

UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA

PROGRAMA DE MAGISTER EN CIENCIAS, MENCIÓN QUÍMICA

“Estudio estructural y energético de la afinidad de unión de inhibidores tipo *orto*-Aminoanilida en complejo a la Metaloenzima HDAC1 mediante 3D-QSAR, *Docking*, Dinámica Molecular y cálculos MMPBSA”

Por

Jorge Francisco González Higuera

Magíster en Ciencias mención Química

Julio 2021



UNIVERSIDAD TECNICA
FEDERICO SANTA MARIA



Departamento
de **Química**



Magíster en Ciencias
Mención **Química**

UNIVERSIDAD TECNICA FEDERICO SANTA MARIA

PROGRAMA DE MAGÍSTER EN CIENCIAS, MENCIÓN QUÍMICA

Estudio estructural y energético de la afinidad de unión de inhibidores tipo *orto*-Aminoanilida en complejo a la Metaloenzima HDAC1 mediante 3D-QSAR, *Docking*, Dinámica Molecular y cálculos MMPBSA.

Tesis de Grado presentada por

Jorge Francisco González Higuera

Como requisito parcial para optar al grado de

Magíster en Ciencias, mención Química.

Director de Tesis: Dr. Lautaro Taborga Morales

Co-Dirección de Tesis: Dr. Jorge Soto Delgado

Valparaíso, Chile

Noviembre 2020

PROGRAMA DE MAGISTER EN CIENCIAS, MENCIÓN QUÍMICA

TÍTULO DE LA TESIS:

Estudio estructural y energético de la afinidad de unión de inhibidores tipo *orto*-Aminoanilida en complejo a la Metaloenzima HDAC1 mediante 3D-QSAR, Docking, Dinámica Molecular y cálculos MMPBSA.

AUTOR:

Jorge Francisco González Higuera

TRABAJO DE TESIS, presentado en cumplimiento parcial de los requisitos para el Grado de Magíster en Ciencias, mención Química de la Universidad Técnica Federico Santa María.

Dra. Beatriz Camara.

Dr. Jaime Mella.

Dr. Claudio Morgado.

Dr. Lautaro Taborga.

Dr. Jorge Soto.

Valparaíso, Chile. Julio de 2021

*En memoria de un gran amigo
Viktor Slisarenko Stachniw*

A la familia, amigos y amigas. A Loreto y Amelia. En especial a todos los que me ayudaron a levantar mi casa, a las semillas, al bosque y su silencio.

AGRADECIMIENTOS

Primero agradecer a cada una de las personas pertenecientes al Departamento de Química UTFSM, que me han dado apoyo incondicional y oportunidades para desarrollarme.

Especialmente al Dr. Lautaro Taborga, por aceptarme como pupilo y ser mi tutor en el proceso. A la Dra. Beatriz Camara por sus importantes aportes en la revisión del proyecto en sus orígenes y que impactaron en los resultados finales. Al Dr. Claudio Morgado por sus contribuciones en la revisión exhaustiva de los manuscritos finales, las genuinas conversaciones y su noble rigurosidad.

Al programa de Magíster en Ciencias Mención Química y a su Director Dr. Luis Espinoza.

A la UTFSM y DGIIP, por el soporte académico y económico para el desarrollo de la Tesis.



Departamento
de **Química**



Magíster en Ciencias
Mención **Química**

Al Dr. Jorge Soto, por recibirme en el Laboratorio de Fármaco-química Molecular de UNAB Viña del mar, y a su proyecto Fondecyt N° 11150988. Recursos que permitieron contar con capacidad computacional necesaria para realizar los estudios de acoplamiento molecular, dinámica molecular y cálculos de energía de enlazamiento.



FONDECYT
Fondo Nacional de Desarrollo
Científico y Tecnológico



Al profesor Dr. Jaime Mella del Laboratorio de Química Medicinal, del instituto de química y bioquímica de la Universidad de Valparaíso, y a su proyecto Fondecyt N°1150121. Recursos que permitieron desarrollar los estudios de relaciones de estructura actividad cuantitativas mediante técnicas 3D-QSAR. Gracias por la confianza de dejarme en su laboratorio, el maní con merkén y el té manzanilla-jengibre-miel.

En General a todas las personas que fueron parte del camino.

LISTA DE ABREVIACIONES Y ACRONIMOS

HDACs	<i>Histonas Deacetilasas</i>
QM	<i>Quantum Mechanics</i>
HF	<i>Hartree-Fock</i>
SCF	<i>Self-consistent field</i>
RESP	<i>Restrained Electro-Static Potential</i>
MEP	<i>Molecular Electrostatic Potential</i>
AM1	<i>Austin Model 1</i>
MM	<i>Molecular Mechanics</i>
FF	<i>Force Fields</i>
AMBER	<i>Assisted Model Building with Energy Refinement</i>
MD	<i>Molecular Dynamics</i>
SASA	<i>Solvent Accessible Surface Area</i>
RMSD	<i>Root-Mean-Square Deviation</i>
MMPBSA	<i>Molecular Mechanics Poisson-Boltzmann Surface Area</i>
CG	<i>Coarse grained</i>
3D-QSAR	<i>3D- QUANTITATIVE STRUCTURE ACTIVITY RELATIONSHIP</i>
CoMFA	<i>Comparative Molecular Field Analysis</i>
CoMSIA	<i>Comparative Molecular Similarity Index Analysis</i>
RG	<i>Receptor guided</i>
LB	<i>Ligand Based</i>
CADD	<i>Computer Aided Drug Design</i>
PLS	<i>Partial Least Squares</i>
LOO	<i>Leave-one-out</i>
TIP3P	<i>Transferable Intermolecular Potential with 3 Points</i>
PME	<i>Particle Mesh Ewald</i>
PDB	<i>Protein Data Bank</i>

CONTENIDO

Agradecimientos	iii
LISTA DE ABREVIACIONES y ACRONIMOS	iv
RESUMEN 7	
SUMMARY 8	
CAPÍTULO 1 - INTRODUCCIÓN.....	9
1.1. Estado del Arte: Inhibidores.....	10
CAPÍTULO 2 - OBJETIVO DE INVESTIGACIÓN	12
2.1. Hipótesis.....	12
2.2. Objetivos Generales.....	12
2.3. Objetivos Específicos	12
CAPÍTULO 3 - METODOLOGIAS.....	13
3.1. Selección de la estructura inicial	14
3.2. Colección del conjunto de datos.....	14
3.3. Acoplamiento molecular	16
3.4. Alineamientos y estudios 3D-QSAR.....	17
3.4.1. Alineamiento basado en ligandos (LB).....	17
3.4.2. Alineamiento Guiado por receptor (RG)	17
3.4.3. Validación interna y análisis PLS: Estudios CoMFA y CoMSIA	17
3.4.4. Validación externa de modelos 3D-QSAR.....	18
3.5. Simulaciones de Dinámica Molecular	19
3.5.1. Análisis de las trayectorias.....	20
3.5.2. Estimación de la afinidad de unión (MMPBSA)	20
CAPÍTULO 4 - RESULTADOS Y DISCUSIONES	22
4.1. Acoplamiento molecular (Docking).....	22
4.2. Resultados de los estudios CoMFA y CoMSIA.....	23
4.3. Alineamiento asistido y mejorado por cristalografía de rayos-x.....	25
4.4. Resultados CoMFA y CoMSIA asistido por difracción de rayos -X.....	26
4.5. Mapas de Contorno CoMFA y CoMSIA.....	29
4.5.1. CoMFA -SE	29
4.5.2. CoMSIA-SEHA.....	30
4.6. Estudios de Dinámica Molecular	33
4.6.1. Estabilidad de la simulación del Complejo HDAC1-inhibidor.....	33

4.6.2.	Cálculos MMPBSA y evaluación de energía de interacción proteína-ligando (aproximada).....	34
4.6.3.	Contribuciones Por-Residuo al ΔG de interacción proteína-ligando (aproximada).....	36
4.6.4.	Análisis sistemático de interacciones no-covalentes específicas.....	40
4.6.5.	Puente de Hidrogeno	41
4.6.6.	Interacción de enlace de hidrógeno entre los residuos del sitio activo de HDAC1 y el compuesto 38.....	41
4.6.7.	Análisis de solvatación del sitio activo del complejo 38	42
CAPÍTULO 5 - CONCLUSIONES.....		45
APENDICES. 47		
Apéndice A1: Estado del arte.		47
A.1.	MODELADO MOLECULAR.....	48
A.1.1.	Estructura electrónica.....	49
A.1.1.1.	Métodos semiempíricos: AM1	49
A.1.1.2.	Estructura electrónica: cargas.....	49
A.1.2.	Métodos clásicos basados en campos de fuerza.....	50
A.1.2.1.	Campos de fuerza.	50
A.1.3.	Mecánica Molecular.....	51
A.1.3.1.	Minimización de la energía.....	51
A.1.4.	Dinámica molecular.....	52
A.1.4.1.	MMPBSA: Método de punto final para cálculo aproximado de $\Delta G_{\text{enlazamiento}}$ 54	
A.1.5.	Diseño de drogas asistido por computadoras	54
A.1.5.1.	Diseño de drogas basado en el receptor.	54
A.1.5.2.	<i>Docking</i> o Acoplamiento molecular.	55
A.1.5.3.	Diseño de fármacos basado en ligando.	55
A.1.5.4.	QSAR	56
A.1.6.	Modelado de farmacóforo.....	57
Apéndice 2: Interacciones no covalentes.		59
A.2.	INTERACCIONES MOLECULARES LIGANDO - RECEPTOR.....	60
A.2.1.	Consideraciones generales.....	60
A.2.2.	Interacciones importantes involucradas en complejos Ligando-Receptor.....	61
A.2.3.	Enlace covalente	62
A.2.4.	Interacciones no - covalentes.....	62
A.2.4.1.	Interacciones iónicas o Electroestáticas.....	62

A.2.4.2.	Interacciones ión-dipolo & dipolo-dipolo.	62
A.2.4.3.	Puentes de Hidrógeno.	63
A.2.4.4.	Complejos de transferencia de carga.	65
A.2.4.5.	Interacciones hidrofóbicas.	65
A.2.4.6.	Interacción π - Cation.	67
A.2.4.7.	Interacciones Metalión - π	68
A.2.4.8.	Puente de Halógeno.....	68
A.2.4.9.	Van der Waals o fuerza de dispersión de London.....	69
A.2.5.	Conclusiones del apéndice.	70
Apéndice A3: Tablas y gráficos.....		71
A.3.	TABLAS Y GRÁFICOS.....	72
A.3.1.	Distancia de entre los átomos de oxígeno O1-O2 y el átomo de Zn ²⁺ , pertenecientes al compuesto 11, durante la simulación.....	72
A.3.2.	Distancia entre los átomos de oxígeno y nitrógeno O1-N3 pertenecientes al compuesto 16 y el átomo de Zn ²⁺ , durante la simulación.....	73
A.3.3.	Distancia entre los átomos de oxígeno y nitrógeno O1-N3 pertenecientes al compuesto 35 y el átomo de Zn ²⁺ , durante la simulación.....	74
A.3.4.	Distancia entre los átomos de oxígeno y nitrógeno O1-N2 pertenecientes al compuesto 35 y el átomo de Zn ²⁺ , durante la simulación.....	75
A.3.4.1.	Fórmulas utilizadas en cálculos de afinidad de unión aproximada, ambos enfoques:	76
A.3.5.	Porcentaje (%) identidad al superponer las cadenas que contienen al sitio activo tanto de HDAC1 como HDAC2 en complejo a compuesto 38 de la serie.....	77
A.3.6.	Tablas de resultados	78
A.3.7.	Distribución de frecuencias de las interacciones no covalentes más comunes observadas en complejos proteína-ligando extraídos de la base de datos de proteínas PDB.	90
REFERENCIAS		91

INDICE DE FIGURAS

FIGURA 1.1. MODELO DE FARMACÓFORO Y LAS ESTRUCTURAS DE LOS MEDICAMENTOS APROBADOS. ADAPTADA REFERENCIA ^[12]	10
FIGURA 1.2. DEFINICIÓN DE LOS ELEMENTOS ESENCIALES DE UNIÓN DE LAS ORTO-AMINOANILIDAS A HDACS. ADAPTADA REFERENCIA ^[12]	11
FIGURA 1.3. TODAS LAS TÉCNICAS DE MODELADO MOLECULAR UTILIZADAS EN ESTA TESIS SON TRATADAS DE MANERA ESPECIAL Y MÁS PROFUNDA EN LOS APARTADOS DEL APÉNDICE A1, PARA QUIENES NO ESTÉN FAMILIARIZADOS CON LOS FUNDAMENTOS DE ESTA DISCIPLINA.....	11
FIGURA 3.1. EL ACOPLAMIENTO ENTRE EL COMPUESTO MÁS ACTIVO Y HDAC1.LA ENZIMA ESTÁ REPRESENTADA EN LA REPRESENTACIÓN DE SUPERFICIE Y EL INHIBIDOR EN MODELO VAN DER WAALS (IZQUIERDA). LA ESTRUCTURA SECUNDARIA SE REPRESENTA EN CINTAS (<i>RIBBONS</i>) Y EL INHIBIDOR SE MUESTRA COMO REPRESENTACIÓN DE PALOS (<i>STICK</i>), EL ÁTOMO DE ZINC COMO REPRESENTACIÓN DE ESFERA METÁLICA (DERECHA)	14
FIGURA 3.2. ANDAMIOS ESTRUCTURALES PARA LAS SERIES A, B, C, D Y E, DE LOS COMPUESTOS EN ESTUDIO.....	16
FIGURA 4.1. (A) ALINEACIÓN DE LOS COMPUESTOS ANCLADOS EN EL SITIO ACTIVO DE HDAC1. (B) ENFOQUE EN CUATRO MOLÉCULAS DEL CONJUNTO DE DATOS SELECCIONADAS PARA SIMULACIONES DE DINÁMICA MOLECULAR, 11 ROJA, 16 NARANJA, 35 CELESTE Y 38 MAGENTA. (C) REPRESENTACIÓN DE POSIBLES INTERACCIONES ENZIMA-LIGANDO EN DOS DIMENSIONES PARA EL COMPUESTO 38. LAS LÍNEAS DISCONTINUAS NEGRAS INDICAN ENLACES DE HIDRÓGENO, LA LÍNEA SÓLIDA VERDE MUESTRA INTERACCIONES HIDROFÓBICAS Y LAS LÍNEAS PUNTEADAS VERDES MUESTRAN INTERACCIONES DE APILAMIENTO π - π	22
FIGURA 4.2. VALIDACIÓN EXTERNA DEL CONJUNTO DE PRUEBA PARA COMFA.....	23
FIGURA 4.3. A LINEAMIENTO RG Y MAPAS DE CONTORNO COMFA-SE AL INTERIOR DEL SITIO ACTIVO DE HDAC1.	24
FIGURA 4.4. (A) ALINEACIÓN LB; (B) ALINEACIÓN RG, (C) CRISTAL 4LY1.PDB Y (D) ALINEACIÓN ASISTIDA.	25
FIGURA 4.5. VALIDACIÓN EXTERNA DE LOS CONJUNTOS DE PRUEBA PARA COMFA (IZQUIERDA) Y COMSIA (DERECHA).	28
FIGURA 4.6. GRÁFICAS DE LOS VALORES PIC_{50} EXPERIMENTALES FRENTE A LOS PREDICHOS PARA EL CONJUNTO DE ENTRENAMIENTO Y PRUEBA DE MOLÉCULAS PARA LOS MODELOS COMFA-SE (A) Y COMSIA-SEHA (B). PREDICCIONES PARA CADA MOLÉCULA EN EL CONJUNTO DE DATOS EN MODELOS COMFA (C) Y COMSIA (B) (SIN VALORES ATÍPICOS).	28
FIGURA 4.7. MAPAS DE CONTORNO COMFA SISTEMAS ESTÉRICO Y ELECTROSTÁTICO, PARA EL COMPUESTO MENOS ACTIVO A Y C RESPECTIVAMENTE; MÁS ACTIVO B Y D RESPECTIVAMENTE.	29
FIGURA 4.8. MAPAS DE CONTORNO COMSIA SISTEMAS ELECTROESTÁTICO Y ESTÉRICO, PARA EL COMPUESTO MENOS ACTIVO A Y C RESPECTIVAMENTE; MÁS ACTIVO B Y D RESPECTIVAMENTE.	31
FIGURA 4.9. MAPAS DE CONTORNO COMSIA SISTEMAS HIDROFÓBICO Y ACEPTOR DE ENLACE DE HIDROGENO, PARA EL COMPUESTO MENOS ACTIVO A Y C RESPECTIVAMENTE; MÁS ACTIVO B Y D RESPECTIVAMENTE.....	32
FIGURA 4.10. RMSDs PARA ÁTOMOS DE <i>BACKBONE</i> DEL RECEPTOR (NEGRO) E INHIBIDORES (ROJO) EN SIMULACIÓN DE DM DE 50 NS. (A) COMPLEJO HDAC1-11 (B) COMPLEJO HDAC1-16 (C) COMPLEJO HDAC1-35 Y (D) COMPLEJO HDAC1-38.	33
FIGURA 4.11. (IZQUIERDA) INSTANTÁNEAS DE CADA INHIBIDOR UTILIZADO PARA LA EXPLORACIÓN DEL SITIO DE UNIÓN HDAC1, EN MD. (DERECHA) CAVIDAD HIDROFÓBICA DE LA HDAC1.....	35
FIGURA 4.12. MAPA DE INTERACCIONES Y DESCOMPOSICIÓN DE ENERGÍA LIBRE POR RESIDUOS DE LOS COMPLEJOS HDAC1-11 (ARRIBA) Y HDAC1-16 (ABAJO), RESPECTIVAMENTE. LOS INHIBIDORES SE MUESTRAN COMO MODELOS CPK DE COLOR AZUL CIAN Y LOS RESIDUOS SE MUESTRAN COMO REPRESENTACIÓN DEL PALO VERDE, EL ÁTOMO DE ZINC COMO REPRESENTACIÓN DE LAS PERLAS METÁLICAS.	38
FIGURA 4.13. MAPA DE INTERACCIONES Y DESCOMPOSICIÓN DE ENERGÍA LIBRE DE RESIDUOS DE HDAC1-35 (ARRIBA) Y HDAC1-38(ABAJO). LOS INHIBIDORES SE MUESTRAN COMO MODELOS CPK DE COLOR AZUL CIAN Y LOS RESIDUOS SE MUESTRAN COMO REPRESENTACIÓN DEL PALO VERDE, EL ÁTOMO DE ZINC COMO REPRESENTACIÓN DE LAS PERLAS METÁLICAS.	39
FIGURA 4.14. METIONINA INTERACTUANDO CON PHE Y TYR. EL MODO DE UNIÓN HIDRÓGENO AROMÁTICO – AZUFRE MET PROPORCIONA UNA ESTABILIZACIÓN ADICIONAL SOBRE INTERACCIONES PURAMENTE HIDROFÓBICAS Y A DISTANCIAS MÁS LARGAS.....	41

FIGURA 4.15. LA CONFORMACIÓN DE UNIÓN E INTERACCIONES DE ENLACE DE HIDRÓGENO DEL COMPUESTO 38 EN EL SITIO ACTIVO DE HDAC1. LOS INHIBIDORES SE MUESTRAN COMO MODELOS CPK DE COLOR AZUL CIÁN Y LOS RESIDUOS COMO REPRESENTACIÓN EN BARRA; EL ÁTOMO DE ZINC COMO REPRESENTACIÓN EN PERLAS METÁLICAS TRANSPARENTES; LOS ENLACES DE HIDRÓGENO SE REPRESENTAN EN LÍNEAS PUNTEADAS DE COLOR CIAN Y SUS DISTANCIAS ESTÁN MARCADAS EN ANGSTROM.	42
FIGURA 5.1. PRINCIPALES RELACIONES ESTRUCTURA-ACTIVIDAD ENCONTRADAS EN ESTE ESTUDIO.....	46
FIGURA A.1.1 ORDEN JERÁRQUICO DE APROXIMACIONES DE MODELADO MOLECULAR A DIFERENTES ESCALAS DE TIEMPO Y LONGITUD. QM: MECÁNICA CUÁNTICA, MM: MECÁNICA MOLECULAR, MD: DINÁMICA MOLECULAR, CG: GRANO GRUESO [67].	48
FIGURA A.1.2. ILUSTRACIÓN DE TÉRMINOS ENERGÉTICOS DE ENLACE (ARRIBA) Y NO-ENLACE (ABAJO) DE UN CAMPO DE FUERZA ^[65]	51
FIGURA A.1.3. ALGORITMO BÁSICO DE LOS MÉTODOS DE MECÁNICA MOLECULAR.....	52
FIGURA A.1.4. ALGORITMO BÁSICO DE LOS MÉTODOS DE DINÁMICA MOLECULAR.	53
FIGURA A.1.5. PASOS DE MODELACIÓN QSAR.	57
FIGURA A.1.6. (A) REPRESENTACIÓN DE UN FARMACÓFORO CON SUS PROPIEDADES QUÍMICAS Y DISTANCIAS ENTRE SUS DIFERENTES REGIONES. HBA; ACEPTOR DE PUENTE DE HIDRÓGENO, RA; ANILLO AROMÁTICO, HY; HIDROFÓBICA, ZB; ENLAZANTE DE ZINC. (B) UNA MOLÉCULA SOLAPADA CON EL FARMACÓFORO.	58
FIGURA A.2.1. EQUILIBRIO ENTRE UN FÁRMACO, UN RECEPTOR, Y UN COMPLEJO FÁRMACO-RECEPTOR.....	60
FIGURA A.2.2. EJEMPLO DE INTERACCIÓN ELECTROSTÁTICA (IÓNICA). LA LÍNEA ONDULADA REPRESENTA LA CAVIDAD DEL RECEPTOR.	62
FIGURA A.2.3. EJEMPLOS INTERACCIONES IÓN-DIPOLO Y DIPOLO-DIPOLO. LA LÍNEA ONDULADA REPRESENTA CAVIDAD DEL RECEPTOR.	63
FIGURA A.2.4. EJEMPLOS DE ENLACES DE HIDRÓGENO. LA LÍNEA ONDULADA REPRESENTA LA CAVIDAD DEL RECEPTOR.....	64
FIGURA A.2.5. DOS EJEMPLOS (A Y B) DE CÓMO LOS ENLACES DE HIDRÓGENO INTRAMOLECULARES PUEDEN IMITAR UN HETEROCICLO BIOISOSTÉRICO.	64
FIGURA A.2.6. EJEMPLO DE UNA INTERACCIÓN DE TRANSFERENCIA DE CARGA. LA LÍNEA ONDULADA REPRESENTA LA CAVIDAD DEL RECEPTOR.....	65
FIGURA A.2.7. FORMACIÓN DE INTERACCIONES HIDROFÓBICAS [DE KOROLKOVAS, A. (1970). ESSENTIALS OF MOLECULAR PHARMACOLOGY, P. 172. WILEY, NUEVA YORK. ESTE MATERIAL FUE ADAPTADO Y SE REPRODUCE CON EL PERMISO DE JOHN WILEY & SONS, INC. Y CON EL PERMISO DE KOPPLE, K. D. 1966. PÉPTIDOS Y AMINOÁCIDOS. ADDISON-WESLEY, READING, MA.....	66
FIGURA A.2.8. EJEMPLOS DE INTERACCIONES HIDROFÓBICAS. (A) GRUPOS ALQUÍLICOS; (B) GRUPOS ARILOS. LA LÍNEA ONDULADA REPRESENTA LA CAVIDAD DEL RECEPTOR.	67
FIGURA A.2.9. POSIBLES DISPOSICIONES DE APILAMIENTO AROMÁTICO. (A) PARALELO CENTRADO EN LA CARA. (B) EN PARALELO DESPLAZADA. (C) FORMA DE T PERPENDICULAR. (D) FORMA DE Y PERPENDICULAR. (E) COMPENSACIÓN PARALELA A TOLUENO ^[105]	67
FIGURA A.2.10. EJEMPLO DE INTERACCIÓN CATION- π . LA LÍNEA ONDULADA REPRESENTA LA CAVIDAD DEL RECEPTOR.	68
FIGURA A.2.11. EJEMPLO DE INTERACCIÓN METALIÓN- π . BENCENO INTERACTUANDO CON ÁTOMO DE SODIO.	68
FIGURA A.2.12. EJEMPLO DE ENLACE HALÓGENO. UN COMPUESTO UNIDO A LA TIROSINA-612 DE LA FOSFODIESTERASA 5. LA LÍNEA ONDULADA REPRESENTA LA CAVIDAD DE LA ENZIMA.	69

INDICE DE TABLAS

TABLA 3.1. ESTRUCTURA Y ACTIVIDADES BIOLÓGICAS DE LOS COMPUESTOS ESTUDIADOS ^[10] . LOS COMPUESTOS EN NEGRITAS 11, 16, 35 Y 38, FUERON USADOS POSTERIORMENTE PARA SIMULACIONES DE DINÁMICA MOLECULAR.	15
TABLA 4.1. PARÁMETROS ESTADÍSTICOS PARA EL MODELO COMFA-SE, ESQUEMA RG. VALIDACIÓN EXTERNA.	23
TABLA 4.2. RESUMEN DE LOS PARÁMETROS DE VALIDACIÓN EXTERNA PARA COMFA.	24
TABLA 4.3. BÚSQUEDA SECUENCIAL PARA LA GENERACIÓN DE LOS MEJORES MODELOS 3D-QSAR ^A . EN NEGRITA LOS MEJORES MODELOS OBTENIDOS MEDIANTE EL ALINEAMIENTO ASISTIDO POR DIFRACCIÓN DE RAYOS-X.	26

TABLA 4.4. RESUMEN DE LOS PARÁMETROS DE VALIDACIÓN EXTERNA PARA COMFA Y COMSIA.	27
TABLA 4.5. VALORES DE ENERGÍA DE UNIÓN Y ENERGÍA DE LOS COMPONENTES INDIVIDUALES CALCULADOS CON EL MÉTODO MM-PBSA PARA HDAC1 EN COMPLEJO CON LAS MOLÉCULAS 11, 16, 35 Y 38 (KCAL/MOL).	35
TABLA 4.6. CONTRIBUCIÓN DE ENERGÍA POR RESIDUO A LA UNIÓN DE LAS MOLÉCULAS 11, 16, 35 Y 38; EN NEGRITA RESIDUOS COMUNES SELECCIONADOS DEL SITIO ACTIVO DEL RECEPTOR HDAC1 (KCAL/MOL).	36
TABLA 4.7. OCUPACIÓN INTERACCIÓN DE HIDRÓGENO ENTRE LOS RESIDUOS ASP170, HIE135 Y GLY143 Y 38.	41
TABLA 4.8. OCUPACIÓN DE ENLACE DE HIDRÓGENO ENTRE ÁTOMOS DEL SOLVENTE E HIDRÓFILOS DE LA MOLÉCULA 38. OCUPACIÓN DE ENLACE DE HIDRÓGENO ENTRE DISOLVENTE Y RESIDUOS HIN172, TYR297 Y HIE135.	43
TABLA A.3.1. BÚSQUEDA DE LOS MEJORES MODELOS DE COMFA Y COMSIA USANDO COMBINACIONES DE CAMPOS PARA RG/X-RAY ASISTIDO.	78
TABLA A.3.2. ACTIVIDADES PREDICHAS Y SUS RESIDUALES DE LOS MODELOS COMFA-SE(RG).	79
TABLA A.3.3. TABLA DE VALORES DE LAS ACTIVIDADES PREVISTAS Y SUS RESIDUOS OBTENIDA A PARTIR DE LOS MODELOS COMFA-SE Y COMSIA-SEHA ALINEAMIENTO ASISTIDO POR DIFRACCIÓN RAYOS-X.	80
TABLA A.3.4. INTERACCIONES NO COVALENTES, ENTRE ÁTOMOS PERTENECIENTE A RESIDUO QUE INTERACCIONA DESDE RECEPTOR Y OTRO PERTENECIENTE AL LIGANDO. DISTANCIA MEDIA, DESVIACIÓN ESTÁNDAR Y ERROR ESTÁNDAR DE LA MEDIA. RECEPTOR HDAC1 Y LIGANDOS (11, 16, 35 Y 38).	81
TABLA A.3.5. OCUPACIÓN DE ENLACE DE HIDRÓGENO ENTRE LOS RESIDUOS HDAC1 Y COMPUESTOS 11, 16, 35 Y 38.	86
TABLA A.3.6. OCUPACIÓN POR ENLACE DE HIDRÓGENO ENTRE LOS ÁTOMOS DEL SOLVENTE Y ÁTOMOS HIDRÓFILOS DE LOS COMPUESTOS 11, 16, 35 Y 38. OCUPACIÓN DE LOS PATRONES HB PARA LOS RESIDUOS DEL SITIO ACTIVO HDAC1 PRESENTES EN DIFERENTES COMPLEJOS (HDC1/11/16/35/38).	87
TABLA A.3.7. ALGORITMO DE BÚSQUEDA Y FUNCIÓN DE PUNTUACIÓN DEL PROGRAMA AUTODOCK.	89

INDICE DE ESQUEMAS

ESQUEMA 3.1. FLUJO DE TRABAJO METODOLÓGICO.	13
---	----

RESUMEN

Antecedentes. La superfamilia de metaloenzimas histonas desacetilasas (HDACs) está ampliamente implicada en la regulación epigenética de la expresión génica. Catalizan la escisión del grupo N-acetilo de los residuos de lisina acetilados situados en las colas de las histonas nucleosómicas centrales H2a, H2b, H3 y H4. La sobreexpresión de la enzima histona desacetilasa da lugar a cromatina hipoacetilada que se vuelve inaccesible para los factores de transcripción. Estas enzimas también regulan el estado de acetilación de numerosas proteínas no histónicas, como los factores de transcripción p53, STAT3 y los receptores de andrógenos y glucocorticoides en el núcleo celular, así como la α -tubulina y la Hsp90 en el citoplasma. La actividad catalítica de las enzimas histonas desacetilasas (HDAC) es directamente relevante en la patogénesis del cáncer y otras enfermedades.

Resultados. En este trabajo, informamos de un estudio de modelado molecular combinado que se realizó sobre una serie de 39 inhibidores de HDAC1 derivados del compuesto CI-994. Validamos los modelos 3D-QSAR mediante la comparación de los dos esquemas principales de alineación, el basado en el ligando (LB) y el guiado por el receptor (RG) que fue asistido por cristalografía de rayos X. El alineamiento guiado por el receptor usando *docking* generó resultados estadísticamente satisfactorios para un modelo CoMFA-SE ($q^2 = 0.560$, $r^2_{\text{test}} = 0.8059$, $r^2_{\text{ncv}} = 0.949$, $N = 4$), este alineamiento luego fue asistido por difracción de rayos X, para construir modelos para los cuales se obtuvieron los siguientes resultados: CoMFA-SE ($q^2 = 0.586$, $r^2_{\text{test}} = 0.7994$, $r^2_{\text{ncv}} = 0.961$, $N = 0.4$) y CoMSIA-SEHA ($q^2 = 0.570$, $r^2_{\text{test}} = 0.7850$, $r^2_{\text{ncv}} = 0.963$, $N=5$). Además, se llevaron a cabo simulaciones de dinámica molecular (MD) de 50 ns en solvente explícito y cálculos de energía de interacción para cuatro diferentes complejos Ligando-Receptor para determinar los modos de unión e identificar residuos cruciales que interactúan vía descomposición por residuos. Las simulaciones MD exploraron más a fondo las interacciones detalladas entre la proteína y los ligandos y los resultados fueron consistentes con los hallazgos obtenidos desde los modelos 3D-QSAR y el docking molecular. Las energías de interacción calculadas por el método MMPBSA mostraron la importancia de la interacción electrostática, van der Waals y la contribución no polar a la energía libre de unión.

Conclusiones. Se ha observado una buena correlación entre los resultados de MD, los estudios de docking y el análisis del mapa de contorno de los modelos 3D-QSAR. El estudio ha identificado los residuos clave en el bolsillo de unión de HDAC1 para la interacción de unión del ligando. Los resultados de este estudio pueden aportar algunas ideas para el desarrollo de nuevos y potentes inhibidores de HDAC1 de estructura general 4-Acetilamino-N-(2'-aminofenil) benzamida.

Palabras clave: HDAC1; HDACi; 4-Acetamido-N-(2-aminofenil) benzamida; epigenética; cáncer; modelado molecular; Docking; 3D-QSAR; simulaciones MD; MMPBSA

SUMMARY

Background. The histone deacetylase (HDACs) superfamily of metalloenzymes is extensively involved in epigenetic regulation of gene expression. They catalyze the cleavage of the N-acetyl group from acetylated lysine residues located on the tails of the core nucleosomal histones H2a, H2b, H3, and H4. However, overexpression of histone deacetylase leads to hypoacetylated chromatin that becomes inaccessible to transcription factors. These enzymes also regulate the acetylation status of numerous nonhistone proteins such as transcription factors p53, STAT3, Androgen and Glucocorticoid receptors in the nucleus, as well as α -tubulin and Hsp90 in the cytoplasm. The catalytic activity of the histone deacetylase (HDAC) enzymes is directly relevant to the pathogenesis of cancer as well as several other diseases.

Results. In this work, we report a combined molecular modeling study was performed on a series of CI-994 derivatives as HDAC1 inhibitors. We validated the 3D-QSAR models by comparison of two major alignment schemes, namely ligand-based (LB) and Receptor-guided (RG) assisted by X-Ray diffraction. The receptor-guided alignment generated statistically satisfactory results for a CoMFA-SE model ($q^2 = 0.560$, $r^2_{\text{test}} = 0.8059$, $r^2_{\text{ncv}} = 0.949$, $N = 4$), this alignment which was then assisted by X-ray diffraction, to construct models for which the following results were obtained: CoMFA-SE ($q^2 = 0.586$, $r^2_{\text{test}} = 0.7994$, $r^2_{\text{ncv}} = 0.961$, $N = 0.4$) and CoMSIA-SEHA ($q^2 = 0.570$, $r^2_{\text{test}} = 0.7850$, $r^2_{\text{ncv}} = 0.963$, $N=5$). Furthermore, a 50-ns molecular dynamics (MD) simulation in explicit solvent and binding free energy calculations were carried out on four different Receptor-Inhibitors complexes to determine the binding modes and to identify crucial interacting residues. The rationality and reliability of molecular docking and 3D-QSAR results were validated by MD simulation. The binding free energies calculated by the MM/PBSA method showed the importance of the Electrostatic, van der Waals interaction and the nonpolar contribution to the binding free energy.

Conclusions. A good correlation between the MD results, docking studies, and the contour map analysis were observed. The study has identified the key amino acid residues in the HDAC1 binding pocket for ligand binding interaction. The results from this study can provide some insights into the development of potent, novel HDAC1 inhibitors of general structure 4-Acetylamino-N-(2'-aminophenyl) benzamide.

Keywords: HDAC1; HDACi; 4-Acetamido-N- (2-aminophenyl) benzamide; epigenetic; cancer; molecular modeling; Docking; 3D-QSAR; MD simulations; MMPBSA

CAPÍTULO 1 - INTRODUCCIÓN

El cáncer puede ser causado por cambios o mutaciones en el ADN [1]. El ADN es la sustancia química en nuestras células que compone nuestros genes. La expresión genética es regulada a través de un control transcripcional mediado por el ARN, promoviendo la síntesis de proteínas y son estas en gran parte las que determinan las características y la función de una célula. Ciertos genes regulan cuándo las células crecen, se dividen, se mantienen vivas, o cuándo mueren, éstos son conocidos como *oncogenes*. El cáncer está asociado a una activación anormal de los oncogenes o dicho de otra forma a la desactivación de los genes supresores de tumores [1].

El conjunto de todas las modificaciones químicas a las cuales está sujeto el ADN de una célula es llamado *epigenoma*. Los cambios epigenéticos son parte del desarrollo normal de un organismo, en efecto, las células de un embrión comienzan con un genoma maestro y a medida que las células se dividen, algunos genes se activan y otros se desactivan, diferenciándose en tipo de tejido y función. Las “etiquetas” o modificaciones químicas que activan o desactivan los genes están en un diálogo de toda una vida y son influenciados por factores como la dieta, exposición a sustancias químicas y medicamentos. De todas las modificaciones químicas posibles, la deacetilación surge como un cambio epigenético que puede ser modulado a través de fármacos inhibiendo a las enzimas Histonas Deacetilasas (HDACs) [2]. En esa línea, la familia de las HDACs se ha convertido en un interesante objetivo biológico contra el cáncer en los últimos años, debido a que está intrínsecamente involucrada en la regulación epigenética de la expresión génica [3]. La sobreexpresión de las histonas deacetilasas está asociada con cromatina que se vuelve inaccesible a los factores de transcripción [3], por lo que la activación anormal de las HDACs está implicada en el silenciamiento epigenético de genes supresores de tumores [4].

Existen cuatro clases distintas de HDACs: clase I (HDAC1, 2, 3 y 8), clase IIa (HDAC4, 5, 7 y 9), clase IIb (HDAC6 y 10) y clase IV (HDAC11), que representan las deacetilasas dependientes de Zn^{+2} . Estas HDACs son responsables de remover los grupos N-acetil desde los residuos de lisina acetilada localizados en las colas de las histonas centrales H2a, H2b, H3 y H4 del nucleosoma [3]. Recientemente, varios estudios han demostrado que ciertos miembros de las familias de las HDACs se expresan de forma aberrante en varios tumores y tienen una función no redundante en el control de los atributos de las células cancerígenas. En particular, las HDACs pertenecientes a la Clase I son consideradas por muchos investigadores como las dianas más relevantes para el tratamiento del cáncer, esto debido a que los inhibidores que poseen actividad contra HDACs 1, 2, 3 y 8, generalmente muestran un fuerte efecto anti proliferativo e inductor de actividad apoptótica. En esta investigación se seleccionó la Histona Deacetilasa 1 como receptor específico, ya que juega un importante rol en varias líneas de células cancerígenas. Cabe destacar que las HDACs de la Clase I se encuentran desreguladas en células y tejidos tumorales en líneas cancerígenas de las más agresivas y con mayor prevalencia a nivel mundial/nacional. Se ha observado sobreexpresión de HDAC1 en varios cánceres humanos, incluyendo el de ovario, mama, próstata, páncreas y leucemia, con actividad sobre la detención del ciclo celular, proliferación, actividad apoptótica y sensibilización para mejoras en resultados de quimioterapia [6].

