



UNIVERSIDAD  
**MAYOR**

para espíritus emprendedores

Facultad de Ciencias

---

**ESCUELA DE  
BIOTECNOLOGÍA**

---

**“Diseño *in silico* de un péptido capaz de interferir con la formación de enlace isopeptídico en el c-terminal de la adhesina Ag I/II de *Streptococcus mutans*”**

Marcelo Emmanuel Jorquera Bolados

Proyecto de Tesis para optar al título de Biotecnólogo

Santiago - Chile  
Año 2023



UNIVERSIDAD  
**MAYOR**

para espíritus emprendedores

Facultad de Ciencias  

---

**ESCUELA DE  
BIOTECNOLOGÍA**

---

**“Diseño *in silico* de un péptido capaz de interferir con la formación de enlace isopeptídico en el c-terminal de la adhesina Ag I/II de *Streptococcus mutans*”**

Marcelo Emmanuel Jorquera Bolados

Proyecto de Tesis para optar al título de Biotecnólogo

Tutor: Dr. Jaime Andrés Rivas Pardo

Santiago - Chile  
Año 2023

## **AGRADECIMIENTOS:**

Me gustaría empezar dando gracias a la Universidad Mayor y a la escuela de Biotecnología por permitir desarrollarme como un buen científico y madurar como persona, siempre otorgando profesores de gran renombre y una verdadera pasión por enseñar y ver a sus alumnos mejorar poco a poco. Entre esos docentes se me es necesario destacar a mi tutor, el Dr. Jaime Andrés Rivas Pardo al cuál admiro muchísimo por no sólo convertirse en mi visión de cómo debería ser un investigador, sino que también un buen profesional, además de siempre estar ahí para aconsejarme cuando era necesario.

Gracias a los miembros del laboratorio, especialmente a Ivana Orellana por enseñarme todo lo que sé sobre dinámicas moleculares y por ser siempre no sólo una muy buena profesora, sino que también una muy buena amiga.

Gracias a mis amigos, tanto a aquellos que conocí en el colegio y a aquellos que conocí durante el camino universitario, gracias por las risas, el cariño y por siempre estar ahí cuando los necesitaba. Una mención especial a Francisco Salazar al que ya no sólo considero como un amigo sino como un hermano al cuál me siento muy afortunado de haber conocido y a Guillermo Albornoz la amistad más linda que pude haber conocido dentro de la universidad.

Muchas gracias a mi pareja Karen Mella, que no importa tan difícil fue el camino siempre estuvo ahí para mí y no hay palabras suficientes en el diccionario como para expresar lo importante que ella es para mí en esta vida.

Finalmente quiero dedicar este trabajo a mis padres ya que sin ellos no estaría aquí en este momento. Gracias por escuchar mis quejas, mis alegrías, mis penas y mis celebraciones, gracias por criarme en el adulto que soy hoy en día y gracias por todo el esfuerzo que han hecho durante estos años. Espero estén felices de ver que ese pequeño niño al que vieron nacer hace 27 años está por dar el siguiente gran paso de su vida... y todo gracias a ustedes.

## INFORMACIÓN DEL CENTRO DE INVESTIGACIÓN

### **Centro de Genómica y Bioinformática – Universidad Mayor**

El desarrollo de esta investigación se hizo dentro de las instalaciones del Centro de Genómica y Bioinformática (CGB) de la Universidad Mayor el cual es liderado por la Dra. Elena Vidal. El CGB está segmentado en 3 unidades de investigación: Laboratorio de Bioinformática, Laboratorio de Genómica Vegetal y Laboratorio de Genómica Microbiana.

El director de la tesis fue el profesor Jaime Andrés Rivas Pardo, líder del grupo de Biología Mecánica perteneciente al Laboratorio de Genómica Microbiana, el cual es dirigido por los profesores Dr. Sebastián Reyes y por el mismo Dr. Rivas Pardo. Cada uno posee una línea de investigación independientes que se reúnen en el estudio y caracterización de mecanismos involucrados en la relación patógeno-hospedero.

En el caso de este trabajo de investigación asociado al Dr. Rivas Pardo, este se enmarcó en la propuesta asociada al entendimiento del mecanismo que gobierna la adhesión bacteriana basada en la estrategia implementada por distintos filamentos de pilina – Mecano-Farmacología de la adhesión bacteriana – proyecto de investigación apoyado por la Agencia Nacional de Investigación y Desarrollo (ANID) mediante el fondo FONDECYT (1221064). Es el, se detallaba como parte de sus objetivos la aplicación de biología computacional acoplada a ingeniería de proteínas y espectroscopía de fuerzas con tal de estudiar en detalle el *pilus* de *Streptococcus mutans*, *Streptococcus pneumoniae*, *Streptococcus pyogenes*, y *Staphylococcus aureus*. Además, el trabajo teórico como experimental incluido en la tesis constó con el apoyo de fondos internos de la Universidad por medio del FDP 2021-20 PEP I-2022019.

## ABREVIATURAS:

- A: Alanina
- Asn; N: Asparagina
- C: Cisteína
- D: Ácido aspártico
- E: Ácido glutámico
- F: Fenilalanina
- G: Glicina
- H: Histidina
- I: Isoleucina
- L: Leucina
- Lys; K: Lisina
- M: Metionina
- P: Prolina
- Q: Glutamina
- R: Arginina
- S: Serina
- SAG: Complejo de aglutininas salivares libres
- SMD: Dinámicas moleculares dirigidas.
- T: Treonina
- V: Valina
- W: Triptófano
- Y: Tirosina
- Gp340: Glicoproteína 340

## ÍNDICE

i. RESUMEN.....	v
ii. SUMMARY.....	vi
1 INTRODUCCIÓN: .....	1
1.1 Caries: causas y efectos en la población .....	1
1.2 Biopelículas orales y su papel en la cariogénesis.....	2
1.2.1 Resistencia a antibióticos, una problemática en potencia.....	3
1.3 Mecanismos de adhesión de <i>Streptococcus mutans</i> .....	5
1.3.1 Enlaces isopeptídicos: .....	8
1.4 Definición de la propuesta.....	10
2 HIPOTESIS Y OBJETIVOS .....	13
2.1 Hipótesis .....	13
2.2 Objetivos.....	13
3 MATERIALES Y MÉTODOS .....	14
3.1 Material utilizado.....	14
3.1.1 Software empleados: .....	14
3.1.2 Equipos: .....	14
3.2 Metodología .....	15
3.2.1 Caracterización mecánica .....	15
3.2.2 Diseño del péptido.....	21
3.2.3 Evaluación de los péptidos .....	28
4 PRESENTACIÓN Y ANÁLISIS DE RESULTADOS .....	29
4.1 Caracterización mecánica e identificación del dominio más influyente.....	29
4.2 Generación del bolsillo para acoplamiento molecular .....	36
4.3 Diseño del péptido.....	37
4.3.1 Iteraciones de péptidos .....	37
4.3.2 Escaneo de alaninas.....	41
4.3.3 Evaluación mecánica de los péptidos .....	45
5 DISCUSIÓN .....	48
6 PROYECCIONES .....	54
7 CONCLUSIONES.....	55
8 REFERENCIAS: .....	56

## i. RESUMEN

Siendo la causa etiológica más prominente en el desarrollo de las caries, *Streptococcus mutans* es una de las bacterias orales Gram-positivo que forman la biopelícula conocida comúnmente como placa bacteriana. Este consorcio de bacterias es el precursor de la colonización patogénica del tejido dental y del desarrollo de las caries. La bacteria *S. mutans* emplea una simple estrategia para adherirse al diente la cual está basada en el empleo de apéndices proteicos conocidos como *pilus* o adhesinas, filamentos de micrones de largo que se proyectan desde la pared celular y que cumplen la función de servir de verdaderas anclas moleculares. La adhesina en cuestión, o Ag I/II, posee una porción C-terminal de tipo elástica— compuesta por los dominios C1, C2 y C3— que no sólo forman parte importante en el proceso de adhesión mismo, sino que también participan en procesos de disipación de energía mecánica. Cada uno de estos dominios posee en su interior un enlace isopeptídico formado por la cadena lateral de un residuo de asparagina y el grupo e-amino saliente de un residuo de lisina. Este enlace isopeptídico une covalentemente hebras beta de la adhesina, formando un complejo estable y altamente resistente al esfuerzo mecánico.

En este trabajo gracias al empleo de estrategias computacionales hemos estudiado la relevancia de los dominios C1, C2 y C3, con el fin de encontrar un objetivo para una posible intervención del proceso mecánico gestionado por esta adhesina. Por medio de dinámicas moleculares dirigidas pudimos describir que al remover el enlace isopeptídico, en el dominio C2 se desencadenan una gran abundancia de eventos mecánicos, lo que lo convierte en un adecuado modelo para intervención peptídica.

A continuación, y mediante herramientas *in silico*, diseñamos y evaluamos una serie de péptidos que buscaban unirse al dominio C2 e interferir la formación natural del enlace isopeptídico entre las cadenas laterales de los residuos Lys1161 y Asn1311. La especificidad de los péptidos fue probada mediante análisis de acoplamiento molecular ciego, para luego estos ser refinados por un protocolo de escaneo de alaninas el cual fue acoplado mediante una metodología propia para funcionar en condiciones *in silico*, para luego evaluar la estabilidad de estos mediante análisis de dinámica molecular dirigida. Nuestro trabajo da los primeros pasos hacia estrategias de desarrollo de moléculas diseñadas capaces de intervenir en la formación de enlaces relacionados a la estabilidad de proteínas que median la adhesión de bacterias Gram-positivas.

## ii. SUMMARY

As the most prominent etiological cause of caries development, *Streptococcus mutans* is one of the many Gram-positive oral bacteria that form the biofilm commonly known as plaque. This biofilm is the precursor to the colonization of dental tissue and the development of caries. *S. mutans* employs a simple strategy to adhere to the tooth based on the use of protein appendages known as pilus or adhesins, micrometric filaments that project from the cell wall and serve as true molecular anchors. The adhesin in question, or Ag I/II, has an elastic C-terminal portion - C1, C2 and C3 domains - which not only play an important part in the adhesion process itself, but also participate in mechanical energy dissipation processes. Each of these domains possesses an isopeptide bond autocatalyzed by the presence of a carboxylic group that promotes the attack of the carbonyl group on the outgoing  $\epsilon$ -amino group of a lysine residue. The isopeptide bond formed covalently binds beta-strands of the adhesin, forming a stable complex that is highly resistant to mechanical stress.

In this work thanks to the use of computational strategies we have studied the relevance of the C1, C2 and C3 domains in order to find a target for a possible intervention of the mechanical process managed by this adhesin. By means of directed molecular dynamics we were able to describe that within the C-terminal domains, the C2 domain corresponds to the most relevant one, since, upon removal of the isopeptide bond, a great abundance of mechanical events is triggered, which makes it a suitable model for peptide intervention.

Next, using *in silico* tools, we designed and evaluated a series of peptides that sought to bind to the C2 domain and hinder the natural formation of the isopeptide bond between the Lys1161 and Asn1311 side chains. The specificity of the peptides was tested by blind molecular docking analysis, and then these were refined by an alanine scanning protocol which was coupled by a proprietary methodology to work under *in silico* conditions, and then evaluated for stability by SMD analysis. Insert areas under the curve here. Our work takes the first steps towards strategies for the development of molecules capable of intervening in the catalysis of stability-related linkages in Gram-positive bacteria, as well as intervening in related adhesion processes.

## 1 INTRODUCCIÓN:

### 1.1 Caries: causas y efectos en la población

Las enfermedades dentales son un problema que como sociedad nos han acompañado desde la antigüedad. Los primeros reportes sobre decaimiento dental datan del tiempo de la civilización de los sumerios, que definían la causa del problema como “gusanos que habitaban en la boca” (1). Por otro lado, evidencia reciente ha demostrado que la civilización del valle del Indo aproximadamente en el año 7000 a.C, fueron de los primeros en practicar odontología usando herramientas rudimentarias de la época (1). Sin embargo, no es hasta el final de la época moderna, siendo a partir de la última década del año 1800, cuando se comenzaron a reportar estadísticas de enfermedades orales de manera oficial. Esto coincide con el desarrollo de los primeros centros de estudios especializados en el entrenamiento de jóvenes estudiantes en el área de la odontología, por parte de la mano de Horace H. Hayden y Chapin A. Harris, pioneros en la salud dental y actualmente conocidos como padres de la odontología (1 – 2).

Hoy en día conocemos que la salud dental de un individuo se puede ver comprometida por diversos factores, incluyendo, la nutrición, la cual tiene estrecha relación con la integridad de la cavidad oral, por ejemplo, la dieta de un individuo tiene un efecto local en la salud bucal de una persona, modificando condiciones de pH y la composición de tanto la saliva como del esmalte (3). Una disbiosis oral, o desequilibrio entre las especies de microorganismos que habitan la cavidad oral, promueve el desarrollo de enfermedades dentales como la periodontitis y las caries, que afectan tanto al tejido dental como a la encía (4). Por ejemplo, una acidificación de la cavidad oral puede entregar un ambiente propicio para el crecimiento y propagación de bacterias acidófilas que producen inflamación local y deterioro de las piezas dentales. Si esto se combina con una dieta rica en azúcares, puede modificarse la composición de la saliva cambia convirtiéndola en un medio idóneo que colabora en el metabolismo de las varias bacterias patogénicas orales (5).

Dentro de las enfermedades dentales que reúnen una acidificación del medio y una alta abundancia de hidratos de carbono se encuentran las caries, padecimiento que ha tomado alta relevancia dado el explosivo aumento de casos reportados en los últimos años (6). Según estudios de la organización mundial de la salud (OMS), sobre el 50% de la población a escala global se encuentra infectada por caries y un 90% las ha padecido en algún punto de su vida, índices que hacen de las caries una de las

enfermedades crónicas más prevalentes a nivel mundial, siendo la principal causa de pérdida de piezas dentales en las personas (7 – 8). Fisiológicamente, las caries se desarrollan como pequeños agujeros acompañados de máculas negras, marrones o blancas en la superficie del diente, generando una sintomatología bastante variada. Entre los síntomas más comunes se puede observar: (i) dolor punzante en la zona afectada, (ii) alto grado de sensibilidad hacia temperaturas y/o texturas, y (iii) variaciones de peso y estatura en infantes. Por otro lado, se ha demostrado que la presencia de caries en ocasiones se puede relacionar con la falta de productividad laboral e incluso depresión (9 – 10).

En términos patológicos, una carie se define como una enfermedad de origen netamente bacteriano, dada la interacción que hay entre el diente y una delgada biopelícula microbiana compuesta de una gran variedad de bacterias acidófilas de tipo Gram positivo (11). La colonización hacia el tejido dental da lugar al decaimiento del diente por la liberación de ácidos provenientes del metabolismo de las bacterias que componen la biopelícula, dada la presencia de azúcares fermentables presentes en la cavidad bucal (12 – 13). La producción de ácidos se hace en todas direcciones pasando a través de los poros del esmalte y la dentina hasta llegar al tejido. Si este proceso es lo suficientemente extendido en el tiempo, y no hay tratamiento de por medio o bien si la higiene bucal no es ideal, termina formándose una cavidad en el diente (14). Actualmente, los tratamientos en contra de las caries no van hacia prevenir el desarrollo de la carie misma, sino que, son procedimientos correctivos que se aplican cuando la carie ya está formada. No obstante, muchas naciones, incluido Chile, poseen programas de higiene bucal y políticas para la salud oral tales como la fluoración del agua potable y dentífricos altos en flúor. Ambos son herramientas que de alguna manera colaboran a revertir o mitigar el efecto de las caries en la población, sin embargo, el tratamiento para las caries se basa en el arrastre mecánico de las bacterias que han colonizado el tejido. En este caso, un especialista, mediante el uso de una cureta remueve parte de la placa bacteriana en el área afectada, lo cual está acoplado al uso de resinas sintéticas que se usan para cubrir por completo en la cavidad, formando lo que se conoce como una tapadura, y, finalmente, en casos en donde la carie está un punto de desarrollo tal que ninguna de las alternativas anteriores podría funcionar, se opta por la remoción de la pieza dental completa (15 – 19).

## **1.2 Biopelículas orales y su papel en la cariogénesis**

La placa dental se forma del conjunto que existe entre la superficie del esmalte con una delgada biopelícula de bacterias que se alojan nativamente en la cavidad bucal (20). Una biopelícula

corresponde a un consorcio organizado de bacterias contenidas dentro de una matriz polimérica de exopolisacáridos que se produce por parte del metabolismo de las mismas bacterias que la componen (21). Además, esta agrupación de bacterias a pesar de ser sensible a antibióticos tradicionales, de manera conjunta en la biopelícula, se vuelven refractarias al uso de antibióticos, haciendo extremadamente difícil tratar la enfermedad una vez formada la biopelícula (22). El consorcio está dominado por bacterias acidogénicas Gram-positivo, caracterizadas por una gruesa pared de peptidoglucano y carentes de membrana lipídica externa, principalmente de los géneros, *Streptococcus*, *Lactobacillus*, *Propionobacterium* y *Actinomyces* (23 – 24). Fuera de que todos los géneros mencionados anteriormente juegan un rol en la cariogénesis – o capacidad de formar caries –, una de las bacterias más relevante dentro del conjunto es *Streptococcus mutans*. Esta bacteria que habita naturalmente la cavidad bucal, y es conocida como la causa etiológica más relevante en el desarrollo de las caries (25). Además, se ha demostrado que *S. mutans* participa de otras enfermedades, tales como la endocarditis infectiva, demostrando mediante exámenes de sangre que aproximadamente un 38% de las bacterias en la sangre de pacientes enfermos pertenecían a *Streptococcus* orales Gram-positivos, proveniente del paso de las bacterias al torrente sanguíneo por un proceso conocido como bacteremia (26 – 27). El potencial cariogénico que tiene *S. mutans* es atribuido a tres características: (i) habilidad de poder sintetizar grandes cantidades de polímeros extracelulares basados en glucanos a partir de sacarosa colaborando a la colonización permanente del tejido y convirtiéndola en el precursor de la biopelícula, (ii) capacidad de transportar y metabolizar una cantidad considerable de hidratos de carbono para transformarlos en ácidos orgánicos, ambiente perfecto para que otras bacterias acidófilas puedan coexistir y (iii) habilidad de sobrevivir en condiciones de estrés y en pH bajos (28 – 29).

### **1.2.1 Resistencia a antibióticos, una problemática en potencia**

El hecho de que *S. mutans* y las demás bacterias Gram-positivas que habitan en la cavidad bucal se organicen en una biopelícula, da muchas ventajas evolutivas, dentro de ellas, su comportamiento refractario al uso de antibióticos. No obstante, probablemente una situación inadvertida de este consorcio de bacterias es la posibilidad de intercambiar genes de resistencias lo cual agrava aún más la situación de manejo de infecciones de microorganismos resistentes (30).

La presencia de patógenos multirresistentes es una realidad que como sociedad llevamos va enfrentando desde hace tiempo. A pesar de que la resistencia a antibióticos no se representa como

una enfermedad a *per se*, el verdadero peso que genera en la salud recae en el hecho de que los antibióticos existentes ya no dan abasto para al tratamiento de infecciones que anteriormente si eran tratables mediante el uso de estas moléculas, transformando estas infecciones como intratables, aumentando incluso el rango de mortalidad asociado (31 – 32). En el caso de la cavidad oral, el tratamiento con antibióticos se usa con tal de evadir la continuidad de la infección usando principalmente penicilina y clindamicina, antibióticos usados como primera opción al tratamiento de infecciones orales (33). Sin embargo, tomando en cuenta las ventajas evolutivas que ofrece la organización de la biopelícula, en conjunto con los crecientes caso de automedicación han hecho que gran parte de las bacterias que habitan en la boca se hayan hecho resistentes al tratamiento con antibióticos (34). Estudios de multiresistencia sobre las bacterias que componen la biopelícula mostró que las bacterias con mayor índice de resistencia corresponden a *S. mutans*, *S. gordonii*, *S. oralis*, *S. sanguinis* y *S. salivarius*, siendo *S. mutans* la bacteria con mayor espectro de multiresistencia (35).

La comunicación entre bacterias juega un rol clave en la resistencia a agentes antimicrobianos, lo cual es facilitada precisamente por la formación de la biopelícula. Cuando las bacterias se organizan en una comunidad, se hace más común la transferencia horizontal de material genético, proceso por el cual un conjunto de bacterias pueden intercambiar genes asociados a resistencia (36).

La resistencia a antibióticos en bacterias se desarrolla en respuesta a las distintas dianas químicas y moleculares que presentan los antibióticos usados en la actualidad: (i) síntesis de proteínas bacterianas; (ii) replicación y reparación de ácidos nucleicos y (iii) enzimas y sustratos relacionados a la biosíntesis de la pared celular (37). Referente al punto uno, los inhibidores de síntesis de elementos proteicos corresponden a los antibióticos en la categoría de los macrólidos y los aminoglucósidos, que funcionan de manera de interferir con la correcta lectura de los codones y causar un término temprano de la traducción en la subunidad 30S del ribosoma (38). A la actualidad existen dos fenotipos resistentes a la actividad de estas clases de antibióticos: El primero, consiste en la modificación química del antibiótico agregando grupos funcionales en secciones estratégicas de la molécula con tal de remover la afinidad que tiene este por el ARN diana (39); mientras que el segundo fenotipo resistente se adquiere mediante la modificación de la diana molecular misma mediante mutaciones puntuales en genes codificantes de las secciones 30S y 16S del ribosoma (40). En el punto 2, entran en juego las fluoroquinolonas que se encargan de atacar ADN girasas y topoisomerasas con tal de intervenir el proceso de replicación del material genético (41). La resistencia contra este tipo de antibióticos se da

mediante la mutación de las zonas de determinación de resistencia a quinolonas (QRDR), las cuales son capaces de cambiar su conformación con tal de adquirir resistencia al alterar la estructura o la función de las proteínas que codifican. estos cambios giran en torno a: Mutaciones puntuales de un nucleótido en el material genético, deleciones, inserciones, reordenamientos y translocaciones (42 – 43). Por último, el punto 3, habla de antibióticos con la capacidad de interrumpir la correcta síntesis del peptidoglucano generando casos de fragilidad osmótica y posible lisis celular como los son los  $\beta$  – lactámicos (44 – 45). La resistencia a esto se adquiere mediante mutaciones en genes que codifican enzimas relevantes al proceso de generación del peptidoglucano o bien promover el flujo del antibiótico hacia afuera de la célula con tal de mantener una concentración intrabacteriana lo suficientemente baja con tal de que la maquinaria bioquímica no se vea afectada (46).

En vista de lo anterior, es posible inferir por qué se dificulta tratar no sólo una carie, sino que también otras enfermedades asociadas a la formación de una biopelícula una vez que el consorcio de bacterias ya está formado. Probablemente, esta es una de las razones por las cuales se busca con tanto ahínco en el estudio de nuevas formas para tratar estos padecimientos, dada lo poco eficiente que han estado siendo los antibióticos con el pasar del tiempo.

### **1.3 Mecanismos de adhesión de *Streptococcus mutans***

Debido al mecanismo de colonización hacia el tejido dental, la adhesión a la superficie resulta ser el paso crítico para el éxito de la formación de la biopelícula y posterior carie. En el caso de *S. mutans*, la adhesión se basa en el empleo de apéndices proteicos conocidos como *pili* o adhesinas, filamentos micrométricos que se proyectan desde la pared celular, que operan como verdaderas anclas moleculares para unirse al tejido (47). En la naturaleza, los *pili* de bacterias han sido reportados de ser capaces de resistir perturbaciones mecánicas de decenas y centenares de pico Newtons, fuerzas que el hospedero impone para constantemente intentar desgarrar a la bacteria de la matriz a la que está unida. El movimiento de la saliva, la lengua o el cepillado son sólo un par de ejemplos de los desafíos desarrollados por el hospedero para desalojar a esta bacteria patógena (48 – 49).

