



Universidad de Concepción
Facultad de Ciencias Químicas
Doctorado en Ciencias con Mención en Química

**CÁLCULOS DE ENERGÍA LIBRE DE UNIÓN DE SISTEMAS
ANFITRIÓN-HUÉSPED Y PROTEÍNA-LIGANDO UTILIZANDO
CARGAS ATÓMICAS DERIVADAS DE LA DENSIDAD
ELECTRÓNICA**

Tesis para optar al grado de Doctor en Ciencias con mención en Química

Por: Duván Esteban González Gómez

Profesor Guía: Dr. Esteban Vöhringer-Martinez

Julio 2022
Concepción, Chile

Contenido

Lista de Figuras	VI
Lista de Tablas	XIV
Resumen	XXIII
1. Introducción	1
1.1. Marco Teórico	7
1.2. Sistemas anfitrión-huésped	17
1.2.1. Anfitriones OA/TEMOA y huéspedes del SAMPL5 .	20
1.2.2. Anfitrión CB7 y huéspedes con núcleo aromático o de adamantano	21
1.3. Sistemas proteína-ligando	24
1.3.1. Lisozima T4 L99A/M102Q y ligandos con núcleo aromático	26
1.4. Hipótesis y Objetivos	28
2. Metodología	30
2.1. Métodos de estructura electrónica	30

2.2.	Mecánica molecular	38
2.2.1.	Parámetros enlazantes	41
2.2.2.	Parámetros no enlazantes	43
2.2.2.1.	Cargas atómicas	45
2.2.2.2.	Parámetros de Lennard-Jones	56
2.3.	Dinámica molecular	60
2.4.	Método Monte Carlo vía cadenas de Markov	61
2.4.1.	Movimientos de desplazamiento y rotación de Me- tropolis Monte Carlo	63
2.5.	Muestreadores	64
2.6.	Ciclo termodinámico	66
2.7.	Cálculos de energía libre	70
2.7.1.	Método de Zwanzing	71
2.7.2.	Integración termodinámica	72
2.7.3.	Bennett Acceptance Ratio (BAR)	73
2.7.4.	Multistate Bennett acceptance ratio (MBAR)	74
2.8.	Corrección de polarización a la energía libre de unión	76
3.	Resultados y Discusión	79
3.1.	Validación de parámetros de simulación utilizados para el cálculo de energía libre de unión	80
3.1.1.	Sistema anfitrión-huésped	80

3.1.2. Sistema proteína-ligando	89
3.2. Validación de parámetros de simulación utilizados para el cálculo de las cargas D-MBIS y energía de polarización pro- medio	93
3.3. Energías libres de unión entre los anfitriones OA/TEMOA y huéspedes del SAMPL5	99
3.4. Energías libres de unión entre CB7 y huéspedes con núcleo aromático	111
3.5. Energías libres de unión entre CB7 y huéspedes con núcleo de adamantano	115
3.6. Energías libres de unión entre la Lisozima T4 doble mutante L99A/M102 y ligandos con núcleo aromático	127
4. Conclusiones	138
Bibliografía	145
A. Anexo: Sistemas OA/TEMOA unidos a huéspedes G1-6	160
B. Anexo: Sistemas CB7 unidos a huéspedes BG1-G6 y AG1-8	163
C. Anexo: Sistemas Lisozima T4 doble mutante L99A/M102Q unido a los ligandos L1-8	175

Lista de Figuras

- 1.1. Enfoques generales para el cálculo de energía libre de unión en los sistemas receptor-ligando. Arriba: Energía libre de unión absoluta, considera la aniquilación del ligando en el ambiente de unión y no unión. La representación en verde del ligando indica que las interacciones electrostáticas y van-der-Waals se encuentran activas y en azul indica que dichas interacciones se han desactivado. Abajo: Energía libre de unión relativa, considera modificaciones *alquímicas* para transformar estructuralmente el ligando A al ligando B, tanto en el ambiente de unión como no unión. La representación en verde indica que las interacciones electrostáticas y van-der-Waals se encuentran activas en ambos ligandos. 10
- 1.2. Estructuras de los anfitriones OA-TEMOA y huéspedes del SAMPL5. Se muestran los microestados más probables de las moléculas anfitrión-huésped de acuerdo al pH experimental y valores de pK_a [1] (ver Tabla A.1 en el Anexo A). Las energías libres de unión de todos los sistemas fueron medidas bajo las mismas condiciones experimentales: pH 11.5, Temperatura de 298 K y disolución reguladora de pH de fosfato de sodio 50 mM. 21
- 1.3. Estructuras del anfitrión CB7 y huéspedes con núcleo aromático y adamantano. Se muestran los microestados más probables de los huéspedes de acuerdo al pH experimental 4.74 y valores de pK_a (ver Tabla B.1 en el Anexo B). Las energías libres de unión de todos los sistemas fueron medidas bajo las mismas condiciones experimentales: pH 4.7, Temperatura de 298 K y solución reguladora de pH de acetato de sodio 50 mM. 23

1.4.	Izquierda -Complejo de lisozima T4 L99A/M102Q-fenol. Se muestra los residuos GLN102, VAL111 y la superficie del ligando. Derecha -Ligandos con los estados de protonación más probable de acuerdo al pH experimental 6.8 y valores de pK_a (ver Tabla C.1 en el Anexo C). Las energías libres de unión de todos los sistemas fueron medidas bajo las mismas condiciones experimentales pH 6.8, Temperatura de 298 K y solución reguladora de pH de fosfato de potasio 50 mM o cloruro de potasio 200 mM.	27
2.1.	Diagrama de flujo utilizado para calcular las cargas atómicas D-MBIS para diferentes configuraciones del huésped junto con el costo de polarización energética del huésped en el estado de unión (complejo) y en solución acuosa (solvente). Los promedios de las energías de polarización corrigen la energía libre calculada por YANK para la aniquilación del huésped en los dos ambientes moleculares.	55
2.2.	Diagrama de flujo computacional utilizado para derivar los parámetros de Lennard-Jones para los ligandos basados en la partición MBIS de la densidad electrónica y considerando un volumen efectivo de un átomo en una molécula y el volumen respectivo del átomo aislado. Se obtuvieron los parámetros del potencial de Lennard-Jones σ y ϵ para todos los átomos en cada huésped y el promedio de los coeficientes para cada tipo de átomo según el campo de fuerza GAFF.	59
2.3.	Estrategias de muestreo. Arriba: intercambio de réplicas hamiltonianas (HREX). Abajo: muestreador con mezclas autoajustable (SAMS). Esquema elaborado por Mey <i>et al.</i> [2] . . .	65

2.4.	Ciclo termodinámico para el cálculo de energía libre de unión absoluta empleando la metodología de doble aniquilación para el receptor en estado de unión y no unión. Esquema adaptado de Lee <i>et al.</i> [3]. La parte superior de la Figura muestra el desacoplamiento para el huésped en estado de unión. Este paso considera una restricción para mantener al huésped en el sitio de unión mientras las interacciones electrostáticas y de van-der-Waals están desactivadas. El desacoplamiento de las interacciones electrostáticas y de van-der-Waals para el huésped en el estado no enlazado se muestra en la parte inferior de la Figura	67
3.1.	Afinidades de unión para el sistema TEMOA-G1 obtenidas con YANK a diferente número de iteraciones utilizando los muestreadores SAMS (izquierda) y HREX (derecha). Se compara un número diferente de estados termodinámicos intermedios (λ) con los determinados automáticamente por el algoritmo "thermodynamic trailblazing" implementado en YANK (60 y 56 estados intermedios para la fase compleja y acuosa, respectivamente).	81
3.2.	Estructura tridimensional de TEMOA y G1 en diferentes conformaciones iniciales para el huésped. La representación en átomos de carbono púrpura y cian corresponden a las conformaciones obtenidas del repositorio de Github (Moble-Lab/benchmarksets [4]) y con la herramienta OMEGA respectivamente. Las energías libres obtenidas con estas dos conformaciones iniciales se reportan en la Tabla 3.1	86

3.3.	Validación de los parámetros de simulación para obtener convergencia en las cargas atómicas MBIS y en la energía de polarización promedio del esquema de trabajo mostrado en el Figura 2.1. La energía de polarización promedio del huésped G1 en el anfitrión TEMOA se muestra para tres trayectorias independientes como función del tiempo total de simulación usando cargas atómicas AM1-BCC (sin actualización de topología) (izquierda), para un número variable de actualizaciones de carga de la topología del huésped con cinco ns de muestreo para cada actualización (centro), y usando cinco actualizaciones y 25 ns de muestreo y variando el radio de corte que define las cargas atómicas para los cálculos de QM/MM que rodean al huésped (derecha). Para los tres parámetros, la convergencia de la energía de polarización se logra en los valores utilizados en el flujo de trabajo de simulación.	96
3.4.	Estructura para los huéspedes del SAMPL5 (G1-6) unidos a los anfitriones OA y TEMOA, nombre y microestado más probable de acuerdo a los valores de pK_a (ver Tabla A.1 en el Anexo A).	100
3.5.	Energía requerida para polarizar los huéspedes en los diferentes ambientes moleculares. Las barras verdes y púrpuras consideran cargas atómicas RESP para los anfitriones OA/TEMOA y D-MBIS para cada uno de los huéspedes. Las barras grises representan el estado de no unión en solución acuosa (los valores corresponden al promedio de los valores esperados para la energía de polarización mostrados en la Tabla 3.3).	108
3.6.	Estructura para los huéspedes con núcleo aromático (BG1-6) unidos al anfitrión CB7, nombre y microestado más probable de acuerdo a los valores de pK_a (ver Tabla B.1 en el Anexo B).	111
3.7.	Estructura para los huéspedes con núcleo de adamantano (AG1-8) unidos al anfitrión CB7, nombre y microestado más probable de acuerdo a los valores de pK_a (ver Tabla B.1 en el Anexo B).	116

- 3.8. Análisis de perturbación adaptado a la simetría (SAPT) de las interacciones en huéspedes con núcleo aromático y adamantano en el nivel SAPT0/jun-ccVDZ. Izquierda: Comparación de las interacciones entre huéspedes con núcleo aromático y de adamantano de una conformación representativa en solución acuosa, incluidas las moléculas de agua que solvatan el grupo de amonio cargado. Los átomos del huésped y las moléculas de agua se colorean según su contribución a la energía de interacción. Las barras en el medio muestran la contribución electrostática (Elec), dispersión (Disp), inducción del anfitrión (A) al huésped (B) (IndAB) e intercambio (Inter) de la energía de interacción total expresada en kcal/mol. Los gráficos circulares de la parte inferior muestran la contribución de cada grupo funcional a la energía de interacción total en porcentaje y la energía de interacción en kcal/mol. Derecha: Las mismas representaciones para dos huéspedes con núcleo de adamantano resaltan la gran contribución de intercambio repulsivo de AG1 y la contribución repulsiva de sus dos grupos metilo. 118
- 3.9. Error absoluto en la afinidad de unión para los huéspedes de núcleo adamantano usando diferentes cargas atómicas y parámetros de Lennard-Jones (LJ). Cargas AM1-BCC/GAFF LJ (barras en verde), cargas D-MBIS / GAFF LJ (barras en gris) y cargas D-MBIS /D-MBIS LJ (barras en púrpura). Se tomaron los parámetros de enlace del campo de fuerza GAFF para el CB7 y huéspedes AG1-G8. 123
- 3.10. Parámetros de Lennard-Jones derivados considerando la partición MBIS de la densidad electrónica polarizada, se muestran los promedios para cada tipo de átomo de los huéspedes con núcleo de adamantano. El radio derivado MBIS al mínimo del potencial se compara con los valores del campo de fuerza GAFF para diferentes tipos de átomos a la izquierda y la profundidad derivada MBIS del potencial a la derecha. . . 125

3.11. Estructura para los ligandos de núcleo aromático (L1-8) unidos a la proteína lisozima T4 L99A/M102Q, nombre y microestado más probable de acuerdo a los valores de pK_a (ver Tabla C.1 en el Anexo C).	128
3.12. Comparación de estructuras APO (rojo) y HOLO (azul) para los ligandos L3 y L6. Se muestra la hélice F para la estructura APO (verde), en el caso del ligando más voluminoso L6 se aprecia mayores diferencias en la conformación de la hélice F. Los ligando L3 y L6 se muestran en amarillo.	136
A.1. Definición del ángulo CH-HH-CG utilizado en las siguientes figuras para analizar la conformación del huésped en el bolsillo de unión (donde CH es el átomo de carbono de la parte inferior del anfitrión, HH es un átomo de hidrógeno de la parte superior del anfitrión, y CG es el átomo de carbono del ión carboxilato del huésped). El ángulo medido CH-HH-CG representa la conformación relativa del grupo funcional principal en el bolsillo de unión. Átomos de oxígeno en rojo, átomos de carbono en gris (excepto los átomos de carbono del huésped en verde), átomos de hidrógeno en blanco y átomos de referencia CH-HH-CG en amarillo.	160
A.2. Izquierda: Gráficos de distribución de ángulos en los sistemas OA-G1/TEMOA-G1 (arriba) y OA-G4/TEMOA-G4 (abajo) usando cargas atómicas RESP/AM1-BCC para el anfitrión-huésped respectivamente. Derecha: Conformaciones representativas para los sistemas OA-G1/TEMOA-G1 (arriba) y OA-G4/TEMOA-G4 (abajo) (los grupos metilo de TEMOA están resaltados en verde y los dos huéspedes coloreados según la distribución que se muestra a la izquierda. Átomos de oxígeno en rojo y átomos de carbono en gris.	161
A.3. Distribución de ángulos que definen la orientación del huésped en los anfitriones variando entre modelo de cargas atómicas RESP/AM1-BCC, RESP/D-MBIS(H) y D-MBIS(A)/D-MBIS(H) para el anfitrión y el huésped respectivamente. Los sistemas OA-G1 y TEMOA-G1 se muestran en la parte superior y los sistemas OA-G4 y TEMOA-G4 en la parte inferior.	162

B.1.	Conformaciones representativas (A-F) adoptadas por el sistema CB7-BG1 durante la simulación de energía libre de unión.	164
B.2.	Conformaciones representativas (A-E) adoptadas por el sistema CB7-BG2 durante la simulación de energía libre de unión.	165
B.3.	Conformaciones representativas (A-C) adoptadas por el sistema CB7-BG3 durante la simulación de energía libre de unión.	166
B.4.	Conformaciones representativas (A-E) adoptadas por el sistema CB7-BG4 durante la simulación de energía libre de unión.	167
B.5.	Conformaciones representativas (A-E) adoptadas por el sistema CB7-BG5 durante la simulación de energía libre de unión.	168
B.6.	Conformaciones representativas (A-C) adoptadas por el sistema CB7-BG6 durante la simulación de energía libre de unión.	169
B.7.	Análisis de Perturbación Adaptado a la Simetría (SAPT) de las interacciones en el sistema CB7-AG8 a nivel SAPT0/juncVDZ. Los átomos del huésped se colorean según su contribución a la energía de interacción. Las barras en el medio muestran la parte electrostática (Elec), dispersión (Disp), inducción desde el anfitrión (A) al huésped (B) (IndAB) e intercambio (Inter) de la energía de interacción total es expresada en kcal/mol. Los gráficos circulares de la parte inferior muestran la contribución de cada grupo funcional a la energía de interacción total en porcentaje y la energía de interacción en kcal/mol.	170
B.8.	Conformaciones representativas (A-C) adoptadas por los huéspedes con núcleo aromático (AG1-8) durante la simulación de energía libre de unión	171

B.9. Función de distribución radial (RDF) entre los átomos de oxígeno del agua y el centro de masa del huésped CB7-AG1 (izquierda) o CB7-AG7 (derecha). Los RDF en azul corresponden a los sistemas descritos con el campo de fuerza GAFF combinado con cargas atómicas RESP y AM1-BCC para el anfitrión y el huésped, respectivamente. Para los RDF grises, se utilizaron cargas atómicas D-MBIS para el anfitrión y huésped. Los RDF en violeta utilizan las mismas cargas atómicas D-MBIS combinadas con los parámetros de Lennard-Jones derivados a partir de la partición MBIS de la densidad electrónica (D-MBIS/LJ-q[D-MBIS]).	174
C.1. Estructura tridimensional para los ligandos L7 y L8 con la respectiva numeración de los átomos empleadas en las simulaciones.	175



Lista de Tablas

3.1.	Afinidades de unión calculadas para diferentes distancias del contraión Na^+ transformado alquímicamente al huésped y diferentes poses del huésped. (Distancias medidas entre el contraión Na^+ y el átomo de carbono del grupo carboxílico cargado negativamente del huésped en Å). Valores de energía libre en kcal/mol	84
3.2.	Afinidades de unión para el sistema lisozima T4 L99A unido al tolueno obtenido con YANK a diferente número de iteraciones utilizando el muestreador SAMS. Los estados termodinámicos intermedios corresponden a 27 y 29 para la fase compleja y acuosa, respectivamente. Valores de energía libre en kcal/mol, el tiempo estimado de la simulación se muestra en horas	90
3.3.	Energías libres de unión para los sistemas OA/TEMOA unidos a los huéspedes del SAMPL5 (G1-6). Experimentales [5] y calculadas usando las cargas atómicas RESP para el OA/TEMOA y cargas atómicas AM1-BCC o D-MBIS para el huésped. Los valores $\langle E_{\text{complejo}}^{\text{pol}} \rangle$ y $\langle E_{\text{solvente}}^{\text{pol}} \rangle$ se obtuvieron con las cargas atómicas RESP/D-MBIS para los anfitriones y huéspedes G1-6, respectivamente. Valores en kcal/mol . . .	102
3.4.	Energías libres de unión para el sistema OA unido a los huéspedes del SAMPL5 (G1-6). Experimentales [5] y calculadas usando cargas atómicas RESP para el anfitrión OA y D-MBIS para los huéspedes (D-MBIS(H)) o cargas atómicas D-MBIS para el anfitrión OA y los huéspedes (D-MBIS). Los valores de $\langle E_{\text{complejo}}^{\text{pol}} \rangle$ y $\langle E_{\text{solvente}}^{\text{pol}} \rangle$ fueron obtenidas con las cargas atómicas D-MBIS para el anfitrión y los huéspedes G1-6. Valores en kcal/mol	110

- 3.5. Energías libres de unión para el CB7 unido a huéspedes con núcleo aromático (BG1-6). Experimentales [5] y calculadas utilizando cargas atómicas RESP para el CB7 y AM1-BCC o cargas atómicas D-MBIS para el CB7 y huéspedes, se consideran los parámetros del campo de fuerza GAFF. Los valores $\langle E_{\text{complejo}}^{\text{pol}} \rangle$ y $\langle E_{\text{solvente}}^{\text{pol}} \rangle$ se obtuvieron con las cargas atómicas D-MBIS para el anfitrión y huéspedes BG1-6. Valores en kcal/mol 113
- 3.6. Energías libres de unión para el CB7 unido a huéspedes con núcleo de adamantano (AG1-8). Experimentales [5] y calculadas utilizando cargas atómicas RESP para el CB7 y AM1-BCC o cargas atómicas D-MBIS para el CB7 y huéspedes y los parámetros del campo de fuerza GAFF. Los valores $\langle E_{\text{complejo}}^{\text{pol}} \rangle$ y $\langle E_{\text{solvente}}^{\text{pol}} \rangle$ se obtuvieron con las cargas atómicas D-MBIS para el anfitrión y huéspedes AG1-8. Valores en kcal/mol 117
- 3.7. Energías libres de unión para el CB7 unido a huéspedes con núcleo de adamantano (AG1-8). Experimentales [5] y calculadas usando las cargas atómicas RESP para el CB7 y AM1-BCC para los huéspedes con parámetros de Lennard-Jones(LJ) del campo de fuerza GAFF (q[AM1-BCC]) o parámetros de LJ derivados de la partición MBIS de la densidad electrónica polarizada del huésped combinada con las cargas atómicas D-MBIS derivadas para el CB7 y huéspedes (LJ-q[D-MBIS]) o usando solo las cargas atómicas D-MBIS del CB7 y las cargas atómicas D-MBIS del huésped derivado en el estado de no unión (LJ-q[D-MBIS(solv)]). Valores en kcal/mol 121

3.8.	Energías libres de unión para la lisozima T4 L99A/M102Q unida a ligandos con núcleo aromático (L1-8). Experimentales [5] y calculadas usando el campo de fuerza AMBERff14SB para el sistema proteico Lisozima T4 L99A/M102Q y variando las cargas atómicas y parámetros de Lennard-Jones para los ligandos. Los valores de afinidad de unión de referencia considera las cargas AM1-BCC para el ligando con parámetros de LJ del campo de fuerza GAFF (q[AM1-BCC]). Adicionalmente se reporta la afinidad de unión usando las cargas atómicas D-MBIS para el ligando derivados en la fase de no unión en conjunto con los parámetros de LJ del campo de fuerza GAFF (q[DMBIS(solv)]) y las cargas atómicas con parámetros de LJ obtenidos a partir de la partición MBIS de la densidad electrónica molecular del ligando en fase acuosa (LJ-q[DMBIS(solv)]). Valores en kcal/mol	129
A.1.	Valores de pK_a para todos los grupos titulables en los huéspedes G1-6. ^a valores de pK_a a 298 K de acuerdo con Haynes, William M.,ed (2015) CRC Handbook of Chemistry and Physics (95th ed.). ^b valores de pK_a a 298 K según el software Chemizalize. ^c No determinado.	160
A.2.	Diferencia absoluta de la carga atómica promedio por átomo entre las cargas atómicas D-MBIS calculadas en anfitrión y en solución acuosa. Valores en $\times 10^{-3}e$	161
A.3.	Energía libre de unión obtenida usando las cargas atómicas D-MBIS del huésped derivadas en el ambiente acuoso (AC) para el estado unido y no unido y usando las cargas para el huésped derivadas cuando está unida al anfitrión (HC) (sin corrección de energía de polarización). Valores en kcal/mol .	162

B.1.	Valores de pK_a para todos los grupos titulables en los huéspedes BG1-6 y AG1-8, respectivamente. ^a valores de pK_a a 298 K de acuerdo con Haynes, William M.,ed (2015) CRC Handbook of Chemistry and Physics (95th ed.). ^b valores de pK_a de acuerdo con Dean, J.A. Handbook of Organic Chemistry. New York, NY: McGraw-Hill Book Co; (1987). ^c valores de pK_a a 298 K de acuerdo al software Chemicalize. ^d valores de pK_a a 298 K de acuerdo con Adamo Fini <i>et al.</i> Acidity constants of sparingly water-soluble drugs from potentiometric determinations in aqueous dimethyl sulfoxide. J. Pharm. Sci., 76(1):48-52, 1987. ^e No determinado.	163
B.2.	Energía libre de unión obtenida para el CB7 unido a los huéspedes BG2 y AG5. Simulaciones CB7-BG2 y CB7-BG2(N) corresponde al huésped en su forma protonada y neutra, respectivamente. CB7-AG5 y CB7-AG5(P) corresponde al huésped en su forma desprotonada y protonada, respectivamente. Valores en kcal/mol	164
B.3.	Porcentaje relativo de la orientación del huésped en las ventanas de simulación con $\lambda = 1$ y asociado con las conformaciones (A-F) mostradas en la Figura B.1. Las simulaciones fueron realizadas usando las cargas atómicas RESP/AM1-BCC y D-MBIS/D-MBIS para el anfitrión y huésped respectivamente. Adicionalmente, se reportan los porcentajes de conformaciones con la presencia de agua en el sitio de unión. . .	164
B.4.	Porcentaje relativo de la orientación del huésped en las ventanas de simulación con $\lambda = 1$ y asociados con las conformaciones (A-E) mostradas en la Figura B.2. Las simulaciones fueron realizadas usando las cargas RESP/AM1-BCC y D-MBIS/D-MBIS para el anfitrión y huésped respectivamente. Adicionalmente, se reportan los porcentajes de conformaciones con la presencia de agua en el sitio de unión.	165

- B.5. Porcentaje relativo de la orientación del huésped en las ventanas de simulación con $\lambda = 1$ y asociado con las conformaciones (A-C) mostradas en la Figura B.3. Las simulaciones fueron realizadas usando las cargas atómicas RESP/AM1-BCC y D-MBIS/D-MBIS para el anfitrión y huésped respectivamente. Adicionalmente, se reportan los porcentajes de conformaciones con la presencia de agua en el sitio de unión. . . 166
- B.6. Porcentaje relativo de la orientación del huésped en las ventanas de simulación con $\lambda = 1$ y asociado con las conformaciones (A-E) mostradas en la Figura B.4. Las simulaciones fueron realizadas usando las cargas atómicas RESP/AM1-BCC y D-MBIS/D-MBIS para el anfitrión y huésped respectivamente. Adicionalmente, se reportan los porcentajes de conformaciones con la presencia de agua en el sitio de unión. . . 167
- B.7. Porcentaje relativo de la orientación del huésped en las ventanas de simulación con $\lambda = 1$ y asociado con las conformaciones (A-E) mostradas en la Figura B.5. Las simulaciones fueron realizadas usando las cargas atómicas RESP/AM1-BCC y D-MBIS/D-MBIS para el anfitrión y huésped respectivamente. Adicionalmente, se reportan los porcentajes de conformaciones con la presencia de agua en el sitio de unión. . . 168
- B.8. Porcentaje relativo de la orientación del huésped en las ventanas de simulación con $\lambda = 1$ y asociado con las conformaciones (A-E) mostradas en la Figura B.6. Las simulaciones fueron realizadas usando las cargas atómicas RESP/AM1-BCC y D-MBIS/D-MBIS para el anfitrión y huésped respectivamente. Adicionalmente, se reportan los porcentajes de conformaciones con la presencia de agua en el sitio de unión. . . 169
- B.9. Peso relativo en % de una específica orientación del huésped (A-C) mostradas en la Figura B.8 observado en las ventanas de simulación con $\lambda = 1$ durante los cálculos de energía libre de unión. Las simulaciones se realizaron utilizando las cargas atómicas RESP/AM1-BCC para el anfitrión y el huésped respectivamente y utilizando los parámetros de LJ del campo de fuerza GAFF. Además, se reportan los porcentajes de conformaciones con presencia de agua en el sitio de unión. . 171

- B.10. Peso relativo en % de una específica orientación del huésped (A-C) mostradas en la Figura B.8 observado en las ventanas de simulación con $\lambda = 1$ durante los cálculos de energía libre de unión. Las simulaciones se realizaron utilizando las cargas atómicas D-MBIS/D-MBIS para el anfitrión y el huésped respectivamente y utilizando los parámetros de LJ del campo de fuerza GAFF. Además, se reportan los porcentajes de conformaciones con presencia de agua en el sitio de unión. 172
- B.11. Peso relativo en % de una específica orientación del huésped (A-C) mostradas en la Figura B.8 observado en las ventanas de simulación con $\lambda = 1$ durante los cálculos de energía libre de unión. Las simulaciones se realizaron utilizando las cargas atómicas D-MBIS/D-MBIS para el anfitrión y el huésped respectivamente y utilizando los parámetros de LJ del campo de fuerza GAFF y los derivados a partir de la partición MBIS de la densidad electrónica para el anfitrión y huésped respectivamente. Además, se reportan los porcentajes de conformaciones con presencia de agua en el sitio de unión. 173
- C.1. Valores de pK_a para todos los grupos titulables en los ligandos L1-8. ^a valores de pK_a a 298 K de acuerdo con Haynes, William M.,ed (2015) CRC Handbook of Chemistry and Physics (95th ed.). ^b valores de pK_a a 298 K según el software Chemizalze. ^c No determinado. 175

Resumen

La predicción de la afinidad de unión por medio de simulaciones computacionales se ha incorporado cada vez más en los proyectos de descubrimiento de fármacos debido a que reducen el número de compuestos que deben ser sintetizados para obtener la potencia y selectividad deseada. Sin embargo, su amplia aplicación está limitada a la precisión de las predicciones en cálculos de energía libre de unión. Las principales fuentes de error se atribuyen a los campos de fuerza utilizados para describir las interacciones moleculares, el muestreo incompleto del espacio configuracional y la descripción correcta del sistema molecular (p. ej., protonación del ligando, estado tautomérico y tratamiento de las condiciones experimentales).

Los sistemas anfitrión-huésped se han utilizado como modelo para mejorar y evaluar la calidad de los campos de fuerza junto a la estrategia para la obtención de afinidades de unión mediante métodos computacionales, porque comparten interacciones similares a las encontradas en los sistemas proteína-ligando. Los anfitriones suelen tener estados de protonación definidos, están conformados por menos de 100 átomos sin considerar los átomos de hidrógeno, poseen una cavidad o hendidura que les permite unirse a otros compuestos de menor tamaño denominados huéspedes y exhiben rigidez estruc-

tural que facilita el muestreo configuracional. El muestreo configuracional es realizado con campos de fuerza clásicos cuyos parámetros de las interacciones intermoleculares fueron evaluados en vacío sin considerar el entorno molecular del anfitrión o el huésped. En esta investigación, se propone un nuevo protocolo que combina cálculos QM/MM y simulaciones de dinámica molecular para derivar los parámetros no enlazantes como las cargas atómicas en sistemas modelo anfitrión-huésped a partir de la partición de la densidad electrónica molecular utilizando el método MBIS (Minimal Basis Iterative Stockholder). Además, el protocolo calcula la energía promedio requerida para polarizar la densidad electrónica del huésped en el ambiente de unión y no unión. Este término suele ignorarse en los campos de fuerza clásicos y es relevante en huéspedes con mayor polarizabilidad.

El protocolo se validó calculando las energías libres de unión en los sistemas acuosos: (i) octa ácido (OA) y tetra-endometil octa ácido (TEMOA) unidos a huéspedes que presentan diferentes grupos funcionales, tales como ión carboxilato, halogenuro, alquino, ciano, amonio cuaternario o nitro con núcleo de adamantano, aromático o cadena alquílica. (ii) Cuburbit[7]urilo (CB7) unidos a huéspedes con núcleo aromático o de adamantano. En los sistemas OA/TEMOA el cálculo de la energía libre de unión empleando cargas atómicas D-MBIS y su posterior corrección por polarización mejoran el rendimiento predictivo en comparación con las cargas atómicas de referencia AM1-BCC. Además, se hace evidente que la energía libre de unión depende

de las cargas atómicas específicas del entorno y la corrección por la energía promedio de polarización, específicamente en sistemas con grupos funcionales polarizables. En los sistemas CB7 el cálculo de energía libre de unión empleando cargas atómicas D-MBIS y su posterior corrección por polarización arrojan mejoras en el rendimiento predictivo para los huéspedes con núcleos aromáticos. El estudio de las interacciones mediante la metodología SAPT para los huéspedes con núcleo de adamantano demostró que en estos huéspedes predominan las interacciones de intercambio y dispersión. Este tipo de interacciones están descritas en los campos de fuerza clásicos por el potencial de Lennard Jones que contiene como parámetros los coeficientes de dispersión C_6 y los radios de van-der-Waals para cada átomo. Al derivar estos parámetros con el método MBIS junto a las cargas D-MBIS para los huéspedes con núcleo de adamantano unidos al CB7 se observó una disminución en el error cuadrático medio (RSME) de $\simeq 2$ kcal/mol en las energías libres de unión.

Finalmente se aplicó el protocolo validado en los sistemas modelo al sistema proteico lisozima T4 L99A/M102Q unido a huéspedes con núcleo aromático. En este caso en particular se ocuparon los parámetros no covalentes derivados en solvente para la fase de unión y no unión. Los resultados empleando el nuevo conjunto de parámetros, no evidencian mejoras en el poder predictivo, respecto a los descriptores estadísticos obtenidos con las cargas AM1-BCC. Sin embargo, para concluir al respecto del rendimiento en

la predicción de energías de afinidad en sistemas proteína-ligando empleando el nuevo protocolo, se requieren estudios adicionales que involucren la derivación de los parámetros de no enlace en ambos ambientes moleculares y consideren otros sistemas proteína-ligando con rangos más amplios en la energía libre de unión. El ligando acetato de bencilo (L6), presentó la mayor desviación de la energía libre de unión, se pudo establecer que el uso de la estructura cristalina HOLO en vez de la APO en las simulaciones mejora la predicción de su afinidad con las cargas atómicas D-MBIS. En futuras investigaciones se propone estudiar el efecto de la estructura inicial de la proteína en la predicción de la energía libre de unión.



1. Introducción

En los últimos años la implementación de simulaciones computacionales se ha convertido en una importante herramienta para el estudio de diversos sistemas físicos, químicos y biológicos. Dentro del espectro de simulaciones computacionales se destacan las empleadas para predecir energías libres de unión en sistemas proteína-ligando debido a su impacto en el sector farmacológico, ya que permite acelerar el desarrollo de fármacos por la reducción del número de compuestos que deberían ser sintetizados [6].

La precisión de estas simulaciones depende del muestreo completo de las configuraciones adoptadas por el complejo proteico, de la descripción correcta del sistema molecular (p. ej., la protonación del ligando, estado tautomérico y los factores experimentales) y del campo de fuerza que describe las interacciones [5]. La flexibilidad y complejidad inherentes a los sistemas proteína-ligando dificultan la cuantificación y el análisis de los errores de predicción a partir de una descripción correcta de las interacciones del sistema molecular. El principal problema son las conformaciones desconocidas del complejo proteína-ligando y cómo las conformaciones del ligando pueden diferir durante el proceso de unión [7]. Para eludir el problema de muestreo del espacio conformacional de la proteína, se han propuesto moléculas relati-

vamente rígidas como anfitriones. Estos anfitriones tienen unos pocos cientos de átomos y tienden a ser más rígidos y menos complejos que las proteínas, lo que facilita el muestreo conformacional. Los sistemas anfitrión-huésped, por lo tanto, sirven como modelo para la predicción de la afinidad de unión absoluta porque comparten eventos de reconocimiento molecular regidos por las mismas fuerzas básicas que los sistemas proteína-ligando y permiten un muestreo más completo del espacio configuracional [8].

El estudio de sistemas moleculares de manera adecuada usando metodologías computacionales requiere la descripción de todas las interacciones de las partículas del sistema. Idealmente se podría recurrir a modelos mecano-cuánticos que describen la interacción de los electrones y núcleos del sistema (mecánica cuántica QM). Pero debido a las dimensiones de los sistemas de estudio y el tiempo de cálculo que ellos demandan, usar modelos mecano-cuánticos no es una opción viable. Como alternativa a estos modelos se emplean los métodos clásicos describiendo las interacciones entre átomos sin considerar a los electrones (mecánica molecular MM) permitiendo abordar sistemas químicos de gran tamaño. Sin embargo, estos modelos no son capaces de describir la formación o rompimiento de enlaces covalentes, ni cambios drásticos en la estructura electrónica. Por tanto, para describir adecuadamente sistemas de gran tamaño que involucran un cambio significativo en la estructura electrónica en alguna región determinada, comúnmente se recurre a métodos híbridos QM/MM [9]. En los métodos QM/MM, la zona

donde ocurre un cambio en la estructura electrónica se describe por QM y el resto del sistema se describe por MM mediante el uso de pseudo-potenciales, cargas eléctricas o campos de fuerza.

Los campos de fuerza clásicos como CHARMM [10], AMBER-GAFF [11], GROMOS [12] y OPLS [13] emplean funciones matemáticas que describen la interacción entre los átomos de un sistema y están formados por dos componentes: (i) las funciones de energía potencial, definidas por un conjunto de ecuaciones que engloban una serie de contribuciones energéticas enlazantes y no enlazantes y (ii) los parámetros empíricos usados por cada uno de los términos. Los términos enlazantes incluyen contribuciones debidas a los enlaces covalentes, ángulos de valencia y ángulos torsionales propios e impropios. Los términos no enlazantes se definen por un término de atracción-repulsión de tipo Lennard-Jones (LJ) para las fuerzas de van-der-Waals y un término coulombico para las interacciones electrostáticas [14]. Comúnmente los campos de fuerza no polarizables asignan un valor estático para la carga atómica y no considera la variación debido al entorno electrostático. La implementación de este tipo de modelos de carga constituye una limitación de los campos de fuerza, debido a que en muchos sistemas la densidad electrónica varía dependiendo del entorno molecular. Implementar un modelo de carga atómica que considere el entorno molecular ha despertado el interés de la comunidad científica en los últimos años, favoreciendo el desarrollo de nuevos campos de fuerza como es el caso de AMOEBA [15], SIBFA [16],

SDFF [17], NEMO [18], entre otros que permiten obtener simulaciones más precisas respecto a los observables de interés. Sin embargo, requieren de un mayor costo computacional y hacen uso de parámetros derivados del ajuste para reproducir propiedades químicas específicas. Otra alternativa para incluir el efecto de la polarización es mediante la metodología propuesta en la presente investigación, la cual, mediante simulaciones de dinámica molecular y QM/MM entrega un nuevo conjunto de cargas y la energía de polarización, que pueden emplearse posteriormente para el cálculo de energías libres de unión en sistemas anfitrión-huésped y proteína-ligando.

Los avances investigativos en la predicción de energías libre de unión suelen evaluarse en los sistemas modelo anfitrión-huésped. Estos sistemas han sido parte de los desafíos de la evaluación estadística del modelado de proteínas y ligandos (SAMPL) que sirven como una oportunidad para que la comunidad científica interesada en predicciones de energía libre de unión, pruebe y mejore los métodos computacionales, algoritmos y campos de fuerza [19, 8, 20]. En uno de los últimos desafíos SAMPL7, Amezcua *et al.* resaltó la importancia de describir correctamente los fenómenos de polarización en los anfitriones y huéspedes que presentan una carga neta asociada ya que puede conducir a una mayor precisión en los cálculos de energía libre de unión [21]. Adicional a la correcta descripción de la polarización es relevante considerar que algunos sistemas presentan interacciones repulsivas de intercambio e interacciones atractivas de dispersión significativas y los campos

de fuerza no polarizables describen estas interacciones mediante el potencial de LJ específico del átomo. Mohebifar *et al.* han demostrado recientemente que los coeficientes atómicos C_6 del potencial de LJ en varios campos de fuerza están sobreestimados en comparación con los derivados de los cálculos químicos cuánticos utilizando el modelo de dipolo de intercambio [22]. La discrepancia entre los parámetros LJ que se encuentran en los campos de fuerza comunes y los de los cálculos químicos cuánticos ha motivado diferentes enfoques para derivarlos. En algunos sistemas anfitrión-huésped se ha demostrado que modestas modificaciones en algunos de los parámetros de LJ conllevan a mejoras significativas en los valores de energía libre de unión calculados [23].

En la presente investigación se evaluó la exactitud en la predicción de energía libre de unión al emplear las cargas atómicas derivadas al particionar la densidad electrónica molecular empleando el método MBIS (Stockholder method using the polarized electron density) en los ambientes de unión y no unión para los huéspedes del SAMPL5 unidos a OA/TEMOA y huéspedes aromáticos unidos a CB7. Para estos sistemas se obtuvo una mejor predicción en la afinidad de unión empleando el nuevo conjunto de cargas atómicas en comparación con las cargas atómicas de referencia. Mediante un análisis de SAPT (Análisis de perturbación adaptado a la simetría) se mostró que en los sistemas CB7 unido a huéspedes con núcleo aromático predominan las interacciones electrostáticas, mientras que en los sistemas CB7 unidos a

huéspedes con núcleo de adamantano predominan las interacciones de intercambio repulsivo y dispersión atractiva. Para los sistemas CB7 unidos a huéspedes con núcleo de adamantano, se complementó el flujo de trabajo computacional, para derivar las cargas atómicas y parámetros de LJ basados en la partición MBIS de la densidad electrónica molecular. Los parámetros de campo de fuerza no enlazantes mejoran la predicción de la afinidad de unión, para los huéspedes con núcleo de adamantano unidos al CB7 donde predominan las interacciones de intercambio repulsivo y las interacciones de dispersión atractiva con el huésped. Finalmente, considerando la baja polaridad del sitio de unión para la doble mutante de la lisozima T4 L99A/M102Q se consideró derivar los parámetros no enlazantes de los ligandos con núcleo aromático solamente en la fase de no unión. En el caso particular del sistema proteico estudiado, los resultados no presentan mejores descriptores estadísticos. Sin embargo, para concluir al respecto de los parámetros de no enlace derivados de la densidad electrónica molecular, se requieren simulaciones adicionales donde se deriven los parámetros no enlazantes en ambos ambientes moleculares.

