

Tesis de Grado

Caracterización estructural y docking molecular en el receptor cannabinoide 2

Spinosa, Mariel Alejandra

2017

Este documento forma parte de las colecciones digitales de la Biblioteca Central Dr. Luis Federico Leloir, disponible en bibliotecadigital.exactas.uba.ar. Su utilización debe ser acompañada por la cita bibliográfica con reconocimiento de la fuente.

This document is part of the digital collection of the Central Library Dr. Luis Federico Leloir, available in bibliotecadigital.exactas.uba.ar. It should be used accompanied by the corresponding citation acknowledging the source.

Cita tipo APA:

Spinosa, Mariel Alejandra. (2017). Caracterización estructural y docking molecular en el receptor cannabinoide 2. Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires. https://hdl.handle.net/20.500.12110/seminario_nBIO001618_Spinosa

Cita tipo Chicago:

Spinosa, Mariel Alejandra. "Caracterización estructural y docking molecular en el receptor cannabinoide 2". Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires. 2017. https://hdl.handle.net/20.500.12110/seminario_nBIO001618_Spinosa

EXACTAS UBA

Facultad de Ciencias Exactas y Naturales



UBA

Universidad de Buenos Aires

Universidad de Buenos Aires
Facultad de Ciencias Exactas y Naturales



Carrera de Ciencias Biológicas

Tesis de Licenciatura

**Caracterización estructural y *docking* molecular
en el receptor cannabinoide 2**

Mariel Alejandra Spinosa

Director de tesis: Claudio N. Cavasotto

Lugar de trabajo: Instituto de Investigación en Biomedicina de Buenos Aires – CONICET – Partner Institute of the Max Planck Society

Buenos Aires, Junio del 2017

Agradecimientos

A los docentes de la FCEyN que han formado parte de cada una de las materias que cursé a lo largo de la carrera, por su compromiso con la educación pública y la formación de profesionales altamente capacitados. Gracias por su generosidad a la hora de transmitir cada uno de sus conocimientos.

A Claudio Cavasotto por permitirme formar parte de su grupo de investigación y por guiarme a lo largo del trabajo de tesis. Además, por su paciencia y buena predisposición a la hora de enseñar.

A mis compañeros de grupo, Damián, Leonardo, Gabriela, Natalia y Romina, por sus consejos y su ayuda durante el desarrollo del trabajo, dedicándome su tiempo. Gracias también a todos los que forman parte del IBioBa por haber hecho mi paso por allí tan agradable.

A mis compañeros de facultad, Conrado, Fede, Belu, Emi, Pau, Meli, Andre, Gabi, Vale, Sofi, y otros, que han compartido conmigo tantas horas de cursada, de estudio en la biblioteca, mates, almuerzos, meriendas, charlas, noches en el bar de deportes y tardes descansando al sol. Gracias por aguantarme, sobre todo en mis momentos de nervios previos a rendir, y por estar presentes siempre que necesité su ayuda.

A mis amigas de toda la vida, Fran, Estefi, Cele, Fer y Nahir, por estar a mi lado desde hace tantos años. Por alegrarse con mis logros y por su contención en los momentos más difíciles. Cada una significa mucho para mí.

A Federico, por haberme acompañado durante estos más de 6 años y haber sido mi gran sostén. Por prestarme su oído y darme las palabras justas. Nunca dudaste de mí ni de mis capacidades, gracias por alentarme a perseguir mis sueños.

A mi familia, principalmente a mis padres, Marta y Alberto, y a hermana Magalí, por su apoyo incondicional y por haberme tolerado durante los días de estudio con mis distintos ánimos. Sin su ayuda no hubiera llegado hasta esta instancia. Son indispensables para mí y les agradezco por siempre haberme dado todo lo que estuviera a su alcance para hacer más llevadero este largo camino.

Índice

Abreviaciones	3
Resumen	4
Introducción	5
Receptores acoplados a proteína G.....	5
Relevancia terapéutica del receptor cannabinoide 2	7
Modelización estructural <i>in silico</i> por homología.....	10
Caracterización de la interacción receptor-ligando.....	12
Objetivos	14
Materiales y métodos	15
Compilación de datos de mutagénesis	15
Modelado por homología	15
I. Alineamiento de secuencias	16
II. Obtención de modelos estructurales de CB2.....	19
Generación de GLL/GDD	22
<i>Docking</i> molecular de proteína rígida.....	23
Resultados	25
Compilación de datos de mutagénesis	25
Alineamiento de secuencias	27
Modelado por homología	29
GLL/GDD.....	35
<i>Docking</i> molecular de pequeña escala para la validación de los modelos	35
Discusión	38
Importancia de la resolución de estructuras experimentales de GPCRs	38
Limitaciones del modelado por homología.....	41
Aminoácidos importantes en el sitio ortostérico de CB2.....	42
Caracterización de la interacción CB2-antagonista/ agonista inverso	42
Caracterización de la interacción CB2-agonista	44
<i>Docking</i> automatizado y validación de los modelos de CB2	49
Conclusiones	54
Referencias	56

Abreviaciones

2-AG: 2-araquidonilglicerol

A2AAR: receptor adenosina A_{2A}

ADRB1: receptor adrenérgico β_1

ADRB2: receptor adrenérgico β_2

CB1: receptor cannabinoide 1

CB2: receptor cannabinoide 2

C-Term: lazo C-terminal

DA: *docking* automatizado

ECL: lazo extracelular (*extracellular loop*)

EF: factor de enriquecimiento

GDD: base de datos de señuelos de GPCR (*GPCR decoy database*)

GLL: biblioteca de ligandos de GPCR (*GPCR ligand library*)

GPCR: receptor acoplado a proteína G

HTD: *high-throughput docking*

HTS: *high-throughput screening*

ICL: lazo intracelular (*intracellular loop*)

LPA1: receptor de ácido lisofosfatídico 1

LSDH: *Ligand-Steered Homology Modelling*

N-Term: lazo N-terminal

S1P1: receptor 1 de esfingosina 1-fosfato

THC: tetrahidrocannabinol

TM: dominio transmembrana

Resumen

En el presente trabajo de tesis, dada la relevancia terapéutica del receptor cannabinoide 2 (CB2) y la falta de una estructura experimental del mismo, se tiene como objetivo el desarrollo de nuevos modelos de CB2 en sus conformaciones agonista y antagonista/ agonista-inverso, junto con su validación en base a datos de mutagénesis y *docking* molecular automatizado. Para lograr dicho objetivo se realizó un modelado por homología basado en un alineamiento de secuencias usando como plantillas estructuras experimentales recientemente resueltas de receptores acoplados a proteína G (GPCRs), teniendo en cuenta el estado conformacional de las mismas (activo o inactivo). La caracterización de la interacción receptor-ligando se llevó a cabo mediante la compilación de datos de mutagénesis sitio-dirigidas del receptor CB2, mientras que la validación de los modelos obtenidos se realizó mediante un *docking* automatizado. Como resultado se obtuvo un modelo de CB2 en su conformación inactiva, utilizando como plantilla el receptor cannabinoide 1 (CB1), unido al agonista inverso SR-144,528, y otros dos modelos de CB2 en su conformación activa, utilizando un abordaje con múltiples plantillas (CB1 y receptor receptor adrenérgico β_1), unido al agonista WIN 55,212-2.

Además, se filtró una base de datos obtenida a partir de un servidor online para construir una biblioteca de ligandos de CB2, obteniendo un total de 996 agonistas, 34 agonistas inversos y 12 antagonistas, que fueron más tarde empleados para la generación de una base de datos señuelo (GDD) para CB2. Tanto la GDD como la biblioteca de ligandos de CB2 (GLL) se utilizaron en el proceso de *docking* automatizado.

Los modelos agonistas así como el modelo antagonista/ agonista inverso obtenidos en este trabajo, cumplen la mayoría de los datos de mutagénesis recolectados, permitiendo una buena caracterización de la pose del ligando y los residuos involucrados en el sitio de unión. En cuanto a la validación de los modelos construidos, evaluada a partir del *docking*, se obtuvieron resultados similares al de los modelos anteriores desarrollados en el grupo (año 2009 y 2013), aunque estos satisfacían los resultados de mutagénesis en menor grado. Dichos modelos eran lo suficientemente precisos para diferenciar ligandos de no ligandos, a pesar de utilizar como moldes las estructuras de GPCRs más lejanos filogenéticamente.

Introducción

Receptores acoplados a proteína G

Los receptores acoplados a proteína G (GPCRs por sus siglas en inglés: *G Protein-Coupled Receptors*) son proteínas integrales de membrana que reconocen numerosos mensajeros como fotones, odorantes, neurotransmisores, ácidos grasos, iones, hormonas y péptidos (Pierce, Premont and Lefkowitz, 2002). Estos estímulos son trasladados como respuestas intracelulares que resultan en comportamientos celulares precisos. El proceso de señalización de los GPCRs está vinculado a varias respuestas fisiológicas y patológicas que afectan el crecimiento y diferenciación celular, el metabolismo, el sistema nervioso, inmune, cardiovascular, endocrino, entre otros (Schöneberg et al., 2004; Li, Huang and Peng, 2005; Lania, Mantovani and Spada, 2006). Es por este motivo que son un blanco atractivo en el descubrimiento de drogas, pudiendo aportar tratamientos novedosos y efectivos a distintas enfermedades neurodegenerativas, oncológicas, cardiovasculares, inmunes y metabólicas (Lappano and Maggiolini, 2011; Blad, Tang and Offermanns, 2012; Guerram, Zhang and Jiang, 2016). Los GPCRs representan la familia con mayor número de *targets* de drogas en el mercado: más del 60% de las drogas existentes actualmente tienen como blanco un GPCR (McCusker et al., 2007).

Con al menos 800 miembros en humanos (Fredriksson et al., 2003), la superfamilia de GPCRs se encuentra clasificada en cinco familias principales (Lagerström and Schiöth, 2008): familia rodopsina (o clase A), la cual es la más numerosa con aproximadamente 300 miembros, familia de secretinas, familia de adhesinas (junto con secretinas conforman la clase B), familia glutamato (clase C) y familia frizzled/sabor (clase F) (Figura 1). La resolución de la estructura cristalina de rodopsina bovina de alta calidad en el año 2000 (Palczewski et al., 2000) confirmó una serie de características estructurales de estos receptores (Venkatakrisnan et al., 2013). Estructuralmente los GPCRs están compuestos por siete α -hélices que atraviesan la membrana celular, conocidos como dominios transmembrana (TMs), las cuales están conectadas por tres lazos intracelulares (ICLs) y tres extracelulares (ECLs) (Figura 2). El extremo N-terminal está ubicado en el lado extracelular y el C-terminal, el cual posee una α -hélice, en el lado intracelular. Los dominios que están cercanos al medio extracelular son generalmente los involucrados en la unión con el ligando, mientras que las regiones intracelulares son las que interactúan de forma directa con proteínas G.

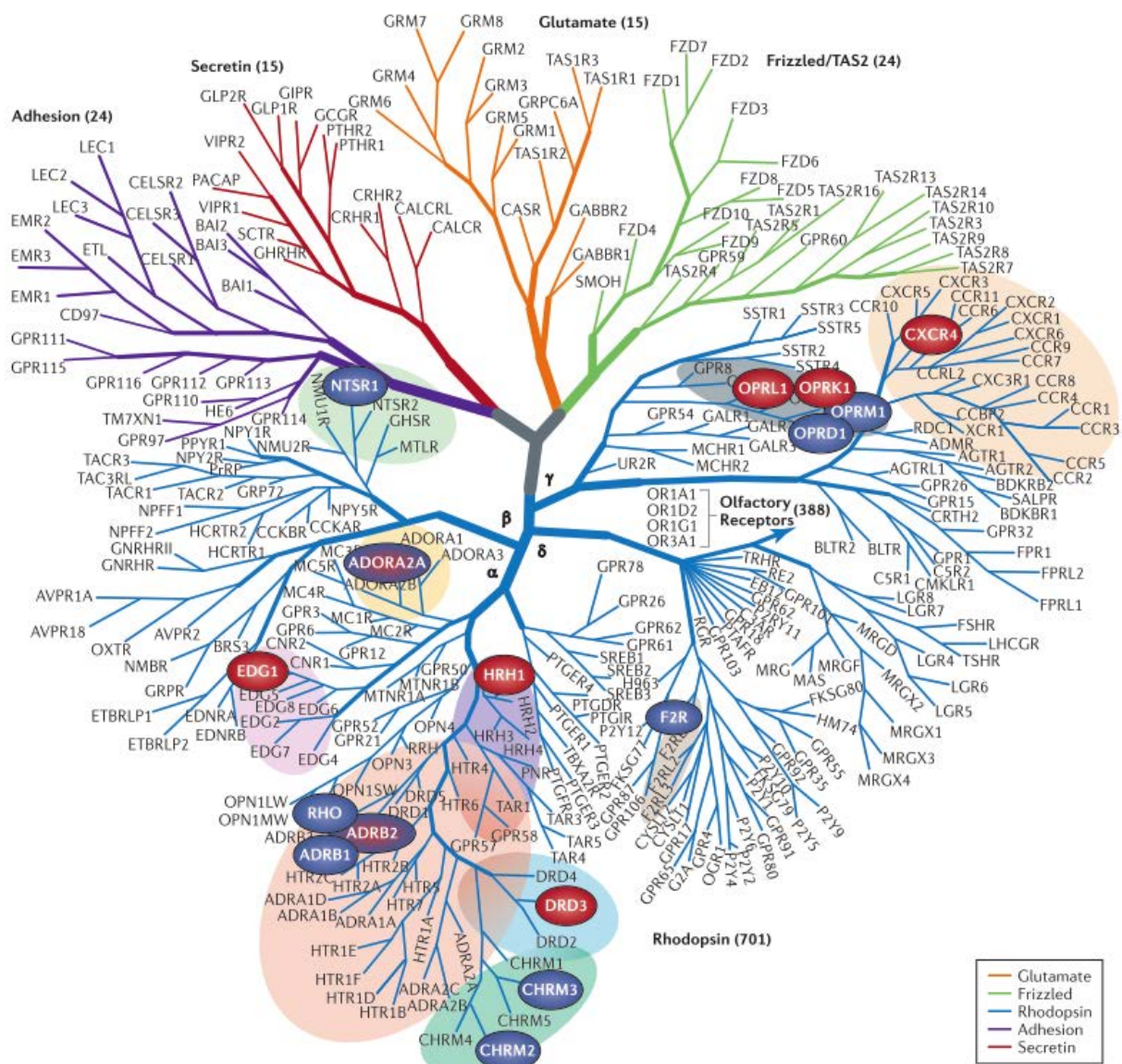


Figura 1. Representación del árbol filogenético de la superfamilia de GPCRs, distinguiéndose las cinco familias: rodopsina, secretinas, adhesinas, glutamato y frizzled/sabor. (De Stevens, R.C. et al. (2013) GPCR Network: A Large-Scale Collaboration on GPCR Structure and Function. *Nature reviews. Drug discovery*, **12**, 25–34).

Los ligandos pueden inducir o estabilizar diferentes estados conformacionales de los dominios, los cuales disparan cascadas de señalización intracelulares controladas por proteínas G. La función de estas proteínas G heterotriméricas está relacionada con la habilidad de la subunidad alfa de cambiar entre una conformación inactiva (unida a GDP) y una activa (unida a GTP) que regula la actividad río abajo de proteínas efectoras (Oldham and Hamm, 2008). En ausencia de un ligando activador, los GPCRs usualmente muestran una actividad basal que es incrementada al unirse un agonista (completo o parcial) (Seifert et al., 2001), reducida al unirse un agonista inverso e

inalterada por un antagonista, el cual bloquea tanto la acción del agonista como del agonista inverso (Rosenbaum, Rasmussen and Kobilka, 2009). También pueden ser modulados por ligandos alostéricos que se unen a un sitio distinto del ortostérico (sitio de unión natural del ligando) o ligandos biotrópicos, los cuales tienen la habilidad de unirse a ambos sitios (ortostérico y alosterico) (Christopoulos and Kenakin, 2002; Valant, Sexton and Christopoulos, 2009).

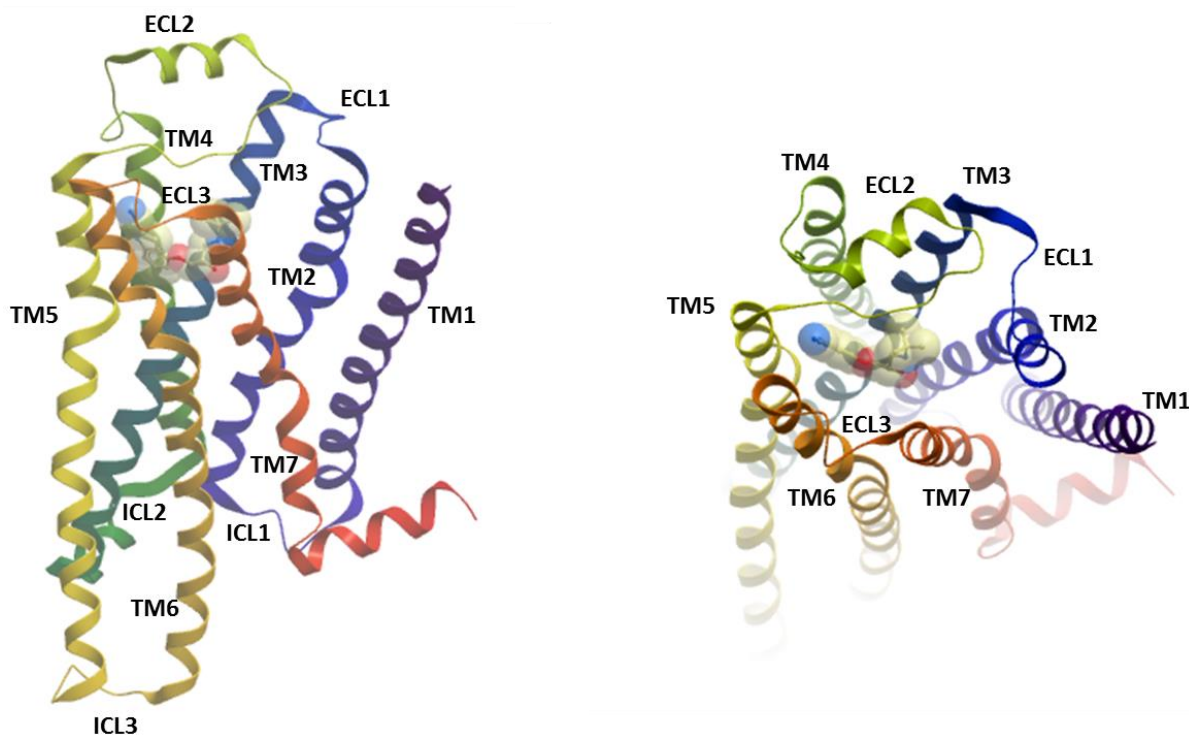


Figura 2. Representación de la estructura 3D de un GPCR. En la misma se muestran las siete α -hélices (TM1-TM7) coloreadas de violeta a rojo, los lazos intracelulares (ICLs) y los lazos extracelulares (ECLs). Además se exhibe el ligando ubicado en el sitio de unión ortostérico, generalmente delimitado por las hélices 3, 5, 6 y 7.

Relevancia terapéutica del receptor cannabinoide 2

El sistema de señalización endocannabinoide en mamíferos está conformado por los mensajeros lipídicos endógenos, anandamida y 2-araquidonilglicerol (2-AG), y por dos GPCRs homólogos (Piomelli, 2003). El receptor cannabinoide 1 (CB1) fue descubierto, clonado y expresado en primer lugar, en el año 1990, y es el GPCR más expresado en el cerebro (Herkenham et al., 1990; Matsuda et al., 1990). Se localiza no sólo en el sistema nervioso central sino también a lo largo del cuerpo. Tres años después fue caracterizado el receptor cannabinoide 2 (CB2) el cual se expresa principalmente en células del sistema inmune y sanguíneas (Munro, Thomas and Abu-Shaar, 1993). Estos receptores además de ser activados por sus ligandos endógenos pueden ser

activados por compuestos naturales (como es el caso de la molécula THC) y por cannabinoides sintéticos (Figura 3). El sistema cannabinoide se encuentra bajo estudio en los últimos años por estar involucrado en una variedad de procesos fisiológicos incluyendo apetito, sensación de dolor, estado de ánimo y memoria (Guindon and Hohmann, 2009; Bermudez-Silva et al., 2010; Micale et al., 2013).

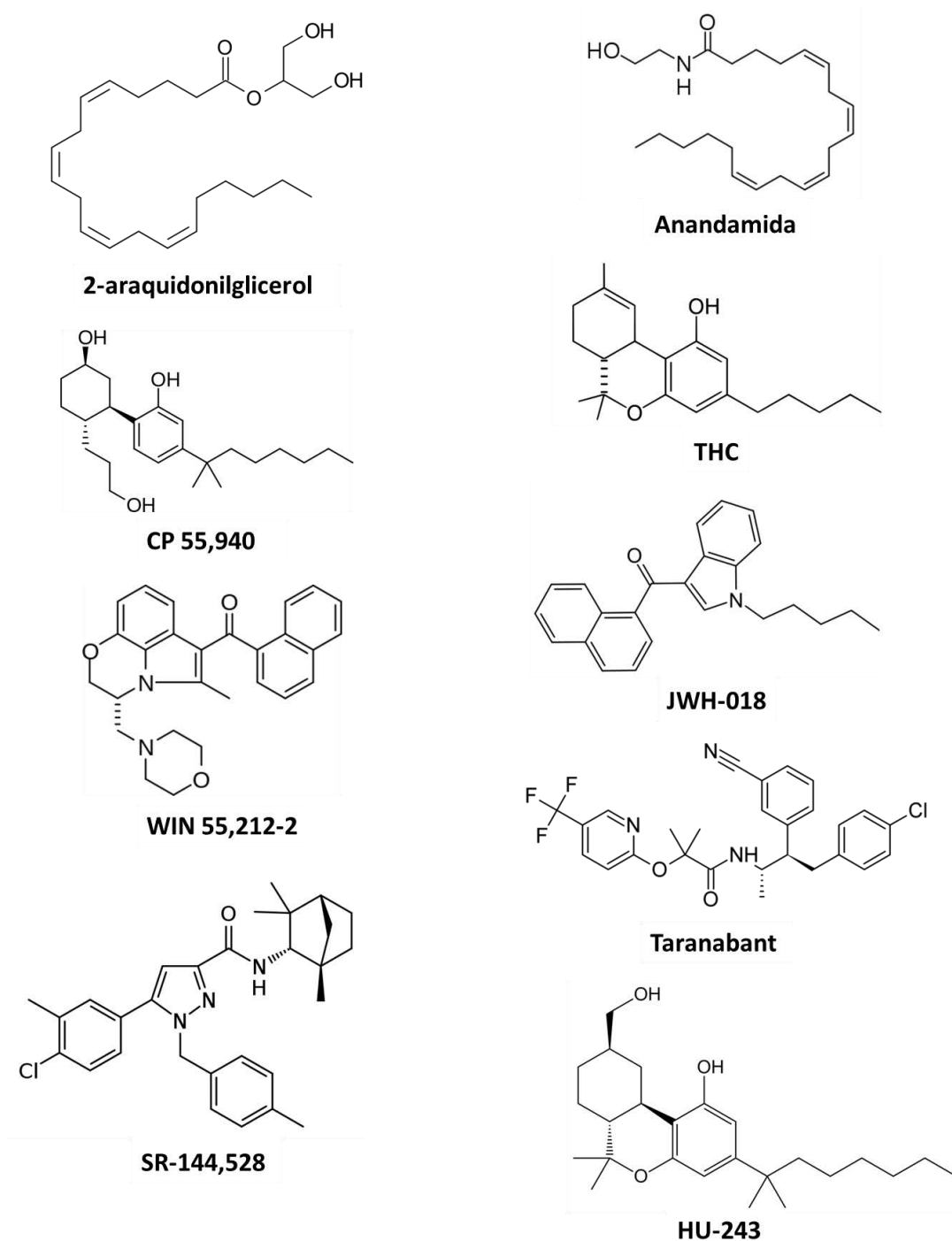


Figura 3. Estructuras químicas de moléculas que se unen a receptores cannabinoide. Las moléculas anandamida y 2-araquidonilglicerol son ligandos endógenos. El tetrahidrocannabinol (THC) es un compuesto natural. Las moléculas CP 55,940, SR-144,528, JWH-018, HU-243, Taranabant y WIN 55,212-2 son ligandos sintéticos.

