



UNIVERSIDAD TÉCNICA DE AMBATO
FACULTAD DE CIENCIAS DE LA SALUD
CARRERA DE LABORATORIO CLÍNICO

TEMA

**“ANÁLISIS INMUNOINFORMÁTICO Y DISEÑO IN SILICO DE
PROTEÍNA MULTIEPITOPOS DEL VIRUS ZIKA PARA
INMUNODIAGNÓSTICO DIFERENCIAL ESPECÍFICO EN ZONAS
ENDÉMICAS”**

Requisito previo para optar por el Título de Licenciada en Laboratorio Clínico

Autora: Balseca Quishpe, Carmen Elena

Tutor: PhD. Salazar Garcés, Luis Fabián

Ambato – Ecuador

Enero – Julio 2026

A. PÁGINAS PRELIMINARES

APROBACIÓN DEL TUTOR

En mi calidad de Tutor del Proyecto de Investigación con el tema:

“ANÁLISIS INMUNOINFORMÁTICO Y DISEÑO IN SILICO DE PROTEÍNA MULTIEPITOPOS DEL VIRUS ZIKA PARA INMUNODIAGNÓSTICO DIFERENCIAL ESPECÍFICO EN ZONAS ENDÉMICAS”, de la Srta. Balseca Quishpe, Carmen Elena, estudiante de la Carrera de Laboratorio Clínico de la Universidad Técnica de Ambato. Considero que reúne los requisitos y méritos suficientes para ser sometido a la evaluación por el jurado examinador designado por el Honorable Consejo Directivo de la Facultad de Ciencias de la Salud.

Ambato, Julio 2026

EL TUTOR

PhD. Salazar Garcés, Luis Fabián

AUTORÍA DEL TRABAJO DE TITULACIÓN

Los criterios emitidos en el Trabajo de Investigación sobre:

“ANÁLISIS INMUNOINFORMÁTICO Y DISEÑO IN SILICO DE PROTEÍNA MULTIEPITOPOS DEL VIRUS ZIKA PARA INMUNODIAGNÓSTICO DIFERENCIAL ESPECÍFICO EN ZONAS ENDÉMICAS”, como también los contenidos, ideas, objetivos y futura aplicación del trabajo de investigación son de exclusiva responsabilidad de mi persona, como autor de este trabajo de grado.

Ambato, Julio 2026

AUTORA

Balseca Quishpe, Carmen Elena

DERECHOS DE AUTOR

Autorizo a la Universidad Técnica de Ambato, para que haga de esta tesis o parte de ella un documento disponible para su lectura, consulta y procesos de investigación.

Cedo los derechos en línea patrimonial de mi tesis con fines de difusión pública: además apruebo la reproducción de este proyecto de investigación, dentro de las regulaciones de la Universidad, siempre y cuando esta reproducción no suponga una ganancia económica y se realice respetando mis derechos de autor.

Ambato, Julio 2026

AUTORA

Balseca Quishpe, Carmen Elena

APROBACIÓN DEL JURADO EXAMINADOR

Los miembros del Tribunal Examinador aprueban el Informe de Investigación, sobre el tema: **“ANÁLISIS INMUNOINFORMÁTICO Y DISEÑO IN SILICO DE PROTEÍNA MULTIEPITOPOS DEL VIRUS ZIKA PARA INMUNODIAGNÓSTICO DIFERENCIAL ESPECÍFICO EN ZONAS ENDÉMICAS”** de Balseca Quishpe, Carmen Elena estudiante de la Carrera de Laboratorio Clínico.

Ambato, Julio 2026

Para constancia firman:

PRESIDENTE/A

1er VOCAL

2do VOCAL

DEDICATORIA

En primer lugar, dedico este trabajo de investigación a Dios, que me ha guiado en cada uno de mis pasos, me ha cuidado y me dado la sabiduría en momentos difíciles permitiéndome cumplir cada meta, avanzar en cada obstáculo y como no, permitiéndome culminar esta importante etapa de mi formación profesional.

A mi mamá, Carmen, quien, con su amor incondicional, sacrificio, apoyo constante y confianza en cada uno de mis sueños, ha sido fundamental en mi formación tanto personal como profesional. Gracias por enseñarme que con constancia, perseverancia y fe es posible alcanzar cualquier meta. Este logro también le pertenece porque cada paso que he dado siempre estuvo a mi lado, brindándome su amor, consejos y apoyo incondicional.

A mis hermanas y hermanos, Franklin, Germania, Silvana, Jona y de manera especial a María Cristina, por su cariño, comprensión y compañía a lo largo de mi vida. Gracias por estar presentes en los momentos más felices, de celebración y éxito, pero también en aquellos de incertidumbre y dificultades en los que los necesite. Su confianza en mí, su apoyo y palabras de aliento siempre me han fortalecido para seguir adelante. Cada experiencia compartida, cada risa, cada logro y cada momento junto a ellos han contribuido de una u otra manera en el esfuerzo y culminación de este logro académico.

Y mi mejor amiga Lizbeth, por estar presente en cada etapa de este proceso, por escucharme, motivarme y brindarme su amistad sincera. Gracias por compartir conmigo las alegrías, preocupaciones y aprendizajes que formaron parte de esta experiencia.

Balseca Quishpe, Carmen Elena

AGRADECIMIENTO

A Dios, por brindarme la vida, salud, fortaleza, sabiduría y valentía para afrontar cada desafío durante la formación académica. Pese a que no todo ha sido fácil, su guía y bendiciones me permiten mantener siempre la fe y ser más optimista y perseverante ante las dificultades.

A mis padres, por creer en mí, apoyarme en mis decisiones y enseñarme con su ejemplo el valor del esfuerzo, la responsabilidad, la confianza y el respeto. Este logro también es el reflejo de todo lo que han sembrado en mi vida.

A mi tutor, PhD Fabián Salazar, por su invaluable apoyo, dedicación y compromiso durante el desarrollo de la presente investigación. Agradezco infinitamente el tiempo brindado, los conocimientos impartidos y la paciencia con la que oriento cada etapa de este proceso académico. Sus observaciones, recomendaciones y enseñanzas contribuyen totalmente al fortalecimiento de este trabajo, permitiéndome expandir mis conocimientos, así como también poder desarrollar áreas nuevas como lo es la biotecnología. Su guía y acompañamiento fueron fundamentales para la culminación exitosa de esta investigación.

A mis amigas incondicionales, Alexandra y Margarita, por su amistad sincera a lo largo de estos años, por las risas, las conversaciones profundas, por los momentos tan lindos que compartimos dentro y fuera de las aulas. Definitivamente gracias a su hermosa amistad, siento que este camino ha sido un poco más fácil y significativo. Llevo conmigo momentos gratos, valiosos y de aprendizaje, gracias por formar parte de mi vida.

También a Luis y a todos los que estuvieron durante mi vida universitaria, los llevare siempre en mi corazón.

Balseca Quishpe, Carmen Elena

ÍNDICE GENERAL

A. PÁGINAS PRELIMINARES	ii
APROBACIÓN DEL TUTOR	ii
AUTORÍA DEL TRABAJO DE TITULACIÓN	iii
DERECHOS DE AUTOR	iv
APROBACIÓN DEL JURADO EXAMINADOR	v
DEDICATORIA	vi
AGRADECIMIENTO	vii
RESUMEN	xiii
SUMMARY	xiv
INTRODUCCIÓN	1
B. CONTENIDO	3
CAPITULO I.	3
1.1. Antecedentes Investigativos.....	3
1.1.1. Virus Zika: estructura, genoma y relevancia clínica	3
1.1.2. Diagnóstico serológico y limitaciones por reactividad cruzada.....	4
1.1.3. Proteínas E y NS1 como blancos antigénicos	6
1.1.4. Inmunoinformática aplicada al diseño de antígenos	7
1.1.5. Constructos multiepítopo en diagnóstico	9
1.1.6. Modelado estructural y validación computacional	11
1.1.7. Optimización génica para expresión heteróloga	13
1.2. OBJETIVOS.....	16
1.2.1. Objetivo general.....	16
1.2.2. Objetivos específicos	16
CAPÍTULO II.....	17
2.1. MATERIALES	17
2.1.1. Recursos humanos.....	17
2.1.2. Recursos tecnológicos.....	17
2.1.3. Recursos institucionales.....	18
2.1.4. Recursos económicos.....	18
2.2. MÉTODOS.....	19
2.2.1. Tipo y diseño de Estudio/Enfoque de la investigación	19

2.2.2. Aspectos éticos.....	19
2.2.3. Variables de la investigación	20
2.2.4. Hipótesis de investigación	22
2.2.5. Descripción de Procedimientos.....	23
CAPÍTULO III.	31
3.1 ANÁLISIS DE RESULTADOS	31
3.1.1. Recuperación y curaduría de secuencias.....	31
3.1.2. Evaluación de la proteína completa	31
3.1.3. Predicciones y selección de epítomos B	33
3.1.4. Evaluación de antigenicidad de epítomos candidatos	34
3.1.5. Evaluación de alergenidad y toxicidad de los epítomos.....	38
3.1.6. Selección final de epítomos para el constructo multiepítomo.....	43
3.1.7. Diseño del constructo multiepítomo	46
3.1.8. Análisis fisicoquímico del constructo	47
3.1.9. Modelado estructural del constructo	48
3.1.10. Refinamiento y validación estructural	50
3.1.11. Optimización génica y análisis de ARNm	52
3.1.12. Análisis BLASTp y verificación de similitud frente a CHIKV y OROV	54
3.1.13. Diseño plasmídico.....	55
3.2. DISCUSIÓN.....	58
3.3. VERIFICACIÓN DE HIPÓTESIS.....	61
CAPITULO IV	63
4.1. CONCLUSIONES.....	63
4.2. RECOMENDACIONES.....	64
REFERENCIAS BIBLIOGRÁFICAS	65
ANEXOS	69

ÍNDICE DE TABLAS

Tabla 1: Optimización de variables.	20
Tabla 2: Resultados del análisis antigénico, alergénico y toxico de la poliproteína del Virus Zika en la secuencia con 1750 aa	32
Tabla 3: Resultados del análisis antigénico, alergénico y toxico de la poliproteína del Virus Zika, en la secuencia con 1675 aa	32
Tabla 4: Epítomos de célula B en la secuencia de 1750 aa. Se coloca los 10 primeros péptidos, siendo estos quienes eran más probables de ser epítomos.	33
Tabla 5: Epítomos de célula B en la secuencia de 1675 aa. Se coloca los 10 primeros péptidos, siendo estos quienes eran más probables de ser epítomos.	34
Tabla 6: Antigenicidad en 20 secuencias de 16 aa. Se ordenó de la puntuación más alta a la más baja puntuación de antigenicidad, de los epítomos B lineales predichos anteriormente.	35
Tabla 7: Antigenicidad en 20 secuencias, considerando el corte a partir de 12 aa iniciales.	36
Tabla 8: Antigenicidad en 20 secuencias, considerando el corte de los 12 aa finales.	37
Tabla 9: Análisis de alergenicidad por epítomos, en 3 herramientas distintas: AllerTOP v2.1 Allergen FP AlgPred. Considerando los epítomos B lineales de 16 aa.	39
Tabla 10: Análisis de alergenicidad por epítomos, considerando el corte de los 12 aa iniciales.	40
Tabla 11: Análisis de alergenicidad por epítomos, considerando el corte de los 12 aa iniciales.	41
Tabla 12: Análisis de 500 aa derivados de la secuencia completa y regiones toxicas.	42
Tabla 13: Epítomos iniciales de 12 aa seleccionados tras el proceso de validación.	44
Tabla 14: Epítomos finales de 12 aa seleccionados tras el proceso de validación.	45
Tabla 15: Secuencias aprobadas tras los filtros, dichas secuencias de 12 aa iniciales y finales.	46
Tabla 16: Análisis fisicoquímico mediante ProtParam.	48

Tabla 17: *Criterios de cumplimiento o rechazo de la Hipótesis Alternativa (H1), en base al diseño constructo proteico multiepitopos de ZIKV generado en la presente investigación.* _____ 61

ÍNDICE DE FIGURAS

Figura 1: Flujo de trabajo.	24
Figura 2: Verificación del registro MF438286.1 mediante BLASTx frente a secuencias proteicas de referencia del virus Zika depositadas en NCBI. El análisis mostró una cobertura de consulta del 100 %, valor E de 0,0, identidad de 99,47 % y puntuación máxima/total de 6802.	31
Figura 3: Ensamblaje del constructo.	47
Figura 4: Estructura tridimensional del constructo obtenida a través del servidor AlphaFold.	49
Figura 5: Rank_1 al Rank_5	49
Figura 6: Gráfica de la cobertura de secuencias homólogas.	50
Figura 7: Predicción de IDDT por posición.	50
Figura 8: Imagen tridimensional refinada generada por GalaxyRefine (Model_1)	51
Figura 9: Puntuaciones obtenidas para los cinco modelos generados por GalaxyRefine	51
Figura 10: Resultado de la herramienta ERRAT obtenido del servidor SAVES v6.1.	51
Figura 11: Resultado de Ramachandran Plot en PROCHEK en el servidor SAVES v6.1.	52
Figura 12: Optimización de codones del constructo multiepítopo para Escherichia coli K12 mediante JCat. Se muestra el ADN optimizado, CAI = 1,0 y contenido GC = 59,78%.	53
Figura 13: Resultados en RNAFold	53
Figura 14: Estructuras secundarias finales de ARNm	54
Figura 15: Análisis de similitud del constructo proteico multiepítopo mediante BLASTp en el servidor NCBI BLAST.	54
Figura 16: Reactividad cruzada con CHIKV.	55
Figura 17: Reactividad cruzada frente OROV.	55
Figura 18: Inserción in silico del gen optimizado a través del vector pET28a (+).	56
Figura 19: Propiedades bioquímicas del plásmido recombinante obtenidas del servidor Benchling.	57

UNIVERSIDAD TÉCNICA DE AMBATO
FACULTAD DE CIENCIAS DE LA SALUD
CARRERA DE LABORATORIO CLÍNICO

**ANÁLISIS INMUNOINFORMÁTICO Y DISEÑO IN SILICO DE PROTEÍNA
MULTIEPITOPOS DEL VIRUS ZIKA PARA INMUNODIAGNÓSTICO
DIFERENCIAL ESPECÍFICO EN ZONAS ENDÉMICAS**

Autora: Balseca Quishpe, Carmen Elena

Tutor: PhD. Salazar Garcés, Luis Fabián

Fecha: Julio 2026

RESUMEN

El virus Zika (ZIKV) constituye un importante problema de salud pública debido a su asociación con complicaciones neurológicas y congénitas, además de la dificultad diagnóstica derivada de su cocirculación con otros arbovirus como Chikungunya (CHIKV) y Oropouche (OROV). La presente investigación, enmarcada en la línea de inmunología, biotecnología y salud pública, tuvo como objetivo diseñar y validar in silico una proteína multiepítipo con potencial inmunodiagnóstico diferencial.

Para ello, se recuperó una secuencia nucleotídica de ZIKV (MF438286.1) desde GenBank del NCBI y fue sometida a traducción mediante la herramienta BLASTx, obteniéndose la secuencia aminoacídica (poliproteína ZIKA ALL27019). Se realizó la predicción de epítomos B lineales mediante ABCPred y la evaluación de antigenicidad, alergenicidad, toxicidad y calidad estructural mediante herramientas inmunoinformáticas.

Como resultado, se obtuvo un constructo multiepítipo de 155 aminoácidos con ausencia de toxicidad y alergenicidad además de ser estable. Estos resultados respaldan su potencial como base científica para ser aplicado en investigaciones futuras orientadas al desarrollo de inmunodiagnóstico diferencial para ZIKV.

PALABRAS CLAVE: VIRUS ZIKA, INMUNOINFORMÁTICA, EPÍTOPOS B LINEALES, PROTEÍNA MULTIEPÍTOPO, INMUNODIAGNÓSTICO DIFERENCIAL, ARBOVIRUS.

TECHNICAL UNIVERSITY OF AMBATO

FACULTY OF HEALTH SCIENCES

CLINICAL LABORATORY CAREER

**IMMUNOINFORMATICS ANALYSIS AND IN SILICO DESIGN OF ZIKA
VIRUS MULTI-EPITOPE PROTEIN FOR SPECIFIC DIFFERENTIAL
IMMUNODIAGNOSIS IN ENDEMIC ÁREAS**

Author: Balseca Quishpe, Carmen Elena

Tutor: PhD. Salazar Garcés, Luis Fabián

Date: July 2026

SUMMARY

The Zika virus (ZIKV) poses a significant public health challenge due to its association with neurological and congenital complications, as well as the diagnostic difficulties arising from its co-circulation with other arboviruses such as Chikungunya (CHIKV) and Oropouche (OROV). This study, situated within the fields of immunology, biotechnology, and public health, aimed to design and validate *in silico* a multi-epitope protein with potential for differential immunodiagnosis.

To this end, a ZIKV nucleotide sequence (MF438286.1) was retrieved from the NCBI GenBank and translated using the BLASTx tool, yielding the corresponding amino acid sequence (ZIKA polyprotein ALL27019). Linear B-cell epitopes were predicted using ABCPred, and the construct was evaluated for antigenicity, allergenicity, toxicity, and structural quality using immunoinformatic tools.

The result was a stable, 155-amino-acid multi-epitope construct that showed no toxicity or allergenicity. These findings support its potential as a scientific basis for future research aimed at developing differential immunodiagnostic methods for ZIKV.

KEYWORDS: ZIKA VIRUS, IMMUNOINFORMATICS, LINEAR B CELL EPITOPES, MULTI-EPITOPE PROTEIN, DIFFERENTIAL IMMUNODIAGNOSIS, ARBOVIRUS.

INTRODUCCIÓN

El virus Zika (ZIKV) es un arbovirus perteneciente al género *Flavivirus* y a la familia *Flaviviridae*, transmitido principalmente por artrópodos del género *Aedes*, especialmente *Aedes aegypti* y *Aedes albopictus*, (1). Su genoma está constituido por ARN monocatenario de sentido positivo, organizado en tres segmentos correspondientes a la cápside (C), premembrana/membrana (prM/M) y envoltura (E), así como en siete proteínas no estructurales denominadas NS1, NS2A, NS2B, NS3, NS4A, NS4B y NS5 siendo estas indispensables en la replicación viral, ensamblaje e interacciones con células hospederas. (2)

La infección por el ZIKV ha sido documentada en 87 países a nivel mundial, siendo Brasil el país con el mayor número de casos de infección por ZIKV reportados en América, y también con el primer caso autóctono registrado en 2015. (3,4) Sin embargo, no fue hasta el 2016 que la World Health Organization (WHO), declaró como emergencia de salud pública, debido a los riesgos asociados a la infección, incluidas las malformaciones congénitas como la microcefalia o complicaciones neurológicas como el Síndrome de Guillain-Barré. Pese a que el último caso registrado del ZIKV a nivel nacional según el Ministerio de Salud Pública fue en el año 2018, la infección continúa siendo motivo de interés, ya que este virus sigue circulando en regiones tropicales y subtropicales manteniendo latente el riesgo de nuevos brotes y así surge la necesidad de fortalecer estrategias de inmunodiagnóstico diferencial. (5,6)

Ante esta problemática, la inmunoinformática emerge como una estrategia eficaz para el diseño racional de antígenos, ya que integra herramientas computacionales que permiten analizar secuencias proteicas virales, seleccionar epítomos con características inmunológicas favorables y evaluar propiedades relacionadas con antigenicidad, alergenicidad, toxicidad, estabilidad estructural y especificidad molecular antes de su validación experimental. (7,8) Además, la inmunoinformática computacional *in silico* también puede ayudar al desarrollo de vacunas mediante un diseño racional de manera eficiente en tiempo y coste. (9)

En este contexto, la presente investigación tuvo como finalidad diseñar y validar *in silico* una proteína multiepítopo específica del virus Zika mediante herramientas de inmunoinformática y biología computacional. Para ello, se identificaron regiones antigénicas conservadas, se seleccionaron epítomos B lineales con propiedades inmunológicas favorables y se evaluaron las características fisicoquímicas, estructurales y de especificidad del constructo diseñado. Con este enfoque se buscó generar una base científica que contribuya al desarrollo futuro de estrategias de inmunodiagnóstico diferencial para ZIKV en zonas donde coexisten múltiples arbovirus.

B. CONTENIDO

CAPITULO I.

