

Revista Latinoamericana de Difusión Científica
Volumen 5 – Número 8
Depósito Legal ZU2019000058 - ISSN 2711-0494

Revista Latinoamericana de Difusión Científica



Volumen 5 - Número 8
Enero – Junio 2023
Maracaibo – Venezuela

Métodos computacionales para estimar la afinidad de un complejo ligando-receptor

DOI: <https://doi.org/10.38186/difcie.58.03>

Martiniano Bello *

RESUMEN

A la fecha se han empleado diferentes métodos basados en la estructura para cuantificar las interacciones receptor-ligando, y a partir de estas predecir la energía libre de asociación que proporcionara un estimado de la afinidad de un compuesto por una diana terapéutica. Entre estos métodos está el acoplamiento molecular y las simulaciones de dinámica molecular en conjunto con métodos de cálculo de energía libre de asociación. El acoplamiento molecular, aunque tiene un alto potencial selectivo posee un éxito limitado en la precisión de la estimación de la energía de solvatación y consideración de cambios en la entropía conformacional. Por lo tanto, se ha recurrido a técnicas computacionales más eficientes que predicen la energía libre de unión de una manera más precisa, como lo son los métodos que combinan mecánica molecular con métodos de cálculo de energía. En este contexto, los métodos MMPBSA y MMGBSA permiten predecir la energía libre de unión usando mecánica molecular y modelos continuos de solvatación implícita. Estas técnicas han facilitado la identificación de diferentes compuestos con alta afinidad por una diana farmacológica. En este artículo científico describiremos las bases fundamentales de los métodos MMPBSA y MMGBSA, así como algunos avances relacionados con el empleo de ambos métodos.

PALABRAS CLAVE: Farmacología, energía libre de unión, dinámica molecular, MMPBSA, acoplamiento molecular.

*Laboratorio de Diseño y Desarrollo de Nuevos Fármacos e Innovación Biotecnológica, Escuela Superior de Medicina, Instituto Politécnico Nacional, Ciudad de México 11340, México. ORCID: <https://orcid.org/0000-0002-9686-0755>. E-mail: bellomartini@gmail.com, mbellor@ipn.mx

Computational methods for estimating the affinity of a ligand-receptor complex

ABSTRACT

To date, different structure-based methods have been used to quantify receptor-ligand interactions, and from these to predict the free energy of association that will provide an estimate of the affinity of a compound for a therapeutic target. Among these methods are molecular docking and molecular dynamics simulations in conjunction with methods for calculating free energy of association. Molecular docking, although it has a high selective potential, has limited success in accurately estimating solvation energy and considering changes in conformational entropy. Therefore, more efficient computational techniques that predict binding free energy more accurately have been used, such as methods that combine molecular mechanics with energy calculation methods. In this context, the MMPBSA and MMGBSA methods allow the prediction of binding free energy using molecular mechanics and continuous implicit solvation models. These techniques have facilitated the identification of different compounds with high affinity for a pharmacological target. In this scientific article, we will describe the essential bases of the MMPBSA and MMGBSA methods, as well as some advances related to the use of both methods.

KEYWORDS: Pharmacology, Binding free energy, Molecular dynamic simulation, MMPBSA, molecular docking.

Introducción

El diseño moderno de nuevos fármacos es un procedimiento que implica varias etapas con una duración de alrededor de 15 años, con un gasto que va de 500 a 1000 millones de dólares. La fase inicial del descubrimiento de nuevos fármacos típicamente consiste en la identificación de la diana farmacológica y de una serie de compuestos con afinidad hacia dicha diana farmacológica. La gran cantidad de estructuras tridimensionales disponibles en el Protein Data Bank (PDB, por sus siglas en inglés), la mejora continua de software y algoritmos para el modelado molecular de las interacciones químicas, ha hecho posible el uso del diseño de nuevos fármacos asistido por computadora, que permite predecir la afinidad de miles de compuestos en una o varias dianas farmacológicas, lo que ha facilitado acelerar la identificación de nuevos compuestos en tiempos relativamente cortos y con bajo costo (Gilson y Zhou, 2007; Hayes y Leonidas, 2010). Aunque

comúnmente para que un compuesto pueda ser considerado un fármaco viable, debe de poseer buenas propiedades farmacocinéticas de absorción, distribución, metabolismo, excreción y toxicidad (ADMET), la propiedad más relevante es una alta selectividad hacia la diana farmacológica de interés.

En general, la mayoría de los métodos empleados para el desarrollo de nuevos fármacos mediante técnicas computacionales se enfocan en predecir la energía libre de unión de un compuesto candidato, y tales métodos pueden dividirse en aquellos que aplican el diseño basado en estructura (en inglés, structure-based design) y el diseño no basado en estructura (en inglés, non-structure-based design).

Con respecto a los métodos no basados en estructura podemos mencionar al método que relaciona de manera cuantitativa la estructura y actividad (en inglés, quantitative structure-activity relationships, QSAR), el cual permite predecir la actividad de nuevos compuestos basado en el conocimiento previo de la actividad de ligandos con alta afinidad para una diana terapéutica en particular (Tropsha, 2010). Este método tiene la desventaja de ser poco predictivo, puesto que considera que moléculas químicamente parecidas tienen propiedades similares, por lo cual muchas veces hace que se propongan compuestos con actividades biológicas diferentes a la esperada.

Lo anterior se debe a un fenómeno denominado acantilados de actividad (Maggiara, 2006; Cruz-Montegudo et al., 2014). Los acantilados de actividad se dan entre pares de compuestos químicamente similares, pero que tienen una diferencia importante en su potencia; por lo tanto, la correcta identificación de dichos acantilados ayuda a mejorar la calidad estadística de los modelos QSAR (Stumpfe et al., 2016).

Los métodos basados en la estructura emplean la información estructural de la macromolécula, la cual comúnmente es una proteína, para obtener información sobre las interacciones macromolécula-ligando, y a partir de la cuantificación de las interacciones termodinámicamente favorables predecir la energía libre de asociación que dará un estimado de la afinidad del compuesto por la diana terapéutica. Entre estos métodos resaltan: el acoplamiento molecular (en inglés molecular docking) y la simulación de dinámica molecular (en inglés molecular dynamics simulation) combinada con métodos de cálculo de energía libre de unión. El método de acoplamiento molecular ha sido utilizado de manera amplia para el tamizaje virtual a gran escala (en inglés, high throughput virtual

screening) de miles de compuestos disponibles en diversas bases de datos dirigidos hacia dianas terapéuticas específicas (Genheden & Ryde, 2015). El acoplamiento molecular es eficiente y de bajo costo, se centra en la identificación de varios compuestos con un alto potencial de ser selectivos para una diana farmacológica específica (Bello et al., 2013; Lavecchia y di Giovanni, 2013); no obstante, la determinación de las energías libres de unión mediante esta metodología generalmente no es de alta precisión (Wang et al., 2016; Li et al., 2010; Hou et al., 2013).

