

Explorando la precisión de AlphaFold en la predicción de estructuras de α -conotoxinas frente a referencias del PDB

Vanessa Martinez Pino

Trabajo de Grado para Optar el Título de Química

Director

Hermisul de Jesús Cano Calle

Químico, PhD

Universidad Industrial de Santander

Facultad de Ciencias

Escuela de Química

Bucaramanga

2026

Agradecimientos

A mi madre por impulsarme en mi formación profesional y acompañarme en cada paso de este camino, y a mi padre, por estar siempre presente y brindarme su respaldo.

A mi director de tesis por su orientación y apoyo durante todo este proceso. Su conocimiento, dedicación y acompañamiento fueron fundamentales para la realización y culminación de este trabajo.

A cada uno de los profesores que conocí durante mi carrera, quienes actuaron como guías y me brindaron el conocimiento necesario para llegar a este punto.

A la Universidad Industrial de Santander por ser mi segundo hogar y espacio de aprendizaje.

Al programa Generación E Excelencia del Gobierno Nacional; gracias a esta oportunidad, hoy puedo ver cumplida esta meta profesional.

A todas aquellas personas que me brindaron la motivación necesaria para perseverar y culminar con éxito esta etapa de mi vida.

Tabla de Contenido

	Pág.
Introducción	12
1. Objetivos	14
1.1. Objetivo General	14
1.2. Objetivos Específicos	14
2. Marco Referencial	15
2.1. Antecedentes	15
2.2. Marco Teórico	17
2.2.1. <i>Las α-conotoxinas</i>	17
2.2.2. <i>Plegamiento de péptidos y el problema computacional</i>	21
2.2.3. <i>Evolución de AlphaFold</i>	23
2.2.4. <i>Bases de datos estructurales</i>	25
2.3. Marco Conceptual	28
3. Metodología	32
3.1. Selección de α -conotoxinas modelo	32
3.2. Generación de modelos estructurales con AlphaFold 3.0	33
3.3. Obtención de modelos comparativos de AFDB y estructuras del PDB	33
3.4. Comparación estructural y análisis cuantitativo	33
3.5. Interpretación de resultados y evaluación del desempeño	34
4. Resultados	35
5. Discusión	48

6. Conclusiones	57
7. Recomendaciones	59
Referencias Bibliográficas	61

Lista de Tablas

	Pág.
Tabla 1. Regiones de los modelos de AlphaFold DataBase conservadas para la obtención de los péptidos maduros empleados en la comparación estructural.....	32
Tabla 2. Características estructurales y de secuencia de las α -conotoxinas seleccionadas para el análisis comparativo.....	33
Tabla 3. Métricas de desempeño estructural y niveles de confianza de AlphaFold 3.0 frente a registros experimentales y predicciones de AlphaFold Database para las alfa-conotoxinas seleccionadas.....	34

Lista de Figuras

	Pág.
Figura 1. Superposiciones estructurales de la α -conotoxina EI. (A) AF3 (azul) vs PDB (verde), (B) AF3 (azul) vs AFDB (naranja), (C) AFDB (naranja) vs PDB (verde).....	36
Figura 2. Superposiciones estructurales de la α -conotoxina GI. (A) AF3 (azul) vs PDB (verde), (B) AF3 (azul) vs AFDB (naranja), (C) AFDB (naranja) vs PDB (verde).....	37
Figura 3. Superposiciones estructurales de la α -conotoxina GIC. (A) AF3 (azul) vs PDB (verde), (B) AF3 (azul) vs AFDB (naranja), (C) AFDB (naranja) vs PDB (verde).....	38
Figura 4. Superposiciones estructurales de la α -conotoxina ImI. (A) AF3 (azul) vs PDB (verde), (B) AF3 (azul) vs AFDB (naranja), (C) AFDB (naranja) vs PDB (verde).....	39
Figura 5. Superposiciones estructurales de la α -conotoxina MII. (A) AF3 (azul) vs PDB (verde), (B) AF3 (azul) vs AFDB (naranja), (C) AFDB (naranja) vs PDB (verde).....	40
Figura 6. Superposiciones estructurales de la α -conotoxina PnIA. (A) AF3 (azul) vs PDB (verde), (B) AF3 (azul) vs AFDB (naranja), (C) AFDB (naranja) vs PDB (verde).....	41
Figura 7. Superposiciones estructurales de la α -conotoxina RgIA. (A) AF3 (azul) vs PDB (verde), (B) AF3 (azul) vs AFDB (naranja), (C) AFDB (naranja) vs PDB (verde).....	42
Figura 8. Superposiciones estructurales de la α -conotoxina SI. (A) AF3 (azul) vs PDB (verde), (B) AF3 (azul) vs AFDB (naranja), (C) AFDB (naranja) vs PDB (verde).....	43
Figura 9. Superposiciones estructurales de la α -conotoxina Vc1.1. (A) AF3 (azul) vs PDB (verde), (B) AF3 (azul) vs AFDB (naranja), (C) AFDB (naranja) vs PDB (verde).....	44
Figura 10. Superposiciones estructurales de la α -conotoxina TxIB. (A) AF3 (azul) vs PDB (verde), (B) AF3 (azul) vs AFDB (naranja), (C) AFDB (naranja) vs PDB (verde).....	45

Lista de Abreviaturas

AF2: AlphaFold 2.0

AF3: AlphaFold 3.0

AFDB: AlphaFold DataBase

CASP: Critical Assessment of Protein Structure Prediction (evaluación Crítica de la Predicción de Estructuras)

KO: Knock-out

MSAs: Multiple Sequence Alignments (Alineaciones de secuencias múltiples)

nAChR: nicotinic Acetylcholine Receptor (receptor nicotínico de acetilcolina)

PDB: Protein Data Bank

pLDDT: Predicted Local Distance Difference Test (Prueba de diferencia de distancia local predicha)

RMN: Resonancia magnética nuclear

RMSD: Root Mean Square Deviation (Desviación cuadrática media)

Glosario

AlphaFold: sistema de inteligencia artificial desarrollado por DeepMind que predice la estructura tridimensional de proteínas a partir de su secuencia de aminoácidos. Utiliza información de MSA y otras características para generar modelos altamente precisos.

CASP: evaluación experimental internacional de métodos de predicción de estructuras de proteínas, realizada cada dos años desde 1994.

ConoServer: base de datos bioinformática especializada en el estudio de conotoxinas (conopéptidos), toxinas producidas por caracoles marinos del género *Conus*.

Método ab initio (o modelado de novo): enfoque computacional para predecir la estructura tridimensional de una proteína utilizando únicamente su secuencia de aminoácidos, sin depender de estructuras conocidas como referencia.

Modelo Knock-out (KO): organismo modificado genéticamente (comúnmente ratones o células) en el que uno o más genes han sido inactivados, permitiendo estudiar su función mediante los efectos de su ausencia.

MSA: conjunto de secuencias de proteínas alineadas para identificar regiones conservadas o similares. Es una herramienta clave utilizada por AlphaFold para mejorar la predicción estructural.

pLDDT: métrica empleada por AlphaFold para estimar la confianza en la predicción estructural a nivel de residuos individuales. Sus valores oscilan entre 0 y 100, donde puntuaciones más altas indican mayor confiabilidad.

PyMOL: software de visualización molecular basado en Python, utilizado para representar, manipular y analizar estructuras tridimensionales de biomoléculas.

RMSD: medida de la desviación promedio entre átomos equivalentes al comparar dos estructuras. Valores cercanos a 0 Å indican alta similitud; valores mayores a 2–3 Å sugieren diferencias estructurales significativas.

UniProt (Universal Protein Resource): base de datos global que proporciona información detallada sobre secuencias de proteínas y sus funciones biológicas.

Resumen

Título: Explorando la precisión de AlphaFold en la predicción de estructuras de α -conotoxinas frente a referencias del PDB*

Autor: Vanessa Martinez Pino**

Palabras Clave: AlphaFold 3.0, α -conotoxinas, PDB, predicción estructural, bioinformática.

Descripción: Las α -conotoxinas, una subfamilia de las conotoxinas, son pequeños péptidos derivados del veneno del caracol marino *Conus* y destacan por su interés farmacológico como antagonistas de canales iónicos y receptores neuronales. Sin embargo, su estudio plantea desafíos debido a su reducido tamaño y la presencia de puentes disulfuro, que son esenciales para su estabilidad y función biológica. En este trabajo se evaluó la precisión de AlphaFold 3.0 (AF3) para predecir la estructura de diez α -conotoxinas, comparando los resultados con estructuras experimentales del Protein Data Bank (PDB) y modelos previos de la AlphaFold Database (AFDB). Los resultados muestran que AF3 alcanza una precisión sobresaliente, con valores de RMSD inferiores a 1,0 Å en el 80% de los casos, respaldados por altos niveles de confianza (pLDDT mayores a 92). Un hallazgo especialmente relevante fue la capacidad de AF3 para corregir errores en la formación de puentes disulfuro presentes en modelos anteriores, logrando mayor exactitud química en toxinas como EI, GI, ImI, MII, PnIA y Vc1.1. Además, el análisis de la α -conotoxina RgIA (RMSD: 3,051 Å) permitió identificar distintas formas conformacionales, lo que sugiere que AF3 tiende a predecir estados de mínima energía compatibles con la flexibilidad observada experimentalmente mediante resonancia magnética nuclear (RMN). En conclusión, AF3 supone un avance significativo en la predicción de la estructura de péptidos pequeños y con alto contenido de cisteínas. Los resultados obtenidos en este trabajo avalan el uso de herramientas bioinformáticas de última generación para el modelado de novo de toxinas marinas, y muestran su utilidad para apoyar estudios estructurales y funcionales posteriores. Además, la capacidad de AF3 para replicar fielmente las características estructurales clave de las α -conotoxinas destaca el valor de estos enfoques para futuras investigaciones en áreas como la química computacional y la biología estructural.

* Trabajo de Grado

**Facultad de Ciencias. Escuela de Química. Director Herminul de Jesús Cano Calle, PhD en Química.

Abstract

Title: Exploring the accuracy of AlphaFold in predicting the structures of α -conotoxins against PDB references*

Author: Vanessa Martinez Pino**

Key Words: AlphaFold 3.0, α -conotoxins, PDB, structural prediction, bioinformatics.

Description: α -Conotoxins, a subfamily of conotoxins, are small peptides derived from the venom of the marine snail *Conus* and are of particular pharmacological interest as antagonists of ion channels and neuronal receptors. However, studying them poses challenges due to their small size and the presence of disulphide bonds, which are essential for their stability and biological function. In this study, the accuracy of AlphaFold 3.0 (AF3) in predicting the structure of ten α -conotoxins was evaluated, comparing the results with experimental structures from the Protein Data Bank (PDB) and previous models from the AlphaFold Database (AFDB). The results show that AF3 achieves outstanding accuracy, with RMSD values below 1,0 Å in 90% of cases, supported by high confidence levels (pLDDT values greater than 92). A particularly relevant finding was AF3's ability to correct errors in disulphide bond formation present in previous models, achieving greater chemical accuracy in toxins such as EI, GI, ImI, MII, PnIA and Vc1.1. Furthermore, analysis of the α -conotoxin RgIA (RMSD: 3,051 Å) allowed the identification of different conformational forms, suggesting that AF3 tends to predict minimum-energy states consistent with the flexibility observed experimentally via nuclear magnetic resonance (NMR). In conclusion, AF3 represents a significant advance in the prediction of the structure of small, cysteine-rich peptides. The results obtained in this study support the use of state-of-the-art bioinformatics tools for the de novo modelling of marine toxins, and demonstrate their usefulness in supporting subsequent structural and functional studies. Furthermore, AF3's ability to faithfully replicate the key structural features of α -conotoxins highlights the value of these approaches for future research in areas such as computational chemistry and structural biology.

* Bachelor Thesis

**Facultad de Ciencias. Escuela de Química. Director Herminul de Jesús Cano Calle, PhD in Chemistry.

Introducción

Las conotoxinas, péptidos ricos en enlaces de disulfuro provenientes del veneno del caracol marino *Conus*, han provocado un gran interés farmacológico debido a su capacidad de interactuar selectivamente con canales iónicos y receptores neuronales. Especialmente, las α -conotoxinas actúan como antagonistas de los receptores nicotínicos de acetilcolina (nAChR) demostrando su potencial en el desarrollo de fármacos. La estabilidad y actividad biológica de estos péptidos está asociada a la estructura tridimensional que adquieren debido a los enlaces de disulfuro presentes (Li et al., 2025). Por lo tanto, la determinación de las estructuras de las α -conotoxinas es importante para el estudio de los mecanismos de interacción y para la investigación y el desarrollo de fármacos.

Tradicionalmente, los métodos usados para el estudio de las estructuras proteicas han sido la cristalografía de rayos X y la espectroscopia de resonancia magnética nuclear (RMN). En el primero, la precisión de la estructura atómica depende de la calidad de los cristales formados; al cristalizar una proteína rígida, con frecuencia, se obtienen cristales de buena calidad, por el contrario, al cristalizar una proteína flexible los cristales que se forman son de mala calidad o resulta difícil que se formen. En cambio, el segundo método analiza proteínas pequeñas en solución y la estructura atómica obtenida es un conjunto de estructuras proteicas donde serán similares en las regiones rígidas y serán diferentes en las regiones flexibles (Berman et al., 2000). Sin embargo, obtener estructuras tridimensionales por métodos experimentales como la cristalografía de rayos X o la resonancia magnética nuclear implica retos técnicos considerables. En cristalografía, la dificultad para conseguir cristales adecuados puede limitar el análisis de ciertos péptidos, mientras que la RMN exige equipos especializados y su aplicabilidad depende de las propiedades de la molécula en estudio. En este escenario, los avances en aprendizaje automático han cobrado cada vez mayor importancia en la biología

estructural, ya que permiten generar modelos tridimensionales de manera rápida, complementando y guiando los estudios experimentales. Con la aparición de AlphaFold, la predicción de estas estructuras ha sido una tarea mas fácil, marcando una brecha entre los métodos tradicionales con el uso de AlphaFold 2 (AF2), al demostrar con el experimento CASP14 una predicción comparable con métodos experimentales como la cristalografía de rayos X (Jumper et al., 2021). Luego de la aparición de AlphaFold 3.0 (AF3), se logró avanzar en la biología estructural, siendo este capaz de predecir la estructura de complejos compuestos por proteínas, ácidos nucleicos, moléculas pequeñas, iones y residuos (Abramson et al., 2024). La aparición de estas herramientas ha sido fundamental para la clasificación y caracterización de conotoxinas beneficiando a la investigación biomédica.

Independientemente de la información disponible en la base de datos de AlphaFold (AFDB), hay limitaciones que ponen en duda su utilidad para el estudio de α -conotoxinas. Se han identificado que las predicciones muchas veces están basadas en secuencias genómicas que no representan a la conotoxina dando como resultado estructuras tridimensionales de biomoléculas que no existen en la naturaleza. Estas inconsistencias menoscaban la fiabilidad de la base de datos, especialmente, teniendo en cuenta que la actividad biológica de estas toxinas está relacionada directamente con su estructura tridimensional y, en particular, con sus puentes de disulfuro. En consecuencia, resulta esencial evaluar si AF3 logra subsanar esas deficiencias. Por lo tanto, es necesario comparar estas predicciones con las estructuras que se encuentran en el Protein Data Bank (PDB) para así determinar si esta nueva versión puede predecir con precisión el plegamiento y la conectividad de disulfuros en contraste con los datos registrados en AFDB.

1. Objetivos

1.1. Objetivo General

Evaluar la precisión estructural de los modelos generados por AF3 para α -conotoxinas seleccionadas mediante la comparación con estructuras experimentales disponibles en el PDB y modelos computacionales previos depositados en la AFDB.

