



**Universidad de Concepción**  
**Facultad de Ciencias Biológicas**

**EXPLORACIÓN CONFORMACIONAL DE rGLUT5 MEDIANTE  
DINÁMICA MOLECULAR DE GRANO GRUESO Y MODELOS  
DE ESTADOS DE MARKOV**

Tesis presentada a la Facultad de Ciencias Biológicas de la Universidad de  
Concepción para optar al grado académico de Magister en Bioquímica y  
Bioinformática

**BASTIAN GIOVANNI LILLO DAPREMONT**

Profesor Guía: Dr. Alexis Marcelo Salas Burgos  
Dpto. de Farmacología, Facultad de Cs. Biológicas  
Universidad de Concepción

Concepción, Chile  
Marzo, 2026

Se autoriza la reproducción total o parcial, con fines académicos, por cualquier medio o procedimiento, incluyendo la cita bibliográfica del documento.

## **AGRADECIMIENTOS**

A quienes integran el Laboratorio de Dinámica Molecular del Cáncer y Ramírez Lab: por su empatía, por el conocimiento compartido y, sobre todo, por esa forma de amistad que solo nace cuando la curiosidad y la humanidad caminan juntas.

A las amistades que esta ciudad extraña y hermosa fue tejiendo sin aviso, recordándome que los lugares nos eligen tanto como nosotros a ellos.

Y a mi familia, que no necesita estar presente para estar siempre

## TABLA DE CONTENIDO

<b>ÍNDICE DE TABLAS .....</b>	<b>v</b>
<b>ÍNDICE DE FIGURAS .....</b>	<b>vi</b>
<b>RESUMEN .....</b>	<b>vii</b>
<b>ABSTRACT .....</b>	<b>viii</b>
<b>INTRODUCCIÓN .....</b>	<b>1</b>
1. <i>Superfamilia de Facilitadores Principales .....</i>	<i>1</i>
2. <i>Características estructurales de los MFS y SP .....</i>	<i>5</i>
3. <i>Modelo de Acceso Alternante (MAA).....</i>	<i>10</i>
4. <i>Simulaciones de Dinámica Molecular.....</i>	<i>14</i>
5. <i>El problema de muestreo .....</i>	<i>18</i>
6. <i>Técnicas de mejoramiento de muestreo .....</i>	<i>19</i>
7. <i>Simulaciones de Dinámica Molecular de Grano Grueso.....</i>	<i>21</i>
8. <i>Campo de fuerza SIRAH.....</i>	<i>25</i>
9. <i>Planteamiento .....</i>	<i>27</i>
<b>HIPÓTESIS.....</b>	<b>30</b>
<b>OBJETIVOS .....</b>	<b>30</b>
<b>MATERIALES Y MÉTODOS .....</b>	<b>31</b>
1. <i>Modelamiento de conformaciones.....</i>	<i>32</i>
2. <i>Construcción de ensamblés solvatados en topología DM-GG .....</i>	<i>36</i>
3. <i>Relajación de ensamblés termodinámicos en condiciones NPT .....</i>	<i>37</i>
4. <i>Muestrear el espacio conformacional de cada ensamble .....</i>	<i>40</i>
5. <i>Selección de Variables Colectivas (CV).....</i>	<i>41</i>
6. <i>Análisis de convergencia estadística .....</i>	<i>42</i>
7. <i>Discretización del espacio conformacional.....</i>	<i>43</i>
8. <i>Transiciones conformacionales mediante modelos de Markov.....</i>	<i>45</i>
<b>RESULTADOS.....</b>	<b>47</b>
1. <i>Validación del descriptor y modelos.....</i>	<i>47</i>
2. <i>Estabilidad termodinámica de los sistemas GG en ensamble NPT .....</i>	<i>54</i>

3. Exploración conformacional mediante DM-GG .....	58
4. Transiciones conformacionales en el espacio explorado .....	64
<b>DISCUSIÓN .....</b>	<b>70</b>
<b>CONCLUSIÓN .....</b>	<b>79</b>
<b>GLOSARIO .....</b>	<b>81</b>
<b>REFERENCIAS .....</b>	<b>88</b>
<b>ANEXOS .....</b>	<b>99</b>

## ÍNDICE DE TABLAS

Tabla 1. Los transportadores MFS humanos y sus funciones. ....	2
Tabla 2. Campos de fuerzas aplicados a DM. ....	15
Tabla 3. Estructuras utilizadas para modelamiento comparativo. ....	31
Tabla 4. Miembros de la familia SP con estructura definida.....	48
Tabla 5. Modelos comparativos para el mecanismo alternante.....	50
Tabla 6. Análisis de convergencia mediante análisis de bloques. ....	58

## ÍNDICE DE FIGURAS

Figura 1. Filogenia de las proteínas pertenecientes a la superfamilia MFS. ....	2
Figura 2. Características estructurales de los transportadores MFS. ....	7
Figura 3. Características estructurales de transportadores SP. ....	9
Figura 4. Modelo de Acceso Alternante y mecanismos en transportadores SP. ....	13
Figura 5. Flujo de trabajo en una simulación de Dinámica Molecular. ....	17
Figura 6. Representación topológica Grano Grueso. ....	22
Figura 7. Línea temporal DM aplicadas a MFS. ....	28
Figura 8. Esquema metodológico propuesto. ....	30
Figura 9. Protocolo para la relajación de los sistemas. ....	36
Figura 10. Evaluación del descriptor estructural. ....	47
Figura 11. Evaluación estructural de los modelos. ....	51
Figura 12. Representaciones atómica y CG de rGLUT5. ....	52
Figura 13. Condiciones termodinámicas en simulaciones CG bajo ensamble NPT. ....	53
Figura 14. Equilibración de simulaciones MD-GC. ....	54
Figura 15. Exploración del espacio conformacional mediante MD-GC 3 $\mu$ s. ....	56
Figura 16. Exploración conformacional con simulaciones extendidas. ....	61
Figura 17. Discretización del espacio conformacional accesible. ....	63
Figura 18. Red de probabilidades de transición en la exploración accesible. ....	66

## RESUMEN

Las simulaciones de Dinámica Molecular (DM) permiten estudiar procesos biológicos a nivel atómico, pero su aplicación al análisis de transiciones conformacionales está limitada por el problema de muestreo, especialmente en proteínas transportadoras con barreras energéticas elevadas. Aunque los métodos fuera del equilibrio y de equilibrio sesgado permiten caracterizar rutas conformacionales, requieren potenciales externos y limitan la interpretación de la dinámica en equilibrio.

En este trabajo se estudiaron las transiciones conformacionales del modelo de acceso alternante del transportador GLUT5 de *Rattus norvegicus* (rGLUT5) mediante simulaciones no sesgadas de Dinámica Molecular de Grano Grueso (DM-GG). Las trayectorias se analizaron mediante variables colectivas basadas en aperturas extracelular e intracelular, y se construyeron modelos de estados de Markov para cuantificar probabilidades de transición.

Las simulaciones permitieron observar transiciones espontáneas entre estados adyacentes del ciclo conformacional. El análisis identificó seis microestados, con transiciones predominantes entre conformaciones exofaciales, ocluidas y endofaciales, pero sin conectividad completa entre los estados extremos. La extensión de las simulaciones y cambios en el integrador no mejoraron la exploración conformacional, indicando que la limitación es energética más que temporal en ausencia de sustrato. Además, la metodología desarrollada es transferible a otros transportadores y constituye una base para estudios futuros con sustrato explícito.

## ABSTRACT

Molecular Dynamics (MD) simulations enable the study of biological processes at the atomic level; however, their application to conformational transitions is limited by sampling issues, particularly in transporter proteins with high energy barriers. Although non-equilibrium and biased equilibrium methods can characterize conformational pathways, they require external potentials and limit the interpretation of equilibrium dynamics.

In this work, conformational transitions of the alternating access model of the GLUT5 transporter from *Rattus norvegicus* (rGLUT5) were studied using unbiased Coarse-Grained Molecular Dynamics (MD-CG) simulations. The trajectories were analyzed using collective variables based on extracellular and intracellular gate openings, and Markov State Models were constructed to quantify transition probabilities.

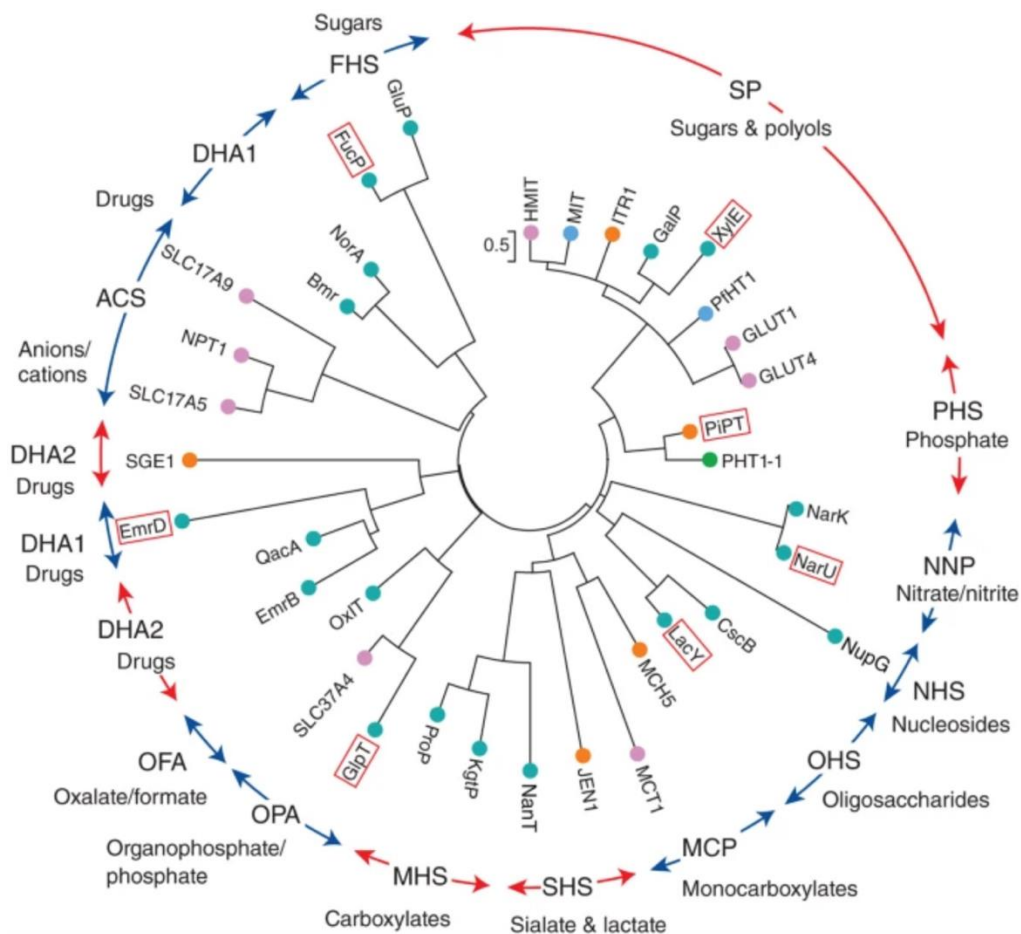
The simulations allowed the observation of spontaneous transitions between adjacent states of the conformational cycle. The analysis identified six microstates, with predominant transitions between outward-facing, occluded, and inward-facing conformations, but without full connectivity between the extreme states. Extending the simulation time and modifying the integrator did not improve conformational sampling, indicating that the limitation is energetic rather than temporal in the absence of substrate. Furthermore, the methodology developed is transferable to other transporters and provides a foundation for future studies incorporating explicit substrate.

## INTRODUCCIÓN

### 1. Superfamilia de Facilitadores Principales

El transportador de fructosa (GLUT5) pertenece a la Superfamilia de Facilitadores Principales, en inglés *Major Facilitator Superfamily* (MFS), una de las superfamilias de transportadores de membrana más antiguas y diversas que se conocen. Con un origen evolutivo estimado en más de 3.000 millones de años inferido a partir de análisis filogenéticos. MFS está presente en bacterias, arqueas y eucariotas, y transporta azúcares, polioles, fármacos, neurotransmisores, aminoácidos, péptidos, lípidos, iones orgánicos e inorgánicos, vitaminas, nucleósidos y nucleótidos (Figura 1) (Drew et al., 2021; Henderson et al., 2019; Law et al., 2008; Wang et al., 2020).

Los transportadores MFS operan mediante tres mecanismos: uniporte, en el que un soluto se desplaza a favor de su gradiente electroquímico; simporte, en el que el ingreso del sustrato se acopla al de un ion, generalmente un protón; y antiporte, en el que dos solutos se intercambian en direcciones opuestas (Deng et al., 2014; Huang et al., 2003; L. Sun et al., 2012). Como consecuencia, transportadores MFS se involucran en procesos fisiológicos diversos y al mismo tiempo, la disfunción de alguno de ellos se asocia a trastornos metabólicos, enfermedades hereditarias y progresión oncológica, siendo principales dianas farmacológicas. La Tabla 1 resume los principales transportadores MFS humanos con relevancia clínica descrita (Marger & Saier, 1993; Zhang et al., 2019).



**Figura 1. Filogenia de las proteínas pertenecientes a la superfamilia MFS.**

Los transportadores se identifican mediante sus acrónimos, y aquellos cuya estructura ha sido caracterizada con mayor profundidad se indican con recuadros rojos. Entre flechas se muestra el sustrato transportado. Los puntos de colores representan los organismos en los que se han identificado los distintos transportadores: verde azulado, bacterias; naranja, hongos; verde, plantas; azul, protozoos; y morado, humanos. Imagen extraída de (Henderson & Baldwin, 2013).

**Tabla 1. Los transportadores MFS humanos y sus funciones.**

\*Se indican ambas clasificaciones (SLC; Comité de Nomenclatura Génica HUGO) y (entre paréntesis) la utilizada por Base de Datos de Clasificación de Transportadores (TCDB).

Familia *	Nombre	Funciones principales	Enfermedades asociadas
SLC2 (2.A.1.1)	Familia de transportadores de azúcares (SP).	Mantener la homeostasis de la glucosa, fructosa y otros azúcares.	Diversos trastornos mendelianos (mutaciones en SLC2A1, SLC2A2 y SLC2A10), hiperuricemia y gota (SLC2A9) y diabetes tipo 2 (desregulación de SLC2A2 /GLUT2 y SLC2A4 /GLUT4). Implicado en el cáncer (especialmente SLC2A1).
SLC15 (2.A.17)	Familia de transportadores de péptidos (PepT o PTR) o familia de	Absorción dietética de péptidos (di- y tripéptidos), reabsorción renal de péptidos y posiblemente	Enfermedades inflamatorias intestinales (SLC15A1 y SLC15A4).

	transportadores de oligopéptidos acoplados a protones (POT).	mantenimiento de la homeostasis de los neuropéptidos cerebrales.	
<b>SLC16 (2.A.1.13)</b>	Familia de transportadores de monocarboxilato (MCT).	Mantener la homeostasis del lactato y de los cuerpos cetónicos, regular el pH celular y la captación de la hormona tiroidea en el cerebro.	Defecto del transportador de lactato eritrocítico (SLC16A1), hiperinsulinemia inducida por el ejercicio (SLC16A1) y retraso psicomotor ligado al cromosoma X (SLC16A8). Implicado en el cáncer.
<b>SLC17 (2.A.1.14)</b>	Familia de simportadores aniones/cationes.	Secreción de urato, carga de glutamato y ATP en vesículas sinápticas y exportación lisosomal de ácidos de azúcar.	Hiperuricemia (SLC17A1 y SLC17A3), enfermedades de almacenamiento lisosomal (SLC17A5) y sordera no sindrómica (SLC17A8).
<b>SLC18 (2.A.1.2)</b>	Transportadores vesiculares de monoamina y acetilcolina.	Carga de neurotransmisores monoaminas (por ejemplo, serotonina) y acetilcolina en vesículas sinápticas.	Los inhibidores de SLC18A2 se utilizan para controlar la Corea de Huntington. SLC18A1 podría estar relacionado con trastornos psicológicos.
<b>SLC19 (2.A.48)</b>	Familia de transportadores de folato/tiamina.	Entrega de folatos y tiamina (vitaminas B) a los tejidos sistémicos.	Resistencia al metotrexato (SLC19A1), anemia megaloblástica sensible a la tiamina (SLC19A2) y enfermedad de los ganglios basales sensible a la biotina (SLC19A3).
<b>SLC21 (2.A.60)</b>	Familia de transportadores de aniones orgánicos portadores de soluto (SLCO) o familia de polipéptidos transportadores de aniones orgánicos (OATP).	Captación en muchos tejidos diferentes de aniones orgánicos anfipáticos, como ácidos biliares, ciertas hormonas y xenobióticos.	Síndrome de Rotor (SLCO1B1 y SLCO1B3) y miopatía inducida por estatinas (SLCO1B1).
<b>SLC22 (2.A.1.19)</b>	Transportadores de aniones orgánicos (OAT), cationes orgánicos (OCT) y zwitteriones/cationes orgánicos (OCTN).	Transporte a través de epitelios de barrera de numerosos sustratos, incluyendo urato, hormonas, neurotransmisores, carnitina y xenobióticos. Funciones importantes en la depuración, pero también otras.	Enfermedad coronaria (SLC22A3), cáncer de próstata (SLC22A3), enfermedades inflamatorias del intestino (SLC22A4 y SLC22A5), artritis reumatoide (SLC22A4), deficiencia sistémica de carnitina (SLC22A5) e hiperuricemia (SLC22A11 y SLC22A12).
<b>SLC33 (2.A.1.25)</b>	Transportador de acetyl-CoA.	Importación de acetyl-CoA en el retículo endoplasmático para la modificación de diversas moléculas.	Papel putativo en la esclerosis lateral amiotrófica (SLC33A1).
<b>SLC37 (2.A.1.4)</b>	Familia de intercambiadores de azúcar-fosfato/fosfato (SPX).	Transporte de fosfatos de azúcar al retículo endoplasmático. Participa en el mantenimiento de la homeostasis de la glucosa en sangre y neutrófilos.	Enfermedades por almacenamiento de glucógeno tipo Ia y Ic (SLC37A4).

<b>SLC40 (2.A.100)</b>	Ferroportina.	Exportación de hierro celular al plasma sanguíneo.	Hemocromatosis tipo 4 (SLC40A1) y enfermedad de la ferroportina (SLC40A1).
<b>SLC43 (2.A.1.44)</b>	Abarca los transportadores de l-aminoácidos (LAT) LAT3 y LAT4.	LAT3 y LAT4 transportan aminoácidos neutros, principalmente de cadena ramificada. Su función in vivo no está clara.	No está claro si existen vínculos con enfermedades, pero SLC43A1 está regulado positivamente en formas de cáncer dependientes de andrógenos.
<b>SLC45 (2.A.2)</b>	Más comúnmente conocida como la familia SLC45.	Los transportadores de azúcar putativos, SLC45A1, pueden llevar a cabo el transporte de azúcar/H <sup>+</sup> en el bulbo raquídeo.	Albinismo óculo cutáneo tipo 4 (SLC45A2), implicado en el cáncer de próstata (SLC45A3).
<b>SLC46 (2.A.1.50)</b>	Familia de transportadores de folato.	Absorción de folato en la dieta y transporte de folato a través del plexo coroideo.	Mal absorción hereditaria de folato (SLC46A1).

A pesar de su relevancia, caracterizar a sus miembros, no es una tarea sencilla: la diversidad estructural, funcional y evolutiva de los transportadores MFS admite múltiples criterios de clasificación, cada uno con su propia lógica. Bases de datos como Pfam (*Protein Families Database*) agrupan proteínas según dominios estructurales conservados, mientras que sistemas de nomenclatura como el HGNC (*HUGO Gene Nomenclature Committee*) los organizan de acuerdo con su función y contexto fisiológico. Ambos criterios son válidos, pero no siempre coinciden. GLUT1 y GLUT5 ilustran bien esta tensión: comparten arquitectura, doce hélices transmembrana, difusión facilitada y dominio *MFS\_1*, pero se clasifican en familias distintas porque transportan sustratos distintos, glucosa y fructosa respectivamente (Deng et al., 2014; Nomura et al., 2015). Un tercer esquema, llamado TCDB (*Transporter Classification Database*), respaldado por la IUBMB (*International Union of Biochemistry and Molecular Biology*), integra criterios funcionales, mecánicos, topológicos y filogenéticos en una jerarquía de cinco niveles (V.W.X.Y.Z) comparable a la clasificación EC (*Enzyme Commission*) de las enzimas, en este sistema, cada transportador recibe un identificador numérico (TC#, *Transporter Classification number*) cuyos niveles describen, en orden, la clase de mecanismo general, la subclase mecánica, la familia

evolutiva, la subfamilia y el sistema específico definido por el sustrato (Saier et al., 2021). GLUT5 se clasifica como TC 2.A.1.1.13: Transportadores impulsados por potencial electroquímico (2), uniportador (A), superfamilia MFS (1), subfamilia de los transportadores de azúcares [*Sugar Porter*, SP (1)] y posición 13 dentro de ella.

Dentro de MFS, los SP (TC 2.A.1.1) son la familia con más miembros e incluyen transportadores centrales para la homeostasis energética, entre ellos los miembros de la familia GLUT (*Glucose Transporter*) (Khakpour et al., 2025; Koepsell, 2020; McComas et al., 2022; Mueckler & Thorens, 2013). La proteína de estudio de este trabajo GLUT5, es un miembro de la familia de los SP y estudios la han vinculado a obesidad, diabetes tipo 2 y sobreexpresión en tumores asociada a la reprogramación metabólica de células cancerosas (Barone et al., 2009; Groenendyk et al., 2022; Wang et al., 2020).

## **2. Características estructurales de los MFS y SP**

El conocimiento estructural de los transportadores MFS se construyó por etapas. Las primeras inferencias, desarrolladas en las décadas de 1980 y 1990, se basaron en perfiles de hidrofobicidad y patrones de distribución aminoacídica, y permitieron proponer una topología de múltiples hélices transmembrana (*Transmembrane Segment*, TM) sin validación experimental directa (Doerge et al., 1998). Esa validación llegó en 2003 con la resolución cristalográfica del simportador lactosa permeasa LacY y del antiportador de glicerol-3-fosfato GlpT, que establecieron la arquitectura de doce hélices TM como rasgo general de la superfamilia (Abramson et al., 2003; Huang et al., 2003). Desde entonces el número de estructuras disponibles creció hasta aproximadamente 250 (Figura 2A). Es importante señalar que este número no refleja proteínas distintas en igual proporción, ya que para un mismo transportador se han resuelto múltiples conformaciones a lo largo del ciclo de

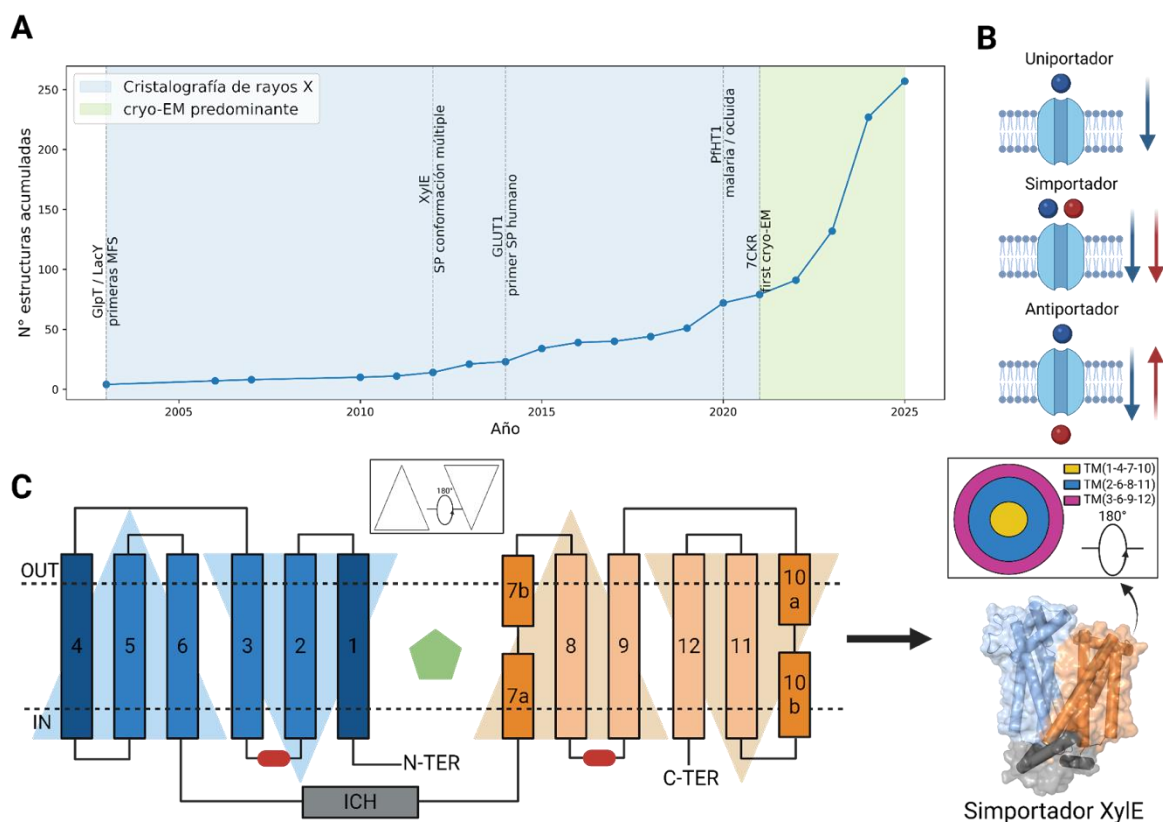
transporte y, en algunos casos, estructuras del mismo estado conformacional en complejo con distintos ligandos. El simportador bacteriano XylE es el ejemplo más representativo de este acervo estructural (Quistgaard et al., 2013; L. Sun et al., 2012).

El análisis de ese conjunto de estructuras reveló un plegamiento conservado. Los transportadores MFS presentan doce TM caracterizado por un plegamiento tipo hélices  $\alpha$ . Estos segmentos además están organizados en dos dominios pseudosimétricos de seis hélices cada uno: el dominio N-terminal (TM1–TM6) y el C-terminal (TM7–TM12) (Drew et al., 2021; Drew & Boudker, 2016; Radestock & Forrest, 2011). Cada dominio contiene dos repeticiones de tres TM relacionadas por una rotación de  $\sim 180^\circ$  paralela al plano de la membrana (Figura 2C).

Desde un punto de vista funcional, las hélices TM1, TM4, TM7 y TM10, llamadas hélices de cavidad, contribuyen a la formación del sitio de unión al sustrato. Las hélices TM2, TM5, TM8 y TM11, hélices basculantes, conforman las paredes laterales de la cavidad con una geometría curva característica tipo "plátano" y son el elemento móvil central durante los cambios conformacionales. Finalmente, las hélices TM3, TM6, TM9 y TM12, hélices de soporte, se localizan en los extremos del núcleo TM y median las interacciones con la bicapa lipídica circundante (Deng et al., 2015; Drew et al., 2021; Quistgaard et al., 2013; Wang et al., 2020).

Una característica relevante involucra a TM7 y TM10. Estas son hélices discontinuas del tipo hélice-péptido-hélice, una disposición energéticamente desfavorable pero funcionalmente relevante, presente también en NhaA, LeuT, GltPh, LacY y los transportadores GLUT (Abramson et al., 2003; Alleva et al., 2022; Deng et al., 2014; Latorraca et al., 2017; Screpanti & Hunte, 2007; Winkelmann et al., 2022).

Como última característica estructural conservada en proteínas MFS se identifica el motivo A ( $Gx_3[D/E][R/K]xGx[R/K][R/K]$ ), una secuencia conservada, abundante en aminoácidos ácidos y básicos, que estabiliza la estructura durante el ciclo de transporte mediante interacciones de tipo puente salino que conectan TM2–TM3 y TM8–TM9 (Drew et al., 2021).

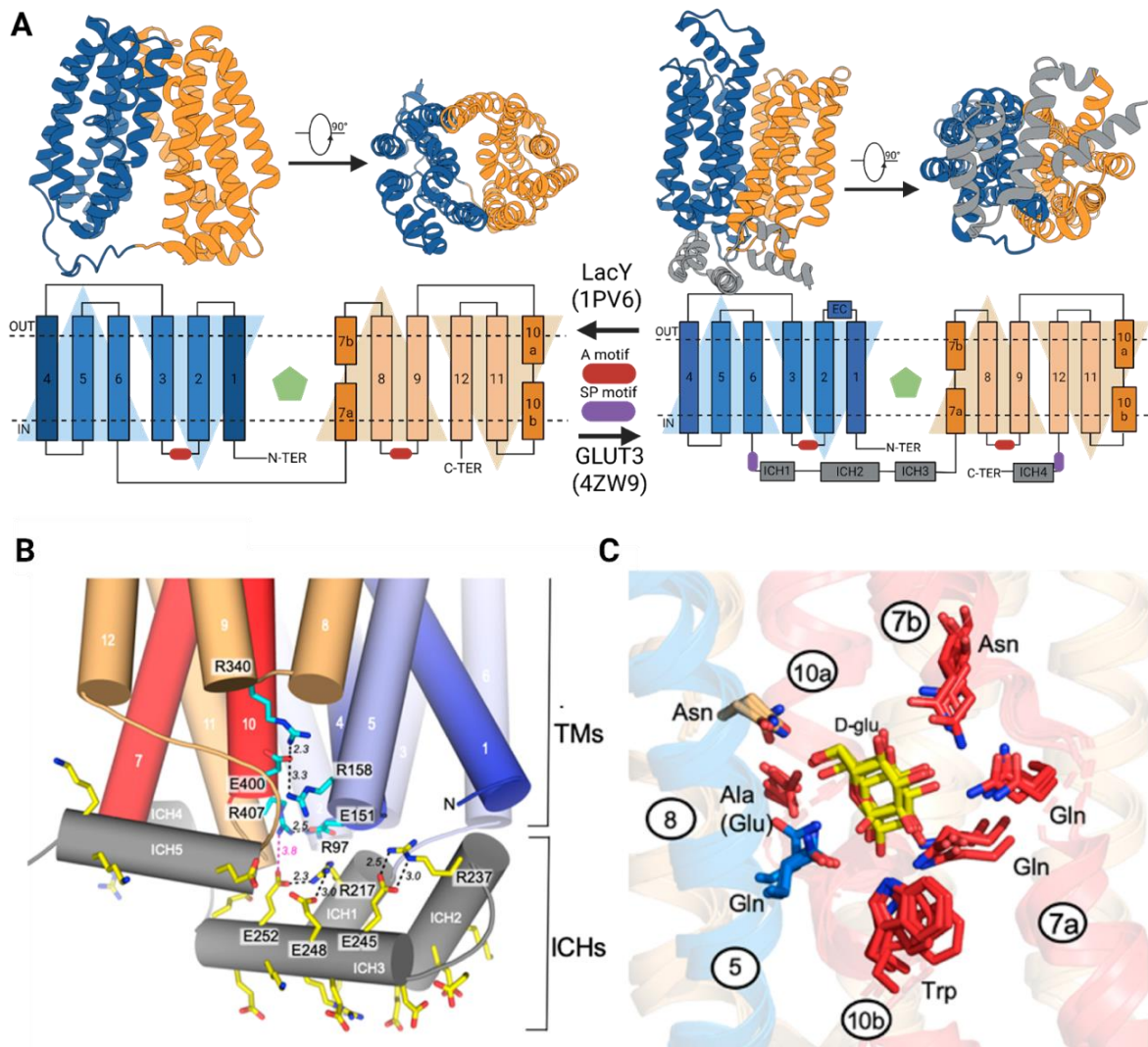


**Figura 2. Características estructurales de los transportadores MFS.**

(A) Línea temporal de hitos en la caracterización estructural de transportadores MFS. El gráfico inicia con las primeras estructuras cristalográficas de miembros de la superfamilia, el transportador de glicerol-3-fosfato GlpT y la permeasa de lactosa LacY en 2003. El número acumulado de estructuras incluye distintas conformaciones o complejos de un mismo transportador. (B) Mecanismos de transporte propuestos para proteínas MFS. (C) Izquierda: esquema del plegamiento canónico de un transportador MFS. El dominio N-terminal (azul) y el dominio C-terminal (naranja) contienen seis segmentos transmembrana cada uno, organizados como repeticiones estructurales invertidas (~180°, triángulos). Se muestra el dominio intracelular (ICH, gris), el motivo conservado A (rojo) y los segmentos TM1, TM4, TM7 y TM10 asociados al reconocimiento y transporte de sustrato. Derecha: estructura del simportador de protones y D-xilosa Xyle (PDB: 4GBY) como ejemplo de transportador MFS (Sun et al., 2012), vista desde el lado citoplasmático. Se destacan tres regiones estructurales: cavidad central (TM1-4-7-10, amarillo), hélices basculantes (TM2-5-8-11, azul) y hélices de soporte (TM3-6-9-12, violeta).

A pesar de esta topología conservada, existen diferencias estructurales relevantes entre las distintas familias de los MFS. Los transportadores SP se distinguen del resto de la superfamilia por la presencia de un haz de hélices intracelulares (ICH, *Intracellular Helices*) que forman el motivo SP en el espacio citoplasmático entre los dominios N- y C-terminal (Figura 3A) (Nomura et al., 2015). Este motivo se estabiliza mediante una red altamente conservada de puentes salinos con los segmentos TM adyacentes (Figura 3B), y se ha demostrado que mutaciones en residuos de las ICH alteran significativamente la cinética del ciclo de transporte (Deng et al., 2014; Drew et al., 2021; Leturque et al., 2009; Mueckler & Thorens, 2013; L. Sun et al., 2012).