1.1. Marco Teórico

La energía libre de unión entre dos moléculas en fases condensadas, $\Delta G_{unión}^0$ se define como la diferencia de energía libre de Gibbs entre estas moléculas en un estado libre (o no unidas) y estado de unión. Puede determinarse experimentalmente a través de técnicas de titulación calorimétrica isotérmica [24] o resonancia magnética nuclear (NMR). Desde el punto de vista computacional existen metodologías que permiten calcular la diferencia de energía libre de unión entre dos moléculas o sistemas receptor-ligando basadas en simulaciones capaces de explorar el espacio de fases. Dentro de las más destacadas se encuentra el método de perturbaciones de energía libre (FEP), integración termodinámica (TI) y Bennet's acceptance ratio (BAR) [25, 26].

La precisión de los métodos mencionados anteriormente depende de la superposición de los espacios de fase de los estados de unión y no unión. Cuando los espacios de fase no muestran gran superposición, se realiza una transformación paulatina mediante un parámetro λ para representar pequeñas alteraciones al sistema y modular la energía potencial, de modo que a medida que λ pasa de 0 a 1, la energía potencial cambia entre el estado inicial y el estado final del sistema. Habitualmente las simulaciones se ejecutan en un conjunto de valores de λ que conectan los estados de unión y no unión. En otras palabras, la vía de transformación se subdivide en una variedad de pa-

tos individuales, donde cada paso implica una transición entre dos valores de λ . El número y el espaciado de los valores de λ se eligen para asegurar una superposición adecuada entre los espacios configuracionales de los estados considerados [27]. Normalmente estos estados intermedios y sus correspondientes valores de λ corresponden a estados no físicos (*alquímicos*), que a menudo implican un cambio en la identidad química de la especie considerada [28].

Los conjuntos de estados intermedios mencionados anteriormente pueden ser estados físicos o *alquímicos* y forman parte del ciclo termodinámico que permite la estimación de energías libres de unión. Dentro de los ciclos termodinámicos, dos amplios enfoques han sido estudiados [29]: la energía libre de unión absoluta y energía libre de unión relativa a un ligando en específico (Figura 1.1 arriba y abajo respectivamente).

La energía libre de unión absoluta evalúa el trabajo reversible de transferir el ligando del sitio de unión a la solución. La determinación computacional directa de la energía libre de unión empleando este enfoque puede ser difícil y computacionalmente intensiva. Por tal motivo se recurre al cálculo de energía libre de unión a partir de una doble aniquilación del ligando formada en el estado de unión y no unión respectivamente. La aniquilación del ligando se interpreta como su desacoplamiento del resto del sistema molecular. Bajo este enfoque la diferencia de energía libre puede calcularse como:

$$\Delta G_{\text{uni3n}} = \Delta G_{\text{solvente}} - \Delta G_{\text{complejo}} \quad (1.1)$$

Donde $\Delta G_{\text{solvente}}$ es la energa libre de hidrataci3n y $\Delta G_{\text{complejo}}$ es la energa libre de aniquilaci3n del ligando del complejo receptor-ligando y del resto del sistema molecular [29].



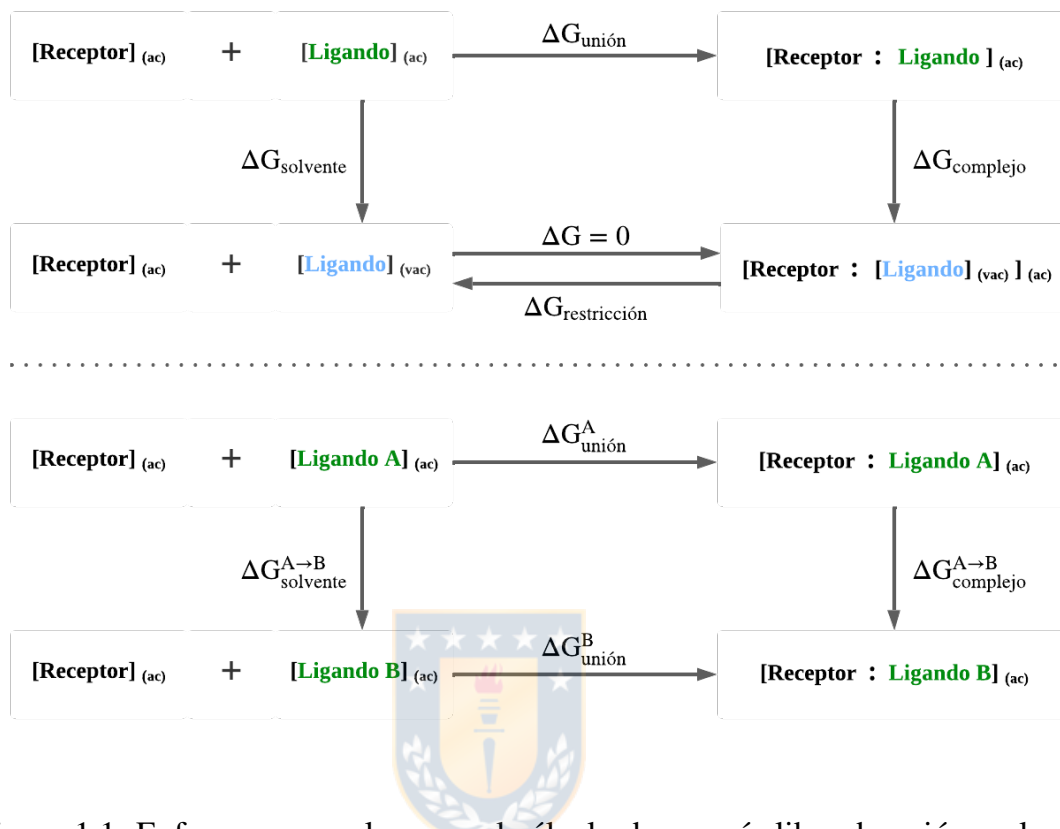


Figura 1.1: Enfoques generales para el cálculo de energía libre de unión en los sistemas receptor-ligando. Arriba: Energía libre de unión absoluta, considera la aniquilación del ligando en el ambiente de unión y no unión. La representación en verde del ligando indica que las interacciones electrostáticas y van-der-Waals se encuentran activas y en azul indica que dichas interacciones se han desactivado. Abajo: Energía libre de unión relativa, considera modificaciones *alquímicas* para transformar estructuralmente el ligando A al ligando B, tanto en el ambiente de unión como no unión. La representación en verde indica que las interacciones electrostáticas y van-der-Waals se encuentran activas en ambos ligandos.

Por otra parte, el enfoque de energía libre de unión relativa, calcula la diferencia de energía libre de unión entre dos diferentes ligandos para el mismo receptor. Estas transformaciones son *alquímicas* y consideran el cambio de energía libre al convertir *alquímicamente* un ligando en otro, primero en

el estado de unión y luego en estado de no unión [27]. En este enfoque la energía libre de unión se calcula mediante la siguiente expresión:

$$\Delta\Delta G_{\text{unión}} = \Delta G_{\text{complejo}}^{A \rightarrow B} - \Delta G_{\text{solvente}}^{A \rightarrow B} \quad (1.2)$$

En la Ecuación 1.2 $\Delta\Delta G_{\text{unión}}$ se refiere a la diferencia de energía libre de unión $\Delta G_{\text{unión}}^A$ y $\Delta G_{\text{unión}}^B$, $\Delta G_{\text{solvente}}^{A \rightarrow B}$ es la energía libre de transformación del ligando A al ligando B en estado acuoso y sin interactuar con el receptor y $\Delta G_{\text{complejo}}^{A \rightarrow B}$ es la energía de transformación del complejo receptor-ligando A al complejo receptor-ligando B. Este enfoque resulta ser muy eficiente si los dos ligandos son similares desde el punto de vista estructural. Sin embargo, puede ser más complicado y presentar grandes problemas de muestreo si ambos ligandos son muy diferentes químicamente o si hay barreras energéticas muy grandes para interconvertir entre los estados conformaciones más estables [30].

A parte del enfoque empleado, uno de los factores principales que afectan la precisión en las simulaciones de energía libre de unión son las descripciones de las interacciones microscópicas en el sistema, las cuales se especifican en el campo de fuerza. Como se mencionó en párrafos anteriores los campos de fuerza son funciones matemáticas que engloban contribuciones energéticas enlazantes y no enlazantes. Las interacciones en los sistemas receptor-ligando son de carácter no covalente por tal motivo los términos no enlazantes

del campo de fuerza presentan mayor relevancia en las predicciones de energía libre de unión. Los campos de fuerza clásicos describen las interacciones electrostáticas en términos de cargas parciales fijas y centradas en los átomos. Si bien las cargas atómicas no son un observable en el contexto de la teoría cuántica, sí dan una descripción simple de la distribución de electrones en una molécula y de las interacciones electrostáticas. Por consecuencia las cargas atómicas sólo pueden determinarse indirectamente midiendo observables fuertemente dependientes de la carga, por ejemplo, el momento dipolar o el gradiente del campo eléctrico.

El uso de campos de fuerza no polarizables conlleva a una descripción inadecuada para algunos sistemas, especialmente en casos donde las moléculas están polarizadas [31]. En un sistema molecular real la polarización juega un rol significativo ya que las cargas pueden fluctuar debido a los reordenamientos moleculares que se llevan a cabo en los sistemas de interés. La comunidad científica ha desarrollado un componente en los campos de fuerza que incluye la polarización. En algunos casos se han creado extensiones de los campos de fuerza clásicos y en otros se han diseñado nuevos campos de fuerza que incorporan un término de polarización desde sus inicios como es el caso de AMOEBA, SIBFA, SDFF, NEMO, entre otros [14]. La inclusión del término de polarización en el campo de fuerza permite obtener simulaciones más precisas respecto a los observables de interés, sin embargo, precisan de un mayor costo computacional.

Según lo expuesto anteriormente la polarización no está definida en los campos de fuerza clásicos, pero puede ser modelada por el tratamiento de cargas atómicas parciales como parámetros efectivos. En general, existen dos estrategias para derivar las cargas atómicas parciales, que también se usan en combinación: (i) ajuste de datos experimentales y (ii) extracción a partir del potencial electrostático molecular (MEP) obtenido de cálculos QM [32]. Es importante tener en cuenta que el último enfoque es un problema matemático sin una solución única, debido a la existencia de infinitas distribuciones de carga que pueden reproducir el potencial electrostático (ESP). El enfoque de ajuste ESP depende de la conformación y del conjunto de bases utilizados, además la carga de los átomos lejanos a la superficie en donde se evalúa el potencial no es característica debido a la baja sensibilidad del potencial en puntos lejanos a la carga. Este problema se asocia generalmente a átomos de carbono. Para abordar este punto frágil de las cargas ESP, Bayly *et al.* [33] presentaron una variante para calcular cargas basadas en el potencial electrostático, introduciendo una función de penalización dentro del proceso de ajuste para obtener las cargas atómicas denominadas RESP. Al introducir esta función de penalización sobre los átomos de carbono para calcular las cargas, se logra reproducir el potencial electrostático molecular y se reduce el problema de intransferibilidad de cargas entre grupos funcionales similares. Sin embargo, las funciones de penalización carecen de significado físico lo que supone uno de los principales problemas de las cargas RESP. En general

las cargas atómicas usadas con mayor frecuencia (AM1-BCC) son derivadas mediante métodos semi-empíricos y tienen la desventaja de incurrir en errores al ser extrapolados a sistemas químicos diferentes.

En forma más rigurosa las cargas atómicas también pueden ser calculadas a partir de una propiedad local como la densidad electrónica. Bajo esta panorámica en 1977 Hirshfeld propuso un modelo de cargas atómicas en el cual la densidad pro-molecular $\sum_{B=1}^N \rho_B^0(\mathbf{r})$ es obtenida superponiendo las densidades atómicas calculadas para pro-átomos neutros $\rho_A^0(\mathbf{r})$ [34]. Sin embargo, este modelo presenta puntos frágiles dado que las cargas atómicas calculadas son cercanas a cero. Los pro-átomos son densidades promediadas esféricamente de átomos neutros aislados, implicando que los resultados dependan en gran medida de la elección de los pro-átomos. Además, estas cargas no reproducen el potencial electrostático molecular. Estos problemas fueron corregidos mediante las cargas atómicas de Hirshfeld iterativas (Hirshfeld-I o HI) haciendo uso de una función de peso iterativa que actualiza la carga de cada pro-átomo hasta hacerlo lo más parecido posible a la carga del átomo en la molécula. Estas cargas tienen una baja dependencia con la conformación utilizada y reproducen el potencial electrostático en vacío lo cual las hace aptas para simulaciones de dinámica molecular [35].

Desafortunadamente, también el método HI tiene sus deficiencias que son más relevantes en sistemas altamente polares. Cuando el método se aplica a óxidos altamente polares, se requiere la densidad promediada esféricamente

del dianión de oxígeno como entrada. Sin embargo, ningún dianión atómico aislado es estable lo que impide el cálculo de la densidad electrónica. Cuando la densidad del dianión se calcula con un conjunto de bases localizadas, la metodología HI sobreestima gravemente los potenciales electrostáticos de los óxidos metálicos altamente polares [36].

Un método que no hace uso de densidades atómicas fue propuesto por Toon Verstraelen y es denominado cargas atómicas MBIS por sus cifras en inglés Minimal Basis Iterative Stockholder. El método MBIS es una alternativa para particionar la densidad electrónica molecular y al no requerir de densidades de iones inestables como entrada se esperaría que no sufra de los mismos problemas que HI [37]. No obstante, tanto las cargas atómicas HI como MBIS se derivan de la densidad electrónica en vacío y no incluye el efecto del solvente, por tanto, pueden presentar deficiencias al ser aplicadas en fases condensadas. Como solución a esta deficiencia se han implementado nuevas estrategias dentro de las cuales se destacan las cargas atómicas dinámicas HI (D-HI) y las cargas atómicas dinámicas MBIS (D-MBIS). Las cargas D-HI y D-MBIS se derivan a partir de simulaciones de dinámica molecular usando métodos QM/MM para describir el ligando (QM) y el resto del sistema molecular con MM. Esta metodología incluye el efecto de polarización de la molécula debido al entorno molecular y provee una mejor descripción de las interacciones electrostáticas y la dinámica conformacional [38, 39, 40].

En el cálculo de energía libre de unión empleando cargas atómicas diná-

micas D-MBIS considera la corrección por polarización asociado a (i) costo energético de polarizar la densidad electrónica del ligando al pasar del vacío a solución (ii) costo energético de polarizar la densidad electrónica del ligando al pasar de la fase acuosa al sitio de unión. Si no se realizan estas correcciones de polarización se asume erróneamente la misma densidad electrónica para el ligando en el vacío y en solución acuosa, y la misma densidad electrónica para el ligando en fase acuosa y en el sitio de unión [41].

En moléculas con grupos funcionales menos polares, priman las interacciones de dispersión y repulsivas de intercambio moduladas mediante los parámetros de LJ en los diferentes campos de fuerza. Existen algunos avances donde se propone derivarlos a partir de la densidad electrónica. Cole *et al.* propuso utilizar la partición AIM de la densidad electrónica para derivar coeficientes de dispersión optimizando los radios de van-der-Waals que reproduzcan algunas densidades de líquidos [42]. Kantonen *et al.* optimizó los parámetros de LJ empleando la partición MBIS y posteriormente se mapea a los parámetros de LJ utilizando un conjunto de parámetros de mapeo mucho más pequeños que el número típico de una clase de átomo en un campo de fuerza [43]. Visher *et al.* propuso derivar los coeficientes de dispersión a partir de una partición AIM de la densidad electrónica utilizando una relación empírica entre el radio de van-der-Waals y la polarizabilidad del átomo [44]. Derivando todos los parámetros del campo de fuerza a partir de los primeros principios que explican la polarización con el método Drude, pudieron

reproducir estimaciones experimentales de las propiedades termodinámicas de líquidos puros de los alcanos. Todos los enfoques mencionados utilizaron cálculos de estructura electrónica de moléculas representativas en el vacío sin tener en cuenta las conformaciones variables y el efecto del entorno en la densidad electrónica. Hasta donde sabemos, ningún método ha derivado parámetros para el potencial de LJ a partir de la densidad de electrones polarizados usando diferentes configuraciones del ligando y el solvente, ni ha analizado su efecto en los cálculos de energía libre de unión receptor-ligando.

Para estudiar la capacidad de predecir energías libres de unión con los diferentes modelos de parámetros no enlazantes, se emplean sistemas referencia anfitrión-huésped y proteína-ligando. Los sistemas de referencia han sido ampliamente estudiados desde el punto de vista experimental y computacional, permitiendo hacer un análisis detallado de los factores y parámetros que influyan en las determinaciones de energía libre de unión. En la siguiente sección se describe los sistemas de interés anfitrión-huésped y proteína-ligando.

1.2. Sistemas anfitrión-huésped

Para la validación de un método computacional sus resultados deben contrastar en buena medida a los valores obtenidos experimentalmente. En el caso de la energía libre de unión es común emplear sistemas anfitrión-huésped. Los anfitriones químicos son moléculas a menudo constituidas por menos de

100 átomos sin considerar los átomos de hidrógeno, con una cavidad o hendidura que les permite unirse a otros compuestos de menor tamaño denominados huéspedes. Los huéspedes presentan diversidad química y exhiben afinidades significativas con el entorno molecular de la cavidad del anfitrión. Los huéspedes suelen presentar mayor flexibilidad y su muestreo conformacional hace parte de los desafíos actuales en los sistema anfitrión-huésped. Los anfitriones unen a sus huéspedes a través de las mismas fuerzas básicas que las proteínas usan para unir sus ligandos lo que los hace un modelo computacional para estudiar uniones no covalentes relevantes en química y biología [5]. Dentro de los anfitriones más estudiados se encuentran la familia de los cucurbit[n]urilos constituidos por monómeros de glicolurilo ($n=5,6,7,8,10$) [45], los GDCC (Gibb Deep Cavity Cavitands) y las ciclodextrinas constituidas por monómeros de glucosa [5]. Estos anfitriones se caracterizan en la mayoría de los casos por tener una estructura rígida que facilita obtener las conformaciones más relevantes y así realizar simulaciones más precisas.

En el marco de buscar mejores metodologías para el cálculo de energías libres de unión en la comunidad científica se han realizado varias series de desafíos a ciegas SAMPL, cuyo objetivo principal es evaluar la capacidad predictiva de la metodología computacional empleada. Dentro de los SAMPL realizados hasta la fecha los sistemas anfitrión-huésped han hecho parte de los desafíos SAMPL3, SAMPL4, SAMPL5, SAMPL6, SAMPL7 y SAMPL8 [1, 6]. Los resultados de estos desafíos permiten deducir que aún no existe un

consenso en relación a la mejor metodología para obtener energías libres de unión en sistemas anfitrión-huésped. Por citar un ejemplo en el SAMPL6 los modelos empíricos (denominados así por el uso de parámetros empíricos adicionales) tienden a obtener un mejor rendimiento que los métodos de primeros principios (llamados también *ab-initio* ya que las únicas entradas de cálculo son constantes físicas). La precisión, sin embargo, mostró una dependencia sustancial del sistema considerado, enfatizando en la necesidad de incluir mayor variedad de huéspedes en posteriores desafíos. Otra de las conclusiones importantes es que se requiere más refinamiento de los parámetros del campo de fuerza, así como un tratamiento mejorado de los efectos químicos (condiciones de soluciones reguladoras de pH, estados de protonación) para mejorar aún más las predicciones [6].

En las siguientes subsecciones se describen los sistemas anfitrión-huésped considerados en la investigación. La familia de GDD y cucurbit[n]urilos son de especial interés debido a que presentan rigidez estructural que facilita la convergencia y el muestro conformacional en los cálculos de energía libre de unión. Las ciclodextrinas no se considerarán en la presente investigación debido a que presentan problemas de convergencia asociados a la flexibilidad de los monómeros de glucosa y dificultades en el muestreo atribuido a que la cavidad de unión presenta dos aperturas no equivalentes [46].

1.2.1. Anfitriones OA/TEMOA y huéspedes del SAMPL5

Los anfitriones más representativos de la familia de los GDCC son el octa ácido (OA) y el tetra-endometil octa ácido (TEMOA), un derivado con cuatro grupos metilo agregados en ubicaciones equivalentes en el portal superior [5]. Estos anfitriones poseen un sitio de unión hidrofóbico profundo en forma de canasta con ocho iones carboxilato responsables de su solubilidad en agua [5] (ver Figura 1.2). OA se presentó como anfitrión en el desafío SAMPL4 [47, 20] y se consideró en desafíos posteriores, mientras que TEMOA se consideró en los desafíos SAMPL5-6. Los huéspedes G1-6 (ver Figura 1.2) fueron propuestos para OA y TEMOA en el SAMPL5 y constituyen una referencia para la familia de los GDCC según Mobley y Gilson [5]. Los sistemas anfitrión-huésped resultantes presentan diversidad química y exhiben un rango significativo de energías libres de unión de -4.25 a -9.37 kcal/mol y -3.90 a -5.94 kcal/mol para OA y TEMOA respectivamente. Los anfitriones OA/TEMOA en combinación con los huéspedes del SAMPL5 también exhiben estabilidad a valores de pH extremos ($\text{pH} = 11.5$), lo que permite un control estricto de los estados de protonación y elimina una fuente de incertidumbre presente en el modelado más desafiante de los sistemas proteína-ligando [48].

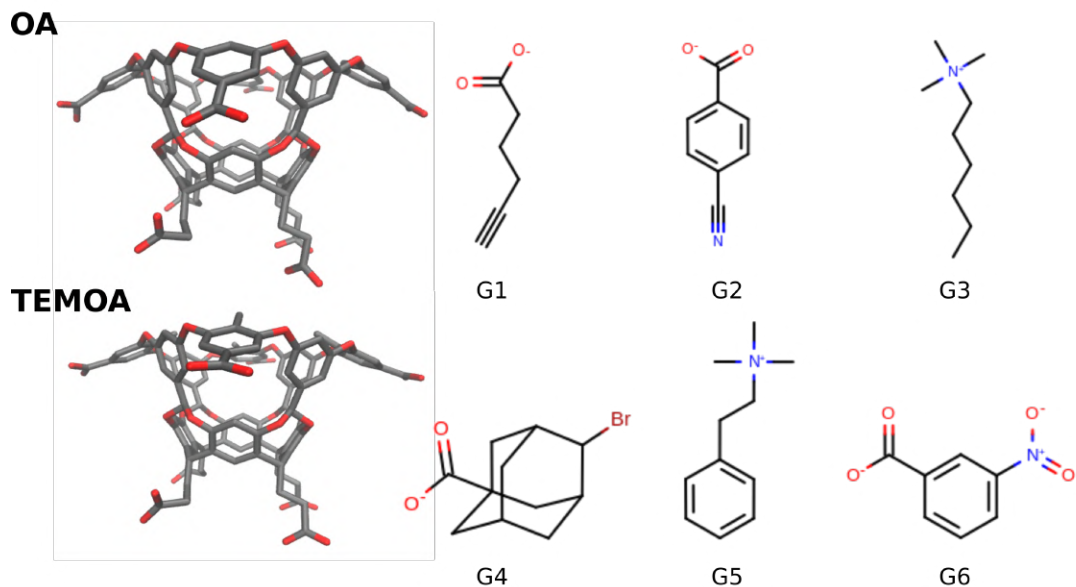


Figura 1.2: Estructuras de los anfitriones OA-TEMOA y huéspedes del SAMPL5. Se muestran los microestados más probables de las moléculas anfitrión-huésped de acuerdo al pH experimental y valores de pK_a [1] (ver Tabla A.1 en el Anexo A). Las energías libres de unión de todos los sistemas fueron medidas bajo las mismas condiciones experimentales: pH 11.5, Temperatura de 298 K y disolución reguladora de pH de fosfato de sodio 50 mM.

1.2.2. Anfitrión CB7 y huéspedes con núcleo aromático o de adamantano

La familia de cucurbit[n]urilos forma una estructura altamente simétrica de monómeros de glicolurilo unidos covalentemente CB[n], donde $n = 5, 6, 7, 8, 10$ indica la cantidad de monómeros presentes en el anfitrión. A medida que aumenta la unidad de glicolurilo se acomodan huéspedes de mayor tamaño progresivamente, pero son consistentes en preferir unir a huéspedes con un núcleo hidrofóbico que encaje en la cavidad interior relativamente no

polar [49]. La cavidad interior, sin embargo, muestra un momento cuadrupolar muy alto que permite interacciones electrostáticas de orden superior [45]. Los anfitriones CB[n] se unen a los huéspedes con un núcleo hidrofóbico y al menos un resto catiónico que interactúa con los átomos de oxígeno de los grupos carbonilo que rodean ambos portales del anfitrión [50]. CB6 es el más antiguo de la familia y ha sido ampliamente estudiado por el trabajo pionero de Mock [51, 52] y Kim [53]. En 2005, el grupo de Isaacs realizó un estudio comparativo de la afinidad de unión de los mismos huéspedes a los anfitriones CB6, CB7 y CB8. En ese estudio estableció que CB7 y CB8 retienen gran parte de las notables características de unión de los complejos CB6 [50]. En contraste con CB8, el anfitrión CB7 exhibe una selectividad muy alta para los huéspedes con núcleo de adamantano [50].

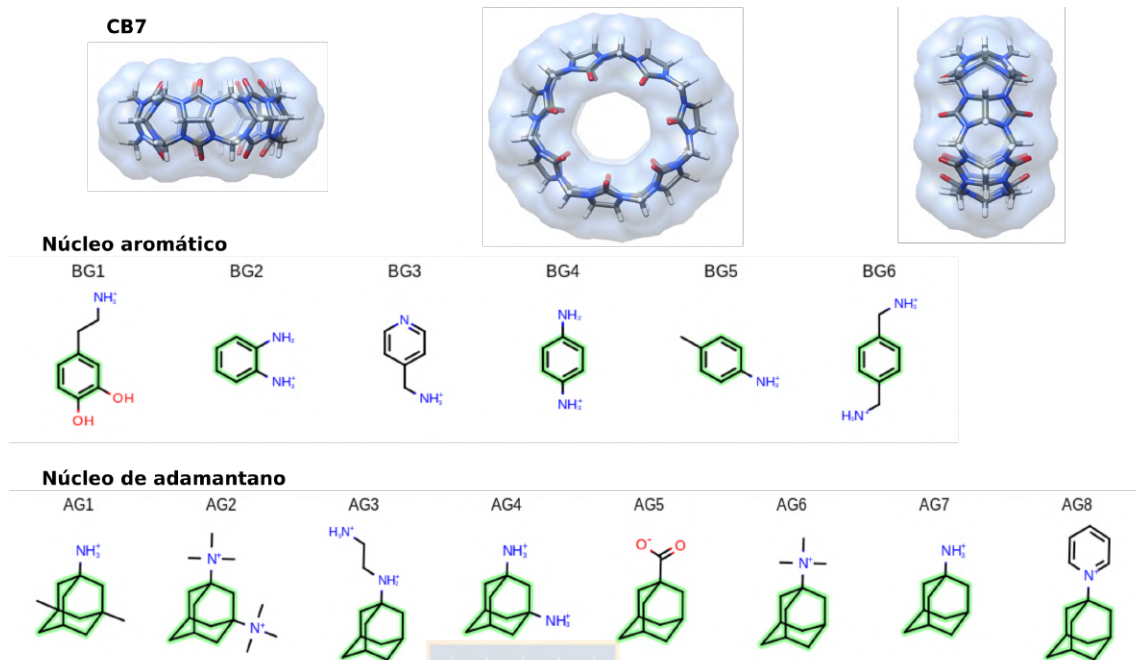


Figura 1.3: Estructuras del anfitrión CB7 y huéspedes con núcleo aromático y adamantano. Se muestran los microestados más probables de los huéspedes de acuerdo al pH experimental 4.74 y valores de pK_a (ver Tabla B.1 en el Anexo B). Las energías libres de unión de todos los sistemas fueron medidas bajo las mismas condiciones experimentales: pH 4.7, Temperatura de 298 K y solución reguladora de pH de acetato de sodio 50 mM.

En esta investigación se estudió el CB7 para evaluar la precisión de los parámetros no enlazantes del campo de fuerza y dos series de huéspedes con núcleo aromático o de adamantano (ver Figura 1.3). CB7 ha sido incluido en varias iniciativas de SAMPL [8, 20] para la predicción de afinidad con una amplia serie de huéspedes. Con núcleo adamantano las energías libres de unión superan incluso a los sistemas proteína-ligando [50, 49, 54, 55]. Las afinidades de unión medidas entre el CB7 y huéspedes con núcleo aromático

y de adamantano cubren un rango de -6,3 a -12,6 kcal/mol y -6,0 a -18,2 kcal/mol, respectivamente [50].

1.3. Sistemas proteína-ligando

Una comprensión profunda de la unión entre una proteína con un ligando impactará en la capacidad de diseñar moléculas con finalidad terapéutica. Un esfuerzo mundial en biología estructural ha llevado a la elucidación de las estructuras de un gran número de sistemas proteína-ligando con resolución atómica, disponibles en Protein Data Bank (PDB) [56]. Además, gracias a esfuerzos experimentales entre la academia y la industria bio-farmacéutica se ha logrado la caracterización de numerosos ligandos unidos a un diverso rango de proteínas. Esto ha permitido proponer simulaciones computacionales para predecir energías libres de unión a partir de estructuras 3D y compararlas frente a las energías libres de unión determinadas experimentalmente [57]. Pese a todos los desarrollos recientes en esta área, la predicción precisa de la energía libre de unión basada en simulaciones computacionales han demostrado ser un gran desafío. Dentro de los desafíos más relevantes se destacan:

- Desafíos de muestreo: Las cadenas laterales del sitio de unión pueden sufrir cambios conformacionales cuando se unen a un ligando. Nor-

malmente estos rearrreglos son lentos y pueden presentar problemas de muestreo para las simulaciones de dinámica molecular. Además, también puede presentarse cambios conformacionales a gran escala, como movimientos de dominio y alostería. Los modos de unión del ligando pueden cambiar con pequeñas modificaciones del ambiente molecular, y/o la presencia de moléculas de agua pueden variar dependiendo del ligando al que se unen.

- **Desafíos del sistema:** Descripción de los posibles cambios en los estados de protonación del ligando y/o residuos durante el proceso de unión. Además, la energía libre de unión es sensible a las condiciones como disoluciones reguladoras de pH, sales y otros factores.
- **Desafíos del campo de fuerza:** Limitación en la descripción adecuada de los parámetros no enlazantes, fenómenos de polarización, interacción del ligando con iones metálicos y algunos grupos funcionales los cuales no cuentan con parámetros en los campos de fuerza.

Los sistemas de referencia para abarcar los desafíos existentes en sistemas proteína-ligando, involucran proteínas pequeñas y sitios de unión simples. En este sentido la lisozima T4 es un sistema proteico referente en los cálculos de energía libre de unión. Al igual que los sistemas anfitrión-huésped la comunidad científica ha concentrado esfuerzos para avanzar en el entendimiento de los procesos de unión entre una proteína y un ligando (compuestos simi-

lares a los fármacos). Los resultados obtenidos son una herramienta bastante prometedora para acelerar el descubrimiento de medicamentos nuevos y reducir el costo del proceso de descubrimiento de fármacos. En este sentido se han realizado una serie de desafíos a ciegas denominados D3R (Drug Design Data Resource) que tienen como objetivo avanzar en tecnologías de acoplamiento, puntuación y métodos de energía libre. Sin embargo, aún existe la necesidad de mejorar los métodos para predecir las poses del ligando y afinidades [58].

1.3.1. Lisozima T4 L99A/M102Q y ligandos con núcleo aromático



Dentro de los sistemas proteicos la lisozima T4 constituye un referente en los estudios de afinidad proteína-ligando, ya que es una proteína relativamente pequeña y ha sido ampliamente estudiada desde el punto de experimental y computacional. Para la lisozima T4 se han reportado particularmente dos sitios de unión denominados L99A asociado con la mutante Leucina 99 a Alanina (cavidad apolar) [59] y L992A/M102Q asociado con la doble mutante Leucina 99 a Alanina/Metionina 102 a Glutamina (cavidad con mayor polaridad respecto a L99A) [60]. El sitio de unión en la cavidad L99A/M102Q (ver Figura 1.4) genera una mutación puntual a lo largo de un borde del sitio de unión, proporcionando una glutamina que introduce polaridad y potencial para enlace de hidrógeno, permitiendo así unirse con una mayor variedad

de grupos funcionales. Como sistema modelo proteína-ligando en el presente proyecto se evaluará la energía libre de unión con diferentes ligandos del sitio de unión L99A/M102Q de la lisozima T4 (Ver Figura 1.4).

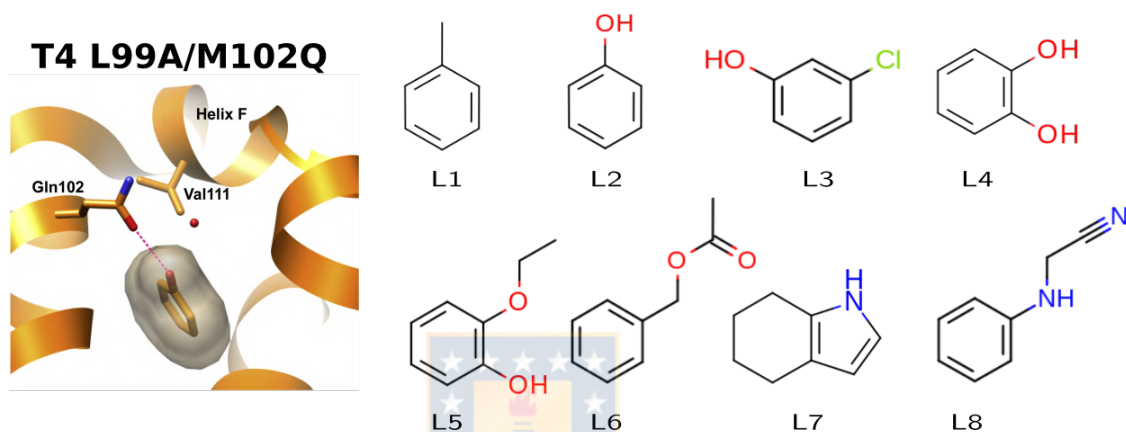


Figura 1.4: Izquierda -Complejo de lisozima T4 L99A/M102Q-fenol. Se muestra los residuos GLN102, VAL111 y la superficie del ligando. Derecha -Ligandos con los estados de protonación más probable de acuerdo al pH experimental 6.8 y valores de pK_a (ver Tabla C.1 en el Anexo C). Las energías libres de unión de todos los sistemas fueron medidas bajo las mismas condiciones experimentales pH 6.8, Temperatura de 298 K y solución reguladora de pH de fosfato de potasio 50 mM o cloruro de potasio 200 mM.

1.4. Hipótesis y Objetivos

Hipótesis

Las cargas atómicas dinámicas (D-MBIS) derivadas de la densidad electrónica molecular describen adecuadamente las interacciones no enlazantes de tipo coulombicas en sistemas anfitrión-huésped y proteína-ligando. El uso de estas cargas en campos de fuerza no polarizables junto con la corrección por la energía de polarización derivadas en el ambiente de unión y no unión conducirán a una mejor predicción de la energía libre de unión en los sistemas de interés.



Objetivo general

Determinar la energía libre de unión para los sistemas referencia anfitrión-huésped y proteína-ligando utilizando cargas atómicas dinámicas derivadas de la densidad electrónica molecular.

Objetivos Específicos

1. Calcular las cargas atómicas dinámicas D-MBIS para los huéspedes y ligandos referencia en el ambiente del solvente.
2. Determinar las cargas atómicas dinámicas D-MBIS para los huéspedes

y ligandos referencia en el ambiente de unión (anfitrión o proteína).

3. Calcular la energía de polarización para el huésped y ligando en el ambiente solvatado (estado de no unión) y en el ambiente de unión.
4. Realizar cálculos de energía libre de unión empleando las cargas atómicas dinámicas D-MBIS para los sistemas referencia anfitrión-huésped.
5. Realizar cálculos de energía libre de unión empleando las cargas atómicas dinámicas D-MBIS para la lisozima T4 L99A/M102Q-ligandos.



2. Metodología

En esta sección se muestran las principales metodologías empleadas en el desarrollo de la investigación. Se presentan algunos conceptos generales relacionados con la derivación de los parámetros no enlazantes del campo de fuerza considerando la densidad electrónica molecular, la dinámica molecular y el método de partición MBIS. Además, se mencionan algunos aspectos relevantes para las predicciones de energía libre de unión.

2.1. Métodos de estructura electrónica

Los métodos de estructura electrónica usan las leyes de la mecánica cuántica para describir explícitamente los electrones y núcleos del sistema mediante la solución de la ecuación de Schrödinger independiente del tiempo:

$$\hat{H}\psi_i = E_i\psi_i \quad (2.1)$$

donde \hat{H} es el operador Hamiltoniano, que opera sobre la función de onda ψ_i y nos entrega el valor propio de la energía E_i del sistema. El Hamiltoniano \hat{H} puede ser separado en dos componentes bajo la aproximación de Born-

Oppenheimer [61, 62], donde se considera que los electrones se mueven en el campo generado por núcleos fijos. Esta aproximación permite separar el término de energía cinética de los núcleos al igual que la correlación del potencial atractivo electrón-núcleo, y la repulsión entre núcleos se considera una constante para una posición fija de estos. De este modo, el Hamiltoniano de la ecuación 2.1 se transforma en usando unidades atómicas:

$$\hat{H}_{el} = - \sum_{i=1}^N \frac{1}{2} \nabla_i^2 - \sum_{i=1}^N \sum_{A=1}^M \frac{Z_A}{r_{iA}} + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{r_{ij}} \quad (2.2)$$

Así, la ecuación de Schrödinger electrónica se puede escribir como:

$$(\hat{H}_{el} + V_N) \psi_{el}(r_i; R_A) = E_{el} \psi_{el}(r_i; R_A) \quad (2.3)$$

Sólo se conocen soluciones exactas de la ecuación 2.3 para sistemas hidrogenoides. Los sistemas poli-electrónicos, no tiene solución analítica. Esto debido a los términos de repulsión entre electrones que aparecen en el Hamiltoniano electrónico. Sin embargo, se proponen métodos matemáticos aproximados, denominados métodos perturbacionales y variacionales, que permiten obtener soluciones a la ecuación de Schrödinger electrónica bastante precisas.

Dentro de los métodos aproximados, uno de lo más empleados para realizar estudios de estructura electrónica es la teoría del funcional de la densidad (DFT) [63]. DFT se fundamenta en que toda la información está contenida en la densidad electrónica $\rho(\vec{r})$. También tiene la ventaja que solamente depende

de tres coordenadas espaciales, a diferencia de la función de onda que depende de $3n$ coordenadas espaciales, donde n representa el número de electrones. La metodología de DFT se basa en dos teoremas desarrollados por Hohenberg y Kohn. El primero de los teoremas prueba que la energía de un sistema electrónico sometido a un potencial externo es un funcional de la densidad electrónica. Por otra parte, el segundo teorema establece que puede obtenerse variacionalmente la densidad electrónica que minimiza el funcional de energía y esta corresponde a la densidad del estado fundamental.