1.2. Objetivos Específicos

Obtener y generar modelos tridimensionales de las α -conotoxinas seleccionadas utilizando AF3, asegurando condiciones de predicción controladas y coherentes para el análisis comparativo.

Recopilar y analizar modelos correspondientes de las mismas toxinas disponibles en AFDB para establecer una base de comparación con versiones anteriores del algoritmo (AlphaFold 2).

Comparar cuantitativamente las estructuras generadas por AF3 con sus correspondientes referencias experimentales y modelos previos, utilizando métricas como: RMSD global y por dominio estructural, identificación y conservación de puentes disulfuro y pLDDT mediante herramientas como PyMol.

Analizar el desempeño de AF3 en función de variables estructurales y bioquímicas, como el tamaño del péptido, la cantidad de puentes de disulfuro, la presencia de hélices alfa y láminas beta, y la complejidad conformacional global.

Determinar las fortalezas y limitaciones de AF3 para predecir estructuras de péptidos pequeños como las α -conotoxinas, generando conclusiones aplicables a estudios futuros en diseño racional de toxinas, farmacología molecular y química computacional.

2. Marco Referencial

2.1. Antecedentes

Recientemente, la comunidad científica se ha enfocado en establecer si los modelos de inteligencia artificial, como AF2, aplican reglas biofísicas o si solo repiten patrones estructurales conocidos. De acuerdo a los hallazgos más recientes, AF2 opera mediante una función de energía biofísica aprendida, la cual le permite explorar el espacio conformacional de manera análoga a un método *ab initio* cuando no se tienen los alineamientos de secuencias múltiples (MSAs). Según Chang y Perez (2025), a partir de este modelo surge un mecanismo de plegamiento al que se le llama “local-first, global-later” (primero lo cercano, luego lo global), en el cual se asignan las interacciones entre los residuos más cercanos en la secuencia, seguido de un ordenamiento en la estructura global.

Aun así, surgen cuestionamientos con esta perspectiva basada en la prioridad local sobre la precisión en péptidos pequeños con alta restricción conformacional, como las α -conotoxinas. La estabilidad de estas moléculas proviene no solo de las interacciones locales sino también de la arquitectura de los puentes disulfuro que incluyen residuos alejados.

Dependiendo de como se formen sus puentes de disulfuro, las α -conotoxinas pueden adoptar tres isómeros: globular, cinta (ribbon) y cuenta (beaded). En la mayoría de α -conotoxinas caracterizadas hasta el momento, el isómero globular (Cys I - Cys III, Cys II - Cys IV) corresponde a la conformación nativa y exhibe la mayor actividad biológica, mostrando antagonismo selectivo sobre distintos subtipos de nAChRs (Akondi et al., 2014; Jin et al., 2019). Debido a esto, este isómero ha sido considerado relevante y resulta útil como referencia para evaluar a AF3 como una herramienta de predicción estructural. No obstante, existen excepciones en donde otras conectividades presentan la mayor actividad biológica, por encima del isómero globular, con la conformación de cinta (Giglio et al., 2025).

Por su parte, AF3 introduce un modelo generativo de difusión que permite incorporar ligandos y ciertas modificaciones post-traduccionales en un mismo proceso de predicción estructural (Abramson et al., 2024). El pasar de un modelo basado en procesamiento de secuencias a uno de difusión, constituye un avance respecto a versiones previas ya que permite incluir un mayor rango de entidades moleculares en el modelado. No obstante, la precisión para representar modificaciones postraduccionales sigue teniendo ciertas limitaciones y aún requiere evaluación adicional. En este sentido, la incorporación de modelos de difusión abre posibilidades interesantes para describir interacciones biomoleculares complejas, aunque es importante analizar su desempeño en función de cada sistema específico y de la evidencia experimental existente (Ramasamy et al., 2025).

Por otro lado, mientras que el PDB registra estructuras determinadas por métodos consistentes como Rayos X y RMN (Berman et al., 2000), la AFDB arroja predicciones basadas en secuencias genómicas que no coinciden con el péptido maduro generando una brecha de datos. Esta divergencia resulta crucial debido a que la actividad biológica de las α -conotoxinas se expresa en el péptido maduro (Mao et al., 2026). En consecuencia, los modelos se deben enfocar, especialmente, en la fidelidad de la conectividad de los puentes disulfuro (Akdal et al., 2022).

Por lo anterior, es de gran importancia analizar la precisión de AF3 frente a las referencias experimentales del PDB. Esto permitirá validar si este cambio generativo logra mitigar los sesgos de plegamiento local observados en AF2 y, a su vez, logra subsanar las inconsistencias en las bases de datos, otorgando fiabilidad a este nuevo modelo en la conectividad de disulfuros con aplicabilidad en la investigación farmacológica de conotoxinas.

2.2. Marco Teórico

2.2.1. *Las α -conotoxinas*

Los caracoles marinos del género *Conus* son depredadores con gran diversidad que han desarrollado una estrategia de caza basada en la inyección un cóctel de toxinas peptídicas. Principalmente, existen tres especies según su presa: caracoles que se alimentan de peces, caracoles que se alimentan de gusanos y caracoles que se alimentan de moluscos, siendo el grupo de los que se alimentan de gusanos el mas extenso, representando el 72% de todas las especies conocidas de este genero (Florice et al., 2025). Estos caracoles utilizan un diente radular hueco especializado para inocular el veneno en sus presas (Akondi et al., 2014).

En 1960 se reconoce a este veneno como una mezcla compleja de componentes biológicamente activos, en particular, péptidos neuroactivos denominados conotoxinas. Estas moléculas se destacan por su potencia y selectividad incomparables hacia una amplia variedad de nAChRs y canales iónicos, lo que las convierte en herramientas fundamentales para la investigación biomédica y el desarrollo de agentes terapéuticos emergentes (Akondi et al., 2014; Jin et al., 2019).

Las conotoxinas son pequeños péptidos, compuestos por entre 10 y 40 residuos de aminoácidos, con una alta proporción de cisteínas que favorecen la formación de enlaces disulfuro. Estos enlaces son fundamentales, ya que aportan estabilidad estructural, resistencia a la degradación proteolítica y definen la actividad biológica de la toxina (Li et al., 2025).

La síntesis de estos péptidos se ve mediada por una secuencia de transcripción que consta de tres regiones: el péptido señal, el pre-pro-péptido y el péptido maduro. La identificación de estas regiones resulta importante para determinar la funcionalidad, síntesis y propiedades de orientación de la conotoxina.

El péptido señal se origina en el ribosoma, se encuentra ubicado en el extremo N-

terminal, actúa como una etiqueta celular que guía al péptido hacia el retículo endoplasmático antes de ser eliminado. El pre-pro-péptido se encuentra entre el péptido señal y el péptido maduro, favorece el plegamiento correcto y evita la toxicidad intracelular antes de su eliminación en el aparato de Golgi. El péptido maduro es la región más importante puesto que se convierte en la conotoxina activa, contiene entre 10 y 30 residuos que conformarán la estructura tridimensional, los puentes de disulfuro y el sitio de interacción con los nAChR (Li et al., 2025).

Además de los puentes disulfuro, las conotoxinas presentan modificaciones postraduccionales frecuentes, difíciles de predecir a partir de la secuencia precursora, como la hidroxilación de la prolina, la sulfonación de la tirosina e hidroxilación del ácido glutámico las cuales favorecen su estabilidad y actividad biológica (Li et al., 2025).

Las conotoxinas se clasifican en familias según su receptor diana, dentro de las familias encontramos las α , μ , ω , etc. No obstante, las α -conotoxinas son las más caracterizadas por su capacidad de antagonizar selectivamente los nAChRs musculares y neuronales (Akondi, 2014; Tae & Adams, 2023). Los nAChRs regulan procesos fisiológicos y patológicos como la transmisión sináptica y el dolor neuropático; debido a que existen 17 subunidades que se combinan en múltiples subtipos, la selectividad de las α -conotoxinas es vital para identificar funciones fisiológicas específicas (Tae & Adams, 2023). Se ha descubierto la funcionalidad de estos subtipos haciendo uso de modelos knock-out (KO). Vincler et al. (2006) demostraron que la α -conotoxina RgIA bloquea el nAChR $\alpha 9\alpha 10$ al usar ratones KO que no tienen ese receptor. Se les administró RgIA a ratones normales con dolor, la toxina bloqueó el receptor $\alpha 9\alpha 10$ y dejaron de sentir dolor, luego, usaron ratones genéticamente modificados y al administrarle la misma α -conotoxina estos no mostraron ningún efecto analgésico. Si esta α -conotoxina actuara sobre otros receptores habría aliviado el dolor sobre los ratones KO, pero no fue así, por lo que se confirmó que ese es su receptor diana.

Las α -conotoxinas se caracterizan por poseer una hélice alfa corta estabilizada por dos puentes de disulfuro. Según su patrón de enlace, pueden adoptar tres isómeros: globular (Cys I-III, Cys II-IV), cinta (Cys I-IV, Cys II-III), cuenta (Cys I-II, Cys III-IV). La conformación principal de las α -conotoxinas es la globular (Tae & Adams, 2023; Giglio et al., 2025).

Se puede tomar como ejemplo la α -conotoxina GI, proveniente del veneno del *Conus geographus*, que tiene como objetivo los receptores musculares. Según estudios, el isómero globular de la α -conotoxina GI presentó mayor actividad frente al isómero cinta al bloquear el potencial de la placa motora, resultando en parálisis en ranas, peces y gatos anestesiados, adicionalmente, se observó una perturbación nula en la presión arterial y la frecuencia cardíaca demostrando una alta selectividad neuromuscular. Se reconoce a esta conotoxina como un marcador para diferenciar sitios de unión en nAChR musculares convirtiéndola en una herramienta farmacológica. Por lo anterior, se confirma que la actividad depende del patrón de los puentes de disulfuro puesto que pequeñas diferencias estructurales cambian directamente la potencia farmacología de las α -conotoxinas. Es fundamental considerar que lo que esta alrededor de la célula juega un papel importante en el reconocimiento de los ligandos, las α -conotoxinas presentan selectividad según la estequiometría del nAChR, dado que el sitio de unión clásico se forma entre una subunidad alfa (+) y una subunidad adyacente (-) (Tae & Adams, 2023).

Si bien se han logrado avances, el estudio de las conotoxinas se ha visto limitado debido a la falta de estrategias de plegamiento para estos péptidos ricos en cisteínas y la complejidad de los ensayos celulares. Ante esto, el enfoque computacional ha adquirido importancia. Tradicionalmente, el descubrimiento de conotoxinas se ha basado en la extracción y aislamiento de los componentes activos del veneno del caracol permitiendo clasificarlas en diversas superfamilias y familias farmacológicas según las secuencias y propiedades funcionales, pero esto requiere de mucho tiempo y trabajo. En consecuencia, se ha ampliado el uso de herramientas bioinformáticas resultando en un método eficiente para la

predicción de conotoxinas (Jin et al., 2019; Li et al., 2025).

La implementación de herramientas de aprendizaje profundo y bases de datos especializadas como UniProt y ConoServer ha sido de gran ayuda para mitigar las limitaciones que existen entorno a la caracterización de conotoxinas. Estos recursos suministran información completa sobre secuencias, funciones y estructuras amparando el desarrollo de modelos predictivos, facilitándole al aprendizaje profundo la predicción de funciones, posibles receptores diana e incluso generar nuevas secuencias. Estos métodos bioinformáticos buscan acelerar el descubrimiento de fármacos y perfeccionar la relación estructura-función, impulsando la transición de las conotoxinas desde ser venenos letales hacia herramientas fundamentales de la medicina moderna (Li et al., 2025; Jin et al., 2019).

2.2.2. Plegamiento de péptidos y el problema computacional

El proceso mediante el cual una secuencia de aminoácidos asume una estructura tridimensional específica se llama plegamiento de péptidos y proteínas. Esto ha sido un desafío para la biología estructural y la química computacional, puesto que resulta complejo predecir la estructura final de cualquier péptido (o proteína) a partir de la secuencia primaria. A esto se le conoce como el “problema del plegamiento de proteínas” (Dill & MacCallum, 2012).

Gracias a sus estudios, Anfinsen (1973) logró establecer que la secuencia de aminoácidos contiene la información necesaria para el plegamiento y que la estructura nativa adoptada por la proteína corresponderá a la que tenga la energía libre de Gibbs más baja. Durante el plegamiento ocurren diversas interacciones fisicoquímicas como las interacciones hidrofóbicas, los puentes de hidrógeno, las interacciones electrostáticas y las fuerzas de Van der Waals, a partir de estas interacciones se origina la estructura secundaria (hélices α , láminas β) y, seguidamente, se genera la estructura terciaria, la cual busca el estado de mínima energía.

Sin embargo, durante la transición de una estructura secundaria a una terciaria puede ocurrir un mal plegamiento (misfolding) en donde la proteína no logra llegar a su estado nativo o, tras alcanzarlo, se reordena de manera errónea. Esto se puede ser a causa de una mutación genética, estrés celular (cambio en el pH, aumento de temperatura o presencia de metales pesados) o fallas durante la síntesis de la proteína. Como mecanismo de control, existen las chaperonas moleculares encargadas de asistir a otras proteínas en su correcto plegamiento, transporte y degradación, reconociendo e interceptando la parte hidrofóbica expuesta de las proteínas mal plegadas. No obstante, cuando este mecanismo falla, los residuos hidrofóbicos exhibidos, debido al plegamiento incorrecto, actúan como pegamento entre proteínas formando agregados tóxicos para las neuronas, entre otros. El misfolding es considerado la causa principal de enfermedades neurodegenerativas como el Alzheimer y el Parkinson (Dobson, 2003).

Desde el punto de vista computacional, el plegamiento consiste en examinar una enorme cantidad de posibles estructuras o formas que puede adoptar la molécula. De la misma manera, la paradoja de Levinthal sostiene que una proteína no puede recorrer al azar todas sus posibles conformaciones, ya que esto tomaría un tiempo astronómicamente largo. Al contrario, el plegamiento está dirigido por una serie de rutas específicas que surgen a raíz de las interacciones locales que van estabilizando y guiando el proceso hacia la estructura nativa. Consecuentemente, surge un modelo de embudo de plegamiento en donde la proteína inicia desplegada y en un estado desordenado. A medida que ocurren las interacciones locales, la energía libre disminuye, lo que conduce al ordenamiento de la estructura tridimensional hasta alcanzar su estado más estable (Levinthal, 1969).

Por otra parte, la predicción de estructuras de péptidos y proteínas ha pasado de usar modelos estáticos a utilizar estrategias dinámicas que permiten explorar cómo cambian y se comportan estas moléculas en diferentes condiciones. Mientras que la mecánica molecular se basa en campos de fuerza para evaluar la energía potencial, la dinámica molecular y los

métodos de Monte Carlo hacen posible simular cómo se mueve y se pliega la estructura primaria. Las simulaciones de dinámica molecular ayudan a ver la estructura final de una proteína y también permiten entender sus movimientos, los estados intermedios y los cambios que son cruciales para su función biológica. Este enfoque dinámico es esencial para mitigar las limitaciones de los métodos basados en búsquedas al azar y ha sido fundamental en el desarrollo de nuevos algoritmos que combinan la física molecular con técnicas de aprendizaje profundo (Karplus & McCammon, 2002).

En contraste, el reciente surgimiento de la inteligencia artificial ha cambiado el enfoque, creando modelos como AlphaFold capaces de predecir estructuras tridimensionales de proteínas con precisión sin hacer simulaciones previas del plegamiento (Jumper et al., 2021). Estas predicciones han logrado tal precisión gracias a datos estructurales recolectados del PDB, al uso de redes neuronales profundas y a la integración de información evolutiva y espacial.