El reconocimiento del sustrato en los SP está determinado por la conservación estructural del sitio de unión, no por identidad de secuencia, que en esta familia es inferior al 20%. Estudios de GLUT3 en complejo con D-glucosa a 1,5 Å mostraron que los elementos clave del sitio de unión se concentran en TM7 y TM10, donde residuos polares establecen una red de puentes de hidrógeno con los grupos hidroxilo del monosacárido (Sun et al., 2012). Este microambiente polar se conserva en GLUT5, XylE, PfHT1 y STP10, lo que refuerza la idea de que la conservación estructural de la cavidad, más que la similitud de secuencia gobierna el reconocimiento de sustratos (Nomura et al., 2015; Paulsen et al., 2019; Qureshi et al., 2020). Sin embargo, la conservación del sitio de unión no garantiza el transporte. En el 2012 se demostró que el transportador XylE tiene afinidades tanto para D-xilosa como D-glucosa, pero solo transporta al primer soluto (Sun et al., 2012). Este hecho demuestra que la selectividad funcional no puede reducirse a la geometría del bolsillo de unión, sino que involucra elementos estructurales adicionales asociados al movimiento coordinado de los TM durante el ciclo de transporte.



**Figura 3. Características estructurales de transportadores SP.**

(A) Representación esquemática del plegamiento de la permeasa de lactosa LacY (PDB: 1PV6) y del transportador de glucosa GLUT3 (PDB: 4ZW9). LacY (izquierda) no presenta hélices intracelulares definidas en la región citoplasmática, mientras que GLUT3 (derecha) muestra un conjunto de hélices intracelulares organizadas, característica distintiva de las proteínas SP. El motivo estructural conservado de los MFS se indica en rojo y el motivo característico de los SP en violeta. (B) Estructura de GLUT3 (PDB: 4ZW9) mostrando residuos que participan en la formación de puentes salinos en la región intracelular. Las cadenas laterales se muestran como palos (amarillo) y los residuos que interactúan con segmentos transmembrana en cian. (C) Comparación del sitio de unión al azúcar en transportadores unidos a D-glucosa: GLUT3 humano, STP10 vegetal (PDB: 6H7D), XylE de *E. coli* (PDB: 4GBZ) y PfHT1 de *P. falciparum* (PDB: 6RW3). Los residuos que establecen enlaces de hidrógeno con la D-glucosa están conservados, con la excepción del residuo en TM10, que corresponde a alanina en las estructuras comparadas y a glutamato en GLUT3.

### 3. Modelo de Acceso Alternante (MAA)

Durante gran parte del siglo XX, el estudio de los transportadores de membrana estuvo dominado por una visión estructural esencialmente estática, consecuencia directa del predominio de la cristalografía de rayos X como principal técnica para la resolución estructural (Cheng, 2018). Si bien esta metodología permitió definir con precisión atómica el plegamiento característico de los transportadores SP, cada estructura cristalográfica representa únicamente una fotografía instantánea correspondiente a un mínimo energético local (Costa et al., 2015; Patodia, 2014). Esta aproximación resulta insuficiente para describir procesos que dependen de diversos arreglos tridimensionales que caracterizan a los diferentes estados conformacionales durante el ciclo de transporte. Por lo tanto, la función de un transportador no se limita a la descripción de una única conformación (Kandt et al., 2007). En ausencia de información dinámica directa, los modelos mecanísticos debieron inferirse indirectamente a partir de un conjunto limitado de estados conformacionales.

Este escenario comenzó a modificarse con la acumulación progresiva de estructuras resueltas en distintos estados conformacionales. Transportadores como Xyle y GLUT1, cristalizados en más de una conformación, permitieron precisar la relación entre estructura, dinámica y función, revelando que estas proteínas no pueden describirse como entidades estáticas, sino como sistemas dinámicos cuya actividad depende de transiciones estructurales definidas que modulan las propiedades cinéticas y termodinámicas del transporte (Iglesias-Fernandez et al., 2017; Jia et al., 2020; Ke et al., 2017).

En este contexto, importante fue el aporte pionero de Jardetzky, quien describió por primera vez el Modelo de Acceso Alternante (MAA) (Jardetzky, 1966). Este modelo se basa en las siguientes proposiciones: (i) en transportadores debe existir una cavidad capaz de alojar al

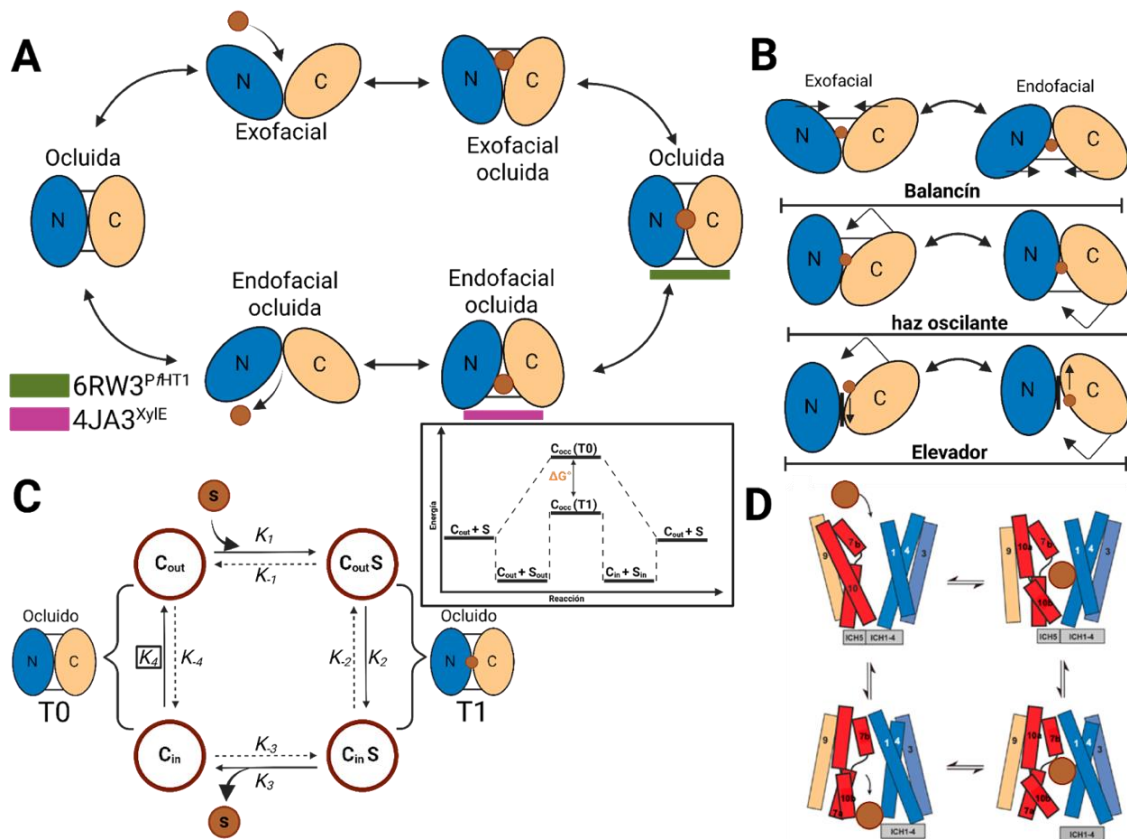
sustrato, (ii) el transportador se pliega en al menos dos conformaciones diferentes, de forma que la cavidad central se exponga alternativamente hacia uno u otro lado de la membrana y (iii) la afinidad del transportador al sustrato difiere entre las conformaciones que exhibe el transportador. De este modo, el MAA plantea que el transporte ocurre mediante una secuencia ordenada de cambios estructurales que alternan la accesibilidad del sitio de unión, evitando una exposición simultánea hacia ambos lados de la bicapa lipídica (Figura 4A) (Drew et al., 2021; Kazmier et al., 2017; Latorraca et al., 2017).

Actualmente se establece que el MAA se describe mediante tres movimientos estructurales o mecanismos: (i) el mecanismo de balancín o *rocker-switch*; (ii) el mecanismo de haz oscilante o *rocking-bundle*; y (iii) el mecanismo tipo ascensor o *elevator* (Drew et al., 2021; Drew & Boudker, 2016) (Figura 4B). Estos representan estrategias estructurales distintas para lograr la alternancia de accesibilidad del sitio de unión.

En el mecanismo *rocker-switch*, característico de los transportadores como GLUT y Xyle, el acceso alternante se logra mediante el balanceo coordinado de los dominios N- y C-terminal alrededor de un eje que atraviesa el sitio central de unión al sustrato (Wisedchaisri et al., 2014). En contraste, en el *rocking-bundle*, típico de transportadores acoplados a sodio con pliegue tipo LeuT, el dominio C-terminal experimenta cambios conformacionales de gran amplitud, mientras que el dominio N-terminal actúa como un andamiaje relativamente rígido (Kazmier et al., 2017; Loland, 2015). Finalmente, en el mecanismo *elevator*, observado en transportadores de glutamato con pliegue tipo GltPh, un dominio completo se desplaza verticalmente a través de la membrana transportando el sustrato, mientras el dominio complementario permanece esencialmente fijo (Moradi et al., 2015).

Independiente del mecanismo utilizado, el MAA involucra tanto reordenamientos estructurales globales como cambios locales altamente coordinados. Los cambios globales corresponden al desplazamiento concertado de los dominios N- y C-terminal, que permite la transición entre un estado exofacial (V), en el cual el sitio de unión se orienta hacia el exterior, y un estado endofacial ( $\Lambda$ ), orientado hacia el citosol. Durante estas transiciones, ambos dominios se comportan como cuerpos rígidos que se desplazan de forma sincronizada (Xiao et al., 2021) (Figura 4A).

Paralelamente, ocurren cambios locales en regiones específicas que estabilizan estados intermedios ocluidos y reorganizan las interacciones que definen el bolsillo de unión (Qureshi et al., 2020). Estas conformaciones metaestables son críticas para asegurar que el sitio de unión no quede simultáneamente accesible desde ambos lados de la membrana y, en la mayoría de los SP, regulan la dinámica del transportador limitando la amplitud de apertura de la cavidad central (Wisedchaisri et al., 2014) (Figura 4C). Dentro de estos cambios locales destacan los denominados residuos compuerta (*gating residues*), que establecen contactos transitorios entre los dominios N y C en estados conformacionales específicos. En los SP, existe consenso en que en la conformación exofacial los segmentos TM4 y TM10b restringen el acceso citoplasmático a la cavidad central, mientras que en la conformación endofacial son TM1 y TM7b los que limitan la apertura hacia el exterior (McComas et al., 2022). Estos cuatro segmentos conforman las hélices compuerta, un elemento estructural central para definir el estado conformacional del transportador (Figura 4D).



**Figura 4. Modelo de Acceso Alternante y mecanismos en transportadores SP.**

(A) Representación esquemática de las conformaciones descritas para transportadores SP dentro del modelo de acceso alternante. Se destacan las estructuras resueltas en conformación ocluida (verde) y endofacial ocluida (magenta). (B) Mecanismos conformacionales propuestos para transportadores de la superfamilia MFS: balancín (*rocker-switch*), haz oscilante (*rocking-bundle*) y elevador (*elevator*). Los transportadores de azúcares (SP) operan mediante el mecanismo de balancín. Las flechas indican la dirección predominante del movimiento de los dominios. (C) Diagrama de King–Altman del ciclo de transporte de glucosa mediado por GLUT1. Se muestran los estados exofacial, ocluido y endofacial, con y sin sustrato. La ruta cinéticamente dominante se indica con líneas continuas. El paso limitante corresponde a la transición del estado apo desde la conformación endofacial hacia la orientación exofacial tras la liberación de glucosa. El recuadro muestra el paisaje de energía libre asociado al ciclo, basado en datos experimentales de Lowe y Walmsley. (D) Esquema de los reordenamientos locales de las hélices TM7b y TM10b durante la transición entre estados exofacial y endofacial, destacando su papel como compuertas conformacionales en transportadores SP.

La discontinuidad de las hélices TM7 y TM10 confiere flexibilidad adicional a estos segmentos, facilitando el reordenamiento de la red de interacciones salinas que acompaña las transiciones conformacionales. De este modo, estas hélices no solo forman parte del núcleo del bolsillo de unión, sino que actúan como elementos móviles que acoplan la unión del sustrato a los cambios estructurales globales (Screpanti & Hunte, 2007). Se ha demostrado

que la unión del monosacárido desde el lado extracelular induce un desplazamiento hacia el interior de TM7b, promoviendo el cierre exofacial y la formación de un estado ocluido que asegura la coordinación de una sola molécula de sustrato por ciclo. Posteriormente, la disociación de TM10b respecto a TM4 permite la ruptura de la red de puentes salinos citoplasmáticos, favoreciendo la transición hacia el estado endofacial y la liberación del sustrato al citosol (Drew et al., 2021; McComas et al., 2022). El retorno del transportador al estado abierto exofacial ocurre a través de un estado vacío-ocluido, que constituye el paso cinéticamente limitante del ciclo y es significativamente más lento que la transición equivalente en presencia de sustrato (Figura 4C) (Mueckler & Thorens, 2013).

A pesar de su valor funcional, el estudio mecanístico del MAA presenta limitaciones importantes. La mayor parte de la evidencia disponible proviene de análisis comparativos entre estructuras cristalográficas que representan mínimos energéticos locales (Andersson et al., 2012; Sonoda et al., 2011). Además, la baja afinidad por el sustrato, la dependencia del entorno lipídico y la naturaleza vectorial del transporte dificultan el abordaje experimental directo de la dinámica conformacional (Drew et al., 2021; Kaback & Guan, 2019). En este contexto, las simulaciones de Dinámica Molecular (DM) emergen como una herramienta fundamental para explorar el paisaje conformacional, permitiendo acceder a información inaccesible mediante técnicas estructurales convencionales.

#### **4. Simulaciones de Dinámica Molecular**

Las simulaciones de Dinámica Molecular (DM) se fundamentan en la segunda ley de Newton, de forma que conocida la masa de cada átomo y las fuerzas que actúan sobre él en un instante dado, es posible calcular su aceleración y actualizar su posición y velocidad. Este proceso se repite de forma iterativa: el intervalo de tiempo entre dos actualizaciones

consecutivas se denomina paso de tiempo o *time step*, y representa el avance mínimo del reloj de la simulación en cada iteración (Karplus & McCammon, 2002).

Para actualizar la posición y velocidad de cada átomo en cada paso, el sistema necesita conocer las fuerzas que actúan sobre él. Estas fuerzas se derivan de la energía potencial del sistema, que depende de las posiciones relativas de todos los átomos, de forma que, a mayor energía potencial en una dirección, mayor es la fuerza que empuja al átomo en sentido contrario. La energía potencial se calcula mediante campos de fuerza moleculares, que son funciones matemáticas parametrizadas diseñadas para reproducir el comportamiento de sistemas químicos reales (González, 2011). Estas funciones incluyen términos que describen interacciones entre átomos enlazados, como estiramientos de enlace, cambios de ángulo y rotaciones de torsión, así como interacciones entre átomos no enlazados, como las fuerzas de Van der Waals y las interacciones electrostáticas. Dependiendo de la naturaleza química del sistema de interés, se han desarrollado distintos campos de fuerza con diferentes filosofías de parametrización (Tabla 2) (Case et al., 2023; Phillips et al., 2005; Van Der Spoel et al., 2005).

**Tabla 2. Campos de fuerzas aplicados a DM.**

AMBER = *Assisted Model Building and Energy Refinement*. GROMACS = *Groningen Machine for Chemical Simulations*. NAMD = *Nanoscale Molecular Dynamics*

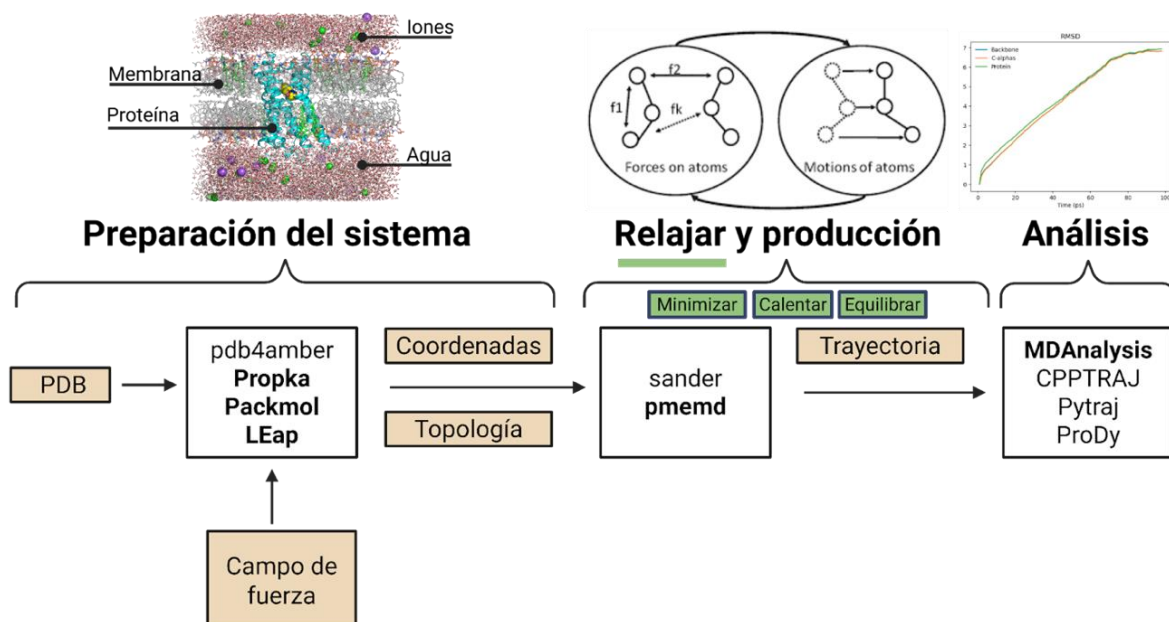
Campo de Fuerza	Software
<b>Proteínas</b>	
ff14SB	AMBER
CHARMM36	GROMACS
OPLS-AA	NAMD
GROMOS 54A7	GROMACS
<b>Lípidos</b>	
Lipid21	AMBER
CHARMM36	GROMACS

<b>Ácidos nucleicos</b>	
AMBER OL15	AMBER
CHARMM36	GROMACS
BSC1	NAMD
<b>Carbohidratos</b>	
GLYCAM_06j	AMBER
GROMOS 53A6 GLYC	GROMACS

Con las fuerzas conocidas, las ecuaciones de movimiento se integran numéricamente mediante algoritmos como *Verlet*, *Leapfrog* o *Velocity-Verlet*, que actualizan posiciones y velocidades en cada paso de tiempo (Fincham, 1992; Grønbech-Jensen & Farago, 2013). El tamaño del paso de tiempo no es arbitrario. Este debe ser suficientemente pequeño para capturar con precisión los movimientos más rápidos del sistema, que corresponden a las vibraciones de enlaces C-H y ocurren en escalas de los femtosegundos. Si el paso fuera demasiado grande, estas vibraciones se integrarían incorrectamente y la simulación se volvería numéricamente inestable. Por esta razón, el *time step* típico en DM atómica convencional se sitúa entre 1 y 2 femtosegundos.

En la práctica, llevar a cabo una simulación de DM consiste en un procedimiento secuencial en el que se construye de manera explícita el sistema a estudiar, tanto en su geometría como en su descripción fisicoquímica (Figura 5). El proceso comienza con una estructura tridimensional de referencia, a partir de la cual se determinan los estados de protonación de los residuos ionizables considerando el contexto estructural y las condiciones del medio. Para proteínas de membrana como GLUT5, esta etapa incluye además la inserción en una bicapa lipídica, solvatación con moléculas de agua y la adición de iones para neutralizar la carga total y reproducir condiciones iónicas fisiológicas. Una vez asignado el campo de fuerza, el sistema se relaja en tres etapas consecutivas: minimización de energía, para eliminar

solapamientos estéricos y tensiones geométricas; calentamiento controlado hasta la temperatura objetivo; y equilibración de las variables termodinámicas relevantes antes de la fase de producción, durante la cual el sistema evoluciona libremente de acuerdo con las ecuaciones clásicas del movimiento (Leach, 1997).



**Figura 5. Flujo de trabajo en una simulación de Dinámica Molecular.**

La figura resume las etapas principales de una simulación de dinámica molecular (DM): preparación del sistema, simulación y análisis de trayectorias. Durante la preparación, una estructura inicial en formato PDB se parametriza para generar la topología molecular de acuerdo con el campo de fuerza seleccionado e incorporar los componentes del sistema (proteína, bicapa lipídica, agua e iones). Posteriormente, el sistema se somete a minimización de energía, calentamiento y equilibración, seguido de la simulación de producción, en la que se integran las ecuaciones de movimiento para generar la trayectoria temporal. Finalmente, las trayectorias se analizan mediante herramientas especializadas. Los programas utilizados en esta tesis se indican en negro.

El resultado es una trayectoria que describe la posición y velocidad de cada átomo a lo largo del tiempo, muestreando el espacio de configuraciones accesibles al sistema. Cuando se aplica al estudio de cambios conformacionales, esta trayectoria representa la exploración del paisaje conformacional (Clarage et al., 1995; Patodia, 2014). Para un transportador SP, el valor de la simulación depende directamente de cuánto de ese paisaje logra explorar la trayectoria, lo que introduce el problema de muestreo que se describe en la sección siguiente.

## 5. El problema de muestreo

El problema de muestreo surge porque las escalas temporales accesibles a una simulación convencional son insuficientes para observar los fenómenos químicos de interés. El paso de integración de una simulación atomística se sitúa en el orden de los femtosegundos, impuesto por las vibraciones más rápidas del sistema, como las oscilaciones de enlaces C-H y los movimientos de ángulos de enlace. Estos movimientos describen bien la dinámica local, pero son irrelevantes para las transiciones conformacionales, ya que los desplazamientos de dominio que gobiernan el mecanismo de acceso alternativo son del orden de  $\sim 100$  Å y ocurren en escalas de microsegundos a milisegundos, entre siete y doce órdenes de magnitud más lentos que el paso de integración. Esto significa que la trayectoria cubre una fracción ínfima del espacio de configuraciones relevante, y los estados conformacionales de interés funcional simplemente no se visitan. (Orellana, 2019; McComas et al., 2022).

Que la trayectoria no cubra ese espacio tiene consecuencias desde la perspectiva de la termodinámica estadística. Caracterizar un sistema requiere que los promedios calculados a lo largo del tiempo de simulación, los promedios temporales, sean equivalentes a los promedios obtenidos considerando simultáneamente todos los estados accesibles ponderados por su probabilidad, los promedios del ensamble. Esta equivalencia, conocida como ergodicidad, exige que la trayectoria explore de forma representativa todo el espacio conformacional relevante. El resultado es que la distribución de Boltzmann, que asigna probabilidades exponencialmente menores a los estados de mayor energía, no se muestrea de forma representativa, y los estados funcionalmente esenciales permanecen inaccesibles no por ausencia de tiempo de cómputo, sino por la naturaleza estocástica e improbable de las transiciones que los conectan (Orellana, 2019; Marzinek et al., 2014; Clarage et al., 1995).

El cambio conformacional entre los estados exofacial, ocluido y endofacial del MAA ejemplifica este desafío con claridad. Alcanzar el estado ocluido requiere el reordenamiento cooperativo de TM7b y TM10b, y la conversión hacia ese estado en ausencia de sustrato es aproximadamente cien veces más lenta que en su presencia (Latorraca et al., 2017; McComas et al., 2022). Esta brecha no se resuelve extendiendo la simulación, sino adoptando estrategias de muestreo específicas, descritas en la sección siguiente.

## 6. Técnicas de mejoramiento de muestreo

Abordar el problema de muestreo exige estrategias que permitan al sistema explorar regiones del paisaje conformacional inaccesibles por DM convencionales o atomísticas. Las soluciones disponibles se dividen en dos grandes categorías, según si introducen o no una fuerza externa dependiente del tiempo (Marzinek et al., 2014).

Las simulaciones fuera del equilibrio aplican fuerzas o potenciales que varían en el tiempo para forzar al sistema a recorrer regiones del paisaje que no visitaría espontáneamente (Kalyaanamoorthy & Chen, 2014; Yang et al., 2019). La dinámica molecular dirigida (SMD, *Steered Molecular Dynamics*) es el ejemplo más utilizado, en donde un potencial armónico móvil empuja al sistema a lo largo de una coordenada de reacción definida *a priori*, permitiendo inducir transiciones conformacionales (Jensen et al., 2007). La dinámica molecular dirigida hacia un objetivo estructural (TMD, *Targeted Molecular Dynamics*) opera de manera análoga, pero usando como referencia la desviación global respecto a una conformación final conocida, lo que la hace especialmente útil cuando se dispone de estructuras exofacial y endofacial bien caracterizadas (McComas et al., 2022; Schlitter et al., 1994).

Las técnicas de equilibrio sesgado resuelven el problema de otra manera. En lugar de empujar al sistema, modifican el paisaje energético para que las transiciones ocurran espontáneamente con mayor probabilidad (Kästner, 2011; Miao et al., 2015). El sistema evoluciona en equilibrio con un Hamiltoniano modificado, y las propiedades termodinámicas del sistema original se recuperan posteriormente mediante reponderación estadística. La dinámica molecular acelerada (aMD) aplica este principio, elevando los mínimos del paisaje para reducir el atrapamiento conformacional (Do et al., 2022). La metadinámica lo aplica de forma adaptativa sobre coordenadas colectivas predefinidas, acumulando un sesgo en las regiones ya exploradas para forzar la exploración de nuevos estados; su ventaja es que permite reconstruir perfiles de energía libre, aunque su resultado depende críticamente de que las coordenadas colectivas describan bien el proceso (Ibrahim & Clark, 2019; Ray & Parrinello, 2023). El muestreo *umbrella* aplica potenciales armónicos estáticos en distintos valores de una coordenada de reacción, permitiendo reconstruir perfiles de energía libre con alta resolución; su limitación es que requiere definir esa coordenada a priori y no puede descubrir rutas no contempladas en el diseño (Govind Kumar et al., 2023; Kulshrestha et al., 2022).

Independientemente de la estrategia de muestreo utilizada, las trayectorias obtenidas requieren herramientas de reconstrucción que permitan extraer información sobre el paisaje conformacional, las rutas de transición y las poblaciones relativas de los estados. Estas herramientas utilizan el espacio conformacional ya explorado sin generar trayectorias adicionales, y son compatibles tanto con simulaciones fuera del equilibrio como con técnicas sesgadas. Su aplicación enfrenta sin embargo un problema de dimensionalidad. Un sistema de  $N$  átomos posee  $3N$  grados de libertad, de modo que describir su dinámica completa es computacionalmente intratable y conceptualmente opaco. La solución es proyectar esa

dinámica sobre un conjunto reducido de coordenadas colectivas (*CVs, collective variables*), variables de baja dimensionalidad que capturan los grados de libertad lentos y funcionalmente relevantes del sistema, como distancias entre compuertas, ángulos de hélices TM o radios de giro, descartando las fluctuaciones rápidas que no contribuyen a la transición de interés (Mitrovic et al., 2023; Qureshi et al., 2020).

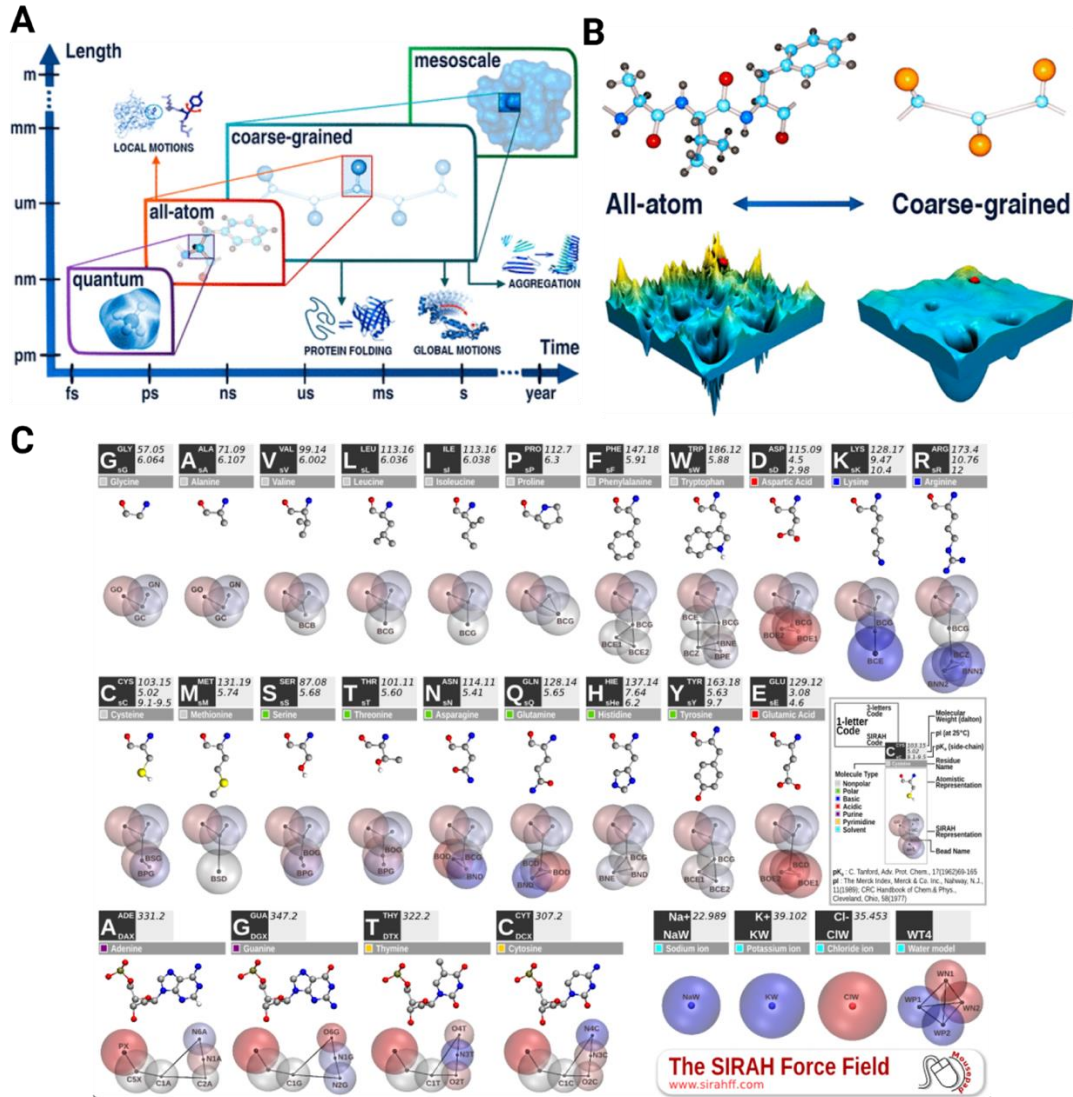
Sobre ese espacio reducido se construyen los métodos de reconstrucción propiamente dichos. El *string method* identifica la ruta de transición más probable entre dos estados como una trayectoria continua en el espacio de CVs, optimizada iterativamente; su variante SMwST (*string method with swarms of trajectories*) evita el cálculo explícito de gradientes usando la evolución promedio de múltiples trayectorias cortas no sesgadas, preservando la dinámica microscópica del sistema (Maragliano et al., 2006; Roux, 2021). Los modelos de estados de Markov (MSM) discretizan el espacio conformacional en estados metaestables mediante clustering en el espacio de CVs y describen la dinámica mediante probabilidades de transición entre ellos, permitiendo estimar poblaciones relativas, escalas temporales y rutas dominantes incluso cuando las trayectorias individuales no capturan transiciones completas (Malmstrom et al., 2014; Nagel & Stock, 2023). Ambos métodos maximizan la información extraíble del muestreo disponible, razón por la cual se combinan frecuentemente con las estrategias descritas anteriormente.

## **7. Simulaciones de Dinámica Molecular de Grano Grueso**

Las simulaciones de Dinámica Molecular de Grano Grueso (DM-GC) constituyen una alternativa a la DM atómica convencional orientada a reducir la complejidad computacional de sistemas biomoleculares de gran tamaño (Garay et al., 2024) (Figura 6A). En este enfoque, múltiples átomos se agrupan en partículas efectivas denominadas *beads*,

cada una representando un conjunto de átomos con propiedades fisicoquímicas colectivas similares, lo que reduce drásticamente el número de grados de libertad del sistema (Barrera et al., 2019).

Al eliminar las fluctuaciones rápidas de naturaleza atómica y conservar únicamente los movimientos colectivos relevantes, el potencial efectivo resultante es más suave, permitiendo integrar la dinámica con pasos de tiempo de 20 a 40 fs, entre 10 y 20 veces mayores que en DM atómica. Dado que el potencial suavizado acelera la dinámica, cada nanosegundo simulado en grano grueso equivale entre 4 y 10 nanosegundos de DM atómica, lo que permite acceder a escalas de cientos de nanosegundos a microsegundos inaccesibles para la DM atómica convencional (Dans et al., 2010) (Figura 6B). Es importante aclarar que las simulaciones DM-GG no constituyen simulaciones fuera del equilibrio ni métodos de equilibrio sesgado: son una aproximación multiescala que genera dinámica emergente dentro de un potencial efectivo de menor resolución y sin perturbaciones externas.



**Figura 6. Representación topológica Grano Grueso.**

(A) Rangos de aplicación de distintas aproximaciones de simulación según su nivel de resolución y escala temporal. Al disminuir el detalle molecular, aumenta la escala espacial y temporal accesible computacionalmente. (B) Comparación conceptual del paisaje energético en modelos atomísticos y de grano grueso. Los modelos atomísticos presentan múltiples mínimos locales separados por barreras energéticas elevadas, mientras que los modelos de grano grueso suavizan el paisaje energético efectivo, reduciendo dichas barreras y facilitando la exploración del espacio conformacional. (C) Esquema de mapeo desde una representación atomística a una representación grano grueso. Cada aminoácido se describe mediante un conjunto reducido de *beads* GG posicionados sobre átomos estructuralmente representativos. Imágenes adaptadas de Barrera et al. (2019) y Kmiecik et al. (2016).

Un aspecto central en las simulaciones DM-GG es el nivel de resolución con que se describe el sistema, lo que depende directamente del número de *beads* utilizados en el esquema de mapeo. Este nivel de resolución se denomina granularidad: a mayor granularidad, más

átomos se agrupan en cada *bead* y menor es el detalle estructural conservado; a menor granularidad, la representación se acerca más a la descripción atomística. No existe un esquema de mapeo universal, ya que la granularidad óptima depende tanto de la naturaleza de la macromolécula como del campo de fuerza empleado y de la pregunta biológica a responder (Kmiecik et al., 2016; Riniker et al., 2012). Un ejemplo ilustrativo es el agua, ya que puede ser modelada como un único *bead* o mediante dos *beads* diferenciados que distinguen el oxígeno de los hidrógenos, según el balance deseado entre eficiencia computacional y fidelidad estructural (Hadley & McCabe, 2012).