Por consiguiente, se ha recurrido a metodologías computacionales más eficientes que permiten predecir la energía libre de unión de una manera más precisa, como lo son los métodos que combinan mecánica molecular con métodos de cálculo de energía. En este contexto, la metodología de Mecánica Molecular Poisson-Boltzmann/Área de Superficie (en inglés, Molecular Mechanics Poisson-Boltzmann Surface Area, MMPBSA) y el de Mecánica Molecular Generalizada-Born/Área de Superficie (en inglés, Molecular Mechanics Generalized-Born Surface Area, MMGBSA), permiten la determinación de la energía libre de unión (ΔG) usando mecánica molecular (Campo de fuerza) y modelos continuos de solvatación implícita (Kollman et al., 2000), que han sido empleados en diferentes dianas terapéuticas y sistemas de transporte de fármacos (Genheden y Ryde, 2015), y está disponible en paquetes computacionales como Amber (Case et al., 2005), Delphi (Rocchia et al., 2001), GROMACS (Van Der Spoel et al., 2005) y Schrödinger (Du et al., 2011), que son de uso amigable, y que ha permitido la identificación de diferentes compuestos con alta afinidad por una diana farmacológica.

En este artículo discutiremos las bases fundamentales de los métodos MMPBSA y MMGBSA (MMPB(GB)SA), así como algunos avances relacionados con la combinación del acoplamiento molecular y el método MMPB(GB)SA, resaltando su importancia con respecto a otros métodos menos eficientes.

1. Cálculo de la energía libre de Gibbs

La constante de afinidad (K_a) de un ligando por una proteína se puede representar mediante la siguiente ecuación:

$$K_a = \frac{[PL]}{[P][L]} \dots \dots \dots (1)$$

donde [P], [L] y [PL] representan concentraciones en equilibrio de la proteína, ligando y del complejo proteína-ligando, respectivamente. Esta definición se basa en el supuesto que la diferencia entre el estado libre y asociada es claramente definida, la cual se cumple solo para complejos con asociaciones fuertes; pero se vuelve un poco imprecisa para asociaciones débiles.

Esta afinidad puede relacionarse con la energía libre de asociación mediante la siguiente ecuación.

$$\Delta G = -RT \ln K_a \dots \dots \dots (2)$$

Donde ΔG representa la energía libre de Gibbs, R es la constante de los gases, T es la temperatura absoluta, y K_a es la constante de afinidad entre dos especies, esta última es la que deseamos estimar mediante estudios computacionales.

2. Métodos de cálculos de energía

El creciente interés por estos métodos surge de que la energía libre de unión entre un receptor y un ligando puede estimarse de manera relativamente eficiente, empleando un ciclo termodinámico que simplifica una pequeña perturbación del espacio conformacional, en contraste con otros cuantificables de interés farmacéutico, como lo son las energías de enlace absolutas y cinéticas de enlace, que necesitan recursos computacionales considerables.

Entre los diferentes métodos que combinan mecánica molecular con métodos de cálculo de energía destacan los métodos de ruta (en inglés, Pathway methods) y los métodos de punto final (en inglés, End-point methods) (Deng y Roux, 2009; Gilson y Zhou, 2007). En el método de ruta, el sistema se convierte de un estado (v.gr., el complejo proteína-ligando) a otro (v.gr., el estado libre de la proteína y el ligando), introduciendo un conjunto de intermediarios no físicos (alquímicos) finitos o infinitesimales en la función de energía (Hamiltoniano) del sistema (Kollman, 1993; Straatsma y McCammon, 1992). El valor de ΔG es determinado a través de las diferencias en la afinidad entre el conjunto de intermediarios alquímicos. Entre los métodos de ruta destacan: Perturbación de la energía libre (en inglés, Free energy perturbation, FEP) y el método de Integración termodinámica

(en inglés, Thermodynamic integration, TI) (Beveridge y Dicapua, 1989). En contraste, en el método de punto final, solo se consideran los estados iniciales (receptor y ligando libres) y finales (complejo receptor-ligando) para determinar el valor de ΔG . Aunque los fundamentos de los métodos FEP y TI fueron postulados hace décadas (Kirkwood, 1935; Zwanzig, 1954), fue en años recientes que empezaron a emplearse para el cálculo del valor de ΔG , debido al incremento en el poder de cómputo, al desarrollo de campos de fuerza más precisos (Cornell et al., 1995; Van Gunsteren., 1996; MacKerell et al., 1998;), la implementación de protocolos para la determinación de la ΔG (Boresch et al., 2003; Boresch y Bruckner, 2011; Lee y Olson, 2006; Mobley et al., 2007), la combinación de dinámica molecular o simulaciones Monte Carlo en solvente explícito, y la reciente implementación de códigos biomoleculares en unidades de procesamiento gráfico o tarjetas gráficas (Harvey et al., 2009; Stone et al., 2011).

Entre los métodos de punto final destacan: Interacción lineal de energía (en inglés, Linear Interaction Energy, LIE) (Aqvist et al., 1994), y el método MMPB(GB)SA (Kollman et al., 2000; Wang et al., 2006; Wang et al., 2001;). Aunque la evaluación de la ΔG en todos estos métodos se determina a través de múltiples conformaciones obtenidas a partir de la simulación molecular, existen diferencias entre ellos. De estos métodos, los métodos de ruta FEP y TI, determinan la energía libre basada en perturbaciones del complejo biológico a partir de una sola simulación molecular; estos métodos realizan un tratamiento exhaustivo de los grados de libertad y emplean modelos de solvatación apropiados; no obstante, el costo computacional de dichos métodos es alto debido a la convergencia lenta del valor de la ΔG . (Meng et al., 2011), sobre todo para el cribado virtual de miles de compuestos.