A diferencia de las proteínas grandes, los péptidos son más difíciles de predecir estructuralmente debido a su flexibilidad y la ausencia de una única conformación estable (Dill & MacCallum, 2012). Su estructura puede variar según el entorno, como el pH o la interacción con membranas, lo que requiere métodos dinámicos para captar estos cambios (Karplus & McCammon, 2002). Ha sido clave comprender estos procesos para el modelado de péptidos mediante herramientas como AlphaFold, con fines terapéuticos, (Jumper et al., 2021), para avanzar en el estudio de enfermedades neurodegenerativas asociadas al mal plegamiento proteico (Dobson, 2003) y para optimizar la ingeniería de proteínas, resolviendo desafíos como la paradoja de Levinthal (Levinthal, 1969).

2.2.3. Evolución de AlphaFold

Durante décadas, entender cómo una secuencia de aminoácidos logra adoptar la estructura tridimensional necesaria para su función ha sido uno de los grandes retos de la biología estructural. En este escenario, la creación de AlphaFold por DeepMind ha marcado un antes y un después, integrando la inteligencia artificial con las ciencias de la vida y revolucionando la forma en que abordamos este problema (Jumper et al., 2021).

Antes de AlphaFold, los científicos se basaban principalmente en el modelado comparativo y en simulaciones de dinámica molecular para predecir estructuras de proteínas. Sin embargo, estos métodos tenían grandes limitaciones, ya que dependían de la existencia de estructuras conocidas y requerían muchos recursos computacionales (Senior et al., 2020). El desarrollo de AlphaFold fue posible gracias al aumento de datos disponibles en el PDB y a los avances en el aprendizaje profundo.

AlphaFold 1 se presentó por primera vez en la competencia CASP13 en 2018. Esta versión usaba redes neuronales profundas y datos de la evolución de las proteínas para predecir cómo se organizan sus residuos, lo que permitió obtener resultados mucho mejores que los métodos tradicionales (Senior et al., 2020). Sin embargo, todavía tenía limitaciones, ya que su diseño era algo rígido y seguía dependiendo en parte de modelos físicos.

AF2 dió el gran salto, presentado en CASP14 en 2020. Esta nueva versión incorporó una tecnología mucho más avanzada, que le permitió predecir la estructura de las proteínas con una precisión sorprendente (Jumper et al., 2021). Gracias a su capacidad para mejorar sus propias predicciones una y otra vez, AlphaFold 2 logró resultados tan buenos que en muchos casos se comparan con los obtenidos en experimentos de laboratorio.

Luego, AlphaFold-Multimer aumentó estas capacidades en la predicción de complejos proteicos, permitiendo modelar interacciones proteína-proteína (Evans et al., 2022). Esto representó un avance hacia el análisis de sistemas biológicos más complejos.

En tiempos más recientes, AlphaFold 3 ha ampliado las posibilidades del modelado estructural, permitiendo analizar no solo proteínas, sino también complejos formados con ADN, ARN y otras moléculas (Abramson et al., 2024). Gracias a distintas innovaciones en su diseño como técnicas de difusión, esta versión logra una mayor precisión y una capacidad mejorada para abordar una variedad más amplia de sistemas biológicos.

AlphaFold ha impactado muchos campos. En la biología estructural, ha hecho que determinar la estructura de las proteínas sea mucho más rápido. En el área de la farmacología, ha ayudado a diseñar nuevos medicamentos y a encontrar posibles blancos terapéuticos (Callaway, 2020). Además, al poner a disposición del público una gran cantidad de estructuras, ha permitido que más personas accedan y contribuyan al conocimiento científico (Varadi et al., 2022).

En resumen, la evolución de AlphaFold puede entenderse en tres grandes etapas: la primera versión marcó el inicio del uso del aprendizaje profundo en este campo, la segunda logró resolver de manera efectiva el problema del plegamiento de proteínas, y la tercera amplió su alcance para incluir sistemas biomoleculares más complejos. Todo este recorrido ha transformado la manera en que la inteligencia artificial contribuye a la investigación científica actual.

2.2.4. Bases de datos estructurales

Las bases de datos que reúnen la información estructural de péptidos y proteínas son esenciales para la biología estructural, la bioinformática y la química computacional. Estas plataformas no solo almacenan y organizan la información tridimensional de las macromoléculas, sino que también hacen posible entender cómo su forma influye en su función y ayudan en el desarrollo de nuevos fármacos y biomateriales.

La idea de almacenar las estructuras tridimensionales de proteínas y otras moléculas nació cuando surgieron técnicas experimentales como la cristalografía de rayos X, la resonancia magnética nuclear (RMN) y, más recientemente, la criomicroscopía electrónica (cryo-EM) (Kühlbrandt, 2014). Dentro de este contexto, el PDB, creado en 1971, se ha convertido en la base de datos más importante a nivel mundial para almacenar información sobre la estructura de macromoléculas (Berman et al., 2000).

El PDB posee datos precisos sobre: coordenadas atómicas tridimensionales, secuencias de aminoácidos, información experimental (resolución, método, condiciones), anotaciones funcionales y estructurales. En la actualidad, esta base de datos es gestionada por el consorcio Worldwide PDB, que incluye organizaciones en Estados Unidos, Europa y Asia, garantizando estándares globales y acceso abierto (wwPDB Consortium, 2019).

Adicionalmente, existe una gran variedad de bases de datos especializadas que amplían el análisis estructural. Las bases de datos primarias son repositorios que contienen datos experimentales originales, el PDB es el principal repositorio estructural.

Existen bases de datos que sirven para validar modelos estructurales y como fuente de datos para entrenamiento de algoritmos. La Electron Microscopy Data Bank se enfoca en estructuras obtenidas por criomicroscopía electrónica (Lawson et al., 2016), mientras que la Biological Magnetic Resonance Data Bank se especializa en datos de RMN (Ulrich et al., 2008).

Asimismo, las bases de datos derivadas o secundarias toman la información del PDB y la enriquecen para hacerla más accesible y útil para los investigadores. Por ejemplo, SCOP organiza las proteínas de acuerdo a su estructura y evolución, CATH las clasifica en distintos niveles según sus características estructurales, y PDBsum ofrece resúmenes visuales de las estructuras (Laskowski et al., 1997). Gracias a estas herramientas, es posible explorar relaciones evolutivas, identificar dominios estructurales y analizar patrones funcionales de las proteínas (Orengo et al., 1997; Murzin et al., 1995).

Las bases de datos de péptidos bioactivos se centran en recopilar información sobre péptidos con funciones biológicas concretas. Por ejemplo, la Antimicrobial Peptide Database reúne péptidos con propiedades antimicrobianas, PeptideAtlas integra datos experimentales de proteómica (Desiere et al., 2006) y BIOPEP se enfoca en péptidos con actividad fisiológica (Minkiewicz et al., 2008). Estas plataformas son fundamentales para el desarrollo de nuevos medicamentos, en particular antibióticos y péptidos con aplicaciones terapéuticas (Wang et al., 2016).

Las bases de datos estructurales han sido clave para que AlphaFold pueda desarrollarse. Gracias a la información almacenada en el PDB, este sistema ha aprendido a predecir con gran precisión la forma de las proteínas (Jumper et al., 2021). Además, la creación de la AlphaFold Protein Structure Database ha hecho posible que tengamos acceso a muchas más estructuras, incluyendo predicciones computacionales (Varadi et al., 2022).

Esta integración ha cambiado profundamente la biología estructural, permitiendo pasar de depender únicamente de métodos experimentales a adoptar un enfoque híbrido en el que los datos y la inteligencia artificial juegan un papel fundamental.

Las bases de datos estructurales desempeñan un papel central en numerosos avances científicos y tecnológicos. Estas herramientas permiten identificar sitios activos para el diseño de fármacos más efectivos, optimizar proteínas mediante ingeniería para mejorar sus propiedades y apoyar desarrollos innovadores en biotecnología, como enzimas industriales y biosensores. Además, son fundamentales para la investigación básica, ya que facilitan el estudio de los mecanismos moleculares, las interacciones entre proteínas y la evolución de sus estructuras (Blundell et al., 2002).

Aunque las bases de datos estructurales son esenciales para la ciencia moderna, todavía enfrentan retos importantes, como el predominio de proteínas solubles sobre proteínas de membrana, variaciones en la precisión y resolución experimental y la falta de información sobre las condiciones dinámicas y fisiológicas reales. Además, el rápido aumento en la

cantidad de datos genera la necesidad de mejorar continuamente su almacenamiento, curación y estandarización.

Aun así, desde la creación del PDB hasta la integración de inteligencia artificial, estas plataformas se han transformado en herramientas fundamentales para la investigación biomolecular. Se estima que su influencia seguirá creciendo, impulsando el desarrollo de nuevas tecnologías, medicamentos y la comprensión de procesos biológicos, especialmente a medida que se adoptan métodos computacionales más avanzados. En definitiva, estas bases de datos no solo gestionan información, sino que también son motor de descubrimiento científico a nivel mundial.

2.3. Marco Conceptual

Las α -conotoxinas son pequeños péptidos neuroactivos que se encuentran en el veneno de caracoles marinos del género *Conus*. Se caracterizan por su alta selectividad y afinidad hacia los nAChRs, lo que les otorga un papel importante en la regulación de procesos fisiológicos específicos. Estas moléculas suelen estar formadas por cadenas cortas de 10 a 30 aminoácidos y contienen un alto número de residuos de cisteína, facilitando la formación de puentes disulfuro que estabilizan su estructura tridimensional. Gracias a estas propiedades, las α -conotoxinas se han convertido en herramientas valiosas para la investigación biomédica y tienen un gran potencial como agentes terapéuticos (Akondi et al., 2014; Jin et al., 2019; Tae & Adams, 2023).

La secuencia genómica se refiere al orden específico de nucleótidos en el ADN que contiene la información necesaria para producir proteínas y péptidos. En las conotoxinas, esta secuencia da lugar a precursores peptídicos que, tras la traducción y maduración, se convierten en péptidos funcionales (Li et al., 2025; Mao et al., 2026). Un elemento clave en este proceso es el péptido señal, una secuencia corta que funciona como una etiqueta celular y dirige la

proteína recién formada hacia el retículo endoplasmático, asegurando su correcto procesamiento y secreción (Li et al., 2025).

Las modificaciones postraduccionales son cambios químicos que se producen en los péptidos después de ser sintetizados por los ribosomas y resultan esenciales para que puedan ejercer su función biológica. Estas modificaciones incluyen, por ejemplo, la hidroxilación de la prolina, la sulfonación de la tirosina y la hidroxilación del ácido glutámico, procesos que influyen en la estabilidad estructural del péptido y en su capacidad para interactuar con los receptores (Akondi et al., 2014; Li et al., 2025). Además, la formación de puentes de disulfuro, que son enlaces covalentes entre residuos de cisteína, ayuda a mantener la estructura tridimensional del péptido y le confiere mayor resistencia frente a condiciones adversas (Jin et al., 2019).

Los isómeros estructurales corresponden a distintas disposiciones espaciales que una misma molécula puede adoptar, manteniendo su composición química intacta. En el caso de las α -conotoxinas, la presencia de diferentes patrones de conectividad de los puentes disulfuro da lugar a estos isómeros, lo que repercute directamente en su actividad biológica y en su especificidad funcional (Tae & Adams, 2023; Giglio et al., 2025).

El receptor diana se refiere a la molécula específica con la que interactúa un ligando para desencadenar su efecto biológico. En el caso de las α -conotoxinas, los principales receptores diana son los nAChRs, integrantes de la familia de los canales iónicos. Estas proteínas transmembrana permiten el paso selectivo de iones a través de la membrana celular (Akondi et al., 2014; Tae & Adams, 2023). Los canales iónicos desempeñan un papel esencial en la transmisión sináptica, proceso fundamental para la comunicación neuronal, y su alteración se ha relacionado con diversas patologías, incluido el dolor neuropático (Tae & Adams, 2023).

Desde el punto de vista estructural, la estructura nativa se define como la conformación tridimensional funcional que una proteína o péptido adopta bajo condiciones fisiológicas. Esta

conformación resulta de la minimización de la energía libre de Gibbs, lo que implica que las proteínas tienden a adoptar el estado energéticamente más estable (Anfinsen, 1973). En este proceso participan diversas interacciones fisicoquímicas, como las fuerzas de Van der Waals, así como interacciones hidrofóbicas y electrostáticas, que en conjunto contribuyen a la estabilidad global de la estructura (Dill & MacCallum, 2012).

La degradación proteolítica consiste en la hidrólisis de proteínas y péptidos mediada por enzimas proteasas, proceso que impacta directamente en su estabilidad y vida media. En este contexto, aquellas estructuras que presentan estabilización mediante puentes disulfuro y modificaciones postraduccionales suelen exhibir una mayor resistencia frente a la degradación proteolítica (Jin et al., 2019).

El análisis del plegamiento y la estructura de proteínas ha sido abordado con herramientas computacionales de última generación. En este ámbito, el aprendizaje profundo, una rama de la inteligencia artificial basada en redes neuronales, permite modelar relaciones complejas en grandes conjuntos de datos y se ha consolidado como una estrategia fundamental para la predicción estructural de biomoléculas (Jumper et al., 2021; Akdel et al., 2022).

En este contexto, AlphaFold se presenta como un sistema basado en aprendizaje profundo, específicamente diseñado para predecir la estructura tridimensional de proteínas a partir de su secuencia de aminoácidos. La primera versión, AlphaFold 1, incorporó redes neuronales para la predicción de contactos entre residuos, superando los enfoques tradicionales en este ámbito (Senior et al., 2020). Posteriormente, AF2 supuso un avance significativo al integrar información evolutiva y arquitecturas de atención profunda, logrando una precisión en la predicción estructural comparable a la obtenida mediante métodos experimentales (Jumper et al., 2021; Callaway, 2020). Recientes investigaciones han ampliado su aplicación, demostrando que también puede ser útil para inferir rutas de plegamiento a partir de la secuencia primaria (Chang & Perez, 2025).

Por otro lado, AlphaFold-Multimer extendió estas capacidades al ámbito de la

predicción de complejos proteína-proteína, permitiendo abordar interacciones moleculares de mayor complejidad (Evans et al., 2022). Más recientemente, AF3 ha incorporado innovaciones tales como la implementación de modelos de difusión, un enfoque probabilístico que permite generar estructuras considerando múltiples configuraciones posibles. Esta estrategia ha facilitado la predicción precisa de interacciones no solo entre proteínas, sino también con ácidos nucleicos y otras biomoléculas (Abramson et al., 2024).

El desarrollo de estos modelos ha sido posible, en gran medida, gracias a la amplia disponibilidad de datos estructurales proporcionados por bases como el PDB, que almacena información tridimensional de macromoléculas (Berman et al., 2000; Bittrich et al., 2024). Asimismo, otras plataformas especializadas, como EMDDataBank y BioMagResBank, han desempeñado un papel fundamental en el entrenamiento y validación de modelos predictivos, al ofrecer recursos estructurales complementarios relevantes para distintas aproximaciones metodológicas (Lawson et al., 2016; Ulrich et al., 2008).

3. Metodología

3.1. Selección de α -conotoxinas modelo

Inicialmente, se eligieron las α -conotoxinas con estructuras disponibles en el PDB, en la búsqueda se usaron palabras clave como “ α -conotoxin”, “Conus”, “toxin”, y se aplicaron filtros por resolución (menor o igual a 2.5 Å), método de determinación (RMN o rayos X) y número de residuos, a su vez se tuvo en cuenta que las estructuras estuvieran presentes en AFDB; las estructuras que escogieron tenían una longitud entre 12 y 18 aminoácidos y una presencia confirmada de los dos puentes de disulfuro.

Una vez se obtuvo el conjunto representativo de α -conotoxinas se realizó una tabla en

la cual se indicó: nombre, código PDB, secuencia, organismos fuente, técnica de determinación, resolución, y número de puentes disulfuro.

