



UNIVERSIDAD DE JAÉN
Centro de Estudios de Postgrado

Trabajo Fin de Máster

ESTUDIO TEÓRICO DE IMIDAS TETRACÍCLICAS Y PIRAZOLONAS COMO INHIBIDORES DE K-RAS

Alumno/a: Hidalgo Ruiz, José Luis

Tutor/a: Prof. D. José Manuel Granadino Roldan

Dpto: Química Física y Analítica

Tutor/a: Prof. D. Jaime Rubio Martínez

Dpto: Química Física (Universidad de Barcelona)

Septiembre, 2016

ESTUDIO TEÓRICO DE IMIDAS TETRACÍCLICAS Y PIRAZOLONAS COMO INHIBIDORES DE K-RAS

Memoria del Trabajo Fin de Máster presentada por José Luis
Hidalgo Ruiz

Jaén, 23 de Septiembre de 2016

Fdo.: José Luis Hidalgo Ruiz



LOS TUTORES DEL TRABAJO FIN DE MÁSTER



Fdo: José Manuel Granadino
Roldán

Profesor del Departamento de
Química Física y Analítica

Fdo: Jaime Rubio Martínez

Profesor del Departamento de
Química Física (Universidad
de Barcelona)

ÍNDICE

1. RESUMEN	7
2. INTRODUCCIÓN	9
2.1. La proteína K-RAS	9
2.2. Dinámica molecular	13
2.2.1. Dinámica molecular clásica (cMD)	13
2.2.1.1. Campo de fuerzas	14
2.2.2. Dinámica molecular escalada (sMD)	15
2.3. Cálculos de energía libre MMPB/GBSA	16
3. OBJETIVOS	19
4. MATERIAL Y MÉTODOS	20
4.1. AMBER	20
4.2. Preparación del sistema	21
4.3. Preparación del ligando	22
4.4. Preparación del complejo	22
4.5. Minimización	23
4.6. Dinámica molecular	23
4.6.1. Dinámica molecular clásica	24
4.6.2. Dinámica molecular escalada	24
4.7. Análisis de componentes principales (PCA)	24
4.8. Cálculo de la energía libre de unión	25
5. RESULTADOS	26
5.1. RMSD	26
5.2. RMSF	27
5.3. Análisis de componentes principales (PCA)	27
5.4. Dinámica molecular	40
5.5. Cálculo de la energía libre de unión	45
6. DISCUSIÓN	55
7. CONCLUSIONES	57
8. BIBLIOGRAFÍA	58
9. RESUMEN DE LAS ASIGNATURAS CURSADAS	63
10. CURRÍCULUM VITAE	66

1. Resumen

La proteína K-RAS pertenece a un grupo de proteínas pequeñas de unión a GTP, conocida como la superfamilia de RAS o GTPasas de tipo RAS. Ésta, consta de 188 aminoácidos con una masa molecular de 21,6 kDa y participa en la transducción de señal intracelular. Esta proteína permanece inactiva hasta que se une a GTP, unión que provoca un cambio en la conformación de la proteína, activándola y de este modo amplificando la actividad GTPasa de la proteína 100.000 veces, aumentando así la proliferación celular de las células cancerosas. En este trabajo se ha realizado el estudio teórico de tres posibles cabezas de serie para la desactivación de la proteína K-RAS mediante simulaciones computacionales de dinámica molecular. Para ello se utiliza la suite de AMBER, que consta de un conjunto de programas para aplicar el campo de fuerza AMBER a biomoléculas, en nuestro caso.

Summary

The K-RAS protein belongs to a group of small GTP-binding proteins, known as the RAS superfamily or GTPases RAS-like. It is built by 188 amino acids with a molecular mass of 21.6 kDa and it is involved in intracellular signal transduction. This protein remains inactive until it binds to GTP, binding which causes a change in the conformation of the protein, activating it and thus amplifying GTPase activity of the protein 100,000 times, increasing cell proliferation of cancer cells. This work was carried out the theoretical study of three possible leads for the deactivation of the K-RAS protein by molecular dynamics computer simulations. For this purpose, the AMBER suite, which consists of a set of programs to implement the AMBER force field to biomolecules, in our case, is used.

2. Introducción

2.1 La proteína K-RAS

Las investigaciones de los últimos años sobre el cáncer revelan que ésta es una enfermedad que implica cambios dinámicos en el genoma. Se han establecido dos clases de genes involucrados en el cáncer, los que tienen mutaciones que producen los oncogenes con ganancia de función dominante de los genes supresores de tumores y por otra parte, los que se relacionan con la pérdida de la función recesiva [1].

El gen humano K-RAS es un homólogo del virus del sarcoma de rata Kristen y el virus del sarcoma murino, descubiertos por Chang y Der en 1982 [2]. De acuerdo con investigaciones de Der, Parada y Weinberg, se ha aceptado que K-RAS es uno de los sensores de primera línea que inician la activación de una serie de moléculas que permiten la transmisión de señales de transducción desde la superficie celular al núcleo, la señalización que afecta por lo tanto la diferenciación celular, el crecimiento, la quimiotaxis, y la apoptosis [3, 4]. Una cascada de transducción de señales iniciada por la forma activada de K-RAS se representa en la Figura 1. Como resultado de estos efectos, K-RAS provoca cambios en el citoesqueleto y consecuentemente afecta a la forma celular, la adhesión y la migración [5, 6].

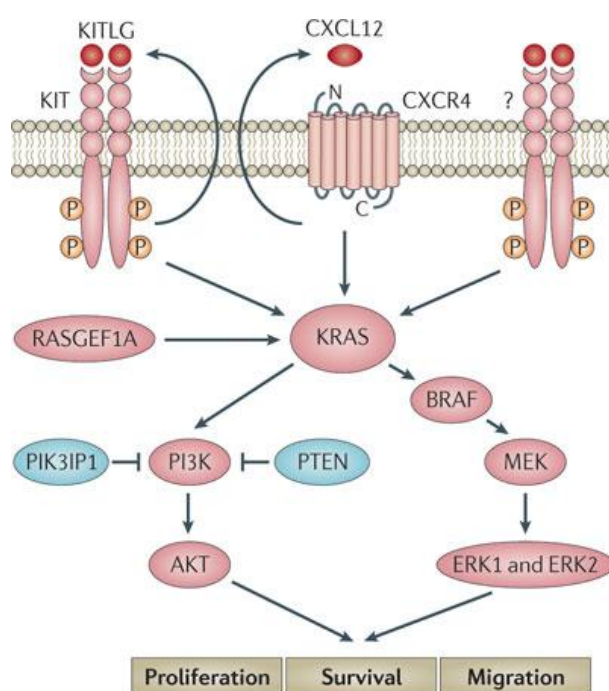


Figura 1. Vía de señalización de la proteína K-RAS [7].

K-RAS pertenece a un grupo de proteínas pequeñas de unión a GTP, conocida como la superfamilia de RAS o GTPasas de tipo RAS. Han sido identificados más de 150 genes de tipo RAS en los genomas de mamíferos [8]. Toda la superfamilia RAS se caracteriza por

la presencia de un dominio G catalítico, pero incluye miembros con especializaciones evolutivas distintas con respecto a diferentes procesos celulares [9]. La subfamilia RAS (RAS, RHO, RAB, ARF, RAC y RAN) incluye las proteínas estudiadas con mayor frecuencia, como Harvey-Ras (H-RAS), neuroblastoma-Ras (N-RAS), y dos variantes de Kristen-RAS (K-RAS), una conocida como K-RAS4A, que se expresa débilmente en células humanas, y la forma dominante, conocida como K-RAS4B, que se expresa mucho más frecuentemente.

El producto del gen K-RAS, la proteína K-RAS, consta de 188 aminoácidos con una masa molecular de 21,6 kDa y participa en la transducción de señal intracelular [10]. Esta proteína permanece inactiva hasta que se une a GTP. Esta unión provoca un cambio en la conformación de la proteína, que implican dos regiones de la proteína, activándola. Estas dos regiones importantes se conocen como *switch 1* (aminoácidos 30-38) y *switch 2* (aminoácidos 59-67), que forman una cascada de efectos, controlando la especificidad de unión de esta GTPasa a sus moléculas efectoras. Este cambio conformacional en la proteína K-RAS afecta sus interacciones con múltiples transductores que se encuentran aguas abajo –proteínas activadoras de GTPasa (GAPs)- y como consecuencia amplifican la actividad GTPasa de la proteína RAS 100.000 veces [11]. Dicho cambio también afecta a las interacciones con los factores de intercambio y liberación de guanina (GEFs/GRFs) promoviendo la liberación de GTP [12]. Las mutaciones de la forma oncogénica de la proteína p21 de RAS perjudican la actividad GTPasa y hacen que la proteína K-RAS no responda a las proteínas GAP. Las formas mutadas de p21 intercambian rápidamente GDP por GTP, que se prefiere como sustrato, lo que induce el estado activo. Tales formas aberrantes de la proteína K-RAS desregulan muchos efectores, lo que afecta a varias vías celulares importantes.

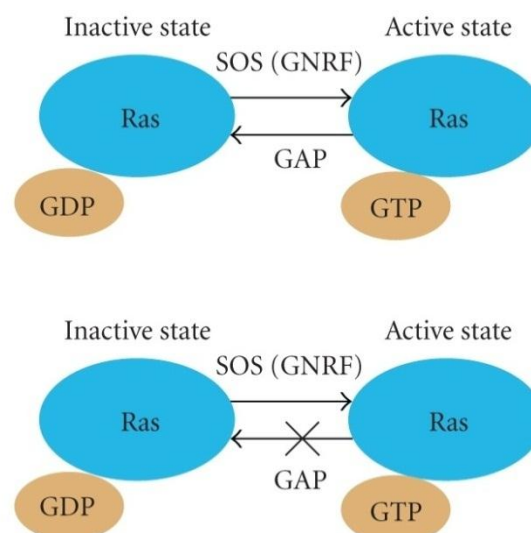


Figura 2. Estados de actividad de la proteína K-RAS. El cambio mutacional en el exón 1 del gen K-RAS lleva a la permanente estado "on".

K-RAS está formado por cuatro dominios. El primer dominio incluye 85 aminoácidos en el extremo N-terminal y es idéntico en las tres formas de RAS (K-RAS, NRAS y HRAS). El segundo dominio contiene 80 aminoácidos, con una identidad de secuencia inferior (70-80%) entre las tres formas de la proteína RAS. Estas regiones son importantes para la función de señalización de la proteína K-RAS y forman conjuntamente el dominio G (aminoácidos 1-165, Figura 3). El dominio G de la proteína K-RAS incluye el *pocket* de unión a GTP, donde los bucles de unión *loop*-fosfato P (aminoácidos 10-16 y 56-59) interactúan con el b-fosfato y c-fosfato de GTP. Los residuos 116-119 y 152-165 interactúan con la base guanina. La región entre los aminoácidos 32 y 40 (zona del núcleo efector) es esencial para las interacciones entre los efectores putativos y BPA. La proteína RAS también contiene una región hipervariable (HVR) en el extremo C-terminal (aminoácidos 165-188/189; el tercer dominio), que guía la modificación postraducciona y determina el anclaje a la membrana plasmática, por tanto esta región tiene un papel importante en la regulación de la actividad biológica de la proteína RAS [14].

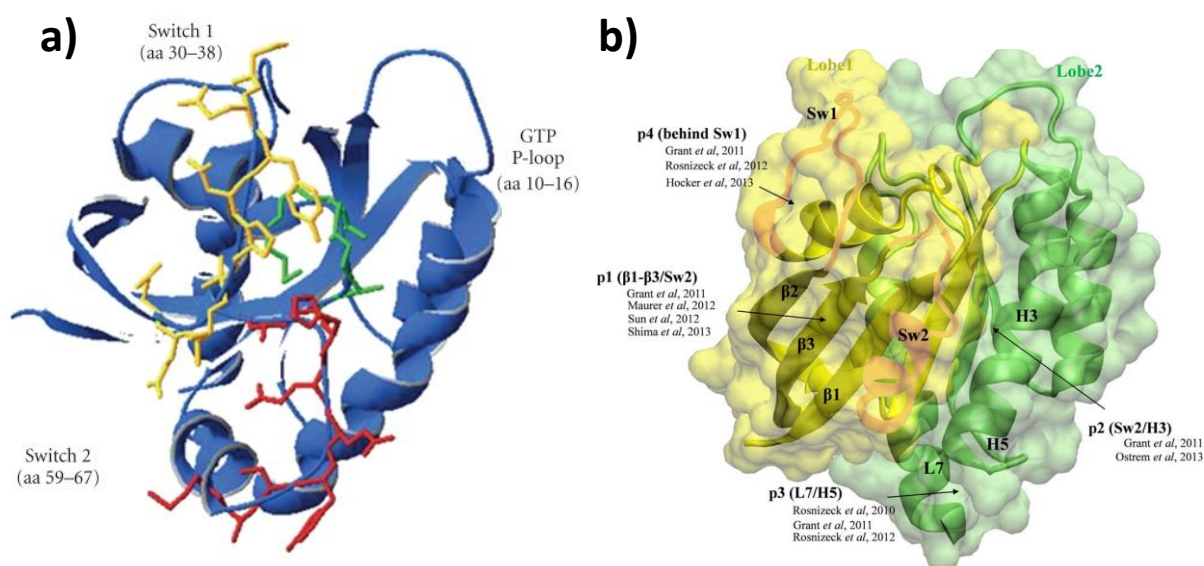


Figura 3. a) Modelo de la proteína K-RAS con dominios importantes resaltados. Modelo de la molécula de 3GFT hecha en Swiss-PdbViewer v4.0.1 (<http://spdbv.vital-it.ch/>). **b)** El dominio catalítico de K-Ras es bi-lóbulo y contiene cuatro sitios de unión a ligandos alostéricos (p1, p2, p3 y p4). El dominio catalítico de K-Ras se muestra en amarillo, el lóbulo 1 (residuos 1-86) y en verde el lóbulo 2 (residuos 87-167) [15].

Los *switches* I y II juegan un papel importante en la unión de los reguladores y efectores. El *loop pocket-P* de unión a fosfato permite la unión temporal de GTP a la proteína RAS. Esta es también la región de la actividad GTPasa, que regula negativamente la proteína RAS a través de la reacción de hidrólisis RAS-GTP y la unión de difosfato de guanosina.

La proteína K-RAS actúa como un interruptor molecular localizado en la membrana plasmática, regulando las múltiples vías de transducción de señales [16]. Se sintetiza en el

citosol, donde es farnesilada por la farnesil transferasa en el residuo de cisteína del motivo carboxi-terminal CAAX (donde C representa cisteína, A es un aminoácido alifático, y X es cualquier aminoácido). El motivo carboxi-terminal forma el último dominio de la proteína K-RAS. El motivo de aminoácidos AAX se escinde por proteasas, mientras que el residuo carboxilo C-terminal de la proteína K-RAS está metilado. La escisión del motivo peptídico AXX y metilación se produce en la superficie citosólica del retículo endoplasmático y está mediado por la enzima convertidora de RAS Rce1 [17]. La farnesilación C-terminal juega un importante papel en la localización de la membrana. En el K-RAS4A, el motivo AXX sufre palmitoilación adicional por palmitoil transferasa, resultando en la orientación adecuada de K-RAS4A a la membrana. Sin embargo, no hay palmitoilación detectable de la predominante K-RAS4B, la cual probablemente llegue a la membrana plasmática a través de un mecanismo dependiente de microtúbulos, evitando así el aparato de Golgi [16, 18]. Se cree que la farnesilación postraduccional y carboximetilación son importantes para la oncogenicidad de la proteína RAS. Se ha demostrado que el tratamiento con inhibidores de la farnesil transferasa inhibe el crecimiento independiente de anclaje tanto en fibroblastos de ratones de ambos con K-RAS transformado como en células tumorales humanas que contienen mutaciones K-RAS y ANR. La transducción de señales de la proteína K-RAS no se produce exclusivamente en la membrana plasmática. La activación de vías de señalización corriente abajo de K-RAS también puede ser activada por señales procedentes de los compartimentos subcelulares, tales como el retículo endoplasmático y el aparato de Golgi [19, 20].

¿Por qué la proteína K-RAS es un importante objeto de estudio?

A menudo se producen mutaciones que activan K-RAS en el proceso carcinogénico. [21]. Una vez que el gen K-RAS muta, adquiere propiedades oncogénicas y parece estar causalmente implicado en el desarrollo de varios cánceres humanos [22, 23]. Se ha observado la pérdida del alelo de K-RAS de tipo salvaje tanto en tumores humanos como de ratón, lo que indica que la ausencia del alelo normal puede facilitar la transformación de una copia del alelo oncogénico de K-RAS [24]. Las mutaciones oncogénicas en el gen K-RAS previenen la hidrólisis de GTP, por tanto, activan las moléculas RAS de forma permanente [25]. Se ha demostrado que la expresión de un gen K-RAS mutante en fibroblastos aumenta la expresión de la metaloproteína 2 (MMP2) en la matriz y mejora la invasión de células del cáncer [26]. La sobreexpresión de esta forma mutada de K-RAS también inhibe la glicosilación de la integrina $\beta 1$, dando como resultado la polarización alterada y aumento de la adhesividad de las células de cáncer de colon. Además, se ha demostrado que la expresión de esta forma oncogénica de la proteína K-RAS se asocia con la sobreexpresión

del antígeno carcinoembrionario (CEA) y con la perturbación de la polarización de las células epiteliales [27].

2.2. Dinámica molecular

La dinámica molecular es una herramienta computacional basada en la Química, Física y Matemáticas para simular, describir y predecir sistemas de comportamiento y las propiedades de los átomos, las moléculas más grandes o en condiciones reales a través de programas de ordenador.

La dinámica molecular consiste, en nuestro caso, en el cálculo de la evolución de varios complejos ligando-proteína a lo largo del tiempo. Para lograr este objetivo, se resuelven las ecuaciones de Newton acopladas para todo el sistema. Ya que sabemos las condiciones iniciales de cada átomo (posición, velocidad y aceleración) y teniendo en cuenta que la dinámica molecular es un método determinista, el estado del sistema se puede predecir en cualquier momento tanto del futuro y como del pasado [28].

2.2.1. Dinámica molecular clásica (cMD)

Los cálculos comienzan con un primer conjunto de coordenadas para el sistema, que se obtiene generalmente a partir de datos de rayos X o RMN, y un conjunto inicial de velocidades que satisfacen la distribución de Maxwell-Boltzmann (Ecuación 1) para la temperatura objetivo, donde la temperatura y las velocidades se relacionan por la Ecuación 2.

$$v_i = \left(\frac{m_i}{2\pi k_B T} \right)^{\frac{1}{2}} \cdot e^{\left(-\frac{m_i \cdot v_i^2}{2k_B T} \right)}$$

Ecuación 1. Distribución de Maxwell-Boltzmann.

$$T = \frac{2}{3k_B} \langle E_k \rangle = \frac{m}{3k_B} \langle v^2 \rangle$$

Ecuación 2. Relación entre la temperatura absoluta y su energía cinética.