1.1. Estado del Arte: Inhibidores

Actualmente, existen cuatro clases estructurales de inhibidores de HDACs: ácidos hidroxámicos, ácidos grasos de cadena corta, grandes péptidos cíclicos y benzamidas [5-9]. El ácido hidroxámico de Suberoylanilida (SAHA, Vorinostat; *Merck*) y Romidepsina (FK-228; *Gloucester Pharmaceuticals*) fueron aprobados para el tratamiento del linfoma cutáneo de células T dentro de los 10 años siguientes a su descubrimiento [7]. El inhibidor pan-HDAC comúnmente utilizado, SAHA, parece inhibir las HDACs de Clase I y IIb, con una actividad significativamente menor contra los HDAC de Clase IIa, y su perfil de inhibición se comparte con otros inhibidores aprobados como Belinostat (PXD-101; *TopoTarget*) y Panobinostat (LBH-589; *Novartis*) [8] ver figura 1.1.

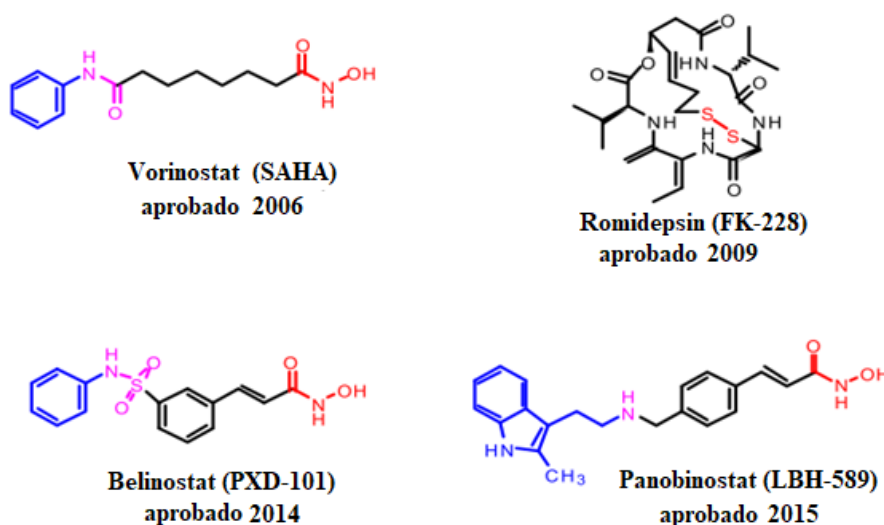


Figura 1.1. Modelo de farmacóforo y las estructuras de los medicamentos aprobados. Adaptada referencia [12].

Los estudios proteómicos muestran que sólo el 10% de todos los sitios de acetilación son sensibles al tratamiento con SAHA y enfatizan aún más en la necesidad de nuevos inhibidores HDACs [9]. El diseño de nuevos inhibidores de HDAC selectivos sobre una clase específica de HDACs es crucial para mejorar las diversas limitaciones que presentan los tratamientos actuales contra el cáncer. El desarrollo de inhibidores selectivos de HDACs es una manera significativa de prevenir los efectos secundarios de los actuales HDACIs basados en ácidos hidroxámicos, que principalmente causan diarrea, fatiga muscular y una baja en las células sanguíneas, teniendo que ser suministrados generalmente en combinación a otros fármacos. Los HDACIs basados en ácidos hidroxámicos no ejercen una selectividad deseada hacia una isoforma específica, mientras que las 2-aminobenzamidas ejercen selectividad sobre HDACs pertenecientes a la Clase I y las isoformas de interés (HDAC 1, 2 y 3); como ejemplo tenemos a Entinostat (MS-275; *Syndax pharmaceuticals Inc*) y Mocetinostat (MGCD0103; *MethylGene Inc*) que están en ensayo clínico fase 2 y 1 respectivamente, que comparten un perfil de inhibición similar [12]. Los fármacos antes mencionados, con selectividad por HDAC1, presentan tanto similitudes farmacofóricas de enlazamiento como perfil de inhibición en HDAC2.

Este trabajo se basó en 39 moléculas inhibidores del tipo benzamida específicamente *orto-aminoanilidas* [10], los cuales comparten un farmacóforo común que está compuesto por un grupo de unión al zinc (rojo), una región de unión generalmente apolar (negro), unidad conectora polar (rosado) y un dominio de reconocimiento de superficie (azul) [5-10,12], ver figura 1.2.

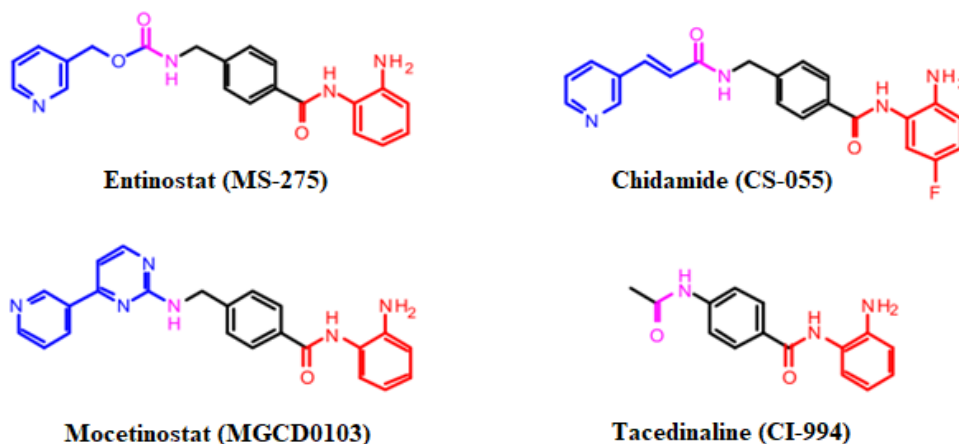


Figura 1.2. Definición de los elementos esenciales de unión de las orto-aminoanilidas a HDACs. Adaptada referencia [12].

A través de este estudio de modelado molecular apoyado de interfaces gráficas se busca explorar y lograr nuevo conocimiento sobre el modo de unión de los inhibidores derivados del fármaco CI-994 en complejo a HDAC1 [10], obteniendo una visualización que facilite potencialmente la comprensión de los mecanismos moleculares relacionados con los inhibidores de HDACs de la Clase I [7-12]. También aportar con información que permita definir preliminarmente los requerimientos estructurales que rigen la selectividad de inhibidores de HDACs de la Clase I derivados de las orto-aminoanilidas (ver figura 1.3). Para de esa forma contribuir en la etapa de conceptualización del proceso de diseño de nuevos inhibidores [11,12].

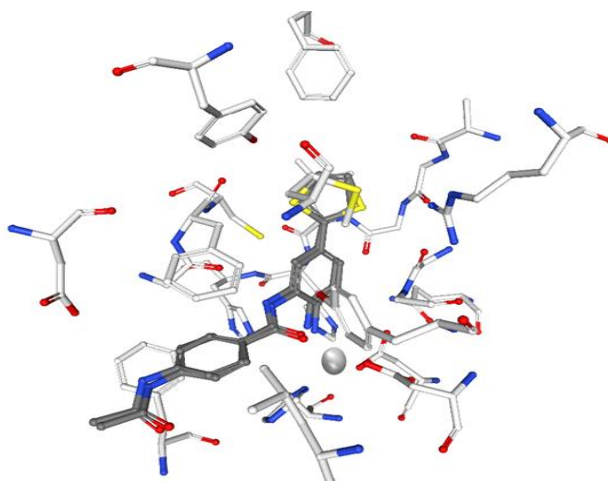


Figura 1.3. Modelación molecular del sitio activo de HDACs en complejo a inhibidor derivado de las orto-aminoanilidas. (Todas las técnicas de modelado molecular utilizadas en esta tesis son tratadas de manera especial y más profunda en los apartados del apéndice A1, para quienes no estén familiarizados con los fundamentos de esta disciplina.)

CAPÍTULO 2 - OBJETIVO DE INVESTIGACIÓN

Entre las preguntas que guían esta investigación destacan las siguientes: (1) ¿Cuál sería el potencial que supondría dirigir la síntesis de este tipo de moléculas a través del diseño de fármacos asistido por computadora?, (2) ¿Podremos obtener nuevos fármacos con mayor actividad y selectividad sobre la modulación de la acción biológica de enzimas HDACs de la clase I a través de esta metodología?, (3) ¿Cuál será la factibilidad sintética de síntesis para su posterior uso en tratamientos contra el cáncer basados en regulación epigenética?, preguntas unificadas en el afán de disminuir indeseados efectos colaterales que causan actuales inhibidores.

En base a la revisión bibliográfica, el desafío de la propuesta de esta tesis de Magister es realizar un estudio estructural y energético del modo de unión de fármacos de tipo *orto*-aminoanilida en complejo a la enzima histona deacetilasa 1 (HDAC1), utilizando para ello la combinación de diversas herramientas computacionales (QM, *Docking*, 3D-QSAR y Dinámica Molecular) e información experimental (Difracción de rayos-X y análisis *in vitro* de citotoxicidad).

2.1. Hipótesis

H1. Desde un estudio de diseño de fármacos asistido por computadora es posible obtener nuevas moléculas basadas en el esqueleto benzamida (tipo molécula CI-994) con afinidad de unión mejorada en el receptor de interés HDAC1 que moléculas *Hits* con actividad biológica informada.

H2. En base a la disponibilidad de la estructura cristalizada de la enzima HDAC1, postulamos que a partir de un alineamiento molecular derivado de estudios de *docking*, es posible obtener relaciones de estructura actividad cuantitativas útiles mediante técnicas 3D-QSAR con valores de $q^2 > 0.5$ y $r^2 > 0.6$.

2.2. Objetivos Generales

Lograr un mejor y nuevo entendimiento estructural del modo de unión de los inhibidores derivados del fármaco CI-994 con la enzima HDAC1, para optimizar compuestos *Hits*.

2.3. Objetivos Específicos

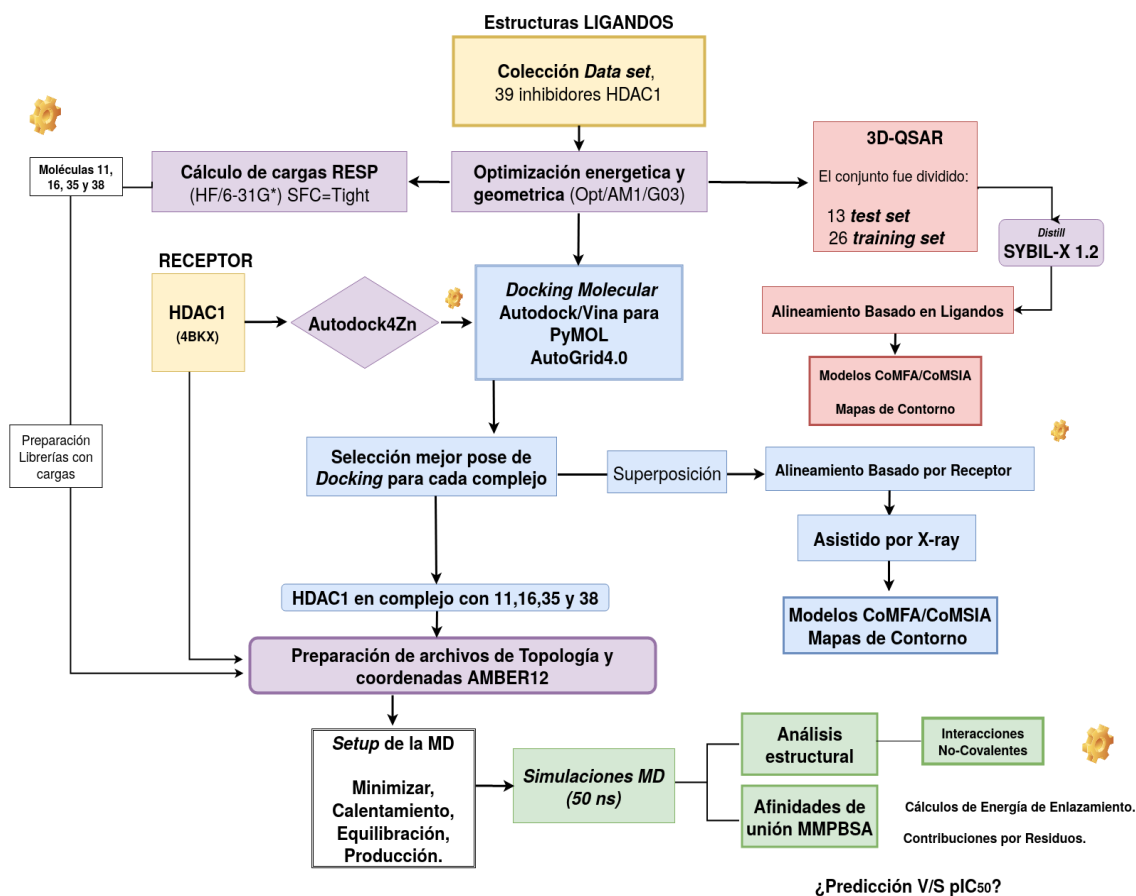
1. Explorar a nivel atómico el sitio activo de la metaloenzima HDAC1 en complejo a 4 inhibidores, para lograr una comprensión mejorada de la estructura y dinámica de las interacciones implicadas en el sistema fármaco-receptor.
2. Integrar diferentes técnicas para potenciar y mejorar el alineamiento de moléculas derivadas del fármaco CI-994, paso clave en la generación de modelos 3D-QSAR estadísticamente confiables.
3. Proponer y diseñar nuevas moléculas.

CAPÍTULO 3 - METODOLOGIAS

La estrategia se basó en la información proveniente desde experimentos de difracción de rayos-X existente tanto para HDAC1 como HDAC2, esta última en complejo a compuestos tipo *orto-aminoanilidas*, sacando ventaja de la resolución que ofrece este tipo de experimentos a nivel de comprensión estructural, para saber en cuales interacciones entre compuestos Hits y aminoácidos del sitio activo del receptor concentrarse.

El estudio de las interacciones no covalentes en el complejo fármaco-receptor juegan un papel clave en el estudio de las HDACs y sus procesos biológicos, un buen entendimiento de estas puede guiar la síntesis de nuevos compuestos, por esta razón se le dan un tratamiento especial en el Apéndice A2. En la presente tesis, se reportan resultados derivados de dichas interacciones en el estado ligado o complejo fármaco-receptor, la metodología es por definición un estudio de modelación molecular combinado (ver esquema 3.1):

- (I) **Mediante el uso de acoplamiento molecular (Docking).** Aproximación que ofrece un resultado rápido que ayuda a entender el proceso de unión.
- (II) **Mediante simulaciones de dinámica molecular MDs.** Cambios conformacionales y energía de enlazamiento, usando método de punto final.
- (III) **Estudios de análisis cuantitativo de la relación estructura-actividad tridimensional 3D-QSAR.** Utilizando alineamiento s moleculares basados en ligando y receptor [11,12]. Proponiendo nuevas moléculas e introduciéndolas al modelo.



Esquema 3.1. Flujo de trabajo metodológico.

3.1. Selección de la estructura inicial

Para lograr realizar I, II y II del apartado anterior es fundamental contar con la estructura cristalina de alta resolución de HDAC1 para humanos (*homo sapiens*), esta se obtuvo desde el banco de datos de proteínas (PDB ID: 4BKX), donde una noción preliminar de los residuos del sitio activo involucrados en la inhibición de HDAC1 fueron reportados [13]. Esta estructura fue tomada como estructura inicial para el acoplamiento molecular y la dinámica molecular (Figura 3.1).

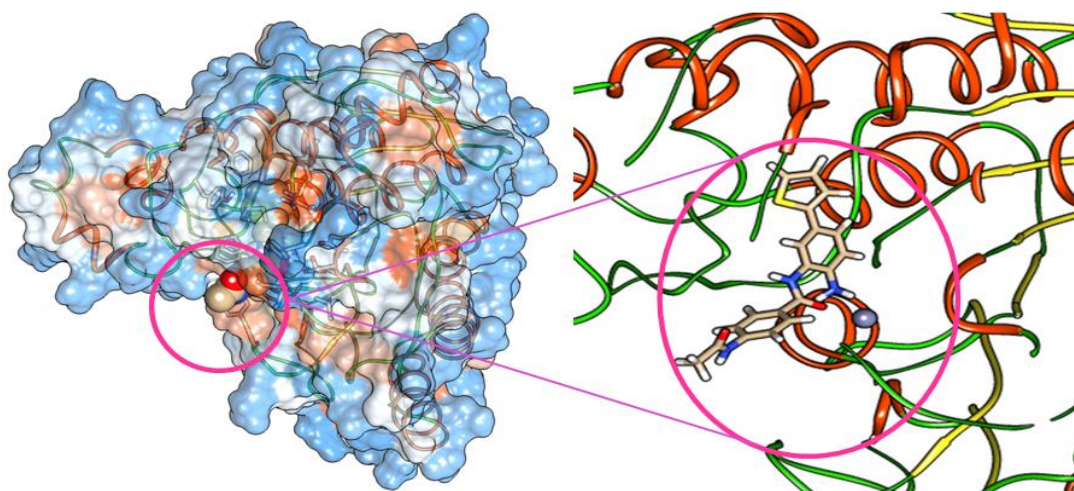


Figura 3.1. El acoplamiento entre el compuesto más activo y HDAC1. La enzima está representada en la representación de superficie y el inhibidor en modelo van der Waals (izquierda). La estructura secundaria se representa en cintas (*ribbons*) y el inhibidor se muestra como representación de palos (*Stick*), el átomo de zinc como representación de esfera metálica (derecha).

3.2. Colección del conjunto de datos

Un total de 39 inhibidores de la enzima HDAC1, derivados del fármaco CI-994, de estructura general 4-Acetilamino-N-(2'-aminofenil) benzamida, reportados por Joey L. Methot & colaboradores, fueron tomados para el estudio (ver Tabla.3.1) [10]. Antes del estudio 3D-QSAR y los cálculos de acoplamiento, todos los inhibidores utilizados en este estudio teórico fueron optimizados energéticamente y geoméricamente en un nivel de teoría semiempírica, usando para ello el método AM1, del programa Gaussian 03 [14]. Para los estudios 3D-QSAR el conjunto de datos se dividió manualmente en un conjunto de entrenamiento para la generación de modelos y un conjunto de pruebas para la validación de modelos.

Los compuestos del conjunto de datos se dividieron inicialmente en 26 para el conjunto de entrenamiento y 13 para el conjunto de prueba, teniendo en cuenta la diversidad estructural (figura 3.2) y la distribución de actividades (Tabla.3.1).

Tabla 3.1. Estructura y actividades biológicas de los compuestos estudiados ^[10].

Compuesto	Sustituyente R	Serie	pIC ₅₀
1	Fenilo	A	7,237
2	2-Piridilo	A	6,509
3	3-Piridilo	A	7,168
4	4-Piridilo	A	7,125
5	2-Tiofenilo	A	6,886
6	2-Benzotiofenilo	A	6,886
7	2-Benzotiazolilo	A	6,456
8	2-Furanilo	A	6,886
9	5-Isoxazolilo	A	7,000
10	3-Aminofenilo	A	6,357
11	4-Aminofenilo	A	7,398
12	Acetilo	A	5,319
13	Hidrocinaamilo	A	6,215
14	Cinaamilo	A	5,921
15	Hidrógeno	B	5,585
16	Fenilo	B	7,319
17	2-Tiofenilo	B	7,187
18	3-Tiofenilo	B	7,284
19	3-Furanilo	B	6,796
20	1-Imidazolilo	B	6,000
21	4-p Piridilo	B	6,187
22	1- Pirrolidinilo	B	6,174
23	Ciclopentflico	B	5,377
24	Flúor	C	6,854
25	NH ₂	C	6,482
26	NMe ₂	C	5,276
27	CH ₂ CN	C	6,602
28	CH ₂ NH ₂	C	5,678
29	CO ₂ CH ₂ Ph	C	5,886
30	CH ₂ CH ₂ CO ₂ H	C	4,886
31	CH ₂ CH ₂ CO ₂ Me	C	5,824
32	Hidrógeno	D	5,620
33	Fenilo	D	7,222
34	2-Tiofenilo	D	7,319
35	R₁:NH₂ , R₂ : H	E	6,244
36	R ₁ :OH , R ₂ : Fenilo	E	7,745
37	R₁:NH₂ , R₂:Fenilo	E	7,553
38	R₁:NH₂, R₂: 2-Tiofeno	E	8,155
39	R ₁ :NH ₂ , R ₂ : 3-Tiofeno	E	8,155

(Los compuestos en negritas 11, 16, 35 y 38, fueron usados posteriormente para simulaciones de dinámica molecular).

Los compuestos del conjunto de pruebas se seleccionaron aleatoriamente, conteniendo compuestos en un rango uniforme de valores de actividad biológica desde baja a alta actividad. La magnitud del logaritmo negativo del valor de IC_{50} (pIC_{50}) van de 4,88 a 8,15 cubriendo un intervalo de más de 3 unidades logarítmicas, que se considera apto para estudios 3D QSAR [15-18].

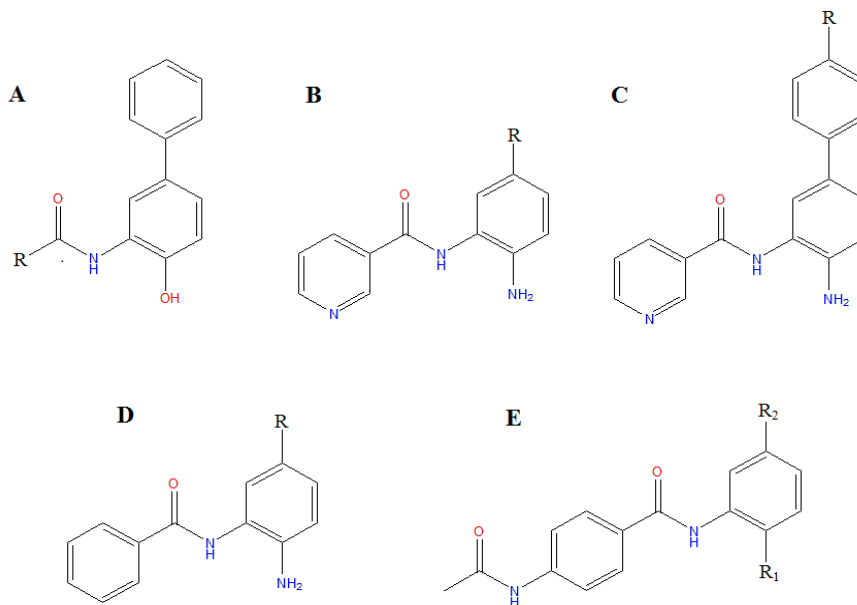


Figura 3.2. Andamios estructurales para las series A, B, C, D y E, de los compuestos en estudio.

3.3. Acoplamiento molecular

Los estudios de acoplamiento se llevaron a cabo utilizando el software AutoDock 4.2.6, AutoDock Tools 1.5.7 [19,20] y AutoDock4Zn. Este último es un campo de fuerza AutoDock mejorado para el acoplamiento de pequeñas moléculas a metaloproteínas de zinc [21]. Se utilizó el *plugin* Autodock/Vina [22] para PyMol [23] para mejorar la realización de cálculos de acoplamiento. La estructura inicial del receptor fue cargada en PyMol [24]. La estructura de la proteína fue preparada removiendo moléculas de agua, agregando cargas Kollman e hidrógenos polares. En primer lugar, los residuos alrededor de 6 Å del ligando co-cristalizado fueron considerados como el sitio de unión para los estudios de *Docking*. Todas las moléculas del conjunto de datos se acoplaron al sitio activo de HDAC1. Luego, usando el programa auxiliar AutoGrid4.0 se generó un mapa de cuadrícula utilizando las coordenadas x , y , z para encontrar un centro de una cuadrícula pequeña, adecuado y lo más representativo para el sitio activo. Las dimensiones de la rejilla se fijaron a partir de 45 x 45 x 45 puntos, con una distancia entre rejillas de 0,375 Å. El número de corridas de *Docking* se fijó en 100. La población en el algoritmo genético fue de 150. Después del acoplamiento, las 100 posiciones de acoplamiento se agruparon (*clustered*) en grupos con desviaciones RMS inferiores a 1,0 Å. Entre todo el grupo de complejos pronosticados por AutoDock, la conformación del grupo más poblado, junto con la conformación de menor energía para el compuesto más activo acoplado al receptor, se utilizó como plantilla para la alineación del acoplamiento.

3.4. Alineamientos y estudios 3D-QSAR

La alineación molecular de compuestos es un paso crucial en el desarrollo de modelos 3D-QSAR [25]. Recientemente se han publicado varios estudios en los que se ha aplicado con éxito la combinación de métodos basados en ligandos y guiados por receptor para el diseño y predicción de compuestos bioactivos [16-18, 24-27].

3.4.1. Alineamiento basado en ligandos (LB)

Todas moléculas del conjunto de datos optimizadas fueron cargadas en el programa Sybyl X 1.2 [28]. Las cargas MMFF94 se aplicaron como cargas parciales. La estadística del modelo depende de la alineación de las moléculas en su conformación bio-activa [24], se asumió entonces que cada molécula se une al sitio activo de manera similar, ya que comparten un andamiaje común. Durante el procedimiento, todas las moléculas del conjunto de datos, se alinean alrededor del núcleo de la benzamida central utilizando la función *Distill* incluida en SybylX 1.2 [28].

3.4.2. Alineamiento guiado por receptor (RG)

Las moléculas del conjunto de datos optimizadas se acoplaron al sitio activo de HDAC1. La conformación de energía más baja y la postura mejor acoplada del compuesto más activo se utilizó como pose de plantilla para la alineación por *Docking*. De esta manera, el resto de las moléculas se superpusieron sobre esta plantilla (dentro del bolsillo de unión), obteniéndose así el alineamiento molecular guiado por el receptor para posteriores análisis, con la finalidad de mostrar y entender cómo estas moléculas podrían unirse al sitio activo [25,26].

3.4.3. Validación interna y análisis PLS: Estudios CoMFA y CoMSIA

En CoMFA, los campos estérico y electrostático se calcularon por separado utilizando un carbono sp³ como átomo sonda con los valores de corte de energía de 30 kcal/mol. Los modelos se generaron utilizando parámetros por defecto [29]. Para generar modelos 3D-QSAR estadísticamente significativos, se utilizó la regresión de mínimos cuadrados parciales (PLS), la cual evalúa el conjunto de entrenamiento correlacionando la variación en sus valores pIC50 con las variaciones en sus campos descriptores CoMFA/CoMSIA. Para analizar la fiabilidad de los modelos generados a partir del análisis PLS, se realizó un análisis de validación cruzada con la metodología "leave-one-out" (LOO) y distancia de muestra (Use SAMPLS); obteniéndose el número óptimo de componentes (N) y el cuadrado del coeficiente de validación cruzada (q²), el cual se calcula según la ecuación (3.1) [29]:

$$q^2 = 1 - \frac{\sum(y_i - y_{pred})^2}{\sum(y_i - \bar{y})^2} \quad (3.1)$$

Donde y_i , \bar{y} , e y_{pred} son valor actual, promedio, y predicho respectivamente de la actividad biológica.

A continuación, se realizó un análisis sin validación cruzada utilizando el número óptimo de componentes N obtenido anteriormente. La validación no cruzada se realizó con un valor de filtro de columna de 2,0 para acelerar el análisis y reducir el ruido.

De este análisis se calcularon el coeficiente de Pearson (r_{ncv}^2), el error estándar de las estimaciones (SEE), F del test de Fisher y la contribución de cada campo [29].

En el análisis de índices de similitud molecular comparativa (CoMSIA) [16,29], se calcularon los descriptores estéricos, electrostáticos, hidrofóbicos, de donante y de aceptor de enlaces de hidrógeno, utilizando un átomo sonda de radio 1,0 Å. Se utilizó un factor de atenuación de 0,30. Se generaron modelos CoMSIA con diferentes combinaciones. De estos modelos desarrollados, se seleccionó un modelo de CoMSIA estadísticamente razonable en términos de q^2 , r^2 y SEE [29].

3.4.4. Validación externa de modelos 3D-QSAR

Para evaluar la capacidad de predicción real de los modelos generados por los análisis CoMFA/CoMSIA, ambos modelos fueron sometidos a criterios de validación externa de acuerdo con la prueba propuesta por Golbraikh y Tropsha [30, 31], el cual utiliza los valores de y_i e y_{pred} de los compuestos del conjunto de prueba y los grafica de dos formas:

- (i) En primer lugar, y_i v/s y_{pred} fijando el intercepto en 0 y obteniendo de la regresión, la recta $y_i = k y_{pred}$, con un coeficiente de correlación r_0^2 que debe ser cercano a 1.
- (ii) En segundo lugar, y_{pred} v/s y_i fijando el intercepto en 0 y obteniendo de la regresión la recta $y_{pred} = k' y_i$, con un coeficiente de correlación $r_0'^2$ que debe ser cercano a 1.

En base lo anterior, se considera predictivo un modelo QSAR, si se cumplen las siguientes condiciones:

$$q^2 > 0,5 \quad (3.2)$$

$$r^2 > 0,6 \quad (3.3)$$

$$\frac{(r^2 - r_0^2)}{r^2} < 0,1 \text{ o } \frac{(r^2 - r_0'^2)}{r^2} < 0,1 \quad (3.4)$$

$$0,85 \leq k \leq 1,15 \text{ o } 0,85 \leq k' \leq 1,15 \quad (3.5)$$

$$|r_0^2 - r_0'^2| < 0,3 \quad (3.6)$$

Se ha demostrado que todos los criterios anteriores son realmente necesarios para evaluar adecuadamente la capacidad de predicción de un modelo QSAR [30,31]. Además, el poder predictivo externo de los modelos 3D-QSAR desarrollados utilizando el *conjunto de prueba* se examinó considerando las métricas r_m^2 [32] como se muestra a continuación:

$$r_m^2 = r^2 (1 - |\sqrt{r^2 - r_0^2}|) \quad (3.7)$$

Donde r^2 y r_0^2 son los coeficientes de correlación cuadrado entre las actividades observadas y predichas del conjunto de prueba, sin y con la intercepción en el origen,

respectivamente. Para una validación de modelo externa significativa, el valor de r_m^2 debe ser superior a 0,5.

Finalmente, los resultados de CoMFA/CoMSIA fueron representados gráficamente por mapas de contorno de campo usando el tipo de campo StDev*Coeff. En los mapas de contorno, los campos moleculares como los campos estérico, electrostático, hidrofóbico, donante y receptor de puente de hidrogeno; definen las regiones favorables o desfavorables de las moléculas alineadas, sugiriendo la modificación necesaria para diseñar nuevas moléculas, con la finalidad de aumentar la actividad de los inhibidores.

3.5. Simulaciones de Dinámica Molecular

Las estructuras acopladas de 4bxx con los compuestos 11, 16, 35 y 38 sirvieron como estructuras de inicio para cuatro simulaciones MD usando el programa Amber 12 [33]. El campo de fuerza de Amber99SB [34] fue usado para parametrizar la enzima. Los archivos de topología y coordenada de la enzima fueron construidos usando el módulo *Leap* y los parámetros de campo de fuerza necesarios para los inhibidores se generaron con el módulo de *Antechamber* basado en el campo de fuerza AMBER generalizado (GAFF) [35,36]. Las distribuciones de las cargas *RESP* para los inhibidores se generaron a partir de la optimización geométrica usando el programa Gaussian 03 [37]. La minimización de cada complejo fue efectuada a través del módulo *sander*. Cada complejo fue disuelto en una caja rectangular de agua usando el modelo TIP3P [38], a una distancia mínima de 2 Å entre el soluto y la caja. El catión divalente de zinc en el modelo fue reemplazado por un catión divalente de zinc en forma de tetraedro que tiene cuatro átomos catiónicos ficticios que rodean al ion de zinc central [39,40]. Los iones de sodio fueron agregados al sistema mediante el reemplazo aleatorio de moléculas de agua para neutralizar el sistema.

La minimización de energía de todo el sistema se llevó a cabo en 500 pasos con el método de descenso más pronunciado, seguido de 500 pasos de minimización vía gradiente conjugado, y un calentamiento de 50 picosegundos (ps) en un equilibrado corto (NVT) a temperatura constante de 300 K usando Termostato Berendsen y 500 ps de presión constante de 1 atm (equilibrado NPT) con restricciones débiles en el complejo. Para mantener los enlaces de hidrógenos restringidos, se utilizó el algoritmo SHAKE [41]. Se realizó una serie de producción por 50 nanosegundos (ns) utilizando el conjunto NPT a 300 K y 10 atm de presión con un paso de tiempo de 2 femtosegundos (fs) y dinámica de Langevin para el control de temperatura [42]. Se registraron coordenadas de la trayectoria cada 2 picosegundos para todas las ejecuciones de MD. Durante las simulaciones de MD, las interacciones de Coulomb de Largo Alcance fueron manejadas usando el método *Particle Mesh Ewald* (PME) [43] y una distancia de corte de 8,0 Å fue usada para las interacciones de Van der Waals.

3.5.1. Análisis de las trayectorias

Los módulos ‘*ptraj*’ and ‘*cpptraj*’ incorporados en AmberTools 12 [44] fueron usados para analizar la desviación de la raíz-cuadrada-media (RMSD) entre pares de estructuras ligando-receptor, en función de la posición media de los átomos de carbono α . Los módulos también se usaron en la medición de distancias entre posiciones atómicas de los ligandos y los residuos del receptor y un análisis de puentes de hidrógenos. Las imágenes y representaciones estructurales fueron preparadas usando PyMOL, VMD [45], y UCSF Chimera [46].

3.5.2. Estimación de la afinidad de unión (MMPBSA)

Las energías de interacción para los complejos proteína-ligando fueron estimadas usando el método MMPBSA y el enfoque de trayectoria única [47-51]. El método no toma en consideración a las moléculas de agua y iones de forma explícita, modelando sus efectos desde la aproximación de un continuo. Los cálculos de energía promedio se realizaron seleccionando 500 instantáneas a intervalos de tiempo definidos a partir de las simulaciones de 50 ns de duración, generando los conjuntos conformacionales para el complejo, receptor y ligando, necesarios para los cálculos energéticos. Los análisis se realizaron utilizando el script MMPBSA.py [52] que genera los archivos para correr los cálculos. La fuerza iónica fue ajustada a 0.1M (istrng=0.1) e idecomp=1. En el método MMPBSA, la energía libre para cada especie (complejo, receptor, ligando) es descompuesta en una energía en fase gas (“entalpía”), una energía libre de solvatación y un término de entropía como se muestra en la ecuación 3.8.

$$G = H - TS = E_{gas} + G_{solv} - TS \quad (3.8)$$

La energía libre G para cada especie se calcula a partir de la ecuación 3.9:

$$G = E_{gas} + G_{solv} - TS = E_{MM} + G_{solv} - TS \quad (3.9)$$

Donde E_{MM} es la energía mecánica molecular en la fase gaseosa, G_{solv} la energía libre de solvatación, y $-TS$, la contribución de la entropía conformacional.

El término E_{MM} consiste en las energías interna E_{int} , electrostática E_{elec} y van der Waals E_{vdW} , de acuerdo con la siguiente ecuación 3.10,

$$E_{MM} = E_{int} + E_{elec} + E_{vdW} = (E_{bond} + E_{angle} + E_{tors}) + E_{elec} + E_{vdW} \quad (3.10)$$

La energía de solvatación, G_{solv} , se compone de una contribución electrostática (polar), $G_{solv,polar}$, y una contribución no polar, $G_{solv,no-polar}$, según ecuación 3.11:

$$G_{solv} = G_{solv,polar} + G_{solv,no-polar} \quad (3.11)$$

Aquí, $G_{solv,pol}$ se estima a partir de una solución de la ecuación lineal Poisson-Boltzmann (PB) y la parte no polar, $G_{solv,no-polar}$ es típicamente estimada por el método del área superficial accesible al solvente (SASA), aunque también podría ser el volumen (SAV).

El término de solvatación no polar ($G_{\text{solv,no-polar}}$) se evalúa a partir de una simple relación lineal según ecuación 3.12,

$$G_{\text{solv,no-polar}} = \gamma \cdot \text{SASA} + \text{offset} \quad (3.12)$$

La parte no polar, $G_{\text{solv,no-polar}}$ es típicamente estimada por el método del área superficial accesible al solvente (SASA), aunque también podría ser el volumen (SAV).

Los modelos modernos de solvatación no polar, usan términos separados para modelar las contribuciones repulsivas (“cavidad”) y atractivas (“dispersión”) de van der Waals, con la finalidad de superar las limitaciones reportadas del enfoque clásico. $G_{\text{solv,no-polar}}$ consiste en dos componentes de interacciones muy diferentes, la energía repulsiva de G_{rep} y la energía atractiva G_{att} (ec. 3.13), las que deben ser modeladas separadamente [55,56],

$$G_{\text{solv,no-polar}} = G_{\text{att}} + G_{\text{rep}} \quad (3.13)$$

Utilizamos el modelo moderno de solvatación no polar, que utiliza la superficie accesible por solvente (SASA) para correlacionar el término repulsivo (hidrofóbico/cavidad), y aplicamos un enfoque de integración de superficie para calcular el término atractivo (dispersión) como [56] (3.14),

$$G_{\text{solv,no-polar}} = G_{\text{disp}} + \gamma \cdot \text{SASA} + \text{offset} \quad (3.14)$$

La entropía de vibración puede ser evaluada, por ejemplo, vía análisis de modos normales. Sin embargo, en trabajos recientes, se ha convertido en una práctica común excluir los términos de entropía (TS) de los cálculos MMPBSA [53]. Pero, esta consideración es “aceptable” sólo cuando se calculan las energías de enlazamiento relativas para comparar ligandos similares en el mismo receptor [54].

Así, la energía de interacción del complejo proteína - ligando puede evaluarse a partir de la siguiente ecuación:

$$\Delta G_{\text{int}} = G_{\text{Complejo}} - G_{\text{Receptor}} - G_{\text{Ligando}} \quad (3.15)$$

Donde G_{complejo} , G_{receptor} , and G_{ligando} denotan la energía libre del complejo, receptor y ligando, respectivamente.

La energía de interacción final es entonces el promedio del valor de ΔG_{int} sobre todos los cuadros seleccionados desde la fase de producción de la trayectoria MD.

Para entender la interacción inhibidor-residuo con más detalle, se utilizó el método de la descomposición de energía por residuo para el cálculo de ΔG_{bind} , [57]. Bajo este enfoque, las energías libres de enlace para cada residuo que interactúa con los inhibidores, se descompone en los respectivos términos de contribución ΔE_{vdW} , ΔE_{elec} , $\Delta G_{\text{solv,polar}}$ y $\Delta G_{\text{solv,no-polar}}$ [58-60].