La obesidad es una enfermedad compleja asociada con un aumento del riesgo de resistencia a la insulina, diabetes tipo 2, hipertensión, enfermedades cardiovasculares e hígado graso. Los niveles de endocannabinoides están aumentados en pacientes obesos comparados con pacientes delgados (Di Marzo et al., 2001). Con el fin de validar la inhibición de la activación de los receptores cannabinoides como un enfoque terapéutico válido para esta enfermedad, es necesario diseñar inhibidores específicos de estos receptores que no produzcan efectos secundarios graves como en el caso de los agonistas inversos Rimonabant y Taranabant (Leite et al., 2009). A su vez, los receptores cannabinoides son posibles blancos para el tratamiento de distintos tipos de cáncer: se han llevado a cabo números estudios donde se obtuvo como resultado una disminución en la proliferación de tumores al ser sometidos a agonistas específicos de CB1/CB2 (De Petrocellis et al., 1998; Qamri et al., 2009; Guindon and Hohmann, 2011). Sin embargo, el potencial terapéutico de los cannabinoides para tratamiento del cáncer se encuentra aún bajo discusión y en proceso de estudios clínicos.

Por otro lado, la activación selectiva de receptores CB2 ha sido propuesta como una estrategia para reducir los efectos secundarios negativos sobre el sistema nervioso central, vistos con agonistas CB1/CB2 no selectivos debido a que CB1 se halla principalmente en este sistema (Ashton, Wright and Tyndall, 2008). Varios agonistas selectivos de CB2 han sido descritos anteriormente (Diaz et al., 2009; Petrov et al., 2013). A pesar de que múltiples estudios preclínicos sugieren que el receptor CB2 es un blanco viable para disminuir tanto las respuestas del dolor agudo como del neuropático (Naguib et al., 2008; Xu et al., 2010), agonistas sintéticos de éste no han avanzado a través de ensayos clínicos. Esto se debe, en parte, a la falta de un completo conocimiento del mecanismo analgésico mediado por CB2 (Atwood et al., 2012). La regulación de CB2 también está implicada en inmunomodulación y neuroprotección, pero el perfil funcional de los ligandos de CB2 induciendo estos efectos no ha sido claramente definido (Cabral and Griffin-Thomas, 2009). La disociación de las subunidades α y $\beta\gamma$ de la proteína G (clase G α_i/o) resultante de la activación de CB2, puede influenciar múltiples sistemas efectores incluyendo adenilato ciclasa, p42/44 MAPK (ERK1/2) o canales iónicos (señalización mediada por Ca²⁺) (Figura 4) (Díaz-Laviada and Ruiz-Llorente, 2005).

Dada la relevancia terapéutica de este receptor, y la necesidad de diseñar potentes ligandos (agonistas y antagonistas/ agonistas inversos), es de suma importancia conocer su estructura y caracterizar el sitio de unión para poder avanzar en proyectos de diseño de fármacos líderes basados en la estructura del receptor.

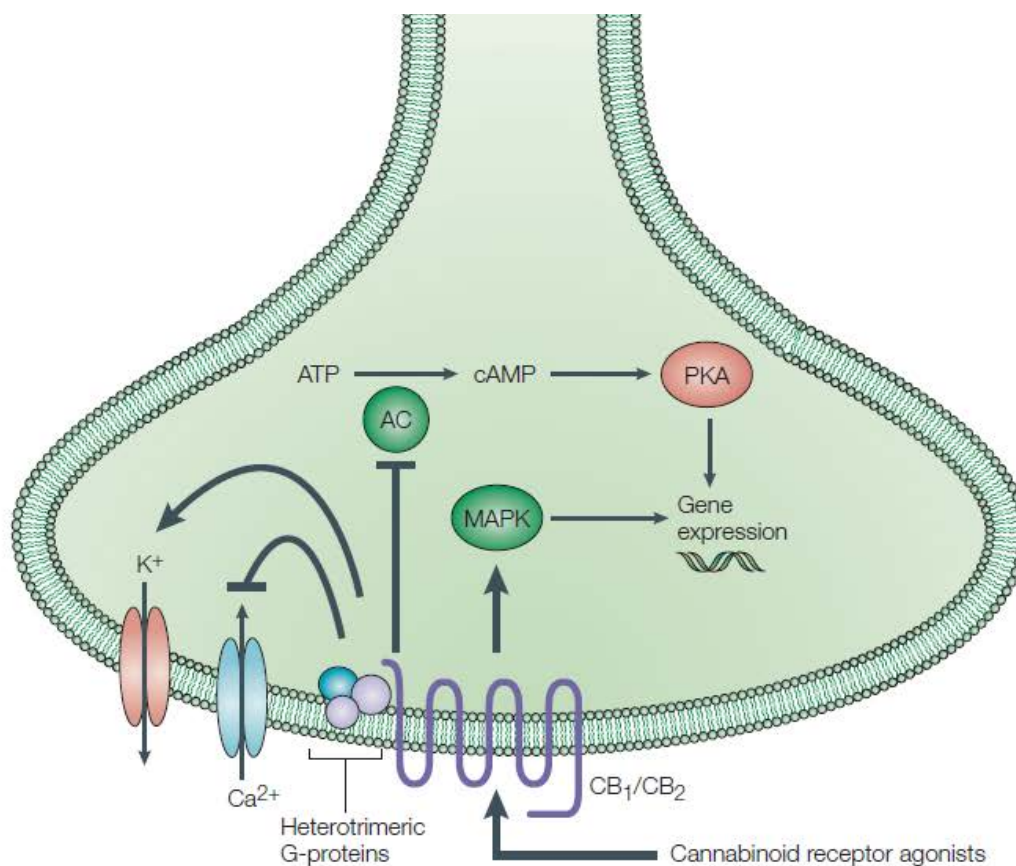


Figura 4. Representación de las principales vías de señalización asociadas con la activación del receptor cannabinoide mediada por agonistas. La activación de ambos receptores, CB1 y CB2, y la posterior estimulación de proteínas $G_{\alpha/\beta}$ heterotriméricas es conocida por estar acoplada a la inhibición de adenilato ciclasa (AC) y a la subsecuente inactivación de la proteína quinasa A (PKA) junto con la estimulación de la proteína quinasa activada por mitógenos (MAPK). Estos eventos intracelulares conducen, entre otros efectos, a la regulación de la expresión de varios genes. Por otro lado, la estimulación de proteínas $G_{\alpha/\beta}$ mediante la activación de CB1 puede llevar a la inhibición de canales de Ca^{2+} voltaje dependientes y a la estimulación de canales rectificadores de K^+ en neuronas. (De Di Marzo, V., Bifulco, M. and De Petrocellis, L. (2004) The endocannabinoid system and its therapeutic exploitation. *Nature reviews. Drug discovery*, **3**, 771–784).

Modelización estructural *in silico* por homología

Con el fin de incorporar la información disponible sobre la estructura del receptor en el proceso de diseño de moduladores y fármacos líderes, se comenzaron a utilizar las estructuras 3D de los complejos proteína-ligando para guiar la optimización de moléculas líderes en aras de aumentar su potencia y selectividad (Congreve, Murray and Blundell, 2005). Esto fue seguido por el desarrollo de métodos computacionales con distintos fines: caracterizar interacciones ligando-proteína, cribado virtual de

bibliotecas químicas (quimiotecas) usando *docking* automatizado (DA) o *high-throughput docking* (HTD) de ligando-receptor (Cavasotto and Orry, 2007), y optimización de moléculas líderes. Esto proveyó una alternativa más racional, económica y rápida que el tradicional *high-throughput screening* (HTS) (Phatak, Stephan and Cavasotto, 2009), y hoy día el diseño de moléculas líderes asistido por computadora está consolidado como el primer paso en el largo proceso de desarrollo de nuevos fármacos (Jorgensen, 2004; Shoichet, 2004).

A pesar de su relevancia terapéutica, los GPCRs representan un gran desafío al momento de su expresión, purificación y cristalización, lo cual explica la baja cantidad de estructuras de alta resolución resueltas a través de las distintas familias. En los últimos años, se han resuelto con mucha dificultad y gran inversión las estructuras cristalinas de aproximadamente 25 GPCRs, lo que expande el horizonte de diseño de fármacos líderes basado en la estructura del receptor. Los receptores cannabinoides pertenecen a la Clase A de GPCRs; muy recientemente (fines del año 2016) ha sido resulta experimentalmente la estructura del receptor cannabinoide 1 (Hua et al., 2016; Shao et al., 2016) mientras que la del receptor cannabinoide 2 no ha sido resuelta aún.

Dado que un elemento imprescindible para el uso de métodos basados en la estructura del receptor es la disponibilidad de la estructura 3D de la proteína diana (Cavasotto, 2011; Salon, Lodowski and Palczewski, 2011), eventualmente en complejo con uno o varios ligandos, ante la ausencia de estructuras experimentales el modelado por homología es la única alternativa viable para obtener una representación 3D de la proteína (Cavasotto and Phatak, 2009; Cavasotto and Palomba, 2015). A partir de la estructura experimental de una proteína (plantilla), se pueden generar modelos estructurales para una secuencia homóloga (blanco) que tenga un ~30% o más de semejanza de secuencia, o que tenga reconocida semejanza estructural (por ejemplo, los GPCRs de clase A tienen una topología similar de siete α -hélices transmembrana, a pesar de la baja semejanza de secuencia entre sus miembros). El proceso de modelado por homología consiste en los siguientes pasos (Cavasotto and Phatak, 2009; Cavasotto, 2011; Palomba and Cavasotto, 2015):

- 1) **Identificación de proteínas con estructuras experimentales 3D relacionadas que puedan usarse como plantillas.** Cuanta más similitud de secuencia haya entre la proteína blanco y la plantilla de estructura conocida, mejor será la calidad del modelo obtenido como resultado. En el caso de no contar con la estructura de una proteína que presente un alto grado de homología con el blanco a modelar, la similitud estructural global por pertenecer a una misma familia puede ser un punto a tener en cuenta al momento de elegir el molde. También es importante tener presente el estado funcional del

receptor: si se quiere obtener un modelo activo es más adecuado buscar una estructura 3D que se haya determinado en presencia de un agonista y, en el caso que se quiera un modelo inactivo, será mejor la estructura de un receptor que se haya cristalizado asociado a un agonista inverso. Esto se debe a que, tras la unión del ligando, muchos sistemas experimentan cambios conformacionales que van desde movimientos locales de cadenas laterales hasta grandes movimientos de dominio.

- 2) **Alineamiento de las secuencias de la plantilla y el blanco.** Este paso asigna a cada uno de los aminoácidos de la proteína blanco el correspondiente en la secuencia plantilla. Muchas veces pueden aparecer en las secuencias inserciones o deleciones, en estos casos no habrá aminoácidos que se correspondan entre ambas secuencias. Dicho alineamiento puede realizarse de a pares o de forma múltiple, donde varias secuencias son alineadas de forma simultánea con el fin de identificar regiones conservadas, predecir sitios funcionales o realizar análisis filogenéticos (Edgar and Batzoglou, 2006).
- 3) **Construcción de un modelo estructural basado en la estructura 3D de la plantilla y el alineamiento.** El modelo crudo puede construirse a partir de una única plantilla (*single-template*) o a partir de varias (*multiple-template*). En el segundo abordaje, el blanco es dividido en varios segmentos y diferentes moldes son utilizados para modelar cada uno de ellos (Larsson et al., 2008).
- 4) **Refinado, validación y evaluación de los modelos.** El refinamiento del modelo es indispensable a fin de obtener una caracterización más precisa del sitio de unión y de determinar con menor error la conformación de las cadenas laterales o residuos. La validación puede llevarse a cabo mediante un *docking* retrospectivo, el cual consiste en tomar una biblioteca de ligandos mezclados con moléculas señuelos (es decir, que no son ligandos de la proteína blanco) y comprobar la habilidad de mi modelo para priorizar las moléculas que son ligandos. Para la evaluación de los modelos pueden utilizarse datos experimentales como por ejemplo datos de mutagénesis sitio dirigidas.

Estos pasos se pueden repetir hasta obtener un modelo satisfactorio (Martí-Renom et al., 2000).

Caracterización de la interacción receptor-ligando

A pesar de compartir una estructura global de siete hélices transmembrana, cada GPCR presenta características estructurales únicas como por ejemplo la plasticidad de las hélices, los aminoácidos presentes en el sitio de unión, diversidad en la posición de

los bolsillos de unión de ligandos y variaciones en los ECLs e ICLs (Katritch, Cherezov and Stevens, 2012). Estas diferencias estructurales pueden ser utilizadas para el diseño de compuestos selectivos de ciertas familias de GPCRs (Congreve and Marshall, 2010), para lo cual es necesario una mayor cantidad de estructuras cristalinas resueltas y el desarrollo de mejores métodos de modelado de estructuras 3D.

Aunque el modelado por homología provee modelos aceptables del receptor, se requieren métodos de refinado para obtener una caracterización más exacta del sitio de unión, ya que pequeños errores pueden comprometer el modelo para uso en DA. En el grupo del Dr. Cavasotto se desarrolló con este propósito el método *Ligand-Steered Homology Modelling* (LSHM), en el cual la información de ligandos es considerada explícitamente para dar forma y optimizar el sitio de unión a través de un algoritmo estocástico basado en *docking* para la minimización global de la energía (Cavasotto and Abagyan, 2004; Cavasotto, Kovacs and Abagyan, 2005; Cavasotto et al., 2008; Phatak, Gatica and Cavasotto, 2010). Es importante destacar que siempre que sea posible, conviene incorporar información extraída de experimentos sobre interacciones específicas entre el receptor y el ligando, como la extraída de datos de mutagénesis; esta información se puede incorporar en el proceso de modelado, por ejemplo, en forma de ligaduras con un potencial armónico. Estas restricciones geométricas ayudan a disminuir los grados de libertad del complejo receptor-ligando. El LSHM es particularmente útil para reducir la incerteza en el modelado del sitio de unión, ya que tanto el ligando como el receptor son considerados flexibles en la etapa de modelado. El método LSHM fue originalmente desarrollado y evaluado en el receptor 1 de la hormona concentradora de melanina (MCHR1), un GPCR blanco de drogas contra la obesidad y sin estructura experimental. Los modelos construidos se usaron en un DA a gran escala seguido por la evaluación experimental en ensayos de competición de las moléculas más promisorias, exhibiendo un rendimiento 10 veces mayor que el del tradicional HTS (Cavasotto et al., 2008). Este método fue aplicado al estudio de la interacción CB2-antagonista/agonistas-inversos (Diaz et al., 2009a) y CB2-agonistas (Diaz et al., 2009b; Petrov et al., 2013). Los modelos de los complejos se utilizaron para guiar la optimización de ligandos y su posterior síntesis química realizada en el laboratorio del Dr. Philippe Diaz (University of Montana, USA).

Considerando entonces: i) la relevancia terapéutica del receptor CB2, ii) la carencia de una estructura cristalina de tal receptor, iii) la aparición reciente de nuevas estructuras experimentales de GPCRs que pueden usarse como plantilla para modelado por homología, se plantea en este trabajo de tesis desarrollar nuevos modelos de CB2 tanto en su conformación agonista como antagonista/ agonista-

inverso, y validar dichos modelos con DA de pequeña escala.

Se proponen por tanto los siguientes **objetivos**:

1. Compilación de datos de mutagénesis del receptor CB2.
2. Obtención de modelos estructurales de CB2 tanto en su conformación agonista como antagonista/ agonista-inverso.
3. Validación de los modelos a través de *docking* de pequeña escala.

Materiales y métodos

Compilación de datos de mutagénesis del receptor CB2

Con el fin de recolectar datos experimentales de CB2 para el posterior análisis y validación de los modelos, se buscaron datos de mutagénesis en distintos artículos publicados en el servidor PubMed (<https://www.ncbi.nlm.nih.gov/pubmed>) y en las bases de datos TinyGRAP (<http://www.cmbi.ru.nl/tinygrap/search/>) y GPCRdb (<http://gpcrdb.org/>) (Edvardsen et al., 2002; Isberg et al., 2016; Munk et al., 2016).

Modelado por homología

La obtención de los modelos estructurales de CB2 en conformación agonista y antagonista/ agonista inverso se llevó a cabo utilizando la metodología representada en el esquema de flujo de la Figura 5, de acuerdo al método de *ligand-steered homology modelling* (Orry and Cavasotto, 2006; Cavasotto et al., 2008; Phatak et al., 2010).

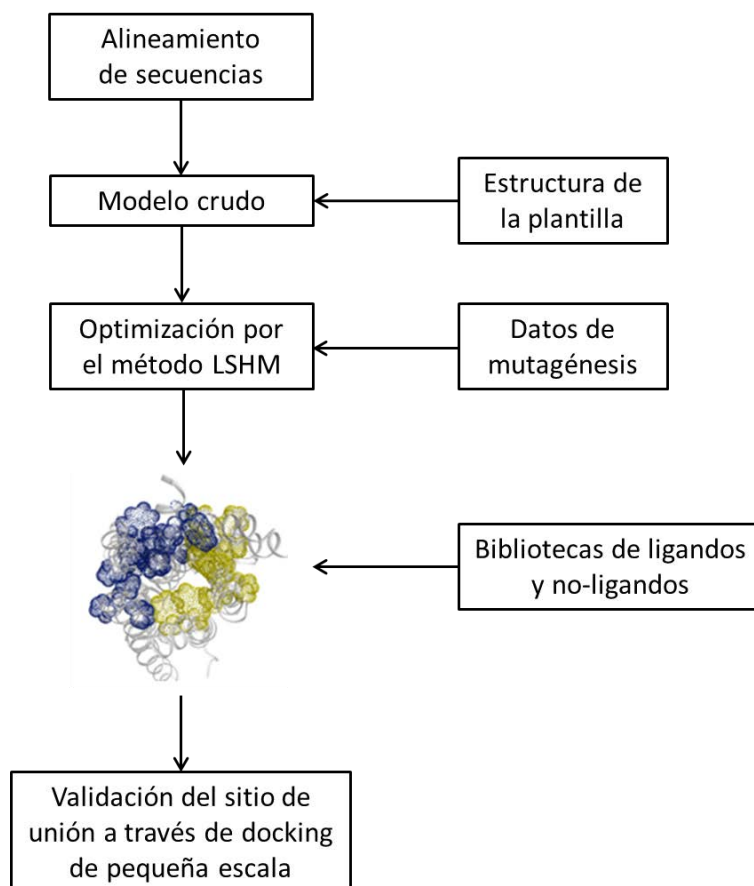


Figura 5. Esquema de flujo de los distintos pasos llevados a cabo para la obtención y validación de los modelos agonistas y antagonistas de CB2.

A lo largo del modelado por homología, se emplearon diversos servidores online (Clustal Omega, PubMed, PubChem, UniProtKB/Swiss-Prot) y programas tanto de libre uso (MODELLER) (Andrej Šali, 1993; Sievers et al., 2011), como comerciales, para el cual el grupo cuenta con la licencia correspondiente (ICM-Pro) (Abagyan, Totrov and Kuznetsov, 1994; ICM, 2012).

Tal como se explicó en la Introducción, el modelado por homología cuenta con varios pasos. En primer lugar, se buscaron y seleccionaron distintas estructuras 3D experimentales de GPCRs filogenéticamente cercanas a CB2 y que, además, contaran con un alto grado de homología en las regiones TM. Luego se realizó el alineamiento de secuencias, el cual fue utilizado como entrada (*input*) para obtener los modelos crudos. Por último, los modelos generados fueron refinados y validados mediante DA. En las siguientes sub-secciones se detallan cada uno de los pasos desarrollados en el presente trabajo.

I. Alineamiento de secuencias

Para obtener alineamientos más exactos, se utilizaron la secuencia de CB2 y la de varios GPCRs de clase A, las cuales se extrajeron de la base de datos UniProtKB/Swiss-Prot (<http://www.uniprot.org/>). El alineamiento múltiple se construyó utilizando el servidor online Clustal Omega (<http://www.uniprot.org/align>) (Sievers et al., 2011) y tiene como ventaja la identificación de áreas similares que pueden estar asociadas con características específicas que han sido más altamente conservadas que otras regiones. Las secuencias que fueron ingresadas en el servidor Clustal Omega corresponden al receptor adrenérgico β_1 (ADRB1, pdb: 2y02), receptor adenosina A_{2A} (A2AAR, pdb: 3qak), receptor 1 esfingosina 1-fosfato (S1P1, pdb: 3v2y), receptor de ácido lisofosfatídico 1 (LPA1, pdb: 4z34), receptor cannabinoide 1 (CB1, pdb: 5tgz) y receptor cannabinoide 2 (CB2).

Una vez obtenido el alineamiento múltiple, se lo inspeccionó visualmente teniendo en cuenta el trabajo de Mirzadegan *et al.* (Mirzadegan et al., 2003) donde se analiza la composición aminoacídica de los GPCRs miembros de la clase A y el grado de conservación de los aminoácidos en cada hélice, y el trabajo de Gonzalez *et al.* (Gonzalez et al., 2012) que estudia las secuencias y estructuras cristalinas de doce GPCRs, para inferir eventuales espacios (*gaps*) en el alineamiento en las hélices 2 y 5.

Por otra parte, para poder utilizar los alineamientos posteriormente en el programa de modelado (software MODELLER, ver en la sección II.) fue necesario convertirlos al formato PIR ya que del servidor Clustal Omega se obtuvieron en formato FASTA (Figura 6). Un ejemplo de alineamiento en formato PIR se muestra en la Figura 7, con

el fin de facilitar la explicación de las partes que lo conforman.

```
>P32835
MSAPAANGEVPTFKLVLVGDGGTGKTTFVKRHLTGEFEKKYIATIGVEVHPLSFYTNFGE
IKFDVWDTAGQEKFGLRDGYINAQCAIIMFDVTSRITYKNVPNWHRDLVRVCENIPIV
LCGNKVDVKERKVKAKTITFHRKKNLQYYDISAKSNYNFEKPFLLWARKLAGNPQLEFVA
SPALAPPEVQVDEQLMQQYQQEMEQTALPLPDEDDADL
>Q9P4E9
-----MAQEVPTFKLVLVGDGGTGKTTFVKRHLTGEFEKKYIATLGVEVHPLGFHTNFGE
LKFDVWDTAGQEKFGLRDGYINGQCGIIMFDVTSRITYKNVPNWHRDLVRVCENIPIV
LCGNKVDVKERKVKAKTITFHRKKNLQYYDISAKSNYNFEKPFLLWARKLVGNPQLEFVA
SPALAPPEVQVDADLMQKYQQEMEQTALPLPDEDDADL
```

Figura 6. Alineamiento de dos secuencias en formato FASTA a modo de ejemplo.

```
>P1;5tgz_a
structureX:5tgz:FIRST :A :LAST:A: cannabinoid receptor 1:::
-----GGRGENFMDIECFMVLNPSQQLAIAVLSLTLGTFTVLENLLV
LCVILHSRSLRCRPSYHFIGSLAVADLLGSVIFVYSFIDFHV FHRKDSRNVFLFKLGGVT
ASFTASVGSFLAAIDRYISIHRLAYKRIVTRPKAVVAFCLMWTIAIVIAVPLLLGWN
EKLQSVCSDFPHIDKTYLMFWIGVSVLLLLFIVYAYMYILWKAHSHAVAKALIVYGSTT
GNTEYTAETIARELADAGYEVDSRDAASVEAGGLFEGFDLVLVLCSTWGDDSIELQDDFI
PLFDSLEETGAQGRKVACFCGDSWEYFCGAVDAIEEKLKNLGAEIVQDGLRIDGDPRA
ARDDIVGWAHDVRGAIIPDQARMIDIELAKTLVLILVLLIICWGPLLAIMVYDVF GKMNKLI
KTVFAFCMLCLLNSTVNPIIYALRSKDLRHAFRSMF--*

>P1;CNR2_CB2
sequence:CNR2_CB2:::::0.00: 0.00
MEECWVTEIANGS-----KDGLDSNPMKDYMILSGPQKTAVAVLCTLLGLLSALENVAV
LYLILSSHQLRRKPSYLFISLAGADFLASVVFACSFVNFHVFHGVDSKAVFLKIGSVT
MTFTASVGSLLLLTAIDRYLCLRYPPSYKALLTRGRALVTLGIMWVLSALVSYLPLMGWTC
CP--RPCSELFPLIPNDYLLSWLLFIAFLFSGIITYGHVWLKAHQHAAAVAS-----
-----
-----LSLDVRLAKTLGLVLAVLLICWFPVLALMAHSLATTLSDQV
KKAFAFCSMLCLINSMVNPVIYALRSGEIRSSAHHCLAH*
```

Figura 7. Alineamiento de secuencias en formato PIR usado como *input* en MODELLER.

