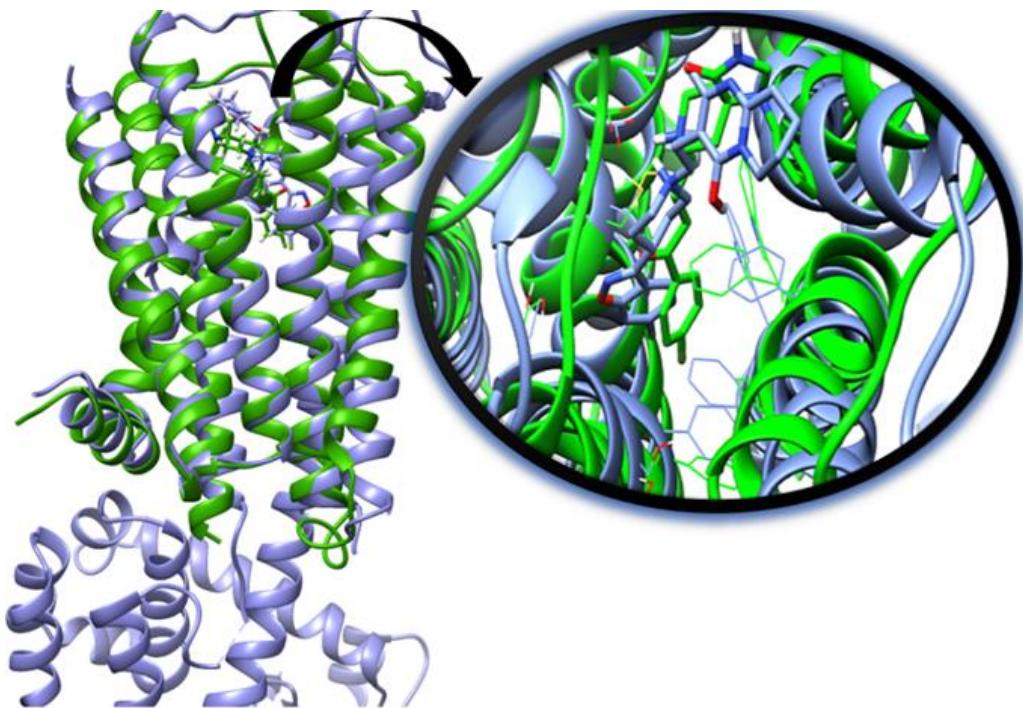


Figure B9: Superposición de la estructura cristalina (celeste) y el modelo de homología (verde) de la forma inactiva del D2R.

Vale la pena señalar que uno de los principales residuos involucrados en el switch de transmisión, I3.40, fue mutado a alanina en la estructura cristalina de D2R (Wang et al., 2018), lo que genera cierta preocupación sobre el estado que realmente representa esta estructura. Mientras que otros interruptores moleculares como el interruptor de bloqueo iónico y el interruptor de bloqueo 3-7 indican que es más probable que el receptor resuelto esté en estado inactivo, hay algunos cambios sutiles en el lado extracelular que sugieren que podría estar en un estado intermedio de activación. Más concretamente, el extremo extracelular de TM5 ya ha experimentado un movimiento hacia adentro en comparación con las estructuras D3R y D4R que se resolvieron en estado inactivo. El movimiento hacia adentro de TM5 podría ser una consecuencia directa de la mutación I3.40A que podría convertir a D2R en un estado activado (constitutivamente) parcial.

Para el modelo D2Ra-Gi unido a agonista, realizamos la comparación con la estructura resuelta de D2R (PDB: 6VMS), cristalizada en su estado activo acoplado a proteína G. Después de la alineación de ambas estructuras, la desviación cuadrática media general (RMSD) entre los átomos del backbones correspondiente resultó ser de 1,16 Å, lo que indica una similitud estructural de nuestro modelo D2R con respecto a la estructura experimental (Figura B10).

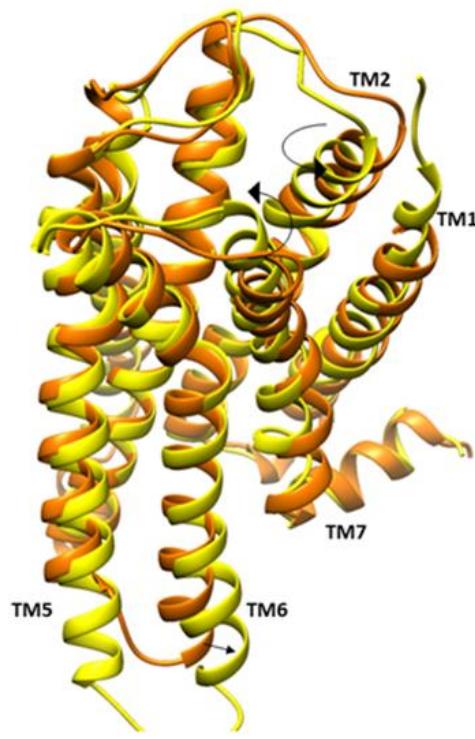


Figura B10: Superposición estructural de estructura cristalina (amarillo) y modelo de homología (naranja) de D2R-Gi en la forma activa.

Luego, analizando los cambios en la activación del receptor, pudimos observar la similitud en los desplazamientos de los residuos iniciados en el sitio de unión del ligando, que se propagan hacia la región intracelular a través de interruptores moleculares, como se puede ver en la Figura B11. Aquí también es claro que la asociación entre los segmentos TM5 y TM6 se fortalece con la activación del receptor, siendo los cambios más notables en los residuos Y5.58 que, al acercarse a TM5, se apila en la parte superior con la L6.37, y el residuo Y5.62 que se aproxima a T6.34 en ambos casos. Los residuos T5.54, I6.36, V6.40, V6.48 y V7.78 se mutaron en la estructura resuelta experimentalmente de D2R-Gi; aunque estos residuos no participan directamente en los interruptores de activación, su mutación puede afectarlos debido a la proximidad a la que encuentran. Además, estos residuos se ubican en los segmentos TM6 y TM7, los cuales se mueven durante la activación como hemos mencionado anteriormente.