MARCO TEÓRICO

1.1. Antecedentes Investigativos

1.1.1. Virus Zika: estructura, genoma y relevancia clínica

El virus Zika (ZIKV) es un arbovirus de ARN monocatenario de sentido positivo perteneciente a la familia *Flaviviridae* y al género *Flavivirus* (10). El mecanismo de transmisión se da principalmente por picadura de mosquitos *Aedes*, especialmente *Aedes aegypti* y *Aedes albopictus*, los cuales actúan como vectores relevantes en regiones tropicales y subtropicales (1). Desde el punto de vista molecular, el genoma del ZIKV codifica una poliproteína que posteriormente es procesada en tres proteínas estructurales, correspondientes a la cápside (C), premembrana/membrana (prM/M) y envoltura (E), así como en siete proteínas no estructurales denominadas NS1, NS2A, NS2B, NS3, NS4A, NS4B y NS5. Entre estas, la proteína de envoltura (E) constituye uno de los principales determinantes de entrada viral y representa un blanco predominante de anticuerpos neutralizantes, debido a su participación en el reconocimiento celular y en los procesos de fusión de membranas (2).

La relevancia clínica del ZIKV se relaciona con su capacidad para producir infecciones sintomáticas y asintomáticas, así como con su asociación a complicaciones neurológicas y congénitas. El virus adquirió importancia epidemiológica mundial a partir de los brotes registrados en América Latina, especialmente en Brasil durante el año 2015, donde las condiciones ecológicas favorecieron una rápida expansión del vector y de la transmisión viral. Esta situación motivó que la Organización Mundial de la Salud declarara una emergencia de salud pública de importancia internacional en 2016, debido a la relación entre la infección por ZIKV, el síndrome congénito por Zika y otras alteraciones neurológicas (3,5).

Aunque la magnitud epidémica observada después de 2016 disminuyó, la transmisión del ZIKV persiste de manera endémica en varios países de la Región de las Américas, lo que mantiene su importancia para los sistemas de vigilancia epidemiológica, esto según Pan American Health Organization [PAHO] (2023)(11). En el contexto nacional, actualmente no se han reportado casos confirmados recientes, siendo los últimos casos notificados en el año 2018 de acuerdo con los registros del Ministerio de Salud Pública(6). Sin embargo, la ausencia de casos confirmados no descarta la posibilidad de infecciones subdiagnosticadas, especialmente en zonas con presencia del vector *Aedes aegypti* y con circulación concomitante de otros arbovirus. Por esta razón, el país mantiene vigilancia epidemiológica activa ante el riesgo de reemergencia de la enfermedad (4,6).

Pese al conocimiento acumulado sobre la estructura genómica, los mecanismos de transmisión y los principales blancos antigénicos del ZIKV, aún no existe un biofármaco, antiviral o vacuna aprobada de uso general que contrarreste de forma específica la infección. Esta limitación ha impulsado el desarrollo de investigaciones orientadas a antivirales, vacunas y herramientas diagnósticas, debido a que la patogénesis viral y los mecanismos de evasión inmunitaria no se comprenden por completo (12). En este sentido, se han descrito candidatos vacunales basados en virus vivos atenuados, virus inactivados, subunidades proteicas, vectores virales, ADN, ARNm e incluso antígenos derivados de la saliva del mosquito. No obstante, la mayoría de estas aproximaciones aún se encuentra en fases clínicas I/II o en etapas preclínicas de evaluación (13). Estos antecedentes evidencian la necesidad de fortalecer estrategias complementarias, particularmente aquellas orientadas al diagnóstico diferencial y a la identificación de blancos antigénicos específicos (12,13).

1.1.2. Diagnóstico serológico y limitaciones por reactividad cruzada

El diagnóstico del ZIKV representa un desafío para los laboratorios clínicos, principalmente en áreas donde existe cocirculación de arbovirus. En fases tempranas de la infección, las pruebas moleculares permiten detectar el material genético viral; sin embargo, la ventana de detección puede ser limitada y depende del tipo de muestra, del tiempo transcurrido desde el inicio de los síntomas y de la carga viral. Por esta razón, los métodos serológicos continúan siendo herramientas relevantes para la

vigilancia epidemiológica y para el diagnóstico complementario, especialmente cuando la detección molecular ya no es posible o cuando se requiere evaluar exposición previa.

No obstante, las pruebas serológicas presentan una limitación importante: la reactividad cruzada entre flavivirus. Esta situación se debe a la similitud estructural y antigénica existente entre proteínas virales conservadas, especialmente en virus relacionados como ZIKV y dengue. Dicha similitud ha sido asociada con el reconocimiento cruzado de anticuerpos, lo cual afecta la especificidad diagnóstica y dificulta la interpretación de resultados en contextos de cocirculación viral (14)

Investigaciones recientes, como la de Katzelnick et al, (2021), han confirmado que los ensayos basados en proteínas completas pueden inducir reconocimiento cruzado debido a epítomos conservados, lo que limita la precisión de los inmunoensayos convencionales y puede incrementar el riesgo de falsos positivos o resultados indeterminados (15)

En América Latina, esta problemática adquiere mayor relevancia debido a la circulación simultánea de ZIKV con otros arbovirus como dengue, Chikungunya y Oropouche. Aunque Chikungunya y Oropouche pertenecen a familias virales distintas, la superposición clínica entre estas infecciones puede dificultar el diagnóstico inicial, particularmente en regiones amazónicas o tropicales donde se han reportado brotes concurrentes (16). Además, la respuesta inmune policlonal generada durante infecciones virales puede interferir en la interpretación de resultados serológicos, aun cuando los virus no pertenezcan al mismo género. Por esta razón, el diseño de antígenos diagnósticos más específicos constituye una necesidad metodológica y clínica para mejorar la diferenciación entre infecciones por arbovirus (17).

En este contexto, la selección precisa de regiones antigénicas se considera un paso crítico para el desarrollo de herramientas diagnósticas. La utilización de proteínas completas puede incrementar la sensibilidad del ensayo, pero también puede comprometer la especificidad cuando existen regiones conservadas compartidas entre virus relacionados. Por el contrario, el uso de péptidos o constructos multiepítomo permite seleccionar regiones más específicas, reduciendo la probabilidad de reconocimiento cruzado y mejorando el desempeño potencial de los inmunoensayos.

Esta estrategia resulta especialmente relevante para estudios preliminares de diseño antigénico, donde las herramientas computacionales permiten filtrar candidatos antes de su validación experimental.

1.1.3. Proteínas E y NS1 como blancos antigénicos

Las proteínas virales E y NS1 han sido ampliamente estudiadas como blancos antigénicos del ZIKV debido a su relevancia en la interacción virus-hospedero y en la respuesta inmune. La proteína E participa en el proceso de unión y entrada del virus a la célula hospedera, por lo que constituye un blanco importante de anticuerpos neutralizantes. Sin embargo, su conservación parcial entre flavivirus puede favorecer fenómenos de reactividad cruzada, especialmente cuando se emplean dominios completos o regiones altamente conservadas en ensayos serológicos. Esta característica obliga a seleccionar cuidadosamente los segmentos antigénicos de la proteína E, priorizando regiones con capacidad de reconocimiento inmunológico y baja similitud con proteínas de virus relacionados (2,14).

Por otra parte, la proteína no estructural 1 (NS1) representa un biomarcador relevante para el diagnóstico y la investigación inmunológica del ZIKV. Esta proteína participa en procesos asociados a replicación viral, interacción con el sistema inmune y patogénesis, además de ser secretada durante la infección, lo que la convierte en un blanco atractivo para el desarrollo de ensayos inmunológicos. En este sentido, Roldan et al., (2021) desarrollaron un ensayo inmunoenzimático competitivo indirecto para la detección temprana de la infección por ZIKV basado en la proteína no estructural 1 (NS1), sin presentar reactividad cruzada significativa. Este hallazgo respalda la utilidad de NS1 como blanco diagnóstico cuando se seleccionan regiones antigénicas adecuadas (18)

De manera complementaria, Yu et al, (2021), demostraron que los anticuerpos monoclonales dirigidos a NS1 pueden conferir protección contra la infección por ZIKV, esto realizado en ratones neonatos C57BL/6. En dicho estudio, los anticuerpos monoclonales 3G2 y 4B8 mostraron mayor eficacia que el anticuerpo monoclonal 4F10, lo que evidencia que no todas las regiones de NS1 y de evaluar su antigenicidad,

alergenicidad, toxicidad y especificidad antes de su aplicación en plataformas diagnósticas (19)

En conjunto, los estudios sobre E y NS1 demuestran que ambas proteínas poseen relevancia como fuentes de epítomos para el diseño de antígenos. No obstante, la selección de regiones debe orientarse hacia segmentos que mantengan capacidad de reconocimiento por anticuerpos y que, al mismo tiempo, reduzcan la similitud con proteínas de otros arbovirus. En este sentido, la proteína NS1 se considera un blanco particularmente útil para el desarrollo de herramientas diagnósticas diferenciales, debido a su exposición inmunológica durante la infección y a la posibilidad de identificar regiones con menor reactividad cruzada (1). A pesar de estos avances, la alta similitud entre flavivirus continúa representando una limitante para el desarrollo de métodos altamente específicos, por lo que el uso de enfoques inmunoinformáticos constituye una alternativa pertinente para optimizar la selección de regiones antigénicas (14)

1.1.4. Inmunoinformática aplicada al diseño de antígenos

La inmunoinformática ha emergido como una estrategia eficaz para el diseño racional de antígenos, ya que integra herramientas computacionales orientadas a identificar, filtrar y priorizar epítomos con potencial inmunogénico. Este enfoque permite analizar propiedades fisicoquímicas, accesibilidad superficial, conservación de secuencia, antigenicidad, alergenicidad, toxicidad y patrones estructurales antes de realizar ensayos experimentales. De acuerdo con Rawat et al, (2023), la inmunoinformática facilita el análisis preliminar de candidatos antigénicos mediante algoritmos predictivos, reduciendo tiempo, costos y recursos asociados al diseño tradicional de proteínas recombinantes (7)

En el caso del ZIKV, la inmunoinformática es particularmente útil debido a la necesidad de seleccionar epítomos que permitan diferenciarlo de otros virus con manifestaciones clínicas similares o con similitud antigénica. La predicción de epítomos B lineales permite identificar regiones peptídicas capaces de ser reconocidas por anticuerpos, lo cual resulta relevante para el diseño de antígenos aplicables en inmunodiagnóstico. Además, estas herramientas permiten estimar si una región

candidata podría presentar reactividad cruzada con otros virus, aspecto directamente relacionado con la especificidad diagnóstica (8).

El diseño *in silico* ofrece ventajas importantes en etapas iniciales de investigación, especialmente cuando se trabaja con agentes infecciosos de interés epidemiológico. Al no requerir manipulación directa de muestras clínicas ni cultivo viral, permite reducir riesgos biológicos y orientar de manera más eficiente la fase experimental posterior. Sin embargo, sus resultados deben interpretarse como predicciones computacionales y no como validaciones definitivas. Por esta razón, todo candidato antigénico diseñado mediante inmunoinformática debe considerarse preliminar hasta ser evaluado mediante ensayos *in vitro*, expresión recombinante, inmunoensayos y análisis de desempeño diagnóstico.

Siguiendo este enfoque, Antonelli et al, (2022) desarrollaron una investigación orientada al diseño *in silico* de una vacuna multiepítipo contra el ZIKV, con el objetivo de identificar regiones antigénicas capaces de inducir respuesta inmune frente a este virus. Como resultado, el constructo diseñado presentó adecuada antigenicidad, estabilidad y ausencia de alergenicidad, además de una cobertura poblacional estimada del 87,86 %. (20). De manera complementaria, Alharbi et al, (2024) desarrollaron una aproximación similar, en la cual se obtuvieron nueve epítomos con características antigénicas, no alergénicas y no tóxicas, los cuales fueron ensamblados en un constructo multiepítipo mediante el uso del enlazador GPGPG y un enlazador adicional EAAAK para unir el adyuvante. Estos estudios demuestran que la inmunoinformática puede orientar la construcción de proteínas sintéticas con características favorables antes de su evaluación experimental (21)

Aunque varios trabajos se han enfocado en el diseño de vacunas, los mismos principios metodológicos pueden adaptarse al desarrollo de antígenos diagnósticos. En este caso, el objetivo principal no es inducir una respuesta inmune protectora, sino diseñar una proteína capaz de ser reconocida de manera específica por anticuerpos generados durante la infección. Por ello, el criterio de selección debe priorizar la antigenicidad, la exposición de epítomos, la ausencia de alergenicidad y toxicidad, así como la baja similitud frente a proteínas de otros virus clínicamente relacionados. Este enfoque permite proponer constructos con potencial utilidad en inmunoensayos serológicos, siempre que sean posteriormente validados en condiciones experimentales.

1.1.5. Constructos multiepítopo en diagnóstico

Los constructos multiepítopo consisten en proteínas sintéticas diseñadas mediante la unión de diferentes regiones antigénicas seleccionadas racionalmente. Estas regiones pueden proceder de una o varias proteínas del mismo agente infeccioso y se ensamblan mediante enlazadores peptídicos que buscan preservar la independencia estructural, flexibilidad y exposición de cada epítopo. Este enfoque ha sido aplicado en estudios *in silico* contra el ZIKV, en los cuales los epítomos candidatos fueron seleccionados considerando antigenicidad, ausencia de alergenicidad, ausencia de toxicidad y estabilidad del constructo final (20,21). En el contexto diagnóstico, esta estrategia permite concentrar regiones inmunológicamente relevantes en una sola molécula, evitando el uso de proteínas completas que podrían contener dominios conservados asociados a reactividad cruzada, una limitación descrita en ensayos serológicos frente a flavivirus relacionados (14,15).

El uso de enlazadores como GPGPG ha sido reportado en diseños multiepítopo debido a su utilidad para separar regiones antigénicas y favorecer la presentación individual de los epítomos dentro de una proteína sintética. Alharbi et al, (2024) emplearon el enlazador GPGPG para ensamblar epítomos antigénicos, no alergénicos y no tóxicos en un constructo multiepítopo contra ZIKV, además de utilizar EAAAK para la unión de un adyuvante.(21). Aunque en aplicaciones diagnósticas no necesariamente se requiere un adyuvante, la selección de enlazadores continúa siendo relevante porque la disposición espacial de las regiones antigénicas puede influir en su reconocimiento por anticuerpos. Esta consideración también ha sido abordada en estudios de vacunología estructural contra ZIKV, como el de Hakami M (2024), donde la organización del constructo, el modelado tridimensional y la estabilidad molecular fueron evaluados como criterios de validación computacional (22)

El potencial diagnóstico de las proteínas multiepitópicas ha sido respaldado por investigaciones orientadas a discriminar infecciones por ZIKV y otros flavivirus. Pereira et al, (2024) diseñaron proteínas multiepitópicas dirigidas a diferenciar entre infecciones por virus Zika y dengue, demostrando una mejora significativa en la especificidad serológica. Este resultado resulta especialmente relevante en contextos endémicos, donde la cocirculación de flavivirus representa una limitante importante para la interpretación de pruebas convencionales (23). De manera complementaria,

Valotto et al, (2026) identificaron mediante mapeo computacional regiones diferenciales en la proteína NS1 entre ZIKV y dengue, lo que refuerza la importancia de seleccionar epítomos específicos como estrategia para optimizar el rendimiento diagnóstico.(24)

En este tipo de aproximaciones, el diseño del constructo debe considerar varios criterios de aceptación. En primer lugar, los epítomos seleccionados deben presentar antigenicidad favorable, lo que sugiere capacidad potencial de reconocimiento inmunológico; este criterio ha sido empleado en estudios inmunoinformáticos para priorizar candidatos multiepítomo frente al ZIKV (7,20).

En segundo lugar, deben ser clasificados como no alergénicos y no tóxicos por herramientas predictivas, reduciendo riesgos en etapas posteriores de expresión y evaluación, tal como se ha considerado en diseños multiepítomo previos (21,22) En tercer lugar, el constructo completo debe presentar propiedades fisicoquímicas y estructurales compatibles con expresión recombinante, estabilidad y manipulación experimental, debido a que la secuencia lineal de epítomos no garantiza por sí sola una adecuada exposición antigénica (22) Finalmente, debe evaluarse su similitud frente a proteínas de virus relacionados, con el fin de reducir la probabilidad de reactividad cruzada, aspecto crítico en el diagnóstico serológico de flavivirus (14,15).

Desde la perspectiva del laboratorio clínico, un antígeno multiepítomo específico podría contribuir al desarrollo de inmunoensayos más precisos, especialmente en regiones donde el diagnóstico diferencial de arbovirosis continúa siendo complejo. Sin embargo, es necesario enfatizar que el diseño computacional constituye una fase preliminar, por lo que la utilidad real de un constructo diagnóstico solo puede establecerse mediante validación experimental. Esta validación deberá incluir expresión recombinante, purificación del antígeno, evaluación con sueros de pacientes confirmados, controles negativos y muestras de individuos infectados por otros arbovirus, además del análisis de sensibilidad, especificidad, valor predictivo positivo, valor predictivo negativo y curvas ROC. Por tanto, el presente enfoque permite proponer un candidato antigénico con potencial inmunodiagnóstico, pero no sustituye la validación clínica posterior, especialmente considerando las limitaciones de especificidad descritas en pruebas serológicas convencionales para ZIKV y otros flavivirus (14,15,23).

1.1.6. Modelado estructural y validación computacional

El modelado estructural constituye una etapa fundamental en el diseño de proteínas sintéticas, debido a que permite evaluar la conformación tridimensional del constructo y la posible exposición de los epítomos seleccionados. En el caso de proteínas multiepítomo, la secuencia primaria no garantiza por sí sola que las regiones antigénicas queden accesibles para el reconocimiento por anticuerpos, por lo que la evaluación estructural se considera un componente necesario dentro del diseño racional de constructos proteicos (22). Por ello, es necesario analizar la estructura tridimensional, la calidad conformacional y la estabilidad teórica del modelo antes de avanzar hacia etapas experimentales, especialmente cuando el constructo se proyecta para aplicaciones inmunológicas o diagnósticas (20,25).

El desarrollo de herramientas basadas en inteligencia artificial ha mejorado de manera significativa la predicción estructural de proteínas. Jumper et al, (2021) introdujeron AlphaFold, un sistema capaz de predecir estructuras tridimensionales con alta precisión en numerosos contextos proteicos (26). Esta herramienta permite evaluar la disposición espacial de constructos diseñados *in silico*, identificar regiones con mayor o menor confianza estructural y orientar la selección o rediseño de secuencias, lo cual ha sido aprovechado en estudios de diseño multiepítomo contra ZIKV (22). En el contexto de un constructo multiepítomo, la predicción estructural resulta crucial porque la accesibilidad de los epítomos puede influir directamente en la sensibilidad y especificidad potencial de un inmunoensayo, particularmente cuando se busca reducir problemas de reconocimiento cruzado asociados a regiones antigénicas conservadas (14,15).

Además del modelado, la validación estructural permite estimar la calidad del modelo generando. Herramientas como gráficos de Ramachandran y ERRAT son utilizadas para analizar la distribución angular de residuos, la calidad de interacciones no enlazantes y la compatibilidad entre la secuencia y el entorno tridimensional predicho, criterios que se integran dentro de los análisis estructurales aplicados al diseño computacional de proteínas multiepítomo (22). Estos análisis no demuestran funcionalidad experimental, pero permiten identificar errores estructurales relevantes y seleccionar modelos con mayor coherencia conformacional antes de una fase experimental (20). En consecuencia, la validación computacional aporta criterios

objetivos para decidir si un constructo puede continuar hacia fases de optimización génica, expresión recombinante y evaluación inmunológica, manteniendo la distinción entre predicción *in silico* y validación biológica posterior (25).

Leal et al, (2025) desarrollaron una investigación orientada al diseño y evaluación de una vacuna multiepítopo contra el Virus del Zika, basada en las proteínas E y NS1 del ZIKV mediante enfoques *in silico* e *in vivo*, empleando herramientas inmunoinformáticas y validaciones experimentales en ratones BALB/c. Los autores reportaron que los constructos presentaron estabilidad estructural, ausencia de toxicidad y elevada antigenicidad, además de inducir respuestas humorales y celulares significativas (25). Estos resultados demuestran que un diseño computacional adecuadamente filtrado puede conservar funcionalidad en etapas experimentales, aunque la extrapolación hacia aplicaciones diagnósticas requiere validaciones específicas orientadas a sensibilidad, especificidad y reconocimiento serológico diferencial (23,25).

De manera similar, Hakami M, (2024) desarrolló un diseño de vacuna multiepítopo mediante inmunoinformática y vacunología estructural contra el ZIKV, integrando epítomos B y T con análisis de docking molecular y simulaciones dinámicas. Los resultados evidenciaron estabilidad, inmunogenicidad y capacidad de interacción con receptores inmunológicos, lo que respalda la utilidad de los análisis estructurales en la evaluación preliminar de constructos proteicos diseñados *in silico* (22). Aunque este tipo de estudios se orienta principalmente a vacunas, sus procedimientos estructurales son relevantes para el diseño de antígenos diagnósticos, ya que permiten evaluar si las regiones seleccionadas se mantienen expuestas, si el constructo presenta una arquitectura plausible y si sus propiedades computacionales justifican avanzar hacia validación experimental (22,23).

