

DIVISIÓN DE CIENCIAS BÁSICAS E INGENIERÍA  
POSGRADO EN CIENCIAS (FÍSICA)

**“MODELO DE LIBERACIÓN DEL VIH-1 DE LA  
MEMBRANA PLASMÁTICA, AUTOENSAMBLAJE DE SU  
RED Y DE CÁPSIDES ICOSAÉDRICAS”**

**TESIS**  
Que para obtener el grado de  
**DOCTOR EN CIENCIAS (FÍSICA)**

**PRESENTA:**  
M. en C. Brian Ignacio Machorro Martínez  
2152801022  
brian31416@gmail.com

**DIRECTOR DE TESIS:**  
Dr. Gustavo Adolfo Chapela Castañares

**JURADO**  
**Presidente:**  
Dr. Rafael Arturo Zubillaga Luna

**Secretario:**  
Dr. Gustavo Adolfo Chapela Castañares

**Vocal:**  
Dr. Jesús Enrique Díaz Herrera  
Dr. José Elías Pérez López  
Dr. Alejandro Gil Villegas Montiel

Iztapalapa, Ciudad de México  
a 6 de marzo de 2023





# Resumen

El VIH es un virus que, hasta la fecha, sigue infectando a miles de personas alrededor del mundo. Actualmente una persona infectada no tiene como único término la muerte, como era hasta hace algunas décadas, sino que puede llevar una vida normal con tratamientos de antiretrovirales. Esto se ha logrado gracias al amplio estudio de este virus. Sin embargo, aunque se conozca en términos generales como se lleva a cabo su ciclo de replicación, los detalles de este proceso siguen siendo desconocidos. En este trabajo, a partir de simulaciones computacionales de dinámica molecular, con modelos moleculares bastante sencillos, se indaga en el proceso de autoensamblaje del VIH. En primer lugar, se replicó la formación de una partícula vírica a partir de sus unidades estructurales. Esto incluye el ensamblaje de los bloques constructores del virus sobre un modelo de membrana lipídica, su involucramiento con ella y su separación de la membrana. Justo como el VIH y otros virus proceden para replicarse. En segundo lugar, se ensambló la red hexagonal que forma el VIH, a partir de sus unidades proteicas, en su proceso de multimerización sobre la parte interna de la membrana lipídica de la célula infectada. Esta red fue ensamblada sin la presencia de elementos como la membrana lipídica, entre otros, que sirven como andamiajes y que intervienen en este proceso del virus. Para este fin se usó un modelo molecular que se basa en la proteína Gag que sirve como bloque constructor (unidad proteica) del virus del VIH. Las principales interacciones atractivas de la proteína Gag son modeladas de dos maneras: con potenciales de pozo cuadrado y con potenciales de Lennard-Jones. En ambos casos se obtuvieron buenos resultados. Finalmente, variando ligeramente el modelo molecular para ensamblar la red, se autoensamblaron las cápsidas icosaédricas T=1, T=3 y T=4, las cuales son estructuras de forma esférica que los virus usan para resguardar su material genético. Esto se logró tan solo con variar un parámetro angular  $\phi$  de la molécula y sin la necesidad de modificar las reglas de interacción entre las mismas.



# Índice general

<b>Resumen</b>	<b>I</b>
<b>1. Introducción</b>	<b>1</b>
1.1. Autoensamblaje . . . . .	1
1.2. Virus . . . . .	2
<b>2. VIH</b>	<b>5</b>
2.1. Origen del VIH . . . . .	5
2.2. Estructura y ciclo de replicación del VIH-1 . . . . .	6
2.3. Estructura de Gag y GagPol . . . . .	9
2.4. Liberación del virión de la membrana plasmática . . . . .	11
2.4.1. Membranas celulares . . . . .	11
2.4.2. Proceso de liberación . . . . .	12
2.5. Red hexagonal en la etapa inmadura . . . . .	13
2.6. Simulación molecular: liberación de viriones y redes abiertas . . . . .	15
2.6.1. Liberación de estructuras tipo cápside de membranas . . . . .	16
2.6.2. Autoensamblaje de redes abiertas . . . . .	17
<b>3. Cápsides icosaédricas</b>	<b>19</b>
3.1. Clasificación de las cápsides icosaédricas . . . . .	19
3.2. Simulación de cápsides icosaédricas . . . . .	22
<b>4. Modelos y Potenciales</b>	<b>23</b>
4.1. Potenciales de interacción . . . . .	23
4.2. Modelos moleculares . . . . .	26
4.2.1. Red hexagonal del VIH-1 en 2D y 3D . . . . .	27

4.2.1.1.	Reglas de interacción . . . . .	29
4.2.2.	Cápsides icosaédricas . . . . .	30
4.2.3.	Liberación del virión de una membrana de doble capa . . . . .	31
4.2.3.1.	Reglas de interacción del sistema Gag-Membrana . . . . .	32
<b>5.</b>	<b>Métodos de simulación</b>	<b>35</b>
5.1.	Física estadística: Conceptos básicos . . . . .	35
5.1.1.	Hipótesis ergódica . . . . .	36
5.2.	Monte Carlo . . . . .	38
5.2.1.	Método de Metropolis . . . . .	39
5.2.2.	Condiciones de frontera y mínima imagen . . . . .	41
5.3.	Dinámica Molecular . . . . .	43
5.3.1.	Integración de las ecuaciones de movimiento: Velocity Verlet . . . . .	44
5.3.2.	Unidades reducidas . . . . .	46
5.4.	Dinámica Molecular de Potenciales Discontinuos . . . . .	47
5.4.1.	Descripción del método . . . . .	47
5.4.2.	Calendario de eventos . . . . .	51
<b>6.</b>	<b>Resultados</b>	<b>55</b>
6.1.	Potencial discontinuo . . . . .	55
6.1.1.	Liberación de un virión de una membrana de doble capa . . . . .	55
6.1.1.1.	Membrana de doble capa . . . . .	55
6.1.1.2.	Sistema Gag-Membrana . . . . .	60
6.1.2.	Red hexagonal del VIH-1 . . . . .	63
6.1.2.1.	Red 2D . . . . .	63
6.1.2.2.	Red 3D . . . . .	66
6.2.	Potencial continuo . . . . .	71
6.2.1.	Rígido vs vibrante . . . . .	71
6.2.2.	Red hexagonal del VIH-1 en 3D . . . . .	74
6.2.2.1.	Sin interacción 2-Fold . . . . .	74
6.2.2.2.	Con interacción 2-Fold . . . . .	75
6.2.3.	Cápsides icosaédricas . . . . .	78
6.2.3.1.	Estabilidad térmica . . . . .	82

6.2.4. Cápsides no-icosaédricas . . . . .	83
<b>7. Conclusiones</b>	<b>87</b>
7.1. Perspectivas . . . . .	89
<b>Bibliografía</b>	<b>91</b>



# Capítulo 1

## Introducción

### 1.1. Autoensamblaje

El auto ensamblaje es un proceso de construcción de estructuras en 2D o 3D formadas por unas cuantas o miles de unidades estructurales llamadas bloques de construcción. Estas unidades estructurales pueden ser partículas, moléculas o algún otro componente que pueda servir como unidad fundamental en el ensamblaje. Los bloques de construcción pueden ser de una o varias especies las cuales, al interactuar entre ellas, son capaces de ensamblar estructuras que pueden tener alguna periodicidad geométrica, como una red o un arreglo cristalino, o por el contrario no tener algún arreglo definido como una membrana lipídica. Las interacciones que median el proceso del autoensamblaje son de diversos tipos. A nivel molecular intervienen interacciones como las de Van der Waals, electrostáticas, hidrofóbicas o puentes de hidrógeno. En un nivel mesoscópico las interacciones pueden ser la atracción gravitatoria, capilaridad, campos eléctricos y magnéticos o interacciones entrópicas.

La clave del autoensamblaje se encuentra en la selección y diseño de los bloques de construcción para formar estructuras deseadas con funciones específicas para aplicaciones de interés. En el área de la nanotecnología y diseño de materiales el autoensamblaje es ampliamente usado ya que mediante este proceso se pueden obtener materiales con propiedades fotónicas<sup>1</sup> y plasmónicas<sup>2</sup> o puede ser usado en la creación de semiconductores.<sup>3</sup> El autoensamblaje de redes abiertas ha sido de interés debido a las aplicaciones que pueden llegar a tener estas redes. Ejemplo de esto es la red Kagome.<sup>4</sup> Esta red está formada por una combinación de triángulos y hexágonos y ha sido ensamblada con distintos mate-

riales como compuestos metal-orgánicos,<sup>5</sup> partículas coloidales<sup>6</sup> o cristales líquidos,<sup>7</sup> esto debido a sus propiedades magnéticas.<sup>4,8</sup> En el área de la síntesis química se desarrollan diversos métodos para, a partir de la combinación de compuestos orgánicos con distintos metales, obtener macrociclos que ensamblen redes con poros de distintos tamaños y formas. Los macrociclos metal-orgánicos MOMs, junto con los marcos metal-orgánicos MOFs que forman estructuras cristalinas, tienen aplicaciones de almacenaje de energía,<sup>9</sup> sensores químicos,<sup>10,11</sup> captura de moléculas<sup>12</sup> o entrega de medicamentos.<sup>5,13</sup>

El autoensamblaje además de sus amplias aplicaciones puede servir para entender procesos biológicos ya que la vida misma está regida por procesos químicos y físicos que desencadenan procesos de autoensamblaje para su desarrollo. El autoensamblaje además de estar presente en la formaciones de bicapas lipídicas<sup>14</sup> se encuentra en el plegamiento y agregación de proteínas<sup>15</sup> o en los filamentos de actina<sup>16</sup> que se ensamblan cumpliendo distintos papeles en función de las necesidades de la célula, entre otras funciones. Desarrollar modelos que imiten algún proceso biológico de autoensamblaje ayuda a entenderlo y posiblemente controlarlo. Esto debido a que al usar un modelo de los bloques constructores se pueden probar fácilmente distintos parámetros para encontrar aquellos que determinan el proceso de autoensamblaje del sistema. Los virus, sin entrar en el debate de si están vivos o no, son un ejemplo de que mientras más se entienda su proceso de replicación más formas de combatirlos se tendrán.

## 1.2. Virus

El virus del tabaco fue el primero en ser descubierto a finales del siglo XIX por Ivanovsky y Beijerinck de manera independiente.<sup>17</sup> Ivanovsky en 1882 demostró que el agente infeccioso, al que denominó como “filter-passing”, era de menor tamaño que una bacteria o un hongo.<sup>18</sup> En 1898 Beijerinck llamo “líquido vivo infeccioso” al virus del tabaco concluyendo que el agente infeccioso se reproducía así mismo en las plantas enfermas.<sup>17,18</sup> No fue hasta la década de los años treinta que se empezó a conocer el aspecto y la naturaleza química de los virus gracias al microscopio electrónico. Actualmente se define a un virus como un organismo extracelular con un ciclo de vida parasitario intracelular.<sup>18</sup> Esto quiere decir que los virus pueden estar de manera libre fuera de una célula, pero necesitan entrar en alguna (infectarla) para replicarse.



Los virus usan ácido nucleico para guardar su información genética. Su genoma se encuentra rodeado por una estructura proteica que lo protege llamada cápside. La cápside está ensamblada por subunidades llamadas capsómeros que están compuestos por una o varias proteínas. Los virus se pueden clasificar dependiendo del ácido nucleico que utilicen para guardar su genoma ya que pueden hacerlo con ADN o ARN y en ambos casos puede ser de una cadena o doble cadena. También se puede caracterizar un virus por la simetría de su cápside, siendo la simetría helicoidal y la icosaédrica las más comunes. También se pueden encontrar virus con cápsides alargadas o cónicas, incluso virus como los bacteriófagos (que infectan bacterias) están compuestos de una cabeza icosaédrica y una cola helicoidal que usan para anclarse a la célula infectada e inyectar su material genético. Además de la cápside algunos virus se encuentran envueltos por una membrana lipídica que les da una protección extra. Esta membrana la toman de la membrana plasmática al liberarse de la célula infectada. Los tamaños que pueden tener los virus van de los 20 a los 300 nanómetros, pero hay algunos que pueden alcanzar hasta los 900 nanómetros.<sup>18</sup> Por regla general mientras más largo sea su ARN o ADN más grande y complejo será el virus. Es clara la gran variedad de virus que existen con distintos tamaños y formas y han ido evolucionado para infectar a todo tipo de células ya sean de animales, hongos, plantas, bacterias e incluso pueden parasitar a otros virus.<sup>19</sup>

Los virus cuando están fuera de la célula, en busca de alguna otra para infectarla, tienen la denominación de partícula vírica o virión.<sup>20</sup> A la célula infectada se le denomina célula hospedera y el virus puede tomar dos caminos después de la infección. El primero es que el virus redireccione la maquinaria de la célula hospedera para replicarse y formar nuevos virus que se liberan de la célula. En este proceso la célula puede morir debido a una gran producción y liberación de nuevos virus o puede mantenerse viva por una producción menor de virus. El segundo camino que toma un virus es el de la latencia.<sup>18</sup> En este caso solo una parte de su genoma se mantiene activo, lo cual impide su replicación. Esta latencia puede ser una característica del propio virus o puede ser causado por la célula. Otra manera de que un virus se mantenga latente se da cuando el genoma del virus se integra al de la célula, permaneciendo inactivo hasta que factores externos, desconocidos en algunos virus, lo reactiven.<sup>18</sup>

El ciclo de replicación de un virus puede variar dependiendo de que tipo de célula infecta o si es un virus de ADN o ARN, pero en general todos siguen los siguientes cinco

pasos:<sup>20</sup>

1. Unión del virión a la célula hospedera
2. Penetración del ácido nucleico del virión a la célula hospedera
3. Replicación del ácido nucleico del virus y síntesis de proteínas estructurales del virus por medio de la maquinaria de la célula hospedera
4. Ensamblado de nuevos virus (ensamblaje de la cápside y empaquetamiento del genoma vírico)
5. Liberación de los virus (viriones) de la célula hospedera

Este ciclo de replicación, mostrado en la figura 1.1, desde que se infecta una célula y se liberan nuevos virus puede tardar de entre 20 a 60 minutos en virus bacteriófagos y de entre 8 a 40 horas en la mayoría de virus animales.<sup>20</sup>

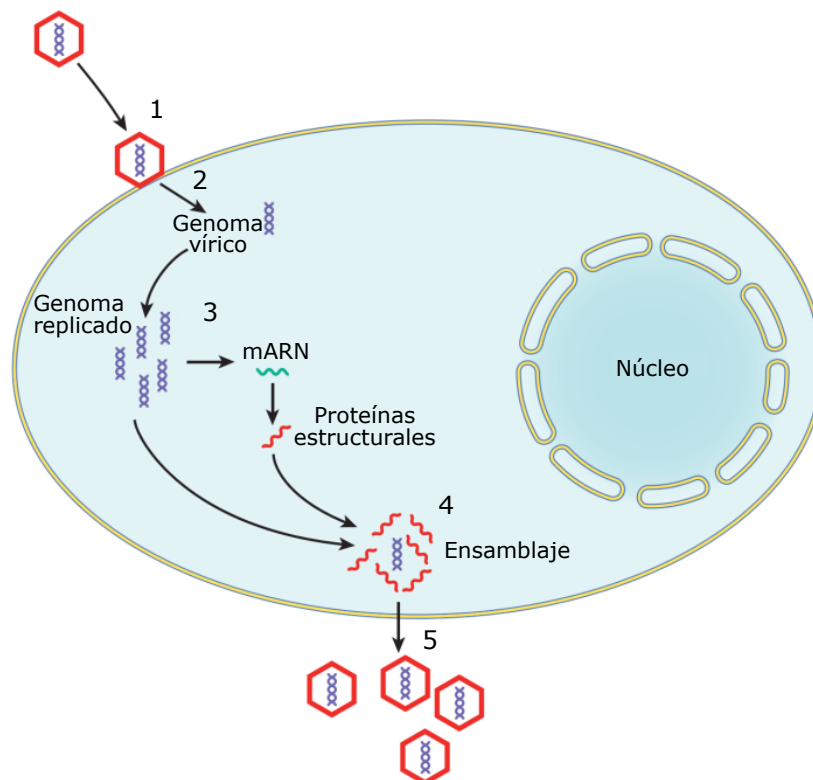


Figura 1.1: El ciclo de replicación de los virus consta de 5 pasos: 1. Anclaje a la célula hospedera. 2. Penetración del genoma vírico. 3. Replicación del ácido nucleico y síntesis de proteínas estructurales. 4. Ensamblaje de la cápside y empaquetamiento del genoma vírico. 5. Liberación de nuevos viriones. Figura modificada de Viruses: Biology, Applications and Control.<sup>18</sup>

# Capítulo 2

## VIH

Este trabajo se enfoca en el virus del VIH por lo que será a partir de este virus que se creen modelos moleculares sencillos para alcanzar dos objetivos. El primero es replicar el ensamblaje de un virión sobre una membrana lipídica. El segundo es ensamblar la red que el VIH forma en una de sus etapas de replicación. A continuación se hablará con más detalle del virus del VIH para entender en que consisten estos objetivos.

El virus de inmunodeficiencia humana (VIH) ataca a los linfocitos T CD4 encargados de activar el sistema inmunitario. Al infectarse del virus este puede mantenerse latente por varios años sin presentar síntomas.<sup>20,21</sup> Una vez que se activa el virus este provoca el debilitamiento del sistema inmune, lo que permite que una variedad de infecciones y cánceres ataquen al organismo sin la posibilidad de defenderse. Cuando se desarrollan estas enfermedades oportunistas se dice que se tiene el síndrome de inmunodeficiencia adquirida (SIDA).<sup>20,22</sup>

### 2.1. Origen del VIH

El VIH es un virus que se clasifica en dos tipos: VIH-1 y VIH-2. El VIH-1 es el causante de la pandemia SIDA, mientras que el VIH-2 es un virus que provoca menos enfermedades y es menos infeccioso. El VIH tiene su origen en los simios africanos. El VIH-1 proviene de los chimpancés y el VIH-2 proviene de los monos.<sup>21</sup> Se cree que alrededor de 1900 se dio el salto inter-especies del virus debido a las costumbres de caza y uso como alimento del chimpancé. Durante las siguientes décadas el virus se esparció lentamente por el continente Africano y alrededor de 1966 el VIH-1 pasó del Congo hacia Haití desde donde, alrededor

de 1969, se diseminó alrededor del mundo entrando a Estados Unidos y a Europa.<sup>23</sup>

En 1981 en Estados Unidos se detectaron los primeros casos de SIDA. Unos años más tarde, en 1983, el equipo de Luc Mongtanier aisló por primera vez el virus del VIH-1<sup>24</sup> y un año después el equipo de Roberto Gallo identificó a este virus como el causante del SIDA.<sup>22</sup> A mediados de los 80's se empezaron a realizar estudios de antirretrovirales (fármacos que impiden la replicación del virus dentro del organismo) para combatir el VIH<sup>25</sup> y en 1987 se aprobó la azidotimidina (AZT) como el primer fármaco contra el VIH-1.<sup>26</sup> Actualmente el desarrollo de una vacuna preventiva contra el VIH-1 se encuentra en su tercera etapa de estudios clínicos en personas de Norte América, América Latina y Europa. Esta vacuna al ser preventiva ayudaría a las personas no infectadas a desarrollar anticuerpos contra muchas de las cepas del virus VIH-1 para evitar su contagio.

## 2.2. Estructura y ciclo de replicación del VIH-1

El VIH-1 es un retrovirus, esto quiere decir que su genoma lo guarda en ARN de una sola cadena el cual necesitará pasarlo a ADN para poder llevar a cabo su ciclo de replicación. En la figura 2.1 se muestra la estructura del virión maduro del VIH-1, listo para infectar una célula. En la figura se observa que el virus guarda dos copias de su ARN las cuales están recubiertas por una proteína llamada nucleocápside. Las dos cadenas de ARN están protegidas por una cápside cónica.<sup>27,28</sup> A su vez, la cápside está rodeada por una membrana lipídica, tomada de la célula infectada al salir de ella. Pegada a la parte interna de la membrana lipídica se encuentra una proteína llamada matriz, de la cual se explicará su función más adelante. Por la parte externa de la membrana se encuentran incrustadas glucoproteínas de envoltura denominadas Env. Estas glucoproteínas Env son las encargadas de reconocer a una nueva célula para infectarla, se unen a la célula y promueven la fusión de las membranas del virus y de la célula para introducir el ARN del virus. Algunos virus como el VIH-1 llevan consigo enzimas necesarias (Transcriptasa inversa, integrasa y proteasa) para llevar a cabo su ciclo de replicación. Por ejemplo, la transcriptasa inversa, también llamada retrotranscriptasa, es la encargada de pasar el ARN de una cadena del virus a ADN de doble cadena. Este paso de retrotranscripción es necesario debido a que el VIH-1 necesita integrar su ADN al de la célula hospedera para poder replicarlo.

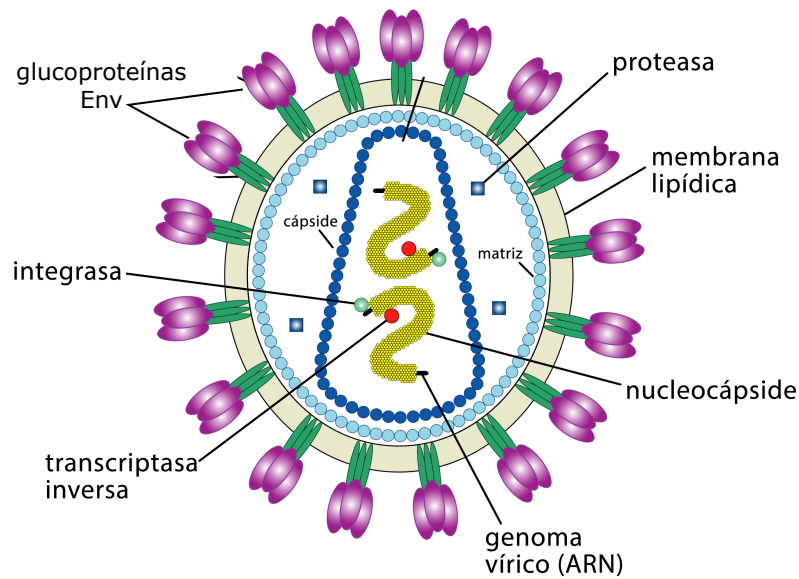


Figura 2.1: Virión maduro del VIH-1 compuesto por una cápside cónica que protege dos copias del ARN del virus. La cápside se encuentra rodeada por una membrana lipídica tomada de la célula infectada. Glucoproteínas Env se encuentran incrustadas alrededor de la membrana lipídica, las cuales sirven para unirse a otra célula para infectarla.

El ciclo de replicación del VIH-1 se da en dos fases. La primera fase, llamada fase temprana, va desde la llegada del virus a la célula hospedera hasta la integración del material genético del virus dentro del núcleo de la célula.<sup>29</sup> En esta fase se da la unión del virus con la célula por medio de las glucoproteínas Env sin que la célula lo reconozca como amenazante. Se da la penetración del virus dentro de la célula, mediante la fusión de las membranas lipídicas del virus y la célula, y comienza la transcripción inversa del ARN vírico a ADN por medio de la enzima transcriptasa inversa. Una vez que se tiene el ADN del virus este se dirige al núcleo de la célula junto con la enzima integrasa la cual se encargará de integrar el ADN del virus al ADN de la célula. A partir de este momento el virus se valdrá de la maquinaria de la célula para replicar su material genético.

La segunda fase de la replicación del VIH-1, llamada fase tardía, comienza con la transcripción de su material genético, dentro del núcleo celular, hasta la liberación del nuevo virión desde la membrana plasmática de la célula hospedera y su maduración.<sup>29</sup> La fase tardía se muestra en la figura 2.2 y transcurre como sigue. El ADN del virus, una vez integrado al de la célula hospedera, se vale de la enzima ARN-polimerasa de la célula para realizar su transcripción genética. En este proceso se sintetiza el ARN que contiene el genoma del virus y mARN (ARN-mensajero) que es donde se codifica la información para generar las proteínas del virus. El ARN del virus, generado durante la

transcripción, es exportado del núcleo al citoplasma, donde se da su traducción de mARN a las proteínas que formarán el virus. Estas proteínas son: Poliproteínas precursoras Gag y GagPol, glucoproteínas de envoltura Env, entre otras proteínas reguladoras y accesorias.

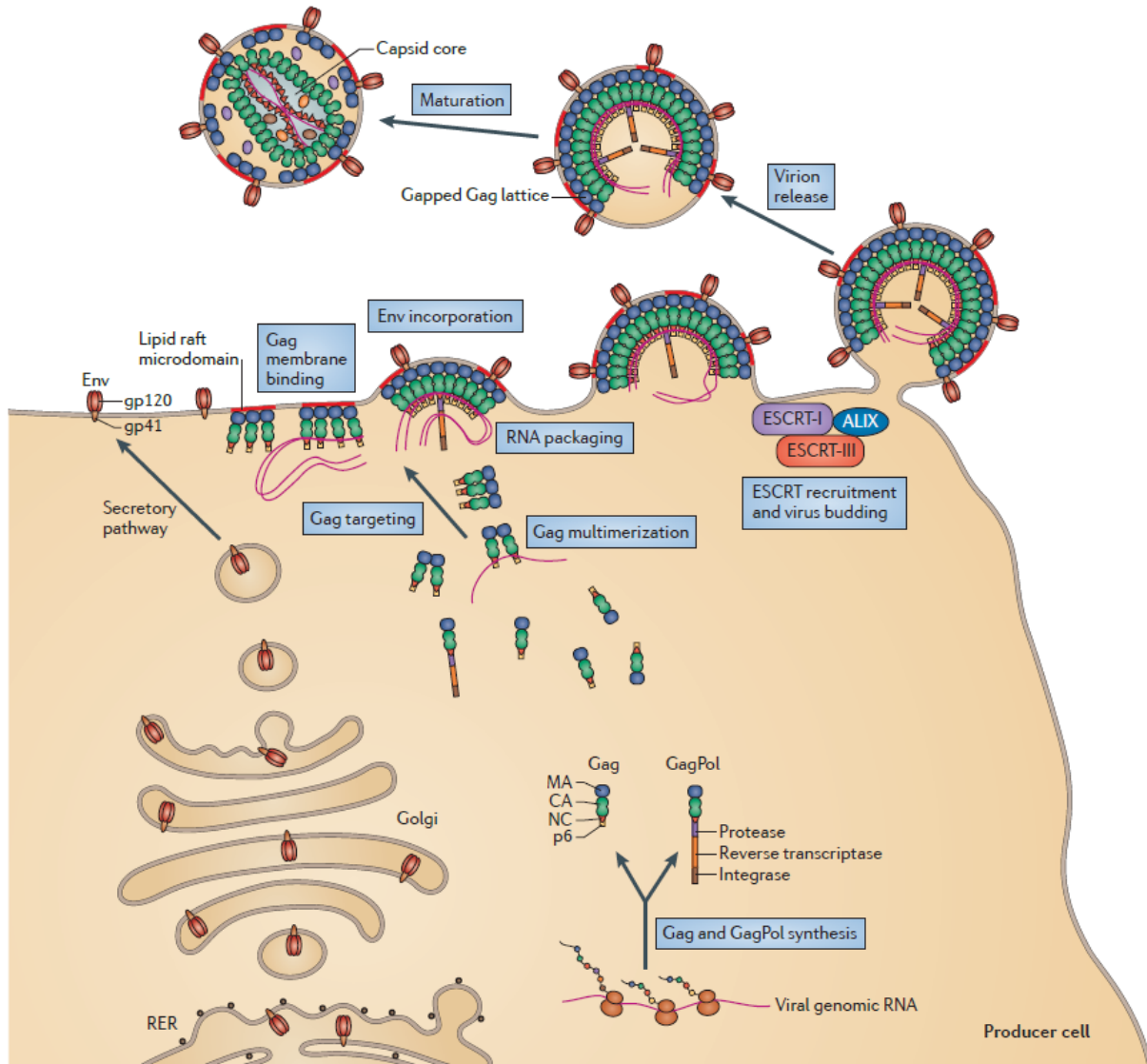


Figura 2.2: Fase tardía del ciclo de replicación del VIH-1. Comienza con la traducción del ADN del virus, integrado al de la célula, para generar el ARN vírico. El ARN es traducido a las poliproteínas Gag y GagPol, conteniendo esta última las enzimas proteasa, transcriptasa inversa e integrasa. Gag y GagPol se dirigen a la parte interna de la membrana plasmática donde se ensamblan y empaquetan el ARN del virus. Las glucoproteínas Env se dirigen a la parte externa de la membrana desde donde se unen a los Gag y GagPol. El ensamblaje de Gag y GagPol en la membrana provoca la formación de una protuberancia esférica. Llegado cierto tamaño de la protuberancia, Gag induce la separación total de la membrana para liberar el virión mediante la reclutación del complejo de clasificación endosomal requerido para el transporte (ESCRT de sus siglas en inglés) de la célula. Una vez liberado el virión, la enzima proteasa corta a Gag y GagPol en las subproteínas que formarán la cápside que protegerá el ARN, con lo que el virión alcanza la maduración. Imagen tomada de Freed<sup>29</sup>

Las poliproteínas precursoras están formadas por las proteínas que serán necesarias para ensamblar la estructura del virus, por ejemplo la cápside, y se diferencian en que

GagPol, además de tener las mismas proteínas que Gag, lleva consigo las enzimas que necesita el virus para su replicación (transcriptasa inversa, integrasa y proteasa). Una vez sintetizadas las poliproteínas Gag y GagPol, estas se dirigen a la membrana plasmática uniéndose a ella mediante interacciones electrostáticas a la vez que se ensamblan entre ellas para formar el nuevo virión. Desde el citosol el ARN, con el genoma del virus, se dimeriza y es llevado por las poliproteínas Gag y GagPol hacia la membrana plasmática para su empaquetamiento. Al mismo tiempo, las glucoproteínas Env se dirigen a la parte exterior de la membrana plasmática desde donde se unen con los poliproteínas Gag y GagPol. Conforme se ensambla el virus en la parte interna de la membrana celular, una protuberancia empieza a aparecer. Esta protuberancia se forma debido a que los Gag y GagPol son más anchos de uno de sus extremos que del otro, por lo que al unirse a la membrana por su extremo más ancho y al ir uniéndose entre ellos, unos al lado de los otros, estos doblan la membrana hacia su extremo más delgado. Esta protuberancia crece de manera esférica, mientras más Gag y GagPol se unen, hasta llegar a cierto tamaño para ser liberada de la célula conteniendo al virus dentro de sí. Esta es la manera en que el virus queda envuelto por la membrana plasmática de la célula hospedera. Una vez formada la partícula vírica (virión), el virus se prepara para su maduración. Esto lo logra a partir de la enzima proteasa, la cual se encarga de cortar las poliproteínas Gag y GagPol en las subproteínas que formarán la cápside que resguardará el ARN del virus. De este modo, el virión del VIH-1 alcanza su maduración y ahora tiene la capacidad de infectar una nueva célula.

### 2.3. Estructura de Gag y GagPol

Las poliproteínas precursoras Gag y GagPol son de plena importancia ya que son las responsables de toda la formación del virión hasta su maduración. Sirven como bloques constructores para ensamblar el virión en la parte interna de la membrana plasmática, se encargan de empaquetar el ARN del virus, inducen el corte de la membrana plasmática para liberarse de la célula hospedera y finalmente, al ser cortado por la proteasa, forma su cápside alcanzando así la maduración. Cada una de estas funciones es llevada a cabo por una parte específica del precursor Gag, las cuales se describirán a continuación

Como se ve en la figura 2.3, el precursor Gag está formado por 4 dominios distintos<sup>29</sup>

matriz (MA), cápside (CA), núcleo cápside (NC) y p6, además de dos péptidos espaciadores (SP1) y (SP2). MA es la encargada de dirigir al Gag a la membrana plasmática ya que la parte superior de este dominio se atrae con los lípidos de la parte interna de la membrana debido a una interacción electrostática entre los residuos básicos de MA y los lípidos ácidos. Además, MA tiene un ácido mirístico (Myr) que es una cadena de catorce carbonos que MA utiliza para anclarse dentro de la membrana. Myr se mantiene dentro de MA y sobresale únicamente cuando MA ya se encuentra unida a la membrana. Finalmente, MA también se encuentra involucrada en la incorporación de las glucoproteínas Env que llegan a la parte externa de la membrana plasmática, por otra vía, e interaccionan con MA desde ese punto.

El dominio CA es el de mayor importancia en la multimerización de Gag. Las primeras interacciones CA-CA se dan desde el citosol, que es la fase líquida que se encuentra entre el núcleo y la membrana plasmática. Llegados a la membrana se organizan en hexágonos formando una red.<sup>30</sup> En la etapa de maduración CA, después de ser cortado por la enzima polimerasa, es el encargado de ensamblar la cápside cónica del VIH-1. La cápside se compone de 216 hexámeros (6 CA formando un hexágono) y 12 pentámeros (5 CA formando un pentágono), siete en su extremo más ancho y 5 en su extremo más delgado.<sup>27</sup> En la figura 2.3 a) se señala como es que CA se subdivide en dos subdominios: dominio N-terminal (NTD) y dominio C-terminal (CTD). La parte CA-NTD se encuentra debajo de MA y la parte CA-CTD, de bajo de CA-NTD, unida al siguiente dominio. La división de CA en subdominios ayuda a la protuberancia que se forma en la membrana plasmática para formar el virión.<sup>31</sup>

NC también participa en la multimerización de Gag, pero su principal función es la de empaquetar ARN vírico. NC está formado por dos dominios a los que se le conoce como dedos de zinc. Estos dedos de zinc son los que reconocen al ARN y se encargan de reclutarlo y llevarlo a la membrana plasmática para su empaquetamiento. Ya en la etapa de maduración NC permanece unido al ARN rodeándolo por completo (Fig. 2.1).

El dominio p6 se encarga de reclutar al complejo de clasificación endosomal requerido para el transporte (ESCRT por sus siglas en inglés), que es un complejo de proteínas citosólicas que intervienen en la última etapa de la división celular, por lo que se encargará de la escisión de la membrana plasmática para liberar al virión de la célula hospedera. Por último, el espaciador SP1 conecta a los dominios MA y NC, este espaciador también



colabora en la multimerización de Gag. El segundo espaciador SP2 conecta a los dominios NC y p6.

La figura 2.3 b) muestra como la poliproteína GagPol está compuesta de la misma manera que Gag, pero esta además incluye, las enzimas proteasa, transcriptasa inversa y polimerasa, en ese orden. Como ya se mencionó, la proteasa se encarga de cortar a Gag y GagPol en las subunidades que formarán la cápside y núcleo cápside, la transcriptasa inversa se encarga de pasar el ARN a ADN y la integrasa integra el ADN del virus al de la célula hospedera. Aunque GagPol lleve consigo las enzimas necesarias para la replicación del virus, esta solo se produce en un 5% respecto de Gag.<sup>29</sup>

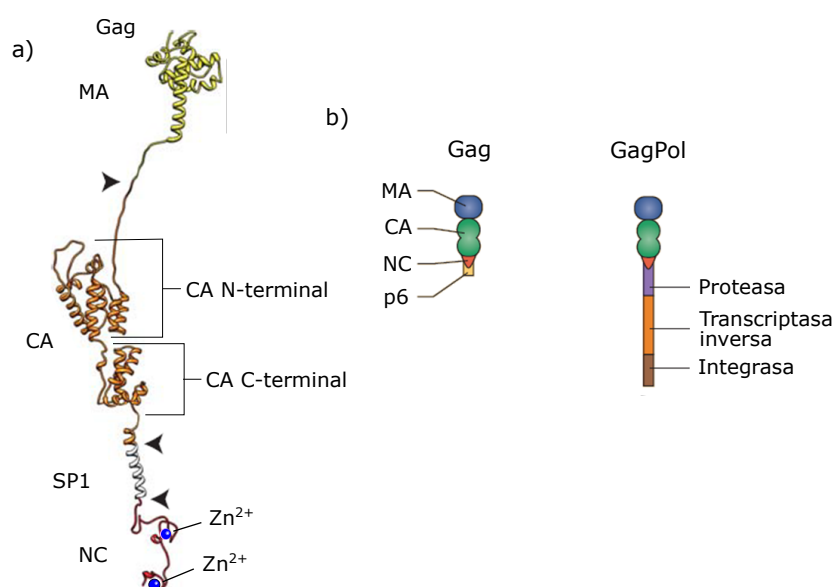


Figura 2.3: Dominios de la poliproteína Gag y GagPol. a) Modelo construido por Sundquist et al.<sup>32</sup> a partir de combinar las estructuras tridimensionales de los dominios del Gag reportadas en el banco de datos de proteínas<sup>33</sup> (PDB por sus siglas en inglés), el cual es una herramienta para explorar estructuras en 3D de macromoléculas biológicas. b) Esquemas de Gag y GagPol que identifican sus dominios y las enzimas en el caso de GagPol.

