



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Maestría y Doctorado en Ciencias Bioquímicas

Estudio computacional del receptor CX3CR1 como posible blanco terapéutico para el tratamiento de la enfermedad de Alzheimer

TESIS

QUE PARA OPTAR POR EL GRADO DE:
Maestro en Ciencias

PRESENTA:

Guillermo David Goode Romero

Laura Domínguez Dueñas
Facultad de Química, UNAM

Marcelino Arciniega Castro
Instituto de Fisiología Celular, UNAM
Rogelio Rodríguez Sotres
Facultad de Química, UNAM

Ciudad de México. Julio de 2019

MIEMBROS DEL JURADO

PRESIDENTE	COSTAS BASÍN MIGUEL ANTONIO
VOCAL	GONZÁLEZ ANDRADE MARTÍN
VOCAL	GARCÍA HERNÁNDEZ ENRIQUE
VOCAL	PASTOR COLÓN CARMEN NINA
SECRETARIO	SOSA PEINADO ALEJANDRO

TUTOR

DOMÍNGUEZ DUEÑAS LAURA

SUSTENTANTE

GOODE ROMERO GUILLERMO DAVID

AGRADECIMIENTOS

AGRADEZCO A LA UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO POR LA FORMACIÓN MULTIDISCIPLINARIA RECIBIDA.

A LA FACULTAD DE QUÍMICA Y A MI TUTORA DRA. LAURA DOMÍNGUEZ POR LA DISPOSICIÓN PLENA, LAS FACILIDADES Y EL APOYO PARA REALIZAR ESTE PROYECTO.

A LOS INTEGRANTES DEL CUBÍCULO F230; Q.F.B. ALEJANDRO ÁLVAREZ HERNÁNDEZ, VÍCTOR H. CALVO JIMÉNEZ, Q. DULCE C. GUZMÁN OCAMPO, HUMBERTO T. HERNÁNDEZ TRUJILLO, Y DE MANERA MUY ESPECIAL AL DR. RODRIGO AGUAYO ORTIZ POR TODOS LOS COMENTARIOS Y SUGERENCIAS PARA LA MEJORA CONTINUA DEL PROYECTO Y DEL DESEMPEÑO EN EL GRUPO DE TRABAJO.

AL DR. MARCELINO ARCINIEGA CASTRO Y DR. ROGELIO RODRÍGUEZ SOTRES POR LAS OBSERVACIONES Y SUGERENCIAS.

A MI MADRE G. SUSANA Y A MI PADRE GUILLERMO POR EL APOYO Y LA CONVIVENCIA DESDE EL INICIO.

A MIS AMIGOS M. RUBÉN SALINAS ESPINOZA, A ISC. EDUARDO E. RODRÍGUEZ GARCÍA Y A MAT. FERNANDO R. RODRÍGUEZ CRUZ POR SU VALIOSA AMISTAD Y COMPAÑÍA.

AL Q.A. HÉCTOR E. RODRÍGUEZ SOLLANO POR LOS RETOS PERSONALES QUE, SIN SU PRESENCIA, NO HUBIERA LLEGADO A LOGRAR.

A LA DIRECCIÓN GENERAL DE CÓMPUTO Y TECNOLOGÍAS DE LA INFORMACIÓN POR EL ACCESO A LA SUPERCOMPUTADORA *HP CLUSTER PLATFORM 3000SL* (MIZTLI).

Y AL CONSEJO NACIONAL DE CIENCIA Y TECNOLOGÍA (CONACYT) POR LA BECA OTORGADA (NO. 631327/482735).

DURANTE LA MAESTRÍA SE REALIZÓ LA PUBLICACIÓN “RELACIONES CUANTITATIVAS ESTRUCTURA ACTIVIDAD/PROPIEDAD EN DOS DIMENSIONES EMPLEANDO EL PROGRAMA R” EN LA REVISTA EDUCACIÓN QUÍMICA, EN COAUTORÍA CON RODRIGO AGUAYO-ORTIZ Y LAURA DOMÍNGUEZ.



“Relaciones cuantitativas estructura-actividad/propiedad en dos dimensiones empleando el programa R”,
Guillermo Goode-Romero, Rodrigo Aguayo-Ortiz y Laura Domínguez
Volumen 30 | Número 2 | Páginas 27-40 | Abril 2019
DOI: [10.22201/eq.18708404e.2019.3.67211](https://doi.org/10.22201/eq.18708404e.2019.3.67211)

RELACIONES CUANTITATIVAS ESTRUCTURA- ACTIVIDAD/PROPIEDAD EN DOS DIMENSIONES EMPLEANDO EL PROGRAMA R

Resumen

Las relaciones cuantitativas estructura-actividad (*QSAR*) y estructura-propiedad (*QSPR*) son modelos matemáticos aplicados a la predicción de actividades biológicas o propiedades de un grupo de compuestos. Estos modelos son generados y validados por análisis estadístico a partir de un grupo de moléculas con una actividad biológica o propiedad conocida. En este trabajo se explica la metodología general para realizar un análisis *QSAR/QSPR* utilizando el lenguaje de programación de *R*, analizando como caso de estudio la predicción del transporte a través de la barrera hematoencefálica.

Palabras clave: QSAR/QSPR, RStudio, Barrera hematoencefálica

TWO-DIMENSIONAL QUANTITATIVE STRUCTURE- ACTIVITY/PROPERTY RELATIONSHIPS USING R SOFTWARE

Abstract

Quantitative structure-activity/property relationships (*QSAR/QSPR*) are mathematical models applied to the prediction of biological activities or properties of a series of compounds. These models are generated and validated by statistical analysis from a group of molecules with a known biological activity or property. This paper describes the general methodology to perform a *QSAR/QSPR* study using the *R* software, employing experimental information of the transport through the blood-brain barrier as a case of study.

Keywords: QSAR/QSPR, RStudio, Blood-brain barrier

Autores: Guillermo Goode-Romero^a, Rodrigo Aguayo-Ortiz^a, and Laura Domínguez^{*a}

^a Facultad de
Química,
Departamento
de Físicoquímica,
Universidad
Nacional
Autónoma de
México, México.
^{*}Autor para
correspondencia:
lauradd@unam.mx

ÍNDICE

ABREVIATURAS, SÍMBOLOS Y UNIDADES	7
I.1 RESUMEN	13
I.2. INTRODUCCIÓN	14
II. ANTECEDENTES	15
II.1. GENERALIDADES	15
II.2. NEUROINFLAMACIÓN	15
II.3. CÉLULAS GLIALES	16
II.3.1. ASTROCITOS	16
II.3.2. NG-2 Y OLIGODENDROCITOS	16
II.3.3. MICROGLÍA	17
II.4. ORIGEN HEMATOPOYÉTICO DE LA MICROGLÍA	17
II.4.1. ACTIVIDAD DE LA MICROGLÍA	17
II.5. CITOCINAS Y QUIMIOCINAS	18
II.5.1. QUIMIOCINAS	18
II.5.2. FRACTALKINA	18
II.5.2.1. EL RECEPTOR DE FRACTALKINA: CX3CR1	19
II.6. RECEPTORES ACOPLADOS A PROTEÍNAS G (GPCR)	20
II.6.1. CLASES DE GPCRS	20
II.7. PROTEÍNAS G	22
II.7.1. PROTEINAS Gi/O	23
II.8. ACTIVACIÓN DE LAS PROTEÍNAS G HETEROTRIMÉRICAS	23
II.8.1. MODELO DE GILMAN DE ACTIVACIÓN DE LAS PROTEÍNAS G HETEROTRIMÉRICAS	23
II.8.2. MODELO DE BOUVIER DE ACTIVACIÓN DE LAS PROTEÍNAS G HETEROTRIMÉRICAS	24

III. JUSTIFICACIÓN, HIPÓTESIS Y OBJETIVOS	25
IV. METODOLOGÍA	26
IV.1. SISTEMAS DE ESTUDIO	26
IV.2. SISTEMAS EN MEMBRANAS DE POPC	27
V. RESULTADOS Y DISCUSIÓN	29
V.1 SISTEMAS DEL RECEPTOR DE REFERENCIA CCR5	29
V.2. SISTEMAS DEL RECEPTOR DE PRUEBA CX3CR1	36
V.2.1. MODELOS CX3CR1 MULTITEMPLADO, ACTIVO, TRANSITIVO Y ESTATIVO	38
V.2.2. MODELO MULTITEMPLADO CON ANTAGONISTAS PRESUNTIVOS	43
V.2.3. MODELO MULTITEMPLADO CON AGONISTAS PRESUNTIVOS	47
V.3. ANÁLISIS DE LOS ESTADOS FUNCIONALES DE CX3CR1 Y LOS EFECTOS DE LOS LIGANDOS	50
V.3.1. PATRONES DE ESTADO FUNCIONAL	50
V.4. MODELOS DE RECEPTOR-PROTEÍNA Gi	53
V.4.1. SUBUNIDAD CX3CR1 DEL COMPLEJO	53
V.4.2. HETEROTRÍMERO Gi	55
VI. CONCLUSIONES Y PERSPECTIVAS	57
VII. REFERENCIAS	58
VIII. ANEXO I: RESULTADOS SUPLEMENTARIOS	63
IX. ANEXO II: INFORMACIÓN COMPLEMENTARIA	77
X. ANEXO III: FUNDAMENTOS TEÓRICOS	82

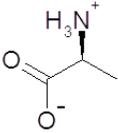
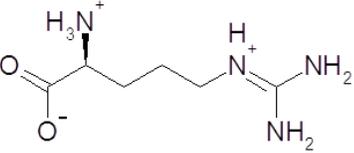
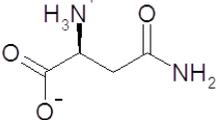
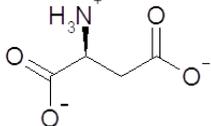
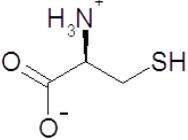
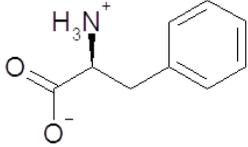
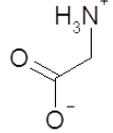
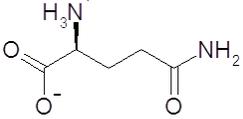
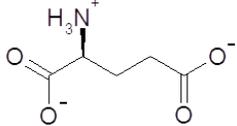
ABREVIATURAS, SÍMBOLOS Y UNIDADES

Letras griegas minúsculas utilizadas

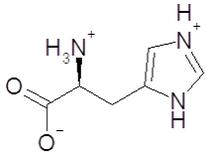
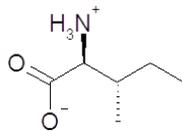
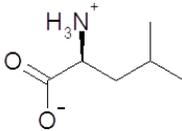
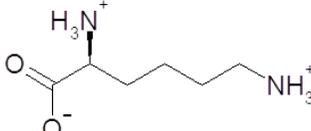
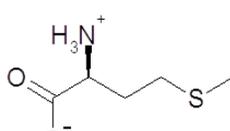
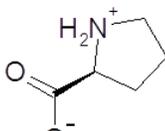
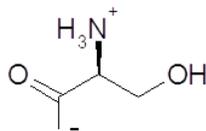
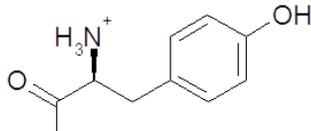
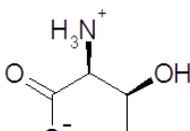
α	Alfa
β	Beta
γ	Gamma
ε	Épsilon
θ	Theta
μ	Mu
ξ	Xi
π	Pi
ρ	Ro
σ	Sigma
τ	Tau
ϕ	Fi
χ	Ji
ψ	Psi

ABREVIATURAS, SÍMBOLOS Y UNIDADES

Aminoácidos

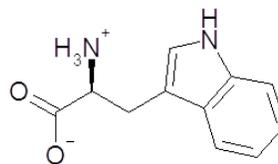
Abreviaturas	Símbolos	Nombre	Estructura predominante a pH=7
Ala	A	Alanina	
Arg	R	Arginina	
Asn	N	Asparagina	
Asp	D	Ácido aspártico	
Cys	C	Cisteína	
Phe	F	Fenilalanina	
Gly	G	Glicina	
Gln	Q	Glutamina	
Glu	E	Ácido glutámico	

ABREVIATURAS, SÍMBOLOS Y UNIDADES

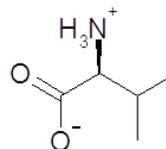
His	H	Histidina	
Ile	I	Isoleucina	
Leu	L	Leucina	
Lys	K	Lisina	
Met	M	Metionina	
Pro	P	Prolina	
Ser	S	Serina	
Tyr	Y	Tirosina	
Thr	T	Treonina	

ABREVIATURAS, SÍMBOLOS Y UNIDADES

Trp W Triptófano



Val V Valina



Abreviatura o acrónimo

Significado

AFKN

Aglicofractalkina, aglicona de CX3CL1 soluble

BHE

Barrera hematoencefálica

CCR5

Receptor 5 de quimiocina con motivo CC; CD195; correceptor de fusión de HIV-1

CHARMM

Campo de fuerza CHARMM (del inglés *Chemistry at Harvard Macromolecular Mechanics*)

CX3CL1

Ligando 1 con motivo CX3C; fractalkina, neurotactina

CX3CR1

Receptor 1 de quimiocina con motivo CX3C; V28; GPR13

EM

Minimización de energía potencial (del inglés *Energy Minimization*)

FF

Campo de fuerza (del inglés *Force Field*)

FKN

Fractalkina, CX3CL1

GDP

Guanosina-5'-difosfato

GPCR

Receptor acoplado a proteína G (del inglés *G protein-coupled receptor*)

GTP

Guanosina-5'-trifosfato

ABREVIATURAS, SÍMBOLOS Y UNIDADES

MD	Dinámica molecular (del inglés <i>Molecular Dynamics</i>)
MRV	Maraviroc; ligando antagonista del receptor CCR5
NPT	Colectivo o ensamble isomolar-isobárico-isotérmico
NTT	Neurotactina, fractalkina, CX3CL1
NVT	Colectivo o ensamble isomolar-isocórico-isotérmico (canónico)
PDB	Base de datos de estructuras experimentales (<i>Protein data bank</i>)
POPC	(2S)-3-Palmitoil-2-Oleílfosfatidilcolina, un fosfolípido membranal del tipo lecitina
R*	Receptor en estado activo
R	Receptor en estado basal o pasivo
R \emptyset	Receptor en estado estativo
R‡	Receptor en estado transitivo, confórmero intermedio
Ras	Proteína G monomérica; subfamilia de GTPasas pequeñas (acrónimo de <i>Rat sarcoma</i>)
QMEAN	Análisis cualitativo de energía del modelo (del inglés <i>Qualitative Model Energy Analysis</i>)
QSAR	Relaciones cuantitativas estructura-actividad (del inglés <i>Quantitative Structure-Activity Relationships</i>)
QSPR	Relaciones cuantitativas estructura-propiedad (del inglés <i>Quantitative Structure-Property Relationships</i>)
UniProt	Base de datos de proteínas (<i>Universal Protein Resource</i>)

Magnitudes y unidades

Å	Ångström (1 Å = 0.1 nm = 10 ⁻¹⁰ m)
d	Distancia
fs	Femtosegundo
n _t	Número de pasos de tiempo
nm	Nanómetro
ns	Nanosegundo
RMSD	Raíz cuadrada de la desviación cuadrática media; raíz del segundo momento ponderado por masas atómicas, centrado en la condición inicial en este estudio
RMSF	Raíz cuadrada de la fluctuación cuadrática media; raíz del segundo momento bivariado ponderado por masas atómicas, centrado en la condición inicial en este estudio
rg	Radio de giro
T	Temperatura
t	Tiempo
φ	Ángulo diedro formado por CA-C-N-CA
χ _i	Ángulos diedros de cadenas laterales
ψ	Ángulo diedro formado por N-CA-C-N

I.1. RESUMEN

La enfermedad de Alzheimer (EA) es una patología compleja con características neurodegenerativas y neuroinflamatorias. Se ha demostrado que los mecanismos involucrados en la respuesta inmunológica en el sistema nervioso central (SNC) tienen una relevancia importante en el inicio y el transcurso de la EA. Las células microgliales residentes en el SNC desempeñan un papel fundamental en la neuroinflamación. El receptor de fractalkina, CX3CR1, es un elemento clave en la activación de una respuesta proinflamatoria o de supervivencia neuronal ante las anomalías propias de la EA. Estudiar el receptor CX3CR1 con técnicas computacionales permite complementar y entender sus características estructurales y dinámicas; para comprender mejor los mecanismos implicados en la EA, así como un punto de partida para el planteamiento de nuevos enfoques de estudio que conduzcan a una eventual propuesta de moléculas con actividad moduladora sobre este receptor, y de esta manera modificar el transcurso de la característica neuroinflamatoria de la EA. Para este trabajo de tesis se recopiló la información disponible para realizar un estudio computacional consistente en modelado y validación de la estructura terciaria, simulación de dinámica molecular clásica del receptor en sistemas membranales diversos, y análisis geométrico de confórmers y rotámeros; con las características necesarias para proponer un mecanismo básico de funcionamiento del receptor CX3CR1, así como sus interacciones con los ligandos conocidos por su actividad experimental. Asimismo, se aplicaron herramientas computacionales para validar y confirmar los resultados expuestos, y concluir satisfactoriamente sobre las características funcionales del receptor CX3CR1, así como proponer perspectivas en el estudio de este sistema y lograr el conocimiento molecular de la EA con las técnicas disponibles.

I.2. INTRODUCCIÓN

Las herramientas computacionales han mostrado ser una buena aproximación para el estudio de los sistemas macromoleculares con detalle atómico. La simulación de dinámica molecular (*MD* por sus siglas en inglés) es un grupo de metodologías que aplican las leyes de la mecánica molecular y la termodinámica estadística, mediante métodos inicialmente desarrollados de manera teórica y algoritmos que simplifican los cálculos computacionales. La simulación de dinámica molecular resuelve las ecuaciones de movimiento de la mecánica clásica a través de pasos de tiempo, así como los cálculos respectivos de fuerzas y energía potencial. La resolución de estas ecuaciones permite conocer las posiciones de cada partícula del sistema para el paso de tiempo siguiente, por lo que para la asignación de velocidades se recurre a una función de densidad de probabilidad de velocidades, para posteriormente calcular las aceleraciones, tanto del paso en proceso como del paso siguiente, a través de diversos algoritmos sencillos. Asimismo, el sistema molecular simulado se considera cerrado, con la propiedad de intercambiar energía con el exterior en un baño virtual de temperatura, para alcanzar un estado de equilibrio térmico. La presión dentro del sistema se calcula mediante modificaciones a las ecuaciones clásicas, y adicionalmente estos parámetros termodinámicos están sujetos a la evaluación de algoritmos que disminuyen la desviación de los valores alcanzados (para mayor detalle, ver el Anexo III: Fundamentos teóricos). De esta manera, se puede conocer el comportamiento aproximado del sistema molecular a través del tiempo simulado, y se pueden conocer detalles aproximados de desplazamientos de componentes del sistema, interacciones entre partículas y cambios conformacionales de segmentos mayores.

Para este trabajo, el receptor CX3CR1 es el principal objeto de estudio mediante técnicas computacionales. Sin embargo, para tener una referencia comparativa sólida, se estudió inicialmente el receptor CCR5, el receptor con mayor similitud a CX3CR1 y con información disponible a nivel atómico. A partir del estudio de CCR5, también se caracterizó la influencia de ligandos, para integrar los resultados y proponer mecanismos y características para el receptor principal CX3CR1. Adicionalmente, para el diseño de los sistemas con ligandos peptídicos se recurrió a la estructura del receptor US28 en complejo con el ligando principal de CX3CR1, para plantear un punto de partida basado en evidencia experimental (las generalidades de los receptores CCR5 y US28 se exponen en el Anexo II: Información complementaria).

II. ANTECEDENTES

II.1. Generalidades

La EA es la causa principal de demencia en personas mayores a nivel mundial.^{1,2} Es una enfermedad neurodegenerativa caracterizada clínicamente por pérdida progresiva de memoria y de las funciones cognitivas. Los depósitos de los péptidos amiloides β 1-42 ($A\beta$ 42) en el parénquima encefálico y en los vasos sanguíneos, así como la hiperfosforilación de la proteína asociada a microtúbulos tau (MAPT), inestabilidad microtubular y neuroinflamación son las características fisiopatológicas distintivas de la enfermedad.^{1,3,4} Los péptidos neurotóxicos $A\beta$ 40 y $A\beta$ 42 son producto de la escisión de la proteína transmembranal APP (proteína precursora de amiloide, por sus siglas en inglés) por la acción del complejo enzimático de la γ -secretasa, con presenilina (PS, PSEN) como subunidad catalítica. Estos péptidos amiloides se acumulan formando oligómeros solubles abundantes en hojas β , y posteriormente se agregan en placas insolubles en el espacio extracelular. La angiopatía cerebral amiloidea, presente en el 80% de los casos de EA, se debe a la deposición de oligómeros amiloides. Los oligómeros solubles $A\beta$ son causantes de neuroinflamación que está estrechamente relacionada con la EA.²

II.2. Neuroinflamación

La inflamación es una respuesta en un tejido cuando se presenta un daño, y puede ser aguda o crónica. La inflamación aguda se caracteriza por una reacción inmediata ante un agente agresor, y consiste básicamente en una respuesta defensiva y posteriormente en el restablecimiento de la homeostasis local. La inflamación crónica, en cambio, se presenta cuando el estímulo nocivo es persistente. La inflamación periférica se manifiesta principalmente por infiltrados leucocitarios y producción de agentes de señalización y regulación. Las células polimorfonucleares, de las cuales los neutrófilos son las más abundantes, son reclutadas al sitio de daño en eventos agudos y las células mononucleares como macrófagos, linfocitos y células plasmáticas componen la respuesta inflamatoria celular crónica.

Las respuestas desencadenadas por la neuroinflamación están reconocidas como factores determinantes en el desarrollo y evolución de las enfermedades neurodegenerativas, incluyendo la EA. De manera análoga a la periferia, en el sistema nervioso central (SNC) la inflamación puede ser aguda o crónica. La inflamación aguda, conocida previamente como gliosis reactiva, se presenta en eventos traumáticos, isquémicos o infecciosos. La inflamación crónica o neuroinflamación, en cambio, tiene características divergentes con su contraparte periférica.⁵ Las células gliales presentes en el sistema nervioso desarrollan las funciones de presentación antigénica, modulación de la inflamación, reclutamiento de células periféricas, inducción de apoptosis, etc.

II.3. Células gliales

Los astrocitos, oligodendrocitos, glías NG-2 y microglías, conocidas en conjunto como células gliales (Figura 2.1), constituyen una fracción importante del SNC de mamíferos, además de las neuronas. Consideradas inicialmente como células de soporte, de “pegamento”, no funcionales, son objeto de extensa investigación.

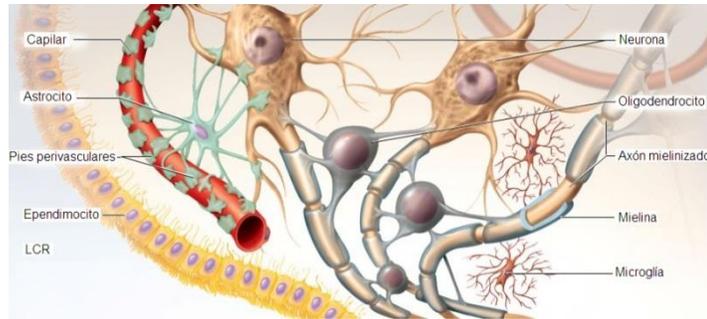


Figura 2.1. Células del SNC. Interacciones de las neuronas con astrocitos, oligodendrocitos y microglías. Los ependimocitos son células epiteliales de los sistemas ventriculares (Modificado de *McGraw-Hill Companies, Inc.*).

II.3.1. Astrocitos

Los astrocitos representan la fracción más abundante de las células gliales encefálicas en un individuo adulto. Entre sus numerosas funciones descritas, destacan el mantenimiento homeostático de iones, agua, nutrientes y desechos, participación cosináptica y contribución en la estructura de la barrera hematoencefálica (BHE).⁶ La ablación astrocítica selectiva en ratas sanas usando α -L-aminoadipato, análogo del L-glutamato, ocasionó principalmente reactividad glial reversible.⁷ En cambio, la supresión de la expresión de proteínas exclusivas astrocíticas resultó en degeneración neuronal, degeneración axonal y parálisis⁸⁻¹¹ y en un subtipo de astrocitos cerebelares causó problemas severos de coordinación motora, similares a los producidos por disfunción cerebelar.

II.3.2. NG-2 y oligodendrocitos

Los oligodendrocitos y las células NG-2, sus precursoras, se consideran una población glial independiente debido a sus características, como son los marcadores receptor α del factor de crecimiento derivado de plaquetas (PDGFR α) y el proteoglicano 4 de condroitina sulfato (CSPG4 o NG2) para los precursores, y los factores de transcripción Olig1, Olig2, Olig3 y SOX10, así como la proteína específica de oligodendocito (OSP), proteína básica de mielina (MBP) y la glicoproteína oligodendrocítica de mielina (MOG) para las células maduras.^{6,12} Las células NG-2 se han identificado como parte fundamental del mantenimiento de los oligodendrocitos mielinizantes a través de la vida adulta.¹³⁻¹⁶ Las células NG-2 presentan alta proliferación, de manera análoga a las células troncales, y son capaces de participar en sinapsis neuronales.^{13,17} La depleción de células NG-2 con citarabina, irradiación X y silenciamiento genético produjo disminución de la señalización de leptina y deterioro del metabolismo energético.¹⁸ El bloqueo de la diferenciación de NG-2 a oligodendrocitos genera deterioro del aprendizaje motor y fallas axonales.^{19,20}

II.3.3. Microglía

Las células microgliales o microglías son las células efectoras inmunológicas de origen mesodérmico en el encéfalo, con los marcadores proteína transmembranal 119 (TMEM119), CD11b-CD45, CX3CR1, entre otros.²¹ Las microglías se infiltran en el tejido neural durante la embriogénesis, y durante el período posnatal desempeñan un papel fundamental en la maduración neuronal. Posterior a la infiltración, se diferencian de una forma ameboide a una morfología ramificada única, estacionaria, que cambia cuando son expuestas a estímulos nocivos. Esta morfología ramificada no es un estado dormante, ya que estas células presentan períodos intermitentes de motilidad y pausa. En respuesta a estímulos asociados a daño y patógenos (*DAMPs*: patrones moleculares asociados a daño; *PAMPs*: patrones moleculares asociados a patógenos), las microglías experimentan una rápida activación dirigiendo sus movimientos de respuesta al microambiente hacia la migración al sitio del insulto.²²

Las microglías activadas liberan al entorno numerosos factores neurotróficos y citocinas capaces de modificar la homeostasis neuronal. Asimismo, expresan una amplia variedad de receptores y moléculas de señalización como respuestas a la actividad neuronal, como el receptor de glutamato AMPAR, purinérgicos P2RX4 y P2Y6, receptor del factor de necrosis tumoral alfa (TNF- α) TNFR1/CD120a, receptor del mediador inflamatorio interleucina 1 (IL-1R), complejo principal de histocompatibilidad tipo II (MHC-II), y receptores de quimiocinas. Adicionalmente son capaces de producir factor de crecimiento nervioso NGF, factor neurotrófico derivado de cerebro (BDNF) y neurotrofina 3.²²

II.4. Origen hematopoyético de la microglía

Las microglías son un tipo de células fagocíticas, macrófagos residentes del SNC que provienen de células progenitoras mieloides residentes en la médula ósea, y constituyen un linaje diferente al de las células fagocíticas mononucleares. Los monocitos, otro linaje mieloides, son células leucocitarias circulantes que forman parte del sistema fagocítico mononuclear.

II.4.1. Actividad de la microglía

En un individuo adulto sano, la autorrenovación microglial es suficiente para mantener la homeostasis del SNC. En condiciones fisiológicas, las microglías contribuyen a la plasticidad neuronal y la función sináptica, controlando su maduración y conectividad. Durante el desarrollo, las microglías remueven las células y cuerpos apoptóticos que resultan en esta etapa, y también contribuyen a la muerte neuronal. La apoptosis neuronal también está presente en adultos. En el caso de los pacientes con EA, las células troncales neurales se encuentran disminuidas,²² lo que deteriora la autorrenovación microglial.

II.5. Citocinas y quimiocinas

Las citocinas son mediadores celulares de naturaleza polipeptídica que regulan la homeostasis de los tejidos a través de acciones locales o por reclutamiento de sistemas externos. La expresión y actividad de las citocinas se incrementa en condiciones de estrés tisular, incluyendo crecimiento rápido, desregulación debida a inflamación crónica y neoplasias, infecciones y eventos traumáticos. En este grupo heterogéneo de mediadores están incluidas interleucinas, factores de necrosis tumoral, interferones, factores estimulantes de colonias, factores de crecimiento, neurotrofinas, neuropoyetinas y quimiocinas. Las citocinas en el SNC pueden ser de producción basal o en condiciones de daño, como la producción de IL-1 β , IL-2 e IL-6 en la EA.²³

II.5.1. Quimiocinas

Las quimiocinas son un tipo de citocinas implicadas en la quimiotaxia y activación leucocitarias. Las quimiocinas han sido clasificadas con base en la posición relativa de sus residuos de cisteína (Cys, C) en el extremo amino terminal. Las α - y β -quimiocinas, con cuatro residuos de cisteína, constituyen la familia más extensa. Las α -quimiocinas contienen un motivo CXC, donde X es un residuo aminoácido variable, y las β -quimiocinas presentan el motivo CC, de residuos adyacentes. La linfotactina y la fractalkina, que no se incluyen en los grupos anteriores, contienen sólo dos residuos de C y otras características distintivas.

II.5.2. Fractalkina

La fractalkina es una δ -quimiocina²⁴ membranal que presenta el motivo CXXXC, o CX3C; específicamente la secuencia CNITC, unido a un dominio tipo mucina altamente glicosilado, anclado en la membrana plasmática. La fractalkina recibió este nombre debido a la analogía de su descubrimiento con un acercamiento a la rama C de la superfamilia de quimiocinas con un fractal, al revelar un patrón autorreferencial de complejidad.²⁵ A esta quimiocina también se denomina CX3CL1, por ser el ligando con el motivo CX3C. La unión a un dominio tipo mucina sólo es compartida con CXCL16. La fractalkina es expresada en la periferia por células NK (*natural killer cells*, de linaje hematopoyético) y por grupos restringidos de neuronas.²⁶ La escisión del dominio extracelular de la fractalkina por la metaloproteinasa ADAM10 (del inglés *a disintegrin and metalloproteinase type 10*), un tipo de enzima α -secretasa, produce polipéptidos solubles de tamaño variable, y bajo condiciones inflamatorias, el corte también es promovido por ADAM17/TACE (Figura 2.2).

La fractalkina posee un dominio citoplasmático muy corto, y parece no interactuar con correceptores. En ratas seniles, la isoforma soluble de fractalkina se encuentra disminuida en la corteza y el hipocampo debido a la pérdida neuronal existente. La forma soluble tiene relación con el grado de tauopatía, o alteración patológica del metabolismo de MAPT; mientras que la forma membranal parece incrementar la patología amiloidea.²⁷

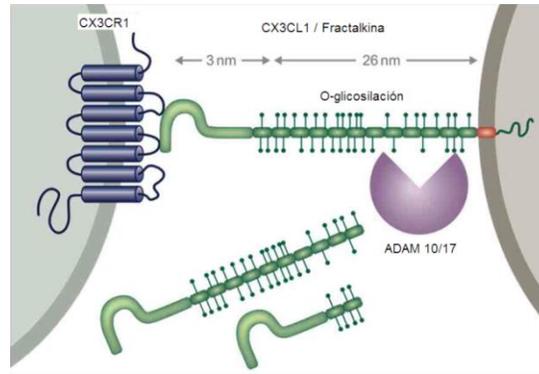


Figura 2.2. CX3CR1 y CX3CL1 en el SNC. Esquemización general de la fractalkina, su interacción con CX3CR1 y sus variantes solubles generadas por las enzimas ADAM. (Modificado de Wolf, *et al.* 2017).²⁶

II.5.2.1. El receptor de fractalkina: CX3CR1

El receptor de fractalkina es un receptor acoplado a proteína G (GPCR por sus siglas en inglés), y se denomina CX3CR1. El CX3CR1 se expresa en linfocitos, monocitos, células NK y células gliales, de expresión particularmente alta en microglías.²⁶

En modelos animales de EA se ha estudiado la influencia de CX3CR1, incluyendo los modelos transgénicos para hTau (Tau silvestre humana) y APP/PSEN1 mutantes. En el primer caso, los ratones transgénicos hTau, carentes del gen CX3CR1 (CX3CR1^{-/-}) presentan hiperfosforilación de tau (p-MAPT) asociada a una alta producción microglial de IL-1 β . En modelos mixtos hTau-APP/PSEN1-CX3CR1^{-/-}, la deficiencia del receptor se manifestó con una reducción de las placas amiloides. En microglías cultivadas, la exposición a A β 42 induce la depleción de la actividad de CX3CR1, producción de IL-6 y TNF- α . Además, la activación microglial incrementa la fagocitosis de las placas amiloides, aunque con un riesgo potencial de secreción de citocinas proinflamatorias relacionadas con un incremento de la tauopatía.

Adicionalmente la deficiencia de CX3CR1 en los modelos transgénicos mixtos previene la pérdida neuronal.²⁸ En los modelos APP/PSEN1 deficientes del ligando CX3CL1 membranal se presenta mayor p-MAPT a pesar de la disminución de la densidad amiloidea. La sobreexpresión de la isoforma soluble (sCX3CL1) en el modelo transgénico Tg4510 mediante vectores adenovirales reduce la tauopatía y previene la neurodegeneración.

Los roedores a los que se les infundió A β 42 en la corteza mediante un lentivirus, así como modelos de ratón con demencia frontotemporal exhibieron una alta activación de la caspasa 3 y una secreción incrementada de la citocina proinflamatoria TNF- α y depósitos de p-MAPT.

Además de las implicaciones de CX3CR1 en la EA, algunas mutaciones en este receptor se han asociado con progresiones distintivas en la EA de inicio tardío. Con estos hallazgos experimentales se relaciona al receptor CX3CR1, un tipo de receptor acoplado a proteína G (GPCR), con la EA.

II.6. Receptores acoplados a proteínas G (GPCR)

Los receptores membranales heptahelicoidales acoplados a proteínas de unión a nucleótidos de guanina (GPCRs) constituyen un grupo amplio de proteínas celulares transmembranales involucradas en la transducción de señales de primeros mensajeros extracelulares y son de gran interés farmacéutico.^{29,30}

La geometría de los GPCRs (Figura 2.3) está conservada en la mayoría de los receptores descritos, especialmente en las siete hélices transmembranales (TMs) y la hélice yuxtamembranal (YXT). Las variaciones principales tienen lugar en las asas extra e intracelulares (ECLs e ICLs respectivamente), así como en los dominios amino (NT) y carboxilo terminal (CT).³¹

Por su diversidad, los GPCRs se clasifican en clases, basadas en la estructura primaria.^{32,33}

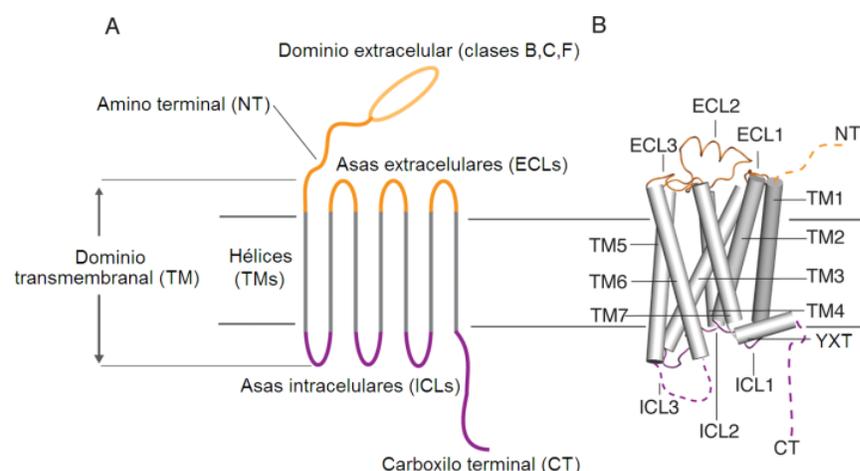


Figura 2.3. Geometría general de un GPCR. A. Dominios geométricos principales. B. Identificadores de los dominios principales. En línea discontinua se representan los segmentos interhelicoidales (Modificado de Latorraca, 2017³¹).