En cuanto al formato PIR, la primera línea de entrada de cada secuencia es la identificadora; luego de “>P1;” se especifica el código de la proteína. Convencionalmente se suele utilizar como código el PDB de la proteína, pero esto no es requisito ya que se puede emplear cualquier otro mientras sea único. De estar separado por espacios o tabulaciones, los únicos caracteres que se leen son los primeros. La segunda línea de cada entrada contiene información necesaria tanto de la secuencia como de la estructura 3D (en el caso de existir). Los distintos campos están separados con “:” y aportan lo siguiente:

- Campo 1: se especifica si existe la estructura 3D para esa secuencia y el método por el cual se obtuvo dicha estructura (“*structureX*”= rayos-X, “*structureN*”= NMR, “*structureM*”= modelo). En caso de no existir una estructura (es el caso de la secuencia blanco) se especifica “*sequence*”.
- Campo 2: se usa para obtener la información estructural de la secuencia. Puede ser el código pdb de la proteína (en la primera entrada de la figura 7, el código pdb= 5tgz), el nombre del archivo con extensión “.pdb” / “.cif” o, incluso, el camino hasta el archivo. En el caso que el primer campo se haya especificado como “*sequence*”, este campo se puede dejar vacío.
- Campos 3 a 6: especifica los identificadores de residuo y cadena para el primer residuo (campos 3-4) y para el último residuo (campos 5-6) de la secuencia. En caso de dejar vacíos estos campos, se lee el archivo de coordenadas indicado en el campo 2 de forma completa. El residuo inicial/ final no estará identificado en el caso que se ingrese “.” (punto) o espacio, mientras que la cadena no va a estar identificada sólo en el caso que se ingrese “.” (punto). Si en el campo 3 se utiliza “*FIRST*” se leen los residuos desde el primero presente en el archivo pdf (para esa cadena). Si en el campo 5 se utiliza “*LAST*” se leen los residuos hasta el final del archivo pdf (para esa cadena).
- Campo 7: indica el nombre de la proteína.
- Campo 8: indica el organismo del cual se obtuvo dicha proteína.
- Campo 9: indica la resolución de la cristalografía.
- Campo 10: indica el factor R de la cristalografía.

En resumen, cuando una secuencia va a ser utilizada junto con información estructural (es el caso de la plantilla) los dos primeros campos son de llenado obligatorio mientras que los demás pueden permanecer vacíos. En el caso de la secuencia blanco, la cual no posee información estructural asociada, sólo es obligatorio llenar el primer campo.

Por último, en la tercera línea se ingresa la secuencia previamente alineada, la cual se da por finalizada con un asterisco (*).

Una vez pasados los alineamientos a formato PIR, se eliminaron aminoácidos del extremo N-terminal (1-12), del extremo C-terminal (316-360) y 11 aminoácidos del lazo intracelular 3 (224-234) de la secuencia de CB2 con el fin de obtener modelos más acordes. El ICL3 se acortó debido a que en varias estructuras de GPCRs existen residuos de este lazo que se pierden en el proceso de cristalografía debido a su longitud. Tanto el ICL3 como el C-terminal no están involucrados en el sitio de unión por lo que al recortarlos no interferirán para alcanzar los objetivos del trabajo. El N-terminal tendría un aporte en el sitio de unión, por este motivo es que sólo se recortó

un segmento inicial que no estaría afectando dicho bolsillo.

II. Obtención de modelos estructurales de CB2

A partir del alineamiento de la secuencia de CB2 (*target*) con las secuencias de las proteínas de estructura 3D conocida (*templates*), se realizó un modelado por homología mediante el uso de MODELLER 9.16 (Andrej Šali, 1993), el cual consiste en ingresar el alineamiento de dos secuencias al programa (*input*) con el fin de obtener un modelo 3D de la proteína blanco (*output*). El mismo contiene la estructura principal y cadenas laterales sin los átomos de hidrógeno, por lo que es un modelo crudo en el sentido que hereda completamente la estructura del esqueleto (*backbone*) de la proteína usada como molde.

El modelado crudo de CB2 en su conformación agonista se realizó a partir de la estructura cristalina de ADRB1 (Warne et al., 2011). Para la conformación antagonista/agonista-inverso se utilizó como molde la estructura de CB1, S1P1 (Hanson et al., 2012) y LPA1 (Chrencik et al., 2015). Dada la alta homología con CB1 (42% de identidad de secuencia), cuya estructura estuvo disponible hace pocos meses, sólo se continuó la optimización del modelo antagonista que fue obtenido a partir de CB1. En todos los casos, las estructuras cristalinas utilizadas como molde fueron descargadas de RSCB Protein Data Bank (<http://www.rcsb.org/pdb>) (Bernstein et al., 1977).

Para realizar estos modelos fue necesario escribir un *script* en lenguaje Python. En la Figura 8 se muestra un *script* utilizado a modo de ejemplo.

```
from modeller import *
from modeller.automodel import *
#from modeller import soap_protein_od

env = environ()
a = automodel(env, alnfile='CB2_CB1_5tgz_PIR.ali',
              knowns='5tgz_a', sequence='CNR2_CB2',
              assess_methods=(assess.DOPE,
                              assess.GA341))

a.starting_model = 1
a.ending_model = 5
a.make()
```

Figura 8. Ejemplo de *script* para la generación de modelos crudos.

En el *script* empleado para la generación de modelos crudos, la primera línea carga el módulo “*automodel*” y lo prepara para su uso. A continuación se crea un objeto “*environ*” el cual contiene el entorno de MODELLER, que incluye valores por defecto para muchos parámetros así como bibliotecas utilizadas para el modelado

comparativo. Luego se crea un objeto “*automodel*” (llamado “a”) y se fijan los parámetros para guiar el proceso de modelado. Con “*alnfile*” se nombra el archivo que contiene el alineamiento blanco-plantilla en formato PIR, “*knowns*” determina cual es el código de la proteína con estructura 3D conocida (*template*) y “*sequence*” define cual es el código de la secuencia blanco en el “*alnfile*” (*target*). Una vez fijados los parámetros, se pidieron los métodos a aplicar para la evaluación de los modelos con “*assess_methods*”. Se define el número de modelos a crear mediante “*automodel.starting_model*” y “*automodel.ending_model*”. Finalmente, se calculan los modelos con la última línea “*automodel.make*”.

En todos los casos se creó un total de cinco modelos que, una vez obtenidos, se los inspeccionó con el fin de seleccionar uno sólo y luego continuar el modelado con ese en particular. El criterio utilizado para dicha selección fue el modelo con el puntaje DOPE más bajo y *Z-score* más adecuado, que para una proteína de aproximadamente 275 aminoácidos es alrededor de -6. El puntaje DOPE se obtuvo a partir de MODELLER, mientras que el *Z-score* se calculó mediante el servidor online ProSA-web (<https://prosa.services.came.sbg.ac.at/prosa.php>) (Sippl, 1993; Wiederstein and Sippl, 2007).

A lo largo del trabajo, para todos los modelos se utilizó la numeración habitual de los aminoácidos y la nomenclatura de Ballesteros y Weinstein (Ballesteros and Weinstein, 1995) quienes definieron un esquema general de dos números, donde el primero denota la hélice en la cual está localizado el aminoácido de interés (va de 1-7) y el segundo asigna la posición relativa al residuo más conservado en la hélice [cfr. (Mirzadegan et al., 2003)], al cual se lo denota X.50 siendo X el número de la hélice. A medida que uno se aleja del aminoácido más conservado, hacia el N-terminal el segundo número disminuye y hacia el C-terminal aumenta.

Para el modelo agonista de CB2 se realizó un abordaje con múltiples plantillas debido a que no hay disponibles estructuras cristalinas en el estado activo con alta homología de secuencia. Mediante el uso del software ICM-Pro (ICM, 2012), se modificaron en el modelo crudo el N-terminal desde el residuo 13 hasta el 30, la hélice 2 entre los aminoácidos 68-97 y el ECL2 junto con la hélice 5 (residuos 171 al 219), tomando como molde la estructura 3D del receptor CB1.

Antes de realizar la optimización de cada modelo, se estableció un puente disulfuro en el ECL2 entre los aminoácidos C174 y C179, en caso que no hubiera sido heredado de la plantilla, en base al estudio experimental realizado por Gouldson y equipo (Gouldson et al., 2000).

Con el software ICM-Pro, se efectuó un *docking* de proteína rígida de los ligandos correspondientes a cada modelo, WIN 55,212-2 para el modelo agonista y SR-144,528

para el antagonista/ agonista-inverso, y se utilizaron las poses de menor energía para la optimización posterior. A partir de estas poses iniciales, se generaron 200 estructuras de cada ligando en el sitio de unión randomizando la posición y orientación del ligando. La posición del centro de masa del ligando se permite que se desplace aleatoriamente dentro de una esfera de 3 Å desde su valor inicial, al igual que la orientación de la molécula en el espacio, donde la amplitud de rotación tiene una distribución gaussiana alrededor de $\theta = 3 \text{ \AA}$ / radio de giro de la molécula (expresado en Å). El refinamiento del sitio de unión se llevó a cabo mediante el método de *ligand steered homology modelling* (Cavasotto et al., 2008; Phatak et al., 2010), en el cual el sitio de unión es optimizado junto con el ligando seleccionado, realizando una minimización global de la energía y manteniendo flexible tanto el receptor como el ligando (Figura 9).

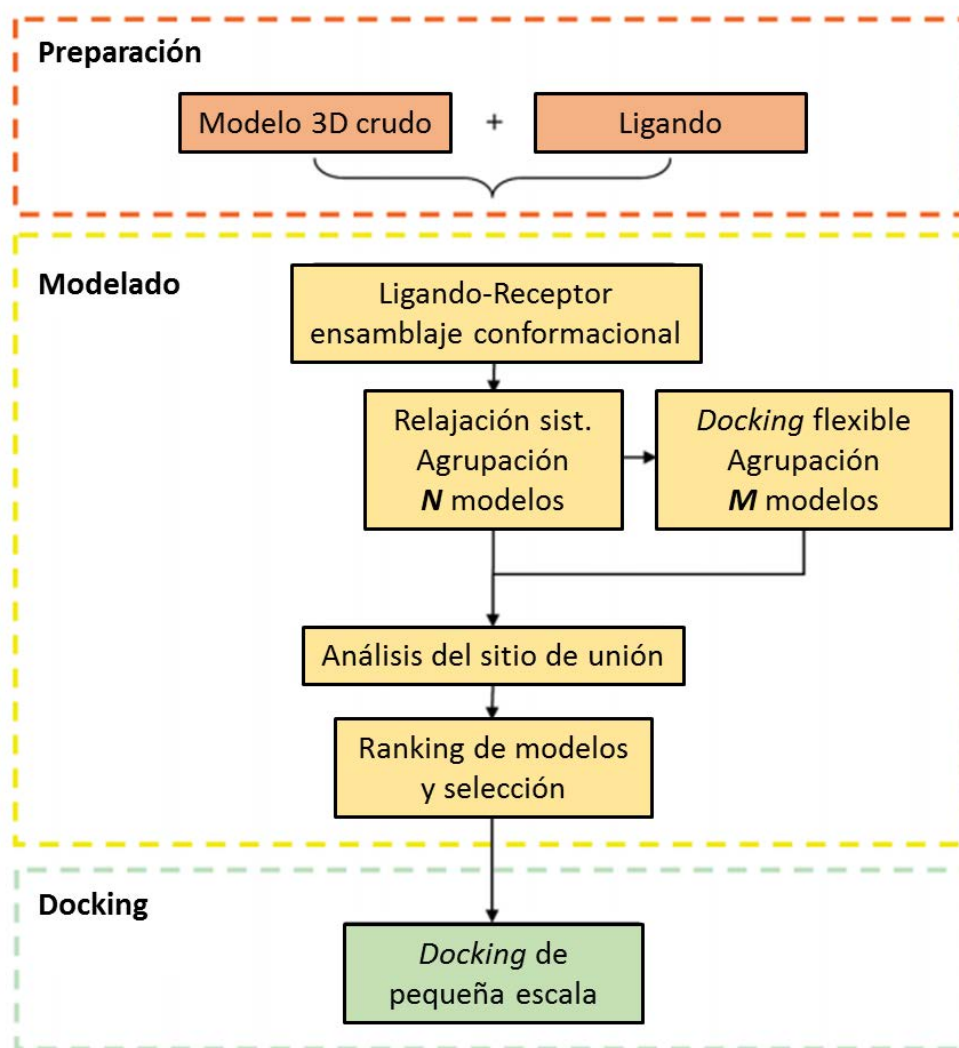


Figura 9. Diagrama de flujo del método de modelado *ligand steered homology modelling* usado en el trabajo.

En este método, se parte de un modelo 3D crudo de la proteína blanco con el ligando ubicado en el sitio de unión. El conjunto de estructuras del complejo proteína-ligando generado por randomización del ligando se someten a una relajación, donde las interacciones de van der Waals se cambian gradualmente de suave a fuerte. Se evitan gradientes de energía altos al comienzo de la minimización para que no se colapse el sitio de unión. Las interacciones receptor-ligando validadas mediante datos experimentales son incorporadas en esta etapa como restricciones de distancias. Las estructuras obtenidas en este paso se agrupan para seleccionar una N cantidad de modelos, los cuales van a ser sujetos a una optimización a través de un *docking* flexible, donde se mantienen flexibles el ligando y el receptor, utilizando una minimización global de la energía basada en el método de Monte Carlo (Metropolis et al., 1953) con minimización local luego de cada salto aleatorio (Abagyan and Totrov, 1994), utilizando 300000 pasos de minimización en total. Las cadenas laterales que apuntan hacia el sitio de unión y que se encuentran a una distancia de 6 Å del ligando son consideradas libres, mientras que al esqueleto (*backbone*) se lo mantiene fijo a excepción de ciertos lazos que se van a considerar flexibles. A las estructuras obtenidas se les realiza un agrupamiento con el fin de seleccionar un número de modelos M menor que diez. Por último, se hace un análisis visual del sitio de unión y se realiza un *ranking* de los modelos obtenidos. A partir de éstos, se seleccionan los modelos estructurales candidatos para realizar la validación mediante un *docking* automatizado de pequeña escala.

Generación de GLL/GDD

Para generar la biblioteca de ligandos de CB2 (GLL), se filtró de una base de datos extraída de PubChem con 2135 ligandos (<https://pubchem.ncbi.nlm.nih.gov/>) agonistas, antagonistas y agonistas inversos que tuvieran una constante de afinidad menor o igual a 1 µM. Para obtener dicha biblioteca, en primer lugar, se descartaron todos aquellos ligandos con una afinidad mayor a 1 µM y aquellas moléculas que fueran péptidos, ya que son de interés sólo las pequeñas moléculas. A continuación, se descartaron los ligandos cuyo ensayo no fue confirmatorio o no presentaba estudio de la actividad biológica, es decir, aquellos que sólo se sabe que son ligandos pero no pueden ser clasificados en agonistas, antagonistas o agonistas inversos. Por último, se eliminaron aquellas moléculas que estuvieran repetidas una o más veces y se clasificó a los ligandos restantes de acuerdo a la actividad biológica obtenida en el ensayo. Las moléculas presentes en las listas finales fueron descargadas en formato "sdf" desde el servidor PubChem.

A partir de la GLL, se generó una biblioteca de moléculas señuelo (GDD), es decir que representan no-ligandos, de acuerdo al método ya desarrollado en el grupo (Gatica and Cavasotto, 2012). Las moléculas no-ligandos generadas por este método se caracterizan por tener propiedades físicas similares a los ligandos, pero estructura química diferente lo cual reduce la probabilidad de generar bibliotecas señuelo que contengan ligandos reales. El diseño racional de la librería de señuelos tiene como objeto evitar el enriquecimiento artificial en el *docking* molecular (Huang, Shoichet and Irwin, 2006; Irwin, 2008; Gatica and Cavasotto, 2012).

Por cada ligando se generaron un total de 39 moléculas señuelo, las cuales poseen similitud ligando-señuelo en seis propiedades físicas (peso molecular, número de enlaces rotatorios, número de aceptores de puente de hidrógeno, número de dadores de puente de hidrógeno, carga neta de la molécula y $\log P$) y diversidad estructural ligando-señuelo.

Docking molecular de proteína rígida

Una vez optimizados los modelos estructurales y obtenidas las bibliotecas de ligandos y no-ligandos, se evaluó la calidad de los modelos por la habilidad de discriminar entre ligandos y moléculas señuelo mediante un *docking* de pequeña escala (Cavasotto, Orry and Abagyan, 2006). Para ello, se utilizó el software ICM-Pro, el cual tiene en cuenta la flexibilidad del ligando en el proceso de *docking*, aunque el receptor es considerado rígido (sin flexibilidad en las cadenas laterales, ni en el esqueleto). La función de energía utilizada durante el proceso de *docking* consiste de:

$$E_{docking} = E_{FFint} + E_{vw} + E_{el} + E_{hb} + E_{hp} \quad (1)$$

donde E_{FFint} es la energía del campo de fuerzas del ligando y E_{vw} , E_{el} , E_{hb} y E_{hp} son los términos potenciales de van der Waals, electrostáticos, puente de hidrógeno, y el término del efecto hidrofóbico, respectivamente. Los últimos cuatro términos son precalculados en una grilla de 0.5 Å de lado para acelerar la evaluación de la energía.

La energía de la molécula libre es minimizada en ausencia del receptor antes del *docking*. Luego el ligando es acoplado en la grilla y la mejor conformación energética es puntuada utilizando una función de puntuación empírica optimizada para maximizar la separación de ligandos y no-ligandos en un conjunto de entrenamiento (ver Ec. 2). Los métodos de puntuación aproximan la energía libre de unión entre un ligando y el receptor. Para lograr un equilibrio entre velocidad y precisión del *docking*, estos métodos de puntuación deben ser rápidos y hacer algunas aproximaciones en los cálculos. Teniendo esto en cuenta, las funciones de puntuación consideran la energía

libre de unión como una función aditiva de diferentes términos de energía que reflejan las distintas contribuciones en la unión ligando-receptor. La función de puntuación de ICM cuya función es discriminar diferentes moléculas consiste de los siguientes términos:

$$E_{\text{score}} = \Delta E_{\text{IntFF}} + T\Delta S_{\text{Tor}} + \alpha_1 \Delta E_{\text{HBond}} + \alpha_2 \Delta E_{\text{HBDsol}} + \alpha_3 \Delta E_{\text{SolEI}} + \alpha_4 \Delta E_{\text{HPhob}} + \alpha_5 Q_{\text{Size}} \quad (2)$$

donde ΔE_{IntFF} incluye las interacciones de van der Waals del ligando con el receptor como así también la energía del campo de fuerza interna del ligando, $T\Delta S_{\text{Tor}}$ es la contribución de la pérdida de entropía conformacional del ligando al unirse al receptor (se asume que es proporcional al número de ángulos torsionales libres), ΔE_{HBond} es el término de puente de hidrógeno, ΔE_{HBDsol} es el término que considera la disrupción de los puentes de hidrógeno con el solvente luego de la unión del ligando, ΔE_{SolEI} es el cambio de energía electrostática de solvatación tras la unión, ΔE_{HPhob} es el término de energía libre asociado al efecto hidrofóbico, y asumido proporcional al área superficial accesible de los átomos hidrofóbicos del receptor y ligando. Q_{Size} es un término de corrección de tamaño para evitar el sesgo hacia ligandos más grandes. Por último, los valores α_1 - α_5 dan a cada interacción el peso apropiado y son optimizados en un patrón diverso de complejos.

Para cada modelo, se fijó una grilla de 20 x 20 x 20 Å que contenga el sitio de unión ortostérico del receptor con el fin de delimitar el espacio. El *docking* y el proceso de puntuación (*scoring*) fue repetido dos veces, conservándose el mejor puntaje para cada compuesto.

El factor de enriquecimiento (EF) a un porcentaje x se calculó como:

$$EF_x = (\text{Hits}_x / N_x) / (\text{Hits}_{\text{total}} / N_{\text{total}}) \quad (3)$$

donde Hits_x es el número de ligandos en el porcentaje x de la base de datos leída, N_x es el número de compuestos a un x% del total de la base de datos, $\text{Hits}_{\text{total}}$ es el valor total de ligandos en la biblioteca y N_{total} es el número total de compuestos en la base de datos. El EF representa la relación (cociente) entre la probabilidad de encontrar un ligando en el x% de la base de datos y la probabilidad de encontrarlo al azar.

Resultados

Compilación de datos de mutagénesis del receptor CB2

Al momento de realizar la búsqueda se tuvo en cuenta datos de mutagénesis aplicados en CB2 para agonistas, antagonistas o agonistas inversos. La recopilación se centró principalmente en datos experimentales de aminoácidos cercanos al sitio de unión de CB2, ya que son los que afectan de forma directa la interacción con el ligando, que estén evaluados con WIN 55,212-2 (agonista de CB2, Tabla 1) y SR-144,528 (agonista inverso de CB2, Tabla 2). Sin embargo, como no existe una gran cantidad de datos experimentales para dichas moléculas, también se tuvo en cuenta mutagénesis estudiadas con otros ligandos.

Brevemente, en los experimentos de mutagénesis sitio dirigidas se lleva a cabo la modificación de un aminoácido particular de la proteína por otro que se quiera evaluar su efecto mediante un cambio de bases en el gen de la proteína. Una vez obtenida la mutación se compara con respecto a la proteína salvaje (*wild type*) el valor de ciertos parámetros, como puede ser la constante de afinidad (K_a), frente a un agonista, antagonista o un agonista inverso en particular.

Tabla 1. Se listan los datos de mutagénesis para WIN 55,212-2 y otros agonistas de CB2, con sus respectivas referencias.