Los métodos LIE y MMPB(GB)SA surgen en un intento de combinar rapidez y precisión, donde el primero considera que la energía libre de unión de un complejo es resultado de la combinación lineal de las energías electrostáticas y de Van der Waals. En tanto que en el método MMPB(GB)SA, la energía libre de un sistema se evalúa a partir de la combinación de mecánica molecular, la energía electrostática, que se estima mediante los modelos de solvatación implícita Poisson-Boltzmann o Generalizada Born, y un término entrópico. Aunque los métodos LIE y MMPB(GB)SA han mostrado resultados similares, estos últimos han sido empleados en una manera exitosa en un mayor número de

investigaciones. En este artículo describiremos en qué consiste el método MMPB(GB)SA, así como algunas aplicaciones recientes empleando ambos métodos.

3. El método MMPB(GB)SA

El método MMPB(GB)SA es empleado para determinar los valores de ΔG relativo, el cual no incluye la contribución entrópica, de un conjunto de ligandos a una diana farmacológica. El método tiene alta reproducibilidad y precisión con un bajo costo computacional comparado con los métodos FEP y TI. Esta reproducibilidad se constata mediante la obtención de valores similares de ΔG por diferentes grupos de investigación, con o sin variación en su implementación, incluido el software y el hardware empleados. En tanto que la precisión se constata por la cercanía de los resultados de ΔG teóricos con los valores experimentales correspondientes. Adicionalmente, el método MMPB(GB)SA puede darnos información detallada de las energías por residuo de los aminoácidos que participan en el reconocimiento molecular (Archontis et al., 2001; Polydoridis et al., 2007; Hayes et al., 2011). Como un método de punto final, el objetivo del método MMPB(GB)SA es determinar el valor de ΔG entre dos estados que representan el estado unido y no unido de dos moléculas solvatadas.

Aunque de manera ideal nos gustaría determinar el valor de ΔG como se muestra:

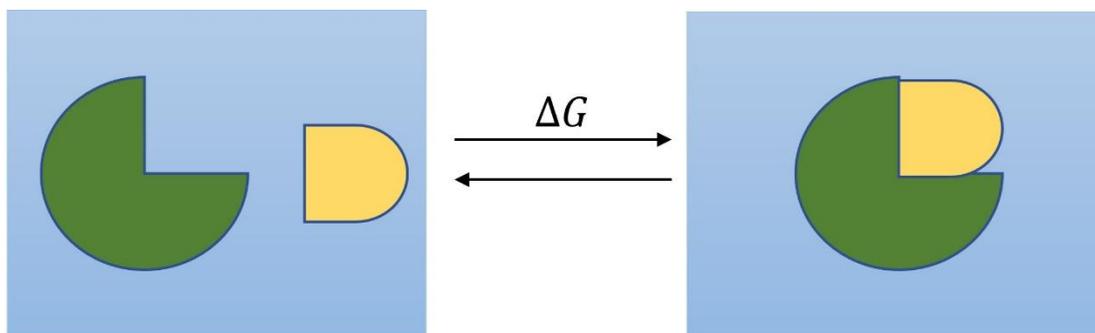


Figura 1. Ciclo termodinámico para la obtención del ΔG considerando el estado libre y asociado del complejo receptor-ligando en agua (medio solvatado).

Sin embargo, el considerar ambos estados como solvatados ocasionaría que gran parte de las contribuciones energéticas provengan de la interacción agua-agua. Adicionalmente, las fluctuaciones en la energía total serían alrededor de un orden de

magnitud mayor que las energías de enlace, ocasionando que estas tarden en alcanzar convergencia. Por lo tanto, una manera más efectiva es realizar el análisis considerando el siguiente ciclo termodinámico:

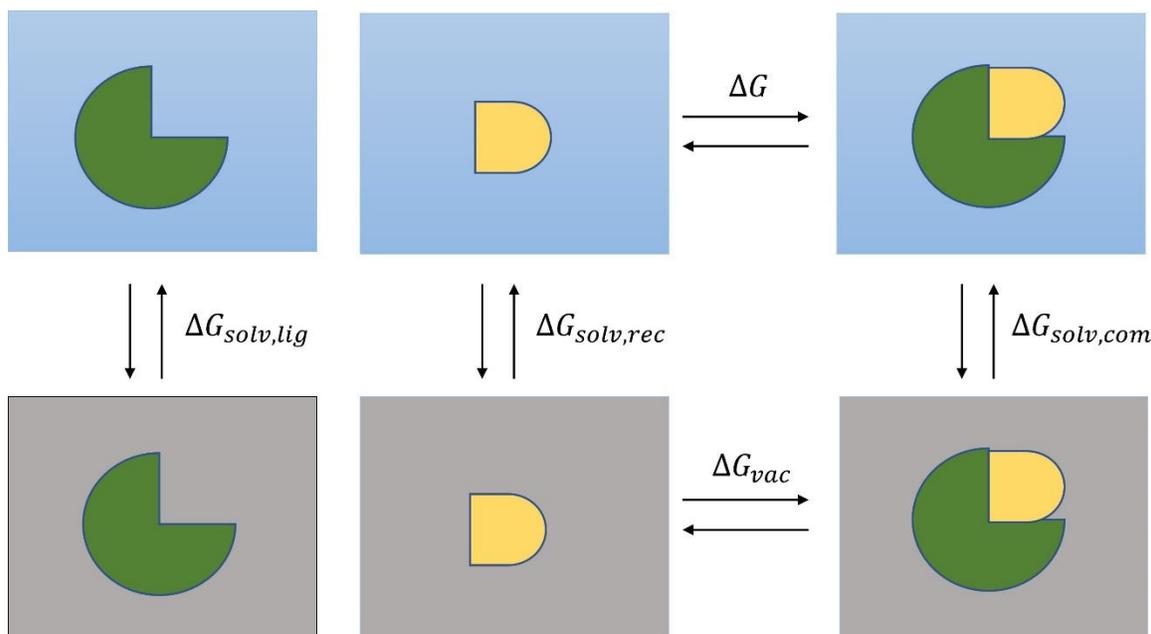


Figura 2. Ciclo termodinámico para la obtención del ΔG considerando el estado libre y asociado del complejo receptor-ligando solvatado y en vacío.

