



UNIVERSIDAD DEL PAPALOAPAN
CAMPUS TUXTEPEC

INGENIERÍA EN BIOTECNOLOGÍA

**"OPTIMIZACIÓN DEL MODELADO *IN SILICO* DE LA INTERACCIÓN PROTEÍNA
CRY-RECEPTOR"**

TESIS

QUE PARA OBTENER EL TÍTULO DE:

Ingeniero en Biotecnología

PRESENTA:

ABIEZER ISAAC CONTRERAS SANTIAGO

Director: Dr. Miguel Ángel Peña Rico

SAN JUAN BAUTISTA TUXTEPEC, OAXACA. 2024



UNIVERSIDAD DEL PAPALOAPAN
CAMPUS TUXTEPEC

INGENIERÍA EN BIOTECNOLOGÍA

ACTA DE REVISIÓN DE TESIS

En la ciudad de San Juan Bautista Tuxtepec, Oaxaca, el día 10 de abril de 2024 a las 16:15 h, los miembros de la comisión revisora de tesis designada por la Jefatura de Carrera de la Ingeniería en Biotecnología se reunieron en la sala de juntas del Instituto de Biotecnología de la Universidad del Papaloapan, con la finalidad de examinar la tesis titulada "Optimización del modelado *in silico* de la interacción proteína Cry-receptor" presentada por el alumno Abiezer Isaac Contreras Santiago, con número de matrícula 16090007, aspirante al título de Licenciatura.

Después de intercambiar opiniones, los miembros de la comisión manifestaron que la tesis **satisface** los requisitos señalados por las disposiciones reglamentarias vigentes, otorgando su **aprobación** para que el aspirante pueda proceder con el proceso de titulación.

Tuxtepec, Oaxaca, a 10 de abril de 2024.

ATENTAMENTE
LA COMISIÓN REVISORA

Dr. Miguel Ángel Peña Rico
Profesor Investigador Titular "A"
Universidad del Papaloapan
Director de Tesis

Dra. Ana Karin Navarro Martínez
Profesor Investigador Titular "B"
Universidad del Papaloapan
Revisor de Tesis

Dra. Karla Viridiana Castro Cerritos
Profesor Investigador Asociado "B"
Universidad del Papaloapan
Revisor de Tesis

Dr. Enrique Villatobos Amador
Profesor Investigador Titular "A"
Universidad del Papaloapan
Revisor de Tesis

Dr. Paul Mauricio Sanchez Ocampo
Profesor Investigador Titular "A"
Universidad del Papaloapan
Revisor de Tesis

Dr. Edgar Baldemar Sepúlveda García
Profesor Investigador Asociado "C"
Universidad del Papaloapan
Revisor de Tesis

Campus Loma Bonita
Av. Ferrocarril S/N, Col. Ciudad Universitaria, Loma
Bonita, Oaxaca C.P. 68400
Tel/Fax: 01 281 872 92 30

www.unpa.edu.mx

Campus Tuxtepec
Circuito Central N° 200, Col. Parque Industrial,
Tuxtepec, Oaxaca. C.P. 68301
Tel/Fax: 01 287 875 9240



UNIVERSIDAD DEL PAPALOAPAN
CAMPUS TUXTEPEC

INGENIERÍA EN BIOTECNOLOGÍA

Tuxtepec, Oaxaca, a 24 de abril de 2024
Oficio No. JCIB/008/04/2024

Lic. Yesenia Barrientos Arenal
Jefa de Servicios Escolares
Universidad del Papaloapan

Con base en el dictamen de la comisión revisora, se autoriza la impresión del trabajo de tesis del alumno **Abiezer Isaac Contreras Santiago** con número de matrícula 16090007 que tiene como título definitivo "Optimización del modelado *in silico* de la interacción proteína Cry-receptor" para ser presentado como trabajo de tesis para obtener el título de Licenciado en Ingeniería en Biotecnología, toda vez que cumple satisfactoriamente con la reglamentación establecida para tal fin.

Dra. Ana Karin Navarro Martínez (Presidente, Universidad del Papaloapan)
Dr. Paul Mauricio Sanchez Ocampo (Secretario, Universidad del Papaloapan)
Dra. Karla Viridiana Castro Cerritos (Vocal, Universidad del Papaloapan)
Dr. Edgar Baldemar Sepúlveda García (Primer Suplente, Universidad del Papaloapan)
Dr. Enrique Villalobos Amador (Segundo Suplente, Universidad del Papaloapan)

Sin más por el momento le envío un cordial saludo.