Donde m_i es la masa de la partícula, k_B es la constante de Boltzmann, T se refiere a la temperatura, v a la velocidad de la partícula y E_k la energía cinética.

Las nuevas coordenadas y velocidades se obtienen resolviendo las ecuaciones de Newton (Ecuación 3) mientras que la temperatura se mantiene constante mediante el método de Langevin [29].

$$\frac{d\vec{v}_i(t)}{dt} = \frac{\vec{F}_i(t)}{m_i} \quad \text{siendo} \quad F_i(t) = - \frac{\delta V(\vec{r}_1(t), \vec{r}_2(t), \dots, \vec{r}_n(t))}{\delta \vec{r}_i(t)}$$

Ecuación 3. Ley de Newton.

Donde i , incluye todos los átomos del sistema ($i = 1, 2, 3, \dots, n$), siendo n el número total de átomos. Las coordenadas atómicas, r_i , y la velocidad, v_i , del átomo, i , con la masa, m_i , se convierten en funciones del tiempo.

El modelo matemático utilizado (campo de fuerzas), se describe cómo el valor de la energía potencial dependiente de la disposición espacial de todos los átomos [30].

2.2.1.1. Campo de fuerzas

El cálculo de la energía de interacción de un sistema molecular requiere un campo de fuerzas, que consiste en dos componentes: un conjunto de funciones que se utilizan para generar las energías potenciales y los parámetros usados en estas funciones. Una ecuación de campo de fuerzas habitual utilizado en la mecánica molecular se muestra en la ecuación 4.

$$U = \sum_{bonds} U^{srt} + \sum_{angles} U^{bend} + \sum_{torsions} U^{for} + \sum_{i,j} U_{i,j}^{LJ} + \sum_{i,j} U_{i,j}^{Coulomb}$$

Ecuación 4. Ecuación generalizada del campo de fuerzas.

Donde los tres primeros términos se corresponden con los términos de enlace y los dos últimos corresponden a términos no enlazantes, que son el término de Lennard-Jones y las contribuciones de Coulomb. El primero de estos términos no enlazantes describe la interacción de van der Waals y el segundo las interacciones electrostáticas [31].

Esta ecuación se puede desarrollar en sus términos y se obtiene la ecuación siguiente:

$$\begin{aligned} U = & \sum_{bonds} \frac{1}{2} k_b (r - r_0)^2 \\ & + \sum_{angles} \frac{1}{2} k_a (\theta - \theta_0)^2 \\ & + \sum_{torsions} \frac{V_n}{2} [1 + \cos(n\phi - \delta)] + \sum_{LJ} 4 \epsilon_{i,j} \left(\frac{\sigma_{i,j}^{12}}{r_{i,j}^{12}} - \frac{\sigma_{i,j}^6}{r_{i,j}^6} \right) + \sum_{Coulomb} \frac{q_i q_j}{r_{ij}} \end{aligned}$$

Ecuación 5. Ecuación generalizada del campo de fuerzas desarrollada.

Como se muestra en la ecuación 5, los términos de energía de enlace de *stretching* (*bonds*) y de energía de enlace angular (*angles*) se representan como una funciones armónicas simples que controlan la longitud de los enlaces covalentes donde r y θ se refieren a los átomos. La energía de torsión (*torsions*) se suele representar por una función

coseno donde ϕ es el ángulo de torsión, δ es la fase, n define el número de mínimo o máximo entre 0 y 2π y V_n determina la altura de la barrera de potencial a la torsión.

Las interacciones de van der Waals entre dos átomos surgen del equilibrio entre las repulsiones y atracciones efectivas. Para representar estas interacciones es común utilizar la ecuación 12-6 de Lennard-Jones (LJ). Por otro lado, el término final de la ecuación se utiliza para describir las interacciones electrostáticas. Generalmente se asigna una carga parcial a cada núcleo y se usa la ley de Coulomb para calcular su contribución a la energía total [32].

2.2.2. Dinámica molecular escalada (sMD)

La dinámica molecular clásica se utiliza para crear una trayectoria de movimiento molecular para todo el sistema. Cuando la simulación se lleva a cabo durante un tiempo infinito, es posible alcanzar condiciones de equilibrio del sistema independientemente del punto de partida. Sin embargo, incluso con los ordenadores modernos, es imposible llevar a cabo simulaciones de biomoléculas complejas por períodos tiempo suficientemente amplios como para observar este comportamiento. Este problema se produce debido que las escalas de tiempo en que algunos movimientos moleculares grandes tienen lugar son demasiado largas, del orden de milisegundos o segundos, mientras que el tiempo de integración de las ecuaciones de movimiento que describen la dinámica molecular son del orden de femtosegundos.

El teorema ergódico, demostrado simultáneamente por George D. Birkhoff y John von Neumann, permite concluir notables resultados sobre el comportamiento medio del sistema durante largos periodos de tiempo [33]. De este modo, Sinko *et al.* [34] teniendo en cuenta este teorema propusieron una modificación de la superficie de energía potencial (PES, de sus siglas en inglés) para aplanar y disminuir la barrera de potencial entre los diferentes estados de transición, ya que el ligando puede quedar atrapado en un mínimo de energía que no es realmente el mínimo absoluto de la dinámica. De este modo, el ligando no puede explorar todo el espacio conformacional, y por tanto no se podría cumplir el teorema de ergodicidad antes comentado.

La PES se modifica escalando la energía potencial $V(\vec{r})$ por un factor λ , que varía entre 0 y 1 (Ecuación 6).

$$V^*(\vec{r}) = \lambda V(\vec{r})$$

Ecuación 6. Modificación de la superficie de energía potencial en sMD.

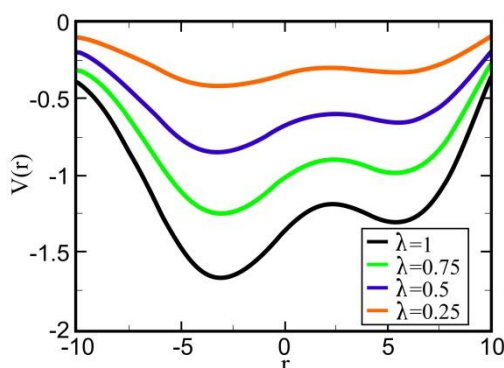


Figura 4. Ilustración esquemática una SMD cambiando el factor λ .

2.3. Cálculos de energía libre MMPB/GBSA

Las energías de mecánica molecular (MM) combinadas con los métodos de la ecuación de Poisson-Boltzmann (PB) o la de Generalized Born (GB) y la superficie de solvatación de zona continua (SA) (MMPBSA y MMGBSA) son enfoques para estimar la energía libre de la unión de pequeños ligandos a macromoléculas biológicas. Estos cálculos se basan en simulaciones de dinámica molecular del complejo receptor-ligando y por lo tanto están a medio camino entre los cálculos empíricos y los métodos de perturbación alquímica [35].

En este método, la energía de unión libre, ΔG_{bind} , se estima a partir de las energías libres de los reactivos (complejo proteína-ligando, a partir de ahora C) y los productos (ligando y receptor, a partir de ahora L y R, respectivamente) de la reacción.

$$\Delta G_{bind} = \langle G_C \rangle - \langle G_L \rangle - \langle G_R \rangle$$

Ecuación 7. Ecuación de la energía libre de unión estimada a partir de las energías libres de los reactivos.

Donde la energía libre G de cada estado C, L y R se calcula mediante la siguiente suma

$$G = E_{bond} + E_{el} + E_{vdW} + G_{pol\ solv} + G_{np\ solv} - TS_{vib}$$

Ecuación 8. Suma de la energía libre de cada estado.

Donde los tres primeros términos son términos energéticos de mecánica molecular estándar de enlace (*bond*, *angles* y *torsions*) (Ecuación 4 y Ecuación 5), interacciones electrostáticas y de van der Waals. La contribución media de la energía libre de solvatación, ΔG_{solv} , está formada por la suma de las contribuciones polares, G_{pol} , y no polares, G_{np} . El término G_{pol} se obtiene resolviendo la ecuación de Poisson-Boltzmann (Ecuación 9) o utilizando el modelo de Generalized Born (Ecuación 10 y Ecuación 11). A su vez, el término

G_{np} se calcula utilizando la ecuación $G_{np} = \alpha SASA + \beta$, suponiendo pues una relación lineal con la superficie accesible al disolvente (SASA), que asigna a cada átomo una zona esférica con un radio determinado. El último término es la temperatura, T , multiplicada por la entropía, S , que se estima por un análisis en modos normales de las frecuencias de vibración.

La aproximación de Poisson Boltzman (PB) describe las interacciones electrostáticas de una proteína (solute) en un disolvente continuo. El soluto se representa como un modelo atómico usando un campo de fuerzas de mecánica molecular, mientras que el disolvente se trata como un continuo sin estructura. El tratamiento continuo representa el soluto como un cuerpo dieléctrico cuya forma está definida por las coordenadas atómicas y los radios de las cavidades atómicas. El soluto, por algunas cargas puntuales, produce un campo electrostático en la región del disolvente. Este campo electrostático puede ser calculado resolviendo la ecuación de PB, mostrada a continuación:

$$\nabla[\epsilon(r)\nabla\phi(r)] = -4\pi\rho(r) - 4\pi\lambda(r) \sum_i z_i c_i \left(\frac{-z_i \phi(r)}{k_B T} \right)$$

Ecuación 9. Ecuación de Poisson-Boltzmann.

Donde $\epsilon(r)$ es la constante dieléctrica, $\phi(r)$ es el potencial electrostático, $\rho(r)$ es la densidad de carga de soluto, $\lambda(r)$ es la función de capa de enmascaramiento Stern, z_i es la carga del ion i , c_i es la densidad del número mayor del ion i , k_B es la constante de Boltzmann, y T es la temperatura; la suma se realiza sobre todos los tipos de iones diferentes.

El modelo de Generalized Born (GB) es una aproximación de la ecuación de PB. Amber tiene muchos modelos GB (ocho hasta ahora), que no sólo difieren ligeramente en su representación matemática sino también en su parametrización. En general, se trata a cada átomo como una esfera de radio R_i con una carga q_i en su centro, relleno por un material con una constante dieléctrica igual a 1. Por otro lado, el sistema estudiado molécula está rodeado por un disolvente de un alto ϵ dieléctrico (80 para el agua a 300 K). El modelo GB utilizado en este trabajo, aproxima G_{pol} de acuerdo con la siguiente ecuación [36],

$$G_{Pol} = -\frac{1}{2} \sum_{ij} \frac{q_i q_j}{f_{GB}(r_{ij}, R_i, R_j)} \left(1 - \frac{\exp[-\kappa f_{GB}]}{\epsilon} \right)$$

Ecuación 10. Aproximación de G_{pol} mediante el modelo Generalized Born.

Donde r_{ij} es la distancia entre los átomos i y j , R_i y R_j son los radios de las moléculas efectivas de Born, y $f_{GB}()$ es la función determinada de sus argumentos. Finalmente, los

efectos de *screening* electrostáticos de la sal (monovalente) se incorporan a través del parámetro de *screening* de Debye-Hückel. Por lo tanto, una aproximación común de f_{GB} es:

$$f_{GB} = \left[r_{ij}^2 + R_i R_j \left(-\frac{r_{ij}^2}{4R_i R_j} \right) \right]^{\frac{1}{2}}$$

Ecuación 11. Aproximación de f_{GB} .

3. Objetivos

El objetivo principal de este trabajo es el estudio, usando diferentes métodos, de cómo la proteína K-RAS interactúa con los diferentes ligandos propuestos. De esta manera se intentará identificar el lugar donde la unión es más probable para cada uno de los ligandos. Estos conocimientos, permitirán saber cuál compuesto es más efectivo para impedir la activación de la proteína, y por tanto detener la proliferación celular de las células cancerosas.

Para alcanzar ese objetivo, se trabaja en los diferentes sistemas de K-RAS con cada ligando. Para ello, los ligandos se fijan a una distancia determinada en diferentes posiciones alrededor de la proteína, y se observa el comportamiento del sistema y se analiza en qué ocasiones se termina formando un complejo estable y cómo sucede esto. Esta distancia es reducida para que el complejo y la proteína puedan interactuar rápido y no necesiten demasiado tiempo, por tanto los cálculos de las dinámicas serán más sencillos. Debido al hecho de que los cálculos de dinámica molecular convencional no pueden mostrar los eventos que requieren un tiempo de muestreo de largo, se realizan cálculos utilizando dinámica molecular escalada, de esta manera se acelera el procesamiento, por lo que se puede observar si el sistema es viable en una escala de tiempo reducido.

Se estudian finalmente los datos de estos cálculos con el fin de analizar el comportamiento de los complejos K-RAS-ligando en las diferentes dinámicas moleculares, y establecer los diferentes sitios de interacción que la proteína podría tener.

4. Material y métodos

4.1. AMBER

AMBER (*Assisted Model Building with Energy Refinement*) es el nombre del paquete de software que simula unos campos de fuerza de dinámica molecular desarrollados originalmente por el grupo de Peter Kollman en la Universidad de California, San Francisco [37]. Más concretamente en nuestro caso, se utiliza el campo de fuerzas GAFF (*Generalized Amber Force Field*), para la preparación de los ligandos, que es compatible con el campo de fuerzas AMBER, ya que puede ser utilizado para proteínas y ácidos nucleicos, y tiene parámetros para la mayoría de moléculas orgánicas compuestas por H, C, N, O, S, P y halógenos. Éste utiliza una forma funcional simple y un número limitado de tipos de átomos, pero incorpora los modelos empírico y heurístico para estimar constantes de fuerza y cargas atómicas parciales [38].

Por otro lado, también se usa el campo de fuerzas ff14SB para la proteína, un campo de fuerzas con los parámetros diedros de los aminoácidos de cadena lateral reajustados a partir del campo de fuerzas ff99SB [39].

La suite de software de AMBER proporciona un conjunto de programas para aplicar el campo de fuerza AMBER a biomoléculas. Está escrito en los lenguajes de programación Fortran 90 y C, con soporte para la mayoría de los sistemas operativos.

Trabaja con el siguiente flujo de trabajo:

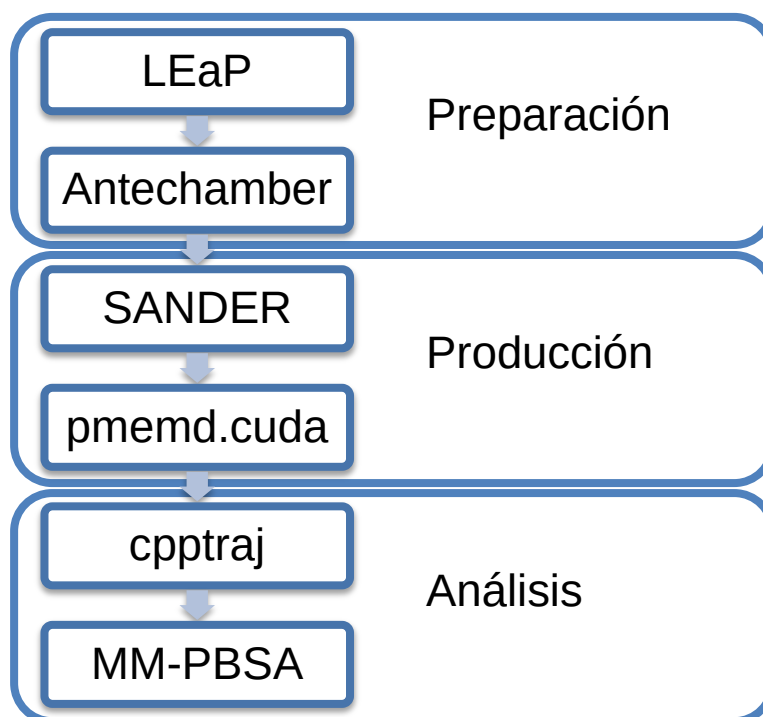


Figura 5. Flujo de trabajo de AMBER.

- LEaP prepara los archivos de entrada para los programas de simulación.
- Antechamber automatiza el proceso de parametrización de moléculas para las cuales no hay parámetros de campo de fuerzas AMBER usando el campo de fuerzas GAFF.
- SANDER (*Simulated Annealing with NMR-Derived Energy Restraints*) es el programa central de simulaciones y se utiliza en el presente trabajo tanto para el proceso de minimización del complejo como el calentamiento del mismo (ver epígrafe 4.6).
- pmemd es una reimplementación de SANDER, diseñado para la computación en paralelo, y se comporta significativamente mejor que SANDER cuando se ejecuta en más de 8-16 procesadores.
 - pmemd.cuda ejecuta las simulaciones en tarjetas gráficas (GPUs) en lugar de en CPUs.
- ptraj analiza los resultados de la simulación numéricamente.
 - cpptraj es una nueva versión de ptraj escrita en C++ para analizar los resultados de la simulación más rápidamente.
- MM-PBSA.py automatiza el cálculo de energía libre de unión a partir de una dinámica molecular según la metodología MMPB(GB)SA.

4.2. Preparación del sistema

Se utilizó el Protein Data Bank (<http://www.rcsb.org>) para obtener la estructura inicial de la proteína K-RAS en su forma activa con el GTP unido a ella (Identificador de PDB: 4DSN, [40]). Se trata de una proteína formada por 189 aminoácidos.



Figura 6. K-RAS.

Esta proteína no necesitó preparación adicional, como ocurre en otros casos en los que ciertas partes de la estructura no pueden ser resueltas, o que necesitan intercambiar algunos aminoácidos para poder cristalizar la misma. El sistema estudiado está formado por una proteína y un ligando, de modo que mientras los parámetros del campo de fuerzas de los aminoácidos que componen la proteína ya están contenidos en la base de datos, por lo tanto, es necesario obtener los parámetros del ligando.

4.3. Preparación del ligando

Los ligandos se dibujan con el programa Avogadro [41], el cual permite la elaboración y además la minimización de las estructuras para adaptar la molécula, de modo que desaparezcan posibles impedimentos estéricos entre átomos, usando el método semiempírico AM1.

Para preparar el sistema, se utiliza el programa de AmberTools 15 de la suite Amber14.

Para obtener los parámetros del ligando (conectividad, geometría, tipo de átomos, cargas, etc.) se usa Antechamber, que asigna los parámetros de enlace necesarios por comparación y calcula los no enlazantes.

