

# **PROYECTO FINAL DE BIOTECNOLOGÍA**

## **PRODUCCIÓN DE LA MAQUINARIA DE FUSIÓN DE MEMBRANAS DEL VIRUS ZIKA EN *E. COLI* PARA EL TESTEO DE UNA VARIANTE ESTABILIZADORA.**

**Toledo, Valentina – LU: 1162729**

**Licenciatura en Biotecnología**

**Tutor:**

**Dr. Dellarole, Mariano, CIBION-CONICET**

**Co-Tutor:**

**Vadillo Martín, UADE**

**Año 2025**

**UADE**



**UNIVERSIDAD ARGENTINA DE LA EMPRESA  
FACULTAD DE INGENIERÍA Y CIENCIAS EXACTAS**

## ÍNDICE

<b>1. AGRADECIMIENTOS.....</b>	<b>3</b>
<b>2. RESUMEN.....</b>	<b>4</b>
<b>3. ABSTRACT.....</b>	<b>5</b>
<b>4. INTRODUCCIÓN.....</b>	<b>6</b>
a. Los Flavivirus como problema sanitario emergente.....	6
b. La proteína E del virus Zika: estructura y función.....	7
c. Mecanismo de entrada y fusión viral mediada por E.....	8
d. Diseño racional y expresión recombinante de variantes estabilizadas de la proteína E.....	10
<b>5. ESTADO DEL ARTE.....</b>	<b>12</b>
a. Estudios estructurales de la proteína E en flavivirus.....	12
b. Limitaciones de las proteínas recombinantes monoméricas.....	12
c. Aplicaciones biomédicas de proteínas E estabilizadas.....	12
<b>6. HIPÓTESIS.....</b>	<b>14</b>
<b>7. OBJETIVOS.....</b>	<b>15</b>
a. Objetivo general.....	15
b. Objetivos específicos.....	15
<b>8. MATERIALES Y MÉTODOS.....</b>	<b>16</b>
a. Diseño in silico y síntesis de la variante doble mutante.....	16
b. Vector de expresión: pET21a(+)......	17
c. Preparación y uso de medios de cultivo LB y 2YT para crecimiento bacteriano..	18
d. Generación de bacterias competentes.....	19
e. Transformación de bacterias: DH5a, C41, BL21.....	20
f. Expresión recombinante inducida con IPTG y escalado de cultivos bacterianos (2 mL, 500 mL, 1 L).....	21
g. Recolección de biomasa, lisis celular y aislamiento de cuerpos de inclusión.....	23
h. Replegado oxidativo.....	24
i. Purificación por afinidad (His-tag) y diálisis.....	26
j. Análisis por SDS-PAGE.....	28
<b>9. RESULTADOS.....</b>	<b>31</b>
a. Diseño racional y validación computacional de variantes de la proteína E mediante predicción de puentes disulfuro y modelado estructural.....	31
b. Validación de expresión recombinante mediante sds page: pruebas de expresión	38
c. Evaluación de pureza y rendimiento.....	44
<b>10. DISCUSIÓN.....</b>	<b>46</b>

<b>11. CONCLUSIONES.....</b>	<b>51</b>
<b>12. BIBLIOGRAFÍA.....</b>	<b>53</b>
<b>13. ANEXOS.....</b>	<b>55</b>

## 1. AGRADECIMIENTOS

Quiero expresar mi profundo agradecimiento a todos quienes de alguna manera participaron y fueron parte fundamental para lograr llegar a cabo este proyecto final.

En primer lugar, deseo agradecer a mi tutor de tesis, el Dr. Mariano Dellarole, por la confianza depositada en mí desde el inicio de este proyecto. Su guía constante, la generosidad con su tiempo y su predisposición a explicar cada concepto con paciencia, fueron pilares fundamentales para la realización de este trabajo. Su apoyo incondicional hizo que el camino de la investigación resultara no solo más claro, sino también mucho más llevadero.

Agradezco también al Instituto CIBION, institución que me brindó el espacio y los recursos necesarios para desarrollar mi tesina, y a la Universidad Argentina de la Empresa (UADE), mi casa de estudios, donde tuve la oportunidad de formarme académicamente y dar los primeros pasos en mi carrera profesional.

A mi familia y amigos, quienes estuvieron presentes en cada etapa, celebrando mis logros y sosteniéndome en los momentos más difíciles. En especial, a mi papá, mi mamá y a mi hermana Paz, quienes no solo siempre me brindaron su apoyo incondicional, sino que también me enseñaron con el ejemplo el valor de la disciplina y el esfuerzo, y me transmitieron la importancia de la pasión por lo que uno hace.

A mis amigos de la facultad, que fueron un sostén fundamental a lo largo de la carrera. A Brisa, amiga y compañera de laboratorio, por su ayuda incansable y por estar siempre presente cuando más lo necesité. Y a Nico, quien me acompañó desde el primer día de cursada hasta el último, siendo no solo un gran amigo, sino también una fuente constante de risas y contención.

A todos ellos, gracias por haber hecho posible este camino.

## 2. RESUMEN

El virus Zika es transmitido a humanos por mosquitos del género *Aedes aegypti* y, desde su irrupción epidémica en el siglo XXI, representa un problema sanitario complejo en países tropicales. La proteína de envoltura E recubre la superficie viral y cumple un rol clave tanto en el ciclo de replicación como en la respuesta inmunitaria. Esta glicoproteína media la entrada viral por fusión de membranas y constituye el principal antígeno reconocido por el sistema inmune, lo que la convierte en un blanco estratégico para el diseño de tratamientos y métodos diagnósticos. En la partícula viral, E se organiza en 180 copias dispuestas como dímeros en simetría icosaédrica. Sin embargo, aislar esta conformación resulta desafiante debido a su tendencia a adoptar estados monoméricos o triméricos con distinta morfología y estabilidad. En este trabajo se diseña una variante estabilizada de E mediante mutagénesis sitio-dirigida *in silico*, incorporando pares de cisteínas orientados a formar puentes disulfuro en la interfase dimerica. Los modelos generados se validan estructuralmente mediante inteligencia artificial (AlphaFold3) y alineamientos tridimensionales. Las secuencias optimizadas se sintetizan y clonan en plásmidos de expresión pET21a(+), y se transforman en cepas competentes de *E. coli* preparadas en el laboratorio. Se evalúan diferentes condiciones experimentales, incluyendo dos cepas de expresión (*E. coli* BL21 y C41), dos medios de cultivo (LB y 2YT) y regímenes de inducción variables. El sistema más eficiente resulta ser la cepa C41(DE3) en medio 2YT con inducción a 37 °C, que permite una expresión robusta de la proteína. Asimismo, se implementa un protocolo de purificación con replegado oxidativo, logrando establecer un esquema de producción de bajo costo para una proteína estructuralmente compleja. En conjunto, este trabajo desarrolla y optimiza una plataforma bacteriana económica y escalable para la producción de proteínas de interés biomédico con estructuras cuaternarias estabilizadas, con potencial aplicación en estudios estructurales y en el desarrollo de candidatos vacunales.

### 3. ABSTRACT

Zika virus is transmitted to humans by *Aedes aegypti* mosquitoes and, since its epidemic emergence in the 21st century, has posed a major public health challenge in tropical regions. The envelope protein E covers the viral surface and plays a central role in both the replication cycle and the immune response. E mediates viral entry through membrane fusion and is the main antigen recognized by the host immune system, making it a strategic target for therapeutic and diagnostic approaches. In the viral particle, 180 copies of E are arranged as dimers in icosahedral symmetry. However, isolating this conformation is challenging due to its intrinsic tendency to adopt monomeric or trimeric states with different morphologies and stabilities. In this work, a stabilized variant of E was designed through *in silico* site-directed mutagenesis by introducing cysteine pairs aimed at forming disulfide bonds at the dimer interface. Structural models were validated using artificial intelligence tools (AlphaFold3) and three-dimensional alignments. Optimized sequences were synthesized, cloned into pET21a(+) expression vectors, and transformed into competent *E. coli* strains generated in the laboratory. Different experimental conditions were evaluated, including two *E. coli* expression strains (BL21 and C41), two culture media (LB and 2YT), and various induction regimes. The most efficient system was C41(DE3) grown in 2YT medium with induction at 37 °C, which enabled robust expression. In addition, an oxidative refolding purification protocol was implemented, establishing a cost-effective workflow for producing this structurally complex protein. Altogether, this study establishes and optimizes a scalable and economical bacterial platform for the production of biomedically relevant proteins with stabilized quaternary structures, providing a foundation for structural studies and the development of vaccine candidates.

## 4. INTRODUCCIÓN

### A. LOS FLAVIVIRUS COMO PROBLEMA SANITARIO EMERGENTE

El presente proyecto se desarrolla en el contexto de un aumento significativo en la incidencia de las Enfermedades Infecciosas Emergentes causadas por virus del género *Flavivirus*. Este género incluye patógenos responsables de enfermedades como la fiebre amarilla, el dengue y la fiebre del Zika, las cuales representan crecientes problemáticas globales que, en la actualidad, están desatendidas. El incremento exponencial de personas infectadas está directamente relacionado con el cambio climático, ya que los mosquitos vectores, pertenecientes al género *Aedes Aegypti*, tienen una distribución muy sensible a las condiciones climáticas tropicales, lo que ha resultado en un aumento de epidemias de gran escala. Los síntomas en la población afectada varían desde cuadros gripales leves hasta para algunos virus fiebre hemorrágica y, en algunos casos, la muerte. En el caso de la fiebre del Zika, aunque muchos infectados no presentan síntomas, los más comunes son fiebre leve, erupción cutánea, conjuntivitis, dolor articular y dolor de cabeza. Las complicaciones más graves incluyen el síndrome de Guillain-Barré y malformaciones congénitas como la microcefalia en bebés de madres infectadas durante el embarazo. Hoy en día la prevención se centra en evitar picaduras de mosquito ya que no hay tratamiento específico contra el Zika, y se recomienda reposo, hidratación y paracetamol para los síntomas.

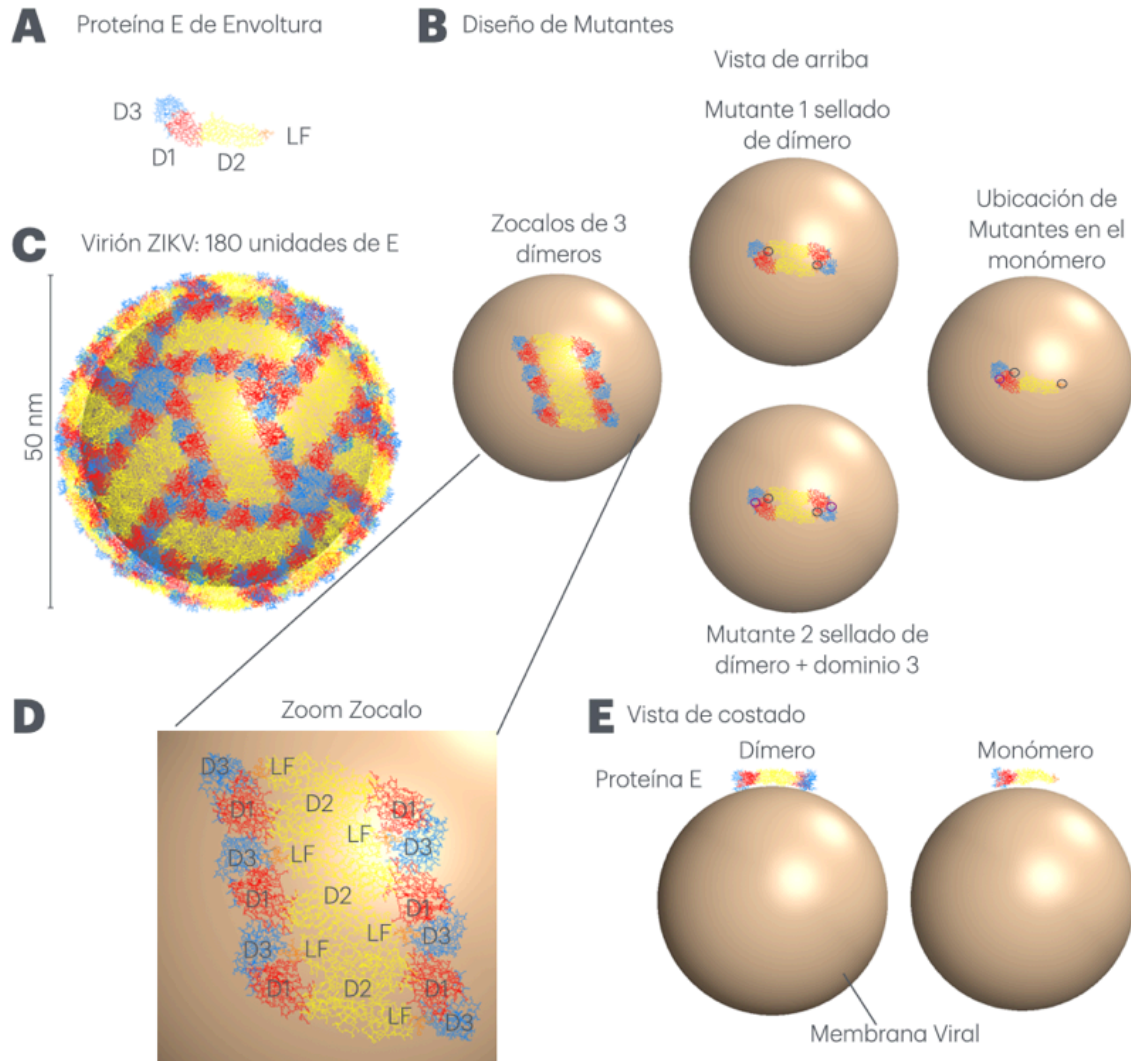
En este contexto global, el laboratorio del Dr. Mariano Dellarole, especializado en Biofísica de Virus, lleva adelante un análisis exhaustivo de las proteínas estructurales E y Pr, que componen la envoltura de estos virus y son esenciales tanto para el proceso de entrada a la célula huésped como para la maduración de las partículas virales infectivas. Estudios recientes (Crampon et al., 2023; Vaney et al., 2022; Rey et al., 2018) han identificado las bases de estabilización de las proteínas E y Pr en el proceso de maduración del virión. A partir de esta información, que revela las zonas clave de la superficie molecular implicadas en la fusión de membranas y la estabilidad de los epítopes, se plantean los objetivos de este proyecto.

## B. LA PROTEÍNA E DEL VIRUS ZIKA: ESTRUCTURA Y FUNCIÓN

Los flavivirus poseen un genoma de 10,7 kb que codifica una sola poliproteína la cual se escinde en una decena de proteínas, siendo las 3 estructurales: C formadora de la cápside, y prM y E, glicoproteínas de envoltura.

En este caso concretamente nos focalizamos en el estudio de la proteína "E" de envoltura, ver Figura 1, la cual se encuentra expuesta en la superficie del virus maduro. Cada protómero de la proteína E está compuesta por ~500 aminoácidos en su versión completa o ~400 aminoácidos en su versión soluble (Figura 1A). Con un peso molecular de 46 kDa, su naturaleza es glicoproteica, con hasta dos asparaginas glicosiladas. E, pertenece al grupo de proteínas de fusión de tipo II dado que su estructura secundaria está compuesta principalmente por láminas beta. Su estructura terciaria está organizada en dos dominios interconectados, llamados D1 (o *central*) y D2 (o *de dimerización*) y un dominio globular llamado D3 (o *cabeza*), los cuales se representan usualmente con los colores primarios, rojo, amarillo y azul, respectivamente. E posee un plegamiento tridimensional complejo el cual se estabiliza mediante seis puentes disulfuro conservados en toda la familia.

E posee la flexibilidad de formar homo-oligómeros diméricos y triméricos. Pero en la superficie viral madura, E se organiza como arreglos diméricos (Figura 1) que *esconden* el bucle hidrofóbico de fusión (Figura 1A coloreado en naranja), el cual se ubica en la punta del dominio 2 (D2). Tres dímeros de E forman un zócalo, 30 zócalos se organizan en un patrón llamado “de espina de pez”, sumando 90 dímeros que empaquetan la totalidad de la superficie viral en 50 nm de diámetro y formando una simetría pseudo icosaédrica (Figura 1C).



**Figura 1:** A) Proteína E de envoltura viral, en conformación monomérica, con sus tres dominios coloreados para DI (en rojo), DII (en amarillo) y DIII (en azul) y el bucle de fusión (LF) en naranja. B) Ubicación estructural de E en la membrana de un virus, color marón, señalado con círculos negros las zonas buscadas en el diseño de las proteínas doble y teta mutantes, indicando cómo se busca sellar la interfaz dimérica de E, es decir interfaz entre protómeros de E y en la zona del LF. C) Arreglo de la proteína E en la superficie de los flavivirus maduros. Tres dímeros forman un zócalo y 30 zócalos recubren la partícula. D) Zoom de cada zócalo de flavivirus, mostrando los dominios y la ubicación del loop de fusión. E. Vista de costado para mostrar cómo se apoya un dímero y un monómero sobre la membrana viral.

La función principal de E es la de fusionar la membrana viral con la membrana diana para concretar la infección, dando inicio al ciclo de replicación.

### C. MECANISMO DE ENTRADA Y FUSIÓN VIRAL MEDIADA POR E

El virus Zika, que entra al torrente sanguíneo luego de la picadura de un mosquito infectado, se adosa de forma inespecífica en glico-receptores de la membrana celular y es luego internalizado para su eliminación por vesículas endosomales. Estas vesículas disminuyen el pH interno mediante canales de tipo bomba de protones, generando cierto grado de acidez que activa enzimas hidrolíticas para la degradación de biomoléculas y/o reciclaje de proteínas. Sin embargo, al bajar el pH en el entorno del virus, se gatilla la fusión entre la membrana del virus y la membrana endosomal, que está naturalmente inducida en el interior de los endosomas. ya que al generarse ese cambio en el entorno, a nivel del pH, E se disocia de dímero a monómero, exponiendo el bucle de fusión, quien se ancla en la membrana endosomal del futuro huésped. Luego, E pasa a un estado trimérico cerrado mediante un cambio conformacional que conlleva a la unión de ambas membranas, formando un poro de fusión.

Luego de la infección, se desarrolla el ciclo de replicación viral que culmina con la salida de partículas recientemente ensambladas, las cuales son sometidas a un proceso de maduración viral dentro y fuera de la célula huésped.