Esta ganancia en escala temporal tiene sin embargo un costo: la reducción de grados de libertad compromete la precisión de propiedades que dependen del detalle atomístico, como energías de interacción específicas, puentes de hidrógeno o la dinámica de cadenas laterales. Por esta razón, los modelos de grano grueso requieren validación contra datos experimentales o simulaciones atomísticas de referencia antes de poder interpretarse con confianza, lo que define el marco de uso responsable de esta aproximación (Kmiecik et al., 2016; Dans et al., 2010).

Independientemente del nivel de granularidad adoptado, todas las simulaciones MD-GC permiten estudiar procesos biomoleculares colectivos inaccesibles para la DM atomística, como grandes cambios conformacionales, reorganización de dominios proteicos o ciclos funcionales completos en proteínas de membrana (Kmiecik et al., 2016). En el contexto del estudio de transportadores MFS, el paisaje energético suavizado facilita la exploración espontánea de estados conformacionales globales que raramente se visitan en simulaciones atomísticas dentro de ventanas temporales razonables, abordando directamente el problema de muestreo.

## 8. Campo de fuerza SIRAH

Los campos de fuerza atomísticos describen cada átomo del sistema de forma explícita, incluyendo sus cargas parciales, radios de van der Waals y términos de enlace, lo que permite una representación química de alta fidelidad, pero a un costo computacional proporcional al número de átomos del sistema. Sin embargo, debido a la filosofía de las DM-GG los campos de fuerza no deben especificar interacciones atomísticas y más bien redefinen el nivel de descripción del sistema mediante los *beads* (Maisuradze et al., 2010; Marrink et al., 2007).

Los campos de fuerza de grano grueso no constituyen una familia homogénea, sino un espectro de aproximaciones que difieren en cómo se agrupa la información atómica y cómo se calibran las interacciones resultantes. El primer criterio de distinción es la granularidad, es decir, cuántos átomos se agrupan en cada *bead*: a mayor granularidad, mayor ganancia computacional pero menor detalle estructural conservado. El segundo es la estrategia de parametrización, que define con qué datos de referencia se calibran las interacciones entre *beads*, ya sean datos termodinámicos experimentales, simulaciones atomísticas de referencia, o ambos. De esta elección depende en gran medida la transferibilidad del campo de fuerza, es decir, su capacidad de describir sistemas distintos a aquellos con los que fue parametrizado sin necesidad de recalibración.

MARTINI ilustra un extremo de este espectro: agrupa en promedio cuatro átomos pesados por *bead* y calibra sus interacciones reproduciendo energías libres de partición entre fases polar y apolar, lo que le otorga amplia transferibilidad para simular membranas y proteínas de membrana (Marrink et al., 2007). Sin embargo, esta estrategia implica que las estructuras secundarias deben mantenerse mediante restricciones externas, lo que limita su aplicabilidad en estudios donde la dinámica conformacional de la proteína es el objeto central de análisis.

SIRAH surge precisamente como una alternativa diseñada para preservar esa información estructural desde el propio esquema de mapeo, sin depender de restricciones adicionales.

SIRAH (*Southamerican Initiative for a Rapid and Accurate Hamiltonian*) es un campo de fuerza de grano grueso diseñado para simulaciones de dinámica molecular de sistemas biomoleculares (Darré et al., 2015). A diferencia de MARTINI, busca preservar la información estructural esencial del sistema manteniendo al mismo tiempo las ventajas computacionales del grano grueso (Barrera et al., 2019; Souza et al., 2021).

El esquema de mapeo de SIRAH no es uniforme, sino que se adapta a la naturaleza fisicoquímica de cada residuo, colocando *beads* en posiciones estructural o funcionalmente relevantes (Klein et al., 2023). En las proteínas, la cadena principal se representa con tres *beads* por residuo correspondientes al nitrógeno del enlace peptídico, el carbono C $\alpha$  y el oxígeno carboxílico, lo que permite una traducción directa entre las representaciones atomística y CG y preserva las restricciones geométricas necesarias para la estabilidad de estructuras secundarias (Darré et al., 2015). Las cadenas laterales se modelan con distintos niveles de simplificación según sus propiedades químicas: los residuos hidrofóbicos alifáticos se representan mediante *beads* neutros cuyo tamaño reproduce el volumen excluido real; los aromáticos, con *beads* de menor diámetro para capturar interacciones de apilamiento; los polares conservan explícitamente los grupos funcionales relevantes para puentes de hidrógeno; y los cargados se representan mediante distribuciones de carga parcial equivalentes a los modelos atómicos (Figura 6C).

Una característica distintiva de SIRAH es su cobertura molecular: incluye parametrización para proteínas, ácidos nucleicos, lípidos, agua e iones, lo que lo convierte en un campo de fuerza autosuficiente para simular sistemas de membrana completos (Garay et al., 2024;

Machado et al., 2019). Esta cobertura es relevante para el presente trabajo, ya que simular GLUT5 en condiciones fisiológicamente representativas requiere describir simultáneamente la proteína, la bicapa lipídica y el entorno acuoso iónico. Adicionalmente, SIRAH incorpora aceleración por GPU en los principales paquetes de simulación, GROMACS y AMBER, y dispone de un conjunto de herramientas de análisis, SIRAH Tools, que facilitan la preparación del sistema, la visualización de trayectorias GG y la reconstrucción de coordenadas atomísticas a partir de la representación de grano grueso (Machado & Pantano, 2016).

## 9. Planteamiento

GLUT5 es un uniportador de la superfamilia MFS, perteneciente a la familia SP, y constituye el principal transportador específico de fructosa en humanos (Nomura et al., 2015). Se expresa predominantemente en el intestino delgado, aunque también está presente en cerebro, tejido adiposo, riñón, testículos y músculo esquelético (Barone et al., 2009; Mueckler & Thorens, 2013). Alteraciones en su expresión se han asociado a obesidad, diabetes tipo 2 y reprogramación metabólica en tumores (Groenendyk et al., 2022), lo que posiciona a GLUT5 como un modelo relevante para el estudio mecanístico del transporte de fructosa.

El plegamiento de GLUT5 consiste en 12 TM organizadas en dos dominios pseudosimétricos que alternan la exposición del sitio de unión mediante un mecanismo rocker-switch, con transiciones entre estados exofacial, ocluido y endofacial gobernadas por movimientos globales de dominio y reorganizaciones locales en TM7 y TM10 (Drew et al., 2021). Estas transiciones ocurren sobre un paisaje de energía libre con múltiples mínimos separados por barreras elevadas, donde los estados ocluidos actúan como cuellos de botella cinéticos inaccesibles para simulaciones atomísticas convencionales (McComas et al., 2022).

Estudios previos han recurrido a simulaciones fuera del equilibrio para inducir estas transiciones, caracterizar rutas conformacionales e identificar el rol modulador del sustrato sobre el paisaje energético (Figura 7) (McComas et al., 2022; Mitrovic et al., 2023; Xu et al., 2025). Sin embargo, estas aproximaciones introducen fuerzas externas que limitan la interpretación termodinámica directa, y el ciclo conformacional completo de GLUT5 no ha sido explorado de manera espontánea bajo un Hamiltoniano no modificado.

Las simulaciones DM-GG constituyen una estrategia complementaria para abordar esta limitación: al reducir los grados de libertad y suavizar el paisaje energético, permiten explorar espontáneamente transiciones globales sin introducir fuerzas externas (Borges-Araújo et al., 2023; Klein et al., 2023; Kmiecik et al., 2016). Un mayor muestreo conformacional no implica por sí mismo comprensión mecánica; por ello, este trabajo integra MD-GC con modelos de estados de Markov para describir cuantitativamente las transiciones entre estados metaestables y sus escalas temporales asociadas.

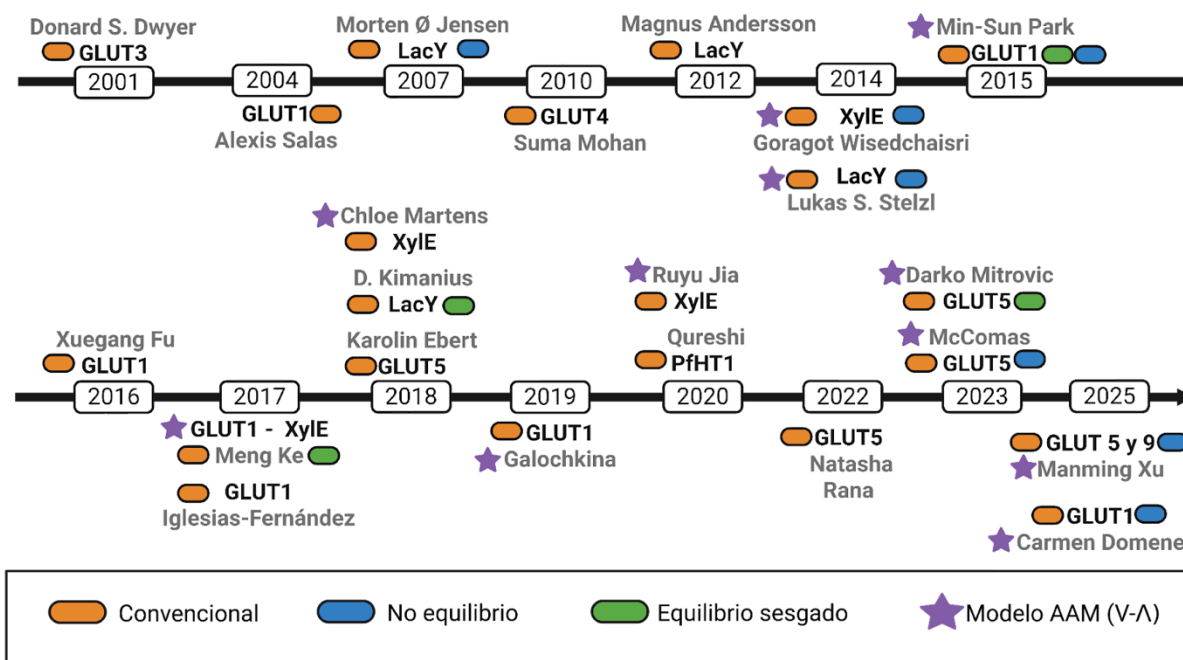


Figura 7. Línea temporal DM aplicadas a MFS.

Línea de tiempo de estudios representativos de DM aplicados a transportadores de azúcares MFS. DM en equilibrio (naranja), fuera del equilibrio (azul) y muestreo sesgado (verde). Los trabajos que contribuyen a la formulación o validación de MAA se indican con una estrella púrpura. Referencias visitadas (Andersson et al., 2012; Domene et al., 2025; Dwyer, 2001; Ebert et al., 2018; Fu et al., 2016; Galochkina et al., 2019; Iglesias-Fernandez et al., 2017; Jensen et al., 2007; Jia et al., 2020; Ke et al., 2017; Khakpour et al., 2025; Kimanius et al., 2018; Liu et al., 2024; Martens et al., 2018; McComas et al., 2022; Mitrovic et al., 2023; Mohan et al., 2010; Park, 2015; Qureshi et al., 2020; Rana et al., 2022; Salas-Burgos et al., 2004; Stelzl et al., 2014; Wisedchaisri et al., 2014; Xu et al., 2025).

## HIPÓTESIS

La hipótesis planteada es:

Simulaciones de Dinámica Molecular de Grano Grueso (DM-GG) no sesgadas permitirán capturar transiciones conformacionales del modelo de acceso alternante en el uniportador de fructosa de *Rattus norvegicus* (rGLUT5).

## OBJETIVOS

**Objetivo general:** Muestrear las transiciones conformacionales asociadas al modelo de acceso alternante del transportador GLUT5 de *Rattus norvegicus* mediante DM-GG.

**Objetivos específicos:**

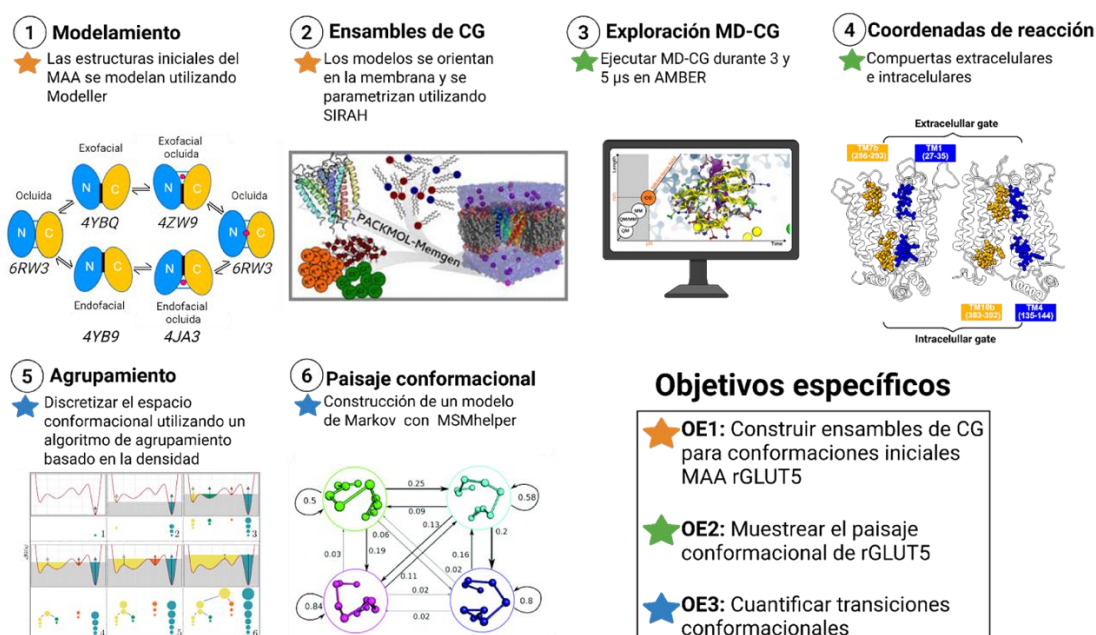
**Objetivo específico 1:** Construir ensamblajes termodinámicos (NPT) de grano grueso para rGLUT5 embebida en membrana de todas las conformaciones iniciales del modelo de acceso alternante.

**Objetivo específico 2:** Explorar el espacio conformacional del modelo de acceso alternante en rGLUT5 mediante DM-GG.

**Objetivo específico 3:** Determinar probabilidades de transición en el espacio conformacional explorado mediante métodos estocásticos.

## MATERIALES Y MÉTODOS

Para observar transiciones conformacionales en rGLUT5, se empleó una estrategia que combina simulaciones DM-GG con herramientas estocásticas basadas en modelos de estados de Markov (MSM). La metodología a continuación detallada involucra la construcción de modelos en sus conformaciones iniciales (exofacial, exofacial ocluida, ocluida, endofacial ocluida y endofacial), el ensemble de dichas conformaciones en topología de grano grueso, la generación de datos para el muestreo conformacional, identificación de microestados a partir de *clustering* y construcción de MSM para analizar las transiciones entre microestados para evaluar probabilidades de transición en el espacio conformacional explorado (Figura 8)



**Figura 8. Esquema metodológico propuesto.**

**(1) Modelado estructural:** Se generaron estructuras iniciales completas con Modeller en las conformaciones exofacial (4YBQ), exofacial ocluida (4ZW9), ocluida (6RW3), endofacial ocluida (4JA3) y endofacial (4YB9) (Tabla 4). **(2) Ensemble coarse-grained (CG):** Las estructuras se orientaron en membrana y se parametrizaron con SIRAH y PACKMOL-Memgen. **(3) Simulaciones DM-GG:** Cada conformación inicial se simuló durante 3 y 5  $\mu$ s en Amber. **(4) Variables colectivas:** Se emplearon descriptores de apertura extracelular e intracelular para monitorear los estados exofacial y endofacial, respectivamente. **(5) Agrupamiento:** El espacio conformacional se discretizó mediante un algoritmo basado en densidad, donde cada clúster representa un microestado del modelo de acceso alternante. **(6) Modelo de Markov:** Se construyó un modelo de Markov oculto con MSMhelper para analizar transiciones entre microestados y reconstruir el ciclo de transporte.

## 1. Modelamiento de conformaciones

El MAA involucra el cambio conformacional del transportador GLUT5 en cinco conformaciones: exofacial, exofacial ocluida, ocluida, endofacial ocluida y endofacial. No obstante, para ninguna isoforma del transportador GLUT5 existe información de coordenadas tridimensionales de sus átomos en los estados conformacionales, por lo que fue necesario el modelamiento comparativo de cada conformación.

Todos los modelos empleados en este estudio se construyeron a partir de la secuencia del transportador GLUT5 de *Rattus norvegicus* (UniProt: P43427), considerando los aminoácidos E7 al V480 y la información estructural correspondiente a los estados iniciales del modelo acceso alternante en miembros de los transportadores de azúcares (SP) (Tabla 3).

**Tabla 3. Estructuras utilizadas para modelamiento comparativo.**

\*CD = ión de cadmio, LU = ión de lutecio, GLC = Alfa-D-glucopiranososa, BGC = Beta-D-glucopiranososa, XYP = Beta-D-xilopiranososa, BNG = Nonil beta-D-glucopiranosido, Fv anticuerpo = Fragmento variable, OLC = (2R)-2,3-dihidroxiopropilo (9Z)-octadec-9-enoato. \*\* NA = No Aplica.

PDB	Nombre	Conformación	Organismo	Resolución	HETATM	Regiones faltantes
4YBQ	GLUT5	exofacial	<i>H. sapiens</i>	3.27 Å	*Fv anticuerpo	giro TM1-TM2
4ZW9	GLUT3	exofacial ocluida	<i>H. sapiens</i>	1.50 Å	*OLC *GLC *BGC	**NA
6RW3	<i>hexose transporter</i>	ocluida	<i>P.falciparum</i>	3.6 Å	*GLC *BGC	ICH 4
4JA3	<i>symporter</i> Xyle	endofacial ocluida	<i>E. coli</i>	3.8 Å	*CD *LU	ICH 4 ICH 5
4YB9	GLUT5	endofacial	<i>B.taurus</i>	3.20 Å	**NA	ICH 5

Para los modelos generados se utilizaron alineamientos estructurales entre rGLUT5 (4YBQ) y la estructura objetivo con *TM-align* y luego, el alineamiento resultante se formateó según las especificaciones de la extensión PIR y se generaron 50 modelos mediante refinamiento rápido utilizando MODELLER para cada conformación (Webb & Sali, 2016; Zhang & Skolnick, 2005).

Se prefirió la alineación estructural por sobre el alineamiento de secuencias, ya que garantiza la conservación de características estructurales. Además, la alineación estructural resultó útil cuando era necesario modelar determinadas partes de una región, como en la región TM1b/bucle TM1-TM2 en estado abierto hacia el exterior (descrito con más detalle abajo), o cuando se necesitaban varias plantillas para una cobertura completa, como en los estados ocluido, ocluido hacia el interior o abierto hacia el interior, ya que la mayoría de la información estructural carente en los transportadores corresponde al plegamiento intracelular (ICH) (Tabla 3).

Finalmente, la selección de los mejores modelos se realizó en función de la cuantificación de la energía DOPE (*Discrete Optimized Protein Energy*), un criterio estadístico que permite evaluar la calidad estructural de modelos tridimensionales al estimar su energía potencial. Los modelos con menor puntuación DOPE fueron seleccionados como candidatos para cada conformación y sometidos a una validación estructural mediante MolProbity (Davis et al., 2018), herramienta que evalúa la calidad geométrica del modelo a través de parámetros como la distribución de ángulos  $\phi/\psi$  en el gráfico de Ramachandran, la conformación de rotámeros de cadenas laterales, desviaciones de enlaces covalentes y ángulos de valencia, y la geometría del *backbone* mediante el análisis CaBLAM. Adicionalmente, se llevó a cabo una revisión

visual detallada con ChimeraX de las características topológicas de los modelos comparativos (Meng et al., 2023).

Dicho lo anterior, los modelos comparativos para cada conformación se modelaron de la siguiente manera:

- **Conformación exofacial:**

Para modelar rGLUT5 en conformación exofacial, se utilizó la estructura cristalográfica 4YBQ (Nomura et al., 2015). En primer lugar, se eliminó el fragmento Fv del anticuerpo estabilizador presente en el cristal, dado que no forma parte del transportador nativo y su presencia interfiere con posteriores simulaciones de dinámica molecular.

La estructura rGLUT5 (4YBQ) presenta una región no resuelta en la cadena A que abarca el extremo C-terminal de la hélice TM1b y el bucle TM1–TM2 (residuos N39–N60). Para modelar esta región faltante, se empleó como plantilla el segmento equivalente de GLUT3 (PDB: 4ZWC), cuya región TM1–TM2 (residuos Y26–P53) se encuentra completamente resuelta (Quistgaard et al., 2013).

Durante la alineación estructural, fue necesario incorporar cinco residuos adicionales de la plantilla de GLUT3 (P54–L58) para garantizar la continuidad geométrica entre el loop modelado y el inicio de la hélice TM2 resuelta en GLUT5 (residuos I33–T37 de 4YBQ).

- **Conformación exofacial ocluida:**

Para el modelamiento de la conformación exofacial ocluida se utilizó la información estructural de GLUT3 humano como plantilla (4ZW9). Afortunadamente hGLUT3 presenta gran resolución (1.50 Å), por lo que todos los plegamientos estructurales estaban presentes en la estructura.

- **Conformación ocluida:**

Para el modelamiento de la conformación completamente ocluida se utilizó la información estructural de PfHT1 como plantilla (6RW3), ya que era la única estructura disponible en esta conformación (Qureshi et al., 2020). Para modelar la hélice intracelular ICH5, que falta parcialmente en la estructura PfHT 6RW3, se modeló utilizando residuos de F467-V480 de rGLUT5 4YBQ.

- **Conformación endofacial ocluida:**

Para el modelamiento de la conformación endofacial parcialmente ocluida se utilizó la información estructural del simportador Xyle (4JA3) como plantilla, ya que era la única estructura disponible en esta conformación (Quistgaard et al., 2013). Para modelar las hélices intracelulares faltantes (ICH4 e ICH5) se utilizaron los residuos A258-R274 y F467-V480 de rGLUT5 4YBQ, respectivamente. Además, se incluyeron en el alineamiento los residuos D253-K257 y M457-T466 para facilitar la continuidad de ICH4 entre ICH3 y TM7, así como la continuidad de ICH5 desde TM12, respectivamente.

- **Conformación endofacial:**

Para el modelamiento de la conformación completamente endofacial se utilizó la información estructural de GLUT5 bovino (4YB9) como plantilla (Nomura et al., 2015). Para modelar la hélice intracelular ICH5, que falta parcialmente en la estructura, se utilizó la información de los residuos F467-V480 de rGLUT5 4YBQ. Además, los residuos M457-T466 (rGLUT5, 4YBQ) también se incluyeron en la alineación para ayudar a la continuidad de ICH5 desde TM12.

## 2. Construcción de ensamblajes solvatados en topología DM-GG

Cada modelo se protonó utilizando el *software* pdb2pqr a pH fisiológico (pH =7.4) y el método de PROPKA para predecir los estados de protonación de los residuos ionizables (Olsson et al., 2011). Además, se incluyó la opción *--keep-chain* para mantener en correcta orientación elementos de estructura secundaria y se configuró el campo de fuerza AMBER mediante la opción *--ff AMBER*, esto para prevenir futuras incompatibilidades con las simulaciones de dinámica molecular, relacionados a la nomenclatura de residuos protonados en AMBER (Case et al., 2023).

Los modelos protonados de rGLUT5 se orientaron e insertaron en una bicapa lipídica simétrica compuesta por fosfatidilcolina (POPC) utilizando la herramienta PACKMOL-Memgen para la generación del entorno membranal (Schott-Verdugo & Gohlke, 2019). Posteriormente, el sistema se solvató con solvente explícito, empleando moléculas de agua de tipo WT4, y se embebió en una caja cúbica de  $160 \times 160 \times 160 \text{ \AA}$  a una concentración final de 0,15 M de NaCl.

Para evitar solapamientos entre los átomos se estableció una distancia mínima de  $10 \text{ \AA}$  entre la proteína y los límites de la caja en las direcciones x, y e z, así como una distancia mínima de  $17,5 \text{ \AA}$  entre las moléculas de agua y cualquier otra entidad del sistema durante la etapa de construcción.

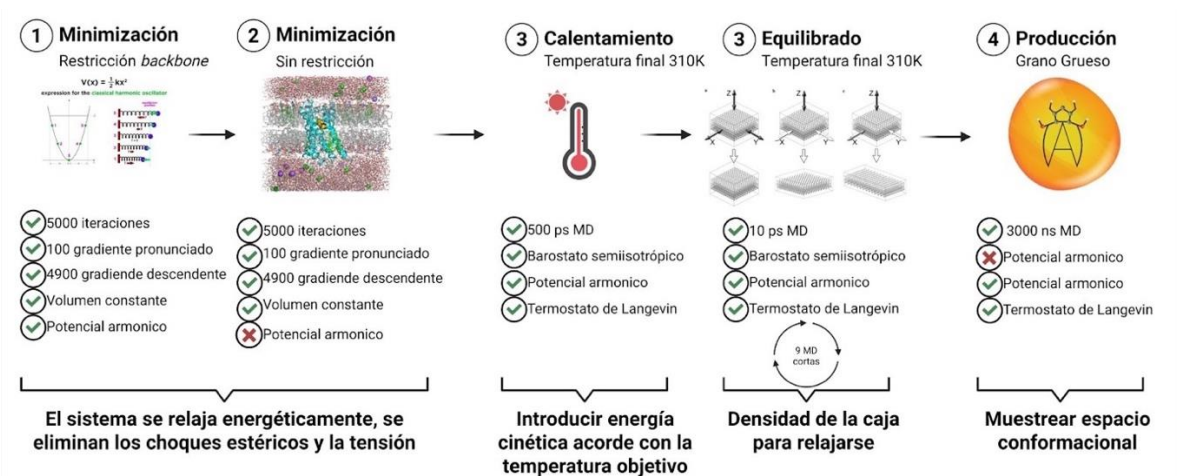
La representación topológica de grano grueso se realizó empleando las parametrizaciones del campo de fuerza *South American Initiative for a Rapid and Accurate Hamiltonian* (SIRAH) (Klein et al., 2023). Con el fin de prevenir inestabilidades estructurales durante las simulaciones, asociadas a la presencia no deseada de moléculas de agua en regiones hidrofóbicas de la bicapa lipídica, se incorporó un parche adicional de solvente de  $27 \text{ \AA}$  de

espesor a lo largo del eje z del sistema, ubicado a una distancia de 0,7 Å respecto a la superficie de la membrana. Esta etapa se realizó utilizando la herramienta *tleap* del paquete AMBER, permitiendo una correcta delimitación de las regiones acuosas y membranales (Case et al., 2023).

Dado que PACKMOL no considera explícitamente la periodicidad del sistema durante la construcción inicial, se implementó un procedimiento posterior de depuración para evitar solapamientos entre moléculas pertenecientes a imágenes periódicas adyacentes. En particular, se eliminaron moléculas de agua cuya distancia a otras moléculas de agua fuera inferior a 10 Å, empleando la herramienta CPPTRAJ, asegurando así la consistencia geométrica del sistema bajo condiciones periódicas (Roe & Cheatham, 2013).

### **3. Relajación de ensambles termodinámicos en condiciones NPT**

Antes de iniciar las simulaciones de dinámica molecular, se llevó a cabo un protocolo de relajación del sistema con el fin de garantizar su estabilidad termodinámica y estructural (Grønbech-Jensen & Farago, 2013). Este protocolo, sugerido por los creadores de SIRAH, incluyó una minimización de energía para eliminar choques estéricos, seguida de un calentamiento gradual hasta la temperatura objetivo (310 K), y una etapa de liberación progresiva de las restricciones aplicadas al sistema (Figura 9).



**Figura 9. Protocolo para la relajación de los sistemas.**

Las etapas de relajación del sistema, incluyendo las fases de minimización energética, calentamiento progresivo y equilibración, así como la posterior fase de producción, se llevaron a cabo siguiendo estrictamente los lineamientos metodológicos recomendados por los desarrolladores del campo de fuerza SIRAH.

Para esto primero se ejecutaron dos minimizaciones, a una de ellas se aplicó un potencial armónico de  $2.4 \text{ kcal} \cdot \text{mol}^{-1} \cdot \text{Å}^{-2}$  durante 5000 ciclos, con un cambio de método de descenso más pronunciado a gradiente conjugado tras los primeros 100 ciclos. Este potencial se aplicó exclusivamente al *backbone*, minimizando cadenas laterales, solvente y lípidos. Esto evita alteraciones en la estructura principal, alcanzando mínimos locales en cadenas laterales. El sistema se mantuvo en volumen constante sin ajuste de presión, registrando la evolución cada 50 pasos y empleando un radio de corte (*cutoff*) de  $12 \text{ Å}$  para las interacciones no enlazadas. La segunda minimización se realizó utilizando los mismos parámetros que la primera, pero omitiendo la aplicación del potencial armónico para minimizar todo el sistema.

Para el calentamiento o *heating*, se ejecutó una simulación dinámica de molecular de ensamble NPT usando termostato de Langevin hasta alcanzar los 310 K, aplicando un potencial armónico de  $2.4 \text{ kcal} \cdot \text{mol}^{-1} \cdot \text{Å}^{-2}$  en el *backbone* (Darden et al., 1993). Se ejecutaron 25,000 pasos de DM con un tiempo de paso de 0.02 ps, registrando la salida cada 100 pasos, guardando reinicios cada 10,000 pasos y escribiendo coordenadas atómicas en cada 100

pasos. El acoplamiento de presión semi isotrópico se combinó con una tensión superficial nula en el plano XY en un tiempo de relajación de presión de 8 ps. Al no imponer fuerzas uniformes que modifiquen la topología de la membrana en los tres ejes, esta puede expandirse o contraerse libremente en su plano, evitando membranas llamadas “aplanadas”.

Dado que en sistemas proteína–membrana la presión no se distribuye de manera isotrópica, a diferencia de lo que ocurre en simulaciones de proteínas solvatadas, resulta fundamental controlar cuidadosamente la evolución del volumen de la caja periódica durante la etapa de equilibración.

Con el objetivo de prevenir cambios abruptos en el volumen y garantizar una relajación progresiva de la bicapa y de la proteína embebida, se implementó un protocolo de equilibración compuesto por nueve etapas secuenciales, cada una de 25.000 pasos de integración, utilizando un paso de tiempo de 0.02 ps. Todas las etapas se ejecutaron en ensamble NPT a 310 K, reanudando cada simulación a partir de la configuración final de la etapa previa. Para mejorar la estabilidad numérica durante esta fase, se empleó un valor elevado del parámetro *skinnb* (5), reduciendo la frecuencia de actualización de la lista de vecinos no enlazados.

En cada etapa de equilibración se aplicó un acoplamiento de presión semi isotrópico, manteniendo tensión superficial nula en el plano de la membrana (xy) y permitiendo la relajación independiente del sistema en la dirección z, con un tiempo de relajación de presión de 8 ps. Este esquema de acoplamiento permite preservar el área lateral de la bicapa mientras se ajusta de manera controlada su espesor, evitando artefactos estructurales como colapsos, expansiones artificiales o inestabilidades en la interfaz proteína–membrana. Durante cada bloque de equilibración, el sistema evolucionó libremente bajo las ecuaciones de

movimiento, registrando energías y observables cada 100 pasos, escribiendo coordenadas con la misma frecuencia y generando archivos de reinicio binarios cada 10.000 pasos. Este procedimiento permitió monitorear la convergencia progresiva del sistema y asegurar condiciones estructurales y termodinámicas estables antes del inicio de la fase de producción.

Finalmente se estableció un radio de corte de 12 Å para las interacciones no enlazadas, mientras que el control de temperatura se mantuvo mediante un termostato de Langevin a 310 K con una frecuencia de colisión de 5 ps<sup>-1</sup>. Con el fin de prevenir despliegues no deseados del esqueleto proteico, se aplicaron restricciones armónicas suaves con una constante de fuerza de 0.24 kcal·mol<sup>-1</sup>·Å<sup>-2</sup> sobre los átomos del *backbone*.

#### **4. Muestrear el espacio conformacional de cada ensamble**

La etapa de producción consistió en la ejecución de dinámica molecular para los sistemas relajados durante 3 μs a 310 K (NPT), utilizando las posiciones cartesianas de los átomos provenientes de la última simulación de relajación. Esta simulación empleó 150 millones de pasos de integración con un paso temporal de 0.02 ps, lo que corresponde a un total de 3 μs de trayectoria.

Durante esta etapa, se mantuvo el acoplamiento de presión semi isotrópico con tensión superficial nula en el plano xy, y se utilizó un radio de corte de 12 Å para las interacciones no enlazadas. La temperatura se controló mediante un termostato de Langevin con una frecuencia de colisión de 5 ps<sup>-1</sup> para garantizar una temperatura constante de 310 K. Finalmente las coordenadas se guardan cada 100 ps, resultando en una frecuencia de muestreo de 0.01 frames/ps.

Para las simulaciones extendidas de hasta 5  $\mu$ s, se empleó el esquema de integración *middle*, un integrador simpléctico recientemente incorporado en AMBER (Case et al., 2023; Z. Sun et al., 2021). En los esquemas de Langevin convencionales, el termostato se aplica al inicio o al final del paso de integración, lo que introduce una asimetría entre las fuerzas deterministas y el ruido estocástico y genera un sesgo en las distribuciones de posición y momento muestreadas. El integrador *middle* resuelve esto dividiendo el paso de integración en tres etapas: primero se actualizan las velocidades con las fuerzas deterministas durante medio paso; luego se aplica el termostato de Langevin sobre esas velocidades ya actualizadas, en el centro del paso; y finalmente se actualizan las posiciones y se completa el medio paso de velocidades restante. Este esquema, denominado *V-O-V*, garantiza que el termostato actúe sobre velocidades que ya incorporan las fuerzas físicas del sistema, preservando la simetría temporal del integrador y produciendo distribuciones estadísticas más cercanas al ensamble canónico correcto.