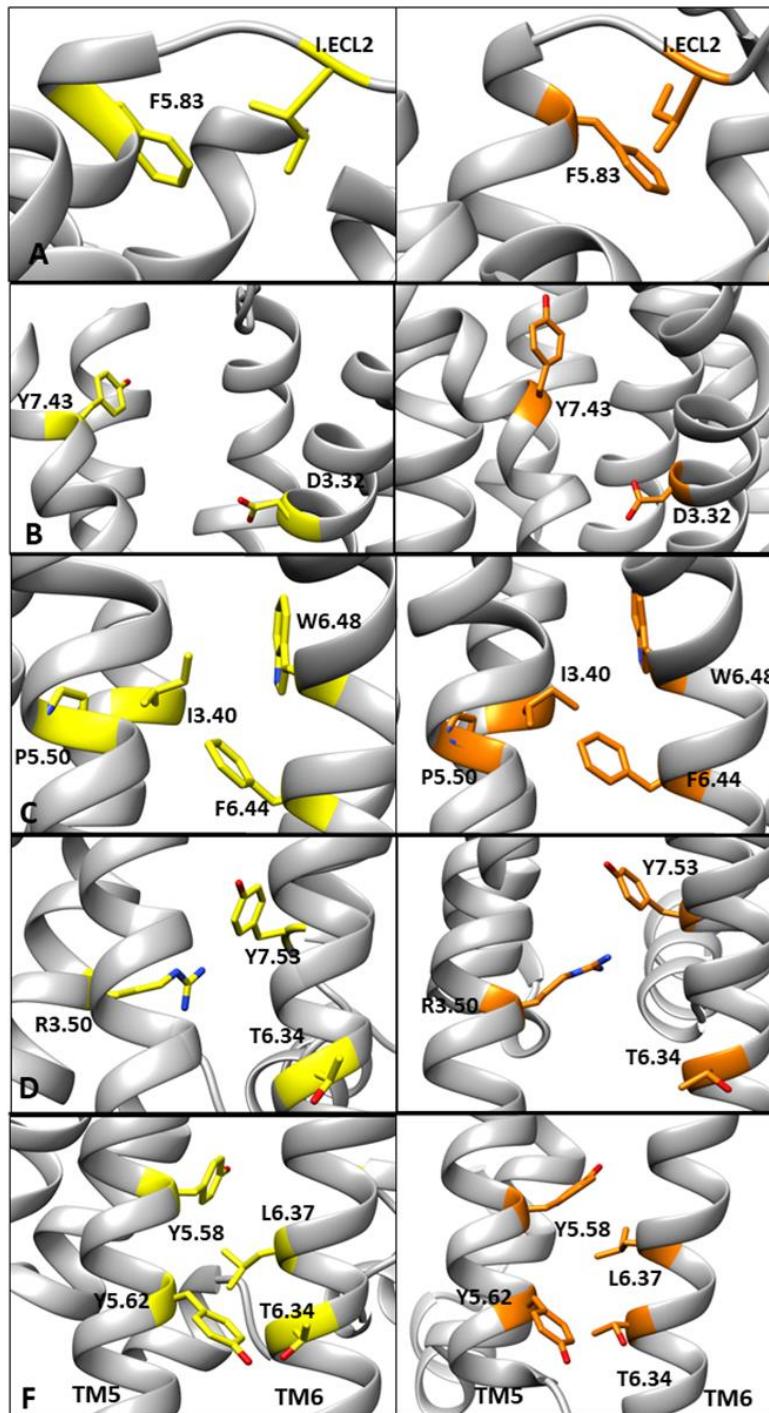


Figura B11: Comparación de los interruptores moleculares en la estructura cristalina de la forma activa (amarillo a la izquierda) y el modelo de homología D2R-Gi (naranja a la derecha). A) Rotación de la parte superior del segmento TM5 durante la activación, interruptor de transmisión, que involucra los residuos F5.83 e Ile de ECL2. B) Interruptor de bloqueo 3-7: Ruptura de la interacción entre Y7.43 y D3.32 que permite el desplazamiento de los segmentos. C) Interruptor de transmisión P5.50-I3.40-F6.44. D) Interruptor de Tiroxina, interacción entre R3.50 (de la secuencia conservada DRY) e Y.753, que no se encuentra en las formas inactivas. F) Desplazamiento de los segmentos TM5 y TM6, en los que se observa la proximidad de los residuos en activación, involucrando el interruptor de transmisión

ANEXO C

Estudio de inhibidores de la enzima Cruzipain y sus interacciones a través de métodos de modelado molecular

Análisis de Densidad electrónica

Para el análisis de densidad electrónica para los complejos con tiosemicarbazonas, se calculó los valores totales de ρ de las interacciones intermoleculares. Como se puede observar en la Figura C1 los valores de densidad electrónica total de los puntos críticos de enlace son similares en ambos pares de compuestos.

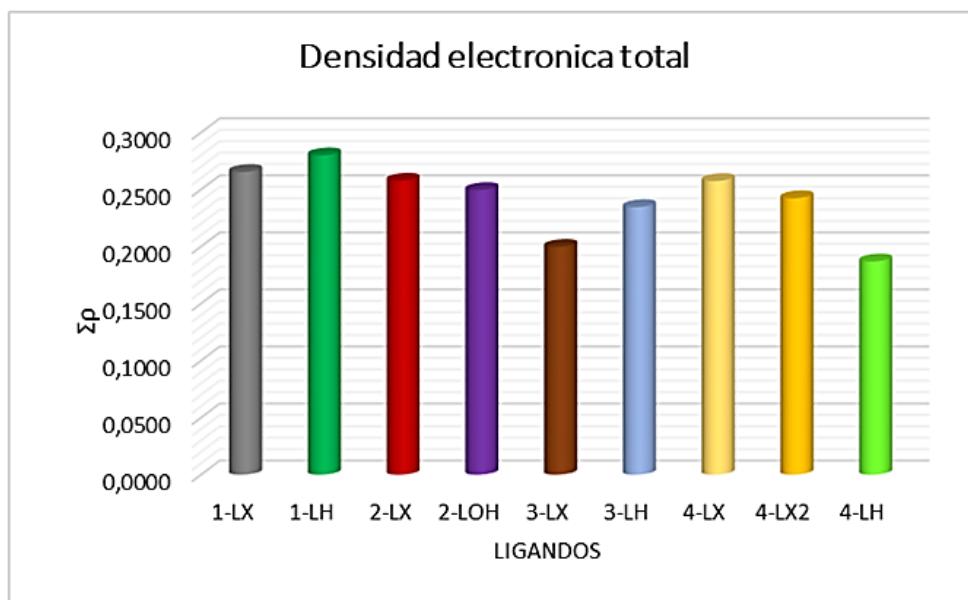


Figura C1: Valores de densidad electrónica total ($\Sigma \rho$) (u.a.) de los puntos críticos de enlace (BCP) de los ligandos con halógeno (X) y sus respectivos análogos (H/OH).