Un aspecto molecular clave de la maduración de las partículas ocurre en el bucle-150 de E, el cual funciona como una tapa que cierra un protómero con otro del dímero de E sellando dicha conformación a pH neutro, pero la cual se abre a pH ácido por repulsión electrostática, permitiendo la disociación y exposición del bucle de fusión de membranas, desencadenando la misma. También se descubrió que a pH ácido, la proteína viral Pr, sella el estado dimérico de E sobre del loop-150 en las nuevas partículas replicadas para evitar fusión de las mismas dentro de la células productoras. Pr se ubica en la zona que queda libre al estar el loop-150 en conformación levantada.

Es decir, los flavivirus utilizan a Pr, para evitar desencadenar la fusión antes de salir de la célula que los produjo. Un sistema vital, ya que en el camino de salida las partículas se exponen a pH ácido.

Estipulamos como hipótesis que es posible mantener la tapa o bucle-150 en conformación cerrada en E, aun a pH ácido, gracias a estrategias de ingeniería de proteínas y que con dicha estabilización del dímero de E, podremos comprender con mayor profundidad las interacciones que gobiernan el ensamblado de partículas virales en flavivirus. De hecho, no se sabe cómo se estabilizan los arreglos de E entre dímeros

y entre zócalos. Además, no se sabe cómo este bucle funciona en diferentes especies de flavivirus, en particular en el Zika.

Para estudiar esta hipótesis, proponemos en el presente proyecto final, la producción de proteínas recombinantes E diseñadas como mutantes selladoras de la tapa o bucle-150, una primera aproximación que establecerá un protocolo de purificación inédito en Argentina que permitirá abrir a futuro, estudios de biofísica estructural por criomicroscopía electrónica.

#### **D. DISEÑO RACIONAL Y EXPRESIÓN RECOMBINANTE DE VARIANTES ESTABILIZADAS DE LA PROTEÍNA E**

El presente trabajo empleará técnicas de ingeniería de proteínas, apoyadas por inteligencia artificial (Mirdita et al., 2022), para estabilizar el dímero de la proteína E y frenar los cambios estructurales asociados con la infección viral. A partir de estas proteínas diseñadas, se realizarán producciones recombinantes con el fin de purificarlas y establecer un sistema de producción de bajo costo. En un futuro posterior al presente trabajo de ingeniería de proteínas, se podría llevar a cabo la producción especializada de las mismas para ensayarlas mediante estudios estructurales y termodinámicos, tales como, técnicas basadas en fluorescencia, calorimetría y criomicroscopía electrónica, para confirmar el potencial aplicable de los constructos proteicos antigénicos de carácter hiperestabilizados. Es decir explotar el sistema con aplicaciones en biomedicina. Particularmente, se evaluará la efectividad de anticuerpos dirigidos a regiones específicas de la proteína E mutante, con la perspectiva de generar una vacuna proteica específica contra Zika.

El objetivo estrictamente académico busca comprender las interacciones que regulan el ensamblaje de partículas virales. Si bien existen otros laboratorios que abordan esta problemática con enfoques similares, no han considerado los últimos hallazgos en los mecanismos de maduración de los flavivirus (Crampon et al., 2023; Vaney et al., 2022; Rey et al. 2018). Sin embargo, sus estudios proporcionan una valiosa base metodológica en cuanto a los protocolos de purificación que se seguirán en esta investigación (Yang et al., 2019; Metz et al., 2019; Dai et al., 2016).

La hipótesis en cuestión nace a partir de un Screening de residuos a mutar a cisteínas en la plataforma MAESTRO (Laimer et al., 2015), desde la estructura cristalográfica de E dimérica del virus Zika y una posterior selección de los pares a mutar luego de evaluar la factibilidad de los resultados obtenidos, con el foco en encontrar las mutantes con mayor estabilidad, así una vez ya establecidos los candidatos, se mandarían a sintetizar a Gene Universal en plásmidos de expresión para sistemas heterólogos como *E. coli*.

El trabajo aquí propuesto incluye la clonación, expresión y purificación de la proteína de envoltura del virus Zika. Los estudios previos han demostrado que los anticuerpos contra los virus Zika y Dengue reconocen epítopes cuaternarios complejos presentes en los homodímeros de la proteína E. Por lo tanto, este proyecto se enfoca en estabilizar dicha conformación dimérica mediante mutagénesis sitio-dirigida, dado que las proteínas expresadas como monómeros no exponen estos epítopes clave.

Una vez que las proteínas sean exitosamente expresadas en cepas bacterianas, se escalará la producción en medios de cultivo del orden de litros y se procederá a la purificación mediante técnicas de replegado oxidativo en cuerpos de inclusión. Esto permitirá obtener suficiente masa de proteínas E mutantes del virus Zika, que serán sometidas a futuro a análisis biofísicos para validar su potencial biotecnológico.

## 5. ESTADO DEL ARTE

### A. ESTUDIOS ESTRUCTURALES DE LA PROTEÍNA E EN FLAVIVIRUS

La proteína E es el principal componente estructural de la envoltura de los flavivirus, incluyendo Zika, Dengue y Fiebre Amarilla. Se organiza en homodímeros que tapizan la superficie del virión maduro y cumple funciones esenciales en el reconocimiento celular, la entrada por endocitosis y la fusión de membranas. Estudios recientes de cristalografía y de criomicroscopía electrónica (Rey et al., 2018; Vaney et al., 2022; Crampon et al., 2023) revelaron las transiciones conformacionales que experimenta E desde su forma inmadura asociada a PrM hasta el estado estable y funcional en el virión maduro. Estos estudios permitieron identificar regiones de interés estructural, tales como el dominio III, implicado en la interacción con receptores, y zonas de glicosilación relevantes para la estabilidad del dímero y la posterior infección viral.

### B. LIMITACIONES DE LAS PROTEÍNAS RECOMBINANTES MONOMÉRICAS

La expresión recombinante de la proteína E en sistemas bacterianos como *E. coli* enfrenta múltiples desafíos. Debido a la ausencia de la capacidad de generar las modificaciones postraduccionales, como la glicosilación, por parte de las bacterias y al entorno reductor de sus citoplasmas, gracias al cual las proteínas tienden a acumularse en cuerpos de inclusión, plegarse de forma incorrecta o mantenerse como monómeros inestables. Esta conformación impide la correcta generación de los epítopes cuaternarios que son reconocidos por anticuerpos neutralizantes, reduciendo significativamente su utilidad como antígeno vacunal o en inmunodiagnóstico. Por este motivo, diversos grupos han propuesto estrategias para estabilizar artificialmente la forma dimérica de E a través de diseño racional y mutagénesis dirigida, seguida de protocolos específicos que aprovechan la fisiología procariota para la expresión y posterior recuperación y purificación de las proteínas en estudio.

### C. APLICACIONES BIOMÉDICAS DE PROTEÍNAS E ESTABILIZADAS

La estabilización de la conformación dimérica de la proteína E abre nuevas posibilidades en el desarrollo de vacunas subunitarias, kits de diagnóstico, aislamiento

de nuevos anticuerpos y estudios estructurales. La exposición adecuada de epítopes conformacionales ha demostrado ser clave para inducir respuestas inmunes protectoras, como evidencian estudios con proteínas E estabilizadas artificialmente (Yang et al., 2019; Dai et al., 2016), así como investigaciones recientes que subrayan la importancia de la estabilidad estructural en la inmunogenicidad de los flavivirus (Pierson & Diamond, 2017; Costello et al., 2024). Asimismo, la obtención de variantes solubles y funcionales permite ensayos de unión con anticuerpos monoclonales, diseño de vacunas basadas en dominios estructurales específicos y generación de inmunógenos para plataformas no convencionales como nanopartículas o vectores virales atenuados.

## 6. HIPÓTESIS

Se plantea que la introducción de mutaciones estabilizadoras en la interfaz dimérica del ectodominio de la proteína E del virus Zika incrementa su estabilidad conformacional y facilita la obtención de proteína soluble mediante producción recombinante en *E. coli*, al combinar un enfoque iterativo de diseño *in silico* utilizando MAESTRO y AlphaFold3 con un método de purificación basado en replegado oxidativo.

## 7. OBJETIVOS

### A. OBJETIVO GENERAL

En este proyecto, buscaremos lograr la producción recombinante de la proteína de envoltura, E, modificada por mutación racional sitio dirigida diseñada a partir de un análisis in-sílico de E en su estado nativo, con el fin de aumentar su estabilidad dimérica. Trabajaremos con E del virus Zika, expresaremos la proteína a partir de clones sintéticos utilizando el sistema de expresión bacteriano E Coli.

Este estudio busca establecer una plataforma de producción de variantes hiperestables de E, centrándose en el diseño in silico y en el establecimiento del protocolo de expresión y purificación de este sistema proteico complejo. El proyecto es de relevante aplicabilidad para el Zika y el resto de los flavivirus, como el Dengue y la Fiebre amarilla.

### B. OBJETIVOS ESPECÍFICOS

Diseñar variantes mutantes de la proteína E. Analizar la secuencia proteica con algoritmos de predicción de puentes disulfuro, validar los candidatos mediante predictores de estructura tridimensional y generar las secuencias para su síntesis y clonaje en plásmidos de expresión.

Preparar materiales para la expresión recombinante. Incluir la preparación de medios de cultivo líquidos y sólidos, la obtención de bacterias competentes por el método de cloruro de calcio, la transformación bacteriana con los plásmidos de interés (DH5 $\alpha$  para stock plasmídico, BL21(DE3) y C41 para expresión proteica mediante choque térmico a 42 °C) y el almacenamiento de clones transformados en stocks de glicerol al 25%, así como la obtención de material plasmídico por miniprep.

Expresar la proteína recombinante E. Realizar pruebas de expresión inducidas con IPTG en el sistema plasmídico pET bajo promotor T7, evaluar diferentes medios de cultivo ricos para comparar el rendimiento y analizar los resultados mediante electroforesis en geles de poliacrilamida (SDS-PAGE).

Purificar la proteína recombinante. Desarrollar un protocolo de replegado oxidativo a partir de cuerpos de inclusión adaptado a un sistema de purificación por afinidad vía etiqueta de histidinas.

## 8. MATERIALES Y MÉTODOS

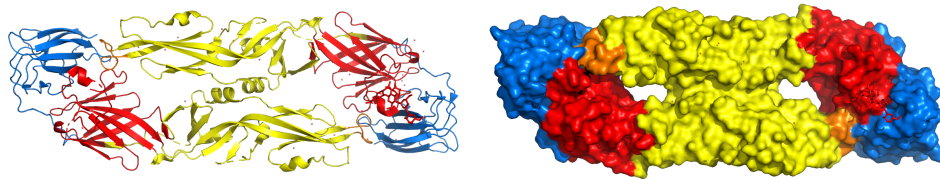
### A. DISEÑO IN SILICO Y SÍNTESIS DE LA VARIANTE DOBLE MUTANTE

La hipótesis propuesta se comienza a explorar mediante un análisis in silico, tomando como punto de partida la estructura cristalográfica de la proteína E del virus Zika. A partir de esta estructura, se buscó identificar variantes estructurales capaces de estabilizar la interfaz dimérica, zona crítica involucrada en la disociación de los protómeros durante el proceso de maduración viral. El objetivo fue diseñar mutaciones que aumentan la robustez del dímero, priorizando aquellas ubicadas en regiones flexibles y con potencial de formar puentes disulfuro entre las sub unidades del monómero.

Para ello, se utilizó el algoritmo **MAESTRO** (Laimer et al., 2015), el cual predice la estabilidad de mutaciones puntuales y permite explorar la introducción de residuos de **cisteína** con potencial de formar enlaces covalentes estabilizantes. A partir de la evaluación de múltiples combinaciones, se seleccionaron las variantes que cumplieran con los criterios estructurales deseados: formación de enlaces disulfuro en sitios clave sin afectar la forma plegada nativa de la proteína.

Las secuencias seleccionadas fueron modeladas mediante la herramienta **AlphaFold3** (Abramson et al., 2024), versión más reciente de esta plataforma basada en Machine Learning, que permite predecir estructuras 3D de proteínas completas a partir de su secuencia primaria, incluyendo interacciones oligoméricas y modificaciones postraduccionales.

Para validar la fidelidad del modelo, se generó una estructura de control correspondiente a la proteína E Wild Type en su forma dimérica, tanto con como sin N-glicosilación en el residuo Asn 346, y se comparó con la estructura cristalográfica original disponible en la base de datos RCSB PDB (PDB ID: 5LBV), revelando una coincidencia estructural notable. Esta comparación sirvió para validar la fidelidad del modelo generado con inteligencia artificial en condiciones nativas, antes de continuar con la predicción de las variantes mutantes (Fig. 2). El alineamiento estructural mostró una concordancia notable, lo que refuerza la confianza en los modelos generados para las variantes mutantes.



**Figura 2:** Estructura Cristalográfica de la proteína E dimérica nativa de Zika, (PDB 5LBV) en la que se basó el análisis llevado a cabo en la plataforma MAESTRO. Representación de tipo cartoon a la izquierda y de tipo superficie molecular a la derecha. Los colores se corresponden con los dominios I II II y el loop de fusión, en rojo amarillo azul y naranja respectivamente. Notar que la zona de interés se corresponde al loop del dominio I que se enfrenta al bucle de fusión entre los dos protómeros.

El pipeline empleado, que consistió en cargar la estructura dimérica en MAESTRO, validar mutantes candidatas, modelarlas con AlphaFold y corroborar mediante alineamiento tridimensional en PyMOL, permitió identificar dos construcciones de interés: una con **dos** mutaciones puntuales y otra con **cuatro** residuos modificados, ambas orientadas a **cerrar la interfase dimérica** mediante nuevos enlaces disulfuro y estabilizar las regiones flexibles de la proteína.

## B. VECTOR DE EXPRESIÓN: pET21A(+)

Para la expresión recombinante de las variantes diseñadas de la proteína E del virus Zika, se utilizó el plásmido pET21a(+), comprado a la empresa Gene Universal, con la secuencia insertada por un sistema de optimización de codones. Este vector está diseñado para su uso en sistemas procariotas, particularmente en *E. coli*, y cuenta con un promotor T7 que permite una expresión proteica altamente eficiente e inducible mediante IPTG, un análogo a la lactosa que reconoce al operador lacUV5, presente en el promotor, que al ser reconocido específicamente por la ARN polimerasa T7 presente en cepas como BL21(DE3) o C41(DE3) desencadena la transcripción del gen río abajo.

El plásmido incluye una secuencia de seis histidinas (His<sub>6</sub>-tag) en el extremo N-terminal, la cual permite la purificación de la proteína de interés por cromatografía de afinidad (Ni-NTA). Adicionalmente, el vector contiene un sitio de clonado múltiple (MCS) que facilita la inserción del gen de interés allí mismo, un sitio de origen de la replicación, y un sistema de selección basado en resistencia a la ampicilina, antibiótico que naturalmente es letal para las bacterias; el mismo va a agregarse en el medio de cultivo, asegurando que el crecimiento microbiano en los cultivos de trabajo esté

únicamente dado por las bacterias que incorporaron correctamente el plásmido. Estas construcciones fueron luego introducidas con fines de transformación en cepas de *E. coli* competentes para la posterior expresión de la proteína.

### **C. PREPARACIÓN Y USO DE MEDIOS DE CULTIVO LB Y 2YT PARA CRECIMIENTO BACTERIANO.**

Como etapa preliminar al inicio de los ensayos experimentales, se procedió a la preparación de los medios de cultivo y soluciones necesarias para el crecimiento y tratamiento bacteriano, además de para las posteriores etapas de transformación y expresión recombinante. En este sentido, se formularon de forma casera medios líquidos LB (Luria-Bertani) y 2YT (Yeast Extract-Tryptone), adecuados para el cultivo de cepas de *E. coli* en diferentes contextos. El medio LB fue utilizado principalmente para la selección de colonias transformadas y el mantenimiento de cultivos, mientras que el medio 2YT, de composición más rica, fue adoptado en etapas posteriores por haber demostrado un crecimiento bacteriano más robusto en las etapas de transformación y una mayor eficiencia en la expresión proteica. Ambos medios se prepararon en condiciones estériles, disolviendo los componentes secos en agua destilada y ajustando el volumen final a un litro. Para los medios sólidos, se añadió agar bacteriológico antes de la esterilización.

En el caso del LB (triptona 10 g/L, extracto de levadura 5 g/L y NaCl 10 g/L), se utilizó como alternativa rica el medio 2YT casero, formulado con triptona 16 g/L, extracto de levadura 10 g/L y NaCl 5 g/L. Cuando se necesitó medio sólido, se agregó 15 g/L de agar a la mezcla, y luego del autoclavado, se vertió el medio caliente en placas de Petri estériles. Las placas se dejaron solidificar a temperatura ambiente y se almacenaron a 4 °C.

Las cepas de *E. coli* utilizadas en este trabajo fueron donadas por el laboratorio Estructura-Función de Proteínas de la Fundación Instituto Leloir y almacenadas a -80 °C en glicerol. Posteriormente, fueron descongeladas y sembradas en los medios mencionados para su transformación, propagación y expresión recombinante.

La esterilización de los medios y demás insumos de laboratorio fue realizada mediante destrucción por calor húmedo, utilizando una autoclave (modelo VZ-500,

marca RAYPA, utilizada en las instalaciones del CIBION-CONICET). El equipo fue operado respetando estrictamente las indicaciones del fabricante y las normas de bioseguridad del laboratorio. Las soluciones fueron autoclavadas a 121 °C durante 20 minutos a 1 atm de presión. Se corroboró previamente el correcto nivel de agua hasta la rejilla metálica interior, y se utilizó agua de osmosis (OSMOSIS 1). La apertura del autoclave se realizó únicamente tras su completo enfriamiento, utilizando guantes térmicos de protección. Al finalizar el ciclo de esterilización, el equipo emitía una alarma sonora que requería el reinicio manual con la tecla “Reset” (R).

Esta combinación de medios y condiciones permitió obtener cultivos bacterianos con buena densidad óptica, fundamentales para el posterior escalado de producción recombinante.

#### **D. GENERACIÓN DE BACTERIAS COMPETENTES**

Con el objetivo de disponer de un stock de células bacterianas listas para procedimientos de transformación y expresión proteica, se generaron células competentes de *E. coli* empleando un protocolo clásico basado en tratamiento con cloruro de calcio ( $\text{CaCl}_2$ ), aplicado a las cepas DH5 $\alpha$ , C41 y BL21 (DE3). Estas fueron utilizadas con distintos fines dentro del proyecto: DH5 $\alpha$  para almacenamiento del stock plasmídico, y C41 y BL21 (DE3) para pruebas de expresión recombinante.

