



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

PROGRAMA DE MAESTRÍA E INGENIERÍA EN MATERIALES

“INTERACCIÓN ENTRE EL DOMINIO DE UNIÓN AL RECEPTOR (RBD) DE LA SUBUNIDAD S1 QUE CORRESPONDE A LA GLICOPROTEÍNA ESPICULAR Y DIFERENTES SUPERFICIES EN SARS-COV-2”

TESIS

**QUE PARA OPTAR POR EL GRADO DE:
MAESTRO EN CIENCIA E INGENIERÍA DE MATERIALES**

**PRESENTA
FIS. GERARDO GUSTAVO RAMÍREZ RUIZ**

Tutor:
Dra. Angélica Estrella Ramos Peña
Instituto de Investigaciones en Materiales

COMITÉ TUTOR:
Dr. Héctor Domínguez Castro
Instituto de Investigaciones en Materiales

Dra. María Beatriz de la Mora Mojica
Instituto de Ciencias Aplicadas y Tecnología Avanzada

Ciudad Universitaria, CD. MX., Julio, 2022



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Jurado asignado

Presidente

Dr. Carlos Ignacio Mendoza Ruiz
Instituto de Investigaciones en Materiales

Primer Vocal

Dra. Angélica Estrella Ramos Peña
Instituto de Investigaciones en Materiales

Segundo Vocal

Dr. Rodolfo Gómez Balderas
Facultad de Estudios Superiores-Cuautitlán

Tercer Vocal

Dra. Xóchitl Pérez Martínez
Instituto de Fisiología Celular

Secretario

Dr. Edgar Núñez Rojas
Universidad Autónoma Metropolitana-Iztapalapa

Sitio donde se desarrollo el trabajo: Instituto de Investigaciones en Materiales, UNAM

Fis. Gerardo Gustavo Ramírez Ruiz

Sustentante

Dra. Angélica Estrella Ramos Peña

Asesora

Este trabajo participó en los siguientes eventos:

- “Interacción entre el dominio de unión al receptor (RBD) de la subunidad S1 que corresponde a la glicoproteína espicular S y una superficie metálica de Cu en SARS-Cov-2.”

XIX Reunión Mexicana de Físicoquímica Teórica

Gustavo Ramírez, Estrella Ramos, Héctor Domínguez, Rodolfo Gómez

UNAM, modalidad virtual, CDMX, 18 y 19 de noviembre del 2021.

Modalidad: Póster

- “Interacción entre el dominio de unión al receptor (RBD) de la subunidad S1 que corresponde a la glicoproteína espicular S y diferentes superficies en SARS-Cov-2.”

Congreso Estatal de Estudiantes de Ciencia e Ingeniería de Materiales

Gustavo Ramírez, Estrella Ramos, Héctor Domínguez, Rodolfo Gómez

UNAM, modalidad virtual, CDMX, 2 a 6 de mayo del 2022.

Modalidad: Oral

Primer Lugar

A mis padres, mi hermana, mis abuelos y todo aquel que me ha apoyado para concluir este importante escalón en mi formación profesional.

"The inconceivable nature of nature."

—Richard P. Feynman.

"No problem is too small or too trivial if we can really do something about it."

—Richard P. Feynman.

"However difficult life may seem, there is always something you can do and succeed at."

—Stephen Hawking.

Agradecimientos

Agradezco al Consejo Nacional de Ciencia y Tecnología (CONACyT) por la beca de maestría otorgada con el número de CVU: 1080413 y al proyecto DGAPA PAPIIT BG100320 por el apoyo económico. Asimismo, agradezco a la UNAM por el apoyo brindado en los recursos de Supercómputo del proyecto LANCAD-UNAM-DGTIC-225.

Expreso un especial agradecimiento a mi familia. A mi papá, el Dr. Gustavo, a mi mamá, la Lic. Betzabel y a mi hermana menor, la Lic. Elisa por su paciencia, su cariño, la privilegiada educación que me brindaron y su constante apoyo en cada uno de los objetivos que me he planteado a lo largo de mi vida, sin el cual ninguno de ellos hubiera sido posible.

Agradezco a esa persona especial por su comprensión, empatía y cariño, que estuvo desde el inicio de esta etapa y al día de hoy se encuentra en la culminación de la misma.

Agradezco a mi asesora, la Dra. Estrella por sus conocimientos, su experiencia profesional, su entusiasmo, su tiempo invertido en cada reunión y su supervisión. Cada uno de estos valiosos ingredientes, que en combinación, me permitieron ir moldeando el proyecto de investigación hasta realizar un trabajo de calidad.

Agradezco a los profesores que han contribuido en mi formación, y que en una excelente muestra de docencia incentivaron mi curiosidad sin la cual no habría decidido tomar este camino.

Agradezco al instituto por la oportunidad otorgada para continuar mi formación académica y concretar otro escalón para llegar a mis objetivos. Particularmente a Oralia L. Jiménez, Maria Teresa Vazquez, Alberto L. Vivas, Alejandro Pompa del personal de computo por el apoyo técnico brindado.

Agradezco al comité tutor por su retroalimentación en cada una de las reuniones realizadas, por su crítica constructiva, que fue esencial durante el desarrollo y consolidación del proyecto de investigación. Así como agradezco al Dr. Rodolfo Gómez Balderas por formar parte del proyecto.

Índice General

	Capítulo 1	1
I. Introducción.....		1
I.1 Objetivos.....		5
I.1.1 Objetivo General.....		5
I.1.2 Objetivos Específicos.....		5
I.2 Hipótesis.....		5
	Capítulo 2	6
II. Marco Teórico.....		6
II.1 Dinámica Molecular.....		7
II.1.1 Ecuaciones de movimiento.....		9
II.1.2 Métodos de diferencias finitas.....		11
II.1.2.1 Algoritmo de Verlet.....		11
II.1.2.2 Algoritmo de velocidad de Verlet.....		12
II.1.2.3 Algoritmo de salto de rana.....		12
II.1.3 Campo de fuerzas.....		13
II.1.4 Condiciones periódicas de frontera.....		19
II.1.4.1 Radio de corte y convención de mínima imagen.....		21
II.1.4.2 Sumas de Ewald.....		23
II.1.5 Minimización de energía.....		27
II.1.6 Termodinámica y Física Estadística.....		29
II.1.6.1 Ensamblés.....		31
II.1.6.1.1 Ensamble NVT.....		34
II.1.7 Equilibración en el ensamble NVT.....		36
II.1.8 Producción de datos.....		38
II.2 Programas utilizados.....		38
II.2.1 GROMACS.....		39
II.2.2 CHARMM-GUI Input Generator.....		39
II.2.3 VMD.....		39
II.3 RMSD.....		39
II.4 Radio de giro.....		40

II.5 Perfil de densidad parcial.....	40
II.6 Función de distribución radial.....	41
II.7 SASA.....	41
II.8 Análisis de puentes de hidrógeno.....	42
II.9 Línea temporal de las estructuras secundarias.....	42
II.9.1 Diagramas de Ramachandran.....	43
II.10 RMSF.....	44
Capítulo 3	45
III. Metodología.....	45
Capítulo 4	47
IV. Resultados y análisis.....	47
IV.1 RMSD.....	47
IV.2 Radio de giro.....	51
IV.3 Puentes de Hidrógeno.....	53
IV.4 Área de la superficie accesible al disolvente (SASA).....	57
IV.5 Función de distribución radial.....	59
IV.6 Perfil de densidad parcial.....	61
IV.7 Frecuencia de residuos en contacto.....	64
IV.8 Línea temporal de la estructura secundaria.....	70
IV.9 RMSF.....	75
V. Conclusiones.....	78
VI. Perspectivas a futuro.....	79
VII. Apéndice.....	80
VIII. Bibliografía.....	83

Capítulo 1

I. Introducción

El nuevo coronavirus responsable del síndrome respiratorio agudo grave (SARS-Cov-2), emergió en diciembre del 2019 en Wuhan, en la provincia de Hubei en China, como un patógeno humano[[1], [2]]. Es el tercer coronavirus en atravesar la barrera entre especies en tan solo dos décadas. Algunos estudios sugieren que el virus mutó del pangolín al humano y que la cepa que infectó al humano posee un 96.2% de similitud en la secuencia de ARN respecto al virus original[3].

El SARS-Cov-2 es un β -coronavirus perteneciente a la familia Coronaviridae (ver fig. I.1), esta familia comprende dos subfamilias: Letovirinae y Orthocoronavirinae. Particularmente, esta última subfamilia consiste de los cuatro géneros: α -coronavirus, β -coronavirus, γ -coronavirus y δ -coronavirus[4]. Tiene un ARN genómico monocatenario de sentido positivo con una longitud de aproximadamente 30 Kb[4], está recubierto por una bicapa lipídica que en conjunto con las proteínas estructurales, conforman el armazón del virus[5].

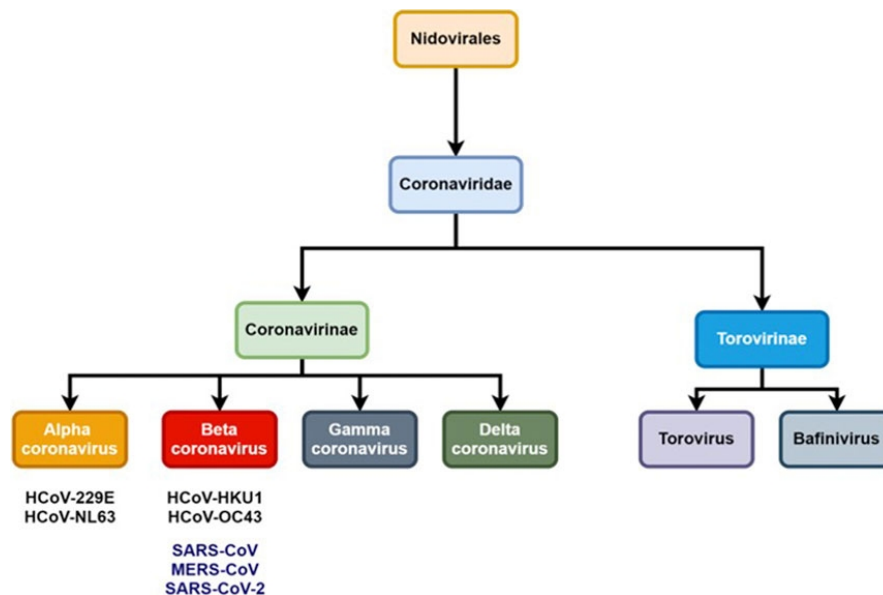


Figura I.1. Representación esquemática de la taxonomía Coronaviridae[0].

El SARS-Cov-2 está compuesto por cuatro proteínas estructurales (ver fig. I.2) conocidas como: Nucleocápside (N), envoltura (E), membrana (M) y glicoproteína espicular (S)[6]. La proteína N es codificada por el gen N, conforma la nucleocápside helicoidal uniéndose a lo largo de todo el genoma viral. Tiene el propósito de camuflajear al virus del sistema inmunitario. La proteína E, es un polipéptido pequeño que se encuentra en cantidades limitadas en la envoltura viral. Durante el ciclo de replicación, se expresa abundantemente dentro la célula infectada, pero solo una pequeña cantidad se

incorpora en la envoltura del viri3n. La mayor3a de la misma se encuentra localizada en el sitio de tr3fico intracelular, como el complejo de Golgi, donde participa en el ensamblado de la part3cula y se considera que es muy importante en la producci3n y maduraci3n del viri3n[6]. La prote3na M, es la prote3na estructural m3s abundante, y la responsable de darle la forma al viri3n. El mon3mero M, que oscila entre 25 y 30 kDa, es una prote3na de membrana que est3 incrustada en la envoltura a trav3s de tres dominios transmembrana. Asimismo, esta prote3na colabora en la fijaci3n de la nucleoc3pside a la membrana de estructuras internas tales como el complejo de Golgi y es la responsable del transporte transmembranal de nutrientes, la liberaci3n del viri3n y la formaci3n de la envoltura[7].

La protuberante glicoprote3na S se proyecta en forma de esp3culas, es un homotr3mero[4] que juega un importante papel en la infecci3n de c3lulas susceptibles al SARS-Cov-2, as3 como en interacciones con otros materiales. La prote3na S tiene un peso molecular de 180 kDa[2], es la mediadora de la entrada del virus SARS-Cov-2 en la c3lula humana que contiene la enzima convertidora angiotencina 2 (ACE2, por sus siglas en ingl3s), que sirve como una c3lula anfitriona de entrada[2]. Esta prote3na comprende la subunidad S1 con el dominio de uni3n al receptor (RBD, por sus siglas en ingl3s), responsable del enlace con el ACE2 y otra subunidad, S2, que permite el v3nculo del virus con la membrana celular despu3s del clivaje S1/S2 (ver fig.I.2.b).

El SARS-Cov-2, pat3geno productor de la enfermedad COVID-19, ha demostrado una contagiosidad y tasa de transmisi3n m3s alta durante la fase presintom3tica[8], en comparaci3n con ejemplares anteriores, pertenecientes a la misma familia, tales como SARS-Cov (2003) y MERS-Cov (2012)[2]. Raz3n por la que se ha convertido en una pandemia y se ha esparcido en m3s de 212 territorios cuya transmisi3n ha tomado lugar en pa3ses tales como: Estados Unidos, Alemania, Francia, Espa3a, Jap3n, Singapore, Corea del Sur, Ir3n e Italia. Reportando globalmente millones de casos y cientos de miles de muertes, con un r3pido incremento en el n3mero de casos[4].

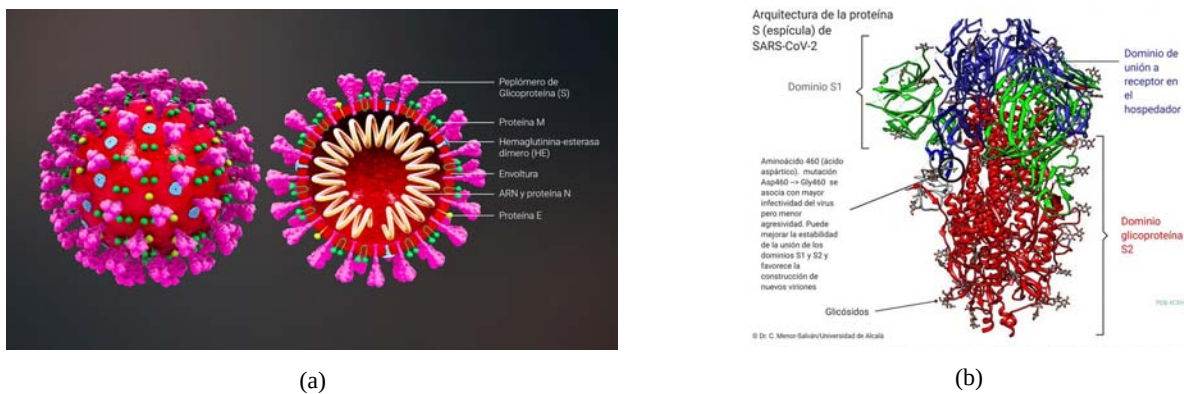


Figura I.2. (a) Secci3n transversal de un coronavirus indicando las prote3nas estructurales del SARS-Cov-2. (b) Estructura de la glicoprote3na esp3cular del SARS-Cov-2[9].

El entendimiento de los modos de transmisi3n de las cepas pand3micas es el primer paso para el desarrollo e implementaci3n de cambios estructurales a las ciudades y medios de transporte para mitigar la propagaci3n de las cepas. Aunque el conocimiento sobre la etiolog3a, patogenia, y la gravedad del COVID-19 ha aumentado considerablemente desde su brote en diciembre del 2019, todav3a hay una falta de consenso con respecto a sus rutas de transmisi3n y su importancia relativa en diferentes entornos. La Organizaci3n Mundial de la Salud (WHO, por sus siglas en ingl3s) y los Centros para el Control y Prevenci3n de enfermedades (CDC, por sus siglas en ingl3s) comparten una posici3n similar en este asunto, afirmando que el virus se espere principalmente por v3a directa o por

contacto cercano, sin descartar la transmisión indirecta por vía de contacto con superficies contaminadas (fómites, en la nomenclatura médica) y transmisiones por aerosoles en espacios cerrados, con poca ventilación y entornos específicos[10].

Mientras que la literatura carece de estudios en los modos de propagación del COVID-19 entre poblaciones humanas, el modelamiento y las simulaciones computacionales han permitido proveer un acercamiento en el esparcimiento del SARS-Cov-2 y de sus interacciones con el entorno.

Por ejemplo, el modelado teórico y la dinámica computacional de fluidos, han sido utilizados para el estudio de la transmisión en espacios interiores (salones de clase[11], aviones[12], autobuses[13], habitaciones[14], clínicas[15]).

En febrero del 2020 se realizó la identificación y caracterización completa del virus con resolución atómica de las estructuras de la glicoproteína espicular que al día de hoy contiene más de 400 estructuras relacionadas con el SARS-Cov-2. El incremento de datos experimentales se ha logrado en parte por técnicas de modelamiento y trayectorias de dinámica molecular (MD, por sus siglas en inglés). Algunos autores han empleado software de modelamiento para incluir características en la estructura del virus a las que no es posible acceder por medios experimentales (dominios transmembranales, por mencionar un ejemplo) y que se utilizan en el desarrollo de simulaciones de dinámica molecular [16]. Asimismo, las trayectorias obtenidas por medio de MD han permitido revelar características dinámicas de interés.

Otros equipos se han enfocado en el estudio teórico de implicaciones biomédicas como los mecanismos en el vínculo del SARS-Cov-2 con la célula receptora humana ACE2 y de la identificación de objetivos potenciales para el desarrollo de vacunas[[17], [18]]. Muchos estudios computacionales han investigado el efecto de pequeños compuestos moleculares con mayor afinidad a la glicoproteína espicular en la interacción RBD-ACE2 para inhibir su enlace[[19], [20]]. Aunque el tamaño del RBD es grande, solo unos pocos residuos se enlazan específicamente al ACE2 (ver tabla I.1), lo que dificulta la búsqueda de moléculas relativamente pequeñas que bloqueen selectivamente todo el RBD[19].

RBD-ACE2

Residuos	Tyr 449	Tyr 453	Asn 487	Tyr 489	Gly 496	Thr 500	Gly 502	Tyr 505
----------	---------	---------	---------	---------	---------	---------	---------	---------

Tabla I.1. Residuos relevantes (Tyr: Tirosina, Asn: Asparagina, Gly: Glicina y Thr: Treonina) etiquetados con su número en el enlace RBD-ACE2[5].

Algunos otros han puesto su atención en la capacidad de supervivencia del SARS-Cov-2 ante entornos hostiles; sin embargo, la información en este campo es limitada. Debido a que trabajar con este patógeno requiere de personal altamente calificado para trabajar bajo condiciones de contención: bioseguridad nivel tres (BSL3, por sus siglas en inglés)[10]. Esto pone en evidencia la utilidad de aproximaciones teóricas que sirvan como guía para futuros experimentos. En julio del 2020 se publicó un artículo en el que se analizó la estabilidad cinética del SARS-Cov-2 al rociar diferentes superficies con una muestra viral de tamaño equivalente en cada caso, sus hallazgos mostraron que el coronavirus era capaz de permanecer viable 72 h en una superficie de acero inoxidable, 48 h sobre una superficie de plástico, hasta 24 h en una superficie de cartón y solo 4 h sobre una superficie de cobre[10]. Estudios más recientes (Noviembre del 2020) reportaron el coeficiente de difusión del virus al alcanzar el 90% de disminución de una muestra sobre diferentes superficies en un intervalo de 4.8 h-9.1 días, dependiendo de la temperatura utilizada (20-40 °C)[10].

La resistencia de este coronavirus en superficies supone un importante motivo para las recomendaciones de las autoridades de la salud en todo el mundo en materia de desinfección de superficies que son manipuladas con mayor frecuencia.

De acuerdo con datos de la OMS, el 80% de las enfermedades contagiosas se propagan a través del contacto con fomites[21]. Existe una carencia en el entendimiento fundamental entre las interacciones de la glicoproteína espicular del SARS-Cov-2 y superficies a nivel fisicoquímico. Esta información, podría ser de utilidad en el diseño de futuros experimentos que involucren virus en superficies, para identificar materiales que actúen como barreras de protección y que incluso desactiven al SARS-Cov-2.

Hasta este momento y, hasta donde es de nuestro conocimiento, las interacciones a nanoescala del virus con superficies solo han sido materia de tres estudios en donde simulaciones de MD fueron utilizadas para su entendimiento. En el primero de ellos, se analizó la interacción de la glicoproteína espicular en un estado de prefusión con superficies de grafito y celulosa, superficies con diferentes propiedades y ampliamente utilizadas en adsorbentes y filtros[8]. Una superficie de grafito tiende a presentar una naturaleza hidrofóbica, una superficie de celulosa es un material con mayor capacidad de generar enlaces de hidrógeno debido a su naturaleza hidrofílica. De acuerdo con sus resultados, la interacción con la superficie de grafito indujo cambios estructurales más grandes en la cadena principal de la glicoproteína S, mientras que la interacción con la superficie de celulosa fue estabilizada por numerosos puentes de hidrógeno, en consecuencia no indujo cambios estructurales en la cadena principal de la glicoproteína S.

En el segundo trabajo, el objeto de estudio fue la interacción entre la glicoproteína espicular en un estado de postfusión con las superficies: Aluminio (Al), dióxido de cobre (CuO_2), Polietileno (PE) y dióxido de silicio (SiO_2). Estas interacciones estuvieron sujetas a diferentes temperaturas. Sus resultados indicaron que la adsorción con mayor afinidad fue sobre la superficie de SiO_2 , asociada a los puentes de hidrógeno generados en la interfase de la interacción. En comparación con la adsorción con menor afinidad que resultó ser sobre la superficie de PE, debido a que presentó la naturaleza más hidrofóbica de entre todas las estudiadas. Por otro lado, sus datos indicaron que la adsorción no se vio significativamente influenciada al incrementar la temperatura[10].

El tercer y más reciente estudio analiza los componentes de la glicoproteína S al interactuar con una superficie de grafeno. En virtud de sus resultados, la glicoproteína S sufrió las deformaciones más abruptas en la cadena principal durante la etapa principal de la adsorción. Observaron cambios en las estructuras secundarias de la subunidad S1 y S2, asociadas con la interacción con grafeno, siendo el RBD la estructura que presentó el comportamiento más rígido durante la adsorción[2].

Por otra parte, otras referencias afirman que simulaciones de MD de proteínas interactuando con grafeno han revelado presentar una fuerte adsorción y cambios remarcables en su estructura secundaria, inducidos por este proceso[[22], [23]].

Teniendo en consideración que el protómero RBD aislado representa una pequeña sección de la glicoproteína S, este fragmento puede representar un modelo útil y menos costoso para modelar la interacción del SARS-Cov-2 con diferentes superficies. Hasta donde es de nuestro conocimiento, estas son las únicas referencias disponibles que permiten comparar la descripción de la interacción tomando en cuenta la proteína S completa y solamente el RBD. Con esta motivación, resulta muy interesante explorar el fenómeno de adsorción del virus tomando como modelo el RBD sobre diferentes superficies. De manera general, el presente trabajo se enfoca en los cambios en la adsorción del protómero RBD en diferentes superficies y el efecto desactivante de las mismas. Se estudió la adsorción en una superficie de cobre en la dirección (100), en una superficie de oro en la dirección (111) y en una superficie de grafito en la dirección (001). De esta manera será posible responder la pregunta planteada desde el inicio de la pandemia acerca de la posible desactivación del virus por contacto con diferentes superficies, además de la gran ventaja que representa el poder modelar el

comportamiento de la proteína S usando únicamente el RBD, lo que representa un considerable ahorro de tiempo de cómputo.

I.1 Objetivos

I.1.1 Objetivo General

- Estudiar la adsorción del protómero (RBD) en SARS-Cov-2 en diferentes conocidas superficies antivirales como: Cu, Au y C_G (grafito) por medio de MD. Con ello, comprender aspectos que puedan favorecer la inhibición de la acción del fragmento RBD de la proteína.

I.1.2 Objetivos Específicos

- Estudiar las trayectorias de MD en cada interacción del RBD con las tres superficies y del RBD aislado mediante: RMSD, radio de giro, puentes de hidrógeno entre estructuras secundarias del RBD, puentes de hidrógeno entre el RBD y moléculas de agua en el disolvente, SASA, perfil de densidad parcial, función de distribución radial, líneas temporales de la estructura secundaria del RBD y RMSF.

I.2 Hipótesis

A partir del estudio teórico de la interacción de diferentes superficies^(a) con el virus^(b) mediante dinámica molecular se espera lograr una propuesta de cambios químicos en la composición del protómero, haciendo que pierda su funcionalidad.

(a) Superficies: Au (111), Cu (100) y C-grafito (001).

(b) Protómero RBD de la subunidad S1, de la proteína S del virus Sars-Cov2.

Capítulo 2

II. Marco Teórico

La física detrás de la dinámica molecular.

En las últimas décadas, el progreso en el desarrollo de dispositivos con mayor poder de cómputo, velocidad y eficiencia ha permitido la optimización de algoritmos que han favorecido el desarrollo de diversas técnicas computacionales al punto de ser consideradas como un complemento a técnicas experimentales debido a la precisión de los cálculos en los resultados del estudio y predicción de fenómenos y propiedades en simulaciones. Particularmente, su implementación en la física computacional ha dado pie a una variedad de contribuciones en todas las ramas de la física así como de otras disciplinas en la ciencia. La eficiencia de la física computacional para plantear la solución de un problema radica en el enfoque interdisciplinario con el que se aborda el fenómeno de estudio, considerando las condiciones del entorno, la escala del sistema, el tiempo del proceso así como de las aproximaciones teóricas más adecuadas para su descripción, por mencionar algunas.

En la figura II.1 puede ser visto el amplio rango de escalas de tiempo y longitud que las simulaciones computacionales cubren. Cualquier método puede llevarse a cabo bajo las subyacentes teorías físicas para describir el modelo de partícula del sistema de estudio. Si se utiliza una descripción de mecánica cuántica (QM, por sus siglas en inglés), los electrones son explícitamente representados en el modelo y la energía de interacción es calculada al resolver la estructura electrónica de las moléculas en el sistema[24]. Las simulaciones de QM pueden llegar a ser muy caras computacionalmente, dependiendo directamente del método elegido, desde métodos de alto nivel como ab-initio hasta métodos semiempíricos y pueden ser utilizados eficientemente en el rastreo de cientos de átomos o menos.

En una descripción del siguiente rango de orden de magnitud, está la mecánica molecular (MM, por sus siglas en inglés), las moléculas son representadas por partículas, representando átomos o grupos de átomos. A cada átomo se le puede asignar una carga eléctrica y una función de energía potencial, de la que se hablará más adelante, con un gran número de parámetros empíricos que son utilizados para el cálculo de interacciones de enlace y de no-enlace en dinámica molecular (MD, por sus siglas en inglés) para calcular las fuerzas que determinan la dinámica del sistema. A diferencia de las simulaciones QM, estas son más rápidas, haciendo que sean el método elegido para la mayoría de estudios de simulaciones moleculares en sistemas biomoleculares en fase condensada[24]; sin embargo, son menos precisas que las simulaciones QM y no pueden simular reordenamientos de enlaces.

Una descripción de sistemas a mayor escala, del orden de micrómetros, requiere de un método distinto. Para ello, los métodos de grano-grosso (CG, por sus siglas en inglés) o mesoescala son utilizados. Los métodos de grano-grosso reducen el número de pares explícitos que son necesarios para el cálculo de la energía y fuerza de un sistema en particular y reducen el tiempo de cómputo, o alternativamente, permiten el estudio de un sistema más grande. Comúnmente, un incremento en la longitud de escala característica en el modelo conllevará a un incremento en el tiempo que podemos utilizar en la

simulación del problema. Por ello, el método de grano-grueso permitirá utilizar un tiempo de paso más grande para cubrir un intervalo de tiempo equivalente al del fenómeno “real” en nuestra simulación[24].

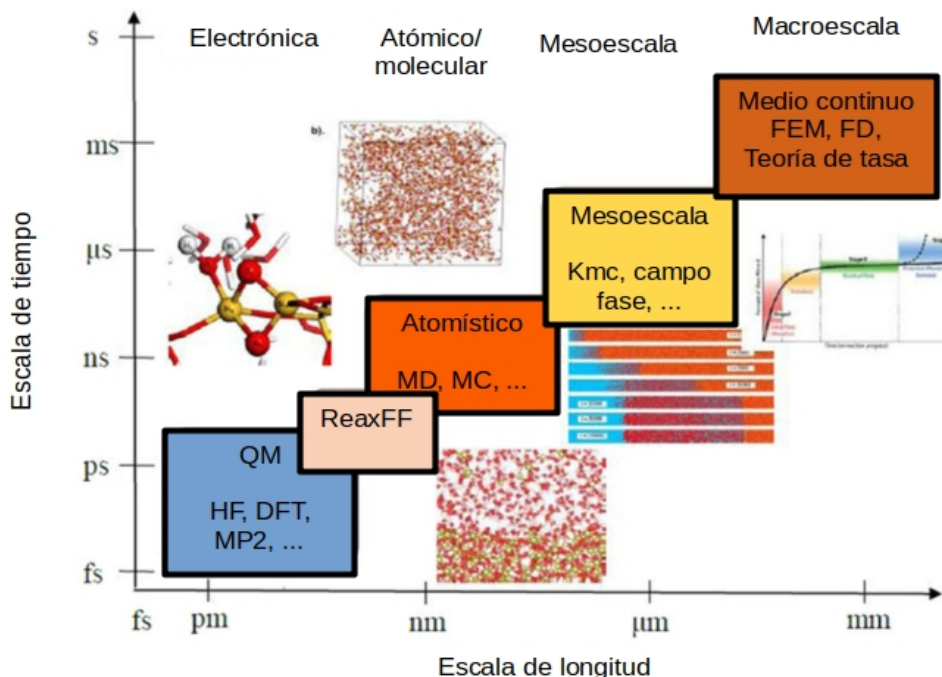


Figura II.1. Aplicabilidad de métodos de simulación para escalas de tiempo y longitud[25].

La velocidad de cómputo es un tema relevante en la descripción de sistemas en fase condensada, dado que usualmente estamos interesados en las propiedades de moléculas o bio-macromoléculas en disolución (sistemas de cientos de miles o millones de átomos). A manera de ilustración, en el caso del cálculo de energías libres o propiedades de transporte a la temperatura ambiente común en un laboratorio, existe una contribución entrópica que afectan el cálculo de las propiedades de interés.

De forma que las simulaciones no solo muestreen estados óptimos individuales, sino que muestreen una distribución adecuada de estados. Un balance adecuado entre el tiempo de computo y la precisión necesaria para la descripción nos dará la técnica adecuada para el estudio del fenómeno de interés.

Por lo que en el resto de este trabajo nos restringiremos solamente al uso de dinámica molecular.

II.1 Dinámica Molecular

De entre la gran variedad de aproximaciones teóricas disponibles, la dinámica molecular es una poderosa herramienta computacional capaz de manejar sistemas con una amplia gama de escalas y proveer un avistamiento de fenómenos nanoscópicos y microscópicos que de otra manera serían muy difíciles de observar o tendrían una justificación limitada si se utilizara alguna de las técnicas experimentales convencionales de caracterización. Esta técnica fue introducida por primera vez en 1959 por Alder and Wainwright[26].

En los sistemas biológicos, la dinámica molecular permite estudiar la evolución de procesos dinámicos y complejos como: cambios conformacionales, estabilidad de una proteína, plegamiento/desplegamiento de una proteína, transporte iónico de sistemas biológicos, reconocimiento molecular, determinación de una estructura por rayos X o RMN, entre otras.

En polímeros biológicos o sintéticos, el cambio de conformación requiere de largos intervalos de tiempo que deben ser capturados para obtener cálculos precisos de sus propiedades.

Las moléculas biológicas exhiben una amplia escala de tiempos (ver fig. II.2) sobre los cuales ocurren procesos específicos. En proteínas, las escalas de tiempo van desde los nanosegundos (ns) hasta los segundos o más, el plegamiento de cadenas hidrofóbicas de aminoácidos puede tomar microsegundos (μs) o más. Mientras que cambios conformacionales más grandes pueden llegar a necesitar de más tiempo[[27], [28]].

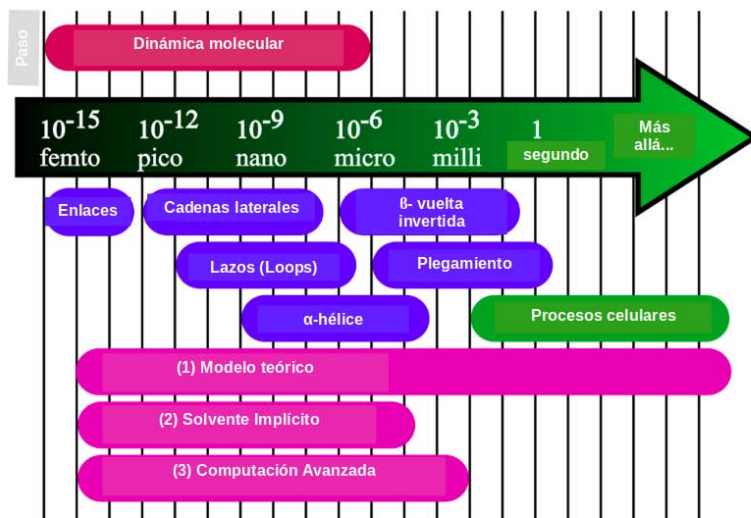


Figura II.2. Escala de tiempos en el cambio de diferentes cambios conformacionales en fenómenos biológicos estudiados con MD[29].

La MD es una manifestación moderna de una antigua idea de la ciencia que establece que el estado de equilibrio y el movimiento de un sistema multiconstituyente pueden ser obtenidos si se conocen las condiciones de frontera y las interacciones entre los componentes del sistema. Las interacciones están definidas por una función de energía potencial, con la que se calcula la energía del sistema. En la figura II.3 se muestra el digrama de flujo del procedimiento que comunmente se utiliza en simulaciones de MD. La simulación comienza desde un estado inicial bajo un conjunto de constricciones termodinámicas definidas por un ensamble, la energía del sistema se actualiza paso por paso. Donde en cada paso las ecuaciones de Newton son resueltas por técnicas de integración basadas en métodos de diferencias finitas, para calcular las fuerzas de los componentes, derivar las propiedades del sistema y construir su trayectoria.

II.1.1 Ecuaciones de movimiento

En el más sencillo de los casos, las ecuaciones de Newton son resueltas para un sistema de N -partículas, especificando el vector de posición $\mathbf{r}(t)$ de una partícula en el espacio a cualquier tiempo t . Lo que es equivalente a especificar tres funciones de tiempo, los componentes del vector $\mathbf{r}(t)$, $\mathbf{r}(t) = (x(t), y(t), z(t))$. Para una partícula i , las ecuaciones de Newton toman la forma:

$$m_i \mathbf{a}_i = \mathbf{F}_i(\mathbf{r}_1, \dots, \mathbf{r}_n, \mathbf{v}_i) \quad (1)$$

Donde m_i es la masa de la partícula i , \mathbf{F}_i es la fuerza ejercida sobre la partícula y \mathbf{a}_i es la aceleración inducida sobre la partícula i .

Estas ecuaciones son referidas como las ecuaciones de movimiento del sistema, su solución está sujeta al conjunto de posiciones iniciales, $\{\mathbf{r}_1(0), \dots, \mathbf{r}_N(0)\}$, y velocidades, $\{\mathbf{v}_1(0), \dots, \mathbf{v}_N(0)\}$.

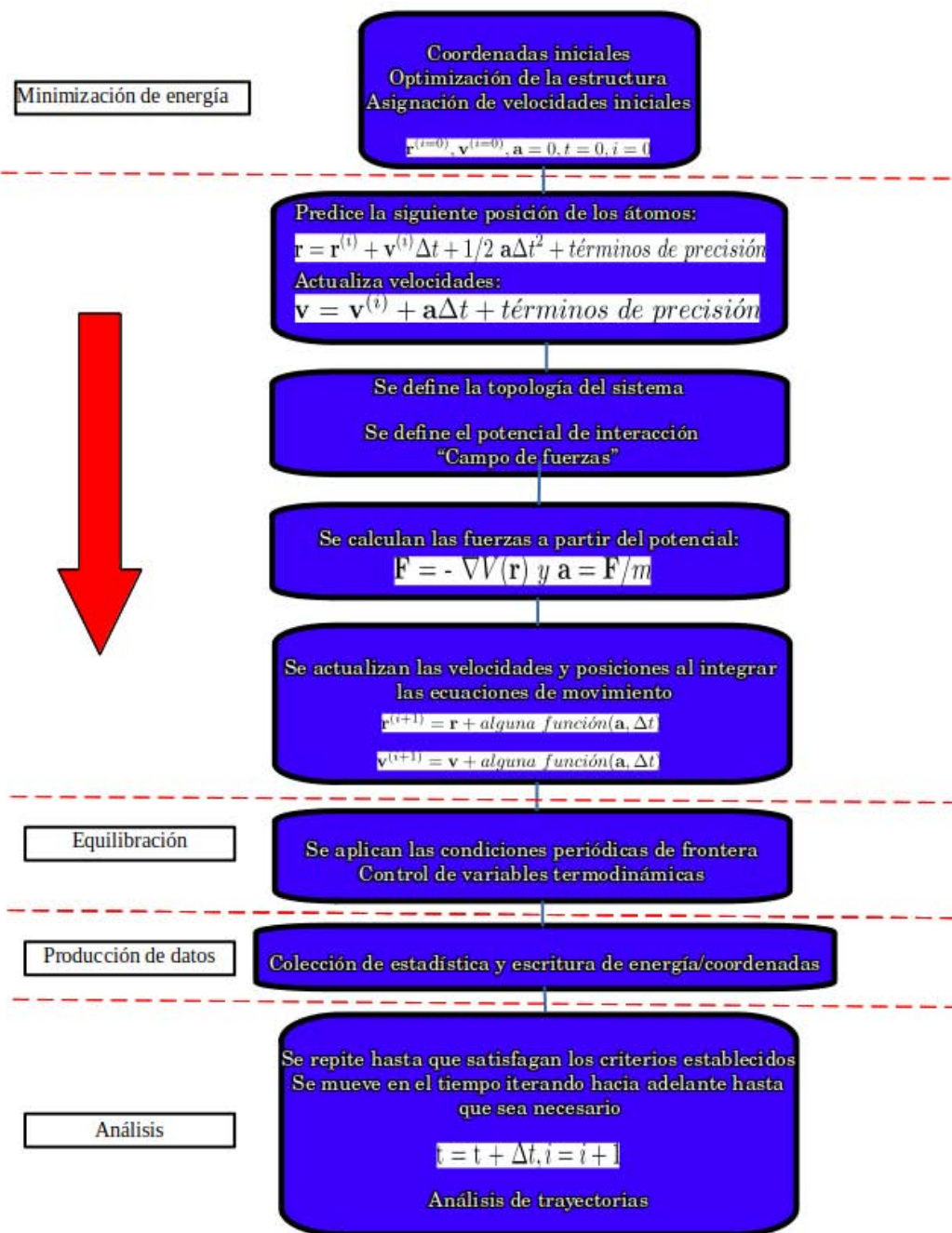


Figura II.3. Diagrama de flujo del procedimiento estándar en simulaciones de dinámica molecular.