Atentamente
Terra uberrima, mens aperta
Bou Lo-tama, chí jí jú



JEFATURA DE INGENIERÍA
EN BIOTECNOLOGÍA
CAMPUS TUXTEPEC

Jacqueline Capataz Tafur
Dra. Jacqueline Capataz Tafur
Jefa de Carrera de Ingeniería en Biotecnología
Universidad del Papaloapan

c.c.p. Dr. Miguel Ángel Peña Rico, Director de tesis, Para su conocimiento
c.c.p. Abiezer Isaac Contreras Santiago, Alumno, Para su conocimiento
c.c.p. Archivo

Héctor López Arjona
Vo.Bo. M.C. Héctor López Arjona
Vice Rector Académico
Universidad del Papaloapan



VICE-RECTORIA
ACADEMICA

Campus Loma Bonita
Av. Ferrocarril S/N, Col. Ciudad Universitaria, Loma
Bonita, Oaxaca C.P. 68400
Tel/Fax: 01 281 872 92 30

www.unpa.edu.mx

Campus Tuxtepec
Circuito Central N° 200, Col. Parque Industrial,
Tuxtepec, Oaxaca. C.P. 68301
Tel/Fax: 01 287 875 9240

HOJA DE ORIGINALIDAD.

El presente trabajo no ha sido aceptado o empleado para el otorgamiento de título o grado diferente o adicional al actual. La tesis es resultado de las investigaciones del autor, excepto donde se indican las fuentes de información consultadas. El autor otorga su consentimiento a la **Universidad del Papaloapan** para la reproducción del documento con el fin del intercambio bibliotecario siempre y cuando se indique la fuente.

AGRADECIMIENTOS

Al Dr. Miguel Ángel Peña Rico, mi director de tesis, por la confianza y sus conocimientos aportados en el desarrollo de esta investigación.

A la Dra. Ana Karín Navarro Martínez, Dr. Paul Mauricio Sánchez Ocampo, Dra. Karla Viridiana Castro Cerritos, Dr. Edgar Baldemar Sepúlveda y al Dr. Enrique Villalobos, miembros de mi comité revisor, por el tiempo, consejos y conocimientos para culminar este trabajo.

Al Dr. Edgar García López, por su asesoramiento en este trabajo, sus conocimientos en bioinformática fueron de gran apoyo para la investigación.

Nuevamente, a la Dra. Ana Karín Navarro Martínez, por el apoyo, consejos, confianza y por permitirme formar parte de su equipo de trabajo desde cuarto semestre y durante mi etapa de licenciatura.

Al CONAHCYT por la beca otorgada, del proyecto #321009. Convocatoria de ciencia básica y/o ciencia de frontera. Modalidad: Paradigmas y controversias de la ciencia 2022.

A Dios por todas sus bendiciones para conmigo y mi familia.

DEDICATORIA

Este trabajo se lo dedico a mis padres por su apoyo, amor, educación, consejos que me han brindado durante mi desarrollo personal y académico, ellos han sido un pilar muy importante en mi vida, y siempre agradeceré todo sacrificio y esfuerzo que han hecho por mí y por mi hermano.

A mis amigos y compañeros de la licenciatura, Ameyalli, Karla, Paola, Luisa, Ángelica, Anahí, Ángel Ortela, también a mis compañeras y amigas de maestría, Carla y Lidia, con quienes pasé y he pasado gratos momentos en la universidad, y me han apoyado y brindado su amistad en momentos buenos y malos.