El procedimiento comenzó con la inoculación de precultivos en medio LB (2 mL) a partir de colonias individuales de cada cepa aisladas anteriormente, incubados durante la noche a 37 °C con agitación constante (250 rpm). Al día siguiente, se utilizó 1 mL de estos cultivos saturados para inocular 500 mL de LB fresco, incubando nuevamente hasta alcanzar una densidad óptica ( $\text{OD}_{600} \sim 0,5$ ), lo cual ocurrió tras aproximadamente entre las horas 3 y 5 de crecimiento en las mismas condiciones.

Las células fueron cosechadas por centrifugación a 5000 rpm durante 10 minutos a 4 °C en tubos estériles enfriados previamente, de 250 mL. El pellet celular resultante se resuspendió en 10 mL de solución fría de  $\text{CaCl}_2$ . Las suspensiones se recolectaron en tubos falcón estériles de 50 mL, y se sometieron a una segunda centrifugación a 2500 rpm durante 5 minutos. Nuevamente se descartó el sobrenadante y se repitió la resuspensión en 10 mL de  $\text{CaCl}_2$ , manteniéndose las muestras en hielo por 30 minutos.

Una tercera centrifugación a 2500 rpm permitió finalmente obtener pellets que fueron resuspendidos en 2 mL de CaCl<sub>2</sub> frío.

Las suspensiones finales se mantuvieron en hielo y, en algunos casos, fueron refrigeradas a 4 °C durante la noche con el fin de potenciar la competencia de las bacterias. Finalmente, las células fueron alicuotadas en volúmenes de 50 µL en tubos estériles preenfriados y almacenadas a –80 °C. Para asegurar su viabilidad a largo plazo, se conservaron stocks adicionales en nitrógeno líquido. Para ello, las alícuotas fueron colocadas en cajas de almacenamiento criogénico, previamente rotuladas, y transferidas con precaución al contenedor de nitrógeno líquido mediante un soporte, asegurando una inmersión rápida que evita la formación de cristales de hielo intracelular.

#### **E. TRANSFORMACIÓN DE BACTERIAS: DH5A, C41, BL21**

Para incorporar el plásmido recombinante en las distintas cepas de *E. coli*, se llevó a cabo una transformación mediante el protocolo clásico de shock térmico a 42°C, que combinado con el tratamiento con cloruro de calcio, facilita la entrada del ADN exógeno a través de poros formados en la membrana plasmática por el cambio de temperatura. Esta etapa del trabajo resultó ser especialmente crítica, ya que implicó varios intentos fallidos y una optimización intensiva del procedimiento para lograr transformaciones eficientes y reproducibles.

Inicialmente, las bacterias competentes (DH5α, C41 y BL21(DE3)), previamente generadas por tratamiento con CaCl<sub>2</sub> y conservadas a –80 °C, fueron descongeladas cuidadosamente sobre hielo. A cada alícuota se le añadió entre 10 y 50 ng de ADN plasmídico y se incubaron en frío durante 15 minutos para favorecer la interacción del ADN con la membrana celular. Luego se procedió con el paso clave correspondiente al shock térmico mediante un baño a 42 °C durante 1 minuto, seguido de una rápida incubación en hielo por otros 10 minutos más.

Acto seguido, se añadió medio LB sin antibiótico (800 µL) a cada tubo con las bacterias transformadas, y se incubaron en agitación a 37 °C durante 45 minutos para permitir la recuperación y expresión del gen de resistencia a antibiótico, lo que en adelante permitió trabajar solo con las células en las que la transformación fue eficiente.

Finalmente, se plaqueó el volumen total sobre placas LB con ampicilina (100 µg/mL) y se incubó toda la noche a 37 °C.

Durante esta instancia, se observó una marcada diferencia en la eficiencia de transformación entre las cepas: las DH5α, que respondieron consistentemente bien al procedimiento, mostrando crecimiento en todas las réplicas, lo que confirmó tanto su competencia como la integridad del plásmido. Sin embargo, tanto las cepas C41 como las BL21(DE3) presentaron deficiencias importantes en cuanto al crecimiento. Se registraron múltiples réplicas sin colonias viables, lo que llevó a cuestionar la competencia celular y la compatibilidad del plásmido con estos backgrounds genéticos.

Tras varios intentos y ajustes en las condiciones (tiempos de incubación, velocidad de agitación, medios de cultivo), finalmente se logró una transformación exitosa en la cepa C41, con una baja cantidad de colonias, pero suficiente para continuar con los ensayos de expresión proteica. Esta recuperación parcial fue clave para la continuación del proyecto, ya que permitió avanzar hacia la producción de las proteínas de interés.

#### **F. EXPRESIÓN RECOMBINANTE INDUCIDA CON IPTG Y ESCALADO DE CULTIVOS BACTERIANOS (2 mL, 500 mL, 1 L).**

Una vez confirmada la transformación exitosa de *E. coli* C41 con el plásmido de interés, se diseñó una estrategia por etapas para poner a punto las condiciones óptimas de expresión recombinante de la proteína. El objetivo fue asegurar niveles adecuados de producción proteica en condiciones que favorecieron su correcta conformación y solubilidad, para su posterior purificación.

La estrategia se dividió en tres etapas de escala creciente: cultivos piloto en tubos de ensayo (2mL), cultivos intermedios en matraz (500 mL) y finalmente cultivos a gran escala (1 L). En todos los casos, la expresión fue inducida con IPTG (isopropil-β-D-1-tiogalactopiranosido), un análogo no metabolizable de la lactosa que activa la transcripción del promotor T7 presente en el vector de expresión que incorporaron las bacterias utilizadas con sistema heterólogo en esta etapa una vez acumulado en el citoplasma bacteriano.

Durante las primeras pruebas de expresión en cultivos de 2 mL, se utilizaron tubos de ensayo con medio LB suplementado con el antibiótico correspondiente y se incubaron a 37 °C hasta alcanzar una densidad óptica de OD600 ~ 0.6. Sin embargo, bajo estas condiciones no se observó crecimiento bacteriano adecuado en las cepas transformadas, lo que si bien no impidió avanzar con la inducción, motivó la reformulación del medio de cultivo. Se optó entonces por utilizar medio 2YT (Yeast Extract-Tryptone), el cual posee una mayor concentración de nutrientes (triptona y extracto de levadura) que el LB, favoreciendo un crecimiento más vigoroso y estable de las bacterias.

La inducción con IPTG en los cultivos se llevó a cabo a una concentración final de 0.5 mM (0.8 µL en 2 mL), seguida de una incubación a 37 °C durante 16 horas, con agitación constante. Esta temperatura de expresión reducida se seleccionó con el objetivo de minimizar la formación de cuerpos de inclusión y favorecer el plegamiento controlado de la proteína recombinante.

Las muestras fueron posteriormente analizadas por SDS-PAGE para confirmar la expresión de la proteína, lo cual permitió validar la efectividad del protocolo en este volumen reducido. A partir de estos resultados positivos, se procedió a escalar el sistema.

Para la siguiente etapa, se realizó un cultivo en matraz de 500 mL en medio 2YT suplementado con antibiótico, inoculado con una colonia previamente ensayada. El cultivo se incubó a 37 °C con agitación a 225 rpm hasta alcanzar una OD600 ~ 0.6. La expresión fue inducida con 200 µL de IPTG (concentración final de 0.5 mM) y el cultivo fue incubado a 37 °C durante 16 horas. Posteriormente, se realizó una centrifugación a 5000 xg durante 10 minutos a 4 °C para recuperar el pellet bacteriano. El procedimiento fue idéntico para el cultivo a escala de 1 L, en el cual se utilizaron 400 µL de IPTG bajo las mismas condiciones de temperatura y agitación.

En todas las etapas se mantuvo la estrategia de control mediante SDS-PAGE para confirmar los niveles de expresión y el tamaño esperado de la proteína. Esta validación fue clave para avanzar con confianza hacia la etapa de purificación.

El sistema de expresión puesto a punto demostró ser robusto y reproducible, con una expresión detectable y específica de la proteína objetivo en condiciones de bajo

costo, utilizando *E. coli* C41 como plataforma bacteriana eficiente y medio 2YT como entorno nutritivo favorable.

#### **G. RECOLECCIÓN DE BIOMASA, LISIS CELULAR Y AISLAMIENTO DE CUERPOS DE INCLUSIÓN.**

Tras la confirmación de la expresión de “E” en los cultivos de 2 mL, 500 mL y 1 L, se procedió a la recuperación de la biomasa por centrifugación a 5000 xg durante 10 minutos a 4 °C. Los pellets bacterianos obtenidos a partir de estos cultivos inducidos fueron reservados a 4 °C para su posterior procesamiento.

La lisis celular se realizó por sonicación, un método físico altamente eficiente para romper células bacterianas mediante la aplicación de ondas ultrasónicas de alta frecuencia. Este procedimiento genera cavitación, un fenómeno que produce burbujas microscópicas en el medio líquido; su colapso provoca ondas de choque capaces de romper las membranas celulares y, por ende, de liberar el contenido intracelular. Para minimizar el calentamiento y evitar la desnaturalización de las proteínas, la sonicación se llevó a cabo en hielo, alternando entre dos sondas (de 35 y 40) con ciclos de 10 segundos de duración cada uno.

Las células se resuspendieron en un buffer de lisis compuesto por 100 mM Tris-HCl pH 6.8, 0.6 M NaCl, 1 mM EDTA y 10 mM 2-mercaptoetanol, en un volumen proporcional al pellet recuperado, manteniendo condiciones homogéneas entre muestras. La sonicación se realizó en el IBioBA (CONICET - Partner Institute of the Max Planck Society), utilizando un equipo de laboratorio específico para este fin. Una vez completado el proceso, las muestras fueron centrifugadas a 13.000 xg durante 20 minutos a 4 °C. De este modo, se separó la fracción soluble del contenido insoluble, donde se encontraban los **cuerpos de inclusión**, agregados densos de proteínas mal plegadas o parcialmente plegadas, que se acumulan en el citoplasma de bacterias durante la sobreexpresión recombinante, generalmente insolubles en condiciones fisiológicas y carentes de actividad biológica nativa. En esta oportunidad alojarán mayoritariamente la proteína recombinante de interés.

A continuación, los cuerpos de inclusión fueron aislados mediante una serie de lavados sucesivos con buffer conteniendo 50 mM Tris-HCl pH 8.0, 2 M urea y 1 %

Triton X-100, con una centrifugación intermedia tras cada paso. Un último lavado se realizó sin urea para reducir la concentración de desnaturalizantes. Finalmente, la fracción purificada fue solubilizada con una solución de 50 mM Tris-HCl pH 8.0, 8 M urea y 10 mM DTT, mediante incubación en agitación durante una hora a temperatura ambiente, o bien overnight a 4 °C, preparándose así para su posterior etapa de replegado oxidativo.

## H. REPLEGADO OXIDATIVO

Una vez aislados los cuerpos de inclusión tras la lisis celular por sonicación y posteriores lavados, se procedió a la recuperación de las proteínas recombinantes de interés mediante un proceso de solubilización en condiciones desnaturalizantes seguido de un replegado oxidativo in vitro, diseñado para restaurar la conformación nativa funcional de la proteína E recombinante del virus Zika, estabilizada mediante mutagénesis dirigida.

Los cuerpos de inclusión fueron resuspendidos inicialmente en 4 mL de buffer de solubilización compuesto por 8 M urea, 100 mM Tris-HCl pH 8,0, y 20 mM  $\beta$ -mercaptoetanol (BME, Sigma-Aldrich) como agente reductor. Esta combinación permite tanto la desnaturalización completa de la proteína como la ruptura de posibles puentes disulfuro incorrectos formados durante la expresión en *E. coli*. El uso de urea en esta concentración rompe eficazmente todas las interacciones no covalentes intra e intermoleculares, permitiendo una resuspensión homogénea de la proteína en su forma desnaturalizada. El BME, por su parte, actúa reduciendo los enlaces disulfuro incorrectos que pudieran haberse generado durante el colapso anómalo de las cadenas polipeptídicas en el entorno citoplásmico.

La mezcla fue incubada durante 30 minutos a temperatura ambiente con agitación orbital constante para favorecer la completa solubilización. Posteriormente, la solución fue centrifugada a 12.000 xg durante 15 minutos a 4°C, y el sobrenadante, que contenía la proteína desnaturalizada en estado reducido, fue cuidadosamente recuperado para la siguiente etapa.

Para inducir el replegado oxidativo, se utilizó un protocolo diseñado en dos fases: primero se disminuyó progresivamente la concentración de urea por dilución

controlada en un buffer de replegado que contiene condiciones redox definidas. El buffer fue preparado con 100 mM Tris-HCl pH 8,5, 0,4 M L-arginina (Sigma-Aldrich) como chaperona química para prevenir agregación de la proteína al replegarse, 2 mM EDTA para quelar iones metálicos que pudieran inducir oxidaciones no específicas, 2 mM GSH (glutación reducido) y 0,2 mM GSSG (glutación oxidado), generando un entorno redox controlado que favorece la correcta formación de enlaces disulfuro nativos. Esta relación GSH:GSSG (10:1) está basada en protocolos optimizados previamente para proteínas con puentes disulfuro intramoleculares e intermoleculares, como es el caso de los dímeros estabilizados por cisteínas introducidas artificialmente.

El volumen total final del sistema de replegado fue de 200 mL, en el cual se introdujeron los 4 mL de proteína desnaturalizada por goteo lento, manteniendo agitación constante a 4°C durante 16 horas. Esta técnica de “dilución gota a gota” permite que la proteína se incorpore progresivamente al entorno de replegado evitando precipitaciones abruptas y favoreciendo una cinética lenta de formación de la estructura terciaria y cuaternaria que fortalecería la generación de la conformación correcta.

Durante el replegado se mantuvo todo el sistema a 4°C, lo cual ayuda a minimizar la agregación no específica y favorece el correcto reordenamiento de las regiones hidrofóbicas internas ya que va a enlentecer el proceso, permitiendo que este se lleve a cabo de una forma más controlada. La presencia de L-arginina, inhibidor de agregación, favoreció el correcto ensamblado del dímero, permitiendo que las cisteínas introducidas en las interfaces inter diméricas formen los puentes disulfuro estables necesarios.

Tras el replegado, para la etapa final de concentración de la proteína, se realizó un proceso de ultrafiltración centrífuga utilizando dispositivos Amicon® Ultra-15 (Merck Millipore, Alemania) con un punto de corte molecular (MWCO) de 30 kDa, bajo centrifugación a  $4.000 \times g$  y 4 °C. Este sistema, basado en una membrana semipermeable de regeneración de celulosa, permite la retención selectiva de proteínas de alto peso molecular mientras elimina sales, urea y péptidos pequeños. El volumen total de 200 mL obtenido en la etapa anterior se concentró hasta alcanzar un volumen final de aproximadamente 2 mL, utilizado posteriormente para la cromatografía de

afinidad. Este paso no solo permitió remover el exceso de sales, urea residual y pequeños péptidos mal plegados, sino que también aumentó la concentración final de la proteína correctamente replegada.

La muestra concentrada fue almacenada transitoriamente a 4°C hasta su análisis por electroforesis en SDS-PAGE, y fue posteriormente purificada mediante cromatografía de afinidad al níquel gracias al His-tag N-terminal presente en la construcción recombinante.

### **I. PURIFICACIÓN POR AFINIDAD (HIS-TAG) Y DIÁLISIS.**

Una vez replegada y concentrada la proteína recombinante, se procedió a su purificación por cromatografía de afinidad utilizando la etiqueta de histidina (His6-tag) ubicada en el extremo N-terminal de la proteína fusión. Esta estrategia permite una purificación específica y de alta eficiencia basada en la capacidad de los residuos de histidina para coordinar cationes metálicos divalentes, como lo es el Ni<sup>2+</sup> o el Co<sup>2+</sup>, inmovilizados sobre una matriz sólida. La técnica empleada fue IMAC (Immobilized Metal Affinity Chromatography), utilizando resina Ni-NTA (nitrilotriacético de níquel) de la marca Qiagen®, un ligando tridentado que coordina fuertemente iones Ni<sup>2+</sup> a través de cuatro sitios de unión: tres grupos carboxilato del NTA y una posición libre que es ocupada por los residuos de histidina del His-tag. Esta interacción es reversible y se puede modular en función de la concentración de imidazol en la fase móvil, ya que este último compite directamente con las histidinas por el sitio de unión al metal.

Para garantizar condiciones óptimas y reproducibles, se prepararon y filtraron todos los buffers a través de membranas de 0,22 µm, y se ajustó el pH con NaOH 1 M o HCl 1 M según necesidad. El buffer de equilibrio y lavado (denominado Buffer A) contenía 50 mM de fosfato de sodio (NaH<sub>2</sub>PO<sub>4</sub>), 300 mM de NaCl y 20 mM de imidazol, con el pH ajustado a 8,0; se prepararon 20 mL de esta solución. El buffer de elución (Buffer B) mantenía la misma composición de fosfato y NaCl, pero contenía 250 mM de imidazol, también con pH 8,0, y se prepararon 10 mL. La inclusión de NaCl a alta concentración estabiliza la proteína en solución y reduce interacciones inespecíficas, mientras que el imidazol a baja concentración en el buffer de lavado permite remover proteínas bacterianas que pudieran haberse unido débilmente a la

matriz, pero no es lo suficientemente fuerte para remover a las proteínas unidas por su His-Tag.

La resina Ni-NTA (Qiagen®, 0.5 mL) fue cargada en una columna de polipropileno con filtro, previamente equilibrada con cinco volúmenes de Buffer A. La muestra concentrada en 2 mL con proteína replegada fue diluida 1:1 con Buffer A para reducir cualquier exceso de urea o sales residuales y aplicada directamente sobre la resina por gravedad, en condiciones de agitación leve para maximizar el contacto. La columna fue incubada durante 30 minutos a 4 °C en condiciones estáticas para favorecer la interacción entre la resina y la proteína His-tag. Posteriormente, se lavó la columna con diez volúmenes de Buffer A así se logró eliminar las proteínas no específicas endógenas de la bacteria que podrían haber estado presentes en la fracción de replegado.

La elución de la proteína de interés se realizó con cinco volúmenes de Buffer B, el cual contenía 250 mM de imidazol, una concentración suficientemente alta como para competir eficazmente con los residuos de histidina y desplazar la proteína unida. Las fracciones de elución se recolectaron en tubos estériles de 1,5 mL en los cuales se fue recolectando eluato de a 0.35 mL aproximadamente y se mantuvieron a 4 °C. Cada fracción de elución fue analizada por electroforesis SDS-PAGE al 12% para verificar la presencia, pureza y tamaño molecular esperado de la proteína recombinante. La fracción con mayor intensidad de banda fue seleccionada para ser utilizada en ensayos posteriores.