4.4. Preparación del complejo

Para preparar el complejo, se utiliza Leap, que es un programa que permite solvatar y neutralizar el sistema, la adición de iones de sodio o cloruro siempre que sea necesario, y también para comprobar que el sistema final utilizado tiene asignados los correspondientes parámetros del campo de fuerzas para todos sus átomos. Esta aplicación crea dos archivos: uno contiene las coordenadas de todos los átomos del sistema, mientras que el otro tiene toda la información necesaria para calcular su energía potencial (esto es, los parámetros del campo de fuerzas para cada uno de los átomos que lo componen). En la mecánica clásica los enlaces no pueden romperse ni crearse, por lo tanto, la topología del sistema no va a cambiar con el tiempo, pero las coordenadas del sistema sí cambian.

Inicialmente, los complejos se introducen en una caja cúbica de moléculas de agua TIP3P [42] con una distancia máxima entre la proteína y el borde de la caja de 15 Å. Se elimina cualquier molécula de agua que se encuentre a menos de 2,2 Å de cualquier átomo del ligando o de la proteína. Finalmente, se añaden los contraiones necesarios con el fin de neutralizar el sistema. Se usan condiciones periódicas de contorno en todas las simulaciones con el método de la partícula de malla Ewald [43] para calcular las

interacciones electrostáticas de largo alcance. Se elige una distancia de corte de 10 Å para el cálculo de las interacciones electrostáticas y de van der Waals (vdW).

4.5. Minimización

El término minimización se refiere a relajar energéticamente las geometrías. Estas minimizaciones energéticas se llevan a cabo utilizando el programa SANDER de la suite AMBER 14. La minimización se realiza utilizando un procedimiento de múltiples etapas, con el fin de eliminar todos los posibles malos contactos, que podrían dar lugar posteriormente a una dinámica molecular con una energía potencial infinita, y relajar progresivamente el sistema: (1) en primer lugar se minimiza sólo la proteína sin tener en cuenta moléculas de agua e iones a través de 1000 pasos usando el algoritmo *steepest descent*, con una restricción de energía constante de 5,0 Kcal/mol Å; (2) a continuación, el sistema es minimizado teniendo en cuenta toda la proteína, manteniendo congelado el resto del sistema, en otros 1000 pasos de nuevo utilizando el algoritmo *steepest descent* con la misma restricción de 5,0 Kcal/mol Å; y finalmente (3) todo el sistema se minimiza sin restricciones a través de 10000 pasos mediante el algoritmo *steepest descent*.

4.6. Dinámica molecular

El sistema se prepara en condiciones de energía cinética cero, por lo tanto, las moléculas no tienen velocidad, solamente energía potencial, lo que significa que se requiere una etapa de acondicionamiento. Dicho acondicionamiento consiste en calentar el sistema desde 0 a 300K a volumen constante para, a continuación, permitir que la densidad del sistema se ajuste de modo que se aproxime a la del agua ejecutando pasos de dinámica molecular en condiciones NPT (esto es, número de átomos, presión y temperatura constantes).

El calentamiento se realiza gradualmente en pasos. Una vez finalizado este proceso, se lleva a cabo el segundo paso, donde se aumenta la densidad del sistema. El último paso del condicionamiento es el punto de partida de la posterior dinámica molecular.

Todas las producciones se llevan a cabo a una temperatura constante de 300 K mediante el acoplamiento del sistema a un baño térmico usando un algoritmo de termostato de Langevin [44, 45] con una frecuencia de colisión de 5 ps⁻¹. Se establece un tiempo de integración de 1 fs durante la adaptación del sistema, y de 2 fs durante la producción, reduciendo a la mitad el tiempo necesario, con las longitudes de enlace de los átomos de hidrógeno limitadas a las longitudes de equilibrio con el algoritmo SHAKE [46].

El proceso de equilibrio comienza por el calentamiento de las estructuras minimizadas a 300 K durante un período de 200 ps a una velocidad constante de $1,5 \text{ K ps}^{-1}$ en condiciones NVT (número de átomos, volumen y temperatura constantes). Después del calentamiento, el equilibrio tiene lugar en 200 ps bajo condiciones NPT (1 atm, 300 K) para aumentar la densidad del sistema, utilizando un algoritmo para mantener la presión constante análogo al de la temperatura de acoplamiento [47]. Por último, el paso de producción se lleva a cabo durante un tiempo de 100 ns a una temperatura constante de 300 K bajo condiciones NVT.

La tarjeta gráfica utilizada para realizar los cálculos de dinámica molecular es una GeForce GTX TITAN X, y para la realización de estos cálculos se utiliza el programa pmemd.cuda, de la suite Amber 14.

4.6.1. Dinámica molecular clásica

Las dinámicas realizadas siempre comienzan con la estructura del sistema inicial preparada con LEaP. Los ligandos se colocan con un programa escrito en Fortran a una distancia de 30 Å respecto de un plano formado por los tres carbonos α de los aminoácidos Lys 176, Lys 177 y Lys 178. Se hace un acondicionamiento del sistema inicial para cada uno de ellos y luego sus respectivos cálculos de dinámica molecular convencional. Como hemos utilizado diferentes velocidades iniciales, será posible ver la evolución del sistema y discernir si hay un patrón común entre ellos.

Se realizan diez cálculos de dinámica molecular de 100 ns para cada uno de los tres ligandos.

4.6.2. Dinámica molecular escalada

Se llevan a cabo cinco cálculos de dinámica molecular de 50 ns cada uno, para cada uno de los tres ligandos, con una constante escalar de 0,8 ($\lambda = 0,8$). Generalmente, los resultados muestran que la dinámica molecular es demasiado rápida y no se da una buena interacción ligando-proteína.

4.7. *Análisis de componentes principales (PCA)*

El análisis de los perfiles de energía libre (ΔG) de los complejos se realiza con el fin de estudiar la evolución del sistema y analizar las diferentes conformaciones que puede explorar el complejo proteína-ligando a partir de las diferentes estructuras de partida.

Para ello, se transforman las coordenadas (xyz) de los tres sistemas, en dos componentes principales (PC1 y PC2) con un punto de coordenadas común entre ellos

como referencia. Para cada dinámica se eligen los residuos de la proteína que no se mueven excesivamente a lo largo de la misma, según el análisis RMSF (ver epígrafe 5.2). Así, en general, los residuos elegidos son del 1 al 168, de modo que, como es común, los residuos finales expuestos al disolvente se eliminan del cálculo en componentes principales.

A continuación, el espacio se define con una matriz de covarianza $3N \times 3N$, donde N representa los átomos $C\alpha$ de la secuencia de residuos de 1 a 168, en la que el análisis se centra respecto de la estructura principal de K-RAS. De este modo se pasa de $3N$ coordenadas a dos coordenadas sin perder demasiada cantidad de información.

Por último, se utilizan las coordenadas del PCA para calcular el perfil de energía libre (ΔG) con el fin de determinar en términos de energía qué puntos de la dinámica molecular tienen menor energía. Los puntos que tengan menor energía serán los puntos por los que el ligando haya pasado un mayor número de veces. De esta manera se construye el gráfico, que representa (ΔG) a través de las coordenadas principales PC1 y PC2, en una escala de color, en la cual el color negro indica la energía mínima relativa ($\Delta G=0$), mientras que el color amarillo indica la máxima energía libre.

4.8. Cálculo de la energía libre de unión

Las energías libres de unión se calculan utilizando las metodologías MMPBSA y MMGBSA incluidas en el paquete Amber 14 mediante el programa mmpbsa.py.

Los valores de las constantes dieléctricas interiores y exteriores se eligieron como 1 y 80, respectivamente. Por otra parte, la aportación no polar G_{np} se estimó utilizando el método LCPO para calcular SASA. Los valores para las constantes α y β se tomaron como $0,005420 \text{ kcal/mol } \text{\AA}^2$ y $0,9200 \text{ kcal/mol}$ en el cálculo MMPBSA y $0,0072 \text{ kcal/mol } \text{\AA}^2$ y $0,0 \text{ kcal/mol}$ en el cálculo MMGBSA. Para el método GB se utiliza la aproximación Onufriev (igb 5) [48].

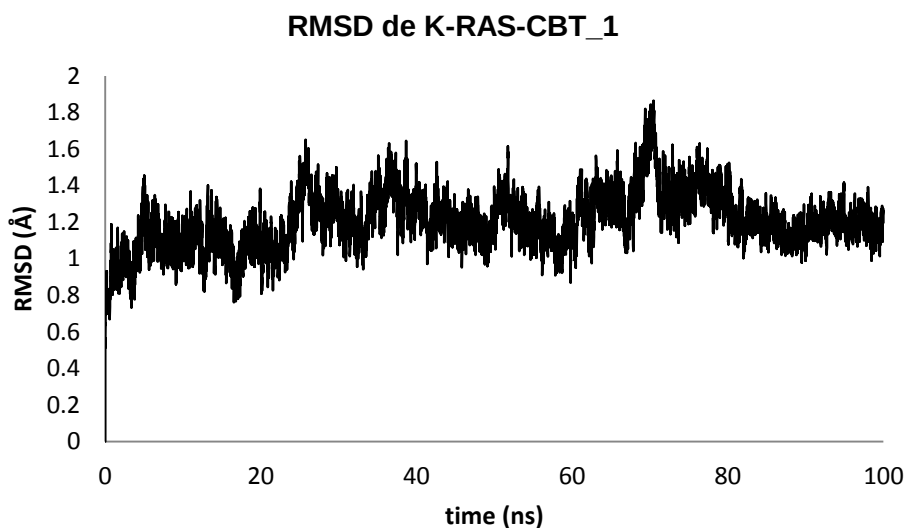
5. Resultados

El complejo proteína-ligando, como ya se ha explicado anteriormente, tenderá, en un tiempo infinito, a encontrar el equilibrio. Para encontrar dicho equilibrio, el ligando se unirá en el sitio de unión que dé como resultado un complejo con la menor energía libre posible, ya que es el estado más estable.

5.1. RMSD

La estabilidad de la estructura de la proteína se caracteriza adicionalmente mediante el cálculo de la desviación cuadrática media (RMSD de sus siglas en inglés) entre las estructuras de la correspondiente simulación y la estructura de partida. Para obtener el RMSD de la estructura se incluyen los residuos 1-170 y se utilizan sólo los átomos C α . La región excluida, el *loop* C-terminal (Ver figura 7), no tiene una estructura definida, por lo que previsiblemente a lo largo de la dinámica los aminoácidos de esta región experimentarán grandes cambios.

Como se puede ver en el ejemplo de la gráfica 1 (dinámica molecular clásica número 1 del ligando CBT) las desviaciones entran dentro de la normalidad, ya que en todos los casos la distancia varía entre 0,8 y 2,0 Å como mucho.

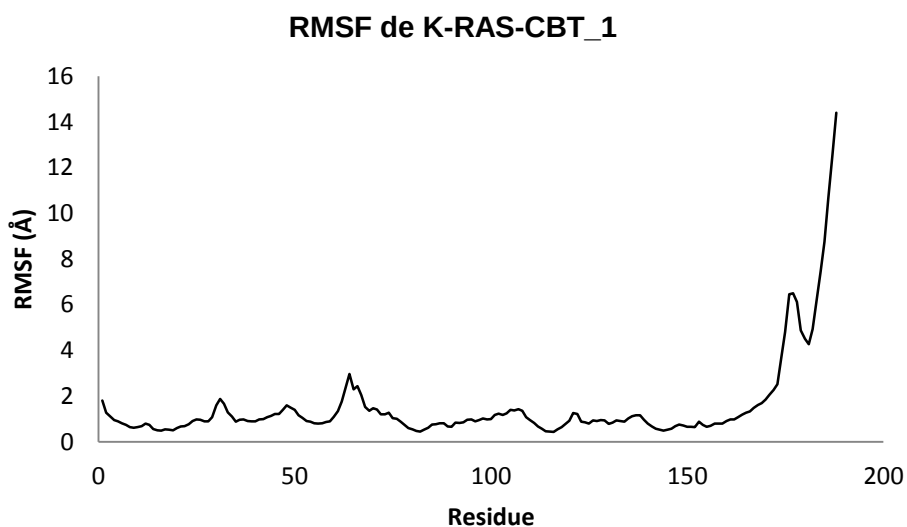


Gráfica 1. Gráfica RMSD de la dinámica 1 del ligando CBT.

5.2. RMSF

Además, la fluctuación dinámica de los residuos de aminoácidos se caracteriza mediante el cálculo de la fluctuación cuadrática media (RMSF de sus siglas en inglés) de átomos C α , durante el tiempo de producción. El análisis de la RMSF permite una identificación de las regiones más flexibles.

En nuestro caso se utilizan los aminoácidos que no fluctúen más de 4 Å respecto de su posición inicial. De acuerdo con este criterio, se eligen los aminoácidos del 1 al 170, ya que a partir del 170 la fluctuación es mayor que la elegida de 4 Å.



Gráfica 2. Gráfica RMSF de la dinámica 1 del ligando CBT.

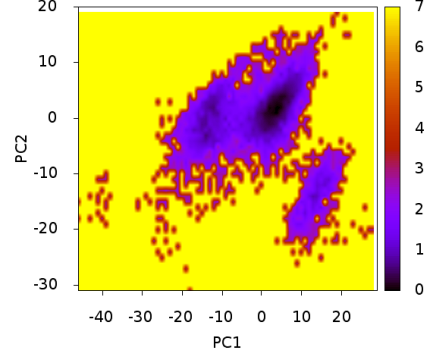
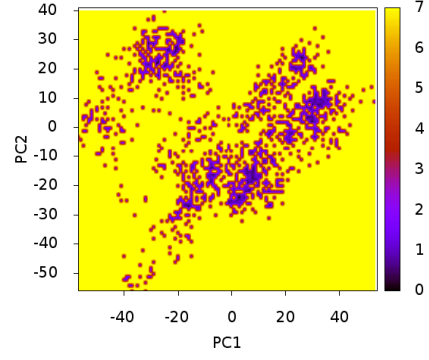
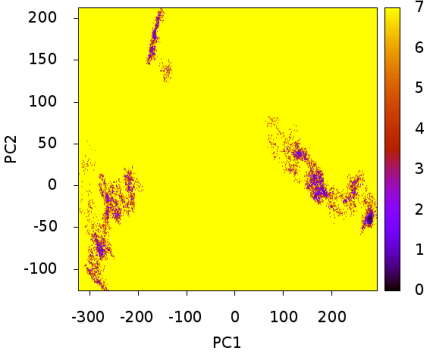
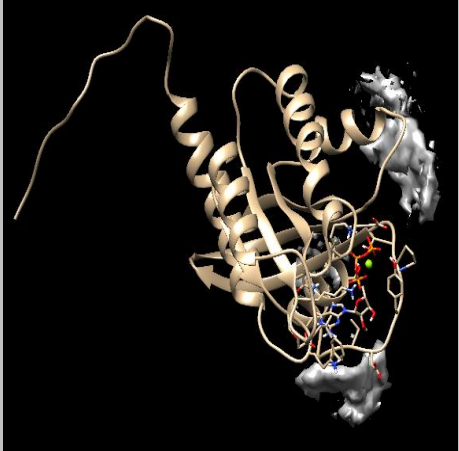
Este procedimiento ha sido seguido para analizar todas las dinámicas, aunque solo se muestran las gráficas RMSD y RMSF de la dinámica molecular clásica número 1 del ligando CBT a modo de ejemplo.

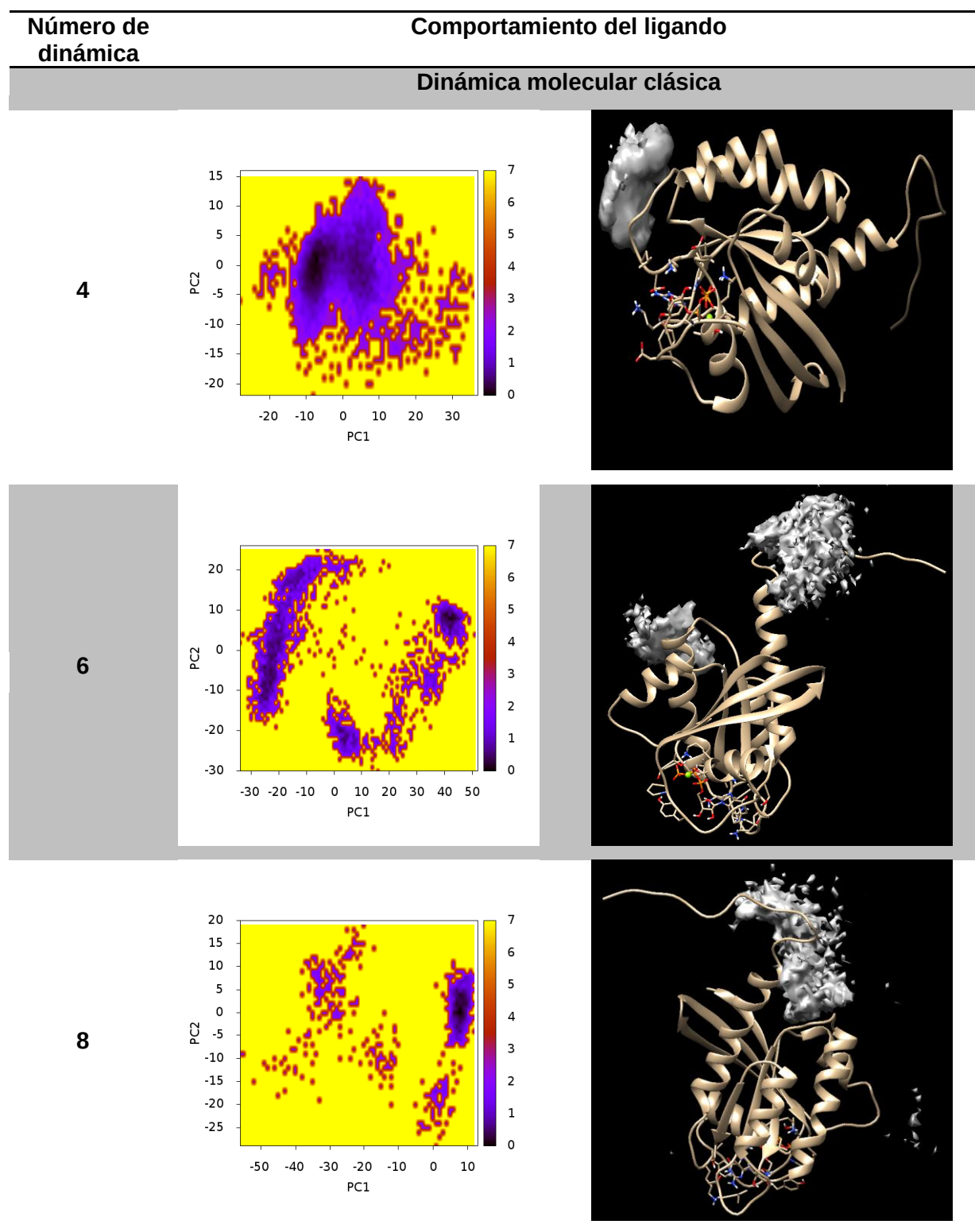
5.3. Análisis de componentes principales (PCA)