I. ÍNDICE GENERAL

II. ÍNDICE DE FIGURAS.....	IX
III. ÍNDICE DE TABLAS.....	XI
IV. ABREVIATURAS.....	XIII
V. RESUMEN.....	XIV
VI. ABSTRACT.....	XV
1.-INTRODUCCIÓN.....	1
2. MARCO TEÓRICO.....	3
2.1 Generalidades.....	3
2.2 Parasporinas (PS).....	4
2.3 Clasificación de las Parasporinas.....	4
2.4 Dominios de las parasporinas.....	5
2.5 Propuestas de mecanismo de acción para las PS.....	6
2.6 Modelado de proteínas.....	8
2.7 Docking Molecular.....	10
2.8 Función de receptores en células normales y de cáncer.....	12
2.9 Función de receptores como dianas terapéuticas.....	13
2.10 Receptores en MCF-7.....	14
3.-JUSTIFICACIÓN.....	15
4.- HIPÓTESIS.....	16
5.- OBJETIVOS.....	17
5.1 Objetivo general:.....	17
5.2 Objetivos particulares:.....	17
6.- DIAGRAMA EXPERIMENTAL.....	18
7.- MATERIALES Y MÉTODOS.....	19
7.1 Búsqueda de receptores.....	19
7.2 Modelos de proteínas.....	19
7.3 Modelado de receptores.....	19
7.3.1 Modelado por homología.....	20
7.3.2 Modelado con Inteligencia Artificial (IA).....	20
7.4 Preparación de proteínas para docking.....	21
7.5 Regiones de interés.....	21
7.6 Docking Molecular.....	22
7.6.1 ClusPro.....	22
7.6.2 HDock.....	23
7.6.3 HEX.....	24

8.-RESULTADOS	25
8.1 Modelos Experimentales	25
8.2 Modelado de proteínas.....	26
8.2.1 Modelado por homología	26
8.2.2 Modelado con IA.....	30
8.3 Docking Molecular.....	33
8.3.1 HDOCK.....	33
8.3.1.1 MODELOS ALPHAFOLD	35
8.3.1.2 MODELOS SWISSMODEL	37
8.3.1.3 DOCKING RECEPTORES HUMANOS	39
8.3.2 ClusPro.....	43
8.3.2.1 MODELOS ALPHAFOLD	43
8.3.2.2 MODELOS SWISS MODEL	45
8.3.2.3 DOCKING RECEPTORES HUMANOS	47
8.3.3 HEX	51
8.3.3.1 MODELOS ALPHAFOLD	51
8.3.3.2 MODELOS SWISS MODEL	52
8.3.3.3 DOCKING RECEPTORES HUMANOS	54
8.3.4 Top 3 receptores para proteínas Cry.....	56
8.3.5 Regiones de interacción.....	59
9. DISCUSIÓN.....	65
10. CONCLUSIONES	70
11. BIBLIOGRAFÍA.....	71

II. ÍNDICE DE FIGURAS

Figura 1. Comparación estructural entre la Parasporina-2, Proteína no toxica 26 kDa y la Parasporina-4 Modelada (Xu <i>et al.</i> , 2014).....	6
Figura 2. Diagrama experimental	18
Figura 3. Interfaz Modelling del servidor en línea SWISS-MODEL.....	20
Figura 4. Interfaz de inicio de Alphafold Database.	21
Figura 5. Interfaz principal del servidor ClusPro.....	22
Figura 6. Sección de carga de trabajo del servidor.....	23
Figura 7. Interfaz principal del servidor HDOCK.....	23
Figura 8. HEX, software de docking molecular.....	24
Figura 9. Pantalla de preparación de docking en HEX.	24
Figura 10. Modelos generados por homología mediante SWISS-MODEL de los receptores APN1 (a,b), APN2 (c,d) y ALP (e,f), paleta de color rainbow (a, c y d) y la paleta de colores de QMEANDisCo (b, d y f), donde las regiones en azul indican las regiones con un mejor valor de QMEANDisCo y las rojas regiones con valores bajos.	28
Figura 11. Gráficos de Ramachandran de los modelos APN1 (a), APN2 (b) y ALP (c), generados con SWISS-MODEL.....	29
Figura 12. Modelos generados por Inteligencia Artificial mediante AlphaFold de los receptores APN1 (a,b), APN2 (c,d) y ALP (e,f), paleta de color rainbow (a, c y d) y paleta de color pLLDT (b, d y f).	31
Figura 13. Gráficos PAE de los modelos APN1 (a), APN2 (b) y ALP (c), generados con AlphaFold y barra de color/error de referencia (d).....	32
Figura 14. Gráficos de Ramachandran de los modelos APN1 (a), APN2 (b) y ALP (c), generados con AlphaFold.	33
Figura 15. Top 3 complejos docking Cry-Receptores de insecto (AlphaFold) arrojados por HDOCK.....	36
Figura 16. Top 3 complejos docking Cry-Receptores de insecto (SWISS-MODEL) arrojados por HDOCK.	38
Figura 17. Top 3 complejos docking Cry1Aa-Receptores humanos arrojados por HDOCK.	40
Figura 18. Top 3 complejos docking Cry1Ac-Receptores humanos arrojados por HDOCK.	42
Figura 19. Top 3 complejos docking Cry-Receptores de insecto (AlphaFold) arrojados por ClusPro.....	44
Figura 20. Top 3 complejos docking Cry-Receptores de insecto (SWISS-MODEL) arrojados por ClusPro.....	46
Figura 21. Top 3 complejos docking Cry1Aa-Receptores humanos arrojados por ClusPro.	48
Figura 22. Top 3 complejos docking Cry1Ac-Receptores humanos arrojados por ClusPro.	50
Figura 23. Complejo docking Cry1Aa-Receptor APN2 (SWISS-MODEL) arrojado por HEX.	51
Figura 24. Top 3 complejos docking Cry-Receptores de insecto (SWISS-MODEL) arrojados por HEX.....	53