El uso de columnas de Ni<sup>2+</sup> en lugar de Co<sup>2+</sup> se justificó en función de su mayor capacidad de unión, siendo más apropiado para capturar proteínas con estructuras parcialmente replegadas o con exposición subóptima del His-tag. Esta elección se consideró adecuada para el sistema expresado en *E. coli*, donde las condiciones oxidativas intracelulares no garantizan una conformación nativa completa en la fracción insoluble, como es habitual en la generación de cuerpos de inclusión. Finalizada la etapa de purificación por afinidad mediante cromatografía con resina de níquel (Ni-NTA), la fracción eluida contenía la proteína recombinante de interés en un buffer con 250 mM de imidazol. Este compuesto, esencial durante la elución por su capacidad de competir con las cadenas laterales de histidina por los sitios de coordinación del níquel, debe ser posteriormente removido ya que puede interferir con

etapas posteriores como el replegado oxidativo, la cuantificación proteica y ensayos funcionales, además de comprometer la estabilidad de la proteína purificada a largo plazo.

Para eliminar el exceso de imidazol y permitir el intercambio de buffer, se procedió a realizar una diálisis por difusión pasiva utilizando una membrana de diálisis estándar Spectra/Por® RC Tubing (Spectrum Labs, Cat. 132645), con un MWCO de 6–8 kDa, ideal para mantener dentro de la misma a las proteínas de alto peso molecular y permitir la difusión libre del imidazol hacia el medio externo. Esta membrana se empleó cortando una longitud de 6.25 cm para dializar un volumen total de 2 mL de muestra, siguiendo la equivalencia provista por el fabricante de 0.32 mL/cm.

La configuración se preparó siguiendo las indicaciones del fabricante, mediante hidratación en agua milliQ y posterior enjuague con buffer de diálisis. La muestra fue cargada cuidadosamente en el interior de la membrana, sellada, y sumergida en un volumen 100 veces superior de buffer de diálisis (compuesto por 20 mM de Tris-HCl pH 8, 150 mM NaCl). La diálisis se llevó a cabo bajo agitación suave (70–90 rpm) durante 16 hs a 4 °C, con un recambio del buffer externo a las 2 horas de iniciado el proceso para asegurar un gradiente de concentración favorable para lograr el objetivo deseado.

El principio físico químico detrás de la diálisis radica en la difusión pasiva de solutos de bajo peso molecular (como el imidazol) a través de una membrana semipermeable, mientras que los solutos de mayor tamaño, como las proteínas, quedan retenidos dentro del compartimento sellado. Este proceso permite, luego del tiempo de equilibrio, obtener una muestra prácticamente libre de pequeños solutos indeseados y lista para ser utilizada en posteriores etapas experimentales como cuantificación, replegado o almacenamiento.

## **J. ANÁLISIS POR SDS-PAGE**

Para monitorear la eficiencia de la expresión recombinante y la pureza de la proteína de interés durante el proceso de producción y purificación de la misma, se empleó la técnica de electroforesis en gel de poliacrilamida en condiciones

desnaturalizantes (SDS-PAGE) a partir de los extractos de proteína anteriormente expresada. Esta técnica permite separar proteínas en función de su peso molecular aparente, ya que el tratamiento previo con SDS, un detergente aniónico, desnaturaliza las proteínas y recubre sus cadenas con carga negativa proporcional a su tamaño, anulando así las diferencias de forma o carga nativa que pueden generar ruido en las lecturas finales de las bandas obtenidas producto de la migración. En presencia de un agente reductor como DTT (incluido en el buffer de carga), los enlaces disulfuro se rompen, asegurando la desnaturalización completa de las proteínas. De este modo, al exponerlas ante un campo eléctrico, las proteínas migran a través de la matriz de acrilamida únicamente en función de su masa, permitiendo una separación precisa y una lectura fiel de los resultados.

La proteína recombinante de interés en este trabajo tiene un peso molecular estimado de aproximadamente 46 kDa. Por ello, se empleó un gel de separación (resolving) al 12% de bisacrilamida, adecuado para proteínas de tamaño dentro del rango de 20 a 100 kDa. La formulación final utilizada para preparar 10 ml de este gel incluyó: 3.2 ml de agua destilada, 4 ml de solución de bis-acrilamida al 30%, 2.6 ml de Tris-HCl 1.5 M pH 7.8, 100  $\mu$ l de SDS al 10%, 100  $\mu$ l de APS al 10% (iniciador de la polimerización) y 10  $\mu$ l de TEMED (catalizador de polimerización). El buffer Tris-HCl a pH 7.8 establece un entorno básico ideal para lograr una resolución óptima de las proteínas de tamaño medio a alto, ya que este pH favorece la ionización del SDS y una migración eficiente en la red polimérica.

Por encima del gel de separación se colocó un gel concentrador (stacking) al 4%, que tiene como función comprimir todas las proteínas en una banda fina al ingresar al gel, mejorando así la resolución. El stacking se preparó con 2.975 ml de agua destilada, 0.67 ml de bis-acrilamida al 30%, 1.25 ml de Tris-HCl 1.5 M pH 6.8, 50  $\mu$ l de SDS al 10%, 50  $\mu$ l de APS al 10% y 5  $\mu$ l de TEMED. El pH ácido del stacking (pH 6.8) crea un sistema de migración diferencial en el cual las proteínas se concentran entre el frente de iones de cloruro (de menor movilidad) y los iones de glicina (de mayor movilidad).

La corrida electroforética se llevó a cabo en un sistema vertical, utilizando como buffer de corrida una solución compuesta por Tris 25 mM, glicina 192 mM y SDS 0.1% (pH ~8.3). Inicialmente se aplicó una tensión de 90 V para favorecer el ingreso

uniforme de las muestras en el stacking, y luego se incrementó a 120–150 V durante la fase de resolución en el gel de separación. Las muestras se mezclaron previamente en una relación 1:1 con buffer de carga (conteniendo SDS,  $\beta$ -mercaptoetanol, glicerol y azul de bromofenol) y se calentaron a 95 °C durante 5 minutos para asegurar su completa desnaturalización.

Una vez finalizada la corrida, los geles fueron teñidos con azul de Coomassie brillante G-250, una tinción de alta sensibilidad que interactúa con residuos básicos como arginina y grupos aromáticos de las proteínas. La decoloración se realizó en solución de metanol y ácido acético para remover el exceso de colorante sin afectar las bandas proteicas. Las muestras evaluadas abarcan desde cultivos inducidos a distintas escalas hasta fracciones purificadas, pasando por lisados celulares, sobrenadantes, pellets y fracciones de elución cromatográfica, permitiendo un monitoreo integral del proceso. El patrón de bandas observado permitió no solo confirmar la expresión de la proteína esperada, sino también evaluar su grado de pureza, presencia de cuerpos de inclusión, y eficiencia de los métodos de purificación empleados.

## 9. RESULTADOS

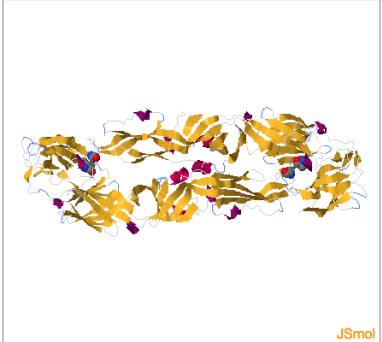
### A. DISEÑO RACIONAL Y VALIDACIÓN COMPUTACIONAL DE VARIANTES DE LA PROTEÍNA E MEDIANTE PREDICCIÓN DE PUENTES DISULFURO Y MODELADO ESTRUCTURAL

Como paso inicial del diseño racional de variantes estabilizadas de la proteína E del virus Zika para llevar a cabo la investigación, se empleó la herramienta MAESTRO (Laimer et al., 2015) para llevar a cabo un screening de la región de la proteína de interés en donde estaba comprendido el loop en estudio, clave para la fusión de membranas en el contexto de la infección viral. Esta plataforma computacional evalúa todos los pares de residuos que se encuentran a una distancia menor a 5 Å entre sus átomos C $\beta$ , simulando su conversión en cisteínas, y calcula para cada combinación un puntaje estructural ( $S_{ss}$ ) y una predicción de cambio en energía libre de estabilización ( $\Delta\Delta G_{pred}$ ). El valor  $S_{ss}$  integra parámetros geométricos y espaciales, representando cuán favorable es la orientación relativa de los residuos para la formación de un puente disulfuro estable, mientras que la  $\Delta\Delta G_{pred}$  estima el costo o beneficio energético que implicaría introducir dicha unión covalente en la estructura. Tal como se describe en la validación original de la herramienta (Laimer et al., 2015), la combinación de ambas métricas pondera la distancia efectiva, el ángulo y la torsión entre residuos para calcular el costo energético global de la formación del puente disulfuro. Los scores obtenidos expuestos en la Figura 3 solo fueron usados como parámetro para los residuos presentes en la zona específica que contenía al loop, dentro de la cual se eligieron los candidatos que con su combinación de valores, mostraron una geometría favorable para llevar a cabo la mutagénesis dirigida y potencial termo estabilizante mayor al incorporar las variaciones en su conformación. Estas predicciones bioinformáticas respaldaron la selección de los residuos específicos para llevar a cabo las mutaciones correspondientes.

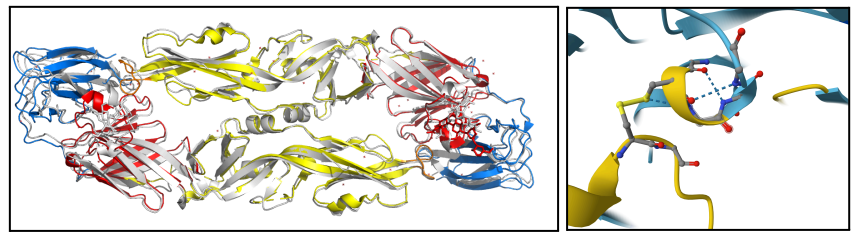
Potential disulfide bonds evaluation (5LVB, biological assembly 1, pH=8)

For this prediction task all residue pairs with a C<sub>β</sub>-C<sub>β</sub> distance closer than 5Å are considered as potential binding partners. The bond score (S<sub>ss</sub>) combines the predicted  $\Delta\Delta G$  value and disulfide bond geometry penalties. The results are sorted by S<sub>ss</sub> from best (most negative score) to worst.

Disulfide bond	$\Delta\Delta G_{pred}$	C <sub>pred</sub>	S <sub>ss</sub>
U270.I.AB(S>);L265.AB(S>...	-0.144	0.742	-2.050
G5.AB(SS),G100.AB(SS)	-0.249	0.766	-2.040
E26.AB(SS),G29.AB(SS)	0.104	0.771	-2.022
G259.AB(SS),E262.AB(SS)	-0.317	0.756	-1.992
R2.AB(SS),S142.AB(SS)	0.301	0.754	-1.966
S146.AB(SS),T366.AB(SS)	0.087	0.750	-1.956
A80.AB(SS),S112.AB(SS)	0.193	0.768	-1.850
M53.AB(SS),K128.AB(SS)	0.723	0.775	-1.815
G102.AB(SS),I152.AB(SS)	0.155	0.754	-1.801
H148.AB(SS),K373.AB(SS)	0.245	0.746	-1.784
L96.AB(SS),K110.AB(SS)	0.298	0.754	-1.755
G182.AB(SS),K301.AB(SS)	-0.078	0.722	-1.711
T115.AB(SS),Q253.AB(SS)	0.545	0.775	-1.688



$\Delta\Delta G_{pred}$  ... total predicted change in stability (kcal/mol).  $\Delta\Delta G_{pred} < 0.0$  indicates a stabilizing mutation. C<sub>pred</sub> ... confidence estimation, given as value between 0.0 (not reliable) and 1.0 (highly reliable). S<sub>ss</sub> ... bond score. For more information see [here](#).



**Figura 3. Plan de trabajo para el diseño racional de mutantes de Zika.** Panel superior: captura de pantalla del output de la aplicación web MAESTRO para la proteína E del virus Zika (PDB 5LVB). La tabla (izquierda) muestra los pares de residuos candidatos a formar puentes disulfuro, junto con el puntaje S<sub>ss</sub>, que integra el cambio de energía libre predicho ( $\Delta\Delta G_{pred}$ ), la confiabilidad estimada (C<sub>pred</sub>) y las penalizaciones geométricas. La imagen (derecha) permite evaluar visualmente la disposición tridimensional de los enlaces disulfuro potenciales. En amarillo se resalta el par correspondiente a la doble mutante G102C/I152C. Panel inferior izquierda: alineamiento estructural entre la proteína E cristalográfica de Zika (PDB 5LVB), coloreada según los dominios I, II y III (rojo, amarillo y azul, respectivamente), y la estructura generada con AlphaFold3, en gris. Panel inferior derecha: inspección estructural en AlphaFold3 para validar la formación de puentes disulfuro interprotoméricos de acuerdo a mutantes propuestas por MAESTRO. Se muestran dos cisteínas mutantes uniendo covalentemente el bucle de fusión (alfa hélice arriba) con el bucle 150 (bucle de abajo).

A partir de las secuencias resultantes del análisis en MAESTRO, se procedió a modelado estructural tridimensional de las variantes seleccionadas mediante la herramienta AlphaFoldv3. En esta etapa, se emplearon alineamientos múltiples de secuencias (MSA) generados automáticamente por MMseqs2, con el objetivo de reducir los tiempos de trabajo computacional manteniendo la precisión en los resultados.

Previo a modelar las variantes mutadas, se desarrolló un análisis de control en el que se predijo la estructura tridimensional de la proteína E del Zika Wild-Type (WT) tanto con, cómo sin, modificaciones postraduccionales (PTMs). En particular, se incorporaron las glicosilaciones típicas descritas para esta proteína, como la

N-glicosilación en la Asparagina 154. Resultando en la formación del dímero gran exactitud, en comparación con la estructura cristalizada, de acuerdo a la inspección visual y a una desviación media cuadrática (RMSD) de 2.47 Å frente al alineamiento de todos los átomos entre ambas estructuras, realizado en PyMOL (Schrödinger, LLC), Figura 3 panel de abajo.

Luego, en la Figura 4 y la Figura 5, exploramos el efecto de modificaciones postraduccionales en E. Recordamos que E es una glicoproteína, la cual se expresa en el retículo endoplasmático y es destinada al exterior de la célula. Para avanzar con una predicción que contemple modificaciones postraduccionales, una opción novedosa en la versión 2 de Alphafold, realizamos una comparación entre las predicciones con y sin modificaciones postraduccionales. Encontramos, no solo que Alphafold predice siempre dímeros bien plegados, con todos sus puentes disulfuro nativos, si no que las predicciones no mostraron diferencias entre las estructuras con y sin modificaciones postraduccionales, con RMSD por debajo de los 2 Å, no mostrado, en los alineamientos entre las diferentes estructuras. Es decir que tanto el plegamiento de los dominios estructurales ni en la superficie antigénica dimérica parece perturbar la presencia de azúcares. Evidencias de otras proteínas E homólogas, como la de fiebre amarilla, sin glicosilaciones, o de algunas variantes de dengue con dos glicosilaciones, habla en este sentido, evolutivamente, los azúcares no parecen en los flavivirus como determinantes significativos desde el punto de vista estructural, sí probablemente desde el punto de vista del tropismo, o el tipo de vector transmisor. Por otra parte, los azúcares son altamente flexibles e hidratables, en línea con una función poco estructural. Este resultado permitió suponer que, al menos desde una perspectiva estructural *in silico*, la ausencia de glicosilaciones en sistemas recombinantes no eucariontes, no comprometería la conformación de la proteína, hipotetizando que la expresión en sistemas bacterianos tiene sentido y sería favorable como estrategia experimental de expresión de E.

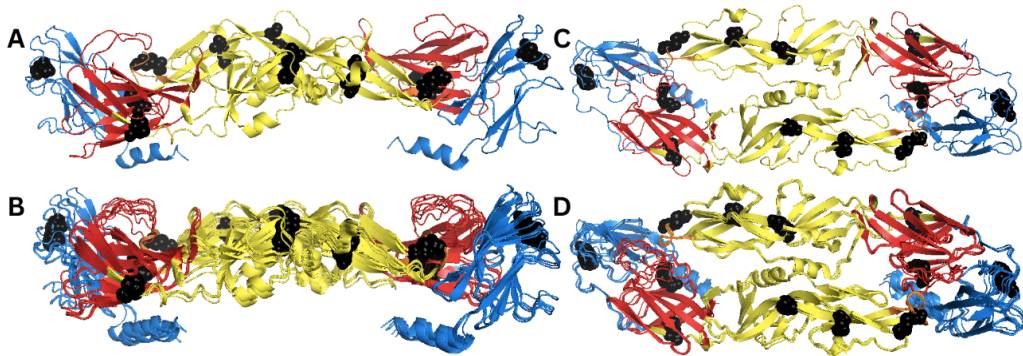


Fig. 4: Proteína E Wild Type A) Uno de los modelos generados visto de perfil B) El alineamiento de los 5 modelos generados visto de perfil C) Uno de los modelos generados visto desde arriba D) El alineamiento de los 5 modelos generados visto desde arriba. Las cadenas laterales de las cisteínas formando puentes disulfuro naturales de E, son representados con esferas negras, siendo 6 por monómero y 12 en total.

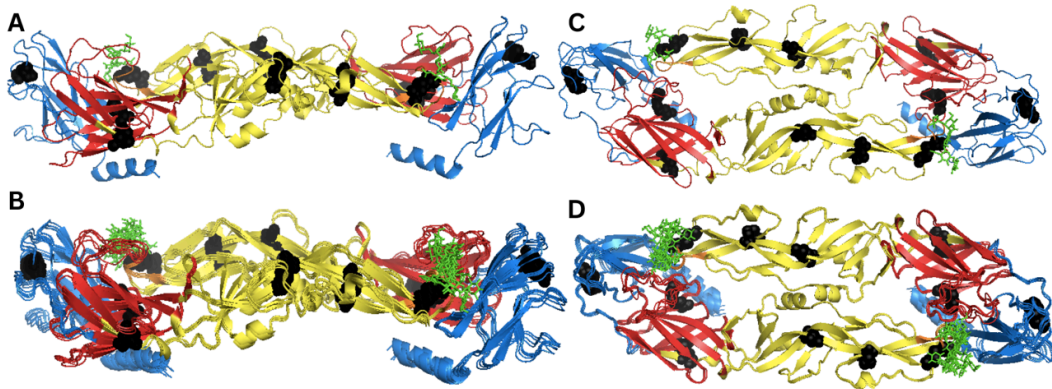


Fig. 5: Proteína E Wild Type con modificación post-traduccional A) Uno de los modelos generados visto de perfil B) El alineamiento de los 5 modelos generados visto de perfil C) Uno de los modelos generados visto desde arriba D) El alineamiento de los 5 modelos generados visto desde arriba. Las cadenas laterales de las cisteínas formando puentes disulfuro naturales de E, son representados con esferas negras, siendo 6 por monómero y 12 en total. En verde se colorean los azúcares covalentemente adosados a la asparagina 154.