## 5. Selección de Variables Colectivas (CV)

La simulación de un sistema molecular genera una gran cantidad de información debido al elevado número de grados de libertad asociados a los movimientos atómicos. Por ello, uno de los objetivos principales de esta tesis es reducir la complejidad de la dinámica de las proteínas a unos pocos parámetros claves, permitiendo descartar los movimientos irrelevantes de GLUT5 y centrarse en aquellos estructuralmente significativos.

No obstante, elegir un sistema de baja dimensión es una tarea compleja, cuyo éxito depende en gran medida de la elección adecuada de la característica estructural. Esta característica debe capturar los movimientos lentos que son fundamentales para describir los cambios conformacionales, a lo que se determinará como variable colectiva (CV) (Nagel & Stock,

2023). La elección de esta CV no es una propuesta promovida de este trabajo y más bien, es una corroboración a las variables utilizadas en el trabajo de Sarah E McComas en el que reconstruyeron el ciclo de transporte en GLUT5 utilizando métodos de no equilibrio (McComas et al., 2022). En particular, se utilizaron la distancia entre el centro de masa (COM) de los segmentos extracelulares de las hélices transmembrana TM1 (residuos 36–43) y TM7 (residuos 295–301) como medida de la apertura de la puerta extracelular. De forma análoga, la apertura de la puerta intracelular se evaluó mediante la distancia entre el COM de los segmentos intracelulares de las hélices transmembrana TM4 (residuos 142–151) y TM10 (residuos 392–400).

## 6. Análisis de convergencia estadística

En dinámica molecular, una pregunta recurrente es si el tiempo de simulación es suficiente para capturar el fenómeno biológico de interés, es decir, si la exploración conformacional ha convergido (Ormeño & General, 2024). Para evaluar la convergencia de los sistemas simulados se empleó el método de promediado por bloques (*block averaging*), que permite estimar la independencia estadística de los datos generados por la simulación (Romo & Grossfield, 2011).

Las definiciones formales del método se encuentran en (Vehtari et al., 2021), pero conceptualmente el procedimiento consiste en dividir la serie temporal de un observable en bloques contiguos de tamaño creciente y calcular la media de cada bloque. La desviación estándar de estas medias se analiza en función del tamaño del bloque. Cuando esta desviación se estabiliza formando una meseta, indica que el tamaño del bloque supera el tiempo de correlación del sistema y que cada bloque puede considerarse estadísticamente independiente. Este punto define la convergencia estadística del observable.

De manera complementaria, se evaluó la función de autocorrelación temporal  $C_{x(\tau)}$  de las distancias intracelulares ( $C_{ic(\tau)}$ ) y extracelulares ( $C_{ec(\tau)}$ ) para estimar el tiempo de autocorrelación integrado, el cual permite calcular el número efectivo de muestras independientes  $N_{eff}$ . Este valor cuantifica cuántas veces la simulación ha explorado de forma estadísticamente independiente el espacio conformacional de interés.

Valores bajos de número de muestras independientes, ausencia de meseta en el análisis por bloques o tiempos de autocorrelación elevados se interpretan como evidencia de muestreo insuficiente, indicando que la simulación permanece atrapada en un subconjunto limitado del espacio conformacional.

Para aplicar este análisis de convergencia estadística, se adaptaron los códigos originalmente desarrollados en Julia, tomando como referencia MolSimToolkit (<https://github.com/m3g/MolSimToolkit.jl>), implementándolos en Python para su integración con el flujo de trabajo de esta tesis.

## 7. Discretización del espacio conformacional

Para discretizar el espacio conformacional continuo en una trayectoria de estados discretos, se empleó un algoritmo de *clustering* geométrico basado en densidades. Se optó por esta metodología ya que no requiere de un número inicial de grupos, como lo hace *kmeans*, si no que utiliza regiones de alta densidad del espacio conformacional como puntos de semilla o *seeds* de los diferentes estados, a los que se asignan sucesivamente puntos de regiones con menor densidad (Sittel & Stock, 2016). De esta manera, se aprovecha la naturaleza jerárquica de la energía libre del espacio conformacional para identificar estados separados por barreras energéticas.

Las definiciones formales del algoritmo están disponibles en <https://github.com/moldyn/Clustering>, pero brevemente la arquitectura del algoritmo consiste:

- **Cálculo de energía libre (FEL):** La energía libre de cada punto del espacio conformacional accesible se estimó en función de la distribución de probabilidades en la coordenada de reacción  $P(x)$ , de acuerdo con la distribución de Boltzmann  $\Delta G(x) = -k_b T \ln P(x)$ , por lo tanto, regiones con mayor densidad de puntos se encuentran separadas por una barrera de aquellas con poca densidad. Dado que la energía libre está definida hasta una constante aditiva, los perfiles se desplazaron de modo que el mínimo global fuese cero.
- **Barrido iterativo del paisaje de energía libre:** se ejecutó un barrido iterativo del paisaje de energía libre, desde valores mínimos hacia máximos del espacio conformacional, con incrementos de  $\Delta G = 0.1 kT$  en cada iteración. En cada umbral de energía  $\check{G}$  se identificaron todos los puntos con energía libre inferior  $G \leq \check{G}$  y se agrupan en *clusters* si cumplían el criterio de proximidad geométrica (0.5 Å). De esta forma, dos puntos con energía inferior al umbral y geoméricamente próximos en el espacio reducido se asignaron al mismo clúster.
- **Identificación semillas de microestados (*seeds*):**  
Conceptualmente, los clústeres representan agrupaciones geométricas similares, pero no necesariamente corresponden a estados de mínima energía libre. En otras palabras, un mismo clúster puede contener varios mínimos locales del paisaje de energía libre. Para identificar estos estados, se utiliza la jerarquía energética inherente al espacio conformacional, donde los mínimos

locales son representados como nodos (*seeds*) separados por regiones de mayor energía.

Para la elección estos *seeds* se utilizaron los siguientes criterios:

- Desconexión geométrica de otros *clusters* a su nivel de energía de aparición
  - No está en su máxima energía libre
  - Población superior a un valor mínimo (500 frames)
- **Asignación de microestados:** Una vez identificados los *seeds*, se asignaron todos los *frames* restantes mediante un proceso de "llenado de cuencas". Los *seeds* identificados en el paso anterior representan solo los puntos de mínima energía en cada cuenca, pero la mayoría de los *frames* de la trayectoria se encuentran en regiones de mayor energía (laderas de las cuencas y barreras energéticas). Para este llenado cada *frame* se asoció al *seed* más cercano geoméricamente hasta que todos los *frames* fueron asignados.

## 8. Transiciones conformacionales mediante modelos de Markov.

Para caracterizar la cinética de los cambios conformacionales de GLUT5 se emplearon Modelos de Estados de Markov (MSM), que describen la dinámica molecular mediante transiciones probabilísticas entre un conjunto finito de estados metaestables discretizados en el espacio conformacional (Malmstrom et al., 2014).

Los MSM se construyen sobre la aproximación markoviana, que establece que el futuro del sistema depende exclusivamente de su estado presente. Esto implica que, una vez transcurrido un intervalo de tiempo suficiente, denominado *lag time*, el sistema ha "olvidado" su historia conformacional previa. Bajo esta condición, la dinámica completa puede resumirse en una matriz de transición: una tabla de probabilidades que indica, para cada

estado posible, con qué probabilidad el sistema se encontrará en cada otro estado tras un intervalo de tiempo  $\tau$ . La elección adecuada del *lag time* es por tanto esencial, ya que debe ser lo suficientemente largo para que la propiedad markoviana se cumpla, pero no tan largo que se pierda resolución temporal en la descripción de los procesos conformacionales.

En este trabajo, los MSM se construyeron y analizaron utilizando *msmhelper*, un paquete de libre acceso implementado en Python, ampliamente empleado en el análisis, validación y visualización de modelos markovianos aplicados a trayectorias de dinámica molecular (Nagel & Stock, 2023).

La validez de la aproximación markoviana se evaluó mediante dos análisis complementarios. Primero, se analizaron las *implied timescales*, escalas temporales asociadas a los modos dinámicos más lentos del sistema: su convergencia al aumentar el *lag time* ( $\tau$ ) verifica que la discretización del espacio conformacional satisface la hipótesis markoviana (Chang et al., 2003). Segundo, se realizaron pruebas de Chapman-Kolmogorov, que evalúan la consistencia interna del modelo comparando las probabilidades de transición observadas a tiempos largos con las predichas por composición sucesiva de matrices de transición a tiempos más cortos; su concordancia confirma que el MSM reproduce fielmente la dinámica efectiva del sistema (Motahari et al., 2025).

## RESULTADOS

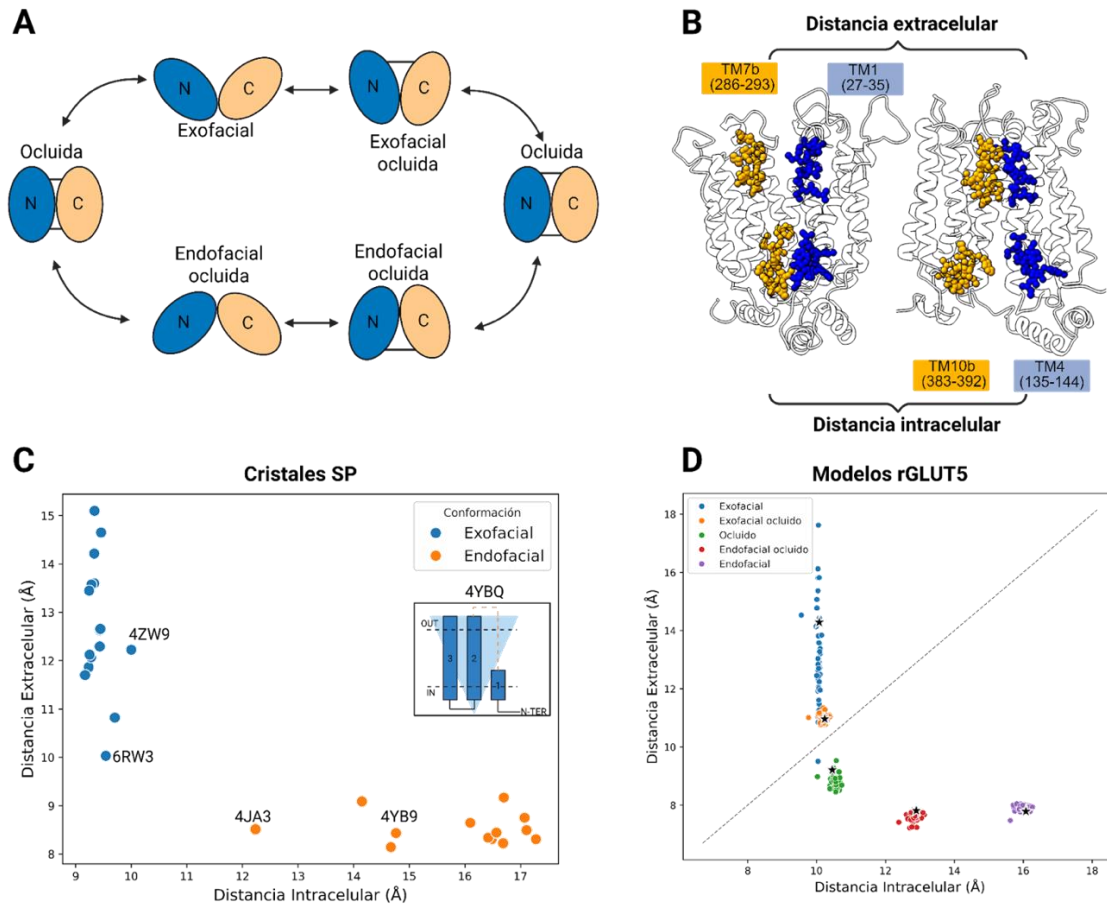
### 1. Validación del descriptor y modelos

Dado que la exploración del espacio conformacional asociado al modelo de acceso alternante en rGLUT5 depende del descriptor utilizado para discriminar estados conformacionales, se realizó una validación de dicho descriptor (Figura 10). Antes de aplicarlo al análisis de las trayectorias de dinámica molecular, se evaluó su capacidad para clasificar distintos estados conformacionales utilizando como referencia las estructuras experimentales disponibles de transportadores pertenecientes a la familia SP. Para este análisis se consideraron un total de 30 estructuras resueltas hasta la fecha.

Dado que los transportadores de la familia SP presentan baja identidad de secuencia, pero arquitectura conservada, se realizó un alineamiento estructural para identificar residuos estructuralmente homólogos entre ellos. Estos residuos fueron utilizados para definir el descriptor (Tabla 4). A partir del alineamiento estructural se identificaron dos grupos conformacionales bien diferenciados. El primero agrupa estructuras con una apertura intracelular que oscila entre 9 y 10 Å y una apertura extracelular de entre 10 y 15 Å, lo que corresponde a conformaciones exofaciales. El segundo grupo reúne estructuras que presentan aperturas intracelulares mayores, entre 12 y 17 Å, mientras que la apertura extracelular es menos variable, con valores entre 8 y 9 Å. Este patrón es consistente con conformaciones endofaciales (Figura 10C).

Para validar esta clasificación se compararon los resultados del análisis con la información estructural reportada en la literatura. En todos los casos, las orientaciones identificadas mediante el descriptor coinciden con las conformaciones asignadas por los autores originales de las estructuras. La única excepción corresponde al transportador de hexosas de

*Plasmodium falciparum* (PDB: 6RW3), descrito como una conformación completamente ocluida. En conjunto, estos resultados indican que el descriptor discrimina correctamente los principales estados conformacionales observados en transportadores de la familia SP.



**Figura 10. Evaluación del descriptor estructural.**

(A) Representación de las conformaciones asociadas al modelo de acceso alternativo. (B) Descriptores estructurales utilizados para caracterizar la apertura de las compuertas extracelular e intracelular. La compuerta extracelular se define mediante la distancia entre residuos representativos de TM1 y TM7, mientras que la compuerta intracelular se define entre residuos de TM4 y TM10. Las distancias corresponden a medidas euclidianas entre los residuos indicados (mostrados como esferas). Como referencia estructural se ilustran las conformaciones exofacial ocluida (PDB: 4ZW9) y endofacial (PDB: 4YB9). (C) Validación del descriptor estructural mediante el análisis de 30 estructuras cristalográficas disponibles para la familia SP. Cada punto corresponde a una estructura proyectada en el espacio conformacional definido por ambos descriptores. El agrupamiento jerárquico mediante HDBSCAN permitió identificar orientaciones conformacionales diferenciadas. Se destacan las estructuras utilizadas como templates en el modelamiento comparativo. La estructura PDB: 4YBQ no se agrupa con los demás estados debido a la ausencia de la región TM1–TM2 en el cristal, lo que impide cuantificar la apertura extracelular mediante el descriptor definido. (D) Aplicación del descriptor para evaluar el grado de apertura en los modelos estructurales utilizados como conformaciones iniciales en el modelo de acceso alternativo de rGLUT5. Cada punto en la gráfica representa los modelos generados para cada estado conformacional. Las estrellas negras indican el modelo seleccionado como representación estructural más adecuada para los análisis posteriores.

**Tabla 4. Miembros de la familia SP con estructura definida.**

HT1 = hexose transporter 1, SPT6= Sugar transport protein 6, SPT10= Sugar transport protein 10, NA = No Aplica, particularmente en 4YBQ ya que no está resuelto ese TM. \* Puerta extracelular. \*\*Puerta intracelular

PDB	Proteína	Organismo	TM1*	TM7*	TM4**	TM10**
4YB9	GLUT5	<i>B.taurus</i>	36-43	295-301	142-151	392-400
4YBQ	GLUT5	<i>R. norvegicus</i>	NA	295-301	142-151	392-400
4GBY	Xyle	<i>E.coli</i>	27-35	294-301	144-153	383-392
4GBZ	Xyle	<i>E.coli</i>	27-35	294-301	144-153	383-392
4JA3	Xyle	<i>E.coli</i>	27-35	294-301	144-153	383-392
4JA4	Xyle	<i>E.coli</i>	27-35	294-301	144-153	383-392
4QIQ	Xyle	<i>E.coli</i>	27-35	294-301	144-153	383-392
6N3I	Xyle	<i>E.coli</i>	27-35	294-301	144-153	383-392
6M20	HT1	<i>P.falciparum</i>	43-51	311-318	145-154	409-418
6M2L	HT1	<i>P.falciparum</i>	43-51	311-318	145-154	409-418
6RW3	HT1	<i>P.falciparum</i>	43-51	311-318	145-154	409-418
4PYP	GLUT1	<i>H.sapiens</i>	29-37	288-295	137-146	385-394
5EQG	GLUT1	<i>H.sapiens</i>	29-37	288-295	137-146	385-394
5EQH	GLUT1	<i>H.sapiens</i>	29-37	288-295	137-146	385-394
5EQI	GLUT1	<i>H.sapiens</i>	29-37	288-295	137-146	385-394
6THA	GLUT1	<i>H.sapiens</i>	29-37	288-295	137-146	385-394
9J2N	GLUT7	<i>H.sapiens</i>	41-49	300-307	149-158	399-308
9G11	SPT6	<i>A.thaliana</i>	38-49	296-303	148-157	400-409
8Y65	GLUT9	<i>H.sapiens</i>	74-82	333-340	182-191	432-441
8Y66	GLUT9	<i>H.sapiens</i>	74-82	333-340	182-191	432-441
7WSM	GLUT4	<i>H.sapiens</i>	41-49	304-311	153-162	401-410
4ZW9	GLUT3	<i>H.sapiens</i>	27-35	286-293	135-144	383-392
4ZWB	GLUT3	<i>H.sapiens</i>	27-35	286-293	135-144	383-392
4ZWC	GLUT3	<i>H.sapiens</i>	27-35	286-293	135-144	383-392
7CRZ	GLUT3	<i>H.sapiens</i>	27-35	286-293	135-144	383-392
7SPS	GLUT3	<i>H.sapiens</i>	27-35	286-293	135-144	383-392
5C65	GLUT3	<i>H.sapiens</i>	27-35	286-293	135-144	383-392
7SPT	GLUT3	<i>H.sapiens</i>	27-35	286-293	135-144	383-392
6H7D	SPT10	<i>A.thaliana</i>	42-53	301-308	153-162	407-416

7AAQ	SPT10	<i>A.thaliana</i>	42-53	301-308	153-162	407-416
7AAR	SPT10	<i>A.thaliana</i>	42-53	301-308	153-162	407-416

Una vez validado, el descriptor se aplicó a los modelos estructurales generados mediante modelamiento comparativo para representar los estados iniciales del ciclo de acceso alternante en rGLUT5. El análisis permitió discriminar de manera consistente las distintas conformaciones modeladas, lo que confirma que el descriptor también es adecuado para caracterizar los estados estructurales del transportador en estudio (Figura 10D). En consecuencia, la selección de las estructuras utilizadas como plantillas para el modelamiento comparativo resulta apropiada para representar los distintos estados conformacionales del ciclo de transporte.

Posteriormente se evaluó la calidad estructural de los modelos generados. Para ello se empleó el potencial estadístico DOPE (*Discrete Optimized Protein Energy*) (Webb & Sali, 2016). Esta evaluación fue necesaria debido a que el modelamiento comparativo se realizó utilizando el protocolo de refinamiento rápido de MODELLER, el cual se basa en un proceso de optimización estocástica que genera múltiples modelos con la misma topología global, pero con diferencias en detalles estructurales locales.

El potencial DOPE se basa en distribuciones estadísticas de distancias interatómicas derivadas de estructuras cristalográficas. Este potencial permite estimar una pseudoenergía asociada a cada modelo y evaluar su compatibilidad estructural con conformaciones proteicas conocidas. En este contexto, valores más bajos de DOPE indican una mayor coherencia energética y una mayor probabilidad de que el modelo sea estructuralmente fiable (Leach, 1997).

Sobre esta base, se seleccionaron como representativos de cada estado conformacional aquellos modelos que presentaron los valores globales de DOPE más bajos. Como se resume en la Tabla 5, los modelos seleccionados presentan puntuaciones DOPE de -62452.02 para la conformación exofacial, -70385.64 para la exofacial ocluida, -61467.34 para la ocluida, -66044.57 para la endofacial ocluida y -64545.64 para la endofacial.

**Tabla 5. Modelos comparativos para el mecanismo alternante.**

\*RMSD = Raíz de la Desviación Cuadrática Media respecto a carbonos alfa del cristal

PDB	Nombre	Conformación	Organismo	RMSD	DOPE	N °Modelo
4YBQ	GLUT5	exofacial	<i>H.sapiens</i>	1.20 Å	-62452.02	2
4ZW9	GLUT3	exofacial ocluida	<i>H.sapiens</i>	0.502 Å	-70385.64	10
6RW3	hexose transporter	ocluida	<i>P.falciparum</i>	0.805 Å	-61467.34	25
4JA3	symporter Xyle	endofacial ocluida	<i>E.coli</i>	0.959 Å	-66044.57	3
4YB9	GLUT5	endofacial	<i>B.taurus</i>	0.154 Å	-64545.64	38

Sin embargo, una evaluación global no permite estimar la calidad estructural de regiones específicas, particularmente aquellas ausentes o pobremente resueltas en las estructuras experimentales utilizadas como plantillas. Por esta razón, el análisis se complementó mediante el cálculo del perfil DOPE por residuo, evaluado sobre los carbonos alfa del *backbone* (Figura 11).

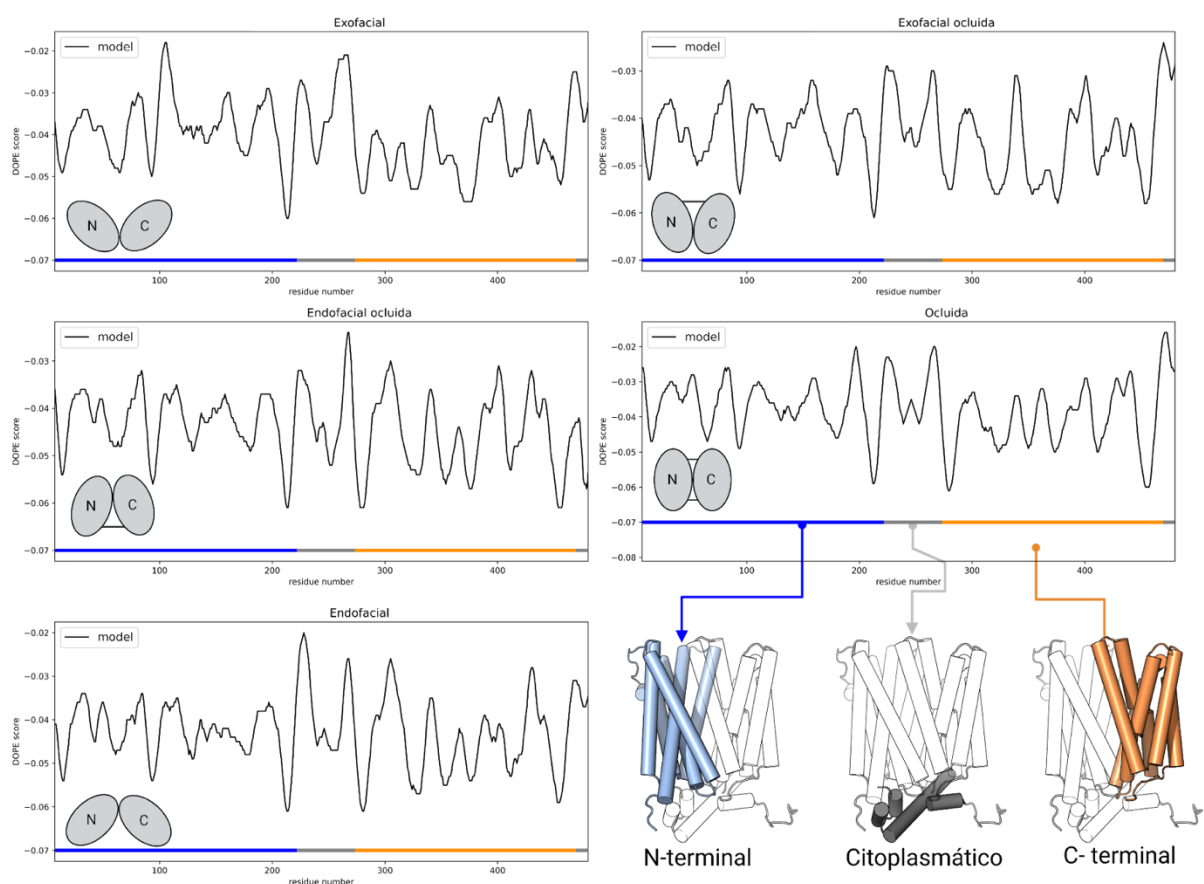
Este análisis local mostró que todas las conformaciones presentan valores DOPE negativos a lo largo de la estructura completa, incluyendo regiones que no están representadas en las estructuras cristalográficas originales, principalmente segmentos del dominio citoplasmático. En conjunto, los resultados del análisis global y local indican que los modelos generados presentan una calidad estructural consistente.

Sin embargo, el criterio DOPE no es suficiente para evaluar la calidad estructural de los modelos generados, ya que corresponde a un potencial estadístico de escala arbitraria útil para comparar modelos de una misma proteína, pero que no proporciona información sobre la corrección geométrica del modelo a nivel atómico. Para abordar esta limitación, se realizó una validación estructural complementaria mediante MolProbity (Davis et al., 2007), una plataforma de validación ampliamente adoptada en biología estructural que integra múltiples criterios geométricos para evaluar la calidad de estructuras proteicas. MolProbity analiza parámetros como la distribución de ángulos  $\phi/\psi$  en el gráfico de Ramachandran, la conformación de rotámeros de cadenas laterales, las desviaciones en longitudes y ángulos de enlaces covalentes, la geometría de carbonos beta y el análisis CaBLAM del *backbone*, permitiendo identificar regiones del modelo con potenciales errores estructurales.

Los resultados MolProbity indican que la calidad estructural de los modelos generados varía según la conformación analizada. La conformación exofacial presentó los mejores indicadores globales, con un 97.84% de residuos en regiones favorecidas del gráfico de Ramachandran, ausencia total de desviaciones en  $C\beta$  (0.00%) y en *CA Geometry* (0.00%), y el valor más bajo de CaBLAM *outliers* (0.60%), único modelo que cumple este criterio. En el extremo opuesto, la conformación ocluida y endofacial ocluida presentaron los indicadores más deficientes, con *outliers* de Ramachandran de 1.51% y 1.54% respectivamente, valores de CaBLAM de 3.6% y 3.2%, y porcentajes de Ramachandran favored por debajo del 94%, todos alejados de los valores ideales. Las conformaciones exofacial ocluida y endofacial mostraron valores intermedios. En todos los modelos se observaron valores elevados de *bad angles* (entre 1.14% y 1.59%) y *poor rotamers* (entre 1.51% y 4.53%). Cabe destacar que en

la totalidad de los modelos no se detectaron errores de quiralidad ni presencia de prolinas cis, indicando que la arquitectura global de la cadena principal fue correctamente modelada.

Estos resultados respaldan que la topología completa del transportador, incluyendo los dominios N-terminal, C-terminal y las regiones citoplasmáticas, fue modelada de manera coherente con el plegamiento característico de los transportadores de la familia SP, a pesar de que ninguna de las conformaciones presentó estándares ideales establecidos para estructuras cristalográficas de alta resolución.



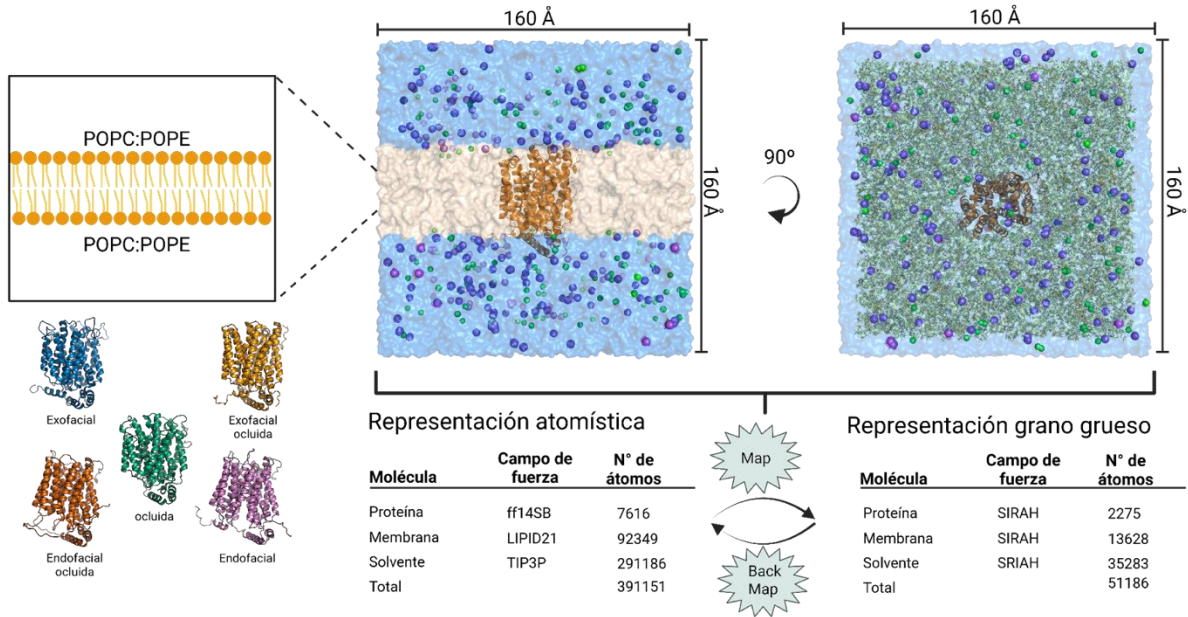
**Figura 11. Evaluación estructural de los modelos.**

Puntaje DOPE por residuo para los mejores modelos evaluados; 2,10,25,3 y 38 que corresponden a exofacial, exofacial ocluida, ocluida, endofacial ocluida y ocluida respectivamente. En azul, gris y naranja se destacan los aminoácidos pertenecientes a los dominios N-terminal, dominio citoplasmático y C-terminal respectivamente.

## **2. Estabilidad termodinámica de los sistemas GG en ensamble NPT**

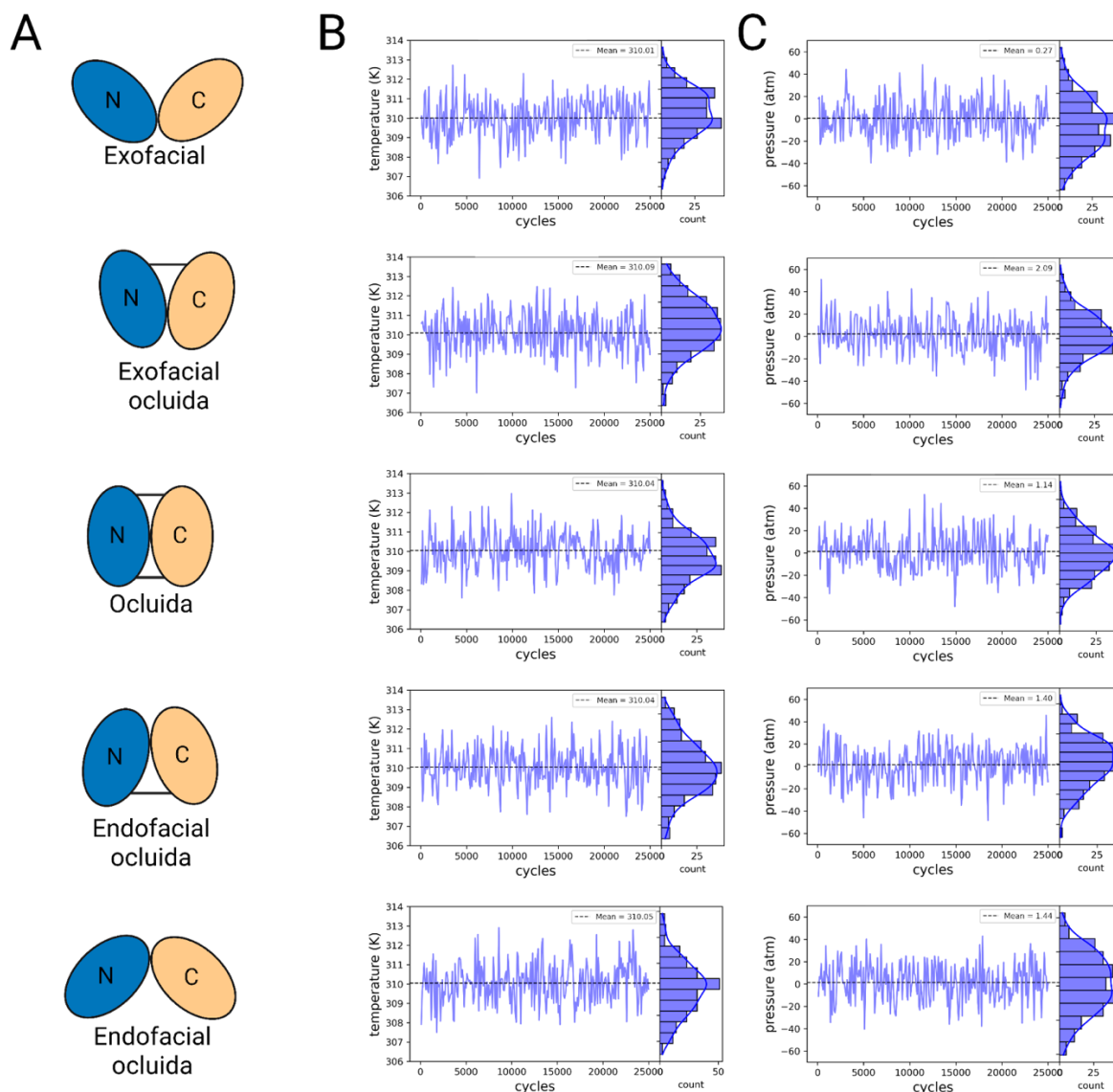
Los modelos seleccionados tras la evaluación de calidad fueron ensamblados en sistemas de simulación y parametrizados utilizando el campo de fuerza GG SIRAH. La construcción de cada sistema se verificó mediante inspección visual, asegurando una solvatación adecuada, la ausencia de agua o iones en la región hidrofóbica de la bicapa y la correcta inserción y orientación de la proteína en la membrana. Los sistemas construidos no presentaron solapamientos estéricos ni interacciones no deseadas, indicando una topología adecuada y el cumplimiento de las distancias mínimas definidas por los radios de van der Waals.