En el presente contexto, el modelado estructural y la validación computacional permiten sustentar la selección del constructo multiepítopo propuesto, debido a que aportan información sobre su conformación, estabilidad teórica y posible exposición de regiones antigénicas (20,22). No obstante, los resultados obtenidos mediante estas herramientas deben interpretarse como evidencia predictiva. La estabilidad estructural, la antigenicidad y la ausencia de toxicidad estimadas por servidores bioinformáticos no equivalen a desempeño diagnóstico demostrado, por lo que el constructo deberá

validarse posteriormente mediante expresión recombinante y evaluación inmunológica frente a muestras o paneles serológicos caracterizados (23,25) Por tanto, la fase computacional debe considerarse como una etapa de priorización racional, cuyo propósito es disminuir el número de candidatos y aumentar la probabilidad de éxito en la fase experimental.

1.1.7. Optimización génica para expresión heteróloga

La optimización génica constituye una etapa complementaria en el diseño de proteínas recombinantes, especialmente cuando el constructo sintético se proyecta para expresión heteróloga en sistemas bacterianos como *Escherichia coli*. En estudios previos de diseño multiepítopo contra ZIKV, la construcción computacional del candidato proteico se plantea como una fase previa a su posible evaluación experimental, por lo que la adaptación del diseño hacia un sistema de expresión representa un componente necesario para su proyección biotecnológica (20,21). Aunque el diseño inicial se realiza a nivel proteico, la producción experimental requiere una secuencia codificante compatible con el organismo hospedero. Para ello, se emplean herramientas de retrotraducción y optimización de codones, las cuales permiten adaptar la secuencia nucleotídica al uso preferencial de codones del sistema de expresión seleccionado, como parte del tránsito entre el diseño *in silico* y la validación experimental posterior (22,25).

Uno de los parámetros más utilizados en este proceso es el índice de adaptación de codones (CAI), el cual permite estimar la compatibilidad de la secuencia optimizada con el repertorio de codones preferidos por el hospedero. De igual manera, el contenido de guanina-citosina (GC) debe mantenerse dentro de un rango adecuado, ya que valores extremos pueden afectar la estabilidad del ADN, la transcripción, la traducción o la eficiencia de síntesis. En constructos sintéticos, también es importante evitar secuencias problemáticas, como sitios de restricción no deseados, motivos repetitivos, señales de terminación prematura o estructuras secundarias altamente estables en el ARNm que puedan interferir con la traducción. Estos criterios son coherentes con los enfoques inmunoinformáticos aplicados al diseño racional de proteínas multiepítopo, en los cuales no solo se seleccionan epítomos por antigenicidad,

alergenicidad y toxicidad, sino que también se proyecta la factibilidad estructural y experimental del constructo diseñado (7,21,22).

La evaluación del plegamiento secundario del ARNm permite estimar la estabilidad teórica de la secuencia optimizada y detectar posibles estructuras que podrían dificultar el inicio de la traducción o reducir la eficiencia de expresión. Aunque estos análisis no garantizan producción proteica exitosa, proporcionan información útil para ajustar el diseño antes de la síntesis génica. En este sentido, la optimización génica debe considerarse una fase de preparación para la validación experimental, en la cual el constructo podrá ser clonado, expresado, purificado y evaluado mediante inmunoensayos. Esta distinción es necesaria porque los resultados computacionales permiten priorizar candidatos, pero no sustituyen la comprobación experimental de antigenicidad, estabilidad, expresión y reconocimiento inmunológico (23,25).

Para aplicaciones diagnósticas, la expresión heteróloga en *Escherichia coli* representa una alternativa viable debido a su bajo costo, disponibilidad, facilidad de manipulación y amplia experiencia en producción de proteínas recombinantes. Sin embargo, este sistema también presenta limitaciones, como la posible formación de cuerpos de inclusión, ausencia de modificaciones postraduccionales e interferencia de la estructura del constructo con la solubilidad. Por ello, el diseño debe incorporar criterios de estabilidad, solubilidad y facilidad de purificación, así como etiquetas que permitan recuperar la proteína sin comprometer los epítomos diagnósticos seleccionados. La necesidad de considerar estos aspectos se relaciona con la evidencia de que los constructos multiepítopo deben evaluarse no solo por su antigenicidad, sino también por su estabilidad estructural, ausencia de toxicidad y viabilidad para fases posteriores de validación (20,22,25).

En conjunto, la optimización génica permite vincular el diseño computacional con una posible fase experimental futura. En el caso de un constructo multiepítopo del ZIKV, la adaptación de codones para *Escherichia coli*, el análisis del contenido GC, la evaluación de la estructura secundaria del ARNm y el diseño plasmídico *in silico* constituyen pasos necesarios para proyectar la expresión recombinante. No obstante, la viabilidad final del constructo dependerá de su síntesis, clonación, expresión, purificación y validación inmunológica frente a muestras clínicas o paneles serológicos caracterizados. Esta validación posterior es indispensable porque el

desempeño diagnóstico no puede inferirse únicamente desde predicciones computacionales, sino que debe demostrarse mediante evaluación serológica, especialmente en contextos donde la reactividad cruzada puede afectar la especificidad de los inmunoensayos (14,15,23).

De esta manera, los antecedentes revisados evidencian que la integración de inmunoinformática, diseño de proteínas multiepítopo, modelado estructural y optimización génica constituye una estrategia prometedora para superar limitaciones de los inmunoensayos convencionales. Esta aproximación resulta pertinente en contextos de cocirculación viral, donde la especificidad diagnóstica es un criterio esencial, debido a los problemas de reconocimiento cruzado descritos en pruebas serológicas frente a flavivirus relacionados (14,15). Por tanto, la presente investigación se orienta al diseño *in silico* de un antígeno multiepítopo del virus Zika con potencial aplicación en inmunodiagnóstico diferencial, considerando criterios predictivos de antigenicidad, seguridad, estabilidad estructural, baja similitud con virus clínicamente relacionados y factibilidad de expresión heteróloga, en concordancia con enfoques previos de diseño multiepítopo e inmunoinformática aplicada al ZIKV (20–23).

1.2. OBJETIVOS

1.2.1. Objetivo general

Diseñar y validar *in silico* una proteína multiepitopos específica con potencial inmunodiagnóstico diferencial, mediante herramientas de inmunoinformática y biología computacional.

1.2.2. Objetivos específicos

- Identificar y seleccionar regiones antigénicas conservadas del virus Zika mediante análisis inmunoinformático, evaluando su especificidad frente a Oropouche y Chikungunya para descartar homología significativa y reducir reactividad cruzada.
- Predecir y filtrar epítomos B lineales con alto potencial antigénico, asegurando ausencia de alergenidad y toxicidad mediante herramientas bioinformáticas especializadas.
- Diseñar y validar *in silico* un constructo proteico multiepitopos optimizado, evaluando sus propiedades fisicoquímicas, estructurales tridimensionales y su adaptación génica para expresión heteróloga en *Escherichia coli*.

CAPÍTULO II.

METODOLOGÍA.

2.1. MATERIALES

El desarrollo del presente estudio requiere fundamentalmente recursos humanos, tecnológicos y de infraestructura digital, dado que se trata de una investigación de tipo computacional (*in silico*) que no involucra experimentación biológica directa ni manipulación de muestras clínicas.

2.1.1. Recursos humanos

El equipo estuvo conformado por la tesista responsable del diseño y ejecución de los análisis bioinformáticos, bajo la supervisión académica de un tutor con experiencia en microbiología, biología molecular y biotecnología aplicada. El acompañamiento metodológico garantizó la correcta aplicación de herramientas inmunoinformáticas, análisis estructural y optimización genética.

2.1.2. Recursos tecnológicos

Se requirió un equipo de cómputo con procesador de capacidad media o superior, memoria RAM mínima de 8 GB y conexión estable a internet de banda ancha. La investigación utilizó principalmente plataformas web de acceso libre, entre ellas NCBI, UniProt, ABCPred, VaxiJen, AllerTOP, ToxinPred2, BLASTp, ProtParam, ColabFold, SAVES (Ramachandran Plot, ERRAT), JCat, RNAfold y NEBcutter.

Asimismo, se emplearon programas de diseño plasmídico como Benchling o SnapGene Viewer para la construcción *in silico* del vector recombinante. El uso de estas herramientas no implica costos de licencia, ya que la mayoría dispone de versiones académicas gratuitas.

Para garantizar la reproducibilidad del estudio, se registraron las fechas de consulta de las bases de datos, los códigos de acceso de las secuencias utilizadas, las versiones de

los servidores bioinformáticos cuando estuvieron disponibles, los parámetros de análisis empleados y los archivos de salida generados en cada fase. Los archivos FASTA de entrada, resultados de BLAST, tablas de epítomos, modelos tridimensionales en formato PDB, reportes de validación estructural, secuencias optimizadas y mapas plasmídicos deberán conservarse como anexos digitales del trabajo de titulación.

2.1.3. Recursos institucionales

La investigación contó con el respaldo de la Facultad de Ciencias de la Salud y la Carrera de Laboratorio Clínico, incluyendo acceso a biblioteca virtual institucional, bases de datos científicas suscritas y espacios físicos para reuniones académicas.

2.1.4. Recursos económicos

Al tratarse de un estudio computacional, no se requieren reactivos de laboratorio ni insumos clínicos. Los costos se limitaron a servicios básicos (electricidad, conectividad) y eventual impresión o encuadernación de documentos finales, lo que convierte a esta investigación en viable y de bajo costo operativo.

2.2. MÉTODOS

2.2.1. Tipo y diseño de Estudio/Enfoque de la investigación

El presente estudio se desarrolló bajo un enfoque cuantitativo, de tipo descriptivo y predictivo mediante un diseño computacional (*in silico*), orientado al análisis inmunoinformático y al diseño racional de una proteína multiepitopos del virus Zika con potencial aplicación en inmunodiagnóstico serológico específico. La metodología seguirá un flujo secuencial estructurado que garantice reproducibilidad y trazabilidad de cada etapa del proceso.

2.2.2. Aspectos éticos

Al tratarse de una investigación que posee un enfoque estrictamente computacional (*in silico*), no involucró participantes humanos directos, manipulación de muestras biológicas, ni experimentación animal. En otras palabras, no se realizaron intervenciones físicas, psicológicas ni sociales sobre ningún individuo.

Las secuencias proteicas utilizadas provienen de bases de datos públicas de acceso abierto (NCBI GenBank, UniProt), depositadas conforme a las normativas internacionales de libre acceso a datos científicos. No se accedió ni procesó información personal identificable de individuos.

En concordancia con los principios de beneficencia y no maleficencia, la investigación se orientó al beneficio potencial de poblaciones en riesgo en zonas endémicas tropicales. Por la naturaleza computacional del estudio, no se requirió ningún formulario de consentimiento informado de participantes. Los resultados generados respetan la propiedad intelectual de las bases de datos y publicaciones científicas consultadas, citando en formato Vancouver todos los recursos utilizados.

2.2.3. Variables de la investigación

Variable Dependiente

Antigenicidad, alergenicidad, toxicidad, estabilidad fisicoquímica, calidad estructural tridimensional, especificidad frente a secuencias homólogas y adaptación génica para expresión heteróloga en *Escherichia coli*.

Variable Independiente

Selección y combinación de epítomos B lineales antigénicos del virus Zika.

La optimización se evidencia en la **Tabla 1**

Tabla 1: Optimización de variables.

Variab les	Definición conceptual	Definición operacional	Dimensiones	Indicadores	Técnica e instrumento
VARIABLE INDEPENDIENTE					
Herramientas de inmunoinformática y biología computacional	Conjunto de herramientas computacionales utilizadas para el análisis de secuencias biológicas y la predicción de propiedades inmunológicas, estructurales y funcionales de proteínas.	Aplicación de plataformas bioinformáticas para la identificación de regiones antigénicas, predicción y selección de epítomos B lineales, evaluación inmunológica, modelado estructural y optimización génica.	Identificación antigénica	Regiones conservadas identificadas en ZIKV	Técnica: análisis bioinformático. Instrumentos: GenBank y BLASTx.
			Predicción de epítomos	Número de epítomos B lineales seleccionados	Técnica: predicción inmunoinformática. Instrumento: ABCPred.
			Evaluación inmunológica	Antigenicidad, alergenicidad y toxicidad	Técnica: análisis inmunoinformático. Instrumentos: VaxiJen v2.0, AllerTOP v2.1, AllergenFP, AlgPred y ToxinPred.
			Modelado y validación estructural	Calidad estructural y estabilidad	Técnica: modelado molecular. Instrumentos:

Proteína multiepítipo específica con potencial inmunodiagnóstico diferencial				d del modelo	PSIPRED, SWISS-MODEL, ERRAT y Ramachandran Plot.
				Optimización génica	CAI y contenido GC Técnica: optimización computacional. Instrumento: JCat.
	VARIABLE DEPENDIENTE				
			Antigenicidad	Puntuación obtenida en VaxiJen v2.0	Técnica: análisis inmunoinformático. Instrumento: VaxiJen v2.0.
	Constructo proteico conformado por múltiples epítomos seleccionados por sus propiedades inmunológicas y especificidad frente al virus Zika.	Antígeno multiepítipo diseñado y validado <i>in silico</i> a partir de epítomos B lineales seleccionados mediante herramientas inmunoinformáticas.	Seguridad biológica	Predicción de alergenidad y toxicidad	Técnica: análisis inmunoinformático. Instrumentos: AllerTOP v2.1, AllergenFP, AlgPred y ToxinPred.
			Propiedades fisicoquímicas	Peso molecular, pI, índice de inestabilidad y GRAVY	Técnica: análisis fisicoquímico. Instrumento: ProtParam.
			Calidad estructural	Factor ERRAT y porcentaje de residuos en regiones favorecidas del gráfico de Ramachandran	Técnica: validación estructural. Instrumentos: ERRAT y Ramachandran Plot.

	Porcentaje de similitud con CHIKV y OROV	Técnica: alineamiento de secuencias. Instrumento: BLASTp.
Especificidad molecular	Cumplimiento de criterios de antigenicidad, estabilidad, seguridad biológica y especificidad	Técnica: integración y análisis de resultados bioinformáticos.
Potencial inmunodiagnóstico diferencial		

Elaborado por: Balseca Quishpe, Carmen Elena

2.2.4. Hipótesis de investigación

Hipótesis alternativa (H1)

El análisis inmunoinformático permite diseñar un constructo proteico multiepítipo del virus Zika con antigenicidad predictiva favorable, ausencia de alergenidad y toxicidad, estabilidad estructural aceptable y baja similitud con regiones antigénicas de virus clínicamente relacionados, lo que respalda su potencial aplicación en inmunodiagnóstico diferencial.

Hipótesis nula (H0)

El análisis inmunoinformático no permite obtener un constructo proteico multiepítipo del virus Zika que cumpla simultáneamente criterios predictivos de antigenicidad, seguridad, estabilidad estructural y especificidad molecular para su potencial aplicación inmunodiagnóstica.

2.2.5. Descripción de Procedimientos

1. Recuperación y curaduría de secuencias

En la primera fase se realizó una búsqueda sistemática de secuencias del virus Zika en la base de datos GenBank del del National Center for Biotechnology Information (NCBI) (<https://www.ncbi.nlm.nih.gov/genbank/>) con fecha de consulta y descarga 18 de marzo del 2026. Se seleccionó la secuencia nucleotídica completa con código de acceso MF438286.1, correspondiente a "Zika virus strain Zika virus/Homo sapiens/Cuba 2017 polyprotein gene, complete cds". A partir del CDS anotado en el registro de GenBank, se recuperó la secuencia aminoacídica traducida de la poliproteína completa, la cual fue descargada en formato FASTA desde la base de datos NCBI Protein, correspondiente al código de acceso proteico ALL27019. Esta información debe mantenerse de forma uniforme en metodología y resultados, debido a que en el manuscrito ya se utiliza dicho código para la evaluación de la proteína completa.

La identidad y completitud del registro seleccionado fueron verificadas mediante BLASTx en el servidor NCBI BLAST (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>), considerando como parámetros principales la cobertura de consulta, el valor E, el porcentaje de identidad y la puntuación total del alineamiento. El análisis permitió confirmar una cobertura de consulta del 100 %, un valor E de 0,0, un porcentaje de identidad de 99,47 % y una puntuación máxima y total de 6802 frente a secuencias proteicas de referencia del ZIKV depositadas en NCBI. Estos criterios respaldaron la selección de la secuencia para los análisis inmunoinformáticos posteriores.

Para la evaluación de especificidad frente a virus clínicamente relacionados, se consideraron secuencias proteicas de referencia de Oropouche virus (OROV) y Chikungunya virus (CHIKV). Las proteínas incluidas para comparación fueron la UYI36405, para OROV mientras que para E1 se seleccionó la secuencia proteica NP_690589.2, correspondiente a structural polyprotein [Chikungunya virus], mientras que para E2 se seleccionó la secuencia AHK22766.1, correspondiente a E2, partial [Chikungunya virus].

De manera general en la **Figura 1**, se evidencia el flujo de trabajo establecido.

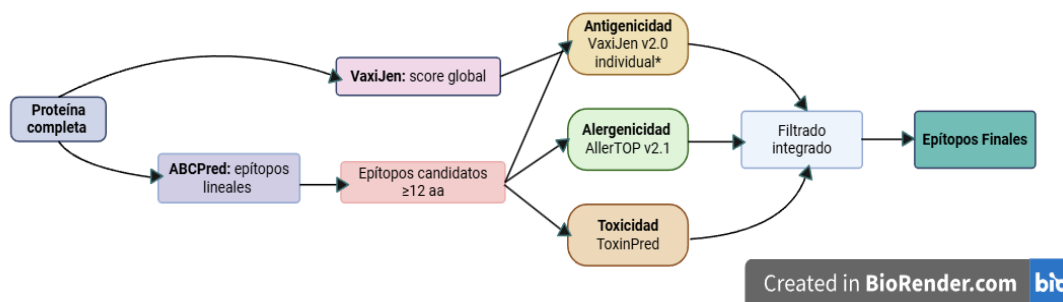


Figura 1: Flujo de trabajo.

Imagen elaborada en BioRender por Balseca Quishpe, Carmen Elena

2. Predicción inicial de epítomos B lineales

Para la selección inicial de epítomos B lineales se utilizó el servidor ABCpred (<http://crdd.osdd.net/raghava/abcpred/> , consultado el 20 de marzo de 2026), herramienta bioinformática basada en redes neuronales recurrentes, empleada para la predicción de epítomos lineales de células B. Debido a que la secuencia aminoacídica descargada en formato FASTA presentó una longitud total de 3425 aminoácidos (aa), esta fue dividida en dos fragmentos para facilitar su procesamiento en los servidores utilizados. El primer fragmento estuvo conformado por 1750 aa, mientras que el segundo comprendió 1675 aa, manteniendo la continuidad de la poliproteína viral y evitando la pérdida de regiones potencialmente antigénicas durante el análisis.

En ABCPred se emplearon como parámetros una longitud peptídica de 16 aa y un umbral de predicción de 0,51, de acuerdo con los resultados reportados en la sección de resultados para las Tablas 3 y 4. Se seleccionaron los 10 epítomos con mayor puntuación de cada fragmento, obteniéndose un total inicial de 20 epítomos lineales de 16 aa. Se aceptaron para la fase de filtrado aquellos péptidos con puntuación igual o superior al umbral establecido y se rechazaron aquellos con puntuación inferior, secuencias duplicadas o regiones que no pudieran ser trazadas dentro de la poliproteína original.

3. Corte de epítomos de 16 aa a 12 aa y justificación metodológica

Cada epítomo inicial de 16 aa fue sometido a un procedimiento de segmentación para obtener dos péptidos derivados de 12 aa: uno correspondiente a los 12 aminoácidos

iniciales y otro correspondiente a los 12 aminoácidos finales. De este modo, a partir de los 20 epítomos iniciales se generaron 40 péptidos lineales de 12 aa para análisis posterior. Este procedimiento se aplicó con el propósito de conservar regiones centrales o terminales con posible capacidad antigénica, reducir la longitud de las secuencias candidatas y facilitar su análisis individual en los servidores de antigenicidad, alergenidad y toxicidad.

La selección de péptidos de 12 aa se justificó porque los epítomos B lineales suelen corresponder a regiones peptídicas cortas reconocibles por anticuerpos, y porque el uso de fragmentos más reducidos permite discriminar con mayor precisión cuáles regiones dentro del epítomo original aportan mayor antigenicidad predictiva. Además, esta segmentación permitió incrementar la resolución del filtrado, al comparar si la región inicial o final de cada péptido de 16 aa presentaba mejores características computacionales para su incorporación en el constructo multiepítomo.