## 2.4. Liberación del virión de la membrana plasmática

### 2.4.1. Membranas celulares

Distintos tipos de lípidos son los encargados de formar membranas que son de suma importancia para la célula. Una membrana lipídica tiene como función aislar a la célula del ambiente a su alrededor, así como de aislar, en su interior, a los distintos orgánulos del citosol. Además de su función protectora, las membranas son indispensables en el

reconocimiento molecular y en el paso de señales, provenientes del exterior, hacia el interior de la célula.

Los lípidos de las membranas se pueden separar en dos grupos: estructurales y no estructurales. Este primer grupo es el encargado de la propia formación de las membranas. Debido a sus cabezas hidrofílicas y a sus colas hidrofóbicas estos lípidos se ensamblan en membranas de doble capa. Los lípidos estructurales también son los responsables de las distintas fases que puede adoptar una membrana. Dependiendo del lípido se pueden formar tres fases distintas: tipo gel, líquido desordenado o líquido ordenado. Caracterizadas cada una por el orden de la cola del lípido y por su difusión traslacional.<sup>14</sup> Por otro lado, el grupo de los lípidos no estructurales se encargan de llevar a cabo las actividades de la membrana de: señalización, paso de mensajes, formación de dominios que reclutan proteínas desde el citosol, así como de procesos de fusión y fisión indispensables para la división celular.

### 2.4.2. Proceso de liberación

El VIH-1 para liberarse de la célula infectada se envuelve en la membrana plasmática formando una especie de vesícula (virión). Las vesículas son usadas por la célula para transportar moléculas entre organelos o para expulsarlas de la célula (exocitosis). En el proceso de formación de las vesículas de la célula participan alrededor de cincuenta moléculas distintas, sin embargo, en el caso del VIH-1, las proteínas Gag son las responsables de inducir este proceso.

La interacción entre el dominio MA del Gag y la parte interna de la membrana plasmática es de plena importancia ya que se ha encontrado que la membrana con la que se envuelve el VIH-1 está enriquecida en esfingolípidos, fosfatidilinositol y colesterol, lípidos que no se encuentran en las mismas concentraciones en otras partes de la membrana plasmática.<sup>34,35</sup> El fosfatidilinositol es el implicado en la atracción electrostática con MA e induce la salida de Myr para anclarse en la membrana.<sup>36</sup> El colesterol es reclutado por la llegada de las proteínas Gag a la membrana y forma micro dominios llamados balsas lipídicas.<sup>37</sup> Estas balsas lipídicas incrementan la selectividad de MA por el fosfatidilinositol y mejoran la unión de Myr con la membrana.<sup>38,39</sup> Se sabe que los lugares por donde se dan contactos célula-célula son ricos en balsas lipídicas y se ha encontrado que la disminución de estas balsas lipídicas es perjudicial para la infectividad del VIH-1.<sup>37</sup>

Conforme MA dirige a las moléculas de Gag hacia la membrana plasmática la multi-

merización de estos empieza a generar una protuberancia sobre la membrana (Fig. 2.4). La multimerización es dirigida principalmente por las interacciones CA-CA de Gag y es apoyada por NC durante el reclutamiento del ARN vírico. La protuberancia es inducida por las moléculas de Gag gracias a que estas se arreglan formando hexágonos.<sup>30,40,41</sup> Modelos de este arreglo hexagonal por parte de CA y SP1, hechos a partir de criotomografías electrónicas en viriones inmaduros de VIH-1, los presentan con una forma cónica,<sup>42</sup> con CA como la base del cono que apunta hacia la membrana y SP1 el vértice que apunta hacia el interior del virión. También, a partir de modelos de grano grueso de CA y SP1 se ha encontrado que el diámetro del hexágono formado por el subdominio CA-NTD es de mayor tamaño que el compuesto por el subdominio CA-CTD.<sup>31</sup> Esto produce una curvatura sobre la red ensamblada por los hexágonos, lo que a su vez contribuye a doblar la membrana plasmática.

Cuando la protuberancia (virión) alcanza cierto tamaño el dominio P6 recluta al complejo ESCRT. Esta maquinaria de la célula se compone de varios complejos de proteínas que se encargan de liberar las vesículas de las membranas y de dividir la membrana plasmática en la formación de células hijas. El VIH-1 hace uso de algunos de los complejos de esta maquinaria como son ESCRT-II y ESCRT-III, siendo este último de los más importantes. Se ha encontrado que ESCRT-III forma filamentos largos que se enrollan en el cuello del virión de la membrana dirigiendo su escisión. También es el encargado de reclutar a la proteína AAA ATPasa la cual provee la energía requerida para la escisión de la membrana.<sup>29</sup> Estudios señalan que el complejo ESCRT se sitúa dentro del virión desde donde dirige la escisión como se ve en la figura 2.4. Por otro lado, otros estudios observan la presencia de los complejos ESCRT solo en el cuello del virión.

Este es el proceso que siguen las moléculas de Gag para ensamblar el virión en la capa interna de la membrana plasmática, concluyendo con su liberación de la membrana para dar paso a su maduración.

## 2.5. Red hexagonal en la etapa inmadura

En la segunda fase de replicación del virus, durante la formación del virión en la membrana plasmática de la célula se da el ensamblaje de la red hexagonal por parte de las poliproteínas Gag. Como ya se mencionó, esta red, que se encuentra en la parte interna

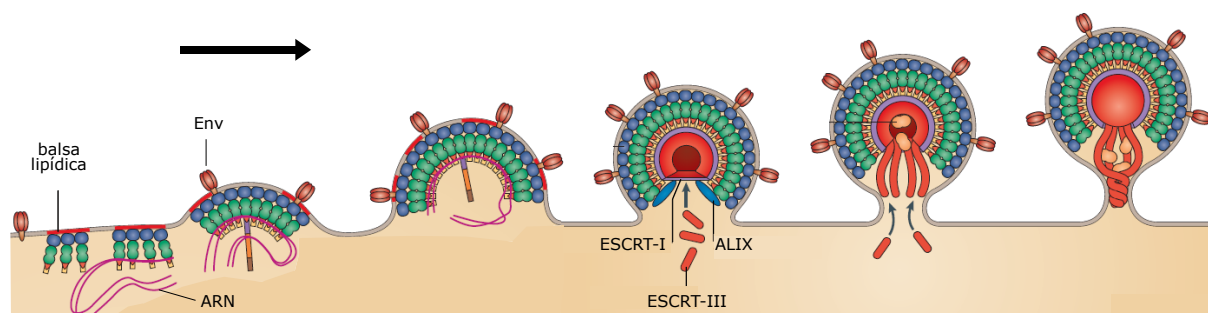


Figura 2.4: Evolución de la formación del virión en la parte interna de la membrana plasmática hasta el reclutamiento del complejo ESCRT para la escisión de la membrana. Imagen modificada de Freed<sup>29</sup>

del virión, salvo la región por donde se da la liberación de la membrana (Fig. 2.4), ayuda a doblar la membrana, lo cual promueve la formación del virión.

Ya desde 1966, a partir de purificar y cristalizar MA, se sabía que este dominio por si solo se ordenaba formando trímeros.<sup>43</sup> Posteriormente, se observó que MA formaba hexámeros de trímeros (Fig. 2.5 b) sobre membranas lipídicas que contenían fosfatidilinositol y colesterol.<sup>40</sup> Con el uso de Tomografías Crio-Electrónicas (TC-E) aplicadas a viriones inmaduros se observó la red hexagonal que presenta el dominio CA de Gag.<sup>41,42,44</sup> De estas observaciones se concluyó que las moléculas de Gag forman una red continua, pero con varios defectos sobre ella, como se ve en la figura 2.5 (c). Más recientemente con estudios de TC-E y métodos de promediado de subtomogramas se reveló la estructura que adopta el dominio CA dentro de la red hexagonal en viriones inmaduros intactos<sup>30</sup>(Fig. 2.5 a). Con los sitios de interacción más importantes identificados en el dominio CA en su etapa inmadura,<sup>45</sup> junto con la estructura de CA, es posible reconstruir una red hexagonal localizando las principales interacciones que intervienen en la formación de la red<sup>30,46</sup>(Fig. 2.5 d). A estas interacciones al coincidir con sitios de simetría de una red hexagonal se les denominaron 2-Fold (2F), 3-Fold (3F) y 6-Fold (6F), indicados en la figura 2.5 (e).

La interacción 2-Fold y 6-Fold se encuentran en la parte CA-CTD, mientras que la 3-Fold se localiza en la parte CA-NTD (Fig. 2.5 a). 2F y 3F son consideradas interacciones inter-hexaméricas ya que forman parejas y ternas, respectivamente, entre moléculas de Gag de hexágonos adyacentes (Fig. 2.5 e). Por el contrario, a la interacción 6F se le considera intra-hexamérica ya que se localiza en el interior del hexágono (Fig. 2.5 e). Esta última interacción no se da directamente entre 6 Gag, sino que interviene otra molécula llamada IP6 que se coloca en el centro del hexágono interaccionando con las 6 moléculas de Gag a la vez. De estas tres interacciones, como se verá más adelante, será

descartada la interacción 6-Fold en la construcción de los modelos de simulación.

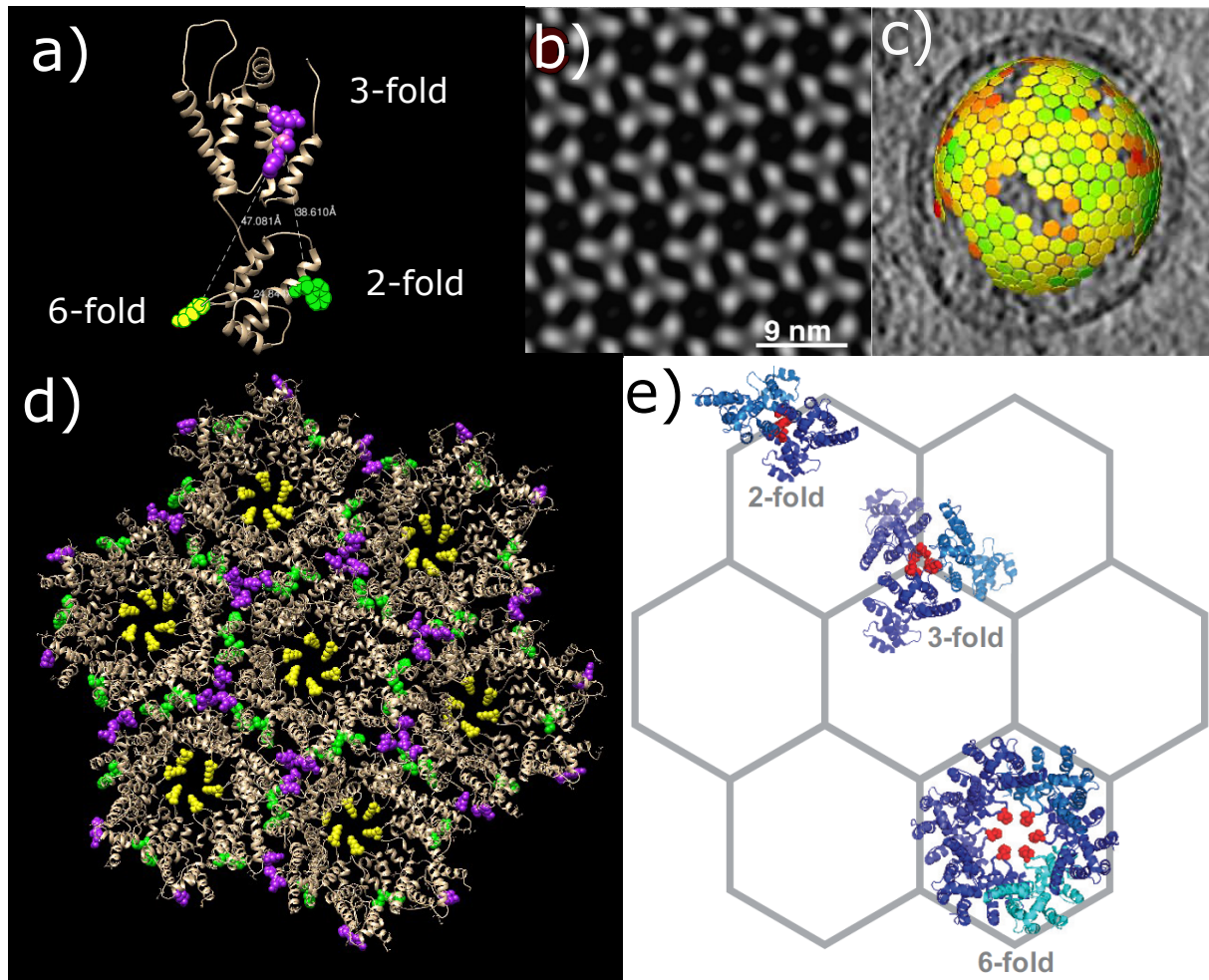


Figura 2.5: Red hexagonal que forma el VIH-1 en su etapa inmadura. a) Estructura del dominio CA en su etapa inmadura, extraída del PDB 4USN,<sup>30</sup> el cual es un archivo del banco de datos de proteínas<sup>33</sup> creado a partir de Tomografías Crio-Eléctricas de viriones inmaduros del VIH-1. b) Microscopía electrónica que muestra la organización de hexámeros de trímeros de MA sobre una membrana lipídica. c) Red hexagonal formada por las poliproteínas Gag en viriones inmaduros obtenidas a partir de criotomografías electrónicas. d) Reconstrucción de la red hexagonal a partir de la estructura del dominio CA (PDB 4USN), en su etapa inmadura. e) Señalización de las interacciones 2-Fold, 3-Fold y 6-Fold dentro de la red hexagonal. b) y c) son tomadas de Alfarhli et al.<sup>40</sup> y Schur,<sup>30</sup> respectivamente.

## 2.6. Simulación molecular: liberación de viriones y redes abiertas

Entender el ciclo de replicación de los virus es de suma importancia para tener mejores formas de combatirlos. El uso de simulaciones moleculares es una herramienta que ayuda en este aspecto, ya que a partir de crear modelos basados en teorías u observaciones experimentales, se contribuye a elucidar la manera en que los virus realizan su ciclo de

replicación. A continuación se mencionarán algunos de los trabajos enfocados en los virus, donde se utilizan modelos de grano grueso, y que están relacionados con los resultados que se mostrarán en este trabajo.

### 2.6.1. Liberación de estructuras tipo cápside de membranas

La liberación del virión de la célula hospedera ha sido de las partes del ciclo de replicación de los virus donde más se han centrado los estudios. Esto se debe a que en este proceso interviene la membrana plasmática, la cual ya ha sido estudiada debido a su importancia en la célula. Por ejemplo, se han hecho simulaciones con dinámica molecular para estudiar la formación de balsas lipídicas en membranas de doble capa.<sup>47</sup> Este trabajo hace uso de un solvente explícito, los lípidos interactúan con un potencial de Lennard-Jones (LJ) de corto alcance y usa un modelo de grano grueso para los lípidos, donde cada 4 átomos del lípido son representados por un centro de interacción. Por otro lado, Cooke y Deserno<sup>48</sup> desarrollaron un modelo libre de solvente para autoensamblar membranas fluidas de doble capa. En este caso también se usa un potencial de LJ (modificado para que el mínimo del potencial sea más amplio) para la interacción entre las colas de los lípidos, el cual simulará la interacción hidrofóbica, pero ahora los lípidos son modelados únicamente por tres esferas, una para la cabeza y dos para la cola del lípido.

Un modelo de membrana, no tan detallado como los anteriores, fue usado por Richard Matthews y Christos Likos<sup>49,50</sup> para estudiar la influencia que tenía sobre partículas parchadas que se agregan en pequeños núcleos, emulando una cápside. La membrana es modelada como una superficie triangular formada por esferas con pequeños parches atractivos a su alrededor. La simplificación de la red permite agregar las partículas tipo cápside sin un gran incremento del tiempo de simulación. Los resultados obtenidos fueron que la interacción con la membrana promueve la formación de los núcleos reduciendo el tiempo de ensamblaje y que una buena deformación de la membrana permite el recubrimiento de los núcleos por parte de esta.

Otros dos trabajos de dinámica molecular, donde usan el modelo de membrana lipídica de Cooke y Deserno,<sup>48</sup> logran la liberación de lo que sería el virión de la membrana. Ruiz-Herrero et al.,<sup>51</sup> primeramente, utiliza una esfera que interacciona directamente con las cabezas de los lípidos. Con este modelo encuentra que conforme se aumenta la fuerza de atracción esfera-lípido se observan las siguientes cuatro trayectorias en la dinámica:

adhesión de la esfera a la membrana, sin envolverse con ella; adhesión y recubrimiento parcial de la esfera por la membrana; recubrimiento completo de la esfera y liberación de la membrana; y por último, envoltura y liberación de la esfera provocando la ruptura de la membrana. Posteriormente, Ruiz-Herrero y Hagan<sup>52</sup> muestran que microdominios en la membrana facilitan el ensamblaje y liberación del virión de esta. En este modelo usan pentámeros formados por varias esferas traslapadas. Los pentámeros interactúan atractivamente entre sí por sus lados y sus vértices e interactúan con la membrana (con las colas de los lípidos) a través de 6 sitios atractivos situados por encima del plano del pentámero (fuera de él) asemejando al ácido mirístico de MA que penetra y se ancla en la membrana. Con este modelo encuentran que con un valor correcto de la fuerza de interacción pentámero-lípido se da el ensamblaje de un icosaedro (cápside) por parte de los pentámeros y que la presencia de un microdominio en la membrana mejora el ensamblaje de los capsómeros y promueve la liberación del virión.

### 2.6.2. Autoensamblaje de redes abiertas

La red hexagonal que se forma en la etapa inmadura del virión del VIH-1 es similar a otro tipo de redes hexagonales como la Kagome. De estas otras redes ya se ha estudiado la manera en la que se autoensamblan, ya sea experimentalmente o con simulaciones numéricas. Por lo que es de interés conocer la manera en que se han auto ensamblado estas otras redes para ayudar a descubrir el proceso de ensamblaje que tienen las moléculas de Gag al formar la red hexagonal inmadura.

Una de las maneras de autoensamblar la Kagome es usando partículas coloidales,<sup>6</sup> que son esferas micrométricas funcionalizadas de tal forma que su ecuador es electrostáticamente repulsivo y sus polos son hidrofólicamente atractivos. Además, estas partículas se pueden funcionalizar con parches atractivos sobre su superficie, de distintos tamaños y en distintas posiciones, lo que les da la posibilidad de ensamblar otro tipo de redes.<sup>53,54</sup> Debido a esto las simulaciones de partículas parchadas es una de las formas de estudiar el ensamblaje de redes abiertas.

Con modelos de partículas parchadas en 2D la Kagome ha sido ensamblada con éxito. Romano y Sciortino<sup>55</sup> implementaron el potencial de Kern-Frenkel,<sup>56</sup> el cual impone restricciones orientacionales entre los parches para ensamblar la red. Chapela et al.<sup>57</sup> usaron potenciales de pozo cuadrado (PC) para los parches atractivos sin imponer restricciones

orientacionales para ensamblar la red. Edlund et al.,<sup>58,59</sup> usaron métodos inversos basados en el principio de incertidumbre para obtener potenciales isotrópicos que pueden ensamblar la Kagome, entre otras redes. Una variación de la Kagome, llamada Kagome deformada por Machorro et al.<sup>60</sup> y Kagome retorcida por Salgado et al.<sup>61</sup> fue ensamblada con partículas parchadas, modeladas con potenciales de PC, y con una mezcla binaria de partículas con potenciales isotrópicos, respectivamente. Otra gran variedad de redes en 2D como la Kagome, panal hexagonal, cuadrada, rectangular, cuadrada truncada y hexagonal truncada fueron ensambladas por el grupo de Truskett<sup>62-64</sup> por medio de potenciales isotrópicos repulsivos obtenidos a partir de una optimización de métodos inversos.

Por otro lado, el reporte de estas redes ensambladas en 3D es escaso, inclusive para la tan estudiada Kagome. Romano<sup>55</sup> usa esferas parchadas para ensamblar la Kagome, permitiéndoles rotar libremente, pero restringe su desplazamiento a un plano. Una red que es más fácil de ensamblar en 3D es la panal hexagonal. Ha sido ensamblada con partículas esféricas con tres parches atractivos equidistantes sobre el ecuador.<sup>65</sup> También, de manera más sencilla, ya que se usan trímeros en lugar de monómeros como bloque constructor, se ensambla esta red por medio de un algoritmo de Reaccion-Difusión que se basa en la distancia entre extremos de los trímeros para interactuar sin tomar en cuenta sus orientaciones.<sup>66</sup>

Es evidente que es más sencillo ensamblar redes en 2D que en 3D, sin embargo la red ensamblada en 3D, sin restricciones en su desplazamiento sobre el espacio,<sup>65</sup> es de simetría hexagonal como la de la red formada por el VIH-1. Aunque se debe tener en cuenta que las moléculas de Gag cuando ensamblan la red están restringidos a moverse por el plano que les impone la membrana de la célula hospedera, lo cual debe facilitarles el ensamblaje.



# Capítulo 3

## Cápsides icosaédricas

Un tercer objetivo que se tiene en este trabajo es el ensamblaje de cápsides icosaédricas. Este objetivo se puede ver como un segundo paso a la formación de la red hexagonal del VIH-1, ya que esta estructura sirve como base para la formación de cápsides cónicas, como la del VIH-1, o icosaédricas. A continuación se hablará de estas últimas para conocerlas.

Las cápsides con forma esférica son muy comunes entre los virus y aquellas que tienen simetría icosaédrica son las más frecuentes. Esto se debe a que como los virus son unos organismos muy simples, que no tienen todos los elementos para subsistir por sí mismos, deben de economizar en las proteínas que tienen que usar para replicarse. Esto se ve reflejado en las proteínas que componen la cápside del virus, donde mientras menos proteínas de distintos tipos se tengan que codificar para ensamblar la cápside se podrán codificar otras proteínas que intervengan en otros procesos que garanticen la replicación del virus. Este proceso de optimización lleva al virus a elegir la simetría icosaédrica ya que es la manera más eficiente de crear una estructura cerrada con la mínima cantidad de subunidades idénticas.

### 3.1. Clasificación de las cápsides icosaédricas

Con base en lo anterior, las cápsides esféricas más pequeñas son aquellas que forman un icosaedro regular de 20 caras triangulares, donde cada cara está compuesta por tres proteínas. De este modo, en cada uno de los 12 vértices del icosaedro las proteínas forman un pentágono llamado pentámero, como se ve en la figura 3.1 a). Esto deja una cápside compuesta por un total de 12 pentámeros con un total de 60 proteínas iguales. Sin em-

bargo, hay virus más grandes que tienen un ácido nucleico que no se puede contener en una estructura de 60 proteínas. Esto lleva a la formación de cápsides más grandes, que de igual manera tendrán una simetría icosaédrica, seguirán compuestas por 12 pentámeros, uno en cada uno de sus vértices, pero ahora estos pentámeros estarán conectados por uno o más hexámeros, es decir, por seis proteínas formando un hexágono.

Caspar y Klug<sup>67</sup> desarrollaron un método para clasificar las cápsides icosaédricas de distintos tamaños, asignándoles un número de triangulación  $T = 1, 3, 4, 7, 9, \dots$ , donde mientras mayor es el número más grande es la cápside. La forma de construir las cápsides icosaédricas se puede ver de la siguiente manera. Si se toman las 20 caras triangulares y se superponen a una red hexagonal, como se ve en la figura 3.1 b), cada triángulo estará formado por cierto número de hexágonos, que dependerá del tamaño del icosaedro que se quiera formar. Se debe cumplir que cada vértice de los caras triangulares debe coincidir con el centro de un hexágono, ya que los vértices formarán 12 pentágonos al doblar y unir las aristas de todas las caras triangulares como se ve en la figura 3.1 b). Estos pentágonos son los responsables de llevar la representación plana de hexágonos a una forma esférica con simetría icosaédrica. La clasificación de Caspar y Klug de las cápsides icosaédricas concuerda con los llamados poliedros Goldberg,<sup>68</sup> que son poliedros formados únicamente por hexágonos y pentágonos.

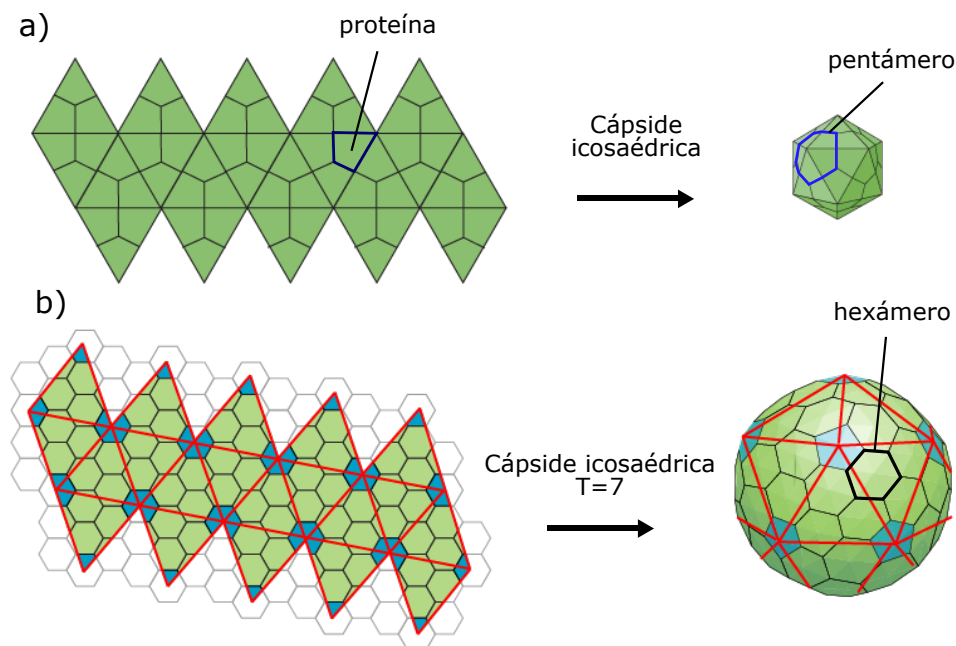


Figura 3.1: Cápsides con simetría icosaédrica. a) Desarrollo plano de un icosaedro. En los virus cada una de las caras triangulares esta compuesta por tres proteínas. b) Mapeo de las caras triangulares sobre una red hexagonal. En los vértices de las caras triangulares se forman los pentágonos encargados de llevar la red hexagonal plana a una forma esférica con simetría icosaédrica.

El número de triangulación está definido como  $T(h, k) = h^2 + hk + k^2$  el cual representa la distancia cuadrada del vértice de una cara triangular al centro de la misma, con  $h$  y  $k$  las coordenadas inclinadas  $60^\circ$  centradas en el vértice del triángulo. De igual manera se pueden ver a  $h$  y  $k$  como números enteros del vector  $C_T = (h, k)$  que une a dos pentágonos adyacentes de la cápside. Con  $h$  el número de saltos en línea recta a través de los hexágonos y  $k$  el número de saltos hacia izquierda o derecha, en caso de ser necesario, para llegar al siguiente pentágono, como se indica en la figura 3.2. De este modo  $k$  puede ser positivo (vuelta derecha), negativo (vuelta izquierda) o cero. Con el número de triangulación se puede saber el número de unidades estructurales (proteínas) dado por  $P = 60T$  y el número de unidades morfológicas (pentámeros y hexámeros) dado por  $M = 10T + 2 = 10(T - 1)$  hexámeros + 12 pentámeros, de las que está compuesta una cápside icosaédrica.

De este modo las primeras cuatro cápsides clasificadas con un número de triangulación son:  $T(1, 0) = 1$  compuesta por 12 pentámeros con un total de  $P = 60(1) = 60$  proteínas. La cápside siguiente es  $T(1, 1) = 3$  compuesta por  $M = 10(3 - 1) + 12 = 20$  hexámeros + 12 pentámeros y  $P = 60(3) = 180$  proteínas. La tercer cápside tiene número de triangulación  $T(2, 0) = 4$  ensamblada por  $M = 10(4 - 1) + 12 = 30$  hexámeros + 12 pentámeros interviniendo  $P = 60(4) = 240$  proteínas. Y la cuarta cápside tiene número de triangulación  $T(2, 1) = 7$  integrada por  $M = 10(7 - 1) + 12 = 60$  hexámeros + 12 pentámeros con un total de  $P = 60(7) = 420$  proteínas.

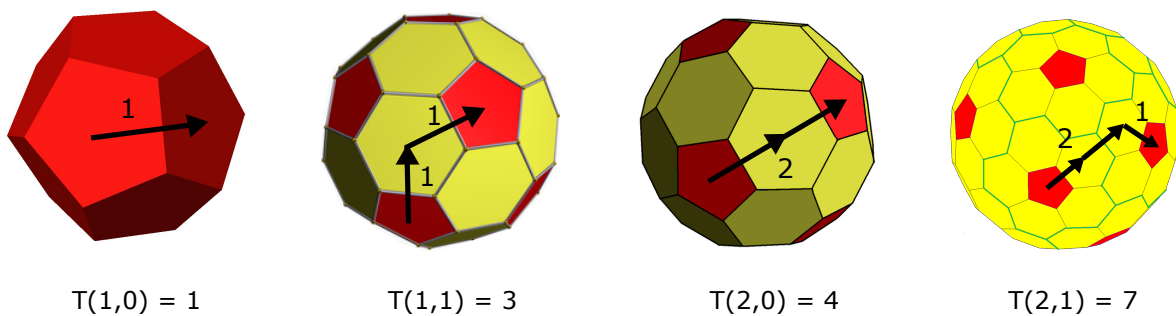


Figura 3.2: Cápsides icosaédricas con número de triangulación  $T=1, 3, 4, 7$ . Las cápsides están formadas por 12 pentámeros, situados en cada uno de los vértices de un icosaedro, y por distinto número de hexámeros dependiendo del número de triangulación. Cada pentámero y hexámero está formado por 5 y 6 proteínas, respectivamente.

## 3.2. Simulación de cápsides icosaédricas

Con el uso de distintas técnicas experimentales se ha podido resolver la estructura de las cápsides de los virus. Sin embargo la dinámica de como se lleva a cabo su ensamblaje es incierto. Debido a esto, las simulaciones moleculares que pretenden reproducir este proceso son variadas. A continuación presentaremos algunas de ellas que lo hacen con modelos desde lo más sencillo, respecto al bloque constructor que usan, hasta los más detallado posible.

Rapaport en una serie de trabajos usa un conjunto de varias esferas para formar un bloque triangular, que interacciona por medio de sitios atractivos en los costados del bloque, con el que ensambla varias cápsides  $T = 1$ <sup>69</sup> y estudia la reversibilidad en el proceso de ensamblaje.<sup>70</sup> También, con bloques trapezoidales, contruidos de la misma manera, de hasta tres especies distintas, con las que se pueden imponer reglas de interacción más específicas, forma cápsides  $T = 1$  y  $T = 3$ .<sup>71</sup> Johnston et al.,<sup>72</sup> también logra ensamblar la cápside  $T = 1$  con pirámides pentagonales rígidas y la cápside  $T = 3$  con una mezcla de pirámides pentagonales y hexagonales. Las bases de las pirámides interaccionan atractivamente mientras que sus ápices interaccionan repulsivamente, lo cual promueve y fija la curvatura de la cápside. En otro trabajo, Nguyen y Brooks<sup>73</sup> han logrado ensamblar de manera correcta hasta la cápside  $T = 13$  a partir de estructuras pentagonales y hexagonales como bloques constructores. Estos pentámeros y hexámeros interaccionan con una serie de reglas específicas propuestas por Kerner<sup>74</sup> las cuales cambian dependiendo del número de triangulación que se quiera ensamblar. Debido a esto es posible ensamblar cápsides de  $T$  altos, sin embargo conforme aumenta el número de triangulación las reglas de interacción también aumentan.

Por otro lado, Baschek et al.<sup>75</sup> también ensambla las cápsides  $T = 1$  y  $T = 3$ , pero en este caso, a diferencia de los anteriores mencionados, usa una esfera con tres parches atractivos distribuidos en su ecuador como bloque constructor. Debido a esto, necesita 5 y 6 de estas moléculas para, primeramente, ensamblar los pentámeros y hexámeros que llevaran a la formación de las cápsides. Este modelo, aunque se diferencia de los anteriores por usar un bloque constructor más cercano a representar una proteína y ensamblar los hexámeros y pentámeros, también requiere de más de una especie de estas moléculas para ensamblar la cápside  $T=3$  de manera correcta.

# Capítulo 4

## Modelos y Potenciales

En este trabajo se buscó la formación de una estructura tipo virión en la superficie de una membrana de doble capa, su envoltura en la membrana y su liberación de ella, justo como el proceso que lleva a cabo el virus del VIH-1. Adicionalmente, se buscó el autoensamblaje de la red hexagonal que las proteínas Gag del VIH-1 forman en su etapa inmadura. Finalmente, se llevó a cabo el autoensamblaje de cápsides icosaédricas a partir del modelo usado para ensamblar la red hexagonal del VIH-1. A continuación se presentarán los modelos moleculares y potenciales de interacción utilizados para llevar a cabo todos los procesos de autoensamblaje mencionados.

### 4.1. Potenciales de interacción

Para modelar los sistemas mencionados se usaron partículas decoradas con parches atractivos de distintos tamaños. La interacción entre las moléculas son sitio-sitio por lo que se tendrán sitios atractivos para los parches y sitios repulsivos para la esfera o disco que contendrá los parches. Las moléculas se modelaron con dos tipos de potenciales, dependiendo del sistema que se iba a modelar. Se usaron potenciales discontinuos (Fig. 4.1b) para el sistemas de la liberación del virión de una membrana y para la formación de la red hexagonal del VIH-1. Mientras que se usaron potenciales continuos (Fig. 4.1c) para el ensamblaje de las cápsides icosaédricas. De este modo, para los sitios atractivos se usaron potenciales de pozo cuadrado (PC) y potenciales de Lennard-Jones (LJ) y para los sitios repulsivos se usaron barreras de potencial infinitas y la parte repulsiva del potencial de LJ.

En la figura 4.1 se muestra el esquema de las moléculas parchadas que se utilizaron en los distintos sistemas. Se muestra un disco repulsivo central decorado con dos parches pequeños y un parche más grande, con sus respectivos alcances de su potencial atractivo  $\lambda$ , para PC, o radio de corte ( $rc$ ), para LJ. En la parte central de la figura 4.1 se muestra el esquema de los potenciales intermoleculares con los que interaccionarían las moléculas. En la parte derecha de la figura 4.1 se muestran los potenciales intramoleculares encargados de mantener los parches en su lugar, respecto del disco central y respecto de los demás parches. A estos confinamientos también se les llama ligaduras y se dice que se liga a dos átomos para mantenerlos a una distancia dada. De este modo, cada ligadura está definida por un potencial centrado a la distancia  $L$ , dada por la longitud de la ligadura (distancia entre los dos átomos), para mantener a dos átomos a esa distancia. Debido a esto, cada una de las partículas que forman la molécula tendrá más de una ligadura para mantener su posición respecto a las demás dentro de la molécula. Los potenciales que se usan para las ligaduras son pozos infinitos, para el caso discontinuo, y osciladores armónicos, para el caso continuo. Debido a que en 3D los parches de la moléculas siempre estarán situados sobre el ecuador de la esfera, la figura 4.1a) representa de igual forma los casos 2D y 3D.