Figura 25. Top 3 complejos docking Cry1Aa-Receptores humanos arrojados por HEX.	54
Figura 26. Top 3 complejos docking Cry1Ac-Receptores humanos arrojados por HEX.	55
Figura 27. Región de interacción entre Cry1Aa y receptor TFRC (docking con HDOCK).	59
Figura 28. Región de interacción entre Cry1Ac y receptor TFRC (docking con HDOCK). ...	59
Figura 29. Región de interacción entre Cry1Aa y receptor TFRC (docking con ClusPro). ...	60
Figura 30. Región de interacción entre Cry1Ac y receptor TFRC (docking con ClusPro). ...	60
Figura 31. Región de interacción entre Cry1Aa y receptor TFRC (docking con HEX).	61
Figura 32. Región de interacción entre Cry1Ac y receptor TFRC (docking con HEX).....	61
Figura 33. Región de interacción entre Cry1Aa y receptor EGFR (docking con HDOCK)....	62
Figura 34. Región de interacción entre Cry1Ac y receptor EGFR (docking con HDOCK)....	62
Figura 35. Región de interacción entre Cry1Ac y receptor EGFR (docking con ClusPro)....	62

III. ÍNDICE DE TABLAS

Tabla 1. Receptores empleados en el trabajo y tipos de cáncer en los que se expresa.	19
Tabla 2. Información general de receptores humanos obtenida del Protein Data Bank.	25
Tabla 3. Información general de los receptores de insecto obtenidos de UniProt.	26
Tabla 4. Datos del templado para cada receptor.	27
Tabla 5. Score de diferentes indicadores estructurales para los modelos por homología. ...	29
Tabla 6. Score de diferentes indicadores estructurales para los modelos por IA.	33
Tabla 7. Valores de referencia de los indicadores LGscore y MaxSub.	34
Tabla 8. Score de los indicadores LGscore y MaxSub para los modelos experimentales. ...	34
Tabla 9. Score de los indicadores LGscore y MaxSub para las proteínas modeladas por homología y con IA.	34
Tabla 10. Scores de docking de Cry1Aa vs. Receptores de insectos generados por IA	35
Tabla 11. Scores de docking de Cry1Ac vs. Receptores de insectos generados por IA	35
Tabla 12. Scores de docking de Cry1Aa vs. Receptores de insectos generados por homología.....	37
Tabla 13. Scores de docking de Cry1Ac vs. Receptores de insectos generados por homología.....	37
Tabla 14. Scores de docking de Cry1Aa vs. Receptores humanos.....	39
Tabla 15. Scores de docking de Cry1Ac vs. Receptores humanos.....	41
Tabla 16. Scores de docking de Cry1Aa vs. Receptores de insectos generados por IA	43
Tabla 17. Scores de docking de Cry1Ac vs. Receptores de insectos generados por IA	43
Tabla 18. Scores de docking de Cry1Aa vs. Receptores de insectos generados por homología.....	45
Tabla 19. Scores de docking de Cry1Ac vs. Receptores de insectos generados por homología.....	45
Tabla 20. Scores de docking de Cry1Aa vs. Receptores humanos.....	47
Para en el docking de Cry1Ac con receptores humanos, los resultados están alojados en la tabla 21, los receptores de mejor desempeño fueron el VIPR seguido del TFRC, este último también fue el segundo mejor complejo con el programa HDOCK (tabla 15).....	49
Tabla 21. Scores de docking de Cry1Ac vs. Receptores humanos.....	49
Tabla 22. Scores de docking de Cry1Aa vs. Receptor APN generado por IA	51
Tabla 23. Scores de docking de Cry1Aa vs. Receptores de insectos generados por homología.....	52
Tabla 24. Scores de docking de Cry1Ac vs. Receptores de insectos generados por homología.....	52
Tabla 25. Scores de docking de Cry1Aa vs. Receptores humanos.....	54
Tabla 26. Scores de docking de Cry1Ac vs. Receptores humanos.....	55
Tabla 27. Tres mejores receptores de insecto para Cry1Aa y Cry1Ac según HDOCK.....	56