Como criterio de aceptación, solo se conservaron los péptidos de 12 aa que cumplieron simultáneamente con los parámetros de antigenicidad, ausencia de alergenidad y ausencia de toxicidad. Se rechazaron los péptidos clasificados como probables no antígenos, alérgicos según el criterio de consenso definido, tóxicos o con inconsistencias de secuencia, tales como caracteres adicionales, signos de puntuación o residuos no estándar.

4. Evaluación de antigenicidad, alergenidad y toxicidad

Los epítomos seleccionados fueron sometidos a un proceso de filtrado secuencial mediante herramientas bioinformáticas especializadas. La antigenicidad se evaluó con VaxiJen v2.0 (<https://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html> , consultado el 18 de marzo de 2026), empleando como modelo de análisis virus y un umbral de aceptación de 0,5. Se consideraron aceptables los péptidos clasificados como “antígeno probable” o aquellos con puntuación igual o superior a 0,5. Se rechazaron los péptidos con puntuación inferior al umbral o clasificados como “probable no antígeno”.

La alergenidad se evaluó mediante AllerTOP v2.1 (https://www.ddg-pharmfac.net/allertop_test/ , consultado el 01 de abril de 2026), AllergenFP v1.1 (<https://ddg-pharmfac.net/AllergenFP/> , consultado el 01 de abril de 2026) y AlgPred

(<https://webs.iitd.edu.in/raghava/algpred/submission.html> , consultado el 01 de abril de 2026), según la longitud permitida por cada herramienta. Para los epítomos de 16 aa, se utilizaron las tres herramientas. Para los péptidos de 12 aa, se emplearon AllerTOP v2.1 y AlgPred, debido a que AllergenFP no permite el análisis de secuencias menores a 16 aa, como se describe en el manuscrito.

Como criterio de consenso para alergenicidad, se aceptaron únicamente aquellos péptidos clasificados como no alergénicos o probablemente no alergénicos en al menos dos herramientas, cuando las tres herramientas fueron aplicables. En el caso de péptidos de 12 aa, al utilizarse dos herramientas, se aceptaron aquellos clasificados como no alergénicos o probablemente no alergénicos en ambas herramientas, o aquellos que, en caso de discrepancia, presentaron simultáneamente antigenicidad elevada y ausencia de toxicidad, quedando marcados para revisión manual. Se rechazaron los péptidos clasificados como alergénicos por la mayoría de herramientas aplicables.

La toxicidad se evaluó mediante ToxinPred (https://webs.iitd.edu.in/raghava/toxinpred/prot_test.php , consultado el 02 de abril de 2026). Para el análisis de la poliproteína completa, la secuencia fue dividida en fragmentos de 500 aa, debido a la limitación de longitud del servidor. En el análisis de epítomos, se empleó un umbral de 0,4 y longitud peptídica de 16 aa, de acuerdo con los parámetros descritos en el manuscrito. Se aceptaron únicamente los péptidos clasificados como no tóxicos y se rechazaron aquellos con predicción de toxicidad positiva.

Finalmente, se seleccionaron como candidatos para el ensamblaje del constructo los epítomos de 12 aa que cumplieron simultáneamente los siguientes criterios: antigenicidad igual o superior a 0,5 en VaxiJen v2.0, clasificación no alergénica según el criterio de consenso establecido y ausencia de toxicidad en ToxinPred. Este criterio permitió reducir el conjunto inicial de péptidos y seleccionar únicamente aquellos con mejores características predictivas para el diseño del constructo multiepítomo.

5. Diseño del constructo multiepítomo

Para el diseño del constructo multiepítomo se fusionaron los epítomos que superaron los filtros de antigenicidad, alergenicidad y toxicidad. Los epítomos seleccionados

fueron ensamblados mediante el enlazador peptídico flexible GPGPG, con el propósito de favorecer la separación estructural entre regiones antigénicas y preservar la exposición individual de cada epítipo. La organización de los péptidos dentro del constructo se estableció considerando la puntuación de antigenicidad obtenida en VaxiJen v2.0, ubicando los epítopos con mayor puntuación en posiciones centrales y los de menor puntuación hacia los extremos, según el criterio metodológico descrito en el manuscrito.

Al inicio de la secuencia se incorporó una metionina (M) como codón de inicio para una eventual expresión recombinante, y al extremo C-terminal se añadió una cola de seis histidinas (HHHHHH) como etiqueta de purificación. Una vez ensamblado, el constructo completo fue reevaluado mediante VaxiJen v2.0, AllerTOP v2.1 y ToxinPred, con el fin de comprobar que la proteína sintética conservara antigenicidad predictiva, ausencia de alergenicidad y ausencia de toxicidad después de la fusión de los epítopos.

Como criterios de aceptación del constructo final, se establecieron los siguientes: antigenicidad global igual o superior a 0,5, clasificación como no alergénico o probablemente no alergénico, ausencia de toxicidad, ausencia de residuos no estándar, conservación de los epítopos seleccionados y presencia correcta de los enlazadores GPGPG y de la etiqueta de histidinas. En caso de que el constructo no cumpliera alguno de estos criterios, se consideraría necesario reorganizar el orden de los epítopos, retirar secuencias problemáticas o reevaluar el diseño.

6. Análisis fisicoquímico

La secuencia resultante del constructo multiepítipo fue evaluada mediante ProtParam, disponible en el servidor ExPASy (<https://web.expasy.org/protparam/> , consultado el 24 de abril de 2026). Se analizaron parámetros fisicoquímicos como número de aminoácidos, peso molecular, punto isoelectrico teórico, composición aminoacídica, índice de inestabilidad, índice alifático y promedio general de hidropaticidad (GRAVY). Estos indicadores permitieron estimar de forma preliminar la estabilidad, comportamiento hidrofóbico/hidrofílico y factibilidad de manipulación del constructo en una fase experimental futura.

Como criterio de aceptación, se consideró favorable un índice de inestabilidad menor de 40, indicativo de estabilidad teórica. Asimismo, se interpretó el valor GRAVY para estimar la tendencia hidrofílica o hidrofóbica del constructo. Los resultados de ProtParam fueron registrados para su posterior comparación con el modelado estructural y la optimización génica.

7. Modelado estructural y validación computacional

La estructura tridimensional del constructo multiepítipo se predijo mediante ColabFold v1.6.1: AlphaFold2 using MMseqs2, disponible en entorno Google Colab (<https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb> , consultado el 29 de abril de 2026). Para el modelado se utilizó como entrada la secuencia aminoacídica completa del constructo en formato FASTA. Se mantuvieron los parámetros predeterminados del servidor, salvo en aquellos campos requeridos para el ingreso de la secuencia, y se registraron los archivos de salida generados, incluyendo el modelo tridimensional en formato PDB y las métricas de confianza estructural disponibles.

8. Refinamiento

Los modelos tridimensionales seleccionados fueron refinados mediante GalaxyWEB/GalaxyRefine, (consultado y descargado el 21 de mayo de 2026) con el propósito de mejorar la calidad estructural antes de su validación. El servidor generó cinco modelos refinados para cada constructo.

9. Validación estructural

Posteriormente, la calidad estructural fue evaluada mediante el servidor SAVES (<https://saves.mbi.ucla.edu/> , consultado el 21 de mayo de 2026), utilizando las herramientas Ramachandran Plot, y ERRAT.

Como criterios de aceptación estructural, se consideró que el modelo debía presentar una proporción mayoritaria de residuos en regiones favorecidas o permitidas del gráfico de Ramachandran, valores de ERRAT compatibles con una calidad estructural aceptable. En caso de observarse valores estructurales deficientes, se consideraría necesario modificar el orden de los epítipos, revisar los enlazadores o generar un nuevo modelo.

10. Optimización génica

La secuencia proteica final del constructo multiepítipo fue retrotraducida y optimizada para expresión heteróloga en *Escherichia coli* cepa K12 mediante el servidor JCat (<https://www.jcat.de/> , consultado el 22 de mayo de 2026). Para este proceso se consideraron como parámetros principales el índice de adaptación de codones (CAI) y el contenido de guanina-citosina (GC), con la finalidad de obtener una secuencia codificante compatible con el sistema de expresión bacteriano. Se registraron como resultados principales el valor de CAI, el porcentaje de GC y la secuencia nucleotídica optimizada.

Como criterios de aceptación, se consideró favorable un CAI cercano a 1,0 y un contenido GC dentro de un rango compatible con expresión en *E. coli*. Además, se verificó la ausencia de elementos que pudieran interferir con la expresión o clonación, como sitios de restricción no deseados, codones de terminación prematura, motivos repetitivos extensos o secuencias que pudieran generar inestabilidad.

11. Evaluación de la estructura secundaria del ARNm

Posteriormente, la secuencia nucleotídica optimizada fue evaluada mediante RNAfold (<http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi> , , consultado el 22 de mayo de 2026), con el propósito de analizar el plegamiento secundario teórico del ARNm y estimar su energía libre mínima. Se registró la estructura secundaria predicha y el valor de energía libre mínima en kcal/mol. Este análisis permitió identificar posibles estructuras secundarias estables que pudieran interferir con la traducción del constructo.

12. Análisis de similitud frente a CHIKV y OROV

De manera complementaria, el constructo proteico diseñado fue sometido a análisis de similitud mediante BLASTp, debido a que la comparación se realizó entre secuencias aminoacídicas. Este análisis permitió contrastar el constructo multiepítipo frente a secuencias proteicas de ZIKV. Mientras que para las comparaciones pareadas frente a proteínas específicas se empleó Global Aligner del servidor BLAST. Para frente a proteínas seleccionadas de virus clínicamente relacionados, como OROV y CHIKV. Para OROV se utilizó la proteína con código de acceso UYI36405 mientras que en CHIKV en E1 se seleccionó la secuencia proteica NP_690589.2, correspondiente a

structural polyprotein [Chikungunya virus], mientras que para E2 se seleccionó la secuencia AHK22766.1, correspondiente a E2, parcial [Chikungunya virus].

13. Diseño Plasmídico y Consolidación Final

Finalmente, el gen optimizado obtenido a partir de JCat será insertado *in silico* en un vector de expresión compatible (pET-28a) que fue descargado y consultado el 28 de mayo de 2026 en SnapGene ([https://www.snapgene.com/plasmids/pet_and_duet_vectors_\(novagen\)/pET-28a\(%2B\)](https://www.snapgene.com/plasmids/pet_and_duet_vectors_(novagen)/pET-28a(%2B))). Se empleó el uso de NEBcutter (<https://nc3.neb.com/NEBcutter/> , consultado el 28 de mayo de 2026), para identificar los sitios de restricción y después mediante Benchling (<https://www.benchling.com/> , consultado el 28 de mayo de 2026). Este procedimiento permitirá obtener un candidato antigénico validado computacionalmente y listo para futuras fases experimentales.

CAPÍTULO III.

RESULTADOS Y DISCUSIÓN

3.1 ANÁLISIS DE RESULTADOS

3.1.1. Recuperación y curaduría de secuencias

Como resultado de la búsqueda realizada en la base de datos GenBank del NCBI, se identificó y seleccionó la secuencia nucleotídica completa con código de acceso MF438286.1, correspondiente a *Zika virus strain Zika virus/Homo sapiens/Cuba 2017 polyprotein gene, complete cds*. A partir del CDS anotado en dicho registro, se recuperó la secuencia aminoacídica traducida de la poliproteína completa, correspondiente al código de acceso proteico ALL27019, la cual fue descargada en formato FASTA desde NCBI Protein.

La verificación del registro seleccionado mediante BLASTx evidenció una cobertura de consulta del 100 %, un valor E de 0,0, un porcentaje de identidad de 99,47 % y una puntuación máxima/total de 6802 frente a secuencias proteicas de referencia del virus Zika depositadas en NCBI. Estos resultados se muestran en la **Figura 2**.

Clústeres

Resumen gráfico

Alineaciones

Taxonomía

Agrupaciones que producen alineaciones significativas

Descargar

Seleccione columnas

Espectáculo

100

☒ seleccionar todo

Se seleccionaron 100 clústeres

GenPept

Gráficos

	Composición del clúster	Ancestro del clúster	Secuencia representativa del clúster	Puntuación máxima	Puntuación total	Portada de la consulta	Valor E	Per. Ident	Longitud de acceso	Adhesión
	Haz clic en el icono para ver el contenido del clúster									
☒	800 miembros, 6 organismos	raíz	poliproteína (virus Zika)	6802	6802	100%	0.0	99.47%	3425	ALL27019.1

Figura 2: Verificación del registro MF438286.1 mediante BLASTx frente a secuencias proteicas de referencia del virus Zika depositadas en NCBI. El análisis mostró una cobertura de consulta del 100 %, valor E de 0,0, identidad de 99,47 % y puntuación máxima/total de 6802.

3.1.2. Evaluación de la proteína completa

Previo a la predicción de epítomos B lineales, se evaluó la proteína completa correspondiente al código de acceso ALL27019. Debido a que la secuencia aminoacídica presentó una longitud total de 3425 aa, esta fue dividida en dos

fragmentos para su procesamiento: un primer fragmento de 1750 aa y un segundo fragmento de 1675 aa.

En el primer fragmento de 1750 aa, VaxiJen v2.0 reportó una puntuación de 0,5586, clasificada como antígeno probable. En la evaluación de alergenicidad, AllerTOP v2.1 y AllergenFP v1.1 clasificaron la secuencia como probablemente no alergénica, mientras que AlgPred la clasificó como no alergénica. El análisis de toxicidad mediante ToxinPred no se ejecutó para este fragmento debido a la longitud de la secuencia, conforme se detalla en la **Tabla 2**.

En el segundo fragmento de 1675 aa, VaxiJen v2.0 reportó una puntuación de 0,4785, clasificada como probable no antígeno. AllerTOP v2.1 clasificó este fragmento como probablemente no alergénico, AllergenFP v1.1 como alérgeno probable y AlgPred como no alergénico. Al igual que en el primer fragmento, ToxinPred no se ejecutó sobre la secuencia completa debido a la longitud del fragmento, como se muestra en la **Tabla 3**.

Para cada servidor se aplicaron los umbrales establecidos según los criterios de análisis de antigenicidad, alergenicidad y toxicidad.

Tabla 2: Resultados del análisis antigénico, alergénico y toxico de la poliproteína del Virus Zika en la secuencia con 1750 aa

ANTIGENICIDAD	ALERGENICIDAD			TOXICIDAD
Vaxijen v2.0	AllerTOP v2.1	AllergenFP v1.1	AlgPred	ToxinPred
0,5586 (ANTÍGENO PROBABLE)	Probablemente No Alérgeno	Probablemente No Alérgeno	No alérgeno	Dada la longitud de la secuencia, no se realizó la corrida.

Elaborado por: Balseca Quishpe, Carmen Elena

Tabla 3: Resultados del análisis antigénico, alergénico y toxico de la poliproteína del Virus Zika, en la secuencia con 1675 aa

ANTIGENICIDAD	ALERGENICIDAD			TOXICIDAD
Vaxijen v2.0	AllerTOP v2.1	AllergenFP v1.1	AlgPred	ToxinPred

0,4785 (probable NO ANTÍGENO)	Probablemente NO ALÉRGENO	Alérgeno probable	No alérgeno	Dada la longitud de la secuencia, no se realizó la corrida
--	---------------------------------	----------------------	----------------	--

Elaborado por: Balseca Quishpe, Carmen Elena

3.1.3. Predicciones y selección de epítomos B

La predicción inicial de epítomos B lineales se realizó mediante ABCPred, utilizando una longitud peptídica de 16 aa y un umbral de predicción de 0,51. Se seleccionaron los diez epítomos con mayor puntuación para cada fragmento de la poliproteína, obteniéndose un total de 20 epítomos candidatos de 16 aa.

En el fragmento de 1750 aa, las puntuaciones más altas fueron de 0,94, correspondientes a las secuencias TVEVQYAGTDGPCKVP, AKVEITPNSPRAEATL, HKKGGEARRSRRAVTLP, GEGRLDPYWGDVKQDL y RECTMPPLSFRAKDGC. También se identificaron epítomos con puntuaciones de 0,93, correspondientes a LDEGVPEDDVDCWCNT, HVTKGSALRSGEGRDL y SLVEDDGPPMREILK, así como epítomos con puntuación de 0,92, correspondientes a LWDVPAPKEVKKGETT y YRTQMKGPPWHSEELEI, como se presenta en la **Tabla 4**.

En el fragmento de 1675 aa, la puntuación más alta fue de 0,98, correspondiente a la secuencia EKRVLKPRWMDARVCS. También se identificaron secuencias con puntuaciones de 0,95: GASLIYTVTRNAGLVK, NRWIEENDHMEDKTP y HGSYEAPTQGSASSLI. Las secuencias RGYISTRVEMGEAAAI y AVEAVNDPRFWALVDK presentaron puntuaciones de 0,94; THASAAQRRGRIGRNP obtuvo 0,93; mientras que NSTHEMYWVSGAKSNT, LSMVGDWLEKRPGAFC y VVLIPEPEKQRSPQDN presentaron puntuaciones de 0,91, como se detalla en la **Tabla 5**.

Tabla 4: Epítomos de célula B en la secuencia de 1750 aa. Se coloca los 10 primeros péptidos, siendo estos quienes eran más probables de ser epítomos.

Rango	Secuencia	Posición inicial	Puntaje
1	TVEVQYAGTDGPCKVP	619	0,94
1	AKVEITPNSPRAEATL	457	0,94

1	HKKGEARRSRRAVTLP	207	0,94
1	GEGRLDPYWGDVKQDL	1565	0,94
1	RECTMPPLSFRAKDGC	1110	0,94
2	LDEGVEPDDVDCWCNT	180	0,93
2	HVTKGSALRSGEGRD	1555	0,93
2	SLVEDDGPPMREILK	1459	0,93
3	LWDVPAPKEVKKGETT	1508	0,92
3	YRTQMKGPWHSEELEI	1056	0,92

Elaborado por: Balseca Quishpe, Carmen Elena

Tabla 5: Epítomos de célula B en la secuencia de 1675 aa. Se coloca los 10 primeros péptidos, siendo estos quienes eran más probables de ser epítomos.

Rango	Secuencia	Posición inicial	Puntaje
1	EKRVLKPRWMDARVCS	340	0,98
2	GASLIYTVTRNAGLVK	755	0,95
2	NRVWIEENDHMEDKTP	1588	0,95
2	HGSYEAPTQGSASSLI	1079	0,95
3	RGYISTRVEMGEAAAI	52	0,94
3	AVEAVNDPRFWALVDK	1195	0,94
4	THASAAQRRGRIGRNP	203	0,93
5	NSTHEMYWVSGAKSNT	986	0,91
5	LSMVGDWLEKRPGAFC	937	0,91
5	VVLIPEPEKQRSPQDN	488	0,91

Elaborado por: Balseca Quishpe, Carmen Elena

3.1.4. Evaluación de antigenicidad de epítomos candidatos

Los 20 epítomos iniciales de 16 aa fueron evaluados mediante VaxiJen v2.0. De acuerdo con el umbral establecido de 0,5, diez secuencias fueron clasificadas como antígeno probable: RECTMPPLSFRAKDGC con 1,7589; LSMVGDWLEKRPGAFC con 1,2522; GEGRLDPYWGDVKQDL con 0,9956; KTHAAAQRRGRIGRNP con 0,8598; AKVEITPNSPRAEATL con 0,8044; RGYISTRVEMGEAAAI con 0,7640;

YRTQMKGPWHSEELEI con 0,7325; LDEGVEPDDVDCWCNT con 0,7325; LWDVPAPKEVKKGETT con 0,5919; y NSTHEMYWVSGAKSNT con 0,5441. Las diez secuencias restantes presentaron valores inferiores al umbral y fueron clasificadas como probable no antígeno, según se muestra en la **Tabla 6**.

Posteriormente, cada epítipo de 16 aa fue segmentado en dos péptidos derivados de 12 aa: uno correspondiente a los 12 aa iniciales y otro a los 12 aa finales, generando un total de 40 péptidos de 12 aa. En el grupo de cortes iniciales, las puntuaciones de antigenicidad más altas fueron GEGRLDPYWGDV con 1,4145; RECTMPPLSFRA con 1,3848; LDEGVEPDDVDC con 1,0707; KTHAAAQRRGRI con 0,9092; RGYISTRVEMGE con 0,8607; AKVEITPNSPRA con 0,8183; LSMVGDWLEKRP con 0,7586; TVEVQYAGTDGP con 0,6891; NSTHEMYWVSGA con 0,5820; HVTKGSALRSGE con 0,5320; y YRTQMKGPWHSE con 0,5117. Estos datos se presentan en la **Tabla 7**.

