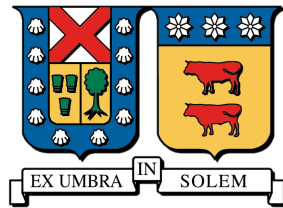


UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA
DEPARTAMENTO DE INGENIERÍA MECÁNICA
VALPARAÍSO - CHILE



Parametrización de un Modelo de Grano Grueso para un Complejo ARN-Proteína Utilizando APBS

GASPAR IGNACIO VUCHETICH DE CHENEY VALDERAS

MEMORIA DE TITULACIÓN PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL MECÁNICO

Profesor Guía: PhD. Christopher Cooper Villagrán
Profesor Correferente: PhD. Simón Poblete Fuentes

Agosto - 2026



CONSTANCIA DE VALIDACIÓN Y CONFIDENCIALIDAD DE MONOGRAFÍA A REPOSITORIO ACADÉMICO

1.- IDENTIFICACIÓN DEL TRABAJO ACADÉMICO

Tipo de monografía (marcar una opción): ☒ Memoria o trabajo de título ☐ Tesis de Postgrado

Título del trabajo: Parametrización de un modelo de grano grueso para un complejo ARN-Proteína utilizando APBS

Nombre del candidato(a): Gaspar Ignacio Vuchetich de Cheney Valderas

Carrera / Grado: Ingeniería Civil Mecánica

Campus: Casa Central Departamento: Departamento de Ingeniería Mecánica

2.- VALIDACIÓN DEL PROFESOR GUÍA/DIRECTOR DE TESIS

Yo Christopher Cooper Villagrán, en mi calidad de profesor(a) guía/director(a) del trabajo académico mencionado anteriormente **DEJO CONSTANCIA** que:

- He revisado esta versión del documento y corresponde a la versión final aprobada del trabajo.
- El trabajo cumple con los requisitos académicos y de formato establecidos por la institución.

3.- EVALUACIÓN DE CONFIDENCIALIDAD POR PROPIEDAD INDUSTRIAL (marcar una opción)

☒ El trabajo **NO contiene** información que amerite confidencialidad y puede ser publicado de inmediato en repositorio con acceso abierto.

☐ El trabajo **CONTIENE** información con potenciales implicancias de propiedad industrial o intelectual y requiere un periodo de confidencialidad (**embargo**) por (**marcar una opción**):

☐ 6 meses ☐ 12 meses ☐ 2 años ☐ 3 años ☐ 5 años ☐ 10 años

Fundamentación de la necesidad de confidencialidad (obligatorio si se solicita embargo):

4.- FIRMAS

Profesor(a) guía o director(a) de memoria o tesis:

Fecha: 17-08-2026

Firma: _____

Estudiante o Candidato(a):

Fecha: 17-08-2026

Firma: _____

Este formulario debe ser insertado como página 2 de la memoria o tesis, completado y firmado por estudiante y profesor(a) antes de la entrega en portal PRISMA de Biblioteca USM.

Resumen

En el presente trabajo se abordó en primer lugar la parametrización de un modelo de grano grueso de ARN viral mediante simulaciones electrostáticas, con el cual posteriormente se modeló el empaquetado del genoma del *virus de la bursitis infecciosa* al interior de su cápside mediante simulaciones de dinámica molecular. Se observó que un potencial de Yukawa junto al modelo de grano grueso empleado representa fielmente la interacción electrostática entre hebras de ARN hasta la distancia de Debye independiente de la orientación relativa, validado por las simulaciones electrostáticas. Sin embargo, se debe tener en consideración APBS no contempla la deformación que pueden sufrir las proteínas al interactuar la cual puede ser especialmente relevante, puesto que por dicha razón se debió excluir del modelo de grano grueso una de ellas. Por otra parte, se realizaron simulaciones de dinámica para el empaquetamiento del ARN del *virus de la bursitis infecciosa* utilizando el modelo parametrizado para diferentes combinaciones de hebras de ARN. Donde se observó una leve replicación del estudio mediante Cryo-EM de densidad radial al interior del virión, además de establecer que la electrostática del sistema a la escala del virus admite la presencia de la proteína VP3. Por otra parte, para 1 hebra su dimerización podría verse dificultada producto de la distancia promedio entre monómeros. Finalmente, se observó que las distancias entre pares de monómeros de una misma cadena no presentan diferencias frente a las distancias inter-cadenas, lo que indica que la electrostática no generó preferencias en cuanto a la dimerización de VP3 entre segmentos de la misma cadena u otras cadenas.

Índice

1	Introducción	9
1.1	Objetivos	9
2	Teoría	11
2.1	Virología	11
2.1.1	Estructura del ARN	11
2.1.2	Estructura Viral	12
2.1.3	Enfermedad de Bursitis Infecciosa	14
2.1.4	Complejo ARN-VP3	15
2.2	Modelos de Grano Grueso	16
2.3	Modelos de Solvente Implícito	17
2.4	Ecuación de Poisson-Boltzmann	18
2.4.1	Potencial de Yukawa o de Coulomb Apantallado	19
2.4.2	Solución Numérica de la Ecuación de Poisson-Boltzmann	20
2.5	Cálculos de Energía mediante la ecuación de Poisson-Boltzmann	21
2.5.1	Energía de Solvatación	22
2.5.2	Energía de Interacción	22
3	Metodología	25
3.1	Implementación de APBS	25
3.1.1	Configuración de APBS	26
3.1.2	Parámetros de APBS Utilizados	27
3.1.3	Simulaciones Aisladas	29
3.2	Gestor de Simulaciones	30
3.3	Implementación de LAMMPS	31
3.3.1	Sistema de Unidades Utilizado	32
3.3.2	Simulación de Inicio (<i>Warm-Up</i>)	33
3.3.3	Simulación de Compresión (<i>Production</i>)	33
3.3.4	Método de Análisis	34
3.4	Modelo de Grano Grueso	36
3.5	Parámetros Utilizados y Métricas de Error	38
3.5.1	Parámetros APBS y Yukawa	38
3.5.2	Parámetros LAMMPS	39
4	Resultados & Análisis	40
4.1	Parametrización	40
4.1.1	Estudio de Malla	40
4.1.2	Simulación Terminal-N VP3	43
4.1.3	Exploración de Alternativas de Grano Grueso	44
4.1.4	Estudio del Potencial de Yukawa	47
4.2	Simulaciones de Compresión de ARN en LAMMPS	54
4.2.1	Compresión de Una Hebra	55
4.2.2	Compresión de Dos Hebras Paralelas	60
4.2.3	Compresión de Dos Hebras Cruzadas	67
5	Conclusiones	70

Índice de Figuras

1	Estructura de hélice simple de ARN versus doble hélice de ADN con sus respectivas bases nitrogenadas. Extraído de Sponk (2010).	11
2	Estructura de un nucleótido de ADN, idéntico en estructura a un nucleótido de ARN con excepción de la timina en lugar de uracilo. Extraído de Alberts (2022).	12
3	Ejemplos de virus de ARN con su respectiva cápside y material genético. Extraído de Alberts (2022).	13
4	Ciclo de vida de un virus de ADN. Extraído de Alberts (2022).	13
5	Ejemplo de dímero de VP3 ayudando a empaquetar dos hebras de ARN. Elaboración propia.	15
6	Diferentes ejemplos de modelos de grano grueso para un <i>tetraloop</i> de ARN, donde A) indica la estructura <i>all-atom</i> y B) a G) indican diversas maneras de simplificar la estructura con grano grueso. Extraído de Sponer et al. (2018)	16
7	Modelos de solvente implícito y explícito. Extraído de Emmett M. Leddin (2021)	17
8	Esquema del método de diferencias finitas para $u = u(x, y)$ con $u \in \Omega$ en una grilla regular de tamaño Δ . Elaboración propia.	20
9	Diagrama del ciclo termodinámico para la energía de interacción de dos moléculas solvatadas. Extraído de Baker (2023b)	21
10	Diagrama de flujo típico para la utilización de APBS en conjunto con PDB2PQR. Extraído de Jurrus et al. (2018)	26
11	Ejemplo de configuración de APBS utilizado en el presente trabajo. (a) primera sección ELEC del archivo de configuración junto a sección READ, (b) segunda sección ELEC del archivo junto a sección PRINT. Elaboración propia.	28
12	Ejemplo de archivo pqr para una simulación aislada. Notar que los átomos pertenecientes a la biomolécula apagada poseen carga y radio cero. Elaboración propia.	30
13	Ejemplo de configuración de LAMMPS donde se indica la separación entre las fases de compresión y relajación para un caso con distinto paso temporal. Elaboración propia.	32
14	Ejemplo de potencial de Yukawa introducido en LAMMPS para una simulación de compresión. Elaboración propia.	34
15	Esquema del modelo de grano grueso empleado respecto a los extremos 3' y 5' de la doble hélice de ARN. Elaboración propia.	36
16	Ejemplo de archivo pqr con un modelo de grano grueso, indicado por las ultimas filas de este.	37
17	Diferencia entre casos <i>FullGeometry</i> (a) y <i>SimpleGeometry</i> (b) respecto a la superficie de exclusión de solvente generada por APBS. Elaboración propia.	37
18	Sistema coordenado definido para el sistema a parametrizar. Elaboración propia.	38
19	Visualización del caso empleado para el estudio de malla.	40
20	Resultados del estudio de malla con extrapolación de Richardson para APBS lineal y no-lineal en rojo.	42
21	Dímero de VP3 (a) y Terminal-N (b) de la proteína simulado. Cabe destacar el entrelazamiento entre proteínas.	43
22	Visualización del caso de estudio utilizado para exploración de alternativas de grano grueso. Elaboración propia.	44
23	Resultados para los modelos <i>all-atom</i> calculados: Yukawa y APBS Lineal/No-Lineal.	45
24	Comparación de resultados de APBS para <i>Coarse Grain</i> Yukawa y APBS.	46
25	Resultados para tipos de <i>Coarse Grain</i> en APBS.	47

26	Visualización de las orientaciones “P” (a) y “N” (b).	47
27	Caso cruzado “many-to-many”, dos hebras de ARN de igual longitud (a) y “one-to-many”, una hebra de ARN corta y otra larga (b). Elaboración propia.	48
28	Energía de interacción normalizada para el caso 1-8 monómeros.	49
29	Diferencia Yukawa-APBS normalizada para el caso 1-8 monómeros.	50
30	<i>Boxplot</i> de la diferencia total absoluta (a) y normalizada (b) sobre todas las rotaciones ω y ambas orientaciones para el caso 1-N monómeros.	51
31	Energía de interacción absoluta para el caso 8-8 monómeros en orientación P. . .	52
32	Diferencia absoluta entre las energías de interacción calculadas por Yukawa y APBS para el caso 8-8 monómeros.	53
33	<i>Boxplot</i> de la diferencia total absoluta (a) y normalizada (b) sobre todas las rotaciones ω y ambas orientaciones para el caso N-N monómeros.	54
35	Renderización de la simulación individual en sus fases de compresión (a) y final de relajación (b).	55
34	Secuencia de compresión para el caso individual. Elaboración propia.	56
36	Densidad radial para el inicio, final de compresión, paso intermedio de relajación y final de etapa de relajación para el caso individual.	56
37	Histograma de distancia entre pares de monómeros excluyendo los 4 vecinos inmediatos para el caso individual.	57
38	Histograma de distancia entre pares lejanos de monómeros, a contar del vecino numero 20 para caso individual.	58
39	Histograma de ángulos para el caso individual (c) con ejemplos de segmentos de 180° (a) y 150° (b).	59
40	Densidad radial para el inicio, final de compresión, paso intermedio de relajación y final de etapa de relajación para el caso individual.	60
41	Secuencia de compresión para el caso doble en paralelo.	61
42	Densidad radial para el inicio, final de compresión, paso intermedio de relajación y final de etapa de relajación para el caso doble en paralelo.	62
43	Histograma de ángulos para el caso doble en paralelo (c) con ejemplos de segmentos de 180° (a) y 150° (b).	63
44	Histograma de distancia entre pares de monómeros excluyendo los 4 vecinos inmediatos para el caso doble en paralelo.	64
45	Histograma de distancia entre pares lejanos de monómeros, a contar del vecino numero 20 para caso doble paralelo.	65
46	Matriz de pares de distancias entre monómeros de una misma cadena para el caso doble paralelo.	66
47	Secuencia de compresión para el caso doble en cruz. Elaboración propia.	67
48	Renderización de la simulación doble cruzada en sus fases de compresión (a) y final de relajación. (b).	68
49	Densidad radial para el inicio, final de compresión, paso intermedio de relajación y final de etapa de relajación para el caso doble cruzado.	68
50	Matriz de pares de distancias entre monómeros de una misma cadena para el caso doble cruzado.	69
51	Energía de interacción para el caso 1-1 orientación P.	77
52	Energía de interacción para el caso 1-1 orientación N.	77
53	Diferencia Yukawa-APBS para el caso 1-1.	78
54	Energía de interacción para el caso 1-2 orientación P.	78
55	Energía de interacción para el caso 1-2 orientación N.	79
56	Diferencia Yukawa-APBS para el caso 1-2.	79

57	Energía de interacción para el caso 1-4 orientación P.	80
58	Energía de interacción para el caso 1-4 orientación N.	80
59	Diferencia Yukawa-APBS para el caso 1-4.	81
60	Energía de interacción para el caso 1-8 orientación N.	81
61	Energía de interacción para el caso 4-4 orientación P.	82
62	Energía de interacción para el caso 4-4 orientación N.	82
63	Diferencia Yukawa-APBS para el caso 4-4.	83
64	Energía de interacción para el caso 8-8 orientación N.	83
65	Histograma de ángulos para el caso doble cruzado (c) con ejemplos de segmentos de 180° (a) y 150° (b).	84
66	Histograma de distancia entre pares de monómeros excluyendo los 4 vecinos inmediatos para el caso doble cruzado.	85
67	Histograma de distancia entre pares lejanos de monómeros, a contar del vecino numero 20 para caso doble cruzado.	85

Índice de Cuadros

1	Resumen de cálculo de factor de conversión de unidades $k_B T$	33
2	Resumen de parámetros de grano grueso por esfera.	36
3	Ejemplo de cargas de grano grueso de ARN.	37
4	Parámetros del solvente.	38
5	Otros parámetros de APBS.	39
6	Opciones del control de simulación de LAMMPS.	39
7	Secuencia de mallas escogida con $r = 5/3$	41
8	Orden de convergencia observado para la secuencia $\Delta = \{0, 500; 0, 387; 0, 300\}$. . .	41
9	Resultados del estudio de malla con error relativo respecto a la extrapolación de Richardson correspondiente.	42
10	Casos simulados para el estudio del potencial de Yukawa. Donde ω representa una rotación en el eje mayor de la hebra móvil como se describió en la subsection 3.5	48
11	Coeficientes de Yukawa utilizados en LAMMPS	54
12	Parámetros utilizados para simulación de compresión para 1 hebra.	55
13	Parámetros utilizados para simulación de compresión para 2 hebras paralelas. . .	61
14	Parámetros utilizados para simulación de compresión para 2 hebras cruzadas. . .	67

Glosario

IBDV: Virus de la bursitis infecciosa o virus de la enfermedad de Gumboro.

Hebra de ARN: Cadena de ARN de doble hélice conformada por dos hebras antiparalelas. Un virión puede contener múltiples cadenas en su interior.

Monómero: Corresponde a la unidad base del modelo de grano grueso parametrizado equivalente a un segmento de ARN de 5 pares de bases nitrogenadas, simplificados en una esfera cargada eléctricamente.

ARN: Ácido desoxirribonucleico. Molécula esencial para la vida que codifica información genética.

Modelo de Grano Grueso: Modelo simplificado de una estructura molecular mayor mediante elementos simples que representan grupos de átomos los cuales interactúan mediante potenciales.

A mi familia, a mis amigos, a mis profesores y a todos quienes hicieron posible este gran logro después de tan largo camino. Gracias totales por todo este tiempo!

"The more clearly we can focus our attention on the wonders and realities of the universe about us, the less taste we shall have for destruction."

— Rachel Carlson

1 Introducción

La infección viral es, fundamentalmente, un fenómeno bioquímico que abarca múltiples escalas físicas, tanto espaciales como temporales; la conformación de macromoléculas al interior de una célula puede tomar ordenes de magnitud más tiempo que el plegamiento individual de las proteínas que las componen. Por otra parte, debido a la presencia de solvente, la interacción electrostática ocurre en el orden de las decenas de ángstroms, mientras que la escala que la gobierna ocurre en el orden de unos cuantos ángstroms debido a los átomos de la molécula y su interacción con el solvente. Debido a esto, producto de limitaciones computacionales, su estudio se ve forzado a separar las escalas, entre los fenómenos que ocurren a escala atómica y los que ocurren a escala de proteína o más, aunque físicamente estén relacionados. Los primeros pueden ser estudiados a través simulaciones de dinámica molecular (MD) —inviable para sistemas grandes— mientras que los segundos mediante modelos de grano grueso —simplificando la interacción atómica— o mediante simulaciones electrostáticas de medio continuo —las cuales son rígidas e incapaces de simular el comportamiento dinámico al interior de un virión. El presente trabajo, por lo tanto, busca reducir esta brecha mediante una parametrización de un modelo de grano grueso mediante simulaciones electrostáticas detalladas, a diferencia de resultados experimentales como suele realizarse en la práctica. De esta manera, busca introducir mediante simulaciones la información de la interacción atómica al modelo de grano grueso. Con esto en mente, el marco de estudio contempla el virus de la enfermedad de bursitis infecciosa, o IBDV, el cual produce una enfermedad homónima en las aves de corral que genera grandes mermas a la producción global y a la calidad de vida de las aves al ser altamente infecciosa y letal (Orakpoghenor et al., 2020). Este corresponde a un virus de doble-hélice de ARN, donde la totalidad de su material genético se encuentra en dos cadenas, sin embargo, se ha observado que las cepas que contienen una copia adicional —es decir 4 cadenas de ARN de doble hebra— resultan más virulentas por razones que son de activo estudio. Como la gran mayoría de los virus, su ciclo de reproducción contempla una etapa de empaquetado, donde el genoma replicado al interior de la célula capturada es introducido en las nuevas capsides virales. Es en esta etapa donde se hipotetiza que el fenómeno descrito con anterioridad respecto a la cantidad de cadenas de ARN en las distintas cepas puede tener su origen. El presente trabajo, por lo tanto, busca estudiar algo acerca del comportamiento del ARN de este virus al empaquetarse al interior de la cápside mediante simulaciones dinámicas con el modelo parametrizado para diferentes cantidades de cadenas.

1.1 Objetivos

A continuación, se indican el objetivo principal y los objetivos específicos que enmarcan el presente trabajo.

Objetivo Principal

Generar una parametrización de un modelo de grano grueso de genoma viral y eventuales proteínas mediante el uso de simulaciones *all-atom* en APBS, incluyendo sus limitaciones, y ponerlo a prueba para un caso biológico.

Objetivo Específicos

1. Optimizar parámetros de cálculo, como mallado y refinamiento, en la resolución numérica de la ecuación de Poisson-Boltzmann a través del *software* APBS.

2. Simular la energía de unión electrostática para el complejo indicado, así como también para sus partes en separado.
3. Parametrizar las cargas de un modelo de grano grueso tal que el comportamiento electrostático resultante se asemeje a la simulación no-simplificada (*all-atom*).

2 Teoría

2.1 Virología

La virología compete el estudio científico de la biología de los virus y enfermedades virales, incluyendo su distribución, bioquímica, fisiología, biología molecular, ecología, evolución y aspectos clínicos de ambos (Nature, 2026). Por consiguiente, a modo de motivación y necesario para entender a cabalidad el presente trabajo, resulta pertinente abordar algunas generalidades de la virología y la estructura viral.

2.1.1 Estructura del ARN

El ARN, o *ácido ribonucleico*, es uno de los bloques fundamentales de la vida en la tierra. Cumple diversas funciones en los organismos relacionadas con la síntesis de proteínas, desde el almacenamiento de información genética como también su transmisión entre las diferentes maquinarias celulares en el interior de esta. Por esta misma razón, es que el ARN cumple un importante rol en los virus puesto que, almacena su material genético o participa directamente en el proceso de replicación.

El ARN se compone de dos piezas fundamentales: las bases nitrogenadas y la columna de Azúcar-Fosfato. Las bases nitrogenadas del ARN son denominadas citosina, guanina, adenina y uracilo y estas codifican la información genética para la producción de proteínas según sea su ordenamiento de forma similar a un código de barras. Estas bases nitrogenadas se conectan a la columna de azúcar-fosfato donde el conjunto formado por el azúcar-fosfato y la base se denomina *nucleótido*. De tal manera que una columna de azúcar-fosfato es capaz de contener cualquiera de las bases nitrogenadas para codificar un mensaje específico en su secuencia (Alberts, 2022).

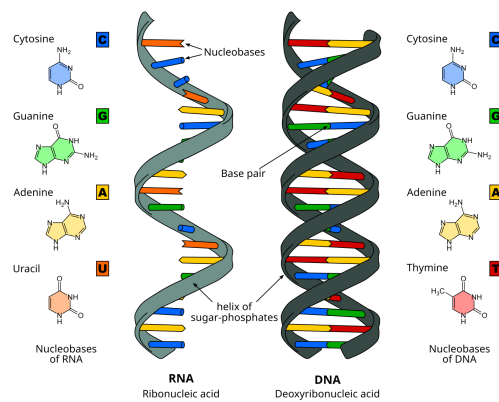


Figura 1: Estructura de hélice simple de ARN versus doble hélice de ADN con sus respectivas bases nitrogenadas. Extraído de Sponk (2010).

Cada nucleótido se enlaza al siguiente mediante el fosfato de la columna, el cual enlaza con el azúcar del siguiente nucleótido en la cadena. En otras palabras, en la columna, el fosfato actúa como una clavija que se acopla a un agujero formado por el azúcar. De esta manera, como se indica en la Figura 2 el ARN puede formar cadenas de longitud variable codificando un mensaje específico en sus bases nitrogenadas. El extremo considerado como el inicio de la cadena es el denominado 5' y el final como 3'. Estos fosfatos que forman parte de la columna vertebral tanto del ADN como del ARN son quienes llevan una gran carga negativa —equivalente a $-1e^-$ — lo

que produce que el ARN sea una molécula altamente cargada, el cual es un factor importante para tomar en cuenta en su modelamiento electrostático como se verá en la Subsección 2.4.

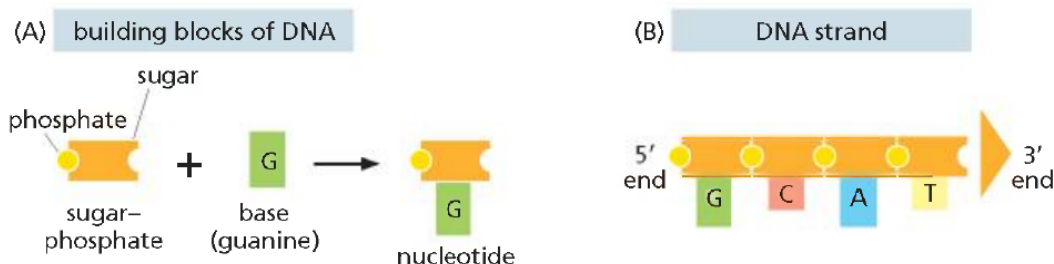


Figura 2: Estructura de un nucleótido de ADN, idéntico en estructura a un nucleótido de ARN con excepción de la timina en lugar de uracilo. Extraído de Alberts (2022).

Cabe destacar que el extremo 5' contiene un fosfato adicional que no está presente en el extremo 3', como se visualiza en la figura. Esto resultará de especial importancia a la hora de plantear el modelo de grano grueso, puesto que los extremos deberán ser tratados con especial cuidado. Finalmente, tanto el ARN como el ADN pueden formar una estructura de doble hélice como se indica en la Figura 1 mediante puentes de hidrógeno entre bases nitrogenadas complementarias, —por ejemplo, citosina con guanina.

2.1.2 Estructura Viral

Todos los virus son conformados por al menos dos piezas fundamentales: material genético viral —ya sea en forma de ARN o ADN— y una *cápside*. La cápside es una coraza proteica que envuelve al material genético del virus y conforma la estructura principal de este. De tal manera que la partícula viral conformada por la cápside y el material genético recibe el nombre de *virión*, donde, un virión es capaz de infectar una célula y producir miles de viriones como descendencia. Esta cápside puede tomar diversas geometrías según sea la familia del virus: cilíndricas, icosaédricas, alargadas, esféricas, entre otras (Alberts, 2022). Debido a su simpleza y falta de maquinaria celular, todos los virus requieren infectar y secuestrar la maquinaria celular de una célula viva de tal manera que fuerza a esta a generar miles de copias de si mismo para luego abandonarla y continuar el ciclo —a excepción de los retrovirus como el VIH. Es por esta razón que el estatus de los virus como seres vivos es poco claro, a pesar de que se reproduzcan y evolucionen como las células vivas.

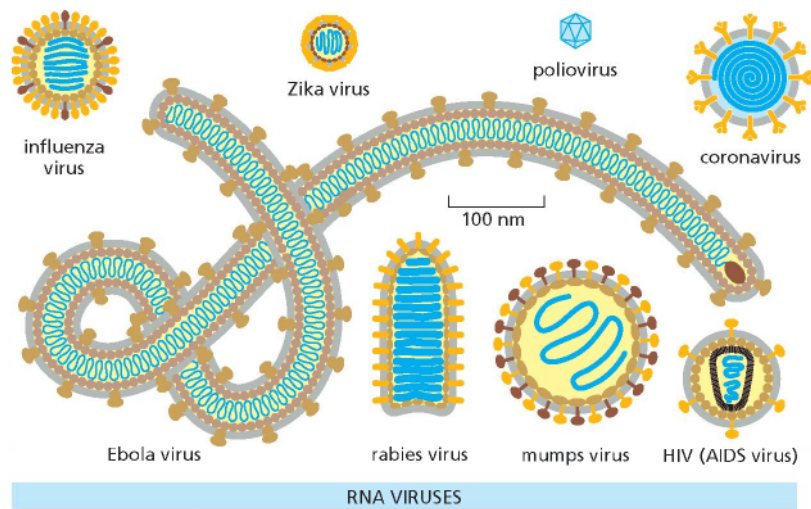


Figura 3: Ejemplos de virus de ARN con su respectiva cápside y material genético. Extraído de Alberts (2022).

El ciclo de reproducción de un virus consta de, a grandes rasgos, 4 etapas las cuales se pueden visualizar en la Figura 4. En primer lugar, ocurre el ingreso a la célula viva y la liberación de su material genético, luego, en segundo lugar, ocurre la replicación del material genético en conjunto con las proteínas virales incluyendo la cápside secuestrando la maquinaria de la célula. En tercer lugar, al interior de la célula se ensamblan las partículas virales empacando el material genético dentro de las cápsides para finalmente, en cuarto lugar, los viriones generados salen de la célula listos para infectar nuevos huéspedes. En este último paso, puede ocurrir que la célula se rompa y muera (Alberts, 2022).

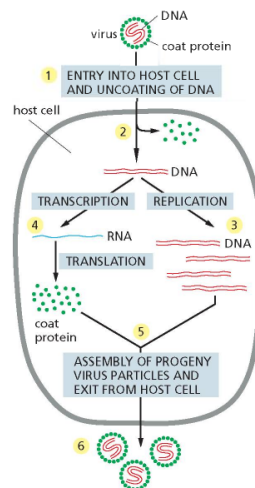


Figura 4: Ciclo de vida de un virus de ADN. Extraído de Alberts (2022).

2.1.3 Enfermedad de Bursitis Infecciosa

El virus de la bursitis infecciosa es un virus de ARN de doble hélice de cápside icosaedro que causa la enfermedad de *bursitis infecciosa* en aves de corral, la cual puede ser letal produciendo una tasa de mortalidad entre el 20% y 50% para las cepas más virulentas. Unas de las medidas de control en la industria para mitigar los efectos de la enfermedad es el uso de vacunas y estrictas medidas de bio-seguridad en los corrales para contener la transmisión virus (Orakpoghenor et al., 2020). Es por estas razones que el virus de la *bursitis infecciosa* es de especial interés para la industria alimenticia, al producir millonarias pérdidas de producción a nivel mundial y reducir la calidad de vida de las aves, incluso tomando en consideración sus medidas de mitigación, lo que incentiva un mayor estudio de las características del virus.

2.1.4 Complejo ARN-VP3

Al interior del virus IBDV existen una serie de proteínas que facilitan su reproducción, una de ellas, denominada VP3, recibe especial atención. Se hipotetiza que esta proteína es capaz de dimerizarse —unirse a otra proteína VP3 formando una estructura mayor— de tal manera de facilitar el empaquetamiento del ARN al interior de la cápside viral.

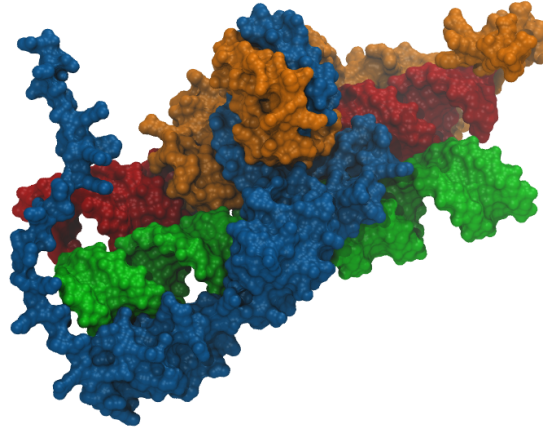


Figura 5: Ejemplo de dímero de VP3 ayudando a empaquetar dos hebras de ARN. Elaboración propia.