Tabla 28. Tres mejores receptores humanos para Cry1Aa y Cry1Ac según HDOCK	56
Tabla 29. Tres mejores receptores de insecto para Cry1Aa y Cry1Ac según ClusPro.....	57
Tabla 30. Tres mejores receptores humanos para Cry1Aa y Cry1Ac según ClusPro	57
Tabla 31. Tres mejores receptores de insecto para Cry1Aa y Cry1Ac según HEX	58
Tabla 32. Tres mejores receptores humanos para Cry1Aa y Cry1Ac según HEX.....	58
Tabla 33. Residuos aminoacídicos de los receptores involucrados en el docking.	63
Tabla 34. Dominios de unión del ligando y receptor involucrados en el docking.....	64

IV. ABREVIATURAS

β-PFT	Toxinas β formadoras de poro
ALP	Fosfatasa alcalina
APN	Aminopeptidasa-N
Bt	<i>Bacillus thuringiensis</i>
CASP	Evaluación crítica de la predicción de la estructura de proteínas (por sus siglas en inglés “Critical Assessment of protein Structure Prediction”)
CD44	Receptor de ácido hialurónico
Cry	Toxinas o proteínas cristal
DL	Aprendizaje profundo (“Deep learning” en inglés)
EGFR	Receptor del factor de crecimiento epidérmico
ErbB	Oncógeno viral de la leucemia eritroblástica
ER	Receptor de estrógeno
FR	Receptor folato
GDTS	Puntuación total de la prueba de distancia global (por sus siglas en inglés “Global Distance Test Total Score”)
GMQE	Estimación de la calidad del modelo global (por sus siglas en inglés “Global Model Quality Estimate”)
GPI	Glicosilfosfatidilinositol
IA	Inteligencia artificial
PAE	Error de alineado predicho
PDB	Protein data bank (base de datos de estructuras proteicas)
pLDDT	prueba de diferencia de distancia local por residuo (por sus siglas en inglés “per-residue local distance difference test”)
PS	Parasporina
TFRC	Receptor de transferrina
TBM	Modelización basada en plantilla
UniProt	The Universal Protein Resource (repositorio de datos de proteínas)
VEGF	Factor de crecimiento vascular endotelial
VIP	Péptido intestinal vasoactivo

V. RESUMEN

Las parasporinas, producidas por *Bacillus thuringiensis* (*B. thuringiensis*), se caracterizan por su actividad citotóxica selectiva hacia células cancerosas en comparación con las células normales, esto las hace potencialmente útiles en el desarrollo de terapias contra el cáncer, se ha observado que estas proteínas pueden reconocer y unirse a receptores específicos en la superficie de las células cancerosas. Existe cierta similitud entre las parasporinas y las proteínas (toxinas) Cry, ambos poseen tres dominios, aunque los dominios están dispuestos de manera diferente en los dos tipos de proteínas, tienen secuencias similares, y las parasporinas tienen un mayor grado de similitud de secuencia con las toxinas Cry de tres dominios que con los otros tipos de toxinas Cry. En el presente trabajo se realizó una búsqueda de receptores sobreexpresados en la línea celular MCF-7 de cáncer de mama, que interactuarán *in silico* con dos proteínas Cry (Cry1Aa y Cry1Ac) mediante docking molecular, utilizando tres programas para evaluar la interacción receptor-ligando, siendo estos, HDOCK, ClusPro y HEX; a su vez para corroborar el desempeño de estos programas, se realizó docking entre las Cry con los receptores de insectos ya reportados para cada Cry, usando técnicas de modelado para obtener las estructuras 3D de estos receptores y empleando dos estrategias para modelar: 1) modelado por homología y 2) modelado con inteligencia artificial. Los mejores modelos fueron los obtenidos por IA, además que el docking de estos los ligó a la Cry correspondiente en literatura. El docking de receptores humanos mostró consistencia entre los 3 programas, siendo el receptor de transferrinas el que obtuvo las mejores puntuaciones, sugiriendo así, emplear estos receptores para pruebas *in silico* con parasporinas, seguido de una validación experimental *in vivo*.