En el grupo de cortes finales, las puntuaciones de antigenicidad más altas correspondieron a MPPLSFRAKDGC con 2,2492; GDWLEKRPGAFC con 1,6716; PAPKEVKKGETT con 1,1609; LDPYWGDVKQDL con 1,0185; AAQRRGRIGRNP con 0,9970; VEPDDVDCWCNT con 0,9272; EARRSRRAVTLP con 0,9162; STRVEMGEAAAI con 0,8373; LKPRWMDARVCS con 0,7801; IEENDHMEDKTP con 0,7340; MKGPWHSEELEI con 0,7084; y EMYWVSGAKSNT con 0,6293. Los resultados completos se muestran en la **Tabla 8**.

Tabla 6: Antigenicidad en 20 secuencias de 16 aa. Se ordenó de la puntuación más alta a la más baja puntuación de antigenicidad, de los epítipos B lineales predichos anteriormente.

SECUENCIA DE 16 AMINOÁCIDOS	SCORE DE ANTIGENICIDAD
RECTMPPLSFRAKDGC	1,7589 (ANTÍGENO PROBABLE)
LSMVGDWLEKRPGAFC	1,2522 (ANTÍGENO PROBABLE)
GEGRLDPYWGDVKQDL	0,9956 (ANTÍGENO PROBABLE)
KTHAAAQRRGRIGRNP	0,8598 (ANTÍGENO PROBABLE)
AKVEITPNSPRAEATL	0,8044 (ANTÍGENO PROBABLE)
RGYISTRVEMGEAAAI	0,7640 (ANTÍGENO PROBABLE)

YRTQMKGPWHSEELEI	0,7325 (ANTÍGENO PROBABLE)
LDEGVEPDDVDCWCNT	0,7325 (ANTÍGENO PROBABLE)
LWDVPAPKEVKKGETT	0,5919 (ANTÍGENO PROBABLE)
NSTHEMYWVSGAKSNT	0,5441 (ANTÍGENO PROBABLE)
AVEAAVNDPRFWALVD	0,4125 (PROBABLE NO ANTÍGENO)
HKKGEARRSRRAVTLP	0,4604 (PROBABLE NO ANTÍGENO)
VVLIPEPEKQRSPQDN	0.1262 (PROBABLE NON-ANTIGEN)
NRVWIEENDHMEDKTP	0,1878 (PROBABLE NO ANTÍGENO)
EKRVLKPRWMDARVCS	0,0670 (PROBABLE NO ANTÍGENO)
HGSYEAPTQGSASSLI	0,0644 (PROBABLE NO ANTÍGENO)
GASLIYTVTRNAGLVK	0,2967 (PROBABLE NO ANTÍGENO)
TVEVQYAGTDGPCKVP	0,2878 (PROBABLE NO ANTÍGENO)
HVTKGSA LRSGEGRD	0,1525 (PROBABLE NO ANTÍGENO)
SLVEDDGPPMREILK	-0,1917 (PROBABLE NO ANTÍGENO)

Elaborado por: Balseca Quishpe, Carmen Elena

Tabla 7: Antigenicidad en 20 secuencias, considerando el corte a partir de 12 aa iniciales.

Secuencia de 16 aminoácidos	Corte de 12 aminoácidos (1-12)	Score de antigenicidad
RECTMPPLSFRAKDGC	RECTMPPLSFRA	1,3848 (ANTÍGENO PROBABLE)
LSMVGDWLEKRPGAFC	LSMVGDWLEKRP	0,7586 (ANTÍGENO PROBABLE)
GEGRLDPYWGDVKQDL	GEGRLDPYWGDV	1,4145 (ANTÍGENO PROBABLE)
KTHAAAQRRGRIGRNP	KTHAAAQRRGRI	0,9092 (ANTÍGENO PROBABLE)
AKVEITPNSPRAEATL	AKVEITPNSPRA	0,8183 (ANTÍGENO PROBABLE)
RGYISTRVEMGEAAAI	RGYISTRVEMGE	0,8607 (ANTÍGENO PROBABLE)
YRTQMKGPWHSEELEI	YRTQMKGPWHSE	0,5117 (ANTÍGENO PROBABLE)
LDEGVEPDDVDCWCNT	LDEGVEPDDVDC	1,0707 (ANTÍGENO PROBABLE)
WDVPAPKEVKKGETT	LWDVPAPKEVKK	0,0696 (probable NO ANTÍGENO)

NSTHEMYWVSGAKSNT	NSTHEMYWVSGA	0,5820 (ANTÍGENO PROBABLE)
AVEAAVNDPRFWALVD	AVEAAVNDPRFW	-0,0004 (probable NO ANTÍGENO)
HKKGEARRSRRAVTL	HKKGEARRSRRA	0,3650 (probable NO ANTÍGENO)
NRVWIEENDHMEDKTP	NRVWIEENDHME	-0,0676 (probable NO ANTÍGENO)
VVLIPEPEKQRSPQDN	VVLIPEPEKQRS	0,2890 (probable NO ANTÍGENO)
EKRVLKPRWMDARVCS	EKRVLKPRWMDA	0,4122 (probable NO ANTÍGENO)
HGSYEAPTQGSASSLI	HGSYEAPTQGSA	0,0666 (probable NO ANTÍGENO)
GASLIYTVTRNAGLVK	GASLIYTVTRNA	0,3955 (probable NO ANTÍGENO)
TVEVQYAGTDGPCKVP	TVEVQYAGTDGP	0,6891 (ANTÍGENO PROBABLE)
HVTKGSALRSGEGRD	HVTKGSALRSGE	0,5320 (ANTÍGENO PROBABLE)
SLVEDDGPPMREILK	SLVEDDGPPMRE	0,1645 (probable NO ANTÍGENO)

Elaborado por: Balseca Quishpe, Carmen Elena

Tabla 8: Antigenicidad en 20 secuencias, considerando el corte de los 12 aa finales.

Secuencia de 16 aminoácidos	Corte de 12 aminoácidos (4-16)	Score de antigenicidad
RECTMPPLSFRAKDGC	MPPLSFRAKDGC	2,2492 (ANTÍGENO PROBABLE)
LSMVGDWLEKRPGAFC	GDWLEKRPGAFC	1,6716 (ANTÍGENO PROBABLE)
GEGRLDPYWGDVKQDL	LDPYWGDVKQDL	1,0185 (ANTÍGENO PROBABLE)
KTHAAAQRRGRIGRNP	AAQRRGRIGRNP	0,9970 (ANTÍGENO PROBABLE)
AKVEITPNSPRAEATL	ITPNSPRAEATL	0,3027 (probable NO ANTÍGENO)
RGYISTRVEMGEAAAI	STRVEMGEAAAI	0,8373 (ANTÍGENO PROBABLE)
YRTQMKGPWHSEELEI	MKGPWHSEELEI	0,7084 (ANTÍGENO PROBABLE)
LDEGVEPDDVDCWCNT	VEPDDVDCWCNT	0,9272 (ANTÍGENO PROBABLE)

WDVPAPKEVKKGETT	PAPKEVKKGETT	1,1609 (ANTÍGENO PROBABLE)
NSTHEMYWVSGAKSNT	EMYWVSGAKSNT	0,6293 (ANTÍGENO PROBABLE)
AVEAAVNDPRFWALVD	AVNDPRFWALVD	0,4820 (probable NO ANTÍGENO)
HKKGEARRSRAVTLP	EARRSRAVTLP)	0,9162 (ANTÍGENO PROBABLE)
NRVWIEENDHMEDKTP	IEENDHMEDKTP	0,7340 (ANTÍGENO PROBABLE)
VVLIPEPEKQRSPQDN	PEPEKQRSPQDN	0,0010 (probable NO ANTÍGENO)
EKRVLKPRWMDARVCS	LKPRWMDARVCS	0,7801 (ANTÍGENO PROBABLE)
HGSYEAPTQGSASSLI	EAPTQGSASSLI	0,0018 (probable NO ANTÍGENO)
GASLIYTVTRNAGLVK	IYTVTRNAGLVK	0,1024 (probable NO ANTÍGENO)
TVEVQYAGTDGPCKVP	QYAGTDGPCKVP	-0,2179 (probable NO ANTÍGENO)
HVTKGSALRSGEGRD	GSALRSGEGRD	-0,0065 (probable NO ANTÍGENO)
SLVEDDGPPMREIILK	DDGPPMREIILK	-0,0566 (probable NO ANTÍGENO)

Elaborado por: Balseca Quishpe, Carmen Elena

3.1.5. Evaluación de alergenicidad y toxicidad de los epítomos

La alergenicidad de los epítomos de 16 aa fue evaluada mediante AllerTOP v2.1, AllergenFP v1.1 y AlgPred. En el análisis de 16 aa, la mayoría de las secuencias presentó al menos una clasificación alergénica en alguna de las herramientas. La secuencia EKRVLKPRWMDARVCS fue clasificada como probablemente no alergénica por AllerTOP v2.1 y AllergenFP v1.1, y como no alergénica por AlgPred. Los resultados completos se presentan en la **Tabla 9**.

Para los péptidos de 12 aa, la evaluación se realizó mediante AllerTOP v2.1 y AlgPred. En los cortes iniciales, las secuencias EKRVLKPRWMDA y HGSYEAPTQGSA fueron clasificadas como probablemente no alergénicas o no alergénicas por ambas herramientas. En los cortes finales, HGSYEAPTQGSA presentó clasificación no alergénica en ambas herramientas, mientras que otras secuencias presentaron resultados discrepantes entre AllerTOP v2.1 y AlgPred. Estos resultados se detallan en las **Tablas 10 y 11**.

El análisis de toxicidad mediante ToxinPred se realizó considerando la división de la poliproteína completa en fragmentos de 500 aa y un fragmento final de 425 aa. En el primer fragmento se identificaron 8 regiones tóxicas y en el tercer fragmento 14 regiones tóxicas. Los demás fragmentos evaluados fueron clasificados como sin regiones tóxicas, según se presenta en la **Tabla 11**.

Tabla 9: Análisis de alergenidad por epítomos, en 3 herramientas distintas: AllerTOP v2.1 Allergen FP AlgPred. Considerando los epítomos B lineales de 16 aa.

Secuencia de 16 aa	AllerTOP v2.1	Allergen FP	AlgPred
RECTMPPLSFRAKDGC	ALÉRGENO probable	ALÉRGENO probable	ALÉRGENO
LSMVGDWLEKRPGAFC	ALÉRGENO probable	Probablemente NO ALÉRGENO	ALÉRGENO
GEGRLDPYWGDVKQDL	Probablemente NO ALÉRGENO	ALÉRGENO probable	ALÉRGENO
KTHAAAQRRGRIGRNP	ALÉRGENO probable	Probablemente NO ALÉRGENO	ALÉRGENO
AKVEITPNSPRAEATL	Probablemente NO ALÉRGENO	Probablemente NO ALÉRGENO	ALÉRGENO
RGYISTRVEMGEAAAI	Probablemente NO ALÉRGENO	Probablemente NO ALÉRGENO	ALÉRGENO
YRTQMKGPWHSEELEI	Probablemente NO ALÉRGENO	ALÉRGENO probable	ALÉRGENO
LDEGVEPDDVDCWCNT	ALÉRGENO probable	ALÉRGENO probable	ALÉRGENO
LWDVPAPKEVKKGETT	Probablemente NO ALÉRGENO	Probablemente NO ALÉRGENO	ALÉRGENO
NSTHEMYWVSGAKSNT	ALÉRGENO probable	ALÉRGENO probable	ALÉRGENO
AVEAAVNDPRFWALVD	ALÉRGENO probable	ALÉRGENO probable	ALÉRGENO
HKKGEARRSRRAVTLF	Probablemente NO ALÉRGENO	Probablemente NO ALÉRGENO	ALÉRGENO
VVLIPEPEKQRSPQDN	ALÉRGENO probable	Probablemente NO ALÉRGENO	ALÉRGENO
NRVWIEENDHMEDKTP	ALÉRGENO probable	ALÉRGENO probable	ALÉRGENO
EKRVLKPRWMDARVCS	Probablemente NO ALÉRGENO	Probablemente NO ALÉRGENO	NO ALÉRGENO
HGSYEAPTQGSASSLI	Probablemente NO ALÉRGENO	Probablemente NO ALÉRGENO	ALÉRGENO
GASLIYTVTRNAGLVK	ALÉRGENO probable	Probablemente NO ALÉRGENO	ALÉRGENO
TVEVQYAGTDGPCKVP	ALÉRGENO probable	ALÉRGENO probable	ALÉRGENO

HVTKGSALRSGEGRD	ALÉRGENO probable	ALÉRGENO probable	NO ALÉRGENO
SLVEDDGPPMREILK	ALÉRGENO probable	Probablemente NO ALÉRGENO	NO ALÉRGENO

***Elaborado por:** Balseca Quishpe, Carmen Elena*

Tabla 10: Análisis de alergenidad por epítomos, considerando el corte de los 12 aa iniciales.

Secuencia de 12 primeros (corte 1-12)	AllerTop v2.1	AlgPred
RECTMPPLSFRA	ALÉRGENO probable	ALÉRGENO
LSMVGDWLEKRP	ALÉRGENO probable	ALÉRGENO
GEGRDLPYWGDV	Probablemente NO ALÉRGENO	ALÉRGENO
KTHAAAQRRGRI	Probablemente NO ALÉRGENO	ALÉRGENO
AKVEITPNSPRA	Probablemente NO ALÉRGENO	ALÉRGENO
RGYISTRVEMGE	ALÉRGENO probable	ALÉRGENO
YRTQMKGPWHSE	ALÉRGENO probable	NO ALÉRGENO
LDEGVEPDDVDC	ALÉRGENO probable	ALÉRGENO
LWDVPAPKEVKK	Probablemente NO ALÉRGENO	ALÉRGENO
NSTHEMYWVSGA	Probablemente NO ALÉRGENO	ALÉRGENO
AVEAAVNDPRFW	ALÉRGENO probable	ALÉRGENO
HKKGEARRSRRA	Probablemente NO ALÉRGENO	ALÉRGENO
VVLIPEPEKQRS	ALÉRGENO probable	ALÉRGENO
NRVWIEENDHME	Probablemente NO ALÉRGENO	ALÉRGENO
EKRVLKPRWMDA	Probablemente NO ALÉRGENO	NO ALÉRGENO
HGSYEAPTQGSA	Probablemente NO ALÉRGENO	NO ALÉRGENO
GASLIYTVTRNA	ALÉRGENO probable	ALÉRGENO
TVEVQYAGTDGP	Probablemente NO ALÉRGENO	ALÉRGENO
HVTKGSALRSGE	ALÉRGENO probable	NO ALÉRGENO
SLVEDDGPPMRE	ALÉRGENO probable	NO ALÉRGENO

***Elaborado por:** Balseca Quishpe, Carmen Elena*

Tabla 11: Análisis de alergenidad por epítomos, considerando el corte de los 12 aa iniciales.

Secuencia de 12 últimos (corte 4-16)	AllerTop v2.1	AlgPred out
MPPLSFRAKDGC	Probablemente NO ALÉRGENO	ALÉRGENO
GDWLEKRPGAFC	Probablemente NO ALÉRGENO	ALÉRGENO
LDPYWGDVKQDL	ALÉRGENO probable	ALÉRGENO
AAQRRGRIGRNP	Probablemente NO ALÉRGENO	ALÉRGENO
ITPNSPRAEATL	ALÉRGENO probable	ALÉRGENO
STRVEMGEAAAI	Probablemente NO ALÉRGENO	ALÉRGENO
MKGPWHSEELEI	Probablemente NO ALÉRGENO	ALÉRGENO
VEPDDVDCWCNT	ALÉRGENO probable	ALÉRGENO
PAPKEVKKGETT	Probablemente NO ALÉRGENO	ALÉRGENO
EMYWVSGAKSNT	ALÉRGENO probable	ALÉRGENO
AVNDPRFWALVD	Probablemente NO ALÉRGENO	ALÉRGENO
EARRSRRAVTLP	Probablemente NO ALÉRGENO	ALÉRGENO
PEPEKQRSPQDN	Probablemente NO ALÉRGENO	ALÉRGENO
IEENDHMEDKTP	Probablemente NO ALÉRGENO	ALÉRGENO
LKPRWMDARVCS	ALÉRGENO probable	NO ALÉRGENO
HGSYEAPTQGSA	Probablemente NO ALÉRGENO	NO ALÉRGENO
IYTVTRNAGLVK	Probablemente NO ALÉRGENO	ALÉRGENO
QYAGTDGPCKVP	ALÉRGENO probable	ALÉRGENO
GSALRSGEGRLD	ALÉRGENO probable	NO ALÉRGENO
DDGPPMREILK	ALÉRGENO probable	ALÉRGENO

Elaborado por: Balseca Quishpe, Carmen Elena

Para el análisis de la toxicidad se utilizó el servidor ToxinPred (https://webs.iitd.edu.in/raghava/toxinpred/prot_test.php), empleando parámetros como: un umbral de 0.4 con cortes de 15 aa de longitud y secuencias con una longitud ≤ 500 aa. Por lo cual, la secuencia completa se dividió 6 secciones de 500 aa y la última sección de 425 aa, demostrado en la **Tabla 12**.

Tabla 12: Análisis de 500 aa derivados de la secuencia completa y regiones tóxicas.

Secuencia	Toxicidad
MVMKNPKKKSGGFRIVNMLKRGVARVSPFGGLKRLPAGLLLGHPIRMVLAILAFLRFTAIPKPSLGLINR WGSVGKKEAMEI IKKFKKDLAAMLRI INARKEKKRRGADTSVGIVGLLLLTTAMAAEVTRRGSAYMYLDR NDAGEAISFP TTLGMNKCYIQIMDLGHMCDATMSYECPLMDEGVEPDDVDCWCNTTSTWVVGTCCHHKKG EARRSRRAVTLPSHSTRKLQTRSQTWLESREYTKHLIRVENWIFRNPGFALAAAAIAWLLGSSTSQKVIY LVMILLIAPAYSIRCIGVSNRDFVEGMSGGTWVDIVLEHGGCVTVMAQDKPTVDIELVTTTVSNMAEVR YCYEASISDMASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWNGCGLFGKGS LVTCAKFACSKMTGK SIQPENLEYRIMLSVHGSQHSGMIVNDTGHE TDENRAKVEITPNSPRAEATLGGFGSLGLDCEPRTGLDF SDLYYLTMMN	8 regiones tóxicas
KHWLVHKEWFHDIPLPWHAGADTGT PHWNNKEALVEFKDAHAKRQTVVVLGSQEGAVHTA LAGALEAEMDGA KGR LSSGHLKCR LKMDKRLKGVSYSLCTAAFTFTKIPAE TLHGTVTVVEVQYAGTDGP CKVPAQMAVDMQTLTPVGRLITANPVITESTENSKMMLELDPFFGDSYIVIGVGEKKITHHWHRS GSTIG KA FEATVRGA KRMAVLGDTAWDFGSVGGALNSLGKGIHQIFGA AFKSLFGGMSWFSQILIGTLLMWLGLN TKNGSISLMCLALGGVLI FLSTAVSADVGCSVDFS KKETRCGTGVFVYNDVEAWRDRIKYHPDSPRRLAA AVKQAWEDGICGISSVSRMENIMWRSVEGELNAILEENG VQLTVVVGSVKNPMWRGPQRLPVPVNE LPHG WKAWGKSHFVRAAKTNSFVVDGDTLKECPLKHRAWNSFLVEDHGFV FHTSVWLKVREDYSLECDPAVI GTAVKGKEAVHSDLGWIES	Ninguna tóxica
EKNDTWRLKRAHLIEMKTCEWPKSHTLWTDGIEESDLIIPKSLAGPLSHH NTREGYRTQMKGPWHSEELIRFEECPG TKVHVEETCGTRGPSLRSTTASGRVIEEWCCRECTMPPLSFR AKDGCWYGMEIRPRKEPESNLVRSMVTAGSTDHMDHFS LGVLVILLMVQEG LKKRMTTKIIISTSM AVL AMILGGFSMSDLAKLAILMGATFAEMNTGGDVAHLALIAAFKVRPALLVSFIFRANWTPRESMLLALASC LLQTAISALEGDLMLVINGFALAWLAIRAMVVPRTDNITLAILAALTPLARGTLLVAWRAGLATCGGFML LSLKGKGSVKKNLPFVMALGLTAVRLVDPINVVGLLLLLTRSGKRSWPPSEVLTAVGLICALAGGFAKADI EMAGPMAAVGLLIVSYVVGSKSVDMYIERAGDITWEKDAEVTGNSPRLDVALDESGDFS LVEDDGPPMRE IILKVVLMTICGMNPIAIPFAAGAWYVYVK	14 regiones tóxicas
TGKRSGALWDVPAPKEVKKGETTDGVYRVMTRRLLGSTQV GVGVMQEGVFHTMWHVTKGSALRSGEGR LDPYWG DVKQDLVSYCGPWKLDAAWDGHSEVQLLAVPPGERA RNIQTLPGIFKTKDGDIGAVALDYPAGTSGSPILDKCGRVIGLYGNGVVIKNGSYVSAITQGRREEETPV ECFEPSMLKKKQLTVLDLHPGAGKTRRVLPEIVREAIKTRLRTVILAPTRVVAEMEEALRGLPVR YMTT AVNVTHSGTEIVDLMCHATFTSRLLQPIRVPNYNLYIMDEAHFTDPSSIAARGYISTRVEMGEAAAI FMT ATPPGTRDAFPDSNSPIMDTEVEVPERAWSSGFDWVTDYSGKTVWFVPSVRNGNEIAACLT KAGKRVIQL SRKTFETEFQKTKHQEWDFVVTTDISEMGANFKADRVIDSRRLCKPVILDGERVILAGPMPVTHASAAQR RGRIGRNPKNPGDEYLYGGGCAETDEDHAWLEARMLLDN	Ninguna tóxica
IYLDGLIASLYRPEADKVAAIEGEFKLRT EQRKTFVELMKRGDLPVWLAYQVASAGITYTDRRWC FDGTNNTIMEDSVPAEVWTRHGEKRV LKPRWMD ARVCS DHAALKSFKEFAAGKRGAAGFVMEALGTLPGHMTERFQEAIDNLAVLMRAETGSRPYKAAAAQLP	Ninguna tóxica