3.2. Generación de modelos estructurales con AF3

Luego, usando el servidor de AF3, se generaron cinco modelos por cada secuencia (al correr una secuencia en el servidor este lo hace automáticamente) con sus valores de confianza asociados (pLDDT), seguidamente, se conservó la predicción más confiable por residuo (según pLDDT) para el análisis comparativo.

3.3. Obtención de modelos comparativos de AFDB y estructuras del PDB

Se descargaron los modelos correspondientes de AlphaFold DB (AFDB) y se verificó que coincidieran con la α -conotoxina reportada en cada estructura experimental seleccionada. Dado que los modelos de AFDB incluyen la secuencia precursora completa de la conotoxina, fue necesario extraer únicamente la región correspondiente al péptido maduro antes de realizar las comparaciones estructurales. Para ello, se tomó como referencia la secuencia depositada en el PDB, identificando los residuos iniciales y finales del péptido maduro y eliminando las regiones señal y propeptídicas utilizando el software BIOVIA Discovery Studio. Los rangos de residuos empleados para cada α -conotoxina se presentan en la Tabla 1. A continuación, se comprobó que la secuencia recortada coincidiera exactamente con la de la estructura experimental utilizada como referencia. También se obtuvieron los archivos .pdb oficiales, tomando en cuenta la calidad estructural, el método experimental empleado y la presencia de modificaciones postraduccionales.

Tabla 1. *Regiones de los modelos de AlphaFold DataBase conservadas para la obtención de los péptidos maduros empleados en la comparación estructural.*

α -Conotoxina	Residuos conservados del modelo AFDB
EI	44–61
GI	50–62
GIC	21–36
ImI	5–16
MII	49–64
PnIA	No requirió recorte
RgIA	20–32
SI	50–62
TxIB	22–37
Vc1.1	50–65

Nota. Los modelos descargados de AlphaFold DB correspondían, en algunos casos, a secuencias precursoras completas. Los residuos indicados corresponden a las regiones conservadas tras el recorte de las secuencias señal y propeptídicas, utilizando como referencia las secuencias de los péptidos maduros reportadas en las estructuras experimentales depositadas en el PDB.

3.4. Comparación estructural y análisis cuantitativo

Se emplearon PyMOL y BIOVIA Discovery Studio como herramientas para realizar comparaciones estructurales y validaciones geométricas. Las métricas que se tuvieron en cuenta: RMSD global entre AF3 y estructuras del PDB, RMSD por regiones (loops, hélices), comparación de puentes disulfuro: presencia, orientación y conectividad, evaluación de pLDDT: distribución de confianza por residuo. Se hicieron las siguientes comparaciones por cada α -conotoxina usando una herramienta de alineación de secuencias por pares (Bittrich et al., 2024): AF3 vs. Estructura PDB, AF3 vs. AFDB y AFDB vs. Estructura PDB.

3.5. Interpretación de resultados y evaluación del desempeño

En esta fase se hizo el análisis de las estructuras y la evaluación del desempeño de AF3, los criterios de evaluación del desempeño fueron: a) Precisión estructural: RMSD < 2.0 Å respecto al PDB, b) Fiabilidad del modelo: pLDDT promedio > 85) Conservación de enlaces disulfuro: coincidencia completa

4. Resultados

Tabla 2. Características estructurales y de secuencia de las α -conotoxinas seleccionadas para el análisis comparativo.

Toxina	Código PDB	Secuencia	N° de residuos	Especie del caracol	Método experimental	Resolución	Puentes de disulfuro**	Subfamilia/loop spacing
ImI	1IM1	GCCSDP RCAWR C	12	<i>Conus imperialis</i>	RMN (20)	N/A*	2 (Cys 2 - Cys 8; Cys 3 - Cys 12)	α 4/3
GI	1NOT	ECCNPA CGRHYS C	13	<i>Conus geographus</i>	Rayos X	1.20 Å	2 (Cys 2 - Cys 7; Cys 3 - Cys 13)	α 3/5
SI	1HJE	ICCNPA CGPKYS C	13	<i>Conus striatus</i>	Rayos X	0.75 Å	2 (Cys 2 - Cys 7; Cys 3 - Cys 13)	α 3/5
MII	1MII	GCCSNP VCHLEH SNLC	16	<i>Conus magus</i>	RMN (20)	N/A*	2 (Cys 2 - Cys 8; Cys 3 - Cys 16)	α 4/7
PnIA	1PEN	GCCSLP PCAANN PDYC	16	<i>Conus pennaceus</i>	Rayos X	1.10 Å	2 (Cys 2 - Cys 8; Cys 3 - Cys 16)	α 4/7
GIC	1UL2	GCCSHP ACAGN NQHIC	16	<i>Conus geographus</i>	RMN (20)	N/A*	2 (Cys 2 - Cys 8; Cys 3 - Cys 16)	α 4/7
EI	1K64	RDPCCY HPTCNM SNPQIC	18	<i>Conus ermineus</i>	RMN (20)	N/A*	2 (Cys 4 - Cys 10; Cys 5 - Cys 18)	α 4/7
RgIA	2JUT	GCCSDP RCRYRC R	13	<i>Conus regius</i>	RMN (20)	N/A*	2 (Cys 2 - Cys 8; Cys 3 - Cys 12)	α 4/3
Vc1.1	2H8S	GCCSDP RCNYDH PEIC	16	<i>Conus victoriae</i>	RMN (20)	N/A*	2 (Cys 2 - Cys 8; Cys 3 - Cys 16)	α 4/7
TxIB	2LZ5	GCCSDP PCRKNH PDLC	16	<i>Conus textile</i>	RMN (20)	N/A*	2 (Cys 2 - Cys 8; Cys 3 - Cys 16)	α 4/7

Nota. Los datos de la columna "Secuencia" corresponden al péptido maduro; en el caso de las

estructuras obtenidas de AlphaFold Database (AFDB), las secuencias fueron recortadas manualmente, excepto por la correspondiente a la α -conotoxina PnIA, para coincidir con el fragmento reportado experimentalmente. Para el método experimental RMN se reporta la cantidad de conformeros que se encuentran registrados en el PDB. El patrón de *loop spacing* indica el número de aminoácidos entre las cisteínas del armazón estructural.

*No aplicable; la técnica de RMN no usa un parámetro de resolución nominal comparable al de la cristalografía de rayos X. **Los puentes de disulfuro registrados están presentes en todas las estructuras modeladas por AF3.

Tabla 3. Métricas de desempeño estructural y niveles de confianza de AF3 frente a registros experimentales y predicciones de AlphaFold Database para las alfa-conotoxinas seleccionadas.

Toxina	Modelo AF3	RMSD			Átomos rechazados para RMSD	Puentes de disulfuros correctos en AFDB	pLDDT	
		AF3 vs PDB	AFDB vs AF3	AFDB vs PDB			AF3	AFDB
EI	1	0,610	2,821	2,887	1 (AF3 vs PDB)	NO (Cys 1-4, Cys 2-3)	96,66	80,70
GI	0	1,060	3,010	3,072	N/A	NO (Cys 1-4, Cys 2-3)	96,85	62,13
GIC	0	0,835	0,540	0,970	1 (AFDB vs AF3)	SI	97,67	85,62
ImI	0	0,458	0,303	0,370	1 en todos los casos	Solo Cys 2-4	96,62	79,65
MII	1	0,756	0,731	1,392	3 (AFDB vs AF3), 2 (AFDB vs PDB)	NO (Cys 1-4, Cys 2-3)	95,13	75,64
PnIA	1	0,445	1,335	1,259	2 (AFDB vs PDB) y (AFDB vs PDB)	NO (Cys 1-4, Cys 2-3)	97,30	69,42
RgIA	4	3,051	1,073	2,762	N/A	SI	92,80	71,96
SI	2	0,643	0,620	0,934	2 (AFDB vs AF3) y (AFDB vs PDB)	SI	97,35	75,60
Vc1.1	1	0,258	0,546	0,199	2 (AF3 vs PDB), 3 (AFDB vs AF3), 6 (AFDB vs PDB)	NO (Cys 1-4, Cys 2-3)	92,28	75,03
TxIB	0	0,428	0,320	0,614	N/A	SI	95,92	78,18

Nota. Los valores de RMSD (Å) corresponden al alineamiento de los carbonos alfa (C α). La columna "Modelo AF3" indica el modelo con el menor valor de RMSD (frente al PDB) y el mayor valor de pLDDT de los 5 modelos que proporcionó AF3. La columna "Átomos rechazados" indica el número de residuos que fueron excluidos del cálculo para optimizar el ajuste global debido a su alta flexibilidad. Las observaciones sobre los puentes disulfuro en la AFDB destacan las discrepancias en la conectividad Cys1-Cys4 y Cys2-Cys3, comparadas con la estructura nativa del PDB. Los puentes de disulfuro fueron asignados correctamente en todas las estructuras obtenidas por AF3. pLDDT = Prueba de diferencia de distancia local predicha; RMSD = Desviación cuadrática media; AF3 = AlphaFold 3.0; AFDB = AlphaFold Database; PDB = Protein Data Bank.

A continuación se presentan los resultados del análisis estructural comparativo de las alfa-conotoxinas evaluadas. Las figuras muestran el alineamiento entre las estructuras obtenidas en AF3, las estructuras obtenidas de AFDB y las estructuras del PDB (usadas como referencia). Para la interpretación de cada imagen tener presente que la estructura de AF3 es de color azul, la estructura del PDB es de color verde y la estructura de AFDB es de color naranja. Asimismo, los valores de RMSD correspondientes a cada superposición se encuentran detallados en la Tabla [2].

Figura 1. Superposiciones estructurales de la α -conotoxina EI. (A) AF3 (azul) vs PDB (verde), (B) AF3 (azul) vs AFDB (naranja), (C) AFDB (naranja) vs PDB (verde)

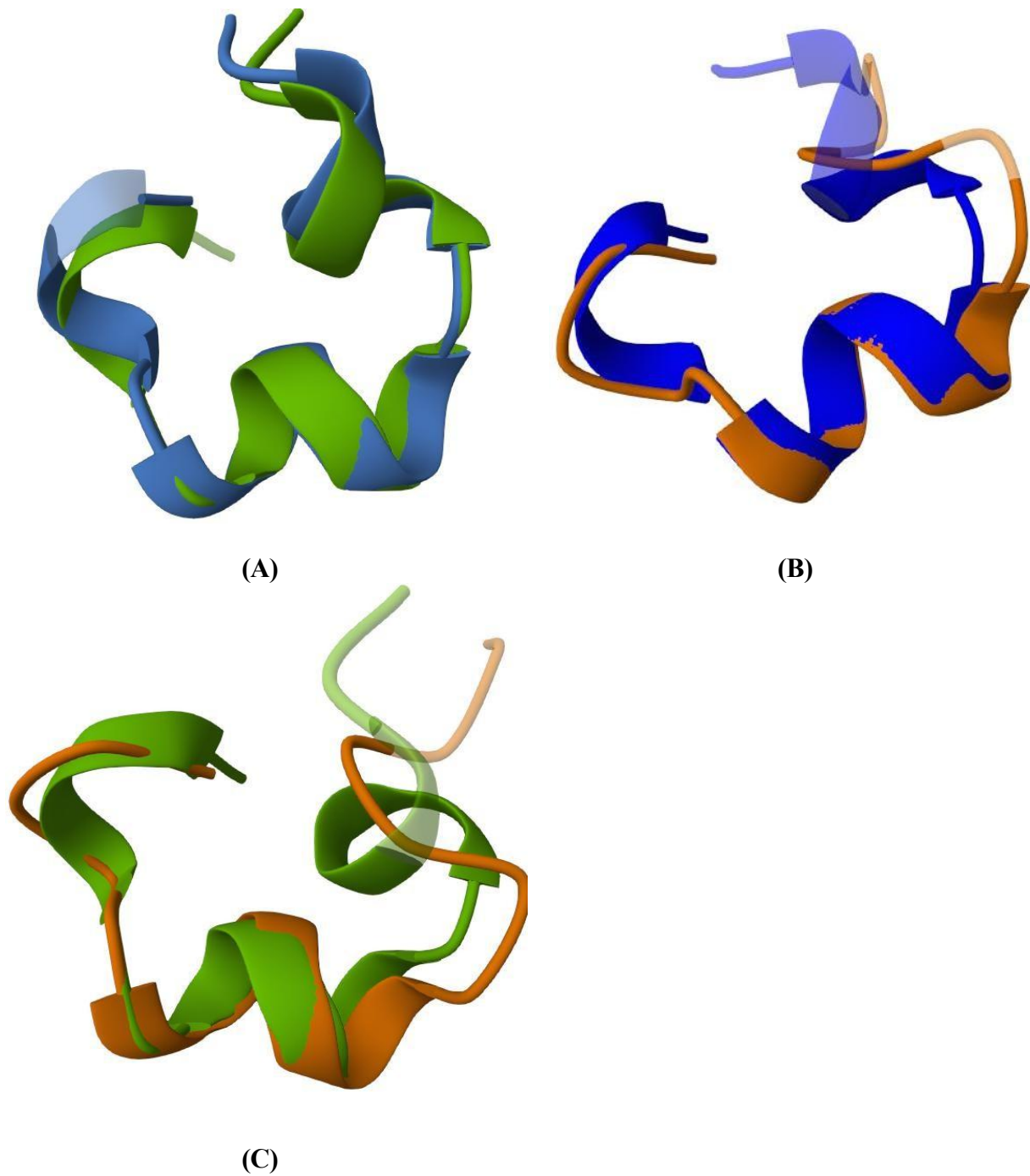


Figura 2. Superposiciones estructurales de la α -conotoxina GI. (A) AF3 (azul) vs PDB (verde), (B) AF3 (azul) vs AFDB (naranja), (C) AFDB (naranja) vs PDB (verde)

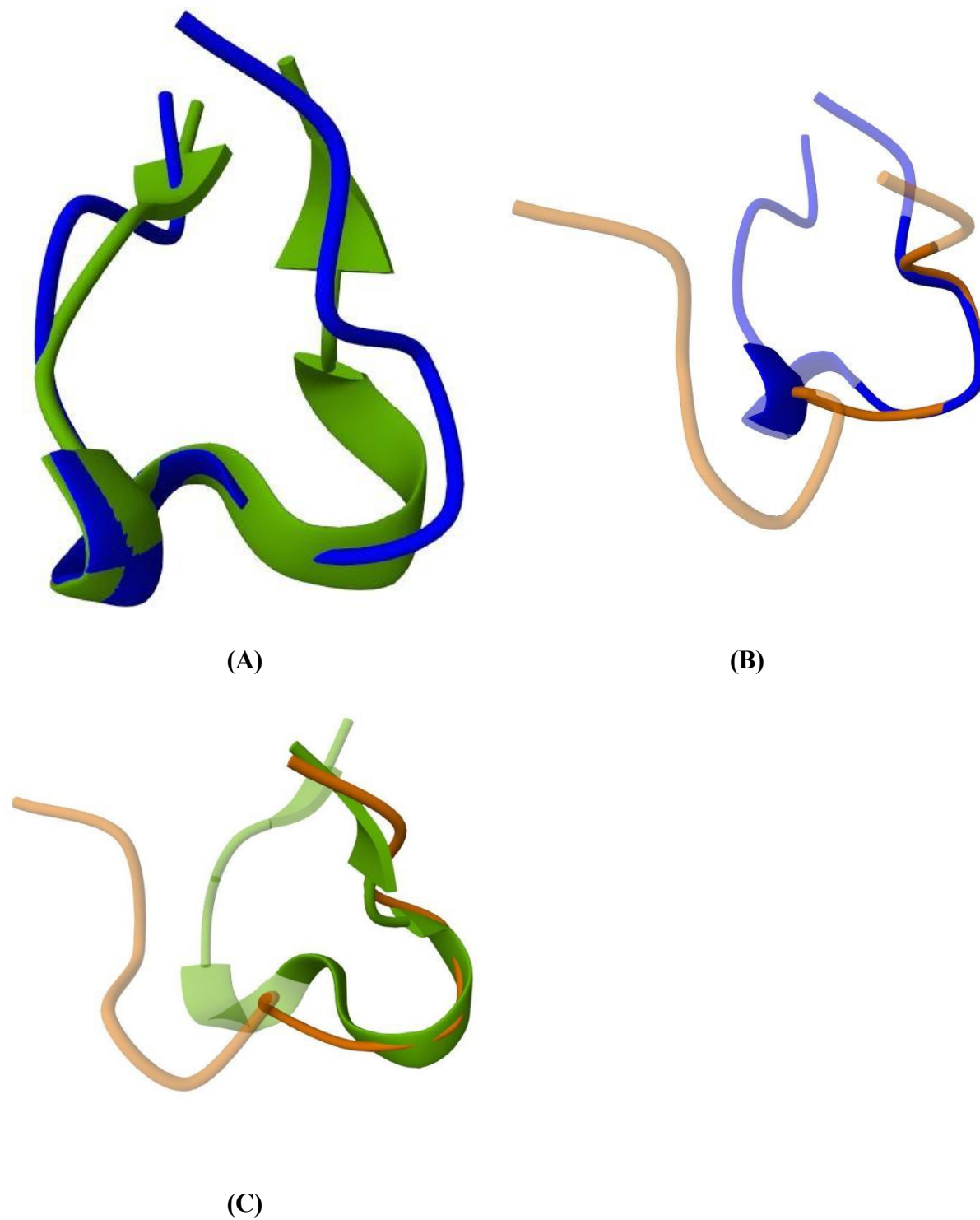


Figura 3. Superposiciones estructurales de la α -conotoxina GIC. (A) AF3 (azul) vs PDB (verde), (B) AF3 (azul) vs AFDB (naranja), (C) AFDB (naranja) vs PDB (verde)