CAPÍTULO 4 - RESULTADOS Y DISCUSIONES

4.1. Acoplamiento molecular (Docking)

El acoplamiento molecular entrega las conformaciones alineadas de entrada para todos los compuestos superpuestos, basándose en el mejor ΔG de unión para cada una de las posturas acopladas. Sin embargo, no siempre la mejor energía representó una conformación inicial estable para el posterior estudio de dinámica molecular y, que al mismo tiempo, estuviera en concordancia con el patrón de unión de moléculas del *conjunto de datos*; para esas moléculas el número de corridas (*Runs*) tuvo que extenderse del análisis propuesto inicialmente de 100 a 250, para encontrar así una posición de acoplamiento de acuerdo con el patrón de unión del resto del conjunto de datos y que tuviera una energía de unión aceptable, a pesar de no ser la más negativa. Este paso, en el que tuvimos que discernir sobre esta acción, resultó ser fundamental, porque la alineación por *Docking* resultó entonces ser estadísticamente más fiable en estudios 3D-QSAR posteriores. Las poses de acoplamiento de los compuestos en estudio (Figura 4.1A) mostraron una concordancia con la forma en que los inhibidores conocidos acceden y se unen en el sitio catalítico de la enzima, donde el átomo de Zn juega un papel fundamental, como se observa estructuralmente en complejos HDAC determinados experimentalmente vía difracción de rayos-x [9,12,13]. Posteriormente, para corroborar si el mapa de interacciones predicho a través de los cálculos de acoplamiento se mantiene en un entorno solvatado y flexible, se realizaron simulaciones MD de 50 ns de duración para complejos entre HDAC1 y cuatro moléculas del conjunto de datos. Estos resultados se representan en la Figura 4.1.

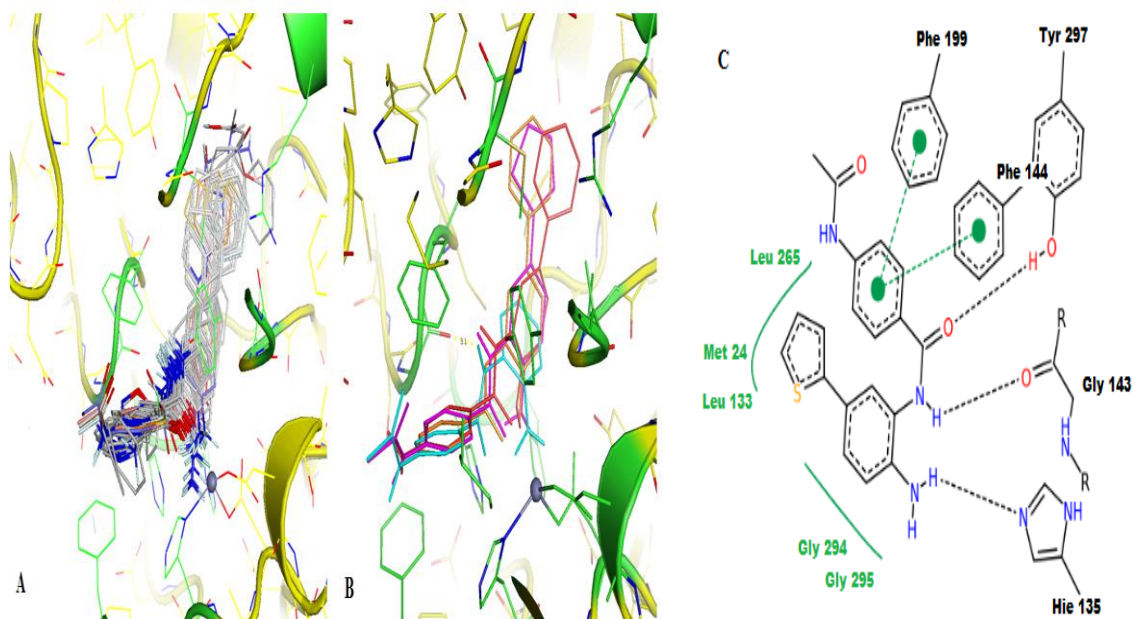


Figura 4.1. (A) Alineación de los compuestos anclados en el sitio activo de HDAC1. (B) Enfoque en cuatro moléculas del conjunto de datos seleccionadas para simulaciones de dinámica molecular, 11 roja, 16 naranja, 35 celeste y 38 magenta. (C) Representación de interacciones enzima-ligando en dos dimensiones para el compuesto 38 obtenidas desde el acoplamiento molecular. Las líneas discontinuas negras indican interacciones de hidrógeno, la línea sólida verde muestra interacciones hidrofóbicas y las líneas punteadas verdes muestran interacciones de apilamiento π - π .

4.2. Resultados de los estudios CoMFA y CoMSIA

Los estudios de CoMFA y CoMSIA se llevaron a cabo utilizando dos esquemas geométricos, basada en ligando (LB) y guiada por el receptor (RG) a través de combinaciones de campo sucesivas con diferentes campos descriptores de forma independiente y en combinación. La alineación LB dio resultados estadísticamente deficientes para los modelos CoMFA y CoMSIA en término del número óptimo de componentes N, sin embargo, fue la primera aproximación para examinar diferentes conjuntos de entrenamiento y prueba. Esta búsqueda permitió encontrar un conjunto de entrenamiento y prueba que luego se usó en la alineación RG, que produjo un modelo significativo para CoMFA considerando las contribuciones estérica y electrostática.

Observando la tabla 4.1, el modelo CoMFA basado en el alineamiento RG, obtenido mediante la combinación de descriptores de campo estérico y electrostático, satisface con valores de $q^2 = 0,560$ y $N = 4$. Este modelo fue validado externamente usando un conjunto de prueba de 12 moléculas, obteniéndose un valor para el coeficiente de correlación r^2_{test} de 0,8059 para el modelo CoMFA-SE. Además, el modelo al forzarse a pasar por el origen es muy parecido a la recta $y = x$ como se definió en la sección 2.12, cumpliendo con todas las condiciones para su validez.

Tabla 4.1. Parámetros estadísticos para el modelo CoMFA-SE, esquema RG. Validación externa.

Modelo	q^2	N	SEP	SEE	r^2_{ncv}	F	Contribución por Campo	
							S	E
CoMFA-SE	0,560	4	0,616	0,210	0,949	102,644	0,567	0,433

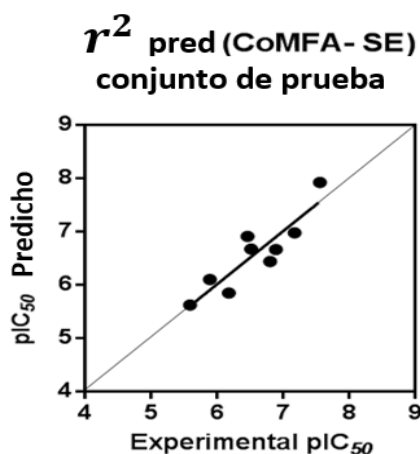


Figura 4.2. Validación externa del conjunto de prueba para CoMFA.

Tabla 4.2. Resumen de los parámetros de validación externa para CoMFA.

Cond.	Parámetro	Valor Umbral	CoMFA
1	q^2	>0,5	0,560
2	r^2	>0,6	0,806
3a	r_0^2	Cercano al valor r^2	0,805
3b	$r_0'^2$	Cercano al valor r^2	0,763
4a	k	$0,85 < k < 1,15$	1,002
4b	k'	$0,85 < k' < 1,15$	0,995
5a	$(r^2 - r_0^2)/r^2$	<0,1	0,001
5b	$(r^2 - r_0'^2)/r^2$	<0,1	0,053
6	$ r_0^2 - r_0'^2 $	<0,3	0,042
7	r_m^2	>0,5	0,782

Los valores de pIC_{50} experimentales, predichos y residuales, para el conjunto de datos completo de este modelo CoMFA-SE (RG) se da en la **Tabla A.4.3** del Apéndice A.3.6.

Los resultados mostraron una relación bastante coherente entre la información obtenida a través de los mapas de contorno, el tamaño medio de las moléculas más activas, tamaño del sitio activo y la relación con los residuos pertenecientes al sitio catalítico del receptor. Ver Figura 4.3.

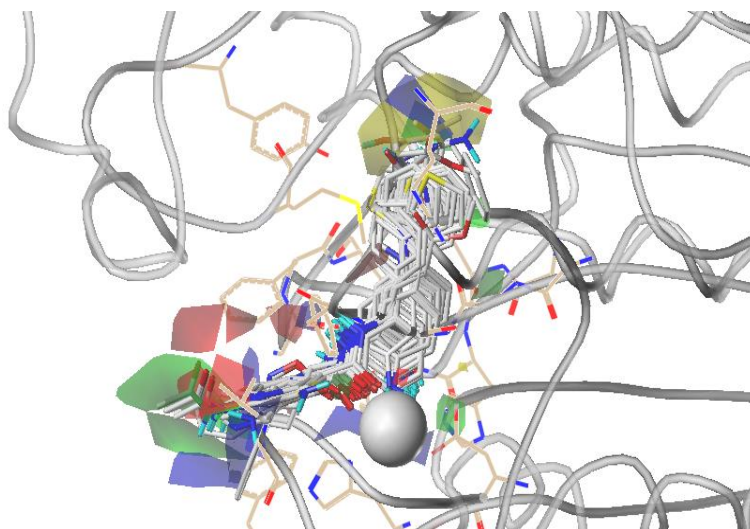


Figura 4.3. A lineamiento RG y mapas de contorno CoMFA-SE al interior del sitio activo de HDAC1.

Se observó que el alineamiento obtenido desde la superposición de las mejores poses de *Docking* si bien muestra un patrón no es muy pulcro, es por eso que se tomó la información suministrada por ambos enfoques geométricos (LB & RG) y se complementó con la entregada por el cristal obtenido por difracción de rayos-X de la proteína HDAC2 (93.22%

de identidad, sitio activo 100% de conservación) en complejo al compuesto 38 de la serie, código PDB: 4LY1 (ver apéndice A 3.5.). Entonces se hizo un nuevo alineamiento mejorado tomando como base la información obtenida hasta ese momento. De este último alineamiento se generaron y validaron los modelos CoMFA y CoMSIA finales de este trabajo. Las consideraciones de este último alineamiento se detallan a continuación.

4.3. Alineamiento asistido y mejorado por cristalografía de rayos-x

En la Figura 4.4A y 4.4B, se observan las alineaciones por los esquemas descritos en las secciones anteriores LB y RG, respectivamente, mostrando similitudes con excepción de dos zonas, la de los sustituyentes aromáticos a la izquierda de las moléculas, encerrados en círculo rojo y diferencias en la zona encerrada en el cuadrado verde.

Esto se debe a que en el esquema basado en el ligando (que muestra una mejor superposición), se considera i) moléculas planas y ii) mínimos energéticos para cada molécula (Figura. 4.4A). Por otro lado, el esquema guiado por el receptor tiene las mejores poses de acoplamiento de cada molécula superpuestas (Figura. 4.4B) tomando en consideración conflictos estéricos, basándose en i) la información de la región del sitio activo circundante y ii) una energía libre de unión (ΔG) obtenida usando la suite AutoDock.

Esta información es coherente (aunque con ciertas diferencias) a un patrón de unión como muestra el cristal del complejo entre HDAC2 y 38 (Fig. 4.4C); Entonces se integra esta información y con eso se construye el alineamiento asistido por el cristal (Fig. 4.4D), buscando mejorar los modelos 3D-QSAR en un nuevo intento.

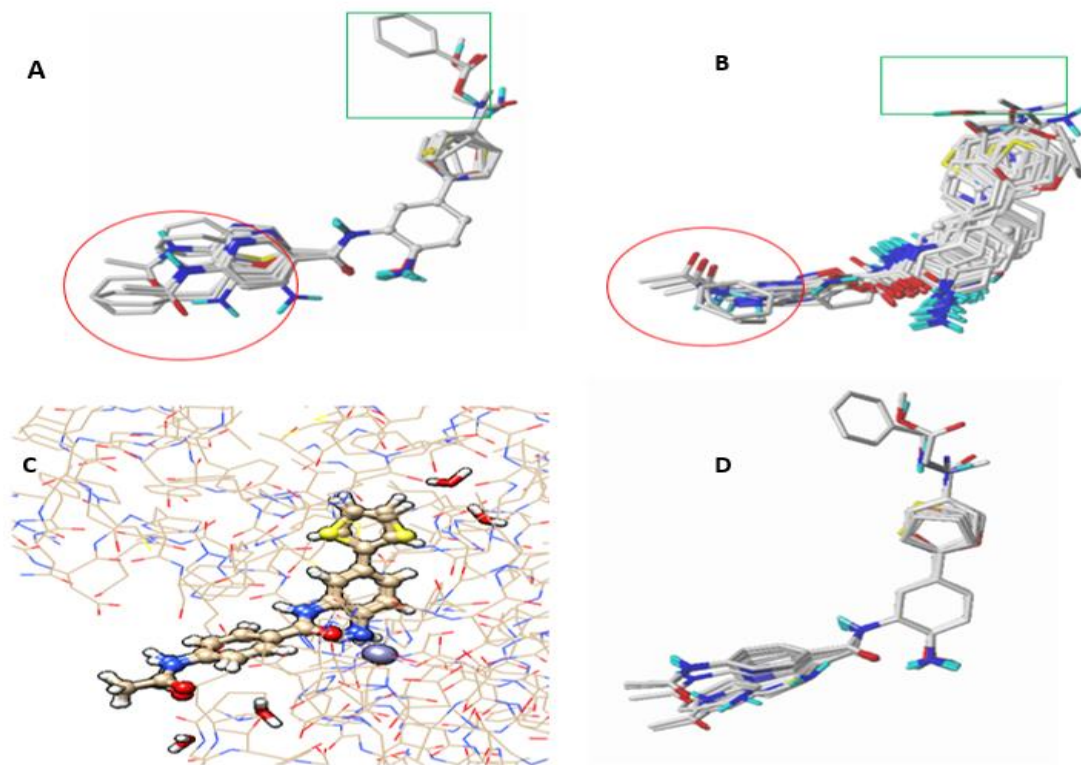


Figura 4.4. (A) Alineación LB; (B) Alineación RG, (C) Cristal 4LY1.pdb y (D) Alineación asistida.

4.4. Resultados CoMFA y CoMSIA asistido por difracción de rayos -X

En la tabla 4.3, se presenta un resumen de los resultados estadísticos para los modelos CoMFA y CoMSIA derivados desde el alineamiento asistido por difracción de rayos-X (Fig. 4.4D). Esta búsqueda se hizo usando diferentes campos descriptores de forma independiente y en combinación. El primer parámetro para evaluar la solidez estadística de un modelo QSAR es el valor de q^2 , que debe ser superior a 0,5. En este contexto el modelo CoMFA-SE obtenido mediante la combinación de descriptores de campo estérico y electrostático mostró un valor óptimo de $q^2 = 0,586$. Por otro lado, el modelo CoMSIA-SEHA obtenido al combinar descriptores estéricos, electrostáticos, hidrofóbicos y de aceptor de puente de hidrógeno mostró un valor óptimo de $q^2 = 0,570$.

Tabla 4.4. Búsqueda secuencial para la generación de los mejores modelos 3d-qsar^a. En negrita los mejores modelos obtenidos mediante el alineamiento asistido por difracción de rayos-x.

Modelo	q^2	N	SEP	SEE	r_{ncv}^2	F	Contribución por Campo					
							S	E	H	D	A	
CoMFA-S	0,645	4	0,542	0,237	0,932	75,204	1					
CoMFA-SE	0,586	4	0,585	0,18	0,961	135,523	0,639	0,361				
CoMSIA-S	0,626	6	0,584	0,296	0,904	31,243	1					
CoMSIA-SE	0,635	11	0,666	0,046	0,998	773,431	0,493	0,507				
CoMSIA-SEH	0,513	4	0,635	0,225	0,939	84,671	0,306	0,319	0,376			
CoMSIA-SEHD	0,505	8	0,708	0,046	0,998	1,083,803	0,266	0,254	0,277	0,202		
CoMSIA-SEHA	0,570	5	0,61	0,18	0,963	108,171	0,214	0,264	0,297			0,226
CoMSIA-SEA	0,501	5	0,658	0,197	0,955	89,439	0,304	0,387				0,309
CoMSIA-SH	0,569	5	0,611	0,227	0,94	66,317	0,436		0,565			
CoMSIA-SA	0,508	10	0,748	0,145	0,981	84,400	0,684					0,316
CoMSIA-SDA	0,503	18	1,063	0,024	1	1,706,125	0,554				0,211	0,235
CoMSIA-EH	0,532	6	0,653	0,156	0,973	121,274		0,498	0,502			
CoMSIA-EHA	0,538	5	0,633	0,216	0,946	73,573		0,328	0,355			0,317

^a q^2 es el coeficiente cuadrado de la validación cruzada del LOO (CV), N es el número óptimo de componentes, SEP es el error estándar de predicción, SEE es el error estándar de estimación del análisis no-CV, r^2 es el coeficiente cuadrado del no-CV, F es el valor de la F-test y S, E, H, D y A son las contribuciones estéricas, electrostáticas, hidrofóbicas, dador y aceptor de enlace hidrógeno, respectivamente.

No obstante, para CoMSIA varios modelos con otras combinaciones de campo presentaron un valor significativo con valores de q^2 mayores al criterio 0,5 y solo por medio del valor de r^2_{test} (de aquí en más simplemente r^2), que evalúa la capacidad de predicción externa del modelo, se permitió la discriminación entre todos los modelos. El número óptimo de componentes (N) también es bajo en los modelos presentados (N = 4 para el CoMFA-SE y N = 5 el modelo CoMSIA-SEHA, en adelante se denominan modelos "CoMFA" y "CoMSIA"). Idealmente, un buen modelo debe tener el menor número posible de componentes (N~10% del conjunto de datos), lo que asegura que las predicciones se basarán en información significativa de las contribuciones de campo, en lugar de un sobre-entrenamiento del modelo.

También hay un balance en las contribuciones de campo (ver Tabla 4.3 valores para S, E, H, D y A), lo que apoya la confiabilidad de las conclusiones obtenidas de cada mapa de contorno.

La Tabla 4.4 presenta un resumen de la validación externa de los modelos CoMFA y CoMSIA. Ambos modelos tienen un valor alto para r^2 de 0,7994 y 0,785 respectivamente, lo que es una indicación de una capacidad predictiva externa adecuada (Ver Figura 4.5).

Tabla 4.5. Resumen de los parámetros de validación externa para CoMFA y CoMSIA.

Condición	Parámetro	Valor Umbral	CoMFA	CoMSIA
1	q^2	>0.5	0,586	0,570
2	r^2	>0.6	0,799	0,785
3a	r_0^2	Cercano al valor r^2	0,798	0,782
3b	$r'_0{}^2$	Cercano al valor r^2	0,723	0,685
4a	k	$0.85 < k < 1.15$	1,001	0,987
4b	k'	$0.85 < k' < 1.15$	0,991	1,010
5a	$(r^2 - r_0^2)/r^2$	<0.1	0,00213	0,00395
5b	$(r^2 - r'_0{}^2)/r^2$	<0.1	0,0958	0,128*
6	$ r_0^2 - r'_0{}^2 $	<0.3	0,0749	0,0971
7	r_m^2	>0.5	0,766	0,741

Los valores de pIC_{50} experimentales, predichos y residuales, tanto para el conjunto de datos completo de los mejores modelos CoMFA y CoMSIA se dan en la **Tabla A.4.6** del Apéndice A.3.6.

Como se mencionó en la sección anterior y según lo expuesto por Golbraikh y Tropsha^[30,31], los valores altos de q^2 y r^2 (condiciones 1 y 2) son necesarias, pero no suficientes para la validación de un modelo. Para que un modelo QSAR tenga una capacidad predictiva fiable, la línea para la actividad experimental v/s actividad predicha debe estar lo más cerca posible de la línea $y = x$. Esto se observa en el cumplimiento de las condiciones [3a o 3b], [4a o 4b], [5a o 5b], y 6 enumeradas en la Tabla 4.4. Finalmente, la condición 7, conocida como métrica r_m^2 , es una medida cuantitativa para determinar la proximidad entre la actividad observada y la predicha para el conjunto de prueba. Los modelos CoMFA y CoMSIA aquí reportados cumplieron todas las condiciones para la validación interna y externa (ver Figura 4.5 y 4.6).

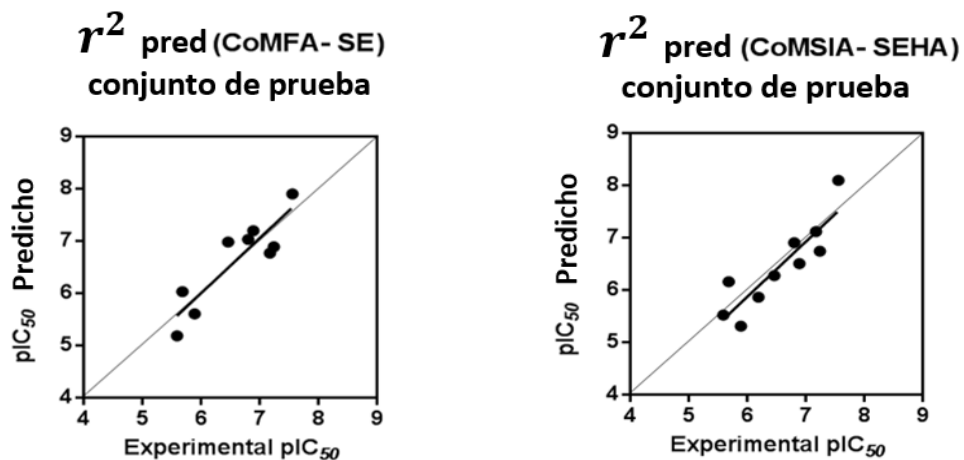


Figura 4.5. Validación externa de los conjuntos de prueba para CoMFA (Izquierda) y CoMSIA (derecha)

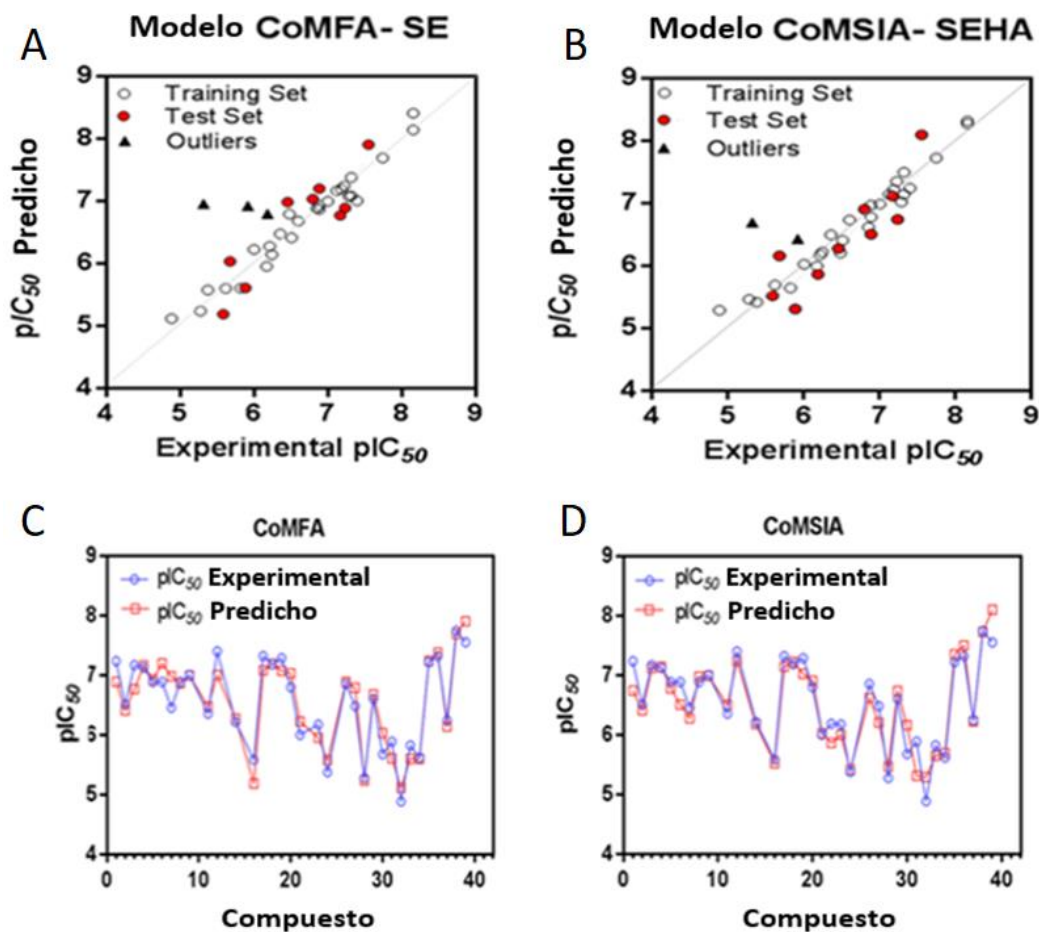


Figura 4.6. Gráficas de los valores pIC_{50} experimentales frente a los predichos para el conjunto de entrenamiento y prueba de moléculas para los modelos CoMFA-SE (A) y CoMSIA-SEHA (B). Predicciones para cada molécula en el conjunto de datos en modelos CoMFA (C) y CoMSIA (B) (sin valores atípicos).

4.5. Mapas de Contorno CoMFA y CoMSIA

4.5.1. CoMFA -SE

En el mapa de contorno estérico de la Figura 4.7A y 4.7B, las regiones verde y amarilla describieron el espacio alrededor de las moléculas donde la introducción de grupos voluminosos podría aumentar o disminuir la actividad respectivamente. Existen dos contornos verdes que son significativos, el primero de ellos está en la zona de los sustituyentes R₂ (Tabla 3.1 y Figura 3.2), orientados hacia la cavidad interna del sitio activo, y que sugiere que un sustituyente aromático (5 o 6 miembros) es una muy buena opción. De hecho, aquellos sustituyentes que carecen de un anillo en esta posición tienen actividades informadas menores al promedio como las moléculas 15, 32 y 35. También queda claro que sobrepasar esta zona verde implica encontrarse con 3 poliedros amarillos, que indican que sustituyentes voluminosos en estas zonas (más allá del poliedro verde) no sería favorable como opción, lo que explicaría las bajas actividades de los compuestos 25 – 31 de la serie. El otro poliedro verde importante es el que está en la zona más expuesta al solvente, donde la sustitución de algún grupo voluminoso sería favorable, de hecho, el menos activo no tiene sustituyentes envueltos por este segundo poliedro, como sí lo está en parte la benzamida terminal de los compuestos más activos de la serie 36 – 39, que eventualmente podría extenderse algo más.

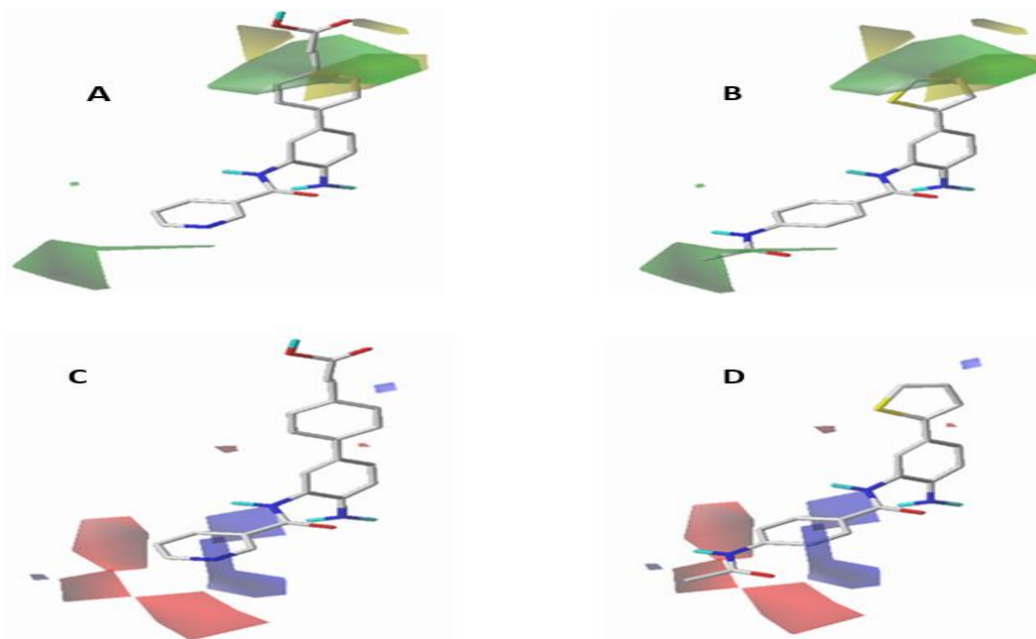


Figura 4.7. (Izquierdas) Mapas de contorno CoMFA-SE, A (estérico) y C (electroestático); para el compuesto menos activo 30. (Derecha) Mapas de contorno CoMFA-SE, B (estérico) y D (electroestático); para compuesto más activo 38.

En los mapas de contorno electrostático del CoMFA de la Figura 4.7C y 4.7D, los contornos azules representan las regiones en las que la carga electropositiva aumentaría la actividad y los contornos rojos representaban las áreas en las que la carga electronegativa aumentaría la actividad.

En esta serie de compuestos, los poliedros rojos están asociados con los átomos electronegativos Nitrógeno y Oxígeno (NH, C=O), pertenecientes a la amida más expuesta al solvente para el compuesto más activo **38**, algo que no ocurre para el menos activo **30**.

Por otra parte, podemos ver un poliedro azul alrededor de la zona de la amida central, para todos los compuestos de la serie, relacionado a la orientación de la densidad de carga (δ^+) de hidrógenos pertenecientes a los grupos NH de la amida central y NH₂ del anillo aminofenil. Otro poliedro a considerar está en las cercanías de la posición 3 del anillo piridínico del compuesto menos activo **30**, lo que explicaría su baja actividad, pues tiene un átomo de Nitrógeno en esta posición. Esto sugiere que una forma de mejorar la actividad de este tipo de derivados sería aumentar la deficiencia electrónica de sustituyentes del anillo en esta posición. Esto podría lograrse insertando grupos de atracción electrónicos en el anillo o insertando grupos como OH, NH o NH₂ que orienten la densidad de carga (δ^+) del átomo de hidrógeno hacia la región azul.

4.5.2. CoMSIA-SEHA

(SE) Estérico y Electroestático. El mapa de contorno electrostático CoMSIA (Figura 4.8A y 4.8B) aporta información diferente, pero complementaria a la que aporta el mapa de contorno electrostático del modelo CoMFA (Figura 4.8C y 4.8D). Un poliedro azul por sobre el anillo en la posición de sustituyentes R o R₂ (Tabla 3.1 y Figura 3.2; dependiendo la serie de compuestos), sugiere sustituyentes que sean deficientes en electrones para los compuestos que presentan un anillo en esta posición. Esto debe ponerse en contraste con la información obtenida desde el mapa CoMSIA estérico, puesto que este poliedro azul se intersecta con el poliedro amarillo de la contribución estérica en gran parte de esta zona, por lo que sustituyentes más bien pequeños con una densidad de carga electropositiva estarían bien en esta zona. El otro poliedro azul está en una zona similar a lo observado en CoMFA, (posición 3 anillo piridina), discutido anteriormente.

Por otra parte, un poliedro rojo estaría íntimamente relacionado a la sustitución de un átomo electronegativo, como el azufre en la posición 2 del tiofeno, lo cual aporta un antecedente adicional para explicar la actividad de **38**.

Al igual que el mapa de contorno estérico CoMFA (Figura 4.7A y 4.7B), el mapa de contorno estérico CoMSIA (Figura 4.8C y 4.8D) muestra una región verde que intersecta e incluso envuelve al anillo en la posición R o R₂ según sea la serie de compuestos (ver Tabla 3.1 y Figura 3.2), lo cual está en concordancia con la mayoría de los compuestos que tienen actividad por sobre el promedio. Aquellos compuestos en que R o R₂ es un átomo de hidrógeno, tienen actividades más bajas.

El otro poliedro verde, al igual que en CoMFA estérico, se orienta a la zona más expuesta al solvente, lo que está relacionado con los compuestos de mayor actividad, como se mencionó anteriormente.

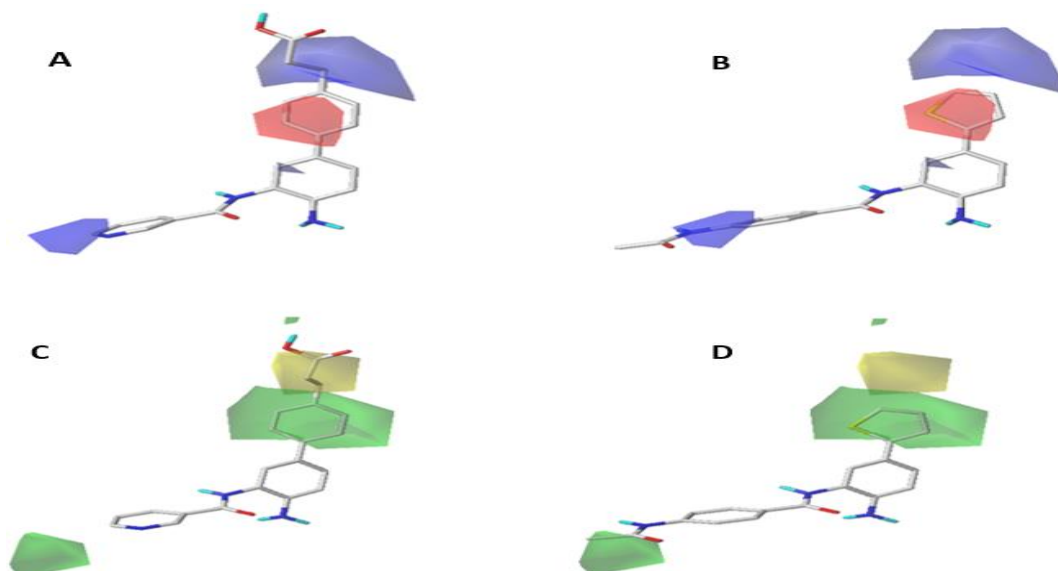


Figura 4.8. (Izquierdas) Mapas de contorno CoMSIA-SE, A (electroestático) y C (estérico); para el compuesto menos activo **30**. (Derecha) Mapas de contorno CoMSIA-SE, B (electroestático) y D (estérico); para compuesto más activo **38**.

(HA) Hidrofóbico y Aceptor de puente de hidrogeno. En los mapas de contorno del campo hidrofóbico (Figura 4.9A y 4.9B), las regiones naranja y blanca indicaron las áreas donde se favorecen las propiedades lipofílicas e hidrofílicas, respectivamente.

Una región blanca sobre posiciones 1,2-del anillo tiofeno del sustituyente en R_2 del compuesto más activo **38** mostró que la introducción de grupos hidrofílicos (en este caso átomo electronegativo como el azufre), podría ser la responsable de la actividad inhibitoria de aquellos compuestos que cuentan con un anillo de estas características, lo cual puede combinarse con el resultado observado en el CoMSIA electroestático (Figura 4.8B). Otro poliedro blanco aparece en la zona de la benzamida terminal (Figura 4.9B) de los compuestos más activos **36-39**, que es la región más expuesta al solvente, lo que deja de manifiesto que la benzamida en esta posición cumple el requerimiento de grupo hidrófilo, aunque pueden probarse otros grupos de similares características.

El poliedro naranja más importante está también en la zona del sustituyente R_2 , pero del lado contrario de la zona hidrófila, un antecedente a considerar en la elección del tipo de anillo y sus sustituciones en estas posiciones. En efecto, se podría probar agregar sustituyentes hidrofóbicos de un tamaño medio como grupos metilo o etilo en estas posiciones como opciones a considerar según sea el anillo al cual están unidos en R_2 (5 o 6 miembros), o hacer de R_2 un anillo fusionado tipo índol e índol sustituido, relacionando este mapa con los poliedros verdes de los mapas estéricos (Fig. 4.7 A y B – Fig 4.8 C y D).

Un poliedro naranja menor cercano a la posición 3 del anillo piridínico avala que en esta zona se favorezcan grupos fenilo o fenilo sustituidos en posición 3 por un grupo de características hidrofóbicas de un tamaño no muy grande como un metilo.

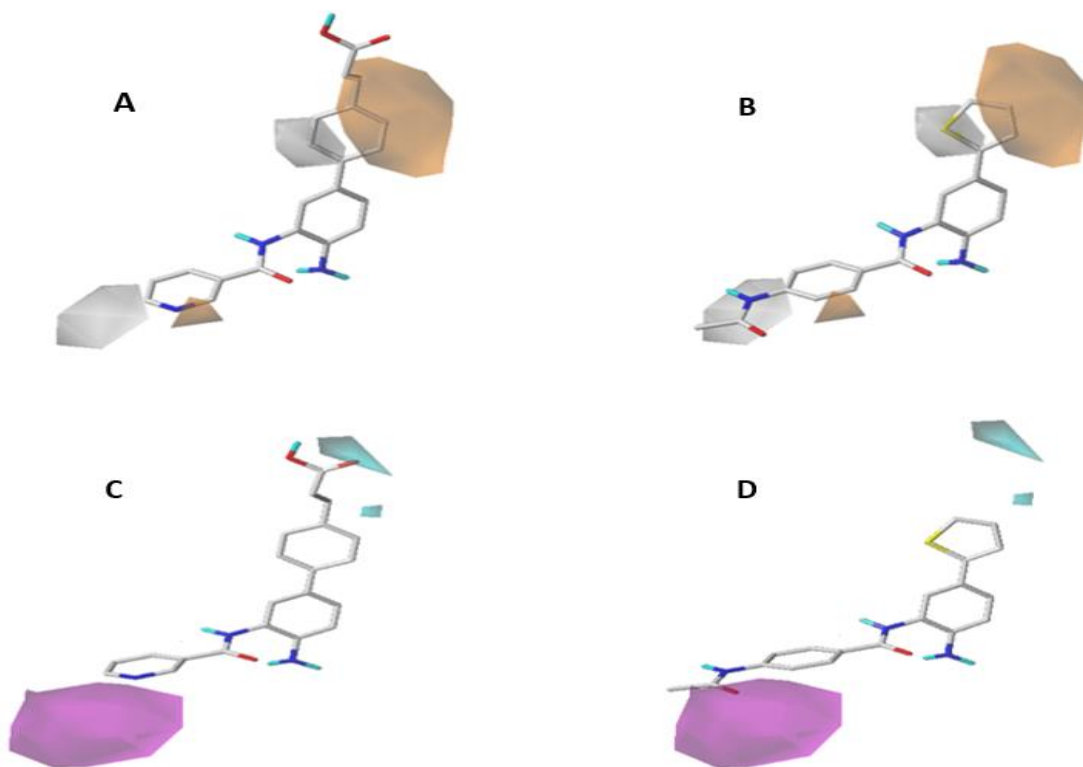


Figura 4.9. (Izquierdas) Mapas de contorno CoMSIA-HA, A (Hidrofóbico) y C (Aceptor interacción de hidrogeno); para el compuesto menos activo 30. (Derecha) Mapas de contorno CoMSIA-SE, B (Hidrofóbico) y D (Aceptor interacción de hidrogeno); para compuesto más activo 38.