La colonización de *S. mutans* al tejido depende de mecanismos dependientes e independientes de sacarosa (50). En el caso del mecanismo dependiente de sacarosa, mediante el uso de glicosiltransferasas se generan glucanos extracelulares a partir de la sacarosa del medio lo que promueve la adhesión al tejido (51), mientras que, el mecanismo independiente requiere de la participación de proteínas de superficie (52). En el caso de *S. mutans* hablamos de Ag I/II, una adhesina multifuncional que pertenece a una familia de proteínas que giran en un rango de entre 140 – 180 kDa de tamaño que asiste en la unión de la bacteria a una matriz (53). En la cavidad oral, Ag I/II interactúa con un complejo glicoproteico de aglutininas salivares conocido como SAG (figura 1) el cual es secretado por las glándulas salivares como parte del fluido salivar. SAG está conformado mayoritariamente por el receptor gp340 el cual contiene secciones repetidas de un dominio *scavenger* rico en cisteína, el que funciona como el sitio de unión para Ag I/II en un proceso de adhesión dependiente de calcio (54 – 56). La adhesina como tal se conforma de múltiples regiones conservadas (Figura 2A). Hacia la región N – terminal se encuentra una zona rica en residuos alanina que forma la región repetida de alanina (A<sub>1-3</sub>), seguido a ésta, se forma un segmento intermedio que se conoce como región variable (V), cercano al C – terminal está la zona rica en prolina, formando la región repetida de prolina (P<sub>1-3</sub>) y por último, la región C – terminal propiamente tal de 3 dominios (C1, C2 y C3) que contiene un motivo consenso LPXTG reconocido por sortasas que se encargan de unir covalentemente a Ag I/II al peptidoglucano de la bacteria (55 – 57). Estructuralmente la región repetida de alanina forma una  $\alpha$  – hélice de gran tamaño que se entrelaza con una hélice formada por la región repetida de prolina formando una estructura inusualmente larga y estrecha. La región variable (V) forma una cabeza globular rica en sábanas –  $\beta$  que se forma justo antes que se entrelacen las regiones repetidas A y P. La sección del C – terminal también siendo globular en su naturaleza se compone por 3 dominios en forma  $\beta$  – sándwiches similares a complejos IgG, siendo ésta la región más conservada en la adhesina, mostrando un porcentaje de identidad de 62% entre diferentes cepas de *S. mutans* haciéndola un atractivo blanco terapéutico (58 – 59).

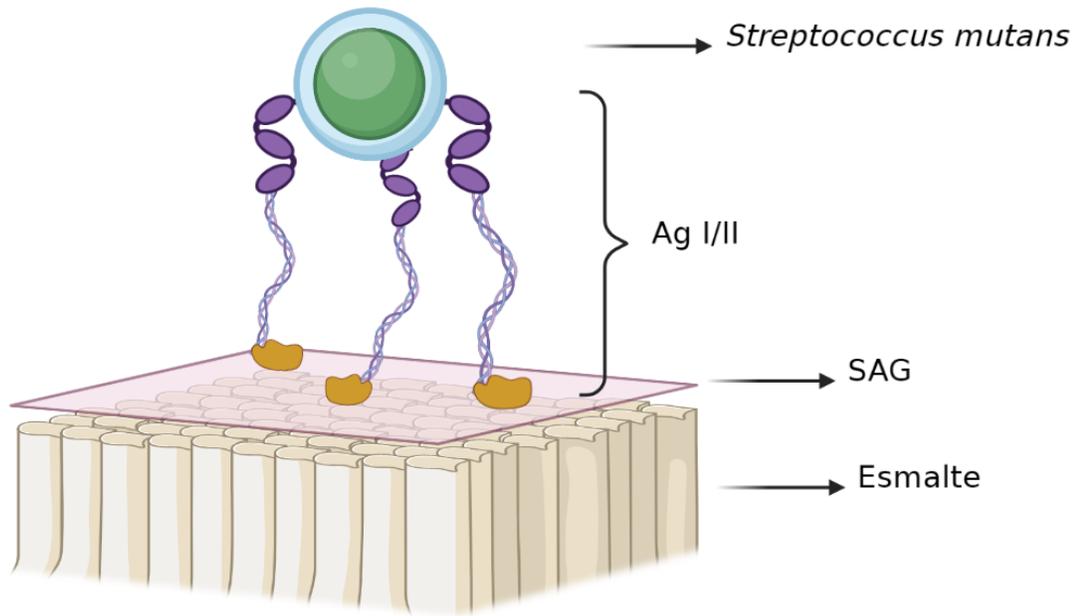


Figura 1: Esquema que muestra el mecanismo que usa *S. mutans* para adherirse al tejido dental. Mediante el uso de apéndices proteicos – *pilus* Ag I/II – *S. mutans* es capaz de adherirse con alta afinidad al complejo de aglutininas salivares libres – SAG – logrando una conformación altamente estable y extremadamente difícil de romper una vez establecida la adhesión promoviendo la colonización por otras bacterias.

La adherencia de *S. mutans* al tejido es un proceso cooperativo que depende de dos secciones de la adhesina. La región  $A_3VP_1$  que es la que da la afinidad hacia la matriz dada la presencia de un sitio de unión a hidratos de carbono que reconoce la región repetida de cisteína del receptor gp340 del SAG formando uniones de carácter multivalente por múltiples estructuras presentes en la proteína de Ag I/II. Experimentos de célula individual asociados a las fuerzas de unión de *S. mutans* a SAG mostraron que cuando múltiples uniones de  $A_3VP_1$  son expuestas a fuerzas, se establece un sistema acoplado y cooperativo capaz de lograr que la fuerza total que se aplica sobre la bacteria sea dividida en el número total de uniones formadas (60). Por otra parte, la porción C – terminal de la proteína, además de poseer la capacidad de unirse al SAG, su verdadera misión recae en darle estabilidad mecánica a la molécula encargándose de disipar perturbaciones mecánicas que amenazan constantemente por desgarrar la unión de *S. mutans* a la superficie del diente (53 – 61).

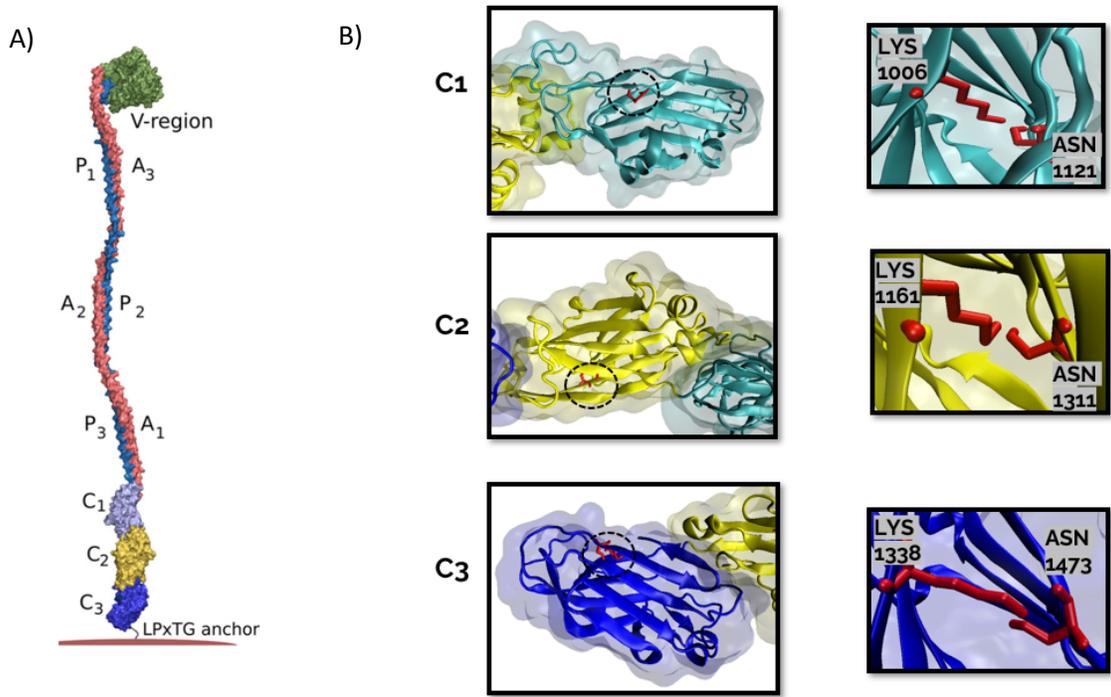


Figura 2: Esquemización de la estructura tridimensional de la adhesina Ag I/II de *Streptococcus mutans* en conjunto con la ubicación de los enlaces isopeptídicos en sus distintos dominios. (A) Modelo de la estructura tridimensional de la adhesina Ag I/II referenciada de Larson y col. (38); (B) Conjunto de figuras preparadas en PyMol mostrando la ubicación de los aminoácidos involucrados en el lace isopeptídico presente en los dominios C1, C2 y C3. El panel izquierdo muestra los 3 dominios del de C – terminal de Ag I/II; C1 (Celeste), C2 (amarillo) y C3 (azul), marcando con un círculo el lugar en dónde se forma el enlace isopeptídico, mientras que el panel de la derecha muestra una ampliación del lugar específico marcado en el panel de la izquierda denotando los aminoácidos específicos que forman el enlace.

### 1.3.1 Enlaces isopeptídicos:

La capacidad estabilizadora que tiene la porción C – terminal de la adhesina, sobre la estructura de la proteína se debe a que en cada uno de sus dominios (C1, C2, C3) presenta enlaces isopeptídicos. Este tipo de enlaces intramoleculares se observa comúnmente en bacterias de tipo Gram positivas, constituido comúnmente por un residuo de lisina y uno de asparagina junto a un residuo de aspartato que se desempeña las funciones de aminoácido catalítico para la formación del enlace. En los dominios C1, C2 y C3 de la porción C – terminal de Ag I/II, estos enlaces se forman entre los residuos, Lys1006 - Asn1121; Lys1161 - Asn1311 y Lys1338 - Asn1473 respectivamente (53). Estos enlaces isopeptídicos,

son similares a los observados en proteínas involucradas en el *pili* de otros microorganismos como SspB de *Streptococcus gordonii*, y Spy2018 de *Streptococcus pyogenes*, donde el átomo NZ de la lisina y el átomo CG de la asparagina, se unen en una reacción que libera amonio (62 – 63). Además, los enlaces de los tres dominios comparten características estructurales similares: (i) Están rodeados de aminoácidos aromáticos e hidrofóbicos, (ii) presentan un residuo de ácido aspártico que forma un enlace de puente de hidrógeno para estabilizar el complejo, (iii) En cada enlace, cercano al ácido aspártico, hay otro residuo de ácido aspártico que se coordina con un ion de  $Ca^{2+}$  (64). Sobre este último residuo, se especula que juega un papel importante en la estabilidad estructural del complejo dada su cercanía al enlace isopeptídico, y que el ion metálico es quelado por múltiples aminoácidos de una cadena lateral para estabilizar la conformación (65 – 66). Los enlaces isopeptídicos de Ag I/II cumplen la función de unir de manera covalente las hebras  $\beta$  de los distintos dominios cerrando las cadenas laterales en una conformación estable y altamente resistente a esfuerzos mecánicos.

Poco se conoce sobre las demás funciones que cumplen estos enlaces dentro de *S. mutans*, pero considerando que la presencia de enlaces isopeptídicos intramoleculares es una característica común en *pili* de otras bacterias Gram positivas, se es posible extrapolar ciertas funciones, entre ellas, la capacidad de otorgar estabilidad tanto termodinámica, como proteolítica. Estas funciones fueron demostradas usando cepas mutantes de *Streptococcus pyogenes*, las que carecían de la capacidad de poder formar un enlace isopeptídico en uno de sus dominios. En ellas se pudo observar que las cepas sin el enlace isopeptídico eran más propensas a sufrir eventos de desplegamiento en sus *pili*, cuando estas eran expuestas a variables de temperatura, siendo 30°C menos estables que las cepas silvestres. En conjunto, al realizar ensayos en los que se trataba a ambas cepas, mutante y silvestre con tripsina, una enzima peptidasa cuya función es romper enlaces peptídicos de las proteínas mediante reacciones de hidrólisis, se notó que, mientras que las cepas silvestres se mantuvieron esencialmente intactas con el pasar del tiempo de reacción, las cepas mutantes se degradaron casi completamente (67 – 68).

En vista de lo dicho anteriormente, se puede concluir que la presencia de enlaces isopeptídicos intramoleculares, tanto en *S. mutans* como en otras bacterias patógenas Gram positivas, juega un papel relevante en la virulencia dada las características que otorgan. Por lo tanto, es posible hipotetizar que aquellas estrategias que intenten interferir con la adhesión de bacterias patógenas podrían estar destinadas en afectar las proteínas del *pilus*. En el caso de *S. mutans* y la colonización del tejido dental,

intentar entorpecer con la función de Ag I/II podría ser clave para evitar la adhesión y la instalación de biopelículas que más tarde se desarrollan hasta formar las caries.

#### 1.4 Definición de la propuesta

El desarrollo de péptidos anti adhesivos es un área de investigación que ha cobrado gran fuerza cuando se habla sobre el tratamiento hacia enfermedades generadas por bacterias, en especial con la creciente crisis en el uso de antibióticos provocada por el constante surgimiento de cepas de bacterias multirresistentes. Una investigación realizada por Rivas-Pardo JA y *cols.* en la que se usaba como modelo de estudio a *Streptococcus pyogenes* demostró que la elaboración de moléculas peptídicas que bloqueen la formación del enlace isopeptídico de esa bacteria, hacen posible interrumpir las cualidades mecánicas del *pilus*, aseveración demostrada en análisis de molécula individual en la que se observó que cuando la adhesina era intervenida con el péptido, se provocaba un cambio radical en las propiedades mecánicas de la proteína mostrando una mayor longitud, lo que equivale a un mayor desplegamiento de la proteína en comparación con una silvestre cuando ambas eran expuestas a esfuerzos mecánicos (69).

A continuación, y por medio de herramientas *in silico* se diseñaron y evaluaron una serie de péptidos con la capacidad de unirse a uno de los dominios del C-terminal del *pilus* de *S. mutans*, con tal de poder intervenir con la formación natural del enlace isopeptídico entre las cadenas laterales de lisina y asparagina del dominio C2 de la porción C – terminal de Ag I/II. El desarrollo del péptido está basado en principios de intervención molecular, lo que se puede definir como un proceso realizado por pequeñas moléculas que se encarguen de entorpecer y/o intervenir un proceso asociado al metabolismo o fisiología de un microorganismo, en este caso con la formación del enlace isopeptídico (figura 3).

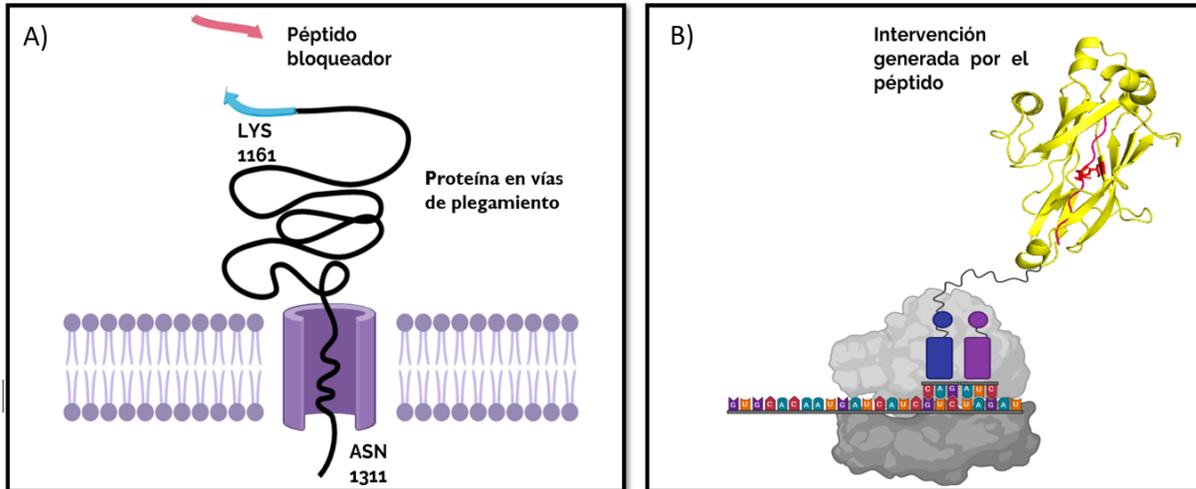


Figura 3: Esquematación del funcionamiento teórico de la propuesta. A) Proteínas de *pilus* son translocadas hacia el espacio periplasmático, lugar donde adoptan su respectiva estructura tridimensional. Es en este lugar donde las proteínas del *pilus* logran formar el enlace isopeptídico. Esta técnica, es una estrategia basada en el desarrollo de un péptido inspirado en la misma estructura de la proteína y que persigue interferir a la proteína en vía de plegamiento. Debido a que aún la proteína está emergiendo del poro—ribosoma o translocón—el péptido diseñado al poseer característica equivalente de la hebra emergente puede unirse y formar un enlace isopeptídico intermolecular. B) Modelo predictivo de cómo funciona la propuesta. Se espera que el péptido sea capaz de atacar a la sección deseada cuando termina la traducción y esta sale del ribosoma para adoptar su conformación tridimensional.

El péptido será diseñado en base a estructuras propias de la proteína para luego ser evaluado mediante herramientas *in silico*. Se intenta imitar y optimizar la conformación que tiene la hebra  $\beta$  complementaria que forma el enlace, buscando diseñar una molécula que sea más afín y estable que su contraparte silvestre.

El diseño constará de dos secciones, siendo la primera de ellas una caracterización mecánica del C – terminal de Ag I/II. La información disponible respecto cómo funciona mecánicamente el C – terminal es relativamente exigua, dado que la mayoría de los experimentos sobre transducción de fuerzas se hacen sobre la región A<sub>3</sub>VP<sub>1</sub>, es por eso, que previo al diseño del péptido se hace necesario saber con claridad cuál es el aporte que da cada enlace isopeptídico dentro de esta porción de la adhesina, a manera de poder saber cuál de ellos es el más relevante respecto a la tarea de estabilizar el proceso de unión. Como segunda etapa, y una vez que se sabe cuál de los dominios es el más influyente en la estabilidad, se procederá a diseñar un péptido estructuralmente mimético a la hebra que forma el

enlace en el dominio identificado con tal de poder maximizar la especificidad del péptido por la sección que se intervendrá y aumentar la probabilidad que de que el péptido forme un enlace isopeptídico entre las hebras silvestres, haciendo posible una unión de carácter irreversible.

## 2 HIPOTESIS Y OBJETIVOS

### 2.1 Hipótesis

La optimización *in silico* de péptidos bloqueadores de enlaces isopeptídicos puede aumentar significativamente su capacidad de unirse a la adhesina Ag I/II de *Streptococcus mutans*.

### 2.2 Objetivos

- **General:**
  - Generar mediante análisis *in silico*, un péptido con más afinidad de acoplamiento en comparación con las estructuras silvestres del C – terminal de Ag I/II.
- **Específicos:**
  - Caracterizar mecánicamente *in silico* la función que cumplen los enlaces isopeptídicos de la porción C – terminal de la proteína Ag I/II mediante dinámicas moleculares.
  - Diseñar y refinar péptidos con tal de optimizar sus energías de unión usando técnicas *in silico* de acoplamiento molecular.
  - Evaluar *in silico* la estabilidad de los péptidos que serán candidatos para intervención molecular por dinámicas moleculares.

## 3 MATERIALES Y MÉTODOS

### 3.1 Material utilizado

#### 3.1.1 Software empleados:

- Campos de fuerza oplsAA y SIRAH: Los campos de fuerza utilizados representan un conjunto de fórmulas matemáticas y parámetros proteicos usado para simulación *in silico* de biomoléculas (70).
- Editor de texto Geany: Editor multipropósito usado para leer archivos de texto independiente del formato (71).
- Editor de texto Kate: Editor utilizado para el reemplazo de múltiples términos de manera simultánea (72).
- GROMACS: Software de uso libre usado para la simulación dinámica de biomoléculas (73).
- Herramienta de trazado de gráficos XMGRACE: Programa de graficado en 2D que permite la fácil visualización de gráficos de línea e histogramas (74).
- Microsoft Office Excel versión 2021: Herramienta usada para análisis a partir de grandes cantidades de datos (75).
- PyRx Autodock Vina: Programa de acoplamiento molecular y estudio de interacciones proteína – proteína (76).
- Python versión 3.10: Lenguaje de programación usado para automatizar ciertas tareas (77).
- Sistema operativo Linux 20.4
- Sistema Operativo Windows 10
- UCSF Chimera: Programa de visualización tridimensional interactiva que permite hacer predicciones de estructuras peptídicas (78).
- Visual Molecular Dynamics (VMD): Programa de visualización molecular para representar y analizar dinámicas moleculares (79)
- Visualizador molecular Pymol: Visualizador multi plataforma de macromoléculas (80).

#### 3.1.2 Equipos:

- Unidad de procesamiento de cálculos de alto rendimiento:

- SAMSUNG SSD 980 PRO 1TB M.2 2280 PCI-Express 4.0 x4 NVMe 1.3c Samsung V-NAND MZ-V8P1T0B/AM
- AMD Ryzen 7 5800X 4.7GHz Turbo 8Core 16 Hilos 105W TDP Pci-e 4.0 Zen 3
- CORSAIR DDR4 32GB 3200Mhz Vengeance RGB Pro 2x 16GB CMW32GX4M2D3200C18
- ASUS ROG Strix GeForce RTX 3080 10GB GDDR6X

## 3.2 Metodología

### 3.2.1 Caracterización mecánica

Dado a que la información respecto al aporte en la estabilidad mecánica que ofrecían los enlaces isopeptídicos dentro de la adhesina Ag I/II es limitada, el primer paso para la implementación de un péptido que interfiera con la formación de enlaces isopeptídicos es determinar cuál de los tres dominios es que aporta en mayor grado a la elasticidad del *pilus*. Este proceso fue hecho mediante un protocolo de dinámicas moleculares dirigidas, que permiten aplicar fuerzas sobre un sistema molecular, imitando los embates mecánicos que experimenta el *pilus* durante la adhesión. Para ello usamos GROMACS, software de uso libre dedicado para realización y análisis de dinámicas moleculares en conjunto a un sistema operativo de Linux, por razones de compatibilidad con las líneas de código utilizadas. Además, para la realización de las dinámicas moleculares se emplearon los campos de fuerza SIRAH y OPLS-AA, los cuales se componen por un conjunto de combinaciones de fórmulas matemáticas y parámetros proteicos que se usan para describir las energías asociadas a una proteína en función de sus coordenadas atómicas. Es importante mencionar que tanto SIRAH como OPLS-AA no son capaces de reconocer la formación de un enlace isopeptídico de por sí, dado esto, las características necesarias para formar un isopéptido entre residuos de lisina y asparagina fueron parametrizadas en el campo de fuerza OPLS-AA, tarea que se hizo con anterioridad por los integrantes del laboratorio.

Con en el objetivo de lograr la caracterización mecánica de los dominios C1, C2 y C3 se siguió el esquema de trabajo representado en la figura 4. Se empleó el modelo cristalográfico de la adhesina Ag I/II, disponible en la base de datos RCSB *Protein Data Bank* bajo el código 3QE5. La unidad asimétrica presente en el cristal de la proteína posee dos cadenas repetidas de la porción que contiene los dominios de C1-C2-C3 de Ag I/II, además de otros componentes – tales como iones y coprecipitadores – que incluidos durante la cristalización (figura 5A). La presencia de estos componentes puede interferir

con el desarrollo de la futura dinámica molecular, por lo que antes de pasar a ese procedimiento se hace necesario removerlos y dejar en el archivo de coordenadas solo la proteína Ag I/II (figura 5B). Para esta tarea se usaron líneas de comando en el visualizador molecular VMD con lo que no sólo se hizo posible limpiar el cristal, sino que también poder escribir un nuevo archivo en formato “.pdb” con la nueva información.

