

# **ESTUDIO DE LA INTERACCIÓN DEL SEGMENTO LLP3 DE LA GLICOPROTEÍNA GP41 DEL VIH CON BIOMEMBRANAS MEDIANTE DINÁMICA MOLECULAR**



**MARINA LODEIRO ARANDA**

**Tutores:** Prof. José Villalaín Boullón / Dr. Emmanuel Fajardo Sánchez

**Departamento:** Bioquímica

**Área:** Bioquímica y Biología Molecular

**Titulación:** Biotecnología, Facultad de Ciencias Experimentales

**Curso académico:** 2015-2016

# Índice:

Resumen. Palabras clave

*Abstract. Keywords*

Listado de abreviaturas

## 1. Introducción

1.1. Los virus: introducción general

1.2. Las proteínas de fusión. Clasificación y características

1.3. Fusión de membranas mediante proteínas de clase I desde el punto de vista proteico y lipídico

1.3.1. Cambios en la glicoproteína

1.3.2. Cambios en la membrana

1.4. SIDA. Insuficiencia de las terapias convencionales para su tratamiento

1.5. VIH: taxonomía, estructura y genoma

1.6. Proteínas de fusión del VIH: gp120 y gp41

1.6.1. Función e interacciones de gp120

1.6.2. Estructura y conformaciones de gp41

1.7. Otras terapias antivirales: inhibidores de fusión de primera, segunda y tercera generación

1.8. Introducción a la simulación de sistemas biológicos por ordenador

1.9. Justificación del presente trabajo

## 2. Antecedentes y objetivos

## 3. Materiales y métodos

3.1. Caracterización del sistema

3.2. Simulación de dinámica molecular

3.3. Modelo de membrana

3.4. Análisis

## 4. Resultados y discusión

## 5. Conclusiones y proyección futura

## 6. Referencias

## Resumen

La proteína de envoltura gp41 del VIH es responsable de la fusión de membranas entre el virus y la célula diana. El endodominio de gp41 contiene segmentos altamente helicoidales e hidrofóbicos, los LLP, de entre los cuales se ha destacado LLP3 con anterioridad. Con el fin de obtener nuevos conocimientos en el mecanismo de la fusión viral, se ha estudiado un péptido de 33 aminoácidos perteneciente a la región LLP3 del endodominio de gp41 en presencia de un modelo de biomembrana de alta movilidad, mediante dinámica molecular. En este trabajo se muestra que el péptido se une e interactúa con los fosfolípidos de membrana mediante interacciones de tipo carga-carga. La inhibición de dicha interacción mediante otro péptido o molécula orgánica podría suponer un avance en la búsqueda de inhibidores de fusión entre membranas mediada por la proteína gp41.

**Palabras clave** VIH, gp41, LLP3, fusión, membrana, dinámica molecular.

## Abstract

*Envelope HIV gp41 protein is responsible for the membrane fusion between the virus and the target cell. The gp41 endodomain contains highly helicoidal and hydrophobic regions, the so called LLPs, among them the LLP3 segment. To obtain new knowledge about the viral fusion mechanism, a 33-residue peptide pertaining to the LLP3 region has been studied in presence of a high mobility biomembrane model, by means of molecular dynamics. Our work shows that the peptide binds to and interacts with the membrane phospholipids by charge interaction. Inhibition of this interaction by other peptides or organic molecules might be a tool to inhibit the membrane gp41-mediated fusion.*

**Keywords** HIV, gp41, LLP3, fusion, membrane, molecular dynamics.

# Abreviaturas

<b>CA</b>	Cápsida	<b>LT</b>	Linfocito T
<b>CCR5</b>	<i>C-C Chemokine Receptor type 5</i>	<b>NAMD</b>	<i>NANoscallable Molecular Dynamics</i>
<b>CD4</b>	<i>Cluster of quadruple Differentiation</i>	<b>NC</b>	Nucleocápsida
<b>CHARMM</b>	<i>Chemistry at HARvard Molecular Mechanics</i>	<b>Nef</b>	<i>Negative regulatory Factor</i>
<b>CHR</b>	<i>C-terminal Heptad Repeats</i>	<b>NHR</b>	<i>N-terminal Heptad Repeats</i>
<b>CTT</b>	<i>C-Terminal Tail</i>	<b>P115-Rho GEF</b>	<i>Rho Guanine nucleotide Exchange Factor</i>
<b>CXCR4</b>	<i>Chemokine (C-X-C Motif) Receptor 4</i>	<b>PR</b>	Proteasa
<b>DCLE</b>	Dicloroetano	<b>PSF</b>	<i>Protein Structure File</i>
<b>DVPC</b>	.Divaleril - fosfatidil Colina	<b>Rev</b>	<i>Regulator of Expression of Virion proteins</i>
<b>DVPG</b>	Divaleril - Fosfatidil Glicerol	<b>RMN</b>	Resonancia Magnética Nuclear
<b>DVPE</b>	Divaleril - Fosfatidil Etanolamina	<b>RMSD</b>	<i>Root-Mean-Square Deviation</i>
<b>DVPS</b>	Divaleril - Fosfatidil Serina	<b>RT</b>	Retrotranscriptasa
<b>EboV</b>	<i>Ebola Virus</i>	<b>SIDA</b>	Síndrome de la Inmunodeficiencia Adquirida
<b>EBV</b>	<i>Epstein-Barr Virus</i>	<b>sp1, 2</b>	<i>Spacer Peptide</i>
<b>Env</b>	Envoltura	<b>Tat</b>	<i>Trans-Activator of Transcription</i>
<b>FP</b>	<i>Fusion Protein</i>	<b>TMC</b>	<i>Trans Monolayer Contact</i>
<b>gp160, gp120, gp41</b>	<i>Glycoprotein 160, 120, 41</i>	<b>TMD</b>	<i>Transmembrane Domain</i>
<b>HMMM</b>	<i>High Mobility Membrane – Mimetics</i>	<b>Vif</b>	<i>Viral Infectivity Factor</i>
<b>HR1, HR2</b>	<i>Heptad Repeats 1, 2</i>	<b>VIH</b>	Virus de la Inmunodeficiencia Humana
<b>IFV</b>	<i>Influenza Virus</i>	<b>VMD</b>	<i>Visual Molecular Dynamics</i>
<b>ILA</b>	<i>Interlamellar Attachment</i>	<b>Vpr</b>	<i>Viral Protein R</i>
<b>IN</b>	Integrasa	<b>Vpu</b>	<i>Viral Protein Unique</i>
<b>KE</b>	<i>Kennedy Epitope</i>		
<b>LLP</b>	<i>Lentivirus Lytic Peptide</i>		

# 1. Introducción

## 1.1 Virus: introducción general

Los virus son parásitos intracelulares obligados. Salvo los bacteriófagos (inyectan su ácido nucleico en el interior de bacterias), necesitan hospedarse en el interior de células, liberando su ácido nucleico y forzando a la maquinaria celular para hacer múltiples copias de sí mismos que puedan continuar el proceso de infección y por tanto perpetuar su existencia. De este modo, la entrada del virus en la célula es un proceso clave en todo el proceso de infección vírica. Se distinguen dos clases de virus con respecto a la presencia o ausencia de envoltura. Los que carecen de ella son llamados virus desnudos, y los que la tienen, virus con envoltura. Esta se encuentra rodeando la cápsida y consta de una bicapa lipídica que contiene proteínas transmembrana y de superficie. Los virus con envoltura la adquieren en el momento en que abandonan por gemación su último hospedador, por lo que su composición es variable en función del tipo celular infectado y del organismo. Los virus se sirven de la fusión de su envoltura con la membrana de la célula que van a infectar, para que se cree un poro por el que la cápsida completa pueda introducirse. Para que este proceso pueda tener lugar es preciso que ambas membranas entren en contacto, proceso que se ve favorecido por la presencia de proteínas de fusión alojadas en la envoltura del virus [1 - 3].

## 1.2 Proteínas de fusión. Clasificación y características

La fusión entre la membrana del virus y de la célula tiene lugar por medio de la interacción de las proteínas virales con la envoltura, la membrana de la célula y receptores celulares. Este proceso está muy conservado en los virus con envoltura, si bien es distinto en función del tipo de glicoproteínas de fusión que el virus posea [1 - 3]. Se distinguen tres clases de proteínas de fusión en función de su estructura: clase I, clase II y clase III [1 - 3] (Figura 1).

El virus Influenza (IFV), el virus de la inmunodeficiencia humana (VIH) o el virus del ébola (EBOV) poseen proteínas de fusión de clase I [1]. Las glicoproteínas de fusión de esta clase se encuentran formando trímeros de dímeros asociados de manera no covalente, alojados perpendicularmente a la membrana. Estos trímeros poseen una estructura tubular altamente estable formada por tres  $\alpha$ -hélices en *coiled-coil*. Como se puede ver en la Figuras 1A y 1D, cada dímero está formado por una subunidad de superficie, cuya función es la unión al receptor de membrana de la célula huésped, y una subunidad transmembrana (TM), que contiene un péptido de fusión N-

terminal, cuya función es la fusión de membranas propiamente dicha. La fusión de membranas tiene lugar debido al cambio de la conformación inicial de *coiled-coil* trimérico en  $\alpha$ -hélice a un ovillo de seis hélices que provoca la yuxtaposición de ambas membranas [1 - 3].

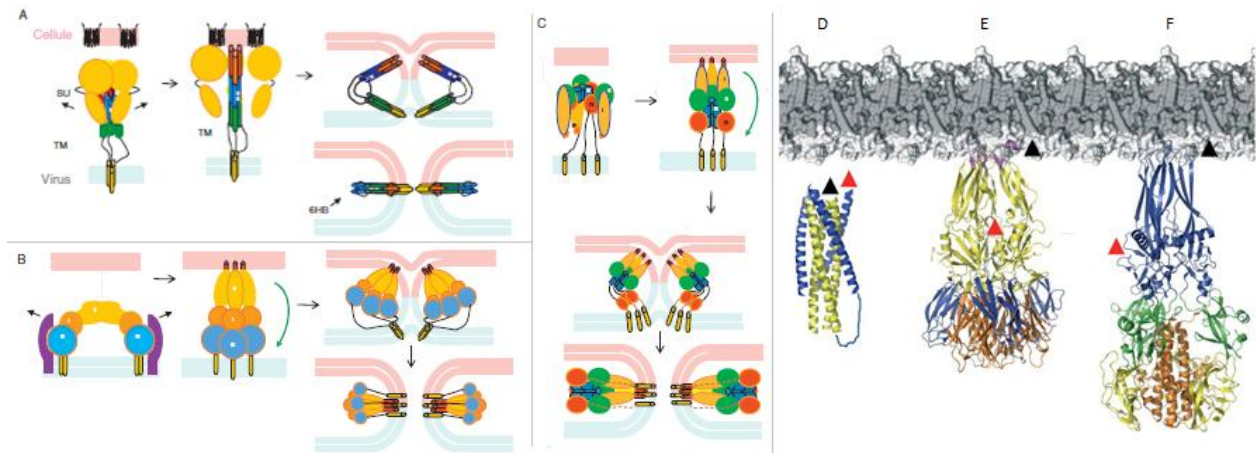


Figura 1. Clasificación de las proteínas de fusión. Clase I (A), clase II (B) y clase III (C). Representación de proteínas de fusión virales de clase I, II y III (D, E y F, respectivamente) en su conformación post-fusión. Las posiciones de anclaje a la membrana mediante el péptido de fusión se indican con los triángulos negros, mientras que los dominios transmembrana se indican con el triángulo rojo. A, B y C adaptado de F.L. Cosset, 2011 [2]; D, E y F adaptado de Weissenhorn et al., 2007 [33].

Las proteínas de fusión de clase II (Figuras 1B y 1E), como las de alfavirus y flavivirus, se caracterizan por estructurarse en dominios en barriles- $\beta$  articulados orientados paralelamente a la membrana. A diferencia de las proteínas de clase I, solamente se organizan en trímeros en el momento de fusión [1] [2]. Contienen un bucle hidrofóbico de fusión interno, cuyo comportamiento con la membrana del hospedador es similar al de los péptidos de fusión de las proteínas de clase I [2, 3].

Por otro lado, las proteínas de fusión de clase III, como el herpes simplex virus (HSV) o el virus de Epstein-Barr (EBV), tienen una estructura combinada. Los trímeros se estabilizan mediante interacciones entre sus  $\alpha$ -hélices centrales, mientras que cada dominio de fusión presenta dos bucles de fusión que se encuentran en el extremo de una hoja  $\beta$  (Figura 1F). También se forma una horquilla de seis hélices  $\alpha$  en su conformación post-fusión, aunque el mecanismo de acción de este tipo de proteínas difiere de los dos anteriores [1 - 3].