Posteriormente se realizó el análisis de densidad electrónica por residuo de la enzima (Figura C2), en estos resultados se puede observar que en los compuestos halogenados generalmente el residuo Met68 (M68) presenta un elevado valor de densidad electrónica, lo que demostraría la posible existencia de un EX con el halógeno del ligando como se planteó en este capítulo. Además, los datos demuestran en los compuestos halogenados los residuos que mayor aporte realizan al valor de densidad electrónica total en todos los casos son Trp26(W26) y Asp 161 (D161), mientras que en los compuestos no halogenados se destaca la His162 (H162).

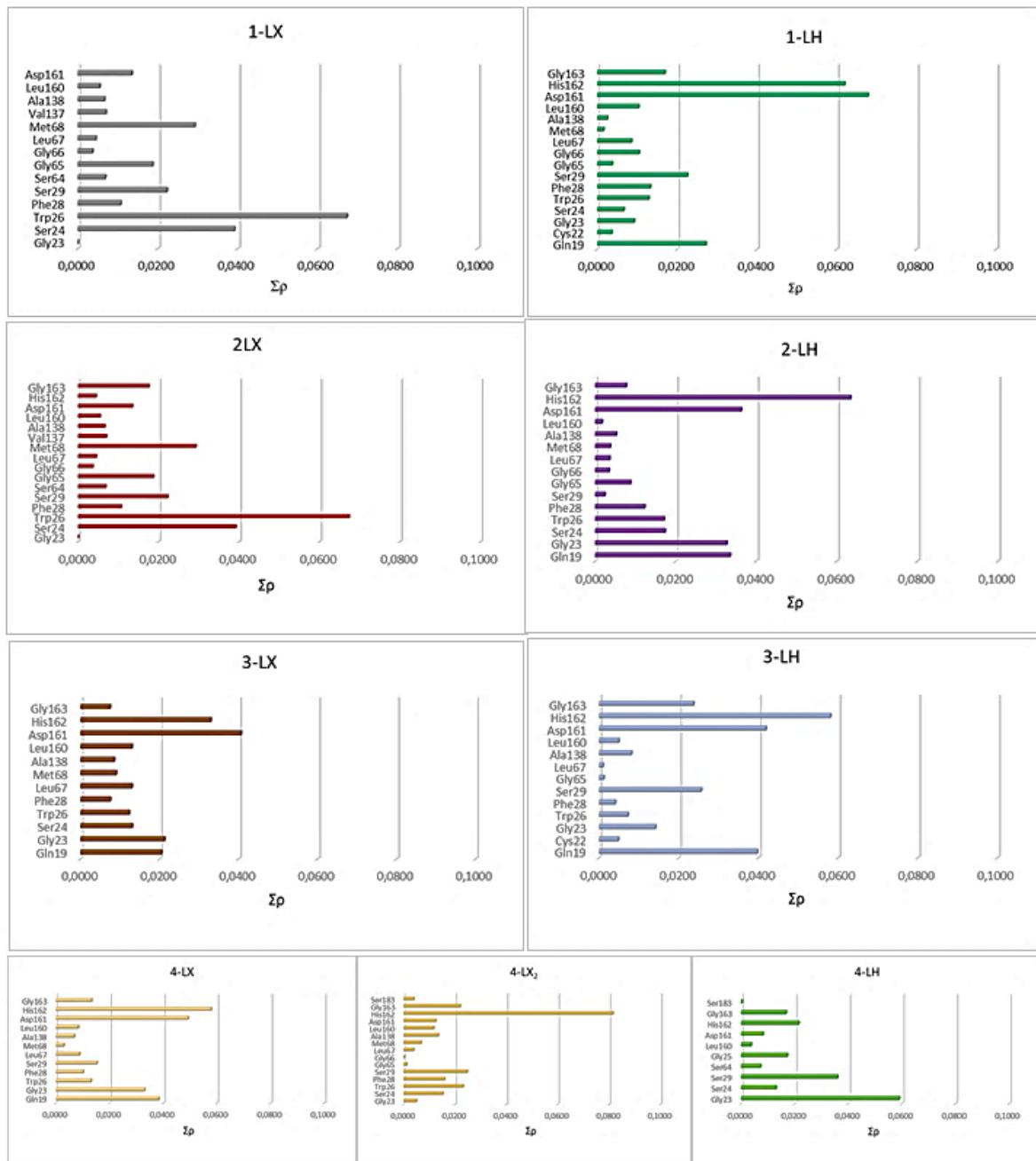


Figura C2: Cada grafico representa suma de los valores de densidad electrónica total (u.a.) de cada residuo de la enzima en los puntos críticos de enlace con cada ligando respectivamente.

Al superponer las estructuras de los mínimos de energía de la Dinámica Molecular de cada par se puede ver que las poses en la que se ubican los compuestos en el sitio de unión de la enzima son similares, pero al analizar los grafos de densidad electrónica se puede ver que forman caminos de enlace con diferentes residuos. En las figuras C3 y C4, se puede observar las poses superpuestas y los grafos de los pares 1-LX/LH y 2-LX/LOH, en los cuales se destaca que solo los halógenos forman un camino de enlace con el azufre de la metionina 68, mientras que en los análogos, tanto el hidrógeno en 1-LH como el grupo hidroxilo en 2-LOH no interaccionan con este residuo.