Con estos datos reforzando el abordaje in silico predictivo, se generaron modelos independientes para distintas variantes formadoras de puentes disulfuro (doble y tetra mutante) y se evaluaron visualmente y estructuralmente para verificar el correcto ensamblaje de la interfase dimérica. Este sistema nos permitió identificar nuestros candidatos mutantes con gran potencial de éxito, siendo el criterio elegido que en todos los modelos predichos por AlphaFold, los puentes disulfuro se formen entre los residuos mutados, sin forzar la geometría ni aplicar modificaciones estructurales. MAESTRO, propone una tabla con más de 300 candidatos, y hemos observado que en varios casos AlphaFold no ha podido formar el puente disulfuro en ninguno de los modelos o solo en algunos, o solo en un lado del dímero y no del otro (No mostrado). El sistema

MAESTRO-AlphaFold nos permite iterar con los candidatos de interés a modo de definir los que mejor se comportan.

En las Figuras 6 y 7 mostramos 5 predicciones de AlphaFold y su correspondiente alineamiento, con los candidatos que logramos una buena predicción. Las estructuras resultantes fueron exportadas en formato .cif y analizadas mediante PyMOL (Schrödinger, LLC), herramienta de visualización molecular que permitió corroborar visualmente la formación de los puentes disulfuro deseados, evidenciados como enlaces covalentes entre residuos cisteína, representados con esferas en color rosa (Figura 6 y Figura 7). Además, se comparó la superposición tridimensional entre el modelo WT y los modelos mutados para verificar que las modificaciones no generan distorsiones globales, concluyendo que las mutaciones seleccionadas no alteraron significativamente el marco estructural general del dímero E. por lo tanto la simulación tridimensional de AlphaFold valida la predicción de MAESTRO y conserva las premisas buscadas para mejorar el carácter de estabilidad sin modificar el carácter antigénico.

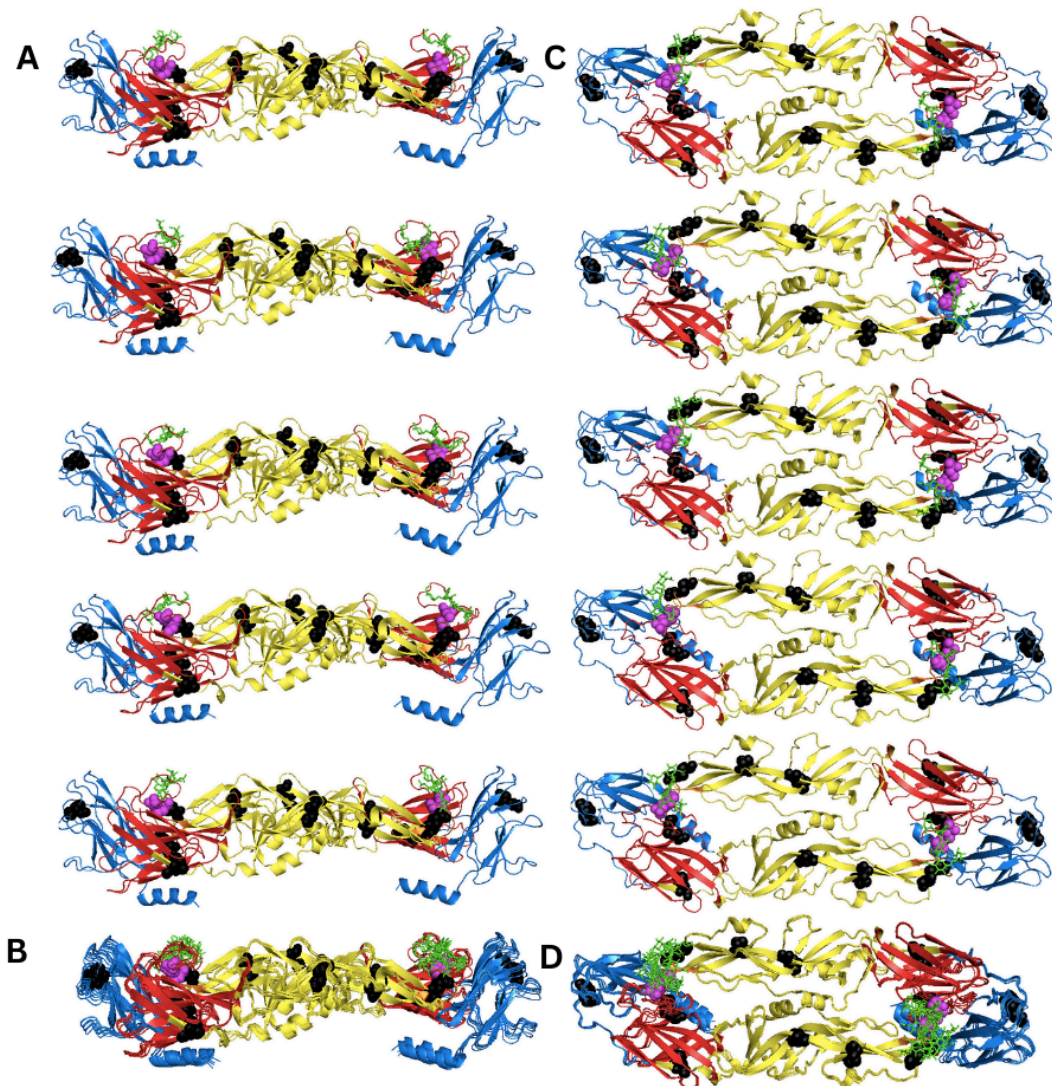


Fig. 6: Proteína E Doble Mutante A) Los 5 modelos generados con los puentes disulfuro nativos (esferas negras) y los puentes disulfuro dados por la mutaciones (esferas rosas) vistos de perfil B) El alineamiento de los 5 modelos generados vistos de perfil C) Los 5 modelos generados con los puentes disulfuro nativos (esferas negras) y los puentes disulfuro dados por la mutaciones (esferas rosas) vistos desde arriba B) El alineamiento de los 5 modelos generados vistos desde arriba. En verde se colorean los azúcares covalentemente adosados a la asparagina 154.

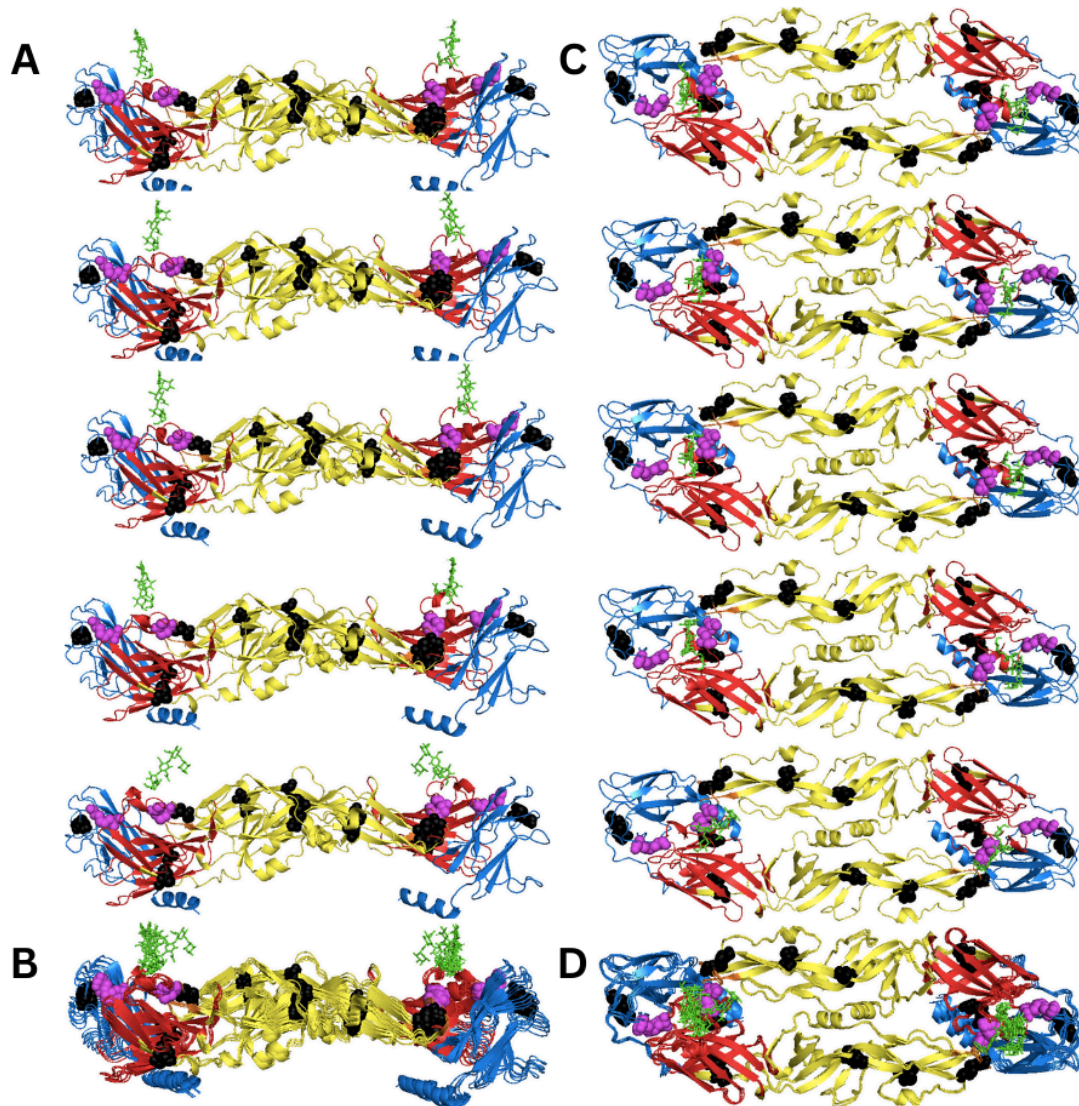


Fig. 7: Proteína E Tetra Mutada A) Los 5 modelos generados con los puentes disulfuro nativos (esferas negras) y los puentes disulfuro dados por la mutaciones (esferas rosas) vistos de perfil B) El alineamiento de los 5 modelos generados vistos de perfil C) Los 5 modelos generados con los puentes disulfuro nativos (esferas negras) y los puentes disulfuro dados por la mutaciones (esferas rosas) vistos desde arriba B) El alineamiento de los 5 modelos generados vistos desde arriba. En verde se colorean los azúcares covalentemente adosados a la asparagina 154.

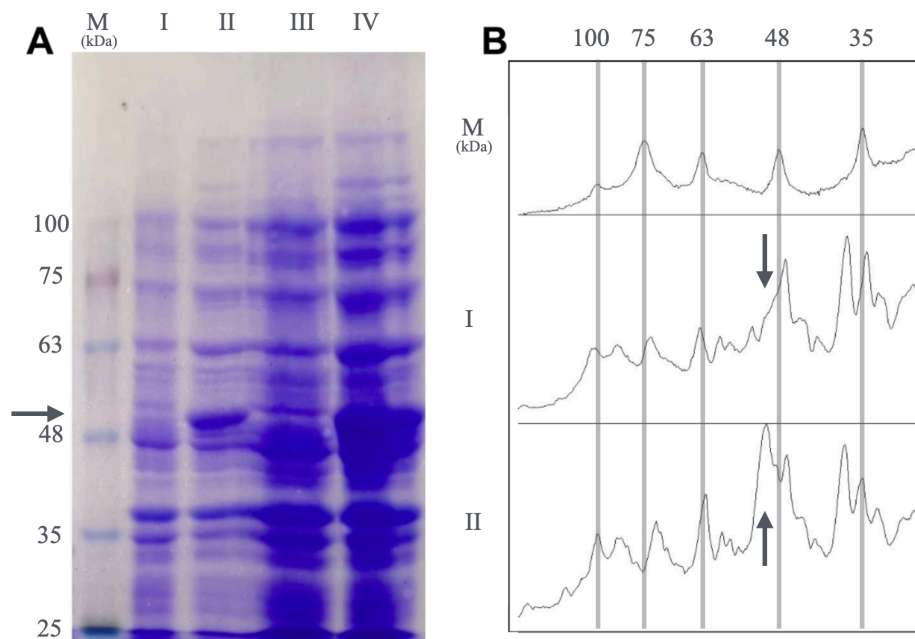
Cabe aclarar que, aunque el análisis *in silico* se llevó a cabo para ambas variantes diseñadas (doble y tetra mutante), únicamente la doble mutante fue utilizada en las etapas experimentales posteriores del proyecto, en concordancia con los recursos disponibles. Con esta base se procedió a sintetizar comercialmente los primeros clones, haciendo uso de optimización de codones. Luego, se utilizarán protocolos de clonación y biología molecular, principalmente, el método de clonación libre de enzimas de restricción, una técnica económicamente ventajosa ya que evita el uso de enzimas de restricción para la producción *in-vitro* de las proteínas en cuestión, para así luego seguir

con los objetivos propuestos en la hipótesis del trabajo y validar empíricamente el método propuesto de ingeniería genética con fines estabilizadores y biotecnológicos.

**B. VALIDACIÓN DE EXPRESIÓN RECOMBINANTE MEDIANTE SDS PAGE:**

**PRUEBAS DE EXPRESIÓN**

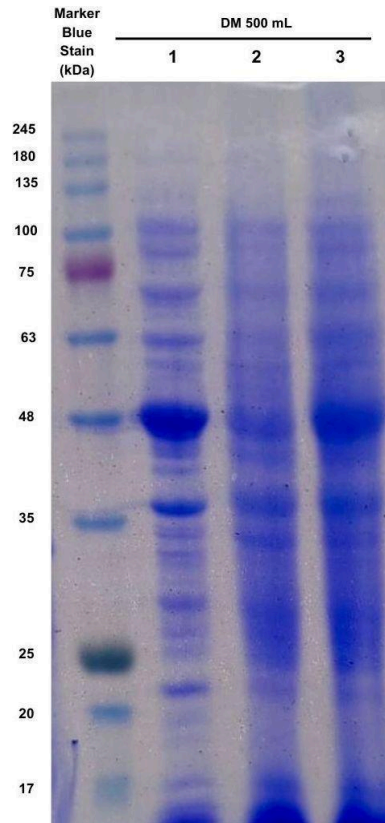
Para evaluar la expresión, solubilidad e integridad de la proteína E del virus Zika (peso molecular teórico ~46 kDa) a lo largo de todo el flujo experimental, se realizaron ensayos de SDS-PAGE bajo condiciones desnaturalizantes. El análisis incluyó pruebas de expresión llevadas a cabo en cultivos a pequeña escala (2 mL) utilizando medio LB, escalados a volúmenes intermedios (500 mL) y grandes (1 L) optimizados con medio 2YT, así como las fracciones resultantes de la lisis celular y la purificación por cromatografía de afinidad.



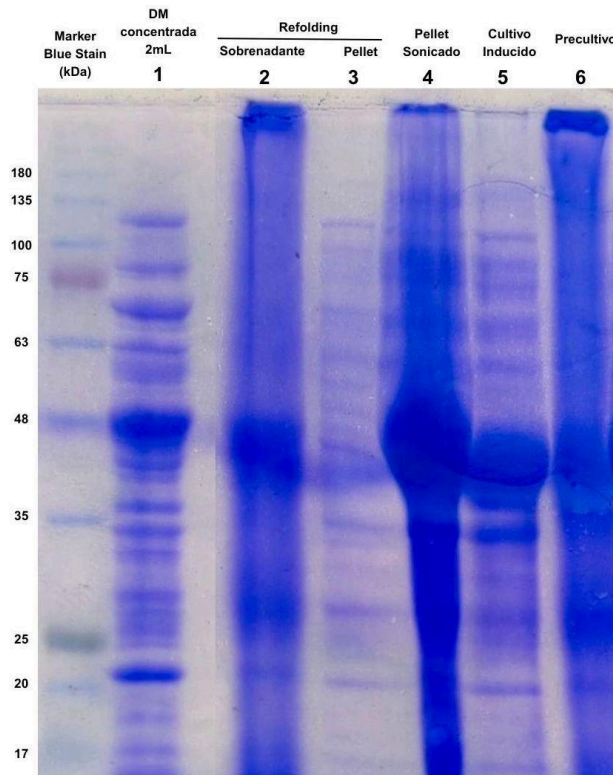
**Figura 8.** Análisis de la prueba de expresión a pequeña escala (2 mL) mediante SDS-PAGE 12 % luego de tratar con Coomassie Blue (0.2 g/100mL) y densitometría generada en ImageJ. (A) Gel representativo mostrando las condiciones de inducción (+IPTG) (calle II y IV) y condiciones control sin inducción (-IPTG)(calles I y III), con un marcador de peso molecular indicado en kDa. (B) Densitometría obtenida con ImageJ que comprende el perfil de picos del marcador de peso molecular Blue Stain (Anexo Figura 2.) y las calle I no inducida y II inducida, evidenciando un pico intenso correspondiente a la proteína recombinante en las condiciones inducidas y su ausencia en los controles sin tratamiento con IPTG. Las flechas marcan la posición de la sobreexpresión proteica.

En la primera prueba de expresión a escala de 2 mL (Figura 8A) se compararon cultivos inducidos con IPTG frente a controles sin inducir. En las muestras inducidas (carriles II y IV) apareció una banda intensa en torno a 46 kDa, ausente en los controles (carriles I y III), lo que indica una sobreexpresión específica de la proteína recombinante en condiciones de inducción. Para reforzar este resultado se generaron perfiles de densitometría a partir de los carriles I y II del gel utilizando la herramienta *Plot Profile* de ImageJ. El perfil del marcador (Figura 8B) se empleó como referencia para ubicar la posición esperada de la proteína de 46 kDa, y los perfiles de las muestras mostraron un pico adicional en las inducidas (+IPTG) ausente en las no inducidas (-IPTG), confirmando la identidad de la banda observada. Fuera de esa diferencia, el resto de los patrones de proteínas resultaron altamente coincidentes. Cabe señalar que la proteína de aproximadamente 46 kDa se visualiza en el SDS-PAGE ligeramente por encima de la marca de 48 kDa, fenómeno atribuible a migración anómala en este tipo de geles, a la presencia de secuencias adicionales del vector o a un plegamiento incompleto asociado a su acumulación en cuerpos de inclusión. Estas variaciones de 1–2 kDa son frecuentes y no implican necesariamente una modificación postraduccional.