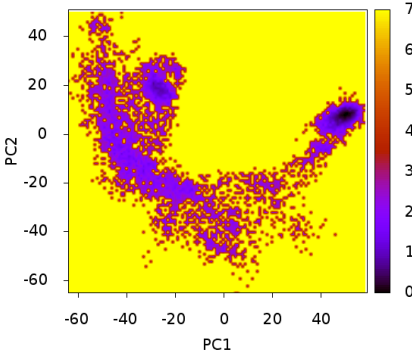
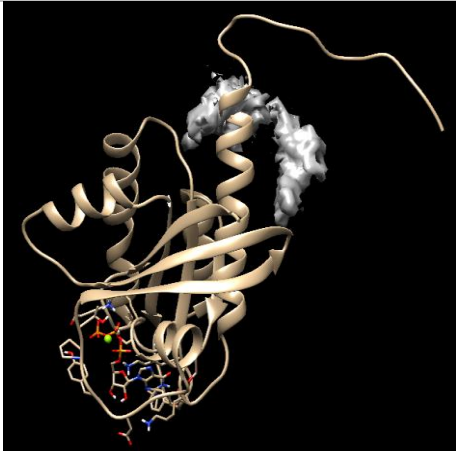
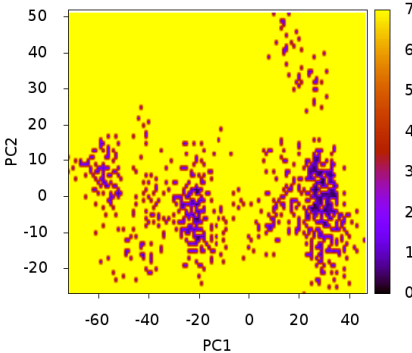
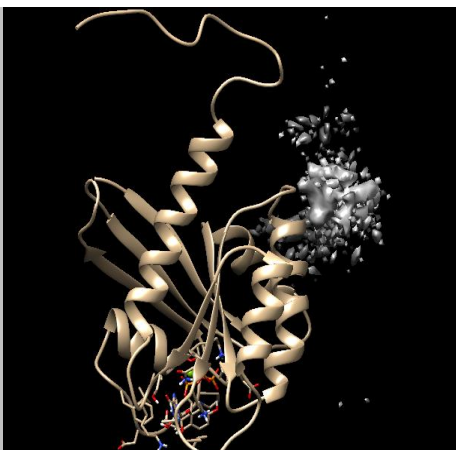
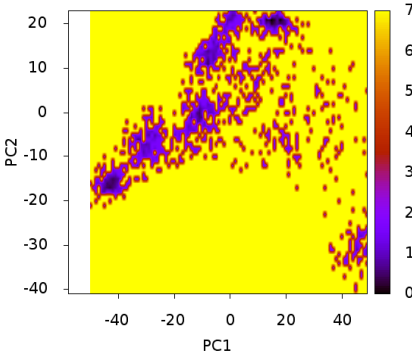
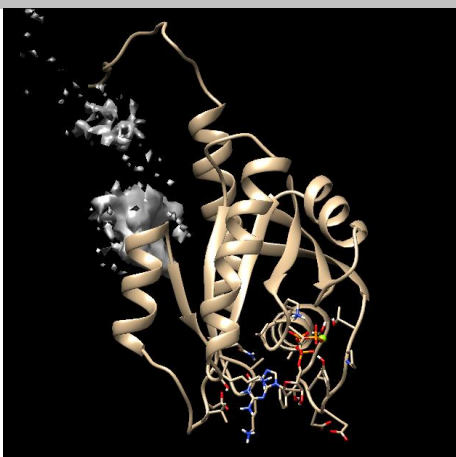
Las siguientes tablas (Tablas 1, 2 y 3) muestran las imágenes del análisis de componentes principales, obtenidas a través de un programa escrito por el Dr. Jaime Rubio de la Universidad de Barcelona en Fortran, y las imágenes de las trayectorias que siguen los ligandos con respecto a la proteína a lo largo de la dinámica. Estas imágenes de las trayectorias han sido calculadas con el programa Chimera [49].

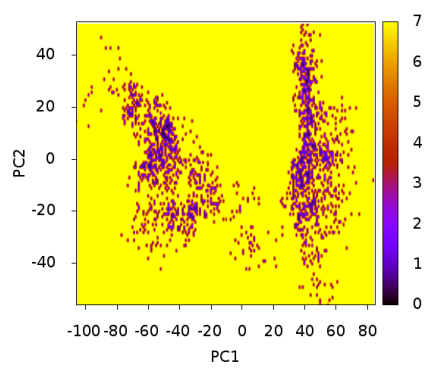
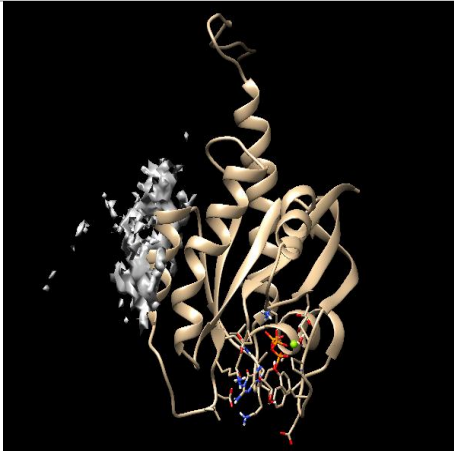
Ligando CBT

Tabla 1. Tabla de PCA e imágenes de trayectoria de las dinámicas del ligando CBT.

Número de dinámica	Comportamiento del ligando	
	Dinámica molecular clásica	
	PCA	Trayectoria del ligando
1		
2		
3		

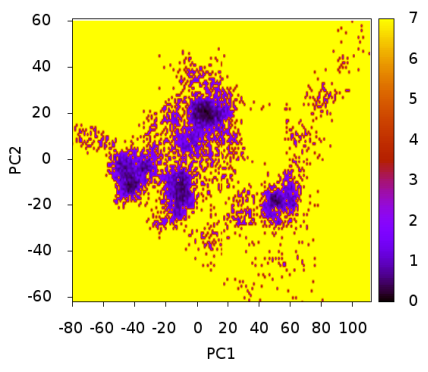
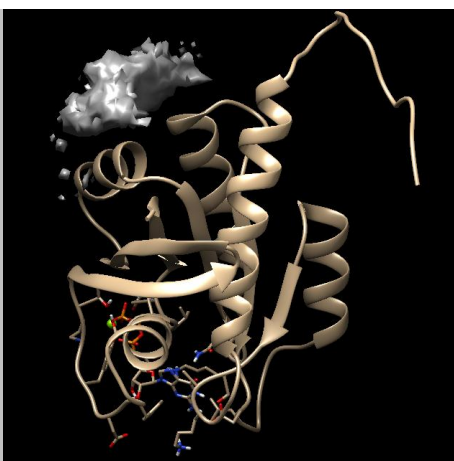


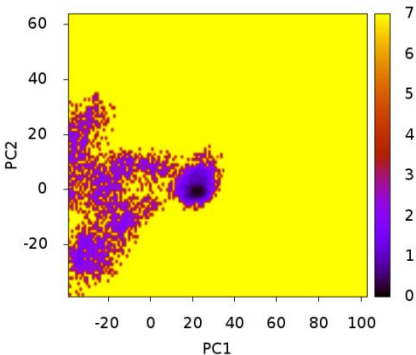
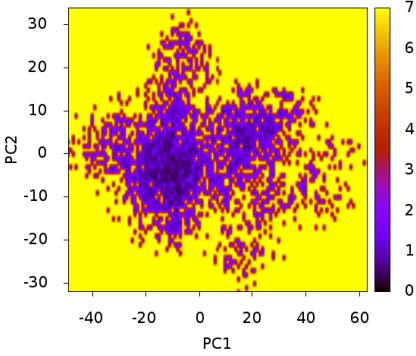
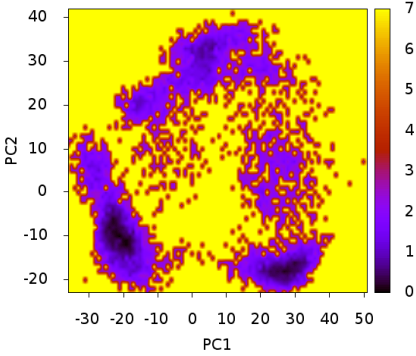
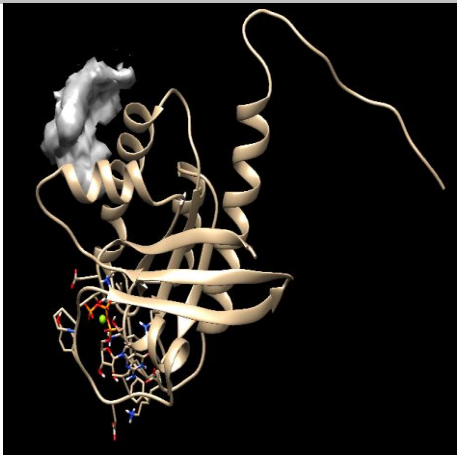
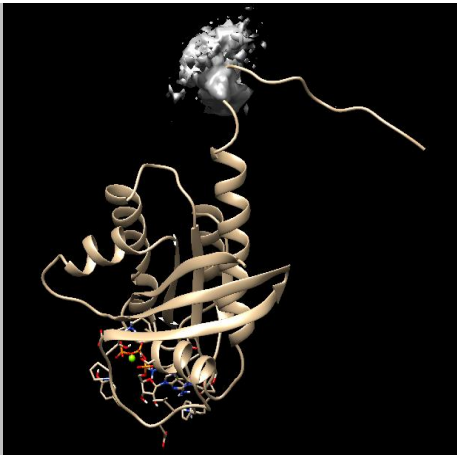
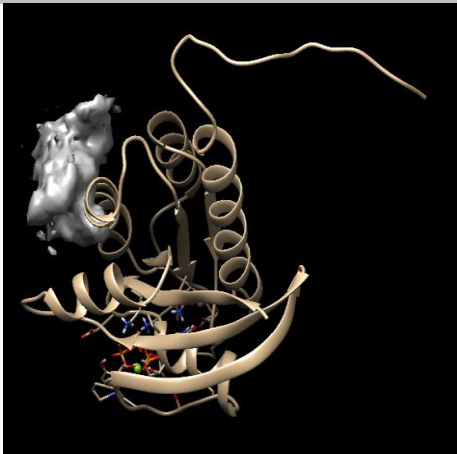
Número de dinámica	Comportamiento del ligando	
Dinámica molecular clásica		
9		
10		
Dinámica molecular escalada		
	PCA	Trayectoria
2		

Número de dinámica	Comportamiento del ligando
Dinámica molecular escalada	
3 	

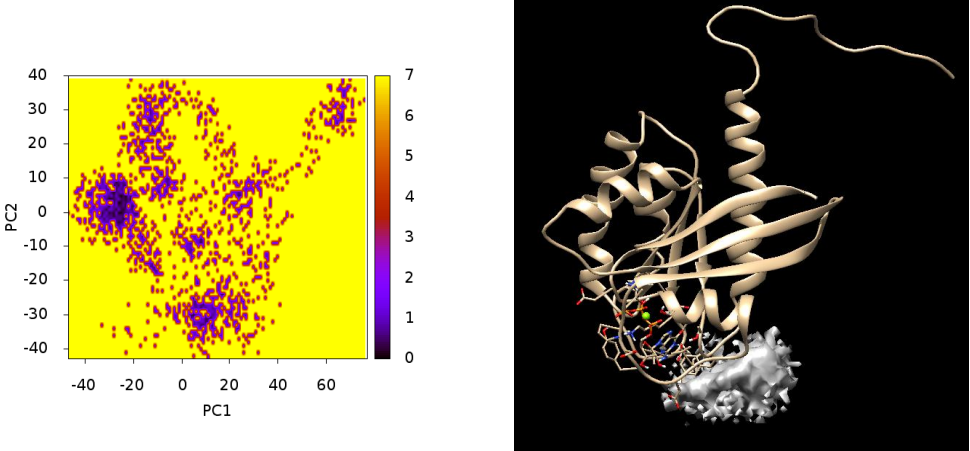
Ligando DAP

Tabla 2. Tabla de PCA e imágenes de trayectoria de las dinámicas del ligando DAP.

Número de dinámica	Comportamiento del ligando
Dinámica molecular clásica	
1 PCA 	Trayectoria 

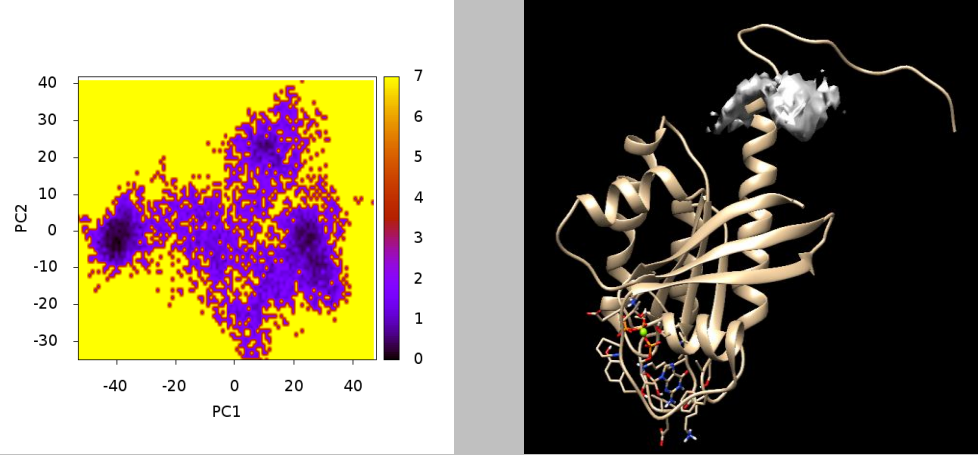
Número de dinámica	Comportamiento del ligando
Dinámica molecular clásica	
2	 PCA plot for dynamics 2. The x-axis is PC1 (ranging from -20 to 100) and the y-axis is PC2 (ranging from -20 to 60). The plot shows a single, dense cluster of points, indicating a stable conformational state. A color scale on the right indicates density from 0 to 7.
5	 PCA plot for dynamics 5. The x-axis is PC1 (ranging from -40 to 60) and the y-axis is PC2 (ranging from -30 to 30). The plot shows a single, dense cluster of points, indicating a stable conformational state. A color scale on the right indicates density from 0 to 7.
6	 PCA plot for dynamics 6. The x-axis is PC1 (ranging from -30 to 50) and the y-axis is PC2 (ranging from -20 to 40). The plot shows multiple distinct clusters of points, indicating conformational changes or multiple stable states. A color scale on the right indicates density from 0 to 7.
	 3D ribbon diagram of a protein structure with a ligand. The protein is shown in a light yellow color, and the ligand is shown in a stick representation with various atoms colored (red, blue, green, orange).
	 3D ribbon diagram of a protein structure with a ligand. The protein is shown in a light yellow color, and the ligand is shown in a stick representation with various atoms colored (red, blue, green, orange).
	 3D ribbon diagram of a protein structure with a ligand. The protein is shown in a light yellow color, and the ligand is shown in a stick representation with various atoms colored (red, blue, green, orange).

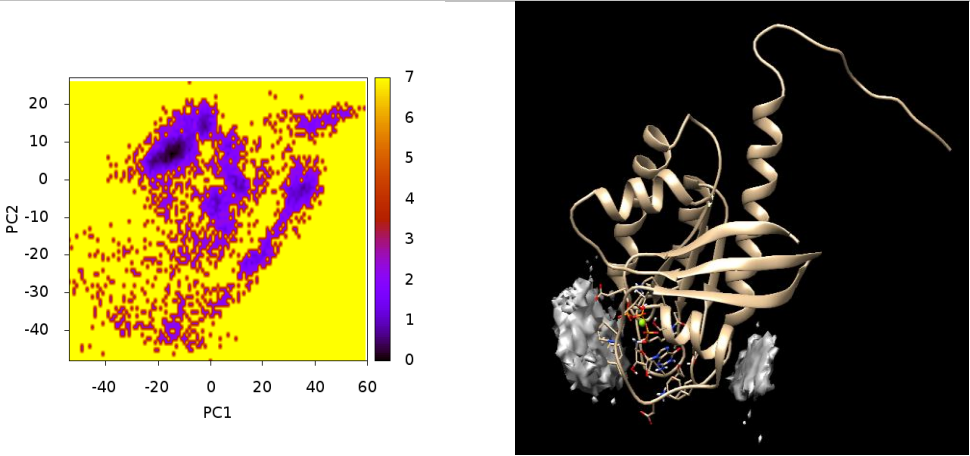
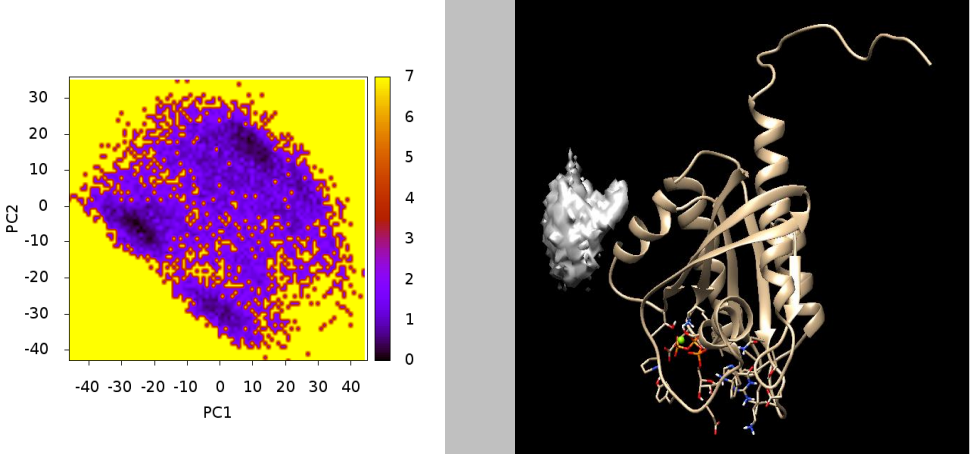
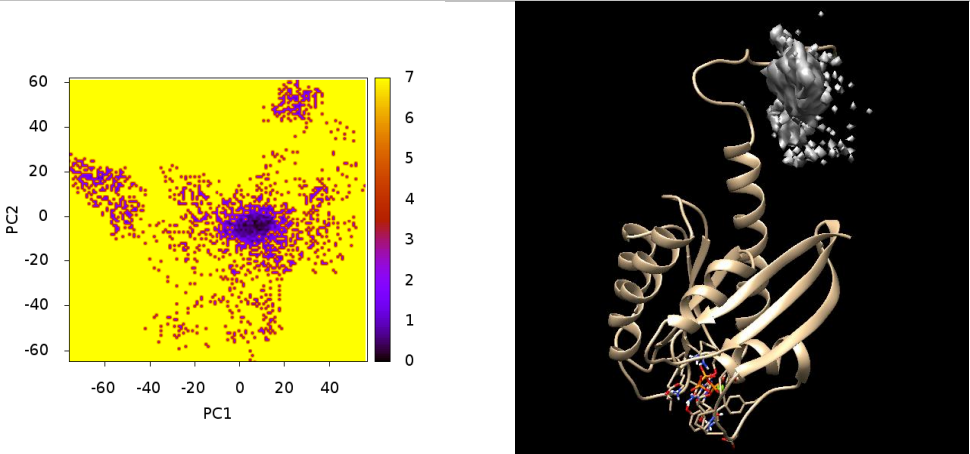
Número de dinámica		Comportamiento del ligando	
Dinámica molecular clásica			
7			
Dinámica molecular escalada			
PCA		Trayectoria	
2			