II.1.2 Métodos de diferencias finitas

Una aproximación estándar en la resolución de ecuaciones diferenciales ordinarias son los métodos de diferencias finitas, donde el tiempo es discretizado en una red finita, el tiempo de paso, δt , es la diferencia entre dos puntos consecutivos en la red. Si se tiene conocimiento de las posiciones, velocidades y alguna otra propiedad dinámica de los átomos del sistema al tiempo t , se calculan las posiciones, velocidades y otras propiedades al tiempo $t + \delta t$ a un cierto grado de precisión.

Las ecuaciones de movimiento son integradas paso por paso, la elección del tiempo de paso, δt , dependerá del método de solución[24].

Los algoritmos de integración más eficientes cuentan con una serie de cinco características importante a considerar:

1. Son rápidos y ocupan poca memoria.
2. Permiten el uso de un δt grande.
3. La trayectoria es reproducible.
4. Satisfacen las leyes de conservación de la energía, momento y son reversibles en el tiempo.
5. Simples y fáciles de programar.

II.1.2.1 Algoritmo de Verlet

Este método es derivado directamente de la expansión de la ecuación (1) en series de Taylor de segundo orden alrededor de $\mathbf{r}(t)$, como sigue:

$$\begin{aligned}\mathbf{r}(t+\delta t) &\approx \mathbf{r}(t) + \delta t \mathbf{v}(t) + \frac{1}{2} \delta t^2 \mathbf{a}(t) + \dots \\ \mathbf{r}(t-\delta t) &\approx \mathbf{r}(t) - \delta t \mathbf{v}(t) + \frac{1}{2} \delta t^2 \mathbf{a}(t) - \dots\end{aligned}\tag{2}$$

Al sumar las expresiones en (2) podemos conocer la posición en el siguiente paso, $\mathbf{r}(t + \delta t)$ a partir de las posiciones actuales, $\mathbf{r}(t)$, las aceleraciones, $\mathbf{a}(t)$ y las posiciones del paso anterior, $\mathbf{r}(t - \delta t)$.

$$\mathbf{r}(t+\delta t) \approx 2\mathbf{r}(t) - \mathbf{r}(t-\delta t) + \delta t^2 \mathbf{a}(t)\tag{3}$$

Este algoritmo es reversible en el tiempo; es decir, $\mathbf{r}(t - \delta t)$ y $\mathbf{r}(t + \delta t)$ son simétricos alrededor del punto de origen. Esto significa que si se corre en una dirección y después en dirección contraria, el sistema va a regresar al punto de inicio. Por ende, la energía total del sistema será conservada, lo que es importante para una dinámica físicamente consistente.

La velocidad no es esencial para el cálculo de las trayectorias. No obstante, permite estimar la energía cinética y la energía total, así como otras propiedades del sistema de interés.

De acuerdo con Störmer[30], a partir de la resta de las expresiones en la ecuación (2), es posible obtener la expresión de la velocidad, como:

$$\mathbf{v}(t) \approx \frac{\mathbf{r}(t+\delta t) - \mathbf{r}(t-\delta t)}{2\delta t} \quad (4)$$

Donde la propagación del error en la expresión (3) es del orden de δt^4 , mientras que en la expresión (4) es del orden δt^2 [30].

II.1.2.2 Algoritmo de velocidad de Verlet

Este algoritmo es una variante del algoritmo de Verlet, propuesto por Swope en 1982[31]. A diferencia del algoritmo anterior donde solo se calcula la posición del siguiente paso de manera explícita, en este algoritmo se puede obtener la evolución de la posición y velocidad de forma explícita, haciendo de este una formulación elegante al contener las dos variables presentes en el espacio fase.

En la implementación más eficiente, actúa sobre un solo tiempo de paso de t a $t + \delta t$, como sigue:

$$\mathbf{v}(t + \frac{1}{2}\delta t) \approx \mathbf{v}(t) + \frac{1}{2}\delta t \mathbf{a}(t) \quad (5.a)$$

$$\mathbf{r}(t + \delta t) \approx \mathbf{r}(t) + \delta t \mathbf{v}(t + \frac{1}{2}\delta t) \quad (5.b)$$

$$\mathbf{v}(t + \delta t) \approx \mathbf{v}(t + \frac{1}{2}\delta t) + \frac{1}{2}\delta t \mathbf{a}(t + \delta t) \quad (5.c)$$

A partir de la expresión (5.a) se lleva la velocidad del tiempo t al $t + \frac{1}{2}\delta t$, utilizando los valores de las aceleraciones al tiempo t . Estas nuevas velocidades son utilizadas para calcular las posiciones del tiempo t al $t + \delta t$ en la expresión (5.b). Finalmente, se obtiene la fuerza para calcular $\mathbf{a}(t + \delta t)$ con la expresión (5.c) que completa la evolución de las velocidades[31].

II.1.2.3 Algoritmo de salto de rana

Este algoritmo es otra modificación del algoritmo de Verlet, cuyo nombre se debe a la forma de su implementación. Las cantidades guardadas son las posiciones actuales, $\mathbf{r}(t)$, y las aceleraciones, $\mathbf{a}(t)$, junto con las velocidades evaluadas en el punto medio del intervalo de tiempo, $\mathbf{v}(t - \frac{1}{2}\delta t)$, cuya expresión para calcular la siguiente posición es idéntica a la expresión (5.b), mientras que la expresión para $\mathbf{v}(t + \frac{1}{2}\delta t)$ puede ser derivada al combinar las ecuaciones (5.a) y (5.c) en el algoritmo de velocidad de Verlet para un paso previo ($t = t - \delta t$).

$$\mathbf{v}(t+\frac{1}{2}\delta t)\approx\mathbf{v}(t-\frac{1}{2}\delta t)+\delta t\mathbf{a}(t) \quad (6.a)$$

$$\mathbf{r}(t+\delta t)\approx\mathbf{r}(t)+\delta t\mathbf{v}(t+\frac{1}{2}\delta t) \quad (6.b)$$

La ecuación (6.a) es implementada primero de forma que las velocidades “saltan” sobre las coordenadas para los siguientes valores de las velocidades (ver fig. II.1.2.2) en el punto medio del intervalo de tiempo, $\mathbf{v}(t + \frac{1}{2}\delta t)$.

Durante este paso, las velocidades actuales puede ser obtenidas como:

$$\mathbf{v}(t)\approx\frac{1}{2}(\mathbf{v}(t+\frac{1}{2}\delta t)+\mathbf{v}(t-\frac{1}{2}\delta t)) \quad (7)$$

La expresión en (7) es necesaria para calcular la energía al tiempo t , así como otras cantidades dependientes de la posición y velocidad en el mismo instante.

La segunda expresión, (6.b), es utilizada para saltar las posiciones por encima de las velocidades. Después, las nuevas aceleraciones son obtenidas para el siguiente paso[24].

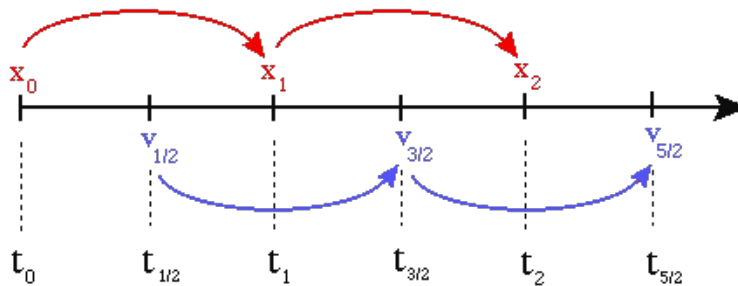


Figura. II.1.2.2. Esquema del algoritmo de salto de rana.

II.1.3 Campo de fuerzas

Un campo de fuerza, en el contexto de las simulaciones computacionales, se refiere a formas funcionales de la energía potencial, utilizadas para describir las interacciones inter- e intramoleculares de una colección de átomos. Donde el funcional y los correspondientes parámetros dentro del mismo, determinarán la energía potencial del sistema. Dichos parámetros se obtienen de experimentos en moléculas aisladas y cálculos precisos de mecánica cuántica. En la práctica, estos parámetros usualmente son refinados al comparar propiedades calculadas por medio de simulaciones computacionales con propiedades experimentales[24].

Los campos de fuerza son usualmente divididos en tres clases. En el presente trabajo solo desglosaremos los campos de fuerza de clase I dado que es la categoría a la que pertenece el utilizado.

Los campos de fuerza de clase I, normalmente tienen una forma funcional del tipo:

$$U(R) = \frac{1}{2} \sum_{\text{enlaces}} k_r (r_{ij} - r_{eq})^2 + \frac{1}{2} \sum_{\text{ángulos}} k_\theta (\theta_{ijk} - \theta_{eq})^2 + \sum_{\text{diedros}} \sum_n k_{\phi,n} (1 + \cos(n\phi_{ijkl} + \delta_n)) + 4 \sum_{i < j}^{\text{átomos}} \epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \sum_{i < j}^{\text{átomos}} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \quad (8)$$

Los primeros tres términos en la ecuación (8) corresponden a interacciones de enlace (intramoleculares), donde los átomos están conectados entre ellos. El primer término es la energía de la suma sobre todos los pares ij de átomos covalentemente enlazados, que son tratados como osciladores armónicos (ver fig. II.1.3.1) con longitud de enlace, r , una longitud de enlace en equilibrio, r_{eq} y cada uno con su propia constante de fuerza, k_r .

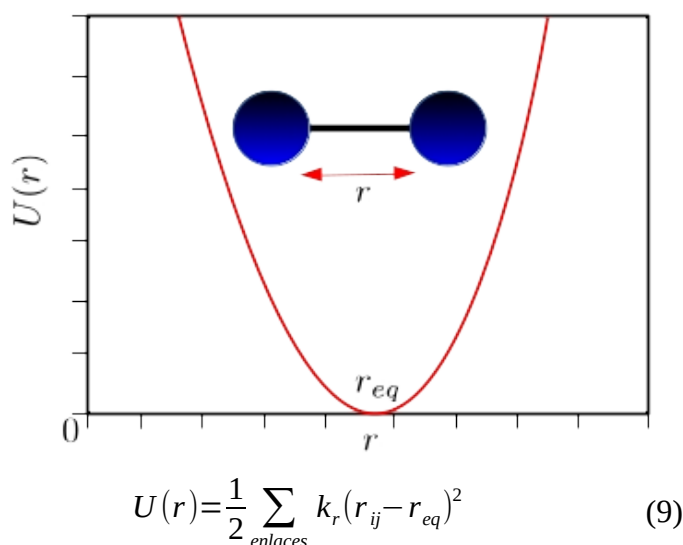
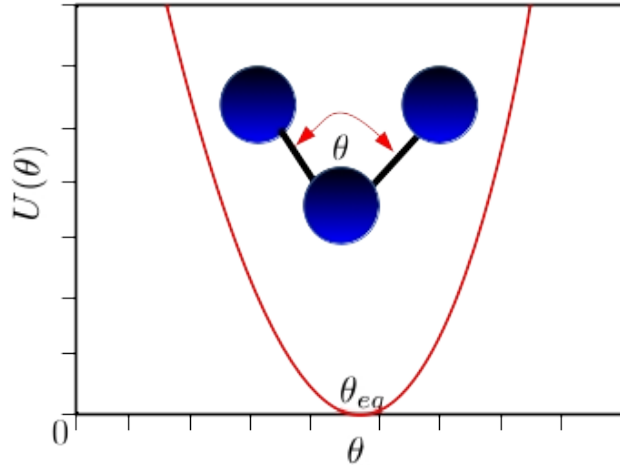


Figura II.1.3.1. Potencial de enlace.

El segundo término es la energía de flexión de todos los ángulos de enlaces covalentes de los átomos vecinos. Hay un término por cada conjunto de tres átomos conectados ijk y el movimiento se modela con un potencial armónico (ver fig. II.1.3.2) en el ángulo de flexión, θ , un ángulo de flexión en equilibrio, θ_{eq} y cada uno tiene su propia constante de flexión, k_θ .

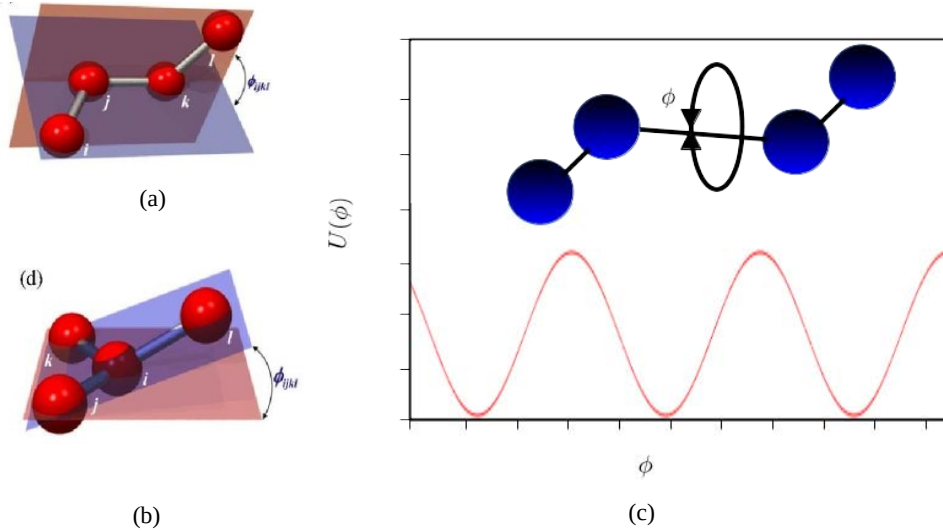


$$U(\theta) = \frac{1}{2} \sum_{\text{ángulos}} k_{\theta} (\theta_{ijk} - \theta_{eq})^2 \quad (10)$$

Figura II.1.3.2. Potencial de ángulos de flexión.

El tercer término es la energía conformacional, asociada con la suma sobre todos los ángulos diedros o de torsión (ϕ_{ijkl}) entre cuatro átomos conectados $ijkl$, donde ϕ_{ijkl} es el ángulo entre los planos ijk y $jkil$ [32]. Existen dos tipos de ángulos de torsión: ángulos de torsión “propios”, alrededor de enlaces que conectan un grupo de átomos (ver fig. II.1.3.3.a) y ángulos de torsión “impropios”, que involucran vecinos alrededor de un átomo central (ver fig. II.1.3.3.b).

Este término es una expansión en funciones trigonométricas (ver fig. II.1.3.3.c) con diferentes valores para la multiplicidad, n , (el número de mínimos en una rotación de 2π alrededor del enlace $j-k$), δ_n es el factor de fase que determina los puntos donde el potencial de torsión pasa por un valor mínimo[33] y $k_{\phi,n}$ es la constante de torsión en cada término n .



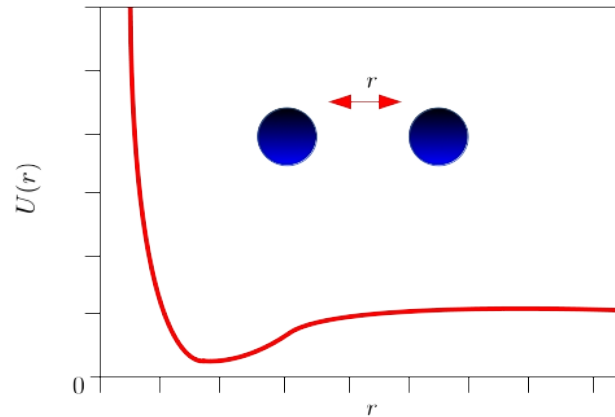
$$U_{\phi} = \sum_{\text{diedros}} \sum_n k_{\phi,n} (1 + \cos(n\phi_{ijkl} + \delta_n)) \quad (11)$$

Figura II.1.3.3. (a) Ilustración de ángulo diedro propio, (b) ángulo diedro impropio y (c) potencial de torsiones.

Los últimos dos términos de la ecuación (8) corresponden a interacciones de no-enlace (intermoleculares e intramoleculares). El cuarto término es la energía debido a interacciones no polares modeladas por el potencial de Lennard-Jones (LJ), también referido como potencial 12-6 (ver fig. II.1.3.4). La naturaleza dinámica del movimiento de los electrones en moléculas permite una distorsión temporal de la distribución electrónica de carga, conllevando a momentos dipolares temporales distintos de cero para la molécula. Estos dipolos instantáneos temporales pueden polarizar moléculas vecinas, llevando a la formación de un dipolo inducido en el vecino. Estas distorsiones en la distribución de carga conllevan a fuerzas y potenciales atractivos llamados fuerzas de dispersión de London (dipolo inducido-dipolo instantáneo) y dipolo-dipolo inducido. En el potencial de LJ estas interacciones son descritas por un término atractivo cuya magnitud varía con la distancia de la forma $\sim -1/r^6$ y depende de la polarizabilidad y energía de ionización de las especies involucradas[27].

De igual forma, el potencial de LJ incluye un término repulsivo de corto alcance donde las distribuciones electrónicas de moléculas conllevan a fuertes repulsiones estéricas. Estas repulsiones, también conocidas como repulsiones de Pauli, deben su nombre al principio de exclusión de Pauli en la mecánica cuántica, el cual prohíbe la presencia de dos o más electrones con los mismos números cuánticos dentro de un estado atómico o molecular en un orbital.

Las repulsiones de Pauli son muy fuertes a distancias pequeñas, pero decaen rápidamente conforme los átomos se separan y el traslape entre sus nubes electrónicas de probabilidad decrece. Las repulsiones de Pauli son típicamente expresadas con un término exponencial de rápido decaimiento o, en el caso del potencial de LJ como $\sim 1/r^{12}$ para distancias menores a $r \sim \sigma$, donde σ , conocido como diámetro de colisión, es la distancia mínima a la cual el potencial de interacción entre dos átomos es cero en un modelo de esferas rígidas (radio de Van der Waals) y ϵ es la profundidad del potencial en el mínimo y r_{ij} es la distancia entre los átomos i y j [[27], [34]].



$$U(r) = 4 \sum_{i < j}^{\text{átomos}} \epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] \quad (12)$$

Figura II.1.3.4. Potencial de Lennard-Jones.

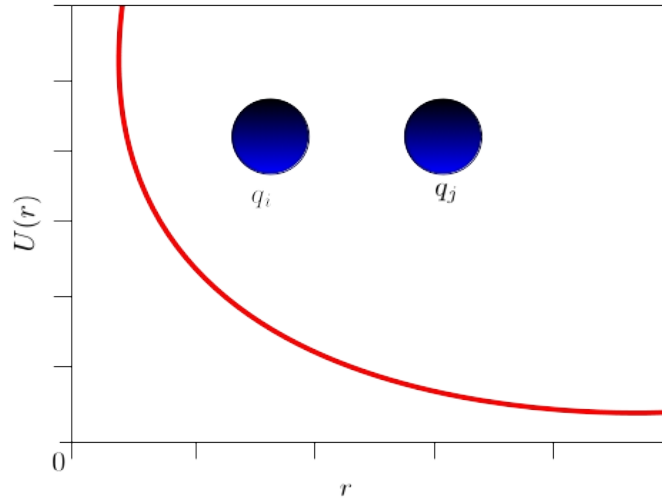
Las interacciones cruzadas para los parámetros del potencial de LJ son usualmente obtenidos por medio de la media aritmética para σ_{ij} y la media geométrica para ϵ_{ij} en las reglas de combinación de Lorentz-Berthelot[24], definidas como:

$$\begin{aligned} \sigma_{ij} &= \frac{1}{2} (\sigma_i + \sigma_j) \\ \epsilon_{ij} &= (\epsilon_i \epsilon_j)^{1/2} \end{aligned} \quad (13)$$

Las interacciones cruzadas también pueden ser determinadas directamente ajustando por medio de un experimento si son relevantes en el modelo, o mediante la fórmula de Slater-Kirkwood[28], dependiendo del campo de fuerzas utilizado.

La repulsión de Pauli, las fuerzas de London, atracciones dipolo-dipolo inducido y dipolo-dipolo son colectivamente conocidas como interacciones de van der Waals[[27], [34]].

El quinto y último término corresponde a la energía necesaria para separar dos cargas a una distancia infinita. Esta contribución corresponde a interacciones de largo alcance ya que se esparce a lo largo del espacio. En campos de fuerza de clase I, este término corresponde al potencial de Coulomb (ver fig. II.1.3.5) que describe la interacción entre las cargas parciales puntuales q_i y q_j , que representan las interacciones electrostáticas entre las moléculas, donde ϵ_0 es la permitividad en el vacío y r_{ij} es la distancia entre ambas cargas.



$$U(r) = q_i \phi(r) = \sum_{i < j}^{\text{átomos}} \frac{q_i q_j}{4 \pi \epsilon_0 r_{ij}} \quad (14)$$

Figura II.1.3.5. Potencial de Coulomb.

Esta aproximación permite describir el potencial electrostático a partir de un conjunto de cargas puntuales y fijas, asignadas a posiciones nucleares, R_j , en la molécula. Los valores de las cargas puntuales son asignados para dar el mejor ajuste para el potencial electrostático real ($\phi[\rho(r)]$) en alguna red de puntos a una distancia mínima de la molécula (ver fig. II.1.3.6). Si $\phi[\rho(r)]$ es el potencial electrostático del núcleo y la función de onda electrónica de la molécula a algún punto de la molécula, determinada por cálculos de mecánica cuántica, el potencial electrostático aproximado para una molécula con N átomos con cargas electrostáticas puntuales q_j es determinado como :

$$U_{\text{aprox}} = q_i \phi_{\text{aprox}} = q_i \sum_{j=1}^N \frac{q_j}{4 \pi \epsilon_0 r_{ij}} \frac{1}{r_{ij}} \approx q_i \phi[\rho(r)] \quad (15)$$

Esta relación solo es válida al aplicarla a distancias mayores al radio de van der Waals de los átomos de la molécula. Mientras que las cargas nucleares son positivas, las cargas puntuales incluyen el efecto neto de las cargas nucleares y electrónicas en la vecindad de un sitio nuclear y pueden ser positivas o negativas. Estas cargas puntuales parciales no son cantidades físicamente observables. No obstante, puede asignarse significado físico a sus valores[27].

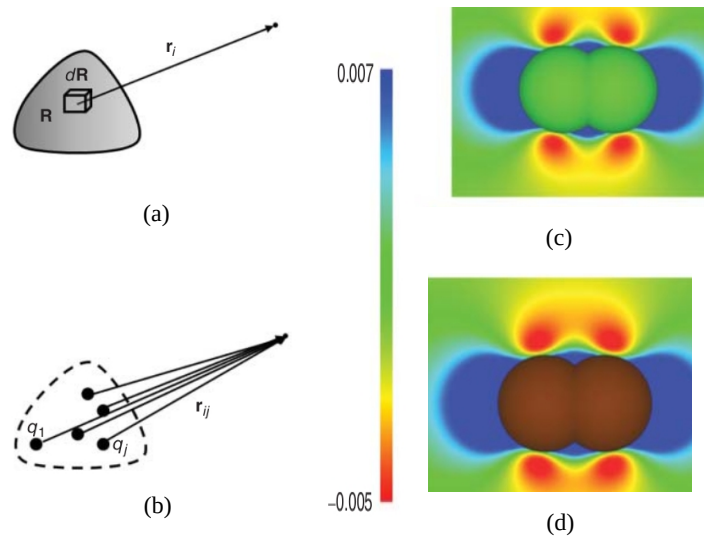


Figura II.1.3.6. Comparación entre un potencial electrostático obtenido en (c), por el modelo de distribución de carga continua en (a) y un potencial electrostático obtenido en (d), por el modelo de distribución de cargas puntuales en (b).

Dependiendo del campo de fuerzas para la descripción de un fenómeno, pueden incluirse términos adicionales, tales como: interacciones entre multipolos fijos, sitios polarizables u otros términos específicos.

Los campos de fuerza tipo II incluyen términos cúbicos al potencial de enlace y definen explícitamente los elementos fuera de la diagonal en la matriz de la constante de fuerza. Mientras que los campos de fuerza tipo III incluyen representaciones más adecuadas para describir las interacciones electrostáticas entre moléculas y la inclusión de polarizabilidad.

II.1.4 Condiciones periódicas de frontera

Las condiciones periódicas de frontera (PBC, por sus siglas en inglés) permiten una estimación más precisa para propiedades de bulto en simulaciones de sistemas finitos, esencialmente para sistemas en nanoescala. En las simulaciones, la limitación del tamaño del sistema o el número de moléculas es abordada implementando PBC a la celda principal de la simulación. La periodicidad implica que todas las interacciones a las que esta sometido un átomo en el bulto están presentes en la celda principal de simulación.

Las PBC también determinan el comportamiento de las moléculas una vez que alcanzan los bordes de la celda de simulación. Específicamente, determinan si las moléculas en el sistema permanecieran o no dentro de los confines de las otras moléculas en la celda de simulación[[27], [35]]. La incorporación de PBC es equivalente a considerar un arreglo donde la celda de simulación se replica en todo el espacio para formar una red infinita de celdas imagen periódicas. Hay dos consecuencias de esta periodicidad:

- Primera, durante el transcurso de una simulación, si un átomo se mueve en la celda principal, su imagen periódica en cada una de las celdas imagen que la rodean se moverá de la misma forma

(ver fig. II.1.4). O en el caso en que un átomo salga de la celda principal, una de sus imágenes entrará inmediatamente por la cara opuesta de la celda principal.

A manera de ilustración, la figura II.1.4 muestra el caso bidimensional de un celda central (sombreada en gris) con celdas imagen a sus alrededores. Las celdas imagen están etiquetadas como A, B, C, etc., con un orden espiral en sentido levógiro desde la celda A, cada una de ellas tiene asociado un vector $\mathbf{n} = (n_x, n_y)$, donde $x, y \in \mathbb{Z}$. Por ejemplo, si las coordenadas del átomo 5 son (x_i, y_i) , las coordenadas del mismo átomo en la celda imagen serán:

$$(x_i, y_i)_n \Rightarrow (x_i, y_i) + (n_x + L_x, n_y + L_y)$$

Por otro lado, los átomos cerca de los bordes de la celda principal están completamente coordinados; es decir, interactúan con los átomos vecinos en las celdas imagen.

- Segunda, para evitar que un solo átomo interactúe con imágenes periódicas en el mismo sistema varias veces, los átomos en la celda principal se encuentran dentro de una distancia (conocida como radio de corte, r_c , de la que se hablará más adelante) límite para interactuar con átomos de una celda imagen adyacente[[24], [28], [30]].

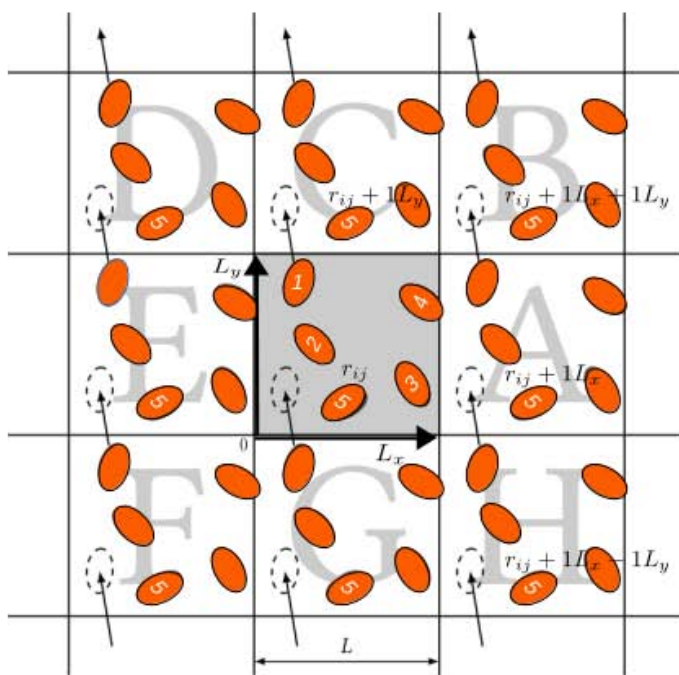


Figura II.1.4. Condiciones periódicas de frontera en dos dimensiones en un sistema con el origen en la celda central (gris) con celdas imagen a sus alrededores[24].

Al incorporar celdas imágenes, el número total de interacciones por pares parece nuevamente tender a infinito. Como sería el caso en la simulación de un sistema macroscópico.

La naturaleza de las fuerzas en sistemas moleculares nos permite limitar el número de cálculos de fuerzas que tenemos para realizar, con el fin de evaluar adecuadamente las fuerzas en las moléculas. Las interacciones de corto alcance decaen rápidamente con la distancia y son tratadas con la convención de mínima imagen, mientras que las interacciones electrostáticas de largo alcance en el sistema, son tratadas de manera separada, por medio de las sumas de Ewald, técnicas de las que se hablará en las siguientes secciones. La introducción de PBC en un sistema periódico infinito facilita la evaluación de estas técnicas[[24], [31]].

II.1.4.1 Radio de corte y convención de mínima imagen

En la sección anterior se discutieron las ventajas e importancia de introducir las PBC en un sistema. Sin embargo, también plantea el siguiente problema. Si se consideran las contribuciones a la energía potencial, derivadas del cálculo de las fuerzas por pares que involucren a un átomo 1, se tendrán $N - 1$ términos en esta suma. En una celda de simulación de N átomos, si solo se consideran las fuerzas de interacciones por pares, se tendrá un total de $N(N - 1)/2$ fuerzas que deben ser evaluadas en cada tiempo de paso. El cálculo de un número infinito de interacciones por pares supone una tarea imposible, en la medida en que disminuye el número de interacciones a considerar, los tiempos de cómputos se reducen[[30], [36]].

Para una función de energía potencial de corto alcance, podemos restringir esta suma considerando el átomo 1 en el centro de una celda del tamaño de la celda principal de simulación como está representada por el cuadrado punteado de la figura II.4.1.1, donde el átomo 1 interactúa con el átomo 2 y las imágenes periódicas (3_D , 4_E y 5_C) más cercanas de los otros $N - 1$ átomos ubicados dentro de la celda principal. Este concepto es conocido como “convención de mínima imagen”.

El radio de corte, r_c , es introducido dentro de la convención de mínima imagen para reducir el número total en los cálculos de la energía potencial y la fuerza. En la figura II.4.1.1 se muestra un círculo punteado, generado por el radio de corte. Para garantizar que el átomo 1 solo interactúe con una de las imágenes de los otros átomos en la simulación, la convención de mínima imagen requiere que el radio de corte sea como máximo $r_c \leq L_{min.}/2$, donde L es la longitud de la dimensión más pequeña de la celda de simulación (ver fig. II.4.1.1)[[27], [30]].

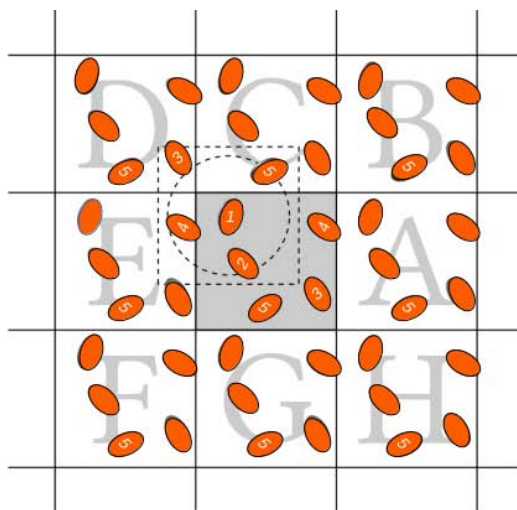


Figura II.4.1.1. Convención de mínima imagen (cuadrado punteado) en un sistema bidimensional. El círculo punteado representa el radio de corte.

Si consideramos la contribución del potencial por pares de corto alcance como: $V(r) = 0$, para $r \geq r_c$, en este caso los átomos que se encuentran en el interior de esta región ($2, 4_E$ y 5_C) contribuyen a la fuerza para 1, mientras que el átomo 3_D no contribuye.

Por ejemplo, en la sección de campos de fuerza se discutieron las interacciones intermoleculares de van der Waals entre dos átomos, que son comúnmente modeladas por el potencial de Lennard Jones.

En la figura II.4.1.2 se muestra el comportamiento del potencial LJ, graficado en unidades reducidas, el potencial LJ y la correspondiente función de fuerza son de corto alcance y decaen rápidamente a distancias grandes.

Es razonable asumir que a una distancia posterior al radio de corte, alrededor de $r_c = 2.5\sigma_{ij}$, el potencial y fuerza se vuelven muy pequeños y no contribuyen apreciablemente.

A $2.5\sigma_{ij}$ el potencial LJ es $U(2.5\sigma_{ij}) = -0.016\sigma_{ij}$ y el potencial se hace más pequeño a distancias más grandes. El despreciar las interacciones entre átomos más allá de esta distancia, tiene un pequeño impacto en la energía potencial total de los átomos, pero reduce drásticamente el número total de interacciones que necesitan ser calculadas.

El corte de interacciones entre dos átomos i y j , separados por una distancia mayor a $2.5\sigma_{ij}$ es implementado reemplazando el potencial de LJ, U_{LJ} , por un potencial truncado, U_{LJ}^T , como:

$$U_{LJ}^T(r_{ij}) = \begin{cases} U_{LJ}(r_{ij}) & \text{si } r_{ij} \leq r_c \\ 0 & \text{si } r_{ij} > r_c \end{cases} \quad (16)$$

De manera que el radio de corte debe ser elegido dentro del intervalo:

$$2.5 \sigma_{ij} \leq r_c \leq L_{min}/2 \quad (17)$$

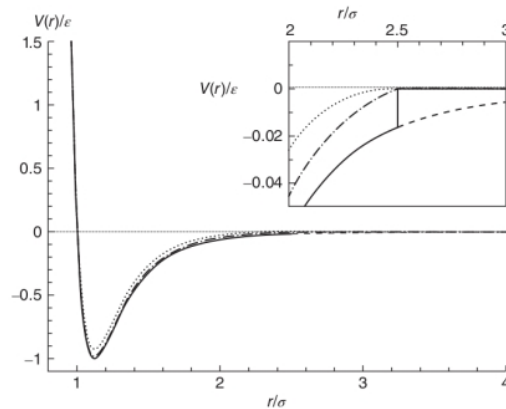


Figura II.4.1.2. Potencial intermolecular de LJ en términos de la distancia reducida r/σ y el potencial reducido U/ϵ . Después de la distancia de corte, $r/\sigma = 2.5$, el valor del potencial puede ser despreciado.

Al introducir el radio de corte, solo energías potenciales y fuerzas entre átomos separados por r_c o menor a él en una celda de simulación de lado L serán consideradas, por lo que el número de vecinos será reducido por un factor de aproximadamente $4\pi r_c^3/3L^3$, que representa un ahorro sustancial.

Sin embargo, para determinar si dos átomos están dentro de la distancia de corte, aún es necesario calcular la distancia entre todos los $N(N - 1)/2$ pares de átomos en la celda imagen mínima[[24], [31], [36]].

II.1.4.2 Sumas de Ewald

Una fuerza de largo alcance es usualmente definida como una en la que la interacción decrece no más rápido que r^{-d} , donde d es la dimensionalidad del sistema. Dentro de esta categoría se encuentran las interacciones carga-carga entre iones que van como $\sim r^{-1}$ y las interacciones dipolo-dipolo que van como $\sim r^{-3}$ [24].

Una fuerza de largo alcance no es truncable por que se extiende a lo largo de todo el espacio. En simulaciones computacionales, su alcance es mayor a la mitad de la longitud del lado con menor dimensión en la celda principal de simulación, lo que representa un problema a considerar. La solución a este problema, consiste en incrementar la longitud de este lado en la celda principal de manera que el apantallamiento por átomos vecinos disminuya el intervalo efectivo de los potenciales. Sin embargo, esta solución no es práctica, ya que implica un esfuerzo computacional que escala como $O(N^2)$ [[24], [31]].

Una eficiente alternativa que permite reducir el costo computacional a $O(N^{3/2})$ es la aproximación de las sumas de Ewald.

Para una mejor comprensión de la utilidad de su implementación, consideremos el fluido iónico de la figura II.1.4.2.1. La convención de mínima imagen correpondería a cortar el potencial en la superficie de un cubo, alrededor de un ión. La naturaleza eléctrica de este cubo será neutra ($\sum_i q_i = 0$), los iones con carga similar tenderán a ocupar las esquinas opuestas del cubo (ver fig II.1.4.2.1). En consecuencia, la imagen periódica de la estructura se superpondrá directamente, dando como resultado una distorsión producida por los iones en lo que debería ser un fluido isotrópico.

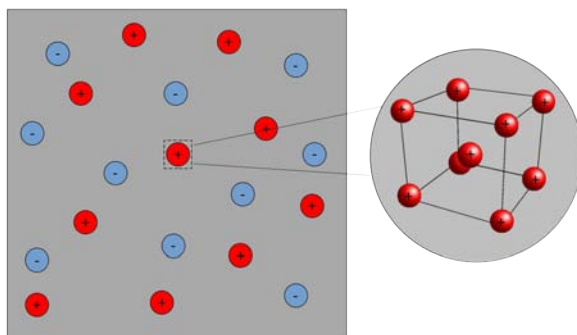


Figura II.1.4.2.1. Ilustración de un fluido iónico con su respectiva convención de mínima imagen.

La alternativa es considerar la interacción del ión con todos los otros iones dentro de la celda principal de simulación y con todos los iones en las celdas imagen periódicas. La energía potencial de un arreglo periódico de N átomos, cada uno con su propia carga, que contiene las contribuciones de todas las réplicas dado que no se impone truncamiento, puede ser escrita en unidades Gaussianas, para hacer la notación más compacta, como:

$$U_{qq} = \frac{1}{2} \sum_{\mathbf{m} \in \mathbb{Z}^3} \sum_{i=1}^N \sum_{j=1}^N \left(\frac{q_i q_j}{|r_{ij} + \mathbf{m}L|} \right) \quad (18)$$

Donde q_i, q_j , son las cargas. La suma sobre $\mathbf{m} = (m_x, m_y, m_z)$, todos los tripletes de enteros, \mathbb{Z}^3 . Para una caja cúbica, $\mathbf{m}L$ representa el centro de cada celda imagen en el arreglo periódico. La prima en la suma indica para $\mathbf{m} = 0$ no se incluye $i = j$ para evitar la auto-interacción.