Posteriormente, se escaló el cultivo de 2 mL a 500 mL para evaluar si la expresión de la proteína se mantenía en mayor volumen mediante SDS-PAGE (Figura 9). En los tres cultivos inducidos (+IPTG; carriles 1, 2 y 3) se observó un patrón de expresión proteica similar al de la escala pequeña y, nuevamente, una banda definida en la posición correspondiente a ~46 kDa, tal como se había visto en los cultivos de 2 mL (Figura 8). Esto demuestra que la producción de la proteína es reproducible a gran escala sin pérdida de rendimiento. En este caso, la banda de sobreexpresión se alineó mejor con el peso molecular esperado, lo que sugiere que el gel anterior pudo haber presentado una migración anómala o que la proteína mantiene interacciones no específicas que afectan su movilidad en condiciones de SDS. Es importante considerar que la carga total de proteína en los carriles también influye en la visualización: en particular, los carriles 2 y 3 que muestran un fondo más uniforme, probablemente debido a una mayor carga proteica que saturó la tinción y redujo la resolución de las bandas individuales.



**Figura 9.** Análisis de la prueba de expresión a mediana escala (500 mL) mediante SDS-PAGE 12 % luego de tratar con Coomassie Blue (0.2 g/100mL). Gel representativo mostrando muestras desde cultivos en condiciones de inducción (+IPTG) (calle 1, 2 y 3), con un marcador de peso molecular indicado en kDa.



**Figura 10.** Perfil SDS-PAGE (12 %) de la proteína “E” doble mutante (~46 kDa) a lo largo del flujo de trabajo: cultivo y lisis, aislamiento de cuerpos de inclusión, refolding y concentración final previo a la purificación por cromatografía de afinidad. Calle 1: proteína concentrada post-refolding (2 mL); Calle 2: sobrenadante soluble tras el refolding; Calle 3: pellet insoluble post-refolding; Calle 4: pellet posterior a la lisis por sonicación (cuerpos de inclusión); Calle 5: cultivo inducido (1 L); Calle 6: pre cultivo no inducido. Marcador en kDa a la izquierda.

Mediante el gel SDS-PAGE (12 %) presentado en la Figura 10 se evidencia la caracterización correspondiente a las fracciones recogidas a lo largo del proceso de aislamiento de cuerpos de inclusión, la solubilización en condiciones desnaturalizantes, el replegado oxidativo y la concentración final de la proteína E doble mutante del virus Zika (~46 kDa). En primer lugar se tomó una muestra luego del proceso de sonicación, método de lisis que al generar ondas de presión y cavitación en la suspensión celular, rompe la membrana y pared bacteriana de forma mecánica, liberando de manera inespecífica todos los componentes intracelulares de la misma. Debido a que los cuerpos de inclusión poseen mayor densidad, se sedimentan durante la centrifugación y permanecen en el pellet, este se evidencia en la Calle 4 de la figura 10, en donde se distingue una banda intensa en torno a los 46 kDa con un patrón de bandas inespecíficas de fondo, correspondientes a las proteínas bacterianas endógenas. El resultado es consistente con la acumulación de la proteína recombinante en cuerpos de inclusión, lo cual es esperable al utilizar un sistema de expresión bacteriano, ya que

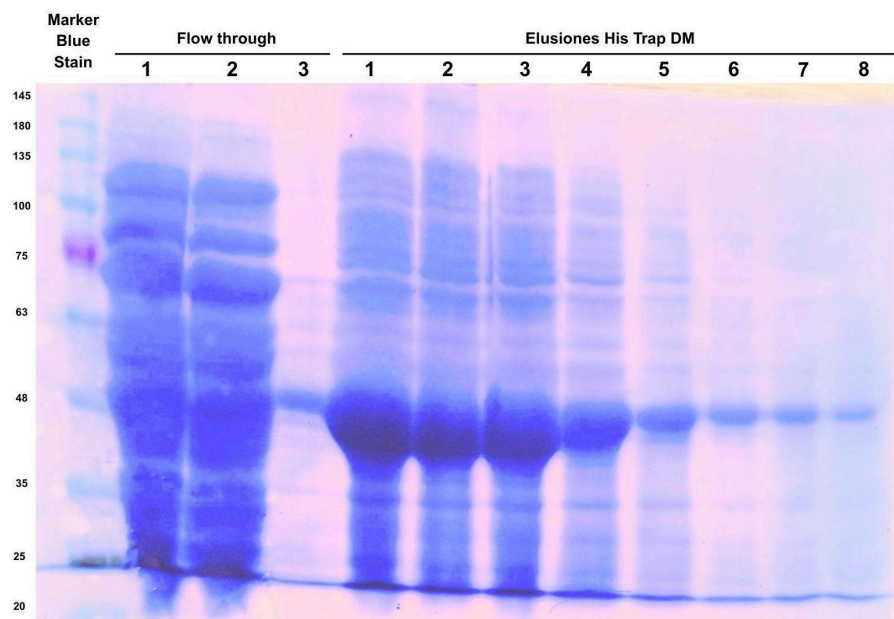
posee bajas capacidades de re-plegamiento de proteínas, por saturación de las chaperonas citosólicas y exposición de residuos hidrofóbicos, así promoviendo la agregación de las proteínas. Posteriormente, este pellet fue solubilizado empleando agentes caotrópicos reductores (urea y  $\beta$ -mercaptoetanol) que desestabiliza las interacciones no covalentes y reduce los enlaces disulfuro mal formados, permitiendo que la proteína adopte un estado completamente desplegado. A continuación, el replegado oxidativo se llevó a cabo mediante dilución controlada en un buffer con arginina y un par redox GSH/GSSG, que actúa modulando el potencial redox de la solución para favorecer la formación ordenada de los puentes disulfuro nativos (en este caso, los generados por las mutaciones G102C/I152C) para mantener la funcionalidad original de la proteína. Tras este paso, se separó la fracción soluble replegada (Calle 2) de los agregados residuales (Calle 3) mediante una centrifugación. En la calle 2 se observa claramente la banda de 46 kDa, con mínima presencia de contaminantes, lo que indica que la proteína objetivo permaneció soluble y correctamente plegada en las condiciones establecidas para el proceso de refolding. Por lo contrario, en la calle 3 prácticamente no se detecta señal a 46 kDa, lo que confirma que la mayor parte de la proteína no precipitó tras el replegado, validando la eficacia de esta etapa.

El concentrado final tras la etapa de ultrafiltración en centrífuga (Calle 1), reducido a 2 mL a partir del volumen post-refolding, exhibe una banda intensa a 46 kDa, con cierto fondo inespecífico atribuible a proteínas bacterianas minoritarias o restos de agentes de refolding que luego serán removidos mediante cromatografía de afinidad. La concentración mediante dispositivos de membrana con corte de peso molecular (MWCO) superior a la masa de la proteína permitió eliminar componentes de bajo peso molecular, aunque parte de las sales y compuestos del buffer pueden contribuir a otorgar una turbidez suave a la fracción.

En cuanto a los controles de expresión, la Calle 5 (cultivo escalado a 1 L inducido (+IPTG) ) presenta la banda de 46 kDa esperada, inmersa en un patrón proteico heterogéneo, característico de un extracto celular total sin clarificación previa ni fraccionamiento. Por su parte, la Calle 6 (pre cultivo no inducido (- IPTG)) carece de la banda específica de la proteína E, pero exhibe zonas de tinción uniforme con una densidad elevada, atribuibles a sobrecarga de biomasa, alta viscosidad por material

genómico no degradado y posible precipitación de proteínas abundantes de *E. coli* que migran de forma anómala.

En conjunto, el análisis secuencial de las fracciones confirma que la proteína recombinante se acumuló inicialmente en cuerpos de inclusión, luego fue solubilizada eficazmente, replegada en condiciones oxidantes que preservaron su solubilidad y su conformación nativa, y concentrada con mínima pérdida antes de la purificación por cromatografía de afinidad.



**Figura 11.** Gel SDS-PAGE con los resultados de la purificación de E Doble Mutante producida en *E. coli*, luego de tratar con Coomassie Blue (0,2 g/100mL). Calles 1, 2 y 3: Flow-Through, Calles 1-8 : fracciones de elución de la cromatografía con 500 mM de imidazol que presentan una banda de aproximadamente 46 kDa coincidente con el tamaño de E Doble Mutante. Marcador de Proteínas [kDa].

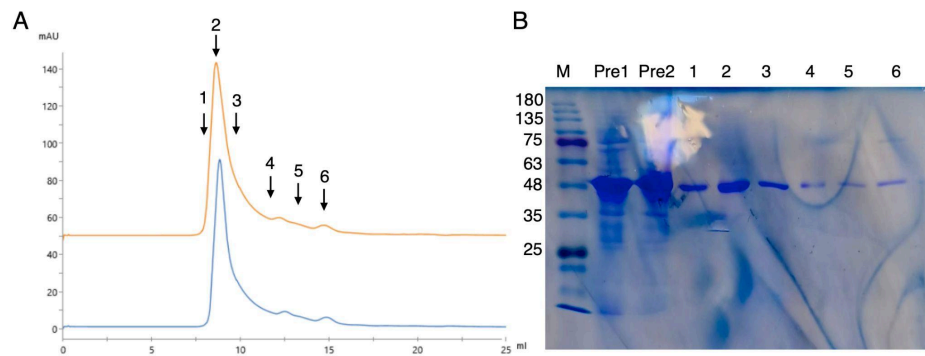
Finalmente, en la purificación por cromatografía de afinidad (His-tag) (Figura 11) los resultados obtenidos indican que en el cultivo de 1L de *E. coli* C41 con la inducción de 1 mM de IPTG durante 16 hs, se mantuvo la correcta expresión de E Doble Mutante que se venía observando a pequeña (2mL) (Figura 8) y a mediana escala (500 mL) (Figura 9). El análisis mediante SDS-PAGE muestra que, al aplicar el lisado celular a la columna de afinidad sin imidazol, la mayoría de las proteínas endógenas son correctamente eliminadas en el flujo no retenido (Figura 11 calles 1, 2 y 3 correspondientes al Flow-Through), en donde no se detecta la banda de interés. En cambio, la proteína recombinante E Doble Mutante fue retenida eficazmente y luego eludida con una solución de 500 mM de imidazol, como se observa en la Figura 11: calles 1-8 correspondientes a las fracciones de elución. Bajo condiciones reductoras

con  $\beta$ -mercaptoetanol, se detecta una banda nítida alrededor de los 46 kDa en el extracto crudo, compatible con el peso molecular estimado de la proteína target y una disminución notable de los contaminantes visibles a medida que avanzan las fracciones de elución. Si bien no se realizó una cuantificación precisa del grado de pureza, la presencia de una única banda definida, sin arrastre visible de proteínas contaminantes en las fracciones de elución, indica que las fracciones seleccionadas sugieren un alto grado de pureza alcanzado en un solo paso de purificación, aunque también se distingue una baja progresiva de la intensidad de la banda de interés, desde las fracciones 5 hasta la 8, lo que se traduce en una caída de rendimiento a lo largo del gradiente de imidazol, lo que es un patrón característico del proceso de elución competitiva con Imidazol, donde la mayor parte de la proteína se libera al alcanzar la concentración óptima de desplazamiento de la resina, lo que sucede en las primeras fracciones (fracciones de la 1 a la 5) ya que en las últimas solo hay cantidades residuales de proteína (Porath et al., 1975)

Al comparar los perfiles de proteínas observados en los geles de SDS-PAGE, se evidencia que el sistema procarionota permitió una expresión eficiente de la proteína que fue fácilmente detectable en el extracto crudo y se recuperó con alta pureza tras la purificación por IMAC.

### C. EVALUACIÓN DE PUREZA Y RENDIMIENTO

Para terminar el proceso de purificación y poder determinar el grado de pureza de la proteína E doble mutante expresada en *E. coli* C41, se llevó a cabo una columna de exclusión molecular (SEC, Superdex 200), a partir de las fracciones de elución del proceso de afinidad Ni<sup>2+</sup>-NTA, Figura 12.



**Figura 12.** Perfil Cromatográfico SEC y Gel SDS-PAGE correspondiente de la proteína E Doble Mutante de Zika. A: perfil de elución monitoreando a 280 nm (mAU) en el AktaPure. Se muestran dos inyecciones, color azul y naranja. B: Gel SDS-PAGE de las fracciones señaladas en A con una flecha.

Se llevó a cabo una columna de exclusión molecular observando un pico de elución principal a los 9 ml (*void volume* de la columna 8 ml). El SDS-PAGE al 12 % bajo condiciones desnaturalizantes y reductoras ( $\beta$ -mercaptoetanol), seguida de tinción con Coomassie Brilliant Blue G-250 de las fracciones correspondientes muestra el grado de pureza. El perfil electroforético de las fracciones de elución reveló una única banda con a la altura de  $\sim 46$  kDa en contraste con el ladder de proteínas, coincidente con el peso molecular teórico de la proteína objetivo. En total se observaron 3 picos, el mayor y dos menores, mostrando un grado heterogéneo del estado oligomérico. La ausencia de bandas adicionales detectables por tinción en las fracciones seleccionadas, sumado a la eliminación completa de la proteína de interés en el flow-through, sugiere un alto grado de pureza aparente ( $>90$  % estimado por inspección visual), alcanzado en un único paso cromatográfico. Se integró el área de los perfiles 280 nm de elución de la SEC y a partir del coeficiente de extinción molecular de E, se pudo cuantificar el rendimiento de la producción en 2 mg por litro de cultivo. Cabe notar que la comparación de perfiles entre las distintas escalas de cultivo (2 mL, 500 mL y 1 L) evidenció que las condiciones optimizadas (medio 2YT, inducción con 0,5 mM IPTG a 37 °C durante 16 hs) mantuvieron la eficiencia de expresión y purificación sin pérdida detectable de pureza ni variaciones significativas en el patrón de recuperación, validando la reproducibilidad del sistema a diferentes volúmenes de producción.

## 10. DISCUSIÓN

El enfoque integrado que combina predicción bioinformática (MAESTRO), modelado estructural (AlphaFold) y experimentación recombinante en *E. coli* ha permitido abordar la hipótesis central de este trabajo: diseñar variantes de la proteína E del virus Zika con puentes disulfuro inter monoméricos que estabilicen la interfase dimerica impidiendo así la fusión de la membrana viral con la célula huésped. Desde la etapa *in silico*, la selección de pares a mutar a cisteína mediante MAESTRO aportó criterios geométricos y termodinámicos (distancias C $\beta$ -C $\beta$ , ángulos y  $\Delta\Delta G$  predicho) que fortalecieron y validaron la decisión experimental. El score S<sub>SS</sub> y la  $\Delta\Delta G$ <sub>pred</sub> proporcionan dos dimensiones complementarias: la primera una evaluación la viabilidad geométrica del enlace disulfuro (ángulo  $\chi$ , distancia y torsión), y la segunda estima el impacto energético sobre la estabilidad local/ global. Los valores obtenidos para G102C/I152C (doble mutante) y para los pares claves de la tetra mutante (por ejemplo G102C/V153C) situaron ambas alternativas dentro de zonas favorables según criterios publicados para ingeniería de proteínas que se enfocan en las variaciones estructurales dadas por los puentes disulfuros (Laimer et al., 2015). Esta selección previa redujo el espacio de diseño y minimizó la probabilidad de introducir mutaciones con costos estructurales adicionales que comprometieran la estabilidad y funcionalidad de la proteína expresada.

Los modelos de AlphaFold generados con y sin PTMs mostraron una conservación de la conformación global del dímero, lo que sugiere que los elementos de plegamiento basales de E están fuertemente codificados en la secuencia aminoacídica y en el acoplamiento coevolutivo utilizado por el algoritmo. Esta observación apoya la hipótesis práctica de que la expresión en *E. coli*, organismo carente de capacidades de generar glicosilaciones complejas, puede producir dominios estructuralmente comparables a los modelos *in silico*, al menos a nivel de la conformación global, aunque con la obvia salvedad de que la presencia/ausencia de PTMs puede modular la dinámica superficial de la proteína, la presentación de epítopes y la estabilidad en condiciones fisiológicas de la proteína (Crampon et al.,

2023; Vaney et al., 2022; Rey et al., 2018). La coincidencia entre la predicción de MAESTRO y la aparición espontánea de puentes disulfuro en todos los modelos de AlphaFold para ambas variantes refuerza la plausibilidad estructural de las mutaciones propuestas.

Desde el punto de vista de la expresión y optimización en *E. coli*. Los resultados experimentales confirman varios principios bien conocidos pero relevantes para la práctica, como la dificultad inicial en la transformación en ciertas cepas (BL21(DE3) vs C41) y la variabilidad en la competencia celular, esto subraya la importancia de optimizar tanto el huésped como las condiciones de cultivo. La elección final de C41 y la transición de LB a 2YT como medio marcó una mejora clara en crecimiento y productividad de los pre cultivos bacterianos que posteriormente se utilizaron para stock plasmídico y expresión proteica. El medio 2YT ,más enriquecido en triptona y extracto de levadura, proporciona mayor disponibilidad de precursores aminoacídicos y factores de crecimiento que favorecen la producción de proteínas heterólogas y, en este caso, permitieron alcanzar la DO600 deseada, en las etapas de inducción, de manera reproducible. La inducción con IPTG a 0.5 mM y la incubación prolongada (~16 hs) a baja temperatura (37 °C) son condiciones claves para favorecer el plegamiento y reducir la formación de cuerpos de inclusión; sin embargo, en este trabajo la proteína continuó acumulándose predominantemente de forma insoluble, lo cual es consistente con la naturaleza de muchas glicoproteínas virales fragmentadas o fragmentos expresados sin el entorno con mecanismos secretorios aportado por las células de naturaleza eucariota.

La formación de cuerpos de inclusión observada (pellet tras sonicación con clara banda a ~46 kDa) y la subsecuente solubilización con 8 M urea + agente reductor ( $\beta$ -mercaptoetanol) siguieron la secuencia de pasos esperada: una desnaturalización completa para retornar a un estado desplegado y, luego, replegarse en un entorno oxidativo controlado. El protocolo de replegado empleado ,dilución gota-a-gota en buffer con L-arginina y en presencia del par redox GSH/GSSG, busca minimizar la agregación y favorecer la formación correcta de puentes disulfuro. La presencia de la banda a 46 kDa en el sobrenadante tras finalizar la etapa de replegado y su ausencia (o debilidad) en el pellet post-refolding indican que fue efectivo en una fracción apreciable de la proteína, y que los agregados que no se resolvieron se descartaron en

el pellet, tal como se esperaba. Técnicamente, la L-arginina actúa como agente anti-agregación al interactuar de forma débil con superficies hidrofóbicas expuestas, mientras que la relación GSH/GSSG define el potencial redox que condiciona la formación en contracara al reordenamiento de los puentes disulfuro; la elección de la relación 10:1 (por ejemplo) busca favorecer reacciones de 'shuffling' que permitan alcanzar el patrón nativo de enlaces.