VI. ABSTRACT

Parasporins, produced by *B. thuringiensis*, are characterized by their selective cytotoxic activity towards cancer cells compared to normal cells, making them potentially useful in the development of cancer therapies. It has been observed that these proteins can recognize and bind to specific receptors on the surface of cancer cells. There is some similarity between parasporins and Cry proteins (toxins). Both possess three domains, although the domains are arranged differently in the two types of proteins. They have similar sequences, and parasporins show a higher degree of sequence similarity to three-domain Cry toxins than to other types of Cry toxins. In this study, a search for overexpressed receptors in the MCF-7 breast cancer cell line was performed, which could interact *in silico* with two Cry proteins (Cry1Aa and Cry1Ac) through molecular docking. Three programs, namely HDOCK, ClusPro, and HEX, were used to evaluate the receptor-ligand interaction. Additionally, to validate the performance of these programs, docking was performed between Cry proteins and insect receptors that were previously reported for each Cry. This was achieved using modeling techniques to obtain the 3D structures of these receptors and employing two strategies, homology modeling, and modeling with artificial intelligence. The best models were obtained using AI, and the docking of these models linked them to the corresponding Cry protein mentioned in the literature. The docking of human receptors showed consistency among the three programs, with the transferrin receptor obtaining the highest scores. This suggests employing these receptors for *in silico* testing with parasporins, followed by experimental validation *in vivo*.

1.-INTRODUCCIÓN

Bacillus thuringiensis (*B. thuringiensis*) es una bacteria gram positiva, facultativamente aeróbica y formadora de esporas, durante su fase de esporulación, produce inclusiones parasporales que contienen varias delta(δ)-endotoxinas insecticidas y ha sido utilizado como bioinsecticida en campos agrícolas durante décadas (Wong *et al.*, 2010). Durante las últimas dos décadas, se ha descubierto que las proteínas parasporales de *B. thuringiensis* exhiben efecto citotóxico en células cancerígenas humanas. En el año 2000, Mizuki *et al.* utilizaron por primera vez la palabra "parasporina" para describir una nueva proteína con una citotoxicidad única. Hoy en día, el término "parasporina" se define como "proteínas parasporales de *B. thuringiensis* y bacterias relacionadas que no son hemolíticas, pero son capaces de matar selectivamente células cancerosas"; desde el descubrimiento de Mizuki, se han descubierto otras parasporinas, todas las cuales presentan propiedades no hemolíticas y anticancerígenas, por lo tanto, el gran reservorio de toxinas de *B. thuringiensis* en la naturaleza sigue siendo objeto de investigación constante porque son buenos candidatos potenciales para el tratamiento del cáncer (Wong, 2010).

El mecanismo de acción de las parasporinas contra las células cancerosas objetivo aún no se comprende completamente. Sin embargo, la información disponible ha revelado que las parasporinas exhiben una variedad de mecanismos de acción para eliminar las células cancerosas (Mendoza-Almanza *et al.*, 2020). Estas proteínas comparten similitudes en su funcionamiento con las toxinas Cry, ya que ambas son altamente específicas hacia ciertos tipos de células, en el caso de las toxinas Cry, se ha establecido que su especificidad depende del reconocimiento de receptores específicos en la membrana celular, como cadherina, aminopeptidasa-N y fosfatasa alcalina, en insectos (Mendoza-Almanza *et al.*, 2020). Así también, existen similitudes entre los receptores de insecto, para las Cry, y los receptores humanos en células de cáncer para las parasporinas, como el caso de APN (N-aminopeptidasa), que podría ser un receptor en el mecanismo de acción de PS2Aa1, ya que la actividad de la parasporina disminuyó considerablemente en una línea celular de cáncer colorrectal, cuando se trataron con un inhibidor de APN, además, se sabe que PS2Aa1 está relacionada con las proteínas Cry, que utilizan APN como receptor en el intestino medio de los insectos (Suarez-Barrera *et al.*, 2022).

