



UNIVERSIDAD NACIONAL DEL SUR

TESIS DOCTOR EN FÍSICA

Cálculos computacionales en macromoléculas:
Rol del eje de simetría quintuple en el Virus del Triatoma
(TrV). Comparación con otros virus icosaédricos.

Juan Francisco Viso

Bahía Blanca

Argentina

2018



UNIVERSIDAD NACIONAL DEL SUR

TESIS DOCTOR EN FÍSICA

Cálculos computacionales en macromoléculas:
Rol del eje de simetría quintuple en el Virus del Triatoma
(TrV). Comparación con otros virus icosaédricos.

Juan Francisco Viso

Bahía Blanca

Argentina

2018

Prefacio

Esta Tesis se presenta como parte de los requisitos para optar al grado Académico de Doctor en Física, de la Universidad Nacional del Sur y no ha sido presentada previamente para la obtención de otro título en esta Universidad u otra. La misma contiene los resultados obtenidos en investigaciones llevadas a cabo en el Departamento de Física de la Universidad Nacional del Sur y en el Instituto de Física del Sur (IFISUR), dependiente del Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) y de la Universidad Nacional del Sur, durante el período comprendido entre el 18 de septiembre del 2012 y el 27 de abril del 2018, bajo la dirección del Doctor Marcelo Daniel Costabel, Profesor Asociado de la Cátedra de Laboratorio I y la codirección del Doctor Diego Marcelo Guérin, Investigador de la Fundación de Biofísica Bizkaia (FBB).

7 de abril de 2017

Juan Francisco Viso

Departamento de Física

UNIVERSIDAD NACIONAL DEL SUR



UNIVERSIDAD NACIONAL DEL SUR
Secretaria General de Posgrado
y Educación Continua

La presente tesis ha sido aprobada el .../.../... , mereciendo la calificación de (.....)

Agradecimientos

Llegando al final de esta travesía que ha sido el Doctorado quiero agradecer a todos los que me ayudaron a lo largo del camino. Espero no olvidarme de ninguno y para los que me conocen sabrán que las palabras son pocas pero el sentimiento es grande.

Al departamento de Física de la UNS y al Instituto de Física del Sur por formarme y darme los medios para realizar este doctorado.

A mi director, **Marcelo**, por introducirme en el mundo de la investigación.

A mis compañeros de oficina, que me soportaron todo el doctorado. A **Fernando** que me enseñó todo lo que sabía y me ayudó en los primeros pasos de este viaje.

A **Néstor**, por sus charlas de ciencia, educación y otros temas.

A **Sebastian** y **Michellina**, que me ayudaron y formaron como docente. A todos los integrantes de *Curiosos*.

A **Ilan** y **Nadia** por transitar este camino juntos y ayudarnos mutuamente aunque sea solo prestando la oreja. A todos los que pasaron por la oficina de becarios.

A mis amigos de la universidad, **Duilio, Pau, Cris y Licho**, por su amistad y por aguantarme en todo este viaje.

A **Maju**, hicimos todo este camino juntos y no se si lo hubiera terminado sin su amistad.

A mis amigos de la infancia, **Rober, Seba, Gaspar, Fede y Sonia**, después de tanto tiempo siempre presentes.

Por último y más importante a **mi familia**. A **mis padres** por darme todo. A **mis hermanos**, a pesar de las peleas, siempre apoyándome.

A **Jobo**.

Resumen

Las cápsides virales tienen un rol fundamental en el proceso de infección viral ya que son las encargadas de censar el ambiente donde se encuentra el virus y también interactuar con la célula infectada. Debido a esto, es de interés comprender como es que desempeñan esta función.

Los virus icosaédricos, como los del género Picornavirales, poseen cápsides con gran simetría. Estas cápsides poseen varios ejes de simetría dobles, triples y quíntuples. Particularmente, se ha propuesto que la cavidad presente en el eje de simetría quíntuple podría ejercer un rol como canal de iones[1]. Sin embargo, esta hipótesis no ha sido corroborada.

Un ejemplo particular de estos virus icosaédricos es el virus del Triatoma (TrV). Este virus es miembro de la familia Dicistroviridae del género Picornavirales y resulta ser un patógeno viral del Triatoma infestans (vinchucas) por lo cual ha sido propuesto como potencial control biológico para estos insectos que son los vectores del Mal de Chagas, enfermedad endémica de América Latina[2].

El objetivo de esta tesis es lograr determinar, a partir de técnicas computacionales, el rol de la cavidad presente en el eje de simetría quíntuple en el virus del Triatoma. Particularmente dilucidar si se comporta como un canal, si es capaz de regular su apertura, bajo qué condiciones se encuentra abierto o cerrado y qué tipo de átomos o moléculas son capaces de atravesarlo. A su vez se busca correlacionar el rol de esta cavidad con el proceso de infección viral.

Con este fin, se utilizarán técnicas computacionales como la Dinámica Molecular para estudiar el comportamiento de la estructura de la cápside bajo distintas condiciones. A su vez se comparará la estructura de TrV con otros virus del género para poder inferir diferencias o similitudes en las funciones de las cavidades presentes en sus ejes de simetría.

Summary

Viral capsids play a fundamental role in the viral infection process since they are in charge of censoring the environment where the virus is found and also interacting with the infected cell. Thus, it is interesting to understand how they perform this function.

Icosahedral viruses, such as those of the genus Picornavirals, have highly symmetric capsids. These capsids have several double, triple and quintuple symmetry axes. In particular, it has been proposed that the cavity present in the five-fold symmetry axis could play a role as an ion channel [1]. However, this hypothesis has not been corroborated.

A particular example of these icosahedral viruses is the Triatoma virus (TrV). This virus is a member of the family Dicistroviridae of the genus Picornavirales and is a viral pathogen of *Triatoma infestans* (vinchucas and is therefore proposed as biological control for these insects that are the vectors of Chagas disease, an endemic disease in Latin America)[2].

The goal of this thesis is to determine, using computational techniques, the role of the cavity present in the fivefold symmetry axis in the Triatoma virus. Particularly, we expect to elucidate if it behaves as a channel, if it is able to regulate its opening, under what conditions it is open or closed and what kind of atoms or molecules are capable of crossing it. At the same time, we seek to correlate the role of this cavity with the viral infection process.

To this end, computational techniques such as Molecular Dynamics will be used to study the behavior of the capsid structure under different conditions. In addition, the structure of TrV will be compared to other viruses of the genus in order to infer differences or similarities in the functions of the cavities present in their symmetry axes.

Certifico que fueron incluidos los cambios y correcciones sugeridas por los jurados.

Firma del Director

Índice general

Resumen	I
Summary	II
Índice de figuras	VI
1. Introducción	1
1.1. Estructura de Proteínas	1
1.2. Canales Iónicos en Proteínas de Membrana	4
1.3. ¿Qué es un Virus?	8
1.3.1. Picornavirus	11
1.3.2. Dicistrovirus	14
1.4. Virus del Triatoma	15
1.5. Efecto Hidrofóbico	16
1.6. Objetivos	18
2. Teoría y Métodos	20
2.1. Dinámica Molecular	20
2.1.1. Aproximación de Born-Oppenheimer	21
2.1.2. Aproximación Clásica	22
2.1.3. Campo de Fuerzas	23
2.1.4. Condiciones de Borde	25
2.1.5. Búsqueda de Vecinos	25
2.1.6. Cálculo de las interacciones de Coulomb	26
2.1.7. Integración de las Ecuaciones de Movimiento	28
2.1.8. Restricciones en las Distancias	29

2.1.9. Termostato y Baróstato	31
2.1.10. Algoritmo de Actualización	35
2.2. Programa de Simulación: GROMACS	35
2.3. Herramientas de Análisis	36
2.4. Cálculo de Energía Libre	37
2.4.1. Potencial de Fuerza Media	39
2.4.2. Umbrella Sampling	40
3. Poro en el Eje de Simetría Quintuple de TrV	42
3.1. Puerta Hidrofóbica	43
3.1.1. Protocolo de la Dinámica Molecular	43
3.1.2. Resultados y Discusión	46
3.2. Energía libre de un Ion a través del Poro	51
3.2.1. Protocolo de muestreo avanzado	52
3.2.2. Resultados y Discusión	53
3.3. Comparación con Otros Virus Icosaédricos	56
3.3.1. Protocolo de la Dinámica Molecular	56
3.3.2. Resultados y Discusión	57
4. Conclusión	63
Bibliografía	66
Publicaciones Científicas	72

Índice de figuras

1.1. Estructura química de un aminoácido	1
1.2. Clasificación de los aminoácidos	2
1.3. Estructuras proteicas	3
1.4. Secuencia aminoacídica y disposición transmembrana de Glicoforina A.	5
1.5. Modelo tridimensional de la porina PdbID: 1A0S.	6
1.6. Estructura de los canales de membrana	7
1.7. Estructura del canal GLIC	8
1.8. Clasificación de los Virus	10
1.9. Estructura de Picornavirus	12
1.10. Estructura de Dicistrovirus	15
1.11. Poro de TrV	17
2.1. Esquematación de los potenciales enlazantes	24
2.2. Algoritmo LINCS	31
2.3. Algoritmo de Dinámica Molecular.	35
3.1. Estructura de la Cápside de TrV	42
3.2. Sistema de simulación	44
3.3. Velocidades de Simulación	46
3.4. Poro de TrV	47
3.5. RMSD y Radio de Giro de TrV	48
3.6. Estructura del Poro de TrV	49
3.7. Estructura del Poro de TrV mutado V3012S	50
3.8. Mapa de densidad electrónica de TrV	52
3.9. Estructura del Poro de TrV con la presencia de un ion Mg^{2+} en su interior.	54
3.10. Mapa de densidad de moléculas de agua en HRV16	59

3.11. Mapa de densidad de moléculas de agua en PoV	59
3.12. Mapa de densidad de moléculas de agua en HAV	60
3.13. Mapa de densidad de moléculas de agua en CrPV	61
3.14. Mapa de densidad de moléculas de agua en BPMV	61
3.15. Mapa de densidad de moléculas de agua en TRsV	62

Capítulo 1

Introducción

1.1. Estructura de Proteínas

El sistema de estudio en este trabajo son los virus, cuya estructura resulta de un ensamblaje multiproteico. Precisamente, las proteínas son componentes químicos de las sustancias orgánicas y son de vital importancia para los organismos ya que cumplen una gran variedad de funciones, tales como: enzimática, inmunológica, estructural, protectora entre otras.

Estructuralmente, las proteínas, son cadenas lineales cuyas unidades básicas son los aminoácidos [3]. Los aminoácidos son compuestos que contienen un grupo amino ($-NH_2$), un grupo ácido o carboxilo ($-COOH$) y un grupo (R) que varía en cada uno de los veinte aminoácidos que existen en la naturaleza (Figura 1.1).

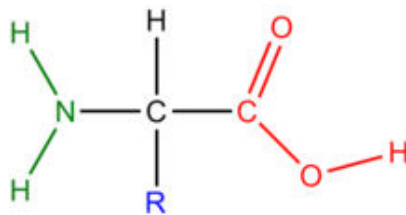


Figura 1.1: Estructura química de un aminoácido. En verde se representa al grupo amino, en rojo al carboxilo y en azul al grupo R.

Según su cadena lateral (R) se los divide en polares y no-polares. Los primeros a su vez se dividen según su carga en positivos, negativos y neutros (Figura 1.2). Según sus características, los distintos aminoácidos pueden interactuar de distintas formas. Por ejemplo, los aminoácidos cargados negativa y positivamente pueden interactuar entre sí

mediante puentes salinos. Así mismo, las cadenas laterales de los aminoácidos también pueden actuar como donores o aceptores de protones formando puentes hidrógeno, que si bien son débiles en solución acuosa, debido a la cantidad que se pueden formar son capaces de estabilizar la estructura proteica. Por otro lado, los aminoácidos No-polares tenderán a interactuar entre sí debido al efecto hidrofóbico. Otro tipo de enlace posible se observa entre pares de cisteínas que debido al sulfuro en el extremo de su cadena lateral pueden formar puentes disulfuro. Finalmente los aminoácidos aromáticos también pueden interactuar por apilamiento.

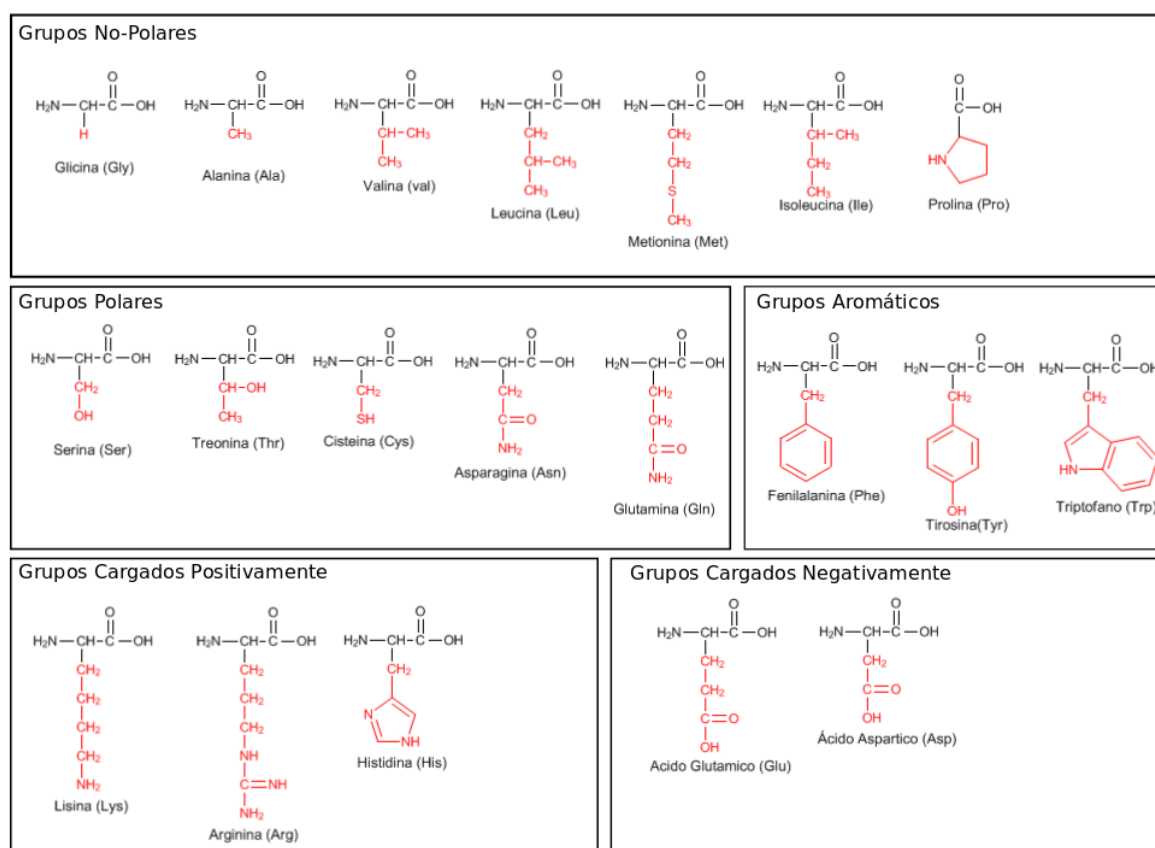


Figura 1.2: Clasificación de los 20 aminoácidos según la característica de su cadena lateral.

La estructura de la proteína se clasifica en cuatro niveles como se observa en la Figura 1.3. El primero se denomina Estructura Primaria y se refiere a la cadena lineal de aminoácidos (residuos) que componen a la proteína. Esta cadena se mantiene unida por los enlaces covalentes denominados puentes peptídicos que unen el grupo carboxilo de un aminoácido con el grupo amino de otro. Los extremos de la cadena polipeptídica se denominan carboxilo-terminal (C-terminal) y amino-terminal (N-terminal) dependiendo el grupo que haya quedado libre en cada extremo.

La estructura secundaria alude a la configuración espacial local de la proteína. Estas estructuras se forman a través de una red de puentes hidrógeno entre los átomos que forman el enlace peptídico. De esta forma se generan estructuras con una geometría definida que puede ser de dos formas distintas: en forma enrollada (α Hélice) o en forma Hoja Plegada β . A su vez estas estructuras se pueden combinar entre sí para formar estructuras más complejas denominadas super-estructura secundaria, algunos ejemplos incluyen la horquilla- β u horquilla- α , el barril- β , etc. En caso de no presentarse ninguna de las dos estructuras anteriores, la proteína presenta una estructura aleatoria denominada lazo o "random coil".

La estructura terciaria se refiere a la configuración tridimensional de las proteínas, es debida a los plegamientos que sufren las moléculas, y se dividen en fibrosas y globulares. El plegamiento hacia estas estructuras esta gobernado por la interacción hidrofóbica de dominios no-polares y luego estabilizada por puentes disulfuro entre las cisteínas, por puentes salinos entre los aminoácidos cargados y/o por puentes hidrógeno.

Por último la estructura cuaternaria no está presente en todas las proteínas y resulta de la combinación de dos o mas cadenas de proteínas, formando un multímero. Los monómeros se asocian entre sí mediante distintas interacciones como pueden ser el puente hidrógeno, puente salino y por interacción hidrofóbica.

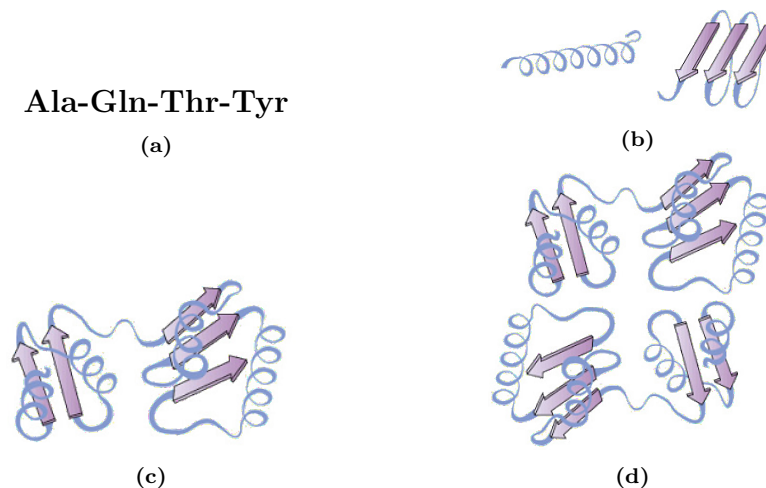


Figura 1.3: Estructuras proteicas: (a) primaria, (b) secundaria, (c) terciaria y (d) cuaternaria.

El conocimiento de las estructuras tridimensionales de las proteínas es de gran importancia debido a que el funcionamiento de las mismas está íntimamente ligado a su estructura. Por ejemplo, las capsides virales, que se utiliza como objeto de estudio en este trabajo, están compuestas por varias copias de una subunidad denominada protómero,

las cuales a su vez están compuestas por varias proteínas. Estos protómeros interactúan entre sí para formar grandes complejos moleculares que tienen como función proteger al material genético del virus y de reconocimiento a la célula que infectará.

Como se mencionó anteriormente, las cavidades presentes en los ejes quintuplos de las cápsides virales han sido propuestas como canales, en forma análoga a los canales de membrana. Precisamente, en el caso de las proteínas de membrana, su estructura está conformada por un arreglo de estructura secundaria con características hidrofóbicas en su exterior y características hidrofílicas en su interior. De esta manera la proteína tenderá a rodearse de los fosfolípidos presentes en la membrana y dejará en su interior un ambiente propicio para permitir la hidratación del poro y así permitir una conexión entre el interior de la célula y su exterior. Este tipo de estructuras se puede formar, por ejemplo, gracias a varias α -hélices paralelas las cuales generan una cavidad en su centro capaz de hidratarse y permitir el pasaje de iones. Precisamente, por esta última característica es que se realiza una descripción más exhaustiva de las mismas en el próximo capítulo.

1.2. Canales Iónicos en Proteínas de Membrana

Un grupo muy estudiado de proteínas son las proteínas que interactúan con las membranas celulares[4]. La estructura de estas proteínas varía según su ubicación en la membrana y su función. Se las clasifica en dos grandes categorías, integrales y periféricas, según sea su interacción con la membrana.

Las proteínas de membrana periféricas, o extrínsecas, no interactúan con el núcleo hidrofóbico de la membrana ya que no poseen región transmembrana; en cambio, interactúan con ella a través de interacciones directas con las cabezas polares de los fosfolípidos o interactúan directamente con otras proteínas de membrana intrínsecas. Un ejemplo de proteínas de membrana periféricas, que se encuentran del lado citosólico de la membrana celular, son las proteínas del citoesqueleto como la espectrina y la actina.

Otro ejemplo de proteínas extrínsecas son las fosfolipasas, que se asocian a los grupos polares de las membranas. Su función es hidrolizar varios enlaces en los grupos polares de los fosfolípidos con el fin de degradar células viejas o dañadas.

Por otro lado, las proteínas integrales, o intrínsecas, tienen uno o más segmentos de su cadena polipeptídica insertos en la membrana lipídica. Estas regiones son ricas en

aminoácidos no-polares, los cuales tienden a interactuar con las cadenas no-polares de los fosfolípidos que forman la membrana mediante interacción hidrofóbica. Estos dominios hidrofóbicos pueden atravesar parcial o completamente la membrana una o múltiples veces. Las regiones transmembrana están formadas por varias α -hélices u hojas- β .

Un ejemplo de proteínas intrínsecas es la Glicoforina. Esta proteína posee un dominio transmembrana formado por una α -hélice que atraviesa toda la membrana (Figura 1.4). Esta α -hélice está formada completamente por aminoácidos hidrofóbicos que interactúan con los fosfolípidos mientras que los grupos amino y carboxilo interactúan entre sí mediante puentes de Hidrógeno. A su vez posee aminoácidos cargados en los extremos de la hélice que impide que la misma se desplace fuera de la membrana, ya que estos interactúan con las cabezas polares de los fosfolípidos.

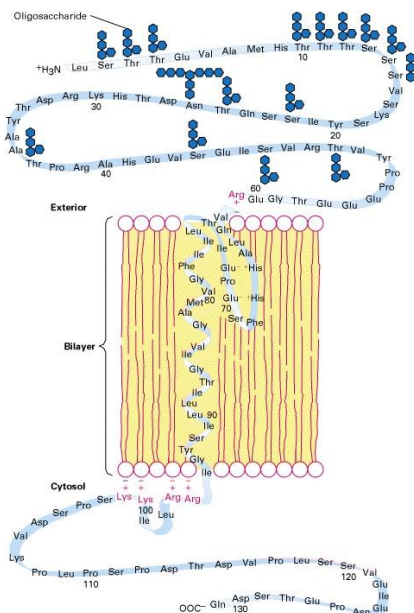


Figura 1.4: Secuencia aminoacídica y disposición transmembrana de Glicoforina A. (Figura sacada de Lodish et al, 2000).

También es posible que las proteínas de membrana presenten más de un dominio que atraviese a la misma. La glicoforina es un caso ya que en la membrana se agrupa en forma de dímeros. Otro ejemplo es el de las porinas, proteínas presentes en la membrana de bacterias gram-negativas como la *E. Coli*. La secuencia de aminoácidos de las porinas no presentan grandes regiones de aminoácidos no-polares que puedan insertarse en la membrana. La resolución por cristalografía de rayos X[5] mostró que las porinas son trímeros donde cada una de sus subunidades forman una estructura de barril con 16 hojas- β que a su vez forman un poro que atraviesa a la membrana (Figura 1.5). En esta estructu-

ra la mitad de cada aminoácido, que tiene características hidrofóbicas apunta hacia la membrana, mientras que el resto, que tiene características principalmente hidrofílicas, se orienta hacia el poro que es capaz de hidratarse.

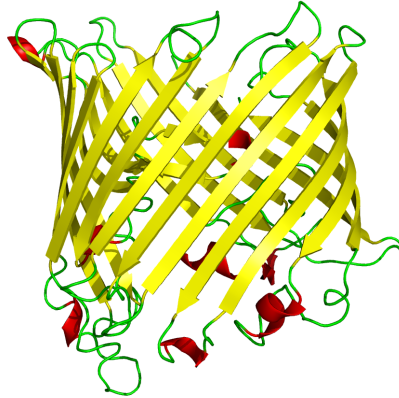


Figura 1.5: Modelo tridimensional de la porina PdbID: 1A0S.

Otro importante grupo de proteínas de membrana intrínsecas, son los canales iónicos[6]. Estas proteínas presentan estructuras macromoleculares complejas cuya función es el transporte de ciertas moléculas entre el exterior y el interior de las células. Estructuralmente forman un poro que atraviesa a la membrana celular, el cual se puede hidratar y permite el flujo de iones a través de la membrana celular. Su utilidad radica en la velocidad con la que los iones son capaces de atravesarlo ($\sim 10^6$ iones por segundo) y en su capacidad de selección. Esta selección, denominada *permeabilidad selectiva*, permite el paso de cierto tipo de iones por el canal, ya sea por su tamaño y/o su carga eléctrica. Otra característica importante de estos canales es que su funcionamiento no requiere el uso de energía metabólica ya que funcionan a favor del gradiente electroquímico, el cual depende de la concentración iónica y del potencial de membrana.