Esta proteína contempla dos zonas importantes denominadas: *core* y *terminal-N*, donde el primero corresponde a la zona donde la proteína se une a una hebra de ARN y la segunda corresponde a la zona donde esta se dimeriza con otra VP3. Cabe destacar que la proteína contempla una tercera zona de menor importancia denominada *terminal-C*, esta zona se ha excluido del estudio puesto que es constituida por un segmento de proteína sin estructura. Para efectos del presente trabajo, se tomó la decisión de eliminar la proteína de las simulaciones entre hebras de ARN, debido a que la unión entre el *core* y las hebras de ARN es fuerte y rígida, así como también la unión entre los VP3 mediante el terminal-N, además, las hebras de ARN del virus poseen un largo mucho mayor a las proteínas VP3 por lo que su efecto electrostático global es menor.

2.2 Modelos de Grano Grueso

Existen fenómenos de interés biológico que ocurren a lo largo de diferentes escalas de tiempo cuyo estudio puede ser realizado a través de simulaciones de dinámica molecular que predigan la conformación que tomaran dichas moléculas bajo diferentes condiciones e interacciones. Sin embargo, algunos de estos fenómenos resultan prohibitivamente costosos o directamente imposibles de simular por completo, como lo son el plegamiento de estructuras de ARN y el movimiento de las subunidades de los ribosomas, ya que pueden abarcar escalas de tiempo del orden de milisegundos o más y millones de átomos individuales (Sponer et al., 2018). En luz de este problema, surgen los modelos de grano grueso como una alternativa viable para el cálculo de estructuras moleculares demasiado complejas para una simulación atomística, reduciendo los grados de libertad de la molécula —en el orden de millones en algunos casos— a una cantidad menor manejable computacionalmente. Estos modelos logran esto reemplazando la complejidad atomística de la molécula por una menor cantidad de esferas las cuales interactúan entre sí por medio de potenciales energéticos efectivos y una cantidad de grados de libertad determinados representados generalmente por barras que conectan las esferas, de tal manera que el modelo de grano grueso capture suficiente información de interés de la estructura molecular y sus interacciones para obtener un resultado. Debido a que las interacciones pueden ser de diferente tipo, y que la estructura de interés puede variar caso a caso, existen diferentes maneras de crear un modelo de grano grueso de una molécula. En la Figura 6 se ejemplifican diversas aproximaciones de grano grueso para un *tetraloop* de ARN.

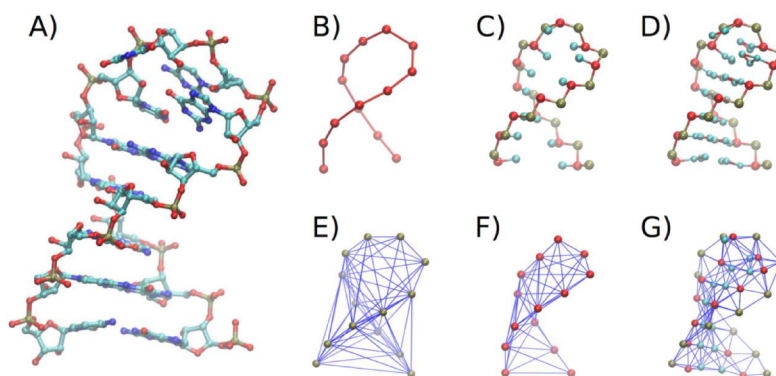


Figura 6: Diferentes ejemplos de modelos de grano grueso para un *tetraloop* de ARN, donde A) indica la estructura *all-atom* y B) a G) indican diversas maneras de simplificar la estructura con grano grueso. Extraído de Sponer et al. (2018)

Existen diversas maneras de parametrizar un modelo de grano grueso, la cual dependerá del fenómeno de interés a estudiar (Sponer et al., 2018). En primer lugar, se puede generar una parametrización de las interacciones entre los sitios de grano grueso mediante simulaciones atomísticas de dinámica molecular, cuya precisión dependerá de la calidad de las simulaciones *all-atom*. En segundo lugar, se puede generar un mapa de las interacciones de grano grueso a partir de resultados experimentales. Observando la frecuencia con la que se observan ciertas estructuras en experimentos, es posible generar una representación de la energía de la configuración mediante el proceso de *inversión de Boltzmann*, el cual relaciona la probabilidad de observar un estado con su energía (Poblete et al., 2018). Finalmente, es posible obtener una parametrización a partir del ajuste de parámetros específicos de funciones de energía de tal manera de reconstruir cantidades experimentales como energías libres.

2.3 Modelos de Solvente Implícito

Dentro de la variedad de maneras de interacción entre moléculas la vía electrostática de medio continuo surge como un mecanismo de especial interés e importancia, especialmente cuando se trata de sistemas biológicos. Esto debido a que las fuerzas electrostáticas son de gran alcance —relativamente— y por ende son quienes suelen iniciar la interacción entre dos moléculas cargadas, y debido a que el solvente donde ocurre la interacción suele poseer iones disueltos que alteran la electrostática del sistema. Además, los sistemas biológicos suelen lidiar con aminoácidos y ácidos nucleicos los cuales contienen una carga eléctrica importante, intensificando la participación de la electrostática en la interacción. Usualmente un sistema biomolecular contempla una o más moléculas cargadas, un solvente el cual por lo general es agua líquida y una cierta cantidad de iones disueltos. Todos estos factores deben ser considerados para una modelación computacional efectiva. Para ello, en los sistemas biomoleculares existen dos grandes categorías: los métodos explícitos y los métodos implícitos (Baker, 2004).

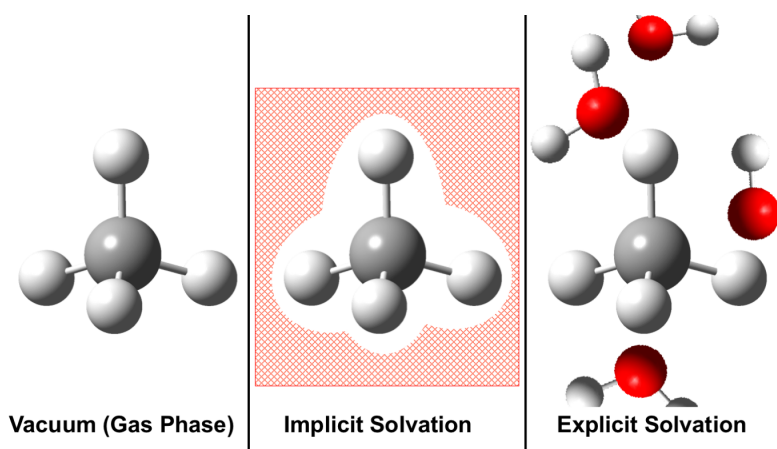


Figura 7: Modelos de solvente implícito y explícito. Extraído de Emmett M. Leddin (2021)

En primer lugar, los métodos de solvente explícito son capaces de entregar una visión muy detallada de la solvatación biomolecular. Logrando esto mediante el modelamiento de las interacciones de los iones móviles del solvente, el solvente y el soluto por medio de campos de fuerza de mecánica molecular (Baker et al., 2008). En otras palabras, se simulan los iones y moléculas del solvente de manera discreta y explícita de tal manera de obtener el comportamiento global de la solvatación. Este alto nivel de detalle acarrea consigo un gran costo computacional puesto que se deben explorar todas las configuraciones posibles en suficiente detalle para obtener valores de energías termodinámicas estables. En segundo lugar, los métodos de solvente implícito surgen como una alternativa de menor costo computacional al enfocarse en el efecto del solvente sobre el soluto como un medio continuo. El soluto —la biomolécula— es representada en detalle atómico mientras que el solvente es representado como el promedio del efecto de este sobre el soluto. De esta manera, la interacción soluto-solvente es representada por la *energía de solvatación* la cual representa el cambio de energía libre de Gibbs al transferir el soluto del vacío al ambiente solvatado. Esta simplificación en el tratamiento del solvente, a costa de introducir cierto grado de error, permite que los métodos de solvente implícito sean ampliamente utilizados para resolver problemas de solvatación biomolecular (Baker et al., 2008) y es, por consiguiente, el método utilizado en el presente trabajo.

2.4 Ecuación de Poisson-Boltzmann

La ecuación de Poisson-Boltzmann (PB) es uno de los modelos implícitos más populares y ha sido ampliamente utilizado para estudiar la electrostática de moléculas en un solvente continuo (Grochowski & Trylska, 2008), modelando el comportamiento del solvente y iones en la proximidad de la molécula.

Su derivación más sencilla, como indica Baker (2004), comienza a partir de la ecuación de Poisson para electrostática general:

$$-\nabla \cdot \epsilon(x) \nabla \phi(x) = \rho(x) \quad \text{para } x \in \Omega \quad (1)$$

con la condición de borde de Dirichlet: $\phi(x) = g(x)$ para $x \in \partial\Omega$.

Donde:

- $\phi(x)$: Potencial electrostático en x .
- $\epsilon(x)$: Constante dieléctrica en x .
- $\rho(x)$: Distribución de carga en x .

Luego, en sistemas biomoleculares, la distribución de carga se separa en dos componentes. La primera consiste en el aporte de las cargas fijas de los átomos de la molécula, denominada ρ_f y la segunda consiste en el aporte de las cargas móviles suspendidas en el solvente alrededor de la molécula, denominada ρ_m . La componente fija consta de cargas puntuales de magnitud Q_i ubicadas en los centros de los átomos de la molécula, x_i , tal que:

$$\rho_f(x) = \frac{4\pi e_c^2}{k_B T} \sum_{i=1}^M Q_i \delta(x - x_i) \quad (2)$$

Por otra parte, la componente de carga móvil modela el efecto promedio de los iones móviles sobre la distribución de carga a partir de una distribución de Boltzmann, de tal manera que para una solución de NaCl —común en sistemas biológicos— es igual a:

$$\rho_m = \bar{\kappa}^2(x) \sinh \phi(x) \quad (3)$$

Donde $\bar{\kappa}$ se denomina *coeficiente de accesibilidad iónica* y captura tanto la intensidad iónica de la solución como la accesibilidad de los iones en el espacio, por lo que es igual a cero en el interior de la cavidad creada por la molécula en el solvente (Baker, 2004). Es equivalente al recíproco de la longitud de Debye (λ_D), la cual representa la distancia característica a la cual persiste el efecto de una carga eléctrica apantallada por una solución electrolítica. Este modelamiento estadístico del efecto de los iones disueltos sobre la biomolécula en la ecuación de Poisson-Boltzmann es uno de los pasos clave que permiten el modelamiento implícito del solvente sin recurrir al alto costo de la dinámica molecular.

Luego, combinando las ecuaciones 1, 2 y 3, y ordenando los términos se obtiene la ecuación de Poisson-Boltzmann No-Lineal:

$$-\nabla \cdot \epsilon(x) \nabla \phi(x) + \bar{\kappa}^2(x) \sinh \phi(x) = \frac{4\pi e_c^2}{k_B T} \sum_{i=1}^M Q_i \delta(x - x_i) \quad \text{para } x \in \Omega \quad (4)$$

Finalmente, como muchas biomoléculas poseen una carga relativamente pequeña, es posible linealizar la ecuación utilizando la aproximación $\sinh \phi \approx \phi$ cuando ϕ es pequeño. Esto da origen a la ecuación de Poisson-Boltzmann Lineal:

$$-\nabla \cdot \epsilon(x) \nabla \phi(x) + \bar{\kappa}^2(x) \phi(x) = \frac{4\pi e_c^2}{k_B T} \sum_{i=1}^M Q_i \delta(x - x_i) \quad \text{para } x \in \Omega \quad (5)$$

Tanto la Ecuación 4 como la Ecuación 5 requieren de una condición de borde de Dirichlet en los bordes del dominio tal que $\phi(x) = g(x)$ para $x \in \partial\Omega$. Esta condición de borde $g(x)$ debe ser capaz de representar el comportamiento del potencial a distancias muy grandes de tal manera de modelar el decaimiento al infinito de este según sea el caso para lo cual existen diversas alternativas (Rocchia, 2005).

2.4.1 Potencial de Yukawa o de Coulomb Apantallado

La ecuación de Poisson-Boltzmann lineal, la Ecuación 5, representa una simplificación de la ecuación de Poisson-Boltzmann No-Lineal para un potencial electrostático relativamente pequeño como el que se encuentra en proteínas y otras moléculas comunes. Reordenando los términos de la Ecuación 5 se obtiene la Ecuación 6.

$$(-\nabla \cdot \epsilon(x) \nabla + \bar{\kappa}^2) \phi(x) = \rho(x) \quad (6)$$

La función de Green para esta ecuación es: $\frac{e^{-\kappa r}}{r}$ por lo que la solución para el potencial generado por una distribución de N cargas puntuales resulta:

$$\phi(r) = \sum_{i,j \neq i}^N q_i \cdot q_j \frac{e^{-\kappa ||r_j - r_i||}}{||r_j - r_i||} \quad (7)$$

Donde κ es el inverso de la distancia de Debye tal que: $\kappa^2 = \frac{I e^2}{\epsilon k_B T}$ con I la fuerza iónica.

Este potencial es denominado como *potencial de Yukawa o de Coulomb apantallado*, y es una aproximación para la interacción electrostática entre pares cargados en la presencia de un solvente electrolítico pues surge naturalmente de la ecuación de Poisson-Boltzmann lineal (Hoppe, 2013) y (Cheng et al., 2011). Representando el efecto del potencial electrostático de Coulomb generado por una carga, pero apantallado por los efectos del solvente y los iones en solución.

2.4.2 Solución Numérica de la Ecuación de Poisson-Boltzmann

La ecuación de Poisson-Boltzmann es una ecuación diferencial parcial elíptica no-lineal que puede ser resuelta mediante diferentes métodos numéricos. Algunos de los más utilizados son: diferencias finitas (FDM), elementos finitos (FEM) y elementos de frontera (BEM) (Baker, 2004) —este último con la desventaja que solo puede ser aplicado a la ecuación lineal a cambio de una mejor representación de la frontera entre el soluto y el solvente. Recientemente, algunos métodos avanzados buscan resolver este y otros problemas mediante una formulación acoplada de FEM y BEM (Guerrero-Montero et al., 2026).

Para efectos del presente trabajo se optó por utilizar el software APBS, el cual fue desarrollado para la resolución de la ecuación de Poisson-Boltzmann aplicada a sistemas biológicos, siendo un código robusto, ampliamente utilizado y maduro (Jurrus et al., 2018). APBS utiliza el método de elementos finitos con una resolución mediante *multigrid* de la ecuación de Poisson-Boltzmann tanto en su formulación lineal como no-lineal para una mejor convergencia, permitiendo además un cálculo en paralelo al descomponer el dominio en regiones de cálculo.

Método de Diferencias Finitas (FDM)

El método de diferencias finitas consiste en aproximar las derivadas de una ecuación diferencial mediante una diferencia definida en un intervalo finito evaluada en múltiples puntos de una malla que cubre todo el dominio de cálculo, la cual usualmente es estructurada y regular (Morton & Mayers, 2005). Estas aproximaciones surgen de la expansión de Taylor de la derivada en un punto x , la cual es truncada a una cierta cantidad de términos induciendo un error que decae conforme la malla se vuelve más fina. La manera en la que se realiza la discretización de la derivada influye en la relación entre el error de esta y el tamaño de la malla.

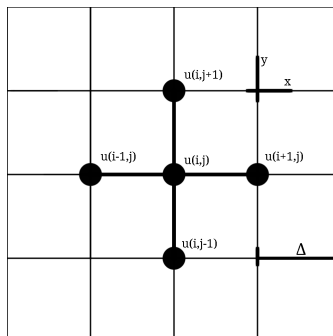


Figura 8: Esquema del método de diferencias finitas para $u = u(x, y)$ con $u \in \Omega$ en una grilla regular de tamaño Δ . Elaboración propia.

Por ejemplo, se puede aproximar la primera derivada de $u = u(x, y)$, ilustrada en la Figura 8, mediante una diferencia central, cuyo error decae conforme al cuadrado del espaciamiento de malla:

$$\frac{\partial u}{\partial x} = \frac{u(i+1, j) - u(i-1, j)}{2\Delta} + O(\Delta^2) \quad (8)$$

Aplicando esta aproximación a la ecuación diferencial a resolver, y planteándola para todos los puntos de la malla teniendo especial cuidado con las condiciones de borde, se puede obtener

un sistema de ecuaciones $\mathbf{A}_{N \times N} \cdot \mathbf{u}_{N \times 1} = \mathbf{f}_{N \times 1}$ el cual puede ser resuelto mediante técnicas de álgebra lineal —si la EDP discretizada es lineal, como la LPBE— o mediante técnicas iterativas como Newton-Rhapson.

El método de diferencias finitas es, por lo tanto, un método flexible y sencillo de implementar que es capaz de enfrentar ecuaciones diferenciales lineales como no-lineales. Esto es a cambio de la desventaja que, la solución deseada o las condiciones de borde, pueden no encontrarse exactamente en un punto de la malla, lo que dificulta su correcta consideración. Esto es especialmente problemático cuando se buscan conservar flujos —como la conservación de masa en fluidos— o cuando se trabaja con cargas puntuales, lo que significa recurrir a interpolaciones complejas de tal manera de relacionar el resultado en los puntos de malla discretos con un dominio continuo. Esto es especialmente relevante al trabajar con la ecuación de PB debido a que: primero, cuenta con cargas puntuales al interior del soluto y segundo, la frontera entre el soluto y solvente rara vez coincide con la malla. Esto implica realizar interpolaciones complejas además de “relajar” los delta-dirac de las cargas lo que introduce un error. Afortunadamente, la documentación de APBS sugiere un mecanismo para lidiar con esto, descrito en la Subsubsección 2.5.1, el cual es aceptado y ampliamente utilizado (Baker, 2023a).

2.5 Cálculos de Energía mediante la ecuación de Poisson-Boltzmann

La ecuación de Poisson-Boltzmann resuelta numéricamente puede ser utilizada para el cálculo de energías libres de diversas configuraciones de proteínas solvatadas en un medio acuoso. Esto resulta de especial importancia debido que las energías libres permiten predecir si una interacción, o configuración específica, es favorable energéticamente o no lo que conlleva a diversas aplicaciones en el campo de la biotecnología.

Para esto, se debe analizar el ciclo termodinámico indicado en la Figura 9 con el cual se pueden dilucidar las dos energías libres de mayor interés para el presente trabajo: la energía de solvatación y la energía de interacción.

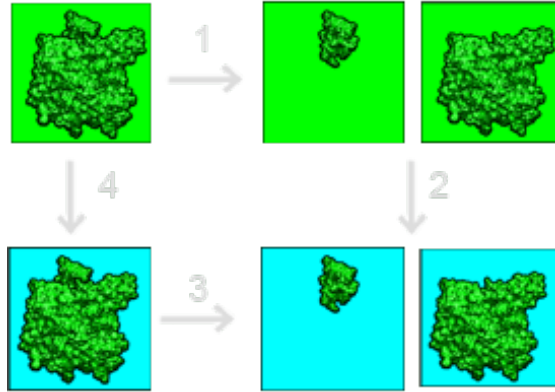


Figura 9: Diagrama del ciclo termodinámico para la energía de interacción de dos moléculas solvatadas. Extraído de Baker (2023b)

En la figura señalada se ilustran 4 posibles configuraciones de dos moléculas cargadas eléctricamente en un medio con permitividad (ϵ) variable, donde el color cian representa la permitividad *bulk* del solvente y el color verde la permitividad presente en el interior de la molécula donde no

hay presencia del solvente y por ende resulta cercana al vacío. Cabe destacar que se entiende el interior de la molécula como aquella región donde el solvente es incapaz de penetrar, es decir, el interior de la superficie de exclusión del solvente (SES) como se indicó en la Subsección 2.4. Por otra parte, las flechas de la figura indican la dirección del cambio de energía libre de Gibbs al pasar de un estado al otro (Baker, 2023b), y debido a que se trata de una variable de estado, la suma de todas las energías libres del ciclo es igual a cero.

2.5.1 Energía de Solvatación

La energía de solvatación representa el cambio de energía libre al introducir la molécula al solvente desde un estado de referencia, donde el estado de referencia posee una permitividad constante en el dominio y el estado en solvente posee igual permitividad de referencia en el interior de la molécula, pero diferente permitividad en el exterior. En otras palabras, es el cambio de energía libre que se produce en presencia del solvente al interactuar con la molécula y viceversa. Esta energía de solvatación posee una componente polar —modelada por la ecuación de Poisson-Boltzmann— y una componente no-polar, la cual para efectos del presente trabajo fue despreciada. Con esto en mente, la energía de solvatación corresponde al cambio de energía libre comprendido por la flecha número cuatro de la Figura 9, la cual traslada la molécula de la referencia al solvente.

$$G_{solv} = \Delta G_4 = G_{sys} - G_{ref} \quad (9)$$

APBS es capaz de calcular eficientemente ambos estados y realizar la resta indicada de tal manera de reportar la energía libre de solvatación en sus resultados sin mayor posprocesamiento, recordando que el estado de referencia corresponde a todo el dominio con una permitividad homogénea y el estado solvatado corresponde a una permitividad variable como se indica en la Figura 9. Los detalles de la configuración al respecto del software se indican en la Subsección 3.1. Cabe destacar que, debido a las dificultades de resolver la ecuación de Poisson-Boltzmann utilizando diferencias finitas como se indicó en la Subsubsección 2.4.2, es de vital importancia cuidar que el mallado sea exactamente idéntico entre ambos estados, puesto que, al realizar la resta, los errores introducidos por la discretización son reducidos (Baker, 2023a). Esto sera de extrema importancia a la hora de calcular energías de interacción.

2.5.2 Energía de Interacción

Para el cálculo de las energías de interacción, que son las de mayor interés para el presente trabajo, utilizando la ecuación de Poisson-Boltzmann se sigue el procedimiento indicado en la documentación de APBS (Baker, 2023b). La energía de interacción se define como el cambio de energía libre ocurrido al llevar dos moléculas desde una gran distancia a estar próximas entre sí en un medio en presencia de solvente, de tal manera que a menor energía de interacción las moléculas poseen mayor afinidad. Un valor de energía de interacción positivo implica, entonces, que las moléculas se repelen entre si mientras que un valor negativo implica que las moléculas se atraen. En vista de esto, se denomina *caso aislado* cuando la molécula se encuentra por si sola y *caso complejo* cuando se encuentran presentes ambas moléculas. Así, se definen en la Figura 9 los cambios de energía libre utilizados para el cálculo de la energía de interacción listados a continuación.

- ΔG_1 : Cambio de energía al llevar las moléculas del caso aislado al caso complejo en el estado de referencia.
- ΔG_2 : Cambio de energía al solvatar las moléculas en el caso complejo, es decir la energía de solvatación del complejo.
- ΔG_3 : Cambio de energía al llevar las moléculas del caso aislado al caso complejo en presencia del solvente. Equivalente a la energía de interacción ΔG_b .
- ΔG_4 : Cambio de energía al solvatar las moléculas en el caso aislado, es decir la energía de solvatación.

Debido a que la energía libre de Gibbs es una variable de estado, la suma de todas las energías libres del ciclo es igual a cero. Recorriendo el ciclo en sentido antihorario se tiene:

$$\Delta G_4 + \Delta G_3 - \Delta G_2 - \Delta G_1 = 0 \quad (10)$$

$$\Delta_b G = \Delta G_3 = \Delta G_2 - \Delta G_4 + \Delta G_1 \quad (11)$$

Por otra parte, según las definiciones indicadas con anterioridad, cada cambio de energía se define como:

$$\Delta G_1 = \Delta_c G_{comp.} - \Delta_c G_{ais.}^a - \Delta_c G_{ais.}^b \quad (12)$$

$$\Delta G_2 = \Delta_s G_{comp.} \quad (13)$$

$$\Delta G_4 = \Delta_s G_{ais.}^a + \Delta_s G_{ais.}^b \quad (14)$$

Donde c indica energía de Coulomb y s energía de solvatación del respectivo caso: aislado o complejo.

Luego, reemplazando lo anterior en la Ecuación 11, se obtiene:

$$\begin{aligned} \Delta_b G &= (\Delta_s G_{comp.} - \Delta_s G_{ais.}^a - \Delta_s G_{ais.}^b) + (\Delta_c G_{comp.} - \Delta_c G_{ais.}^a - \Delta_c G_{ais.}^b) \\ &= \Delta \Delta_s G + \Delta \Delta_c G \end{aligned} \quad (15)$$

Esta energía de interacción se puede separar en las componentes de solvatación y de coulomb: $\Delta \Delta_s G$ y $\Delta \Delta_c G$ respectivamente. La primera contempla el cambio en la energía libre de solvatación polar al llevar las moléculas del caso aislado al complejo, y la segunda captura el cambio en la energía libre Coulómbica al realizar la misma operación. En conjunto, contemplan la totalidad de la energía libre considerada por el modelo, necesaria para llevar las moléculas del caso aislado al complejo en presencia de un solvente (Baker et al., 2008) dado que la componente no-polar ha sido despreciada.

De esta manera, el resultado de las simulaciones con APBS son las energías de solvatación indicadas en la Ecuación 15 — $\Delta_s G_{comp.}$, $\Delta_s G_{ais.}^a$, y $\Delta_s G_{ais.}^b$. La componente coulómbica $\Delta \Delta_c G$ debe ser calculada mediante el procedimiento indicado a continuación. Finalmente, utilizando ambos resultados, se procede a calcular la energía de interacción mediante la Ecuación 15.

Contribución Coulómbica

Como se vio con anterioridad para el cálculo de la energía de interacción entre dos moléculas resulta necesario el cálculo de una componente solvatada $-\Delta\Delta G_s$ y otra coulómbica $-\Delta\Delta G_c$. La primera es calculada con la ayuda de APBS como se describió en la sección anterior, tanto para los casos aislados como complejos, mientras que la segunda debe ser calculada manualmente por reglas de la electrostática elemental al tratarse de un caso de referencia con permitividad constante. Con esto en mente, se recuerda la ley de coulomb para la energía electrostática contenida en una nube de partículas en un medio de permitividad relativa ϵ_r .

$$\Delta G_c(x_i) = \frac{1}{4\pi\epsilon_0\epsilon_r} \sum_{i,i \neq j}^N \frac{q_i \cdot q_j}{||x_i - x_j||} \quad (16)$$

De esta manera, iterando sobre todos los pares de átomos de la molécula con su respectiva carga, se pueden obtener las energías coulómbica tanto para los casos aislados como complejos a utilizar en la Ecuación 15.

Energía de Interacción con Potencial de Yukawa

Por otra parte, recordando el potencial de Yukawa definido en la Subsubsección 2.4.1, es posible emplearlo para el cálculo de la energía de interacción entre dos moléculas directamente. Yukawa al ser una solución de la ecuación de Poisson-Boltzmann Lineal es, en sí mismo, una aproximación de la energía de solvatación de la molécula. De esta manera, observando el ciclo termodinámico de la Figura 9, el potencial de Yukawa entrega el valor de la energía libre directamente para un caso solvatado (Hoppe, 2013) sin necesidad de utilizar la Ecuación 15 sino que siguiendo lo indicado por la Ecuación 17 a continuación.

$$\Delta G_{bind.} = \Delta G_3 = \Delta_{yuk.} G_{comp.} - \Delta_{yuk.} G_{ais.}^a - \Delta_{yuk.} G_{ais.}^b \quad (17)$$

$$\Delta G_s(x_i) = \frac{1}{4\pi\epsilon_0\epsilon_r} \sum_{i,i \neq j}^N \frac{q_i \cdot q_j \cdot e^{-\kappa r}}{r} \quad (18)$$

Donde las energías libres son el resultado de aplicar la Ecuación 18 a la colección de átomos de la molécula de cada caso, o en su defecto, a las esferas de grano grueso del modelo. Lo que respecto a la implementación computacional es prácticamente idéntico a lo realizado por la Ecuación 16 detallada con anterioridad, salvo los parámetros adicionales de Yukawa.

3 Metodología

Como se indicó con anterioridad el objetivo del presente trabajo es parametrizar un modelo de grano grueso mediante simulaciones electrostáticas de la ecuación de Poisson-Boltzmann realizadas por APBS. Para esto, resultó necesario realizar múltiples simulaciones electrostáticas a diferentes distancias y configuraciones de tal manera de generar una curva de energía de interacción según la distancia entre dos moléculas. De esta manera, resulta posible comparar la curva de energía de interacción entregada por el modelo de grano grueso con la curva entregada por la simulación *all-atom* de APBS, la cual, al resolver la ecuación de Poisson-Boltzmann representa con mayor fidelidad la realidad física y por ende constituye el caso base utilizado que debe ser aproximado por el modelo de grano grueso.