El término $1/r$ decae muy lentamente y se requiere de muchos valores de \mathbf{m} (de las celdas imagen) en la suma para converger a valores del potencial electrostático en el sitio de la molécula i ; es decir, a potenciales de largo alcance la suma en la expresión (18) es condicionalmente convergente (su resultado depende del orden en el que se agreguen los términos) como las series armónicas¹. Si se consideran las celdas imagen en una secuencia con un orden de proximidad desde la celda central. Por ejemplo, el primer término tiene $|\mathbf{m}| = 0$, esto es, $\mathbf{m} = (0, 0, 0)$; el segundo término, $|\mathbf{m}| = 1$, comprende las 6 celdas laterales centradas en: $(\pm L, 0, 0)$, $(0, \pm L, 0)$, $(0, 0, \pm L)$ y $|\mathbf{m}| = 2$ está asociado con los vectores que pueden ser expresados como: $\mathbf{m} = (\pm 2L, 0, 0)$, $(0, \pm 2L, 0)$, $(0, 0, \pm 2L)$, $(\pm L, 0, \pm L)$, $(\pm L, \pm L, 0)$, $(0, \pm L, \pm L)$ [36]. De forma que conforme añadimos más términos, nuestro sistema infinito es construido en capas esféricas conformadas por las celdas imagen de simulación (ver fig. II.1.4.2.2).

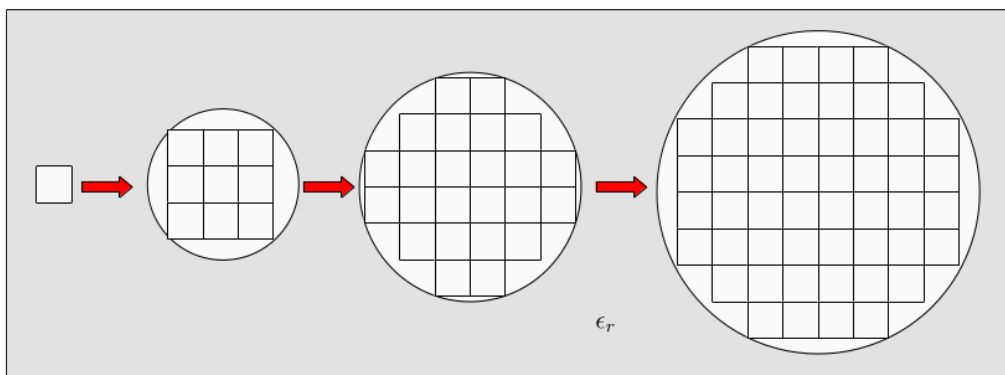


Figura II.1.4.2.2. Ordenamiento de celdas imagen para hacer la suma convergente.

Al utilizar esta aproximación debe especificarse el medio que rodea a la esfera; es decir, la permitividad dieléctrica relativa (ϵ_r).

El potencial en una esfera que rodea un conductor como un metal ($\epsilon_r = \infty$) y en una rodeada por el vacío son diferentes:

¹ Una serie $S = \sum_{n=1}^{\infty} a_n$ es condicionalmente convergente si la suma finita $\sum_n a_n$ converge, pero la serie $\sum_n |a_n|$ diverge. También es bien conocido que la suma de una serie condicionalmente convergente depende del orden de la suma.

$$U_{qq}(\epsilon_r = \infty) = U_{qq}(\epsilon_r = 1) - \frac{2\pi}{3L^3} \left| \sum_i q_i r_i \right|^2 \quad (19)$$

Esta ecuación aplica en el límite de una esfera (de celdas) suficientemente grande. En el vacío, la esfera tiene una capa dipolar en su superficie: el último término en la expresión (19) se cancela. Para la esfera en un conductor no existe tal capa[24].

El método de Ewald transforma la ecuación (18) en una expresión que involucra una suma de dos series absolutas y rápidamente convergentes en el espacio real y recíproco (Apéndice). Esta conversión se logra representando cada carga puntual como la densidad de carga gaussiana en la ecuación (20), representada como una función delta en la figura II.1.4.2.3 por cada carga puntual.

$$\rho_{G_i}(\mathbf{x}) = q_i \left(\frac{\alpha^3}{\pi^{1/2}} \right) \exp(-\alpha^2 x^2) \quad (20)$$

Donde el parámetro α controla el ancho de la distribución y \mathbf{x} es la posición relativa al centro de la distribución.

Esta distribución apantalla la interacción entre cargas vecinas. Las interacciones apantalladas ahora son de corto alcance. También se agrega una distribución de carga del mismo signo que la carga original y la misma forma que la de la distribución $\rho_{G_i}(\mathbf{x})$ (ver fig. II.1.4.2.3). Esta distribución de cancelación reduce el potencial general al debido al conjunto original de cargas. La distribución canceladora se suma en el espacio recíproco; es decir, se suman las transformadas de Fourier (Apéndice) de las distribuciones canceladoras (una por cada carga original) y el total se vuelve a transformar al espacio real.

Por lo que el potencial apantallado total se calcula sumando todas las contribuciones de los átomos en la celda central y todas las de sus imágenes en la red del espacio real de las celdas imagen.

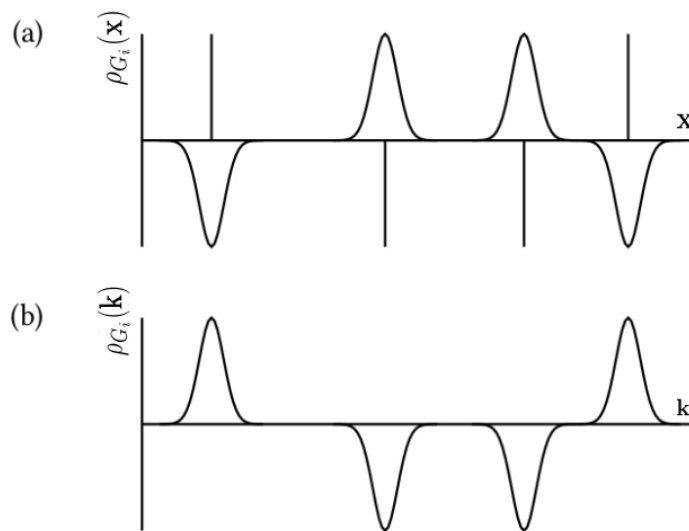


Figura II.1.4.2.3. Distribución de carga en la suma de Ewald. (a) Cargas puntuales originales y distribuciones gaussianas apantalladoras en el espacio real. (b) Distribuciones de cancelación en el espacio de Fourier.

Con el fin de reducir el costo computacional de la interacción de largo alcance de Coulomb, solo se necesita reconocer que cualquier función a caracterizar como de largo alcance en el espacio real se vuelve de corto alcance en el espacio recíproco (Fourier). Más aún, dado que se trata de un sistema periódico, la representación de Fourier es totalmente adecuada.

La suma de Ewald está basada en la reorganización de esta suma de réplicas en capas esféricas concéntricas, asumiendo neutralidad de carga $\sum_j q_j = 0$.

$$\begin{aligned}
U_{qq} = & \sum_{1 \leq i < j \leq N} q_i q_j \left[\sum_m \frac{\text{erfc}(\alpha |r_{ij} + L\mathbf{m}|)}{|r_{ij} + L\mathbf{m}|} \right. \\
& \left. + \frac{1}{\pi L} \sum_{m \neq 0} \left(\frac{1}{m^2} \right) \exp\left(-\frac{\pi^2 m^2}{\alpha^2 L^2} + \frac{2\pi i}{L} \mathbf{m} \cdot \mathbf{r}_{ij}\right) \right] \\
& \frac{1}{2} \left[\sum_{m \neq 0} \left(\frac{\text{erfc}(\alpha L m)}{L m} + \frac{1}{\pi L m^2} \exp\left(-\frac{\pi^2 m^2}{\alpha^2 L^2}\right) \right) - \frac{2\alpha}{\pi^{1/2}} \sum_{j=1}^N q_j^2 \right. \\
& \left. + \frac{2\pi}{3L^3} \left| \sum_{j=1}^N q_j \mathbf{r}_j \right|^2 \right]
\end{aligned} \tag{21}$$

Donde $\text{erfc}(x) = 1 - \text{erf}(x)$ es la función complementaria y $\text{erf}(x) = \frac{2}{\pi} \int_0^x e^{-t^2} dt$ es la función error

$$\text{Si } x \gg 1 \rightarrow \text{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dt$$

La suma reordenada incluye el parámetro α , cuyo valor debe ser determinado para maximizar la precisión numérica; α puede ser elegido de modo que los términos $\exp(-\alpha^2 L^2)$ son despreciables. Por lo que la expresión (21) se convierte en:

$$\begin{aligned}
U_{qq} = & \sum_{1 \leq i < j \leq N} \frac{q_i q_j \text{erfc}(\alpha r_{ij})}{r_{ij}} - \frac{\alpha}{\pi^{1/2}} \sum_{j=1}^N q_j^2 \\
& + \frac{1}{2\pi L} \sum_{m \neq 0} \left[\frac{1}{m^2} \exp\left(-\frac{\pi^2 m^2}{L^2 \alpha^2}\right) \left| \sum_{j=1}^N q_j \exp\left(\frac{2\pi i}{L} \mathbf{m} \cdot \mathbf{r}_j\right) \right|^2 \right]
\end{aligned} \tag{22}$$

Los primeros dos términos en la expresión (22) corresponden a términos en el espacio real y son ahora de corto alcance, por lo que un radio de corte esférico, dentro de un intervalo $r_c < L/2$, puede ser implementado con PBC. Mientras que el tercer término correspondiente a una suma sobre el espacio de Fourier, será también truncable después de un determinado número de términos[[24], [36]].

La suma del momento dipolar en el último término de la expresión (21) ha sido eliminado debido a que el caparazón más externo está rodeado por un medio conductor. Si este término fuera incluido, equivaldría a colocar el sistema en el vacío, como se discutió en la expresión (19).

Un valor típico en el parámetro α es $\alpha = 5/L$. Se elige un radio de corte esférico en esta suma de forma que $m \equiv |\mathbf{m}| \leq m_c$, donde m_c suele ser $m_c = 5$.

Imponiendo un límite en el espacio real de r_c , conlleva a un error del orden de $\exp(-\alpha^2 r_c^2)$; truncando la suma sobre el espacio de Fourier a m_c produce un error del orden de $\exp(-\pi^2 m_c^2 / \alpha^2 L^2)$ [36]. Por lo tanto, para obtener errores de magnitud similar en la contribución en el espacio real y recíproco de U_{qq} , hacemos:

$$m_c = \frac{\alpha^2 r_c L}{\pi} \quad (23)$$

Para el caso en que $r_c = L/2$, si $\alpha = 5/L$ obtenemos: $m_c = \frac{25}{2\pi} \approx 4$

En resumen, al reorganizar las sumas de interacciones sobre imágenes periódicas del sistema, el método de Ewald es capaz de incorporar fronteras periódicas de una manera que evita la truncación explícita.

II.1.5 Minimización de energía

Esta primera etapa de la simulación permite llevar al sistema a una configuración más favorable en donde la distancia los átomos es óptima y los traslapes entre átomos con la misma posición son eliminados, para evitar el cálculo de fuerzas espurias debido a potenciales infinitos. De esta manera, el sistema se encontrará en un menor estado energético. Por medio de algoritmos específicos se garantiza que el sistema se encuentre en un mínimo energético local de la superficie de potencial.

La minimización de una estructura ocurre en dos pasos. Primero, debe definirse la ecuación que describe la energía del sistema en función de sus coordenadas (campo de fuerzas) y evaluarse para una determinada configuración. Se pueden construir funciones que también incluyan restricciones externas para sesgar la minimización.

A continuación, la conformación se ajusta para reducir el valor de la función de interés. Se puede encontrar un mínimo después de un ajuste o puede requerir miles de iteraciones, según la naturaleza del algoritmo, la forma de la función del campo de fuerzas y el tamaño del sistema. Por lo tanto, la eficiencia de la minimización se mide tanto por el tiempo para evaluar la función, como por el número de iteraciones de ajustes estructurales de la minimización[37].

De entre los algoritmos disponibles para la minimización de energía, por ser de interés particular, se ilustrará la implementación del algoritmo “Steepest Descent”.

En “steepest descent”, la dirección de búsqueda se toma simplemente como la pendiente. Después de cada dirección de búsqueda, la dirección anterior se reemplaza con el gradiente en el nuevo punto y se repite la búsqueda. La figura II.1.5.1.a muestra la ruta de minimización seguida por un enfoque de descenso más pronunciado en una función cuadrática. Cada búsqueda produce una nueva dirección que es perpendicular al gradiente anterior. Sin embargo, usar el gradiente actual como la siguiente dirección conduce a oscilaciones en el camino hacia el mínimo. Aunque cada dirección sucesiva es ortogonal a la dirección anterior, el camino al mínimo no es particularmente eficiente.

Para mejorar su eficiencia se elimina la búsqueda y solo se actualiza la posición cada vez que el punto de prueba a lo largo del gradiente tiene una energía más baja. Con ello, la minimización comienza desde el mismo punto que en la figura II.1.5.1.a, pero cada búsqueda usa como máximo dos evaluaciones de función, si el punto de prueba tiene una energía más alta, el tamaño de paso se ajusta hacia abajo y se genera un nuevo punto de prueba (ver fig. II.1.5.1.b).

En este caso los pasos son más erráticos, pero se llega al mínimo en aproximadamente en 12 iteraciones. La enorme ventaja que esto presenta es el evitar búsquedas exhaustivas y el número total de evaluaciones de funciones es solo el 10-20% del utilizado por el método inicial.

Steepest descent, es el método más robusto para generar una estructura de menor energía sin importar cuál sea la función o dónde comience. Por lo tanto, este algoritmo es comunmente utilizado en un inicio cuando las configuraciones están lejos del mínimo, en otras palabras, cuando las pendientes son grandes[37].

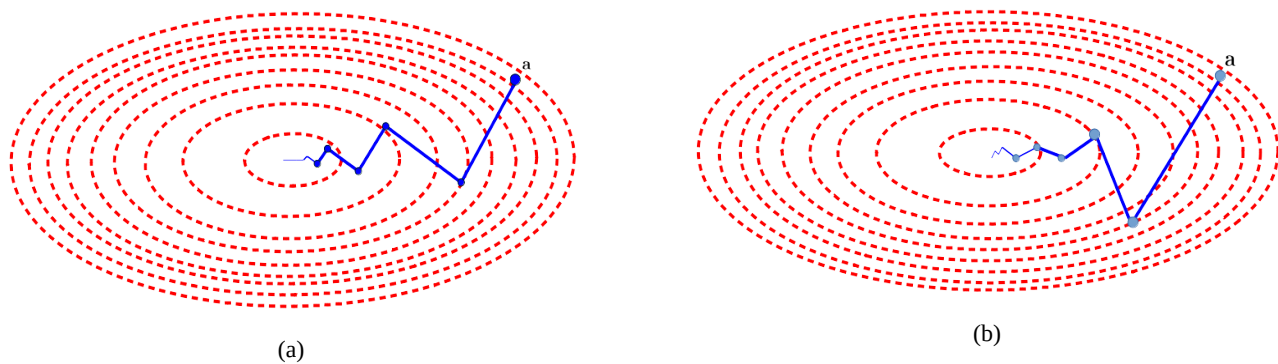


Figura II.1.5.1. Trayectoria de minimización a través del algoritmo “steepest descent”. (a) Utilizando solamente búsqueda de dirección con el gradiente, iniciando desde el punto a, requiere de aprox. 8 evaluaciones de la función por cada búsqueda. (b) Sin utilizar búsquedas de dirección, la minimización total es 5 veces más rápida, en promedio cada iteración requiere de 1.2 evaluaciones de la función.

En GROMACS, la paquetería de DM utilizada de la que se hablará más adelante, permite implementar este algoritmo. Para ello, se define el vector \mathbf{r} como el vector de todas las $3N$ coordenadas. Inicialmente, debe ser dado un desplazamiento máximo h_0 . De acuerdo con el manual, un valor comunmente utilizado es $h_0 = 0.01 \text{ nm}$.

Primero se calcula la energía potencial y las fuerzas. Donde las nuevas posiciones son calculadas por:

$$\mathbf{r}_{n+1} = \mathbf{r}_n + \frac{\mathbf{F}_n}{\max(|\mathbf{F}_n|)} h_n \quad (24)$$

Donde h_n es el desplazamiento máximo en el paso n y $\mathbf{F}_n = -\nabla U_n$ es la fuerza o el negativo del gradiente del potencial U_n . La notación $\max(|\mathbf{F}_n|)$ representa el valor escalar más grande de la fuerza en cualquier átomo. Posteriormente, las fuerzas y energía son calculadas nuevamente, como:

Si $U_{n+1} \leq U_n$, las nuevas posiciones son aceptadas y $h_{n+1} = 1.2h_n$

Si $U_{n+1} \geq U_n$, las nuevas posiciones son rechazadas y $h_n = 0.2h_n$

El algoritmo se detiene cuando el número especificado de evaluaciones de la fuerza ha sido completado (100), o cuando el máximo de los valores absolutos de los componentes de la fuerza es menor al valor especificado ϵ [32].

II.1.6 Termodinámica y Física Estadística

La termodinámica es una teoría autoconsistente y una que no hace alusión a la estructura microscópica de la materia. Un sistema termodinámico es un sistema macroscópico y siempre se divide en el sistema de estudio y sus alrededores.

En termodinámica, los parámetros que definen un estado termodinámico, como la presión P , el volumen V , la temperatura T o la masa total M , son cantidades medibles que se pueden obtener experimentalmente. Un estado termodinámico se especifica proporcionando los valores de todos los parámetros termodinámicos necesarios para que la descripción de un sistema sea completa.

En termodinámica, la ecuación de estado de un sistema es una relación entre parámetros termodinámicos que describe su variación de un estado de equilibrio a otro.

Por lo que, si P , V y T son los parámetros termodinámicos fundamentales de un sistema, la ecuación de estado tomará la forma general:

$$f(P, V, T) = 0 \quad (25)$$

Una consecuencia de esta ecuación es que solo existen dos parámetros termodinámicos independientes en un estado de equilibrio. Además, una función de estado bajo una transformación termodinámica depende solamente del estado inicial y final de la transformación y no de una trayectoria particular entre ambos estados.

Por otro lado, las leyes de la mecánica clásica establecen que la evolución de un sistema puede ser totalmente determinada (en ambas direcciones) si se conocen las posiciones y velocidades de todas las partículas en un solo instante de tiempo. Por lo que la mecánica clásica comprende toda la información necesaria para predecir las propiedades de un sistema en cualquier instante[31].

La dinámica molecular (MD), permite el tratamiento de sistemas en macroescala, en los que se necesita de soluciones numéricas y las trayectorias son generadas para extraer observables macroscópicos.

También proporciona un amplio detalle en la visualización de movimientos de átomos individuales en el sistema, por lo que provee una especie de ventana al mundo microscópico.

Sin embargo; para llegar a ello, primero debió solucionarse la inconsistencia entre la termodinámica y las leyes microscópicas de movimiento.

Primero, los sistemas macroscópicos tienen una enorme cantidad de grados de libertad y los sistemas del mundo real se caracterizan por interacciones no lineales. El carácter no lineal de las fuerzas en estos sistemas vuelve inviable la solución analítica de las ecuaciones de movimiento. Alternativamente, se puede proponer una solución numérica para las ecuaciones de Newton en una computadora; sin embargo, la memoria requerida para guardar solamente un punto del espacio fase para un sistema del orden del número de Avogadro (6×10^{23}), excede la capacidad actual de poder de cómputo, así como la del futuro cercano.

Por lo que, mientras la mecánica clásica encapsula toda la información necesaria para predecir las propiedades de un sistema, el problema de extraer esa información es aparentemente irresoluble[31].

Aunado al problema del tamaño de sistemas macroscópicos, la segunda ley de la termodinámica establece una flecha unidireccional hacia adelante en la dirección del tiempo, comunmente conocida como la dirección en la que la entropía se incrementa.

Esta unicidad en la dirección del tiempo está en desacuerdo con las leyes mecánicas microscópicas, que debido a su naturaleza matemática, son reversibles en el tiempo. Esta paradójica situación es conocida como la paradoja de Loschmidt's[31].

El desarrollo de la física estadística fue justamente el que dio pie al puente entre el mundo clásico microscópico, descrito por las leyes de movimiento de Newton y los observables macroscópicos. La contribución más notable de la Física Estadística es el haber permitido comprender que las propiedades observables macroscópicas de un sistema no dependen significativamente del movimiento detallado de cada partícula en un sistema macroscópico, sino de los promedios que no involucran los detalles microscópicos. Por lo tanto, al aplicar las leyes mecánicas microscópicas de manera estadística, a partir del concepto de ensamble, se puede proporcionar un vínculo entre las teorías microscópicas y macroscópicas de la materia.

Este concepto también conduce a procedimientos para calcular muchos otros observables macroscópicos, incluyendo propiedades estructurales, dinámicas y termodinámicas.

II.1.6.1 Ensembles

La idea de que los observables macroscópicos de un sistema no son sensibles a detalles microscópicos específicos es la base del concepto de ensemble, originalmente introducido por Gibbs. Formalmente, un ensemble es una colección de sistemas que satisfacen un estado termodinámico particular o tienen un conjunto de propiedades macroscópicas idénticas, de modo que cada sistema se encuentra en un estado microscópico único, conocido como microestado y definido por un punto en el espacio fase. El espacio fase consta de todos los posibles microestados disponibles para un sistema de N partículas. El enfoque del espacio fase considera distribuir la energía entre los estados de los N átomos de todo el sistema. Los puntos se distribuyen según una distribución de probabilidad $\rho(\Gamma)$ y permiten describir su evolución en el tiempo.

La derivación de funciones de distribución de probabilidad para sistemas que interactúan se obtuvo hasta después del descubrimiento de los principios de la mecánica cuántica. La introducción de estados discretos en la teoría de la mecánica cuántica condujo a expresiones para las distribuciones de probabilidad en sistemas que interactúan. Expresiones que en límite clásico, se tradujeron en expresiones de la física estadística que son de utilidad para la MD.

Existen dos enfoques para calcular la distribución más probable de microestados en el espacio fase. El primero, desarrollado inicialmente por Boltzmann, se denomina representación en el espacio- μ [[38], [39]]. Este enfoque solo es válido para sistemas de átomos no interactuantes.

El segundo, denominado espacio- Γ (espacio fase), basado en un sistema de N átomos, donde las posiciones y momentos pueden considerarse como coordenadas en un espacio multidimensional, este espacio tiene $6N$ dimensiones[[31], [24], [30]].

La energía suele ser obtenida a partir de la formulación Hamiltoniana de la mecánica clásica debido a su compatibilidad con el espacio fase al encontrarse descrita explícitamente en términos de las posiciones y momentos. Adicionalmente, la formulación Hamiltoniana es totalmente equivalente a la formulación Newtoniana de la mecánica clásica, resultando de especial utilidad en la mecánica estadística.

Un ensemble contiene solo aquellos microestados que son consistentes con un conjunto dado de observables macroscópicos. En consecuencia, los microestados de un ensemble son un subconjunto estricto de todos los posibles puntos del espacio de fase o están agrupados más densamente en ciertas regiones del espacio de fase y menos densamente en otras[24].

En la teoría de ensembles clásica, cada observable macroscópico de un sistema está directamente conectado a una función microscópica de las coordenadas y momentos del sistema. Consideremos alguna propiedad A_{obs} (podría ser la energía potencial) como una función $A(\Gamma)$. Donde denotamos Γ como un punto particular en el espacio de fase, que se simula bajo las condiciones de un ensemble (como volumen constante y temperatura constante) especificadas por X . En cada momento de la simulación, el valor de esta cantidad es $A(\Gamma; X)$, definida como:

$$A_{obs} = \frac{1}{Z} \sum_{\lambda=1}^Z A(X_{\lambda}) \equiv \int d\Gamma \rho(\Gamma, t) A(X) \equiv \langle A(X) \rangle \quad (26)$$

Donde Z es el número de miembros en el ensamble y la equivalencia intermedia es una generalización de la suma discreta en el espacio fase.

Por lo tanto, necesitamos describir cuantitativamente cómo se distribuyen los sistemas en un ensamble en el espacio de fase en cualquier momento. En general, una función de distribución de ensamble (función de distribución de espacio fase) está determinada por las variables termodinámicas elegidas (NPT, NVT, etc.), denotadas por la notación: ρ_{NPT} , ρ_{NVT} , o en general, ρ_{ens} [24].

La evolución temporal de la densidad del espacio de fase $\rho_{ens}(\Gamma)$ cambiará sin alterar (crear o destruir) ningún sistema; es decir, los promedios de ensambles se pueden realizar en cualquier momento. Estos ensambles se conocen como ensambles de equilibrio. Esta afirmación está sostenida por el teorema de Liouville, que es esencialmente una ley de conservación para la densidad de probabilidad, establece que la derivada total con respecto al tiempo, puede expresarse como:

$$\frac{d}{dt}\rho(\Gamma(t), t) = \frac{\partial}{\partial t}\rho(\Gamma(t), t) + \{\rho(\Gamma(t), t), H(\Gamma(t), t)\} = 0 \quad (27)$$

La ecuación (27) asegura que podemos realizar promedios sobre el ensamble en cualquier momento porque se conserva la fracción de miembros del conjunto en el tiempo como se prueba en la referencia[31].

Si A es un observable en equilibrio, el ensamble promedio en la expresión (26), el resultado será independiente del tiempo, en cuyo caso, la distribución del ensamble de un sistema en equilibrio termodinámico no es temporalmente dependiente; es decir, $\partial\rho/\partial t = 0$.

En este caso ($\partial\rho/\partial t = 0$), la ecuación de Liouville en la expresión (27) se reduce a:

$$\{\rho(\Gamma(t)), H(\Gamma(t))\} = 0 \quad (28)$$

La solución general a la ecuación (28) es cualquier función del Hamiltoniano, $H(\Gamma(t))$:

$$\rho(\Gamma) = \frac{1}{Z} P(\Gamma) \quad (29)$$

Donde $P(\Gamma)$ es una función de probabilidad en el espacio fase y Z es la denominada función de partición. Definida como:

$$Z = \int d\Gamma P(H(\Gamma)) \quad (30)$$

Esta función es una de las cantidades centrales en la mecánica estadística. Es una medida del número de estados microscópicos accesibles en el espacio fase dentro de un ensamble determinado. Cada ensamble tiene una función de partición particular que depende de las variables de control que definen ese ensamble[31].

Otros observables en equilibrio pueden ser obtenidos de acuerdo con:

$$A_{obs} = \langle A \rangle = \frac{1}{Z} \int d\Gamma A(\Gamma) P(H(\Gamma)) \quad (31)$$

Cada sistema evoluciona bajo leyes microscópicas de movimiento a partir de una condición inicial diferente, de modo que en cualquier momento, cada sistema tiene un estado microscópico único definido en el espacio fase[30]. Cada variable termodinámica se ha incluido en las ecuaciones de movimiento para las variables mecánicas. Las variables termodinámicas se introducen de forma que garanticen que el sistema satisfaga la distribución de probabilidad.

En una simulación de MD o en un sistema observado experimentalmente, las mediciones se realizan en un solo sistema durante un periodo de tiempo.

Las predicciones de un ensamble permiten conectar las probabilidades estadísticas clásicas con las observaciones experimentales o de MD si se satisface la llamada hipótesis ergódica (del griego: ergon [trabajo] + hodos [camino]). La hipótesis ergódica establece que el promedio de un ensamble de una cantidad ($\langle A(X) \rangle$) es igual al promedio temporal de esa cantidad ($\langle A(X) \rangle_t$) en un solo sistema si la observación se realiza durante un intervalo de tiempo lo suficientemente largo[30].

Matemáticamente, la hipótesis ergódica presenta las siguientes equivalencias:

$$A_{obs} \equiv \langle A(X) \rangle = \langle A(\Gamma(t); X) \rangle_t = \langle A(X) \rangle_t \quad (32)$$

$$A_{obs} = \int A(\Gamma) P(\Gamma; X) d\Gamma = \lim_{t_{obs} \rightarrow \infty} \frac{1}{t_{obs}} \int_0^{t_{obs}} A(\Gamma(t); X) dt = \frac{1}{\tau_{obs}} \sum_{\tau=1}^{\tau_{obs}} A(\Gamma(t); X)$$

En otras palabras, la evolución temporal de cualquier sistema individual muestrea la hipersuperficie constante del ensamble de forma suficientemente representativa para que el promedio temporal y el promedio del ensamble representados por la integral de espacio de fase en la expresión (32) sean idénticos [[30], [40], [41]]. Por lo que, si los tiempos de simulación de la dinámica molecular son demasiado cortos para el fenómeno de estudio, el muestreo de la hipersuperficie del espacio de fase no será suficientemente representativo del comportamiento del sistema y el resultado de la simulación podría no corresponder a las observaciones experimentales[30].

Al elegir un ensamble de interés, los observables macroscópicos se calculan realizando promedios sobre los sistemas del ensamble. Los ensambles se pueden definir para una amplia variedad de situaciones termodinámicas.

El ejemplo más sencillo es un sistema aislado de sus alrededores. De igual forma, los ensambles también describen sistemas en contacto con baños de calor, sistemas en contacto con depósitos de

partículas, sistemas acoplados a mecanismos de control de presión como pistones mecánicos y otras combinaciones.

Dichos ensambles son útiles para determinar propiedades estáticas como temperatura, presión, energía libre, estructura promedio, etc[31].

II.1.6.1.1 Ensamble NVT

El ensamble canónico o ensamble isocórico-isotérmico tiene como variables de control, el número de partículas N constante, volumen V constante y temperatura T constante, que caracterizan a un sistema en contacto térmico con una fuente de calor infinito. Cuando un sistema entra en contacto con una fuente de calor infinita, su energía fluctuará de manera que la temperatura se mantendrá constante.

En este ensamble, el Hamiltoniano no se conserva debido a que un sistema intercambia energía con sus alrededores, comunmente llamado, baño térmico. En otras palabras, el intercambio de energía cambia el número de microestados accesibles. En consecuencia, obedece la distribución de Boltzmann como veremos en la siguiente sección.

En termodinámica, el potencial termodinámico que a partir de la ecuación fundamental y las transformaciones de Legendre describe de manera natural un sistema en función de estas variables de control es la energía libre de Helmholtz ($A(N, V, T)$), definida como:

$$\begin{aligned} A(N, V, T) &= U(N, V, S(T)) - TS(N, V, T) \\ &= U(N, V, S(N, V, T)) - \frac{\partial U}{\partial S} S(N, V, T) \end{aligned} \quad (33)$$

Consideremos un proceso en el que N , V y T cambien en una cantidad infinitesimal definida por: dN , dV , dT , respectivamente. Dado que A es una función punto¹ el cambio en la energía libre de Helmholtz puede expresarse como una ecuación diferencial exacta, dA , definida como:

$$dA = \left(\frac{\partial A}{\partial N} \right)_{V, T} dN + \left(\frac{\partial A}{\partial V} \right)_{N, T} dV + \left(\frac{\partial A}{\partial T} \right)_{N, V} dT \quad (34)$$

¹ función en la que el valor que toma un punto dado es independiente de la trayectoria para alcanzar dicho punto. Aquellas funciones que no son de punto, diferenciales inexactas, no son variables termodinámicas[42].

Como:

$$\begin{aligned}
 A &= U - TS \quad y \quad dU = TdS - PdV + \mu dN \\
 dA &= dU - d(TS) = dU - SdT - TdS \\
 &= TdS - PdV + \mu dN - SdT - TdS \\
 &= -PdV + \mu dN - SdT
 \end{aligned} \tag{35}$$

Por lo que al comparar las ecuaciones (34) y (35), las variables termodinámicas obtenidas de las derivadas parciales de A, son:

$$\mu = \left(\frac{\partial A}{\partial N} \right)_{V,T}, \quad P = - \left(\frac{\partial A}{\partial V} \right)_{N,T}, \quad S = - \left(\frac{\partial A}{\partial T} \right)_{N,V} \tag{36}$$

Estas relaciones definen la termodinámica básica del ensamble canónico.

Formalmente una diferencial es exacta si su integral de línea a lo largo de una curva en la región donde se encuentra definida, sea independiente de la elección particular de dicha curva[43].

De acuerdo con su significado físico, una transformación termodinámica de un sistema desde un estado inicial hasta un estado final, se lleva a cabo a través de una trayectoria reversible. Por lo que el trabajo necesario para efectuar esta transformación es igual al cambio en la energía libre de Helmholtz, ΔA .

Es importante notar que aunque la termodinámica no brinda información sobre la escala de tiempo en la que el proceso ocurre, el problema queda resuelto con la introducción de ensambles como se discutió en la sección pasada. Con base en el concepto de energía libre, la termodinámica permite determinar si el proceso es favorable al haber un descenso de la energía libre o desfavorable, indicado por un incremento en la energía libre[31].

Como se mencionó en la sección anterior, cada ensamble tiene una función de partición particular definida en función de sus variables de control. En el ensamble canónico, el número total de microestados accesibles, definidos por su función de partición ($Q(N, V, T)$) tiene la forma:

$$Q(N, V, T) = \frac{1}{N! h^{3N}} \int d\Gamma e^{-\beta H(\Gamma)} \tag{37}$$

Donde $\beta = 1/k_B T$ y h es la constante de Planck, donde $h^{3N} = (\Delta x)^{3N} (\Delta p)^{3N}$.

Para vincular las relaciones que definen la termodinámica básica del ensamble canónico en la expresión (36) y los estados microscópicos de un sistema en términos del Hamiltoniano ($H(\Gamma)$), se emplea la función de partición de este ensamble en la ecuación (37) a través de la ecuación (33), esta dado por:

$$A(N, V, T) = -k_B T \ln Q(N, V, T) \quad (38)$$

Lo que puede comprobarse al ver que la expresión (38) es solución de la ecuación $A = U - TS$.

De las ecuaciones en (36), los observables termodinámicos están dados en términos de la función de partición como:

$$\begin{aligned} \mu &= -k_B T \left(\frac{\partial \ln Q}{\partial N} \right)_{V, T} \\ P &= k_B T \left(\frac{\partial \ln Q}{\partial V} \right)_{N, T} \\ S &= k_B \ln Q + k_B T \left(\frac{\partial \ln Q}{\partial T} \right)_{N, V} \\ U &= - \left(\frac{\partial}{\partial \beta} \ln Q \right)_{N, V} \end{aligned} \quad (39)$$

Las variables de control en el ensamble canónico son particularmente útiles en la descripción de fenómenos de superficie como la adsorción, motivo por el cual este fue el ensamble elegido en el presente trabajo para la etapa de equilibración durante el protocolo de MD.

II.1.7 Equilibración en el ensamble NVT

Antes de comenzar una simulación, necesitamos no solo las posiciones sino también las velocidades. En la sección II.1.5, discutimos el papel de la etapa de minimización de energía para proveer solamente un conjunto de posiciones energéticamente favorables con respecto al conjunto de posiciones iniciales, como punto de partida. Por lo que las velocidades iniciales deben ser asignadas, usualmente esto es hecho al asignar velocidades iniciales a átomos de forma que se llega a la temperatura deseada a través de una adecuada distribución como punto de inicio, dicha distribución es la distribución de Maxwell-Boltzmann (MB) definida como:

$$P(p) = \left(\frac{\beta}{2\pi m} \right)^{3/2} \exp(-\beta p^2 / 2m) \quad (40)$$

El proceso de asignación de velocidades no es relevante dado que la distribución MB surge naturalmente de las ecuaciones de movimiento, una consecuencia del ensamble canónico que discutimos anteriormente. La razón de ello es que el momento del centro de masa de la caja de

simulación es conservado por las ecuaciones de movimiento de Newton. Para evitar un desplazamiento de la caja de simulación de su posición inicial, se suele asignar un valor de cero al centro de masa de la caja simulación, removiendo la velocidad del centro de masa de todas las partículas posterior a la asignación aleatoria de las velocidades.

Adicionalmente a las velocidades asignadas de acuerdo con la distribución MB, en el ensamble canónico se necesita implementar un mecanismo que permita añadir o remover calor del sistema que se aproxime a una adecuada equipartición de las energías cinética y potencial. Dicho mecanismo se conoce como termostato[27].

De entre las distintas categorías en que se han implementado los algoritmos de termostato desarrollados, los deterministas y estocásticos son de los más utilizados. Muchos de los termostatos globales se pueden convertir en variantes locales acoplando termostatos separados a cada partícula en el sistema en lugar de tener un solo termostato para todo el sistema. También hay varios métodos empleados por los algoritmos de los termostatos para controlar la temperatura. Algunos termostatos funcionan cambiando la escala de velocidades fuera de las ecuaciones de movimiento de la MD. Otros, incluyen colisiones estocásticas entre el sistema y un baño térmico implícito de partículas, o explícitamente incluyen grados de libertad adicionales en las ecuaciones de movimiento que tienen el efecto de un baño térmico externo[27].

Un ejemplo de este último caso es el termostato Langevin, que simula el acoplamiento del sistema de interés a un baño térmico. Aquí, las ecuaciones de movimiento se modifican de dos maneras: primero, se introduce un término de "fuerza aleatoria", este es el elemento estocástico del algoritmo. En segundo lugar, se agrega una "fuerza de fricción" determinista, proporcional a las velocidades de las partículas. La fuerza de estos términos y la temperatura prescrita están conectadas por el teorema de fluctuación-disipación¹.

Este termostato es particularmente efectivo para equilibrar grados de libertad fijos, dado que cada grado de libertad es acoplado a un baño térmico de manera independiente, debido a que permite reemplazar los grados de libertad del baño térmico con una memoria integral y el término de fuerza estocástica mencionado anteriormente, dejando solamente las coordenadas de interés para ser tratadas de forma explícita, por lo que es menos demandante computacionalmente y accesible al estudio de fenómenos que requieren mayores tiempos de escala[[24], [31]].