Método de Kohn-Sham

El método propuesto por Kohn-Sham [64] considera un sistema electrónico de referencia no interactuante, cuya densidad electrónica del estado fundamental es igual a la del sistema real, según la relación:

$$E[\rho] = T_s[\rho] + V_{ne}[\rho] + J[\rho] + E_{xc}[\rho] \quad (2.4)$$

donde $T_s[\rho]$ corresponde a la energía cinética del sistema electrónico no interactuante, $V_{ne}[\rho]$ es la energía de interacción núcleo-electrón, $J[\rho]$ representa la energía de interacción electrón-electrón y $E_{xc}[\rho]$ corresponde al término de intercambio y correlación, que puede expresarse como:

$$E_{xc}[\rho] = (T[\rho] - T_s[\rho]) + (V_{ee}[\rho] - J[\rho]) \quad (2.5)$$

donde $T[\rho]$, V_{ee} corresponden a la energía cinética del sistema electrónico y al potencial interelectrónico respectivamente. Esta aproximación aprovecha el conocimiento de una expresión exacta para la densidad electrónica en términos de orbitales:

$$\rho(\vec{r}) = \sum_i^n |\phi_i(\vec{r})|^2 \quad (2.6)$$

donde $\phi(\vec{r})$ corresponde a los orbitales monoelectrónicos ortonormales, que son soluciones de la siguiente ecuación:

$$\left[-\frac{1}{2}\nabla^2 + v_s(\vec{r}) \right] \phi_i(\vec{r}) = \varepsilon_i \phi_i(\vec{r}) \quad (2.7)$$

donde $v_s(\vec{r})$ es el denominado potencial de Kohn-Sham. Este corresponde a la suma del potencial externo, es decir, el potencial coulombico de los núcleos, el potencial electrostático clásico (también conocido como potencial de Hartree) y el potencial de intercambio y correlación, que no es más que la derivada de $E_{xc}[\rho(\vec{r})]$ respecto $\rho(\vec{r})$:

$$v_s(\vec{r}) = V_{ext}(\vec{r}) + \int \frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r}' + v_{xc}([\rho]; \vec{r}) \quad (2.8)$$

$$v_{xc}([\rho]; \vec{r}) = \frac{\delta E_{xc}}{\delta \rho(\vec{r})} \quad (2.9)$$

Resolviendo la ecuación 2.7 se pueden obtener los orbitales $\phi_s(\vec{r})$ que pro-

porcionan la densidad electrónica deseada, donde cada uno de estos orbitales es una combinación lineal de funciones base. En el método de Kohn-Sham, no se conoce la expresión exacta del potencial de intercambio-correlación $v_{xc}([\rho]; \vec{r})$. Por lo tanto, se hace necesario emplear potenciales aproximados (y en algunos casos parámetros ajustados) que permitan incluir su contribución al potencial externo.

El funcional de correlación-intercambio

Actualmente, no existe una forma de obtener la expresión exacta para el funcional de intercambio y correlación. Por tanto, se usan aproximaciones basadas en soluciones de sistemas sencillos e incluso el uso de parámetros.

Las aproximaciones de la densidad local (LDA) son una clase de aproximación al funcional de intercambio-correlación. Estas dependen únicamente del valor de la densidad electrónica en cada punto del espacio y no de derivadas de la densidad o de los orbitales. Para un sistema sencillo, como un gas de electrones uniforme, se puede encontrar una solución casi exacta empleando la siguiente expresión:

$$E_{xc}^{LDA}[\rho] = \int \epsilon_{xc}^{LDA}[\rho] \rho(\vec{r}) d\vec{r} \quad (2.10)$$

La energía de intercambio y correlación puede considerarse como la suma de la energía de intercambio y correlación.

$$E_{xc}^{LDA}[\rho] = \varepsilon_x^{LDA}[\rho] + \varepsilon_c^{LDA}[\rho] \quad (2.11)$$

donde ε_x^{LDA} representa la energía de intercambio generada por el funcional de intercambio [65, 66] y ε_c^{LDA} representa la energía de correlación generada por el funcional de correlación. Debido a su complejidad, este último se ha calculado a través métodos numéricos. Los siguientes funcionales pueden ocuparse para determinar la energía de intercambio y correlación respectivamente:

$$E_x^{LDA}[\rho] = -\frac{3}{4} \left(\frac{3\rho(\vec{r})}{\pi} \right)^{1/3} \quad (2.12)$$

$$E_c^{VWN}[\rho] = A \left[\ln \frac{x^2}{X(x)} + \frac{2b}{Q} P - \frac{bx_0}{X(x_0)} \left(\ln \frac{(x-x_0)^2}{X(x)} + \frac{2(b+2x_0)}{Q} P \right) \right] \quad (2.13)$$

En la ecuación 2.13 [67] A es una constante fijada por la expansión de alta densidad, x_0 , b , c corresponden a parámetros obtenidos por el método de mínimos cuadrados respecto a los valores reportados por *Ceperley y Alder* de

la energía de correlación [67], x , Q , $X(x)$, P corresponden a:

$$x = \left(\frac{3}{4\pi\rho(\vec{r})} \right)^{1/6} \quad X(x) = x^2 + bx + c$$

$$Q = (4c - b^2)^{1/2} \quad P = \arctan \frac{Q}{2x + b}$$

Al combinar ambos funcionales, se obtiene la aproximación de la densidad local (LDA). De esta forma, cada funcional depende únicamente del valor de la densidad en cada punto. LDA asume que la densidad es homogénea. Debido a esto, LDA tiende a subestimar la energía de intercambio y sobreestimar la energía de correlación. Para corregir esta tendencia, es común expandir en términos del gradiente de densidad (*Generalised Gradient Approximations*, GGA) para tener en cuenta la falta de homogeneidad de la densidad electrónica real. Dentro de los funcionales GGA se destacan:

- Funcional de intercambio de Becke [68]:

$$E_x^{B88}[\rho] = E_x^{LDA}[\rho] - \beta \int \rho^{4/3} \frac{x^2}{1 + 6\beta x \text{sen}^{-1} h x} d\vec{r} \quad (2.14)$$

Donde x y β corresponde a:

$$x = \frac{\nabla\rho}{\rho^{4/3}} \quad \beta = 0,0042$$

- Funcional de correlación de Lee-Yang-Parr [69]:

$$E_x^{LYP}[\rho] = -a \int \frac{\rho}{1+d\rho^{-1/3}} dr - ab \int \omega \rho^2 \left[\frac{3}{10} (3\pi^2)^{2/3} \rho^{8/3} - \left(\frac{1}{24} + \frac{7}{72} \delta \right) |\nabla \rho|^2 \right] d\vec{r} \quad (2.15)$$

Donde ω , δ , a , b , c y d corresponden a:

$$\omega = \frac{e^{-c\rho^{-1/3}}}{1+d\rho^{-1/3}} \rho^{-11/3} \quad \delta = c\rho^{-1/3} + \frac{d\rho^{-1/3}}{1+d\rho^{-1/3}}$$

$$a = 0,04918 \quad b = 0,132 \quad c = 0,2533 \quad d = 0,349$$



- Funcional híbrido de tres parámetros de Becke [70]:

$$\begin{aligned} E_{xc}^{B3} &= E_{xc}^{LDA} + a_0(E_x^{HF} - E_x^{LDA}) + a_x(E_x^{B88} - E_x^{LDA}) + a_c(E_c - E_c^{VWN}) \\ &= a_0 E_x^{HF} + a_x E_x^{B88} + (1 - a_0 - a_x) E_x^{LDA} + a_c E_c + (1 - a_c) E_c^{VWN} \end{aligned} \quad (2.16)$$

Donde $a_0 = 0,20$, $a_x = 0,72$ y $a_c = 0,81$. Mediante el método de HF, se calcula el funcional de intercambio exacto E_x^{HF} usando los orbitales de Kohn-Sham. Para el cálculo E_c , se puede incluir un funcional de correlación, siendo uno de

los más utilizados el funcional E_c^{LYP} que da como resultado el método llamado B3LYP, uno de los funcionales híbridos más comunes. Para la aplicación en las distintas áreas de la química los métodos DFT han demostrado proveer resultados más exactos ya que consideran los efectos de la correlación e intercambio a un bajo costo computacional. Por tanto, las derivaciones de las densidades electrónicas en la presente investigación consideran el funcional híbrido B3LYP.

2.2. Mecánica molecular

La mecánica molecular se basa en las leyes de la mecánica clásica aplicada al estudio de sistemas moleculares, en los cuales sus interacciones se modelan mediante potenciales analíticos sencillos. Las funciones que describen las interacciones moleculares se les denomina campo de fuerza y engloban dos componentes. El primer componente ayuda a definir la estructura covalente de la molécula, el cual está asociado con la interacción entre átomos vecinos. Este componente incluye contribuciones debidas a uniones covalentes, ángulos de enlace y torsiones propias e impropias (parámetros enlazantes especificados en la sección 2.2.1.). El segundo componente describe la interacción no covalente, generadas por las interacciones entre las distribuciones electrónicas que rodean los diferentes átomos. A cortas distancias, las interacciones son principalmente de repulsión debido a la superposición entre las

nubes electrónicas de cada átomo. A largas distancias, hay diferentes tipos de interacciones entre las cuales, se pueden destacar dos. En primer lugar, están las interacciones de dispersión, producidas por las fluctuaciones en las distribuciones de carga de los grupos. En segundo lugar, están las interacciones electrostáticas, que surgen de la interacción de las distribuciones de carga sobre cada molécula o parte de ella. En los campos de fuerza, generalmente se describen las interacciones de dispersión y repulsión por el potencial de Lennard-Jones y las interacciones electrostáticas con un término coulombico (parámetros no enlazantes especificados en la sección 2.2.2.).

$$E_{enlazante} = E_{enlace} + E_{ángulos} + E_{diedros} + E_{impropios} \quad (2.17)$$

$$E_{no-enlazante} = E_{electrostática} + E_{van-der-Waals} \quad (2.18)$$

La mayoría de los campos de fuerza clásicos, usan cargas derivadas de cálculos de estructura electrónica, destacándose los campos de fuerza CHARMM [10], AMBER-GAFF [11], ECEPP [71], GROMOS [12], entre otros. Dependiendo del campo de fuerza empleado, las cargas atómicas varían debido a que ocupan diferentes estrategias para su derivación. Por ejemplo el campo de fuerza CHARMM utiliza las cargas Merz-Kollman [72, 73] obtenidas de cálculos mecano-cuánticos a nivel MP2/6-31G* [74] y el campo de fuerza AMBER99 utiliza las cargas RESP las cuales se obtienen a partir de cálculos

mecano-cuánticos a nivel HF/6-31G* [11, 75]. Existen también campos de fuerza que emplean cargas atómicas empíricas, que se han obtenido principalmente del ajuste para reproducir algunas propiedades de los líquidos. Tal es el caso del campo OPLS [13] y GROMOS [12]. Las cargas atómicas empíricas también varían de acuerdo a la estrategia usada para derivarlas. En el caso específico de OPLS, se emplean cargas atómicas empíricas ajustadas para reproducir las propiedades de los líquidos orgánicos como la entalpía de vaporización, las capacidades caloríficas y la densidad. GROMOS emplea cargas atómicas ajustadas para reproducir las entalpías de hidratación o solvatación y el calor de vaporización en fase condensada.

El uso de estos campos de fuerza conlleva a una descripción inadecuada para algunas aplicaciones, especialmente en casos donde las moléculas están polarizadas. En un sistema molecular real la polarización juega un rol significativo, ya que las cargas pueden fluctuar debido a los reordenamientos moleculares que exhiben los diferentes sistemas. Esto ha motivado a la comunidad científica para desarrollar campos de fuerza que incluyan la polarización. En algunos casos, se han creado extensiones de los campos de fuerza clásicos y en otros se han diseñado nuevos campos de fuerza que incorporan un término de polarización desde sus inicios. Este es el caso de AMOEBA [76], SIBFA [15], SDFF [16], NEMO [17], entre otros. La inclusión del término de polarización en el campo de fuerza permite obtener simulaciones más precisas respecto a los observables de interés, sin embargo, precisan de un mayor

costo computacional y emplean mayor cantidad de parámetros.

Bajo esta panorámica en la presente investigación se derivaron nuevos parámetros no enlazantes que consideren las fluctuaciones de la densidad electrónica molecular en el proceso de unión no covalente. Estos nuevos parámetros junto con el costo energético de polarizar la densidad electrónica se espera que mejoren las predicciones de energía libre de unión en los sistemas anfitrión-huésped y proteína-ligando de interés. En las siguientes subsecciones se describen en más detalle los parámetros enlazantes y no enlazantes del campo de fuerza, haciendo hincapié en las diferentes estrategias empleadas para describir las cargas atómicas y los parámetros de Lennard-Jones.

2.2.1. Parámetros enlazantes

- Energía de tensión de los enlaces:

Para describir la energía de tensión de los enlaces es posible recurrir a un potencial armónico mediante la ecuación de Hooke:

$$E_{enlace} = \sum_{enlace} \frac{1}{2} k_r (r_{ij} - r_{eq})^2 \quad (2.19)$$

La sumatoria se extiende sobre todos los enlaces que se han definido en el sistema. k_r es la constante de fuerza asociada al estiramiento de los átomos i y j , r_{ij} es la distancia de enlace para una geometría particular y r_{eq} es la distancia donde el potencial de enlace es cero.

- Energía de flexión:

La energía de flexión producida por dos enlaces con un átomo en común los cuales forman un ángulo y también puede ser representada por una expresión tipo oscilador armónico:

$$E_{\text{ángulos}} = \sum_{\text{ángulos}} \frac{1}{2} k_{\theta} (\theta_{ijk} - \theta_{eq})^2 \quad (2.20)$$

La sumatoria se extiende sobre todos los ángulos que se han definido en el sistema. k_{θ} es la constante de fuerza θ_{ijk} es el ángulo de enlace de una de una geometría particular y r_{eq} es el ángulo donde el potencial formado por los ángulos es mínimo.

- Energía de ángulos diedros:

Para cada ángulo en una molécula existe un plano, cuando dos ángulos tienen un enlace en común forman un ángulo de torsión ϕ_{ijkl} . Para la energía del ángulo diedro se utiliza una función periódica sinusoidal que es continua para todo ángulo. La energía de los ángulos diedros puede expresarse como:

$$E_{\text{diedros}} = \sum_{\text{diedros}} \frac{V_n}{2} [1 + \cos(n\phi_{ijkl} - \gamma)] \quad (2.21)$$

La suma es sobre todos los ángulos diedros que se definen en el sistema, V_n es la barrera potencial de cada término de la función, n es la

periodicidad del ángulo que determina el número de picos y pozos en el potencial, ϕ_{ijkl} es el ángulo de torsión de una cierta geometría y γ es la fase del ángulo.

- Energía de ángulos impropios:

Está asociada con la energía de restauración que impide que el grupo se mueva fuera del plano. La forma común de representar un ángulo diedro impropio es similar a los ángulos diedros, sin embargo, el ángulo formado por ω_{ijkl} no son átomos unidos contiguamente, sino átomos unidos formando una estructura triangular.

$$E_{impropios} = \sum_{impropios} \frac{V_n}{2} [1 + \cos(n\omega_{ijkl} - \gamma)] \quad (2.22)$$

2.2.2. Parámetros no enlazantes

- Electrostático:

Las interacciones electrostáticas pueden representarse por la ley de Coulomb mediante las cargas atómicas interactuantes. Las cargas atómicas pueden variar dependiendo del campo de fuerza utilizado, debido a que se emplean diferentes metodologías para extraer las cargas atómicas. El término electrostático se representa mediante la siguiente expresión.

$$E_{electrostático} = \sum_{j>i} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \quad (2.23)$$

donde q_i y q_j representan las cargas utilizadas por el campo de fuerza para cada tipo de átomo, r_{ij} indica la distancia que existe entre el átomo i y el átomo j y ϵ_0 es la constante dieléctrica del medio.

- van-der-Waals :

Este término es usado para evitar la fusión de átomos no unidos con cargas opuestas como también para representar interacciones atractivas a corta distancia producto de fuerzas dispersivas. Este término puede ser representado por un potencial de Lennard-Jones el cual está compuesto por una suma de dos contribuciones. La primera es una contribución repulsiva para las distancias cortas entre los átomos y una parte atractiva para distancias largas. La expresión que representa este término es la siguiente:

$$E_{van\ der\ Waals} = \sum_{j>i} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] \quad (2.24)$$

donde la suma se realiza sobre todos los átomos con $j \neq i$, ϵ_{ij} es la profundidad del potencial, σ_{ij} es la distancia finita en que el potencial entre partículas es cero y r_{ij} es la distancia entre partículas. En la ecuación 2.24 el término $1/r_{ij}^{12}$, se refiere a la parte repulsiva del potencial que

describe la repulsión de Pauli a cortas distancias debido a la superposición de nubes electrónicas, y $1/r_{ij}^6$ corresponde a la parte atractiva de dispersión.

2.2.2.1. Cargas atómicas

Existe una amplia variedad de cargas atómicas, usualmente derivadas mediante protocolos empíricos y mecano-cuánticos. Las cargas atómicas pueden hacer contribuciones dominantes en la energía potencial, especialmente en moléculas polares y compuestos iónicos. Además, son críticas para simular la geometría, energía de interacción y la reactividad. En todos los campos de fuerza, las cargas atómicas son utilizadas para determinar las interacciones electrostáticas representadas por la ecuación de Coulomb 2.23. A continuación, se describe con más detalles algunos aspectos relevantes considerados en la investigación respecto a las cargas atómicas.

Cargas atómicas ESP

Las cargas atómicas denominadas *Electrostatic Potential* (ESP) [77], son obtenidas mediante cálculos de estructura electrónica. El potencial electrostático se obtiene para una unidad positiva de carga en un punto r , para una región determinada de la molécula. Este es calculado de acuerdo a la ecuación:

$$V_i^0(r) = \sum_A \frac{Z_A}{|r - R_A|} - \sum_{mn} P_{mn} \int \frac{\phi_m \phi_n}{|r - r'|} dr \quad (2.25)$$

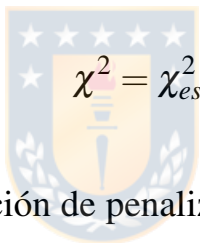
donde el primer término es la sumatoria sobre todos los núcleos y el segundo término sobre todas las funciones base utilizadas para el cálculo. P_{mn} es el elemento de matriz mn asociado a la matriz de densidad de la función de onda para la molécula, obtenida a través del cálculo de estructura electrónica; ϕ_m y ϕ_n son funciones base utilizadas. Al obtener el potencial electrostático de un cálculo de estructura electrónica, se realiza un ajuste de mínimos cuadrados tratando de reproducir el potencial obtenido por la ecuación 2.25, utilizando cargas atómicas centradas en los núcleos. El potencial obtenido se evalúa a través de una grilla de puntos, según la ecuación 2.26, donde w_i es un factor estadístico de peso por cada punto en la grilla. Para distintos modelos de carga ESP, la descripción de esta función es variable.

$$R = \sum_i^m w_i(r) \left[V_i^0(r) - \sum_j^{n-1} q_j r_{ij}^{-1} \right]^2 \quad (2.26)$$

Cargas atómicas RESP

El enfoque de ajuste ESP depende del conjunto de bases utilizados. Además, la carga de los átomos lejanos a la superficie en donde se evalúa el potencial, no es característica debido a la baja sensibilidad del potencial

en puntos lejanos a la carga. Este problema se asocia generalmente a átomos de carbono. Como consecuencia de esta problemática, se producen valores de carga atómica no transferibles entre grupos funcionales en moléculas relacionadas. Para abordar este punto frágil de las cargas ESP, Bayly *et al.* [33] presentaron una variante para calcular cargas basadas en el potencial electrostático, introduciendo una función de penalización (ecuación 2.27) dentro del proceso de ajuste para obtener las cargas atómicas denominadas RESP.


$$\chi^2 = \chi_{esp}^2 + \chi_{rstr}^2 \quad (2.27)$$

Al introducir esta función de penalización sobre los átomos de carbono para calcular las cargas, se logra reproducir el potencial electrostático molecular y se reduce el problema de cargas intransferibles entre grupos funcionales similares. Sin embargo, las funciones de penalización carecen de significado físico, lo que supone uno de los principales problemas de las cargas RESP.

La principal fortaleza de estas cargas derivadas del potencial electrostático, es que se reproducen ciertas propiedades de interacción intermolecular con un simple potencial aditivo compuesto por dos partes. La elección inicial de la función de penalización fue un potencial armónico simple:

$$\chi_{rstr}^2 = a \sum_j ((q_{0j} - q_j)^2) \quad (2.28)$$

donde a es el factor de escala que determinar la fuerza de la restricción y q_0 es la carga objetivo de la restricción.

Cargas atómicas AM1-BCC

El método de carga atómica AM1-BCC [78] ajusta un conjunto inicial de cargas AM1 que no emulan el potencial electrostático ESP HF/6-31G* alrededor de una molécula, en un nuevo conjunto de cargas que, si lo hacen, usando términos de corrección aditiva simple aplicado a las cargas AM1. Este modelo de cargas genera así rápidamente cargas atómicas de alta calidad adecuadas para las simulaciones en sistemas condensados.

La carga AM1-BCC $q_{AM1-BCC}^j$ para un átomo j se obtiene sumando la corrección de carga neta q_{BCC}^j a la carga atómica de población AM1 q_{AM1}^j :

$$q_{BCC}^j = q_{AM1}^j + q_{BCC}^j \quad (2.29)$$

La corrección de carga neta q_{BCC}^j depende sólo de los vecinos inmediatos enlazados al átomo j y está dado por:

$$q_{BCC}^j = \sum_{\alpha=1}^{nb} T_{j\alpha} B_{\alpha} \quad (2.30)$$

donde $T_{j\alpha}$ es la matriz de conectividad de enlace, B_{α} es la corrección de carga para el tipo de enlace α , y la suma se ejecuta en todos los átomos vecinos unidos al átomo j . En general las cargas atómicas AM1-BCC son usadas con bastante frecuencia en cálculos de energía libre de unión, sin embargo, tienen la desventaja de incurrir en errores al extrapolarlas a otros sistemas químicos más complejos.



Cargas atómicas de Hirshfeld

Las cargas atómicas de Hirshfeld se basan en el método de átomos en moléculas (AIM), el cual plantea que es posible obtener las cargas atómicas mediante la partición de la densidad electrónica molecular [34]. En primer lugar, se obtiene una función de densidad pro-molecular, la cual corresponde a la sumatoria de las funciones de densidad de los átomos aislados constituyentes de la molécula para cierto radio r , según la siguiente expresión:

$$\rho^{pro}(r) = \sum_i \rho_i^{at}(r) \quad (2.31)$$

donde $\rho_i^{at}(r)$ es la función de densidad del átomo aislado, la cual se

determina para todos los tipos de átomos que forman la molécula. Luego, se define un factor de peso o contribución $w_i(r)$, que da cuenta del aporte de la densidad del pro-átomo $\rho_i^{at}(r)$ a la densidad pro-molecular $\rho^{pro}(r)$ mediante la siguiente expresión:

$$w_i(r) = \frac{\rho_i^{at}(r)}{\rho^{pro}(r)} \quad (2.32)$$

Una vez obtenido el factor de peso $w_i(r)$, se determina la densidad AIM. Esta considera cuánta densidad molecular real $\rho^{mol}(r)$ se le puede asociar a cada átomo, según la ecuación:

$$\rho_i^{AIM}(r) = w_i(r)\rho^{mol}(r) \quad (2.33)$$

Por último, considerando el número atómico, se obtiene la carga de Hirshfeld bajo la siguiente expresión:


$$q_i = Z_i - \int \rho^{AIM}(r) dr \quad (2.34)$$

Existen algunos problemas con las cargas de Hirshfeld, entre los cuales se destacan: las cargas son muy pequeñas y tienden a ser prácticamente cero, se usan pro-átomos fijos (densidades promediadas esféricamente de átomos neutros aislados) y además no reproducen el potencial elec-

trostático molecular.

Cargas atómicas de Hirshfeld-Iterativas

Las deficiencias del método de Hirshfeld motivó para que Bultinck *et al.* [79] desarrollaran un método iterativo para obtener las cargas de Hirshfeld, el cual denominaron cargas atómicas de Hirshfeld-I (HI). El método se basa en la actualización de la carga del pro-átomo hasta hacerla lo más parecida a la carga AIM.


$$W_A^i(r) = \frac{\rho_A^{i-1}(r)}{\rho_{mol}^{i-1}(r)} \quad (2.35)$$

En el esquema HI, los pro-átomos son interpolaciones lineales entre densidades de iones aislados, de modo que las poblaciones AIM y las poblaciones pro-atómicas coinciden. Las cargas de HI obtenidas no son cercanas a cero, no dependen de la elección del pro-átomo, tienen baja dependencia con la conformación utilizada y reproducen el potencial electrostático molecular en vacío.

Cargas atómicas MBIS

Las cargas MBIS (*Minimal Basis Iterative Stockholder*) fueron propuestas por Toon Verstraelen [36]. Las cargas MBIS utilizan una definición

distinta para la densidad del pro-átomo $\rho_i^{at}(r)$ en comparación con el método de Hirshfeld:

$$\rho_i^{at} = \sum_{i=1}^{m_A} \rho_{Ai}^0(r) \quad (2.36)$$

donde m_A es el número de funciones de Slater (número de capas del átomo) y ρ_{Ai}^0 se define como:

$$\rho_{Ai}^0 = N_{Ai} f_{Ai} = \frac{N_{Ai}}{\sigma_{Ai}^3 8\pi} \exp\left(-\frac{|r - R_A|}{\sigma_{Ai}}\right) \quad (2.37)$$

En la anterior expresión, N_{Ai} es la población, σ_{Ai} es el ancho de cada capa atómica y ambos son variables libres, que se refinan iterativamente hasta que los parámetros del pro-átomo no cambien significativamente. f_{Ai} son funciones normalizadas de tipo Slater de la densidad.

Las cargas MBIS mejoran algunas deficiencias presentes en las cargas de Hirshfeld-I. En el esquema de Hirshfeld-I, los pro-átomos son interpolaciones lineales entre densidades de iones aislados, de modo que las poblaciones AIM y las poblaciones pro-atómicas coincidan. En sistemas con átomos con carga negativa considerable ($-2 < q_A < -1$), esto conduce a pro-átomos que se interpolan entre aniones aislados y dianiones. Ningún dianión atómico aislado es estable y por tanto se impide el cálculo de su densidad electrónica. Por otra parte, el modelo MBIS, no

requiere de densidades de iones inestables como entrada y se esperaría que no sufra los mismos problemas que Hirshfeld-I.

Cargas atómicas dinámicas MBIS

En los métodos anteriores no se considera la polarización de la densidad electrónica del ligando debido a la interacción con el ambiente molecular, ya sea unido al receptor o en solución. Para considerar la polarización, E. Vöhringer [9] ha propuesto una metodología mediante la cual utiliza simulaciones de dinámica molecular y el método QM/MM donde el ligando es descrito mediante QM y el resto del sistema molecular descrito mediante MM. La polarización se describe a través de la variación de las cargas MBIS debido a las fluctuaciones en el ambiente molecular, dando como resultado las cargas atómicas dinámicas MBIS (D-MBIS). Se espera que estas cargas describan de mejor forma las interacciones electrostáticas del sistema, debido a que consideran el promedio de la polarización realizada por el ambiente molecular.

El procedimiento para calcular las cargas D-MBIS es mediante simulaciones de dinámica molecular en conjunto con la metodología QM/MM como una medida para explorar conformaciones posibles en un periodo de tiempo. Por tanto, cada ciertos pasos de la simulación se realiza un cálculo de QM/MM donde se obtiene la densidad electrónica del ligando y se particiona con la metodología MBIS para obtener las cargas MBIS.

Este cálculo se repite k veces de acuerdo con el número total de pasos de la simulación, posteriormente se toman el promedio de las cargas sobre cada átomo considerando todas las cargas obtenidas durante la simulación. En un siguiente paso, las cargas son reemplazadas en el campo de fuerza y se vuelve a iniciar el proceso hasta convergencia de las cargas. Ver Figura 2.1 donde se muestra el diagrama de flujo para el proceso de obtención de cargas D-MBIS, en los sistemas Anfitrión-Huésped en el sitio de unión (A-H) y del huésped en solución acuosa (H).



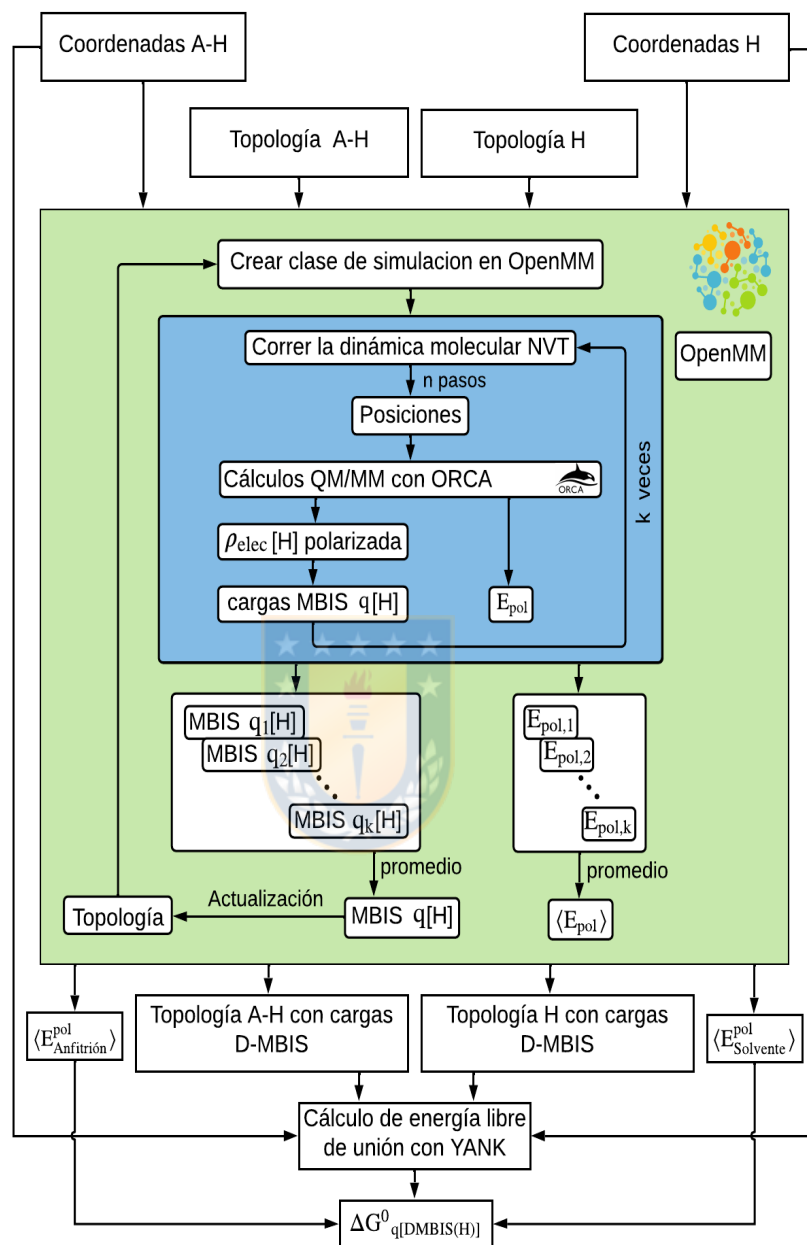


Figura 2.1: Diagrama de flujo utilizado para calcular las cargas atómicas D-MBIS para diferentes configuraciones del huésped junto con el costo de polarización energética del huésped en el estado de unión (complejo) y en solución acuosa (solvente). Los promedios de las energías de polarización corrigen la energía libre calculada por YANK para la aniquilación del huésped en los dos ambientes moleculares.

2.2.2.2. Parámetros de Lennard-Jones

El otro término importante en las interacciones no covalentes, está relacionado con las interacciones de dispersión e intercambio de repulsión descrito por los parámetros de Lennard-Jones (LJ). Mohebifar *et al.* [80] ha mostrado recientemente que los coeficientes atómicos C_6 en varios campos de fuerza están sobreestimados comparados con los derivados desde la mecánica cuántica. La discrepancia entre estos parámetros ha motivado a estudiar nuevos esquemas para derivarlos.

En algunos de los sistemas de interés predominan las interacciones dispersivas e intercambio de repulsión (ver sección de resultados) y se esperaría obtener mejores resultados en la predicción de energía libre de unión al describir adecuadamente estas interacciones. Al igual que las cargas atómicas dinámicas, se propone una metodología para derivar los parámetros de LJ usando simulaciones de dinámica molecular y cálculos QM/MM para obtener la densidad electrónica polarizada en cada configuración. La metodología propuesta, resumida en la Figura 2.2, particiona la densidad electrónica con el método MBIS y obtiene un volumen efectivo de la densidad AIM para cada átomo en la molécula $\langle r^3 \rangle_{aim} = \int r^3 \rho_{aim} d\tau$. Posteriormente se hace un promedio de $\langle r^3 \rangle_{aim}$ para cada átomo con todas las configuraciones muestreadas para cada ciclo. El proceso considera a su vez un cálculo de cargas MBIS para ca-

da configuración muestreada y su respectivo promedio que es reemplazado en cada actualización. Al finalizar las actualizaciones, se obtiene un valor promedio final para $\langle r^3 \rangle_{aim}$ usado para derivar los parámetros de LJ. Se aplicó la relación propuesta por Becke y Johnson [81], posteriormente por Tkatchenko [82], que permite derivar los coeficientes de dispersión C_6 para cada átomo en la molécula según la ecuación:

$$C_6 = \left(\frac{\langle r^3 \rangle_{aim}}{\langle r^3 \rangle_{aislado}} \right)^2 C_{6,aislado} \quad (2.38)$$

donde $C_{6,aislado}$ corresponde al coeficiente de dispersión del átomo aislado tomado desde la publicación de Chu [83] y $\langle r^3 \rangle_{aislado}$ al valor esperado para el átomo aislado. Los radios de van-der-Waals de cada átomo en la molécula se calculan con la relación recientemente introducida por Fedorov y Tkatchenko [84, 85]:

$$R_{rdw} = 2,54\alpha^{1/7} \quad (2.39)$$

donde α es la polarizabilidad estática del átomo en la molécula, obtenido escalando la polarizabilidad del átomo aislado recomendado por Chu *et al.* [83] con factor $\frac{\langle r^3 \rangle_{aim}}{\langle r^3 \rangle_{aislado}}$.

Se conserva el tipo de átomo de GAFF para los huéspedes, pero asignando nuevos parámetros de Lennard-Jones promediando los valores de

todos los átomos en el huésped con el mismo tipo de átomo según el campo de fuerza GAFF. Con el nuevo conjunto de parámetros de LJ para los tipos de átomos del huésped, las cargas atómicas D-MBIS se derivan nuevamente para el huésped en el estado de unión y solvatado junto con sus energías de polarización.



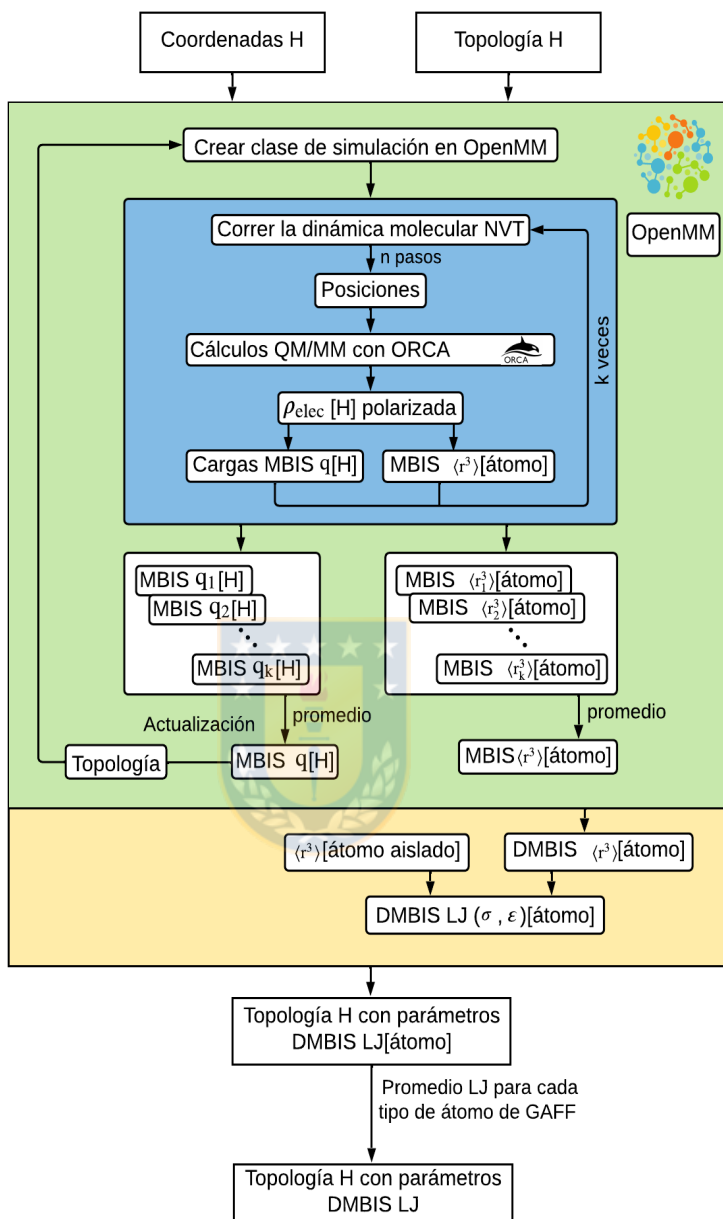


Figura 2.2: Diagrama de flujo computacional utilizado para derivar los parámetros de Lennard-Jones para los ligandos basados en la partición MBIS de la densidad electrónica y considerando un volumen efectivo de un átomo en una molécula y el volumen respectivo del átomo aislado. Se obtuvieron los parámetros del potencial de Lennard-Jones σ y ϵ para todos los átomos en cada huésped y el promedio de los coeficientes para cada tipo de átomo según el campo de fuerza GAFF.

2.3. Dinámica molecular

La dinámica molecular es una técnica de simulación computacional de tipo determinista, que describe los movimientos de todas las partículas que componen el sistema. Está basada en las leyes de la mecánica clásica, donde se integran las ecuaciones de movimiento con el fin de generar sucesivas configuraciones, permitiendo así obtener una secuencia temporal de la evolución del sistema. En simulaciones de dinámica molecular se calculan las trayectorias de los átomos, esto es, sus posiciones y momentos, generalmente en intervalos de 1 a 2 fs. El cálculo de las trayectorias en una DM se realiza a través de la integración numérica de las ecuaciones de movimiento dadas por la segunda ley de Newton:

$$\vec{F}_i(\vec{r}_i, \dots, \vec{r}_n) = m_i \ddot{\vec{r}}_i = m_i \frac{d^2 \vec{r}_i}{dt^2} \quad (2.40)$$

donde \vec{F}_i es la fuerza que actúa sobre el átomo i , m_i es la masa del átomo que se considera constante y $\ddot{\vec{r}}_i$ es la aceleración.

Dada una conformación inicial, las velocidades iniciales se asignan aleatoriamente a cada átomo, satisfaciendo una distribución de Maxwell-Boltzmann a cierto valor de temperatura. Si se conocen las posiciones y las velocidades iniciales solo resta conocer la fuerza que actúa sobre cada átomo para calcular la posición en el siguiente instante de tiempo $t_0 + dt$.