3.1 Implementación de APBS

Como se mencionó con anterioridad, APBS, o *Adaptative Poisson-Boltzmann Solver*, es un software especializado para la resolución de la ecuación de Poisson-Boltzmann ya sea en su forma lineal (5) o no-lineal (4) mediante diferencias finitas utilizando multigrid. Es, por ende, altamente paralelizable y eficiente, además de ser un código maduro y probado (Jurrus et al., 2018) para el cálculo de energías de solvatación ofreciendo una forma amigable de resolver la ecuación de PB para estructuras arbitrarias. APBS utiliza archivos **pqr** para generar el mallado y asignar los átomos que definen la molécula. Estos archivos son estructuralmente similares a los **pdb** utilizados en bioinformática, con la salvedad de incluir la carga y radio de cada átomo incluyendo los hidrógenos faltantes del **pdb**. Para generar un archivo **pqr** a partir de un archivo **pdb** existe la herramienta PDB2PQR el asigna cargas y radios a cada átomo y además añade los hidrógenos. De esta manera, un flujo generalizado para la utilización de APBS es el indicado en la Figura 10 el cual contempla el uso de PDB2PQR. Una vez generado el archivo **pqr** para el caso a simular, APBS requiere de un archivo de configuración **.in** donde se configuran y señalan todas las variables relacionadas con la simulación electrostática.

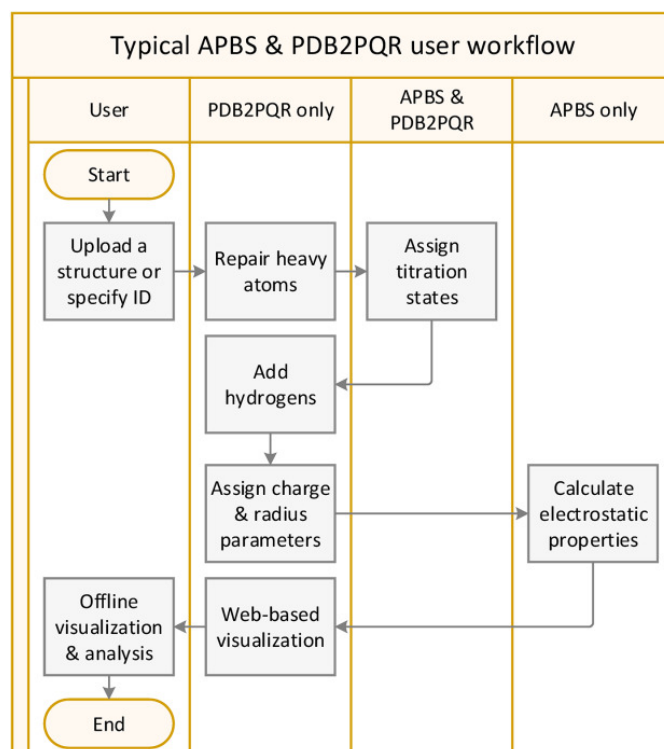


Figura 10: Diagrama de flujo típico para la utilización de APBS en conjunto con PDB2PQR. Extraído de Jurrus et al. (2018)

En el presente trabajo, debido a que la visualización no resultó necesaria al trabajar con curvas de energías libres, el flujo de APBS termina con el cálculo de las propiedades electrostáticas.

3.1.1 Configuración de APBS

APBS posee la flexibilidad necesaria para ajustar todos los parámetros de la simulación electrostática, incluyendo el mallado, condiciones de borde y dimensiones de la caja computacional, propiedades del solvente y otros. Es pertinente, entonces, realizar un pequeño resumen de las opciones disponibles y detallar los parámetros utilizados para las simulaciones de APBS ejecutadas. Con esto en mente, la configuración de APBS se divide en tres secciones generales que serán detalladas a continuación.

READ

Esta sección indica la entrada de datos al programa. Para efectos del presente trabajo, esto es un archivo `pqr`, aunque APBS tenga la capacidad de leer distintos tipos de archivos como carga eléctrica, potencial dieléctrico, entre otros. La sintaxis, en este caso, sigue la forma:

```

READ
mol pqr {dirección al archivo pqr}
END

```

ELEC

Esta sección indica todos los parámetros pertinentes a una simulación de solvatación polar, por lo que es la más extensa y detallada. Un solo archivo `.in` de configuración puede tener más de una sección ELEC, lo que resulta de utilidad a la hora de calcular energías de solvatación.

Como se detalló en la Subsubsección 2.5.1, la energía de solvatación requiere del cálculo de un estado de referencia y un estado solvatado, siendo la diferencia de energía libre entre ambos la energía de solvatación. Debido a consideraciones computacionales detalladas en la Subsubsección 2.4.2, es importante que la malla de ambos estados sea idéntica. El uso de dos secciones ELEC simplifica esta tarea al permitir calcular ambos estados en una misma malla con facilidad. Así, las entradas principales a la sección ELEC siguen el formato:

```
ELEC name {nombre de la sección}
{modo}
{parámetros del modo}
END
```

Donde se permite definir un nombre a la sección ELEC —lo cual será aprovechado para el cálculo de las energías de solvatación más adelante— y un modo con sus parámetros correspondientes. Con esto en mente, para efectos del presente trabajo para el control fino de la simulación, se empleó el modo *mg-manual*. El cuál indica a APBS realizar la simulación mediante multigrid y el método de diferencias finitas (2.4.2), exponiendo al usuario una multitud de parámetros de simulación relevantes que serán detallados más adelante.

PRINT

Finalmente, esta sección permite realizar operaciones aritméticas sobre los parámetros calculados en las secciones ELEC ejecutadas con antelación. De esta manera, es posible calcular directamente en APBS la energía de solvatación e imprimirla a la consola para luego ser utilizada en el cálculo de la energía de interacción entre ambas moléculas. Así, la sintaxis de la sección es:

```
PRINT {variable a manipular} {nombre op. nombre ...} END
```

Así, la configuración de APBS toma la forma indicada en la Figura 11.

3.1.2 Parámetros de APBS Utilizados

Como se observa en la Figura 11, la sección ELEC posee una multitud de parámetros que controlan la simulación. De estos, los más relevantes a mencionar son:

- `ion`: Controla las características de los iones disueltos en el solvente.
- `dime`: Determina las dimensiones de la caja computacional.
- `grid`: Determina el espaciamiento de malla.
- `pdie`: Determina la constante dieléctrica para el interior de la molécula. Tomado como 2,0.
- `sdie`: Determina la constante dieléctrica para el solvente. Tomado como 78,54.
- `lpbe`: Determina si APBS resuelve la ec. de PB lineal (`lpbe`) o no-lineal (`npbe`).

```

elec name solv
mg-manual
dime 289 321 513
grid 0.3 0.3 0.3
gcent mol 1
mol 1
npbe
bcfl mdh
ion charge 1 conc 0.150 radius 2.0
ion charge -1 conc 0.150 radius 2.0
pdie 2.0000
sdie 78.5400
srfm smol
chgm spl2
sdens 10.00
srad 1.4
swin 0.30
temp 298.15
calcenergy total
calcforce no
#write pot dx R0_2rna_solv
end

elec name vacc
mg-manual
dime 289 321 513
grid 0.3 0.3 0.3
gcent mol 1
mol 1
npbe
bcfl mdh
pdie 2.0000
sdie 2.0000
srfm smol
chgm spl2
sdens 10.00
srad 1.4
swin 0.30
temp 298.15
calcenergy total
calcforce no
#write pot dx R0_2rna_vacc
end

print elecEnergy solv - vacc end

```

Figura 11: Ejemplo de configuración de APBS utilizado en el presente trabajo. (a) primera sección ELEC del archivo de configuración junto a sección READ, (b) segunda sección ELEC del archivo junto a sección PRINT. Elaboración propia.

- **srad**: Determina el radio de las moléculas del solvente para generar la superficie de exclusión de solvente alrededor de la molécula. Tomado como 1.4 para un solvente acuoso.
- **bcfl**: Determina las condiciones de borde a aplicar a la caja computacional.

ion

Como se mencionó con anterioridad, el parámetro **ion** especifica las características de los iones disueltos en el solvente, y si estos deben ser incluidos en el cálculo. Su sintaxis es:

```
ion charge {carga} conc {concentracion} radius {radio}
```

Para efectos del presente estudio, para las simulaciones solvatadas de APBS se emplean dos especies iónicas de tal manera de representar un medio biológico.

```
ion charge 1 conc 0.150 radius 2
ion charge -1 conc 0.150 radius 2
```

dime & grid

Los parámetros **dime** y **grid** en conjunto controlan el tamaño de la caja computacional, así como también la resolución de cálculo al especificar el espaciamiento entre nodos y su cantidad en los tres ejes cartesianos. De esta manera, escogiendo un espaciamiento adecuado en **grid** y conociendo el tamaño de lo que se requiera calcular, se puede obtener la cantidad de nodos a incluir en **dim**. De esta manera, sigue la sintaxis:

```
dim nx ny nz
grid delta_x delta_y delta_z
```

Donde para efectos de simpleza, se utiliza un espaciamiento idéntico para las tres direcciones cartesianas. Es relevante mencionar que la elección de **nx**, **ny** y **nz** no es del todo libre, sino que debe ajustarse a la Ecuación 19.

$$n = c2^{l+1} + 1 \quad (19)$$

Donde c es un número entero y l el nivel de la jerarquía utilizada por el *solver* de *multigrid*, usualmente tomado como 4. Para efectos de la automatización de las simulaciones esto resulta especialmente relevante, pues APBS automáticamente puede ajustar el n proveído por el usuario a uno que se ajuste a la ecuación indicada, pero puede ocurrir que deje parte de la molécula fuera del dominio de cálculo, o bien, esta quede muy próxima a uno de los extremos de la caja computacional lo cual, por razones que se verán a continuación, afectaría negativamente el resultado del cálculo.

bcbf

El parámetro **bcbf** controla las condiciones de borde de la simulación por lo que requiere de especial atención. APBS permite múltiples elecciones, siendo la utilizada por el presente trabajo la del tipo “Múltiple Debye-Hückel” (mdh). Para el uso correcto de esta condición de borde se debe asegurar que la molécula a simular esté a al menos de 10 [Å] de cualquiera de las paredes de la caja computacional e idealmente no más allá de 20 [Å] para no desperdiciar memoria innecesariamente, esto según resultados encontrados experimentalmente. Con esto en mente, el gestor de simulaciones construido detallado en la Subsección 3.2 determina el tamaño de **nx**, **ny** y **nz** de tal forma que la molécula quede a alrededor de 15 [Å] de los bordes asegurando que se sigan las reglas requeridas por APBS para la elección de estos valores como se detalló con anterioridad.

3.1.3 Simulaciones Aisladas

Como se mencionó en la Subsubsección 2.5.2, para el cálculo de energías de interacción es necesario el cálculo de energías de solvatación de casos aislados, esto es, cada biomolécula del complejo simulada por separado. Sin embargo, debido a las consideraciones del mallado necesarias para APBS indicadas en las secciones 2.4.2 y 2.5.1, resulta de especial importancia que la malla de los casos complejos y aislados sea equivalente puesto que pequeñas diferencias en la interpolación de las cargas o la superficie de solvente excluido introducen un error significativo al cálculo final de la energía de interacción.

Con esto en mente, se optó por mantener un mismo **pqr** para todas las simulaciones APBS necesarias para la energía de interacción de una configuración geométrica, apagando y activando átomos según sea necesario. Dicho de otra manera, para generar la simulación aislada de la biomolécula “A” se utilizó el mismo **pqr** que el caso complejo, con la salvedad de asignar la carga y el radio de todos los átomos pertenecientes a la otra biomolécula —“B”— a cero. De esta manera, todos los átomos de la biomolécula “B” son ignorados por APBS al, primero, no tener un radio que genere una superficie excluida de solvente y, segundo, no tener carga. Así, la biomolécula “B” es apagada y vuelta invisible a APBS permitiendo simular la biomolécula “A” de manera aislada manteniendo el mismo mallado que el caso complejo. En la Figura 12 se puede observar cómo se ve esto en un archivo **pqr**.

ATOM	3953	H1'	RC	E	40	21.967	-24.439	-45.204	0.2029	1.2870
ATOM	3954	H4'	RC	E	40	24.175	-23.441	-47.425	0.1174	1.3870
ATOM	3955	H5	RC	E	40	18.033	-21.573	-47.221	0.1928	1.4590
ATOM	3956	H42	RC	E	40	15.897	-22.340	-46.343	0.4234	0.6000
ATOM	3957	H41	RC	E	40	15.757	-23.818	-45.383	0.4234	0.6000
ATOM	3958	C5'	RC	F	1	-1.392	-23.564	-49.758	0.0000	0.0000
ATOM	3959	O5'	RC	F	1	-2.338	-24.542	-49.372	0.0000	0.0000
ATOM	3960	C4'	RC	F	1	-2.027	-22.513	-50.631	0.0000	0.0000
ATOM	3961	O4'	RC	F	1	-2.568	-23.148	-51.832	0.0000	0.0000
ATOM	3962	C3'	RC	F	1	-3.224	-21.796	-50.034	0.0000	0.0000

Figura 12: Ejemplo de archivo pqr para una simulación aislada. Notar que los átomos pertenecientes a la biomolécula apagada poseen carga y radio cero. Elaboración propia.

3.2 Gestor de Simulaciones

Debido a que el cálculo de la energía de interacción a una determinada distancia requiere de al menos tres simulaciones electrostáticas —casos complejo, aislado (a) y aislado (b)— y que por cada curva de energía de interacción se determinó un total de 9 puntos para caracterizarla correctamente, la cantidad de simulaciones APBS escala rápidamente al orden de los cientos sobre todo cuando se requieren diferentes configuraciones geométricas. Adicionalmente, la energía de interacción requiere, además, del cálculo de la componente coulombica correspondiente como se indicó en la Subsubsección 2.5.2 y del cálculo de la distancia entre paredes excluyentes de solvente (SES) de ambas moléculas, junto con realizar cálculos para el modelo de grano grueso en equivalentes condiciones. Para gestionar esta complejidad adicional del problema estudiado, se desarrolló el programa *apbsHelper* de tal manera de facilitar y estandarizar el cálculo y posterior análisis de las curvas de energía de interacción entre moléculas bajo diferentes modelos, configuraciones geométricas y otros parámetros de manera agnóstica al PDB en cuestión. Para lograr esto, el programa se compone de cuatro clases con diferentes responsabilidades y jerarquía: **ConfigHandler**, **CaseHandler**, **SimulationHandler** y **HandlerFunction**. De esta manera, las responsabilidades de cada clase se resumen en:

1. **ConfigHandler**: Ingiera y verifica la configuración **yaml** del usuario. Crea objetos **CaseHandler**, compila y reporta resultados.
2. **CaseHandler**: Configura y ejecuta en paralelo múltiples instancias de **SimulationHandler** pertenecientes a un mismo caso, incluyendo las simulaciones aisladas y complejas.
3. **SimulationHandler**: Configura y ejecuta una sola instancia de APBS, o Yukawa, por completo con parámetros específicos. Genera archivos requeridos en su carpeta de trabajo y calcula la energía de interacción respectiva. Recupera resultados de la simulación si existe el respectivo archivo json generado en una ejecución pasada.
4. **HandlerFunction**: Utilizada por **SimulationHandler**, define la estructura mínima para la función **handle()** creada por el usuario. Es decir, cada objeto *handler* definido por el usuario para su estudio debe ser una instancia de esta clase. Genera archivos PDB y PQR, aplicando Grano Grueso en los casos pertinentes. Debe generar un PDB con ambas moléculas (Complex), y uno para cada molécula por separado (Left y Right) puesto que son requeridos por Nanoshaper para el cálculo de la distancia entre paredes.

A grandes rasgos, entonces, el flujo del programa es el siguiente:

1. **ConfigHandler** ingiere la configuración **yaml** entregada por el usuario la cual define los casos a correr y sus parámetros, comprobando sintaxis y creando la estructura general de carpetas.

2. **ConfigHandler** genera objetos de **CaseHandler** en base a la configuración del usuario, donde cada caso corresponde a un conjunto de puntos relacionados entre sí con parámetros en común. Por ejemplo, un ángulo de rotación a diferentes distancias, o viceversa, utilizando la ecuación lineal de PB. Otro caso podría, por ejemplo, realizar la misma configuración geométrica, pero resolviendo la ecuación no-lineal de PB o utilizar un mallado diferente. Cada caso posee tres conjuntos de objetos **SimulationHandler** de simulaciones: Complejos, Aislados Derecho y Aislados Izquierdo.
3. Cada objeto **CaseHandler** luego genera sus respectivos objetos **SimulationHandler**, donde cada uno de ellos es responsable de configurar y ejecutar una simulación APBS, o Yukawa, de inicio a fin, junto al cálculo de la energía de interacción y la distancia entre paredes mediante Nanoshaper.
4. **CaseHandler** configura y ejecuta en paralelo sus objetos **SimulationHandler**, comenzando por las simulaciones aisladas y luego las complejas. Cada **SimulationHandler** por consiguiente ejecuta su *pipeline* desde configuración a ejecución y término.
5. Para la configuración de cada instancia de **SimulationHandler**, en primer lugar, se ejecuta la función `handle()` del objeto **HandlerFunction** entregado por el usuario, la cual genera los `pdb` y `pqr` del caso base modificados según sea el caso, aplicando el modelo de Grano Grueso al archivo `pqr` en caso de ser requerido. De esta manera, el usuario puede definir su propio *handler* que realice las operaciones geométricas y de grano grueso que sean necesarias para el caso de estudio, teniendo la posibilidad de entregar libremente los parámetros de configuración necesarios en el archivo `yaml`.
6. Terminadas todas las simulaciones de todos los casos, **ConfigHandler** compila los resultados y crea un archivo csv incluyendo los metadatos pertinentes para su posterior análisis.

Además, el programa expone *logs* a la consola y es capaz de recuperar casos incompletos si se detiene antes de lo esperado para continuar su ejecución desde donde terminó. Se incluye un ejemplo de configuración de *apbsHelper* en el Anexo B.

3.3 Implementación de LAMMPS

Una vez parametrizado el modelo de grano grueso, se siguió con estudiar la compactación del ARN del virus en su cápside mediante simulaciones simplificadas de dinámica para diferentes configuraciones de ARN de tal manera de probar el comportamiento de la parametrización. En otras palabras, comenzando desde una hebra de ARN completamente estirada se simuló como esta hebra se compacta hasta alcanzar el tamaño final del interior del virus de tal manera de dilucidar la configuración que allí podría tomar. Para realizar esto, fue necesario separar en dos etapas las simulaciones: *warm-up* y *production*, las cuales fueron realizadas en LAMMPS.

LAMMPS es código para la realización de simulaciones de dinámica molecular creado originalmente por *Sandia Corporation* del Departamento de Energía de EE. UU. en 1995 (Sandia Corporation, 2026). Su intención de diseño busca ser un código para la modelación de materiales flexible y modificable, capaz de ejecutar grandes simulaciones moleculares eficientemente en paralelo —esta flexibilidad ha permitido que LAMMPS sea extendido por diversos investigadores a lo largo de los años. Es por esta razón, que se escogió dicho código para la simulación dinámica del modelo de grano grueso parametrizado mediante APBS al permitir un control fino de los parámetros del modelo manteniendo la eficiencia computacional.

De esta manera, se buscó simular mediante dinámica molecular la configuración que adquiere el ARN una vez comprimido al interior de la cápside del virus, de tal manera de evaluar si el modelo de grano grueso parametrizado produce resultados comparables con la realidad biológica del virus a pesar de sus limitaciones.

LAMMPS expone al usuario una gran cantidad de parámetros que permiten controlar por completo la simulación realizada, por lo que se adjunta un ejemplo de configuración utilizado en el Anexo C, sin embargo, a continuación, se presentan las principales configuraciones utilizadas respecto al control de la simulación.

1. `units`: Controla el sistema de unidades utilizado por LAMMPS. En este caso, corresponde a Lenard-Jones.
2. `boundary`: Controla el tipo de condición de borde, puede ser “p”, periódica, o “f”, fija.
3. `read_data`: Indica el archivo `.dat` a utilizar como condición inicial.
4. `dump`: Controla cada cuantos pasos y que información escribir al archivo `.lammptraj` de resultado.
5. `timestep`: Controla el paso temporal de la simulación.
6. `run`: Controla cuantos pasos temporales ejecutar.
7. `thermo_style`: Controla los parámetros de simulación impresos a la consola.

Así, mediante la manipulación de estos parámetros es posible definir el progreso de la simulación detalladamente y las variables a almacenar. Esto fue utilizado en gran parte para controlar el paso temporal y otros parámetros de las fases de compresión y posterior relajación de cada caso, como se indica en la Figura 13. En cuanto a las variables adicionales a las posiciones de los monómeros y sus variables termodinámicas, se calculó también la energía potencial del sistema de tal manera de monitorear y evaluar la estabilidad de la simulación.

```
run 37500000          # Compress for 37.500.000 steps
variable rad equal "35" # Freeze sphere in place
timestep 0.0001        # Reduce ts ←
run 350000000         # Run other 300.000.000 steps
```

Figura 13: Ejemplo de configuración de LAMMPS donde se indica la separación entre las fases de compresión y relajación para un caso con distinto paso temporal. Elaboración propia.

Las opciones que definen los parámetros del modelo de grano grueso, como sus coeficientes de elasticidad e interacción, son indicados en la Subsección 3.4.

3.3.1 Sistema de Unidades Utilizado

Como se mencionó en la sección anterior, LAMMPS tiene la capacidad de trabajar bajo diferentes sistemas de unidades según sea el caso. Para el presente trabajo, se utilizó el sistema de unidades reducidas de Lenard-Jones —indicado en la configuración de LAMMPS como `units lj`. Esto significa que la energía se encuentra medida relativa a unidades de $k_B T$ donde T es la temperatura modelada por la simulación en Kelvin y las distancias medidas en nanómetros. Así, resulta necesario realizar una conversión de unidades para ingresar los parámetros parametrizados con APBS a LAMMPS. Para estos efectos, todas las simulaciones realizadas en LAMMPS fueron

evaluadas con temperatura $T = 300$ [K], como se indica en el Cuadro 1. De esta manera, utilizando el factor de conversión indicado, es posible transferir los resultados de la parametrización realizada con APBS al sistema de unidades utilizado por LAMMPS.

Cuadro 1: Resumen de cálculo de factor de conversión de unidades $k_B T$.

Parámetro	Valor	Unidades
T	300	[K]
k_B	$1,380649 \cdot 10^{-23}$	$[\frac{J}{K}]$
$k_B T$	$4,141947 \cdot 10^{-21}$	$[\frac{J}{kT}]$

3.3.2 Simulación de Inicio (*Warm-Up*)

Todos los casos estudiados en LAMMPS constan de dos simulaciones. En primer lugar, una vez generado la cadena de ARN mediante un script, se procede a relajarla con potenciales suaves. Esto es necesario debido a que el *script* que genera la condición inicial de ARN no genera un caso físicamente plausible, sino que genera una columna recta de ARN sin consideración a los ángulos entre monómeros ni los parámetros de rigidez del polímero. Así, con el fin de evitar problemas de convergencia en la simulación de producción y de obtener una condición inicial física, se procede con una corta simulación previa denominada *warm-up*. Esta simulación tiene el propósito de relajar la estructura generada por el *script* inicial y llevarla a un caso físicamente plausible y estable, considerando también los parámetros de rigidez del modelo de grano grueso, antes de realizar cualquier compresión asegurando la estabilidad de las simulaciones posteriores.

Esta simulación consiste en dos fases:

1. Relajación con potencial entre pares *soft*.
2. Relajación con potencial entre pares Lenard-Jones.

La primera consiste en una relajación breve de tal manera de evitar interferencias entre los monómeros creados por el *script*, y la segunda busca relajar la estructura hacia un estado físicamente plausible para así generar la condición inicial de la simulación de compresión. Debido a que esta simulación busca relajar la estructura, se utilizaron condiciones de borde periódicas.

3.3.3 Simulación de Compresión (*Production*)

Una vez ejecutada la simulación inicial, se prepara y ejecuta la simulación de compresión de tal manera de modelar el empaquetamiento del ARN a la cápside del virus. La cual, a diferencia de la simulación anterior, utiliza condiciones de borde fijas por lo que se debe tener especial cuidado en establecer los límites de la caja computacional de tal manera que todos los monómeros de la condición inicial estén dentro de ella.

En esta etapa, se introduce el potencial de Yukawa entre los pares de monómeros según lo parametrizado mediante APBS y el modelo de grano grueso descrito en la Subsección 3.4, con el cuidado de convertir las unidades como se indicó en la Subsubsección 3.3.1 dando como resultado lo que se indica en la Figura 14.

De esta manera, la simulación de compresión contempla el potencial de Yukawa para las interacciones entre pares de monómeros no enlazados directamente. LAMMPS modela el potencial

```
# Yukawa Potential T = 300
pair_style yukawa 1.27282 6.0
pair_coeff * * 70.91960 # Default interaction
pair_coeff 1 1 70.91960 # Bulk-Bulk interaction
pair_coeff 1 2 63.82764 # Bulk-Cap interaction
pair_coeff 2 2 57.44487 # Cap-Cap interaction
```

Figura 14: Ejemplo de potencial de Yukawa introducido en LAMMPS para una simulación de compresión. Elaboración propia.

de Yukawa según la forma indicada en la Ecuación 20 (Sandia Corporation, 2026), para lo cual es necesario el cálculo del coeficiente “A” en las unidades descritas con anterioridad según el modelo de grano grueso e indicadas en la Figura 14.

$$E = A \frac{e^{-\kappa r}}{r}, r < r_c \quad (20)$$

La simulación de compresión se puede dividir en dos etapas:

1. Compresión hasta el radio final del virus.
2. Relajación hasta alcanzar un estado estacionario.

Para modelar la compresión de las hebras de ARN se utiliza un muro potencial esférico Lenard-Jones únicamente repulsivo descrito en la Ecuación 21, el cual fuerza al polímero a compactarse hasta adquirir el tamaño final del virus. El cual —utilizando datos de Cryo-Em extraídos de Mata et al. (2018)— fue fijado en 25 [nm] de radio.

$$LJ126(r; \sigma; \epsilon) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right] \text{ con } r < r_{cutoff} = 2^{1/6} \sigma \quad (21)$$

Cabe destacar que, según las pruebas empíricas llevadas a cabo sobre el muro potencial, el polímero de ARN se mantuvo como máximo a $r_{max} = R - \sigma$ [nm] del centro de la esfera potencial, donde R es el radio de esta, σ el parámetro correspondiente al potencial y r_{max} la máxima distancia que toma el polímero del centro. Así, para lograr que el polímero logre una compresión equivalente al radio del virus de 25 [nm] visto en el Cryo-Em, la esfera potencial de la simulación debió adquirir un radio final equivalente a $R = 25 + \sigma$ [nm]. Una vez finalizada la etapa de compresión y el polímero alcance el radio establecido para el virus, se congela la posición del muro potencial y se permite que el sistema evolucione en ese estado comprimido hasta alcanzar un estado estacionario.

3.3.4 Método de Análisis

Debido a que LAMMPS entrega las ubicaciones de todos los monómeros de las hebras resulta necesario definir un método de análisis de tal manera de identificar si la etapa de relajación alcanzó un estado estacionario además de visualizar las estructuras que se puedan generar. Para esto, se establecieron los histogramas detallados a continuación. Adicionalmente, para obtener una mayor cantidad de datos y suavizar las perturbaciones térmicas, para cada histograma se consideraron 100 *frames* de datos.

1. Histograma de Densidad Radial
2. Histograma de Distancias entre Monómeros

3. Histograma de Distancias entre Vecinos Lejanos
4. Histograma de Ángulos
5. Matriz de Pares

Histograma de Densidad Radial

El presente histograma comienza como un histograma de las distancias los monómeros al origen del sistema —el cual se encuentra en el centro del casquete potencial— el cual después es normalizado según el centro de cada *bin* al cuadrado y el ancho de estos. De esta manera, se convierte la frecuencia en bruto a una medida de monómeros por nm^3 .

$$\rho(r) = \frac{\text{freq}(r)}{\Delta r \cdot r_{1/2}^2}$$

Luego, se sobrepone sobre el histograma generado el muro potencial utilizado y los datos de densidad radial empíricos obtenidos mediante Cryo-EM según los resultados vistos por Mata et al. (2018).

Histograma de Distancias entre Monómeros

Para construir el presente histograma se calcularon todas las distancias entre pares de monómeros para luego construir el histograma de distancias entre estos, sin embargo, se tomó la precaución de excluir de dicho calculo los 4 vecinos inmediatos de cada monómero, pues estos forman parte de los enlaces que conforman la cadena: estiramiento, rotación, ángulo y diedro. Además, se normalizó la frecuencia por el ancho de cada bin y el centro de estos al cuadrado.

Histograma de Distancias entre Vecinos Lejanos

Resulta interesante establecer si las distribuciones de distancias varia si se trata de monómeros de una misma hebra versus entre una hebra y las demás, en conjunto con establecer si existen diferencias respecto a las distancias entre los vecinos lejanos de cada monómero. De esta manera, se puede establecer el nivel de torsión de una cadena si, por ejemplo, los vecinos lejanos se encuentran próximos a un monómero puesto que esto indicaría si la cadena se dobló de tal manera que monómeros lejanos en índice se encuentran próximos en el espacio. Para esto, se decidió excluir los 20 vecinos de cada monómero para la construcción del histograma puesto que el perímetro de la esfera al final de la compresión es aproximadamente equivalente a 40 monómeros de largo. Por último, se normalizo el histograma de manera similar a los anteriores.