En el ensamble (N, V, T) la energía fluctúa de forma que la temperatura se mantiene constante. La magnitud de estas fluctuaciones es un punto clave durante la etapa de equilibración. Por ello, la fase de equilibración debe ser monitoreada evaluando la evolución de la temperatura del sistema, así como de sus energías cinética y potencial para garantizar que alcancen un valor promedio constante durante la simulación.

Una vez que las energías cinética y potencial fluctúan alrededor de valores constantes y alguna otra propiedad clave se mantiene estable en el tiempo, se dice que la etapa de equilibración ha llegado a su fin. En general, si una de las propiedades aún presenta fluctuaciones o cambios abruptos durante la simulación se considera como una señal de que la equilibración no se ha completado aún[27].

¹ El teorema de fluctuación-disipación es un resultado general de la física estadística que cuantifica la relación entre las fluctuaciones en un sistema en equilibrio termodinámico y la respuesta del sistema a la aplicación de perturbaciones.

El principio que guía el acoplamiento de variables externas a variables microscópicas se basa en una adecuada distribución de las variables microscópicas de acuerdo con la función del ensamble utilizado posterior a la equilibración. Por ejemplo, después de la operación de un termostato simulado en un sistema con volumen constante, al llegar al equilibrio, las variables microscópicas tienen una distribución que corresponde a la función de distribución del ensamble canónico[30].

Durante la equilibración, la temperatura es acoplada a las ecuaciones de movimiento por medio de las variables mecánicas microscópicas v_i . Donde la temperatura se relaciona con la energía cinética promedio de los átomos (velocidad promedio de los átomos) en la simulación por medio del teorema de equipartición:

$$\frac{3}{2} N k_B T = \left\langle \frac{1}{2} \sum_{i=1}^N m_i v_i^2 \right\rangle \quad (41)$$

II.1.8 Producción de datos

La producción de datos es la etapa posterior a la etapa de equilibración en la MD, una vez que el sistema se ha llevado a un estado de equilibrio definido por el ensamble utilizado en la etapa anterior (temperatura, presión, etc.), comienza la producción de datos. Durante esta etapa se realiza la captura de datos para su posterior análisis. Dependiendo del sistema y las condiciones elegidas el tiempo apropiado para culminar esta etapa estará ligado a un adecuado muestreo de las configuraciones más probables en el ensamble de equilibrio de interés.

De acuerdo con la literatura, se sugiere que nunca se realice la producción de datos inmediatamente después de realizar un cambio en las condiciones del sistema fuera de equilibrio[27]. Si se comienza desde un sistema con una configuración menos estable, será necesario emplear más tiempo en la etapa de equilibración como se mencionó en la sección de equilibración, para alcanzar las configuraciones espaciales más relevantes.

Qué información capturar y con qué frecuencia debe hacerse es otro punto clave en esta etapa. Una captura muy frecuente de datos podría llevar a problemas de almacenamiento al saturar la memoria. Para diversas aplicaciones, lo ideal es capturar las energías y trayectorias cada cierto tiempo en el intervalo de la producción, de forma que se pueda ir monitoreando su evolución. Por ejemplo, en el caso de un análisis estructural, esta estrategia es utilizada junto con un análisis de energías, ya que las energías típicamente utilizan menos memoria y pueden ser capturadas con mayor frecuencia[27].

II.2 Programas utilizados

Los programas empleados en el presente trabajo han sido utilizados innumerables veces en el pasado por otros investigadores en problemas relacionados (molécula-superficie)[44], así como su precisión ha sido validada a través de su comparación con trabajos experimentales.

A continuación se hace mención de las características generales de GROMACS, VMD y CHARMM-GUI.

II.2.1 GROMACS

GROMACS es una versátil paquetería para realizar cálculos y simulaciones de dinámica molecular a partir de las ecuaciones de movimiento de Newton en sistemas con cientos y hasta miles de partículas.

Es una paquetería de licencia libre, disponible bajo la GNU LGPL, por sus siglas en inglés. Está esencialmente diseñada para biomoléculas como proteínas, lípidos y ácidos nucleicos que tienen muchas interacciones de enlace, que usualmente dominan las simulaciones. De acuerdo con su página, distintos grupos científicos utilizan esta paquetería incluso para investigar en sistemas no biológicos y polímeros.

GROMACS es compatible con todos los algoritmos que usualmente son esperados en la implementación de algoritmos de MD; sin embargo, también cuenta con características que hacen de esta paquetería una herramienta sobresaliente en su campo[32].

Como dato curioso, aunque algunas fuentes indican que GROMACS es un acrónimo del lugar donde fue creado (GROningen Machine for Chemical Simulations), la página oficial indica que GROMACS es un nombre propio, por lo que no es un acrónimo de nada.

II.2.2 CHARMM-GUI Input Generator

CHARMM-GUI es una interfaz gráfica de usuarios para CHARMM desarrollada por el equipo de investigación del Dr. Im en la Universidad de Lehigh, Bethlehem. El proyecto fue iniciado por dos asistentes de investigación de pregrado, Vidyashankara G. Iyer y Sunhwan Jo[45]. Dentro de las diferentes modalidades con las que cuenta, CHARMM-GUI Input Generator facilita la construcción de sistemas para correr cálculos de MD así como de superficies, principal motivo por el que este software fue utilizado.

Un dato interesante de esta herramienta es que la licencia era de paga previo al año 2013, fecha en la que el Dr. Martin Karplus ganó el premio Nobel de Química, razón por la que el Dr. Im y su equipo liberaron la licencia para que todo aquél con intereses académicos pudiera utilizarla[46].

II.2.3 VMD

VMD es una herramienta diseñada para modelación, visualización y análisis de sistemas biológicos como proteínas, ácidos nucleicos, complejos de bicapas lipídicas, entre otros. VMD provee una amplia gama de métodos para renderizar y colorear moléculas en su representación: puntos simples y líneas, esferas CPK, VDW, etc. VMD tiene la capacidad de ser utilizado para animar y analizar trayectorias de simulaciones de MD[47].

II.3 RMSD

El parámetro de desviación cuadrática media (RMSD, por sus siglas en inglés), es una cantidad que permite caracterizar los cambios estructurales en la evolución temporal de una configuración inicial

tomada como referencia. Por ejemplo, el cambio de la estructura de una proteína a un tiempo t_2 respecto a su configuración de referencia al tiempo t_1 .

Usualmente se calcula respecto a la cadena principal de la proteína (conocida como “backbone” en inglés) conformada por los átomos C_α .

En GROMACS, este parámetro se define como:

$$RMSD(t_1, t_2) = \left[\frac{1}{M} \sum_{i=1}^N m_i \| \mathbf{r}_i(t_1) - \mathbf{r}_i(t_2) \|^2 \right]^{1/2} \quad (42)$$

Donde $M = \sum m_i$ y $\mathbf{r}_i(t)$ es la posición del átomo i al tiempo t .

II.4 Radio de giro

El radio de giro es un parámetro que permite medir el grado de compactación presente en una estructura[32]. Por ejemplo, en el caso de proteínas permite cuantificar el plegamiento de su estructura terciaria durante la simulación. En GROMACS el radio de giro se define como:

$$R_g = \left(\frac{\sum_i \| \mathbf{r}_i \|^2 m_i}{\sum_i m_i} \right)^{1/2} \quad (43)$$

Donde \mathbf{r}_i es la posición del átomo i y m_i es la masa del átomo i .

II.5 Perfil de densidad parcial

El perfil de densidad parcial permite conocer la composición de la caja de simulación, es decir, la concentración de las especies involucradas. El perfil de densidad parcial divide la caja de simulación del sistema en rebanadas de igual tamaño a lo largo de una dirección del marco de referencia. De manera que permite conocer la densidad de los componentes en el sistema en cada rebanada de la caja de simulación.

En GROMACS, el perfil de densidad parcial se define como:

$$\rho_n(z) = \frac{1}{L_x L_y \Delta z} \sum_{i=1}^N n_i(z) \quad (44)$$

Donde Δz es el ancho de la rebanada de la caja de simulación, $n_i(z)$ es el número de partículas presentes en la rebanada n , L_x y L_y son las dimensiones en los ejes “x” y “y”, respectivamente.

II.6 Función de distribución radial

La función de distribución radial (RDF, por sus siglas en inglés) permite caracterizar las correlaciones espaciales entre pares de moléculas en simulaciones de sistemas equilibrados. Dichas correlaciones pueden ser medidas experimentalmente para proveer detalles acerca del promedio de la organización de las moléculas. La RDF describe la organización local promediada esféricamente alrededor de cualquier átomo.

En GROMACS, la RDF se define como:

$$g_{AB}(r) = \frac{\langle \rho_B(r) \rangle}{\langle \rho_B(r) \rangle_{local}} = \frac{1}{\langle \rho_B(r) \rangle_{local}} \frac{1}{N_A} \sum_i^{N_A} \sum_j^{N_B} \frac{\delta r_{ij} - r}{V(r)} \quad (45)$$

Donde $\langle \rho_B(r) \rangle$ es la densidad de las partículas B a una distancia r alrededor de partículas A y $\langle \rho_B(r) \rangle_{local}$ es la densidad de partículas B promediadas sobre todos los caparazones esféricos alrededor de partículas A con radio r_{max} , usualmente $r_{max} = L/2$. El término de la suma indica el número de pares ij que están separados una distancia r . Estos pares son determinados al considerar caparazones esféricos con volumen $V(r)$ y normalizando con la densidad promedio local $\langle \rho_B(r) \rangle_{local}$ en todo el volumen y el número total de partículas A (N_A).

II.7 SASA

El análisis del área de la superficie accesible al disolvente (SASA, por sus siglas en inglés) permite identificar las regiones de la proteína con mayor y menor exposición al disolvente debida a la naturaleza hidrofóbica o hidrofílica de los residuos presentes. Para ello, se define una esfera de prueba cuyo radio es igual al radio de la molécula del disolvente (en GROMACS el radio de la molécula del agua es de 0.14 nm), los átomos de la proteína se definen a partir de esferas con su correspondiente radio de van der Waals. En la figura II.7.1 se ilustra la base de este análisis. La idea consiste en deslizar la esfera de prueba sobre la superficie de la proteína generando una envolvente dibujada desde el centro de la esfera de prueba, aquellas regiones en las que la esfera de prueba pueda acomodarse se consideran parte del área de la superficie accesible al disolvente.

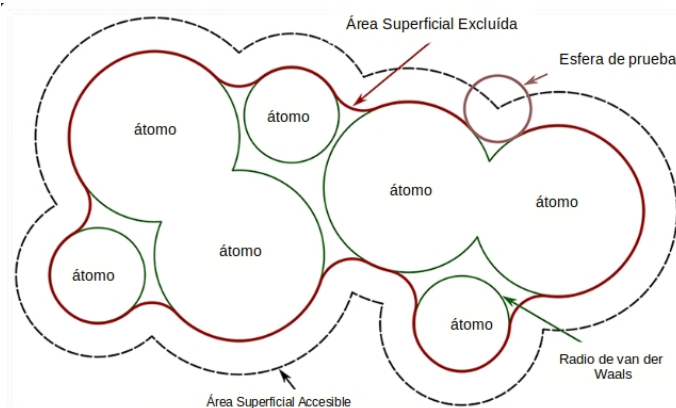


Figura II.7.1. Esquema del área de la superficie accesible al disolvente.

II.8 Análisis de puentes de hidrógeno

Los puentes de hidrógeno son considerados como un tipo “especial” de interacciones no covalentes, dado que, por sus características, estos no pueden ser clasificados como interacciones meramente electrostáticas (interacciones entre momentos n-polares) o de dispersión (efectos cuánticos).

Puesto que no siempre están involucrados átomos electronegativos en la formación de los puentes de hidrógeno, una mejor y más moderna definición es aquella propuesta por Weinhold y Klein, donde un puente de hidrógeno es “un enlace químico fraccional entre ácidos y bases de Lewis $A-H\cdots B$ (que involucra hidrógeno), asociado a una resonancia parcial intermolecular $A-H\cdots B \leftrightarrow A:\cdots H-B$ comúnmente originado en la interacción entre un par libre n_B de una base de Lewis y un híbrido de antienlace σ^*_{AH} de un ácido de Lewis” [48]. Adicionalmente, a este tipo de interacciones se le suele asociar una fuerza de interacción de entre 0.6 a 45 kcal/mol [48].

II.9 Línea temporal de las estructuras secundarias

La línea temporal de las estructuras secundarias es una herramienta que permite identificar la composición de la estructura terciaria de una proteína durante la simulación de MD.

En VMD este análisis se realiza por medio del algoritmo: STRIDE por sus siglas en inglés (STRucture IDentification), que evalúa la información de la energía de los enlaces de H junto con la información de los ángulos de Ramachandran¹ (φ y ψ) derivados estadísticamente de la cadena principal de la proteína de acuerdo con sus coordenadas. Durante la asignación de la estructura secundaria el inicio y fin de cada una es determinado por un factor de peso en el producto de las energías de enlaces de H y de las probabilidades de los ángulos de torsión que corresponden por ejemplo, a α -hélices u *hojas*- β .

La figura II.9.1 muestra el código de colores utilizado en VMD para identificar cada una de las posibles estructuras secundarias presentes. Los lazos (T: Turn) se muestran en color verde, las *hojas*- β (E: Extended Configuration) en color amarillo, los puentes aislados (B: Isolated Bridge) en color verde lima, las α -hélices en color magenta, las 3_{10} -hélices (G) en color azul, las π -hélices (I) en color rojo, mientras que el resto de estructuras que no se encuentren dentro de estas categorías pertenecen a vueltas (C: Coil) en color blanco.

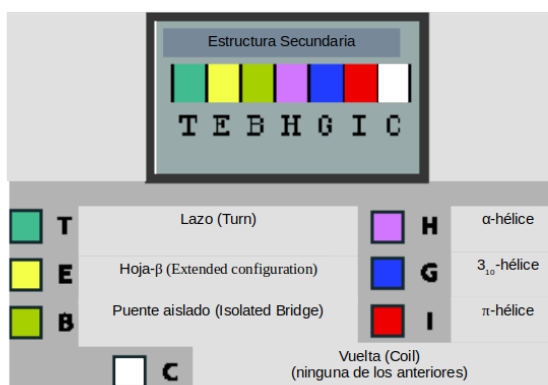


Figura II.9.1. Código de colores de VMD para la línea temporal de estructuras secundarias[47].

¹ Ángulos de torsión en los residuos de un polipéptido.

Una representación alternativa de la línea temporal para comprender la evolución de la estructura terciaria se consigue a través de la compilación de diagramas de Ramachandran de la proteína para cada tiempo t de la trayectoria.

II.9.1 Diagramas de Ramachandran

Los diagramas de Ramachandran son una representación bidimensional de la distribución de los posibles valores de los ángulos de torsión φ y ψ en los residuos de un polipéptido. Un posible tercer ángulo de torsión es ω . En la figura II.9.1.1.a se muestran las rotaciones alrededor de la cadena principal de un polipéptido que describen los ángulos: φ , ψ y ω en los enlaces: $N-C_\alpha$, $C_\alpha-C$ y $C-N$, respectivamente. Al conjunto de estos tres ángulos también se les conoce como ángulos de Ramachandran.

El enlace peptídico es incapaz de rotar debido al enlace parcial doble que presenta, su planaridad es de 180° en casi todos los enlaces peptídicos de la cadena principal. En consecuencia, cada enlace peptídico define un plano con 6 átomos y solo se toman en consideración los ángulos φ y ψ definidos en el intervalo $[-\pi, \pi]$.

En otras palabras, la rotación de la cadena principal de una proteína puede ser descrita como la rotación de un plano de enlace peptídico relativo a otro.

Los ángulos de torsión proveen la flexibilidad para que la cadena principal de una proteína adopte una conformación. De forma que, si se conocen los tres ángulos de Ramachandran será posible predecir el plegamiento de la estructura terciaria de una proteína.

En los diagramas de Ramachandran los átomos son tratados como esferas rígidas definidas por el correspondiente radio de van der Waals. Las combinaciones de los ángulos φ y ψ que resulten en colisiones entre las esferas corresponden a conformaciones impedidas estéricamente o no permitidas. Por el contrario, aquellas combinaciones que no resultan en colisiones corresponden a conformaciones permitidas o estructuras secundarias. En la figura II.9.1.1.b se muestra un ejemplo de un diagrama de Ramachandran donde las regiones blancas corresponden a conformaciones donde los átomos en el polipéptido se aproximan más que la suma de sus radios de van der Waals (son regiones estéricamente impedidas o conformaciones no permitidas).

Las regiones rojas corresponden a conformaciones donde no hay colisiones estéricas, son conformaciones permitidas (α -hélice, hoja- β) y las regiones azules corresponden a conformaciones permitidas donde el radio de van der Waals utilizado en los cálculos es un poco menor al usual, por lo que los átomos pueden aproximarse ligeramente.

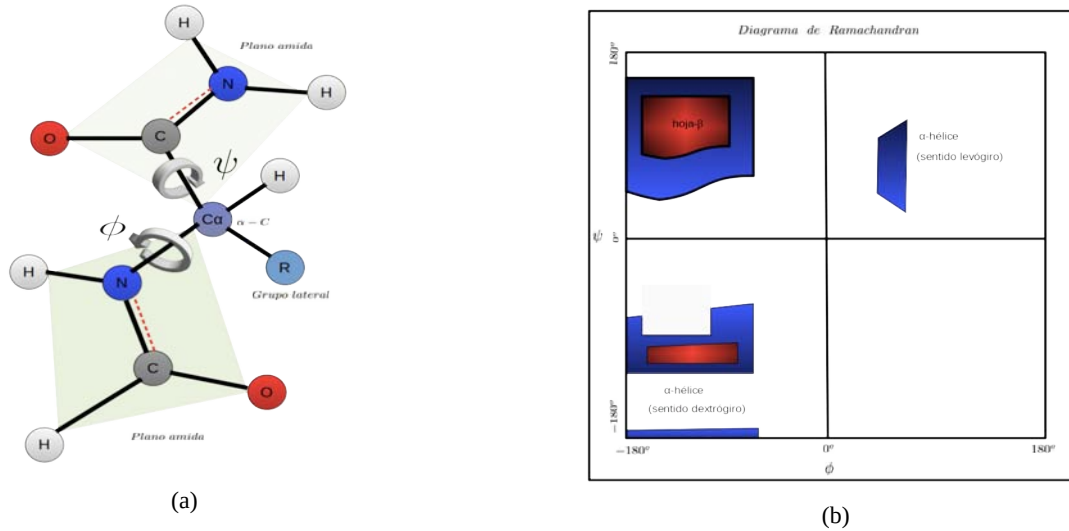


Figura II.9.1.1. (a) Representación de los ángulos de Ramachandran en un polipéptido. (b) Representación de un diagrama de Ramachandran.

II.10 RMSF

El parámetro de fluctuación cuadrática media (RMSF, por sus siglas en inglés), es una cantidad que permite caracterizar los cambios estructurales de cada residuo (también puede ser átomo o partícula) en la evolución temporal de la posición de cada residuo en una configuración inicial tomada como referencia. Esta configuración inicial de referencia será la posición promediada en el tiempo de la misma partícula i .

La diferencia con el parámetro RMSD definido anteriormente es que el RMSF es promediado sobre el tiempo, asignando un valor (factor de peso) para cada partícula i , mientras que el RMSD toma el promedio sobre todas las partículas.

En GROMACS, este parámetro se define como:

$$RMSF_i = \left[\frac{1}{T} \sum_{t_j=1}^T |r_i(t_j) - r_i^{ref}|^2 \right]^{1/2} \quad (46)$$

Donde T es el tiempo sobre el cual se va a promediar.

Capítulo 3

III. Metodología

Por medio de MD se estudió la interacción del protómero RBD de la subunidad S1 que corresponde a la glicoproteína espicular S y una superficie metálica de Cu, una superficie metálica de Au y una superficie de carbono grafito en SARS-Cov-2. El modelo seleccionado de la glicoproteína espicular S del SARS-Cov-2 es el 6vsb.pdb, proporcionado por la base de estructuras biológicas: Protein Data Bank (PDB). Posteriormente, con la asistencia del software SWISSADD se identificaron los enlaces y aminoácidos que conciernen al fragmento RBD para un adecuado aislamiento de su estructura. La dinámica molecular fue ejecutada con el software GROMACS, el software VMD como visualizador.

En materia del protocolo seguido para lograr una ejecución eficiente con dinámica molecular en el estudio de este fenómeno de adsorción, se dividió el sistema en tres etapas, la primera de ellas correspondió a la aplicación del protocolo usual de MD sobre el fragmento RBD con el objetivo de obtener la estructura energéticamente favorable y mejor descrita del protómero RBD. Este procedimiento se dividió en cinco bloques principales.

El primero de ellos comprendió la elección del campo de fuerzas a utilizar para la descripción de las interacciones entre los componentes involucrados en el sistema de estudio, en este caso: el protómero RBD, las superficies y agua, para los cuales el campo de fuerzas CHARMM36 fue el seleccionado. Aunque este campo de fuerzas no tiene parámetros adecuados para la descripción de interacciones con Cu y Au, existen artículos en la literatura en los que con base en cálculos de DFT se proponen parámetros para el Cu y Au que resultan en una descripción más precisa de interacciones entre superficies de Cu, Au y biomoléculas, así como con agua, por lo que aprovechando estos parámetros el campo de fuerzas CHARMM36 resulta adecuado para la descripción del fenómeno de estudio. En este mismo bloque se realizó la construcción de la geometría del sistema, con la asistencia de VMD y su consola tkl podemos correr un script que nos permita rotar la posición del RBD de forma que los aminoácidos del extremo expuesto se encuentren orientados en la dirección de la superficie para la futura interacción y guardar esas coordenadas. Para la construcción de la caja de simulación, debido a que se utilizaron condiciones de frontera periódicas, las dimensiones se definieron de acuerdo con la convención de mínima imagen para evitar interacciones del protómero con sus imágenes periódicas virtuales. Después, se definió la topología del sistema en donde se encuentran las interacciones de enlace que incluye enlaces, ángulos de enlace, ángulos diedros propios, ángulos diedros impropios, así como interacciones de antienlace o de van der Waals. Seguido de ello se realizó la solvatación del sistema en la caja de simulación con el modelo de agua TIP3P para simular el entorno acuoso de las gotitas de flush, enseguida se realizó la neutralización de la carga eléctrica del sistema por medio de iones con carga opuesta para evitar el cálculo de fuerzas espureas.

El segundo bloque comprendió una minimización energética del sistema a través del algoritmo steepest descent con la finalidad de eliminar traslapes entre átomos y moléculas, encontrar posiciones y distancias óptimas entre moléculas y obtener una configuración energética más favorable. El tercer bloque estuvo constituido por la equilibración del sistema; en general, esta etapa se lleva a cabo en función de una propiedad de interés dentro de un ensamble estadístico a partir de la configuración previamente minimizada. En este caso, el fenómeno de estudio es la adsorción del protómero con la superficie, por lo que la implementación del ensamble NVT también conocido como isotérmico-isocórico resulta adecuada para este trabajo.

El penúltimo y cuarto bloque comprendió la producción de resultados con la configuración previamente minimizada y equilibrada a la temperatura deseada (298 K). Esta configuración final fue la utilizada como configuración inicial en la tercera etapa para la interacción con la superficie de cobre. La segunda etapa correspondió a la construcción de la superficie de cobre para la cual fue necesario determinar el plano para clivar la celda unitaria y de ese modo generar la correspondiente superficie, cuya cantidad de capas estará definida por el radio de corte del campo de fuerzas utilizado (1.2 nm en CHARMM36), para culminar con la tercera etapa, nuevamente se aplicó el protocolo de dinámica molecular a la interacción entre el protómero RBD y cada superficie con la diferencia de que en este punto el quinto bloque correspondiente al análisis de resultados fue desarrollado en función de las propiedades y características que permitan identificar la pérdida de actividad biológica en el protómero RBD: RMSD, radio de giro, puentes de hidrógeno, SASA, perfil de densidad parcial, función de distribución radial, número de contactos entre RBD-superficie durante la adsorción y línea de tiempo de estructuras secundarias y RMSF.

Además de la caracterización descrita, el análisis permitió el estudio de la evolución conformacional de los residuos más importantes en el vínculo con la célula receptora humana ACE2 (dentro de nuestro modelo estaban presentes 6 de los 12 receptores) durante la interacción con la superficie. Somos conscientes de que este no era uno de los objetivos propuestos, ni parte de la hipótesis planteada; Sin embargo, este análisis puede ser relevante para investigaciones futuras que puedan continuar en esta dirección.

Capítulo 4

IV. Resultados y análisis

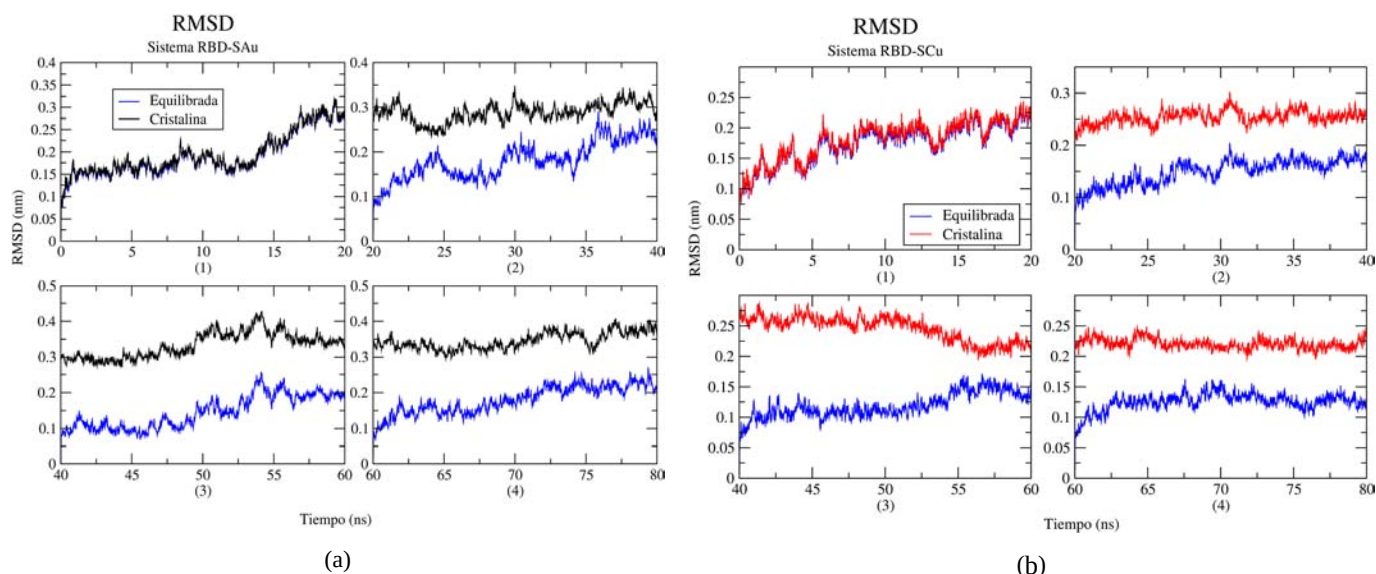
En la literatura se ha visto que estos materiales son capaces de inhibir el SARS-Cov-2, incluso en qué medida son capaces de afectar a la proteína S del virus. Además de los análisis rutinarios en MD de este estudio tales como RMSD, radio de giro, SASA, entre otros, se han incluido RMSF y líneas temporales de la estructura secundaria del RBD con el objetivo de sacar el mayor provecho posible a los resultados generados por las simulaciones de MD. A continuación se empezará con el análisis de los resultados obtenidos en el RMSD.

IV.1 RMSD

La raíz de la desviación cuadrática media (RMSD: Root Mean Square Deviation) de la diferencia en las posiciones de la configuración inicial con respecto a la configuración final, es un parámetro que puede ser usado como un indicador de la estabilidad global de la estructura en la adsorción del protómero RBD sobre la superficie. Una diferencia grande representa un cambio importante en el esqueleto de la proteína, cambios grandes podrían llevar a la desestabilización de la proteína.

En la figura IV.1.1 se exhibe el comportamiento del RMSD como función del tiempo, distribuido en cuatro bloques de 20 ns en orden ascendente de izquierda a derecha, comenzando en el cuadro superior izquierdo. Este parámetro se obtuvo tomando como referencia la estructura cristalina del RBD; es decir, posterior a la minimización de energía con respecto a la estructura instantánea del RBD en la producción de datos para cada sistema. Los tres fenómenos analizados han sido etiquetados de la siguiente manera: Complejo RBD aislado (azul marino), complejo RBD y superficie de oro, cobre y Carbono grafito son, respectivamente: SAu (negro), SCu (rojo) y SC (verde), es importante tomar en cuenta que a lo largo de toda la tesis usaremos este código de color, para facilitar el análisis.

Incluir el RMSD obtenido tomando como referencia la estructura equilibrada del RBD (azul) permitirá conocer la estabilidad de su estructura en la interacción con cada una de las superficies, como se describirá en el siguiente párrafo.



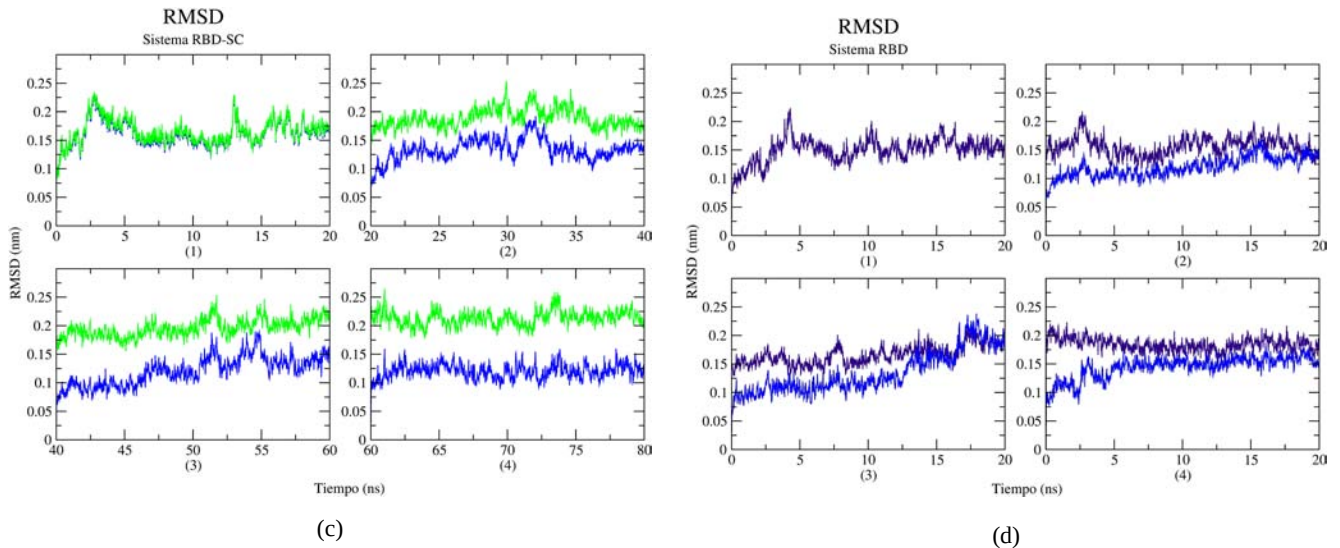


Figura IV.1.1 Gráfica del RMSD como función del tiempo de la simulación para la interacción del RBD relativo a la estructura cristalina en (a) Superficie de SAu (negro), (b) SCu (rojo), (c) SC (verde) y (d) RBD aislado relativo a la estructura cristalina (azul marino). Se muestra el RMSD de la estructura equilibrada (azul) como referencia.

Incluir el RMSD obtenido tomando como referencia la estructura equilibrada del RBD (azul) permite conocer la estabilidad de su estructura en la interacción con cada una de las superficies, como se describirá en el siguiente párrafo.

De acuerdo con el comportamiento del RMSD con respecto a las estructuras equilibrada y cristalina (fig. IV.1.1), existe un traslape de ambas gráficas durante el primer bloque de 20 ns en todos los casos (ver fig. IV.1.1 a.1, b.1, c.1 y d.1), esto sugiere que el esqueleto del RBD es muy estable en este intervalo de la simulación. Esta estabilidad se ve perturbada a partir del segundo bloque en los cuatro casos (ver fig. IV.1.1 a.2, b.2, c.2 y d.2). Con una diferencia aproximada en la evolución inicial y final de ambas estructuras de referencia de 0.2 nm a 0.05 nm (ver fig. IV.1.1 a.2), 0.125 nm a 0.1 nm (ver fig. IV.1.1 b.2), 0.1 nm a 0.047 nm (ver fig. IV.1.1 c.2) y de 0.08 nm a 0.01 nm (ver fig. IV.1.1.d.2), de 0.2 nm a 0.15 nm (ver fig. IV.1.1 a.3), 0.2 nm a 0.08 nm (ver fig. IV.1.1 b.3), 0.1 nm a 0.065 nm (ver fig. IV.1.1 c.3) y de 0.08 nm a 0.01 nm (ver fig IV.1.1.d.3), y de 0.275 nm a 0.15 nm (ver fig. IV.1.1 a.4), 0.15 nm a 0.1 nm (ver fig. IV.1.1 b.4), 0.125 nm a 0.1 nm (ver fig. IV.1.1 c.4) y de 0.124 nm a 0.025 nm (ver fig. IV.1.1.d.4).

El valor de la diferencia en la evolución inicial y final de la discrepancia del RMSD relativo entre ambas estructuras de referencia disminuyó desde 0.053 nm durante el segundo bloque (ver fig. IV.1.1 c.2) a 0.035 nm en el tercer bloque (ver fig. IV.1.1 c.3) y terminó en 0.025 nm en el cuarto bloque (ver fig. IV.1.1 c.4) a lo largo de la interacción con SC. En la interacción con SCu este valor aumentó de 0.025 nm en el segundo bloque (ver fig. IV.1.1 b.2) a 0.120 nm en el tercer bloque (ver fig. IV.1.1 b.3) y volvió a disminuir a 0.05 nm en el cuarto bloque (ver fig. IV.1.1 b.4). En la interacción con SAu este valor fue de 0.15 nm en el segundo bloque (ver fig. IV.1.1 a.2), disminuyó a 0.05 nm en el tercer bloque (ver fig. IV.1.1 a.3) e incrementó a 0.125 nm en el último bloque (ver fig. IV.1.1 c.4). En el RBD aislado, este valor fue de 0.07 nm en el segundo bloque (ver fig. IV.1.1.d.2), se mantuvo igual en el tercer bloque (ver fig. IV.1.1.d.3) e incrementó a 0.09 nm en el último bloque (ver fig. IV.1.1.d.4).

Dado que para todos los casos los cambios son más pequeños, estos datos indican una mayor estabilidad en la estructura del protómero en la interacción con SCu en el segundo bloque, y SC en el tercer y cuarto bloque.

Podemos concluir que sí hay cambios estructurales en el RBD guiados por la interacción con las tres superficies. En cada caso, conforme la diferencia del RMSD entre ambas gráficas disminuye, aumenta la estabilidad de la estructura aislada del RBD o en la interacción del RBD con las tres superficies. Por tanto, la estructura del RBD se mantuvo más estable al estar aislada. En presencia de una superficie la estructura del RBD se mantuvo más estable al interactuar con SC, seguida de la interacción con SCu y finalmente la interacción con SAu reflejó una menor estabilidad en la estructura del RBD.

Es importante analizar el comportamiento del RBD sobre las superficies en forma global y comparar en cuál de las tres hubo mayores cambios en el esqueleto con respecto a la del RBD aislado, en la figura IV.1.2 se muestra el RMSD tomando como referencia la estructura cristalina con respecto a la instantánea.

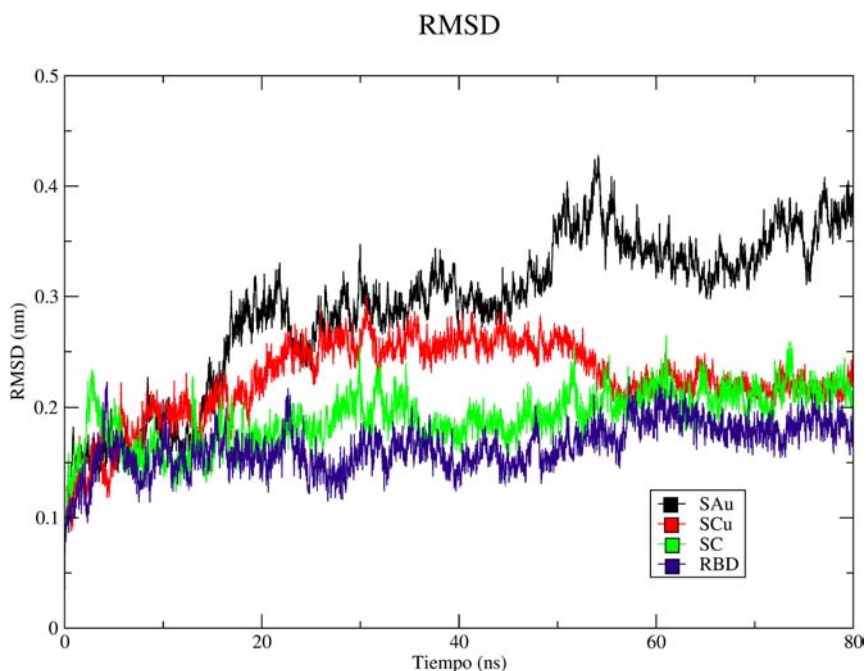


Figura IV.1.2. Gráfica del RMSD como función del tiempo para el RBD aislado (azul marino) y para la interacción del RBD con cada superficie: SAu (negro), SCu (rojo) y SC (verde).

La figura IV.1.2 muestra el RMSD relativo a la estructura cristalina como función del tiempo para el caso de la interacción del RBD con cada una de las superficies. Marcando el inicio de la adsorción en 10 ns^1 puede observarse que el ΔRMSD entre el RBD aislado y adsorbido es de $\sim 0.2 \text{ nm}$ para SAu y SCu y de $\sim 0.175 \text{ nm}$ para SC, indicando pequeños cambios estructurales durante la adsorción. En el caso de SC, el RMSD permanece constante al inicio de la adsorción, pero crece continuamente en el caso de las otras dos superficies. El $\Delta\text{RMSD} > 0$ es un indicador de la inestabilidad de la estructura en la interacción con la superficie. En la adsorción, el cambio más abrupto se observa en $t \sim 15 \text{ ns}$, $\sim 17 \text{ ns}$ y $\sim 2.5 \text{ ns}$ para SAu, SCu y SC, respectivamente. La adsorción del RBD en cada superficie se estabilizó a tiempos mayores a $t \sim 57 \text{ ns}$, este valor permitirá establecer una etapa final de adsorción, como se verá más adelante. De manera global, se estima un ΔRMSD de 0.275 nm , 0.137 nm , 0.112 nm y 0.095 nm en la interacción con SAu, SCu, SC y RBD solo, respectivamente. Además, el comportamiento global del RMSD en el RBD aislado es predominantemente más estable durante la simulación en comparación con los RMSD del RBD en interacción con las superficies, por lo que las interacciones con las superficies perturban la estabilidad de la cadena principal del RBD y esos cambios conformacionales

¹ La justificación de esta elección se explicará en el apartado del número de contactos.

están siendo realmente inducidos por la interacción con cada superficie, cuya magnitud en los cambios es totalmente dependiente de la naturaleza de la superficie presente en la interacción. Para facilitar la apreciación de los cambios estructurales en la figura IV.1.3 se presenta la estructura del RBD en el comienzo de la etapa inicial y al culminar la etapa final de la adsorción sobre las tres superficies.