Empleando este esquema, el valor de ΔG en presencia de solvente se determina de la siguiente manera.

$$\Delta G = \Delta G_{vac} + \Delta G_{sol,com} - (\Delta G_{sol,lig} + \Delta G_{sol,rec}) \dots \dots \dots (1)$$

Donde ΔG_{vac} corresponde al valor de ΔG determinado en el vacío, mientras que $\Delta G_{sol,com}$, $\Delta G_{sol,lig}$ y $\Delta G_{sol,rec}$ corresponden al ΔG del complejo, ligando y el receptor, respectivamente, determinado en presencia de solvente. Los valores de ΔG_{solv} se determinan mediante la resolución de la ecuación de Poisson-Boltzmann (PB) o de Generalizada Born para cada uno de los tres estados. Como resultado se obtiene la contribución electrostática considerando la constante dieléctrica del exterior e interior del receptor ($G_{elec,c=80}$, y $G_{elec,c=1}$), y las contribuciones hidrofóbicas (ΔG_{hidrof}) de la ΔG_{solv} :

$$\Delta G_{solv} = G_{elec,c=80} - G_{elec,c=1} + \Delta G_{hidrof} \dots \dots \dots (2)$$

En tanto que el valor de ΔG_{vac} se obtiene a través del valor promedio de la energía de interacción (ΔE_{MM}) entre el receptor y ligando, y considerando la contribución entrópica (ΔS), si es requerido para el análisis.

$$\Delta G_{vac} = E_{MM} - T \cdot \Delta S \dots \dots \dots (3)$$

A su vez E_{MM} es la sumatoria de las interacciones unidas (en inglés, non-bonded interactions) o covalentes y las interacciones no unidas (en inglés, non-bonded interactions): interacciones electrostáticas (elec) y de Van der Waals (vdw).

$$E_{MM} = E_{bonded} + E_{elec} + E_{vdw} \dots \dots \dots (4)$$

Estas contribuciones se calculan a partir de las estructuras atómicas de las especies libres y formando el complejo empleando el campo de fuerza.

En tanto que el ΔS contiene las contribuciones traslacionales, rotacionales y vibracionales. Las contribuciones traslacionales y rotacionales se calculan mediante expresiones de la mecánica estadística estándar; mientras que la contribución vibracional se determina mediante modos normales (Karplus y Kushick 1981; Tidor y Karplus 1994; Brooks et al., 1995). En teoría se pueden emplear diferentes protocolos para realizar estudios empleando el método MMPB(GB)SA. No obstante, se ha puesto atención a una serie de detalles empleados para realizar los cálculos.

En general, un protocolo consiste en estos tres pasos. El primer paso es realizar estudios de simulaciones de dinámica molecular en solvente explícito para obtener una serie de conformaciones de complejos receptor-ligando; en esta fase es importante que las conformaciones seleccionadas estén en equilibrio. En la segunda fase, se determinan las energías de interacción E_{MM} y los valores de ΔG_{solv} sobre el grupo de conformaciones obtenidos en el estudio de dinámica molecular. Para este paso es necesario remover las moléculas de agua y reemplazarlas por un continuo dieléctrico. Finalmente, se determinan las contribuciones entrópicas ΔS para el conjunto de estructuras, y dicho valor entrópico se

combina con el ΔG para obtener el valor final que se describe en la ecuación 2. Aunque hemos descrito como se realizan los estudios con el método MMPB(GB)SA empleando solo una simulación de dinámica molecular (en inglés single-trajectory approximation) (Gohlke y Case, 2004); también se emplea el método de tres simulaciones (en inglés, three-trajectory approximation), en el que se realizan simulaciones independientes para el ligando, receptor y complejo receptor-ligando. Esta última es considerada cuando el sistema a evaluar experimenta cambios conformacionales significativos en su reconocimiento molecular, llevando a mejores estimaciones de la energía libre de unión que aquellas determinadas usando solo una simulación de dinámica molecular (Yang et al., 2009).

3.1. Aplicaciones del método MMPB(GB)SA

La combinación de acoplamiento molecular, dinámica molecular y determinación de la energía libre de unión empleado el método MMPB(GB)SA nos permiten abordar muchos problemas tanto en biofísica estructural como en el diseño de fármacos. Con estas técnicas se pueden modelar estructuras biológicas complejas, calcular las energías libres de asociación, e identificar los principales contactos (residuos de aminoácidos) que contribuyen de manera importante en la formación del complejo biológico. En esta sección describiremos algunas aplicaciones del método MMPB(GB)SA que se han empleado para determinar la afinidad para una variedad de sistemas biológicos.

- Ejemplo 1: **Desarrollo de fármacos dirigidos a combatir el cáncer de mama**

Las anomalías en los niveles de expresión del receptor del factor de crecimiento epidérmico humano tipo 1 (EGFR, por sus siglas en inglés: epidermal growth factor receptor) y tipo 2 (HER2, por sus siglas en inglés: human epidermal growth factor receptor 2) se encuentran sobreexpresados en muchos tipos de cáncer humano. Por lo consiguiente, el diseño de compuestos con actividad inhibitoria dual de EGFR/HER2 es una estrategia anticancerígena reconocida. Estudios experimentales demostraron que algunos derivados de lapatinib, un compuesto que ha demostrado actividad dual inhibitoria de EGFR/HER2 (Rusnak et al., 2001; Konecny et al., 2006), con modificación en el grupo metilsulfoniletilamino, demostraron que el compuesto 2i mostró una potente actividad inhibitoria contra las células cancerosas que sobreexpresan EGFR/HER2, y también una

mayor actividad inhibitoria utilizando un KitC EGFR/HER2 (IC_{50} de 0.5 y 4.1 nM para HER2 y EGFR, respectivamente) que lapatinib (IC_{50} de 2.7 y 7.9 nM para HER2 y EGFR, respectivamente) (Lyu et al., 2014). Debido a la falta de información estructural experimental para la unión de estos ligandos, se realizaron estudios de acoplamiento y dinámica molecular en solvente explícito usando Amber 16 (Bello et al., 2020) para cada complejo EGFR-inhibidor o HER2-inhibitor, con dinámicas moleculares de producción de 100 nanosegundos (ns), escala de tiempo suficiente para observar la relajación y equilibrio de este tipo de sistema. Basado en el tiempo de simulación en la que los complejos alcanzaron el equilibrio, se determinaron las conformaciones predominantes mediante un análisis de agrupamientos (análisis de clústers) para determinar las interacciones predominantes durante la dinámica molecular. Adicionalmente, se estimaron los valores de afinidad de los complejos empleando el método MMPB(GB)SA. Basados en sus resultados dilucidaron cómo una ligera modificación de lapatinib contribuye a un aumento en la afinidad de ambos receptores por el compuesto 2i en comparación con el compuesto lapatinib, que está de acuerdo con los resultados experimentales. Adicionalmente, el análisis de componentes principales (PCA, por sus siglas en inglés: Principal Component Analysis) permitió observar que el acoplamiento de 2i en EGFR/HER2 está relacionado con una reducción en la movilidad conformacional, que también puede contribuir a la mejora de la afinidad por este compuesto en comparación con lapatinib.