El docking molecular (o acoplamiento molecular), es un procedimiento de simulación por computadora para predecir la conformación de un complejo receptor-ligando, donde el receptor suele ser una proteína o una molécula de ácido nucleico (ADN o ARN) y el ligando puede ser una molécula pequeña u otra proteína, otra manera de definirlo es, como un proceso de simulación donde se estima la posición de un ligando en un sitio de unión predicho o predefinido (Dias & De Azevedo, 2008). El docking de proteínas, es decir, la predicción de la estructura de un complejo proteína-proteína a partir de las estructuras de las proteínas individuales, ha evolucionado significativamente desde sus primeros días al incorporar funciones de energía más adecuadas y aprovechar el creciente conocimiento sobre las estructuras e interacciones de las proteínas, el conocimiento actual de los principios de interacción de proteínas es mucho mayor que antes, lo que ayuda a diseñar enfoques de acoplamiento mejores. El espectacular progreso en el hardware informático ha desempeñado un papel importante, abriendo nuevas formas de pensar la modelización de las interacciones de proteínas y a menudo permitiendo la implementación de ideas antiguas pero inviables en ese momento (Vakser, 2014).

La falta de información sobre posibles receptores involucrados en el mecanismo de acción de parasporinas conduce al uso de herramientas y metodologías para develar a estas proteínas involucradas, como lo es el docking molecular, debido a la similitud funcional y estructural entre las parasporinas con las proteínas Cry, en este trabajo se analizaron las interacciones Cry con receptores sobreexpresados en la línea MCF-7 para identificar *in silico* moléculas específicas de la célula cancerosa con la capacidad de interactuar con las proteínas Cry.

2. MARCO TEÓRICO

2.1 Generalidades

Bacillus thuringiensis (*B. thuringiensis*) es una bacteria omnipresente Gram positiva, con forma de bastoncillo y esporulante que se ha aislado en todo el mundo de una gran diversidad de ecosistemas, incluidos el suelo, el agua e insectos muertos (Palma *et al.*, 2014). Es una bacteria formadora de endosporas que se caracteriza por la presencia de un cristal de proteína dentro del citoplasma de la célula esporulante. Las proteínas que constituyen dicho cristal son tóxicas para algunos insectos, lo que explica el uso extensivo de *B. thuringiensis* como insecticida biológico (De Maagd *et al.*, 2001). Estas toxinas, identificadas como crystal (Cry) y citolíticas (Cyt), también conocidas como δ -endotoxinas, al inicio de la esporulación y durante la fase de crecimiento estacionario como inclusiones cristalinas parasporales (Palma *et al.*, 2014). Una vez ingeridos por los insectos, estos cristales se solubilizan en el intestino medio, las toxinas son proteolíticamente activadas por las proteasas del intestino medio, ya en su forma activa se unen a receptores específicos ubicados en la membrana celular del insecto, lo que conduce a la alteración celular y la muerte del insecto (Palma *et al.*, 2014).

Las toxinas Cry se clasifican en cuatro familias proteicas (Velásquez Cardona *et al.*, 2018): **1)** Toxinas Cry de tres dominios (Cry-3d), es una familia que agrupa la mayoría de toxinas producidas por *B. thuringiensis* y se caracterizan por presentar tres dominios en su estructura; se han reconocido ampliamente por su actividad insecticida contra lepidópteros, coleópteros, dípteros entre otros órdenes de insectos; **2)** La familia ETX/MTX, relacionadas estructural y secuencialmente con la toxinas ETX de *Clostridium epsilon* y con la toxina MTX de *Lysinibacillus sphaericus*; **3)** La familia de las toxinas Bin (binarias), que incluye proteínas que de forma individual no exhiben efecto tóxico, pero sí cuando interactúan con otra proteína y; **4)** La familia de las toxinas Cyt que presentan actividad hemolítica (Velásquez Cardona *et al.*, 2018).

Existen proteínas Cry que no presentaron actividad insecticida, pero que demostraron tener citotoxicidad sobre ciertas líneas de células de cáncer (Okumura *et al.*, 2008). Con este descubrimiento se estableció un nuevo grupo de proteínas Cry a las que se les denominó Parasporinas (PS), las cuales pueden clasificarse en diferentes familias según la homología de sus aminoácidos. Una característica común en las PS es la necesidad de una activación para

volverse biológicamente activa, este proceso puede realizarse de manera enzimática mediante tratamiento con proteinasa K y tripsina en ambiente alcalino o por algún otro método (Melo *et al.*, 2014).