Estos canales tienen diversas funciones en la célula, tales como: definir los potenciales de reposo en las membranas celulares, permitir el flujo de iones Ca^{+2} , controlar el volumen celular, entre otras. Para lograr realizar estas funciones adecuadamente, los canales deben contar con un mecanismo de apertura y cierre que les permita controlar el flujo de iones. Uno de los mecanismos que poseen para tal fin son las puertas hidrofóbicas. Este mecanismo de apertura y cierre puede ser provocado por diversas señales externas como una diferencia de voltaje, por neurotransmisores, una deformación mecánica o algún otro estímulo químico. Gracias a este mecanismo, los canales son capaces de enviar señales a través de la membrana celular.

Las puertas hidrofóbicas logran controlar la apertura del canal gracias a que impiden el ingreso de moléculas de agua en una región del canal, lo que impide que los iones lo atraviesen ya que estos deben estar hidratados para lograr atravesar el canal [7]. Estas puertas, en los canales, se consiguen debido a la presencia de secuencias de aminoácidos con características no-polares o hidrofóbicas en alguna sección del canal. El efecto hidrofóbico impide el ingreso de moléculas de agua en regiones donde es estéricamente posible que las mismas ingresen. Luego debido a alguna señal externa se logra la apertura del canal y de esta manera se permite la hidratación del mismo y el pasaje de iones.

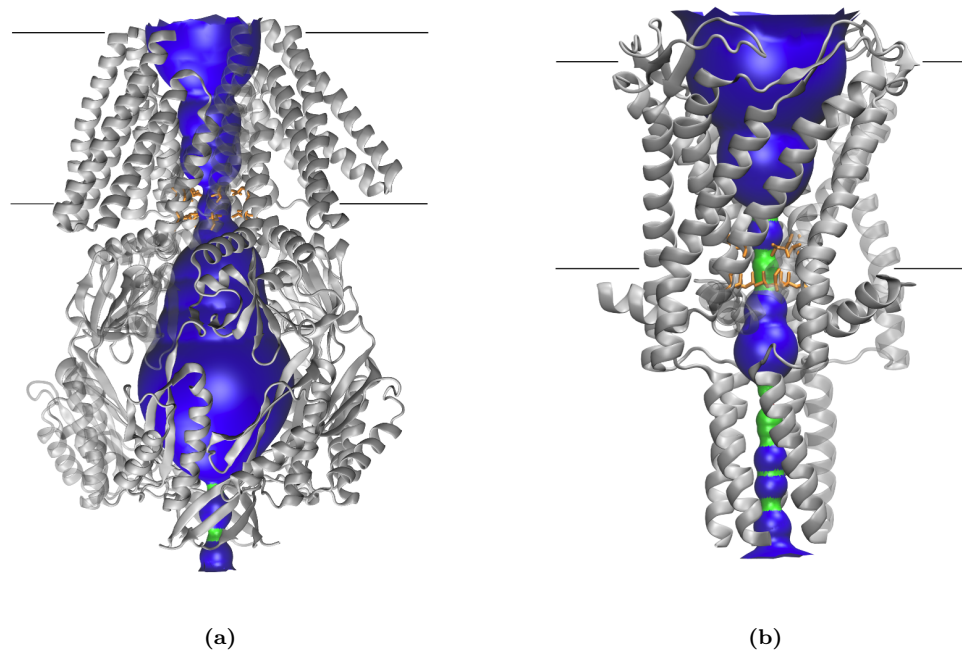


Figura 1.6: Estructura de los canales mecano-sensible de procariotas y la ubicación de sus puertas hidrofóbicas. (a) MscS (pdbID: 2OAU) y (b) MscL (pdbID: 2OAR). Las estructuras de los canales están representadas como *NewCartoon* en plateado. Los aminoácidos hidrofóbicos que forman la puerta hidrofóbica están representados en *licorice* en naranja. La superficie azul y verde representa el radio del canal. Las líneas horizontales marcan la posición de las membranas.

Este mecanismo de puertas hidrofóbicas se ha observado en diversos canales de membrana. Inicialmente se creyó que los canales mecano-sensibles de procariotas, que se abren frente a la tensión en la membrana, se encontraban abiertos ya que el canal tenía un diámetro de $\sim 5 \text{ \AA}$ [8]. Sin embargo utilizando Dinámica Molecular se observó que el mismo estaba cerrado hidrofóbicamente [9, 10]. En la figura 1.6 se puede observar las estructuras de dos de estos canales, resaltando la región donde se ubica su puerta hidrofóbica.

En canales pentaméricos regulados por ligando se ha demostrado que un cinturón de aminoácidos hidrofóbicos crean una barrera energética en contra del movimiento de las

aguas y de los iones de sodio a través del poro [11] (Figura 1.7). También se ha descubierto que para la proteína GLIC (Gloeobacter Ligand-gated Ion Channel) el costo de energía libre necesario para hidratar la puerta hidrofóbica es de ~ 11 kcal/mol, mientras que el necesario para que un ion atraviese la puerta es solo 4 kcal/mol mayor [7]. Esto indica que el mayor costo energético está dado por la hidratación de la puerta. Esto muestra que el proceso de hidratación del poro es clave para que un ion logre atravesar el canal. Este mecanismo, clave en la funcionalidad de estos complejos moleculares, es el que proponemos como partícipe del desensamblaje de una cápside viral.

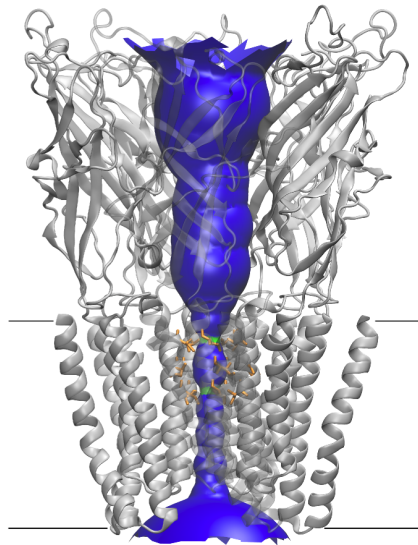


Figura 1.7: Estructura del canal pentamérico regulado por ligando GLIC (pdbID: 4NPQ). La estructura del canal está representada como *NewCartoon* en plateado. Los aminoácidos hidrofóbicos que forman la puerta hidrofóbica están representados en *licorice* en naranja. La superficie azul y verde representa el radio del canal. Las líneas horizontales marcan la posición de la membrana.

1.3. ¿Qué es un Virus?

Los virus[12] son pequeños parásitos intracelulares. Estos pequeños elementos genéticos móviles dependen de la compleja maquinaria metabólica de la célula huésped para poder propagarse. Su principal función es infectar a la célula y expresar su genoma para que la maquinaria celular pueda transcribirlo y traducirlo.

Están compuestos por un genoma de ADN o ARN y envueltos por una capa proteica. Esta capa proteica, que rodea y protege al material genético, se denomina cápside. La cápside cumple el doble rol de proteger el material genético del virus de cualquier agente que quiera dañarlo y de reconocer receptores específicos en la célula huésped a la cual se

unirá durante el proceso de infección. Algunos virus poseen, además, una bicapa lipídica derivada de la membrana de la célula huésped que es rica en glicoproteínas de membrana propias del virus. Toda la partícula viral, genoma con sus proteínas asociadas, cápside y, en algunos casos, la bicapa lipídica, se denomina Virión.

Los virus son clasificados según su morfología, su composición química, su método de replicación y su genoma. El comité Internacional de Taxonomía de Virus (ICTV por el acrónimo en inglés) los clasifica de manera similar a la clasificación de los organismos celulares, es decir, dividido en Orden, Familia, Género y Especie.

Por otro lado, la clasificación de Baltimore[13] los agrupa según el tipo de genoma y el método de replicación. En este caso, los virus de ARN, que componen el 70 % de los virus conocidos, pueden dividirse en cuatro grupos según la cantidad de cadenas y el sentido de la misma, que determina si el ARN puede funcionar como ARN mensajero(ARNm) o no. De esta forma se obtienen los siguientes grupos: ARN bicatenario (dsRNA, por sus siglas en inglés) cuyas cadenas son complementarias entre sí. Solo la cadena positiva que es capaz de actuar como ARNm es liberada durante la replicación. ARN monocatenario Positivo (+ssRNA) el cual puede funcionar como ARNm. ARN monocatenario Negativo (-ssRNA) que no puede funcionar como ARNm, por lo cual requiere primero ser convertido en positivo a través de una ARN polimerasa viral. Estos dos grupos son capaces de reproducirse en el citoplasma de la célula, ya que el ARNm puede acceder directamente a los ribosomas del huésped. ARN monocatenario retrotranscripto (ssRNA-RT), estos virus poseen una proteína denominada transcriptasa reversa (RT) la cual puede transcribir al ARN en ADN circular. Este ADN luego se une al ADN de la célula huésped para su posterior transcripción.

Los virus de ADN se dividen en tres grupos. ADN bicatenario (dsDNA), que tiene que acceder al núcleo de la célula para acceder a la ADN polimerasa durante el ciclo de reproducción celular. ADN monocatenario (ssDNA), estos virus deben formar una estructura capaz de replicarse, es decir una estructura bicatenaria. Esto lo logran gracias a la ADN polimerasa de la célula huésped. Por último, ADN bicatenario retrotranscripto (dsDNA-RT), estos virus poseen ADN con el cual generan ARNm y un ARN pregenómico el cual sirve como modelo para que la transcriptasa reversa replique el ADN. De esta manera se definen 7 grupos que se detallan en la figura 1.8.

Finalmente, según su morfología, los virus se pueden clasificar en dos estructuras.

Clasificación de los virus			
Grupo		Familia	Ejemplo
dsDNA	Virus Bicatenario de ADN	Adenoviridae Herpesviridae Poxviridae	Human Mastadenovirus C Human Herpevirus 1 Variola Virus
ssDNA	Virus Monocatenario de ADN	Parvoviridae	Primate Bacoparvovirus
dsRNA	Virus Bicatenario de ARN	Reoviridae	Rotavirus A
+ssRNA	Virus Monocatenario Positivo de ARN	Picornaviridae Togaviridae	Enterovirus Rhinovirus Cricket Paralysis Virus Triatoma Virus
-ssRNA	Virus Monocatenario Negativo de ARN	Orthomyxoviridae Rhabdoviridae	Influenza A Virus Alfalfa Dwarf Virus
ssRNA-RT	Virus Monocatenario Retrotranscrito de ARN	Retroviridae	Human Immunodeficiency Virus 1
dsDNA-RT	Virus Monocatenario Retrotranscrito de ADN	Hepadnaviridae	Hepatitis B Virus

Figura 1.8: Clasificación de los Virus según Baltimore.

Los virus helicoidales y los icosaédricos. Los primeros tienen simetría helicoidal donde la subunidad proteica o protómero se ensambla en un arreglo helicoidal que rodea el material genético. De esta manera se generan tubos rígidos o flexibles. Estas estructuras son clasificadas según su longitud, ancho, inclinación de la hélice y número de protómeros por vuelta de la hélice.

Por otro lado los virus icosaédricos forman cápsides virales compuestas por 20 caras triangulares formando un icosaédro. Estas estructuras presentan varios ejes de simetría dobles, triples y quintuples. En estos virus los protómeros se agrupan en arreglos oligoméricos denominados capsómeros. Luego, los capsómeros generan la cápside icosaédrica y según su arreglo permite la clasificación de las cápsides.

Las cápsides virales juegan un rol fundamental en el proceso de infección viral. Esto se debe a que las cápsides son las encargadas del reconocimiento e interacción de los virus con las células infectadas. Las cápsides inician la infección al interactuar con receptores presentes en la superficie externa de la membrana celular. Estos receptores pueden ser diversas moléculas, desde proteínas, carbohidratos y glicolípidos [14].

1.3.1. Picornavirus

Un orden particular de virus es el *Picornavirales*[15] perteneciente al grupo de virus +ssRNA. Estos virus reciben este nombre por ser virus pequeños (*pico*) y por tener RNA en su material genético. Este orden se divide a su vez en cinco familias: Picornavirus, que infectan a vertebrados, e Iflavirus, Dicistrovirus, Marnavirus, que infectan a invertebrados, y Secovirus, que infectan a plantas y algas.

Estos virus se caracterizan por tener su material genético rodeado por una cápside proteica con simetría icosaédrica de aproximadamente 30 nm de diámetro. Esta cápside tiene un rol crucial tanto en la protección del virus como en su interacción con el medio. Por esta razón es importante conocer cuál es la función y la estructura de las proteínas que forman al virus.

Particularmente para el caso de las cápsides icosaédricas observadas en los virus pertenecientes a la Familia Picornavirus, el virus posee cuatro proteínas (proteínas virales o VP) VP1, VP2, VP3 y VP4 que interactúan entre sí formando un protómero (Figura 1.9a y 1.9b), el cual es la subunidad simétrica que forma una cápside. Este protómero interactúa con otros cuatro protómeros para formar un pentámero. A su vez 12 copias del pentámero interactúan entre sí para formar una cápside icosaédrica cuasi T=3 (Figura 1.9c y 1.9d). Todas las proteínas tienen una estructura similar formada por un barril- β constituido por 8 laminas. Estas laminas están unidas por segmentos de *random coil*. Cada una de estas proteínas tiene un peso de 30 kDa aproximadamente.

Varias de estas cápsides poseen una estructura característica común que se denomina cañón[16] (Figure 1.9c y 1.9d). Esta estructura se encuentra alrededor del eje quíntuple y se hipotetiza que es el sitio de unión del receptor celular, debido a la dificultad que le presenta a los anticuerpos para unirse a él. Esto es de gran importancia porque marca el inicio del proceso de infección viral, el acoplamiento con las células. Esto fue confirmado para rinovirus[17, 18], poliovirus [19, 20, 21] y coxsacie A21 virus[22].

Otra característica de estas cápsides es la presencia de una cavidad hidrofóbica ubicada en el barril- β de VP1, al cual se le pueden unir pequeñas moléculas hidrofóbicas (~ 300 Da). Esta cavidad estabiliza la estructura del virus, razón por la cual es el objetivo de ciertos compuestos antivirales[23]. Por otro lado, se ha propuesto que una competencia entre el factor hidrofóbico capaz de unirse a esta cavidad y el receptor que se une en el cañón, son los responsables de iniciar el proceso de desmantelamiento de la cápside[24, 25].

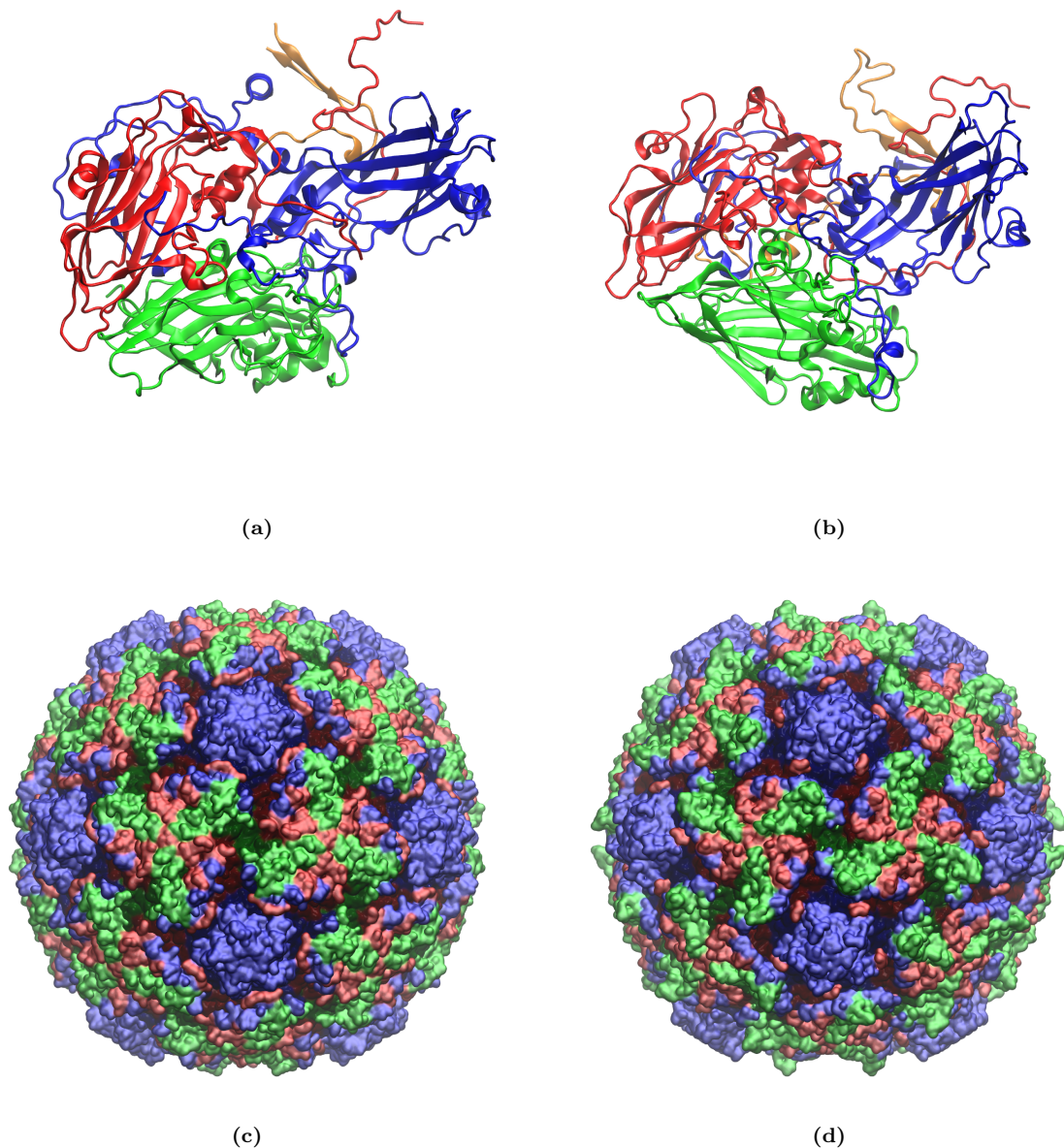


Figura 1.9: Estructura secundaria del protómero del Virus de Rhinovirus 16 (a) y del protómero de Poliovirus (b) coloreados por cadenas, azul VP1, verde VP2, rojo VP3 y naranja VP4. Superficie de las cápsides de Rhinovirus 16 (c) y Poliovirus (d), las regiones más oscuras corresponden a zonas más profundas en las cápsides denominadas cañón.

Este proceso de desensamblaje no ha sido explicado en su totalidad. Se hipotetiza que uno de los primeros eventos que ocurriría es la pérdida de VP4[26, 27]. Una de las posibilidades de que ocurra este fenómeno es que el extremo N-terminal de VP4 unido covalentemente a un grupo miristoil interactúe de algún modo con la membrana de la célula y abandone la cápside en conjunto con el genoma viral[28]. Un camino posible por el cual VP4 pueda salir de la cápside es a través del eje de simetría quintuple. Antes de que esto ocurra, la cápside debe desestabilizarse para aumentar considerablemente el

radio del poro presente en el eje quíntuple para poder dar lugar a que una molécula del tamaño de VP4 sea capaz de atravesarlo.

Este poro observado en el eje de simetría quíntuple podría actuar como un canal que conecta el exterior de la cápside con el interior[1]. Incluso, se observa la presencia de iones metálicos en diversas posiciones a lo largo del canal. En rhinovirus y coxsackievirus estos iones están cerca de la superficie exterior, mientras que en el enterovirus bovino están en el interior del canal. Como se mencionó anteriormente, se ha propuesto que este canal sea el camino por el cual se externaliza el ARN durante el proceso de infección, de manera similar a como se sabe que ocurre con el ADN del bacteriófago ϕ -X174[29].

Por otro lado, los extremos N-terminal de VP1 se encuentran cerca de la superficie del poro. En el caso de coxsackievirus A9[30], se observa la presencia del extremo N-terminal en esta región lo cual es consistente con la hipótesis de que el extremo N-terminal de VP1 sale a través del canal.

Los extremos N-terminal de VP3, también se encuentran cerca del extremo interior del canal. Es posible que el extremo N-terminal de VP4, en conjunto con el grupo miristoil, estabilice el barril- β de VP3, lo cual es esencial para el ensamblaje de la cápside[31].

Teniendo en cuenta todas estas características, se propone, como un posible mecanismo, que el primer evento del desensamblaje sea la interacción de los grupos miristoil de VP4 con la membrana de la célula huésped a través del canal presente en el eje de simetría quíntuple. Este evento llevaría a la pérdida de VP4 a través del canal y la exposición de los extremos N-terminal de VP1. De esta manera, se prepararía al canal para la salida del ARN[32]. Sin embargo, todo este proceso depende de que una molécula como VP4 sea capaz de atravesar el poro del eje quíntuple, el cual tiene un radio muy reducido.

Otra hipótesis del mecanismo de liberación del ARN, se basa en el estudio de las cápsides vacías (sin su genoma)[33, 34]. En estos trabajos se observan desplazamientos en los protómeros y un aumento del diámetro de la cápside. Esto provoca una ruptura en las interacciones entre pentámeros, particularmente en el eje de simetría doble. Debido a la presencia de estos agujeros en el eje de simetría doble se propone que la expulsión del ARN ocurre a través del mismo.

Por otro lado, se ha hipotetizado que el poro presente en el eje quíntuple de los virus icosaédricos podría llegar a actuar como un canal iónico que intercambie iones entre el exterior y el interior de la cápside[1]. Sin embargo, esta hipótesis no ha sido comprobada

rigurosamente, por lo cual actualmente se desconoce la funcionalidad de este poro.

1.3.2. Dicistrovirus

Los Dicistrovirus son otra familia dentro del orden Picornavirales que infectan artrópodos, principalmente insectos[35]. Como en el caso de los virus pertenecientes a la familia Picornavirus, los Dicistrovirus también forman una cápside icosaédrica a partir de 60 copias de sus proteínas virales. Sin embargo, no todos los virus de esta familia presentan la proteína VP4 en la estructura de la cápside. Tal es el caso del virus del *Triatoma* (TrV)[36] que en su estructura cristalográfica no se observa la presencia de VP4, lo cual fue atribuido a que la misma se encuentra en el interior de la cápside con una estructura desordenada.

Actualmente solo se resolvieron dos estructuras cristalográficas de Dicistrovirus, una de ellas es TrV (Figura 1.10a y 1.10c) y la otra Cricket Paralysis Virus (CrPV)[37] (Figura 1.10b y 1.10d). En ambas estructuras se pueden observar dos diferencias estructurales fundamentales con los Picornavirus, las cuales son la ausencia del cañón y la cavidad hidrofóbica.

Debido a estas diferencias entre las estructuras de las cápsides de Picornavirus y Dicistrovirus, se desconoce como los últimos interactúan con los receptores. Para el caso de *Drosophila C Virus*, se sabe que la entrada del virus a la célula se lleva a cabo por endocitosis mediada por clatrina[38]. Luego de la entrada del virus a la célula, el mismo se desensambla y libera su ARN en el citoplasma[39].

En TrV, se ha comparado la estructura de cápsides vacías, las cuales se forman luego de liberar su material genético, y cápsides en conjunto con su material genético[40]. Este estudio muestra que no hay un cambio sustancial entre ambas estructuras. Por otro lado, muestra la existencia de estructuras formadas por dos pentámeros enfrentados que se cree que se generan luego de la liberación del ARN. Todo esto podría indicar que el proceso de desensamblaje y la consiguiente liberación del genoma ocurre a través de una apertura parcial de la cápside. Particularmente, el ARN podría escapar a través de un poro abierto en las zonas periféricas de los pentámeros. También se ha observado que este virus es estable a pH ácidos pero bajo condiciones alcalina comienza el proceso de liberación de su ARN. Este fenómeno ocurre entre valores de pH de 8 a 9 [41].