- Ejemplo 2: **Efectos de las mutaciones en la predicción del valor de ΔG**

Las mutaciones son frecuentes en las proteínas, lo cual repercute en el mapa de interacciones y afinidad entre la proteína y un fármaco, por lo que establecer el impacto de las mutaciones en el valor de ΔG es de gran importancia. En un estudio reciente se evaluó la capacidad predictiva del método MMPB(GB)SA empleando diferentes proteínas conteniendo mutaciones (Yu et al., 2021). Estos estudios se realizaron considerando las siguientes estrategias computacionales: diferentes longitudes de simulaciones de dinámica molecular, diferentes valores de las constantes dieléctricas, y consideración o no de los efectos entrópicos en el cálculo del valor de ΔG . Los resultados de este estudio señalaron que una simulación de dinámica molecular relativamente larga, por ejemplo, de 100 ns, beneficia la precisión de la predicción tanto para el método MMGBSA y MMPBSA. También

se observó que los sistemas con un mayor número de mutaciones mostraron una mejor correlación entre los valores teóricos y experimentales, indicando que el algoritmo es más sensible a cambios más significativos, es decir, a un número considerable de mutaciones. En general esta investigación demostró que para proteínas con múltiples mutaciones es necesario un ajuste conformacional para refinar el microambiente en los sistemas mutados y así poder valores ΔG confiables, lo cual explica por qué se obtuvieron mejores resultados en simulaciones de dinámica molecular más largas.

- Ejemplo 3: **Predicción del ΔG empleando el método MMGBSA en servidor web**

El empleo del método MMPB(GB)SA para la predicción del valor de ΔG está implementado en diferentes paquetes computacionales como Amber (Case et al., 2005) y GROMACS (Van Der Spoel et al., 2005); no obstante, su empleo requiere cierto tipo de conocimiento de manejo de plataformas UNIX, bioinformática y quimioinformática para construir los sistemas a simular, lo cual es un reto incluso para los expertos; por lo tanto, se han desarrollado servidores web como HawkDock (Weng et al., 2019) y farPPI (Wang et al., 2019) que ayudan en la predicción del valor de ΔG para un complejo proteína-proteína empleando el método MMPB(GB)SA a partir de un complejo generado mediante estudios de acoplamiento molecular. Mas recientemente, se desarrolló otro servidor web llamado fastDRH (Wang et al., 2022) diseñado para predecir valor de ΔG para un complejo proteína-ligando combinando el acoplamiento molecular, y el método MMPB(GB)SA con estrategias mejoradas a las empleadas en HawkDock y farPPI. No obstante, un par de inconvenientes de fastDRH comparado con el cálculo del valor del ΔG mediante la combinación de simulaciones de dinámica molecular con el método MMPB(GB)SA, es que no considera cambios conformacionales en el complejo proteína-ligando, lo cual es muy importante cuando se exploran sistemas muy flexibles; y que trunca o recorta el sistema a evaluar para acelerar la velocidad de cálculo.

- Ejemplo 4: **Predicción del valor de ΔG entre métodos de ruta y de punto final**

La fosfodiesterasa 5 (PDE5A, por sus siglas en inglés) es una enzima que participa en la conversión de guanosín monofosfato cíclico a guanosín monofosfato. El sildenafil cuyo nombre comercial es Viagra es empleado en el tratamiento de la disfunción eréctil, y es un inhibidor selectivo de la PDE5A. En un estudio reciente se

identificaron algunos inhibidores potentes de la PDE5A empleando el método de ruta FEP y de punto final MMPB(GB)SA (Wu et al., 2022). Este estudio demostró que el método FEP fue capaz de proveer en general mejores predicciones que el método MMPB(GB)SA, las cuales fueron más consistentes con los valores experimentales. Este resultado está en concordancia con reportes que señalan que los métodos de ruta son más capaces de proporcionar predicciones más consistentes que los métodos de punta final, aunque tienen la desventaja de un costo computacional elevado debido a la convergencia lenta del valor de ΔG . (Meng et al., 2011), costo que se ve acentuado en el cribado virtual de miles de compuestos.

- Ejemplo 5: **Implementaciones en el método MMPB(GB)SA para mejorar su capacidad predictiva**

Aunque el método MMPB(GB)SA se ha empleado ampliamente para la predicción del valor de ΔG , en la práctica es un método en disputa, ya que es dependiente del sistema en estudio. En un estudio reciente, se desarrolló una versión novedosa denominada MMPB(GB)SA de dieléctrico atómico variable (variable atomic dielectric, por sus siglas en inglés) VAD- MMPB(GB)SA (Wang et al., 2021). Con esta estrategia se asignan constantes dieléctricas variables directamente a los átomos de proteína y el ligando, permitiendo una precisión mejorada en los cálculos del valor de ΔG . VAD-MM/GBSA también demostró ser mejor que Prime MMGBSA, que es una versión del método MMPB(GB)SA implementado en el software Schrödinger. Basado en lo anterior es posible que el método VAD-MMPB(GB)SA con poca demanda computacional reemplace al método MMPB(GB)SA presente en el software AMBER. Adicionalmente, existe un servidor en línea disponible para realizar predicciones del ΔG empleando el método VAD-MMGBSA.

Otro reto en el método MMPB(GB)SA es la determinación correcta del componente entrópico, el cual también es sistema dependiente y en muchos casos en lugar de mejorar el poder predictivo del método lo empeora. En un estudio reciente realizaron la implementación del método denominado de interacción entrópica (Interaction entropy, IE, por sus siglas en inglés) (Zhong et al., 2020). El desempeño de esta implementación fue evaluado considerando 176 complejos proteína-ligando y proteína-proteína de la familia de la proteína Bcl-2 (B-cell lymphoma 2, por sus siglas en inglés), que participa de manera

importante en el desarrollo normal de las células y es uno de los principales reguladores de la vía apoptótica celular. La incorporación de la IE fue capaz de mejorar la capacidad predictiva del método MMPB(GB)SA, proporcionando mejores resultados para los sistemas proteína-proteína que para proteína-ligando.