Realizando un tratamiento clásico para la energía cinética y potencial mediante el teorema de trabajo y energía, además de asumir un sistema con fuerzas conservativas (suma de energía cinética T y potencial V es constante), la fuerza sobre cada partícula se puede obtener a partir del siguiente gradiente de potencial:

$$\vec{F}_i(\vec{r}_i, \dots, \vec{r}_n) = -\nabla_i V(\vec{r}_i, \dots, \vec{r}_n) \quad (2.41)$$

La energía potencial V en un sistema molecular es requerida para hacer simulaciones de DM y se puede obtener mediante la mecánica molecular descrita anteriormente o métodos basados en la mecánica cuántica.

2.4. Método Monte Carlo vía cadenas de Markov

Las simulaciones de Monte Carlo (MC) se utilizan para modelar la probabilidad de diferentes resultados en un proceso que no se puede predecir fácilmente debido a la intervención de variables aleatorias. Los métodos Monte Carlo vía cadenas de Markov (MCMC) comprenden una clase de algoritmos para la distribución de probabilidad. Los métodos MCMC surgen de la necesidad de simular el comportamiento de variables aleatorias y de estimar parámetros de las funciones de probabilidad de las mismas. La mayoría de estas técnicas tienen un enfoque estadístico bayesiano, donde la inferencia

se realiza sobre lo que se denomina una función *a posteriori*, denominada $\pi(\theta|x)$.

Las siglas MCMC vienen marcadas por las cadenas de Markov y por la integración Monte Carlo. La inferencia bayesiana, en multitud de ocasiones necesita integrar sobre distribuciones de dimensión muy elevada (en muchas ocasiones, con cientos de parámetros). Existen métodos numéricos aproximados que producen buenas soluciones, pero que no se escalan bien con la dimensión, siendo en muchos casos, computacionalmente intratables. En términos generales, la aplicación de técnicas MCMC consta de dos pasos:

1. Generar una muestra X_1, \dots, X_n mediante una cadena de Markov cuya distribución estacionaria sea la buscada.
2. Tomar medidas muestrales (integración Monte Carlo) y realizar inferencias sobre la muestra anteriormente mencionada.

Para implementar los métodos MCMC se requiere un algoritmo para generar cadenas de Markov, cuya distribución estacionaria sea una distribución *a posteriori*. Dentro de los procedimientos más relevantes, se destacan el algoritmo de Metropolis-Hastings y el muestreo de Gibbs.

2.4.1. Movimientos de desplazamiento y rotación de Metropolis Monte Carlo

Los ligandos a menudo quedan atrapados por barreras cinéticas y choques estéricos, lo que limita su configuración espacial. El intercambio de estados termodinámicos facilita que el ligando escape de estas barreras, cambiando a un nuevo estado termodinámico. Sin embargo, el intercambio puede ocurrir con un estado en el que el ligando no está cerca del mismo volumen de espacio.

En la presente investigación se hace uso del software YANK para el cálculo de la energía libre de unión, este paquete mejora el muestreo del ligando al proponer movimientos de rotación y desplazamiento de MC sin dejar el mismo estado. Estos movimientos MC son acciones separadas y se evalúan como movimientos independientes. El movimiento de rotación MC se propone rotando todas las partículas en el ligando por un cuaternión uniforme normalizado de cuatro vectores [18]. El desplazamiento de MC propone un movimiento en que los átomos del ligando se trasladan en cada dimensión cartesiana, mediante un número aleatorio en el dominio $[0, 1]$ y luego escalado por una pequeña distancia de perturbación σ . La posición propuesta final se evalúa luego como el movimiento MC. El ligando en este caso se traduce como una unidad singular, por lo que cada partícula se traslada la misma distancia en cada dirección.

2.5. Muestreadores

Para poder estimar la energía libre de unión de los sistemas de interés, se requiere muestrear los diferentes estados termodinámicos λ considerados en el ciclo termodinámico. Existen diferentes esquemas para realizar muestreos en las simulaciones computacionales, destacándose las estrategias de: (i) Réplicas independientes (ii) Intercambio de réplicas (iii) Muestreador con mezcla auto-ajustable (iv) No equilibrio [2]. En la presente investigación, se enfocó en el muestreo con intercambio de réplicas hamiltonianas (HREX) y muestreo con mezcla auto-ajustable (SAMS) esquematizados la Figura 2.3 en la parte superior e inferior respectivamente.

En el enfoque HREX [86] se propagan múltiples estados termodinámicos simultáneamente, permitiendo el intercambio periódico entre los estados termodinámicos en dos pares de réplicas. En este enfoque, cada estado termodinámico λ puede sufrir intercambios periódicos con valores vecinos de λ entre dos réplicas. Esta forma de intercambio de réplicas se basa en ideas desarrolladas a partir de simulaciones Monte Carlo y emplean el criterio de aceptación Metropolis-Hastings para los intercambios. El conjunto generado de todas las réplicas sigue tomando muestras con una distribución de Boltzmann, por lo que este enfoque está implementado en varios paquetes de software.

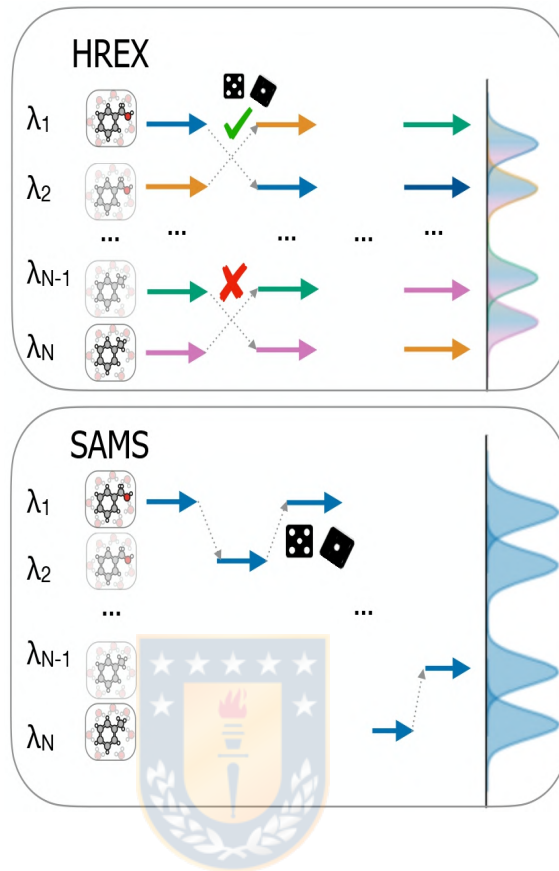


Figura 2.3: Estrategias de muestreo. Arriba: intercambio de réplicas hamiltonianas (HREX). Abajo: muestreador con mezclas auto-ajustable (SAMS). Esquema elaborado por Mey *et al.* [2]

Por otra parte (SAMS) [87], también conocido como muestreo de mezcla ajustado de manera óptima, se utilizó para muestrear toda la gama de estados termodinámicos λ , lo que permite el cálculo rápido de los cálculos de energía libre. SAMS utiliza una réplica en un conjunto ampliado [88] para muestrear múltiples estados termodinámicos, adaptados sobre la marcha [87] para lograr las probabilidades objetivo deseadas para cada estado. El esquema de actualización de estado "global-jump" permite que el muestreador visite cual-

quier estado termodinámico.

La ventaja que tiene SAMS sobre HREX es que se realizan cálculos de energía libre de unión a menor costo computacional y logra predecir energías libres de unión comparables con las demás técnicas de muestreo. HREX tiene una mayor precisión, pero a un costo computacional mucho mayor.

2.6. Ciclo termodinámico

Para realizar el cálculo de energía libre de unión absoluta, se puede calcular la diferencia de energía libre entre los estados de unión y no unión según el esquema superior de la Figura 1.1. Sin embargo esta alternativa es computacionalmente intensiva y no es posible para sistemas biomoleculares en escalas de tiempo razonables. Una alternativa es recurrir a estados *alquímicos*, en el cual sus integrales de espacio de fase deben tener suficiente solapamiento y permitan conectar los dos estados físicos de interés. En la presente investigación, se usa la metodología de doble aniquilación del huésped en el estado de unión y no unión (ver Figura 2.4). La aniquilación del huésped puede asociarse con el desacoplamiento de las interacciones electrostáticas y de van-der-Waals respecto al resto del sistema molecular.

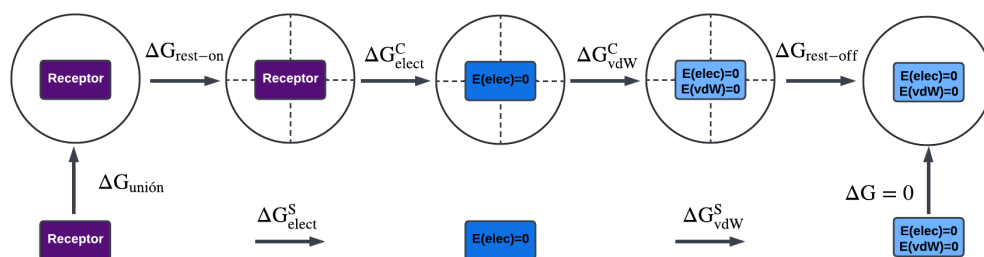


Figura 2.4: Ciclo termodinámico para el cálculo de energía libre de unión absoluta empleando la metodología de doble aniquilación para el receptor en estado de unión y no unión. Esquema adaptado de Lee *et al.* [3]. La parte superior de la Figura muestra el desacoplamiento para el huésped en estado de unión. Este paso considera una restricción para mantener al huésped en el sitio de unión mientras las interacciones electrostáticas y de van-der-Waals están desactivadas. El desacoplamiento de las interacciones electrostáticas y de van-der-Waals para el huésped en el estado no enlazado se muestra en la parte inferior de la Figura

La transformación superior de la Figura 2.4, requiere la aplicación de una restricción para que el ligando se encuentre cerca del sitio de unión durante las transformaciones *alquímicas*. Las restricciones limitan la región del espacio conformacional que debe integrarse en la mayoría de los estados *alquímicos*, acelerando la convergencia. Dentro de las restricciones más relevantes se destacan las restricciones esféricas (*harmonic, flat-bottom*) y restricciones de orientación (Boresch, RMSD). La elección del tipo de restricción depende en gran medida del sistema de estudio. En los sistemas anfitrión-huésped se aplicó una restricción tipo armónico [89] entre los centroides de dos grupos de átomos que pertenecen al anfitrión y el huésped respectivamente. La expresión de energía asociada a una restricción armónica está dada por:

$$E = \lambda_{restricción} \left(\frac{K}{2} \right) r^2 \quad (2.42)$$

donde K es una constante de fuerza, r es la distancia entre los dos grupos de centroides y $\lambda_{restricción}$ es un factor de escala que controla la fuerza de la restricción. Los grados de libertad del huésped disminuyen al aplicar este tipo de restricciones, sin embargo, la constante que restringe una distancia permite al método MCMC explorar diferentes espacios configuracionales del huésped en el sitio de unión.

En el sistema proteico, se ocupó una restricción tipo Boresch [90], que restringe en gran medida el número de orientaciones o modos de unión que el ligando debe visitar durante los estados *alquímicos*. Esto puede causar que la estimación de energía libre esté sesgada si se realiza una simulación con un modo de unión no representativo del sistema real. La aplicación de este tipo de restricción no permite que el método MCMC realice movimientos rotacionales ni traslacionales sobre el ligando. Sin embargo, en el sistema proteico seleccionado, el sitio de unión y modos de unión han sido ampliamente estudiados tanto experimental como computacionalmente, por tanto, es una opción viable ocupar este tipo de restricciones. Las restricciones tipo Boresch restringen el modo de unión del huésped al restringir una distancia, dos ángulos y tres diedros entre tres átomos del receptor y tres átomos del ligando.

Más precisamente, la expresión para ésta restricción está dado por:

$$\begin{aligned}
 E = \lambda_{restricción} [& K_r/2 (|r_3 - l_1| - r_{aA0})^2 + K_{\theta_{aA}/2} (AN(r_2, r_3, l_1) - \theta_{A0})^2 + \\
 & K_{\theta_{aB}/2} (AN(r_3, l_1, l_2) - \theta_{B0})^2 + K_{\phi_A/2} hav(DI(r_1, r_2, r_3, l_1) - \phi_{A0})^2 + \\
 & K_{\phi_B/2} hav(DI(r_2, r_3, l_1, l_2) - \phi_{B0})^2 + K_{\phi_C/2} hav(DI(r_3, l_1, l_2, l_3) - \phi_{C0})^2]
 \end{aligned}
 \tag{2.43}$$

donde *hav* es una función Haversine $\frac{1-\cos(x)}{2}$, AN y DI son abreviaciones para el ángulo y diedro formado entre los átomos seleccionados del sistema receptor-ligando. r_1, r_2, r_3 son las coordenadas de los tres átomos del receptor, l_1, l_2, l_3 son las coordenadas de los tres átomos del ligando. K_r es la constante de fuerza para la distancia restringida $|r_3 - l_1|$, r_{aA0} es la distancia de equilibrio de $|r_3 - l_1|$. $K_{\theta_{aA}}$ y $K_{\theta_{aB}}$ son las constantes de fuerza para los ángulos $AN(r_2, r_3, l_1)$ y $AN(r_3, l_1, l_2)$, θ_{A0} y θ_{B0} el ángulo de equilibrio de $AN(r_2, r_3, l_1)$ y $AN(r_3, l_1, l_2)$. K_{ϕ_A} , K_{ϕ_B} y K_{ϕ_C} son las constantes de fuerza para los diedros $DI(r_1, r_2, r_3, l_1)$, $DI(r_2, r_3, l_1, l_2)$ y $DI(r_3, l_1, r_2, l_3)$ respectivamente, ϕ_{A0} , ϕ_{B0} y ϕ_{C0} son las torsiones de equilibrio de $DI(r_1, r_2, r_3, l_1)$, $DI(r_2, r_3, l_1, l_2)$ y $DI(r_3, l_1, r_2, l_3)$ respectivamente. $\lambda_{restricción}$ es un factor de escala que puede ser usado para modular la fuerza de la restricción.

El ciclo termodinámico implementado en YANK, calcula la energía libre necesaria para liberar la restricción en un volumen específico y se considera en el cálculo de la energía libre de unión. Por tanto, la energía libre de unión

empleando las cargas atómicas D-MBIS, las energías promedio de polarización de los estados de unión y no unión y la energía asociada a la restricción aplicada, queda expresada como:

$$\Delta G_{unión} = \Delta G_{solvente} - \Delta G_{complejo} - \Delta G_{restricción} \quad (2.44)$$

Donde $\Delta G_{solvente}$ es la energía libre para desacoplar al huésped en la fase de no unión (involucra los términos ΔG_{elect}^S y ΔG_{vdW}^S de la Figura 2.4). $\Delta G_{complejo}$ es la energía libre para desacoplar al huésped en la fase de unión (involucra los términos ΔG_{elect}^C y ΔG_{vdW}^C de la Figura 2.4). $\Delta G_{restricción}$ es la energía de corrección estándar asociada con la restricción para mantener al huésped en el sitio de unión durante las transformaciones *alquímicas* (involucra los términos $\Delta G_{rest-on}$ y $\Delta G_{rest-off}$ de la Figura 2.4).

2.7. Cálculos de energía libre

Los cálculos de energía libre son relevantes debido a que permiten estudiar diferentes propiedades, como afinidades de unión de ligandos a proteínas, energías libres de hidratación, coeficientes de partición, constantes de asociación/disociación, entre otros. En esta sección, se revisan los principales métodos para calcular la energía libre de unión, basadas en simulaciones capaces de explorar el espacio de fases.

2.7.1. Método de Zwanzing

El método de Zwanzing [91], también es conocido como perturbación de energía libre, es el más conocido para calcular la diferencia de energía libre en simulaciones computacionales. La energía libre de unión empleando el método de Zwanzing, puede calcularse mediante la siguiente expresión:

$$\Delta G_{ij} = G_j - G_i = -\frac{1}{\beta} \ln \langle e^{-\beta(U_j(\vec{q}) - U_i(\vec{q}))} \rangle_i \quad (2.45)$$

El método se basa en el muestreo de las configuraciones del estado i y su subsiguiente evaluación de las configuraciones en el Hamiltoniano de j . En la ecuación 2.45 $\beta = \frac{1}{k_B T}$, k_B es la constante de Boltzmann, T es la temperatura absoluta, $U_j(\vec{q})$ y $U_i(\vec{q})$ son las energías internas a lo largo de un espacio de coordenadas \vec{q} y $\langle \rangle_i$ es el promedio sobre el colectivo del estado i . La precisión del método depende de la superposición de los espacios de fase de los estados i y j . En el caso de que la superposición sea pequeña, las configuraciones generadas en el estado i se identifican como microestados de alta energía cuando se evalúan con el Hamiltoniano del estado j y contribuyen poco a la media exponencial. Por tanto, se sabe que el enfoque converge lentamente y solo es manejable para la estimación de ΔG entre dos estados que muestran un gran superposición en el espacio de fase [92, 93].

Cuando los espacios de fase no muestran gran superposición, se realiza una transformación paulatina introduciendo un parámetro λ en el Hamilto-

niano $H_\lambda(\vec{q}) = (1 - \lambda)H_i(\vec{q}) + \lambda H_j(\vec{q})$. λ representa pequeñas alteraciones al sistema y modula la energía potencial del sistema, de modo que a medida que λ pasa de 0 a 1, la energía potencial cambia entre el estado inicial y el estado final del sistema.

2.7.2. Integración termodinámica

Otra de las metodologías más empleadas para el cálculo de energía libre de unión, que desde sus orígenes utilizó vías *alquímicas* es la integración termodinámica (TI) [94]. Los cálculos de TI estiman la diferencia de energía libre entre los dos estados A y B acoplados a través de un parámetro λ que sirve como una coordenada adicional no espacial. Bajo esta metodología, λ permite calcular la diferencia de energía libre entre los estados mediante la expresión.

$$\Delta G_{ij} = \int_0^1 \left\langle \frac{\partial H}{\partial \lambda} \right\rangle_\lambda d\lambda \quad (2.46)$$

En la Ecuación 2.46 $\lambda = 0$ en el estado i , $\lambda = 1$ en el estado j y $\langle \rangle_\lambda$ indica un promedio sobre el colectivo basado en un valor del parámetro λ . La ventaja de esta metodología es que las perturbaciones pueden ser ajustadas por el parámetro λ , llevando una mejor convergencia de los valores de energía libre.

2.7.3. Bennett Acceptance Ratio (BAR)

La diferencia de la energía potencial en ambas direcciones entre dos estados, puede ser usada para estimar la energía libre. Esto proporciona una mejora sustancial comparado con otros métodos. Esta diferencia de energía potencial para la misma configuración \vec{q} de los estados diferentes a lo largo del camino es $\Delta U_{ij}(\vec{q})$. Hay una relación directa entre las distribuciones de energía potencial $\Delta U_{ji}(\vec{q})$ con las configuraciones del estado i $\Delta U_{ij}(\vec{q})$ con las configuraciones del estado j . Por tanto, utilizando la información de ambos estados i y j , se puede aplicar un procedimiento estadísticamente óptimo para obtener una mejor estimación de la energía libre. Este procedimiento fue propuesto por el físico Charles H. Bennett y en su honor se establece el nombre de este método como Bennett Acceptance ratio (BAR) [95]. Bennett proporcionó la siguiente relación para la diferencia de energía libre:

$$\Delta G_{ij} = -\ln k_B T \frac{\langle E_i \rangle}{\langle E_j \rangle} = k_B T \ln \frac{\langle \alpha(\vec{q}) e^{[-\beta \Delta U_{ij}(\vec{q})]} \rangle_j}{\langle \alpha(\vec{q}) e^{[-\beta \Delta U_{ji}(\vec{q})]} \rangle_i} \quad (2.47)$$

Esta expresión es válida para cualquier función $\alpha > 0$ para todas las configuraciones \vec{q} . Bennett usó el cálculo variacional para encontrar el cambio de $\alpha(\vec{q})$ que minimiza la variación de la energía libre, obteniendo una función implícita de ΔG que puede resolverse numéricamente:

$$\sum_{i=1}^{n_i} \frac{1}{1 + e^{(\ln(n_i/n_j) + \beta \Delta U_{ij} - \beta \Delta G)}} - \sum_{j=1}^{n_j} \frac{1}{1 + e^{(\ln(n_j/n_i) + \beta \Delta U_{ji} - \beta \Delta G)}} = 0 \quad (2.48)$$

donde n_i y n_j son los números de configuraciones muestreadas para cada estado. Se ha demostrado que el método BAR es más eficiente que el método Zwanzig [92, 96] y de integración termodinámica, debido a la inclusión de información para la transformación en sentido $i \rightarrow j$ y $j \rightarrow i$.

2.7.4. Multistate Bennett acceptance ratio (MBAR)

El método de MBAR [97] es una extensión del método BAR, ya que permite evaluar los datos de todos los estados (no sólo de dos estados como en el caso de BAR) y así predecir de mejor manera la energía libre. MBAR requiere la energía de cada configuración muestreada, evaluada en cada estado en el que se desea calcular la energía libre y luego el número de muestras extraídas de cada estado. A partir de ahí, se estima la energía libre de cada estado a través de la siguiente ecuación autoconsistente:

$$\hat{f}_i = -\ln \sum_{j=1}^K \sum_{n=1}^{N_j} \frac{e^{-u_i(x_{j,n})}}{\sum_{k=1}^K N_k e^{[\hat{f}_k - u_k(x_{j,n})]}} \quad (2.49)$$

donde f es la estimación de energía libre del estado i , K es el número de estados muestreados, U es la energía potencial, N es el número de muestras de un estado dado y x es la configuración muestreada. Debido a que esto requiere

autoconsistencia, es un problema sub-especificado tal como se presenta. Sin embargo, en realidad sólo importan las diferencias en las energías libres entre estados, por lo que se elige un estado como referencia, como $f_0 = 0$. Esta ecuación tampoco está condicionada a que el estado i sea parte de los estados muestreados K , lo que permite la estimación de la energía libre en cualquier estado termodinámico.



2.8. Corrección de polarización a la energía libre de unión

Para la descripción adecuada de las interacciones del huésped con el entorno molecular, se tomó en cuenta la densidad de electrones del huésped polarizado al calcular las cargas atómicas D-MBIS ver Figura 2.1. La polarización del huésped en el ambiente de unión y en ambiente acuoso está asociado a un costo de energía potencial que siempre es positivo y contrarresta a la energía libre obtenida en cada una de los ambientes moleculares. Se calculó la energía libre de unión considerando un esquema de doble desacoplamiento y empleando el ciclo termodinámico mostrado en la parte superior de la Figura 2.4. En principio, si la energía de polarización para el huésped es la misma en el estado de unión y acuoso, esta contribución se cancelaría. Pero, como se muestra en la sección de resultados, el costo de la energía de polarización difiere entre los dos entornos moleculares. Esta diferencia contribuye a la energía libre de enlace y generalmente no se considera en los cálculos de energía libre de unión empleando campos de fuerza no polarizables.

El costo energético asociado con la polarización del huésped en estado de unión $E_{complejo}^{pol}$ y acuoso $E_{solvente}^{pol}$ se determinó con un cálculo de estructura electrónica adicional según el protocolo descrito anteriormente, ver Figura 2.1. En este cálculo, se usó el Hamiltoniano de estructura electrónica estándar

dar en vacío \hat{H}_{vac} y se calculó su valor esperado usando la función de onda polarizada del huésped obtenida con los cálculos QM/MM en el ambiente de unión $\Psi_{complejo}$ o acuoso $\Psi_{solvente}$. El costo de polarización energética es la diferencia entre este valor esperado y la energía del campo autoconsistente obtenida en el cálculo de un solo punto empleando el mismo Hamiltoniano:

$$E_{complejo}^{pol} = \langle \Psi_{complejo} | \hat{H}_{vac} | \Psi_{complejo} \rangle - \langle \Psi_{vac} | \hat{H}_{vac} | \Psi_{vac} \rangle \quad (2.50)$$



$$E_{solvente}^{pol} = \langle \Psi_{solvente} | \hat{H}_{vac} | \Psi_{solvente} \rangle - \langle \Psi_{vac} | \hat{H}_{vac} | \Psi_{vac} \rangle \quad (2.51)$$

La energía de polarización promedio $\langle E_{complejo}^{pol} \rangle$ y $\langle E_{solvente}^{pol} \rangle$ de cada huésped en los dos ambientes moleculares se obtiene promediando 25 configuraciones del protocolo de simulación (ver Figura 2.1) y representa el valor promedio del colectivo. Como consecuencia, para el cálculo de la energía libre de enlace $\Delta G_{unión}$ se consideran dos contribuciones relacionadas con el costo de la energía de polarización. Por tanto, el cálculo de la energía libre de unión empleando la ecuación 1.1 y considerando el valor promedio de la energía de polarización para los dos ambientes moleculares, puede reescribirse como:

$$\Delta G_{unión} = \Delta G_{solvente} - \langle E_{solvente}^{pol} \rangle - \Delta G_{complejo} + \langle E_{complejo}^{pol} \rangle - \Delta G_{restricción} \quad (2.52)$$

La ecuación 2.52 se emplea para reportar las energías libres de unión de los sistemas de interés, corregidas por el promedio de la energía de polarización derivada en los ambientes moleculares de unión y no unión. Las correcciones por polarización se hacen una vez YANK entregue las respectivas energías libres. Además, es importante resaltar que en aquellos sistemas donde sólo se consideró los parámetros de no enlace con la derivación D-MBIS del huésped en estado acuosa y su respectivo reemplazo en el estado de unión, los términos $\langle E_{solvente}^{pol} \rangle$ y $\langle E_{complejo}^{pol} \rangle$ se anularían.

3. Resultados y Discusión

A continuación, se muestran los resultados más relevantes del proyecto de investigación. En primera instancia se discuten los protocolos de simulación empleados para calcular la energía libre de unión en sistemas anfitrión-huésped y proteína ligando. Posteriormente y considerando los protocolos validados previamente se reportan y analizan los resultados obtenidos para las energías libres de unión en los sistemas: (i) anfitrión-huésped -octa ácido (OA) y tetra-endometil octa ácido(TEMOA) con los huéspedes del SAMPL5, -cucurbit[7]urilo (CB7) con huéspedes de núcleo aromático y de adamantano. (ii) proteína-ligando -lizosima T4 L99A/M102Q y ligandos con núcleo aromático. Los resultados de energía libre de unión consideran diferentes conjuntos de parámetros no enlazantes y se comparan ocupando los descriptores estadísticos: error cuadrático medio (RMSE), error absoluto medio (MAE), coeficiente de correlación de Pearson (R) y coeficiente kendall (τ).

3.1. Validación de parámetros de simulación utilizados para el cálculo de energía libre de unión

3.1.1. Sistema anfitrión-huésped

El método computacional empleado para el cálculo de energía libre de unión deber ser un método razonable en cuanto al tiempo de simulación y nivel de precisión, para garantizar que los resultados no estén sesgados. Las estrategias computacionales empleadas para el cálculo de afinidad dependen en gran medida de las condiciones y parámetros seleccionados para la simulación. Para validar la metodología en el cálculo de energía libre de unión se estudió el sistema TEMOA-G1 y se analizaron los siguientes aspectos: conformación inicial del huésped, número de iteraciones, número de estados termodinámicos, modelo de agua, efecto de la fuerza iónica y dependencia del muestreador ocupado (HREX, SAMS).

Todas las simulaciones fueron realizadas inicialmente con el software YANK versión 0.23.7 usando 1000 pasos de minimización, masa de hidrógeno tres una y tiempo de paso de cuatro fs, y una restricción tipo armónica para mantener el huésped en la cavidad de unión del anfitrión durante las transformaciones *alquímicas* y una equilibración automática antes del

cálculo de energía libre de unión. Las configuraciones iniciales y los parámetros del campo de fuerza GAFF fueron obtenidos del repositorio GitHub MobleyLab/benchmarksets [4]. Estos parámetros consideran las cargas atómicas RESP para el anfitrión y cargas AM1-BCC para el huésped además del modelo de agua TIP4-Ew.

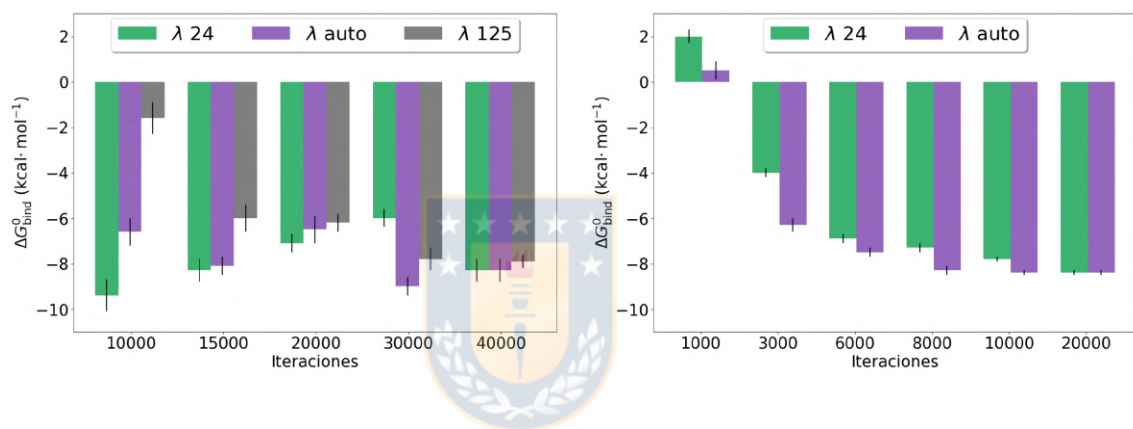


Figura 3.1: Afinidades de unión para el sistema TEMOA-G1 obtenidas con YANK a diferente número de iteraciones utilizando los muestreadores SAMS (izquierda) y HREX (derecha). Se compara un número diferente de estados termodinámicos intermedios (λ) con los determinados automáticamente por el algoritmo "thermodynamic trailblazing" implementado en YANK (60 y 56 estados intermedios para la fase compleja y acuosa, respectivamente).

Primero se evaluó el método de muestreo, combinado con el número de estados termodinámicos óptimos que proporciona una convergencia en la energía libre de unión en el sistema TEMOA-G1 en el menor tiempo posible y con recursos computacionales limitados. Para evaluar el costo computacional se consideró el tiempo que lleva obtener los resultados en una GPU

(RTX2080). Se varió el número de iteraciones con tres conjuntos de estados termodinámicos λ 24, λ 125 (Estos números de estados son propuestos en los ejemplos de ejecución de YANK para sistemas anfitrión-huésped) y λ *auto* corresponde a estados termodinámicos derivado para garantizar una superposición óptima del espacio de fases (*thermodynamic trailblazing*). Para el método de muestreo SAMS, 40000 iteraciones presentaron resultados convergentes. En el método de muestreo HREX, las energías libres de unión no varían significativamente después de 8000 iteraciones ver Figura 3.1. En el método SAMS la energía libre de unión empleando los estados termodinámicos λ 25 presenta grandes fluctuaciones al cambiar el número de iteraciones mientras que con λ *auto* y λ 125 las variaciones son menores. Sin embargo, con λ *auto* ocupa menos tiempo computacional. En contra parte, en el método HREX, el valor de afinidad de unión converge al mismo valor, pero a mayor número de iteraciones. Aunque HREX es más preciso, su tiempo computacional es aproximadamente diez veces mayor que el método SAMS: considerando 8000 y 40000 iteraciones junto con λ *auto* los tiempos de simulación son 41 horas y 520 horas para HREX y SAMS respectivamente.

Para continuar con el estudio de la validación se seleccionó el método SAMS que entrega energías libres de unión en el intervalo de 1-2 días en una GPU y precisa de un menor costo computacional que el método HREX. De los protocolos estudiados en SAMS la validación se hará con λ 125 para garantizar la convergencia en la energía libre unión. Una vez fijado el protocolo

de simulación para predecir las energías libres de unión de los sistemas de interés se puede usar la herramienta *thermodynamic trailblazing* implementada en YANK para generar los estados termodinámicos y reducir aún más el costo de simulación.

Durante la validación de los parámetros de simulación, se observó que en algunas simulaciones las energías libres de unión difieren significativamente dependiendo de la selección del contraión Na^+ que es transformado alquímicamente en paralelo con el huésped 5-hexinoato (G1). En general, cuando se ocupa la metodología de aniquilación de las interacciones y si el huésped está cargado positiva o negativamente, deben aniquilarse en paralelo las interacciones del contraión respectivo para mantener la neutralidad del sistema. Específicamente, la distancia del contraión al huésped G1 cargado negativamente condujo a cambios significativos en los valores calculados, incluso cuando las posiciones iniciales del huésped y del solvente se mantuvieron constantes (ver tabla 3.1). Se intercambiaron las coordenadas del ión Na^+ transformado alquímicamente con otro ión con carga positiva en la caja de simulación a diferentes distancias del átomo de carbono del grupo carboxílico con carga negativa del huésped, manteniendo el resto de las posiciones atómicas iniciales. Las energías libres de unión evidencian una dependencia pronunciada de la distancia del contraión transformado alquímicamente al huésped (ver tabla 3.1), que alcanza el valor de referencia de -8.0 ± 0.3 kcal/mol a distancias superiores a 25 Å. Además, también se consideró un

sistema donde se generaron nuevas posiciones iniciales para el solvente y los iones y se seleccionó las coordenadas del ión transformado alquímicamente como la más lejana. La coincidencia de la energía libre de unión obtenida con el valor de referencia confirma que los resultados son independientes de la configuración inicial de los iones y las moléculas de agua en el sistema.

Tabla 3.1: Afinidades de unión calculadas para diferentes distancias del contraión Na^+ transformado alquímicamente al huésped y diferentes poses del huésped. (Distancias medidas entre el contraión Na^+ y el átomo de carbono del grupo carboxílico cargado negativamente del huésped en Å). Valores de energía libre en kcal/mol

Simulación	$d(\text{Na}^+ - \text{H})$	$\Delta G_{\text{unión}}^0$
Misma configuración diferente contraión	11.3	-3.5 ± 0.5
	22.7	-5.3 ± 0.5
	27.9	-8.0 ± 0.4
Diferentes posiciones iniciales de los iones	34.4	-7.9 ± 0.4
Diferente pose del huésped y posición de los iones	34.0	-8.2 ± 0.5

La versión YANK 0.24.1 introdujo un mejor tratamiento de las interacciones de largo alcance, lo que nos motivó a recalcular la afinidad de unión con la selección de una distancia superior a 25 Å del contraión que se transforma alquímicamente respecto al huésped y los mismos parámetros de simulación. El valor obtenido de -8.3 ± 0.3 kcal/mol se asemeja al resultado anterior dentro de los márgenes de error, lo que sugiere que el efecto de las interacciones

electrostáticas de largo alcance es menor. Interesantemente, las simulaciones con la nueva versión de YANK, con una distancia del contraión Na^+ al huésped de 11.3 Å, dan como resultado una energía libre de unión de -6.4 ± 0.3 kcal/mol más cercana al valor de referencia. Por lo tanto, mejorar la interacción electrostática de largo alcance en la versión 0.24.1 reduce el efecto de la distancia del ion transformado alquímicamente al huésped en la afinidad de unión. Sin embargo, todavía se desvía significativamente del valor de -8.3 kcal/mol. Para las posteriores simulaciones se consideró la nueva versión de YANK 0.24.1 para una mayor validación y producción debido a un mejor tratamiento de las interacciones electrostáticas de largo alcance, además de la distancia superior a 25 Å del contraión que se transforma alquímicamente respecto al huésped.

Las energías libres de unión también deberían ser independientes de la posición inicial del huésped en el anfitrión. Se abordó el efecto de la posición inicial del huésped en el anfitrión generando una conformación alternativa a partir de su representación de cadena SMILES a través de la biblioteca OMEGA en OpenEye Toolkit (versión 2018.10.1) (ver Figura 3.2). Las diferencias en la afinidad de unión fueron solo de 0.05 kcal/mol, lo que confirma que las energías libres de unión obtenidas con YANK para este sistema no depende de la conformación inicial del huésped. La razón principal podría estar relacionada con el algoritmo de YANK, que promueve la rotación y la traslación del huésped en el anfitrión en cada iteración. Los movimientos de rotación

y traslación de Monte Carlo permiten explorar y muestrear más orientaciones relativas en el sitio de unión que las simulaciones de dinámica molecular estándar y reducen la dependencia de la configuración inicial.

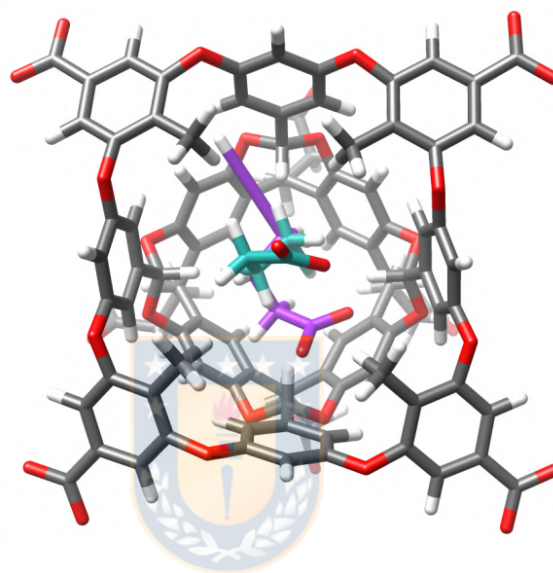


Figura 3.2: Estructura tridimensional de TEMOA y G1 en diferentes conformaciones iniciales para el huésped. La representación en átomos de carbono púrpura y cian corresponden a las conformaciones obtenidas del repositorio de Github (Mobley-Lab/benchmarksets [4]) y con la herramienta OMEGA respectivamente. Las energías libres obtenidas con estas dos conformaciones iniciales se reportan en la Tabla 3.1

También se evaluó cambiar la masa atómica de los átomos de hidrógeno a una una con un paso de tiempo de dos fs manteniendo todos los demás parámetros. La afinidad de unión fue -7.9 ± 0.7 kcal/mol sin diferencia estadísticamente significativa a la reportada previamente. Sin embargo, el tiempo computacional se duplica, por lo que se mantuvo el tiempo de paso más

grande junto a la masa de los átomos de hidrógeno en tres una para las simulaciones posteriores.

Con base en el análisis discutido anteriormente, se eligió la metodología que emplea el muestreador SAMS con 40000 iteraciones y detección automática de estados λ intermedios y la característica de seleccionar el contraión *alquímico* más lejano del huésped junto con un paso de tiempo de cuatro fs y la masa atómica de hidrógeno de tres una para simulaciones de producción. Esta combinación de parámetros y métodos proporciona una energía libre de enlace anfitrión-huésped aproximadamente en 41 horas en una GPU RTX2080.

Finalmente, también se estudió el efecto de incluir iones en el sistema para imitar la fuerza iónica experimental y el cambio del modelo de agua (TIP3P y TIP4P-Ew) en las energías libres de unión anfitrión-huésped calculadas. Se agregaron tres y dos iones de sodio y cloruro adicionales en las simulaciones con el anfitrión y la solución acuosa, respectivamente, para imitar la concentración experimental de 10 mM de la solución reguladora de pH. Se debe notar que debido a la carga neta del anfitrión TEMOA, la cantidad de iones Na^+ en el sistema de simulación es ocho para mantener un sistema neutral. Las energías libres de unión calculadas fueron -3.7 ± 0.4 y -3.5 ± 0.3 kcal/mol utilizando el modelo de agua TIP3P con y sin solución amortiguadora de pH, respectivamente. Curiosamente, los resultados con el modelo de agua TIP3P muestran solo un ligero efecto en solución reguladora de pH.