II.6.1. Clases de GPCRs

Los GPCRs descritos pertenecen a las clases A, B1, B2, C, F, gustativos y otros (Tabla 2.1).³² Dentro de cada clase, los receptores poseen similitud estructural alta.

Tabla 2.1. Clases de GPCR basadas en alineamientos de sus estructuras primarias.

Clase	Receptor típico	Subtipos principales
A	Rodopsina	Receptores de proteínas, algunos polipéptidos, sensoriales, esteroideos, aminérgicos, alicarboxílicos, lipídicos, nucleotídicos.
B1	Receptor de secretina	Receptores de calcitonina, factor liberador de corticotropina, glucagon, parathormona y péptido intestinal vasoactivo.
B2	Receptores de adhesión	ADGRs
C	Receptores de glutamato	Receptores sensibles a calcio, aminoacídicos, gustativos tipo 1.
F	Receptores <i>frizzled</i>	FZDs
Gustativos tipo 2	Receptores TAS2	TAS2Rs (de gusto amargo)

El receptor de fractalkina CX3CR1 (Figura 2.4) está clasificado como receptor tipo A debido a su similitud estructural con la rodopsina.³²

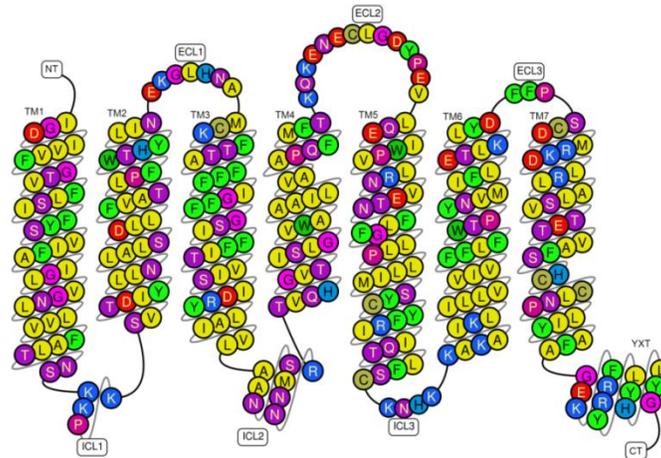


Figura 2.4. Estructura primaria del receptor de fractalkina CX3CR1, con la distribución de residuos de aminoácidos en los motivos transmembranales (TM1 a TM7), la hélice yuxtamembranal (YXT) y los terminales amino (NT) y carboxilo (CT).³²

La respuesta de los GPCRs ante los estímulos se basa principalmente en los cambios de estado funcional que pueden experimentar. Se conocen al menos seis estados funcionales de los GPCRs, cuatro estados de tendencia a no señalar y dos de tendencia a señalar (Figura 2.5). Un estado funcional inactivo corresponde con un receptor incapaz de señalar con los sistemas transductores intracelulares, como las proteínas G. Un estado estativo constituye una funcionalidad quiescente, generalmente sin actividad y con una restricción al cambio de estado. Un estado basal o pasivo describe los receptores que en ausencia de ligandos pueden adoptar otros estados funcionales de manera espontánea. Cuando un GPCR es capaz de adoptar estados activos y señalar en ausencia de ligandos activadores o agonistas, se dice que es un receptor con actividad constitutiva o intrínseca. A este tipo de receptores se les puede anular su función constitutiva mediante la interacción con agentes agonistas inversos, y dirigirlos a un estado inactivo. Cuando un GPCR no posee actividad intrínseca, donde el cambio de estado basal a activo no está favorecido, el estado estativo y el estado inactivo pueden coincidir. En cuanto a los estados capaces de señalar, se encuentran los estados transitivos que incluyen todos aquellos cambios que experimenta el receptor en una activación incipiente, es un estado intermedio o metaestado. Es en el estado transitivo donde se han encontrado las interacciones determinantes con los transductores, principalmente las proteínas G. Un receptor que se ha activado, unido al transductor correspondiente, presenta un estado funcional activo, que se propaga y se convierte en el respectivo estado activo de las proteínas transductoras. Finalmente, cuando ha tenido lugar la señalización intracelular, el receptor sufre una modificación que anula su actividad, mediante interacciones con nuevas proteínas o modificaciones covalentes, para entrar a un estado de desactivación permanente.

La geometría de los GPCRs permite la formación de una cavidad interhelicoidal que atraviesa la membrana de un lado al otro, y se ha encontrado el paso de moléculas de agua y el ingreso de un ion sodio.³⁴ Durante el proceso de activación funcional, se describió la interrupción del paso de moléculas de agua a través del poro interhelicoidal.³⁵ El cambio de estado también involucra una tétrada de residuos de aminoácidos interactuantes, conocida como compuerta iónica, la cual puede presentarse en un estado de oclusión del poro, o cerrado, cuando uno o dos residuos básicos interactúan con uno o dos ácidos; o un estado de apertura del poro, por ruptura de las interacciones iónicas.³⁶

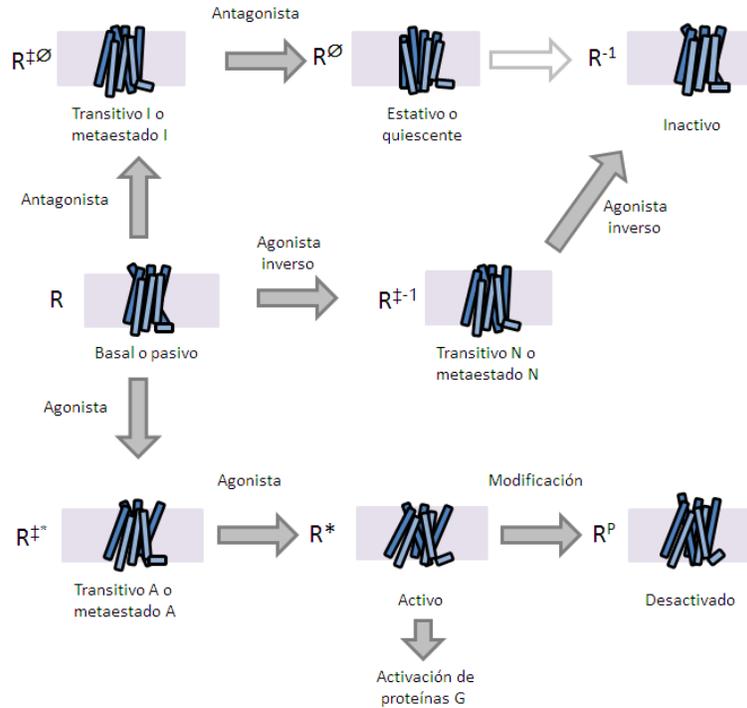


Figura 2.5. Esquematación de los estados funcionales de los GPCRs, con la nomenclatura utilizada.

II.7. Proteínas G

Las proteínas de unión a nucleótidos de guanina constituyen una familia de GTPasas monoméricas o pequeñas y heterotriméricas. Las proteínas G monoméricas están implicadas de manera importante en la señalización de estímulos extracelulares dirigidos principalmente al citoesqueleto, al tráfico de vesículas y al crecimiento y diferenciación celular. Las proteínas heterotriméricas están compuestas por subunidades denominadas α , β y γ , que son las proteínas transductoras de los complejos membranales GPCR-proteína G.³⁷

Las subunidades $G\alpha$ poseen N- y/o S-acilaciones en el extremo amino y dos dominios principales de actividad, el dominio alfa-helicoidal (α AH) y el dominio tipo *Ras* (α Ras, por la proteína G monomérica *Ras*, descrita por primera vez en un sarcoma de rata).^{38,39} Los dominios α AH y α Ras están unidos por dos asas conectoras, y forman una cavidad de alojamiento para el

nucleótido de guanina. Cuando el nucleótido alojado es Mg[GTP], la proteína se encuentra en un estado activo, y cuando éste se hidroliza por la tríada catalítica formada por los residuos Ser47, R178 y D200, inducida por la acción de activadores de proteínas GTPasas (*GTPase-activating proteins*, GAPs); la subunidad se inactiva y se detiene el proceso de señalización.

Las subunidades $G\beta$ tienen una estructura secundaria conocida como propela beta, un motivo poligonal regular de láminas beta.

Las subunidades γ presentan una prenilación en el dominio carboxilo terminal que la mantiene adyacente a la membrana, e interactúa directamente con la subunidad $G\beta$.

Una característica distintiva entre las subunidades $G\alpha$ de diferentes clases es la sensibilidad a toxinas bacterianas con actividad de ADP-ribosiltransferasa directa (toxina Pertussis, PTX de *Bordetella pertussis*) o indirecta (toxina del cólera (CTX), y toxinas termolábiles (LTs) de *Escherichia coli* y *Bacillus cereus*).³⁷

El receptor CX3CR1 transduce en gran medida mediante proteínas G_i .

II.7.1. Proteínas $G_{i/o}$

Las proteínas de clase $G_{i/o}$, al ser activadas, bloquean la producción del segundo mensajero AMP cíclico (cAMP) mediante la interacción de la subunidad $G\alpha$ -Mg[GTP] con la enzima transmembranal adenilato ciclasa. El heterodímero $\beta\gamma$ a su vez interactúa con varias enzimas membranales como proteínas cinasas C (PKCs, Figura 2.6).

II.8. Activación de las proteínas G heterotriméricas

La activación de las proteínas G mediante el GPCR activo asociado se debe a un cambio conformacional propagado desde el receptor, desencadenado por el ligando, hasta el segmento ICL2, que adquiere conformación helicoidal,³⁸ y propagado a las subunidades de la proteína G. Este cambio conformacional final favorece el recambio de GDP unido a la subunidad α , por Mg[GTP], que induce la interacción de la proteína G con sus respectivas proteínas diana. La capacidad de activación de las proteínas G se conoce como GEF (*guanine nucleotide exchange factor*), por lo que al receptor membranal activo asociado se le considera un GEF.

II.8.1. Modelo de Gilman de activación de las proteínas G heterotriméricas

El modelo de Gilman propone que la activación de la proteína G induce la disociación del complejo $G\alpha\beta\gamma$ - en la subunidad $G\alpha$ -Mg[GTP] y el heterodímero $G\beta\gamma$ ⁴⁰ (Figura 2.7).

II.8.2. Modelo de Bouvier de activación de las proteínas G heterotriméricas

El modelo del grupo de Bouvier propone que la activación de la proteína G favorece un cambio conformacional no disociativo⁴¹ seguido de la apertura de la cavidad del nucleótido, por distanciamiento de los dominios α AH y α Ras, permitiendo el intercambio de los nucleótidos de guanina (Figura 2.8).

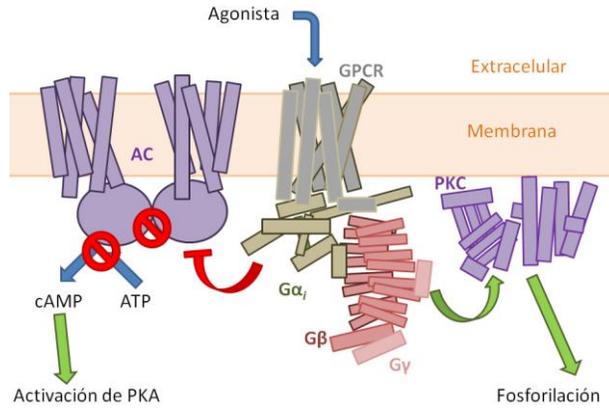


Figura 2.6. Transducción de la señal de un agonista a través de un GPCR acoplado a una proteína Gi. Al activarse el GPCR (color gris) y posteriormente la proteína Gi, la subunidad G α_i activa (color café) inhibe la actividad de la enzima adenilato ciclasa (AC), que cataliza la transformación de ATP a AMP cíclico (cAMP). Este efecto disminuye la concentración de cAMP en el citoplasma. Por otra parte, las subunidades G $\beta\gamma$ activadas (color rosa) pueden inducir la actividad de la proteína cinasa C (PKC). Esta activación permite la fosforilación por PKC de múltiples sustratos.

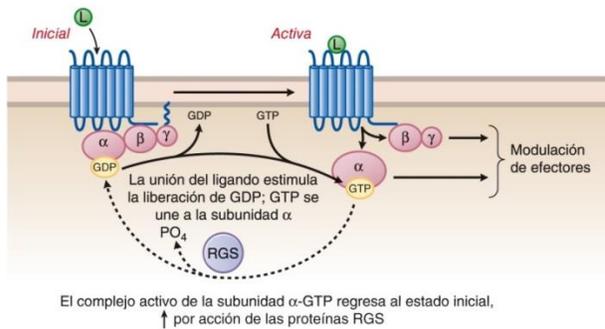


Figura 2.7. Modelo de Gilman de activación de las proteínas G.^{40,42} El estado inicial también se denomina basal o pasivo.

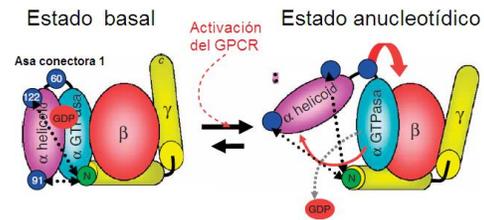


Figura 2.8. Modelo del grupo de Bouvier⁴¹ de activación de las proteínas G.

Estos dos modelos experimentales de activación proponen características claramente diferenciadas en las subunidades de las proteínas G. Una parte de este trabajo se centró en el estudio del complejo CX3CR1-Gi, y se consideraron los mecanismos de los dos modelos de activación de proteínas G para identificar algunas señales o indicadores de estos comportamientos en los estudios computacionales que involucran esta proteína.

III.1. JUSTIFICACIÓN

El receptor CX3CR1 es miembro de la superfamilia más grande de proteínas blanco de numerosos fármacos, y su influencia definitoria en el desarrollo y progresión de la neuroinflamación ha sido sustentada por una cantidad creciente de evidencia experimental *ex vivo* e *in vivo*. Por otra parte, las metodologías computacionales de dinámica molecular han demostrado ser una aproximación sólida y confiable para estudiar el comportamiento de los sistemas biológicos a escala atómica. Por estos motivos, el estudio dinámico y conformacional de este receptor con las capacidades disponibles de cómputo lo convierte en un sistema atractivo para contribuir al conocimiento molecular de la enfermedad de Alzheimer.

III.2. HIPÓTESIS

El receptor CX3CR1 como miembro de la clase A de GPCRs, comparte patrones generales con los integrantes de esta clase, referentes a cambios conformacionales y funcionales que permitirán caracterizarlo estructural y dinámicamente. Su estudio con agentes con actividades agonista y antagonista, así como con la proteína Gi, facilitará la dilucidación de los aspectos conformacionales más relevantes.

III.3. OBJETIVOS

1. Conocer la estabilidad dinámica del receptor CX3CR1 en un modelo de sistema membranal.
2. Caracterizar los tres estados funcionales descritos para este receptor.
3. Identificar sus interacciones con ligandos bioactivos.
4. Proponer moléculas pequeñas como posibles ligandos del receptor.

IV. METODOLOGÍA

El procedimiento general que se siguió para el estudio del receptor CX3CR1 se esquematiza en la Figura 4.1.

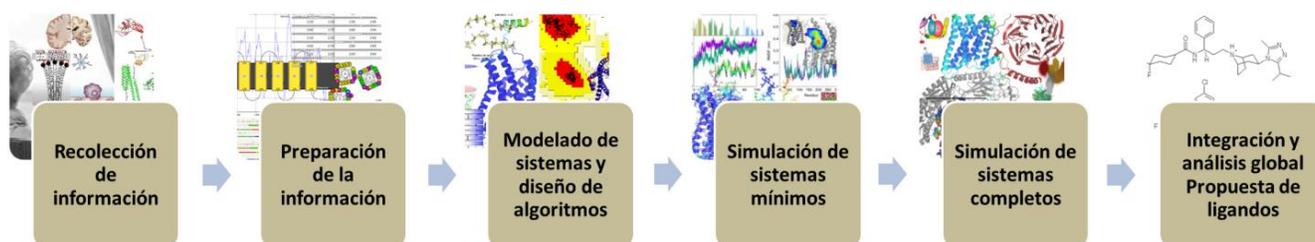


Figura 4.1. Procedimiento general para el estudio computacional de CX3CR1.

IV.1. Sistemas de estudio

Los sistemas de receptores modelados y simulados se muestran en la Tabla 4.1. El receptor CCR5 se utilizó como referencia del receptor CX3CR1 debido a su alta similitud estructural, y por ser el miembro con mayor información a nivel molecular (ver Anexo II). Se modelaron tres sistemas del receptor CCR5 a partir de su estructura depositada en el *Protein Data Bank*,⁴³ once sistemas del receptor de prueba CX3CR1 en diferentes condiciones, y finalmente cuatro sistemas de CX3CR1 en complejo con el heterotrímero Gi (ver Anexo II).

Tabla 4.1. Sistemas de receptores simulados y sus funciones principales.

Sistema	Detalles
CCR5-MRV	El receptor CCR5 se cristalizó en su estado estativo (PDB: 4MBS), que fungió como referencia para el análisis del ligando maraviroc (MRV) como antagonista (ver Anexo II).
CCR5-mrv	El receptor CCR5 con la variante desprotonada del antagonista (mrv) permitió caracterizar los cambios en la dinámica ante esta modificación.
CCR5	El aporreceptor CCR5 permitió conocer los cambios relativos al receptor estativo.
CX3CR1	Receptor de prueba, de interés principal por su función en las microglías.
CX3CR1*	Modelo activo del receptor de prueba CX3CR1 (<i>GPCR database</i>).
CX3CR1‡	Modelo transitivo o intermedio del receptor de prueba CX3CR1 (<i>GPCR database</i>).
CX3CR1∅	Modelo estativo del receptor de prueba CX3CR1 (<i>GPCR database</i>).
CX3CR1-MRV	Receptor de prueba con ligando maraviroc, posicionado como en la estructura de referencia (PDB: 4MBS).
CX3CR1-E6130	Receptor de prueba con el antagonista experimental E6130, en correspondencia de grupos funcionales con la estructura de referencia (PDB: 4MBS).
CX3CR1-AZD8797	Receptor de prueba con el antagonista AZD8797, posicionado según sus interacciones.
CX3CR1-LG1	Receptor de prueba con el ligando propuesto LG1.
CX3CR1-AFKN	Receptor de prueba con la variante aglicona de fractalkina (AFKN; PDB: 4XT1).
CX3CR1-FKN ^{G2}	Receptor de prueba con la variante glicosilada G2 de fractalkina (ver Anexo II).
CX3CR1-FKN ^{G2S2}	Receptor de prueba con la variante glicosilada G2S2 de fractalkina, de la cual se conoce su actividad agonista (ver Anexo II).
G-CX3CR1	Receptor de prueba, asociado a proteína Gi (PDB: 3SN6 y 5KDO; ver Anexo II).
G-CX3CR1-MRV	Receptor de prueba, asociado a proteína Gi y unido a maraviroc.
G-CX3CR1-AFKN	Receptor de prueba, asociado a proteína Gi y unido a la aglicona de fractalkina (AFKN).

IV.2. Sistemas en membranas de POPC (1O-palmitoil-2O-oleílfosfatidilcolina)

Los pasos de la metodología (ver Anexos II y III) para cada sistema de receptor se describen en la Tabla 4.2.

Tabla 4.2. Metodología general aplicada a los sistemas de receptor.

Proceso	Descripción
Orientación	Rototraslación del receptor a través del eje z.
Modelados particulares	Modelados, alineamientos de FKN, maraviroc y proteína Gi a partir de estructuras experimentales.
Modelado de membrana	Construcción de los modelos membranales con POPC (ver el Anexo II).
Solvatación	Solvatación y adición de iones sodio y cloruro hasta concentración 0.15 M
Minimización de energía potencial (EM)	Por optimización geométrica con el algoritmo <i>Steepest Descent</i> (ver Anexo III).
Equilibrio NVT	Dos equilibrios con restricciones de posición decrecientes.
Equilibrio NPT	Tres equilibrios con restricciones de posición decrecientes.
Producción de dinámica molecular (MD)	Con tiempo de simulación de 400 ns.
Análisis de desplazamientos y conformaciones	Cálculo de RMSD, RMSF, radio de giro, distancias mínimas, estructura secundaria, correlación cruzada e información mutua torsionales.
Análisis de componentes principales	Para los movimientos principales de los receptores, y en su caso, para la subunidad G α .
Análisis de densidad de moléculas de agua	Caracterización de la cantidad de moléculas de agua que interactúan con los receptores en posiciones específicas.

Se orientaron los modelos de GPCR con el servidor *Orientations of Proteins in Membranes (OPM) database*,⁴⁴ para el posicionamiento del modelo membranal, y posteriormente se utilizó el servidor *CHARMM-GUI*^{45,46} para preparar cada sistema con el modelo de membrana,^{47,48} la solvatación con el modelo de agua TIP3 y la adición aleatoria de iones Na⁺ y Cl⁻ en el bulto acuoso hasta 0.15 M, una concentración equivalente a la que se encuentran las membranas plasmáticas biológicas.

Los ligandos no peptídicos maraviroc, AZD8797, E6130, y el propuesto LG1, se parametrizaron con el campo de fuerza CHARMM36 en el servidor CHARMM-GUI, y los diecisiete sistemas (Tabla 4.1) se simularon 400 ns con el programa GROMACS 5.0.7.⁴⁹

Los glicanos en los ligandos peptídicos se modelaron en el servidor CHARMM-GUI, a partir de la información de la base de datos *Glycan Fragment DB (GFDB)*.⁵⁰

A partir de las trayectorias de 400 ns se calcularon valores de la raíz de la desviación cuadrática media (RMSD) respecto a la estructura inicial para la cadena principal de las hélices transmembranales (TMs), la raíz de la fluctuación cuadrática media (RMSF) respecto a la estructura inicial, para las cadenas laterales de los aminoácidos, radio de giro para toda la región transmembranal, estructura secundaria, distancias mínimas entre

residuos interactuantes, particularmente la tétada de residuos conocida como compuerta iónica y el sitio de unión ortostérico, y análisis de componentes principales, con el fin de identificar los modos principales y fluctuaciones esenciales de los receptores. Estos cálculos se realizaron empleando GROMACS 5.0.7. Los ángulos de inclinación de las hélices transmembranales respecto al vector ortogonal a la membrana, los ángulos diedros ϕ , ψ y χ s de los residuos de aminoácidos para cálculos estadísticos, y la cantidad de moléculas de agua al interior del receptor, con *MDAnalysis*.^{51,52} La correlación cruzada torsional con la paquetería *Bio3d* de R^{53–55} y la información mutua torsional a partir de los valores de diedros, para la detección de movimientos concertados a distancia, con el código de *MutInf*.⁵⁶ Adicionalmente se analizaron las cavidades formadas en el receptor y en la membrana circundante con *trj_cavity*,⁵⁷ en particular la cavidad extracelular de unión a ligando, tomando como base las interacciones del receptor de referencia CCR5 con el ligando maraviroc en la estructura de PDB con clave 4MBS.⁵⁸

V. RESULTADOS Y DISCUSIÓN

En las simulaciones MD de los sistemas de CCR5 y CX3CR1 se estudiaron las estructuras secundarias, desplazamientos de segmentos o grupos de residuos, análisis de componentes principales, análisis de distancias y moléculas de agua, y propagación de información mutua torsional.

V.1 Sistemas del receptor de referencia CCR5.

Los sistemas de referencia constituidos por el receptor CCR5 se simularon en tres condiciones a partir de la estructura experimental en estado estativo con el antagonista maraviroc (Figura 5.1). Este estado conformacional es característico cuando el ligando no induce activación ni inactivación, sino un estado de bloqueo ante otros ligandos.

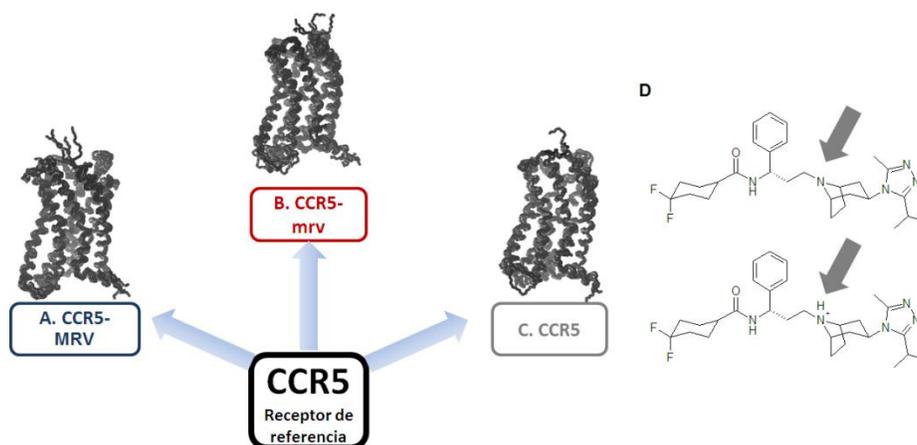


Figura 5.1. Sistemas de referencia con el receptor CCR5. Las conformaciones provienen del análisis de agrupamiento de distribución exponencial, con $\Delta\text{RMSD}=0.15$ Å. (A) Sistema CCR5 \emptyset con ligando maraviroc protonado (MRV), del cual se simularon dos réplicas. (B) Sistema CCR5 \emptyset con ligando maraviroc desprotonado (mrv) con una réplica. La relevancia del estado de protonación impacta de manera fundamental en las interacciones con la proteína, así como en los cálculos sobre el subsistema que representa el ligando. (C) Sistema CCR5 \emptyset sin ligando o aporreceptor, con dos réplicas. (D) Estructuras del antagonista maraviroc en los dos estados de protonación más probables, mrv y MRV.

En el sistema CCR5 \emptyset -MRV, el análisis de estructura secundaria mostró que el receptor adoptó una conformación poco ordenada en el segmento estructural ICL2, en oposición a la conformación α -helicoidal que adquirió el receptor en el sistema sin ligando, misma que se presentó de manera muy favorecida en el sistema CCR5 \emptyset -mrv. Esta conformación helicoidal en ICL2 se espera en receptores que se encuentran en un estado activo (unido a agonista) o basal (con actividad intrínseca no nula), pero no estativo.³⁸ Otro hallazgo fue el plegamiento β de ECL2, el asa extracelular más extensa, en el sistema CCR5 \emptyset -MRV, debida principalmente a la proximidad de este segmento con el ligando. En cambio, en el aporreceptor no se observó este plegamiento, siendo más estable una estructura tipo hélice 3-10 (Figura 5.2, A-C).

El ligando experimentó desplazamientos distintivos en los sistemas CCR5[∅]-MRV y CCR5[∅]-mrv, en la proximidad de la tétrada de residuos Y37^{1.39}, W86^{2.60}, Y251^{6.51} y E283^{7.39} correspondiente al sitio de unión ortostérico (ver anexo II para la nomenclatura de Ballesteros-Weinstein). La presencia del grupo amino terciario protonable en MRV favorece la interacción de éste con la cadena lateral de E283^{7.39}, como está reportado en la estructura experimental.⁵⁸ La especie desprotonada del ligando promueve el distanciamiento del carboxilo del residuo E283^{7.39} (Figura 5.2, D-F). Esta tétrada mostró desplazamientos mayores en el receptor unido a mrv, lo que sugiere la tendencia a un desequilibrio dinámico.

Los valores de la raíz cuadrada de las fluctuaciones cuadráticas media, centradas en la estructura inicial (RMSF) de las cadenas laterales de los receptores, comparados con el aporreceptor, tienen diferencias notables en determinadas regiones, que es evidente con los valores de Δ RMSF (Figura 5.2, G); como es evidente en el segmento ICL2 de CCR5[∅]-MRV. Esta amplia fluctuación corresponde con la estructura secundaria desordenada del asa, típica de una conformación estativa. En contraste, los valores de Δ RMSF son considerablemente menores en el segmento ICL2 del aporreceptor debido a su plegamiento helicoidal en CCR5[∅], más favorecido en ausencia del antagonista. El maraviroc protonado presentó escasa movilidad a través del tiempo de simulación, como se observa en el gráfico de la raíz de la media cuadrática, o RMS centrado en el origen; RMS⁰ (Figura 5.2, H), que está determinada principalmente por los grupos rotables; de manera característica los grupos fenilo e isopropilo.

Al analizar los movimientos colectivos de los sistemas, el receptor en el sistema CCR5[∅]-MRV mostró restricciones en su movilidad global y el aporreceptor CCR5[∅] tuvo mayor número de grados de libertad disponibles para los desplazamientos (Figura 5.3 A). El sistema CCR5[∅]-MRV tuvo mayor movilidad relativa en la región extracelular que en la región intracelular; siendo ésta última la que establece contacto estrecho con las proteínas de señalización, como se observa en los modos principales de la cadena principal (Figura 5.3 B y C; de izquierda a derecha) con los modos asociados al primero, segundo y tercer autovectores respectivamente. En estos sistemas de CCR5 la movilidad extracelular está estrechamente influida por la presencia del MRV. Los sistemas CCR5[∅]-mrv y CCR5[∅] mostraron una movilidad considerablemente mayor en las respectivas regiones intracelulares, como se espera en receptores no estativos. En cuanto a los movimientos de las cadenas laterales, en las regiones extracelulares (flechas color rojo) se observan valores mayores que en las regiones intracelulares (flechas color azul) en el espacio esencial bidimensional (Figura 5.3D). Los sistemas tuvieron comportamientos notablemente distintos al comparar los valores de RMSF esenciales en los sistemas con ligando CCR5[∅]-MRV y CCR5[∅], lo que demuestra las diferencias que establecen la presencia del ligando y la estatividad relativa del sistema CCR5[∅]-MRV. La excepción la constituye el inicio del dominio C terminal (CT), dada su naturaleza espacialmente desordenada. Las regiones de mayores desplazamientos que se observaron en los modos principales facilitaron las fluctuaciones de las cadenas laterales.

La distancia de separación entre el residuo R126^{3.50} ubicado en la hélice TM3 y parte del motivo conservado DRY, y Y297^{7.53} en la hélice TM7 y constituyente del motivo NPXXY y de una compuerta iónica atípica, puede

sugerir el estado funcional del receptor.⁵⁹ Una distancia suficientemente pequeña para la formación de una interacción no covalente entre un protón en el grupo guanidinio de la arginina y el átomo de oxígeno del hidroxilo fenólico de la tirosina permite la proximidad de las hélices TMs, favoreciendo una estructura central compacta (ver Anexo I: Resultados suplementarios). Si la distancia entre R126^{3.50} y Y297^{7.53} es mayor, se permite el alejamiento de las hélices TM3 y TM7, favoreciendo la formación del poro o canal de agua en el centro del receptor.³⁵ En el sistema CCR5[∅]-MRV esta distancia tuvo una magnitud de tendencia constante desde la conformación inicial; que aunque no fue lo suficientemente pequeña para la interacción ion-dipolo, se caracterizó por permanecer relativamente invariante a lo largo de la simulación, permaneciendo en un estado mecánicamente estático. En el sistema CCR5[∅] en cambio, esta distancia disminuyó respecto a la conformación inicial, y se mantuvo de esta forma la mayor parte del tiempo de simulación. Dado que la compuerta iónica en CCR5 no establece interacciones ion-ion sino ion-dipolo, esta variación de distancia entre estos residuos, y por ende entre TM3 y TM7 es relativamente más débil, y se espera en un estado basal o pasivo de un receptor con actividad constitutiva o intrínseca, ya que le permite adoptar espontáneamente varios estados funcionales en la ausencia de un ligando, como se conoce para el receptor CCR5.³⁶

Debido a las discrepancias del sistema con maraviroc desprotonado respecto a la información disponible sobre los GPCRs, y a la evidencia experimental del estado protonado de este ligando en el receptor, con pKa en medio acuoso de 7.3 (ver Anexo I), el sistema CCR5[∅]-mrv se descartó para los análisis posteriores; ya que la observación de este estado no es viable fuera de una simulación computacional, y muestra que el estado de protonación del ligando es de importancia fundamental para el estudio de este receptor y sus interacciones con el maraviroc.

Al interior del poro intramembranal, los residuos R126^{3.50}, Y214^{5.58}, R235^{6.35} y Y297^{7.53} constituyen la compuerta iónica relacionada con el paso de moléculas de agua a través del GPCR³⁵ (Figura 5.4). En el sistema CCR5[∅]-MRV, las distancias entre estos residuos se mantuvieron relativamente constantes; con distancias menores entre R126^{3.50}-Y214^{5.58} y R235^{6.35}- Y297^{7.53} que sugiere una interacción entre el grupo guanidinio de los residuos de arginina y el fenol del residuo de tirosina de estos pares. La distancia R235^{6.35}-Y297^{7.53} no indica interacciones; y en particular la distancia aproximada de 1 nm entre los residuos R126^{3.50}-Y297^{7.53} permite describir al poro en un estado quiescente en esta región. En contraste para el aporreceptor CCR5[∅], tres de las distancias entre estos residuos variaron a lo largo de la simulación. La distancia R126^{3.50}-Y297^{7.53} disminuyó hacia los 100 ns de simulación, y se alejaron los pares R126^{3.50} y Y214^{5.58} para permitir la interacción principal y conducir a un cierre de la compuerta del poro.

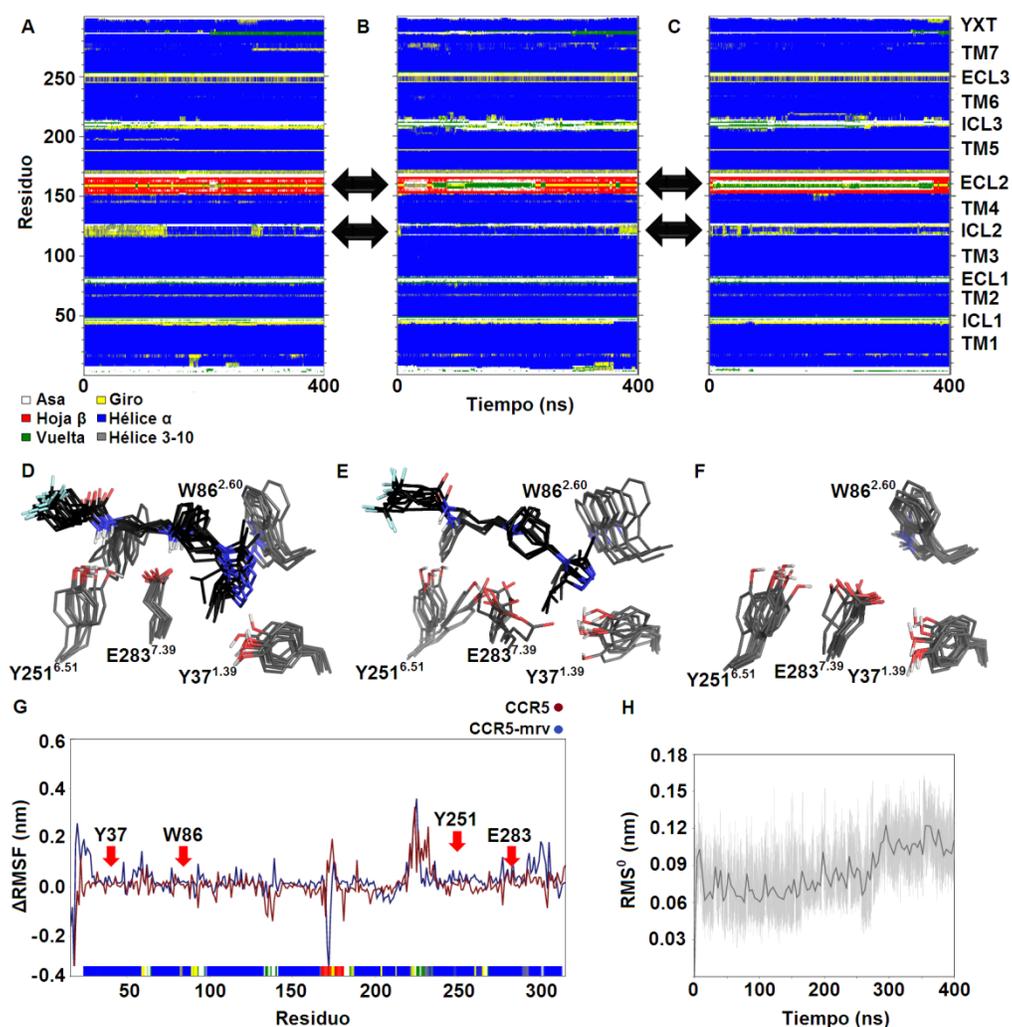


Figura 5.2. Sistemas de referencia con el receptor CCR5^Δ, con maraviroc protonado (MRV), maraviroc desprotonado (mrv) y sin ligando, respectivamente. (A, B, C). Estructuras secundarias de CCR5^Δ a través de los 400 ns de simulación. Se identifican del lado derecho los segmentos geométricos y en la parte inferior los motivos estructurales en código de colores. (D, E, F). Sitio de unión ortostérico del ligando de referencia que incluye la tétrada Y37^{1.39}, W86^{2.60}, Y251^{6.51} y E283^{7.39}, mostrando seis configuraciones tomadas de la trayectoria. (G). Valores de ΔRMSF de cadenas laterales del receptor a través del tiempo, respecto al sistema CCR5^Δ. (H). Valores de RMS⁰ del maraviroc protonado en su sitio de unión, donde la línea marcada central corresponde a los valores suavizados, y el contorno claro representa los valores precisos de RMS⁰.