- Ejemplo 6: **Implementaciones en el protocolo MMPB(GB)SA para mejorar su capacidad predictiva en complejos TCR-HLA**

El complejo entre el receptor de células T (TCR por sus siglas en inglés) y su ligando endógeno, el antígeno peptídico leucocitario humano (pHLA, por sus siglas en inglés) es un componente importante del sistema inmunológico del cuerpo que permite diferenciar células propias en condiciones normales o en el curso de una enfermedad metabólica. El reconocimiento molecular de pHLA por parte del TCR promueve la activación de células T, que apoyan en la destrucción y erradicación de células enfermas. En este contexto se han diseñado diferentes moléculas TCR que están dirigidos a unir pHLA específicos presentes en células enfermas, como por ejemplo de cáncer, pero sin afinidad para la gran mayoría de pHLA presentados por células sanas, lo cual evita efectos tóxicos (Oates et al., 2013). Para lograr la especificidad del complejo TCR-pHLA, se requiere que el pHLA tenga una alta afinidad por el TCR en el rango picomolar.

En este contexto, Crean et al. (2022), exploraron diferentes protocolos que les permitieron evaluar de manera confiable el valor de ΔG para complejos entre variantes del receptor de células. Para este estudio consideraron diferentes complejos TCR-pHLA con 3 a 14 mutaciones presentes en pHLA. Sus resultados señalaron que una serie de propiedades pueden ser consideradas para el estudio de estos sistemas: constante dieléctrica interna de entre 4 y 8, simulaciones de dinámica molecular cortas de entre 4 a 5 ns, solvatación explícita en la superficie de interacción proteína-proteína, y considerar el TCR truncado o recortado para el cálculo de la entropía. Estos resultados podrían aplicarse para el desarrollo y optimización de TCR con potencial terapéutico.

Conclusión

MMPB(GB)SA es un método de cálculo de energía libre de punto final que permite calcular la energía libre de unión absoluta con un modesto costo computacional. Aunque

inicialmente MMPB(GB)SA se presentó como un método que no contenía ningún parámetro de ajuste, a la fecha el método incluye varios parámetros como la constante dieléctrica, la energía no polar, los radios utilizados para las constantes de solvatación implícita (PB y GB), y la inclusión de la entropía, que pueden ajustarse dependiendo del sistema a estudiar. No obstante, el método involucra varias aproximaciones termodinámicas; por ejemplo, el método ignora la entropía conformacional del complejo, así como la entropía de las moléculas de agua involucradas en la asociación molecular. Por consiguiente, debido a estas aproximaciones, el método MMPB(GB)SA es más aplicable para clasificar las afinidades de unión entre un grupo de ligandos y un receptor específico, que para determinar las energías libres de unión absolutas que puedan reproducir energías libres de unión experimentales; sin embargo, en varios casos el método ha permitido reproducir datos experimentales.

Referencias

Aqvist, J.; Medina, C.; Samuelsson, J. E. (1994). New Method for Predicting Binding-Affinity in Computer-Aided Drug Design. *Protein Eng.* 7 (3), 385–391. doi: 10.1093/protein/7.3.385

Archontis, G.; Simonson T.; Karplus M. (2001). Binding free energies and free energy components from molecular dynamics and Poisson-Boltzmann calculations. Application to amino acid recognition by aspartyl-tRNA synthetase. *J Mol Biol.*, 306 (2), 307-327. doi: 10.1006/jmbi.2000.4285

Bello, M.; Martinez-Archundia, M.; Correa-Basurto, J. (2013). Automated docking for novel drug discovery. *Expert Opinion on Drug Discovery*, 8, 821-834. doi: 10.1517/17460441.2013.794780

Bello, M.; Guadarrama-García, C.; Rodríguez-Fonseca, R. A (2020). Dissecting the molecular recognition of dual lapatinib derivatives for EGFR/HER2. *J Comput Aided Mol Des.* 2020 Mar;34(3):293-303. doi: 10.1007/s10822-019-00270-4

Beveridge, D. L.; Dicapua, F. M. (1989). Free-Energy Via Molecular Simulation - Applications to Chemical and Biomolecular Systems. *Annu. Rev. Biophys. Biomol. Struct.*, 18, 431-492. doi: 10.1146/annurev.bb.18.060189.002243

Boresch, S.; Bruckner, S. (2011). Avoiding the van der Waals endpoint problem using serial atomic insertion. *Journal of Computational Chemistry*, 32, 11, 2449-2458. doi: 10.1002/jcc.21829

Boresch, S.; Tettinger, F.; Leitgeb, M.; Karplus, M. (2003). Absolute binding free energies: A quantitative approach for their calculation. *Journal of Physical Chemistry B*, 107, 35, 9535-9551. DOI:10.1021/JP0217839

Brooks, B.R.; Janezic, D.; Karplus M. (1995). Harmonic-analysis of large systems. 1. Methodology. *J. Comput. Chem.*, 16, 1522-1542. <https://doi.org/10.1002/jcc.540161209>

Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. (2005). The Amber biomolecular simulation programs. *Journal of Computational Chemistry*, 26(16), 1668-1688. doi: 10.1002/jcc.20290

Cornell, W.D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman P. A. (1995). A second generation forcefield for the simulation of proteins, nucleic acids, and organic molecules. *Journal of the American Chemical Society*, 117, 19, 5179-5197. <https://doi.org/10.1021/ja00124a002>

Crean, R. M.; Pudney, C. R.; Cole, D. K.; Van der Kamp, M. W. (2022). Reliable in silico ranking of engineered therapeutic TCR binding affinities with MMPB/GBSA. *Journal of chemical information and modeling*, 62(3), 577-590. doi: 10.1021/acs.jcim.1c00765

Cruz-Monteagudo, M.; Medina-Franco, J. L.; Pérez-Castillo, Y.; Nicolotti, O.; Cordeiro, M. N., et al. (2014). Activity cliffs in drug discovery: Dr. Jekyll or Mr. Hyde? *Drug Discovery Today*, 19, 1069-1080. doi: 10.1016/j.drudis.2014.02.003