Residuo	Modificación en CB2	Referencia
C89 (2.59)	C89S: no hubo cambios en la afinidad de CP55940.	(Mercier et al., 2010)
K109 (3.28)	K109A: no hubo efecto significativo en la unión de WIN55,212-2. K109AS112G: sin efecto en afinidad de WIN55,212-2.	(Tao et al., 1999)
V113 (3.32)	V113E: se pierde la unión de CP55940. V113L: sin cambios en afinidad de unión de CP55940.	(Feng et al., 2014)
F183 (ECL2)	F183Q: se pierde la unión de CP55940. F183V: se pierde la unión de CP55940.	(Feng et al., 2014)
Y190 (5.39)	Y190F: sin efecto en la afinidad de WIN55,212-2. Y190I: se pierde la unión de CP55940.	(McAllister et al., 2002)
L192 (5.41)	L192A: sin efecto en la afinidad de unión de CP55940 L192S: pérdida de la unión de CP55940.	(Alqarni et al., 2014)
S193 (5.42)	S193G: sin cambios en la afinidad de WIN55,212-2.	(Gouldson

		et al., 2000)
W194 (5.43)	W194Y: se reduce 8 veces afinidad de WIN55,212-2. W194F: se pierde la unión de WIN55,212-2. W194A: se pierde la unión de WIN55,212-2.	(Zhang et al., 2011)
F197 (5.46)	F197V: se reduce 14 veces la afinidad de unión de WIN55,212-2.	(Song et al., 1999)
C257 (6.47)	C257A: sin cambios en la afinidad de WIN55,212-2. C257S: sin cambios en la afinidad de WIN55,212-2. C257K, C257I, C257L, C257D, C257Y: se pierde la unión específica de CP55940.	(Pei et al., 2008)
W258 (6.48)	W258A: se pierde la afinidad por WIN55,212-2. W258F: se pierde la unión de WIN55,212-2. W258C: pérdida de la unión de CP55940.	(Nebane et al., 2008)
P260 (6.50)	P260C: se pierde la capacidad de unión de CP55940.	(Nebane et al., 2008)
F281 (7.35)	F281K: pérdida de la unión de CP55940.	(Feng et al., 2014)
C284 (7.38)	C284A: sin cambios en la afinidad de WIN55,212-2. C284S: se mantiene la afinidad de unión de CP55940.	(Mercier et al., 2010)
S285 (7.39)	S285A: se reduce 13 veces la afinidad de unión de HU-243.	(Rhee, 2002)

Tabla 2. Se listan los datos de mutagénesis para SR-144,528 y otros agonistas inversos de CB2, con sus respectivas referencias.

Residuo	Modificación en CB2	Referencia
V113 (3.32)	V113E: se pierde la afinidad por SR-144,528. V113L: sin cambios en la afinidad de SR-144,528.	(Feng et al., 2014)
S161 (4.53)	S161A: pérdida de la unión de SR-144,528.	(Gouldson et al., 2000)
V164 (4.56)	V164I: sin cambios en la afinidad de SR-144,528.	(Gouldson et al., 2000)
S165 (4.57)	S165A: se pierde la afinidad de SR-144,528.	(Gouldson et al., 2000)
C174 (4.66)	C174S: pérdida de la unión de SR-144,528.	(Gouldson et al., 2000)
C175 (4.67)	C175S: se pierde la afinidad por SR-144,528	(Gouldson

		et al., 2000)
R177 (ECL2)	R177S: sin efectos en la afinidad de SR-144,528.	(Gouldson et al., 2000)
C179 (ECL2)	C179S: se pierde la unión de SR-144,528.	(Gouldson et al., 2000)
F183 (ECL2)	F183Q: se pierde la afinidad por SR-144,528. F183V: pérdida de la unión de SR-144,528.	(Feng et al., 2014)
L192 (5.41)	L192A: se reduce 4 veces la afinidad de SR-144,528. L192S: se pierde la unión de SR-144,528.	(Alqarni et al., 2014)
S193 (5.42)	S193G: sin cambios en la afinidad de SR-144,528.	(Gouldson et al., 2000)
W194 (5.43)	W194Y: se reduce 4 veces la afinidad de SR-144,528. W194A: pérdida de la unión de SR-144,528. W194F: se pierde la unión de SR-144,528.	(Zhang et al., 2011)
F281 (7.35)	F281K: se pierde la afinidad por SR-144,528.	(Feng et al., 2014)
C284 (7.38)	C284AC288A y C284SC288S: se anula la interacción entre AM1336 y CB2.	(Mercier et al., 2010)

A partir de estas tablas con datos de mutagénesis se analizará la interacción receptor-ligando y se evaluará la calidad de los modelos obtenidos.

Alineamiento de secuencias

En la Figura 10 se muestra el alineamiento obtenido a partir del servidor Clustal Omega. En la parte inferior de las secuencias alineadas, se observan cilindros que representan la longitud de las siete α -hélice del receptor S1P1 (pdb: 3v2y). Aquellas partes sin cilindros indican los lazos, ya sea intra o extracelulares, y el N- o C-terminal. Además, se obtuvo la identidad de secuencia entre CB2 y cada uno de los receptores utilizados en el alineamiento (Tabla 3). La misma se calcula como el número de residuos similares (en términos de propiedades) en el alineamiento dividido por el largo de la secuencia más corta y multiplicado por 100%.

El alineamiento obtenido se inspeccionó de forma visual y se corroboró que es acorde al trabajo de Gonzalez *et al.* donde ciertos receptores presentan espacios (*gaps*) en las hélices transmembrana 2 y 5, siendo CB2, CB1, S1P1 y LPA1 uno de estos casos (Gonzalez et al., 2012). De acuerdo a la superposición de las estructuras cristalinas de los GPCRs, el residuo L201 de CB2 se lo consideró como 5.50 debido a

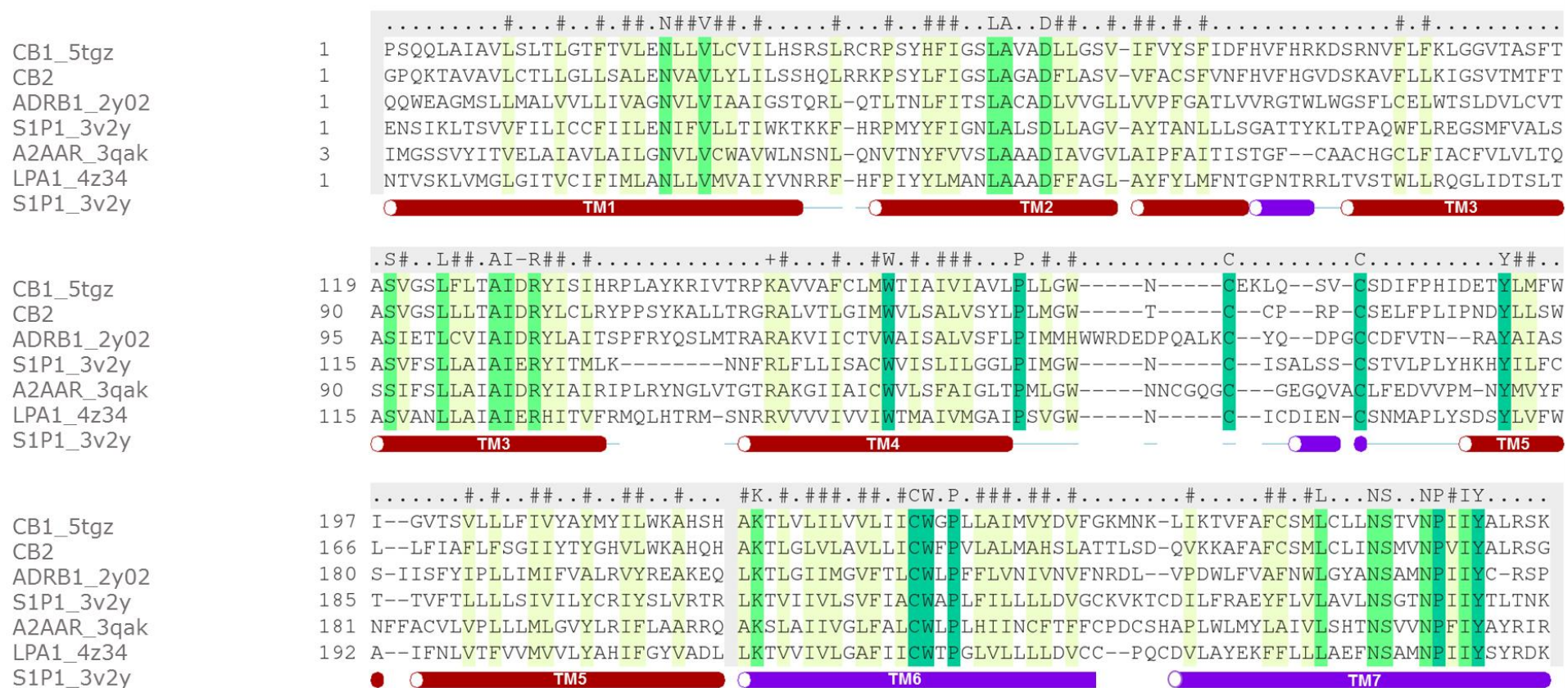


Figura 10. Alineamiento múltiple de los distintos GPCRs de clase A: CB2, CB1, S1P1, A2AAR, ADRB1 y LPA1. A la izquierda se indica a que receptor pertenece cada secuencia junto con su pdb correspondiente. En color verde se resaltan los aminoácidos conservados en las seis secuencias, mientras que en amarillo se resaltan los aminoácidos conservados pero en un número menor de secuencias. Los dominios transmembrana de la estructura de S1P1 se muestran en forma de cilindros. Tanto el C-terminal como el N-terminal se omitieron para una mejor visualización del alineamiento (son regiones poco conservadas).

la ausencia de prolina en la hélice 5. A su vez, se confirmó que ciertos motivos de aminoácidos que están muy conservados a lo largo de los GPCRs de Clase A se encuentran correctamente alineados. Este es el caso de los aminoácidos Gly y Asn en hélice 1, Leu y Asp en hélice 2, el residuo Cys y el motivo DRY (AspArgTyr) en la hélice 3, de los residuos Trp y Pro en la hélice 4, Pro y Tyr en hélice 5, los aminoácidos Phe, Trp y Pro de la hélice 6 y el motivo NPXXY (Asn, Pro y Tyr) en la hélice 7 (Mirzadegan et al., 2003). La conservación de residuos claves en cada hélice asegura un alineamiento aceptable a pesar de una baja similitud global entre algunas secuencias de GPCRs.

Tabla 3. Identidad o similitud de secuencia de CB2 con cada receptor utilizado en el alineamiento múltiple.

	CB1	S1P1	LPA1	ADRB1	A2AAR
ID secuencia global	42%	27%	23%	28%	22%
ID secuencia región TM	50%	34%	30%	35%	29%

Modelado por homología

A partir del alineamiento de secuencias se crearon los siguientes modelos crudos:

1. CB2 en su conformación agonista, usando como plantilla el receptor ADRB1 en su conformación activa (pdb: 2y02). (Figura 11)
2. CB2 en su conformación antagonista/ agonista inverso, usando como plantilla el receptor CB1 en su conformación inactiva (pdb: 5tgz). (Figura 12)
3. CB2 en su conformación antagonista/ agonista inverso, usando como plantilla el receptor S1P1 en su conformación inactiva (pdb: 3v2y).
4. CB2 en su conformación antagonista/ agonista inverso, usando como plantilla el receptor LPA1 en su conformación inactiva (pdb: 4z34).

En un principio, dado que no existía la estructura cristalina de CB1, se comenzó a modelar CB2 en su conformación inactiva a partir de las estructuras de S1P1 y LPA1. Sin embargo, dado que hace unos pocos meses se publicó la estructura cristalina de CB1 en un complejo con el antagonista AM6538 (molécula con similitud estructural a SR-144,528), se descartaron estos modelos.

Inicialmente, el proceso que realiza el programa MODELLER para generar los modelos crudos es transferir características espaciales, como distancias C α -C α , puentes de hidrógeno, ángulos diedros de la cadena principal y las cadenas laterales, desde la plantilla hacia la secuencia blanco. Es decir que extrae restricciones espaciales de la proteína utilizada como molde y luego las satisface en la

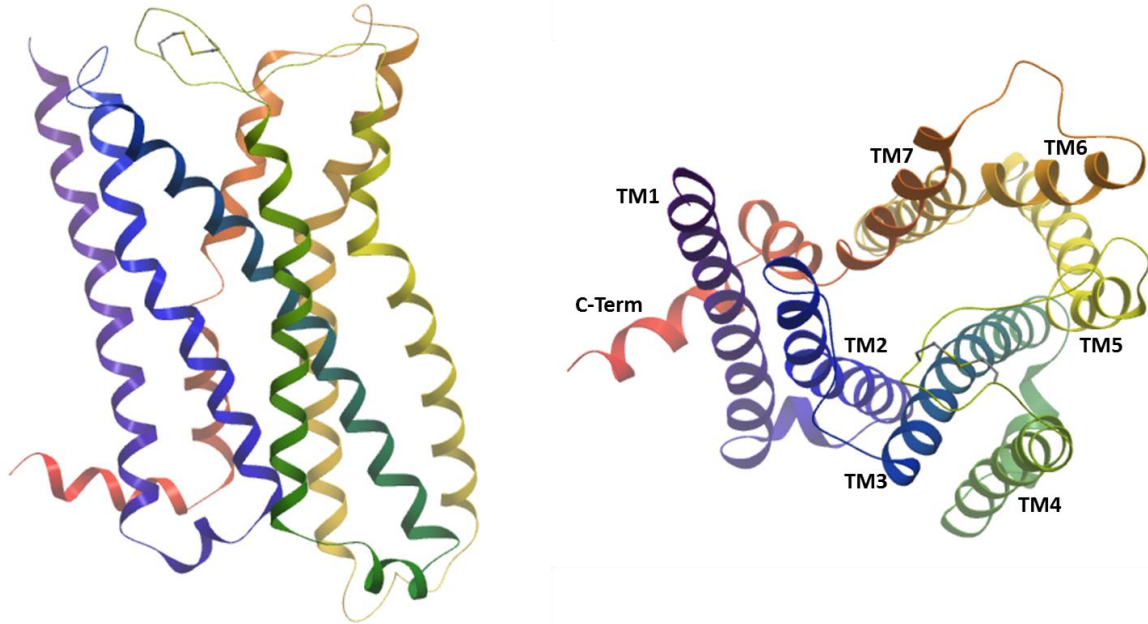


Figura 11. Representación del modelo crudo de CB2 en su conformación agonista, usando como plantilla la estructura cristalina del receptor ADRB1 en su conformación activa (pdb: 2y02). A la izquierda se observa una vista a través de la membrana plasmática y a la derecha se muestra el modelo desde una vista extracelular.

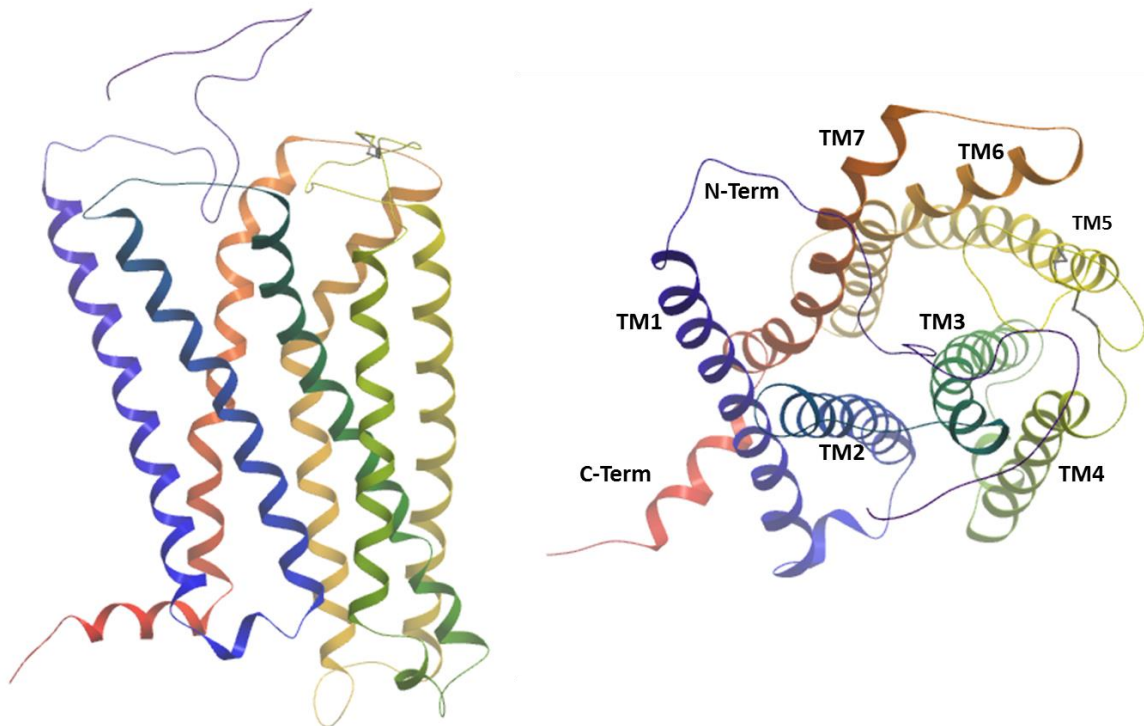


Figura 12. Representación del modelo crudo de CB2 en su conformación antagonista/agonista-inversa, usando como plantilla la estructura cristalina del receptor CB1 en su conformación inactiva (pdb: 5tgz). A la izquierda se observa una vista a través de la membrana plasmática y a la derecha se muestra el modelo desde una vista extracelular.

proteína blanco, en este caso CB2, de la mejor forma posible. Finalmente, el modelo se obtiene optimizando la función objetivo en el espacio cartesiano. La optimización se lleva a cabo mediante el método de función objetivo variable, empleando métodos de gradientes conjugados y dinámica molecular (Braun and Go, 1985).

El Z-score de cada modelo se comparó con los de todas las proteínas obtenidas experimentalmente, presentes en RSCB Protein Data Bank (Figura 13, gráficos obtenidos a partir del servidor online ProSA-web).

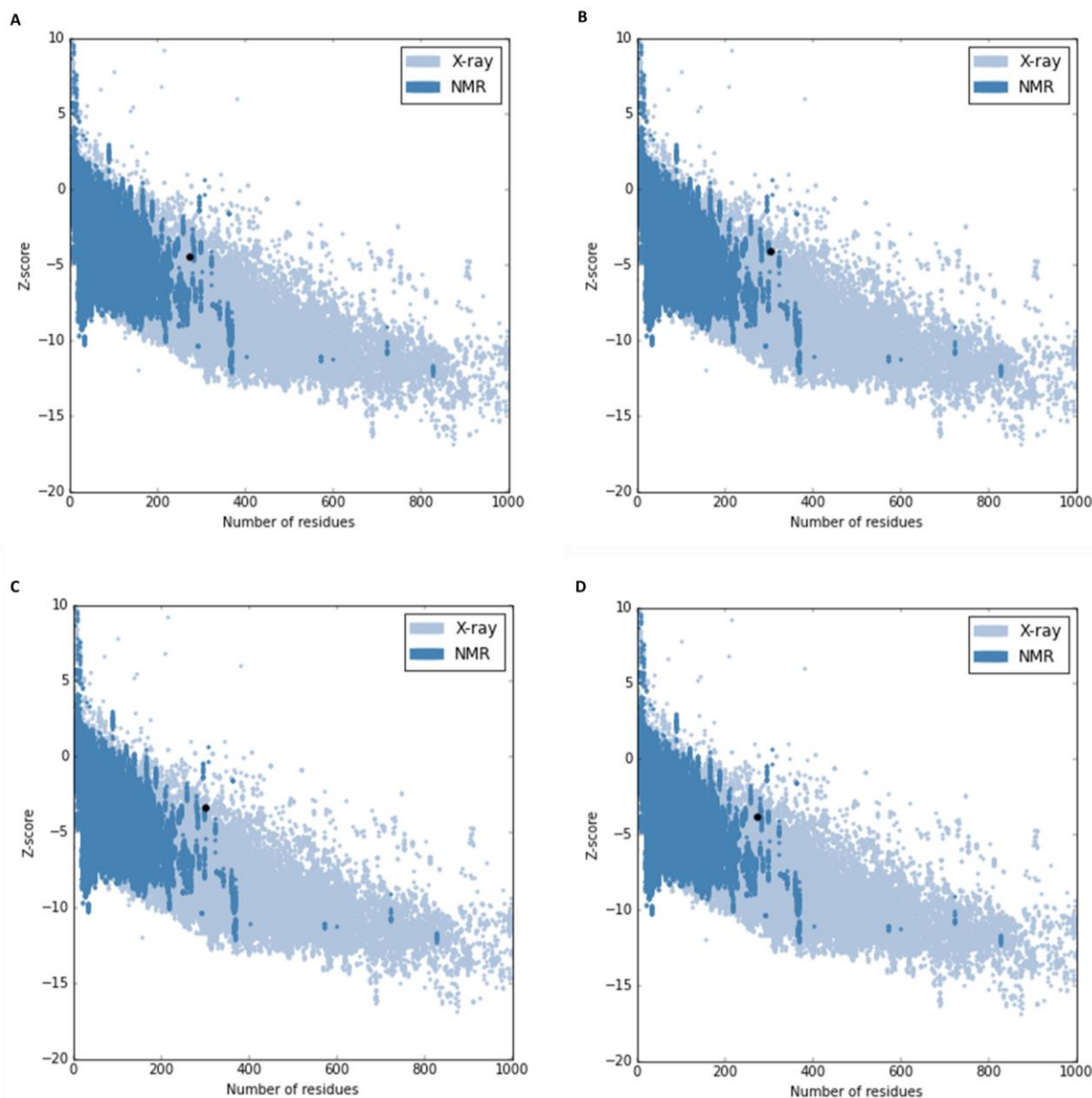


Figura 13. Z-score de los modelos crudos seleccionados indicado con un punto negro en cada gráfico. A) Z-score del modelo con el molde de ADRB1= -4,51. B) Z-score del modelo con el molde de CB1= -4,02. C) Z-score del modelo con el molde de LPA1= -3,06. D) Z-score del modelo con el molde S1P1= -3,8.

El *Z-score* indica la calidad global del modelo, midiendo la desviación de la energía total de la estructura con respecto a una distribución de energía derivada de conformaciones aleatorias (Sippl, 1993; Wiederstein and Sippl, 2007). Si el *Z-score* de un modelo se encuentra por fuera del rango esperable, entonces la calidad del mismo no es tan buena.

Dado que todos los modelos obtenidos presentaron un *Z-score* dentro de los parámetros esperados, los modelos crudos que se seleccionaron fueron de acuerdo al menor valor DOPE obtenido al realizar dichos modelos en el programa MODELLER (Tabla 4).

Tabla 4. Valores DOPE obtenidos para cada modelo. Los modelos seleccionados se marcaron en negrita.

Modelo	ADRB1	CB1	LPA1	S1P1
1	-34895	-38479	-37380	-35172
2	-34772	-38286	-37648	-34822
3	-35222	-38444	-37094	-34688
4	-34559	-38402	-37612	-35253
5	-34814	-38331	-37468	-34654

En cuanto al modelo agonista, las hélices 2 y 5 se modificaron usando la estructura de CB1 como molde debido a que CB2 presenta un espacio (*gap*) en el alineamiento con respecto a ADRB1 y dos espacios en la hélice 5 con respecto a A2AAR. Además, en CB2 se observa la ausencia del aminoácido prolina en las posiciones 2.58 y 5.50, residuo conservado en varios GPCRs clase A que produce un quiebre en la α -hélice (Gonzalez et al., 2012). Estas características se pueden observar en la Figura 11, donde el modelo crudo de CB2 en su conformación agonista heredó las hélices 2 y 5 de ADRB1 que presentan un quiebre (debido al aminoácido prolina) y un giro más amplio (debido a las inserciones de aminoácidos) en las mismas.