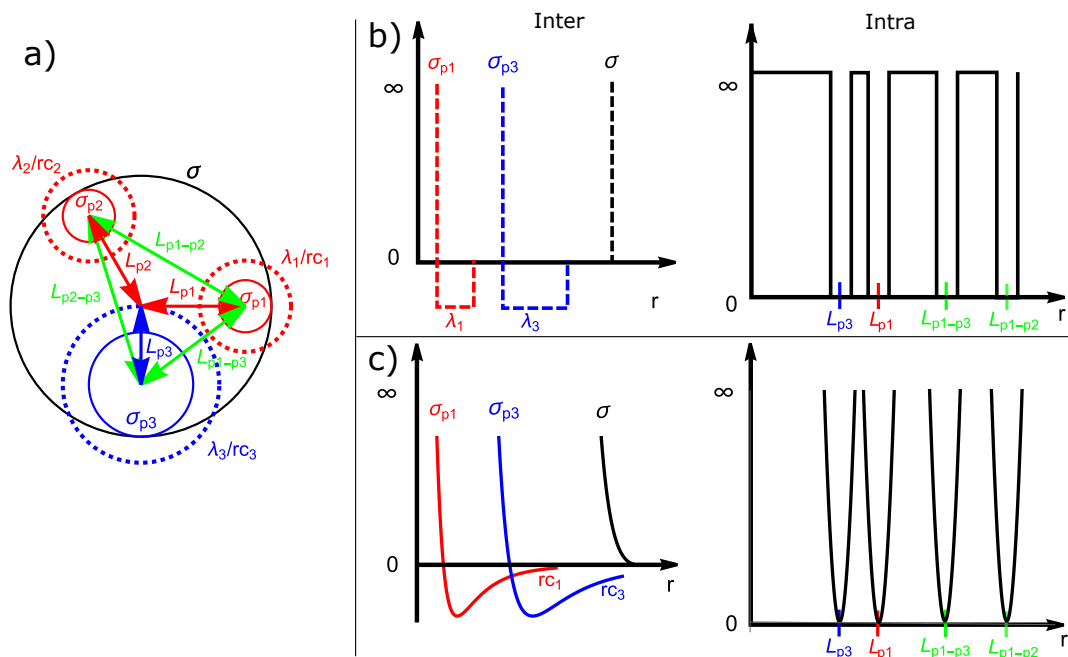


Figura 4.1: Potenciales de interacción entre moléculas con sitios atractivos. a) Modelo de una molécula con parches atractivos. Las ligaduras encargadas de mantener la posición de los átomos de la molécula se representan con flechas. b) Representación esquemática de los potenciales intermoleculares e intramoleculares en su versión discontinua. c) Versión continua de los potenciales intermoleculares e intramoleculares.

De este modo, la interacción entre dos moléculas  $m$  y  $n$  compuestas por un número de esferas  $n_E$  (o número de discos  $n_D$ ) y con número de parches  $np$  estará dado por:

$$U(m, n) = \sum_{i=1}^{n_E} \sum_{j=1}^{n_E} U_{EE}(r_{ij}) + \sum_{i=1}^{np} \sum_{j=1}^{np} U_{pp}(r_{ij}) \quad (4.1)$$

Donde  $U_{EE}(r_{ij})$  es el potencial repulsivo con el que interactúan las esferas de las moléculas  $m$  y  $n$ , que como ya se mencionó se tendrán dos casos, uno discontinuo y uno continuo. Por lo que se tendrá una barrera de potencial dada por:

$$U_{EE}(r_{ij}) = \begin{cases} \infty, & \text{si } r_{ij} \leq \sigma_{ij} \\ 0, & \text{si } r_{ij} > \sigma_{ij} \end{cases} \quad (4.2)$$

con  $r_{ij}$  la distancia entre los centros de las esferas  $i$  e  $j$  de las moléculas  $m$  y  $n$ , y con  $\sigma_{ij} = (\sigma_i + \sigma_j)/2$  el diámetro promedio de las esferas  $i$  e  $j$ . O será la parte repulsiva del potencial de LJ desplazada hacia arriba  $U_{LJD}(r_{ij})$  dada por:

$$U_{EE}(r_{ij}) = U_{LJD}(r_{ij}) = \begin{cases} U_{LJ}(r_{ij}) + \epsilon_{ij}, & \text{si } r_{ij} \leq r_{c_{ij}}, \\ 0, & \text{si } r_{ij} > r_{c_{ij}}. \end{cases} \quad (4.3)$$

con  $r_{c_{ij}} = 2^{1/6} \sigma_{ij}$  el radio de corte y  $U_{LJ}(r_{ij})$  el potencial de LJ dado por:

$$U_{LJ}(r_{ij}) = 4\epsilon_{ij}((\sigma_{ij}/r_{ij})^{12} - (\sigma_{ij}/r_{ij})^6), \quad (4.4)$$

con  $\epsilon_{ij} = \sqrt{\epsilon_i \epsilon_j}$  la media geométrica de las profundidades de los potenciales,  $\sigma_{ij}$  el diámetro promedio y  $r_{ij}$  las distancia entre las esferas  $i$  e  $j$ .

Siguiendo con el segundo término de la ecuación 4.1, se tiene que  $U_{pp}(r_{ij})$  es el potencial con el que interactuarán atractivamente los parches de las moléculas  $m$  y  $n$ . Nuevamente se tendrán dos casos. El caso discontinuo donde se utiliza un potencial de PC dado por:

$$U_{pp}(r_{ij}) = U_{PC}(r_{ij}) = \begin{cases} \infty, & \text{si } r_{ij} \leq \sigma_{ij} \\ -\epsilon_{PC}, & \text{si } \sigma_{ij} < r_{ij} \leq \lambda \\ 0, & \text{si } r_{ij} > \lambda \end{cases} \quad (4.5)$$

con  $\epsilon_{PC}$  la profundidad y  $\lambda$  el alcance del PC. Para el caso continuo se usa el potencial de LJ dado por la ecuación (4.4).

Los potenciales anteriores son los potenciales intermoleculares. A continuación se pre-

sentan los potenciales intramoleculares que son los encargados de mantener la geometría de la molécula. Para el caso discontinuo se confina a las partículas entre barreras de potencial infinitas dadas por:

$$U_{IN}(r_{ij}) = \begin{cases} \infty, & \text{si } r_{ij} \leq L - \delta/2 \\ 0, & \text{si } L - \delta/2 < r_{ij} < L + \delta/2 \\ \infty, & \text{si } r_{ij} \geq L + \delta/2 \end{cases} \quad (4.6)$$

con  $r_{ij}$  la distancia entre las dos partículas que se ligarán,  $L$  la longitud de la ligadura y  $\delta$  la amplitud de las barreras de confinamiento. Mientras que para el caso continuo se usa un oscilador armónico para el confinamiento de las partículas dado por:

$$U_{IN}(r_{ij}) = \frac{1}{2}k(L - r_{ij})^2 \quad (4.7)$$

En la figura 4.1 se muestran los esquemas de los potenciales intramoleculares e intermoleculares en sus dos versiones: discontinuos (Fig. 4.1b) y continuos (Fig. 4.1c). En el caso donde las moléculas no contengan parches, como es el caso del sistema de la liberación del virión de la membrana que se verá a continuación, su interacción será solo con el término  $U_{EE}$  de la ecuación (4.1), ya sea de forma repulsiva (ec. 4.2) o de forma atractiva (ec. 4.5).

## 4.2. Modelos moleculares

Para los modelos se usó de base el dominio CA de la proteína Gag del VIH-1 ya que es el que tiene el mayor papel en la multimerización de Gag. En este dominio hay 3 interacciones importantes: 2-Fold, 3-Fold, y 6-Fold. Por medio de las interacciones 2F y 3F las moléculas de Gag se unen entre ellas y forman dímeros y trímeros, respectivamente. En cambio, con la interacción 6F no hay interacción directa entre moléculas de Gag, sino que interactúan con la molécula IP6, la cual se rodea de 6 Gag para formar un hexámero quedando ella en el centro. Por simplicidad, como la interacción 6F requiere de una molécula adicional a Gag, esta será descartada. En su lugar se usará la interface de interacciones intra-hexaméricas, formada por regiones no continuas en la base del dominio CA, que Lingappa et al.<sup>45</sup> también considera importante para la multimerización junto con la interacción 6F. Esta interacción se da en dos sitios distintos de Gag de tal modo



que una proteína Gag atrae a dos más, una por su derecha y otra por su izquierda, en un ángulo tal que les será posible cerrar un hexágono al unirse las proteínas necesarios.

De este modo, para los modelos se considerarán las tres interacciones siguientes: 2-Fold, 3-Fold y la interface de interacciones intra-hexaméricas, a la que se llamará IH. En la figura 4.2 se muestra el procedimiento que se realizó para generar un modelo en 2D que pudiera ensamblar la red hexagonal del VIH-1 en su etapa inmadura. Se partió del hexámero que forman los dominio CA de Gag hasta obtener un modelo lo más sencillo posible formado por un disco con distintos parches atractivos.

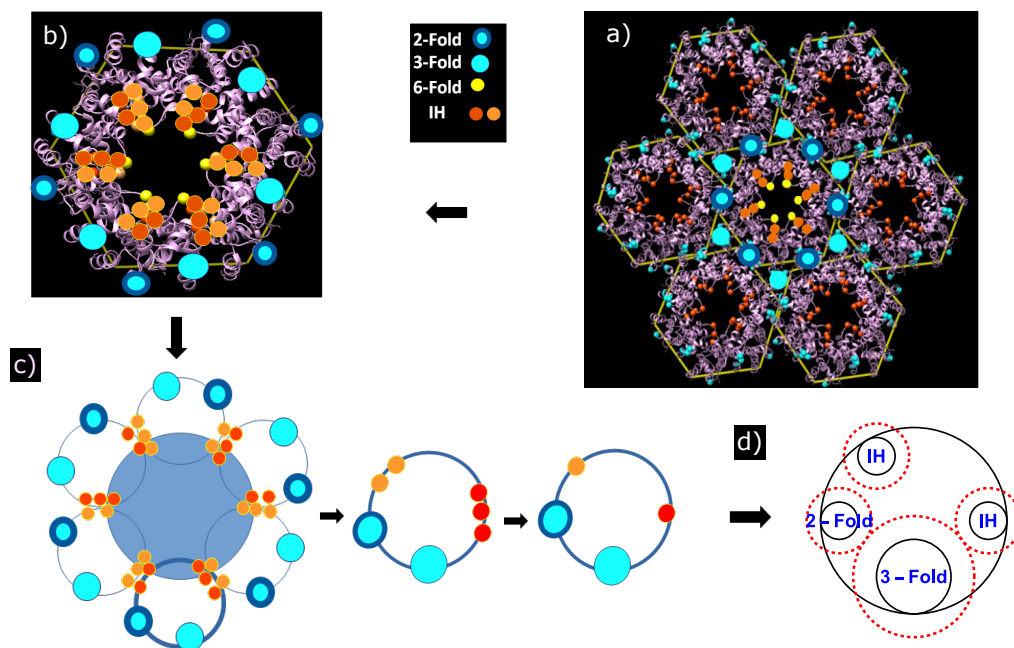


Figura 4.2: Proceso de construcción del modelo de parches atractivos basado en las principales interacciones que intervienen en la formación de la red inmadura del VIH-1. a) Construcción de la red hexagonal del VIH-1 con el dominio CA a partir del PDB 4USN,<sup>30</sup> el cual se obtuvo de Tomografías Crio-Electrónicas de viriones inmaduros del VIH-1. b) Selección de un solo hexámero marcando los sitios donde se ubican las interacciones 2-Fold, 3-Fold, 6-Fold e IH. En la interacción 6-Fold no se muestra la molécula IP6 que se sitúa al centro del hexágono. c) Representación simplificada del hexámero sustituyendo los dominios CA por discos. d) Modelo final en 2D de la molécula parchada para ensamblar la red hexagonal del VIH-1 descartando la interacción 6-Fold.

#### 4.2.1. Red hexagonal del VIH-1 en 2D y 3D

La figura 4.3 muestra los modelos usados para ensamblar la red hexagonal. El modelo en 2D está compuesto por un disco repulsivo de diámetro  $\sigma = 1$ , mientras que los modelos en 3D están compuestos por dos esferas repulsivas traslapadas de diámetros  $\sigma_1 = \sigma_2 = 1$ , separadas una distancia  $d = 0.5\sigma_1$  desde sus centros. Las esferas/discos son decoradas con parches atractivos alrededor de sus ecuadores/perímetro. Las interacciones 2F, 3F e

IH que son principalmente electrostáticas, son modeladas con potenciales de PC (caso discontinuo) y LJ (caso continuo) por medio de parches atractivos. Para el caso continuo en 3D (fig 4.3 (b-c)), se usan parches de dos tamaños distintos. Tanto la interacción 2F e IH son representadas con parches pequeños de diámetro  $\sigma_{par} = 0.15$  que son posicionados a una distancia  $L_{par} = 0.425$  desde el centro de la esfera que las contiene, de tal forma que los parches se mantengan tangentes por dentro a la esfera. Este tamaño solo permite la formación de parejas<sup>76</sup> entre las moléculas cuando interactúan a través de estos parches. La interacción restante 3F, es representada con un parche más grande de diámetro  $\sigma_{ter} = 0.45$  posicionado a una distancia  $L_{ter} = 0.275$  del centro de la esfera que lo contienen manteniéndose tangente por dentro a la esfera. Este parche solo interacciona con parches de la misma especie y forman ternas entre las moléculas cuando se unen. La posición de los parches sobre el ecuador son como sigue. La interacción IH esta compuesta por dos parches, con las características ya mencionadas, IH1 e IH2 posicionados a un ángulo de  $\theta_{IH1} = 0^\circ$  y  $\theta_{IH2} = 120^\circ$ , respectivamente. Los parches restantes 2F y 3F son situados en un ángulo de  $\theta_{2F} = 180^\circ$  y  $\theta_{3F} = 270^\circ$ , respectivamente. En el caso continuo, los parches 2F e IH tendrán los mismos valores para sus potenciales de  $\epsilon = 6$ ,  $\sigma_{IH1i} = \sigma_{IH2i} = \sigma_{2Fi} = 0.15$  y  $r_c = 0.45$ , con  $i = 1, 2$  subíndice que identifica a la esfera que pertenecen, mientras que para el parche 3F serán de  $\epsilon = 6$ ,  $\sigma_{3Fi} = 0.45$  y  $r_c = 1$ .

Para el caso discontinuo (PC), el diámetro de todos los parches es de  $\sigma = 0.1$  y todos se posicionan a una distancia  $L = 0.45$  del centro del disco/esfera para mantenerse por dentro del disco/esfera. También, todos los parches tendrán una profundidad del pozo de  $\epsilon = 1$ . En este caso el alcance  $\lambda$  del PC será lo que dirija la formación de parejas y ternas. De este modo, en 3D los parches 2F e IH tendrán un valor de  $\lambda_{par} = 0.15$ , para formar parejas y el parche 3F un valor de  $\lambda_{ter} = 0.45$  para formar ternas. En 2D los valores del alcance del pozo cambian a  $\lambda_{par} = 0.2$  y  $\lambda_{ter} = 0.4$ . En cuanto a la posición de los parches sobre el ecuador/perímetro, esta se mantiene igual que en el caso continuo. En 2D, con el modelo M1, será el único caso donde solo se usarán los potenciales discontinuos. En la figura 4.3 a) se muestra el modelo M1 con las características de los parches del caso discontinuo, con parches del mismo tamaño. En cambio, en 3D los modelos M2 y M3 si se simularán con las dos versiones del potencial, continuo y discontinuo. Sin embargo, por simplicidad, en la figura 4.3 (b-c) solo se muestran los modelos M2 y M3 en su versión de potencial continuo, con parches de dos tamaños distintos.

Como se ve en la figura 4.3 el modelo M1 no incluye la interacción 2F, mientras que en 3D hay dos versiones, una que no incluye la interacción 2F (M2) y otra que si lo hace (M3). Además, En los modelos en 3D las esferas traslapadas representan los dos subdominios CA-NTD y CA-CTD que componen el dominio CA. En estos subdominios la interacción 3F se encuentra en CA-NTD, mientras que las interacciones 2F e IH se ubican en CA-CTD. Sin embargo, con el objetivo de obtener la red hexagonal más rápidamente, en todos los modelos en 3D las tres interacciones son colocadas en ambas esferas de la molécula.

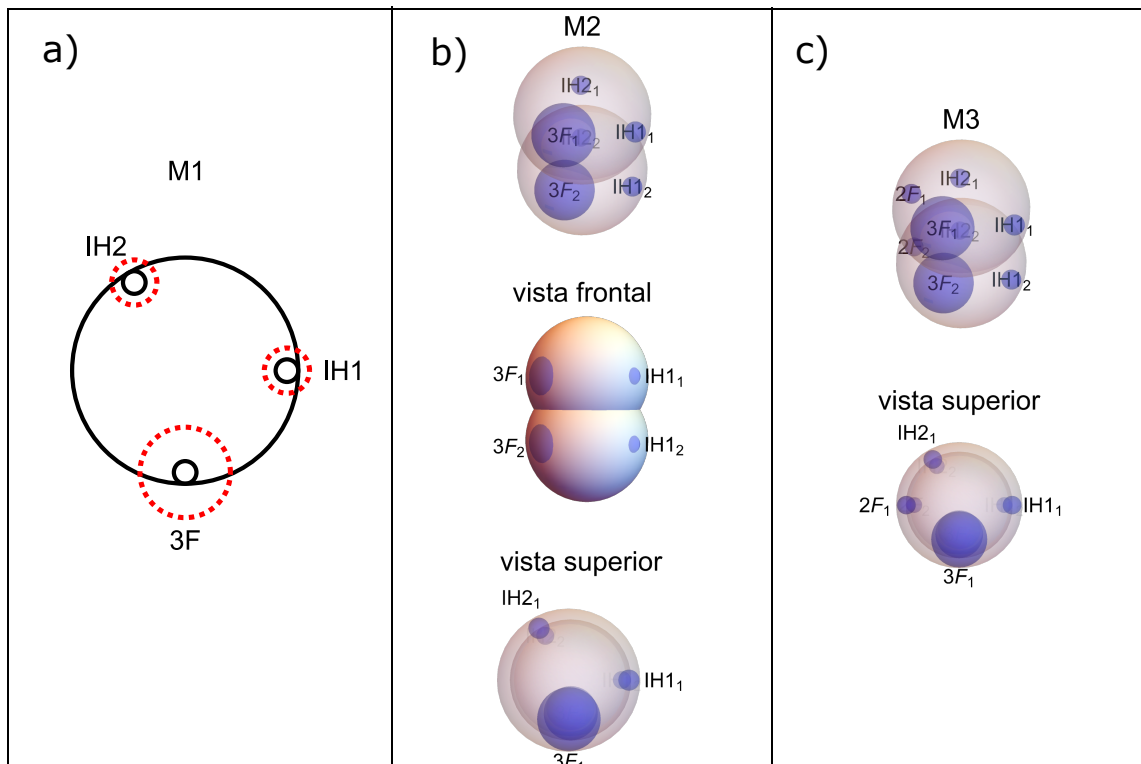


Figura 4.3: Modelos moleculares que forman la red hexagonal del VIH-1 en su etapa inmadura. Los parches atractivos representan las interacciones 2F, 3F e IH del dominio CA de la proteína Gag. Los parches son situados sobre el ecuador de las esferas a un ángulo de  $0^\circ$ ,  $120^\circ$ ,  $180^\circ$ , y  $270^\circ$  para  $IH1_i$ ,  $IH2_i$ ,  $2F_i$  y  $3F_i$ , respectivamente. El subíndice  $i = 1, 2$  distingue a las dos esferas de la molécula. a) Único modelo en 2D que usa la versión de potenciales discontinuos. (b-c) Modelos en 3D presentados con la versión continua del potencial, donde los parches son de dos tamaños distintos. Los modelos M1 y M2 no incluyen las interacción 2F, mientras que el modelo M3 si la incluyen. Los modelos en 3D se muestran con cierta opacidad para tener una visualización de los parches atractivos que se sitúan tangentes por dentro a las esferas de las moléculas. En la vista frontal del modelo M2 no hay opacidad y los parches sobresalen solo por motivos de visualización.

#### 4.2.1.1. Reglas de interacción

La reglas de interacción entre los parches atractivos de dos moléculas distintas se darán de la misma manera en los 3 modelos. No habrá interacción entre los parches de la esfera 1 (E1) con los parches de la esfera 2 (E2) de dos moléculas distintas, solo habrá interacción

entre los parches de la misma especie de esfera E1 con E1 y E2 con E2. Entre los parches pertenecientes a la misma especie de esfera E1-E1 y E2-E2 las interacciones serán como sigue. El parche 3F solo interacciona atractivamente con parches de su misma especie 3F-3F y con nadie más. Lo mismo sucede con el parche 2F, solo interacciona con su misma especie 2F-2F y con nadie más. Los parches IH solo interaccionan atractivamente entre ellos de manera cruzada IH1-IH2, mientras que en los casos IH1-IH1 e IH2-IH2 no habrá interacción.

### 4.2.2. Cápsides icosaédricas

La figura 4.4 muestra los dos moldes usados para ensamblar las cápsides icosaédricas. Estos modelos consisten de dos esferas repulsivas traslapadas de diámetros distintos  $\sigma_1 = 1$  y  $\sigma_2 < \sigma_1$  separados una distancia  $d = 0.5\sigma_1$  entre sus centros. La diferencia de diámetros produce un ángulo  $\phi$  entre la recta que une el ecuador de ambas esferas y la recta que baja perpendicular al ecuador de la esfera de diámetro  $\sigma_1$  (vista frontal Fig.4.4a). Con un ángulo  $\phi > 0$ , el ensamblaje de las moléculas producirá una curvatura hacia la esfera de menor tamaño y dependiendo del valor de este ángulo se pueden obtener estructuras cerradas. Para el modelo M4, mostrado en la figura 4.4a), en la esfera de diámetro  $\sigma_1 = 1$  se incluyen los parches IH1<sub>1</sub>, IH2<sub>1</sub> y 3F<sub>1</sub>, con las mismas posiciones sobre el ecuador que en los modelos para formar la red hexagonal de  $\theta_{IH1} = 0^\circ$ ,  $\theta_{IH2} = 120^\circ$  y  $\theta_{3F} = 270^\circ$ , respectivamente. En este modelo también se descarta la interacción 2F. Para la esfera de menor tamaño  $\sigma_2 < \sigma_1$  se usan los mismo parches que en la esfera 1, con la misma posición sobre el ecuador, pero sus diámetros son escalados con el diámetro de su esfera  $\sigma_2$ . Esto deja a los diámetros de los parches con valores de  $\sigma_{IH1} = \sigma_{IH2} = \sigma_{2F} = 0.15\sigma_2$  y  $\sigma_{3F} = 0.45\sigma_2$ . De igual manera, se escala su posición L para que permanezcan tangentes por dentro a la esfera, con lo que se tiene que  $L_{IH1} = L_{IH2} = L_{2F} = 0.425\sigma_2$  y  $L_{3F} = 0.275\sigma_2$ . El segundo modelo M5, mostrado en la figura 4.4b) es idéntico al modelo M4, con la diferencia de diámetros de las esferas repulsivas  $\sigma_2 < \sigma_1$ , pero ahora si se agrega el parche 2F en ambas esferas.

Para estos modelos en ambos casos solo se hará uso de los potenciales continuos de LJ para las interacciones atractivas y repulsivas.

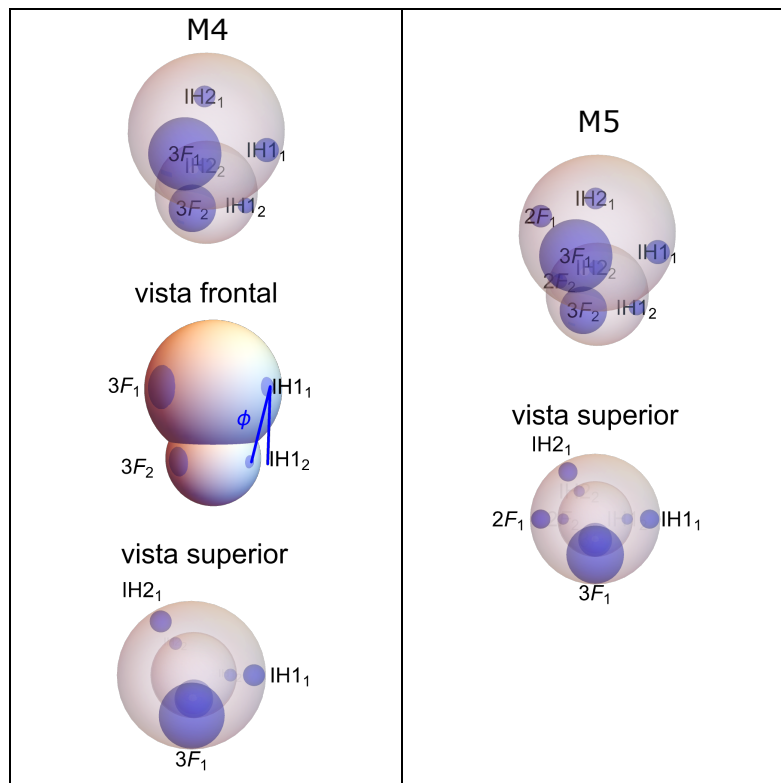


Figura 4.4: Modelos para ensamblar cápsides icosaédricas. Los parches  $IH1_i, IH2_i, 2F_i, 3F_i$  son situados sobre el ecuador a un ángulo de  $0^\circ, 120^\circ, 180^\circ$ , y  $270^\circ$ , respectivamente. El subíndice  $i = 1, 2$  indica a que esfera pertenece el parche. a) Modelo que no incluye la interacción  $2F$ . b) En este modelo el parche  $2F$  si se incluye. Los modelos se muestran con cierta opacidad para visualizar los parches atractivos que se sitúan tangentes por dentro a las esferas que los contiene. En la vista frontal del modelo M4 no hay opacidad y los parches sobresalen solo por motivos de visualización.

### 4.2.3. Liberación del virión de una membrana de doble capa

Para modelar este sistema se tiene que modelar tanto la membrana de doble capa como las proteínas Gag que se encargarán de formar el virión. El modelo de la membrana que se usará está basado en el modelo de grano grueso libre de solvente de Cooke et al.<sup>48</sup> La molécula que se usa para modelar el lípido se muestra en la figura 4.5 a), la cual consta de una cadena de tres esferas tangentes de diámetros  $\sigma_L = 1$ , para las tres esferas. La primer esfera representa la cabeza hidrofílica y las dos esferas restantes representan la cola hidrofóbica del lípido, como se ve en la figura 4.5. Adicionalmente, se propone un modelo, aún más sencillo, para modelar una membrana de doble capa. En la figura 4.5 b) se muestra la molécula para este fin, la cual solo consta de una cadena de dos esfera tangentes, una esfera para la cabeza y una esfera para la cola del lípido.

El modelo de grano grueso que se usará para modelar la proteína Gag estará formado por tres esferas duras de distinto tamaño como se ve en la figura 4.5 c). En el modelo la esfera más grande corresponde al dominio MA, con un diámetro de  $\sigma_{MA} = 1$ , la esfera

siguiente corresponde al dominio CA, con un diámetro de  $\sigma_{CA} = 0.8\sigma_{MA}$  y la esfera más pequeña corresponde al dominio NC, con un diámetro de  $\sigma_{NC} = 0.6\sigma_{MA}$ . En este modelo se omiten el dominio p6 así como los dos espaciadores SP1 y SP2 del Gag. Esto se debe a que la intención no es tener una esfera para cada dominio del Gag, sino que se busca lograr una geometría tipo cónica, con una parte ancha por un extremo y una más delgada por el otro. El propósito de esta geometría es que al irse ensamblando las moléculas Gag sobre la membrana, unas a lado de las otras, le induzcan una curvatura que provoque su envolvimiento conforme más y más moléculas Gag se unan.

#### 4.2.3.1. Reglas de interacción del sistema Gag-Membrana

En este sistema los modelos del lípido y el Gag no interactúan atractivamente por medio de parches, sino directamente con las esferas que los componen. Las reglas de interacción entre las moléculas se darán de la siguiente manera. Entre los lípidos solo habrá interacción atractiva entre las colas, mientras que las interacciones cabeza-cabeza y cabeza-cola serán repulsivas. En cuanto a las moléculas Gag las esferas que la componen solo interactuarán de forma atractiva con las esferas de su misma especie, es decir, CA-CA y NC-NC. Salvo MA-MA que será repulsiva. Mientras que las interacciones cruzadas, de distinta especie, serán repulsivas. Por último, las interacciones entre lípido y Gag serán repulsiva entre cada uno de sus componentes, salvo la única atractiva que se dará entre MA y la cabeza de los lípidos, MA-cabeza.

En estos modelos se usan únicamente los potenciales discontinuos de PC. Los alcances de los PC de los lípidos será de  $\lambda = 2.5\sigma_L$ . Por su parte el Gag tendrá alcances de  $\lambda_{MA} = 1.5\sigma_{MA}$ ,  $\lambda_{CA} = 1.5\sigma_{CA}$  y  $\lambda_{NC} = 1.5\sigma_{NC}$ .

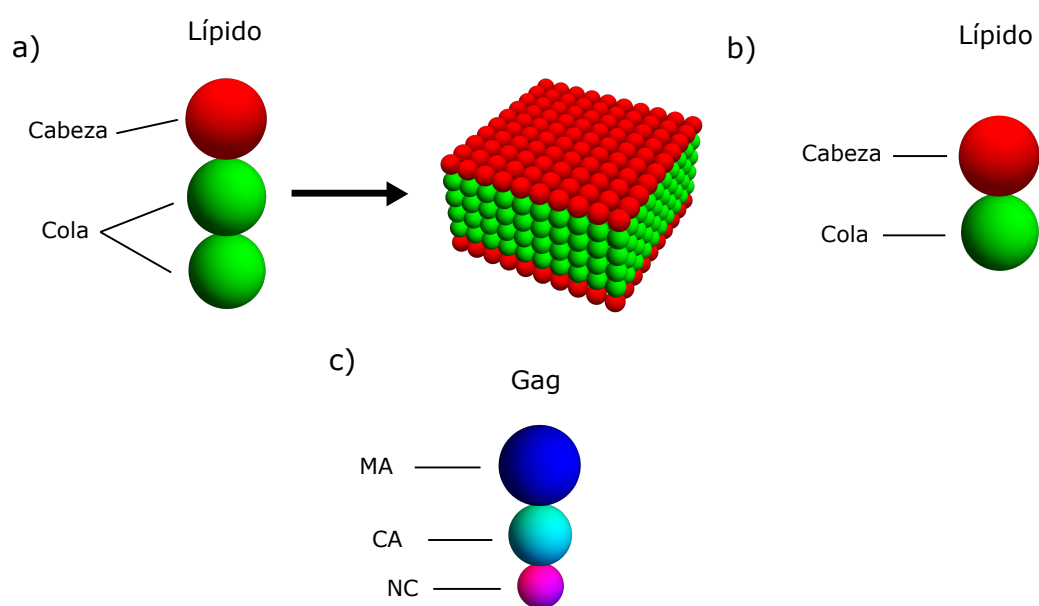


Figura 4.5: Moléculas usadas para modelar la liberación del virión de la membrana plasmática. a) Modelo del lípido para formar una membrana de doble capa. La esfera roja representa la cabeza hidrofílica y las esferas verdes la cola hidrofóbica. b) Modelo más simplificado del lípido para formar una doble capa. En este caso solo se usan dos esferas, una para la cabeza y otra para la cola. c) Molécula que modela el Gag del VIH-1 compuesta por tres esferas tangentes de distintos diámetros.





# Capítulo 5

## Métodos de simulación

La simulación molecular es una herramienta que permite estudiar sistemas a escala micro y nanométrica como partículas coloidales o proteínas. Esta consiste en modelar las partículas o moléculas y llevar a cabo simulaciones computacionales de su movimiento en conjunto. Para esto se deben considerar reglas específicas que tratan de replicar sus condiciones naturales para poder estudiar su comportamiento de manera correcta.

### 5.1. Física estadística: Conceptos básicos

Para entender como se aplica la dinámica molecular es necesario hablar de mecánica estadística. La física o mecánica estadística<sup>77,78</sup> establece una conexión entre la descripción termodinámica de un sistema microscópico con un sistema macroscópico. Esta conexión se hace esencialmente por medio de la función de partición de dicho sistema. La función de partición da información de cómo se encontrará el sistema una vez que llegue a un estado de equilibrio. Esto se hace por medio de un conteo de los estados del sistema para ciertas restricciones termodinámicas como número de partículas, volumen y energía constante. Este conteo de los estados a los que el sistema puede acceder con las restricciones dadas se hace con la suposición de que todos los micro estados posibles son igualmente probables.

La función de partición para un sistema en el ensamble canónico con número de partículas  $N$ , volumen  $V$  y temperatura  $T$  fijos está dado por:

$$Z(N, V, T) = \text{Tr} e^{-\beta H}, \quad (5.1)$$

donde  $H$  es el Hamiltoniano del sistema y  $\beta = T^{-1}$  (tomando la constante de Boltzman

como 1). También se define a  $Tr$  como la traza sobre todos los estados en un sistema cuántico, o clásicamente se define como la integral sobre el espacio fase

$$Tr \equiv \frac{1}{N!h^{dN}} \prod_{\alpha=1}^N \int d\mathbf{p}^{\alpha} d\mathbf{r}^{\alpha}, \quad (5.2)$$

con  $h$  la constante de Planck y  $d$  el número de dimensiones.

Para este ensamble la probabilidad de ocurrencia de un punto dado en el espacio fase, o lo que es lo mismo de un estado del sistema, es

$$P = \frac{e^{-\beta H}}{Z(N, V, T)}. \quad (5.3)$$

Con esto es posible obtener propiedades macroscópicas del sistema promediando sobre el ensamble. Por ejemplo, al no estar fija la energía, el promedio de esta es el promedio del Hamiltoniano sobre el ensamble  $\langle E \rangle \equiv \langle H \rangle = TrPH$ , lo que resulta en

$$\langle E \rangle = \frac{\prod_{\alpha=1}^N \int d\mathbf{p}^{\alpha} d\mathbf{r}^{\alpha} e^{-\beta[\sum_i p_i^2/2m_i + U(\mathbf{r}^{\alpha})]} [\sum_i p_i^2/2m_i + U(\mathbf{r}^{\alpha})]}{\prod_{\alpha=1}^N \int d\mathbf{p}^{\alpha} d\mathbf{r}^{\alpha} e^{-\beta[\sum_i p_i^2/2m_i + U(\mathbf{r}^{\alpha})]}}, \quad (5.4)$$

donde se sustituyó la parte cinética y potencial del Hamiltoniano  $H = K + U$ .

La ecuación (5.4) es la manera en la que se pueden calcular observables del sistema de partículas, al promediar a la observable sobre el ensamble  $\langle A \rangle = TrPA$ , con  $A$  la observable de interés. Sin embargo, la función de partición no siempre se puede obtener para todos los sistemas, por lo que se hace uso de otros métodos para estudiar los sistemas.

### 5.1.1. Hipótesis ergódica

Un promedio sobre todos los posibles estados de un sistema es un promedio sobre el ensamble. Pero esta no es la forma en la que usualmente se piensa cuando se habla de promediar una variable de un sistema. En la mayoría de experimentos se hacen una serie de mediciones en ciertos intervalos de tiempo y después se promedian esas mediciones. De hecho, esta es la idea detrás de la Dinámica Molecular (DM), ya que se calcula numéricamente la evolución temporal del sistema y se promedia la variable de interés sobre un tiempo lo suficientemente largo.<sup>77</sup> Por ejemplo, se puede considerar la densidad promedio

$\rho_i(r)$  de un fluido a una distancia  $r$  de un determinado átomo  $i$ . La densidad instantánea dependerá de las posiciones  $\mathbf{r}_j$  de todas las partículas  $j$ , y conforme pase el tiempo las coordenadas de las partículas cambiarán de acuerdo a las ecuaciones de movimiento de Newton. Por lo tanto, la densidad alrededor de la partícula  $i$  cambiará con el tiempo. Si se conocieran todas las posiciones  $\mathbf{r}^N(0)$  y momentos  $\mathbf{p}^N(0)$  iniciales de las partículas, en principio, se conocería la evolución temporal de la densidad  $\rho_i(r, \mathbf{r}^N(0), \mathbf{p}^N(0))$ . En una simulación de dinámica molecular se calcula la densidad promediada en el tiempo  $\overline{\rho_i(r)}$  de un sistema con  $N$  partículas, volumen  $V$  y energía  $E$  como:

$$\overline{\rho_i(r)} = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t dt' \rho(r, t'). \quad (5.5)$$

Al escribir esta ecuación se ha asumido que, para tiempos suficientemente grandes, el promedio temporal es independiente de las condiciones iniciales. Aunque esto en general no siempre se cumple, se asumirá que una vez fijados  $N, V$  y  $E$  el promedio temporal no dependerá de las condiciones iniciales. Si esto se cumple, entonces la densidad promedio no se verá afectada si se promedia sobre varias condiciones iniciales diferentes. Es decir, si se hacen varias simulaciones de DM de un sistema de partículas con distintas coordenadas y momentos iniciales, pero con los mismos valores de  $N, V$  y  $E$  se tendrá lo siguiente

$$\overline{\rho_i(r)} = \frac{\sum_{\text{condiciones iniciales}} \left( \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t dt' \rho(r, \mathbf{r}^N(0), \mathbf{p}^N(0), t') \right)}{\text{número de condiciones iniciales}}. \quad (5.6)$$

Si se consideran solo las condiciones iniciales compatibles con los valores fijos de  $N, V$  y  $E$ , entonces la suma sobre las condiciones iniciales se puede reemplazar por una integral

$$\frac{\sum_{\text{condiciones iniciales}} f(\mathbf{r}^N(0), \mathbf{p}^N(0))}{\text{número de condiciones iniciales}} \rightarrow \frac{\int_E d\mathbf{r}^N d\mathbf{p}^N f(\mathbf{r}^N(0), \mathbf{p}^N(0))}{\Omega(N, V, E)}, \quad (5.7)$$

donde  $f$  es una función arbitraria de las condiciones iniciales y  $\Omega(N, V, E) = \int_E d\mathbf{r}^N d\mathbf{p}^N$ . Un promedio del espacio fase corresponde al promedio sobre el ensamble y se denota por  $\langle \dots \rangle$  para diferenciar del promedio temporal denotado con una barra. Ahora, si se intercambia el orden del promedio temporal y del promedio sobre el ensamble se tendrá

$$\overline{\rho_i(r)} = \frac{1}{t} \int_0^t dt' \langle \rho(r, \mathbf{r}^N(0), \mathbf{p}^N(0), t') \rangle_{NVE}. \quad (5.8)$$

El promedio sobre el ensamble de la ecuación anterior no depende del tiempo  $t'$ . Entonces, promediar sobre todas las coordenadas iniciales del espacio fase es equivalente a promediar sobre la evolución temporal de las coordenadas del espacio fase. Debido a esto se puede quitar el promedio temporal de la ecuación (5.8) convirtiéndose en

$$\overline{\rho_i(r)} = \langle \rho(r) \rangle_{NVE}. \quad (5.9)$$

Esta ecuación enuncia que, si se quiere calcular el promedio de una cantidad en función de las coordenadas y momentos de un sistema de partículas, se puede obtener dicha cantidad al promediarla en el tiempo (aproximación con DM) o al promediarla sobre el ensamble (aproximación con MC). Hay que tener en cuenta que la ecuación (5.9) no se cumple en general, Por ejemplo, hay sistemas como vidrios o fases metaestables que no son ergódicas. Aun así, se asume que la hipótesis ergódica como se muestra en la ecuación (5.9) es válida para los sistemas que se estudian en simulaciones computacionales.