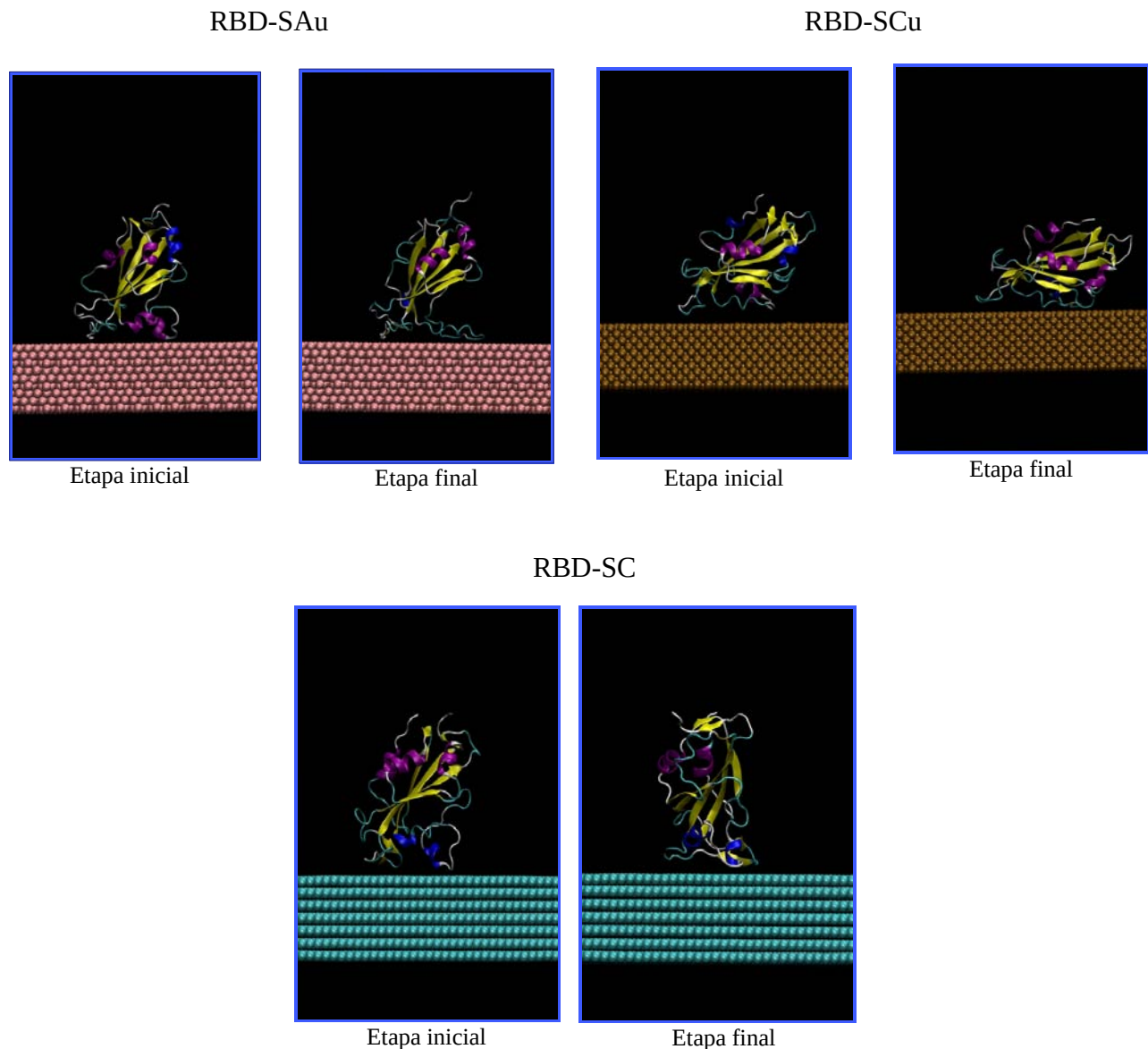


Figura IV.1.3. Imagen de la estructura del RBD al comenzar la etapa inicial y al terminar la etapa final de adsorción sobre cada una de las tres superficies. Se muestra la configuración del RBD al inicio y al final del proceso de adsorción. a) SCu, b) SAu y c) SC.

Los Δ RMSD encontrados sugieren cambios significativos en la estructura del RBD inducidos por las superficies metálicas (SAu y SCu), por otra parte, el protómero mostró un comportamiento más rígido al no verse afectado de manera significativa por los cambios inducidos por SC. Siendo el Δ RMSD de esta última interacción el más similar al de la estructura aislada del RBD. Este análisis indica una interacción más fuerte con las superficies metálicas en comparación con el grafito.

Estos resultados están en sincronía con los hallazgos reportados en los artículos mencionados en la introducción [8] donde la estructura del RBD mostró el comportamiento más rígido en comparación

con otras estructuras en la glicoproteína S al interactuar con grafito. Asimismo, la estabilidad de la estructura del RBD al interactuar con superficies y la magnitud de los Δ RMSD obtenidos se encuentran en acuerdo con el RMSD de resultados previos donde nuevamente la estructura del RBD fue la que menos cambios tuvo en este parámetro en comparación con otras estructuras dentro de la glicoproteína S[1].

IV.2 Radio de giro

El radio de giro es una medida de la compacidad en una estructura que nos permite evaluar su desplegamiento durante la trayectoria de la simulación [32]. Si la proteína se despliega en mayor grado en una superficie, podemos decir que la relajación estructural es mayor y esto tiene como consecuencia una mayor desestabilización de la proteína y por tanto, una posible inhibición de su actividad biológica.

En la figura IV.2.1 se observa la evolución temporal del radio de giro del RBD al interactuar las tres superficies y del RBD aislado. Los incrementos en el radio de giro son otro indicador de la inestabilidad de la estructura del RBD debidas a la interacción con las superficies.

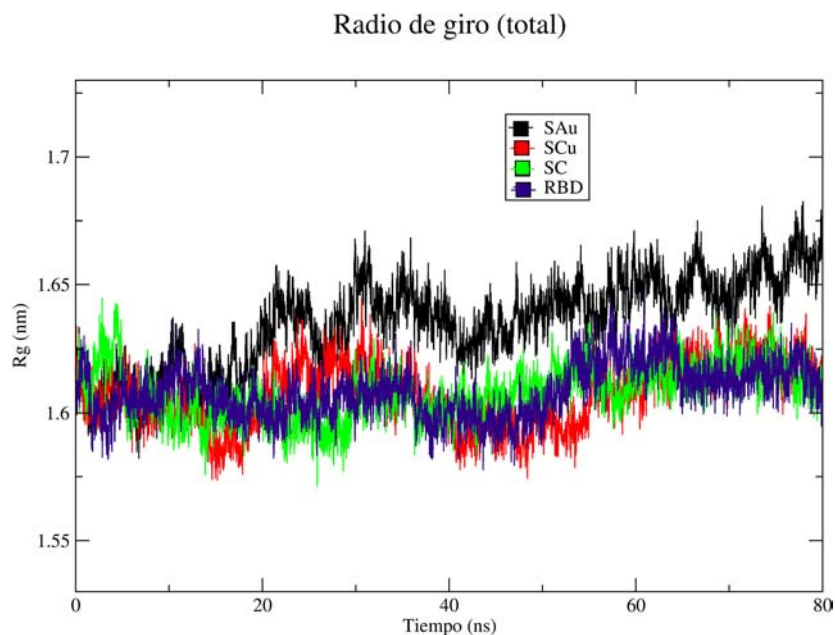


Figura IV.2.1. Radio de giro del RBD aislado (azul marino) y del RBD al interactuar con cada superficie: SAu (negro), SCu (rojo) y SC (verde) como función del tiempo de la simulación.

En el caso de SAu, el valor promedio del radio de giro se incrementa en 0.055 nm con respecto a su valor inicial, mientras que en el caso de SCu el incremento es de tan solo 0.01 nm respecto al valor observado inicialmente. En SC el valor promedio del radio de giro se mantuvo constante a lo largo del tiempo, indicando que el RBD no tiene una interacción eficiente con la superficie de carbono. Mientras que en el RBD aislado, el valor promedio del radio de giro también se incrementó en 0.03 nm respecto a su valor inicial. En SAu, las fluctuaciones del radio de giro son globalmente crecientes, pero, a $t > 40$ ns las fluctuaciones del radio de giro son más pequeñas, indicando que al principio los cambios en la compactación de la estructura eran grandes, pero al final esta compactación disminuyó. En SCu la

fluctuación más grande se ve alrededor de $t \sim 25$ ns y vuelve a presentarse a $t \sim 74$ ns, mientras que muestra dos mínimos de similar magnitud en $t \sim 15$ ns y uno más extenso en el rango $t \sim 40-55$ ns. Estos incrementos están relacionados con una relajación en la empaquetación de la configuración 3-D de las estructuras secundarias que conforman el RBD.

En la referencia [49] se ha probado que las α -hélices actúan como una barrera de plegamiento en proteínas, por lo que la tasa de plegamiento depende de la longitud efectiva L_{eff} que es la longitud total de regiones no helicoidales de una cadena de proteínas.

Un ejemplo de esto se muestra en la figura IV.2.2 donde el RBD está representado por listones con los colores característicos de las estructuras secundarias existentes en la interacción con SAu ((IV.2.2.a), (IV.2.2.b)) y con SCu ((IV.2.2.c),(IV.2.2.d)). La figura denota cambios conformacionales de una α -hélice y una hoja-2- β en $t \sim 18$ ns (IV.2.2.a) a una loop y un puente aislado β (IV.2.2.b) respectivamente a $t \sim 22$ ns en la interacción con SAu donde el radio de giro aumenta. Aunque la α -hélice es una estructura que no favorece el plegamiento del RBD, es una estructura que presenta patrones regulares de puentes de hidrógeno, por lo que presenta una estructura más compacta que un loop (segmento con conformaciones no regulares), como en el caso de la transformación de la hoja-2- β a puente aislado β que, en conjunto, incrementan el radio de giro en esta transición.

Otro ejemplo se observa en la interacción con SCu durante el primer mínimo (IV.2.2.c) en $t \sim 15$ ns y durante el primer máximo (IV.2.2.d) en $t \sim 25$ ns del radio de giro. La figura IV.2.2.c exhibe 2 α -hélices, 2 3_{10} -hélices, 1 hoja β -5, 1 hoja β -3 y 1 hoja β -2. Mientras que la figura IV.2.2.d muestra 3 α -hélices, 1 3_{10} -hélice, 1 hoja β -5, y una hoja β -2. Por lo que este incremento en las α -hélices y disminución de las hojas- β es consistente con el incremento del valor del radio de giro, así como de su disminución en el mínimo mencionado.

Con excepción de la fluctuación en $t \sim 5$ ns, SC muestra un radio de giro con un comportamiento más uniforme, durante la simulación. Esto significa que las estructuras secundarias que conforman la configuración 3-D del RBD no sufrieron un cambio de conformación en el tiempo.

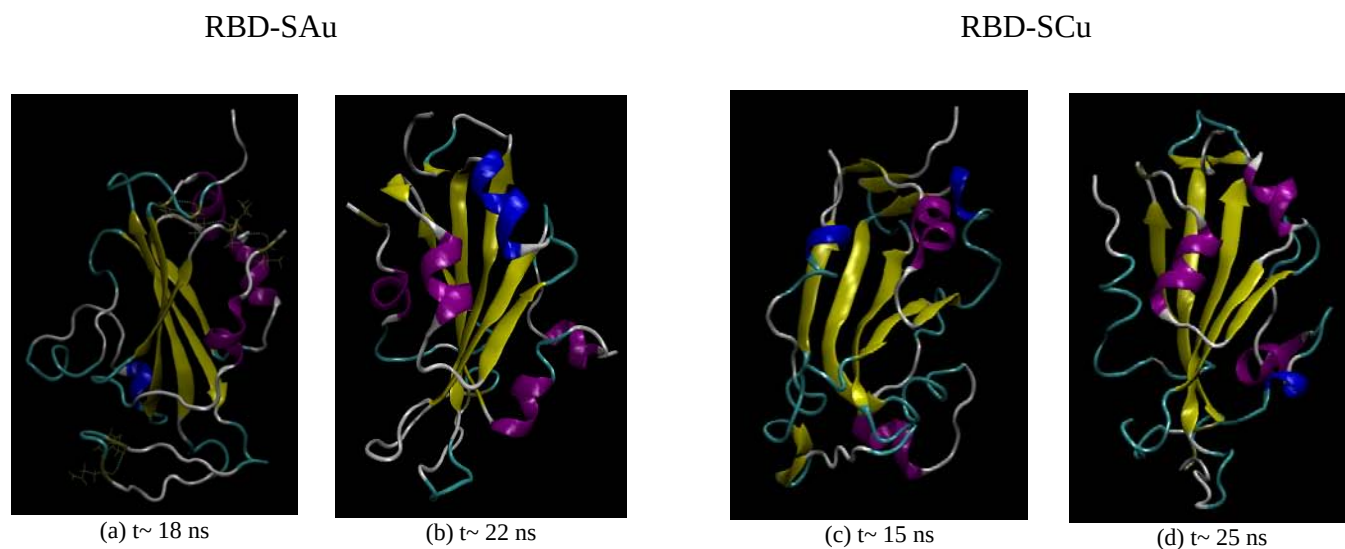


Figura IV.2.2. Estructura del RBD en la representación de listones y código de color de estructura secundaria al interactuar con SAu en (a) $t \sim 18$ ns y después en (b) $t \sim 22$ ns y con SCu en (c) $t \sim 15$ ns y después en (d) $t \sim 25$ ns.

En el RBD aislado el radio de giro muestra un comportamiento constante durante la mayor parte de la simulación, incluso más estable que con SC durante el último intervalo ($t \sim 65$ ns a 80 ns) de la simulación, lo que es un indicador de que la estructura terciaria del RBD alcanzó una configuración más compacta en la parte final. La perturbación en la tendencia del radio de giro es una consecuencia de la relajación en el empaquetamiento de la estructura terciaria del RBD como respuesta a la interacción con cada superficie.

El RMSD y el radio de giro son parámetros relacionados con la integridad estructural y la flexibilidad de la proteína para llevar a cabo su función biológica. Un valor alto del RMSD indica cambios en su estructura principal (esqueleto) que se traduce en una baja estabilidad debida a la interacción, del mismo modo, el radio de giro nos habla de la evolución temporal de la compactación del protómero que es dependiente de las transiciones en las estructuras secundarias que lo conforman. Asimismo, la flexibilidad se define como la capacidad de una proteína para adaptarse (fluctuaciones de átomos) en su entorno sin perder su función principal. En este contexto, la flexibilidad y adaptabilidad son sinónimos[50]. Las fluctuaciones en estos parámetros están directamente asociadas con un impacto en la funcionalidad del protómero.

En este sentido, el RBD aislado, mostró un bajo grado de flexibilidad al presentar un radio de giro que permaneció muy estable y con un valor bajo durante el intervalo simulado.

Por otro lado, el encuentro con SAu significó el mayor cambio inducido en la cadena principal del RBD, y el hecho de que la descompactación de su estructura 3-D haya sido la más notable, sugiere un mayor grado de flexibilidad en la estructura del RBD, que podría conllevar a la inactivación de su función biológica principal.

En contraste, la respuesta observada del RMSD y el radio de giro en la interacción con SC sugieren un menor grado de flexibilidad por parte del RBD al mostrar un comportamiento más rígido, de acuerdo con lo observado en otros estudios en el comportamiento del RBD dentro de la proteína S al interactuar con SC, lo que indica que el RBD es una sección muy rígida en comparación con el resto de la estructura[2].

De acuerdo con el comportamiento observado en el RMSD y el radio de giro, la interacción con SCu indica un caso intermedio en el grado de estabilidad presente respecto a SAu y SC, que son opuestos (cambios muy pronunciados para SAu y cambios poco significativos para SC).

IV.3 Puentes de Hidrógeno

Los puentes de Hidrógeno contribuyen a definir la estructura terciaria del protómero, las fluctuaciones en este parámetro indican cambios estructurales en el protómero y como consecuencia, representan una forma indirecta de analizar su estabilidad frente a una perturbación (presencia de un disolvente, una superficie, entre otras), que conllevan a cambios conformacionales en el RBD.

Uno de los criterios para contabilizar enlaces de hidrógeno de un cálculo de dinámica molecular se basa en las siguientes dos condiciones: el ángulo de enlace donador-Hidrógeno-aceptor debe ser menor o igual a 30° y la distancia al átomo aceptor debe ser menor o igual a 0.35 nm[32].

La dependencia temporal del número total de puentes de Hidrógeno (N_H) dentro de las estructuras secundarias del protómero RBD aislado y al interactuar con cada superficie es presentada en la figura IV.3.1.

Número de enlaces de H

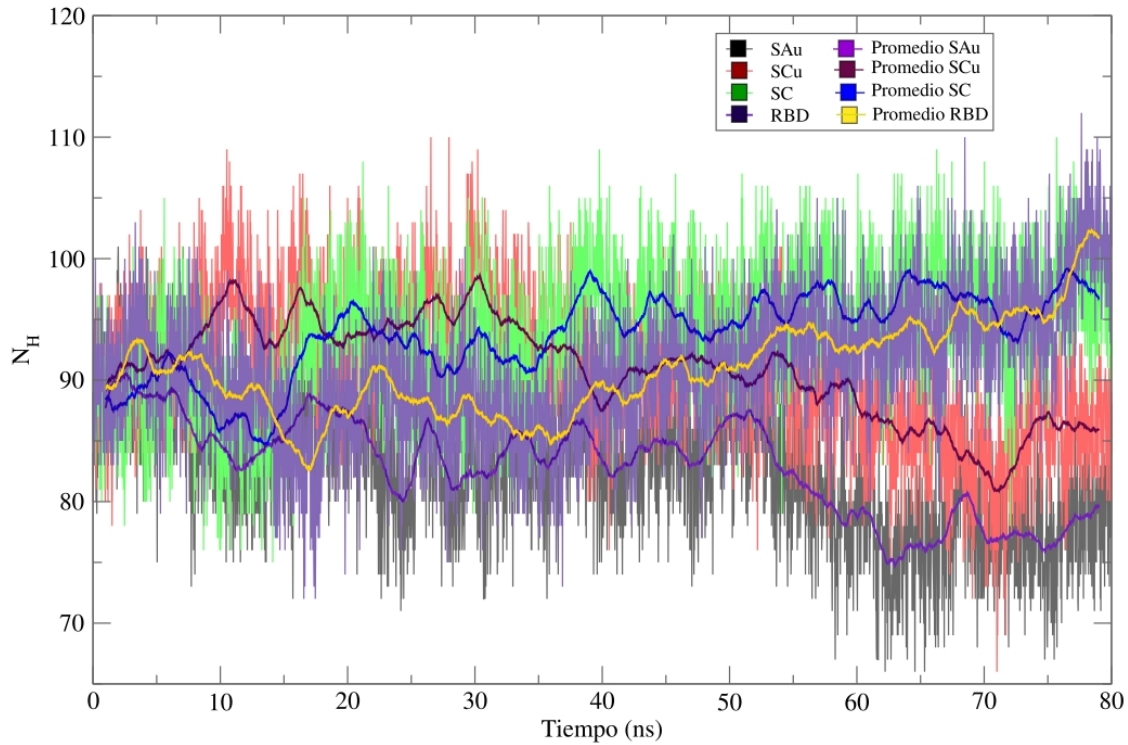


Figura IV.3.1. Evolución temporal del número de enlaces de Hidrógeno formados dentro de la estructura del RBD aislado (azul marino) y dentro de la estructura del RBD en presencia de SAu (negro), SCu (rojo) y SC (verde) con sus correspondientes promedios temporales para SAu (morado), SCu (café), SC (azul) y RBD aislado (amarillo).

En la figura IV.3.1 se puede observar que las fluctuaciones para las tres superficies son más o menos las mismas al inicio del proceso, para SAu el primer mínimo aparece en 0.3 ns, para SCu en 0.2 ns y en SC en 0.52 ns. Esto pone en evidencia que no es fácil distinguir el inicio de la adsorción con este parámetro, dado que son muy cercanos. En la última etapa de la adsorción sí se puede distinguir entre lo que sucede con las superficies metálicas con respecto a lo que sucede sobre la superficie de grafito, dado que para este último como en el caso del RBD aislado no hay fluctuaciones importantes. En SC, el número total de enlaces de Hidrógeno (N_H) en el interior del prótomo RBD tiene una reducción alrededor de $t \sim 10$ ns, se incrementa y mantiene una tendencia constante a partir de $t \sim 20$ ns. Esto también es consistente con la estabilidad observada en el radio de giro al mantener una estructura compacta. En el caso opuesto, la adsorción en SAu indica un comportamiento estable en el N_H hasta $t \sim 20$ ns donde ocurre una declinación del mismo que visto desde el punto de vista de la estabilidad sugiere el desplegamiento del RBD, justificado por el aumento del valor en el radio de giro. Mientras que el resto de la tendencia decreciente en N_H empata con el desplegamiento más pronunciado mostrado en el radio de giro.

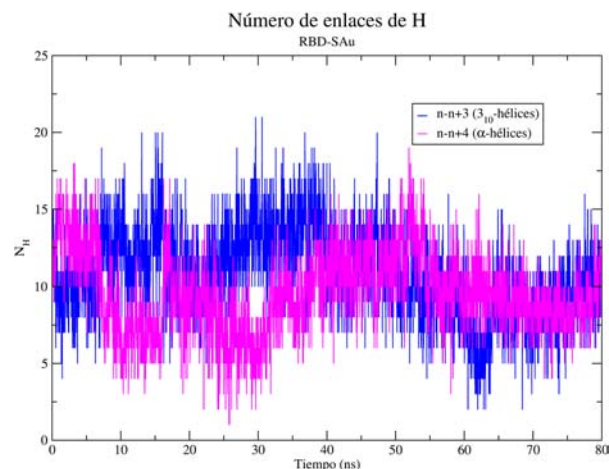
El número total de enlaces de hidrógeno en SCu se incrementa en $t \sim 10$ ns y mantiene un comportamiento constante hasta $t \sim 30$ ns donde ocurre una disminución de N_H , conservando una tendencia decreciente en el resto de la simulación. Esto explica la disminución en la compactación de la estructura del RBD reflejada en el radio de giro y que, aunque es decreciente no es tan pronunciada como en el caso de SAu.

Por último, en el RBD aislado, el N_H disminuye en $t \sim 12$ ns, se incrementa a partir de $t \sim 21$ ns y conserva un incremento constante en el intervalo restante, consistente con la compactación de la estructura del RBD aislado al final de la simulación mostrada en el radio de giro.

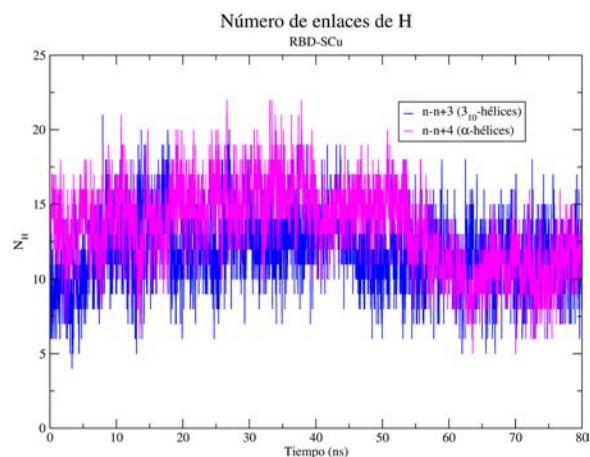
En esta gráfica podemos observar que se puede distinguir con claridad una tendencia para el RBD aislado y al interactuar con cada superficie, y el grafito se distingue claramente de las dos superficies metálicas, N_H es mayor para SC y N_H menor para SAu y SCu, SAu presenta el menor N_H de todos los estudiados.

En comparación con la notable diferencia debida a la presencia de las superficies metálicas, la interacción con SC es la que muestra la tendencia más parecida a la del RBD aislado, pudiéndose apreciar que al interactuar con SC, el N_H se ve favorecido con excepción de los últimos 10 ns de la simulación donde el promedio oscila alrededor de un promedio similar al del RBD aislado. Estas observaciones sugieren un ligero aumento de la estabilidad en la estructura del RBD al favorecer el N_H dentro de las estructuras secundarias del RBD al interactuar con SC, mientras que exhiben una pérdida en la integridad estructural del RBD reflejado en la disminución de N_H al interactuar con las superficies metálicas.

La figura IV.3.2 muestra en detalle la descomposición de enlaces de hidrógeno en estructuras secundarias que también contribuyen en la formación de estructuras helicoidales dentro del RBD al interactuar con cada superficie: SAu (a), SCu (b) y SCg (c) y dentro del RBD aislado (d). En SAu, la gráfica IV.3.2.a revela las contribuciones debidas a enlaces de hidrógeno entre residuos n y $n+4$, responsables de la formación de α -hélices y de los enlaces de hidrógeno entre residuos n y $n+3$ asociados con 3_{10} -hélices en las que se presentan más fluctuaciones. Sin embargo, los puentes de hidrógeno en las 3_{10} -hélices tienen un valor final que oscila alrededor del valor inicial. En la interacción con SCu (figura IV.3.2.b), para las α -hélices la formación de puentes de hidrógeno oscila alrededor del mismo valor, a partir de $t \sim 40$ ns se ve un descenso en N_H y al final de la dinámica se observa un valor aún más bajo. En el caso de las 3_{10} -hélices se ven favorecida hasta $t \sim 50$ ns se puede detectar la última etapa de adsorción y el valor de N_H oscila alrededor del valor inicial, indicando que la adsorción tampoco tiene efecto sobre los residuos $n-n+3$. Para SC (figura IV.3.2.c), la tendencia denota un comportamiento estable de ambos tipos de estructuras secundarias. En el RBD aislado (figura IV.3.2.d), también se observa un comportamiento estable en los dos tipos de estructuras secundarias durante la simulación. Con esto, se comprueba que la interacción con el carbono no es suficientemente fuerte para provocar cambios notables en estructurales helicoidales.



(a)



(b)

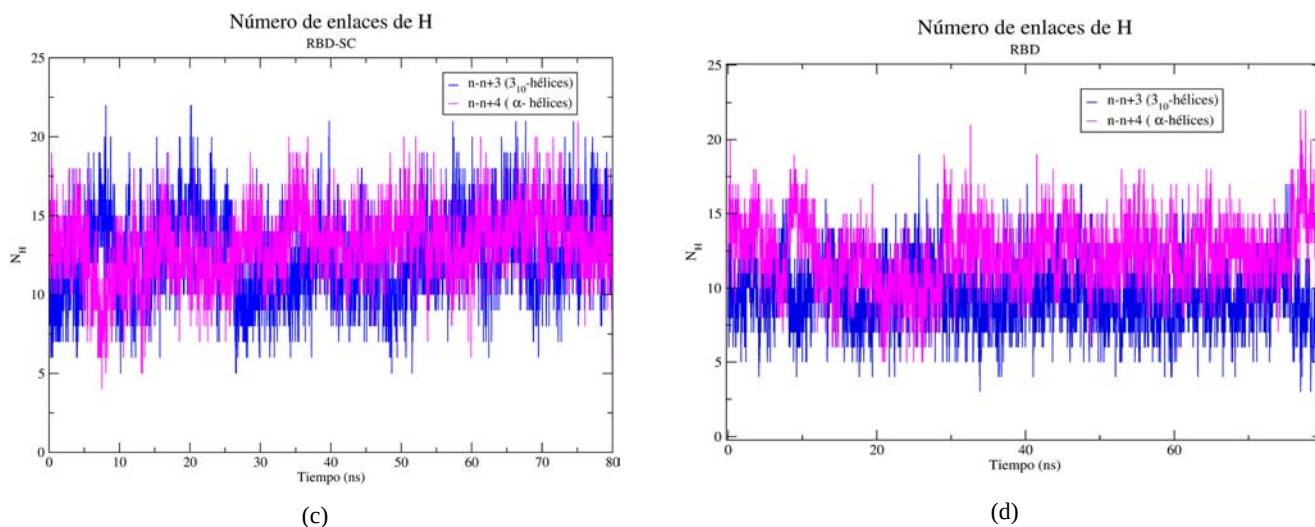


Figura IV.3.2. Evolución temporal de la descomposición en las contribuciones de enlaces de H que surgen de las interacciones entre residuos separados por 2 (azul) y 3 (magenta) residuos del RBD al interactuar con cada superficie: SAu (a), SCu (b) y SC (c) y en el RBD aislado (d).

La evolución temporal del N_H entre el RBD aislado y el RBD al interactuar con cada superficie y las moléculas de agua en el disolvente $(H_2O)_{dis}$ también representan un factor relevante en la estabilización de la estructura del RBD. Este N_H también es afectado por el proceso de adsorción, como se muestra en la figura IV.3.3.

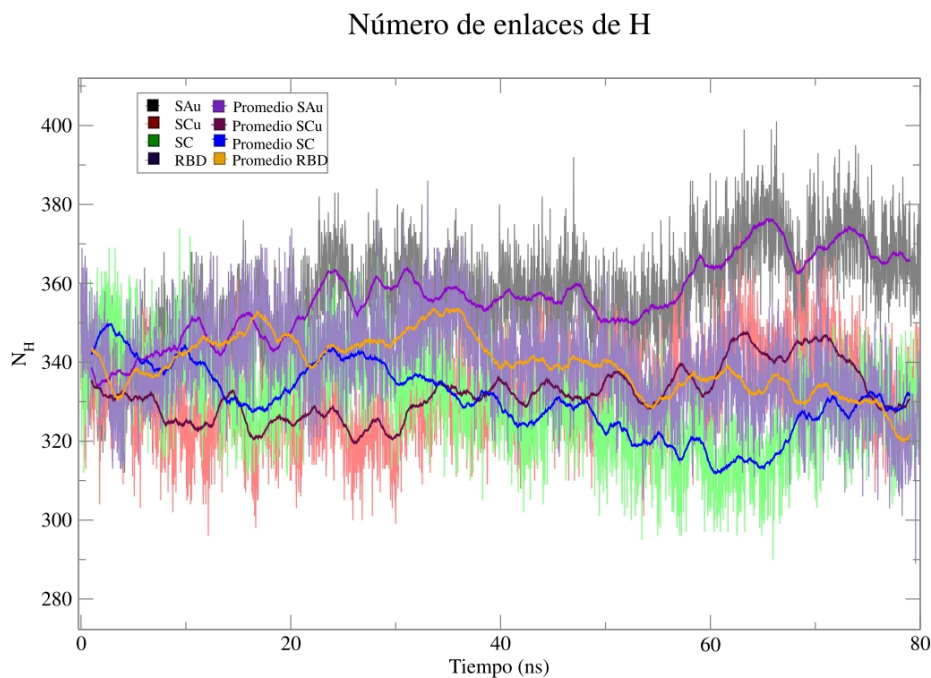


Figura IV.3.3. Evolución temporal del número de enlaces de Hidrógeno formados entre las moléculas de agua del disolvente y la estructura aislada del RBD (azul marino) y entre las moléculas de agua del disolvente y la estructura del RBD en presencia de SAu (negro), SCu (rojo) y SC (verde) y sus correspondientes promedios para SAu (morado), SCu (café), SC (azul) y el RBD aislado (amarillo).

En la figura IV.3.3 se distingue la predominante tendencia creciente en el caso de la interacción con las superficies metálicas y la tendencia decreciente en la interacción con la superficie de grafito así como en el del RBD aislado. En SC, el N_H entre el RBD y $(H_2O)_{dis}$ tiene una reducción alrededor de $t \sim 15$ ns, se incrementa y mantiene una tendencia decreciente a partir de $t \sim 20$ ns hasta $t \sim 65$ ns donde vuelve a incrementarse y termina oscilando alrededor de un promedio constante. En SAu, el N_H indica un comportamiento estable durante el intervalo de $t \sim 22$ ns a $t \sim 44$ ns, posterior a ello ocurre un incremento y el resto de la gráfica permanece constante. El N_H en SCu se reduce en el intervalo a partir de $t \sim 8$ ns hasta $t \sim 12$ ns, mantiene una tendencia creciente en su mayor parte hasta $t \sim 72$ ns donde ocurre un descenso y converge alrededor de un promedio similar al de SC.

En todos los casos se observa la influencia de la interacción con una superficie y una consistente proporcionalidad inversa entre el incremento de N_H entre RBD- $(H_2O)_{dis}$ y el incremento de N_H en el interior del RBD. En el caso de la interacción con las superficies metálicas explica la disminución de N_H intramoleculares del RBD al estar siendo reemplazados por enlaces de hidrógeno entre RBD- $(H_2O)_{dis}$, favorecidos por los residuos polares en las cadenas laterales del RBD y la configuración extendida del RBD adquirida durante la adsorción. La interacción con la SC explica la configuración compacta del RBD adquirida durante la adsorción. La disminución de N_H entre RBD- $(H_2O)_{dis}$ favorece el incremento de N_H intramoleculares del RBD debido a los residuos hidrofóbicos en las cadenas laterales y estructuras secundarias del RBD que tienden a empaquetarse entre sí.

El reemplazo de enlaces de hidrógeno entre RBD- $(H_2O)_{dis}$ con enlaces de hidrógeno intramoleculares del RBD es más determinante dependiendo de la superficie sobre la que el RBD es adsorbido.

IV.4 Área de la superficie accesible al disolvente (SASA)

El área de la superficie accesible al disolvente permite conocer el grado de estabilidad o compacidad de una proteína al interactuar con un disolvente. Como ya se mencionó antes, si el disolvente es agua, SASA permite además conocer la naturaleza hidrofóbica o hidrofílica de la proteína: Un valor creciente de SASA indica una tendencia hidrofílica por parte de la proteína, lo que se traduce como una pérdida en su estabilidad estructural o un incremento en el desplegamiento de la proteína. Por el contrario, un comportamiento decreciente en SASA indica una tendencia hidrofóbica por parte de la proteína, lo que significa mayor estabilidad en su estructura o un plegamiento constante en la estructura de la proteína.

En la figura IV.4.1 se muestra el SASA como función del tiempo para cada uno de los sistemas de interés, RBD-SAu (negro), RBD-SCu (rojo), RBD-SC (verde) y RBD aislado (azul marino).

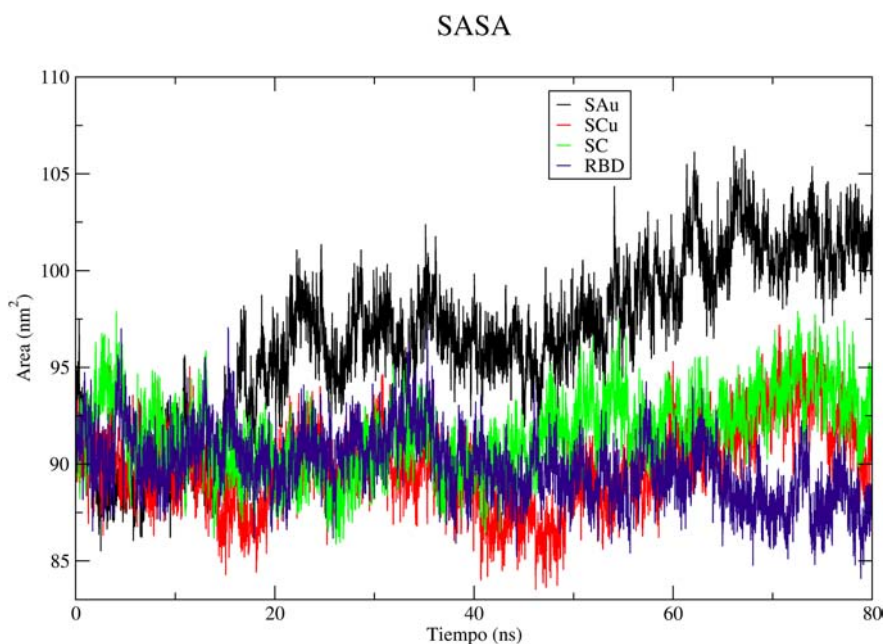


Figura IV.4.1. Área de disolvente accesible como función del tiempo en el RBD al interactuar con la superficie: SAu (negro), SCu (rojo) y SC (verde) y del RBD aislado (azul marino).

En el intervalo 0-10 ns el comportamiento de SASA oscila alrededor de un valor similar en los cuatro casos. En el intervalo de 15-20 ns, el comportamiento de SASA es parecido en SCu y SC (oscila alrededor del mismo valor), mientras que la tendencia es la misma, pero las oscilaciones ocurren a un valor mayor para el RBD aislado en ese mismo intervalo. Para SAu en 15 ns inicia el incremento cuya tendencia oscila alrededor de un valor constante en el intervalo de 20-40 ns como en el caso del RBD aislado y SCu no muestra cambios en este intervalo. El intervalo de 40-50 ns muestra su mínimo absoluto en el valor de SASA para SCu, permanece constante en SC así como para el RBD aislado y en SAu no hay cambios respecto al intervalo anterior. En el intervalo final (60 -80 ns), el valor promedio final de SASA fluctúa alrededor del valor promedio inicial para SCu y SC, en contraste, el valor promedio final aumentó respecto al inicial para SAu y disminuyó para el RBD aislado.

De acuerdo con la descripción anterior, el aumento en el valor de SASA para SAu, está asociado con regiones hidrofílicas predominantes del protómero que favorecieron la descompactación de la estructura terciaria del RBD durante la simulación. Este comportamiento es consistente con aquel observado en el correspondiente radio de giro (ver fig. IV.2.1), cuya justificación estaba basada de entre las razones discutidas, por la pérdida de estructuras secundarias regulares y que a su vez se vio reflejado en la mayor reducción del número de puentes de hidrógeno intramoleculares del RBD de la figura IV.3.1 y del mayor incremento del número de puentes de hidrógeno entre RBD-(H₂O)_{dis} de la figura IV.3.3. De manera complementaria, el conteo del número de puentes de hidrógeno de la figuras IV.3.1 y IV.3.3 fortalecen este análisis y permite sostener que la interacción RBD-SAu es la más fuerte y la que afecta más la estabilidad estructural del RBD.