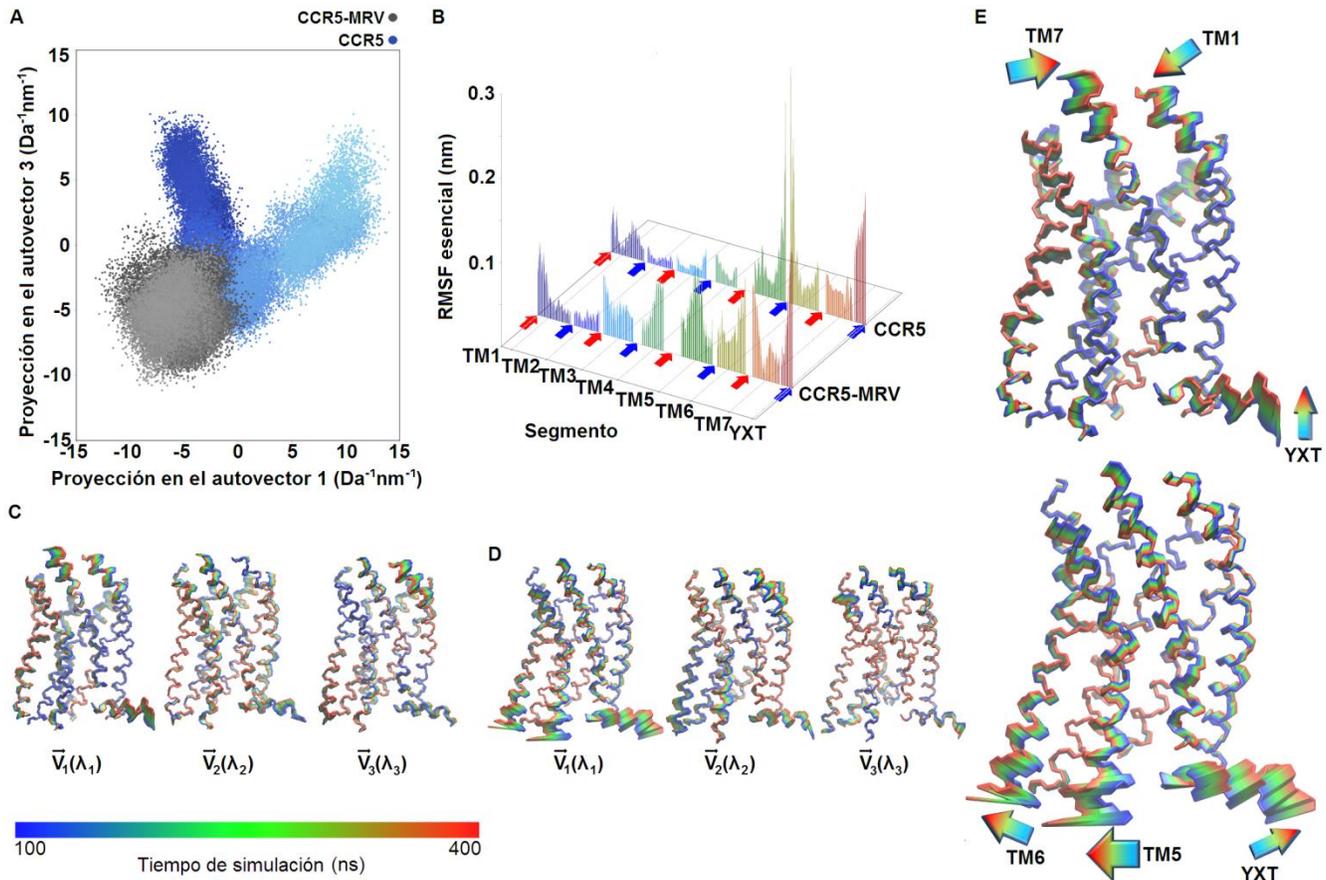


Figura 5.3. Análisis de componentes principales de los receptores CCR5^Δ-MRV y CCR5^Δ. (A) Proyecciones de las trayectorias euclidianas de CCR5^Δ-MRV y CCR5^Δ en la matriz de autovectores del mismo aporreceptor. El degradado de color indica el tiempo; inicio (oscuro), término (claro). (B) Valores de RMSF esencial para las cadenas laterales del receptor en los dos sistemas. Las flechas color rojo señalan los segmentos en contacto con el entorno extracelular y las flechas color azul a los segmentos en el entorno intracelular. (C) Modos principales del sistema CCR5^Δ-MRV para los autovectores 1 a 3 (de izquierda a derecha; 29%, 11% y 6% de la varianza) respectivamente. (D) Modos principales del aporreceptor CCR5^Δ para los autovectores 1 a 3 (27%, 12% y 7% de la varianza) respectivamente; calculados de 100 ns a 400 ns de la simulación. (E) Desplazamientos esenciales mayores de CCR5^Δ-MRV (arriba) y CCR5^Δ (abajo) señalados con flechas de color en relación con el tiempo. Es notable el movimiento de mayor varianza, o movimiento hipercedástico, en la región intracelular de TM5 y TM6 y en la hélice yuxtamembranal (YXT) en el sistema del aporreceptor.

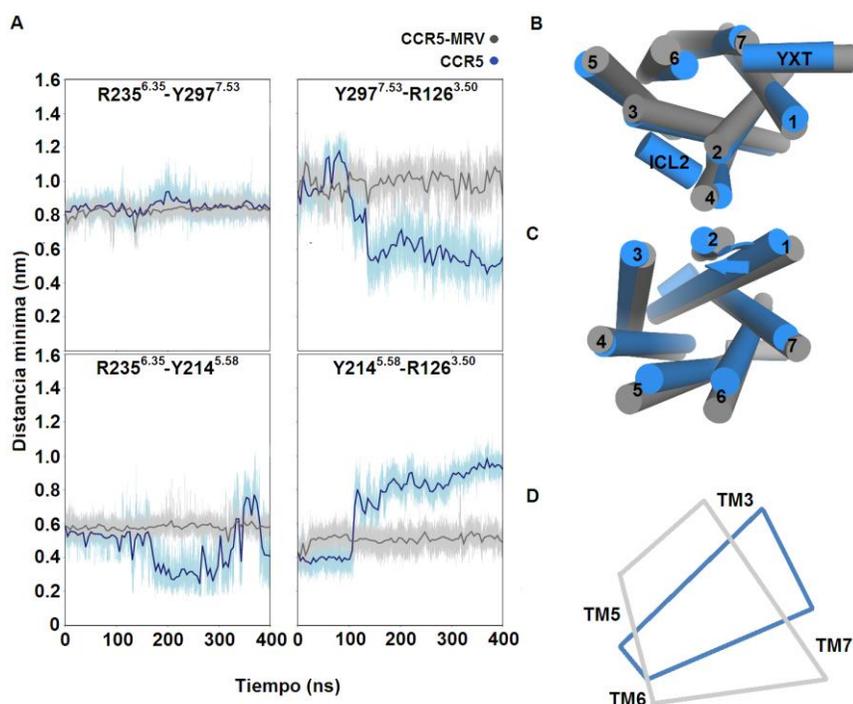


Figura 5.4. Compuerta iónica. (A). Distancias mínimas entre R126^{3.50}, Y214^{5.58}, R235^{6.35} y Y297^{7.53} en los sistemas de CCR5, conocida como compuerta iónica, a través del tiempo de simulación. El sistema CCR5^Δ-mrVH se caracterizó por un comportamiento mecánicamente estático y el sistema CCR5^Δ por un acercamiento notable de los residuos R126^{3.50}-Y297^{7.53} al separarse R126^{3.50} y Y214^{5.58}. Superposición de las TMs a partir de las estructuras helicoidales medias, (B) vista inferior o intracelular y (C) vista superior o extracelular. (D) Representación de cuadrilátero escalada de la compuerta iónica a los 300 ns de simulación.

Estos hallazgos muestran que el sistema CCR5^Δ-mrVH permaneció relativamente estático, o antagonizado en la simulación, y el sistema CCR5^Δ experimentó una relativa oclusión del poro, interpretada desde la comparación de las distancias interhelicoidales relativas (Figura 5.4.B y C) y las distancias mínimas entre los residuos a los 300 ns de simulación (Figura 5.4D); en congruencia con los resultados de Yuan, *et al.*³⁵ Para corroborar el estado del poro, se analizaron las trayectorias de moléculas de agua al interior del canal dividiéndolo en cuatro secciones longitudinales. Se generó una región extracelular donde se localiza el residuo E283^{7.39} interactuante con ligando, una segunda región media-extracelular hidrofóbica, una tercera región media-intracelular donde se ubica un residuo ácido conservado⁶⁰ y que se ha descrito con capacidad de coordinar un ion sodio,³⁴ el residuo D76^{2.50} para CCR5; y la región intracelular donde se encuentra la compuerta iónica previamente descrita y la región que permite la unión con los heterotrímeros de la proteína G (Figura 5.5A). La densidad de agua en el entorno de los receptores mostró mayor cantidad del solvente en CCR5^Δ-MRV respecto a CCR5^Δ (Figura 5.5, B y C), lo cual se cuantificó al contar las moléculas de agua en las cuatro secciones, donde el sistema estativo CCR5^Δ-MRV mostró numerosos intervalos de tiempo con un flujo acuoso

continuo (Figura 5.5.D), que puede relacionarse con la apertura del poro. El aporreceptor experimentó bloqueos parciales y totales al flujo acuoso particularmente en la región S3, y una solvatación y residencia abundante en la región 4 (Figura 5.5.E), en concordancia con los hallazgos obtenidos para la compuerta iónica.

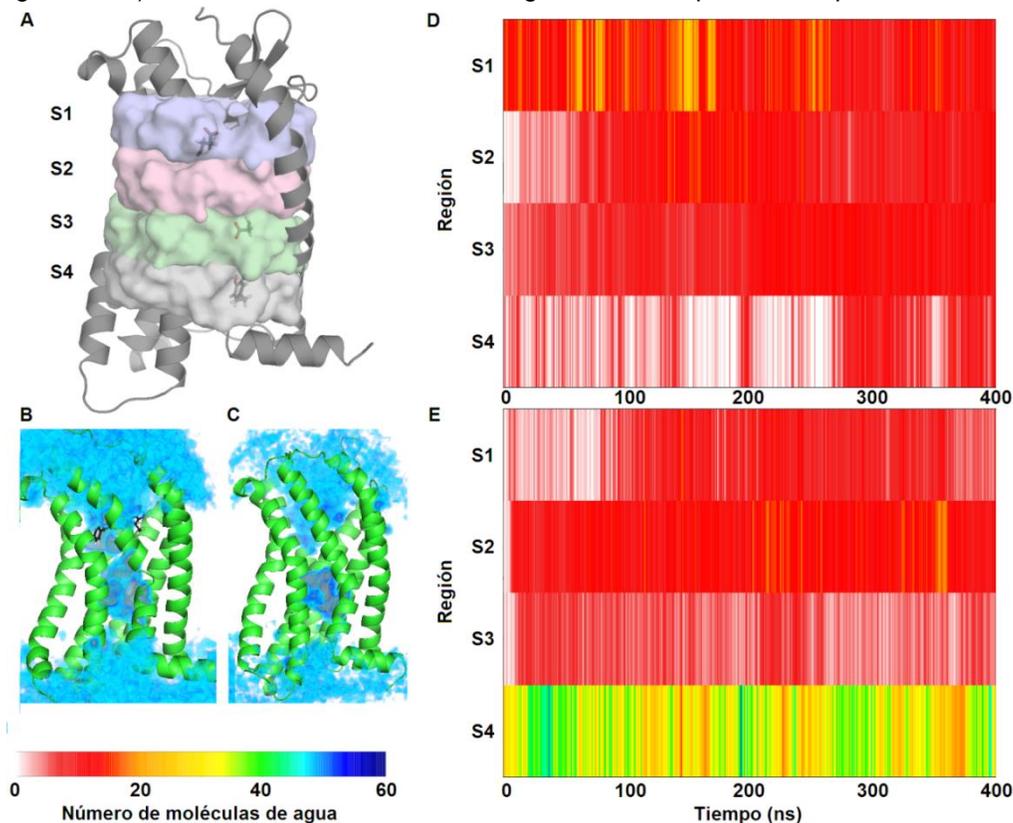


Figura 5.5. Poro interhelicoidal para el receptor CCR5. (A). Regiones S1, S2, S3 y S4 consideradas para el análisis de moléculas a través del receptor en los sistemas de receptor-ligando y aporreceptor. (B y C) Densidad media de agua calculada para CCR5^o-MRV y CCR5^o respectivamente. (D y E) Análisis cuantitativo de moléculas de agua en el poro del receptor en las cuatro regiones definidas, en escala de color de 0 a 60 moléculas para CCR5^o-MRV y para CCR5^o respectivamente.

La energía de unión estimada del maraviroc protonado se calculó mediante el muestreo umbeliforme (del inglés *Umbrella Sampling*, ver Anexo III) y el método de análisis por histogramas ponderados (*WHAM* por sus siglas en inglés). Del potencial de fuerza media asociado al jalón para el ligando en el sistema CCR5^o-MRV (Figura 5.6), al analizar el segmento con las pendientes mínimas se obtiene un valor medio estimado de energía de unión de 35.71 ± 0.04 kcal/mol.

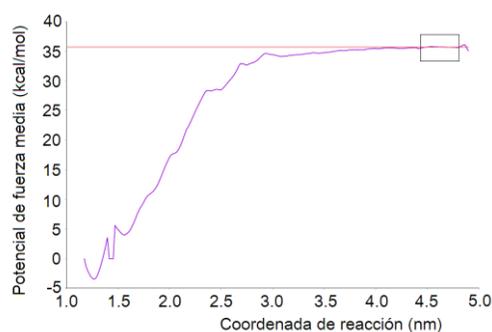


Figura 5.6. Potencial de fuerza media para la disociación del maraviroc del receptor CCR5 en el sistema CCR5 \emptyset -MRV (línea color morado). El segmento final de la curva con pendientes mínimas ($n=24$ pares ordenados de datos) se señala en color negro, y el valor medio estimado de 35.7 kcal/mol con un error estándar de la media de 0.04 kcal/mol en color rojo.

El estudio de los tres sistemas del receptor CCR5 permitió caracterizar algunos aspectos relevantes de dos de sus estados funcionales; el estado estativo ante la presencia del antagonista maraviroc, y un estado en ausencia de ligando, con tendencia al cambio, y que corresponde con un colectivo de conformaciones transicionales para salir del estado estativo, con relación probable con una actividad intrínseca de CCR5.

V.2. Sistemas del receptor de prueba CX3CR1

El estudio del receptor CX3CR1 se abordó de tres maneras distintas: Se compararon los hallazgos del modelo multitemplado, construido por enhebrado (ver Anexos I y II) con los modelos activo, transitivo y estativo esperados (Figura 5.7), de acuerdo al conocimiento de conformaciones experimentales de los GPCRs.³² Posteriormente los hallazgos del modelo multitemplado en complejos independientes con los ligandos evaluados (Tabla 5.1) se compararon con los modelos de los tres estados funcionales, y finalmente una comparación global (Figura 5.8).

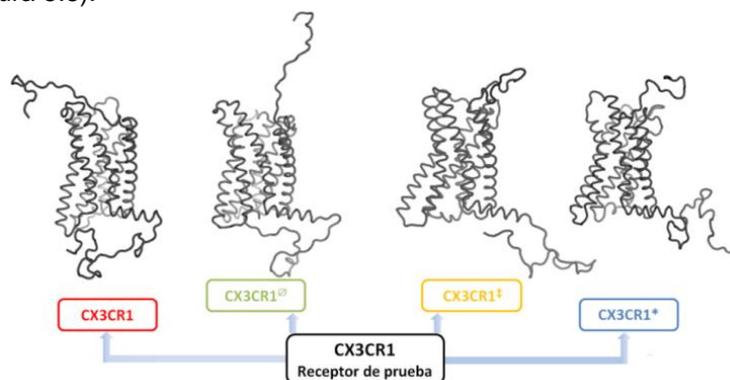


Figura 5.7. Esquema del modelo multitemplado (CX3CR1) y los tres estados funcionales del receptor (estativo CX3CR1 \emptyset , transitivo CX3CR1 \pm y activo CX3CR1 $*$).

Tabla 5.1. Ligandos evaluados en el modelo multitemplado CX3CR1.

Antagonista presuntivo		Agonista presuntivo (Dominio de quimiocina)			
Maraviroc (MRV) CX3CR1-MRV		Aglicofractalkina ^d CX3CR1-AFKN	Aglicona		
E6130 ^a CX3CR1-E61		Fractalkina, variante 1 ^e CX3CR1-FKN ^{G2}			
AZD8797 ^b CX3CR1-AZD					
Ligando 1 (LG1) ^c CX3CR1-LG1		Fractalkina, variante 2 ^e CX3CR1-FKN ^{G2S2}			
Código de glicanos:		N-Acetil-β-D-2-desoxi-2-glucosamina	α-D-Manosa	β-D-Galactosa	Sialato

Notas: a. El ligando E6130 tiene actividad experimental como antagonista ortostérico.⁶¹ b. El ligando AZD8797 es un antagonista alostérico.⁶² c. El ligando 1 se propuso a partir de la información obtenida de las interacciones de los sistemas CCR5^Δ-MRV y CX3CR1-MRV. d. Se consideró la aglicofractalkina, la aglicona del dominio de quimiocina de CX3CL1 (CDC por sus siglas en inglés), como punto de partida en las simulaciones con agonistas debido a que se conoce a nivel molecular la interacción de un GPCR con este complejo.⁶³ (Ver Anexos I y II). e. Debido a la variación de los carbohidratos biantenales complejos,^{64,65} se evaluaron las glicoformas G2 y G2S2 (Ver Anexo II).

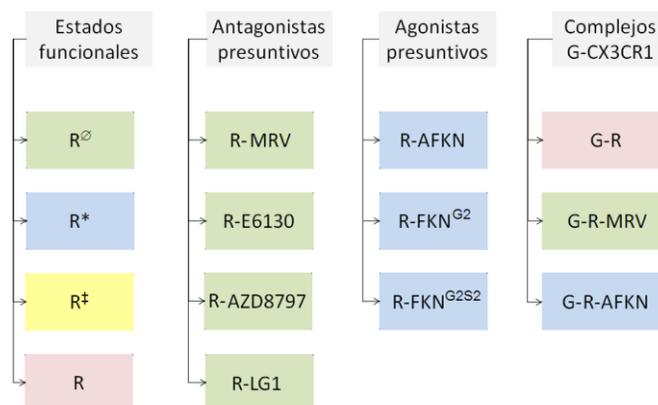


Figura 5.8. Esquema de los diferentes sistemas del receptor CX3CR1, abreviado como R (ver Tabla de abreviaturas o acrónimos). Los colores agrupan los sistemas con base en sus características esperadas. Estativo (verde), activo (azul), transitivo (amarillo) y basal (rosa).

V.2.1. Modelos CX3CR1 multitemplado, activo, transitivo y estativo

Como se menciona en sección IV, el modelo principal de CX3CR1 se obtuvo a partir de la información experimental de receptores con mayor similitud en estructura primaria, por lo que la comparación con los estados esperados activo, transitivo y estativo fue necesaria para determinar las características generales de este modelo multitemplado.

De las estructuras secundarias a través del tiempo para los sistemas, el modelo multitemplado CX3CR1 presentó inicialmente el plegamiento beta en ECL2 dadas las estructuras base para su modelado. Este segmento se desestabilizó hacia los 150 ns de simulación debido a la ausencia de ligando (Figura 5.9A). El plegamiento helicoidal de ICL2 se estabilizó en el modelo activo CX3CR1* en todo el intervalo (Figura 5.9B), y en CX3CR1‡ este segmento tuvo un orden relativo (Figura 5.9C) en comparación con CX3CR1∅ (Figura 5.9D). La modificación torsional debida a las rupturas de helicidad en TM2 en CX3CR1 y CX3CR1* fue leve comparada con CX3CR1‡ y CX3CR1∅. Los valores de RMSF para las cadenas laterales resaltan la diferencia de estados, en particular la fluctuación baja del segmento ECL3 del modelo transitivo y la movilidad mayor en ICL2 en el modelo estativo (Figura 5.9E). Los residuos equivalentes en el sitio de unión ortostérico constituido por Y38^{1.39}, W87^{2.60}, F113^{3.36}, R191^{4.54}, Y247^{6.51}, E279^{7.39} y F283^{7.43} tuvieron una alta fluctuación en los modelos multitemplado y transitivo, teniendo cierta estabilidad dinámica en los modelos activo y estativo (Figura 5.9 F-I). Ante la ausencia de ligandos, los residuos Y247^{6.51} y E279^{7.39} formaron una interacción prevaeciente en CX3CR1 y CX3CR1∅, infrecuente, alrededor de 10% del tiempo de simulación para CX3CR1* y ocasional en CX3CR1‡.

En el análisis de componentes principales los estados CX3CR1 y CX3CR1∅ se sitúan próximos en el espacio esencial, en concordancia con la naturaleza del modelado; y el estado CX3CR1* se presentó

inaccesible para estos dos sistemas en el tiempo de simulación de 400 ns (Figura 5.10A). Los valores de RMSF esenciales para cadenas laterales del dominio transmembranal son muy pequeños para CX3CR1 y mayores para los tres modelos funcionales (Figura 5.10B), que puede representar la influencia de los modelos estativos en el modelado de CX3CR1. Se distinguen los valores menores en TM4, TM5 y YXT para el modelo activo, los valores mayores en TM1, TM2 y TM3 en el modelo transitivo y en TM6 y TM7 para el modelo estativo. Los modos principales de los cuatro sistemas revelaron desplazamientos clave para caracterizar el estado, en particular el amplio movimiento de la región intracelular de TM6 (ICL3-TM6) en CX3CR1* debido probablemente a una transición fuera del estado activo; así como una traslación frontal de YXT, mientras que en CX3CR1 \emptyset el movimiento de vaivén de TM1 y TM2 y un giro endomembranal de YXT hacia atrás. El modelo transitivo presentó flexiones pronunciadas en las hélices TM5 y TM6, y el modelo multitemplado conservó la estatividad inferida a partir de las fluctuaciones del sitio ortostérico y los valores de RMSF esenciales (Figura 5.10, C-I). La ausencia del ligando en estos sistemas limitó en cierta medida la comparación detallada entre los tres estados funcionales, ya que permite una posible convergencia de conformaciones.

La compuerta iónica en los modelos CX3CR1 tuvo una diferencia importante entre CX3CR1* y los otros tres modelos, debido al cierre del poro acuoso a través de la interacción R127^{3.50}-Y293^{7.53}, mientras que la mayor distancia R127^{3.50}-Y293^{7.53} se presentó en CX3CR1 \emptyset , y en el modelo multitemplado fue disminuyendo gradualmente esta apertura al favorecer la interacción R127^{3.50}-Y211^{5.58} (Figura 5.11A). El cierre del poro acuoso para el modelo activo concordó con el aporreceptor de referencia CCR5. Sin embargo, estos sistemas no son totalmente comparables y el comportamiento del estado activo CX3CR1 se replanteó en los sistemas con ligandos agonistas. Las diferencias en las posiciones de los extremos intracelulares de TM5 y TM6 en los cuatro modelos sugieren que es una característica distintiva del estado funcional. Los diagramas bivariados de RMSD del dominio transmembranal (TMD) respecto a la distancia entre las regiones ICL1-TM2 e ICL3-TM6 (Figura 5.11D) permitieron la segregación de los modelos multitemplado y transitivo; pero en CX3CR \pm y CX3CR1 \emptyset se identificó un aparente cambio de estado. Además, se confirma la compacidad relativa del poro en los modelos multitemplado y activo respecto a los otros dos. Al analizar los histogramas bivariados de los ángulos diedros χ_1 y χ_2 de Y293^{7.53}, se observa que predominan tres rotámeros, dos de ellos compartidos en los modelos multitemplado y activo, con la cadena lateral hacia el poro y hacia TM7, y sólo uno en el modelo estativo. La moda histográfica en el modelo transitivo es intermedia entre CX3CR1* y CX3CR1 \emptyset , mientras que el modelo multitemplado mostró una bimodalidad importante y distintiva (Figura 5.11E).

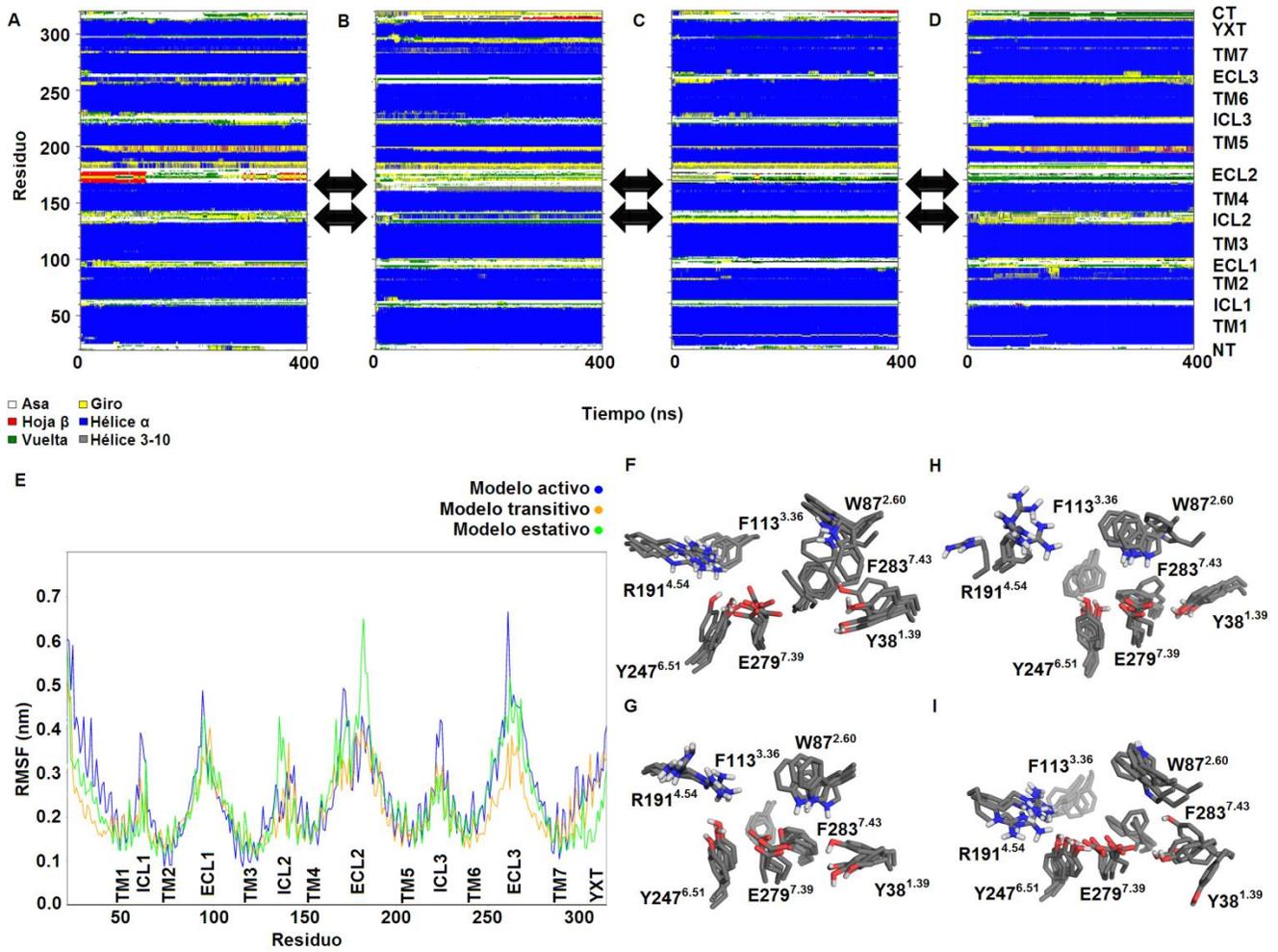


Figura 5.9. Análisis de estructura secundaria y sitio ortostérico para el receptor en modelo multitemplado (CX3CR1), activo (CX3CR1*), transitivo (CX3CR1‡) y estatico (CX3CR1Ⓢ) respectivamente. (A, B, C y D) Estructuras secundarias a través del tiempo para los cuatro modelos. (E) Valores de RMSF de cadenas laterales de los tres sistemas. (F, G, H e I) Residuos del sitio de unión equivalente del maraviroc: Y38^{1.39}, W87^{2.60}, F113^{3.36}, R191^{4.54}, Y247^{6.51}, E279^{7.39} y F283^{7.43} para los cuatro modelos.

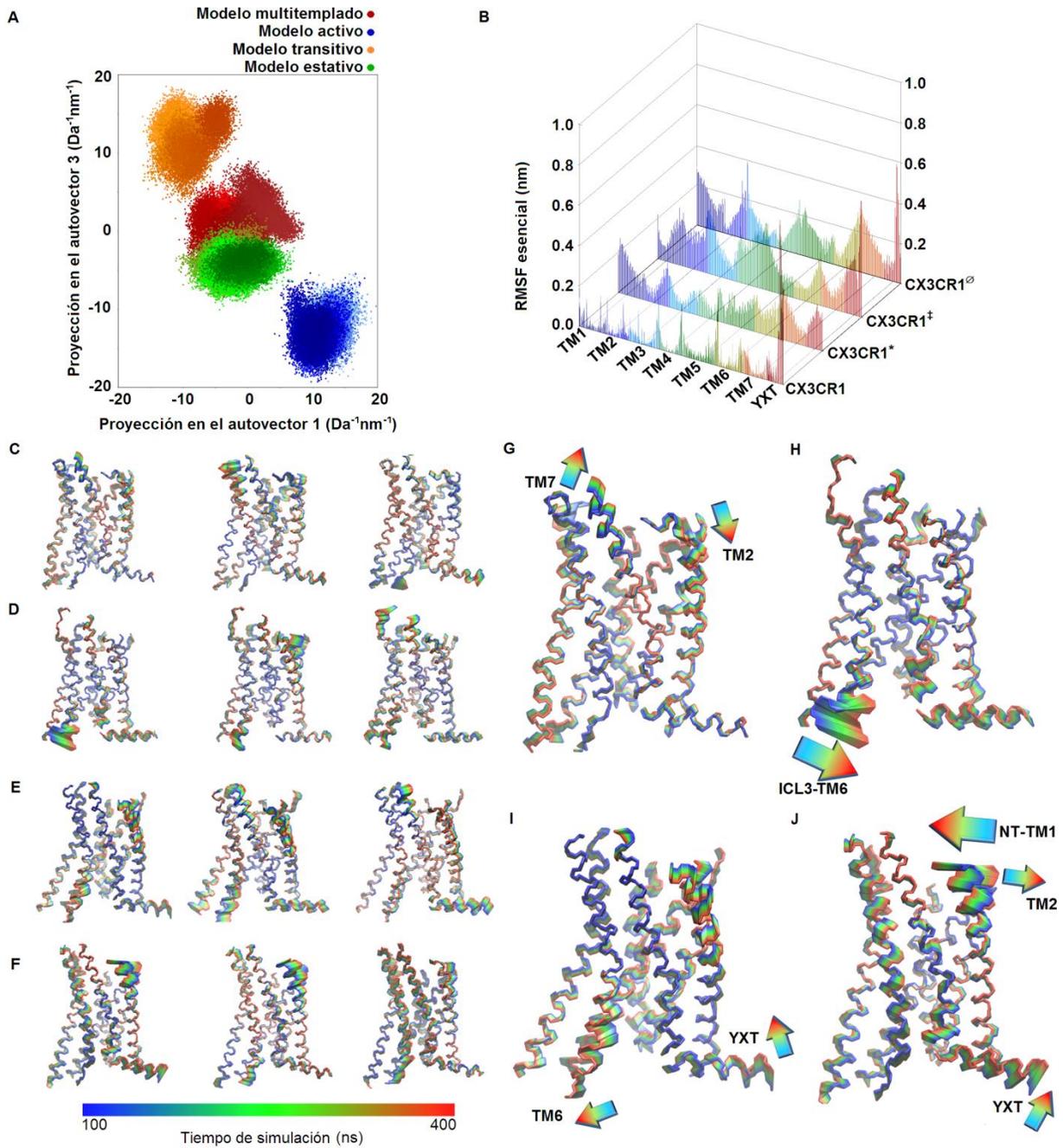


Figura 5.10. Análisis de componentes principales para el receptor CX3CR1, modelos multitemplado, activo, transitivo y estativo. (A) Proyecciones de las trayectorias euclidianas en la matriz de autovectores del modelo multitemplado, mostrando las diferencias de locaciones esenciales accesibles para cada estado. (B) Valores de RMSF esencial de las cadenas laterales del dominio transmembranal para los modelos CX3CR1, CX3CR1*, CX3CR1+ y CX3CR1 $^{\ominus}$. Modos principales de (C) CX3CR1, (D) CX3CR1*, (E) CX3CR1+ y (F) CX3CR1 $^{\ominus}$. (G, H, I, J) Movimientos esenciales mayores en CX3CR1, CX3CR1*, CX3CR1+ y CX3CR1 $^{\ominus}$, señalados con flechas.