## 1.3 Fusión de membranas mediante proteínas de clase I desde el punto de vista proteico y lipídico

### 1.3.1 Cambios en la glicoproteína

Durante el proceso de fusión en proteínas de clase I, tanto la glicoproteína viral como la membrana sufren cambios estructurales y conformacionales [1 - 6] (Figura 2). En cuanto a la proteína de fusión, para que este proceso pueda darse, la subunidad de superficie interactúa con el receptor de la célula huésped, activando a la subunidad transmembrana. Como se muestra en la Figura 2 (Figuras 2.A.2, 2.C, 2.D y 2.E), tras su activación, el ectodominio cambia de su conformación nativa (inactiva) a un intermediario en pre-horquilla, en el que el péptido de fusión (FP) es expuesto al medio mediante la formación de un *coiled-coil* extendido (forma activa). En su conformación nativa, el péptido de fusión yace oculto en el núcleo del trímero, que es hidrofóbico. En la formación de la pre-horquilla, el péptido de fusión se extiende y se proyecta hacia la parte N-terminal de la proteína, permitiendo su introducción en la membrana del hospedador, dada su naturaleza

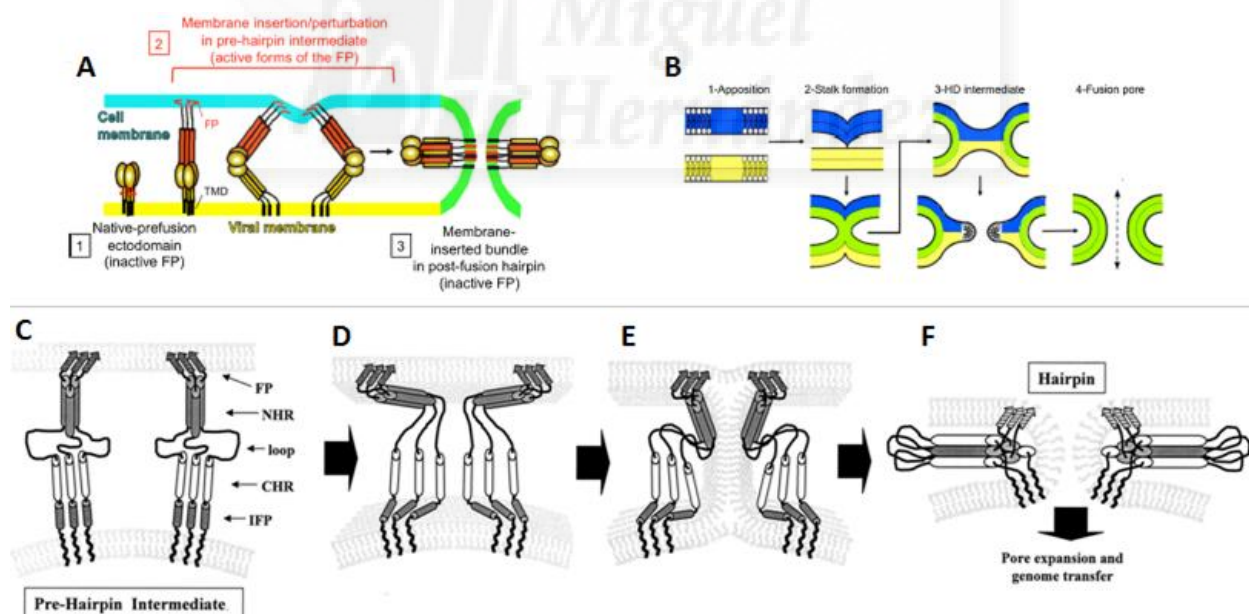


Figura 2. Proceso de fusión de membranas inducido por glicoproteínas virales y modelo de interacción de la subunidad transmembrana con membranas después de su activación tras la unión de la subunidad de superficie a sus receptores. (A) Modelo general de acción de glicoproteínas de fusión de clase I. (B) Fusión de membranas desde el punto de vista lipídico. (C) Liberación del péptido de fusión y anclaje en la membrana. (D) Cambio conformacional. (E) Formación del tallo, hemifusión de membranas. (F) Formación del ovillo de seis hélices y apertura del poro. A y B Adaptadas de B. Apellániz et al., 2014 [1]; C, D, E y F adaptadas de Shnaper et al., 2004 [32].

hidrofóbica (Figura 2.A.1). Esta nueva estructura, estirada y filamentosa, es altamente inestable, por ello tiende al colapso, formando la estructura en horquilla trimérica de baja energía en la que ambos extremos están juntos, yuxtaponiendo las membranas viral y celular (Figuras 2.A.3 y 2.F). Esto se puede deber a que el dominio N-terminal extendido en *coiled-coil* crea unos surcos hidrofóbicos muy conservados en los que las héptadas del C-terminal pueden empaquetarse de manera hidrofóbica y antiparalela, formando el ovillo de seis hélices. En el transcurso de este colapso, la energía liberada por el mismo posibilita el vencimiento de las fuerzas que estabilizan la bicapa lipídica y, por tanto, la unión y fusión de ambas membranas, formándose el poro.

### 1.3.2 Cambios en la membrana

Desde el punto de vista lipídico, para que la fusión pueda tener lugar, es necesario en primera instancia que las dos bicapas se aproximen lo suficiente entre sí como para que las dos hemicapas externas puedan interactuar la una con la otra para formar un poro, venciendo las fuerzas de repulsión que las mantienen naturalmente separadas. Este proceso puede verse detallado en la Figura 3 [1 – 3, 7]. Además, es preciso desestabilizar la membrana, que se encuentra formando una estructura lamelar cohesionada por las fuerzas hidrofóbicas que mantienen las cadenas acíclicas no polares aisladas del medio polar. Se pueden vencer estas fuerzas mediante la intervención de proteínas de fusión, que pueden desestabilizar la membrana por medio de su interacción con los lípidos de dos maneras. La primera se trata de la formación de un poro proteico que atraviesa la membrana. La segunda se corresponde a la formación de un poro lipídico donde las proteínas estabilizan, mantienen el poro abierto y disminuyen la energía necesaria para que el proceso tenga lugar [3]. De acuerdo con el segundo tipo de poro de fusión, se describe el modelo de fusión del tallo –del inglés, *stalk*– (Siegel, 1999) en el que las dos hemicapas que están en contacto (*cis* o proximales) se fusionan, dando lugar al tallo (Figura 3.C). Posteriormente, las dos hemicapas opuestas (*trans* o distales) contactan entre sí, dando lugar al intermediario de hemifusión o TMC (*Trans Monolayer Contact*) (Figuras 3.D y 3.H) que una vez que se rompe, da lugar al poro de fusión, también conocido como lugar de acoplamiento interlamelar o ILA (*Interlamelar Attachment*) [3, 7] (Figuras 3.E y 3.I).



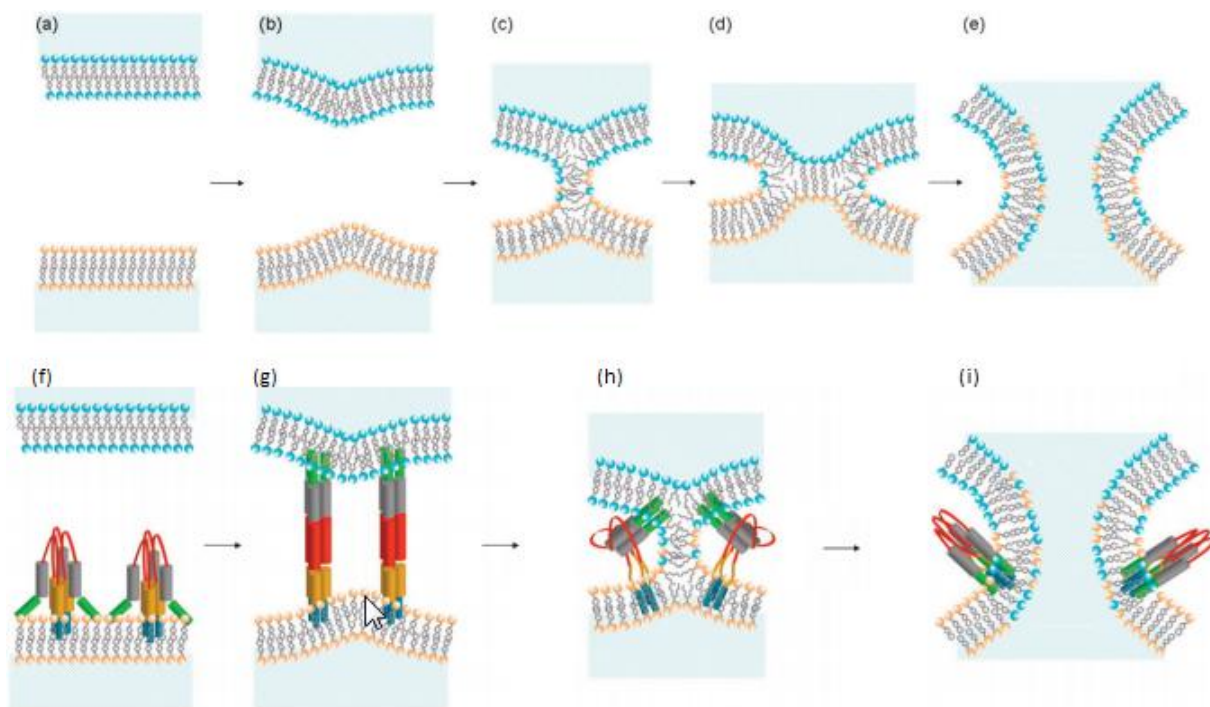


Figura 3. Proceso de fusión de membranas, según la teoría del tallo. (a) estado pre-fusión y (f) estado pre-fusión en el que el péptido de fusión de la glicoproteína viral se encuentra secuestrado, en conformación nativa. (b) Protrusión puntual en las membranas que minimiza la repulsión electrostática y permite la aproximación de las hemicapas *cis* y (g) conformación extendida de la pre-horquilla, que provoca la protrusión en las membranas. (c) Formación del tallo. (d) *Transmonolayer-contact*. (h) Colapso de la pre-horquilla, formación del tallo. (e) y (i) Abertura del poro. Adaptada de H. Marsden et al., 2010 [8].

#### 1.4 SIDA. Insuficiencia de las terapias convencionales para su tratamiento

El Síndrome de InmunoDeficiencia Adquirida (SIDA) surgió a partir de una primera infección por el virus de la inmunodeficiencia humana de tipo 1 (VIH-1), ocurrida en África a causa de una zoonosis entre primates y humanos. Los síntomas principales de esta infección varían en función del estadio de la enfermedad. En un primer momento (2-4 semanas), el VIH causa fiebre, erupciones, dolor muscular, articular, de cabeza y malestar en general. En la fase de latencia del virus, que puede durar varias décadas, el afectado no padece síntomas directamente relacionados con la infección. En etapas avanzadas de la enfermedad, el SIDA se caracteriza por la pérdida rápida de peso, fiebre recurrente, cansancio extremo, diarrea, llagas en la boca, el ano o los genitales, neumonía, pérdidas de memoria y depresión [8, 9]. Desde su origen, 78 millones de personas se han infectado de VIH y 39 millones han fallecido a causa de esta enfermedad. En 2013 había 35 millones de casos de SIDA y causó 1,5 millones de muertes [10].

El tratamiento convencional del SIDA se ha centrado en una terapia antirretroviral. Esta, suele centrarse en procesos específicos del ciclo vital del virus, como pueden ser la entrada, la retrotranscripción, la integración genómica o la maduración de los viriones. Para unos resultados más eficientes y evitar la aparición de cepas resistentes a uno de los tratamientos, suelen aplicarse terapias combinadas, que suponen un buen impedimento de la evolución viral [11]. El problema asociado con este tipo de tratamientos es que, debido a que se aplican post-infección, solamente disminuyen la cantidad de fallecidos a causa de patologías relacionadas con el SIDA, mientras que el número de afectados y el número de nuevas infecciones siguen aumentando [12]. Es por esta razón que en los últimos 10 años se han estado investigando y desarrollando nuevas terapias de prevención frente a la infección del VIH, centrándose en el bloqueo de la entrada del virus en la célula. No obstante, este tipo de inhibidores no es tan eficaz como cabría esperar, por lo que es necesario profundizar más en el tema.

## 1.5 VIH: taxonomía, estructura y genoma

El proceso de fusión es característico de todos los virus que tienen proteínas de fusión de clase I y también está muy conservado. De entre ellos y con el fin de investigar sobre su fusión, este trabajo se centra en el VIH-1, concretamente la cepa HXB2, que se usará como referencia en la numeración de aminoácidos. El VIH pertenece al grupo VI de la clasificación de Baltimore [14]: virus que transcriben en reverso. Taxonómicamente, se incluye dentro del género *Lentivirus*, dentro de la subfamilia *Orthoretroviridae*, perteneciente a su vez a la familia *Retroviridae*. El virus maduro tiene un diámetro de unos 100 nm, tiene una envoltura que consta de un fragmento de la membrana de la última célula huésped que infectó, que presenta espículas proteicas en su superficie (Figura 4A). Contiene dos copias del genoma de (+)ARN monocatenario de 9kb en el interior de una cápsida cónica. El genoma codifica para proteínas estructurales, enzimáticas, reguladoras y accesorias [3, 4, 8, 13].

Las proteínas estructurales están codificadas en dos poliproteínas por los genes *gag* y *env* [4, 8] (Figuras 4A y 4B). La poliproteína Gag, una vez procesada, da lugar a las proteínas matriz (MA, p17), que estabiliza la membrana viral; cápsida (CA, p24), y nucleocápsida (NC, p7) que forman la cápsida cónica que contiene el genoma; p6, que participa en el empaquetamiento y liberación de los nuevos viriones y los péptidos espaciadores sp1 y sp2. Por otro lado, *env* codifica para gp160, proteína que se escinde en gp120 y gp41, las proteínas que se encuentran ancladas formando trímeros en la envoltura viral, permaneciendo unidas de manera no covalente. La proteína gp120 se une a los

receptores de la superficie de la célula diana, siendo por tanto la determinante del tipo celular infectado, mientras que la proteína gp41 se encarga de la fusión de la membrana celular y de la envoltura viral [15, 16].

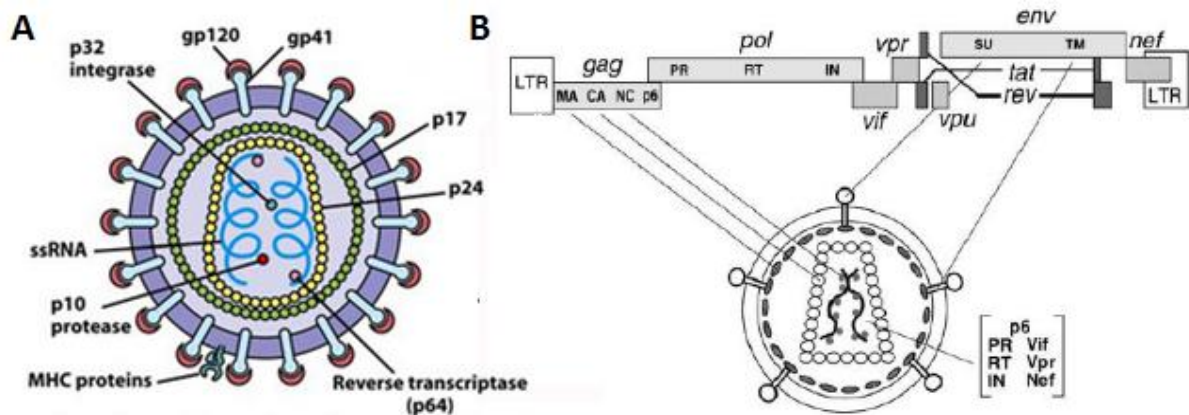


Figura 4. (A) Representación del virión maduro de HIV y (B) de su genoma. (A) Adaptado de M.J. Kozal et al., 1996 [34] y (B) adaptado de J. Barroso, 2011 [4].