En SCu, a pesar de que el valor de SASA no presenta cambios tan significativos como los de SAu, la interacción RBD-SCu también mostró un comportamiento con incrementos y decrementos mayores a los mostrados por SC. Los incrementos asociados con regiones mayormente hidrofílicas que favorecieron la reducción en la compactación de la estructura del RBD explican el incremento del radio de giro. De igual forma, para SCu los decrementos, asociados con contribuciones mayormente

hidrofóbicas, favorecieron la estabilidad de la estructura del RBD en esos intervalos (ver fig. IV.2.1). Asimismo, el conteo del número de puentes de hidrógeno intramoleculares del RBD en la figura IV.3.1 junto con el conteo del número de puentes de hidrógeno entre RBD-(H₂O)_{dis} en la figura IV.3.3 complementa este análisis y permite sostener que la interacción RBD-SCu es la segunda menos intensa y que afecta en menor medida la estabilidad estructural del RBD.

En SC el valor de SASA osciló alrededor del mismo valor durante toda la simulación; sin embargo, fue el comportamiento más estable respecto al observado en las dos superficies metálicas analizadas. Con este resultado, podemos suponer contribuciones hidrofóbicas asociadas con la permanencia de la compactación de las estructuras secundarias en la estructura terciaria del RBD, esta afirmación también queda respaldada por el comportamiento menos fluctuante en el radio de giro (ver fig. IV.2.1) y con el conteo de puentes de hidrógeno intramoleculares del RBD, que se mantuvo constante (ver fig. IV.3.1), así como en el conteo de puentes de hidrógeno entre RBD-(H₂O)_{dis}, con valores más bajos (ver fig. IV.3.3) durante la simulación de los cuatro casos analizados.

Para el RBD aislado el parámetro de SASA osciló alrededor del mismo valor durante casi toda la simulación, con excepción del tramo final (65-80 ns) dónde volvió a disminuir, lo que deja en evidencia la influencia de las tres superficies sobre el RBD y la transición de la estructura del RBD aislado en el tramo final a un estado más compacto de las estructuras secundarias en la estructura terciaria del RBD. Afirmación sustentada por la disminución en el comportamiento del radio de giro (ver fig. IV.2.1), del conteo de puentes de hidrógeno intramoleculares del RBD, que aumenta en el tramo final (ver fig. IV.3.1) y que es consistente con la disminución del conteo de puentes de hidrógeno entre RBD-(H₂O)_{dis} en el mismo tramo.

IV.5 Función de distribución radial

La función de distribución radial (RDF, por sus siglas en inglés) permite conocer la probabilidad de encontrar una partícula a una determinada distancia de una partícula de referencia. La variación de la RDF como función de la distancia medida desde una partícula de referencia tiene una envolvente característica para cada fase, por lo que esta función brinda información de las fases involucradas en un sistema. Otra forma de interpretar esta función es a partir de dos grupos distintos constituyentes del sistema y considerando a uno de ellos como el grupo de referencia. Por ejemplo, si el primer grupo estuviese integrado por las moléculas de agua y los átomos de la superficie conformaran el segundo grupo, la RDF mostraría la distribución de las moléculas de agua en función de una determinada distancia con respecto a los átomos de la superficie. Si el primer grupo estuviese definido por los átomos de la superficie y el segundo por los átomos de la proteína, la RDF mostraría la distancia a la cuál están ocurriendo las primeras interacciones superficie-proteína.

En la figura IV.6.1 se muestra la RDF para dos interacciones distintas, la superficie siempre será el grupo de referencia y se verá su interacción con moléculas de agua en presencia del RBD (negro), con moléculas de agua en ausencia del RBD (verde) y con los átomos del RBD (rojo), para SAu (ver fig. IV.6.1.a), SCu (ver fig. IV.6.1.b) y SC (ver fig. IV.6.1.c), respectivamente.

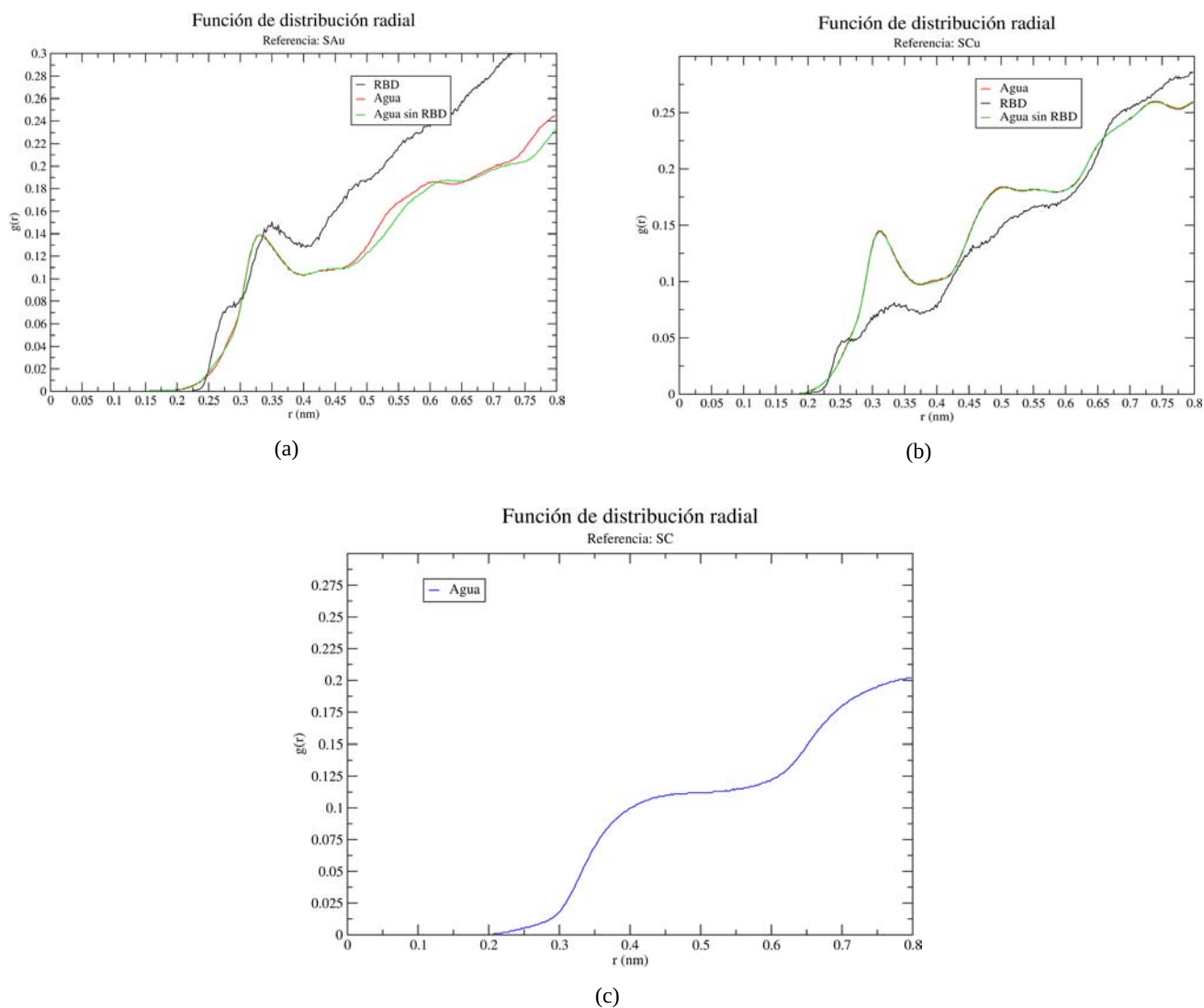


Figura IV.6.1. Función de distribución radial con el segundo grupo el disolvente sin el RBD (verde), el disolvente con el RBD (rojo) y el RBD (negro), tomando las superficies (a) SAu, (b) SCu y (c) SC como el grupo de referencia.

En SAu, la RDF (rojo) y la RDF (verde) de la figura IV.6.1.a muestran un primer pico a ~ 0.33 nm y un segundo pico a ~ 0.60 nm. En SCu, la RDF (rojo) y la RDF (verde) de la figura IV.6.1.b muestran la primera capa de hidratación a ~ 0.32 nm y la segunda, menos prominente a $0.72 \sim$ nm respecto a la superficie. En SC la contribución de los otros átomos de carbono unidos al átomo de carbono de referencia están apantallando las correlaciones por pares respecto a los átomos del segundo grupo, por lo que no es posible apreciar la primera capa de hidratación a partir de su RDF. Tampoco resulta un parámetro adecuado para caracterizar sistemas donde la superficie tiene átomos enlazados covalentemente.

En ambos casos, estos dos picos corresponden a la primera de hidratación, esta información es consistente con la información descrita en la siguiente sección.

Notemos que para la adsorción en SAu, en la RDF de la figura IV.6.1.a (negro) se observan las primeras interacciones a ~ 0.225 nm entre el RBD y SAu, esta función presenta un primer hombro en el rango 0.26-0.30 nm y un pico más alto en el intervalo 0.31-0.37 nm. En SCu, la RDF de la figura IV.6.1.b (negro) muestra que la primera interacción RBD-superficie ocurre a ~ 0.20 nm, la RDF tiene

un primer hombro en el rango 0.265-0.28 nm y un pico más pronunciado en el rango 0.31-0.38 nm. En SC nuevamente no es posible apreciar estas primeras interacciones debido a que la contribución de los otros átomos de carbono unidos al carbono de referencia están apantallando las correlaciones por pares respecto a los átomos del segundo grupo, por lo que no es posible apreciar la primera capa de hidratación a partir de su RDF. Por este motivo, en la siguiente sección solo en este caso se utilizará el perfil de densidad parcial para caracterizar las primeras interacciones RBD-superficie, así como la primera capa de hidratación.

Para las superficies metálicas, la adsorción del RBD (negro) y la primera capa de hidratación (rojo) ocurren en el mismo valor (ver fig. IV.6.1.a y fig. IV.6.1.b), esto sugiere una orientación u ordenamiento preferencial de los residuos del RBD en esa región, influenciada tanto por la interacción con la superficie correspondiente, como por la primera capa de hidratación.

De acuerdo con este análisis, las interacciones RBD-superficie, se presentan en el intervalo 0-0.35 nm, en esta región es posible monitorear la adsorción de los residuos del RBD en cada superficie.

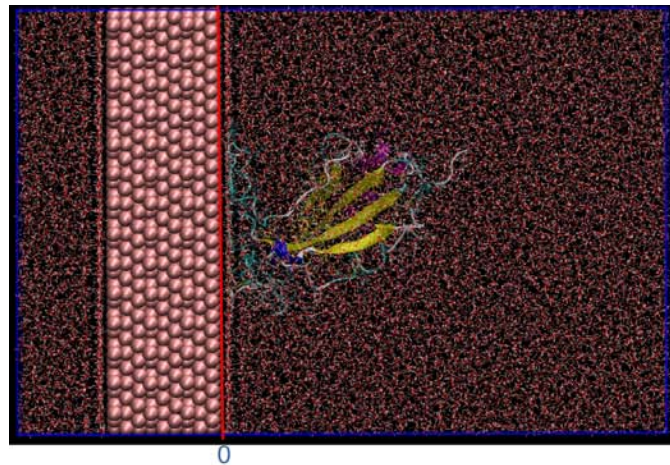
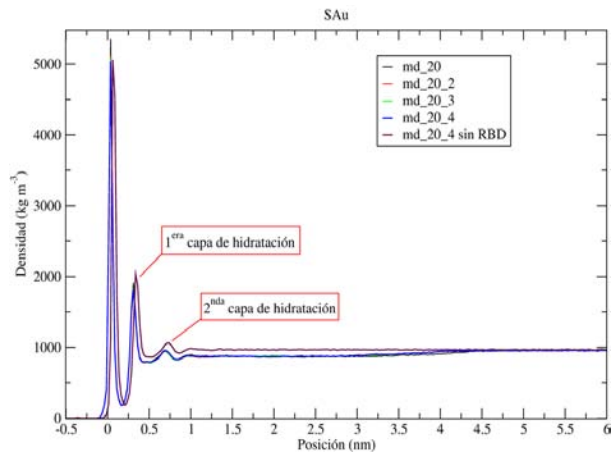
IV.6 Perfil de densidad parcial

El perfil de densidad parcial proporciona la distribución de moléculas o átomos de interés al dividir la caja de simulación en rebanadas de un determinado ancho a lo largo de un eje axial en la caja de simulación y considerando un punto como el origen.

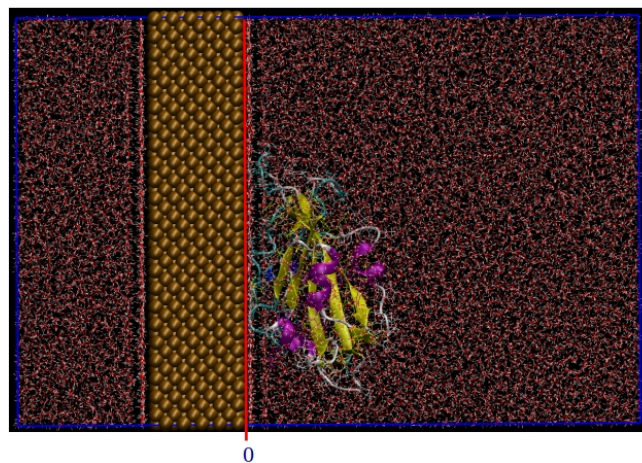
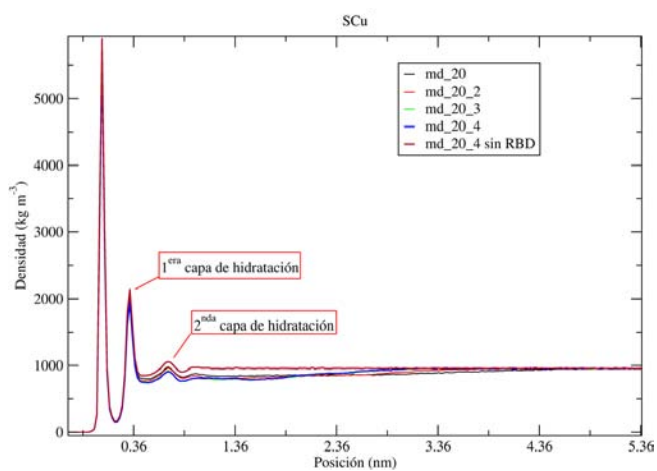
Si el arreglo de las moléculas de agua alrededor de una superficie o proteína sigue un patrón regular, el perfil de densidad exhibirá picos en función de la posición de las moléculas de agua, indicando el ancho de cada pico que está asociado a regiones con moléculas de agua que presentan una distribución regular en la caja de simulación, a dichas regiones se les conoce como capas de hidratación/solvatación o esferas de coordinación.

En el caso presentado en este trabajo, el perfil de densidad parcial se calculó tomando como origen la superficie del sistema (RBD-superficie) en la caja de simulación (ver fig. IV.5.1) para las moléculas de agua que conformaron el disolvente en cada sistema: 21,775 moléculas de agua para RBD-SAu (IV.5.1.a); 22,383 moléculas de agua para RBD-SCu (IV.5.1.b) y 21,439 moléculas de agua para RBD-SC (IV.5.1.c). Así como para los sistemas superficie sin RBD en la caja de simulación y el perfil de densidad parcial del RBD en el sistema RBD-SC. En todos los casos, la caja fue dividida en 500 rebanadas a lo largo del eje z, cada rebanada tiene un ancho de 0.025 nm debido a sus dimensiones (8.19 x 8.19 x 12.5 nm).

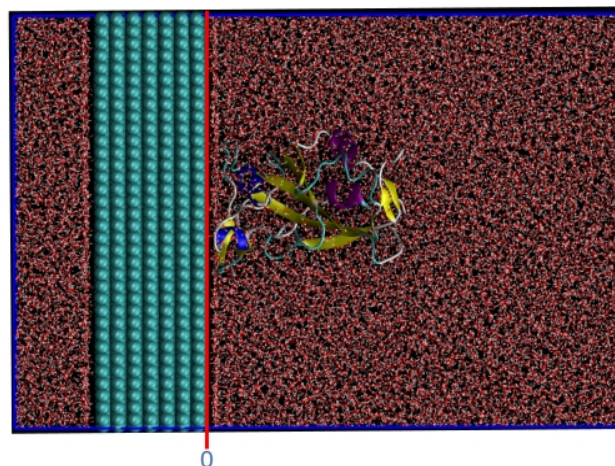
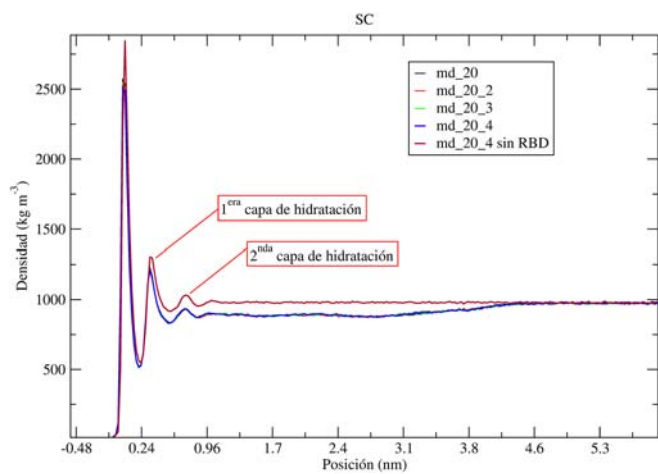
En los tres casos (RBD-superficie), la densidad del bulto de agua fluctúa alrededor de un valor fijo para una distancia mayor a 3.8 nm desde la superficie. El primer pico de izquierda a derecha en la figura IV.5.1 corresponde al cambio de densidad de la superficie con respecto a la del agua, la presencia del segundo y tercer pico, que corresponden a la primera y segunda capa de hidratación, indican una rápida reestructuración de las moléculas de agua desde el primer bloque de simulación (md_20) y mantienen un ordenamiento invariable en los bloques posteriores (md_20_2, md_20_3 y md_20_4) de la correspondiente interfase, comportamiento que también fue observado en los tres sistemas (superficie sin RBD).



(a)



(b)



(c)

Figura IV.5.1. Perfil de densidad a lo largo del eje z en la caja de simulación con el RBD solvatado en agua y las superficies: SAu (a), SCu (b) y SC (c).

En la figura IV.5.1.a, la primera y segunda capa de hidratación del sistema RBD-SAu se observan a ~ 0.33 nm con un ancho de ~ 0.13 nm y ~ 0.65 nm con un ancho de ~ 0.27 nm de SAu, respectivamente. En la figura IV.5.1.b, la primera y segunda capa de hidratación del sistema RBD-SCu corresponden en ~ 0.32 nm con un ancho de ~ 0.13 nm y en ~ 0.68 nm, con un ancho de ~ 0.25 nm de SCu, en ese orden. En la figura IV.5.1.c, las dos capas de hidratación del sistema RBD-SC corresponden a ~ 0.32 nm con un ancho de ~ 0.18 nm y ~ 0.70 nm con un ancho de ~ 0.22 nm de SC, en el mismo orden. Existe una separación de ~ 0.39 nm, ~ 0.37 nm y ~ 0.38 nm entre el máximo de ambas capas de hidratación en SAu, SCu y SC, respectivamente.

Para la adsorción sobre SC, en el perfil de densidad parcial del RBD de la figura IV.5.2, se observan las primeras interacciones RBD-SC a ~ 0.23 nm, lo que completa la caracterización de la sección anterior.

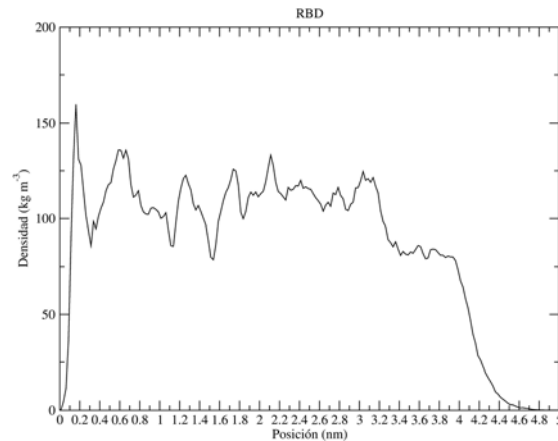


Figura IV.5.2. Perfil de densidad del RBD a lo largo del eje z en la caja de simulación para el sistema RBD-SC.

Retomando la descripción anterior de la figura IV.5.1, tanto la primera como la segunda capa de hidratación se configuran a distancias estimadas muy similares para los tres casos (RBD-superficie), pudiendo asumir que la naturaleza de las interacciones superficie-agua son semejantes.

La configuración de las moléculas de agua interfaciales es ligeramente perturbada por la presencia del protómero RBD, como se muestra en la figura IV.5.1, donde el perfil de densidad parcial del agua sobre cada superficie en presencia del RBD en el último bloque (md_20_4) es comparado con el perfil de densidad parcial del agua sobre cada superficie en ausencia del RBD en el último bloque (md_20_4 sin RBD).

De acuerdo con la figura IV.5.1, la primera capa de hidratación es más densa para SCu (ver fig. IV.5.1.a) con respecto a SAu (ver fig. IV.5.1.b) y SC (IV.5.1.c) con la menor densidad. En los tres casos, la segunda capa de hidratación presenta una densidad mucho menor que la primera. Esto se puede ver más claramente en la figura IV.5.3.a en SAu, en la figura IV.5.3.b en SCu y en la figura IV.5.3.c en SC.

En términos de hidrofobia o hidrofilia, SC mostró la naturaleza más hidrofóbica al mostrar una mayor distancia entre la superficie y la primera capa de hidratación, lo que puede interpretarse como mayor repulsión superficie-agua, misma que se vio reflejada en el valor más pequeño de la densidad. Por otra parte, SCu mostró la primera capa de hidratación con la mayor densidad.

En relación con lo discutido en secciones pasadas, el hecho de que el protómero RBD haya tenido un mayor desplegamiento y una mayor adsorción sobre SAu que sobre SC, como se verá más adelante, puede interpretarse como una mejor interacción que los residuos en contacto con la superficie de SC que tienen una naturaleza con carácter hidrofílico predominante, ya que una superficie hidrofóbica tiene más puntos accesibles para el adsorbato, por lo que de ser el adsorbato de naturaleza predominantemente hidrofóbica, esta superficie sería aquella sobre la cuál la adsorción habría sido la mayor, lo cual está en desacuerdo con lo observado. Estos datos se complementan con la siguiente sección que también está relacionada con la distribución de átomos en el sistema.

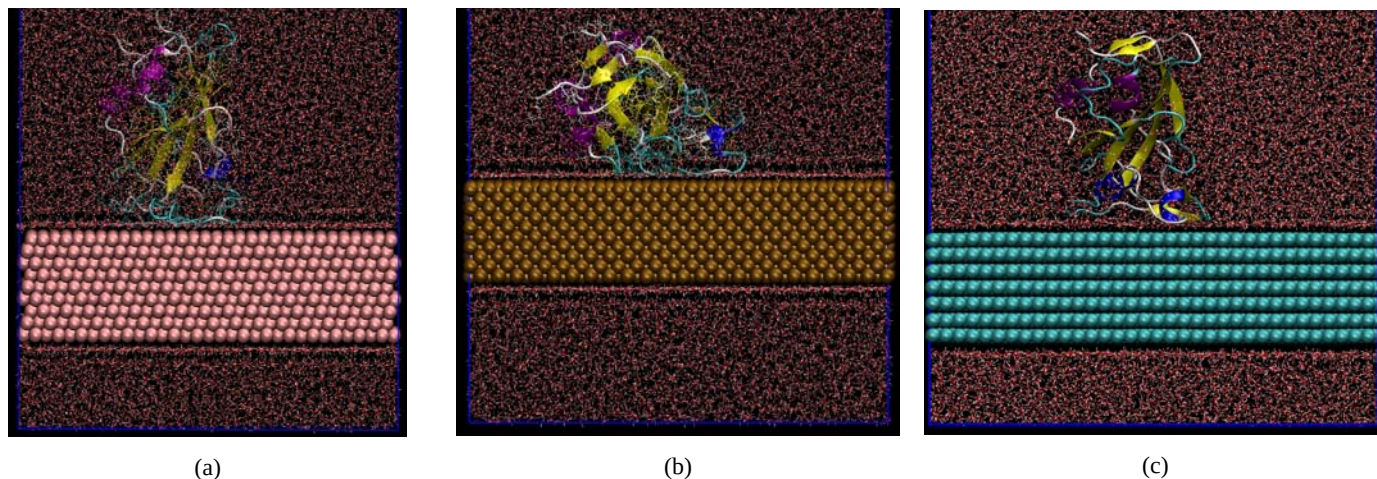


Figura IV.5.3. Primera y segunda capa de hidratación en la caja de simulación de (a) SAu, (b) SCu y (c) SC.

IV.7 Frecuencia de residuos en contacto

El proceso de la adsorción del RBD brinda información de los residuos del RBD más cercanos a la superficie. Dicho fenómeno estará favorecido cuando exista mayor adherencia por parte del adsorbato, o menos favorecido en el caso de una menor adherencia. La adsorción estará influenciada por la naturaleza polar de los residuos del protómero más próximos a la superficie con respecto a la naturaleza polar de la superficie al estar ambos en contacto con las moléculas de agua en el disolvente, por ende, este análisis permite identificar los residuos del RBD con mayor frecuencia de contactos en la interacción con cada superficie, respectivamente. En el proyecto que aquí se estudia, la adsorción fue monitoreada para la identificación de contactos directos entre ambos grupos. Se consideró un contacto directo cuando la distancia entre cualquier residuo del RBD y un átomo de la superficie era menor a 0.35 nm.

En la figura IV.7.1 se muestra la gráfica del número de contactos entre los residuos del RBD y cada una de las superficies analizadas: SAu (negro), SCu (rojo) y SC (verde) dentro de los 0.35 nm mencionados anteriormente.

Residuos en contacto < 3.5 Å

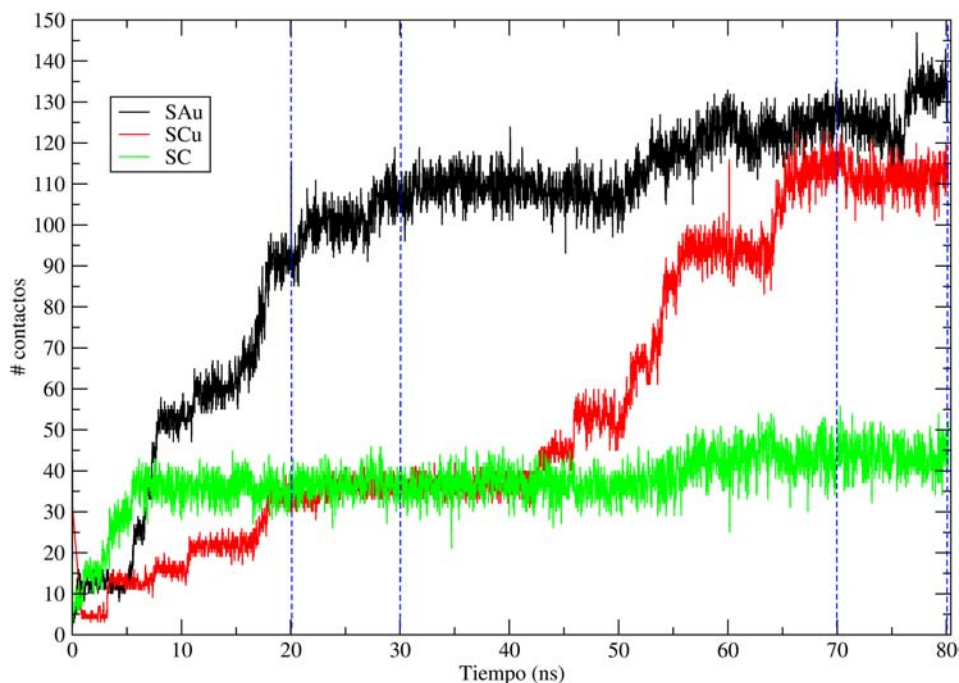


Figura IV.7.1. Gráfica de contactos RBD-superficie dentro de 0.35 nm para SAu (negro), SCu (rojo) y SC (verde).

Para SAu, en $t \sim 5$ ns el número de contactos se incrementa drásticamente hasta 55 contactos y continúa creciendo hasta $t \sim 20$ ns donde el número de contactos comienza a estabilizarse, manteniendo esa tendencia hasta $t \sim 40$ ns, a partir de este punto ocurre una ligera reducción en el número de contactos en el intervalo de $t \sim 40$ -50 ns, alcanzando un valor de 130 contactos en $t \sim 60$ ns y oscilando alrededor del mismo valor hasta $t \sim 75$ ns donde finalmente el número de contactos fluctúa de forma constante el resto de la simulación.

En SCu, la tendencia exhibe un crecimiento escalonado desde $t \sim 3$ ns hasta $t \sim 20$ ns con ~ 38 contactos, mostrando una menor cantidad de contactos respecto a las otras dos superficies hasta este punto. Después, se observa una meseta desde $t \sim 20$ ns hasta $t \sim 44$ ns con ~ 40 contactos, punto en donde se observa un incremento hasta $t \sim 55$ ns con un valor ~ 95 contactos que continúa en el intervalo de $t \sim 55$ -65 ns. Por último, la gráfica oscila en torno a un valor estable de 110 contactos en el intervalo ($t \sim 70$ -80 ns).

En SC, a diferencia del caso de las superficies metálicas, la gráfica reporta un incremento hasta $t \sim 5$ ns con ~ 38 contactos, en el intervalo $t \sim 20$ -42 ns el número de contactos fluctúa alrededor del mismo valor que en SCu y se mantiene hasta $t \sim 55$ ns donde ocurre un ligero incremento a ~ 40 contactos, oscilando alrededor de ese valor el resto de la simulación.

De acuerdo con la descripción de la adsorción del RBD sobre SAu en el párrafo anterior, los residuos del RBD interactuaron con mayor frecuencia durante la adsorción. Para entender mejor la razón de esto, recordemos los análisis discutidos en la sección de perfil de densidad parcial y en la sección de la función de distribución radial de pares para los últimos 20 ns de la simulación, en donde vimos que las primeras interacciones entre RBD-SAú ocurrían alrededor de ~ 0.225 nm y que la primera capa de hidratación tenía una densidad intermedia que también se interpretaba como un grado intermedio de hidrofobicidad respecto al observado en las otras dos superficies. Además, la primera capa de

hidratación se encontraba en el mismo intervalo que el pico en la RDF de los grupos RBD-SAu, que sugería un ordenamiento preferencial de los residuos del RBD influenciado por las moléculas de agua.

Análogamente, la descripción de la adsorción del RBD sobre SCu mostró ser menor que la encontrada en SAu, aunque con incrementos más abruptos que los observados en SAu, cuya explicación está asociada con la orientación horizontal del RBD adquirida durante el intervalo de $t \sim 40-60$ ns.

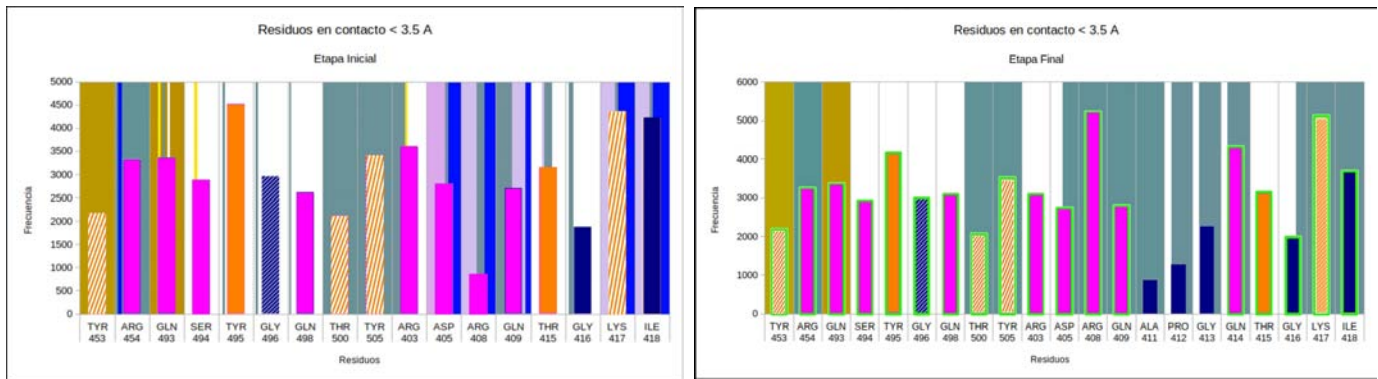
Como en la adsorción de SAu, los análisis discutidos en la sección de perfil de densidad parcial y en la sección de la función de distribución radial para los últimos 20 ns, permiten justificar la adsorción del RBD sobre esta superficie. Las primeras interacciones entre RBD-SCu ocurrían alrededor de ~ 0.20 nm, la primera capa de hidratación tenía la mayor densidad, lo que se traducía como la superficie con menor carácter hidrofóbico en comparación con las otras dos superficies. Por lo que al haber más moléculas de agua en la primera capa de hidratación, ubicada en el mismo intervalo que el pico en la RDF de los grupos RBD-SCu, el ordenamiento de los residuos del RBD que conllevaran a una orientación geométrica preferencial fue menos favorecida, razón por la cual el número de contactos en la figura IV.7.1 fue menor durante la adsorción.

En el tercer caso, la descripción de la adsorción sobre SC mostró un comportamiento monótono en el número de contactos entre RBD-SC, así como la superficie con menor número de contactos al interactuar con los residuos del RBD. La justificación de este comportamiento está asociada con la orientación vertical del RBD que predominó durante la simulación.

Como en los casos anteriores, el correspondiente perfil de densidad parcial permite complementar la información sobre la adsorción para comprender mejor la interacción del RBD con la SC. Las primeras interacciones entre RBD-SC ocurrían alrededor de ~ 0.23 nm (ver fig. IV.5.2), la primera capa de hidratación tenía la menor densidad, lo que se interpretaba como la superficie con mayor carácter hidrofóbico en comparación con las otras dos superficies. En consecuencia, al haber un menor número de moléculas de agua en la primera capa de hidratación, (ubicada en el mismo intervalo que el perfil de densidad parcial del RBD-SC) la superficie tendría disponible un mayor número de sitios accesibles para su adsorción; sin embargo, el hecho de que el número de contactos en la figura IV.7.1 haya sido el menor de los tres estudiados, sugiere que la escasa orientación geométrica por parte de los residuos del RBD es debido a la naturaleza polar de los mismos, como se verá más adelante.

Con la finalidad de discutir estos resultados en mayor detalle, el proceso de adsorción se dividió en dos etapas, definidas en intervalos que mostraran un comportamiento estable. La etapa inicial correspondiente al contacto del RBD con cada superficie, quedó definida en el intervalo $t \sim 20-30$ ns y la etapa final quedó definida en el intervalo $t \sim 70-80$ ns, posterior a los cambios estructurales observados en el RBD sobre cada superficie, como se indica en la figura IV.7.1.

En la figura IV.7.2 se muestra la gráfica de barras de la frecuencia de contactos de los residuos involucrados durante la etapa inicial (ver fig. IV.7.2.a) y final (ver fig. IV.7.2.b) del proceso de adsorción para SAu. Clasificados según su naturaleza polar en: hidrofílicos (rosa), hidrofóbicos (azul marino) y anfipáticos (naranja)[50]. El color del fondo en cada residuo corresponde al código de colores discutido anteriormente para clasificar las estructuras secundarias a las que los residuos pertenecen. El análisis de la estructura secundaria de los residuos se presentará en una sección posterior. Las barras resaltadas con un borde verde en la etapa final representan los residuos encontrados también en la etapa inicial del proceso y aquellas barras con un patrón rayado pertenecen a residuos relevantes en el vínculo con la célula humana receptora ACE2, que se encuentran presentes en nuestro modelo del RBD. Estos residuos son de especial interés en la inhibición de la función principal del RBD, por lo que se analizarán más adelante.



(a)

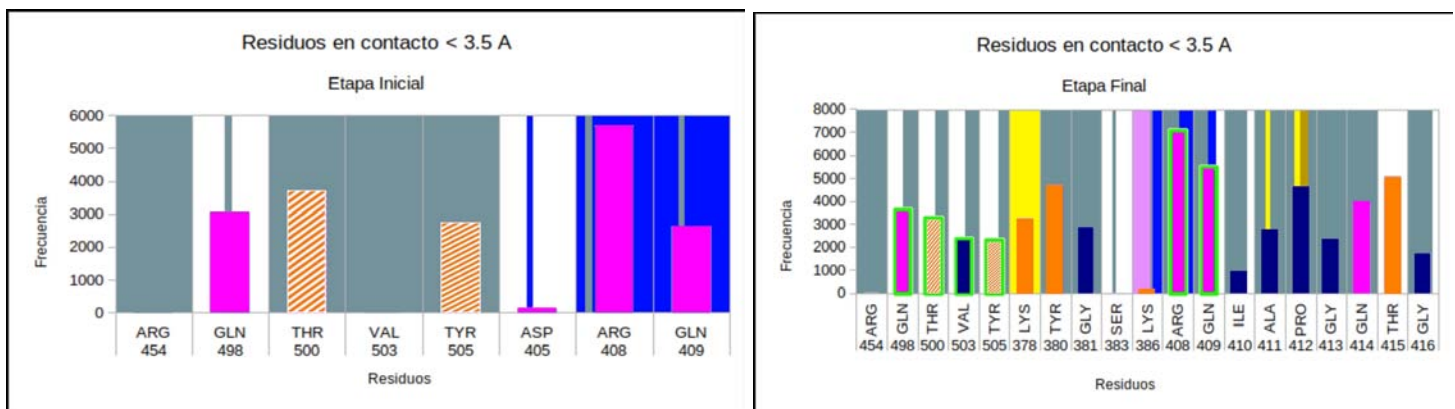
(b)



Figura IV.7.2. Gráfica de barra de la frecuencia de contactos de los residuos: hidrofóbicos (azul marino), hidrofílicos (rosa) y anfipáticos (naranja) con SAu dentro de 0.35 nm en la etapa inicial (a) y final (b) del proceso de adsorción.