Para el refinado con el método *ligand steered homology modelling* se usaron como ligandos WIN 55,212-2 y SR-144,528 ya que son moléculas voluminosas, razonablemente rígidas y, además, presentan una excelente afinidad (en el orden nm). Para obtener conformaciones iniciales de estas moléculas, se efectuó un *docking* de proteína rígida con ICM-Pro. Durante la generación de las 200 estructuras receptor-ligando, la amplitud de desplazamiento aleatorio de los ligandos se fijó en 3 Å por dos razones: en primer lugar para evitar la generación de conformaciones poco útiles con

el ligando fuera del bolsillo de unión y, en segundo lugar, porque el interés principal es generar modos de interacción alternativos del ligando con residuos vecinos del sitio de unión. A continuación, se llevó a cabo el refinamiento del sitio de unión para cada uno de los modelos ya que el aporte de flexibilidad al receptor es esencial para predecir de forma más exacta la orientación del ligando y las interacciones dentro del sitio de unión.

En la Figura 14 se muestran los modelos estructurales con sus ligandos correspondientes ubicados en el sitio de unión ortostérico y en la Figura 15 se puede observar la posición obtenida de WIN 55,212-2 y SR-144,528 en sus respectivos modelos, una vez realizado el método de LSHM. Debido a que hubo dos conformaciones de WIN 55,212-2 que resultaron con una buena puntuación, se exhiben ambas conformaciones (A y B). Es para destacar que el *docking* de SR-144,528 tiene similitud con la pose de AM6538 en la estructura cristalina de CB1 (pdb: 5tgz).

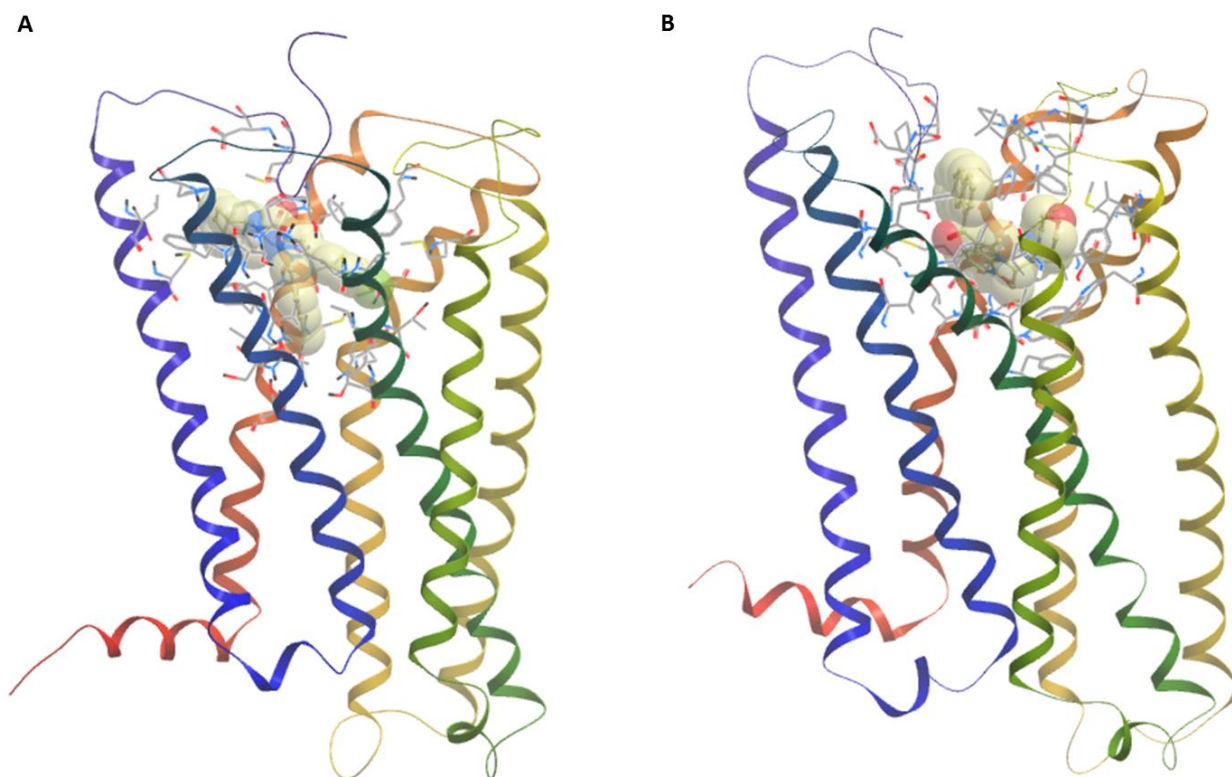


Figura 14. A) Representación del modelo de CB2 en su conformación antagonista/ agonista-inversa, con el agonista inverso SR-144,528 ubicado en el sitio de unión ortostérico. **B)** Representación del modelo de CB2 en su conformación agonista, con el ligando WIN 55,212-2 ubicado en el sitio de unión ortostérico. Además, en ambas figuras se muestran los residuos involucrados en el sitio de unión.

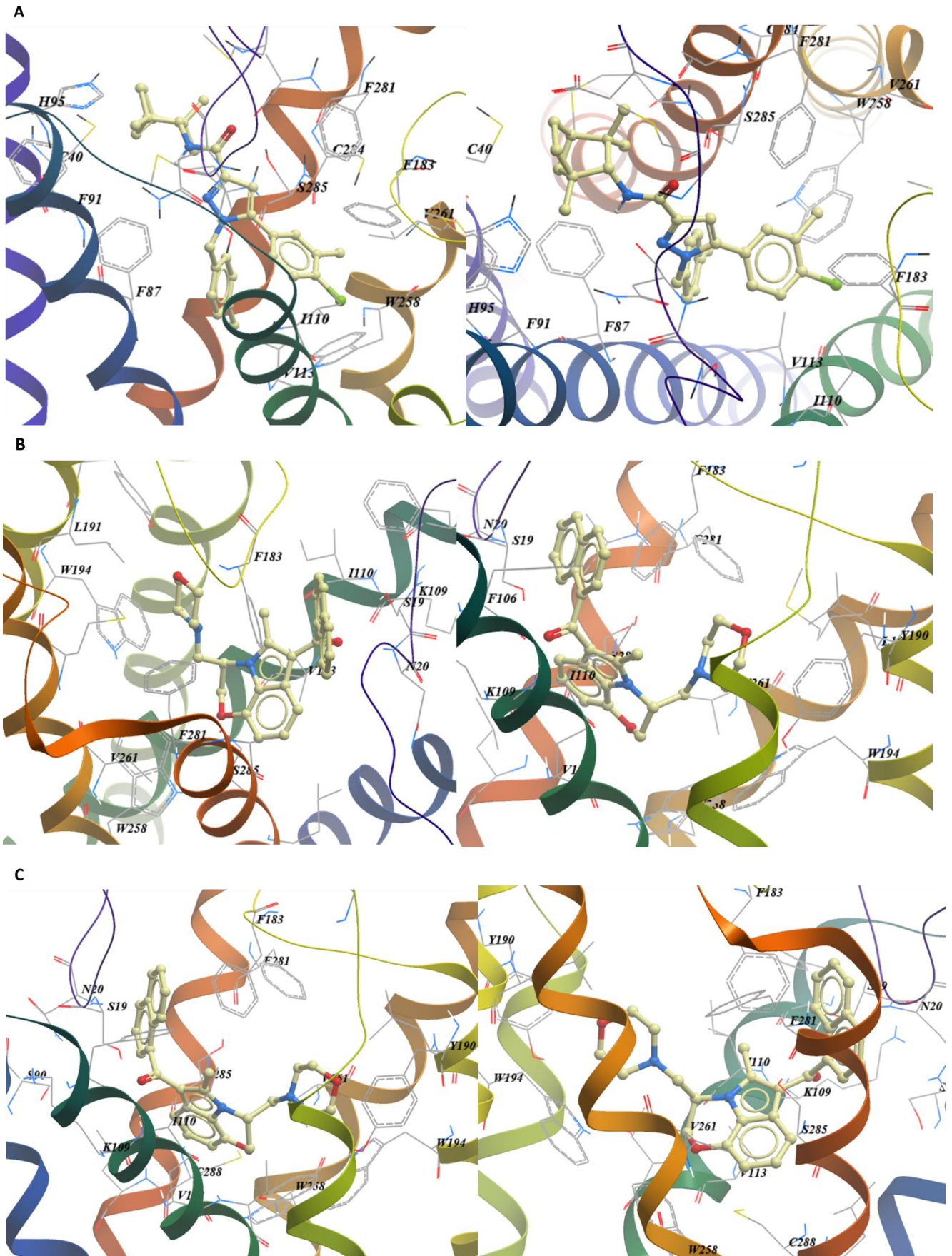


Figura 15. **A)** Representación de la pose del agonista inverso SR-144,528 ubicado en el sitio de unión ortostérico en CB2. **B)** Representación de la pose A del agonista WIN 55,212-2 ubicado en el bolsillo de unión. **C)** Representación de la conformación B del agonista WIN 55,212-2 ubicado en el sitio de unión. En las tres figuras se muestran los residuos del sitio de unión ortostérico de CB2.

No se tiene en cuenta la energía libre de cada modelo a la hora de compararlos entre sí ya que el sistema no se encuentra equilibrado en una membrana y, por lo tanto, no pasa a ser tan relevante. Para evaluar la calidad de los modelos obtenidos se realiza tanto un *docking* automatizado como una evaluación visual teniendo en cuenta los datos experimentales de mutagénesis sitio dirigidas. Los modelos de mejor calidad típicamente exhiben un factor de enriquecimiento más alto en los procedimientos de *docking*.

GLL/GDD

Para construir la GDD, primero se realizó un agrupamiento (*clustering*) de los agonistas en base a la similitud estructural de las moléculas, generando así una biblioteca diversa de 100 agonistas (a partir de 996 ligandos agonistas). Luego, para cada ligando se generó un set de moléculas señuelo (39 moléculas señuelo por ligando). Por último, se descartaron las moléculas señuelo duplicadas.

Este proceso se repitió para la biblioteca de ligandos antagonista/ agonista inverso pero, en este caso, para generar la GDD se utilizó la totalidad de ligandos antagonistas y agonistas inversos (46) presentes en la GLL.

Los resultados obtenidos en la generación de la GLL y GDD se presentan en la Tabla 5.

Tabla 5. Cantidad de ligandos de CB2 y moléculas señuelo para cada una de las bibliotecas construidas junto con el total de moléculas presentes en las bases de datos utilizadas para el posterior *docking* automatizado en modelo agonista y antagonista/ agonista inverso.

Librería	Agonistas	Antagonistas	Ago. inversos
GLL	100	34	12
GDD	3900	1794	
Total moléculas	4000	1840	

Docking molecular de pequeña escala para validación de los modelos

En esta etapa, la calidad de los modelos es de suma importancia, ya que en el procedimiento de *docking* se lleva a cabo una discriminación de los posibles ligandos presentes en una base de datos. En este caso, las bases de datos que se proveyeron se obtuvieron de fusionar la GLL con su GDD correspondiente, siendo el número total de moléculas 4000 para la base de datos del modelo agonista y 1840 para el modelo antagonista/ agonista inverso.

En la Tabla 6 y 7 se exhiben los resultados del *docking* automatizado realizado para el modelo antagonista/ agonista-inverso y para los modelos agonistas de CB2. También se llevó a cabo un *docking* en los modelos del grupo desarrollados años atrás (Diaz et al., 2009a; Petrov et al., 2013), 2009 modelo antagonista/ agonista-inverso y 2013 modelo agonista, con el fin de comparar la calidad de los mismos. Mientras que el modelo antagonista/ agonista-inverso del 2009 fue desarrollado a partir de la estructura cristalina del receptor adrenérgico β_2 (ADRB2) en su conformación inactiva (Cherezov et al., 2007), el modelo agonista (2013) se realizó a partir de un abordaje de múltiples plantillas, utilizando como molde la estructura 3D de A2AAR en el estado activo (Xu et al., 2011) y la estructura de S1P1 para el ECL2 y TM5.

Tabla 6. Número de antagonistas recobrados a distintos porcentajes de la base de datos leída para los modelos de CB2 en su conformación antagonista/ agonista-inverso.

% DB	2	5	10	20	50
Modelo 2009	2	4	7	16	31
Modelo nuevo	1	4	6	15	33

Tabla 7. Número de agonistas recobrados a distintos porcentajes de la base de datos leída para los modelos de CB2 en su conformación agonista.

% DB	2	5	10	20	50
Modelo 2013	0	3	5	24	59
Modelo A	1	4	9	22	55
Modelo B	0	1	2	15	48

A partir de los resultados obtenidos se calculó para cada modelo el factor de enriquecimiento en los distintos porcentajes de la base de datos leída (Tablas 8 y 9). Para ello, además del número de ligandos recobrados, se tuvo en cuenta el número de compuestos leídos para cada porcentaje y el número total de ligandos (*hits*) y compuestos en cada una de las bases de datos.

Tabla 8. Factor de enriquecimiento a distintos porcentajes de la base de datos leída, para los modelos de CB2 en su conformación antagonista/ agonista-inverso.

EF	2%	5%	10%	20%	50%
Modelo 2009	2,17	1,74	1,52	1,74	1,35
Modelo nuevo	1,09	1,74	1,30	1,63	1,43

Tabla 9. Factor de enriquecimiento a distintos porcentajes de la base de datos leída, para los modelos de CB2 en su conformación antagonista/ agonista-inverso.

EF	2%	5%	10%	20%	50%
Modelo 2013	0	0,60	0,50	1,20	1,18
Modelo A	0,50	0,80	0,90	1,10	1,10
Modelo B	0	0,20	0,20	0,75	0,96

No es de sorprender un bajo EF a un 2% de la base de datos leída ya que la GDD aporta mucha exigencia al *docking* en comparación con una biblioteca randomizada, debido a que los compuestos señuelo tienen propiedades físicas similares a la de los ligandos.

Un ligando podría no ser detectado como tal (es decir, tener un puntaje bajo) durante el proceso de *docking* por varias razones:

- Por la generación de choques (*clashes*) con el receptor debido a que los contactos proteína-ligando no son lo suficientemente fuertes como para mantener el ligando en una posición correcta.
- En el caso que la posición del ligando sea la correcta pero los contactos ligando-proteína no sean óptimos debido a fallas en la estructura del receptor (por ejemplo, por una mala calidad del modelo).
- Ciertos ligandos se unen al receptor a través de moléculas de agua o de algún ion particular. Si la estructura no cuenta con estas moléculas, entonces esos ligandos no serán unidos y no obtendrán un buen puntaje.
- Debido a errores en el estado de ionización del ligando o del receptor.
- Pueden no ser reconocidos como ligandos debido a un muestreo conformacional insuficiente.
- Debido a fallas en la función de puntuación del *software* utilizado.

Discusión

Importancia de la resolución de estructuras experimentales de GPCRs

En el modelado de GPCRs ocurre que, a pesar de tener un bajo porcentaje de similitud de secuencias (entre 20-30%), éstas proteínas presentan una alta conservación de la estructura global en la forma de siete hélices transmembrana lo cual permite obtener modelos con buenas aproximaciones. Sin embargo, si la similitud de secuencias en partes funcionales (como por ejemplo, el sitio de unión) no es muy alta, como es el caso de GPCRs filogenéticamente lejanos, en el proceso de modelado por homología se obtendrán modelos con mayor incerteza no sólo en la conformación de las cadenas laterales sino también en la posición del esqueleto de la proteína (*backbone*). Es por esta razón que, a pesar de la difícil tarea que implica obtener estructuras cristalinas, es de suma importancia la resolución experimental de distintas familias de GPCRs, ya que una mayor representación en las cinco familias permitiría avanzar hacia el descubrimiento de drogas basado en la estructura de receptores que podrían llegar a ser importantes blancos terapéuticos. Dado que en el presente todavía faltan muchas estructuras por resolver, es trascendente el desarrollo de métodos y modelos más precisos que permitan obtener, al menos, una representación del sitio de unión lo más exacta posible.

Si se comparan los modelos obtenidos en este trabajo con modelos de CB2 publicados años atrás (Xie, Chen and Billings, 2003; Tuccinardi et al., 2006), se puede afirmar que estos últimos son más imprecisos debido que hasta el año 2007 sólo se contaba con la estructura cristalina del receptor Rhodopsina Bovina, el cual no es tan cercano a CB2 en la filogenia. Recién en los últimos años han sido descubiertas las estructuras cristalinas de un mayor número de GPCRs, entre ellos algunos que tienen mayor similitud de secuencia con CB2: S1P1 (Hanson et al., 2012) y LPA1 (Chrencik et al., 2015) en sus conformaciones inactivas unidas a un antagonista.

La reciente resolución de las estructuras de S1P1 y LPA1 fue de gran ayuda en el comienzo del trabajo para poder obtener modelos mejorados de CB2 en su conformación antagonista/ agonista-inverso, debido a la cercanía filogenética y a la similitud de secuencia. Esto último se puede notar en el caso de las hélices 2 y 5 (Figura 16), donde las tres secuencias no presentan el residuo Prolina conservado en muchos otros GPCRs y, además, poseen la misma cantidad de deleciones, es decir que no poseen espacios (*gaps*) entre ellas (Gonzalez et al., 2012). Estas similitudes permitieron hipotetizar que tanto la hélice 2 como la 5 del modelo de CB2 serían α -hélices canónicas (hélices con ángulos de giro de 100° , $\sim 3,6$ residuos por vuelta), igual

a las presentes en LPA1 y S1P1. Una vez publicada la estructura cristalina de CB1 a finales del 2016, los modelos obtenidos a partir de S1P1 y LPA1 han pasado a un segundo plano ya que CB1 está relacionado de forma mucho más directa a CB2 y, por lo tanto, permitiría llegar a un modelo más preciso y de mejor calidad. Tal cual lo esperado para CB2, la estructura de CB1 también presentó las α -hélices canónicas en los dominios transmembrana 2 y 5. En cuanto al N-terminal, se puede ver una clara diferencia entre las tres estructuras cristalinas de S1P1, LPA1 y CB2 (Figura 17, A). Mientras que en CB1 forma un lazo en forma de V que se introduce hacia el sitio de unión, en S1P1 y LPA1 forma una α -hélice que recubre el sitio ortostérico siendo la longitud de la misma variable entre estas dos estructuras. Además, se puede observar que en CB1 el sitio ortostérico es más amplio debido a un corrimiento hacia fuera de las hélices 1 y 2. Si se comparan los ECL, se ve un ECL2 muy similar en las tres estructuras (lo cual podría deberse al puente disulfuro altamente conservado) pero, en el caso de CB1, se ve un ECL1 sin región α -hélice y un ECL3 más corto debido a un alargamiento de la hélice 7 (Figura 17, B). Todas estas diferencias y similitudes deben ser tenidas en cuenta a la hora de seleccionar la proteína molde. Es por este motivo que, al momento de continuar el modelado de CB2 en su conformación inactiva, se cree más adecuada la elección del modelo obtenido a partir de CB1.

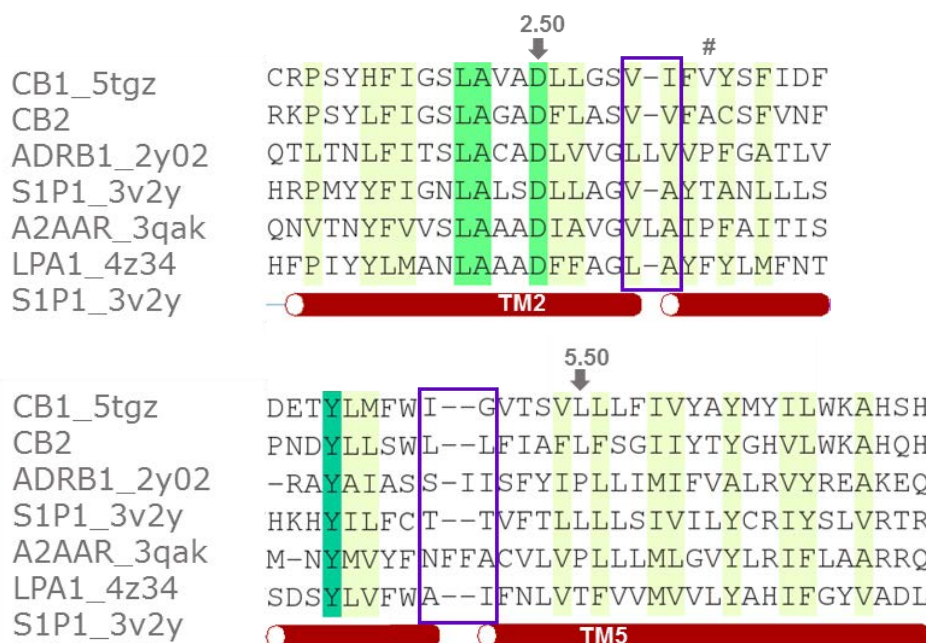


Figura 16. Alineamiento de las secuencias de TM2 y TM5 de los GPCRs CB2, CB1, S1P1, A2AAR, ADRB1 y LPA1. Con un recuadro violeta se destacan las posiciones en las cuales LPA1, S1P1 y CB2 poseen la misma cantidad de deleciones. En la hélice 2 se marca con un numeral (#) la posición en la cual estos tres receptores no presentan el residuo prolina, mientras que en la hélice 5 dicha posición es la 5.50. Los dominios TM de la estructura de S1P1 se muestran en forma de cilindros.

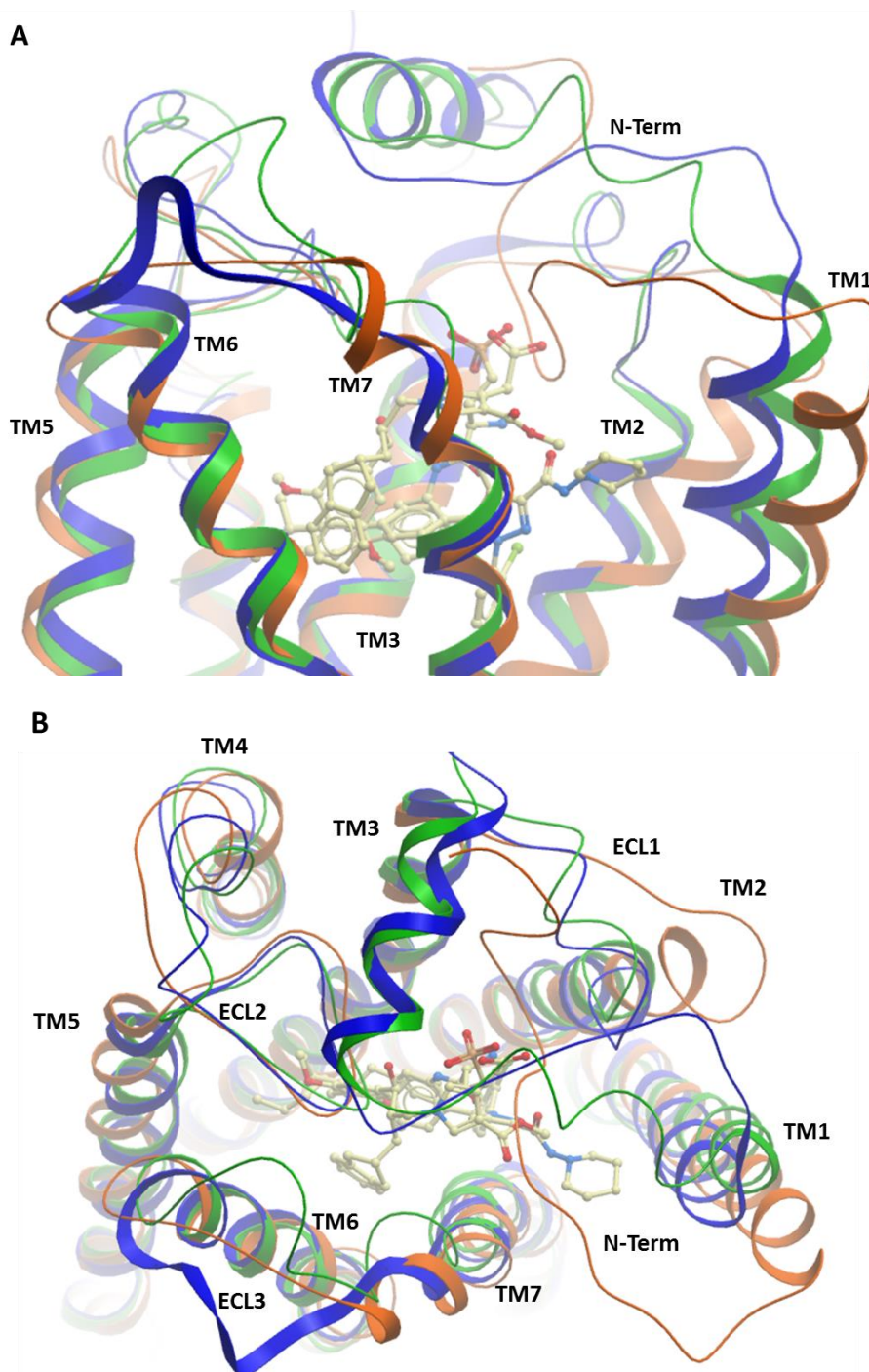


Figura 17. Superposición de las estructuras experimentales de CB1 (color naranja), S1P1 (color verde), y LPA1 (color azul). **A)** Vista lateral. En la misma se observan las diferencias en el extremo N-terminal y las hélices 1 y 2 de CB1 despalzadas hacia fuera. **B)** Vista superior (desde lado extracelular). En CB1 se observa ECL1 sin región α -hélice y una TM7 más larga llevando a un ECL3 más corto. ECL2 semejante en los tres receptores.