La representación GG redujo el número de partículas del sistema de 391,191 átomos en la descripción atomística a 51,186 partículas en el modelo GG, lo que corresponde a una disminución del 86,9%. Esta reducción permitió mejorar la eficiencia computacional y extender las escalas temporales accesibles (Figura 12). Respecto a la relajación de los sistemas, la temperatura se mantuvo controlada alrededor de 310 K en todas las conformaciones, con valores promedio de 310K. De manera consistente, la presión mostró fluctuaciones características del ensamble NPT, con promedios cercanos a 1 atm (Figura 13), indicando un control termodinámico adecuado durante la fase de equilibración.



**Figura 12. Representaciones atómica y CG de rGLUT5.**

El panel izquierdo muestra el esquema de la membrana asimétrica utilizada. Se presentan las diferentes conformaciones del transportador rGLUT5: abierto hacia el exterior, ocluido hacia el exterior, ocluido hacia el interior y abierto hacia el interior. El sistema completo (160 Å × 160 Å) se muestra en el centro de la imagen con rotación de 90°. Las tablas comparativas detallan los campos de fuerza empleados (ff14SB para proteína, LIPID21 para membrana, TIP3P para solvente en el sistema atómico; SIRAH para todos los componentes en grano grueso) y el número de átomos en cada representación. La reducción significativa en el número de partículas (de 391,151 a 51,186) demuestra la eficiencia computacional del modelo grano grueso manteniendo las características estructurales esenciales del sistema membrana-proteína.

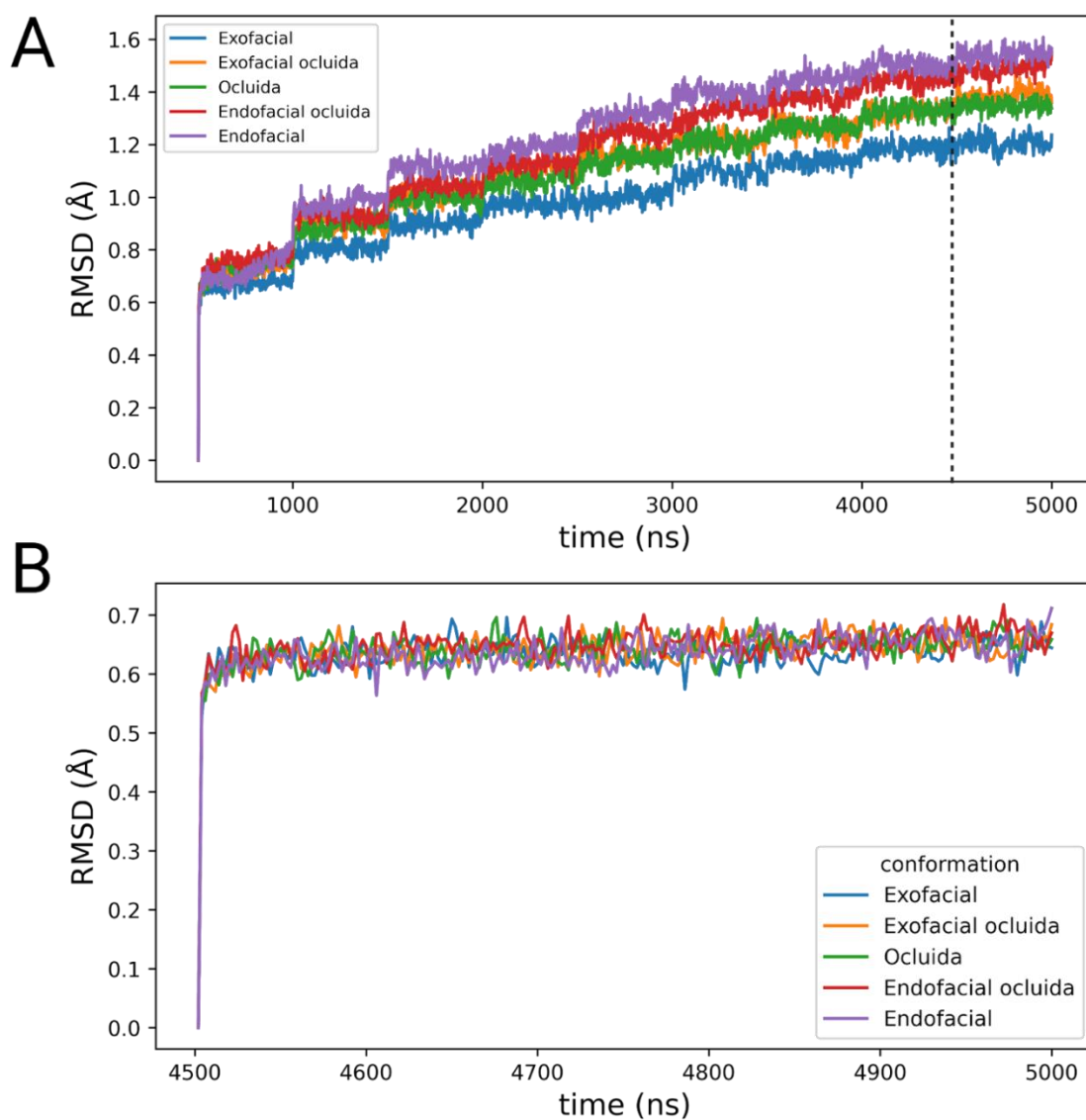


**Figura 13. Condiciones termodinámicas en simulaciones GG bajo ensemble NPT.**

(A) Esquema de los estados conformacionales analizados, mostrando el ensemble termodinámico utilizado. (B) y (C) monitoreo de temperatura y presión durante las simulaciones de equilibración respectivamente. Línea azul continua indica las fluctuaciones, línea negra punteada indica el promedio de la medición respectiva y los histogramas laterales representan la distribución KDE de los valores. Se utilizó un termostato de Langevin y un barostato de Berendsen. Los datos representan a la última simulación de equilibración.

Las simulaciones de equilibrio también fueron exitosas (Figura 14). El análisis del RMSD del *backbone* demostró un aumento inicial asociado a la relajación estructural temprana, seguido de una estabilización en el último ciclo de equilibrio. En esta fase, las trayectorias oscilaron alrededor de valores constantes sin tendencias sistemáticas de deriva, lo que sugiere un comportamiento estacionario compatible con un estado equilibrado bajo las restricciones

aplicadas. En conjunto, estos resultados indican que los sistemas en representación CG presentan estabilidad estructural y termodinámica, constituyendo condiciones iniciales adecuadas para el inicio de las simulaciones de producción orientadas a la exploración del paisaje conformacional de rGLUT5.



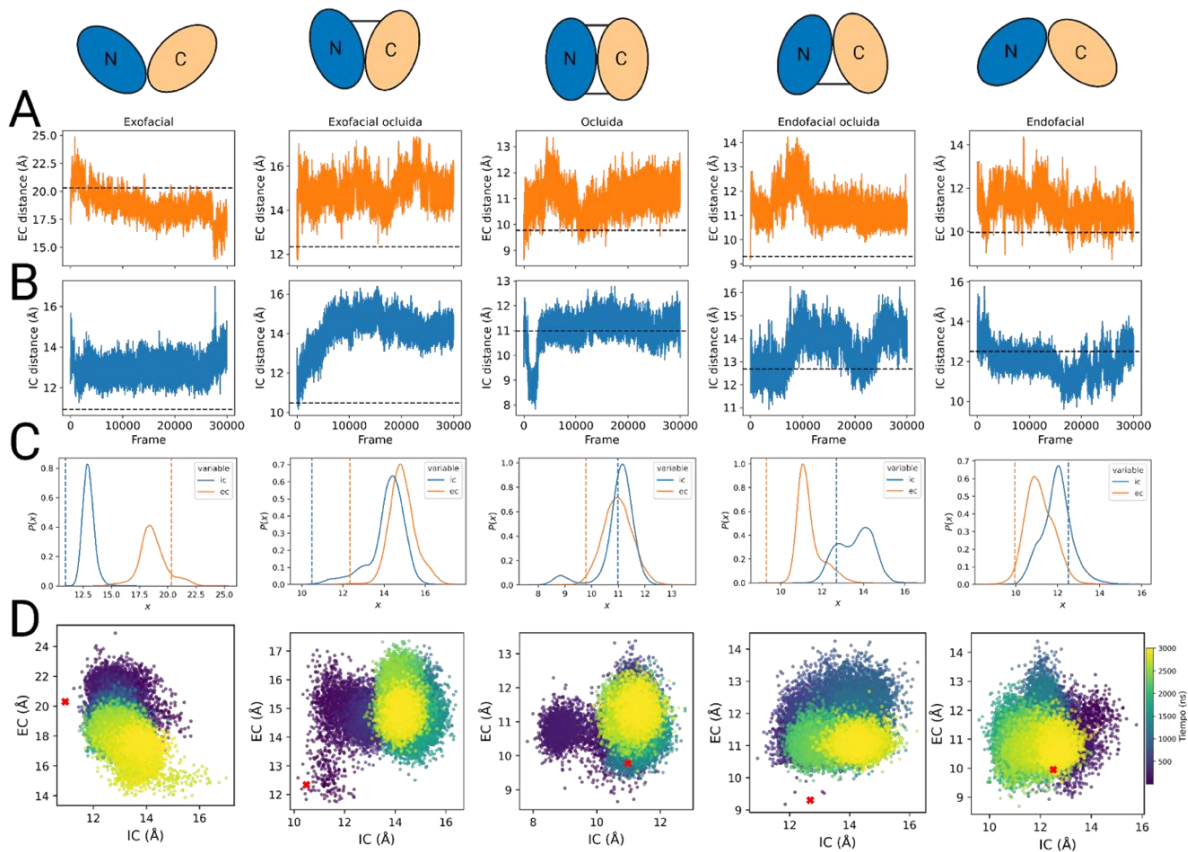
**Figura 14. Equilibración de simulaciones MD-GC.**

(A) Evolución temporal del RMSD del *backbone* para las cinco conformaciones del MAA (exofacial, exofacial ocluida, ocluida, endofacial ocluida y endofacial), calculado sobre las nueve etapas de equilibración consecutivas bajo restricción armónica. La línea vertical discontinua indica el inicio de la novena etapa. (B) RMSD del *backbone* durante la novena equilibración, mostrando la estabilización de las trayectorias al alcanzar el comportamiento estacionario.

### 3. Exploración conformacional mediante DM-GG

La exploración conformacional se refiere a la capacidad de un sistema de muestrear distintas configuraciones estructurales dentro de su paisaje energético durante una simulación. Esta exploración no depende únicamente de la duración de la trayectoria, sino también de la posibilidad del sistema de transitar entre diferentes estados conformacionales separados por barreras energéticas. En sistemas con paisajes energéticos con múltiples mínimos locales separados por barreras energéticas grandes, una simulación extensa puede explorar alrededor de un único mínimo local, sin acceder a otras regiones del espacio conformacional. Para identificar las regiones exploradas por cada simulación, se analizó la distribución de las configuraciones en el plano definido por las distancias de las compuertas intracelular en las trayectorias (IC [142:151-392:400]) y extracelular (EC [36:43-295:301]).

Como se observa en la Figura 15 y Tablas S1–S2, la conformación exofacial presentó una distancia EC promedio de  $18.70 \pm 0.51 \text{ \AA}$ , inferior al valor cristalográfico inicial ( $20.30 \text{ \AA}$ ), con un rango de variación amplio ( $13.89\text{--}24.88 \text{ \AA}$ ) que contrasta con el rango más estrecho de IC ( $11.27\text{--}17.00 \text{ \AA}$ ), indicando que el movimiento dominante corresponde al cierre de la compuerta extracelular. En la conformación exofacial ocluida, ambas compuertas aumentaron respecto al control: EC de  $12.34$  a  $14.89 \pm 0.12 \text{ \AA}$  e IC de  $10.48$  a  $14.17 \pm 0.32 \text{ \AA}$ . Los rangos de variación similares en ambas compuertas (EC:  $11.75\text{--}17.35 \text{ \AA}$ ; IC:  $10.16\text{--}16.39 \text{ \AA}$ ) sugieren un comportamiento simétrico durante la simulación. La conformación ocluida mantuvo valores cercanos al control en ambas compuertas (EC:  $10.98 \pm 0.14 \text{ \AA}$ ; IC:  $11.01 \pm 0.17 \text{ \AA}$ ), con rangos más estrechos que los estados exofaciales (EC:  $8.65\text{--}13.38 \text{ \AA}$ ; IC:  $7.82\text{--}12.78 \text{ \AA}$ ), indicando una exploración conformacional más acotada.



**Figura 15. Exploración del espacio conformacional mediante MD-GC 3  $\mu$ s.**

(A) Evolución temporal de la distancia de la compuerta extracelular (EC) en cada conformación simulada. (B) Evolución temporal de la compuerta intracelular (IC) en cada conformación simulada. Las líneas punteadas representan valores obtenidos desde estructuras cristalográficas usadas para el modelamiento de cada conformación. (C) Distribuciones de las distancias IC (azul) y EC (naranja) para cada conformación. Las líneas punteadas indican los valores correspondientes a las estructuras cristalográficas iniciales. (D) Proyección del espacio conformacional explorado en función de las distancias intracelular (IC) y extracelular (EC). Los puntos están coloreados según el tiempo de simulación, indicando la progresión temporal del muestreo. Los marcadores rojos señalan la posición inicial de cada estructura cristalográfica.

En la conformación endofacial ocluida, la mayor amplitud de variación en IC (10.93–16.27 Å) respecto a EC (9.18–14.24 Å), junto con los promedios de  $11.36 \pm 0.19$  Å y  $13.60 \pm 0.27$  Å respectivamente, indica que las principales fluctuaciones ocurren en la compuerta intracelular, con tendencia al cierre extracelular y apertura hacia el citoplasma.

Finalmente, la conformación endofacial presentó rangos similares en ambas compuertas (EC: 8.69–14.38 Å; IC: 9.59–15.77 Å) y promedios comparables (EC:  $11.15 \pm 0.22$  Å; IC:  $11.97 \pm 0.23$  Å), indicando una dinámica equilibrada entre ambas compuertas. En conjunto, estos resultados muestran que las simulaciones iniciadas desde distintas conformaciones exploran regiones diferenciadas del espacio definido por las distancias IC y EC. Las conformaciones exofaciales presentan variaciones más amplias dominadas por cambios en la compuerta extracelular, mientras que los estados ocluidos exhiben rangos de fluctuación más restringidos. Por su parte, las conformaciones endofaciales muestran una mayor contribución de las fluctuaciones en la compuerta intracelular.

Si bien la proyección en el plano IC–EC permite identificar las regiones del espacio conformacional exploradas por cada simulación, esta representación no revela información sobre la evolución temporal del muestreo. En particular, no permite determinar si las distribuciones observadas corresponden a configuraciones que se estabilizan durante la simulación o si continúan modificándose a medida que se incorporan nuevas estructuras a la trayectoria, es decir, ¿Las distribuciones del descriptor (EC, IC) cambian si extendemos la simulación?. Con el objetivo de evaluar el grado de muestreo efectivo de las trayectorias de dinámica molecular, se aplicó el análisis de promediado por bloques (*block averaging*) a los descriptores estructurales EC e IC para cada una de las conformaciones analizadas. A partir de este análisis se estimó el tiempo de autocorrelación integrado  $\tau_{\text{INT}}$  y el número efectivo de muestras estadísticamente independientes  $N_{\text{eff}}$  (Romo & Grossfield, 2011).

En la conformación exofacial, el descriptor EC presentó un tiempo de autocorrelación integrado de  $\tau_{\text{INT}}=4804.68$ , correspondiente a un número efectivo de muestras  $N_{\text{eff}}=6.24$ . En contraste, el descriptor IC mostró un menor tiempo de correlación  $\tau_{\text{INT}}=634.65$ , lo que resulta

en mayor número efectivo de muestras  $N_{\text{eff}}=47.27$ . Los resultados del análisis se encuentran en la Tabla 6 y Anexos.

**Tabla 6. Análisis de convergencia mediante análisis de bloques.**

IC= Distancia intracelular, EC= Distancia extracelular.  $\tau$  representa el tiempo de correlación, de forma que  $\frac{N}{\tau}$  representa el número de frames independiente ( $N_{\text{eff}}$ ), para  $N=30000$

Tiempo ( $\mu\text{s}$ )	Descriptor	Media ( $\text{\AA}$ )	$\tau$ (frames)	$\frac{N}{\tau}$	SEM ( $\text{\AA}$ )
<b>Exofacial</b>					
3	IC	12.99	234.3	47	0.07
3	EC	18.88	2599.8	6	0.50
5	IC	13.44	1123.26	7	0.24
5	EC	19.59	1770.91	4	1.06
<b>Exofacial ocluida</b>					
3	IC	14.17	2616.97	7	0.32
3	EC	14.89	547.89	28	0.12
<b>Ocluida</b>					
3	IC	11.01	1192.02	14	0.17
3	EC	10.98	10135.81	17	0.13
5	IC	11.29	1510.00	11	0.20
5	EC	11.28	2721.52	6	0.32
<b>Endofacial ocluida</b>					
3	IC	13.6	2103.15	10	0.26
3	EC	11.36	1883.85	10	0.18
<b>Endofacial</b>					
3	IC	11.97	1689.78	10	0.23
3	EC	11.15	1244.4	10	0.21
5	IC	12.51	556.13	12	0.15
5	EC	10.24	594.48	11	0.16

En la conformación exofacial ocluida, el descriptor EC presentó un tiempo de autocorrelación integrado de  $\tau_{\text{INT}}=1088.84$ , correspondiente a un número efectivo de muestras  $N_{\text{eff}}=27.55$ . En

contraste, el descriptor IC mostró un mayor tiempo de correlación  $\tau_{\text{INT}}=4024.49$ , lo que resulta en un menor número efectivo de muestras  $N_{\text{eff}}=7.45$ .

En la conformación ocluida, el descriptor EC presentó un tiempo de autocorrelación integrado de  $\tau_{\text{INT}}=1788.15$ , correspondiente a un número efectivo de muestras  $N_{\text{eff}}=16.78$ . De manera similar, el descriptor IC presentó un tiempo de correlación de  $\tau_{\text{INT}}=2108.68$ , lo que corresponde a un número efectivo de muestras  $N_{\text{eff}}=14.23$ .

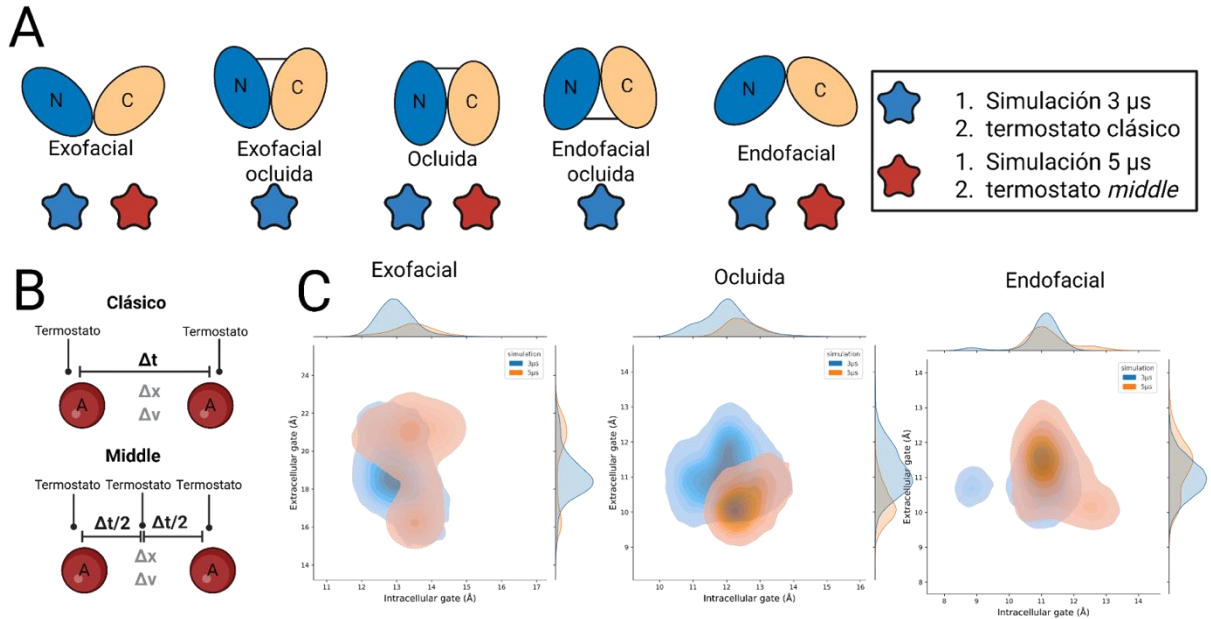
En la conformación endofacial ocluida, el descriptor EC presentó un tiempo de autocorrelación integrado de  $\tau_{\text{INT}}=2909.37$ , correspondiente a un número efectivo de muestras  $N_{\text{eff}}=10.31$ . Por su parte, el descriptor IC mostró un tiempo de correlación de  $\tau_{\text{INT}}=3151.58$ , lo que corresponde a un número efectivo de muestras  $N_{\text{eff}}=9.52$ .

Finalmente, en la conformación endofacial, el descriptor EC presentó un tiempo de autocorrelación integrado de  $\tau_{\text{INT}}=3105.55$ , correspondiente a un número efectivo de muestras  $N_{\text{eff}}=9.66$ . De manera similar, el descriptor IC mostró un tiempo de correlación de  $\tau_{\text{INT}}=3018.46$ , lo que corresponde a un número efectivo de muestras  $N_{\text{eff}}=9.94$ .

En conjunto, los valores estimados de  $\tau_{\text{INT}}$  y  $N_{\text{eff}}$  y evidencian la presencia de correlación temporal en los descriptores EC e IC a lo largo de todas las conformaciones analizadas. Como consecuencia, el número efectivo de muestras independientes es considerablemente menor que el número total de configuraciones presentes en las trayectorias, variando entre aproximadamente 6 y 47 según el descriptor y la conformación considerada. Estas diferencias reflejan variaciones en las escalas temporales de las fluctuaciones capturadas por cada descriptor y sugieren que el muestreo del espacio conformacional asociado aún es limitado dentro de la escala temporal de las simulaciones analizadas.

En función de estos resultados, se extendieron las trayectorias iniciadas desde las conformaciones exofacial, ocluida y endofacial hasta 5  $\mu$ s, incorporando además el termostato *middle* recientemente implementado en Amber (Figura 16A y Anexos) (Case et al., 2025). Sin embargo, la extensión de las simulaciones a 5  $\mu$ s no significó una mejora en la exploración conformacional respecto a las trayectorias de 3  $\mu$ s, observándose únicamente variaciones moderadas en los rangos de apertura analizados (Figura 16C). En la conformación exofacial, la distancia EC aumenta desde  $18.70 \pm 1.27$  Å hasta  $19.59 \pm 2.18$  Å, acompañado de un incremento en la desviación estándar, lo que sugiere una mayor amplitud en la exploración de la compuerta extracelular. En la conformación ocluida, ambas distancias permanecen prácticamente constantes, con valores de EC entre 10.98 y 11.28 Å y de IC entre 11.01 y 11.29 Å, indicando una alta estabilidad estructural. En contraste, la conformación endofacial muestra una disminución moderada en EC (de 11.15 a 10.24 Å) y un aumento en IC (de 11.97 a 12.51 Å), consistente con una mayor apertura hacia el lado intracelular.

Al aplicar el método de convergencia, las simulaciones de 5  $\mu$ s tampoco demostraron aumentar el número efectivo de muestras independientes en comparación con simulaciones de 3  $\mu$ s (Tabla 6). Esto sugiere que el muestreo permanece restringido a cuencas específicas del paisaje energético, lo que sugiere que las barreras cinéticas entre los estados del modelo de acceso alternante no son superadas en las escalas temporales exploradas.



**Figura 16. Exploración conformacional con simulaciones extendidas.**

(A) Esquema general de las simulaciones realizadas para el muestreo del ciclo conformacional, indicando los estados iniciales considerados (exofacial, exofacial ocluida, ocluida, endofacial ocluida y endofacial), la duración de las trayectorias (3  $\mu$ s y 5  $\mu$ s) y el tipo de termostato empleado. Con el objetivo de evaluar la ocurrencia de transiciones entre los estados exofacial y ocluido, las simulaciones iniciadas desde las conformaciones exofacial, ocluida y endofacial fueron extendidas hasta 5  $\mu$ s, incorporando el esquema de termostatización middle. (B) Representación esquemática de las diferencias entre el termostato clásico y el termostato middle, destacando la aplicación del acoplamiento térmico en relación con los pasos de integración. (C) Distribuciones de densidad en el espacio reducido definido por las distancias intracelulares (IC) y extracelulares (EC) para simulaciones de 3  $\mu$ s y 5  $\mu$ s, mostrando el efecto de la extensión temporal y del esquema de termostatización sobre el muestreo conformacional en los estados exofacial, ocluido y endofacial.

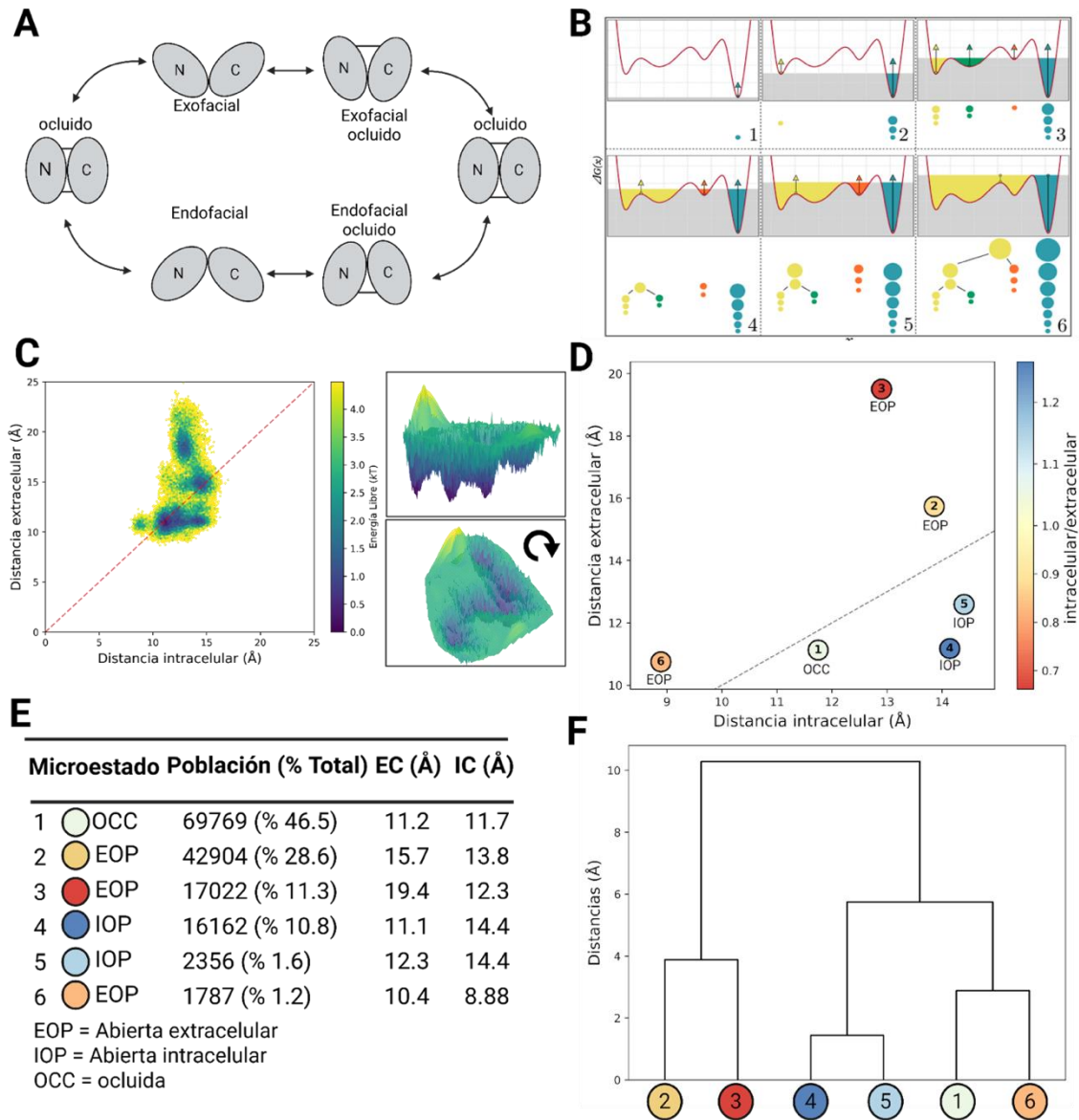
#### 4. Transiciones conformacionales en el espacio explorado

Con el objetivo de cuantificar las probabilidades de transición entre estados conformacionales, se construyeron Modelos de Estados de Markov (MSM). Dado que la propiedad markoviana se define en espacios discretos, el espacio continuo proyectado en las coordenadas IC–EC fue discretizado mediante un algoritmo de *clustering* geométrico basado en densidad (ver metodología). Esto permitió estimar el paisaje de energía libre a partir de las densidades de puntos proyectados en el plano IC–EC, de modo que las regiones de alta densidad son mínimos locales de energía libre (Figura 17C).

Como lo muestra la Figura 17D se identificaron seis microestados en el espacio conformacional accesible, en donde cada uno representa una conformación diferente caracterizada por valores IC y EC diferenciales. El microestado 1 corresponde al estado más poblado del sistema, con 69769 *frames*, lo que equivale al 46.5 % del total de las trayectorias concatenadas. Este estado representa una conformación ocluida. El segundo microestado más poblado corresponde al estado 2, asociado a una conformación abierta hacia el lado extracelular, con 42 904 *frames* (28.6 %). El estado 3 presenta 13 022 *frames* (11.3 %) y corresponde a una conformación extracelular más abierta que la observada en el estado 2. Por su parte, el estado 4, asociado a una conformación abierta hacia el lado intracelular, contiene 16 162 *frames* (10.8 %).

Los microestados menos poblados corresponden a los estados 5 y 6. El estado 5 presenta 2 356 *frames* (1.6 %) y describe una conformación intracelular abierta, aunque con una apertura extracelular ligeramente más restringida. Finalmente, el estado 6 corresponde a un estado ocluido con una apertura intracelular más estrecha en comparación con el estado ocluido predominante (estado 1).

A partir de la discretización del espacio conformacional se estimó la matriz de probabilidad de transición entre microestados utilizando un modelo de estados de Markov. En este marco, la dinámica del sistema se describe mediante una matriz de transición dependiente del tiempo de retardo ( $\tau$ ). La elección de  $\tau$  es crítica para la construcción del modelo, ya que la propiedad markoviana implica que la evolución futura del sistema depende únicamente del estado presente y no de la historia previa del sistema. Es decir, dado que el sistema está en el microestado A, ¿cuál es la probabilidad de que en  $\tau$  pasos esté en B?



**Figura 17. Discretización del espacio conformacional accesible.**

(A) Diagrama esquemático del ciclo conformacional de GLUT5. (B) Metodología de agrupamiento aplicada a discretizar el espacio conformacional (C) Paisaje de energía libre (FEL) del espacio conformacional accesible. (D-F) Mapeo conformacional del modelo de acceso alternativo. Cada esfera corresponde al *frame* representativo del agrupamiento. EC e IC corresponden a la apertura extracelular e intracelular respectivamente.

Para determinar un  $\tau$  adecuado se calcularon las escalas de tiempo implícitas (*implied timescales*) en función del tiempo de retardo  $\tau$  (Anexo). A medida que  $\tau$  aumenta, las escalas de tiempo más lentas tienden a estabilizarse, lo que indica una aproximación al régimen markoviano. En este caso, se observa que las escalas de tiempo principales comienzan a

mostrar un comportamiento aproximadamente estacionario a partir de valores de  $\tau$  cercanos a 1500 *frames*.