En SAu, la adsorción en la etapa inicial (ver fig. IV.7.2.a) mostró 17 residuos en contacto, mayormente dominada por aminoácidos de naturaleza hidrofílica (8 aminoácidos) y los aminoácidos Lys 417, Ile 418 y Tyr 495, como los tres con mayor frecuencia de contactos. En la etapa final (IV.7.2.b) el número de residuos en contacto se incrementó en cuatro y continuó la tendencia mayoritaria de aminoácidos hidrofílicos (9 aminoácidos), con Arg408, Gln414 y Lys417 aquellos con mayor frecuencia de contactos.

En la figura IV.7.3 se muestra la gráfica de barras de la frecuencia de contactos de los residuos involucrados durante la etapa inicial y final del proceso de adsorción para SCu, utilizando la misma clasificación que en la figura anterior.



(a)

(b)



Figura IV.7.3. Gráfica de barra de la frecuencia de contactos de los residuos: hidrofóbicos (azul marino), hidrofílicos (rosa) y anfipáticos (naranja) con SCu dentro de 0.35 nm en la etapa inicial (a) y final (b) del proceso de adsorción.

En SCu, la adsorción en la etapa inicial (ver fig. IV.7.3.a) tuvo menor contribución respecto a la etapa inicial de SAu, registrando 8 residuos en contacto, de los cuales 1 fue hidrofóbico y 5 hidrofílicos (Arg454, Arg408, Gln498, Gln409, Asp405). En la etapa final (ver fig. IV.7.3.b) se encontraron 19 residuos en contacto de los cuales 6 de ellos estuvieron en la etapa inicial y a diferencia de la etapa final en SAu, en este caso se presentó una menor diferencia entre la predominancia de aminoácidos de naturaleza hidrofóbica (7 aminoácidos) con respecto a los de naturaleza hidrofílica (6 aminoácidos), permaneciendo presentes los residuos Thr500 y Tyr505 del RBD-ACE2 encontrados en la etapa inicial.

En la figura IV.7.4 se muestra la gráfica de barras de la frecuencia de contactos de los residuos involucrados durante la etapa inicial y final del proceso de adsorción para SC, utilizando la clasificación previamente descrita.

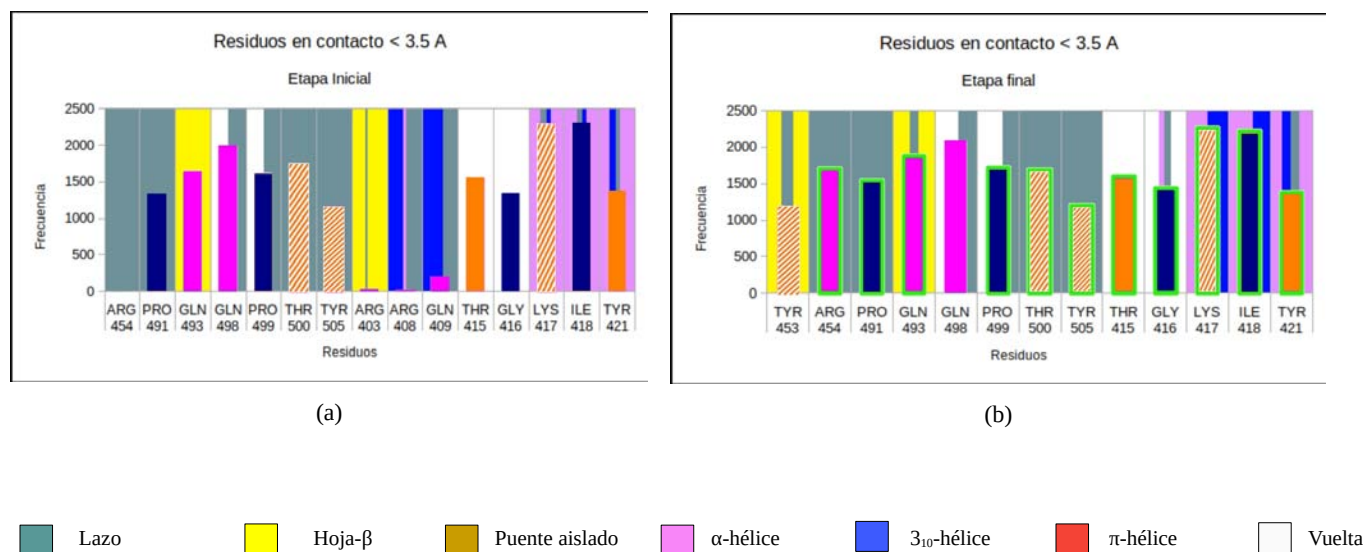


Figura IV.7.4. Gráfica de barra de la frecuencia de contactos de los residuos: hidrofóbicos (azul marino), hidrofílicos (rosa) y anfipáticos (naranja) con SC dentro de 0.35 nm en la etapa inicial (a) y final (b) del proceso de adsorción.

En SC, la adsorción en la etapa inicial (ver fig. IV.7.4.a) tuvo mayor contribución respecto a SCu pero menor contribución respecto a SAu, registrando 15 residuos en contacto, de los cuales 4 fueron hidrofóbicos, 6 hidrofílicos y los residuos clave (Tyr453, Tyr505, Thr500 y Lys417)¹. En la etapa final (ver fig. IV.7.4.b) se encontraron 13 residuos en contacto, de los cuales 11 estuvieron en la etapa inicial y predominaron los residuos anfipáticos (6 residuos) y tan solo 3 residuos hidrofílicos, dejando el resto dentro de la categoría hidrofóbica (4 residuos).

Estos resultados confirman las aseveraciones descritas en el párrafo final del análisis del perfil de densidad y aquellas realizadas al final del análisis de frecuencia de contactos de la figura IV.7.1, pues justifican el ordenamiento originado por la orientación preferencial de los residuos del RBD.

Por ejemplo, en SC explica la razón de la escasa orientación geométrica por parte de los residuos del RBD debido a la naturaleza hidrofílica en los residuos de contacto, así como por las predominantes contribuciones de aquellos residuos anfipáticos cuyo carácter hidrofílico podría estar contribuyendo a que no ocurra la adsorción, mientras que el carácter hidrofóbico del mismo podría estar contribuyendo a la adsorción como se había venido sospechando. De igual manera, la presencia predominante de residuos hidrofóbicos y anfipáticos es la responsable de desfavorecer la adsorción del RBD en la SCu. Eso también explica el balance entre residuos hidrofílicos e hidrofóbicos presentes durante la adsorción

¹ Estos cuatro presentes en el vínculo RBD-ACE2.

sobre SAu permitió que la orientación geométrica fuera favorecida y que el RBD se adsorbiera en mayor medida que en las otras dos superficies.

La adsorción de proteínas en superficies sólidas es irreversible y por lo tanto conduce a la inmovilización de proteínas en la fase de la superficie dado que ya no son libres de difundirse[51]. Esto es muy importante en este trabajo, dado que un cambio estructural podría conllevar a la pérdida en el funcionamiento biológico del RBD para el vínculo con la célula receptora humana ACE2.

De acuerdo con la figura IV.7.2 en RBD-SAu hay abundancia de arginina (3 residuos), al ser de los residuos que más se repiten. La frecuencia de contactos también es un indicador de la fuerte adsorción de los residuos sobre la SAu. Los residuos Lys e Ile fueron aquellos con la mayor frecuencia de contactos y que, junto con Arg fueron los que se conservaron en ambas etapas de la adsorción. En estudios teóricos de simulaciones de MD de polipéptidos en SAu, la adsorción más fuerte sobre SAu ha sido encontrada en los residuos: arginina (Arg), treonina (Thr), serina (Ser), prolina (Pro) e isoleucina (Ile)[52]. Por otra parte, la glicina (Gly) ha mostrado tener una débil adsorción sobre SAu[53].

Este resultado es consistente con lo encontrado en estudios relacionados donde una respuesta que se ha observado anteriormente para polipéptidos, es que la principal contribución de residuos adsorbidos sobre SAu proviene de aminoácidos polares más que de aminoácidos hidrofóbicos, donde aún con una composición balanceada entre residuos polares e hidrofóbicos, los átomos en contacto cercano con SAu han sido principalmente aquellos polares de las cadenas laterales, como ocurriría también con aquellos de naturaleza anfipática (Lys)[54].

Algunos de los residuos que presentan la adsorción más fuerte sobre la SAu en el RBD no están presentes en nuestro modelo, por este motivo, no se percibe su presencia en la adsorción. Otra explicación para esta ausencia, está asociada con los residuos de los alrededores que previenen, a los residuos de fuerte adsorción generar contactos con la SAu.

Además, los residuos: Tyr453, Gly496, Thr500, Tyr505 y Lys417 del RBD se hicieron presentes en ambas etapas de adsorción. Como se mencionó en la introducción, son residuos que participan en las interacciones específicas entre el RBD y el receptor ACE2 durante la fase inicial de infección de la célula humana[5].

Algunos estudios de simulaciones de MD sugieren que diversas proteínas no se adsorben fácilmente sobre una SCu[53]. En otros, sus resultados indican que la adsorción de un péptido (conformado por 14 residuos, incluyendo seis glicinas (Gly), dos lisinas (Lys) y seis histidinas (Hys) al final del mismo) sobre una SCu es fuerte para histidina (Hys) y presenta una débil adsorción para los residuos glicina (Gly) y lisina (Lys)[55]. Experimentalmente y en simulaciones computacionales de DFT la Arg ha presentado una fuerte adsorción sobre SCu[56].

En acuerdo con las observaciones anteriores, la figura IV.7.3 muestra que Arg fue uno de los residuos más repetidos y con mayor frecuencia de contactos durante ambas etapas de adsorción. De igual forma, los aminoácidos Gly y Lys tuvieron un número similar que los correspondientes a Arg, aunque solo durante la etapa final de adsorción y cuya reducida frecuencia de contactos confirma su reducida capacidad para adsorberse sobre SCu.

Como se mencionó anteriormente sobre SC, los residuos hidrofóbicos estuvieron más presentes durante la etapa final, mientras que los residuos hidrofílicos predominaron durante la etapa inicial (ver fig. IV.7.4). Algunos trabajos resaltan la importancia de efectos cooperativos para conducir residuos hidrofílicos hacia SC como ocurre con aquellos hidrofóbicos. En general estos últimos muestran interacciones más fuertes con la superficie hidrofóbica[57]. En conformidad con esto y aunque residuos

hidrofóbicos como Pro e Ile fueron dos de los más repetidos durante ambas etapas de adsorción, la dominante presencia de residuos anfipáticos (Lys, Tyr, Thr), en conjunto con residuos hidrofílicos (Gln, Arg) que presentaron una alta frecuencia de contactos impidieron que la adsorción fuera favorecida.

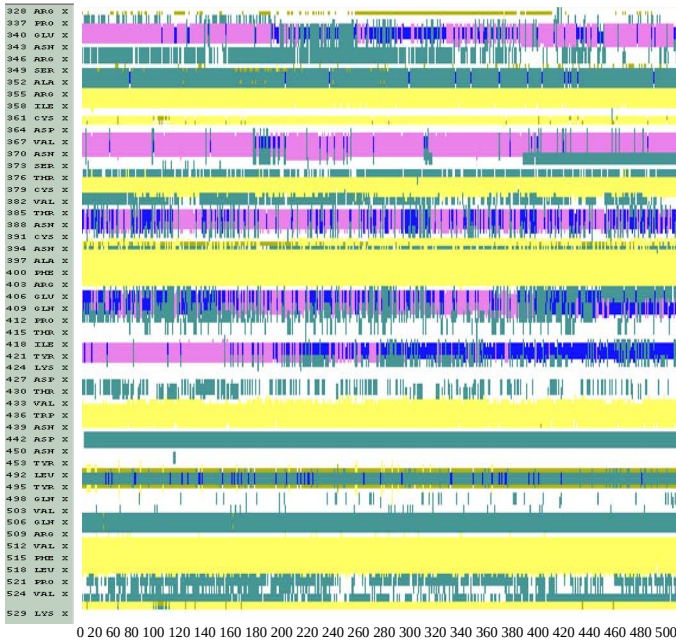
Otros trabajos indican que la adsorción sobre SC es más fuerte en presencia de triptófano (Trp)[57].

IV.8 Línea temporal de la estructura secundaria

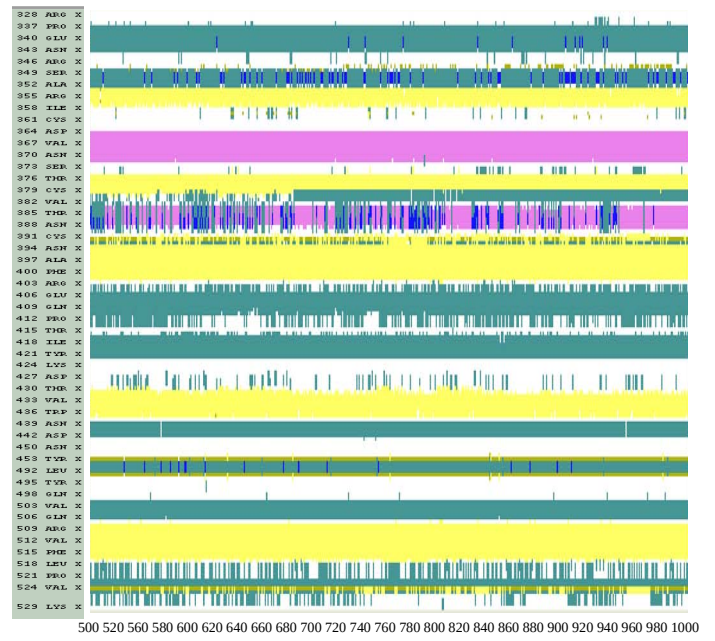
El cambio en la estructura secundaria del RBD es un indicador relevante en el proceso de plegamiento o desplegamiento del RBD. La asistencia de la línea temporal en VMD permitió analizar el cambio en las estructuras secundarias del RBD, una herramienta que se encuentra basada en el algoritmo STRIDE[47], que combina la información de la energía de los enlaces de hidrógeno con información derivada de los ángulos de torsión de la cadena principal del protómero.

En la figura IV.8.1 se presenta la variación de la estructura secundaria en el tiempo para el protómero RBD sobre SAu y para el RBD sobre SCu durante la etapa inicial y final de adsorción, respectivamente. La figura IV.8.2 presenta la variación de la estructura secundaria en el tiempo para el protómero RBD sobre SC y para el RBD aislado durante la etapa inicial y final de adsorción, respectivamente. Durante las dos etapas de adsorción en las tres superficies, las figuras IV.7.2, IV.7.3 y IV.7.4 muestran que los residuos involucrados en la adsorción que fueron en su mayoría asociados a estructuras secundarias de tipo lazo y vueltas (con excepción de Tyr453 mayormente ubicado en hoja- β y Lys417 mayormente ubicado en una α -hélice). Los dos primeros residuos (Tyr 453 y Lys417) tienen posiciones más restringidas en comparación con el resto de los residuos ubicados en estructuras secundarias de tipo lazo y vueltas que tienen posiciones sin restricciones y que consecuentemente son de las partes más flexibles en la estructura de una proteína como se verá en la última sección. Dependiendo de esta flexibilidad intrínseca la transición entre estructuras secundarias permitirá convertir a las estructuras como las α -hélices y 3_{10} -hélices. Es ampliamente aceptado que las transiciones de α -hélices a 3_{10} -hélices son pasos importantes en el plegamiento y desplegamiento de hélices[58]. Este tipo de estructuras secundarias son comunes de encontrar en la superficie de proteínas y típicamente sobresalen en el disolvente. En consecuencia, proporcionan sitios convenientes para el reconocimiento de proteínas o para la unión de ligandos[50]. Un hecho que se puede corroborar al observar que casi en su totalidad los residuos involucrados en las interacciones con el receptor humano ACE2 se encuentran en este tipo de estructuras secundarias.

RBD-SA_u

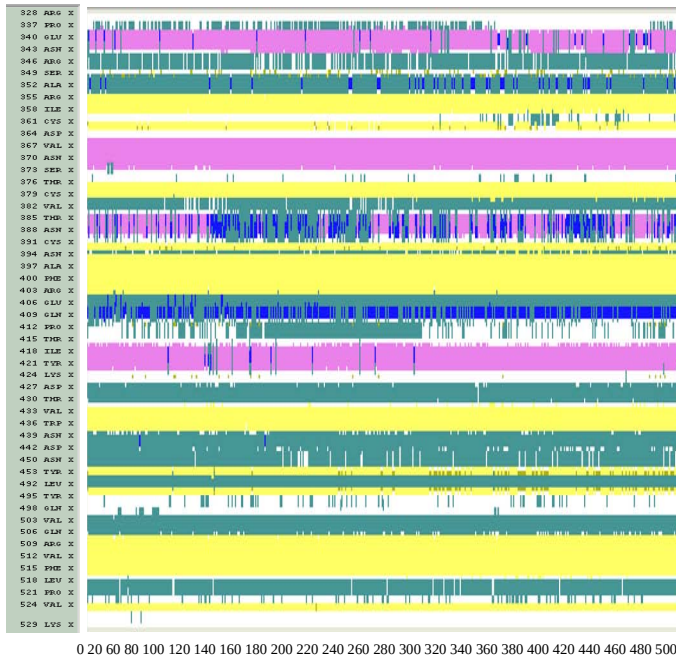


(a)

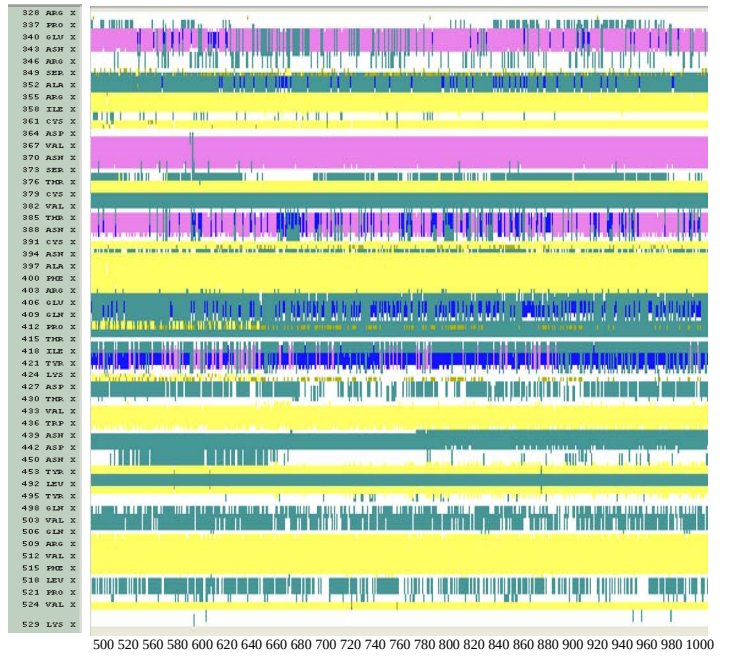


(b)

RBD-SC_u



(a)

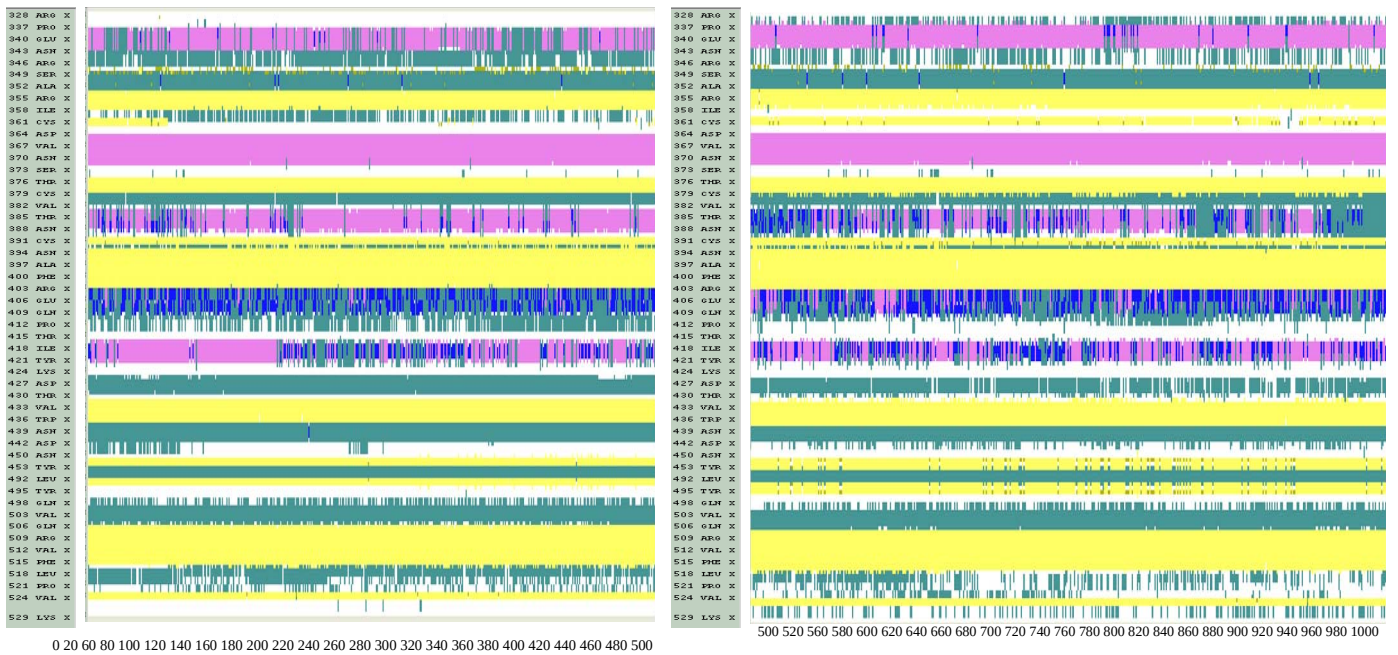


(b)

Lazo
 Hoja-β
 Puente aislado
 α-hélice
 3₁₀-hélice
 π-hélice
 Vuelta

Figura IV.8.1. Evolución temporal de la estructura secundaria del RBD sobre SA_u y del RBD sobre SC_u durante la etapa inicial (a) y final (b) de adsorción, respectivamente. El eje horizontal representa el tiempo y el eje vertical representa la secuencia de residuos.

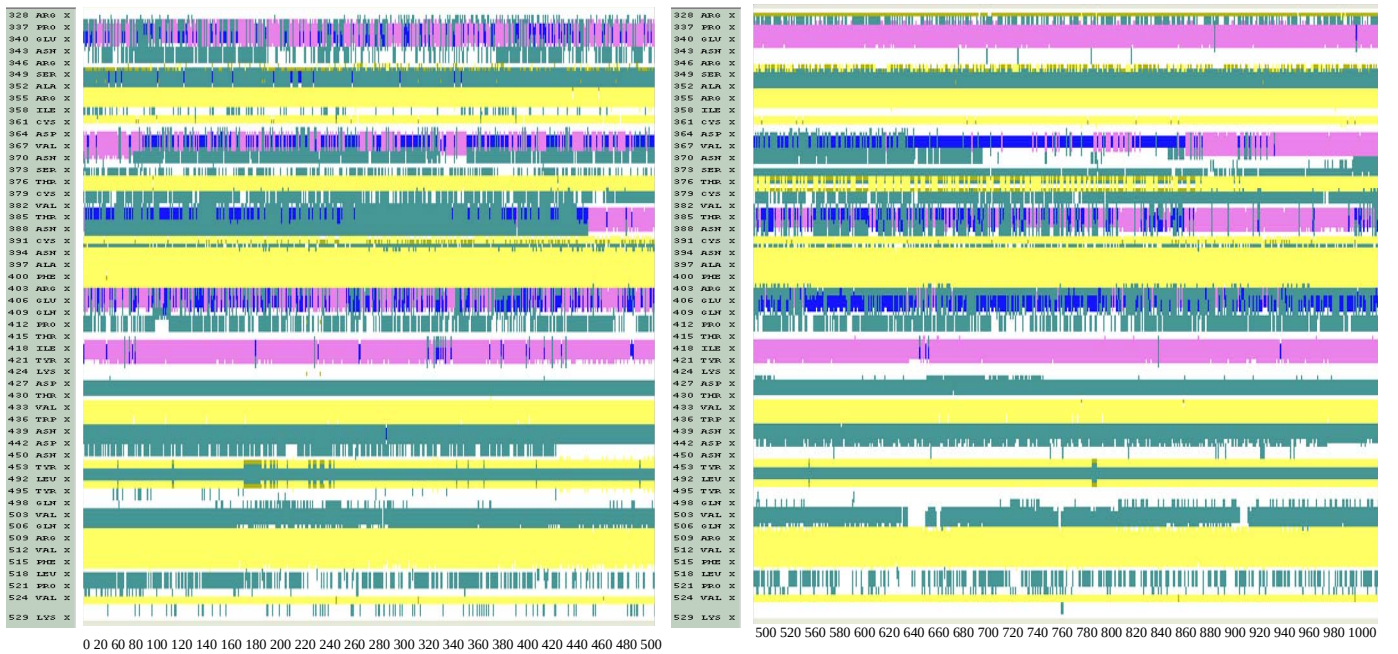
RBD-SC



(a)

(b)

RBD



(a)

(b)

Lazo
 Hoja-β
 Puente aislado
 α-hélice
 3₁₀-hélice
 π-hélice
 Vuelta

Figura IV.8.2. Evolución temporal de la estructura secundaria del RBD sobre SC y del RBD aislado durante la etapa inicial (a) y final (b) de adsorción, respectivamente. El eje horizontal representa el tiempo y el eje vertical representa la secuencia de residuos.

La estructura secundaria del RBD denota abundantes cambios como puede apreciarse en las figuras IV.8.1 y IV.8.2, no solo durante su relajación en agua sino también durante el proceso de adsorción sobre las tres superficies. Esto se ha observado anteriormente en estudios teóricos de la adsorción del RBD sobre superficies como: grafito, celulosa y grafeno[[8], [2]].

En la tabla IV.8.1 se compara el contenido y composición de las estructuras secundarias de tipo hélices y hojas- β del RBD en su conformación inicial con las correspondientes estructuras en la conformación final del RBD aislado y del RBD al interactuar con las tres superficies.

La estructura terciaria del RBD en su conformación inicial está principalmente integrada por dos α -hélices, tres 3_{10} -hélices, una hoja- β de 3 hebras, una hoja- β de 2 hebras y una hoja- β de 5 hebras.

Durante la adsorción sobre las tres superficies, la α -hélice (2: Lys417-Ile418-Ala419-Asp420-Tyr421-Asn422) desaparece y se convierte en un lazo, mientras que una nueva α -hélice (3: Tyr365-Ser366-Val367-Leu368-Tyr369-Asn370-Ser371-Ala372) aparece. La adsorción sobre las superficies metálicas produce la creación de una nueva α -hélice (4: Pro384-Thr385-Lys386-Leu387-Asn388-Asp389), estructura que aparece nuevamente en la adsorción sobre SCu pero con un residuo menos de diferencia (Pro384-Thr385-Lys386-Leu387-Asn388).

La situación es distinta con el RBD aislado relajado en agua. En concordancia con el RBD adsorbido sobre las tres superficies, se origina la α -hélice (3: Tyr365-Ser366-Val367-Leu368-Tyr369-Asn370) también y se conservan las dos α -hélices (1: Phe338-Gly339-Glu340-Val341-Phe342-Asn343 y 2: Lys417-Ile418-Ala419-Asp420-Tyr421-Asn422) presentes en la conformación inicial de RBD.

Por lo que de la comparación del RBD adsorbido sobre las tres superficies con su análogo relajado en agua, se puede concluir que la adsorción sobre las tres superficies es la responsable de la desaparición de la α -hélice (2: Lys417-Ile418-Ala419-Asp420-Tyr421-Asn422) y que la adsorción sobre las superficies: SCu y SC es la causante de la estabilización de la α -hélice (1: Phe338-Gly339-Glu340-Val341-Phe342-Asn343), estructura que también aparece en la conformación del RBD relajado en agua. La nueva α -hélice (3: Tyr365-Ser366-Val367-Leu368-Tyr369-Asn370) se origina de la relajación en el agua más que por la adsorción sobre las superficies, aunque sufre un ligero incremento de dos residuos en la longitud de su extensión (Tyr365-Ser366-Val367-Leu368-Tyr369-Asn370-Ser371-Ala372).

La hoja- β principal de 5 hebras es estabilizada tras la adsorción porque es preservada durante la relajación del RBD y la adsorción del RBD sobre las tres superficies.

En la conformación inicial del RBD, la hoja- β de 2 hebras (1: Gln493-Ser494 y Leu452-Tyr453) y la hoja- β de 3 hebras (1: Cys361-Val362, Cys391-Phe392 y Val524-Cys525) no son alteradas durante la relajación del RBD y la adsorción sobre SCu, mientras que se ven afectadas por la adsorción sobre las otras dos superficies, pasando de hoja- β de 2 hebras a puentes aislados (Gln493 y Tyr453) y de hoja- β de 3 hebras a puentes aislados (Val524 y Phe392) en la adsorción sobre SAu e incrementando levemente su longitud en la hoja- β de 2 hebras (1: Gln493-Ser494-Tyr495 y Tyr451-Leu452-Tyr453) y cambiando de hoja- β de 3 hebras a puentes aislados (Phe392) en la adsorción sobre SC.

Notemos que las 3_{10} -hélices aparecen como subestructuras transitorias durante las simulaciones, transformándose en lazos, vueltas o fragmentos extremales de α -hélices. Un ejemplo de esto fue la α -hélice (2: Lys417-Ile418-Ala419-Asp420-Tyr421-Asn422) en la conformación inicial del RBD que se mantuvo durante la relajación del RBD en agua pero que mostró este comportamiento transitorio a la 3_{10} -hélice (4: Ile418-Ala419-Asp420-Tyr421) durante su adsorción sobre SC. Otro ejemplo se observa en la 3_{10} -hélice (2: Leu387-Asn388-Asp389) en la conformación inicial del RBD y que se conserva durante la relajación en agua del RBD, pero que se transforma en la α -hélice (4: Pro384-Thr385-Lys386-Leu387-Asn388-Asp389 y 4: Pro384-Thr385-Lys386-Leu387-Asn388) durante la adsorción sobre SAu y SCu, respectivamente.

Una excepción es la 3₁₀-hélice (3: Val407-Arg408-Gln409) en la conformación inicial del RBD cuya estabilización es debida a la adsorción sobre SC como se aprecia en la tabla IV.8.1.

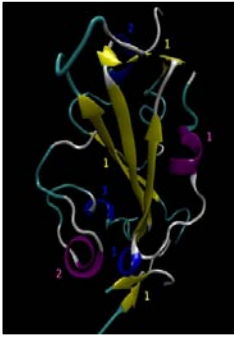
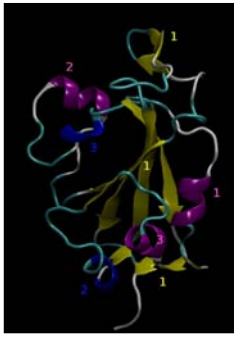
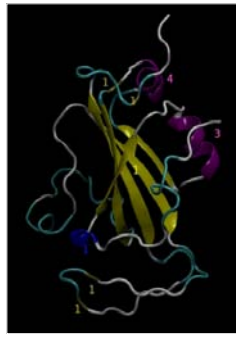

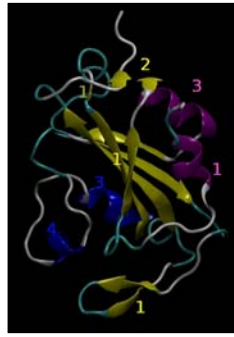
	RBD inicial	RBD final	RBD-SA _u	RBD-SC _u	RBD-SC
α-hélices	1: Phe338-Gly339-Glu340-Val341-Phe342-Asn343 2: Lys417-Ile418-Ala419-Asp420-Tyr421-Asn422	1: Phe338-Gly339-Glu340-Val341-Phe342-Asn343 3: Tyr365-Ser366-Val367-Leu368-Tyr369-Asn370 2: Lys417-Ile418-Ala419-Asp420-Tyr421-Asn422	3: Tyr365-Ser366-Val367-Leu368-Tyr369-Asn370-Ser371-Ala372 4: Pro384-Thr385-Lys386-Leu387-Asn388-Asp389	1: Phe338-Gly339-Glu340-Val341-Phe342-Asn343 3: Tyr365-Ser366-Val367-Leu368-Tyr369-Asn370-Ser371-Ala372 4: Pro384-Thr385-Lys386-Leu387-Asn388	1: Phe338-Gly339-Glu340-Val341-Phe342-Asn343 3: Tyr365-Ser366-Val367-Leu368-Tyr369-Asn370-Ser371-Ala372
3 ₁₀ -hélices	1: Val350-Tyr351-Ala352 2: Leu387-Asn388-Asp389 3: Val407-Arg408-Gln409	2: Leu387-Asn388-Asp389 3: Val407-Arg408-Gln409	1: Val350-Tyr351-Ala352	4: Ala419-Asp420-Tyr421 1: Leu425 (Puente aislado) 1: Pro412 (Puente aislado)	3: Gly404-Asp405-Glu406-Val407-Arg408-Gln409 4: Ile418-Ala419-Asp420-Tyr421
Hoja-β (3 hebras)	1: Cys361-Val362, Cys391-Phe392, Val524-Cys525	1: Cys361-Val362, Cys391-Phe392, Val524-Cys525	1: Val524 (Puente aislado) 1: Phe392 (Puente aislado)	1: Cys361-Val362, Cys391-Phe392, Val524-Cys525	Phe392(Puente aislado)
Hoja-β (2 hebras)	1: Gln493-Ser494, Leu452-Tyr453	1: Gln493-Ser494, Leu452-Tyr453 Leu335 (Puente aislado)	1: Gln493 (Puente aislado) 1: Tyr453 (Puente aislado)	1: Gln493-Ser494, Leu452-Tyr453	2: Cys361-Val362, Val524-Cys525 1: Tyr451-Leu452-Tyr453, Gln493-Ser494-Tyr495
Hoja-β (5 hebras)	1: Asn354-Arg355-Lys356-Arg357-Ile358-Ser359-Thr376-Phe377-Lys378-Cys379, Asn394-Val395-Tyr396-Ala397-Asp398-Ser399-Phe400-Val401-Ile402, Cys432-Val433-Ile434-Ala435-Trp436-Asn437, Tyr508-Arg509-Val510-Val511-Val512-Leu513-Ser514-Phe515-Glu516	1: Asn354-Arg355-Lys356-Arg357-Ile358, Thr376-Phe377-Lys378, Asn394-Val395-Tyr396-Ala397-Asp398-Ser399-Phe400-Val401-Ile402-Arg403, Cys432-Val433-Ile434-Ala435-Trp436-Asn437, Pro507-Tyr508-Arg509-Val510-Val511-Val512-Leu513-Ser514-Phe515-Glu516	1: Asn354-Arg355-Lys356-Arg357-Ile358, Thr376-Phe377-Lys378-Cys379, Asn394-Val395-Tyr396-Ala397-Asp398-Ser399-Phe400-Val401-Ile402, Gly431-Cys432-Val433-Ile434-Ala435-Trp436-Asn437, Tyr508-Arg509-Val510-Val511-Val512-Leu513-Ser514-Phe515-Glu516	1: Asn354-Arg355-Lys356-Arg357-Ile358, Thr376-Phe377-Lys378, Asn394-Val395-Tyr396-Ala397-Asp398-Ser399-Phe400-Val401-Ile402-Arg403, Cys432-Val433-Ile434-Ala435-Trp436-Asn437, Pro507-Tyr508-Arg509-Val510-Val511-Val512-Leu513-Ser514-Phe515-Glu516	1: Asn354-Arg355-Lys356-Arg357-Ile358, Thr376-Phe377-Lys378-Cys379, Asn394-Val395-Tyr396-Ala397-Asp398-Ser399-Phe400-Val401-Ile402-Arg403, Cys432-Val433-Ile434-Ala435-Trp436-Asn437, Pro507-Tyr508-Arg509-Val510-Val511-Val512-Leu513-Ser514-Phe515-Glu516
Ilustración					

Tabla IV.8.1. Estructuras secundarias presentes del RBD en su conformación inicial, del RBD aislado relajado en agua y del RBD en su configuración final posterior a su adsorción sobre las tres superficies.

IV.9 RMSF

La raíz de la fluctuación cuadrática media (RMSF: Root Mean Square Fluctuation) de la diferencia en las posiciones de cada residuo i de una configuración final con respecto a las posiciones de cada residuo i de una configuración inicial de referencia, es un parámetro que permite cuantificar los cambios en la flexibilidad de los residuos del RBD y que puede ser utilizado como un indicador de las regiones con los residuos más flexibles dentro del prótomo RBD como consecuencia de su relajación en un disolvente, o en un disolvente al interactuar con una superficie. Una fluctuación grande del RMSF en un residuo representa una alta flexibilidad en su estructura, mientras que una fluctuación pequeña del RMSF en un residuo representa una alta rigidez.

En la figura IV.9.1 se muestra el RMSF para la interacción del RBD con las tres superficies y del RBD aislado durante los intervalos mismos intervalos definidos para la etapa inicial y final de adsorción, correspondientemente.

En el RBD aislado, la figura IV.9.1 muestra que durante el intervalo de la etapa inicial de adsorción (ver fig. IV.9.2.1), la región con mayor flexibilidad pertenece al intervalo de residuos de 343 a 348 (pico a de color azul) que corresponden a la vuelta (Asn343-Ala344-Thr345-Arg346-Phe347-Ala348), mientras que en la etapa final de adsorción (ver fig. IV.9.2.2) en la figura IV.9.1, el dominio entre los residuos 443 y 451 (pico b de color azul) que corresponden a la vuelta (Ser443-Tyr449-Asn450-Tyr451) fue aquella con la mayor flexibilidad exhibida. La gráfica del RMSF del RBD al adsorberse sobre las tres superficies indica que la flexibilidad del RBD se vio alterada al interactuar sobre las tres superficies. En la etapa inicial de adsorción sobre SAu (ver fig. IV.9.2.3), la región del RBD con mayor flexibilidad fue en el intervalo de los residuos de 406 a 410 (pico a de color negro) contenidos en la 3_{10} -hélice (Glu406-Val407-Arg408) y el inicio de la vuelta (Gln409-Ile410), entre el intervalo de residuos de 336 a 338 (pico a de color rojo) en el final de la vuelta (Cys336-Pro337) y el inicio de un lazo en Phe338 (ver fig. IV.9.2.5) en la etapa inicial de adsorción sobre SCu y en el intervalo de los residuos de 343 a 348 (pico a de color verde), con Asn343 el residuo final de la α -hélice, perteneciente al lazo (Ala344-Thr345-Arg346-Phe347) y Ala348 en el inicio de un puente aislado (ver fig. IV.9.2.7), en la etapa inicial de adsorción sobre SC. Correspondientemente, en la etapa final de adsorción sobre SAu, la región del RBD con mayor flexibilidad fue entre los residuos 338 y 341 (pico b de color negro) pertenecientes al lazo (Phe338-Gly339-Glu340-Val341: ver fig. IV.9.2.4), entre los residuos de 336 a 341 (pico b de color rojo) que corresponden al final de la vuelta (Cys336-Pro337) y el inicio de la α -hélice (Phe338-Gly339-Glu340-Val341), en la etapa final de adsorción sobre SCu (ver fig. IV.9.2.6) y entre los residuos de 355 a 358 (pico b de color verde) contenidos en la hoja- β (Arg355-Lys356-Arg357-Ile358), en la etapa final de adsorción sobre SC (ver fig. IV.9.2.6.8).