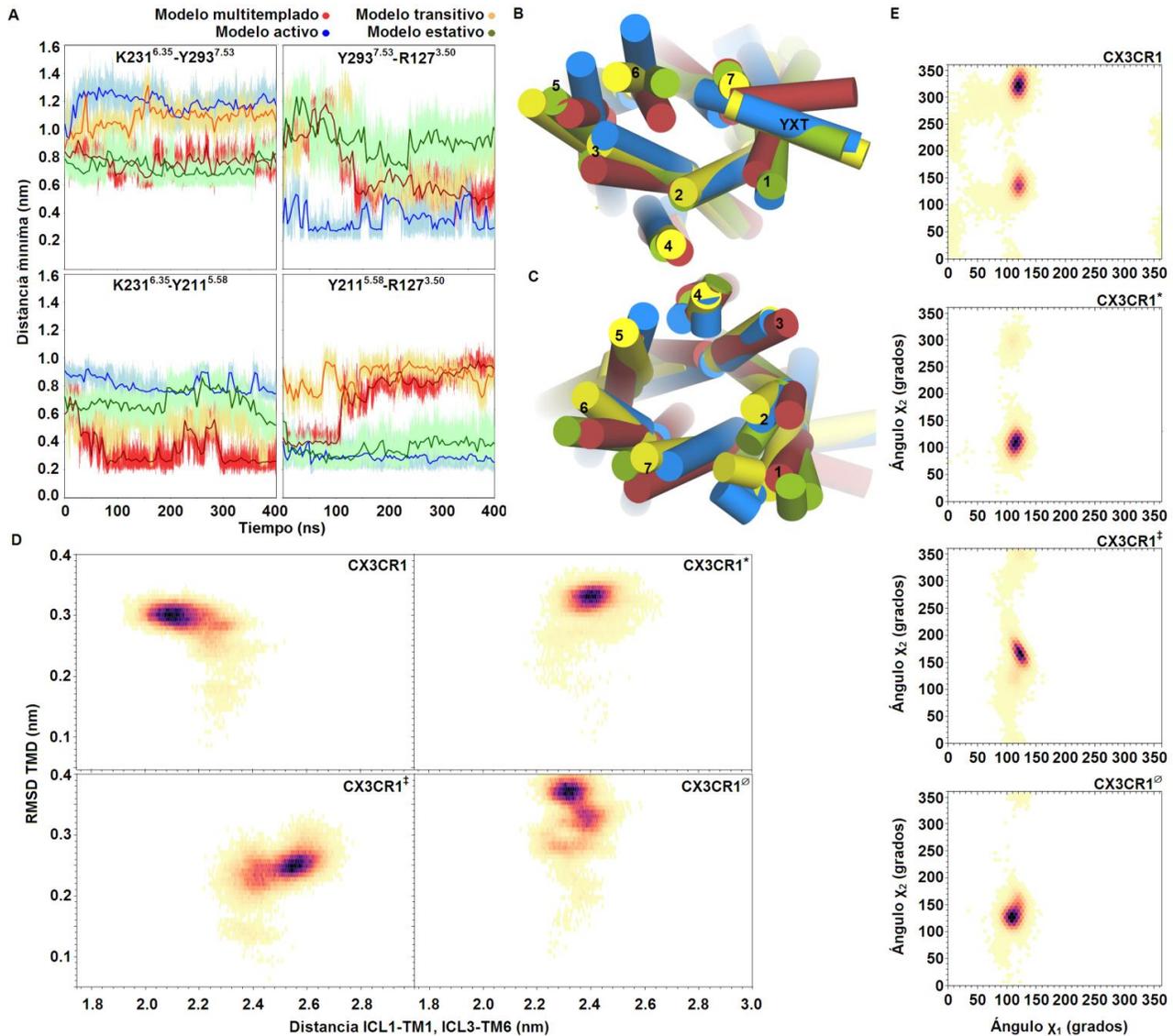


Figura 5.11. (A) Distancias de interacción de la compuerta iónica. En CX3CR1 y CX3CR1[∅] los pares K231^{6.35}-Y293^{7.53} y K231^{6.35}-Y211^{5.58} establecieron contactos cercanos, manteniendo la compuerta Y293^{7.53}-R127^{3.50} abierta al menos en los primeros 100 ns de simulación. En CX3CR1[†] los pares K231^{6.35}-Y293^{7.53} y K231^{6.35}-Y211^{5.58} se mantuvieron con interacciones menores, permitiendo el cierre de la compuerta Y293^{7.53}-R127^{3.50}. Superposición de las hélices (B) con vista superior o extracelular y (C) inferior o intracelular. (D) Histogramas bivariados para las distancia del extremo intracelular de TM2 (ICL1-TM2) y el extremo intracelular de TM6 (ICL3-TM6) con el valor de RMSD del dominio transmembranal (TMD). Con una flecha en el diagrama del modelo multitemplado, se propone la trayectoria de estas variables en los tres cambios de estado. (E) Histogramas bivariados para la ocurrencia de los ángulo diedros χ_1 y χ_2 de Y293^{7.53} en los tres sistemas. Para CX3CR1, la distribución es bivariada, con los dos pares predominantes con el grupo lateral hacia el poro y hacia la TM7.

V.2.2. Modelo multitemplado con antagonistas presuntivos

Para los sistemas de CX3CR1 con ligandos presuntamente antagonistas (Figura 5.12), la característica común en los cuatro sistemas fue la estabilización del motivo beta plegado en ECL2 por la presencia del ligando, como se observó en los sistemas de referencia de CCR5. Un hallazgo notable en el sistema con el ligando E6130 fue el desplegamiento del extremo extracelular de TM1 en ocho residuos, que no se propagó hacia el dominio intramembranal. El plegamiento helicoidal en ICL2 se desfavoreció en los 400 ns de simulación para los sistemas con los antagonistas E6130 y AZD8797, y predominó en el sistema con maraviroc (Figura 5.12A). Este segmento intracelular se plegó hacia la mitad de la simulación, lo que puede implicar al menos dos eventos. Dado que un estado funcional es un colectivo de conformaciones,⁶⁶ el plegamiento helicoidal puede responder a la conversión hacia otros conformeros del colectivo estativo, o a un cambio de estado funcional. Adicionalmente, se presentó una extensión helicoidal en la región ECL2-TM5 hacia el final de la simulación en los sistemas con AZD8797 y LG1. El sistema con E6130 ocasionó una ruptura helicoidal en TM7 cercana al sitio ortostérico (Figura 5.12B). El ligando AZD8797 indujo una considerable ruptura helicoidal en TM5 (Figura 5.12C), distante del sitio alostérico que está formado por los residuos D77^{2.50}; el residuo conservado de interacción con el catión sodio; W244^{6.48}, H285^{7.45}, N289^{7.49}, P290^{7.50} y Y293^{7.53}; cuatro de ellos implicados en la formación del complejo con sodio, el residuo conservado P290^{7.50} y la tirosina que integra la compuerta iónica. En el sistema con LG1, la hélice TM6 extendió el plegamiento helicoidal hacia ECL3 (Figura 5.12D); mientras que en los otros tres sistemas este segmento tuvo mayor variabilidad en el plegamiento. En el análisis de componentes principales de la cadena principal transmembranal se agruparon los cuatro sistemas con ligandos en locaciones muy similares, respecto al modelo multitemplado (Figura 5-13A), donde el sistema con E6130 tuvo un mayor distanciamiento del modelo multitemplado, y el sistema con LG1 dirigiéndose hacia la locación del modelo transitivo. Los modos principales del dominio TMD mostraron resultados opuestos para los desplazamientos de la hélice YXT para los sistemas con MRV y LG1; de descenso hacia la interfase intracelular, y de ascenso intramembranal, respectivamente (Figura 5-13, B-E). Adicionalmente la flexión de las hélices TM5 y TM6 es equiparable a la flexión obtenida con el modelo transitivo CX3CR1[‡]. El probable cambio de estado de CX3CR1-LG1 hacia la activación se podría confirmar con un tiempo de simulación más prolongado, ya que los cambios de estado en los GPCRs se ha reportado en escalas de tiempo de simulación de microsegundos.^{35,38}

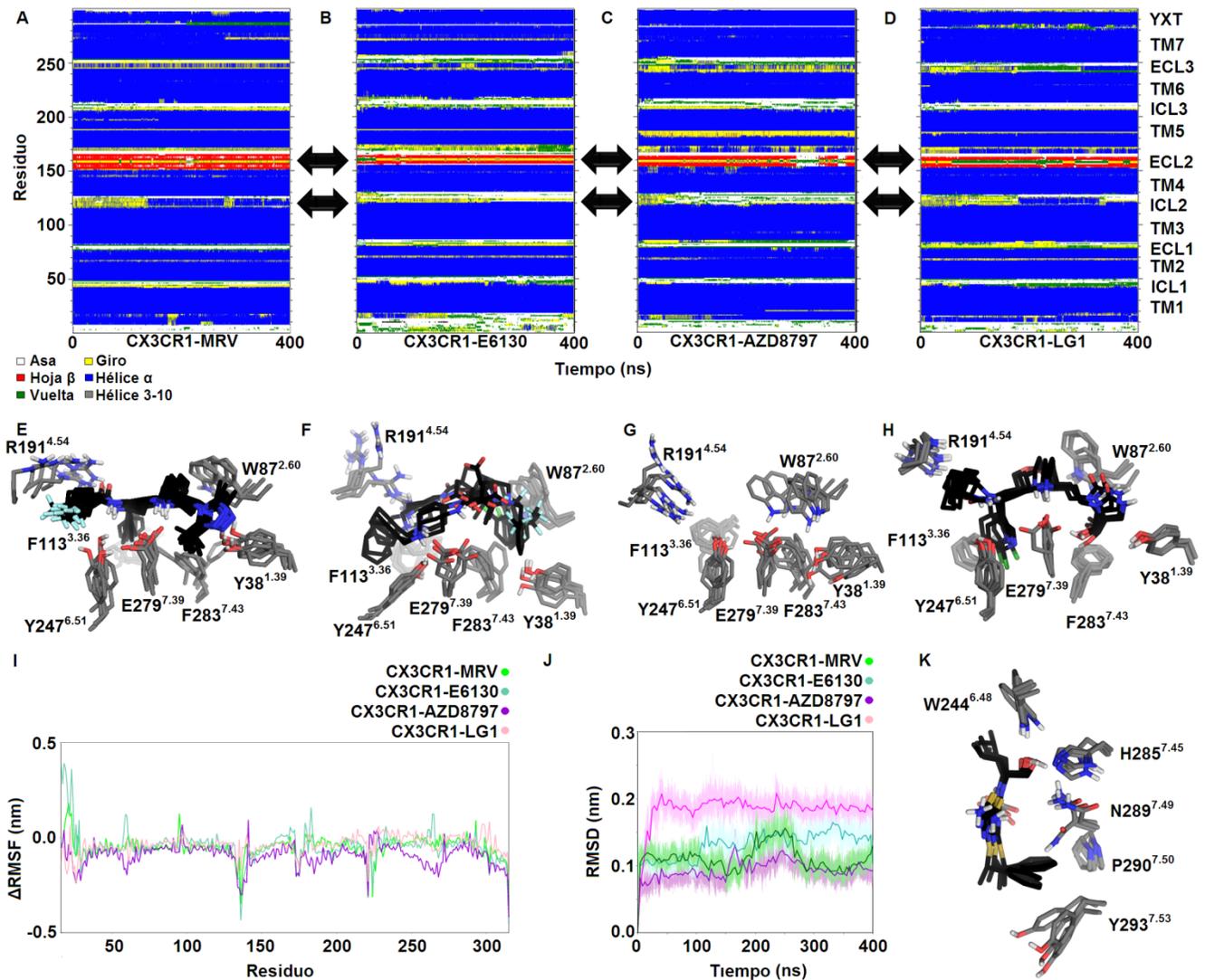


Figura 5.12. Análisis de estructura secundaria y del sitio de unión para los estados estativos. Estructura secundaria a través del tiempo para (A) CX3XR1-MRV, (B) CX3CR1-E6130, (C) CX3CR1-AZD8797 y (D) CX3CR1-LG1. (E-H) Sitio de unión de los ligandos, respectivamente. (I) Diferencia de RMSF para las cadenas laterales de los sistemas respecto al modelo multitemplado CX3CR1. (J) Valores de RMSD a través del tiempo de la cadena principal de los residuos del sitio alostérico; W244^{6.48}, H285^{7.45}, N289^{7.49}, P290^{7.50} y Y293^{7.53}. (K) Vista frontal del sitio alostérico con el ligando AZD8797.

Al analizar la compuerta iónica en los cuatro sistemas, la interacción principal de R127^{3.50}-Y293^{7.53} fue más estrecha en los sistemas con maraviroc y LG1, en oposición con la diferencia de desplazamientos principales de la hélice YXT (Figura 5.14A). En general, los sistemas tuvieron una tendencia estativa, con dos excepciones; el distanciamiento de K231^{6.35}-Y211^{5.58} en el sistema CX3CR1-E6130, y de K231^{6.35}-Y293^{7.53} en el sistema CX3CR1-LG1. Las regiones intracelulares de las hélices TM5 y TM7 tuvieron la máxima separación media en el sistema con LG1 (Figura 5.14 B y C), con tendencia aparente al estado transitivo. Los dos sistemas con los

antagonistas E6130 y AZD8797 no difieren notablemente en sus conformeros medios (Figura 5.14D), por lo que la trayectoria de los dos sistemas estativos podría converger aun teniendo dos tipos distintos de antagonistas. Al evaluar las variables para analizar el cambio de estado, los cuatro sistemas manifestaron un aumento en la distancia de separación de las hélices y el RMSD transmembranal, favoreciendo la apertura del poro intrahelicoidal, siendo mayor el incremento en los sistemas con E6130 y LG1 (Figura 5-14E). Las distribuciones de los diedros χ_1 y χ_2 son semejantes en los sistemas CX3CR1, CX3CR1-E6130 y CX3CR1-LG1; mientras que en el sistema con AZD8797 las rotaciones en el sitio alostérico se encuentran fuertemente restringidos. Estas distribuciones son multimodales, predominando la mayor variación de χ_2 respecto a χ_1 ; $0^\circ < \chi_2 < 360^\circ$ y $\chi_2 < 135^\circ$. El sistema CX3CR1-E6130 con cuatro modas leptocúrticas, es decir, con alta frecuencia, presentó cuatro rotámeros predominantes que constituyen cuatro conformeros del colectivo estativo para este sistema. En el caso de los sistemas CX3CR1-LG1 y CX3CR1-MRV, las modas alrededor de $(120^\circ, 345^\circ)$ son mesocúrticas respecto al sistema con E6130; y en el segundo el rotámero asociado es prácticamente inaccesible para el sistema (Figura 5.14F).

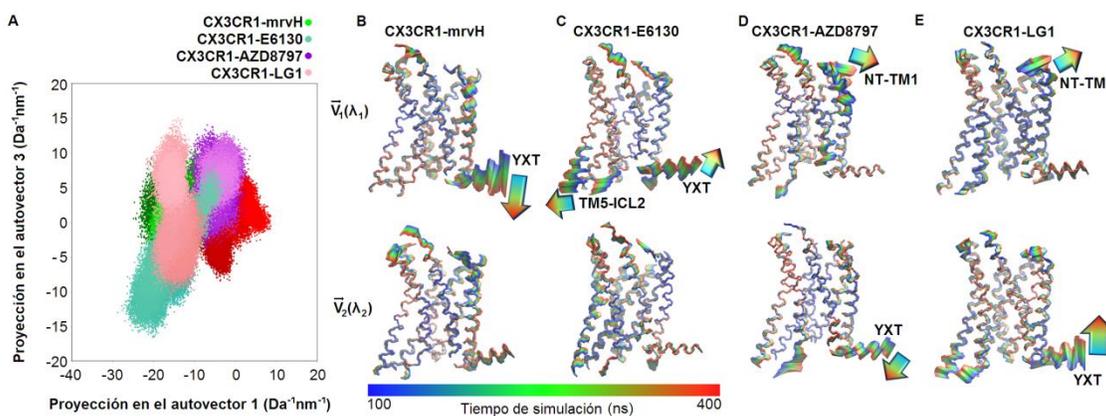


Figura 5.13. Análisis de componentes principales de los sistemas con antagonistas presuntivos. (A) Proyecciones de las trayectorias de la cadena principal de los sistemas CX3CR1-MRV, CX3CR1-E6130, CX3CR1-AZD8797 y CX3CR1-LG1 sobre la matriz de autovectores del modelo multitemplado. Modos principales de los sistemas (B) CX3CR1-MRV, (C) CX3CR1-E6130, (D) CX3CR1-AZD8797 y (E) CX3CR1-LG1.

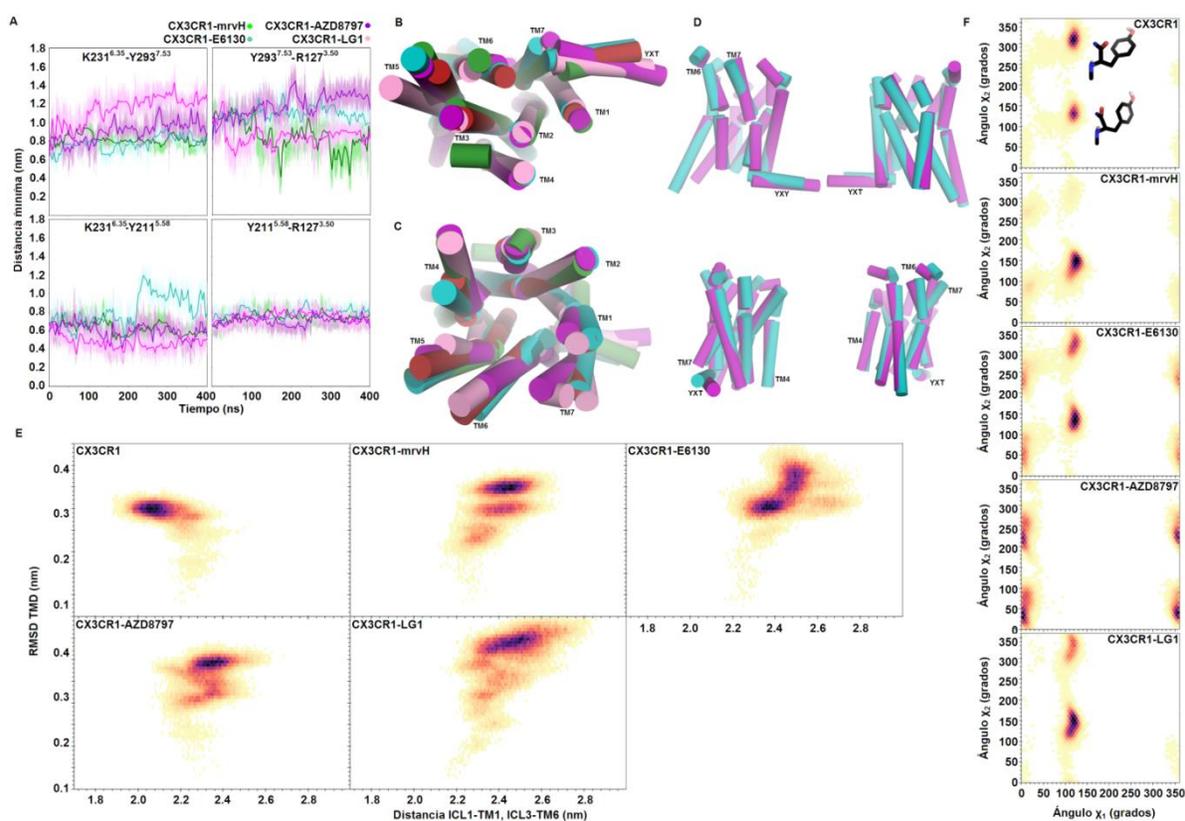


Figura 5.14. Distancias de los residuos en la compuerta iónica. (A) Interacciones entre los cuatro residuos. (B) Vista inferior y (C) vista superior de la superposición de los cinco dominios TMD. (D) Comparación de los dos sistemas con el antagonista ortostérico E6130 y el antagonista alostérico AZD8797. (E) Histogramas bivariados de RMSD del dominio TMD y la distancia ICL1-TM2 e ICL3-TM6. Al analizar estas distribuciones a través del tiempo, los cuatro sistemas con ligando aumentaron los valores de ambas variables. (F) Histogramas para los ángulos diedros χ_1 y χ_2 de Y293^{7.53}. Los tres sistemas con ligandos ortostéricos se comportaron de manera similar en sus distribuciones, y el sistema con AZD8797 fue distinto debido a su ubicación en el receptor.

En el cálculo de la energía libre estimada de disociación del maraviroc, el jalón del ligando se situó en configuraciones muy inestables, y que convergieron a una locación similar al sistema inicial, lo que ocasionó la no intersección de los histogramas (Ver anexo I), y una estimación de energía libre de aproximadamente 10 kcal/mol, que no es comparable con la estimación de energía libre del sistema de referencia CCR5-MRV (Figura 5.15). Este comportamiento de caídas en el gráfico de PMF se debe fundamentalmente a la interacción del dominio amino terminal NT con el ligando durante su jalón por la coordenada de reacción, lo que dificultó sustancialmente la estimación inestegada. En el sistema de referencia CCR5-MRV, el dominio NT está ausente desde la estructura experimental y no se presentó este fenómeno. El resultado de PMF y la estimación de la asíntota que representa la energía libre de disociación se puede precisar con un aumento del tamaño de la caja de simulación y un extensivo muestreo umbeliforme, lo que implica un costo computacional elevado, y equiparar al omitir el dominio NT de CX3CR1.

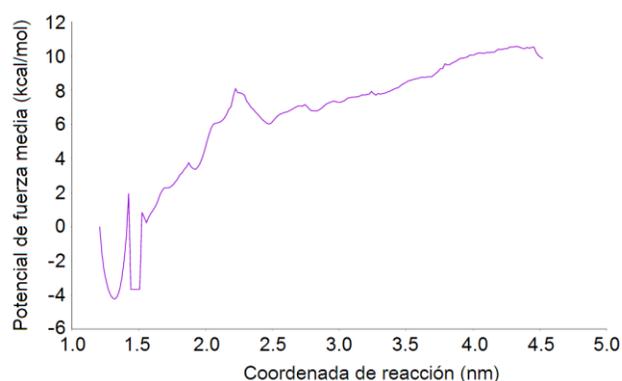


Figura 5.15. Potencial de fuerza media para el maraviroc en CX3CR1.

V.2.3. Modelo multitemplado con agonistas presuntivos

El agonista endógeno en su forma soluble sCX3CL1, se comparó inicialmente con el sistema testigo US28-AFKN, del receptor viral US28 (ver Anexo I para resultados de este sistema testigo); y posteriormente se analizaron los sistemas de receptor-aglicona de sCX3CL1 (CX3CR1-AFKN); y las variantes glicosiladas receptor-sCX3CL1 (CX3CR1-FKN^{G2} y CX3CR1-FKN^{G2S2}, ver Anexo II para la descripción de los glicanos).

El receptor en los tres sistemas formó el motivo de dos hojas beta en el segmento ECL2 (Figura 5-16 A-C), coincidiendo con los sistemas antagonizados, y notablemente en el sistema CX3CR1-FKN^{G2} ocurrió una extensión de la hélice YXT por plegamiento helicoidal hacia el CT. Además, en el sistema CX3CR1-FKN^{G2S2} se presentó el plegamiento helicoidal en ICL2, esperado en un sistema con un agonista, y ausente con los dos sistemas antagonizados CX3CR1-E6130 y CX3CR1-AZD8797. Este plegamiento ocurrió desde el inicio de la simulación, de manera equiparable con el sistema CX3CR1*. Las interacciones del ligando en las tres variantes coincidieron en la región extracelular de NT-TM1, TM2-ECL1-TM3, ECL2 y TM5-ECL3-TM3 del receptor (Figura 5-16D). Al comparar los valores de RMSD de estos segmentos entre los tres sistemas, se presentó una diferencia pequeña de casi 0.3 nm hacia el final de la simulación, pero de hasta 0.4 nm para CX3CR1-FKN^{G2S2} (Figura 5-16E). Asimismo, los modos principales de esta región y el ligando son notablemente distintos, con una movilidad extrema en el NT y un desplegamiento del CT de AFKN en el sistema CX3CR1-AFKN (Figura 5-16F). En contraste, el sistema glicosilado CX3CR1-FKN^{G2} no mostró desplegamiento ni fluctuaciones extremas en el NT del receptor, y en el sistema con glicosilación G2S2, el aumento de RMSD se debió de manera importante a la fractalkina.

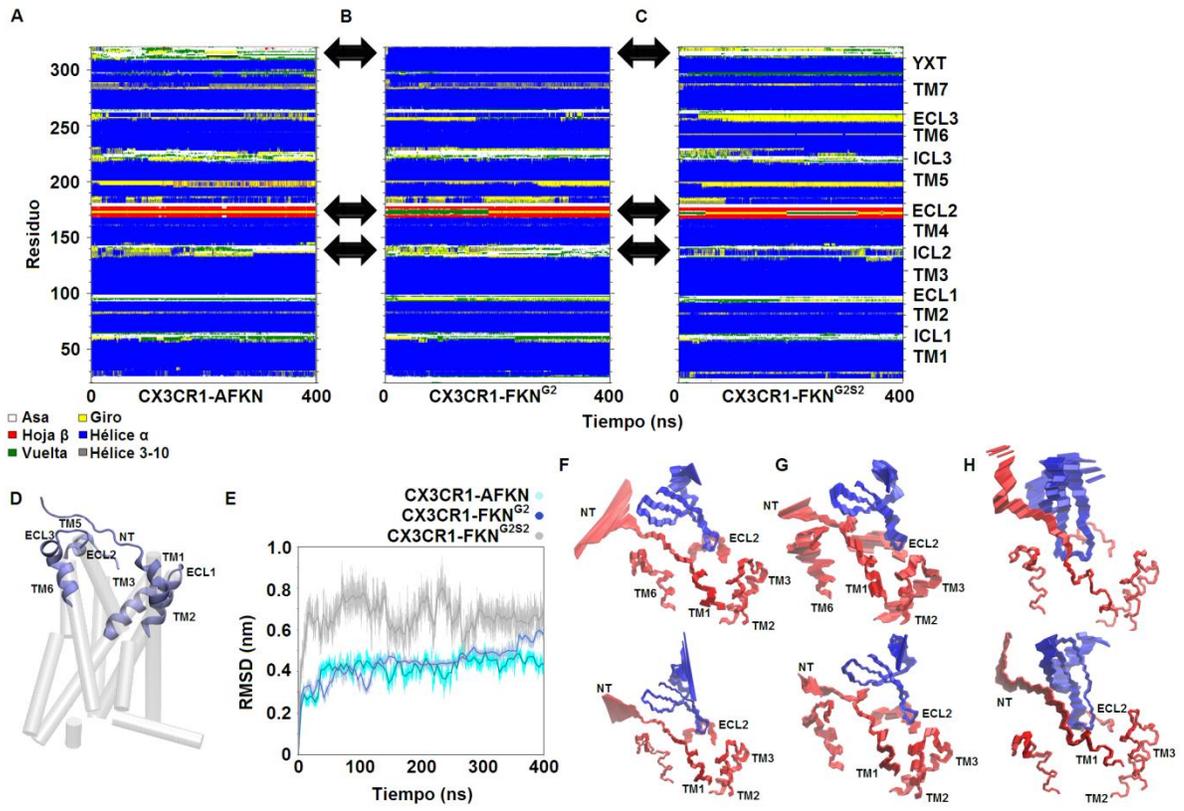


Figura 5.16. Receptores en estado activo o en complejo con agonistas. (A-C) Estructuras secundarias de los receptores en los sistemas CX3CR1-AFKN, CX3CR1-FKN^{G2} y CX3CR1-FKN^{G2S2} respectivamente. (D) Región NT-TM1, TM2-ECL1-TM3, ECL2, TM7-ECL3-TM7 de contactos del ligando en el receptor. (E) RMSD de la cadena principal de los residuos de la región de contacto. (F, G, H) Modos principales del ligando y su región de contacto en las simulaciones.

En el análisis de componentes principales, los sistemas con AFKN y FKNG2 se ubicaron notablemente distantes del sistema CX3CR1, y próximos a la locación del modelo activo CX3CR1* (Figura 5.17A). Para el receptor con FKN^{G2S2}, la ubicación relativa de la proyección en el subespacio esencial, más que una tendencia divergente, tiene la característica de un estado relacionado al activo y al transitivo; suposición sostenida con la gran inclinación de los extremos intracelulares de TM5 y TM6, como lo muestran los modos principales (Figura 5.17 B-D). Los modos principales incluyen los desplazamientos de acercamiento y alejamiento del extremo celular de TM5 y TM6, el ascenso de la hélice YXT hacia la membrana, y alejamientos de los extremos extracelulares de TM1 y TM7 respecto al sitio ortostérico. El distanciamiento de TM5 y TM6 coincidió en estos casos con el ascenso de YXT y viceversa.

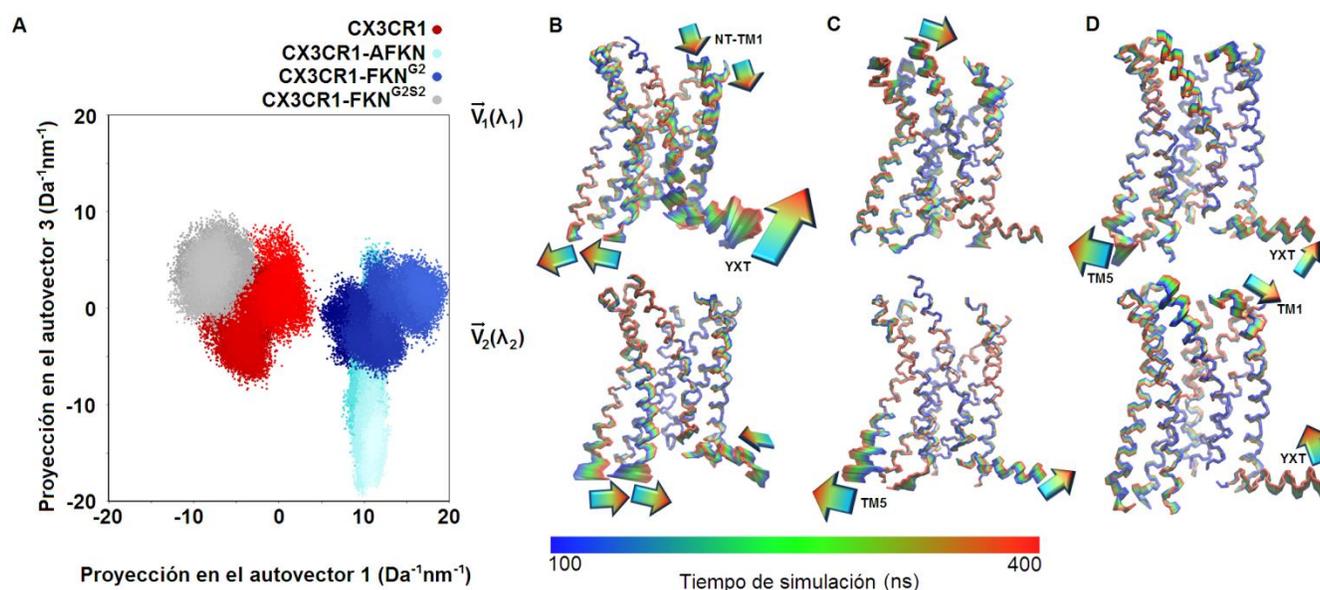


Figura 5.17. Análisis de componentes principales para los sistemas activos. (A) Proyecciones de las trayectorias de la cadena principal de los sistemas CX3CR1-AFKN, CX3CR1-FKN^{G2} y de CX3CR1-FKN^{G2S2} sobre la matriz de autovectores del modelo multitemplado CX3CR1. El sistema CX3CR1-FKN^{G2S2}, la glicofoma extendida, se localiza hacia la región de los sistemas del ligando LG1 y el modelo transitivo. Modos principales de (B) CX3CR1-AFKN, presentando desplazamientos mínimos de TM5, TM6 y YXT (C) CX3CR1-FKN^{G2}, la versión trunca del glicano, con mayor predominio de separación de TM5 y TM6; y (D) CX3CR1-FKN^{G2S2}.

El sistema CX3CR1-FKN^{G2S2} tuvo el acercamiento R127^{3.50}-Y293^{7.53} más prolongado durante la simulación, seguido del sistema CX3CR1-AFKN. El alejamiento Y211^{5.58}-K2316.35, y la ausencia de una aproximación de los tres pares restantes sugieren un cambio conformacional importante en TM5 y TM6, separándose mutuamente y del resto de las TMs. Las distancias menores en el sistema CX3CR1-FKN^{G2S2} coinciden con una activación incipiente del receptor, de manera semejante al receptor con el glicano G2. El comportamiento del sistema con la aglicona es análogo en cuanto a la tendencia, pero presentó una distancia aproximadamente 0.2 nm mayor entre el par R127^{3.50}-Y293^{7.53} (Figura 5.17A). Estos fenómenos, aunados con la interpretación de los modos principales y las proyecciones esenciales, cualifican a las tres isoformas de CX3CR1 como ligandos con actividades tipo agonista; probablemente pertenecientes a colectivos convergentes de estado funcional, o al menos estrechamente relacionados. Resalta la extensión de YXT hacia CT. También se visualiza el segmento helicoidal ICL2 en CX3CR1-FKN^{G2S2} (Figura 5-18 B y C). Sin embargo, si este plegamiento es inherente a la activación del receptor CX3CR1, o si es una variante del colectivo, se puede esclarecer al aumentar el tiempo de simulación y al realizar estudios adicionales con las proteínas G. Una diferencia importante entre los sistemas con AFKN y FKN^{G2S2} es la distribución de los ángulos diedros para el residuo Y293 (Figura 5-18D), que se aproximan a ser mutuamente excluyentes, y también contrastando con los sistemas CX3CR1 y CX3CR1-FKN^{G2}, que comparten ambos rotámeros, y adicionalmente éste último tiene una distribución

cuatrimodal. Las distribuciones histográficas del RMSD del dominio TMD respecto a la distancia TM1-TM6 ubica los tres sistemas con ligando en una región claramente separada del receptor CX3CR1, con valores altos de RMSD en congruencia con los movimientos principales observados de TM5 y TM6; y donde CX3CR1-FKN^{G2S2} tiende a crecer en esta característica (Figura 5-18E).

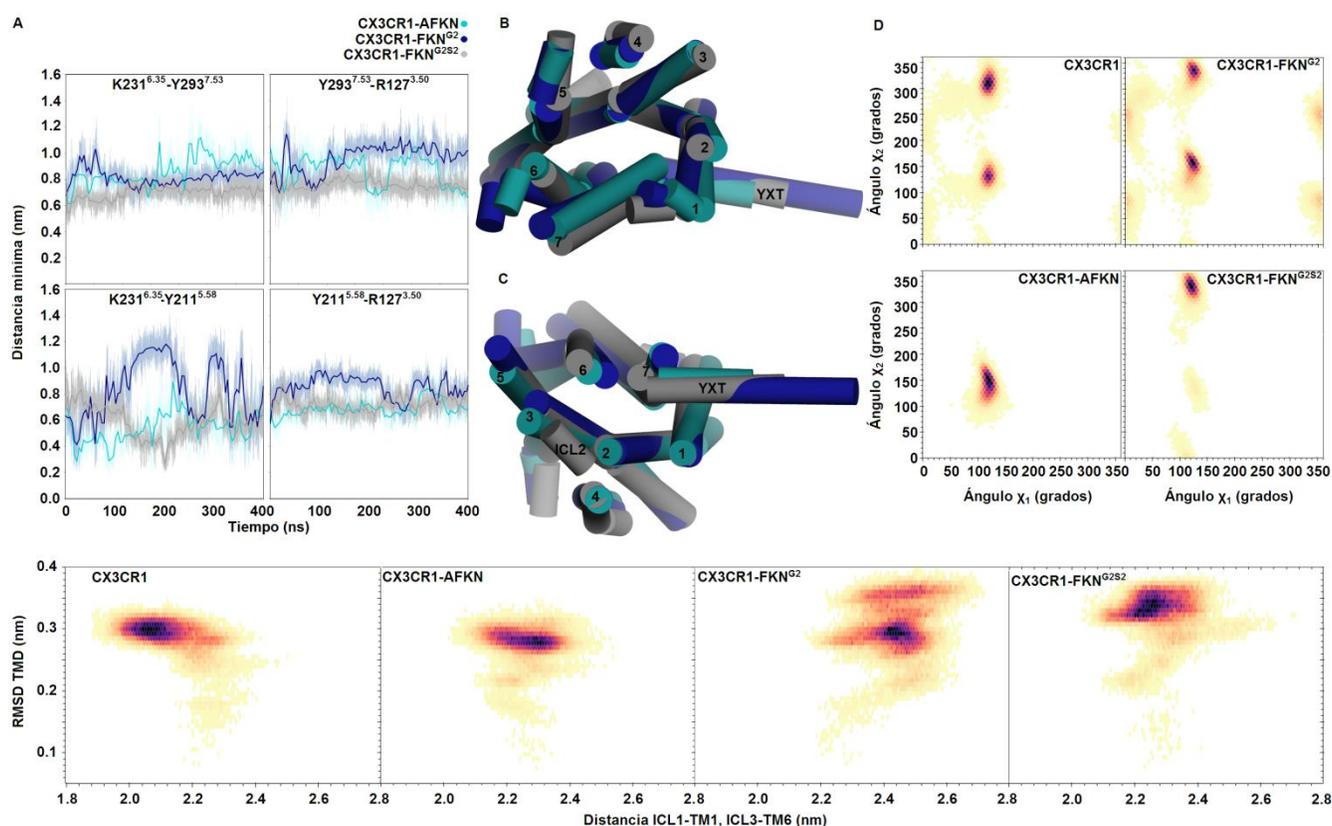


Figura 5.18. Características de los residuos de la compuerta iónica. (A) Distancias entre la tétroda de residuos. (B) Superposición de los tres sistemas con ligando peptídico, desde una vista extracelular y (C) intracelular. (D) Distribuciones histográficas de los ángulos diedros χ_1 y χ_2 de Y293^{7.53}. (E) Histogramas bivariados de RMSD del dominio TMD respecto a la distancia intracelular de separación TM1-TM6.