Para el modelo de agua TIP4P-Ew, el impacto de la solución reguladora de pH es más significativo que el modelo de agua TIP3P, lo que resulta en -4.0 ± 0.4 y -8.3 ± 0.3 kcal/mol con y sin solución amortiguadora de pH, respectivamente. Una posible explicación de la diferencia entre los dos modelos de agua es la constante dieléctrica, que podría ser determinante en esta alta concentración de iones [98].

Estudios previos de TEMOA-G1 en el desafío SAMPL5 usando métodos similares y el modelo TIP3P reportan valores idénticos con y sin solución amortiguadora de pH (-4.2 ± 0.4 y -4.2 ± 0.0 kcal/mol) [99]. Los mismos autores informan que teniendo en cuenta la fuerza iónica experimental en la simulación con métodos similares y el mismo campo de fuerza para todos los huéspedes en TEMOA, se redujo el RMSE de 3.5 a 1.9 kcal/mol. Pal *et al.* informa un cambio de energía libre vinculante de -7.5 ± 0.1 a -4.8 ± 0.1 kcal/mol en TEMOA-G1 después de tener en cuenta la fuerza iónica usando el campo de fuerza OPLSA2005 para el anfitrión y el huésped combinado con el modelo de agua AGBNP2 y una reducción de RMSE de 6.7 a 2.1 kcal/mol [100].

Con base en los resultados obtenidos y la observación previa de todos los huéspedes en los anfitriones OA y TEMOA, se decidió usar el modelo de agua TIP3P debido a su compatibilidad con el campo de fuerza GAFF y tener en cuenta la fuerza iónica que imita a la condición experimental para calcular la energía libre de unión en el resto de los sistemas anfitrión-huésped

de interés.

3.1.2. Sistema proteína-ligando

Los sistemas proteína-ligando exhiben mayores desafíos computacionales que los sistemas anfitrión-huésped. Por tanto, se decidió validar una metodología para el cálculo de energía libre de unión en sistemas proteicos. Se seleccionó como sistema modelo la lisozima T4 L99A unido al tolueno (L1) debido a que este es un sistema ampliamente estudiado y se ha reportado en la literatura energías libres de unión empleando el software YANK [101]. Si bien la lisozima T4 L99A/M102Q es el sistema de interés, se optó por seleccionar la lisozima T4 L99A porque presenta menor flexibilidad y se pueden evaluar con mayor objetividad la influencia de la variación de parámetros en la energía libre de unión. Se espera que el protocolo validado para Lisozima T4 L99A-tolueno sea transferible a los sistemas T4 L99A/M102Q-ligandos aromáticos (ver Figura 1.4) ya que la diferencia de los sistemas proteicos se centran solo en la mutación M102Q. Los archivos iniciales se extrajeron de la información soporte publicada por Gill *et al.* [102], se usaron algunos de los parámetros estudiados en el sistema anfitrión-huésped y se varió el número de iteraciones (ver Tabla 3.2). Se ocupó la masa de hidrógeno tres una, tiempo de paso cuatro fs, muestreador SAMS, los campos de fuerza ff99sb-ildn [103] y GAFF para la proteína y el ligando respectivamente, modelo de cargas AM1-BCC para el tolueno, modelo de agua TIP3P, detección auto-

mática de estados termodinámicos (27 y 29 para la fase compleja y acuosa respectivamente) y restricción tipo Borech [104] entre 3 átomos de la proteína (Carbonos alfa de los residuos LEU118, VAL111, GLN102) y tres átomos del tolueno (Carbono del grupo metilo y Carbonos en la posición orto y para).

Tabla 3.2: Afinidades de unión para el sistema lisozima T4 L99A unido al tolueno obtenido con YANK a diferente número de iteraciones utilizando el muestreador SAMS. Los estados termodinámicos intermedios corresponden a 27 y 29 para la fase compleja y acuosa, respectivamente. Valores de energía libre en kcal/mol, el tiempo estimado de la simulación se muestra en horas

Número de iteraciones	$\Delta G_{\text{unión}}$	tiempo
10000	-6.0 ± 0.2	10.35
20000	-5.9 ± 0.1	20.71
30000	-6.0 ± 0.1	32.31
40000	-6.0 ± 0.2	43.08
50000	-6.4 ± 0.2	54.05
80000	-6.3 ± 0.1	88.14

Las energías libres de unión obtenidas (ver Tabla 3.2) reflejan una rápida convergencia incluso después de las 10000 iteraciones. Esta tendencia puede explicarse por el el tipo de restricción aplicada. Las restricciones de orientación tipo Borech restringen una distancia, dos ángulos y tres torsiones en un intento de mantener el ligando en una pose de unión específica [104]. En los sistemas proteicos suelen emplearse este tipo de restricciones donde los

pasos MCMC de traslaciones y rotaciones no se realizan. El valor convergido, sin embargo, se aleja del valor reportado por Gill *et al* [102] de -4.2 ± 0.04 kcal/mol. Para conocer la divergencia de estos resultados, se realizó una simulación con todos los parámetros reportados en la publicación el cual usa la masa del hidrógeno en una uma, un tiempo de paso de dos fs, HREX como muestreador (26 y 21 estados para la fase compleja y acuosa respectivamente) y 5000 iteraciones. Se obtuvo una energía libre de unión de -3.8 ± 0.1 kcal/mol similar al valor de referencia. Como se demostró anteriormente en el caso de TEMOA-G1 el cambio de la masa de hidrógeno y el tiempo de paso no influye en el cálculo de energía libre de unión. Se asoció por tanto la divergencia al muestreador empleado y la cantidad de estados termodinámicos. Para validar esta hipótesis se realizó la misma simulación reportada por Gill *et al* [102] pero incluyendo los estados termodinámicos generados de forma automática con YANK (27 y 29 para la fase compleja y acuosa respectivamente) y se obtuvo una afinidad de unión de -6.2 ± 0.1 kcal/mol con un tiempo de simulación de 31.53 horas. Este resultado corrobora que el resultado de la energía es independiente del muestreador empleado, sin embargo, el costo computacional empleando el muestreador HREX es mayor.

De los resultados de esta validación se concluye que es posible obtener energías libres de unión reproducibles para el sistema proteico de interés independientes del tipo de muestreador empleado y dependiente de la cantidad y espaciado de los estados termodinámicos considerados en la transformación

alquímica. Además, empleando el muestreador SAMS se logra converger a una energía libre de unión absoluta en un menor tiempo computacional.

Para las simulaciones del sistema lisozima T4 L99A/M102Q y huéspedes de núcleo aromático se considerarán los archivos extraídos del repositorio de Github MobleyLab benchmarksets [4]. Estos sistemas fueron preparados considerando la estructura APO (estructura inactiva, sin unirse a un ligando) de la proteína a partir del pdb 1LGU (T4 lisozima mutante L99A/M102Q) y las posiciones iniciales de todos los ligandos obtenidas directamente de las estructuras cristalinas con T4 lisozima L99A/M102Q. Los parámetros para la proteína se obtienen del campo de fuerza AMBERff14SB y de los ligandos del campo de fuerza GAFF v1.7 empleando el modelo de cargas AM1-BCC. Los sistemas fueron solvatados empleando el modelo de agua TIP3P con iones K^+ y Cl^- para emular la fuerza iónica, con parámetros iónicos específicos de TIP3P de Joung y Cheatham [105]. Los cálculos de energía libre se realizaron empleando el muestreador SAMS, con 20000 iteraciones, masa de hidrógeno de tres una, tiempo de paso de cuatro fs, detección automática de los estados termodinámicos, pasos de equilibración fijados por el software YANK y restricción tipo Boresch [104].

3.2. Validación de parámetros de simulación utilizados para el cálculo de las cargas D-MBIS y energía de polarización promedio

En la transferencia de una molécula del vacío a las fases condensadas, su densidad electrónica se distorsiona, lo que resulta en su polarización. La energía requerida para polarizar la densidad de electrones para una posición fija de los núcleos es la energía de polarización y siempre es positiva. Esta energía generalmente no se considera en el cálculo de energías libres de unión o energías libres de hidratación con campos de fuerza no polarizables, como ya lo señaló Swope *et al.* [106, 107]. En investigaciones previas del grupo de investigación se demostró que tener en cuenta esta energía de polarización en los cálculos de energía libre de hidratación de la base de datos FreeSolv da como resultado una buena concordancia con los valores experimentales cuando se combina con cargas atómicas MBIS [108]. Sin embargo, las simulaciones de energía libre de unión absoluta a través de rutas *alquímicas* han descuidado en su mayoría una posible contribución de la polarización a la afinidad de unión. Estas simulaciones computacionales calculan el costo de energía libre de desacoplar la interacción del huésped de su entorno en dos sistemas moleculares diferentes: el estado unido y el no unido (rodeado de moléculas solventes). Al despreciar la contribución de la polarización se asume que

ambos estados comparten el mismo valor resultando en su cancelación en el ciclo termodinámico de energía libre. Esta aproximación puede no ser siempre precisa, especialmente cuando el huésped posee una gran polarizabilidad y el sitio de unión es hidrofóbico. Aquí, se evaluó si esta aproximación es válida para los sistemas anfitrión-huésped.

Después de que el huésped se polarice por el solvente o el anfitrión, las interacciones electrostáticas con su entorno molecular diferirán de su estado no polarizado, por ejemplo, en vacío. Para mantener una descripción compatible con un campo de fuerza no polarizable, se asignaron cargas atómicas para el huésped, que describe mejor la nueva distribución de la densidad de electrones polarizados. Se eligió el método de partición MBIS de la densidad de electrones para calcular las cargas atómicas debido a sus características (ver referencia [109] para una discusión detallada) y los resultados anteriores del grupo de investigación para las energías libres de hidratación Free-Solv [108]. Debido a la distribución heterogénea de átomos en los sistemas anfitrión-huésped, se utilizó la metodología QM/MM para obtener cargas atómicas MBIS a partir de la densidad electrónica polarizada del huésped. Estos cálculos de estructura electrónica incluyen cargas atómicas del entorno molecular circundante representado por el campo de fuerza en el Hamiltoniano electrónico, lo que da como resultado una densidad electrónica polarizada.

La energía de polarización y las cargas atómicas dependerán de la posición instantánea de los átomos y presentarán fluctuaciones debido a su diná-

mica en un colectivo canónico. Estamos interesados en las cargas atómicas MBIS, que describen mejor las interacciones promedio y la energía de polarización promedio del conjunto del huésped. Por lo tanto, se creó un nuevo protocolo de simulación que genera nuevas configuraciones del sistema con simulaciones de dinámica molecular clásica a nivel de campo de fuerza usando el software OpenMM [110] y realiza un cálculo QM/MM con ORCA 4.2.1 [111] usando las nuevas posiciones del sistema. Luego, las cargas atómicas MBIS se obtienen a partir de la densidad de electrones polarizados del huésped y la energía de polarización se calcula como la diferencia del valor esperado de energía del Hamiltoniano de vacío aplicado a la función de onda polarizada y la función de onda en el vacío (consulte la sección de Metodología). Después de un extenso muestreo del espacio de configuración, se hace un promedio de las energías de polarización obtenidas y las diversas cargas atómicas MBIS. Estas últimas reemplazan las antiguas cargas del huésped en el campo de fuerza, iniciando un nuevo ciclo de simulaciones de dinámica molecular y cálculos QM/MM. Finalmente, se repitió la actualización de carga en el campo de fuerza hasta que se alcance la convergencia (ver Figura 2.1 esquema del flujo de trabajo de simulación).

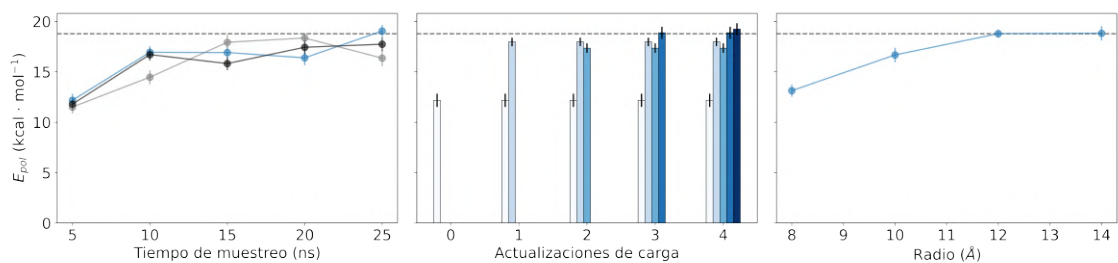


Figura 3.3: Validación de los parámetros de simulación para obtener convergencia en las cargas atómicas MBIS y en la energía de polarización promedio del esquema de trabajo mostrado en el Figura 2.1. La energía de polarización promedio del huésped G1 en el anfitrión TEMOA se muestra para tres trayectorias independientes como función del tiempo total de simulación usando cargas atómicas AM1-BCC (sin actualización de topología) (izquierda), para un número variable de actualizaciones de carga de la topología del huésped con cinco ns de muestreo para cada actualización (centro), y usando cinco actualizaciones y 25 ns de muestreo y variando el radio de corte que define las cargas atómicas para los cálculos de QM/MM que rodean al huésped (derecha). Para los tres parámetros, la convergencia de la energía de polarización se logra en los valores utilizados en el flujo de trabajo de simulación.

Para obtener energías de polarización promediadas por conjuntos representativos y cargas atómicas D-MBIS, se validaron los parámetros del flujo de trabajo de simulación. La energía de polarización resultó ser más sensible en la validación de parámetros del flujo de trabajo. Como sistema de prueba, se seleccionó el huésped G1 en el anfitrión TEMOA y se abordó primero el tiempo de muestreo y luego el número total de actualizaciones de carga. Aumentando el tiempo de muestreo de 5 a 25 ns con las cargas atómicas AM1-BCC iniciales, se observaron fluctuaciones en la energía de polarización entre diferentes trayectorias para tiempos de muestreo más largos (ver Figura 3.3). Para reducir estas variaciones y mantener un costo computacional reducido,

se decidió establecer el tiempo de muestreo en cinco ns y analizar el efecto de actualizar las cargas atómicas del huésped en la topología. Después de la primera actualización de carga del huésped en la topología, la energía de polarización promedio sufre un cambio considerable (consulte la Figura 3.3). En esta primera actualización, las cargas del huésped también varían significativamente, alterando la configuración del entorno molecular y la energía de polarización. Sin embargo, después de cinco actualizaciones, la energía de polarización promedio converge a un valor constante. Esto va acompañado de cambios marginales en cada actualización de topología que dan como resultado la energía de polarización promediada del conjunto convergente y las cargas atómicas D-MBIS. Como señaló Pan *etal*[112], la energía de polarización también podría depender del radio utilizado para seleccionar las cargas atómicas del entorno molecular para los cálculos QM/MM. La selección usa la distancia al huésped como criterio y la combina con una función de cambio que escala los cargas a cero comenzando 2\AA antes del radio de corte[112]. Usando cinco actualizaciones y 25 ns de muestreo por actualización, se estudió su dependencia para varios radios de selección de 8 a 14\AA . Un radio de 12\AA conduce a una energía de polarización promedio convergente para el huésped G1 en el anfitrión TEMOA (consulte la Figura 3.3).

No se llevó a cabo un análisis comparativo extenso del método de estructura electrónica utilizado en los cálculos QM/MM. El conjunto de base triple zeta (def2-TZVP) combinado con el funcional B3LYP proporciona un enfo-

que sólido y computacionalmente eficiente para obtener la densidad de electrones moleculares a partir de los cálculos QM/MM. Aunque los funcionales locales (GGA) son computacionalmente menos costosos, sufren el error de autointeracción que conduce a momentos dipolares pequeños y una descripción inexacta de la distribución de electrones en la molécula [113, 114, 115]. Los funcionales híbridos describen mejor la distribución, y un conjunto base triple zeta representa un buen balance computacional entre la precisión y el costo computacional.

Según el análisis anterior, el cálculo de las cargas atómicas D-MBIS se realizan con cinco actualizaciones de carga en la topología utilizando 25 configuraciones para los cálculos QM/MM tomados de trayectorias de cinco ns (corte de distancia de 12 Å). En paralelo se derivara la energía libre de polarización promedio en ambos ambientes moleculares y empleando el esquema de trabajo validado (ver Figura 2.1).

Estos parámetros no covalentes del campo de fuerza derivados a partir de la densidad electrónica particionada con el método MBIS, junto con otros parámetros del campo de fuerza GAFF y estructuras iniciales, sirven como entrada para el cálculo de la energía libre de unión en el estado de unión y no unión en YANK para los diferentes sistemas de interés. Las correcciones de polarización a la energía libre de unión se hacen una vez YANK finalice la simulación (ver ecuación 2.52) y corresponden a las energías reportadas en las siguientes secciones.

3.3. Energías libres de unión entre los anfitriones OA/TEMOA y huéspedes del SAMPL5

Con el protocolo de simulación validado para las afinidades anfitrión-huésped descritas en la sección anterior, se estudió la variación de la energía libre de unión de Gibbs cuando se reemplazan las cargas atómicas AM1-BCC en el campo de fuerza GAFF por el nuevo conjunto de cargas atómicas D-MBIS. Después de incluir la corrección de polarización, se compararon los resultados con los valores de referencia experimentales. Primero, se mantuvo las cargas RESP en el anfitrión reemplazando solo las cargas AM1-BCC del huésped por el nuevo conjunto de cargas atómicas D-MBIS ($\Delta G_{q[\text{DMBIS}(\text{H})]}^{\circ}$), y luego se calcularon las cargas atómicas D-MBIS para el anfitrión seguido de una nueva derivación de cargas para el huésped ($\Delta G_{q[\text{DMBIS}]}^{\circ}$).

La Tabla 3.3 resume los valores experimentales de la energía libre de unión de 6 huéspedes (ver Figura 3.4 unidos a OA/TEMOA y los resultados calculados usando dos conjuntos de cargas atómicas para el huésped, AM1-BCC o D-MBIS con su respectiva corrección de polarización. Las últimas dos columnas muestran el promedio de la energía de polarización asociadas con el costo energético de polarizar la densidad electrónica del huésped en estado de unión o en solución acuosa. Todas las energías libres de unión calculadas con las cargas atómicas D-MBIS dan cuenta de la diferencia en la energía de po-

larización del huésped en el anfitrión y en solución acuosa. Estas diferencias van desde 0.1-5.0 y 0.5-6.8 kcal/mol en los sistemas OA/TEMOA respectivamente. En sistemas donde la diferencia es mayor, el aporte a la energía libre de unión absoluta es significativo. Esta contribución usualmente se omite al emplear campos de fuerza no polarizables en cálculos de energía libre.

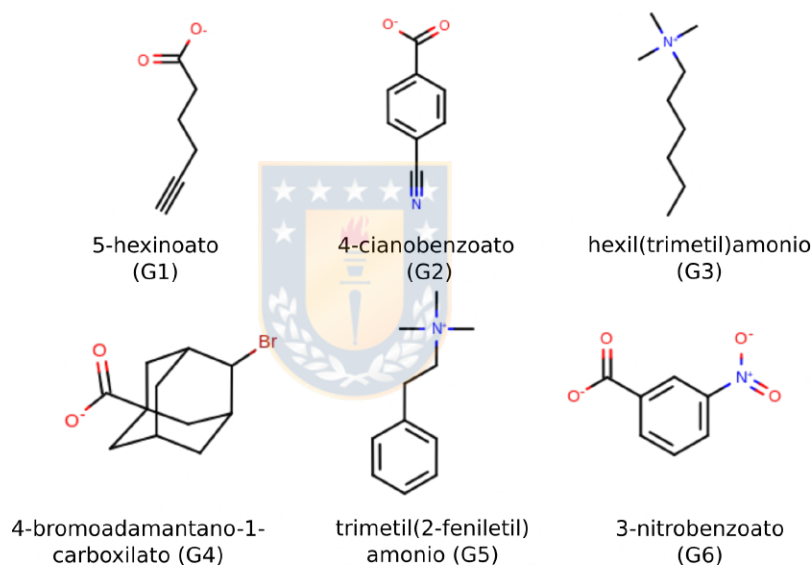


Figura 3.4: Estructura para los huéspedes del SAMPL5 (G1-6) unidos a los anfitriones OA y TEMOA, nombre y microestado más probable de acuerdo a los valores de pK_a (ver Tabla A.1 en el Anexo A).

Las energías libres de unión de Gibbs calculadas, incluyendo la corrección por la energía de polarización y cargas atómicas D-MBIS $\Delta G_{q[\text{DMBIS}(\text{H})]}^{\circ}$ tienen un error cuadrático medio (RMSE) de 2.1 ± 0.3 ligeramente menor al obtenido usando las cargas atómicas RESP/AM1-BCC para el OA y huésped

des respectivamente. Los otros descriptores estadísticos como el coeficiente de correlación R o τ no muestran una diferencia significativa entre los dos conjuntos de carga. Para el anfitrión TEMOA, el RMSE baja de 1.9 ± 0.3 kcal/mol obtenido con las cargas atómicas AM1-BCC a 1.4 ± 0.2 kcal/mol cuando se usó el conjunto de carga D-MBIS. Los otros descriptores estadísticos mencionados también evidencian mejor concordancia con los valores de referencia experimentales. El valor del coeficiente de Kendall τ igual a uno evidencia que las cargas atómicas D-MBIS con corrección de polarización ordenan correctamente los huéspedes de menor a mayor energía libre unión. En aplicaciones de descubrimiento de fármacos, esta clasificación correcta de menor a mayor energía libre de unión facilita explorar diversas variaciones estructurales para un ligando y su posible afinidad a una proteína.

Tabla 3.3: Energías libres de unión para los sistemas OA/TEMOA unidos a los huéspedes del SAMPL5 (G1-6). Experimentales [5] y calculadas usando las cargas atómicas RESP para el OA/TEMOA y cargas atómicas AM1-BCC o D-MBIS para el huésped. Los valores $\langle E_{\text{complejo}}^{\text{pol}} \rangle$ y $\langle E_{\text{solvente}}^{\text{pol}} \rangle$ se obtuvieron con las cargas atómicas RESP/D-MBIS para los anfitriones y huéspedes G1-6, respectivamente. Valores en kcal/mol

Sistema	$\Delta G_{\text{exp}}^{\circ}$	$\Delta G_{\text{q[AM1-BCC]}}^{\circ}$	$\Delta G_{\text{q[DMBIS(H)]}}^{\circ}$	$\langle E_{\text{complejo}}^{\text{pol}} \rangle$	$\langle E_{\text{solvente}}^{\text{pol}} \rangle$
OA					
OA-G1	-5.04 ± 0.01	-7.5 ± 0.2	-3.5 ± 0.5	13.4 ± 0.3	14.0 ± 0.2
OA-G2	-4.25 ± 0.01	-7.8 ± 0.3	-7.5 ± 0.4	18.5 ± 0.4	15.1 ± 0.2
OA-G3	-5.06 ± 0.01	-7.1 ± 0.2	-8.2 ± 0.3	3.4 ± 0.1	2.5 ± 0.1
OA-G4	-9.37 ± 0.00	-12.2 ± 0.3	-10.6 ± 0.8	18.9 ± 0.4	20.9 ± 0.5
OA-G5	-4.50 ± 0.00	-3.4 ± 0.2	-3.5 ± 0.2	4.8 ± 0.1	4.7 ± 0.2
OA-G6	-5.33 ± 0.00	-8.4 ± 0.2	-4.1 ± 0.5	13.6 ± 0.3	18.6 ± 0.3
RMSE		2.6 ± 0.3	2.1 ± 0.3		
MAE		2.5 ± 0.3	1.9 ± 0.3		
R		0.8 ± 0.4	0.7 ± 0.4		
τ		0.4 ± 0.4	0.4 ± 0.4		
TEMOA					
TEMOA-G1	-5.24 ± 0.05	-3.8 ± 0.3	-5.5 ± 0.4	18.9 ± 0.3	13.8 ± 0.1
TEMOA-G2	-5.04 ± 0.03	-7.8 ± 0.2	-5.2 ± 0.5	15.0 ± 0.3	15.8 ± 0.3
TEMOA-G3	-5.94 ± 0.12	-7.9 ± 0.2	-8.7 ± 0.2	3.6 ± 0.1	2.6 ± 0.1
TEMOA-G4	-2.38 ± 0.02	-2.5 ± 0.4	-1.8 ± 0.9	23.2 ± 0.5	19.6 ± 0.4
TEMOA-G5	-3.90 ± 0.02	-4.0 ± 0.2	-2.4 ± 0.3	4.0 ± 0.1	4.5 ± 0.1
TEMOA-G6	-4.52 ± 0.02	-7.5 ± 0.2	-3.3 ± 0.5	11.5 ± 0.3	18.3 ± 0.3
RMSE		1.9 ± 0.3	1.4 ± 0.2		
MAE		1.6 ± 0.2	1.1 ± 0.2		
R		0.7 ± 0.3	0.9 ± 0.1		
τ		0.6 ± 0.3	1.0 ± 0.0		

Con el fin de comprender los diferentes resultados obtenidos con las cargas atómicas D-MBIS para los huéspedes en los anfitriones OA y TEMOA, se analizaron las conformaciones de unión del mismo huésped en los dos anfi-

triones. Las variaciones en la energía libre de unión del mismo huésped unido a dos anfitriones pueden asociarse solo con las diferentes interacciones con el entorno molecular o también puede contener una contribución proveniente de las diferentes orientaciones del huésped en los dos anfitriones [116, 117].

Se abordaron diferentes orientaciones posibles en los modos de enlace para los huéspedes 5-hexinoato (G1) y 4-bromoadamantano-1-carboxilato (G4) en los dos anfitriones. Estos huéspedes presentan una considerable diferencia en la energía libre de unión entre los dos anfitriones usando las cargas atómicas AM1-BCC. Ambos huéspedes en TEMOA se alinean con el eje de simetría del anfitrión. Por el contrario, en el anfitrión OA, están ligeramente inclinados en la cavidad del anfitrión, lo que permite la presencia de moléculas de agua adicionales en la apertura superior del anfitrión (ver Figura A.2 en el Anexo A). Los grupos metilo reducen la abertura superior del TEMOA y alinean al huésped respecto al eje de simetría del anfitrión restringiendo sus conformaciones y limitando el acceso de moléculas de agua a la cavidad. Se concluye que los huéspedes presentan diferentes orientaciones en los dos anfitriones OA/TEMOA y están expuestos a diferentes entornos moleculares. Interesantemente, el promedio de la energía de polarización de G1 refleja la diferencia del entorno molecular en ambos anfitriones.

La variación en las conformaciones de los huéspedes en los respectivos anfitriones conduce a diferencias significativas en su afinidad. Por tanto, un cambio en los conjuntos de carga que describen las interacciones electrostáti-

cas AM1-BCC a D-MBIS en principio podría alterar la conformación preferida del huésped en el respectivo anfitrión. Sin embargo, para los huéspedes G1 y G4 discutidos anteriormente, no se observaron cambios significativos en sus conformaciones de unión al reemplazar el conjunto de cargas AM1-BCC por D-MBIS (ver Figura A.3 en el Anexo A). Por tanto, las cargas atómicas que describen las interacciones contribuyen significativamente a las diferencias en la energía libre de unión. Para estudiar esta contribución, se analizaron los sistemas anfitrión-huésped OA-G1, OA-G6 y TEMOA-G6, que presentan una diferencia notable en la afinidad de unión de 4-5 kcal/mol entre los dos conjuntos de carga. Para el huésped G1, las diferencias surgen principalmente de la variación en las interacciones electrostáticas del huésped que se originan a partir de las cargas atómicas D-MBIS de $\sim 0.2 e$ mayores en el átomo de carbono alquino. Para el 3-nitrobenzoato (G6) las variaciones más significativas entre los dos conjuntos de carga son $\sim 0.4e$ para el átomo de nitrógeno del grupo del grupo nitro, $\sim 0.2 e$ para sus átomos de oxígeno y $\sim 0.2 e$ para el átomo de carbono aromático unido al grupo nitro. En los tres casos (OA-G1, OA-G6 y TEMOA-G6), los resultados obtenidos con cargas atómicas D-MBIS, incluida la corrección de energía de polarización, se acercan más al valor experimental.

Junto con la diferencia absoluta de valores de carga atómica, una diferencia crucial entre nuestra metodología y las cargas atómicas AM1-BCC es que las cargas atómicas D-MBIS son específicas del entorno y pueden variar

entre el estado de unión y no unión del huésped. Para probar el efecto de usar las cargas atómicas específicas del entorno molecular, se seleccionaron los huéspedes 4-cianobenzoato (G2) y hexil(trimetil)amonio (G3) en el anfitrión OA. G2 presenta la diferencia de carga atómica más prominente entre el estado de unión y no unión y G3 la más pequeña (ver Tabla A.2 en el Anexo A). Para ambos huéspedes, se repitieron los cálculos de energía libre usando las mismas cargas atómicas derivadas en solución acuosa para el estado de unión y no unión sin corregir por la energía de polarización. Se obtuvo una energía libre de unión de -3.6 kcal/mol para el huésped G2, que es 3.9 kcal/mol más positivo que el valor reportado en la Tabla 3.3, mientras G3 presentó una desviación de 0.4 kcal/mol. Además, se calculó la afinidad de unión usando las mismas cargas atómicas derivadas en el estado de unión para ambos ambientes y sin corrección por la energía de polarización. Se obtuvo una energía libre de unión de -6.3 kcal/mol que es 1.2 kcal/mol más positivo que el valor reportado en la Tabla 3.3, para G3 hay una desviación de 0.8 kcal/mol (ver Tabla A.3 en el Anexo A). La diferencia entre las cargas atómicas derivadas en el ambiente de unión y no unión para los huéspedes G2 y G3 se puede explicar por la naturaleza química de los huéspedes. G2 porta dos grupos funcionales polares (ciano y ácido carboxílico desprotonado) y exhibe mayor variación en la energía libre de unión al considerar solo las cargas derivadas en un ambiente molecular. G3 presenta un catión de amonio cuaternario con baja polarizabilidad y no presenta cambios significativos al considerar las car-

gas derivadas en un solo ambiente molecular. A partir de este análisis, se hace evidente que la energía libre de unión depende de las cargas atómicas específicas del entorno y la corrección por la energía promedio de polarización, especialmente en sistemas con grupos funcionales polarizables.

Para analizar cómo la diferencia del promedio de la energía de polarización entre los anfitriones y el solvente contribuye a la energía libre de unión, se exploró su dependencia de la naturaleza química del huésped y el ambiente molecular: solución acuosa y anfitriones OA o TEMOA. Basados en la naturaleza química del huésped, se espera que los huéspedes catiónicos (G3, trimetil(2-feniletil)amonio (G5)) presenten energías de polarización menores que los huéspedes de carácter aniónico con mayor polarizabilidad. El promedio de la energía de polarización mostrada en la Figura 3.5 confirma esta expectativa para los tres entornos moleculares. Solo existen variaciones menores entre los valores de los anfitriones o la solución acuosa en los huéspedes G3 y G5. Las diferencias entre los dos entornos moleculares son más significativas para los sistemas aniónicos G1, G2, G4 y 3-nitrobenzoato (G6). El promedio de la energía de polarización aumenta con el tamaño del huésped, como se espera debido a su mayor polarización. Por otra parte, adicional al grupo carboxilo los compuestos G1, G2, G4 y G6 presentan otros grupos funcionales polarizables. El huésped con núcleo de adamantano contiene bromo y presenta los valores más destacados, mientras G1 con un alquino y G2 con un grupo ciano muestran valores similares. Interesantemente, los valores di-

fieren considerablemente entre ambos anfitriones y la solución acuosa. La diferencia entre el anfitrión OA y TEMOA para G1 se relaciona principalmente con la orientación del grupo carboxilo dentro del anfitrión discutido anteriormente. En OA, G1 está completamente solvatado en una conformación ligeramente inclinada, lo que resulta en un promedio de la energía de polarización similar a la obtenida en solución acuosa. La apertura restringida por los grupos metilo del anfitrión TEMOA impiden el acceso al agua, resultando en una energía de polarización más considerable que el OA o la solución acuosa. G4 presenta una variación similar de las conformaciones del huésped en los dos anfitriones y la cavidad del TEMOA menos expuesta al solvente. Se observó una diferencia considerable en la energía de polarización entre el anfitrión y la solución acuosa para G6, posiblemente originada por el grupo nitro y la resonancia entre el grupo carboxilo y el grupo nitro. Se concluye que en los huéspedes aniónicos, la diferencia en el promedio de la energía de polarización entre el estado de unión y no unión contribuye considerablemente a la afinidad de unión y depende de la naturaleza química del huésped, dirigida principalmente por la polarizabilidad. Para los huéspedes catiónicos, la contribución a la afinidad de unión está en el rango de 1-2 kcal/mol.

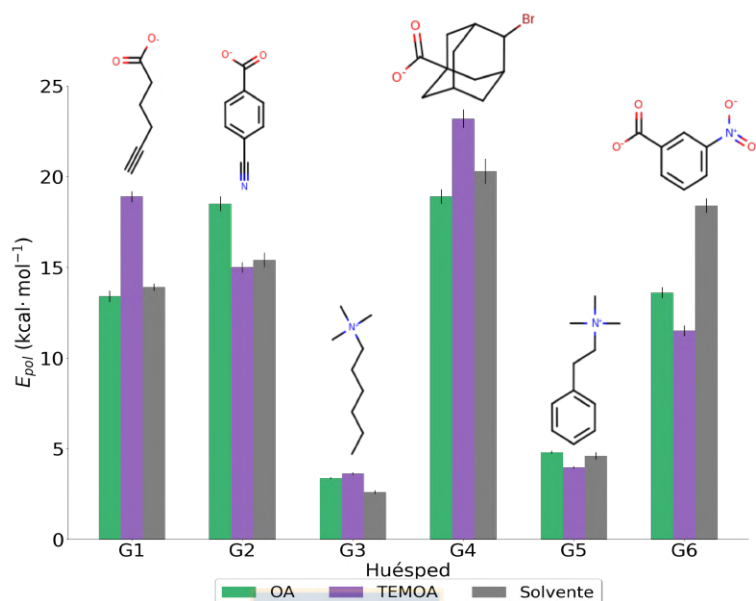


Figura 3.5: Energía requerida para polarizar los huéspedes en los diferentes ambientes moleculares. Las barras verdes y púrpuras consideran cargas atómicas RESP para los anfitriones OA/TEMOA y D-MBIS para cada uno de los huéspedes. Las barras grises representan el estado de no unión en solución acuosa (los valores corresponden al promedio de los valores esperados para la energía de polarización mostrados en la Tabla 3.3).

Finalmente, se exploró si derivar las cargas atómicas D-MBIS para el anfitrión en solución acuosa en lugar de usar las cargas RESP proporcionarían mejores resultados. Se eligió el anfitrión OA donde las cargas atómicas D-MBIS para el huésped no mejoraron los resultados obtenidos con el conjunto de cargas atómicas AM1-BCC. Además, el ordenamiento de los huéspedes de acuerdo a su afinidad fue insatisfactorio. Con las nuevas cargas atómicas D-MBIS para el anfitrión OA, se recalcularon las cargas atómicas D-MBIS para los huéspedes y el promedio de energía de polarización del

huésped en el anfitrión o en solución acuosa. Las energías libres de unión de Gibbs calculadas, incluyendo la corrección por la energía de polarización $\Delta G_{q[\text{DMBIS}]}^{\circ}$ se resumen en la Tabla 3.4. La mayoría de los descriptores estadísticos reflejan un mejor rendimiento respecto a los valores experimentales, reduciendo el RMSE a 1.8 ± 0.2 kcal/mol y aumentando el coeficiente de Kendall τ a 0.6 ± 0.4 . **Este resultado evidencia que las cargas atómicas D-MBIS mejoran considerablemente los resultados obtenidos con las cargas atómicas RESP/AM1-BCC para el anfitrión y huésped respectivamente, además la distribución de carga del anfitrión también altera la energía libre de unión.**



Tabla 3.4: Energías libres de unión para el sistema OA unido a los huéspedes del SAMPL5 (G1-6). Experimentales [5] y calculadas usando cargas atómicas RESP para el anfitrión OA y D-MBIS para los huéspedes (D-MBIS(H)) o cargas atómicas D-MBIS para el anfitrión OA y los huéspedes (D-MBIS). Los valores de $\langle E_{\text{complejo}}^{\text{pol}} \rangle$ y $\langle E_{\text{solvente}}^{\text{pol}} \rangle$ fueron obtenidas con las cargas atómicas D-MBIS para el anfitrión y los huéspedes G1-6. Valores en kcal/mol

Sistema	$\Delta G_{\text{exp}}^{\circ}$	$\Delta G_{\text{q[DMBIS(H)]}}^{\circ}$	$\Delta G_{\text{q[DMBIS]}}^{\circ}$	$\langle E_{\text{complejo}}^{\text{pol}} \rangle$	$\langle E_{\text{solvente}}^{\text{pol}} \rangle$
OA					
OA-G1	-5.04 ± 0.01	-3.5 ± 0.5	-2.5 ± 0.5	12.1 ± 0.2	14.6 ± 0.3
OA-G2	-4.25 ± 0.01	-7.5 ± 0.4	-5.9 ± 0.5	14.3 ± 0.3	15.4 ± 0.3
OA-G3	-5.06 ± 0.01	-8.2 ± 0.3	-4.2 ± 0.3	4.2 ± 0.1	2.6 ± 0.1
OA-G4	-9.37 ± 0.00	-10.5 ± 0.8	-11.0 ± 0.7	18.8 ± 0.5	20.8 ± 0.3
OA-G5	-4.50 ± 0.00	-3.5 ± 0.2	-2.1 ± 0.2	6.1 ± 0.1	4.8 ± 0.1
OA-G6	-5.33 ± 0.00	-4.1 ± 0.5	-6.6 ± 0.5	17.6 ± 0.3	18.6 ± 0.3
RMSE		2.1 ± 0.3	1.8 ± 0.2		
MAE		1.9 ± 0.3	1.7 ± 0.2		
R		0.7 ± 0.4	0.8 ± 0.4		
τ		0.4 ± 0.4	0.6 ± 0.4		

3.4. Energías libres de unión entre CB7 y huéspedes con núcleo aromático

En primera instancia se estudiaron los sistemas CB7 unidos a los huéspedes con núcleo aromático BG1-8 (ver Figura 3.6). Los huéspedes presentan un carácter catiónico con pequeñas diferencias en la energía de polarización entre el estado de unión y no unión [118], por tanto contribuye $\simeq 1$ kcal/mol a la energía libre de unión (ver Tabla 3.5). Solo 2-(3,4-dihidroxifenil)etilamonio (BG1) presenta una diferencia de 4.0 kcal/mol. En este caso el costo de polarización es mayor en el estado de unión debido a la fuerte polarización de los grupos hidroxilo expuestos a los grupos carbonilo que bordean los dos portales simétricos del CB7. Además, los enlaces de hidrógeno intramoleculares entre los dos grupos hidroxilo compiten con los intermoleculares generados con moléculas de agua en estado de no unión.