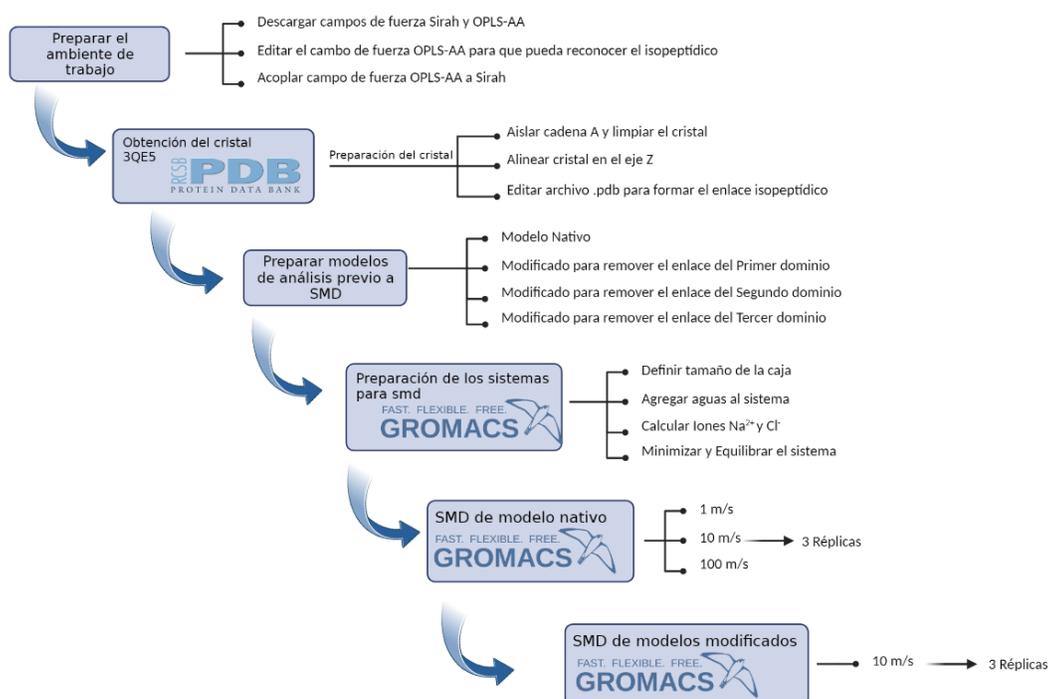


Figura 4: Esquema del flujo de actividades seguido para lograr la caracterización mecánica de los dominios C1, C2 y C3 de la porción C – terminal de la proteína Ag I/II.

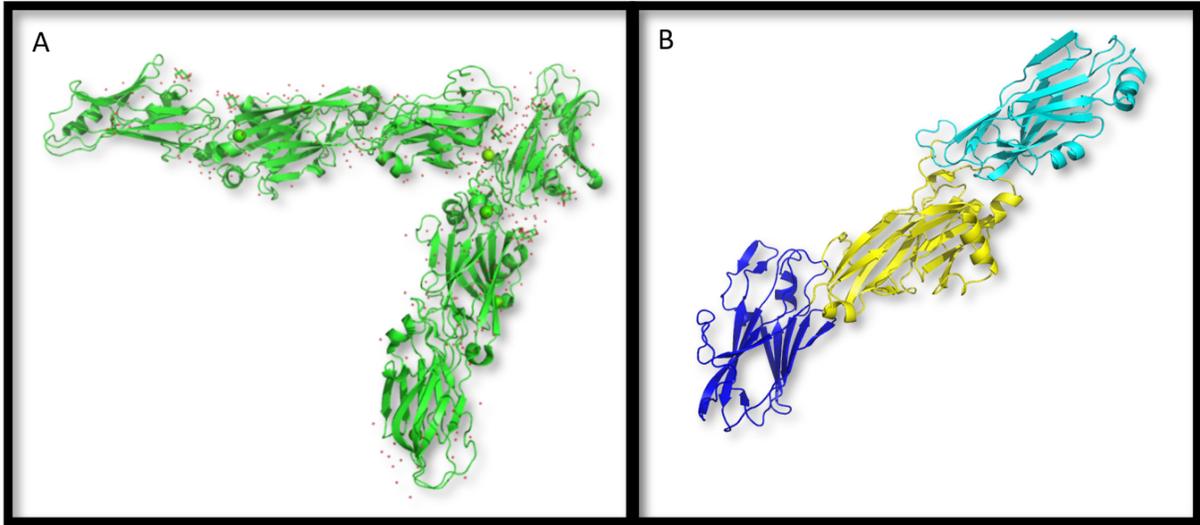


Figura 5: Comparación antes y después del proceso de limpieza que se le da al cristal previo al desarrollo de la dinámica molecular con tal de aislar una sola de las cadenas además de las aguas e iones presentes en el cristal 3QE5. A) Cristal 3QE5 descargado desde *Protein Data Bank*; B) Cadena A del cristal 3QE5 luego de usar los comandos en VMD para limpiarlo. Los dominios están marcados con colores, siendo; C1 de color celeste; C2 de color amarillo y C3 de color azul.

La proteína del nuevo archivo de coordenadas se encuentra en un plano de ejes XYZ; para el desarrollo de la dinámica se hace necesario que este sea alineado y centrado en el eje Z en base a su centro de masa, con tal de que, al momento de aplicar las fuerzas para estirar la proteína, esto se haga solamente acorde a ese eje en cuestión con la finalidad de optimizar recursos computacionales en la fase de estiramiento (revisar caja S1 del aparatado de anexo para la línea de comando utilizada). Con tal de que los campos de fuerza que se usaron para realizar la dinámica pudieran reconocer un enlace isopeptídico, el código de los aminoácidos que forman el enlace fueron modificados. En este caso, el código de la lisina se modificó desde “LYS” a “LYY”, mientras que la asparagina se pasó de “ASN” a “ASA”, nomenclatura que fue decidida previamente por consenso en el laboratorio cuando se parametrizaron los datos de formación del enlace isopeptídico dentro del campo de fuerza oplsAA. Esto se realizó usando el editor de texto Geany, modificando los residuos de lisina y asparagina en el archivo de coordenadas que participan en la formación de los tres enlaces isopeptídicos presentes en la proteína.

Para determinar cuál de los enlaces isopeptídicos era el más relevante en cuanto a dar estabilidad mecánica en la proteína se establecieron 4 modelos de estudio: (i) Modelo silvestre que mantenía la capacidad de formar sus tres enlaces isopeptídicos; (ii) sistema modificado para no formar el enlace

Lis1006 - Asn1121; (iii) sistema modificado para no formar el enlace Lis1161 - Asn1311; (iv) Sistema modificado para no formar el enlace Lys1338 - Asn1473.

El archivo de entrada para el desarrollo de la dinámica molecular corresponde al de la cadena polipeptídica preparada como se detalló en párrafos anteriores, el cual debe ser transformado a un formato de archivo que pueda ser leído por el programa GROMACS. Este archivo denominado como topología, contiene todos los atributos constantes de cada átomo que compone la proteína, incluyendo: Ángulos, torsiones, enlaces, distancias entre átomos y diversas interacciones que se encuentran parametrizadas dentro de los campos de fuerza utilizados. Una vez se crea la topología, el paso siguiente es generar el sistema que contendrá a nuestro modelo de estudio. Para ello, se generó un espacio que pueda simular de la mejor forma posible un ambiente biológico, que además contenga un conjunto de leyes físico químicas y energéticas con tal de poder predecir como se comportaría la proteína al momento de aplicar una magnitud que amenace con desplegar la molécula, espacio que se definirá como sistema acuoso de aquí en adelante. Parte importante de la creación de este espacio es definir las dimensiones de este, dado que al llegar a la fase de estiramiento se espera que la proteína alcance una magnitud nanométrica mayor al inicial producto de las fuerzas ejercidas sobre esta. En caso de que la cadena polipeptídica fuera a salir del sistema acuoso el proceso fallaría. El cálculo de las dimensiones se hizo usando la siguiente formula:

$$( \textit{Aminoácidos totales} - \textit{Contribución de los enlaces} ) \times 4$$

$$(486 - (115 + 150 + 138)) \times 4$$

$$(486 - 403) \times 4$$

$$83 \times 4 = 332 \approx 340$$

Ecuación 1: Cálculo de la longitud del sistema en el eje Z usando como ejemplo al modelo silvestre.

El número de aminoácidos totales se obtuvo con la información presente en el archivo del cristal, mientras que la contribución de los enlaces se calcula verificando cuantos aminoácidos existen entre la lisina y la asparagina que forman el enlace en cada dominio de Ag I/II, dado que se consideran como secciones inextensibles en la proteína por lo que no deberían afectar en la longitud final. Cabe recalcar que los datos que se pueden observar en la fórmula 1 corresponden al cálculo que se hizo para el

sistema silvestre, es decir aquel que toma en consideración todos los enlaces, aun así, la única diferencia entre el resto de los modelos de estudio es que para cada uno se remueve la contribución del enlace que se estudia en el momento, por ejemplo, para el sistema del modelo en donde se remueve el primer enlace isopeptídico, sólo se toma la consideración la contribución del segundo y tercer enlace al momento de aplicar la fórmula. El cálculo especificado se hace para definir la longitud del sistema sólo en base al eje z, decisión no trivial que se hace en base a dos puntos: La proteína al punto en dónde se construye el sistema está alineada en el eje z y al momento se aplicará la perturbación para desplegar la proteína se hará en torno a ese eje particular. Esto hace que el sistema sea más pequeño en cuanto a sus dimensiones totales y de esta manera poder optimizar recursos computacionales. El espacio creado, debe ser llenado ahora con moléculas de agua con tal de que la proteína no se encuentre en el vacío. El sistema que se usa para realizar las dinámicas corresponde a un sistema híbrido con dos tipos de moléculas de agua: Atomística (SOL) que rodean a la proteína con tal de que los cálculos del comportamiento de la proteína sean más de mejor resolución y grano grueso (WT4) que ocupan el resto del espacio (Figura 6). Al igual que la decisión sobre las dimensiones de la caja el hecho de utilizar un sistema híbrido se hace con tal de ahorrar recursos computacionales y reducir el tiempo de análisis. La estación de trabajo calcula de manera autónoma el número de cada tipo de aguas que irán al sistema, valor que debe ser luego recuperado con líneas de código con tal de poder seguir con el proceso.

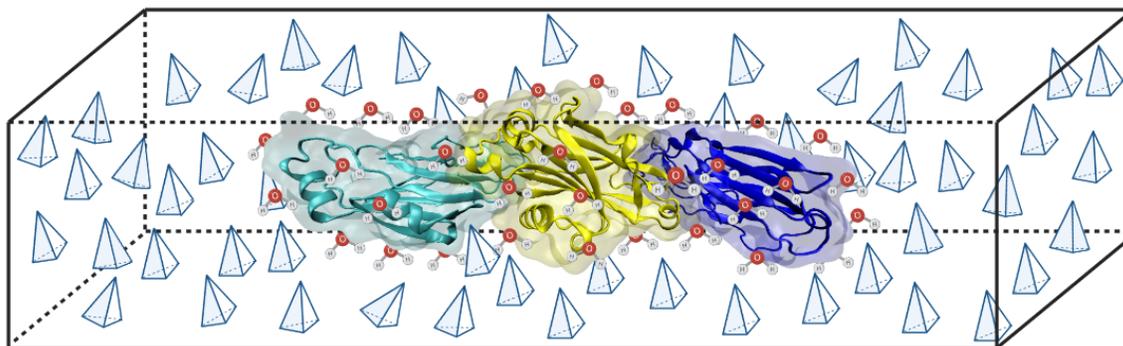


Figura 6: Representación gráfica del sistema construido. Las moléculas de agua que rodean directamente a la proteína representan las aguas tipo atomísticas mientras que las pirámides que completan el resto del espacio corresponden a aguas de tipo "grano grueso".

Una vez recuperado el número de aguas es necesario calcular el número de iones que tendrá el sistema con tal de poder neutralizar la carga neta que tiene la proteína, valor que entrega la terminal al momento de generar la topología. Se agregan iones sodio ( $\text{Na}^{2+}$ ) e iones cloro ( $\text{Cl}^-$ ), acorde a las siguientes dos fórmulas:

$$Total\ 1 = \frac{(N^{\circ}\ de\ aguas\ SOL) \times Concentración}{56}$$

$$Total\ 1 \pm \left( \frac{Carga\ total}{2} \right)$$

Ecuación 2: Fórmula adaptada previamente en el laboratorio usada para calcular el número de iones sodio y cloro que van dentro del sistema. Se usa el número de aguas atomísticas (SOL) dado la relevancia que poseen al interaccionar directamente con la proteína, en conjunto con la concentración de sales presentes en el cristal al momento de la purificación (de no saber el valor se usa un valor general de 0,15), además de un valor constante de 56.

El cálculo anterior entrega 2 resultados, y el cómo definir el valor que le corresponde a cada ion depende de la carga misma del sistema. En caso de que la carga neta del sistema sea de valor negativo, el valor mayor que entrega la fórmula correspondería a iones de sodio mientras que el menor al cloro para neutralizar usando iones de carga positiva, y por otra parte si el sistema tuviera carga neta de valor positivo la definición de iones se hace al revés. Para casos específicos de este experimento la carga neta de los sistemas era negativa, por lo que se usó el valor mayor para los iones sodio mientras que el otro se usó para los iones cloro. Cabe recalcar que el cálculo de iones se hace sólo en base a las aguas atomísticas dado a que estas rodean a la proteína. Un punto importante respecto a la preparación del sistema acuoso, es que como se realizó bajo los parámetros del campo de fuerza oplsAA, el archivo se escribe en términos bajo la nomenclatura asociada a este campo de fuerza, términos que no pueden ser leídos por el campo de fuerza SIRAH al momento de preparar los archivos para la realización de la SMD. Para resolver lo anterior, se usó el editor de texto Kate, el cual permite reemplazar de manera rápida los términos del campo de fuerza oplsAA por aquellos compatibles con SIRAH, conversión que puede ser vista en la tabla suplementaria S1.

Una vez eso terminado tenemos todas las partes que componen nuestro sistema acuoso, la proteína, aguas e iones, por lo que se puede pasar al siguiente paso que es de minimizar. El minimizado se encarga de llegar hasta una conformación favorable de energía que se denomina el mínimo de energía local y lo hace aplicando el algoritmo de Montecarlo, buscando posiciones de átomos al azar en la proteína hasta encontrar al conformero más estable y de menor energía (81). Luego, el sistema deber pasar al proceso de equilibrado, donde se aplican variables de presión y de temperatura hasta un punto en donde la proteína alcance un estado estacionario y deseable al momento de seguir colectando

información. Adicionalmente, una vez termina el proceso de equilibrado es necesario informarle al programa los puntos de estiramiento por los cuáles se aplica la variable energética las cuáles pueden ser a velocidad o fuerza constantes dependiendo del caso, e independiente de la variable que se elija, la energía se aplica en direcciones contrarias una a la otra con tal de poder estirar la proteína. En este caso los puntos de estiramiento definidos corresponden al primer y último carbono  $\alpha$  de la secuencia peptídica de Ag I/II.

En un inicio, las dinámicas moleculares dirigidas se realizaron aplicando magnitudes de velocidad constante empleando el modelo de la proteína silvestre a velocidades de 1, 10 y 100 m/s. Esta exploración a diferentes velocidades funciona como un barrido que permite determinar cuál de las magnitudes entregaba la mayor calidad de resultados en una menor escala de tiempo real. Una vez identificada la mejor velocidad para trabajar, los SMD de tanto el modelo silvestre como de los modelos modificados para no formar enlace se hacen en base a ese parámetro usando 3 réplicas por cada modelo.

La dinámica molecular entrega dos resultados siendo el primero dos archivos organizados por columnas, uno representando la extensión de la proteína en nanómetros y otro con las fuerzas asociadas al proceso de desplegamiento que se entregan en unidades de kJ/mol/nm. Estas últimas son más tarde transformadas a unidades de fuerza en la escala de los pico Newton (pN), proceso que se realiza simplemente dividiendo la columna por un factor de conversión de 1,661. El segundo es un archivo que contiene todos los números asociados al proceso de la dinámica el cuál fue transformado usando líneas de código con tal de poder tener un archivo de trayectoria con el cual se puede visualizar como se resolvió la dinámica a lo largo del proceso, haciendo posible ver de manera directa el desplegamiento de la proteína. El análisis de cuál de los dominios corresponde al más influyente, se hizo analizando la trayectoria y el gráfico en conjunto, con tal de poder ver en cuál estado se registraban menores magnitudes de fuerza y mayores eventos de desplegamiento cuándo se removían los enlaces isopeptídicos. Una vez identificado el enlace más relevante se procede al diseño del péptido.

### **3.2.2 Diseño del péptido**

El diseño del péptido se realiza mediante el uso de tres programas principales: Python, UCSF Chimera y PyRx autodock Vina. Mediante el uso de estas tres herramientas fuimos capaces de generar listas de péptidos generados al azar que se llamaron como iteraciones, las cuáles fueron convertidas a un

formato .pdb mediante una predicción de estructura tridimensional que finalmente fueron evaluados mediante técnicas de acoplamiento molecular (Figura 7).

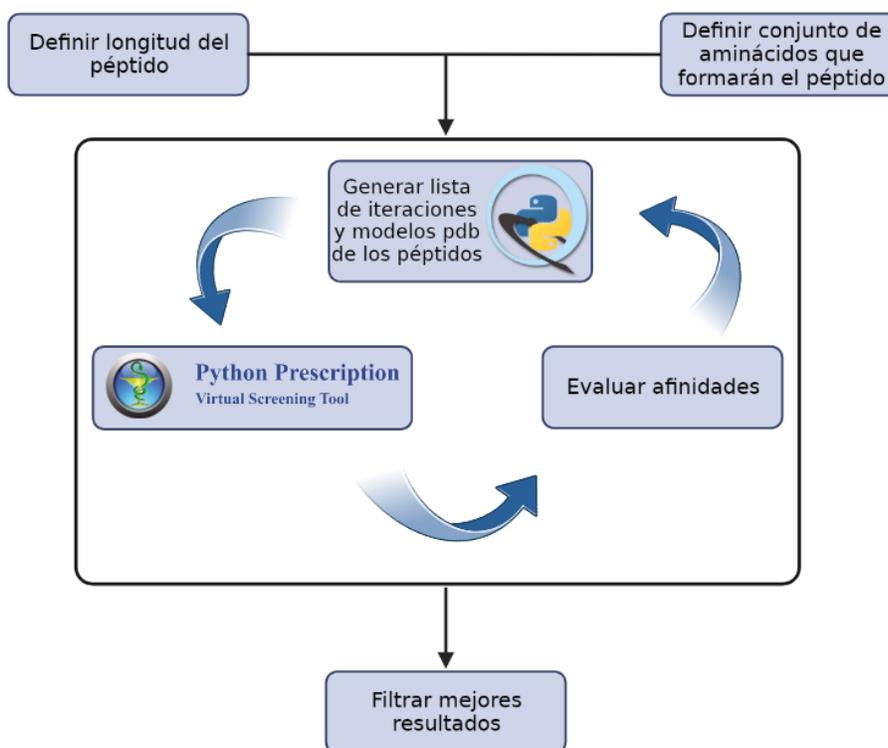


Figura 7: Esquema del flujo de actividades seguido para generar la lista de péptidos que serán evaluados mediante análisis de acoplamiento molecular.

El proceso de diseño del péptido parte inicialmente con definir la longitud y los aminoácidos con los que se construirá la molécula. La longitud del péptido se decide en torno al número de aminoácidos que constituyen la hebra  $\beta$  que contiene la asparagina con la que se forma el enlace de manera nativa en Ag I/II, usando como referencia el dominio definido en el paso anterior.

Una vez identificada la longitud del péptido se usa el código de Python para genera una lista de péptidos posibles en base a un conjunto de aminoácidos definidos por el usuario (figura 8).

```

from itertools import product
import random

secuencia = {'1': ['R', 'H', 'K', 'D', 'E', 'S', 'T', 'N', 'Q', 'C', 'G', 'P', 'A', 'V', 'I', 'L', 'M', 'F', 'Y', 'W'],
             '2': ['R', 'H', 'K', 'D', 'E', 'S', 'T', 'N', 'Q', 'C', 'G', 'P', 'A', 'V', 'I', 'L', 'M', 'F', 'Y', 'W'],
             '3': ['R', 'H', 'K', 'D', 'E', 'S', 'T', 'N', 'Q', 'C', 'G', 'P', 'A', 'V', 'I', 'L', 'M', 'F', 'Y', 'W'],
             '4': ['R', 'H', 'K', 'D', 'E', 'S', 'T', 'N', 'Q', 'C', 'G', 'P', 'A', 'V', 'I', 'L', 'M', 'F', 'Y', 'W'],
             '5': ['N'],
             '6': ['R', 'H', 'K', 'D', 'E', 'S', 'T', 'N', 'Q', 'C', 'G', 'P', 'A', 'V', 'I', 'L', 'M', 'F', 'Y', 'W'],
             '7': ['F'],
             '8': ['V'],
             '9': ['L'],
             '10': ['V']}

```



1	2	3	4	5	6	8	7	9	10
X	X	X	X	<b>N</b>	X	X	X	X	X
R	R	R	R		R	R	R	R	R
H	H	H	H		H	H	H	H	H
K	K	K	K		K	K	K	K	K
D	D	D	D		D	D	D	D	D
E	E	E	E		E	E	E	E	E
S	S	S	S		S	S	S	S	S
T	T	T	T		T	T	T	T	T
N	N	N	N		N	N	N	N	N
Q	Q	Q	Q		Q	Q	Q	Q	Q
C	C	C	C		C	C	C	C	C
G	G	G	G		G	G	G	G	G
P	P	P	P		P	P	P	P	P
A	A	A	A		A	A	A	A	A
V	V	V	V		V	V	V	V	V
I	I	I	I		I	I	I	I	I
L	L	L	L		L	L	L	L	L
M	M	M	M		M	M	M	M	M
F	F	F	F		F	F	F	F	F
Y	Y	Y	Y		Y	Y	Y	Y	Y
W	W	W	W		W	W	W	W	W

Figura 8: Representación gráfica de como funciona el código utilizado en python para generar la lista de péptidos. Con el código de python ilustrado fuimos capaces de generar una lista de péptidos generados al azar en base a todas las combinatorias posibles que se pueden formar en base a la longitud del péptido y el conjunto de aminoácidos definidos por cada posición.

El funcionamiento del código consiste en asignar un número a cada carácter del péptido y a cada número otorgarle un conjunto de aminoácidos. De esta manera mediante el uso de la librería “itertools”, se generarán todas las iteraciones posibles de péptidos en base a los aminoácidos definidos en cada posición. Se dejará como término fijo un aminoácido de asparagina cuya posición será definida dependiendo del dominio descrito en el paso de caracterización. Adicionalmente, para reducir la cantidad de iteraciones que arroja el código se decidió dejar como términos fijos los 4 últimos aminoácidos del péptido usando una secuencia FVLV. Esto se decidió usando como referencia una investigación en la que se demostró la existencia de una secuencia peptídica dentro del mismo C-terminal de Ag I/II con la capacidad de inhibir competitivamente la adhesión, dentro del cual la

secuencia FVLV juega un importante papel en la inhibición (82). Al terminar la ejecución del código se genera con un archivo de texto con todas las iteraciones generadas, de la cual se hizo una selección al azar de 500 péptidos los cuales fueron transformados a formato .pdb con el software UCSF Chimera, el cuál hará una predicción de cómo sería la estructura tridimensional de péptido en base a su secuencia, con tal de poder realizar el acoplamiento molecular (figuras 9, 10). Adicionalmente en esa sección, se preparan los modelos control que servirán para evaluar la efectividad de los péptidos de la lista, los cuáles consisten en dos imitaciones de la hebra  $\beta$  complementaria silvestre una generada en Chimera y la otra aislada del cristal de Ag I/II para que conserve su configuración estructural, proceso que se hace usando las funciones de Pymol con las cuales se permite descargar secciones específicas de un modelo proteico en formato .pdb y por último una cadena de poli glicina.