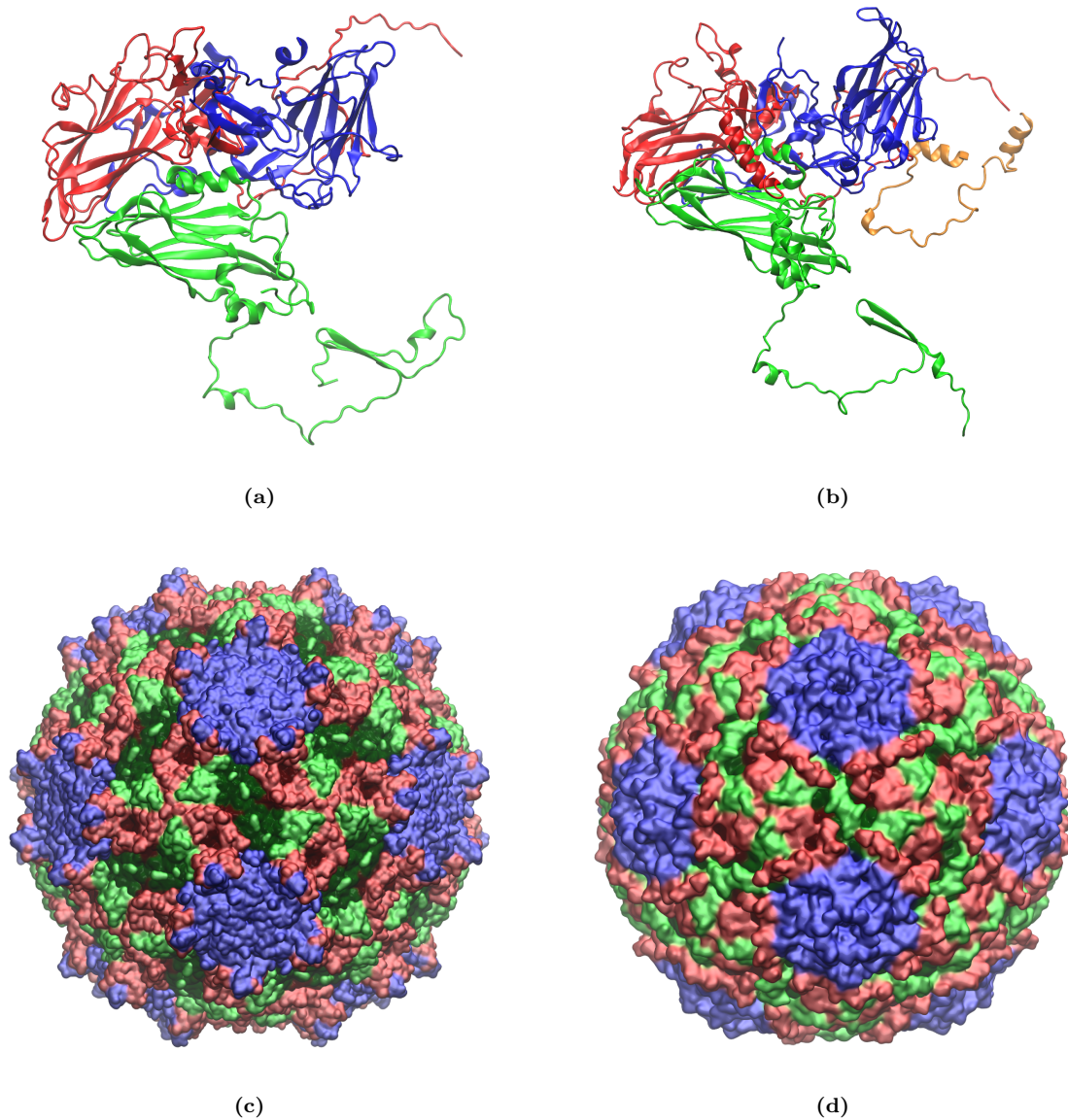


Figura 1.10: Estructura secundaria del protómero del Virus del Triatoma (a) y del protómero de Cricket Paralysis Virus (b) coloreados por cadenas, azul VP1, verde VP2, rojo VP3 y naranja VP4. Superficie de las cápsides del Virus de Triatoma (c) y de Cricket Paralysis Virus (d). En estas estructuras se observa la ausencia del cañón característico de los Picornavirus.

1.4. Virus del Triatoma

El virus del Triatoma o TrV es un virus del orden Picornavirales, de la familia Dicistrovirus. Esta familia se caracteriza por ser virus que infectan a los insectos. Particularmente TrV es un patógeno natural del *Triatoma Infestans*[2] mejor conocido como Vinchuca. Debido a esto, es de interés el estudio de TrV para ser utilizado como método de control biológico de estos insectos que son el principal vector de la enfermedad Trypanosomiasis

Americana, mejor conocida como Mal de Chagas.

Esta enfermedad es un riesgo a la salud en Latinoamérica, donde es una enfermedad endémica que afecta alrededor de 8 millones de personas, de las cuales el 30-40% desarrollan cardiomiopatía y/o megasíndromes digestivos[42]. El agente de la enfermedad es el protozoo *Trypanosoma Cruzi*, el cual infecta al insecto. Luego, se transmite al humano a través de la materia fecal del triatomino ya que mientras el mismo se alimenta de la sangre de la persona también defeca infectando a través de la herida.

La estructura de la cápside de TrV[36] (Figura 1.10c) está formada, como se mencionó anteriormente por 60 copias del protómero formado por VP1-2-3 (Figura 1.10a). Como todas las capsides icosaédricas, TrV posee varios ejes de simetría doble, triple y quintuple. Estos últimos, son particularmente interesantes ya que en este eje se encuentra una estructura que podría actuar como un poro que comunique el interior del virus con el exterior.

Este poro se encuentra en el centro de la subunidad pentamérica. Esta estructura es particularmente interesante ya que, como se menciona anteriormente, se ha observado que es capaz de mantener su estructura incluso si no está formando parte de la cápside al interactuar con otro pentámero[40]. Esto indica que el pentámero tiene un cierto grado de integridad estructural.

El poro mencionado está compuesto por seis anillos de aminoácidos, formados a su vez por las cinco copias simétricas de los protómeros que forman cada eje quintuple. De esta manera se genera una estructura tubular de aproximadamente 30 Å integrada por los seis anillos de glutamina 1128, valina 1166 y treonina 1167 pertenecientes a VP1, y glutamina 3014, valina 3012 y treonina 3010 pertenecientes a VP3 (Figura 1.11).

A diferencia de algunos canales iónicos de membrana, el poro presente en TrV no posee aminoácidos cargados. Sin embargo posee dos anillos de valinas que poseen propiedades no-polares. Particularmente, el poro presenta su región más angosta al nivel de las valinas 3012. Esta región podría actuar como un mecanismo de regulación de la apertura de la cavidad gobernado por el efecto hidrofóbico.

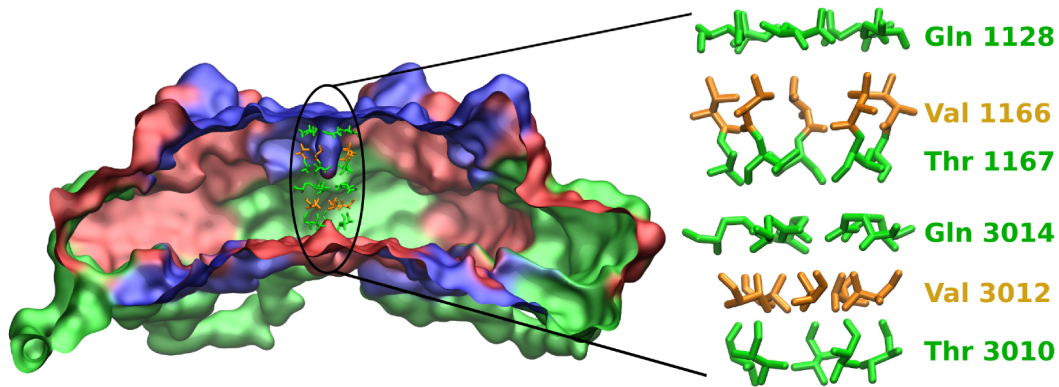


Figura 1.11: Aminoácidos que forman el poro de TrV. (Izquierda) Pentón de TrV visto de lado. Las proteínas están coloreadas azul para VP1, verde para VP2 y rojo para VP3. (Derecha) Aminoácidos que forman el poro presente en el eje quintuple. Los aminoácidos están coloreados en verde los polares y en naranja los no-polares.

1.5. Efecto Hidrofóbico

El denominado Efecto Hidrofóbico o Interacción Hidrofóbica[43], es el fenómeno responsable de mecanismos de regulación de la apertura en distintos canales proteicos presentes en las células. Este mecanismo de regulación se lo ha observado en nanotubos de carbono[44], regulando su apertura a través de la transición entre un estado hidratado a otro deshidratado. También se observa que una pequeña variación en el tamaño del poro puede llevar a la hidratación del mismo.

Se relaciona, a este efecto, con la repulsión que aparentan tener las moléculas no-polares al interactuar con el agua. Sin embargo, realmente las moléculas no-polares se ven atraídas por el agua, pero esta atracción no es tan fuerte como la atracción que sienten las moléculas de agua entre sí. Esta fuerte atracción se debe a la red de enlaces de hidrógeno que el agua es capaz de generar. Cada molécula es capaz de generar cuatro enlaces de hidrógeno con cuatro moléculas de agua vecinas. Esta fuerte atracción entre las moléculas de agua induce la segregación de las moléculas no-polares, lo cual provoca una atracción efectiva entre ellas.

Desde el punto de vista termodinámico [45], se relaciona al efecto hidrofóbico con el cambio de la energía libre de solvatación molar (G^0):

$$\Delta G^0 = \Delta h^0 - T\Delta s^0 \quad (1.1)$$

donde h^0 es la entalpía molar y s^0 es la entropía molar. En este proceso, el cambio en la

entalpía representan un cambio en la fuerza de interacción entre las moléculas mientras que los cambios en la entropía representan un cambio en el ordenamiento de las mismas. Debido a que el proceso de solvatación de moléculas no-polares en agua no es favorable, el cambio en la energía libre será positivo.

Cuando una pequeña molécula no-polar es disuelta en agua [46, 47], el agua se ordena a su alrededor formando puentes de hidrógeno con sus vecinos, sin que la red de puentes de hidrógeno se vea alterada fuertemente. Por lo tanto no se observará un gran cambio en la entalpía. Debido a esto, el aumento en la energía libre de solvatación, se debe a la disminución de la entropía causada por el ordenamiento de las moléculas de agua alrededor de la molécula no-polar. Por lo que el cambio en la energía libre de solvatación es entrópico y no entálpico.

Por otro lado, cuando el soluto no-polar es de mayor volumen, la interface entre el solvente y el soluto provoca que el agua no pueda formar todos los puentes de hidrógeno que es capaz de formar. Esto causa un aumento en la energía libre de solvatación por un mecanismo prioritariamente entálpico. Estas dos situaciones presentan una variación en el cambio de la energía libre de solvatación. De todos modos ambos mecanismos son responsables de la segregación de los solutos no-polares del agua, provocando así la hidrofobicidad.

Las características estructurales del poro de TrV dan sustento a la posibilidad de que el mismo funcione como una conexión entre el medio exterior y el interior de la cápside. Particularmente, la presencia de una región angosta rica en aminoácidos no-polares, dan sustento a la existencia de un mecanismo de regulación dominado por el efecto hidrofóbico. Esta característica es similar a la observada en proteínas de membrana.

1.6. Objetivos

Debido a que una de las tantas características de las proteínas es la de formar canales por donde se puedan transportar moléculas, y ya que no se ha observado la existencia de este tipo de estructuras en las cápsides virales, en la presente tesis se analizará la posibilidad de la existencia de dichos canales proteicos en las cápsides virales. Particularmente, se estudiará el rol de la cavidad presente en el eje de simetría quíntuple de la cápside icosaédrica del Virus del Triatoma perteneciente al genero Picornavirales. El

objetivo es descubrir si este virus posee canales proteicos que se encarguen del transporte de sustancias entre su interior y su exterior. Actualmente se desconoce la existencia de este tipo de mecanismos en los virus.

De acuerdo a lo presentado, se exponen los objetivos de la presente tesis.

Objetivo General

- Estudiar mediante simulación computacional por Dinámica Molecular la relación Estructura-Función de la cápside del Virus del Triatoma. En especial el rol del poro presente en el eje de simetría quintuple.

Objetivos Específicos

- Determinar si el poro de TrV tiene función de canal.
- Determinar si el poro posee algún mecanismo de regulación de su apertura.
- Determinar que tipo de moléculas son capaces de atravesarlo.
- Determinar similitudes y diferencias entre la estructura del eje quintuple con la estructura de otros virus pertenecientes al género.

Capítulo 2

Teoría y Métodos de Simulación Computacional

2.1. Dinámica Molecular

La Dinámica Molecular se basa en la resolución numérica de la ecuación de Newton para un sistema de miles o millones de partículas. Este método permite hallar información estructural y dinámica de arreglos proteicos que sería impracticable con otros métodos como la mecánica cuántica.

Al resolver la ecuación de Newton numéricamente se obtienen sucesivas configuraciones de la trayectoria de las partículas, es decir las posiciones y velocidades de todas las partículas para cada instante de tiempo. Luego si este muestreo es lo suficientemente extenso, el espacio de fases será muestreado adecuadamente, particularmente las configuraciones de menor energía. Teniendo en cuenta además la hipótesis ergódica, que dice que los promedios temporales son iguales a los promedios del ensamble:

$$\lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_{t=0}^{\tau} A(\mathbf{p}^N(t), \mathbf{r}^N(t)) dt = \iint d\mathbf{p}^N d\mathbf{r}^N A(\mathbf{p}^N, \mathbf{r}^N) \rho(\mathbf{p}^N, \mathbf{r}^N) \quad (2.1)$$

donde A es una propiedad física del sistema, \mathbf{p} y \mathbf{r} son los momentos y posiciones de las partículas y ρ la densidad de probabilidad. En conjunto esto permitirá obtener valores para distintas propiedades macroscópicas del sistema (energía libre, capacidad calorífica, etc.) a partir de los promedios temporales que se obtengan de la dinámica molecular.

Sin embargo, una limitación que se observa es la falta de detalle. A nivel clásico es

imposible obtener información de los electrones de los átomos, lo cual puede dar valiosa información acerca de los enlaces entre los mismos. Más aún, se podría preguntar que tan confiable es la estructura que se va a obtener resolviendo esta ecuación, válida para sistemas macroscópicos, siendo que se está empleando para el estudio de sistemas microscópicos. Para obtener información de la estructura molecular a nivel atómico, se debe resolver la ecuación de Schrödinger, con la cual es posible obtener información de la estructura electrónica de una molécula. Sin embargo, a medida que el sistema crece la resolución de esta ecuación incrementa en complejidad e incluso se vuelve imposible de resolver a menos que se utilicen computadoras con gran capacidad de cálculo. Incluso así, el tiempo que puede llevar resolver un sistema de cientos de átomos puede llegar a ser impracticable en escalas de tiempo humano.

Para poder subsanar el problema de la confiabilidad de los resultados, se utilizará una de las aproximaciones más importante de la Dinámica Molecular: la Aproximación de Born-Oppenheimer.

2.1.1. Aproximación de Born-Oppenheimer

El Hamiltoniano de un sistema de partículas viene dado por:

$$\mathcal{H} = T_N + T_e + V_{NN} + V_{ee} + V_{Ne} \quad (2.2)$$

Donde $T_N = \Sigma p_N^2/2M_N$ y $T_e = \Sigma p_i^2/2m$ son las energías cinéticas de los núcleos y los electrones de las partículas, respectivamente, y V_{NN}, V_{ee} y V_{Ne} las energías potenciales de interacción núcleo-núcleo, electrón-electrón y núcleo-electrón respectivamente.

La aproximación de Born-Oppenheimer[48, 49] propone que se puede considerar al lento movimiento de los núcleos atómicos independientemente del rápido movimiento de sus electrones. Esta aproximación separa a la función de onda $\Phi(\mathbf{R}, \mathbf{r})$ en una parte electrónica $\psi(\mathbf{R}, \mathbf{r})$ y en otra parte nuclear $\Psi(\mathbf{R})$, donde \mathbf{R} es la posición de los núcleos y \mathbf{r} la posición de los electrones. Luego, considerando únicamente la parte electrónica y manteniendo la posición del núcleo fija se obtiene:

$$[T_e + V_{ee} + V_{Ne}]\psi(\mathbf{R}, \mathbf{r}) = E_e(\mathbf{R})\psi(\mathbf{R}, \mathbf{r}) \quad (2.3)$$

Teniendo en cuenta la ecuación 2.3 y planteando la ecuación de Schrödinger para Φ se

obtiene:

$$\begin{aligned} \psi(\mathbf{R}, \mathbf{r})[T_N + V_{NN} + E_e]\Psi(\mathbf{R}) &= \psi(\mathbf{R}, \mathbf{r})E\Psi(\mathbf{R}) \\ &- \sum \frac{-\hbar^2}{2M_N} [\Psi(\mathbf{R})\nabla_{\mathbf{R}}^2\psi(\mathbf{R}, \mathbf{r}) + 2\nabla_{\mathbf{R}}\Psi(\mathbf{R})\nabla_{\mathbf{R}}\psi(\mathbf{R}, \mathbf{r})] \end{aligned} \quad (2.4)$$

Multiplicando esta ecuación por $\Phi(\mathbf{R}, \mathbf{r})^*$ e ignorando el último término[49] se llega a la ecuación de Born-Oppenheimer:

$$[T_N + \epsilon(\mathbf{R})]\Psi(\mathbf{R}) = E\Psi(\mathbf{R}) \quad (2.5)$$

Donde $\epsilon(\mathbf{R}) = V_{NN}(\mathbf{R}) + E_e(\mathbf{R})$. Esta es la ecuación de Schrödinger para los núcleos que tiene en cuenta la energía de interacción con sus electrones.

2.1.2. Aproximación Clásica

La aproximación de Born-Oppenheimer dice que podemos tratar al núcleo bajo un potencial efectivo dado por:

$$V = V_{NN}(\mathbf{R}) + E_e(\mathbf{R}) \quad (2.6)$$

Luego, se puede aproximar a la dinámica del sistema de partículas resolviendo la ecuación de Newton para cada uno de los núcleos:

$$\mathbf{F}_i = m_i\mathbf{a}_i \quad (2.7)$$

Donde \mathbf{a}_i es la aceleración de la partícula i -ésima, m_i su masa y \mathbf{F}_i la fuerza a la que está sometida. Esta fuerza se puede calcular a través del gradiente de una función potencial V que dependerá de la posición de todas las N partículas del sistema.

$$\mathbf{F}_i = -\nabla_i V(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) \quad (2.8)$$

La forma funcional del potencial y todos los parámetros que lo definen se lo denomina “Campo de Fuerza”.

2.1.3. Campo de Fuerzas

El campo de fuerzas es una de las herramientas más importantes de la Dinámica Molecular ya que posee toda la información sobre la interacción entre las distintas partículas del sistema. Debido al número de átomos interactuantes en los sistemas estudiados, los potenciales utilizados son funciones simples (como la ley de Hooke) para poder acelerar el tiempo de cálculo. Además, a pesar de la sencillez del modelo, la precisión lograda con los campos de fuerza puede llegar a ser tan buena como algunos cálculos de la mecánica cuántica. Aún así, siempre se debe tener en cuenta que este método no podrá dar información concerniente a la distribución electrónica de una molécula.

El campo de fuerza se divide en dos tipos de interacciones, interacciones enlazantes y no-enlazantes. Las primeras modelan los enlaces entre átomos vecinos y las segundas las interacciones coulombicas y de Van der Waals entre átomos más alejados. De esta manera el campo de fuerza toma la forma:

$$V = \sum_{\text{enlaces}} \frac{k_{b,i}}{2} (l_i - l_{i0})^2 + \sum_{\text{angulos}} \frac{k_{\theta,i}}{2} (\theta_i - \theta_{i,0})^2 + \sum_{\text{torsiones}} \frac{\nu_n}{2} (1 + \cos(n\omega - \gamma)) + \sum_{i=1} \sum_{j=i+1} (4\epsilon [(\frac{\sigma_{ij}}{r_{ij}})^{12} - (\frac{\sigma_{ij}}{r_{ij}})^6] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}) \quad (2.9)$$

Los tres primeros términos en la ecuación anterior son los términos enlazantes. El primero tiene en cuenta el estiramiento del enlace y esta modelado por un potencial elástico, de esta manera el enlace oscilara en torno a una longitud natural de enlace. El segundo término, también modelado como un oscilador, representa las oscilaciones en el ángulo formado por tres partículas. Por último, el tercer termino representa los cambios de energía cuando un enlace rota sobre si mismo (Estos enlaces están representados esquemáticamente en la figura 2.1).

Los últimos términos de la ecuación 2.9 representan las interacciones no-enlazantes entre pares de átomos que encuentren distanciados mas allá de tres enlaces o en otra molécula. Para la interacción electrostática se utiliza el potencial de Coulomb y para la de Van der Waals el potencial de Lennard-Jones.

Todas las constantes que aparecen en la ecuación 2.9 (constantes elásticas, posiciones de equilibrio, constantes de Van der Waals, etc) son los parámetros definidos en el Campo de Fuerza. Estos parámetros son definidos de manera tal que el modelo logre reproducir

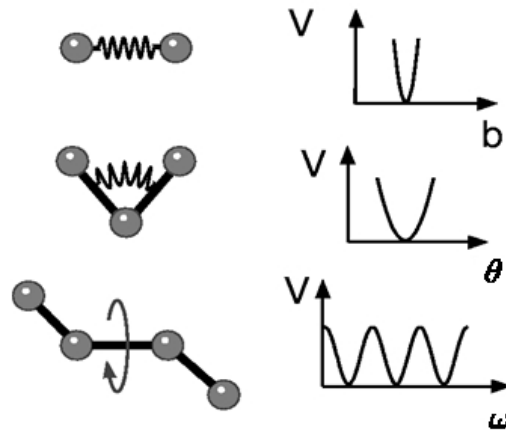


Figura 2.1: Esquemización de los potenciales enlazantes. De arriba a abajo: potencial de enlace, de ángulo y de torsión.

propiedades estructurales o de interés. El hecho que algunos campos de fuerza no logren reproducir con exactitud alguna característica en particular, no es un fallo, sino que se debe ser precavidos en la elección del mismo dependiendo de la propiedad que se desee estudiar.

Los distintos campos de fuerza existentes se los pueden clasificar según el nivel de detalle que utilizan para representar todos los átomos del sistema. Hay campos de fuerza que poseen parámetros para todos los átomos (*all-atom*), los campos de fuerza AMBER[50] y OPLS[51] son ejemplos de estos. Otros modelan un grupo de átomos en conjunto, por ejemplo, modelan al carbono de las cadenas carbonadas en conjunto con sus hidrógenos. Este tipo de campo de fuerzas se lo denomina de átomo unido o *united-atom*. Un ejemplo de este tipo de campos de fuerza es el GROMOS[52]. Esta aproximación puede ser llevada más lejos modelando grupos más grandes de átomos con una sola partícula que posee las propiedades del conjunto, se los denomina de grano grueso o *coarse-grain*. Ejemplos de estos son el MARTINI[53] y el SIRAH[54].

El objetivo de reducir el nivel de detalle en los campos de fuerza de grano grueso es disminuir el número de partículas y así aumentar la velocidad de simulación. Una opción intermedia es utilizar un enfoque híbrido, es decir, utilizar dos campos de fuerza en simultaneo, uno de detalle atómico y otro de grano grueso. De esta manera se simulan con gran detalle las zonas de interés y con menor detalle los alrededores.

2.1.4. Condiciones de Borde

Cuando se desea simular un sistema como una proteína, luego de definir los campos de fuerza a utilizar y antes de agregar las moléculas de agua que hidratarán a la proteína, se debe elegir la forma y tamaño de la caja en la que estará confinado el sistema. Luego esta caja será rellena con las moléculas de agua. Estas cajas de simulación pueden variar en forma y tamaño dependiendo el sistema que se desee simular. Particularmente uno de los criterios de elección es elegir una caja de manera que se minimice la cantidad de moléculas de agua que son necesarias para llenar a la misma, y de esta manera disminuir el número de partículas a simular. Entre los tipos de caja posibles se encuentran la cúbica, triclinica, octaedro truncado y dodecaedro rómbico.

Un problema que surge al momento de simular al sistema dentro de la caja es el tratamiento de los bordes. Las partículas en los extremos de la caja se encontrarán rodeadas por un lado por otras partículas pero por otro lado por vacío. Para evitar este problema se utilizan condiciones de borde periódicas. Es decir, que las partículas que atraviesen una de las caras de la caja, aparecerán por la cara opuesta. De esta manera se simula un sistema periódico con una simetría similar a la cristalográfica.

Teniendo en cuenta estas condiciones, al momento de elegir el tamaño de la caja se debe tener cuidado de no elegir una caja muy pequeña y que la proteína empiece a interactuar con su propia imagen periódica. Para evitar esto hay que tener en cuenta las distancias de corte de las interacciones no-enlazantes y el tamaño de la caja de simulación.

2.1.5. Búsqueda de Vecinos

Como se mencionó anteriormente, los campos de fuerzas se dividen en dos interacciones, las enlazantes y las no-enlazantes. En el caso de las segundas, que tienen en cuenta las interacciones de Van der Waals y Coulomb, las interacciones se dan entre todos los pares de átomos. Esto puede ser costoso computacionalmente para sistemas muy grandes. Por otro lado, muchas de estas interacciones son mínimas cuando las partículas se encuentran alejadas una de otras debido a que las fuerzas decaen como $1/r^2$, $1/r^7$, $1/r^{13}$, etc. Para acelerar este proceso se define una distancia hasta la cual se considerará efectiva la interacción entre pares de partículas, es decir, una distancia de corte (o cut-off).

Al definir un cut-off para las interacciones no-enlazantes, surge el problema de en-

contrar las partículas que se encuentran a esta distancia de corte de la partícula a la cual se desea calcular la fuerza. Para ello se definen varios esquemas para encontrar los “vecinos” de una partícula. Algunos de ellos son Group y Verlet[55].

Ambos esquemas, encuentran los grupos de partículas que se hallan a la distancia del cut-off y generan la lista de vecinos de la cual luego se calculará el potencial. Esta lista se redefinirá cada una cierta cantidad de pasos de integración. Durante los pasos donde la lista de vecinos no se redefine, se corre el riesgo de que partículas entren o salgan de la región definida por el cut-off, introduciendo así errores en el cálculo. El esquema de verlet resuelve este problema definiendo una distancia de cut-off para la cual se buscará primeros vecinos y otra distancia menor para la cual se calculará el potencial. De esta manera queda definido un cascarón donde hay partículas que se encuentran en la lista de primeros vecinos pero no se tienen en cuenta para el cálculo del potencial, pero si estas partículas se acercan a la segunda distancia del cut-off comenzarán a tenerse en cuenta y viceversa.