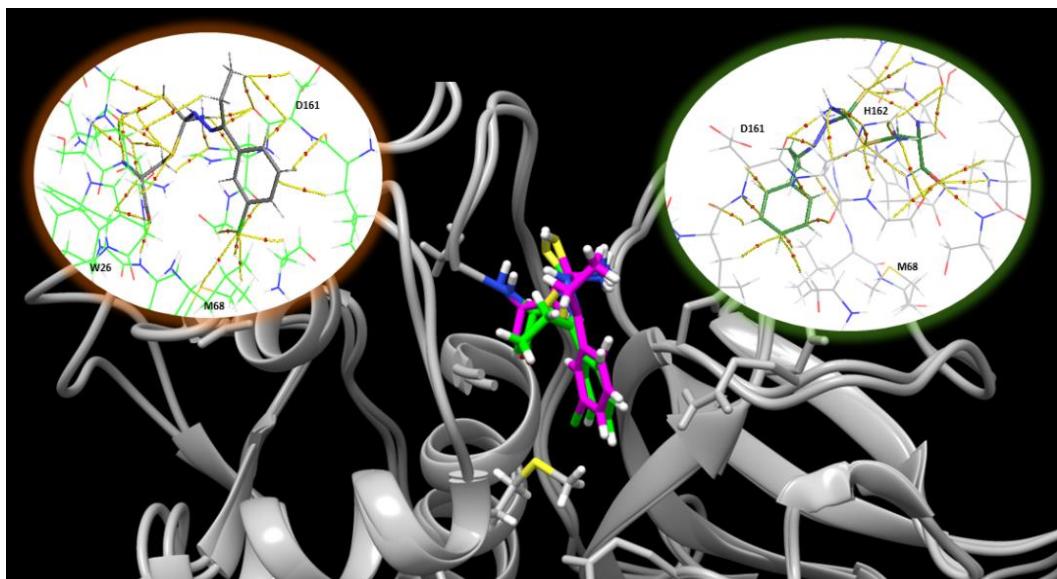


Figura C3: Estructura superpuesta del mínimo de energía del par 1-LX/LH, el compuesto halogenado (magenta) con su análogo no halogenado (verde). Resaltados con un círculo sombreado los grafos moleculares de cada ligando con los residuos seleccionados para el cálculo QTAIM, a la izquierda 1-LX y a la derecha el grafo molecular de 1-LH.

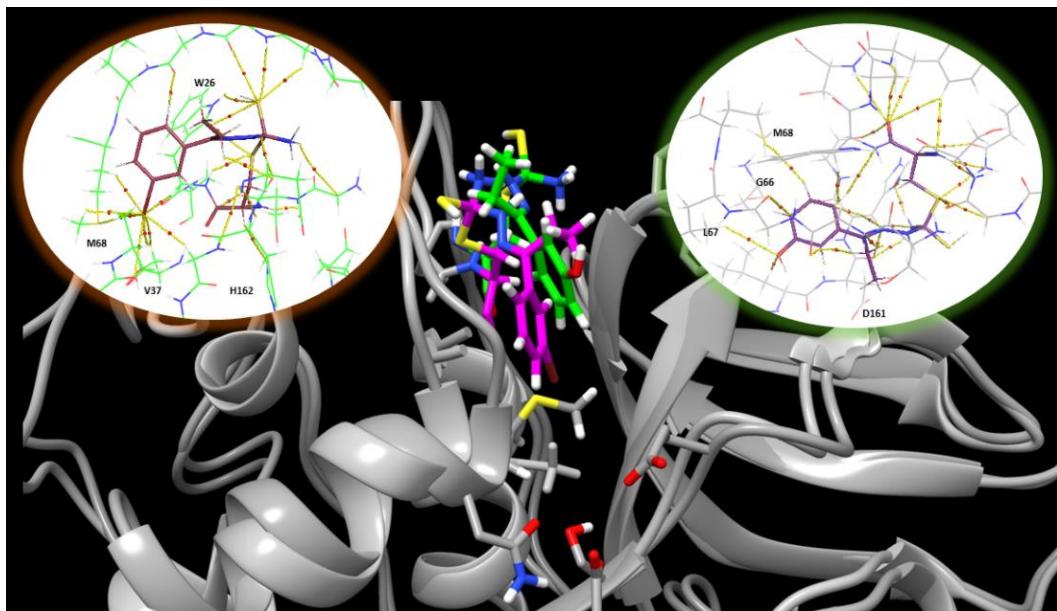


Figura C4: Estructura superpuesta del mínimo de energía del par 2-LX/LOH, el compuesto halogenado (magenta) con su análogo no halogenado (verde). Resaltados con un círculo sombreado los grafos moleculares de cada ligando con los residuos seleccionados para el cálculo QTAIM, a la izquierda 2-LX y a la derecha el grafo molecular de 2-LOH.

Tiempo de vida media y porcentaje de ocupación de EHs

Se calculó el *porcentaje de ocupación o de ocupación* de los EH que forma el ligando en el sitio activo de la enzima a lo largo de las simulaciones de dinámica molecular para los pares de complejos 1-LX/LH y 2-LX/LOH; lo que determina cuánto tiempo perdura el enlace de hidrógeno formado sobre el número total de “frames” multiplicado por 100. Luego, para un análisis más profundo se calculó también el tiempo de vida media de los Enlace de hidrogeno, “lifetime” (LT); LT es la longitud de tiempo que una interacción está presente, lo que nos permite comparar aquellas que son más relevantes en cada caso.

En el caso del complejo 1-LX el resultado de LT demuestra que el puente de hidrogeno Lig25@O-Ser29@NH posee el mayor tiempo de vida media, 10 frames; mientras que para 1-LH el puente de hidrogeno Asp161@O-Lig25@N3H5 posee el mayor tiempo de vida media de 33 frames, seguido por la interacción entre Asp161@O-Lig25@N3H6 que dura 30 frames. Los datos obtenidos de los cálculos de porcentaje de ocupación se correlacionan con estos resultados, en 1-LX el 34% de la simulación hay un puente de hidrogeno entre el O del Ligando 25 y el Nitrógeno del residuo Ser 29; un 28% de la simulación se mantiene un puente de hidrogeno entre el oxígeno del Asp161 y el N3 del Ligando. Mientras que para 1-LH hay un puente de hidrogeno durante un 36% de la simulación entre el O del ligando y el N del residuo

Ser29; un 28 % de la simulación se mantiene además el puente entre el O del Asp 161 y el N3 del Ligando25.