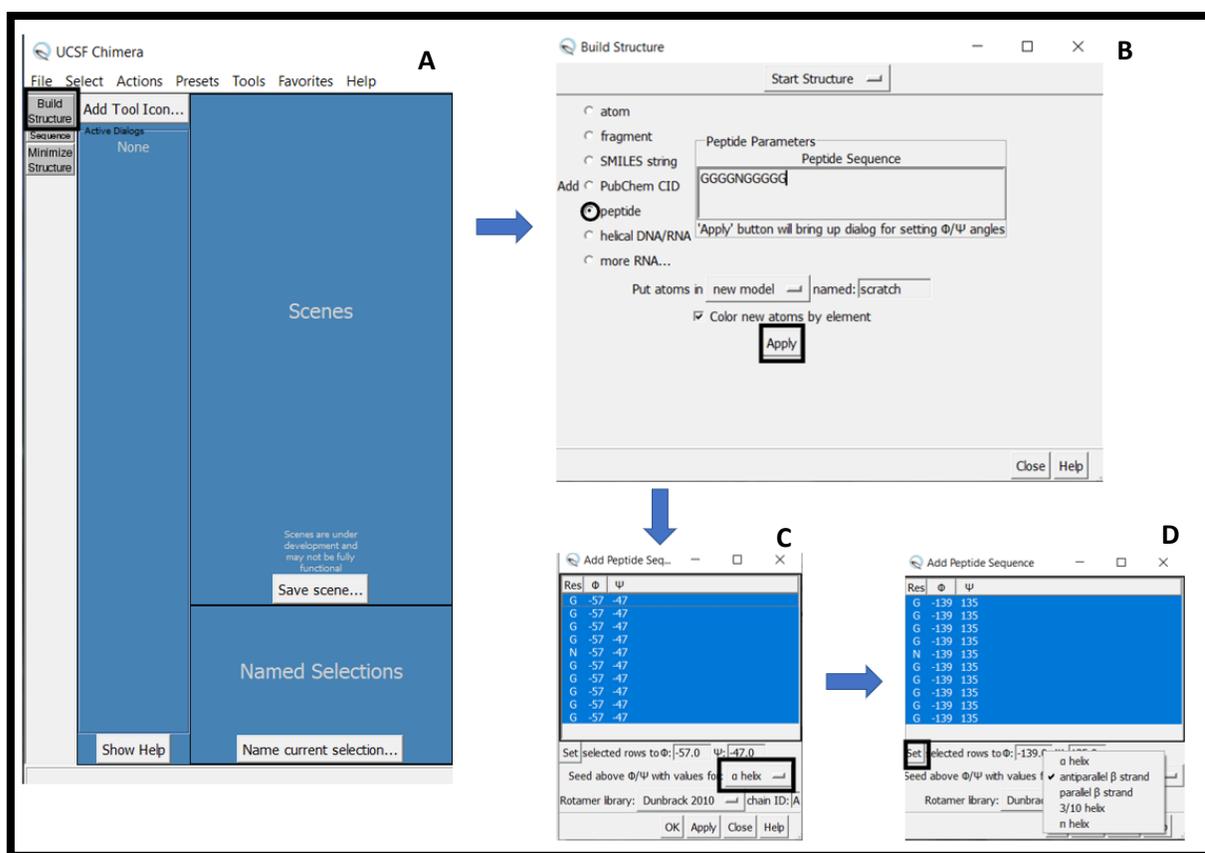


Figura 9: Camino utilizado en el software UCSF Chimera para generar un archivo .pdb en base a la predicción de estructura tridimensional a partir de un péptido en formato texto. (A) Ventana en dónde inicia el proceso; (B) ventana desplegada luego elegir opción "build structure" que pide como entrada la secuencia escrita del péptido que se desea generar; (C – D) ventana desplegable en dónde se elige el tipo de cadena peptídica que se desea generar con los ángulos asociados a ésta.

El proceso que se sigue luego de abrir el software de Chimera empieza apretando el botón “*Build Structure*” que desplegará una nueva ventana mostrando las diferentes que opciones de predicción de estructuras que tiene el programa. Para casos de este trabajo le damos a “*peptide*”, lo que liberó una caja de texto en la que debemos poner la secuencia que conforma el péptido que se desea crear para luego apretar el botón “*apply*”. Esto desplegará otra ventana en las que se muestran los ángulos  $\phi$  y  $\psi$  en conjunto con el tipo de estructura secundaria que adopta el péptido al momento de generarse, en este caso, se elige una hebra antiparalela y se la da al botón “*set*”. Luego de fijar esos parámetros se aprieta el botón “*apply*” en la ventana donde se definió el ángulo y la estructura, para que el programa empiece con la predicción del modelo tridimensional el cuál se verá reflejado en la ventana principal del programa (Figura 10). Generada ya la estructura, se usó la función “*minimize structure*” fijada con los parámetros basales del sistema para hacer un minimizado del péptido y permitir estabilizar la conformación de este. Finalmente, se guarda el archivo como formato .pdb y se guarda para el futuro acoplamiento molecular.

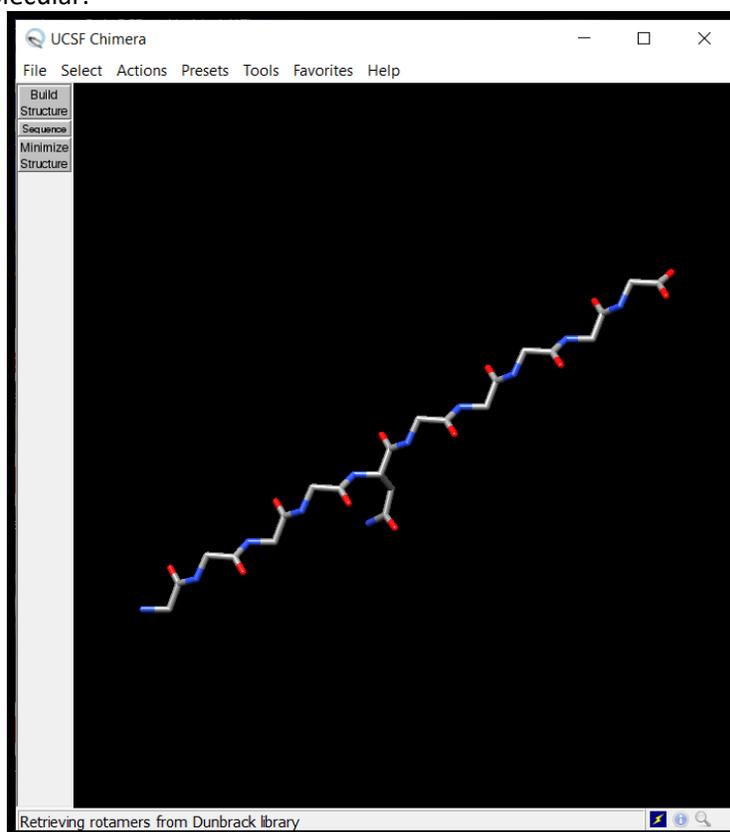


Figura 10: Forma en el que el software Chimera entrega los resultados de la predicción de estructura secundaria a partir de una secuencia peptídica. Al terminar el proceso de generación, UCSF Chimera entrega como resultado final el péptido en formato de visualización “lycorice”, el cuál puede ser guardado como formato pdb o pasar por otros procedimientos disponibles en el software dependiendo de las necesidades del usuario.

El desarrollo del acoplamiento molecular se realizó con el software PyRx, una herramienta de visualización virtual ampliamente usado en el desarrollo de fármacos mediante análisis predictivos de como una molécula se uniría a un sitio de unión o receptor específico, el cual requiere de dos archivos para poder funcionar, una macromolécula que en este caso sería el cristal de Ag I/II y un archivo ligando que serían los péptidos que se desean evaluar. Previo a ello, el cristal de Ag I/II tuvo que ser editado con tal de que poder abrir la sección de la hebra  $\beta$  que se desea intervenir, dado que por la configuración estructural que tiene naturalmente, dicha sección se encuentra protegida, formándose una especie de bolsillo que funcionará como el sitio de unión para el péptido. El bolsillo en cuestión se hizo usando un protocolo de SMD similar al aplicado durante la caracterización mecánica, el cual se usó para desplegar la sección que protege a la cadena que se deseaba intervenir formando así el espacio necesario como para que el péptido se pudiera unir.

Luego de ello, el archivo del bolsillo, los 500 péptidos, y los controles a analizar fueron subidos al software con tal de poder fijar los parámetros de análisis. Importante a mencionar es que el programa de PyRx tiene un parámetro especial para los archivos de ligando bajo el nombre de grados de torción, el cual define usando una escala numérica las libertades estructurales que puede tomar el péptido para unirse al bolsillo haciendo la diferencia entre un acoplamiento rígido y uno flexible. En este caso, se realizó un acoplamiento flexible usando preliminarmente 10 grados de torsión excepto para el control, aislado del cristal el cuál se dejó con 0 grados de torción. Luego de seleccionar los elementos con los que se realizará el acoplamiento el sistema pide generar un espacio de análisis que teóricamente funciona de la misma manera que la caja formada previo al desarrollo de las SMD. Cuando se delimita el espacio, el programa fija un área específica de lugares en dónde se puede unir o no la molécula que se evalúa. En el caso de esta investigación se trabaja con un acoplamiento molecular ciego, por lo que, al momento de definir el espacio de análisis, en vez de delimitar un área específica, se fija como espacio a toda la proteína. Luego de ello, se fijó el nivel de exhaustividad con el cuál el programa desarrolla el proceso, el cuál fue igual a 10. Finalizado el proceso, se entregan dos resultados, un archivo con las coordenadas en dónde está ubicado el péptido en el espacio y una tabla con las afinidades asociadas al proceso la cual se entrega en unidades de kcal/mol. Las afinidades se entregan en valores negativos los cuáles entre más negativo sea el valor, más afín es el ligando por el sitio en donde ocurrió la unión, lo que se traduce como la energía liberada durante el proceso, y en contraste la energía necesaria para romper la unión.

De los resultados entregados, se hará una selección de los péptidos que hayan presentado valores de afinidad más altos en comparación con los controles, con tal de poder refinarlos. El proceso se hace realizando un proceso denominado como escaneo de alanina el cual fue acondicionado para que funcionará en condiciones *in silico*, con tal de evaluar que aminoácidos dentro de los mejores resultados contribuyen de manera positiva y negativa a la afinidad final. La premisa de esta metodología consiste en mutar los distintos aminoácidos que forman el péptido por una alanina uno a la vez formándose distintas variaciones dependiendo de la longitud que tiene el péptido, variaciones con las que se hará un acoplamiento molecular de la misma naturaleza del anterior con tal de ver el efecto que tiene la remoción de uno de los aminoácidos en la afinidad del proceso (figura 11).

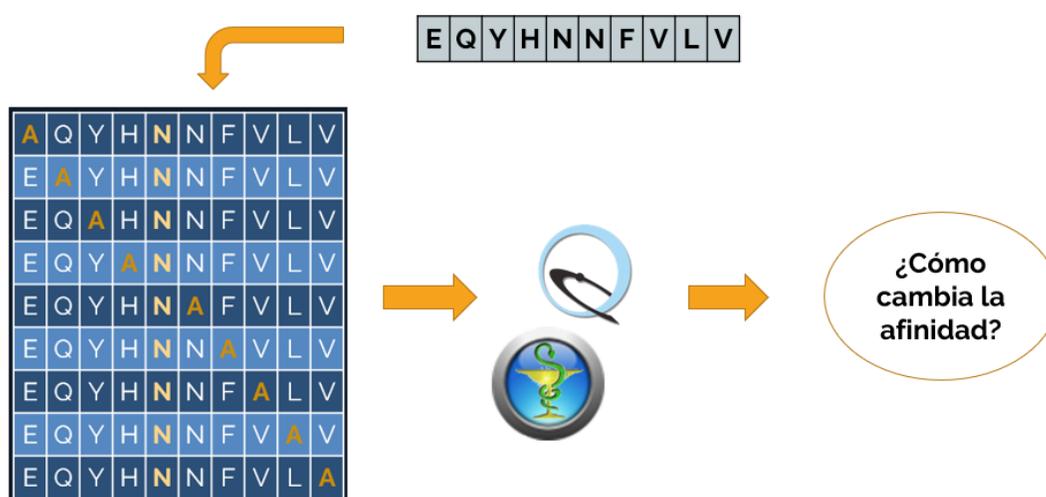


Figura 11: Esquema mostrando el cómo se realizó el escaneo de alaninas. Las distintas variaciones del péptido evaluado fueron generadas manualmente para luego predecir su estructura tridimensional mediante Chimera, estructuras que luego fueron evaluadas mediante acoplamiento molecular para verificar como variaba la afinidad final del proceso.

Como bien se explica en la figura 11 las variaciones se generan en Chimera y son evaluadas en PyRx con tal de ver como varía la afinidad. Para facilitar la visualización se hicieron gráficas en plantillas de Excel en dónde se deja la afinidad original del péptido como parámetro basal. Dependiendo de los resultados, cuando se reconoce que la ausencia de un aminoácido no causa variaciones en la afinidad basal, o bien se presente un valor más alto en comparación, se usará el código de Python para iterar la posición por los diferentes 20 aminoácidos y se hará un acoplamiento molecular para evaluar con que aminoácido en esa posición se logra el mayor grado de afinidad.

### 3.2.3 Evaluación de los péptidos

Con la evaluación finalizada, se descargaron los mejores conformeros tanto de los péptidos como de las hebras control con tal de poder construir modelos ligando-proteína para análisis de SMD. Usando el visualizador molecular Pymol se subieron los archivos del bolsillo y los conformeros simultáneamente para formar un sólo archivo que contenga las coordenadas espaciales de ambas moléculas el cual fue descargado en formato .pdb que funcionará como archivo de entrada para la dinámica molecular.

El procedimiento para formar el sistema sigue el mismo patrón que el mostrado en el punto 3.2.1 con algunas diferencias puntuales: la primera es que el tamaño de la caja en la línea de código se define como X10; Y10 y Z150 (proceso destacado en la figura S2). La longitud del eje Z se redujo ya que no se espera un desplegamiento de la proteína y tampoco que el péptido fuera a salir de las dimensiones definidas. Respecto al momento en el que se definen los puntos de estiramiento, se definió el primer carbono  $\alpha$  de la proteína como punto de estiramiento 1 y el último carbono  $\alpha$  del ligando como punto de estiramiento 2. El proceso luego de ello prosigue de la misma manera ejecutando una dinámica a una velocidad constante de 10 m/s hasta el punto en donde se pueda ver que el ligando se desprende de la proteína. El posterior análisis se hace realizando gráficos de la misma manera a como se definió anteriormente para evaluar el contraste de las fuerzas registradas justo al momento del desprendimiento de la cadena y las áreas bajo la curva relacionadas al proceso entre los péptidos generados y las hebras control para ver el comportamiento de ambas antes un desafío mecánico.

## 4 PRESENTACIÓN Y ANÁLISIS DE RESULTADOS

### 4.1 Caracterización mecánica e identificación del dominio más influyente

Con tal de saber cuál velocidad era la más conveniente para realizar los análisis en base al tiempo y resolución de análisis se hizo un análisis con 3 velocidades diferentes, lo que arrojó los siguientes resultados (figura 12):

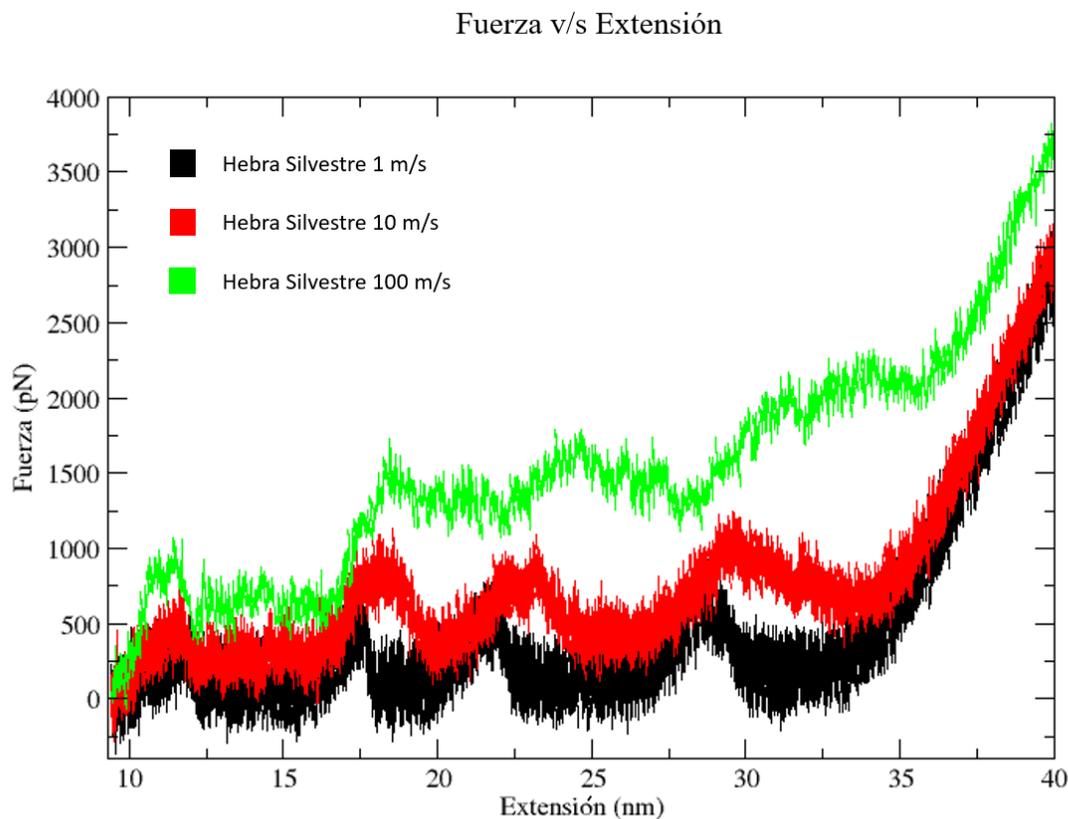


Figura 12: Gráfico de Fuerza v/s Extensión que muestra la diferencia que hubo entre las dinámicas moleculares obtenidas a velocidades constantes de 1 m/s (trayectoria color negro); 10 m/s (trayectoria color rojo) y 100 m/s (trayectoria color verde). El barrido de velocidades permitió evaluar cuál de las velocidades evaluadas daba la mayor calidad de datos en base al tiempo que tardó cada SMD.

En la figura 12 se muestran tres trayectorias pertenecientes a un sistema de cadena silvestre de Ag I/II evaluando las tres velocidades mencionadas en la descripción de la figura 12. El gráfico muestra en el eje X la extensión total que alcanzó la proteína durante el proceso y en el eje Y, las fuerzas asociadas al desplegamiento de las distintas partes de la adhesina. Cada máximo mostrado en las trayectorias representa el momento en el que una sección de Ag I/II es desplegada durante el desafío mecánico al

que se le expone, lo que finalmente llega a un proceso de sobre estiramiento en el que la proteína ya no se puede seguir siendo extendida, proceso que se explicará en detalle más adelante. La trayectoria color negro que representa la SMD hecha a 1 m/s muestra 4 máximos bien definidos en 11,25 nm; 17,5 nm; 21,5 nm y 28,75 nm respectivamente con un sobre estiramiento que comienza al momento de alcanzar 35 nm de extensión, registrando un máximo de fuerzas de aproximadamente 1500 pN. La dinámica hecha a 10 m/s muestra fundamentalmente los mismos máximos observados en el caso anterior referente al nivel de extensión en la que se resolvió cada una, con la diferencia de que las fuerzas registradas son de una mayor magnitud registrando un máximo de fuerza en 3500 pN y un sobre estiramiento registrado también en los 35 nm de extensión, con la diferencia de que la extensión total luego de este es menor que lo visto en el caso anterior. En el caso de la dinámica resuelta a 100 m/s, sólo se pueden observar dos máximos bien resueltos uno a los 11,25 nm y otro a 18,75 nm, similar a los dos primeros máximos vistos a 10 m/s, con una fuerza máxima registrada en 2250 pN aproximadamente. Estos resultados se usaron para descartar esta magnitud para futuros análisis ya que la resolución no es buena, y las fuerzas registradas no son de confianza.

En base al tiempo de cómputo que tardó cada dinámica en resolverse por completo, a 1 m/s fueron nueve días, a 10 m/s fue de 1 día y por último a 100 m/s el tiempo de análisis fue de dos horas. En base a estos resultados, se puede observar que la dinámica hecha a 10 m/s es considerablemente más rápida que la realizada a 1 m/s. Tomando en cuenta esto, en conjunto con lo que se mencionó anteriormente, se decide realizar el resto de las dinámicas moleculares a velocidad constante de 10 m/s ya que muestra una buena resolución de los datos siendo altamente similares a su contraparte más lenta y en un menor tiempo de análisis.

Al definir con que velocidad se haría el resto de los análisis, si realizaron los análisis entre el gráfico y la trayectoria con tal de reconocer en detalle los eventos de desplegamiento se muestran en los gráficos.

### Fuerza v/s Extensión

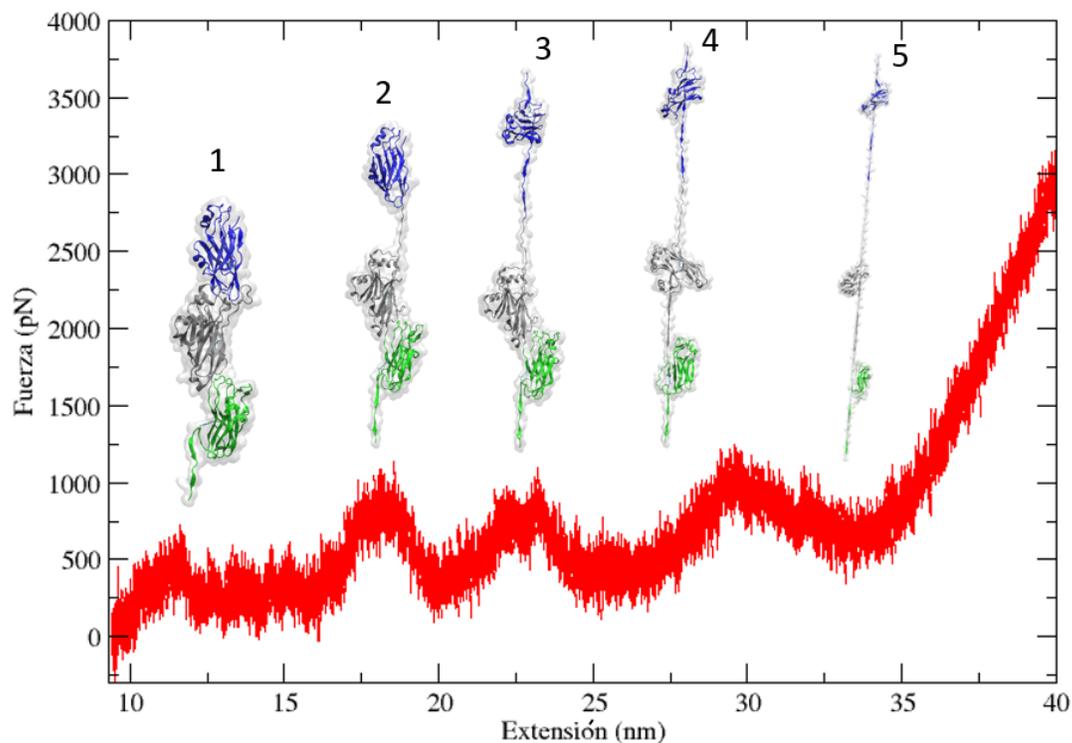


Figura 13: Gráfico de SMD a 10 m/s de Ag I/II silvestre mostrando los eventos de desplegamiento (1-4) que reflejan los máximos a lo largo del proceso de estiramiento. Los máximos vistos en la gráfica reflejan el momento en el que una estructura de la proteína es desplegada producto de la perturbación mecánica.