Todas las gráficas del RMSF muestran un comportamiento anómalo en el intervalo de los residuos de 455 a 490 debido a que nuestro modelo de la estructura del RBD no tiene residuos en ese intervalo.

Además, todas las gráficas del RMSF sugieren que la flexibilidad de los residuos en el RBD se ve alterada de forma no uniforme, con particular énfasis en los prominentes picos descritos anteriormente que en su mayoría pertenecen a estructuras secundarias de tipo lazos y vueltas. En la literatura se ha visto que estructuras secundarias como las vueltas y lazos tienen una mayor flexibilidad que las estructuras secundarias que presentan un orden regular tales como hélices y hojas- β [59]. Por otro lado, la pérdida en la flexibilidad de la estructura del RBD aislado es más notoria en la etapa final de adsorción sobre las tres superficies, donde las gráficas muestran valores más pequeños del RMSF.

Simultáneamente, algunas regiones flexibles manifiestan cambios en la estructura secundaria y por lo tanto, contribuyen al RMSD global de la proteína, así como el incremento o disminución del radio de giro.

Los resultados proveen evidencia de que la interacción del RBD con estas superficies en este intervalo de simulación modifica la flexibilidad de los segmentos críticos para la interacción del RBD con la

célula receptora humana ACE2 y por tanto, irrumpe en la organización espacial atómica del esqueleto y la cadena lateral en los residuos clave, así como en su capacidad para adaptarse a su entorno, lo que podría conllevar a la pérdida de su función biológica principal.

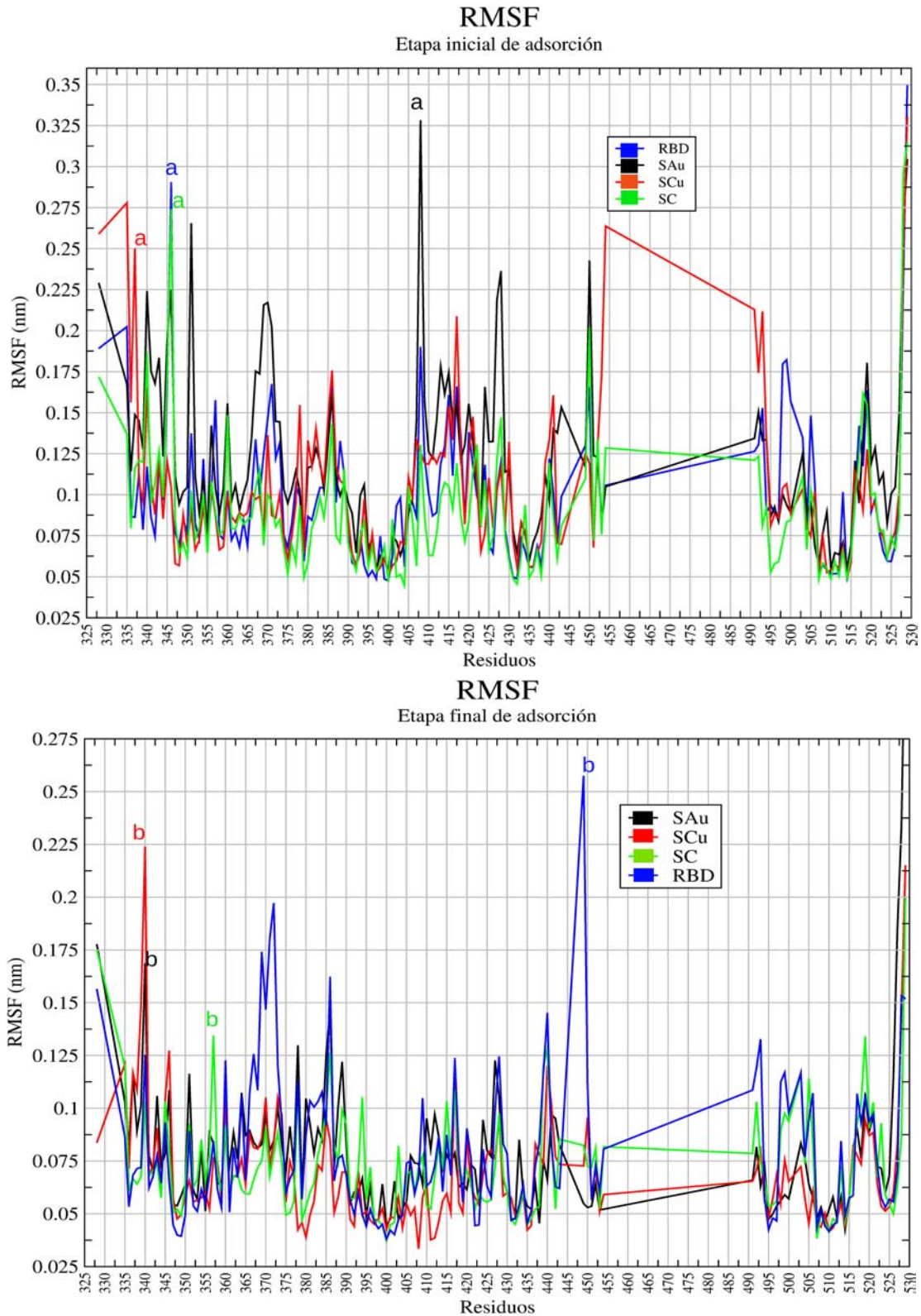


Figura IV.9.1. Gráfica del RMSF como función de los residuos del RBD de la simulación en las etapas inicial y final de adsorción del RBD sobre SAu (negro), del RBD sobre SCu (rojo), del RBD con SC (verde) y para el RBD aislado (azul).

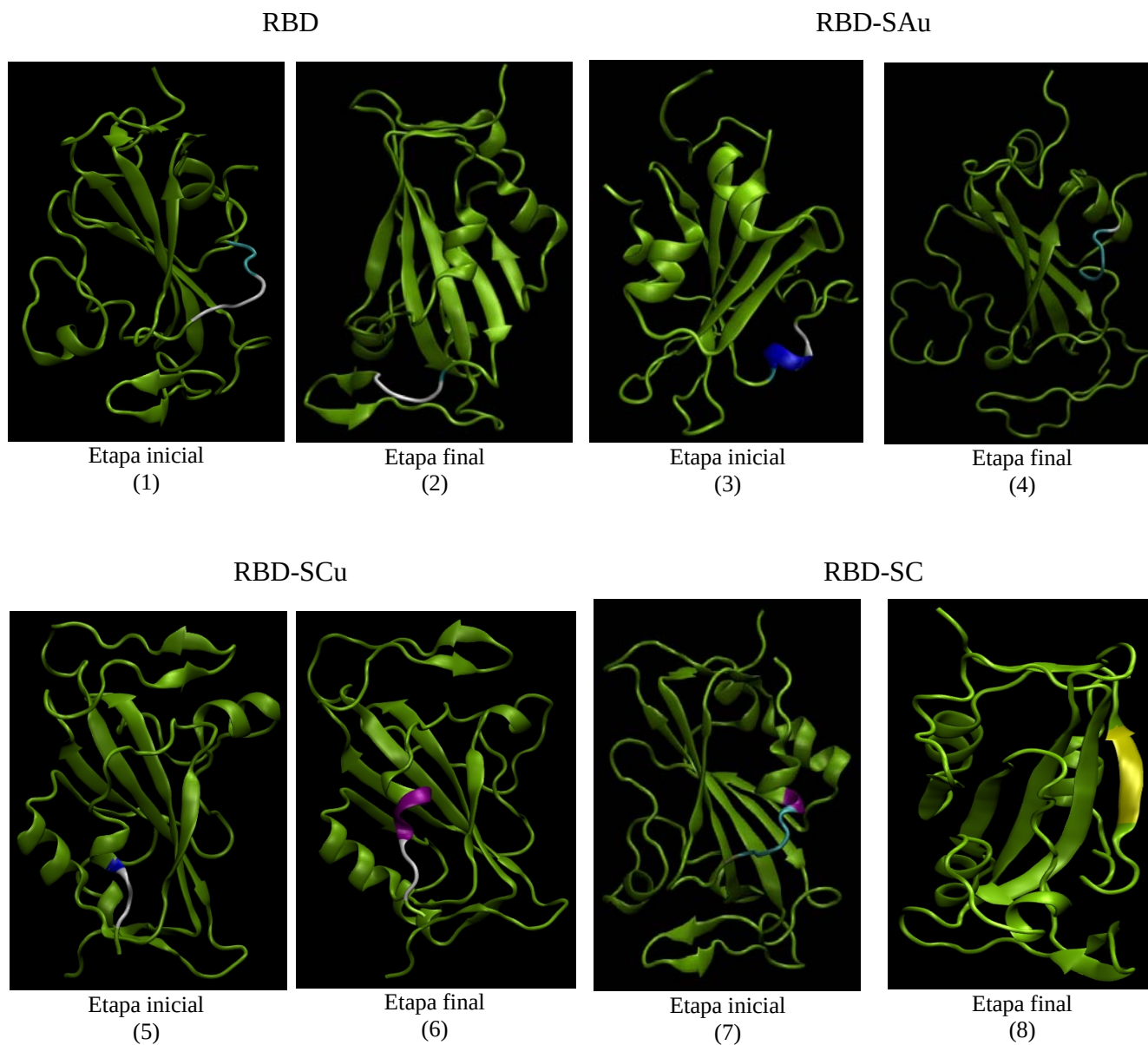


Figura IV.9.2. Estructuras secundarias de los residuos con las regiones más flexibles en el RBD aislado ((1) y (2)) y del RBD sobre las tres superficies en la etapa inicial y final de adsorción sobre SA_u ((3) y (4)), sobre SC_u ((5) y (6)) y sobre SC ((7) y (8)).

V. Conclusiones

El presente trabajo representa la etapa final de adsorción del virus con una superficie, en donde el elemento más externo de la proteína con mayor exposición al entorno (glicoproteína S) interactúa con cada superficie y provee una aproximación de la afinidad del virión con cada una de las superficies estudiadas. Se describió mediante MD el efecto de la adsorción del RBD sobre tres superficies, en comparación con el RBD aislado.

En orden descendente, el RMSD mostró cambios más grandes en la estructura del RBD al adsorberse sobre SAu, que sobre SCu y sobre SC en contraste con el RMSD del RBD aislado. De acuerdo con el radio de giro, el plegamiento del RBD fue favorecido en la adsorción sobre SC, se mantuvo de manera compacta al estar aislada, mientras que su estructura sufrió un desplegamiento al adsorberse sobre las superficies metálicas. En orden descendente, la descompactación fue más grande en presencia de SAu, que de SCu y que de SC. En orden descendente, el conteo de puentes de hidrógeno determinó que la participación más importante de puentes de hidrógeno intramoleculares del RBD ocurrió en presencia de SC, que del RBD aislado, que sobre SCu y que sobre SAu. Por otro lado, la participación de puentes de hidrógeno entre RBD-(H₂O)_{dis} fue mayor en la adsorción sobre SAu, SCu y SC. En conjunto, explican la estabilidad observada y favorecida del radio de giro en la adsorción sobre las tres superficies. El SASA permitió determinar que el RBD tuvo un mayor grado de exposición al disolvente al interactuar con SAu y SCu en comparación con SC donde la exposición al disolvente fue muy similar a la del RBD aislado. La RDF permitió definir la distancia entre las primeras interacciones RBD-superficie, así como identificar la distancia a la que se encontraban las capas de hidratación sobre SAu y SCu. El perfil de densidad parcial permitió identificar la primera capa de hidratación, así como las primeras interacciones RBD-SC y corroboró la información obtenida en la RDF de las superficies metálicas.

Este parámetro también determinó a la SC como la superficie más hidrofóbica, seguida de SAu y SCu, con el valor más grande en la densidad de la primera capa de hidratación. La frecuencia de residuos en contacto mostró que la adsorción del RBD sobre SAu fue la más favorecida, seguida de SCu y SC, debido a la naturaleza polar de los residuos involucrados durante la etapa inicial y final de adsorción. La línea temporal de la estructura secundaria exhibió cambios conformacionales específicos en la estructura inicial del RBD como consecuencia de la interacción con las tres superficies y no solamente por efecto de la relajación de la estructura del RBD aislado en su configuración final. En conjunto con la frecuencia de contactos, la línea temporal de estructura secundaria mostró que los residuos involucrados en la interacción RBD-ACE2 permanecieron predominantemente en estructuras secundarias de tipo lazo y tipo vuelta, en ambas etapas de adsorción. Finalmente, en orden descendente, el RMSF mostró que la flexibilidad de los residuos del RBD fue más afectada durante la etapa final de adsorción sobre SAu, que sobre SCu y SC, en comparación con la flexibilidad de los residuos del RBD aislado, incluyendo los residuos de la interacción RBD-ACE2.

Estos resultados proveen evidencia de que la interacción del RBD con las superficies estudiadas en el intervalo propuesto de simulación modifica la flexibilidad del RBD y de los segmentos críticos para la interacción RBD-ACE2 y por tanto, irrumpe en la organización espacial atómica del esqueleto y la cadena lateral en el RBD y en los residuos clave para la interacción RBD-ACE2, así como en su capacidad para adaptarse a su entorno, lo que podría conllevar a la pérdida de su función biológica principal.

En lo que concierne a las posibles implicaciones prácticas de estos resultados, es importante remarcar que este trabajo es una simplificación de un escenario real, dado que no se consideran los efectos de adsorción de todo el virus sobre las superficies, interacciones dominadas por fuerzas de largo alcance.

Además, dado que el disolvente juega un papel fundamental en las interacciones, usar como modelo de disolvente medios biológicos como la saliva, así como sus interacciones con las superficies resultarían de gran interés.

VI. Perspectivas a futuro

La pandemia puso en evidencia la importancia de estudiar y entender todo lo relacionado con el virus SARS-Cov-2, existen muchas vías de ataque para entender un problema tan complicado y complejo. Una de las posibles formas para abordar este problema mediante la físicoquímica teórica es la MD, así como la química computacional. En este trabajo se eligió la MD como herramienta debido al tamaño de los sistemas a estudiar, dentro de las posibilidades de abordaje en esta misma corriente se proponen las siguientes líneas de ataque:

- Utilizar MD fuera de equilibrio para estudiar la desorción del RBD sobre las distintas superficies.
- Utilizar otro campo de fuerzas y comparar los resultados con el campo de fuerzas utilizado (CHARMM36).
- Utilizar métodos de QM/MM para el estudio de regiones específicas (residuos presentes en el vínculo RBD-ACE2) en la adsorción del RBD sobre las superficies.
- Realizar la caracterización cuantitativa de la adsorción del RBD sobre las superficies utilizadas.
- Realizar la caracterización en la adsorción sobre otras superficies.
- Utilizar un modelo que contenga la estructura completa de la glicoproteína S.

VII. Apéndice

Espacio recíproco y transformada de Fourier

El espacio recíproco es un espacio generado a partir de vectores, digamos \mathbf{a}^* , \mathbf{b}^* y \mathbf{c}^* relacionados con los vectores \mathbf{a} , \mathbf{b} y \mathbf{c} que generan el espacio real. El espacio recíproco asociado al espacio real es el espacio real. En otras palabras, ambos espacios son uno recíproco del otro. Están relacionados por la transformada de Fourier, de forma que el espacio recíproco es también conocido como espacio de Fourier.

La transformada de Fourier es una técnica matemática que mapea una función $f(r)$ en el espacio real en una función conjugada $g(k)$ donde r y k están definidas en los espacios real y recíproco, respectivamente. Las variables en el espacio de Fourier tienen dimensiones recíprocas a las dimensiones de las variables en el espacio real.

La transformada de Fourier de una función $f(r)$, denotada como $F[f(r)]$, está definida como:

$$F[f(r)] = g(k) = \frac{1}{(2\pi)^{3/2}} \int f(r) e^{-ik \cdot r} dr \quad (47)$$

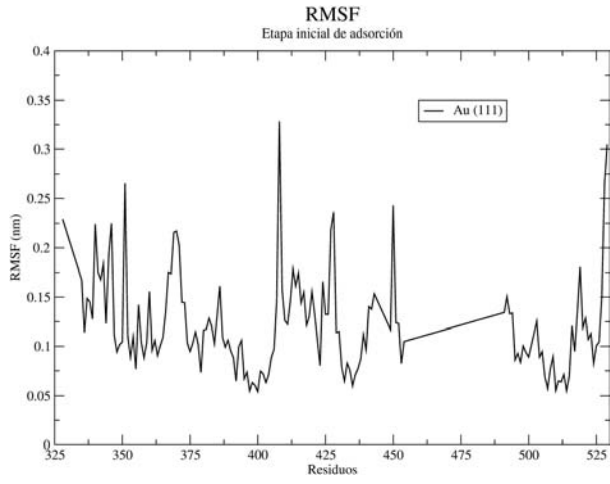
Donde la integral está definida sobre el intervalo de las tres dimensiones cartesianas.

La transformada de Fourier recíproca mapea $g(k)$ de regreso a $f(r)$, técnica particularmente útil cuando $f(r)$ no se puede obtener directamente, pero sí es posible determinar $g(k)$. Una vez obtenida $g(k)$, es transformada de regreso al espacio real para obtener la función $f(r)$.

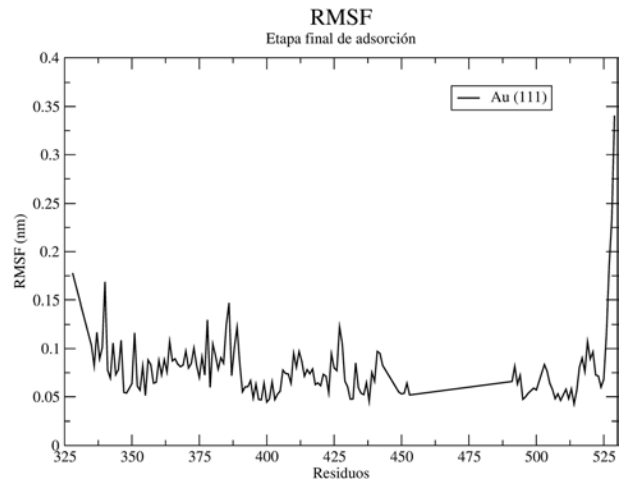
El concepto del espacio recíproco es clave en el entendimiento de los patrones de difracción de rayos X generados por un haz incidente de rayos X sobre un cristal. Dicho patrón está compuesto por puntos generados por un conjunto de planos en el cristal, donde cada punto representa un vector en el espacio recíproco. De forma que el espacio recíproco permite explicar fenómenos dentro de la Cristalografía que dentro del espacio real no sería posible explicarlos adecuadamente.

RMSF

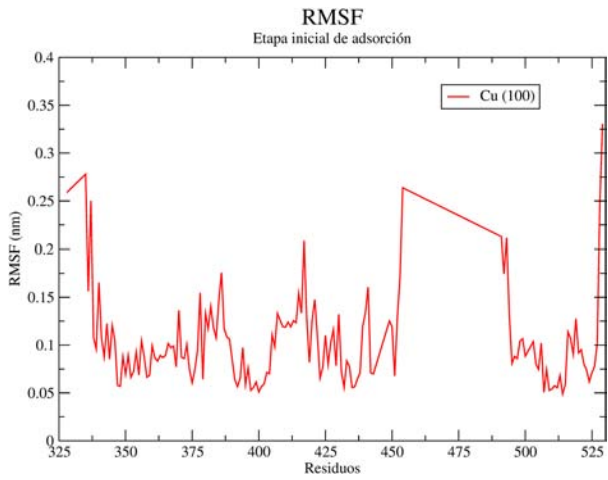
Se muestran las gráficas individuales de la sección del RMSF como función de los residuos del RBD de la simulación en las etapas inicial (a) y final (b) de adsorción del RBD sobre SAu (negro), para las etapas inicial (c) y final (d) de adsorción del RBD sobre SCu (rojo), para las etapas inicial (e) y final (f) de adsorción del RBD con SC (verde) y en el mismo intervalo definido para las etapas inicial (g) y final (h) de adsorción sobre las tres superficies para el RBD aislado (azul).



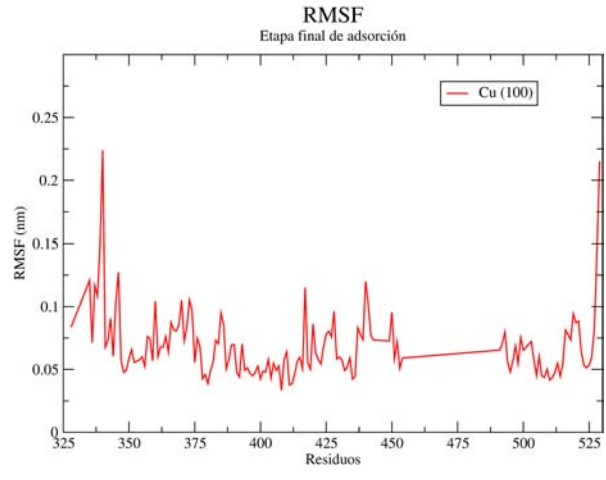
(a)



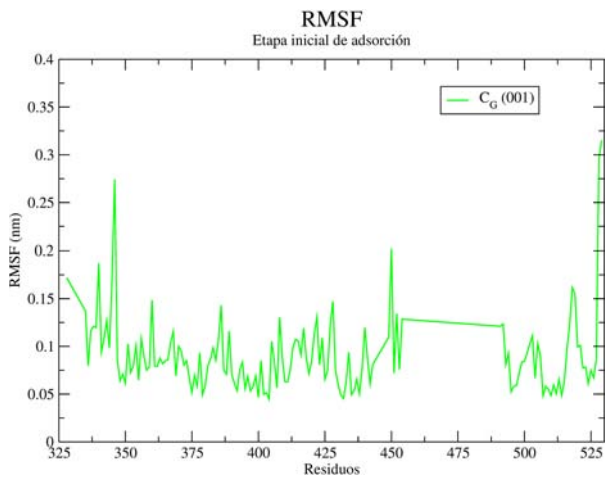
(b)



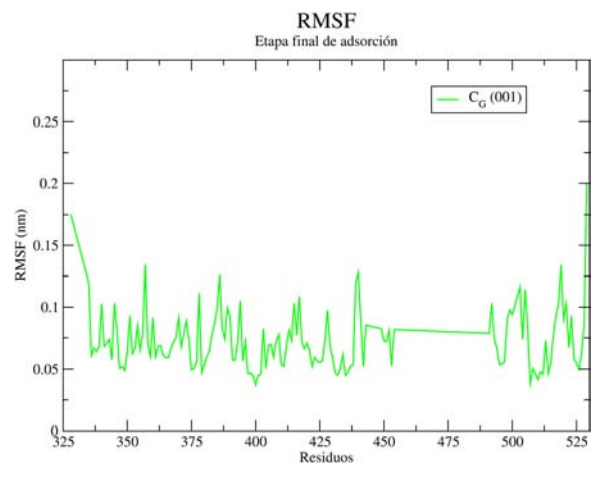
(c)



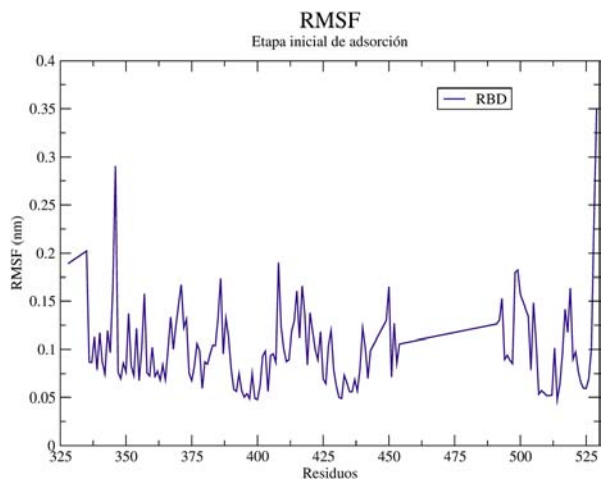
(d)



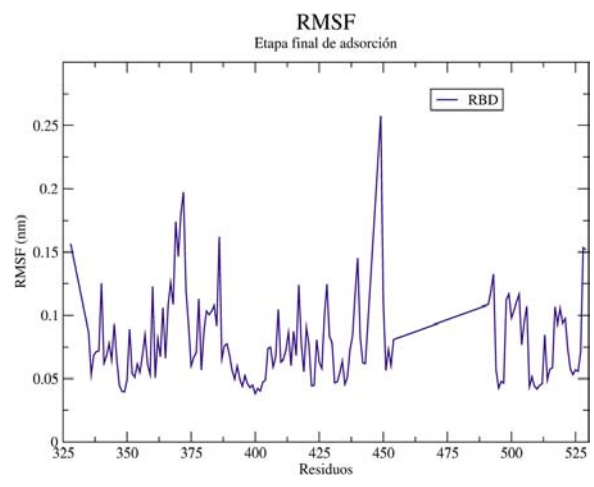
(e)



(f)



(g)



(h)

VIII. Bibliografía

- [1]. N. Zhu, D. Zhang, W. Wang, L. Xingwang, Y. Bo, S. Jingdong, Z. Xiang, H. Baoying, S. Weifeng, L. Roujian, N. Peihua, Z. Faxian, M. Xuejun, W. Dayan, X. Wenbo, W. Guizhen, G. G. F. and T. Wenjie, "A Novel Coronavirus from Patients with Pneumonia in China, 2019", *New England Journal of Medicine*, vol. 382, Feb. 2020.
- [2]. Z. Benková, M.N.D.S. Cordeiro, "Structural behavior of monomer of SARS-CoV-2 spike protein during initial stage of adsorption on graphene", *Elsevier*, vol. 22, Aug. 2021.
- [3]. L. L. Maldonado, A. M. Bertelli, L. Kamenetzky, "Molecular features similarities between SARS-CoV-2, SARS, MERS and key human genes could favour the viral infections and trigger collateral effects", *Scientific Reports*, vol. 11, Feb. 2021.
- [4]. G. Zhu, C. Zhu, Y. Zhu, and F. Sun, "Minireview of progress in the structural study of SARS-CoV-2 proteins", *Current Research in Microbial Sciences*, vol. 1, Sept. 2020.
- [5]. J. Lan, J. Ge, J. Yu, S. Shan, H. Zhou, S. Fan, Q. Zhang, X. Shi, Q. Wang, L. Zhang, and X. Wang, "Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor", *Nature*, vol. 581, Mar. 2020.
- [0]. M. K. Kathiravan, S. Radhakrishnan, V. Namasivayam, and S. Palaniappan, "An Overview of Spike Surface Glycoprotein in Severe Acute Respiratory Syndrome–Coronavirus", *Frontiers in Molecular Biosciences*, vol. 18, May. 2021.
- [6]. Comisión Directiva Sociedad Argentina de Virología, "INFORME SARS COV-2", *aam*, vol. 1, Mar. 2020.
- [7]. B. N. Fields, D. M. Knipe, P. M. Howley, "Fields Virology", *Wolters Kluwer Health/Lippincott Williams & Wilkins*, 2013.
- [8]. D. C. Malaspina and J. Faraudo, "Computer simulations of the interaction between SARS-CoV-2 spike glycoprotein and different surfaces", *Biointerphases*, vol. 15, Oct. 2020.
- [9]. Universidad de Alcalá, "Introducción a la biología molecular del coronavirus SARS-CoV-2", . [] . : <http://www3.uah.es/chemevol/index.php/introduccion-a-la-biologia-molecular-del-coronavirus-sars-cov-2/>.
- [10]. G. Hamid, Y. Hessam, H. F. Elham, and M. Yaghoub, "Interactions of SARS-CoV-2 with inanimate surfaces in built and transportation environments", *Sustainable Cities and Society*, vol. 72, Sept. 2021.
- [11]. M. Abuhegazy, K. Talaat, O. Anderoglu, and S. V. Poroseva, "Numerical investigation of aerosol transport in a classroom with relevance to COVID-19", *Physics of Fluids*, vol. 32, Oct. 2020.
- [12]. P. S. Desai, N. Sawant, and A. Keene, "On COVID-19-safety ranking of seats in intercontinental commercial aircrafts: A preliminary multiphysics computational perspective", *Building Simulation*, vol. 14, Feb. 2021.
- [13]. Z. Zhang, T. Han, K. H. Yoo, J. Capecelatro, A. L. Boehman, and K. Maki, "Disease transmission through expiratory aerosols on an urban bus", *Physics of Fluids*, vol. 33, Jan. 2021.
- [14]. C. Sun and Z. Zhai, "The efficacy of social distance and ventilation effectiveness in preventing COVID-19 transmission", *Sustainable Cities and Society*, vol. 62, Nov. 2020.

- [15]. Y. Zhou and S. Ji, "Experimental and numerical study on the transport of droplet aerosols generated by occupants in a fever clinic", *Building and Environment*, vol. 187, Jan. 2021.
- [16]. H. Woo, S. J. Park, Y. K. Choi, T. Park, M. Tanveer, Y. Cao, N. R. Kern, J. Lee, M. S. Yeom, T. I. Croll, C. Seok, and W. Im, "Developing a Fully Glycosylated Full-Length SARS-CoV-2 Spike Protein Model in a Viral Membrane", *J. Phys. Chem. B*, vol. 124, Aug. 2020.
- [17]. J. He, H. Tao, Y. Yan, S. Huang, and Y. Xiao, "Molecular Mechanism of Evolution and Human Infection with SARS-CoV-2 ", *Viruses*, vol. 12, Apr. 2020.
- [18]. E. S. Brielle, D. Schneidman-Duhovny, M. Linial, "The SARS-CoV-2 Exerts a Distinctive Strategy for Interacting with the ACE2 Human Receptor ", *Viruses*, vol. 12, Apr. 2020.
- [19]. Y. Han, P. Král, "Computational Design of ACE2-Based Peptide Inhibitors of SARS-CoV-2", *ACS Nano*, vol. 14, Apr. 2020.
- [20]. J. Wang, "Fast Identification of Possible Drug Treatment of Coronavirus Disease-19 (COVID-19) through Computational Drug Repurposing Study", *Journal of Chemical Information and Modeling*, vol. 60, Apr. 2020.
- [21]. N. A. Megahed and E. M. Ghoneim, "Antivirus-built environment: Lessons learned from Covid-19 pandemic", *Sustainable Cities and Society*, vol. 61, Oct. 2020.
- [22]. J. Chen, X. Wang, C. Dai, S. Chen, and Y. Tu, "Adsorption of GA module onto graphene and graphene oxide: A molecular dynamics simulation study", *Physica E: Low-dimensional Systems and Nanostructures*, vol. 62, Aug. 2014.
- [23]. Z. Daohui, L. Libo, H. Daohang, and J. Zhou, "Molecular dynamics simulations of conformation changes of HIV-1 regulatory protein on graphene", *Applied Surface Science*, vol. 377, July. 2016.
- [24]. M. P. Allen and D. J. Tildesley, "Computer Simulation of Liquids" , Oxford University Press, 2017.
- [25]. K. N. Grewa and W. K.S. Chiu, "A review of modeling and simulation techniques across the length scales for the solid oxide fuel cell", *Journal of Power Sources*, vol. 199, Feb. 2012.
- [26]. C. Holm, "Simulation Methods in Physics 1" , University of Stuttgart, 2013.
- [27]. E. Braun, J. Gilmer, H. B. Mayes, D. L. Mobley, J. I. Monroe, S. Prasad, and D. M. Zuckerman, "Best Practices for Foundations in Molecular Simulations", *Living J Comput Mol Sci.* , vol. 1, Nov. 2018.
- [28]. T. Schlick, "Molecular Modeling and Simulation: An Interdisciplinary Guide" , Springer, 2006.
- [29]. D. E. Tanner, "Lengthening the timescale reach of molecular dynamics", , , . 2012.
- [30]. S. Alavi, "Molecular Simulations: Fundamentals and Practice" , Wiley-VCH, 2020.
- [31]. M. E. Tuckerman, "Statistical Mechanics: Theory and Molecular Simulations" , Oxford University Press, 2010.
- [32]. L. A. Hess, and van der Spoel, "GROMACS 2021.5 Manual (2021.5)", 2022. [Online]. Available: <https://doi.org/10.5281/zenodo.5849961>.
- [33]. A. R. Leach, "Molecular Modelling: Principles and Applications", Prentice Hall, pp. 174, 2001.
- [34]. , "Specific Interactions " , 2020. [Online]. Available: https://chem.libretexts.org/Bookshelves/Physical_and_Theoretical_Chemistry_Textbook_Maps/

[Supplemental Modules \(Physical and Theoretical Chemistry\)/Physical Properties of Matter/Atomic and Molecular Properties/Intermolecular Forces/Specific Interactions.](#)

- [35]. D. Frenkel and B. Smit , "Understanding Molecular Simulation : From algorithms to applications" , Academic Press, 2002.
- [36]. D. C. Rapaport, "The Art of Molecular Dynamics Simulation" , Cambridge University Press, 2010.
- [37]. D. H. J. Mackay, A. J. Cross, and A. T. Hagler, "Prediction of Protein Structure and the Principles of Protein Conformation", Springer Science & Business Media, pp. 317-358, 2012.
- [38]. L. Boltzmann, "Lectures on Gas Theory" , Dover Publications, 1964.
- [39]. P. Ehrenfest and T. Ehrenfest, "The Conceptual Foundations of the Statistical Approach in Mechanics" , Dover Publications, 2002.
- [40]. T. L. Hill, "Statistical Mechanics : Principles and selected applications" , Dover Publications, 2020.
- [41]. A. I. Kinchin and G. Gamow, "Mathematical Foundations of Statistical Mechanics" , Martino Publishing, 2014.
- [42]. L. García and C. Scherer, "Introducción a la termodinámica clásica" , Trillas, 1990.
- [43]. H. B. Callen, "Thermodynamics" , John Wiley & Sons, 1985.
- [44]. M. Ozboyaci, D. B. Kokh, S. Corni, and R. C. Wade, "Modeling and simulation of protein–surface interactions: Achievements and challenges", Quarterly Reviews of Biophysics, vol. 49, Jan. 2016.
- [45]. J. Sunhwan, K. Taehoon, I. G. Vidyashankara, and I. Wonpil, "CHARMM-GUI: A web-based graphical user interface for CHARMM", Journal of Computational Chemistry, vol. 29, Jun. 2008.
- [46]. Legigh , " CHARMM-GUI Provides a Close-up Look at Bacteria ", 2021. [Online]. Available: <https://www2.lehigh.edu/news/charmm-gui-provides-a-close-up-look-at-bacteria>.
- [47]. NIH Biomedical Research Center for Macromolecular Modeling and Bioinformatics, "VMD User's Guide", 2016. [Online]. Available: https://www.ks.uiuc.edu/Research/vmd/allversions/what_is_vmd.html.
- [48]. F. Weinhold, R.A. Klein, "What is a hydrogen bond? Mutually consistent theoretical and experimental criteria for characterizing H-bonding interactions" , , vol. 110, Feb.. 2012.
- [49]. E. N. Baker, R. E. Hubbard, "Hydrogen Bonding in Globular Proteins" , Massey University, 1984.
- [50]. G. A. Petsko and D. Ringe, "Protein Structure and Function " , Oxford University Press, 2004.
- [51]. T. A. HORBETT, B. D. RATNER, J. M. SCHAKENRAAD, F. J. SCHOEN, "Biomaterials Science: An Introduction to Materials in Medicine" , Biomaterials Science, 1996.
- [52]. J. M. Acres, A. V. Verde, "Investigating the Specificity of Peptide Adsorption on Gold Using Molecular Dynamics Simulations", Biomacromolecules, vol. 10, July. 2009.
- [53]. S. Chen, B. Wang, W. Huang, S. Gao, "Molecular dynamics simulation of the interaction between Maps and Cu(100) surface", 2010 2nd Conference on Environmental Science and Information Application Technology, ESIAT 2010, vol. 3, Jul.. 2010.

- [54]. R. Braun, M. Sarikaya, K. Schulten, "Genetically engineered gold-binding polypeptides: structure prediction and molecular dynamics", *Journal of biomaterials science. Polymer edition*, vol. 13, Mar.. 2002.
- [55]. Z. Yang, Y Zhao, "Adsorption of His-tagged peptide to Ni, Cu and Au (100) surfaces: Molecular dynamics simulation", *Engineering Analysis with Boundary Elements*, vol. 31, Feb.. 2007.
- [56]. R. Totani, C. Méthivier, D. Costa, T. J. Inizan, C. M. Pradier, V. Humblot, , "Binding and 2D organization of arginine on Cu(1 1 0)", *Applied Surface Science*, vol. 509, Apr.. 2020.
- [57]. J. Katoch, K. S. Nyon, K. Zhifeng, B. L. Farmer, R. R. Naik, S. A. Tatulian, M. Ishigami, "Structure of a Peptide Adsorbed on Graphene and Graphite", *Nano Letters*, vol., May. 2012.
- [58]. N. Ghani, R. Karjiban, M. Basri, N. Faujan, W. Z. Lim, "Unveiling Amyloid- β 1–42 Interaction with Zinc in Water and Mixed Hexafluoroisopropanol Solution in Alzheimer's Disease", *International Journal of Peptide Research and Therapeutics*, vol. 23, Dec.. 2016.
- [59]. D. S. Saputri, L. Songling, R. J. E. Floris, Z. Xu, H. S. Ismanto, A. Davila, S. Teraguchi, K. Katoch, M. S. Daron, "Flexible, Functional, and Familiar: Characteristics of SARS-CoV-2 Spike Protein Evolution", *Frontiers in Microbiology*, vol. 11, Sept.. 2020.