Deng, Y.; Roux, B. (2009). Computations of standard binding free energies with molecular dynamics simulations. *J Phys Chem B.*, 113 (8), 2234-2246. doi: 10.1021/jp807701h

Du, J.; Sun, H.; Xi, L.; Li, J.; Yang, Y.; Liu, H.; Yao, X. (2011). Molecular modeling study of Checkpoint Kinase 1 inhibitors by multiple docking strategies and Prime/MMGBSA. *Journal of Computational Chemistry*, 32, 13, 2800-2808. DOI: [10.1002/jcc.21859](https://doi.org/10.1002/jcc.21859)

Genheden, S.; Ryde, U. (2015). The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin Drug Discov.*, 10(5), 449-461. doi: 10.1517/17460441.2015.1032936

Gilson, M. K.; Zhou, H. X. (2007). Calculation of protein-ligand binding affinities. *Annu Rev of Biophys Biomol Struct*, 36, 21-42. doi: 10.1146/annurev.biophys.36.040306.132550

Harvey, M. J.; Giupponi, G.; De Fabritiis, G. (2009). ACEMD: Accelerating biomolecularsimulations in the microsecond time scale. *Journal of Chemical Theory & Computation*, 5, 6, 1632-1639. doi: 10.1021/ct9000685

Hayes, J. M.; Leonidas, D. D. (2010). Computation as a tool for glycogen phosphorylase inhibitor design. *Mini Reviews in Medicinal Chemistry*, 10, 12, 1156-1174. doi: 10.2174/1389557511009011156

Hayes, J. M.; Skamnaki, V. T.; Archontis, G.; Lamprakis, C.; Sarrou, J.; Bischler, N.; Skaltsounis, A. L.; Zographos, S. E.; Oikonomakos, N. G. (2011). Kinetics, in silico docking,

molecular dynamics, and MM-GBSA binding studies on prototype indirubins, KT5720, and staurosporine as phosphorylase kinase ATP-binding site inhibitors: The role of water molecules examined. *Proteins.*, 79 (3), 703-719. doi: 10.1002/prot.22890

Hou, X. B.; Du, J. T.; Zhang, J.; Du, L. P.; Fang, H.; Li, M. Y (2013). How to Improve Docking Accuracy of Autodock4.2: A Case Study Using Different Electrostatic Potentials. *J. Chem. Inf. Model.*, 53, 188–200. doi: 10.1021/ci300417y

Karplus, M.; Kushick, J. N. (1981). Method for estimating the configurational entropy of macromolecules. *Macromolecules*, 14, 2, 325-332. <https://doi.org/10.1021/ma50003a019>

Kirkwood, J. G. (1935). Statistical mechanics of fluid mixtures. *Journal of Chemical Physics*, 3, 5, 300-313. <https://doi.org/10.1063/1.1749657>

Kollman, P. (1993). Free energy calculations: Applications to chemical and biochemical phenomena. *Chem. Rev.*, 93 (7), 2395-2417. <https://doi.org/10.1021/cr00023a004>

Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S. H.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E. (2000). Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. *Acc. Chem. Res.*, 33 (12), 889-897. <https://doi.org/10.1021/ar000033j>

Konecny, G. E.; Pegram, M. D.; Venkatesan, N.; Finn, R.; Yang, G. (2006). Activity of the dual kinase inhibitor lapatinib (GW572016) against HER-2-overexpressing and trastuzumab-treated breast cancer cells. *Cancer Res* 66:1630-1639. doi: 10.1158/0008-5472.CAN-05-1182

Lavecchia, A.; di Giovanni, C. (2013). Virtual screening strategies in drug discovery: A critical review. *Current Medicinal Chemistry*, 20, 2839-2860. doi: 10.2174/09298673113209990001

Lee, M. S.; Olson, M. A. (2006). Calculation of absolute protein-ligand binding affinity using path and endpoint approaches. *Biophysical Journal*, 90, 3, 864-877. doi: 10.1529/biophysj.105.071589

Li, Y.; Shen, J.; Sun, X.; Li, W.; Liu, G.; Tang, Y. (2010). Accuracy Assessment of Protein-Based Docking Programs against RNA Targets. *J. Chem. Inf. Model*, 50, 1134–1146. doi: 10.1021/ci9004157

Lyu, A.; Lei, F.; Shaohua, G. (2014). Design and synthesis of Lapatinib derivatives containing a branched side chain as HER1/HER2 targeting antitumor drug candidates. *Eur J Med Chem.*, 87, 631-642. doi: 10.1016/j.ejmech.2014.10.006.

MacKerell, A.D.; Bashford, D.; Bellott M.; Dunbrack, Jr.; R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F.T.K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhorn, B.; Reiher, III, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. (1998). All-atom empirical potential for molecular modeling

and dynamics studies of proteins. *Journal of Physical Chemistry B*.102., 18, 3586-3616. <https://doi.org/10.1021/jp973084f>

Maggiore, G. M. (2006). On outliers and activity cliffs-why QSAR often disappoints. *Journal of Chemical Information and Modeling*, 46, 1535. <https://doi.org/10.1021/ci060117s>

Meng, Y.; Sabri Dashti, D.; Roitberg, A. E (2011). Computing Alchemical Free Energy Differences with Hamiltonian Replica Exchange Molecular Dynamics (H-REMD) Simulations. *J. Chem. Theory Comput.*, 7, 2721–2727. doi: [10.1021/ct200153u](https://doi.org/10.1021/ct200153u)

Mobley, D. L.; Graves, A. P.; Chodera, J. D.; McReynolds, A. C.; Shoichet, B. K.; Dill, K. A. (2007). Predicting absolute binding free energies to a simple model site. *Journal of Molecular Biology.*, 371, 4, 1118-1134. doi: 10.1016/j.jmb.2007.06.002

Oates, J.; Jakobsen, B.K. (2013). ImmTACs: Novel bi-specific agents for targeted cancer therapy. *Oncoimmunology*. 1, 2(2):e22891. doi: 10.4161/onci.22891

Polydoridis, S.; Leonidas, D. D.; Oikonomakos, N. G.; Archontis G. (2007). Recognition of ribonuclease A by 3'-5'-pyrophosphate-linked dinucleotide inhibitors: A molecular dynamics/continuum electrostatics analysis. *Biophys J.*, 92(5), 1659-1672. doi: 10.1529/biophysj.106.093419

Rocchia, W.; Alexov, E.; Honig, B. (2001). Extending the applicability of the nonlinear Poisson-Boltzmann equation: Multiple dielectric constants and multivalent ions. *J. Phys. Chem. B.*, 105, 6507-6514. DOI:10.1021/JP010454Y

Rusnak, D.W.; Lackey, K.; Affleck, K. (2001). The effects of the novel, reversible epidermal growth factor receptor/ErbB-2 tyrosine kinase inhibitor, GW2016, on the growth of human normal and tumor-derived cell lines in vitro and in vivo. *Mol Cancer Ther.*, 1(2), 85-94.