Tabla C1: Resultados de LT para el complejo 1-LX

	Máximo LT	Promedio LT	Total (frames)
GLN_19@NE2-LIG_25@N-H	1	1	5
GLY_23@O-LIG_25@N-H	3	1,0962	114
GLY_23@O-LIG_25@N2-H4	5	1,2752	139
SER_64@O-LIG_25@N2-H4	5	1,584	198
ASP_161@O-LIG_25@N3-H6	34	3,6055	4478
HIE_162@ND1-LIG_25@N3-H6	14	1,972	3871
ASP_161@O-LIG_25@N3-H7	26	3,6078	4701
HIE_162@ND1-LIG_25@N3-H7	11	1,9507	3084
ASP_161@O-LIG_25@N3-H8	26	3,9428	5717
HIE_162@ND1-LIG_25@N3-H8	12	1,9508	3211
LIG_25@O-PHE_28@N-H	2	1,0239	343
LIG_25@O-SER_29@N-H	10	1,5889	6813

Tabla C2: Resultados de LT para el complejo 1-LH

	Máximo LT	Promedio LT	Total (frames)
LIG_25@N-GLN_19@NE2-HE21	1	1	1
GLN_19@NE2-LIG_25@N-H	1	1	8
GLY_23@O-LIG_25@N-H	5	1,2481	322
GLY_23@O-LIG_25@N2-H7	1	1	1
ASP_161@O-LIG_25@N3-H5	33	3,8531	4932
HIE_162@ND1-LIG_25@N3-H5	10	1,9127	2957
ASP_161@O-LIG_25@N3-H6	30	3,9857	5576
HIE_162@ND1-LIG_25@N3-H6	9	1,8966	3027
ASP_161@O-LIG_25@N3-H8	27	3,8469	4647
HIE_162@ND1-LIG_25@N3-H8	13	1,9051	3452
LIG_25@O-PHE_28@N-H	3	1,0419	348
LIG_25@O-SER_29@N-H	13	1,6787	7168
LIG_25@O-SER_29@OG-HG	8	2,2222	60

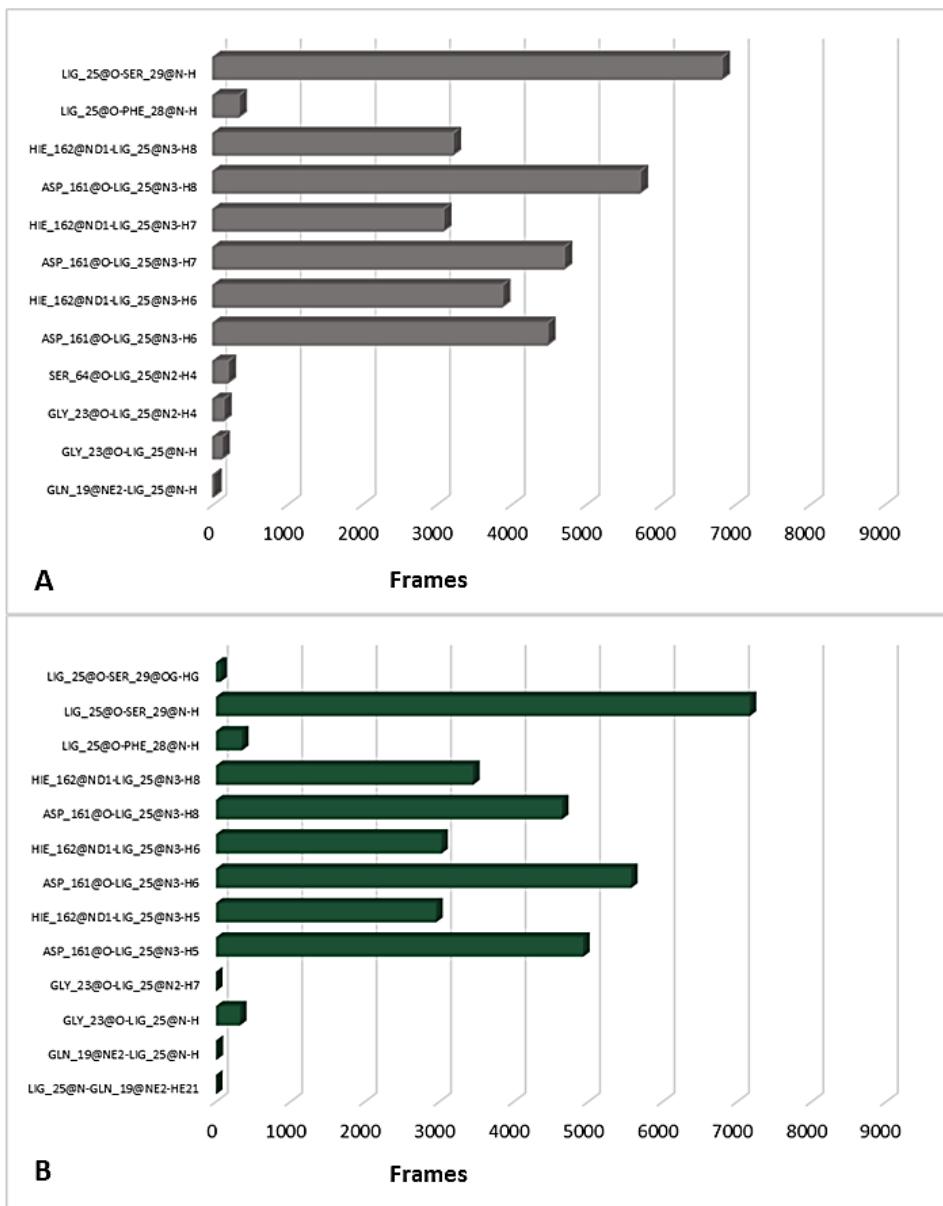


Figura C5: Grafico del tiempo de vida media durante la trayectoria de Dinámica Molecular de las interacciones de 1-LX (A) y 1-LH (B).

Tabla C3: Resultados de Porcentaje de Ocupación de EH para el complejo 1-LX

	Frames	Porcentaje de ocupancia (%)	Distancia promedio (Å)	Angulo promedio (°)
LIG_25@O-SER_29@H	6813	34,07	2,9019	156,1936
ASP_161@O- LIG_25@H8	5717	28,58	2,8374	155,3332
ASP_161@O- LIG_25@H7	4701	23,51	2,8387	155,3878
ASP_161@O-LIG_25@H6	4478	22,39	2,8386	155,807
HIE_162@ND1-LIG_25@H6	3871	19,35	2,9034	158,926
HIE_162@ND1- LIG_25@H8	3211	16,05	2,9024	158,7168
HIE_162@ND1- LIG_25@H7	3084	15,42	2,9042	158,889
LIG_25@O- PHE_28@H	343	1,71	2,904	141,7716
SER_64@O- LIG_25@H4	198	0,99	2,8953	155,2765
GLY_23@O-LIG_25@H4	139	0,69	2,8953	158,2119
GLY_23@O - LIG_25@H	114	0,57	2,9025	141,8623
GLN_19@NE2 -LIG_25@H	5	0,03	2,9367	144,7084