Las proteínas con actividad enzimática son producto del gen *pol*: la proteasa (PR), la retro-transcriptasa (RT) y la integrasa (IN) (Figuras 4A y 4B), necesarias para que pueda tener lugar el ciclo vital del virus [4, 8]. La proteasa escinde las poliproteínas en las subunidades maduras que desempeñan una función concreta; la retro-transcriptasa es la enzima encargada de retro-transcribir el genoma de ARN de polaridad positiva a un intermediario de ADN, susceptible de ser integrado en el genoma de la célula hospedadora gracias a la intervención de la integrasa. Una vez integrado, el virus se denomina provirus.

Los genes reguladores son *tat* y *rev* [4, 8] (Figura 4B). El gen *tat* codifica para una proteína que se une al ADN del provirus y promueve su transcripción, mientras que el gen *rev* da lugar a una proteína que se encarga del transporte del ARN viral del núcleo al citoplasma.

Por último, los genes accesorios del VIH son *nef*, *vif*, *vpu* y *vpr* [4, 8] (Figura 4B). El producto proteico de *nef* tiene como función principal aumentar la infectividad del virus y disminuir el reconocimiento y destrucción de las células infectadas reduciendo la expresión de receptores de membrana y del complejo mayor de histocompatibilidad. El producto de *vif* se encarga de mantener la infectividad, degradando inhibidores celulares del ciclo viral. El producto del gen *vpu* tiene como objeto la disminución de la expresión de los receptores CD4, induciendo su degradación a nivel del retículo endoplásmico. Por último, la proteína generada por *vpr* tiene como función la asistencia en la integración del genoma [4].

## 1.6 Proteínas de fusión del VIH: gp120 y gp41

### 1.6.1 Función e interacciones de gp120

El VIH ataca al sistema inmunitario del huésped mediante la infección y destrucción de los linfocitos T (LT) y los macrófagos [8]. El virus infecta estas células mediante su unión a la superficie del LT y posterior fusión de la envoltura viral y la membrana celular, permitiendo la entrada de la cápsida dentro del compartimento celular (Figura 5). El VIH se une a la célula mediante la interacción entre la glicoproteína de envoltura gp120 y el receptor de membrana CD4, que tiene un dominio de alta afinidad por gp120 [3]. Esta unión produce un cambio conformacional en el complejo gp120/CD4 que permite la interacción de gp120 con el correceptor celular, que suele ser CCR5 o CXCR4. Una vez que gp120 se encuentra unido al receptor CD4 y a su correceptor, deja de interactuar con gp41, que queda expuesto al medio siendo susceptible de cambiar a la conformación en pre-horquilla extendiéndose y proyectando su extremo N-terminal hacia la membrana del LT [3, 4] [16 - 18].

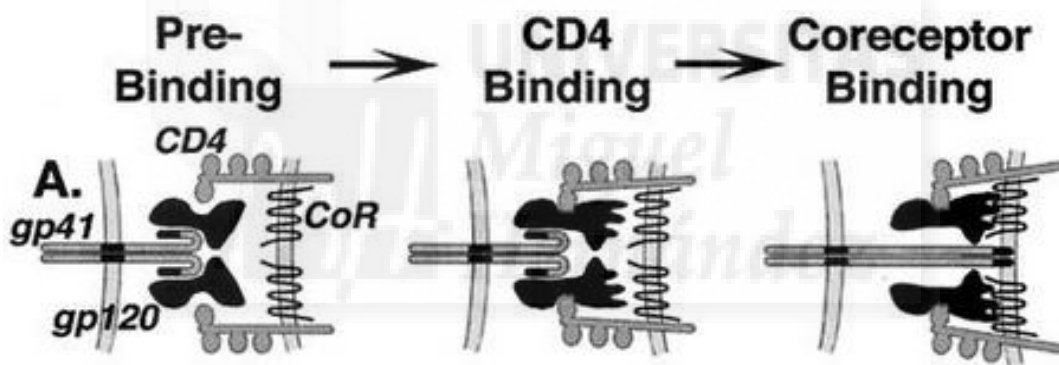


Figura 5. Modelo representativo de la interacción entre gp120 y CD4 y su co-receptor. Solo se muestran dos de las tres subunidades del trímero por comodidad. Adaptada de K. Salzwedel et al., 2000.

### 1.6.2 Estructura y conformaciones de gp41

Por otro lado, gp41 (513 – 856) [15, 19] tiene la secuencia altamente conservada, aunque su estructura y conformación nativas no se conocen con claridad. Por el momento, se conocen dos posibles conformaciones para gp41. Una de ellas, la conformación clásica y más conocida, posee un dominio transmembrana (*transmembrane domain*, TMD) [1, 5, 11] (Figura 6); la otra posibilidad, una conformación con múltiples TMD, se ha estado investigando los últimos años y, al parecer, puede haber como mínimo tres segmentos que atraviesen la membrana. Las regiones exactas que constituyen los nuevos TMD todavía permanecen inciertas [15, 17, 23].

En el modelo de un solo TMD, el N-terminal se corresponde con el ectodominio, es decir, la región que se encuentra hacia el exterior de la bicapa (Figuras 6 y 7A). Los primeros aminoácidos del ectodominio constituyen el péptido de fusión o FP. Se trata de una región muy hidrofóbica, en la cual los residuos del aminoácido glicina confieren la flexibilidad necesaria para que pueda introducirse en la membrana. Se sabe que el FP se inserta en forma de  $\alpha$ -hélice, aunque una vez insertado podría cambiar a conformación de hoja  $\beta$  induciendo la desestabilización de la bicapa [3]. También hay estudios que sugieren que podría insertarse en forma de hoja  $\beta$  [1, 24] y completar el proceso de fusión en conformación de  $\alpha$ -hélice. Esta incertidumbre en cuanto a la conformación del péptido de fusión frente a la membrana puede ser debida a las condiciones de experimentación, ya sean la temperatura del sistema, el método de detección y análisis y la composición lipídica [24].

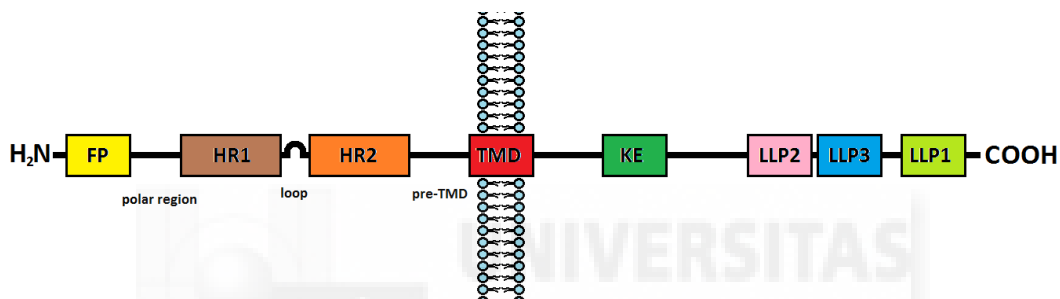


Figura 6. Representación esquemática de los dominios de la glicoproteína gp41 de VIH (conformación clásica).

Seguidamente del FP se encuentra una pequeña región polar precediendo una hélice constituida por dos repeticiones de héptadas, HR-1 y HR-2, también llamadas repeticiones de héptadas del N-terminal (NHR) y del C-terminal (CHR), respectivamente [1, 3]. Las héptadas se encuentran separadas por un pequeño bucle que posee dos cisteínas susceptibles de formar un puente disulfuro. Son elementos clave en la formación de la estructura filamentosa en *coiled-coil* de la pre-horquilla [1, 5]. La última región descrita del ectodominio de gp41 se trata del dominio pre-transmembrana (pre-TMD). Se trata de un dominio muy conservado, rico en triptófanos, que posee una secuencia de unión a colesterol, por lo que se le asocia la inducción de estructuras abultadas o curvadas en las membranas [1, 3]. La región transmembrana clásica se localiza entre los residuos 684 y 706 en la cepa HXB2. Constituye una región muy hidrofóbica y conservada [1, 5, 15, 19].

El endodominio de gp41, también conocido como cola citoplásmica C-terminal (CTT), se ha mantenido en un plano secundario mientras se investigaba a fondo el ectodominio [20, 21]. Por eso, muchos de los aspectos relacionados con su función biológica se desconocen todavía. Asimismo, se sabe que es inusualmente largo: tiene aproximadamente unos 150 aminoácidos [20 – 22]. Se le

asignan funciones de regulación de la fusogenicidad, replicación, infectividad, transmisión, citopatogenicidad, desprendimiento de la envoltura, ensamblaje de componentes víricos... [20].

El CTT contiene cuatro dominios característicos: el epítipo de Kennedy o KE (724 – 745) y tres pequeñas regiones llamadas “lentivirus lytic peptide” o LLP, i.e., LLP1 (828 – 852), LLP2 (768 – 788) y LLP3 (789 – 815) [16, 19] (Figuras 6 y 8A). Los LLP son tres segmentos anfipáticos y helicoidales, que parece que interactúan con la membrana interna de la envoltura del VIH o la de la célula. Se sabe que LLP1 y LLP2 tienen efecto citolítico en procariontas y eucariontas [20]. También se conoce que los LLP interactúan con proteínas. Por ejemplo, LLP1 interactúa con calmodulina y LLP3 con p115-RhoGEF, un factor intercambiador de guanina [20].

En la secuencia de los LLP, hay tres veces más arginina que en el resto del promedio de proteínas, mientras que hay tres veces menos lisina que en el promedio. Esto significa que se conserva la arginina en gran medida, incluso frente a mutaciones conservativas del tipo Arg→Lys [16], que tienen carga muy similar y la diferencia de masa es mínima. No obstante, este hecho puede ser debido a que la arginina tiene como cadena lateral un grupo guanidinio, que es susceptible de formar más puentes de hidrógeno que el grupo amino que tiene la lisina. También, la arginina se une mejor al grupo fosfato de los lípidos que la lisina, hecho enormemente importante en la interacción de proteínas con membranas [16].

Mediante ensayos con mutantes Arg→Lys [16], se ha concluido que la arginina en LLP1 es necesaria para la fusogenicidad normal de gp41, así como su expresión e incorporación a la membrana, la replicación del virus, cinética e infectividad. Por otro lado, la arginina en LLP2 es necesaria para la incorporación de gp41 a la membrana y para la fusogenicidad célula-célula [16].

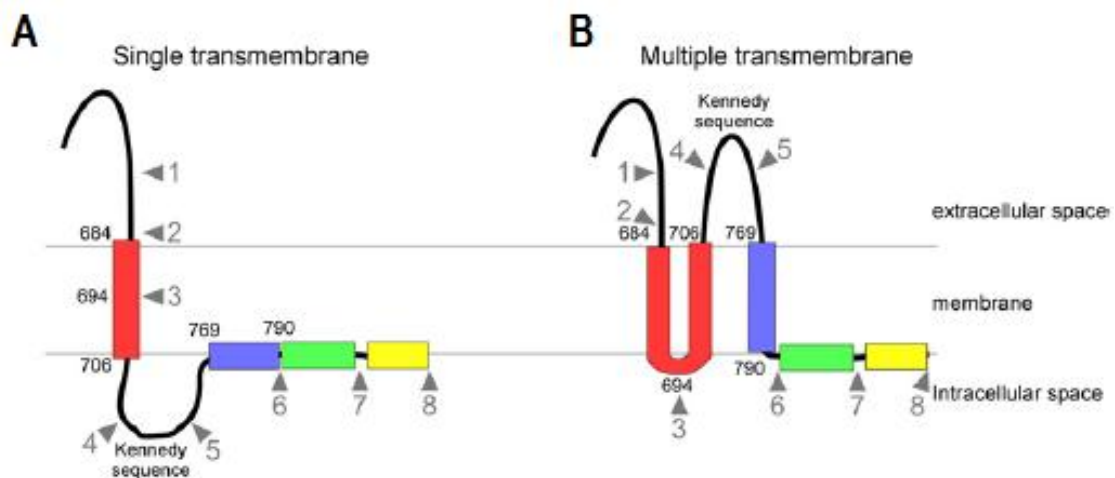


Figura 7. Modelos estructurales teóricos propuestos para el endodominio de gp41, para uno (A) o múltiples (B) dominios transmembrana. TM, LLP2, LLP3 y LLP1 se muestran en rojo, azul, verde y amarillo, respectivamente. Adaptado de L. Shujun et al., 2010 [15].

Por otra parte, el segmento correspondiente a LLP3 contiene una cremallera de leucinas, con 5 leucinas, y una región rica en aromáticos, puesto que presenta 4 triptófanos y 2 tirosinas [20, 22].

LLP3 tiene carga positiva, interacciona fuertemente con lípidos de carga negativa y también con lípidos zwitteriónicos [20, 22]. Se inserta en la membrana e induce perturbaciones cuando se trata de lípidos neutros o de carga positiva, ya que la interacción entre LLP3 y lípidos de carga negativa es tan fuerte que el péptido se adhiere a la superficie de la membrana y por tanto no se integra [20]. Teniendo en cuenta la capacidad de interacción de LLP3 con membranas cabría suponer que esta región puede interactuar con la membrana en sinergia con otras regiones, colaborando activamente en el proceso de fusión para la entrada y la salida [20].