En los mapas de contorno de aceptor de puente de hidrógeno (Figura. 4.8 C y D), los contornos de color magenta indican zonas donde aceptores de puente de hidrógeno son favorables y color cián (celeste) donde aceptores de puente de hidrogeno son desfavorables (o donantes serían favorables). Este mapa muestra un poliedro magenta grande cerca de la posición del carbonilo de la amida expuesta al solvente, lo que significa que la incorporación de átomos aceptores de enlace de hidrógeno en estas posiciones es favorable. Entonces, sería interesante probar otras combinaciones que cumplan con este requisito. De hecho, los compuestos más activos de la serie **36-39** posicionan los átomos de oxígeno desde los grupos carbonilo en esta zona.

En la zona superior dos poliedros color cián indicarían que grupos dadores de puente de hidrógeno en estas posiciones sería beneficioso, la diferencia de electronegatividad entre el Carbono y el Hidrogeno podría bastar para formar una interacción de hidrogeno débil, aunque esta inclusión debe ser meticulosa e integrando la información obtenida de los mapas de contorno de las otras contribuciones, el entorno de aminoácidos del sito activo del receptor y la información que nos pueda proveer el estudio de los aspectos dinámicos de estas interacciones, por medio de simulaciones de dinámica molecular (MD).

4.6. Estudios de Dinámica Molecular

4.6.1. Estabilidad de la simulación del Complejo HDAC1-inhibidor

Los resultados del *docking* pueden considerarse como un estado instantáneo, son ampliamente aceptados porque proporcionan un resultado rápido de la interacción entre nuevas moléculas propuestas y su afinidad con el receptor inmóvil, sin embargo, no se consideran concluyentes porque la unión *in vivo* de inhibidores a una enzima es un proceso dinámico. Por esta razón, se realizaron simulaciones de 50 ns de duración para cuatro compuestos: las moléculas 11, 16, 35 y 38 del conjunto de datos, para investigar aspectos dinámicos del modo de unión y probar además la estabilidad de la conformación acoplada. La estabilidad conformacional del complejo enzima-inhibidor, fue medida siguiendo los valores del RMSD de los átomos del *Backbone* de la proteína y los átomos pertenecientes a cada ligando, tomando como referencia la estructura inicial aportada por *docking* para cada complejo (ver Figura 4.10.).

Los resultados muestran que solo tres simulaciones muestran estabilidad relativa a su RMSD. La excepción fue la molécula 11, en la Figura 4.10A, donde se puede observar la existencia de al menos tres estados conformacionales.

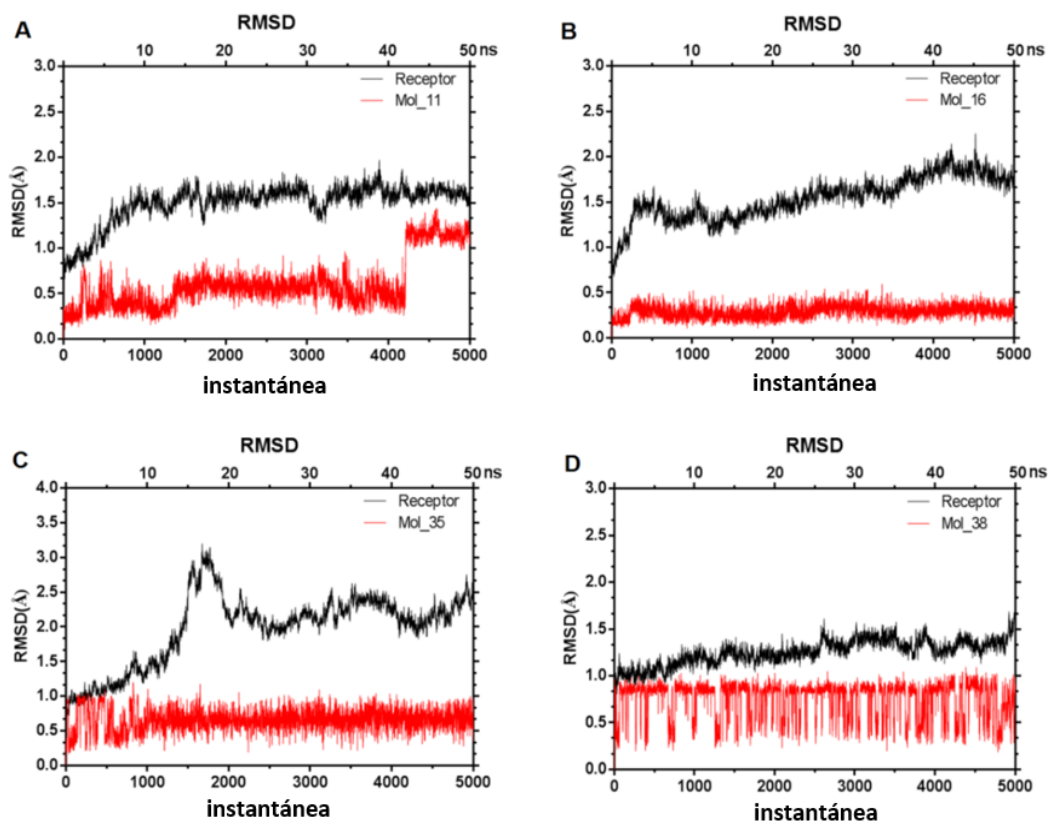


Figura 4.10. RMSDs para átomos de *backbone* del receptor (negro) e inhibidores (rojo) en simulación de dm de 50 ns. (A) complejo HDAC1-11 (B) complejo HDAC1-16 (C) complejo HDAC1-35 y (D) complejo HDAC1-38.

Revisando visualmente la simulación del complejo HDAC1-11, los primeros 15 ns representan el estado ligado, la parte central (15-42 ns) una conformación intermedia donde se ha formado un puente de hidrógeno entre la TYR297 y el carbonilo del ligando; finalmente alrededor de la instantánea 4200 (~42 ns) el inhibidor se aleja del átomo Zn^{+2} (ver Apéndice A.3.1), siendo desplazado por moléculas de agua circundantes. Se observó entonces que la solvatación era desfavorable a la afinidad de unión en el complejo 11 y esto fue confirmado posteriormente por cálculos de MMPBSA. Por esta razón, el espacio de muestreo conformacional efectivo para los cálculos MMPBSA de energía libre para la molécula 11 se encuentra entre las conformaciones 1-1500 (primeros 15 ns de las simulaciones), teniendo cuidado de mantener el conjunto de muestras en 500 instantáneas, idéntico número de extracciones al de los otros complejos en estudio. La fluctuación para el ligando en el complejo HDAC1-38 (D), es debido a un enlace de libre rotación que queda expuesto al solvente.

4.6.2. Cálculos MMPBSA y evaluación de energía de interacción proteína-ligando (aproximada)

La afinidad de unión de los compuestos 11, 16, 35 y 38 fue evaluada por el método MMPBSA y la descomposición por residuos de la energía libre de enlazamiento aproximada, para identificar las principales fuerzas que impulsan dicha asociación molecular y examinar las diferencias y similitudes entre los cuatro complejos.

Este análisis muestra que la energía libre de interacción aproximada ΔG proteína-ligando, sin considerar la contribución entrópica, es energéticamente favorable para todos los complejos. La Tabla 4.5 ilustra las contribuciones individuales a la energía libre de enlazamiento de las componentes van der Waals y las interacciones electrostáticas, así como también las contribuciones a la energía libre de solvatación polar y no polar, desde los enfoques con y sin dispersión.

En primer lugar, la energía de solvatación no polar se calculó con una relación lineal a la superficie accesible por solvente (SASA). Segundo, se utilizó un enfoque reciente que divide la energía de solvatación no polar en una contribución repulsiva y otra atractiva (Ver Apéndice A.3.4.1).

Para las cuatro moléculas, los resultados sugirieron que la formación del complejo es favorecida por las interacciones intermoleculares electrostáticas y de van der Waals, así como por el componente no polar repulsivo de la energía libre de solvatación ($\Delta G_{SAS(rep)}$). Por otro lado, que el componente no polar atractivo ($\Delta G_{disp(att)}$) y la energía libre de solvatación polar fueron desfavorables para la unión del ligando al receptor.

Se puede observar que existe una cierta tendencia en los valores de cada contribución, sin embargo, es evidente que la molécula **35** rompe con la tendencia, ya que carece de un grupo aromático en la posición R_2 (ver figura 4.11 izquierda), el cual se encuentra presente e incrustado en la cavidad hidrofóbica interna (ver figura 4.11 derecha) para las moléculas 11, 16 y 38. Esta podría ser la razón, por la cual el compuesto **35** muestra un valor para la componente electrostática que esta favorecida en desmedro de la componente de van der Waals.

Tabla 4.7. Valores de energía de unión y energía de los componentes individuales calculados con el método MM-PBSA para HDAC1 en complejo con las moléculas 11, 16, 35 y 38 (kcal/mol).

Inhibidor	ΔE_{vdw}	ΔE_{elec}	ΔG_{PB}	$\Delta G_{SAS(rep)}$	$\Delta G_{disp(att)}$	ΔG_{bind}		pIC ₅₀
						no-disp ^a	disp ^b	
Molécula 11	-37,2(0,19)	-76,9(0,62)	69,6(0,25)	-30,6(0,04)	49,6(0,05)	-75,1	-25,5(0,47)	7,4
Molécula 16	-32,3(0,24)	-88,8(0,39)	67,8(0,20)	-29,9(0,06)	47,5(0,07)	-83,2	-35,7(0,28)	7,32
Molécula 35	-19,4(0,23)	-98,0(0,50)	68,1(0,38)	-23,4(0,06)	36,4(0,08)	-72,7	-36,3(0,25)	6,24
Molécula 38	-35,5(0,21)	-85,5(0,39)	62,8(0,30)	-31,1(0,04)	50,1(0,05)	-89,4	-39,3(0,23)	8,16

NOTA: Todas las energías se promedian en 500 instantáneas a intervalos de tiempo definidos a partir de las simulaciones MD de 50 ns de duración, y se expresan en kcal/mol. El error estándar de la media se indica entre paréntesis. Ver Apéndice A.3.4.1

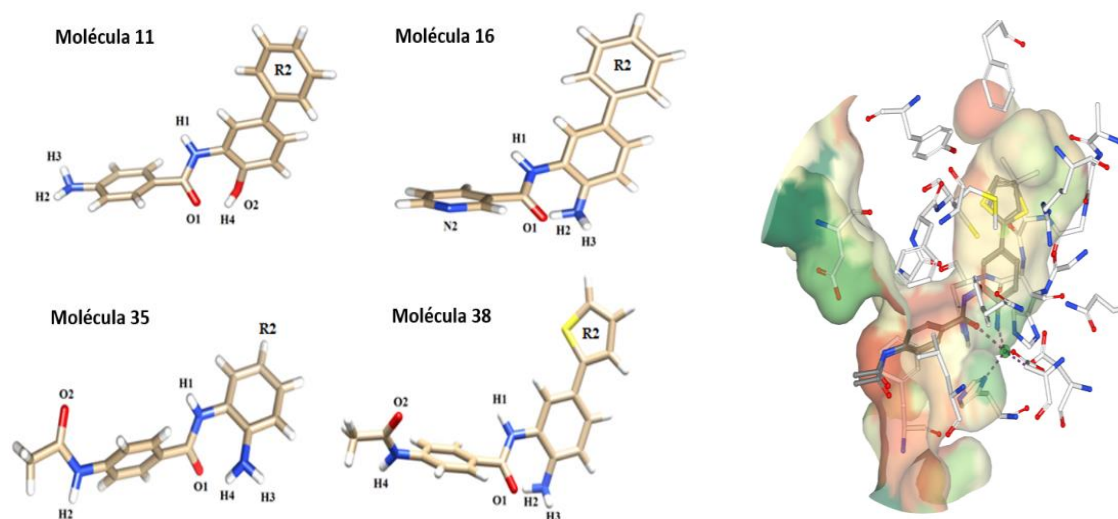


Figura 4.11. (Izquierda) Instantáneas de cada inhibidor utilizado para la exploración del sitio de unión HDAC1, en MD. (Derecha) Cavity hidrofóbica de la HDAC2 en complejo a 38.

Además, es consistente con el hecho de mostrar la contribución para $\Delta G_{SAS(rep)}$ menos negativa entre todos los compuestos, teniendo en cuenta que este término está relacionado con la cavidad o la superficie accesible al solvente. No obstante, esto último se compensa un poco en el componente de la contribución atractiva de la energía de interacción de solvatación no polar $\Delta G_{disp(att)}$. La energía libre de solvatación polar es desfavorable para todos los complejos.

El análisis de las dinámicas mostró que todos los complejos están estabilizados por una combinación y balance fino de interacciones polares y no polares con átomos de residuos pertenecientes al sitio activo de HDAC1 y moléculas de agua, mientras que el núcleo central de benzamida está estabilizado por el átomo de Zn⁺² tetra-coordinado con distancias interatómicas $d \leq 4 \text{ \AA}$ (ver Apéndice A.3.1 – A.3.4). Basándose en estos valores de energía de

enlazamiento, la afinidad estimada de los inhibidores estudiados con el sitio activo HDAC1 puede describirse mediante el siguiente orden:

MOL 38 > MOL 35 > MOL 16 > MOL 11; para el método con dispersión $\Delta G_{\text{bind,disp}}$

MOL 38 > MOL 16 > MOL 11 > MOL 35; para el método sin dispersión $\Delta G_{\text{bind,no-disp}}$.

En ambos casos la tendencia de afinidad no está totalmente de acuerdo con los datos de pIC_{50} *in vitro*, en el que el orden observado es, MOL 38 > MOL 11 > MOL 16 > MOL 35.

Entre los dos enfoques para estimar la contribución $\Delta G_{\text{solv,npol}}$, la aproximación sin dispersión fue capaz de predecir el orden para el compuesto con mayor y menor afinidad, lo que resultó en línea con los datos experimentales, mostrando diferencias en el orden central, MOL 16 > MOL 11 *v/s in vitro*; MOL 11 > MOL 16.

4.6.3. Contribuciones Por-Residuo al ΔG de interacción proteína-ligando (aproximada)

La descomposición de la energía de enlazamiento o interacción por residuos muestra que los complejos entre HDAC1 y las moléculas 11, 16, 35 y 38 se estabilizan con un gran número de residuos similares. Este estudio permitió identificar los residuos que están presentes en todos los complejos: MET 24, LEU 133, HIE 135, PHE144, CYS 145, PHE 199, LEU 265, GLY 295 y el átomo Zinc tratado como residuo ZN 371.

Observamos que las interacciones electrostáticas definen la máxima eficacia de la unión de ligando y están en línea con los resultados de los cálculos de MMPBSA (Tablas 4.5 y 4.6). Estos resultados demuestran que la contribución electrostática domina el total de ΔG_{bind} en las metaloenzimas dependientes Zn.

Tabla 4.8. Contribución de energía por residuo a la unión de las moléculas 11, 16, 35 y 38; en negrita residuos comunes seleccionados del sitio activo del receptor HDAC1 (kcal/mol).

Residuo	mol_11	mol_16	mol_35	mol_38
TYR 18		-0,02	-0,04	
MET 24	-1,01	-0,54	-0,16	-1,51

Programa de Magíster en Ciencias mención Química

ARG 28	-0,15		-1,02	-0,6
PHE 103	-0,1	-0,03		-0,08
LEU 133	-0,45	-1,68	-0,53	-1,21
HIE 134	-0,36	-2,61	-0,18	
HIE 135	-2,1	-0,39	-1,3	-1,11
GLY 143		-0,3		
PHE 144	-1,16	-1,69	-2,05	-1,69
CYS 145	-0,69	-0,18	-0,2	-0,04
ASP 170*			-1,89	
HIN 172*				-1,36
PHE 199	-0,46	-0,23	-0,4	-0,26
ASP 258*				
LEU 265	-0,85	-0,51	-1,01	-0,77
GLY 293	-0,23	-0,07	-0,24	
GLY 294		-0,62	-1	-0,44
GLY 295	-0,83	-0,74	-0,62	-1,33
TYR 297				-1,78
ZN 371	-12,98	-13,59	-13,56	-12,38
ΔG_{total}	-21,38	-23,16	-24,16	-24,58

Las Figuras 4.12 y 4.13 muestran gráficas comparativas de la contribución a la energía libre aproximada de enlazamiento (izquierda) aportada por cada residuo del sitio activo en interacción no covalente con los compuestos 11, 16, 35 y 38 (Derecha). Se han seleccionado instantáneas representativas para cada complejo donde se ilustra el tipo de interacción no covalente que está involucrada entre aminoácidos y el ligando.

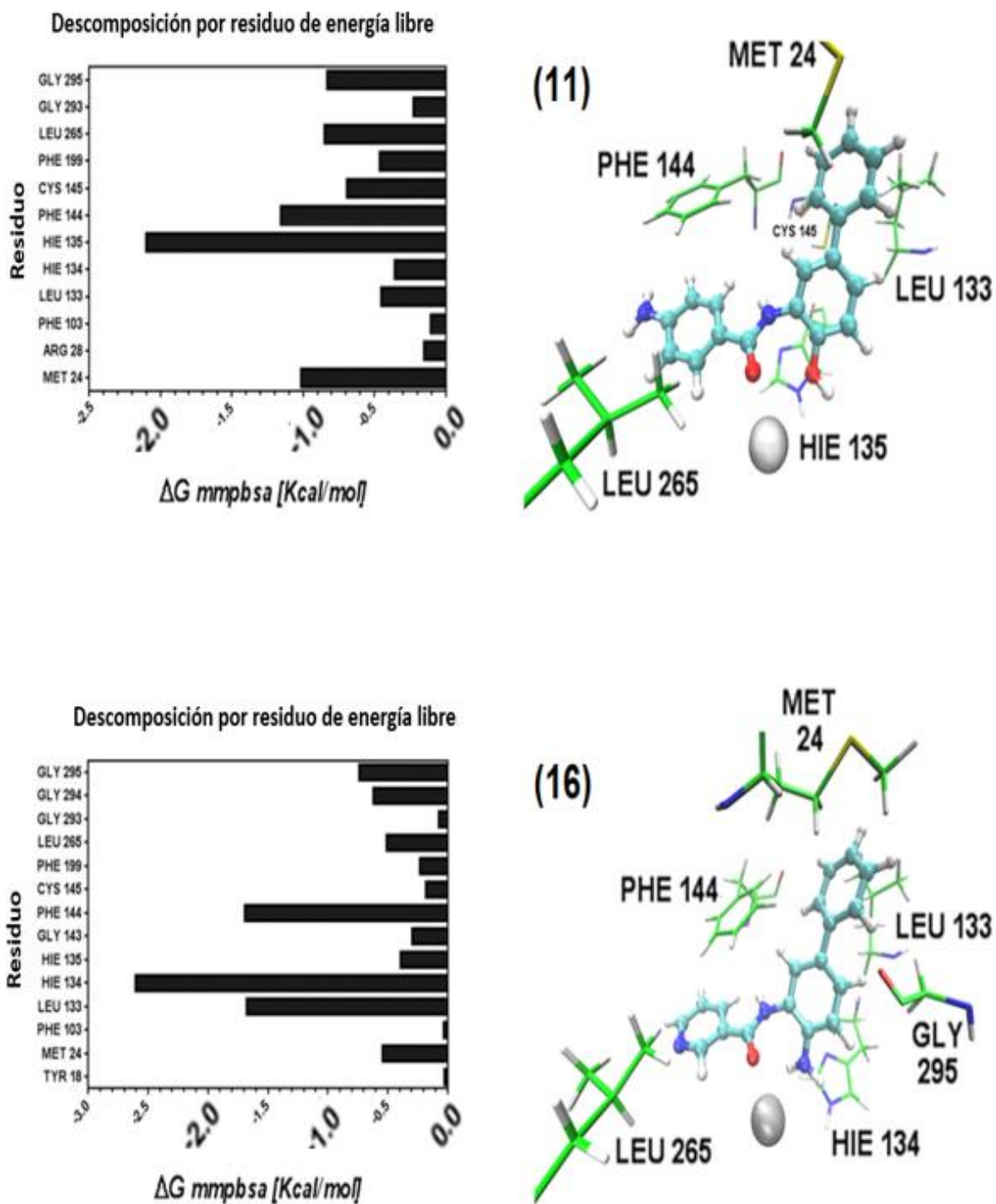


Figura 4.12. Mapa de interacciones y descomposición de energía libre por residuos de los complejos HDAC1-11 (arriba) y HDAC1-16 (abajo), respectivamente. Los inhibidores se muestran como modelos CPK de color azul cian y los residuos se muestran como representación del palo verde, el átomo de zinc como representación de las perlas metálicas.

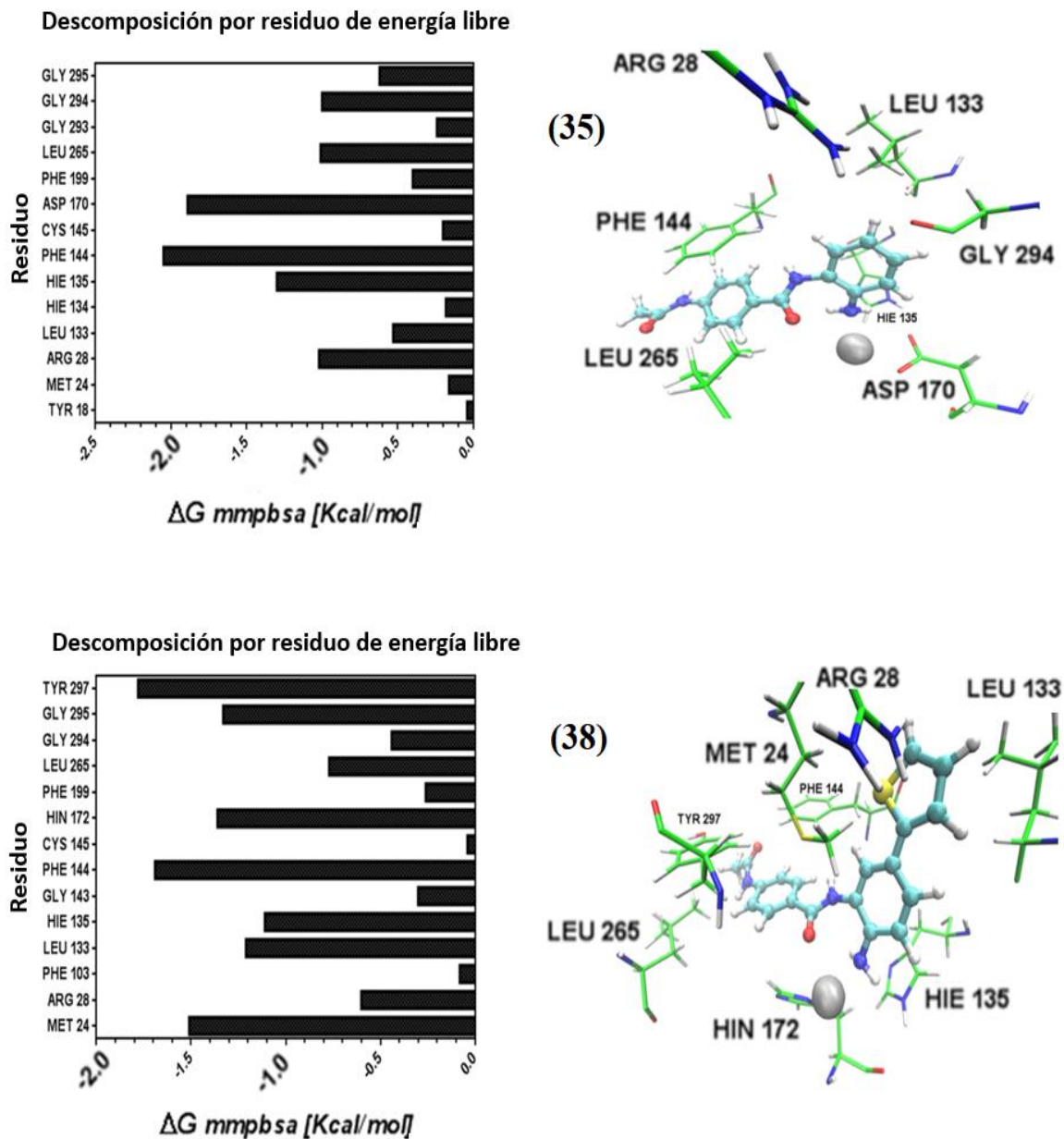


Figura 4.13. Mapa de interacciones y descomposición de energía libre de residuos de HDAC1-35 (arriba) y HDAC1-38(abajo). Los inhibidores se muestran como modelos CPK de color azul cian y los residuos se muestran como representación del palo verde, el átomo de zinc como representación de las perlas metálicas.

Un análisis más detallado de las interacciones presentes en cada complejo se da en la siguiente sección, en la cual se analizan las más relevantes y de mayor presencia en cada simulación, con énfasis en los residuos en común que están presentes y aportan en la formación de cada uno de los cuatro complejos estudiados.

4.6.4. Análisis sistemático de interacciones no-covalentes específicas.

El análisis de las simulaciones MD permite tener una mejor apreciación de las interacciones intermoleculares entre HDAC1 y cuatro inhibidores del conjunto de datos, esta información puede contribuir a mejorar el diseño basado en la estructura y la optimización de inhibidores de HDAC1. Las interacciones no covalentes y distancias características de diferentes contactos observados en la simulación se muestran en la Tabla A.4.9 de la sección de Apéndice A.3.6 para los complejos 11, 16, 35 y 38.

Entre las interacciones que se observan con más frecuencia se encuentran las conocidas y ampliamente utilizadas en el diseño de ligandos, como los contactos *hidrofóbicos* y *apilamiento* π (π -stacking). Estos son seguidos por *interacciones de hidrógeno* débiles y *apilamiento de amida*. Los patrones de enlaces de hidrógeno clásicos tienen un tratamiento especial en la siguiente sección.

Las interacciones hidrofóbicas son la principal fuerza motriz en las interacciones entre los receptores de drogas. No es sorprendente que la Leucina, seguida de la Metionina (ambas contienen grupos metilo) y las cadenas laterales de Glicina sean las que con mayor frecuencia participan en interacciones hidrofóbicas. Aunque la *Metionina* está clasificada como residuo hidrofóbico, estudios recientes muestran que la interacción aromática C-H \cdots S (MET) produce una energía de estabilización adicional de entre -1 a -1,5 kcal/mol en comparación con una interacción puramente hidrofóbica^[61].

La presencia de la interacción débil de carbono aromático C-H \cdots S (MET) prevalece en las estructuras metaloproteicas de HDACs conocidas, en efecto desde las simulaciones puede confirmarse que estabiliza las interacciones de unión ligado-receptor^[62].

Las interacciones que involucran anillos aromáticos son omnipresentes en los sistemas químicos y biológicos y pueden considerarse un caso especial de interacciones hidrofóbicas (Ver A.2.2). Todas las interacciones de apilamiento se observan entre los aminoácidos pertenecientes a la enzima; fenilalanina Phe144/ tirosina Tyr297/ histidina Hie135 y anillos aromáticos en el ligando.

Las interacciones que involucran interacciones de hidrógeno débiles fueron frecuentes debido a que los ligandos contienen hidrógenos provenientes de carbonos aromáticos como buenos dadores, y la proteína buenos aceptores para estas interacciones, principalmente oxígenos/nitrógenos de la cadena lateral de la Glicina y la Metionina. Además, la interacción de enlaces de hidrógeno débiles se combina con el apilamiento de amidas, que sucede cuando los anillos aromáticos del ligando interactúan con la amida de la proteína y el anillo aromático de la proteína interactúa con las amidas de los ligandos. La Glicina fue el residuo de aminoácido que con mayor frecuencia participó en el apilamiento de amida cara a cara, este resultado es un antecedente a tener en cuenta cuando se incluyen ligandos que contienen amidas, como es el caso del compuesto **38**, que tiene una amida terminal y otra central (ver figura 4.14).

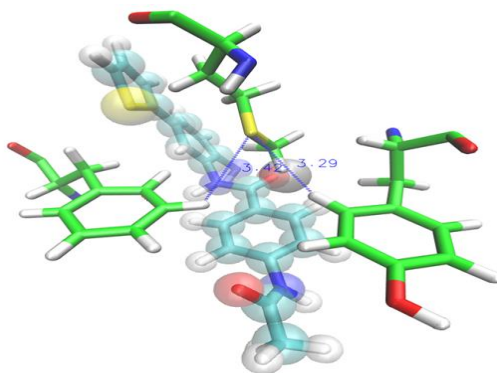


Figura 4.14. Metionina interactuando con PHE y TYR. El modo de unión hidrógeno aromático - azufre MET proporciona una estabilización adicional sobre interacciones puramente hidrofóbicas y a distancias más largas.

4.6.5. Puente de Hidrogeno

La contribución de un enlace de hidrógeno depende del entorno local: un enlace de hidrógeno expuesto a solvente contribuye significativamente menos a la energía de interacción neta que el mismo enlace de hidrógeno en una bolsa hidrofóbica incrustada en la cavidad del receptor. Para obtener una mejor comprensión de la naturaleza dinámica del enlace de hidrógeno, en la figura 4-13 se genera una representación esquemática de los patrones de enlace de hidrógeno. Por otra parte, para reflejar la frecuencia y el tipo de enlace de hidrógeno que ocurre se crearon tablas donde se organiza esta información. A modo de resumen se muestran sólo los resultados para la molécula **38**, la más activa, para los otros ligandos ver Tabla A.4.10 el Apéndice A 3.6.

4.6.6. Interacción de enlace de hidrógeno entre los residuos del sitio activo de HDAC1 y el compuesto 38

En la Tabla 4.7, se muestran los resultados de los patrones de enlace de hidrógeno para mostrar la frecuencia y el tipo de enlace de hidrógeno que ocurre durante el período de simulación. Asp170, Hie135 y Gly143 mostraron una mayor ocupación del enlace de hidrógeno.

Tabla 4.11. Ocupación Interacción de hidrógeno entre los residuos Asp170, Hie135 y Gly143 y 38.

DONOR :res@atom	ACCEPTOR:res@atom	%ocupación	d [Å]	α [°]
:Asp170@OD2	:LIG38@H2	38,64	2,817 (0,11)	150,78 (11,01)
:Gly143@O	:LIG38@H1	35,10	2,897 (0,07)	36,15 (8,84)
:Asp170@OD1	:LIG38@H2	11,50	2,939 (0,05)	133,62 (5,77)
:Hie135@ND1	:LIG38@H3	5,66	2,915 (0,06)	114,17 (11,43)

*O Desviación estándar /res: Numero del residuo. atom: ID del átomo en el archivo PDB . d : distancia y α : ángulo promedios.

También se extrajeron instantáneas representativas desde la trayectoria para el monitoreo de parámetros geométricos y el análisis del comportamiento de los enlaces de hidrógeno al usar el modelo de tres centros, una instantánea representativa se presenta en la Fig. 4.15.

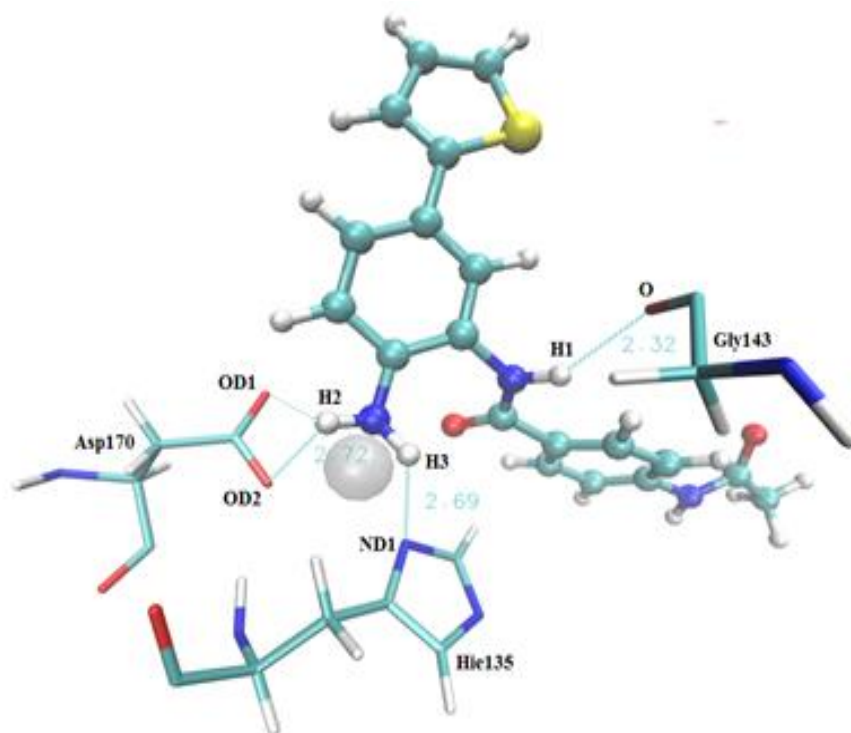


Figura 4.15. La conformación de unión e interacciones de enlace de hidrógeno del compuesto 38 en el sitio activo de HDAC1. Los inhibidores se muestran como modelos CPK de color azul cian y los residuos como representación en barra; el átomo de zinc como representación en perla metálica transparente; los enlaces de hidrógeno se representan en líneas punteadas de color cian y sus distancias están marcadas en angstrom.

4.6.7. Análisis de solvatación del sitio activo del complejo 38

Las moléculas de agua juegan un papel crucial en la unión proteína-ligando. De hecho, la unión de ligando a proteína implica la ruptura de enlaces de hidrógeno entre aminoácidos con moléculas de agua y la formación de nuevos enlaces de hidrógeno entre la proteína y el ligando.

La ocupación del enlace de hidrógeno (HB%) durante la simulación da una estimación aproximada de la ocupación y organización de la primera capa de solvatación alrededor de los átomos de soluto según un criterio. En este análisis, se consideró el HB% para moléculas de agua incluidas en un radio de corte de 3,0 Å a los aminoácidos del sitio activo y el ligando, con una presencia superior al 10%. Estos valores sugieren una participación soluto-solvente de los átomos hidrófilos de soluto, los cuales pueden ser átomos pertenecientes al ligando o aminoácidos. Además, valores superiores al 100% indicarían la formación de más de un enlace de hidrógeno a través de la simulación alrededor de un átomo de soluto, sugiriendo la presencia de más de una molécula de agua circundante.

Un % HB significativo en el rango de ~0-100 % representa una molécula de agua parcialmente presente (ver Tabla 4.8). Una tabla completa incluyendo el complejo entre HDAC1 y las moléculas 11,16 y 35 para el estudio se muestra en la Tabla A.4.6 del Apéndice 3.6.

Tabla 4.12. Ocupación de enlace de hidrógeno entre átomos del solvente e hidrófilos de la molécula 38. Ocupación de enlace de hidrógeno entre disolvente y residuos Hin172, Tyr297 y Hie135,

Donor:res@atom	Aceptor:res@atom	%ocupa	d [Å]	α [°]
:LIG@O2	Solventacceptor	192,60	2,797 (0,12)	73,10 (48,80)
Solventdonor	:LIG@H3	67,24	2,858 (0,08)	119,68 (8,26)
solventdonor	:LIG@H4	15,28	2,909 (0,06)	34,32 (20,26)
:Hin172@NE2	Solventacceptor	342,00	2,834 (0,10)	68,83 (50,51)
:Tyr297@OH	Solventacceptor	188,96	2,828 (0,11)	120,28 (36,02)
:Hie135@ND1	Solventacceptor	94,64	2,882 (0,08)	72,00 (47,70)
Solventdonor	:Tyr297@HH	94,48	2,828 (0,11)	46,89 (46,63)

*O Desviación estándar /res: Numero del residuo. atom: ID del átomo en el archivo PDB. d : distancia y α : ángulo promedios.

Los resultados sugieren que la contribución dinámica de las moléculas de agua involucradas en interacciones de hidrogeno que exhiben un HB % sobre 90 %, tienen como componente principal el movimiento vibracional/rotacional, pues una vez formado la interacción de hidrógeno se pierde en gran medida la traslación y se mantienen esa posición durante casi toda la simulación (~95 %). En los residuos Hin172, Tyr297 y Hie135, usando VMD, se puede ver como las moléculas de agua permanecen unidas al sitio hidrófilo con sólo la posibilidad de rotaciones limitadas que eventualmente rompen la interacción de hidrogeno con el aminoácido y lo reforman inmediatamente después (ver Tabla 4.8).

El caso más representativo es la solvatación del residuo Hin172, en el que la energía libre de solvatación polar fue favorable para la unión del ligando, residuo con un % HB fue de 342 % y que visualmente está rodeado por al menos 3 moléculas de solvente (ver Figura. 4.16) a lo largo de gran parte de la simulación.

La importancia de la solvatación polar, tal como se ha revelado aquí, pone de manifiesto que la inspección estructural para identificar las contribuciones clave a las afinidades de unión debe combinarse/complementarse con los cálculos de contribución y descomposición por residuo de la energía libre de interacción (ver Figuras 4.12 y 4.13) y la solvatación polar obtenida por métodos como MMPBSA u otros (ver Figura 4.17).