Tabla C4: Resultados de Porcentaje de Ocupación de EH para el complejo 1-LH

	Frames	Porcentaje de ocupacion (%)	Distancia promedio (Å)	Angulo promedio (°)
LIG_25@O-SER_29@H	7168	35,84	2,9004	158,3377
ASP_161@O- LIG_25@H6	5576	27,88	2,8409	156,3398
ASP_161@O-LIG_25@H5	4932	24,66	2,8411	156,0052
ASP_161@O- LIG_25@H8	4647	23,24	2,8433	156,433
HIE_162@ND1 - LIG_25@H8	3452	17,26	2,9072	160,0836
HIE_162@ND1- LIG_25@H6	3027	15,14	2,9073	159,8631
HIE_162@ND1- LIG_25@H5	2957	14,79	2,9062	160,1124
LIG_25@O- PHE_28@H	348	1,74	2,9032	142,2078
GLY_23@O- LIG_25@H	322	1,61	2,8912	143,9794
LIG_25@O-SER_29@HG	60	0,3	2,8142	162,0006
GLN_19@NE2-LIG_25@H	8	0,04	2,9689	143,6915
LIG_25@N-GLN_19@HE21	1	0,01	2,9756	142,0417
GLY_23@O-LIG_25@H7	1	0,01	2,9805	158,1636

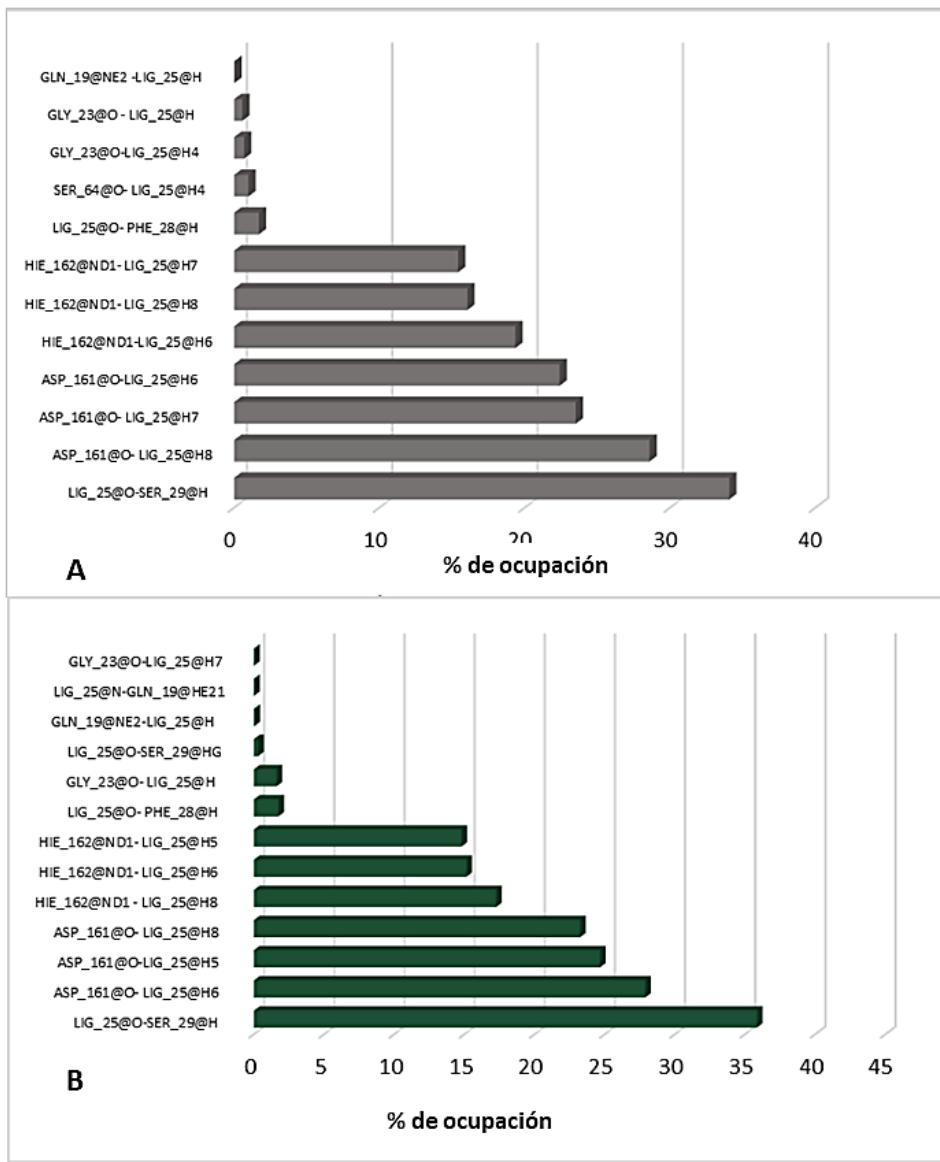


Figura C6: Grafico de porcentaje de ocupación durante la trayectoria de Dinámica Molecular de los complejos 1-LX (A) y 2-LH (B)

Analizando el par 2-LX/LOH, se puede observar que para el compuesto halogenado la interacción con mayor vida media es Lig25@O-Ser29@NH que se mantiene 14 frames consecutivos, mientras que para 2-LOH el puente de hidrogeno Asp161@O-Lig25@N3H4 es el que posee mayor tiempo de vida media durante la simulación, un total de 14 frames. Observando los porcentajes de ocupación de los EHs en 2-LX el 42 % de la simulación se mantiene el puente de hidrogeno entre la O del Ligando25 y el N del residuo Serina 29 , mientras

que para 2-LOH durante el 20% de la simulación se mantienen los puentes de hidrogeno entre el O del residuo Asp161 y el nitrógeno del Ligando.