Histograma de Ángulos

El presente histograma es generado separando cada hebra en segmentos de 11 monómeros y calculando el ángulo que conforma dicho segmento entre el primer, quinto y undécimo monómero. De esta manera, es posible establecer el nivel de compresión de cada hebra en el espacio al interior de la cápside, donde un histograma sesgado hacia ángulos menores indica un mayor nivel de compresión debido a que la hebra se encontraría más doblada.

Matriz de Pares

Finalmente, se construyó una visualización de la matriz de pares de distancias entre monómeros. Para esto, en primer lugar, se calculó dicha matriz de distancia entre pares para los monómeros de una hebra en particular, para la cual se calculó la distancia promedio entre pares durante los 100 *frames* utilizados. Luego, los pares que se encontraron a menos de 8 $[nm]$, esto es, aproximadamente la distancia máxima de dimerización de la proteína VP3 según lo indicado en la

Subsubsección 2.1.4, fueron coloreados. De esta manera, resulta posible visualizar la distribución de la hebra en el espacio y si esta contiene zonas donde resulta plausible que la proteína se acople a las cadenas de ARN.

3.4 Modelo de Grano Grueso

El modelo de grano grueso a aplicar al ARN resulta de suma importancia puesto que su parametrización y estudio en contraste con las simulaciones electrostáticas *all-atom* es un pilar central para el presente trabajo. Con esto en mente, se optó por utilizar un modelo de grano grueso el cual condensa la información de 5 pares de nucleótidos de ARN por una esfera cargada eléctricamente; recordando que el ARN en cuestión pertenece a un virus de ARN de doble hélice y por consiguiente consiste en dos hebras de ARN antiparalelas entrelazadas entre sí. Para esta esfera se empleó un radio de 11,185 Å y una carga eléctrica equivalente a la cantidad de fosfatos contenidos en los 5 pares de nucleótidos, recordando que cada nucleótido aporta un fosfato, por último, su posición se ubica en el centro de masa de los 5 pares de nucleótidos a simplificar.

Cuadro 2: Resumen de parámetros de grano grueso por esfera.

Ítem	Parámetros por esfera
Cantidad de nucleótidos	10 (5 pares)
Radio	11,185 Å
Centro	Centro de masa del grupo de nucleótidos.
Carga	Equivalente a la cantidad de fosfatos en el grupo.

Cabe destacar que debido a lo expuesto en la Subsubsección 2.1.1 respecto a la estructura del ARN la cantidad de grupos fosfatos por nucleótido puede variar según su posición. Los nucleótidos del extremo 3' del cualquiera de las hebras de ARN no poseen un grupo fosfato, por lo que no introducen carga a la esfera de grano grueso que los contenga. Dicho de otra manera, las esferas de grano grueso de los extremos de las hebras tendrán una carga inferior al *bulk* debido a los extremos 3' en los bordes de la doble hélice de ARN.

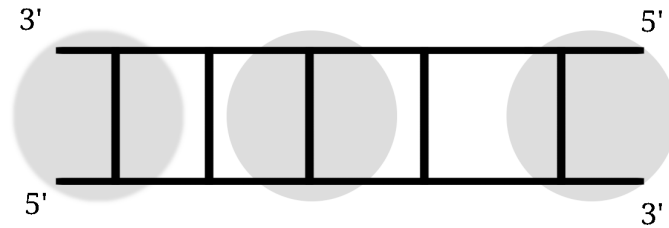


Figura 15: Esquema del modelo de grano grueso empleado respecto a los extremos 3' y 5' de la doble hélice de ARN. Elaboración propia.

Como se indica en la Figura 15 las esferas de grano grueso de los extremos contienen un extremo 3' y otro 5', debido a la doble hélice, por consiguiente, estas esferas poseen 9 fosfatos mientras que el *bulk* posee 10 fosfatos de tal manera que la carga de ambas esferas será de $-9e$ y $-10e$ respectivamente. Así, observando progresivamente la carga de una doble hebra de ARN de grano grueso, se obtiene lo indicado en el Cuadro 3. Cabe destacar que, si se trata de una cadena de tan solo 5 pares de nucleótidos, la esfera de grano grueso tendrá una carga de $-8e$ puesto que contendrá por sí sola los dos extremos 3' de la doble hélice.

Cuadro 3: Ejemplo de cargas de grano grueso de ARN.

Índice	Rango de Nucleótidos	Carga [e]
1	1:5 y 15:20	-9
2	5:10 y 10:15	-10
3	10:15 y 5:10	-10
4	15:20 y 1:5	-9

Finalmente, a modo de implementación del modelo de grano grueso se modifica el `pqr` a simular de tal manera de reemplazar por ceros la carga de todos los átomos presentes, para luego añadir al final del archivo las posiciones, cargas y radios de las esferas de grano grueso como se indica en la Figura 16. Los radios, por otra parte, de los átomos originales del `pqr` podrán ser reemplazados por ceros, siendo identificado como un caso *SimpleGeometry* o bien manteniéndolos sin modificar — caso *FullGeometry*. Esto se realiza debido a que APBS genera la superficie de exclusión de solvente en base a estos radios y de esta forma se puede apreciar el efecto de la geométrica en el cálculo de las energías mediante APBS manteniendo la carga concentrada.

```

ATOM 2920 H42 RC F 40 -14.661 2.602 -5.490 0.0000 0.0000
ATOM 2921 H41 RC F 40 -15.175 0.968 -5.928 0.0000 0.0000
ATOM 90000 H41 RC D 0 -12.694 0.000 -0.000 -8.0000 11.185
ATOM 90001 H41 RC C 0 11.436 48.070 0.539 -9.0000 11.185
ATOM 90002 H41 RC C 1 13.871 33.863 -5.316 -10.000 11.185
ATOM 90003 H41 RC C 2 14.430 21.281 4.241 -10.000 11.185
ATOM 90004 H41 RC C 3 12.161 7.035 -3.438 -10.000 11.185
ATOM 90005 H41 RC C 4 17.576 -6.562 3.132 -10.000 11.185
ATOM 90006 H41 RC C 5 10.748 -21.113 -1.540 -10.000 11.185
ATOM 90007 H41 RC C 6 20.788 -33.979 -0.285 -10.000 11.185
ATOM 90008 H41 RC C 7 13.452 -48.593 2.651 -9.0000 11.185

```

Figura 16: Ejemplo de archivo `pqr` con un modelo de grano grueso, indicado por las ultimas filas de este.

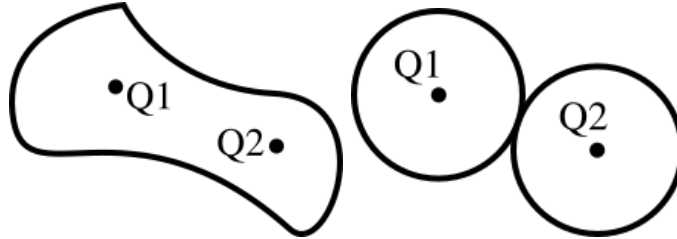


Figura 17: Diferencia entre casos *FullGeometry* (a) y *SimpleGeometry* (b) respecto a la superficie de exclusión de solvente generada por APBS. Elaboración propia.

Se puede observar con mejor claridad la diferencia entre los casos *FullGeometry* y *SimpleGeometry* en la Figura 17. Donde el primero mantiene la geometría real de la superficie de exclusión de solvente de la molécula, pero concentrando las cargas en los centros de las esferas de grano grueso, mientras que el segundo simplifica la superficie de solvente excluido reemplazándola por esferas con carga centralizada. Por otra parte, para la realización de las simulaciones de dinámica molecular una vez parametrizado el modelo, se definieron los coeficientes de elasticidad del modelo los cuales contienen la información acerca de la rigidez del ARN.

3.5 Parámetros Utilizados y Métricas de Error

Con el objetivo de llevar a cabo la parametrización del modelo de grano grueso se optó por utilizar como parámetro principal la energía de interacción entre dos hebras de ARN bajo diferentes orientaciones simuladas con ayuda de APBS. Para lo cual se definió un sistema coordenado local indicado en la Figura 18. Además, para simplificar la modelación y configuración de los casos, se mantuvo una hebra de ARN fija mientras que la otra —denominada hebra móvil— recibe los ajustes geométricos de la simulación.

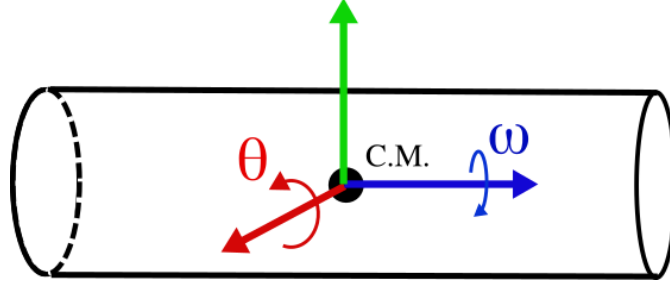


Figura 18: Sistema coordenado definido para el sistema a parametrizar. Elaboración propia.

3.5.1 Parámetros APBS y Yukawa

En primer lugar, se presentan los parámetros utilizados por APBS y Yukawa para los cálculos de energía de interacción presentes en la sección de parametrización del modelo de grano grueso en el Cuadro 4.

Cuadro 4: Parámetros del solvente.

Ítem	Valor	Unidad
Radio Atómico	2	Å
Carga	± 1	$\bar{e}$
Concentración Molar	0,150	M
Distancia de Debye (κ Yukawa)	7,8566 (0,1273)	Å(1/Å)
Constante Dieléctrica	78,54	-

Cabe destacar que el factor κ utilizado en el potencial de Yukawa corresponde al inverso de la distancia de Debye según lo indicado en la Subsubsección 2.4.1, indicado entre paréntesis en el cuadro previo. Cabe destacar que para las posteriores simulaciones realizadas en LAMMPS este debió ser convertido a las unidades de $[1/nm]$. Se establecieron, además, un par de parámetros para caracterizar y comparar los resultados de la parametrización mediante el potencial de Yukawa los cuales son indicados en las ecuaciones 22 y 23 a continuación.

$$e(r, \omega, \theta) = G_{bind.}^{yukawa}(r, \omega, \theta) - G_{bind.}^{apbs}(r, \omega, \theta) \quad (22)$$

$$E(\omega, \theta) = \int_{\forall r} |G_{bind.}^{yukawa}(r, \omega, \theta) - G_{bind.}^{apbs}(r, \omega, \theta)| dr = \int_{\forall r} |e(r, \omega, \theta)| dr \quad (23)$$

De esta manera, la primera métrica corresponde a la distancia entre las energías de solvatación de ambos métodos a diferentes distancias, y la segunda, al integrar sobre la distancia, funciona

Cuadro 5: Otros parámetros de APBS.

Ítem	Valor	Unidad
Constante Dieléctrica del Solute (ϵ_p)	2	-
Temperatura	298,150	K
Radio de Sonda del Solvente	1,4	Å
Tamaño de Malla	0,300	Å

como una métrica que permite comparar la diferencia absoluta entre el modelo de grano grueso y el caso base de APBS según la orientación de las hebras en un solo valor.

3.5.2 Parámetros LAMMPS

LAMMPS expone una multitud de parámetros al usuario por lo que se incluye un ejemplo de configuración en el Anexo C. De todas maneras, a continuación en el Cuadro 6 se detallan los principales parámetros correspondientes a las simulaciones de compresión.

Cuadro 6: Opciones del control de simulación de LAMMPS.

Parámetro	Valor	Unidades
Velocidad de Compresión	$1 \cdot 10^{-5}$	[nm/paso]
Radio Final	35	[nm]
Paso Temporal	0,001	[-]
Pasos por Frame	100000	[pasos/frame]

Los detalles del modelo de grano grueso respecto a sus coeficientes introducidos en LAMMPS se detallan en la Subsección 3.4 y en el Anexo C.

4 Resultados & Análisis

La primera sección de la presente sección de resultados y análisis consiste en el estudio realizado con APBS para la parametrización del modelo de grano grueso, incluyendo el estudio de malla realizado. La segunda sección, por otra parte, contemplan los casos estudiados en LAMMPS del modelo parametrizado.

4.1 Parametrización

4.1.1 Estudio de Malla

En primer lugar, se realizó un estudio de malla en APBS de tal manera de establecer el balance adecuado entre la fidelidad numérica y los recursos computacionales disponibles. Para esto, se realizaron varias simulaciones de un mismo caso en APBS con diferentes tamaños de malla para la resolución de la ecuación lineal y no-lineal de Poisson-Boltzmann con cuyos resultados se realizó una extrapolación de Richardson para ambos métodos de tal manera de estimar el porcentaje de error por la discretización según el procedimiento indicado por Feldman (2021). El mallado escogido como resultado de este estudio de malla fue utilizado en todas las simulaciones realizadas en APBS.

Para el estudio de malla se utilizó el caso que requiera el mayor nivel de detalle en el mallado: ARN más proteína VP3 adosada. Se escogió este caso debido a que no solo contiene la elevada carga del ARN, sino que también los detalles de la estructura de la proteína requieren de un mallado tal que puedan ser capturados fielmente. Como parámetro de comparación se optó por la energía de solvatación $G_{solv.}$.

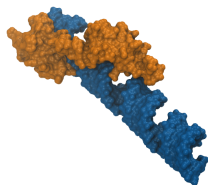


Figura 19: Visualización del caso empleado para el estudio de malla.

La extrapolación de Richardson requiere que las mallas se relacionen mediante una razón, la cual es utilizada para estimar el límite que tomaría la variable con una malla infinitamente fina siguiendo lo delineado en la Ecuación 24 y en la Ecuación 25.

$$f_{h=0} \approx f_1 + \frac{f_1 - f_2}{r^2 - 1} \quad (24)$$

Donde $f_{h=0}$ es un estimado de la variable calculada con una malla infinitamente fina, f_1 corresponde a la variable bajo la malla más fina de la secuencia y f_2 la variable bajo la malla más gruesa.

Adicionalmente a estimar el comportamiento de la variable en el límite del mallado, es posible estimar el orden de convergencia estimado a partir de los resultados de tres mallas consecutivas, según la Ecuación 26. Este orden de convergencia observado difiere del teórico del método

numérico en este caso debido a las condiciones de borde utilizadas, y se ha observado en la práctica que es cercano a 1.

$$\Delta_i = \frac{\Delta_0}{r^i} \quad (25)$$

$$p = \ln\left(\frac{f_3 - f_2}{f_2 - f_1}\right) / \ln(r) \quad (26)$$

Una vez definido el caso de estudio, se definieron las mallas a estudiar para lo cual se optó por la secuencia delineada a continuación en el Cuadro 7 con $r = 5/3$.

Cuadro 7: Secuencia de mallas escogida con $r = 5/3$.

N	i	r^i	Δ [Å]
4	0,0	1,000	0,500
3	0,5	1,291	0,387
2	1,0	1,667	0,300
1	2,0	2,778	0,180

Cabe destacar que, debido a limitaciones computacionales, el mallado de 0,180 [Å] para el método lineal falló, por lo que los cálculos de extrapolación y estimación del orden de convergencia se realizaron con las mallas $\Delta = \{0, 500; 0, 387; 0, 300\}$ a modo de realizar una comparación fiel entre ambos métodos computacionales, a pesar de que el método no-lineal se haya ejecutado exitosamente en la malla de 0,180 [Å].

Resultados

Se puede observar en el Cuadro 8 que el orden de convergencia para las mallas indicadas es cercano a 1 —menor que el orden del método numérico el cual es 2— debido a lo anteriormente mencionado por lo que se encuentra dentro de los valores esperados para APBS según resultados experimentales.

Cuadro 8: Orden de convergencia observado para la secuencia $\Delta = \{0, 500; 0, 387; 0, 300\}$.

Mallado [Å]	Orden Observado	
	PB No-Lineal	PB Lineal
0,500 - 0,387 - 0,300	1,058	1,210

Por otra parte, se puede observar que el error relativo indicado en el Cuadro 9 disminuye conforme la malla se hace más fina de la forma esperada, y que el ultimo mallado disponible para ambos métodos —lineal y no-lineal— de 0,300 [Å] posee un error relativo ubicado entre 1,821% y 0,803%. Debido a que se realizarán comparaciones entre métodos lineales y no lineales, se escoge esta última malla disponible para ambos. En otras palabras, **el mallado escogido corresponde a 0,300 [Å] con un error estimado de 1,821%.**

Finalmente, observando la Figura 20 se puede observar que el método no-lineal calcula una energía de solvatación consistentemente inferior en magnitud al método lineal, condición que se repite para su correspondiente extrapolación de Richardson. Esto es esperable debido a los altos niveles de carga del ARN, puesto que si se observa la Ecuación 4 esta posee un seno hiperbólico el cual es linealizado para la ecuación lineal. Esto implica que, debido a que el seno hiperbólico

Cuadro 9: Resultados del estudio de malla con error relativo respecto a la extrapolación de Richardson correspondiente.

Δ [Å]	$G_{solv.}$ [kJ/mol]		Err. Rel (Richardson)	
	PB No-Lineal	PB Lineal	PB No-Lineal	PB Lineal
0,180	-11862,991	ERR	0,493%	ERR
0,300	-12019,792	-14460,106	1,821%	0,803%
0,387	-12173,962	-14558,605	3,127%	1,489%
0,500	-12438,689	-14741,374	5,370%	2,763%

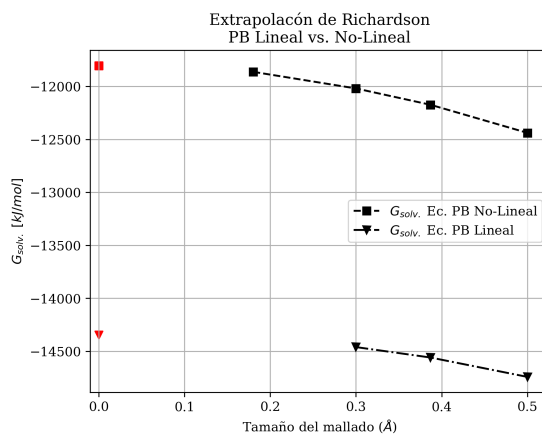


Figura 20: Resultados del estudio de malla con extrapolación de Richardson para APBS lineal y no-lineal en rojo.

crece a mayor velocidad que la recta $y = x$, para valores elevados de carga como las presentes en la columna fosfatada del ARN, la ecuación lineal estimará un nivel de energía inferior. En conclusión, se estableció en base a estos resultados que el método a utilizar como caso base para la parametrización será el método no-lineal de APBS con un mallado de 0,300 [Å].

4.1.2 Simulación Terminal-N VP3

Se realizó una simulación exploratoria del terminal-N de la proteína VP3 del complejo con los parámetros de malla definidos de tal manera de dilucidar la manera de realizar un modelo de grano grueso para dicha interacción.



Figura 21: Dímero de VP3 (a) y Terminal-N (b) de la proteína simulado. Cabe destacar el entrelazamiento entre proteínas.

Sin embargo, como se puede observar en la Figura 21, ambas proteínas se encuentran entrelazadas de tal manera que resulta imposible separarlas limpiamente. Se simuló la energía de interacción separando en las dos direcciones en torno al eje que une los centros de masa de ambos terminales, sin embargo, no se lograron resultados concluyentes por lo que el terminal-N así como también el *core* de la proteína VP3 fueron descartados para efectos del modelo de grano grueso. Este resultado implica que los cambios conformacionales de las proteínas momentos antes de realizar la unión pueden ser extremadamente relevantes, considerando que las simulaciones realizadas por APBS son rígidas y por lo tanto no reflejan el cambio de forma de las proteínas al interactuar, por lo que la forma y estructura del **pdb** a simular requiere una vía de separación entre ambas moléculas despejada.

4.1.3 Exploración de Alternativas de Grano Grueso

En primer lugar, una vez definido el mallado y el método base a emplear como punto de comparación, se exploraron alternativas respecto al modelo de grano grueso a utilizar, incluyendo comparaciones respecto al efecto de la geometría y distribución de carga en APBS a modo explorar si estos parámetros correspondientes a un modelo de grano grueso tienen un impacto en la solución mediante el método de solvente implícito.

Para establecer una comparación entre las alternativas de grano grueso se optó por el caso límite en cuanto a la energía de interacción entre dos cadenas de ARN: dos hebras paralelas. Al mantener las cadenas fosfatadas enfrentadas a lo largo de toda la longitud de la hebra, la energía de interacción contenida es máxima. Luego, para efectos de la parametrización es necesario dilucidar la curva de energía de interacción entre ambas hebras de ARN a diferentes distancias, por lo que se simuló ambas hebras a 9 distancias diferentes abarcando un rango entre ≈ 5 [Å] hasta ≈ 40 [Å]. En esta primera etapa, se estudiaron las curvas de interacción entre ambas hebras de ARN para diferentes longitudes de nucleótidos, o en su defecto, esferas de grano grueso. Debido a que la energía de interacción aumenta conforme aumenta la longitud de la cadena, se normalizó la energía de interacción por N^2 donde N es la cantidad de esferas de grano grueso equivalentes a la longitud de ARN en cuestión. En otras palabras, N equivale a dividir la cantidad de nucleótidos por 5 puesto que cada esfera de grano grueso agrupa a 5 pares de nucleótidos según lo indicado en la Subsección 3.4. De esta manera, es posible comparar el comportamiento de la energía de interacción conforme aumenta la longitud de la cadena y por ende observar si el comportamiento de la parametrización de grano grueso es consistente.

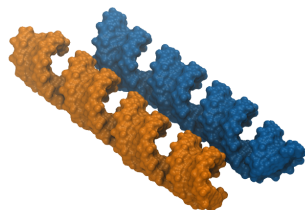


Figura 22: Visualización del caso de estudio utilizado para exploración de alternativas de grano grueso. Elaboración propia.

Resultados

En primer lugar, se simuló los casos *all-atom* —esto es, con resolución atómica— para APBS Lineal y No-Lineal y el potencial de Yukawa dando como resultado lo indicado en la Figura 23. Se observa que las energías de interacción son positivas, lo que implica una interacción repulsiva. Esto es esperable debido a que el ARN contiene una fuerte carga negativa por lo que existe una repulsión entre ellas. Luego, se puede observar que, para las 4 longitudes de cadena estudiadas, APBS Lineal sobrestimó el nivel de energía de interacción respecto a APBS No-Lineal, lo cual, según lo indicado en la Subsubsección 4.1.1, resulta esperable debido al alto nivel de carga presente en el sistema. Esto a diferencia del potencial de Yukawa cuya energía de interacción se asemeja a la energía dada por APBS No-Lineal para hebras mayores a 5 pares de

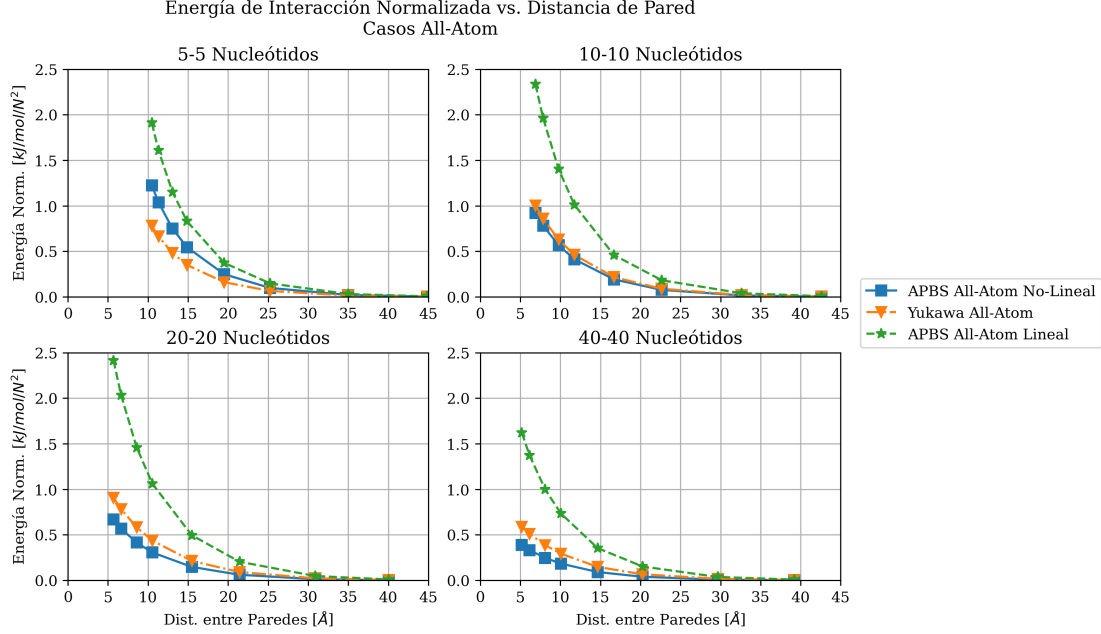


Figura 23: Resultados para los modelos *all-atom* calculados: Yukawa y APBS Lineal/No-Lineal.

nucleótidos de longitud. Esto implica que, a pesar de que el potencial de Yukawa es un método sencillo que surge de la solución a la ecuación lineal de Poisson-Boltzmann, en los casos vistos entregó resultados comparables al método más detallado, aunque sobrestimando ligeramente la energía de interacción.

Por otra parte, se puede observar una cierta simetría en los casos APBS Lineal: el caso de 5 pares de nucleótidos se asemeja al caso con 40 pares de nucleótidos mientras que el caso con 10 pares se asemeja al caso con 20 pares. Esto se puede deber a la geometría resultante al truncar las hebras: al tener una geometría helicoidal, es posible que el mallado resultante entre 5 pares de nucleótidos se asemeje al de 40 pares, pero no al de 10 pares, siendo este último semejante al caso con 20 pares en su lugar. Es, por lo tanto, esperable que este patrón se repita conforme se alargue la hebra más allá de los casos vistos. Sin embargo, este comportamiento no se observa con fuerza en los resultados entregados por APBS No-Lineal y el potencial de Yukawa donde se observa que la energía normalizada por la longitud disminuye conforme se alargan las hebras. Esto implica que la energía de interacción entre las hebras aumenta de manera no lineal conforme se alargan las hebras. Este fenómeno no ocurre para los resultados entregados por APBS Lineal, lo que sugiere que este efecto se relaciona con la no-linealidad de la ecuación de Poisson-Boltzmann, y que además puede ser capturado en cierta manera por el potencial de Yukawa.

Luego, se comparó el efecto de la geometría sobre los resultados entregados por APBS, indicado en la Figura 24. Para esto, se simuló dos casos de grano grueso en APBS: *FullGeometry* y *SimpleGeometry* —detallados en la Subsección 3.4. El primero consiste en mantener la geometría *all-atom* de las moléculas, pero concentrando las cargas en la misma ubicación que el centro de la esfera de grano grueso, mientras que el segundo consiste en reemplazar toda la geometría y cargas por esferas.

Se puede observar que el efecto de la geometría es relevante a la hora de realizar el cálculo en

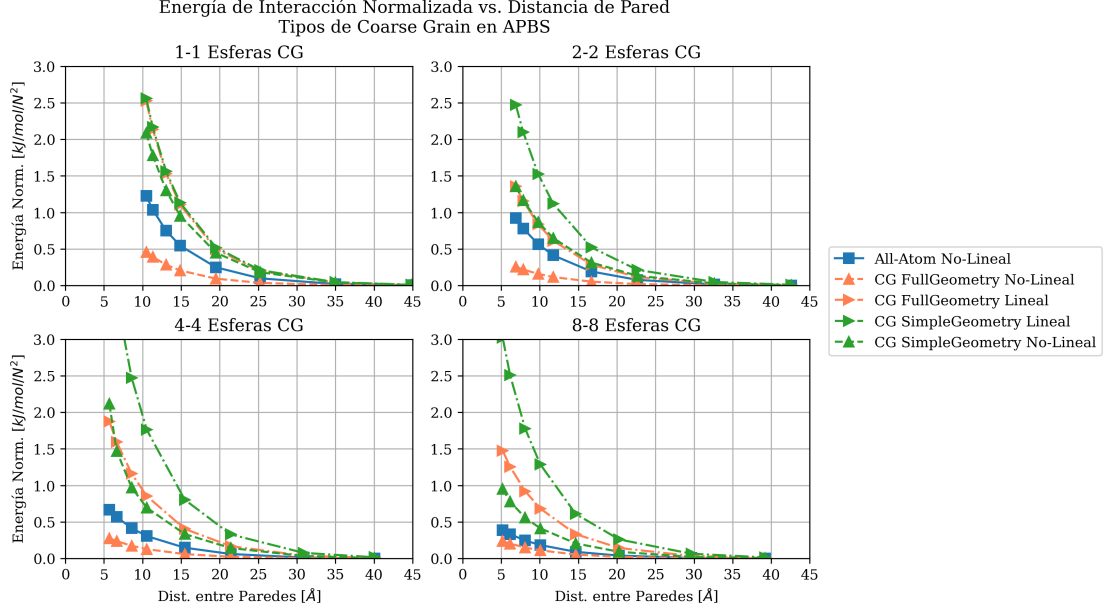


Figura 24: Comparación de resultados de APBS para *Coarse Grain* Yukawa y APBS.