Por otro lado, CB2 en su conformación agonista se realizó a partir de un abordaje con múltiples plantillas ya que al día de hoy no existen estructuras cristalinas en

conformación activa que tengan una alta homología de secuencia con CB2. Inicialmente se seleccionó como plantilla la estructura de ADRB1 porque, de las estructuras cristalinas de GPCRs presentes en conformación activa, es una de las que presenta mayor similitud de secuencia con CB2 a pesar de no ser un receptor tan cercano filogenéticamente. Dado que en el año 2013 en el grupo se realizó un modelado de CB2 a partir de la estructura de A2AAR en su conformación activa (pdb: 3qak), para este modelado no se tuvo en cuenta a dicho receptor. El N-terminal, junto con las hélices 2 y 5, se modelaron a partir de CB1 debido a que en esas regiones es donde existe una menor conservación de secuencias entre CB2 y ADRB1.

Limitaciones del modelado por homología

Durante el proceso de optimización de los modelos, el paso de minimización de la energía global tiene como objetivo aliviar la tensión estructural derivada del reemplazo de residuos no conservados en el proceso de modelado por homología. Fue de importancia que este paso se realizara con el ligando situado en el bolsillo de unión, ya que al minimizar los modelos crudos sin ligando se colapsaba dicho bolsillo. Esto se debe a que, en comparación con otros GPCRs de Clase A, CB1 presenta un sitio ortostérico más amplio, por lo que al minimizar el modelo sin ligando el espacio del bolsillo va a tender a reducirse y de esa forma minimizar la energía global. Al realizar la minimización con el ligando situado en el bolsillo de unión, se logra conservar el espacio inicial del sitio heredado de la estructura usada como molde.

Es un paso habitual tratar de incorporar información sobre la pose de unión del ligando utilizado en el proceso de refinado del sitio ortostérico del receptor. Sin embargo, al momento de situar los ligandos WIN 55,212-2 y SR-144,528 en los respectivos modelos 3D obtenidos, debido a la falta de información experimental y a que no hay un consenso sobre los residuos involucrados en el sitio de unión de CB2, no se pudo fijar ninguna ligadura en el complejo receptor-ligando (por ejemplo, puente de hidrógeno) con el fin de reducir los grados de libertad.

Debido a la falta de cálculos precisos de la energía de unión y a limitaciones en el modelado del sistema (como por ejemplo, la falta de fosfolípidos de la membrana celular y de moléculas de agua, datos de mutagénesis insuficientes y ausencia de evidencia experimental sólida que pueda ser utilizada como restricciones estructurales) se cree que es mejor validar los modelos obtenidos en el trabajo a través de datos experimentales del receptor y un *docking* de pequeña escala.

Aminoácidos importantes en el sitio ortostérico de CB2

En cuanto a la información experimental recolectada en el presente trabajo, Feng *et al.* (Feng et al., 2014) demostró la importancia de V113^{3.32} y F183^{ECL2} en la unión de SR-144,528 y CP 55,940 a CB2. El estudio de ciertos aminoácidos de la hélice 4, dió como resultado que la modificación de S161^{4.53} y S165^{4.57} por alanina, llevaría a una pérdida de afinidad por SR-144,528 (Gouldson et al., 2000). Dado que la hélice 4 por lo general no se encuentra implicada en el sitio de unión de los GPCRs de Clase A, este cambio podría atribuirse a un efecto alostérico. La pérdida de afinidad de SR-144,528 al mutar C174^{4.66} o C179^{ECL2} por una serina se debe a que no se puede formar el puente disulfuro del ECL2, el cual parece ser importante en el ensamblaje del sitio de unión ortostérico. En cuanto al residuo Y190^{5.39}, estaría implicado en la unión de CP 55,940 ya que al cambiarlo por un aminoácido alifático (isoleucina) se pierde la unión. Al modificarlo por fenilalanina, otro residuo aromático, el ligando WIN 55,212-2 mantiene la unión, por lo que la presencia de un residuo aromático en esa posición sería de importancia (McAllister et al., 2002). Por otra parte, W194^{5.43} parecería tener implicancia en la unión de WIN 55,212-2 y SR-144,528 al sitio de unión, ya que al modificar este aminoácido por fenilalanina o alanina se da una pérdida de la unión (Zhang et al., 2011). La modificación del aminoácido W258^{6.48} por alanina o fenilalanina dan como resultado una pérdida de la unión de WIN 55,212-2, por lo que sería un aminoácido importante en el sitio de unión ortostérico (Nebane et al., 2008). En el mismo trabajo se estudió el residuo P260^{6.50} y se obtuvo que al mutarlo por cisteína da como resultado una pérdida de la unión de CP 55,940. Al estudiar el residuo F281^{7.35}, se observó que su reemplazo por un aminoácido cargado positivamente (lisina) produce la pérdida de unión tanto de SR-144,528 como de CP 55,940 (Feng et al., 2014). Estos resultados permiten hipotetizar que la presencia de un aminoácido aromático, o al menos sin carga, en esa posición es de importancia para la unión de esos ligandos.

Caracterización de la interacción CB2-antagonista/ agonista inverso

Al observar el modelo de CB2 en la conformación antagonista/ agonista-inverso se puede decir que los residuos que estarían involucrados en el sitio de unión ortostérico serían de las hélices 1, 2, 3, 6 y 7, junto con algunos del N-terminal y del ECL2. Los residuos pertenecientes a las hélices 4 y 5 que fueron identificados como relevantes para la unión de SR-144,528 (S161^{4.53}, S165^{4.57}, C175^{4.67}, L192^{5.41} y W194^{5.43}) no se encuentran dentro de distancias de contacto, por lo que estos aminoácidos podrían

tener contribuciones indirectas en la unión a través de la estabilización estructural de la conformación de esas hélices. Si se compara éste modelo con la estructura experimental de CB1, en el último caso el bolsillo de unión está también conformado por la hélice 5. Al ser AM6538 (antagonista junto al cual se cristalizó a CB1) un ligando que posee una cadena carbonada de mayor longitud que SR-144,528, es esta característica la que le permite interactuar con la hélice 5 de forma directa.

Si se analiza el modelo obtenido en función de los datos de mutagénesis recolectados (Tabla 2): la presencia de un aminoácido no polar alifático en la posición 113 (valina) es de importancia para que los grupos aromáticos del agonista inverso se ubiquen en la parte inferior del bolsillo de unión. La mutación por un residuo cargado, como lo es el ácido glutámico, generaría una repulsión e impediría que el ligando se pueda unir. Al modificar C175^{4,67} por una serina, se observa una pérdida de la unión de SR-144,528. Por estar lejos del sitio de unión, se puede inferir que en realidad esa modificación estaría afectando la estructura del ECL2 lo cual llevaría a la pérdida de la afinidad del ligando ya que ciertos residuos de ese lazo están implicados en la unión. Éste es el caso de la fenilalanina 183 que se encuentra haciendo una interacción π - π de tipo T con uno de los grupos aromáticos del ligando (es decir, los anillos aromáticos interaccionarían de forma perpendicular). Al cambiar este aminoácido por una valina (residuo alifático) o por una glutamina (residuo polar) esta interacción se pierde. En cuanto a los aminoácidos estudiados de la hélice 7, se vio que la mutación de la fenilalanina 281 por una lisina (aminoácido polar cargado) lleva a una pérdida de unión, lo cual condice con el modelo obtenido ya que ese residuo se encuentra cubriendo la parte superior del bolsillo y formando, junto al aminoácido F183, una zona hidrofóbica. Por otro lado, los aminoácidos C284^{7,38} y C288^{7,42} serían de importancia para la unión de otro agonista inverso de CB2, AM1336. Si bien en el modelo la cisteína 288 apunta más hacia el bolsillo que la de la posición 284, ambas podrían estar involucradas en la unión.

Otros residuos que serían de importancia para la unión de SR-144,528 para los cuales no existe información experimental son: Asn20, Met22, Phe87^{2,57}, Phe91^{2,61}, Ile110^{3,29}, Trp258^{6,48}, Val261^{6,51}, Ser285^{7,39}, Leu289^{7,43}. En el caso de la serina 285, podría formar un puente de hidrógeno con el oxígeno del grupo carbonilo de SR-144,528, ya que la distancia entre el hidrógeno del grupo hidroxilo y el átomo de oxígeno es de 3.4 Å (Figura 18).

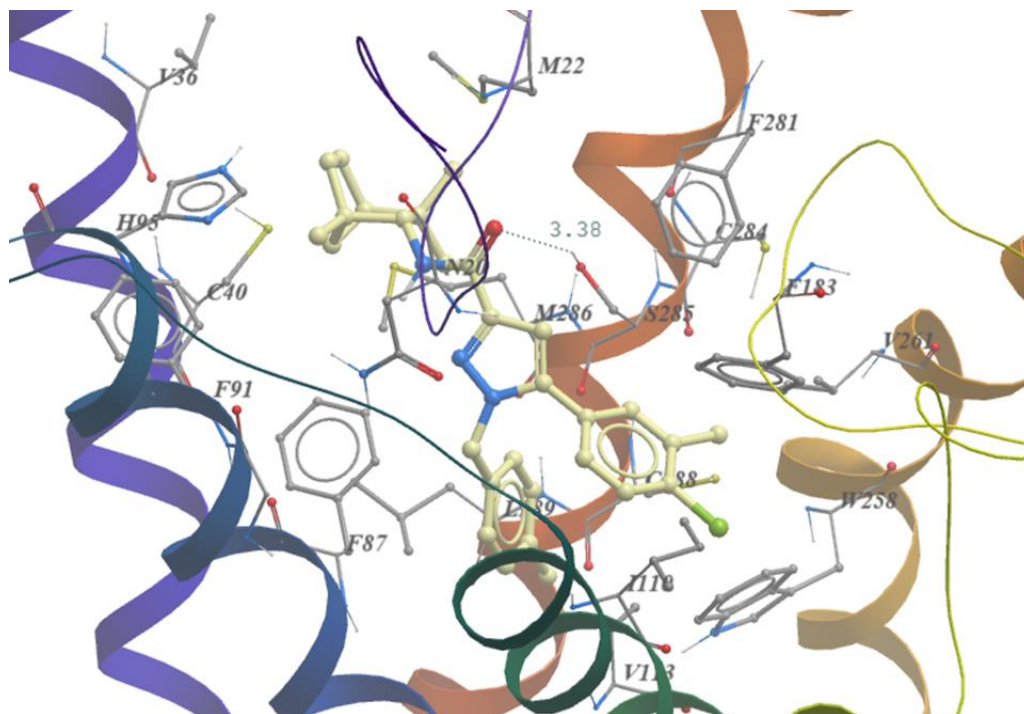


Figura 18. Modelo de CB2 en la conformación antagonista/ agonista inversa con la pose obtenida para el agonista inverso SR-144,528 y la distancia (Å) marcada entre éste y Ser285.

Caracterización de la interacción CB2-agonista

Si se analiza el modelo A de CB2 en la conformación agonista se observa que el bolsillo de unión está formado por las hélices 2, 3, 5, 6 y 7, junto con el N-terminal y el ECL2. Varios residuos involucrados en el sitio de unión del modelo obtenido tienen información experimental que los sustentan. Tal es el caso de la V113^{3,32}, donde se vio que al mutarla por un ácido glutámico se pierde la afinidad del agonista CP 55,940 pero al modificarla por otro aminoácido alifático (leucina) la afinidad se mantenía. En el caso del modelo este residuo se encuentra en la parte inferior del bolsillo, formando una zona hidrofóbica junto a otros aminoácidos y permitiendo el posicionamiento del grupo indol de WIN 55,212-2. En cuanto al ECL2, existen datos de mutagénesis del residuo F183 asociado al ligando CP 55,940: al mutarlo por una valina o por una glutamina se da una pérdida en la unión. Siendo que en el modelo éste residuo y el grupo naftaleno de WIN 55,212-2 se encuentran formando una interacción π - π de tipo *stacking* (es decir que los grupos aromáticos interaccionan ubicándose de forma paralela), la mutación por un aminoácido que no sea aromático llevaría a la pérdida de esta interacción, concordando con lo que se obtuvo para el agonista estudiado por Feng *et al.* (Feng *et al.*, 2014).

De la hélice 5 se estudiaron varios residuos a lo largo de distintos trabajos. Al modificar la tirosina 190 por una fenilalanina, no se observaron cambios en la afinidad

de WIN 55,212-2, pero al cambiar este residuo por una isoleucina (residuo alifático que no posee anillo aromático) se da una pérdida en la unión de CP 55,940. A pesar que en el modelo éste residuo se encuentra un poco alejado del sitio de unión (ya que apunta más hacia la hélice 4), por su volumen podría de alguna forma estar aportando hidrofobicidad en esa zona para la inserción del grupo morfolino, como también podría estar estabilizando la posición del residuo W194^{5.43}. En ambos casos, dichos aportes se perderían al colocar en esa posición un aminoácido que no sea aromático. En el estudio realizado por Alqarni *et al.* donde se mutó la lisina 192 (Alqarni *et al.*, 2014) obtuvieron resultados que no coinciden con el modelo obtenido. El reemplazo de L192^{5.41} por un aminoácido polar, como es la serina, da una pérdida en la unión de CP 55,940; en nuestro modelo dicho aminoácido se encuentra apuntando hacia fuera del bolsillo de unión. Sin embargo, en el modelo A se halla la L191^{5.40} apuntando hacia el sitio de unión y parecería estar involucrada en la unión de WIN 55,212-2, ya que está formando interacciones de van der Waals con el grupo morfolino. Además, este dato de mutagénesis no es coherente con los datos obtenidos para Y190^{5.39} y W194^{5.43} (como se discutirá a continuación) debido a que en una vuelta entera de una α -hélice canónica hay aproximadamente 3,6 residuos, por lo que un aminoácido que se encuentra a dos residuos de distancia de otro estará en la dirección opuesta en la hélice (tal como se obtuvo en el modelo para la lisina 191). Sobre la serina 193, su modificación por una glicina no afecta la afinidad del ligando, lo cual es acorde al modelo ya que dicho residuo no aporta una interacción con WIN 55,212-2. Tal como se dijo anteriormente, el residuo W194^{5.43} parecería ser de importancia para la unión de WIN 55,212-2: al mutarlo por una tirosina se observó una reducción en la afinidad, mientras que al mutarlo por una fenilalanina o alanina se ve una pérdida en la unión. Esto lleva a pensar que, o bien el triptófano es un residuo importante en el sitio de unión por su volumen y aromaticidad, o bien está involucrado en un puente de hidrógeno con el ligando (por los cambios de afinidad que se ven al modificarlo). En el modelo obtenido éste residuo estaría aportando interacciones hidrofóbicas con el grupo morfolino del ligando y, por lo tanto, al modificarlo por un residuo aromático de menor volumen (como es la tirosina o fenilalanina) o por uno alifático, esta interacción se perdería. En la hélice 5 también se estudió la F197^{5.46} y se vio que su modificación por una valina causa una reducción de 14 veces en la afinidad de WIN 55,212-2 por CB2. Al observar el modelo, este residuo apunta hacia fuera del bolsillo y, si se compara su posición con la del aminoácido correspondiente en los demás receptores incluidos en el alineamiento que no cuentan con inserciones en la hélice 5, se ve que ésta coincide: en ninguno de los casos el residuo de la posición 5.46 apunta hacia el sitio de unión (Figura 19, A). Sólo ocurre esto en el caso de los receptores ADRB1 y

A2AAR, los cuales cuentan con inserciones (de 1 y 2 aminoácidos respectivamente) en este paso transmembrana y, como resultado, tienen una vuelta más amplia en esa zona (Figura 19, B).

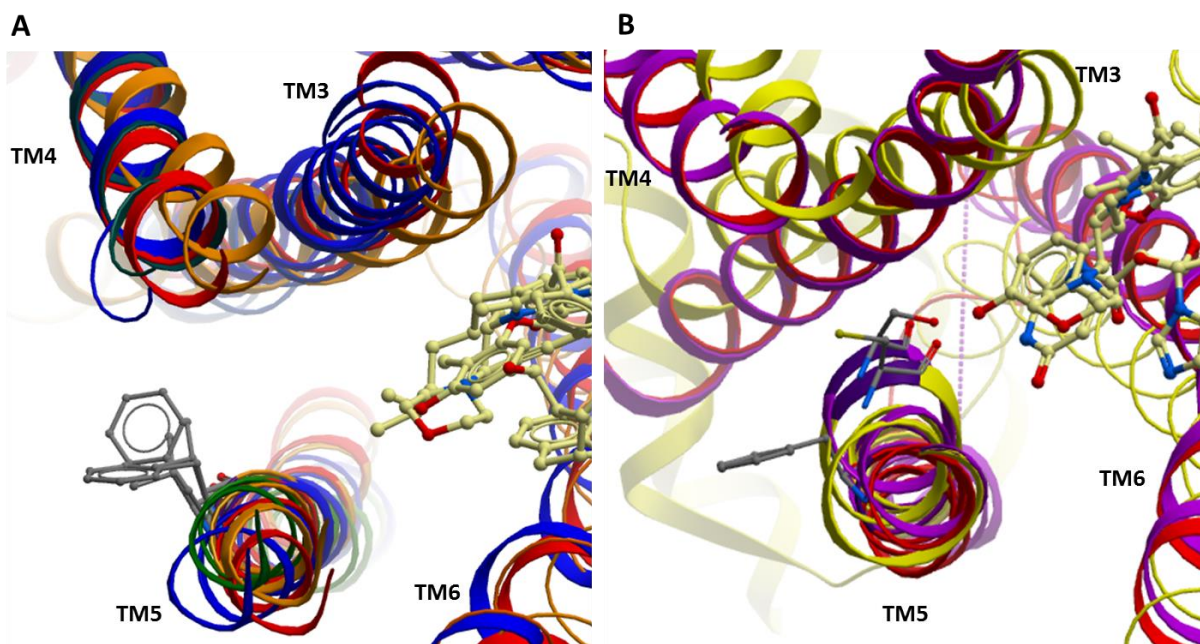


Figura 19. A) Superposición de la estructura de CB1, S1P1, LPA1 y el modelo A de CB2, con los residuos de la posición 5.46 desplegados. **B)** Superposición de la estructura de ADRB1 y A2AAR con los residuos de la posición 5.46 exhibidos. En ambas figuras se muestran los respectivos ligandos.

Cuando el residuo C257^{6,47}, que en el modelo obtenido se encuentra apuntando hacia la hélice 7, es mutado por una lisina, isoleucina, leucina, ácido aspártico o tirosina se produce una pérdida de unión del ligando CP 55,940. Al mutarlo por una alanina o serina, no hay cambios en la afinidad de WIN 55,212-2. Estos resultados experimentales conciben con lo obtenido en el modelo: la mutación por un residuo no conservativo de mayor volumen en esa posición llevaría a una desestabilización de la conformación por choques con los residuos de la hélice 7, mientras que el reemplazo por un aminoácido de tamaño similar (como es la alanina o serina) no produciría dicho efecto. Se estudió además el triptófano 258 y se vio que al mutarlo por una alanina o fenilalanina se produce la pérdida de unión de WIN 55,212-2, mientras que la mutación por una cisteína se estudió con CP 55,940 dando el mismo resultado. Dado que en el modelo dicho triptófano se encuentra formando una zona hidrofóbica en la parte inferior del bolsillo junto con F117^{3,36} y W194^{5,43}, su modificación por un aminoácido alifático o uno aromático de menor volumen llevarían a una pérdida de contacto

(interacciones de van der Waals) del residuo con el ligando. Al mutar el residuo P260^{6.50} por cisteína, a pesar de estar lejos del sitio de unión, se vio una pérdida de la unión de CP 55,940 lo cual puede atribuirse a un efecto alostérico ya que el aminoácido prolina suele aportar quiebres en las α -hélices, como ya se mencionó anteriormente. El aminoácido F281^{7.35} es de importancia en esa posición ya que al introducir un aminoácido cargado (lisina) se pierde la afinidad por CP 55,940. En el modelo se observa que dicho aminoácido forma interacciones π - π de tipo T con el residuo F183 (ECL2), por lo que podría estar estabilizando su posición y como consecuencia la interacción con el ligando. En último lugar, la serina 285 sería de importancia para la unión del agonista HU-243 ya que al reemplazarla por una alanina se da una reducción de 13 veces en la afinidad. Dicho aminoácido se encuentra formado parte del bolsillo de unión en el modelo A.

Entre los residuos que no existe información experimental al día de la fecha y en el modelo obtenido parecerían ser de importancia se encuentran: Ser19 (N-terminal), Phe106^{3.25}, Ile110^{3.29}, Ile186 (ECL2), Leu191^{5.40}, Leu289^{7.44}. En este modelo, la serina 19 podría formar un puente de hidrógeno con el grupo carbonilo de WIN 55,212-2 ya que el hidrógeno del grupo hidroxilo de éste residuo se encuentra a 2.9 Å de distancia del oxígeno del carbonilo (Figura 20).