ETLETIMLLGLLGTVSLGIFVLMRNKGIGKMGFGMVTLGASAWLMWLSEIEPARIACVLIVVFLLLVVL
 IPEPEKQSPQDNQMAIIIMVAVGLLGLITANELGWLERTKSDLSHLMGRREEGATMGFSMDIDLRPASA
 WAIYAALTTFITPAVQHAVTTSYNNYSLMAMATQAGVLFMGKGMPFYAWDFGVPLLMIGCYSQTLPLTL
 IVAIILLVAHYMYLIPGLQAAAAARAAQKRTAAGIMKNPVVDGIVVTDIDTMTIDPQVEKKMGQVLLMAVA
 VSSAILSRTAWGWGEAGALITAATSTLWEGSPNKYWNSSSTATSLCNIFRG

SYLAGASLIYTVTRNAGLVK

Ninguna
 toxica

RRGGGTGETLGEKWKARLNQMSALEFYSSKSGITEVCREEARRALKDGVATGGHAVSRGSAKLRWLVER
 GYLQPYGKVIDLGCGRGWSYYAATIRKVQEVKGYTKGGPGHEEPVLVQSYGWNIVRLKSGVDVFHMAAE
 PCDTLLCDIGESSSSPEVEEARTLRVLSMVGDWLEKRPAGFCIKVLCPYTSTMMETLERLQRRYGGGLVR
 VPLSRNSTHEMYWVSGAKSNTIKSVSTTSQLLLGRMDGPRRPVKYEEDVNLGSGTRAVVSCAEAPNMKII
 GNRIERIRSEHAETWFFDENHPYRTWAYHGSYEAPTQGSASSLINGVVRLLSKPWDVVTGVTGIAMTDTT
 PYGQQRVFKEKVDTRVPDPQEGTRQVMSMVSSWLWKELGKHKRPRVCTKEEFINKVRSNAALGAIFEEEEK
 EWKTAVEAVNDPRFWALVDKEREHLRGECQSCVYNMMGKREKKQGEFGKAKGSRAIWM

WLGARFLEFE

Ninguna
 toxica

ALGFLNEDHWMGRENSGGGVEGLGLQRLGYVLEEMSRIPGGRMYADDTAGWDTRISRFDLNEALITNQM
 EKGHRALALAI IKYTYQNKVVKVL RPAEKGKTVM DII SRQDQRGSGQVVTYALNTFTNLVVQLIRNMEAE
 EVLEMQDLWLLRRSEKVTNWLQSNQWDRLKRMVSGDDCVVKPIDDRFAHALRFLNDMGKVRKDTQEWKP
 STGWDNWEVFPFCSHHFNKLHLKDGRSIVVPCRHQDELIGRARVSPGAGWSIRETACLAKSYAQMWQLLY
 FHRRDLRLMANAICSSVPVDWVPTGRTTWSIHGKGEMMTEDMLVVWNRVWIEENDHMEDKTPVTKWTDI
 PYLGKREDLWCGSLIGHRPRTTWAENIKNTVMVRRIIGDEEKYMDYLSTQVRYLGEEGSTPGVL

Elaborado por: Balseca Quishpe, Carmen Elena

3.1.6. Selección final de epítomos para el constructo multiepítomo

Luego de integrar los resultados de antigenicidad, alergenicidad y toxicidad, se seleccionaron los péptidos de 12 aa que cumplieron simultáneamente los criterios establecidos para el diseño del constructo. En los cortes iniciales, las secuencias que conservaron antigenicidad probable, clasificación no alergénica o probablemente no alergénica y ausencia de toxicidad fueron KTHAAAQRRGRI, AKVEITPNSPRA, TVEVQYAGTDGP y NSTHEMYWVSGA, entre otras evaluadas en la **Tabla 13**.

En los cortes finales, las secuencias que cumplieron los criterios establecidos incluyeron PAPKEVKKGETT, EARRSRRAVTLP, STRVEMGEAAAI, IEENDHMEDKTP y MKGPWHSEELEI, como se muestra en la **Tabla 14**.

Tabla 13: Epítomos iniciales de 12 aa seleccionados tras el proceso de validación.

Secuencia de 12 epítomos primeros (corte 1-12)	Antigenicidad (VaxiJen v2.0)	Alergenicidad (AllerTOP v2.1)	Toxicidad (ToxinPred)
GEGRLDPYWGDV	1,4145 (ANTÍGENO PROBABLE).	Probablemente NO ALÉRGENO	TOXICO
RECTMPPLSFRA	1,3848 (ANTÍGENO PROBABLE)	ALÉRGENO probable	TOXICO
LDEGVEPDDVDC	1,0707 (ANTÍGENO PROBABLE).	ALÉRGENO probable	TOXICO
KTHAAAQRRGRI	0,9092 (ANTÍGENO PROBABLE).	Probablemente NO ALÉRGENO	NO TOXICO
AKVEITPNSPRA	0,8183 (ANTÍGENO PROBABLE),	Probablemente NO ALÉRGENO	NO TOXICO
RGYISTRVEMGE	0,8607 (ANTÍGENO PROBABLE).	ALÉRGENO probable	NO TOXICO
LSMVGDWLEKRP	0,7586 (ANTÍGENO PROBABLE)	ALÉRGENO probable	NO TOXICO
TVEVQYAGTDGP	(0,6891) (ANTÍGENO PROBABLE)	Probablemente NO ALÉRGENO	NO TOXICO
NSTHEMYWVSGA	0,5820 (ANTÍGENO PROBABLE).	Probablemente NO ALÉRGENO	NO TOXICO
HVTKGSALRSGE	0,5320 (ANTÍGENO PROBABLE).	ALÉRGENO probable	NO TOXICO
YRTQMKGPWHSE	0,5117 (ANTÍGENO PROBABLE)	ALÉRGENO probable	NO TOXICO
EKRVLKPRWMDA	0,4122 (probable NO ANTÍGENO).	Probablemente NO ALÉRGENO	NO TOXICO
GASLIYTVTRNA	0,3955 (probable NO ANTÍGENO)	ALÉRGENO probable	NO TOXICO
HKKGEARRSRRA	0,3650 (probable NO ANTÍGENO).	Probablemente NO ALÉRGENO	NO TOXICO
VVLIPEPEKQRS	0,2890 (probable NO ANTÍGENO).	ALÉRGENO probable	NO TOXICO
SLVEDDGPPMRE	0,1645 (probable NO ANTÍGENO).	ALÉRGENO probable	NO TOXICO
LWDVPAPKEVKK	0,0696 (probable NO ANTÍGENO).	Probablemente NO ALÉRGENO	NO TOXICO
HGSYEAPTQGSA	0,0666 (probable NO ANTÍGENO).	Probablemente NO ALÉRGENO	NO TOXICO
AVEAAVNDPRFW	-0,0004 (probable NO ANTÍGENO).	ALÉRGENO probable	NO TOXICO
NRVWIEENDHME	-0,0676 (probable NO ANTÍGENO).	Probablemente NO ALÉRGENO	NO TOXICO

Elaborado por: Balseca Quishpe, Carmen Elena

Tabla 14: Epítomos finales de 12 aa seleccionados tras el proceso de validación.

Secuencia de 12 epítomos últimos (corte 4-16)	Antigenicidad (VaxiJen v2.0)	Alergenicidad (AllerTOP v2.1)	Toxicidad (ToxinPred)
MPPLSFRAKDGC	2,2492 (ANTÍGENO PROBABLE)	Probablemente NO ALÉRGENO	TOXICO
GDWLEKRPGAFC.	1,6716 (ANTÍGENO PROBABLE)	Probablemente NO ALÉRGENO	TOXICO
PAPKEVKKGETT	1,1609 (ANTÍGENO PROBABLE)	Probablemente NO ALÉRGENO	NO TOXICO
LDPYWGDVKQDL	1,0185 (ANTÍGENO PROBABLE)	ALÉRGENO probable	NO TOXICO
AAQRRGRIGRNP	0,9970 (ANTÍGENO PROBABLE)	Probablemente NO ALÉRGENO	NO TOXICO
VEPDDVDCWCNT	0,9272 (ANTÍGENO PROBABLE)	ALÉRGENO probable	NO TOXICO
EARRSRRAVTLP	0,9162 (ANTÍGENO PROBABLE)	Probablemente NO ALÉRGENO	NO TOXICO
STRVEMGEAAAI	0,8373 (ANTÍGENO PROBABLE)	Probablemente NO ALÉRGENO	NO TOXICO
LKPRWMDARVCS	0,7801 (ANTÍGENO PROBABLE)	ALÉRGENO probable	NO TOXICO
IEENDHMEDKTP	0,7340 (ANTÍGENO PROBABLE)	Probablemente NO ALÉRGENO	NO TOXICO
MKGPWHSEELEI	0,7084 (ANTÍGENO PROBABLE)	Probablemente NO ALÉRGENO	NO TOXICO
EMYWVSGAKSNT	0,6293 (ANTÍGENO PROBABLE)	ALÉRGENO probable	NO TOXICO
AVNDPRFWALVD	0,4820 (probable NO ANTÍGENO)	Probablemente NO ALÉRGENO	NO TOXICO
ITPNSPRAEATL	0,3027 (probable NO ANTÍGENO)	ALÉRGENO probable	NO TOXICO
IYTVTRNAGLVK	0,1024 (probable NO ANTÍGENO)	Probablemente NO ALÉRGENO	NO TOXICO
EAPTQGSASSLI	0,0018 (probable NO ANTÍGENO)	Probablemente NO ALÉRGENO	NO TOXICO
PEPEKQRSPQDN	0,0010 (probable NO ANTÍGENO)	Probablemente NO ALÉRGENO	NO TOXICO
DDGPPMREILK	-0,0566 (probable NO ANTÍGENO)	ALÉRGENO probable	NO TOXICO
GSALRSGEGRDL	-0,0065 (probable NO ANTÍGENO)	ALÉRGENO probable	NO TOXICO
QYAGTDGPCKVP	-0,2179 (probable NO ANTÍGENO)	ALÉRGENO probable	NO TOXICO

Elaborado por: Balseca Quishpe, Carmen Elena

Finalmente, la **Tabla 15** consolidó nueve epítomos aprobados para el ensamblaje del constructo multiepítomo:

Tabla 15: Secuencias aprobadas tras los filtros, dichas secuencias de 12 aa iniciales y finales.

Epítomos	Antigenicidad (VaxiJen v2.0)	Alergenicidad (AllerTOP v2.1)	Toxicidad (ToxinPred)
PAPKEVKKGETT	1,1609 (ANTÍGENO PROBABLE)	Probablemente NO ALÉRGENO	NO TOXICO
EARRSRRAVTLP	0,9162 (ANTÍGENO PROBABLE)	Probablemente NO ALÉRGENO	NO TOXICO
KTHAAAQRRGRI	0,9092 (ANTÍGENO PROBABLE).	Probablemente NO ALÉRGENO	NO TOXICO
STRVEMGEAAAI	0,8373 (ANTÍGENO PROBABLE).	Probablemente NO ALÉRGENO	NO TOXICO
AKVEITPNSPRA	0,8183 (ANTÍGENO PROBABLE),	Probablemente NO ALÉRGENO	NO TOXICO
IEENDHMEDKTP	0,7340 (ANTÍGENO PROBABLE).	Probablemente NO ALÉRGENO	NO TOXICO
MKGPWHSEELEI	0,7084 (ANTÍGENO PROBABLE).	Probablemente NO ALÉRGENO	NO TOXICO
TVEVQYAGTDGP	(0,6891) (ANTÍGENO PROBABLE)	Probablemente NO ALÉRGENO	NO TOXICO
NSTHEMYWVSGA	0,5820 (ANTÍGENO PROBABLE).	Probablemente NO ALÉRGENO	NO TOXICO

Elaborado por: Balseca Quishpe, Carmen Elena

3.1.7. Diseño del constructo multiepítomo

Los nueve epítomos aprobados fueron ensamblados mediante el enlazador peptídico flexible GPGPG. La distribución se realizó considerando la puntuación de antigenicidad, ubicando los epítomos con mayor puntuación hacia la región central y los de menor puntuación hacia los extremos. Al inicio de la secuencia se incorporó una metionina (M) y al extremo C-terminal se añadió una cola de seis histidinas (HHHHHH).

La secuencia final del constructo multiepitótopo fue:

MNSTHEMYWVSGAGPGPGMKGPWHSEELEIGPGPGAKVEITPNSPRAGPGPGKTHAAAQRRGRIGPGPGGEARRSRRAVTLPGPGPAPKEVKKGETTGPGPGSTRVEMGEAAAIGPGPGIEENDHMEDKTPGPGPGTVEVQYAGTDGPHHHHHH

Posteriormente, el constructo ensamblado fue reevaluado mediante VaxiJen v2.0, AllerTOP v2.1 y ToxinPred. Los resultados obtenidos fueron: VaxiJen v2.0: 0,7422, clasificado como antígeno probable; AllerTOP v2.1: probablemente no alergénico; y ToxinPred: no tóxico. La organización esquemática y secuencia del constructo se presenta en la **Figura 3**.

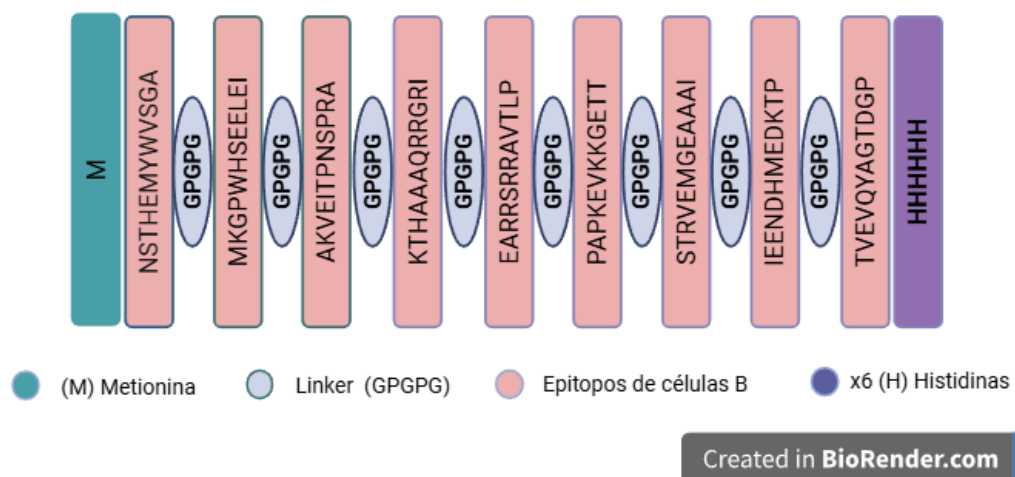


Figura 3: Ensamblaje del constructo.

Imagen elaborada en BioRender por Balseca Quishpe, Carmen Elena

3.1.8. Análisis fisicoquímico del constructo

El constructo final fue evaluado mediante ProtParam del servidor ExPASy. El análisis reportó una longitud de 155 aminoácidos, un punto isoeléctrico teórico de 6,80 y un peso molecular de 15841,53 Da. La fórmula molecular estimada fue C681H1060N216O213S5.

La vida media estimada fue de 30 horas en reticulocitos de mamíferos *in vitro*, mayor a 20 horas en levadura *in vivo* y mayor a 10 horas en *Escherichia coli in vivo*. El índice

de inestabilidad fue de 37,12, clasificado por el servidor como estable. Estos resultados se presentan en la **Tabla 16**.

Tabla 16: Análisis fisicoquímico mediante ProtParam.

Parámetro	Resultado
Número de aminoácidos	155
Punto isoeléctrico teórico	6,80
Peso molecular	15841,53
Formula	C ₆₈₁ H ₁₀₆₀ N ₂₁₆ O ₂₁₃ S ₅
Vida media estimada	• 30 horas (reticulocitos de mamíferos, <i>in vitro</i>) • >20 horas (levadura, <i>in vivo</i>) • >10 horas (<i>Escherichia coli</i>, <i>in vivo</i>)
Índice de inestabilidad (II)	37.12 (Se clasifica como estable)
Índice GRAVY	-0,977

Elaborado por: Balseca Quishpe, Carmen Elena

3.1.9. Modelado estructural del constructo

La estructura tridimensional del constructo fue predicha mediante ColabFold v1.6.1: AlphaFold2 using MMseqs2. El modelo tridimensional final se muestra en la **Figura 4**. Además, el servidor generó cinco modelos estructurales, ordenados como Rank_1, Rank_2, Rank_3, Rank_4 y Rank_5, los cuales se presentan en la **Figura 5**.

Las gráficas de error alineado predicho mostraron zonas de menor incertidumbre en regiones cercanas a la diagonal, mientras que otras zonas presentaron mayor incertidumbre relativa. La cobertura de secuencias homólogas utilizada para la construcción del alineamiento múltiple se presenta en la **Figura 6**.

La predicción de IDDT por posición mostró que los modelos Rank_1, Rank_2 y Rank_3 presentaron valores principalmente entre 40 y valores cercanos a 80, mientras que los modelos Rank_4 y Rank_5 mostraron valores inferiores, cercanos a 20–35. Estos resultados se presentan en la **Figura 7**.

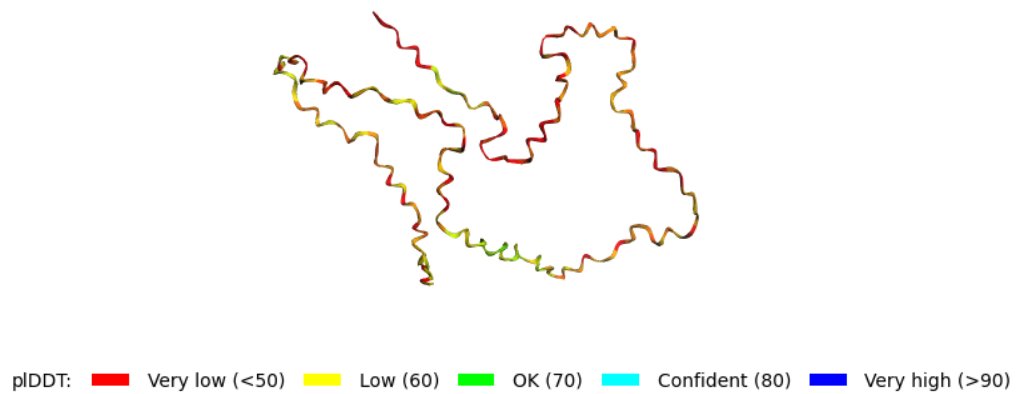


Figura 4: Estructura tridimensional del constructo obtenida a través del servidor AlphaFold.

En la **Figura 5**, se presentan los 5 modelos posibles generados por el servidor AlphaFold, ordenados desde Rank_1 a Rank_5, con la finalidad de evaluar la calidad estructural predicha del constructo. Las gráficas corresponden al error alineado predicho, donde la línea diagonal de color azul representa una baja incertidumbre en relación entre los residuos cercanos de la secuencia, mientras que las zonas de color rojo indican mayor incertidumbre. Este comportamiento puede ser dado a posibles interacciones entre residuos, las cuales pueden estar asociadas con procesos de plegamiento y conformación de la proteína.