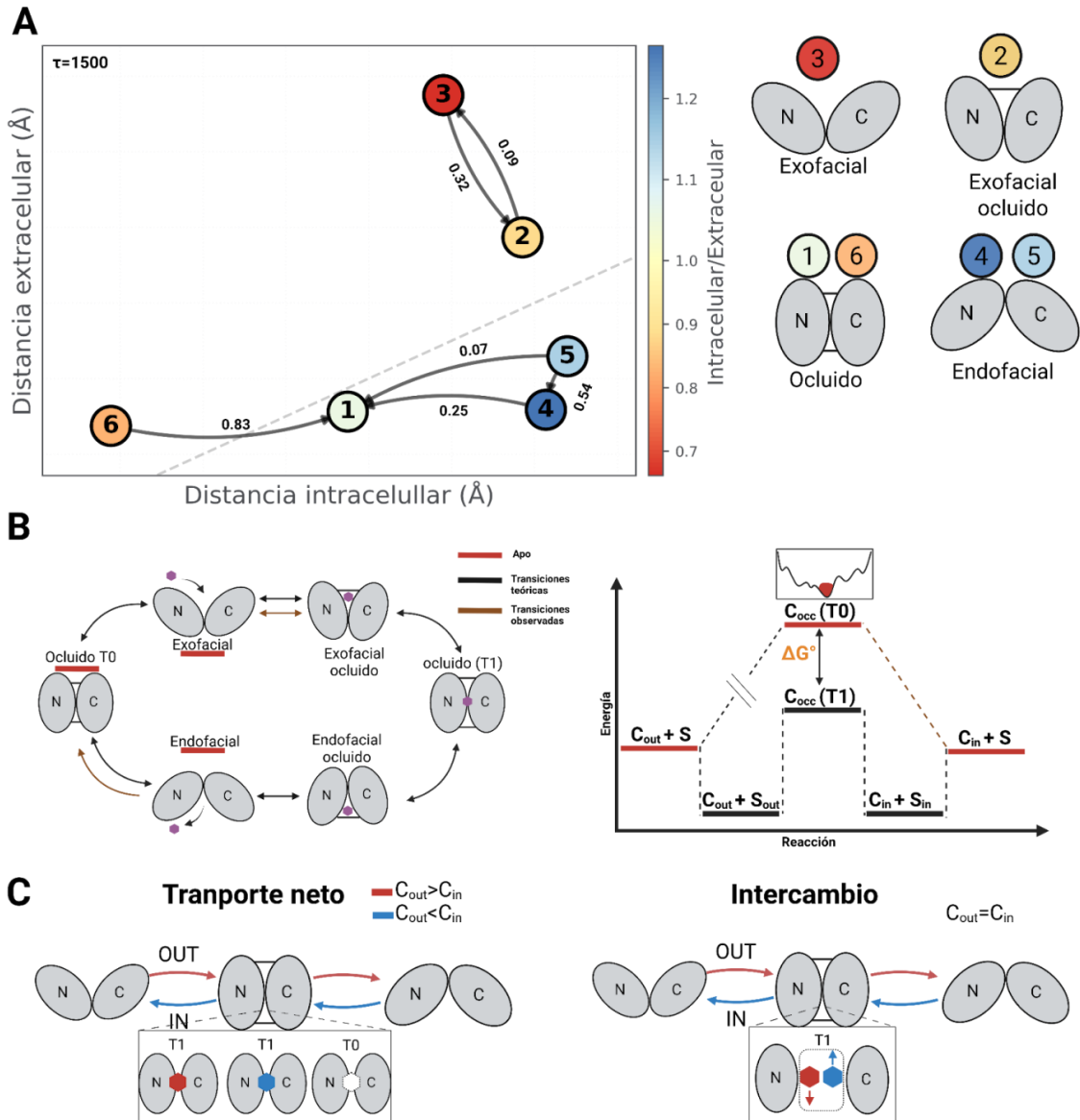
Con base en este comportamiento, se seleccionó un tiempo de retardo de  $\tau=1500$  *frames* para la construcción del modelo de estados de Markov. La validez de esta elección fue posteriormente evaluada mediante la prueba de Chapman–Kolmogorov (Anexo), que compara las probabilidades de permanencia predichas por el modelo con aquellas obtenidas directamente de las trayectorias de dinámica molecular.

Como se observa, las probabilidades estimadas por el modelo reproducen de manera consistente la evolución temporal observada en las simulaciones para los seis estados metastables identificados (S1–S6), lo que respalda la consistencia interna del modelo y su capacidad para describir la dinámica lenta del sistema en la escala temporal considerada.

Habiendo definido el *lag time* adecuado, las probabilidades de transición entre microestados fueron estimadas utilizando un *lag time* de 1500 pasos que se representa en la Figura 18A. Entre las transiciones observadas, destaca la transición desde el estado 3 hacia el estado 2 con una probabilidad de 0.32, mientras que la transición inversa (2→3) presenta una probabilidad menor de 0.09.

En el caso de los estados ocluidos, se observa una alta probabilidad de transición desde el estado 6 hacia el estado 1 ( $P_{6\rightarrow 1}=0.83$ ), lo que indica una fuerte conexión cinética entre ambos estados. Por otra parte, se identificaron transiciones entre estados con apertura intracelular, incluyendo la transición desde el estado 5 hacia el estado 4 con una probabilidad de 0.54. Asimismo, se observaron transiciones desde el estado 5 hacia el estado 1 ( $P_{5\rightarrow 1}=0.07$ ) y desde el estado 4 hacia el estado 1 ( $P_{4\rightarrow 1}=0.25$ ).

En conjunto, los resultados confirman una exploración ineficiente del espacio conformacional explorado, sin embargo, en las escalas utilizadas en este trabajo fue posible determinar transiciones conformacionales entre conformaciones adyacentes del mecanismo de acceso alternante. En particular, las conformaciones exofaciales y exofaciales ocluidas presentan conectividad reducida y aparecen prácticamente desacopladas de la red principal de transiciones. En contraste, las conformaciones ocluidas y endofaciales muestran mayor interconectividad y mayores probabilidades de transición entre sí, indicando que dominan la dinámica observable en las escalas temporales exploradas (Figura 18B). El análisis de las probabilidades de transición revela además una asimetría cinética. En los estados exofaciales, la transición desde una conformación abierta hacia una parcialmente ocluida es más probable que el proceso inverso, lo que sugiere una tendencia hacia la oclusión parcial sin progresar hacia estados completamente cerrados. En cambio, las conformaciones endofaciales muestran mayor propensión a transitar hacia estados ocluidos, evidenciando una preferencia cinética por la oclusión desde configuraciones endofaciales.



**Figura 18. Red de probabilidades de transición en la exploración accesible.**

(A) Red de transiciones del MSM proyectada sobre el plano IC/EC. Cada nodo corresponde a un microestado de la discretización conformacional y el color refleja el cociente IC/EC como descriptor estructural. Las flechas indican transiciones cinéticamente relevantes a *lag time*  $\tau = 1500$ , con sus probabilidades asociadas. El MSM fue validado mediante *implied timescales* y test de Chapman-Kolmogorov (Anexo). La figura derecha esquematiza la conformación de cada microestado. (B) Aporte de esta tesis al modelo de acceso alternativo. Las simulaciones DM-GG no sesgadas caracterizan transiciones entre conformaciones exofaciales, endofaciales y ocluidas en estado apo. El estado ocluido post-translocación (T0) no fue accesible dentro de las escalas temporales simuladas, consistente con la elevada barrera energética descrita en la literatura. (C) Proyección metodológica. El marco cinético obtenido orienta el diseño de simulaciones futuras: el intercambio puede abordarse con sustrato en ambos compartimentos, favoreciendo transiciones holo locales sin requerir el estado ocluido apo, mientras que el transporte neto requiere gradientes de concentración y unión secuencial del sustrato. La ubicación del sustrato emerge, así como variable clave para determinar qué componentes del MAA son explorables computacionalmente.

## DISCUSIÓN

Nuestras células contienen miles de proteínas que realizan diversas funciones. El paradigma central de la biología estructural establecía que el plegamiento tridimensional de una proteína está codificado en la secuencia (Wright & Dyson, 2015). Sin embargo, con el avance de diversas técnicas experimentales, tales como: la cristalografía de rayos X de resolución temporal de alto rendimiento, la criomicroscopía electrónica (cryo-EM) y la resonancia magnética nuclear (RMN), junto con técnicas complementarias como la dispersión de rayos X de ángulo pequeño (SAXS) y la transferencia de energía de resonancia de Förster (FRET), fue posible resolver y obtener información de sistemas cada vez más complejos, muchos de ellos en más de una configuración espacial o conformación, lo que permitió expandir esta visión clásica hacia un cambio paradigmático que centra el dinamismo estructural de las proteínas en la actividad biológica (Jiang & Kalodimos, 2017; Kimanius et al., 2015; Okamoto & Sako, 2017; Ourmazd, 2019). Por lo tanto, lejos de ser estructuras estáticas, ahora está claro que las proteínas se comportan más bien como entidades dinámicas (Henzler-Wildman & Kern, 2007). Comprender estos cambios conformacionales es útil para el diseño de fármacos, la ingeniería de proteínas y la caracterización de deficiencias mecánicas en proteínas mutantes dada una condición fisiopatológica (Orellana, 2019).

Si bien los enfoques experimentales existentes son eficientes en captar algunas de estas conformaciones, no permiten capturar la totalidad del paisaje conformacional. Es por esto por lo que uno de los deberes de la química computacional es desarrollar nuevas metodologías que permitan dilucidar estos fenómenos conformacionales. Una de las herramientas más útiles es la DM, en la cual se cuantifica la disposición espacial de los átomos en el tiempo a partir de la integración de las leyes de Newton; es por esto por lo que

algunos la denominan "microscopio computacional" (Dror et al., 2012; Karplus & McCammon, 2002).

A pesar de las positivas contribuciones de esta técnica al entendimiento de numerosos fenómenos bioquímicos asociados a proteínas, aplicarla a la comprensión de cambios conformacionales es una tarea que requiere explorar o muestrear dichas conformaciones. Esto parece obvio, pero ¿cómo podemos determinar algo que no observamos? Dependiendo del grado de reorganización, que puede ir desde el movimiento de unos pocos aminoácidos hasta dominios completos, el tiempo requerido para observar dichas transiciones involucra simulaciones en escalas entre  $\mu\text{s}$  -  $\text{ms}$ , e incluso segundos, lo que representa un costo computacional elevado (Shaw et al., 2009; Sweet et al., 2013).

Aparte de estos aspectos técnicos, existe un problema de muestreo (*sampling problem*) que limita la exploración de todas las conformaciones (Orellana, 2019; Yang et al., 2019). Esto se debe a que, en procesos biológicos mediados por proteínas, cada conformación ocupa un mínimo local, de forma que el paisaje energético (FEL) de un transportador de membrana, por ejemplo, contiene múltiples mínimos locales separados por barreras energéticas (Costa et al., 2015; McComas et al., 2022; Mitrovic et al., 2023). La exploración del espacio de fases en cada mínimo local supondría entonces una perturbación que otorgue la energía necesaria para superar dichas barreras y así explorar todo el paisaje conformacional.

Afortunadamente, existen variantes de la dinámica molecular que permiten explorar de forma más exhaustiva el espacio conformacional. Estas técnicas de mejoramiento (*Enhanced Molecular Dynamics*) integran potenciales adicionales a la integración numérica; ejemplos de ellas son *umbrella sampling*, *targeted molecular dynamics* y *accelerated molecular dynamics* (Govind Kumar et al., 2023; Miao et al., 2015; Schlitter et al., 1994). Sin embargo,

al introducir un potencial externo en la exploración, no queda claro si las conformaciones determinadas existen en el equilibrio químico. Por otro lado, otra estrategia para la exploración conformacional consiste en disminuir los grados de libertad del sistema, reduciendo así el costo computacional de las simulaciones. En el presente trabajo se utilizó simulaciones de Dinámica Molecular de Grano Grueso (DM-GG) para responder a la pregunta: ¿es posible muestrear transiciones conformacionales en el transportador de fructosa de *Rattus norvegicus* (rGLUT5) a partir del monitoreo de los reordenamientos globales? La elección de esta técnica, por sobre otras de mejoramiento del muestreo, radica en el interés por observar el fenómeno de acceso alternante en condiciones de equilibrio, evitando introducir sesgos. Además, la accesibilidad computacional de la técnica permite reducir los grados de libertad del sistema y, con ello, el costo computacional asociado (Klein et al., 2023).

La elección de rGLUT5 no fue arbitraria, y se justifica por los siguientes argumentos. El primero es su relevancia fisiológica: su asociación con trastornos metabólicos como obesidad, diabetes tipo 2 y cáncer lo posiciona como un modelo adecuado para estudios conformacionales (Barone et al., 2009; Groenendyk et al., 2022; Wang et al., 2020). El segundo se debe a la disponibilidad de información estructural. La superfamilia de facilitadores principales (MFS), en particular los miembros de la familia *Sugar Porter* (SP), se caracteriza por agrupar transportadores en múltiples conformaciones, lo que facilitó el modelamiento comparativo de las conformaciones que explican el modelo de acceso alternante: exofacial, exofacial ocluida, ocluida, endofacial ocluida y endofacial (Drew et al., 2021). En tercer lugar, se eligió rGLUT5 por sobre otros transportadores porque una pregunta similar a la de este trabajo fue planteada por el equipo de Sarah E. McComas, quienes la

respondieron desde otra aproximación metodológica, acoplando el *string method with swarms of trajectories* al muestreo conformacional con *targeted molecular dynamics* en rGLUT5 (McComas et al., 2022).

Es importante mencionar que, a pesar de ser una pregunta similar, existen matices diferentes. Para el equipo de McComas, la pregunta específica fue: ¿cómo podría ocurrir la transición entre dos estados conocidos?, mientras que este trabajo plantea: ¿cómo es la exploración conformacional de cada conformación en simulaciones que no introducen sesgos? Una segunda aclaración metodológica corresponde a la justificación del uso de modelos de estado de Markov como método estocástico para observar transiciones conformacionales (Nagel & Stock, 2023). A diferencia de otras técnicas de reconstrucción, como el *string method with swarms of trajectories*, los modelos de estados de Markov no identifican caminos de transición de mínima energía dados dos estados; su aplicación es más modesta: dadas dos conformaciones  $(i, j)$ , ¿cuál es la probabilidad de llegar a  $j$  estando previamente en la conformación  $i$ ? Esto se ajusta a la metodología de grano grueso, ya que la precisión con que se calcula la energía libre en este nivel de representación no es la adecuada: al agrupar varios grados de libertad en un solo sitio de interacción, calcular energía libre en esta topología sería formalmente incorrecto.

Respecto a los resultados, el presente trabajo corroboró la utilidad de las aperturas intracelulares y extracelulares como descriptores globales capaces de discriminar conformaciones en el modelo de acceso alternante para todos los transportadores de la familia SP. Si bien esta aproximación es útil para la clasificación de conformaciones en topología de grano grueso, podría resultar una limitante por sí sola para cuantificar energías libres en sistemas atómicos. Esta limitación se debe a que la caracterización de las barreras

energéticas entre estados metaestables requiere coordenadas colectivas adicionales, capaces de distinguir con mayor resolución entre conformaciones estructuralmente similares, pero cinéticamente distintas.

En este contexto, estudios recientes han abordado este desafío de manera más explícita. Un trabajo publicado en 2023 combinó análisis de coevolución con herramientas de aprendizaje automático para identificar contactos estructurales específicos de estado que median los cambios conformacionales en GLUT5, demostrando que la topología del paisaje energético está determinada en gran medida por redes de interacciones locales dependientes del estado conformacional. En particular, este estudio identificó contactos estructurales críticos que no pueden capturarse mediante descriptores geométricos globales. Entre los más relevantes se encuentran: (1) la red de puentes salinos entre las hélices TM4–TM10b, cuya ruptura progresiva es determinante para la transición desde el estado ocluido al estado de cara interna; (2) las interacciones entre TM7b y TM1, que coevolucionaron para estabilizar el estado ocluido mediante el contacto entre Tyr297–Asn40, residuos que modulan el cierre extracelular; y (3) los contactos entre TM5–TM8, que estabilizan las conformaciones de cara externa a través de interacciones específicas como Asn157–Glu337. Estas interacciones representan eventos moleculares discretos cuya presencia o ausencia define la identidad funcional de cada estado conformacional, y que permanecen invisibles para coordenadas colectivas basadas únicamente en distancias de apertura global.

En línea con el trabajo de Mitrovic et al. (2023), en esta tesis se evaluó la incorporación de dichos contactos estructurales con el fin de mejorar la discriminación del paisaje conformacional; sin embargo, su implementación se vio limitada por la falta de disponibilidad pública de los algoritmos y modelos utilizados en el estudio original. La

aplicación de este enfoque habría requerido el desarrollo y entrenamiento de modelos supervisados adicionales, como redes neuronales convolucionales para clasificación de estados conformacionales y máquinas de soporte vectorial (SVM) para la extracción de contactos específicos de estado (Mitrovic et al., 2023), lo que excedía los objetivos y el alcance metodológico de esta tesis.

No obstante, esta limitación abre una línea concreta de trabajo futuro. Una propuesta metodológicamente viable consistiría en incorporar como variables colectivas adicionales las distancias mínimas entre los pares de residuos identificados por Mitrovic et al. (2023) como críticos para cada transición conformacional, específicamente los contactos TM4–TM10b para la transición ocluida→endofacial ocluida, los contactos TM7b–TM1 para la estabilización del estado ocluido, y los contactos TM5–TM8 para la discriminación de la conformación exofacial. Estas variables colectivas basadas en contactos locales podrían combinarse con los descriptores geométricos globales ya empleados en esta tesis, permitiendo una caracterización más precisa de las barreras energéticas entre estados y una estimación más robusta del paisaje de energía libre a lo largo del ciclo de transporte completo.

Un segundo resultado relacionado con la exploración conformacional mediante grano grueso es la ausencia de la transición completa del modelo de acceso alternante. En particular, utilizando simulaciones de 3  $\mu$ s fue posible observar transiciones conformacionales espontáneas entre estados adyacentes del ciclo: los modelos de estados de Markov identificaron seis microestados, con transiciones predominantes entre conformaciones exofaciales y exofaciales ocluidas, y entre conformaciones ocluidas y endofaciales, con probabilidades de hasta 0.83 entre subestados ocluidos y 0.32 entre estados exofaciales. A pesar de haberse extendido la duración de las simulaciones, acoplado un termostato middle

para capturar con mayor precisión las energías cinéticas y, por tanto, las posiciones cartesianas de los átomos en el sistema, no fue posible observar transiciones entre conformaciones ubicadas en los extremos del modelo de acceso alternante, de forma que las conformaciones exofaciales quedan 'desconectadas' de las conformaciones ocluidas y, aún más, de las endofaciales. Esta observación es consistente con estudios de energía libre que muestran que, en ausencia de sustrato, el estado exofacial de GLUT5 se ubica en un pozo energético menos profundo, separado por barreras energéticas elevadas de los estados ocluidos y endofaciales (McComas et al., 2022). Este comportamiento difiere del observado en presencia de fructosa, donde las barreras energéticas disminuyen y los estados exofaciales se conectan, facilitando el cambio conformacional de acuerdo con el modelo de acceso alternante (Ebert et al., 2018; McComas et al., 2022; Mitrovic et al., 2023; Rana et al., 2022). Lo anterior pareciera ser propio de los facilitadores, ya que se han observado comportamientos similares en otros miembros de la familia SP, reportados para Xyle y GLUT1 mediante técnicas que introducen sesgo (Jia et al., 2020; Liu et al., 2024; Martens et al., 2018). Desde esta perspectiva, la ausencia de conectividad continua observada en este trabajo no contradice el mecanismo funcional de GLUT5, sino que es consistente con la dinámica conformacional esperada para el estado apo del transportador.

Afortunadamente, durante la ejecución de este trabajo se diseñó una metodología automatizada que involucra el modelamiento de conformaciones, el ensamble en topología de grano grueso y la ejecución de simulaciones para el muestreo conformacional. Esto permite que el diseño y la ejecución de simulaciones futuras sea rápida y accesible. Una de estas podría ser la simulación considerando sustratos; sin embargo, esto depende de las preguntas que se intenten responder. Incorporar el sustrato no debe ser un esfuerzo orientado

a forzar transiciones superando la barrera energética, sino una oportunidad para cuestionar el mecanismo de transporte: ¿cómo cambia la estabilidad de los estados?, ¿cómo se acopla la unión del sustrato con el cambio conformacional?, ¿la selectividad al sustrato depende de la conformación?, ¿el mecanismo de transporte difiere cuando existe diferencia de potencial químico respecto a cuando el sistema se encuentra en equilibrio? En la práctica, esto podría significar estudiar el transporte neto, en donde el mecanismo de acceso alternante involucra al menos cinco conformaciones. En cambio, en el intercambio acelerado, que ocurre cuando los potenciales químicos se encuentran en equilibrio, se trataría de determinar si la presencia simultánea de sustrato favorece transiciones conformacionales que no requieran el retorno del transportador vacío (Domene et al., 2025; Iglesias-Fernandez et al., 2017; Naftalin, 2018). Es importante aclarar que el intercambio acelerado no indica incompatibilidad con el modelo de acceso alternante, sino que cuestiona el mecanismo secuencial de las transiciones conformacionales (exofacial → ocluida → endofacial → ocluida → exofacial).

Otra proyección relevante de este trabajo se relaciona con la composición lipídica de la membrana. Los sistemas simulados consideran una bicapa simplificada compuesta exclusivamente por fosfatidilcolina (POPC), un fosfolípido de carga neta neutra, lo cual constituye una aproximación ampliamente utilizada en estudios computacionales de proteínas de membrana (Martens et al., 2018). Sin embargo, la membrana intestinal donde se expresan los transportadores GLUT5 presenta una composición significativamente más compleja, incluyendo glucolípidos, colesterol, esfingomielina y fosfatidilserina. La evidencia experimental sugiere que esta complejidad lipídica no es meramente estructural, sino que cumple un rol funcional clave. En particular, estudios en liposomas reconstituidos han demostrado que GLUT5 es activo únicamente dentro de un rango específico de fluidez de

membrana, y que altas concentraciones de fosfolípidos cargados negativamente, como la fosfatidilserina, pueden inducir un “arresto” del transportador en una conformación endofacial (Suades et al., 2023).

En este contexto, resulta relevante proyectar la incorporación de modelos de membrana más complejos en futuras simulaciones, que incluyan componentes como colesterol, esfingomielina, fosfatidilserina y glucolípidos en proporciones representativas del entorno intestinal. Este enfoque permitiría evaluar de manera más precisa el impacto de la composición lipídica sobre la estabilidad conformacional y la dinámica del transportador, avanzando hacia una descripción más fisiológicamente realista del mecanismo de transporte.

Finalmente, si bien la DM-GG permitió reducir significativamente el costo computacional y explorar transiciones globales entre conformaciones adyacentes, esta aproximación implica limitaciones inherentes respecto dinámicas moleculares atomísticas. La reducción de los grados de libertad mediante la agrupación de átomos en sitios de interacción (*beads*) conlleva una pérdida de resolución química que dificulta la caracterización precisa de las interacciones locales que definen los cambios conformacionales, tales como puentes de hidrógeno e interacciones electrostáticas. Para superar estas limitaciones, una estrategia metodológica sería el uso de simulaciones de retromapeado (*backmapping*), en las que se recupera la información atómica a partir de la geometría de los *beads*, permitiendo refinar la descripción de las interacciones locales y calcular energías libres con mayor precisión (Machado et al., 2016).

## CONCLUSIÓN

El principal aporte de este trabajo es el desarrollo de una metodología automatizada y reproducible para ensamblar, parametrizar y simular conformaciones del modelo de acceso alternante en transportadores de la familia SP, aplicada aquí a rGLUT5. Este marco integra modelamiento comparativo, ensamble en topología de grano grueso y simulaciones de DM-GG con modelos de estados de Markov, y está diseñado para ser transferible a otros miembros de la familia sin modificaciones sustanciales al protocolo.

La representación en topología CG redujo el número de partículas del sistema en un 86.9%, extendiendo las escalas temporales accesibles sin modificar el Hamiltoniano del sistema ni introducir fuerzas externas. En simulaciones de 3  $\mu\text{s}$  fue posible observar transiciones conformacionales espontáneas entre estados adyacentes del ciclo: los modelos de estados de Markov identificaron seis microestados, con transiciones predominantes entre conformaciones exofaciales y exofaciales ocluidas, y entre conformaciones ocluidas y endofaciales, con probabilidades de hasta 0.83 entre subestados ocluidos y 0.32 entre estados exofaciales. Las conformaciones exofaciales y endofaciales, sin embargo, aparecieron desacopladas entre sí, sin conectividad cinética continua a lo largo del ciclo completo.

La extensión de las simulaciones a 5  $\mu\text{s}$  y la incorporación del integrador *middle* no produjeron mejoras observables en la exploración conformacional: los rangos de apertura de las compuertas variaron menos de 0.3 Å respecto a las simulaciones de 3  $\mu\text{s}$  y el número efectivo de muestras independientes no aumentó. Esto indica que la restricción no es temporal sino energética: en ausencia de sustrato, las barreras cinéticas entre los estados extremos del ciclo no son superadas bajo dinámica no sesgada, comportamiento consistente

con el paisaje energético reportado para el estado apo de rGLUT5 y otros miembros de la familia SP.

En conjunto, la hipótesis se acepta de manera parcial: las simulaciones de grano grueso capturan transiciones globales entre conformaciones adyacentes, pero no logran reconstruir el ciclo conformacional completo en ausencia de sustrato. El marco metodológico desarrollado sienta las bases para incorporar sustrato explícito y descriptores conformacionales de mayor resolución en trabajos futuros, avanzando hacia una descripción completa del mecanismo de transporte de rGLUT5.

## GLOSARIO

- Acceso alternante, modelo de (MAA): modelo conceptual que describe el mecanismo de transporte de solutos a través de la membrana mediante transiciones secuenciales entre conformaciones que exponen el sitio de unión alternativamente hacia el exterior celular o el citoplasma, sin acceso simultáneo desde ambos lados.
- Antiportador: proteína transportadora que media el intercambio acoplado de dos solutos en direcciones opuestas a través de la membrana.
- *Bead*: partícula efectiva utilizada en simulaciones de grano grueso que representa un conjunto de átomos con propiedades fisicoquímicas colectivas similares.
- *Block averaging* (promediado por bloques): método estadístico que evalúa la convergencia de una simulación dividiendo la trayectoria en bloques de tamaño creciente y analizando la estabilidad de la desviación estándar entre bloques. Permite estimar el tiempo de autocorrelación integrado y el número efectivo de muestras independientes.
- Campo de fuerza (*force field*): conjunto de funciones matemáticas y parámetros empíricos que describen las interacciones interatómicas o entre *beads* en una simulación de dinámica molecular, incluyendo términos enlazantes y no enlazantes.
- Chapman-Kolmogorov: Prueba de validación de los modelos de estados de Markov que compara las probabilidades de transición predichas por el modelo con las observadas directamente en las trayectorias de simulación a distintas escalas temporales.
- *Clustering* geométrico basado en densidad: algoritmo de agrupamiento que identifica microestados del espacio conformacional utilizando regiones de alta densidad como semillas, sin requerir un número inicial de grupos. Permite aprovechar la naturaleza

jerárquica del paisaje de energía libre para identificar estados separados por barreras energéticas.

- Conformación endofacial: estado conformacional de un transportador en el que el sitio de unión al sustrato está orientado hacia el citoplasma.
- Conformación exofacial: estado conformacional de un transportador en el que el sitio de unión al sustrato está orientado hacia el espacio extracelular.
- Conformación ocluida: estado conformacional intermedio del modelo de acceso alternante en el que el sitio de unión queda inaccesible desde ambos lados de la membrana.
- Coordenadas colectivas (CVs, *collective variables*): variables de baja dimensionalidad que describen los grados de libertad lentos y funcionalmente relevantes de un sistema biomolecular, como desplazamientos relativos entre dominios o distancias entre compuertas. Permiten proyectar la dinámica sobre un espacio reducido e interpretable.
- DOPE (*Discrete Optimized Protein Energy*): potencial estadístico basado en distribuciones de distancias interatómicas derivadas de estructuras cristalográficas, utilizado para evaluar la calidad estructural de modelos proteicos generados por modelamiento comparativo. Valores más negativos indican mayor coherencia estructural.
- Dinámica molecular (DM): técnica computacional que simula el movimiento de los átomos de un sistema integrando numéricamente las ecuaciones de movimiento de Newton en función del tiempo, permitiendo describir la evolución microscópica del sistema.
- Dinámica molecular de grano grueso (DM-GG): variante de la dinámica molecular en la que múltiples átomos se agrupan en partículas efectivas denominadas *beads*, reduciendo

los grados de libertad del sistema y extendiendo las escalas temporales accesibles sin modificar el Hamiltoniano ni introducir fuerzas externas.

- Dinámica molecular dirigida (SMD, *Steered Molecular Dynamics*): simulación fuera del equilibrio que impone un potencial armónico móvil sobre una coordenada de reacción definida a priori, forzando al sistema a recorrer regiones del espacio conformacional que no visitaría espontáneamente.
- Dinámica molecular dirigida (TMD, *Targeted Molecular Dynamics*): simulación fuera del equilibrio que sesga el sistema mediante una métrica global de similitud respecto a una conformación final conocida, facilitando la exploración de reordenamientos de gran escala.
- Ensamble NPT: condiciones de simulación en las que el número de partículas (N), la presión (P) y la temperatura (T) se mantienen constantes, correspondiente al ensamble isotérmico-isobárico.
- Espacio conformacional: conjunto de configuraciones estructurales accesibles a una proteína, determinado por su paisaje de energía libre.
- Espacio de fases: espacio matemático de dimensión  $6N$  definido por las posiciones y momentos de todos los átomos de un sistema, cuya exploración describe la evolución microscópica del sistema durante una simulación.
- Estado apo: estado de una proteína en ausencia de sustrato.
- Familia de los transportadores de azúcares (SP, *Sugar Porter*): familia de transportadores de la superfamilia MFS clasificada como TC 2.A.1.1, que incluye los transportadores de hexosas GLUT. Se distingue del resto de MFS por la presencia del motivo SP formado por hélices intracelulares conservadas.

- GLUT5 (SLC2A5): transportador facilitador de fructosa perteneciente a la familia SP, expresado predominantemente en intestino delgado. Clasificado como TC 2.A.1.1.13 en TCDB. Ha sido vinculado a obesidad, diabetes tipo 2 y sobreexpresión tumoral.
- Granularidad: grado de agrupamiento atómico en un modelo de grano grueso. A mayor granularidad, más átomos se agrupan en cada *bead* y menor es el detalle estructural conservado; a menor granularidad, la representación se acerca más a la descripción atomística.
- Hamiltoniano: función matemática que describe la energía total de un sistema como la suma de la energía cinética y la energía potencial.
- HGNC (*HUGO Gene Nomenclature Committee*): organismo que organiza las proteínas del genoma humano según criterios de nomenclatura e interpretación fisiológica.
- HMMs (*Hidden Markov Models*, modelos ocultos de Markov): modelos estadísticos utilizados por Pfam para identificar dominios proteicos conservados mediante alineamientos múltiples de secuencias.
- *Implied timescales* (escalas de tiempo implícitas): escalas temporales características asociadas a los modos dinámicos más lentos de un sistema, utilizadas para determinar el *lag time* adecuado en la construcción de modelos de estados de Markov.
- *Integrator middle*: integrador en el que el termostato de Langevin se aplica en el centro del paso de integración, sobre velocidades que ya incorporan las fuerzas físicas del sistema, produciendo distribuciones estadísticas más cercanas al ensamble canónico correcto.
- *Lag time* (tiempo de retardo,  $\tau$ ): intervalo temporal utilizado en la construcción de modelos de estados de Markov para estimar las probabilidades de transición entre estados, seleccionado cuando las escalas de tiempo implícitas del sistema se estabilizan.

- Microestado: agrupación discreta de configuraciones estructurales similares identificada mediante *clustering* en el espacio de coordenadas colectivas, utilizada como unidad básica en la construcción de modelos de estados de Markov.
- Modelamiento comparativo: método de predicción de estructura proteica que construye un modelo tridimensional a partir de la homología con estructuras experimentales conocidas utilizadas como plantillas.
- MODELLER: programa de modelamiento comparativo utilizado para generar modelos estructurales tridimensionales de proteínas a partir de alineamientos con plantillas experimentales.
- Modelos de estados de Markov (MSM, *Markov State Models*): marco estadístico que describe la dinámica conformacional de un sistema mediante probabilidades de transición entre estados metaestables discretizados, asumiendo que la evolución futura del sistema depende únicamente del estado presente.
- $N_{\text{eff}}$  (número efectivo de muestras independientes): estimación del número de configuraciones estadísticamente independientes en una trayectoria de dinámica molecular, calculado a partir del tiempo de autocorrelación integrado. Un  $N_{\text{eff}}$  mayor indica un muestreo más eficiente del espacio conformacional.
- Paisaje de energía libre (FEL, *Free Energy Landscape*): representación de la energía libre del sistema en función de coordenadas colectivas, en la que los mínimos locales corresponden a estados conformacionales estables y las barreras entre ellos determinan las tasas de transición.
- PACKMOL-Memgen: herramienta utilizada para ensamblar proteínas en bicapas lipídicas y construir sistemas de simulación proteína-membrana.

- Pfam (*Protein Families Database*): base de datos que agrupa proteínas por dominios estructurales conservados, identificados mediante alineamientos múltiples y modelos ocultos de Markov.
- POPC (palmitoil-oleoil-fosfatidilcolina): fosfolípido utilizado para construir la bicapa lipídica en los sistemas de simulación de esta tesis.
- Problema de muestreo (*sampling problem*): limitación de las simulaciones de dinámica molecular para explorar todo el espacio conformacional relevante, consecuencia de las barreras cinéticas que separan los estados metaestables y de la brecha entre las escalas temporales del sistema y las accesibles computacionalmente.
- PROPKA: método utilizado para predecir los estados de protonación de los residuos ionizables de una proteína a pH fisiológico.
- rGLUT5: isoforma del transportador GLUT5 de *Rattus norvegicus*, utilizada en este trabajo como modelo de estudio del mecanismo de acceso alternante en transportadores de azúcares.
- *Rocker-switch*: mecanismo de acceso alternante en el que los dominios N- y C-terminal del transportador se balancean coordinadamente alrededor de un eje que atraviesa el sitio de unión, exponiendo la cavidad alternativamente hacia el exterior o el citoplasma. Es el mecanismo característico de los transportadores SP.
- SIRAH (*Southamerican Initiative for a Rapid and Accurate Hamiltonian*): campo de fuerza de grano grueso que adapta el esquema de mapeo a la naturaleza fisicoquímica de cada residuo, priorizando la conservación de la topología molecular y la geometría estructural. Incluye parametrización para proteínas, lípidos, ácidos nucleicos, agua e iones.

- Simportador: proteína transportadora en la que el ingreso del sustrato se acopla al de un ion, generalmente un protón, en la misma dirección.
- Superfamilia MFS (*Major Facilitator Superfamily*, Superfamilia de Facilitadores Principales): superfamilia de transportadores de membrana con origen evolutivo estimado en más de 3.000 millones de años, presente en bacterias, arqueas y eucariotas, que opera mediante uniporte, simporte y antiporte y transporta una amplia variedad de sustratos.
- TCDB (*Transporter Classification Database*): base de datos respaldada por la IUBMB que clasifica transportadores de membrana integrando criterios funcionales, mecánicos, topológicos y filogenéticos en una jerarquía de cinco niveles comparable a la clasificación EC de las enzimas.
- *TM-align*: herramienta de alineamiento estructural de proteínas basada en la puntuación *TM-score*, utilizada en esta tesis para alinear rGLUT5 con las plantillas estructurales de cada conformación.
- $\tau_{\text{INT}}$  (tiempo de autocorrelación integrado): medida de la correlación temporal en una trayectoria de dinámica molecular que indica cuántos pasos deben transcurrir para obtener una muestra estadísticamente independiente. Valores elevados de  $\tau_{\text{INT}}$  indican que el sistema explora lentamente el espacio conformacional, reduciendo la eficiencia estadística del muestreo.
- Uniportador: proteína transportadora que facilita el desplazamiento pasivo de un soluto a favor de su gradiente electroquímico sin acoplamiento a otras especies.