En la purificación por IMAC, el patrón observado, ausencia de la banda objetivo en el flow-through y recuperación predominante en las primeras fracciones de elución con 250 mM imidazol de la proteína en estudio, es congruente con una captura efectiva y una elución competitiva típica (Porath et al., 1975). La caída de intensidad en fracciones tardías se debe a que la mayor parte del material con afinidad robusta queda desalojado en los primeros volúmenes de elución; lo que no indica pérdida en términos de pureza, sino el agotamiento de la carga estable útil de la resina. El uso de imidazol en el buffer de lavado (~20 mM) fue suficiente para reducir uniones débiles y mejorar la selectividad del proceso.

Respecto a la evaluación de pureza y rendimiento, el uso de SDS-PAGE combinado con densitometría (ImageJ) es válido como método semi-cuantitativo para estimar la fracción relativa de proteína objetivo en cada muestra.

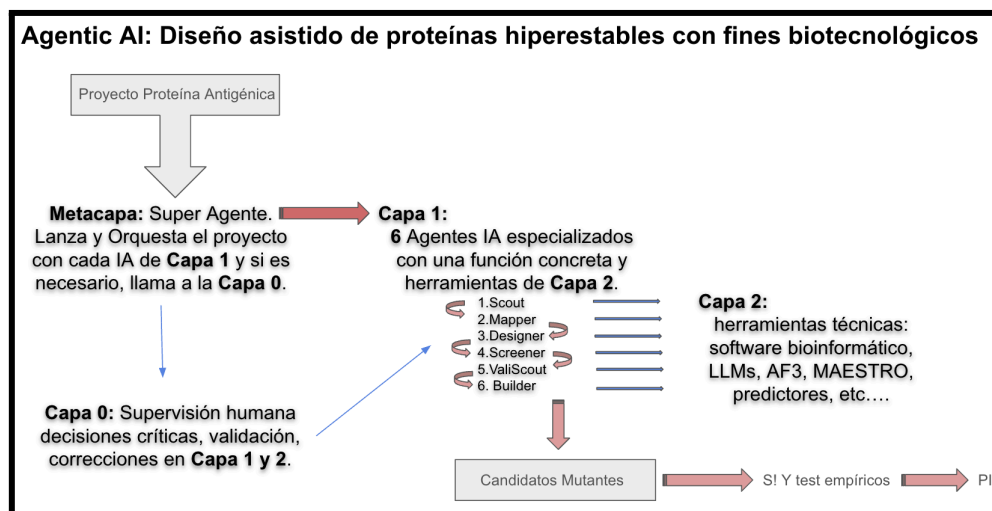
Las limitaciones del estudio incluyen, en primer lugar, el alcance experimental limitado a la doble mutante: aunque tanto la doble como la tetra mutante fueron modeladas in silico y sintetizadas, sólo la doble fue evaluada en esta ocasión. En segundo lugar, la ausencia de la posibilidad de generación de modificaciones postraduccionales en la plataforma bacteriana impone una distancia interpretativa respecto a la antigenicidad nativa de la proteína, que puede diferir de la observada en el virus completo. Esto se debe a que la falta de glicosilaciones podría alterar el reconocimiento por anticuerpos cuyos ligandos interaccionan con dichos azúcares, aunque no necesariamente afecte la funcionalidad estructural intrínseca de la proteína estabilizada.

Adicionalmente, la caracterización estructural se apoya en modelos computacionales y análisis biofísicos básicos (SDS-PAGE, perfiles densitométricos); para una validación final sería necesario incorporar técnicas de mayor resolución como espectrometría de masas, DSC para estabilidad térmica y preferentemente

criomicroscopía o cristalografía para una validación de la estructura cristalina definitiva.

Finalmente, las observaciones prácticas sugieren vías de mejora inmediata: coexpresión con chaperonas (p. ej. GroEL/ES, DnaK/DnaJ) o el uso de cepas con ambiente citosólico oxidante (BL21 Origami, BL21 Shuffle), lo que podría facilitar la formación de puentes disulfuro in-vivo y reducir cuerpos de inclusión; alternativas como la expresión en periplasma o bajar la temperatura en la etapa de replegado para lograr un proceso más controlado podría aumentar la solubilidad, así mismo en el replegado, se puede optimizar la relación redox y la concentración de arginina o bien ensayar gradientes de urea y on-column refolding para aumentar el rendimiento de la proteína soluble.

Por último, a partir del esquema de la Figura 12, el desarrollo planteado en este trabajo podría integrarse como un flujo de trabajo basado en protocolos experimentales clásicos de ingeniería de proteínas, que incluyen tanto las etapas de diseño in silico como las etapas experimentales de expresión y purificación, dentro de un marco de agentes basados en inteligencia artificial como oportunidad de automatización del proceso y su consecuente optimización.



**Figura 12.** Esquema conceptual de un sistema basado en agentes de inteligencia artificial aplicado al diseño asistido de variantes estructurales de proteínas hiper-estables para aplicaciones biotecnológicas. La *Meta capa* actúa como super agente coordinando el workflow planteado, activando a los agentes especializados que comprenden la *Capa 1*, que cumplen funciones específicas apoyándose en las herramientas bioinformáticas contenidas en la *Capa 2*. La *Capa 0* corresponde a la supervisión humana crítica para la toma de decisiones y la validación de los pasos que se están siguiendo. Por último el sistema va a generar proteínas mutantes candidatas que van a avanzar hacia las pruebas de expresión y posteriores pasos de purificación, etapas que asimismo pueden ser automatizadas mediante plataformas robóticas de cultivo celular y columnas cromatográficas.

A la cabeza del sistema, aparece la Meta Capa como “super agente”, que es el punto de partida que va a dar dirección a todo el proyecto, coordinando a los agentes especializados contenidos en la Capa 1, activandolos o no, según necesidad y lo hará comunicándose con la Capa 0, en donde subyace la supervisión humana con el criterio necesario para validar los caminos que va tomando el sistema, lo que asegura que las decisiones automatizadas se vayan validando.

La Capa 1 comprende seis agentes de IA que tienen tareas concretas: desde el Scout, encargado de hacer una investigación bibliográfica y de repositorios de secuencias, hasta el Builder, responsable de seleccionar a las mutantes candidatas. Estos agentes se apoyan en las herramientas bioinformáticas de la Capa 2, donde residen los programas técnicos que se utilizaron para el proyecto, tales como MAESTRO, AlphaFold3, predictores de estabilidad, optimizadores de codones o incluso LLMs que permiten interpretar y generar hipótesis a partir de grandes volúmenes de información.

El output del sistema son efectivamente las mutantes candidatas que se priorizaron y potencialmente podrían ser ensayadas directamente mediante las pruebas de expresión experimentales para posteriormente purificarlas. En este punto se abre la posibilidad de automatizar tanto la fase de expresión y purificación inicial mediante plataformas robóticas de cultivo bacteriano como los sistemas de cromatografía para mayor grado de purificación en paralelo, como lo demuestran Schreiber et al. (2017), quienes desarrollaron una plataforma de alto rendimiento para optimizar expresión de péptidos antimicrobianos en diferentes sistemas, y Stracquadiano et al. (2023), que propusieron un esquema de bajo costo y accesible para llevar a cabo el screening rápido de variantes usando robótica de manejo de líquidos.

En conjunto, estos avances y la figura ilustra cómo el trabajo aquí desarrollado, basado en la predicción de mutantes más estables, su modelado estructural, la validación por expresión y purificación, podría escalarse hacia un pipeline semi-automatizado gobernado por IA, donde los pasos in silico y experimentales se retroalimentan continuamente. Esta transición refleja una tendencia creciente en biotecnología: utilizar IA no solo como herramienta de análisis, sino también como agente coordinador autónomo de proyectos de ingeniería de proteínas con aplicaciones biomédicas y biotecnológicas.

## 11. CONCLUSIONES

En este trabajo se lograron integrar herramientas bioinformáticas de predicción estructural de proteínas y aproximaciones experimentales de biología molecular para el diseño, expresión, replegado y purificación de una variante modificada de la proteína E del virus Zika, introduciendo puentes disulfuro inter monoméricos con el fin de evaluar su viabilidad estructural y su producción en sistemas bacterianos.

La validación de la selección de las mutaciones se basó en un análisis exhaustivo de las posibles candidatas en MAESTRO, que evaluó todas las posibles combinaciones de residuos a mutar a cisteína dentro de un radio de interacción compatible con la formación de puentes disulfuro ( $<5 \text{ \AA}$  entre  $C\beta$ ), ponderando criterios geométricos (ángulo de torsión  $\chi$ , alineación de orbitales de azufre) y termodinámicos ( $\Delta\Delta G_{\text{pred}}$ ). La doble mutante G102C/I152C mostró un  $S_{SS}$  negativo indicativo de geometría favorable y una  $\Delta\Delta G_{\text{pred}}$  cercana a cero, sugiriendo que la modificación no impondría un coste energético significativo en el plegamiento global. La coherencia entre esta predicción y el modelado estructural posterior con AlphaFold, mostró la formación espontánea del puente disulfuro en todas las predicciones generadas, reforzando la plausibilidad estructural del diseño.

A nivel experimental, se optimizaron factores críticos para la expresión recombinante exitosa en *E. coli* como plataforma de expresión heteróloga. La elección de la cepa C41(DE3) y del medio 2YT permitieron un crecimiento robusto y una expresión detectable tras inducción con IPTG a 0,5 mM a 37 °C por 16 hs. No obstante, la proteína recombinante se acumuló mayoritariamente en forma insoluble, conformando cuerpos de inclusión. Esta localización es consistente con la expresión bacteriana de proteínas virales con dominios de membrana o pliegues complejos que, en ausencia de un entorno secretor eucariota y de glicosilaciones específicas, tienden a agregarse en el citoplasma; por lo que parece útil valerse de un protocolo de purificación proteica desde los mismos agregados formados.

La recuperación de proteína soluble requirió la implementación de un protocolo de solubilización con urea 8 M y agente reductor, seguido de un replegado oxidativo

en condiciones controladas. El uso de L-arginina como agente antiagregante y un sistema redox GSH/GSSG optimizado permitió que una fracción significativa de la proteína adoptará un estado soluble compatible con su posterior purificación. La ausencia de proteína en el pellet post-refolding y su presencia predominante en el sobrenadante confirmaron la eficacia de este proceso.

La etapa de purificación por cromatografía de afinidad (IMAC con resina Ni-NTA) bajo condiciones desnaturizantes y elución con imidazol 250 mM produjo una fracción final de alta pureza aparente. El perfil observado, retención efectiva, mínima pérdida en flow-through, y recuperación concentrada en fracciones iniciales de elución, es congruente con el mecanismo competitivo del imidazol desplazando la interacción coordinada entre el grupo imidazol de la histidina y el ion Níquel presente en la matriz de la columna.

La pureza se corroboró por SDS-PAGE, donde la banda de 46 kDa correspondiente a la variante de E doble mutante se observó como única señal predominante en la fracción eluida. El análisis densitométrico de las bandas, aunque semi-cuantitativo, confirmó un enriquecimiento significativo respecto a las fracciones previas. No obstante, la determinación de rendimiento absoluto requeriría cuantificación por métodos adicionales (Bradford, BCA,  $Abs_{280}$ ), si bien luego de pasar la muestra final por la columna de exclusión molecular e integrando el área de los perfiles obtenidos al medir la  $Abs_{280}$  se logró estimar que la producción de proteína es aproximadamente 2 mg/L de cultivo.

En conjunto, este trabajo demuestra que el diseño racional asistido por predicción estructural es capaz de generar variantes de la proteína E con puentes disulfuro estables y conformacionalmente viables, y que estas pueden ser expresadas, replegadas y purificadas en sistemas bacterianos con una pureza suficiente para análisis posteriores. Las limitaciones derivadas de la ausencia de PTMs y del sistema de expresión bacteriano podrían abordarse en futuros trabajos mediante expresión en sistemas eucariotas, coexpresión de chaperonas o estrategias de purificación adicionales.

Este estudio sienta así las bases para la producción de variantes estabilizadas de la proteína E que podrían tener aplicaciones en el desarrollo de antígenos más robustos para ensayos inmunológicos o como candidatos de vacunas, siempre que se

complementa con estudios estructurales de alta resolución y caracterización funcional frente a anticuerpos neutralizantes específicos del virus Zika.

## 12. BIBLIOGRAFÍA

- Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., ... Jumper, J. M. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, 630(8016), 493–500. <https://doi.org/10.1038/s41586-024-07487-w>
- Costello, A., et al. (2024). Advances in structure-guided design of flavivirus vaccines. *mSphere*, 9(2), e00961-23. <https://doi.org/10.1128/msphere.00961-24>
- Crampon, E., Covernton, E., Vaney, M. C., Dellarole, M., Sommer, S., Sharma, A., Haouz, A., ... Rey, F. A. (2023). New insight into flavivirus maturation from structure/function studies of the yellow fever virus envelope protein complex. *mBio*, 14(5), e00706-23. <https://doi.org/10.1128/mbio.00706-23>
- Dai, L., Song, J., Lu, X., Deng, Y. Q., Musyoki, A. M., Cheng, H., Zhang, Y., ... Gao, G. F. (2016). Structures of the Zika virus envelope protein and its complex with a flavivirus broadly protective antibody. *Cell Host & Microbe*, 19(5), 696–704. <https://doi.org/10.1016/j.chom.2016.04.013>
- Gonzalez, G., Hannigan, B., & DeGrado, W. F. (2014). A real-time all-atom structural search engine for proteins. *PLoS Computational Biology*, 10(7), e1003750. <https://doi.org/10.1371/journal.pcbi.1003750>
- Laimer, J., Hofer, H., Fritz, M., Wegenkittl, S., & Lackner, P. (2015). MAESTRO—Multi agent stability prediction upon point mutations. *BMC Bioinformatics*, 16(1), 116. <https://doi.org/10.1186/s12859-015-0548-6>
- Laimer, J., Hiebl-Flach, J., Lengauer, D., & Lackner, P. (2016). MAESTROweb: A web server for structure-based protein stability prediction. *Bioinformatics*, 32(9), 1414–1416. <https://doi.org/10.1093/bioinformatics/btw009>
- Metz, S. W., Thomas, A., Brackbill, A., Forsberg, J., Miley, M. J., Lopez, C. A., Lazear, H. M., Tian, S., & de Silva, A. M. (2019). Oligomeric state of the ZIKV E protein defines protective immune responses. *Nature Communications*, 10(1), 4606. <https://doi.org/10.1038/s41467-019-12549-4>
- Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., & Steinegger, M. (2022). ColabFold: Making protein folding accessible to all. *Nature Methods*, 19(6), 679–682. <https://doi.org/10.1038/s41592-022-01488-1>

- Pierson, T. C., & Diamond, M. S. (2017). The emergence of Zika virus and its new clinical syndromes. *Journal of Virology*, 91(18), e00229-17. <https://doi.org/10.1128/jvi.00229-17>
- Porath, J., Carlsson, J., Olsson, I., & Belfrage, G. (1975). Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature*, 258(5536), 598–599. <https://doi.org/10.1038/258598a0>
- Rey, F. A., Stiasny, K., Vaney, M. C., Dellarole, M., & Heinz, F. X. (2018). The bright and the dark side of human antibody responses to flaviviruses: Lessons for vaccine design. *EMBO Reports*, 19(2), 206–224. <https://doi.org/10.15252/embr.201745302>
- Schreiber, C., Müller, H., Birrenbach, O., Solenthaler, A., & Hunziker, P. (2017). A high-throughput expression screening platform to optimize the production of antimicrobial peptides. *Microbial Cell Factories*, 16(1), 29. <https://doi.org/10.1186/s12934-017-0642-9>
- Stracquadano, G., et al. (2023). Democratizing the rapid screening of protein expression for materials development. *Molecular Systems Design & Engineering*, 8(1), 76–88. <https://doi.org/10.1039/D2ME00140A>
- Vaney, M. C., Dellarole, M., Duquerroy, S., Medits, I., Tsouchnikas, G., Rouvinski, A., England, P., Stiasny, K., Heinz, F. X., & Rey, F. A. (2022). Evolution and activation mechanism of the flavivirus class II membrane-fusion machinery. *Nature Communications*, 13(1), 3718. <https://doi.org/10.1038/s41467-022-31235-7>
- Yang, C., Zeng, F., Gao, X., Zhao, S., Li, X., Liu, S., Li, N., Deng, C., Zhang, B., & Gong, R. (2019). Characterization of two engineered dimeric Zika virus envelope proteins as immunogens for neutralizing antibody selection and vaccine design. *The Journal of Biological Chemistry*, 294(27), 10638–10648. <https://doi.org/10.1074/jbc.RA119.007548>

### 13. ANEXOS

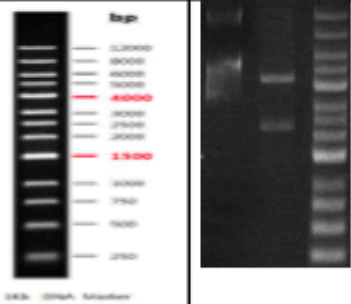
Gene Synthesis QA Report			
Gene Name	02 ZV G102C I152C	Order ID	7A05295-2
Lot#	TF87718	Cloning Vector	pET-21a(+)
Length (bp)	1269bp	Cloning Sites	NdeI(CATATG)-XhoI(CTCGAG)
Quality Control			
Test Items	Specifications	Results	
Sequencing Alignment	Sequencing data consistent with target	Pass	
Vector Sequence	Flanking sequence of cloning sites are correct	Pass	
Restriction Digest	Insert size is correct and no contaminated bands	Pass	
ORF Across Junction	Correct and consistent with target	Pass	
PCR Amplification	Correct and no contaminated bands	Pass	
Endotoxin Level	Verified, <0.1 EU/μg (Endo-Free Preps Only)	N/A	
Appearance	Clear, no foreign particles	Pass	
DNA Purity	Purity (A 260/A280 = 1.8 - 2.0)	Pass	
Restriction Digestion			
		<p>Gene name: 02 ZV G102C I152C</p> <p>Clone ID#: TF87718</p> <p>RES: ApaI-XhoI</p>	
<b>Construct Map:</b>			

Figura 1. Reporte de control de calidad “QC” de la síntesis del constructo genético utilizado para la expresión de la proteína de interés, pedido y generado por Gene Universal.