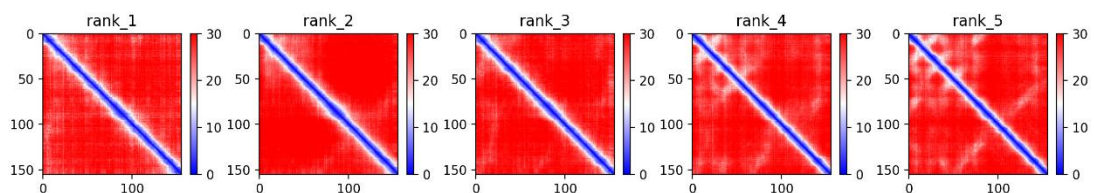


Figura 5: Rank_1 al Rank_5

Mientras que en la **Figura 6** se evidencia la cobertura de las secuencias homólogas empleadas para la construcción del alineamiento múltiple de secuencias utilizado durante la predicción estructural.

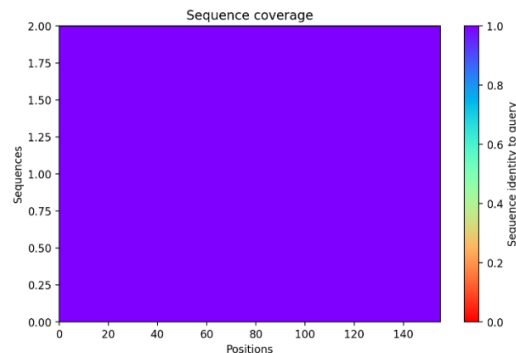


Figura 6: Gráfica de la cobertura de secuencias homólogas.

Finalmente, en la **Figura 7**, se visualiza la predicción IDDT por posición para los 5 modelos generados por AlphaFold, donde el Rank_1, Rank_2 y Rank_3 muestran valores principalmente de 40 a rangos cercanos a 80, siendo estos, los de mayor interés y confianza. Por el contrario, los modelos Rank_4 y Rank_5 presentan valores inferiores, cercanos a 20–35, lo que indica menor confiabilidad estructural.

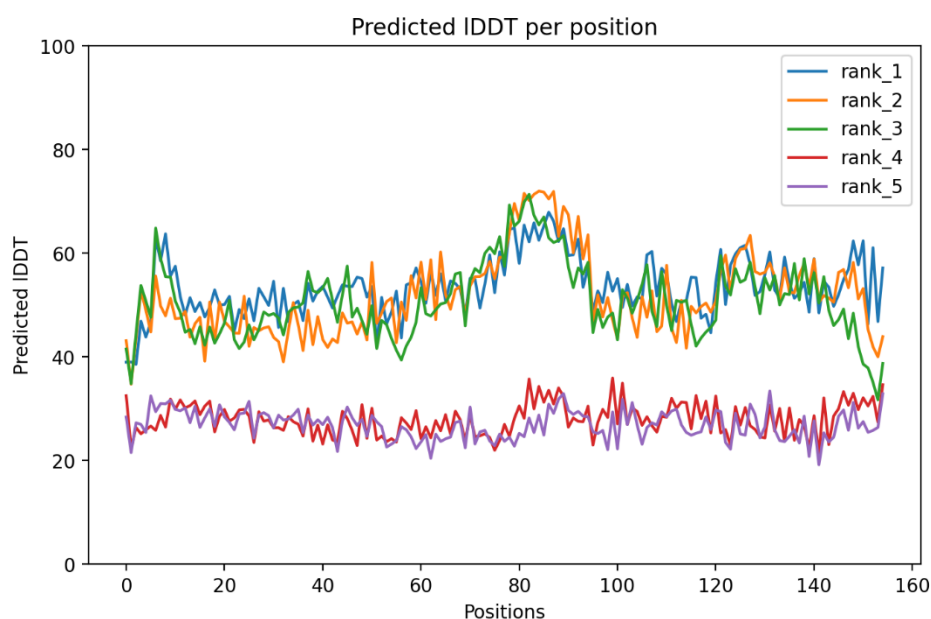


Figura 7: Predicción de IDDT por posición.

3.1.10. Refinamiento y validación estructural

El modelo estructural seleccionado fue refinado mediante GalaxyRefine, obteniéndose el modelo refinado presentado en la **Figura 8** y en la **Figura 9** las puntuaciones de calidad estructural proporcionadas por la plataforma.

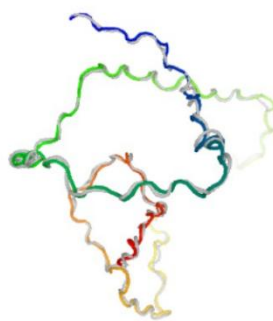


Figura 8: Imagen tridimensional refinada generada por GalaxyRefine (Model_1)

Modelo	GDT-HA	RMSD	MolProbity	Marcador del choque	Pobres rotámeros	Rama favorecido
Iniciales	1.0000	0.000	3.527	18.5	7.2	13.7
MODELO 1	0.7645	0.864	0.658	0.5	0.0	99.3
MODELO 2	0.7694	0.878	0.866	1.4	0.9	98.0
MODELO 3	0.7726	0.858	0.940	1.8	0.9	99.3
MODELO 4	0.7935	0.832	0.500	0.0	0.9	98.0
MODELO 5	0.7661	0.877	0.658	0.5	0.0	98.7

Figura 9: Puntuaciones obtenidas para los cinco modelos generados por GalaxyRefine

Posteriormente, el archivo correspondiente al Modelo 1 fue evaluado en el servidor SAVES v6.1 mediante las herramientas ERRAT y PROCHECK (Ramachandran Plot). En ERRAT se obtuvo un Overall Quality Factor de 51,3514, como se muestra en la **Figura 10**. El gráfico reportó regiones con diferente nivel de riesgo estructural, representadas en colores rojo, amarillo y verde



Figura 10: Resultado de la herramienta ERRAT obtenido del servidor SAVES v6.1.

En PROCHECK se reportaron los resultados obtenidos de Ramachandran Plot, donde el 98% de los residuos están localizados en regiones favorecidas (A, B, L), mientras que el 2% se localizan en regiones adicionalmente permitidas. Expuesto en la **Figura 11**.

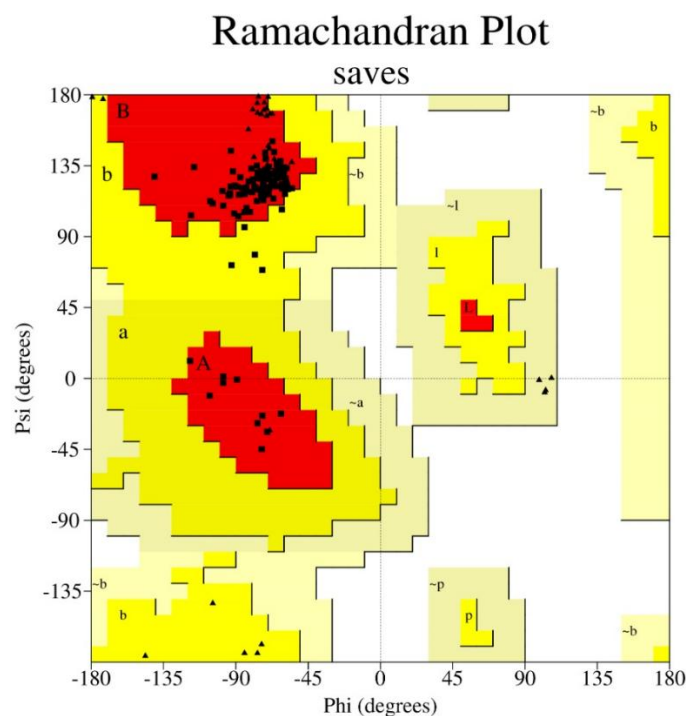


Figura 11: Resultado de Ramachandran Plot en PROCHECK en el servidor SAVES v6.1.

3.1.11. Optimización génica y análisis de ARNm

La secuencia proteica final fue retrotraducida y optimizada para expresión heteróloga en *Escherichia coli* cepa K12 mediante JCat. El servidor generó una secuencia de ADN optimizada de 450 pb, correspondiente al constructo proteico de 155 aa. El índice de adaptación de codones (CAI) obtenido fue de 1,0, mientras que el contenido de GC de la secuencia optimizada fue de 59.78 %, como se muestra en la **Figura 12**.

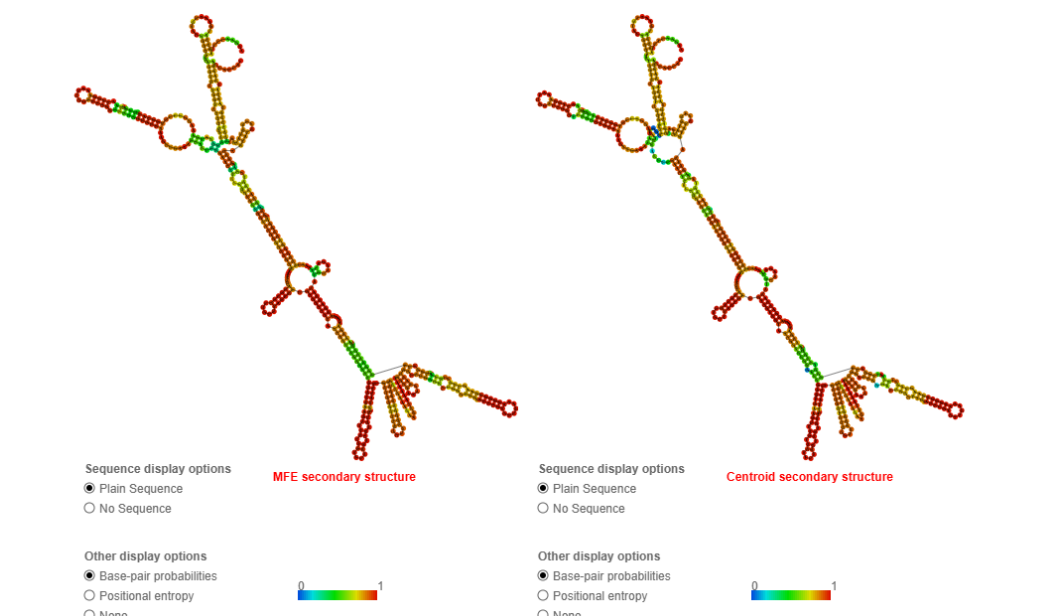


Figura 14: Estructuras secundarias finales de ARNm

3.1.12. Análisis BLASTp y verificación de similitud frente a CHIKV y OROV

El constructo proteico multiepítipo fue sometido a análisis de similitud mediante BLASTp. El análisis general mostró dos agrupaciones con alineamientos significativos. El primer clúster correspondió a una secuencia descrita como FliC quimérico [construcción sintética], con puntuación máxima de 48,9, puntuación total de 48,9, cobertura de consulta de 62 %, valor E de 0,003, identidad de 40,82 %, longitud de acceso de 530 y adhesión UNP53947.1. El segundo clúster correspondió a un unnamed protein product, partial [Boreogadus s...], con puntuación máxima de 45,1, puntuación total de 45,1, cobertura de consulta de 86 %, valor E de 0,046, identidad de 30,71 %, longitud de acceso de 248 y adhesión CAL8236413.1, como se presenta en la **Figura 15**.

Agrupaciones que producen alineaciones significativas									
seleccionar todo 2 clústeres. GenPept Gráficos Árbol de distancias de resultados Alineación múltiple Visor MSA									
	Composición del clúster	Ancestro del clúster	Secuencia representativa del clúster	Puntuación máxima	Puntuación total	Portada de la consulta	Valor E	Per. Ident	Longitud de acceso
☒	1miembro(s), 1organismo(s)	other sequences	FliC quimérico [construcción sintética]	48.9	48.9	62%	0.003	40,82%	530
☒	1miembro(s), 1organismo(s)	bony fishes	unnamed protein product, partial [Boreogadus s...]	45.1	45.1	86%	0.046	30,71%	248

Figura 15: Análisis de similitud del constructo proteico multiepítipo mediante BLASTp en el servidor NCBI BLAST.

En la comparación frente a CHIKV, se evaluaron constructos sintéticos correspondientes a las proteínas E1 y E2. Para E1 se observó una similitud de 21,00 %, con puntuación de 93,0. Para E2 se observó una similitud de 22,00 %, con puntuación de 105. Estos resultados se presentan en la **Figura 16**.

A

Descripción	Puntaje	Porcentaje de identificación	Adhesión
producto proteico no identificado	93.0	21.00%	Consulta_700665

B

Descripción	Puntaje	Porcentaje de identificación	Adhesión
producto proteico no identificado	105	22.00%	Consulta_7067805

Figura 16: Reactividad cruzada con CHIKV.
A: constructo sintético E1.
B: constructo sintético E2.

En la comparación frente a OROV, se identificó una similitud de 26,00 %, con puntuación de 109, como se muestra en la **Figura 17**.

Descripción	Puntaje	Porcentaje de identificación	Adhesión
producto proteico no identificado	109	26.00%	Consulta_295999

Figura 17: Reactividad cruzada frente OROV.

3.1.13. Diseño plasmídico

A partir del gen optimizado obtenido en JCat se empleó el uso de NEBcutter (<https://nc3.neb.com/NEBcutter/>, consultado el 28 de mayo de 2026), para identificar los sitios de restricción y después mediante Benchling (<https://www.benchling.com/>, consultado el 28 de mayo de 2026), el gen optimizado se insertó con éxito en el vector de expresión pET28a(+), dicho vector se descargó de SnapGene ([https://www.snapgene.com/plasmids/pet_and_duet_vectors_\(novagen\)/pET-28a\(%2B\)](https://www.snapgene.com/plasmids/pet_and_duet_vectors_(novagen)/pET-28a(%2B)), consultado y descargado el 28 de mayo de 2026) utilizando los sitios de

en la **Figura 19**, se muestran las propiedades bioquímicas del plásmido.

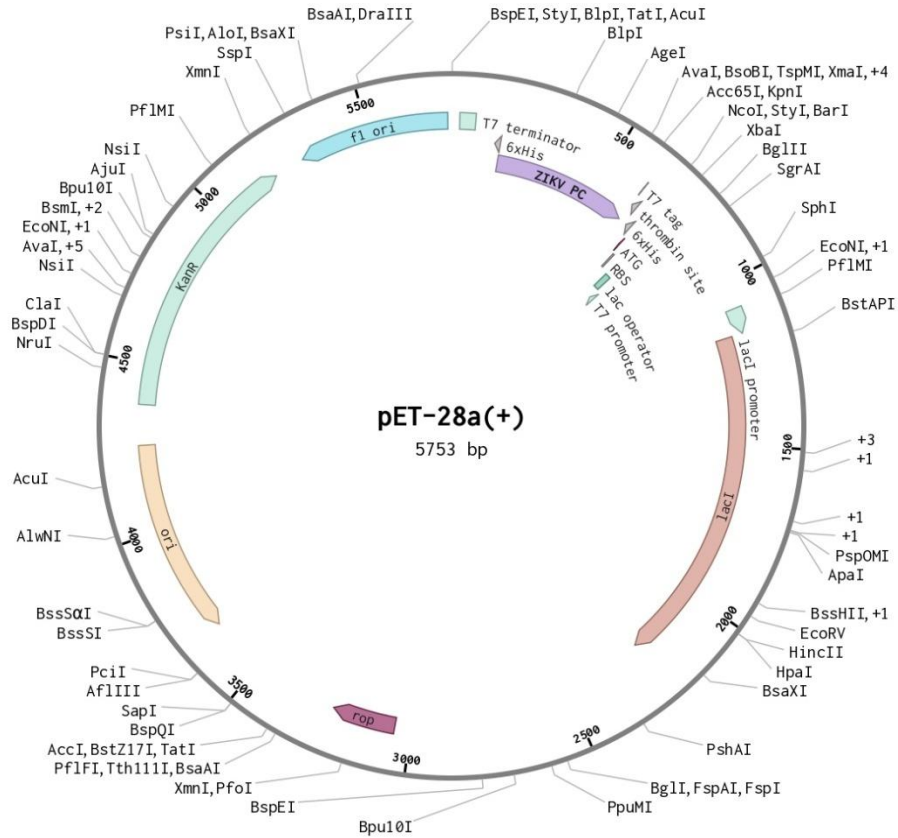


Figura 18: Inserción in silico del gen optimizado a través del vector pET28a (+).

Properties	
Position	158-622 (forward)
Summary	MNST...HHHH 155 AAs
Molecular weight ⓘ	15841.53 Da
Isoelectric point (pI)	6.80
Extinction coefficient	
Cys fully reduced	13980.00 M ⁻¹ cm ⁻¹
Abs 0.1% (1 g/l)	0.882
Cys fully oxidized	13980.00 M ⁻¹ cm ⁻¹
Abs 0.1% (1 g/l)	0.882
Instability index	37.12 (stable)

Figura 19: Propiedades bioquímicas del plásmido recombinante obtenidas del servidor Benchling.

3.2. DISCUSIÓN

Este estudio tuvo como finalidad un análisis inmunoinformático para el diseño *in silico* de una proteína multiepitopos del virus Zika, orientada para el inmunodiagnóstico diferencial específico en zonas endémicas. Este enfoque resulta relevante dado a que, en regiones cocirculan otros arbovirus, y en consecuencia produce un desafío diagnóstico (3). En este contexto el empleo de inmunoinformática, como lo es la selección de epítomos antigénicos, no alérgicos y no tóxicos, constituyen una estrategia crucial para el inmunodiagnóstico diferencial (14,27).

Para el estudio se consideró únicamente la proteína no estructural 1 (NS1), una glicoproteína implicada en la replicación viral, la evasión inmune y la patogénesis (24). Esta proteína es un blanco diagnóstico clave dado que sus propiedades altamente específicas difieren en cada flavivirus, lo que minimiza los resultados de reactividad cruzada entre los distintos grupos de flavivirus (28). Por ello, el blanco de interés seleccionado fue la proteína no estructural NS1, varios estudios como Roldan et al, (2021) emplean esta proteína para estudios experimentales como el desarrollo de ensayos inmunoenzimático, demostrando así el interés y viabilidad de la proteína (18).

Mientras que para la predicción de epítomos B, en el estudio se tomó en consideración únicamente los epítomos B lineales, excluyendo a los epítomos B conformacionales, debido a que los epítomos lineales pueden sintetizarse fácilmente y utilizarse en la producción de anticuerpos específicos (29). No obstante estudios como de Roy et al, (2025) desarrollaron un estudio donde detalla la predicción de epítomos tanto lineales como conformacionales en las proteínas E y NS1 (1). Por otro lado, Prasasty et al, (2019) quienes no solo trabajaron con células B sino también con células T, donde de igual manera realizó la predicción de epítomos de células B lineales y conformacionales. (30). Existen numerosos métodos para la predicción de los epítomos lineales, en el presente estudio se empleó ABCPred para desarrollarlo a comparación de los estudios de Prasasty y Roy, quienes emplearon BepiPred-2.0, IEBB, Discotope 2.0 y ElliPro entre otras redes neuronales.

Para el diseño del constructo multiepitopos, los criterios prioritarios considerados fueron: la selección de epítomos con adecuada capacidad antigénica, ausencia de

alergenicidad y de toxicidad. Para ello, se recuperó la secuencia nucleotídica completa con código de acceso MF438286.1, [Zika virus strain Zika virus/Homo sapiens/Cuba 2017] [polyprotein gene, complete cds], que posteriormente, se descargó en formato FASTA la secuencia aminoacídica traducida de la proteína NS1 con código de acceso ALL27019. A partir de esta secuencia se realizó un análisis mediante el servidor ABCPred, permitiendo identificar epítomos B lineales de 16 aminoácidos con mayor potencial inmunológico. Posteriormente, dichos epítomos fueron fragmentados en péptidos de 12 aa iniciales y finales, con el propósito de optimizar la precisión del filtrado inmunoinformático. Los péptidos obtenidos se evaluaron utilizando VaxiJen v2.0, AllerTOP v2.1, AlgPred y ToxinPred, seleccionándose únicamente aquellos clasificados como antigénicos, no alergénicos y no tóxicos. De los cuales solo 9 epítomos B lineales cumplieron estos criterios mencionados. En donde se demuestra que este constructo ensamblado se clasifica como antígeno probable, probablemente no alergénico y no tóxico. Estos hallazgos guardan concordancia con lo descrito por Antonelli et al, (2022), quienes reportaron constructos multiepitopos con elevada antigenicidad y ausencia de toxicidad (20). Asimismo, Hakami M, (2024) y Leal et al, (2025) documentaron perfiles inmunológicos favorables mediante estrategias inmunoinformáticas aplicadas al virus Zika (22,25).