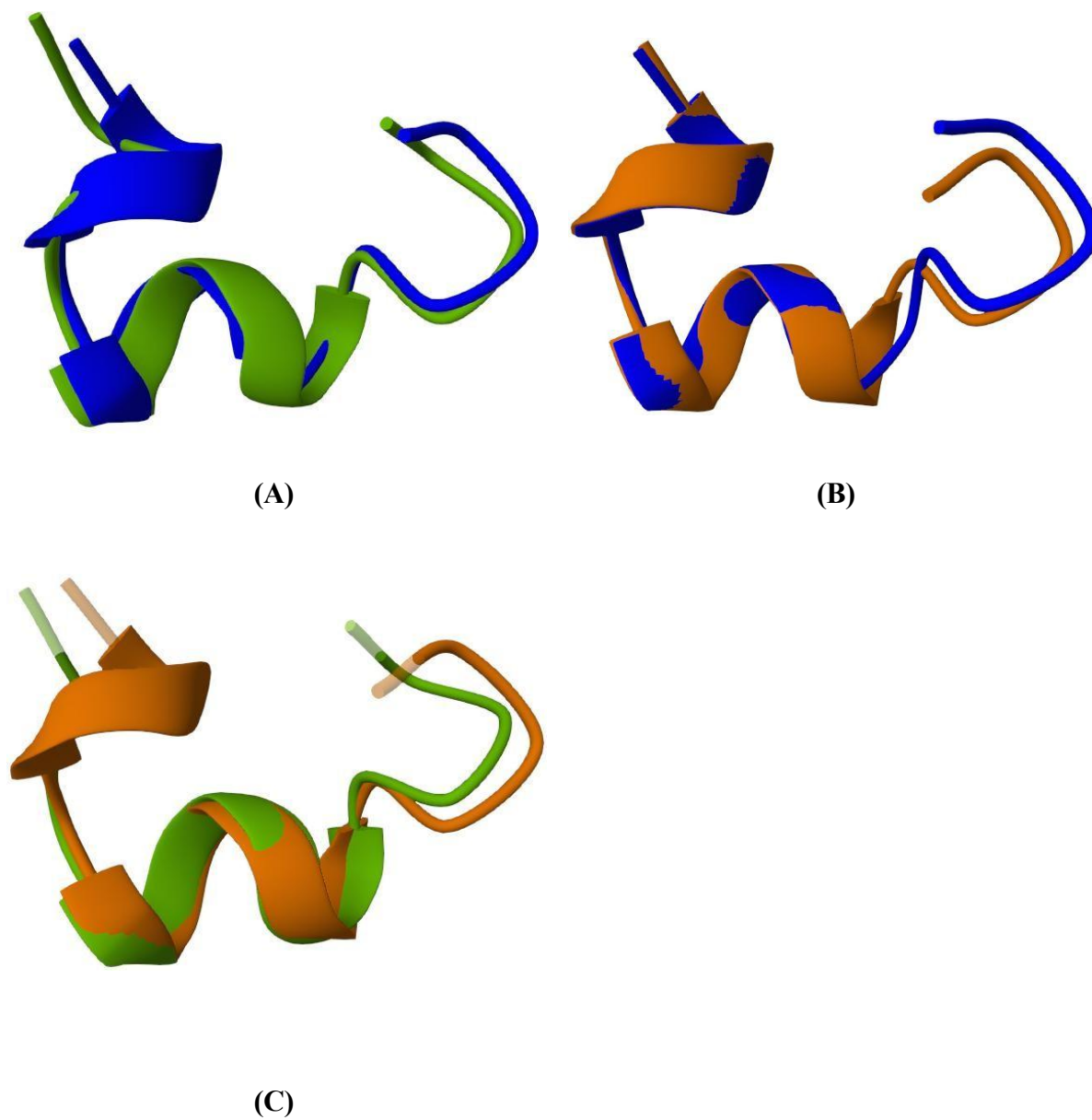


Figura 4. Superposiciones estructurales de la α -conotoxina ImI. (A) AF3 (azul) vs PDB (verde), (B) AF3 (azul) vs AFDB (naranja), (C) AFDB (naranja) vs PDB (verde)

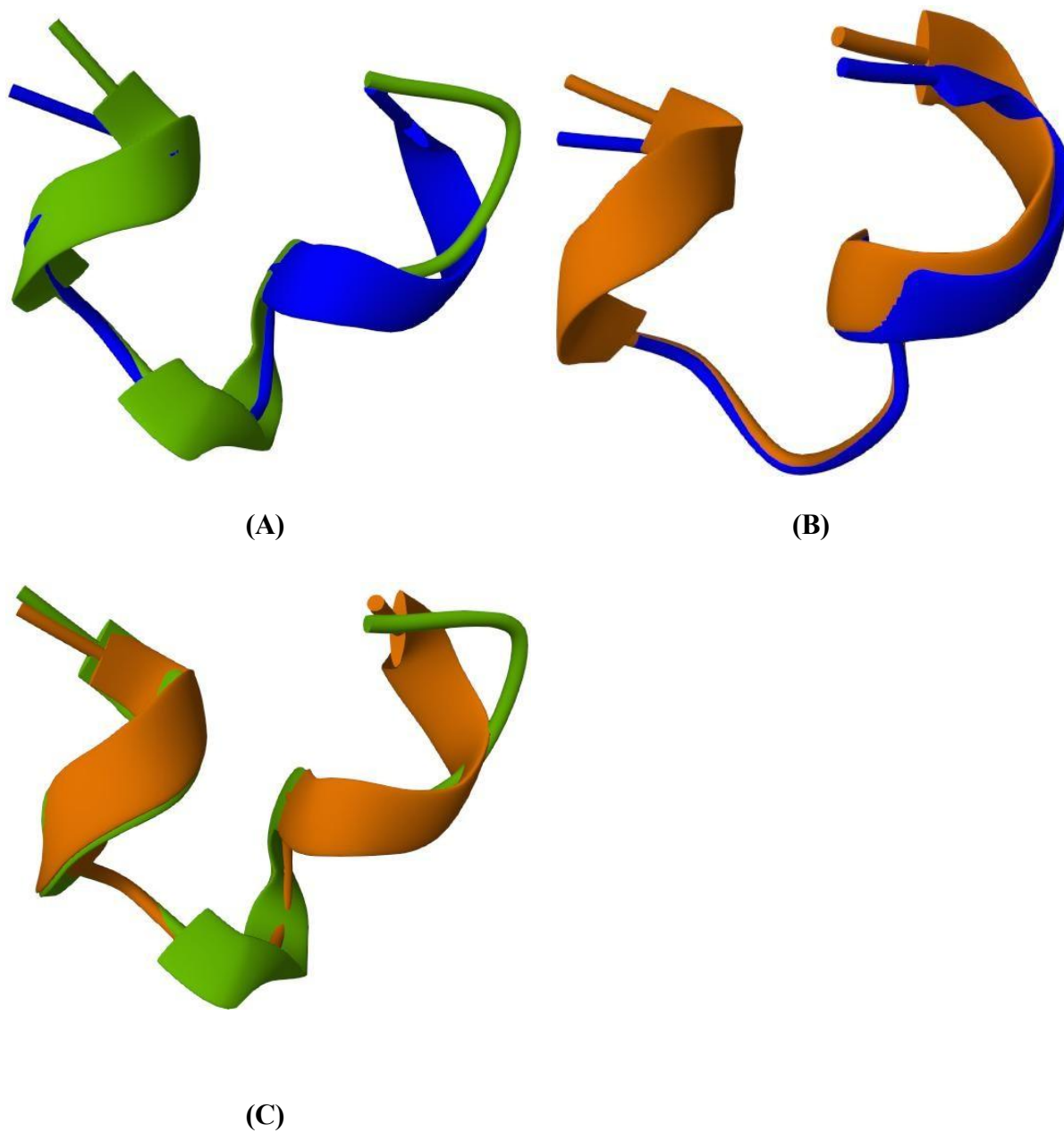


Figura 5. Superposiciones estructurales de la α -conotoxina MII. (A) AF3 (azul) vs PDB (verde), (B) AF3 (azul) vs AFDB (naranja), (C) AFDB (naranja) vs PDB (verde)

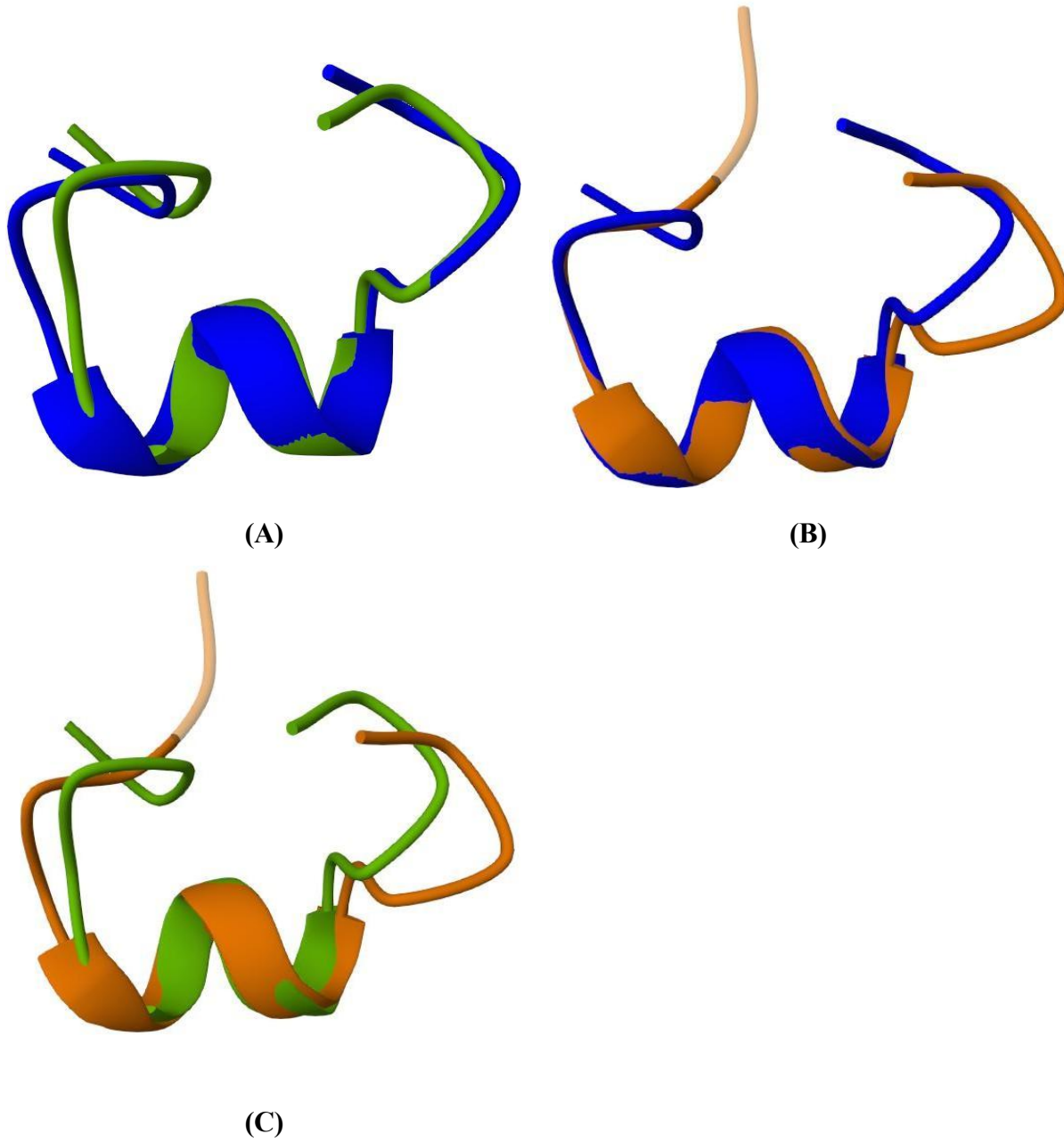


Figura 6. Superposiciones estructurales de la α -conotoxina PnIA. (A) AF3 (azul) vs PDB (verde), (B) AF3 (azul) vs AFDB (naranja), (C) AFDB (naranja) vs PDB (verde)