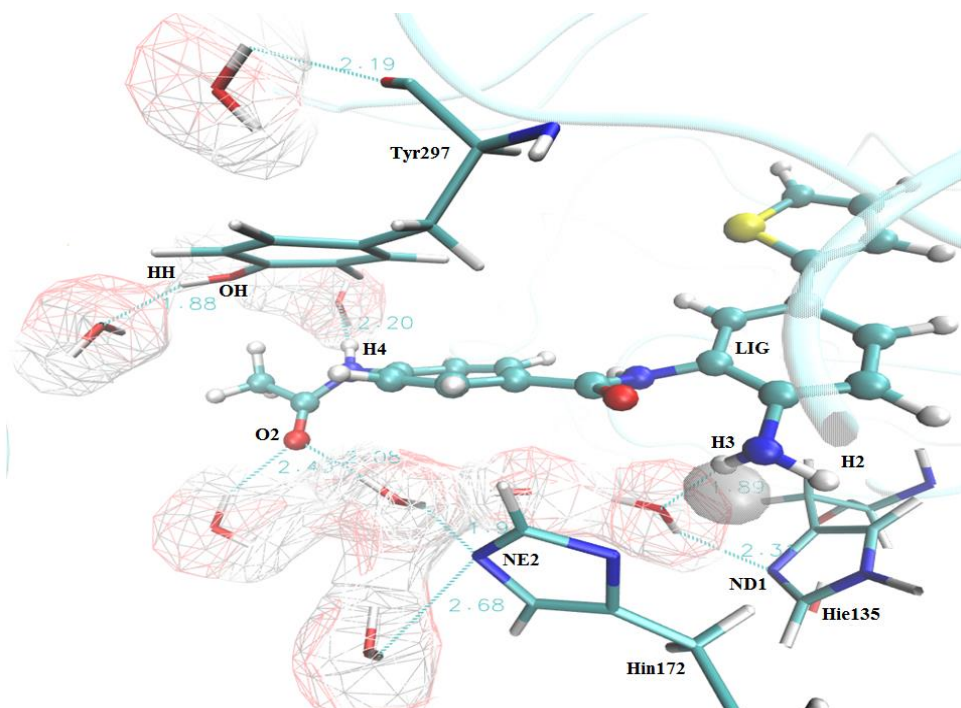


Figura 4.16. La conformación de unión y las interacciones de unión de hidrógeno del compuesto 38 y los residuos en el sitio activo y las moléculas de agua. Los enlaces de hidrógeno se representan en líneas punteadas de color cian y sus distancias están marcadas en angstrom.

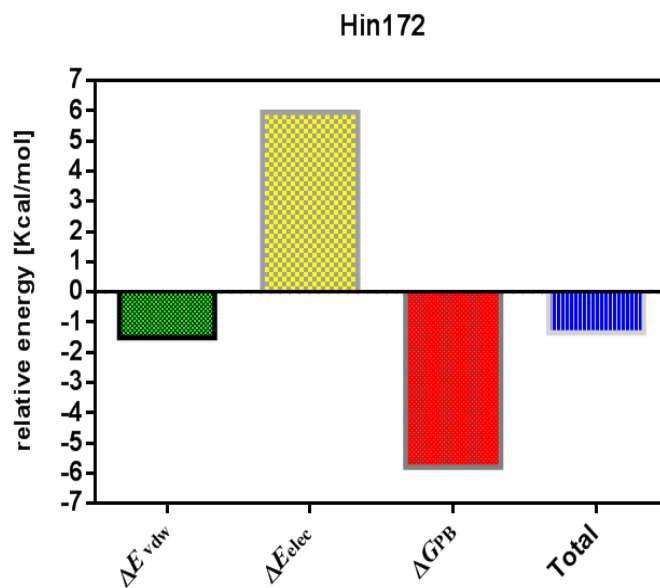


Figura 4.17. Las contribuciones desde las interacciones de Van der Waals ($\Delta E_{vdw} = -1.51$) y electrostáticas ($\Delta E_{elec} = 5.95$), así como la solvatación polar ($\Delta G_{pb} = -5.80$) del residuo hin 172, el total se da en la tabla 3.4. Los errores estándar de las estimaciones de energía libre, así como de los componentes individuales, no son mayores de 0,06 kcal/mol.

CAPÍTULO 5 - CONCLUSIONES

En este estudio, se obtuvo novedosa información acerca del mecanismo de unión de inhibidores de andamiaje de benzamida a la enzima HDAC1, lográndose una comprensión mejorada sobre este sistema ligando-receptor, fundamentada principalmente en información estructural y energética. Se obtuvo información sobre cuáles son los aminoácidos responsables en el mecanismo de unión y por ende inhibición de HDAC1, así como tipos de interacciones no covalente que estarían operando en la formación de los complejos.

Los resultados de *Docking* fueron consistentes con los resultados de cristalografía de difracción rayos-X para inhibidores basados en benzamida en complejo a HDACs, destacando (i) El alineamiento del conjunto de datos dentro del receptor que se obtuvo desde el acoplamiento y posteriormente mejorado por resultados de cristalografía; (ii) La identificación de interacciones no covalentes que serían responsables de la estabilización del complejo ligando-receptor, las cuales fueron posteriormente contrastadas con las observadas desde análisis de cuatro dinámicas moleculares para cuatro complejos.

Los cálculos MMPBSA de energía libre para estos complejos mostraron que la interacción electrostática proporciona la fuerza más importante para la unión de la metaloenzima dependiente de Zn^{+2} y el inhibidor durante el acoplamiento. En la exploración de métodos de punto final para evaluar la energía de enlazamiento ligando-receptor el método MMPBSA sin considerar los términos de dispersión, fue más asertivo relativo a los datos *in vitro*. Las discrepancias observadas entre enfoques (con y sin dispersión) pueden atribuirse al hecho de que la molécula **35** carece de un grupo aromático en la posición R_2 , el cual está presente en las moléculas **11**, **16** y **38**; lo que explicaría porque la molécula **35** tiene un componente electrostático compensado, por medio de la disminución en el componente de Van der Waals y diferencias con la tendencia de la solvatación no polar repulsiva y atractiva. Ambos enfoques sin embargo coinciden en que las principales contribuciones al proceso de unión son las electrostáticas, Van der Waals y solvatación no polar repulsiva (SASA/*cavity*).

Lo más útil del método MMPBSA es el análisis de descomposición de energía por residuos, el cual muestra que los complejos entre HDAC1 y las moléculas 11, 16, 35 y 38 del conjunto de datos se estabilizan con un gran número de residuos en común. Este estudio permitió identificar que residuos están presentes en interacciones para los 4 complejos estudiados y que aportan a la energía libre aproximada de enlazamiento.

Adicionalmente, en el presente trabajo, se utilizó una serie de inhibidores derivados del compuesto CI-994 como inhibidores de HDAC1 para desarrollar modelos 3D-QSAR. Se desarrollaron y validaron modelos comparativos de CoMFA y CoMSIA, utilizando los dos esquemas de alineación principales, basado en ligando y guiado por receptor. La alineación guiada por receptor generó resultados estadísticamente satisfactorios para un modelo CoMFA-SE ($q^2 = 0,560$, $r^2_{\text{test}} = 0,8059$, $r^2_{\text{ncv}} = 0,949$, $N = 4$). Los mapas de contorno CoMFA-SE del conjunto de datos alineado dentro del receptor, nos proporcionó información que permite tener en cuenta el entorno fisicoquímico y posibles conflictos estéricos a la hora de interpretar mapas de contorno y proponer mejoras a fármacos tipo *orto*-aminoanilida.

La alineación guiada por receptor fue asistida por la difracción de rayos-X, para construir nuevos modelos mejorados, de estos seleccionamos dos, los cuales mostraron los siguientes resultados CoMFA-SE ($q^2 = 0.586$, $r^2_{\text{test}} = 0.7994$, $r^2_{\text{ncv}} = 0.961$, $N = 4$) y CoMSIA-SEHA ($q^2 = 0.570$, $r^2_{\text{test}} = 0.7850$, $r^2_{\text{ncv}} = 0.963$, $N = 5$).

Los resultados expuestos en esta tesis en forma de figuras/esquemas/gráficos/tablas provenientes de los modelos 3D-QSAR, el acoplamiento molecular y las simulaciones MD, proporcionaron información que ha sido integrada y utilizada para proponer ideas para el diseño racional de novedosos y más potentes inhibidores contra isoformas específicas de HDACs de la clase I (ver figura 5.1) para la futura síntesis de compuestos antitumorales altamente eficaces basados en el andamiaje de benzamida.

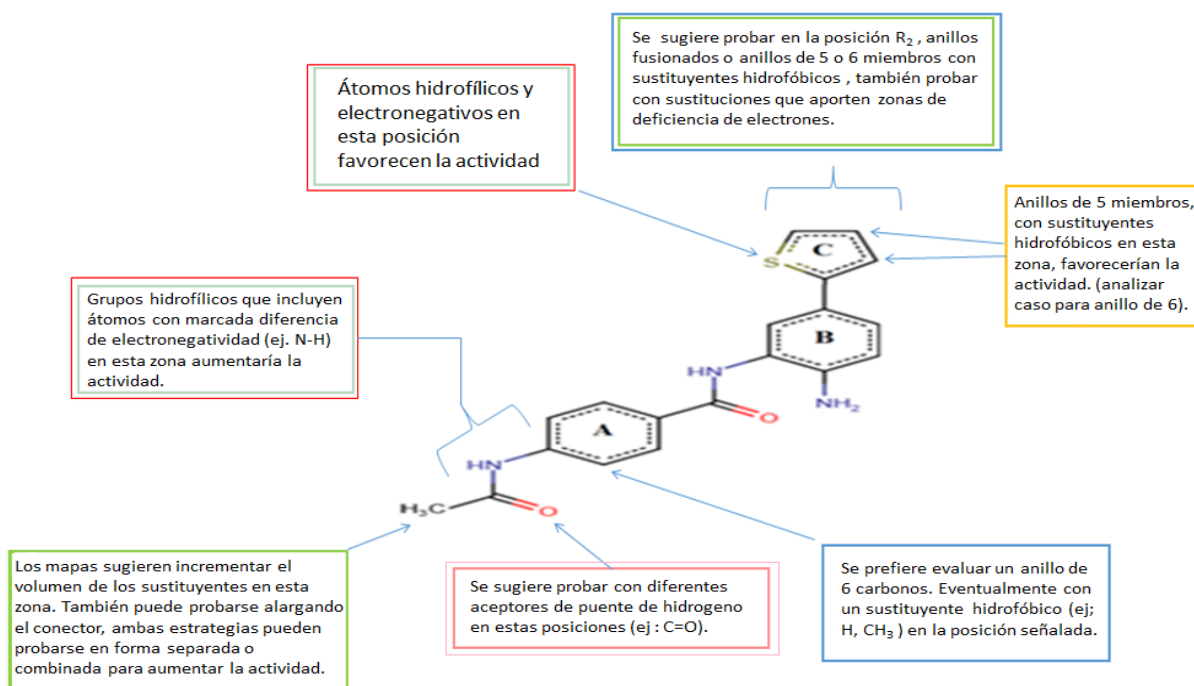


Figura 5.1. Principales relaciones estructura-actividad encontradas en este estudio e ideas para la etapa de conceptualización y diseño *a posteriori* de nuevas moléculas basadas en el andamiaje de benzamida.

APENDICES.

APÉNDICE A1: ESTADO DEL ARTE.

«Me gustaría subrayar con fuerza mi convicción de que la era de los químicos computacionales, en la que no cientos si no miles de químicos trabajarán con ordenadores en lugar de hacerlo en el laboratorio para incrementar muchos aspectos de la información sobre productos químicos, ya está próxima. Sólo hay un obstáculo, a saber, que alguien debe pagar por el tiempo de cálculo».

«...cuanto más exactos son los cálculos, más conceptos tienden a desvanecerse en el aire».

- Robert Sanderson Mulliken

(7 de junio de 1896 - 31 de octubre de 1986)

A.1. MODELADO MOLECULAR.

La química computacional permite a los investigadores experimentar *in silico* mediante simulaciones por ordenador un proceso biológico o químico de interés. Los modelos computarizados de procesos químicos ofrecen mayor resolución a costa de una menor precisión. La química computacional tiene por objeto la determinación teórica de las propiedades de las moléculas a partir de las ecuaciones químico cuánticas o clásicas del movimiento. Si bien los enfoques computacionales *ab initio* tienen la capacidad de modelar y obtener propiedades para moléculas pequeñas comparables a las obtenidas experimentalmente, su aplicabilidad se vuelve muy limitada a medida que aumenta el tamaño del sistema. También las escalas de tiempo que se pueden sondear utilizando dinámicas moleculares *ab initio* son pequeñas, aunque es un campo en constante crecimiento. En consecuencia, para macromoléculas, la aplicación de enfoques mecánico cuánticos (QM) está severamente limitada. Por este motivo, uno necesita recurrir a métodos basados en la mecánica clásica para estudiar fenómenos con escalas de longitud y tiempo más grandes (ver figura 1.1). En efecto, cuando se combina la Mecánica Molecular (MM) con Dinámica Molecular (MD), se logra simular el problema de N-cuerpos para los N-átomos de un sistema como el ligando-receptor, calculando el movimiento de los átomos según la física newtoniana y las descripciones empíricas de interacciones atómicas, obteniendo una visión atomística de la interacción entre un fármaco y un receptor biológico [63 - 66].

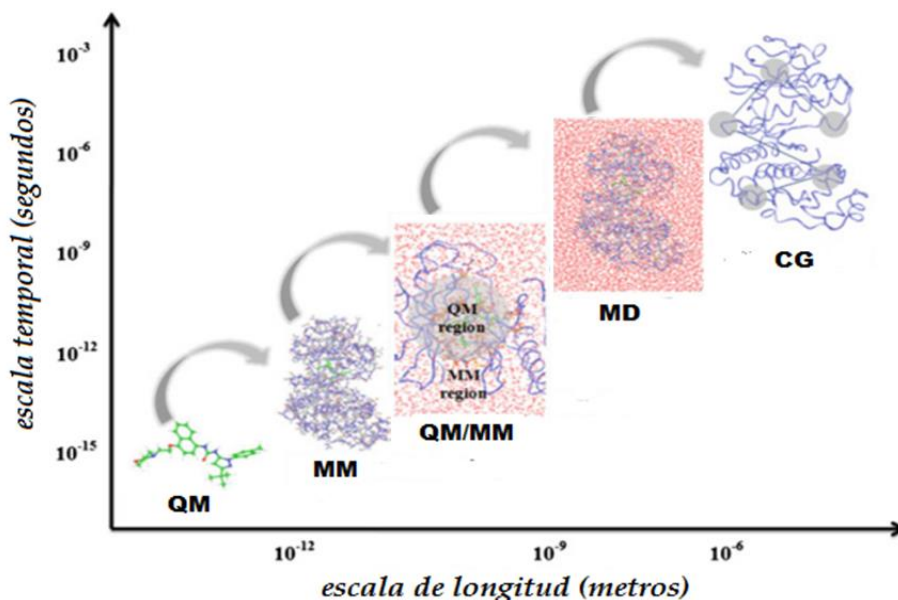


Figura A.1.1 Orden jerárquico de aproximaciones de modelado molecular a diferentes escalas de tiempo y longitud. QM: mecánica cuántica, MM: mecánica molecular, QM/MM: híbridos, MD: dinámica molecular, CG: grano grueso [67].

En este apartado, se presentan varios métodos computacionales que pueden aplicarse al estudio de sistemas moleculares como es el sistema Fármaco-Receptor, abordable desde diferentes escalas de longitud y tiempo de manera complementaria. Además, discutimos los principios básicos de los enfoques basados en ligando y guiados por receptor, para el diseño de fármacos asistidos por computadora.

A.1.1. Estructura electrónica

En el área de las moléculas pequeñas, los cálculos de estructura electrónica han jugado un papel muy importante y han producido una gran cantidad de datos de alta confiabilidad obtenidos al emplear herramientas computacionales fundamentadas en los principios de la Mecánica Cuántica (QM). Según ésta, toda la información posible sobre un sistema molecular puede obtenerse a partir de una función de onda, ψ , que se obtiene resolviendo la ecuación de onda de Schrödinger [66]. Los enfoques de la química cuántica son necesarios para modelar con precisión los sistemas a escala atómica y, lo que es más importante, para obtener información de la estructura electrónica de un sistema molecular [66 - 68].

Los métodos de estructura electrónica y la dinámica molecular se han utilizado de forma complementaria para investigar el sistema Ligando-Receptor [67]. Los métodos de estructura electrónica *semiempíricos* o *ab initio* usualmente se usan para optimizar la geometría de sistemas moleculares pequeños como los fármacos, obteniéndose parámetros tales como las constantes de fuerza de enlaces, longitudes/ángulos de equilibrio, así como cargas parciales de los átomos de un inhibidor para su posterior uso en la configuración de una simulación MD [66 - 70].

En las siguientes secciones nos enfocaremos sólo en los métodos de estructura electrónica que fueron usados para la optimización de geometría y obtención de información de la estructura electrónica de inhibidores de estructura de *orto*-aminoanilida que forman parte de este estudio.

A.1.1.1. Métodos semiempíricos: AM1

Aquí damos una mirada al enfoque AM1 (Austin Model 1), que se utiliza para lograr predicciones químicas significativas, a partir de cálculos computacionalmente poco costosos para la optimización de la geometría de los inhibidores [66]. AM1 es un modelo semiempírico basado en el formalismo del método HF que hace ciertas aproximaciones, donde principalmente omite las integrales de repulsión inter-electrónica entre pares de electrones y obtiene algunos parámetros a partir de datos experimentales [64-66].

A.1.1.2. Estructura electrónica: cargas.

La derivación de cargas atómicas y la construcción de una librería de campos de fuerza para pequeñas moléculas, como los inhibidores, son pasos clave a la hora de desarrollar un campo de fuerza necesario para llevar a cabo un análisis estructural y energético de la interacción proteína-ligando utilizando la mecánica molecular [67-70]. En este trabajo se calculan las cargas RESP para los cuatro ligandos de las simulaciones. El programa RESP ajusta el Potencial Electroestático Molecular (MEP), calculado mecánico-cuánticamente en superficies moleculares utilizando un modelo de carga puntual centrado en el átomo. La teoría más efectiva basada en métodos de estructura electrónica que se seleccionó para hacer esto, fue la teoría de "campo auto consistente" *ab initio* (SCF), que integra el método HF y un conjunto de bases adecuado, por ejemplo 6-31G*, para una descripción representativamente más precisa de los orbitales [65, 66]. El conjunto de bases establecido para obtener una cierta precisión para una propiedad como la carga de los átomos presentes en una molécula, suele ser diferente del requerido para la precisión correspondiente a la

energía de una molécula. Por lo tanto, el conjunto de bases para las propiedades, como la carga, debe adaptarse añadiendo funciones hasta que se obtenga la precisión deseada [65].

A.1.2. Métodos clásicos basados en campos de fuerza.

Dentro de las técnicas clásicas, la energía potencial molecular de un sistema químico se expresa sólo en función de las posiciones de los núcleos y se evalúa a través de expresiones analíticas clásicas y parámetros ajustados que incorporan información electrónica promediada. El conjunto de expresiones o fórmulas analíticas para cada una de las fuerzas involucradas con el movimiento atómico de un sistema molecular constituye lo que se denomina un campo de fuerza o FF [64-67].

A.1.2.1. Campos de fuerza.

En el modelado molecular, la función de un campo de fuerza es modelar la energía potencial de un sistema atómico. Un FF incluye las formas funcionales de la energía potencial debido a interacciones atómicas que incluyen el estiramiento o compresión de enlaces más allá de sus longitudes y ángulos de equilibrio, los efectos de la torsión sobre los enlaces simples, las atracciones de van der Waals o las repulsiones de los átomos cercanos, y las interacciones electrostáticas entre las cargas parciales en una molécula debido a los enlaces polares [64-70]. La forma funcional puede obtenerse con métodos analíticos o empíricos, y los parámetros del campo de fuerza, como las constantes de fuerza y los valores de distancias de equilibrio, pueden obtenerse reproduciendo datos experimentales o cálculos de químico-cuánticos [64,65,67-70]. Existen diversos FF en la bibliografía, siendo AMBER el FF utilizado en este trabajo [64]. Para un FF, la función $U_{\text{sis}}(r,p)$ describe la energía potencial molecular de un sistema que depende de las coordenadas r del sistema y de los parámetros p del campo de fuerza que se emplee, en este caso AMBER. Dicha función se expresa como muestra la ecuación (A1.1) [64].

$$\begin{aligned}
 U(r,p) = & \sum_{stretch} \frac{1}{2} K_{AB} (R_{AB} - R_{e,AB})^2 + \sum_{bend} \frac{1}{2} K_{ABC} (\theta_{ABC} - \theta_{e,ABC})^2 \\
 & + \sum_{dihedros} \frac{U_0}{2} (1 - \cos(n(\chi - \chi_e))) + \sum_{impropias} \frac{K}{2 \sin^2 \psi_e} (\cos \psi - \cos \psi_e)^2 \quad (A.1.1) \\
 & \sum_{van\ der\ Waals} \left(\frac{C_{AB}^{12}}{R_{AB}^{12}} - \frac{C_{AB}^6}{R_{AB}^6} \right) + \frac{1}{4\pi\epsilon_0} \sum_{Cargas} \frac{Q_A Q_B}{R_{AB}} + \sum_{H-bonds} \left(\frac{C_{AB}^{12}}{R_{AB}^{12}} - \frac{C_{AB}^{10}}{R_{AB}^{10}} \right)
 \end{aligned}$$

El primer término hace referencia a la energía de interacción por *stretching* entre pares de átomos enlazados entre sí; el segundo a la energía requerida por el *bending* del ángulo formado por tres átomos; el tercer término representa la energía torsional de rotación en torno a un enlace diedro; el cuarto, el potencial impropio debido al ángulo fuera del plano; el quinto y sexto término describen las interacciones no enlazantes entre átomos definidos por un término de atracción-repulsión de tipo Lennard-Jones para las fuerzas de Van der Waals y un término Coulómbico para interacciones electrostáticas [64, 65]; el último término adicional representa los puentes de hidrógeno. Un esquema de dichas interacciones se muestra en la figura A1.2.

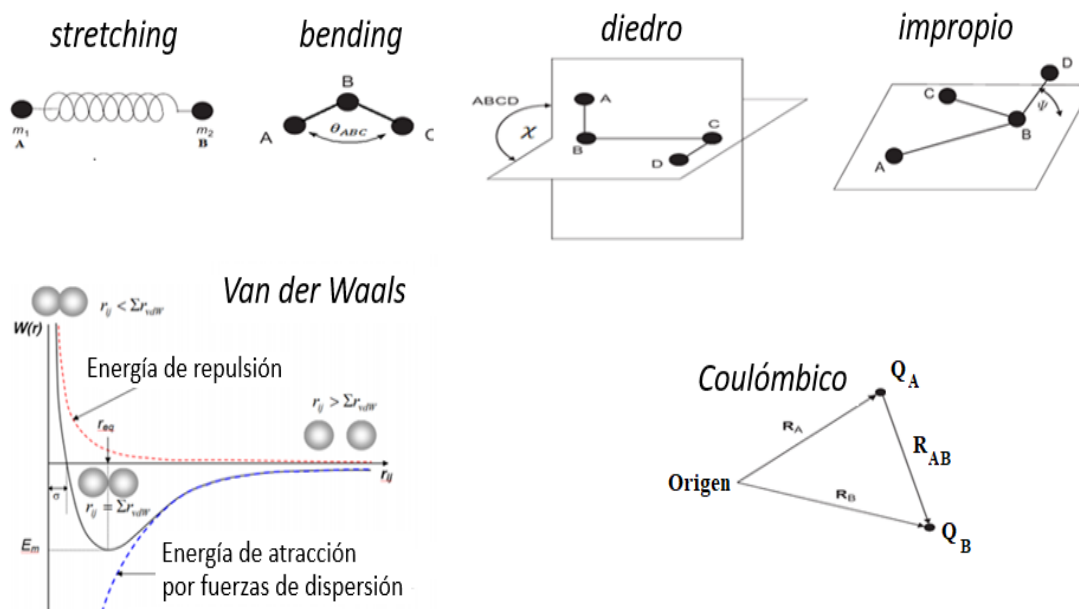


Figura A.1.2. Ilustración de términos energéticos de enlace (arriba) y no-enlace (abajo) de un campo de fuerza [65].

A.1.3. Mecánica Molecular.

En la Mecánica Molecular (MM) se aplica la mecánica de Newton en una molécula o un sistema molecular para modelar su estructura detallada y las propiedades físicas mediante el cálculo de la energía de una molécula en términos de las interacciones enlazantes y no-enlazantes. Los métodos MM atómicos tratan moléculas como bolas unidas por resortes en el que cada átomo es una sola partícula con un radio asignado (típicamente el radio de van der Waals), polarizabilidad y carga neta constante (generalmente derivado de cálculos químico cuánticos y experimentos). Por medio de la minimización de esta función potencial del FF, los métodos de MM predicen geometrías en el equilibrio y energías relativas [64-66].

A.1.3.1. Minimización de la energía

Este método trata de encontrar una o más configuraciones estables de una molécula, a partir de una optimización de la energía potencial del sistema, como se ilustra en la figura A1.3. La minimización suele utilizarse con tres finalidades principalmente: (i) la obtención de la geometría más estable de un sistema; (ii) el cálculo de la energía relativa entre las estructuras de equilibrio de diferentes especies; (iii) la obtención de una estructura de partida que no presente interacciones fuertemente desfavorables para una simulación de dinámica molecular [64,66,68].

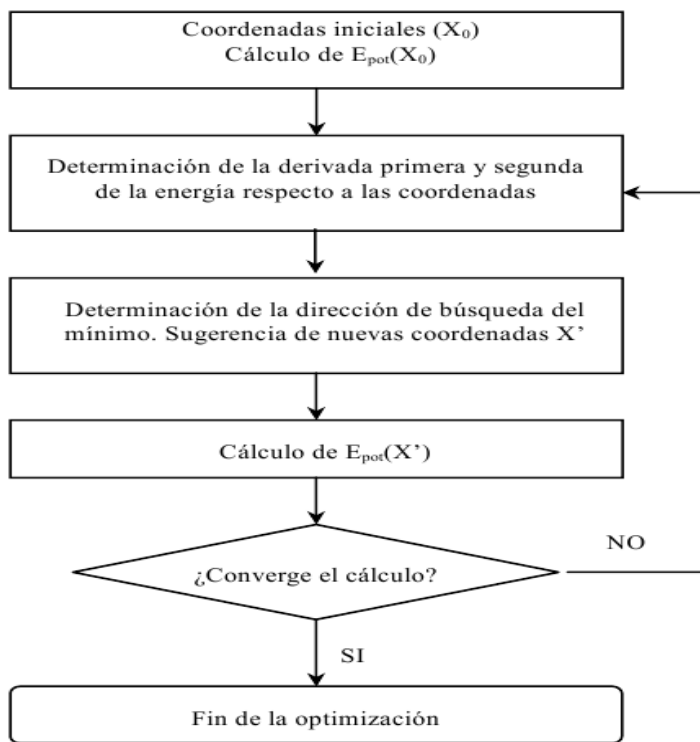


Figura A.1.3. Algoritmo básico de los métodos de Mecánica Molecular.

Los métodos más utilizados hoy son el de descenso más pronunciado (*steepest descent*) y el de gradiente conjugado (*conjugate gradients*), basados en gradientes, los cuales hacen uso de la primera derivada de la función de energía potencial respecto a la posición de los átomos [64,65].

A.1.4. Dinámica molecular

La simulación de dinámica molecular (MD) es un método de modelado *in silico* para el estudio de movimientos de partículas (principalmente átomos). Una simulación MD, comienza con una conformación que es un mínimo energético, donde a cada átomo se asigna una serie de componentes de velocidad v_x , v_y , v_z que son elegidas al azar a partir de una distribución de Maxwell-Boltzmann a una temperatura asignada. Luego, se aplica la segunda ley de Newton a cada átomo, donde cada componente de la fuerza sobre el átomo i se calcula según la ecuación (A.1.2), donde U se obtiene a partir del FF de la Mecánica Molecular. Se integra numéricamente la segunda ley de Newton para obtener la posición de cada átomo en los tiempos Δt , $2\Delta t$, $3\Delta t$, ... El paso de integración también resulta ser importante en el caso de macromoléculas, el cual tiene que ser menor que el movimiento más rápido del sistema [64-66].

Las nuevas configuraciones se obtienen solucionando la ecuación diferencial contenida en la segunda ley de Newton principalmente en su forma potencial (A.1.2):

$$\frac{-dU}{dx} = m \frac{d^2x}{dt^2} \quad (\text{A.1.2})$$

Donde X son las coordenadas nucleares, m la masa y U la energía potencial del sistema.

Estas ecuaciones no se pueden resolver analíticamente debido a que la fuerza ejercida sobre cada átomo depende de la posición de los otros, dando como resultado un movimiento acoplado de partículas. Por esto, se usan métodos de diferencias finitas en base a la fuerza que actúa en cada átomo, coordenada y velocidad en el paso anterior y desde este se calculan las futuras coordenadas, asumiendo que dentro de este intervalo la fuerza será constante. Dentro de los algoritmos de integración numérica de las ecuaciones de Newton, es muy común el uso del algoritmo de *leap-frog* o "salto de rana", que es una modificación del algoritmo de *Verlet* [64, 65, 68,69]. Lo anterior, se resume en un diagrama ilustrado en la Figura A.1.4.

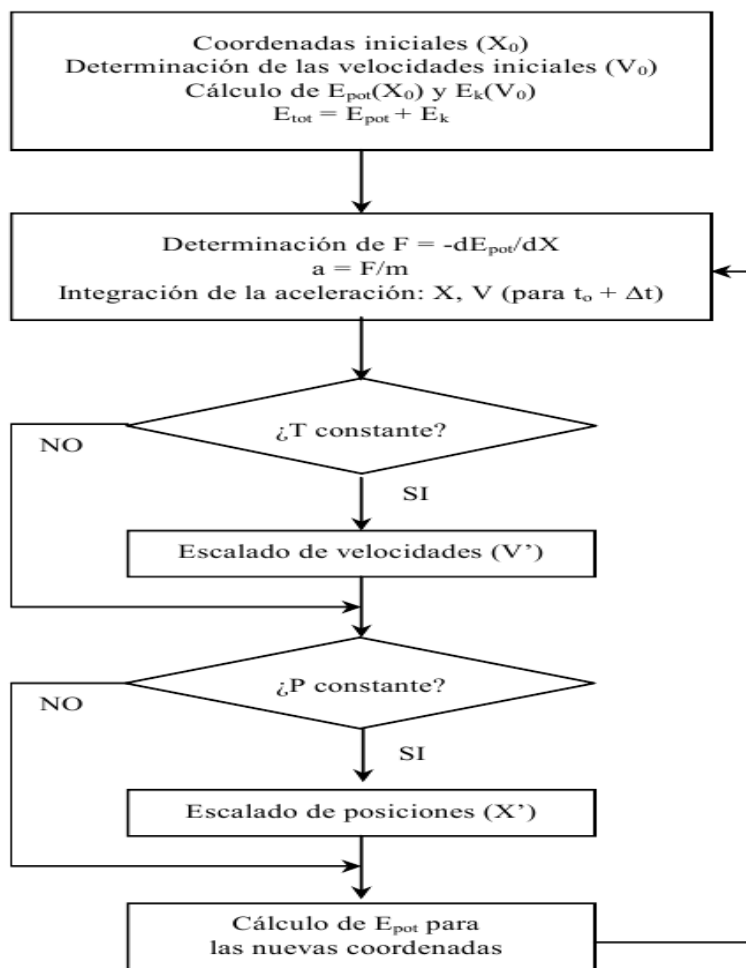


Figura A.1.4. Algoritmo básico de los métodos de dinámica molecular.

Las posiciones y cantidades de movimiento de todas las partículas de un sistema definen un estado microscópico. Las posiciones y cantidades de movimiento de todas las partículas en el sistema son adjudicados como coordenadas de un espacio $6N$ dimensional también llamado *espacio de fase* [70]. En resumen, la evolución de un sistema con el tiempo, corresponde a una trayectoria en el espacio de fase y puede determinarse mediante la resolución de las ecuaciones de movimiento de Newton basado en la energía potencial.

A.1.4.1. MMPBSA: Método de punto final para cálculo aproximado de $\Delta G_{\text{enlazamiento}}$

MMPBSA es un método que combina la mecánica molecular y solventes continuos [47]. Hipotéticamente, un evento de unión entre receptor y ligando se divide en dos etapas: asociación en la fase gaseosa y disolución en la fase acuosa [48]. Mientras que el *single point* de energía en fase gaseosa se evalúa con un campo de fuerza clásico, en este caso AMBER, el término de disolución se cuantifica utilizando el modelo de disolvente implícito Poisson-Boltzmann, (PB) que es el núcleo de este método.

La configuración de los cálculos de MMPBSA comienza con la creación de archivos de parámetros/topologías despojados de moléculas de agua e iones, utilizando los parámetros recomendados para el método de análisis solicitado [49]. Posteriormente, se generan conjuntos conformacionales de complejos, receptores y ligandos mediante la extracción automatizada de instantáneas de la simulación MD en un intervalo especificado por el usuario. Las instantáneas son creadas de acuerdo al enfoque de trayectoria única [50], obteniéndose desde la trayectoria específica para cada complejo en estudio. En el paso final del proceso de configuración, se crean los archivos necesarios para realizar los cálculos de energía con el script `mm_pbsa.pl` de AMBER. Los cálculos pueden ser iniciados usando los scripts *batch* y *shell* proporcionados en cualquier sistema de computación en el que AMBER esté disponible [49-51].

A.1.5. Diseño de drogas asistido por computadoras

En las últimas dos décadas, los enfoques *in silico* se han convertido en una parte integral de prácticamente todos los programas racionales (academia e industria) de descubrimiento de fármacos [71-73]. Los enfoques racionales en el descubrimiento de fármacos se clasifican tradicionalmente como basados en estructura (guiados por receptor: RG) y basados en análogos (basado en ligando: LB) [72,73]. En ambos casos los enfoques se han aplicado para obtener información sobre las características de unión del fármaco con el objetivo biológico siguiendo los principios del diseño de drogas asistido por computadora (CADD), que se enfoca en la comprensión de tres factores esenciales para el diseño de fármacos: i) las características que hacen que una macromolécula sea un receptor farmacológico, ii) las propiedades que distinguen a una molécula pequeña para que actúe como un fármaco y iii) las interacciones que facilitan un ajuste óptimo de una molécula como fármaco en un objetivo susceptible de ser medicado [71-73].

A.1.5.1. Diseño de drogas basado en el receptor.

El diseño de fármacos basado en estructuras (o diseño directo de fármacos) se basa en el conocimiento de la estructura tridimensional del objetivo biológico obtenido a través de métodos como la cristalografía de rayos X o la espectroscopia de RMN [71]. El diseño de fármacos basado en la estructura es el diseño y la optimización de una estructura química con el objetivo de identificar un compuesto adecuado para las pruebas clínicas: un fármaco candidato. Se basa en el conocimiento de la estructura tridimensional de la droga y cómo su forma y carga hacen que interactúe con su objetivo biológico, lo que finalmente provoca un

efecto médico [72,73]. La explosión de información genómica, proteómica y estructural ha proporcionado cientos de nuevos objetivos y oportunidades para el descubrimiento de compuestos en el futuro [73].

A.1.5.2. *Docking* o Acoplamiento molecular.

El *Docking* se lleva a cabo utilizando un algoritmo informático automatizado que determina la unión de un compuesto al sitio activo de una proteína [74]. Esto incluye determinar la orientación del compuesto, su geometría conformacional y la puntuación. Aunque todos los compuestos (moléculas pequeñas) se someten a cambios de conformación durante el acoplamiento, la proteína se mantiene rígida en una geometría fija en la mayoría de los casos. Algunos programas facilitan la alteración en la conformación del sitio activo, permitiendo un acoplamiento más flexible. En este trabajo se utiliza el programa de acoplamiento Autodock 4.6, que contiene dos componentes clave a saber: *el algoritmo de búsqueda y la función de puntuación* (ver Apéndice A.10). El algoritmo de búsqueda posiciona las moléculas en una multitud de ubicaciones, orientaciones y conformaciones dentro del sitio activo [75]. La elección del algoritmo de búsqueda determina la minuciosidad del programa para verificar las posibles posiciones de la molécula y el tiempo empleado [75,76]. Por otro lado, las funciones de puntuación son métodos matemáticos aproximados rápidos utilizados para predecir la fuerza de la interacción no covalente (también denominada afinidad de unión) entre dos moléculas después de que se hayan acoplado. Las orientaciones moleculares con átomos colocados en forma razonable se anotan y las otras se descartan. Los modos energéticamente favorables de unión de ligandos se almacenan como poses diferentes. Estas poses que representan las interacciones proteína-ligando se puntúan en términos de entalpías o energías libres de unión [76].

A.1.5.3. Diseño de fármacos basado en ligando.

El diseño de fármacos basado en ligandos ha surgido en paralelo a los enfoques basados en la estructura [75]. Estos enfoques se complementan, ya que en este enfoque se desconoce la estructura del objetivo biológico, pero se cuenta con información de un conjunto de inhibidores activos para el objetivo biológico [75]. Estas moléculas se pueden usar para derivar un modelo de farmacóforo que define las características estructurales mínimas necesarias que una molécula debe poseer para unirse al receptor. En otras palabras, se puede re-construir un modelo del objetivo biológico basándose en el conocimiento de lo que se une a él, y este modelo a su vez puede usarse para diseñar nuevas entidades moleculares que interactúen con el objetivo [77,78]. Los modelos QSAR y de farmacóforos son los métodos más importantes en el diseño de fármacos basados en ligandos [77]. Una Relación Estructura-Actividad Cuantitativa, permite derivar una correlación entre las propiedades calculadas de las moléculas y su actividad biológica determinada experimentalmente. Estas relaciones QSAR a su vez pueden usarse para predecir la actividad de nuevos análogos. Han transcurrido más de cincuenta años desde que Corwin Hansch [79] fundó el campo del modelo de Relaciones Cuantitativas Estructura-Actividad (QSAR). Inicialmente conceptualizado como la extensión lógica de la química-física orgánica, el modelado QSAR ha crecido, diversificado y evolucionado desde su aplicación a pequeñas series de compuestos análogos usando métodos de regresión relativamente simples, hasta el análisis de grandes conjuntos de datos que comprenden miles de diversas estructuras moleculares

usando una amplia variedad de técnicas estadísticas y de aprendizaje automático. Estos modelos predictivos filtran rápidamente grandes bases de datos para identificar nuevas moléculas cabezas de serie con una actividad biológica mejorada y mayor potencia, abriendo así el camino a nuevos tipos de estructuras para la investigación de fármacos [77].

A.1.5.4. QSAR

El desarrollo racional de un ligando con determinada afinidad por un sitio de unión se puede abordar mediante técnicas computacionales y estadísticas que relacionen la actividad biológica de una molécula con sus características estructurales mediante lo que se conoce como estudios de relaciones estructura-actividad cuantitativos (QSAR) [75].

Como se observa en la Figura A1.5, el primer paso de la modelación QSAR es la colección y preparación de un set de datos que deben obtenerse desde información experimental existente sobre un conjunto de moléculas de interés farmacológico, de las cuales se cuenta con sus estructuras geométrica 3D previamente optimizadas y datos de su actividad biológica informada. En el estudio QSAR, las variaciones de actividad/toxicidad/propiedad biológica dentro de una serie de compuestos se correlacionan con los cambios en un grupo de características de las moléculas que se calculan y son denominadas *descriptores* [75, 77, 80]. Los descriptores moleculares son información química que está codificada dentro de las estructuras moleculares, sirven como variables independientes de un modelo QSAR y son colectivamente responsables de una actividad particular de la molécula [80]. Esta actividad debiese seguir una distribución uniforme, dividiéndose el conjunto de compuestos en: *conjunto de entrenamiento y conjunto de prueba*.