2.1.6. Cálculo de las interacciones de Coulomb

La interacción coulombica, como se muestra en la ecuación 2.9, viene dada por:

$$V_C = \sum_{i=1} \sum_{j=i+1} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \quad (2.10)$$

Debido a las características de este potencial y al hecho de usar distancias de corte para truncar el alcance de los mismos, se han desarrollado varios métodos para calcular las sumatorias. Entre los métodos desarrollados para aumentar la velocidad y la precisión de los cálculos se encuentran el Campo de Reacción, Sumatorias de Ewald, Particle Mesh Ewald (PME), entre otros.

Considerando que el sistema se encuentra dentro de una caja de simulación cúbica de lado L y si se tienen en cuenta las imágenes periódicas del sistema, la interacción de Coulomb, por ser de largo alcance, debería tener en cuenta también la interacción de estas imágenes. Teniendo esto en consideración la ecuación 2.10 quedaría:

$$V_C = \frac{f}{2} \sum_{n \in \mathbb{Z}^3} \sum_{i,j=1}^N \frac{q_i q_j}{r_{ij,n}} \quad (2.11)$$

Donde la prima indica que los términos con $i=j$ deben ser omitidos cuando $(n_x, n_y, n_z) = (0, 0, 0)$. La distancia $r_{ij,n}$ representa la distancia real y no la distancia mínima entre las imágenes. Esta suma es condicionalmente convergente pero su convergencia es muy lenta.

Para acelerar este proceso se utiliza la Sumatoria de Ewald[56] que consiste en reescribir la sumatoria presentada en la ecuación 2.11 como la suma de dos sumatorias cuya convergencia sea más rápida que la original. El método empleado se encuentra explicado en detalle por Deserno y Holm [57]. Utilizando la identidad:

$$\frac{1}{r} = \frac{f(r)}{r} + \frac{1-f(r)}{r} \quad (2.12)$$

y eligiendo apropiadamente la función $f(r)$ se logra separar en dos términos las dos propiedades del potencial de Coulomb, su rápida variación para r pequeños y su lento decaimiento para r grande. Teniendo esto en cuenta, $\frac{f(r)}{r}$ debería ser despreciable más allá de un cierto valor de corte y $\frac{1-f(r)}{r}$ debería variar lentamente para todo valor de r , de manera que su transformada de Fourier se la pueda representar con pocos vectores \mathbf{k} . Con este fin se elige que $f(r)$ sea igual a la función error complementaria $erfc(r)$

De esta manera, la energía electrostática quedaría de la forma:

$$E = E^{(r)} + E^{(k)} + E^{(s)} + E^{(d)} \quad (2.13)$$

donde los términos de la energía del espacio real $E^{(r)}$, del espacio recíproco $E^{(k)}$, la autoenergía $E^{(s)}$ y la corrección dipolar $E^{(d)}$ vienen dadas por:

$$E^{(r)} = \frac{f}{2} \sum_{n \in Z^3} \sum_{i,j=1}^l q_i q_j \frac{erfc(\alpha |\mathbf{r}_{ij,n}|)}{|\mathbf{r}_{ij,n}|} \quad (2.14)$$

$$E^{(k)} = \frac{f}{2L^3} \sum_{\mathbf{k} \neq 0} \frac{4\pi}{k^2} e^{-k^2/4a^2} |\tilde{\rho}(\mathbf{k})|^2 \quad (2.15)$$

$$E^{(s)} = -\frac{af}{\sqrt{\pi}} \sum_i q_i^2 \quad (2.16)$$

$$E^{(d)} = \frac{2\pi f}{(1+2\epsilon')L^3} \left(\sum_i q_i \mathbf{r}_i \right)^2 \quad (2.17)$$

donde $\tilde{\rho}(\mathbf{k})$ es la transformada de Fourier de la densidad de carga:

$$\tilde{\rho}(\mathbf{k}) = \int_{V_L} d^3r \rho(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{r}} = \sum_{j=1}^N q_j e^{-i\mathbf{k}\cdot\mathbf{r}_j} \quad (2.18)$$

El parámetro α , denominado parámetro de Ewald, permitirá ajustar la contribución del espacio real y el espacio recíproco.

De esta manera se logra dividir a la ecuación inicial 2.11 en cuatro términos cuya convergencia es más rápida. Los dos primeros términos convergen exponencialmente, lo cual permite escoger una distancia de corte pequeña sin perder precisión en el resultado.

El algoritmo resultante es de $\mathcal{O}(N^2)$ lo cual es muy costoso computacionalmente. Debido a esto es que se generó otro algoritmo que implementa la Transformada Rápida de Fourier (FFT) para calcular el término 2.15, de esta manera se logra reducir la velocidad del algoritmo a $\mathcal{O}(N \cdot \log(N))$ [58] este último método se denomina Particle Mesh Ewald (PME).

2.1.7. Integración de las Ecuaciones de Movimiento

Una vez definido el potencial el siguiente paso es resolver la ecuación 2.7 para todas las partículas del sistema. El problema que surge ahora es que este sistema de ecuaciones solo tiene solución analítica para un máximo de dos partículas y en dinámica molecular se debe resolverlo para miles. Debido a esto es que se debe utilizar un método numérico para resolver el problema.

Un método numérico para integrar esta ecuación de movimiento es el algoritmo de Verlet[55]. El cual puede deducirse a partir de los siguientes desarrollos en Taylor:

$$\begin{aligned} x(t + \Delta t) &= x(t) + v(t)\Delta t + \frac{1}{2}a(t)\Delta t^2 + \dots \\ x(t - \Delta t) &= x(t) - v(t)\Delta t + \frac{1}{2}a(t)\Delta t^2 + \dots \end{aligned} \quad (2.19)$$

Luego, reemplazando una de las expansiones en la otra, obtenemos el algoritmo de Verlet para hallar la posición.

$$x(t + \Delta t) = 2x(t) - x(t - \Delta t) + a(t)\Delta t^2 + O(\Delta t^4) \quad (2.20)$$

Este algoritmo tiene la ventaja, frente a una expansión en Taylor, que el error local es de orden $O(\Delta t^4)$.

Otro método equivalente que se utiliza en Dinámica Molecular es el *leap-frog*[59]. Este método calcula las velocidades cada medio paso de tiempo. Para poder hallar las ecuaciones de evolución de la posición y velocidad, definimos a la velocidad como:

$$\begin{aligned} v(t + \frac{1}{2}\Delta t) &= \frac{x(t + \Delta t) - x(t)}{\Delta t} \\ v(t - \frac{1}{2}\Delta t) &= \frac{x(t) - x(t - \Delta t)}{\Delta t} \end{aligned} \quad (2.21)$$

Utilizando estas definiciones y la ecuación 2.20 obtenemos el integrador *leap-frog*.

$$\begin{aligned} x(t + \Delta t) &= x(t) + v(t + \frac{1}{2}\Delta t)\Delta t \\ v(t + \frac{1}{2}\Delta t) &= v(t - \frac{1}{2}\Delta t) + a\Delta t \end{aligned} \quad (2.22)$$

Este integrador es útil debido a su reversibilidad, se obtiene el mismo resultado si se avanza n pasos de tiempo y luego se vuelven otros n pasos, y por su naturaleza simpléctica, conserva la energía del sistema levemente perturbada.

Al momento de utilizar estos integradores hay que prestar especial cuidado al valor del paso de tiempo Δt que utilizaremos. El paso de tiempo tiene que ser lo suficientemente pequeño como para poder muestrear los movimientos más rápidos que ocurren en nuestro sistema. Como regla general $\omega\Delta t \ll 2fs$, donde ω es la frecuencia de oscilación de una partícula.

2.1.8. Restricciones en las Distancias

El tamaño del paso de tiempo Δt es una cuestión muy importante en la Dinámica Molecular. Un paso de tiempo grande nos permite hacer simulaciones más largas en menos tiempo de cálculo pero corremos el riesgo de perdernos movimientos entre las partículas que ocurren en escalas de tiempo menores al Δt o incluso puede ocurrir la aparición de inestabilidades en la estructura como dos átomos muy cerca el uno del otro. Al mismo tiempo, Δt muy pequeños nos aseguran la calidad de nuestra simulación ya que muestrean adecuadamente los movimientos más rápidos pero causan que los tiempos de simulación sean extremadamente largos.

Usualmente se utiliza un paso de tiempo que sea la décima parte del período de vibración más corto en nuestro sistema[60]. Particularmente en moléculas flexibles los movimientos más rápidos se dan en los enlaces con Hidrógenos. El enlace C-H tiene un período aproximado de 10 fs por lo que el paso de tiempo debería ser de 1 fs. Debido a que este valor es muy pequeño y que el movimiento de estos enlaces no son de mucho interés, ya que no tiene gran efecto en los movimientos generales del sistema, se desarrollaron métodos que permiten aumentar el paso de tiempo.

Estos métodos consisten en mantener fijas ciertas coordenadas del sistema, particularmente ciertas longitudes de enlaces. Para lograr esto se utilizan restricciones holonómicas y el método de los multiplicadores de Lagrange para poder resolver el problema.

El algoritmo SHAKE es uno de dichos métodos[61]. Esta técnica resuelve el algoritmo de verlet (Ecuación 2.20) sometido a k restricciones que pueden ser en la longitud del enlace, en el ángulo de enlace, etc. Las restricciones tienen la forma:

$$F_x = \lambda_k \frac{\delta \sigma_k}{\delta x} \quad (2.23)$$

Y particularmente para una restricción en un enlace:

$$\sigma_{ij} = (\mathbf{r}_i - \mathbf{r}_j)^2 - d_{ij}^2 = 0 \quad (2.24)$$

Debido a esto, la ecuación de verlet queda de la forma:

$$\mathbf{r}_i(t + \Delta t) = 2\mathbf{r}_i(t) - \mathbf{r}_i(t - \Delta t) + \frac{F_i(t)}{m_i} \Delta t^2 + \frac{\lambda_k \mathbf{r}_{ij}(t)}{m_i} \Delta t^2 \quad (2.25)$$

Recordando que las coordenadas $\mathbf{r}'_i(t + \Delta t)$ sin perturbar se pueden obtener de la ecuación 2.20, las nuevas coordenadas pueden escribirse como:

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}'_i(t + \Delta t) + \frac{\lambda_k \mathbf{r}_{ij}(t)}{m_i} \Delta t^2 \quad (2.26)$$

Este algoritmo actualiza las coordenadas sin perturbar a las coordenadas perturbadas a través del último término de la ecuación.

Finalmente queda hallar los multiplicadores λ_k que satisfacen todas las restricciones simultáneamente. Este procedimiento puede resolverse analíticamente para algunos casos sencillos pero para sistemas más complejos no es posible. El método de SHAKE resuelve

de a una restricción por vez. Debido a que este procedimiento puede provocar que al satisfacer una restricción se viole otra, es necesario iterar repetidas veces este proceso hasta llegar a una tolerancia aceptable.

Otro método existente para resolver este problema es el SETTLE[62], el cual resuelve el problema analíticamente para pequeñas moléculas. Este método es utilizado para restringir a las moléculas de agua. Por último el algoritmo LINCS[63] resuelve el problema, sin iterar, en dos pasos. Estos pasos son aplicados luego de una actualización del sistema sin restricciones, al igual que en SHAKE. El primer paso define en cero las proyecciones de los nuevos enlaces sobre los viejos. El segundo paso corrige la longitud de los enlaces debido a la rotación de los mismos (Ver Figura 2.2). Este algoritmo presenta la ventaja que acelera el tiempo de simulación entre 3 y 4 veces, sin embargo no es recomendable para realizar restricciones en los ángulos de los enlaces. Aún así estas técnicas aplicadas a la longitud del enlace permite aumentar el paso de tiempo a 2 fs.

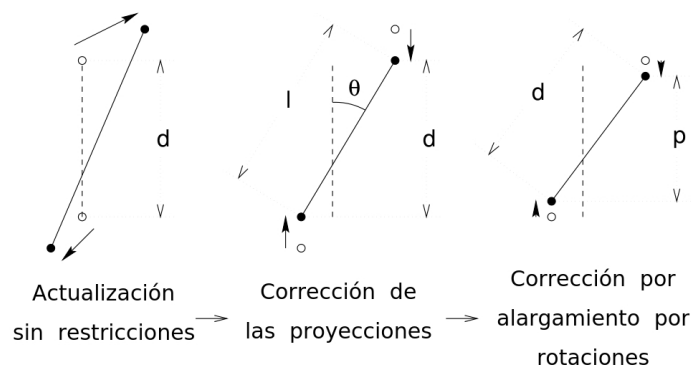


Figura 2.2: Pasos del algoritmo de LINCS. La línea punteada representa a los enlaces viejos de longitud d y la línea continua a los nuevos.

2.1.9. Termostato y Baróstato

Las simulaciones obtenidas por Dinámica Molecular permiten muestrear el espacio de configuraciones de nuestro sistema en condiciones de volumen y energía constante, es decir, muestrean adecuadamente el ensamble microcanónico. Sin embargo, en el laboratorio los experimentos son llevados a cabo en menor medida a volumen y temperatura constante (ensamble canónico) y en mayor medida en temperatura y presión constante (isotérmico-isobárico). Para lograr que las simulaciones sean realizadas bajo estas condiciones se desarrollaron distintos algoritmos. Los que mantienen constante la temperatura lo logran reajustando a cada paso de la dinámica las velocidades del sistema, las cuales

determinan la energía cinética y esta, a su vez, define a la temperatura. Por otro lado los que mantienen constante la presión, lo hacen reajustando las posiciones de todos los átomos. De esta manera alteran el volumen de la caja en la que se encuentra el sistema simulado, alterando consecuentemente la presión.

Termostato de Berendsen

La temperatura (T) de un sistema se la relaciona con la energía cinética promedio (\overline{K}) a través de:

$$\overline{K} = \frac{3}{2} N k_B T \quad (2.27)$$

Donde N es el número de partículas del sistema y k_B la constante de Boltzmann. Luego, si la velocidad de las partículas en un determinado instante t es multiplicada por un factor λ :

$$\Delta T = \frac{1}{2} \sum_{i=1}^N \frac{2 m_i (\lambda v_i^2)}{3 N k_B} - \frac{1}{2} \sum_{i=1}^N \frac{2 m_i v_i^2}{3 N k_B} \quad (2.28)$$

$$\Delta T = (\lambda^2 - 1) T(t) \quad (2.29)$$

$$\lambda = \sqrt{T_{nueva} / T(t)} \quad (2.30)$$

Usando este resultado se puede calcular el valor de λ para ajustar las velocidades a la temperatura deseada.

El termostato de Berendsen[64] reajusta las velocidades pero simulando que el sistema este ligeramente ligado a un baño térmico a la temperatura deseada $T_{baño}$. De esta manera el cambio de temperatura del sistema es proporcional a la diferencia de temperatura con el baño.

$$\frac{dT(t)}{dt} = \frac{1}{\tau} (T_{baño} - T(t)) \quad (2.31)$$

El parámetro τ determina que tan ligado esta el sistema al baño térmico. Este método nos da un decaimiento exponencial hacia la temperatura deseada. De esta manera, para cada paso de tiempo el parámetro λ va a estar dado por:

$$\lambda^2 = 1 + \frac{\Delta t}{\tau} \left(\frac{T_{baño}}{T(t)} - 1 \right) \quad (2.32)$$

Este termostato tiene la ventaja de ser estable y fácil de implementar. Sin embargo, los

promedios canónicos generados no son estrictamente correctos. Aún así es ampliamente utilizado en simulaciones.

Termostato de reescalo de las velocidades

Este termostato, conocido como v-rescale[65], es muy similar al de Berendsen pero corregido de manera que muestree adecuadamente el ensamble canónico y genere una correcta distribución de velocidades. Esta corrección se logra agregando un término estocástico quedando la variación de energía cinética como:

$$dK = \frac{dt}{\tau}(K_{baño} - K) + 2\sqrt{\frac{K K_{baño}}{N_l} \frac{dW}{\sqrt{\tau}}} \quad (2.33)$$

Donde el primer término es exactamente igual al del termostato de Berendsen y el segundo término es una componente estocástica dada por el ruido de Weiner (dW) y donde N_l es el número de grados de libertad.

Baróstato de Berendsen

El algoritmo de Berendsen[64] para mantener la presión constante es similar al que se utiliza para mantener la temperatura constante. En este caso se acopla al sistema a un “baño” a la presión deseada cuya variación viene dada por:

$$\frac{dP(t)}{dt} = \frac{1}{\tau}(P_{baño} - P(t)) \quad (2.34)$$

Donde τ es la constante de acoplamiento. De esta manera, análogamente a lo que sucede en el termostato, se reajusta la caja de simulación con un factor μ , de tal manera que:

$$\mu = V_{nuevo}/V(t) \quad (2.35)$$

O lo que es equivalente, reajustar las coordenadas por $\mu^{1/3}$.

Finalmente, recordando que la compresibilidad isotérmica viene dada por:

$$\kappa = -\frac{1}{V} \left(\frac{\partial V}{\partial P_T} \right) \quad (2.36)$$

Esta definición en conjunto con las ecuaciones (2.34) y (2.35) definen a μ como:

$$\mu = 1 - \kappa \frac{\Delta t}{\tau} (P_{\text{baño}} - P(t)) \quad (2.37)$$

Al igual que en el caso de la temperatura, este termostato es estable y sencillo de implementar pero no muestrea adecuadamente el ensamble.

Baróstato de Parrinello-Rahman

El baróstato de Parrinello-Rahman[66] fue desarrollado de manera que este muestree adecuadamente el ensamble. Este método representa a los vectores que definen la caja de simulación con una matriz \mathbf{b} que sigue la siguiente ecuación:

$$\frac{d\mathbf{b}^2}{dt^2} = V \mathbf{W}^{-1} \mathbf{b}'^{-1} (P - P_0) \quad (2.38)$$

Donde V es el volumen de la caja, y \mathbf{W} es una matriz que determina la intensidad del acoplamiento. Por otro lado, modifica la ecuación de movimiento de las partículas agregando un termino resistivo:

$$\frac{dr_i^2}{dt^2} = \frac{\mathbf{F}_i}{m_i} - M \frac{dr_i}{dt} \quad (2.39)$$

$$\mathbf{M} = \mathbf{b}^{-1} \left[\mathbf{b} \frac{d\mathbf{b}'}{dt} + \frac{d\mathbf{b}}{dt} \mathbf{b}' \right] \mathbf{b}'^{-1} \quad (2.40)$$

La matriz inversa \mathbf{W}^{-1} tiene la información sobre la intensidad del acoplamiento y como se deforma la caja de simulación. Está definido a partir de compresibilidad isotérmica y la constante de tiempo de la siguiente manera:

$$(\mathbf{W}^{-1})_{ij} = \frac{4\pi^2 \beta_{ij}}{3\tau^2 L} \quad (2.41)$$

Este baróstato muestrea adecuadamente el ensamble pero tiene dificultades en llevar al sistema al equilibrio si este se encuentra lejos del equilibrio. En estos casos conviene estabilizar el sistema con el baróstato de Berendsen.

2.1.10. Algoritmo de Actualización

Teniendo en cuenta todos los aspectos mencionados anteriormente se define el algoritmo de actualización que se llevará a cabo en cada paso de la dinámica molecular (Figura 2.3)

Algoritmo de actualización
Dado: - Posición \mathbf{r} de todos los átomos a un tiempo t - Velocidad \mathbf{v} de todos los átomos a un tiempo $t-\Delta t$ - Aceleraciones \mathbf{F}/m en todos los átomos a tiempo t .
<ol style="list-style-type: none"> 1. Calcular los factores de ajuste λ y μ de acuerdo con las ecuaciones (2.32) y (2.37). 2. Actualizar y ajustar las velocidades $\mathbf{v}' = \lambda(\mathbf{v} + \mathbf{a}\Delta t)$ 3. Calcular las nuevas posiciones sin restricciones. $\mathbf{r}' = \mathbf{r} + \mathbf{v}'\Delta t$ 4. Aplicar los algoritmos de restricción a las coordenadas. $\mathbf{r}' \rightarrow \mathbf{r}''$. 5. Corregir las velocidades debido a las restricciones. 6. Ajustar las coordenadas y la caja: $\mathbf{r} = \mu\mathbf{r}''$; $\mathbf{b} = \mu\mathbf{b}$.

Figura 2.3: Algoritmo de Dinámica Molecular.

2.2. Programa de Simulación: GROMACS

GROMACS (GRoningen MACHine for Chemical Simulations)[67] es uno de los programas desarrollados para realizar Dinámica Molecular. A pesar de la existencia de otros programas capaces de obtener resultados similares como NAMD[68], CHARMM[69] entre otros, GROMACS presenta una serie de ventajas que resultaron definitivas al momento de elegir que programa usar para el desarrollo de esta tesis. A continuación se listan las ventajas que presenta Gromacs:

- Cuenta con una continua optimización de los algoritmos, lo cual le otorga un alto desempeño en comparación con otros programas.
- A partir de una de las últimas versiones (4.6), Gromacs puede ser utilizado con las placas de video de NVIDIA (Graphics Processing Unit o GPU) para lograr una mayor aceleración en la simulación al ser programado en CUDA. Esto permite un aumento sustancial en los tiempos de simulación, lo cual es determinante al simular sistemas muy extensos como lo es una cápside viral.

- El programa es amigable al usuario. Todos los archivos de topología y de parámetros están escritos en formato claro. Esto es importante particularmente al momento de utilizar parámetros que no son propios del programa, como en el caso de querer agregar un campo de fuerzas externo.
- Por otro lado, al poseer la capacidad de correr en paralelo y utilizando las GPU, esto permite la realización de simulaciones de grandes sistemas de partículas, en tiempos reducidos y en computadoras de bajo costo comercial al compararlas con un Cluster.
- Gromacs es un software de distribución libre y gratuito.

2.3. Herramientas de Análisis

Una vez obtenida la simulación por dinámica molecular, el siguiente paso es extraer la información deseada de la simulación. Para ello, los programas como Gromacs, cuentan con diversas subrutinas para calcular distintas propiedades, tales como la energía, temperatura, presión, desplazamientos, distancias, radios de giro, estructuras secundarias y más. A continuación se detallarán algunas de ellas.

Al momento de analizar las simulaciones, en primer lugar se corrobora la integridad de la estructura. Es decir que aún mantenga la estructura terciaria de la proteína. Es posible que la estructura se desplace levemente de la estructura original pero se espera que llegue a una estructura estacionaria. Para evaluar la convergencia de la estructura se analiza el desplazamiento cuadrático medio o RMSD (por sus siglas en inglés). El RMSD mide que tan parecidas son dos estructuras y se define como:

$$\sqrt{\frac{\sum_{i=1}^{N_{\text{átomos}}} d_i^2}{N_{\text{átomos}}}} \quad (2.42)$$

Donde $N_{\text{átomos}}$ es el número de átomos de la estructura y d_i es la distancia entre el átomo i de las dos estructuras. Usualmente para evaluar convergencia de estructuras proteicas se evalúa el RMSD sobre los carbonos alpha de la proteína. Este resultado puede ser obtenido fácilmente utilizando la subrutina de gromacs *g_rms*.

Por otro lado, es posible medir el radio de giro de la proteína. Este parámetro nos describe como es la distribución de las partículas alrededor de un eje. De esta manera se

puede estudiar si hubo algún cambio conformacional importante que haya podido alterar sustancialmente esta propiedad. Se lo define como:

$$R_g^2 = \frac{1}{2N^2} \sum_{i,j} (\mathbf{r}_i - \mathbf{r}_j)^2 \quad (2.43)$$

Donde N es el número de partículas y \mathbf{r} la posición de cada partícula. Gromacs cuenta con una rutina (*g-gyrate*) capaz de calcular esta magnitud.

También es posible medir la densidad de moléculas de agua a lo largo de un eje utilizando la rutina de gromacs *g-densmap*. Esta rutina mide la densidad de moléculas de agua promedio a lo largo de toda una simulación en una zona particular del sistema. Particularmente, es posible definir una sección cilíndrica y contar las moléculas de agua presentes en esa región promediada en la dirección azimutal y también a lo largo de toda la simulación. De esta manera se obtiene un mapa de la densidad en función de la coordenada radial y la dirección vertical.

Finalmente, se puede medir el radio de un canal presente en una proteína utilizando el programa Hole[70]. Este programa busca el máximo radio de una esfera en un punto \mathbf{p} sin que se solape con la superficie de Van der Waals de los átomos que forman a la proteína. Esto lo logra buscando el mínimo:

$$R(p) = \min_i^{N_{\text{átomos}}} [|\mathbf{x}_i - \mathbf{p}| - VdW_i] \quad (2.44)$$

Donde $N_{\text{átomos}}$ es el número de átomos de la proteína, \mathbf{x}_i es la posición de cada átomo de la proteína y VdW el radio de Van der Waals de cada átomo. A su vez se realiza una simulación utilizando el método Monte Carlo en combinación con Simulated Annealing con el fin de maximizar el radio posible que pueda presentar el poro. Este proceso se repite para distintos planos paralelos separados por una pequeña distancia a lo largo del poro, obteniendo finalmente el radio a lo largo de toda la estructura.

2.4. Cálculo de Energía Libre

La energía libre[60] es una magnitud termodinámica de gran importancia, ya que nos indica la cantidad de trabajo que un sistema termodinámico es capaz de realizar. O dicho de otro modo, el trabajo necesario para ir de un estado a otro en un sistema. Esta cantidad

se suele expresar de dos formas distintas según en que ensamble se esté trabajando: la energía libre de Helmholtz para un ensamble donde se mantengan constantes la cantidad de partículas, el volumen y la temperatura (NVT), y la energía libre de Gibbs para sistemas en donde se mantiene constante la presión en lugar del volumen (NPT).