## 5.2. Monte Carlo

La ecuación (5.4) de la sección anterior enuncia que se puede calcular el valor de una observable  $A$  de un sistema de partículas al promediarla sobre el ensamble

$$\langle A \rangle = \frac{\int d\mathbf{p}^N d\mathbf{r}^N e^{-\beta H(\mathbf{r}^N, \mathbf{p}^N)} A(\mathbf{r}^N, \mathbf{p}^N)}{\int d\mathbf{p}^N d\mathbf{r}^N e^{-\beta H(\mathbf{r}^N, \mathbf{p}^N)}}, \quad (5.10)$$

con  $N$  es el número total de partículas. El Hamiltoniano  $H = K + U$  y la observable están en términos de las coordenadas y momentos de las partículas. Si se quisiera evaluar numéricamente esta ecuación en el espacio de configuraciones, digamos por la regla de Simpson, se tendría que generar una malla de puntos a lo largo de los ejes de coordenadas. Si se toman  $m$  puntos equidistantes a lo largo de los ejes coordenados, el número de integrales que se tendrían que evaluar sería de  $m^{nd}$ , con  $nd$  el número de dimensiones. Aun si se toma un número bajo de puntos de  $m = 5$  para un sistema de 100 partículas en 3 dimensiones, el número de integrales que se tendrían que evaluar es del orden de  $10^{210}$ . Cálculos de esta magnitud son imposibles de realizar debido al tiempo que tomaría terminarlos. Debido a esto es necesario evaluar con otro método las integrales de la ecuación (5.10). Una mejor forma de realizar estos promedios es con el método de Monte Carlo tomando en cuenta la importancia del muestreo.

El método simple de Monte Carlo consiste en evaluar una integral  $I$ :

$$I = \int_a^b dx f(x), \quad (5.11)$$

la cual, en lugar de usar el método de cuadraturas, se escribirá de la siguiente manera

$$I = (b - a) \langle f(x) \rangle, \quad (5.12)$$

Donde  $\langle f(x) \rangle$  es el promedio no ponderado de  $f(x)$  en el intervalo  $[a, b]$ . Con el método simple de MC de muestreo aleatorio, este promedio es determinado por evaluar  $f(x)$  en un gran número de valores de  $x$  distribuidos aleatoriamente en el intervalo  $[a, b]$ . Si  $L$  denota el número de puntos evaluados, entonces, si  $L \rightarrow \infty$  este procedimiento dará el valor correcto de  $I$ . Este método sigue siendo de poca utilidad para calcular promedios del tipo de la ecuación (5.10) porque la mayor parte del tiempo se gasta en puntos donde el factor de Boltzman  $fb = \exp[-\beta U(\mathbf{r})]$  es despreciable. Por lo tanto, es mejor hacer un muestreo en lugares donde el factor de Boltzman no sea despreciable. Una elección de un mejor muestreo es la idea principal detrás del muestreo introducido por Metropolis *et al.*<sup>79</sup> en 1953.

### 5.2.1. Método de Metropolis

Algo importante sobre el método de Monte Carlo es que sólo se considera el espacio de configuraciones al eliminar la parte de los momentos. Al quitar la parte de la dinámica, este método solo se puede usar para sistemas que estén en equilibrio. Además, la mayoría de las veces no se necesita saber el valor de una sola integral sino el valor de la razón de dos integrales

$$\langle A \rangle = \frac{\int d\mathbf{r}^N e^{-\beta U(\mathbf{r}^N)} A(\mathbf{r}^N)}{\int d\mathbf{r}^N e^{-\beta U(\mathbf{r}^N)}}. \quad (5.13)$$

Al escribir la función de partición como  $Z \equiv \int e^{-\beta U(\mathbf{r}^N)}$  la razón  $e^{-\beta U(\mathbf{r}^N)}/Z$  de la ecuación anterior, al igual que en la ecuación (5.3), es la densidad de probabilidad de encontrar al sistema en una configuración  $\mathbf{r}^N$

$$N(\mathbf{r}^N) \equiv \frac{e^{-\beta U(\mathbf{r}^N)}}{Z}, \quad (5.14)$$

y al igual que en (5.4), el valor promedio de la observable será  $\langle A \rangle = \text{Tr} N(\mathbf{r}^N) A$ .

Ahora supóngase que es posible generar puntos en el espacio de configuraciones de acuerdo a la distribución de probabilidad  $N(\mathbf{r}^N)$ . Esto quiere decir que, en promedio, el número de puntos  $n_i$  generados por unidad de volumen alrededor de  $\mathbf{r}^N$  es igual a  $LN(\mathbf{r}^N)$ , con  $L$  el número total de puntos generados. Si se utiliza lo anterior, la ecuación (5.13) se puede escribir como:

$$\langle A \rangle \approx \frac{1}{L} \sum_i^L n_i A(\mathbf{r}^N), \quad (5.15)$$

donde los puntos  $n_i$  son generados por medio de una caminata aleatoria en el espacio de configuraciones donde el factor de Boltzman no es despreciable.

La forma en la que se genera la caminata aleatoria es como sigue. Inicialmente se prepara al sistema en una configuración  $\mathbf{r}^N$ , el cual se denota por  $o$  (old), con un factor de Boltzman  $\exp[-\beta U(o)]$ . Después, se genera una configuración de prueba  $\mathbf{r}'^N$ , denotada por  $n$  (new), al realizar un desplazamiento aleatorio  $\Delta$  a  $\mathbf{r}^N$ . Esta nueva configuración tendrá un factor de Boltzman  $\exp[-\beta U(n)]$ . Ahora se debe decidir si se acepta o se rechaza la nueva configuración. Aunque hay varias formas de decidir si se acepta o se rechaza la configuración  $\mathbf{r}'^N$ , en el esquema de Metropolis la probabilidad de aceptación de ir de una configuración  $o$  a una configuración  $n$  está dada por

$$\text{acc}(o \rightarrow n) = \exp[-\beta(U(n) - U(o))] < 1. \quad (5.16)$$

Para decidir si se acepta o se rechaza la configuración de prueba se genera un número aleatorio, denotado por  $\text{Ranf}$  (del inglés “Random number uniform”), con una distribución uniforme en un intervalo  $[0, 1]$ . El movimiento de prueba se aceptará si  $\text{Ranf} < \text{accp}(o \rightarrow n)$  y se rechazará en el caso contrario. Esta regla garantiza que la probabilidad de aceptar un movimiento de prueba de  $o$  a  $n$  sea en efecto igual a la probabilidad de aceptación  $\text{acc}(o \rightarrow n)$ .

De este modo, un programa básico de Monte Carlo consistiría de los siguientes pasos principales:

1. Se selecciona una partícula aleatoriamente y se calcula su energía  $U(\mathbf{r}^N)$ .
2. Se deslaza a la partícula,  $\mathbf{r}' = \mathbf{r} + \Delta$ , y se calcula su nueva energía  $U(\mathbf{r}'^N)$ .

3. Aceptar el movimiento de  $\mathbf{r}$  a  $\mathbf{r}'$  con probabilidad

$$acc(o \rightarrow n) = \text{mín} (1, \exp(-\beta[U(\mathbf{r}'^N) - U(\mathbf{r}^N)])) .$$

### 5.2.2. Condiciones de frontera y mínima imagen

Una simulación se realiza con unos cientos o miles de partículas dentro de una caja de simulación. Es decir, con un sistema finito. Pero está claro que un sistema de este tipo no tiene el tamaño de aquel al que intenta emular, por ejemplo un líquido o un gas. En un sistema real, como un líquido contenido, sólo una pequeña fracción de los átomos está cerca de las paredes del contenedor. Estos átomos experimentan condiciones diferentes a las del resto, pero esto no afecta el comportamiento de la gran cantidad de átomos al interior. En una simulación de unas 1000 partículas en tres dimensiones, alrededor de 500 partículas estarán cerca de las paredes dejando pocas partículas al interior. Es claro que un sistema con tan pocas partículas al interior no representarán adecuadamente a un sistema real. No se puede tan sólo incrementar el número de partículas debido al enorme gasto de tiempo de cómputo y de memoria para almacenamiento de datos que se requeriría. Debido a esto, la representación de un sistema real se tiene que hacer de otra manera.

Un sistema que este contenido, pero a la vez libre de paredes físicas, se obtiene introduciendo condiciones periódicas de frontera.<sup>80</sup> La periodicidad en la frontera es equivalente a tener un sistema infinito de réplicas de la región de simulación, como se ve en el esquema de la figura 5.1. Aun así, la estadística del sistema está restringida al  $N$  número de partículas de la región de simulación. Esto se logra de la siguiente manera. Si una partícula sale por una de las caras de la caja de simulación, inmediatamente es introducida por la cara opuesta. Si se tiene una caja cúbica de longitud  $L$ , la reinserción de partículas se da bajo la siguiente regla. Para el eje  $x$

- Si  $r_{ix} \geq L_x$ , entonces reemplazar por  $r_{ix} - L_x$ ;
- Si  $r_{ix} \leq 0$ , entonces reemplazar por  $r_{ix} + L_x$ ,

con  $r_{ix}$  la componente  $x$  de la posición de la partícula  $i$ . Se procede de la misma forma para los tres ejes.

La implementación de las condiciones de frontera implica la necesidad de especificar claramente cuando se dará la interacción de las partículas. Es decir, una partícula cerca

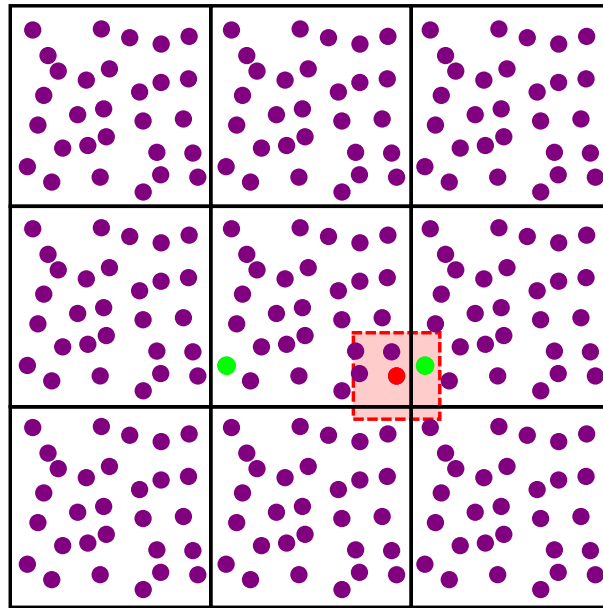


Figura 5.1: Condiciones periódicas de frontera en 2D. La caja de simulación original se replica alrededor de esta. Para definir si hay interacción entre dos partículas se elige la distancia mínima entre la partícula  $i$  (disco rojo) y la partícula  $j$  (disco verde) de la misma caja o de la caja adyacente.

de una pared ¿interactuará o no? con las partículas de la celda adyacente. Como se ve en la celda central de la figura 5.1, que representa la región original de simulación, la partícula en rojo tiene un rango de interacción que cubre a partículas de su misma caja, pero también alcanza una partícula de la caja adyacente. Para cubrir esos casos se usa lo que se llama mínima imagen entre las partículas. La mínima imagen se aplica de la siguiente manera. Se elijen dos partículas  $i$  y  $j$ , para definir si la partícula  $j$  está dentro del rango de interacción de la partícula  $i$  se debe de tomar la mínima de las siguientes distancias: la distancia con la partícula  $j$  de la misma caja o la distancia con la imagen de la partícula  $j$  de la caja de alrededor. En la simulación esto se determina por medio la longitud de la caja de simulación. Si en el eje  $x$  la distancia entre las partículas  $i$  y  $j$  es mayor que  $L_x/2$  entonces la distancia con la imagen de la partícula  $j$  en la caja de al lado será menor. Para tomar la distancia correcta entre los pares de partículas se sigue la siguiente regla

- Si  $r_{ijx} > L_x/2$ , entonces reemplazar por  $r_{ij} - L_x$ ;
- Si  $r_{ijx} < -L_x/2$ , entonces reemplazar por  $r_{ij} + L_x$ ,

con  $r_{ijx}$  la componente  $x$  de la distancia entre las partículas  $i$  y  $j$ . La regla se aplica a las tres componentes de la distancia.



Se debe tener en cuenta que el uso de las condiciones periódicas de frontera limitará el uso de interacciones a aquellas que no tengan un rango más allá de la mitad de la caja de simulación.

### 5.3. Dinámica Molecular

A diferencia del método de Monte Carlo que se usa para estudiar sistemas en equilibrio y calcular propiedades que no dependan de la evolución temporal del sistema. En las simulación con DM, al incluir las posiciones y velocidades de las partículas, se puede seguir la evolución del sistema hasta que este alcance el equilibrio. Entonces se puede calcular las propiedades termodinámicas del sistema incluyendo propiedades que sean dependientes del tiempo. Además, con DM se pueden calcular propiedades de transporte y es un excelente método para aproximar un gran rango de materiales.

En una simulación de DM se selecciona un modelo que consiste en un sistema de  $N$  partículas. Se resuelven las ecuaciones de movimiento de Newton para este sistema hasta que sus propiedades dejen de cambiar significativamente en el tiempo y solo oscilen alrededor de un valor fijo, es decir, hasta que el sistema alcance el equilibrio. Conforme transcurre la simulación, se miden las propiedades de interés para obtener su valor promedio una vez que el sistema esté en equilibrio.

Para medir una observable en una simulación de DM, lo primero que se tiene que hacer es poder expresar dicha observable en términos de las posiciones y velocidades de las partículas del sistema. Por ejemplo, una forma de calcular la temperatura en un sistema clásico de muchas partículas a partir de sus velocidades es a través del teorema de equipartición de la energía. Como cada grado de libertad traslacional contribuye a la energía cinética con  $\frac{1}{2}mv_\alpha^2 = K_B T/2$ , donde  $v_\alpha$  es la  $\alpha$  componente de la velocidad de una partícula, la temperatura instantánea de un sistema de  $N$  partículas con  $n$  grados de libertad estará dada por:

$$T(t) = \sum_{i=1}^N \frac{m_i v_i^2(t)}{nk_B}. \quad (5.17)$$

Las fluctuaciones relativas en la temperatura serán del orden de  $1/\sqrt{n}$ . Para un sistema de  $N = 500$  partículas en 3 dimensiones las fluctuaciones en el cálculo de la temperatura serían del orden del 2.6%. Por otro lado, si lo que se quiere es mantener una  $T$ , constante

la temperatura instantánea  $T(t)$  se puede ajustar escalando las velocidades con el factor  $\sqrt{T/\overline{T}}$ . También se puede mantener una temperatura constante por medio de algún termostato. Como puede ser el de Andersen<sup>81</sup> que somete a la partículas a colisiones estocásticas con un baño térmico o el termostato de cadenas de Nose-Hoover que usa un Hamiltoniano extendido con coordenadas y momentos artificiales<sup>82</sup>

Un programa simple de dinámica molecular en el que se tienen sus principales características está formado como sigue:

1. Se leen los parámetros con los que se iniciará el sistema, por ejemplo: número de partículas, temperatura, presión, paso en el tiempo  $\Delta t$ .
2. Se inicializa el sistema. Se dan las posiciones y velocidades iniciales.
3. Se calculan las fuerzas sobre todas las partículas (al tiempo  $t = t_0$ ). Aquí se inicia el ciclo principal de cálculos.
4. Se integran las ecuaciones de movimiento de Newton. Se calculan las nuevas posiciones y velocidades después de un tiempo  $t = t_0 + \Delta t$ . Con las nuevas posiciones se puede calcular la temperatura actual, la energía potencial y cinética, teniendo en cuenta que la energía total se debe mantener constante para todo tiempo  $t$ . Aquí termina el ciclo principal y se regresa al paso anterior para calcular nuevamente la fuerza con las posiciones al tiempo  $t = t_0 + \Delta t$ .
5. Después de terminar varias veces el ciclo de los pasos 3 y 4 hasta alcanzar la cantidad de tiempo deseada, se calculan e imprimen los promedios de las cantidades medidas.

Estos pasos son los constituyentes básicos de una simulación de DM. Con el cálculo de las fuerzas, el paso en lo que se gasta mayor tiempo de cómputo. Hay varias formas de economizar tiempo en este cálculo. Más adelante se hablará de uno de los que se usa en este trabajo.

### 5.3.1. Integración de las ecuaciones de movimiento: Velocity Verlet

Hay varias formas de integrar las ecuaciones de movimiento para encontrar las nuevas posiciones y velocidades de las partículas. Dos de los más comunes usan los algoritmos de

Verlet y Velocity Verlet.<sup>77</sup> Para obtener estos algoritmos se parte de una serie de Taylor de las coordenadas alrededor del tiempo  $t$

$$r(t + \Delta t) = r(t) + v(t)\Delta t + \frac{F(t)}{2m}\Delta t^2 + \frac{\Delta t^3}{3!}\ddot{r} + O(\Delta t^4), \quad (5.18)$$

$$r(t - \Delta t) = r(t) - v(t)\Delta t + \frac{F(t)}{2m}\Delta t^2 - \frac{\Delta t^3}{3!}\ddot{r} + O(\Delta t^4), \quad (5.19)$$

con  $F(t)$  las fuerzas entre las partículas. Al sumar las ecuaciones anteriores se obtiene

$$r(t + \Delta t) + r(t - \Delta t) = 2r(t) + \frac{F(t)}{m}\Delta t^2 + O(\Delta t^4), \quad (5.20)$$

$$r(t + \Delta t) \approx 2r(t) - r(t - \Delta t) + \frac{F(t)}{m}\Delta t^2, \quad (5.21)$$

donde  $\Delta t$  es el paso en el tiempo en la DM. Si ahora se restan las ecuaciones (5.18) y (5.19) se obtendrá la velocidad

$$r(t + \Delta t) - r(t - \Delta t) = 2v(t)\Delta t + O(\Delta t^3), \quad (5.22)$$

$$v(t) = \frac{r(t + \Delta t) - r(t - \Delta t)}{2\Delta t} + O(\Delta t^2). \quad (5.23)$$

Las ecuaciones (5.21) y (5.23) forman el algoritmo de Verlet. Como se puede ver de (5.21), no se usan las velocidades para calcular las nuevas posiciones, sino que se usan las posiciones actuales  $r(t)$  y las anteriores  $r(t - \Delta t)$  para conocer las nuevas posiciones. Además, las nuevas posiciones y velocidades no se calculan al mismo tiempo, ya que primero se necesita conocer las posiciones siguientes  $r(t + \Delta t)$  para poder calcular, tan sólo, las velocidades actuales  $v(t)$ .

Este inconveniente se elimina al usar el algoritmo Velocity Verlet, el cual se obtiene como sigue. Se usa la ec.(5.23) para obtener  $r(t - \Delta t) = r(t + \Delta t) - 2v(t)\Delta t$ , al sustituir esta expresión en (5.21) y resolver para  $r(t + \Delta t)$  se obtiene

$$r(t + \Delta t) = r(t) + v(t)\Delta t + \frac{F(t)}{2m}\Delta t^2. \quad (5.24)$$

Ahora se usa la ec.(5.23) para obtener  $v(t + \Delta t)$ , que queda como

$$v(t + \Delta t) = \frac{r(t + 2\Delta t) - r(t)}{2\Delta t}, \quad (5.25)$$

se usa (5.21) para obtener  $r(t + 2\Delta t) = 2r(t + \Delta t) - r(t) - F(t + \Delta t)\Delta t^2/m$  y se sustituye en la ecuación (5.25) para quedar

$$v(t + \Delta t) = \frac{r(t + \Delta t) - r(t)}{\Delta t} + \frac{1}{2m}F(t + \Delta t)\Delta t, \quad (5.26)$$

y al usar (5.24) para sustituir  $r(t + \Delta t)$  en (5.26) se obtiene finalmente la expresión para las nuevas velocidades

$$v(t + \Delta t) = v(t) + \frac{F(t + \Delta t) + F(t)}{2m}\Delta t. \quad (5.27)$$

Las ecuaciones (5.24) y (5.27) forman el algoritmo Velocity Verlet. Con este algoritmo las velocidades se calculan en el mismo paso en el tiempo  $\Delta t$ , sólo se requiere calcular las fuerzas  $F(t + \Delta t)$  inmediatamente después de haber movido las partículas para obtener las nuevas velocidades.

### 5.3.2. Unidades reducidas

Es conveniente escribir las cantidades físicas en términos de unidades reducidas. Esto es escoger convenientemente las unidades de energía, longitud y masa y escribir todas las demás cantidades en términos de estas unidades. Hacer esto trae varios beneficios, como trabajar con números que estén cerca de la unidad y no con cantidades muy pequeñas relacionadas con escalas atómicas lo cual llevaría a errores numéricos en los cálculos. Otra razón para usar este tipo de unidades tiene que ver con el escalamiento, ya que un sólo modelo puede representar a todo un grupo de sistemas, y una vez calculadas las propiedades físicas en unidades reducidas estas se pueden escalar a las unidades apropiadas para el problema de interés.

Una de las elecciones que se usa para definir las unidades básicas es tomar  $\sigma$ ,  $\epsilon$  y  $m$  como las unidades de longitud, energía y masa, respectivamente. Donde  $\sigma$  es el diámetro de las partículas y  $\epsilon$  la profundidad del potencial de interacción. Con estas unidades básicas se derivan otras unidades, por ejemplo la unidad de tiempo queda como  $\sigma\sqrt{m/\epsilon}$ , y la unidad de temperatura como  $\epsilon/k_B$ . En términos de estas nuevas unidades se obtiene las cantidades reducidas, representadas con un asterisco, como distancia  $r^* = r/\sigma$ , energía

$U^* = U/\epsilon$ ; densidad  $\rho^* = \rho\sigma^{nd}$ , con  $nd$  el número de dimensiones y  $\rho = N_m/V$ , con  $N_m$  el número de moléculas; temperatura  $T^* = Tk_B/\epsilon$  con  $kb$  la constante de Boltzman y  $T = 2Ke/k_bndN$ , con  $Ke$  la energía cinética; y por último tiempo  $t^* = t\sqrt{\epsilon/\sigma^2m}$ . En la sección de resultados se usan las unidades reducidas, pero serán utilizadas sin el asterisco.

## 5.4. Dinámica Molecular de Potenciales Discontinuos

Hasta el momento no se ha hablado de qué tipo de potenciales se pueden usar en cada método, ya sea MC o DM. La razón es que se puede usar cualquier potencial que se adecúe a las necesidades del sistema de estudio. Las fuerzas  $F$  sobre todas las partículas, que se necesita para resolver las ecuaciones de movimiento, se calculan de manera distinta para potenciales discontinuos y continuos. Se tiene que tener en consideración el radio de corte del potencial para el caso continuo. En el presente trabajo se utilizan tanto potenciales continuos como discontinuos para representar los parches atractivos de las moléculas. Para el caso discontinuo se usan potenciales de pozo cuadrado y además, para integrar las ecuaciones de movimiento, se consideran las colisiones entre los pares de partículas del sistema. Este método de dinámica molecular de potenciales discontinuos (DMPD), basado en los tiempos de colisión de las partículas, fue desarrollado por Alder y Wainwright<sup>83</sup> y se describe a continuación

### 5.4.1. Descripción del método

En un sistema de partículas de diámetros iguales que interactúan por medio de un potencial de pozo cuadrado:

$$U = \begin{cases} \infty & r \leq \sigma_1 \\ U_0 & \sigma_1 < r < \sigma_2 \\ 0 & r \geq \sigma_2, \end{cases} \quad (5.28)$$

con  $r$  la distancia entre los centros de un par de partículas,  $U_0$  la profundidad del pozo,  $\sigma_1$  el diámetro de las partículas y  $\sigma_2$  el alcance del pozo de potencial. Este potencial permite que la secuencia de eventos del sistema de partículas sea descrita por una serie de colisiones entre dos cuerpos. Esto es debido a que una partícula solo experimentará cambios en su velocidad cuando esté a una distancia  $\sigma_1$  o  $\sigma_2$  de otra partícula. Además,

nunca habrá, en un sistema finito, más de dos partículas que cambien sus velocidades a la vez.

Una vez preparada la inicialización del sistema con número de partículas, temperatura, densidad, posiciones y velocidades iniciales, se calcula exactamente el tiempo en el cual se dará la primera colisión del sistema. Se considerará como una colisión a cualquier interacción atractiva o repulsiva entre las partículas. El tiempo de colisión se obtiene al calcular el tiempo que le toma a un par de partículas encontrarse a una distancia  $\sigma_1$  o  $\sigma_2$  de sus centros. Si dos partículas,  $i$  y  $j$ , tienen posiciones iniciales  $\mathbf{r}_{i0}$  y  $\mathbf{r}_{j0}$  y velocidades  $\mathbf{v}_i$  y  $\mathbf{v}_j$ , después de un tiempo  $t$ , el cuadrado de la separación de las partículas será

$$(\mathbf{r}_i - \mathbf{r}_j)^2 = (\mathbf{r}_{i0} - \mathbf{r}_{j0})^2 + 2t(\mathbf{r}_{i0} - \mathbf{r}_{j0}) \cdot (\mathbf{v}_i - \mathbf{v}_j) + t^2(\mathbf{v}_i - \mathbf{v}_j)^2. \quad (5.29)$$

Entonces, el tiempo requerido para que se dé una colisión atractiva o repulsiva es

$$t_{ij}^{(\alpha)} = \frac{-b_{ij} \pm (b_{ij}^2 - v_{ij}^2 C_{ij}^{(\alpha)})^{1/2}}{v_{ij}^2}, \quad (5.30)$$

donde

$$\begin{aligned} \mathbf{r}_{ij} &= \mathbf{r}_{i0} - \mathbf{r}_{j0} & b_{ij} &= \mathbf{r}_{ij} \cdot \mathbf{v}_{ij} \\ \mathbf{v}_{ij} &= \mathbf{v}_i - \mathbf{v}_j & C_{ij}^\alpha &= r_{ij}^2 - \sigma_\alpha^2. \end{aligned}$$

$t_{ij}^{(1)}$  es el tiempo requerido para una colisión repulsiva, la cual se da al chocar las partículas y  $t_{ij}^{(2)}$  es el tiempo requerido para una colisión atractiva con el pozo de potencial.  $C_{ij}^{(1)}$  siempre debe ser positiva, de otro modo habría un traslape entre las partículas. En cambio,  $C_{ij}^{(2)}$  puede ser positiva o negativa dependiendo de si la separación de las partículas está dentro o fuera del rango del pozo atractivo.

Para conocer en qué momento se da la primera colisión es necesario calcular el tiempo de colisión de todos los pares de partículas del sistema. Al calcularse los tiempos de colisión estos se clasifican mediante la siguiente secuencia.

I  $b_{ij} < 0$  (Centros se aproximan.)

(a)  $C_{ij}^{(2)} < 0$  ( $r_{ij}$  está dentro del rango atractivo.)

(1)  $b_{ij}^2 - v_{ij}^2 C_{ij}^{(1)} > 0$  (Partículas colisionan. Ocurre colisión repulsiva.)

$$t_{ij}^{(1)} = \frac{-b_{ij} - (b_{ij}^2 - v_{ij}^2 C_{ij}^{(1)})^{1/2}}{v_{ij}^2}.$$

(2)  $b_{ij}^2 - v_{ij}^2 C_{ij}^{(1)} < 0$  (Partículas no colisionan. Ocurre colisión atractiva.)

$$t_{ij}^{(2)} = \frac{-b_{ij} + (b_{ij}^2 - v_{ij}^2 C_{ij}^{(2)})^{1/2}}{v_{ij}^2}.$$

(b)  $C_{ij}^{(2)} > 0$  ( $r_{ij}$  está fuera del rango atractivo.)

(1)  $b_{ij}^2 - v_{ij}^2 C_{ij}^{(2)} > 0$  (Ocurre colisión atractiva; captura.)

$$t_{ij}^{(2)} = \frac{-b_{ij} - (b_{ij}^2 - v_{ij}^2 C_{ij}^{(2)})^{1/2}}{v_{ij}^2}.$$

(2)  $b_{ij}^2 - v_{ij}^2 C_{ij}^{(2)} < 0$  (No ocurre ninguna colisión.)

II  $b_{ij} > 0$  (Centros se alejan.)

(a)  $C_{ij}^{(2)} < 0$  ( $r_{ij}$  está dentro del rango atractivo. Ocurre colisión atractiva.)

$$t_{ij}^{(2)} = \frac{-b_{ij} + (b_{ij}^2 - v_{ij}^2 C_{ij}^{(2)})^{1/2}}{v_{ij}^2}.$$

(b)  $C_{ij}^{(2)} > 0$  ( $r_{ij}$  está fuera del rango atractivo. No ocurre ninguna colisión.)

Al calcular los tiempos de colisión, un tiempo máximo  $t_m$  es asignado. Este tiempo es demasiado corto como para permitir que cualquier par de partículas cambien su posición por más de  $(\frac{L}{2} - \sigma_2)$ , siendo  $L$  la longitud de la caja cuadrada o cúbica de simulación. Cualquier tiempo  $t_{ij}^{(\alpha)}$  mayor que el tiempo máximo es rechazado. Como suele suceder en sistemas poco densos, si no se encuentran tiempos  $t_{ij}^{(\alpha)}$  menores que  $t_m$ , todas las partículas son movidas por un tiempo igual a  $t_m$  y se realiza una nueva búsqueda de tiempos de colisión. Una vez encontrados todos los  $t_{ij}^{(\alpha)}$  menores que  $t_m$  se termina lo que se denomina un ciclo largo. Con todos los tiempos de colisión encontrados, se mueven todas las partículas por el  $t_{ij}^{(\alpha)}$  más corto encontrado. Ahora un par de partículas estarán separadas por una distancia  $\sigma_1$  o  $\sigma_2$ , por lo que, si se considera un sistema de partículas

de igual masa, este par de partículas experimentarán un cambio, igual y opuesto, en sus velocidades. El cambio de velocidades  $\Delta \mathbf{v}$ , dependerá del tipo de colisión como sigue:

(1) Ocurre colisión de partículas

$$\Delta \mathbf{v}_i = -\Delta \mathbf{v}_j = \frac{-\mathbf{r}_{ij} b_{ij}}{\sigma_1^2}$$

(2) Colisión atractiva

(a)  $C_{ij}^{(2)} > 0$                       Captura

$$\Delta \mathbf{v}_i = -\Delta \mathbf{v}_j = \frac{-\mathbf{r}_{ij}}{2\sigma_2^2} \left[ \left( \frac{4\sigma_2^2 U_0}{m} + b_{ij}^2 \right)^{1/2} + b_{ij} \right]$$

(b)  $C_{ij}^2 < 0$

(I)  $b_{ij}^2 > \frac{4\sigma_2^2 U_0}{m}$                       Disociación

$$\Delta \mathbf{v}_i = -\Delta \mathbf{v}_j = \frac{-\mathbf{r}_{ij}}{2\sigma_2^2} \left[ - \left( \frac{-4\sigma_2^2 U_0}{m} + b_{ij}^2 \right)^{1/2} + b_{ij} \right]$$

(II)  $b_{ij}^2 < \frac{4\sigma_2^2 U_0}{m}$                       Rebote

$$\Delta \mathbf{v}_i = -\Delta \mathbf{v}_j = \frac{-\mathbf{r}_{ij} b_{ij}}{\sigma_2^2},$$

donde el valor de  $C_{ij}^2$  debe ser calculado antes de que se mueva el par de partículas. Estos resultados son obtenidos de la conservación del momento y la energía para dos partículas con potencial de pozo cuadrado.

La diferencia entre el  $t_{ij}^{(\alpha)}$  más corto que acaba de transcurrir y el siguiente  $t_{ij}^{(\alpha)}$  más grande es el tiempo necesario para que se dé la siguiente colisión. Esto a menos que las partículas que acaban de colisionar se involucren inmediatamente en otra colisión. Debido a esto, una vez que se cambian las velocidades, se deben buscar colisiones en las que puedan estar involucradas solo el par de partículas que acaban de colisionar. A esto se le llama ciclo corto. Todos los  $t_{ij}^{(\alpha)}$  de la lista calculada en el ciclo largo deben disminuir por el  $t_{ij}^{(\alpha)}$  que acaba de transcurrir. También los  $t_{ij}^{(\alpha)}$  que involucren a las partículas que recién colisionaron deben ser eliminados de la lista. Los nuevos  $t_{ij}^{(\alpha)}$  recién calculados en



los ciclos cortos, deben ser agregados a la lista siempre y cuando sean menores al tiempo máximo establecido  $t_m$ .

En un ciclo largo el número de pares de partículas que tienen que ser examinados para encontrar posibles colisiones es de  $N(N - 1)/2$ , mientras que en un ciclo corto solo se deben examinar  $2N - 3$  pares. En los ciclos cortos hay un ahorro de tiempo de cómputo lo cual es de ayuda en especial para sistemas grandes. Hay otras formas de ahorrar tiempo de cómputo. La usada en esta DMPD se ve en la siguiente sección.

### 5.4.2. Calendario de eventos

Con el uso de potenciales escalón se requiere guardar una gran cantidad de información al calcular los tiempos de colisión de las partículas. Esto requiere un manejo adecuado de toda esa información para tener una simulación correcta. El ahorro de tiempo de cómputo también es indispensable para tener una simulación más eficiente. Una manera de ahorrar tiempo de cómputo en el cálculo de las colisiones es mediante el uso de celdas. Este método reduciría el número de pares de partículas en los que buscar por nuevas colisiones, después de un cambio de velocidades, a un valor menor e independiente de  $N$ .

Suponiendo que toda la región de simulación está dividida en celdas cuya longitud sobrepase el valor del diámetro de las partículas  $\sigma_1$ . Es claro que la búsqueda de colisiones de una partícula se dará sólo en su propia celda y en celdas adyacentes, lo cual representa un ahorro en los tiempos de cálculo. Lo único que se necesita hacer es tener un seguimiento de a que celda pertenece cada partícula. Esto se hace mediante listas que contienen a las partículas pertenecientes a cada celda. Todo esto implica la introducción de un nuevo evento el cual se dará cuando una partícula sale de una celda y entra a otra. Este evento señalará el momento exacto en el que una partícula cruzará de una celda a otra y por qué cara lo hará.

Otra forma de ahorrar tiempo de cómputo es con el uso de un tiempo local asociado a cada partícula. Cuando se de una colisión, sólo las partículas de celdas adyacentes son de interés y será sólo a ellas a quienes se mueva por el tiempo de la reciente colisión. De este modo, sólo en algunas ocasiones será necesaria una actualización de todo el sistema.