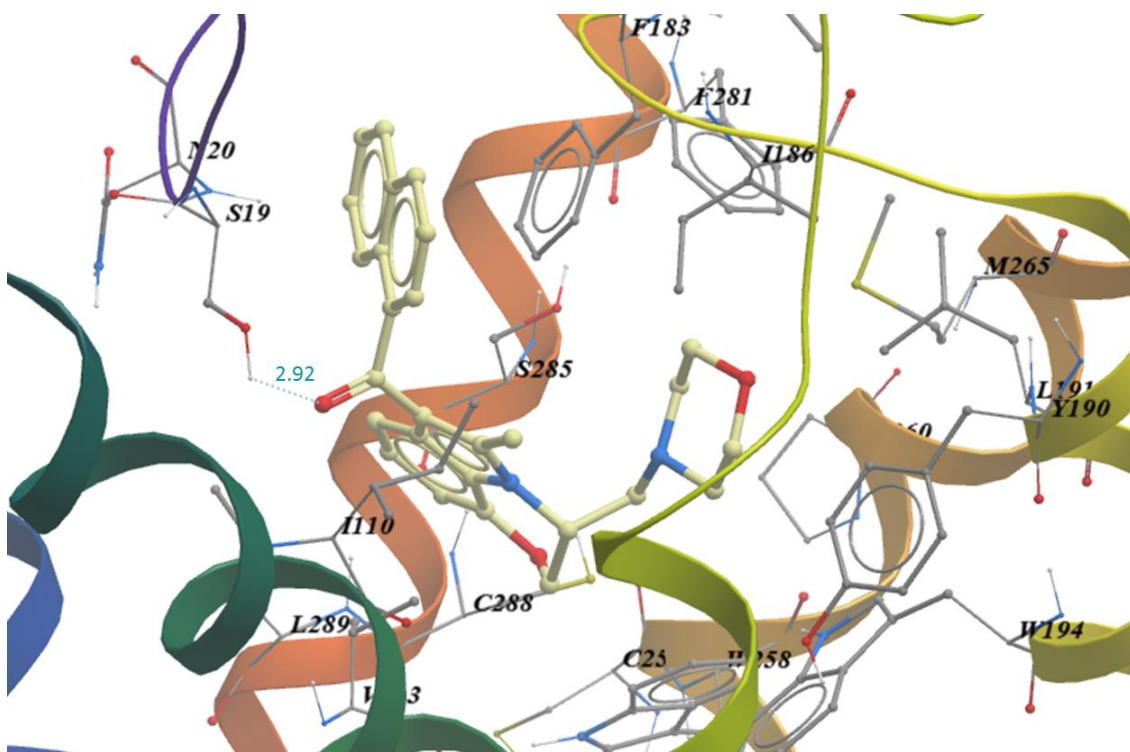


Figura 20. Modelo A de CB2 en conformación agonista con la pose obtenida para el ligando WIN 55,212-2 y la distancia (2,92 Å) marcada entre el O del grupo carbonilo y el H del grupo hidroxilo de Ser19.

El modelo B de CB2 en conformación agonista es similar al modelo A en función de los residuos involucrados en el sitio de unión (Figura 21). Presenta algunas diferencias si se lo analiza teniendo en cuenta los datos experimentales, como es el caso del residuo F183^{ECL2} que en el modelo B no se encuentra posicionado de forma totalmente paralela al grupo naftaleno de WIN 55,212-2 (interacción π - π menos favorable). El residuo triptófano 190 se encuentra en la misma posición que en el modelo A pero, al estar corrido más hacia la hélice 5 el grupo morfolino del ligando, se da una interacción de van der Waals más directa. Lo mismo ocurre para el aminoácido W194^{5.43}: el grupo morfolino se encuentra más cercano en este modelo. El W258^{6.48} también se encuentra en la misma posición y está más cercano al grupo indol de WIN 55,212-2 debido a la posición que adoptó el ligando, con éste grupo situado un poco más profundo en el bolsillo de unión. El aminoácido F281^{7.35} puede interactuar de forma directa con el residuo F183 mediante una interacción π - π de tipo T. Es de notar que en la asparagina 20, el H del grupo amino que se encuentra formando el enlace peptídico está posicionado justo en el centro de un anillo del grupo naftaleno de WIN 55,212-2, dando lugar a un puente de hidrógeno descrito anteriormente en otras proteínas y que sería una interacción de importancia (Figura 22) (Perutz, Fermi and Abraham, 1986).

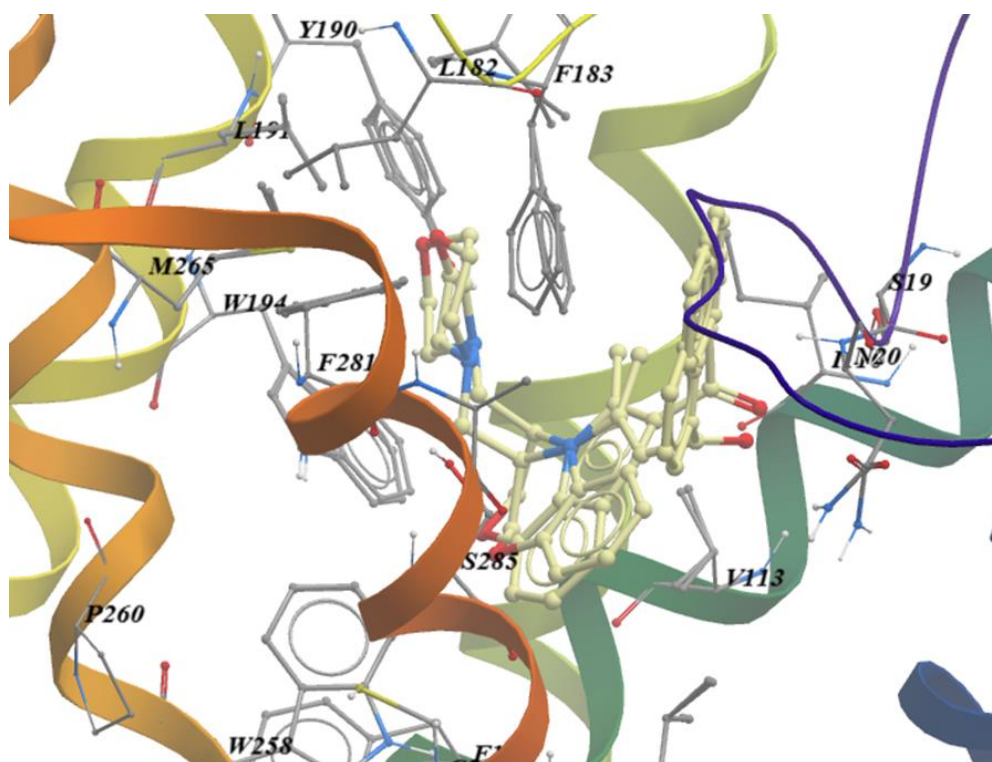


Figura 21. Superposición de los modelos A y B de CB2 en conformación agonista, con las poses obtenidas para el ligando WIN 55,212-2. Además se muestran los residuos importantes en el sitio de unión con el fin de comparar las conformaciones obtenidas.

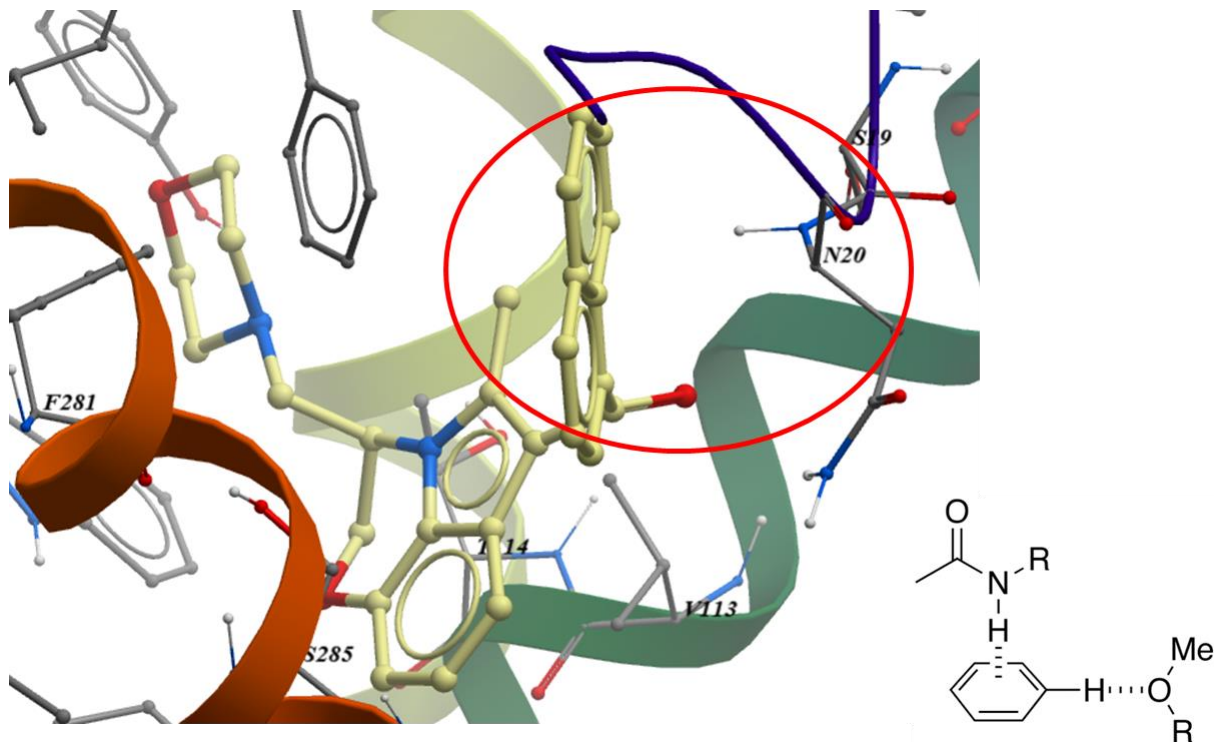


Figura 22. Modelo B de CB2 en conformación agonista (hélice 1 borrada para una mejor visualización) con WIN 55,212-2 ubicado en el sitio de unión. Se marca el H de la Asn20 apuntando directamente hacia un anillo aromático del grupo naftaleno del ligando, lo que daría lugar a la formación de un puente de hidrógeno. En el extremo inferior derecho se esquematiza la interacción de un átomo de hidrógeno con densidad de carga positiva, debido a su enlace con un átomo de N (átomo muy electronegativo), estableciendo una interacción de puente de hidrógeno con un anillo aromático.

Docking automatizado y validación de los modelos de CB2

En el procedimiento de *docking* automatizado se obtuvo como resultado valores de EF bajos para todos los modelos. Sin embargo, estos valores se encuentran dentro de los esperados para un GPCR, aún más si se tiene en cuenta la implementación de una GDD en las bases de datos correspondientes. Si se usara una biblioteca randomizada de compuestos en vez de una GDD, donde las moléculas son no-ligandos con propiedades físicas similares a los ligandos pero químicamente distintas, seguramente se obtendrían valores de EF más altos. Además, el hecho que la relación sea 39:1 (es decir, por cada 39 no ligandos hay un sólo ligando) no favorece que los ligandos sean los que obtengan mejores puntajes por una razón estadística. De los 39 no-ligandos es probable que alguno de ellos tenga un buen puntaje y eso afecta de forma directa al EF.

En cuanto al modelo antagonista/ agonista inverso se obtuvieron resultados muy similares entre el modelo obtenido en este trabajo y el del 2009 para los EF a distintos porcentajes de la base de datos leída. Esto no era de esperar ya que el modelo del 2009 se realizó a partir de la estructura cristalina de ADRB2, receptor más lejano filogenéticamente y con menor similitud de secuencia en comparación con CB1, que fue el receptor que se utilizó como molde para el modelo actual. Éste resultado evidencia que el modelo del 2009 es lo suficientemente bueno como para que los ligandos agonistas/ antagonistas inversos de CB2 se posicionen en el sitio de unión y obtengan un buen puntaje. Sin embargo, probablemente las poses adquiridas por los ligandos en este modelo no sean las adecuadas, ya que el bolsillo de unión se encuentra desplazado más hacia la hélice 4 al compararlo con el modelo obtenido a partir de CB1 (Figura 23). Dado que el bolsillo de unión en ambos modelos es lo suficientemente largo, aquellos ligandos y no-ligandos que posean una estructura química alargada y no tan ramificada podrán situarse, obteniendo buenos puntajes. Se puede ver, además, que los residuos que estarían involucrados en el sitio de unión se encuentran en posiciones similares, presentándose las mayores diferencias en el ECL2 y las hélices 2 y 5, donde el modelo 2009 heredó de ADRB2 hélices que presentan una vuelta más amplia. Además, la hélice 2 de éste modelo presenta un quiebre que lleva a que los residuos más cercanos a la parte extracelular se ubiquen más hacia dentro del bolsillo de unión, cerrando el paso hacia la hélice 1 (ayudaría al corrimiento del bolsillo hacia la TM5). En el modelo antagonista/ agonista inverso actual, TM2 y TM5 presentan α -hélices canónicas tal cual las halladas en la estructura cristalina de CB1. En la hélice 3 la principal diferencia es que el residuo lisina 109^{3,28} se encuentra formando parte del sitio de unión en el modelo 2009, mientras que en el otro modelo no. Tanto el residuo F183 como el F281^{7,35} se encuentran más alejados del sitio de unión en el caso del modelo realizado a partir de ADRB2.

Por otro lado, los ligandos utilizados para el *docking* y el posterior proceso de *ligand steered homology modelling* fueron distintos: mientras que para el modelo del presente trabajo se utilizó el agonista inverso SR-144,528, en el modelo 2009 se empleó un agonista inverso de CB2 sintetizado en ese trabajo.

Si se analizan los resultados obtenidos en el proceso *docking* para los modelos agonistas se podría decir que el modelo A sería mejor en comparación con el modelo B y el modelo del 2013. Tanto en el 2, 5 y 10% de la base de datos leída, el número de ligandos recobrados fue mayor, lo cual lleva a obtener mejores EF. En la Figura 24 se compara el modelo A con el modelo 2013, el cuál fue realizado a partir de A2AAR y la estructura de S1P1 para el ECL2 y TM5. En el caso del modelo 2013, el sitio de unión parecería ser más profundo y menos alargado debido a que el TM1 y TM2 se

encuentran más hacia dentro. Las mayores diferencias se observan en el TM2: el modelo del 2013 presenta la hélice con una vuelta más amplia heredada de la estructura cristalina de A2AAR dando como resultado que las fenilalaninas 87, 91 y 94 apunten hacia el sitio de unión. En el modelo A y B, como la TM2 se modeló a partir de la estructura de CB1, dicha vuelta no está presente y estos residuos apuntan hacia TM1. Además, el modelo 2013 tiene la tirosina 190 apuntando hacia fuera del bolsillo de unión, por lo que no puede involucrarse en la unión del ligando.

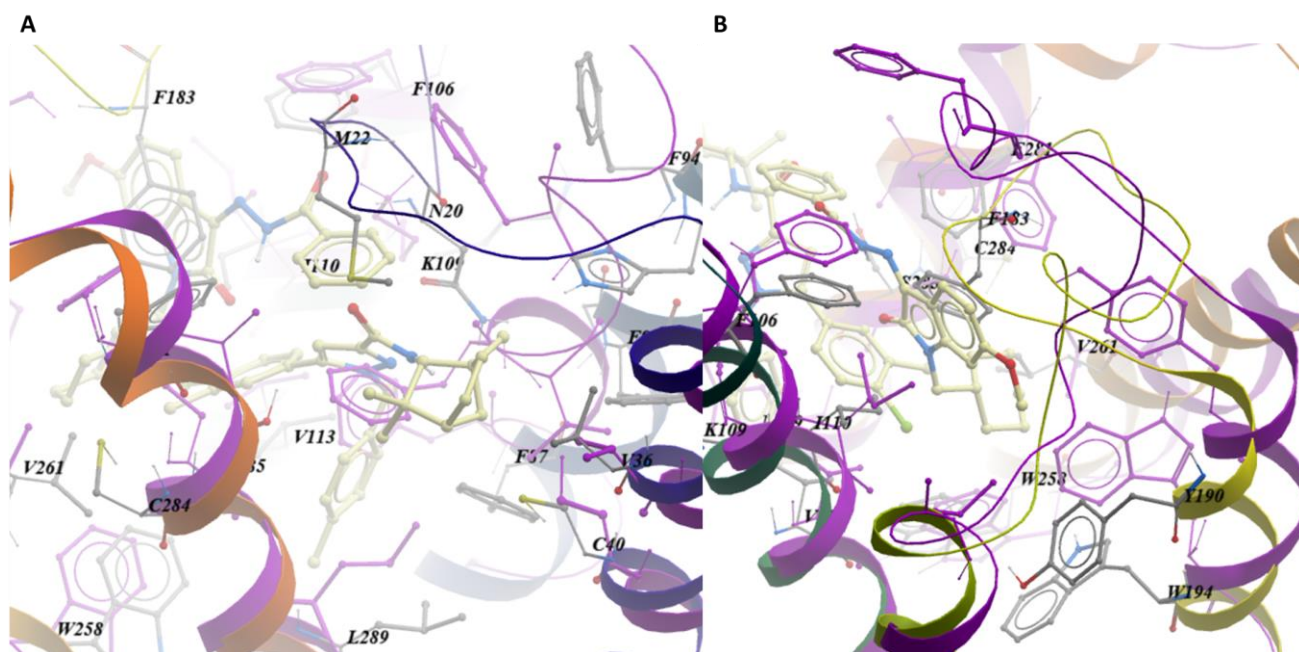


Figura 23. Superposición de los modelos 2009 (coloreado en púrpura) y 2017 (hélices coloreadas del azul al naranja) de CB2 en conformación antagonista/ agonista inverso, con la pose sus respectivos agonistas inversos. **A)** Se pueden observar las diferencias en el TM2 y el agonista inverso SR-144,528 localizado más hacia la hélice 1. **B)** Se puede visualizar las diferencias en TM5 y ECL2, además del agonista inverso del modelo 2009 ubicado más hacia la hélice 4.

Algo similar pasa con el W258^{6,48} que se halla más alejado del sitio de unión. Tanto el residuo F183 como la de F281^{7,35} del modelo 2013 se encuentran en posiciones más desfavorables para interactuar con el ligando, pudiendo aportar en el *docking* para la obtención de un menor número de ligandos recobrados a bajos porcentajes de la base de datos leída. Una diferencia notable entre el modelo 2013 y el modelo A/B es la posición de WIN 55,212-2: mientras que en los modelos A y B se halla en la parte superior, manteniendo interacciones de tipo π - π con F183, en el modelo 2013 el grupo naftaleno se ubica en la parte inferior del bolsillo, perdiéndose esta interacción. Sin embargo, éste grupo podría estar formando interacciones de tipo π - π con el residuo

fenilalanina 117 de la hélice 3, el cual está situado en la parte inferior del bolsillo (coincidiendo con el modelo A y B). En cuanto al grupo morfolino, el del modelo 2013 se ubica en la parte superior del bolsillo, pudiendo mantener interacciones de van der Waals con algunos residuos de esa zona.

En cuanto al modelo B, a pesar de tener varios residuos con conformaciones similares a las del modelo A, se obtuvo un menor número de ligandos recobrados en los porcentajes más bajos de la base de datos leída. Podría deberse a que la posición de F183 resulta menos favorable para la interacción con los ligandos agonistas. Sin embargo, como los resultados obtenidos son bastante similares se puede decir que las diferencias obtenidas en el EF no son significativas como para determinar al modelo B como uno de mala calidad.

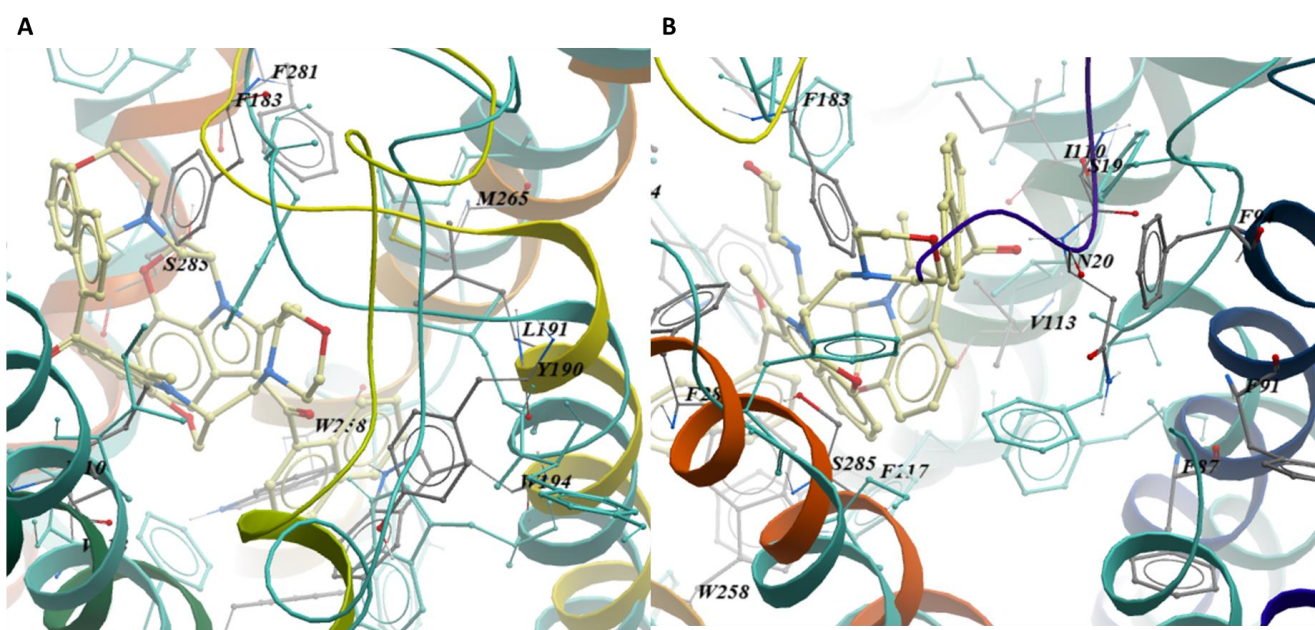


Figura 24. Superposición de modelo A de CB2 (hélices coloreadas del azul al naranja) y el modelo 2013 (coloreado en celeste), con la respectiva pose de WIN 55,212-2. **A)** Se puede ver la diferencia en la posición de F183, Y190, W258 y F281. **B)** Para una mejor visualización se borró parte de TM1. Se observa las diferencias en la TM2 y nuevamente las posiciones de F183 y F281.

Con el fin de poder diferenciar cuál de todos los modelos obtenidos presenta una mejor aproximación del sitio de unión, sería una opción para continuar este trabajo realizar un nuevo *docking* automatizado en un software distinto, por ejemplo Glide (Friesner et al., 2004; Halgren et al., 2004), para independizarse de la función de puntuación. Si los resultados continúan dando similares, entonces las fallas no

estarían en la función de puntuación del programa utilizado, sino en alguno de los demás factores mencionados anteriormente.

Conclusiones

En el presente trabajo, se lograron desarrollar y obtener dos modelos agonistas y un modelo antagonista/ agonista inverso del GPCR de clase A CB2. Es de extrema importancia seleccionar adecuadamente las estructuras cristalinas que se utilizarán como moldes, así como un correcto alineamiento sustentado por análisis filogenéticos. Mientras que los modelos agonistas se realizaron a partir de un abordaje múltiple, empleando la estructura experimental de ADRB1 en su conformación activa y la de CB1, el modelo antagonista/ agonista inverso se modeló a partir de la estructura cristalina de CB1 recientemente resuelta. Los modelos fueron refinados y analizados de manera visual para corroborar que cumplieran con los datos experimentales recolectados de trabajos anteriores. Se obtuvo una gran concordancia entre los datos de mutagénesis y la pose obtenida para los ligandos WIN 55,212-2 (agonista selectivo de CB2) y SR-144,528 (agonista inverso selectivo de CB2), pudiendo caracterizar la interacción receptor-ligando y los residuos importantes en el sitio de unión. En cuanto a los modelos de CB2 en su conformación agonista, se observó que el ligando se posiciona entre las hélices 2, 3, 5, 6 y 7 y, en el caso del modelo en su conformación antagonista/ agonista inverso el sitio ortostérico estaría conformado por las hélices 1, 2, 3, 6 y 7. Entre los residuos que formarían parte del bolsillo de unión, se vio que es de importancia la presencia de aminoácidos aromáticos en la parte inferior del sitio de unión y que la modificación por aminoácidos con carga neta distinta de cero lleva a una pérdida en la unión de los ligandos. Parecería que CB2 tiene involucrado en su sitio de unión aminoácidos del N-terminal y el ECL2, como Phe183 el cual está sustentado por datos experimentales. Al igual que en otros GPCRs, se encuentra conservado el puente disulfuro presente en el ECL2 y su ausencia produce cambios conformacionales que llevan a la pérdida de afinidad de los ligandos.