APBS puesto que todos los casos con geometría simplificada sobrestimaron considerablemente la energía respecto al caso base de APBS No-Lineal *All-Atom*. En cuanto a los casos con geometría completa y cargas concentradas, se obtuvieron resultados algo más cercanos al caso base, aunque esta vez subestimando la energía. Esto sugiere que el efecto de la geometría en el mallado es relevante para efectos de la energía de interacción aún más que la no-linealidad de la ecuación. De todas maneras, como punto de comparación, se escogió el caso con geometría completa y cargas concentradas como representación de realizar el grano grueso en APBS en sus formas lineal y no-lineal al ser presentar estos la menor diferencia con el caso base. Finalmente, se realizó una comparación de todas las alternativas de grano grueso disponibles en la Figura 25, donde el grano grueso realizado en APBS corresponde a una geometría detallada y cargas concentradas.

Se puede observar con claridad que para hebras de más de 4 pares de nucleótidos de longitud el potencial de Yukawa con grano grueso genera resultados excelentes. Esto es sorpresivo puesto que implica que el modelo más sencillo: Yukawa con esferas de grano grueso, fue capaz de modelar con gran nivel de detalle el modelo más complejo: Poisson-Boltzmann No-Lineal en lo que respectan las energías de interacción para el caso visto. Esta tendencia se rompió para hebras más cortas inferiores a 4 pares de nucleótidos donde ninguno de los modelos de grano grueso caracteriza satisfactoriamente al caso base de forma consistente. Sin embargo, debido a que la parametrización utilizada será empleada en hebras largas de ARN —correspondientes a casos biológicos— esto no resulta dañino. Es por esta última razón, y los resultados indicados en la Figura 25, que se escoge el potencial de Yukawa como el mejor potencial de grano grueso para la energía de interacción entre dos hebras largas de ARN, respaldado por la simulación numérica de la ecuación de Poisson-Boltzmann No-Lineal.

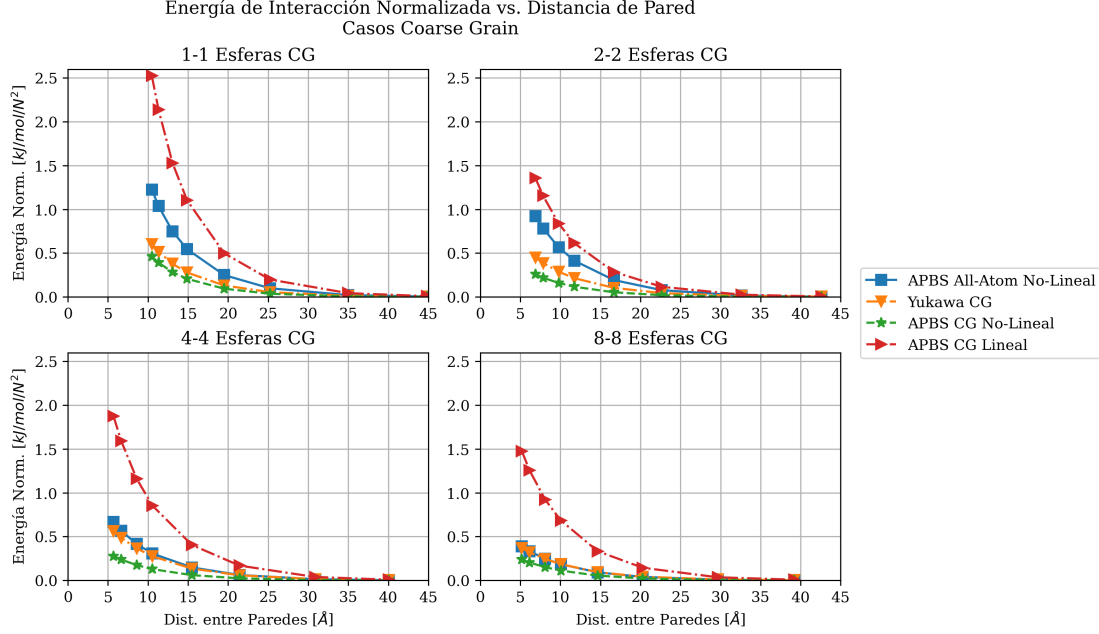


Figura 25: Resultados para tipos de *Coarse Grain* en APBS.

4.1.4 Estudio del Potencial de Yukawa

Una vez establecido el potencial de Yukawa como el utilizado para el modelo de grano grueso se siguió con un estudio detallado de las cotas de error de este para los casos esperables de interacción entre dos hebras de ARN. Con esto en mente se definieron dos orientaciones de estudio para las hebras: paralelas y cruzadas, llamadas orientaciones P y N respectivamente, indicadas en la Figura 26. De esta manera se logra acotar la diferencia entre la energía de interacción entregada por el modelo de grano grueso y la entregada por la simulación *all-atom* en APBS según la orientación relativa de las hebras.



Figura 26: Visualización de las orientaciones “P” (a) y “N” (b).

Por otra parte, debido a que la estructura del ARN del virus presenta una doble hélice, según lo indicado en la Subsubsección 2.1.3, y que por lo tanto su superficie varía a lo largo de su longitud, se simuló cuatro rotaciones a lo largo de su eje principal. Así se modeló la interacción entre ambas cadenas de ARN a diferentes niveles de desfase puesto que, debido a la geometría helicoidal, una rotación completa en su eje equivale a modelar por separado distintos segmentos

de una cadena de mayor longitud; capturando así por completo las rugosidades superficiales de la molécula.

Cuadro 10: Casos simulados para el estudio del potencial de Yukawa. Donde ω representa una rotación en el eje mayor de la hebra móvil como se describió en la Subsección 3.5

	1-N Monómeros	N-N Monómeros
Orientación P	2 Hebras Paralelas $\omega = \{0, 90, 180, 270\}$ $N = \{1, 2, 4, 8\}$	2 Hebras Paralelas $\omega = \{0, 90, 180, 270\}$ $N = \{4, 8\}$
Orientación N	2 Hebras Cruzadas $\omega = \{0, 90, 180, 270\}$ $N = \{1, 2, 4, 8\}$	2 Hebras Cruzadas $\omega = \{0, 90, 180, 270\}$ $N = \{4, 8\}$

Finalmente, para las orientaciones P y N, se varió la longitud de la hebra de ARN móvil, primero manteniendo la hebra fija a una longitud de 1 monómero —5 nucleótidos— mientras que la hebra móvil varió entre 1, 2, 4 y 8 monómeros de largo. Este caso se denominó “one-to-many” o 1-N. Luego, se simuló con dos hebras de igual longitud de 4 y 8 monómeros recibiendo este caso el nombre “many-to-many” o N-N. De esta manera, se repartieron las simulaciones del estudio: 2 orientaciones —P y N— con dos casos cada una —1-N y N-N. Donde, además, cada orientación contempla la rotación de la hebra móvil sobre su eje principal como si mencionó con anterioridad. Esto queda resumido en el Cuadro 10. Cabe destacar que las energías de interacción para los casos 1-N monómeros fueron normalizadas por N para realizar comparaciones a diferentes longitudes de cadena.



Figura 27: Caso cruzado “many-to-many”, dos hebras de ARN de igual longitud (a) y “one-to-many”, una hebra de ARN corta y otra larga (b). Elaboración propia.

Resultados 1-N Monómeros

En primera instancia se presentan los resultados de las energías de interacción según la distancia entre pared del caso 1-8 monómeros en orientación paralela en la Figura 28. Debido a la gran cantidad de figuras, se presentan los resultados de las energías de interacción según distancia para los casos 1-1, 1-2, 1-4 monómeros en ambas orientaciones y para el caso 1-8 monómeros en orientación cruzada en el Anexo D. Se puede observar en la que el modelo de grano grueso con el potencial de Yukawa bajo los parámetros de carga definidos en la Subsección 3.4 y la distancia de Debye definida en la Subsubsección 4.1.3, entrega una muy buena aproximación de la energía de interacción calculada mediante APBS para el rango estudiando —con la diferencia máxima presentándose para $\omega = 0^\circ$.

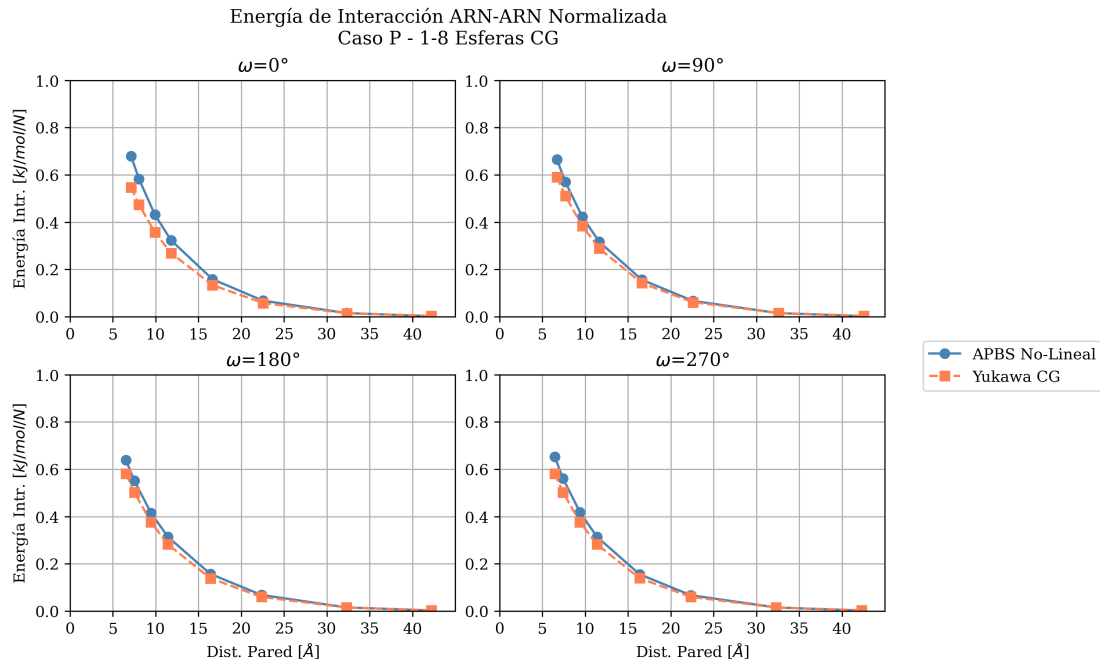


Figura 28: Energía de interacción normalizada para el caso 1-8 monómeros.

El gráfico indicado corresponde al caso 1-8 monómeros, sin embargo, se observó en los casos 1-1, 1-2 y 1-4 monómeros un comportamiento similar, entregando el modelo de grano grueso buenos resultados con excepción del caso 1-1 monómeros. Esto último se puede deber que, al tratarse de dos cadenas de longitud muy pequeña, las rugosidades superficiales, que son simplificadas por el modelo de grano grueso pero que, si son consideradas por APBS, toman mayor importancia.

Por otra parte, se puede observar en la Figura 29 que la diferencia absoluta entre la energía de interacción presenta leves diferencias conforme se trate de la orientación paralela o cruzada de las hebras, con la primera presentando la mayor diferencia. Esto es esperable puesto que las hebras en su orientación paralela poseen mayor energía de interacción respecto a la orientación cruzada, sin embargo, el comportamiento de dicha diferencia es consistente respecto a la orientación de la hebra alcanzando un máximo de aproximadamente 1,5 [kJ/mol] a una distancia de pared de 6 [Å] y $\omega = 0^\circ$. Observando la Figura 28 se observa que la energía de interacción normalizada según APBS para ese punto alcanza un valor de 0,7 [kJ/mol/N], la cual, al escalar por N, se

aproxima a 5,6 [kJ/mol]. Por consiguiente, se estima una diferencia máxima relativa en 27% a una distancia de 6 [Å]. cabe destacar que este escenario de 1-8 monómeros no corresponde a un caso físico real, sino que su objetivo es el estudio del comportamiento de la energía a diferentes longitudes de hebra sin considerar las interacciones cruzadas y determinar su consistencia.

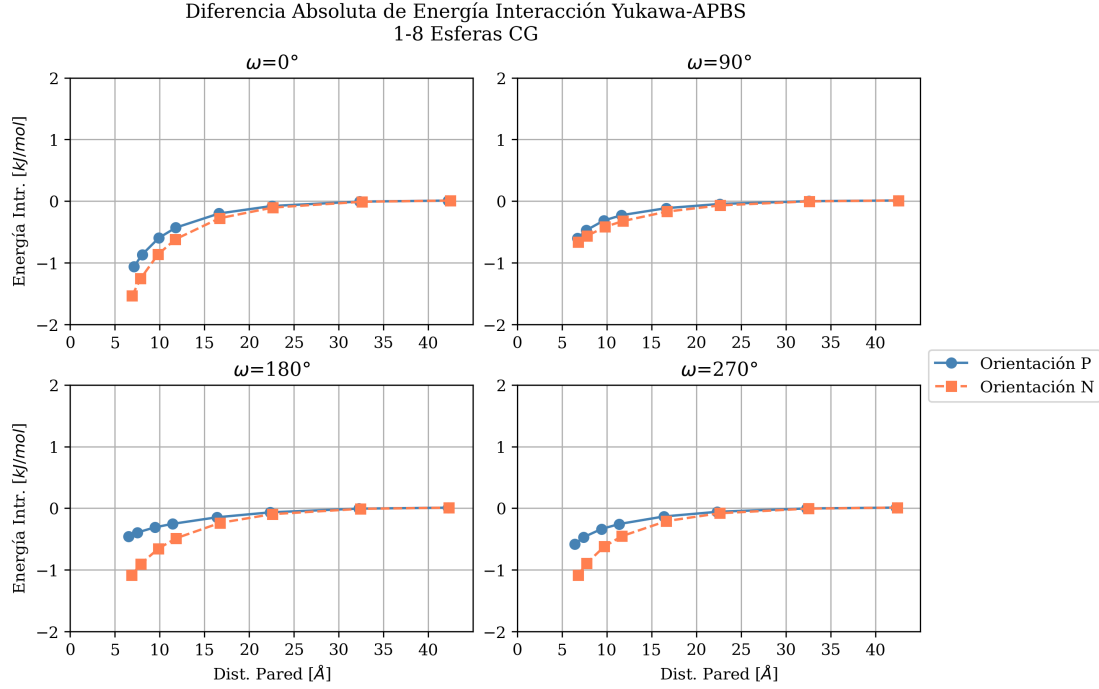


Figura 29: Diferencia Yukawa-APBS normalizada para el caso 1-8 monómeros.

Finalmente, se puede observar el comportamiento general de la diferencia entre el modelo de grano grueso y lo entregado por APBS a diferentes longitudes de cadena y ambas orientaciones en la Figura 30, donde el rango de la caja corresponde al rango en donde se encuentra la diferencia total para las 4 rotaciones en ω .

Se puede observar con claridad que la diferencia total absoluta se mantiene aproximadamente constante a medida que aumenta la longitud de la cadena, pero que al normalizar por la longitud de esta —y por ende considerar el hecho que el aumentar el largo de la hebra aumenta la energía de interacción— la diferencia cae abruptamente conforme la cadena se hace más larga. Adicionalmente, se puede observar en la subfigura (b) que el rango del caso 1-1 es mayor lo que implica que la rugosidad superficial de las hebras de ARN tiene un gran impacto sobre la energía de interacción para estos casos con hebras cortas, especialmente si se trata del caso paralelo.

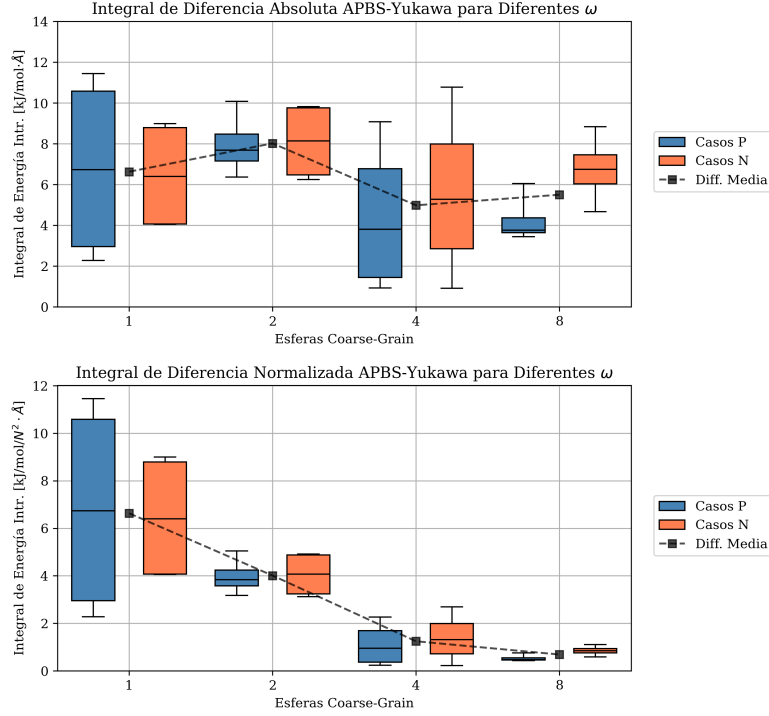


Figura 30: *Boxplot* de la diferencia total absoluta (a) y normalizada (b) sobre todas las rotaciones ω y ambas orientaciones para el caso 1-N monómeros.

Resultados N-N Monómeros

Observando la Figura 31 se puede notar que la energía de interacción de Yukawa coincide notablemente con la energía calculada por APBS. A pesar de la simpleza del modelo, este logra describir con gran certeza la interacción entre ambas hebras de ARN de igual longitud donde la diferencia máxima, correspondiente a $\omega = 270^\circ$ como se indica en la Figura 32, es de 4 [kJ/mol]. Considerando que la energía de interacción total en dicha ubicación corresponde a ≈ 31 [kJ/mol], se tiene una diferencia relativa máxima en torno al 13% —la cual decae rápidamente a distancias ligeramente mayores.

De forma similar a la sección anterior, las figuras correspondientes a los casos 4-4 monómeros y la energía de interacción del caso 8-8 monómeros en orientación N, se presentan en el Anexo D. De todas maneras, se puede observar en la Figura 32 que la energía calculada por el potencial de Yukawa para la orientación N posee menor se asemeja aún más a APBS. Esto es esperable puesto que, como se mencionó con anterioridad, el caso paralelo enfrenta las superficies de las hebras de ARN sobre una mayor longitud por lo que los efectos de la rugosidad superficial —perdidos por el modelo de grano grueso— tienen una mayor preponderancia sobre la energía de interacción.

Luego, observando la Figura 33 se puede notar un comportamiento similar a lo visto en la Figura 30, donde la diferencia de energía total absoluta mantiene un nivel relativamente constante y donde el efecto de las rotaciones en ω es menor conforme la hebra de ARN aumenta su longitud. Por otra parte, al normalizar la diferencia en energía de interacción por N^2 —es decir por el total de interacciones entre monómeros— se puede notar que esta diferencia disminuye notablemente al trabajar con hebras de 8 monómeros de largo junto al efecto de las rotaciones en ω .

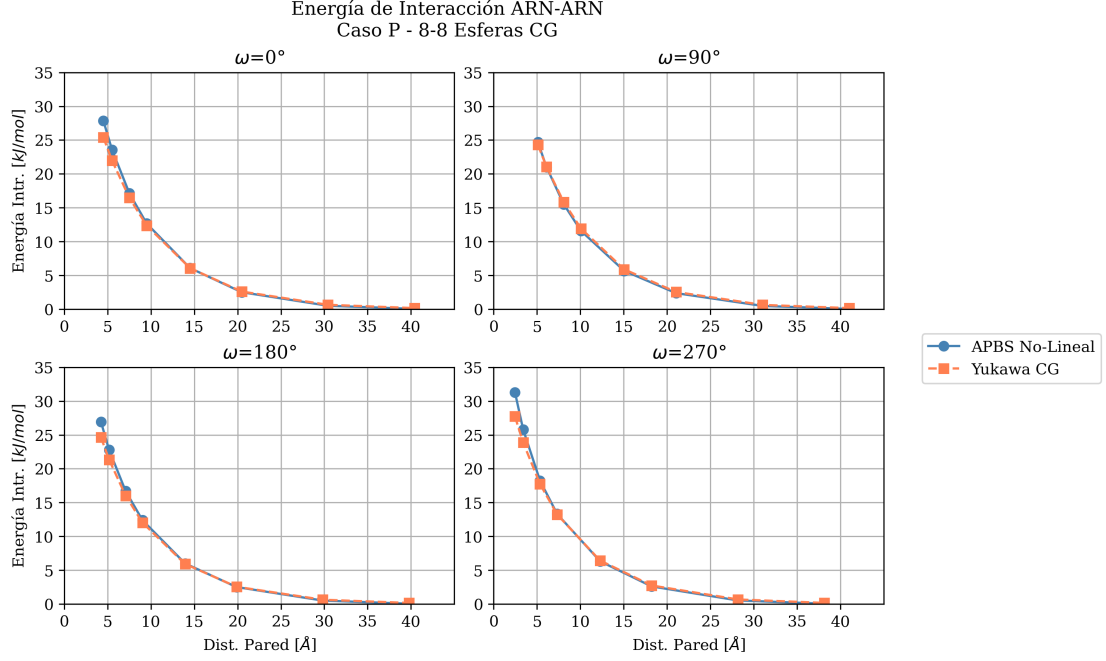


Figura 31: Energía de interacción absoluta para el caso 8-8 monómeros en orientación P.

En base a estos resultados, y a lo visto en la sección anterior, es posible concluir que a medida que aumenta el largo de la hebra, los efectos de rugosidad superficial de las hebras de ARN se vuelven despreciables implicando que el modelo de grano grueso es capaz de capturar la mayor parte de la interacción entre hebras de ARN para cadenas de gran longitud.

Por otra parte, es posible estimar bajo que distancia se puede considerar una hebra infinitamente larga bajo los parámetros de Yukawa establecidos, para esto, se puede comparar la interacción entre dos monómeros a la mínima distancia simulada con la interacción entre los mismos monómeros a la longitud de las hebras de ARN.

$$E \propto \frac{e^{-\kappa r_0}}{r_0} \quad (27)$$

Utilizando $r_0 = 5 \text{ [Å]}$ y $\kappa = 0,1273 \text{ [1/Å]}$, se obtiene un valor de $E \propto 2,6$. Luego, considerando que las hebras poseen una longitud de $\approx 100 \text{ [Å]}$, se obtiene un valor de $E \propto 3,0 \cdot 10^{-4}$. Esto implica que la interacción entre los extremos de las hebras es del orden de 10.000 veces menor que la máxima energía simulada y, por lo tanto, despreciable. Por consiguiente, es posible considerar que para hebras de 8-8 se está bajo el régimen de hebra infinitamente larga de tal manera que los resultados vistos respecto al comportamiento del modelo de grano grueso debiesen mantenerse para hebras de una longitud arbitrariamente larga.

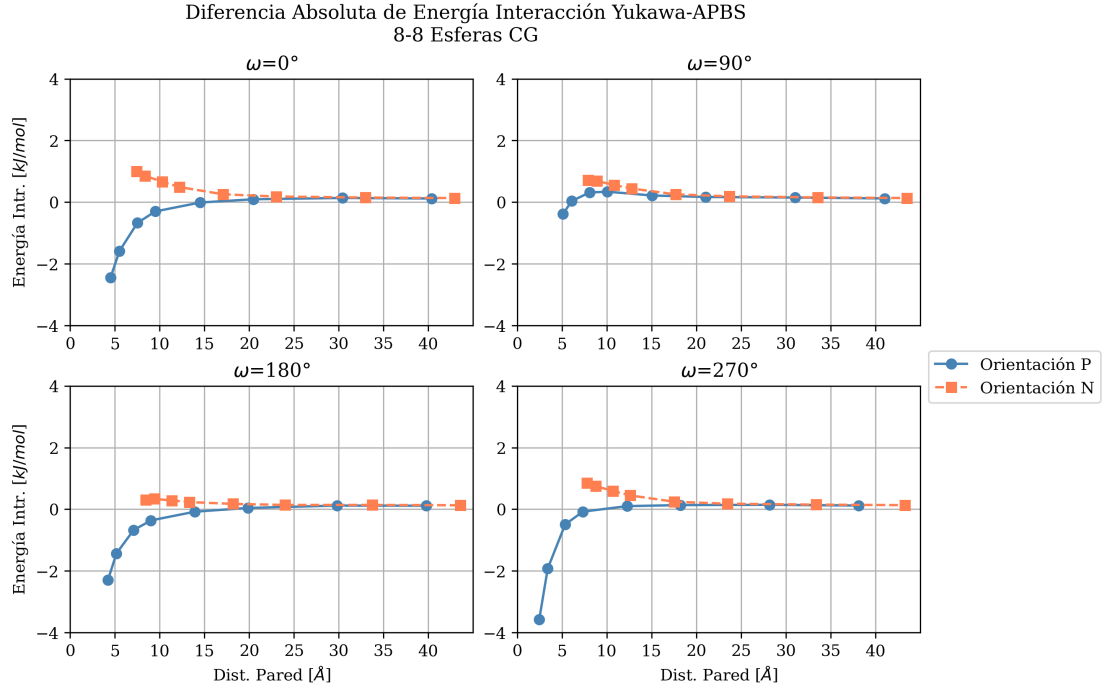


Figura 32: Diferencia absoluta entre las energías de interacción calculadas por Yukawa y APBS para el caso 8-8 monómeros.

Finalmente, se puede concluir que la distancia mínima de validez del modelo de grano grueso para orientaciones arbitrarias corresponde a la distancia de Debye —cercana a los 8 [nm]— obteniendo resultados similares a las simulaciones electrostáticas *all-atom* con APBS donde esta limitante corresponde a la distancia mínima bajo la cual resulta posible despreciar la componente apolar del solvente.

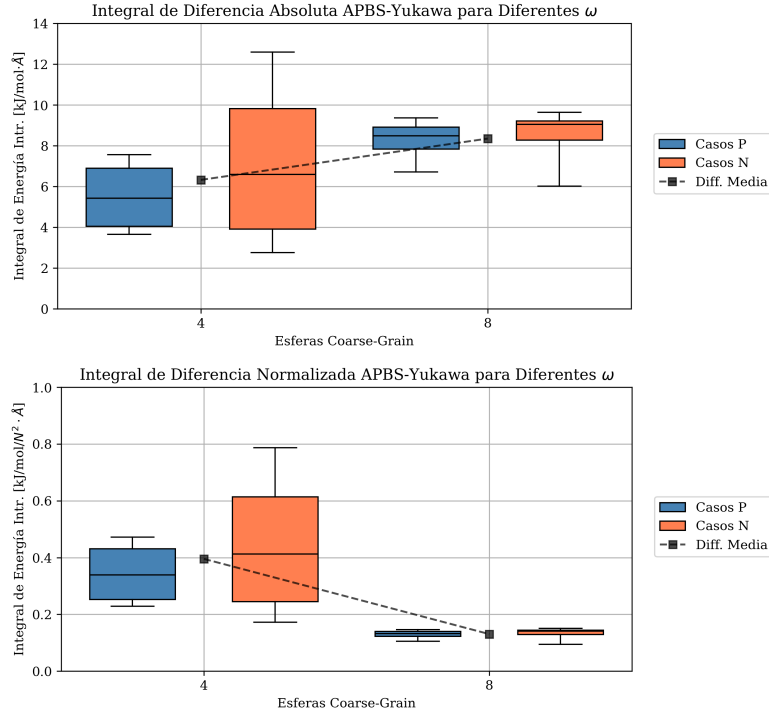


Figura 33: *Boxplot* de la diferencia total absoluta (a) y normalizada (b) sobre todas las rotaciones ω y ambas orientaciones para el caso N-N monómeros.

4.2 Simulaciones de Compresión de ARN en LAMMPS

Una vez caracterizada la interacción de Yukawa para el ARN se modeló la manera en que este se empaqueta al interior de la cápside del virus de tal manera de establecer si la parametrización dada entrega resultados consistentes con el nivel de compresión que presenta el ARN en el virus.

Por otra parte, cabe destacar que se convirtieron las unidades del potencial de Yukawa al sistema de unidades utilizado por LAMMPS, mediante el factor indicado en la Subsubsección 3.3.1, dando como resultado lo indicado en la Cuadro 11, utilizando un factor $\kappa = 1,273$ [nm] según lo indicado en la Subsección 3.5 y una distancia de corte de 6 [nm] puesto que, según lo visto en los resultados de APBS, la interacción entre las hebras decae a prácticamente cero a 4,5 [nm].

Cuadro 11: Coeficientes de Yukawa utilizados en LAMMPS

Interacción	Carga [$\bar{e}$]	Coeficiente LAMMPS [$kT \cdot nm$]
Bulk-Bulk	-10 & -10	70,91960
Bulk-Extremo	-10 & -9	63,82764
Extremo-Extremo	-9 & -9	57,44487

Como se describió en la Subsubsección 3.3.3, la compresión de las hebras es modelada mediante un muro potencial Lenard-Jones esférico cuyo radio se va reduciendo progresivamente, donde una vez alcanzado el tamaño final el muro se detiene en su lugar iniciando la fase de

relajación. Para todas las simulaciones de compresión realizadas se utilizaron los parámetros $\sigma = 10$ y $\epsilon = 5.0$ para el muro potencial. Estos parámetros fueron definidos arbitrariamente con el resguardo que no sacrifiquen la estabilidad de la simulación puesto que el muro se ajustó empíricamente tal que la posición radial final del polímero refleje los resultados vistos por Mata et al. (2018).