Straatsma, T. P.; McCammon J. A. (1992). Computational alchemy. *Annual Review of Physical Chemistry.*, 43, 407-435. <https://doi.org/10.1146/annurev.pc.43.100192.002203>

Stone, J. E., Hardy, D. J., Isralewitz, B., y Schulten, K. (2011). Chapter 16: GPU algorithms for molecular modeling. In: *Scientific Computing with Multicore & Accelerators*, Dongarra J.; Bader J.A & Kurzak J., pp. (351-371), Chapman & Hall/CRC Press, ISBN 9781439825365.

Stumpfe, D.; Hu, Y.; Dimova, D., Bajorath, J. (2016). Recent Progress in Understanding Activity Cliffs and Their Utility in Medicinal Chemistry. *J Med Chem.*, 57: 18- 28. <https://doi.org/10.1021/jm401120g>

Suzuki, T., Motohashi, H., Yamamoto, M. (2013). Toward Clinical Application of the Keap1–Nrf2 Pathway. *Trends Pharmacol. Sci.*, 34, 340–346. doi: 10.1016/j.tips.2013.04.005

Tidor, B.; Karplus, M. (1994). The contribution of vibrational entropy to molecular association – the dimerization of insulin. *J Mol Biol.*, 238(3), 405-414. doi: 10.1006/jmbi.1994.1300

Tropsha, A. (2010). Best practices for QSAR model development, validation, and exploitation. *Molecular Informatics*, 29, 476-488. doi: 10.1002/minf.201000061

Van Gunsteren, W. F., Billeter, S. R., Eising, A. A., Hünenberger, P. H., Krueger, P., Mark, A. E., Scott, W. R. P., y Tironi, I. G. (1996). *Biomolecular Simulation: The GROMOS96 Manual and User Guide*; vdf Hochschulverlag AG an der ETH Zürich and BIOMOS b.v. Zürich, Groningen.

Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A.E.; Berendsen, H.J. (2005). GROMACS: fast, flexible, and free. *J. Comput. Chem.*, 26, 1701-1718. doi: 10.1002/jcc.20291

Wang, E.; Fu, W.; Jiang, D.; Sun, H.; Wang, J.; Zhang, X.; Weng, G.; Liu, H.; Tao, P.; Hou, T. (2021). VAD-MM/GBSA: A Variable Atomic Dielectric MM/GBSA Model for Improved Accuracy in Protein-Ligand Binding Free Energy Calculations. *J Chem Inf Model*. 28, 61(6), 2844-2856. doi: 10.1021/acs.jcim.1c00091

Wang, J.M.; Hou, T.J.; Xu, X.J. (2006). Recent advances in free energy calculations with a combination of molecular mechanics and continuum models. *Curr. Comput. Aided Drug Des.*, 2 (3), 287-306. DOI: 10.2174/157340906778226454

Wang, Z.; Sun, H.; Yao, X.; Li, D.; Xu, L.; Li, Y.; Tian, S.; Hou, T. (2016). Comprehensive Evaluation of Ten Docking Programs on a Diverse Set of Protein-Ligand Complexes: The Prediction Accuracy of Sampling Power and Scoring Power. *Phys. Chem. Chem. Phys.*, 18, 12964-12975. DOI: 10.1039/c6cp01555g

Wang, Z.; Pan, H.; Sun, H.; Kang, Y.; Liu, H.; Cao, D.; Hou, T. (2022). fastDRH: a webserver to predict and analyze protein-ligand complexes based on molecular docking and MM/PB(GB)SA computation. *Brief Bioinform.* 20, 23(5):bbac201. doi: 10.1093/bib/bbac201

Wang, Z.; Wang, X.; Li, Y.; et al (2019). farPPI: a webserver for accurate prediction of protein-ligand binding structures for small-molecule PPI inhibitors by MM/PB(GB)SA methods. *Bioinformatics*. 35, 1777-9. doi: 10.1093/bioinformatics/bty879

Weng, G.; Wang, E.; Wang, Z.; Liu, H.; Zhu, F.; Li, D.; Hou, T. (2019). HawkDock: a web server to predict and analyze the protein-protein complex based on computational docking and MM/GBSA. *Nucleic Acids Res.* 2;47(W1), W322-W330. doi: 10.1093/nar/gkz397

Wu, D.; Zheng, X.; Liu, R.; Li, Z.; Jiang, Z.; Zhou, Q.; Huang, Y.; Wu, X. N.; Zhang, C.; Huang, Y. Y.; Luo, H. B. (2022). Free energy perturbation (FEP)-guided scaffold hopping. *Acta Pharm Sin B*. 2(3), 1351-1362. doi: 10.1016/j.apsb.2021.09.027

Yang, C. Y.; Sun, H.; Chen, J.; Nikolovska-Coleska, Z.; Wang, S. (2009). Importance of ligand reorganization free energy in protein-ligand binding affinity prediction. *J Am Chem Soc.*, 131(38), 13709-13721. doi: 10.1021/ja9039373

Yu, Y.; Wang, Z.; Wang, L.; Tian, S.; Hou, T.; Sun, H. (2022). Predicting the mutation effects of protein-ligand interactions via end-point binding free energy calculations: strategies and analyses. *J Cheminform.* 20, 14(1), 56. doi: 10.1186/s13321-022-00639

Zhong, S.; Huang, K.; Luo, S.; Dong, S.; Duan, L. (2020). Improving the performance of the MM/PBSA and MM/GBSA methods in recognizing the native structure of the Bcl-2 family using the interaction entropy method. *Physical Chemistry Chemical Physics.* 22(7), 4240-4251. doi: 10.1039/c9cp06459a

Zwanzig, R. W. (1954). High-temperature equation of state by a perturbation method. I. Nonpolar gases. *Journal of Chemical Physics.* 8, 1420-1426. <https://doi.org/10.1063/1.1740409>