También se llevó a cabo un *docking* automatizado de pequeña escala con el fin de validar los modelos obtenidos y evaluar la calidad de los mismos. A pesar que los resultados obtenidos no fueron los esperados, ya que se esperaba obtener valores de EF más altos, se cree que son lo suficientemente buenos para distinguir ligandos de no ligandos. Los bajos EF obtenidos pueden deberse a la exigencia otorgada por el uso de una GDD en la base de datos utilizada en el proceso de *docking*. En cuanto a los modelos agonistas, los resultados obtenidos demostraron que el modelo A sería mejor que el B y el modelo 2013. El modelo antagonista/ agonista-inverso dio resultados similares al desarrollado en el 2009.

Los modelos realizados en éste trabajo resultan útiles a la hora de guiar la

optimización de ligandos específicos de CB2, para su posterior síntesis química y estudio en sistemas biológicos. Estos aportes contribuyen a poder conocer mejor la estructura y conformaciones adquiridas por CB2 en sus distintos estados.

Referencias bibliográficas

- Abagyan, R. and Totrov, M. (1994) Biased Probability Monte Carlo Conformational Searches and Electrostatic Calculations for Peptides and Proteins. *Journal of Molecular Biology*, **235**, 983–1002.
- Abagyan, R., Totrov, M. and Kuznetsov, D. (1994) ICM - a New Method For Protein Modeling and Design - Applications to Docking and Structure Prediction From the Distorted Native Conformation. *Journal of Computational Chemistry*, **15**, 488–506.
- Alqarni, M., Myint, K.Z., Tong, Q., Yang, P., Bartlow, P., Wang, L., et al. (2014) Examining the critical roles of human CB2 receptor residues Valine 3.32 (113) and Leucine 5.41 (192) in ligand recognition and downstream signaling activities. *Biochemical and Biophysical Research Communications*, **452**, 334–339.
- Andrej Šali. (1993) MODELLER A Program for Protein Structure Modeling. *Comparative protein modelling by satisfaction of spatial restraints.*, 779–815.
- Ashton, J.C., Wright, J.L. and Tyndall, J.M.M. and J.D.A. (2008) Cannabinoid CB1 and CB2 Receptor Ligand Specificity and the Development of CB2-Selective Agonists. *Current Medicinal Chemistry*, **15**, 1428–1443.
- Atwood, B.K., Wager-Miller, J., Haskins, C., Straiker, A. and Mackie, K. (2012) Functional selectivity in CB(2) cannabinoid receptor signaling and regulation: implications for the therapeutic potential of CB(2) ligands. *Molecular Pharmacology*, **81**, 250–263.
- Ballesteros, J.A. and Weinstein, H. (1995) [19] Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. *Methods in Neurosciences*, **25**, 366–428.
- Bermudez-Silva, F.J., Viveros, M.P., McPartland, J.M. and Rodriguez de Fonseca, F. (2010) The endocannabinoid system, eating behavior and energy homeostasis: The end or a new beginning? *Pharmacology Biochemistry and Behavior*, **95**, 375–382.
- Bernstein, F.C., Koetzle, T.F., Williams, G.J., Meyer, E.F., Brice, M.D., Rodgers, J.R., et al. (1977) The Protein Data Bank. *European Journal of Biochemistry*, **80**, 319–324.
- Blad, C.C., Tang, C. and Offermanns, S. (2012) G protein-coupled receptors for energy metabolites as new therapeutic targets. *Nature reviews. Drug discovery*, **11**, 603–19.
- Braun, W. and Go, N. (1985) Calculation of protein conformations by proton-proton distance constraints. A new efficient algorithm. *Journal of Molecular Biology*, **186**,

611–626.

- Cabral, G.A. and Griffin-Thomas, L. (2009) Emerging role of the cannabinoid receptor CB2 in immune regulation: therapeutic prospects for neuroinflammation. *Expert reviews in molecular medicine*, **11**, e3.
- Cavasotto, C.N. (2011) Homology models in docking and high-throughput docking. *Current Topics in Medicinal Chemistry*, **11**, 1528–1534.
- Cavasotto, C.N. and Abagyan, R.A. (2004) Protein flexibility in ligand docking and virtual screening to protein kinases. *Journal of Molecular Biology*, **337**, 209–225.
- Cavasotto, C.N., Kovacs, J.A. and Abagyan, R.A. (2005) Representing Receptor Flexibility in Ligand Docking through Relevant Normal Modes. *Journal of the American Chemical Society*, **127**, 9632–9640.
- Cavasotto, C.N. and Orry, A.J.W. (2007) Ligand docking and structure-based virtual screening in drug discovery. *Current topics in medicinal chemistry*, **7**, 1006–1014.
- Cavasotto, C.N., Orry, A.J.W. and Abagyan, R. (2006) Receptor Flexibility in Ligand Docking. *Handbook of Theoretical and Computational Nanotechnology* (eds M. Rieth), & W. Schommers), pp. 218–257. American Scientific Publishers.
- Cavasotto, C.N., Orry, A.J.W., Murgolo, N.J., Czarniecki, M.F., Kocsi, S.A., Hawes, B.E., et al. (2008) Discovery of novel chemotypes to a G-protein-coupled receptor through ligand-steered homology modeling and structure-based virtual screening. *Journal of Medicinal Chemistry*, **51**, 581–588.
- Cavasotto, C.N. and Palomba, D. (2015) Expanding the horizons of G protein-coupled receptor structure-based ligand discovery and optimization using homology models. *Chemical Communications*, **51**, 13576–13594.
- Cavasotto, C.N. and Phatak, S.S. (2009) Homology modeling in drug discovery: current trends and applications. *Drug Discovery Today*, **14**, 676–683.
- Cherezov, V., Rosenbaum, D.M., Hanson, M.A., Rasmussen, S.G.F., Thian, F.S., Kobilka, T.S., et al. (2007) High-resolution crystal structure of an engineered human b2-adrenergic G protein-coupled receptor. *Science*, **318**, 1258–1265.
- Chrencik, J.E., Roth, C.B., Terakado, M., Kurata, H., Omi, R., Kihara, Y., et al. (2015) Crystal Structure of Antagonist Bound Human Lysophosphatidic Acid Receptor 1. *Cell*, **161**, 1633–1643.
- Christopoulos, A. and Kenakin, T. (2002) G protein-coupled receptor allostereism and complexing. *Pharmacological reviews*, **54**, 323–374.
- Congreve, M. and Marshall, F. (2010) The impact of GPCR structures on pharmacology and structure-based drug design. *British Journal of Pharmacology*, **159**, 986–996.
- Congreve, M., Murray, C.W. and Blundell, T.L. (2005) Structural biology and drug discovery. *Drug discovery today*, **10**, 895–907.

- Díaz-Laviada, I. and Ruiz-Llorente, L. (2005) Signal transduction activated by cannabinoid receptors. *Mini reviews in medicinal chemistry*, **5**, 619–30.
- Díaz, P., Phatak, S.S., Xu, J., Astruc-Díaz, F., Cavasotto, C.N. and Naguib, M. (2009a) 6-Methoxy-N-alkyl Isatin Acylhydrazone Derivatives as a Novel Series of Potent Selective Cannabinoid Receptor 2 Inverse Agonists: Design, Synthesis, and Binding Mode Prediction. *Journal of Medicinal Chemistry*, **52**, 433–444.
- Díaz, P., Phatak, S.S., Xu, J., Fronczek, F.R., Astruc-Díaz, F., Thompson, C.M., et al. (2009b) 2,3-Dihydro-1-benzofuran derivatives as a series of potent selective cannabinoid receptor 2 agonists: Design, synthesis, and binding mode prediction through ligand-steered modeling. *ChemMedChem*, **4**, 1615–1629.
- Edgar, R.C. and Batzoglou, S. (2006) Multiple sequence alignment. *Current Opinion in Structural Biology*, **16**, 368–373.
- Edvardsen, O., Reiersen, A.L., Beukers, M.W. and Kristiansen, K. (2002) tGRAP, the G-protein coupled receptors mutant database. *Nucleic Acids Res*, **30**, 361–363.
- Feng, Z., Alqarni, M.H., Yang, P., Tong, Q., Chowdhury, A., Wang, L., et al. (2014) Modeling, molecular dynamics simulation, and mutation validation for structure of cannabinoid receptor 2 based on known crystal structures of GPCRs. *Journal of Chemical Information and Modeling*, **54**, 2483–2499.
- Fredriksson, R., Lagerström, M.C., Lundin, L.-G. and Schiöth, H.B. (2003) The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Molecular pharmacology*, **63**, 1256–72.
- Friesner, R.A., Banks, J.L., Murphy, R.B., Halgren, T.A., Klicic, J.J., Mainz, D.T., et al. (2004) Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *Journal of Medicinal Chemistry*, **47**, 1739–1749.
- Gatica, E.A. and Cavasotto, C.N. (2012) Ligand and Decoy Sets for Docking to G Protein-Coupled Receptors. *Journal of Chemical Information and Modeling*, **52**, 1–6.
- Gonzalez, A., Cordoní, A., Caltabiano, G. and Pardo, L. (2012) Impact of Helix Irregularities on Sequence Alignment and Homology Modeling of G Protein-Coupled Receptors. *ChemBioChem*, **13**, 1393–1399.
- Gouldson, P., Calandra, B., Legoux, P., Kernéis, A., Rinaldi-Carmona, M., Barth, F., et al. (2000) Mutational analysis and molecular modelling of the antagonist SR 144528 binding site on the human cannabinoid CB2 receptor. *European Journal of Pharmacology*, **401**, 17–25.
- Guerram, M., Zhang, L.-Y. and Jiang, Z.-Z. (2016) G-protein coupled receptors as

- therapeutic targets for neurodegenerative and cerebrovascular diseases. *Neurochemistry International*, **101**, 1–14.
- Guindon, J. and Hohmann, A.G. (2009) The endocannabinoid system and pain. *CNS & neurological disorders drug targets*, **8**, 403–21.
- Guindon, J. and Hohmann, A.G. (2011) The endocannabinoid system and cancer: Therapeutic implication. *British Journal of Pharmacology*, **163**, 1447–1463.
- Halgren, T.A., Murphy, R.B., Friesner, R.A., Beard, H.S., Frye, L.L., Pollard, W.T., et al. (2004) Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *Journal of Medicinal Chemistry*, **47**, 1750–1759.
- Hanson, M. a., Roth, C.B., Jo, E., Griffith, M.T., Scott, F.L., Reinhart, G., et al. (2012) Crystal structure of a lipid G protein-coupled receptor. *Science*, **335**, 851–5.
- Herkenham, M., Lynn, A.B., Little, M.D., Johnson, M.R., Melvin, L.S., de Costa, B.R., et al. (1990) Cannabinoid receptor localization in brain. *Proceedings of the National Academy of Sciences of the United States of America*, **87**, 1932–6.
- Hua, T., Vemuri, K., Pu, M., Qu, L., Han, G.W., Wu, Y., et al. (2016) Crystal Structure of the Human Cannabinoid Receptor CB1. *Cell*, **167**, 750–762.e14.
- Huang, N., Shoichet, B.K. and Irwin, J.J. (2006) Benchmarking sets for molecular docking. *Journal of Medicinal Chemistry*, **49**, 6789–6801.
- ICM. (2012) ICM.
- Irwin, J.J. (2008) Community benchmarks for virtual screening. *Journal of Computer-Aided Molecular Design*, **22**, 193–199.
- Isberg, V., Mordalski, S., Munk, C., Rataj, K., Harps??e, K., Hauser, A.S., et al. (2016) GPCRdb: An information system for G protein-coupled receptors. *Nucleic Acids Research*, **44**, D356–D364.
- Jorgensen, W.L. (2004) The many roles of computation in drug discovery. *Science (New York, N.Y.)*, **303**, 1813–8.
- Katritch, V., Cherezov, V. and Stevens, R.C. (2012) Diversity and modularity of G protein-coupled receptor structures. *Trends in Pharmacological Sciences*, **33**, 17–27.
- Lagerström, M. and Schiöth, H. (2008) Structural diversity of G protein-coupled receptors and significance for drug discovery. *Nat Rev Drug Discov*, **7**, 339–57.
- Lania, A.G., Mantovani, G. and Spada, A. (2006) Mechanisms of disease: Mutations of G proteins and G-protein-coupled receptors in endocrine diseases. *Nature clinical practice. Endocrinology & metabolism*, **2**, 681–693.
- Lappano, R. and Maggiolini, M. (2011) G protein-coupled receptors: novel targets for drug discovery in cancer. *Nature reviews. Drug discovery*, **10**, 47–60.

- Larsson, P., Wallner, B., Lindahl, E. and Elofsson, A. (2008) Using multiple templates to improve quality of homology models in automated homology modeling. *Protein Science*, **17**, 990–1002.
- Leite, C.E., Mocelin, C.A., Petersen, G.O., Leal, M.B. and Thiesen, F. V. (2009) Rimonabant: An antagonist drug of the endocannabinoid system for the treatment of obesity. *Pharmacological Reports*, **61**, 217–224.
- Li, S., Huang, S. and Peng, S.B. (2005) Overexpression of G protein-coupled receptors in cancer cells: involvement in tumor progression. *Int J Oncol*, **27**, 1329–1339.
- Martí-Renom, M. a, Stuart, a C., Fiser, a, Sánchez, R., Melo, F. and Sali, a. (2000) Comparative protein structure modeling of genes and genomes. *Annual review of biophysics and biomolecular structure*, **29**, 291–325.
- Di Marzo, V., Goparaju, S.K., Wang, L., Liu, J., Bátkai, S., Járαι, Z., et al. (2001) Leptin-regulated endocannabinoids are involved in maintaining food intake. *Nature*, **410**, 822–825.
- Matsuda, L. a, Lolait, S.J., Brownstein, M.J., Young, a C. and Bonner, T.I. (1990) Structure of a cannabinoid receptor and functional expression of the cloned cDNA. *Nature*, **346**, 561–564.
- McAllister, S.D., Tao, Q., Barnett-Norris, J., Buehner, K., Hurst, D.P., Guarnieri, F., et al. (2002) A critical role for a tyrosine residue in the cannabinoid receptors for ligand recognition. *Biochemical Pharmacology*, **63**, 2121–2136.
- McCusker, E.C., Bane, S.E., O'Malley, M.A. and Robinson, A.S. (2007) Heterologous GPCR expression: A bottleneck to obtaining crystal structures. *Biotechnology Progress*, **23**, 540–547.
- Mercier, R.W., Pei, Y., Pandarinathan, L., Janero, D.R., Zhang, J. and Makriyannis, A. (2010) HCB2 Ligand-interaction landscape: Cysteine residues critical to biarylpyrazole antagonist binding motif and receptor modulation. *Chemistry and Biology*, **17**, 1132–1142.
- Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. and Teller, E. (1953) Equation of state calculations by fast computing machines. *Journal Chemical Physics*, **21**, 1087–1092.
- Micale, V., Di Marzo, V., Sulcova, A., Wotjak, C.T. and Drago, F. (2013) Endocannabinoid system and mood disorders: Priming a target for new therapies. *Pharmacology and Therapeutics*, **138**, 18–37.
- Mirzadegan, T., Benko, G., Filipek, S. and Palczewski, K. (2003) Sequence analyses of G-protein-coupled receptors: Similarities to rhodopsin. *Biochemistry*, **42**, 2759–2767.
- Munk, C., Isberg, V., Mordalski, S., Harps??e, K., Rataj, K., Hauser, A.S., et al. (2016)

- GPCRdb: the G protein-coupled receptor database ??? an introduction. *British Journal of Pharmacology*, 2195–2207.
- Munro, S., Thomas, K.L. and Abu-Shaar, M. (1993) Molecular characterization of a peripheral receptor for cannabinoids. *Nature*, **365**, 61–65.
- Naguib, M., Diaz, P., Xu, J.J., Astruc-Diaz, F., Craig, S., Vivas-Mejia, P., et al. (2008) MDA7: a novel selective agonist for CB2 receptors that prevents allodynia in rat neuropathic pain models. *British Journal of Pharmacology*, **155**, 1104–1116.
- Nebane, N.M., Hurst, D.P., Carrasquer, C.A., Qiao, Z., Reggio, P.H. and Song, Z.H. (2008) Residues accessible in the binding-site crevice of transmembrane helix 6 of the cb2 cannabinoid receptor. *Biochemistry*, **47**, 13811–13821.
- Oldham, W.M. and Hamm, H.E. (2008) Heterotrimeric G protein activation by G-protein-coupled receptors. *Nature reviews. Molecular cell biology*, **9**, 60–71.
- Orry, A.J.W. and Cavasotto, C.N. (2006) Ligand-docking-based homology model of the Melanin-Concentrating Hormone 1 receptor. *231st Meeting of the American Chemical Society* p. Atlanta, GA.
- Palczewski, K., Kumasaka, T., Hori, T., Behnke, C.A., Motoshima, H., Fox, B.A., et al. (2000) Crystal structure of rhodopsin: A G protein-coupled receptor. *Science*, **289**, 739–745.
- Palomba, D. and Cavasotto, C.N. (2015) Protein Structure Modeling in Drug Design. *In Silico Drug Discovery and Design: Theory, Methods, Challenges, and Applications* (ed C.N. Cavasotto), pp. 215–248. CRC Press, Taylor & Francis Group.
- Pei, Y., Mercier, R.W., Anday, J.K., Thakur, G.A., Zvonok, A.M., Hurst, D., et al. (2008) Ligand-Binding Architecture of Human CB2 Cannabinoid Receptor: Evidence for Receptor Subtype-Specific Binding Motif and Modeling GPCR Activation. *Chemistry and Biology*, **15**, 1207–1219.
- Perutz, M.F., Fermi, G. and Abraham, D.J. (1986) Hemoglobin as a receptor of drugs and peptides: X-ray studies of the stereochemistry of binding. *Journal of the ...*, **108**, 1064–1078.
- De Petrocellis, L., Melck, D., Palmisano, A., Bisogno, T., Laezza, C., Bifulco, M., et al. (1998) The endogenous cannabinoid anandamide inhibits human breast cancer cell proliferation. *Proc Natl Acad Sci U S A*, **95**, 8375–8380.
- Petrov, R.R., Knight, L., Chen, S.R., Wager-Miller, J., McDaniel, S.W., Diaz, F., et al. (2013) Mastering tricyclic ring systems for desirable functional cannabinoid activity. *European Journal of Medicinal Chemistry*, **69**, 881–907.
- Phatak, S.S., Gatica, E.A. and Cavasotto, C.N. (2010) Ligand-steered modeling and docking: A benchmarking study in class A G-protein-coupled receptors. *Journal of Chemical Information and Modeling*, **50**, 2119–2128.

- Phatak, S.S., Stephan, C.C. and Cavasotto, C.N. (2009) High-throughput and in silico screenings in drug discovery. *Exp. Opin. Drug Discov.*, **4**, 947–959.
- Pierce, K.L., Premont, R.T. and Lefkowitz, R.J. (2002) Seven-transmembrane receptors. *Nature reviews. Molecular cell biology*, **3**, 639–650.
- Piomelli, D. (2003) The molecular logic of endocannabinoid signalling. *Nature reviews. Neuroscience*, **4**, 873–884.
- Qamri, Z., Preet, A., Nasser, M.W., Bass, C.E., Leone, G., Barsky, S.H., et al. (2009) Synthetic cannabinoid receptor agonists inhibit tumor growth and metastasis of breast cancer. *Molecular cancer therapeutics*, **8**, 3117–29.
- Rhee, M.H. (2002) Functional role of serine residues of transmembrane dopamin VII in signal transduction of CB2 cannabinoid receptor. *Journal of veterinary science (Suwon-si, Korea)*, **3**, 185–191.
- Rosenbaum, D.M., Rasmussen, S.G.F. and Kobilka, B.K. (2009) The structure and function of G-protein-coupled receptors. *Nature*, **459**, 356–63.
- Salon, J. a, Lodowski, D.T. and Palczewski, K. (2011) The Significance of G Protein-Coupled Receptor Crystallography for Drug Discovery. *Pharmacological Reviews*, **63**, 901–937.
- Schöneberg, T., Schulz, A., Biebermann, H., Hermsdorf, T., Römpler, H. and Sangkuhl, K. (2004) Mutant G-protein-coupled receptors as a cause of human diseases. *Pharmacology & Therapeutics*, **104**, 173–206.
- Seifert, R., Wenzel-Seifert, K., Gether, U. and Kobilka, B.K. (2001) Functional differences between full and partial agonists: evidence for ligand-specific receptor conformations. *The Journal of pharmacology and experimental therapeutics*, **297**, 1218–1226.
- Shao, Z., Yin, J., Chapman, K., Grzemska, M., Clark, L., Wang, J., et al. (2016) High-resolution crystal structure of the human CB1 cannabinoid receptor. *Nature*, 1–20.
- Shoichet, B.K. (2004) Virtual screening of chemical libraries. *Nature*, **432**, 862–865.
- Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., et al. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular systems biology*, **7**, 539.
- Sippl, M.J. (1993) Recognition of errors in three-dimensional structures of proteins. *Proteins: Structure, Function, and Bioinformatics*, **17**, 355–362.
- Song, Z.H., Slowey, C. a, Hurst, D.P. and Reggio, P.H. (1999) The difference between the CB(1) and CB(2) cannabinoid receptors at position 5.46 is crucial for the selectivity of WIN55212-2 for CB(2). *Molecular pharmacology*, **56**, 834–840.
- Tao, Q., McAllister, S.D., Andreassi, J., Nowell, K.W., Cabral, G.A., Hurst, D.P., et al. (1999) Role of a conserved lysine residue in the peripheral cannabinoid receptor

- (CB2): evidence for subtype specificity. *Molecular Pharmacology*, **55**, 605–613.
- Tuccinardi, T., Ferrarini, P.L., Manera, C., Ortore, G., Saccomanni, G. and Martinelli, A. (2006) Cannabinoid CB2/CB1 selectivity. Receptor modeling and automated docking analysis. *Journal of Medicinal Chemistry*, **49**, 984–994.
- Valant, C., Sexton, P.M. and Christopoulos, A. (2009) Orthosteric/allosteric bitopic ligands: going hybrid at GPCRs. *Molecular interventions*, **9**, 125–135.
- Venkatakrishnan, A.J., Deupi, X., Lebon, G., Tate, C.G., Schertler, G.F. and Babu, M.M. (2013) Molecular signatures of G-protein-coupled receptors. *Nature*, **494**, 185–194.
- Warne, T., Moukhametzianov, R., Baker, J.G., Nehmé, R., Edwards, P.C., Leslie, A.G.W., et al. (2011) The structural basis for agonist and partial agonist action on a β 1-adrenergic receptor. *Nature*, **469**, 241–244.
- Wiederstein, M. and Sippl, M.J. (2007) ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Research*, **35**.
- Xie, X.Q., Chen, J.Z. and Billings, E.M. (2003) 3D structural model of the G-protein-coupled cannabinoid CB2 receptor. *Proteins: Structure, Function and Genetics*, **53**, 307–319.
- Xu, J.J., Diaz, P., Astruc-Diaz, F., Craig, S., Munoz, E. and Naguib, M. (2010) Pharmacological characterization of a novel cannabinoid ligand, MDA19, for treatment of neuropathic pain. *Anesthesia and Analgesia*, **111**, 99–109.
- Xu, F., Wu, H., Katritch, V., Han, G.W., Jacobson, K.A., Gao, Z.-G., et al. (2011) Structure of an Agonist-Bound Human A2A Adenosine Receptor. *Science*, **332**, 322–327.
- Zhang, Y., Xie, Z., Wang, L., Schreiter, B., Lazo, J.S., Gertsch, J., et al. (2011) Mutagenesis and computer modeling studies of a GPCR conserved residue W5.43(194) in ligand recognition and signal transduction for CB2 receptor. *International Immunopharmacology*, **11**, 1303–1310.