4.2.1 Compresión de Una Hebra

En primer lugar, se estableció un caso base de 1 hebra de ARN de 560 monómeros de largo, la cual es comprimida mediante la metodología descrita en la Subsubsección 3.3.3 hasta el tamaño biológico del virus —de 25 [nm] al interior de su cápside en base a los resultados entregados por Mata et al. (2018). En el Cuadro 12 se presenta un resumen de los parámetros utilizados. Cabe destacar que la longitud de hebra simulada corresponde al tamaño biológico del ARN del virus como se indica en la Subsubsección 2.1.3 correspondiente a 2800 pares de bases.

Cuadro 12: Parámetros utilizados para simulación de compresión para 1 hebra.

Parámetro	Valor
Longitud	560 monómeros
Pasos Compresión	34.500.000
Pasos Relajación	630.000.000
Paso Temporal	0,001

Resultados

En la Figura 34 se presentan los resultados de la etapa de compresión del ARN en el plano XY, donde se puede observar un aplastamiento de la hebra contra el muro potencial a medida que la hebra es comprimida. Esto puede ocurrir debido a que la velocidad de compresión resulta más rápida que el tiempo que requiere el ARN para relajarse, al no propagarse el efecto del muro hacia la sección media de la hebra. Se espera que esto no afecte los resultados pues la fase de relajación posterior a la compresión cumple el rol de propagar este efecto y así alcanzar un estado termodinámico probable. Se puede apreciar, además, el gran nivel de compresión que alcanza el ARN al interior de la cápside del virus al comparar el tamaño inicial del muro potencial con el final.



Figura 35: Renderización de la simulación individual en sus fases de compresión (a) y final de relajación (b).

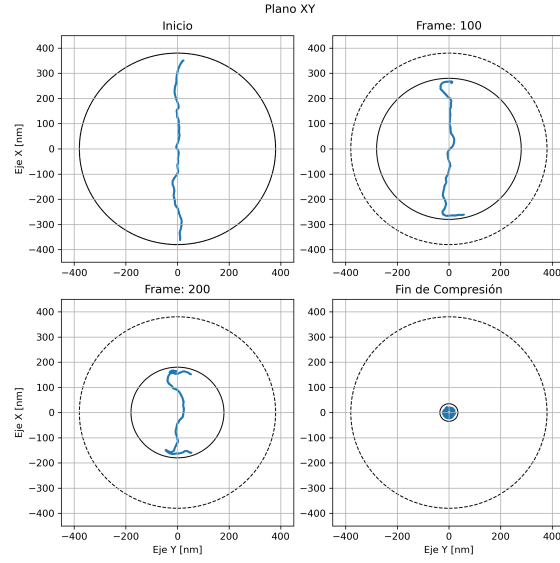


Figura 34: Secuencia de compresión para el caso individual. Elaboración propia.

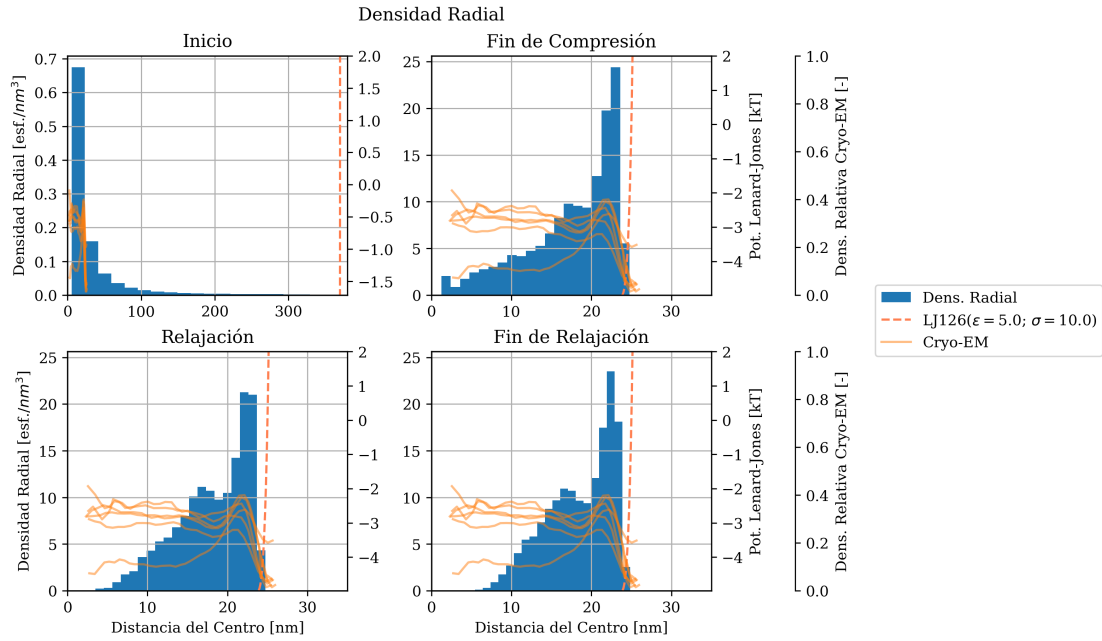


Figura 36: Densidad radial para el inicio, final de compresión, paso intermedio de relajación y final de etapa de relajación para el caso individual.

Luego, se puede observar el histograma de densidad radial para la presente simulación para el inicio y final de compresión, y para el punto medio y final de la etapa de relajación en la Figura 36. En primer lugar, se puede notar que la diferencia entre las dos capturas de la etapa de relajación es prácticamente nula, por lo que la simulación alcanzó un estado estacionario. En segundo lugar,

se observa que los monómeros de ARN se estabilizan en el borde del casquete dejando la región interior libre de este. Esto es una diferencia importante respecto a los resultados del Cryo-EM, a pesar de que incluye el perfil de densidad observado para un virión con 1 hebra de ARN. Esto se puede deber a que la simulación presente solo contempla el material genético del virus y no el resto de sus proteínas las cuales si podrían ser observadas por el Cryo-EM permitiendo llenar el vacío observado en el histograma. Por otra parte, se logra notar un pequeño abultamiento de monómeros en a una distancia aproximada de 18 [nm] del centro, los cuales forman un valle a una distancia de 20 [nm], para luego acumularse justo antes de la pared potencial. Esta acumulación y el pequeño valle previo se logra observar tenuemente en el Cryo-EM, donde se observa una mayor densidad cercano a los 25 [nm] y un tenue valle previo.

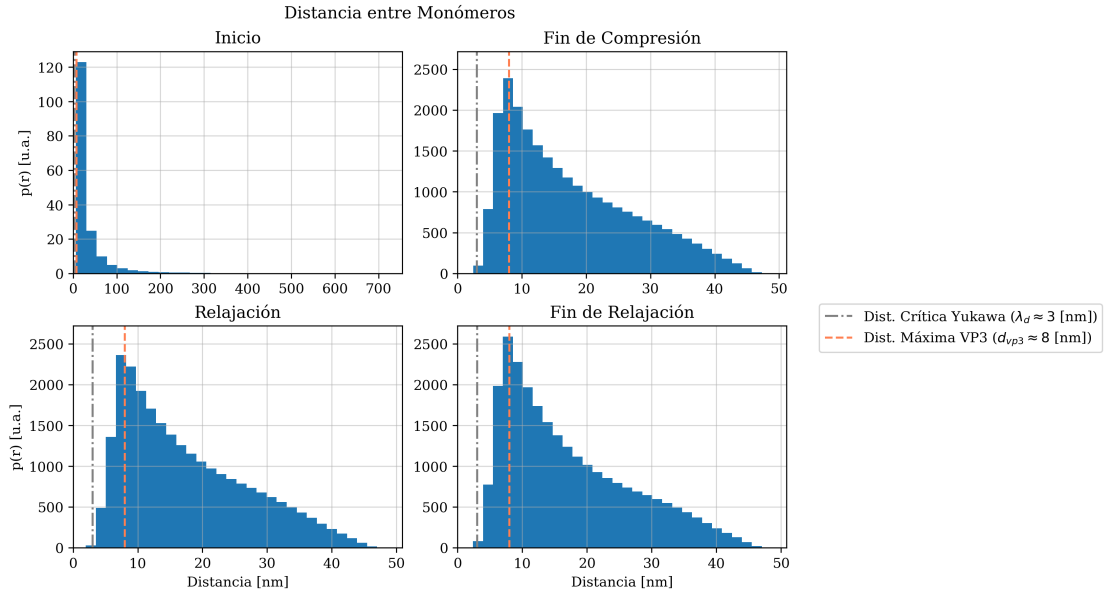


Figura 37: Histograma de distancia entre pares de monómeros excluyendo los 4 vecinos inmediatos para el caso individual.

Luego, se observa en la Figura 37 que todos los monómeros se encuentran más allá de la distancia crítica de Yukawa —establecida en la Subsección 4.1— lo que indica que el modelo de Yukawa para el presente caso es válido según lo establecido. Por otra parte, se observa que el grueso de los monómeros se encuentra a una distancia superior a 8 [nm], correspondiente a aproximadamente la distancia máxima de dimerización entre dos VP3, esto implica que existe espacio suficiente entre los monómeros producto de la electrostática como para admitir la presencia de VP3 entre las hebras según la simulación.

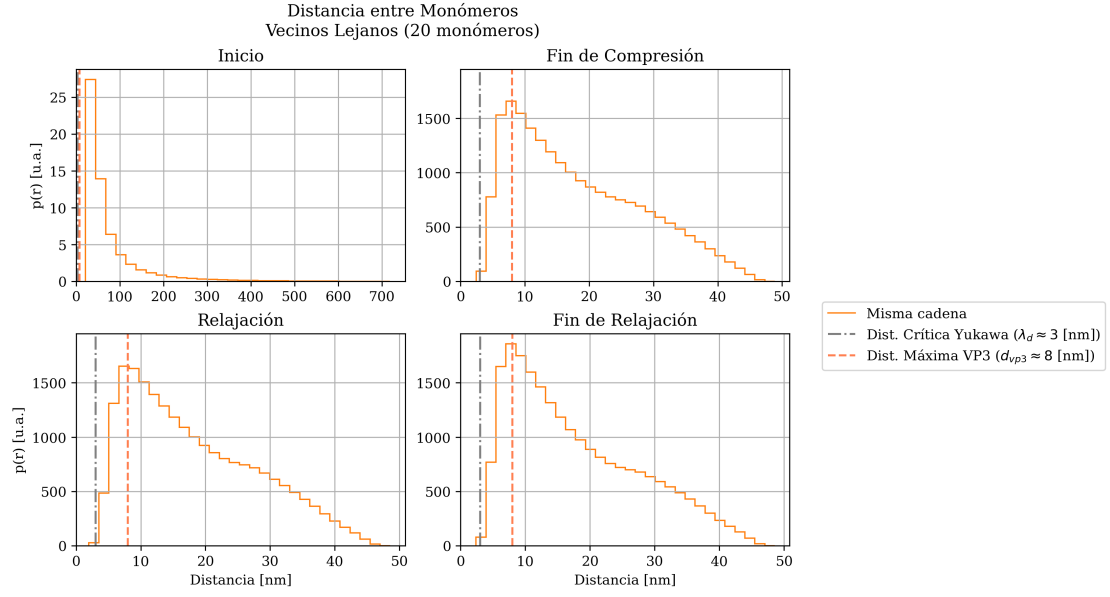


Figura 38: Histograma de distancia entre pares lejanos de monómeros, a contar del vecino numero 20 para caso individual.

Por otra parte, como se observa en la Figura 38, las distancias entre vecinos lejanos de ARN se distribuyen de manera similar a lo visto en la Figura 37 con una leve diferencia en el rango de 20 [nm].

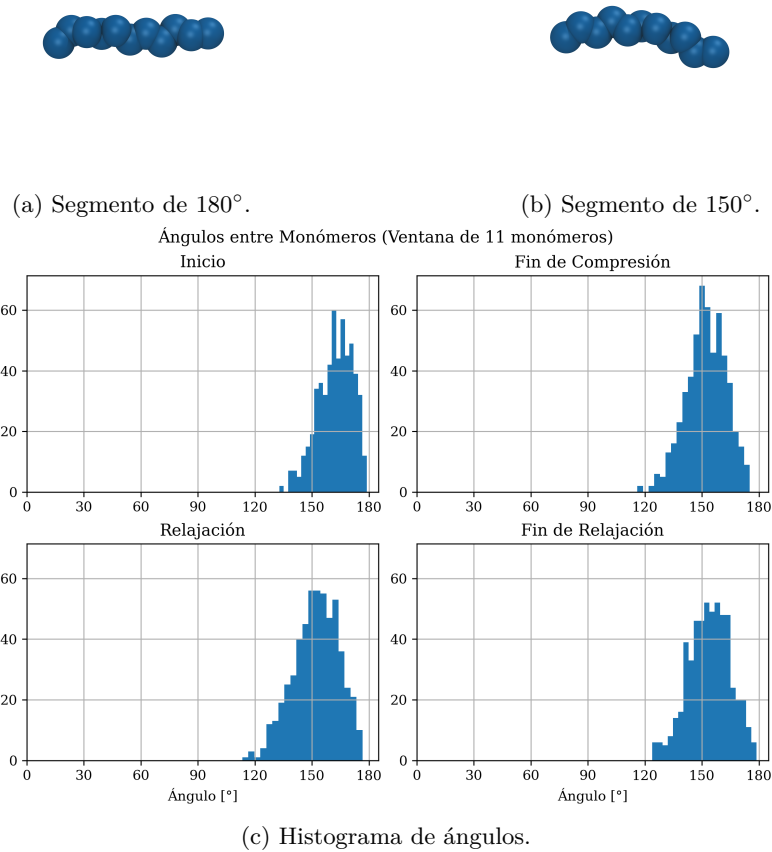


Figura 39: Histograma de ángulos para el caso individual (c) con ejemplos de segmentos de 180° (a) y 150° (b).

Luego, se observa en la Figura 39 que los ángulos del ARN se agrupan en torno a los 155° lo que implica que el ARN mantiene un nivel de curvatura apreciable dado el nivel de compresión como se indica en las subfigura (b).

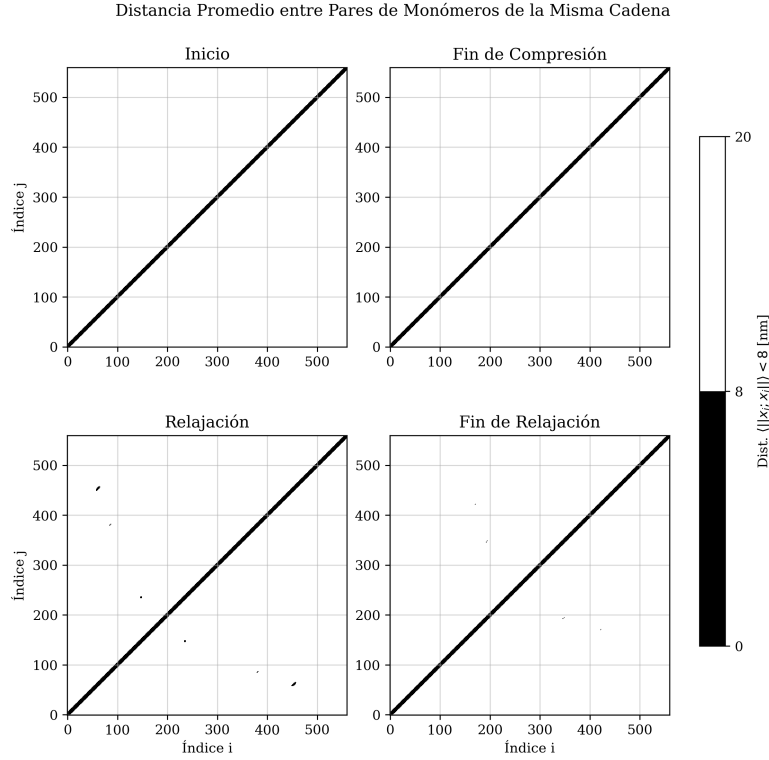


Figura 40: Densidad radial para el inicio, final de compresión, paso intermedio de relajación y final de etapa de relajación para el caso individual.

Finalmente, se puede observar en la Figura 40 que, para este caso, prácticamente no existen monómeros de la misma hebra que se encuentren en promedio a una distancia inferior a 8 [nm], esto podría implicar que, para los casos con una sola hebra de ARN, la proteína VP3 podría tener dificultades para dimerizarse con otros segmentos de la cadena. Sin embargo, cabe destacar que el efecto de las mismas proteínas sobre la electrostática podría jugar un rol importante que altere este resultado.

4.2.2 Compresión de Dos Hebras Paralelas

Se realizó una simulación de compresión de dos hebras inicialmente en paralelo para observar el comportamiento de los parámetros estudiados cuando aumenta la cantidad de hebras al interior de la cápside. Los parámetros de simulación se resumen en el Cuadro 13.

En este caso, fue necesario reducir el paso temporal en la etapa de relajación puesto que se encontraron problemas de estabilidad con un paso mayor debido a la rigidez del muro. A pesar de utilizar un paso temporal menor y una menor cantidad de pasos de relajación, la simulación alcanzó un estado estacionario satisfactoriamente.

Resultados

Cuadro 13: Parámetros utilizados para simulación de compresión para 2 hebras paralelas.

Parámetro	Valor
Longitud	560 & 640 monómeros
Pasos Compresión	41.000.000
Pasos Relajación	350.000.000
Paso Temporal Compresión	0,001
Paso Temporal Relajación	0,0001

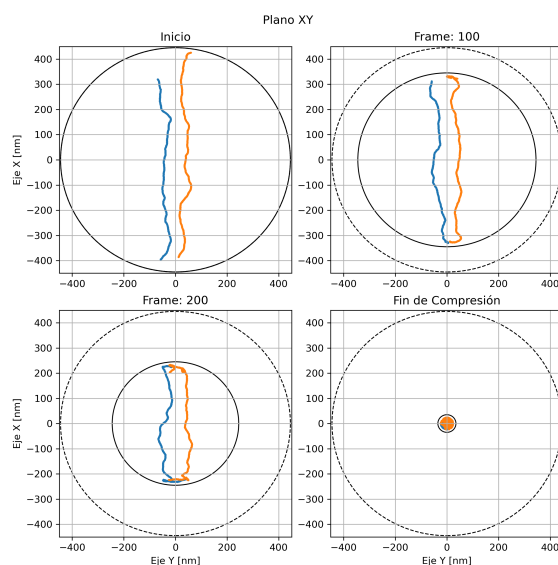


Figura 41: Secuencia de compresión para el caso doble en paralelo.

En primer lugar, se puede observar en la Figura 42 una estructura semejante a la del caso anterior, donde es notable la presencia del valle anterior al abultamiento de monómeros previo al muro potencial. Este valle, además, resulta más profundo y delimitado en este caso. Cabe destacar que, con dos hebras, los monómeros se extienden hasta el centro del casquete potencial asemejándose de mejor manera al resultado visto en el Cryo-EM, aunque se nota una tendencia a una menor densidad hacia el centro.

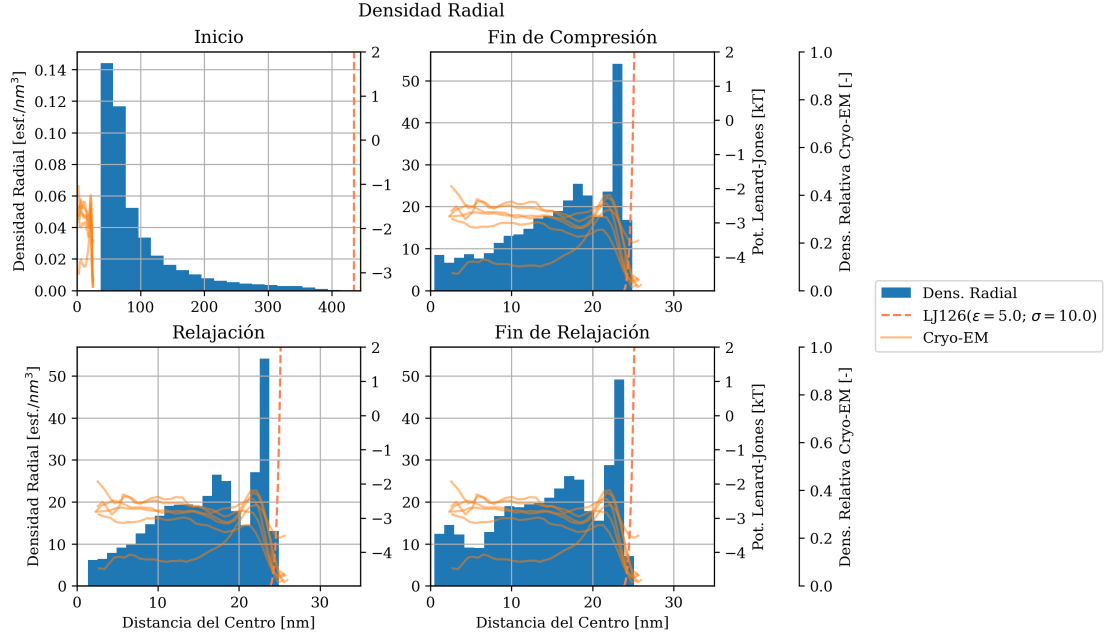


Figura 42: Densidad radial para el inicio, final de compresión, paso intermedio de relajación y final de etapa de relajación para el caso doble en paralelo.

En cuanto al histograma de ángulos se puede apreciar una tendencia similar al caso anterior, donde la mayoría de los ángulos que conforman las hebras corresponden a 150° , sin embargo, en el presente caso, la distribución se asemeja a simple vista de mejor manera a una distribución normal. Cabe destacar la diferencia entre las distribuciones entre el final de la compresión y el final de la relajación, donde en el primero se observa una campana ancha que pareciera aumentar su curtosidad al relajarse las estructuras. Esto se puede deber a que el ARN, al ser un polímero relativamente rígido, al relajarse adquiere su estado más favorable lo que reduce los ángulos extremos dada la compresión tanto sobre la media como por debajo de esta.

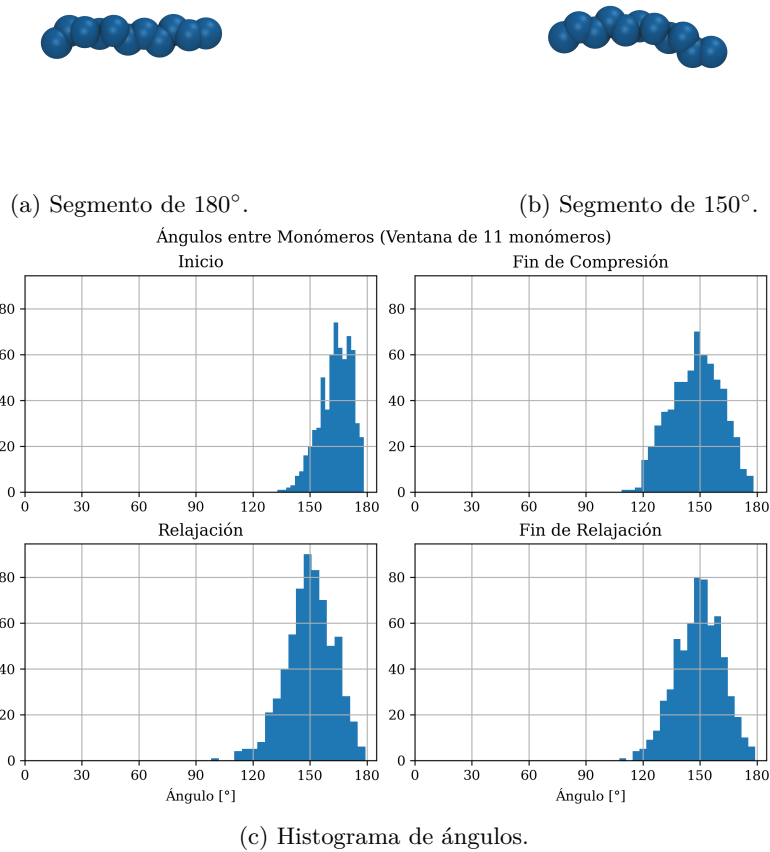


Figura 43: Histograma de ángulos para el caso doble en paralelo (c) con ejemplos de segmentos de 180° (a) y 150° (b).

Luego, se puede notar en la Figura 44 una distribución de distancias entre monómeros similar a el caso anterior, donde parte de los monómeros se encuentra a una distancia tal que admita la presencia de la proteína VP3 y donde se cumple el criterio de distancia para mantener las cotas establecidas para la interacción de Yukawa. Se puede observar, además, que existe una mayor presencia de monómeros entre 15 y 30 [nm] respecto al caso anterior, lo que se explica por la presencia de dos hebras de ARN lo que fuerza a algunos monómeros a mantener una mayor distancia entre si puesto que el abultamiento de monómeros en las paredes del muro potencial se vuelve inviable.

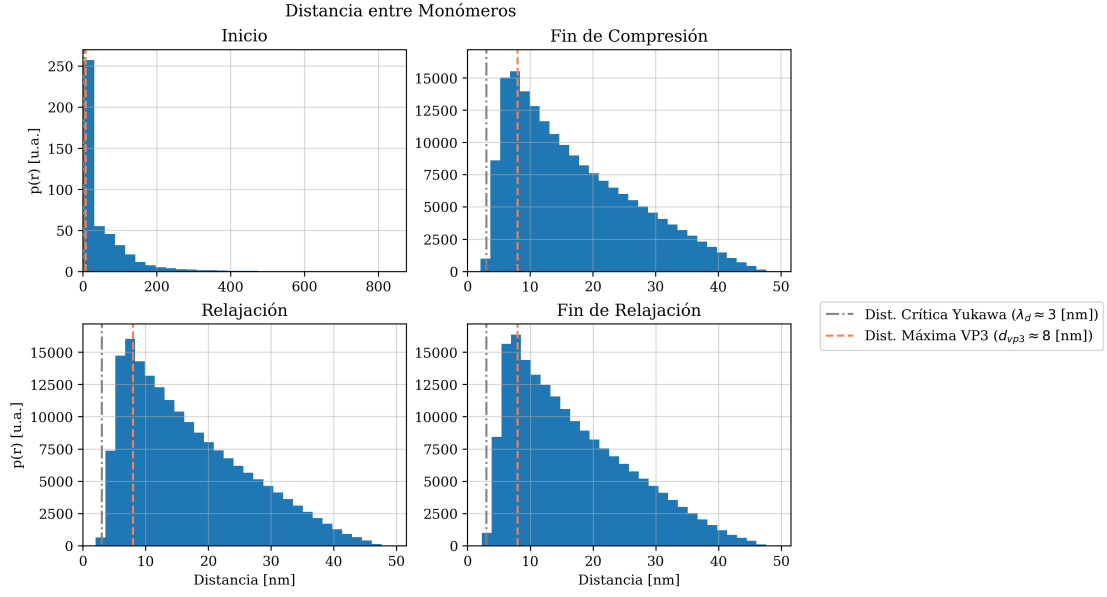


Figura 44: Histograma de distancia entre pares de monómeros excluyendo los 4 vecinos inmediatos para el caso doble en paralelo.

Por otra parte, observando la Figura 45, se puede notar que no existe una variación en la forma de las distribuciones de distancias entre las distancias de una misma cadena consigo misma en comparación a esta cadena frente a las demás. Esto implica que, el nivel de compresión es tal que para un monómero no existe una distinción clara entre si otro monómero próximo es parte de su misma cadena o no. En otras palabras, los monómeros se encuentran a distancias semejantes entre si independiente de la cadena a la que pertenezcan. Esto significa que, para propósitos de la dimerización de la proteína VP3, podría no haber una clara preferencia en si esta se dimeriza con segmentos de una misma cadena o no.

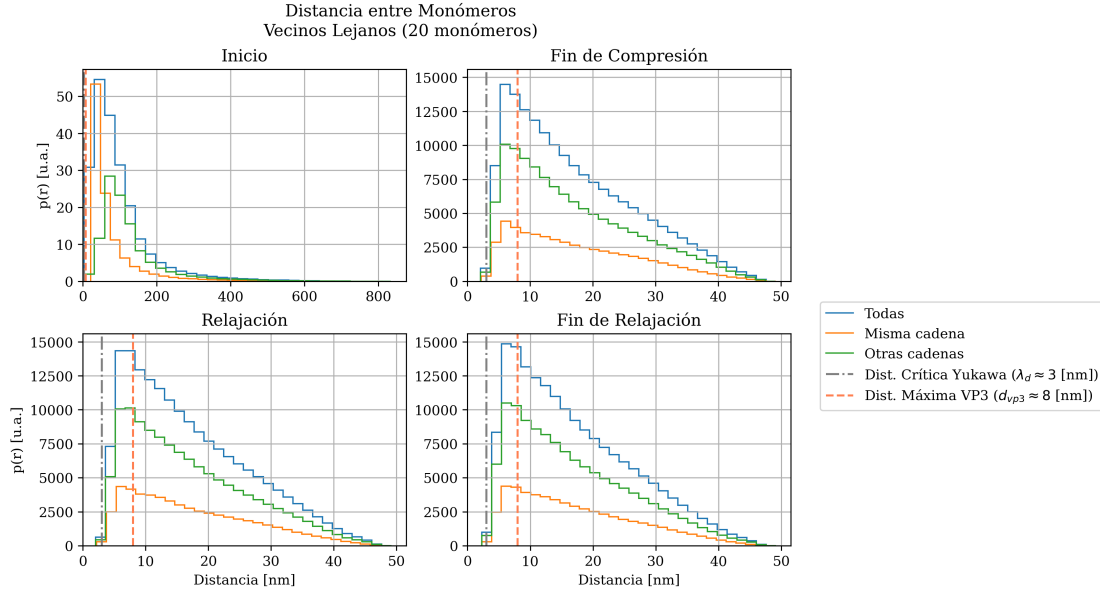


Figura 45: Histograma de distancia entre pares lejanos de monómeros, a contar del vecino numero 20 para caso doble paralelo.