Conocer el tiempo de todos los eventos y saber cuál será el siguiente evento en ocurrir, ya sea una colisión o un cruce de celda, y que partículas están involucradas, implica la existencia de un calendario de eventos. Dicho calendario debe organizar todos los fu-

turos eventos y además debe ser fácilmente modificable, ya que al ocurrir una colisión, nuevas colisiones deben ser añadidas al calendario y otras deben ser borradas al quedar invalidadas.

El calendario que rige el sistema contiene una lista de futuras colisiones y cruces de celda, además de eventos correspondientes a la medida de diferentes cantidades en ciertos intervalos de tiempo. El calendario está regido por una estructuración de datos en forma de un árbol binario. Cada evento es un nodo del árbol. Se usan tres punteros para colocar a los eventos en el árbol, los cuales descienden hacia izquierda y derecha y otro que viene desde arriba. De este modo cada nodo apunta a otros dos nodos descendentes, uno a la izquierda y otro a la derecha. Los nodos están ordenados en el tiempo de tal modo que el evento de la izquierda está programado para ocurrir primero que el nodo actual del que desciende, mientras que el nodo de la derecha ocurrirá después del nodo actual del que desciende.

Para tener una eliminación rápida de eventos relacionados, cada nodo está ligado a dos listas circulares. Si se trata de una colisión se tendrá una lista para cada una de las partículas involucradas. Estas listas guardan los demás nodos colisión en los que aparecen estas partículas. Por otro lado, si se trata de un cruce de celda se tendrán dos listas distintas para la misma partícula.

De este modo, cada nodo tiene la información de que evento se trata y cuándo ocurrirá. Si es una colisión, se especifican las partículas involucradas. Si es un cruce de celda, se especifica la partícula y la cara de la celda que se va a cruzar. En la figura 5.2 se muestra una representación de un pequeño árbol de eventos. En la representación del árbol se omiten las listas circulares y en su lugar se muestran los cuatro punteros, a los lados del nodo, que se usan para fijar el nodo dentro de esas listas.

El manejo del árbol se lleva a cabo mediante cinco funciones. Una función se encarga de inicializar la estructura del árbol y crea las listas circulares. Otra función coloca a los eventos en el lugar correcto dentro del árbol y los liga dentro de las listas circulares. Otra función determina que evento será el próximo en ocurrir y quita del árbol a ese evento, así como todos aquellos eventos, ahora inválidos, en los que están involucradas estas mismas partículas. Las últimas dos funciones se encargan de remover los eventos del árbol y de reordenar los punteros después de la remoción de los nodos. Todos los detalles de estas funciones y del calendario de eventos con estructura de árbol se pueden encontrar en el

libro de Rapaport "The Art of Molecular Dynamics Simulation"<sup>80</sup> donde se presenta este método.

En resumen, la DMPD está construida para potenciales tipo escalón. Usa el método del cálculo de tiempos de colisión de las partículas e implementa un calendario de eventos. Este calendario ayuda a manejar de forma optimizada toda la información sobre los tipos de eventos, lo cual reduce los tiempos de cómputo de la simulación. Todo un programa de DMPD con estas características fue desarrollado por Chapela et al.<sup>84</sup> Este programa cuenta con una biblioteca de moléculas clasificadas por formas a las que se les puede decorar con sitios atractivos de potenciales de pozo cuadrado. Este programa está pensado para realizar investigaciones enfocadas al autoensamblaje y fue el que se usó para los sistemas con potenciales discontinuos.

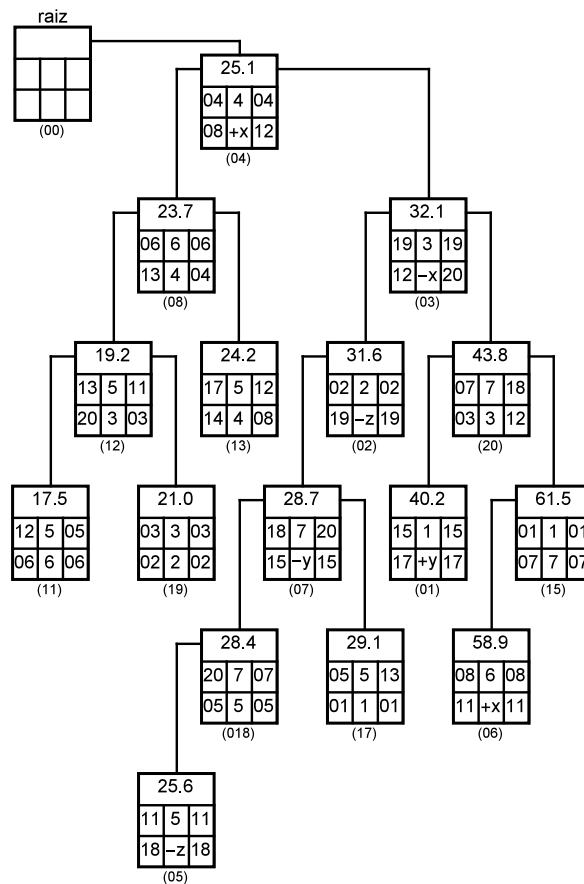


Figura 5.2: Esquema de un calendario de eventos organizado como un árbol binario.



# Capítulo 6

## Resultados

La presentación de los resultados se dividirá en dos partes. En la primera parte se presentarán los resultados obtenidos con el potencial discontinuo de pozo cuadrado, aquí se incluye la liberación del virión de la membrana y la red hexagonal en 2D y 3D. En la segunda parte se presentarán los resultados correspondientes al potencial continuo de Lennard-Jones, donde se incluye nuevamente la red hexagonal en 2D y 3D y las cápsides icosaédricas. Todos los resultados que se presentarán fueron obtenidos en el ensamble NVT.

### 6.1. Potencial discontinuo

#### 6.1.1. Liberación de un virión de una membrana de doble capa

Como primera parte se realizó el autoensamblaje de la membrana de doble capa que se utilizaría en el sistema Gag-Membrana para simular la formación y liberación de un virión de una membrana lipídica. A continuación se presentan estos resultados.

##### 6.1.1.1. Membrana de doble capa

Para ensamblar la membrana de doble capa se usó el modelo L1 (Fig. 2.1a). El ensamblaje se realizó en dos formas. Primero se construyó con los lípidos un arreglo muy cercano al de una doble capa. En la figura 6.1 a) se observa este arreglo, el cual consta de dos capas de lípidos, en un arreglo cuadrado, separadas una distancia igual al diámetro de las esferas del lípido y con una pequeña separación entre los lípidos en el plano y-z. Este arreglo se usó como configuración inicial en las corridas de simulación, las cuales

se dejaron correr varios bloques de colisiones. Hay que recordar que para los potenciales discontinuos se usa la DMPD donde la evolución del sistema se da por medio de colisiones entre las partículas. Para estos sistemas se corrieron bloques de colisiones, donde cada bloque consta de 25 millones de colisiones.

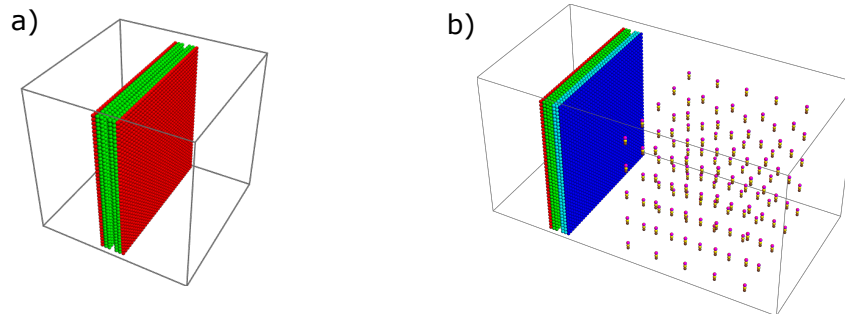


Figura 6.1: a) Configuración inicial para formar rápidamente una membrana de doble capa. Se colocan dos capas de lípidos separadas una pequeña distancia entre sí b) Configuración inicial del sistema Gag-Membrana, a la izquierda se coloca la membrana y a la derecha las moléculas Gag en un arreglo cuadrado. Se utiliza el arreglo cercano a la doble capa en los lípidos para que la simulación se centre en el ensamblaje del virión y no en el ensamblaje de la membrana.

Se hicieron varias corridas con distintos valores de densidad y temperatura para encontrar una buena red. Las densidades que se reportarán corresponden a densidades atómicas dadas por  $\rho = N_{at}/V$ , con  $N_{at}$  el número de átomos del que se componen las moléculas. También se usaron dos valores para el alcance del pozo con el que interaccionan atractivamente las colas de los lípidos, de  $\lambda = \sigma 1.5$  y  $\lambda = \sigma 2.5$ , con  $\sigma = 1$  el diámetro de las esferas del lípido. En la figura 6.2 se muestran los resultados principales de estas corridas. Se muestran cuatro isócoras, de  $\rho = 0.1$  y  $\rho = 0.15$  para  $\lambda = 1.5$  y, de nuevo, de  $\rho = 0.1$  y  $\rho = 0.15$  para  $\lambda = 2.5$ . Como se puede observar, para las isócoras de  $\rho = 0.1$ , donde la caja tiene una longitud de  $45\sigma$ , la doble capa que se mantiene en un mismo plano, mientras que para las de  $\rho = 0.15$ , al reducir la longitud de la caja a  $39.3\sigma$ , las membranas son capaces de doblarse para mantener la membrana de doble capa. Las filas no tienen la misma temperatura, pero sí indican cuatro configuraciones distintas con las que termina la membrana al ir subiendo la temperatura. En la primera fila de abajo hacia arriba, a la temperatura más baja, las corridas terminan en una doble capa rígida. Al subir de temperatura, en la segunda fila se observa una región donde las capas de la membrana tienen una reducción de su área debido a que se ordenan ligeramente en su plano, lo cual produce hoyos en la membrana. Esto solo se observa en las isócoras de  $\rho = 0.1$ . Para las isócoras más densas la membrana resultante, después de su disminución de área, seguirá siendo mayor al área del plano y-z de la caja de simulación, por lo tanto la membrana

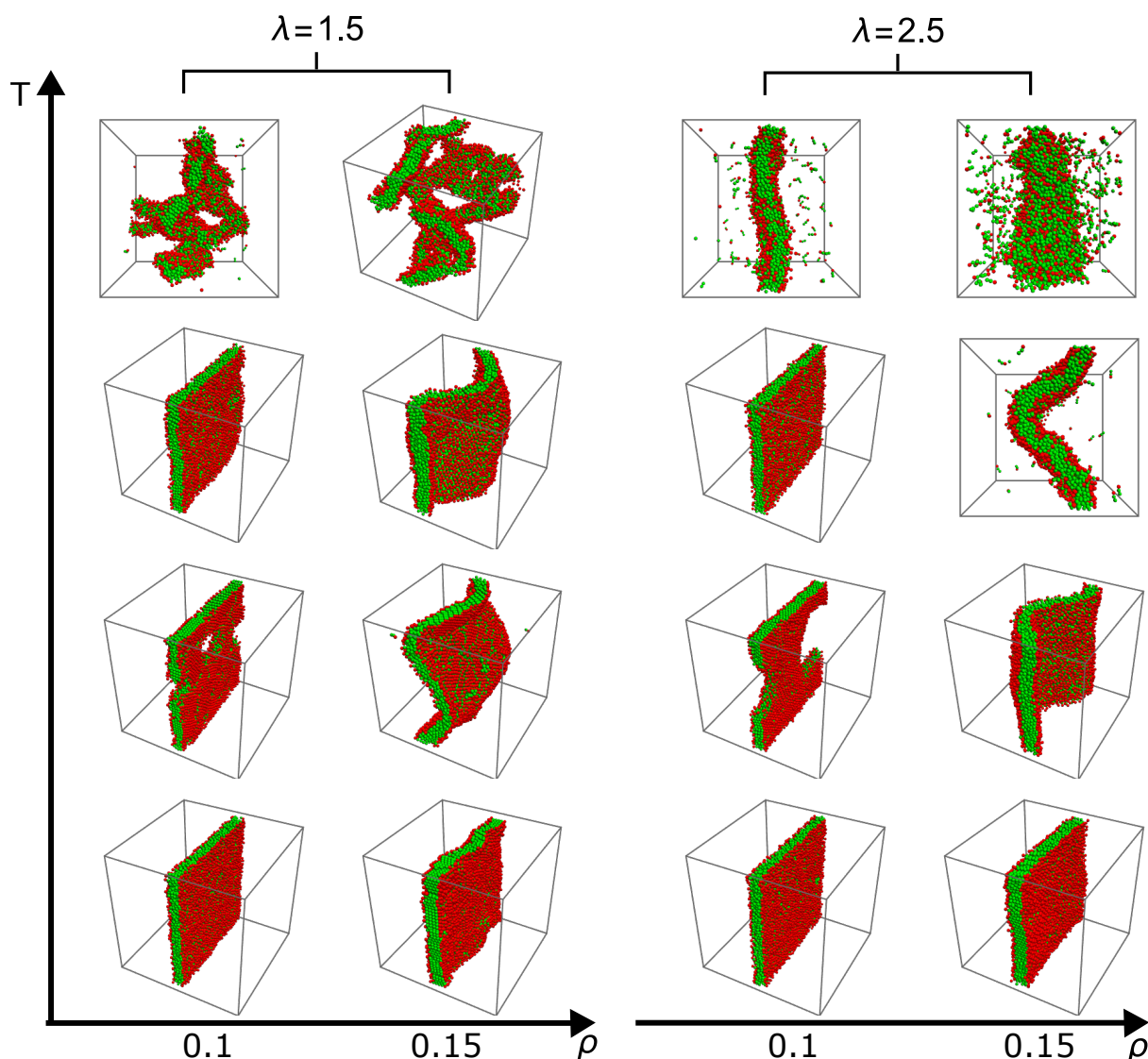


Figura 6.2: Membranas obtenidas con el modelo L1. Se presentan cuatro isócoras, dos para el alcance de  $\lambda = 1.5$  y dos para  $\lambda = 2.5$ . Las filas no necesariamente indican una misma temperatura, sino cuatro regiones distintas donde se encuentran cuatro configuraciones distintas de la membrana: membrana completa rígida, completa con hoyos, completa fluida y desintegración de la membrana. Las membranas están formadas por 3042 lípidos y fueron corridas 50 bloques de colisiones.

no presentará hoyos. En la tercer región al subir la temperatura, la membrana resultante es nuevamente una que ocupa la totalidad de la caja para  $\rho = 0.1$ . Esta membrana ya no es rígida sino fluida ya que los lípidos presentan una mayor movilidad dentro de la membrana. Para  $\rho = 0.15$ , nuevamente no se presentan cambios respecto a la región anterior y se sigue manteniendo una membrana con curvas. En la última región, en las temperaturas más altas, empieza a desintegrarse la membrana. La diferencia importante aquí es que para  $\lambda = 1.5$  las membranas, sin importar la densidad, se rompen. Mientras que para  $\lambda = 2.5$  las membranas se desintegran perdiendo lípidos poco a poco.

Con el modelo L2, el cual se compone de solo dos esferas (Fig 2.1b), se usó el mismo

método de ensamblaje de usar como configuración inicial una estructura cercana a la doble capa. En la figura 6.3 se muestran los principales resultados de este modelo. En este caso también se usan dos alcances del pozo de  $\lambda = \sigma 1.5$  y  $\lambda = \sigma 2.5$ , pero ahora solo se usa una densidad de  $\rho = 0.07$  para las isócoras. A esta densidad la membrana se mantiene en un solo plano sin doblarse. Al igual que con el modelo L1 se observan cuatro regiones dentro de la isócora: membrana rígida completa, membrana con huecos, membrana completa fluida y por último su desintegración.

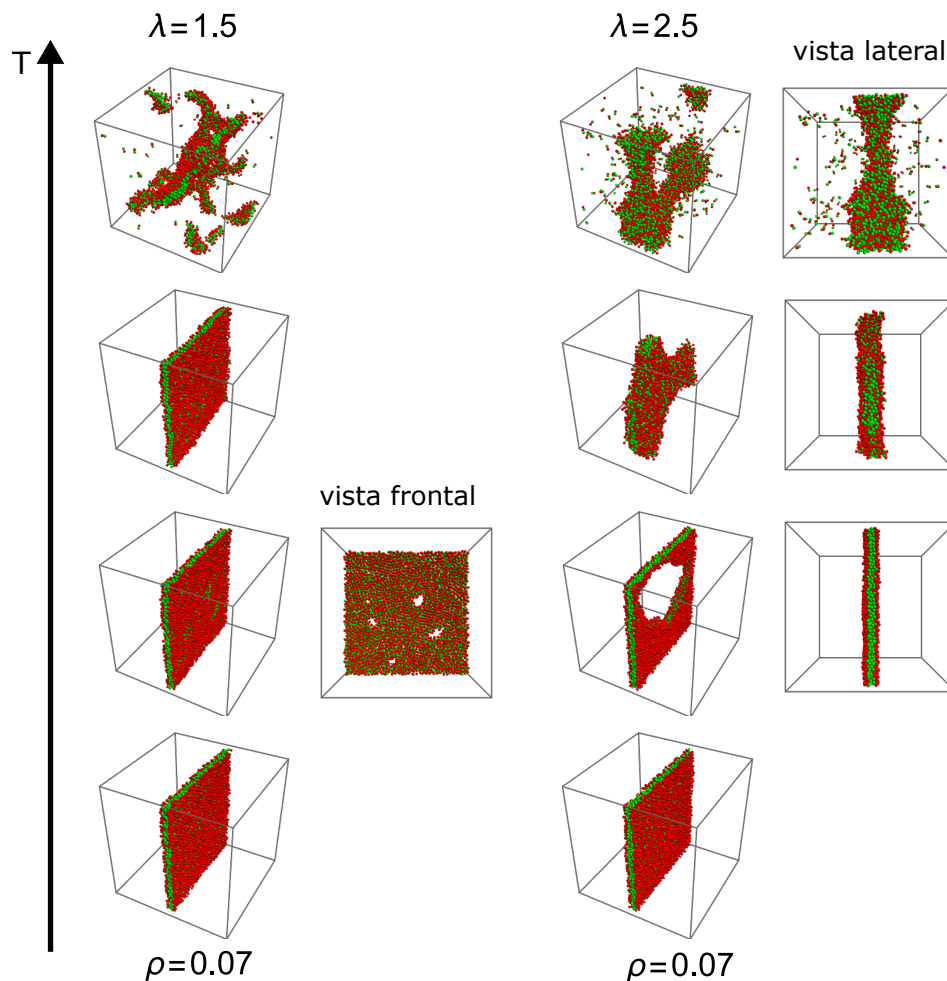


Figura 6.3: Membranas obtenidas con el modelo L2 compuesto de solo dos esferas. Se muestran dos isócoras, para  $\lambda = 1.5$  y  $\lambda = 2.5$ , a una misma densidad. Las membranas se componen de 3042 lípidos y fueron corridas por 50 bloques de colisiones.

En este modelo los distintos alcances del pozo cuadrado introducen una nueva diferencia en las membranas. Para  $\lambda = 1.5$  el comportamiento es igual que en el modelo L1. En cambio, para  $\lambda = 2.5$  en la región de membrana fluida la membrana se ensancha. Esto se ve claramente en la vista lateral de la membrana en la tercera fila, de abajo hacia arriba, de la figura 6.3. Las colas cortas de los lípidos producen una membrana delgada, pero el



largo alcance del potencial permite que las dos capas puedan tener una separación mas amplia entre ellas que la que tendrían con un alcance corto. Este espacio entre las dos capas no se mantiene libre, sino que es ocupado por otros lípidos. Esto provoca el ensanchamiento de la membrana, lo que a su vez produce una reducción de su área dejando de cubrir la totalidad del plano y-z.

La segunda manera de obtener las membranas fue utilizando uno arreglo cuadrado como configuración inicial como se ve en el extremo izquierdo de la figura 6.4. De esta forma se asegura que los modelos L1 y L2 son capaces de autoensamblar una membrana de doble capa. En la figura 6.4 se muestran las configuraciones de la membrana en el proceso de ensamblaje hasta la formación de la doble capa. En la parte superior se observa la evolución del ensamblaje del modelo L1 y en la parte inferior la del modelo L2. Estas corridas se llevaron acabo con 2916 lípidos a una densidad y temperatura de  $\rho = 0.11$  y  $T = 2.3$  para L1 y a  $\rho = 0.07$  y  $T = 0.3$  para L2. Al modelo L1 le tomó 30 bloques de colisiones ensamblar la membrana, lo que corresponde a 42.1 horas de tiempo de simulación, mientras que a L2 le tomo un total de 60 bloques, lo cual corresponde a 12.6 horas. Al modelo L2 le toma mayor número de colisiones ensamblar la membrana debido a su alcance de las colas de  $\lambda = 1.5$  menor que el del modelo L1 de  $\lambda = 2.5$ . Sin embargo, su tiempo de simulación es menor debido a la partícula que tiene de menos esta molécula.

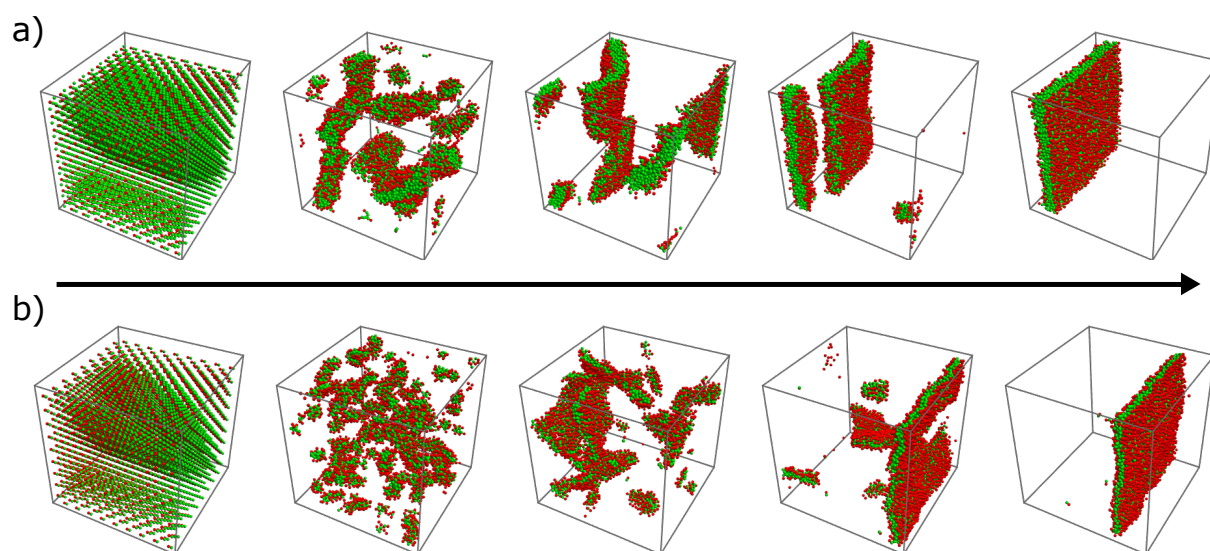


Figura 6.4: Evolución del autoensamblaje de la membrana de doble capa partiendo de un a configuración FCC de los lípidos. a) Modelo L1 compuesto por tres esferas tangentes, una para la cabeza y dos para la cola del lípido. b) Modelo L2 compuesto de solo dos esferas tangentes.

### 6.1.1.2. Sistema Gag-Membrana

Para simular la formación y liberación de un virión se usaron las membranas obtenidas en la sección anterior y se les agregaron las moléculas Gag modeladas con el modelo G1. En la figura 6.1 b) se muestra la configuración con que se iniciaron las corridas de este sistema. En una caja rectangular, del lado izquierdo se inicia con los lípidos en el arreglo cercano a la doble capa y del lado derecho se colocan las moléculas Gag en un arreglo cuadrado.

En la figura 6.5 se muestra un ejemplo de como actúa este sistema. Debido a la interacción atractiva entre el dominio MA del Gag y las cabezas de los lípidos, las moléculas Gag se pegan a la membrana. Cuando las Gag se reúnen en pequeños grupos, estos empiezan a doblar ligeramente la membrana formando una pequeña protuberancia en el extremo opuesto de la membrana, como se ve en la figura 6.5 b). En esta misma figura se puede ver como las moléculas Gag, por las condiciones periódicas de frontera, se pegan por ambos lados de la membrana. Para evitar esto de una manera sencilla, cada capa de la membrana se formó con una especie diferente de lípido, como se ve en la figura 6.4 b), donde la cabeza de una especie de lípido se pinta en azul y la cabeza de la otra especie se pinta en rojo. Entonces, solo se permite la interacción atractiva de las moléculas de Gag con la capa que queda de su lado.

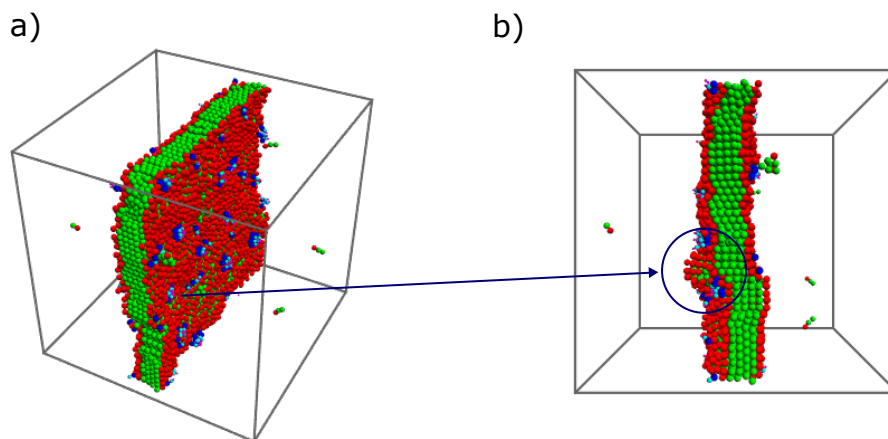


Figura 6.5: Sistema Gag-Membrana. Cuando las moléculas Gag se pegan a la membrana se agrupan entre ellos y pueden producir protuberancias en el extremo opuesto de la membrana como se observa en b).

A causa del ancho de la membrana que se produce con el modelo L1, los Gag, que son de menor longitud que un lípido, no pueden doblar por si solos la membrana para recubrirse con ella. Debido a esto la formación de un virión con el modelo de lípido L1 y el modelo de Gag G1 no fue posible.

Sin embargo con el modelo de lípido L2 que forma una membrana más delgada los Gag G1 fueron capaces de ensamblarse en la membrana, envolverse con ella y finalmente liberarse del resto de la membrana. En la figura 6.6 se muestra la evolución de la formación del virión y su liberación. Esta corrida se llevó a cabo con 3042 lípidos y 125 moléculas de Gag a una densidad de  $\rho = 0.074$ , de tal forma que la membrana recubriera por completo el plano y-z de la caja sin doblarse por si misma. La temperatura usada fue de  $T = 0.28$  y se requirieron 260 bloques de colisiones para completar la liberación del virión, lo cual corresponde a 81.6 horas de tiempo de simulación. En la figura 6.6 h) la membrana se muestra con cierta opacidad para observar las moléculas Gag que forman el virión. Se ensamblan 4 viriones compuestos por 47, 43, 19 y 15 moléculas Gag. El mas grande, que se pinta de distinto color, es el que se libera de la membrana, rodeado por ella. El segundo más grande permanece unido a la membrana al igual que los dos más pequeños.

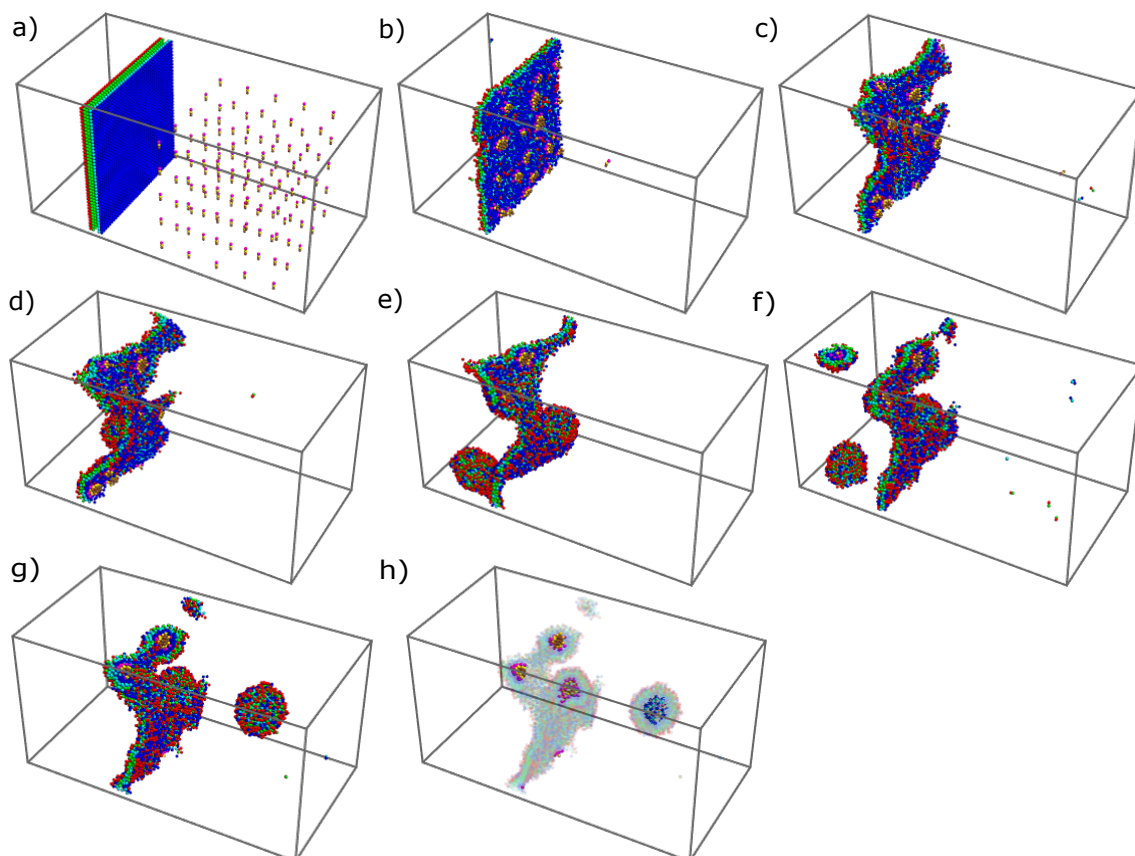


Figura 6.6: Evolución del autoensamblaje, a lo largo de 260 bloques de colisiones, de la formación y liberación de un virión de una membrana de doble capa. a) Configuración inicial del sistema Gag-Membrana. b) Después de 20 bloques de colisiones casi la totalidad de los Gags se encuentran pegados a la membrana. c) Conforme los Gags se unen en grupos más grandes, estos producen pequeñas protuberancias en toda la red lo que provoca que deje de ocupar la totalidad del plano de la caja. (d-f) Los Gags terminan por reunirse en 4 grandes grupos, de los cuales el mayor forma una gran protuberancia hasta finalmente separarse de la membrana. (g-h) Muestran otra vista del virión liberado con la membrana con cierta opacidad para observar los Gags que forman viriones compuestos de 47, 43, 19 y 15 Gags. El virión de 47 Gags se pinta de distinto color.

La corrida mostrada en la figura 6.6 corresponde a la temperatura  $T = 0.28$  de una isócora de  $\rho = 0.074$ . En esta isócora la formación de viriones se encontró en un rango de temperaturas de  $T = 0.26$  a  $T = 0.32$ . Este rango de temperaturas corresponde a la región donde se forma una membrana completa (que ocupa todo el plano) fluida, donde los lípidos tienen una buena movilidad dentro de la membrana, lo que le facilita curvarse. Por debajo de estas temperaturas se presenta las otras dos regiones de las membranas: rígida completa y membrana con hoyos. En estos casos los Gag se agrupaban en la membrana, pero son incapaces de doblarla como se ve en la figura 6.7a). Por arriba de este rango de temperaturas se siguen formando viriones, pero en este caso los Gag no se ensamblan en la membrana, ya que a estas temperaturas más altas la membrana se rompe (Fig. 6.3). Lo que sucede es que los Gag se agrupan y se comienzan a rodear con los lípidos que se encuentran libres, es decir, sin formar una membrana, hasta ensamblar un virión. En la figura 6.7 c) se muestra la configuración final de la última temperatura  $T = 0.39$  de la isócora donde se observan 4 viriones compuestos por 35, 34, 33 y 19 moléculas Gag.

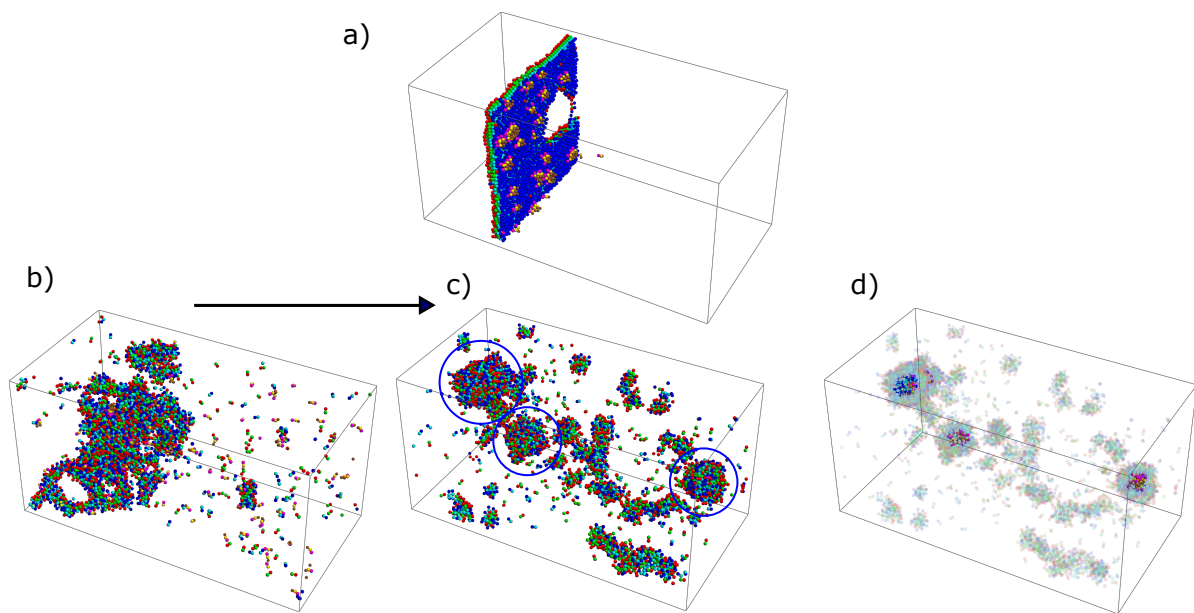


Figura 6.7: Configuraciones del sistema Gag-Membrana fuera de la región de temperaturas donde se forman correctamente los viriones. a) Membrana rígida con hoyos a  $T = 0.2$  que no permite su deformación por parte de los Gags. (b-c) Membrana a  $T = 0.39$  donde a pesar de que la membrana se rompe después del primer bloque de colisiones (b) los Gags se agrupan y se rodean con los lípidos para formar viriones (c). d) Misma configuración que (c), pero con opacidad en los lípidos. Se observan grupos de Gags de 35, 34, 33 y 19 Gags.

## 6.1.2. Red hexagonal del VIH-1

### 6.1.2.1. Red 2D

Para formar la red hexagonal se usó la molécula M1 (4.3 a) con los parches atractivos correspondientes a la interacción 3-Fold e IH, sin incluir el 2-Fold. Se corrieron dos isócoras a  $\rho = 2$  dentro del mismo rango de temperaturas de  $T = 0.07$  a  $T = 0.21$  con un incremento de  $\Delta T = 0.02$ , con un posición de los parches desde el centro del disco duro de  $L = 0.4$  y  $L = 0.45$ , respectivamente. Las corridas consistieron de 200 moléculas, se iniciaron en un arreglo cúbico centrado en la cara (FCC) y se dejaron correr por 50 bloques de colisiones, lo cual corresponde a 5.7 horas de tiempo de simulación. Para la primera isócora la red se ensambló en una sola temperatura de  $T = 0.13$ . En cambio, para la segunda isócora la red se encontró en un rango de temperaturas más amplio de  $T = 0.13$  a  $T = 0.17$ . De estas tres temperaturas, donde se encontró una mejor formación de la red, continua y sin defectos, es en la temperatura más alta. En la figura 6.8 se muestra la red resultante a esta temperatura. Mientras que en las dos temperaturas menores la red que se ensambló presenta defectos y no es continua. En la temperatura siguiente de  $T=0.19$  la red desaparece por completo y solo se encuentran dos hexágonos separados al final de la corrida. A esta temperatura las parejas entre los parches 2F ya no son lo suficientemente estables como para mantener formados a los hexágonos. Por su parte, las ternas formadas por los parches 3F, que se encargan de unir los hexágonos entre si, son más estables y se mantienen sin dificultad a esta temperatura, sin embargo una red ya no pueden crecer debido a la falta de hexágonos.