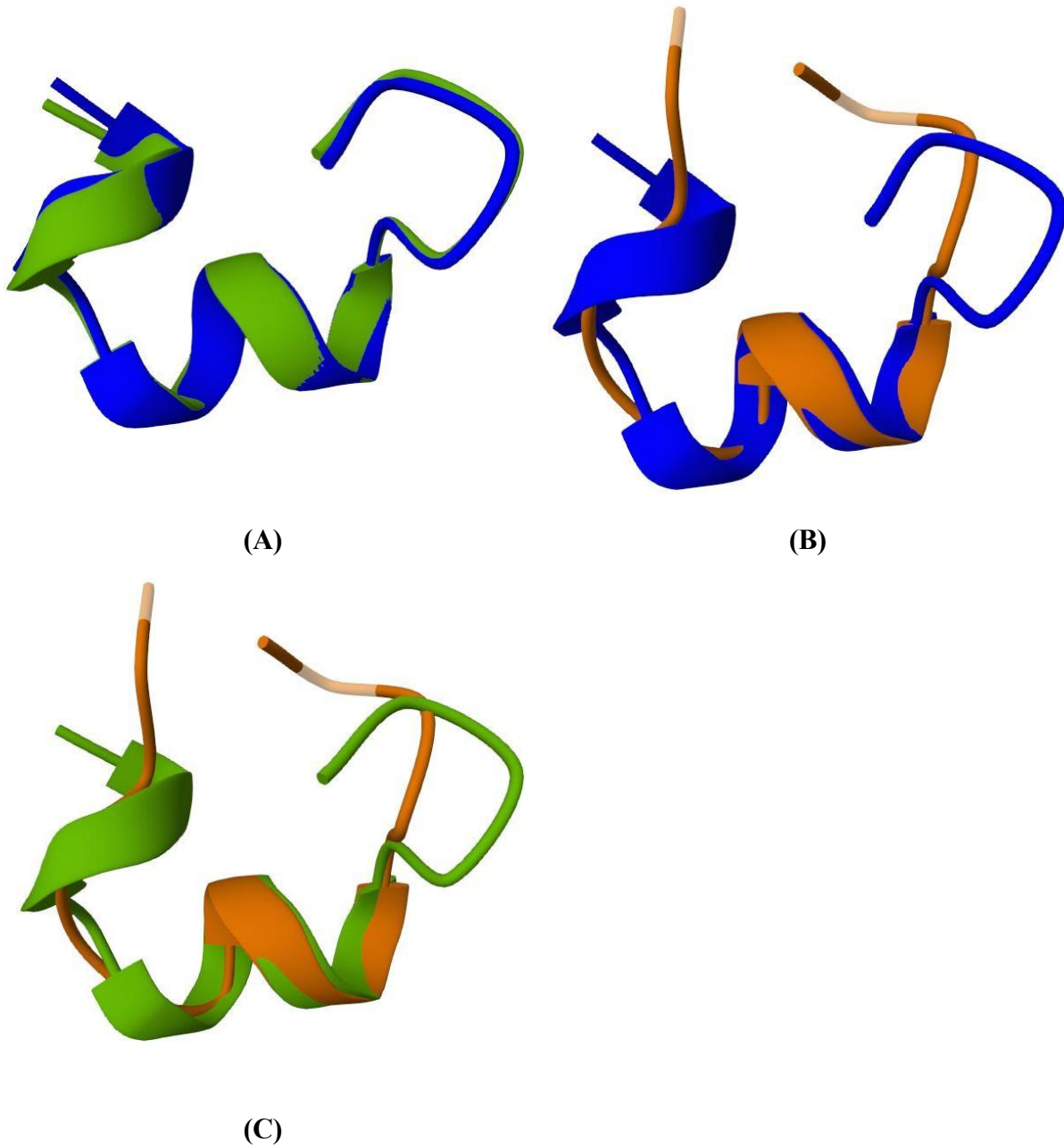


Figura 7. Superposiciones estructurales de la α -conotoxina RgIA. (A) AF3 (azul) vs PDB (verde), (B) AF3 (azul) vs AFDB (naranja), (C) AFDB (naranja) vs PDB (verde)

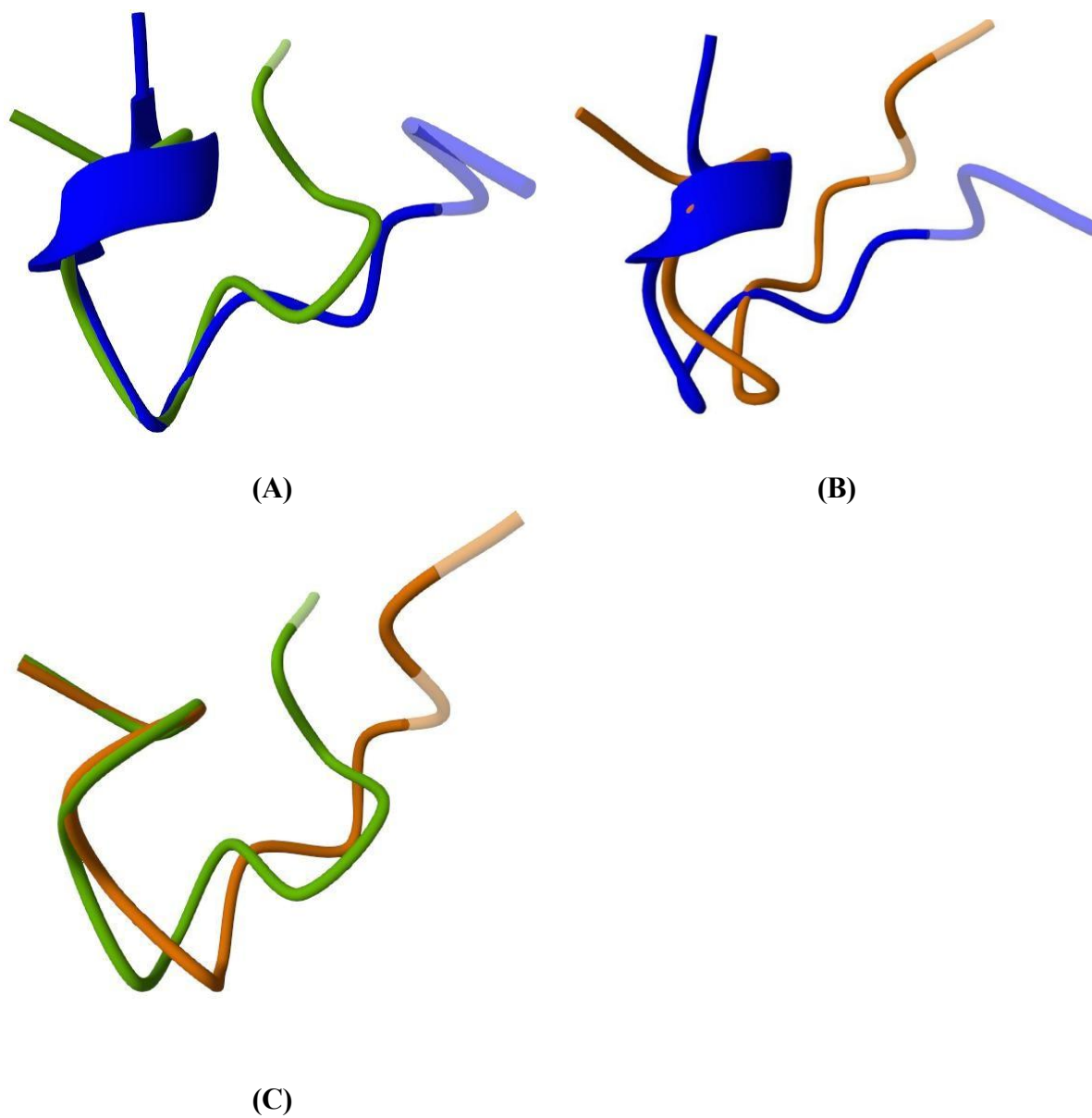
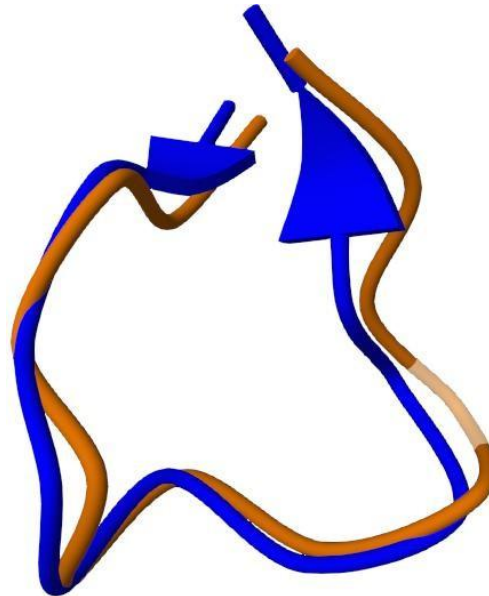


Figura 8. Superposiciones estructurales de la α -conotoxina SI. (A) AF3 (azul) vs PDB (verde), (B) AF3 (azul) vs AFDB (naranja), (C) AFDB (naranja) vs PDB (verde)



(A)



(B)



(C)

Figura 9. Superposiciones estructurales de la α -conotoxina Vc1.1. (A) AF3 (azul) vs PDB (verde), (B) AF3 (azul) vs AFDB (naranja), (C) AFDB (naranja) vs PDB (verde)

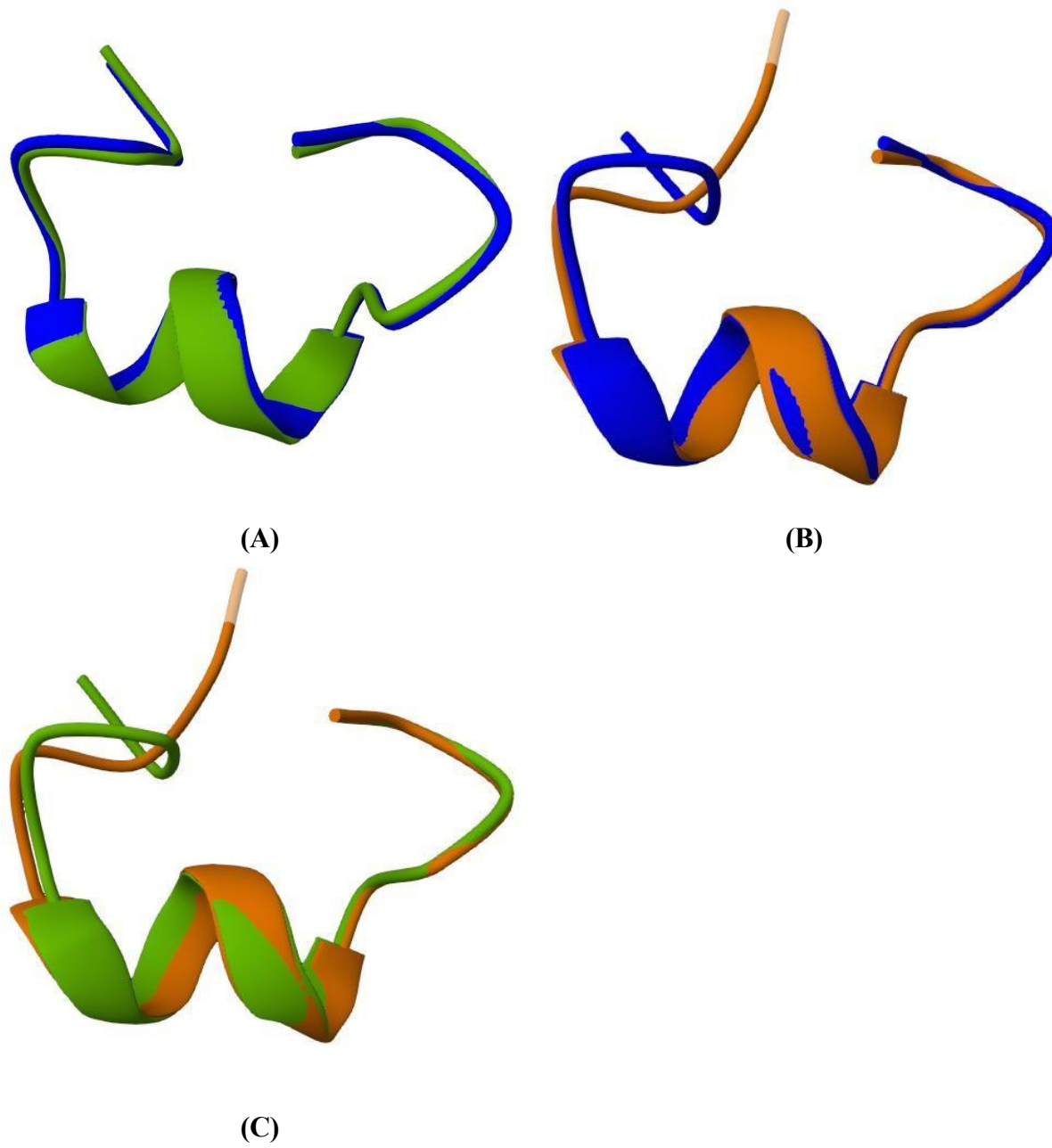
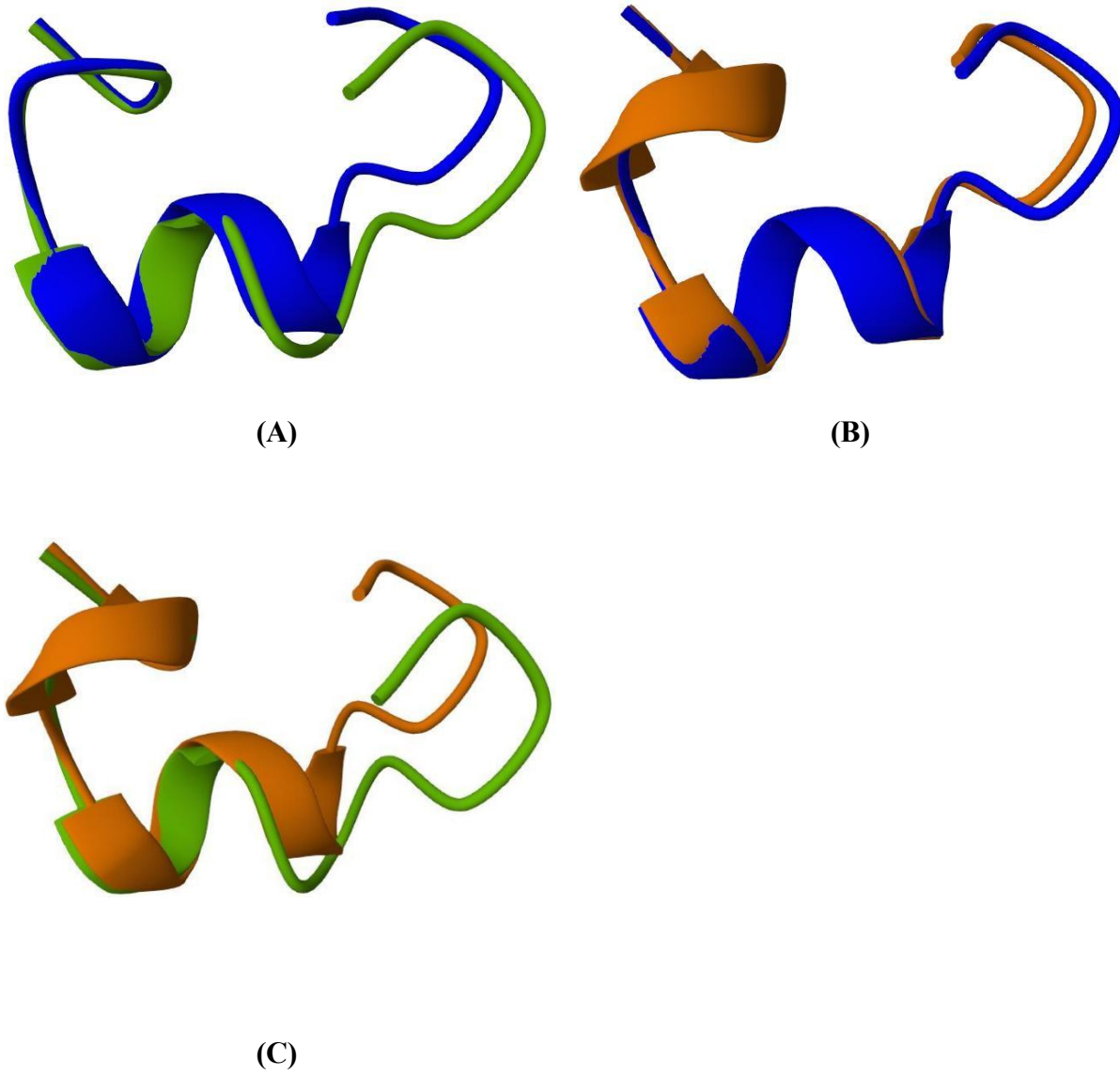


Figura 10. Superposiciones estructurales de la α -conotoxina TxIB. (A) AF3 (azul) vs PDB (verde), (B) AF3 (azul) vs AFDB (naranja), (C) AFDB (naranja) vs PDB (verde)



5. Discusión

En la tabla 1 se puede observar que las α -conotoxinas escogidas presentan longitudes comprendidas en un rango de 12 a 18 residuos. En general, los péptidos más cortos tienen menos grados de libertad conformacional, lo que favorece una mayor precisión en la predicción de su estructura. Por el contrario, los péptidos más largos tienden a ser más flexibles, lo que puede dificultar la precisión del modelado estructural. Además, el espacio entre los bucles (loop spacing) es un factor clave en la topología de estas moléculas, ya que bucles más largos incrementan la flexibilidad y la complejidad del modelado. Así, se esperaría que AF3 obtenga mejores resultados al predecir la estructura de α -conotoxinas con bucles cortos, como ImI, GI, SI y RgIA, mientras que podría enfrentar mayores dificultades con aquellas que presentan bucles más extensos, como MII, PnIA, GIC, EI, Vc1.1 y TxIB.