Finalmente, se puede notar en la Figura 46 que, a diferencia del caso individual, se observan diferentes regiones de una misma cadena donde resulta plausible una dimerización de la proteína VP3. En primer lugar, se puede notar que existe una semejanza entre el final de la compresión y el final de la relajación de la estructura, donde algunas zonas de la primera etapa se mantienen a una distancia promedio inferior a 8 [nm]. Esto indica que, las regiones demarcadas al final de la relajación guardan una relación con cómo se encontró la estructura al finalizar la compresión. En segundo lugar, se puede observar dos franjas entre los índices i 150 y 250, en las cuales ambos índices aumentan en conjunto. Esto implica que el segmento $i = [150, 250]$ se encuentra a menos de 8 [nm] en promedio del segmento $j = [80, 150]$. En otras palabras, en dicha zona, existen dos segmentos próximos de una misma hebra de ARN en sentido paralelo, donde el ARN se plegó de tal manera que ambos segmentos resultaron próximos en el espacio formando un bucle. Este mismo comportamiento se puede observar entre los índices $i = [500, 560]$, sin embargo, mientras un índice aumenta el otro disminuye por lo que dichos segmentos se encuentran en sentido antiparalelo formando probablemente un semi-bucle.

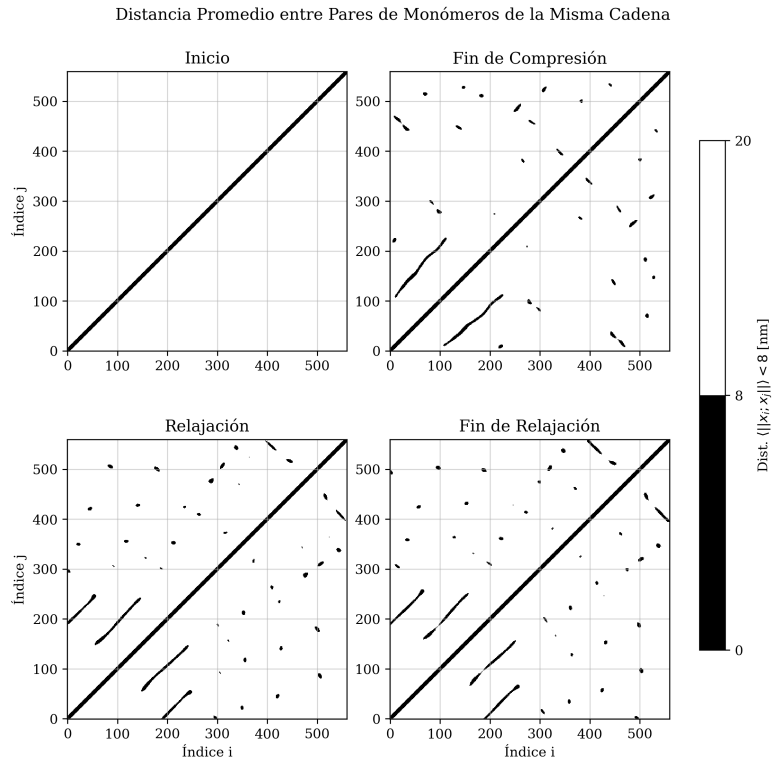


Figura 46: Matriz de pares de distancias entre monómeros de una misma cadena para el caso doble paralelo.

4.2.3 Compresión de Dos Hebras Cruzadas

Se realizó una simulación de compresión de dos hebras de ARN con las hebras cruzadas en un inicio, como se indica en la Figura 47, de tal manera de estudiar el impacto de la condición inicial sobre los parámetros analizados. Para esto, se establecieron las mismas condiciones que para la simulación anterior, las que se describen en el Cuadro 14.

Cuadro 14: Parámetros utilizados para simulación de compresión para 2 hebras cruzadas.

Parámetro	Valor
Longitud	560 & 640 monómeros
Pasos Compresión	41.000.000
Pasos Relajación	350.000.000
Paso Temporal Compresión	0,001
Paso Temporal Relajación	0,0001

Resultados

Los resultados de la presente simulación fueron similares a lo visto en la simulación interior, indicando que los resultados vistos son razonablemente independientes de la condición inicial de las hebras, como se puede observar en la Figura 49. Por esta razón, se ponen a disposición las demás figuras de la presente simulación en el Anexo E. En la figura indicada, se puede notar que la distribución de la densidad radial es similar a la vista en el caso paralelo, siendo destacable la presencia del valle a ≈ 20 [nm] del centro y la tendencia a menor densidad de monómeros hacia el centro del casquete potencial.

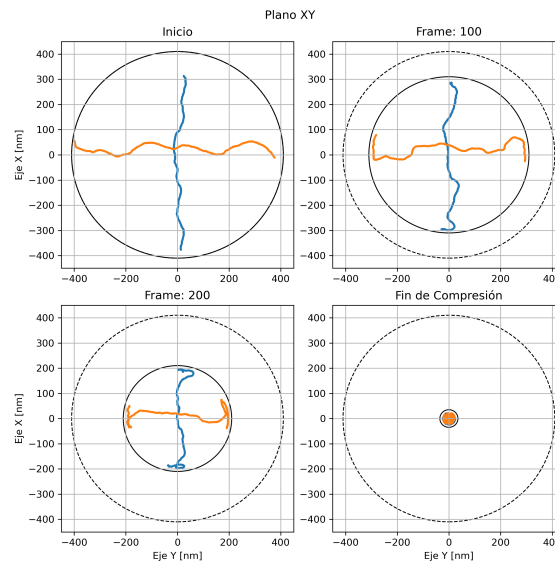


Figura 47: Secuencia de compresión para el caso doble en cruz. Elaboración propia.



Figura 48: Renderización de la simulación doble cruzada en sus fases de compresión (a) y final de relajación. (b).

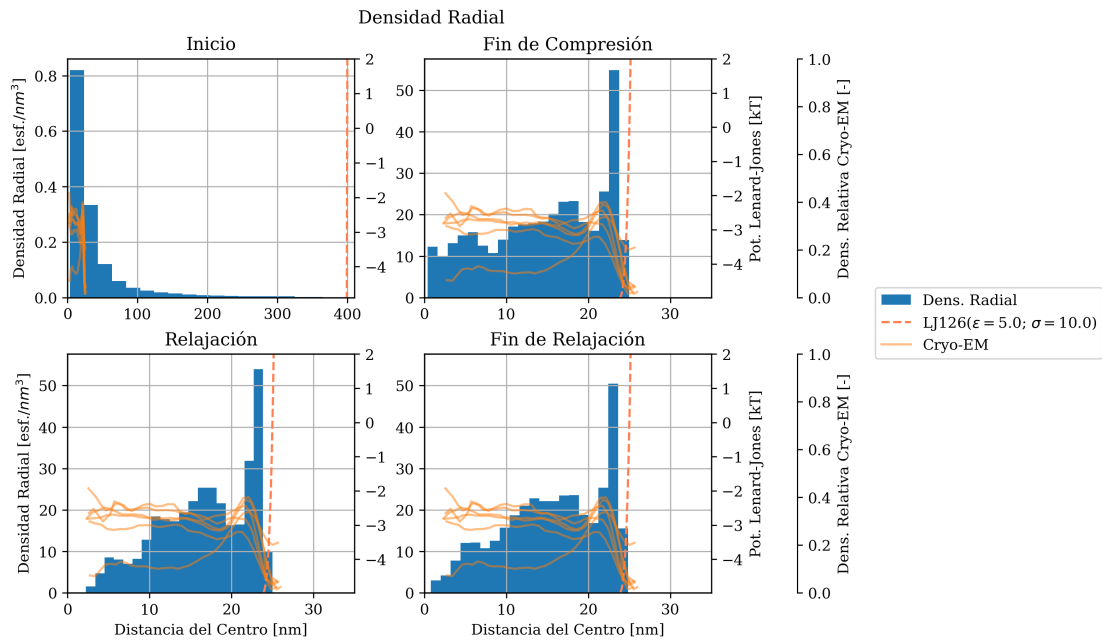


Figura 49: Densidad radial para el inicio, final de compresión, paso intermedio de relajación y final de etapa de relajación para el caso doble cruzado.

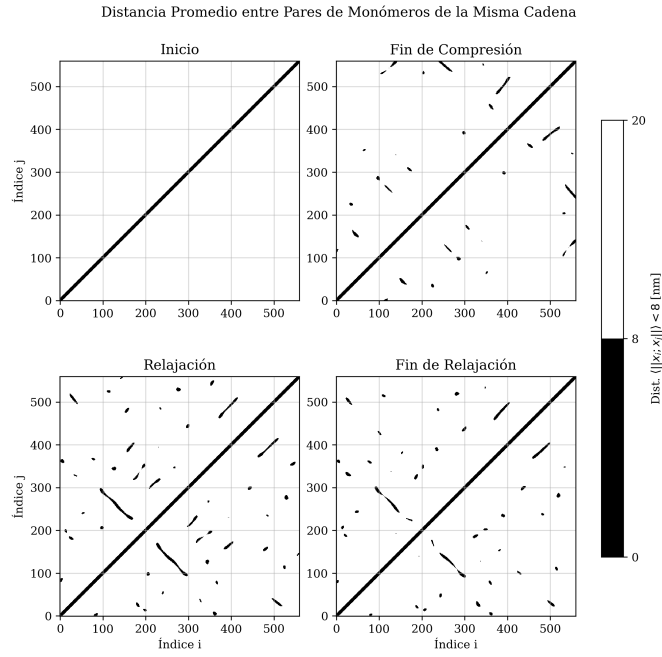


Figura 50: Matriz de pares de distancias entre monómeros de una misma cadena para el caso doble cruzado.

Finalmente, se puede notar en la Figura 50 que, existen regiones donde resulta plausible que la proteína VP3 dimerizada se acople a las cadenas de ARN, similar a lo visto en la simulación previa. Sin embargo, la disposición de las regiones marcadas difiere, lo que indica que esta disposición específica depende en gran medida de la condición inicial de las hebras, más no la existencia de dichas zonas.

5 Conclusiones

Las conclusiones del presente trabajo se pueden separar en dos secciones: los resultados de la parametrización mediante APBS, y los resultados vistos en las simulaciones de compresión. Respecto a la parametrización, en primer lugar, se puede concluir que el potencial de Yukawa con el modelo de grano grueso indicado en la Subsección 3.4 representa fielmente la energía de interacción calculada mediante APBS entre dos hebras de ARN infinitamente largas hasta una distancia mínima equivalente a la distancia de Debye del solvente —a partir de la cual los efectos no polares comienzan a tomar mayor protagonismo— independiente de su orientación relativa. Sin embargo, se debe tomar precaución que las hebras sean de una longitud tal que los efectos de la rugosidad superficial sean despreciables. En segundo lugar, a partir de lo visto en la Subsubsección 4.1.3, se puede concluir que los efectos de la geometría de la superficie de exclusión de solvente son relevantes respecto al cálculo de la energía de interacción calculada por APBS, al menos tan relevantes como la distribución de carga al interior de la molécula. De forma general, para efectos de la parametrización de modelos de grano grueso mediante APBS, resulta de especial importancia que las vías de separación entre ambas moléculas para el efecto de la energía de interacción sean despejadas, y se debe tener en consideración que APBS no contempla deformaciones geométricas de las proteínas al interactuar las cuales pueden ser especialmente relevantes. Por otra parte, respecto a lo visto en las simulaciones de compresión, en primer lugar, se puede establecer que existen estructuras compartidas entre las simulaciones con 1 y 2 hebras tales como el valle previo al muro potencial, la distancia entre pares de monómeros y que los ángulos de las hebras rondan los 150° en todos los casos. Además, se logró establecer una cierta independencia a las condiciones iniciales para el caso con 2 hebras de ARN. En base a esto, se puede concluir que se logró replicar levemente la densidad radial indicada por el Cryo-EM del virus IBDV, con algunas diferencias notables como una tendencia a acumular ARN en las paredes del muro potencial en lugar de distribuirse uniformemente por el virión. Esto puede ser explicado con que las simulaciones presentes solo consideran material genético, mientras que los datos del Cryo-EM contemplan otros elementos.

En segundo lugar, se puede concluir que la electrostática del sistema permite que las hebras de ARN se encuentren a una distancia tal que la proteína VP3, ya sea su dímero o en solitario, pueda acoplarse a las cadenas de ARN mediante el núcleo al contar con espacio suficiente entre las cadenas para los casos vistos.

Cabe destacar que también se observó que cuando existe solo una cadena de ARN el acople de la proteína a dos segmentos de esta puede ser más compleja en comparación al caso doble, puesto que los monómeros se encuentran en promedio a una distancia mayor que la máxima admitida por el dímero de la proteína. Sin embargo, estos resultados contemplan únicamente el impacto electrostático de las hebras de ARN por lo que el efecto propio de la proteína puede ser relevante. De todas maneras, en ambos casos, la distancia entre pares de monómeros es tal que admite la presencia de proteínas VP3 en el sistema.

En tercer lugar, en base a los resultados entregados por los histogramas de distancia, es posible que no exista una preferencia fuerte por la proteína VP3 a dimerizarse con segmentos de su misma hebra de ARN en comparación a segmentos de otras cadenas puesto que la distribución de los monómeros no presenta marcadas diferencias entre estos casos.

Por último, posibles vías de exploración para trabajos futuros resultan en realizar más simulaciones para observar que ocurre con 3, 4 o 5 hebras de ARN al comprimirse hasta el tamaño del virus. Luego, podría ser de interés realizar una parametrización de grano grueso como la del

presente trabajo entre el ARN y la cápside del virus de tal manera de modelar la interacción entre el ARN y el muro potencial de mejor manera, así como también el efecto electrostático de la proteína VP3. Finalmente, se podría estudiar la energía de interacción entre fragmentos de ARN con y sin la proteína VP3 adosada bajo diferentes orientaciones y condiciones de tal manera de refinar la parametrización realizada.

Bibliografia

- Alberts, B. (2022). *Molecular Biology of the Cell* (Seventh edition). W. W. Norton & Company.
- Baker, N. A. (2004). PoissonBoltzmann Methods for Biomolecular Electrostatics. In *Methods in Enzymology* (pp. 94–118, Vol. 383). Elsevier. [https://doi.org/10.1016/S0076-6879\(04\)83005-2](https://doi.org/10.1016/S0076-6879(04)83005-2)
- Baker, N. A. (2023a). APBS - Adaptive Poisson-Boltzmann Solver APBS 3.4.1 documentation. Retrieved March 29, 2026, from <https://apbs.readthedocs.io/en/latest/index.html>
- Baker, N. A. (2023b). Binding energies APBS 3.4.1 documentation. Retrieved March 29, 2026, from <https://apbs.readthedocs.io/en/latest/using/examples/binding-energies.html>
- Baker, N. A., Dong, F., & Olsen, B. (2008). Computational Methods for Biomolecular Electrostatics. In *Methods in Cell Biology* (pp. 843–870, Vol. 84). Elsevier. [https://doi.org/10.1016/S0091-679X\(07\)84026-X](https://doi.org/10.1016/S0091-679X(07)84026-X)
- Cheng, H.-B., Cheng, L.-T., & Li, B. (2011). Yukawa-field approximation of electrostatic free energy and dielectric boundary force. *Nonlinearity*, 24(11), 3215–3236. <https://doi.org/10.1088/0951-7715/24/11/011>
- Emmett M. Leddin. (2021). Gaussian Input Files — Computational Chemistry Resources. Retrieved July 21, 2026, from <https://emleddin.github.io/comp-chem-website/Otherguide-gaussian-input.html>
- Feldman, J. (2021, November). C.1: Richardson Extrapolation. Retrieved March 29, 2026, from [https://math.libretexts.org/Bookshelves/Calculus/CLP-2_Integral_Calculus_\(Feldman_Rechnitzer_and_Yeager\)/04%3A_Appendices/4.03%3A_C%3A_More_About_Numerical_Integration/4.3.01%3A_C.1%3A_Richardson_Extrapolation](https://math.libretexts.org/Bookshelves/Calculus/CLP-2_Integral_Calculus_(Feldman_Rechnitzer_and_Yeager)/04%3A_Appendices/4.03%3A_C%3A_More_About_Numerical_Integration/4.3.01%3A_C.1%3A_Richardson_Extrapolation)
- Grochowski, P., & Trylska, J. (2008). Continuum molecular electrostatics, salt effects, and counterion binding A review of the PoissonBoltzmann theory and its modifications. *Biopolymers*, 89(2), 93–113. <https://doi.org/10.1002/bip.20877>
- Guerrero-Montero, M., Bosy, M., & Cooper, C. D. (2026). A self-adjusting FEM-BEM coupling scheme for the nonlinear Poisson-Boltzmann equation. *Computer Physics Communications*, 324, 110158. <https://doi.org/10.1016/j.cpc.2026.110158>
- Hoppe, T. (2013). A simplified representation of anisotropic charge distributions within proteins. *The Journal of Chemical Physics*, 138(17), 174110. <https://doi.org/10.1063/1.4803099>
- Jurrus, E., Engel, D., Star, K., Monson, K., Brandi, J., Felberg, L. E., Brookes, D. H., Wilson, L., Chen, J., Liles, K., Chun, M., Li, P., Gohara, D. W., Dolinsky, T., Konecny, R., Koes, D. R., Nielsen, J. E., HeadGordon, T., Geng, W., ... Baker, N. A. (2018). Improvements to the `APBS` biomolecular solvation software suite. *Protein Science*, 27(1), 112–128. <https://doi.org/10.1002/pro.3280>
- Mata, C. P., Mertens, J., Fontana, J., Luque, D., Allende-Ballester, C., Reguera, D., Trus, B. L., Steven, A. C., Carrascosa, J. L., & Castón, J. R. (2018). The RNA-Binding Protein of a Double-Stranded RNA Virus Acts like a Scaffold Protein (S. López, Ed.). *Journal of Virology*, 92(19), e00968–18. <https://doi.org/10.1128/JVI.00968-18>
- Morton, K. W., & Mayers, D. F. (2005). *Numerical solution of partial differential equations* (2nd ed). Cambridge University Press.
- Nature. (2026, April). Virology - Latest research and news — Nature. Retrieved April 16, 2026, from <https://www.nature.com/subjects/virology>
- Orakpoghenor, O., Oladele, S. B., & Abdu, P. A. (2020). Infectious Bursal Disease: Transmission, Pathogenesis, Pathology and Control - An Overview. *World's Poultry Science Journal*, 76(2), 292–303. <https://doi.org/10.1080/00439339.2020.1716652>

- Poblete, S., Bottaro, S., & Bussi, G. (2018). A nucleobase-centered coarse-grained representation for structure prediction of RNA motifs. *Nucleic Acids Research*, *46*(4), 1674–1683. <https://doi.org/10.1093/nar/gkx1269>
- Rocchia, W. (2005). Poisson-boltzmann equation boundary conditions for biological applications. *Mathematical and Computer Modelling*, *41*(10), 1109–1118. <https://doi.org/10.1016/j.mcm.2005.05.006>
- Sandia Corporation. (2026, April). LAMMPS Documentation (4 Jul 2026 version) LAMMPS documentation. Retrieved July 13, 2026, from <https://docs.lammps.org/>
- Sponer, J., Bussi, G., Krepl, M., Baná, P., Bottaro, S., Cunha, R. A., Gil-Ley, A., Pinamonti, G., Poblete, S., Jureka, P., Walter, N. G., & Otyepka, M. (2018). RNA Structural Dynamics As Captured by Molecular Simulations: A Comprehensive Overview. *Chemical Reviews*, *118*(8), 4177–4338. <https://doi.org/10.1021/acs.chemrev.7b00427>
- Sponk, F. D. R.-D. s. S. /. *. (2010, March). Comparación de una cadena sencilla de ARN con una doble hélice de ADN con sus correspondientes bases nitrogenadas. Retrieved July 21, 2026, from https://commons.wikimedia.org/wiki/File:Difference_DNA_RNA-EN.svg

Anexo

Anexo A: Configuración de APBS

```
1      read
2          mol pqr R0_2rna.pqr
3      end
4      elec name solv
5          mg-manual
6          dime 289 321 513
7          grid 0.3 0.3 0.3
8          gcent mol 1
9          mol 1
10         npbe
11         bcfl mdh
12         ion charge 1 conc 0.150 radius 2.0
13         ion charge -1 conc 0.150 radius 2.0
14         pdie 2.0000
15         sdie 78.5400
16         srfm smol
17         chgm spl2
18         sdens 10.00
19         srad 1.4
20         swin 0.30
21         temp 298.15
22         calcenergy total
23         calcforce no
24         #write pot dx R0_2rna-solv
25     end
26
27     elec name vacc
28         mg-manual
29         dime 289 321 513
30         grid 0.3 0.3 0.3
31         gcent mol 1
32         mol 1
33         npbe
34         bcfl mdh
35         pdie 2.0000
36         sdie 2.0000
37         srfm smol
38         chgm spl2
39         sdens 10.00
40         srad 1.4
41         swin 0.30
42         temp 298.15
43         calcenergy total
44         calcforce no
45         #write pot dx R0_2rna-vacc
46     end
47
48     print elecEnergy solv - vacc end
49     quit
```

Anexo B: Ejemplo de configuración de *apbsHelper*

```
1 #Example Config File
2 name: rna_many_to_many_8-cg
3
4 config_only: false #Only configure the simulations, do not run APBS
5 log_level: info
6 max_threads: 4 #Max amount of parallel simulations
7
8 cases_to_run:
9     # # APBS All Atom: P Cases
10     - theta-0-omega-0-apbsAllAtom_8-cg
11
12     # # APBS All Atom: N Cases
13     - theta-90-omega-0-apbsAllAtom_8-cg
14
15     # Yukawa Coarse Grain: P Cases
16     - theta-0-omega-0-yukawaCg_8-cg
17
18     # Yukawa Coarse Grain: N Cases
19     - theta-90-omega-0-yukawaCg_8-cg
20
21 # Pdb handler class. Must be subclass of HandlerFunction
22 pdb_handler: src.handlers.rna_study_handler.many_to_many
23 pdb_directory: pdbFiles
24 core_pdb_filename: rna/8-cg_2rna.pdb
25
26 # Global simulation settings (Overriden by case settings if present)
27 global_sim_options:
28     mesh_size: 0.3
29     yukawa: false
30     coarse_grain: false
31     linear_pb: false
32     keep_complex_dx: false
33
34 global_handler_options:
35     base: # Parameters under base apply to all simulations of case.
36         cg_geometry: SimpleGeometry
37         cg_charges: [-10]
38 # Define number of simulations in case.
39 # Each item in list is one simulation under the case.
40 variants:
41     distance: [0, 33, 34, 36, 38, 43, 49, 59, 69]
42
43 # Cases definition
44 cases:
45     - name: theta-0-omega-0-apbsAllAtom_8-cg
46       sim_options:
47         coarse_grain: false
48         yukawa: false
49       handler_options:
50         base:
51             rotation_theta: 0
52             rotation_omega: 0
53
54     - name: theta-90-omega-0-apbsAllAtom_8-cg
55       sim_options:
56         coarse_grain: false
57         yukawa: false
58       handler_options:
59         base:
```

```

60         rotation_theta: 90
61         rotation_omega: 0
62
63     - name: theta-0-omega-0-yukawaCg_8-cg
64       sim_options:
65         coarse_grain: true
66         yukawa: true
67       handler_options:
68         base:
69           rotation_theta: 0
70           rotation_omega: 0
71
72     - name: theta-90-omega-0-yukawaCg_8-cg
73       sim_options:
74         coarse_grain: true
75         yukawa: true
76       handler_options:
77         base:
78           rotation_theta: 90
79           rotation_omega: 0

```

Anexo C: Ejemplo de configuración de LAMMPS

```
1 #####
2 # LAMMPS script for a polymer
3 #####
4
5 # Box and units (use LJ units and periodic boundaries)
6 units lj # use lennard-jones (i.e. dimensionless) units
7 atom_style molecular # atoms with bonds and angles
8
9 #boundary p p p # all boundaries are periodic
10 boundary f f f
11
12 # Pair interactions require lists of neighbours to be calculated
13 neighbor 1.9 bin
14 neigh_modify every 1 delay 1 check yes
15 special_bonds lj/coul 0.0 0.0 0.0 # Exclusion
16 comm_modify cutoff 4.0
17
18 # READ "start" data file
19 read_data last_centered.dat
20
21 # Generate regular RESTART files to store state of simulation
22 restart 5000000 chkpnt.restart
23
24 # Reset timestep
25 reset_timestep 0
26
27 # Define groups
28 group all type 1 #(atom type 1 is group 'all')
29
30 # Dump configurations
31
32 dump dum1 all custom 100000 dump.RNP.lammpstrj id mol type x y z ix iy iz
33
34 # every 1000 time steps, each atom's id, type, x,y,z coords and x,y,z image flages
35 # are output.
36 # image flags count how many times each atom has moved through a periodic boundary
37
38 # Set up interactions — start with 'soft' potentials
39 # Define a spherical region
40 variable rad equal "445 - step*0.00001" # Define el RADIO de la esfera
41 region my_sphere sphere 0 0 0 v_rad side in
42 fix my_wall all wall/region my_sphere lj126 5.0 10.0 11.22
43 fix infofile all print 100000 "$t ${rad}" file info.dat screen no
44
45 ### Between bonded atoms
46 bond_style harmonic
47 # For style harmonic, specify:
48 # * bond type
49 # * K (energy/distance^2)
50 # * R0 (distance)
51 # Remember that the distance is measured in nm.
52 # The original K is 0.9054358554067355 .
53 bond_coeff 1 90.54358554067355 1.6751745377096114
54
55 ### angles
56 # For style cosine/squared, specify:
57 # angle type
58 # K (energy)
59 # theta0 (degrees)
```

```

59 # cos theta0 = -0.4930903891239791
60 angle_style cosine/squared
61 angle_coeff 1 43.60265790665352 119.543907079
62 # this choice gives a polymer with the same persistence length as DNA
63
64 ## dihedrals
65 dihedral_style quadratic
66 # K
67 # phi0 . No dihedral type according to manual
68 dihedral_coeff 1 16.732963818 146.02232849922407
69
70 ## Between non-bonded atoms
71 pair_style lj/expand 0.561231024
72 pair_modify shift yes # option to ensure energy is calculated correctly
73 pair_coeff * * 1.0 0.5 2.371
74
75 # Yukawa Potential T = 300
76 pair_style yukawa 1.27282 6.0
77 pair_coeff * * 70.91960 # Default interaction
78 pair_coeff 1 1 70.91960 # Bulk-Bulk interaction
79 pair_coeff 1 2 63.82764 # Bulk-Cap interaction
80 pair_coeff 2 2 57.44487 # Cap-Cap interaction
81
82 # Set up fixes
83 variable seed equal 1 # a seed for the thermostat
84
85 fix 1 all nve # NVE integrator
86 fix 2 all langevin 1.0 1.0 1.0 ${seed} # langevin thermostat
87
88 ##### Output thermodynamic info (temperature, energy, pressure, etc.) #####
89 thermo 10000
90 thermo_style custom step temp epair emol press vol
91
92
93 ##### Output thermodynamic info to file #####
94 variable t equal step
95 variable mytemp equal temp
96 variable myepair equal epair
97 #fix mythermofile all print 1000 "$t ${mytemp} ${myepair}" file thermo-output.dat
98 screen no
99
100 ##### Set up a compute for R_g
101 compute myRG2compute all gyration
102 ##### and write it to a file
103 variable RG2 equal c_myRG2compute
104 fix myRG2file all print 100000 "$t ${RG2}" file radius_of_gyration_squared.dat
105 screen no
106
107 # Compute PE
108 variable myPE equal pe
109 fix myPE all print 10000 "$t ${myPE}" file pe-energy.dat screen no
110
111 # set timestep of integrator
112 timestep 0.001
113
114 # run integration for a number of steps to equilibrate with soft potentials
115
116 run 41000000 # Compress for 41.000.000 steps
117 variable rad equal "35" # Freeze sphere in place
118 timestep 0.0001 # Reduce ts
119 run 350000000 # Run other 300.000.000 steps

```

```
119 | write_restart RNP.restart
120 |
121 | # write a final restart file
122 | write_restart final.restart
123 | write_data data.*
124 | write_dump all xyz final.xyz
```

Anexo D: Figuras Adicionales APBS

Casos 1-N

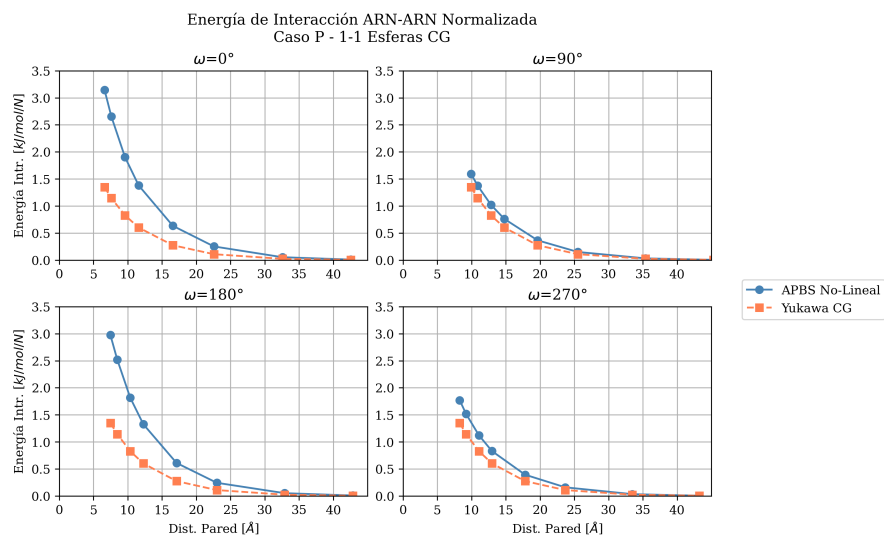


Figura 51: Energía de interacción para el caso 1-1 orientación P.