Por otro lado, en el modelo de varios TMD, el segmento original atraviesa la bicapa dos veces y existe un tercero que corresponde, según las predicciones, a LLP2 [15] o a un fragmento del epítipo de Kennedy [17]. El equipo de Liu et. al., en 2010 (Figura 7), predice el TMD previamente conocido de 23 aminoácidos en la posición 684 – 706 de gp41 en la cepa HXB2. No obstante, un análisis de unión de anticuerpos contra KE y LLP2, epítipos teóricamente localizado en el endodominio, resulta positivo en el compartimento extracelular, contradiciendo por tanto la teoría de un único dominio transmembrana, pues para que KE y LLP2 puedan encontrarse en el exterior, gp41 tiene que sufrir un cambio conformacional que suponga como mínimo una inserción en la membrana por parte de otro segmento de la cola intracitoplásmica [15].

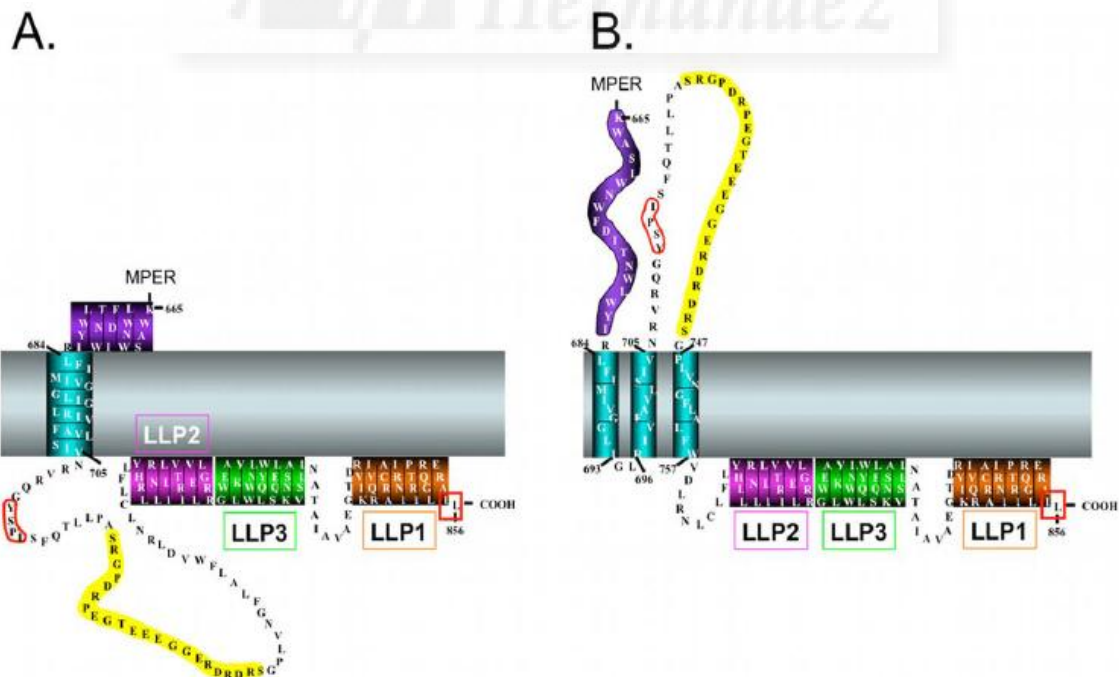


Figura 8. Diferentes modelos estructurales para el endodominio de gp41. (A) Modelo tradicional con un solo TMD en  $\alpha$ -hélice. (B) Modelo alternativo con tres TMD en hoja  $\beta$ , en el que el KE está expuesto al medio extracelular. Adaptado de J.D. Steckbeck et al., 2010 [17].

Por otro lado, el equipo de Steckbeck et. al. en 2010 apunta a que la conformación de gp41, concretamente de su CTT, varía en función de la proximidad que tenga el virión a la célula hospedadora (Figura 8). Es decir, sugieren que la conformación nativa de gp41 tiene un solo TMD, pero al aproximarse el VIH a una célula LT, se producen cambios en la conformación de la glicoproteína que suponen la exposición, aunque transitoria, del KE al medio extracelular. De este modo, un fragmento del propio KE, LLP2 o un segmento intermedio entre ambos dominios [23] constituiría un tercer TMD.

## **1.7 Otras terapias antivirales: inhibidores de fusión de primera, segunda y tercera generación**

Puesto que se sabe que el ectodominio de gp41 es esencial para la fusión de membranas, se han diseñado inhibidores de la fusión complementarios a regiones específicas del mismo, con la finalidad de impedir que pueda desempeñar su función, evitando por tanto la fusión de membranas y la formación del poro de entrada [11, 12, 16]. A fecha de hoy, ha habido tres generaciones de inhibidores de fusión para el VIH. El primer inhibidor diseñado fue *Maraviroc* (Selzentry, Celsentri de Pfizer). Se trata de un antagonista del co-receptor CCR5. *Enfuvirtide* o T20 (Fuzeon de Roche/Trimeris), *SJ-2176* y *C34* son de primera generación. Se trata de pequeños péptidos análogos de HR2, cuyo modo de acción consiste en la competición con HR2 para unirse a HR1. De este modo, se impide el empaquetamiento de HR2 en los surcos hidrofóbicos de HR1, bloqueando a su vez la formación del paquete de seis hélices subsiguiente al colapso de la estructura en pre-horquilla a la horquilla post-fusión (Figura 9). Este bloqueo consigue inhibir la dispersión del virus vía célula-célula y constituye un avance muy útil si se usan en co-terapia con otros fármacos anti-VIH, pero la baja biodisponibilidad que presentan estos péptidos y la corta vida media en plasma limitan su uso e hicieron necesario el desarrollo de inhibidores de segunda generación [11].

A partir de la secuencia de T20, se diseñó T1249, un péptido de mayor longitud que su predecesor y con una secuencia ligeramente diferente, consiguiendo una respuesta mejor incluso en pacientes en los que la terapia con T20 fue infructuosa. El problema de T1249 es que presenta estructura en *random-coil*, por lo que es sensible a la degradación por proteasas y se une con menor afinidad a HR1 al carecer de la estructura en  $\alpha$ -hélice necesaria [11].



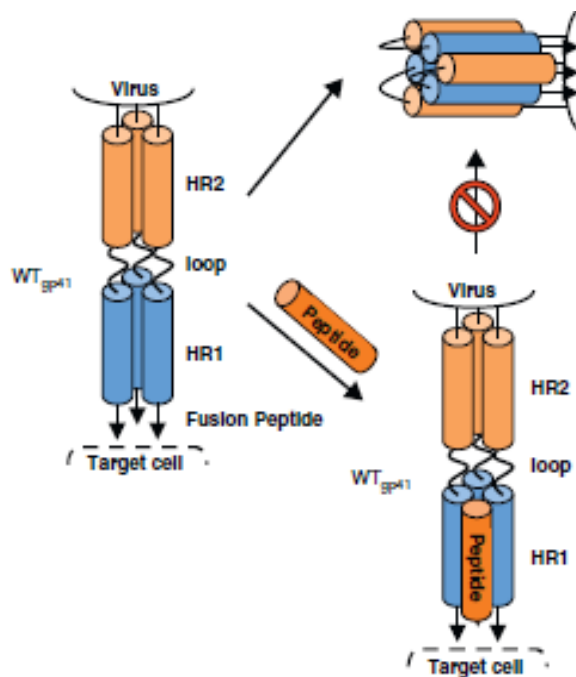


Figura 9. Parte superior: cambio conformacional de gp41 de su estructura extendida (pre-horquilla) al ovillo de seis hélices, susceptible de ser inhibido, en el que se empaqueta HR2 hidrofólicamente en los surcos de HR1. Parte inferior: inhibición del colapso por parte de un péptido análogo a HR2. Adaptado de B. Berkhout et al., 2012 [11].

En la tercera generación de inhibidores de fusión, se pretendió mejorar la vida media en plasma y la estructura en  $\alpha$ -hélice, añadiendo puentes salinos y diaminoalcano en cada giro de la hélice. Se diseñaron *Cp32M*, *Sifuvirtude* y *T2635*, que presentan una mayor estabilidad térmica, vida media en plasma y potencia antiviral. Se incrementa la resistencia a proteasas y peptidasas añadiendo grupos funcionales en el extremo N-terminal y en el C-terminal. También se mejoran las propiedades farmacocinéticas y la vida media del péptido en suero mediante otras modificaciones y su unión a proteínas de plasma. De este modo, se ha estudiado en ratones un nuevo inhibidor conjugado a través de un agente de entrecruzamiento de polietilenglicol con un transportador pentasacárido que une antitrombina, con una vida media *in vivo* prolongada y una mayor actividad antiviral. En humanos, este nuevo inhibidor tendría una vida media de 120h, lo que implicaría únicamente una inyección por semana [11].

## 1.8 Introducción a la simulación de sistemas biológicos por ordenador

El estudio de los sistemas biológicos constituidos por membranas resulta arduo y complejo. Esto es debido en parte a la gran diversidad en estructura y composición de las membranas de distintos organismos, orgánulos y compartimentos intracelulares, puesto que existen miles de especies únicas

de lípidos, con propiedades y comportamientos característicos, que se presentan en proporción variable junto con otros componentes como proteínas [25, 28]. Además, la interacción de proteínas con membranas resulta difícil de estudiar, ya que generalmente las secuencias que interaccionan con lípidos o se insertan en el interior de la bicapa suelen ser muy hidrofóbicas y agregan en solución. De este modo, estudiar directamente en células procesos de interacción proteína-lípido resulta muy difícil, debido a que las técnicas de estudio actuales (espectrometría, cristalografía) carecen de la resolución espacial (escala atómica) y temporal (nanosegundos) necesarias para una observación directa de este tipo de fenómenos [26].

Por este motivo, técnicas computacionales como el modelado molecular suponen un gran avance, ya que se pueden conseguir simulaciones a gran escala temporal, permitiendo observar la evolución del sistema a largo plazo en periodos de tiempo más breve, obteniendo representaciones más realistas. Actualmente estas técnicas se usan en paralelo con las experimentales para ensayos descriptivos o predictivos [25, 27].

Las simulaciones mediante dinámica molecular se fundamentan en la mecánica clásica y estadística. Requieren un equipo de hardware rápido y capaz de llevar a cabo cálculos matemáticos muy complejos. Las interacciones se calculan entre átomos, mediante el uso de parámetros que definen el campo de fuerza, en función de la variación de energía potencial. En cada paso de la simulación, se determina la variación de energía de cada átomo con respecto a sus coordenadas. El resultado de una simulación de dinámica molecular es un archivo de trayectoria de pocos micro o nanosegundos que incluye información acerca de los movimientos e interacciones que pueden ocurrir en un sistema bajo unas condiciones determinadas, incluyendo aquellas que no son accesibles experimentalmente [25 - 27].

Sin embargo, las membranas biológicas son muy diferentes a los sistemas modelo, puesto que contienen muchas proteínas, están sometidas a fuerzas de no-equilibrio y a los efectos del citoesqueleto, que no se reflejan en las simulaciones [27].

Por otro lado, la mayoría de los procesos biológicos de interés suceden en la escala de los microsegundos o milisegundos, por lo que las simulaciones por dinámica molecular se ven limitadas por los cortos intervalos de tiempo requeridos para una correcta integración y descripción de la trayectoria molecular. No obstante, la mejora de los algoritmos matemáticos, la disponibilidad de hardware más rápido y el refinamiento de los campos de fuerza disponibles hacen que la distancia que separa las simulaciones y los experimentos sea cada vez menor [27].

## 1.9 Justificación del presente trabajo

Se sabe que el ectodominio de gp41 juega un papel esencial en la fusión de las membranas durante la entrada del VIH pero estudios recientes [3, 11, 12, 15 – 17, 20, 23, 27] apuntan a que la región C terminal (teóricamente en el endodominio de gp41) puede intervenir también en la fusión, dada la dualidad de estructura que esta glicoproteína presenta, así como la capacidad de unión a membranas que tiene el CTT.

Por estos motivos, para este trabajo se ha considerado que un estudio en profundidad sobre la interacción entre la región LLP3 y la membrana podría resultar esclarecedor y aportar más luz a los pocos rasgos conocidos del endodominio de esta glicoproteína. El diseño de un inhibidor de la interacción entre LLP3 y la membrana podría abrir una vía para el desarrollo de nuevas terapias con mayor eficacia que las actualmente conocidas.



## 2. Antecedentes y objetivos

En la última década del siglo pasado se ha estado estudiando de manera exhaustiva el ectodominio y la región transmembrana de gp41, mientras que su endodominio, su estructura y sus funciones permanecían en segundo plano, con una importancia menor. No obstante, recientemente muchos grupos de investigación están trabajando en la tarea de descubrir nuevos aspectos del endodominio de la glicoproteína de fusión del VIH y, por tanto, haciendo surgir nuevas necesidades, nuevos descubrimientos y más dudas.

El equipo de investigación del profesor José Villalaín ha estado investigando el VIH, sus glicoproteínas y su interacción con membranas mediante técnicas espectroscópicas (fluorescencia, RMN e infrarrojo) y calorimetría [35 - 39]. En 2006, se realizó un estudio para determinar las regiones más membranotrópicas de gp41, concluyendo que el FP, una pequeña región del NHR inmediatamente consecutiva al FP, parte del bucle, el dominio pretransmembrana, el TMD y los LLP tienen características membranotrópicas en mayor o menor medida, mientras que el CHR y el KE no las poseen. En 2007, se demostró que la primera parte de NHR interacciona específicamente con lípidos con carga negativa, induce desestructuración y agregaciones en liposomas, por lo que se induce que participa en sinergia con otras regiones membranotrópicas en la fusión de membranas. Por último, en 2008 se realizó un estudio de LLP3 y se determinó que afecta al comportamiento de fase de la membrana y que induce desestructuración y agregación en liposomas. Sugirieron que podría formar parte de un bucle externo junto con KE, pudiendo tener una función moduladora en la fusión viral o la gemación, en sinergia con otras regiones de gp41.