V.3. Análisis de los estados funcionales de CX3CR1 y los efectos de los ligandos

V.3.1. Patrones de estado funcional

El análisis de los once sistemas de CX3CR1 permitió la caracterización de los estados funcionales y los ligandos. La proyección de las trayectorias en la matriz de autovectores del modelo multitemplado permitió una agrupación de los sistemas como estativos, estativos en transición, transitivos, activos-transitivos y activos (Figura 5.19A). El ligando antagonista E6130 induce el estado quiescente en un estado no activo, mientras que los ligandos LG1 y MRV favorecen un estado quiescente en un estado funcional de transición. El ligando

en particular en el entorno oxidante que circunda las placas seniles y las neuronas afectadas por las lesiones de la EA, y el segundo en la región que se separa en la transición y la activación de la funcionalidad. Los residuos cercanos entre sí, F114^{3.37}, F118^{3.41} e I120^{3.43}, también se relacionaron de manera sobresaliente, entre ellos y con D16 y V143 en ICL2, implicación que podría significar un mecanismo de oclusión del poro si se considera que son residuos hidrofóbicos y voluminosos.

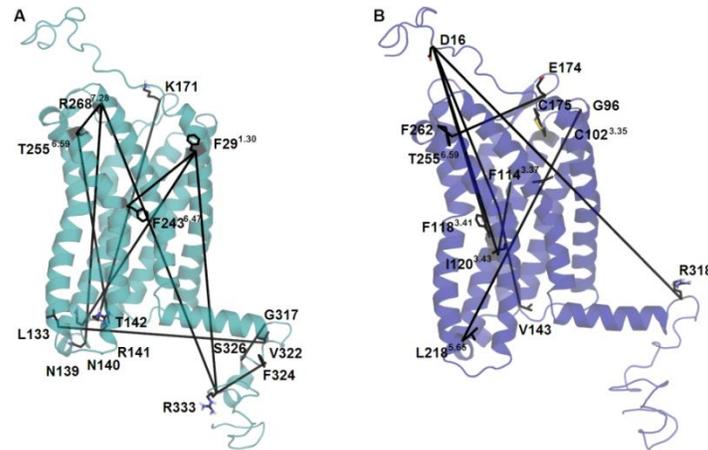


Figura 5.20. Información mutua torsional para los sistemas (A) CX3CR1-E6130 y (B) CX3CR1-FKN^{G2S2}. Los valores de información mutua fueron mayores a 0.35 para cada sistema.

Con los hallazgos estructurales y dinámicos de los once sistemas, se propone un mecanismo de activación para el receptor de quimiocina CX3CR1. A partir del estado funcional correspondiente a un conformero del colectivo basal o pasivo, los segmentos TM1-ICL1 y TM5-ICL3-TM6 se separan del resto de las demás TMs usando las divisiones (GX2G)^{1.46-1.49}, y (GX3P)^{5.46-5.50}; siendo más acentuado el movimiento en el segundo motivo, y generando la característica flexión intracelular del estado transitivo, que corresponde con el ingreso del CT de la proteína G subunidad α .^{38,59,68} Adicionalmente, el desplazamiento de TM3-ICL2 hacia la compuerta iónica favorece el plegamiento helicoidal en ICL2 (Figura 5.21A). Posteriormente, la aproximación de TM1-ICL1 hacia las TMs induce una flexión de YXT hacia fuera del bulto hidrofóbico membranal y una separación relativa de NT-TM1 (Figura 5.21B). Finalmente, el distanciamiento de NT-TM1 permite el cambio de inclinación de TM2 y su acercamiento al sitio ortostérico, además de un movimiento endomembranal de YXT hacia la región de colas hidrocarbonadas en la membrana, así como un movimiento de TM5-ICL3-TM6 para cerrar la compuerta iónica, característica del estado activo (Figura 5.21C).

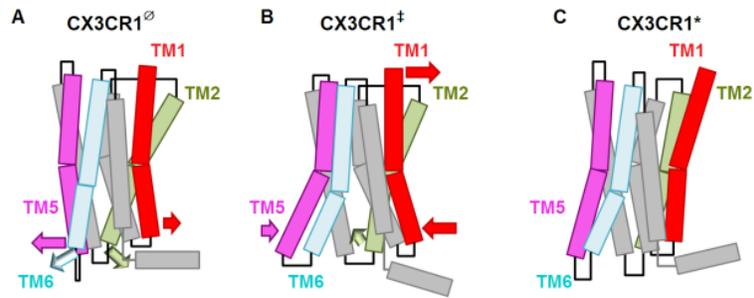


Figura 5.21. Mecanismo de activación propuesto para el receptor CX3CR1. Las flechas indican los desplazamientos principales en cada colectivo funcional.

V.4. Modelos de receptor-proteína Gi

V.4.1. Subunidad CX3CR1 del complejo

El receptor en complejo con el heterotrímero Gi, unido a GDP, presentó varias características nuevas, ausentes en los sistemas anteriores (Figura 5.22). El sistema G-CX3CR1-MRV tuvo una gran accesibilidad a desplazamientos en el subespacio esencial bidimensional, al contrario del sistema CX3CR1-MRV. El evidente ensanchamiento del receptor en este sistema es congruente con una apertura aparente del poro acuoso, característico de un estado no activo de acuerdo con el trabajo de Yuan, *et al.*³⁵ El desplazamiento del extremo intracelular de TM7 es también un componente nuevo del receptor.

Las distancias entre residuos de la tétada en la compuerta iónica confirman el ensanchamiento y el estado al que se dirige el sistema G-CX3CR1-MRV, al presentar separaciones máximas en los cuatro pares (Figura 5-23). De una manera característica, el par Y293^{7.53}-R127^{3.50} tiene separación mayor en el complejo con maraviroc, intermedia en el complejo sin ligando y menor en G-CX3CR1-AFKN. Debido al tiempo de simulación, y de la escala de tiempo reportada para las interacciones evidentes de los complejo G-R, la caracterización de un flujo acuoso no se alcanzó a observar en estos sistemas (Figura 5.24).

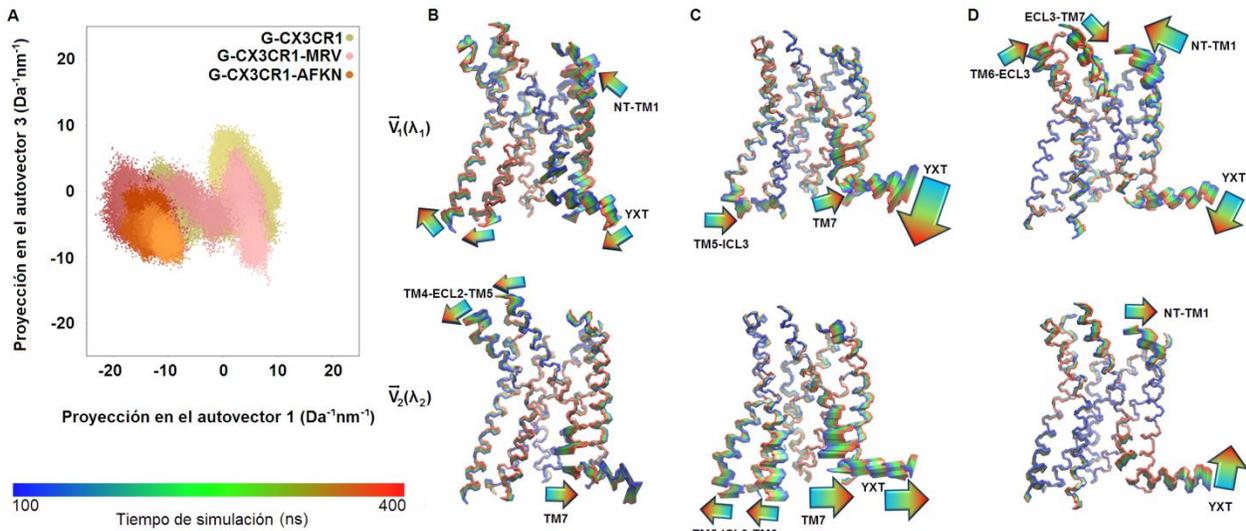


Figura 5.22. Análisis de componentes principales de CX3CR1 en los complejos con proteína G. (A) Proyección de las trayectorias sobre el sistema CX3CR1. El sistema G-CX3CR1-MRV tuvo un acceso amplio a locaciones en el subespacio esencial bidimensional, a comparación de G-CX3CR1-AFKN y semejante a G-CX3CR1. Aunque estos sistemas no son comparables con los sistemas sin proteína G, el comportamiento de G-CX3CR1-mrvH revela hallazgos novedosos. Los modos principales de (B) G-CX3CR1, (C) G-CX3CR1-MRV y (D) G-CX3CR1-AFKN muestran también el movimiento de separación de TM7, la hélice que contiene el motivo conservado NPXXY, Y293^{7,53}. Adicionalmente, el receptor en G-CX3CR1-MRV experimentó un ensanchamiento evidente, y la separación de los segmentos intracelulares TM5-ICL3-TM6, TM7 y YXT, inducido claramente por el heterotrímero G.

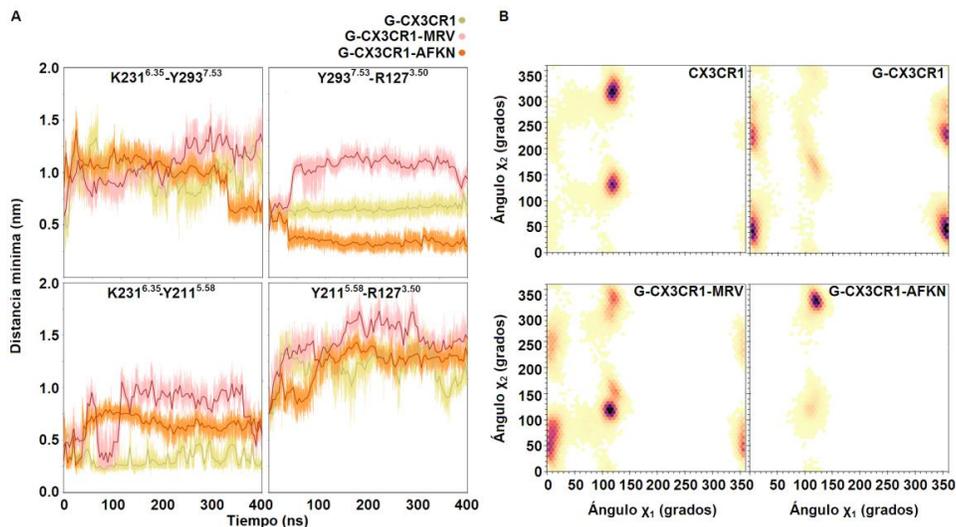


Figura 5.23. Características de la compuerta iónica en los sistemas complejos con proteína Gi. (A) Distancias de la tétrada R127^{3,50}, Y211^{5,58}, K231^{6,35} y R293^{7,53}. (B) Distribuciones histográficas de los ángulos diedros χ_1 y χ_2 de R293^{7,53}. Estas distribuciones no son comparables con los sistemas sin proteína G, pero son claramente distintos entre sí. El complejo G-CX3CR1 presentó una distribución análoga al sistema CX3CR1[‡].

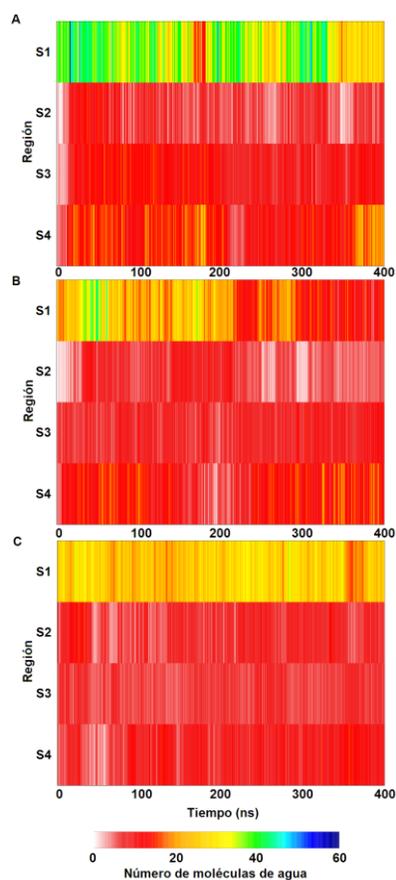


Figura 5.24. Moléculas de agua a través del poro acuoso interhelicoidal para (A) G-CX3CR1, (B) G-CX3CR1-MRV y (C) G-CX3CR1-AFKN.

V.4.2. Heterotrímero Gi

La activación de un GPCR en estado basal mediante su interacción con un agonista, o de manera probabilística al visitar los respectivos microestados asociados, transduce la señal extracelular hacia algunos de los heterotrímeros de proteínas G. Las proteínas G están constituidas por una subunidad $G\alpha$ (GNA), una GTPasa anclada a la membrana por dos acilaciones; una subunidad $G\beta$ (GNB) con estructura beta propela; y una pequeña subunidad $G\gamma$ (GNG) S-prenilada en el carboxilo terminal. Los tipos descritos de proteínas G que interactúan con CX3CR1 son G_i y G_z ,⁶⁹ siendo más abundante y conocido el heterotrímero G_i . El mecanismo de activación de los heterotrímeros G se ha descrito detalladamente de manera experimental,^{40,41,70,71} y se reportado la presencia de complejos preformados previo a la activación del sistema y el reclutamiento del heterotrímero hacia el receptor activo.⁶⁸ El mecanismo de activación de la subunidad $G\alpha$ implica la separación de dos dominios llamados α -helicoidal (GAH) y tipo Ras (Gras), en un movimiento tipo bisagra. Se ha demostrado que la presencia de un antagonista reduce la probabilidad y espontaneidad de la proteína G para este evento. Debido a la prolongada escala de tiempo necesaria para inducir y transducir la activación, de microsegundos a milisegundos, únicamente se estudiaron los complejos G_i -CX3CR1, G_i -CX3CR1-MRV y G_i -CX3CR1-AFKN.

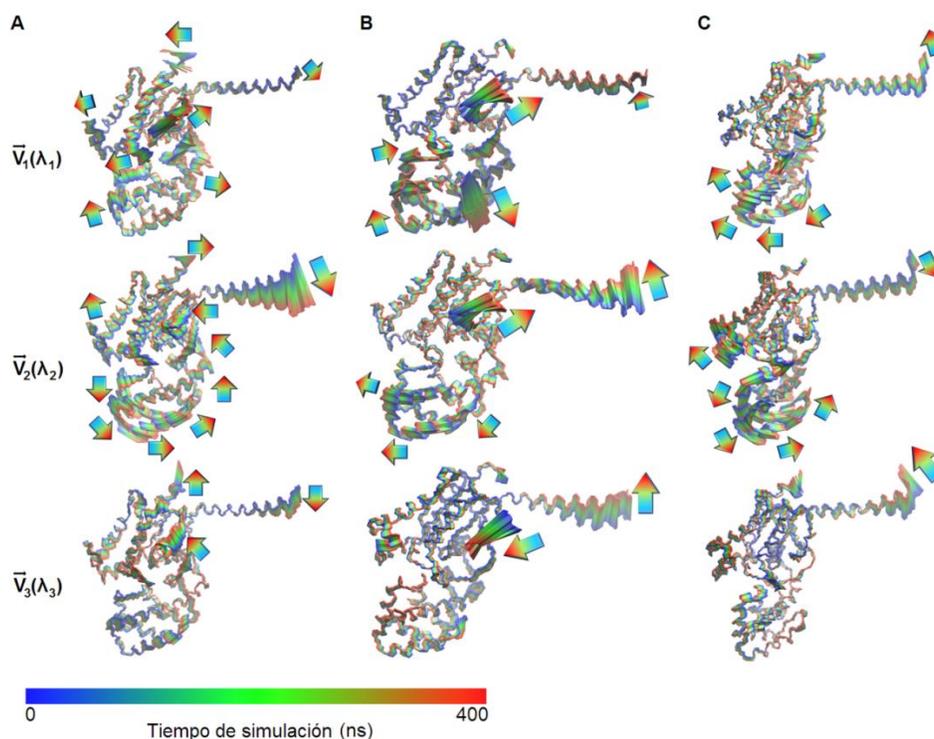


Figura 5.25. Análisis de componentes principales de la subunidad G α i para el sistema (A) G-CX3CR1, (B) G-CX3CR1-AFKN y (C) G-CX3CR1-MRV. En el sistema G-CX3CR1, G α i presenta un retraimiento de la hélice CT (59% de varianza) y una tendencia del dominio AH a separarse de Aras (13% de varianza), aunque no se llevó a cabo este desplazamiento. En G-CX3CR1-AFKN, complejo del agonista, la hélice CT se aproxima a la región intracelular del receptor de manera predominante (61%), y el dominio AH no presentó componentes relevantes del movimiento de bisagra en los 400 ns de simulación. En G-CX3CR1-MRV los desplazamientos de la hélice CT aproximándose al receptor (30%) y alejándose de él (15%) no son comparables con el sistema sin proteína G, pero es congruente con la actividad de un antagonista, en cuanto que se presentan las dos características a diferencia del sistema con AFKN, donde sólo predomina el ascenso de la hélice CT.

VI. CONCLUSIONES, OBSERVACIONES IMPORTANTES Y PERSPECTIVAS

El modelo propuesto para el receptor CX3CR1 fue adecuado en su geometría, para la caracterización estructural y dinámica del modelo y de los diferentes estados funcionales, así como sus interacciones con los ligandos antagonistas y agonistas evaluados. Los hallazgos con los ligandos, con los modelos funcionales, así como la comparación entre sistemas y estados condujeron a una propuesta del mecanismo de activación del receptor, que involucra la separación de los extremos intracelulares de TM1, TM5 y TM6, y el reacomodo de los segmentos extracelulares de TM1 y TM2. Asimismo, se identificaron algunas generalidades y particularidades de los estados funcionales del GPCR, como son un poro acuoso abierto y una proximidad entre hélices, que caracterizan al estado estativo; un reacomodo importante de las regiones intracelulares TM1-ICL2, TM5-ICL3-TM6 y TM3-ICL2 principalmente, que describen un estado transitivo; y la oclusión del poro acuoso y un rearreglo del sitio ortostérico, notable para el estado activo. Estos hallazgos permiten predecir varias características de los sistemas de estudio.

Al tratarse de sistemas complejos y con múltiples variables, el estudio preciso para caracterizar e identificar los eventos relacionados con los estados y sus cambios requieren mayor tiempo de simulación, en escalas de microsegundos a milisegundos. Sin embargo, con los tiempos de simulación evaluados en este trabajo se consiguieron obtener características suficientes para proponer una descripción de la dinámica funcional de este receptor, con información creciente sobre su influencia en neuroinflamación. Al integrar todos los resultados de este estudio, se propone al receptor CX3CR1 como un blanco terapéutico potencial para el tratamiento de la enfermedad de Alzheimer. Asimismo, se propone una comparación exhaustiva de este receptor con la estructura experimental cuando se encuentre disponible para su estudio.

VII. REFERENCIAS

- (1) Kumar, A.; Singh, A.; Ekavali. *Pharmacol. Reports* **2015**, 67 (2), 195–203.
- (2) Thériault, P.; Elali, A.; Rivest, S. *Alzheimers. Res. Ther.* **2015**, 7 (41), 1–10.
- (3) Tellechea, P.; Pujol, N.; Esteve-Belloch, P.; Echeveste, B.; García-Eulate, M. R.; Arbizu, J.; Riverol, M. *Neurologia* **2015**, 799, 1–10.
- (4) Arendt, T.; Stieler, J. T.; Holzer, M. *Brain Res. Bull.* **2016**, 126, 238–292.
- (5) Streit, W. J.; Mrak, R. E.; Griffin, W. S. T. *J. Neuroinflammation* **2004**, 4, 1–4.
- (6) Jäkel, S.; Dimou, L. *Front. Cell. Neurosci.* **2017**, 11 (24), 1–17.
- (7) Khurgel, M.; Koo, A.; Ivy, G. *Glia* **1996**, 16, 351–358.
- (8) Cui, W.; Allen, N.; Skynner, M.; Gusterson, B.; Clark, A. *Glia* **2001**, 34, 272–282.
- (9) Cahoy, J.; Emery, B.; Kaushal, A.; Foo, L.; Zamanian, J.; Christopherson, K. *J Neurosci* **2008**, 28, 264–278.
- (10) Schreiner, B.; Romanelli, E.; Liberski, P.; Ingold-Heppner, B.; Sobottka-Brillout, B.; Hartwig, T. *Cell Rep* **2015**, 12, 1377–1384.
- (11) Takamiya, Y.; Kohsaka, S.; Toya, S.; Otani, M.; Tsukada, Y. *Brain Res* **1988**, 466, 201–210.
- (12) Hirzel, A. Abcam.
- (13) Dimou, L.; Götz, M. *Physiol. Rev* **2014**, 94, 709–737.
- (14) Rivers, L.; Yong, K.; Rizzi, M.; Jamen, F.; Psachoulia, K.; Wade, A. *Nat Neurosci* **2008**, 11, 1392–1401.
- (15) Psachoulia, K.; Jamen, F.; Young, K.; Richardson, W. *Neuron Glia Biol* **2009**, 5, 57–67.
- (16) Simon, C.; Götz, M.; Dimou, L. *Glia* **2011**, 59, 869–881.
- (17) Bergles, D.; Roberts, J.; Somogyi, P.; Jahr, C. *Nature* **2000**, 405, 187–191.
- (18) Djogo, T.; Robins, S.; Schneider, S.; Kryszkaya, D.; Liu, X.; Mingay, A. *Cell Metab* **2016**, 23, 797–810.
- (19) McKenzie, I.; Ohayon, D.; Li, H.; de Faria, J.; Emery, B.; Tohyama, K. *Science (80-.)*. **2014**, 346, 318–322.
- (20) Schneider, S.; Gruart, A.; Grade, S.; Zhang, Y.; Kroger, S.; Kirchhoff, F. *Glia* **2016**, 64, 2201–2218.
- (21) Hirzel, A. Abcam.
- (22) Wake, H.; Moorhouse, A.; Nabekura, J. *Neuron Glia Biol.* **2011**, 7 (1), 47–53.
- (23) Hopkins, S. J.; Rothwell, N. J. *Trends* **1995**, 18, 83–88.
- (24) Karlström, S.; Nordvall, G.; Sohn, D.; Hettman, A.; Turek, D.; Åhlin, K.; Kers, A.; Claesson, M.; Slivo, C.; Lo-Alfredsson, Y.; Petersson, C.; Bessidskaia, G.; Svensson, P. H.; Rein, T.; Jerning, E.; Malmberg, Å.; Ahlgen, C.; Ray, C.; Vares, L.; Ivanov, V.; Johansson, R. *J. Med. Chem.* **2013**, 56, 3177–3190.
- (25) Bazan, J. F.; Bacon, K. B.; Hardiman, G.; Wang, W.; Soo, K.; Rossi, D.; Greaves, D.; Zlotnik, A.; Schall, T. J. *Lett. to Nat.* **1997**, 385, 640–644.
- (26) Wolf, Y.; Yona, S.; Kim, K.; Jung, S. *Front. Cell. Neurosci.* **2013**, 7 (26), 1–9.
- (27) Merino, J.; Muñetón-Gómez, V.; Álvarez, M.; Toledano-Díaz, A. *Curr. Alzheimer Res.* **2016**, 13 (4), 403–412.

- (28) Fuhrmann, M.; Bittner, T.; Jung, C. K. E.; Burgold, S.; Page, R. M.; Mitteregger, G.; Haass, C.; Laferla, F. M.; Kretschmar, H. *Nat. Neurosci.* **2010**, *13* (4), 411–413.
- (29) Constanzi, S.; Wang, K. In *G Protein-Coupled Receptors - Modeling and Simulation*; Filizola, M., Ed.; New York, 2014; p 228.
- (30) Venkatakrisnan, A. J.; Flock, T.; Este, D.; Oates, M. E.; Gough, J.; Babu, M. M. *Curr. Opin. Struct. Biol.* **2014**, *27*, 129–137.
- (31) Latorraca, N. R.; Venkatakrisnan, A. J.; Dror, R. O. *Chem. Rev.* **2017**, *117*, 139–155.
- (32) Isberg, V.; Mordalski, S.; Munk, C.; Rataj, K.; Harpsøe, K.; Hauser, A. S.; Vroiling, B.; Bojarski, A. J.; Vriend, G.; Gloriam, E. *Nucleic Acids Res.* **2017**, *44*, 356–364.
- (33) Katritch, V.; Cherezov, V.; Stevens, R. *Annu. Rev. Pharmacol. Toxicol.* **2013**, *53*, 531–556.
- (34) Liu, W.; Chun, E.; Thompson, A. A.; Chubukov, P.; Xu, F.; Katritch, V.; Han, G. W.; Roth, C. B.; Heitman, L. H.; IJzerman, A. P.; Cherezov, V.; Stevens, R. C. *Science (80-)*. **2012**, *337* (6091), 232–236.
- (35) Yuan, S.; Filipek, S.; Palczewsky, K.; Vogel, H. *Nat. Commun.* **2014**, *5* (4733), 1–10.
- (36) Flanagan, C. A. In *Pharmacology & Therapeutics of Constitutively Active Receptors*; Elsevier Inc., 2014; Vol. 70, pp 215–263.
- (37) Mccudden, C. R.; Hains, M. D.; Kimple, R. J.; Siderovski, D. P.; Willard, F. S. *Cell. Mol. Life Sci.* **2005**, *62*, 551–577.
- (38) Rasmussen, S. G. F.; Devree, B. T.; Zou, Y.; Kruse, A. C.; Chung, K. Y.; Kobilka, T. S.; Thian, F. S.; Chae, P. S.; Pardon, E.; Calinski, D.; Mathiesen, J. M.; Shah, S. T. A.; Lyons, J. A. *Nature* **2011**, *477* (7366), 549–555.
- (39) Alexander, N. S.; Preininger, A. M.; Kaya, A. I.; Stein, R. A.; Hamm, H. E.; Meiler, J. *Nat. Publ. Gr.* **2013**, *21* (1), 56–63.
- (40) Gilman, A. G. *Annu. Rev. Biochem.* **1987**, *56*, 615–649.
- (41) Galés, C.; van Durm, J. J. J.; Schaak, S.; Pontier, S.; Percherancier, Y.; Paris, H.; Audet, M.; Bouvier, M. *Nat. Struct. Mol. Biol.* **2006**, *13* (9), 778–786.
- (42) Blumenthal, D.; Garrison, J. In *Las Bases Farmacológicas de la Terapéutica de Goodman & Gilman*; Brunton, L., Chabner, B., Knollmann, B., Eds.; México, D.F., 2012; pp 41–71.
- (43) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235–242.
- (44) Lomize, M. A.; Lomize, A. L.; Pogozheva, I. D.; Mosberg, H. I. *Bioinformatics* **2006**, *22* (5), 623–625.
- (45) Jo, S.; Kim, T.; Iyer, V. G.; Im, W. *J. Comput. Chem.* **2008**, *29*, 1859–1865.
- (46) Lee, J.; Cheng, X.; Swails, J. M.; Yeom, M. S.; Eastman, P. K.; Lemkul, J. A.; Wei, S.; Buckner, J.; Jeong, J. C.; Qi, Y.; Jo, S.; Klauda, J. B.; Widmalm, G.; Im, W. *J. Chem. Theory Comput.* **2016**, *12*, 405–413.
- (47) Klauda, J. B.; Venable, R. M.; Freites, J. A.; O'Connor, J. W.; Tobias, D. J.; Mondragon-Ramirez, C.; Vorobyov, I.; MacKerell Jr., A. D.; Pastor, R. W. *J Phys Chem B.* **2010**, *114* (23), 7830–7843.
- (48) Venable, R. M.; Sodt, A. J.; Rogaski, B.; Rui, H.; Hatcher, E.; MacKerell Jr., A. D.; Pastor, R. W.; Klauda,

- J. B. *Biophys J.* **2016**, *107* (1), 134–145.
- (49) van der Spoel, D.; Lindahl, B.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. *J. Comp. Chem.* **2005**, *26*, 1701–1719.
- (50) Jo, S.; Im, W. *Nucleic Acids Res.* **2013**, *41* (D1), D470–D474.
- (51) Gowers, R. J.; Linke, M.; Barnoud, J.; Reddy, T. J. E.; Melo, M. N.; Seyler, S. L.; Dotson, D. L.; Buchoux, S.; Kenney, I. M.; Beckstein, O. *Proc. Python Sci. Conf.* **2016**, No. Scipy, 98–105.
- (52) Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. *J. Comput. Chem.* **2011**, *32*, 2319–2327.
- (53) Grant, B. J.; Rodrigues, A. P. C.; Elsayy, K. M.; Mccammon, J. A.; Caves, L. S. D. *Bioinformatics* **2006**, *22* (21), 2695–2696.
- (54) Skjærven, L.; Yao, X.; Scarabelli, G.; Grant, B. J. *Bioinformatics* **2014**, *15*, 399.
- (55) R-Development-Core-Team. R Foundation for Statistical Computing: Vienna, Austria 2008.
- (56) McClendon, C. L.; Friedland, G.; Mobley, D. L.; Jacobson, M. P. *J Chem Theory Comput* **2010**, *5* (9), 2486–2502.
- (57) Paramo, T.; East, A.; Garzón, D.; Ulmschneider, M. B.; Bond, P. J. *J. Chem. Theory Comput.* **2014**, *10* (5), 2151–2164.
- (58) Tan, Q.; Zhu, Y.; Li, J.; Chen, Z.; Han, G. W.; Kufareva, I.; Li, T.; Ma, L.; Fenalti, G.; Li, J.; Zhang, W.; Xie, X.; Yang, H.; Jiang, H.; Cherezov, V.; Liu, H.; Stevens, R. C.; Zhao, Q.; Wu, B. *Science (80-)*. **2013**, *341* (6152), 1387–1390.
- (59) Rosenbaum, D. M.; Rasmussen, S. G. F.; Kobilka, B. K. *Nature* **2009**, *459* (7245), 356–363.
- (60) Ballesteros, J. A.; Weinstein, H. In *Methods in Neurosciences*; Sealfon, S. C., Ed.; Academic Press, 1995; pp 366–428.
- (61) Wakita, H.; Yanagawa, T.; Kuboi, Y.; Imai, T. *Mol. Pharmacol.* **2017**, *92* (5), 502–509.
- (62) Cederblad, L.; Rosengren, B.; Ryberg, E.; Hermansson, N. *Biochem. J.* **2016**, *473*, 641–649.
- (63) Burg, J. S.; Ingram, J. R.; Venkatakrishnan, A. J.; Jude, K. M.; Dukkupati, A.; Feinberg, E. N.; Angelini, A.; Waghray, D.; Dror, R. O.; Ploegh, H. L.; Garcia, C. *Science (80-)*. **2015**, *347* (6226), 1113–1117.
- (64) Unverzagt, C.; André, S.; Seifert, J.; Kojima, S.; Fink, C.; Srikrishna, G.; Freeze, H.; Kayser, K.; Gabius, H.-J. *J. Med. Chem.* **2002**, *45* (2), 478–491.
- (65) Wang, P.; Wang, H.; Gai, J.; Tian, X.; Zhang, X.; Lv, Y.; Jian, Y. *Sci. Rep.* **2017**, *7* (40301), 1–13.
- (66) Dror, R.; Arlow, D. H.; Borhani, D. W.; Jensen, M. Ø.; Piana, S.; Shaw, D. E. *Proc. Natl. Acad. Sci.* **2009**, *106* (12), 4689–4694.
- (67) Dror, R. O.; Arlow, D. H.; Maragakis, P.; Mildorf, T. J.; Pan, A. C.; Xu, H.; Borhani, D. W.; Shaw, D. E. *Proc Natl Acad Sci USA* **2011**, *108* (46), 18684–18689.
- (68) Galés, C.; Rebois, V.; Hogue, M.; Trieu, P.; Breit, A.; Hébert, T. E.; Bouvier, M. *Nat. Methods* **2005**, *2*, 177–184.
- (69) Casey, J.; Fongq, K. W.; Simonlj, I.; Gilman, G. *J. Biol. Chem.* **1990**, *265* (4), 2383–2390.
- (70) Gilman, A. G.; Rall, T. W. *J. Biol. Chem.* **1968**, *243* (22), 5872–5881.

- (71) Gilman, A. G. *Proc. Natl. Acad. Sci.* **1970**, 67 (1), 305–312.
- (72) NCBI. PubChem database <https://www.pubchem.ncbi.nlm.nih.gov/>.
- (73) Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; Thompson, J. D.; Higgins, D. G.; McWilliam, H.; Remmert, M.; So, J. *Mol. Syst. Biol.* **2011**, 7 (539), 1–6.
- (74) Kyte, J.; Doolittle, R. F.; Diego, S.; Jolla, L. *J. Mol. Biol.* **1982**, 157, 105–132.
- (75) Eisenberg, D.; Weiss, R. M.; Terwilliger, T. C. *Nature* **1982**, 299, 371–374.
- (76) Services, B. EMBOSS.
- (77) Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M. R.; Appel, R. D.; Bairoch, A. In *The Proteomics Protocols Handbook*; Walker, J. M., Ed.; Humana Press, 2005; pp 571–607.
- (78) Kelley, L. A.; Mezulis, S.; Yates, C. M.; Wass, M. N.; Sternberg, M. J. E. *Nat. Protoc.* **2015**, 10 (6), 845–858.
- (79) Tsirigos, K. D.; Peters, C.; Shu, N.; Lukas, K.; Elofsson, A. *Nucleic Acids Res.* **2015**, 43 (Web Server issue), 401–407.
- (80) Zhang, Y. *BMC Bioinformatics* **2008**, 9 (40), 1–8.
- (81) Šali, A.; Blundell, T. L. *J. Mol. Biol.* **1993**, 234, 779–815.
- (82) Laskowski, R. A. *J. Appl. Crystallogr.* **1993**, 26, 283–291.
- (83) Lomize, A. L.; Pogozheva, I. D.; Lomize, M. A.; Mosberg, H. I. *Protein Sci.* **2006**, 15, 1318–1333.
- (84) Ng, H. W.; Laughton, C. A.; Doughty, S. W. *J. Chem. Inf. Model.* **2014**, 54 (2), 573–581.
- (85) Umehara, H.; Bloom, E. T.; Okazaki, T.; Domae, N.; Imai, T. *TRENDS Immunol.* **2001**, 22 (22), 602–607.
- (86) Perry, V. H.; Holmes, C. *Nat. Rev. Neurol.* **2014**, 10, 217–224.
- (87) Consortium, T. U. *Nucleic Acids Res.* **2019**, 47 (D1), D506–D515.
- (88) Wasmuth, E. V.; Lima, C. D. *Nucleic Acids Res.* **2016**, 45 (November 2016), 1–12.
- (89) Jazar, R. N. In *Advanced Dynamics, Rigid Body, Multibody, and Aerospace Applications*; John Wiley & Sons, Inc.: Hoboken, 2011; pp 887–932.
- (90) Abraham, M.; Hess, B.; van der Spoel, D.; Lindahl, E. The GROMACS development teams at the Royal Institute of Technology and Uppsala University: Uppsala 2014.
- (91) Rabal-Gracia, O. Tesis doctoral: Herramientas de cribado virtual aplicadas a inhibidores de tirosina quinasas. Contribución al desarrollo del programa PRALINS para el diseño de quimiotecas combinatorias.
- (92) Jordaan, J. A.; Živanović, R. *Measurement* **2006**, 39, 451–457.
- (93) Maxwell, J. C. *Philosophical Transactions of the Royal Society of London*. Londres 1867, pp 49–88.
- (94) Haaser, N. B.; LaSalle, J. P.; Sullivan, J. A. In *Análisis matemático 1*; Editorial Trillas, S.A.I: México, D.F., 1970.
- (95) Landau, L.; Lifshitz, E. In *Curso Abreviado de Física Teórica: Mecánica y Electrodinámica*; MIR: Moscú, 1971; p 57,58.
- (96) Darden, T.; York, D.; Pedersen, L.; Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, 98 (10089), 1–5.