En el panel derecho de la figura 6.8 se muestra la energía potencial ya equilibrada de las distintas temperaturas de la segunda isócora. Las temperaturas donde se forma la red concuerdan con los valores más bajos de la energía potencial. Además un rasgo interesante es que para tener una buena formación de red, que sea continua sin defectos, se debe estar a una temperatura cercana al punto donde la red desaparece. Este hecho se ve claramente en el salto de la energía potencial de  $T = 0.17$  (red hexagonal) a  $T = 0.19$  (fluido de ternas).

Se llevaron a cabo tres variaciones de la posición de los parches de la molécula con la intención de encontrar algún otro tipo de estructura que pudiera formar la red. Sin embargo, lo que se encontró fue la gran capacidad que tiene esta red de ser ensamblada con distintas variaciones de la molécula ideal. A continuación se describe este hecho.

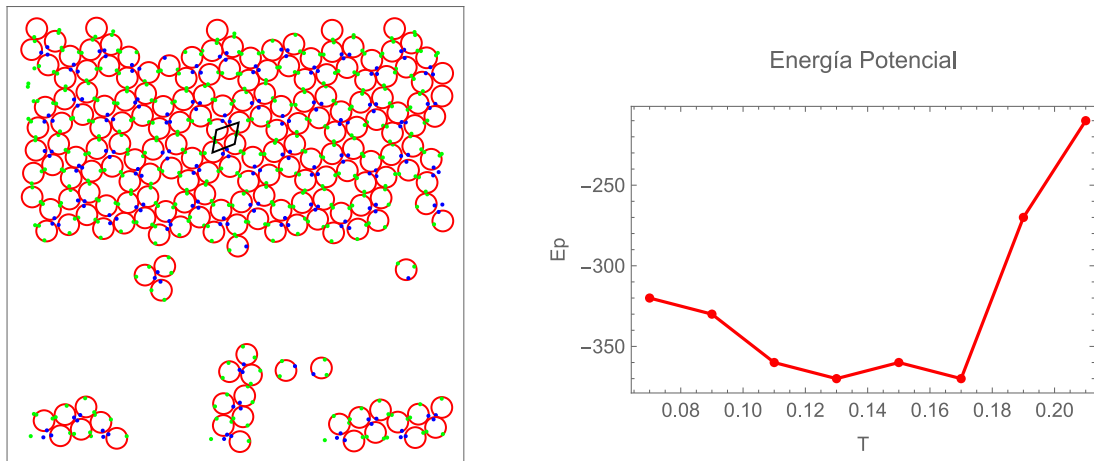


Figura 6.8: Red hexagonal del VIH-1 autoensamblada con el modelo M1, sin incluir el parche 2F. La red corresponde a la temperatura  $T = 0.17$  de una isócora a  $\rho = 2$ . En la parte derecha se muestra la energía potencial ya equilibrada de las distintas temperaturas de la isócora.

En todas las variaciones de la molécula original lo que se mueve son los ángulos que le dan posición a los parches alrededor del disco duro. Los demás parámetros de la molécula permanecen sin cambios y se usa  $L = 0.45$  para la posición radial de los parches. Todas las corridas fueron hechas con los mismos parámetros de densidad, temperatura, número de moléculas y número de colisiones que de las corridas anteriores con la molécula original. La primera variación de la molécula se realizó en el primero y segundo parches, correspondientes a la interacción IH1 y IH2, que están en  $0^\circ$  y  $120^\circ$  respectivamente. Se cambiaron sus posiciones a  $-5^\circ$  y  $125^\circ$ ,  $-10^\circ$  y  $130^\circ$  y a  $-15^\circ$  y  $135^\circ$ . Así, en estas tres nuevas moléculas, las posiciones de los parches que están encargados de formar el hexágono que tenían una separación de  $120^\circ$  grados pasan a tener separaciones de  $130^\circ$ ,  $140^\circ$  y  $150^\circ$ . El tercer parche, que corresponde a la interacción 3F, permanece a  $270^\circ$  en los tres casos, como en la molécula original. Lo que se encontró con estas tres moléculas fue que todas ensamblaron la red hexagonal en un rango de temperaturas de  $T = 0.14 - 0.18$ , para las primeras dos, y en un rango de  $T = 0.12 - 0.16$ , para la tercera.

El segundo tipo de variación de ángulos se realizó solo en el segundo parche IH2 situado a  $120^\circ$  en la molécula original. Su posición se varió desde  $80^\circ$  hasta  $180^\circ$  en incrementos de  $10^\circ$ . Se pasó específicamente por  $108^\circ$ , ya que este ángulo corresponde al de los ángulos interiores de un pentágono, por lo que este ángulo debería poder formar pentágonos en la red. En esta ocasión se dejaron fijos los parches IH1 y 3F situados en  $0^\circ$  y  $270^\circ$ . De estas diez variaciones de ángulo solo en  $100^\circ$ ,  $170^\circ$  y  $180^\circ$ , no se formó red alguna. De las 7 variaciones restantes en 5 se formó la misma red de hexágonos que forma el VIH-1,



incluso en la variación de  $108^\circ$  donde se esperaba la formación de pentágonos. En la figura 6.9 se muestra la red con la variación a  $108^\circ$  así como las demás variaciones a  $130^\circ$ ,  $140^\circ$ ,  $150^\circ$  y  $160^\circ$  donde se forma la red hexagonal.

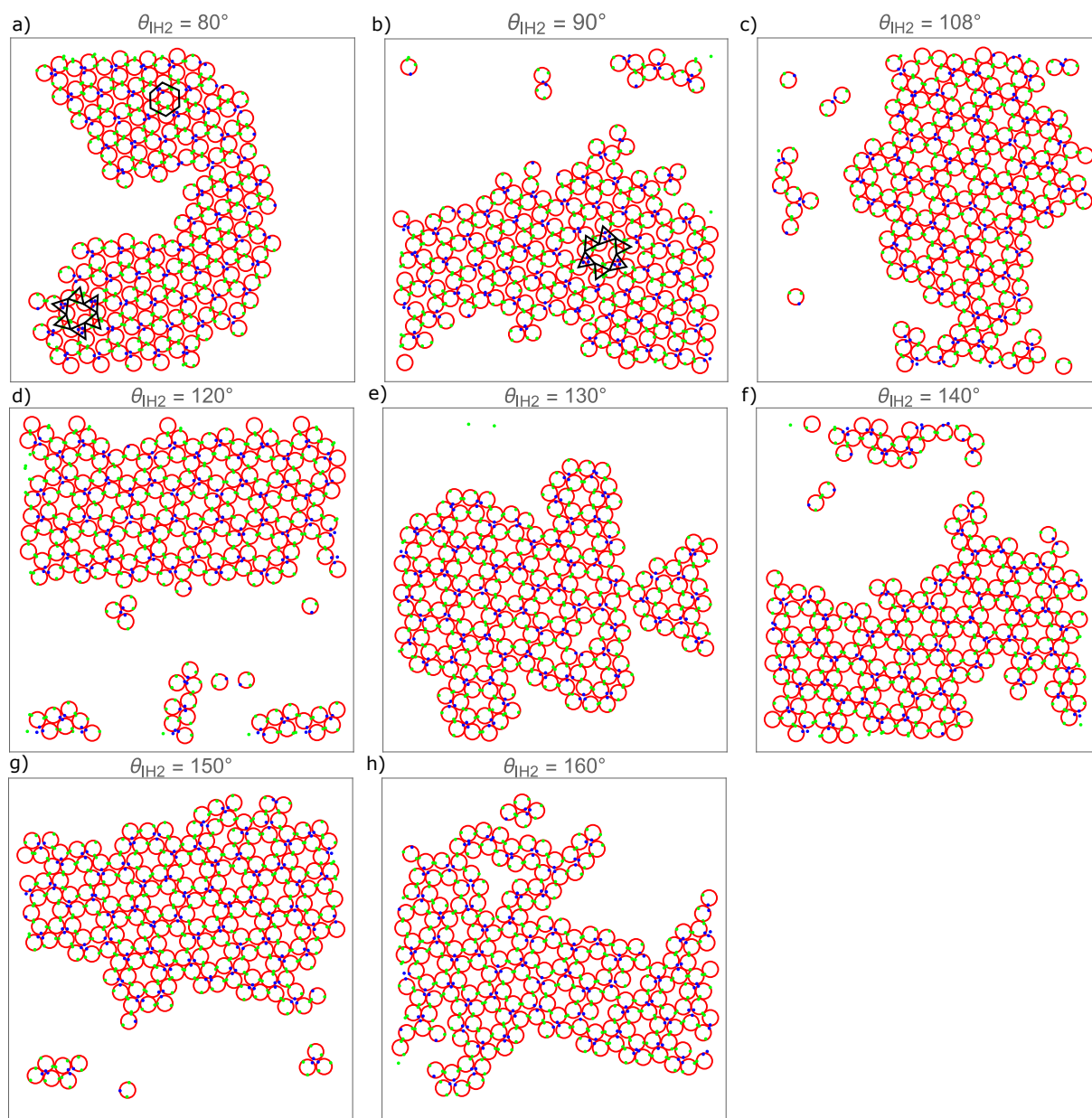


Figura 6.9: Redes obtenidas al variar la posición del parche IH2 que originalmente se encuentra a  $120^\circ$ , mientras que los parches IH1 y 3F permanecen en su misma posición.

Las dos variaciones restantes, donde se formó red, corresponden a los ángulos  $80^\circ$  y  $90^\circ$ . Estas redes difieren de las anteriores en que en estos casos no se forman parejas entre las moléculas, sino que tanto los parches terna como los parches pareja forman ternas de moléculas. Esto se debe a que cuando una molécula atrae a otra por cada uno de sus parches pareja, estas dos moléculas se unen a su vez entre sí por el parche pareja que les queda. De esta manera se forma una terna con 6 parches parejas. Esta característica de

formar solo ternas provoca que la variación de  $80^\circ$  forme una red donde se pueden ver dos regiones, una compacta promovida por los parches pareja y su alcance de  $\lambda = 0.2$ , y una abierta promovida por los alcances de  $\lambda = 0.4$  de los parches terna. En la figura 6.9 a) se indica la región compacta que se acerca a una red hexagonal y la región abierta que tiende a formar la red Kagome Deformada. La red Kagome Deformada es una variación de la red Kagome, caracterizada por tener dos parches que forman parejas separados  $150^\circ$  el uno del otro en lugar de  $180^\circ$  como la Kagome. En la variación de  $90^\circ$  también se ven dos regiones debido a los alcances distintos de los parches pareja y terna, pero la red claramente corresponde a la red Kagome Deformada en sus dos regiones, como se ve en la figura 6.9 b).

El tercer tipo de variación de ángulos se realizó en el parche 3F situado a  $270^\circ$ , moviéndolo desde  $180^\circ$  hasta  $270^\circ$ . Se corrieron diez ángulos distintos empezando en  $180^\circ$ ,  $200^\circ$  y  $240^\circ$ , y a partir de ahí se hicieron incrementos de  $5^\circ$  hasta llegar a los  $270^\circ$ . Al mantenerse fijos los parches IH1 e IH2 no se afectaría la formación de los hexágonos, sino más bien la unión entre ellos. El resultado fue que solo cuando el tercer parche se encuentra en  $180^\circ$  no se forma red alguna. Para todos los demás ángulos, desde  $200^\circ$  a  $270^\circ$ , siempre se ensambla una red hexagonal bien definida, continua y sin defectos. La diferencia entre estas redes es la manera en la que se unen los hexágonos. En los casos donde el parche está en  $200$  y  $270$  grados, entre cuatro moléculas de dos hexágonos adyacentes se forma un rombo como se indica en la figura 6.10 a) y en 6.8 a). Cuando se aumenta el ángulo, desde los  $200^\circ$ , o se disminuye, desde los  $270^\circ$ , esas mismas cuatro moléculas se abren poco a poco hasta formar un cuadrado cuando llegan a los  $245^\circ$ , como se indica en la figura 6.10b).

### 6.1.2.2. Red 3D

Inicialmente se usó una cadena de tres esferas, similar al modelo del Gag en 3D, pero con las esferas del mismo diámetro, cada una con parches en su ecuador en la misma disposición que en 2D. Con esta molécula en una corrida de 500 moléculas, en un arreglo FCC como configuración inicial, a una densidad atómica de  $\rho = 1$  la mayor formación de hexágonos que se encontró fue de 17, pero sin formar una sola red, sino fragmentos de 9, 3, 2 y 1 hexágonos. Esto se debe a que los fragmentos de red formados deben rotar y orientarse de manera correcta para unirse, pero estos movimientos se ven obstruidos por



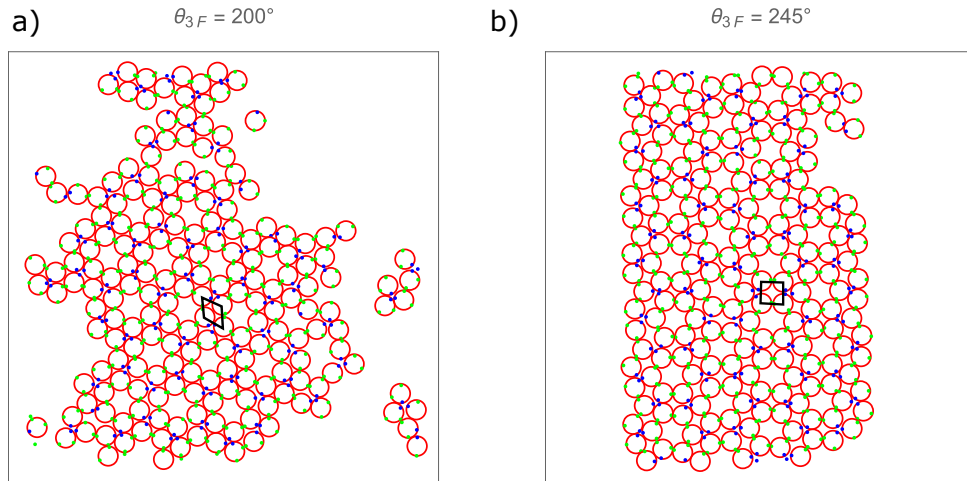


Figura 6.10: Redes obtenidas al variar la posición del parche 3F que originalmente se encuentra a  $270^\circ$ , mientras que los parches IH1 e IH2 permanecen en su misma posición.

el mismo largo de la molécula. Para densidades más bajas, la rotación no se obstruye, pero poder orientarse y unirse correctamente le toma mucho tiempo a los fragmentos de red. Esto provoca que se necesite una corrida con un tiempo de simulación (número de colisiones) mayor. En la figura 6.11 a) se muestra esta corrida con las moléculas que forman los hexágonos pintadas en azul claro. En el panel izquierdo de la figura se ve cómo las moléculas se alinean de buena forma gracias a los parches en el ecuador y en el panel derecho se muestra el conteo de los hexágonos conforme avanza la simulación.

En el virus del VIH-1 aunque el Gag tiene una longitud considerable debido a sus 6 dominios, no presenta estos problemas. Esto se debe a que el ensamblaje hexagonal se da una vez que las proteínas Gag se encuentren pegados a la parte interna de la membrana de la célula infectada, la cual le sirve como andamiaje.<sup>31</sup> En este momento las proteínas Gag se mantienen alineados entre si y solo se trasladan en el plano que les impone la membrana lipídica.

Para mejorar la red se creó una molécula donde las tres esferas estuvieran traslapadas un radio de distancia entre ellas, manteniendo los parches en el ecuador de cada esfera. Con esta molécula en una simulación de 500 moléculas a una densidad atómica de  $\rho = 0.45$  se obtuvo una formación de 36 hexágonos. Esta molécula traslapada muestra una gran mejora a la versión tangente, lo que permitió ensamblar una red de tamaño considerable, aunque aún sin ser continua debido a la formación de algunos defectos durante su ensamblaje. En la figura 6.11 b) se muestra el resultado de esta corrida junto con el conteo de hexágonos a lo largo de la corrida.

El mejoramiento de la red con la molécula traslapada llevó a la formación de la molécu-

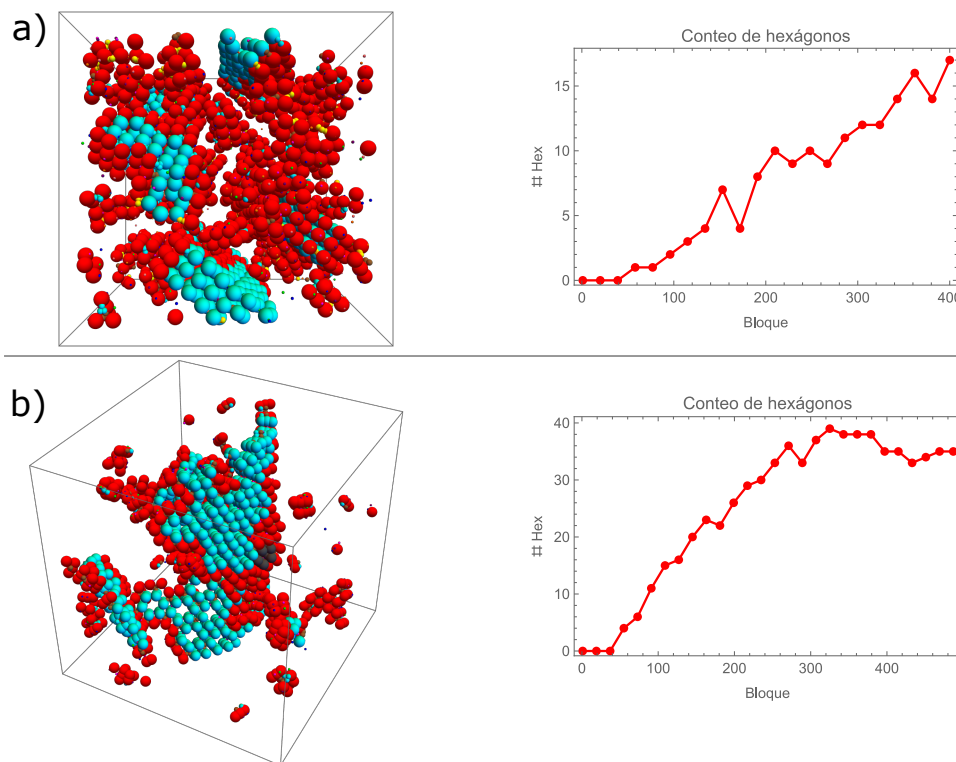


Figura 6.11: Hexágonos formados con partículas de 3 esferas con parches atractivos en sus ecuadores. a) Se usa una molécula compuesta por tres esferas tangentes. b) Se usa una molécula compuesta por tres esferas traslapadas, reduciendo así su longitud. En la parte derecha se grafica el conteo de la formación de hexágonos a lo largo de la corrida. Cada bloque corrido en la simulación consta de 25 millones de colisiones entre las partículas.

la M2 (Fig. 4.3 b), ahora solo con dos esferas traslapadas, igualmente por un radio de distancia. Con esta molécula en una corrida de 500 moléculas con una densidad de  $\rho = 0.46$ , se obtuvo una red continua formada por 57 hexágonos como se muestra en la figura 6.12 a). En el panel derecho de la figura se muestra el conteo de los hexágonos formados conforme avanza la simulación. Este conteo de hexágonos muestra como el modelo de dos esferas traslapadas no solo es el que ensambla una red continua sino que lo hace más rápidamente en comparación con los otros dos modelos. Después de 100 bloques de colisiones con el modelo M2 ya había cerca de 30 hexágonos formados, mientras que para los modelos anteriores no se alcanzaron a formar más de 15 hexágonos. Respecto al tiempo de simulación, este modelo aparentemente es más rápido que el modelo de esferas tangentes, debido a su disminución de átomos por molécula. Al modelo traslapado M2 le toma 8 horas completar los primeros 50 bloques de colisiones, mientras que al modelo de esferas tangentes le tomó 12 horas. Al modelo M2 le toma alrededor de 400 bloques llegar a su máximo número de hexágonos, lo cual le toma un tiempo de 97 horas. Sin embargo, este tiempo no difiere mucho de las 103 horas que le toma al modelo tangente completar 400

bloques. Esto se debe a que como el modelo M2 forma rápidamente un gran número de hexágonos, el cálculo de las colisiones entre las partículas se hace más lento debido a la cercanía entre ellas. Hecho que no sucede con el modelo tangente por el bajo número de hexágonos.

Es notable como esta red presenta una curvatura en algunas regiones, esto es debido a que el tamaño de la red ensamblada no es conmensurable con la caja de simulación. Este efecto se puede disminuir haciendo una corrida con una densidad más baja o se puede acentuar si se aumenta la densidad. En la figura 6.12 b) se muestra el caso de subir la densidad, donde se observa un aumento en la curvatura que puede alcanzar la red. De hecho, debido al tamaño de la caja, la red fue capaz de inducir defectos, en forma de dos heptágonos y un pentágono, que le ayudaron a aumentar su curvatura y seguir creciendo en el reducido espacio.

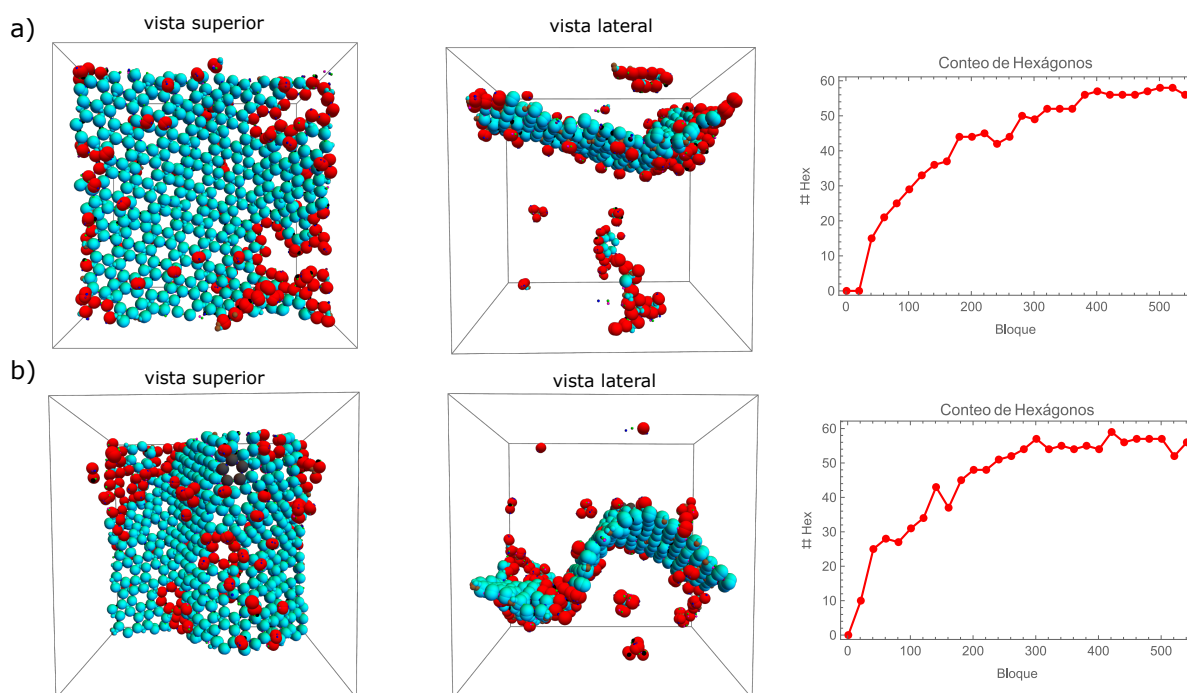


Figura 6.12: Red hexagonal del VIH-1 en 3D autoensamblada con el modelo M2, compuesta por dos esferas traslapadas, que no incluye la interacción 2F. a) Red compuesta por 57 hexágonos a una densidad  $\rho = 0.46$  y a  $T = 0.17$ . b) Red compuesta por 58 hexágonos a una densidad  $\rho = 0.61$  y a  $T = 0.17$ . En el extremo derecho se muestra la gráfica con el conteo de hexágonos a lo largo de la corrida, donde cada bloque consta de 25 millones de colisiones entre las partículas.

Como se mencionó en la sección de modelos moleculares, la molécula traslapada M2 representa los subdominios CA-NTD y CA-CTD del Gag del VIH-1. Aunque en el Gag la interacción 3F se encuentre solo en CA-NTD y las interacciones IH y 2F solo estén en CA-CTD, que en el modelo M2 los tres parches que representan estas interacciones se

coloquen en ambas esferas es de gran importancia. Esto se debe a que en las simulaciones las moléculas interactúan libremente sin la intervención de andamiajes que las orienten como es el caso del Gag.<sup>31</sup> En primer lugar, las esferas traslapadas por si mismas ayudan a dar una orientación al crecimiento de la red sobre un mismo plano. Hecho que se complica si se intenta usar únicamente una esfera para ensamblar la mayoría de las redes que ya se han ensamblado en 2D. En segundo lugar, colocar los tres parches en ambas esferas ayuda a tener una única orientación en el crecimiento de la membrana, como se ejemplifica en el panel c) de la figura 6.13.

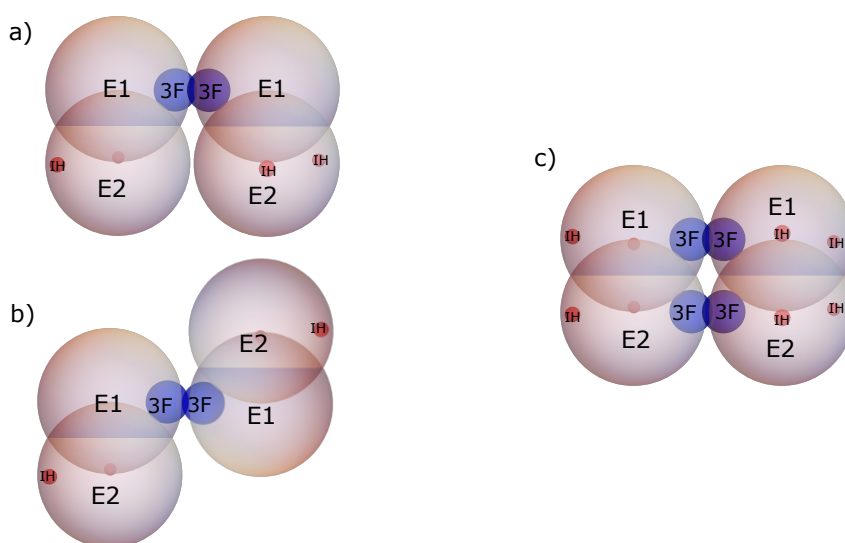


Figura 6.13: Planos en los que puede crecer la red dependiendo de si los parches se ponen solo en una de las esferas (a-b) o si se ponen en ambas esferas (c).

El código desarrollado para correr la dinámica molecular de eventos es muy versátil. Se pueden generar distintas moléculas y preparar corridas con distintos parámetros, lo que lo convierte en una buena herramienta para desarrollar un modelo molecular y aplicarlo a algún problema de interés. Sin embargo, los modelos en 3D usados aquí tienen un tiempo de simulación muy grande debido al número de partículas que forman la molécula, ya que cada parche atractivo es una partícula más de la molécula. De este modo, las simulaciones de 500 moléculas presentadas en la figura 6.12, son simulación de 4000 partículas a las que aplanar su energía potencial le toma 4 días a la densidad más baja y 3.25 días a la densidad más alta. Debido a estos tiempos altos de simulación los siguientes resultados son obtenidos con el paquete de simulación HOOMD-Blue<sup>85,86</sup> para correr en GPUs y acelerar las simulaciones.

## 6.2. Potencial continuo

Al usar el paquete de simulación HOOMD<sup>85,86</sup> además de correr en GPUs permite usar potenciales continuos y modelar las moléculas como rígidas o vibrantes (con ligaduras como hasta ahora). Al tener estas dos formas de modelar la molécula es necesario saber cuál de las dos tiene un mejor desempeño en el autoensamblaje de la red y en los tiempos de simulación.

### 6.2.1. Rígido vs vibrante

Para hacer la comparación en 2D se usa la molécula M1, pero se sustituyen los alcances del potencial de PC de  $\lambda = 0.2$  y  $\lambda = 0.4$  por los diámetros  $\sigma = 0.15$  y  $\sigma = 0.45$  de los potenciales de LJ de los parches que forman parejas y ternas, respectivamente. La posición radial  $L$  de los parches también se modifica para mantenerlos tangentes por dentro al disco repulsivo de diámetro  $\sigma = 1$  que los contiene. Se realizaron corridas de 50 millones de pasos en el tiempo (PT), con un paso de  $dt = 0.001$ , para sistemas de 196 moléculas a una densidad atómica de  $\rho = 2$ , para ambos casos, rígido y vibrante.

Para el caso rígido se encontró la formación de la red continua y sin defectos en un rango de temperaturas de  $T=0.68$  a  $T=0.78$ , con  $T=0.68$  la temperatura donde se encontró la mejor red formada por 23 hexágonos. Para el caso vibrante la red se ensambló en un rango de  $T=0.64$  a  $T=0.7$ , con  $T=0.66$  la temperatura donde se encontró la mejor red, formada por 22 hexágonos. En la figura 6.14 se muestran las redes para el caso rígido y vibrante en las temperaturas de mejor formación mencionadas. La figura 6.14 también muestra la evolución de la energía potencial y el conteo de los hexágonos formados a lo largo de la corrida para ambos casos. De la figura 6.14 d) se puede observar que al caso vibrante le toma de 2 a 4 millones de pasos más llegar al mismo número de hexágonos que al caso rígido, esto solo dentro de los primeros 10 millones de pasos, lo cual se refleja en un ritmo de crecimiento similar al caso rígido, solo que ligeramente tardío. También se observa como, para los 20 millones de pasos, ambas redes alcanzan su tamaño máximo de 22 y 23 hexágonos para el caso vibrante y rígido, respectivamente, alrededor del cual se mantienen oscilando durante el resto de la corrida. Este comportamiento del crecimiento de la red se refleja en la energía potencial, figura 6.14 c), la cual descende y se aplanan al mismo ritmo para ambos modelos, solo se observa un corrimiento hacia valores menos

negativos para el caso vibrante. Este corrimiento hacia arriba se debe a la contribución energética de las ligaduras del modelo vibrante.

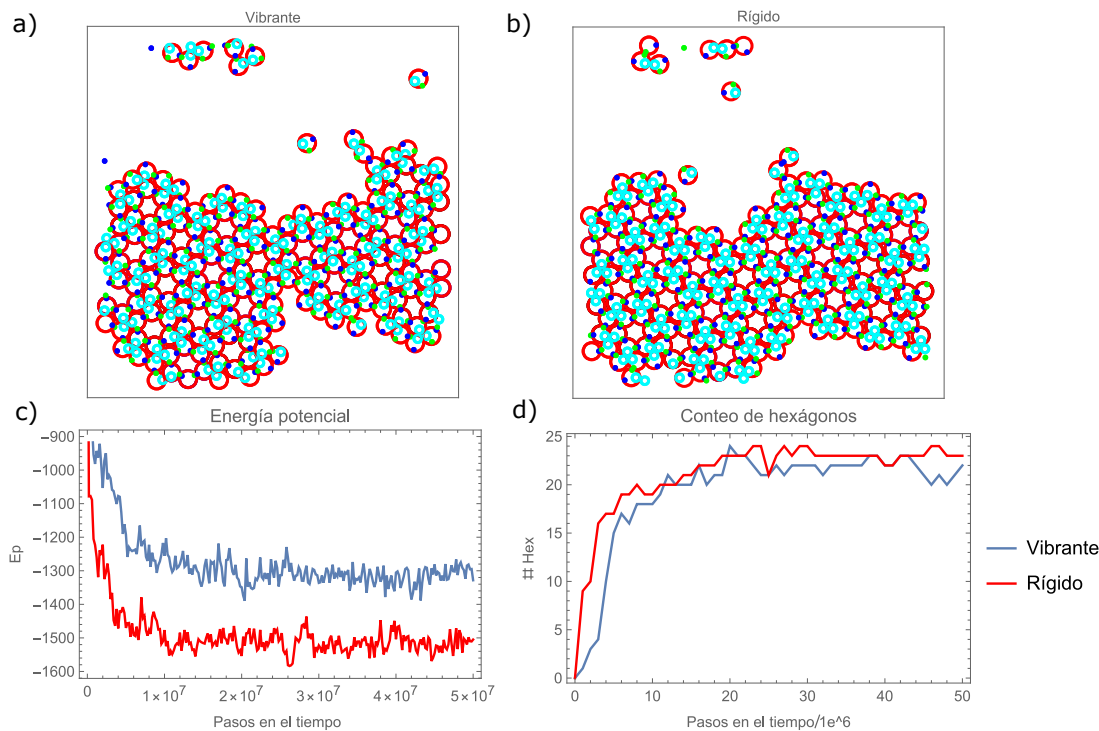


Figura 6.14: Comparación modelo rígido vs vibrante. Se muestran las redes que se obtienen con cada uno de los modelos y se comparan la energía potencial y el conteo de los hexágonos formados a lo largo de la corrida.

Para la comparación de los sistemas rígido y vibrante en 3D se usó el modelo M2 en simulaciones de 120 millones de pasos en el tiempo, con un paso de  $dt = 0.001$ , para sistemas de 216 moléculas a una densidad atómica de  $\rho = 0.512$ . Para el caso rígido se exploró en un rango de temperaturas de  $T = 0.62$  a  $T = 0.74$ . Se encontró que por arriba de  $T = 0.72$  ya no hay formación de hexágonos y a  $T = 0.66$  se encontró uno de los mejores ensamblajes de la red con un total de 9 hexágonos, lo cual se muestra en la figura 6.15 b). Para el caso vibrante se exploró en el mismo rango de temperaturas, pero ahora se encontró que a partir de  $T = 0.68$  el número de hexágonos es menor de 4, para el final de la corrida, y a partir  $T = 0.72$  ya no hay formación de hexágonos. En la figura 6.15 a) se muestra la mejor red formada por el modelo vibrante, encontrada en  $T = 0.66$  y compuesta por un total de 17 hexágonos. En los paneles c y d de la figura 6.15 se muestra la evolución de la energía potencial y el conteo de hexágonos de los modelos rígido y vibrante. Para el caso vibrante, se observa un crecimiento continuo del número de hexágonos a lo largo de toda la corrida. En cambio, en el modelo rígido, durante largos lapsos de la simulación no hay formación de nuevos hexágonos, lo cual da como resultado

el bajo número de hexágonos al final de la simulación.

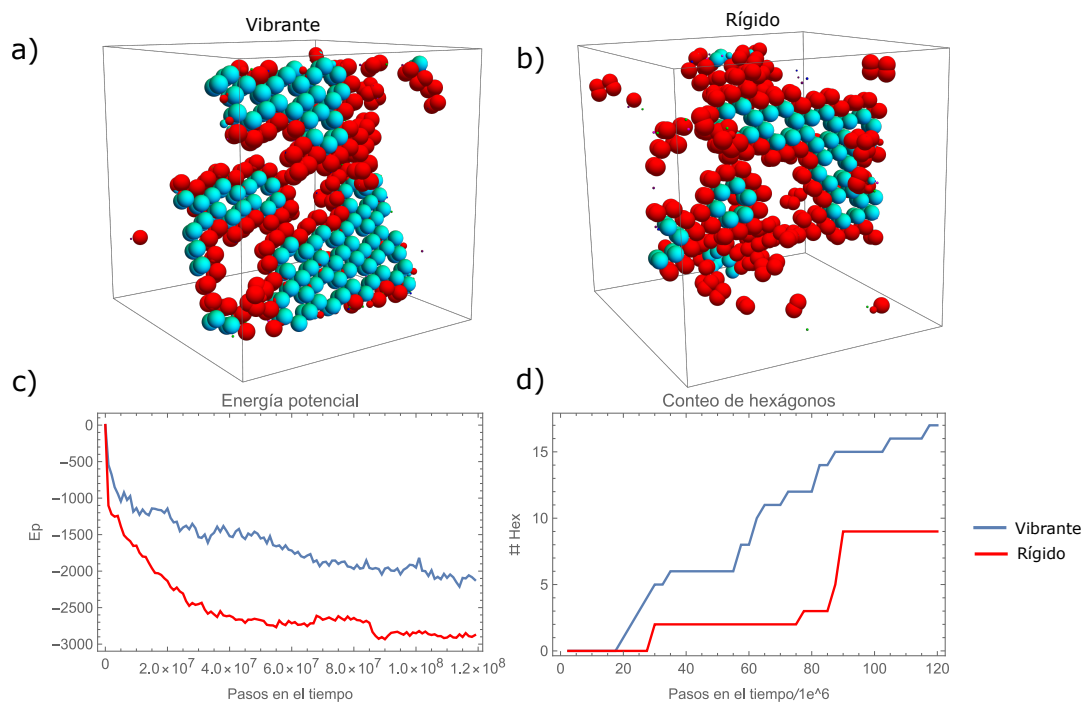


Figura 6.15: Comparación de los modelos rígido y vibrante en 3D. Se compara la energía potencial y el número de hexágonos formado de las mejores corridas de cada modelo.