## REFERENCIAS

- Abramson, J., Smirnova, I., Kasho, V., Verner, G., Kaback, H. R., & Iwata, S. (2003). Structure and Mechanism of the Lactose Permease of *Escherichia coli*. *Science*, 301(5633), 610-615. <https://doi.org/10.1126/science.1088196>
- Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2002). *The molecular biology of the cell* (Vols. 1-4). <https://hero.epa.gov/reference/1060032/>
- Alleva, C., Drew, D., & Delemotte, L. (2022). Molecular modeling investigation of proton coupling in Xyle. *Biophysical Journal*, 121(3), 468a. <https://doi.org/10.1016/j.bpj.2021.11.456>
- Andersson, M., Bondar, A.-N., Freitas, J. A., Tobias, D. J., Kaback, H. R., & White, S. H. (2012). Proton-coupled dynamics in lactose permease. *Structure* (London, England: 1993), 20(11), 1893-1904. <https://doi.org/10.1016/j.str.2012.08.021>
- Barone, S., Fussell, S. L., Singh, A. K., Lucas, F., Xu, J., Kim, C., Wu, X., Yu, Y., Amlal, H., Seidler, U., Zuo, J., & Soleimani, M. (2009). Slc2a5 (Glut5) is essential for the absorption of fructose in the intestine and generation of fructose-induced hypertension. *The Journal of Biological Chemistry*, 284(8), 5056-5066. <https://doi.org/10.1074/jbc.M808128200>
- Barrera, E. E., Machado, M. R., & Pantano, S. (2019). Fat SIRAH: Coarse-Grained Phospholipids To Explore Membrane-Protein Dynamics. *Journal of Chemical Theory and Computation*, 15(10), 5674-5688. <https://doi.org/10.1021/acs.jctc.9b00435>
- Bonati, L., Rizzi, V., & Parrinello, M. (2020). Data-Driven Collective Variables for Enhanced Sampling. *The Journal of Physical Chemistry Letters*, 11(8), 2998-3004. <https://doi.org/10.1021/acs.jpcllett.0c00535>
- Borges-Araújo, L., Patmanidis, I., Singh, A. P., Santos, L. H. S., Sieradzan, A. K., Vanni, S., Czaplewski, C., Pantano, S., Shinoda, W., Monticelli, L., Liwo, A., Marrink, S. J., & Souza, P. C. T. (2023). Pragmatic Coarse-Graining of Proteins: Models and Applications. *Journal of Chemical Theory and Computation*, 19(20), 7112-7135. <https://doi.org/10.1021/acs.jctc.3c00733>
- Case, D. A., Aktulga, H. M., Belfon, K., Cerutti, D. S., Cisneros, G. A., Cruzeiro, V. W. D., Forouzes, N., Giese, T. J., Götz, A. W., Gohlke, H., Izadi, S., Kasavajhala, K., Kaymak, M. C., King, E., Kurtzman, T., Lee, T.-S., Li, P., Liu, J., Luchko, T., ... Merz, K. M. Jr. (2023a). AmberTools. *Journal of Chemical Information and Modeling*, 63(20), 6183-6191. <https://doi.org/10.1021/acs.jcim.3c01153>
- Chang, H. S., Fard, P. J., Marcus, S. I., & Shayman, M. (2003). Multitime scale Markov decision processes. *IEEE Transactions on Automatic Control*, 48(6), 976-987. <https://doi.org/10.1109/TAC.2003.812782>
- Cheng, Y. (2018). Membrane protein structural biology in the era of single particle cryo-EM. *Current Opinion in Structural Biology, Cryo electron microscopy: the impact of the cryo-EM revolution in biology • Biophysical and computational methods - Part A*, 52, 58-63. <https://doi.org/10.1016/j.sbi.2018.08.008>

- Clarage, J. B., Romo, T., Andrews, B. K., Pettitt, B. M., & Phillips, G. N. (1995). A sampling problem in molecular dynamics simulations of macromolecules. *Proceedings of the National Academy of Sciences*, 92(8), 3288-3292. <https://doi.org/10.1073/pnas.92.8.3288>
- Costa, M. G. S., Batista, P. R., Bisch, P. M., & Perahia, D. (2015). Exploring Free Energy Landscapes of Large Conformational Changes: Molecular Dynamics with Excited Normal Modes. *Journal of Chemical Theory and Computation*, 11(6), 2755-2767. <https://doi.org/10.1021/acs.jctc.5b00003>
- Dans, P. D., Zeida, A., Machado, M. R., & Pantano, S. (2010). A Coarse Grained Model for Atomic-Detailed DNA Simulations with Explicit Electrostatics. *Journal of Chemical Theory and Computation*, 6(5), 1711-1725. <https://doi.org/10.1021/ct900653p>
- Darden, T., York, D., & Pedersen, L. (1993). Particle mesh Ewald: An  $N \cdot \log(N)$  method for Ewald sums in large systems. *The Journal of Chemical Physics*, 98(12), 10089-10092. <https://doi.org/10.1063/1.464397>
- Darré, L., Machado, M. R., Brandner, A. F., González, H. C., Ferreira, S., & Pantano, S. (2015). SIRAH: A Structurally Unbiased Coarse-Grained Force Field for Proteins with Aqueous Solvation and Long-Range Electrostatics. *Journal of Chemical Theory and Computation*, 11(2), 723-739. <https://doi.org/10.1021/ct5007746>
- Deng, D., Sun, P., Yan, C., Ke, M., Jiang, X., Xiong, L., Ren, W., Hirata, K., Yamamoto, M., Fan, S., & Yan, N. (2015). Molecular basis of ligand recognition and transport by glucose transporters. *Nature*, 526(7573), 391-396. <https://doi.org/10.1038/nature14655>
- Deng, D., Xu, C., Sun, P., Wu, J., Yan, C., Hu, M., & Yan, N. (2014). Crystal structure of the human glucose transporter GLUT1. *Nature*, 510(7503), 121-125. <https://doi.org/10.1038/nature13306>
- Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X., Murray, L. W., Arendall, W. B., Snoeyink, J., Richardson, J. S., & Richardson, D. C. (2007). MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Research*, 35(Web Server), W375-W383. <https://doi.org/10.1093/nar/gkm216>
- Do, H. N., Wang, J., Bhattarai, A., & Miao, Y. (2022). GLOW: A Workflow Integrating Gaussian-Accelerated Molecular Dynamics and Deep Learning for Free Energy Profiling. *Journal of Chemical Theory and Computation*, 18(3), 1423-1436. <https://doi.org/10.1021/acs.jctc.1c01055>
- Do, P.-C., Lee, E. H., & Le, L. (2018). Steered Molecular Dynamics Simulation in Rational Drug Design. *Journal of Chemical Information and Modeling*, 58(8), 1473-1482. <https://doi.org/10.1021/acs.jcim.8b00261>
- Doege, H., Schürmann, A., Ohnimus, H., Monser, V., Holman, G. D., & Joost, H. G. (1998). Serine-294 and threonine-295 in the exofacial loop domain between helices 7 and 8 of glucose transporters (GLUT) are involved in the conformational alterations during the transport process. *Biochemical Journal*, 329(Pt 2), 289-293. <https://doi.org/10.1042/bj3290289>

- Domene, C., Wiley, B., Gonzalez-Resines, S., & Naftalin, R. J. (2025). Insight into the Mechanism of d-Glucose Accelerated Exchange in GLUT1 from Molecular Dynamics Simulations. *Biochemistry*, 64(4), 928-939. <https://doi.org/10.1021/acs.biochem.4c00502>
- Drew, D., & Boudker, O. (2016). Shared Molecular Mechanisms of Membrane Transporters. *Annual Review of Biochemistry*, 85(Volume 85, 2016), 543-572. <https://doi.org/10.1146/annurev-biochem-060815-014520>
- Drew, D., North, R. A., Nagarathinam, K., & Tanabe, M. (2021). Structures and General Transport Mechanisms by the Major Facilitator Superfamily (MFS). *Chemical Reviews*, 121(9), 5289-5335. <https://doi.org/10.1021/acs.chemrev.0c00983>
- Dror, R. O., Dirks, R. M., Grossman, J. P., Xu, H., & Shaw, D. E. (2012). Biomolecular Simulation: A Computational Microscope for Molecular Biology. *Annual Review of Biophysics*, 41(Volume 41, 2012), 429-452. <https://doi.org/10.1146/annurev-biophys-042910-155245>
- Ebert, K., Ewers, M., Bisha, I., Sander, S., Rasputniac, T., Daniel, H., Antes, I., & Witt, H. (2018). Identification of essential amino acids for glucose transporter 5 (GLUT5)-mediated fructose transport. *Journal of Biological Chemistry*, 293(6), 2115-2124. <https://doi.org/10.1074/jbc.RA117.001442>
- Fincham, D. (1992). Leapfrog Rotational Algorithms. *Molecular Simulation*, 8(3-5), 165-178. <https://doi.org/10.1080/08927029208022474>
- Garay, P. G., Machado, M. R., Verli, H., & Pantano, S. (2024). SIRAH Late Harvest: Coarse-Grained Models for Protein Glycosylation. *Journal of Chemical Theory and Computation*, 20(2), 963-976. <https://doi.org/10.1021/acs.jctc.3c00783>
- González, M. A. (2011). Force fields and molecular dynamics simulations. *École Thématique de La Société Française de La Neutronique*, 12, 169-200. <https://doi.org/10.1051/sfn/201112009>
- Govind Kumar, V., Polasa, A., Agrawal, S., Kumar, T. K. S., & Moradi, M. (2023). Binding affinity estimation from restrained umbrella sampling simulations. *Nature Computational Science*, 3(1), 59-70. <https://doi.org/10.1038/s43588-022-00389-9>
- Groenendyk, J., Stoletov, K., Paskevicius, T., Li, W., Dai, N., Pujol, M., Busaan, E., Ng, H. H., Boukouris, A. E., Saleme, B., Haromy, A., Cui, K., Hu, M., Yan, Y., Zhang, R., Michelakis, E., Chen, X.-Z., Lewis, J. D., Tang, J., ... Michalak, M. (2022). Loss of the fructose transporter SLC2A5 inhibits cancer cell migration. *Frontiers in Cell and Developmental Biology*, 10. <https://doi.org/10.3389/fcell.2022.896297>
- Grønbech-Jensen, N., & Farago, O. (2013). A simple and effective Verlet-type algorithm for simulating Langevin dynamics. *Molecular Physics*, 111(8), 983-991. <https://doi.org/10.1080/00268976.2012.760055>
- Hadley, K. R., & McCabe, C. (2012). Coarse-grained molecular models of water: A review. *Molecular Simulation*, 38(8-9), 671-681. <https://doi.org/10.1080/08927022.2012.671942>

- Hao, P.-Y. (2009). Interval regression analysis using support vector networks. *Fuzzy Sets and Systems*, Theme: Learning, 160(17), 2466-2485. <https://doi.org/10.1016/j.fss.2008.10.012>
- Henderson, R. K., Fendler, K., & Poolman, B. (2019). Coupling efficiency of secondary active transporters. *Current Opinion in Biotechnology*, 58, 62-71. <https://doi.org/10.1016/j.copbio.2018.11.005>
- Henzler-Wildman, K., & Kern, D. (2007). Dynamic personalities of proteins. *Nature*, 450(7172), 964-972. <https://doi.org/10.1038/nature06522>
- Huang, Y., Lemieux, M. J., Song, J., Auer, M., & Wang, D.-N. (2003). Structure and Mechanism of the Glycerol-3-Phosphate Transporter from *Escherichia coli*. *Science*, 301(5633), 616-620. <https://doi.org/10.1126/science.1087619>
- Ibrahim, P., & Clark, T. (2019). Metadynamics simulations of ligand binding to GPCRs. *Current Opinion in Structural Biology, Macromolecular Assemblies. Theory and simulation: Demystifying GPCRs*, 55, 129-137. <https://doi.org/10.1016/j.sbi.2019.04.002>
- Iglesias-Fernandez, J., Quinn, P. J., Naftalin, R. J., & Domene, C. (2017). Membrane Phase-Dependent Occlusion of Intramolecular GLUT1 Cavities Demonstrated by Simulations. *Biophysical Journal*, 112(6), 1176-1184. <https://doi.org/10.1016/j.bpj.2017.01.030>
- Jardetzky, O. (1966). Simple Allosteric Model for Membrane Pumps. *Nature*, 211(5052), 969-970. <https://doi.org/10.1038/211969a0>
- Jensen, M. Ø., Yin, Y., Tajkhorshid, E., & Schulten, K. (2007). Sugar Transport across Lactose Permease Probed by Steered Molecular Dynamics. *Biophysical Journal*, 93(1), 92-102. <https://doi.org/10.1529/biophysj.107.103994>
- Jia, R., Martens, C., Shekhar, M., Pant, S., Pellowe, G. A., Lau, A. M., Findlay, H. E., Harris, N. J., Tajkhorshid, E., Booth, P. J., & Politis, A. (2020). Hydrogen-deuterium exchange mass spectrometry captures distinct dynamics upon substrate and inhibitor binding to a transporter. *Nature Communications*, 11, 6162. <https://doi.org/10.1038/s41467-020-20032-3>
- Jiang, Y., & Kalodimos, C. G. (2017). NMR Studies of Large Proteins. *Journal of Molecular Biology, John Kendrew's 100th Anniversary Special Edition*, 429(17), 2667-2676. <https://doi.org/10.1016/j.jmb.2017.07.007>
- Kaback, H. R., & Guan, L. (2019). It takes two to tango: The dance of the permease. *Journal of General Physiology*, 151(7), 878-886. <https://doi.org/10.1085/jgp.201912377>
- Kalyaanamoorthy, S., & Chen, Y.-P. P. (2014). Modelling and enhanced molecular dynamics to steer structure-based drug discovery. *Progress in Biophysics and Molecular Biology*, 114(3), 123-136. <https://doi.org/10.1016/j.pbiomolbio.2013.06.004>
- Kandt, C., Ash, W. L., & Peter Tieleman, D. (2007). Setting up and running molecular dynamics simulations of membrane proteins. *Methods*, 41(4), 475-488. <https://doi.org/10.1016/j.ymeth.2006.08.006>
- Karplus, M., & McCammon, J. A. (2002). Molecular dynamics simulations of biomolecules. *Nature Structural Biology*, 9(9), 646-652. <https://doi.org/10.1038/nsb0902-646>

- Kästner, J. (2011). Umbrella sampling. *WIREs Computational Molecular Science*, 1(6), 932-942. <https://doi.org/10.1002/wcms.66>
- Kazmier, K., Claxton, D. P., & Mchaourab, H. S. (2017). Alternating access mechanisms of LeuT-fold transporters: Trailblazing towards the promised energy landscapes. *Current Opinion in Structural Biology*, 45, 100-108. <https://doi.org/10.1016/j.sbi.2016.12.006>
- Ke, M., Yuan, Y., Jiang, X., Yan, N., & Gong, H. (2017). Molecular determinants for the thermodynamic and functional divergence of uniporter GLUT1 and proton symporter XyleE. *PLOS Computational Biology*, 13(6), e1005603. <https://doi.org/10.1371/journal.pcbi.1005603>
- Khakpour, A., Ghiabi, S., Babaheydari, A. K., Mirahmadi, S. A., Baziyar, P., Heidari-Soureshjani, E., & Horestani, M. K. (2025). Discovering the therapeutic potential of Naringenin in diabetes related to GLUT-4 and its regulatory factors: A computational approach. *Chemical Physics Impact*, 10, 100784. <https://doi.org/10.1016/j.chphi.2024.100784>
- Kimanius, D., Pettersson, I., Schluckebier, G., Lindahl, E., & Andersson, M. (2015). SAXS-Guided Metadynamics. *Journal of Chemical Theory and Computation*, 11(7), 3491-3498. <https://doi.org/10.1021/acs.jctc.5b00299>
- Klein, F., Soñora, M., Helene Santos, L., Nazareno Frigini, E., Ballesteros-Casallas, A., Rodrigo Machado, M., & Pantano, S. (2023). The SIRAH force field: A suite for simulations of complex biological systems at the coarse-grained and multiscale levels. *Journal of Structural Biology*, 215(3), 107985. <https://doi.org/10.1016/j.jsb.2023.107985>
- Kmiecik, S., Gront, D., Kolinski, M., Wieteska, L., Dawid, A. E., & Kolinski, A. (2016). Coarse-Grained Protein Models and Their Applications. *Chemical Reviews*, 116(14), 7898-7936. <https://doi.org/10.1021/acs.chemrev.6b00163>
- Koepsell, H. (2020). Glucose transporters in brain in health and disease. *Pflügers Archiv - European Journal of Physiology*, 472(9), 1299-1343. <https://doi.org/10.1007/s00424-020-02441-x>
- Kulshrestha, A., Punnathanam, S. N., & Ayappa, K. G. (2022). Finite temperature string method with umbrella sampling using path collective variables: Application to secondary structure change in a protein. *Soft Matter*, 18(39), 7593-7603. <https://doi.org/10.1039/D2SM00888B>
- Latorraca, N. R., Fastman, N. M., Venkatakrisnan, A. J., Frommer, W. B., Dror, R. O., & Feng, L. (2017). Mechanism of Substrate Translocation in an Alternating Access Transporter. *Cell*, 169(1), 96-107.e12. <https://doi.org/10.1016/j.cell.2017.03.010>
- Law, C. J., Maloney, P. C., & Wang, D.-N. (2008). Ins and outs of major facilitator superfamily antiporters. *Annual Review of Microbiology*, 62, 289-305. <https://doi.org/10.1146/annurev.micro.61.080706.093329>
- Leach, A. R. (1997). *Molecular Modelling: Principles and Applications*.
- Lehninger, A. L. (with Internet Archive). (2005). *Lehninger principles of biochemistry*. New York : W.H. Freeman. [http://archive.org/details/lehningerprincip0000lehn\\_04ed](http://archive.org/details/lehningerprincip0000lehn_04ed)

- Leturque, A., Brot-Laroche, E., & Le Gall, M. (2009). GLUT2 mutations, translocation, and receptor function in diet sugar managing. *American Journal of Physiology. Endocrinology and Metabolism*, 296(5), E985-992. <https://doi.org/10.1152/ajpendo.00004.2009>
- Liu, Z., Cao, X., Ma, Z., Xu, L., Wang, L., Li, J., Xiao, M., Jiang, X., Liu, Z., Cao, X., Ma, Z., Xu, L., Wang, L., Li, J., Xiao, M., & Jiang, X. (2024). Enhanced Sampling Molecular Dynamics Simulations Reveal Transport Mechanism of Glycoconjugate Drugs through GLUT1. *International Journal of Molecular Sciences*, 25(10). <https://doi.org/10.3390/ijms25105486>
- Loland, C. J. (2015). The use of LeuT as a model in elucidating binding sites for substrates and inhibitors in neurotransmitter transporters. *Biochimica et Biophysica Acta (BBA) - General Subjects, Structural biochemistry and biophysics of membrane proteins*, 1850(3), 500-510. <https://doi.org/10.1016/j.bbagen.2014.04.011>
- Machado, M. R., Barrera, E. E., Klein, F., Sónora, M., Silva, S., & Pantano, S. (2019). The SIRAH 2.0 Force Field: Altius, Fortius, Citius. *Journal of Chemical Theory and Computation*, 15(4), 2719-2733. <https://doi.org/10.1021/acs.jctc.9b00006>
- Machado, M. R., & Pantano, S. (2016). SIRAH tools: Mapping, backmapping and visualization of coarse-grained models. *Bioinformatics*, 32(10), 1568-1570. <https://doi.org/10.1093/bioinformatics/btw020>
- Maisuradze, G. G., Liwo, A., & Scheraga, H. A. (2009). Principal Component Analysis for Protein Folding Dynamics. *Journal of Molecular Biology*, 385(1), 312-329. <https://doi.org/10.1016/j.jmb.2008.10.018>
- Malmstrom, R. D., Lee, C. T., Van Wart, A. T., & Amaro, R. E. (2014). Application of Molecular-Dynamics Based Markov State Models to Functional Proteins. *Journal of Chemical Theory and Computation*, 10(7), 2648-2657. <https://doi.org/10.1021/ct5002363>
- Maragliano, L., Fischer, A., Vanden-Eijnden, E., & Ciccotti, G. (2006). String method in collective variables: Minimum free energy paths and isocommittor surfaces. *The Journal of Chemical Physics*, 125(2), 024106. <https://doi.org/10.1063/1.2212942>
- Marger, M. D., & Saier, M. H. (1993). A major superfamily of transmembrane facilitators that catalyse uniport, symport and antiport. *Trends in Biochemical Sciences*, 18(1), 13-20. [https://doi.org/10.1016/0968-0004\(93\)90081-W](https://doi.org/10.1016/0968-0004(93)90081-W)
- Martens, C., Shekhar, M., Borysik, A. J., Lau, A. M., Reading, E., Tajkhorshid, E., Booth, P. J., & Politis, A. (2018). Direct protein-lipid interactions shape the conformational landscape of secondary transporters. *Nature Communications*, 9(1), 4151. <https://doi.org/10.1038/s41467-018-06704-1>
- Marzinek, J. K., Bond, P. J., Lian, G., Zhao, Y., Han, L., Noro, M. G., Pistikopoulos, E. N., & Mantalaris, A. (2014). Free Energy Predictions of Ligand Binding to an  $\alpha$ -Helix Using Steered Molecular Dynamics and Umbrella Sampling Simulations. *Journal of Chemical Information and Modeling*, 54(7), 2093-2104. <https://doi.org/10.1021/ci500164q>

- McComas, S. E., Mitrovic, D., Alleva, C., Bonaccorsi, M., Drew, D., & Delemotte, L. (2022). Determinants of sugar-induced influx in the mammalian fructose transporter GLUT5 (p. 2022.06.17.495601). *bioRxiv*. <https://doi.org/10.1101/2022.06.17.495601>
- Meng, E. C., Goddard, T. D., Pettersen, E. F., Couch, G. S., Pearson, Z. J., Morris, J. H., & Ferrin, T. E. (2023). UCSF ChimeraX: Tools for structure building and analysis. *Protein Science*, 32(11), e4792. <https://doi.org/10.1002/pro.4792>
- Miao, Y., Feixas, F., Eun, C., & McCammon, J. A. (2015). Accelerated molecular dynamics simulations of protein folding. *Journal of Computational Chemistry*, 36(20), 1536-1549. <https://doi.org/10.1002/jcc.23964>
- Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G. A., Sonnhammer, E. L. L., Tosatto, S. C. E., Paladin, L., Raj, S., Richardson, L. J., Finn, R. D., & Bateman, A. (2021). Pfam: The protein families database in 2021. *Nucleic Acids Research*, 49(D1), D412-D419. <https://doi.org/10.1093/nar/gkaa913>
- Mitrovic, D., McComas, S. E., Alleva, C., Bonaccorsi, M., Drew, D., & Delemotte, L. (2023). Reconstructing the transport cycle in the sugar porter superfamily using coevolution-powered machine learning. *eLife*, 12, e84805. <https://doi.org/10.7554/eLife.84805>
- Moradi, M., Enkavi, G., & Tajkhorshid, E. (2015). Atomic-level characterization of transport cycle thermodynamics in the glycerol-3-phosphate:phosphate antiporter. *Nature Communications*, 6(1), 8393. <https://doi.org/10.1038/ncomms9393>
- Motahari, H., Ghanbarzadeh Noudehi, M., Jamali, T., & Jafari, G. R. (2025). Chapman-Kolmogorov test for estimating memory length of two coupled processes. *Scientific Reports*, 15(1), 11603. <https://doi.org/10.1038/s41598-025-92238-8>
- Mueckler, M., & Thorens, B. (2013). The SLC2 (GLUT) family of membrane transporters. *Molecular Aspects of Medicine, The ABCs of membrane transporters in health and disease (SLC series)*, 34(2), 121-138. <https://doi.org/10.1016/j.mam.2012.07.001>
- Naftalin, R. J. (2018). A critique of the alternating access transporter model of uniport glucose transport. *Biophysics Reports*, 4(6), 287-299. <https://doi.org/10.1007/s41048-018-0076-9>
- Nagel, D., & Stock, G. (2023). msmhelper: A Python package for Markov state modeling of protein dynamics. *Journal of Open Source Software*, 8(85), 5339. <https://doi.org/10.21105/joss.05339>
- Nanni, L., Lumini, A., & Brahmam, S. (2014). An Empirical Study of Different Approaches for Protein Classification. *The Scientific World Journal*, 2014(1), 236717. <https://doi.org/10.1155/2014/236717>
- Nomura, N., Verdon, G., Kang, H. J., Shimamura, T., Nomura, Y., Sonoda, Y., Hussien, S. A., Qureshi, A. A., Coincon, M., Sato, Y., Abe, H., Nakada-Nakura, Y., Hino, T., Arakawa, T., Kusano-Arai, O., Iwanari, H., Murata, T., Kobayashi, T., Hamakubo, T., ... Drew, D. (2015a). Structure and mechanism of the mammalian fructose transporter GLUT5. *Nature*, 526(7573), 397-401. <https://doi.org/10.1038/nature14909>

Nomura, N., Verdon, G., Kang, H. J., Shimamura, T., Nomura, Y., Sonoda, Y., Hussien, S. A., Qureshi, A. A., Coincon, M., Sato, Y., Abe, H., Nakada-Nakura, Y., Hino, T., Arakawa, T., Kusano-Arai, O., Iwanari, H., Murata, T., Kobayashi, T., Hamakubo, T., ... Drew, D. (2015b). Structure and mechanism of the mammalian fructose transporter GLUT5. *Nature*, 526(7573), 397-401. <https://doi.org/10.1038/nature14909>

Okamoto, K., & Sako, Y. (2017). Recent advances in FRET for the study of protein interactions and dynamics. *Current Opinion in Structural Biology, Cryo electron microscopy: exciting advances in CryoEM herald a new era in structural biology • Biophysical methods: behind the scenes of the cryo-EM revolution*, 46, 16-23. <https://doi.org/10.1016/j.sbi.2017.03.010>

Olsson, M. H. M., Søndergaard, C. R., Rostkowski, M., & Jensen, J. H. (2011). PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions. *Journal of Chemical Theory and Computation*, 7(2), 525-537. <https://doi.org/10.1021/ct100578z>

Orellana, L. (2019). Large-Scale Conformational Changes and Protein Function: Breaking the in silico Barrier. *Frontiers in Molecular Biosciences*, 6. <https://doi.org/10.3389/fmolb.2019.00117>

Ourmazd, A. (2019). Cryo-EM, XFELs and the structure conundrum in structural biology. *Nature Methods*, 16(10), 941-944. <https://doi.org/10.1038/s41592-019-0587-4>

Park, M.-S. (2015). Molecular Dynamics Simulations of the Human Glucose Transporter GLUT1. *PLOS ONE*, 10(4), e0125361. <https://doi.org/10.1371/journal.pone.0125361>

Patodia, S. (2014). Molecular Dynamics Simulation of Proteins: A Brief Overview. *Journal of Physical Chemistry & Biophysics*, 4(6). <https://doi.org/10.4172/2161-0398.1000166>

Paulsen, P. A., Custódio, T. F., & Pedersen, B. P. (2019). Crystal structure of the plant symporter STP10 illuminates sugar uptake mechanism in monosaccharide transporter superfamily. *Nature Communications*, 10(1). <https://doi.org/10.1038/s41467-018-08176-9>

Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kalé, L., & Schulten, K. (2005). Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, 26(16), 1781-1802. <https://doi.org/10.1002/jcc.20289>

Pulawski, W., Jamroz, M., Kolinski, M., Kolinski, A., & Kmiecik, S. (2016). Coarse-Grained Simulations of Membrane Insertion and Folding of Small Helical Proteins Using the CABS Model. *Journal of Chemical Information and Modeling*, 56(11), 2207-2215. <https://doi.org/10.1021/acs.jcim.6b00350>

Quistgaard, E. M., Löw, C., Moberg, P., Trésaugues, L., & Nordlund, P. (2013). Structural basis for substrate transport in the GLUT-homology family of monosaccharide transporters. *Nature Structural & Molecular Biology*, 20(6), 766-768. <https://doi.org/10.1038/nsmb.2569>

Qureshi, A. A., Suades, A., Matsuoka, R., Brock, J., McComas, S. E., Nji, E., Orellana, L., Claesson, M., Delemotte, L., & Drew, D. (2020). The molecular basis for sugar import in malaria parasites. *Nature*, 578(7794), 321-325. <https://doi.org/10.1038/s41586-020-1963-z>

- Radestock, S., & Forrest, L. R. (2011). The Alternating-Access Mechanism of MFS Transporters Arises from Inverted-Topology Repeats. *Journal of Molecular Biology*, 407(5), 698-715. <https://doi.org/10.1016/j.jmb.2011.02.008>
- Rana, N., Aziz, M. A., Oraby, A. K., Wuest, M., Dufour, J., Abouzid, K. A. M., Wuest, F., West, F. G., Rana, N., Aziz, M. A., Oraby, A. K., Wuest, M., Dufour, J., Abouzid, K. A. M., Wuest, F., & West, F. G. (2022). Towards Selective Binding to the GLUT5 Transporter: Synthesis, Molecular Dynamics and In Vitro Evaluation of Novel C-3-Modified 2,5-Anhydro-D-mannitol Analogs. *Pharmaceutics*, 14(4). <https://doi.org/10.3390/pharmaceutics14040828>
- Ray, D., & Parrinello, M. (2023). Kinetics from Metadynamics: Principles, Applications, and Outlook. *Journal of Chemical Theory and Computation*, 19(17), 5649-5670. <https://doi.org/10.1021/acs.jctc.3c00660>
- Riniker, S., Allison, J. R., & Gunsteren, W. F. van. (2012). On developing coarse-grained models for biomolecular simulation: A review. *Physical Chemistry Chemical Physics*, 14(36), 12423-12430. <https://doi.org/10.1039/C2CP40934H>
- Roe, D. R., & Cheatham, T. E. I. (2013). PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *Journal of Chemical Theory and Computation*, 9(7), 3084-3095. <https://doi.org/10.1021/ct400341p>
- Romo, T. D., & Grossfield, A. (2011). Block Covariance Overlap Method and Convergence in Molecular Dynamics Simulation. *Journal of Chemical Theory and Computation*, 7(8), 2464-2472. <https://doi.org/10.1021/ct2002754>
- Roux, B. (2021). String Method with Swarms-of-Trajectories, Mean Drifts, Lag Time, and Commitor. *The Journal of Physical Chemistry A*, 125(34), 7558-7571. <https://doi.org/10.1021/acs.jpca.1c04110>
- Suades, A., Qureshi, A., McComas, S. E., Coinçon, M., Rudling, A., Chatzikyriakidou, Y., Landreh, M., Carlsson, J., & Drew, D. (2023). Establishing mammalian GLUT kinetics and lipid composition influences in a reconstituted-liposome system. *Nature Communications*, 14, 4070. <https://doi.org/10.1038/s41467-023-39711-y>
- Schlitter, J., Engels, M., & Krüger, P. (1994). Targeted molecular dynamics: A new approach for searching pathways of conformational transitions. *Journal of Molecular Graphics*, 12(2), 84-89. [https://doi.org/10.1016/0263-7855\(94\)80072-3](https://doi.org/10.1016/0263-7855(94)80072-3)
- Schott-Verdugo, S., & Gohlke, H. (2019). PACKMOL-Memgen: A Simple-To-Use, Generalized Workflow for Membrane-Protein-Lipid-Bilayer System Building. *Journal of Chemical Information and Modeling*, 59(6), 2522-2528. <https://doi.org/10.1021/acs.jcim.9b00269>
- Screpanti, E., & Hunte, C. (2007). Discontinuous membrane helices in transport proteins and their correlation with function. *Journal of Structural Biology*, 3rd International Conference on Structure, Dynamics and Function of Proteins in Biological Membranes, 159(2), 261-267. <https://doi.org/10.1016/j.jsb.2007.01.011>

Seal, R. L., Braschi, B., Gray, K., Jones, T. E. M., Tweedie, S., Haim-Vilmovsky, L., & Bruford, E. A. (2023). Genenames.org: The HGNC resources in 2023. *Nucleic Acids Research*, 51(D1), D1003-D1009. <https://doi.org/10.1093/nar/gkac888>

Shaw, D. E., Dror, R. O., Salmon, J. K., Grossman, J. P., Mackenzie, K. M., Bank, J. A., Young, C., Deneroff, M. M., Batson, B., Bowers, K. J., Chow, E., Eastwood, M. P., Ierardi, D. J., Klepeis, J. L., Kuskin, J. S., Larson, R. H., Lindorff-Larsen, K., Maragakis, P., Moraes, M. A., ... Towles, B. (2009). Millisecond-scale molecular dynamics simulations on Anton. *Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis, SC '09*, 1-11. <https://doi.org/10.1145/1654059.1654126>

Shen, W., Zhou, T., & Shi, X. (2023). Enhanced sampling in molecular dynamics simulations and their latest applications—A review. *Nano Research*, 16(12), 13474-13497. <https://doi.org/10.1007/s12274-023-6311-9>

Sittel, F., & Stock, G. (2016). Robust Density-Based Clustering To Identify Metastable Conformational States of Proteins. *Journal of Chemical Theory and Computation*, 12(5), 2426-2435. <https://doi.org/10.1021/acs.jctc.5b01233>

Sonoda, Y., Newstead, S., Hu, N.-J., Alguel, Y., Nji, E., Beis, K., Yashiro, S., Lee, C., Leung, J., Cameron, A. D., Byrne, B., Iwata, S., & Drew, D. (2011). Benchmarking Membrane Protein Detergent Stability for Improving Throughput of High-Resolution X-ray Structures. *Structure*, 19(1), 17-25. <https://doi.org/10.1016/j.str.2010.12.001>