Como se mencionó en párrafos anteriores, los máximos reflejados en los gráficos de fuerza v/s extensión representan el momento en que una sección de la proteína es desplegada producto de la perturbación mecánica aplicada. En el caso de la proteína en su forma silvestre, estos eventos fueron enumerados del 1 al 5, cada uno representando lo siguiente: (1) C – terminal del dominio C3 marcado en 750 pN; (2) N – terminal del dominio C2 en 1125 pN aproximadamente; (3) N y C – terminal del dominio C1 simultáneamente aproximadamente en 1090 pN y (4) Parte del C – terminal de C2 a 1250 pN.

Luego del análisis de la cadena se procedió a realizar las SMD correspondientes a los modelos modificados para que no formaran el enlace isopeptídico partiendo por C1 (figura 14).

### Fuerza v/s Extensión

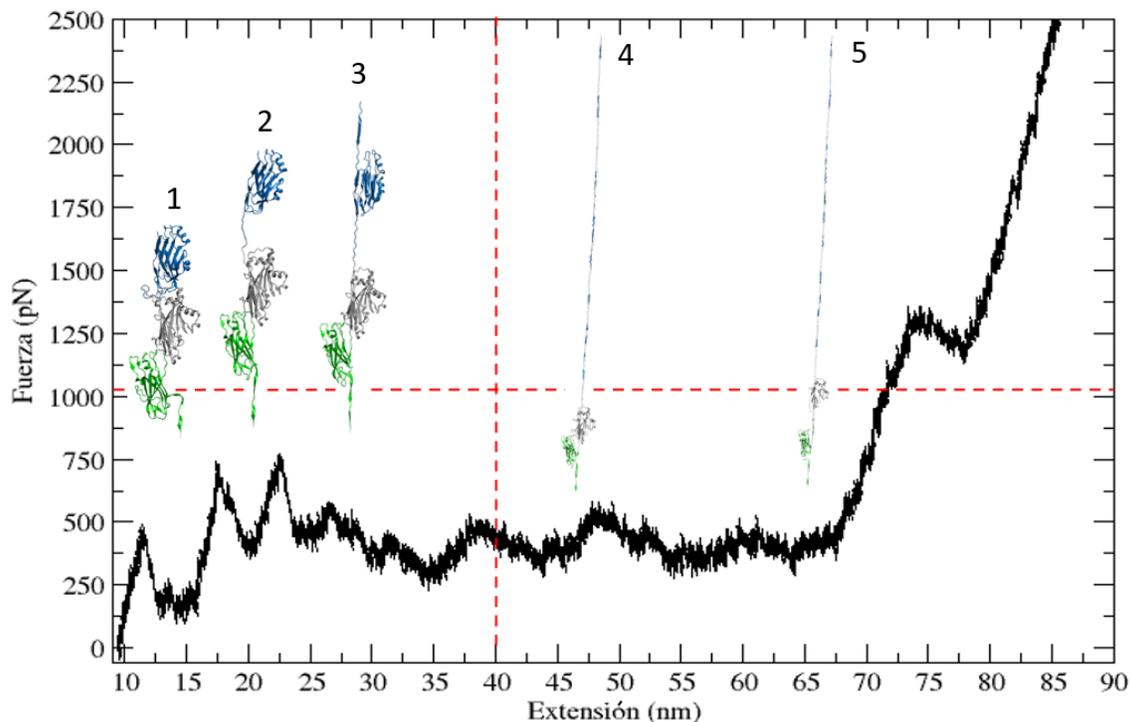


Figura 14: Dinámica molecular realizada a 10 m/s sobre un modelo modificado del C – terminal de Ag I/II para no formar el enlace isopeptídico en el primer dominio de la proteína. Se muestran con líneas entrecortadas color rojo los valores máximos de fuerza y de extensión mostrados por la proteína silvestre para denotar los efectos de la remoción de un enlace isopeptídico en la proteína.

Los eventos de desplegamiento se resolvieron de la siguiente manera: (1) C – terminal del dominio C3 en 500 pN; (2) C – terminal de C1 a los 750 pN; (3) N – terminal de C1 ligeramente por encima de los 750 pN; (4) evento de desplegamiento progresivo de todo el dominio C1 que marca una fuerza máxima de aproximadamente 625 pN y (5) N – terminal de C2 a los 1375 pN. En comparación con la cadena silvestre se puede ver que las fuerzas asociadas al desplegamiento disminuyeron y que la extensión final que adquiere Ag I/II al final de proceso también es mayor, demostrando que el enlace isopeptídico efectivamente no se formó. De la misma manera se puede ver que en el momento que empieza a desplegarse la primera sección de C1 (figura 14, evento 2) todo lo que le prosigue corresponde eventos que culminan en el desplegamiento completo del dominio C1. Aun así, fuera de que las fuerzas en general disminuyeron, se puede observar que el último evento de desplegamiento registrado, perteneciente a estructuras del C2 registró fuerzan aún mayores que lo visto en la trayectoria del control (figura 13).

Los eventos de desplegamiento correspondientes a la intervención del enlace isopeptídico del dominio C2 fueron (figura 15): (1) C – terminal de C3 a los 500 pN; (2) C – terminal de C1 aproximadamente en 875 pN; (3) N- terminal de C1 a 875 pN; (4) N – terminal de C2 en conjunto de un desplegamiento más pronunciado del C – terminal de C1 marcado en aproximadamente 900 pN; (5) Comienzo del desplegamiento progresivo de C2 registrando el mismo valor de fuerza que el evento anterior y (6) Desplegamiento completo de C2 en conjunto con desplegamientos adicionales de estructuras del dominio C1 en la marca de 875 pN. La ausencia del enlace isopeptídico del dominio C2 presenta características que lo hacen sobresalir en comparación con los demás modelos evaluados. Primeramente, se puede observar que además del hecho que las fuerzas registradas son más bajas que las registradas en el control, este presenta una mayor cantidad de eventos de desplegamiento, los cuales no sólo pertenecen a estructuras de C2 sino que también a otras estructuras de C1, dando a entender una posible relación estructural entre ambos dominios, demostrado por el hecho que en conjunto al desplegamiento progresivo de C2, en C1 se obtuvieron desplegamientos más prolongados además de un visible cambio conformacional en la estructura terciaria del dominio.

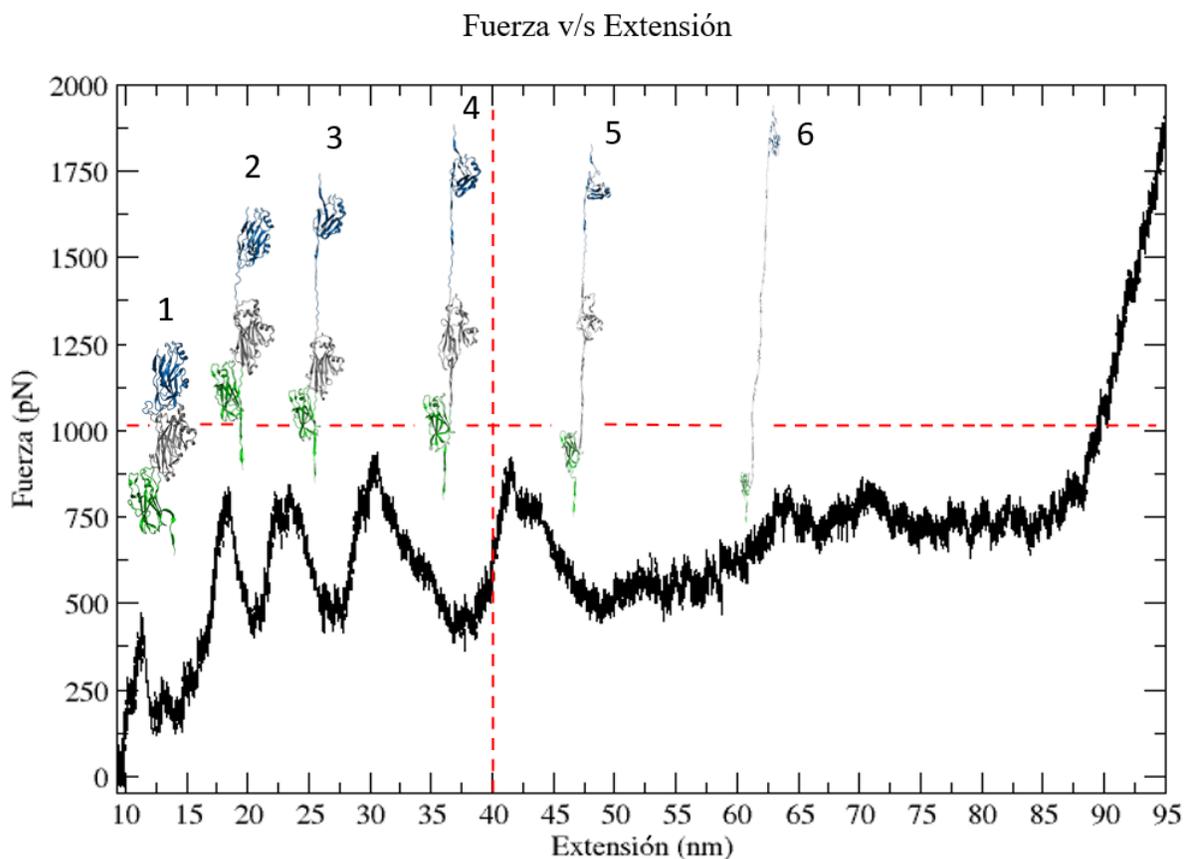


Figura 15: Dinámica molecular realizada a 10 m/s sobre un modelo modificado del C – terminal de Ag I/II para que no forme el enlace isopeptídico en su segundo dominio. Se muestran con líneas entrecortadas color rojo los valores máximos de fuerza y de extensión mostrados por la proteína silvestre para denotar los efectos de la remoción de un enlace isopeptídico en la proteína.

Finalmente, los eventos de desplegamiento correspondientes al modelo C3 modificado fueron los siguientes (figura 16): (1) C – terminal de C3 a los 500 pN; (2) Desplegamiento prolongado del C – terminal y una pequeña sección del N-terminal de C3 que se observa justo después que del primer máximo en aproximadamente 375 pN como un pequeño evento de desplegamiento; (3) C – terminal de C1 en aproximadamente 800 pN; (4) Desplegamiento progresivo de lo que resta del dominio C3 registrando una fuerza máxima de no más de 600 pN y (5) que corresponde a los dos últimos máximos representados en el gráfico siendo estos el desplegamiento del C – terminal de C2 a 1250 pN y el N – terminal del mismo dominio en los 1500 pN. La ruta de desplegamiento que sigue Ag I/II cuando se remueve el tercer enlace isopeptídico es relativamente similar a la mostrada en la figura 14, observándose disminución en las fuerzas totales y logrando una extensión final similar a la mostrada en la figura 15. Nuevamente se puede observar que las estructuras C2 resolvieron eventos de desplegamientos en altos niveles de fuerza los cuales difieren de los mostrados en la remoción del primer enlace isopeptídico dado que tanto el C – terminal y el N – terminal de C2 se resolvieron en los últimos momentos de la SMD.

### Fuerza v/s Extensión

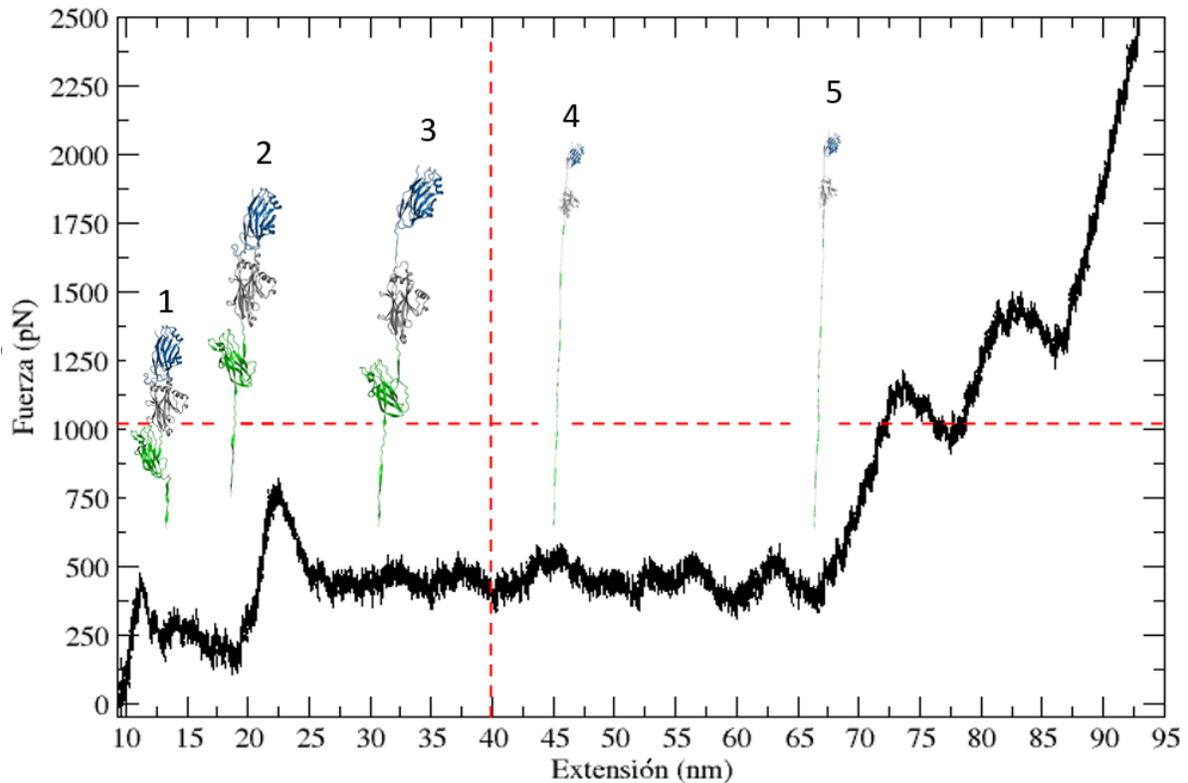


Figura 16: Dinámica molecular realizada a 10 m/s sobre un modelo modificado del C – terminal de Ag I/II para que no forme el enlace isopeptídico en su tercer dominio. Se muestran con líneas entrecortadas color rojo los valores máximos de fuerza y de extensión mostrados por la proteína silvestre para denotar los efectos de la remoción de un enlace isopeptídico en la proteína

Con los resultados obtenidos, se puede observar la importancia que tienen los enlaces isopeptídicos en Ag I/II, dado que en la ausencia de ellos la proteína se hace extensible a menores magnitudes de fuerza. Adicionalmente, se puede ver que cuando los dominios de son expuestos a estrés mecánico estos tienden a rotar y/o cambiar su estructura con tal de establecer una conformación más estable ante la perturbación, lo que fue observado en el estado nativo de la proteína como en los dominios intervenidos en los dominios C1 y C2.

En vista de lo anterior, los tres enlaces podrían servir el propósito de funcionar como dianas para la intervención molecular, aun así, se decide seguir con el protocolo usando como objetivo al segundo enlace. La remoción del enlace de C2 mostró varias características que lo hicieron especialmente

atractivo como candidato para intervención. Primero, el hecho de que el modelo asociado haya mostrada la mayor longitud al final de proceso; Segundo, se resolvieron una mayor cantidad de eventos de desplegamiento en comparación con los otros modelos y tercero, los resultados obtenidos por la intervención de C1 y C3 mostraron que secciones adicionales de C2 fueron desplegadas a un rango mayor de fuerza producto que a falta de un enlace, las diferentes estructuras de Ag I/II comienzan a recibir magnitudes más altas de fuerza lo cual no puede ser visto en un estado nativo. Es por ello que para seguir con el protocolo se decide usar como objetivo al enlace presente en segundo dominio del C – terminal de Ag I/II.

#### 4.2 Generación del bolsillo para acoplamiento molecular

Para el desarrollo del acoplamiento molecular, el primer paso fue formar el espacio al cual se le conoce como bolsillo, al cual el péptido se debería unir. Previo a este se hizo una pequeña edición al cristal del C – terminal en la cual se removió por completo el tercer dominio de la proteína usando las facilidades que ofrece Pymol, dado que la presencia de este sólo sería un interferente al momento de formar el espacio para el sitio de unión.

El bolsillo requerido para hacer el acoplamiento molecular se hizo en base a un protocolo de SMD similar al seguido para la caracterización mecánica el cuál arrojó el siguiente resultado (figura 17):

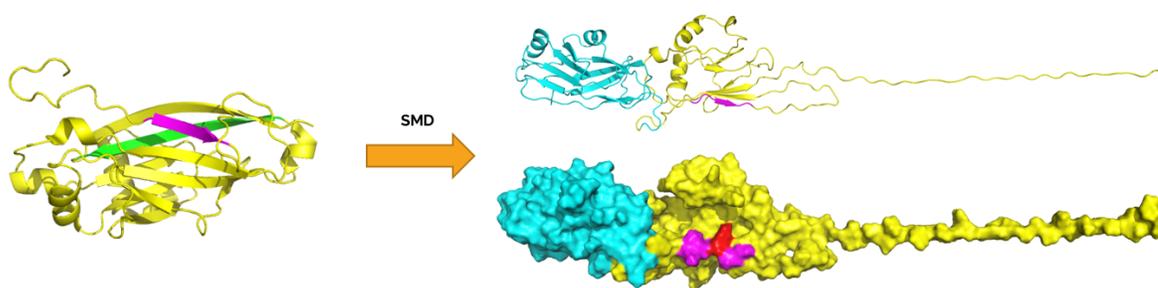


Figura 17: Resultado obtenido posterior a la dinámica molecular para generar el bolsillo necesario para realizar el acoplamiento molecular. (A) En color morado se destaca la cadena con la lisina 1161 mientras que en verde la cadena con la asparagina 1311. (B) Representación en cartoon (superior) y surface (inferior) de la separación de la cadena que protegía la hebra  $\beta$ , liberando el área necesaria como para que el péptido pueda unirse al bolsillo, denotándose en color rojo la lisina 1161.

A diferencia de las SMD realizadas para la caracterización mecánicas en este caso los puntos se estiramiento se definieron de manera que el punto 1 fuera el conjunto de carbonos  $\alpha$  que comprendían

el inicio de la proteína hasta el hasta el carbono  $\alpha$  de donde inicia la cadena de la Lys1121 y como punto 2 el último carbono  $\alpha$  de la secuencia.

Dada la naturaleza de estos puntos de estiramiento que se definieron durante la preparación de la dinámica molecular fue posible separar la sección que protegía la sección necesaria para realizar la intervención molecular sin necesidad de tener que desplegar la proteína en su totalidad. En el momento en el que se logró este estado el modelo fue exportado como un archivo .pdb y se guardó como archivo .pdb para el desarrollo del acoplamiento molecular.

### **4.3 Diseño del péptido**

#### **4.3.1 Iteraciones de péptidos**

Una vez decidido que el uso del segundo dominio como objetivo para intervención molecular se hizo un análisis de la hebra  $\beta$  complementaria que contienen la Asn que forma el enlace nativamente en el dominio C2 para definir la longitud que tendrá el péptido, y en conjunto también se hizo un análisis comparativo entre los enlaces de los demás dominios con tal de ver si existían patrones comunes en cuanto a la composición de estas cadenas y de esta manera poder reducir el número de péptidos totales que se generará mediante el código de Python (figura 19).

Analizando la hebra  $\beta$  complementaria se verifica que está compuesta por un conjunto 12 aminoácidos con la asparagina que forma el enlace en la quinta posición. En más detalle, la sección de la hebra que interacciona directamente con la cadena que contiene la lisina que forma el enlace corresponde sólo a los 10 primeros aminoácidos que la componen. En base a esto se define una plantilla de un péptido de 10 aminoácidos dejando como término fijo una asparagina en la quinta posición con tal de imitar la composición neta de la hebra original como bien puede ser visto en la figura 18.

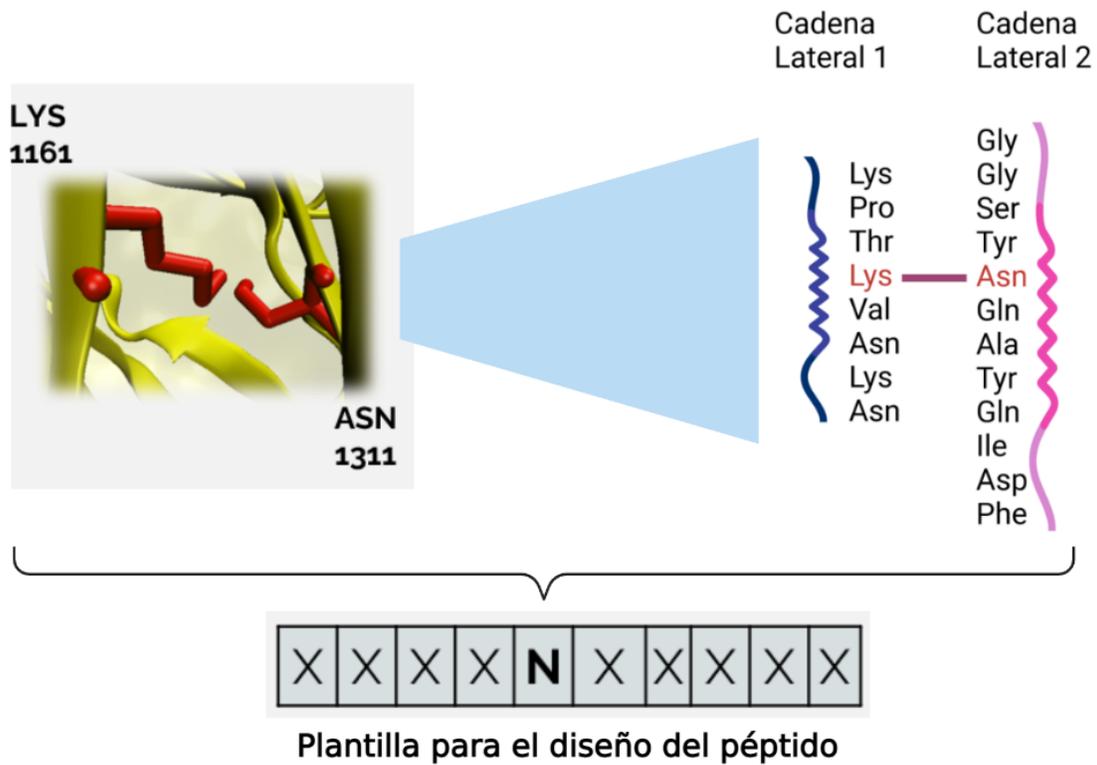


Figura 18: Esquematzación de los áminácidos que conforman la hebra  $\beta$  complementaria que forma el enlace isopeptídico dentro del segundo dominio del C – terminal de Ag I/II que se usó para determinar el número de aminácidos que forman el péptido a diseñar. Se deja como término fijo la asparagina que forma el enlace en quinto lugar, mientras que las “x” se consideran como los términos que serán iterados por distintos aminoácidos usando la línea de código de Python.

```

from itertools import product
import random

secuencia = {'1': ['R', 'H', 'K', 'D', 'E', 'S', 'T', 'N', 'Q', 'C', 'G', 'P', 'A', 'V', 'I', 'L', 'M', 'F', 'Y', 'W'],
            '2': ['R', 'H', 'K', 'D', 'E', 'S', 'T', 'N', 'Q', 'C', 'G', 'P', 'A', 'V', 'I', 'L', 'M', 'F', 'Y', 'W'],
            '3': ['R', 'H', 'K', 'D', 'E', 'S', 'T', 'N', 'Q', 'C', 'G', 'P', 'A', 'V', 'I', 'L', 'M', 'F', 'Y', 'W'],
            '4': ['R', 'H', 'K', 'D', 'E', 'S', 'T', 'N', 'Q', 'C', 'G', 'P', 'A', 'V', 'I', 'L', 'M', 'F', 'Y', 'W'],
            '5': ['N'],
            '6': ['R', 'H', 'K', 'D', 'E', 'S', 'T', 'N', 'Q', 'C', 'G', 'P', 'A', 'V', 'I', 'L', 'M', 'F', 'Y', 'W'],
            '7': ['F'],
            '8': ['V'],
            '9': ['L'],
            '10': ['V']}

```

Figura 19: Línea de código utilizada para generar la primera lista de péptidos candidatos para intervención molecular. Mediante el uso de éste código se fue capaz de generar una lista de péptidos en base al conjunto de aminácidos definido para el acoplamiento molecular.