Por otro lado, la comparación entre las estructuras predichas por AF3 y las referencias del PDB requiere considerar ciertos aspectos metodológicos. Cuando las estructuras experimentales se obtienen mediante RMN, la evaluación es más compleja, ya que este método genera un conjunto de conformaciones que reflejan la dinámica de la molécula en solución, en lugar de una sola estructura estática. En cambio, las estructuras determinadas por difracción de rayos X corresponden a una conformación fija, lo que facilita una comparación más directa con los modelos generados por AF3. Por lo tanto, parte de las diferencias observadas entre las predicciones de AF3 y las referencias experimentales puede deberse no solo a posibles limitaciones del modelo, sino también a la propia variabilidad estructural de las α -conotoxinas en solución.

La tabla muestra que, a pesar de las diferencias en longitud, secuencia y origen de las α -conotoxinas, todas comparten una arquitectura muy conservada, caracterizada principalmente por la presencia de enlaces disulfuro y patrones específicos en el

espaciamiento de los bucles. Estas características estructurales ofrecen un contexto favorable para la predicción con AF3. No obstante, es importante considerar que factores como la flexibilidad de los bucles, la longitud de cada péptido y el tipo de método experimental utilizado como referencia (como RMN o cristalografía) pueden introducir variaciones que influyen en la precisión de las predicciones del modelo.

Esta tabla demuestra que, aunque las α -conotoxinas presentan variabilidad en longitud, secuencia y origen, mantienen una arquitectura altamente conservada dominada por enlaces disulfuro y patrones de loop spacing característicos. Estas restricciones estructurales sugieren un entorno favorable para la predicción mediante AF3. Sin embargo, factores como la flexibilidad de los loops, la longitud del péptido y la naturaleza del método experimental de referencia (RMN vs cristalografía) introducen fuentes de variabilidad que deben considerarse al interpretar la precisión del modelo.

La Tabla 2 revela que AF3 logra una notable precisión al predecir la estructura de las α -conotoxinas, con valores de RMSD generalmente por debajo de 1 Å en comparación con las estructuras experimentales del PDB, y superando de manera constante a los modelos previos de la AFDB. Este desempeño se ve respaldado por altos valores de pLDDT, lo que refleja una gran confianza en las predicciones obtenidas. Un aspecto especialmente relevante es la correcta identificación de los puentes disulfuro en todos los modelos generados por AF3, a diferencia de las discrepancias observadas en los modelos de AFDB, lo que resalta la capacidad de AF3 para captar restricciones estructurales clave. Sin embargo, existen casos particulares, como el de RgIA, en los que se observa una desviación estructural considerable, lo que sugiere limitaciones relacionadas con la flexibilidad conformacional o con características propias del péptido. Además, la exclusión de ciertos residuos en el cálculo del RMSD resalta la influencia de regiones altamente dinámicas, lo que indica que parte de la variabilidad observada responde tanto a la naturaleza intrínseca de estas moléculas como a posibles limitaciones del modelo.

En el caso de la toxina EI, se observó una diferencia importante entre las herramientas computacionales analizadas. Aunque la estructura disponible en la AlphaFold Database (AFDB) mostraba un valor de confianza pLDDT de 80,70, el alineamiento con el modelo experimental evidenció un RMSD elevado de 2,821 Å debido a una incorrecta formación de los puentes disulfuro. En cambio, el modelo generado por AF3 corrigió este problema, logrando una precisión mucho mayor, con un RMSD de solo 0,61 Å respecto al PDB. Este resultado destaca que AF3 tiene una mayor sensibilidad para captar la química de las cisteínas y resolver correctamente la arquitectura de estos enlaces, superando las limitaciones de autoconfianza que presentaban las versiones anteriores del algoritmo.

El análisis de la toxina GI puso en evidencia la capacidad de AF3 para diferenciar entre distintas formas conformacionales y encontrar el estado de menor energía. Mientras el modelo de la AFDB, con un pLDDT de 62,13, predijo una configuración tipo cinta—relacionada experimentalmente con baja actividad biológica en las α -conotoxinas—, AF3 logró identificar la conformación globular nativa, con una divergencia de 3,010 Å respecto a la base de datos. Este hallazgo es especialmente relevante para el estudio, ya que muestra que AF3 prioriza la predicción del estado funcionalmente activo de la molécula, en lugar de otras formas posibles que no representan la estructura nativa.

En el caso de la toxina GIC, el estudio evidenció un proceso de refinamiento progresivo en la predicción de la estructura. Aunque el modelo de la AFDB ya ofrecía una muy buena coincidencia con la estructura experimental (RMSD de 0,970 Å), AF3 consiguió mejorar aún más esta precisión, reduciendo la desviación a 0,835 Å. La pequeña diferencia observada entre los modelos de AF3 y AFDB (0,540 Å) sugiere que ambos algoritmos tienden a converger en el caso de plegamientos especialmente estables, y que la nueva versión no solo confirma los logros previos, sino que los lleva a un nivel de detalle aún mayor, acercándose a la resolución atómica.

Los resultados de las toxinas ImI, PnIA, Vc1.1 y TxIB validan la fiabilidad de AF3 en

la obtención de precisiones consistentes, con RMSD menores a 1,0 Å (ej. 0,458 Å en ImI). En estos casos, se excluyeron uno o dos átomos durante el cálculo del RMSD, al estar localizados dichos residuos en los extremos N y C-terminales (como la Gly1 y Cys2 en la Vc1.1), regiones de alta flexibilidad, su rechazo técnico no compromete la integridad del núcleo estructural, demostrando que la predicción del armazón central de las α -conotoxinas es prácticamente idéntica a la realidad experimental.

El caso de la α -conotoxina RgIA se puede considerar una excepción, la racha de RMSD de 1.0 Å se pierde al obtener un RMSD de 3,051 Å para AF3 vs PDB. Debido a que la estructura de esta α -conotoxina fue calculada experimentalmente mediante RMN, se demuestra que posee una alta flexibilidad en sus loops. Por otro lado, AF3 evidencia, con un valor de pLDDT de 92,80, la alta confianza que tiene de haber encontrado una estructura termodinámicamente estable. Se considera que la falla no está en el plegamiento primario debido a que la conectividad de los puentes de disulfuro fue predicha correctamente, al contrario, puede deberse a una divergencia conformacional. Mientras que AF3 predijó un conformero estable energéticamente con alta confianza, la estructura experimental del PDB refleja un conformero distinto naturalmente estable. Esto comprueba la idea de una isomería conformacional.

En la Figura 1A se aprecia un notable solapamiento entre la estructura predicha por AF3 (azul) y la estructura experimental (verde), con una coincidencia especialmente marcada en el núcleo estructural estabilizado por los puentes disulfuro. Las diferencias menores se localizan principalmente en las regiones periféricas, lo que sugiere que corresponden a segmentos con mayor flexibilidad, evidenciando que AF3 logra una representación altamente precisa del plegamiento tridimensional en comparación con la estructura experimental del PDB. Por otro lado, la Figura 1B muestra una discrepancia más pronunciada entre las predicciones de AF3 (azul) y AFDB (naranja), particularmente en la orientación de los bucles y en la conformación global del péptido lo cual es consistente con su valor de RMSD alto y

errores en la asignación de puentes disulfuro. Finalmente, la Figura 1C evidencia que el modelo de AFDB (naranja) no logra reproducir de manera adecuada la estructura experimental (verde), presentando diferencias más notorias que las observadas con AF3 y afectando tanto la geometría general como la disposición de segmentos estructurales clave.

En la Figura 2A se observa una buena concordancia global entre la estructura predicha por AF3 (azul) y la experimental (verde), especialmente en el núcleo estructural. Sin embargo, a diferencia de lo observado para EI, en este caso se identifican desviaciones más notables en regiones específicas, principalmente en los segmentos tipo loop. Aunque se trata de un péptido corto y con un espaciamiento reducido entre los bucles, la predicción no logra el mismo grado de precisión que en otras α -conotoxinas. Esto sugiere que existen factores adicionales, como la organización particular de los loops, que también pueden influir en el rendimiento del modelo. Asimismo, se discute que al ser la estructura experimental resuelta por rayos X, este péptido puede adoptar una conformación antinatural al cristalizarse, de manera que la estructura predicha (estable) por AF3 puede que no coincida con la real. No obstante, la resolución de este método (1,20 Å) sugiere una predicción excelente. En la Figura 2B, la comparación entre AF3 (azul) y AFDB (naranja) revela diferencias marcadas tanto en la conformación global como en la orientación de los segmentos estructurales que se pueden asociar a la asignación incorrecta de los puentes disulfuro. Por último, la Figura 2C evidencia que el modelo de AFDB (naranja) no logra reproducir fielmente la estructura experimental (verde), mostrando desviaciones significativas a lo largo de la conformación completa.

En la Figura 3A se observa una adecuada superposición global entre la estructura predicha por AF3 (azul) y la estructura experimental (verde), con una coincidencia particularmente destacada en el núcleo estructural estabilizado por los puentes disulfuro. No obstante, se identifican desviaciones localizadas en regiones tipo loop, especialmente en el extremo más flexible del péptido. Por su parte, la Figura 3B evidencia una similitud relativamente alta entre las predicciones de AF3 (azul) y AFDB (naranja) en comparación con

los casos analizados previamente (EI, GI); aunque persisten diferencias en la orientación de ciertos segmentos, el plegamiento global resulta ampliamente comparable. Finalmente, la Figura 3C muestra una superposición razonablemente buena entre el modelo de AFDB (naranja) y la estructura experimental del PDB (verde), si bien se observan desviaciones en las regiones periféricas. Este resultado puede explicarse por la correcta asignación de los puentes disulfuro en el modelo de AFDB para esta toxina, lo que pone de manifiesto la importancia fundamental de estas interacciones en la determinación de la estructura global.

En la Figura 4A se aprecia una superposición prácticamente perfecta entre la estructura predicha por AF3 (azul) y la estructura experimental (verde), con una notable coincidencia tanto en el núcleo estructural como en las regiones periféricas. Esto confirma la alta precisión de AF3 en la predicción del plegamiento de alfa-conotoxinas de menor longitud. Por otro lado, la Figura 4B muestra una diferencia estructural clara entre AF3 (azul) y AFDB (naranja), especialmente en la orientación del péptido y en la conformación de los loops, lo cual podría atribuirse, en parte, a una asignación incompleta de los puentes disulfuro. Finalmente, la Figura 4C evidencia que el modelo de AFDB (naranja) presenta discrepancias respecto a la estructura experimental (verde), aunque estas no son tan marcadas como en otros casos (EI, GI). En conjunto, estos resultados refuerzan la idea de que AF3 proporciona una mejora constante en la precisión estructural, especialmente en la correcta representación de restricciones químicas clave.

En la Figura 5A se observa una buena correspondencia global entre la estructura predicha por AF3 (azul) y la estructura experimental (verde), especialmente en el núcleo estabilizado por los puentes disulfuro. No obstante, se identifican desviaciones claras en las regiones tipo loop, sobre todo en los extremos del péptido, lo que pone de manifiesto la influencia de la flexibilidad conformacional en la precisión del modelo. Por otro lado, la Figura 5B revela una diferencia estructural notable entre AF3 (azul) y AFDB (naranja), principalmente en la conformación de los loops y en la orientación general de la cadena

polipeptídica. La Figura 5C evidencia que el modelo de AFDB (naranja) presenta discrepancias notables respecto a la estructura experimental (verde), en particular en las regiones periféricas y segmentos tipo loop, probablemente asociadas a errores en la asignación de los puentes disulfuro y a una menor capacidad para capturar detalles estructurales finos. Si bien AF3 no elimina por completo las limitaciones inherentes a la predicción de regiones flexibles, representa una mejora significativa en la modelación de péptidos con mayor complejidad estructural.

La Figura 6A confirma la notable precisión de AF3 en la predicción de la estructura de PnIA, mostrando una coincidencia casi perfecta entre el modelo predicho (azul) y la estructura experimental (verde). Esta alta fidelidad se refleja también en un sobresaliente valor de pLDDT (97,30), lo que indica que AF3 es especialmente eficaz para modelar estructuras altamente rígidas y con abundancia de cisteínas.

Por otro lado, las Figuras 6B y 6C muestran pequeñas diferencias en las regiones terminales más flexibles (N y C-terminal). Si bien la hélice central de PnIA se mantiene bien definida en todos los modelos, la estructura de AFDB (naranja) presenta leves desviaciones en la posición de los extremos respecto tanto a AF3 como al PDB. Este resultado pone de manifiesto que AF3 no solo reproduce de manera precisa la arquitectura general del plegamiento, sino que también introduce un refinamiento geométrico superior en las zonas más flexibles, logrando acercar la predicción aún más a la estructura observada experimentalmente.

En la Figura 7A, el valor de RMSD de 3,051 Å se manifiesta en diferencias notables en la amplitud de los loops entre la predicción de AF3 (azul) y la estructura experimental del PDB (verde), aunque el núcleo estructural y la conectividad de los puentes disulfuro se mantienen conservados. Por su parte, la Figura 7B evidencia una alta similitud visual entre AF3 y AFDB, con un RMSD de 1,073 Å, lo que indica una fuerte concordancia entre ambos modelos computacionales. Las discrepancias observadas respecto a la estructura experimental

pueden estar asociadas a las distintas conformaciones que la RMN capta en solución, así como a las propias limitaciones de los modelos de predicción para reflejar la diversidad estructural de la molécula. Estudios recientes señalan que AlphaFold suele converger hacia conformaciones dominantes o promedio, lo que puede limitar la representación de estados alternativos identificados experimentalmente (Ramasamy et al., 2025; Wang et al., 2025). Por ello, los resultados obtenidos para RgIA deben interpretarse como evidencia de una diferencia conformacional entre los modelos computacionales y la estructura experimental, sin que los datos actuales permitan determinar con certeza el origen de dicha discrepancia.

La Figura 8A muestra una coincidencia prácticamente total entre la predicción de AF3 y la estructura experimental del PDB, lo que confirma la alta fiabilidad del software en péptidos que presentan puentes disulfuro bien definidos, respaldado además por un destacado valor de pLDDT de 97,35. En la Figura 8B, el RMSD de 0,620 Å entre AF3 y AFDB se traduce en pequeñas diferencias geométricas, especialmente en las regiones de los loops. Aunque la AFDB (con un pLDDT de 75,60) ofrece una representación adecuada, AF3 consigue un refinamiento aún mayor en la estructura, corrigiendo las leves desviaciones presentes en la base de datos pública, como se observa en la Figura 8C. Este caso pone de relieve que AF3 aporta una precisión superior en la orientación de los giros estructurales, un aspecto clave para el reconocimiento molecular de este tipo de toxinas.

La Figura 9 refleja el nivel de precisión más alto alcanzado en este estudio. El modelo generado por AF3 (Figura 9A) mostró una fidelidad estructural sobresaliente frente al PDB, con un RMSD de solo 0,258 Å. Aunque la AFDB también presentó un RMSD bajo (0,199 Å), se detectó una deficiencia importante en la correcta formación de los puentes disulfuro, ya que las conexiones Cys 1-4 y Cys 2-3 corresponden al isómero cinta, el cual no presenta actividad biológica considerable. Además, fue necesario excluir seis átomos en el modelo de la AFDB en comparación con AF3, donde solo fueron dos, lo que resalta que AF3 no solo logra una

mayor precisión en la estructura principal, sino que también mantiene una coherencia química integral.