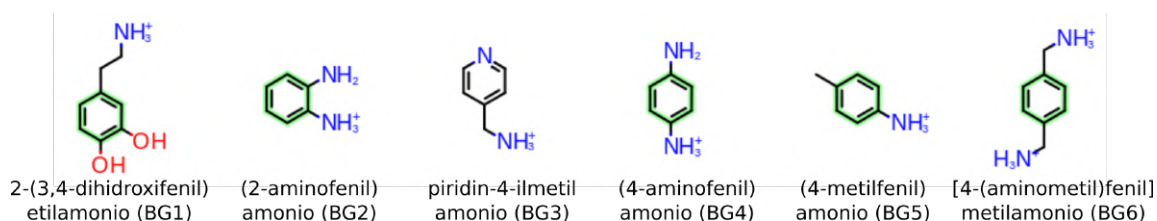


Figura 3.6: Estructura para los huéspedes con núcleo aromático (BG1-6) unidos al anfitrión CB7, nombre y microestado más probable de acuerdo a los valores de pK_a (ver Tabla B.1 en el Anexo B).

Por otra parte las cargas atómicas D-MBIS derivadas para el anfitrión y el huésped predicen la energía libre de unión mejor que el conjunto de cargas atómicas AM1-BCC, evidenciado por valores más pequeños de RMSE, MAE y un incremento del coeficiente de Kendall τ (ver Tabla 3.5). La mejora en los parámetros estadísticos empleando las cargas D-MBIS confirma los resultados obtenidos previamente para los anfitriones OA/TEMOA con 6 diferentes huéspedes. El huésped (4-aminofenil)amonio (BG4) presenta la mayor desviación respecto al valor experimental. Una posible explicación podría estar asociada a que ambos grupos amonio podrían cambiar en su estado de protonación a diferentes niveles de exposición al solvente. El valor de pK_a del (2-aminofenil)amonio (BG2) está cerca del valor de pH experimental de 4.74 (ver Tabla B.1 en el Anexo B) y sugiere que una forma neutra para el huésped puede coexistir en solución acuosa. Se realizaron cálculos de energía libre de unión de la forma neutra con las cargas atómicas D-MBIS, pero mostró una desviación mayor del valor experimental (ver Tabla B.2 en el Anexo B). Cuando el valor de pK_a está cerca del pH experimental, coexisten múltiples estados de protonación del huésped y la unión al anfitrión también podría inducir a cambios en el estado de protonación. Los cambios de protonación de los huéspedes BG2 y BG4 no se consideran en este trabajo e incluirlos en simulaciones de pH constante está fuera del alcance de esta tesis.

Tabla 3.5: Energías libres de unión para el CB7 unido a huéspedes con núcleo aromático (BG1-6). Experimentales [5] y calculadas utilizando cargas atómicas RESP para el CB7 y AM1-BCC o cargas atómicas D-MBIS para el CB7 y huéspedes, se consideran los parámetros del campo de fuerza GAFF. Los valores $\langle E_{\text{complejo}}^{\text{pol}} \rangle$ y $\langle E_{\text{solvente}}^{\text{pol}} \rangle$ se obtuvieron con las cargas atómicas D-MBIS para el anfitrión y huéspedes BG1-6. Valores en kcal/mol

Sistema	$\Delta G_{\text{exp}}^{\circ}$	$\Delta G_{\text{q[AM1-BCC]}}^{\circ}$	$\Delta G_{\text{q[DMBIS]}}^{\circ}$	$\langle E_{\text{complejo}}^{\text{pol}} \rangle$	$\langle E_{\text{solvente}}^{\text{pol}} \rangle$
CB7-BG1	-6.3 ± 0.1	-10.1 ± 0.4	-5.1 ± 0.4	14.7 ± 0.2	10.7 ± 0.2
CB7-BG2	-6.7 ± 0.1	-5.5 ± 0.2	-6.8 ± 0.3	7.3 ± 0.2	8.4 ± 0.2
CB7-BG3	-7.6 ± 0.1	-5.4 ± 0.4	-8.3 ± 0.4	9.1 ± 0.2	10.1 ± 0.2
CB7-BG4	-8.6 ± 0.1	-8.1 ± 0.3	-4.9 ± 0.5	6.0 ± 0.2	7.1 ± 0.1
CB7-BG5	-9.4 ± 0.1	-8.9 ± 0.4	-8.1 ± 0.3	6.0 ± 0.1	6.3 ± 0.1
CB7-BG6	-12.6 ± 0.1	-7.3 ± 0.5	-10.0 ± 0.4	6.2 ± 0.1	5.4 ± 0.1
RMSE		2.9 ± 0.6	2.0 ± 0.5		
MAE		2.3 ± 0.6	1.6 ± 0.4		
R		0.0 ± 0.5	0.7 ± 0.3		
τ		-0.1 ± 0.5	0.5 ± 0.4		

Para cerciorarse que el conjunto de cargas atómicas D-MBIS mejoran la descripción de las interacciones anfitrión-huésped, hay que asegurarse que las diferencias observadas en la energía libre de unión entre los dos conjuntos de cargas no se originen por conformaciones diferentes. El análisis conformacional realizado (ver Figuras B.1-B.6 y Tablas B.3-B.8 del Anexo B) confirma que no hay variación significativa en las conformaciones adoptadas por los huéspedes al modificar el conjunto de cargas atómicas.

Por tanto, se concluye que las cargas atómicas D-MBIS para el CB7

y los huéspedes BG1-6 mejoran las predicciones de afinidad de unión, con excepción de BG4. Además, los huéspedes catiónicos BG2-6 poseen pequeñas diferencias en la energía de polarización y contribuyen $\simeq 1$ kcal/mol a la energía libre de unión.



3.5. Energías libres de unión entre CB7 y huéspedes con núcleo de adamantano

Después de los resultados observados en la predicción de energía libre entre el CB7 y los huéspedes BG1-6, se centra el interés en los cálculos de energía libre de unión para los huéspedes con núcleo de adamantano AG1-8 (ver Figura 3.7) unidos al anfitrión CB7. Se espera una mejora en los descriptores estadísticos tal como se han obtenido en los anteriores sistemas usando cargas atómicas D-MBIS para el anfitrión y el huésped, sin embargo las predicciones de la energía libre de unión presentan peores descriptores estadísticos ver Tabla 3.6. De manera similar a los resultados obtenidos para los huéspedes aromáticos BG2-6, para los huéspedes con núcleo de adamantano AG1-8, tienen una diferencia entre la energía polarización en el estado de unión y no unión de $\simeq 1$ kcal/mol, excepto para el huésped 1-adamantil(2-aminoiletil)amonio (AG3) donde la diferencia es de 2.7 kcal/mol.

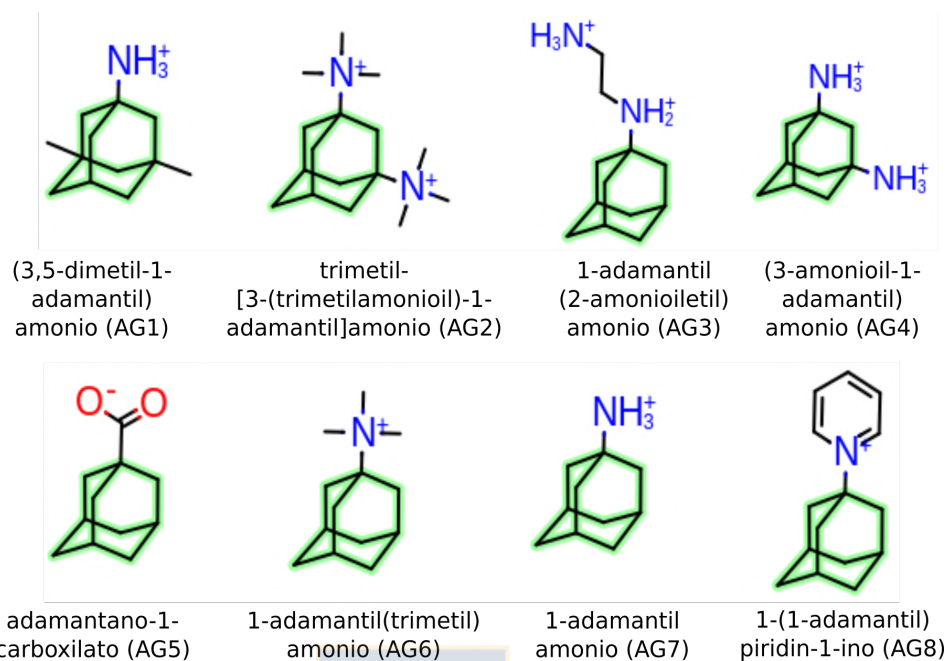


Figura 3.7: Estructura para los huéspedes con núcleo de adamantano (AG1-8) unidos al anfitrión CB7, nombre y microestado más probable de acuerdo a los valores de pK_a (ver Tabla B.1 en el Anexo B).

Para entender el bajo rendimiento de las cargas atómicas D-MBIS en los huéspedes con núcleo de adamantano comparados con los huéspedes de núcleo aromático, se analizó que tipo de interacción predomina en ambos sistemas anfitrión-huésped con la teoría de perturbación adaptada a la simetría (SAPT0). Se tomaron conformaciones representativas de las simulaciones de dinámica molecular y se incluyó las moléculas de agua de coordinación para el huésped (4-metilfenil)amonio (BG5) y 1-adamantilamonio (AG7) que representan a cada grupo de huéspedes con núcleo aromático y de adamantano respectivamente (ver Figura 3.8).

Tabla 3.6: Energías libres de unión para el CB7 unido a huéspedes con núcleo de adamantano (AG1-8). Experimentales [5] y calculadas utilizando cargas atómicas RESP para el CB7 y AM1-BCC o cargas atómicas D-MBIS para el CB7 y huéspedes y los parámetros del campo de fuerza GAFF. Los valores $\langle E_{\text{complejo}}^{\text{pol}} \rangle$ y $\langle E_{\text{solvente}}^{\text{pol}} \rangle$ se obtuvieron con las cargas atómicas D-MBIS para el anfitrión y huéspedes AG1-8. Valores en kcal/mol

Sistema	$\Delta G_{\text{exp}}^{\circ}$	$\Delta G_{\text{q[AM1-BCC]}}^{\circ}$	$\Delta G_{\text{q[DMBIS]}}^{\circ}$	$\langle E_{\text{complejo}}^{\text{pol}} \rangle$	$\langle E_{\text{solvente}}^{\text{pol}} \rangle$
CB7-AG1	-6.0 ± 0.1	-6.8 ± 0.3	-10.1 ± 0.7	7.9 ± 0.1	6.5 ± 0.1
CB7-AG2	-6.6 ± 0.1	-3.0 ± 0.2	-2.1 ± 0.5	3.6 ± 0.1	3.4 ± 0.1
CB7-AG3	-18.2 ± 0.1	-19.4 ± 0.7	-24.0 ± 0.6	15.5 ± 0.2	12.8 ± 0.1
CB7-AG4	-11.3 ± 0.1	-4.2 ± 0.5	-12.0 ± 0.4	7.9 ± 0.1	7.8 ± 0.1
CB7-AG5	-11.6 ± 0.1	-17.0 ± 0.5	-17.7 ± 0.6	19.6 ± 0.2	20.2 ± 0.4
CB7-AG6	-16.7 ± 0.1	-22.4 ± 0.2	-23.5 ± 0.3	3.4 ± 0.1	2.5 ± 0.1
CB7-AG7	-17.2 ± 0.1	-20.6 ± 0.4	-22.7 ± 0.3	6.4 ± 0.1	6.0 ± 0.1
CB7-AG8	-16.8 ± 0.8	-22.4 ± 0.4	-24.7 ± 0.2	2.7 ± 0.1	2.0 ± 0.1
RMSE		4.6 ± 0.6	5.6 ± 0.6		
MAE		4.1 ± 0.7	5.2 ± 0.7		
R		0.9 ± 0.1	0.9 ± 0.0		
τ		0.5 ± 0.3	0.7 ± 0.2		

Las interacciones de intercambio en el voluminoso huésped AG7 son mayores en comparación con las del huésped BG5. AG7 también presenta una mayor contribución de las interacciones de dispersión comparadas con la interacción electrostática favorable en BG5 (relación dispersión (Disp)/ electrostática (Elec): 0.36 (BG5) vs 0.68 (AG7)). Las moléculas de agua tienen una participación significativa en la energía de interacción de BG5, posiblemente porque su menor tamaño permite un mejor acceso al portal de entrada

del CB7. El núcleo de adamantano en el voluminoso AG7 contribuye un 26 % a la energía de interacción total, mientras que el anillo aromático contribuye solo un 18 %.

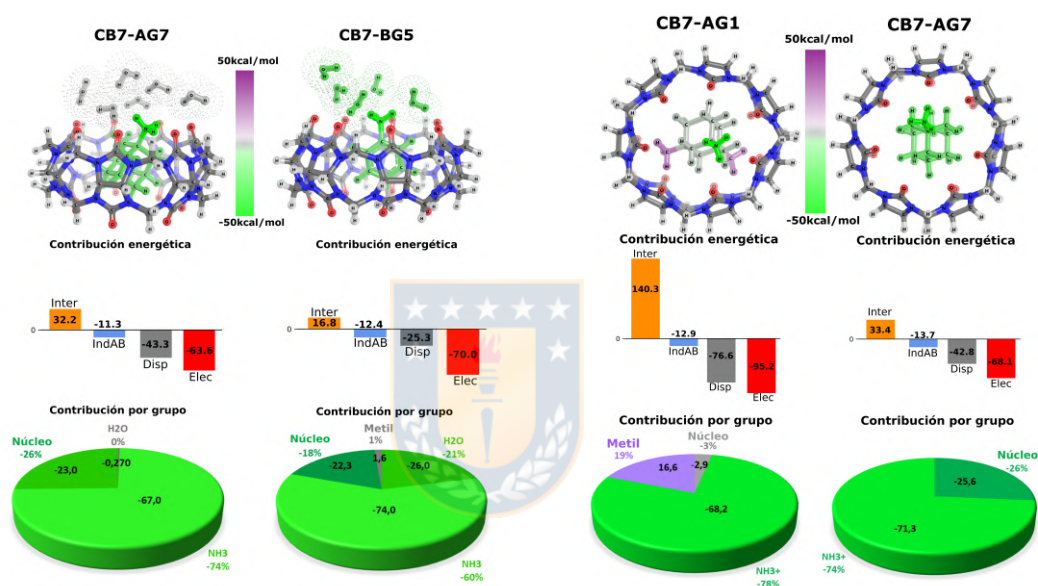


Figura 3.8: Análisis de perturbación adaptado a la simetría (SAPT) de las interacciones en huéspedes con núcleo aromático y adamantano en el nivel SAPT0/jun-ccVDZ. Izquierda: Comparación de las interacciones entre huéspedes con núcleo aromático y de adamantano de una conformación representativa en solución acuosa, incluidas las moléculas de agua que solvatan el grupo de amonio cargado. Los átomos del huésped y las moléculas de agua se colorean según su contribución a la energía de interacción. Las barras en el medio muestran la contribución electrostática (Elec), dispersión (Disp), inducción del anfitrión (A) al huésped (B) (IndAB) e intercambio (Inter) de la energía de interacción total expresada en kcal/mol. Los gráficos circulares de la parte inferior muestran la contribución de cada grupo funcional a la energía de interacción total en porcentaje y la energía de interacción en kcal/mol. Derecha: Las mismas representaciones para dos huéspedes con núcleo de adamantano resaltan la gran contribución de intercambio repulsivo de AG1 y la contribución repulsiva de sus dos grupos metilo.

Si el origen de la gran desviación en la energía libre de unión de los huéspedes con núcleo de adamantano corresponde a interacciones repulsivas y de dispersión incorrectamente descritas, este tipo de interacciones deberían dominar en los huéspedes con grandes desviaciones en la afinidad de unión. La energía libre de unión calculada de (3,5-dimetil-1-adamantil)amonio (AG1) que posee dos grupos metilo se desvía en 4.1 kcal/mol del valor obtenido con el conjunto de cargas D-MBIS. La composición de la energía total de interacción para este huésped (ver Figura 3.8) muestra una contribución de intercambio repulsivo significativamente mayor en comparación con AG7 combinada con una mayor contribución de dispersión en comparación con la interacción electrostática atractiva (relación dispersión (Disp)/electrostática (Elec): 0.80 (AG1) frente a 0.63 (AG7)). La gran contribución repulsiva se origina en la superposición de la densidad de electrones del anfitrión y del huésped que apunta a interacciones estéricas desfavorables. Los dos grupos metilo en AG1 contribuyen significativamente a esta interacción repulsiva, como lo demuestra la partición F-SAPT de la interacción total por grupos que se muestra en la parte inferior de la Figura 3.8. El huésped 1-(1-adamantil)piridin-1-ino (AG8) también presenta una gran desviación en la energía libre de enlace de 7.9 kcal/mol. En comparación con la AG7, este huésped posee un grupo piridina con carga positiva en lugar del grupo amonio. Curiosamente, revela una interacción repulsiva similar a AG7, pero una mayor contribución de las interacciones atractivas de dispersión (relación dispersión (Dis)/electrostática(Elst):

0.86 (AG8) frente a 0.80 (AG1)); ver Figura B.7 del Anexo B. Este análisis permite deducir que los huéspedes con núcleo de adamantano tienen una mayor contribución de la interacción repulsiva de intercambio e interacción atractiva de dispersión. Una posible explicación de las grandes desviaciones obtenidas en la energía libre de unión, podría ser una sobre estimación de la interacción de dispersión por los coeficientes C_6 demasiado grandes en el campo de fuerza empleado (GAFF), como informó recientemente Rowley[119, 80].

Del análisis SAPT, concluimos que los huéspedes con núcleo de adamantano tienen una gran contribución de interacción de intercambio repulsivo y una considerable interacción de dispersión atractiva. Los huéspedes AG1 y AG8 representan casos límite donde la repulsión o la interacción de dispersión gobiernan las interacciones en comparación con AG7. Según este análisis, la mejora de los parámetros en los campos de fuerza no polarizables que representan las interacciones de intercambio repulsivo y la interacción de dispersión atractiva podría conducir a una mejora en la predicción de la afinidad de unión para los huéspedes con núcleo de adamantano.

Tabla 3.7: Energías libres de unión para el CB7 unido a huéspedes con núcleo de adamantano (AG1-8). Experimentales [5] y calculadas usando las cargas atómicas RESP para el CB7 y AM1-BCC para los huéspedes con parámetros de Lennard-Jones(LJ) del campo de fuerza GAFF (q[AM1-BCC]) o parámetros de LJ derivados de la partición MBIS de la densidad electrónica polarizada del huésped combinada con las cargas atómicas D-MBIS derivadas para el CB7 y huéspedes (LJ-q[D-MBIS]) o usando solo las cargas atómicas D-MBIS del CB7 y las cargas atómicas D-MBIS del huésped derivado en el estado de no unión (LJ-q[D-MBIS(solv)]). Valores en kcal/mol

Sistema	$\Delta G_{\text{exp}}^{\circ}$	$\Delta G_{\text{q[AM1-BCC]}}^{\circ}$	$\Delta G_{\text{LJ-q[D-MBIS]}}^{\circ}$	$\Delta G_{\text{LJ-q[D-MBIS(solv)]}}^{\circ}$
CB7-AG1	-6.0 ± 0.1	-6.8 ± 0.3	-6.3 ± 0.3	-4.0 ± 0.6
CB7-AG2	-6.6 ± 0.1	-3.0 ± 0.2	-2.5 ± 0.4	0.8 ± 0.4
CB7-AG3	-18.2 ± 0.1	-19.4 ± 0.7	-16.3 ± 0.4	-20.5 ± 0.3
CB7-AG4	-11.3 ± 0.1	-4.2 ± 0.5	-9.4 ± 0.5	-10.5 ± 0.5
CB7-AG5	-11.6 ± 0.1	-17.0 ± 0.5	-14.0 ± 0.5	-9.5 ± 0.5
CB7-AG6	-16.7 ± 0.1	-22.4 ± 0.2	-19.2 ± 0.3	-14.0 ± 0.2
CB7-AG7	-17.2 ± 0.1	-20.6 ± 0.4	-18.1 ± 0.3	-18.6 ± 0.3
CB7-AG8	-16.8 ± 0.1	-22.4 ± 0.4	-19.5 ± 0.3	-15.0 ± 0.3
RMSE		4.6 ± 0.6	2.4 ± 0.3	3.2 ± 0.9
MAE		4.1 ± 0.7	2.1 ± 0.4	2.5 ± 0.6
R		0.9 ± 0.1	0.9 ± 0.1	0.9 ± 0.0
τ		0.5 ± 0.3	0.6 ± 0.2	0.9 ± 0.1

Con base en la gran contribución de intercambio repulsivo y las interacciones de dispersión de los huéspedes con un núcleo de adamantano descrito anteriormente, se derivaron los parámetros LJ en solución acuosa para reemplazar los parámetros de GAFF en el flujo de trabajo computacional empleado. Se usaron los mismos parámetros descritos anteriormente para la derivación de las cargas atómicas D-MBIS empleando cálculos QM/MM

y simulaciones de dinámica molecular para obtener la densidad de electrones polarizados del huésped en cada configuración. Después de particionar la densidad electrónica con el método MBIS, se calculó un volumen efectivo de la densidad AIM para cada átomo en la molécula $\langle r^3 \rangle_{\text{aim}} = \int r^3 \rho_{\text{aim}} d\tau$. Se promediaron los valores de $\langle r^3 \rangle_{\text{aim}}$ para cada átomo en todas las configuraciones muestreadas para cada ciclo. Después de 5 ciclos de actualizaciones de topología, los últimos valores promedio de $\langle r^3 \rangle_{\text{aim}}$ se usan para derivar los parámetros de LJ (σ y ϵ) en conjunto con el valor esperado de $\langle r^3 \rangle_{\text{aislado}}$ para el átomo aislado (ver 2.2). Se derivan los dos parámetros del potencial de LJ para todos los átomos en cada huésped con el flujo de trabajo computacional considerando varias configuraciones en solución acuosa y se promedia los coeficientes para cada tipo de átomo en el campo de fuerza GAFF. Con los parámetros de LJ obtenidos, se calcula un nuevo conjunto de cargas atómicas D-MBIS para cada huésped y posteriormente se repiten los cálculos de energía libre utilizando el nuevo conjunto de parámetros de campo de fuerza no enlazantes. Curiosamente, la cuarta columna de la tabla 3.7 muestra que las afinidades de unión de D-MBIS para los 8 huéspedes principales de adamantano concuerdan mejor con los valores experimentales, reduciendo el RMSE y MAE por un factor de dos, manteniendo un coeficiente de correlación de Pearson R de 0.9 y un valor casi igual de Kendall τ .

Para la mayoría de los huéspedes, los parámetros del campo de fuerza derivados desde la partición MBIS reproducen las afinidades de enlace me-

jor que los que usan cargas atómicas D-MBIS con parámetros LJ del campo de fuerza GAFF, excepto para el huésped (3-amoniol-1-adamantil)amónio (AG4) como se muestra en la Figura 3.9. AG4 posee dos grupos de amonio cargados en el núcleo de adamantano, donde pueden ocurrir cambios en el estado de protonación dependiendo de la exposición al solvente de los dos grupos. Pero como se mencionó anteriormente, incluir cambios de estado de protonación en las simulaciones está fuera del alcance de este trabajo.

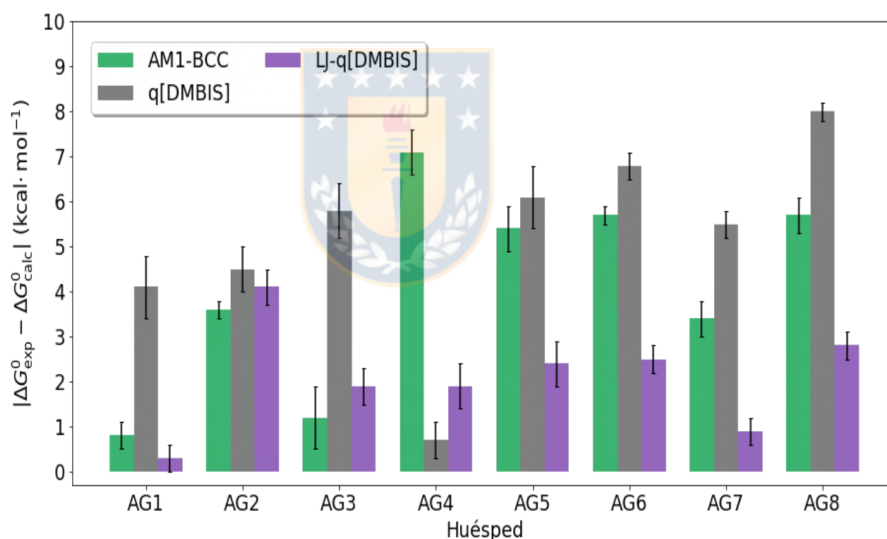


Figura 3.9: Error absoluto en la afinidad de unión para los huéspedes de núcleo adamantano usando diferentes cargas atómicas y parámetros de Lennard-Jones (LJ). Cargas AM1-BCC/GAFF LJ (barras en verde), cargas D-MBIS / GAFF LJ (barras en gris) y cargas D-MBIS /D-MBIS LJ (barras en púrpura). Se tomaron los parámetros de enlace del campo de fuerza GAFF para el CB7 y huéspedes AG1-G8.

Antes de que se pueda concluir que los parámetros de no enlace del campo de fuerza derivados con la partición MBIS mejoran la descripción de las in-

teracciones, tenemos que asegurar que las conformaciones adoptadas por los huéspedes sean similares al emplear diferentes parámetros del campo de fuerza. El análisis de las orientaciones relativas más probables que puede optar el huésped durante la simulación con las interacciones electrostáticas y van-der-Waals activas (estados termodinámicos con $\lambda = 1$) (ver Figura B.8 y Tablas B.9-B.11 del Anexo B) muestra una similitud en las orientaciones relativas para los huéspedes AG1-8 unido al CB7 y empleando diferentes conjuntos de carga y parámetros de LJ empleados. Sin embargo, los parámetros de no enlace alteran la función de distribución radial de las moléculas de agua (ver Figura B.9 en el Anexo B). La introducción de los parámetros de LJ reduce la primera capa de solvatación del huésped debido a una mayor interacción repulsiva. Al mismo tiempo, aumenta la probabilidad de encontrar moléculas de agua a distancias superiores a 2.5 Å.

Para entender los resultados asociados a la derivación de los nuevos parámetros no enlazantes, se comparan los parámetros de LJ del campo de fuerza GAFF con los obtenidos considerando la partición MBIS. Los valores de σ de los átomos de hidrógeno h_n y h_x unidos al átomo de nitrógeno del grupo amonio aumentan significativamente con el método de partición MBIS y explican la diferencia observada en la función de distribución radial del solvente. Para los otros tipos de átomos, los cambios en σ son menores a excepción de los átomos de oxígeno y el átomo de carbono en el grupo carboxilo. Adicionalmente se observa que la profundidad del potencial ϵ obtenido de la

partición MBIS es menor que la de GAFF. Las diferencias son más pronunciadas para los tipos de átomos electronegativos n4, na y o. En general, los átomos de hidrógeno derivados con la partición MBIS presentan valores de ϵ más pequeños que los del campo de fuerza GAFF, especialmente para ha. (ver Figura 3.10).

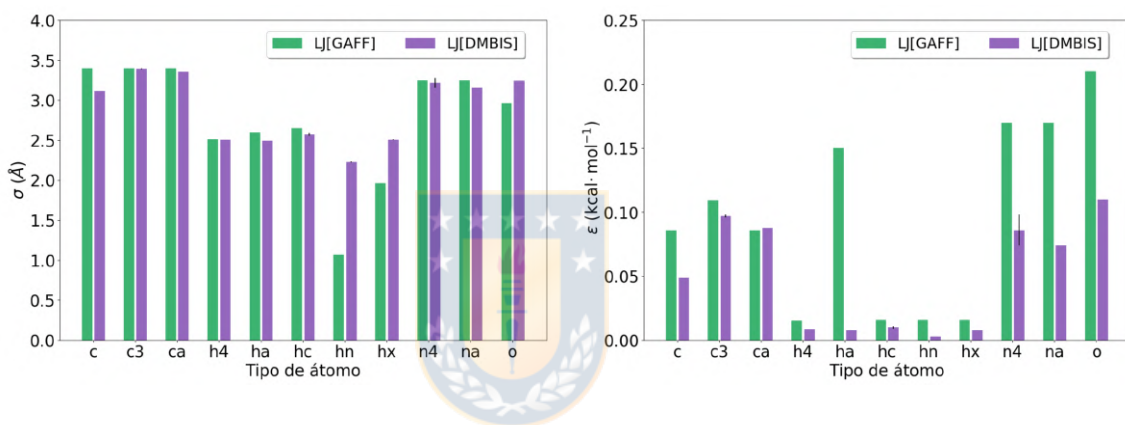


Figura 3.10: Parámetros de Lennard-Jones derivados considerando la partición MBIS de la densidad electrónica polarizada, se muestran los promedios para cada tipo de átomo de los huéspedes con núcleo de adamantano. El radio derivado MBIS al mínimo del potencial se compara con los valores del campo de fuerza GAFF para diferentes tipos de átomos a la izquierda y la profundidad derivada MBIS del potencial a la derecha.

En base a los resultados obtenidos para los sistemas anfitrión-huésped estudiados, se llega a la conclusión de que la partición MBIS de la densidad electrónica polarizada proporciona cargas atómicas y parámetros de LJ que mejoran las predicciones de energía libre de unión.

Finalmente, se intenta reducir el costo computacional de tener que realizar

cálculos de estructura electrónica en el estado unión y no unión del huésped, para ello se emplea el mismo conjunto de carga atómica D-MBIS de la solución acuosa para el cálculo de energía libre en el estado de unión del anfitrión CB7. Aunque las energías libres de unión obtenidas presentan una mayor desviación respecto a los valores experimentales, todavía muestran una buena correlación y un orden correcto en la energía de afinidad (ver Tabla 3.7). Solo adamantano-1-carboxilato (AG5) presenta una gran desviación debido a su carácter aniónico y la diferencia significativa entre la energía de polarización en los dos entornos moleculares, de acuerdo con nuestros resultados anteriores[118]. El pK_a de AG5 está muy cerca del valor de pH experimental y también podrían surgir desviaciones en la energía libre de unión calculada respecto al valor experimental debido a la contribución de los diferentes estados de protonación del huésped y un posible cambio de protonación en el proceso de unión.

3.6. Energías libres de unión entre la Lisozima T4 doble mutante L99A/M102 y ligandos con núcleo aromático

Con el protocolo de simulación validado se calculan las energías libres de unión para los sistemas T4 lisozima L99A/M102Q con ligandos de núcleo aromático L1-8 (ver Figura 3.11). El sitio de unión L99A/M102Q presenta una cavidad con mayor polaridad respecto a L99A por la presencia de glutamina que introduce polaridad y potencial para enlace de hidrógeno [60]. Sin embargo, el resto de los residuos presentes en la cavidad tienen características apolares. Según la experiencia adquirida con sistemas anfitrión-huésped estudiados previamente, la derivación de los parámetros no enlazantes del campo de fuerza derivados de la densidad electrónica molecular del huésped solo en la fase acuosa mejoran la predicción de la energía libre de unión respecto al sistema de referencia. Adicionalmente este protocolo reduce el costo computacional comparado con los sistemas donde se hace una derivación de los parámetros no enlazantes en el ambiente de unión y no unión. Bajo esta panorámica se optó por derivar las cargas atómicas D-MBIS y parámetros de LJ para los ligandos L1-8 solo en solución acuosa y reemplazarlos en la fase de unión para su posterior cálculo de energía libre. La Tabla 3.8 resume los valores experimentales de la energía libre de unión de los 8 ligandos unidos

al sitio de unión L99A/M102Q de la lisozima T4 y los resultados usando el campo de fuerza AMBERff14SB [120] para el sistema proteico Lisozima T4 L99A/M102Q y variando las cargas atómicas y parámetros de Lennard-Jones para los ligandos. Los valores de afinidad de unión de referencia consideran las cargas AM1-BCC para el ligando con parámetros de LJ del campo de fuerza GAFF (q[AM1-BCC]). Adicionalmente se considera la afinidad de unión usando las cargas atómicas D-MBIS para el ligando derivados en la fase de no unión en conjunto con los parámetros de LJ del campo de fuerza GAFF (q[D-MBIS(H)]) y las cargas atómicas con parámetros de LJ obtenidos a partir de la partición MBIS de la densidad electrónica molecular del huésped en fase acuosa (LJ-q[D-MBIS(H)]).

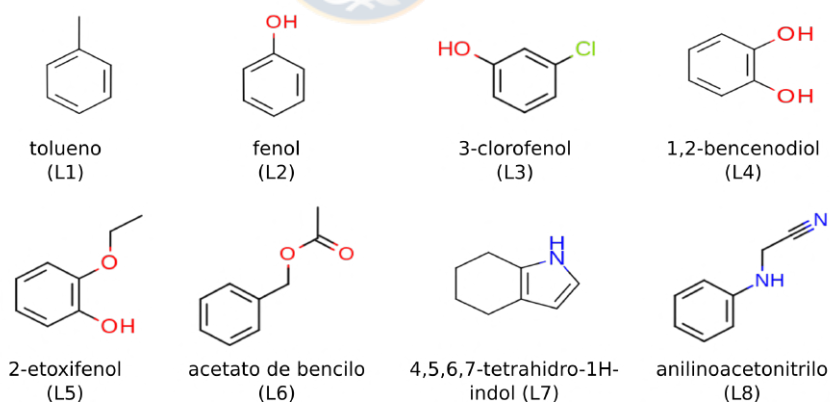


Figura 3.11: Estructura para los ligandos de núcleo aromático (L1-8) unidos a la proteína lisozima T4 L99A/M102Q, nombre y microestado más probable de acuerdo a los valores de pK_a (ver Tabla C.1 en el Anexo C).

Tabla 3.8: Energías libres de unión para la lisozima T4 L99A/M102Q unida a ligandos con núcleo aromático (L1-8). Experimentales [5] y calculadas usando el campo de fuerza AMBERff14SB para el sistema proteico Lisozima T4 L99A/M102Q y variando las cargas atómicas y parámetros de Lennard-Jones para los ligandos. Los valores de afinidad de unión de referencia considera las cargas AM1-BCC para el ligando con parámetros de LJ del campo de fuerza GAFF (q[AM1-BCC]). Adicionalmente se reporta la afinidad de unión usando las cargas atómicas D-MBIS para el ligando derivados en la fase de no unión en conjunto con los parámetros de LJ del campo de fuerza GAFF (q[DMBIS(solv)]) y las cargas atómicas con parámetros de LJ obtenidos a partir de la partición MBIS de la densidad electrónica molecular del ligando en fase acuosa (LJ-q[DMBIS(solv)]). Valores en kcal/mol

Sistema	$\Delta G_{\text{exp}}^{\circ}$	$\Delta G_{\text{q[AM1-BCC]}}^{\circ}$	$\Delta G_{\text{q[DMBIS(solv)]}}^{\circ}$	$\Delta G_{\text{LJ-q[DMBIS(solv)]}}^{\circ}$
Lisozima-L1	-4.93	-3.6 ± 0.2	-3.9 ± 0.2	-3.3 ± 0.3
Lisozima-L2	-5.24	-6.5 ± 0.3	-5.0 ± 0.3	-4.7 ± 0.2
Lisozima-L3	-5.51	-6.2 ± 0.3	-4.1 ± 0.3	-2.6 ± 0.3
Lisozima-L4	-4.16 ± 0.03	-2.9 ± 0.4	-2.8 ± 0.4	-2.2 ± 0.3
Lisozima-L5	-4.02 ± 0.03	-2.7 ± 0.4	-2.7 ± 0.4	-2.8 ± 0.4
Lisozima-L6	-4.48 ± 0.03	-3.1 ± 0.4	-0.8 ± 0.3	-0.2 ± 0.3
Lisozima-L7	-4.61 ± 0.09	-2.9 ± 0.4	-0.3 ± 0.3	-2.3 ± 0.2
Lisozima-L8	-5.5 ± 0.2	-6.3 ± 0.4	-3.7 ± 0.4	-3.7 ± 0.4
RMSE		1.3 ± 0.1	2.3 ± 0.5	2.3 ± 0.5
MAE		1.2 ± 0.1	1.9 ± 0.4	2.0 ± 0.4
R		0.8 ± 0.2	0.4 ± 0.2	0.4 ± 0.3
τ		0.9 ± 0.2	0.6 ± 0.2	0.5 ± 0.3

Los 4 descriptores estadísticos no muestran una mejorara al variar los parámetros no enlazantes del campo de fuerza respecto a los sistemas de referencia con los dos conjuntos de resultados q[D-MBIS(H)] y (LJ-q[D-MBIS(H)]) al incluir las cargas atómicas D-MBIS comparado con las cargas atómicas y parámetros de LJ derivados de la partición MBIS de la densidad

electrónica molecular para el ligando en fase de no unión. Es importante acotar que el rango de energías libres para los sistemas seleccionados es muy estrecho, dificultando el análisis del rendimiento de los parámetros de no enlace derivados considerando el nuevo protocolo. Dentro del conjunto de ligandos evaluados los que presentan mayor divergencia en la energía libre de unión respecto a la data de referencia son los ligandos 3-clorofenol (L3), acetato de bencilo (L6), 4,5,6,7-tetrahidro-1H-indol (L7) y anilinoacetonitrilo (L8). Las diferencias pueden estar asociadas a que el nuevo conjunto de parámetros no enlazantes altere las conformaciones que puede optar el ligando, la conformación de los residuos o la presencia de moléculas de agua dentro del sitio de unión. Estos ligandos en general no sufren cambios conformacionales significativos por la rigidez del núcleo aromático y la restricción orientacional aplicada. En el caso de los residuos se analizó el RMSD para los átomos de carbono alfa de los residuos 84-162 (residuos más cercanos al sitio de unión) y se encontró que la variación está en el rango de 0-1 Å para los ligandos L3, L6, L7 y L8. Estas pequeñas variaciones en el RMSD indican que no hay cambios conformaciones en la cadena principal de los residuos cercanos al sitio de unión.

Para el ligando L3 al cambiar las cargas AM1-BCC del ligando por las cargas D-MBIS se observa una diferencia de 2.1 kcal/mol. Esta diferencia está asociada a los cambios en las magnitudes de la carga: las diferencias absolutas más relevantes son de $\sim 0.31 e$, $0.21 e$ y $0.19 e$ para el átomo de carbono

en posición *ipso*, el átomo de oxígeno y el átomo de carbono en posición *orto* respectivamente. Se analizó el porcentaje relativo de estados termodinámicos donde se evidencia la presencia de moléculas de agua en el sitio de unión a un radio de 5 Å y la distancia entre el hidrógeno del -OH del ligando L3 y el oxígeno del grupo carbonilo del residuo GLN102 en un rango de 1.5 a 2.0 Å (distancia donde se genera un puente de hidrógeno). No hay variación significativa en los porcentajes relativos asociados con los puentes de hidrógeno, mientras que el porcentaje de conformaciones con moléculas de agua en el sitio de unión pasa de 35.0 % a 3.7 % con los modelos de carga AM1-BCC y D-MBIS respectivamente. La variación en las moléculas de agua en el sitio de unión podría ser una de las razones que fundamentan la variación en la energía libre de unión. Al comparar las cargas atómicas D-MBIS con las cargas D-MBIS derivadas en conjunto con los parámetros de LJ no se encontraron diferencias significativas. Los parámetros de LJ del campo de fuerza GAFF y los derivados considerando la metodología planteada en la presente investigación no muestran cambios relevantes, salvo para el átomo de H del grupo funcional alcohol, que pasa de tener valores de $\sigma = 0.0000$ Å y $\epsilon = 0.0000$ kcal/mol en el campo de fuerza GAFF a valores de $\sigma = 1.2100$ Å y $\epsilon = 0.0025$ kcal/mol derivados de la partición MBIS. Al considerar el cambio de los modelos de carga D-MBIS a LJ-D-MBIS se observan cambios en los porcentajes relativos de conformaciones con moléculas de agua en el sitio de unión de 3.7 % a 0.0 % y de enlaces por puente de hidrógeno de 40.7 % a

4.5 %, estos cambios pueden estar asociados al aumento del radio de van-der-Waals que disminuye la probabilidad para generar puentes de hidrógeno en una distancia de 1.5 a 2.0 Å y de encontrar moléculas de agua en el sitio de unión.