De este modo y siguiendo con la línea de investigación del grupo, el objetivo del presente trabajo de fin de grado es la determinación y caracterización de las interacciones de un péptido correspondiente al dominio LLP3 de gp41 con una membrana de composición lipídica heterogénea, mediante una simulación *in silico*. En el caso de que se determine la existencia de una interacción específica en lípidos en general o algún lípido en particular, se podría proceder al diseño de un inhibidor de dichas interacciones con el fin último de encontrar un antiviral específico para este virus.

## 3. Materiales y Métodos

### 3.1 Caracterización del sistema

La secuencia del dominio LLP3 de la glicoproteína gp41, tiene 32 aminoácidos de longitud y es la siguiente: LLGRRGWVFLKYWWNLLQYWSQELKSSAVSLL. Con el fin de dotar a la secuencia de estructura secundaria en  $\alpha$ -hélice, se utilizó el programa PyMol (Figura 10A) [29].

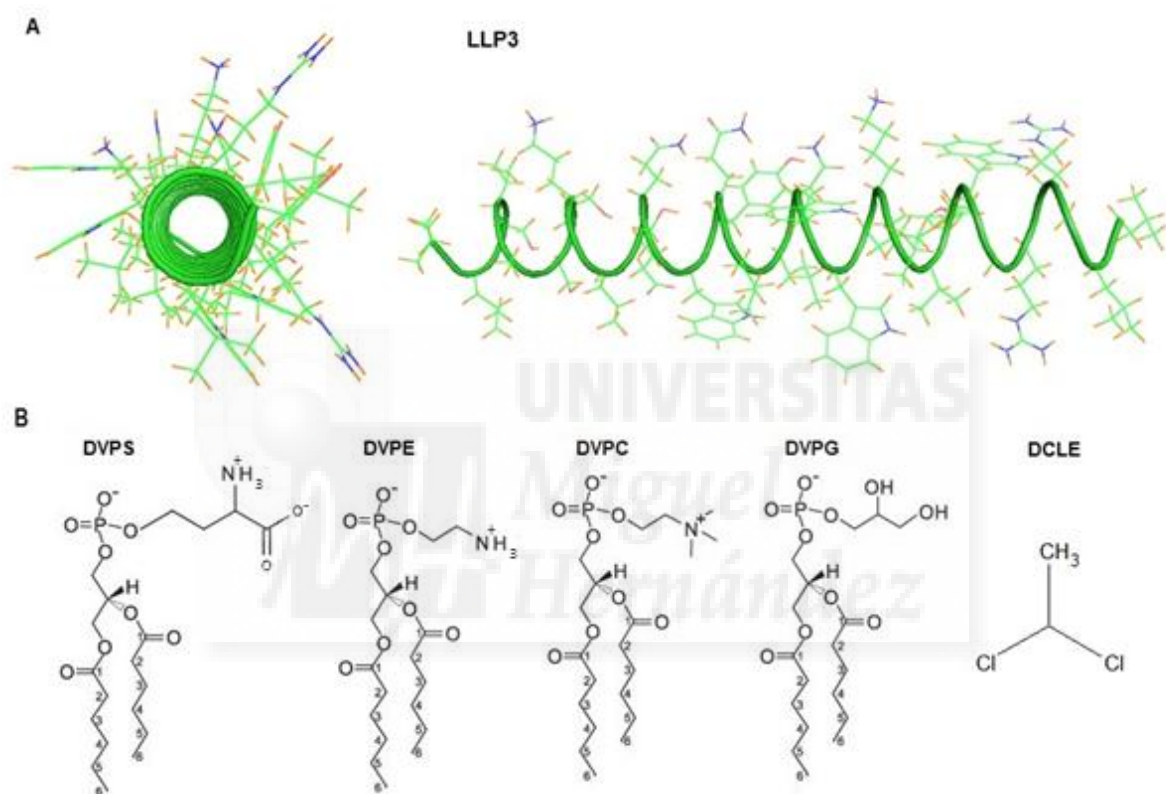


Figura 10. (A)  $\alpha$ -Hélice construida a partir de la secuencia primaria de LLP3 utilizando Pymol [29]. (B) Estructura molecular de los lípidos DVPS, DVPE, DVPC y DVPG a pH fisiológico (todos ellos se han derivado de POPS, POPE, POPC y POPS para realizar la HMMM) y del DCLE (dicloroetano) presente en la interfase de la bicapa.

Las dos capas de la membrana se mantienen unidas entre sí mediante una interacción de tipo hidrofóbico entre las cadenas alifáticas de los lípidos que las configuran. La longitud de dichas cadenas determina la velocidad de la simulación, pues a mayor longitud de las mismas, la interacción hidrofóbica entre las hemicapas será más fuerte. Este hecho implica que el deslizamiento lipídico entre la capa superior y la inferior será más lento también, teniendo como consecuencia un aumento en el tiempo necesario de simulación para lograr determinar las interacciones objeto de estudio.

Por tanto, para que el proceso de simulación sea lo más corto posible, pero siendo fiel en todo momento a la realidad, se utilizan, en lugar de bicapas convencionales, los sistemas modelo HMMM (*High Mobility Membrane – Mimetics*; miméticos de membranas de alta movilidad). Estos modelos tienen dos rasgos característicos. En primer lugar, los lípidos que forman la bicapa son de cadena corta (derivados de lípidos de cadena larga) y, en segundo lugar, hay una interfase de naturaleza orgánica presente entre las dos bicapas. Estos dos factores aceleran el movimiento de los lípidos (deslizamiento, rotación, etc), por lo que el tiempo de simulación se reduce notablemente (2 – 3 veces menor).

La construcción de una membrana biológica para llevar a cabo la simulación de la interacción péptido-lípido se llevó a cabo de manera rápida y sencilla gracias al uso de la aplicación web <http://www.charmm-gui.org> [Figura 11] [30].

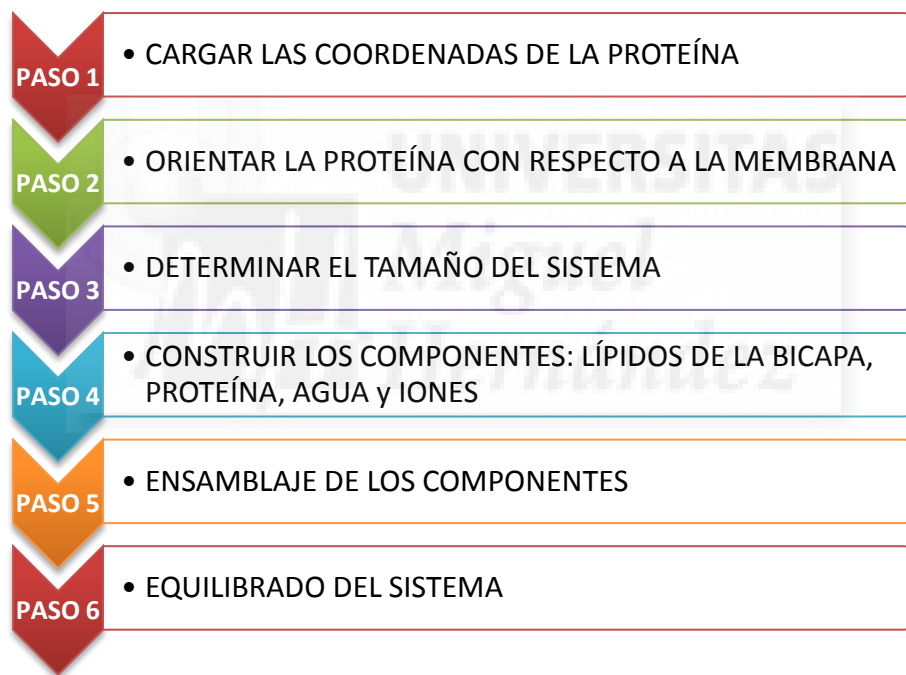


Figura 11. Esquema del proceso de trabajo que sigue la aplicación web charmm-gui para la obtención del sistema proteína-membrana.

### 3.2 Simulación de dinámica molecular

Hemos llevado a cabo simulaciones de dinámica molecular usando NAMD2.9 (*NANoscalable Molecular Dynamics*) [40], el campo de fuerza CHARMM36 para los lípidos [41] y el campo de fuerza general CHARMM para LLP3 [42], obtenidos de [http://mackerell.umaryland.edu/charmm\\_ff.shtml](http://mackerell.umaryland.edu/charmm_ff.shtml). Para el agua, hemos usado el modelo TIP3P [43]. Todas las simulaciones se han llevado a cabo con un

número constante de partículas en condiciones NPT a 1 atm y 310K. El lapso de tiempo por paso de simulación escogido ha sido 2 fs. Se ha utilizado el método de Nosé-Hoover Langevin para mantener la presión constante [44, 45]. La temperatura se ha mantenido constante mediante la dinámica de Langevin, con un coeficiente de amortiguación ( $\gamma$ ) de  $0,5 \text{ ps}^{-1}$ . Para calcular la interacción a largo alcance de los sistemas, se ha usado el método de la malla de Ewald para partículas estándar en condiciones límite periódicas. Las interacciones de no-enlace se restringen a  $12 \text{ \AA}$ , con una función de suavizado aplicada tras los  $10 \text{ \AA}$  [46]. Antes de la simulación y con el fin de eliminar los contactos atómicos desfavorables, los sistemas modelo de membrana tuvieron un proceso de equilibrado de 1 ns, después de 50000 pasos de minimización. La trayectoria de los sistemas lipídicos se ha calculado durante 90 ns. Todas las simulaciones se han llevado a cabo en un entorno electrostáticamente neutro con un número apropiado de iones de sodio/cloro.

### 3.3 Modelo de membrana

Hemos utilizado un modelo mimético de membrana lipídica con alta movilidad (HMMM), compuesto de DVPC / DVPE / DVPS / DVPG (divaleril-fosfatidilserina, divaleril-fosfatidiletanolamina, divaleril-glicero-3-fosfofoserina, divaleril-fosfatidilglicerol) en una relación molar de 3: 1: 1: 1. Estos, derivados de POPC, POPE, POPS y POPG (palmitoiloleil-fosfatidilserina, palmitoiloleil-fosfatidiletanolamina, palmitoiloleil-glicero-3-fosfofoserina y palmitoiloleil-fosfatidilglicerol), conservan la cabeza polar del precursor y los seis primeros carbonos de sus cadenas acilo. Como interfase orgánica que acelere el proceso de simulación, se ha utilizado DCLE (dicloroetano). El modelo consta de 300 moléculas de lípido (150 lípidos en cada hemicapa) en una caja rectangular. El sistema de derivados DVPC / DVPE / DVPS / DVPG estaba rodeado de 56502 moléculas de agua [47]. La Figura 10 presenta la estructura y numeración de carbono de las moléculas consideradas en este estudio, LLP3 (Figura 10A), los derivados de POPS, POPE, POPC, POPG y DCLE (Figura 10B). La bicapa se localiza en el plano  $xy$  y la normal de la bicapa es paralela al eje  $z$ . Inicialmente, la caja de simulación tenía las dimensiones de 8,9 nm en las direcciones  $x$  e  $y$ , y 8 nm en la dirección  $z$ . La altura de la caja de simulación y las áreas de sección transversal podían variar independientemente una de otra. Los modelos de membrana se han preparado utilizando el servidor web CHARMM-Gui (<http://www.charmm-gui.org>, [30]). El plug-in de generación automática de PSF de VMD [31] se ha utilizado para obtener los archivos PSF y PDB necesarios para que NAMD pueda construir el sistema completo.

### 3.4 Análisis

La superficie de la membrana se define por la capa de los átomos de los grupos fosfato de las cabezas de los fosfolípidos y se orienta paralela al plano  $xy$ . El espesor de la bicapa se ha definido como la distancia media entre los átomos de fósforo de los fosfolípidos de las hemicapas opuestas, mientras que el centro de la bicapa,  $z = 0$ , se ha definido por el centro de masa de los átomos de las moléculas de fosfato de los fosfolípidos. El centro de masa, distancia  $z$ , de la molécula LLP3 con respecto al centro de la bicapa se ha utilizado para estudiar su distribución espacial en la membrana. Se han usado VMD [31] y Pymol [29] para la visualización y el análisis. El grosor de la membrana, el centro de masas y el RMSD (*Root-Mean-Square Deviation*) se han obtenido mediante el plugin de VMD Membplugin [48] mientras que los perfiles de densidad de masas se han calculado utilizando el plugin de VMD Density Profile Tool [49].





## 4. Resultados y discusión

El péptido LLP3 es una región del CTT de gp41 que consta de 15 aminoácidos de longitud y tiene estructura helicoidal. LLP3 está cargado positivamente e interacciona con lípidos zwitteriónicos y de carga negativa, pudiendo inducir perturbaciones en la membrana e insertarse. Para caracterizar molecularmente su interacción con la membrana, se ha utilizado una simulación de dinámica molecular mediante la aproximación del modelo HMMM.

El modelo HMMM estaba compuesto de 4 tipos de fosfolípidos de cadena corta, DVPC y DVPE, zwitteriónicos y DVPS y DVPG, cargados negativamente, en la proporción 2:1:1:1. Para asegurar que, en el caso de que hubiera un proceso de unión, este no se viera influido por el hecho de haber situado el péptido en contacto con la membrana (o en su interior) al inicio de la simulación, LLP3 se ubicó en medio de la capa de agua sin tocar la superficie de la membrana, a 15 Å de la superficie (Figura 12A). Como se observa en el último paso de simulación (Figura 12B), la interacción específica de LLP3 con la superficie de la membrana nos ha permitido la caracterización del proceso de unión, determinar los residuos involucrados en la unión y describir la localización y orientación de LLP3 en la membrana.