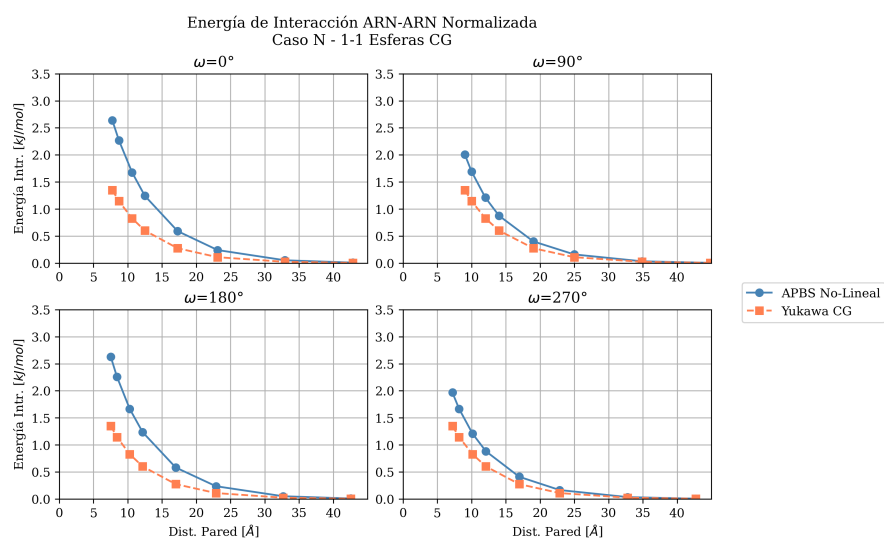
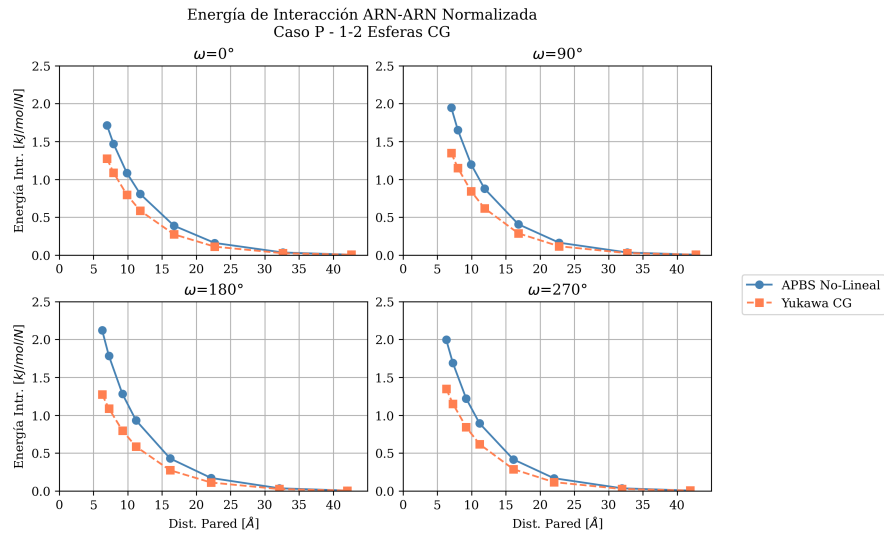
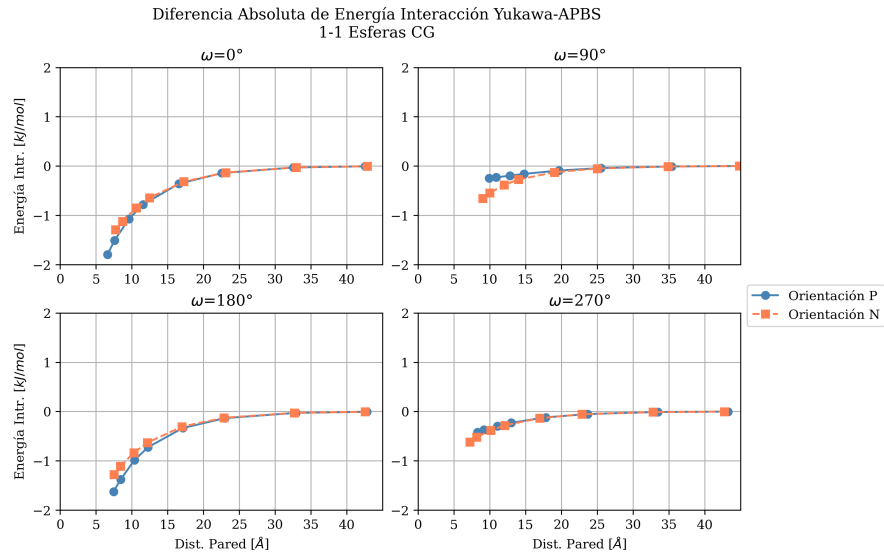


Figura 52: Energía de interacción para el caso 1-1 orientación N.



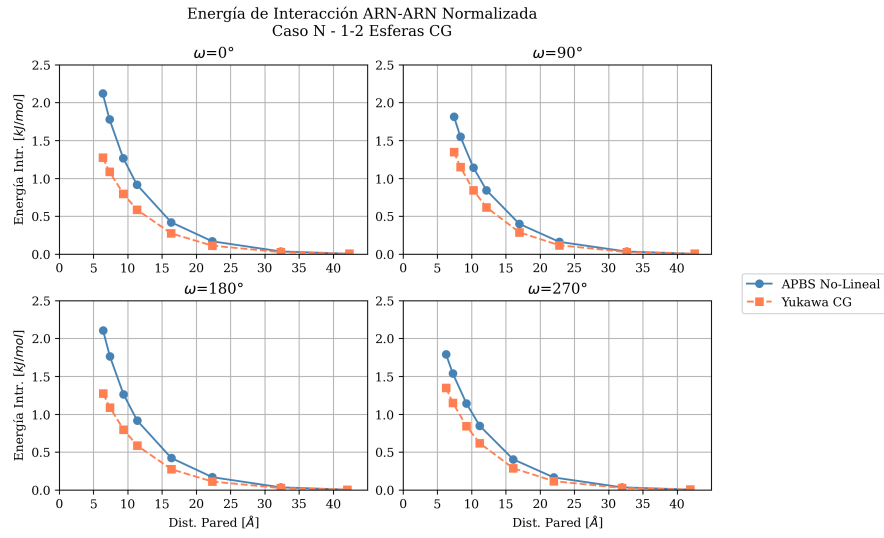


Figura 55: Energía de interacción para el caso 1-2 orientación N.

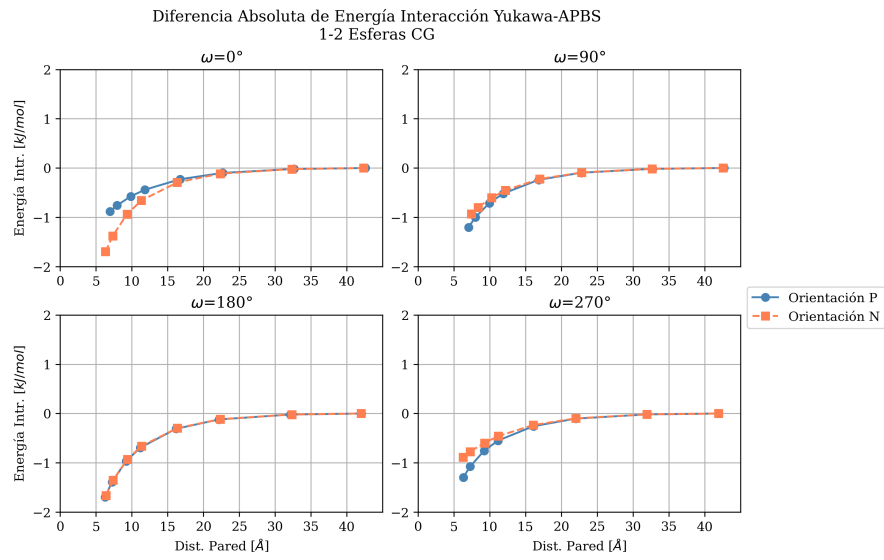


Figura 56: Diferencia Yukawa-APBS para el caso 1-2.

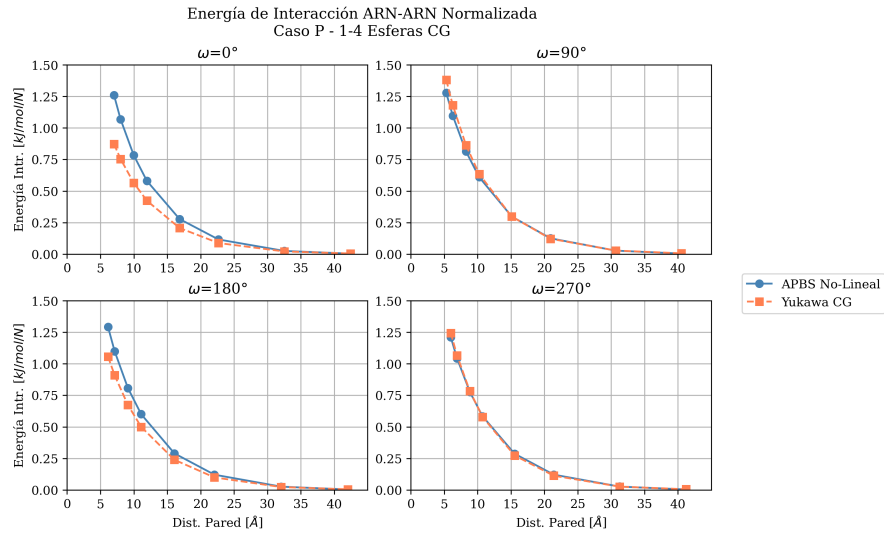


Figura 57: Energía de interacción para el caso 1-4 orientación P.

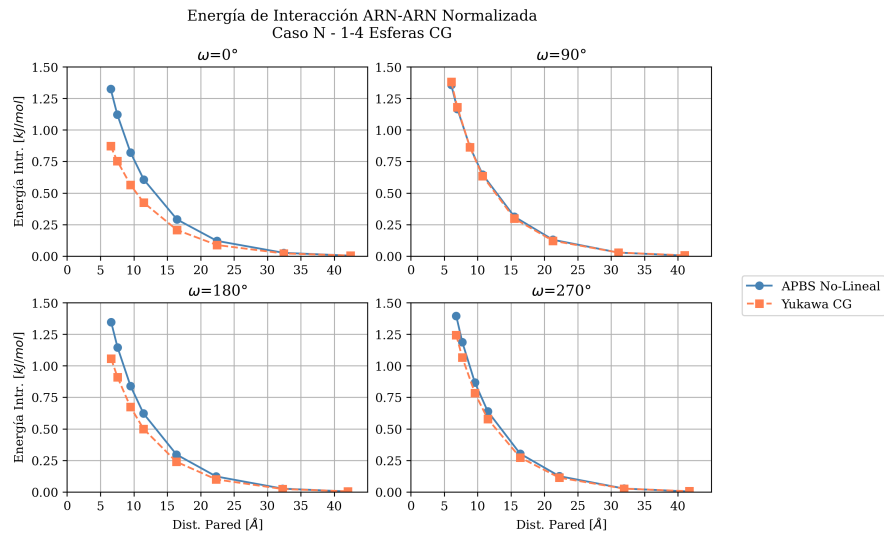
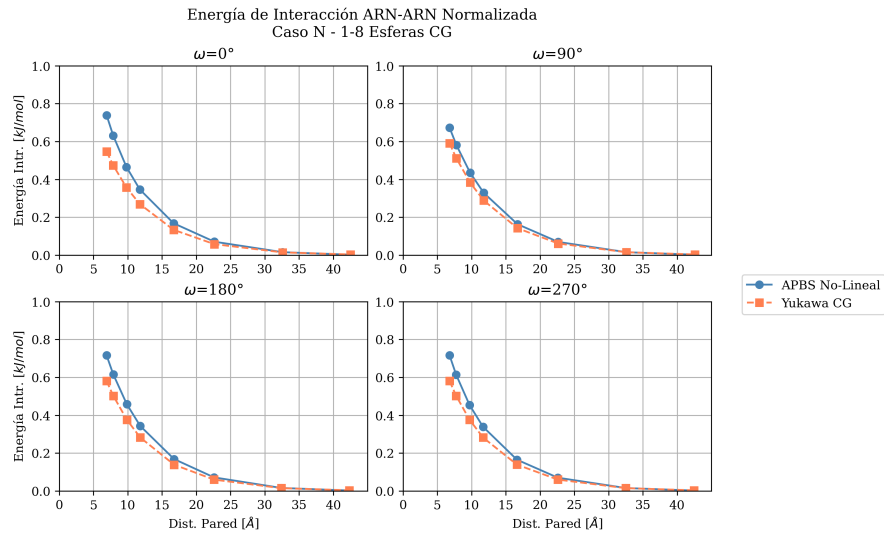
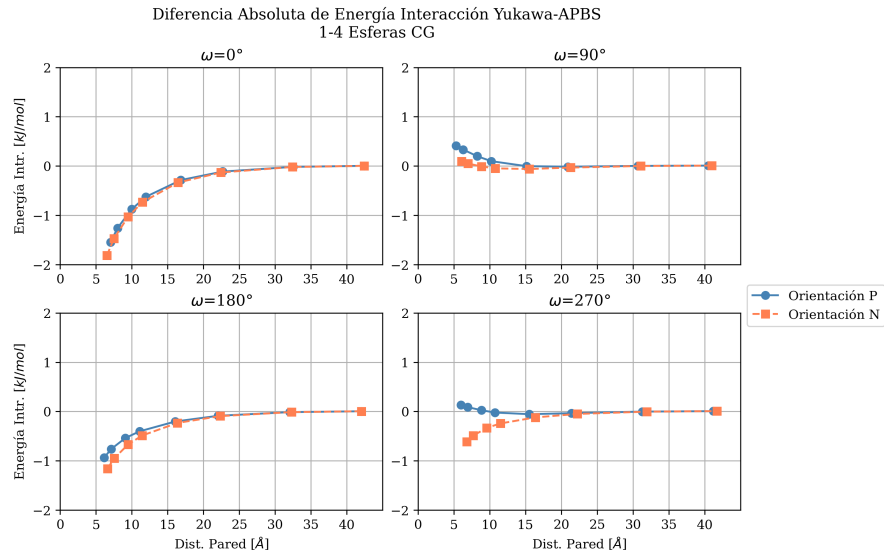


Figura 58: Energía de interacción para el caso 1-4 orientación N.



Casos N-N

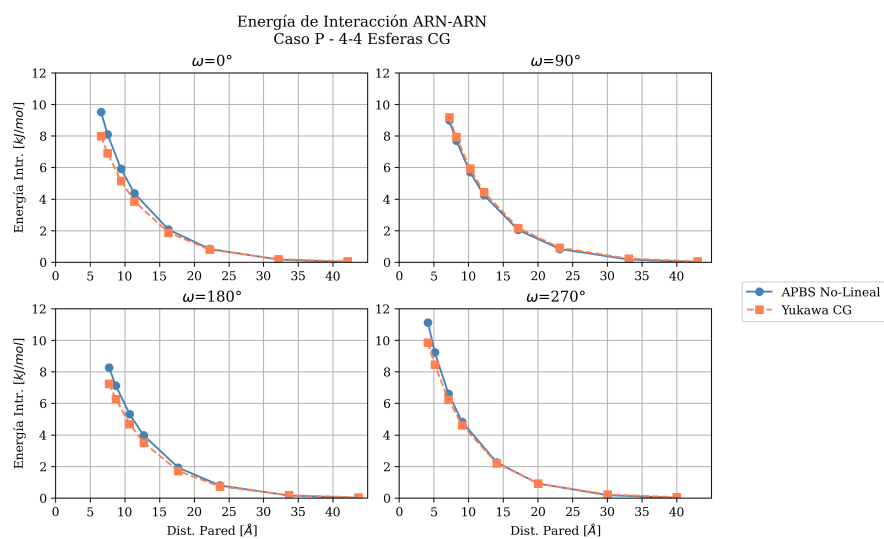


Figura 61: Energía de interacción para el caso 4-4 orientación P.

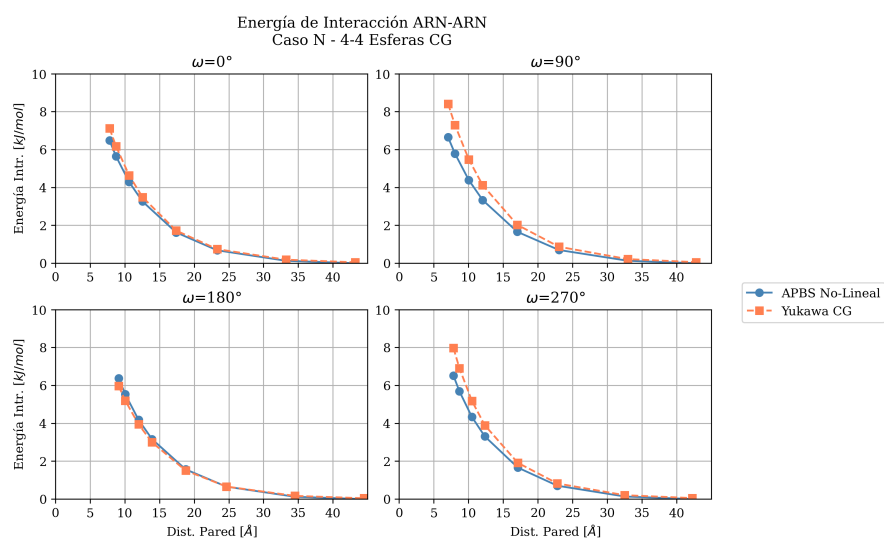


Figura 62: Energía de interacción para el caso 4-4 orientación N.

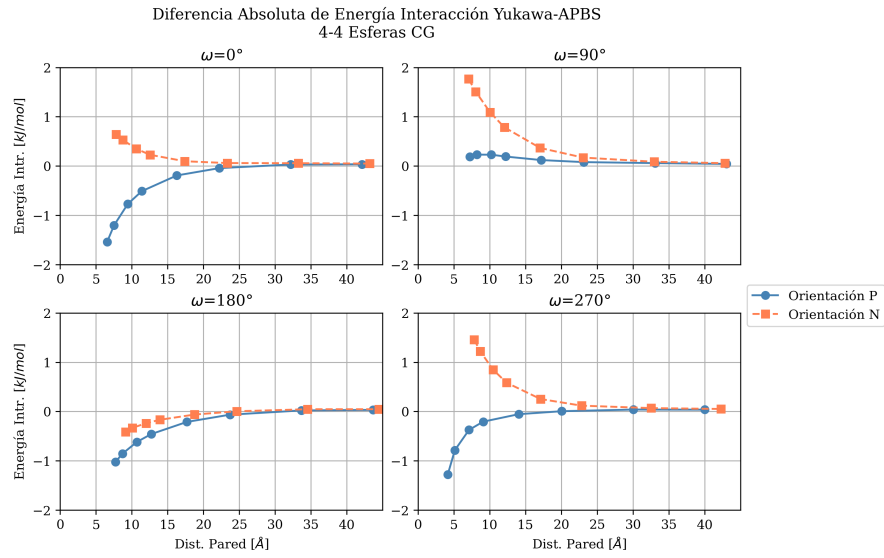


Figura 63: Diferencia Yukawa-APBS para el caso 4-4.

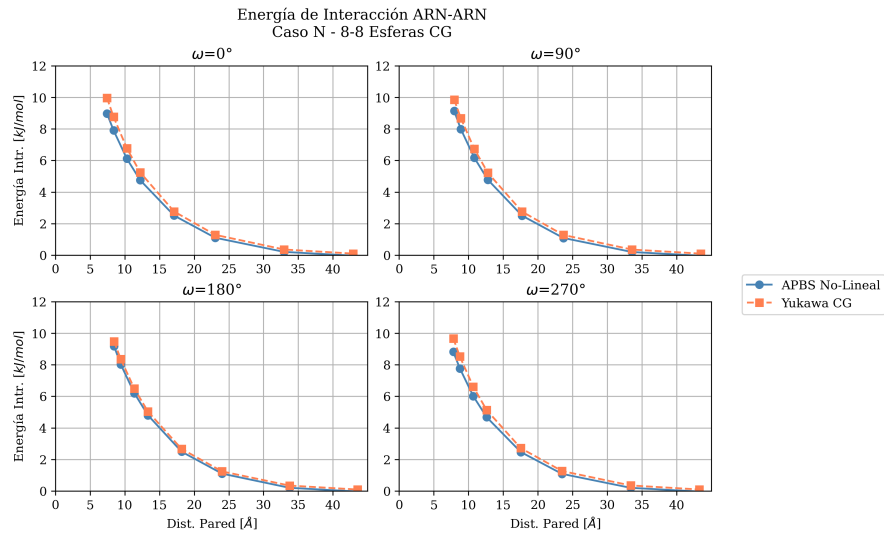


Figura 64: Energía de interacción para el caso 8-8 orientación N.

Anexo E: Figuras Adicionales LAMMPS

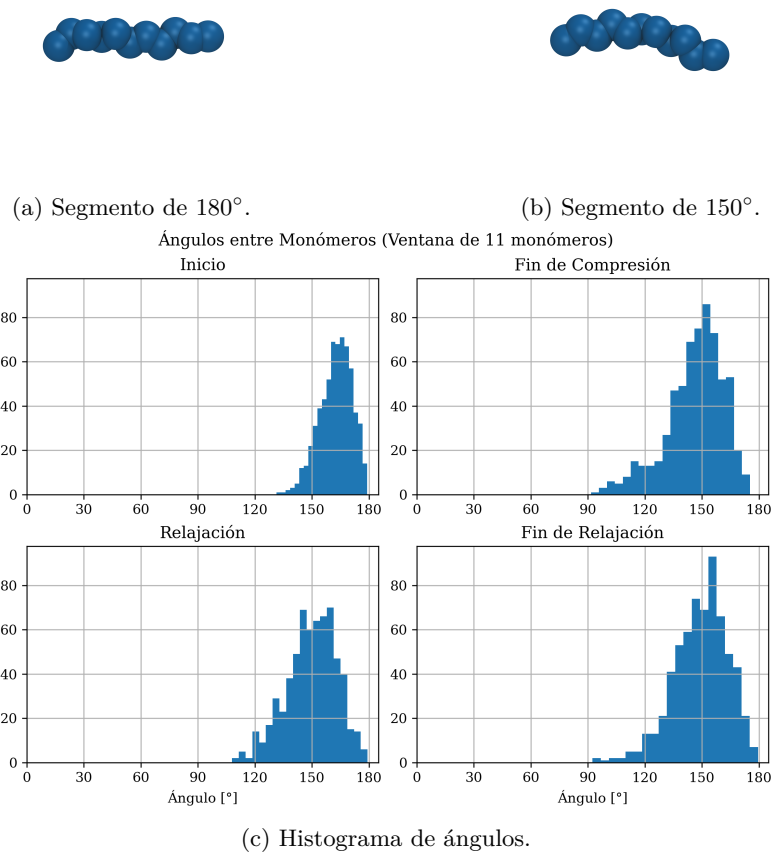


Figura 65: Histograma de ángulos para el caso doble cruzado (c) con ejemplos de segmentos de 180° (a) y 150° (b).

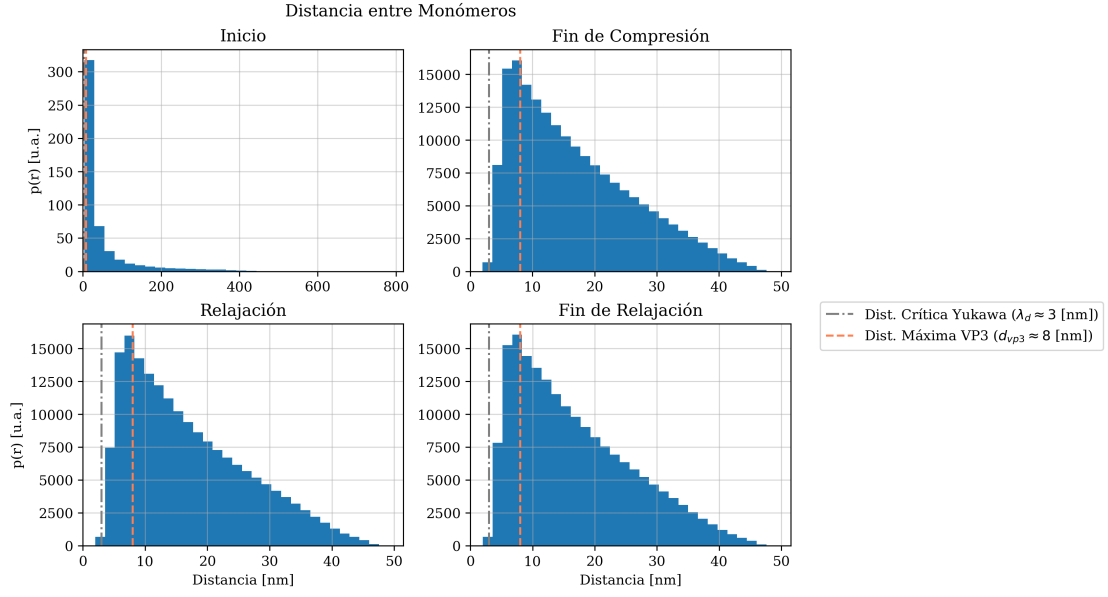


Figura 66: Histograma de distancia entre pares de monómeros excluyendo los 4 vecinos inmediatos para el caso doble cruzado.

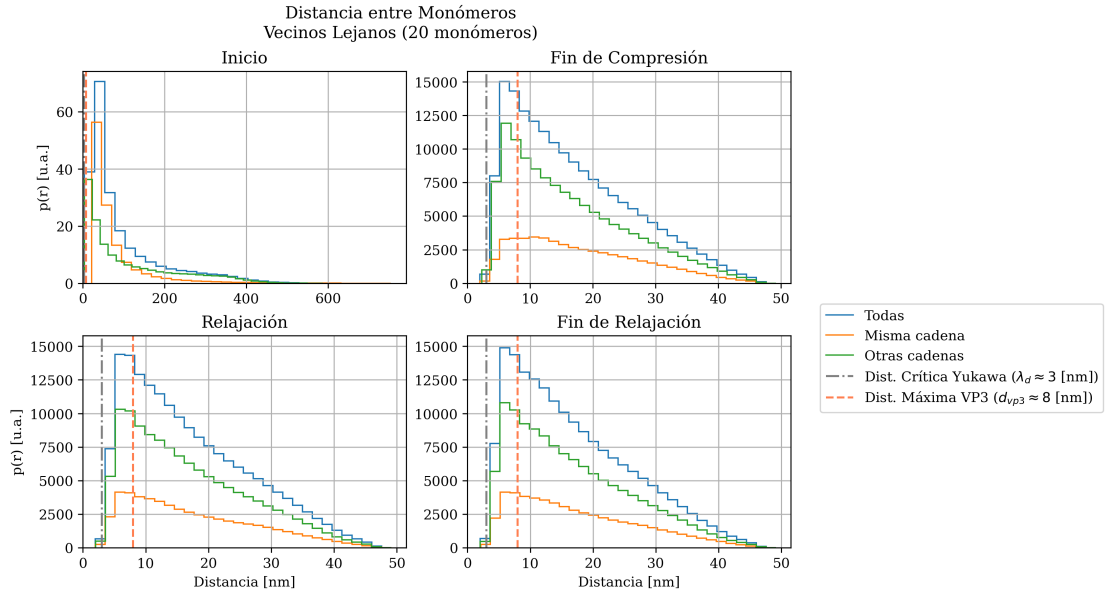


Figura 67: Histograma de distancia entre pares lejanos de monómeros, a contar del vecino numero 20 para caso doble cruzado.