El sistema con ligando L6 es el que más fluctúa al cambiar el conjunto de cargas atómicas y parámetros de Lennard-Jones y se ha reportado en la literatura dificultades en las simulaciones para predecir su afinidad de unión [121]. Las diferencias absolutas más marcadas entre las cargas atómicas AM1-BCC y D-MBIS son de $\sim 0.52 e$, $\sim 0.16 e$ y $\sim 0.27 e$ para los átomos de carbono y oxígeno del grupo carbonilo, y el átomo de carbono unido covalentemente al átomo de carbono del grupo carbonilo, respectivamente. Estas diferencias pueden ocasionar cambios en el ambiente molecular: el porcentaje relativo de conformaciones con moléculas de agua a un radio de 5 Å en el sitio de unión pasa de 23.8 % al 3.0 %. Mientras que el porcentaje relativo de puentes de hidrógeno generados entre el hidrógeno del grupo funcional $-\text{RCOONH}_2$ del residuo GLN102 y el oxígeno del grupo carbonilo del ligando L6 pasa de 0.0 % a 12.1 %. La comparación de las cargas atómicas y parámetros de LJ en las simulaciones con los modelos D-MBIS y LJ-D-MBIS no muestran diferencias significativas, ni en el porcentaje de conformaciones con moléculas de agua encontradas en el sitio de unión, sin embargo, el porcentaje relativo de los enlaces por puente de hidrógeno pasa de 12.1 % a 0.0 % al considerar los modelos D-MBIS y LJ-D-MBIS respectivamente.

Por otra, parte L7 también sufre cambios importantes en la energía libre de unión al variar el modelo de carga y parámetros de Lennard-Jones. Las diferencias absolutas más marcadas al comparar las cargas AM1-BCC y D-MBIS son de $\sim 0.24 e$ para el átomo de nitrógeno y $\sim 0.24 e$, $\sim 0.29 e$, $\sim 0.2 e$, ~ 0.25 para los átomos de carbono $e C_1, C_3, C_5, C_8$, respectivamente (ver Figura C.1 con numeración en el Anexo C). El análisis de la simulación muestra que el porcentaje relativo de enlace por puente de hidrógeno entre el oxígeno del grupo carbonilo y el hidrógeno de la amina aromática no presenta grandes cambios (pasa de 33.3 % a 34.8 % para las cargas AM1-BCC y D-MBIS respectivamente). Con respecto al porcentaje de conformaciones con moléculas de agua presentes en el sitio de unión a un radio de 5 Å tampoco se observan cambios significativos y pasa de 0.0 % a 4.3 %. La comparación de las cargas atómicas y parámetros de LJ en las simulaciones D-MBIS y LJ-D-MBIS no muestran diferencias notorias en las cargas atómicas D-MBIS. Los parámetros de LJ no sufren grandes cambios excepto para el átomo de hidrógeno enlazado al nitrógeno, que pasa de tener el valor de $\sigma = 0.600 \text{ \AA}$ en el campo de fuerza GAFF a $\sigma = 1.296 \text{ \AA}$ derivado de la partición MBIS. Estas diferencias ocasionan cambios en el porcentaje relativo de enlaces por puente de hidrógeno pasando de 34.8 % a 9.5 % y además se presentan cambios drásticos en el porcentaje de conformaciones con moléculas de agua en el sitio de unión pasando de un porcentaje relativo de 4.3 % a 47.6 %.

Para el ligando L8 al igual que los ligandos voluminosos L6 y L7, se ob-

servan cambios importantes en la energía libre de unión al variar el modelo de carga y parámetros de Lennard-Jones. Las diferencias absolutas más marcadas al comparar las cargas AM1-BCC y D-MBIS son de $\sim 0.20 e$, $\sim 0.23 e$ para los átomos de nitrógeno de la amina y el ciano, respectivamente y $\sim 0.36 e$, $\sim 0.25 e$ para los átomos de carbono C₇ y C₈, respectivamente. (ver Figura C.1 con numeración en el Anexo C). El análisis de la simulación muestra que el porcentaje relativo de enlace por puente de hidrógeno entre el oxígeno del grupo carbonilo y el hidrógeno de la amina secundaria no presenta grandes cambios (pasa de 11.5 % a 13.3 % para las cargas AM1-BCC y D-MBIS respectivamente). Con respecto al porcentaje de conformaciones con moléculas de agua presentes en el sitio de unión a un radio de 5 Å pasa de 0.0 % a 6.7 %. La comparación de las cargas atómicas y parámetros de LJ en las simulaciones D-MBIS y LJ-D-MBIS no muestran diferencias notorias en las cargas atómicas D-MBIS. Los parámetros de LJ no sufren grandes cambios excepto para el átomo de hidrógeno enlazado al nitrógeno, que pasa de tener el valor de $\sigma = 0.600 \text{ \AA}$ en el campo de fuerza GAFF a $\sigma = 1.296 \text{ \AA}$ derivado de la partición MBIS. Estas diferencias ocasionan cambios en el porcentaje relativo de enlaces por puente de hidrógeno pasando de 13.3 % a 0.0 % y además se presentan ligeros cambios en la cantidad de conformaciones con moléculas de agua que se observan en el sitio de unión pasando de un porcentaje relativo de 6.7 % a 2.9 %.

Otro aspecto importante para analizar es la conformación inicial del sis-

tema proteico[122]. El protocolo estudiado considera la estructura APO para la proteína en conjunto con una posición del ligando lo más cercana a la estructura cristalina. En la Figura 3.12 se comparan las estructuras HOLO (estructura activa, unida al ligando) y APO para el sistema proteico T4 L99A/M102Q unida a los ligandos L3 y L6. Los cambios más relevantes se observan con el ligando más voluminoso L6 que pueden ocasionar mayores cambios conformaciones de la hélice F (residuos 108 a 114) resaltada en verde. Para el ligando L6 se realizaron simulaciones adicionales considerando la estructura inicial HOLO en el cálculo de energía libre de unión, con el protocolo AM1-BCC y D-MIS obteniendo resultados de -6.4 ± 0.4 kcal/mol y -4.2 ± 0.4 kcal/mol respectivamente. Estos resultados difieren en -3.3 kcal/mol y -3.4 kcal/mol respecto a las energías libres de unión obtenidos con la estructura APO en los modelos de carga AM1-BCC y D-MBIS respectivamente. La energía libre de unión obtenida con la estructura HOLO y modelo cargas D-MBIS es más cercana al valor experimental, indicando que tanto la estructura inicial como las cargas D-MBIS son relevantes al momento de describir adecuadamente las interacciones de unión en este tipo de sistemas. En futuras investigaciones podrían incluirse las estructuras HOLO en los demás sistemas para analizar el efecto que tiene la estructura inicial combinada con las cargas atómicas D-MBIS.

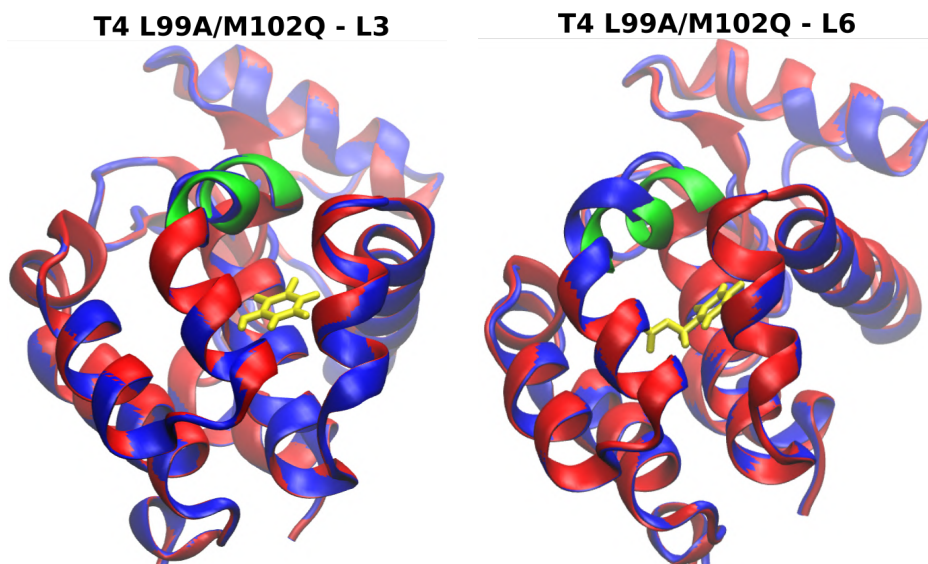


Figura 3.12: Comparación de estructuras APO (rojo) y HOLO (azul) para los ligandos L3 y L6. Se muestra la hélice F para la estructura APO (verde), en el caso del ligando más voluminoso L6 se aprecia mayores diferencias en la conformación de la hélice F. Los ligandos L3 y L6 se muestran en amarillo.

Finalmente es importante destacar que tanto las cargas D-MBIS como los parámetros de Lennard-Jones fueron derivados solo en la fase acuosa. Esta decisión fue tomada en base a los resultados obtenidos para el CB7 unido a huéspedes con núcleo de adamantano, donde los descriptores estadísticos mejoraron respecto al sistema de referencia y la derivación del nuevo conjunto de cargas atómicas D-MBIS y parámetros de LJ empleaban menos recursos computacionales que su contraparte LJ-q[D-MBIS]. Las cargas de los huéspedes en el CB7 y en solución acuosa no deberían variar significativamente ya que en el sitio de unión los grupos funcionales polares prefieren estar

orientados hacia el solvente. El ambiente molecular en el sistema proteico es más hidrofóbico y los grupos funcionales polares en el sitio de unión tendrán menos moléculas de agua al rededor del sitio de unión, en comparación con la fase acuosa donde estarán permanente interactuando con el solvente. Bajo esta panorámica las cargas atómicas D-MBIS y parámetros de LJ D-MBIS se esperaría que presentaran mayor diferencia al comparar los ambientes moleculares de unión y no unión. Por tanto, antes de concluir al respecto de las cargas atómicas D-MBIS y la derivación de los parámetros de LJ sería necesario hacer una derivación de las cargas atómicas D-MBIS y parámetros de LJ en ambos ambientes moleculares.



4. Conclusiones

En este proyecto se propone un nuevo protocolo de simulación que combina cálculos QM/MM y simulaciones de dinámica molecular para derivar parámetros no enlazantes a partir de la partición de la densidad electrónica molecular con el método MBIS junto con el costo energético promedio de polarizar la densidad electrónica en el ambiente molecular de unión y no unión. Posteriormente, se calcula la energía libre de unión empleando los nuevos parámetros del campo de fuerza, la estrategia de doble aniquilación (uso de estados *alquímicos* para apagar las interacciones electrostáticas y de van-der-Waals del huésped), la restricción tipo armónica, el muestreador SAMS y parámetros óptimos para obtener una energía libre de unión convergida con menos recursos computacionales en el software YANK. El poder predictivo de la energía libre de unión empleando los diferentes conjuntos de parámetros no enlazantes se comparó por medio de los descriptores estadísticos RMSE, MAE, R y τ obtenidos empleando parámetros no enlazantes de referencia (cargas atómicas RESP/AM1-BCC para los distintos sistemas anfitrión-huésped respectivamente y parámetros de Lennard-Jones del campo de fuerza GAFF) y los valores experimentales reportados.

En los sistemas OA/TEMOA se consideró la derivación de las cargas ató-

micas D-MBIS. En primera instancia las cargas atómicas D-MBIS derivadas para el huésped y la inclusión del costo de polarización mejoran la predicción de las afinidades de unión para los 6 huéspedes unidos al anfitrión TEMOA en comparación con el conjunto de cargas AM1-BCC. En contra parte, en el sistema OA la inclusión de las cargas D-MBIS y el costo energético de polarización para los 6 huéspedes no presentan una mejora significativa en los descriptores estadísticos empleados. Se realizaron cálculos de afinidad de unión adicionales con la derivación de las cargas atómicas D-MBIS para el OA y cada huésped junto con la energía de polarización promedio, obteniéndose una mejor concordancia con los valores experimentales. Para racionalizar las diferencias en la energía libre de unión, se analizaron los cambios conformacionales adoptados por el huésped durante la simulación empleando distintos modelos de carga atómica. Los huéspedes no varían significativamente las orientaciones que pueden adoptar y por tanto los cambios en la energía libre de unión están directamente asociados con la modificación de las interacciones con el ambiente molecular por las cargas atómicas D-MBIS. También se analizó el cambio conformacional del huésped al variar de anfitrión. En este caso se encontraron diferencias significativas, explicadas por la presencia de grupos metilo en el portal superior del TEMOA que reduce los grados de libertad del huésped y explica por ejemplo la diferencia en la energía libre de unión para el voluminoso huésped 4-bromoadamantano-1-carboxilato (G4). Finalmente se comparó la energía de polarización promedio en diferen-

tes ambientes moleculares. Los huéspedes con presencia del ión carboxilato como grupo funcional principal 5-hexinoato (G1), 4-cianobenzoato (G2), G4 y 3-nitrobenzoato (G6) presentan valores más elevados en la energía promedio de polarización tanto en la fase de unión como no unión. Los huéspedes con catión de amonio cuaternario como grupo funcional principal hexil(trimetil)amonio (G3) y trimetil(2-feniletil)amonio (G5) tienen menores energías promedios de polarización en ambos ambientes moleculares. Los huéspedes G1, G2, G4 y G6 presentan energías de polarización diferentes al comparar los ambientes moleculares solvente, OA, TEMOA y por tanto la corrección por polarización contribuye considerablemente a la afinidad de unión. Para los huéspedes catiónicos G3 y G5 las energías de polarización en los ambientes moleculares son similares y contribuyen a la energía libre de unión solo en un rango de 1-2 kcal/mol. **De los anteriores resultados se puede concluir que para los sistemas OA/TEMOA unidos a huéspedes del SAMPL5 las cargas atómicas D-MBIS mejoran considerablemente la predicción de las energías libres de unión en comparación con las cargas atómicas RESP/AM1-BCC para el anfitrión y huésped respectivamente.**

Para el anfitrión CB7 unido a huéspedes con núcleo aromático (BG1-6) y adamantano (AG1-8), se evaluó la energía libre de unión considerando las cargas atómicas D-MBIS para el CB7 y los diferentes huéspedes. Para los sistemas CB7 unidos a los huéspedes BG1-6 se obtuvo una mejor predicción en la energía libre de unión, salvo para el huésped BG4 que presenta la mayor

variación respecto al valor experimental. BG4 posee dos grupos amonio que pueden cambiar su estado de protonación de acuerdo a su nivel de exposición del solvente. La metodología empleada no considera cambios de protonación lo que podría ser una razón por la cual se observa esta desviación. La mayoría de los huéspedes con núcleo aromático son de carácter catiónico y no presentan otros grupos funcionales polarizables, por tanto, la corrección por polarización a la afinidad de unión es $\simeq 1$ kcal/mol para los huéspedes BG2-6. Para el huésped 2-(3,4-dihidroxifenil)etilamonio (BG1) la corrección por polarización es mayor debido a la presencia de dos grupos funcionales -OH que le confieren polarizabilidad.

Para los sistemas CB7 unidos a los huéspedes AG1-8 el uso de las cargas atómicas D-MBIS presentan peores descriptores estadísticos. Para entender el bajo rendimiento se realizó un análisis de las interacciones con la metodología SAPT0 para los sistemas representativos (4-metilfenil)amonio (BG5) y 1-adamantilamonio (AG7). Este análisis arrojó que las interacciones electrostáticas predominan en los huéspedes con núcleo aromático, mientras que los huéspedes con núcleo de adamantano presentan mayor contribución de interacciones repulsivas y de dispersión. Estas interacciones están descritas en los campos de fuerzas mediante el potencial de Lennard-Jones. En el marco de este trabajo se propone un protocolo similar a la derivación de las cargas atómicas D-MBIS para derivar nuevos parámetros de Lennard-Jones para los sistemas CB7 unidos a los huéspedes AG1-8. Las energías libres de unión

incluyendo los nuevos parámetros no covalentes MBIS (cargas y parámetros de Lennard-Jones) mejoran la predicción de la energía libre de unión y disminuye el RMSE aproximadamente en 2 kcal/mol. Las correcciones por polarización para los huéspedes con núcleo AG1-8 son \simeq 1 kcal/mol, excepto para 1-adamantil(2-aminoetil)amonio (AG3) donde es 2.7 kcal/mol. Al igual que en los sistemas OA/TEMOA se realizó un análisis conformacional y se encontró que no hay diferencias apreciables en el cambio de la orientación del huésped al considerar los diferentes modelos de carga atómica para los sistemas CB7 unido a los huéspedes BG1-6 y AG1-8. Para los huéspedes BG2 y AG5 con valores de pK_a cercanos al pH experimental, diferentes estados de protonación podrían contribuir a la afinidad de unión al CB7, que no se consideran en este trabajo. La coexistencia de múltiples estados de protonación del huésped y sus cambios al unirse al CB7 podría explicar parte de las desviaciones observadas para la energía libre de unión respecto al valor experimental. **Se puede concluir que para los sistemas CB7 los parámetros MBIS derivados de la densidad electrónica polarizada proporcionan cargas atómicas y parámetros de LJ que mejoran las predicciones de la energía libre de unión.**

Por último, se realizó el cálculo de parámetros no enlazantes para ligandos con núcleo aromático (L1-8) unidos a la lisozima T4 L99A/M102Q. En este caso en particular se ocuparon los parámetros no enlazantes derivados en solvente para la fase de unión y no unión. **Los resultados empleando el nuevo**

conjunto de parámetros no enlazantes no evidencian mejoras en los descriptores estadísticos. Se encontró que los huéspedes con mayor variación en la energía libre de unión respecto al sistema de referencia son los ligandos 3-clorofenol (L3), acetato de bencilo (L6), 4,5,6,7-tetrahidro-1H-indol (L7) y anilinoacetnitrilo (L8). Estas variaciones están asociadas a las diferencias en los parámetros no enlazantes que alteran las interacciones del ligando en sus diferentes ambientes moleculares como por ejemplo la cantidad de moléculas de agua que pueden ingresar al sitio de unión y los enlaces de hidrógeno generados entre el residuo GLN102 que provee carácter polar al sitio de unión y los ligandos. Además, es importante acotar que se consideraron estructuras APO para el cálculo de las afinidades de unión que pueden variar respecto al sistema HOLO que considera la estructura de la proteína unida a un ligando en particular. Se compararon las estructuras iniciales APO y HOLO para los ligandos L3, L6, L7 y L8 y se encontraron diferencias considerables en las conformaciones de la hélice F (residuos 108 a 114), especialmente para los sistemas más voluminosos. Para el ligando L6, que presenta mayor diferencia entre las conformaciones APO y HOLO se realizaron simulaciones adicionales de la energía libre de unión empleando la estructura inicial HOLO. Los resultados muestran diferencias notorias en las energías libres de unión empleando los diferentes modelos de carga atómica y el empleo de las cargas atómicas D-MBIS arroja como resultado una energía libre de unión más cercana al valor experimental comparado con el sistema de referencia.

Estos resultados motivan a que se realicen investigaciones futuras donde se inicien los cálculos computacionales desde la estructura HOLO. Finalmente, otro aspecto importante a considerar es la derivación de los parámetros no enlazantes en los ambientes moleculares de unión y no unión, debido a que los residuos alrededor del sitio de unión en su gran mayoría son de carácter apolar y por tanto se espera que los parámetros difieran de los derivados en la fase de solvente.



Bibliografía

- [1] Jian Yin, Niel M. Henriksen, David R. Slochower, Michael R. Shirts, Michael W. Chiu, David L. Mobley, and Michael K. Gilson. Overview of the SAMPL5 host–guest challenge: Are we doing better? *J. Comput. Aided Mol. Des.*, 31(1):1–19, 2017.
- [2] Antonia S. J. S. Mey, Bryce K. Allen, Hannah E. Bruce McDonald, John D. Chodera, David F. Hahn, Maxmillian Kuhn, Julien Michel, David L. Mobley, Levi N. Naden, Samarjeet Prasad, Andrea Rizzi, Jenke Scheen, Michael R. Shirts, Gary Tresadern, and Huafeng Xu. Best practices for alchemical free energy calculations [article v1.0]. *Living J. Comp. Mol. Sci.*, 2(1):18378, 2020.
- [3] Juyong Lee, Florentina Tofoleanu, Frank C. Pickard, Gerhard König, Jing Huang, Ana Damjanović, Minkyung Baek, Chaok Seok, and Bernard R. Brooks. Absolute binding free energy calculations of CBClip host–guest systems in the SAMPL5 blind challenge. *J. Comput. Aided Mol. Des.*, 31(1):71–85, 2017.
- [4] David L. Mobley and David Slochower. Mobleylab/benchmarksets: Version 1.2. *Zenodo.*, 2017. DOI:10.5281/zenodo.839047.
- [5] David L Mobley and Michael K Gilson. Predicting Binding Free Energies: Frontiers and Benchmarks. *Annu. Rev. Biophys.*, 46(1):531–558, 2017.
- [6] Andrea Rizzi, Steven Murkli, John N. McNeill, Wei Yao, Matthew Sullivan, Michael K. Gilson, Michael W. Chiu, Lyle Isaacs, Bruce C. Gibb, David L. Mobley, and John D. Chodera. Overview of the sampl6 host–guest binding affinity prediction challenge. *J. Comput. Aided Mol. Des.*, 32(10):937–963, 2018.
- [7] Martin Amezcua, Léa El Khoury, and David L. Mobley. SAMPL7 Host–Guest Challenge Overview: assessing the reliability of polariza-

- ble and non-polarizable methods for binding free energy calculations. *J. Comput. Aided Mol. Des.*, 35(1):1–35, 2021.
- [8] Hari S. Muddana, C. Daniel Varnado, Christopher W. Bielawski, Adam R. Urbach, Lyle Isaacs, Matthew T. Geballe, and Michael K. Gilson. Blind prediction of host–guest binding affinities: a new SAMPL3 challenge. *J. Comput. Aided Mol. Des.*, 26(5):475–487, 2012.
- [9] Esteban Vöhringer-Martinez, Fernanda Duarte, and Alejandro Toro-Labbé. How does pin1 catalyze the cis–trans prolyl peptide bond isomerization? a qm/mm and mean reaction force study. *J. Phys. Chem. B*, 116(43):12972–12979, 2012.
- [10] A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, and M. Karplus. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B*, 102(18):3586–3616, 1998.
- [11] Junmei Wang, Romain M. Wolf, James W. Caldwell, Peter A. Kollman, and David A. Case. Development and testing of a general amber force field. *J. Comput. Chem.*, 25(9):1157–1174, 2004.
- [12] Chris Oostenbrink, Alessandra Villa, Alan E. Mark, and Wilfred F. Van Gunsteren. A biomolecular force field based on the free enthalpy of hydration and solvation: The gromos force-field parameter sets 53a5 and 53a6. *J. Comput. Chem.*, 25(13):1656–1676, 2004.
- [13] William L. Jorgensen, David S. Maxwell, and Julian Tirado-Rives. Development and testing of the opls all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.*, 118(45):11225–11236, 1996.
- [14] Piotr Cieplak, François-Yves Dupradeau, Yong Duan, and Junmei Wang. Polarization effects in molecular mechanical force fields. *J. Phys.: Condens. Matter*, 21(33):333102, 2009.

- [15] Nohad Gresh, G. Andrés Cisneros, Thomas A. Darden, and Jean-Philip Piquemal. Anisotropic, polarizable molecular mechanics studies of inter and intramolecular interactions and ligand macromolecule complexes. a bottom up strategy. *J. Chem. Theory Comput.*, 3(6):1960–1986, 2007.
- [16] Kim Palmo, Noemi G. Mirkin, and Samuel Krimm. Spectroscopically determined force fields for macromolecules. 2. saturated hydrocarbon chains. *J. Phys. Chem. A*, 102(32):6448–6456, 1998.
- [17] Asbjørn Holt, Jonas Boström, Gunnar Karlström, and Roland Lindh. A nemo potential that includes the dipole-quadrupole and quadrupole-quadrupole polarizability. *J. Comput. Chem.*, 31:1583–91, 2010.
- [18] Ken Shoemake. Iii.6 - uniform random rotations. In DAVID KIRK, editor, *Graphics Gems III (IBM Version)*, pages 124–132. Morgan Kaufmann, San Francisco, 1992.
- [19] Matthew T. Geballe, A. Geoffrey Skillman, Anthony Nicholls, J. Peter Guthrie, and Peter J. Taylor. The SAMPL2 blind prediction challenge: introduction and overview. *J. Comput. Aided Mol. Des.*, 24(4):259–279, 2010.
- [20] Hari S. Muddana, Andrew T. Fenley, David L. Mobley, and Michael K. Gilson. The SAMPL4 host–guest blind prediction challenge: an overview. *J. Comput. Aided Mol. Des.*, 28(4):305–317, 2014.
- [21] David L. Mobley, Martin Amezcuca, Jay Ponder, Yuriy Khalak, Yigitkan Eken, Nuno Almeida, Lyle Isaacs, Bruce Gibb, Katy Kellett, and Dylan Serrilon. The SAMPL7 host-guest challenge virtual workshop. *Zenodo.*, 2020. DOI:10.5281/zenodo.3674155.
- [22] Mohamad Mohebifar, Erin R. Johnson, and Christopher N. Rowley. Evaluating Force-Field London Dispersion Coefficients Using the Exchange-Hole Dipole Moment Model. *J. Chem. Theory Comput.*, 13(12):6146–6157, 2017.
- [23] Jian Yin, Andrew T. Fenley, Niel M. Henriksen, and Michael K. Gilson. Toward Improved Force-Field Accuracy through Sensitivity

- Analysis of Host-Guest Binding Thermodynamics. *J. Phys. Chem. B*, 119(32):10145–10155, 2015.
- [24] Michael M. Pierce, C.S. Raman, and Barry T. Nall. Isothermal titration calorimetry of protein–protein interactions. *Methods*, 19(2):213 – 221, 1999.
- [25] A. Kukol. *Molecular Modeling of Proteins*. Methods in Molecular Biology. Springer New York, 2017.
- [26] Charles H Bennett. Efficient estimation of free energy differences from monte carlo data. *J. Comput. Phys.*, 22(2):245 – 268, 1976.
- [27] C. Chipot and A. Pohorille. *Free Energy Calculations: Theory and Applications in Chemistry and Biology*. Springer Series in Chemical Physics. Springer Berlin Heidelberg, 2007.
- [28] Pavel V. Klimovich, Michael R. Shirts, and David L. Mobley. Guidelines for the analysis of free energy calculations. *J. Comput. Aided Mol. Des.*, 29(5):397–411, 2015.
- [29] Andrew Pohorille, Christopher Jarzynski, and Christophe Chipot. Good practices in free-energy calculations. *J. Phys. Chem. B*, 114(32):10235–10253, 2010.
- [30] Shuai Liu, Yujie Wu, Teng Lin, Robert Abel, Jonathan P. Redmann, Christopher M. Summa, Vivian R. Jaber, Nathan M. Lim, and David L. Mobley. Lead optimization mapper: automating free energy calculations for lead optimization. *J. Comput. Aided Mol. Des.*, 27(9):755–770, 2013.
- [31] Raiker Witter, Margit Möllhoff, Frank-Thomas Koch, and Ulrich Sternberg. Fast atomic charge calculation for implementation into a polarizable force field and application to an ion channel protein. *J. Chem.*, 2015:14, 2015.
- [32] Sereina Riniker. Fixed-charge atomistic force fields for molecular dynamics simulations in the condensed phase: An overview. *J. Chem. Inf. Model.*, 58(3):565–578, 2018.

- [33] Christopher I. Bayly, Piotr Cieplak, Wendy Cornell, and Peter A. Kollman. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the resp model. *J. Phys. Chem.*, 97(40):10269–10280, 1993.
- [34] F. L. Hirshfeld. Bonded-atom fragments for describing molecular charge densities. *Theor. Chem. Acc.*, 44(2):129–138, 1977.
- [35] Patrick Bultinck, Christian Van Alsenoy, Paul W. Ayers, and Ramon Carbó-Dorca. Critical analysis and extension of the hirshfeld atoms in molecules. *J. Chem. Phys.*, 126(14):144111, 2007.
- [36] Toon Verstraelen, Steven Vandenbrande, Farnaz Heidar-Zadeh, Louis Vanduyfhuys, Veronique Van Speybroeck, Michel Waroquier, and Paul W. Ayers. Minimal basis iterative stockholder: Atoms in molecules for force-field development. *J. Chem. Theory Comput.*, 12(8):3894–3912, 2016.
- [37] T. Verstraelen, P. W. Ayers, V. Van Speybroeck, and M. Waroquier. Hirshfeld-e partitioning: Atom charges with an improved trade-off between robustness and accurate electrostatics. *J. Chem. Theory Comput.*, 9(5):2221–2225, 2013.
- [38] A. Lara, M. Riquelme, and E. Vöhringer-Martinez. Partition coefficients of methylated dna bases obtained from free energy calculations with molecular electron density derived atomic charges. *J. Comput. Chem.*, 39(22):1728–1737, 2018.
- [39] David Adrian Saez and Esteban Vöhringer-Martinez. A consistent s-adenosylmethionine force field improved by dynamic hirshfeld-i atomic charges for biomolecular simulation. *J. Comput. Aided Mol. Des.*, 29(10):951–961, 2015.
- [40] Esteban Vöhringer-Martinez, Toon Verstraelen, and Paul W. Ayers. The influence of ser-154, cys-113, and the phosphorylated threonine residue on the catalytic reaction mechanism of pin1. *J. Phys. Chem. B*, 118(33):9871–9880, 2014.

- [41] Maximiliano Riquelme, Alejandro Lara, David L. Mobley, Toon Verstraelen, Adelio R. Matamala, and Esteban Vöhringer-Martinez. Hydration free energies in the freesolv database calculated with polarized iterative hirshfeld charges. *J. Chem. Inf. Model.*, 58(9):1779–1797, 2018.
- [42] Daniel J. Cole, Jonah Z. Vilseck, Julian Tirado-Rives, Mike C. Payne, and William L. Jorgensen. Biomolecular Force Field Parameterization via Atoms-in-Molecule Electron Density Partitioning. *J. Chem. Theory Comput.*, 12(5):2312–2323, 2016.
- [43] Sophie M. Kantonen, Hari S. Muddana, Michael Schauerl, Niel M. Henriksen, Lee-Ping Wang, and Michael K. Gilson. Data-Driven Mapping of Gas-Phase Quantum Calculations to General Force Field Lennard-Jones Parameters. *J. Chem. Theory Comput.*, 16(2):1115–1127, 2020.
- [44] Koen M. Visscher and Daan P. Geerke. Deriving Force-Field Parameters from First Principles Using a Polarizable and Higher Order Dispersion Model. *J. Chem. Theory Comput.*, 15(3):1875–1883, 2019.
- [45] Khaleel I. Assaf and Werner M. Nau. Cucurbiturils: from synthesis to high-affinity binding and catalysis. *Chem. Soc. Rev.*, 44:394–418, 2015.
- [46] Kenneth A. Connors. The stability of cyclodextrin complexes in solution. *Chem. Rev.*, 97(5):1325–1358, 1997.
- [47] Corinne L. D. Gibb and Bruce C. Gibb. Binding of cyclic carboxylates to octa-acid deep-cavity cavitand. *J. Comput. Aided Mol. Des.*, 28(4):319–325, 2014.
- [48] Matthew R. Sullivan, Punidha Sokkalingam, Thong Nguyen, James P. Donahue, and Bruce C. Gibb. Binding of carboxylate and trimethylammonium salts to octa-acid and TEMOA deep-cavity cavitands. *J. Comput. Aided Mol. Des.*, 31(1):21–28, 2017.
- [49] Liping Cao, Marina Šekutor, Peter Y. Zavalij, Kata Mlinarić-Majerski, Robert Glaser, and Lyle Isaacs. Cucurbit[7]uril guest pair with an at-

- tomolar dissociation constant. *Angew. Chem. Int. Ed.*, 53(4):988–993, 2014.
- [50] Simin Liu, Christian Ruspic, Pritam Mukhopadhyay, Sriparna Chakrabarti, Sriparna Chakrabarti, Peter Y. Zavalij, and Lyle Isaacs. The Cucurbit[n]uril Family: Prime Components for Self-Sorting Systems. *J. Am. Chem. Soc.*, 127(45):15959–15967, 2005.
- [51] William L. Mock and Neng Yang Shih. Organic ligand-receptor interactions between cucurbituril and alkylammonium ions. *J. Am. Chem. Soc.*, 110(23):4706–4710, 1988.
- [52] William L. Mock and Neng Yang Shih. Structure and selectivity in host-guest complexes of cucurbituril. *J. Org. Chem.*, 51(23):4440–4446, 1986.
- [53] Dongmok Whang, Jungseok Heo, and Joo Hyun Park. A molecular bowl with metal ion as bottom: Reversible inclusion of organic molecules in cesium ion complexed cucurbituril. *Angew. Chem*, 37(1):78–80, 1998.
- [54] Sarvin Moghaddam, Yoshihisa Inoue, and Michael K Gilson. Host-guest complexes with protein-ligand-like affinities: computational analysis and design. *J. Am. Chem. Soc.*, 131(11):4012–4021, 2009.
- [55] Mikhail V. Rekharsky, Tadashi Mori, Cheng Yang, Young Ho Ko, N. Selvapalam, Hyunuk Kim, David Sobransingh, Angel E. Kaifer, Simin Liu, Lyle Isaacs, Wei Chen, Sarvin Moghaddam, Michael K. Gilson, Kimoon Kim, and Yoshihisa Inoue. A synthetic host-guest system achieves avidin-biotin affinity by overcoming enthalpy–entropy compensation. *Proc. Natl. Acad. Sci. U. S. A.*, 104(52):20737–20, 2007.
- [56] Helen M. Berman, John Westbrook, Zukang Feng, Gary Gilliland, T. N. Bhat, Helge Weissig, Ilya N. Shindyalov, and Philip E. Bourne. The Protein Data Bank. *Nucleic Acids Res.*, 28(1):235–242, 2000.
- [57] Kaushik Raha and Kenneth M. Merz. Chapter 9 calculating binding free energy in protein–ligand interaction. volume 1 of *Annual Reports in Computational Chemistry*, pages 113–130. Elsevier, 2005.

- [58] Conor D. Parks, Zied Gaieb, Michael Chiu, Huanwang Yang, Chenghua Shao, W. Patrick Walters, Johanna M. Jansen, Georgia McGaughey, Richard A. Lewis, Scott D. Bembenek, Michael K. Ameriks, Tara Mirzadegan, Stephen K. Burley, Rommie E. Amaro, and Michael K. Gilson. D3R grand challenge 4: blind prediction of protein–ligand poses, affinity rankings, and relative binding free energies. *J. Comput. Aided Mol. Des.*, 2020(2):99–119, 2020.
- [59] Andrew Morton, Walter A. Baase, and Brian W. Matthews. Energetic origins of specificity of ligand binding in an interior nonpolar cavity of t4 lysozyme. *Biochem*, 34(27):8564–8575, 1995.
- [60] Alan P. Graves, Ruth Brenk, and Brian K. Shoichet. Decoys for docking. *J. Med. Chem.*, 48(11):3714–3728, 2005.
- [61] István. Mayer. *Simple Theorems, Proofs, and Derivations in Quantum Chemistry*. Mathematical and Computational Chemistry. Springer, Boston, MA, 2003.
- [62] Brian T. Sutcliffe and R. Guy Woolley. On the quantum theory of molecules. *J. Chem. Phys.*, 137(22):22A544, 2012.
- [63] Robert G. Parr. *Density Functional Theory of Atoms and Molecules*. Mathematical and Computational Chemistry. Springer, Dordrecht, 1980.
- [64] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140:A1133–A1138, 1965.
- [65] P. A. M. Dirac. Note on exchange phenomena in the thomas atom. *Math. Proc. Camb. Philos. Soc.*, 26(3):376–385, 1930.
- [66] J. C. Slater. A simplification of the hartree-fock method. *Phys. Rev.*, 81:385–390, 1951.
- [67] S. H. Vosko and L. Wilk. Influence of an improved local-spin-density correlation-energy functional on the cohesive energy of alkali metals. *Phys. Rev. B*, 22:3812–3815, 1980.

- [68] A. D. Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A*, 38:3098–3100, 1988.
- [69] Chengteh Lee, Weitao Yang, and Robert G. Parr. Development of the colle-salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B*, 37:785–789, 1988.
- [70] Axel D. Becke. Density-functional thermochemistry. iii. the role of exact exchange. *J. Chem. Phys.*, 98(7):5648–5652, 1993.
- [71] Yelena Arnautova, Anna Jagielska, and Harold Scheraga. A new force field (ecepp-05) for peptides, proteins, and organic molecules. *J. Phys. Chem. B*, 110:5025–44, 2006.
- [72] Brent H. Besler, Kenneth M. Merz Jr., and Peter A. Kollman. Atomic charges derived from semiempirical methods. *J. Comput. Chem.*, 11(4):431–439, 1990.
- [73] U. Chandra Singh and Peter A. Kollman. An approach to computing electrostatic charges for molecules. *J. Comput. Chem.*, 5(2):129–145, 1984.
- [74] K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, and A. D. MacKerell Jr. Charmm general force field: A force field for drug-like molecules compatible with the charmm all-atom additive biological force fields. *J. Comput. Chem.*, 31(4):671–690, 2010.
- [75] Wendy D. Cornell, Piotr Cieplak, Christopher I. Bayly, Ian R. Gould, Kenneth M. Merz, David M. Ferguson, David C. Spellmeyer, Thomas Fox, James W. Caldwell, and Peter A. Kollman. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.*, 117(19):5179–5197, 1995.
- [76] Johnny C. Wu, Gaurav Chatree, and Pengyu Ren. Automation of amoeba polarizable force field parameterization for small molecules. *Theor. Chem. Acc.*, 131(3):1138, 2012.

- [77] P K Weiner, R Langridge, J M Blaney, R Schaefer, and P A Kollman. Electrostatic potential molecular surfaces. *Proc. Natl. Acad. Sci. U. S. A.*, 79(12):3754–3758, 1982.
- [78] Araz Jakalian, David B. Jack, and Christopher I. Bayly. Fast, efficient generation of high-quality atomic charges. am1-bcc model: Ii. parameterization and validation. *J. Comput. Chem.*, 23(16):1623–1641, 2002.
- [79] Patrick Bultinck, Christian Van Alsenoy, Paul W. Ayers, and Ramon Carbó-Dorca. Critical analysis and extension of the hirshfeld atoms in molecules. *J. Chem. Phys.*, 126(14):144111, 2007.
- [80] Mohamad Mohebifar, Erin R. Johnson, and Christopher N. Rowley. Evaluating Force-Field London Dispersion Coefficients Using the Exchange-Hole Dipole Moment Model. *J. Chem. Theory Comput.*, 13(12):6146–6157, 2017.
- [81] Axel D Becke and Erin R Johnson. Exchange-hole dipole moment and the dispersion interaction: high-order dispersion coefficients. *J. Chem. Phys.*, 124(1):14104, 2006.
- [82] Alexandre Tkatchenko and Matthias Scheffler. Accurate Molecular Van Der Waals Interactions from Ground-State Electron Density and Free-Atom Reference Data. *Phys. Rev. Lett.*, 102(7), 2009.
- [83] X. Chu and A. Dalgarno. Linear response time-dependent density functional theory for van der Waals coefficients. *J. Chem. Phys.*, 121(9):4083–4088, 2004.
- [84] Dmitry V. Fedorov, Mainak Sadhukhan, Martin Stöhr, and Alexandre Tkatchenko. Quantum-Mechanical Relation between Atomic Dipole Polarizability and the van der Waals Radius. *J. Phys. Chem. Lett.*, 121(18), 2018.
- [85] Alexandre Tkatchenko, Dmitry V. Fedorov, and Matteo Gori. Fine-structure constant connects electronic polarizability and geometric van-der-waals radius of atoms. *J. Phys. Chem. Lett.*, 12(39):9488–9492, 2021.