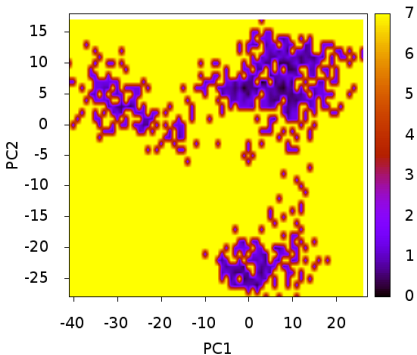
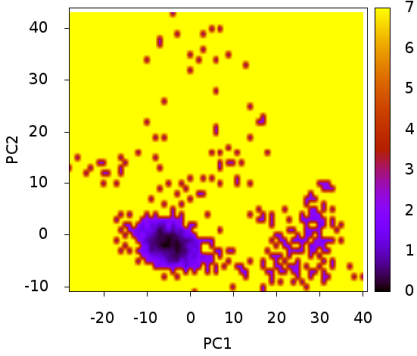
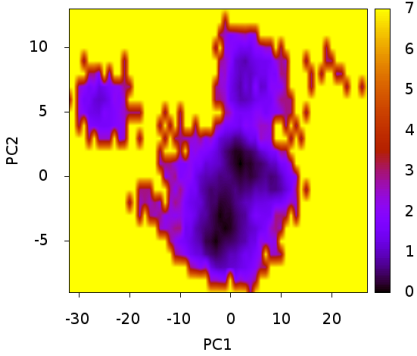
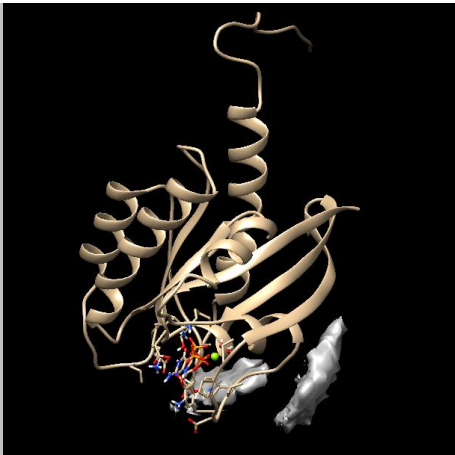
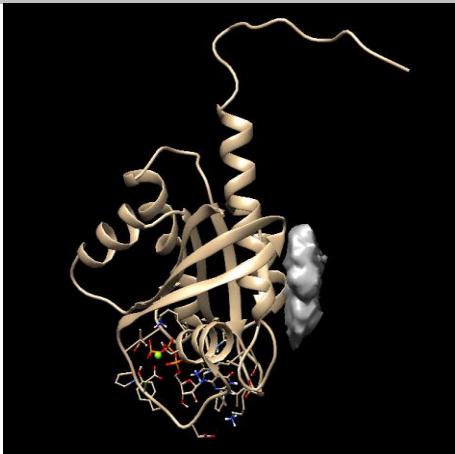
Número de dinámica	Comportamiento del ligando
Dinámica molecular escalada	
4	

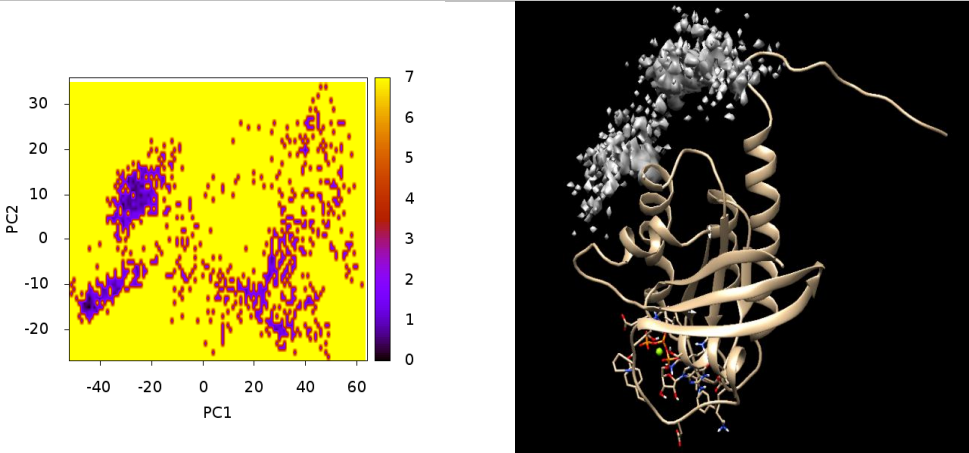
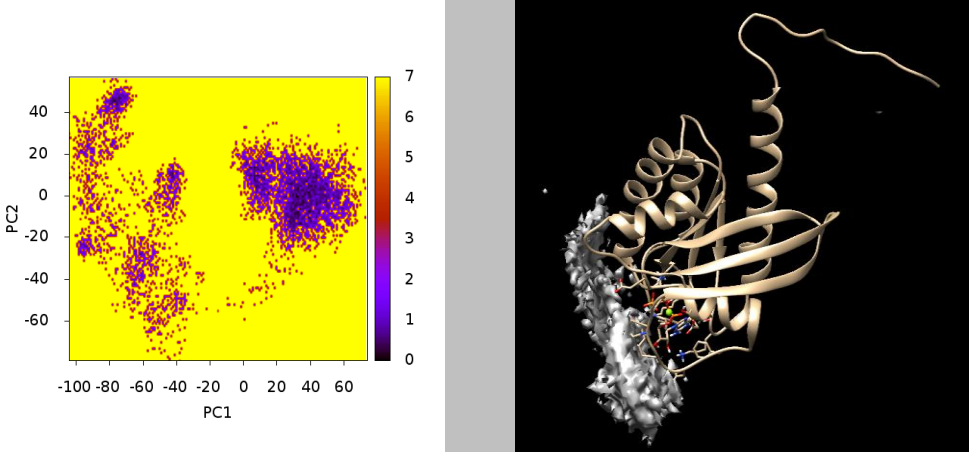
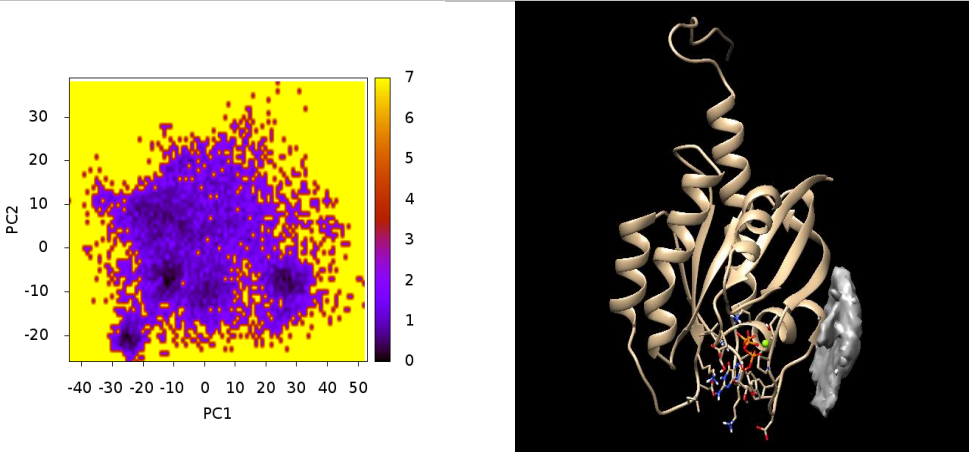
Ligando ITC

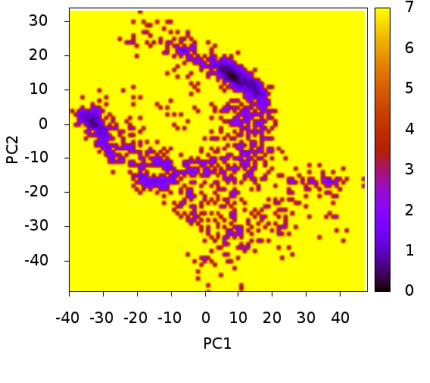
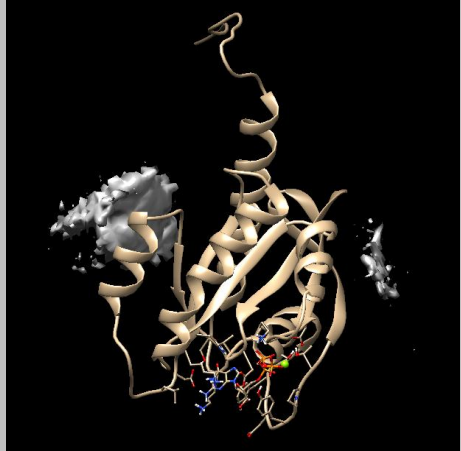
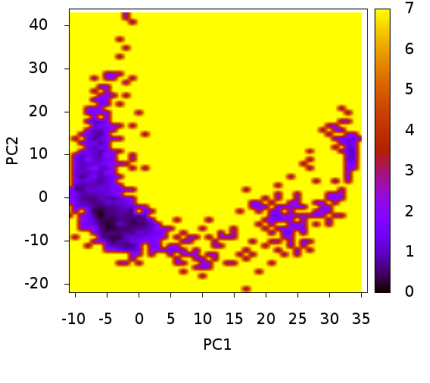
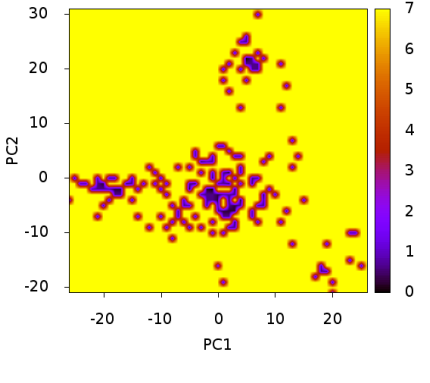
Tabla 3. Tabla de PCA e imágenes de trayectoria de las dinámicas del ligando ITC.

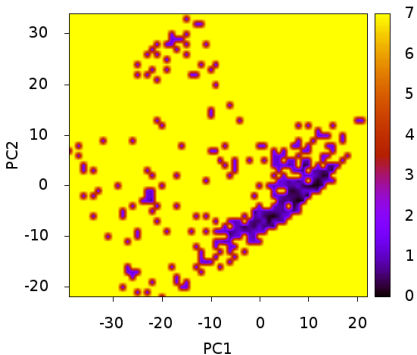
Número de dinámica	Comportamiento del ligando
Dinámica molecular clásica	
	PCA Trayectoria
1	

Número de dinámica	Comportamiento del ligando
Dinámica molecular clásica	
2	
3	
4	

Número de dinámica	Comportamiento del ligando
Dinámica molecular clásica	
5	 PCA plot for dynamics 5. The x-axis is PC1 (ranging from -40 to 20) and the y-axis is PC2 (ranging from -25 to 15). The plot shows a dense cluster of points with a color scale from 0 (purple) to 7 (yellow). The points are concentrated in two main regions: one around PC1 = -20, PC2 = 5 and another around PC1 = 10, PC2 = 5.
6	 PCA plot for dynamics 6. The x-axis is PC1 (ranging from -20 to 40) and the y-axis is PC2 (ranging from -10 to 40). The plot shows a dense cluster of points with a color scale from 0 (purple) to 7 (yellow). The points are concentrated in two main regions: one around PC1 = -10, PC2 = 0 and another around PC1 = 30, PC2 = 0.
7	 PCA plot for dynamics 7. The x-axis is PC1 (ranging from -30 to 20) and the y-axis is PC2 (ranging from -5 to 10). The plot shows a dense cluster of points with a color scale from 0 (purple) to 7 (yellow). The points are concentrated in two main regions: one around PC1 = -20, PC2 = 5 and another around PC1 = 0, PC2 = 5.
	 3D ribbon diagram of a protein structure with a ligand. The protein is shown in a tan ribbon representation, and the ligand is shown in a stick representation with a color scale from 0 (purple) to 7 (yellow). The protein structure is complex, with multiple helices and loops. The ligand is located in the binding pocket of the protein.
	 3D ribbon diagram of a protein structure with a ligand. The protein is shown in a tan ribbon representation, and the ligand is shown in a stick representation with a color scale from 0 (purple) to 7 (yellow). The protein structure is complex, with multiple helices and loops. The ligand is located in the binding pocket of the protein.
	 3D ribbon diagram of a protein structure with a ligand. The protein is shown in a tan ribbon representation, and the ligand is shown in a stick representation with a color scale from 0 (purple) to 7 (yellow). The protein structure is complex, with multiple helices and loops. The ligand is located in the binding pocket of the protein.

Número de dinámica	Comportamiento del ligando
Dinámica molecular clásica	
8	
9	
10	

Número de dinámica	Comportamiento del ligando	
Dinámica molecular escalada		
	PCA	Trayectoria
2		
3		
4		

Número de dinámica	Comportamiento del ligando
Dinámica molecular escalada	
5	
	

Como se puede apreciar en las tablas 1, 2 y 3, el análisis de componentes principales desvela los mínimos de energía según el lugar donde se haya encontrado el ligando. Por otro lado, las imágenes de las trayectorias de los ligandos realizadas con Chimera, representan una malla tridimensional en la que cada punto de la malla tiene un valor proporcional al número de veces que ha estado el ligando en ese punto. Para cada una de las imágenes se elige un porcentaje del tiempo que queremos que se represente, de modo que la malla de puntos se vea lo más clara posible.

5.4. Dinámica molecular

A continuación se hace un análisis de los sitios de unión de cada ligando con la proteína que se dan en cada dinámica, para ello se tiene en cuenta la nomenclatura de la figura 7.

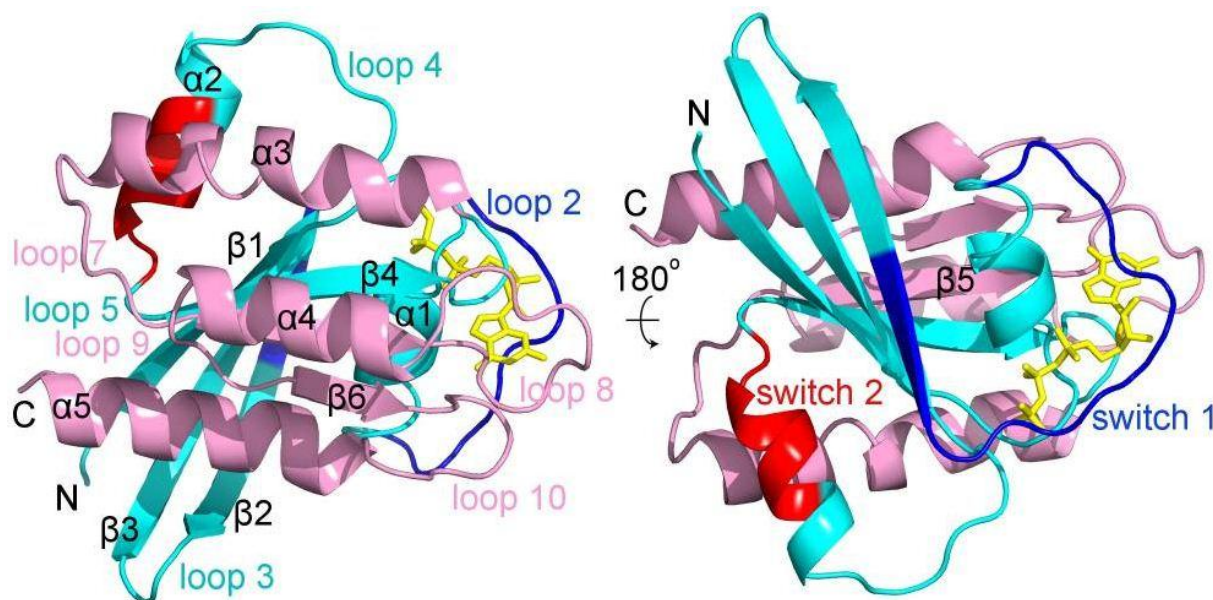


Figura 7. Los *loops* N-terminal y C-terminal se muestran en cian y en rosa respectivamente, mientras que los *switches* 1 y 2 lo hacen en azul y rojo. El GTP se muestra en amarillo [50].

Ligando CBT

En la siguiente tabla (Tabla 4) se muestra el comportamiento del ligando CBT en cada una de las dinámicas, tanto clásicas como escaladas.

Tabla 4. Comportamiento del ligando CBT.

Número de dinámica	Comportamiento del ligando
Dinámica molecular clásica	
1	El ligando queda unido a la hélice $\alpha 5$ por la parte exterior de la proteína durante 65 ns.
2	El ligando se une al <i>loop</i> C-terminal durante 17 ns.
3	El ligando interactúa con el <i>loop</i> 2 (<i>switch</i> 1) a los 30 ns, y permanece estable durante 25 ns. A los 65 ns se une a la hélice $\alpha 2$ (<i>switch</i> 2) durante los últimos 35 ns de la dinámica.
4	El ligando queda unido al <i>loop</i> 8 durante 66 ns.
5	El ligando se mantiene unido al <i>loop</i> 10 durante 32 ns y a la hélice $\alpha 5$ durante los últimos 10 ns.
6	El ligando interactúa con el final de la hélice $\alpha 3$ durante 25 ns.
7	El ligando no interactúa con la proteína en ningún momento de la dinámica.
8	El ligando se une al <i>loop</i> 7 y con el <i>loop</i> C-terminal durante 13 ns.
9	El ligando interactúa con la hélice $\alpha 5$ y con el <i>loop</i> C-terminal durante 45 ns.
10	El ligando interactúa con el final de la hélice $\alpha 3$ durante 14 ns.

Dinámica molecular escalada	
Número de dinámica	Comportamiento del ligando
1	El ligando no interactúa con la proteína en ningún momento de la dinámica.
Dinámica molecular escalada	
2	El ligando se une a la hélice $\alpha 4$ durante 16 ns.
3	El ligando se une a la hélice $\alpha 4$ durante 20 ns.
4	El ligando no interactúa con la proteína en ningún momento de la dinámica.
5	El ligando no interactúa con la proteína en ningún momento de la dinámica.