La energía libre de Helmholtz de un sistema puede ser calculada como:

$$A = -k_B T \ln Q \quad (2.45)$$

donde k_B es la constante de Boltzmann y Q es la función partición que viene dada por:

$$Q_{NVT} = \frac{1}{N!} \frac{1}{h^{3N}} \iint d\mathbf{p}^N d\mathbf{r}^N e^{-\frac{\mathcal{H}(\mathbf{p}^N, \mathbf{r}^N)}{k_B T}} \quad (2.46)$$

donde N es el número de partículas, h la constante de Planck y \mathcal{H} el Hamiltoniano del sistema. Reemplazando la ecuación 2.46 en 2.45 y utilizando las siguientes identidades:

$$1 = \frac{1}{(8\pi^2 V)^N} \iint d\mathbf{p}^N d\mathbf{r}^N e^{-\frac{E(\mathbf{p}^N, \mathbf{r}^N)}{k_B T}} e^{\frac{E(\mathbf{p}^N, \mathbf{r}^N)}{k_B T}} \quad (2.47)$$

$$\rho(\mathbf{p}^N, \mathbf{r}^N) = \frac{e^{-\frac{E(\mathbf{p}^N, \mathbf{r}^N)}{k_B T}}}{\frac{1}{N!} \frac{1}{h^{3N}} \iint d\mathbf{p}^N d\mathbf{r}^N e^{-\frac{E(\mathbf{p}^N, \mathbf{r}^N)}{k_B T}}} \quad (2.48)$$

donde se reemplazó al Hamiltoniano por la Energía E y $\rho(\mathbf{p}^N, \mathbf{r}^N)$ es la densidad de probabilidad, se llega a la expresión para la energía libre:

$$A = k_B T \ln \left(\iint d\mathbf{p}^N d\mathbf{r}^N e^{-\frac{E(\mathbf{p}^N, \mathbf{r}^N)}{k_B T}} \rho(\mathbf{p}^N, \mathbf{r}^N) \right) \quad (2.49)$$

De esta última expresión se puede observar que los estados de mayor energía son los que más aportan a la integral y, de esta manera, al valor final de la energía libre. Esto acarrea un problema para poder determinar estos valores utilizando el método de la Dinámica Molecular, ya que esta muestrea adecuadamente estados de baja energía. De esta manera, la energía libre y otras propiedades entrópicas obtenidas de trayectorias de dinámica molecular no serán muy precisas.

Debido a esta dificultad, diversas metodologías han sido desarrolladas para mejorar el muestreo de las regiones de alta energía a lo largo de una dinámica molecular. El método de *Potencial de Fuerza Media* nos permite determinar la energía libre a lo largo de una

coordinada de reacción específica. Por otro lado el método de muestreo avanzado *Umbrella Sampling* permite muestrear adecuadamente regiones desfavorables energéticamente.

2.4.1. Potencial de Fuerza Media

Una forma de calcular la energía libre es empleando el método de Potencial de Fuerza Media[71] (PMF por sus siglas en ingles), el cual nos permite hallar el cambio de energía libre en función de alguna coordenada intra o intermolecular, denominada coordenada de reacción, como por ejemplo la distancia entre dos grupos de átomos. De esta manera es posible calcular, por ejemplo, el cambio de energía libre que ocurre cuando un ion atraviesa un canal en una proteína[72]. En este caso la coordenada de reacción es coincidente con el eje del canal.

El nombre de este método se debe a la relación existente entre el promedio de las fuerzas y el trabajo realizado[73]. El PMF, $\mathcal{W}(\mathbf{x}_1)$, es la energía libre requerida para posicionar una partícula en la posición x_1 , relativa a un estado de referencia y viene dado por:

$$e^{-\beta\mathcal{W}(\mathbf{x}_1)} = C \int d^3\mathbf{x}_2 \dots d^3\mathbf{x}_N e^{-\beta U(\mathbf{x}_1, \dots, \mathbf{x}_N)} \quad (2.50)$$

donde se integran todas las coordenadas salvo la coordenada de la partícula 1, C es una constante fijada por el estado de referencia y $\beta = 1/k_B T$. Luego, el gradiente $-\nabla_1 \mathcal{W}(\mathbf{x}_1)$ es la fuerza media ejercida sobre la partícula 1 por el resto de las partículas.

Para un fenómeno de transporte es conveniente definir una coordenada de reacción (ξ) que parametrize el camino recorrido, $\mathbf{x}(\xi)$, por la partícula de interés. En términos de esta coordenada, $\mathcal{W}(\xi(\mathbf{x}_1))$ toma la forma:

$$e^{-\beta[\mathcal{W}(\xi(\mathbf{x}_1)) - \mathcal{W}(\xi_0(\mathbf{x}_1))]} = C \frac{\int d^3\mathbf{x}_2 \dots d^3\mathbf{x}_N \delta(\mathbf{x}(\xi) - \mathbf{x}_1) e^{-\beta U(\mathbf{x}_1, \dots, \mathbf{x}_N)}}{\int d^3\mathbf{x}_2 \dots d^3\mathbf{x}_N \delta(\mathbf{x}(\xi_0) - \mathbf{x}_1) e^{-\beta U(\mathbf{x}_1, \dots, \mathbf{x}_N)}} \quad (2.51)$$

donde ξ_0 es un punto de referencia y C es arbitrario. Por otro lado, a través de la función densidad de probabilidad promedio $\langle \rho(\mathbf{x}) \rangle$ definida por:

$$\langle \rho(\xi) \rangle = \frac{\int d^3\mathbf{x}_1 \dots d^3\mathbf{x}_N \delta(\xi(\mathbf{x}_j) - \xi) e^{-\beta U(\mathbf{x}_j)}}{\int d^3\mathbf{x}_1 \dots d^3\mathbf{x}_N e^{-\beta U(\mathbf{x}_j)}} \quad (2.52)$$

donde la coordenada de reacción ξ típicamente no depende de todas las coordenadas \mathbf{x}_i ,

se puede definir a $\mathcal{W}(\mathbf{x})$ como:

$$\mathcal{W}(\xi) = \mathcal{W}(\xi_0) - k_B T \ln \frac{\langle \rho(\xi) \rangle}{\langle \rho(\xi_0) \rangle} \quad (2.53)$$

donde $\mathcal{W}(\mathbf{x}_0)$ y $\langle \rho(\mathbf{x}_0) \rangle$ son variables arbitrarias. Finalmente, a través, del muestreo del espacio de fases utilizando la Dinámica Molecular es posible determinar la densidad de probabilidad promedio y luego, utilizando la ecuación 2.53 encontrar el PMF. Sin embargo, como se planteó en la Sección 2.4 las regiones con valores de energía alta son los más importantes para determinar correctamente a la energía libre y al mismo tiempo son los que presentan una ocupación mínima, por lo que las funciones de densidad de probabilidad encontradas no son confiables. Por esta razón, es necesario utilizar una metodología de muestreo avanzado en conjunto con PMF para muestrear adecuadamente las regiones de alta energía. Uno de estos métodos se denomina Umbrella sampling.

2.4.2. Umbrella Sampling

El método de Umbrella Sampling[74, 75] soluciona el problema de la falta de muestreo en las regiones de alta energía, agregando un potencial $w(\mathbf{r}^N)$ el cual fuerza a que las partículas se mantengan en la región de alta energía. De esta manera se realizan simulaciones con un nuevo potencial perturbado $U(\mathbf{r}^N) + w(\mathbf{r}^N)$ el cual favorece el muestreo en la región deseada. La forma del potencial normalmente utilizada es la de un potencial armónico.

$$w(\mathbf{r}^N) = \frac{1}{2}k(\mathbf{r}^N - \mathbf{r}_0^N)^2 \quad (2.54)$$

Este potencial restringe el movimiento de las partículas en torno a la posición de equilibrio \mathbf{r}_0^N . De este modo, las configuraciones que se encuentren lejos de la posición de equilibrio serán menos favorables, promoviendo así el muestreo en torno a la zona de interés. Normalmente es necesario utilizar más de una perturbación para lograr muestrear a lo largo de toda la coordenada de reacción (ξ). A cada una de estas perturbaciones se las denomina ventanas y cada una de ellas está caracterizada por una constante de fuerza y por una posición de equilibrio. Luego, la densidad de probabilidad perturbada en la i -ésima ventana viene dada por:

$$\langle \rho(\xi) \rangle_i^{Perturbada} = e^{-\beta w_i(\xi)} \langle \rho(\xi) \rangle_i^{Sin Perturbar} \langle e^{-\beta w_i(\xi)} \rangle \quad (2.55)$$

por lo tanto, la densidad de probabilidad sin perturbar esta dada por:

$$\langle \rho(\xi) \rangle_i^{\text{Sin Perturbar}} = e^{\beta w_i(\xi)} \langle \rho(\xi) \rangle_i^{\text{Pertubada}} \langle e^{-\beta w_i(\xi)} \rangle^{-1} \quad (2.56)$$

Finalmente, reemplazando en la ecuación 2.53 se obtiene el PMF para la i -ésima ventana:

$$\mathcal{W}_i(\mathbf{x}) = \mathcal{W}(\xi_0) - k_B T \ln \frac{\langle \rho(\xi) \rangle_i}{\langle \rho(|x_{i_0}) \rangle - w_i(\xi) + F_i} \quad (2.57)$$

$$\text{con : } e^{-\beta F_i} = \langle e^{-\beta w_i(\xi)} \rangle \quad (2.58)$$

Utilizando esta última ecuación es posible determinar el PMF para cada ventana a lo largo de la coordenada de reacción. Todos los términos de la ecuación 2.57, salvo el F_i , pueden ser determinados de la simulación de la ventana. El término F_i representa la energía libre necesaria para introducir la perturbación. Determinar adecuadamente el valor de este término para cada una de las ventanas es fundamental para poder reconstruir el PMF a lo largo de toda la coordenada de reacción. Uno de los métodos utilizados con el fin de determinar los F_i es el *Weighted Histogram Analysis Method* (WHAM)[76].

Capítulo 3

Poros en el Eje de Simetría Quíntuple de TrV

El objetivo de estudio de la presente tesis, es la cápside del virus del Triatoma (Figura 3.1a). Como se ha mencionado anteriormente, esta cápside presenta varios ejes de simetría, de los cuales prestaremos especial atención al poro presente en el eje de simetría quíntuple ubicado en el centro de la subunidad pentamérica que forma la cápside (Figura 3.1b).

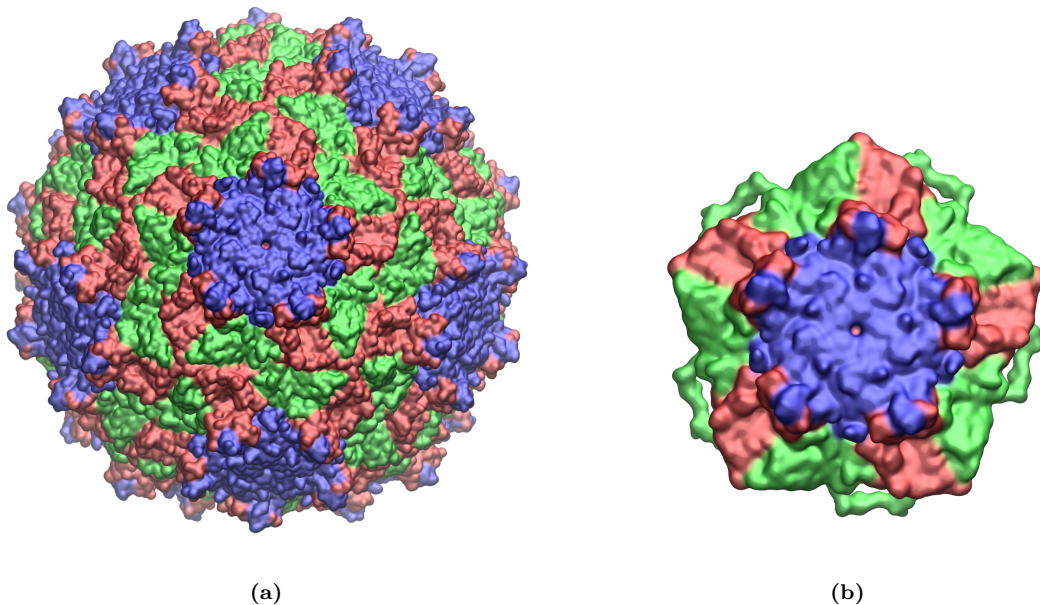


Figura 3.1: Estructura de la Cápside de TrV, en azul, verde y rojo las proteínas virales VP1,2 y3: (a) Cápside completa, (b) Pentámero.

3.1. Puerta Hidrofóbica

Para poder estudiar el rol del poro presente en el eje de simetría quíntuple, el primer paso propuesto fue estudiar el proceso de hidratación del mismo. Esto se debe a que es una condición fundamental para el funcionamiento de un canal que este se encuentre hidratado. Se ha observado en otro trabajo[7] que la energía libre necesaria para que un ion atraviese un canal viene dada en gran parte por la energía necesaria para hidratar al canal. Debido a esto, en esta sección, se propone estudiar el comportamiento de las moléculas de agua a lo largo de la cavidad presente en el eje de simetría quíntuple de TrV.

3.1.1. Protocolo de la Dinámica Molecular

La cápside del virus del *Triatoma* está constituida por 60 copias del protómero formado por las proteínas virales VP1-2-3, generando así una estructura cuaternaria de aproximadamente 360.000 átomos (Figura 3.1a). Debido al gran tamaño de la cápside viral, los tiempos de simulación se ven disminuidos, ralentizando y dificultando el estudio del sistema. Para lograr agilizar el proceso se realizó la dinámica de un único pentón aislado (Figura 3.1b) y luego se compararon los resultados con una simulación de la cápside completa.

Todas las simulaciones de la cápside de TrV fueron realizadas a partir de la estructura cristalográfica (pdb ID: 3NAP[36]) obtenida de la base de datos ViperDB[77].

Para realizar las dinámicas se utilizó el programa GROMACS[67] y una combinación de campos de fuerza para realizar una aproximación multi-escala. Para armar el sistema donde se simuló la cápside completa, se generaron tres niveles de detalle. Esto se logró simulando con detalle atómico a la proteína, usando el campo de fuerza Gromos53a6[52], y una capa de agua de un 1 nm de espesor a su alrededor con el modelo de agua SPC[78]. Luego se utilizó el campo de fuerza Sirah[54] para simular una segunda capa de moléculas de agua de grano-grueso (WT4[79, 80]) con un espesor de 2.5 nm y una última capa de aguas con menor detalle aún (WLS) (Figura 3.2a). Utilizando este método se logra disminuir notablemente el número de partículas del sistema sin perder detalle en el área de interés, y de esta manera acelerar el proceso de simulación.

En el caso de las simulaciones del pentón aislado, se utilizó el campo de fuerza Gro-

mos53a6 para la proteína y el SPC para las aguas en el entorno del pentón en un cascaron de 1 nm. Luego el resto de la caja de simulación fue rellenada con aguas de grano-grueso WT4 (Figura 3.2b).

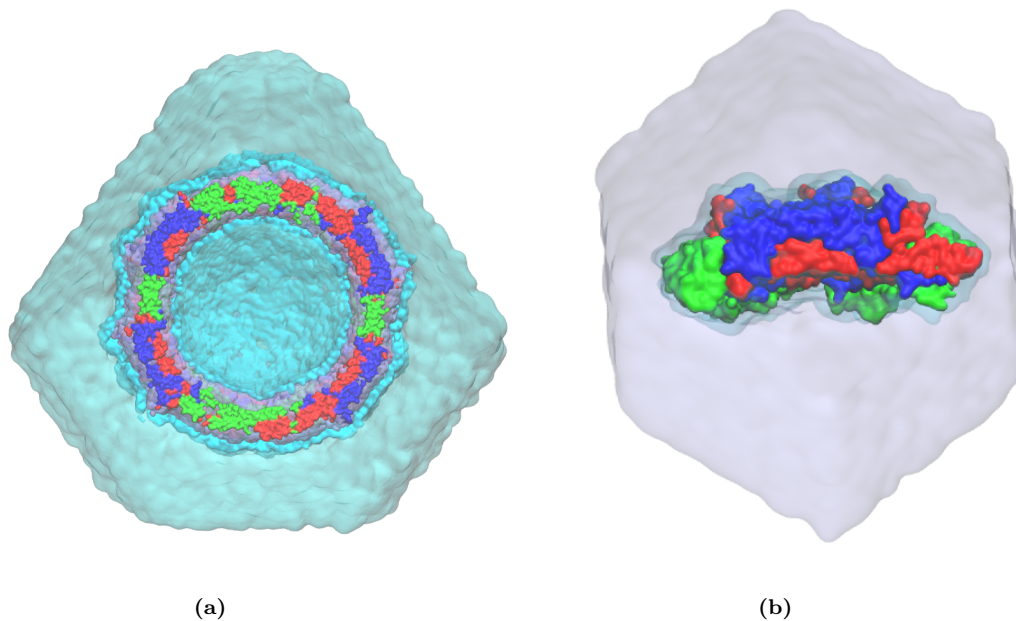


Figura 3.2: Corte transversal del sistema de simulación, (a) cápside completa y (b) pentámero aislado. En azul, verde y rojo las superficies de las proteínas virales VP1, 2 y 3. Seguido de las capas de solvente SPC, WT4 y WLS representadas como superficies transparentes.

El sistema se construyó usando una caja de simulación con forma de un dodecaedro rómbico, ya que de esta manera se disminuye el número de moléculas de agua periféricas que se colocan en el sistema. Luego, aleatoriamente se reemplazan algunas moléculas de agua por iones para simular una solución fisiológica. Las concentraciones de sales utilizadas fueron de 0.15 mol/litro de NaCl y 0.015 mol/litro de MgCl_2 .

El algoritmo de Verlet fue utilizado para buscar primeros vecinos y para también hacer uso de la aceleración por GPU. Las interacciones de largo alcance de Coulomb se calcularon usando el método de Particle-Mesh Ewald y para las interacciones de Lennard-Jones se utilizó el Cut-Off. Cuando se requería que los sistemas estuvieran en contacto con un reservorio térmico o de presión, se utilizaron los algoritmos de Berendsen. Se restringió las distancias de los enlaces de hidrógeno utilizando el algoritmo LINCS, lo cual permite utilizar un paso de integración de 2 fs.

Una vez armadas las cajas de simulación de cada sistema, con sus aguas y la concentración de sales deseadas, el siguiente protocolo fue utilizado para equilibrar al sistema

antes de comenzar a realizar las simulaciones de producción. En primer lugar se realizaron una serie de minimizaciones de energía, estas fueron alternándose entre los algoritmos de *Steepest Descent* y *Conjugate Gradient* hasta que el parámetro de minimización EMTOL llegara al valor deseado de 1000 KJ/molK . El objetivo de este paso es corregir cualquier contacto que pueda haber entre las distintas estructuras que forman el sistema y que puedan provocar una energía extremadamente alta.

El siguiente paso en la equilibración es permitir que las moléculas de agua se relajen alrededor de la proteína. Para ello se permite que las aguas se muevan libremente mientras que se restringe el movimiento de la proteína. Este proceso se lo denomina restricción de posiciones ('Position Restraint' o PR). En esta etapa se permitió al sistema interactuar con un reservorio térmico a 310 K con una constante de tiempo de 0.2 ps. Esta simulación se la extendió por 500 ps.

El último paso en la equilibración es la termalización, la cual lleva al sistema de una temperatura inicial de 100 K a una temperatura final de 310 K. Esta simulación, al igual que el PR se la realizó a volumen constante con el sistema conectado a un reservorio térmico cuya constante de tiempo era de 1 ps. En los primeros 500 ps, las velocidades de las partículas se las extrajo de una distribución de velocidades de Maxwell a 100 K y el reservorio se encontraba a 200 K. Luego los siguientes 500 ps se le permitió al sistema interactuar con un reservorio a 310 K, permitiendo al sistema llegar a la temperatura final deseada.

Finalizado el protocolo anterior, se comienza a hacer las simulaciones de producción. En estas simulaciones se permite al sistema interactuar con un reservorio de temperatura y otro de presión, de esta manera el ensamble de simulación es a temperatura y presión constante. La temperatura fue mantenida a 310 K con una constante de tiempo de 0.5 ps y la presión a 1 bar y su constante de tiempo era de 0.4 ps.

En caso de ser necesario realizar alguna mutación de aminoácidos se utilizó el programa pymol[81] para lograrlo. En tal caso se seleccionó el rotámero del mutante cuya cadena lateral tuviera el menor contacto con el resto de la proteína.

Siguiendo este protocolo se generó una simulación de 50 ns de la estructura cristalográfica de la cápside completa (TrV_Full.WT). Por otro lado se obtuvieron 100 ns de simulación del pentón aislado (TrV_WT). Finalmente, se obtuvo un conjunto de simulaciones del pentón aislado con una duración de 20 ns a 100 ns donde se mutaron todas las

valinas 3012 por cinco aminoácidos iguales (Serina, Asparagina, Treonina y Glicina).

Por otro lado, se repitieron todas las simulaciones del pentón aislado cambiando el modelo de agua al SPCE[82]. Debido a que varios de los resultados obtenidos están relacionados con la interacción del agua con la cápside viral, se deseó ver si los fenómenos observados eran producto de la parametrización del campo o eran propios de las características de la cápside.

El interés en trabajar con el pentámero aislado radica en que este sistema tiene un menor número de partículas, lo cual se traduce en simulaciones más rápidas, de mayor duración (ver figura 3.3) y por ende mejor estadística en los resultados. Por otro lado, se ha observado que el pentámero es capaz de mantener su estructura luego del desmantelamiento de la cápside al interactuar con otro pentámero[40]. Por esta razón, se considera que los resultados del pentámero aislado son comparables a los de la cápside completa.

	Nº de partículas	Velocidad de Simulación
TrV Cápside Atomística	~3.300.000	?
TrV Cápside Multiescala	~1.300.000	~1 <i>ns/día</i>
TrV Pentámero Atomístico	~900.000	~2 <i>ns/día</i>
TrV Pentámero Multiescala	~170.000	~10 <i>ns/día</i>

Figura 3.3: Velocidades de simulación para distintos sistemas del virus de TrV. Todas las simulaciones fueron realizadas en una computadora con procesadores i7 de 4-6 núcleos y placas de video GeForce 780 o 980.

3.1.2. Resultados y Discusión

La estructura estudiada fue la cavidad presente en el eje de simetría quíntuple. Como se menciono anteriormente, esta cavidad está compuesta por los seis anillos de glutamina 1128, valina 1166 y treonina 1167 pertenecientes a VP1, y glutamina 3014, valina 3012 y treonina 3010 pertenecientes a VP3 (Figura 3.4).

Una vez armados los dos sistemas descritos en la sección de protocolo con aguas SPC, TrV_Full_WT y TrV_WT, se procedió a su análisis. El primer paso fue la comparación de ambas estructuras para ver si los resultados de la segunda simulación son confiables. Para ello se estudió el RMSD, el radio de giro y la dinámica de ambas estructuras, particularmente en los bordes del pentámero aislado y en el eje quíntuple de ambas simulaciones.

Observando los gráficos 3.5a y 3.5b se puede observar que ambas estructuras llegan a converger en valores no muy lejanos de la estructura de referencia. Si bien la estructura

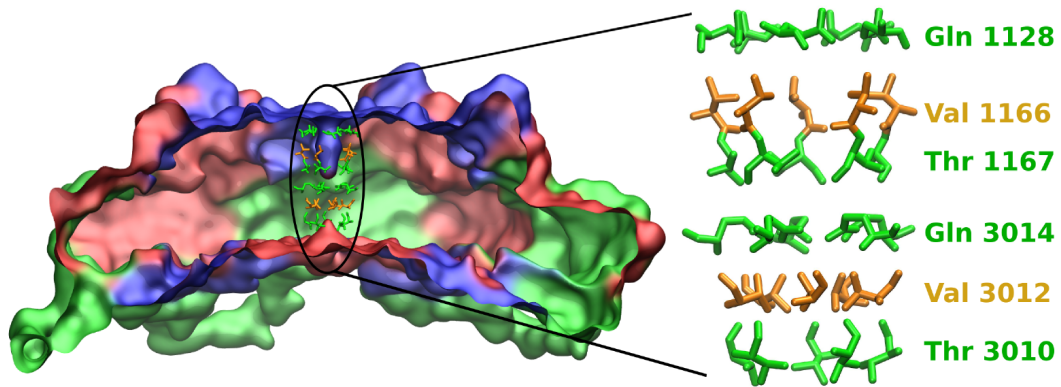


Figura 3.4: Aminoácidos que forman el poro de TrV. (Izquierda) Pentón de TrV visto de lado. Las proteínas están coloreadas azul para VP1, verde para VP2 y rojo para VP3. (Derecha) Aminoácidos que forman el poro presente en el eje quíntuple. Los aminoácidos están coloreados en verde los polares y en naranja los no-polares.

de TrV_WT llega a valores levemente mayores, estos valores todavía son aceptables. Lo mismo se puede observar en las figuras 3.5c y 3.5d donde se observan los radios de giro de ambas estructuras. Ambos valores permanecen constantes en la mayoría de la simulación y nos demuestra que la forma tridimensional de ambas estructuras no cambió sustancialmente. Es decir, la cápside no se desensambló y el pentámero mantuvo su estructura cuaternaria.