En el análisis fisicoquímico se caracterizó mediante el servidor ProtParam de Expasy, evidenciando que la secuencia proteica tiene 155 aa, pI de 6.8, un peso molecular de 15841.53, la formula molecular $C_{681}H_{1060}N_{216}O_{213}S_5$ e un índice GRAVY de -0,977. Además, la vida media estimada de la vacuna artificial fue de 10 h, 30 min y 20 h para *Escherichia coli*, levaduras y reticulocitos de mamíferos, respectivamente. Estos hallazgos coinciden con lo reportado por Alharbi et al, (2024), Antonelli et al, (2022) y Hakami M, (2024) quienes también describieron perfiles fisicoquímicos estables y propiedades hidrofílicas favorables en constructos multiepitópicos diseñados mediante herramientas inmunoinformáticas aplicadas al virus Zika (20–22)

En cuanto a la validación estructural se empleó el servidor AlphaFold 2, para ver un modelo tridimensional del constructo, obteniéndose una conformación proteica estable y coherente, 5 modelos de Rank_1 al Rank_5 y la predicción IDDT. AlphaFold2 constituye una de las herramientas más precisas para predicción estructural basada en inteligencia artificial (24,26). Posteriormente, el refinamiento mediante GalaxyRefine

permitió mejorar la calidad conformacional del modelo. La validación estructural mediante SAVES mostró resultados favorables en Ramachandran Plot con el 98% de los residuos localizados en regiones favorables y un 2% en regiones adicionalmente permitidas y ERRAT de 51,3514, evidenciando adecuada estabilidad conformacional, calidad global y compatibilidad entre la secuencia aminoacídica y la estructura tridimensional predicha. Estos hallazgos coinciden con lo reportado por Alharbi et al, (2024) y Antonelli et al, (2022) en constructos multiepitopos diseñados mediante herramientas inmunoinformáticas similares (20,21).

La Herramienta de Adaptación de Codones de Java (JCat) se utilizó para optimizar los codones, considerando los parámetros como el índice de adaptación de codones (CAI), que presento un valor de 1.0 y el contenido de GC con un valor de 59,78 %. Estudios como de Alharbi et al, (2024), Hakami M, (2024) y Qiao et al, (2026), donde de igual manera coinciden con valores aceptables y considerados dentro de rango, es decir, estos resultados indican que la secuencia optimizada presenta características compatibles con una adecuada traducción y estabilidad genética, favoreciendo su posible producción experimental (21,22,31).

En cuanto a la reactividad cruzada, el constructo tras ser sometido a análisis de similitud mediante BLASTp frente a al constructo proteico de OROV, evidenció dos agrupaciones con alineamientos. El primer clúster correspondió a una secuencia descrita como FliC quimérico [construcción sintética], con puntuación máxima de 48,9, puntuación total de 48,9, cobertura de consulta de 62 %, valor E de 0,003, identidad de 40,82 %, longitud de acceso de 530 y adhesión UNP53947.1. El segundo clúster correspondió a un unnamed protein roducto, partial [Boreogadus s...], con puntuación máxima de 45,1, puntuación total de 45,1, cobertura de consulta de 86 %, valor E de 0,046, identidad de 30,71 %, longitud de acceso de 248 y adhesión CAL8236413.1. Mientras que en los constructos proteicos E1 y E2 de CHIKV, se observó una similitud de 21,00 %, con puntuación de 93,0 para E1 y para E2 se observó una similitud de 22,00 %, con puntuación de 105. En este contexto, Zimmerman et al, (2020) y Pereira et al, (2024) destacaron la importancia del diseño de proteínas multiepitópicas específicas para mejorar la discriminación serológica del virus Zika frente a otros flavivirus, respaldando la utilidad de estrategias inmunoinformáticas orientadas a incrementar la especificidad diagnóstica (14,23).

Finalmente, para el diseño plasmídico la secuencia optimizada resultante de JCat se clonó *in silico* usando Snapgene, y se expresó en el vector de expresión pET-28a (+) que contenía, kanamicina y MSC, generando el mapa plasmídico final mediante software especializado Benchling que presentó un pI de 6.80, un peso molecular de 15841.53 Da y Instability index de 37.12 siendo este estable. Este procedimiento permitió obtener un candidato antigénico validado computacionalmente y listo para futuras fases experimentales. La utilización del sistema pET28a (+) representa una estrategia frecuentemente empleada en estudios de expresión recombinante debido a su eficiencia, estabilidad y capacidad de inducir altos niveles de producción proteica. Como es el caso del estudio de Qiao et al, (2026), que implementaron el mismo vector, pero diferentes sitios de restricción HindIII y EcoRI, y con un enfoque distinto, dado que se empleó para el desarrollo de una vacuna multiepitópica contra *Pasteurella multocida* ovina, donde se obtuvo un constructo recombinante final de 6201 pb (31). Así como también, otros estudios como de Alharbi et al, (2024), que trabajó específicamente en ZIKV emplearon estrategias similares de clonación utilizando vectores de la familia pET y enzimas de restricción para la expresión recombinante de proteínas multiepitópicas, destacando la utilidad de estos sistemas para futuras aplicaciones inmunodiagnósticas y biotecnológicas (21).

3.3. VERIFICACIÓN DE HIPÓTESIS

Conforme a los resultados obtenidos en la presente investigación, evidenciados en la **Tabla 17**, se acepta la Hipótesis alternativa (H1), en la cual se planteó que el análisis inmunoinformático permitiría diseñar un constructo proteico multiepitópo del virus Zika con, ausencia o baja probabilidad de alergenicidad y ausencia de toxicidad, estabilidad estructural aceptable y baja similitud con regiones antigénicas de virus clínicamente relacionados, respaldando su potencial aplicación en inmunodiagnóstico diferencial. Y, por ende, se rechaza la hipótesis nula (H0), la cual era la negación de la H1.

Tabla 17: Criterios de cumplimiento o rechazo de la Hipótesis Alternativa (H1), en base al diseño constructo proteico multiepitopos de ZIKV generado en la presente investigación.

Criterio de H1	Resultado obtenido	Cumple
----------------	--------------------	--------

Construceto multiepítopo	Se diseñó un construceto de 155 aa, en base a 9 epítomos B lineales junto con enlazadores GPGPG, una metionina y una cola de 6 histidinas.	Sí
Antigenicidad predictiva favorable	Mediante el servidor VaxiJen v2.0 se obtuvo una puntuación de 0.7422 clasificándolo como Antígeno probable.	Sí
Ausencia o baja probabilidad de alergenicidad	Se predijo como Probablemente no alergénico mediante el servidor AllerTOP v2.1.	Sí
Ausencia de toxicidad	Mediante el servidor ToxinPred, se determinó que era No toxico	Sí
Estabilidad estructural aceptable	Índice de inestabilidad obtenido 37.12 (<40)	Sí
	Peso molecular de 15841,53 Da	Sí
	Punto Isoeléctrico de 6.80	Sí
	Factor de calidad de ERRAT de 51.354	Sí
	Ramachandran Plot: con el 98% de los residuos están localizados en regiones favorecidas	Sí
Baja similitud con regiones antigénicas de virus clínicamente relacionados	CHIKV, E1: similitud de 21,00% con puntuación de 93,0.	Sí
	CHIKV, E2: similitud de 22,00 %, con puntuación de 105	
	OROV: similitud de 26,00 %, con puntuación de 109.	Sí

Elaborado por: Carmen Elena Balseca Quishpe

CAPITULO IV

CONCLUSIONES Y RECOMENDACIONES

4.1. CONCLUSIONES

- Se identificaron regiones antigénicas conservadas del virus Zika a partir de la poliproteína con el código de acceso ALL27019 mediante herramientas inmunoinformáticas. Los análisis de similitud realizados frente a la proteína E1 y E2 del CHIKV y la proteína de OROV, evidenciaron porcentajes de similitud bajos, siendo de 21,00 % con una puntuación de 93,0 para la proteína E1, el 22,00 % con una puntuación de 105 y el 26,00 % con una puntuación de 109 para OROV, indicando una reducida homología con estos arbovirus y respaldando la especificidad del constructo para disminuir el riesgo de reactividad cruzada en aplicaciones diagnósticas.
- Se predijeron y seleccionaron epítomos B lineales del virus Zika con alto potencial antigénico mediante herramientas inmunoinformáticas especializadas, los cuales fueron filtrados considerando criterios de seguridad biológica. Los análisis realizados permitieron descartar secuencias con potencial alergénico o tóxico, asegurando la selección de epítomos candidatos adecuados para la construcción del diseño multiepítomo, con características favorables para su posible aplicación en el desarrollo de estrategias inmunodiagnósticas.
- Se diseñó y validó *in silico* un constructo multiepítomo de 155 aa mediante el ensamblaje de los epítomos seleccionados utilizando enlazadores GPGPG. Su distribución fue con los de mayor puntuación de antigenicidad en el centro y los de menor puntuación a los extremos. El constructo presentó una puntuación de antigenicidad de 0,7422, fue predicho como no tóxico y probablemente no alergénico, además de exhibir propiedades fisicoquímicas compatibles con estabilidad proteica y una conformación estructural aceptable, evidenciada por un índice de inestabilidad de 37,12, un factor de calidad ERRAT de 51,3514 y un 98 % de residuos localizados en regiones favorecidas del gráfico de

Ramachandran. Asimismo, la optimización génica presento un CAI de 1,0 y contenido GC de 59,78%, parámetros compatibles con una potencial expresión heteróloga en *Escherichia coli*.

- En función de los resultados obtenidos, se concluyó que la integración de herramientas de inmunoinformática y biología computacional permitieron diseñar y validar *in silico* una proteína multiepitopo específica del virus Zika, compuesta por epítopos B lineales seleccionados por características predictivas favorables de antigenicidad, seguridad biológica, estabilidad estructural y especificidad molecular, respaldando su potencial aplicación en futuras investigaciones orientadas al desarrollo de inmunodiagnóstico diferencial.

4.2.RECOMENDACIONES

- Considerando que el presente trabajo de investigación se desarrolló desde un enfoque *in silico* empleando herramientas inmunoinformáticas, se recomienda realizar la síntesis del gen optimizado y su posterior clonación en el vector pET28(+), con la finalidad de validar experimentalmente el diseño del constructo propuesto y así dar continuidad en la producción recombinante de una proteína multiepitopos.
- Para validar la antigenicidad y especificidad del constructo diseñado mediante inmunoensayos empleando muestras biológicas, se recomienda utilizar sueros positivos para ZIKV, así como para otros arbovirus como OROV y CHIKV, con la finalidad de determinar la capacidad del constructo de ser reconocido por anticuerpos dirigidos contra ZIKV y verificar si existen reacciones cruzas con los otros arbovirus.
- Se recomienda evaluar experimentalmente la expresión y estabilidad del constructo en sistemas heterólogos basados en *Escherichia coli* con la finalidad de validar los parámetros favorables obtenidos durante los análisis bioinformáticos y corroborar su factibilidad para la producción de una proteína recombinante en condiciones experimentales.

REFERENCIAS BIBLIOGRÁFICAS

1. Roy RR, Tadmalkar N, Deshpande GR, Atre NM, Shil P, Sapkal G. Identification of B-cell epitopes of Indian Zika virus strains using immunoinformatics. *Front Immunol.* 27 de febrero de 2025;16. doi:10.3389/fimmu.2025.1534737 PubMed PMID: 40083545.
2. Feng Y. Recent advances in the study of zika virus structure, drug targets, and inhibitors. *Front Pharmacol.* 1 de julio de 2024;15. doi:10.3389/FPHAR.2024.1418516/FULL
3. Rabe IB, Hills SL, Haussig JM, Walker AT, Santos T dos, Martin JLS, et al. A Review of the Recent Epidemiology of Zika Virus Infection. *Am J Trop Med Hyg.* 7 de mayo de 2025;112(5):1026-35. doi:10.4269/AJTMH.24-0420 PubMed PMID: 39933180.
4. Pan American Health Organization [PAHO]. Zika: Análisis por país [Internet]. 2026 [citado 19 de mayo de 2026]. Disponible en: <https://www.paho.org/es/arbo-portal/zika-datos-analisis/zika-analisis-por-pais>
5. World Health Organization [WHO]. World Health Organization [WHO] [Internet]. 2025 [citado 6 de mayo de 2026]. Virus de Zika. Disponible en: <https://www.who.int/es/news-room/fact-sheets/detail/zika-virus>
6. Ministerio de Salud Pública. ENFERMEDADES TRANSMITIDAS POR VECTORES [Internet]. marzo de 2025 [citado 19 de mayo de 2026]. Disponible en: <https://www.salud.gob.ec/wp-content/uploads/2025/03/ENFERMEDADES-TRANSMITIDAS-POR-VECTORES-SE-10-2025-OK.pdf>
7. Rawat SS, Keshri AK, Kaur R, Prasad A. Immunoinformatics Approaches for Vaccine Design: A Fast and Secure Strategy for Successful Vaccine Development. *Vaccines* 11(2), 221. 18 de enero de 2023;11(2). doi:10.3390/vaccines11020221
8. Gomase VS, Sharma R, Dhamane SP. Immunoinformatics Approach for Optimization of Targeted Vaccine Design: New Paradigm in Clinical Trials and

- Healthcare Management. *Rev Recent Clin Trials*. 18 de julio de 2025;20. doi:10.2174/0115748871374235250702065617 PubMed PMID: 40676801.
9. Parvizpour S, Pourseif MM, Razmara J, Rafi MA, Omid Y. Epitope-based vaccine design: a comprehensive overview of bioinformatics approaches. *Drug Discov Today*. 1 de junio de 2020;25(6):1034-42. doi:10.1016/j.drudis.2020.03.006 PubMed PMID: 32205198.
 10. Giraldo MI, Gonzalez-Orozco M, Rajsbaum R. Pathogenesis of Zika Virus Infection. *Annual Review of Pathology: Mechanisms of Disease*. 24 de enero de 2023;18(Volume 18, 2023):181-203. doi:10.1146/annurev-pathmechdis-031521-034739 PubMed PMID: 36151059.
 11. Pan American Health Organization [PAHO]. Actualización Epidemiológica - Dengue, chikunguña y Zika. Pan American Health Organization [PAHO] [Internet]. junio de 2023 [citado 6 de mayo de 2026]. Disponible en: <https://www.paho.org/es/documentos/actualizacion-epidemiologica-dengue-chikunguna-zika-10-junio-2023>
 12. Lunardelli VAS, Apostolico JDS, Fernandes ER, Santoro Rosa D. Zika virus— an update on the current efforts for vaccine development. *Human Vaccines e Immunotherapeutics*. 4 de marzo de 2021;17(3):904-8. doi:10.1080/21645515.2020.1796428 PubMed PMID: 32780659.
 13. Wang Y, Ling L, Zhang Z, Marin-Lopez A. Current Advances in Zika Vaccine Development. *Vaccines*, 10 (11), 1816. 27 de octubre de 2022;10(11). doi:10.3390/VACCINES10111816
 14. Zimmerman MG, Wrammert J, Suthar MS. Cross-Reactive Antibodies during Zika Virus Infection: Protection, Pathogenesis, and Placental Seeding. *Cell Host Microbe*. 8 de enero de 2020;27(1):14-24. doi:10.1016/j.chom.2019.12.003 PubMed PMID: 31917957.
 15. Katzelnick LC, Zambrana JV, Elizondo D, Collado D, Garcia N, Arguello S, et al. Dengue and Zika virus infection in children elicit cross-reactive protective and enhancing antibodies that persist long term. *Sci Transl Med*. 6 de octubre

- de 2021;13(614):eabg9478. doi:10.1126/scitranslmed.abg9478 PubMed PMID: 34613812.
16. Rodriguez-Morales AJ, Drexler JF. Re-emergence of Oropouche virus in Brazil and Latin America. *Lancet Infect Dis.* 1 de febrero de 2025;25(2):137-9. doi:10.1016/S1473-3099(24)00672-8 PubMed PMID: 39423836.
 17. Pan American Health Organization [PAHO]. Epidemiological Update for Dengue and other Arboviruses. Pan American Health Organization [PAHO] [Internet]. enero de 2026 [citado 6 de mayo de 2026]. Disponible en: https://ais.paho.org/ha_viz/Arbo/Arbo_Bulletin_2024.asp?env=pri
 18. Roldan JS, Cassola A, Castillo DS. Development of a novel NS1 competitive enzyme-linked immunosorbent assay for the early detection of Zika virus infection. *PLoS One.* 1 de agosto de 2021;16(8):e0256220. doi:10.1371/JOURNAL.PONE.0256220 PubMed PMID: 34403457.
 19. Yu L, Liu X, Ye X, Su W, Zhang X, Deng W, et al. Monoclonal antibodies against zika virus ns1 protein confer protection via fccreceptor-dependent and-independent pathways. *mBio.* 1 de enero de 2021;12(1):1-18. doi:10.1128/MBIO.03179-20;SUBPAGE:STRING:FULL PubMed PMID: 33563822.
 20. Antonelli ACB, Almeida VP, de Castro FOF, Silva JM, Pfrimer IAH, Cunha-Neto E, et al. In silico construction of a multiepitope Zika virus vaccine using immunoinformatics tools. *Scientific Reports* 12(1). 7 de enero de 2022;12(1):53-. doi:10.1038/s41598-021-03990-6 PubMed PMID: 34997041.
 21. Alharbi M, Alshammari A, Alsabhan JF, Alzarea SI, Alshammari T, Alasmari F, et al. A novel vaccine construct against Zika virus fever: insights from epitope-based vaccine discovery through molecular modeling and immunoinformatics approaches. *Front Immunol.* 2024;15. doi:10.3389/fimmu.2024.1426496 PubMed PMID: 39050858.
 22. Hakami MA. An immunoinformatics and structural vaccinology approach to design a novel and potent multi-epitope base vaccine targeting Zika virus. *BMC*

Chemistry 18(31). 13 de febrero de 2024;18(1):31-. doi:10.1186/S13065-024-01132-3

23. Pereira SH, Sá Magalhães Serafim M, Moraes T de FS, Zini N, Abrahão JS, Nogueira ML, et al. Design, development, and validation of multi-epitope proteins for serological diagnosis of Zika virus infections and discrimination from dengue virus seropositivity. *PLoS Negl Trop Dis*. 1 de abril de 2024;18(4):e0012100. doi:10.1371/JOURNAL.PNTD.0012100 PubMed PMID: 38635656.
24. Valotto RS, Júnior PCML, Machado-de-Ávila R, Campos F V., Costa DS, Guimarães MCC, et al. Consensus-based computational mapping of NS1 antigenic determinants in dengue and Zika viruses to improve diagnostic specificity. *BMC Infect Dis*. 1 de diciembre de 2026;26(1):380-380. doi:10.1186/s12879-025-12291-6 PubMed PMID: 41580589.
25. Leal LRS, de Sena MGAM, Invenção M da CV, de Moura IA, Jesus ALS de, de Sousa GF, et al. Development and Immunogenicity Assessment of a Multi-Epitope Antigen Against Zika Virus: An In Silico and In Vivo Approach. *Vaccines (Basel)*. 26 de diciembre de 2025;14(1):31. doi:10.3390/vaccines14010031
26. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 26 de agosto de 2021;596(7873):583-9. doi:10.1038/s41586-021-03819-2 PubMed PMID: 34265844.
27. José de Jesús LR, Lucía OC, Yolanda Margarita SC, Jorge Joel RM. Uso de la inmunoinformática en el diseño de vacunas basadas en epítomos. *Sociedades Rurales, Producción y Medio Ambiente [Internet]*. 30 de julio de 2025 [citado 27 de mayo de 2026];25(49):97-126. Disponible en: <https://sociedadesruralesojs.xoc.uam.mx/index.php/srpma/article/view/539>
28. Fisher R, Lustig Y, Sklan EH, Schwartz E. The Role of NS1 Protein in the Diagnosis of Flavivirus Infections. *Viruses* 2023, Vol 15,. 18 de febrero de 2023;15(2). doi:10.3390/V15020572 PubMed PMID: 36851784.

29. Ras-Carmona A, Lehmann AA, Lehmann P V., Reche PA. Prediction of B cell epitopes in proteins using a novel sequence similarity-based method. *Scientific Reports* 2022 12:1. 12 de agosto de 2022;12(1):13739-. doi:10.1038/s41598-022-18021-1 PubMed PMID: 35962028.
30. Prasasty VD, Grazzolie K, Rosmalena R, Yazid F, Ivan FX, Sinaga E. Peptide-Based Subunit Vaccine Design of T- and B-Cells Multi-Epitopes against Zika Virus Using Immunoinformatics Approaches. *Microorganisms*. 1 de agosto de 2019;7(8):226. doi:10.3390/MICROORGANISMS7080226 PubMed PMID: 31370224.
31. Qiao Y, Wu A, Li H, Zhuang Y, Fu Q, Yang L, et al. Design of a Multi-Epitope Vaccine Against *Ovine Pasteurella multocida* Using Immunoinformatics Strategies. *Microorganisms* 2026, Vol 14,. 12 de marzo de 2026;14(3). doi:10.3390/MICROORGANISMS14030656

ANEXOS