El modelado QSAR típicamente describe estructuras moleculares en términos de los descriptores y luego correlaciona estos descriptores moleculares con actividades observadas utilizando diversos métodos estadísticos. El método QSAR predice una cierta propiedad de una molécula a partir de su estructura como una expresión matemática en la forma de (ec. A.1.3):

$$y = m_1x_1 + m_2x_2 + \dots + C \quad (\text{A.1.3})$$

Donde, y es la propiedad pronosticada (variable dependiente) y x_1, x_2, \dots, x_N ; son descriptores.

QSAR usa descriptores que son un solo número que describe algún aspecto de la molécula, como el peso molecular, el número de átomos, los índices topológicos, etc. Los coeficientes m_1, m_2, \dots, m_N ; en la ecuación QSAR son pesos de los descriptores obtenidos mediante el uso de varios métodos de ajuste de curvas. Las actividades y propiedades modeladas por QSAR se conocen como variables dependientes (y) del modelo QSAR [75, 80, 81].

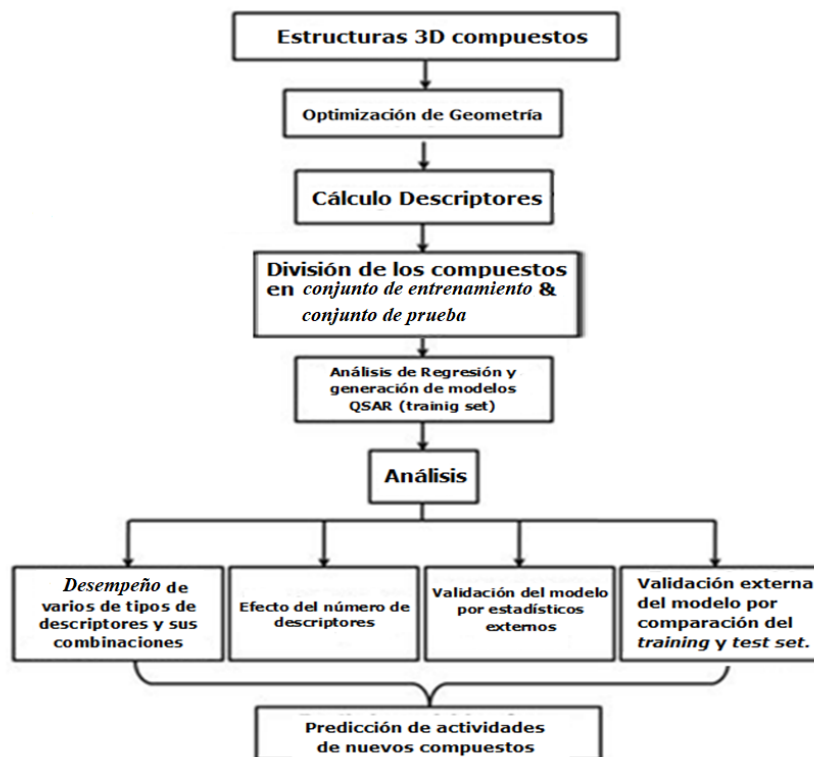


Figura A.1.5. Pasos de modelación QSAR.

Para la correlación estructura-actividad, la técnica de mínimos cuadrados parciales (PLS) fue la utilizada en esta tesis, debido a que presenta ventajas comparativas frente a otros métodos, incluida la capacidad de manejar más variables descriptoras que compuestos, resultados biológicos múltiples, al mismo tiempo que proporciona una mayor precisión predictiva y un riesgo mucho menor de correlación aleatoria [82]. PLS es una técnica importante, introducida por Hermann y Svante Wold [82-84], que produce una ecuación lineal para describir o predecir diferencias en los valores de una o más propiedades a partir de las diferencias en los valores de otras propiedades. Las propiedades descritas o predichas se denominan 'variables dependientes' o, en la literatura PLS, 'bloque Y'. Las propiedades de descripción o predicción se denominan 'variables independientes' o 'bloque X'.

El poder predictivo de un modelo QSAR se puede verificar luego a través de medidas estadísticas tales como el coeficiente de correlación entre los valores reales y los pronosticados [84]. Por otra parte, los modelos deben ser también evaluados por medio de una validación externa sobre el conjunto de prueba, en este trabajo adoptamos la propuesta por Golbraikh y Tropsha [30,31].

A.1.6. Modelado de farmacóforo.

Últimamente el modelado de farmacóforos ha adquirido una gran importancia como enfoque basado en ligando debido a su simplicidad. De acuerdo con la definición de IUPAC [87]: "Un farmacóforo es el conjunto de características estéricas y electrónicas que son necesarias para asegurar las interacciones supramoleculares óptimas con una estructura específica de un objetivo biológico y para activar (o bloquear) su respuesta biológica".

Un farmacóforo no representa ninguna molécula real, sino que representa el patrón de interacción común de un grupo de compuestos con su objetivo, detectados luego de un paso clave conocido como alineamiento (ver Figura A.1.6) [88,89].

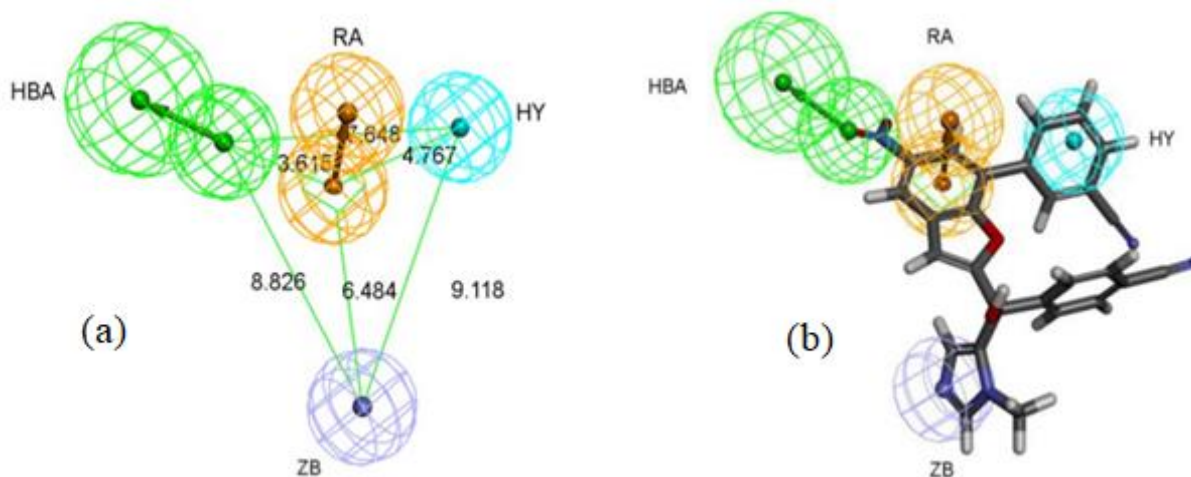


Figura A.1.6. (a) Representación de un farmacóforo con sus propiedades químicas y distancias entre sus diferentes regiones. HBA; aceptor de puente de hidrógeno, RA; anillo aromático, HY; hidrofóbica, ZB; enlazante de zinc. (b) Una molécula solapada con el farmacóforo.

APÉNDICE 2: INTERACCIONES NO COVALENTES.

*“La maquinaria interna de la vida, la química de las partes, es algo hermoso.
Toda la vida está interconectada con el resto de la vida”.*

-Richard P. Feynman

A.2. INTERACCIONES MOLECULARES LIGANDO - RECEPTOR.

A.2.1. Consideraciones generales.

Para apreciar los mecanismos de la acción de fármacos, es importante comprender las fuerzas de interacción que unen ligandos a sus receptores. La fuerza motriz para la interacción fármaco-receptor puede considerarse como un estado de baja energía del complejo fármaco-receptor (ver figura A.2.1).

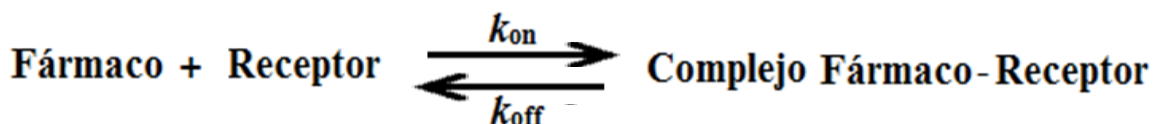


Figura A.2.1. Equilibrio entre un fármaco, un receptor, y un complejo fármaco-receptor.

Donde k_{on} es la constante de velocidad para la formación del complejo fármaco-receptor, que depende de las concentraciones del fármaco y el receptor, y k_{off} es la constante de velocidad para la descomposición del complejo, que depende de la concentración del complejo fármaco-receptor. La actividad biológica de un fármaco está relacionada con su afinidad por el receptor, es decir, la estabilidad del complejo fármaco-receptor. Esta estabilidad se mide comúnmente por lo difícil que es que el complejo se disocie, la cual se representa por K_d . La constante de disociación K_d de complejo fármaco-receptor en equilibrio queda representada por (ec. A.2.1):

$$K_d = \frac{[\text{fármaco}][\text{receptor}]}{[\text{complejo fármaco-receptor}]} \quad (\text{A.2.1})$$

Dado que K_d es una constante de disociación, cuanto menor es la K_d , mayor es la concentración del complejo fármaco-receptor, lo que se traduce a una mayor estabilidad del complejo, producto de una mayor afinidad del fármaco por el receptor. Los fármacos y receptores existen como un conjunto de confórmers en solvente [90]. Por lo tanto, para formar un complejo solvatado con el receptor, la molécula de fármaco necesita desplazar las moléculas de disolvente que ocupan el sitio de unión del receptor [91]. Esto es posible solo cuando las interacciones entre los fármacos y el receptor son más fuertes que sus interacciones individuales con las moléculas de disolvente [90]. La formación del complejo ligando-receptor es entrópicamente desfavorable e induce una pérdida en los grados de libertad conformacional, rotacional y traslacional, tanto del fármaco como del sitio activo del receptor en específico, por lo que se espera que la pérdida entrópica sea compensada por contactos entálpicos favorables, es decir, principalmente interacciones no-covalentes.

La actividad de una molécula pequeña (fármaco) se inicia entonces por su interacción a nivel atómico con la macromolécula (receptor u objetivo). Esta asociación se estabiliza mediante una plétora de interacciones intermoleculares fármaco-receptor que son de naturaleza no covalente en la mayoría de los casos [90]. Estas interacciones se forman espontáneamente entre los átomos con una disminución de la energía libre (ΔG). La interacción de un fármaco con el sitio de unión de un receptor depende de la

complementariedad de ajuste entre las dos moléculas, tal como se establece en la Hipótesis de Llave Cerradura (*Lock and Key*) por Emil Fischer [92].

A.2.2. Interacciones importantes involucradas en complejos Ligando-Receptor.

Las interacciones involucradas en el complejo fármaco-receptor [90-92] son las mismas fuerzas experimentadas por todas las moléculas orgánicas que interactúan, e incluyen: enlaces covalentes, interacciones iónicas (electrostáticas), interacciones ión-dipolo y dipolo-dipolo, enlaces de hidrógeno, interacciones por transferencia de carga, interacciones hidrofóbicas, interacciones, enlaces de halógenos e interacciones de van der Waals.

Las interacciones débiles generalmente son posibles solo cuando las superficies moleculares son cercanas y complementarias, es decir, la fuerza de la unión depende de la distancia. Como se describió anteriormente, la formación espontánea de un enlace entre átomos ocurre con una disminución en la energía libre, es decir, una interacción no covalente ocurrirá solo cuando hay un ΔG negativo. El cambio en la energía libre de enlace está relacionado con la constante de equilibrio de enlace (K_{eq}) de acuerdo con la ecuación (A2.2). Por lo tanto, a una temperatura fisiológica (37°C), los cambios en la energía libre de unas pocas kilocalorías por mol pueden tener un efecto importante en el establecimiento de las interacciones que determinan la formación o no del complejo.

$$\Delta G^0 = -RT \ln K_{eq} \quad (A2.2)$$

Una interacción particular puede valer diferentes cantidades de energía libre dependiendo de las estructuras moleculares involucradas. La multiplicidad de interacciones en un complejo proteína-ligando es un compromiso entre fuerzas atractivas y repulsivas, los procesos de solvatación, las interacciones de largo alcance y los cambios conformacionales (que a menudo se descuidan).

Sería deseable que las interacciones observadas sean aditivas; sin embargo, las interacciones moleculares tienden a comportarse de una manera altamente no aditiva [90]. Además, es muy fácil ser engañado por las interacciones fármaco-receptor en las estructuras cristalinas, que presentan vistas estáticas de las interacciones y no tienen en cuenta el costo de energía del desplazamiento de las moléculas de agua del sitio de unión [90,91]. En general, las interacciones no covalentes formadas entre un fármaco y un receptor son débiles por naturaleza; en consecuencia, los efectos producidos son reversibles. Debido a esto, un fármaco se vuelve inactivo tan pronto como disminuye su concentración en los fluidos extracelulares, generalmente por metabolismo. Con frecuencia, es deseable que el efecto del fármaco dure solo durante un tiempo limitado, de modo que la acción farmacológica pueda finalizar. A veces, sin embargo, el efecto producido por un medicamento debe persistir, e incluso ser irreversible. En las siguientes sub-secciones, los diversos tipos de interacciones fármaco-receptor se discuten brevemente. Estas interacciones son aplicables a todos los tipos de receptores, incluidas las enzimas y el ADN.

A.2.3. Enlace covalente

El enlace covalente es el enlace más fuerte, con valores entre -30 a -230 kcal/mol en estabilidad, entre los más observados en química orgánica y biomoléculas. Rara vez se forma por una interacción fármaco-receptor, excepto con enzimas y ADN.

A.2.4. Interacciones no - covalentes.

A.2.4.1. Interacciones iónicas o Electroestáticas.

Para los receptores de proteínas a pH fisiológico (significa generalmente tomado para pH 7.4, el pH de la sangre), los grupos básicos tales como las cadenas laterales de amino de arginina, lisina y, en mucha menor medida, histidina, están protonados y, por lo tanto, proporcionan un ambiente catiónico. Los grupos ácidos, tales como las cadenas laterales de ácido carboxílico de ácido aspártico y ácido glutámico, se desprotonan para dar grupos aniónicos. Los fármacos y receptores se atraerán mutuamente siempre que tengan cargas opuestas. Esta interacción iónica puede ser efectiva a distancias más largas que las requeridas para otros tipos de interacciones, y pueden persistir por más tiempo. Una interacción iónica simple puede proporcionar un $\Delta G^0 = -5$ kcal/mol, que disminuye por el cuadrado de la distancia entre las cargas. Si esta interacción se ve reforzada por otras interacciones simultáneas, la interacción iónica se vuelve más fuerte ($\Delta G^0 = -10$ kcal/mol) y persiste más tiempo.

El medicamento antidepresivo **pivagabina** (Tonerg®) se usa como ejemplo de una molécula que puede participar hipotéticamente en una interacción iónica con un residuo de arginina (Figura A.2.2).

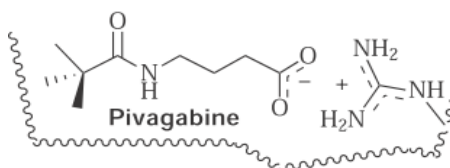


Figura A.2.2. Ejemplo de interacción electrostática (iónica). La línea ondulada representa la cavidad del receptor.

A.2.4.2. Interacciones ión-dipolo & dipolo-dipolo.

Como resultado de la mayor electronegatividad de átomos como oxígeno, nitrógeno, azufre y halógenos en relación con la del carbono, los enlaces C-X en fármacos y receptores, donde X es un átomo electronegativo, tendrán una distribución asimétrica de electrones; esto produce dipolos electrónicos. Estos dipolos en una molécula de fármaco pueden ser atraídos por iones (interacción ión-dipolo) u otros dipolos (interacción dipolo-dipolo) en el receptor, siempre que las cargas de signo opuesto estén alineadas correctamente. Debido a que la carga de un dipolo es menor que la de un ión, una interacción dipolo-dipolo es más débil que una interacción ión-dipolo. En la Figura A.2.3, la droga contra insomnio **zaleplon** (Sonata®) se usa para demostrar estas interacciones, que pueden proporcionar un ΔG^0 desde -1 a -7 kcal/mol.

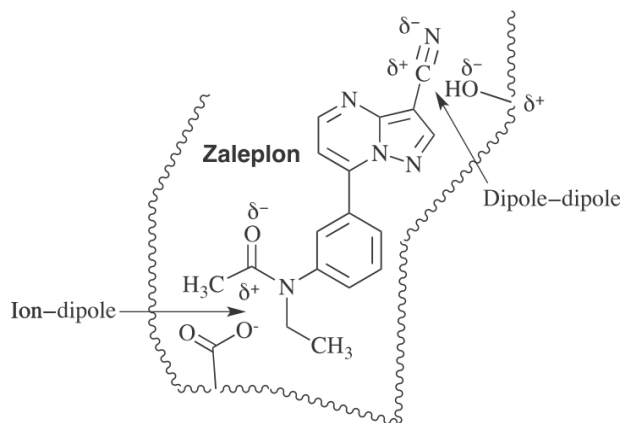


Figura A.2.3. Ejemplos interacciones ión-dipolo y dipolo-dipolo. La línea ondulada representa cavidad del receptor.

A.2.4.3. Puentes de Hidrógeno.

Los *enlaces de hidrógeno* son un tipo de interacción dipolo-dipolo formado entre el protón de un grupo X-H, donde X es un átomo electronegativo y uno o más átomos electronegativos (Y) que contienen un par de electrones no enlazados. Los enlaces de hidrógeno más significativos se producen en moléculas donde X e Y son N y O y, en menor medida, F [94,95]; También se han descrito casos especiales interesantes de enlaces de hidrógeno débiles para X = C [93,94]. X elimina la densidad de electrones del hidrógeno por lo que este último tiene una carga positiva parcial, que es fuertemente atraída por los electrones no enlazantes de Y. La interacción se denota como una línea de punteada, -X-H ... Y-, para indicar que todavía existe un enlace covalente entre X y H, pero que la interacción entre H e Y también ocurre. En esta descripción, X se denomina donador de enlace de hidrógeno e Y es el aceptor de enlace de hidrógeno. Cuando X e Y son equivalentes en electronegatividad y grado de ionización, el protón se puede compartir por igual entre los dos grupos, es decir, -X ... H ... Y-, denominado enlace de hidrógeno de baja barrera [94,97]. Los valores medios de la distancia de enlace de hidrógeno entre átomos electronegativos de diferentes grupos funcionales han sido ampliamente estudiados en la literatura [96].

El enlace de hidrógeno es exclusivo del hidrógeno porque es el único átomo que puede llevar una carga positiva a pH fisiológico mientras permanece unido covalentemente a una molécula, y que también es lo suficientemente pequeño para permitir el acercamiento de un segundo átomo electronegativo. Hay enlaces de hidrógeno intramoleculares e intermoleculares; los primeros son más fuertes, por ejemplo, el ácido salicílico utilizado en los remedios para la eliminación de verrugas (Figura A.2.4). El enlace de hidrógeno intramolecular es una propiedad importante de las moléculas, este puede tener un efecto significativo en los enfoques de modificación de compuestos cabeza de serie [95,97,98].

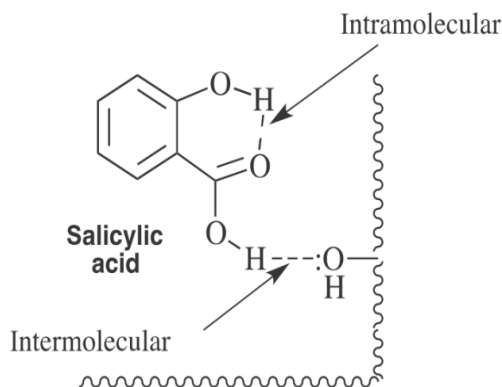


Figura A.2.4. Ejemplos de enlaces de hidrógeno. La línea ondulada representa la cavidad del receptor.

La conformación bioactiva de una molécula es la conformación óptima de la molécula cuando se une a su receptor. Cuando hay grupos donantes y aceptores de enlace de hidrógeno en un compuesto que tienen la posibilidad de interactuar para formar un anillo intramolecular de cinco a siete miembros, esas interacciones producirán una conformación estable que puede aproximarse o no a la conformación bioactiva.

El enlace de hidrógeno intramolecular produce estructuras que pueden considerarse bioisómeros de compuestos bicíclicos, un tipo de andamio a considerar (Figura A.2.5) [98]. Algunos enlaces de hidrógeno intramoleculares son lo suficientemente fuertes como para persistir en el agua [97,98].

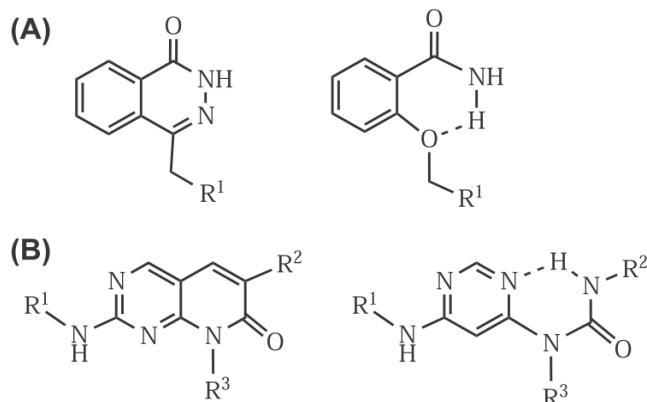


Figura A.2.5. Dos ejemplos (A y B) de cómo los enlaces de hidrógeno intramoleculares pueden imitar un heterociclo bioisostérico.

Debido a que el enlace de hidrógeno intramolecular elimina un donador y un aceptor del resto de la molécula, aumenta su lipofilidad y la permeabilidad de la membrana y disminuye su solubilidad acuosa. Esto puede tener un impacto significativo en la farmacocinética.

La estimación de la contribución energética de los enlaces de hidrógeno es a veces problemática, ya que en la solución acuosa la formación de un enlace de hidrógeno requiere la de-solvatación tanto del donante como del aceptor. Los enlaces de hidrógeno de mayor energía se producen en el caso de los enlaces de hidrógeno reforzados con carga.

Como ejemplo, en muchos casos, un puente de sal es realmente una combinación de enlaces de hidrógeno y una interacción electrostática (iónica). El ΔG^0 para enlaces de hidrógeno puede estar entre -1 y -7 kcal/mol, pero usualmente está en el rango de -3 a -5 kcal/mol. Las afinidades de unión aumentan en aproximadamente un orden de magnitud por enlace de hidrógeno.

A.2.4.4. Complejos de transferencia de carga.

Cuando una molécula (o grupo) que es un buen donador de electrones entra en contacto con una molécula (o grupo) que es un buen aceptor de electrones, el donante puede transferir parte de su carga al aceptor. Esto forma un complejo de transferencia de carga, que, en efecto, es una interacción molecular dipolo-dipolo. La energía potencial de esta interacción es proporcional a la diferencia entre el potencial de ionización del donante y la afinidad electrónica del aceptor.

Los grupos de donadores contienen electrones π , tales como alquenos, alquinos y residuos aromáticos con sustituyentes donadores de electrones, o grupos que contienen un par de electrones no unidos, tales como oxígeno, nitrógeno y restos de azufre. Los grupos aceptores contienen orbitales π deficientes en electrones, tales como alquenos, alquinos y restos aromáticos que tienen sustituyentes que atraen electrones y protones débilmente ácidos. Hay grupos en los receptores que pueden actuar como donantes de electrones, como el anillo aromático de tirosina o el grupo carboxilato del aspartato y otros actuar como aceptores de electrones, como la cisteína. Además, existen residuos que pueden actuar como donantes y aceptores de electrones, tales como histidina, triptófano y asparagina. El fungicida, **clorotalonil**®, se usa en la figura A.2.6 como un ejemplo hipotético de una interacción de transferencia de carga con una tirosina. El ΔG^0 para las interacciones de transferencia de carga también puede variar de -1 a -7 kcal/mol.

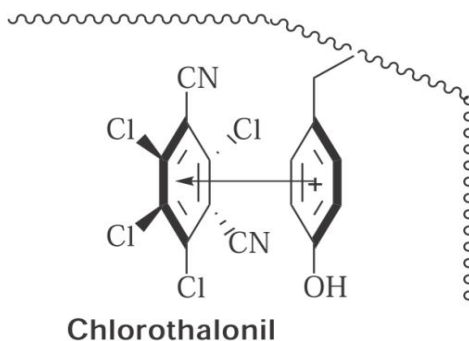


Figura A.2.6. Ejemplo de una interacción de transferencia de carga. La línea ondulada representa la cavidad del receptor.

A.2.4.5. Interacciones hidrofóbicas.

En presencia de una molécula o región no polar de una molécula, las moléculas de agua circundantes se orientan a sí mismas y, por lo tanto, se encuentran en un estado de mayor energía que cuando solo están presentes otras moléculas de agua.

Cuando dos grupos no polares, tal como un grupo lipofílico de un fármaco y un grupo no polar en el receptor, cada uno rodeado por moléculas de agua ordenadas, se aproximan,

estas moléculas de agua se desordenan en un intento de asociarse entre sí. Este aumento en la entropía, por lo tanto, da como resultado una disminución en la energía libre ($\Delta G = \Delta H - T \Delta S$), que estabiliza el complejo fármaco-receptor. Esta estabilización se conoce como interacción hidrofóbica (ver figura. A.2.7).

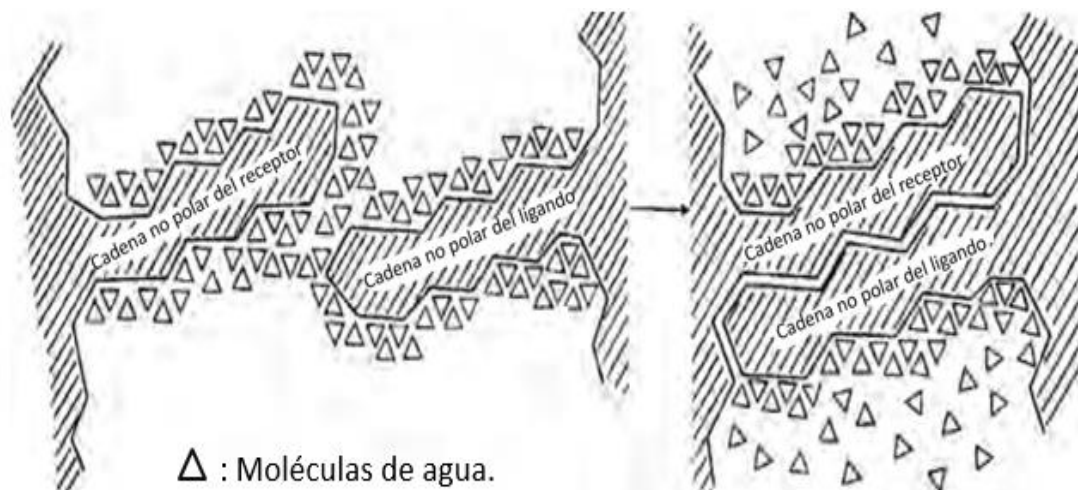


Figura A.2.7. Formación de interacciones hidrofóbicas [de korolkovas, a. (1970). *Essentials of molecular pharmacology*, p. 172. Wiley, Nueva York. Este material fue adaptado y se reproduce con el permiso de John Wiley & Sons, Inc. Y con el permiso de Kopple, K. D. 1966. *Péptidos y aminoácidos*. Addison-Wesley, Reading, MA.

En consecuencia, esta no es una fuerza atractiva de dos grupos no polares "disolviéndose" uno en el otro, sino, más bien, es la energía libre disminuida del grupo no polar debido a la entropía aumentada de las moléculas de agua circundantes. Jencks ha sugerido que las fuerzas hidrofóbicas pueden ser el factor individual más importante responsable de las interacciones intermoleculares no covalentes en la solución acuosa [99]. La adición de un único grupo metilo que puede ocupar un bolsillo de unión al receptor mejora la unión por -1.5 kcal/mol. Esta interacción hidrofóbica se conoce como interacción de un "metilo mágico" [100]. En la figura A.2.8(a), el anestésico tópico **butamben** se representa en una interacción hidrofóbica hipotética con un grupo isoleucina.

Otro tipo de interacción hidrofóbica, llamada interacción de apilamiento $\pi - \pi$, involucra dos grupos arilo [101-103]. Las interacciones de apilamiento (*Stacking*) $\pi - \pi$ más comunes implican una disposición paralela de anillos aromáticos [104], en la que los electrones π interactúan en una disposición cara a cara [105]. En la figura A.2.8 (b), el anillo de fenilo del medicamento anticonvulsivo **lacosamida** (Vimpat®) se muestra en una hipotética interacción de apilamiento π con un receptor de fenilalanina.

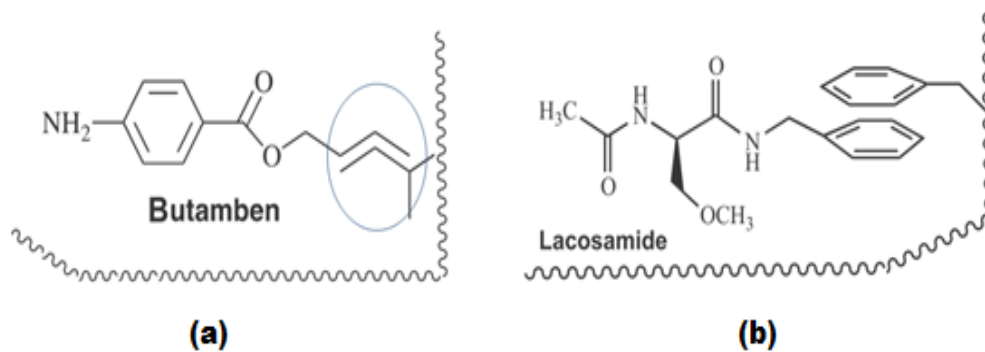


Figura A.2.8. Ejemplos de interacciones hidrofóbicas. (a) Grupos alquílicos; (b) Grupos arilos. La línea ondulada representa la cavidad del receptor.

Alternativamente, es posible una disposición en forma de T (interacción borde a cara), en la cual el borde de un anillo aromático forma una T con la cara del otro anillo aromático. Cuando un anillo es electrónicamente deficiente y el otro es rico en electrones, entonces los procesos de transferencia de carga se vuelven importantes también.

Las interacciones entre los anillos aromáticos están bien documentadas, el 60% de los residuos aromáticos están implicados en las interacciones aril-aril. Predominan la disposición de apilamiento en paralelo en forma de T y la disposición de apilamiento en paralelo desplazada (b y c en figura. A.2.9 [106]). Mientras que en las proteínas la disposición de apilamiento desplazado en paralelo se observa más a menudo, las dos disposiciones son de energía de interacción similar, con un ΔG^0 entre -1.6 a -2.4 kcal/mol, y se cree que es una combinación de dispersión van der Waals e interacción electrostática.

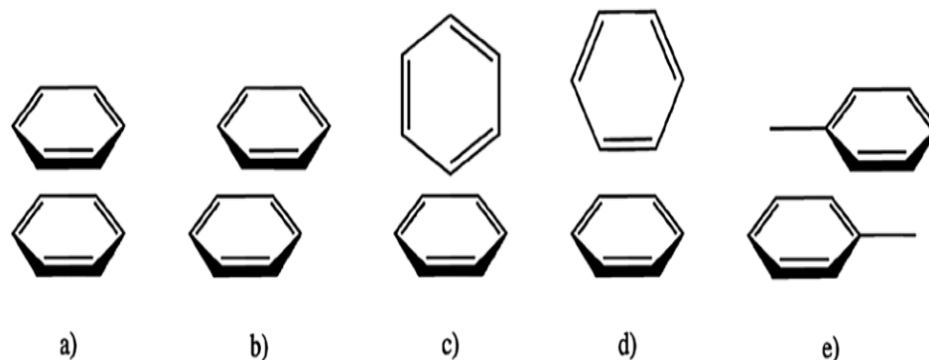


Figura A.2.9. Posibles disposiciones de apilamiento aromático. (a) paralelo centrado en la cara. (b) en paralelo desplazada. (c) forma de t perpendicular. (d) forma de y perpendicular. (e) compensación paralela a tolueno [105].

A.2.4.6. Interacción π - Cation.

Estas interacciones son muy comunes en la estructura de las proteínas y también pueden explotarse para las interacciones fármaco-receptor [107,108]. En proteínas, el grupo aromático más común involucrado en una interacción catiónica es el triptófano (aunque la fenilalanina, tirosina y la histidina también participan), y el catión más común es la arginina (aunque la

lisina también es importante) [109,110]. Un grupo catiónico en un fármaco puede experimentar una interacción catión- π con un grupo aromático en el receptor, o viceversa. La figura A.2.10 es un ejemplo de una interacción catiónica hipotética entre el ion amonio de **lisdexanfetamina** (Vyvans®), un fármaco para el trastorno por déficit de atención con hiperactividad y un residuo de triptófano. Las interacciones ΔG^0 para la interacción catión- π pueden estar entre, -0.5 y -7 kcal/mol, pero usualmente están en el rango de -1 a -5 kcal/mol.

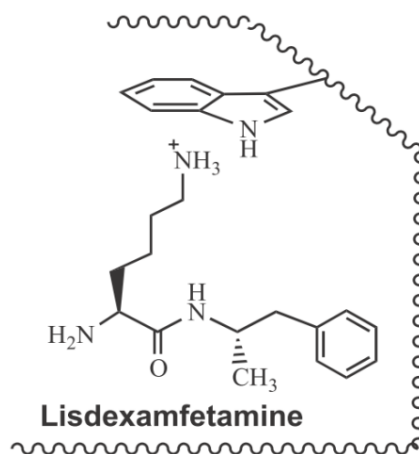


Figura A.2.10. Ejemplo de interacción catión- π . La línea ondulada representa la cavidad del receptor.

A.2.4.7. Interacciones Metalión - π

La interacción metalión - π es una interacción molecular no covalente entre la cara de un sistema π rico en electrones, con un catión adyacente (ver Figura. A.2.11), pero esta vez un metal (ej. K^+ , Na^+). La energía de la interacción catión - π es del mismo orden de magnitud del enlace de hidrógeno o del puente salino, y juega un rol importante en el reconocimiento molecular, en algunos casos, resulta ser la fuerza intermolecular decisiva. Algunas características del sistema de estudio en particular influyen en la fuerza del enlace, como por ejemplo: la naturaleza del catión, los sustituyentes en el sistema π , así como el solvente [111,112].

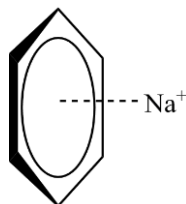


Figura A.2.11. Ejemplo de interacción metalión- π . Benceno interactuando con átomo de sodio.

A.2.4.8. Puente de Halógeno.

Ahora se ha establecido bien que un átomo de halógeno unido covalentemente puede actuar como un aceptor de electrones (ácido de Lewis) para someterse a un enlace de halógeno con un átomo donante rico en electrones, como O, N o S [113-116]. Sobre la base de estructura

cristalina y de datos basados en mecánica cuántica/mecánica molecular, se encontró que muchas de las distancias de enlace de halógeno a oxígeno (o nitrógeno) eran iguales o menores que la suma de los respectivos radios de van der Waals, lo que indica la formación de un enlace de halógeno (similar a un enlace de hidrógeno) [117,118]. La fuerza de estas interacciones está en el orden $H \approx I > Br > Cl \gg F$. La interacción es causada por la anisotropía de la densidad de electrones en el halógeno, que resulta de una región con carga positiva en la parte del lado posterior del átomo de halógeno a lo largo del eje de enlace R-X. Estas interacciones pueden gobernar la conformación de las moléculas en el sitio de unión de las proteínas. Se diseñaron una serie de inhibidores de la fosfodiesterasa 5 con átomos de flúor, cloro, bromo y yodo. Estos compuestos fueron incorporados en estudios sobre la fuerza relativa de enlaces de halógeno. Se analizaron dichos resultados observando que existe una buena correlación entre las energías de enlace calculadas y la actividad de las moléculas [119]. Las interacciones predichas entre el átomo de halógeno y el átomo de oxígeno fenólico de Tyr-612 fueron validados por cristalografía de rayos X, como se muestra en la figura A.2.12. El ΔG^0 para la unión halógena puede estar entre -1 y -15 kcal/mol [119,120], pero por lo general está en el rango -1 a -5 kcal/mol.

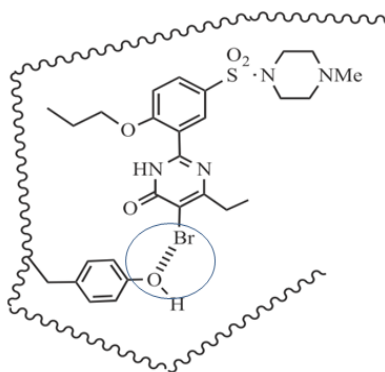


Figura A.2.12. Ejemplo de enlace halógeno. Un compuesto unido a la tirosina-612 de la fosfodiesterasa 5. La línea ondulada representa la cavidad de la enzima.

A.2.4.9. Van der Waals o fuerza de dispersión de London.

Los átomos en moléculas no polares pueden tener una distribución no simétrica temporal de la densidad de electrones, lo que da como resultado la generación de un dipolo temporal. A medida que átomos de diferentes moléculas (como un fármaco y un receptor) se aproximan, los dipolos temporales de una molécula inducen dipolos opuestos en la molécula que se aproxima. En consecuencia, resultan atracciones intermoleculares, conocidas como fuerzas de Van der Waals. Estas fuerzas universales débiles solo se vuelven significativas cuando hay un contacto superficial cercano de los átomos; sin embargo, cuando existe una complementariedad molecular, resultan numerosas interacciones atómicas (cada interacción contribuye aproximadamente -0,5 kcal/mol a la ΔG_0), lo que puede sumar un componente significativo de unión al receptor del fármaco. Otras interacciones débiles también pueden contribuir a la unión del ligando al receptor [121,122].

A.2.5. Conclusiones del apéndice.