- [86] John D. Chodera and Michael R. Shirts. Replica exchange and expanded ensemble simulations as gibbs sampling: Simple improvements for enhanced mixing. *J. Chem. Phys.*, 135(19):194110, 2011.
- [87] Zhiqiang Tan. Optimally Adjusted Mixture Sampling and Locally Weighted Histogram Analysis. *J. Comput. Graph. Stat.*, 26(1):54–65, 2017.
- [88] A. P. Lyubartsev, A. A. Martsinovski, S. V. Shevkunov, and P. N. Vorontsov-Velyaminov. New approach to monte carlo calculation of the free energy: Method of expanded ensembles. *J. Chem. Phys.*, 96(3):1776–1783, 1992.
- [89] Kai Wang, John D. Chodera, Yanzhi Yang, and Michael R. Shirts. Identifying ligand binding sites and poses using gpu-accelerated hamiltonian replica exchange molecular dynamics. *J. Comput. Aided Mol. Des.*, 27(12):989–1007, 2013.
- [90] Stefan Boresch, Franz Tettinger, Martin Leitgeb, and Martin Karplus. Absolute binding free energies: A quantitative approach for their calculation. *J. Phys. Chem. B*, 107(35):9535–9551, 2003.
- [91] Robert W. Zwanzig. High-temperature equation of state by a perturbation method. i. nonpolar gases. *J. Chem. Phys.*, 22(8):1420–1426, 1954.
- [92] Nandou Lu, Jhumpa Adhikari, and David A. Kofke. Variational formula for the free energy based on incomplete sampling in a molecular simulation. *Phys. Rev. E*, 68:026122, 2003.
- [93] Nandou Lu, David A. Kofke, and Thomas B. Woolf. Improving the efficiency and reliability of free energy perturbation calculations using overlap sampling methods. *J. Comput. Chem.*, 25(1):28–40, 2004.
- [94] John G. Kirkwood. Statistical mechanics of fluid mixtures. *J. Chem. Phys.*, 3(5):300–313, 1935.
- [95] Charles H Bennett. Efficient estimation of free energy differences from monte carlo data. *J. Chem. Phys.*, 22(2):245–268, 1976.

- [96] Michael R. Shirts and Vijay S. Pande. Comparison of efficiency and bias of free energies computed by exponential averaging, the bennett acceptance ratio, and thermodynamic integration. *J. Chem. Phys.*, 122(14):144107, 2005.
- [97] Michael R. Shirts and John D. Chodera. Statistically optimal analysis of samples from multiple equilibrium states. *J. Chem. Phys.*, 129(12):124105, 2008.
- [98] Jian Yin, Niel M. Henriksen, Hari S. Muddana, and Michael K. Gilson. Bind3P: Optimization of a Water Model Based on Host–Guest Binding Data. *J. Chem. Theory Comput.*, 14(7):3621–3632, 2018.
- [99] Florentina Tofoleanu, Juyong Lee, Frank C Pickard IV, Gerhard König, Jing Huang, Minkyung Baek, Chaok Seok, and Bernard R. Brooks. Absolute binding free energies for octa-acids and guests in SAMPL5. *J. Comput. Aided Mol. Des.*, 31(1):108–118, 2017.
- [100] Rajat Kumar Pal, Kamran Haider, Divya Kaur, William Flynn, Junchao Xia, Ronald M Levy, Tetiana Taran, Lauren Wickstrom, Tom Kurtzman, and Emilio Gallicchio. A combined treatment of hydration and dynamical effects for the modeling of host-guest binding thermodynamics: the SAMPL5 blinded challenge. *J. Comput. Aided Mol. Des.*, 31(1):29–44, 2017.
- [101] Bing Xie, Trung Hai Nguyen, and David D. L. Minh. Absolute Binding Free Energies between T4 Lysozyme and 141 Small Molecules: Calculations Based on Multiple Rigid Receptor Configurations. *J. Chem. Theory Comput.*, 13(6):2930–2944, 2017.
- [102] Samuel C. Gill, Nathan M. Lim, Patrick B. Grinaway, Ariën S. Rustenburg, Josh Fass, Gregory A. Ross, John D. Chodera, and David L. Mobley. Binding modes of ligands using enhanced sampling (blues): Rapid decorrelation of ligand binding modes via nonequilibrium candidate monte carlo. *J. Phys. Chem. B*, 122(21):5579–5598, 2018.
- [103] Kresten Lindorff-Larsen, Stefano Piana, Kim Palmo, Paul Maragakis, John L. Klepeis, Ron O. Dror, and David E. Shaw. Improved side-chain

- torsion potentials for the amber ff99sb protein force field. *Proteins: Struct., Funct., Bioinf.*, 78(8):1950–1958, 2010.
- [104] David L. Mobley, John D. Chodera, and Ken A. Dill. On the use of orientational restraints and symmetry corrections in alchemical free energy calculations. *J. Chem. Phys.*, 125(8):084902, 2006.
- [105] In Suk Joung and Thomas E. Cheatham. Determination of Alkali and Halide Monovalent Ion Parameters for Use in Explicitly Solvated Biomolecular Simulations. *J. Phys. Chem. B*, 112(30):9020–9041, 2008.
- [106] William C. Swope, Hans W. Horn, and Julia E. Rice. Accounting for Polarization Cost When Using Fixed Charge Force Fields. II. Method and Application for Computing Effect of Polarization Cost on Free Energy of Hydration. *J. Phys. Chem. B*, 114(26):8631–8645, 2010.
- [107] William C. Swope, Hans W. Horn, and Julia E. Rice. Accounting for Polarization Cost When Using Fixed Charge Force Fields. I. Method for Computing Energy. *J. Phys. Chem. B*, 114(26):8621–8630, 2010.
- [108] Maximiliano Riquelme, Alejandro Lara, David L. Mobley, Adelio R. Matamala, and Esteban Vöhringer-Martinez. Hydration Free Energies in the FreeSolv Database Calculated with Polarized Iterative Hirshfeld Charges. *J. Chem. Inf. Model*, 58(9):1779–1797, 2018.
- [109] Farnaz Heidar-Zadeh, Paul W. Ayers, Toon Verstraelen, Ivan Vinogradov, Esteban Vöhringer-Martinez, and Patrick Bultinck. Information-Theoretic Approaches to Atoms-in-Molecules: Hirshfeld Family of Partitioning Schemes. *J. Phys. Chem. A*, 122(17):4219–4245, 2018.
- [110] Peter Eastman, Jason Swails, John D. Chodera, Robert T. McGibbon, Yutong Zhao, Kyle A. Beauchamp, Lee-Ping Wang, Andrew C. Simonett, Matthew P. Harrigan, Chaya D. Stern, Rafal P. Wiewiora, Bernard R. Brooks, and Vijay S. Pande. OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLoS Comput. Biol.*, 13(7):e1005659, 2017.
- [111] Frank Neese. Software update: the orca program system, version 4.0. *WIREs. Comput. Mol. Sc*, 8(1):e1327, 2018.

- [112] Xiaoliang Pan, Edina Rosta, and Yihan Shao. Representation of the QM Subsystem for Long-Range Electrostatic Interaction in Non-Periodic Ab Initio QM/MM Calculations. *Molecules*, 23(10):2500, 2018.
- [113] Diptarka Hait, Yu Hsuan Liang, and Martin Head-Gordon. Too big, too small, or just right? A benchmark assessment of density functional theory for predicting the spatial extent of the electron density of small chemical systems. *J. Chem. Phys.*, 154(7):074109, 2021.
- [114] Diptarka Hait and Martin Head-Gordon. How accurate are static polarizability predictions from density functional theory? An assessment over 132 species at equilibrium geometry. *Phys. Chem. Chem. Phys.*, 20(30):19800–19810, 2018.
- [115] Diptarka Hait and Martin Head-Gordon. How Accurate Is Density Functional Theory at Predicting Dipole Moments? An Assessment Using a New Database of 200 Benchmark Values. *J. Chem. Theory Comput.*, 14(4):1969–1981, 2018.
- [116] Paulius Mikulskis, Daniela Cioloboc, Milica Andrejić, Sakshi Khare, Joakim Brorsson, Samuel Genheden, Ricardo A. Mata, Pär Söderhjelm, and Ulf Ryde. Free-energy perturbation and quantum mechanical study of SAMPL4 octa-acid host–guest binding energies. *J. Comput. Aided Mol. Des.*, 28(4):375–400, 2014.
- [117] Octav Caldararu, Martin A. Olsson, Majda Misini Ignjatović, Meiting Wang, and Ulf Ryde. Binding free energies in the SAMPL6 octa-acid host–guest challenge calculated with MM and QM methods. *J. Comput. Aided Mol. Des.*, 32(10):1027–1046, 2018.
- [118] Duván González, Luis Macaya, and Esteban Vöhringer-Martinez. Molecular environment-specific atomic charges improve binding affinity predictions of sampl5 host–guest systems. *J. Chem. Inf. Model.*, 61(9):4462–4474, 2021.
- [119] Evan T. Walters, Mohamad Mohebifar, Erin R. Johnson, and Christopher N. Rowley. Evaluating the London Dispersion Coefficients of

Protein Force Fields Using the Exchange-Hole Dipole Moment Model. *J. Phys. Chem. B*, 122(26):6690–6701, 2018.

- [120] James A. Maier, Carmenza Martinez, Koushik Kasavajhala, Lauren Wickstrom, Kevin E. Hauser, and Carlos Simmerling. ff14sb: Improving the accuracy of protein side chain and backbone parameters from ff99sb. *J. Chem. Theory Comput.*, 11(8):3696–3713, 2015.
- [121] Sarah E. Boyce, David L. Mobley, Gabriel J. Rocklin, Alan P. Graves, Ken A. Dill, and Brian K. Shoichet. Predicting Ligand Binding Affinity with Alchemical Free Energy Methods in a Polar Model Binding Site. *J. Mol. Biol.*, 394(4):747–763, 2009.
- [122] Y. Khalak, G. Tresadern, M. Aldeghi, H. M. Baumann, D. L. Mobley, B. L. de Groot, and V. Gapsys. Alchemical absolute protein–ligand binding free energies for drug design. *Chem. Sci.*, 12:13958–13971, 2021.



A. Anexo: Sistemas OA/TEMOA unidos a huéspedes G1-6

Tabla A.1: Valores de pK_a para todos los grupos titulables en los huéspedes G1-6. ^a valores de pK_a a 298 K de acuerdo con Haynes, William M.,ed (2015) CRC Handbook of Chemistry and Physics (95th ed.). ^b valores de pK_a a 298 K según el software Chemizalze. ^c No determinado.

Huésped	pK_a
G1	4.54 ^b
G2	3.55 ^a
G3	ND ^c
G4	3.77 ^b
G5	ND ^c
G6	3.46 ^a

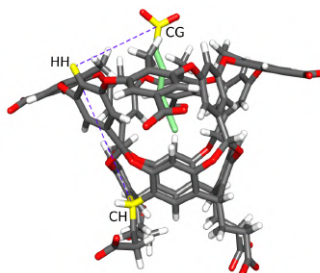


Figura A.1: Definición del ángulo CH-HH-CG utilizado en las siguientes figuras para analizar la conformación del huésped en el bolsillo de unión (donde CH es el átomo de carbono de la parte inferior del anfitrión, HH es un átomo de hidrógeno de la parte superior del anfitrión, y CG es el átomo de carbono del ión carboxilato del huésped). El ángulo medido CH-HH-CG representa la conformación relativa del grupo funcional principal en el bolsillo de unión. Átomos de oxígeno en rojo, átomos de carbono en gris (excepto los átomos de carbono del huésped en verde), átomos de hidrógeno en blanco y átomos de referencia CH-HH-CG en amarillo.

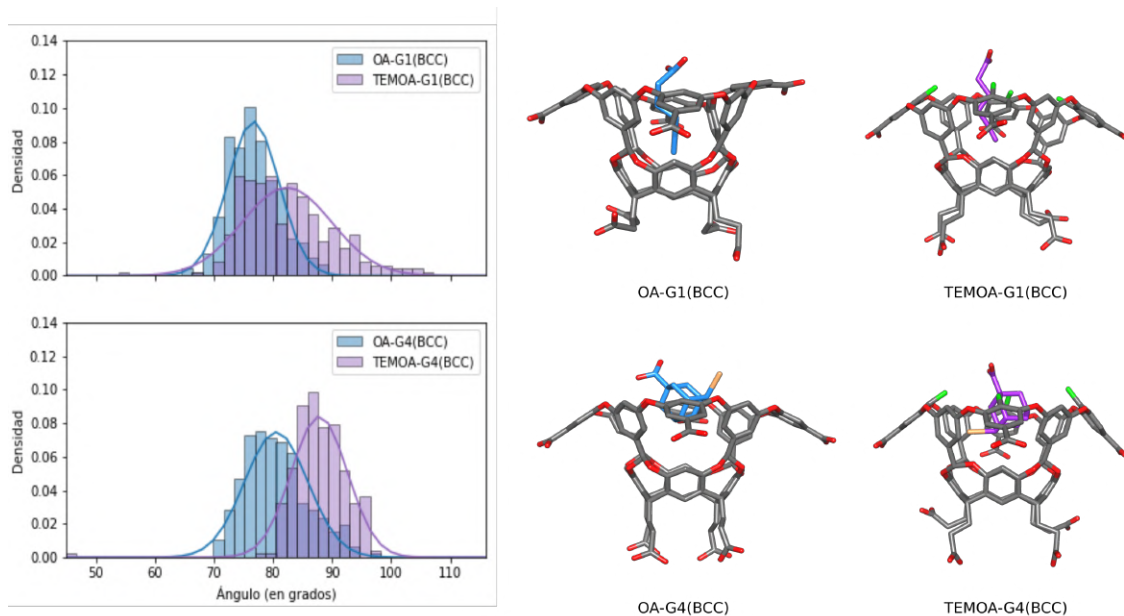


Figura A.2: Izquierda: Gráficos de distribución de ángulos en los sistemas OA-G1/TEMOA-G1 (arriba) y OA-G4/TEMOA-G4 (abajo) usando cargas atómicas RESP/AM1-BCC para el anfitrión-huésped respectivamente. Derecha: Conformaciones representativas para los sistemas OA-G1/TEMOA-G1 (arriba) y OA-G4/TEMOA-G4 (abajo) (los grupos metilo de TEMOA están resaltados en verde y los dos huéspedes coloreados según la distribución que se muestra a la izquierda. Átomos de oxígeno en rojo y átomos de carbono en gris.

Tabla A.2: Diferencia absoluta de la carga atómica promedio por átomo entre las cargas atómicas D-MBIS calculadas en anfitrión y en solución acuosa. Valores en $\times 10^{-3}e$

	OA	TEMOA	OA
	DMBIS(H)	DMBIS(H)	DMBIS
Huésped			
G1	1.2 \pm 0.6	1.0 \pm 0.2	1.0 \pm 0.2
G2	2.7 \pm 0.3	0.8 \pm 0.3	2.2 \pm 0.2
G3	0.11 \pm 0.01	0.14 \pm 0.01	0.11 \pm 0.01
G4	0.24 \pm 0.03	0.4 \pm 0.1	0.25 \pm 0.03
G5	0.18 \pm 0.04	0.21 \pm 0.05	0.24 \pm 0.07
G6	0.8 \pm 0.1	1.0 \pm 0.1	1.0 \pm 0.1

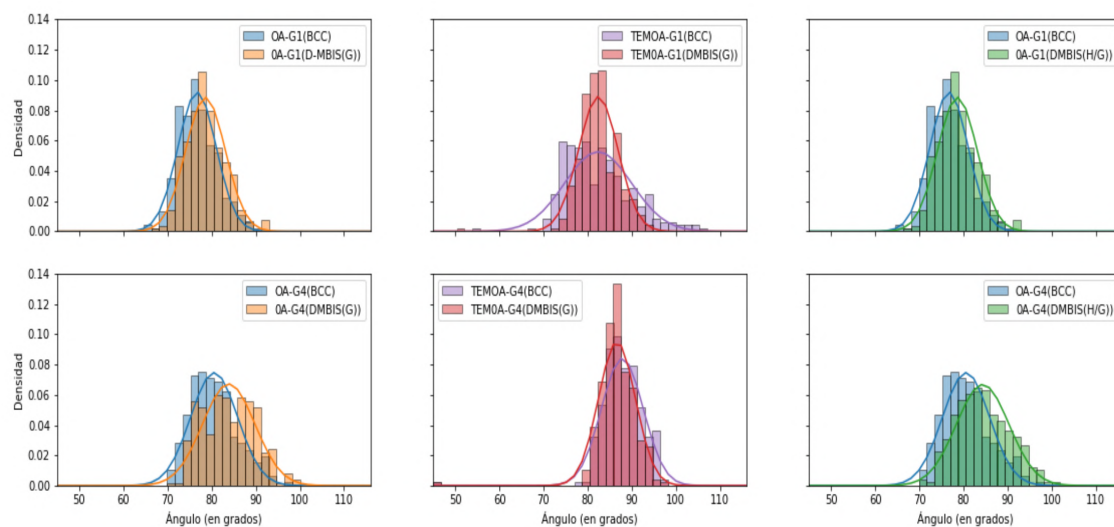


Figura A.3: Distribución de ángulos que definen la orientación del huésped en los anfitriones variando entre modelo de cargas atómicas RESP/AM1-BCC, RESP/D-MBIS(H) y D-MBIS(A)/D-MBIS(H) para el anfitrión y el huésped respectivamente. Los sistemas OA-G1 y TEMOA-G1 se muestran en la parte superior y los sistemas OA-G4 y TEMOA-G4 en la parte inferior.

Tabla A.3: Energía libre de unión obtenida usando las cargas atómicas D-MBIS del huésped derivadas en el ambiente acuoso (AC) para el estado unido y no unido y usando las cargas para el huésped derivadas cuando está unida al anfitrión (HC) (sin corrección de energía de polarización). Valores en kcal/mol

Sistema	ΔG_{AC}°	ΔG_{HC}°
OA		
OA-G2	-3.6 ± 0.3	-6.3 ± 0.3
OA-G3	-7.8 ± 0.2	-7.4 ± 0.2

B. Anexo: Sistemas CB7 unidos a huéspedes BG1-G6 y AG1-8

Tabla B.1: Valores de pK_a para todos los grupos titulables en los huéspedes BG1-6 y AG1-8, respectivamente. ^a valores de pK_a a 298 K de acuerdo con Haynes, William M., ed (2015) CRC Handbook of Chemistry and Physics (95th ed.). ^b valores de pK_a de acuerdo con Dean, J.A. Handbook of Organic Chemistry. New York, NY: McGraw-Hill Book Co; (1987). ^c valores de pK_a a 298 K de acuerdo al software Chemicalize. ^d valores de pK_a a 298 K de acuerdo con Adamo Fini *et al.* Acidity constants of sparingly water-soluble drugs from potentiometric determinations in aqueous dimethyl sulfoxide. J. Pharm. Sci., 76(1):48-52, 1987. ^e No determinado.

Huésped Núcleo aromático	pK_a	Huésped Núcleo de adamantano	pK_a
BG1	(1)8.90 (2)10.6 ^a	AG1	10.30 ^a
BG2	(1) < 2 (2)4.47 ^b	AG2	ND ^e
BG3	(1)3.15 (2)9.06 ^c	AG3	(1)7.27 (2)10.53 ^c
BG4	(1)3.29 (2)6.08 ^b	AG4	(1)8.92 (2)10.67 ^c
BG5	5.08 ^a	AG5	5.00 ^d
BG6	(1)8.96 (2)9.58 ^c	AG6	ND ^e
		AG7	10.1 ^a
		AG8	ND ^e

Tabla B.2: Energía libre de unión obtenida para el CB7 unido a los huéspedes BG2 y AG5. Simulaciones CB7-BG2 y CB7-BG2(N) corresponde al huésped en su forma protonada y neutra, respectivamente. CB7-AG5 y CB7-AG5(P) corresponde al huésped en su forma desprotonada y protonada, respectivamente. Valores en kcal/mol

Sistema	$\Delta G_{q[\text{AM1-BCC}]}^{\circ}$	$\Delta G_{q[\text{DMBIS}]}^{\circ}$
CB7-BG2	-5.5 ± 0.2	-6.8 ± 0.3
CB7-BG2(N)	-7.3 ± 0.1	-3.6 ± 0.3
CB7-AG5	-17.0 ± 0.7	-17.7 ± 0.6
CB7-AG5(P)	-21.4 ± 0.2	-20.7 ± 0.2

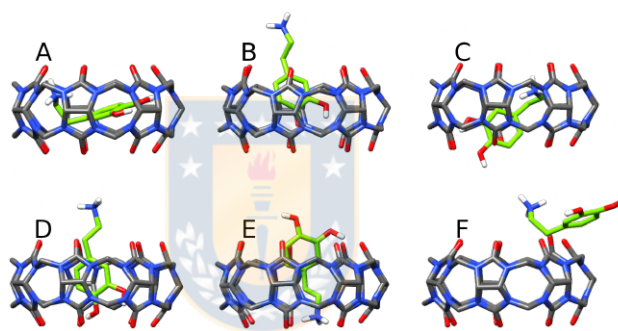


Figura B.1: Conformaciones representativas (A-F) adoptadas por el sistema CB7-BG1 durante la simulación de energía libre de unión.

Tabla B.3: Porcentaje relativo de la orientación del huésped en las ventanas de simulación con $\lambda = 1$ y asociado con las conformaciones (A-F) mostradas en la Figura B.1. Las simulaciones fueron realizadas usando las cargas atómicas RESP/AM1-BCC y D-MBIS/D-MBIS para el anfitrión y huésped respectivamente. Adicionalmente, se reportan los porcentajes de conformaciones con la presencia de agua en el sitio de unión.

CB7-BG1								
Cargas Atómicas	Ventanas totales con $\lambda = 1$	Orientación del huésped						Ventanas con moléculas de agua dentro del sitio de unión
		A	B	C	D	E	F	
RESP/AM1BCC	27	4	7	7	37	30	15	11
D-MBIS/D-MBIS	27	0	7	4	30	44	15	22

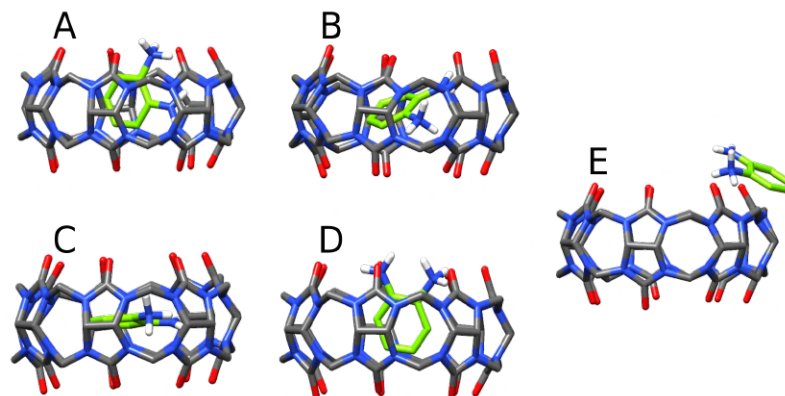


Figura B.2: Conformaciones representativas (A-E) adoptadas por el sistema CB7-BG2 durante la simulación de energía libre de unión.

Tabla B.4: Porcentaje relativo de la orientación del huésped en las ventanas de simulación con $\lambda = 1$ y asociados con las conformaciones (A-E) mostradas en la Figura B.2. Las simulaciones fueron realizadas usando las cargas RESP/AM1-BCC y D-MBIS/D-MBIS para el anfitrión y huésped respectivamente. Adicionalmente, se reportan los porcentajes de conformaciones con la presencia de agua en el sitio de unión.

CB7-BG2							
Cargas atómicas	Ventanas totales con $\lambda = 1$	Orientación del huésped					Ventanas con moléculas de agua dentro del sitio de unión
		A	B	C	D	E	
RESP/AM1BCC	39	54	3	0	33	10	13
D-MBIS/D-MBIS	38	13	3	29	37	18	19

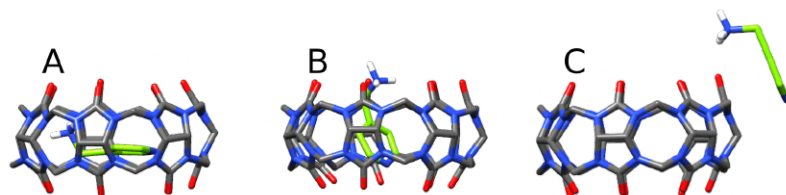


Figura B.3: Conformaciones representativas (A-C) adoptadas por el sistema CB7-BG3 durante la simulación de energía libre de unión.

Tabla B.5: Porcentaje relativo de la orientación del huésped en las ventanas de simulación con $\lambda = 1$ y asociado con las conformaciones (A-C) mostradas en la Figura B.3. Las simulaciones fueron realizadas usando las cargas atómicas RESP/AM1-BCC y D-MBIS/D-MBIS para el anfitrión y huésped respectivamente. Adicionalmente, se reportan los porcentajes de conformaciones con la presencia de agua en el sitio de unión.

CB7-BG3					
Cargas atómicas	Ventanas totales con $\lambda = 1$	Orientación del huésped			Ventanas con moléculas de agua dentro del sitio de unión
		A	B	C	
RESP/AM1BCC	28	21	79	0	0
D-MBIS/D-MBIS	30	27	60	13	13

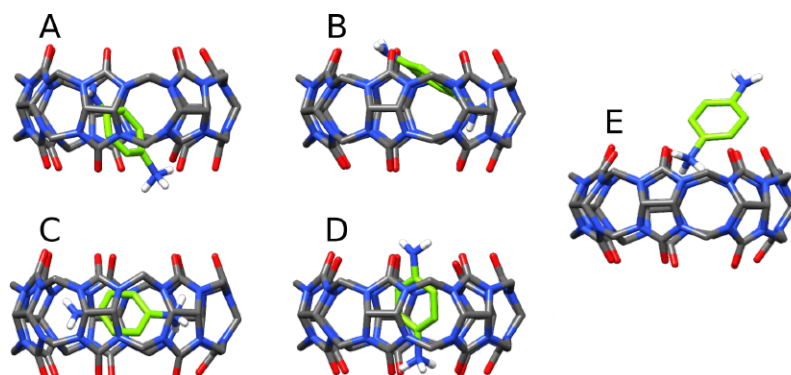


Figura B.4: Conformaciones representativas (A-E) adoptadas por el sistema CB7-BG4 durante la simulación de energía libre de unión.

Tabla B.6: Porcentaje relativo de la orientación del huésped en las ventanas de simulación con $\lambda = 1$ y asociado con las conformaciones (A-E) mostradas en la Figura B.4. Las simulaciones fueron realizadas usando las cargas atómicas RESP/AM1-BCC y D-MBIS/D-MBIS para el anfitrión y huésped respectivamente. Adicionalmente, se reportan los porcentajes de conformaciones con la presencia de agua en el sitio de unión.

CB7-BG4							
Cargas atómicas	Ventanas totales con $\lambda = 1$	Orientación del huésped					Ventanas con moléculas de agua dentro del sitio de unión
		A	B	C	D	E	
RESP/AM1BCC	41	12	2	7	76	2	7
D-MBIS/D-MBIS	42	17	2	2	64	14	14

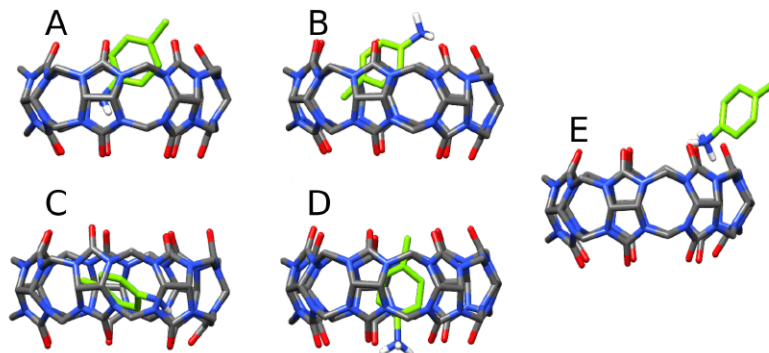


Figura B.5: Conformaciones representativas (A-E) adoptadas por el sistema CB7-BG5 durante la simulación de energía libre de unión.

Tabla B.7: Porcentaje relativo de la orientación del huésped en las ventanas de simulación con $\lambda = 1$ y asociado con las conformaciones (A-E) mostradas en la Figura B.5. Las simulaciones fueron realizadas usando las cargas atómicas RESP/AM1-BCC y D-MBIS/D-MBIS para el anfitrión y huésped respectivamente. Adicionalmente, se reportan los porcentajes de conformaciones con la presencia de agua en el sitio de unión.

CB7-BG5							
Cargas atómicas	Ventanas totales con $\lambda = 1$	Orientación del huésped					Ventanas con moléculas de agua dentro del sitio de unión
		A	B	C	D	E	
RESP/AM1BCC	34	35	3	3	53	6	6
D-MBIS/D-MBIS	45	22	0	13	38	27	22

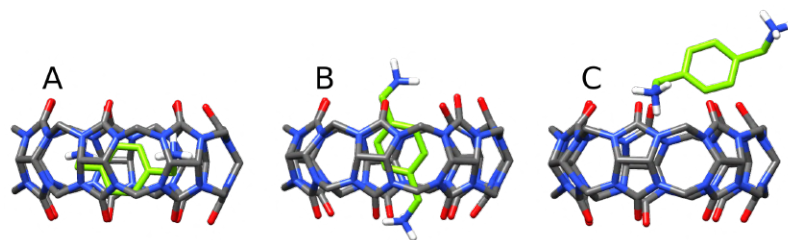
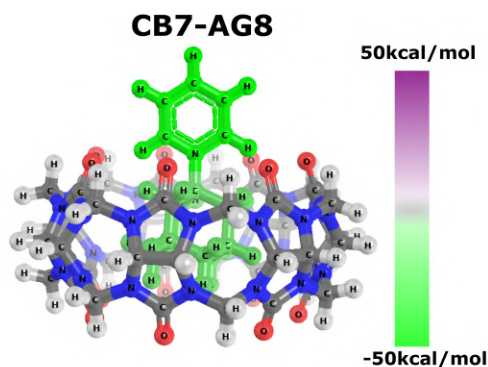


Figura B.6: Conformaciones representativas (A-C) adoptadas por el sistema CB7-BG6 durante la simulación de energía libre de unión.

Tabla B.8: Porcentaje relativo de la orientación del huésped en las ventanas de simulación con $\lambda = 1$ y asociado con las conformaciones (A-E) mostradas en la Figura B.6. Las simulaciones fueron realizadas usando las cargas atómicas RESP/AM1-BCC y D-MBIS/D-MBIS para el anfitrión y huésped respectivamente. Adicionalmente, se reportan los porcentajes de conformaciones con la presencia de agua en el sitio de unión.

CB7-BG6					
Cargas atómicas	Ventanas totales con $\lambda = 1$	Orientación del huésped			Ventanas con moléculas de agua dentro del sitio de unión
		A	B	C	
RESP/AM1BCC	46	2	89	9	9
D-MBIS/D-MBIS	43	2	74	23	21



Contribución energética

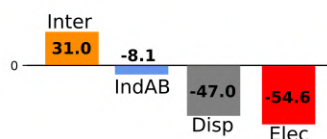


Figura B.7: Análisis de Perturbación Adaptado a la Simetría (SAPT) de las interacciones en el sistema CB7-AG8 a nivel SAPT0/jun-ccVDZ. Los átomos del huésped se colorean según su contribución a la energía de interacción. Las barras en el medio muestran la parte electrostática (Elec), dispersión (Disp), inducción desde el anfitrión (A) al huésped (B) (IndAB) e intercambio (Inter) de la energía de interacción total es expresada en kcal/mol. Los gráficos circulares de la parte inferior muestran la contribución de cada grupo funcional a la energía de interacción total en porcentaje y la energía de interacción en kcal/mol.

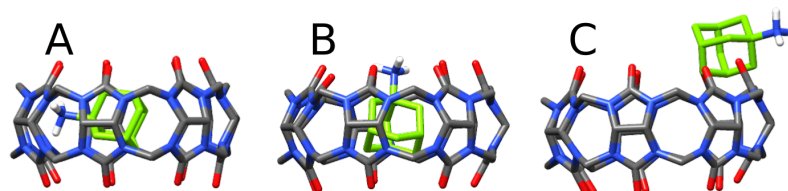


Figura B.8: Conformaciones representativas (A-C) adoptadas por los huéspedes con núcleo aromático (AG1-8) durante la simulación de energía libre de unión

Tabla B.9: Peso relativo en % de una específica orientación del huésped (A-C) mostradas en la Figura B.8 observado en las ventanas de simulación con $\lambda = 1$ durante los cálculos de energía libre de unión. Las simulaciones se realizaron utilizando las cargas atómicas RESP/AM1-BCC para el anfitrión y el huésped respectivamente y utilizando los parámetros de LJ del campo de fuerza GAFF. Además, se reportan los porcentajes de conformaciones con presencia de agua en el sitio de unión.

RESP/AM1-BCC					
Sistema	Ventanas totales con $\lambda = 1$	Orientación del huésped			Ventanas con moléculas de agua dentro del sitio de unión
		A	B	C	
CB7-AG1	28	39	43	18	18
CB7-AG2	38	0	58	42	29
CB7-AG3	29	3	90	7	7
CB7-AG4	34	0	97	3	3
CB7-AG5	28	4	82	14	14
CB7-AG6	22	0	86	14	14
CB7-AG7	33	6	91	3	3
CB7-AG8	28	0	89	11	0

Tabla B.10: Peso relativo en % de una específica orientación del huésped (A-C) mostradas en la Figura B.8 observado en las ventanas de simulación con $\lambda = 1$ durante los cálculos de energía libre de unión. Las simulaciones se realizaron utilizando las cargas atómicas D-MBIS/D-MBIS para el anfitrión y el huésped respectivamente y utilizando los parámetros de LJ del campo de fuerza GAFF. Además, se reportan los porcentajes de conformaciones con presencia de agua en el sitio de unión.

D-MBIS/D-MBIS					
Sistema	Ventanas totales con $\lambda = 1$	Orientación del huésped			Ventanas con moléculas de agua dentro del sitio de unión
		A	B	C	
CB7-AG1	49	41	57	2	2
CB7-AG2	38	0	45	55	29
CB7-AG3	45	0	89	11	9
CB7-AG4	50	0	100	0	0
CB7-AG5	27	0	74	26	19
CB7-AG6	39	3	92	5	8
CB7-AG7	31	3	94	3	3
CB7-AG8	39	0	90	10	5

Tabla B.11: Peso relativo en % de una específica orientación del huésped (A-C) mostradas en la Figura B.8 observado en las ventanas de simulación con $\lambda = 1$ durante los cálculos de energía libre de unión. Las simulaciones se realizaron utilizando las cargas atómicas D-MBIS/D-MBIS para el anfitrión y el huésped respectivamente y utilizando los parámetros de LJ del campo de fuerza GAFF y los derivados a partir de la partición MBIS de la densidad electrónica para el anfitrión y huésped respectivamente. Además, se reportan los porcentajes de conformaciones con presencia de agua en el sitio de unión.

D-MBIS/LJ-q[D-MBIS]					
Sistema	Ventanas totales con $\lambda = 1$	Orientación del huésped			Ventanas con moléculas de agua dentro del sitio de unión
		A	B	C	
CB7-AG1	43	14	81	5	5
CB7-AG2	44	0	41	48	36
CB7-AG3	34	0	91	9	6
CB7-AG4	41	0	100	0	0
CB7-AG5	41	0	90	10	7
CB7-AG6	35	0	86	14	14
CB7-AG7	43	7	88	5	5
CB7-AG8	39	0	79	18	10

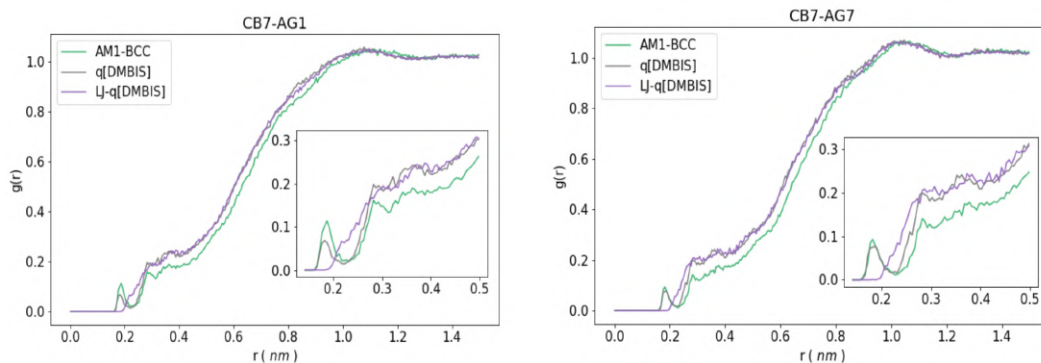
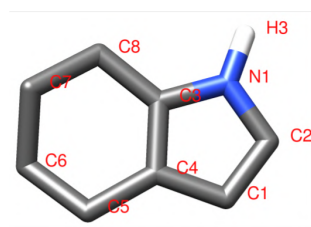


Figura B.9: Función de distribución radial (RDF) entre los átomos de oxígeno del agua y el centro de masa del huésped CB7-AG1 (izquierda) o CB7-AG7 (derecha). Los RDF en azul corresponden a los sistemas descritos con el campo de fuerza GAFF combinado con cargas atómicas RESP y AM1-BCC para el anfitrión y el huésped, respectivamente. Para los RDF grises, se utilizaron cargas atómicas D-MBIS para el anfitrión y huésped. Los RDF en violeta utilizan las mismas cargas atómicas D-MBIS combinadas con los parámetros de Lennard-Jones derivados a partir de la partición MBIS de la densidad electrónica (D-MBIS/LJ-q[D-MBIS]).

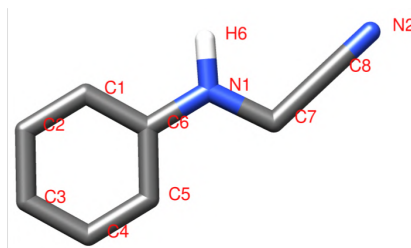
C. Anexo: Sistemas Lisozima T4 doble mutante L99A/M102Q unido a los ligandos L1-8

Tabla C.1: Valores de pK_a para todos los grupos titulables en los ligandos L1-8. ^a valores de pK_a a 298 K de acuerdo con Haynes, William M.,ed (2015) CRC Handbook of Chemistry and Physics (95th ed.). ^b valores de pK_a a 298 K según el software Chemizalize. ^c No determinado.

Ligando	pK_a
L1	ND ^c
L2	9.99 ^a
L3	10.1 ^a
L4	(1)9.34 (2)12.6 ^a
L5	10.1 ^b
L6	ND ^c
L7	ND ^c
L8	1.72 ^b



L7



L8

Figura C.1: Estructura tridimensional para los ligandos L7 y L8 con la respectiva numeración de los átomos empleadas en las simulaciones.