**Detailed Sequence of the Whole Construct:**

```

1 TGGCGAATGG GACGCGCCCT GTAGCGGCGC ATTAAGCGCG GCGGGTGTGG
TGGTTACGGC
61 CAGCGTGACC GCTACACTTG CCAGCGCCCT AGCGCCCGCT CCTTTCGCTT TCTTCCCTTC
121 CTTTCTCGCC ACGTTCGCCC GCTTTCCCCG TCAAGCTCTA AATCGGGGGC TCCCTTTAGG
181 GTTCCGATTT AGTGCTTTAC GGCACCTCGA CCCCAGAAAA CTTGATTAGG GTGATGGTTC
241 ACGTAGTGGG CCATCGCCCT GATAGACGGT TTTTCGCCCT TTGACGTTGG AGTCCACGTT
301 CTTTAATAGT GGACTCTTGT TCCAAACTGG AACAACTC AACCCATCTC CGGTCTATTC
361 TTTTGATTTA TAAGGGATTT TGCCGATTTG GGCCTATTGG TTAAGAAATG AGCTGATTTA
421 ACAAAAATTT AACGCGAATT TTAACAAAAT ATTAACGTTT ACAATTTGAG GTGGCAGCTT

3661 TCCAGTCGGG AAACCTGTGG TGCCAGCTGC ATTAATGAAT CGGCCAACGC GCGGGGAGAG
3721 GCGGTTTGGC TATTGGGCGC CAGGGTGGTT TTTCTTTTCA CCAGTGAGAC GGGCAACAGC
3781 TGATTGCCCT TCACCGCCTG GCCCTGAGAG AGTTGCAGCA AGCGGTCCAC CCGTGGTTGC
3841 CCCAGCAGGC GAAAACTCTG TTTGATGGTG GTTAACGGCG GGATATAACA TGAGCTGTCT
3901 TCGGTATCGT CGTATCCAC TACCAGATA TCCGCACCAA CGCGCAGCCC GGACTCGGTA
3961 ATGGCGCGCA TTGCGCCAG CGCCATCTGA TCGTTGGCAA CCAGCATCGC AGTGGGAACG
4021 ATGCCCTCAT TCAGCATTG CATGGTTTGT TGAAAACCGG ACATGGCACT CCAGTCGCCCT
4081 TCCCGTTCCG CTATCGGCTG AATTTGATTG CGAGTGAGAT ATTTATGCCA GCCAGCCAGA
4141 CGCAGACGCG CCGAGACAGA ACTTAATGGG CCCGTAACA GCGCGATTGG CTGGTGACCC
4201 AATGCGACCA GATGCTCCAC GCCCAGTCGC GTACCGTCTT CATGGGAGAA AATAACTCTG
4261 TTGATGGGTG TCTGGTCAGA GACATCAAGA AATAACGCCG GAACATTAGT GCAGGCAGCT
4321 TCCACAGCAA TGCCATCTG GTCATCCAGC GGATAGTTAA TGATCAGCCC ACTGACCGCT
4381 TGCGCGAGAA GATTGTGCAC CGCCGCTTTA CAGGCTTCGA CGCCGCTTCG TTCTACCATC
4441 GACACCACCA CGCTGGCACC CAGTTGATCG GCGCGAGAT TAATCGCCGC GACAATTTGC
4501 GACGCGCGCT GCAGGGCCAG ACTGGAGGTG GCAACGCCAA TCAGCAACGA CTGTTTGCCC
4561 GCCAGTTGTT GTGCCACGCG GTTGGGAATG TAATTCAGCT CCGCCATCGC CGCTTCCACT
4621 TTTTCCCGCG TTTTCGAGA AACGTGGCTG GCCTGGTTCA CCACGCGGGA AACGGTCTGA
4681 TAAGAGACAC CGGCATACTG TGCGACATCG TATAACGTTA CTGGTTTCACT ATTCACCACC
4741 CTGAATTGAC TCTCTCCGG GCGCTATCAT GCCATACCGC GAAAGGTTTT GCGCCATTCG
4801 ATGGTGTCCG GGATCTCGAC GCTCTCCCTT ATGCGACTCC TGCATTAGGA AGAGCCACG
4861 TAGTAGGTTG AGGCCGTTGA GCACCGCCGC CGCAAGGAAT GGTGCATGCA AGGAGATGGC
4921 GCCCAACAGT CCCCAGGCCA CGGGGCCTGC CACCATACCC ACGCCGAAAC AAGCGCTCAT
4981 GAGCCCGAAG TGCGGAGCCG GATCTTCCCC ATCGGTGATG TCGGCATGAT AGGCCCCAGC
5041 AACCCGACCT GTGGCGCCGG TGATGCGCGC CACGATGCGT CCGGCCTAGA GATCAGGAT
5101 CTCGATCCCG CGAAATTAAT ACGACTCACT ATAGGGGAAT TGTGAGCGGA TAACAATTCC
5161 CCTCTAGAAA TAATTTTGTG TAACTTTAAG AAGGAGATAT ACATATGATC CGCTGTATTG
5221 GTGTTAGTAA TCGCATTTT GTTGAAGGTA TGAGTGCCGG CACCTGGGTG GATGTTGTTT
5281 TGGAACATGG TGGTTGTGTG ACCGTGATGG CCCAGGATAA ACCGACCGTT GATATTGAAC
5341 TGGTTACCAC CACCGTGAGC AATATGGCCG AAGTGCCGAG TTATTGCTAT GAAGCAAGCA
5401 TTAGCGATAT GGCAAGTAT AGTCGTTGTC CGACCCAGGG CGAAGCATAT CTGGATAAAC
5461 AGAGTGATAC CCAGTATGTT TGTAACGTA CCCTGGTTGA TCGTGGCTGG TGCAATGGCT
5521 GTGGTCTGTT TGCCAAAGGT AGTCTGGTGA CCTGCGCAAA ATTTGCCTGC AGTAAAAAGA
5581 TGACCGGTAA AAGTATTCAG CCGGAAAAATC TGGAAATATCG CATTATGCTG AGCGTGCATG
5641 GTAGTCAGCA TAGCGGTATG TGTGTTAATG ATACCGGCCA TGAACCGAT GAAAATCGTG
5701 CAAAAGTTGA AATTACCCCG AATAGTCCGC GCGCAGAAGC AACCCCTGGT GGTGTTGGTA
5761 GTCTGGGTCT GGATTGTGAA CCGCGTACC GTCCTGGATT TAGTGATCTG TATTATCTGA
5821 CCATGAATAA TAAGCATTGG CTGGTTCATA AAGAATGGTT TCATGATATT CCGTCCCGT
5881 GGCATGCAAG TGCAATGACC GGTACCCCGC ATTGGAATAA TAAGGAAGCA CTGGTGGAAAT
5941 TCAAAGATGC ACATGCCAAA CGCCAGACCG TTGTGGTGCT GGGTAGTCAG GAAGGCGCAG
6001 TGATACCGC ACTGGCCGGT GCACTGGAAG CAGAAATGGA TGCGCCAAA GGTCTGCTGA
6061 GTAGCGGCCA TCTGAAATGC CGCCTGAAAA TGGATAAACT GCGCCTGAAA GGCCTGAGCT
6121 ATAGTCTGTG CACCGCAGCC TTTACCTTTA CCAAATTC GGCCGAAACC CTGCATGGCA
6181 CCGTTACCCT TGAAGTTGAG TATGCCGGTA CCGATGGCCC GTGTAAGTG CCGCCGAGA
6241 TGCCCGTTGA TATGCAGACC CTGACCCCGG TGGGTGCTCT GATTACCGCA AATCCGGTTA
6301 TTACCGAAAG CACCGAAAT AGCAAAATGA TGCTGGAAC GGATCCGCCG TTTGGCGATA
6361 GTTATATTGT GATTGGTGTG GCGGAAAAGA AAATTACCCA TCATTGGCAT CGTAGCGGTA
6421 GTGAAAATCT GTATTTTCAG GGCCATCATC ATCATCACCA TTAGCTCGAG CACCACCACC
6481 ACCACCCTG AGATCCGGCT GCTAACAAAG CCCGAAAGGA AGCTGAGTTG GCTGTGCCA
6541 CCGCTGAGCA ATAAGTAGCA TAACCCCTTG GGGCCTCAA ACGGGTCTTG AGGGGTTTTT
6601 TGCTGAAAGG AGGAACTATA TCCGGAT

```

//

```

481 TCGGGGAAAT GTGCGCGGAA CCCCTATTTG TTTATTTTTT TAAATACATT CAAATATGTA
541 TCCGCTCATG AGACAATAAC CCTGATAAAT GCTTCAATAA TATTGAAAAA GGAAGAGTAT
601 GAGTATTCAC CATTTCCTGT TCGCCCTTAT TCCCTTTTTT GCGGCATTTT GCCTTCCTGT
661 TTTTGTCTAC CCAGAAACGC TGGTGAAAGT AAAAGATGCT GAAGATCAGT TGGGTGCACG
721 AGTGGGTTAC ATCGAACTGG ATCTCAACAG CGGTAAGATC CTTGAGAGTT TTCGCCCGA
781 AGAACGTTTT CCAATGATGA GCACTTTTAA AGTTCTGCTA TGTGGCGCGG TATTATCCCG
841 TATTGACGCC GGGCAAGAGC AACTCGGTCC CGCATACAC TATTCTCAGA ATGACTTGGT
901 TGAGTACTCA CCAGTCACAG AAAAGCATCT TACGGATGGC ATGACAGTAA GAGAATTATG
961 CAGTGCTGCC ATAACCATGA GTGATAACAC TCGGCCAAC TTACTTCTGA CAACGATCGG
1021 AGGACCGAAG GAGCTAACCG CTTTTTTGCA CAACATGGGG GATCATGTAA CTCGCCTTGA
1081 TCGTTGGGAA CCGGAGCTGA ATGAAGCCAT ACCAAACGAC GAGCGTGACA CCACGATGCC
1141 TGAGCAATG GCAACAACGT TCGCAGAACT ATTAACCTGG GAACACTTGA CTCTAGCTTC
1201 CCGGCAACAA TTAATAGACT GGATGGAGGC GGATAAAGTT GCAGGACCAC TTCTGCCTC
1261 GGCCTTCCG GCTGGCTGGT TTATTGCTGA TAAATCTGGA GCCGGTGAGC GTGGGTCTCG
1321 CGGTATCATT GCAGCACTGG GGCAGATGG TAAGCCCTCC CGTATCGTAG TTATCTACAC
1381 GACGGGGAGT CAGGCAACTA TGGATGAACG AAATAGACAG ATCGCTGAGA TAGGTGCCCTC
1441 ACTGATTAAG CATTGGTAAC TGTCAGACCA AGTTTACTCA TATACTATT AGATTGATTT
1501 AAAACTTCAT TTTTAATTTA AAAGGATCTA GGTGAAGATC CTTTTTGATA ATCTCATGAC
1561 CAAAATCCCT TAACGTGAGT TTTCTGTTCA CTGAGCGTCA GACCCCGTAG AAAAGATCAA
1621 AGGATCTTCT TGAGATCCTT TTTTTCTGCG CGTAATCTGC TGCTTGCAAA CAAAAAACC
1681 ACCGCTACCA GCGGTGGTTT GTTTGCCGGA TCAAGAGCTA CCAACTCTTT TTCCGAAGGT
1741 AACTGGCTTC AGCAGAGCGC AGATACCAAA TACTGTCTTT CTAGTGTAGC CGTAGTTAGG
1801 CCACCACTTC AAGAACTCTG TAGCACCGCC TACATACCTC GCTCTGCTAA TCCTGTTACC
1861 AGTGGCTGCT GCCAGTGCGC ATAAGTCGTG TCTTACCCTG TTGGACTCAA GACGATAGTT
1921 ACCGGATAAG GCGCAGCGGT CGGGCTGAAC GGGGGTTTCC TGCACACAGC CCAGTTGGA
1981 GCGAACGACC TACACCGAAC TGAGATACCT ACAGCGTGAG CTATGAGAAA GCGCCACGCT
2041 TCCCGAAGGG AGAAAGCGCG ACAGGTATCC GGTAAAGCGGC AGGGTCCGGA CAGGAGAGCG
2101 CACGAGGGAG CTTCCAGGGG GAAACGCCTG GTATCTTTAT AGTCCTGTCT GGTTCGCCA
2161 CCTCTGACTT GAGCGTCGAT TTTTGTGATG CTCGTCAGGG GGGCGGAGCC TATGGAAAAA
2221 CGCAGCAAC GCGGCTTTT TACGGTTCTT GGCCTTTTGC TGGCCTTTTG CTCACATGTT
2281 CTTTCTGCTG TTATCCCCTG ATTCTGTGGA TAACCGTATT ACCGCTTTTG AGTGAGCTGA
2341 TACCGCTCGC GCGAGCCGAA CGACCGAGCG CAGCGAGTCA GTGAGCGAGG AAGCGGAAGA
2401 GCGCTGATG CGGTATTTTC TCCTTACGCA TCTGTCCGGT ATTTACACACC GCATATATGG
2461 TGCACTCTCA GTACAATCTG CTCTGATGCC GCATAGTTAA GCCAGTATAC ACTCCGCTAT
2521 CGCTACGTGA CTGGGTCATG GCTGCGCCCG GACACCCGCC AACACCCGCT GACGCGCCCT
2581 GACGGGCTTG TCTGCTCCCG GCATCCGCTT ACAGACAAGC TGTGACCCTC TCCGGGAGCT
2641 GCATGTGTCA GAGGTTTTCA CCGTCATCAC CGAAACGCGC GAGGCAGCTG CGGTAAGCT
2701 CATCAGCGTG GTCGTGAAGC GATTACACAG TGCTGCCTG TTCATCCGCG TCCAGCTCGT
2761 TGAGTTTTCTC CAGAAGCGTT AATGTCTGGC TTCTGATAAA GCGGGCCATG TTAAGGGCGG
2821 TTTTTTCTG TTTGGTCACT GATGCCTCCG TGTAAGGGGG ATTTCTGTTC ATGGGGGTAA
2881 TGATACCGAT GAAACGAGAG AGGATGCTCA CGATACGGGT TACTGATGAT GAACATGCC
2941 GGTTACTGGA ACGTTGTGAG GGTAAACAAC TGGCGGTATG GATGCGGCGG GACCAGAGAA
3001 AAATCACTCA GGGTCAATGC CAGCGCTTCC TTAATACAGA TGTAGGTGTT CCACAGGGTA
3061 GCCAGCAGCA TCCTGCGATG CAGATCCGGA ACATAATGGT GCAGGGCGCT GACTTCCGCG
3121 TTTCCAGACT TTACGAAACA CGGAAACCGA AGACCATTCA TGTGTGTGCT CAGGTCCGAG
3181 ACGTTTTGCA GCAGCAGTGC CTTACGTTT GCTCGCGTAT CGGTGATTCA TTCTGCTAAC
3241 CAGTAAGGCA ACCCCGCCAG CCTAGCCGGG TCCTCAACGA CAGGAGCAGC ATCATGCGCA
3301 CCCGTGGGGC CGCCATGCCG GCGATAATGG CCTGCTTCTC GCCGAAACGT TTGGTGGCGG
3361 GACCAGTGAC GAAGGCTTGA GCGAGGGCGT GCAAGATTCC GAATACCGCA AGCGACAGGC
3421 CGATCATCGT CGCGCTCCAG CGAAAGCGGT CCTCGCCGAA AATGACCCAG AGCGCTGCCG
3481 GCACCTGTCC TACGAGTTGC ATGATAAAGA AGACAGTCAT AAGTGCGGCG ACGTATGTA
3541 TGCCCCGCGC CCACCGAAG GAGCTGACTG GGTGAAGGC TCTCAAGGGC ATCGGTGCGA
3601 ATCCCGGTGC CTAATGAGTG AGCTAACTTA CATTAAATTG GTTGCCTCA CTGCCGCTT

```

Figura 2. Secuencia de ADN de la variante de la proteína E doble mutante.

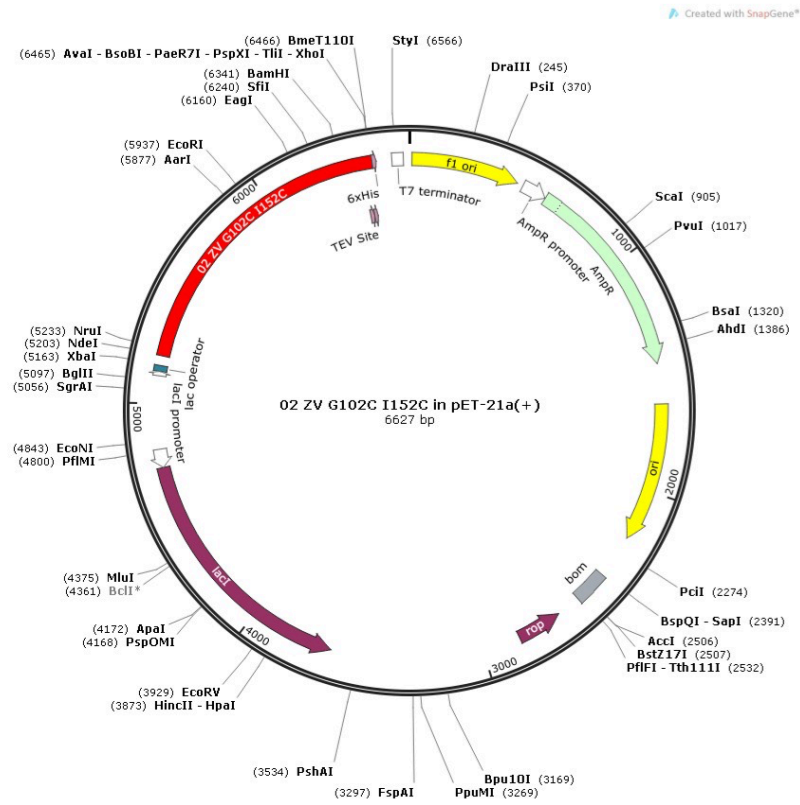


Figura 3.. Mapa del vector de expresión pET-21a(+)

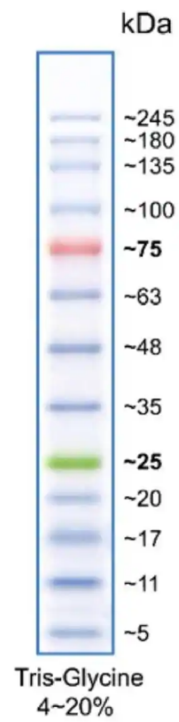


Figura 4. Marcador de peso molecular Blue Stain Protein Marker (5-245 kDa) utilizado en un gel de SDS- PAGE al 12% Tris-glicina.