Tabla C5: Resultados de LT de los Enlaces de Hidrogeno para el complejo 2-LX

	Maximo LT	Promedio LT	Total (frames)
GLN_19@OE1-LIG_25@N-H	3	1,1226	119
GLN_19@NE2-LIG_25@N-H	1	1	1
GLY_23@O-LIG_25@N-H	1	1	2
ASP_161@O-LIG_25@N2-H6	4	1,1293	1293
CYX_63@O-LIG_25@N3-H4	2	1,5	3
ASP_161@O-LIG_25@N3-H4	8	1,3382	641
HIE_162@ND1-LIG_25@N3-H4	5	1,6429	46
CYX_63@O-LIG_25@N3-H5	1	1	1
ASP_161@O-LIG_25@N3-H5	6	1,286	616
HIE_162@ND1-LIG_25@N3-H5	7	2,1724	63
GLY_23@O-LIG_25@N3-H8	1	1	1
CYX_63@O-LIG_25@N3-H8	2	1,6667	5
ASP_161@O-LIG_25@N3-H8	6	1,3503	636
HIE_162@ND1-LIG_25@N3-H8	6	1,7714	62
LIG_25@O-PHE_28@N-H	1	1	11
LIG_25@O-SER_29@N-H	14	1,8409	8400

Tabla C6: Resultados de LT de los Enlaces de Hidrogeno para el complejo 2-LOH

	Máximo LT	Promedio LT	Total (frames)
GLN_19@OE1-LIG_25@N-H	1	1	1
GLN_19@NE2-LIG_25@N-H	2	1,025	41
GLY_23@O-LIG_25@N-H	3	1,063	422
GLY_23@O-LIG_25@N2-H6	4	1,2109	178
SER_64@O-LIG_25@N2-H6	2	2	2
LEU_160@O-LIG_25@O1-H8	7	1,3447	394
ASP_161@O-LIG_25@N3-H4	14	1,6458	3685
HIE_162@ND1-LIG_25@N3-H4	9	1,5311	3422
ASP_161@O-LIG_25@N3-H5	9	1,7152	3818
HIE_162@ND1-LIG_25@N3-H5	8	1,566	3392
LIG_25@N1-LIG_25@N3-H9	1	1	1
ASP_161@O-LIG_25@N3-H9	10	1,6817	3836
HIE_162@ND1-LIG_25@N3-H9	9	1,5182	3384
LIG_25@O-PHE_28@N-H	2	1,0075	135
LIG_25@O-SER_29@N-H	5	1,1592	1711
LIG_25@O-SER_29@OG-HG	10	2,1145	960
LIG_25@N-GLN_19@NE2-HE21	1	1	6

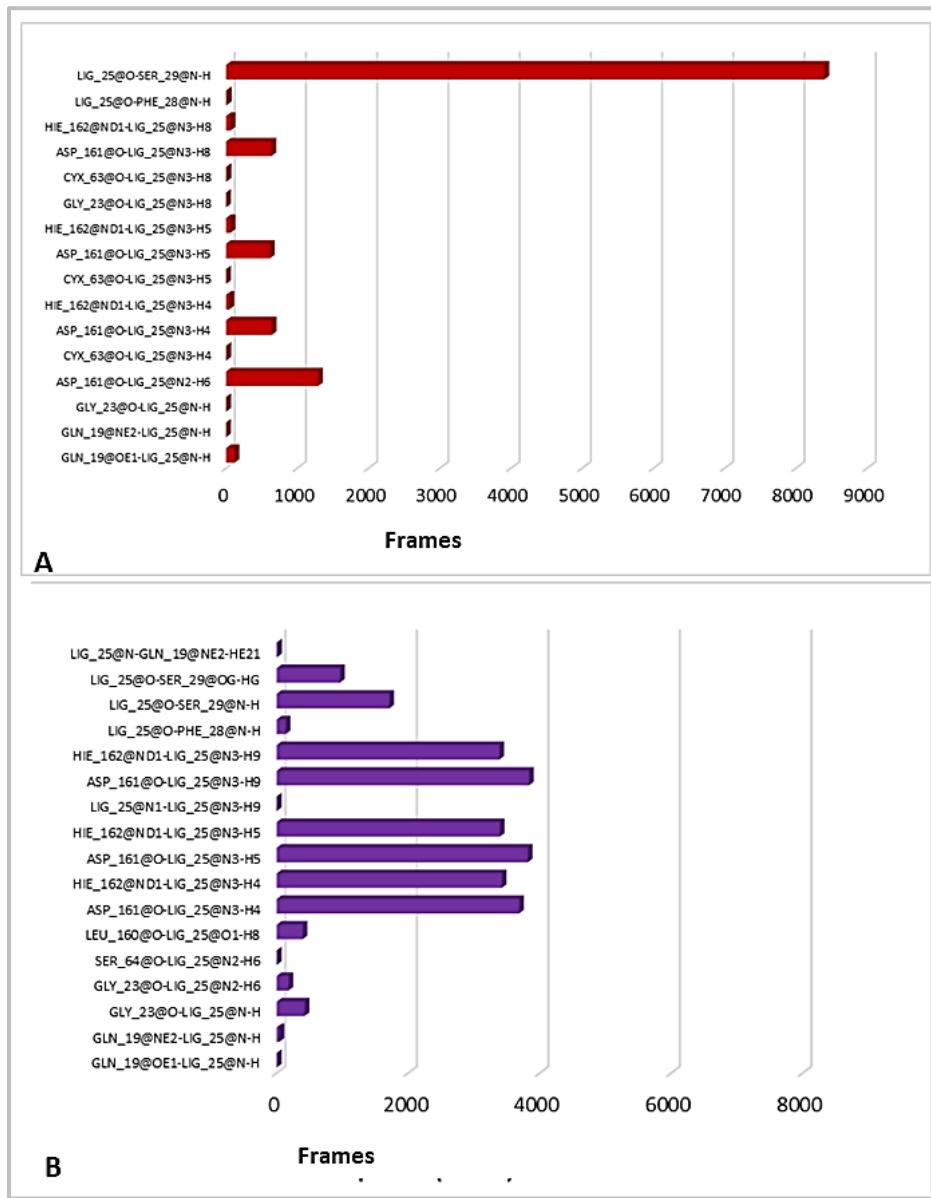


Figura C7: Grafico del tiempo de vida media durante la trayectoria de Dinámica Molecular de las interacciones de 2-LX (A) y 2-LOH (B)

Tabla C8: Resultados de Porcentaje de Ocupación de los Enlaces de Hidrogeno para el complejo 2-LX