Por otra lado, se estudió las dinámicas de ambas simulaciones para ver que no ocurriera ningún cambio sustancial en la estructura y finalmente se estudió el comportamiento del poro presente en el eje quíntuple. En el caso de la cápside completa se analizaron los 12 poros y en el caso del pentámero el único presente.

Debido a que se observó el mismo comportamiento en ambas simulaciones y a que la simulación del pentámero aislado resulta un sistema de un número considerablemente menor de partículas (~ 170.000 contra $\sim 1.300.000$ átomos), se decidió continuar todos los estudios con esta simulación ya que permite aumentar considerablemente los tiempos de simulación.

Además, a partir de lo observado en la figura 3.5a, se consideró que el sistema convergió, es decir llegó a una estructura estable, a partir de los 10 ns. Teniendo esto en cuenta, se realizarán todos los análisis siguientes utilizando los 90 ns finales de la simulación TrV_WT.

De esta simulación, se puede observar que el poro presente en el eje quíntuple presenta una zona deshidratada, donde no penetra ninguna molécula de agua. Esto se observa

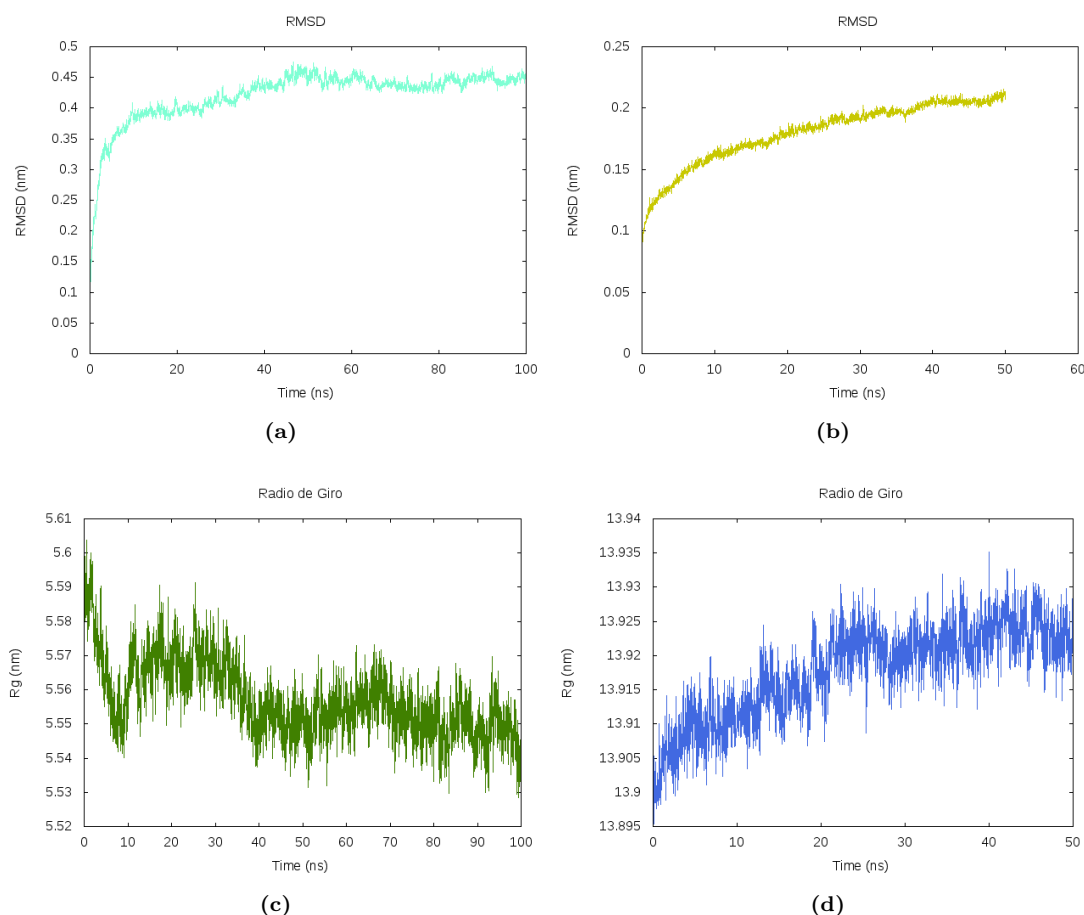


Figura 3.5: RMSD del backbone y Radio de giro de TrV: (a) y (c) pentámero y (b) y (d) Cápside completa. En los gráficos de RMSD se ve la convergencia de las estructuras. En conjunto con el radio de giro se puede ver que las estructuras permanecen estables a lo largo de la simulación sin una variación considerable de su estructura tridimensional.

claramente en la figura 3.6C, donde se puede ver una vista lateral del poro del eje quíntuple de TrV. Como se aprecia en la figura, la región deshidratada coincide con la posición de la Valina 3012 que es un aminoácido con carácter no-polar o hidrofóbico. Esto se comparó con el mapa de densidad promedio de moléculas de agua, calculado con el *g_densmap* utilizando los últimos 90 ns. En el mismo se puede observar (Figura 3.6B) que esta ausencia de moléculas de agua no es un evento fortuito de un único instante en la simulación sino que es un fenómeno que ocurre a lo largo de toda la simulación.

A su vez, se puede ver (Figura 3.6A) que esta región hidrofóbica coincide con la región de menor radio del poro. Utilizando el programa HOLE se puede observar que el radio del poro en la región más angosta puede alcanzar valores mayores a 1.4 \AA . Teniendo en cuenta que el radio de Van der Waals de una molécula de agua es de $\sim 1.4 \text{ \AA}$, la ausencia de moléculas de agua en esta región no es necesariamente por un impedimento estérico sino por las características no-polares de los aminoácidos del entorno. Debido a esto, se

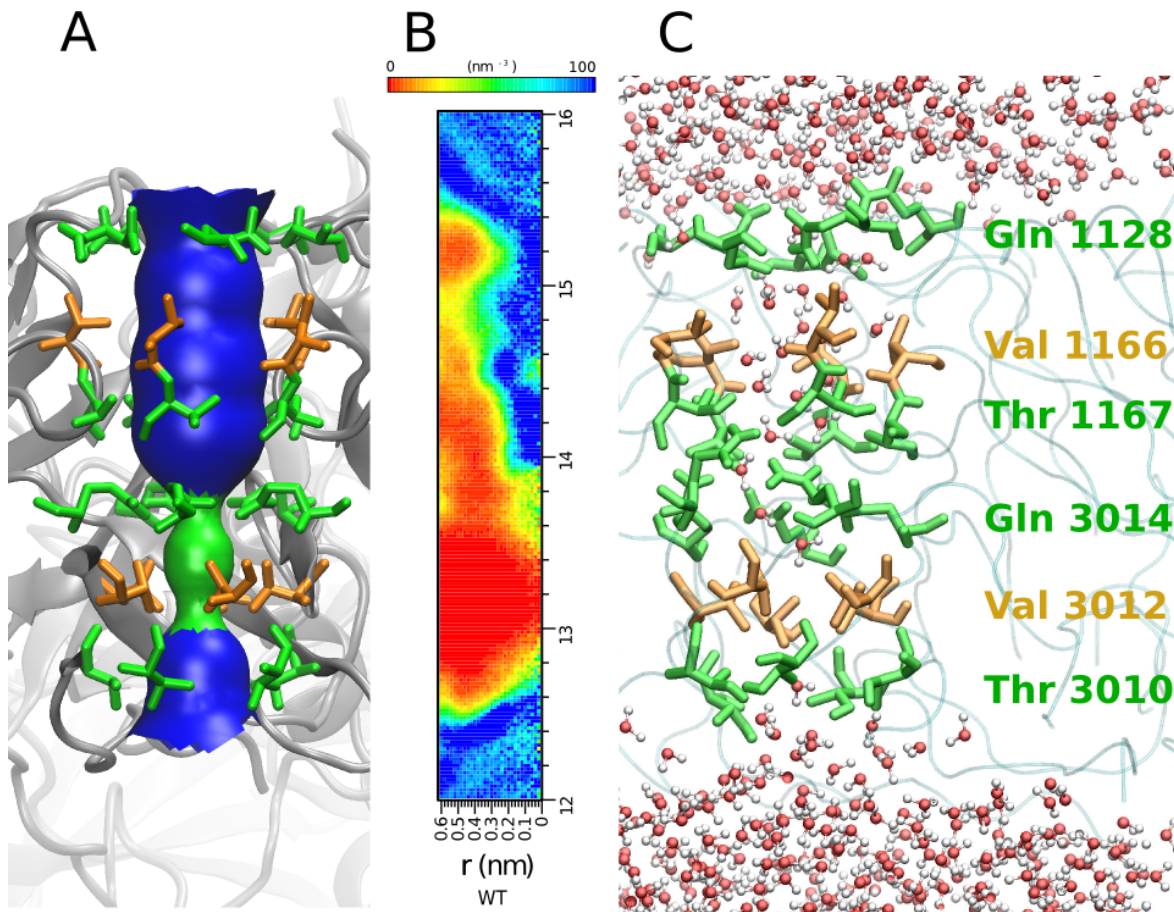


Figura 3.6: Estructura del poro presente en el eje quíntuple de TrV. (A) Radio del poro representado en una superficie, donde la región verde corresponde a la de menor radio. (B) Mapa de densidad de moléculas de agua en el poro en función del radio del poro y la posición vertical. (C) Vista lateral del poro de TrV en la simulación del pentámero aislado, en verde se muestran los aminoácidos con carácter polar y en naranja los de carácter no-polar.

postula la existencia de un mecanismo de puerta hidrofóbica que regule la apertura y cierre del poro. Este tipo de mecanismos es comúnmente estudiado en canales iónicos de proteínas de membrana, pero nunca ha sido observado en cápsides virales.

Para confirmar esta hipótesis se realizaron varias simulaciones donde las valinas 3012 fueron mutadas por otros aminoácidos (Glicina, Serina, Treonina y Asparagina). Se eligió a la glicina por su reducida cadena lateral, el fin era observar si aumentando el espacio disponible para las moléculas de agua estas entrarían. Por otro lado, las mutaciones con serina, treonina y asparagina fueron realizadas para ver si la imposibilidad de que las moléculas de agua penetraran en el sistema era por motivos estéricos o debido a la característica no-polar de la valina. Particularmente, la serina y la treonina poseen una cadena lateral de similar longitud con la de la valina, mientras que la asparagina posee una cadena de mayor longitud.

Tanto la mutación de treonina como la asparagina, no presentaron hidratación completa de la cavidad. Sin embargo, una minuciosa inspección de la simulación y de la estructura formada por estos aminoácidos muestra que el poro siempre se encuentra cerrado estéricamente por las cadenas laterales de estos aminoácidos. Por otro lado, la mutación por glicina presenta una hidratación de sus alrededores. No obstante, la presencia de las glicinas provocan un cambio conformacional en el poro generando un cierre del mismo en la región ocupada por VP1. Esto indicaría que la presencia de las valinas tiene una implicancia estructural en el poro además de regular su apertura.

Por último, la mutación por serina, cuya cadena lateral es similar en longitud a la de la valina pero posee características polares, presentó una completa hidratación del poro sin alterar la estructura del pentámero (Figura 3.7). En este caso los mapas de densidad fueron creados con los últimos 80 ns de simulación, ya que el RMSD de esta simulación muestra convergencia a partir de los 20 ns y además este instante coincide con la hidratación del poro.

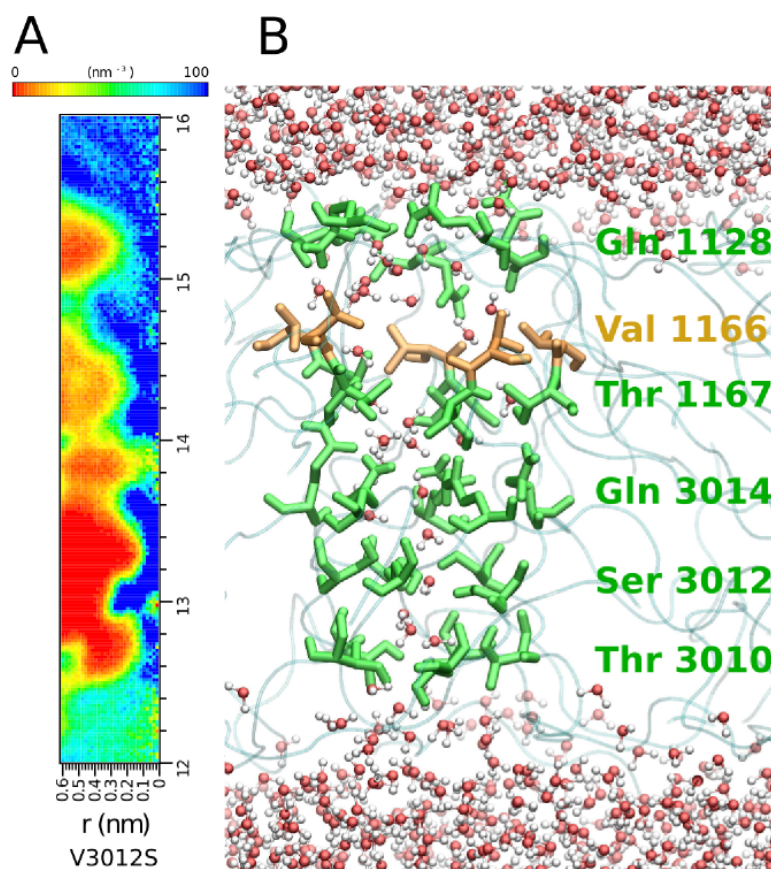


Figura 3.7: Estructura del poro presente en el eje quíntuple de TrV al mutar la valina 3012 por una serina. (A) Mapa de densidad de moléculas de agua en el poro en función del radio del poro y la posición vertical. (B) Vista lateral del poro de TrV mutado, en verde se muestran los aminoácidos con carácter polar y en naranja los de carácter no-polar.

Todos los resultados mencionados anteriormente para el pentámero aislado con aguas SPC fueron repetidos cambiando el modelo de aguas al SPCE. Todas las simulaciones del pentámero y sus mutaciones muestran el mismo comportamiento observado para las simulaciones donde se utilizó el modelo de aguas SPC.

Basado en estos resultados, se propone que el eje quíntuple de TrV presenta un poro capaz de comunicar el medio exterior de la cápside con su interior y además posee una puerta hidrofóbica situada a la altura de las valinas 3012. Esta característica no ha sido observada en cápsides virales pero sí ha sido estudiada en canales iónicos de membrana y otros sistemas no biológicos [83, 44, 84, 85].

3.2. Energía libre de un Ion a través del Poro

Como se mencionó en la sección anterior, TrV presenta un mecanismo de regulación de apertura en el poro presente en su eje de simetría quíntuple. Este mecanismo de puerta hidrofóbica provoca la deshidratación del poro en la región de las valinas 3012. Debido a que para que un poro actúe como canal de alguna molécula es necesario que el mismo se encuentre completamente hidratado, en esta sección estudiaremos bajo que condiciones el poro es capaz de hidratarse completamente.

En trabajos previos, donde se estudió la estructura de TrV por cristalografía, se presentó una densidad electrónica en la región del eje quíntuple que fue omitida al momento de resolver la estructura [36]. En este mapa (Figura 3.8) se muestra una densidad, a la altura de las glutaminas 3014, aproximadamente esférica, la cual postulan podría deberse a un ion y particularmente al de un ion Mg^{2+} ya que la solución donde fue creado el cristal era rica en dicho ion.

Partiendo de esta información se estudió la factibilidad de que un ion Mg^{2+} penetre en el poro. Para ello se calculó la variación energía libre de un ion Mg^{2+} a lo largo del poro de TrV. Esto se logró calculando el Potencial de Fuerza Media (PMF) del ion a lo largo de la coordenada de reacción definida por el eje quíntuple de TrV utilizando el método de Umbrella Sampling.

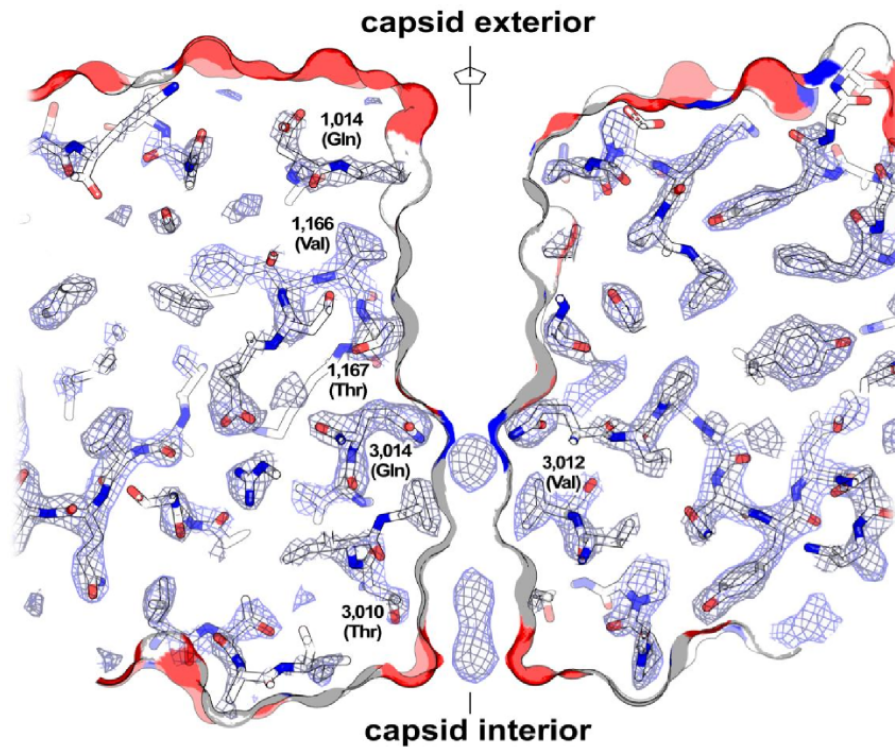


Figura 3.8: Mapa de densidad electrónica de TrV de las proteínas VP1 y VP3 en la región del eje de simetría quíntuple (Material suplementario[36]).

3.2.1. Protocolo de muestreo avanzado

Para calcular el PMF de un ion Mg^{2+} a lo largo de la coordenada de reacción definida por el eje de simetría quíntuple de TrV, se utilizó el método de muestreo avanzado de Umbrella Sampling. Con este fin se alineó al pentámero de manera que el eje quíntuple coincidiera con el eje Z. Luego se generaron 29 ventanas de simulación donde el ion Mg^{2+} fue colocado a lo largo del eje Z separado por 1 \AA entre ventana y ventana. A continuación, se agregaron 11 ventanas más (para llegar a un total de 40 ventanas) para mejorar el muestreo en regiones particulares del poro.

La configuración inicial para estas simulaciones fue la configuración final de la simulación del pentámero aislado (TrV_WT), presentada en la sección anterior. Utilizando esta configuración, se colocó al ion Mg^{2+} en la posición deseada en la coordenada de reacción. Luego se realizó una minimización de energía usando el algoritmo de Steep Descent para relajar al sistema frente al agregado del ion. A continuación, se comenzó la dinámica para realizar el muestreo de la coordenada de reacción.

Con el fin de muestrear adecuadamente la región deseada, se utilizó el Pull-Code de Gromacs con un potencial armónico para restringir el movimiento del ion. Esto es nece-

sario ya que de lo contrario el ion abandonaría la posición deseada debido a la repulsión que sentiría por estar en una región de alta energía.

Tanto la minimización de energía como la dinámica fueron realizadas con los mismos parámetros que las dinámicas de la sección anterior, salvo por la restricción presente sobre el ión Mg^{2+} . El Pull-Code utilizado fue el de Umbrella con la geometría de posición. De esta manera, el potencial armónico tiraba entre los centros de masa del ion y el de los aminoácidos que forman al poro, forzando al ion a mantenerse en su posición inicial. La restricción en el movimiento solo fue aplicada en la dirección del eje, el cual se hizo coincidir con el eje Z del sistema. La constante de fuerza utilizada fue de $3000 \text{ KJ/mol} \cdot \text{nm}^2$ en las primeras 29 ventanas. Debido a que ciertas regiones del poro presentan barreras de mayor energía, es más desfavorable que el ion pueda muestrearlas adecuadamente. Por esta razón, 11 ventanas fueron agregadas para mejorar el muestreo en dichas regiones. Para lograr esto, las constantes de fuerza fueron aumentadas a $6000 \text{ KJ/mol} \cdot \text{nm}^2$ y $10.000 \text{ KJ/mol} \cdot \text{nm}^2$, de este modo el ion permanece en la región desfavorable. Cada una de estas ventanas fueron simuladas por 10 ns.

Una vez obtenidas todas las simulaciones, se utilizó la subrutina de Gromacs *g-wham* [86] para calcular la energía libre del ion a lo largo de la coordenada de reacción.

Por otro lado, para corroborar los resultados obtenidos con PMF, en algunas de las ventanas utilizadas anteriormente se permitió la evolución temporal del ion libre de la fuerza armónica introducida por el Pull-Code. De esta manera se pudo observar el comportamiento del ion libre dentro del poro. En estos casos el protocolo fue el mismo que en las simulaciones de TrV_WT y fueron extendidas hasta 100 ns.

3.2.2. Resultados y Discusión

Luego de obtener todas las simulaciones de las 40 ventanas de Umbrella Sampling, se utilizó el programa *g-wham* para obtener el Potencial de Fuerza Media. De esta manera se logró obtener el perfil de Energía Libre de Gibbs de un Ion Magnesio atravesando el poro del eje de simetría quíntuple de TrV (Figura 3.9A). En esta figura se observa que el perfil de energía libre presenta un máximo a la altura de las Valinas 3012 que coincide con la posición de la puerta hidrofóbica postulada en la sección anterior. A su vez se advierte la presencia de un mínimo a la altura de las Glutaminas 3014. Este mínimo coincide con la posición de la densidad electrónica omitida en la figura 3.8, lo cual corrobora la hipótesis

de la presencia de magnesio en esa región. Además, la amplitud del máximo presente en las valinas 3012, indica que el ingreso del ion hacia el mínimo ocurre desde el exterior de la cápside ya que la barrera energética a atravesar es mucho menor.

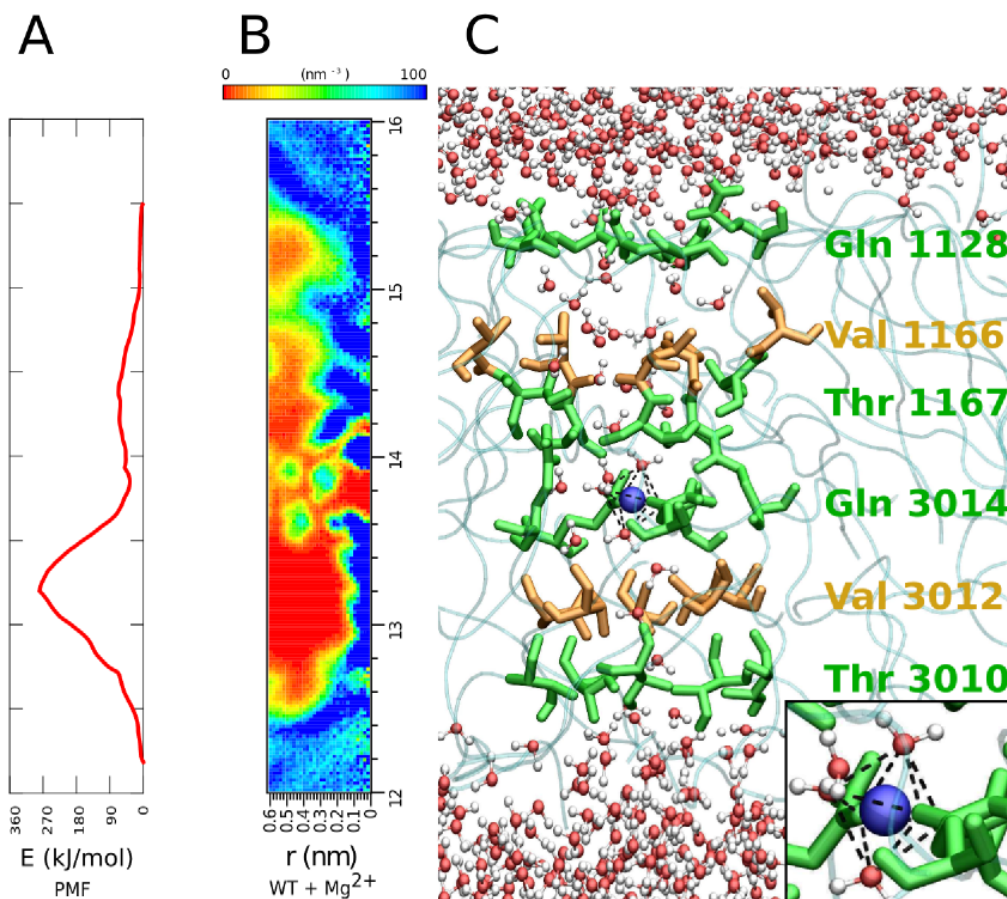


Figura 3.9: Estructura del Poro de TrV con la presencia de un ion Mg^{2+} en su interior. (A) Energía Libre de Gibbs del Mg^{2+} en función de la coordenada de reacción. (B) Mapa de densidad de moléculas de agua en el poro en función del radio del poro y la posición vertical. (C) Vista lateral del poro de TrV en la simulación del pentámero aislado con la presencia de un ion Mg^{2+} a la altura de la Glutamina 3014. En la sección inferior derecha se muestra aumentada la coordinación octaédrica del Magnesio con el oxígeno. En verde se muestran los aminoácidos con carácter polar y en naranja los de carácter no-polar.