De la comparación en 2D se ve que tanto en el modelo rígido como en el modelo vibrante el número de hexágonos que forma la red crece de manera similar a lo largo de la corrida y que a ambos modelos les toma 20 millones de PT llegar al tamaño máximo de red que pueden formar. Por otro lado, las diferencias las encontramos en el tiempo de simulación de estos dos modelos. Al modelo rígido le toma 34 minutos llegar a 20 millones de PT, haciéndolo con un promedio de pasos por segundo (TPS) de 9743. En cambio, al modelo vibrante le toma solo 20 minutos alcanzar los 20 millones de PT, haciéndolo con un TPS de 16619. Estos resultados indican que tanto el modelo rígido como el modelo vibrante llegan al mismo tamaño de red en el mismo número de PT, pero que el modelo vibrante lo hace 1.7 veces más rápido.

En 3D las diferencias entre los modelos rígido y vibrante son más notorias, ya que al final de las corridas, el modelo vibrante alcanza un tamaño de red de casi el doble de hexágonos comparado con su versión rígida. Esto es debido a que el modelo rígido muestra una dificultad mayor para ensamblar hexágonos, no solo respecto de su contraparte vibrante en 3D, sino también de su versión rígida en 2D. Además, en cuanto a los tiempos de simulación, el modelo rígido está en desventaja con la versión vibrante. Al caso rígido le toma 4.88 horas completar 120 millones PT, mientras que a la versión vibrante le toma

2.83 horas completar el mismo número de pasos. Nuevamente el modelo vibrante es 1.72 veces más rápido que el modelo rígido.

Otra comparación entre los modelos rígido y vibrante se hizo con el ensamblaje de la cápside icosaédrica  $T=1$ . Para este sistema, del cual se hablará en detalle más adelante, en una simulación con 180 moléculas a una misma densidad atómica de  $\rho = 0.512$ , la versión vibrante forma la cápside en 25 millones de PT, mientras que la versión rígida lo hace en 40 millones de PT. Nuevamente la versión vibrante tiene ventaja sobre la rígida.

De la comparación de los modelo rígido y vibrante, el que tuvo un mejor desempeño en la formación de hexágonos fue el modelo vibrante, además de ser más rápido. Debido a esto, el ensamblaje de la red hexagonal del VIH-1 en 3D con un número mayor de moléculas y las cápsides icosaédricas, usando el potencial continuo de LJ, se harán con modelos moleculares vibrantes.

## 6.2.2. Red hexagonal del VIH-1 en 3D

### 6.2.2.1. Sin interacción 2-Fold

La red hexagonal es ensamblada exitosamente por el modelo M2. En los primeros dos paneles de la Figura 6.16 se muestra el resultado de una simulación de 512 moléculas a una temperatura de  $T = 0.66$ , con densidad de  $\rho = 0.512$ . Para este y todos los demás modelos siguientes se parte de un arreglo cuadrado como configuración inicial. En la figura 6.16 a) se ve claramente la red con los hexágonos pintados en azul, para una mejor visualización, mientras que las moléculas que no forman parte de algún hexágono son pintadas en rojo. En el panel intermedio, se muestra una vista lateral de la red donde pareciera que hay varios segmentos de red, pero en realidad se trata de una sola red compuesta por 55 hexágonos atravesando la caja de simulación gracias a las condiciones periódicas de frontera.

En la figura 6.16 c) se muestra el conteo de hexágonos a lo largo de las corridas de una isócora usando el modelo M2. La densidad de la isócora es de  $\rho = 0.512$  y la temperatura va de  $T = 0.58$  a  $T = 0.7$  con incrementos de  $\Delta T = 0.02$ . En la gráfica se ve claramente como solo dos corridas,  $T = 0.64$  y  $T = 0.66$ , superan los 50 hexágonos formados. En ambas temperaturas desde los  $340 \times 10^6$  de pasos, de los  $400 \times 10^6$  totales que se corrieron, las redes alcanzan su máximo de 55 hexágonos y desde ese punto solo suben y bajan en una unidad su conteo de hexágonos. En la siguiente temperatura de  $T = 0.68$  el conteo de



hexágonos llega solo hasta 46 para el final de la corrida. En esta temperatura es a partir de los  $80 \times 10^6$  de pasos que empieza un crecimiento continuo del número de hexágonos. Este crecimiento no solo se da más tarde sino también más lento que las temperaturas anteriores. Para el final de la corrida el conteo sigue aumentando, pero necesitará de más pasos en el tiempo para alcanzar un valor cercano a las dos temperaturas anteriores. Para la temperatura final de  $T = 0.7$  la simulación termina solo con dos hexágonos, sin siquiera estar unidos entre ellos. Para las temperaturas más bajas de  $T = 0.58$  a  $T = 0.62$  la formación de hexágonos tiene un buen ritmo de crecimiento desde el inicio, pero en ninguno de los tres casos se logra alcanzar un valor cercano a los 55 hexágonos de las dos mejores corridas. En los 3 casos desde los  $290 \times 10^6$  de pasos el conteo llega a su máximo de 32, 48 y 47 hexágonos para  $T = 0.58$ ,  $T = 0.6$  y  $T = 0.62$ , respectivamente.

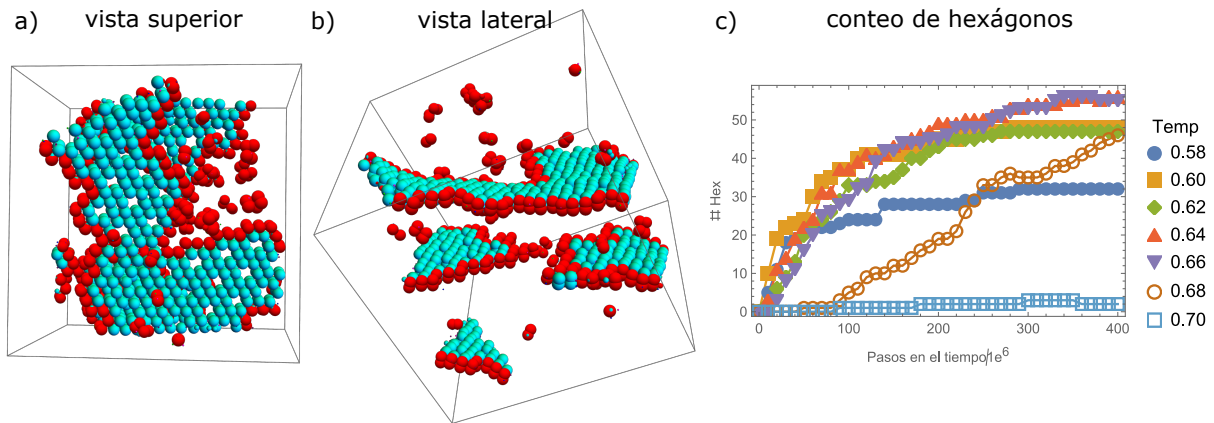


Figura 6.16: Red hexagonal del virión del VIH-1 en su etapa inmadura ensamblada por el modelo M2 sin la interacción 2-Fold. **a)** Vista superior de una red compuesta por 55 hexágonos con parámetros:  $\rho = 0.512$ ,  $T = 0.66$  y 512 moléculas. **b)** Vista lateral de la misma red. **c)** Conteo de hexágonos a lo largo de 7 corridas a la misma densidad y número de moléculas de  $\rho = 0.512$  y  $nmol = 512$ , con temperaturas que van de  $T = 0.58$  a  $T = 0.7$ .

Del conteo de hexágonos resalta que hay una pequeña región de temperaturas donde la red formada alcanza el mayor tamaño posible hasta estabilizar su crecimiento. Por arriba de esta región de temperaturas, primero se torna más lento el crecimiento de la red, y posteriormente se llega a un límite de temperatura donde ya no se formarán hexágonos. Por debajo de esta región la red no alcanza su tamaño óptimo.

### 6.2.2.2. Con interacción 2-Fold

Con el modelo M3, donde si se incluye la interacción 2-Fold, también se ensambla la red de hexágonos. En la figura 6.17 a) se muestra una red compuesta por 57 hexágonos.

En esta red además de pintar las moléculas en azul claro (formadoras de hexágonos) y rojo (no formadoras de hexágonos) algunas moléculas están pintadas en azul fuerte. Este grupo de moléculas tampoco forman parte de algún hexágono, pero se encuentran delimitando huecos que se hayan creado en el interior de la red. Estos huecos se forman por la falta de unas pocas moléculas para completar los hexágonos, formando huecos pequeños, o por la falta de varias moléculas, formando huecos más grandes como se ve en la red.

De este modo, los defectos que aparecen en la red impidiendo su continuidad no son más que huecos provocados por la ausencia de hexágonos en el interior de la red. Esto concuerda con estudios realizados sobre la estructura que adoptan las proteínas Gag en viriones inmaduros.<sup>41,42,44</sup> En estos estudios se observa una sola red interconectada de hexágonos con varios defectos causados por la falta de proteínas Gag y no por Gag desordenados o ensamblados de manera incorrecta. Estos defectos ayudan a liberar la tensión que se produce al doblarse la red para formar el virión esférico. En una red hexagonal la inserción de pentágonos es necesaria para permitir una curvatura sobre ella. Sin embargo, para el caso del VIH-1 Brigs et al.<sup>44</sup> concluye que los defectos no tienen una preferencia por adoptar formas pentagonales, sino más bien adoptan geometrías diversas, como es el caso de la red de la figura 6.17.

En la figura 6.17 b) se muestra el conteo de hexágonos para el modelo M3 en una isócora a la misma densidad que en el modelo M2, de  $\rho = 0.64$ . Debido a la inclusión del parche 2F las temperaturas donde se da una mejor formación de las redes aumentan con respecto al modelo M2, por lo que las temperaturas sobre la isócora ahora van de de  $T = 0.76$  a  $T = 0.9$  con incrementos de  $\Delta T = 0.02$ . Los resultados son similares a los del modelo M2. Nuevamente se observa una región de temperaturas, ahora más amplia, de  $T = 0.78$  a  $T = 0.86$  donde las redes alcanzan un número mayor de hexágonos de entre 54 a 59 hexágonos. Por arriba de esta región se vuelve a observar un decaimiento en el ritmo de crecimiento de la red hasta llegar al punto donde no habrá formación de hexágonos. Por debajo de esta región, los pequeños grupos de hexágonos al unirse entre ellos lo hacen de manera incorrecta sin la posibilidad de acomodarse para poder crecer una red más grande.

Dentro de la región de temperaturas donde se estabilizan las redes con un número mayor de hexágonos, en las temperaturas  $T = 0.78$  y  $T = 0.8$  se forma una red con defectos. Como la de la figura 6.17 a) que corresponde a  $T = 0.8$ . En las siguientes

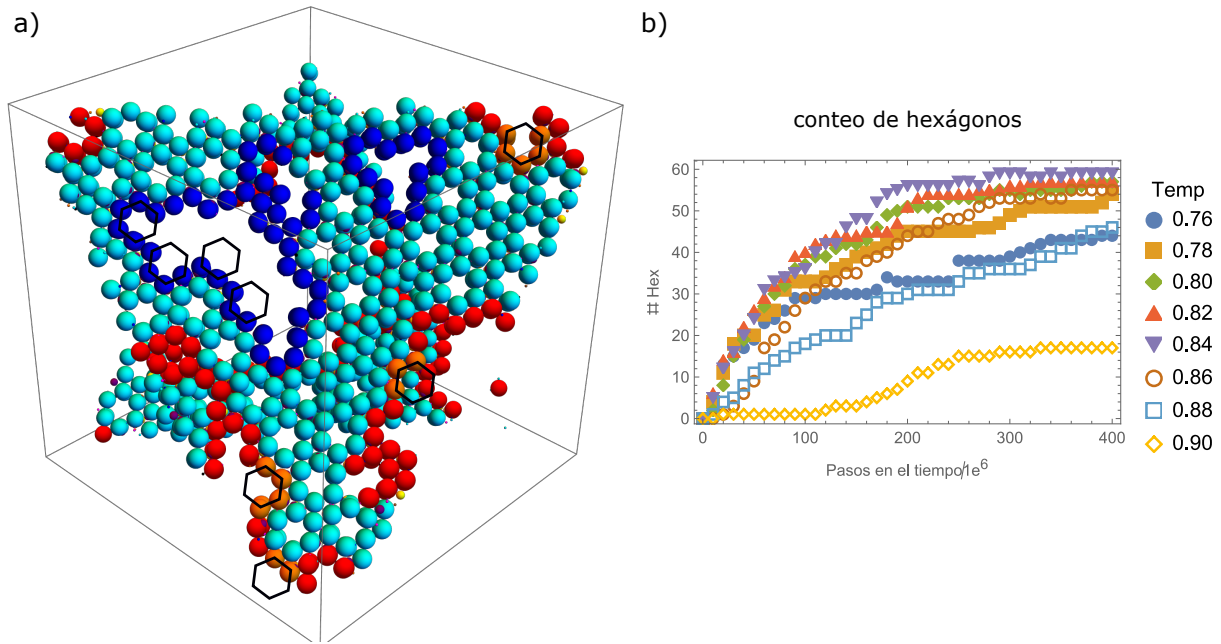


Figura 6.17: Red hexagonal del virión del VIH-1 en su etapa inmadura ensamblada por el modelo M3 con la interacción 2-Fold. **a)** vista superior de una red compuesta por 57 hexágonos con parámetros:  $\rho = 0.64$ ,  $T = 0.8$  y  $nmol = 512$ . En azul fuerte se pintan las moléculas que delimitan los defectos de la red. En naranja se pintan algunas moléculas que componen hexágonos incompletos en los extremos de la red. Líneas negras marcan algunos hexágonos incompletos en los extremos de la red y en los defectos. **b)** Conteo de hexágonos a lo largo de 8 corridas a la misma densidad y número de moléculas de  $\rho = 0.64$  y  $nmol = 512$ , con temperaturas que van de  $T = 0.76$  a  $T = 0.9$ .

temperaturas de  $T = 0.82$  a  $T = 0.86$  las redes que se forman son continuas y libres de defectos. Este comportamiento también se presenta en el modelo M2, ya que de las dos temperaturas que incluyen la región de mejor formación, la red formada a  $T = 0.64$  tiene defectos, pero la red formada a  $T = 0.66$  (Fig. 6.16 (a-b)) no los presenta. El proceso de crecimiento de la red se da a partir de la unión de pequeños grupos de hexágonos formados en los primeros millones de pasos. Posteriormente la red resultante de este proceso va ensamblando hexágonos en su periferia a partir de las moléculas que están libres. Es en este proceso inicial de crecimiento de la red donde se forman los defectos. Sin embargo, en las temperaturas más altas de la región de formación de las redes, los hexágonos tienen una capacidad mayor de acomodarse por lo que les es posible no dejar huecos en el proceso inicial del crecimiento de la red.

A partir de la alineación de subtomogramas obtenidos de Tomografías Crio-Electrónicas (TC-E) de viriones inmaduros<sup>41,44,87</sup> se observa que los hexágonos de los extremos de la red, incluyendo los defectos, tienen una correlación cruzada menor que los demás hexágonos. A partir del procesamiento y análisis de los datos de las TC-E, Tan et al.<sup>88</sup> concluyen que esto se debe a que los extremos de la red del VIH-1 están formados por

hexágonos incompletos.

Ambas redes formadas por los modelos M2 y M3 se encuentran rodeadas por moléculas pintadas en rojo que están unidas a toda la periferia de la red. Lo mismo sucede con los defectos en el interior de la red (moléculas pintadas en azul Fig. 6.17 a)). Estas moléculas son pintadas en rojo y azul por no pertenecer a un hexágono. Pero para el caso específico de estas moléculas unidas a la red, en concordancia con Tan et al.,<sup>88</sup> se puede decir que son hexágonos incompletos a los que les faltan una o más moléculas. De este modo, las redes ensambladas con los modelos M2 y M3 concuerdan con lo observado en viriones inmaduros. En la figura 6.17 a) se pinta en color naranja cuatro hexágonos incompletos de 5, 4, 3, y 2 moléculas en los extremos de la red. En el defecto más grande de la red también se marcan algunos hexágonos incompletos.

### 6.2.3. Cápsides icosaédricas

Para ensamblar las cápsides icosaédricas se usan los modelos M4 y M5 donde las esferas repulsivas tienen un diferencia en sus diámetros formando un ángulo  $\phi$  entre ellas (Fig. 4.4 a)). El valor de este ángulo dictará el grado de curvatura que tendrá la red hexagonal. Para obtener estructuras cerradas como las cápsides es necesaria la inserción de pentágonos en la red. Los modelos M2 y M3 a temperaturas bajas formaban unos pocos pentágonos. Estos pentágonos a temperaturas más altas desaparecen, sin embargo la curvatura producida por el ángulo  $\phi$  es capaz de estabilizar los pentágonos a temperaturas altas. Ya desde la red ensamblada con el modelo con potencial de PC (Fig. 6.12 b)) se observaba una red con una gran curvatura gracias a la formación de un pentágono y otros defectos. De este modo, el mismo modelo molecular, M4 y M5, puede formar pentágonos y hexágonos, por lo que solo dependerá del ángulo  $\phi$  el tamaño de la cápside que se quiera formar.

En la Figura 6.18 se muestran las cápsides obtenidas con el modelo M4 que no incluye la interacción 2-Fold. Los hexágonos formados son pintados de cian, los pentágonos de negro y las moléculas que no forman parte de algún polígono son pintadas de rojo la esfera grande y verde la esfera pequeña. Las cápsides formadas son  $T = 1$ ,  $T = 3$  y  $T = 4$  con número de capsómeros de  $C_1 = 12$  pentámeros,  $C_3 = 20$  hexámeros + 12 pentámeros y  $C_4 = 30$  hexámeros + 12 pentámeros, respectivamente. Los ángulos requeridos para ensamblar la cápsides correctamente son  $\phi_{T1} = 12.7^\circ$ ,  $\phi_{T3} = 7.4^\circ$  y  $\phi_{T4} = 6.2^\circ$ .

En la Figura 6.18 también se muestra media cápside  $T = 7$  donde se observa como los

hexámeros y pentámeros están bien posicionados, lo cual indica que el ángulo de  $\phi_{T7} = 4.6^\circ$  podría formar bien la cápside. Sin embargo no fue posible cerrar una cápside con el número de capsómeros correctos. Esto se debe a que para los números de triangulación  $T(h, k)$  con  $k > 0$  como  $T(2, 1) = 7$ ,  $T(3, 1) = 13$  o  $T(3, 2) = 19$  las cápsides son enantiomorfas y se pueden ensamblar en su forma izquierda (levógira) y derecha (dextrógira). Es decir, una cápside se puede ensamblar de dos formas, por girar a la izquierda y recorrer  $k$  capsómeros para llegar de un pentágono a otro o por girar a la derecha para el mismo motivo. Al no imponer una restricción en la simulación para elegir una de estas dos rutas la cápside  $T = 7$  combina sus dos enantiómeros, lo cual impide la formación correcta de la cápside.

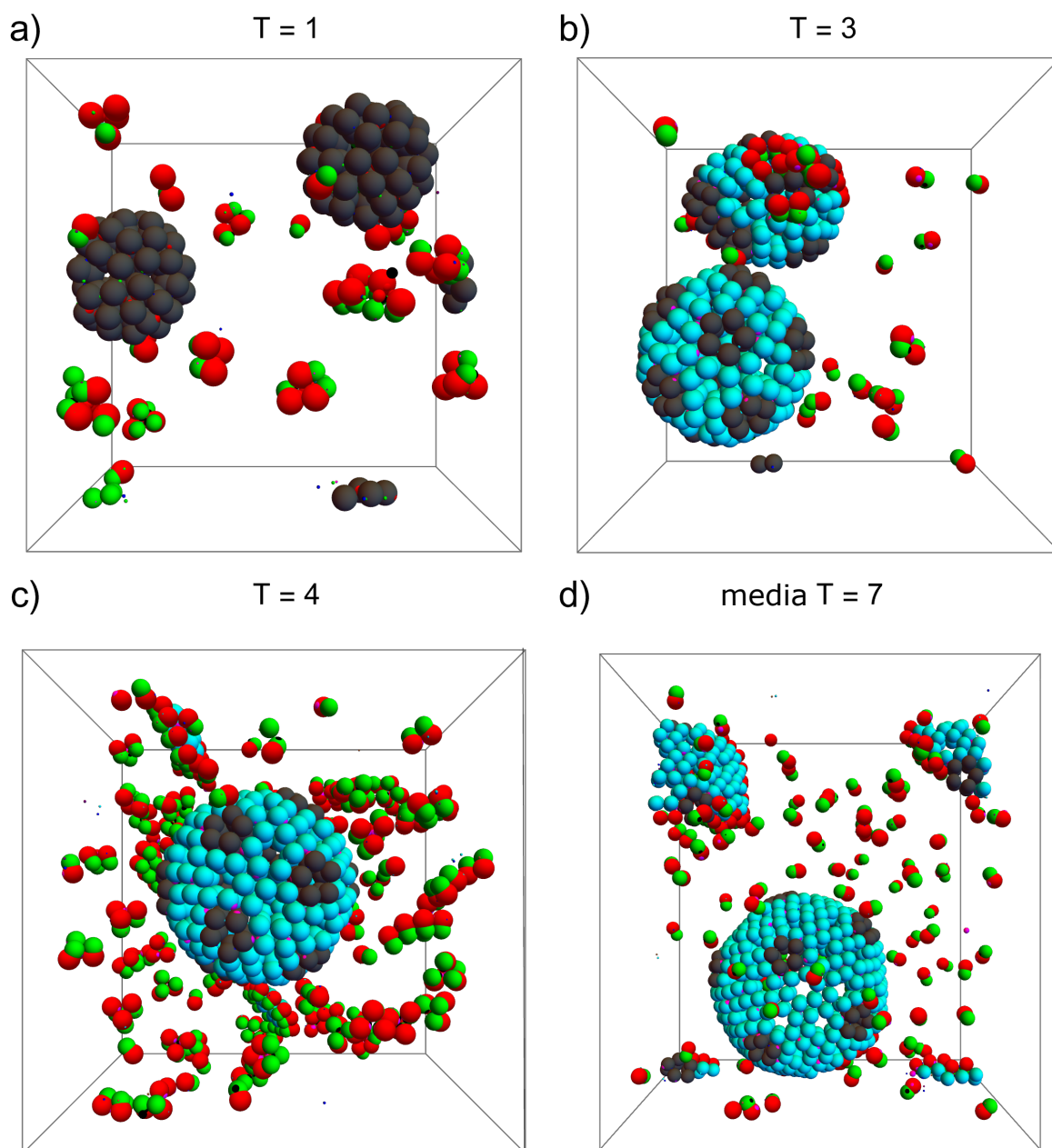


Figura 6.18: Cápsides icosaédricas ensambladas con el modelo M4. En todos los casos se usa una densidad de  $\rho = 0.512$ . **a)** Cápsides  $T=1$ , formadas solo por pentágonos, compuestas por 60 moléculas cada una en una corrida de 180 moléculas. **b)** Cápside  $T = 3$  compuestas por 180 moléculas en una corrida de 343 moléculas. Esta cápside está acompañada por una estructura cerrada que no tiene la distribución de hexágonos y pentágonos correcta. **c)** Cápside  $T = 4$  compuestas por 240 moléculas en una corrida de 512 moléculas. **d)** Cápside  $T = 7$  incompleta compuesta por 49 hexágonos y 9 pentágonos de los 60 hexágonos y 12 pentágonos requeridos.

Para encontrar los ángulos  $\phi_T$  que ensamblan las cápsides se buscó alrededor de un ángulo inicial  $\phi_i$ . El ángulo inicial se determinó a partir de ensamblar manualmente algunos capsómeros contiguos usando de base un dodecaedro para  $T = 1$  y un icosaedro truncado para  $T = 3$ . Durante el proceso se ajustó el diámetro de las esferas de las

moléculas para mantenerlas tangentes con sus vecinas y así obtener los diámetros que determinarían el ángulo  $\phi_i$ . Una vez encontrado el ángulo  $\phi_T$  y ensamblado correctamente las cápsides, se usó ese mismo ángulo para ensamblar las cápsides con el modelo M5. Para determinar el ángulo  $\phi_{T1}$  se calculó un ángulo inicial  $\phi_i = 12.7^\circ$  y fue en ese ángulo donde se ensambló la cápside. Para  $T=3$  se calculó un ángulo inicial  $\phi_i = 6.6^\circ$ , lo cual corresponde a un diámetro de la segunda esfera de  $\sigma_2 = 0.885$ . Se buscó de  $\sigma_2 = 0.86$  hasta  $\sigma_2 = 0.92$  con un incremento de  $\Delta\sigma = 0.01$ . Esto corresponde a buscar en un rango de ángulos que va de  $5.1^\circ$  a  $8^\circ$ , rango dentro del cual se encontró  $T=3$  ( $\phi_{T3} = 7.4^\circ$ ) y  $T=4$  ( $\phi_{T4} = 6.2^\circ$ ). Para  $T = 7$  se calculó un ángulo de  $4^\circ$  y se buscó hacia arriba hasta los  $5^\circ$  con un delta de  $0.02^\circ$ .

Además de formar las cápsides con el modelo M5, que si incluye la interacción 2-Fold, se usaron dos modelos moleculares más. Estos modelos son idénticos a los modelos M4 y M5, pero se diferencian en que su parche 3-Fold tiene una profundidad del pozo de  $\epsilon = 4$  mientras que los demás parches se mantienen con la profundidad de  $\epsilon = 6$ . Esta diferencia está motivada en la medición de las fuerzas de atracción de las interacciones 2-Fold, 3-Fold e IH del Gag del VIH. Estos datos se obtuvieron con PDBePISA,<sup>89</sup> una herramienta que calcula la energía libre de Gibbs de las interfaces que se producen entre los sitios de interacción de las proteínas. El resultado de esta medición fue que la interacción 3-Fold es tres veces más débil que las interacciones 2-Fold e IH. Debido a esto se cambió el valor de  $\epsilon$  para el parche 3-Fold. Sin embargo como los modelos M4 y M5 tienen profundidad de  $\epsilon = 6$  lo más que se consiguió disminuir el valor de epsilon, y mantener la formación de parejas y ternas a la misma temperatura, fue a  $\epsilon = 4$ .

En la Tabla 6.1 se muestran los datos de densidad, número de moléculas y rangos de temperaturas para los que se forman las cápsides utilizando los cuatro modelos. Los modelos con el cambio en la profundidad del potencial serán identificados como  $M4_\epsilon$  (sin 2-Fold) y  $M5_{2F\epsilon}$  (con 2-Fold). Al observar los datos de la tabla se encuentran algunos hechos interesantes. De la sexta columna de la tabla se ve como es que conforme la cápside es más grande el rango de temperaturas donde se formaron se reduce. Las últimas tres columnas de la tabla indican la temperatura, pasos en el tiempo y tiempo de simulación de la cápside que se ensambló más rápidamente dentro de su rango de temperaturas de formación. Con estos datos se pueden hacer varias observaciones. La más clara es que cuanto más grande es la cápside se necesitan más pasos en el tiempo para formarla. También es notable que

No. Triang.	$\phi_T$	$\rho$	No. mol	Modelo	Rango Temp	Temp	PT/10 <sup>6</sup>	t (hr:min)
$T = 1$	12.7	0.512	180	M4	0.65 - 0.71	0.69	25	0:38
		0.64		M5	0.84 - 0.9	0.84	30	0:48
		0.512		M4 <sub><math>\epsilon</math></sub>	0.67 - 0.75	0.69	10	0:14
		0.64		M5 <sub>2F<math>\epsilon</math></sub>	0.81 - 0.89	0.87	15	0:22
$T = 3$	7.4	0.512	343	M4	0.68 - 0.695	0.695	315	10:09
		0.64		M5	0.86 - 0.88	0.88	147.5	8:30
		0.512		M4 <sub><math>\epsilon</math></sub>	0.61 - 0.62	0.62	190	9:40
		0.64		M5 <sub>2F<math>\epsilon</math></sub>	0.775 - 0.79	0.79	102	3:35
$T = 4$	6.2	0.512	512	M4	0.695	0.695	430	18:16
		0.64		M5	0.88 - 0.89	0.88	522	27:35
		0.512		M4 <sub><math>\epsilon</math></sub>	0.61	0.61	245	11:15
		0.64		M5 <sub>2F<math>\epsilon</math></sub>	0.765 - 0.77	0.77	180	7:55

Cuadro 6.1: Modelos que forman cápsides icosaédricas. M4 (sin 2-Fold), M5 (con 2-Fold), M4 <sub>$\epsilon$</sub>  (sin 2-Fold con  $\epsilon$  menor para el parche 3-Fold) y M5<sub>2F $\epsilon$</sub>  (con 2-Fold y con  $\epsilon$  menor en 3-Fold). En las columnas se presenta el número de triangulación; ángulo que forma la cápside ( $\phi_T$ ); densidad atómica ( $\rho$ ) y número de moléculas (No. mol) de la simulación; modelo usado; rango de temperaturas donde se ensambla la cápside (Rango Temp). En las últimas 3 columnas se presenta la temperatura (Temp); pasos en el tiempo (PT) y tiempo (t) de la cápside que se ensambló más rápido dentro de su rango de temperaturas de formación.

para cada número de triangulación las cápsides formadas con los modelos M4 <sub>$\epsilon$</sub>  y M5<sub>2F $\epsilon$</sub>  se ensamblan más rápidamente que su contra parte M4 y M5. Otro hecho remarcable es que para las cápside  $T = 3$  y  $T = 4$  los modelos M5 y M5<sub>2F $\epsilon$</sub> , que incluyen el parche 2-Fold, son de 1.2 a 2.1 veces más rápidos que los modelos M4 y M4 <sub>$\epsilon$</sub>  que no lo incluyen. Para el caso de T1 esta observación se cumple a la inversa Ahora son las versiones sin el parche 2-Fold, M4 y M4 <sub>$\epsilon$</sub> , las que son 1.2 y 1.5 veces más rápidas que su versión M5 y M5<sub>2F $\epsilon$</sub>  que si lo tienen. Por último, para los números de triangulación  $T = 3$  y  $T = 4$  la temperatura a la que se forma la cápside más rápidamente se da a temperaturas muy similares para los cuatro modelos. Nuevamente las cápsides  $T = 1$  se comportan de manera diferente, siendo solamente la temperatura del modelo M4 la que se acerca a los otros números de triangulación.

### 6.2.3.1. Estabilidad térmica

Para probar qué tan estables son las cápsides formadas con los modelos, se eligió la cápside T=3 ensamblada con el modelo M4 <sub>$\epsilon$</sub>  a la temperatura de  $T = 0.61$  para ser sometida a un aumento y disminución de la temperatura. Se tomó la última configuración de la simulación como punto de partida y se corrieron seis simulaciones por 50 millones de pasos en el tiempo a las temperaturas siguientes: 0.6, 0.5, 0.4, 0.3, 0.2 y 0.1. En todos estos casos las cápsides se mantuvieron ensambladas sin inconvenientes.

El aumento de temperatura se hizo de dos maneras. Primero se aumentó la tempe-



ratura de manera cuasiestática aplicando incrementos de temperatura de  $\Delta T = 0.02$ . Se tomó la última configuración de la temperatura anterior como punto de partida en cada incremento de temperatura. La segunda manera de incrementar la temperatura fue a base de choques térmicos, como en el caso de la disminución. Se aumentó la temperatura en múltiplos de  $\Delta T = 0.01$  empezando todas las simulaciones a la misma temperatura de  $T = 0.61$ . Estas simulaciones nuevamente se corrieron por 50 millones de pasos en el tiempo. Para el caso del choque térmico la cápside mantuvo su estructura estable hasta  $T = 0.69$ , desintegrándose por completo en la siguiente temperatura. Para el caso del incremento cuasiestático la cápside se mantuvo estable a una temperatura mayor de  $T = 0.73$ .

#### 6.2.4. Cápsides no-icosaédricas

Durante la búsqueda de las cápsides icosaédricas se encontraron otro tipo de estructuras. Estas estructuras se cierran sin dejar huecos, aislando su interior del exterior, por lo que también se les puede llamar cápsides. Estas cápsides, al igual que las icosaédricas, cuentan con 12 pentámeros, pero ahora no se colocan en los vértices de un icosaedro sino que se distribuyen de distintas maneras. El número de hexámeros dependerá del tamaño de la cápside, que a su vez depende del ángulo  $\phi$  de la molécula.

En la Figura 6.19 se muestran las cápsides no icosaédricas encontradas. Un hecho interesante de estas cápsides es que, sin importar el número de hexámeros del que esté compuesta la cápside, siempre se guarda una simetría en el acomodo de los pentámeros y hexámeros. Esto se ve fácilmente en los paneles b) y e) de la Figura 6.19. En el panel b) se observa una cápside alargada con un anillo zigzagueante de hexámeros en el ecuador, mientras que en los polos se concentran los pentámeros. En el panel e) se muestra una cápside con forma de tetraedro con un hexámero en cada una de sus cuatro caras. Estos acomodos simétricos entre pentámeros y hexámeros son más fácil de identificar en las primeras cinco cápsides, pero las 4 restantes también presentan arreglos simétricos entre pentámeros y hexámeros.

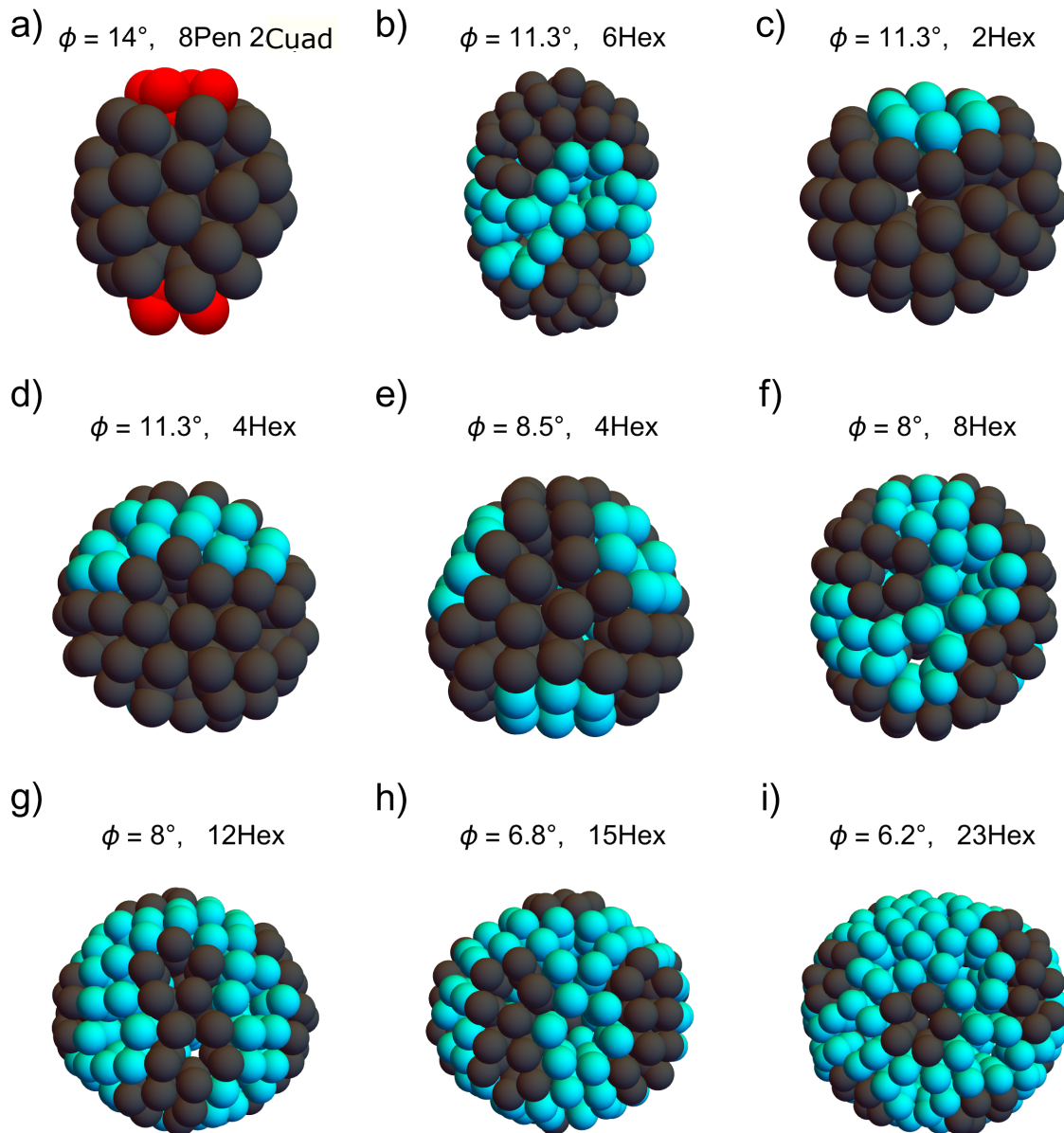


Figura 6.19: Cápsides no icosaédricas ensambladas con el modelo M4. a) Cápside más pequeña encontrada formada por 8 pentámeros y dos cuadrados, pintados en rojo, uno en cada polo. (b - i). Cápsides de distintos tamaños ensambladas con un número diferente de hexámeros, pero manteniendo, en todos los casos, 12 pentámeros. En todos los casos se muestra el ángulo  $\phi$  de la molécula que se usó.