-
- (97) Essmann, U.; Perera, L.; Darden, M. L. B.; Lee, H.; Pedersen, L. G.; Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103* (8577), 8577–8593.
- (98) Kolmogorov, A. N.; Fomin, S. V. In *Elementos de la teoría de funciones y del análisis funcional*; MIR: Moscú, 1972; p 130,131.
- (99) Aleksandrovich, Y.; Syromyatnikov, V. N. In *Phase Transitions and Crystal Symmetry*; Springer Science & Business Media, 2012; p 12.
- (100) Reifenberg, R. Inter-Molecular Forces:Keesom Force <https://www.physics.purdue.edu/~reifebr>.
- (101) Gross, J.; Engineering, C. P.; Ag, B.; Gic, P. Q. *Am. Inst. Chem. Eng. J.* **2005**, *51* (9), 2556–2568.
- (102) van Oss, C. J. *J. Mol. Recognit.* **2003**, *16*, 177–190.
- (103) Schlick, T. *Molecular Modeling and Simulation*; Springer: New York, 2002.
- (104) Vikas; Sharma, M.; Rajput, S. *Mol. Simul.* **2007**, *33* (12), 1017–1022.
- (105) Schrödinger, L. Schrödinger: Scientific leader in life sciences and materials research <https://www.schrodinger.com/>.

VIII. ANEXO I: Resultados suplementarios

VIII.1. Receptor de referencia CCR5

VIII.1.1. Simulación MD de CCR5

Los gráficos de desviaciones para cada posición (Figura 8.1) muestran que los receptores tienen comportamientos diferentes en la cadena principal y en las cadenas laterales (Figura 8.1A). El sistema con el ligando protonado tuvo una fluctuación muy alta en el segmento ECL2, en contacto estrecho con MRV; fenómeno que no se observó en el sistema con maraviroc desprotonado (Figura 8.1B). De manera comparativa, al analizar la diferencia de RMSF del sistema CCR5-MRV respecto a CCR5, el segmento ICL2 tiene fluctuaciones mayores debido a la pérdida de estructura helicoidal típica de los receptores en estados no activos (Sección V.2.2, Figura 5.9). Adicionalmente la región ICL3, interactuante con proteína G, tuvo cambios escasos respecto al sistema CCR5 (Figura 8.1C).

El receptor en el sistema CCR5 mostró una compacidad similar al sistema CCR5-MRV al encontrarse en un estado inicial estativo que no cambió funcionalmente de manera notable. En contraste, el sistema CCR5-mrv experimentó una compactación mayor a los otros dos sistemas, contrastando con las características de un receptor en complejo con un antagonista, por lo que se deduce que este sistema es atípico debido al estado de protonación del maraviroc. Además de la estructura experimental protonada,⁵⁸ el ligando tendría asociada una abundancia relativa esperada muy baja debido a su valor de pKa en medio acuoso ($pK_a=7.3$).⁷² Ante esta característica, el ligando que se localiza en el sitio de unión ortostérico y próximo a moléculas de agua circundantes, y que presenta interacciones importantes con el residuo E283 a través de la subestructura de tropinio (Figura 8.2) se consideró que el estado del ligando como mrv sin protonar es de especiación improbable en el entorno extracelular.

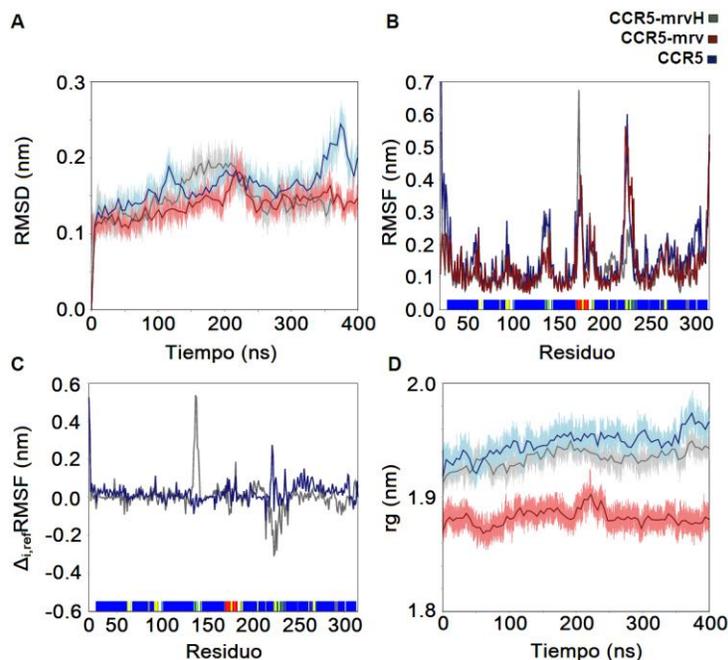


Figura 8.1. Desviaciones del receptor de referencia en los tres sistemas CCR5-MRV, CCR5-mrv y CCR5. (A) Gráfico de RMSD de la cadena principal a través del tiempo de simulación. (B) Gráfico de RMSF de las cadenas laterales. (C) Diferencia de RMSF respecto al aporreceptor CCR5. (D) Gráfico de radios de giro a través del tiempo.

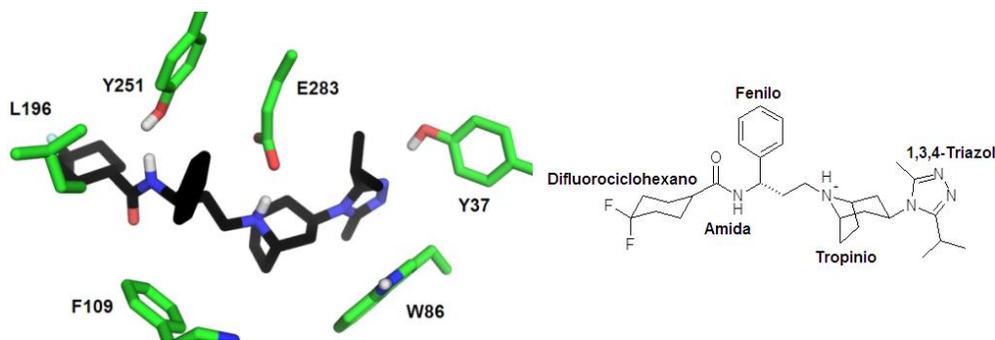


Figura 8.2. Interacción del ligando MRV con el residuo E283 en el sitio ortostérico. La interacción iónica principal se lleva a cabo entre E283 y el subgrupo tropinio del ligando. Otras interacciones relevantes incluyen los puentes de hidrógeno entre Y37 y el subgrupo 1,3,4-triazol de MRV, y entre Y251 y el grupo amida del maraviroc. También se reconocen las interacciones hidrofóbicas entre L196 y el anillo de difluorociclohexano, de F109 y el grupo fenilo, y entre W86 y el biciclo de tropinio.

VIII.2. Receptor de prueba CX3CR1

VIII.2.1. Modelado de CX3CR1, CX3CR1 \emptyset , CX3CR1 $\#$ y CX3CR1*

Similitud con estructuras experimentales

Las proteínas con mayor similitud al receptor CX3CR1 en estructura primaria (Tabla 8.1) forman parte también del subgrupo de ligandos peptídicos de la superfamilia de GPCRs clase A.

Tabla 8.1. Similitud primaria entre CX3CR1 y otros receptores de clase A.

Proteína	Descripción	Similitud con CX3CR1
CCR4	Receptor principal para CCL4/MIP-1 (del inglés <i>macrophage inflammatory protein 1</i>).	42.3%
CCR2	Receptor principal para CCL2/MCP-1 (del inglés <i>monocyte chemoattractant protein 1</i>).	38.1%
CCR5	Receptor principal para CCL5/RANTES (del inglés <i>Regulated upon Activation, Normal T Cell Expressed and Secreted</i>). Correceptor de la proteína de envoltura gp120 de HIV-1 con tropismo hacia macrófagos y células linfoides CD4+ (variante viral R5).	37.6%
US28	Receptor de herpesvirus 5 para varias quimiocinas.	34.6%
CXCR4	Receptor principal para CXCL4/SDF-1 (del inglés <i>stromal-derived factor 1</i>). Correceptor de la proteína de envoltura gp120 de HIV-1 con tropismo principal hacia células linfoides CD4+ (variante viral X4).	31.9%

VIII.2.2. Análisis de la estructura primaria

La similitud entre estructuras primarias se realiza con los algoritmos de Clustal Omega⁷³ y es útil para seleccionar estructuras molde y comparativas (Figura 8.3).

CCR5	1	MDYQVSSPIYDINY-YTSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMLVILILINCK	59
CX3C1	1	MDQFPESVTENFEYDDLAECYIGDIVVFGTVFLSIFYSVIFAIGLVGNLLVVFALTNK	60
		** . * : : * : * * : : : : * : * : * : * : * : * : * : * : * : * : *	
CCR5	60	RLKSMTDIYLLNLAISDLFLLVFPWAHYAAAQWDFGNTMCQLLTGLYIFGFFSGIFFI	119
CX3C1	61	KPKSVTDIYLLNALSDDLFPVATLFPWTHYLINKEGLHNMCKFTTAFFIFGFFSGIFFI	120
		: * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : *	
CCR5	120	ILLTIDRYLAVVHAVFALKARTVTFGVVTSVITWVAVFASLP6IIFTRSQKEGLHYTCS	179
CX3C1	121	TVISIDRYLAIVLAANSNNRTVQHGVTISLGVWAAAILVAAPQFMFTKQKE---NECL	176
		: : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : *	
CCR5	180	SHFPYSQYQFQWKNFQTLKIVILGLVPLLVVICYSYGILKTLRLCRNEKRRHRAVRLIFT	239
CX3C1	177	GDYPEVLQEIWVLRNVETNLFGLFLLPLLINSYCFRIIQTFLFCKNHKK-AKAIKLILL	235
		.. * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : *	
CCR5	240	IMIVYFLWAPYNIWLLNFTQEFFGLNCCSSNRDQAMQVTELTGMTHCCINPIIYAF	299
CX3C1	236	VVIVFFLWTPYNYMIFLETCLKLYDFFPSCDMRKDLRLALSVTETVAFSHCCLNPLIYAF	295
		: * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : *	
CCR5	300	VGEKFRNYLLVFFQKHIAKRFCKCSI-FQQEAPF--RASS-----VYTRSTGEQEISV	350
CX3C1	296	AGEKFRNYLYHLYGKCLAVLCGRSVHVDFFSSSESQRSRHGSLVSSNFTYHTSDGDALLL	355
		. * * * * . * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : *	
CCR5	351	GL	352
CX3C1	356	--	355

Figura 8.3. Alineamiento de estructura primaria de CCR5 y CX3CR1, con 136 posiciones idénticas y 126 posiciones similares. La similitud es de 37.6 %.

VIII.2.3. Análisis de estructura secundaria

Índice de hidropatía

El índice de hidropatía (hpi) o índice de afinidad por agua es el valor de la diferencia de energía libre de transferencia de un medio acuoso al estado vapor, y escalado por un factor lineal (Ecuaciones 8.1A y 8.1B).⁷⁴

$$\text{hpi} \propto \Delta_{\text{aq} \rightarrow \text{v}} G_{\text{res}}^0 \quad (8.1A)$$

$$\text{hpi} = \left(\frac{\text{mol}}{\text{kJ}} \right) \Delta_{\text{aq} \rightarrow \text{v}} G_{\text{res}}^0 + \frac{\text{mol}}{\text{kJ}} \quad (8.1B)$$

Momento hidrofóbico

El momento hidrofóbico bidimensional μ_{H} (Ecuación 8.2A) para un motivo helicoidal es el vector definido como la serie de la diferencia de la posición del átomo CA respecto al punto medio de la hélice, escalada con las diferencias de energía libre de transferencia de un medio apolar a un medio polar. El cálculo de μ_{H} se realiza para un intervalo definido de residuos, con centro en el residuo de interés, y para el cual se asocia la media aritmética $\langle \mu_{\text{H}} \rangle$ (Ecuación 8.2B).⁷⁵

$$\vec{\mu}_{\text{H},\text{res}} = \vec{r}_{\text{res}} - (\nabla \times \{\vec{r}_{\text{hel}}\}|_0) \quad (8.2A)$$

$$\langle \vec{\mu}_{\text{H},\text{res}} \rangle = \frac{1}{n_{\text{res}}} \sum_{i=\text{res}-5}^{\text{res}+5} \sum_{j=\text{res}-5}^{\text{res}+5} \|\vec{r}_{ij} - (\nabla \times \{\vec{r}_{\text{hel}}\}|_0)\| \quad (8.2B)$$

La geometría del modelo se determinó al asignar los motivos de estructura secundaria como asas y hélices, y definir los segmentos amino terminal (NT), hélices transmembranales (TM1 a TM7), hélice yuxtamembranal (YXT), asas intracelulares (ICL1 a ICL3) y extracelulares (ECL1 a ECL3); y el extremo carboxilo terminal (CT). La estructura primaria de CX3CR1 obtenida de *UniProt* con 355 residuos de aminoácidos se analizó para detectar regiones de residuos con características hidrofóbicas a través de los índices de hidropatía y momento hidrofóbico en la región posterior a la séptima hélice transmembranal (Figuras 8.4 y 8.5).

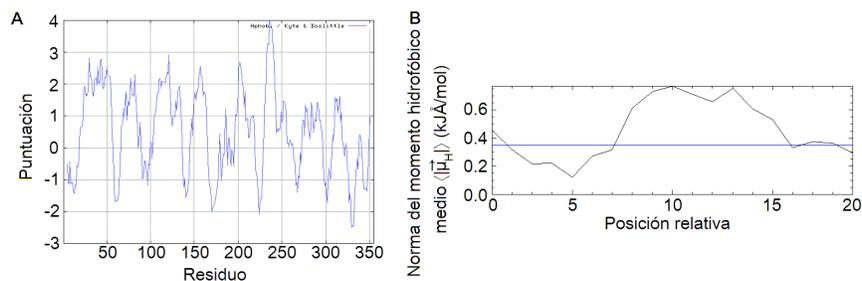


Figura 8.4. Análisis bidimensional de la estructura primaria de CX3CR1. (A) Gráfico de Kyte-Doolittle de índices de hidropatía.⁷⁶ B. Gráfico de la norma del momento hidrofóbico medio para el segmento [290,315].⁷⁷

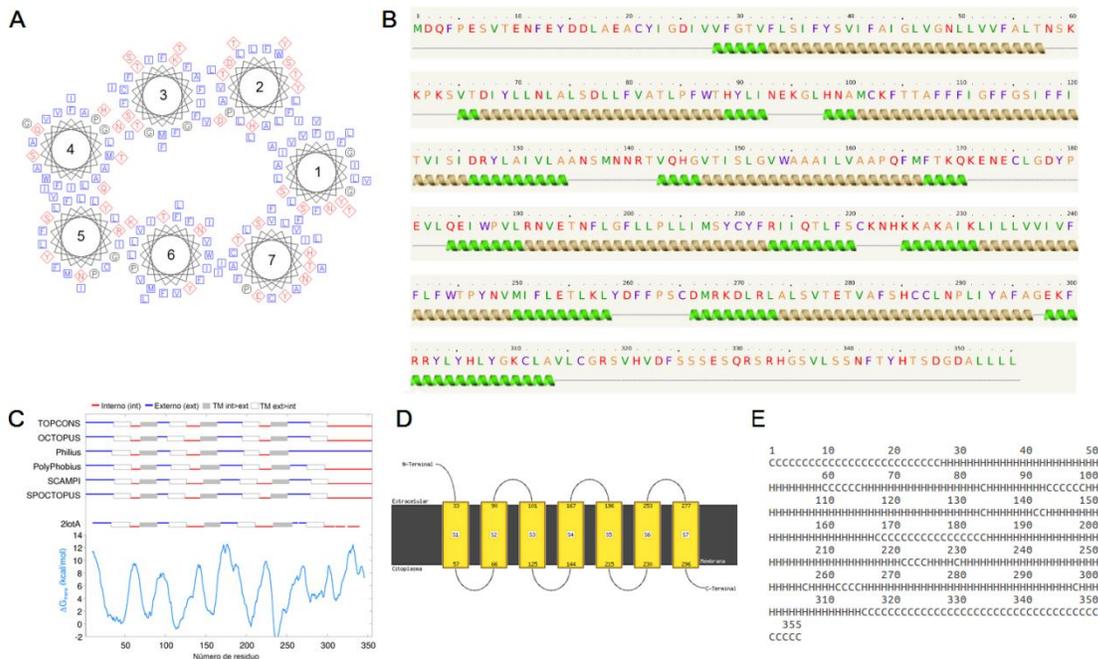


Figura 8.5. Topografía y estructura secundaria estimada para el receptor CX3CR1. (A) Diagramas de rueda que muestran la locación relativa estimada de los residuos de aminoácidos de la región helicoidal.⁷⁶ (B) Estructura secundaria predicha por el servidor Phyre2.⁷⁸ (C) Análisis predictivos de estructura primaria obtenidos con varios algoritmos.⁷⁹ (D) Esquema de geometría del receptor. (E) Estructura secundaria modal, C: asa, H: hélice.

VIII.2.4. Análisis de estructura terciaria

Enhebrado estructural

El servidor GPCR-I-TASSER (del inglés *Iterative Threading Assembly Refinement*)⁸⁰ emplea un método de predicción de estructura y función de proteínas. Identifica moldes de estructuras experimentales depositadas en PDB con el enhebrado LOMETS (*Local Meta-Threading Server*) genera modelos en \mathbb{R}^3 con puntuaciones altas. Además incluye el programa REMO (del inglés *reconstruct atomic model*) es una herramienta multifuncional para la optimización de los puentes de hidrógeno de la cadena principal.

Modelados de CX3CR1

Se realizaron tres grupos de modelos de estructura terciaria en los servidores GPCR-I-TASSER, I-TASSER⁸⁰ y con la aplicación MODELLER 9.17⁸¹ con cinco, cinco y diez estructuras respectivamente. Para la validación se analizaron los ángulos diedros ϕ y ψ y se graficaron los diagramas de Ramachandran (Figura 8.6), donde las regiones en color rojo son aquellas que corresponden a los valores más favorecidos de ϕ y ψ , las regiones de color amarillo representan torsiones favorecidas, las regiones color amarillo claro opaco son regiones de poco favorecimiento,

y las regiones color blanco no están favorecidas y se mencionan los residuos en color rojo. De cada grupo se escogió un modelo con los mejores estadísticos de validación. Los dos modelos multitemplado de los servidores (Figura 8.6 A y B) tuvieron un patrón similar de distribuciones de torsión; con múltiples valores atípicos. Sin embargo, estos modelos tuvieron los valores más bajos de RMSD respecto a CCR5 al comparar sus cadenas principales ajustadas (Tabla 8.2), siendo el modelo del servidor GPCR-I-TASSER el seleccionado para la comparación con los dos restantes. De este ajuste, se observó que los dos modelos de servidores son similares (Tabla 8.3), y la conformación del modelo de MODELLER 9.17 fue divergente aunque su diagrama de Ramachandran tuviera escasos valores atípicos (Figura 8.6C). Además, el segmento yuxtamembranal en este último modelo está carece del plegamiento esperado para un GPCR. Con este análisis, se eligió el modelo multitemplado que se realizó en el servidor GPCR-I-TASSER.

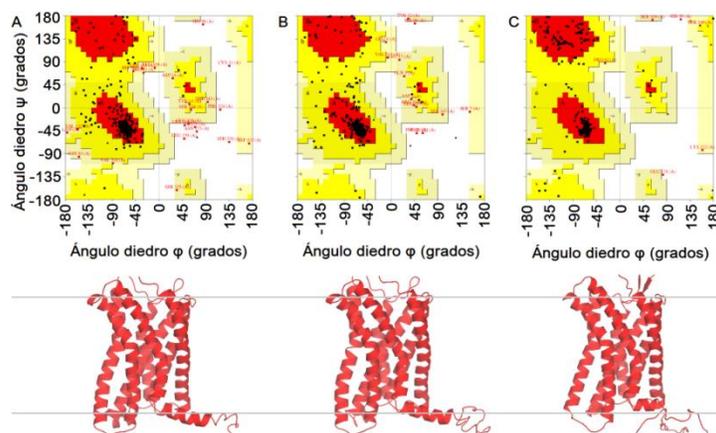


Figura 8.6. Diagramas de Ramachandran para tres modelos de CX3CR1 mostrando la topometría diédrica calculada con PROCHECK⁸² y representación de estructuras secundarias. (A) Modelo multitemplado de GPCR-I-TASSER con 97.2% de residuos en regiones favorecidas. (B) Modelo multitemplado de I-TASSER con 99.1% de residuos en regiones favorecidas. (C) Modelo mixto de MODELLER9.17 con 99.4% de residuos en regiones favorecidas.

Tabla 8.2. Valores de RMSD (nm) de la cadena principal del dominio transmembranal de los tres modelos multitemplado respecto a la región transmembranal del receptor de referencia CCR5.

GPCR-I-TASSER	I-TASSER	MODELLER9.17
0.39	0.48	0.61

Tabla 8.3. Relación de valores de RMSD (nm) de la cadena principal del dominio transmembranal para los tres modelos multitemplado.

	GPCR-I-TASSER	I-TASSER	MODELLER9.17
GPCR-I-TASSER	0		
I-TASSER	0.23	0	
MODELLER 9.17	0.71	0.72	0

La topometría del receptor define los segmentos geométricos con base en sus mediciones y cálculos estructurales y energéticos (Tabla 8.4). La topometría se obtuvo con el servidor PPM de *Orientations of Proteins in Membranes Database*.^{44,83}

Tabla 8.4. Topometría para el modelo multitemplado seleccionado, del servidor GPCR-I-TASSER.

Segmento	Residuos embebidos	Residuos transmembranales con estructura helicoidal
TM1	30-57	32-54
TM2	67-92	67-92
TM3	97, 99-125	100-125
TM4	145-168	146-167
TM5	186-212	187-212
TM6	231-257, 260-262	231-257
TM7	271, 273-296, 300	273-296

Notas: El grosor hidrofóbico es de 3.28 ± 0.070 nm, con un estimado de energía libre de transferencia de -69.8 kcal/mol y una inclinación relativa de $3^\circ \pm 2^\circ$.

VIII.3. Réplicas del modelo multitemplado

El modelo multitemplado se basó en CCR5 (PDB: 4MBS) y US28 (PDB: 4XT1). De dos réplicas simuladas para este sistema, se obtuvieron comportamientos distintos (Figura 8.7), teniendo la segunda réplica un comportamiento divergente a la primera réplica, utilizada como comparativo de los demás sistemas. Esto se debió probablemente a la salida del estado funcional en el que se encontraba inicialmente. Esta suposición se confirma con el confórmero medio (Figura 8.8A) y los modos principales de la segunda réplica (Figura 8.8B), particularmente la separación de la TM1 en la parte extracelular, lo que sugiere nuevamente que el receptor CX3CR1 posee actividad constitutiva o intrínseca.

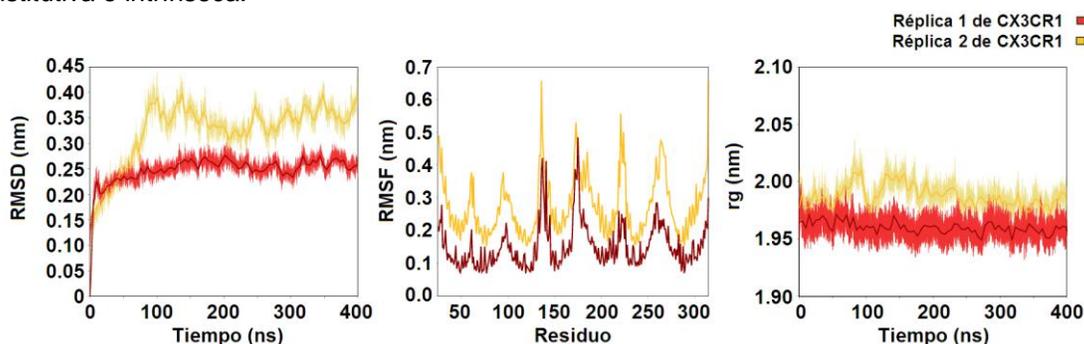


Figura 8.7. Comparación de las dos réplicas de simulación MD de CX3CR1. (A) Gráfico de RMSD de la cadena principal a través del tiempo. (B) Gráfico de RMSF por residuo. (C) Gráfico de radio de giro de la región transmembranal en el tiempo de simulación.

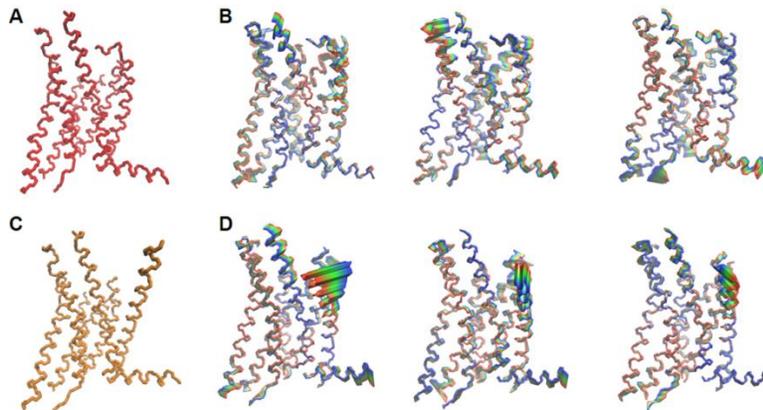


Figura 8.8. Estructuras notables de cada réplica de CX3CR1. (A y C) Estructuras promedio de la cadena principal para las réplicas 1 y 2 respectivamente. (B y D) Los tres primeros modos principales de las réplicas 1 y 2 respectivamente.

VIII.4. Sistemas funcionales CX3CR1 multitemplado, estativo, transitivo y activo

Los desplazamientos en los tres estados del receptor CX3CR1 tuvieron comportamientos diferentes. El modelo multitemplado CX3CR1 se mostró con movimientos distintivos de la cadena principal respecto a los otros dos sistemas y fluctuaciones mayores en ICL2 e ICL3, lo que sugiere que el estado funcional del modelo no corresponde con los estados activo y estativo aunque el modelado se basó principalmente en CCR2 y CCR5 en estado estativo. Sin embargo, los primeros cien nanosegundos de simulación MD permitieron a este sistema acceder a un estado de equilibrio mecánico que se prolongó hasta el final de la simulación (Figura 8.9A). Los tres sistemas tuvieron fluctuaciones altas en segmentos determinados: El estado de CX3CR1 en ICL2 e ICL3, el estado CX3CR1* en ECL1, ICL3 y ECL3, y el estado CX3CR1 en ECL1 y ECL3 (Figura 8.9B). Este hallazgo muestra que los integrantes de estos estados funcionales tienen conformaciones distintivas.

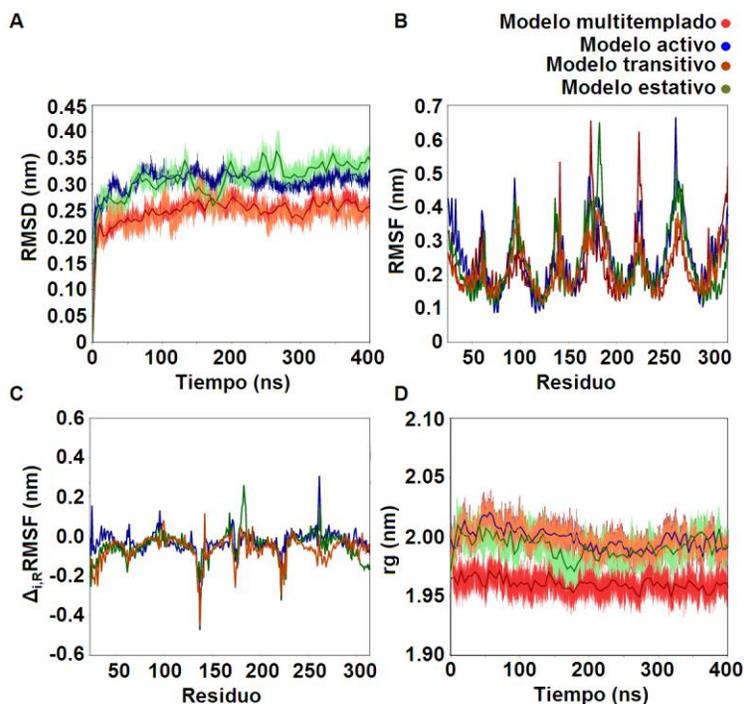


Figura 8.9. Desplazamientos del receptor de prueba CX3CR1 en los tres estados: basal, estativo, transitivo y activo.

Los segmentos helicoidales se determinaron con base en el análisis de helicidad (Figura 8.10). Los ángulos de inclinación para las hélices transmembranales TM1 a TM7, considerando las rupturas de helicidad por residuos Pro o Gly, y la hélice yuxtamembranal se calcularon a través del tiempo de simulación y sus densidades estimadas (Figura 8.11), mostrando algunas coincidencias importantes entre los sistemas.

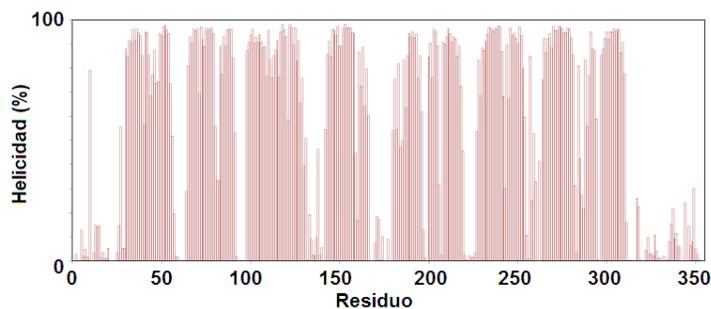


Figura 8.10. Porcentajes de helicidad para los residuos del modelo CX3CR1.

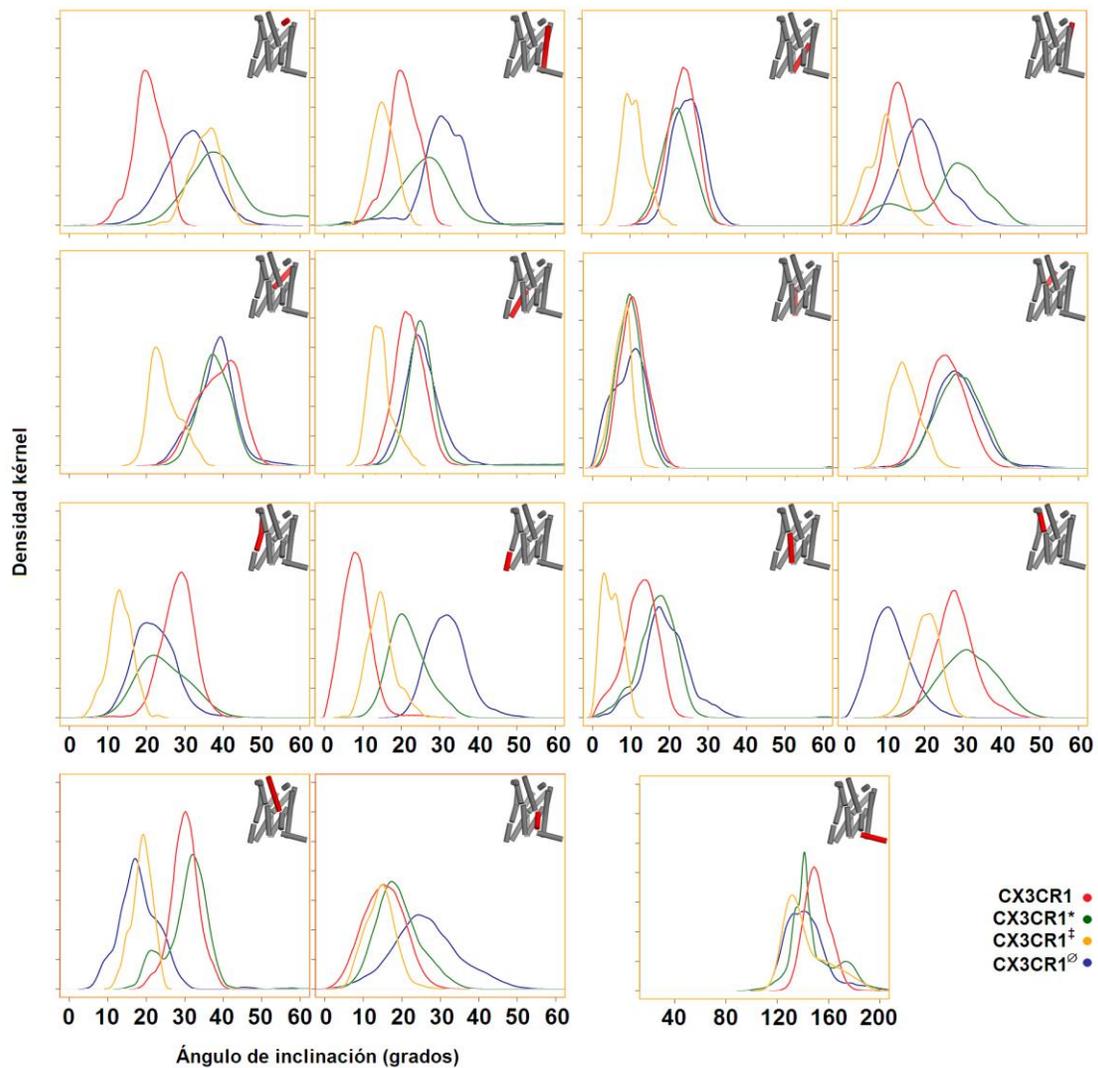


Figura 8.11. Ángulos de inclinación de los segmentos helicoidales para los sistemas multitemplado, activo, transitivo y estativo.

Los sistemas con antagonistas coincidieron en gran parte de las densidades angulares de inclinación, y con aproximación al sistema estativo CX3CR1 \emptyset . El segmento extracelular TM1 tiene densidades bajas en comparación con el sistema CX3CR1 \emptyset , dada la ausencia de ligando en éste. También en el segmento intracelular de TM5, la inclinación es baja, como se espera en los sistemas antagonizados; y la inclinación de YXT es predominantemente obtusa, y destacando la variabilidad en esta característica en los sistemas, en especial CX3CR1-AZD8797 y CX3CR1 \emptyset (Figura 8.12).