Basados en el resultado anterior, se realizó una serie de simulaciones con la presencia de un ion Mg^{2+} libre a lo largo de la coordenada de reacción. Como se esperaba de lo observado en el gráfico de energía libre (Figura 3.9A), los iones que se ubicaban en posiciones cercanas a los extremos del poro, rápidamente abandonaban el canal en periodos menores a 2 ns. Por otro lado, los iones colocados en la cercanía del mínimo se coordinaban con las glutaminas 3014.

En esta simulación se observó que el ion permaneció en la región inicial y se coordinó con tres moléculas de agua y con los oxígenos de las cadenas laterales de tres glutaminas

3014 para formar una estructura octaédrica. Esta estructura resultó ser muy estable a lo largo de los 100 ns de simulación ya que en ningún momento el ion abandonó esta región así como tampoco lo hicieron las moléculas de agua a las que se coordinó inicialmente. En conjunto con este fenómeno, se observó que la región hidrofóbica de las valinas 3012 se encontraba hidratada. Esto es de especial interés ya que presenta un posible mecanismo por el cual regular la apertura del poro. Dicho fenómeno, se puede observar claramente en la figura 3.9B donde se muestra el mapa de densidad de moléculas de agua, construido con los últimos 80 ns de simulación. Aquí se observa que la región previamente deshidratada se encuentra completamente solvatada. La pequeña región que se observa completamente deshidratada a la altura de las glutaminas 3014 coincide con la posición del magnesio. Por otro lado, se puede observar una cadena continua de moléculas de agua que comunica el exterior con el interior de la cápside (Figura 3.9C).

Esta última simulación, donde se permitió la libre evolución del ion magnesio, también fue realizada en una cápside completa para comparar los resultados obtenidos. En este caso se colocó un ion en la posición coincidente con el mínimo de energía libre en cada uno de los 12 ejes de simetría quíntuple y se estudió su evolución. Los resultados fueron los esperados, donde se observó que algunos ejes de simetría presentaban el mismo comportamiento que la simulación con el pentón aislado y el Mg^{2+} . En algunos casos, si bien el ion se coordinó en la misma región la coordinación variaba levemente, se coordinaba con 4 moléculas de agua y dos glutaminas, con 4 glutaminas y dos moléculas de agua y otras variantes estructurales. Algunas de estas variaciones no lograban la hidratación del poro. De todas maneras, el evento más observado fue el mismo que se observó en la simulación del pentámero aislado.

Por otro lado, simulaciones similares fueron realizadas con otros iones como Na^+ y Cl^- para ver si presentaban comportamientos similares. En el caso de los iones Cl^- , estos rápidamente abandonan al poro, mostrando que no se trata de una zona favorable para los mismos. Por otro lado, los iones Na^+ mostraban mayor afinidad por permanecer dentro del poro pero no tanto como el Mg^{2+} .

Finalmente, un análisis del movimiento de las moléculas de agua que hidrataban a la puerta hidrofóbica muestra que estas moléculas presentan una movilidad reducida. Particularmente, una vez que llegan a la posición que permite la hidratación del canal, permanecen en dicha posición a lo largo de toda la simulación. Debido a esta baja movi-

lidad es poco probable que el canal actúe como un canal de aguas. Tampoco es probable que permita el paso de iones, ya que el ion Mg^{2+} también presenta poca movilidad y, como se mencionó anteriormente, los iones deben atravesar los canales en conjunto con su esfera de hidratación[7], que por lo expuesto anteriormente respecto al movimiento de las aguas es poco probable. Debido a todo esto se propone que el canal podría actuar como una cadena transportadora de protones.

Basados en estas observaciones, se puede concluir que la presencia de iones magnesio en el poro de TrV, particularmente en el mínimo de energía libre el cual coincide con la posición de las glutaminas 3014, juegan un rol fundamental en la completa hidratación y consecuente apertura de la puerta hidrofóbica presente en las valinas 3012. A su vez el perfil de energía libre nos indica que el ingreso de los iones ocurre desde el exterior de la cápside ya que de lo contrario, los iones deberían atravesar una barrera energética mayor. Por último, la cadena continua de moléculas de agua podría indicar la existencia de un mecanismo de transporte de protones entre el exterior y el interior de TrV.

3.3. Comparación con Otros Virus Icosaédricos

Debido a la puerta hidrofóbica observada en la cápside de TrV y cómo está podría regular el intercambio entre el interior del virus y sus alrededores, se estudió si esta característica era común al resto de los virus de su orden. En primer lugar se buscaron las cápsides virales pertenecientes al orden Picornavirales, cuyas estructuras tridimensionales hayan sido resueltas, y se observó si estas presentaban en su estructura un arreglo de aminoácidos capaz de funcionar como puerta hidrofóbica. Luego se eligieron algunas estructuras representativas de cada familia y se estudió el eje quíntuple de cada estructura por dinámica molecular.

3.3.1. Protocolo de la Dinámica Molecular

Las estructuras de las cápsides virales fueron obtenidas de la base de datos ViperDB y luego estudiadas con el programa VMD[87]. Las dinámicas moleculares de las estructuras seleccionadas fueron realizadas con los mismos protocolos detallados en la sección 4.1.1 para las simulaciones de pentones aislados.

3.3.2. Resultados y Discusión

Estructuras de las familia de Picornavirus, Dicistrovirus y Secovirus fueron estudiadas. Las familias Iflavirus y Marnavirus no tienen ninguna estructura resuelta cristalográficamente, por esta razón no se incluyen en este estudio. De la familia Picornavirus se estudió la estructura de los siguiente virus: Human Enterovirus 71 (3VBS), Human Rhinovirus 16 (1AYM), Foot and Mouse Disease Virus (4GH4), Equine Rhinitis A Virus (2WFF), Theiler Murine Encephalomyelitis Virus (1TMF), Mengo Encephalomyelitis Virus (2MEV), Seneca Valley Virus (3CJI), Poliovirus (1HXS) y Hepatitis A Virus (4QPI). De la familia Dicistrovirus se estudió las únicas dos estructuras resueltas: Cricket Paralysis Virus (1B35) y Triatoma Virus (3NAP). Por último de la familia Secovirus se estudiaron las estructuras de: Cowpea Mosaic Virus (1NY7), Bean Pod Mottle Virus (1BMV), Grapevine Fanleaf Virus (2Y7T) y Tobacco Ringspot Virus (1A6C).

Todas estas estructuras fueron estudiadas en VMD para observar la composición de aminoácidos presentes en el eje de simetría quíntuple. A partir de este análisis se pudo observar ciertas características en común entre todos los virus. En primer lugar, el poro presente en el eje de simetría quíntuple esta compuesto por varios anillos de aminoácidos, cada uno de ellos formado por sus cinco copias simétricas. La cantidad de anillos varía de estructura a estructura desde 6 a 27 aminoácidos. Las propiedades de los mismos también varían, no todos presentan aminoácidos cargados, característica normal en los canales iónicos de membrana. Todos los virus, salvo los Secovirus, tienen su poro formado por las proteínas virales VP1 y VP3 y en algunos casos por VP4. En el caso de los Secovirus, el poro solo se encuentra formado únicamente por VP1.

La característica más interesante observada es que, en todos los virus pertenecientes a las familias Picornavirus y Dicistrovirus, la región más angosta del poro ($\sim 1.4\text{-}4 \text{ \AA}$ de radio) está formada por uno o más anillos hidrofóbicos. Como se observó en TrV, los radios son lo suficientemente grandes para el paso de moléculas de agua pero la propiedad no-polar de los aminoácidos presentes podría evitar el pasaje de agua. Debido a esto, es posible que todos estos virus presenten el mismo mecanismo de regulación de su apertura a través de una puerta hidrofóbica. En el caso de los Secovirus, no se observó la presencia de un anillo hidrofóbico lo suficientemente angosto como para poder crear un mecanismo de regulación de este estilo.

A continuación, se eligieron algunos virus representativos y se realizaron dinámicas

moleculares de los mismos para corroborar si efectivamente las constricciones observadas pueden actuar como puerta hidrofóbica y para observar en el caso de los Secovirus que la ausencia de estos mecanismo permite el paso de moléculas de agua. Los virus elegidos fueron Poliovirus (PoV), Hepatitis A Virus (HAV) y Human Rhinovirus 16 (HRV16) de la familia Picornavirus, Cricket Paralysis Virus (CrPV) de los Dicistrovirus y Bean Pod Mottle Virus (BPMV) y Tobacco Ringspot Virus (TRsV) de la familia Secovirus. Todos los virus fueron simulados por 100 ns salvo los Secovirus que fueron simulados por 50 ns. En cada caso se estudió los RMSD de cada simulación y se eligió el período de simulación donde la misma llegó a una estructura estable. A partir de estas porciones de simulación (80 ns para las simulaciones cuya extensión era de 100 ns y 40 ns para las cuales su extensión era de 50ns) se construyeron los mapas de densidad de moléculas de agua utilizando el programa *g_dens*.

En las figuras figuras 3.10 y 3.11 se puede observar los mapas de densidad de moléculas de agua para los poros presentes en el eje quíntuple de HRV16 y PoV. Ambos mapas muestran una región deshidratada que coincide con una de las regiones más angostas de la cavidad. Sin embargo, el radio observado en ambos canales es suficiente para que pueda entrar una molécula de agua. Dado que esta región coincide con la ubicación de los anillos de aminoácidos hidrofóbicos, la causa de que las moléculas de agua no logren entrar en esta región es debido al carácter no-polar de la misma. Se puede observar que estas regiones son de mayor tamaño que en TrV al estar compuestas por tres anillos, en el caso de HRV16 (Valina 3006, Valina 3004 y Leucina 3002), o por dos anillos, en el caso de PoV (Valina 3004 y Leucina 3002).

También se puede ver que ambos virus presentan una región angosta en la parte superior del poro, lindando con el exterior del virus. Esta región coincide con la ubicación de un anillo de histidinas. En el caso de Rhinovirus (3.10A), esta región es hidratada a lo largo de la simulación, indicando que las aguas son capaces de penetrar en el poro y aumentar levemente el radio del mismo. Sin embargo, en el caso de Poliovirus (3.11A), esta región se encuentra deshidratada a lo largo de toda la simulación. Al estudiar la dinámica en esta región, se observa que la cavidad se encuentra obstruida por las cadenas laterales de las histidinas, por lo que el cierre es estérico. Al ocurrir este cierre, las aguas entran en el poro, rodeando estos aminoácidos, como se puede ver en el mapa de densidad para $r=0.6$ nm.

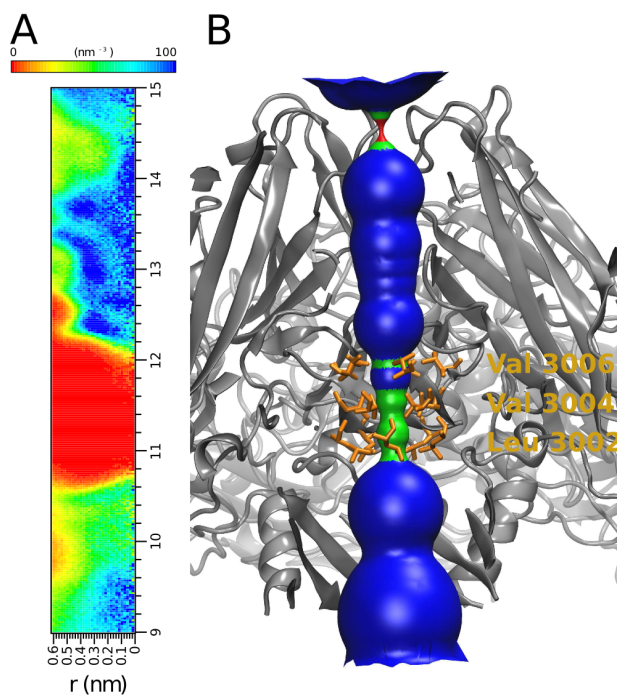


Figura 3.10: (A) Mapa de densidad de moléculas de agua en HRV16 en función del radio del poro y el radio de la cápside. (B) Estructura del poro de HRV16. La cápside se representa en *NewCrtoon* en gris. La superficie azul, verde y roja representa el radio del poro calculado con el programa Hole. Se resaltan los aminoácidos hidrofóbicos que forman al poro en la región deshidratada.

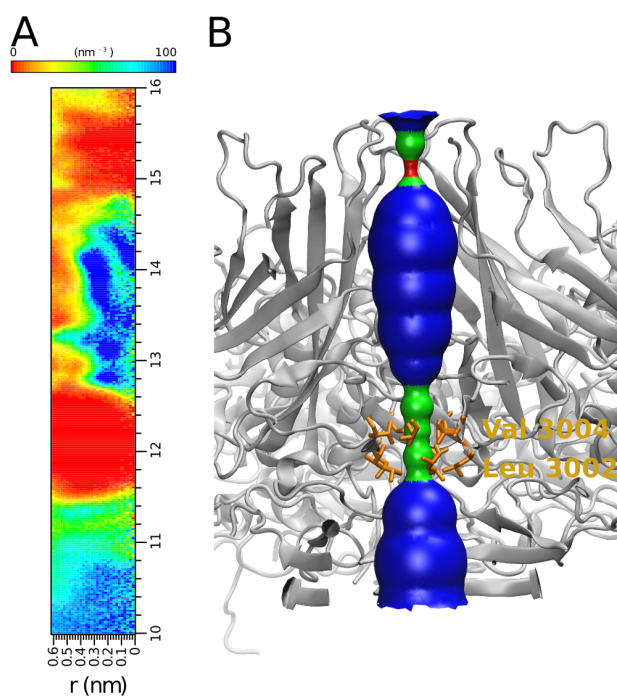


Figura 3.11: (A) Mapa de densidad de moléculas de agua en PoV en función del radio del poro y el radio de la cápside. (B) Estructura del poro de PoV. La cápside se representa en *NewCrtoon* en gris. La superficie azul, verde y roja representa el radio del poro calculado con el programa Hole. Se resaltan los aminoácidos hidrofóbicos que forman al poro en la región deshidratada.

En la figura 3.12 se observa el mapa de densidad para el HAV, también perteneciente a la familia de Picornavirus. Al igual que HRV16 y PoV, el virus de la hepatitis presenta una región en el poro que se encuentra completamente deshidratada a lo largo de toda la simulación. De igual manera, esta región coincide con la presencia de un anillo de aminoácidos no-polares que provocan la deshidratación del poro a pesar de haber espacio suficiente para el ingreso de moléculas de agua.

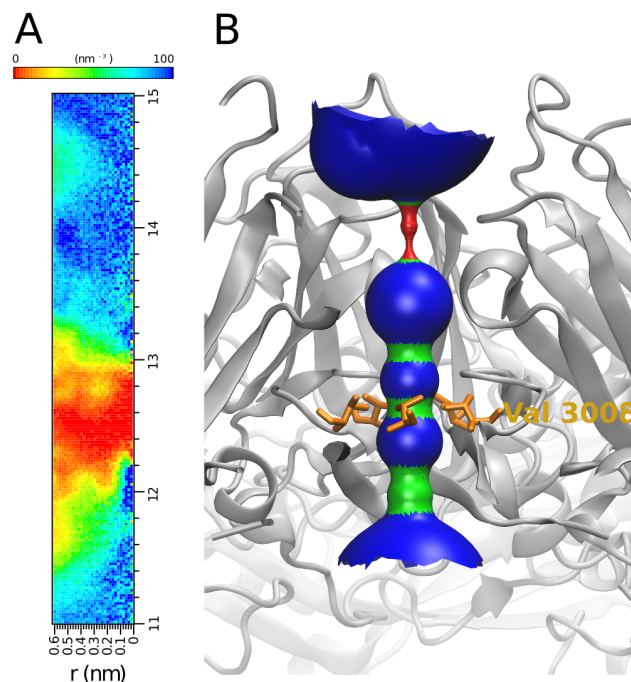


Figura 3.12: (A) Mapa de densidad de moléculas de agua en HAV en función del radio del poro y el radio de la cápside. (B) Estructura del poro de HAV. La cápside se representa en *NewCrtoon* en gris. La superficie azul, verde y roja representa el radio del poro calculado con el programa Hole. Se resaltan los aminoácidos hidrofóbicos que forman al poro en la región deshidratada.

Por otro lado, en la figura 3.13 se puede observar el poro presente en CrPV junto con el mapa de densidad. Este virus pertenece a la familia de Dicistrovirus como TrV y como era esperado presente una región deshidratada en su poro que coincide con la ubicación de tres anillos hidrofóbicos formados por Valina 1174, Leucina 3014 y Cisteina 3012.

Por último, en las figuras figuras 3.14 y 3.15 se muestran los mapas de densidad para los virus pertenecientes a la familia Secovirus. Ambos mapas muestran una completa hidratación del canal. Esto era lo esperado debido a la ausencia de aminoácidos hidrofóbicos en las regiones más angostas del poro. Debido a esto, los poros presentes en los virus pertenecientes a esta familia, deben poseer otro tipo de mecanismo de regulación de su apertura.

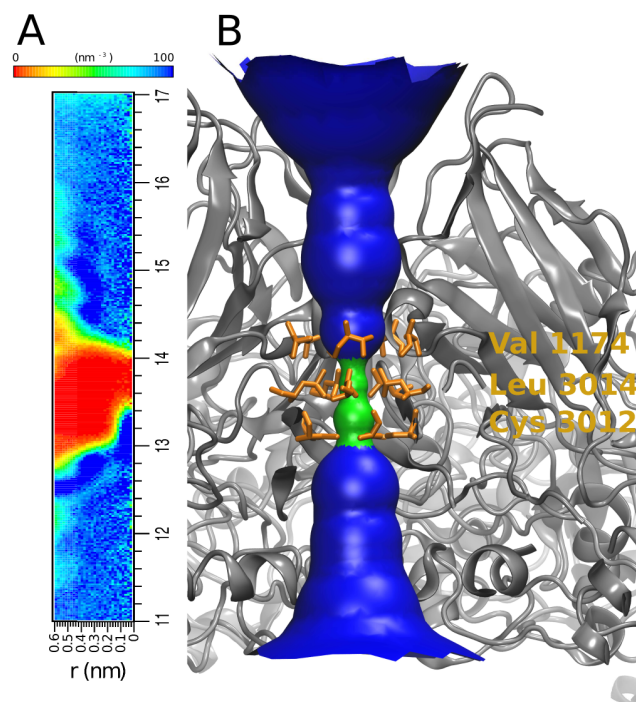


Figura 3.13: (A) Mapa de densidad de moléculas de agua en CrPV en función del radio del poro y el radio de la cápside. (B) Estructura del poro de CrPV. La cápside se representa en *NewCrtoon* en gris. La superficie azul, verde y roja representa el radio del poro calculado con el programa Hole. Se resaltan los aminoácidos hidrofóbicos que forman al poro en la región deshidratada.

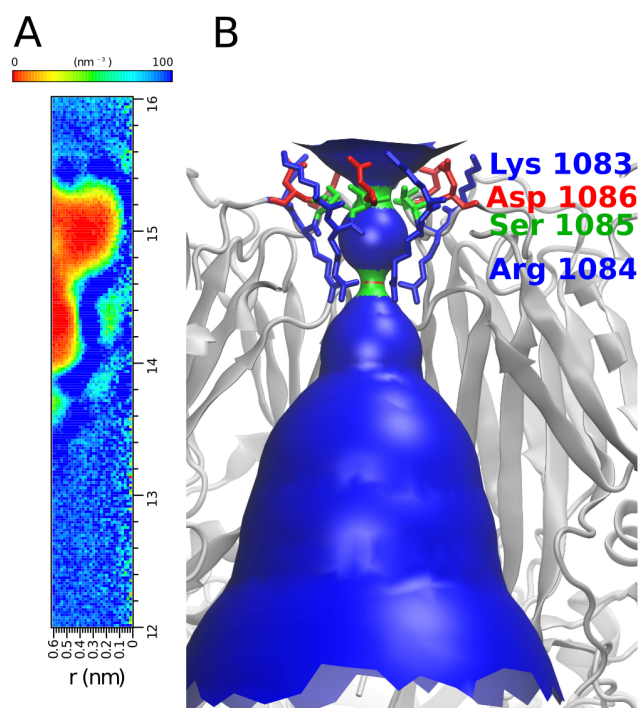


Figura 3.14: (A) Mapa de densidad de moléculas de agua en BPMV en función del radio del poro y el radio de la cápside. (B) Estructura del poro de BPMV. La cápside se representa en *NewCrtoon* en gris. La superficie azul, verde y roja representa el radio del poro calculado con el programa Hole. Se resaltan los aminoácidos que forman al poro en la región más angosta coloreados según sus características. En rojo los cargados negativamente y en azul los positivos.

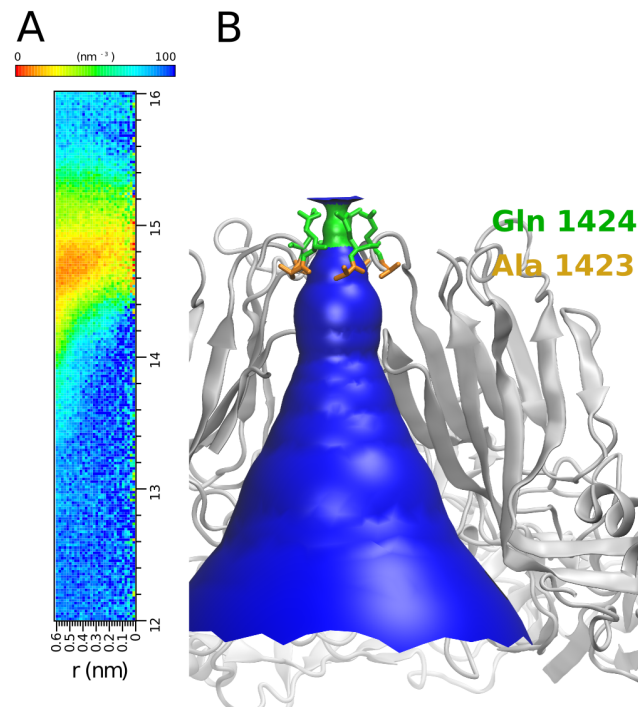


Figura 3.15: (A) Mapa de densidad de moléculas de agua en TRsV en función del radio del poro y el radio de la cápside. (B) Estructura del poro de TRsV. La cápside se representa en *NewCrtoon* en gris. La superficie azul, verde y roja representa el radio del poro calculado con el programa Hole. Se resaltan los aminoácidos que forman al poro en la región más angosta coloreados según sus características. En verde los polares y en naranja los no-polares.

A partir de lo observado podemos concluir que todos los virus Picornavirales que afectan a animales (Picornavirus y Dicistrovirus) poseen un mecanismo de regulación de la apertura del poro presente en el eje quíntuple. Este mecanismo es regulado gracias al efecto hidrofóbico producido por los aminoácidos no-polares presentes en la región más angosta del poro. Por otro lado los virus que afectan a las plantas (Secovirus) poseen un poro más pequeño que no posee aminoácidos no-polares por lo que, si existe algún mecanismo de regulación de apertura del poro, este no es el de puerta hidrofóbica.

Capítulo 4

Conclusión

A lo largo de la tesis se ha intentado dilucidar el rol de la cavidad presente en el eje de simetría quíntuple de la cápside icosaédrica del Virus del Triatoma. Particularmente, el papel que cumple como posible canal para el intercambio de partículas entre el interior del virus y su exterior. Este tipo de estructuras protéicas que regulan el paso de partículas entre dos medios son ampliamente estudiados en membranas celulares, como por ejemplo los canales iónicos, sin embargo no ha sido así en las cápsides virales. En esta tesis se confirma la existencia de canales en cápsides virales como fue propuesto, pero no comprobado, décadas atrás por *Kalko et al.* Con este fin, se identificó, a través de simulaciones computacionales, el proceso de hidratación y regulación de la apertura del canal.

Las simulaciones del pentámero de TrV permiten identificar la presencia de una puerta hidrofóbica, que regula la apertura del canal en el interior de la cavidad presente en el eje quíntuple. Esta región se encuentra ubicada a la altura de las Valinas 3012 pertenecientes a VP3. Las propiedades hidrofóbicas de las Valinas en conjunto con el reducido radio del canal en esa región, permiten crear una puerta que controla el pasaje de partículas.

A su vez, se demostró que la inserción de un ion magnesio dentro del canal cumple un rol clave en la consecuente apertura de la puerta hidrofóbica. Este ion es capaz de ingresar en el canal desde el exterior del virus y llegar hasta el nivel del anillo de aminoácidos definido por las Glutaminas 3014. Una vez ubicado en esta región, interactúa con los oxígenos de las cadenas laterales de las glutaminas y con el de las moléculas de agua del entorno, formando una estructura octaédrica estable. Esta estructura se ubica adyacente al anillo definido por las Valinas 3012 donde se encuentra la puerta hidrofóbica. La presencia de este ion logra que las moléculas de agua ingresen en la región previamente

deshidratada y, de esta manera, se logre la apertura del canal. Gracias a la presencia del ion y la apertura del canal se consigue una cadena estable de moléculas de agua que conectan de forma continua al exterior de la cápside con el interior.