Con la primera ejecución del código se generó una lista de más de cuatro millones de iteraciones posibles, de la cuál se hizo una selección al azar de 500 péptidos los cuales se les hizo una predicción de su estructura secundaria usando UCSF Chimera para ser probados por análisis de acoplamiento molecular en conjunto con las tres hebras control: (i) Un péptido mimético a la hebra complementaria de C2 de secuencia GSYENQAYQI el cuál fue luego construido en Chimera; (ii) Otro péptido mimético, en forma de la misma hebra complementaria que forma el enlace en el cristal original, la cuál fue removida de este usando el software de Pymol, hebra que se denomina como hebra cristalográfica; (iii) un péptido de secuencia GGGGNGGGGG al cuál se denominó como poliG. Los resultados obtenidos por las hebras control fueron los siguientes:

Tipo de Hebra	Afinidad kcal/mol
Control Chimera	-8,2
Control Cristalográfico	-7,5
Poli Glicina	-6,9

Tabla 1: Afinidades mostradas luego del acoplamiento realizado con los controles

Luego de tener las afinidades de las hebras control se procedió a filtrar el resultado de las afinidades obtenidas de las 500 evaluaciones hechas usando como parametro de separación la afinidad obtenida por la hebra control recreada en Chimera, obteniendo un conjunto de 20 peptidos con afinidades por sobre -8 que cumplían con la cualidad específica de entrar al bolsillo y que la asparagina estuviera bien orientada hacia la lisina de la hebra a intervenir (tabla 2):

N°	Secuencia	Afinidad (kcal/mol)
<b>Silvestre</b>	<b>GSYENQAYQI</b>	<b>-8,2</b>
<b>1</b>	<b>TQWNNFVLV</b>	<b>-8,4</b>
<b>2</b>	<b>FGARNQFVLV</b>	<b>-8,5</b>
<b>3</b>	<b>GFVENTFVLV</b>	<b>-8,5</b>
4	EYADNNFVLV	-8,2
5	EYFENTFVLV	-8,5
<b>6</b>	<b>DWIENSFVLV</b>	<b>-8,2</b>
7	QQAENNFVLV	-8,2
8	TFARNTFVLV	-8,2
9	QNLHNQFVLV	-8,1
10	SYADNCFVLV	-8,1
11	TEADNSFVLV	-8,1
12	QYWHNTFVLV	-8,1
13	YFFDNCFVLV	-8,1
14	ESVENTFVLV	-8,1
15	YTADNQFVLV	-8
<b>16</b>	<b>EYAENTFVLV</b>	<b>-8,5</b>
17	FFYHNCFVLV	-8,3
<b>18</b>	<b>QFWDNNFVLV</b>	<b>-8,4</b>
<b>19</b>	<b>YYWENEFVLV</b>	<b>-8,5</b>
<b>20</b>	<b>FGAENNFVLV</b>	<b>-8,4</b>

Tabla 2: Tabla con las afinidades recuperadas del análisis del barrido de 500 péptidos. Los péptidos marcados en caracteres azules son aquellos que destacaron por presentar afinidades mayores al control siendo candidatos al procedimiento de refinamiento por escaneo de alaninas.

Dentro de la lista obtenida, se puede observar que 8 de ellos presentan afinidades mayores al control siendo estas mayores a -8,2. En base a estos resultados se decide tomar esos péptidos para aplicar el procedimiento de escaneo de alaninas. Sumado a esto, en base a los resultados vistos en la figura 20 se puede ver que cada uno los 8 péptidos descritos entraron en la misma posición al bolsillo, haciendo de estos resultados viables para seguir con los análisis.

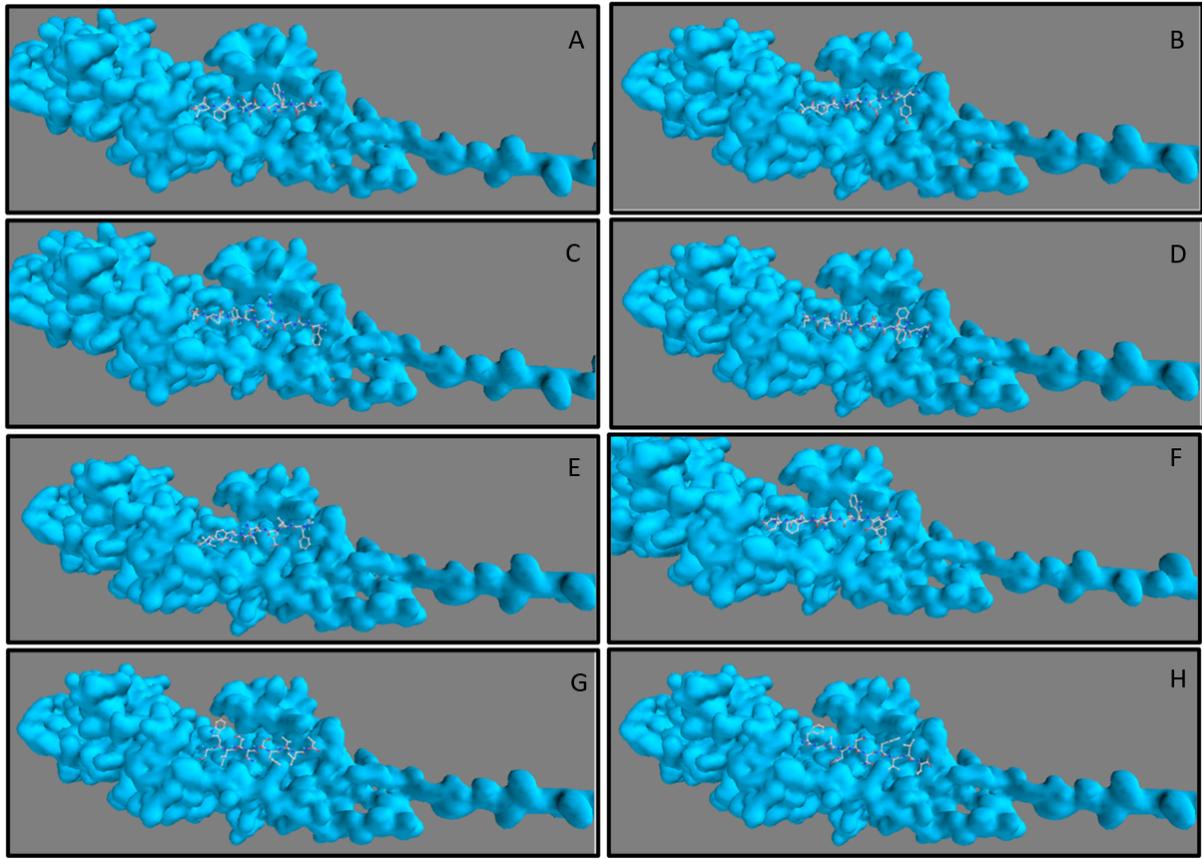


Figura 20: Imagen obtenida de la ventana de visualización de PyRx referente a los resultados obtenidos del acoplamiento molecular respecto a las posiciones en las que los péptidos entraron al bolsillo. Según la numeración de la tabla 2; (A) péptido N°1; (B) péptido N°2; (C) péptido N°3; (D) péptido N°6; (E) péptido N°16; (F) péptido N°18; (G) péptido N°19; (H) péptido N°20.

#### 4.3.2 Escaneo de alaninas

El escaneo de alaninas se usó como una base para poder refinar aquellos péptidos que presentaron un buen valor de afinidad identificando aquellos aminoácidos con un rol en la afinidad mostrada durante en el acoplamiento. Los resultados de estos análisis están representados a continuación en la figura 21.

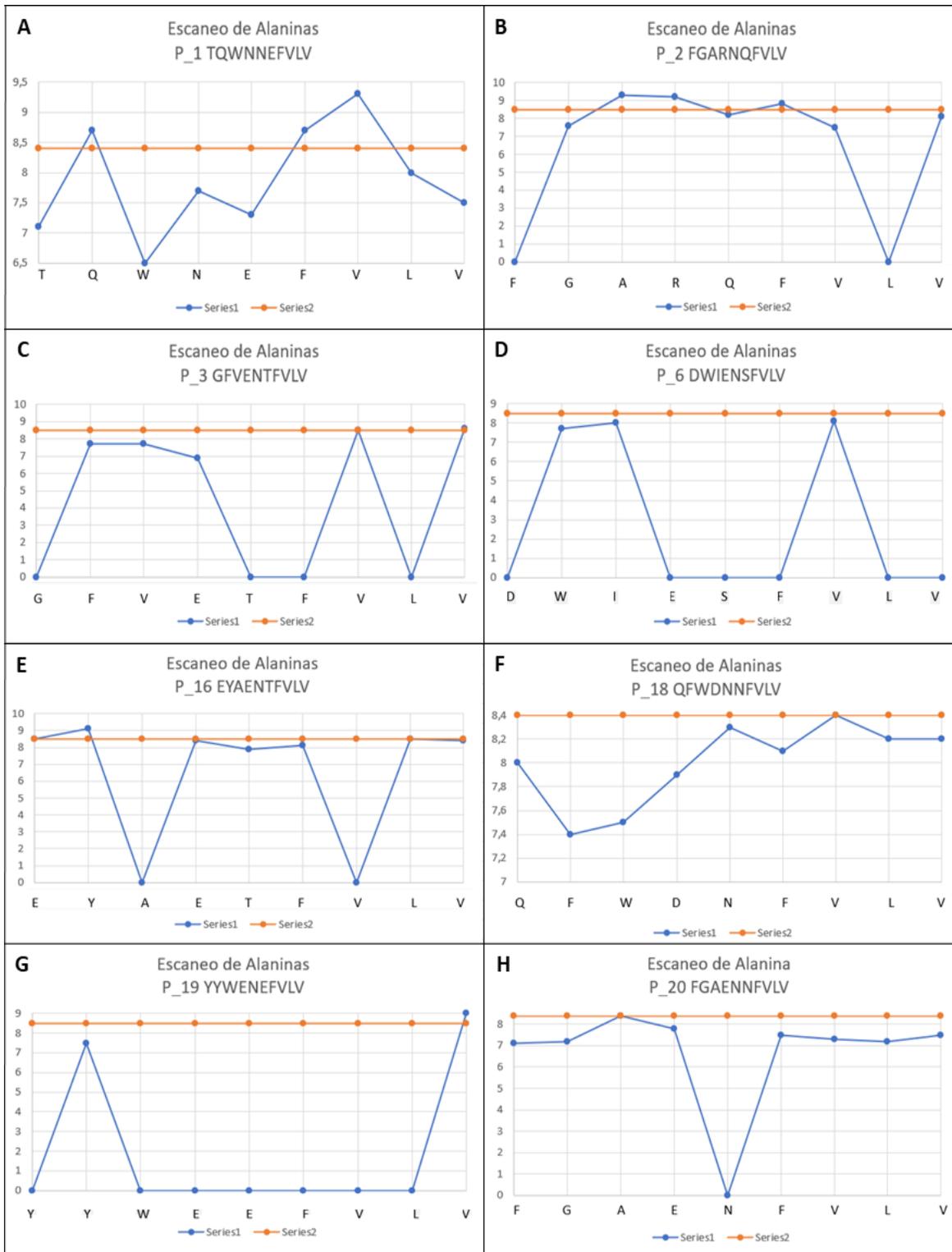


Figura 21: Gráficos obtenidos del escaneo de alaninas aplicado a los 8 péptidos encontrados en el paso anterior. Cada gráfico muestra en una línea de color naranja una afinidad basal que se define como la afinidad mostrada

por el péptido inicialmente, las cuales pueden ser vistas en la tabla 2, mientras que la recta azul muestra la variación de la afinidad cuándo cada uno de los aminoácidos que constituyen el péptido se reemplazan por una alanina mostrando así cuál de los aminoácidos no juegan un rol determinante en la afinidad, dependiendo si la recta azul está por encima o por debajo de la afinidad basal. ; (A) péptido N°1; (B) péptido N°2; (C) péptido N°3; (D) péptido N°6; (E) péptido N°16; (F) péptido N°18; (G) péptido N°19; (H) péptido N°20.

En cada uno de los péptidos intervenidos se puede identificar al menos un aminoácido cuya ausencia en la secuencia no afecta en la afinidad. En P1 corresponde a la Val en la posición 8; P2 a la Ala en la tercera posición y a la Arg en la cuarta posición; P3 a la Val en la décima posición; P6 a la Val en la octava posición; en P16 a la Tyr en la tercera posición; P18 a la Val de la octava posición; P19 a la Val de la décima posición y por último P20 a la Ala de la tercera posición. Mediante este análisis se fue capaz de identificar aquellos aminoácidos que no jugaban un papel importante en la afinidad del péptido y por tanto aquellas posiciones modificables para refinar el valor de afinidad obtenido en primera instancia.

Con estos resultados se procedió a generar un conjunto de archivos .pdb en base al sitio en dónde se identificaron aumentos de afinidad acorde a lo visto en los gráficos de la figura anterior. Las variantes se generaron reemplazando el aminoácido identificado por el resto de los aminoácidos existentes generando 18 ó 19 variantes dependiendo del caso, por ejemplo en el caso de P\_N°1 se generaron 18 variantes dado que el aminoácido de Ala fue probado durante el proceso de escaneo y Val era el aminoácido que tenía en manera nativa.

P_N°1	TQWNNEF <b>V</b> LV
Variante 1	TQWNNEF <b>L</b> LV
Variante 2	TQWNNEF <b>I</b> LV
Variante 3	TQWNNEF <b>M</b> LV
Variante 4	TQWNNEF <b>F</b> LV
Variante 5	TQWNNEF <b>W</b> LV
Variante 6	TQWNNEF <b>Y</b> LV
Variante 7	TQWNNEF <b>N</b> LV
Variante 8	TQWNNEF <b>C</b> LV
Variante 9	TQWNNEF <b>Q</b> LV
Variante 10	TQWNNEF <b>S</b> LV
Variante 11	TQWNNEF <b>T</b> LV
Variante 12	TQWNNEF <b>D</b> LV
Variante 13	TQWNNEF <b>E</b> LV
Variante 14	TQWNNEF <b>R</b> LV
Variante 15	TQWNNEF <b>H</b> LV
Variante 16	TQWNNEF <b>K</b> LV
Variante 17	TQWNNEF <b>G</b> LV
Variante 18	TQWNNEF <b>P</b> LV

Tabla 4: Tabla ejemplo de como se forman las variantes posterior al escaneo de alaninas. Variantes generadas en base al peptido N°1 generando 18 iteraciones dada la sustitución de la octava Valina (destacada en azul) por los 18 aminoácidos existentes restantes.

El proceso se realizó con los demás péptidos para luego evaluar sus afinidades en PyRx repitiendo los mismos parámetros de análisis descritos en el apartado de metodología. De los datos obtenidos a partir de las 164 variantes generadas se destacan los siguientes resultados

<b>P_N°1</b>	<b>TQWNNEFVLV</b>	Afinidad kcal/mol
Variante 5	TQWNNEFFLV	-10
Variante 7	TQWNNEFILV	-9,8
Variante 11	TQWNNEFMLV	-9,5
Variante 17	TQWNNEFHLV	-10,7

<b>P_N°2</b>	<b>FGARNQFVLV</b>	Afinidad kcal/mol
Variante 4	FGVRNQFVLV	-9,5
Variante 8	FGNRNQFVLV	-9,6
Variante 11	FGSRNQFVLV	-9,7
Variante 13	FGDRNQFVLV	-9,6
Variante 16	FGHRNQFVLV	-9,6
Variante 19	FGAWNQFVLV	-9,8
Variante 23	FGAYNQFVLV	-10

<b>P_N°18</b>	<b>QFWDNNFVLV</b>	Afinidad kcal/mol
Variante 10	QFWDNNFSLV	-9,8
Variante 11	QFWDNNFTLV	-10
Variante 13	QFWDNNFELV	-10,5
Variante 17	QFWDNNFGLV	-9,9

<b>P_N°19</b>	<b>YYWENEFVLV</b>	Afinidad kcal/mol
Variante 1	YYWENEFVLL	-9,9

Tabla 5: Resultados destacados de la evaluación de afinidad de las diferentes variantes generadas en base a los resultados de el escaneo de alaninas denotando en color azul la posición en dónde se hizo la iteración. Los péptidos N° 1, 2, 18, 19 fueron aquellos en los se observaron los mejores resultados en cuanto a aumento de afinidades superando un umbral de -9,5 kcal/mol.

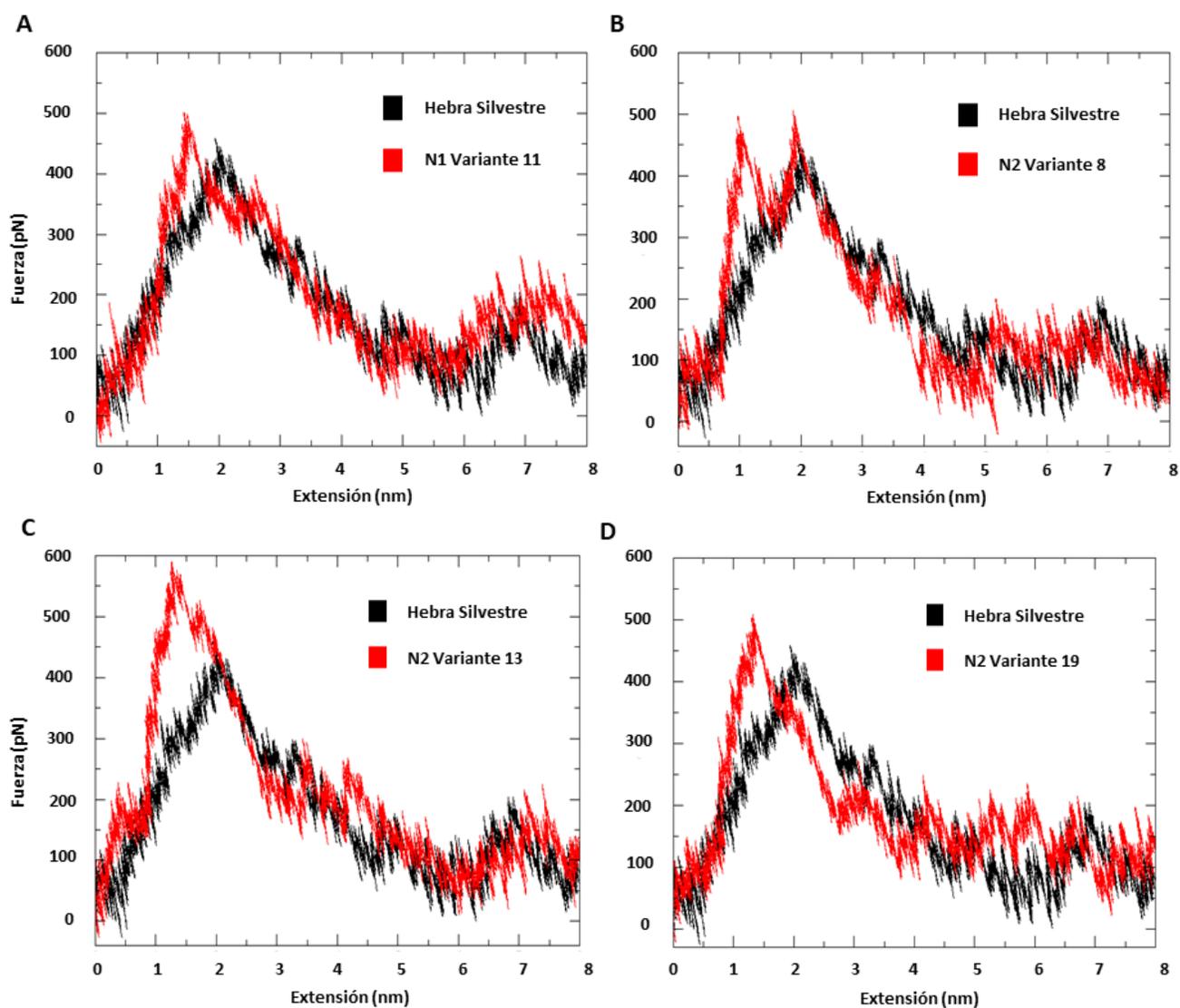
Con los resultados obtenidos se puede comprobar el escaneo de alaninas dio en gran parte buenos resultados logrando obtener 17 nuevos péptidos con mayor afinidad en comparación a sus versiones antes de pasar por el proceso alcanzado mejoras aproximadas del 25% en los mejores resultados mostrados por la variante 17 del péptido N°1 y la variante 13 del péptido número 2 e incluso en los resultados más bajos se alcanzaron mejoras que rodearon el 10%.

#### 4.3.3 Evaluación mecánica de los péptidos

Con tal de evaluar la estabilidad de los péptidos encontrados se aplicó un protocolo basado en dinámicas moleculares con el cual se fuerza al péptido a separarse del bolsillo a modo de poder medir

las fuerzas asociadas al proceso comparando los resultados obtenidos con un control, para lo cual se usó la cadena silvestre generada en Chimera. Como bien se explicó en el apartado de metodología, la dinámica se realizó a una velocidad de 10 m/s y se detuvo el proceso al momento que se identifica que el péptido evaluado se desprende del bolsillo.

De la evaluación de las 17 variantes de péptidos encontradas por el proceso de escaneo de alaninas se destacan los 6 resultados que se pueden observar en la figura 22. Las restantes se descartaron ya que las trayectorias mostraban resultados menores o iguales al control. Los gráficos obtenidos mediante las SMD ligando – proteína, se resuelven de manera que los máximos representados en la curva de la trayectoria representan el momento en el que el ligando se desprende del bolsillo al cuál se encuentra adherido producto del acoplamiento molecular.



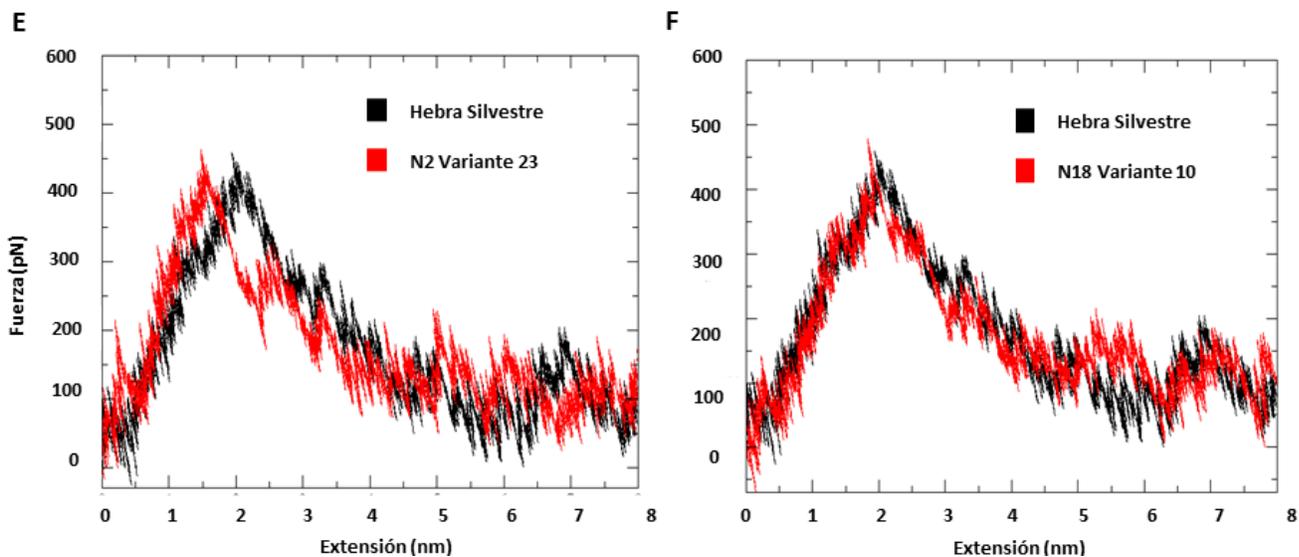


Figura 22: Resultados destacados obtenidos de las dinámicas moleculares ligando proteína. En todos los casos la curva de color rojo representa la trayectoria del ligando que se prueba en el momento mientras que la de color negro a la trayectoria de la cadena control silvestre generada en Chimera, siendo el máximo de la trayectoria el momento en el que péptido se desprende del bolsillo. (A) N1 Variante 1; (B) N2 Variante 8; (C) N2 Variante 13; (D) N2 Variante 19; (E) N2 Variante 23; (F) N18 Variante 18.