En la Figura 10A, AF3 demostró una precisión destacada, alcanzando un RMSD de 0,428 Å y un pLDDT de 95,92, lo que lo posiciona como el modelo más cercano a la estructura experimental. Es importante señalar que, aunque la AFDB logró predecir correctamente la conectividad de los puentes disulfuro en esta toxina, su desviación con respecto al PDB fue algo mayor (0,614 Å). La marcada similitud entre ambos modelos computacionales (RMSD de 0,320 Å en la Figura 10B) indica que la TxIB presenta un plegamiento especialmente estable, el cual puede ser capturado por ambas versiones del software, aunque AF3 logra alcanzar un nivel óptimo de resolución atómica.

6. Conclusiones

En conclusión, AF3 demuestra un desempeño consistentemente superior al de los modelos de la AFDB en la predicción de α -conotoxinas, alcanzando precisiones sub-Ángstrom ($\text{RMSD} \leq 1,0 \text{ \AA}$) en el 80% de los casos analizados. Esta ventaja no se limita únicamente al aspecto geométrico, sino que también abarca la dimensión química, ya que AF3 logró asignar correctamente la conformación globular de los puentes disulfuro en las α -conotoxinas EI, GI, ImI, MII, PnIA y Vc1.1, donde la base de datos previa mostraba conectividades de tipo cinta. Estos resultados ponen de manifiesto que el algoritmo ha avanzado considerablemente en su capacidad para comprender y modelar las restricciones covalentes de las cisteínas.

El análisis comparativo confirma que los altos valores de pLDDT obtenidos por AF3 (>92) son un indicador confiable de la precisión estructural, a diferencia de lo observado en la AFDB, donde valores de confianza moderados no siempre garantizaban una topología adecuada. Se concluye que la habilidad de AF3 para considerar el contexto químico de los ligandos y los enlaces disulfuro convierte a la métrica pLDDT en una herramienta interna de validación especialmente útil para el diseño racional de péptidos en farmacología molecular.

Se concluye que la evaluación de la precisión de AF3 está estrechamente ligada al método experimental utilizado para obtener las referencias en el PDB. Mientras que la cristalografía de rayos X, por su carácter estático, permite comparaciones más directas, el caso de la RgIA muestra que desviaciones elevadas de RMSD ($\approx 3,0 \text{ \AA}$) en estructuras determinadas por RMN no necesariamente representan un fallo del modelo, sino que reflejan la existencia de diferentes conformaciones en equilibrio. AF3 tiende a predecir las formas de mínima energía, que pueden coexistir con la flexibilidad dinámica capturada por RMN en solución.

Se concluye que, si bien la arquitectura de las α -conotoxinas es altamente conservada, aspectos como la longitud del péptido y el espaciamiento de los bucles influyen en el grado de

refinamiento alcanzado por el modelo. En toxinas con una estructura especialmente rígida, como PnIA y TxIB, AF3 logra niveles óptimos de precisión. Por el contrario, las regiones N y C-terminales, debido a su mayor flexibilidad, constituyen la principal fuente de variabilidad. Sin embargo, al excluir técnicamente estas regiones del cálculo del RMSD, se confirma que el armazón central funcional se mantiene prácticamente idéntico a la estructura experimental.

En definitiva, los resultados de este estudio consolidan a AF3 como una herramienta muy valiosa para la química computacional estructural, ya que es capaz de generar modelos tridimensionales que muestran una alta concordancia con las estructuras experimentales disponibles. La identificación precisa del plegamiento globular en la α -conotoxina GI destaca la habilidad del modelo para reproducir rasgos estructurales clave en este tipo de péptidos. Así, los modelos producidos por AF3 ofrecen un punto de partida sólido para futuros análisis computacionales, siempre y cuando se tenga en cuenta las limitaciones propias de los métodos de modelado estructural.

7. Recomendaciones

Se sugiere a quienes trabajan en bioinformática y farmacología molecular considerar a AlphaFold 3 (AF3) como la herramienta de referencia para el modelado *de novo* de conotoxinas. Gracias a su precisión para corregir errores en la formación de puentes disulfuro, que a menudo se encuentran en bases de datos estáticas como la AFDB, AF3 ofrece modelos iniciales más confiables para estudios de acoplamiento molecular y simulaciones de dinámica molecular, facilitando así un avance más sólido en la investigación de estas moléculas.

Se recomienda que, al emplear modelos computacionales como referencia en estudios de RMN, no se descarten automáticamente aquellos modelos que presenten un RMSD elevado (mayor a 2,0 Å) si cuentan con un valor alto de pLDDT y una conectividad covalente

adecuada. En estos casos, es preferible realizar un análisis de posibles isómeros conformacionales, reconociendo que el modelo predicho podría reflejar un estado de mínima energía que complementa la diversidad estructural observada experimentalmente en solución. Sería pertinente realizar un estudio de las estructuras experimentales obtenidas por RMN y por rayos X para la α -conotoxina RgIA de manera que se pueda entender mejor el comportamiento de la toxina en ambos medios.

Para lograr métricas de precisión que reflejen mejor la calidad del núcleo funcional de los péptidos, se aconseja estandarizar la exclusión de los residuos terminales (N y C-terminal) al calcular el RMSD. Tal como se evidenció en este estudio, la elevada flexibilidad de aminoácidos como la glicina inicial puede aumentar artificialmente el margen de error, incluso en modelos cuyo armazón central es prácticamente idéntico a la estructura nativa. Esta práctica permitirá una evaluación más justa y precisa de la calidad estructural en estudios de modelado de péptidos.

Dada la alta precisión demostrada por AF3 en el modelado de monómeros como lo son las α -conotoxinas, se sugiere que futuros estudios exploren su capacidad para predecir complejos conotoxina-receptor, como los nAChRs. Evaluar el desempeño del algoritmo en estos sistemas permitiría comprobar si la exactitud sub-Ángstrom observada en este trabajo también se refleja en la predicción de los sitios de unión y en la estimación precisa de las energías de interacción biológica.

Se aconseja a la comunidad académica llevar a cabo una revisión crítica de los modelos depositados en la AFDB especialmente en el caso de péptidos pequeños y con alto contenido de cisteínas. Los resultados de este estudio muestran que las bases de datos masivas pueden incluir isómeros que no son biológicamente activos, por lo que es recomendable volver a predecir estas secuencias empleando las versiones más recientes del algoritmo. Así, se refuerza la integridad química y la calidad de las investigaciones basadas en estos modelos.

Referencias Bibliográficas

Abramson, J., Adler, B., Dunger, J., Evans, R., Gaztañaga, P., Goldin, V., Jumper, J., Koos, K., Lamm, M., Licata, A. M., Ma, A., Moltzau, N. J., Steinegger, M., Thrapp, N., Wiles, T., Wong, K., Wu, Z., Xiao, H., & Hassabis, D. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, *630*(8016), 493–500. <https://doi.org/10.1038/s41586-024-07487-w>

Akdel, M., Pires, D. E., Pardo, E. P., Jänes, J., Zalevsky, A. O., Mészáros, B., Bryant, P., Good, L. L., Laskowski, R. A., Ferrer-Lucia, R., Alverson, A. J., Strand, N. S., Cheese, R., Pereira, J., Schwede, T., & Beltrao, P. (2022). A structural biology community assessment of AlphaFold2 applications. *Nature Structural & Molecular Biology*, *29*(11), 1056–1067. <https://doi.org/10.1038/s41594-022-00849-w>

Akondi, K. B., Muttenthaler, M., Dutertre, S., Kaas, Q., Craik, D. J., Lewis, R. J., & Alewood, P. F. (2014). Discovery, synthesis, and structure–activity relationships of conotoxins. *Chemical Reviews*, *114*(11), 5815–5847. <https://doi.org/10.1021/cr400401e>

Anfinsen, C. B. (1973). Principles that govern the folding of protein chains. *Science*, *181*(4096), 223–230. <https://doi.org/10.1126/science.181.4096.223>

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., & Bourne, P. E. (2000). The Protein Data Bank. *Nucleic Acids Research*, *28*(1), 235–242. <https://doi.org/10.1093/nar/28.1.235>

Bittrich, S., Segura, J., Duarte, J. M., Burley, S. K., & Rose, Y. (2024). RCSB Protein Data Bank: Exploring protein 3D similarities via comprehensive structural alignments. *Bioinformatics*, *40*(6), btae370. <https://doi.org/10.1093/bioinformatics/btae370>

Blundell, T. L., Jhoti, H., & Abell, C. (2002). High-throughput crystallography for lead

discovery in drug design. *Nature Reviews Drug Discovery*, 1(1), 45–54.
<https://doi.org/10.1038/nrd706>

Callaway, E. (2020). ‘It will change everything’: DeepMind’s AI makes gigantic leap in solving protein structures. *Nature*, 588(7837), 203–204. <https://doi.org/10.1038/d41586-020-03348-4>

Chang, L., & Perez, A. (2025). Rapid estimation of protein folding pathways from sequence alone using AlphaFold2. *Nature Communications*, 17(1), 170. <https://doi.org/10.1038/s41467-025-66870-x>

Desiere, F., Deutsch, E. W., King, N. L., Nesvizhskii, A. I., Mallick, P., Eng, J., & Aebersold, R. (2006). The PeptideAtlas project. *Nucleic Acids Research*, 34, D655–D658. <https://doi.org/10.1093/nar/gkj040>

Dill, K. A., & MacCallum, J. L. (2012). The protein-folding problem, 50 years on. *Science*, 338(6110), 1042–1046. <https://doi.org/10.1126/science.1219021>

Dobson, C. M. (2003). Protein folding and misfolding. *Nature*, 426(6968), 884–890. <https://doi.org/10.1038/nature02261>

Evans, R., O’Neill, M., Pritzel, A., Antropova, N., Senior, A., Green, T., & Hassabis, D. (2022). Protein complex prediction with AlphaFold-Multimer. *bioRxiv*. <https://doi.org/10.1101/2021.10.04.463034>

Flores, C. M. C., Kaas, Q., Barghi, N., & Lluisma, A. O. (2025). Diversity and novelty of venom peptides in vermivorous cone snails. *Marine Drugs*, 23(7), 266. <https://doi.org/10.3390/md23070266>

Giglio, M. L., Flórez-Salcedo, P., Azam, L., Watkins, M., Koch, T. L., Basgall-De la Rosa, E., Douglass, A. D., McIntosh, J. M., Olivera, B. M., & Gajewiak, J. (2025). An N-terminally elongated peptide from *Conus rolani* Defines a New Class of Ribbon α -Conotoxins Targeting Muscle nAChRs. *The FASEB Journal*, 39(12), e70698. <https://doi.org/10.1096/fj.202500721RR>

Jin, A. H., Muttenthaler, M., Dutertre, S., Himaya, S. W. A., Kaas, Q., Craik, D. J., Lewis, R. J., & Alewood, P. F. (2019). Conotoxins: Chemistry and biology. *Chemical Reviews*, *119*(21), 11510–11549. <https://doi.org/10.1021/acs.chemrev.9b00338>

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., & Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, *596*(7873), 583–589. <https://doi.org/10.1038/s41586-021-03819-2>

Karplus, M., & McCammon, J. A. (2002). Molecular dynamics simulations of biomolecules. *Nature Structural Biology*, *9*(9), 646–652. <https://doi.org/10.1038/nsb0902-646>

Kühlbrandt, W. (2014). The resolution revolution. *Science*, *343*(6178), 1443–1444. <https://doi.org/10.1126/science.1251652>

Laskowski, R. A., Hutchinson, E. G., Michie, A. D., Wallace, A. C., Jones, M. L., & Thornton, J. M. (1997). PDBsum. *Trends in Biochemical Sciences*, *22*(12), 488–490. [https://doi.org/10.1016/S0968-0004\(97\)01140-7](https://doi.org/10.1016/S0968-0004(97)01140-7)

Lawson, C. L., Baker, M. L., Best, C., Bi, C., Dougherty, M., Feng, P., & Chiu, W. (2016). EMDatabank unified data resource. *Nucleic Acids Research*, *44*(D1), D396–D403. <https://doi.org/10.1093/nar/gkv1126>

Levinthal, C. (1969). How to fold graciously. En *Mossbauer Spectroscopy in Biological Systems* (pp. 22–24). University of Illinois Press.

Li, R., Yu, J., Ye, D., Liu, S., Zhang, H., Lin, H., & Deng, K. (2025). Conotoxins: Classification, Prediction, and Future Directions in Bioinformatics. *Toxins*, *17*(2), 78. <https://doi.org/10.3390/toxins17020078>

Mao, K. L., Fu, J. X., Chen, J., Liao, Y. L., Huang, M. L., Chen, Z., Lin, L. M., Shi, Q., & Gao, B. M. (2026). From marine predator to pharmacology: Conotoxin diversity, discovery, and therapeutic potential. *Zoological Research*, *47*(2), 378–403. <https://doi.org/10.24272/j.issn.2095-8137.2025.553>

Minkiewicz, P., Dziuba, J., Iwaniak, A., Dziuba, M., & Darewicz, M. (2008). BIOPEP

database. *Journal of AOAC International*, 91(4), 965–980.

Murzin, A. G., Brenner, S. E., Hubbard, T., & Chothia, C. (1995). SCOP. *Journal of Molecular Biology*, 247(4), 536–540. [https://doi.org/10.1016/S0022-2836\(05\)80134-2](https://doi.org/10.1016/S0022-2836(05)80134-2)

Ramasamy, P., Zuallaert, J., Martens, L., & Vranken, W. F. (2025). Assessing the relation between protein phosphorylation, AlphaFold3 models and conformational variability. bioRxiv, 2025-04. <https://doi.org/10.1101/2025.04.14.648669>

Orengo, C. A., Michie, A. D., Jones, S., Jones, D. T., Swindells, M. B., & Thornton, J. M. (1997). CATH. *Structure*, 5(8), 1093–1108. [https://doi.org/10.1016/S0969-2126\(97\)00260-8](https://doi.org/10.1016/S0969-2126(97)00260-8)

Senior, A. W., Evans, R., Jumper, J., Kirkpatrick, J., Sifre, L., Green, T., & Hassabis, D. (2020). Improved protein structure prediction. *Nature*, 577(7792), 706–710. <https://doi.org/10.1038/s41586-019-1923-7>

Tae, H. S., & Adams, D. J. (2023). Nicotinic acetylcholine receptors and α -conotoxins. *Pharmacological Research*, 191, 106747. <https://doi.org/10.1016/j.phrs.2023.106747>

Ulrich, E. L., Akutsu, H., Doreleijers, J. F., Harano, Y., Ioannidis, Y. E., Lin, J., & Markley, J. L. (2008). BioMagResBank. *Nucleic Acids Research*, 36, D402–D408. <https://doi.org/10.1093/nar/gkm957>

Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., & Velankar, S. (2022). AlphaFold Protein Structure Database. *Nucleic Acids Research*, 50(D1), D439–D444. <https://doi.org/10.1093/nar/gkab1061>

Vincler, M., Wittenauer, S., Parker, R., Ellison, M., Olivera, B. M., & McIntosh, J. M. (2006). Molecular mechanism for analgesia involving specific antagonism of $\alpha 9\alpha 10$ nicotinic acetylcholine receptors. *PNAS*, 103(47), 17880–17884. <https://doi.org/10.1073/pnas.0608715103>

Wang, G., Li, X., & Wang, Z. (2016). APD3 antimicrobial peptide database. *Nucleic Acids Research*, 44(D1), D1087–D1093. <https://doi.org/10.1093/nar/gkv1278>

Wang, H., Meng, X., & Zhang, Y. (2025). Biomolecular interaction prediction: The era of AI. *Advanced Science*, 12(38), e09501. <https://doi.org/10.1002/advs.202509501>

wwPDB Consortium. (2019). Protein Data Bank global archive. *Nucleic Acids Research*, 47(D1), D520–D528. <https://doi.org/10.1093/nar/gky949>