Las fuerzas que predominan en la interacción fármaco-receptor son mayoritariamente interacciones no covalentes. Debido a que éstas son generalmente débiles, el cooperativismo por varios tipos de interacciones es crítico. Una vez que ha tenido lugar la primera interacción entre la droga y el receptor, se pierde la entropía asociada a las traslaciones del ligando. Esto resulta en una pérdida de entropía, mucho más baja en la formación de una segunda interacción. Con tantas de estas interacciones, simplemente mirar un modelo de una proteína con un ligando acoplado resultado de *Docking* puede ser engañoso, además de tener en cuenta que el estado ligado solo da parte de la información, por eso es preciso fomentar las simulaciones de MD en solvente explícito en este tipo de sistemas y otros de interés.

La de-solvatación, por ejemplo, juega un papel muy importante, particularmente con los grupos funcionales ionizados o polares donde habrá un gran término de de-solvatación desfavorable, además de los cambios entrópicos que tienen, como se mencionó anteriormente, influencia significativa. El efecto de este cooperativismo da como resultados que interacciones bastante débiles puedan combinarse para producir una interacción fuerte y, debido a que están involucrados varios tipos de interacciones diferentes, puede darse como resultado la selectividad en las interacciones fármaco-receptor. Un cuadro resumen de las interacciones no covalentes más frecuentes en macromoléculas para sistemas ligando-receptor presentes en el banco de datos de proteínas se presenta en el Apéndice A.3.7 ^[123].

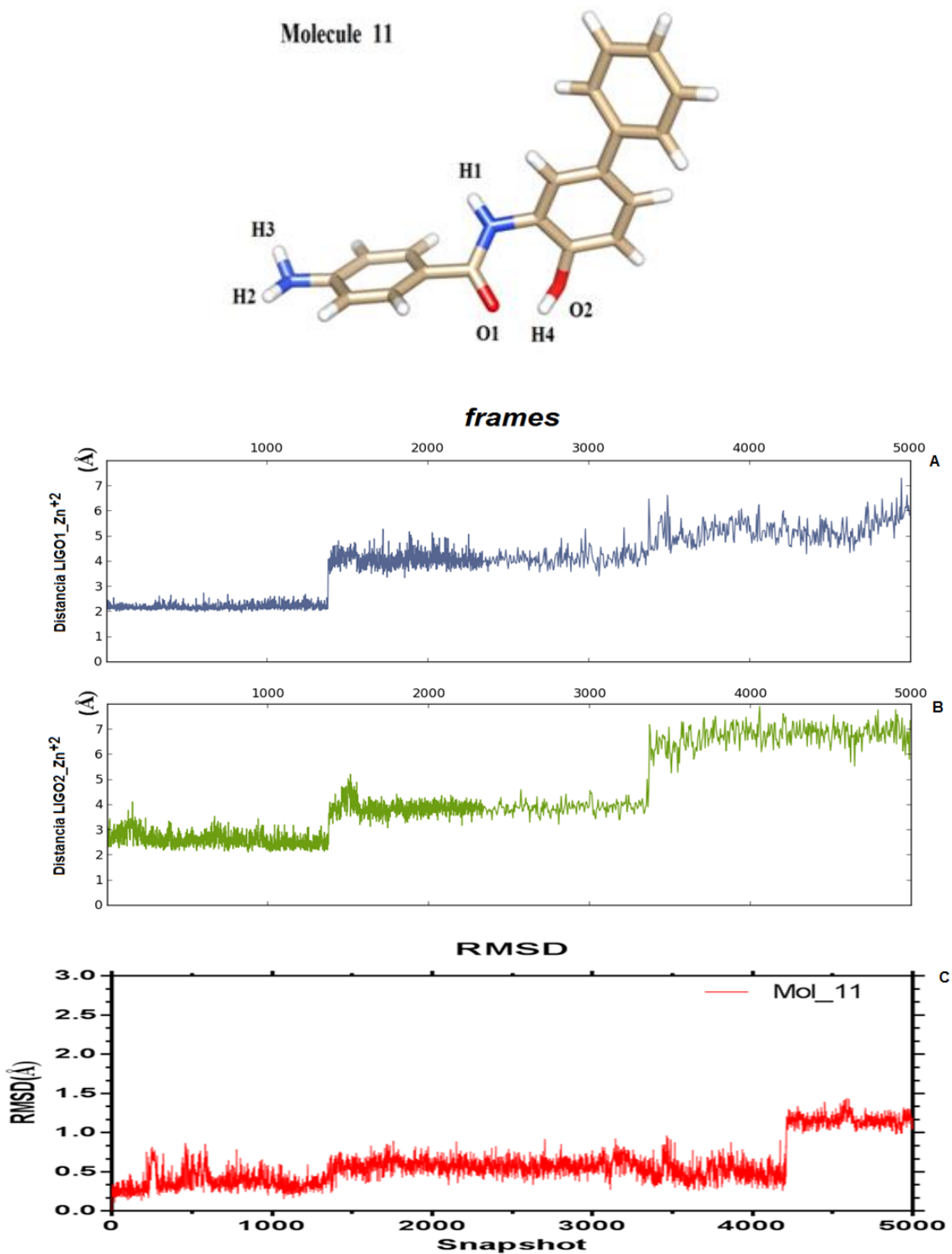
APÉNDICE A3: TABLAS Y GRÁFICOS.

Ponte a descubrir el “para hacer qué estás hecho”, y luego entrégate con pasión a hacerlo.

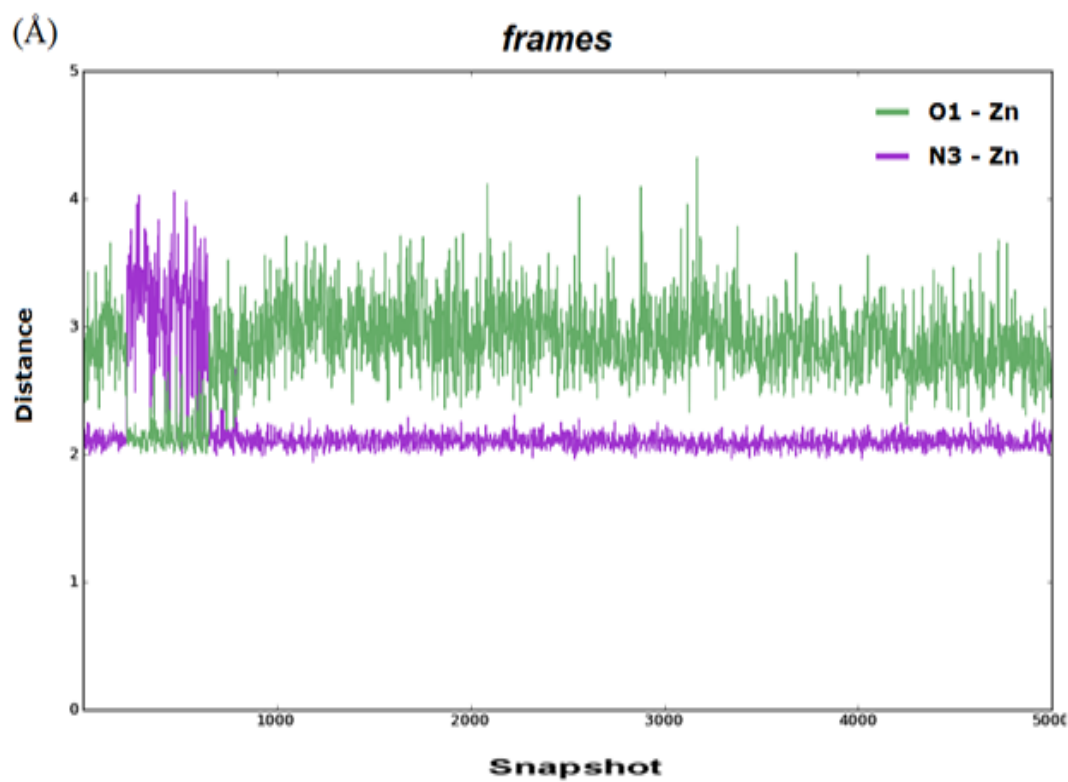
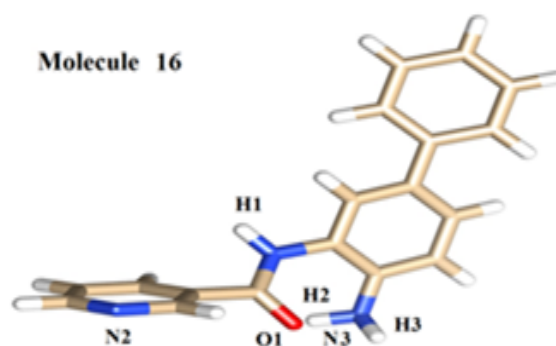
Martin Luther King Jr.

A.3. TABLAS Y GRÁFICOS.

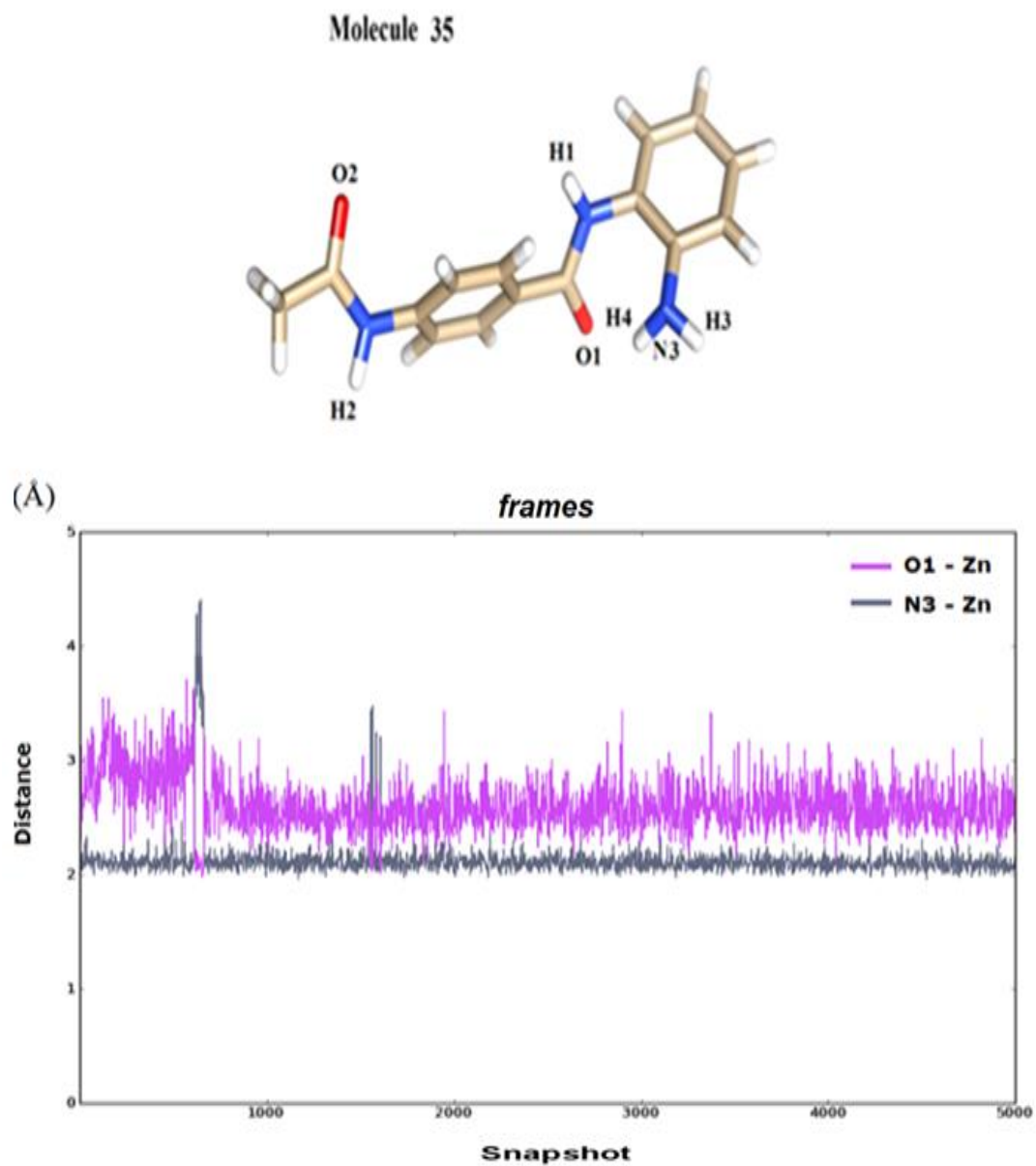
A.3.1. Distancia de entre los átomos de oxígeno O1-O2 y el átomo de Zn^{+2} , pertenecientes al compuesto 11, durante la simulación.



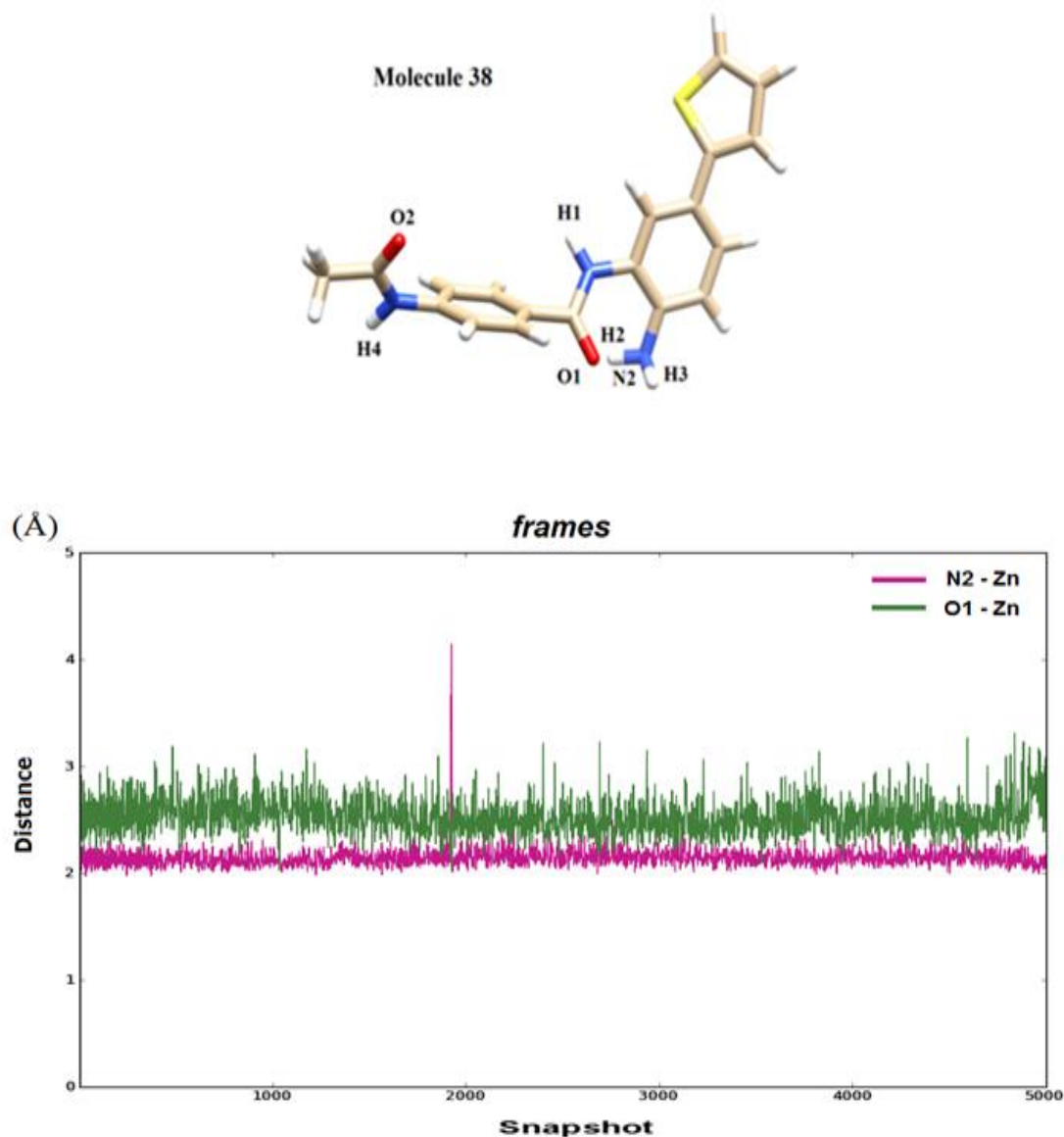
A.3.2. Distancia entre los átomos de oxígeno y nitrógeno O1-N3 pertenecientes al compuesto 16 y el átomo de Zn^{+2} , durante la simulación.



A.3.3. Distancia entre los átomos de oxígeno y nitrógeno O1-N3 pertenecientes al compuesto 35 y el átomo de Zn^{+2} , durante la simulación.



A.3.4. Distancia entre los átomos de oxígeno y nitrógeno O1-N2 pertenecientes al compuesto 35 y el átomo de Zn^{+2} , durante la simulación.



En resumen, en B, el RMSD del ligando (Mol 16) es de $0,29 \pm 0,07$ Å y la metaloenzima tiene un valor de $1,54 \pm 0,24$ Å, y en C, el RMSD del ligando (Mol 35) es de $0,66 \pm 0,17$ Å y la metaloenzima es de $1,98 \pm 0,52$ Å. En D, la amida terminal de las moléculas 38 del conjunto de datos, que contienen un enlace de rotación libre, permiten la existencia de isómeros conformacionales que dan un valor RMSD para el ligando de $0,75 \pm 0,20$ Å, pero para el metaloenzima es de $1,25 \pm 0,14$ Å.

A.3.4.1. Fórmulas utilizadas en cálculos de afinidad de unión aproximada, ambos enfoques:

$$\Delta G_{\text{bind}} = \Delta E_{\text{elec}} + \Delta E_{\text{vdw}} + {}^{\text{a,b}}\Delta G_{\text{solv.}}$$

$$\Delta G_{\text{solv}} = \Delta G_{\text{polar}} + \Delta G_{\text{np}} .$$

$${}^{\text{a}}\Delta G_{\text{solv,no-disp}} = \Delta G_{\text{PB}} + \Delta G_{\text{SAS(rep)}},$$

$${}^{\text{b}}\Delta G_{\text{solv,disp.}} = \Delta G_{\text{PB}} + \Delta G_{\text{SAS(rep)}} + \Delta G_{\text{disp(att)}},$$

Universidad Técnica Federico Santa María
Departamento de Química
A.3.6. Tablas de resultados

Tabla A.3.1. Búsqueda de los mejores modelos de CoMFA y CoMSIA usando combinaciones de campos para RG/X-ray asistido.

Modelo	q ²	N	SEP	SEE	r ² _{ncv}	F	Field Contributions					
							S	E	H	D	A	
CoMFA-S	0,645	4	0,542	0,237	0,932	75.204	1					
CoMFA-E	0,038	3	0,872	0,504	0,679	16.195		1				
CoMFA-SE	0,586	4	0,585	0,18	0,961	135.523	0,639	0,361				
CoMSIA-S	0,626	6	0,584	0,296	0,904	31.243	1					
CoMSIA-E	0,299	4	0,761	0,328	0,87	36.755		1				
CoMSIA-H	0,439	9	0,775	0,201	0,962	47.991			1			
CoMSIA-D	-0,026	1	0,864	0,755	0,217	6910,00					1	
CoMSIA-A	0,133	2	0,811	0,638	0,464	10368,00						1
CoMSIA-SE	0,635	11	0,666	0,046	0,998	773.431	0,493	0,507				
CoMSIA-SEH	0,513	4	0,635	0,225	0,939	84.671	0,306	0,319	0,376			
CoMSIA-SEHD	0,505	8	0,708	0,046	0,998	1.083.803	0,266	0,254	0,277	0,202		
CoMSIA-SEHA	0,570	5	0,61	0,18	0,963	108.171	0,214	0,264	0,297			0,226
CoMSIA-SED	0,483	20	1,252	0,016	1.000	3.651.851	0,365	0,348		0,287		
CoMSIA-SEA	0,501	5	0,658	0,197	0,955	89.439	0,304	0,387				0,309
CoMSIA-SEDA	0,462	20	1,277	0,019	1.000	2.575.136	0,287	0,283		0,206		0,224
CoMSIA-SH	0,569	5	0,611	0,227	0,94	66.317	0,436		0,565			
CoMSIA-SD	0,487	10	0,764	0,221	0,957	35.500	0,713			0,287		
CoMSIA-SA	0,508	10	0,748	0,145	0,981	84.400	0,684					0,316
CoMSIA-SHD	0,467	14	0,899	0,045	0,999	630.966	0,293		0,44	0,267		
CoMSIA-SHA	0,453	13	0,875	0,052	0,998	507.222	0,34		0,406			0,254
CoMSIA-SDA	0,503	18	1,063	0,024	1	1.706.125	0,554			0,211		0,235
CoMSIA-SHDA	0,364	15	1,026	0,038	0,999	847.493	0,303		1,722	0,215		0,162
CoMSIA-EH	0,532	6	0,653	0,156	0,973	121.274		0,498	0,502			
CoMSIA-ED	0,092	4	0,866	0,318	0,878	39.543		0,66		0,34		
CoMSIA-EA	0,448	7	0,727	0,158	0,974	101.783		0,555				0,445
CoMSIA-EHD	0,365	20	1,388	0,019	1.000	2.550.691		0,373	0,377	0,25		
CoMSIA-EHA	0,538	5	0,633	0,216	0,946	73.573		0,328	0,355			0,317
CoMSIA-EDA	0,345	20	1,409	0,024	1.000	1.622.875		0,421		0,273		0,306
CoMSIA-EHDA	0,419	5	0,71	0,205	0,952	82.438		0,279	0,31	0,149		0,263
CoMSIA-HD	0,242	13	1.030	0,121	0,99	94.810			0,64	0,36		
CoMSIA-HA	0,357	10	0,855	0,161	0,977	68.689			0,585			0,415
CoMSIA-HDA	0,236	5	0,813	0,311	0,888	33.370			0,491	0,204		0,305
CoMSIA-DA	0,101	2	0,826	0,619	0,494	11.719				0,631		0,369
CoMSIA-ALL	0,484	17	1,022	0,023	1.000	2.036.825	0,207	0,211	0,236	0,177		0,169

Tabla A.3.2. Actividades predichas y sus residuales de los modelos CoMFA-SE(RG).

Molécula	Experimental pIC ₅₀	Predicho pIC ₅₀	Residual
1	7,2366	7,1090	0,128
2	6,5086	6,6730	-0,164
3	7,1675	6,9780	0,190
4	7,1249	7,1220	0,003
5	6,8861	6,6630	0,223
6	6,8861	6,9350	-0,049
7	6,4559	6,9090	-0,453
8	6,8861	6,8940	-0,008
9	7	6,9980	0,002
10	6,3565	6,7090	-0,35
11	7,3979	6,5430	0,85
12	5,3188	6,9850	-1,67
13	6,2147	6,2930	-0,08
14	5,9208	5,9570	-0,04
15	5,585	5,6220	-0,04
16	7,3188	6,7970	0,52
17	7,1871	7,2000	-0,01
18	7,284	7,0730	0,21
19	6,7959	6,4410	0,35
20	6	5,9320	0,07
21	6,1871	6,3210	-0,13
22	6,1739	5,8460	0,33
23	5,3768	5,7360	-0,36
24	6,8539	6,4290	0,42
25	6,4815	6,6840	-0,20
26	5,2757	5,1830	0,09
27	6,6021	6,6600	-0,06
28	5,6778	5,6770	0,00
29	5,8861	6,1040	-0,22
30	4,8861	4,7920	0,09
31	5,8239	6,0170	-0,19
32	5,6198	5,5550	0,06
33	7,2218	6,6090	0,61
34	7,3188	7,1940	0,12
35	6,2441	6,3110	-0,07
36	7,7447	7,8220	-0,08
37	7,5528	7,9240	-0,37
38	8,1549	8,0830	0,07
39	8,1549	8,3330	-0,18

CoMFA-SE training set:

1,4,6,8,9,10,13,14,16,17,18,20,21,23,24,25,26,27,28,30,31,32,34,35,36,38,39.

test set: **2,3,5,7,11,12,15,19,22,29,33,37;***outliers:* **11,12,33.**

Tabla A.3.3. Tabla de valores de las actividades previstas y sus residuos obtenida a partir de los modelos CoMFA-SE y CoMSIA-SEHA alineamiento asistido por difracción rayos-X.

Molecul	Experimental pIC ₅₀	CoMFA		CoMSIA	
		Predicho pIC ₅₀	Residual	Predicho pIC ₅₀	Residual
1	7,2366	6,889	0,3476	6,743	0,4936
2	6,5086	6,41	0,0986	6,41	0,0986
3	7,1675	6,769	0,3985	7,117	0,0505
4	7,1249	7,163	-0,0381	7,143	-0,0181
5	6,8861	6,928	-0,0419	6,782	0,1041
6	6,8861	7,196	-0,3099	6,505	0,3811
7	6,4559	6,98	-0,5241	6,278	0,1779
8	6,8861	6,865	0,0211	6,974	-0,0879
9	7	6,994	0,006	6,991	0,009
10	6,3565	6,471	-0,1145	6,498	-0,1415
11	7,3979	7,003	0,3949	7,243	0,1549
12	5,3188	6,954	-1,6352	6,695	-1,3762
13	6,2147	6,274	-0,0593	6,176	0,0387
14	5,9208	6,92	-0,9992	6,431	-0,5102
15	5,585	5,186	0,399	5,522	0,063
16	7,3188	7,082	0,2368	7,143	0,1758
17	7,1871	7,189	-0,0019	7,224	-0,0369
18	7,284	7,074	0,21	7,024	0,26
19	6,7959	7,033	-0,2371	6,905	-0,1091
20	6	6,222	-0,222	6,029	-0,029
21	6,1871	6,795	-0,6079	5,863	0,3241
22	6,1739	5,954	0,2199	5,998	0,1759
23	5,3768	5,574	-0,1972	5,419	-0,0422
24	6,8539	6,884	-0,0301	6,619	0,2349
25	6,4815	6,79	-0,3085	6,208	0,2735
26	5,2757	5,236	0,0397	5,464	-0,1883
27	6,6021	6,675	-0,0729	6,736	-0,1339
28	5,6778	6,03	-0,3522	6,16	-0,4822
29	5,8861	5,607	0,2791	5,312	0,5741
30	4,8861	5,116	-0,2299	5,288	-0,4019
31	5,8239	5,604	0,2199	5,647	0,1769
32	5,6198	5,6	0,0198	5,695	-0,0752
33	7,2218	7,24	-0,0182	7,351	-0,1292
34	7,3188	7,377	-0,0582	7,498	-0,1792
35	6,2441	6,138	0,1061	6,222	0,0221
36	7,7447	7,689	0,0557	7,726	0,0187
37	7,5528	7,9	-0,3472	8,098	-0,5452
38	8,1549	8,14	0,0149	8,271	-0,1161
39	8,1549	8,407	-0,2521	8,319	-0,1641

CoMFA-SE y CoMSIA-SEHA

training set: 2,4,5,8,9,10,11,13,16,17,18,20,22,23,24,25,26,27,30,31,32,33,34,35,36,38,39;

test set: **1,3,6,7,12,14,15,19,21,28,29,37;**

CoMFA outliers:**12,14,21.**

CoMSIA outliers:**12,14.**

Tabla A.3.4. Interacciones no covalentes, entre átomos perteneciente a residuo que interacciona desde receptor y otro perteneciente al ligando. Distancia media, Desviación Estándar y Error estándar de la media. Receptor HDAC1 y Ligandos (11, 16, 35 y 38).

Complejo HDAC1/11								
Label	Interaction		Res Aminoacid	ProteinAtom	LigandAtom	Distance Mean	Std. Deviation	Std. Error of Mean
	Hydrophobic							
:1@C17,H14 :24@SD	Sulfur-Aromatic		MET	S	C,H aromático	4,4	0,44	0,0098
:1@C16 :24@CE	Aromatic-Alifatico		MET	C Aliph	C Aromático	4,0	0,39	0,0087
:1@C18 :24@CG	Aromatic-Alifatico		MET	C Aliph	C Aromático	4,1	0,43	0,0096
:1@C6 :135@CB	Aromatic-Alifatico		HIE	C Aliph	C Aromático	4,7	0,90	0,0202
:1@C18 :145@CA	Aromatic-Alifatico		CYS	C Aliph	C Aromático	4,2	0,36	0,0081
:1@C19 :145@CB	Aromatic-Alifatico		CYS	C Aliph	C Aromático	4,1	0,47	0,0106
:1@C18 :144@CB	Aromatic-Aromatic		PHE	C Aromático	C Aromático	4,5	0,48	0,0107
:1@C19 :144@CB	Aromatic-Aromatic		PHE	C Aromático	C Aromático	4,3	0,34	0,0076
π-Stacking								
:1@C9 :135@HE1	Aromatic-Aromatic		HIE	H	C Aromático	4,0	0,84	0,0188
:1@C9 :144@HE2	Aromatic-Aromatic		PHE	H	C Aromático	4,4	1,05	0,0235
:1@C10 :144@HZ	Aromatic-Aromatic		PHE	H	C Aromático	5,2	1,15	0,0256
:1@H8 :144@CE2	Aromatic-Aromatic		PHE	C Aromático	H	4,6	0,91	0,0204
:1@H9 :144@CZ	Aromatic-Aromatic		PHE	C Aromático	H	5,0	1,23	0,0274
Amidestacking								
:1@C7 :135@NE2	Aromatic-Aromatic		HIE	N arom	C amide	4,4	0,45	0,0101
Hydrogenbonding								
:1@H1 :135@ND1	N-H ... N		HIE	N arom	N-H	3,6	0,83	0,0185
Label	Indirectstabilization HB	Res Aminoacid	ProteinAtom	ProteinAtom	Distance Mean	Std. Deviation	Std. Error of Mean	
:135@HE2 :170@OD2	N-H ...O	HIE-ASP	N-H	O-	3,9	1,87	0,0419	
:135@HE2 :170@OD1	N-H ...O	HIE-ASP	N-H	O-	4,6	0,91	0,0204	

Complejo HDAC1/16							
Label	Interaction		ProteinAtom	LigandAtom	Distance Mean	Std. Deviation	Std. Error of Mean
	Hydrophobic(aromatic)*	Res Aminoacid					
:1@C16 :24@SD	C arom S (CH3)	MET	S	C	4,9	0,49	0,00686
:1@C17 :24@SD	C arom S (CH3)	MET	S	C	4,9	0,47	0,0067
:1@C16 :24@CE	Aromatic-Alifatico	MET	C	C	3,7	0,34	0,00481
:1@C17 :24@CG	Aromatic-Alifatico	MET	C	C	4,5	0,48	0,00675
Hydrophobic(stack)							
:1@C14 :133@CB	Aromatic-Alifatico	LEU	C	C	5,3	0,35	0,005
:1@C15 :133@CD1	Aromatic-Alifatico	LEU	C	C	4,5	1,03	0,0146
:1@C16 :133@CG	Aromatic-Alifatico	LEU	C	C	4,1	0,52	0,0074
:1@C17 :133@CD1	Aromatic-Alifatico	LEU	C	C	4,4	0,52	0,00742
:1@C4 :134@CB	Aromatic-Alifatico	HIE	C	C	3,7	0,62	0,00883
Hydrogenbonding							
:1@H2-134@ND1	N-H N	HIE	N-H	N	2,4	0,47	0,00659
π-Stacking							
:1@C10 :144@HE1	Aromatic-Aromatic	PHE	C Aromático	H	4,6	1,25	0,0177
:1@C11 :144@HE1	Aromatic-Aromatic	PHE	C Aromático	H	4,2	1,12	0,0158
:1@C9 :144@HE1	Aromatic-Aromatic	PHE	C Aromático	H	5,2	1,86	0,0263
:1@C3 :144@HE2	Aromatic-Aromatic	PHE	C Aromático	H	4,6	1,3	0,0183
:1@H4 :144@CE2	Aromatic-Aromatic	PHE	H	C Aromático	3,8	1,13	0,016
Amidestacking							
:1@C7-144@CZ	π -- C=O	PHE	C Aromático	C=O (amide)	4,8	1,18	0,0166
:1@N1-144@HZ	C-H ... N	PHE	H Aromatic	N	5,2	1,48	0,0209
:1@H1-144@CZ	N-H ... π	PHE	C Aromático	H	4,6	0,96	0,0136
Complejo HDAC1/35							
Label	Interaction		ProteinAtom	LigandAtom	Distance Mean	Std. Deviation	Std. Error of Mean
	Hydrophobic(stack)	Res Aminoacid					
:1@C11 :135@HD2	Aromatic-Alifatico	HIE	H (N-H)	C arom	3,9	0,97	0,0138

Programa de Magíster en Ciencias mención Química

:1@C15 :135@CB	Aromatic-Alifatico	HIE	C aliph	C arom	4,9	0,58	0,0082
	N-H ...π (HB)*						
:1@H1 :135@ND1	N-H ... π	HIE	N-H	N arom	5,0	0,98	0,0138
:1@H3 :135@NE2	N-H ... π	HIE	N-H	N arom	3,4	0,40	0,00565
:1@H4 :135@NE2	N-H ... π	HIE	N-H	N arom	3,1	0,79	0,0112
	π-Stacking						
:1@H7_ :144@CE1	Aromatic-Aromatic	PHE	C arom	C-H arom	3,3	0,57	0,00802
:1@H8_ :144@CD2	Aromatic-Aromatic	PHE	C arom	C-H arom	3,4	0,48	0,00671
:1@C2_ :144@HE2	Aromatic-Aromatic	PHE	C-H arom	C arom	3,8	0,50	0,00712
:1@C5_ :144@HZ	Aromatic-Aromatic	PHE	C-H arom	C arom	3,7	0,46	0,00653
	Amidestacking						
:1@H1 - :144@CD2	N-H ... π	PHE	C arom	H , (N-H amide)	3,6	0,47	0,00658
	Weakhydrogenbonding						
:1@N1 - :144@HD2	C-H ... N	PHE	C-H arom	N (amide)	3,9	0,50	0,00706
	Weakhydrogenbonding						
:1@O2 - :144@HZ	C-H ... O	PHE	C-H arom	O (amide)	4,2	1,23	0,0175
:1@N2 - :144@HZ	C-H ... N	PHE	C-H arom	N (amide)	4,1	0,62	0,00874
	Hydrophobic(stack)						
:1@C3 - :265@CD1	Aromatic-Alifatico	LEU	C aliph	C arom	4,1	0,51	0,00715
:1@C4 - :265@CD2	Aromatic-Alifatico	LEU	C aliph	C arom	4,1	0,51	0,00714
:1@C13 :294@CA	Aromatic-Alifatico	GLY	C aliph	C arom	3,9	0,26	0,00372
	Amidestacking						
:1@C13 - :294@C	π -C=O	GLY	C=O	C arom	3,6	0,44	0,00623
	Weakhydrogenbonding						
:1@H13 - :294@O	C-H --- O	GLY	O (amide)	C-H arom	3,3	0,41	0,00577
:1@H14 - :294@O	C-H --- O	GLY	O (amide)	C-H arom	3,6	0,22	0,00311
Label	Indirectstabilization HB	Res Aminoacid	ProteinAtom	ProteinAtom	Distance Mean	Std. Deviation	Std. Error of Mean
:297@OH :28@HH12	H-O ... H	TYR-ARG	OH	H (N-H)	3,2	0,48	0,0068
:297@OH :28@HH22	H-O ... H	TYR-ARG	OH	H (N-H)	3,3	0,55	0,0077
:294@O :28@HH12	O ... H-N	GLY-ARG	O (amide)	H (N-H)	2,4	0,75	0,0106

Universidad Técnica Federico Santa María
Departamento de Química

:294@O :28@HH22	O ... H-N	GLY-ARG	O (amide)	H (N-H)	2,3	0,42	0,0059
:295@O :28@HH11	O ... H-N	TYR-ARG	O (amide)	H (N-H)	3,4	0,91	0,0129
:295@O :28@HH12	O ... H-N	TYR-ARG	O (amide)	H (N-H)	3,2	0,83	0,0118
Hydrogenbonding							
:172@N :135@HE2	N-H...N	HIN-HIE	N	N-H	4,1	1	0,0142
Complejo HDAC1/38							
Interaction							
Label	Hydrophobic	Res	ProteinAtom	LigandAtom	Distance Mean	Std. Deviation	Std. Error of Mean
:24@SD - :1@C7	sulfur-aromatic	MET	S	C arom	4,5	0,59	0,0084
:24@SD - :1@C12	sulfur-aromatic	MET	S	C arom	5,0	0,45	0,0064
:1@S1 - :24@CG	sulfur-Alifatico	MET	C aliph	S	3,9	0,44	0,0062
:1@C7 - :24@CE	Aromatic-Alifatico	MET	C aliph	C arom	4,0	0,41	0,0058
:1@C19 - :265@CD1	Alifatico-Aromatic	LEU	C aliph	C aliph	5,3	1,26	0,0178
:1@C19 - :265@CD2	Alifatico-Aromatic	LEU	C aliph	C aliph	5,3	0,97	0,0137
:1@C16 - :265@CD1	Aromatic-Alifatico	LEU	C aliph	C arom	5,6	1,32	0,0186
:1@C16 - :265@CD2	Aromatic-Alifatico	LEU	C aliph	C arom	5,5	1,25	0,0177
:1@C7 - :24@CG	Aromatic-Alifatico	MET	C aliph	C arom	4,5	0,40	0,0057
:1@C17 - :24@CE	Aromatic-Alifatico	MET	C aliph	C arom	4,4	0,66	0,0094
:1@C3 - :295@CA	Aromatic-Alifatico	GLY	C aliph	C arom	3,8	0,28	0,0040
:1@C4 - :295@CA	Aromatic-Alifatico	GLY	C aliph	C arom	3,7	0,23	0,0032
Weakhydrogenbonding							
:1@O1 - :24@HE1	C-H --- O	MET	C-H	O (amide)	3,8	0,63	0,0090
:1@O1 - :24@HE2	C-H --- O	MET	C-H	O (amide)	3,8	0,64	0,0090
:1@O1 - :24@HE3	C-H --- O	MET	C-H	O (amide)	3,8	0,65	0,0092
:1@N1 - :24@HE1	C-H --- N	MET	C-H	N (amide)	3,9	0,73	0,0103
:1@O2 - :144@HZ	C-H --- O	PHE	C-H arom	O (amide)	4,1	1,58	0,0223
:1@H6 - :295@N	C-H --- N	GLY	N (N-H)	H arom	3,6	0,26	0,0036
:1@O2 - :297@HE2	C-H --- O	TYR	C-H arom	O (amide)	3,7	1,21	0,0172
:1@H10 - :24@O	(arom)C-H...O S/ π	MET	O (amide)	H arom	3,8	0,47	0,0067

Programa de Magíster en Ciencias mención Química

:1@C1 - :24@SD	S --- C=O π-Stacking	MET	S	C=O (amide)	4,9	0,54	0,0077
:1@C4 - :135@HD2	Aromatic-Aromatic	HIE	C-H arom	C arom	3,9	1,27	0,0179
:1@H2 - :135@CE1	Aromatic-Aromatic	HIE	C (arom)	H (N-H)	3,7	0,52	0,0074
:1@C7 - :144@CB	Aromatic-Alifatico	PHE	C arom	C arom	4,7	0,35	0,0050
:1@H7 - :144@HB2	Aromatic-Alifatico	PHE	C-H arom	C-H arom	2,7	0,35	0,0049
:1@C16 - :297@HD2	Aromatic-Aromatic	TYR	C (arom)	C (arom)	3,7	0,45	0,0064
:1@C17 - :297@HD2	Aromatic-Aromatic	TYR	C (arom)	C (arom)	3,7	0,51	0,0071
:1@C16 - :297@HE2	Aromatic-Aromatic	TYR	C (arom)	C (arom)	4,0	0,56	0,0079
Hydrogenbonding/π-stack							
:1@H3 - :135@NE2	N-H --- π	HIE	N arom	H (N-H)	3,2	1,02	0,0144
:1@H2 - :135@ND1	N-H --- π	HIE	N arom	H (N-H)	3,8	0,74	0,0104
:1@C4 - :295@H	N-H --- π	GLY	H (N-H)	C arom	3,4	0,49	0,0069
Amide π-stacking							
:1@N3 - :297@HE2	C-H --- N	TYR	N (N-H)	C-H arom	3,8	0,51	0,0073
:1@C18 - :297@OH	O --- C=O	TYR	O (O-H)	C=O (amide)	4,2	0,70	0,0099

Tabla A.3.5. Ocupación de enlace de hidrógeno entre los residuos HDAC1 y compuestos 11, 16, 35 y 38.