Ligando DAP

En la siguiente tabla (Tabla 5) se muestra el comportamiento del ligando DAP en cada una de las dinámicas, tanto las clásicas como las escaladas.

Tabla 5. Comportamiento del ligando DAP.

Comportamiento del ligando	
Número de dinámica	
Dinámica molecular clásica	
1	El ligando fluctúa entre la hélice $\alpha 2$ (<i>switch 2</i>) y el <i>loop 7</i> durante 75 ns.
2	El ligando queda unido al final de la hélice $\alpha 3$ durante 92 ns.
3	El ligando no interactúa con la proteína en ningún momento de la dinámica.
4	El ligando no interactúa con la proteína en ningún momento de la dinámica.
5	El ligando queda unido al <i>loop</i> C-terminal durante 97 ns.
6	El ligando se une al final de la hélice $\alpha 3$ durante 83 ns.
7	El ligando va al <i>loop 2</i> (<i>switch 1</i>) a los 36 ns, y permanece durante 24 ns, a continuación deja de interactuar hasta los 76 ns, cuando vuelve a la misma posición durante otros 24 ns hasta el final de la dinámica.
8	El ligando no interactúa con la proteína en ningún momento de la dinámica.
9	El ligando interactúa con la hélice $\alpha 4$ durante 40 ns.
10	El ligando no interactúa con la proteína en ningún momento de la dinámica.
Dinámica molecular escalada	
1	El ligando no interactúa con la proteína en ningún momento de la dinámica.
2	El ligando interactúa al final de la hélice $\alpha 5$ durante 14 ns.
3	El ligando no interactúa con la proteína en ningún momento de la dinámica.
4	El ligando se une al <i>loop 2</i> (<i>switch 1</i>) durante 19 ns.
5	El ligando no interactúa con la proteína en ningún momento de la dinámica.

Ligando ITC

En la siguiente tabla (Tabla 6) se muestra el comportamiento del ligando ITC en cada una de las dinámicas.

Tabla 6. Comportamiento del ligando ITC.

Comportamiento del ligando	
Número de dinámica	
Dinámica molecular clásica	
1	El ligando se une a la hélice $\alpha 5$ durante 67 ns.
2	El ligando interactúa con el <i>loop 2</i> (<i>switch 1</i>) durante 14 ns y 72 ns con el

	loop 10.
3	El ligando se une a la hélice $\alpha 2$ (<i>switch 2</i>) durante 90 ns.
Número de dinámica	Comportamiento del ligando
Dinámica molecular clásica	
4	El ligando interactúa con el <i>loop</i> C-terminal durante 17 ns.
5	El ligando se une a la hélice $\alpha 4$ durante 10 ns.
6	El ligando se une al <i>loop 2</i> (<i>switch 1</i>) durante 54 ns.
7	El ligando queda unido a la hélice $\alpha 5$ durante 85 ns.
8	El ligando interactúa con el <i>loop 4</i> y el <i>loop</i> C-terminal durante 13 ns.
9	El ligando queda cerca del <i>loop 8</i> durante 19 ns y después con el <i>loop 10</i> durante 39 ns.
10	El ligando queda unido al <i>loop 2</i> (<i>switch 1</i>) durante 77 ns. Además, en este caso interactúa con la hélice $\alpha 1$ y con la lámina $\beta 2$.
Dinámica molecular escalada	
1	El ligando no interactúa con la proteína en ningún momento de la dinámica.
2	El ligando interactúa con la hélice $\alpha 4$ durante 22 ns.
3	El ligando queda unido al final de la hélice $\alpha 3$ durante 22 ns.
4	El ligando interactúa con la hélice $\alpha 2$ (<i>switch 2</i>) durante 4 ns.
5	El ligando queda cerca de la hélice $\alpha 2$ (<i>switch 2</i>) durante 7 ns.

Analizando las dinámicas se puede apreciar que los ligandos interactúan con el *loop* C-terminal en cinco de ellas, aunque esta posición no se considera relevante, ya que el que los ligandos interactúen con ella se puede deber a que está muy expuesto, aunque en realidad este *loop* no cumple función en la proteína.

Respecto al *loop 2*, el cual es el *switch 1*, se aprecia que el ligando CBT interactúa con él una vez, el ligando DAP dos veces y el ligando ITC hasta tres veces, con lo que los ligandos interactúan un total de seis veces. De este modo, este *loop* se puede considerar un sitio de unión probable.

Por su parte, los *loops 4, 7, 8 y 10* no son buenos sitios de unión ya que los ligandos interactúan una, dos, dos y tres veces respectivamente. Esto se debe a que tienen otros sitios de unión mejores cerca –la hélice $\alpha 2$ en el caso del *loop 4*, las hélices $\alpha 3$ y $\alpha 4$ en el caso del *loop 7* y el *loop 2* en el caso de los *loops 8 y 10*-, y por lo tanto los ligandos prefieren unirse a estos otros.

De igual forma, la hélice $\alpha 1$ y la lámina $\beta 2$ no se pueden considerar buen sitio de unión ya que solamente el ligando ITC interactúa una vez en estas posiciones, y por un tiempo muy corto.

Por otra parte, la hélice $\alpha 2$, la mayor parte de ella considerada el *switch 2*, interactúa una vez con el ligando CBT, una vez más con el ligando DAP y tres veces con el ligando ITC.

Por último, en cuanto a las hélices $\alpha 3$, $\alpha 4$ y $\alpha 5$ son sitios de unión probables, ya que se unen un total de cinco veces a cada hélice.

Haciendo un análisis por ligando, tal como se puede ver en las tablas 4 a 6 se observa que el ligando CBT interactúa en once de los quince procesos a los que ha sido sometido. Del mismo modo, el ligando DAP interactúa con la proteína en ocho de las quince dinámicas. Sin embargo, el ligando ITC se une a la proteína en catorce de las quince dinámicas, lo que a priori, antes de hacer un análisis de los cálculos de energía libre, hace pensar que éste va a ser un buen ligando, ya que tiene una mayor tendencia a unirse a la proteína que los ligandos CBT y DAP.

La tabla 7 muestra los enlaces de hidrógeno que forman los ligandos con los respectivos aminoácidos y la posición en la que se encuentran estos, y además, cuando ha sido posible determinarlo, se muestran qué átomos concretos son los que forman estos enlaces de hidrógeno.

Tabla 7. Enlaces de hidrógeno formados por los ligandos con los aminoácidos de la proteína.

Ligando	Número de dinámica	Sitio de unión	Enlaces de hidrógeno	
			Aminoácido	Átomo
CBT	1	Hélice $\alpha 5$	Arg 161	
			Arg 164	
			Lys 165	
			Glu 168	
			Lys 169	
CBT	2	Loop C-terminal	Lys 182	
			Thr 183	
			Lys 184	
CBT	3	Loop 2	Gln 25	HE21
			Asn 26	HD21
			Asp 30	OD2
		Hélice $\alpha 2$	Ala 66	
			Met 67	
CBT	4	Loop 8	Gln 70	HE21
			Tyr 71	OH
			Pro 121	
			Ser 122	
			Arg 123	
			Thr 124	
			Val 125	
CBT	5	Loop 10	Asp 126	
			Thr 148	HG1
			Gln 150	OE1
CBT	6	Hélice $\alpha 1$	Gln 22	HE22
		Hélice $\alpha 3$	Glu 98	
			Lys 101	
			Arg 102	

Enlaces de hidrógeno				
Ligando	Número de dinámica	Sitio de unión	Aminoácido	Átomo
CBT	8	Loop 7	Lys 104	HZ1, HZ2
			Ser 106	HG
CBT	9	Loop C-terminal	Asp 173	OD1, OD2, CG
		Hélice α5	Lys 169	
		Loop C-terminal	Lys 172	
			Asp 173	
			Gly 174	
			Lys 175	
			Lys 176	
CBT	10	Hélice α3	Lys 101	
CBT	sMD 2	Hélice α4	Gln 131	
			Ala 134	
			Arg 135	
CBT	sMD 3	Hélice α4	Asp 132	
			Leu 133	
			Tyr 137	
DAP	1	Hélice α2	Met 67	
			Gln 70	HE21, HE22, CD
			Arg 73	HE, HH11, HH12, HH22, NH1,
			Thr 74	HG1
Loop 7			Lys 104	O
			Asp 105	OD1, OD2
			Glu 107	H
DAP	2	Hélice α3	Hip 95	
			Glu 98	OE1
			Gln 99	HE22, OE1
			Arg 102	HE, HH22
			Asp 105	OD2
DAP	5	Loop C-terminal	Lys 175	
			Lys 176	
			Ser 181	
			Lys 182	
DAP	6	Hélice α3	Hie 94	
			Hip 95	
			Glu 98	
			Gln 99	
			Arg 102	
DAP	7	Loop 2	Gln 25	
			Asn 26	ND2, OD1, HD21, HD22
			Hip 27	HE2
DAP	9	Hélice α4	Leu 133	
			Arg 135	
			Ser 136	
			Tyr 137	
DAP	sMD 2	Hélice α5	Lys 165	
DAP	sMD 4	Loop 2	Phe 28	

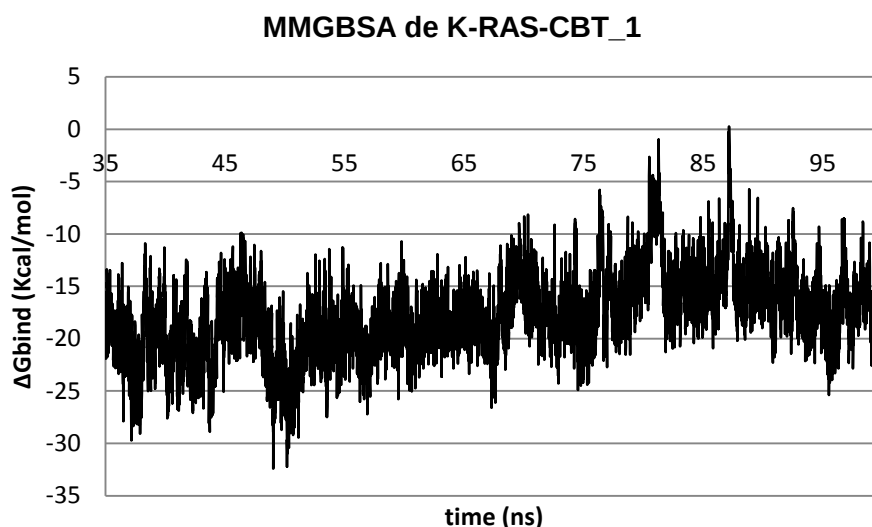
Ligando	Número de dinámica	Sitio de unión	Enlaces de hidrógeno	
			Aminoácido	Átomo
ITC	1	Hélice $\alpha 5$	Lys 167	HZ1, HZ2, HZ3, NZ
			Lys 169	HZ3
			Ser 171	HG
ITC	2	Loop 2	Asn 26	HD22
		Loop 10	Lys 117	HZ1, HZ3
			Ser 122	HG
ITC	3	Hélice $\alpha 2$	Thr 148	O
			Ala 66	
			Met 67	
ITC	4	Loop C-terminal	Lys 180	
ITC	5	Hélice $\alpha 4$	Arg 135	
			Ser 136	
			Asn 26	O, HD22, ND2
ITC	6	Loop 2	Phe 28	H
			Asp 154	
			Arg 161	
ITC	7	Hélice $\alpha 5$	Lys 165	
			Tyr 64	
			Ala 66	
ITC	8	Loop 4	Ile 187	
		Loop C-terminal	Lys 117	HZ2
		Loop 8	Ser 122	HG
ITC	9	Loop 10	Lys 147	HZ1
		Loop 2	Asn 26	HD21, HB3
		Hélice $\alpha 1$	Ile 24	O
ITC	10	Lámina $\beta 2$	Gln 25	HE21, HE22, OE1
			Arg 41	O, H
			Lys 42	HZ1, HZ2, HZ3
ITC	SMD 2	Hélice $\alpha 4$	Gln 43	H
			Arg 135	
			Hie 94	
ITC	SMD 3	Hélice $\alpha 3$	Hip 95	
			Glu 98	
			Arg 102	
ITC	SMD 4	Hélice $\alpha 2$	Met 67	
			Gln 70	
			Tyr 71	
ITC	SMD 5	Hélice $\alpha 2$	Thr 74	
			Ala 66	
			Gln 70	HE 22

Todos los átomos indicados en la tabla pertenecen a la cadena lateral de la proteína, mientras que en los aminoácidos que no tienen indicados los átomos, el enlace de hidrógeno se produce en el esqueleto central.

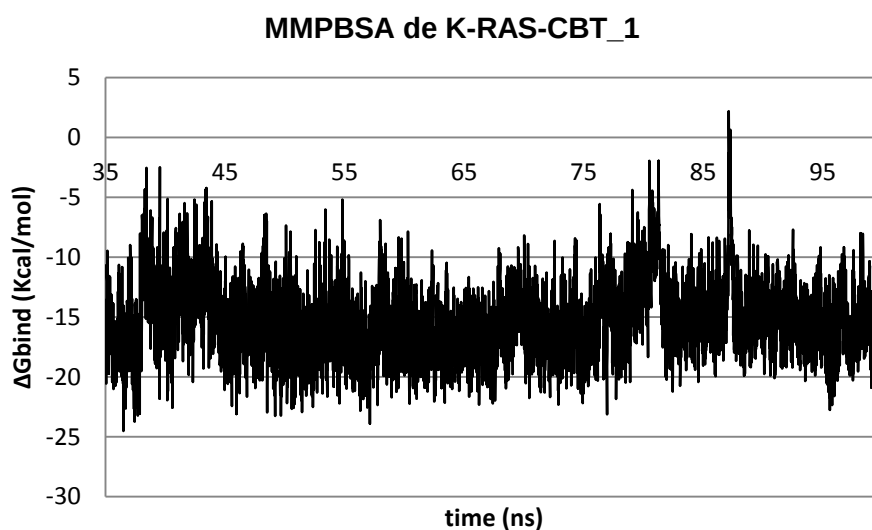
5.5. Cálculo de la energía libre de unión

La energía libre de unión total (ΔG_{bind}) se representa como una función de la trayectoria, con el fin de analizar la convergencia en la simulación y la estabilidad del complejo. Para calcular esta ΔG_{bind} se usan los métodos MMGBSA y MMPBSA.

Como ejemplo, las gráficas 3 y 4 muestran las energías libres de enlace calculadas en función del tiempo en el que el ligando CBT está unido a la proteína en la simulación 1.



Gráfica 3. Gráfica del método MMGBSA de la dinámica 1 del ligando CBT.



Gráfica 4. Gráfica del método MMPBSA de la dinámica 1 del ligando CBT.

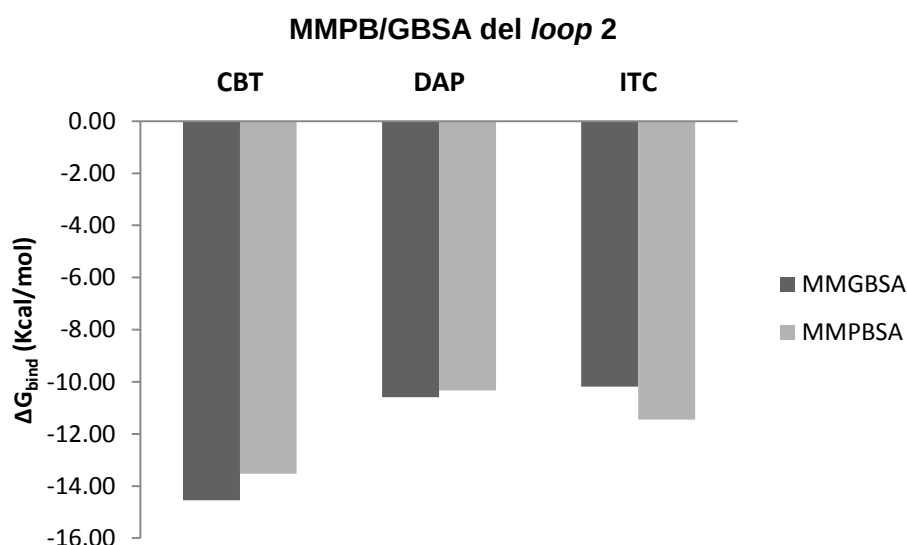
Indicar que, aunque generalmente los valores calculados por estos métodos dan energías libres sobreestimadas, éstos que son útiles con fines comparativos. Por otro lado, el término entrópico no se calcula debido a que tiene un alto coste computacional.

En este apartado se van a calcular las energías libres medias de enlace de cada sitio de unión interesante, comparadas por ligando mediante los métodos MMGBSA y MMPBSA. De este modo, según lo discutido en el apartado 5.4 se calcularán las energías libres de enlace del *loop* 2, y de las hélices $\alpha 2$, $\alpha 3$, $\alpha 4$ y $\alpha 5$.

Comenzando por el *loop* 2, se calculan las energías libres medias de cada una de las dinámicas, mediante cada uno de los dos métodos, que interactúan con este *loop* en el momento de la interacción. Por tanto, para el cálculo de estas energías, en este caso usamos las dinámicas CBT 3, DAP 7, la dinámica molecular escalada 4 del ligando DAP, ITC 2, ITC 6 e ITC 10. De este modo se obtienen los datos siguientes (Tabla 8) y se construye la gráfica (Gráfica 5)

Tabla 8. Energías libres de unión del *loop* 2 calculadas mediante los métodos MMGBSA y MMPBSA expresadas en Kcal/mol.

	MMGBSA	MMPBSA
Ligando	ΔG_{bind}	ΔG_{bind}
CBT	-14.54	-13.52
DAP	-10.59	-10.33
ITC	-10.18	-11.45



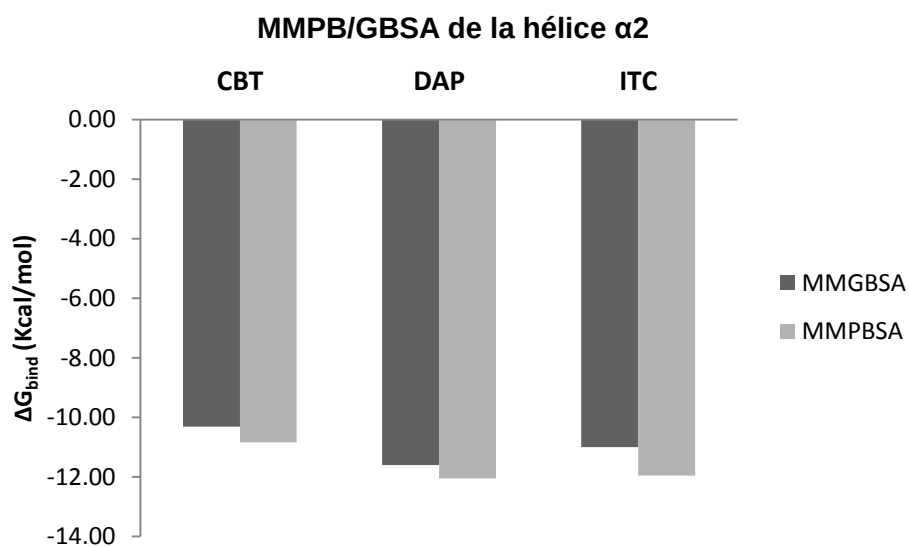
Gráfica 5. Gráfica de las energías libres de unión del *loop* 2 calculadas mediante los métodos MMGBSA y MMPBSA.