Souza, P. C. T., Alessandri, R., Barnoud, J., Thallmair, S., Faustino, I., Grünwald, F., Patmanidis, I., Abdizadeh, H., Bruininks, B. M. H., Wassenaar, T. A., Kroon, P. C., Melcer, J., Nieto, V., Corradi, V., Khan, H. M., Domański, J., Javanainen, M., Martinez-Seara, H., Reuter, N., ... Marrink, S. J. (2021). Martini 3: A general purpose force field for coarse-grained molecular dynamics. *Nature Methods*, 18(4), 382-388. <https://doi.org/10.1038/s41592-021-01098-3>

Sterpone, F., Derreumaux, P., & Melchionna, S. (2015). Protein Simulations in Fluids: Coupling the OPEP Coarse-Grained Force Field with Hydrodynamics. *Journal of Chemical Theory and Computation*, 11(4), 1843-1853. <https://doi.org/10.1021/ct501015h>

Sun, L., Zeng, X., Yan, C., Sun, X., Gong, X., Rao, Y., & Yan, N. (2012). Crystal structure of a bacterial homologue of glucose transporters GLUT1–4. *Nature*, 490(7420), 361-366. <https://doi.org/10.1038/nature11524>

Sun, Z., Kalhor, P., Xu, Y., & Liu, J. (2021). Extensive numerical tests of leapfrog integrator in middle thermostat scheme in molecular simulations †. *Chinese Journal of Chemical Physics*, 34(6), 932-948. <https://doi.org/10.1063/1674-0068/cjcp2111242>

Sweet, J. C., Nowling, R. J., Cickovski, T., Sweet, C. R., Pande, V. S., & Izaguirre, J. A. (2013). Long Timestep Molecular Dynamics on the Graphical Processing Unit. *Journal of Chemical Theory and Computation*, 9(8), 3267-3281. <https://doi.org/10.1021/ct400331r>

Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., & Berendsen, H. J. C. (2005). GROMACS: Fast, flexible, and free. *Journal of Computational Chemistry*, 26(16), 1701-1718. <https://doi.org/10.1002/jcc.20291>

- Vehtari, A., Gelman, A., Simpson, D., Carpenter, B., & Bürkner, P.-C. (2021). Rank-Normalization, Folding, and Localization: An Improved  $\hat{R}$  for Assessing Convergence of MCMC (with Discussion). *Bayesian Analysis*, 16(2), 667-718. <https://doi.org/10.1214/20-BA1221>
- Wang, S. C., Davejan, P., Hendargo, K. J., Javadi-Razaz, I., Chou, A., Yee, D. C., Ghazi, F., Lam, K. J. K., Conn, A. M., Madrigal, A., Medrano-Soto, A., & Saier, M. H. (2020). Expansion of the Major Facilitator Superfamily (MFS) to include novel transporters as well as transmembrane-acting enzymes. *Biochimica Et Biophysica Acta. Biomembranes*, 1862(9), 183277. <https://doi.org/10.1016/j.bbamem.2020.183277>
- Webb, B., & Sali, A. (2016). Comparative Protein Structure Modeling Using MODELLER. *Current Protocols in Bioinformatics*, 54, 5.6.1-5.6.37. <https://doi.org/10.1002/cpbi.3>
- Winkelmann, I., Uzdavinyis, P., Kenney, I. M., Brock, J., Meier, P. F., Wagner, L.-M., Gabriel, F., Jung, S., Matsuoka, R., von Ballmoos, C., Beckstein, O., & Drew, D. (2022). Crystal structure of the Na<sup>+</sup>/H<sup>+</sup> antiporter NhaA at active pH reveals the mechanistic basis for pH sensing. *Nature Communications*, 13(1), 6383. <https://doi.org/10.1038/s41467-022-34120-z>
- Wisedchaisri, G., Park, M.-S., Iadanza, M. G., Zheng, H., & Gonen, T. (2014). Proton-coupled sugar transport in the prototypical major facilitator superfamily protein XyleE. *Nature Communications*, 5(1), 4521. <https://doi.org/10.1038/ncomms5521>
- Wright, P. E., & Dyson, H. J. (2015). Intrinsically disordered proteins in cellular signalling and regulation. *Nature Reviews Molecular Cell Biology*, 16(1), 18-29. <https://doi.org/10.1038/nrm3920>
- Xiao, Q., Sun, B., Zhou, Y., Wang, C., Guo, L., He, J., & Deng, D. (2021). Visualizing the nonlinear changes of a drug-proton antiporter from inward-open to occluded state. *Biochemical and Biophysical Research Communications*, 534, 272-278. <https://doi.org/10.1016/j.bbrc.2020.11.096>
- Xu, M., Jiang, J., Gao, L., Alyemni, S. O., & Haider, S. (2025). Structural Insights into the Substrate Egress Pathways Explains Specificity and Inhibition of Human Glucose Transporters (GLUT1 and GLUT9). *ACS Pharmacology & Translational Science*, 8(6), 1778-1790. <https://doi.org/10.1021/acsptsci.5c00209>
- Yang, Y. I., Shao, Q., Zhang, J., Yang, L., & Gao, Y. Q. (2019). Enhanced sampling in molecular dynamics. *The Journal of Chemical Physics*, 151(7), 070902. <https://doi.org/10.1063/1.5109531>
- Zhang, Y., & Skolnick, J. (2005). TM-align: A protein structure alignment algorithm based on the TM-score. *Nucleic Acids Research*, 33(7), 2302-2309. <https://doi.org/10.1093/nar/gki524>
- Zhang, Y., Zhang, Y., Sun, K., Meng, Z., & Chen, L. (2019). The SLC transporter in nutrient and metabolic sensing, regulation, and drug development. *Journal of Molecular Cell Biology*, 11(1), 1-13. <https://doi.org/10.1093/jmcb/mjy052>

## ANEXOS

### Anexo 1. Reporte estructural MolProbity para conformación exofacial

	Obtenido (%)	Ideal (%)
Poor rotamers	1.51	< 0.3
Favored rotamers	92.95	<98
Ramachandran outliers	0.43	<0.05
Ramachandran favored	97.84	>98
C $\beta$ deviations >0.25Å	0.00	0
Bad bonds	0.05	0
Bad angles	1.14	<0.1
Cis Prolines	0.00	<5
CaBLAM outliers	0.60	<1.0
CA Geometry outliers	0.00	<0.5

### Anexo 2. Reporte estructural MolProbity para conformación exofacial ocluida

	Obtenido (%)	Ideal (%)
Poor rotamers	1.76	< 0.3
Favored rotamers	91.18	<98
Ramachandran outliers	0.86	<0.05
Ramachandran favored	97.62	>98
C $\beta$ deviations >0.25Å	2.06	0
Bad bonds	0.08	0
Bad angles	1.59	<0.1
Cis Prolines	0.00	<5
CaBLAM outliers	1.5	<1.0
CA Geometry outliers	0.64	<0.5

### Anexo 3. Reporte estructural MolProbity para conformación ocluida

	Obtenido (%)	Ideal (%)
Poor rotamers	4.53	< 0.3
Favored rotamers	87.91	<98
Ramachandran outliers	1.51	<0.05
Ramachandran favored	92.44	>98
C $\beta$ deviations >0.25Å	3.67	0
Bad bonds	0.00	0
Bad angles	1.49	<0.1
Cis Prolines	0.00	<5
CaBLAM outliers	3.6	<1.0
CA Geometry outliers	1.06	<0.5

### Anexo 4. Reporte estructural MolProbity para conformación endofacial ocluida

	Obtenido (%)	Ideal (%)
Poor rotamers	2.54	< 0.3
Favored rotamers	87.31	<98
Ramachandran outliers	1.54	<0.05
Ramachandran favored	93.83	>98
C $\beta$ deviations >0.25Å	2.06	0
Bad bonds	0.00	0
Bad angles	1.33	<0.1
Cis Prolines	0.00	<5
CaBLAM outliers	3.2	<1.0
CA Geometry outliers	1.91	<0.5

### Anexo 5. Reporte estructural MolProbity para conformación endofacial

	Obtenido (%)	Ideal (%)
Poor rotamers	4.04	< 0.3
Favored rotamers	90.40	<98
Ramachandran outliers	0.87	<0.05
Ramachandran favored	95.01	>98
C $\beta$ deviations >0.25Å	0.23	0
Bad bonds	0.03	0
Bad angles	1.36	<0.1
Cis Prolines	0.00	<5

CaBLAM outliers	3.0	<1.0
CA Geometry outliers	1.70	<0.5

### Anexo 6. Estadísticos descriptivos de las aperturas intracelular (IC) y extracelular (EC) por conformación DM-GG 3 $\mu$ s

EC= Distancia Extracelular. IC = Distancia Intracelular. \*representa la media y su desviación estándar calculadas por el método del bloque

Conformación	*promedio EC (Å)	EC rango (Å)	*promedio IC (Å)	IC rango (Å)
Exofacial	18.70 $\pm$ 0.51	13.89–24.88	12.99 $\pm$ 0.07	11.27–17.00
Exofacial Ocluida	14.89 $\pm$ 0.12	11.75–17.35	14.17 $\pm$ 0.32	10.16–16.39
Ocluida	10.98 $\pm$ 0.14	8.65–13.38	11.01 $\pm$ 0.17	7.82–12.78
Endofacial Ocluida	11.36 $\pm$ 0.19	9.18–14.24	13.60 $\pm$ 0.27	10.93–16.27
Endofacial	11.15 $\pm$ 0.22	8.69–14.38	11.97 $\pm$ 0.23	9.59–15.77

### Anexo 7. Comparación de aperturas entre promediados de simulaciones (3 $\mu$ ) y sus respectivas plantillas estructurales (controles).

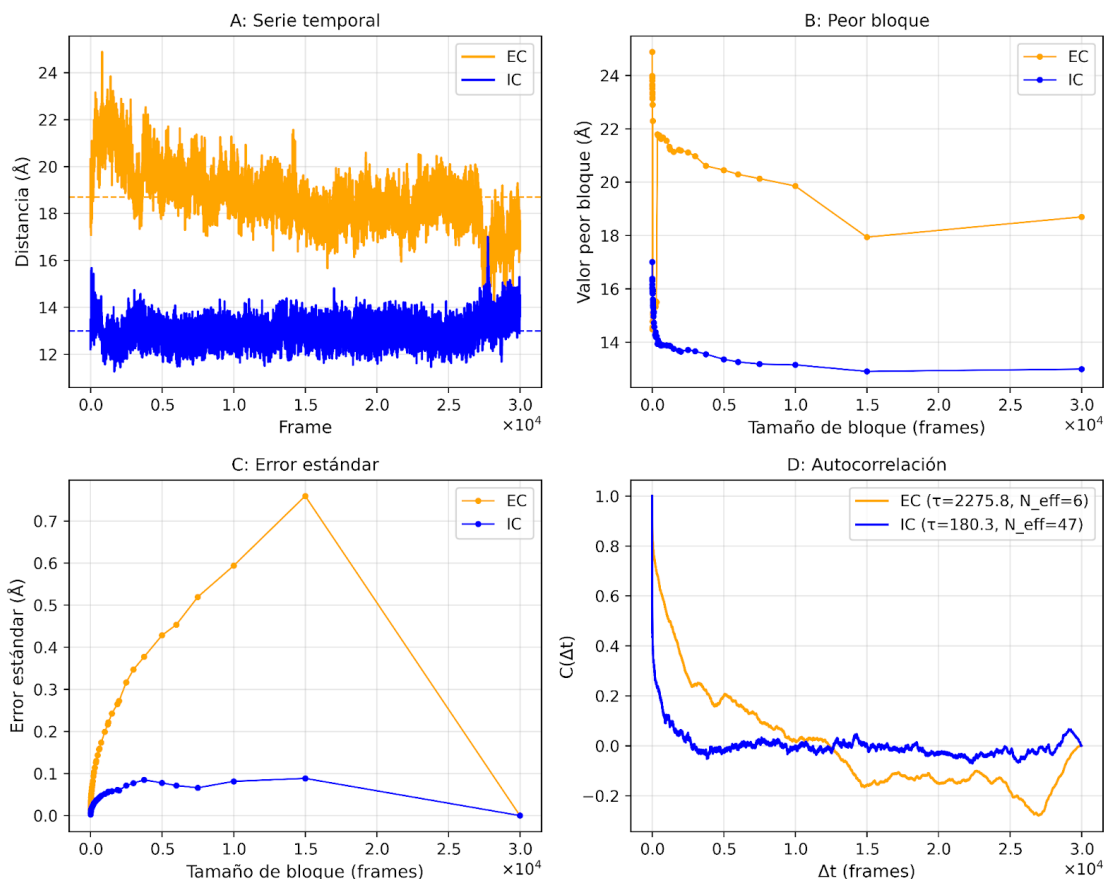
\*Corresponden a las medias de los controles  $\Delta$ EC = Diferencia entre la media de la simulación y controles

Conformación	EC (Å)*	EC sim (Å)	$\Delta$ EC (Å)	IC (Å)*	IC sim (Å)	$\Delta$ IC (Å)
Exofacial	20.30	18.70	-1.60	10.93	12.99	+2.06
Exofacial Ocluida	12.34	14.89	+2.55	10.48	14.17	+3.69
Ocluida	9.78	10.98	+1.20	10.98	11.01	~0
Endofacial Ocluida	9.31	11.36	+2.05	12.68	13.60	+0.92
Endofacial	9.95	11.15	+1.20	12.51	11.97	-0.54

## Anexo 8. Comparación de rangos de aperturas entre simulaciones 3 $\mu$ s y 5 $\mu$ s

\*Corresponden a simulaciones 3 $\mu$ s. \*\*Corresponden a simulaciones 5 $\mu$ s

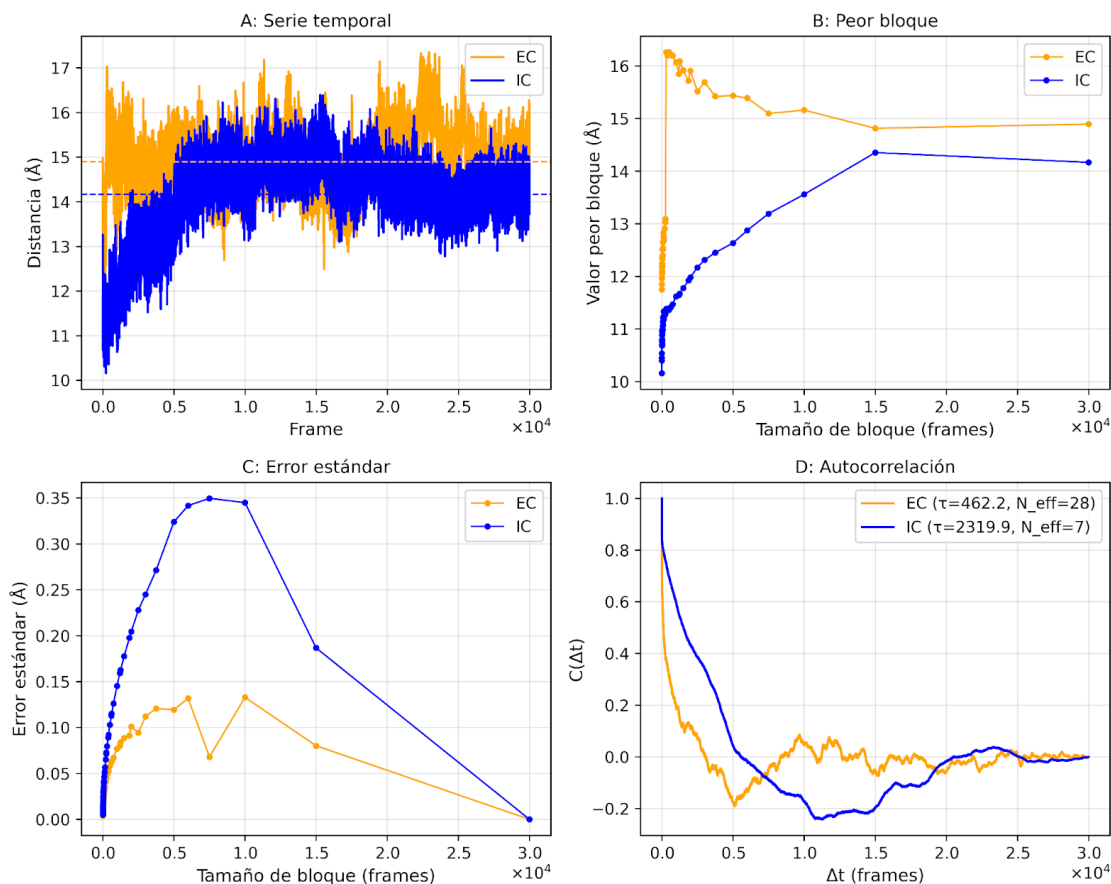
Conformación	EC rango (Å)*	IC rango (Å)*	EC rango (Å)**	IC rango (Å)**
Exofacial	13.89–24.88	11.27–17.00	14.83–23.95	11.02–15.97
Ocluida	8.65–13.38	7.82–12.78	8.10–13.94	9.61–14.28
Endofacial	8.69–14.38	9.59–15.77	8.40–12.42	10.83–15.09



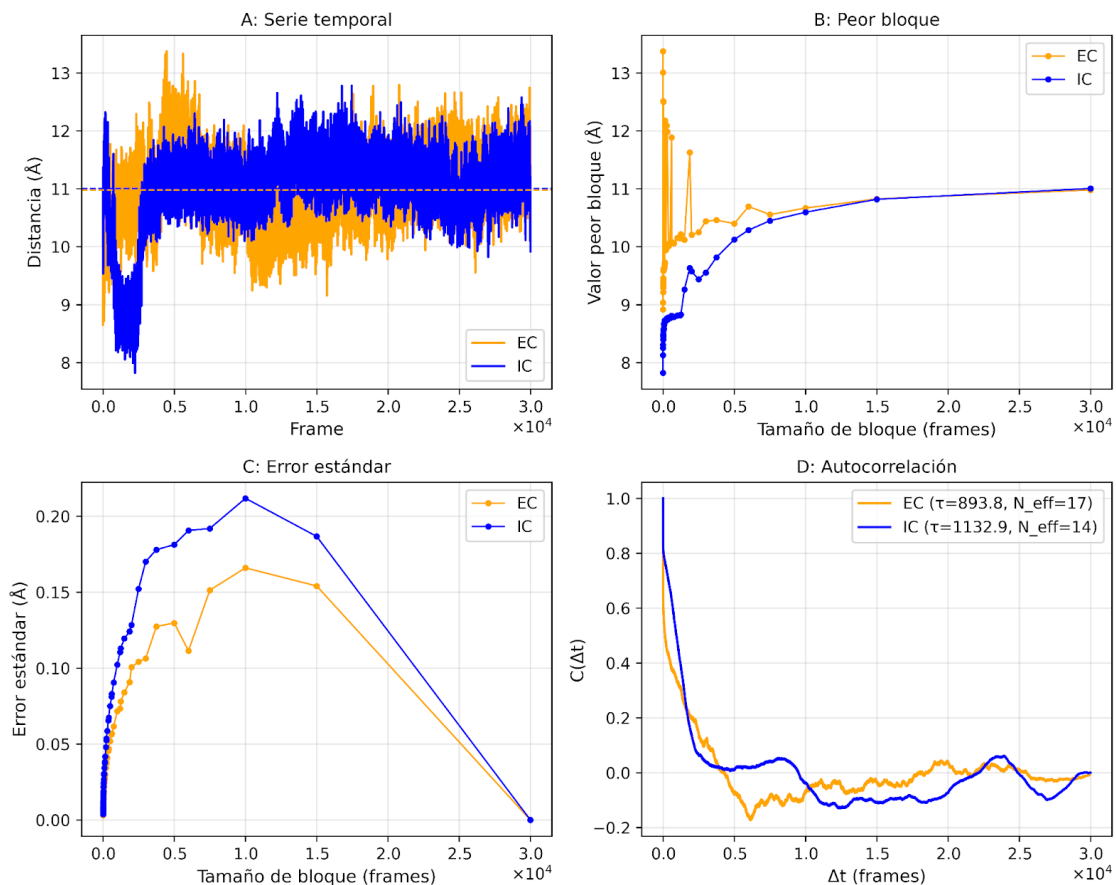
## Anexo 9. Análisis de convergencia en conformación exofacial. Se muestran los resultados del análisis de *block averaging* y autocorrelación para las distancias exofacial (EC, naranja) e intracelular (IC, azul).

(A) Serie temporal de las distancias EC e IC. Las líneas punteadas representan al valor de cada variable del cristal (B) Valor del peor bloque para distintos tamaños de bloque, mostrando la máxima desviación respecto al promedio. (C) Error estándar de la media calculado como función del tamaño de bloque. (D) Función de

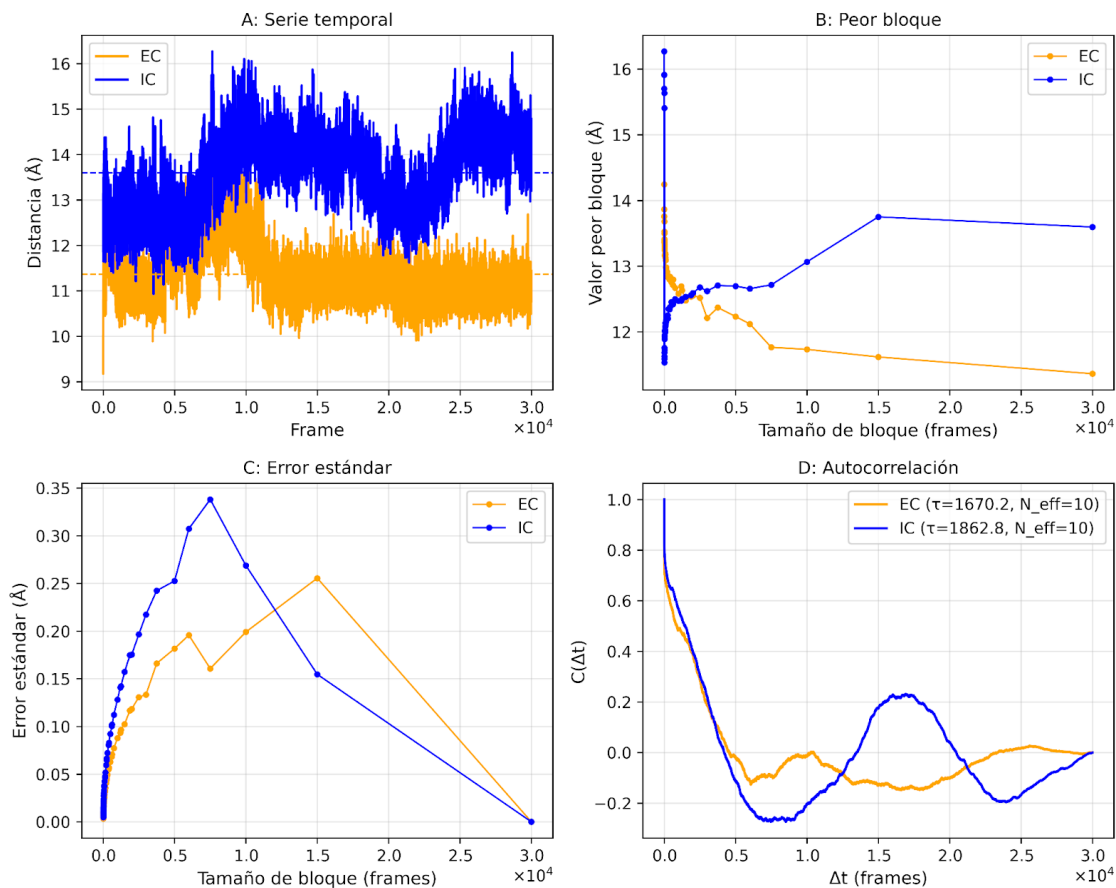
autocorrelación para EC e IC. La leyenda incluye el tiempo de autocorrelación estimado ( $\tau$ ) y el número efectivo de observaciones  $N_{\text{eff}}$ .



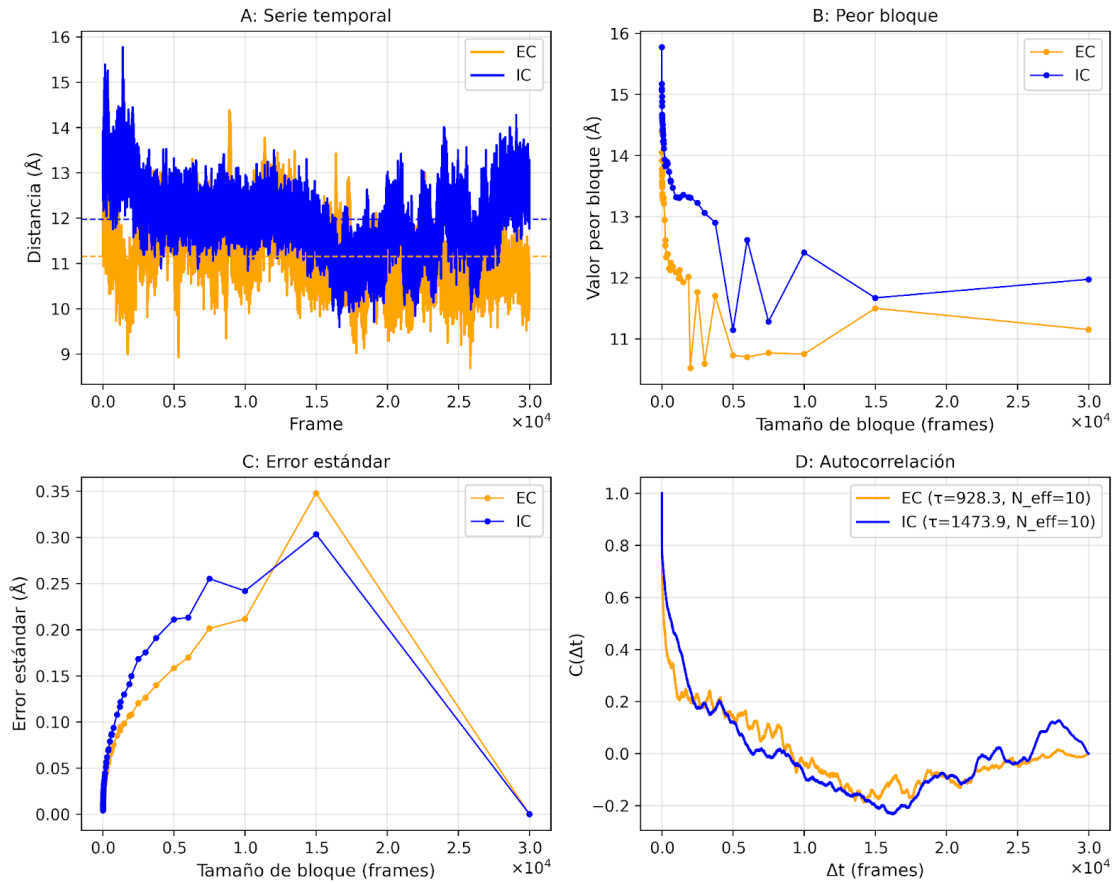
**Anexo 10. Análisis de convergencia en conformación exofacial ocluida. Se muestran los resultados del análisis de *block averaging* y autocorrelación para las distancias exofacial (EC, naranja) e intracelular (IC, azul). (A) Serie temporal de las distancias EC e IC. Las líneas punteadas representan al valor de cada variable del cristal (B) Valor del peor bloque para distintos tamaños de bloque, mostrando la máxima desviación respecto al promedio. (C) Error estándar de la media calculado como función del tamaño de bloque. (D) Función de autocorrelación para EC e IC. La leyenda incluye el tiempo de autocorrelación estimado ( $\tau$ ) y el número efectivo de observaciones  $N_{\text{eff}}$ .**



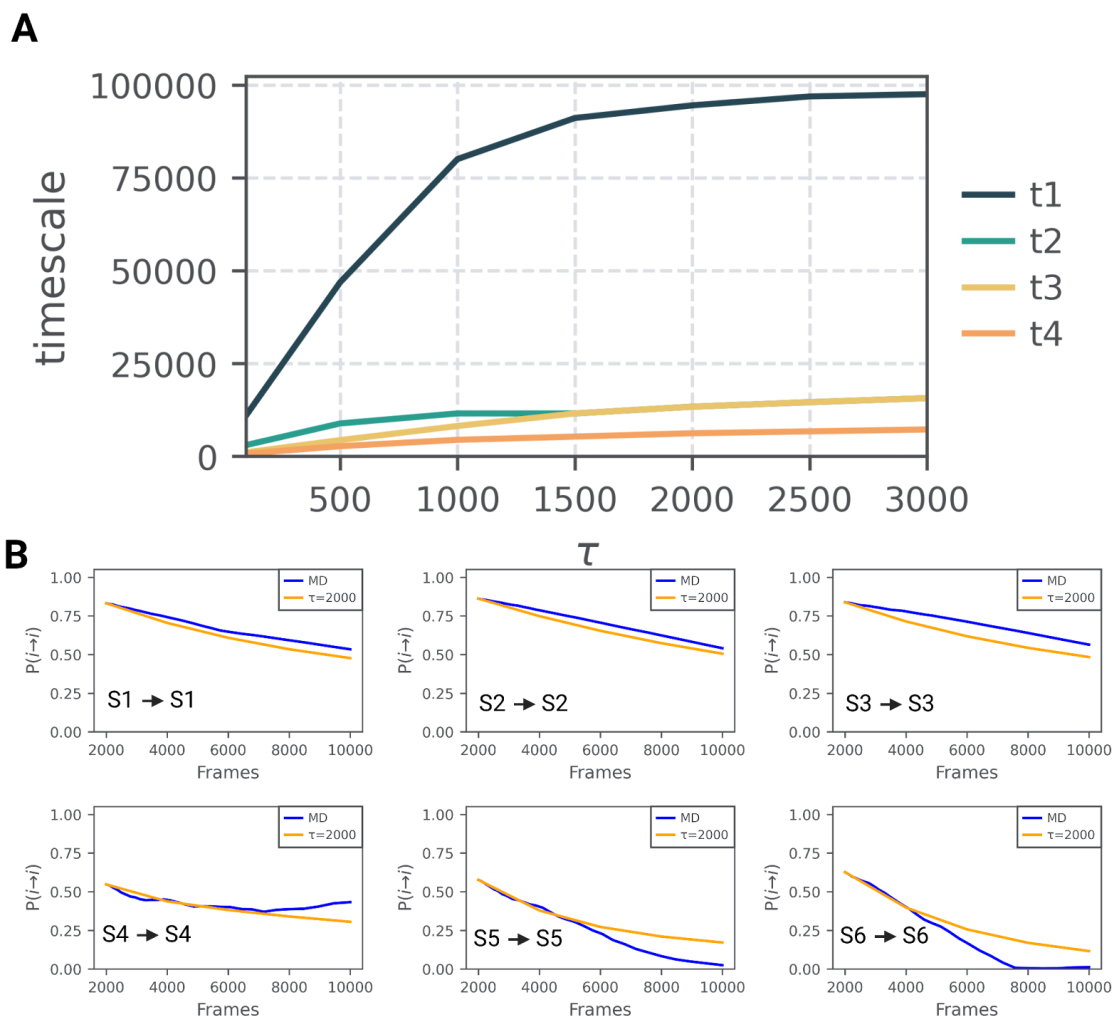
**Anexo 11. Análisis de convergencia en conformación ocluida. Se muestran los resultados del análisis de *block averaging* y autocorrelación para las distancias exofacial (EC, naranja) e intracelular (IC, azul).** (A) Serie temporal de las distancias EC e IC. Las líneas punteadas representan al valor de cada variable del cristal (B) Valor del peor bloque para distintos tamaños de bloque, mostrando la máxima desviación respecto al promedio. (C) Error estándar de la media calculado como función del tamaño de bloque. (D) Función de autocorrelación para EC e IC. La leyenda incluye el tiempo de autocorrelación estimado ( $\tau$ ) y el número efectivo de observaciones  $N_{eff}$ .



**Anexo 12. Análisis de convergencia en conformación endofacial ocluida. Se muestran los resultados del análisis de *block averaging* y autocorrelación para las distancias exofacial (EC, naranja) e intracelular (IC, azul). (A) Serie temporal de las distancias EC e IC. Las líneas punteadas representan al valor de cada variable del cristal (B) Valor del peor bloque para distintos tamaños de bloque, mostrando la máxima desviación respecto al promedio. (C) Error estándar de la media calculado como función del tamaño de bloque. (D) Función de autocorrelación para EC e IC. La leyenda incluye el tiempo de autocorrelación estimado ( $\tau$ ) y el número efectivo de observaciones  $N_{eff}$ .**



**Anexo 13. Análisis de convergencia en conformación endofacial. Se muestran los resultados del análisis de *block averaging* y autocorrelación para las distancias exofacial (EC, naranja) e intracelular (IC, azul). (A) Serie temporal de las distancias EC e IC. Las líneas punteadas representan al valor de cada variable del cristal (B) Valor del peor bloque para distintos tamaños de bloque, mostrando la máxima desviación respecto al promedio. (C) Error estándar de la media calculado como función del tamaño de bloque. (D) Función de autocorrelación para EC e IC. La leyenda incluye el tiempo de autocorrelación estimado ( $\tau$ ) y el número efectivo de observaciones  $N_{eff}$ .**



**Anexo 14. Validación del modelo de estados de Markov (MSM).** (A) Escalas de tiempo implícitas (*implied timescales*) calculadas en función del tiempo de retardo  $\tau$ . La convergencia progresiva de las escalas de tiempo más lentas al aumentar  $\tau$  indica la aproximación al régimen markoviano y permite identificar un rango adecuado de tiempos de retardo para la construcción del MSM. (B) Prueba de Chapman–Kolmogorov para los seis estados metastables identificados (S1–S6). Se muestra la comparación entre las probabilidades de permanencia obtenidas directamente de las trayectorias de dinámica molecular (MD, azul) y las predichas por el modelo de Markov construido con  $\tau = 1500$ frames (naranja). La concordancia entre ambas curvas respalda la consistencia interna del modelo y su capacidad para reproducir la dinámica lenta del sistema en la escala temporal considerada.