Finalmente, un análisis del resto de los virus del orden Picornavirales mostró que los virus pertenecientes a la Familias Picornavirus, que infectan a vertebrados, y Dicistrovirus, que infectan a insectos, comparten la presencia de un canal en el eje de simetría quíntuple de sus cápsides, el cual regula su apertura a través de una puerta hidrofóbica formada por anillos de aminoácidos no-polares. Por otro lado, los virus pertenecientes a la Familia Secovirus, que infectan a plantas, no muestran la presencia de un mecanismo de regulación de apertura de su poro a través del mecanismo de puerta hidrofóbica.

Palabras Finales

En resumen, esta tesis tiene como objetivo responder la incógnita de: ¿Existe algún mecanismo de transporte en el eje quíntuple del Virus del Triatoma? Y de ser así, ¿De qué manera regula su apertura? y ¿Qué tipo de moléculas son capaces de atravesarlo?

En conclusión, mediante el trabajo realizado a lo largo de esta tesis, se puede afirmar que existe una región presente en el eje de simetría quíntuple de TrV que es capaz de comunicar el interior del virus con su exterior. Este canal regula su apertura a través de una puerta hidrofóbica formada por un anillo de cinco Valinas. El mismo es capaz de abrirse gracias al ingreso de un ion Mg^{+2} dentro del canal, el cual al interactuar con un anillo de glutaminas próximas a la puerta hidrofóbica, permiten el ingreso de moléculas de agua en la puerta. Estas moléculas de agua generan una cadena continua de moléculas de agua.

De estos resultados nacen nuevas interrogantes respecto de este canal: ¿Qué tipo de partículas son capaces de atravesarlo? ¿y con qué fin? Basado en estas preguntas se elaboró la hipótesis de que la cadena de aguas formado a lo largo de la puerta hidrofóbica en conjunto con el ion Mg^{+2} crean el ambiente propicio para la existencia de una cadena transportadora de protones, la cual podría causar la desestabilización de la cápside viral.

Debido a que para poder estudiar un mecanismo de transporte de protones como el Grothuss[88] es necesario observar la ruptura de enlaces covalentes, la Dinámica Molecular ya no es una herramienta adecuada para continuar el estudio. Para continuar con la

comprobación de esta hipótesis se propuso una colaboración con las Dras. Patricia Belleli, María Marta Branda y el Dr. Sergio Pantano.

En esta colaboración, en primer lugar, se simuló a la estructura del canal hidratado debido a la presencia del Mg^{+2} en conjunto con las aguas responsables de la hidratación del mismo utilizando el paquete de programas Gaussian[89] y la Teoría del Funcional Densidad (DFT). De estos resultados[90] se pudo confirmar la hipótesis de que el canal funciona como una cadena transportadora de protones mediante el mecanismo de Grotthuss. Posteriormente, se realizaron simulaciones de grano grueso utilizando el campo de fuerzas Sirah. En estas dinámicas se simuló el egreso de protones de la cápside viral colocando un exceso de carga negativa en el interior de la misma. La presencia de este exceso de carga negativa, provocó una expansión del radio de la cápside debido a una repulsión electrostática. Esta expansión causó la apertura de la cápside en los bordes que unen a los pentámeros entre sí, dejando al pentón intacto. Este fenómeno es consistente con lo observado experimentalmente en cápsides vacías[33, 34]. De este modo, la apertura de la cápside en estas regiones, daría lugar a la externalización del material genético del virus, comenzando así el proceso de infección.

De esta manera, el trabajo presentado en esta tesis, logra proponer un posible mecanismo de desensamblaje de la cápside viral del virus del Triatoma. Este mecanismo es de suma importancia durante el proceso de infección viral. El mismo consiste en la regulación del pasaje de protones a través del poro presente en el eje quintuple. Esta regulación se logra gracias a una puerta hidrofóbica en su interior, la cual controla su apertura gracias a la presencia de un ion magnesio dentro del poro. Luego, la apertura del canal provoca el flujo de protones desde el interior hacia el exterior por medio de un mecanismo tipo Grothuss cuando el virus se encuentra en un medio básico. Finalmente, el egreso de los protones, provoca un desbalance en la carga eléctrica lo cual causa la apertura de la cápside en las regiones donde se ubican los ejes dobles, permitiendo al ARN abandonar la cápside e iniciando el proceso de infección celular. De esta manera, se presenta en esta tesis por primera vez, utilizando como modelo a la estructura de TrV y por medio de la simulación computacional, el mecanismo de desensamblaje de una cápside viral.

Bibliografía

- [1] S. G. Kalko, R. Cachau, and A. Silva, “Ion channels in icosahedral virus: a comparative analysis of the structures and binding sites at their fivefold axes,” *Biophysical journal*, vol. 63, no. 4, p. 1133, 1992.
- [2] O. Muscio, J. LaTorre, and E. Scodeller, “Small nonoccluded viruses from triatomine bug *triatoma infestans* (hemiptera: Reduviidae),” *Journal of Invertebrate Pathology*, vol. 49, no. 2, pp. 218 – 220, 1987.
- [3] E. De Robertis, *Fundamentos de Biología Celular y Molecular de De Robertis*. El Ateneo, 2007.
- [4] H. Lodish, A. Berk, and S. Zipursky, *Molecular Cell Biology*. W. H. Freeman, 4th ed., 2000.
- [5] D. Forst, W. Welte, T. Wacker, and K. Diederichs, “Structure of the sucrose-specific porin scry from salmonella typhimurium and its complex with sucrose,” *Nature Structural Biology*, vol. 5, no. 1, pp. 37–46, 1998.
- [6] B. Hille, *Ionic channels of excitable membranes*. Sinauer Associates, 1984.
- [7] F. Zhu and G. Hummer, “Drying transition in the hydrophobic gate of the GLIC channel blocks ion conduction,” *Biophysical Journal*, vol. 103, no. 2, pp. 219 – 227, 2012.
- [8] R. B. Bass, P. Strop, M. Barclay, and D. C. Rees, “Crystal structure of escherichia coli mscs, a voltage-modulated and mechanosensitive channel,” *Science*, vol. 298, no. 5598, pp. 1582–1587, 2002.
- [9] A. Anishkin and S. Sukharev, “Water dynamics and dewetting transitions in the small mechanosensitive channel mscs,” *Biophysical Journal*, vol. 86, no. 5, pp. 2883 – 2895, 2004.
- [10] M. Sotomayor and K. Schulten, “Molecular dynamics study of gating in the mechanosensitive channel of small conductance mscs,” *Biophysical Journal*, vol. 87, no. 5, pp. 3050 – 3065, 2004.
- [11] O. Beckstein and M. S. P. Sansom, “A hydrophobic gate in an ion channel: the closed state of the nicotinic acetylcholine receptor,” *Physical Biology*, vol. 3, no. 2, p. 147, 2006.
- [12] H. R. Gelderblom, “Structure and Classification of Viruses,” in *Medical Microbiology. 4th edition* (S. Baron, ed.), ch. 41, Galveston (TX): University of Texas Medical Branch at Galveston, 1996.

- [13] D. Baltimore, "Expression of animal virus genomes," *Bacteriological Reviews*, vol. 35, no. 3, pp. 235–241, 1971.
- [14] D. J. Evans and J. W. Almond, "Cell receptors for picornaviruses as determinants of cell tropism and pathogenesis," *Trends in microbiology*, vol. 6, no. 5, p. 198–202, 1998.
- [15] M. G. Rossmann, "Picornavirus structure overview," in *Molecular Biology of Picornavirus*, pp. 27–38, American Society of Microbiology, 2002.
- [16] M. G. Rossmann, E. Arnold, J. W. Erickson, E. A. Frankengerger, J. P. Griffith, H.-J. Hecht, J. E. Johnson, G. Kamer, M. Luo, A. G. Mosser, R. R. Rueckert, B. Sherry, and G. Vriend, "Structure of a human common cold virus and functional relationship to other picornaviruses," *Nature*, vol. 317, pp. 145–153, Sep 1985.
- [17] P. R. Kolatkar, J. Bella, N. H. Olson, C. M. Bator, T. S. Baker, and M. G. Rossmann, "Structural studies of two rhinovirus serotypes complexed with fragments of their cellular receptor," *EMBO Journal*, vol. 18, no. 22, pp. 6249–6259, 1999.
- [18] N. H. Olson, P. R. Kolatkar, M. A. Oliveira, R. H. Cheng, J. M. Greve, A. McClelland, T. S. Baker, and M. G. Rossmann, *Structure of a human rhinovirus complexed with its receptor molecule*, vol. 90, pp. 507–511. 1993.
- [19] D. M. Belnap, B. M. McDermott, D. J. Filman, N. Cheng, B. L. Trus, H. J. Zuccola, V. R. Racaniello, J. M. Hogle, and A. C. Steven, "Three-dimensional structure of poliovirus receptor bound to poliovirus," *PNAS*, vol. 97, no. 1, pp. 73–78, 2000.
- [20] Y. He, V. D. Bowman, S. Mueller, C. M. Bator, J. Bella, X. Peng, T. S. Baker, E. Wimmer, R. J. Kuhn, and M. G. Rossmann, "Interaction of the poliovirus receptor with poliovirus," *Proceedings of the National Academy of Sciences*, vol. 97, no. 1, pp. 79–84, 2000.
- [21] L. Xing, K. Tjarnlund, B. Lindqvist, G. G. Kaplan, D. Feigelstock, R. H. Cheng, and J. M. Casanovas, "Distinct cellular receptor interactions in poliovirus and rhinoviruses," *The EMBO Journal*, vol. 19, no. 6, pp. 1207–1216, 2000.
- [22] C. Xiao, C. M. Bator, V. D. Bowman, E. Rieder, Y. He, B. Hébert, J. Bella, T. S. Baker, E. Wimmer, R. J. Kuhn, and M. G. Rossmann, "Interaction of coxsackievirus A21 with its cellular receptor, ICAM-1," *Journal of virology*, vol. 75, no. 5, pp. 2444–51, 2001.
- [23] T. J. Smith, M. J. Kremer, M. Luo, G. Vriend, E. Arnold, G. Kamer, M. G. Rossmann, M. A. McKinlay, G. D. Diana, and M. J. Otto, "The site of attachment in human rhinovirus 14 for antiviral agents that inhibit uncoating," *Science*, vol. 233, no. 4770, pp. 1286–1293, 1986.
- [24] M. A. Oliveira, R. Zhao, W.-M. Lee, M. J. Kremer, I. Minor, R. R. Rueckert, G. D. Diana, D. C. Pevear, F. J. Dutko, M. A. McKinlay, and M. G. Rossmann, "The structure of human rhinovirus 16," *Structure*, vol. 1, no. 1, pp. 51–68, 1993.

- [25] M. G. Rossmann, “Viral cell recognition and entry,” *Protein Science*, vol. 3, no. 10, pp. 1712–1725, 1994.
- [26] R. L. Crowell and L. Philipson, “Specific alterations of coxsackievirus b3 eluted from hela cells,” *J Virol*, vol. 8, pp. 509–515, Oct 1971.
- [27] K. Lonberg-Holm and B. D. Korant, “Early interaction of rhinoviruses with host cells,” *J Virol*, vol. 9, pp. 29–40, Jan 1972.
- [28] V. L. Giranda, B. A. Heinz, M. A. Oliveira, I. Minor, K. H. Kim, P. R. Kolatkar, M. G. Rossmann, and R. R. Rueckert, “Acid-induced structural changes in human rhinovirus 14: possible role in uncoating.,” *Proc Natl Acad Sci U S A*, vol. 89, pp. 10213–10217, Nov 1992.
- [29] R. McKenna, D. Xia, P. Willingmann, L. L. Hagg, S. Krishnaswamy, M. G. Rossmann, N. H. Olson, T. S. Baker, and N. L. Incardona, “Atomic structure of single-stranded dna bacteriophage [phi]x174 and its functional implications,” *Nature*, vol. 355, pp. 137–143, Jan 1992.
- [30] E. Hendry, H. Hatanaka, E. Fry, M. Smyth, J. Tate, G. Stanway, J. Santti, M. Maaronen, T. Hyypä, and D. Stuart, “The crystal structure of coxsackievirus a9: new insights into the uncoating mechanisms of enteroviruses,” *Structure*, vol. 7, pp. 1527–1538, 2017/07/07 1999.
- [31] C. E. Fricks and J. M. Hogle, “Cell-induced conformational change in poliovirus: externalization of the amino terminus of vp1 is responsible for liposome binding.,” *J Virol*, vol. 64, pp. 1934–1945, May 1990.
- [32] M. S. Smyth and J. H. Martin, “Picornavirus uncoating,” *Mol Pathol*, vol. 55, pp. 214–219, Aug 2002.
- [33] D. Garriga, A. Pickl-Herk, D. Luque, J. Wruss, J. R. Castón, D. Blaas, and N. Verdaguer, “Insights into minor group rhinovirus uncoating: The x-ray structure of the hrv2 empty capsid,” *PLoS Pathog*, vol. 8, p. e1002473, Jan 2012.
- [34] M. Bostina, H. Levy, D. J. Filman, and J. M. Hogle, “Poliovirus RNA is released from the capsid near a twofold symmetry axis,” *J. Virol.*, vol. 85, pp. 776–783, Jan 2011.
- [35] S. Valles, Y. Chen, A. Firth, D. Guérin, Y. Hashimoto, S. Herrero, J. de Miranda, E. Ryabov, and I. R. Consortium, “Ictv virus taxonomy profile: Dicistroviridae,” *Journal of General Virology*, vol. 98, pp. 355–356, 2017.
- [36] G. Squires, J. Pous, J. Agirre, G. S. Rozas-Dennis, M. D. Costabel, G. A. Marti, J. Navaza, S. Bresanelli, D. M. A. Guérin, and F. A. Rey, “Structure of the *Triatoma virus* capsid,” *Acta Crystallographica Section D*, vol. 69, no. 6, pp. 1026–1037, 2013.

- [37] J. Tate, L. Liljas, P. Scotti, P. Christian, T. Lin, and J. E. Johnson, “The crystal structure of cricket paralysis virus: the first view of a new virus family,” *Nature structural biology*, vol. 6, no. 8, pp. 765–774, 1999.
- [38] S. Cherry and N. Perrimon, “Entry is a rate-limiting step for viral infection in a drosophila melanogaster model of pathogenesis,” *Nature Immunology*, vol. 5, no. 1, pp. 81–87, 2004.
- [39] S. Cherry and N. Silverman, “Host-pathogen interactions in drosophila: new tricks from an old friend,” *Nature Immunology*, vol. 7, no. 9, pp. 911–917, 2006.
- [40] J. Agirre, G. Goret, M. LeGoff, R. Sánchez-Eugenia, G. A. Marti, J. Navaza, D. M. A. Guérin, and E. Neumann, “Cryo-electron microscopy reconstructions of triatoma virus particles: a clue to unravel genome delivery and capsid disassembly,” *Journal of General Virology*, vol. 94, no. 5, pp. 1058–1068, 2013.
- [41] J. Snijder, C. Uetrecht, R. Rose, R. Sanchez-Eugenia, G. Marti, J. Agirre, D. Guerin, G. Wuite, A. Heck, and W. Roos, “Probing the biophysical interplay between a viral genome and its capsid,” *Nature Chemistry*, vol. 5, pp. 502–509, 6 2013.
- [42] A. R. Jr, A. Rassi, and J. A. Marin-Neto, “Chagas disease,” *The Lancet*, vol. 375, no. 9723, pp. 1388 – 1402, 2010.
- [43] W. Kauzmann, “Some factors in the interpretation of protein denaturation,” *Advances in Protein Chemistry*, vol. 14, pp. 1–63, 1959.
- [44] G. Hummer, J. C. Rasaiah, and J. P. Noworyta, “Water conduction through the hydrophobic channel of a carbon nanotube,” *Nature*, vol. 414, pp. 188–190, Nov 2001.
- [45] K. A. Dill, S. Bromberg, and D. C. Stigter, *Molecular Driving Forces : Statistical Thermodynamics in Chemistry and Biology*. Garland Science, 2003.
- [46] D. Chandler, “Two faces of water,” *Nature*, vol. 417, pp. 491–491, 2002.
- [47] D. Chandler, “Interfaces and the driving force of hydrophobic assembly,” *Nature*, vol. 437, pp. 640–647, 2005.
- [48] M. Born and J. R. Oppenheimer, “Zur quantentheorie der molekeln,” *Annalen der Physik*, vol. 389, no. 20, pp. 457–484, 1927.
- [49] F. Schwalb, *Quantum Mechanics*. Springer, 4th ed., 2007.
- [50] J. W. Ponder and D. A. Case, “Force fields for protein simulations,” in *Protein Simulations*, vol. 66 of *Advances in Protein Chemistry*, pp. 27–85, Academic Press, 2003.

- [51] G. A. Kaminski, R. A. Friesner, J. Tirado-Rives, and W. L. Jorgensen, "Evaluation and reparameterization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides," *The Journal of Physical Chemistry B*, vol. 105, no. 28, pp. 6474–6487, 2001.
- [52] C. Oostenbrink, A. Villa, A. E. Mark, and W. F. Van Gunsteren, "A biomolecular force field based on the free enthalpy of hydration and solvation: The gromos force-field parameter sets 53a5 and 53a6," *Journal of Computational Chemistry*, vol. 25, no. 13, pp. 1656–1676, 2004.
- [53] S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman, and A. H. de Vries, "The MARTINI force field: Coarse grained model for biomolecular simulations," *The Journal of Physical Chemistry B*, vol. 111, no. 27, pp. 7812–7824, 2007.
- [54] L. Darré, M. R. Machado, and S. Pantano, "Coarse-grained models of water," *Wiley Interdisciplinary Reviews: Computational Molecular Science*, vol. 2, no. 6, pp. 921–930, 2012.
- [55] L. Verlet, "Computer experiments on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules," *Physical Review*, vol. 159, no. 1, pp. 98–103, 1967.
- [56] P. P. Ewald, "Die berechnung optischer und elektrostatischer gitterpotentiale," *Annalen der Physik*, vol. 369, no. 3, pp. 253–287, 1921.
- [57] M. Deserno and C. Holm, "How to mesh up ewald sums. I. a theoretical and numerical comparison of various particle mesh routines," *The Journal of Chemical Physics*, vol. 109, no. 18, pp. 7678–7693, 1998.
- [58] T. Darden, D. York, and L. Pedersen, "Particle Mesh Ewald: An NlogN method for Ewald sums in large systems," *The Journal of Chemical Physics*, vol. 98, no. 12, pp. 10089–10092, 1993.
- [59] R. W. Hockney and J. W. Eastwood, *Computer Simulations Using Particles*. Taylor & Francis, Inc., 1988.
- [60] A. R. Leach, *Molecular Modelling: Principles and Applications*. Prentice Hall, 2001.
- [61] J. P. Ryckaert, G. Ciccotti, and H. J. Berendsen, "Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes," *Journal of Computational Physics*, vol. 23, no. 3, pp. 327–341, 1977.
- [62] S. Miyamoto and P. A. Kollman, "SETTLE: an analytical version of the SHAKE and RATTLE algorithm for rigid water models," *Journal of Computational Chemistry*, vol. 13, no. 8, pp. 952–962, 1992.
- [63] B. Hess, H. Bekker, and H. J. Berendsen, "LINCS: A linear constraint solver for molecular simulations," *Journal of Computational Chemistry*, vol. 18, no. 12, pp. 1463–1472, 1997.

- [64] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, “Molecular dynamics with coupling to an external bath,” *The Journal of Chemical Physics*, vol. 81, no. 8, pp. 3684–3690, 1984.
- [65] G. Bussi, D. Donadio, and M. Parrinello, “Canonical sampling through velocity rescaling,” *The Journal of Chemical Physics*, vol. 126, no. 1, p. 014101, 2007.
- [66] M. Parrinello and A. Rahman, “Polymorphic transitions in single crystals: a new molecular dynamics method,” *Journal of Applied Physics*, vol. 52, no. 12, pp. 7182–7190, 1981.
- [67] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. J. C. Berendsen, “Gromacs: Fast, flexible, and free,” *Journal of Computational Chemistry*, vol. 26, no. 16, pp. 1701–1718, 2005.
- [68] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, and K. Schulten, “Scalable molecular dynamics with namd,” *Journal of Computational Chemistry*, vol. 26, no. 16, pp. 1781–1802, 2005.
- [69] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, “Charmm: A program for macromolecular energy, minimization, and dynamics calculations,” *Journal of Computational Chemistry*, vol. 4, no. 2, pp. 187–217, 1983.
- [70] O. Smart, J. Neduvélil, X. Wang, B. Wallace, and M. Sansom, “HOLE: a program for the analysis of the pore dimensions of ion channel structural models,” *Journal of Molecular Graphics*, vol. 14, no. 6, pp. 354–360, 1996.
- [71] J. G. Kirkwood, “Statistical mechanics of fluid mixtures,” *The Journal of Chemical Physics*, vol. 3, no. 5, pp. 300–313, 1935.
- [72] B. Roux, T. Allen, S. Bernèche, and W. Im, “Theoretical and computational models of biological ion channels,” *Quarterly Reviews of Biophysics*, vol. 37, no. 1, p. 15–103, 2004.
- [73] D. Chandler, *Introduction to Modern Statistical Mechanics*. Oxford University Press, 1987.
- [74] G. Torrie and J. Valleau, “Nonphysical sampling distributions in Monte Carlo free-energy estimation - Umbrella sampling,” *Journal of Computational Physics*, vol. 23, pp. 187–199, Feb. 1977.
- [75] J. P. Valleau and G. M. Torrie, *A Guide to Monte Carlo for Statistical Mechanics: 2. Byways*, pp. 169–194. Boston, MA: Springer US, 1977.
- [76] S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, and P. A. Kollman, “THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method,” *J. Comput. Chem.*, vol. 13, no. 8, pp. 1011–1021, 1992.
- [77] M. Carrillo-Tripp, C. M. Shepherd, I. A. Borelli, S. Venkataraman, G. Lander, P. Natarajan, J. E. Johnson, C. L. Brooks, and V. S. Reddy, “VIPERdb2: an enhanced and web API enabled relational database for structural virology,” *Nucleic Acids Research*, vol. 37, no. 1, pp. D436–D442, 2009.

- [78] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, and J. Hermans, *Interaction Models for Water in Relation to Protein Hydration*, pp. 331–342. Dordrecht: Springer Netherlands, 1981.
- [79] L. Darré, A. Tek, M. Baaden, and S. Pantano, “Mixing atomistic and coarse grain solvation models for md simulations: Let wt4 handle the bulk,” *Journal of Chemical Theory and Computation*, vol. 8, no. 10, pp. 3880–3894, 2012.
- [80] H. C. Gonzalez, L. Darré, and S. Pantano, “Transferable mixing of atomistic and coarse-grained water models,” *The Journal of Physical Chemistry B*, vol. 117, no. 46, pp. 14438–14448, 2013.
- [81] Schrödinger, LLC, “The PyMOL molecular graphics system, version 1.8.” November 2015.
- [82] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, “The missing term in effective pair potentials,” *The Journal of Physical Chemistry*, vol. 91, no. 24, pp. 6269–6271, 1987.
- [83] O. Beckstein and M. S. Sansom, “Liquid–vapor oscillations of water in hydrophobic nanopores,” *Proceedings of the National Academy of Sciences*, vol. 100, no. 12, pp. 7063–7068, 2003.
- [84] J. C. Rasaiah, S. Garde, and G. Hummer, “Water in nonpolar confinement: from nanotubes to proteins and beyond,” *Annu. Rev. Phys. Chem.*, vol. 59, pp. 713–740, 2008.
- [85] P. Aryal, M. S. Sansom, and S. J. Tucker, “Hydrophobic gating in ion channels,” *Journal of molecular biology*, vol. 427, no. 1, pp. 121–130, 2015.
- [86] J. S. Hub, B. L. de Groot, and D. van der Spoel, “g_wham—A free weighted histogram analysis implementation including robust error and autocorrelation estimates,” *Journal of Chemical Theory and Computation*, vol. 6, no. 12, pp. 3713–3720, 2010.
- [87] W. Humphrey, A. Dalke, and K. Schulten, “VMD – Visual Molecular Dynamics,” *Journal of Molecular Graphics*, vol. 14, pp. 33–38, 1996.
- [88] C. J. T. De Grotthuss, *Mémoire sur la décomposition de l’eau: et des corps qu’elle tient en dissolution à l’aide de l’électricité galvanique*. 1805.
- [89] M. J. Frisch *et al.*, “Gaussian 09, revision b.01,” *Gaussian 09, Revision B.01*, Gaussian, Inc., Wallingford CT, 2009.
- [90] J. F. Viso, P. Belelli, M. Machado, H. González, S. Pantano, M. J. Amundarain, F. Zamarreño, M. M. Branda, D. M. Guerin, and M. D. Costabel, “Multiscale modelization in a small virus: Mechanism of proton channeling and its role in triggering capsid disassembly,” *PLOS Computational Biology*, 2018.

Publicaciones Científicas

1. **Multiscale modelization in a small virus: Mechanism of proton channeling and its role in triggering capsid disassembly.** Viso J.F., Belelli P., Machado M., González H., Pantano S., Amundarain M.J., Zamarreño F., Branda M.M., Guérin D.M.A. and Costabel M.D. *PLOS Computational Biology*, 2018.