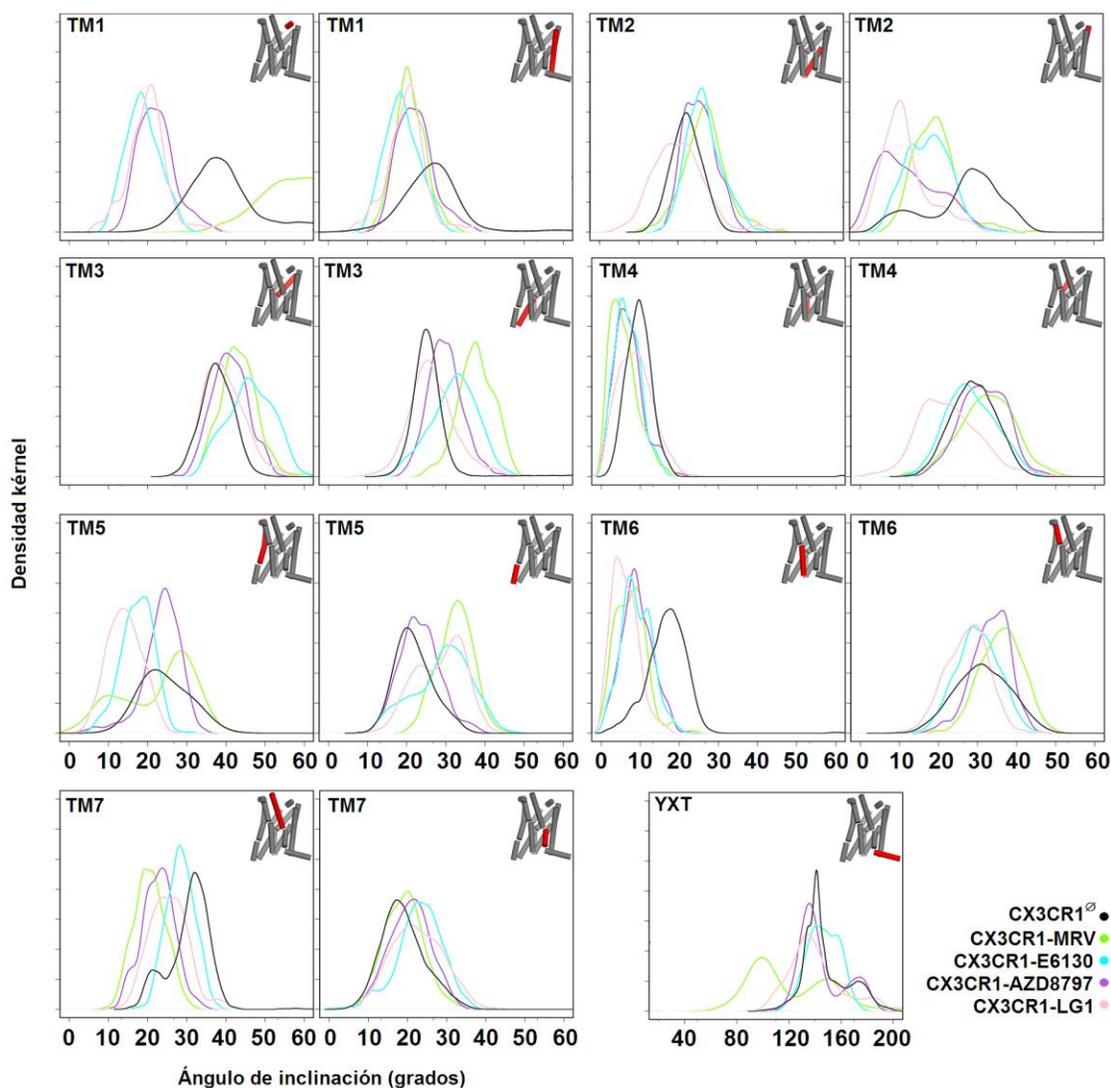


Figura 8.12. Ángulos de inclinación de los segmentos helicoidales para los sistemas estativos CX3CR1[∅], CX3CR1-MRV, CX3CR1-E6130, CX3CR1-AZD8797 y CX3CR1-LG1.

Para los sistemas con agonistas, las densidades angulares sólo coincidieron en su mayoría en el sistema CX3CR1-FKN^{G2S2} con el modelo activo CX3CR1*, confirmando la influencia del carbohidrato en la actividad del péptido agonista. En estos sistemas destaca el corrimiento de las modas angulares hacia valores más altos para ambos segmentos de TM1; el desplazamiento del segmento intracelular de TM5 hacia valores mayores. Los sistemas CX3CR1-AFKN y CX3CR1-FKN^{G2} tuvieron los comportamientos más extremos, como la amplia fluctuación angular del segmento extracelular de TM2, la escasa inclinación de TM3 y la pronunciada flexión intramembranal de YXT en CX3CR1-AFKN, mientras que la exagerada inclinación de las regiones

intracelulares de TM5 y TM6 plantean un probable cambio de estado para el sistema CX3CR1-FKN^{G2}, con una flexión de YXT extrema (Figura 8.13).

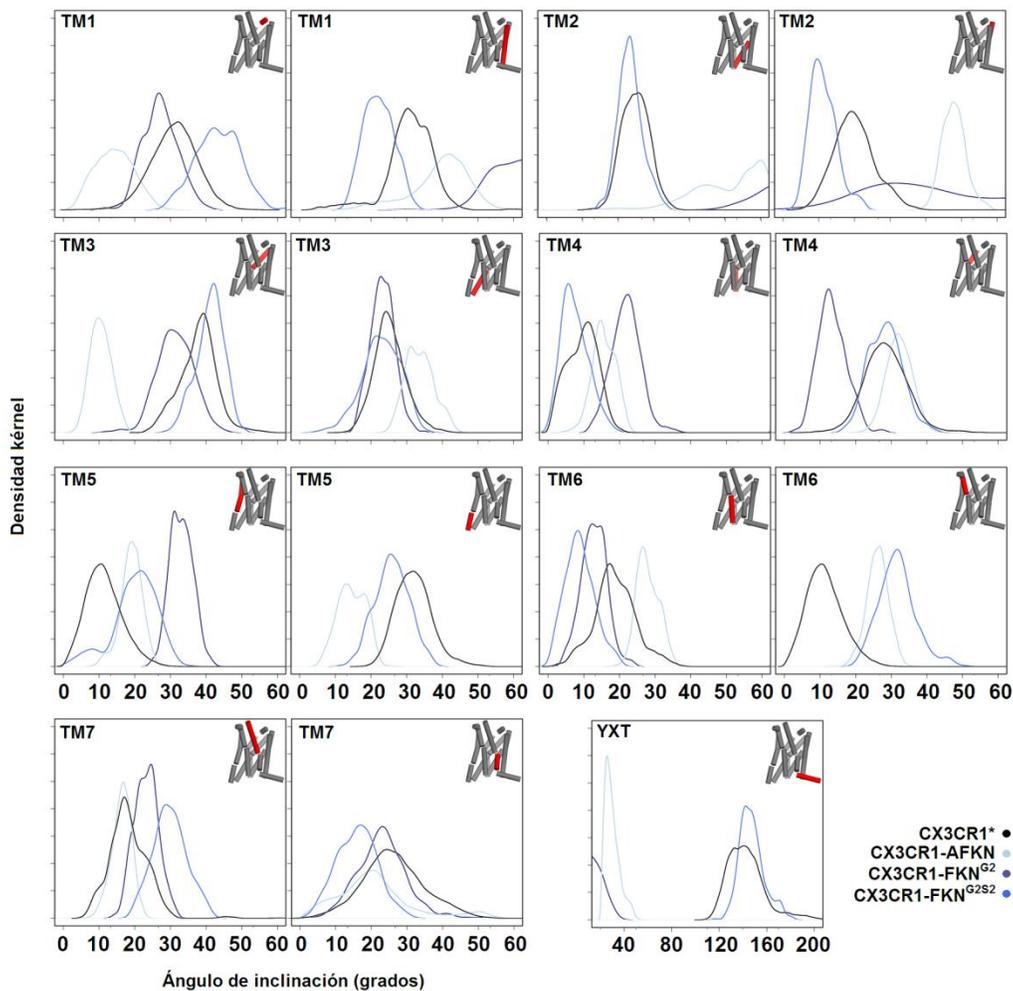


Figura 8.13. Ángulos de inclinación de las hélices de los sistemas CX3CR1*, CX3CR1-AFKN, CX3CR1-FKN^{G2} y CX3CR1-FKN^{G2S2}. Las distribuciones de ángulos comparadas con el modelo activo, donde el sistema CX3CR1-FKN^{G2S2} coincide en ocho segmentos, en particular los segmentos intracelulares de TM1 y TM5.

VIII.5. Sistemas membranales

Las membranas plasmáticas de las células son de composición heterogénea y presentan variaciones múltiples en La (2S)-3-O-palmitil-2-O-oleílfosfatidilcolina (POPC, Figura 8.14) o 3-palmitil-2-oleil-lecitina es uno de los fosfolípidos más abundantes en las membranas eucariotes. La anchura membranal de una bicapa de POPC es aproximadamente 3.9 a 4.0 nm y permite mayor movilidad en el subespacio esencial⁸⁴ (Figura 8.15).

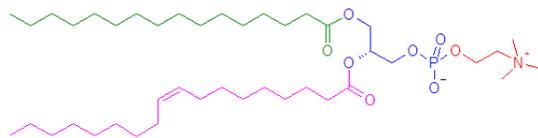


Figura 8.14. Estructura de POPC. Se muestran los residuos en color. Rojo: colina. Azul: ácido fosfatídico. Verde: ácido palmítico. Magenta: ácido oleico.

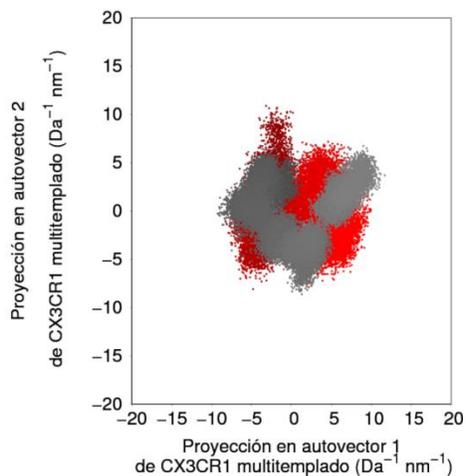


Figura 8.15. Espacio esencial bidimensional de los autovectores 1 y 3 para el receptor CX3CR1 en una membrana de POPC y una membrana multicomponente que incluye esfingomielinas NSM y LSM, y POPE.

VIII.6. Correlación cruzada

La correlación cruzada calculada para los átomos C α de los residuos de CX3CR1-E6130 y CX3CR1-FKN^{G2S2} no es abundante en desplazamientos de corto alcance, por ejemplo, entre residuos conectados o muy cercanos; y no permite la discriminación inicial de los movimientos concertados a distancia, aunque sus mapas coloreados muestran diferencias importantes (Figura 8.16).

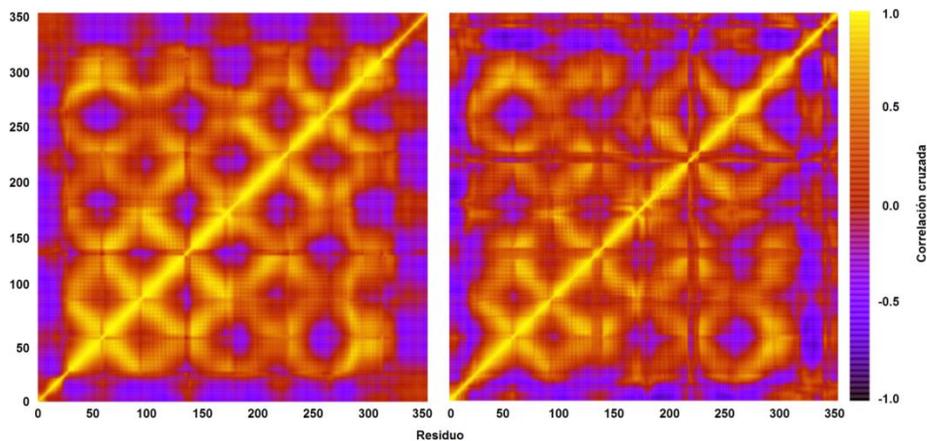


Figura 8.16. Matrices de correlación cruzada para los residuos en los sistemas (A) CX3CR1-E6130 y (B) CX3CR1-FKN^{G2S2}.

Por esta cantidad de datos que interfieren con la examinación de alosterismo, se optó por el cálculo de la información mutua torsional con el objetivo de esclarecer las relaciones presentes en los sistemas del receptor CX3CR1.

VIII.7. Simulación MD de receptor viral US28 con AFKN

El sistema testigo del receptor activo US28 en complejo con aglicofractalkina presentó varias características comunes con los sistemas activos, como el plegamiento beta en ECL2, la direccionalidad de los extremos intracelulares TM5 y TM6, así como una escasa presencia de moléculas de agua en el poro (Figura 8.17).

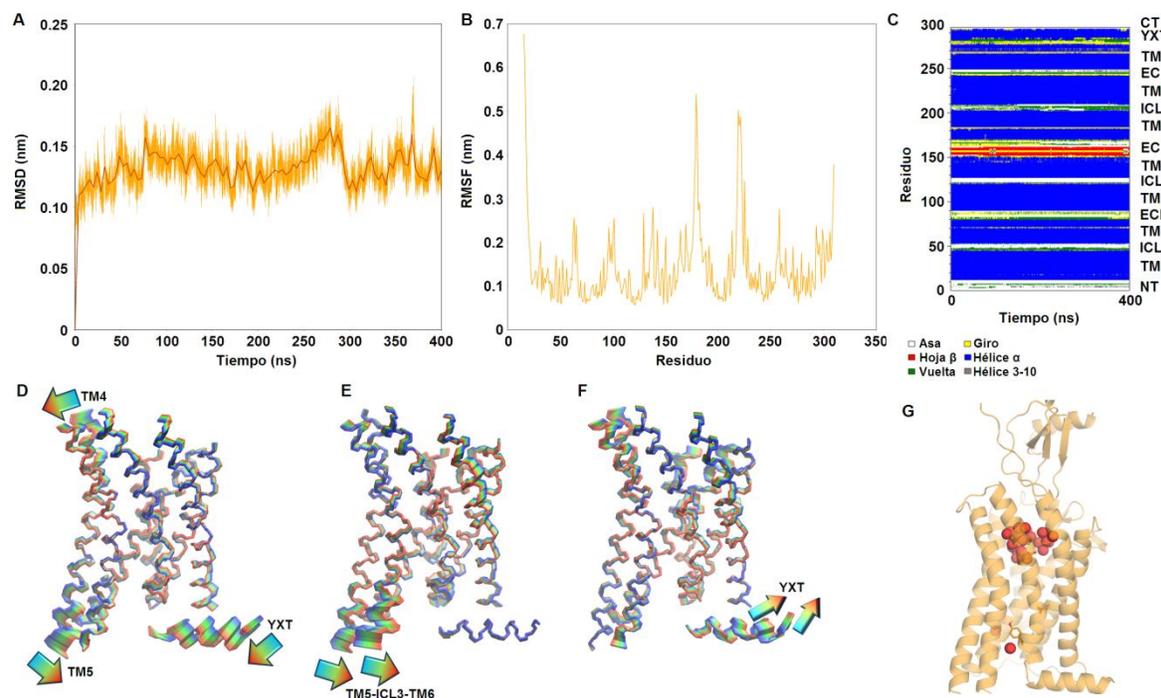


Figura 8.17. Características dinámicas del sistema US28*-AFKN. (A) RMSD de la cadena principal transmembranal. (B) RMSF de las cadenas laterales. (C) Estructura secundaria. (D-F) Modos principales correspondientes con los tres autovectores de mayor magnitud, respectivamente. (G) Representación acumulada de moléculas de agua en el poro interhelicoidal.

IX. ANEXO II: Información complementaria

IX.1. Fractalkina

La fractalkina o neurotactina (CX3CL1, FKN, NTT) es una integrina bitópica tipo I que está constituida por diferentes dominios estructurales y funcionales. Los dominios estructurales son el dominio extracelular, el dominio transmembranal y un pequeño dominio intracelular. Los dominios funcionales son el dominio de quimiocina (sCX3CL1) o fractalkina soluble, en el extremo amino, y el dominio tipo mucina de glicosilación múltiple (Figura 9.1A). La fractalkina membranal participa en las sinapsis inmunológicas con linfocitos, macrófagos y microglías interactuando con las integrinas $\alpha 4\beta 1$ y $\alpha v\beta 3$ en presencia o ausencia del receptor CX3CR1 (Figura 9.1B).

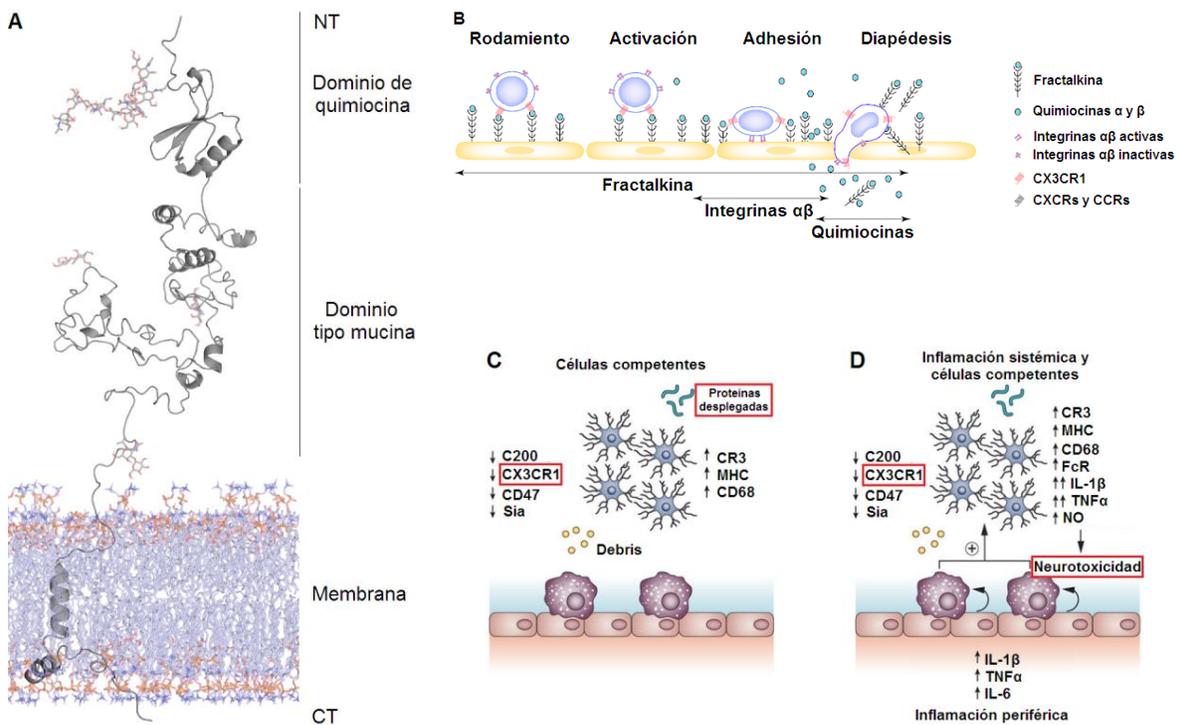


Figura 9.1. La quimiocina tipo δ transmembranal fractalkina (mCX3CL1). (A) Modelo de mCX3CL1 construido por enhebrado a partir del dominio de quimiocina (PDB: 4XT1 y 4XT3). La estructura de este dominio se conoce con detalle. (B) Funciones de CX3CL1 en la quimiotaxis y migración leucocitaria (Modificado de Umehara H., *et al*).⁸⁵ (C) Algunas funciones neuroinmunológicas de la fractalkina (Modificado de Perry V. H. & Holmes, C.).⁸⁶

IX.1.2. Carbohidrato biantenal complejo

Las glicosilaciones son modificaciones postraduccionales comunes en las proteínas y desempeñan funciones importantes en el reconocimiento de estado y naturaleza de una proteína o incluso de una célula. Las glicosiltransferasas catalizan estas modificaciones, y ocurren comúnmente en residuos de serina, treonina y asparagina. El dominio sCX3CL1 posee al menos una N-glicosilación en Asn33,⁸⁷ la cual puede variar dependiendo de numerosos factores. Los residuos de monosacáridos comúnmente encontrados son N-acetil-2-desoxi-2-D-glucosamina (GlcNAc), D-manosa (Man), D-galactosa (Gal), N-acetil-2-desoxi-2-D-galactosamina (GalNAc), L-fucosa (Fuc) y ácido siálico (Sia o NeuNAc). El motivo ramificado bis[Man-(α -1,3; α -1,6)]-Man-(α -1,4)-GlcNAc-(β -1,4)-GlcNAc-(β -1,ND1)-Asn es el oligosacárido base⁶⁴ (Figura 9.2A), a partir del cual se extiende la glicosilación. El carbohidrato biantenal complejo (Figura 9.2 B y C) consiste en el núcleo oligosacárido extendido en los dos residuos de manosa distales. Imagen de la fractalkina soluble (sCX3CL1), o dominio de quimiocina de CX3CL1 (CDC, Figura 9.2D).

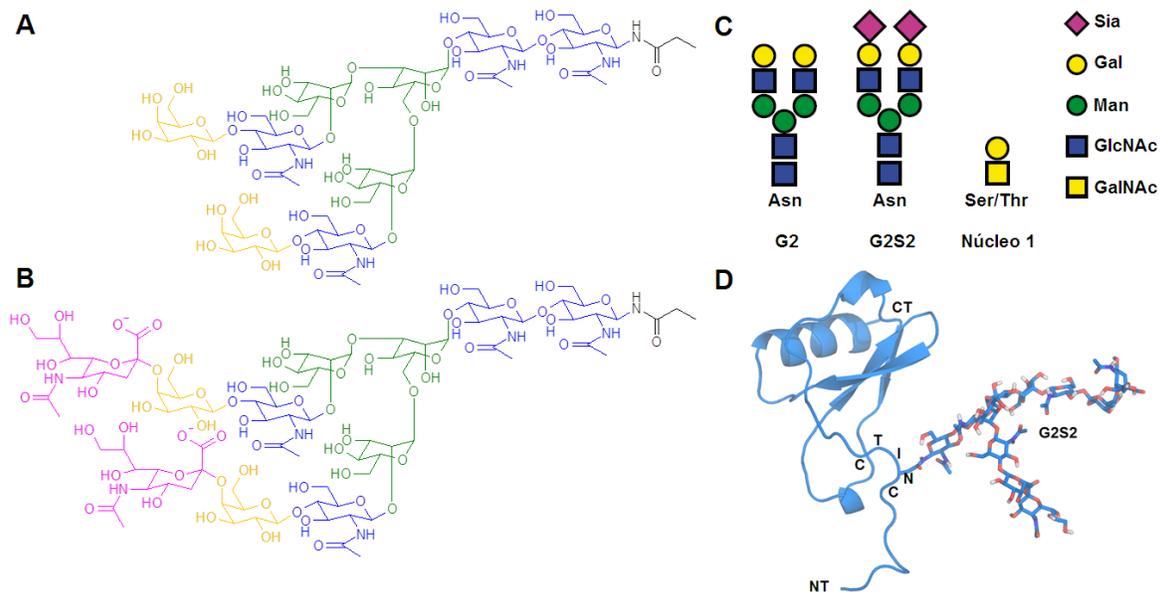


Figura 9.2. Glicosilación de fractalkina. (A) Estructura del glicano G2, modelado en el sistema CX3CR1-FKN^{G2}. (B) Estructura del glicano G2S2 del sistema CX3CR1-FKN^{G2S2}. (C) Esquema representativo de los glicanos. (D) Imagen de sCX3CL1 o CDC con el carbohidrato biantenal complejo.

IX.2. Receptor US28 y aglicofractalkina

El receptor homólogo de GPCR US28 es un receptor de β -quimiocina, tipo CC, de herpesvirus humano 5 (HHV5).⁸⁷ Sus ligandos incluyen a CCL2/MCP-1, CCL3/MIP-1- α , CCL5/RANTES (*Regulated upon Activation, Normal T Cell Expressed and Secreted*) y CX3CL1/FKN (Figura 9.3). Este receptor induce endocitosis independiente del ligando, y se asocia con proteínas de tipo $G_{q/11}$. Los receptores US28 y CX3CR1 comparten 34.615%⁷³ de similitud, con 126 residuos idénticos y 132 residuos similares; gran parte de ellos en las TMs.

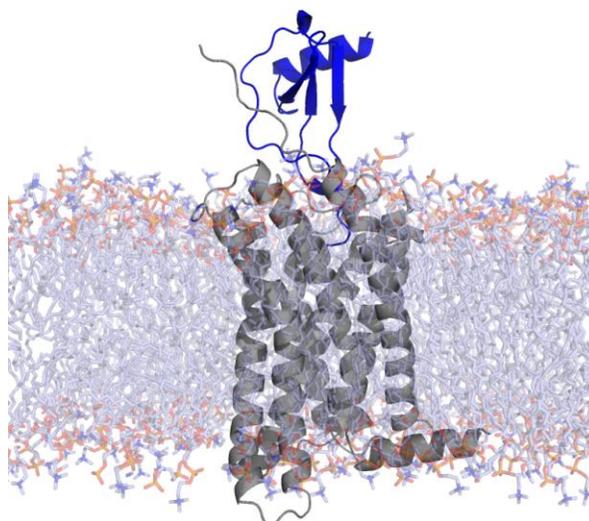


Figura 9.3. Estructura del complejo US28-AFKN en membrana, tomada de la estructura cristalográfica 4XT1 de 2.89 Å de resolución.⁶³

IX.3. Receptor CCR5 y maraviroc

El receptor CCR5 (también conocido como CD195, cúmulo de diferenciación 195; Figura 9.4) es un GPCR para β -quimiocinas como CCL3/MIP-1- α , CCL4/MIP-1- β y CCL5/RANTES.⁸⁷ El receptor CCR5 induce la quimiotaxia de linfocitos T de memoria (CD3+,CD4+,CD45RO+), linfocitos B (CD19+), células NK (CD16/56+), eosinófilos y basófilos. Adicionalmente, es uno de los dos correceptores descritos de la glicoproteína 120 (gp120) de envoltura del virus de inmunodeficiencia humana tipo 1 (HIV-1).⁵⁸ El tropismo de HIV-1 en células CD4+ se debe a que la gp120 reconoce e interactúa estrechamente con el CD4/TCR (*T-cell receptor*) como ligando; y CCR5 o CXCR4 como correceptores. Este proceso conduce a la fusión del HIV-1 a la membrana plasmática y a la posterior endocitosis. El maraviroc es uno de los antagonistas de CCR5, un inhibidor de la entrada del HIV-1 a células no infectadas.

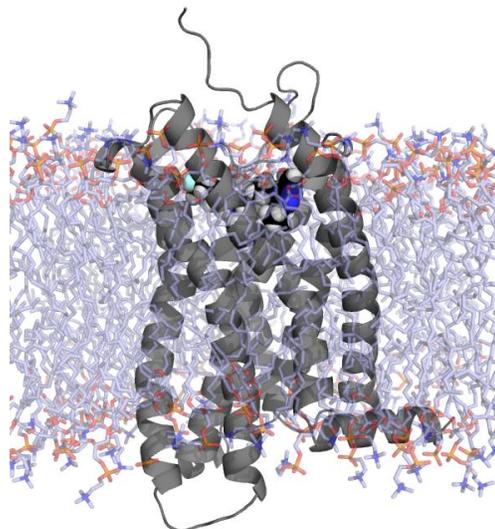


Figura 9.4. Estructura del receptor de referencia CCR5, obtenida del cristal PDB: 4MBS con resolución de 2.71 Å.⁵⁸

IX.4. Numeración de los receptores acoplados a proteínas G

El sistema de numeración propuesto por Ballesteros-Weinstein consiste en identificar los residuos conservados en las TMs de los GPCRs de neurotransmisores, y considerarlos una referencia. Para cualquier GPCR, la numeración de los residuos en las TMs se establece en relación con los residuos conservados para realizar comparaciones congruentes. Los residuos de referencia se asignan arbitrariamente como el número de TM y la cifra 50, separados por un punto. De esta manera, a los residuos localizados hacia el extremo amino les corresponden cifras menores a 50, y aquellos ubicados hacia el extremo carboxilo tienen identificadores mayores a 50⁶⁰ (Tabla 9.1).

Tabla 9.1. Numeración de Ballesteros-Weinstein para GPCRs.

Número de hélice TM	Residuo conservado	Identificador	Equivalencia en CX3CR1	Notas
1	N	N1.50	N49 ^{1.50}	
2	D	D2.50	D77 ^{2.50}	Residuo ligante de Na ⁺
3	R	R3.50	R127 ^{3.50}	Parte del motivo DRY
4	W	W4.50	W154 ^{4.50}	
5	P	P5.50	P203 ^{5.50}	
6	P	P6.50	P246 ^{6.50}	
7	P	P7.50	P290 ^{7.50}	Parte del motivo NPXXY

IX.5. Proteínas G heterotriméricas

La geometría de las proteínas G heterotriméricas (Figura 9.11) está conservada en los diferentes subtipos.

IX.5.1. Proteínas G_{i/o}

Las proteínas de clase G_{i/o}, al ser activadas, bloquean la producción del segundo mensajero AMP cíclico (cAMP) mediante la interacción de la subunidad G_{αi}-Mg[GTP] con la enzima transmembranal adenilato ciclasa. El heterodímero βγ a su vez interactúa con varias enzimas membranales como proteína cinasas C (PKC). Son sensibles a la inactivación en R178 por la toxina de *Vibrio cholerae* (CTX) y en C351 por la toxina de *Bordetella pertussis* (PTX).⁸⁸

IX.5.2. Proteínas G_s

Las proteínas de clase G_s estimulan la actividad catalítica de la adenilato ciclasa, mediante su interacción con la subunidad G_{αs}-Mg[GTP], aumentando la concentración de cAMP. Son sensibles a inactivación en R201 por CTX y la toxina termolábil de *Escherichia coli* (LT).⁸⁸

IX.5.3. Proteínas $G_{q/11}$

Las proteínas de clase $G_{q/11}$ estimulan la actividad catalítica de las fosfolipasas C membranales (PLC), mediante la interacción de la subunidad $G\alpha_{q/11}$ -Mg[GTP], de la cual se producen los segundos mensajeros diacilgliceroles (DAG) y *myo*-inositol-1,4,5-trifosfato (IP3) a partir de los fosfolípidos fosfatidilinositol-4,5-difosfato (PIP2) membranales. Son susceptibles de inhibición por CTX en R183.⁸⁸

IX.5.4. Proteínas $G_{12/13}$

Las proteínas de clase $G_{12/13}$ interactúan con proteínas G monoméricas de tipo Rho y están implicadas en el remodelamiento del citoesqueleto. Son resistentes a CTX y PTX.⁸⁸

IX.5.5. Proteínas G_z

Las proteínas de clase G_z comparten similitud con las proteínas $G_{i/o}$, pero son de expresión predominante en el sistema nervioso central y son resistentes a la inactivación por PTX⁶⁹ pero sensibles a la ADP-ribosilación en R179 por CTX.⁸⁸

IX.5.6. Transducinas o proteínas G_t

Las proteínas de clase G_t o T al ser estimuladas por el GEF, activan la fosfodiesterasa 6 de cGMP citosólica. Estas proteínas están asociadas a la rodopsina. Son sensibles a PTX, CTX y LT.⁸⁸

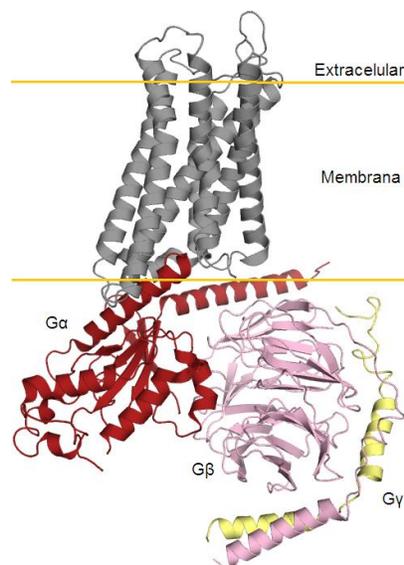


Figura 9.11. Geometría general de las proteínas G. En esta imagen se representa el heterotrímero G_i (PDB: 3SN6 y 5KDO), conformado por las isoformas GNAI1 ($G\alpha$), GB1 ($G\beta$) y GNG2 ($G\gamma$).

X. ANEXO III: Fundamentos teóricos

X. Dinámica molecular

La simulación de dinámica molecular (*MD*, del inglés *Molecular Dynamics*) calcula los desplazamientos, las fuerzas y la energía cinética y potencial del sistema mecánico que constituyen las moléculas a través de un intervalo de tiempo determinado. La simulación MD resuelve las ecuaciones clásicas de movimiento (ecuación 10.1) de N átomos interactuantes por un método numérico.

$$-\nabla E_{p,i} = \vec{F}_i = m_i \frac{\partial^2 \vec{r}_i}{\partial t^2} \quad (10.1)$$

X.1. Espacios de simulación

La resolución de la ecuación 9.1 para un sistema aislado \mathcal{U} con N partículas comienza en un espacio métrico lineal normado euclídeo de dimensión \mathbb{R}^{3N} , conocido como espacio de configuración $Q(r)$, que describe los tres componentes de posición r de cada partícula, $\vec{r}_i = \{x_i, y_i, z_i\}$. Al considerar a \mathcal{U} un sistema aislado, se impone la restricción de energía constante $dE_{\mathcal{U}}=0$. Para calcular las propiedades mecánicas del sistema se realiza la transformación al espacio de configuración ampliado $Q^v(r,v)$, con $\dim(Q^v) = \mathbb{R}^{6N}$ que describe las tres componentes de posición y de velocidad de cada partícula. Esta transformación es un haz o fibrado tangente de Q , que es la suma de todos los espacios tangentes a Q .

Adicionalmente, para acceder al cálculo de energía potencial se recurre a la transformación al espacio de fase $\Phi(r,p)$, que es el fibrado cotangente de Q , el cual describe el momento lineal \vec{p} y la posición \vec{r} de las N partículas de \mathcal{U} y constituye las trayectorias solución del sistema. Cada punto $(dr_i, dp_i) \in \Phi$ constituye un microestado Ω_i , que describe la posición y el momento para cada componente $\{i_N\}$.

X.2. Ergodicidad

Un sistema mecánico es ergódico si el único conjunto invariante de medida no nula de la variedad de energía potencial del espacio de fase es toda la variedad de energía potencial. El conjunto invariante es la posición de las N partículas, y es de medida no nula cuando $d\vec{r} \neq 0$. Con la restricción $dE_{\mathcal{U}}=0$, cuando $t \rightarrow \infty$, el sistema accede a todos los microestados $\{\Omega_N\}$ congruentes con las restricciones macroscópicas, y se dice que el sistema cumple la hipótesis ergódica. Además, la evolución temporal de \mathcal{U} en $\Phi(r,p)$ es una variedad topológica, y su isosuperficie corresponde con una distribución de sistemas \mathcal{U} en un instante dado, constituyendo un colectivo o ensamble. El colectivo contiene múltiples sistemas con la misma restricción. En cualquier instante t , la cantidad de sistemas en un microestado particular del colectivo es igual a su probabilidad. Para el cálculo de

las propiedades del sistema que son extrapolables a los casos macroscópicos, se requieren los parámetros asociados a las contribuciones vibracionales, rotacionales, traslacionales y electrónicos de las partículas, los cuales están contenidos en el campo de fuerza (*FF*, del inglés *Force field*).

Restricciones holonómicas

Cuando la distancia entre las coordenadas de dos átomos enlazados es invariante con la traslación y rotación de estos, se dice que se encuentran en un espacio homogéneo e isotrópico, el espacio euclídeo. En el espacio Q , las restricciones L son holonómicas si reducen L veces los grados de libertad del sistema, siendo los grados de libertad de este sistema restringido $6N - L$. Cada restricción holonómica define un subespacio métrico donde los posibles movimientos del subsistema pueden ocurrir. La condición L aplica una fuerza de restricción en cada partícula, para mantenerla en el subespacio correspondiente.⁸⁹ Durante la simulación de equilibrio mecánico inicial, se restringen sólo las posiciones de las cadenas principales de los polipéptidos, simulando una fuerza que limita su movimiento.⁹⁰ En diferentes etapas de la simulación también pueden restringirse sólo los átomos enlazados a hidrógeno para disminuir el costo computacional.