	Frames	Porcentaje (%)	Distancia promedio (Å)	Angulo promedio (°)
LIG_25@O-SER_29@H	8400	42	2,9028	161,147
ASP_161@O-LIG_25@H6	1293	6,46	2,9176	147,2766
ASP_161@O-LIG_25@H4	641	3,21	2,8345	151,1325
ASP_161@O-LIG_25@H8	636	3,18	2,8293	152,1091
ASP_161@O-LIG_25@H5	616	3,08	2,8273	152,1576
GLN_19@OE1-LIG_25@H	119	0,6	2,9173	157,1407
HIE_162@ND1- LIG_25@H5	63	0,32	2,8919	158,2313
HIE_162@ND1- LIG_25@H8	62	0,31	2,8701	156,8168
HIE_162@ND1-LIG_25@H4	46	0,23	2,8777	155,4109
LIG_25@O-PHE_28@H	11	0,06	2,8989	142,5962
CYX_63@O-LIG_25@H8	5	0,03	2,7889	154,7441
CYX_63@O- LIG_25@H4	3	0,01	2,8411	153,1142
GLY_23@O-LIG_25@H	2	0,01	2,8077	157,7012
GLY_23@O-LIG_25@H8	1	0,01	2,749	161,2275
CYX_63@O- LIG_25@H5	1	0,01	2,9343	156,6101
GLN_19@NE2-LIG_25@H	1	0,01	2,9483	154,0306

Tabla C9: Resultados de Porcentaje de Ocupación de los Enlaces de Hidrogeno para el complejo 2-LOH

	Frames	Porcentaje (%)	Distancia promedio (Å)	Angulo promedio (°)
ASP_161@O- LIG_25@H9	3836	20,19	2,8468	154,532
ASP_161@O-LIG_25@H5	3818	20,09	2,8426	154,5702
ASP_161@O-LIG_25@H4	3685	19,39	2,8447	154,5591
HIE_162@ND1- LIG_25@H4	3422	18,01	2,8975	157,4102
HIE_162@ND1-LIG_25@H5	3392	17,85	2,9	157,6362
HIE_162@ND1- LIG_25@H9	3384	17,81	2,8996	157,7422
LIG_25@O- SER_29@H	1711	9,01	2,9259	158,8215
LIG_25@O -SER_29@HG	960	5,05	2,8152	159,6158
GLY_23@O-LIG_25@H	422	2,22	2,8847	141,6686
LEU_160@O- LIG_25@H8	394	2,07	2,7589	158,5232
GLY_23@O- LIG_25@H6	178	0,94	2,9001	151,1512
LIG_25@O-PHE_28@H	135	0,71	2,9184	145,1727
GLN_19@NE2-LIG_25@H	41	0,22	2,9417	144,3396
LIG_25@N-GLN_19@HE21	6	0,03	2,9047	142,381
SER_64@O- LIG_25@H6	2	0,01	2,9258	162,7805
LIG_25@N1- LIG_25@H9	1	0,01	2,6085	137,4942
GLN_19@OE1-LIG_25@H	1	0,01	2,9748	136,3051

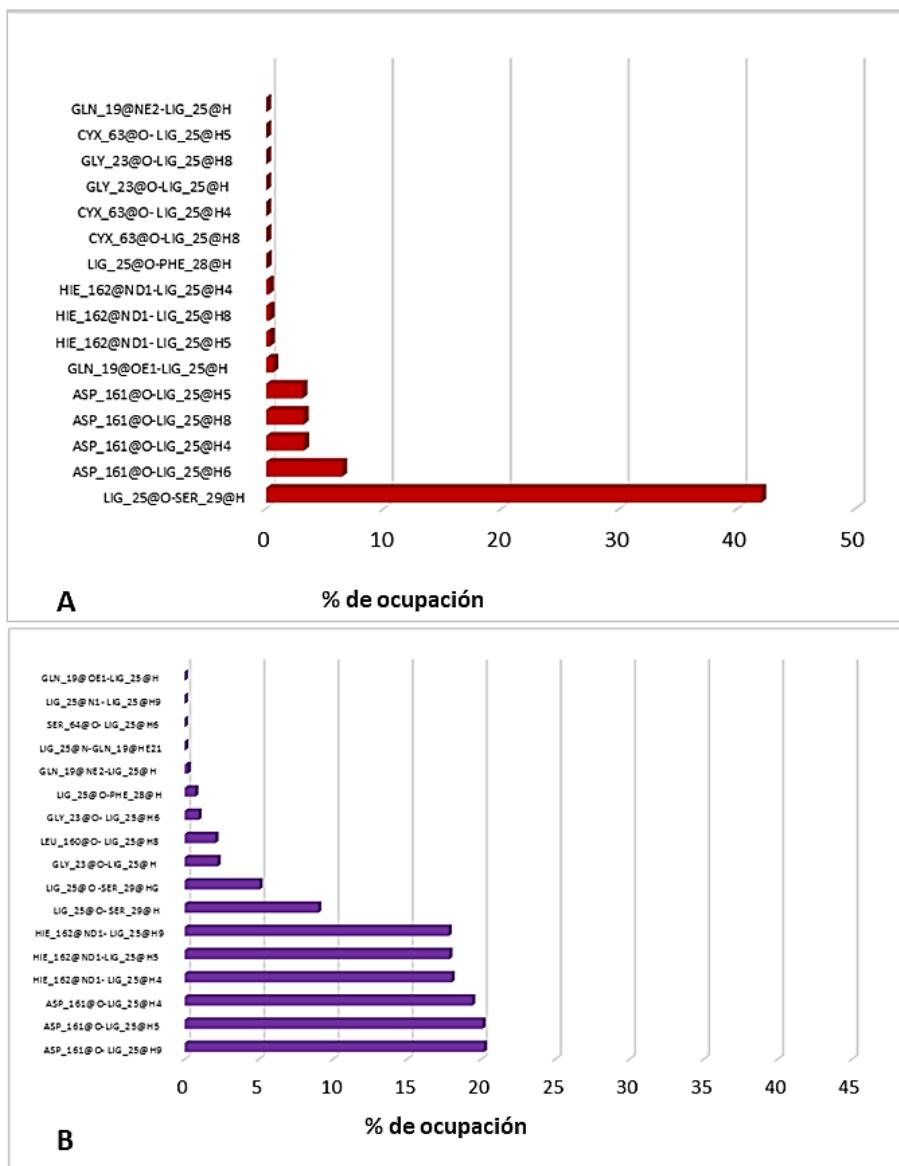


Figura C8: Grafico de porcentaje de ocupación durante la trayectoria de Dinamica Molecular de los complejos 2-LX (A) y 2-LOH (B)

Finalmente, del análisis del tiempo de vida y el porcentaje de ocupación de los EHs se pudo comprobar que para el caso de los compuestos halogenados la interacción más estable durante la dinámica es con la Ser 29, mientras que para el caso de los análogos no halogenados presenta mayor relevancia la interacción con el Asp 161.