11				
DONOR :res@atom	ACCEPTOR :res@atom	%ocupa	d [Å]	α [°]
:Asp170@OD1	:LIG12@H4	78,70	2,791 (0,13)	69,07 (30,29)
:Asp170@OD2	:LIG12@H4	58,20	2,696 (0,13)	47,50 (24,56)
:LIG12@O1	:Tyr297@HH	19,05	2,803 (0,11)	18,80 (12,98)
:Asp258@OD1	:LIG12@H4	6,25	2,766 (0,11)	37,84 (14,03)
16				
DONOR :res@atom	ACCEPTOR :res@atom	%ocupa	d [Å]	α [°]
:Asp170@OD2	:LIG17@H3	48,00	2,812 (0,11)	146,57 (11,72)
:Asp170@OD2	:LIG17@H2	48,00	2,812 (0,11)	65,73 (20,64)
:Hie134@ND1	:LIG17@H3	24,02	2,920 (0,06)	98,71 (21,51)
:Hie134@ND1	:LIG17@H2	24,02	2,920 (0,06)	41,71 (26,19)
:Gly143@O	:LIG17@H1	17,14	2,907 (0,07)	31,08 (8,31)
:LIG17@O1	:Tyr297@HH	14,72	2,807 (0,12)	31,90 (27,85)
:Asp170@OD1	:LIG17@H3	14,02	2,939 (0,05)	128,37(9,40)
:Asp170@OD1	:LIG17@H2	14,02	2,939 (0,05)	81,61 (6,58)
35				
DONOR :res@atom	ACCEPTOR :res@atom	%ocupa	d [Å]	α [°]
:Asp170@OD2	:LIG35@H4	95,22	2,779 (0,09)	30,94 (11,13)
:Asp170@OD2	:LIG35@H3	95,22	2,779 (0,09)	134,08 (6,44)
:Asp170@OD1	:LIG35@H4	45,72	2,914 (0,06)	80,27 (6,46)
:Asp170@OD1	:LIG35@H3	45,72	2,914 (0,06)	140,72 (8,06)
:Gly143@O	:LIG35@H1	14,18	2,917 (0,06)	30,86 (7,96)
38				
DONOR :res@atom	ACCEPTOR :res@atom	%ocupa	d [Å]	α [°]
:Asp170@OD2	:LIG38@H3	38,64	2,817 (0,11)	60,85 (21,75)
:Asp170@OD2	:LIG38@H2	38,64	2,817 (0,11)	150,78 (11,01)
:Gly143@O	:LIG38@H1	35,10	2,897 (0,07)	36,15 (8,84)
:Asp170@OD1	:LIG38@H3	11,50	2,939 (0,05)	80,80 (7,52)
:Asp170@OD1	:LIG38@H2	11,50	2,939 (0,05)	133,62 (5,77)
:Hie135@ND1	:LIG38@H3	5,66	2,915 (0,06)	114,17 (11,43)
:Hie135@ND1	:LIG38@H2	5,66	2,915 (0,06)	27,21 (13,43)

Tabla A.3.6. Ocupación por enlace de hidrógeno entre los átomos del solvente y átomos hidrófilos de los compuestos 11, 16, 35 y 38. Ocupación de los patrones HB para los residuos del sitio activo HDAC1 presentes en diferentes complejos (HDC1/11/16/35/38).

11				
DONOR :res@atom	ACCEPTOR :res@atom	%ocupa	d [Å]	α [°]
solventdonor	:LIG@H3	72,75	2,891 (0,07)	89,93 (48,29)
solventdonor	:LIG@H2	72,75	2,891 (0,07)	86,44 (48,54)
solventdonor	:LIG@H4	24,30	2,838 (0,09)	106,92 (13,36)
:LIG@O1	solventacceptor	54,60	2,766 (0,13)	77,25 (49,99)
:LIG@O2	solventacceptor	48,60	2,838 (0,09)	90,83 (34,65)
:Hin172@O	solventacceptor	244,90	2,787 (0,12)	69,79 (47,04)
:Hin172@NE2	solventacceptor	228,10	2,808 (0,10)	67,80 (49,33)
:Gly294@O	solventacceptor	127,50	2,807 (0,12)	72,63 (47,62)
:Hie135@ND1	solventacceptor	67,00	2,877 (0,09)	71,14 (44,82)
:Asp170@OD1	solventacceptor	22,80	2,836 (0,12)	114,07 (29,78)
solventdonor	:Tyr297@HH	14,55	2,840 (0,10)	30,30 (31,19)
:Asp258@OD1	solventacceptor	9,50	2,886 (0,08)	130,22 (37,97)
:Asp258@OD2	solventacceptor	5,50	2,901 (0,09)	118,94 (38,71)
16				
DONOR :res@atom	ACCEPTOR :res@atom	%ocupa	d [Å]	α [°]
:LIG@N2	solventacceptor	139,84	2,852 (0,09)	70,59 (48,48)
solventdonor	:LIG@H3	34,90	2,863 (0,08)	19,33 (9,39)
solventdonor	:LIG@H2	34,90	2,863 (0,08)	115,22 (7,53)
solventdonor	:LIG@H1	9,26	2,926 (0,06)	98,71 (11,39)
:LIG@O1	solventacceptor	5,84	2,785 (0,12)	71,84 (47,69)
:Hin172@NE2	solventacceptor	244,12	2,813 (0,10)	67,96 (49,35)
:Hin172@O	solventacceptor	144,24	2,798 (0,12)	69,84 (46,66)
:Hie135@ND1	solventacceptor	66,72	2,884 (0,08)	71,40 (48,12)
:Asp258@OD2	solventacceptor	17,56	2,751 (0,12)	69,81 (45,84)
solventdonor	:Hie135@HE2	10,26	2,901 (0,07)	48,56 (29,54)

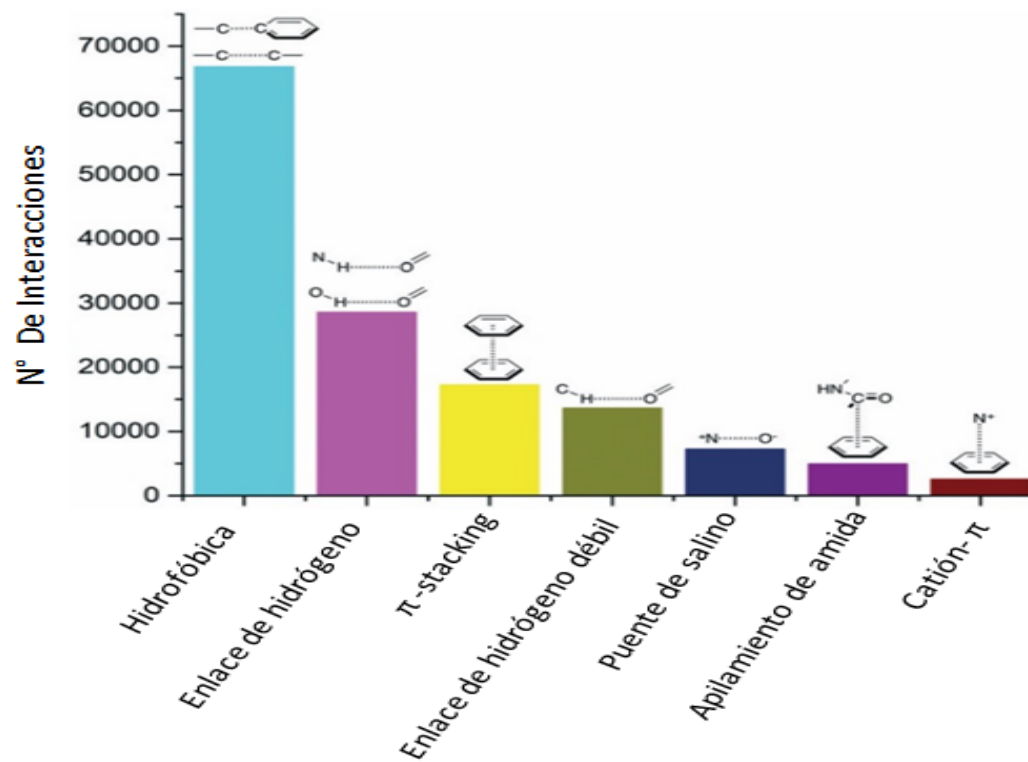
Universidad Técnica Federico Santa María
Departamento de Química

35				
DONOR :res@atom	ACCEPTOR :res@atom	%ocupa	d [Å]	α [°]
:LIG@O2	solventacceptor	233,64	2,791 (0,12)	72,60 (48,19)
solventdonor	:LIG@H4	46,74	2,875 (0,08)	117,44 (6,62)
solventdonor	:LIG@H3	46,74	2,875 (0,08)	17,77 (8,66)
solventdonor	:LIG@H2	29,52	2,907 (0,07)	26,79 (20,02)
:Hin172@NE2	solventacceptor	318,96	2,813 (0,10)	68,02 (50,65)
:Hin172@O	solventacceptor	246,20	2,784 (0,12)	73,14 (48,02)
:Hie135@ND1	solventacceptor	107,64	2,884 (0,08)	73,55 (47,93)
solventdonor	:Arg28@HH22	36,74	2,870 (0,08)	117,04 (29,02)
solventdonor	:Arg28@HH21	36,74	2,870 (0,08)	46,09 (31,11)
:Asp170@OD2	solventacceptor	26,64	2,727 (0,12)	69,57 (51,04)
solventdonor	:Arg28@HH12	5,82	2,874 (0,08)	63,73 (44,59)
solventdonor	:Arg28@HH11	5,82	2,874 (0,08)	97,79 (51,83)
solventdonor	:Hie135@HE2	5,74	2,907 (0,07)	48,28 (33,28)
38				
DONOR:res@atom	ACCEPTOR :res@atom	%ocupa	d [Å]	α [°]
:LIG@O2	solventacceptor	192,60	2,797 (0,12)	73,10 (48,80)
solventdonor	:LIG@H3	67,24	2,858 (0,08)	119,68 (8,26)
solventdonor	:LIG@H2	67,24	2,858 (0,08)	17,56 (9,33)
solventdonor	:LIG@H4	15,28	2,909 (0,06)	34,32 (20,26)
solventdonor	:LIG@H1	6,62	2,933 (0,05)	104,23 (9,29)
:Hin172@NE2	solventacceptor	342,00	2,834 (0,10)	68,83 (50,51)
:Tyr297@OH	solventacceptor	188,96	2,828 (0,11)	120,28 (36,02)
:Hie135@ND1	solventacceptor	94,64	2,882 (0,08)	72,00 (47,70)
solventdonor	:Tyr297@HH	94,48	2,828 (0,11)	46,89 (46,63)
solventdonor	:Arg28@HH12	64,24	2,863 (0,08)	126,98 (5,41)
solventdonor	:Arg28@HH11	64,24	2,863 (0,08)	20,90 (10,48)
solventdonor	:Arg28@HH22	43,14	2,860 (0,08)	126,47 (12,91)
solventdonor	:Arg28@HH21	43,14	2,860 (0,08)	30,25 (14,09)
:Gly295@O	solventacceptor	26,04	2,785 (0,11)	75,98 (50,27)
solventdonor	:Hie135@HE2	5,24	2,887 (0,08)	40,21 (27,56)

Tabla A.3.7. Algoritmo de búsqueda y función de puntuación del programa AutoDock.

Software	Algoritmo de búsqueda (<i>Search algorithm</i>)	Función de puntuación (<i>Scoring function</i>)
AutoDock	Lamarckian Genetic Algorithm (LGA)	$\Delta G = \Delta G_{vdw} \sum_{ij} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) + \Delta G_{nbond} \sum_{ij} E(t) \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right) + \Delta G_{elec} \sum_{ij} \frac{q_i q_j}{\epsilon(r_{ij}) r_{ij}} + \Delta G_{tor} N_{tor} + \Delta G_{solv} \sum_{ij} (S_i V_j + S_j V_i) e^{(-r_{ij}^2 / 2\sigma^2)}$

A.3.7. Distribución de frecuencias de las interacciones no covalentes más comunes observadas en complejos proteína-ligando extraídos de la base de datos de proteínas PDB.



REFERENCIAS

- [1] Lund, A. H., & van Lohuizen, M. (2004). Epigenetics and cancer. *Genes & development*, 18(19), 2315-2335.
- [2] Gregoretto IV, Lee YM, Goodson HV. Molecular evolution of the histone deacetylase family: functional implications of phylogenetic analysis. *J. Mol. Biol.* 338(1), 17-31 (2004).
- [3] Kouzarides T. Chromatin modifications and their functions. *Cell* 128 (4), 693-705 (2007).
- [4] Baylin SB, Ohm JE. Epigenetic gene silencing in cancer-mechanism for early oncogenic pathway addiction? *Nat. Rev. Cancer* 6(2), 107-116 (2006).
- [5] Bolden, J. E., Peart, M. J., & Johnstone, R. W. (2006). Anticancer activities of histone deacetylase inhibitors. *Nature reviews Drug discovery*, 5(9), 769.
- [6] Manal, M., Chandrasekar, M. J. N., Priya, J. G., & Nanjan, M. J. (2016). Inhibitors of histone deacetylase as antitumor agents: a critical review. *Bioorganic chemistry*, 67, 18-42.
- [7] Mottamal, M., Zheng, S., Huang, T. L., & Wang, G. (2015). Histone deacetylase inhibitors in clinical studies as templates for new anticancer agents. *Molecules*, 20(3), 3898-3941.
- [8] Dickinson, M., Johnstone, R. W., & Prince, H. M. (2010). Histone deacetylase inhibitors: potential targets responsible for their anti-cancer effect. *Investigational new drugs*, 28(1), 3-20.
- [9] Gryder, B. E., Sodji, Q. H., & Oyelere, A. K. (2012). Targeted cancer therapy: giving histone deacetylase inhibitors all they need to succeed. *Future medicinal chemistry*, 4(4), 505-524.
- [10] Methot, J. L., Chakravarty, P. K., Chenard, M., Close, J., Cruz, J. C., Dahlberg, W. K., ... & Harsch, A. (2008). Exploration of the internal cavity of histone deacetylase (HDAC) with selective HDAC1/HDAC2 inhibitors (SHI-1: 2). *Bioorg Med Chem Lett*, 18, 973-978.
- [11] Muddassar, M., et al. "Receptor guided 3D-QSAR: a useful approach for designing of IGF-1R inhibitors." *Bio Med Research International* 2008 (2008).
- [12] Abdizadeh, T., Ghodsi, R., & Hadizadeh, F. (2017). 3D-QSAR (CoMFA, CoMSIA) and Molecular Docking Studies on Histone Deacetylase 1 Selective Inhibitors. *Recent patents on anti-cancer drug discovery*, 12(4), 365-383.
- [13] Millard, C. J., Watson, P. J., Celardo, I., Gordiyenko, Y., Cowley, S. M., Robinson, C. V., & Schwabe, J. W. (2013). Class I HDACs share a common mechanism of regulation by inositol phosphates. *Molecular cell*, 51(1), 57-67.
- [14] Frisch, M. J. E. A., Trucks, G. W., Schlegel, H., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., ... & Millam, J. M. (2008). *Gaussian 03, revision C. 02*.

- [15] Mannhold, R., Krogsgaard-Larsen, P., & Timmerman, H. (2008). QSAR: Hansch analysis and related approaches (Vol. 1). John Wiley & Sons.
- [16] Mella, J., Villegas, F., Morales-Verdejo, C., Lagos, C. F., & Recabarren-Gajardo, G. (2017). Structure-activity relationships studies on weakly basic N-arylsulfonylindoles with an antagonistic profile in the 5-HT₆ receptor. *Journal of Molecular Structure*, 1139, 362-370.
- [17] Merz Jr, K. M., Ringe, D., & Reynolds, C. H. (Eds.). (2010). *Drug design: structure-and ligand-based approaches*. Cambridge University Press.
- [18] López-Lira, C., Alzate-Morales, J. H., Paulino, M., Mella-Raipaán, J., Salas, C. O., Tapia, R.T., & Soto-Delgado, J. (2018). Combined molecular modelling and 3D-QSAR study for understanding the inhibition of NQO1 by heterocyclic quinone derivatives. *Chemical Biology & Drug Design*, 91(1), 29–38.
- [19] Morris, G. M., Goodsell, D. S., Pique, M. E., Lindstrom, W., Huey, R., Forli, S., & Olson, A. J. (2009). AutoDock Version 4.2; Updated for Version 4.2. 6.
- [20] Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., & Olson, A. J. (2009). AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of computational chemistry*, 30(16), 2785-2791.
- [21] 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.
- [22] Trott, O., & Olson, A. J. (2010). AutoDockVina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry*, 31(2), 455-461.
- [23] Seeliger, D., & de Groot, B. L. (2010). Ligand docking and binding site analysis with PyMOL and Autodock/Vina. *Journal of computer-aided molecular design*, 24(5), 417-422.
- [24] Schrödinger, L. L. C. (2015). PyMOL: The PyMOL Molecular Graphics System, Version 1.7. 6.6. Schrödinger LLC, NY, USA. URL <https://www.pymol.org>.
- [25] Patel, B. D., & Ghate, M. D. (2015). 3D-QSAR studies of dipeptidyl peptidase-4 inhibitors using various alignment methods. *Medicinal Chemistry Research*, 24(3), 1060-1069.
- [26] Muddassar, M., Pasha, F. A., Chung, H. W., Yoo, K. H., Oh, C. H., & Cho, S. J. (2008). Receptor guided 3D-QSAR: a useful approach for designing of IGF-1R inhibitors. *BioMed Research International*, 2008.
- [27] Balasubramanian, P. K., Balupuri, A., Kang, H. Y., & Cho, S. J. (2017). Receptor-guided 3D-QSAR studies, molecular dynamics simulation and free energy calculations of Btk kinase inhibitors. *BMC systems biology*, 11(2), 6.
- [28] Tripos International (1699). SYBYL-X 1.2. South Hanley Rd., St. Louis, Missouri, 63144, USA.

- [29] Andrades, J., Campanini, J., Vásquez, D., Silvestri, C., Morales, C., Romero, J., & Mella, J. (2015). A combined CoMFA and CoMSIA 3D-QSAR study of benzamide type antibacterial inhibitors of the FtsZ protein in drug-resistant *Staphylococcus aureus*. SAR and QSAR in Environmental Research, 26(11), 925-942.
- [30] Golbraikh, A.; Tropsha, A. Beware of q^2 ! J. Mol. Graph. Model. 2002,20, 269-276.
- [31] Tropsha, A. Best Practices for QSAR Model Development, Validation, and Exploitation. Mol. Inform. 2010,29, 476-488.
- [32] Roy, K.; Chakraborty, P.; Mitra, I.; Ojha, P. K.; Kar, S.; Das, R. N. Some case studies on application of " $r(m)^2$ " metrics for judging quality of quantitative structure-activity relationship predictions: emphasis on scaling of response data. J. Comput. Chem. 2013,34, 1071-1082.
- [33] Case, D. A., Cheatham, T. E., 3rd, Darden, T., Gohlke, H., Luo, R., Merz, K. M., Jr., Onufriev, A., Simmerling, C., Wang, B. & Woods, R. J. (2005). The Amber biomolecular simulation programs. Journal of Computational chemistry 26(16), 1668-1688.
- [34] Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A. & Simmerling, C. (2006). Comparison of multiple Amber force fields and development of improved protein backbone parameters. Proteins: Structure, Function, and Bioinformatics, 65(3), 712-725.
- [35] Wang, J., Wang, W., Kollman P. A.; Case, D. A. (2006). Automatic atom type and bond type perception in molecular mechanical calculations. Journal of molecular graphics and modelling, 25(2), 247-260.
- [36] Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., & Case, D. A. (2004). Development and testing of general amber force field. Journal of computational chemistry, 25(9), 1157-1174.
- [37] Cornell, W. D., Cieplak, P., Bayly, C. I., & Kollmann, P. A. (1993). Application of RESP charges to calculate conformational energies, hydrogen bond energies, and free energies of solvation. Journal of the American Chemical Society, 115(21), 9620-9631.
- [38] Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. The Journal of chemical physics 79(2), 926-935.
- [39] Pang, Y. P., Xu, K. U. N., El Yazal, J., & Prendergast, F. G. (2000). Successful molecular dynamics simulation of the zinc-bound farnesyltransferase using the cationic dummy atom approach. Protein Science, 9(10), 1857-1865.
- [40] Pang, Y.P. (1999). Novel zinc protein molecular dynamics simulations: Steps toward antiangiogenesis for cancer treatment. Molecular modeling annual 5(10), 196-202.
- [41] Van Gunsteren, W. F. & Berendsen, H. J. C. (1977). Algorithms for macromolecular dynamics and constraint dynamics. Molecular Physics 34(5), 1311-1327.

- [42] Loncharich, R. J., Brooks, B. R., & Pastor, R. W. (1992). Langevin dynamics of peptides: The frictional dependence of isomerization rates of N-acetylalanyl-N'-methylamide. *Biopolymers*, 32(5), 523-535.
- [43] Darden, T., York, D., & Pedersen, L. (1993). Particle mesh Ewald: An N · log (N) method for Ewald sums in large systems. *The Journal of chemical physics*, 98(12), 10089-10092.
- [44] Roe, D. R., & Cheatham III, T. E. (2013). PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *Journal of chemical theory and computation*, 9(7), 3084-3095.
- [45] Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD: visual molecular dynamics. *Journal of molecular graphics*, 14(1), 33-38.
- [46] 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.
- [47] Kollman, P. A., Massova, I., Reyes, C., Kuhn, B., Huo, S., Chong, L., ... & Donini, O. (2000). Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Accounts of chemical research*, 33(12), 889-897.
- [48] Reddy MR, Erion MD (eds) (2001) cálculos de energía libre en el diseño racional de fármacos. Kluwer / PlenumPress, Nueva York
- [49] Srivastava HK, Sastry GN (2012) MD investigation on a series of HIV protease inhibitors: assessing the performance of MM-PBSA and MM-GBSA approaches. *J Chem Inf Model* 52:3088–3098
- [50] Homeyer, N., & Gohlke, H. (2013). FEW: a workflow tool for free energy calculations of ligand binding. *Journal of computational chemistry*, 34(11), 965-973. Computer aided drug desing.
- [51] Liu, H., & Hou, T. (2016). CaFE: a tool for binding affinity prediction using end-point free energy methods. *Bioinformatics*, 32(14), 2216-2218.
- [52] Miller III, B. R., McGee Jr, T. D., Swails, J. M., Homeyer, N., Gohlke, H., & Roitberg, A. E. (2012). MMPBSA.py: an efficient program for end-state free energy calculations. *Journal of chemical theory and computation*, 8(9), 3314-3321.
- [53] Gentile, F., Deriu, M. A., Licandro, G., Prunotto, A., Danani, A., & Tuszynski, J. A. (2015). Structure based modeling of small molecules binding to the TLR7 by atomistic level simulations. *Molecules*, 20(5), 8316-8340.
- [54] Wolf, R. M., (2012). Amberlite: Some AMBER-tools Based Utilities. *AmberTools12: Reference Manual* (pp. 124 - 125). Basle, Switzerland: Novartis Institutes for Biomedical Research (NIBR).

- [55] Tan, C., Tan, Y. H., & Luo, R. (2007). Implicit nonpolar solvent models. *The Journal of Physical Chemistry B*, 111(42), 12263-12274.
- [56] Wang, C., Nguyen, P. H., Pham, K., Huynh, D., Le, T. B. N., Wang, H., ...&Luo, R. (2016). Calculating protein–ligand binding affinities with MMPBSA: Method and error analysis. *Journal of computational chemistry*, 37(27), 2436-2446.
- [57] Gohlke, H. & Case, D. A. (2004). Converging free energy estimates: MM-PB(GB)SA studies on the protein-protein complex Ras-Raf. *Journal of Computational chemistry*, 25(2), 238-250.
- [58] Kar, P., Lipowsky, R., & Knecht, V. (2013). Importance of polar solvation and configurational entropy for design of antiretroviral drugs targeting HIV-1 protease. *The Journal of Physical Chemistry B*, 117(19), 5793-5805.
- [59] Sixto-López, Y., Bello, M., Rodríguez-Fonseca, R. A., Rosales-Hernández, M. C., Martínez-Archundia, M., Gómez-Vidal, J. A., & Correa-Basurto, J. (2017). Searching the conformational complexity and binding properties of HDAC6 through docking and molecular dynamic simulations. *Journal of Biomolecular Structure and Dynamics*, 35(13), 2794-2814.
- [60] Wang, C., Greene, D. A., Xiao, L., Qi, R., &Luo, R. (2017). Recent Developments and Applications of the MMPBSA Method. *Frontiers in molecular biosciences*, 4, 87.
- [61] Valley, C. C., Cembran, A., Perlmutter, J. D., Lewis, A. K., Labello, N. P., Gao, J., & Sachs, J. N. (2012). The methionine-aromatic motif plays a unique role in stabilizing protein structure. *Journal of Biological Chemistry*, 287(42), 34979-34991.
- [62] Wilcken, R., Zimmermann, M. O., Lange, A., Zahn, S., Kirchner, B., & Boeckler, F. M. (2011). Addressing methionine in molecular design through directed sulfur–halogen bonds. *Journal of chemical theory and computation*, 7(7), 2307-2315.
- [63] Cramer CJ (2004) *Fundamentos de Química Computacional: teorías y modelos*, segunda ed. Wiley, Chichester
- [64] *Molecular Modelling for Beginners – 2° Edition* Wiley, Alan Hinchliffe
- [65] *Introduction to Computational Chemistry – 2° Edition* Wiley, Frank Jensen.
- [66] Levine IN (2001) *química cuántica*, 5ª edición. Prentice Hall, Reino Unido
- [67] Badrinarayan, P., Choudhury, C., & Sastry, G. N. (2015). Molecular modeling. In *Systems and Synthetic Biology* (pp. 93-128). Springer, Dordrecht. Chapter 6.
- [68] Frenkel D, B Smit (2002) *Comprendiendo las simulaciones moleculares: desde los algoritmos a las aplicaciones*, 2ª ed. vol 1, Serie Ciencia Computacional Academic Press, San Diego
- [69] Rapaport DC (2004) *The art of MD simulation*, 2nd edn. Cambridge University Press, New York

- [70] McQuarrie, DA (2000). *Statistical Mechanics* (revised 2nd ed.) University Science Books. pp. 91-128. ISBN 978-1-891389-15-3.
- [71] Leach, Andrew R.; Harren, Jhoti (2007). *Structure-based Drug Discovery*. Berlin: Springer. ISBN 1-4020-4406-2.
- [72] Fuente: <https://www.nature.com/subjects/structure-based-drug-design>.
- [73] Anderson, A. C. (2003). The process of structure-based drug design. *Chemistry & biology*, 10(9), 787-797.
- [74] Stahl M, Rarey M (2001) Detailed analysis of scoring functions for virtual screening. *J Med Chem*44:1035-1042.
- [75] Young, D. C. (2009). *Computational drug design: a guide for computational and medicinal chemists*. John Wiley & Sons.
- [76] Cross JB, Thompson DC, Rai BK, Baber JC, Fan KY, Hu Y, Humblet C (2009) Comparison of molecular docking programs: pose prediction and virtual screening accuracy. *J Chem Inf Model* 49:1455-1474
- [77] Cherkasov, A., Muratov, E. N., Fourches, D., Varnek, A., Baskin, I. I., Cronin, M., ... & Consonni, V. (2014). QSAR modeling: where have you been? Where are you going to? *Journal of medicinal chemistry*, 57(12), 4977-5010.
- [78] Cramer, R. D. (2012). The inevitable QSAR renaissance. *Journal of computer-aided molecular design*, 26(1), 35-38.
- [79] Hansch, C., Maloney, P. P., Fujita, T., & Muir, R. M. (1962). Correlation of biological activity of phenoxyacetic acids with Hammett substituent constants and partition coefficients. *Nature*, 194(4824), 178.
- [80] Todeschini, R., Consonni, V., & Mannhold, R. (2000). *Methods and principles in medicinal chemistry*. Kubinyi H, Timmerman H (Series eds) *Handbook of molecular descriptors*. Wiley-VCH, Weinheim.
- [81] Helguera, A. M., Combes, R. D., González, M. P., & Cordeiro, M. N. D. S. (2008). Applications of 2D descriptors in drug design: a DRAGON tale. *Current topics in medicinal chemistry*, 8(18), 1628-1655.
- [82] Wold, S., Ruhe, A., Wold, H., & Dunn, III, W. J. (1984). The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses. *SIAM Journal on Scientific and Statistical Computing*, 5(3), 735-743.
- [83] Wold, S., Martens, H., & Wold, H. (1983). The multivariate calibration problem in chemistry solved by the PLS method. *Matrix pencils*, 286-293.
- [84] Wold, S. (1991). Validation of QSAR's. *Molecular Informatics*, 10(3), 191-193.

- [85] Wermuth, C. G., Ganellin, C. R., Lindberg, P., & Mitscher, L. A. (1998). Glossary of terms used in medicinal chemistry (IUPAC Recommendations 1998). *Pure and Applied Chemistry*, 70(5), 1129-1143.
- [86] Wermuth, C. G. (2006). Pharmacophores: historical perspective and viewpoint from a medicinal chemist. *Methods and Principles in Medicinal Chemistry*, 32, 3.
- [87] Güner, O. F. (Ed.). (2000). Pharmacophore perception, development, and use in drug design (Vol. 2). Internat'l University Line.
- [88] Bissantz, C., Kuhn, B., & Stahl, M. (2010). A medicinal chemist's guide to molecular interactions. *Journal of medicinal chemistry*, 53(14), 5061-5084.
- [89] Silverman, R. B., & Holladay, M. W. (2014). *The organic chemistry of drug design and drug action*. Academic press
- [90] Dill, K. A. (1997). Additivity principles in biochemistry. *Journal of Biological Chemistry*, 272(2), 701-704.
- [91] Levitt, M., & Park, B. H. (1993). Water: now you see it, now you don't. *Structure*, 1(4), 223-226.
- [92] Fischer, E. Ber. (1894). *Dt. Chem. Ges.* 27, 2985-2993.
- [93] Pierce, A. C., terHaar, E., Binch, H. M., Kay, D. P., Patel, S. R., & Li, P. (2005). CH•••O and CH•••N Hydrogen Bonds in Ligand Design: A Novel Quinazolin-4-ylthiazol-2-ylamine Protein Kinase Inhibitor. *Journal of medicinal chemistry*, 48(4), 1278-1281.
- [94] Cleland, W. W., & Kreevoy, M. M. (1994). Low-barrier hydrogen bonds and enzymatic catalysis. *SCIENCE-NEW YORK THEN WASHINGTON-*, 1887-1887.
- [95] Kuhn, B., Mohr, P., & Stahl, M. (2010). Intramolecular hydrogen bonding in medicinal chemistry. *Journal of medicinal chemistry*, 53(6), 2601-2611.
- [96] Menear, K. A., Adcock, C., Alonso, F. C., Blackburn, K., Copsey, L., Drzewiecki, J., ...& Lence, C. F. (2008). Novel alkoxybenzamide inhibitors of poly (ADP-ribose) polymerase. *Bioorganic& medicinal chemistry letters*, 18(14), 3942-3945.
- [97] Furet, P., Caravatti, G., Guagnano, V., Lang, M., Meyer, T., & Schoepfer, J. (2008). Entry into a new class of protein kinase inhibitors by pseudo ring design. *Bioorganic & medicinal chemistry letters*, 18(3), 897-900.
- [98] Jansma, A.; Zhang, Q.; Li, B.; Ding, Q.; Uno, T.; Bursulaya, B; Liu, Y.; Furet, P.; Gray, N. S.; Geierstanger, B. H. Verification of a designed intramolecular hydrogen bond in a drug scaffold by nuclear magnetic resonance spectroscopy. *J. Med. Chem.* 2007, 50, 5875-5877.
- [99] Jencks, W. P. (1987). *Catalysis in chemistry and enzymology*. Courier Corporation.
- [100] Leung, C. S., Leung, S. S., Tirado-Rives, J., & Jorgensen, W. L. (2012). Methyl effects on protein-ligand binding. *Journal of medicinal chemistry*, 55(9), 4489-4500.

- [101] Burley, S. K., & Petsko, G. A. (1985). Aromatic-aromatic interaction: a mechanism of protein structure stabilization. *Science*, 229, 23-29.
- [102] Salonen, L. M., Ellermann, M., & Diederich, F. (2011). Aromatic rings in chemical and biological recognition: energetics and structures. *Angewandte Chemie International Edition*, 50(21), 4808-4842
- [103] Raju, R. K., Bloom, J. W., An, Y., & Wheeler, S. E. (2011). Substituent effects on non-covalent interactions with aromatic rings: Insights from computational chemistry. *Chem Phys Chem*, 12(17), 3116-3130.
- [104] McGaughey, G. B., Gagné, M., & Rappé, A. K. (1998). π -Stacking interactions alive and well in proteins. *Journal of Biological Chemistry*, 273(25), 15458-15463.
- [105] Tsuzuki, S., Honda, K., Uchimaru, T., Mikami, M., & Tanabe, K. (2002). Origin of attraction and directionality of the π/π interaction: model chemistry calculations of benzene dimer interaction. *Journal of the American Chemical Society*, 124(1), 104-112.
- [106] Martinez, C. R., & Iverson, B. L. (2012). Rethinking the term "pi-stacking". *Chemical Science*, 3(7), 2191-2201.
- [107] Mecozzi, S., West, A. P., & Dougherty, D. A. (1996). Cation- π interactions in aromatics of biological and medicinal interest: electrostatic potential surfaces as a useful qualitative guide. *Proceedings of the National Academy of Sciences*, 93(20), 10566-10571.
- [108] Gallivan, J. P., & Dougherty, D. A. (1999). Cation- π interactions in structural biology. *Proceedings of the National Academy of Sciences*, 96(17), 9459-9464.
- [109] Ma, J. C., & Dougherty, D. A. (1997). The cation- π interaction. *Chemical reviews*, 97(5), 1303-1324.
- [110] Frontera, A., Quinonero, D., & Deya, P. M. (2011). Cation- π and anion- π interactions. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 1(3), 440-459.
- [111] Reddy, A. S., & Sastry, G. N. (2005). Cation [$M = H^+, Li^+, Na^+, K^+, Ca^{2+}, Mg^{2+}, NH_4^+$, and NMe_4^+] interaction with the aromatic motifs of naturally occurring amino acids: a theoretical study. *The Journal of Physical Chemistry A*, 109(39), 8893-8903.
- [112] Reddy, A. S., Zipse, H., & Sastry, G. N. (2007). Cation- π interactions of bare and coordinatively saturated metal ions: contrasting structural and energetic characteristics. *The Journal of Physical Chemistry B*, 111(39), 11546-11553.
- [113] 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.
- [114] Auffinger, P., Hays, F. A., Westhof, E., & Ho, P. S. (2004). Halogen bonds in biological molecules. *Proceedings of the National Academy of Sciences of the United States of America*, 101(48), 16789-16794.

- [115] Ibrahim, M. A. (2011). Molecular mechanical study of halogen bonding in drug discovery. *Journal of computational chemistry*, 32(12), 2564-2574.
- [116] Hardegger, L. A., Kuhn, B., Spinnler, B., Anselm, L., Ecabert, R., Stihle, M., ...& Plancher, J. M. (2011). Systematic investigation of halogen bonding in protein–ligand interactions. *Angewandte Chemie International Edition*, 50(1), 314-318.
- [117] 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.
- [118] Lu, Y., Shi, T., Wang, Y., Yang, H., Yan, X., Luo, X., ...& Zhu, W. (2009). Halogen bonding a novel interaction for rational drug design? *Journal of medicinal chemistry*, 52(9), 2854-2862.
- [119] Xu, Z., Liu, Z., Chen, T., Chen, T., Wang, Z., Tian, G., ...& Wang, G. (2011). Utilization of halogen bond in lead optimization: a case study of rational design of potent phosphodiesterase type 5 (PDE5) inhibitors. *Journal of medicinal chemistry*, 54(15), 5607-5611.
- [120] Lu, Y. X., Zou, J. W., Wang, Y. H., Jiang, Y. J., & Yu, Q. S. (2007). Ab initio investigation of the complexes between bromobenzene and several electron donors: some insights into the magnitude and nature of halogen bonding interactions. *The Journal of Physical Chemistry A*, 111(42), 10781-10788
- [121] Paulini, R., Müller, K., & Diederich, F. (2005). Orthogonal multipolar interactions in structural chemistry and biology. *Angewandte Chemie International Edition*, 44(12), 1788-1805.
- [122] Bartlett, G. J., Newberry, R. W., VanVeller, B., Raines, R. T., & Woolfson, D. N. (2013). Interplay of hydrogen bonds and $n \rightarrow \pi^*$ interactions in proteins. *Journal of the American Chemical Society*, 135(49), 18682-18688.
- [123] de Freitas, R. F., & Schapira, M. (2017). A systematic analysis of atomic protein–ligand interactions in the PDB. *Med Chem Comm*, 8(10), 1970-1981.