En la Figura 6.19a) se muestra una cápside más pequeña que la cápside  $T = 1$  y por tal motivo solo cuenta con 8 pentámeros. Esta cápside ensambló dos cuadrados, uno en cada polo, para poder cerrarse por completo. Esto fue posible, por un lado, por el ángulo  $\phi = 14^\circ$ , mayor que el de la cápside  $T = 1$  de  $\phi_{T1} = 12.7^\circ$  el cual induce una curvatura mayor. Y por otro lado, se debe a la capacidad del modelo de formar pentágonos e incluso cuadrados a partir de un modelo diseñado para formar hexágonos. Esta capacidad está

basada en el hecho de que las esferas repulsivas pueden rotar alrededor de ellas mismas manteniendo a los parches IH, formadores de parejas, a la distancia del mínimo de su potencial, dada por  $d_{min} = 2^{1/6}\sigma_{IH}$ . Si la distancia entre los parches es igual a la distancia del mínimo del potencial, como se muestra en la Figura 6.20, esta rotación corresponde a una variación del ángulo que da posición a los parches sobre el ecuador de  $\theta = \pm 11.6^\circ$ . Esto quiere decir que el parche situado a  $120^\circ$  para formar hexágonos queda muy cerca del ángulo para formar pentágonos ( $108^\circ$ ) debido a la rotación  $\theta$ , lo cual le permite formar hexágonos y pentágonos. No se debe olvidar que la aparición de pentágonos solo sucede cuando la red plana se intenta formar a temperaturas bajas y cuando se le induce una curvatura a la red por medio del ángulo  $\phi$ .

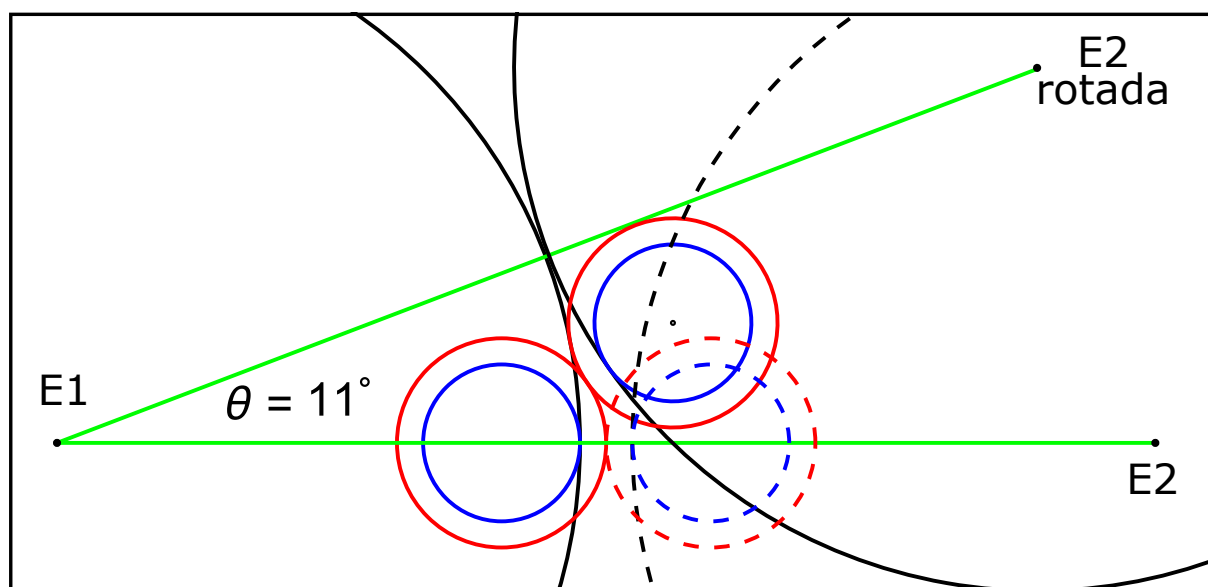


Figura 6.20: Esquema de la rotación de las esferas repulsivas alrededor de los parches pareja IH, lo cual les permite formar pentágonos. Si los parches (pintados en azul) siempre se mantienen a la distancia de su mínimo de potencial (pintado en rojo), los centros de la esfera E1 y la esfera E2 rotada, después de rotar, alcanzan un ángulo de  $\theta = 11^\circ$  respecto de la esfera punteada sin rotar E2.



# Capítulo 7

## Conclusiones

El resultado principal de este trabajo es que a partir de modelos sencillos se reproducen estructuras de una gran complejidad. Es de destacar que modelos de grano grueso con un mayor detalle estructural del dominio CA del Gag del VIH-1, logran formar en 2D la red hexagonal,<sup>90</sup> pero en 3D, a pesar de formar hexámeros y pentámeros, no obtienen estructuras cerradas.<sup>91</sup> También hay modelos sencillos que logran ensamblar hasta la cápside T=3,<sup>75</sup> pero requieren de tres especies moleculares distintas y reglas de interacción complejas adaptadas para cada caso. En cambio, los modelos presentados aquí requieren de una sola especie molecular como bloque constructor y no se necesitan cambiar las reglas de interacción para ensamblar una red hexagonal o una cápside icosaédrica.

El modelo de esferas traslapadas presentado en este trabajo, en si mismo demuestra ser capaz de formar fácilmente una red en 3D, lo cual le da la posibilidad de ensamblar otras redes en 3D. Respecto a la red del VIH-1, los modelos moleculares utilizados aquí demuestran que la estructura hexagonal se puede formar sin la necesidad de la interacción 6F, la proteína IP6 ni de otros dominios que dan estabilidad a la red en el virión. Otro hecho importante es que la red obtenida con modelo M2 se puede ensamblar sin la necesidad de la interacción 2-Fold del dominio CA del Gag del VIH-1 en su etapa inmadura. Esto es de resaltar ya que esta interacción se mantiene en los hexámeros al pasar de la etapa inmadura a la etapa madura (cápside) del virus. Hecho que no sucede con las interacciones 3-Fold y 6-Fold.

La inserción de pentágonos en una red hexagonal permite curvar la red, pero además el bloque constructor morfológicamente debe tener las condiciones para permitir la curvatura, es decir, debe tener una forma cónica. Cuando los modelos tienen un ángulo  $\phi = 0$

(M2 y M3) los pentágonos no aparecen en la red, pero si  $\phi > 0$  (M4 y M5) los pentágonos se estabilizan adecuadamente dentro de la red. Al medir los diámetros de los dominios del Gag del VIH-1, usando la reconstrucción del Gag completo hecha por Sundquist et al. (figura 1c),<sup>32</sup> se encuentra su forma cónica con un ángulo de  $\phi_{gag} = 2.2$ . Pak et al.<sup>31</sup> usando un modelo de grano grueso encuentra que cuando el dominio CA forma hexámeros en la etapa inmadura, el diámetro del hexágono formado por su N-terminal es mayor que el del formado por su parte C-terminal. Nuevamente se observa una forma cónica. Pack concluye que el dominio CA induce una curvatura en la red hexagonal, hecho que ayuda a doblar la membrana celular que formará el virión del VIH-1. De este modo, la capacidad de los modelos introducidos aquí de formar pentágonos y hexágonos a partir de la misma unidad estructural, junto con la forma cónica de la molécula, hace posible la formación de cápsides icosaédricas. Además se encuentra que la única variable necesaria para elegir entre una u otra cápside es el ángulo  $\phi$ .

Por otra parte, la imposibilidad de los modelos M4 y M5 de formar la cápside  $T = 7$  se debe a no tener control sobre que enantiómero debe ser ensamblado. Además en trabajos de simulación donde se eligen entre enantiómeros y se usan polígonos como bloques constructores de las cápsides<sup>73</sup> la efectividad de formación de las cápsides  $T \geq 7$  se reduce a menos del 10%. En consecuencia, con los resultados presentados aquí se reconoce que las cápsides pequeñas ( $T \leq 4$ ) pueden ser ensambladas libremente, pero las cápsides mayores necesitan de andamiajes<sup>92</sup> para su correcto ensamblaje. Para el virus del VIH estos andamiajes no son más que su material genético y la membrana lipídica de la célula infectada.<sup>31</sup> Estos elementos dan el soporte necesario para formar primeramente el virión inmaduro y posteriormente ensamblar cápsides de gran tamaño.

Con respecto al sistema Gag-Membrana para modelar la formación de un virión envuelto por la membrana, hay dos puntos importantes que señalar. El primero es que se logró el objetivo de ensamblar un virión a partir de unidades tipo Gag y no con estructuras más elaboradas tipo pentámero.<sup>52</sup> Además, se muestra que unidades estructurales tipo cónicas son suficientes para doblar la membrana. El segundo punto tiene que ver con los modelos de membrana lipídica. En primer lugar, se consigue autoensamblar una membrana con un potencial de pozo cuadrado en lugar de un potencial continuo, que fue el que se usó como referencia.<sup>48</sup> En segundo lugar, se consiguió ensamblar una membrana de doble capa con un modelo más sencillo que el de Cooke et al.,<sup>48</sup> compuesto solo por

dos esferas tangentes para representar un lípido. Y no solo se consigue una membrana de doble capa con este modelo más sencillo, sino que es la que se usa para ensamblar y liberar el virión, lo cual demuestra su completa funcionalidad.

## 7.1. Perspectivas

Los resultados presentados en este trabajo son la base de lo que se podría obtener con los sistemas presentados. Ya que se puede continuar obteniendo resultados en cada uno de los sistemas simulados a partir de modificaciones en los modelos o con distintas estrategias de simulación. Por lo que a continuación se enuncia brevemente el trabajo futuro con estos sistemas.

En el sistema Gag-Membrana, el paso siguiente es sustituir las interacciones atractivas isotrópicas entre los dominios de la molécula Gag por los parches atractivos 2F, 3F e IH. Esto con el objetivo de obtener un virión que forme una red hexagonal con las moléculas Gag que lo compongan, asemejándose aun más al virión del VIH-1.

El sistema de dos esferas traslapadas del mismo diámetro formó relativamente fácil la red hexagonal del VIH-1. Debido a esto se buscará ensamblar otras redes abiertas en 3D utilizando el modelo de esferas traslapadas con parches atractivos.

Finalmente, para el sistema de las cápsides icosaédricas se buscará ensamblar la cápside T=7 así como las siguientes cápsides T=9,13,... Para este fin se usará un cascaron esférico dentro del cual interaccionaran las moléculas que formaran la cápside. Ya sea que las moléculas se desplacen sobre la superficie interna del cascaron o libremente dentro él. El objetivo del cascaron es que sirva como andamiaje y facilite la formación de cápsides con números de triangulación grandes. Además, se calcularán los diagramas de fases de las cápsides empezando con las formadas hasta el momento.





# Bibliografía

- [1] A.-P. Hynninen, J. Thijssen, E. Vermolen, M. Dijkstra, and A. Van Blaaderen, “Self-assembly route for photonic crystals with a bandgap in the visible region,” *Nature Materials*, vol. 6, pp. 202–205, mar 2007.
- [2] J. A. Fan, C. Wu, K. Bao, J. Bao, R. Bardhan, N. J. Halas, V. N. Manoharan, P. Nordlander, G. Shvets, and F. Capasso, “Self-assembled plasmonic nanoparticle clusters,” *Science*, vol. 328, no. 5982, pp. 1135–1138, 2010.
- [3] B. Sun and H. Siringhaus, “Solution-processed zinc oxide field-effect transistors based on self-assembly of colloidal nanorods,” *Nano letters*, vol. 5, pp. 2408–13, 01 2006.
- [4] J. Atwood, “A molecular toolkit for magnetism,” *Nature materials*, vol. 1, pp. 91–2, 11 2002.
- [5] N. Ahmad, H. Hassan, G. Zhang, K. Hecke, and F. Verpoort, “Direct synthesis of the 2d copper(ii) 5-prop-2-ynoxysophthlate mof: Comment on “surface functionalization of porous coordination nanocages via click chemistry and their application in drug delivery”,” *Advanced Materials*, vol. 31, 05 2019.
- [6] Q. Chen, S. C. Bae, and S. Granick, “Directed self-assembly of a colloidal kagome lattice,” *Nature*, vol. 469, pp. 381–384, Jan 2011.
- [7] B. Glettner, F. Liu, X. Zeng, M. Prehm, U. Baumeister, M. Walker, M. Bates, P. Boesecke, G. Ungar, and C. Tschierske, “Liquid-crystalline kagome,” *Angewandte Chemie (International ed. in English)*, vol. 47, pp. 9063–6, 11 2008.

- [8] B. Moulton, J. Lu, R. Hajndl, S. Hariharan, and M. Zaworotko, “Crystal engineering of a nanoscale kagome lattice,” *Angewandte Chemie (International ed. in English)*, vol. 41, pp. 2821–4, 08 2002.
- [9] L. Wang, Y. Han, X. Feng, J. Zhou, P. Qi, and B. Wang, “Metal–organic frameworks for energy storage: Batteries and supercapacitors,” *Coordination Chemistry Reviews*, vol. 307, pp. 361–381, 2016. *Chemistry and Applications of Metal Organic Frameworks*.
- [10] I. Stassen, N. Burtch, A. Talin, P. Falcaro, M. Allendorf, and R. Ameloot, “An updated roadmap for the integration of metal–organic frameworks with electronic devices and chemical sensors,” *Chem. Soc. Rev.*, vol. 46, pp. 3185–3241, 04 2017.
- [11] Z. Meng, R. Stolz, L. Mendecki, and K. Mirica, “Electrically-transduced chemical sensors based on two-dimensional nanomaterials,” *Chemical Reviews*, vol. 119, pp. 478–598, 01 2019.
- [12] L. Mendecki, M. Ko, X. Zhang, Z. Meng, and K. Mirica, “Porous scaffolds for electrochemically controlled reversible capture and release of ethylene,” *Journal of the American Chemical Society*, vol. 139, pp. 17229–17232, 11 2017.
- [13] D. Zhao, S. Tan, D. Yuan, W. Lu, Y. Rezenom, H. Jiang, L.-Q. Wang, and H.-C. Zhou, “Surface functionalization of porous coordination nanocages via click chemistry and their application in drug delivery,” *Advanced materials (Deerfield Beach, Fla.)*, vol. 23, pp. 90–3, 01 2011.
- [14] G. Meer, D. Voelker, and G. Feigenson, “Membrane lipids: Where they are and how they behave,” *Nature reviews. Molecular cell biology*, vol. 9, pp. 112–24, 03 2008.
- [15] A. Finkelstein and O. Ptitsyn, *Protein Physics A Course of Lectures*. Academic Press, 2002.
- [16] T. Svitkina, “Ultrastructure of the actin cytoskeleton,” *Current opinion in cell biology*, vol. 54, pp. 1–8, 02 2018.
- [17] S. D. Kundg and S. F. Yang, *Discoveries in Plant Biology*. World Scientific, 1998.
- [18] H. P. Harper, *Viruses: Biology, Applications and Control*. Garland Science, 2012.

- [19] B. la Scola, C. Desnues, I. Pagnier, C. Robert, L. Barrassi, G. Fournous, M. Merchat, M. Suzan-Monti, P. Forterre, E. Koonin, and D. Raoult, “The virophage as a unique parasite of the giant mimivirus,” *Nature*, vol. 455, pp. 100–4, 10 2008.
- [20] M. T. Madugan, J. M. Martinko, K. S. Bender, D. H. Buckley, and D. A. Stahl, *BROCK. Biología de los organismos*. Pearson Educations, 2015.
- [21] R. Delgado, “Características virológicas del vih,” *Enfermedades Infecciosas y Microbiología Clínica*, vol. 29, no. 1, pp. 58–65, 2011.
- [22] M. Popovik, M. G. Samgadharan, E. Read, and R. C. Gallo, “Detection, isolation and continuous production of cytopathic retroviruses (htlv-iii) from patients with aids and pre-aids,” *Science*, vol. 224, pp. 497–500, May 1984.
- [23] M. T. P. Gilbert, A. Rambaut, G. Wlasiuk, T. J. Spira, A. E. Pitchenik, and M. Wrobel, “The emergence of hiv/aids in the americas and beyond,” *Proceedings of the National Academy of Sciences*, vol. 104, no. 47, pp. 18566–18570, 2007.
- [24] S. Wain-Jobson, J. Vartanian, and N. Henry, “Lav revisited origin of the early hiv-1 isolates from institut pasteur,” *Science*, vol. 250, pp. 961–965, May 1991.
- [25] H. Mitsuya, K. J. Weinhold, P. A. Furman, H. S. Clair, M. A. Marty, S. N. Lerman, and R. C. Gallo, “3'-azido-2'-deoxythymidine (bw a509u): An antiviral agent that inhibits the infectivity and cytopathic effect of human t-lymphotropic virus type iii/lymphadenopathy-associated virus in vitro,” *Medical Sciences*, vol. 82, pp. 7096–7100, Oct 1985.
- [26] M. A. Fischl, D. D. Richman, M. H. Grieco, M. S. Gottlieb, P. A. Volberding, O. L. Laskin, J. M. Leedom, J. E. Groopman, D. Mildvan, and R. T. Schooley, “The efficacy of azidothymidine (azt) in the treatment of patients with aids and aids-related complex. a double-blind, placebo-controlled trial,” *The New England journal of medicine*, vol. 317 4, pp. 185–91, 1987.
- [27] S. Li, C. P. Hill, W. I. Sundquist, , and J. T. Finch, “Image reconstructions of helical assemblies of the hiv-1 ca protein,” *Nature*, vol. 407, pp. 409–413, Sep 2000.
- [28] G. Zhao, J. R. Perilla, E. L. Yufenyuy, X. Meng, B. Chen, J. Ning, J. Ahn, A. M. Gronenborn, K. Schulten, C. Aiken, and P. Zhang, “Mature hiv-1 capsid structure

- by cryo-electron microscopy and all-atom molecular dynamics,” *Nature*, vol. 497, pp. 643–646, May 2013.
- [29] E. O. Freed, “Hiv-1 assembly, realise and maturation,” *Nature Reviews*, vol. 13, pp. 484–496, Aug 2015.
- [30] F. K. M. Schur, W. J. H. Hagen, M. Rumlová, T. Ruml, B. Müller, H.-G. Kräusslich, and J. A. G. Briggs, “Structure of the immature hiv-1 capsid in intact virus particles at 8.8 Å resolution,” *Nature*, vol. 517, pp. 505–508, 01 2015.
- [31] A. J. Pak, J. M. A. Grime, P. Sengupta, A. K. Chen, A. E. P. Durumeric, A. Srivastava, M. Yeager, J. A. G. Briggs, J. Lippincott-Schwartz, and G. A. Voth, “Immature hiv-1 lattice assembly dynamics are regulated by scaffolding from nucleic acid and the plasma membrane,” *Proceedings of the National Academy of Sciences*, vol. 114, no. 47, pp. E10056–E10065, 2017.
- [32] W. Sundquist and H.-G. Kräusslich, “Hiv-1 assembly, budding, and maturation,” *Cold Spring Harbor perspectives in medicine*, vol. 2, p. a006924, 07 2012.
- [33] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne, “The Protein Data Bank,” *Nucleic Acids Research*, vol. 28, pp. 235–242, 01 2000. (<http://www.rcsb.org>).
- [34] B. Brügger, B. Glass, P. Haberkant, I. Leibrecht, F. T. Wieland, and H.-G. Kräusslich, “The hiv lipidome: A raft with an unusual composition,” *Proceedings of the National Academy of Sciences*, vol. 103, no. 8, pp. 2641–2646, 2006.
- [35] R. Chan, P. D. Uchil, J. Jin, G. Shui, D. E. Ott, W. Mothes, and M. R. Wenk, “Retroviruses human immunodeficiency virus and murine leukemia virus are enriched in phosphoinositides,” *Journal of virology*, vol. 88, no. 22, p. 11228–11238, 2008.
- [36] J. S. Saad, J. Miller, J. Tai, A. Kim, R. H. Ghanam, and M. F. Summers, “Structural basis for targeting hiv-1 gag proteins to the plasma membrane for virus assembly,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 30, p. 11364–11369, 2006.

- [37] A. Ono and E. O. Freed, “Plasma membrane rafts play a critical role in hiv-1 assembly and release,” *Proceedings of the National Academy of Sciences*, vol. 98, no. 24, pp. 13925–13930, 2001.
- [38] A. Alfadhli, A. Still, and E. Barklis, “Analysis of human immunodeficiency virus type 1 matrix binding to membranes and nucleic acids,” *Journal of virology*, vol. 83, no. 23, p. 12196–12203, 2009.
- [39] P. Provitera, R. El-Maghrabi, and S. Scarlata, “The effect of hiv-1 gag myristoylation on membrane binding,” *Biophysical Chemistry*, vol. 119, no. 1, pp. 23–32, 2006.
- [40] A. Alfadhli, R. L. Barklis, and E. Barklis, “Hiv-1 matrix organizes as a hexamer of trimers on membranes containing phosphatidylinositol-(4,5)-bisphosphate,” *Virology*, vol. 387, no. 2, pp. 466–472, 2009.
- [41] L.-A. Carlson, A. de Marco, H. Oberwinkler, A. Habermann, J. Briggs, H.-G. Kräusslich, and K. Grünewald, “Cryo electron tomography of native hiv-1 budding sites,” *PLoS pathogens*, vol. 6, p. e1001173, 11 2010.
- [42] E. Wright, J. Schooler, H. Ding, C. Kieffer, C. Fillmore, W. Sundquist, and G. Jensen, “Electron crytomography of immature hiv-1 virions reveals the structure of the ca and sp1 gag shells,” *The EMBO journal*, vol. 26, pp. 2218–26, 04 2007.
- [43] C. P. Hill, D. Worthylake, A. M. Bancroft, D. P. and Christensen, and W. I. Sundquist, “Crystal structures of the trimeric human immunodeficiency virus type 1 matrix protein: implications for membrane association and assembly,” *Proceeding of the National Academy of Sciences of the USA*, vol. 93, p. 3099–3104, Apr 1996.
- [44] J. Briggs, J. Riches, B. Glass, V. Lux, G. Zanetti, and H.-G. Kräusslich, “Structure and assembly of immature hiv,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, pp. 11090–5, 07 2009.
- [45] J. R. Lingappa, J. C. Reed, M. Tanaka, K. Chutiraka, and B. A. Robinson, “How hiv-1 gag assembles in cells: Putting together pieces of the puzzle,” *Virus research*, vol. 193, pp. 89–107, 2014.

- [46] L.-A. Carlson, Y. Bai, S. C. Keane, J. A. Doudna, and J. H. Hurley, “Reconstitution of selective hiv-1 rna packaging in vitro by membrane-bound gag assemblies,” *eLife*, vol. 5, p. e14663, jun 2016.
- [47] S. J. Risselada, H. J. and. Marrink, “The molecular face of lipid rafts in model membranes,” *Proceeding of the National Academy of Sciences of the USA*, vol. 105, no. 45, pp. 17367–17372, 2008.
- [48] I. R. Cooke and M. Deserno, “Solvent-free model for self-assembling fluid bilayer membranes: Stabilization of the fluid phase based on broad attractive tail potentials,” *The Journal of Chemical Physics*, vol. 123, no. 22, pp. 224710–224710–13, 2005.
- [49] R. Matthews and C. N. Likos, “Influence of fluctuating membranes on self-assembly of patchy colloids,” *Phys. Rev. Lett.*, vol. 109, p. 178302, Oct 2012.
- [50] R. Matthews and C. N. Likos, “Dynamics of self-assembly of model viral capsids in the presence of a fluctuating membrane,” *The journal of physical chemistry. B*, vol. 117, p. 8283–8292, Jun 2013.
- [51] T. Ruiz-Herrero, E. Velasco, and M. F. Hagan, “Mechanisms of budding of nanoscale particles through lipid bilayers,” *The Journal of Physical Chemistry B*, vol. 116, no. 32, pp. 9595–9603, 2012.
- [52] T. Ruiz-Herrero and M. F. Hagan, “Simulations show that virus assembly and budding are facilitated by membrane microdomains,” *Biophysical Journal*, vol. 108, no. 3, pp. 585–595, 2015.
- [53] Q. Chen, E. Diesel, J. K. Whitmer, S. C. Bae, E. Luijten, and S. Granick, “Tri-block colloids for directed self-assembly,” *Journal of the American Chemical Society*, vol. 133, no. 20, pp. 7725–7727, 2011.
- [54] Q. Chen, J. Yan, J. Zhang, S. C. Bae, and S. Granick, “Janus and multiblock colloidal particles,” *American Chemical Society*, vol. 28, pp. 13555–13561, Sep 2012.
- [55] F. Romano and F. Sciortino, “Two dimensional assembly of triblock janus particles into crystal phases in the two bond per patch limit,” *Soft Matter*, vol. 7, pp. 5799–5804, 2011.

- [56] N. Kern and D. Frenkel, “Fluid–fluid coexistence in colloidal systems with short-ranged strongly directional attraction,” *Journal of Chemical Physics*, vol. 118, p. 9882, 2003.
- [57] G. A. Chapela, O. Guzmán, J. A. Martínez-González, P. Díaz-Leyvaa, and J. Quintana-H, “Self-assembly of kagome lattices, entangled webs and linear fibers with vibrating patchy particles in two dimensions.,” *Soft Matter*, vol. 10, no. 45, pp. 9167–9176, 2014.
- [58] E. Edlund, O. Lindgren, and M. Jacobi, “Designing isotropic interactions for self-assembly of complex lattices,” *Physical review letters*, vol. 107, p. 085503, 08 2011.
- [59] E. Edlund, O. Lindgren, and M. Jacobi, “Using the uncertainty principle to design simple interactions for targeted self-assembly,” *The Journal of chemical physics*, vol. 139, p. 024107, 07 2013.
- [60] B. Machorro-Martínez, O. Guzmán, and G. Chapela, “Phase diagrams of extended and deformed kagome lattices,” *Physica A: Statistical Mechanics and its Applications*, vol. 585, p. 126397, 09 2021.
- [61] D. Salgado-Blanco and C. Mendoza, “Non-additive simple potentials for pre-programmed self-assembly,” *Soft Matter*, vol. 11, pp. 889–897, 09 2014.
- [62] A. Jain, J. R. Errington, and T. M. Truskett, “Dimensionality and design of isotropic interactions that stabilize honeycomb, square, simple cubic, and diamond lattices.,” *Physical Review X*, vol. 4, no. 3, p. 031049, 2014.
- [63] W. Piñeros, M. Baldea, and T. Truskett, “Designing convex repulsive pair potentials that favor assembly of kagome and snub square lattices,” *The Journal of Chemical Physics*, vol. 145, p. 054901, 07 2016.
- [64] B. Lindquist, R. Jadrich, and T. Truskett, “Inverse design for self assembly via on-the-fly optimization,” *The Journal of Chemical Physics*, vol. 145, p. 111101, 09 2016.
- [65] Z.-W. Li, Y.-L. Zhu, Z.-Y. Lu, and Z.-Y. Sun, “A versatile model for soft patchy particles with various patch arrangements,” *Soft Matter*, vol. 12, pp. 741–749, 2016.

- [66] M. Johnson, “Modeling the self-assembly of protein complexes through a rigid-body rotational reaction-diffusion algorithm,” *The Journal of Physical Chemistry B*, vol. 122, pp. 11771–11783, 09 2018.
- [67] D. L. D. Caspar and A. Klug, “Physical principles in the construction of regular viruses,” *Cold Spring Harbor Symposia on Quantitative Biology*, vol. 27, pp. 1–24, 1962.
- [68] M. Goldberg, “A class of multi-symmetric polyhedra,” *Tohoku Mathematical Journal, First Series*, vol. 43, pp. 104–108, 1937.
- [69] D. C. Rapaport, J. E. Johnson, and J. Skolnick, “Supramolecular self-assembly: molecular dynamics modeling of polyhedral shell formation,” *Computer Physics Communications*, vol. 121, pp. 231–235, 1999.
- [70] D. C. Rapaport, “Studies of reversible capsid shell growth,” *Journal of Physics: Condensed Matter*, vol. 22, no. 10, pp. 104–115, 2010.
- [71] D. C. Rapaport, “Modeling capsid self-assembly: design and analysis,” *Physical Biology*, vol. 7, p. 045001, dec 2010.
- [72] I. G. Johnston, A. A. Louis, and J. P. K. Doye, “Modelling the self-assembly of virus capsids,” *Journal of Physics: Condensed Matter*, vol. 22, p. 104101, feb 2010.
- [73] H. Nguyen and C. Brooks, “Generalized structural polymorphism in self-assembled viral particles,” *Nano letters*, vol. 8, pp. 4574–81, 01 2009.
- [74] R. Kerner, “Classification and evolutionary trends of icosahedral viral capsids,” *Computational and Mathematical Methods in Medicine*, vol. 9, pp. 175–181, 2008.
- [75] J. Baschek, H. Klein, and U. Schwarz, “Stochastic dynamics of virus capsid formation: Direct versus hierarchical self-assembly,” *BMC biophysics*, vol. 5, p. 22, 12 2012.
- [76] A. B. Gutiérrez, B. I. Machorro-Martínez, J. Quintana, J. C. Armas-Pérez, P. Mendoza, J. M. E. Lucero, and G. A. Chapela, “Hiv-1 immature virion and other networks formation with simple patchy disks,” *Molecular Physics*, vol. 120, no. 22, p. e2129759, 2022.



- [77] D. Frenkel and B. Smit, *Understanding Molecular Simulations: From Algorithms to Applications*. Academic Press, 2 ed., 2002.
- [78] D. A. McQuarrie, *Statistical Mechanics*. University Science Books, 2 ed., 2000.
- [79] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, “Equation of State Calculations by Fast Computing Machines,” *Journal of Chemical Physics*, vol. 21, no. 6, pp. 1087–1092, 1953.
- [80] D. C. Rapaport, *The Art of Molecular Dynamics Simulation*. Cambridge University Press, 2004.
- [81] H. C. Andersen, “Molecular dynamics simulations at constant pressure and/or temperature,” *The Journal of Chemical Physics*, vol. 72, no. 4, pp. 2384–2393, 1980.
- [82] G. Martyna, M. Klein, and M. Tuckerman, “Nosé-hoover chains: The canonical ensemble via continuous dynamics,” *The Journal of Chemical Physics*, vol. 97, pp. 2635–2643, 08 1992.
- [83] B. J. Alder and T. E. Wainwright, “Studies in molecular dynamics. i. general method,” *The Journal of Chemical Physics*, vol. 31, no. 2, pp. 459–466, 1959.
- [84] G. A. Chapela, S. E. Martínez-Casas, and J. Alejandre, “Molecular dynamics for discontinuous potentials: I. general method and simulation of hard polyatomic molecules,” *Molecular Physics*, vol. 53, no. 1, pp. 139–159, 1984.
- [85] J. Anderson, C. Lorenz, and A. Travesset, “General purpose molecular dynamics simulations fully implemented on graphics processing units,” *Journal of Computational Physics*, vol. 227, pp. 5342–5359, 05 2008.
- [86] J. Glaser, T. Nguyen, J. Anderson, P. Lui, F. Spiga, J. Millan, D. C. Morse, and S. Glotzer, “Strong scaling of general-purpose molecular dynamics simulations on gpus,” *Computer Physics Communications*, vol. 192, 12 2014.
- [87] F. K. M. Schur, M. Obr, W. J. H. Hagen, W. Wan, A. J. Jakobi, J. M. Kirkpatrick, C. Sachse, H.-G. Kräusslich, and J. A. G. Briggs, “An atomic model of hiv-1 capsid-sp1 reveals structures regulating assembly and maturation,” *Science (New York, N.Y.)*, vol. 353, p. 506—508, July 2016.

- [88] A. Tan, A. Pak, D. Morado, G. Voth, and J. Briggs, “Immature hiv-1 assembles from gag dimers leaving partial hexamers at lattice edges as potential substrates for proteolytic maturation,” *Proceedings of the National Academy of Sciences*, vol. 118, p. e2020054118, 01 2021.
- [89] E. Krissinel and K. Henrick, “Inference of macromolecular assemblies from crystalline state,” *Journal of Molecular Biology*, vol. 372, no. 3, pp. 774–797, 2007. ‘Protein interfaces, surfaces and assemblies’ service PISA at the European Bioinformatics Institute. ([http://www.ebi.ac.uk/pdbe/prot\\_int/pistart.html](http://www.ebi.ac.uk/pdbe/prot_int/pistart.html)).
- [90] B. Chen and R. Tycko, “Simulated self-assembly of the hiv-1 capsid: Protein shape and native contacts are sufficient for two-dimensional lattice formation,” *Biophysical Journal*, vol. 100, no. 12, pp. 3035–3044, 2011.
- [91] X. Qiao, J. Jeon, J. Weber, F. Zhu, and B. Chen, “Mechanism of polymorphism and curvature of hiv capsid assemblies probed by 3d simulations with a novel coarse grain model,” *Biochimica et Biophysica Acta*, vol. 1850, no. 11, pp. 2353–2367, 2015.
- [92] L. Liljas, “Virus assembly,” *Current opinion in structural biology*, vol. 9, pp. 129–34, 03 1999.





Casa abierta al tiempo  
UNIVERSIDAD AUTÓNOMA METROPOLITANA

ACTA DE DISERTACIÓN PÚBLICA

No. 00056  
Matrícula: 2152801022

Modelo del liberación de VIH-1 de la membrana plasmática, autoensamblaje de su red y de cápsides icosaédricas.

En la Ciudad de México, se presentaron a las 10:00 horas del día 6 del mes de marzo del año 2023 en la Unidad Iztapalapa de la Universidad Autónoma Metropolitana, los suscritos miembros del jurado:

DR. RAFAEL ARTURO ZUBILLAGA LUNA  
DR. JESUS ENRIQUE DIAZ HERRERA  
DR. JOSE ELIAS PEREZ LOPEZ  
DR. ALEJANDRO GIL VILLEGAS MONTIEL  
DR. GUSTAVO ADOLFO CHAPELA CASTAÑARES



BRIAN IGNACIO MACHORRO MARTINEZ  
ALUMNO

Bajo la Presidencia del primero y con carácter de Secretario el último, se reunieron a la presentación de la Disertación Pública cuya denominación aparece al margen, para la obtención del grado de:

DOCTOR EN CIENCIAS (FISICA)  
DE: BRIAN IGNACIO MACHORRO MARTINEZ

y de acuerdo con el artículo 78 fracción IV del Reglamento de Estudios Superiores de la Universidad Autónoma Metropolitana, los miembros del jurado resolvieron:

**APROBAR**

REVISÓ  
MTRA. ROSALIA SERRANO DE LA PAZ  
DIRECTORA DE SISTEMAS ESCOLARES

Acto continuo, el presidente del jurado comunicó al interesado el resultado de la evaluación y, en caso aprobatorio, le fue tomada la protesta.

DIRECTOR DE LA DIVISIÓN DE CBI  
DR. ROMAN LINARES ROMERO

PRESIDENTE  
DR. RAFAEL ARTURO ZUBILLAGA LUNA

VOCAL  
DR. JESUS ENRIQUE DIAZ HERRERA

VOCAL  
DR. JOSE ELIAS PEREZ LOPEZ

VOCAL  
DR. ALEJANDRO GIL VILLEGAS MONTIEL

SECRETARIO  
DR. GUSTAVO ADOLFO CHAPELA CASTAÑARES