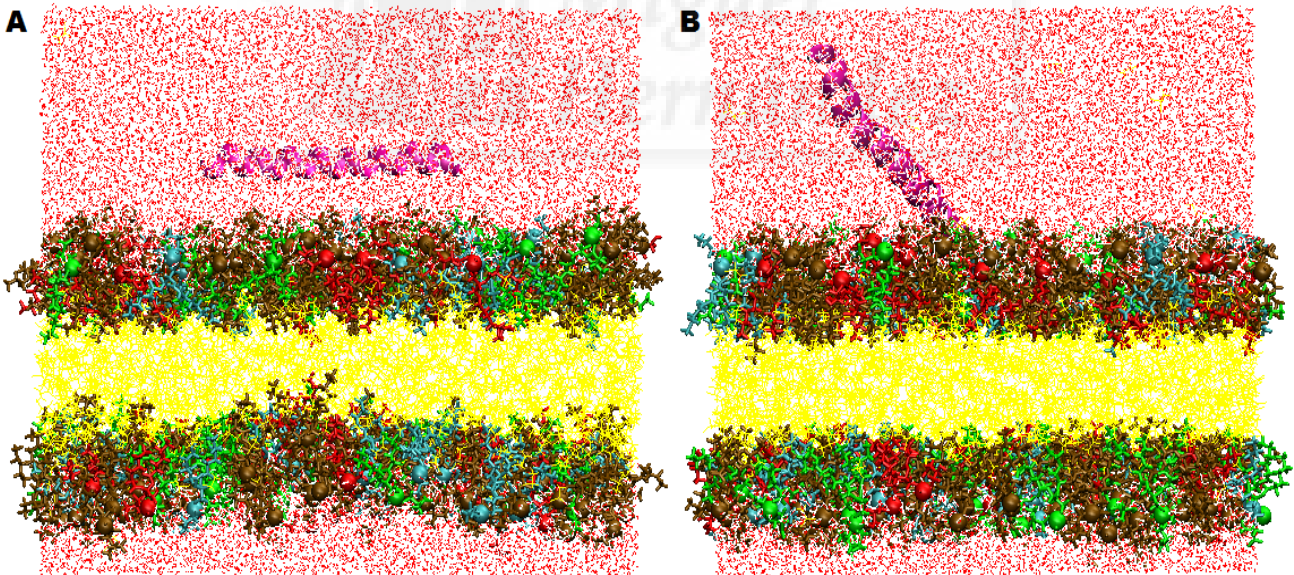


Figura 12. Representación del sistema al inicio (A) y al final (B) de la simulación. En rosa, LLP3. Formando la membrana: en marrón, DVPC; en rojo, DVPG; en verde, DVPS, en azul, DVPE; en amarillo DCLE. Rodeando todo el sistema, moléculas de agua (en rojo y blanco).

Con el fin de conocer la estabilidad del sistema, se han calculado el grosor de la membrana, el centro de masas de LLP3 y de los fosfolípidos de la membrana y el RMSD de LLP3 (Figuras 13 a 15). Se entiende por grosor de la membrana la distancia promedio entre los átomos de fósforo del grupo fosfato de las cabezas polares de los fosfolípidos que configuran las dos capas, Z+ y Z-. La figura 13 representa el grosor de la membrana durante todo el proceso de simulación (85 ns aproximadamente). En ella se ve que el grosor de la membrana simulada tiende a ser constante, con un ligero cambio durante los primeros 25 ns, que puede ser debido a reestructuraciones de los lípidos que la configuran. El hecho de que no haya variaciones significativas a nivel global a partir de los 25 ns, sugiere que los fosfolípidos se equilibraron al principio de la simulación y que el sistema se estabilizó tempranamente alcanzando un estado estacionario, confirmando que la metodología de simulación era adecuada. Por tanto, la aproximación HMMM para simulaciones con membranas es capaz de alcanzar un sistema estable.

Por otro lado, la oscilación del centro de masas de LLP3, la membrana y los fosfolípidos se muestra en la figura 14. El centro de masas de LLP3 osciló entre 42 Å y 27 Å durante toda la simulación, mostrando una evidente dispersión entre las distancias máxima y mínima. Por otro lado, el centro de masas de los lípidos de los fosfolípidos de cada hemicapa exhibió una menor dispersión durante toda la simulación, entre 21 Å y 19 Å para los fosfolípidos de Z+ y entre -18 Å y -19 Å para los

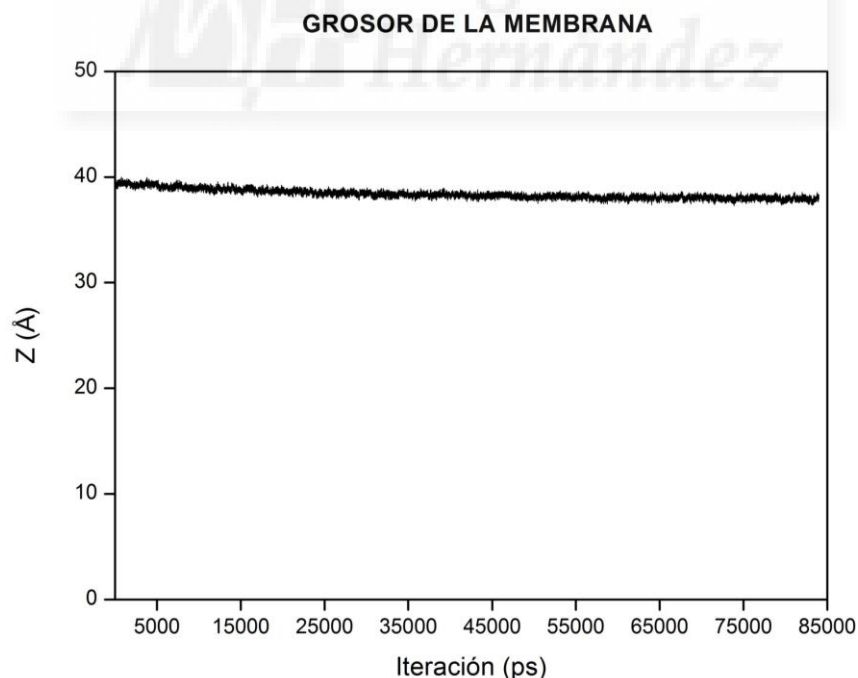


Figura 13. Representación gráfica de la evolución del grosor de la membrana en Angstroms, durante el proceso completo de simulación.

fosfolípidos de Z-, lo que no supone una variación importante. Al igual que el grosor de la membrana, su centro de masas permanece constante debido a que el sistema es estable. Cabe destacar que el péptido difunde libremente en agua y la unión no depende de su configuración inicial. Sin embargo, se puede observar que el centro de masas del péptido se localiza siempre sobre el nivel de los átomos de fosfato. Este dato podría implicar que el péptido no se inserta en la membrana pero se acerca a la superficie (Figura 14B). Puesto que LLP3 posee estructura en  $\alpha$ -hélice, cuando esta cambia su orientación, su centro de masas oscila considerablemente (10 – 15 Å) entorno a la distancia inicial a la bicapa, 15 Å. Inicialmente, el péptido fue situado paralelo a la superficie de la membrana, pero a lo largo de la simulación, LLP3 ha rotado de su posición inicial hasta encontrar una orientación más estable al cabo de unos 62 ns, en la que el extremo N-terminal está en contacto con los fosfolípidos de la hemicapa Z+.

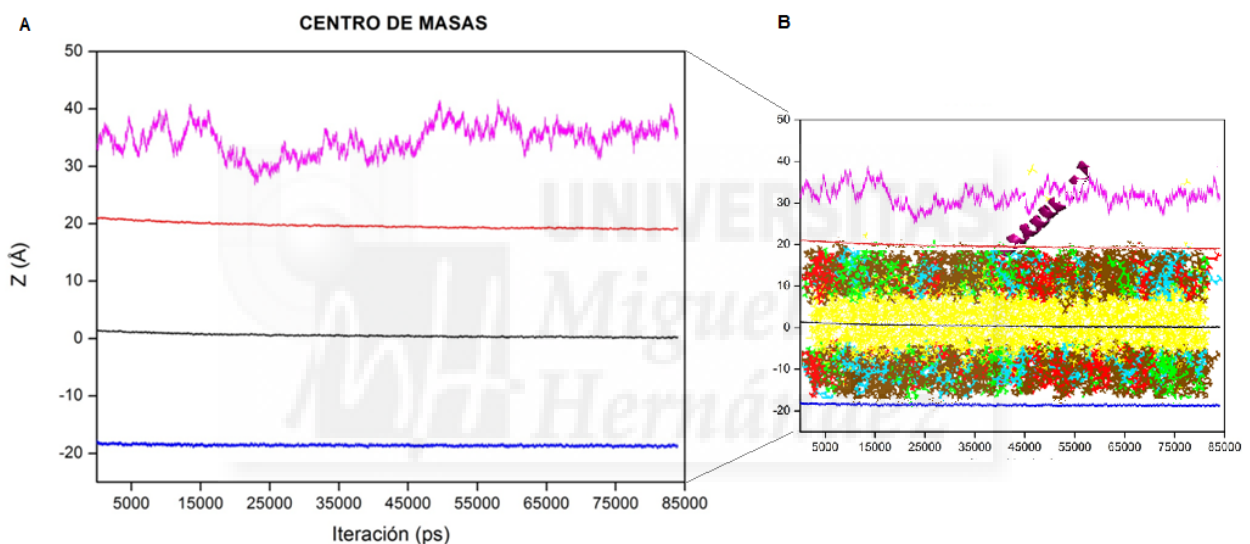


Figura 14. (A) Centro de masas de la membrana (rojo para los fosfolípidos de la hemicapa superior o Z+, azul para los fosfolípidos de la hemicapa inferior o Z-, negro para el promedio de ambos) y del péptido LLP3, en fucsia. (B) Superposición de una captura del sistema al gráfico del centro de masas, con el fin de ilustrar mejor a qué corresponde cada capa. DVPC (marrón), DVPG (rojo), DVPS (verde), DVPE (azul), DCLE (amarillo) y LLP3 (fucsia).

Por otro lado, se han calculado los perfiles de densidad de masa para todos los componentes del sistema (DVPC, DVPS, DVPG, DVPE, LLP3, el DCLE y el agua). La figura 15 muestra cómo se distribuyen los componentes durante los últimos 3 ns. Como se aprecia en la figura 15, DVPC, DVPG, DVPE y DVPS yacen entre las capas de agua y DCLE, mostrando dos partes diferenciadas, una correspondiente a las cadenas hidrocarbonadas y la otra a las cabezas polares, fosfato más colina, glicerol, etanolamina y serina, respectivamente. LLP3 se localiza inmerso en la capa de agua. Es interesante destacar que el péptido se encontraba inicialmente a 25 Å de distancia de la superficie de

la membrana y orientado paralelamente. No obstante, como se observa en la figura 15, LLP3 cambia su posición durante el proceso de simulación, aproximándose a la superficie de la bicapa, lo que indicaría la capacidad del péptido de unirse o interactuar con los fosfolípidos que forman la superficie de la membrana.

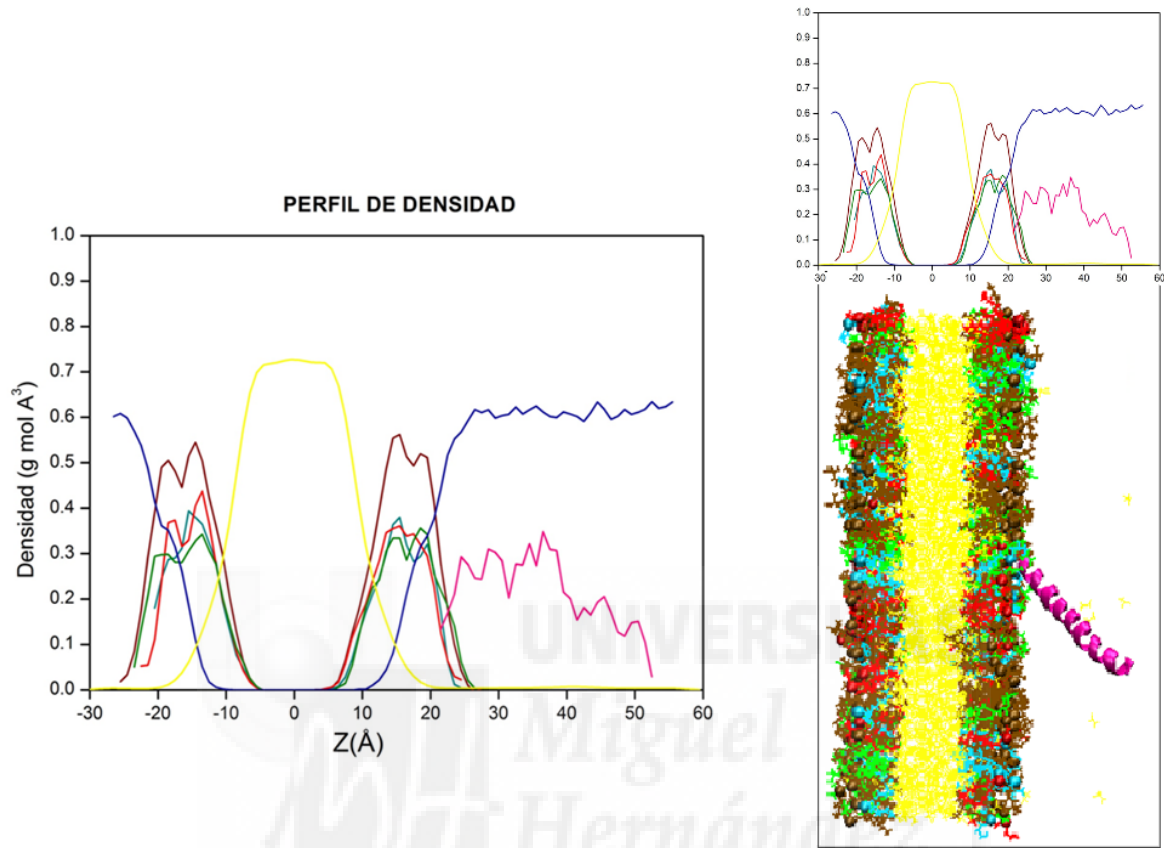


Figura 15. Perfil de densidades de todos los componentes del sistema simulado. Para una mejor visualización, se han multiplicado DVPC por 2, DVPS, DVPE, DVPG por 4 y LLP3 por 10. En la derecha, se adjunta una captura del sistema simulado, para ilustrar con mayor claridad la localización de los componentes: DVPC (marrón), DVPG (rojo), DVPS (verde), DVPE (azul), DCLE (amarillo) y LLP3 (fucsia).