Como se puede apreciar en la gráfica (Gráfica 5), la unión K-RAS-CBT es la que tiene una menor energía libre de unión, lo que significa que el ligando CBT tiene mayor afinidad por el *loop* 2 que los ligandos DAP e ITC.

En cuanto a la hélice $\alpha 2$, se siguen los mismos pasos que para calcular las energías libres de unión del *loop* 2. De este modo se usan las dinámicas CBT 3, DAP 1, ITC 3, y las dinámicas moleculares escaladas de ITC 4 y 5.

Tabla 9. Energías libres de unión de la hélice $\alpha 2$ calculadas mediante los métodos MMGBSA y MMPBSA expresadas en Kcal/mol.

	MMGBSA	MMPBSA
Ligando	ΔG_{bind}	ΔG_{bind}
CBT	-10.31	-10.84
DAP	-11.59	-12.05
ITC	-11.00	-11.95

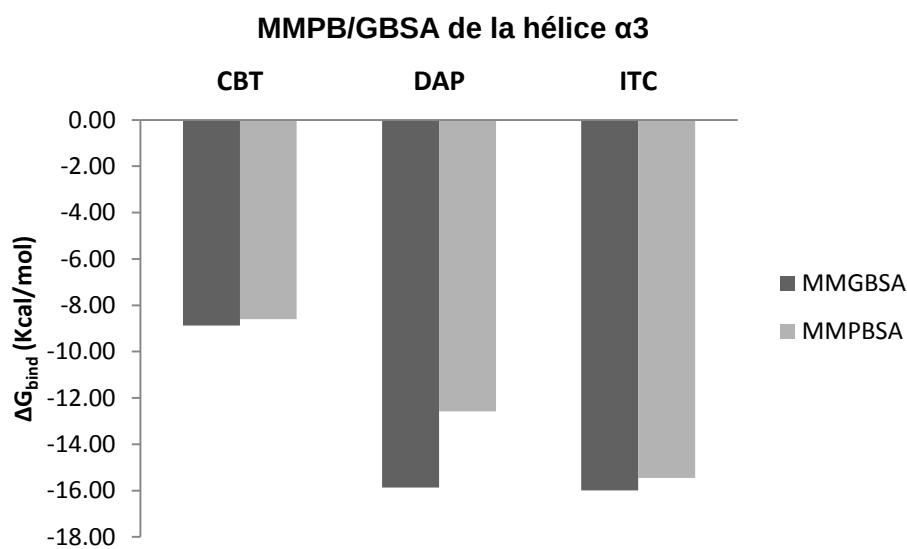
**Gráfica 6.** Gráfica de las energías libres de unión de la hélice $\alpha 2$ calculadas mediante los métodos MMGBSA y MMPBSA.

En este caso, los complejos que necesitan una menor energía para formarse serían K-RAS-DAP y K-RAS-ITC, aunque como se puede ver (ver Tabla 9), los valores de energía libre se encuentran muy próximos entre sí.

Para la hélice $\alpha 3$ se usan las dinámicas CBT 6, CBT 10, DAP 2, DAP 6 y la dinámica molecular escalada número 3 del ligando ITC.

Tabla 10. Energías libres de unión de la hélice $\alpha 3$ calculadas mediante los métodos MMGBSA y MMPBSA expresadas en Kcal/mol.

	MMGBSA	MMPBSA
Ligando	ΔG_{bind}	ΔG_{bind}
CBT	-8.87	-8.60
DAP	-15.86	-12.57
ITC	-15.99	-15.46



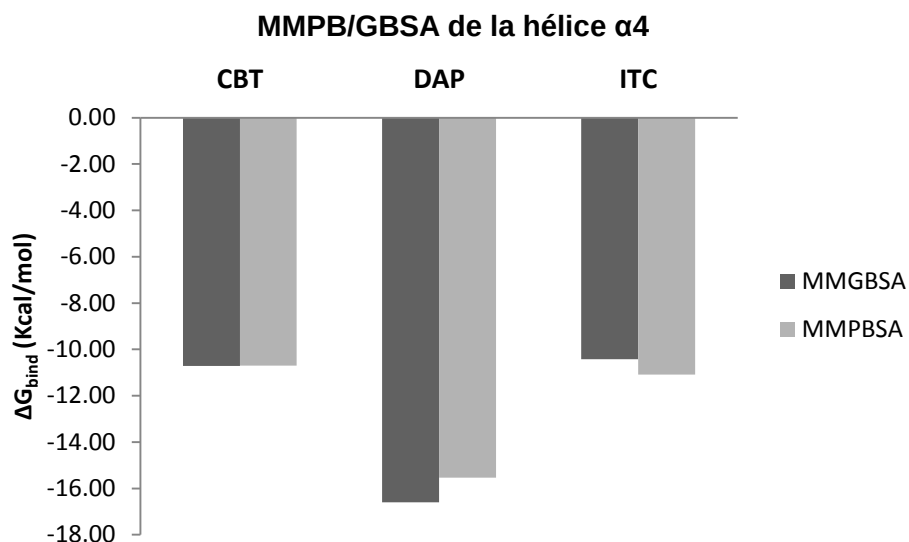
Gráfica 7. Gráfica de las energías libres de unión de la hélice $\alpha 3$ calculadas mediante los métodos MMGBSA y MMPBSA.

Como se puede apreciar (ver Gráfica 7), CBT es ligando que menor afinidad tiene por la hélice $\alpha 3$, mientras que los ligandos DAP e ITC tienen una afinidad parecida.

En el caso de la hélice $\alpha 4$, las dinámicas utilizadas con la 2 y la 3 de las dinámicas moleculares escaladas del ligando CBT, DAP 9, ITC 5 y la dinámica molecular escalada 2 de ITC.

Tabla 11. Energías libres de unión de la hélice $\alpha 4$ calculadas mediante los métodos MMGBSA y MMPBSA expresadas en Kcal/mol.

	MMGBSA	MMPBSA
Ligando	ΔG_{bind}	ΔG_{bind}
CBT	-10.72	-10.70
DAP	-16.59	-15.53
ITC	-10.43	-11.09



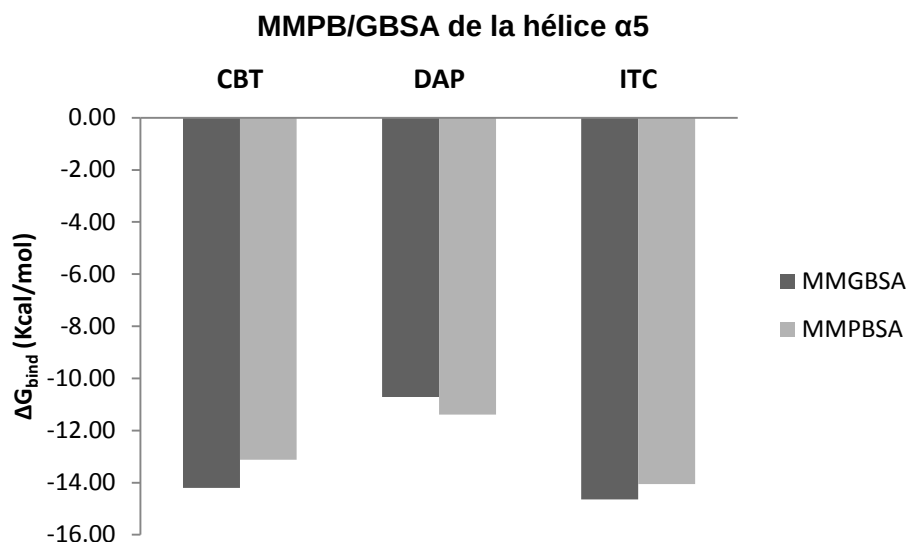
Gráfica 8. Gráfica de las energías libres de unión de la hélice $\alpha 4$ calculadas mediante los métodos MMGBSA y MMPBSA.

A la vista de los resultados (ver Gráfica 8), se ve con claridad que el ligando DAP tiene una mayor tendencia a unirse a la hélice $\alpha 4$ que los ligandos CBT e ITC.

Por último, para la hélice $\alpha 5$ se usan las dinámicas CBT 1, CBT 2, la dinámica molecular escalada número 2 de DAP y las dinámicas ITC 1 e ITC 7.

Tabla 12. Energías libres de unión de la hélice $\alpha 5$ calculadas mediante los métodos MMGBSA y MMPBSA expresadas en Kcal/mol.

	MMGBSA	MMPBSA
Ligando	ΔG_{bind}	ΔG_{bind}
CBT	-14.20	-13.12
DAP	-10.71	-11.39
ITC	-14.64	-14.05



Gráfica 9. Gráfica de las energías libres de unión de la hélice $\alpha 5$ calculadas mediante los métodos MMGBSA y MMPBSA.

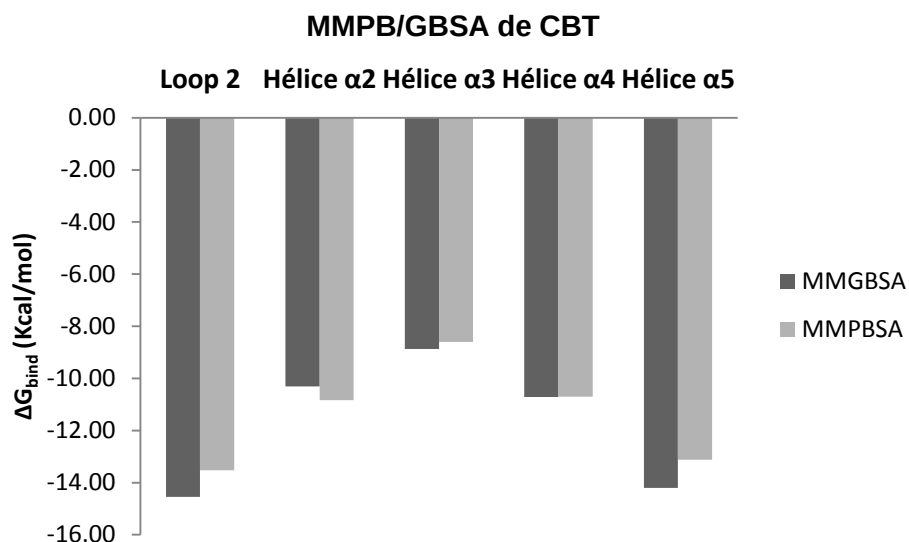
Al contrario que sucedía con la hélice $\alpha 4$, en este caso los complejos que presentan una menor energía libre de unión si se une por la hélice $\alpha 5$ son los formados por los ligandos CBT e ITC, mientras que DAP es el ligando que menor afinidad presenta por esta hélice.

A continuación se hace un análisis de los ligandos en función de las energías libres de enlace de los sitios de unión.

En primer lugar, se muestran los datos (Tabla 13) y la gráfica (Gráfica 10) construida a partir de estos datos del ligando CBT.

Tabla 13. Energías libres de unión del ligando CBT a cada sitio de unión calculadas mediante los métodos MMGBSA y MMPBSA expresadas en Kcal/mol.

	MMGBSA	MMPBSA
	ΔG_{bind}	ΔG_{bind}
Loop 2	-14.54	-13.52
Hélice $\alpha 2$	-10.31	-10.84
Hélice $\alpha 3$	-8.87	-8.60
Hélice $\alpha 4$	-10.72	-10.70
Hélice $\alpha 5$	-14.20	-13.12



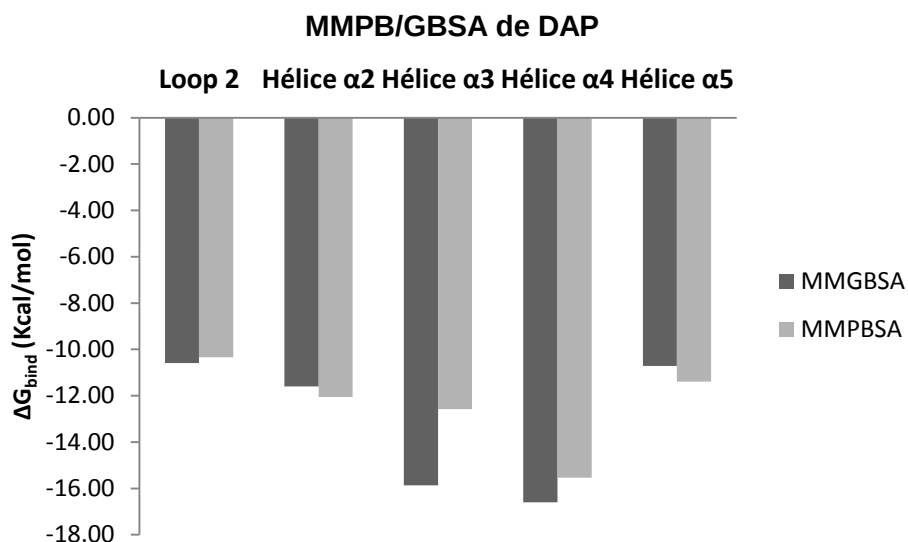
Gráfica 10. Gráfica de las energías libres de unión del ligando CBT calculadas mediante los métodos MMGBSA y MMPBSA.

Como se puede apreciar (ver Gráfica 10), el ligando CBT tiene mayor afinidad por el *loop* 2 y por la hélice α5, ya que la energía libre del complejo formado a través de estos sitios de unión es menor que las energías de las otras tres hélices.

Seguidamente, se ve la tabla (Tabla 14) con las energías libres de unión para el ligando DAP.

Tabla 14. Energías libres de unión del ligando DAP a cada sitio de unión calculadas mediante los métodos MMGBSA y MMPBSA expresadas en Kcal/mol.

	MMGBSA	MMPBSA
	ΔG_{bind}	ΔG_{bind}
Loop 2	-10.59	-10.33
Hélice α2	-11.59	-12.05
Hélice α3	-15.86	-12.57
Hélice α4	-16.59	-15.53
Hélice α5	-10.71	-11.39



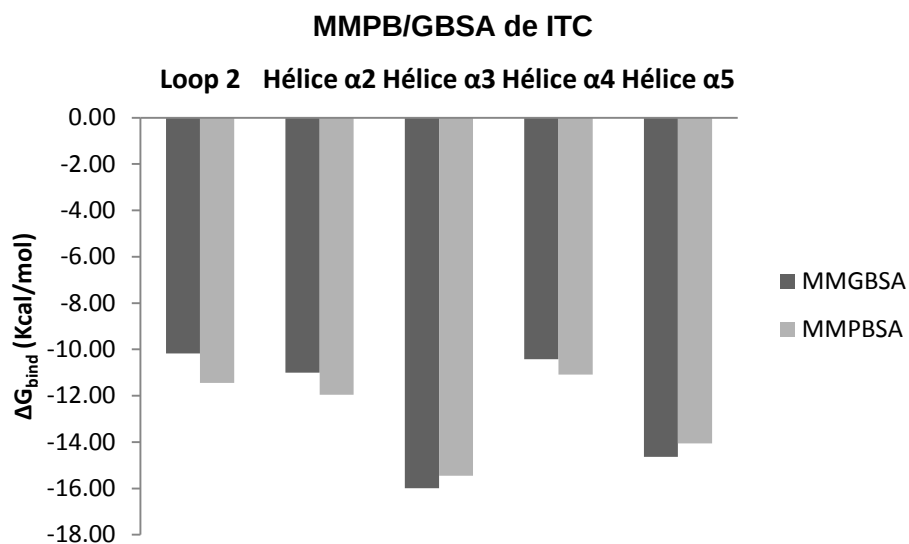
Gráfica 11. Gráfica de las energías libres de unión del ligando DAP calculadas mediante los métodos MMGBSA y MMPBSA.

A diferencia del ligando CBT, el ligando DAP tiende a unirse preferentemente a la hélice $\alpha 4$, seguido con una energía de enlace ligeramente mayor, de la hélice $\alpha 3$, mientras que las energías de enlace con el *loop 2* y las hélices $\alpha 2$ y $\alpha 5$ suben ligeramente.

Por último, las energías libres de enlace con cada sitio de unión del ligando ITC son las siguientes (Tabla 15):

Tabla 15. Energías libres de unión del ligando ITC a cada sitio de unión calculadas mediante los métodos MMGBSA y MMPBSA expresadas en Kcal/mol.

	MMGBSA	MMPBSA
	ΔG_{bind}	ΔG_{bind}
Loop 2	-10.18	-11.45
Hélice $\alpha 2$	-11.00	-11.95
Hélice $\alpha 3$	-15.99	-15.46
Hélice $\alpha 4$	-10.43	-11.09
Hélice $\alpha 5$	-14.64	-14.05



Gráfica 12. Gráfica de las energías libres de unión del ligando ITC calculadas mediante los métodos MMGBSA y MMPBSA.

En este último caso, se ve (ver Gráfica 12) que el ligando ITC suele unir a las hélices $\alpha 3$ y $\alpha 5$ con una diferencia de energía apreciable respecto a la energía necesaria para unir con el *loop 2* y con las hélices $\alpha 2$ y $\alpha 4$.

6. Discusión

Las simulaciones de dinámica molecular, tanto clásica como escalada, de los ligandos CBT, DAP e ITC con la proteína K-RAS se realizan con el objetivo de comprobar teóricamente cuál de ellos es un mejor ligando a la hora de desactivar dicha proteína, analizando los posibles sitios de unión de la misma.

Los resultados obtenidos mediante dinámica molecular clásica sugieren que los ligandos se unen a la proteína generalmente cuando ha transcurrido un tiempo inferior al de los cálculos de la dinámica molecular escalada. Esto es debido a que al subir la barrera de potencial, el mínimo de energía de unión es demasiado alto y por lo tanto los ligandos no son suficientemente estables cuando llegan a uno de estos mínimos.

El análisis RMSD refleja que la proteína es estable durante toda la dinámica molecular.

Por otro lado, si se observa el análisis de componentes principales se pueden ver los mínimos de energía en cada proceso. Esto, combinado con las imágenes de la trayectorias seguidas por los ligandos en cada una de las dinámicas realizadas con el programa Chimera arroja luz a cerca de en qué sitios de unión han interactuado los compuestos.

Analizando las dinámicas moleculares se observa que la proteína K-RAS tiene hasta cinco sitios de unión que pueden resultar interesantes: estos son, el *loop* 2, la hélice $\alpha 2$, la hélice $\alpha 3$, la hélice $\alpha 4$ y la hélice $\alpha 5$. Por tanto se realiza un análisis de cada uno de estos sitios de unión respecto de cada ligando, calculando la energía libre de unión (ΔG) de los complejos proteína-ligando formados mediante los métodos MMGBSA y MMPBSA.