Dentro de los resultados obtenidos de las SMD ligando – proteína se da especial relevancia a la variante 13 del péptido N°2 (figura 22C), ya que a pesar de que todos los ligandos representados superan el umbral de los 450 pN aproximadamente fijado por la hebra silvestre, fue justamente la variante 13 del péptido N°2 aquel que marcó el nivel de fuerza más alta antes de desprenderse casi llegando a los 600 pN. En base esto, se dispone de un ligando que no sólo es más afín por la proteína Ag I/II sino que también, mecánicamente más estable, haciendo de este un buen candidato para intervención molecular al momento de que los datos de este trabajo sean probados in vitro.

## 5 DISCUSIÓN

La presencia de enlaces isopeptídicos juegan un papel de alta importancia en la virulencia y potencial patológico de bacterias Gram positivo, haciéndolas resilientes a actividad proteolítica además de otorgar resistencias a desafíos térmicos y mecánicos (83). En el caso de *Streptococcus mutans*, bacteria definida como la causa etiológica principal en el desarrollo de las caries, este fenómeno representa una de las respuestas hacia el por qué las caries son tan difíciles de combatir una vez formado el complejo.

En este trabajo se desarrolló la idea de generar un péptido con la capacidad de intervenir la formación de uno de los enlaces isopeptídicos de la adhesina Ag I/II de *S. mutans*, a modo de solución preventiva al desarrollo de las caries, produciendo una molécula más afín que las estructuras silvestres del *pilus* de Ag I/II y por tanto más estable a modo de lograr se establezca una unión de carácter irreversible.

Para saber que enlace intervenir dentro del dominio C – terminal, se hizo una caracterización mecánica para identificar el dominio más influyente en cuanto a la estabilidad mecánica/estructural del C – terminal de Ag I/II. En este trabajo se presentan las primeras evidencias del aporte energético que otorgan los enlaces isopeptídicos cuando Ag I/II es expuesta a un desafío mecánico. Mediante esto se observó que los dominios C1 y C2 de Ag I/II presentaron las características más relevantes. En ausencia de C1 vemos que las fuerzas totales del proceso disminuyen presentando los menores valores entre las demás simulaciones, y en C2 además de que las fuerzas asociadas al proceso disminuyeron, también se logró una mayor extensión neta de la proteína la cuál llegó hasta aproximadamente 95 nm. En documentación previa respecto a la estructura cristalográfica del C – terminal de Ag I/II, se explica que los dominios C1 y C2 juegan un papel crucial durante la adhesión de *S. mutans* a la superficie del diente, en gran parte por la presencia de sitios de unión a glucosa que reconocen a como sitios de unión a estructuras altamente glicosiladas como lo es el SAG (53). Tomando esto en cuenta, es posible relacionar que la importancia que tienen los dominios C1 y C2 en el proceso de adhesión no sólo está dado por la presencia de los sitios de unión mencionados sino que también por el aporte mecánico que otorgan los enlaces presentes en esos dominios, jugando un posible rol en la estabilización del proceso unión por parte de la región C – terminal. El punto anterior también puede ser usado para darle validez al uso del péptido como interferente del proceso de adhesión, dada la capacidad que tienen los péptidos antimicrobianos para desestabilizar estructuras proteicas (84), dado el hecho que el péptido diseñado en este trabajo tendrá como función prevenir la formación del enlace isopeptídico lo que causa irregularidades en el plegamiento y adquisición de estructura secundaria en el C – terminal, lo

que a su vez causaría no sólo un desequilibrio en la transducción de fuerzas en la adhesina sino que se le prevería a esta de un sitio de unión al SAG. Adicionalmente, un resultado que se debe destacar fue el como los dominios del C – terminal rotaron y/o cambiaron su conformación estructural al enfrentarse a un desafío mecánico. Este fenómeno fue demostrado en otra bacteria tipo Gram - positivo, *Staphylococcus epidermidis* usando también métodos *in silico* (85). En la investigación se demostró que cuando *S. epidermidis* era expuesta a estrés mecánico ésta rotaba los dominios de su adhesina con tal de mitigar el impacto que provocaba dicha perturbación en su estructura tridimensional adquiriendo una conformación más estable. El fenómeno de rotación se vió en 2 instancias de éste trabajo: En la proteína silvestre y cuando se intervino la formación de los enlaces en los dominios C1 y C2. Considerando el caso de estudio de *Staphylococcus epidermidis* citado anteriormente, su adhesina es capaz de adoptar diferentes conformaciones dependiendo de si la fuerzas aplicadas sobre ella era de baja o alta escala, no es raro haber observado rotación en los dominios de Ag I/II como respuesta adaptativa, además tomando en consideración la capacidad que tienen las proteínas de pasar por cambios conformacionales con tal de ser estructuralmente más estables (86) es posible afirmar que tanto los dominios C1 y C2 juegan un importante papel en la estabilidad estructural de la proteína especialmente cuando se es expuesta a grandes estímulos de energéticos. Mediante esto se puede concluir que la estabilidad mecánica de Ag I/II viene no sólo de la presencia de los enlaces isopeptídicos que se encargan de disipar fuerzas elásticas sino que también por los cambios conformacionales reportados durante las dinámicas, los cuales son aún más estables dada la presencia de los enlaces. En base a esto se puede decir que la estabilidad del proceso de adhesión depende en gran parte por los anteriorers dos dominios, ya que fuera de no haber hecho experimentos de adhesión hacia una matriz, las simulaciones moleculares cuando se hacen con el suficiente cuidado, son capaces de predecir con alto grado de exactitud el comportamiento de proteínas mecanoestables (87).

El diseño del péptido a *per se* dependía de dos partes: Primero, decidir que enlace se desea atacar dada la naturaleza mimética de la molécula, con tal de que la estructura del péptido pueda ser reconocida con la proteína con tal garantizar una buena afinidad (88) y segundo definir en qué posición irá la asparagina que va de la mano con el primero punto. La decisión de elegir la asparagina como aminoácido fijo para formar el enlace isopeptídico y no la lisina no es trivial ya que como se demostró en la investigación de Rivas *y cols.* (69) el momento que se realiza la intervención es crucial para el éxito de esta propuesta cuando se migre de un ambiente *in silico* a uno *in vitro* e incluso *in vivo*. Como bien se mencionó en el apartado de introducción los enlaces isopeptídicos de Ag I/II se forman entre

aminoácidos de lisina y asparagina, y, una característica en común entre ellos es que la lisina siempre está río arriba de la asparagina. Esto quiere decir que cuando la adhesina está en vías de expresión, en otras palabras, cuando la cadena polipeptídica sale desde el ribosoma para adquirir su estructura terciaria en la membrana plasmática previo al mismo anclaje de la adhesina al peptidoglucano (89), la primera parte en salir hacia el espacio extracelular es precisamente la sección con el aminoácido de lisina. De esta manera el péptido debería ser capaz de unirse a esa sección específica antes que ocurra el plegamiento de la proteína y por ende antes que se exprese la sección con la asparagina. Siguiendo con el diseño del péptido se mencionó que no sólo se dejó como término fijo la Asn en la quinta posición sino que también, una secuencia FVLV en los últimos cuatro términos de este. La inclusión de esta secuencia dio buenos resultados en primera instancia ya que las afinidades encontradas en el primer barrido dieron cercanas e incluso mayores al control generado en Chimera jugando un buen rol en reducir los términos totales arrojados por código usado en python. Sin embargo, una vez aplicado el proceso de escaneo de alaninas se vio en algunos casos que cuando se removía un aminoácido de la secuencia FVLV la afinidad subía. Por lo general, los reemplazos que se hicieron en esa área específica fueron la mayoría de los casos por aminoácidos con cadenas laterales de carácter hidrofóbico como la misma valina u otros aminoácidos con anillos aromáticos. Esto se relaciona con una característica estructural de los isopéptidos, siendo estos rodeados por residuos hidrofóbicos y aromáticos (65). De igual manera este reemplazo no ocurrió en todos los casos, habiendo ocasiones donde incluso la falta de esta secuencia generaba una pérdida total de la afinidad hacia el bolsillo.

Respecto al proceso de refinamiento de péptidos, este se hizo adecuando a condiciones *in silico* una metodología de escaneo de alaninas, que generalmente se usa realizando mutaciones sitio dirigidas para evaluar la funcionalidad o estabilidad de una proteína bajo condiciones definidas dependiendo del diseño experimental que se desea seguir (91). Para la correcta realización de los objetivos, se desarrolló una metodología simple y libre de costo para evaluar la afinidad relacionada a interacciones proteína – proteína basada en los mismos principios teóricos del escaneo de alaninas que se realiza *in vitro*. La opción que más se asemeja a lo desarrollado en este documento es una herramienta bioinformática conocida DrugScore<sup>PPI</sup> (91) una plataforma online libre de registro que requiere un archivo de pdb que funciona como la proteína diana y una secuencia peptídica que funciona como el ligando control al que se aplica el proceso de escaneo con tal de poder identificar residuos críticos en complejos proteína – proteína. Sin embargo, fuera de las similitudes que comparten en el resultado final que se entrega, DrugScore<sup>PPI</sup> presenta 2 limitaciones, la primera es que no permite delimitar un espacio de análisis

realizando sólo acoplamiento ciego y la segunda es que limita el número de aminoácidos que conforman el ligando a 7 residuos, lo que hace que los análisis que se pueden realizar en la plataforma sean relativamente restringidos. En cambio, dada la naturaleza análoga de la metodología desarrollada en este trabajo permite hacer escaneos de alaninas con mayor libertad y totalmente ajustable. Además, el proceso propuesto por DrugScore<sup>PPI</sup> se detiene al momento de entregar una tabla con las diferencias de afinidades asociadas a la remoción de cada aminoácido, mientras que el protocolo propuesto en este trabajo incluye una fase de refinamiento, por lo cual no sólo se identifican aquellos aminoácidos que juegan un papel relevante en la afinidad sino que también aquellos que no, siendo estos candidatos para probar que otro aminoácido en esa posición sea capaz de subir la afinidad neta de los ligandos. Un punto negativo que sí posee el escaneo de alaninas realizado es que es medianamente azaroso dada la cantidad de pasos asociados a este, aun así, el protocolo se encuentra en constante mejora buscándose incluso una forma de poder automatizarlo.

En cuanto a los resultados obtenidos por las dinámicas ligando – proteína, y como se bien se mencionó en el apartado de resultados, de los 6 de los 17 complejos evaluados registraron magnitudes de fuerzas mayores en comparación al control realizado con la cadena silvestre, siendo el mejor de ellos perteneciente a la variante 13 del péptido N°2 posterior al refinamiento. Este péptido, fuera de ser más afín que la cadena silvestre, en cuanto a las afinidades netas del proceso este no fue el que marcó el valor de afinidad más alto dentro del conjunto. Recordar que las afinidades mostradas por un proceso de acoplamiento molecular se representan con valores negativos ya que representan la energía liberada durante la unión del ligando al sitio de unión, lo que también se traduce como la energía necesaria para comprometer dicha unión (92). Teniendo en cuenta lo anterior se esperaría que a afinidades más altas las fuerzas de desprendimiento reportadas fueran más altas, lo que no se vio en los experimentos realizados, como lo es el contraste que existe entre la variante 13 del péptido N°2 y la variante 17 del péptido N°1, el cuál como se ve en la tabla 5, fue aquel que marcó el mayor valor de afinidad, pero que aun así las fuerzas registradas durante la SMD no fueron capaces de superar a la cadena silvestre. En relación a esto, existe documentación que explica que la fuerza por la que un ligando se adhiere a una superficie no depende necesariamente de la afinidad que haya registrado durante un acoplamiento molecular (93). Esto quiere decir que la composición que tiene la variante 13 del péptido N°2 fuera de no marcar un valor tan alto de afinidad en comparación con los demás péptidos presentados en la tabla 5, presentando casos dónde incluso péptidos de la misma afinidad mostraron fuerzas de adhesión diferentes, lo que da a entender que la variante 13 del péptido N°2

tiene características en su composición aminoácida que le dan ventajas por sobre el resto de las cadenas. Como se explicó anteriormente, aminoácidos hidrofóbicos y aquellos con anillo aromático juegan un papel de alta importancia al momento de formar los isopeptidos en condiciones naturales, analizando la composición de la cadena se compone por FGDRNQFVLV, siendo el cambio que generó la variante un Asp en la tercera posición. Anteriormente se mencionó que el aminoácido de Asp es importante en el desarrollo del futuro del enlace isopeptídico, coordinando este, y formando puentes de hidrógeno con la cadena complementaria, lo que podría dar una explicación a la mayor afinidad. En cuanto al mayor grado de fuerza de unión, la respuesta podría estar dentro de los aminoácidos de Phe al principio y al final de la cadena. Dentro de las muchas interacciones que ocurren entre complejos proteína – proteína se encuentra una llamada interacción del catión  $\pi$ , la cual representa una fuerte unión de carácter no covalente que suele suceder entre cadenas laterales de aminoácidos aromáticos como los la Phe, Tyr y el Trp, lo que provee de una superficie de potencial electrostático negativo la cual es altamente afín por aminoácidos catiónicos (94 – 95). Tomando en consideración la composición de la cadena complementaria Lys1161 que se busca intervenir, TKVNK, la cual es mayoritariamente positiva por la presencia de los dos aminoácidos de Lys que presenta, es posible confirmar que la gran fuerza que se genera durante la SMD con la variante 13 del péptido N°2 se debe a una acción conjunta entre el aminoácido de Asp generando puentes de hidrógeno y los dos aminoácidos de Phe generando interacciones de cationes  $\pi$ . En cuanto al por qué no se notó quizás un valor de afinidad esperado, la respuesta puede darse por medio del mismo algoritmo de acoplamiento utilizado, ya que Vina no incorpora en su función de puntuación interacciones de cationes  $\pi$  (96 – 97).

El desarrollo de péptidos anti adhesivos representa un área de vasto interés en la ciencia hoy en día ya que representa una forma de combatir la creciente crisis que existe referente a la resistencia de antibióticos lo que dificulta el tratamiento de varias enfermedades, en este caso las caries (97 – 98). El desarrollo de este péptido presenta varios puntos positivos. El hecho de usar como diana a la región C – terminal y no la sección A3VP1 de la adhesina, va de la mano con el hecho de que la estructura de la región C – terminal a diferencia de la región variable, se encuentra conservada hasta en un 62% de las demás cepas de *S. mutans*. Adicionalmente, el intervenir el correcto plegamiento de esta estructura permite remover la resistencia proteolítica que poseen estas bacterias y de esta manera poder impedir la adhesión y la colonización, hecho que ya ha sido demostrado en otras bacterias gram positivas (100 – 101). En el caso específico de *S. mutans* el interferir con el plegamiento de la región C – terminal de Ag I/II implica intervenir la unión multivalente que tiene la bacteria a la superficie. Esta particular unión,

como bien se mencionó en el apartado de introducción, genera que la fuerza total que se aplica en la bacteria se divida en el número total de uniones que se generan entre la bacteria y la superficie. Al intervenir aunque sea una de las uniones, hace que las demás secciones reciban mayores magnitudes de fuerza o bien que esta sea degradada por proteasas dada la pérdida de la resistencia proteolítica, lo que hará que los distintos sitios de unión de la proteína reciban aun más mayores magnitudes de fuerzas dada el desequilibrio que se provocará en la transducción de las fuerzas, y si a esto sumamos que los *pili* de *S. mutans* de por sí pueden ser desprendidos a bajos índices energéticos, podría significar el desprendimiento completo de la bacteria de la superficie del diente.

## 6 PROYECCIONES

Los resultados presentados en este trabajo abren las puertas para un posible tratamiento preventivo al desarrollo de las caries producto de la colonización por la bacteria *Streptococcus mutans*. A un futuro, dada la naturaleza *in silico* de los resultados obtenidos, éstos deben ser extrapolados para comenzar una fase de experimentación *in vitro* con tal de poder evaluar el potencial biológico de la propuesta. Dentro de ella, se tiene que evaluar la efectividad del péptido en ensayos de molécula individual con tal de simular como actuaría este *in vivo*, resultados que según las mismas simulaciones hechas dentro del desarrollo de esta propuesta deberían dar buenos resultados. En cuanto al potencial biotecnológico del péptido este podría ser acoplado a un dentífrico o a algún tipo de enjuague bucal a modo de funcionar como una manera profiláctica de tratar el desarrollo de las caries.

Este proyecto da los primeros pasos hacia estrategias de desarrollo de moléculas capaces de intervenir la catalización de enlaces relacionados a la estabilidad de bacterias Gram - positivo, logrando de igual manera intervenir procesos de adhesión relacionados.

## 7 CONCLUSIONES

- Mediante el uso de herramientas *in silico* se fue capaz de generar una molécula que mostró ser más estable mecánicamente y más afín en comparación a las estructuras silvestres del C – terminal de Ag I/II, cumpliendo así lo propuesto en la hipótesis del trabajo.
- Se logró realizar de las primeras caracterizaciones mecánicas respecto a cómo funcionan los enlaces isopeptídicos del C – terminal de Ag I/II, logrando de igual manera identificar el más relevante.
- El protocolo propuesto de escaneo de alaninas y refinamiento de péptidos representó una rápida y efectiva manera para encontrar péptidos con buen potencial para realizar intervención molecular no sólo para este proyecto, sino que también para otros desarrollos de esta índole.
- El péptido generado cumplió con las características propuestas, formando así buenos cimientos para poder extrapolar estos resultados a condiciones *in vitro*.

## 8 REFERENCIAS:

1. Hussain A, Khan FA. History of dentistry. *Archives of Medicine and Health Sciences*. 2014 Jan 1;2(1):106.
2. Marthaler TM. Changes in dental caries 1953–2003. *Caries research*. 2004;38(3):173-81.
3. Touger-Decker R, Van Loveren C. Sugars and dental caries. *The American journal of clinical nutrition*. 2003 Oct 1;78(4):881S-92S.
4. Rosier BT, Marsh PD, Mira A. Resilience of the oral microbiota in health: mechanisms that prevent dysbiosis. *Journal of dental research*. 2018 Apr;97(4):371-80.
5. Lăzureanu PC, Popescu F, Tudor A, Stef L, Negru AG, Mihăilă R. Saliva pH and Flow Rate in Patients with Periodontal Disease and Associated Cardiovascular Disease. *Medical Science Monitor: International Medical Journal of Experimental and Clinical Research*. 2021;27:e931362-1.
6. Bagramian RA, Garcia-Godoy F, Volpe AR. The global increase in dental caries. A pending public health crisis. *Am J dent*. 2009 Feb 1;22(1):3-8.
7. Vos T, Barber RM, Bell B, Bertozzi-Villa A, Biryukov S, Bolliger I, Charlson F, Davis A, Degenhardt L, Dicker D, Duan L. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *The lancet*. 2015 Aug 22;386(9995):743-800.
8. Petersen PE, Bourgeois D, Ogawa H, Estupinan-Day S, Ndiaye C. The global burden of oral diseases and risks to oral health. *Bulletin of the world health organization*. 2005;83:661-9.
9. Mishu MP, Hobdell M, Khan MH, Hubbard RM, Sabbah W. Relationship between untreated dental caries and weight and height of 6-to 12-year-old primary school children in Bangladesh. *International journal of dentistry*. 2013 Apr 10;2013.
10. Featherstone JD. Dental caries: a dynamic disease process. *Australian dental journal*. 2008 Sep;53(3):286-91.
11. Pitts NB, Zero DT, Marsh PD, Ekstrand K, Weintraub JA, Ramos-Gomez F, Tagami J, Twetman S, Tsakos G, Ismail A. Dental caries. *Nature reviews Disease primers*. 2017 May 25;3(1):1-6.
12. Marsh PD. Microbial ecology of dental plaque and its significance in health and disease. *Advances in dental research*. 1994 Jul;8(2):263-71.
13. Loesche WJ. Role of *Streptococcus mutans* in human dental decay. *Microbiological reviews*. 1986 Dec;50(4):353-80.

14. Dye BA, Thornton-Evans G, Li X, Iafolla T. Dental caries and tooth loss in adults in the United States, 2011-2012. Hyattsville, MD, USA: US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics; 2015 May 1.
15. Banerjee A. *Applications of scanning microscopy in the assessment of dentine caries and methods for its removal* (Doctoral dissertation, King's College London (University of London)).
16. Berkowitz RJ. Causes, treatment and prevention of early childhood caries: a microbiologic perspective. *Journal-Canadian Dental Association*. 2003 May 1;69(5):304-7.
17. Bader JD, Shugars DA. What do we know about how dentists make caries-related treatment decisions?. *Community dentistry and oral epidemiology*. 1997 Feb;25(1):97-103.
18. Zero DT, Fontana M, Martínez-Mier EA, Ferreira-Zandoná A, Ando M, González-Cabezas C, Bayne S. The biology, prevention, diagnosis and treatment of dental caries: scientific advances in the United States. *The Journal of the American Dental Association*. 2009 Sep 1;140:25S-34S.
19. Mass E, Zilberman DU. Clinical and radiographic evaluation of partial in carious exposure of permanent molars. *Pediatric dentistry*. 1993 Jul;15(4):257.
20. de Alencar CR, de Oliveira GC, Tripodi CD, Gonçalves PS, Ionta FQ, Honorio HM, Oliveira TM, Rios D. Dental Plaque Disclosing as an Auxiliary Method for Professional Dental Prophylaxis in Early Childhood. *International Journal of Clinical Pediatric Dentistry*. 2019 May;12(3):189.
21. Høiby N, Bjarnsholt T, Givskov M, Molin S, Ciofu O. Antibiotic resistance of bacterial biofilms. *International journal of antimicrobial agents*. 2010 Apr 1;35(4):322-32.
22. De la Fuente-Núñez C, Reffuveille F, Fernández L, Hancock RE. Bacterial biofilm development as a multicellular adaptation: antibiotic resistance and new therapeutic strategies. *Current opinion in microbiology*. 2013 Oct 1;16(5):580-9.
23. Munson MA, Banerjee A, Watson TF, Wade WG. Molecular analysis of the microflora associated with dental caries. *Journal of clinical microbiology*. 2004 Jul;42(7):3023-9.
24. Marsh PD. Dental plaque as a biofilm: the significance of pH in health and caries. *Compendium of continuing education in dentistry* (Jamesburg, NJ: 1995). 2009 Mar 1;30(2):76-8.
25. Daboor SM, Masood FS, Al-Azab MS, Nori EE. A review on streptococcus mutans with its diseases dental caries, dental plaque and endocarditis. *Indian J Microbiol Res*. 2015;2(2):76-82.
26. Nomura R, Nakano K, Nemoto H, Fujita K, Inagaki S, Takahashi T, Taniguchi K, Takeda M, Yoshioka H, Amano A, Ooshima T. Isolation and characterization of Streptococcus mutans in heart valve and dental plaque specimens from a patient with infective endocarditis. *Journal of medical microbiology*. 2006 Aug 1;55(8):1135-40.