En la figura 16 se representa la variación del RMSD para LLP3, que determina el cambio de posición de los átomos del péptido durante los 85 ns de simulación. Inicialmente y hasta los 45 ns, el RMSD sugiere inestabilidad en la proteína, debido a que se valoran todas las posiciones posibles. A partir de entonces, se establece una interacción entre LLP3 y la superficie de la bicapa. Dicha interacción estabiliza el péptido, disminuyendo los valores que toma el RMSD. Con el fin de obtener más información acerca de la interacción entre LLP3 y la membrana, se han calculado los enlaces de hidrógeno totales, específicos por lípido y específicos por aminoácido y por lípido (Figuras 17 a 20). En la figura 17 se representa gráficamente los enlaces de hidrógeno totales durante los 85 ns de simulación. Inicialmente y debido a la orientación inicial de LLP3, paralela a la superficie de la

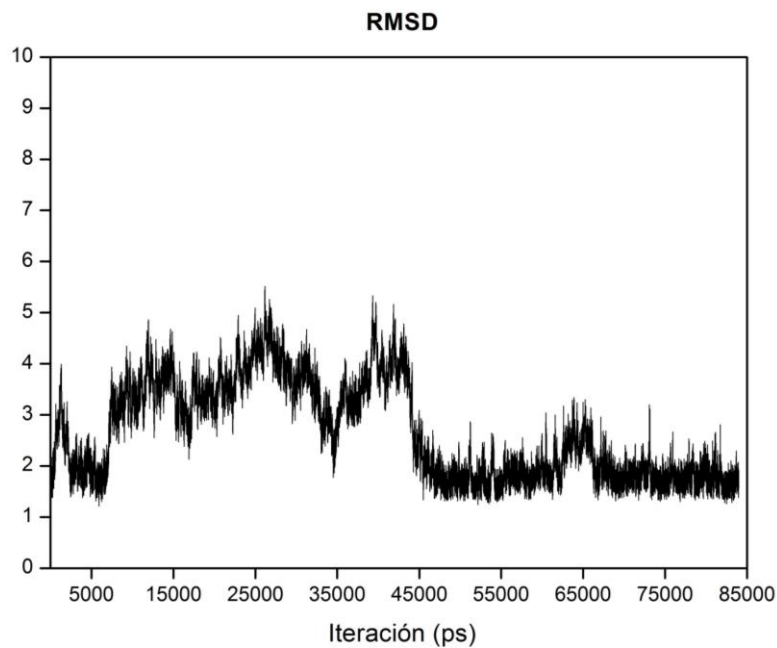


Figura 16. RMSD del péptido LLP3 durante todo el proceso de simulación.

membrana, el número total de enlaces de hidrógeno posibles es mayor que conforme avanza la simulación, ya que el péptido se inclina, acercándose a la membrana. De este modo, en la figura 17 se aprecia cómo, a partir de los 80 ns aproximadamente, el número máximo de enlaces de hidrógeno disminuye y tiende a permanecer constante, puesto que la interacción de LLP3 con la membrana se estabiliza. Puesto que la estructura global de LLP3 unido a la membrana permanece generalmente inalterada durante la simulación, las cadenas laterales del péptido son entonces las principales involucradas en la unión y por tanto responsables de su estabilidad.

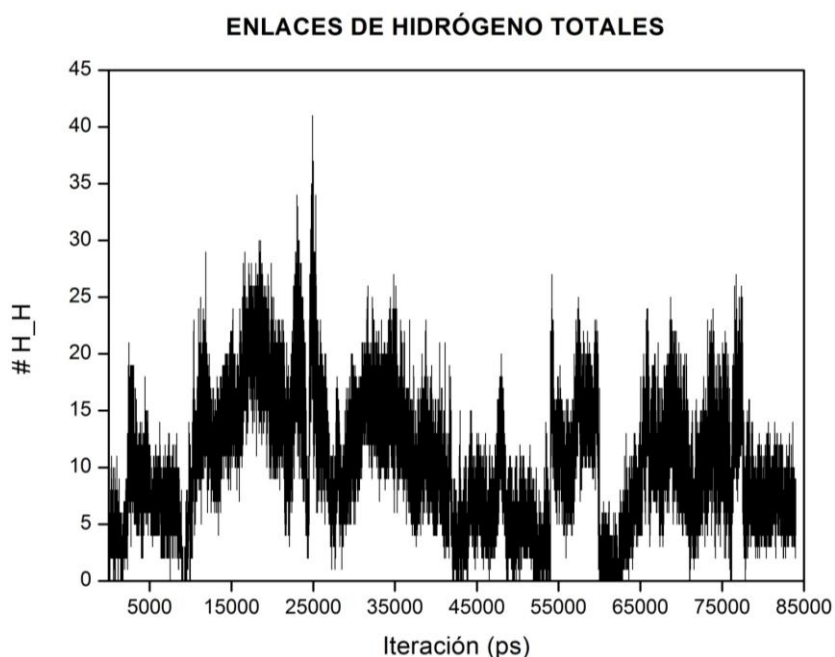


Figura 17. Enlaces de hidrógeno totales durante todo el proceso de simulación.

La figura 18 muestra los enlaces de hidrógeno específicos para cada lípido durante los últimos 6 ns de simulación. Así, se aprecia que el péptido tiene mayor interacción con los lípidos DVPG, DVPC y, en menor medida, con DVPS. Además no hay interacción con DVPE. Recordando la estructura de cada lípido a pH fisiológico (Figura 10), sabemos que DVPG tiene una carga negativa, proveniente del grupo fosfato. DVPS tiene dos cargas negativas por parte de los grupos fosfato y carboxilo, de las cuales una se neutraliza con la carga positiva del grupo amino. Por otro lado, DVPC y DVPE tienen carga neta cero debido a que se neutraliza la carga negativa de la cabeza polar con la carga positiva de la amina en ambos casos. Sabiendo que los grupos cargados tienen mayor dificultad para presentar este tipo de interacción, entonces es comprensible que sean los lípidos derivados de POPG los que forman prioritariamente enlaces de hidrógeno con LLP3.

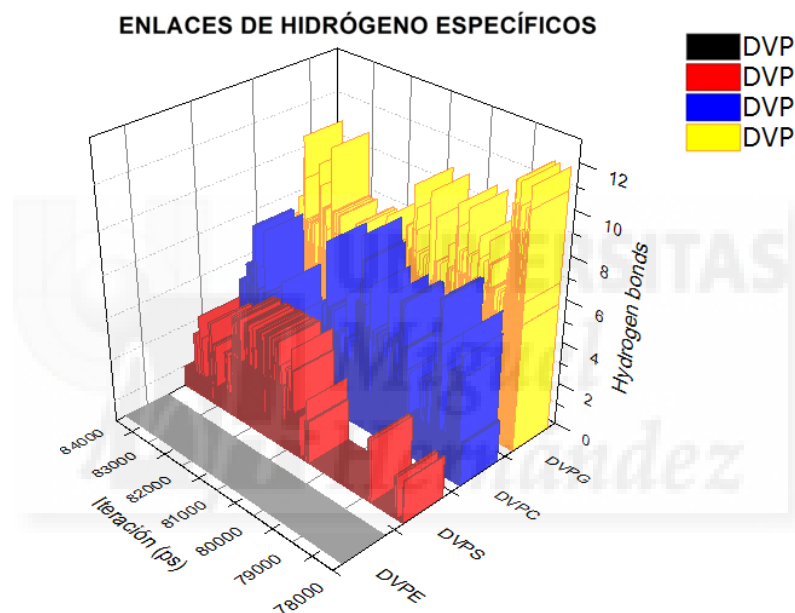


Figura 18. Enlaces de hidrógeno totales, específicos para cada lípido (últimos 6000 ps de simulación).

Como se aprecia en las figuras 19 y 20, los residuos que interaccionan mayoritariamente con los lípidos de la membrana son Arg5, Trp8, Leu2, Leu3, Gly4 y Arg6. De todos ellos, el principal es la Arg5, tal como se puede observar en las figuras 19 y 20. Estas observaciones acentúan el hecho de que los residuos de arginina, sobre todo la Arg5, son responsables de la unión del péptido, estabilizan la estructura helicoidal del péptido unido y la carga es el único contribuyente a la unión.

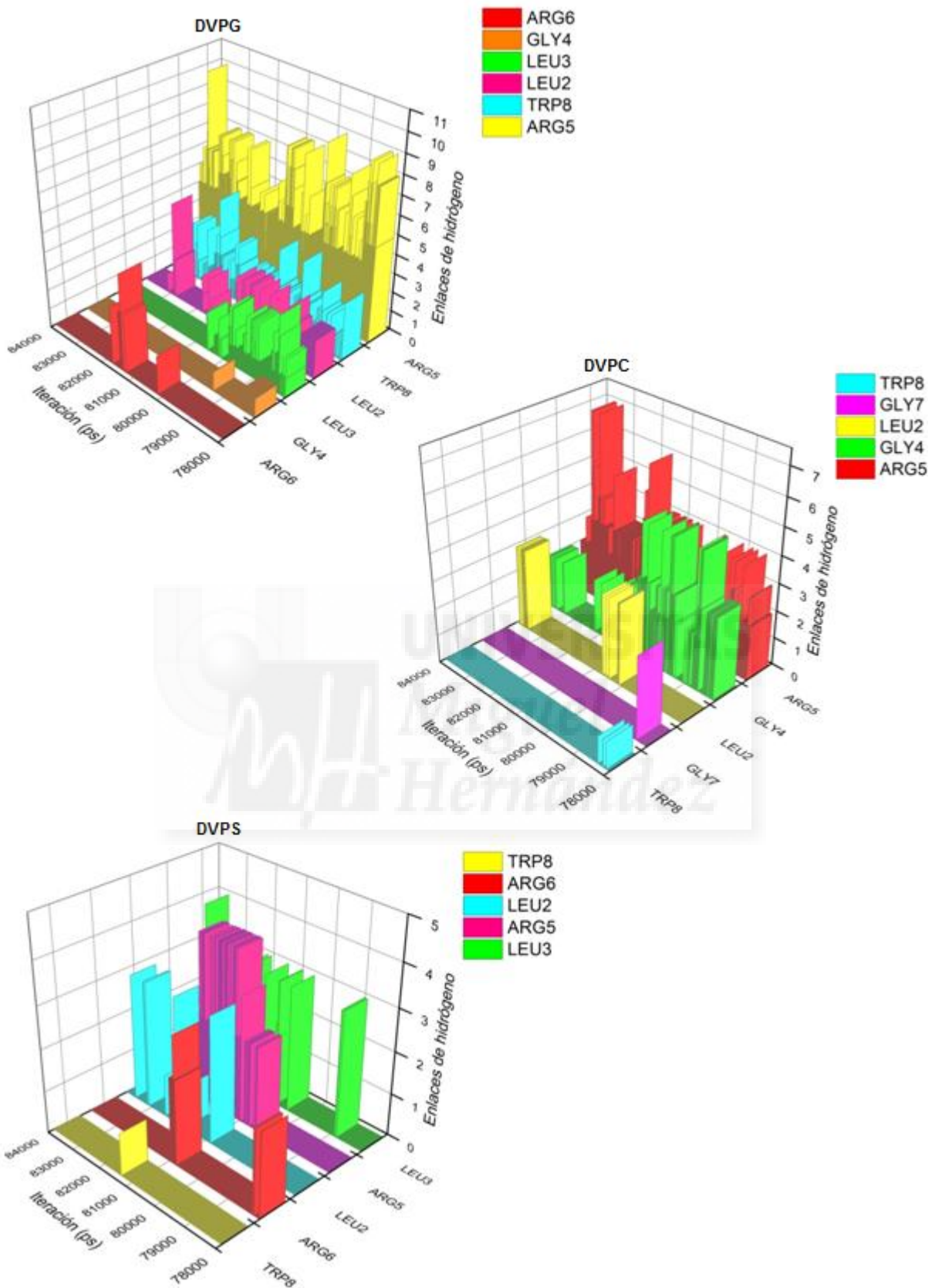


Figura 19. Número de enlaces de hidrógeno formados entre aminoácidos concretos de LLP3 y cada tipo de lípido, durante los últimos 6000 pasos de simulación (Nótese que no se dan enlaces de hidrógeno entre LLP3 y PE en este lapso de tiempo).

## ENLACES DE HIDRÓGENO

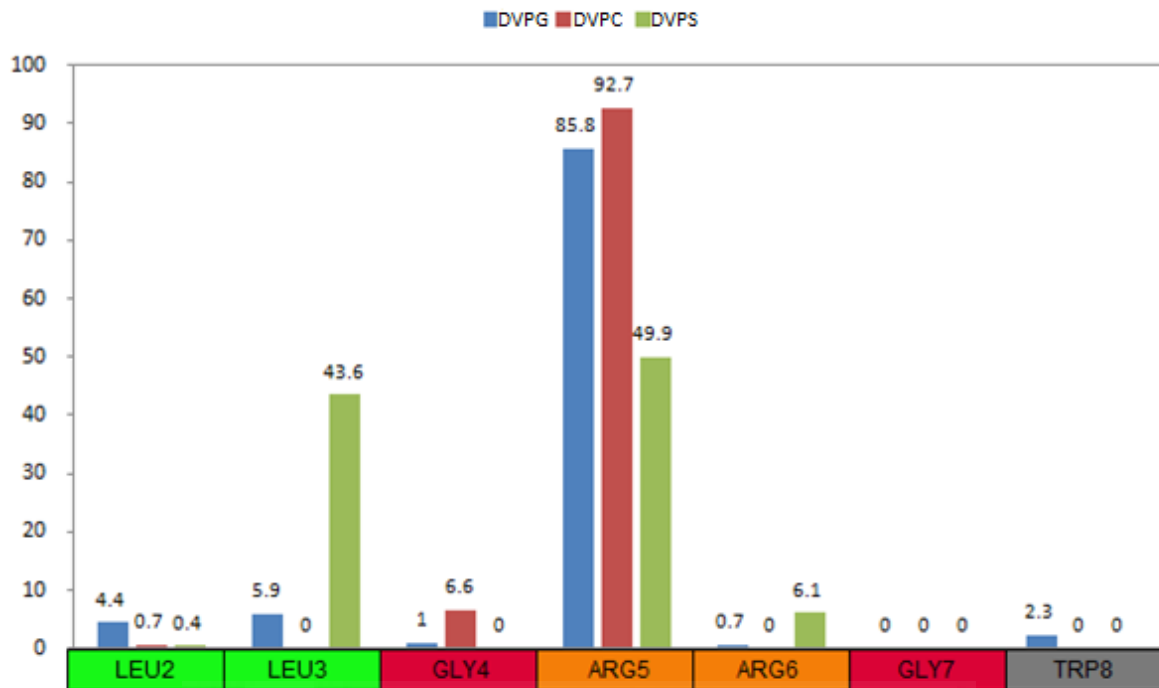


Figura 20. Gráfico que muestra la cantidad de enlaces de hidrógeno formados entre los primeros 7\* aminoácidos de secuencia de LLP3 y los lípidos de la membrana, en tanto por ciento.