De este modo, se ve que el ligando que tiene mayor afinidad por el *loop* 2 es CBT ya que la energía libre del complejo K-RAS-CBT formado por la unión del ligando a través de este *loop* es menor que las energías de los complejos formados por los ligandos DAP e ITC. De igual forma, el complejo que se forma con mayor facilidad con la hélice $\alpha 2$ es el formado por el ligando DAP, los complejos K-RAS-DAP y K-RAS-ITC que se forman a través de la hélice $\alpha 3$ tienen unas energías libres parecidas, pero menores que la del complejo K-RAS-CBT. A su vez, el ligando DAP es el que más interactúa con la hélice $\alpha 4$, mientras que los ligandos CBT e ITC forman complejos con energías similares más fácilmente con la hélice $\alpha 5$.

Se hace además un análisis de cada ligando, y su interacción con cada sitio de unión. Viendo los datos obtenidos de los cálculos MMPB/GBSA (Tabla 10) del ligando CBT se observa que los complejos formados por este ligando con el *loop* 2, así como con la

hélice $\alpha 5$ tienen una menor energía de enlace, por lo que es más probable que este compuesto interactúe con la proteína a través de estos sitios de unión. De igual modo, el ligando DAP se unirá preferentemente a las hélices $\alpha 2$ o $\alpha 4$, mientras que el ligando ITC lo hará con las hélices $\alpha 3$ o $\alpha 5$.

El hecho de que todos los ligandos tengan tendencia a unirse a todos los sitios de unión se debe a que todos los compuestos tienen estructuras muy parecidas y por lo tanto es de esperar que su modo de unión sea similar.

Por lo tanto, tanto en el análisis hecho tomando como referencia cada sitio de unión, como en el hecho tomando como referencia cada ligando, se llega a la misma conclusión.

7. Conclusiones

En primer lugar, es más efectivo el uso de la dinámica molecular clásica comparado con el uso de la dinámica molecular escalada.

Por otro lado, las gráficas de RMSD y RMSF demuestran que la proteína se mantiene estable durante los procesos, por lo que se pueden realizar los cálculos de manera fiable.

Finalmente, a la vista de los resultados obtenidos, los tres compuestos tienen buena afinidad por la proteína K-RAS aunque dependiendo de qué sitio de unión sea el que lleve a posterior inactivación de la proteína deberemos emplear uno u otro, por lo que en estudios posteriores se debería estudiar qué sitio de unión consigue desactivar la proteína de una manera más efectiva y usar el ligando que mejor interaccione con ese sitio de unión.

8. Bibliografía

1. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell*. 2000, 100, 57-70.
2. Chang EH, Gonda MA, Ellis RW, et al. Human genome contains four genes homologous to transforming genes of Harvey and Kirsten murine sarcoma viruses. *Proceedings of the National Academy of Sciences of the United States of America*. 1982, 79 (16), 4848–4852.
3. Der CJ, Cooper GM. Altered gene products are associated with activation of cellular ras genes in human lung and colon carcinomas. *Cell*. 1983, 32 (1), 201–208.
4. Parada LF, Weinberg RA. Presence of a Kirsten murine sarcoma virus ras oncogene in cells transformed by 3-methylcholanthrene. *Molecular & Cellular Biology*. 1983, 3 (12), 2298–2301.
5. Esser D, Bauer B, Wolthuis RMF, Wittinghofer A, Cool RH, Bayer P. Structure determination of the ras-binding domain of the Ral-specific guanine nucleotide exchange factor Rlf. *Biochemistry*. 1998, 37 (39), 13453–13462.
6. Zuber J, Tchernitsa OI, Hinzmann B, et al. A genome-wide survey of RAS transformation targets. *Nature Genetics*. 2000, 24 (2), 144–152.
7. Gilbert D, Rapley E, Shipley J. Testicular germ cell tumours: predisposition genes and the male germ cell niche. *Nature Reviews Cancer*, 2011, 11, 278-288.
8. Wennerberg K, Rossman KL, Der CJ. The Ras superfamily at a glance. *Journal of Cell Science*. 2005, 118 (5), 843–846.
9. Paduch M, Jeleń F, Otlewski J. Structure of small G proteins and their regulators. *Acta Biochimica Polonica*. 2001, 48 (4), 829–850.
10. McGrath JP, Capon DJ, Smith DH. Structure and organization of the human Ki-ras proto-oncogene and a related processed pseudogene. *Nature*. 1983, 304 (5926), 501–506.
11. Gideon P, John J, Frech M, et al. Mutational and kinetic analyses of the GTPase-activating protein (GAP)-p21 interaction: the C-terminal domain of GAP is not sufficient for full activity. *Molecular and Cellular Biology*. 1992, 12 (5), 2050–2056.
12. Giglione C, Parrini MC, Baouz S, Bernardi A, Parmeggiani A. A new function of p120-GTPase-activating protein: prevention of the guanine nucleotide exchange factor-stimulated nucleotide exchange on the active form of Ha-Ras p21. *The Journal of Biological Chemistry*. 1997, 272 (40), 25128–25134.

13. Scheffzek K, Ahmadian MR, Kabsch W, et al. The Ras-RasGAP complex: structural basis for GTPase activation and its loss in oncogenic ras mutants. *Science*. 1997, 277 (5324), 333–338.
14. Hancock JF, Prior IA. Electron microscopic imaging of Ras signaling domains. *Methods*. 2005, 37 (2), 165–172.
15. Prakash P, Hancock JF, Gorfe AA. Binding hotspots on K-Ras: consensus ligand binding sites and other reactive regions from probe-based molecular dynamics analysis. *Proteins*. 2015, 83 (5), 898-909.
16. Hancock JF, Magee AI, Childs JE, Marshall CJ. All ras proteins are polyisoprenylated but only some are palmitoylated. *Cell*. 1989, 57 (7), 1167–1177.
17. Rajalingam K, Schreck R, Rapp UR, Albert Š. Ras oncogenes and their downstream targets. *Biochimica et Biophysica Acta*. 2007, 1773 (8), 1177–1195.
18. Thissen JA, Gross JM, Subramanian K, Meyer T, Casey PJ. Prenylation-dependent association of Ki-Ras with microtubules. Evidence for a role in subcellular trafficking. *The Journal of Biological Chemistry*. 1997, 272 (48), 30362–30370.
19. Hingorani SR, Petricoin EF, III, Maitra A, et al. Preinvasive and invasive ductal pancreatic cancer and its early detection in the mouse. *Cancer Cell*. 2003, 4 (6), 437–450.
20. Fivaz M, Meyer T. Reversible intracellular translocation of KRAS but not HRas in hippocampal neurons regulated by C Ca^{2+} /calmodulin. *Journal of Cell Biology*. 2005, 170 (3), 429–441.
21. Wakamatsu N, Devereux TR, Hong H-HL, Sills RC. Overview of the Molecular Carcinogenesis of Mouse Lung Tumor Models of Human Lung Cancer. *Toxicologic pathology*. 2007, 35 (1), 75-80.
22. McCoy MS, Toole JJ, Cunningham JM, et al. Characterization of a human colon/lung carcinoma oncogene. *Nature*. 1983, 302 (5903), 79–81.
23. Kranenburg O. The KRAS oncogene: past, present, and future. *Biochimica et Biophysica Acta*. 2005, 1756 (2), 81–82.
24. Hegi ME, Devereux TR, Dietrich WF, et al. Allelotype analysis of mouse lung carcinomas reveals frequent allelic losses on chromosome 4 and an association between allelic imbalances on chromosome 6 and K-ras activation. *Cancer Research*. 1994, 54 (23), 6257–6264.

25. Barbacid M. Ras genes. *Annual Review of Biochemistry*. 1987, 56, 779–827.
26. Liao J, Wolfman JC, Wolfman A. K-Ras regulates the steady-state expression of matrix metalloproteinase 2 in fibroblasts. *The Journal of Biological Chemistry*. 2003, 278 (34), 31871–31878.
27. Yan Z, Deng X, Chen M, et al. Oncogenic c-Ki-ras but not oncogenic c-Ha-ras up-regulates CEA expression and disrupts basolateral polarity in colon epithelial cells. *The Journal of Biological Chemistry*. 1997, 272 (44), 27902–27907.
28. Severin Lupala C. Discovery of new selective antagonists of G-protein coupled receptors of therapeutic interest. *Universitat Politècnica de Catalunya- Doctoral Thesis; Barcelona* 2015, 34-61.
29. Schlick T. *Molecular Modeling and Simulation*. Springer. 2002, 480.
30. Gavathiotis E, Suzuki M, L. Davis M, et al. BAX activation is initiated at a novel interaction site. *Nature* 2008, 455, 1076-1081.
31. De Vivo M, Massetti M, Bottegoni G, et al. Role of Molecular Dynamics and Related Methods in Drug Discovery. *J. Med. Chem.* 2016, 59 (9), 4035-4061.
32. Gonzalez Ma. *Force fields and molecular dynamics simulations*. EDP Sciences. 2011, 12, 169–200.
33. Moore CC. Ergodic theorem, ergodic theory, and statistical mechanics. *Proceedings of the National Academy of Sciences of the United States of America*. 2015;112(7):1907-1911.
34. Sinko W, Miao Y, de Oliveira C. A. F, et al. Population Based Reweighting of Scaled Molecular Dynamics. *J. Phys. Chem.* 2013, 117 (42), 12759-12768.
35. Genheden S, Ulf Ryde. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities, *Expert Opin. Drug Discov*, 2015, Vol 10 (5), 449-461.
36. Onufriev A, Bashford D, Case DA. Exploring protein native states and large-scale conformational changes with a modified generalized Born model. *Proteins*, 2004, 55, 383–394.
37. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM Jr, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. "A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules". *J. Am. Chem. Soc.*, 1995, 117 (19), 5179–5197.

38. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. "Development and testing of a general AMBER force field". *Journal of Computational Chemistry*, 2004, 25, 1157-1174.
39. Maier JA, Martinez M, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.*, 2015, 11 (8), 3696–3713.
40. Maurer T, Garrenton LS, Oh A, Pitts K, Anderson DJ, Skelton NJ, Fauber BP, Pan B, Malek S, Stokoe D, Ludlam MJ, Bowman KK, Wu J, Giannetti AM, Starovasnik MA, Mellman I, Jackson PK, Rudolph J, Wang W, Fang G. Small-molecule ligands bind to a distinct pocket in Ras and inhibit SOS-mediated nucleotide exchange activity. *Proc.Natl.Acad.Sci.USA*, 2012, 109, 5299-5304.
41. <http://www.avogadro.cc>
42. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML, Comparison of simple potential functions for simulating liquid water, *J. Chem. Phys.*, 1983, 79, 926–935.
43. Darden T, York D, Pedersen L, Particle mesh Ewald: An $N \log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.*, 1993, 98, 10089–10092.
44. Uberuaga BP, Anghel M, Voter AF. Synchronization of trajectories in canonical molecular-dynamics simulations: observation, explanation, and exploitation. *J. Chem. Phys.*, 2004, 120, 6363–6374.
45. Sindhikara DJ, Kim S, Voter AF, Roitberg AE. Bad seeds sprout perilous dynamics: stochastic thermostat induced trajectory synchronization in biomolecules. *J Chem Theory Comput.*, 2009, 5, 1624–1631.
46. Ryckaert JP, Ciccotti G, Berendsen H. Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of nalkanes, *J. Comput. Phys.* 1977, 23, 327–341.
47. Berendsen HJC, Postma JPM, Gunsteren WF, DiNola VA, Haak JR. Molecular dynamics with coupling to an external bath, *J. Chem. Phys.*, 1984, 81, 3684–3690.
48. Onufriev A, Bashford D, Case DA. Exploring protein native states and largescale conformational changes with a modified generalized born model. *Protein Struct. Funct. Bioinf.*, 2004, 55, 383–394.

49. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem.* 2004, 25 (13), 1605-12.
50. Grant BJ, Lukman S, Hocker HJ, Sayyah J, Brown JH, McCammon JA, Gorfe AA. Novel allosteric sites on ras for lead generation. *PLoS One*, 2011, 6 (10), e25711.

9. Resumen de las signaturas cursadas

Seguridad en el laboratorio y experimentación animal (Obligatoria, 4 créditos ECTS)

En este curso hemos estudiado la dinámica de seguridad a seguir en un laboratorio, como las diferentes normas de seguridad y la normativa vigente que se aplica para mantener un nivel de seguridad adecuado en el laboratorio, cuyo objetivo trata de evitar accidentes. Además también se impartió la normativa que regula el manejo y la experimentación con animales de laboratorio y la experimentación humana (bioética).

Genómica, bioinformática y biología de sistemas (Optativa, 4 créditos ECTS)

En esta asignatura hemos estudiado las diferentes herramientas informáticas básicas para realizar un estudio molecular. Se utilizaron las diferentes bases de datos disponibles la web para realizar una recopilación de una serie de parámetros de gen elegido (la secuencia codificante, SNPs, ortólogos y parálogos, etc.). Entre las bases de datos utilizadas encontramos GenBank, NCBI, Expasy, etc.

Además también hemos utilizado diferentes programas y aplicaciones que nos permitieran realizar estudios de secuencias de nucleótidos y aminoácidos. Entre los que encontramos Bioedit, Clustal W2, Blast, etc.

Por otra parte también hemos utilizado diferentes paquetes estadísticos para procesar datos entre los que destacó R y el programa de modelización de sistemas complejos Vensim.

Regulación de la expresión génica (Optativa, 4 créditos ECTS)

En esta asignatura se trataron muchos temas que afectan a la expresión génica, tales como función y estructura de la ARN polimerasa, transcripción, etc. También se estudiaron el ADN repetitivo, ARN de interferencia y los transposones.

Ingeniería genética avanzada (Optativa, 4 créditos ECTS).

En este curso se profundizó en la metodología que permite modificar genéticamente células, organismos diversos o plantas. Para ello, se trató la base conceptual de éste conjunto de técnicas, su aplicación a la clonación y la obtención de organismos modificados genéticamente, dándose a conocer sus aplicaciones en biotecnología y biomedicina. Finalmente, se analizaron las implicaciones éticas de estas metodologías, y la normativa nacional y comunitaria vigente.

Biología molecular de enfermedades infecciosas (Optativa, 4 créditos ECTS)

En esta asignatura se estudiaron los diferentes agentes patógenos que causan enfermedades infecciosas, separándolo en dos bloques en función del agente que lo provocaras (virus y bacterias). En la parte teórica estudiamos los mecanismos moleculares por los cuales estos agentes son capaces de infectar y vivir en el hospedador y las estrategias replicativas virales. También el tratamiento de éstas y su repercusión a nivel mundial.

Biotecnología diagnóstica (Optativa, 4 créditos ECTS)

En este curso se han estudiado las herramientas biotecnológicas para la detección de de mutaciones tanto en el ámbito medico como en investigación criminal. También hemos estudiado las diferentes técnicas de citología: citometría de flujo, inmunoistoquímica PCR en tiempo real, etc. En la parte de secuenciación masiva estudiamos los tipos de secuenciadores que existen en el mercado y cuales se ajustan mejor a ciertas necesidades a la hora de secuenciar distintos tipos de muestras. En la parte práctica del curso se realizó un ensayo inmunohistoquímico en muestras control y de cáncer de mama y una PCR en tiempo real.

Biotecnología industrial y bioprocesos (Optativa, 4 créditos ECTS).

En esta asignatura se ha estudiado el concepto de bioproceso y su desarrollo en el campo de la biotecnología. También estudiamos los conceptos de biorreactores, biología y bioquímica de microorganismos, para conformar procesos completos de transformación biológica de materias primas en productos. Se incluyeron la integración de procesos y equipos, su instrumentación y control, así como el escalado de los mismos, como herramientas imprescindibles para la puesta en marcha real de cualquier proceso. Además se dio una visión aplicada de estos procesos a la industria.

Patología molecular y celular. (Optativa, 4 créditos ECTS)

En esta asignatura se ha estudiado la base biológica y la patología de enfermedades humanas, como el cáncer donde se estudiaron sus mecanismos moleculares, oncogenes, genes supresores de tumores, angiogénesis y muerte celular. También se estudiaron patología provocada por el estrés oxidativo y nitrosativo. Por otra parte su tratamiento (marcadores de tumores, quimioterapia) y terapia génica.

Biología molecular y celular del envejecimiento (Optativa, 4 créditos ECTS)

En esta asignatura hemos estudiado las bases moleculares y celulares que conducen a la degeneración y muerte celular provocada por el envejecimiento. En esta materia hemos estudiado las bases genéticas y ambientales culpables del deterioro progresivo, ilustrado por ejemplos. También se estudiaron diferentes enfermedades neurodegenerativas como el Parkinson, el Alzheimer y enfermedad de Huntington.

Biotecnología de producción y purificación de proteínas recombinantes (4 créditos ECTS)

La asignatura enseña la metodología de producción y purificación de proteínas recombinantes y nativas, tanto para el uso en investigación como para sus posibles aplicaciones biotecnológicas, ya sean producción de anticuerpos, hormonas, factores de crecimiento, etc. Se proporcionan conocimientos generales sobre la purificación de proteínas, técnicas para la preparación de las muestras biológicas, métodos de purificación por cromatografía en FPLC, producción masiva de proteínas mediante 58 ADN recombinante y estrategias cromatográficas para la purificación de proteínas recombinantes y purificación de anticuerpos.

Proteómica y dinámica de proteínas (4 créditos ECTS)

En esta asignatura se estudian los fundamentos, técnicas y aplicaciones de la proteómica. Se empieza impartiendo conocimientos sobre el plegamiento, transporte, recambio, maduración, degradación de proteínas e interacciones de las proteínas con su ligando. A continuación, se proporcionan conceptos generales sobre técnicas proteómicas de separación, identificación y cuantificación, así como la aplicación de la proteómica en el diagnóstico de enfermedades, descubrimiento de fármacos, industria y agricultura.

Enzimología aplicaciones biotecnológicas y clínicas (4 créditos ECTS)

En esta asignatura se estudian los fundamentos teóricos y prácticos de la enzimología y las aplicaciones de las enzimas en biotecnología y biomedicina. Aunque se encuentra dentro del módulo Biotecnología, puede considerarse una asignatura común a los dos módulos Biotecnología y Biomedicina. En la primera parte del curso se profundiza en la teoría y práctica de las determinaciones enzimáticas. En una segunda parte se estudia la utilización de las enzimas en la industria biotecnológica, tanto a nivel agroalimentario como industrial. En una tercera parte se estudia la aplicación de las enzimas en la clínica, su uso en el diagnóstico de enfermedades y como dianas terapéuticas de fármacos.