27. Mansour TR, Alam Y, Dahbour L, Alnemari A, Jumaa M, Schroeder JL. Streptococcus mutans: a potential risk factor in recurrent hemorrhagic stroke. *Cureus*. 2017 May 19;9(5).
28. Hillman JD. Genetically modified *Streptococcus mutans* for the prevention of dental caries. *Antonie Van Leeuwenhoek*. 2002 Aug;82(1):361-6.
29. Lemos JA, Palmer SR, Zeng L, Wen ZT, Kajfasz JK, Freires IA, Abranches J, Brady LJ. The biology of *Streptococcus mutans*. *Microbiology spectrum*. 2019 Jan 18;7(1):7-1.
30. Huddleston JR. Horizontal gene transfer in the human gastrointestinal tract: potential spread of antibiotic resistance genes. *Infection and drug resistance*. 2014;7:167.
31. Mohammadali M, Davies J. Antimicrobial resistance genes and wastewater treatment. *Antimicrobial Resistance in Wastewater Treatment Processes*. 2017 Dec 12:1-3.
32. Rossolini GM, Arena F, Pecile P, Pollini S. Update on the antibiotic resistance crisis. *Current opinion in pharmacology*. 2014 Oct 1;18:56-60.
33. Narita M, Shibahara T, Takano N, Fujii R, Okuda K, Ishihara K. Antimicrobial susceptibility of microorganisms isolated from periapical periodontitis lesions. *The Bulletin of Tokyo Dental College*. 2016;57(3):133-42.
34. Bennadi D. Self-medication: A current challenge. *Journal of basic and clinical pharmacy*. 2013 Dec;5(1):19.
35. Loyola-Rodriguez JP, Ponce-Diaz ME, Loyola-Leyva A, Garcia-Cortes JO, Medina-Solis CE, Contreras-Ramire AA, Serena-Gomez E. Determination and identification of antibiotic-resistant oral streptococci isolated from active dental infections in adults. *Acta Odontologica Scandinavica*. 2018 May 19;76(4):229-35.
36. Schroeder M, Brooks BD, Brooks AE. The complex relationship between virulence and antibiotic resistance. *Genes*. 2017 Jan 18;8(1):39.
37. Dessen A, Di Guilmi AM, Vernet T, Dideberg O. Molecular mechanisms of antibiotic resistance in gram-positive pathogens. *Current Drug Targets-Infectious Disorders*. 2001 May 1;1(1):63-77.
38. Vázquez-Laslop N, Mankin AS. How macrolide antibiotics work. *Trends in biochemical sciences*. 2018 Sep 1;43(9):668-84.
39. Fischetti VA, Novick RP, Ferretti JJ, Portnoy DA, Rood JI. *Gram-positive pathogens*. ASM Press; 2006.
40. Fourmy D, Recht MI, Blanchard SC, Puglisi JD. Structure of the A site of *Escherichia coli* 16 S ribosomal RNA complexed with an aminoglycoside antibiotic. *Science*. 1996 Nov 22;274(5291):1367-71.

41. Hooper DC. Mechanisms of action and resistance of older and newer fluoroquinolones. *Clinical infectious diseases*. 2000 Aug 1;31(Supplement\_2):S24-8.
42. Spratt BG. Resistance to antibiotics mediated by target alterations. *Science*. 1994 Apr 15;264(5157):388-93.
43. Kotb DN, Mahdy WK, Mahmoud MS, Khairy RM. Impact of co-existence of PMQR genes and QRDR mutations on fluoroquinolones resistance in Enterobacteriaceae strains isolated from community and hospital acquired UTIs. *BMC infectious diseases*. 2019 Dec;19(1):1-8.
44. Van Heijenoort J. Assembly of the monomer unit of bacterial peptidoglycan. *Cellular and Molecular Life Sciences CMLS*. 1998 Apr;54(4):300-4.
45. Tooke CL, Hinchliffe P, Bragginton EC, Colenso CK, Hirvonen VH, Takebayashi Y, Spencer J.  $\beta$ -Lactamases and  $\beta$ -Lactamase Inhibitors in the 21st Century. *Journal of molecular biology*. 2019 Aug 23;431(18):3472-500.
46. Walsh C. Molecular mechanisms that confer antibacterial drug resistance. *Nature*. 2000 Aug;406(6797):775-81.
47. Persat A, Nadell CD, Kim MK, Ingremeau F, Siryaporn A, Drescher K, Wingreen NS, Bassler BL, Gitai Z, Stone HA. The mechanical world of bacteria. *Cell*. 2015 May 21;161(5):988-97.
48. Biais N, Ladoux B, Higashi D, So M, Sheetz M. Cooperative retraction of bundled type IV pili enables nanonewton force generation. *PLoS biology*. 2008 Apr;6(4):e87.
49. Maier B, Potter L, So M, Seifert HS, Sheetz MP. Single pilus motor forces exceed 100 pN. *Proceedings of the National Academy of Sciences*. 2002 Dec 10;99(25):16012-7.
50. Oho T, Yu H, Yamashita Y, Koga T. Binding of salivary glycoprotein-secretory immunoglobulin A complex to the surface protein antigen of *Streptococcus mutans*. *Infection and Immunity*. 1998 Jan 1;66(1):115-21.
51. Tao L, Tanzer JM. Novel sucrose-dependent adhesion co-factors in *Streptococcus mutans*. *Journal of dental research*. 2002 Jul;81(7):505-10.
52. Matsumoto-Nakano M. Role of *Streptococcus mutans* surface proteins for biofilm formation. *Japanese Dental Science Review*. 2018 Feb 1;54(1):22-9.
53. Larson MR, Rajashankar KR, Crowley PJ, Kelly C, Mitchell TJ, Brady LJ, Deivanayagam C. Crystal Structure of the C-terminal Region of *Streptococcus mutans* Antigen I/II and Characterization of Salivary Agglutinin Adherence Domains\*♦. *Journal of Biological Chemistry*. 2011 Jun 17;286(24):21657-66.

54. Ligtenberg AJ, Veerman EC, Nieuw Amerongen AV, Mollenhauer J. Salivary agglutinin/glycoprotein-340/DMBT1: a single molecule with variable composition and with different functions in infection, inflammation and cancer.
55. Madsen J, Mollenhauer J, Holmskov U. Gp-340/DMBT1 in mucosal innate immunity. *Innate immunity*. 2010 Jun;16(3):160-7.
56. Purushotham S, Deivanayagam C. The calcium-induced conformation and glycosylation of scavenger-rich cysteine repeat (SRCR) domains of glycoprotein 340 influence the high affinity interaction with antigen I/II homologs. *Journal of Biological Chemistry*. 2014 Aug 8;289(32):21877-87.
57. Brady LJ, Maddocks SE, Larson MR, Forsgren N, Persson K, Deivanayagam CC, Jenkinson HF. The changing faces of *Streptococcus* antigen I/II polypeptide family adhesins. *Molecular microbiology*. 2010 Jul;77(2):276-86.
58. Larson MR, Rajashankar KR, Patel MH, Robinette RA, Crowley PJ, Michalek S, Brady LJ, Deivanayagam C. Elongated fibrillar structure of a streptococcal adhesin assembled by the high-affinity association of  $\alpha$ - and PPII-helices. *Proceedings of the National Academy of Sciences*. 2010 Mar 30;107(13):5983-8.
59. Brady LJ, Crowley PJ, Ma JK, Kelly C, Lee SF, Lehner T, Bleiweis AS. Restriction fragment length polymorphisms and sequence variation within the spaP gene of *Streptococcus mutans* serotype c isolates. *Infection and immunity*. 1991 May;59(5):1803-10.
60. Sullan RM, Li JK, Crowley PJ, Brady LJ, Dufrêne YF. Binding forces of *Streptococcus mutans* P1 adhesin. *ACS nano*. 2015 Feb 24;9(2):1448-60.
61. Proft T, Baker EN. Pili in Gram-negative and Gram-positive bacteria—structure, assembly and their role in disease. *Cellular and molecular life sciences*. 2009 Feb;66(4):613-35.
62. Forsgren N, Lamont RJ, Persson K. Two intramolecular isopeptide bonds are identified in the crystal structure of the *Streptococcus gordonii* SspB C-terminal domain. *Journal of molecular biology*. 2010 Apr 2;397(3):740-51.
63. Walden M, Crow A, Nelson MD, Banfield MJ. Intramolecular isopeptide but not internal thioester bonds confer proteolytic and significant thermal stability to the *S. pyogenes* pilus adhesin Spy0125. *Proteins: Structure, Function, and Bioinformatics*. 2014 Mar;82(3):517-27.
64. Hall M, Nylander Å, Jenkinson HF, Persson K. Structure of the C-terminal domain of AspA (antigen I/II-family) protein from *Streptococcus pyogenes*. *FEBS Open Bio*. 2014 Jan 1;4:283-9.

65. Nylander Å, Forsgren N, Persson K. Structure of the C-terminal domain of the surface antigen SpaP from the caries pathogen *Streptococcus mutans*. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications*. 2011 Jan 1;67(1):23-6.
66. Forsgren N, Lamont RJ, Persson K. Two intramolecular isopeptide bonds are identified in the crystal structure of the *Streptococcus gordonii* SspB C-terminal domain. *Journal of molecular biology*. 2010 Apr 2;397(3):740-51.
67. Kang HJ, Baker EN. Intramolecular isopeptide bonds give thermodynamic and proteolytic stability to the major pilin protein of *Streptococcus pyogenes*. *Journal of Biological Chemistry*. 2009 Jul 31;284(31):20729-37.
68. Alegre-Cebollada J, Badilla CL, Fernández JM. Isopeptide bonds block the mechanical extension of pili in pathogenic *Streptococcus pyogenes*. *Journal of Biological Chemistry*. 2010 Apr 9;285(15):11235-42.
69. Rivas-Pardo JA, Badilla CL, Tapia-Rojo R, Alonso-Caballero Á, Fernández JM. Molecular strategy for blocking isopeptide bond formation in nascent pilin proteins. *Proceedings of the National Academy of Sciences*. 2018 Sep 11;115(37):9222-7.
70. Siu SW, Pluhackova K, Böckmann RA. Optimization of the OPLS-AA force field for long hydrocarbons. *Journal of Chemical theory and Computation*. 2012 Apr 10;8(4):1459-70.
71. Pajankar A, Pajankar A. Introduction to Python and Digital Image Processing. *Raspberry Pi Image Processing Programming: Develop Real-Life Examples with Python, Pillow, and SciPy*. 2017:25-39.
72. Marginean A, Barr ET, Harman M, Jia Y. Automated transplantation of call graph and layout features into Kate. *In Search-Based Software Engineering: 7th International Symposium, SSBSE 2015, Bergamo, Italy, September 5-7, 2015, Proceedings 7 2015* (pp. 262-268). Springer International Publishing.
73. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJ. GROMACS: fast, flexible, and free. *Journal of computational chemistry*. 2005 Dec;26(16):1701-18.
74. Cowan R, Grosdidier G. Visualization tools for monitoring and evaluation of distributed computing systems. *In Proc. of the International Conference on Computing in High Energy and Nuclear Physics, Padova, Italy 2000* Feb 7.
75. Harney AS, Arwert EN, Entenberg D, Wang Y, Guo P, Qian BZ, Oktay MH, Pollard JW, Jones JG, Condeelis JS. Real-Time Imaging Reveals Local, Transient Vascular Permeability, and Tumor Cell

- Intravasation Stimulated by TIE2hi Macrophage-Derived VEGFA Imaging the Tumor Microenvironment of Metastasis. *Cancer discovery*. 2015 Sep 1;5(9):932-43.
76. Dallakyan S, Olson AJ. Small-molecule library screening by docking with PyRx. *Chemical biology: methods and protocols*. 2015:243-50.
77. Van Rossum G. An introduction to Python. Drake FL, editor. Bristol: Network Theory Ltd.; 2003 Apr.
78. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of computational chemistry*. 2004 Oct;25(13):1605-12.
79. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *Journal of molecular graphics*. 1996 Feb 1;14(1):33-8.
80. DeLano WL. Pymol: An open-source molecular graphics tool. *CCP4 Newsl. Protein Crystallogr*. 2002 Mar;40(1):82-92.
81. Li Z, Scheraga HA. Monte Carlo-minimization approach to the multiple-minima problem in protein folding. *Proceedings of the National Academy of Sciences*. 1987 Oct;84(19):6611-5.
82. Kelly CG, Younson JS, Hikmat BY, Todryk SM, Czisch M, Haris PI, Flindall IR, Newby C, Mallet AI, Ma JK, Lehner T. A synthetic peptide adhesion epitope as a novel antimicrobial agent. *Nature biotechnology*. 1999 Jan;17(1):42-7.
83. Hendrickx A, Budzik JM, Oh SY, Schneewind O. Architects at the bacterial surface—sortases and the assembly of pili with isopeptide bonds. *Nature Reviews Microbiology*. 2011 Mar;9(3):166-76.
84. Bahar AA, Ren D. Antimicrobial peptides. *Pharmaceuticals*. 2013 Nov 28;6(12):1543-75.
85. Melo MC, Gomes DE, Bernardi RC. Molecular origins of force-dependent protein complex stabilization during bacterial infections. *Journal of the American Chemical Society*. 2022;
86. Beckloff N, Laube D, Castro T, Furgang D, Park S, Perlin D, Clements D, Tang H, Scott RW, Tew GN, Diamond G. Activity of an antimicrobial peptide mimetic against planktonic and biofilm cultures of oral pathogens. *Antimicrobial agents and chemotherapy*. 2007 Nov;51(11):4125-32.
87. Sedlak SM, Schendel LC, Gaub HE, Bernardi RC. Streptavidin/biotin: tethering geometry defines unbinding mechanics. *Science advances*. 2020 Mar 25;6(13):eaay5999.
88. Sedlak SM, Schendel LC, Gaub HE, Bernardi RC. Streptavidin/biotin: tethering geometry defines unbinding mechanics. *Science advances*. 2020 Mar 25;6(13):eaay5999.

89. Vengadesan K, Narayana SV. Structural biology of Gram-positive bacterial adhesins. *Protein science*. 2011 May;20(5):759-72.
90. Moreira IS, Fernandes PA, Ramos MJ. Hot spots—A review of the protein–protein interface determinant amino-acid residues. *Proteins: Structure, Function, and Bioinformatics*. 2007 Sep;68(4):803-12.
91. Križ ½ger DM, Gohlke H. DrugScorePPI webserver: fast and accurate in silico alanine scanning for scoring protein–protein interactions. *Nucleic acids research*. 2010 Jul 1;38(suppl\_2):W480-6.
92. Pagadala NS, Syed K, Tuszynski J. Software for molecular docking: a review. *Biophysical reviews*. 2017 Apr;9(2):91-102.
93. Kuo SC, Lauffenburger DA. Relationship between receptor/ligand binding affinity and adhesion strength. *Biophysical journal*. 1993 Nov 1;65(5):2191-200.
94. Dougherty DA. Cation- $\pi$  interactions involving aromatic amino acids. *The Journal of nutrition*. 2007 Jun 1;137(6):1504S-8S.
95. Ma JC, Dougherty DA. The cation-  $\pi$  interaction. *Chemical reviews*. 1997 Aug 5;97(5):1303-24.
96. Quiroga R, Villarreal MA. Vinardo: A scoring function based on autodock vina improves scoring, docking, and virtual screening. *PloS one*. 2016 May 12;11(5):e0155183.
97. Quiroga R, Villarreal MA. Vinardo: A scoring function based on autodock vina improves scoring, docking, and virtual screening. *PloS one*. 2016 May 12;11(5):e0155183.
98. Yu K, Lo JC, Yan M, Yang X, Brooks DE, Hancock RE, Lange D, Kizhakkedathu JN. Anti-adhesive antimicrobial peptide coating prevents catheter associated infection in a mouse urinary infection model. *Biomaterials*. 2017 Feb 1;116:69-81.
99. Alaki SM, Burt BA, Garetz SL. The association between antibiotics usage in early childhood and early childhood caries. *Pediatric dentistry*. 2009 Jan 15;31(1):31-7.
100. Contreras F, Rivas-Pardo JA. Interfering with the Folding of Group A Streptococcal pili Proteins. In *Group A Streptococcus 2020* (pp. 347-364). Humana, New York, NY.
101. Rivas-Pardo JA, Badilla CL, Tapia-Rojo R, Fernández JM. Isopeptide-Blocker Impairs the Mechanics of Recently Translated Pili Proteins.

## ANEXO

```
## Comandos en La Tk consola de VMD para limpiar y centrar La proteína en su
centro de masa

# Comando para sólo extraer La cadena A del cristal 3QE5
> Set all [atomselect top "chain A"]
> $all writepdb ChainA.pdb

# Abrir pdb ChainA.pdb
# Comando para limpiar La cadena A
> Set all1 [atomselect top "protein"]
> $all1 writepdb protein.pdb

## Set de comandos para centrar La proteína en su centro de masas en el eje z
> set all [ atomselect top "protein"]
> $all writepdb pan.pdb
> $all moveby [vecinvert [measure center $all]]
> measure center $all
> $all moveby {0 0 0}
# Probar moviendo en el eje y, hasta centrar en el eje Z
> $all move [transaxis y 0]
> $all writepdb pili_limpia_centrada_Z.pdb
```

Caja S1: Conjunto de líneas de código utilizadas en la terminal de comandos de VMD para limpiar y centrar la proteína en el eje z.

```

# Generación del archivo de topología
$gmx pdb2gmx -f I10.pdb -o i102.gro

# Generada La topología, incluir Las siguientes líneas en el archivo, específicamente luego de La línea 21
## Inclusión necesaria para que el campo de fuerza sepa que tipo de aguas tiene que usar
    #include "./sirah.ff/hybsol_comb3.itp"
    #include "./sirah.ff/solv.itp"

# Generación de La caja y cálculo de aguas del sistema
$gmx editconf -f pili.gro -o pili_box.gro -box x y z -c
$gmx solvate -cp pili_box.gro -cs spc216.gro -o pili_shell.gro -shell 1
$gmx solvate -cp pili_shell.gro -cs ./sirah.ff/wt4tip3p.gro -o pili_sol.gro

# Conteo del número de aguas en el sistema
$grep -c OW pili_sol.gro; grep -c WP1 pili_sol.gro
# Agregar Los resultados a La topología

$gmx grompp -f ./sirah.ff/tutorial/2/GPU/em_HYBSOL.mdp -p topol.top -c
pili_sol.gro -o pili_sol.tpr -maxwarn 1

# Agregar iones sodio y cloro acorde al cálculo hecho en base a aguas atomísticas
$gmx genion -s pili_sol.tpr -o pili_ion.gro -np [Número de iones sodio] -pname NaW
-nn [Número de iones cloro] -nname CLW
## Reemplazar aguas WT4 cuando pregunte La consola

# Contar aguas nuevamente, pero esta vez con el archivo generado ion.gro
$grep -c OW pili_ion.gro; grep -c WP1 pili_ion.gro
# Reemplazar Los valores obtenidos más Los iones en el archivo de topología

# Crear un archivo .psf
$./sirah.ff/tools/g_top2psf.pl -i topol.top -o pili_ion.psf
## EL archivo psf se hace manera de convertir La topología en un archivo que pueda interpretar Los #include mencionados anteriormente, considerando cadenas, moléculas de solvente, ligando u otras moléculas agregadas al sistema

# El archivo psf es necesario para La creación del archivo index el cuál posee todos números asociados al sistema de La dinámica molecular.}
# Previo a ello, se deben reemplazar Los códigos oplS del archivo psf ya que dicha nomenclatura puede leída sólo por el campo de fuerza oplSAA
# Los reemplazos asociados se encuentran en La tablaS1 adjunta

# Creación del archivo index
$echo -e "r WT4 | r NaW | r CLW\nq\n " | gmx make_ndx -f pili_ion.gro -o
pili_ion.ndx

```

```

# Comenzar proceso de minimizado
$gmx grompp -f ./sirah.ff/tutorial/2/GPU/em_HYBSOL.mdp -p topol.top -po em.mdp -
n pili_ion.ndx -c pili_ion.gro -o pili_em.tpr
$gmx mdrun -v -deffnm pili_em > EM.log &

# Terminado el minimizado comenzar el proceso de equilibrado
$gmx grompp -f sirah.ff/tutorial/2/GPU/eq_HYBSOL2.mdp -p topol.top -po eq.mdp -n
pili_ion.ndx -c pili_em.gro -r pili_em.gro -o pili_eq.tpr
$gmx mdrun -v -deffnm pili_eq > EQ.log &

# Previo a La dinámica molecular colocar Los número correspondiente a Los puntos
de estiramiento al final del archivo index
[pull1]
Carbono alfa 1
[pull2]
Carbono alfa 2

# Modificar el archivo de configuración de La dinámica para a correr a velocidad
constante
;Pull code
pull                               = yes
pull-ncoords                        = 1
pull-ngroups                        = 2
pull-group1-name                    = pull1
pull-group2-name                    = pull2
pull-coord1-groups                  = 1 2
pull-coord1-geometry                = distance
pull-coord1-dim                     = N N Y
pull_coord1_start                   = yes
pull_coord1_rate                    = 0.01      ; magnitud de velocidad en nm/ps
pull-print-components               = yes
pull_coord1_k                       = 1000      ; kJ mol-1 nm-2
## Pegar Las variables al final del archivo de configuración

# Comenzar La dinámica molecular
$gmx grompp -f ./sirah.ff/tutorial/2/GPU/md_HYBSOL2.mdp -p topol.top -po md.mdp -
n pili_ion.ndx -c pili_eq.gro -o pili_md.tpr
$gmx mdrun -v -deffnm pili_md &> MD.log &

# Transformar Los datos para ver trayectoria en vmd
$gmx trjconv -s pili_em.tpr -f pili_md.xtc -o pili_md_pbc.xtc -n pili_ion.ndx -
ur compact -center -pbc mol
# Cuando La consola lo pida especificar que cargue La proteína y luego el sistema
$vmd pili_ion.psf pili_ion.gro pili_md_pbc.xtc -e
./sirah.ff/tools/sirah_vmdtk.tcl

# Aislar columnas de Los archivos de estiramiento
$awk '{print $2}' pili_md_pullx.xvg > pili_pullx.xvg
$awk '{print $2}' pili_md_pulld.xvg > pili_pulld.xvg

```

```
## Transformar La data del estiramiento de fuerza en pN y reemplazar Los datos
```

```
# Juntas ambas columnas
```

```
$paste pili_pullx.svg pili_pullf.svg > pili_pullsFX.svg
```

```
# Visualizar La data en xmgrace
```

```
$xmgrace pili_pullsFX.svg
```

CajaS2: Códigos utilizados para el desarrollo de la dinámica molecular

opls_272	O2
opls_271	C
opls_155	HO
opls_154	OH
opls_140	H1
opls_157	CT
opls_236	O
opls_241	H
opls_136	CT
opls_240	H
opls_246	CT
opls_223B	CT
opls_308	CT
opls_303	N2
opls_302	CA
opls_301	H
opls_292	CT

opls_145	CA
opls_166	C
opls_168	H2
opls_202	S
opls_505	CT
opls_511	NB
opls_287	N
opls_290	H
opls_293B	CT
opls_135	CT
opls_137	CT
opls_235	C
opls_238	N
opls_224B	CT
opls_237	N
opls_239	N
opls_245	CT
opls_274	CT
opls_307	CT
opls_304	H
opls_300	N2
opls_158	CT

opls_149	CT
opls_146	HA
opls_167	OH
opls_210	CT
opls_209	CT
opls_507	CC
opls_508	CW
opls_506	CR
opls_116	OW
opls_503	NA
opls_117	HW
opls_500	C*
opls_514	CW
opls_501	CB
opls_502	CN
opls_283	CT
opls_504	H

Tabla S1: Tabla usada para reemplazar los términos opls presentes en el archivo “.psf”. El proceso de reemplazo se hizo con tal de que los parámetros del campo de fuerza SIRAH fueran capaces de leer correctamente el archivo y de esta manera poder continuar con el desarrollo de la SMD