\*Debido al sistema de numeración de VMD, el primer aminoácido de la cadena es nombrado como 2.



## 5. Conclusiones y proyección futura.

En este trabajo, se ha estudiado mediante dinámica molecular la interacción de un péptido derivado de la región LLP3 de la proteína gp41 del virus del SIDA con membranas modelo de alta movilidad. La proteína gp41 resulta fundamental en el proceso de entrada del virus VIH en la célula del huésped. Determinar el mecanismo de fusión y la obtención de inhibidores de dicha fusión son elementos clave en la lucha contra esta infección.

Durante la realización de este trabajo, se ha demostrado la existencia de una unión específica entre el péptido estudiado y la membrana simulada. Una vez que se establece dicha interacción, se mantiene constante a lo largo de toda la simulación. Se ha observado además que la interacción se realiza principalmente mediante el aminoácido Arg5. Esto significa que se trata de una unión por carga, descartando las de tipo hidrofóbico.

Puesto que se ha determinado que el aminoácido Arg5 es el principal responsable de la interacción entre este péptido y la membrana, la inhibición de dicha interacción mediante otro péptido o molécula orgánica posibilitaría su utilización como posible inhibidor de la fusión entre membranas mediada por la proteína gp41. Estos datos nos pueden ayudar en nuestra comprensión del mecanismo molecular de la fusión y la búsqueda de más y mejores inhibidores.

## 6. Referencias.

- [1] B. Apellániz, N. Huarte, E. Largo, J.L. Nieva. The three lives of viral fusion peptides. *Chemistry and Physics of Lipids*, 181 (2014) 40-55.
- [2] F.L. Cosset, D. Lavillete. Cell entry of enveloped viruses. *Advances in Genetics*, VOL. 73 (2011), 121-183.
- [3] M.R. Moreno. Elucidación de las regiones membranotrópicas de la glicoproteína gp41 del HIV-1 y su implicación en el proceso de fusión. Búsqueda de inhibidores de la entrada del HIV-1. 2007, Tesis Doctoral, Universidad Miguel Hernández de Elche.
- [4] J. Barroso. Mecanismos moleculares implicados en la infección por el VIH-1. 2012, Tesis Doctoral, Universidad de la Laguna, Tenerife.
- [5] S.G. Peisajovich, Y. Shai. Viral fusion proteins: multiple regions contribute to membrane fusion. *Biochimica et Biophysica Acta*, 1614 (2003) 122-129.
- [6] M.R. Moreno, M. Giudici, J. Villalaín. The membranotropic regions of the endo and ecto domains of HIV gp41 envelope glycoprotein. *Biochimica et Biophysica Acta* 1758 (2006) 111–123.
- [7] H.R. Marsden, I. Tomatsu, A. Kros. Model systems for membrane fusion. *Chemical Society Reviews* (2011), 40, 1572-1585
- [8] L. Banks, A. Carbone, H.E. Blum, E. Cesarman . IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. IARC Monographs, Volume 100 (B), pages 215-253
- [9] <https://www.aids.gov/hiv-aids-basics/hiv-aids-101/signs-and-symptoms/> Signos y síntomas del SIDA.
- [10] <http://www.who.int/gho/hiv/en/> Prevalencia del SIDA en todo el mundo. Estadísticas.
- [11] B. Berkhout, D. Eggink, R.W. Sanders. Is there a future for antiviral fusion inhibitors? *Current Opinion in Virology* (2012), 2:50-59.
- [12] M. Gochin, G. Zhou. Amphipathic properties of HIV-1 gp41 fusion inhibitors. *Current Topics in Medicinal Chemistry* (2011), 11 (24): 3022-3032.
- [13] L. Carrasco, J.M. Almendral. *Virus Patógenos*. 1ª ed.2006. Ed: hélice y fundación BBVA. Madrid. ISBN: 84-934106-0-8

- [14] D. Baltimore. Expression of Animal Virus Genomes. *Bacteriological Reviews* (1971), 35 (3): 235 - 241
- [15] S. Liu, N. Kondo, Y. Long, D. Xiao, A. Iwamoto, Z. Matsuda. Membrane topology analysis of HIV-1 envelope glycoprotein gp41. *Retrovirology* (2010), 7:100.
- [16] A.S. Khulmann, J.D. Steckbeck, T.J. Strugeon, J.K. Craigo, R.C. Montelaro. Unique functional properties of conserved arginine residues in the Lentivirus Lytic Peptide domains of the C-terminal tail of HIV-1 gp41. *The Journal of Biological Chemistry* (2014), VOL. 289, NO. 11, pp. 7630-7640.
- [17] J.D. Steckbeck, C. Sun, T.J. Sturgeon, R.C. Montelaro. Topology of the C-terminal tail of HIV-1 gp41: differential exposure of the Kennedy Epitope on cell and viral membranes. *PLoS ONE* 5(12): e15261
- [18] K. Salzedel, E.A. Berger. Cooperative subunit interactions within the oligomeric envelope glycoprotein of HIV-1: Functional complementation of specific defects in gp120 and gp41. *Proceedings of the National Academy of the United States of America* (2000), VOL.97, NO. 23. 12794–12799.
- [19] B.T. Korber, B.T. Foley, C.L. Kuiken, S.K. Pillai, J.G. Sodroski. Numbering positions in HIV relative to XHB2CG. *Human Retroviruses and AIDS*, 1998, páginas III-102-111.
- [20] M.R. Moreno, A.J. Pérez-Berná, J. Guillén, J. Villalaín. Biophysical characterization and membrane interaction of the most membranotropic region of the HIV-1 gp41 endodomain. *Biochimica et Biophysica Acta* 1778 (2008), 1298-1307.
- [21] J.D. Setckbeck, C. Sun, T.J. Sturgeon, R.C. Montelaro. Detailed topology mapping reveals substantial exposure of the “cytoplasmic” C-terminal tail (CTT) sequences in HIV-1 env proteins at the cell surface. *PLoS ONE* (2013), vol 8, issue 5, e65220.
- [22] Y. Kliger, Y. Shai. A leucine zipper-like sequence from the cytoplasmic tail of the HIV-1 envelope glycoprotein binds and perturbs lipid bilayers. *Biochemistry* (1997), 36, 5157-5169.
- [23] M. Hollier, N.J. Dimmock. The C-terminal tail of the gp41 transmembrane envelope glycoprotein of HIV-1 clades A, B, C, and D may exist in two conformations: An analysis of sequence, structure, and function. *Virology* 337 (2005) 284 – 296.
- [24] R. M. Epanand. Further insights into the properties of the HIV gp41 fusion domain: commentary on the article of A. L. Lai et al. *Journal of Molecular Biology*, Volume 418, Issues 1–2, 20 April 2012, Pages 1–2.

- [25] Bennett, W. F. D., & Tieleman, D. P. (2013). Computer simulations of lipid membrane domains. *Biochimica et Biophysica Acta*, 1828(8), 1765–76.
- [26] Ulmschneider, J. P., Andersson, M., & Ulmschneider, M. B. (2011). Determining peptide partitioning properties via computer simulation. *Journal of Membrane Biology*, 239(1-2), 15–26. 8
- [27] Shaikh, S. a., Li, J., Enkavi, G., Wen, P. C., Huang, Z., & Tajkhorshid, E. (2013). Visualizing functional motions of membrane transporters with molecular dynamics simulations. *Biochemistry*, 52(4), 569–587.
- [28] Li, J., Wen, P.-C., Moradi, M., & Tajkhorshid, E. (2015). Computational characterization of structural dynamics underlying function in active membrane transporters. *Current Opinion in Structural Biology*, 31, 96–105.
- [29] The PyMOL Molecular Graphics System, Version 1.7.4 Schrödinger, LLC.
- [30] E.L. Wu, X. Cheng, S. Jo, H. Rui, K.C. Song, E.M. Davila-Contreras, Y. Qi, J. Lee, V. Monje-Galvan, R.M. Venable, J.B. Klauda, W. Im, CHARMM-GUI Membrane Builder toward realistic biological membrane simulations, *J Comput Chem*, 35 (2014) 1997-2004 ] <http://www.charmm-gui.org>.
- [31] Humphrey, W., Dalke, A. and Schulten, K., "VMD - Visual Molecular Dynamics", *J. Molec. Graphics*, 1996, vol. 14, pp. 33-38. <http://www.ks.uiuc.edu/Research/vmd/>
- [32] S. Shnaper, K. Sackett, S.A. Gallo, B. Blumenthal, Y. Shai. The C- and N-terminal regions of glycoprotein 41 ectodomain fuse membranes enriched and not enriched with cholesterol, respectively. *The Journal of Biological Chemistry* (2004), 279, 18526-18534.
- [33] W. Weissenhorn, A. Hinz, Y. Gaudin. Virus membrane fusion. *FEBS Letters* (2007), VOL. 581, Issue 11, 2150–2155
- [34] M.J. Kozal, N. Shah, N. Shen, R. Yang, R. Fucini, T.C Merigan, D.D Richman, D. Morris, E. Hubbell, M. Chee, T.R. Gingeras. Extensive polymorphisms observed in HIV-1 clade B protease gene using high-density oligonucleotide arrays. *Nature Medicine* (1996) 2(7):753-9.
- [35] Contreras, L. M., Aranda, F. J., Gavilanes, F., González-Ros, J. M., & Villalaín, J. (2001). Structure and interaction with membrane model systems of a peptide derived from the major epitope region of HIV protein gp41: implications on viral fusion mechanism. *Biochemistry*, 40(10), 3196–207.

- [36] Moreno, M. R., Pascual, R., & Villalaín, J. (2004). Identification of membrane-active regions of the HIV-1 envelope glycoprotein gp41 using a 15-mer gp41-peptide scan. *Biochimica et Biophysica Acta - Biomembranes*, 1661, 97–105.
- [37] Pascual, R., Contreras, M., Fedorov, A., Prieto, M., & Villalaín, J. (2005). Interaction of a peptide derived from the N-heptad repeat region of gp41 Env ectodomain with model membranes. Modulation of phospholipid phase behavior. *Biochemistry*, 44(43), 14275–88.
- [38] Pascual, R., Moreno, M. R., & Villalaín, J. (2005). A peptide pertaining to the loop segment of human immunodeficiency virus gp41 binds and interacts with model biomembranes: implications for the fusion mechanism. *Journal of Virology*, 79(8), 5142–5152.
- [39] Moreno, M. R., Guillén, J., Pérez-Berná, A. J., Amorós, D., Gómez, A. I., Bernabeu, Á., & Villalaín, J. (2007). Characterization of the interaction of two peptides from the N terminus of the NHR domain of HIV-1 gp41 with phospholipid membranes. *Biochemistry*, 46(37), 10572–10584.
- [40] J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kale, K. Schulten, Scalable molecular dynamics with NAMD, *J Comput Chem*, 26 (2005) 1781-1802.
- [41] J.B. Klauda, R.M. Venable, J.A. Freites, J.W. O'Connor, D.J. Tobias, C. Mondragon-Ramirez, I. Vorobyov, A.D. Mackerell, Jr., R.W. Pastor, Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types, *J Phys Chem B*, 114 (2010) 7830-7843.
- [42] K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, A.D. Mackerell, Jr., CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields, *J Comput Chem*, 31 (2010) 671-690.
- [43] M. Ozu, H.A. Alvarez, A.N. McCarthy, J.R. Grigera, O. Chara, Molecular dynamics of water in the neighborhood of aquaporins, *Eur Biophys J*, 42 (2013) 223-239.
- [44] S.E. Feller, Zhang, Y., Pastor, R.W.J., Constant pressure molecular dynamics simulation: The Langevin piston method, *J. Chem. Phys.*, 103 (1995) 4613-4621.
- [45] G.J. Martyna, Tobias, D.J., Klein, M.L., Constant pressure molecular dynamics algorithms, *J. Chem. Phys.*, 101 (1994) 4177-4189.

- [46] M. Patra, M. Karttunen, M.T. Hyvonen, E. Falck, P. Lindqvist, I. Vattulainen, Molecular dynamics simulations of lipid bilayers: major artifacts due to truncating electrostatic interactions, *Biophys J*, 84 (2003) 3636-3645.
- [47] K. Murzyn, T. Rog, G. Jezierski, Y. Takaoka, M. Pasenkiewicz-Gierula, Effects of phospholipid unsaturation on the membrane/water interface: a molecular simulation study, *Biophys J*, 81 (2001) 170-183.
- [48] R. Guixa-Gonzalez, I. Rodriguez-Espigares, J.M. Ramirez-Angueta, P. Carrio-Gaspar, H. Martinez-Seara, T. Giorgino, J. Selent, MEMBPLUGIN: studying membrane complexity in VMD, *Bioinformatics*, 30 (2014) 1478-1480.
- [49] T. Giorgino, Computing 1-D atomic densities in macromolecular simulations: the Density Profile Tool for VMD, *Computer Physics Communication*, 185 (2014) 317-322.