X.3. Procesos deterministas y estocásticos

La simulación de dinámica molecular involucra distintos procesos que implican la resolución de ecuaciones deterministas y estocásticas.

Un proceso $\{S_N\}$ es determinista si para cada estado, la transición s_i a s_{i+1} no depende de eventos aleatorios, y el estado s_{i+1} es el único sucesor de s_i .

Un proceso $\{S_N\}$ es estocástico si para cada estado s_i , existe una probabilidad diferente de cero de transitar a cada estado sucesivo $s_{i+1}, s_{j+1}, s_{k+1}$, etc.

X.4. Campos de fuerza (FF)

El campo de fuerza define el tratamiento de átomos, enlaces, ángulos e interacciones. Asigna a cada átomo un tipo en función de su hibridación, carga y valencia. A cada tipo le corresponde un grupo de parámetros (constantes de fuerza vibracional, rotacional y traslacional; radios atómicos, carga eléctrica, masa, etc.). Cada FF define una ecuación de energía potencial.⁹¹

X.5. Minimización de energía potencial por optimización geométrica

Las condiciones iniciales de un sistema termodinámico implican un equilibrio mecánico localizado en algún mínimo de la variedad de energía potencial. En la búsqueda de soluciones del

sistema de ecuaciones de energía potencial en función de las posiciones atómicas, se encuentran los puntos críticos energéticos de la variedad, donde $-\nabla E_{p,i}$ es cero. Uno de los métodos más utilizados en mecánica molecular es el método iterativo *Steepest Descent* (SD).⁹¹ Para aplicar el algoritmo se deben especificar las condiciones de la minimización de la energía (Tabla 10.1).

Tabla 10.1. Condiciones de la minimización de la energía del sistema.

Variable	Valor o argumento
Radio iónico de corto alcance	1.0 nm
Radio de van der Waals de corto alcance	1.0 nm
Magnitud de la fuerza máxima	<1000 kJnm/mol
Solvente	Explícito, agua con Na ⁺ y Cl ⁻ , 0.15 M

Algoritmo *Steepest Descent* (SD)

La dirección de descenso (ecuación 10.2) corresponde a la posición donde el potencial sea menor. El máximo desplazamiento d_{MAX} es definido inicialmente.

$$\vec{r}_{n+1} = \vec{r}_n + \max(d) \frac{\vec{F}_n(r)}{\max(|\vec{F}_n(r)|)} \quad (10.2)$$

Si $E_{p;n+1} < E_{p;n}$, se acepta la nueva posición. Si $E_{p;n+1} \geq E_{p;n}$ se rechaza la nueva posición.

El algoritmo SD es determinista, ya que no depende de funciones de variables aleatorias.

X.6. Colectivo canónico o isomolar-isocórico-isotérmico (NVT)

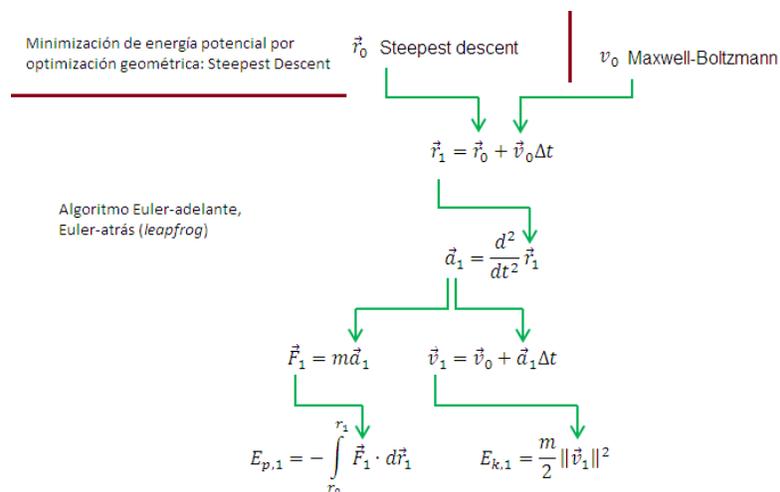
Bajo la suposición de que el promedio temporal de una propiedad del sistema es el promedio sobre el colectivo, la simulación del colectivo NVT con las condiciones especificadas (Tabla 10.2) determina la asignación de las velocidades iniciales y en consecuencia, la temperatura del sistema. El colectivo NVT simula un reservorio o baño de temperatura que interactúa con el sistema de interés. El reservorio de temperatura es un sistema con una capacidad térmica de valor tal que sus potenciales termodinámicos son invariantes ante la transferencia de calor. Para el sistema de interés, cada partícula en un microestado del colectivo tiene asociado un número de ocupación $\{a_k\}$, tal que la ocupancia A en cada microestado es la suma sobre a_k ; y a la distribución más probable de los números de ocupación se le conoce como distribución de Boltzmann. Las restricciones del colectivo NVT son que el sistema tiene volumen AV , cantidad de sustancia AN , energía total U y restricciones holonómicas L en la cadena principal de la proteína.

Tabla 10.2. Condiciones de las simulaciones NVT.

Variable	Valor o argumento
Integrador	SD mediante método <i>Leap-frog</i>
Número de pasos	5000
Paso de tiempo (dt)	2 fs = 0.002 ps
Tiempo de simulación	10 ps
Asignación de velocidades	Distribución de Maxwell-Boltzmann
Restricciones	Holonómicas para cadena principal
Temperatura de referencia (T_0)	310 K

Método de resolución de Euler-delante, Euler-detrás (*Leap-frog*)

Es un método iterativo para el problema con valores iniciales $r(\vec{0})$ de la ecuación 10.1. Con $r(\vec{0})$ y Δt , que depende del paso de tiempo, se calcula una posición subsecuente \vec{r}_1 a partir de ecuaciones iterativas para posición, velocidad y aceleración (Figura 10.1).⁹² La resolución del método es determinista, pero no sus variables.

**Figura 10.1.** Diagrama del método *leap-frog*.

Distribución de velocidades de Maxwell-Boltzmann

En el modelo del gas ideal, las velocidades de las partículas puntuales son descritas por la densidad de probabilidad de velocidades de Maxwell-Boltzmann⁹³ (Ecuación 10.3), una distribución derivada de la mecánica estadística de Boltzmann.

$$f(\|\vec{v}\|) = \sqrt{\frac{2}{\pi}} \left(\frac{m}{k_B T}\right)^3 \|\vec{v}\|^3 e^{-\frac{m\|\vec{v}\|^2}{2k_B T}} \quad (10.3)$$

Donde la velocidad media es la esperanza de v (Ecuación 10.4), y la temperatura del sistema es la media colectiva de la energía cinética (Ecuación 10.5).

$$\langle v \rangle = E(v) = \int_0^{\infty} v f(v) dv \quad (10.4)$$

$$T = \frac{2}{3N_{df} k_B} \left\langle \sum_{i=1}^N m_i \|v_i\|^2 \right\rangle \quad (10.5)$$

Dada la naturaleza aleatoria de la variable v de la función de distribución de Maxwell-Boltzmann, la asignación de velocidades iniciales y la energía cinética de los átomos del sistema son procesos estocásticos, por lo que el sistema está sujeto a fluctuaciones térmicas.

X.7. Colectivo isomolar-isobárico-isotérmico (NPT)

El colectivo canónico representa las propiedades macroscópicas N , V y T . Para calcular la presión del sistema, y a través de ésta conocer la entalpía H ; se realiza el cambio de variable V a P donde $f(V)$ es convexa ($f''(V) > 0, \forall V \in \mathcal{U}$) y suave, por lo que $f'(V)$ es monótona y por lo tanto homeomorfa con $f(V)$. La transformada de Legendre es el morfismo $L: f(V) - f'(V)V$; donde V y $f'(V)=P$ son las variables conjugadas. Solo es aplicable para dos funciones diferenciables f_1, f_2 tales que $f'_1 = (f'_2)^{-1}$. Entonces f_1, f_2 son homeomorfas mediante la transformada de Legendre. Las condiciones para el colectivo NPT se definen inicialmente (Tabla 10.3).

Tabla 10.3. Condiciones de las simulaciones NPT.

Variable	Valor o argumento
Integrador	SD mediante método <i>Leap-frog</i>
Número de pasos	5000
Paso de tiempo (dt)	2 fs = 0.002 ps
Tiempo de simulación	10 ps
Asignación de velocidades	Nula
Restricciones	Holonómicas para cadena principal
Presión de referencia (P_0)	1 bar
Compresibilidad isotérmica del agua	$4.5 \times 10^{-5} \text{ bar}^{-1}$

La presión del sistema debe llegar al equilibrio antes de realizar la simulación de producción MD. La presión se calcula inicialmente como el tensor \mathbf{P} de orden 3 (ecuación 10.6) a partir del tensor virial Ξ (Ecuación 10.7).⁹⁰

$$\mathbf{P} = \frac{2}{V} \left[\left(\sum_{i=1}^N E_{k,i} \right) - \Xi \right] \quad (10.6)$$

$$\Xi = -\frac{1}{2} \sum_{i < j} (\vec{r}_{ij} \otimes \vec{F}_{ij}) \quad (10.7)$$

El escalar presión es la traza del tensor \mathbf{P} , la suma de los elementos diagonales del tensor.

X.8. Dinámica molecular

Método de resolución de Verlet

El teorema de Taylor (Ecuación 10.8) establece que si una función tiene derivadas continuas hasta la de orden $n+1$ en un intervalo dado, la función se aproxima con una serie de n términos y un residuo R_n asociado al error de la aproximación.⁹⁴

$$r(t) = \sum_{i=0}^n \frac{(\Delta t)^i}{i!} \frac{d^i}{dt^i} r(t_0) + R_n \quad (10.8)$$

La ecuación 10.8 es la serie de Taylor de r alrededor de t_0 .

En la simulación MD, las ecuaciones de Newton son resueltas simultáneamente en pasos cortos de tiempo (típicamente comprendido entre 1 y 10 fs) mediante el algoritmo de Verlet, el cual emplea dos aproximaciones en serie de Taylor truncadas al cuarto término (ecuaciones 10.9a y 10.9b) para las posiciones del paso nuevo ($t+\Delta t$) y el previo ($t-\Delta t$) respectivamente. El sistema mantiene la presión y la temperatura en los intervalos requeridos, y las coordenadas atómicas son escritas en un archivo de salida cada cierto período. Las coordenadas como función de t representan la trayectoria del sistema.

$$r(t) + r(\Delta t) = \sum_{i=0}^3 \frac{(\Delta t)^i}{i!} \frac{d^i}{dt^i} r(t_0) \quad (10.9a)$$

$$r(t) - r(\Delta t) = \sum_{i=0}^3 \frac{(-\Delta t)^i}{i!} \frac{d^i}{dt^i} r(t_0) \quad (10.9b)$$

El algoritmo de Verlet es casi exacto, siendo el error de truncación de la serie del orden de $(\Delta t)^4$.⁹¹

La simulación MD se realiza considerando los enlaces como osciladores armónicos, con una oscilación simple con un grado de libertad. Al eje internuclear le corresponde una posición al equilibrio en la cual la energía vibracional es mínima, y la desviación de esta posición genera una fuerza (Ecuación 10.10) que retorna el enlace al punto de equilibrio. La energía potencial del enlace considerado un resorte según la Ley de Hooke (Ecuación 10.11) y la energía cinética asociada (Ecuación 10.12) generan la ecuación diferencial de movimiento (Ecuación 10.13).

$$\vec{F} = -\nabla_r U = k(\vec{r} - \vec{r}_{eq}) \quad (10.10)$$

$$E_p = \frac{k(\|\vec{r} - \vec{r}_{eq}\|)^2}{2} \quad (10.11)$$

$$E_k = \frac{m\|\vec{v}\|^2}{2} \quad (10.12)$$

$$\frac{d^2}{dt^2} \vec{r} + \frac{k}{m} \vec{r} = r'' + \omega^2 r = 0 \quad (10.13)$$

La cual tiene raíces complejas conjugadas, y la solución (Ecuación 10.14) denota que el enlace tiene un movimiento armónico.^{90,95}

$$r = C_1 \cos \omega t + C_2 \sin \omega t \quad (10.14)$$

La energía potencial rotacional (Ecuación 10.15) y la energía cinética (debido a la invarianza traslacional, Ecuación 10.16) definen la energía interna U del sistema (Ecuación 10.17).

$$E_{rot} = \frac{m\|\vec{r}\|^2 \omega^2}{2} \quad (10.15)$$

$$E_{tras} = E_k = \frac{m\|\vec{v}\|^2}{2} \quad (10.16)$$

$$U = \sum_{i=1}^N E_{p,i} + \sum_{i=1}^N E_{k,i} \quad (10.17)$$

Intervalo de tiempo de integración

El intervalo de tiempo de integración (τ) define la simulación correcta en la trayectoria. Con valores pequeños de τ la trayectoria de la simulación es más congruente con el resultado de la integración analítica. Con valores grandes de τ los átomos pueden ocupar las mismas coordenadas y producen falla de la simulación.⁹¹

Temperatura y termostato de Berendsen

El algoritmo de Berendsen o termostato de Berendsen (Ecuación 10.18) es una corrección que previene la desviación de la temperatura del sistema. Esta desviación decae exponencialmente con la constante τ .⁹⁰

$$\frac{dT}{dt} = \frac{T - T_0}{\tau}, \quad T = T_0 \left(1 - e^{-\frac{t}{\tau}}\right) \quad (10.18)$$

Presión y barostato de Parrinello-Rahman

En el barostato de Parrinello-Rahman la matriz del vector de la caja de simulación \mathbf{B} obedece la nueva ecuación de movimiento (Ecuación 10.19).

$$\frac{d^2\mathbf{B}}{dt^2} = \mathbf{V}\mathbf{W}^{-1}(\mathbf{B}^T)^{-1}(\mathbf{P} - P_{ref}) \quad (10.19)$$

Donde \mathbf{W} es un parámetro matricial que es función de la compresibilidad isotérmica y la constante τ_P .

Con el barostato de Parrinello-Rahman se modifican las ecuaciones de movimiento (Ecuación 10.20). La constante τ_P debe ser mayor a la constante τ , ya que si el sistema se encuentra fuera del equilibrio resulta en oscilaciones muy grandes.⁹⁰

$$\frac{d^2\vec{r}}{dt^2} = \frac{\vec{F}}{m} - \mathbf{M} \frac{d\vec{r}}{dt} \quad (10.20)$$

Donde \mathbf{M} es una matriz de modificación de las matrices \mathbf{B} .

Límites del sistema

Las condiciones periódicas en la frontera (*periodic boundary conditions, PBC*) simulan un sistema infinito, donde la celda de simulación está rodeada de réplicas de sí misma, evitando desviaciones de los átomos limítrofes respecto a los átomos centrales del sistema.

Sumas de Ewald y *PME*

Está basado en la expresión coulombica de la energía potencial electrostática, y define las interacciones de las partículas en la celda de simulación con sus correspondientes partículas en *PBC*.

En una celda B con N cargas puntuales $\{q_N\}$ en posiciones respectivas $\{\vec{r}_N\}$, definida por tres vectores $\vec{e}_1, \vec{e}_2, \vec{e}_3$, la carga q_i en \vec{r}_i tiene una interacción coulombica con q_j en \vec{r}_j ($i \neq j$), así como con sus imágenes periódicas en $\vec{r}_j + n_1 \vec{e}_1 + n_2 \vec{e}_2 + n_3 \vec{e}_3$. La carga q_i también interactúa con sus propias imágenes periódicas en $\vec{r}_i + n_1 \vec{e}_1 + n_2 \vec{e}_2 + n_3 \vec{e}_3$.⁹⁶ La energía electrostática de Ω (ecuación 10.21) tiene la serie infinita exterior cuya convergencia depende de los valores de n_1, n_2 y n_3 .⁹⁷

$$E(r_1, \dots, r_N) = \frac{1}{2} \sum_{n \neq 0} \sum_i \sum_{j \neq i} \frac{q_i q_j}{\|\vec{r}_i - \vec{r}_j + \vec{n}\|} \quad (10.21)$$

El método de sumas de Ewald (ecuación 10.22) transformó la serie infinita por dos términos que convergen absolutamente, una suma directa en el espacio lineal aritmético \mathbb{R}^3 (E_{dir}) y una suma recíproca en el espacio de Fourier (E_{rec}). Las interacciones no enlazantes y no periódicas no se calculan y se engloban en un tercer término de corrección (E_{cor}).

$$E = E_{\text{dir}} + E_{\text{rec}} + E_{\text{cor}} \quad (10.22)$$

$$E_{\text{dir}} = \sum_{\vec{n}} \frac{1 - \text{erf}(\beta \|\vec{r} + \vec{n}\|)}{\|\vec{r} + \vec{n}\|}$$

$$E_{\text{rec}} = \frac{1}{\pi V} \sum_{\vec{m}} \frac{1}{\|\vec{m}\|^2} e^{-\left(\frac{\pi \|\vec{m}\|}{\beta}\right)^2} e^{2\pi i \vec{m} \cdot \vec{r}}$$

Donde $\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$ es la función error de x.

El primer sumando equivale físicamente a rodear cada carga puntual en el sistema por una distribución, típicamente normal, neutralizante, de cargas de igual magnitud y signo contrario. Éste término es responsable de las interacciones de corto alcance.

El segundo sumando compensa la distribución neutralizante del primer término, mediante una distribución compleja de cargas de signo opuesto a las del espacio real.⁹¹ La base del espacio lineal de tres dimensiones son los vectores ortogonales de posición $\hat{i}, \hat{j}, \hat{k}$ ⁹⁸ y la base del espacio de Fourier son los vectores de onda $m_1 = 2\pi \frac{j \times k}{(i \times j) \cdot k}$, $m_2 = 2\pi \frac{k \times i}{(i \times j) \cdot k}$ y $m_3 = 2\pi \frac{i \times j}{(i \times j) \cdot k}$.⁹⁹

Las aproximaciones de la suma de Ewald, basadas en mallas (*particle mesh Ewald, PME*) aceleran la solución del método. En lugar de considerar una densidad continua de carga, se

aproximan las cargas distribuidas en una malla⁹¹ sobre la celda de simulación, discretizando los valores y aplicando la transformada rápida de Fourier (FFT, un método rápido de cálculo de la transformada discreta de Fourier). Se simula un cubo de 0.16 nm de lado en el espacio de Fourier para la suma recíproca.

Interacciones de van der Waals

Estas interacciones pueden ser dipolo-dipolo, dipolo-dipolo inducido o interacciones de Debye, y dipolo instantáneo-dipolo inducido o interacciones de London.

Las interacciones dipolo-dipolo se consideran como de orientación fija o de media angular. Las interacciones de orientación fija tienen cuatro variables, la distancia de separación entre los dipolos y tres ángulos, y la energía de la interacción (Ecuación 10.23) depende de la posición de los átomos en el sistema. La energía de interacción de media angular asume las fluctuaciones medias ponderadas de los dipolos en función de la temperatura (Ecuación 10.24).¹⁰⁰

$$U(\theta_1, \theta_2, \theta_3, r_{ij}) = \frac{(\vec{p}_i \cdot \vec{r}_{ij})(\vec{p}_j \cdot \vec{r}_{ij})}{4\pi\epsilon\epsilon_0 \|\vec{r}_{ij}\|^3} f(\theta_1, \theta_2, \theta_3) \quad (10.23)$$

$$U(r_{ij}) = \langle U(\theta_1, \theta_2, \theta_3, r_{ij}) \rangle = \frac{\iiint U(\theta_1, \theta_2, \theta_3, r_{ij}) e^{-\frac{U(\theta_1, \theta_2, \theta_3, r_{ij})}{k_B T}} d\theta_1 d\theta_2 d\theta_3}{\iiint f(\theta_1, \theta_2, \theta_3) e^{-\frac{U(\theta_1, \theta_2, \theta_3, r_{ij})}{k_B T}} d\theta_1 d\theta_2 d\theta_3} \propto \frac{1}{r^6} \quad (10.24)$$

La energía de las interacciones de Debye (Ecuación 10.25) considera el campo eléctrico (\vec{E}) y una media ponderada del ángulo debido a la posición del dipolo.

$$U(\theta, r_{ij}) = \frac{f(\vec{E})f(\theta)}{r_{ij}^6} \quad (10.25)$$

La energía de las interacciones de London (Ecuación 10.26) no incluye momentos dipolares, sino una función de las primeras energías de ionización de los átomos i, j involucrados.

$$U(\theta, r_{ij}) = \frac{f(E_{ion,1,i}, E_{ion,1,j})f(\alpha_i, \alpha_j)}{r_{ij}^6} \quad (10.26)$$

Los tres tipos de energías de van der Waals se engloban aproximadamente en el término de grado 6 en la ecuación de Lennard-Jones¹⁰¹ (Figura 10.2).

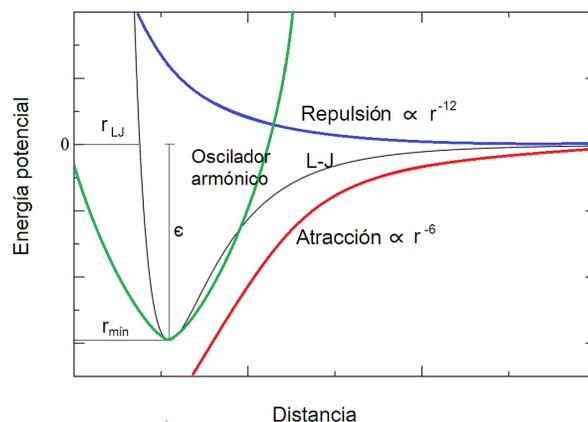


Figura 10.2. Gráfica de la energía potencial de atracción y repulsión, y su aproximación por la ecuación de Lennard-Jones (L-J).

Interacciones hidrofóbicas

La atracción hidrofóbica es la fuerza no covalente y no electrostática más importante entre partículas apolares. En un medio acuoso, la atracción hidrofóbica se debe a la energía libre de los enlaces de hidrógeno entre las moléculas de agua del entorno, siendo la energía de la atracción hidrofóbica igual a la energía libre de cohesión en el agua.¹⁰² Componentes de estas interacciones se incluyen en los términos constantes de la ecuación de Lennard-Jones. El efecto hidrofóbico entre el solvente y las proteínas se aproxima mediante las ecuaciones diferenciales estocásticas langevianas.

Dinámica langeviana

La aplicación del algoritmo Euler-delante Euler-detrás para un número grande de moléculas en el sistema ralentiza considerablemente el cálculo numérico. La simplificación del cálculo mediante ecuaciones diferenciales estocásticas, particularmente las ecuaciones de Langevin agilizan en gran medida el cómputo. Las ecuaciones de Langevin asumen que el comportamiento de las partículas del sistema es un proceso gaussiano estocástico. La función gaussiana estocástica $N(\{r\}_n)$ es multivariada, ya que incluye un gran número de coordenadas de partículas, y estacionaria porque la media y la varianza de la función a través del tiempo de simulación son invariantes. Para las interacciones que involucran un gran número de moléculas de solvente, la ecuación de Langevin (10.27) simula la viscosidad η y la fricción de manera aproximada.^{103,104}

$$\sum_{i=1}^N m_i D_t^2 \mathbf{R} = -\nabla U(\mathbf{R}) - \eta D_t \mathbf{R} + \sqrt{2\eta k_B T} N(\{r\}_n) \quad (10.27)$$

X.9. Análisis de las simulaciones

Desviación respecto a la posición inicial: RMSD

Una parte del análisis de las simulaciones de MD se basa en las desviaciones de la posición del átomo i -ésimo respecto a una referencia, en este caso a las coordenadas iniciales a través del tiempo de simulación. La raíz cuadrada de la desviación cuadrática media, RMSD (*Root-mean-square deviation*, Ecuación 10.28) es la raíz cuadrada del segundo momento centrado en las posiciones iniciales y ponderado con las masas atómicas, es decir, la raíz cuadrada de la varianza centrada en las posiciones iniciales y ponderada con masas atómicas y expresa los desplazamientos medios de las N partículas a través del tiempo.

$$\text{RMSD} = \sqrt{\frac{\sum_{i=1}^N m_i \|r_i(t) - r_i(t_0)\|^2}{\sum_{i=1}^N m_i}} \quad (10.28)$$

Se considera el sistema en un estado de equilibrio mecánico cuando los valores de RMSD en función del tiempo tienen oscilaciones mínimas; cuando la amplitud de la función periódica tiende a un mínimo. Debido a la naturaleza de las gráficas de RMSD en función del tiempo, se puede elegir el intervalo de tiempo aproximado al cual se considera el sistema equilibrado.

Codesviación respecto a la posición inicial: RMSF

La raíz cuadrada de la fluctuación cuadrática media, RMSF (*Root-mean-square fluctuation*, Ecuación 10.29) es la raíz cuadrada del segundo momento bivariado centrado en las posiciones iniciales, y ponderado con masas atómicas y en la partición de tiempo analizado, o la raíz cuadrada de la covarianza centrada en las posiciones iniciales, y ponderada con masas atómicas y en la partición de tiempo analizado. En otras palabras, es la codesviación de las posiciones a través del tiempo y para cada partícula.¹⁰⁵

$$\text{RMSF} = \sqrt{\frac{\sum_{i=1}^N \sum_t^{n_t} (\|r_i(t) - r_i(t_0)\|^2)}{n_t \sum_{i=1}^N m_i}} \quad (10.29)$$

El análisis de RMSF suele realizarse desde el tiempo t en que el sistema ha alcanzado un equilibrio dinámico hasta el final de la simulación para evitar fluctuaciones de gran magnitud. Dependiendo del sistema, este valor de t puede variar considerablemente, y debe encontrarse de forma particular.

Radio de giro

El radio de giro de un subsistema (Ecuación 10.30a) es la distancia media ponderada de los átomos al baricentro. El baricentro o centro de masa (Ecuación 10.30b) es el punto donde las posiciones relativas ponderadas por masa de los componentes suman cero. El radio de giro es un indicador del desplegamiento o compacidad de la estructura.

$$rg(t) = \sqrt{\frac{\sum_{i=1}^N [m_i (\|\vec{r}_i - \vec{r}_{bar}\|)^2]}{\sum_{i=1}^N m_i}} \quad (10.30a)$$

$$\vec{r}_{bar} = \frac{\sum_{i=1}^N m_i \vec{r}_i}{\sum_{i=1}^N m_i} \quad (10.30b)$$

Análisis de estructura secundaria

La estructura secundaria de una proteína se refiere a disposiciones particularmente estables de los aminoácidos que dan lugar a patrones estructurales repetitivos. El análisis de estructuras secundarias a través del tiempo de simulación consiste en el cálculo de los ángulos diedros ϕ y ψ de la cadena principal de la proteína.

Análisis de componentes principales

El análisis de componentes principales es un método de reducción de dimensionalidad que permite detectar los movimientos que determinan la mayor varianza del subsistema. A partir de la matriz de covarianza Σ centrada en la media y ponderada (Ecuación 10.30) o no ponderada, se calcula la estructura media, y posteriormente se diagonaliza (Ecuación 10.31) para obtener los autovalores y autovectores del sistema.

$$\Sigma_{3N \times 3N} = \left(\frac{\sum_{i=1}^N \sum_{j \neq i} \sqrt{m_i m_j} ((r_i(t) - \bar{r}_i)) ((r_j(t) - \bar{r}_j))}{\sum_{i=1}^N \sum_{j \neq i} \sqrt{m_i m_j}} \right) \quad (10.30)$$

$$\mathbf{V}^T \Sigma \mathbf{V} = \text{diag}(\Sigma) = \Lambda \quad (10.31)$$

Donde cada elemento λ_{ii} de la matriz diagonal Λ es el autovalor de Σ y cada columna \mathbf{V}_j es el autovector asociado a la transformación. Un autovector, eigenvector o vector propio \mathbf{V} es todo vector no nulo que al ser transformado por un operador no cambia su dirección (Ecuación 10.32).

$$\mathbf{A} \mathbf{V} = \lambda \mathbf{V} \quad (10.32)$$

Donde \mathbf{A} es una transformación lineal.

Los autovalores se suelen ordenar de manera decreciente, y los de mayor magnitud corresponden con movimientos colectivos. La dimensionalidad del eigensistema, $3N$, puede reducirse al considerar algunos de los autovectores, con frecuencia los asociados a los mayores desplazamientos. La proyección de la trayectoria euclídea sobre todo o una parte del eigensistema (Ecuación 10.33) genera el subespacio esencial, generalmente de dimensión mucho menor a $3N$, por ejemplo, 1, 2 o 3.

$$\mathbf{Q}_{3N \times 3N} \mathbf{V}_{3N \times 1} = \mathbf{E}_{3N \times 1} \quad (10.33)$$

La comparación de subespacios esenciales entre trayectorias es útil para conocer las locaciones accesibles para cada sistema dinámico.

Correlación cruzada e información mutua

La trayectoria puede contener desplazamientos atómicos con tendencias determinadas. Cuando la posición de al menos un par de partículas cambia de manera concertada, se dice que están correlacionadas por la transformación dinámica. Una manera de cuantificar la correlación es mediante el cálculo de la correlación cruzada R (Ecuación 10.34). La correlación cruzada se calcula sobre dos funciones de cambio de ángulo a través del tiempo; la primera permaneciendo al tiempo t y la segunda al tiempo $t+dt$. A diferencia de la correlación simple que se calcula sobre diferentes valores de la misma función, la correlación cruzada se calcula sobre diferentes valores de una función fija y una función móvil. La matriz de correlación cruzada \mathbf{R} para todos los pares de residuos se calcula sobre el espacio de distancias de enlace, ángulos y torsiones, y no sobre el espacio euclídeo.

$$R = (\varphi_1 * \varphi_2)(t) = \sum_{t=0}^{n_t} \varphi_1(t) \varphi_2(t + dt) \quad (10.34)$$

La entropía de Shannon $H_2(x)$ de una variable aleatoria x es una medida de su incertidumbre (Ecuación 10.35). La entropía relativa es una medida de la distancia entre dos distribuciones de probabilidad, y es una medida de la ineficiencia de asumir que una variable tiene distribución de probabilidad $p(x)$ cuando tiene $q(x)$ (Ecuación 10.36).

$$H_2(x) = - \sum_{x \in \text{Dom}(X)} p(x) \log_2 p(x) \quad (10.35)$$

$$H_2(p \parallel q) = \sum_x p(x) \log_2 \frac{p(x)}{q(x)} \quad (10.36)$$

La información mutua I es una medida de la cantidad de información que una variable aleatoria contiene acerca de otra variable aleatoria. Describe la correlación entre grados de libertad y proporciona una medida de la cantidad de información asociada a un grado de libertad que se obtiene al conocer algo del otro grado de libertad.⁵⁶ La información mutua se conoce mediante la divergencia de Küllback-Leiber D_{KL} (Ecuación 10.37), y determina la reducción de la entropía de una variable aleatoria dado el conocimiento que se tiene de la otra.

$$I(\phi_i; \phi_j) = \sum_x p(\phi_i, \phi_j) \log_2 \frac{p(\phi_i, \phi_j)}{p(\phi_i)p(\phi_j)} = D_{KL}(p(\phi_i, \phi_j) \parallel p(\phi_i)p(\phi_j)) \quad (10.37)$$

Para las simulaciones MD, la información mutua cuantifica la correlación entre los confórmers en diferentes sitios, basándose en cálculos entrópicos para analizar los colectivos conformacionales. Al igual que el cálculo del coeficiente de correlación bivariado, la información mutua se calcula en el espacio BAT, enfocándose en los valores de los ángulos diedros.

La entropía torsional (Ecuación 10.38) está relacionada con la información mutua. Bajo la hipótesis nula de independencia, la distribución conjunta de dos variables ϕ_1, ϕ_2 es igual al producto de sus distribuciones marginales (Proposición 10.39). Dado que en ocasiones la información mutua de dos variables aleatorias cualesquiera es no nula, el cálculo de $MutInf$ ⁵⁶ emplea un estimador corregido de entropía histográfica sobre particiones adaptativas (Ecuación 10.40).

$$S = \sum_i^n \int_0^{2\pi} p(\phi_i) \ln p(\phi_i) d\phi_i - \sum_i^n \sum_j^n \int_0^{2\pi} \int_0^{2\pi} p(\phi_i, \phi_j) \ln \frac{p(\phi_i, \phi_j)}{p(\phi_i)p(\phi_j)} d\phi_i d\phi_j \quad (10.38)$$

$\forall i, j \in n \mid n = \dim(T); T < BAT$

$$H_0: f(\phi_i, \phi_j) = f(\phi_i) \cap f(\phi_j) \quad (10.39)$$

$$I = \sum_{i=1}^r \frac{n_i}{N} \left(\ln N - F(n_i) - \frac{(-1)^{n_i}}{n_i + 1} \right) + \sum_{j=1}^s \frac{n_j}{N} \left(\ln N - F(n_j) - \frac{(-1)^{n_j}}{n_j + 1} \right) - \sum_{j=1}^{rs} \frac{n_{ij}}{N} \left(\ln N - F(n_{ij}) - \frac{(-1)^{n_{ij}}}{n_{ij} + 1} \right) \quad (10.40)$$

Donde $F(n_i) = \frac{\Gamma'(n_i)}{\Gamma(n_i)}$ es la función digamma.

Este estimador asume que los histogramas provienen de un proceso Poisson ($n_{ij} \ll N$). Adicionalmente, para evitar el artefacto de $I \neq 0$ dado que $\phi_1 \perp \phi_2$, se realizan permutaciones de los datos muestrales. Asimismo, para evitar falsos positivos, se descartan mediante un método bayesiano. Si se asume que la información mutua proviene de una distribución normal, esto es, $I \sim N\{E(I), \text{Var}(I)\}$; la probabilidad de que la información mutua calculada sea menor al valor

esperado si las variables torsionales φ_1, φ_2 son independientes se elimina si es mayor al valor de significancia (Proposición 10.41).

$$p(I(\varphi_1, \varphi_2) < (E\{I(\varphi_1, \varphi_2)\} | \varphi_1 \perp \varphi_2) > \alpha) \quad (10.41)$$

Matrices de distancias

La matriz de distancias mínimas entre dos residuos de aminoácidos A_i, A_j en una proteína se define como la distancia más pequeña entre los pares de átomos $i \in A_i, j \in A_j$. En la gráfica de la matriz se pueden encontrar el establecimiento de interacciones dependientes de la distancia.⁹⁰

Energía de unión

Muestreo *Umbrella* o umbeliforme

De la ecuación de la energía de Keesom (Ecuación 10.22), la relación de ángulos define la configuración espacial de un dipolo. La ecuación de Keesom puede reescribirse en relación con la función de configuración q (Ecuaciones 10.42 y 10.43).

$$U(\theta_1, \theta_2, \theta_3, r_{ij}) = \frac{(\vec{p}_i \cdot \vec{r}_{ji})(\vec{p}_j \cdot \vec{r}_{ij})}{4\pi\epsilon\epsilon_0 \|\vec{r}_{ij}\|^3} f(\theta_1, \theta_2, \theta_3) = \frac{(\vec{p}_i \cdot \vec{r}_{ji})(\vec{p}_j \cdot \vec{r}_{ij})}{4\pi\epsilon\epsilon_0} q \quad (10.42)$$

$$q = \frac{f(\theta_1, \theta_2, \theta_3)}{\|\vec{r}_{ij}\|^3} \quad (10.43)$$

Para un subsistema dinámico con un dipolo, la energía de Helmholtz asociada en el colectivo NVT (Ecuación 10.44) incluye la integral de configuración Z (Ecuación 10.45) relacionada directamente con la función de partición canónica q .

$$A_\mu = A_\mu^0 - k_B T \ln \frac{Z}{Z^0} \quad (10.44)$$

$$Z = \iiint_{\Theta} e^{-\frac{U(\theta_1, \theta_2, \theta_3, r_{ij})}{k_B T}} d\theta_1 d\theta_2 d\theta_3 \quad (10.45)$$

Al realizar histogramas de las posiciones del centro de masa a través de la coordenada de reacción para cada una de las configuraciones, estos adquieren una forma acampanada o umbeliforme, con curtosis y asimetría variables. El traslape de cada histograma con aquellos adyacentes es necesario para la disminución del sesgo en la estimación de la energía libre.