Cada una de las PS que integran a las 6 familias poseen actividad sobre una o más líneas diferentes de cáncer. Se sugiere que estas líneas celulares sensibles a estas proteínas poseen un receptor específico único y/o sobreexpresado que tiene como ligando a la PS (Kayatama *et al.*, 2007), receptor que a la fecha no se ha logrado identificar. En este contexto, nuestro grupo se dio a la tarea de encontrar PS de diversas cepas de *B. thuringiensis*. De esta búsqueda, se han hallado dos PS, la A34 y AX2, con significativa actividad citotóxica sobre la línea celular MCF-7 sin causar daño a las células de origen no tumoral. Una de las premisas del grupo de investigación es dilucidar tanto el mecanismo de acción como los actores que participan en la actividad anticancerígena de estas proteínas.

2.2 Parasporinas (PS)

Son proteínas presentes en la inclusión parasporal de las cepas no insecticidas de *B. thuringiensis*, que no tiene actividad hemolítica, pero tiene actividad citocida, matando preferentemente las células cancerosas (Ekino *et al.*, 2014). La actividad citopática desencadenada por PS se ha estudiado desde las primeras observaciones utilizando ensayos *in vitro*. Su efecto se ha confirmado en diferentes células cancerosas como HeLa (cáncer de cuello uterino), HepG2 (carcinoma hepatocelular) y Molt-4 (leucemia linfoblástica aguda) donde se aplicaron diferentes subtipos de PS y produjeron altos niveles de toxicidad (Melo *et al.*, 2020).

2.3 Clasificación de las Parasporinas

El Comité de Clasificación y Nomenclatura de Parasporinas (Committee of Parasporin Classification and Nomenclature), ha agrupado las PS en seis familias (PS1-PS6). La nomenclatura inicia con el tipo de proteína (PS en el caso de parasporinas), después del tipo de proteína se le asignan cuatro rangos según el grado de identidad de aminoácidos por pares entre ellas (Ammons *et al.*, 2015). Se utilizan números arábigos para el primer y cuarto rango, y se asignan letras mayúsculas y minúsculas para el segundo y tercer rango, respectivamente. Este criterio no implica la toxicidad de las proteínas o incluso el modo de acción, solo se basa en la identidad de la secuencia de aminoácidos (Palma *et al.*, 2014).

De esta manera, a las proteínas que comparten menos del 45% de identidad por pares se les asigna un rango primario diferente (un número arábigo, por ejemplo, PS1 y PS2); a dos proteínas que comparten menos del 78% de identidad por pares se les asigna un rango secundario diferente (una letra mayúscula, por ejemplo, PS1A); a las proteínas que comparten menos del 95% de identidad por pares se les asigna un rango terciario diferente (una letra minúscula, por ejemplo, PS1Aa y PS1Ab); y, finalmente, para diferenciar entre proteínas que comparten más del 95% de identidad por pares, se asigna un rango cuaternario (un número arábigo, por ejemplo, PS1Aa1 y PS1Aa2) (Palma *et al.*, 2014).

2.4 Dominios de las parasporinas

De manera general, las parasporinas poseen 3 dominios proteicos, guardando similitud estructural con las proteínas Cry. El dominio I de estas, es sugerido como dominio de interacción, con proteínas GPI ancladas y con carbohidratos debido a la presencia de residuos aromáticos, este es el único dominio en donde se encuentran α -hélices (Xu *et al.*, 2014).

El dominio II se encuentra conformado por hojas beta, en el caso de la parasporina 2, se forma una hoja beta de cuatro hebras cerca del límite con el dominio I, asociado a las hélices de este por interacciones hidrofóbicas, la superficie interna de la hoja beta y la subestructura beta hairpin anfipática conforman un núcleo hidrofóbico (Akiba *et al.*, 2009). De las hebras beta separadas de las PS2 y PS4 modelada, se propuso que jugaban un rol en el proceso transmembrana, además, la PS2 posee residuos de serina y treonina en las hebras beta adyacentes al beta hairpin, secuencias ricas estos residuos se han hallado en otras β -PFT (toxinas β formadoras de poro) y se ha propuesto su participación en el proceso transmembrana (Xu *et al.*, 2014).

El dominio III, tomando como referente a la PS2, se compone de una lámina beta antiparalela de tres hebras y otra, pero de dos hebras, empaquetadas como β -sandwich (Akiba *et al.*, 2009). Debido a que durante la digestión proteolítica el extremo C-terminal de la PS2 puede ser eliminado, se expone parte del núcleo hidrofóbico que se encuentra dentro del sándwich β , creando una pequeña superficie hidrofóbica a lo largo de las hebras β . Se ha propuesto que la activación del extremo C-terminal sería necesaria para llevar a cabo la oligomerización (Xu *et al.*, 2014).

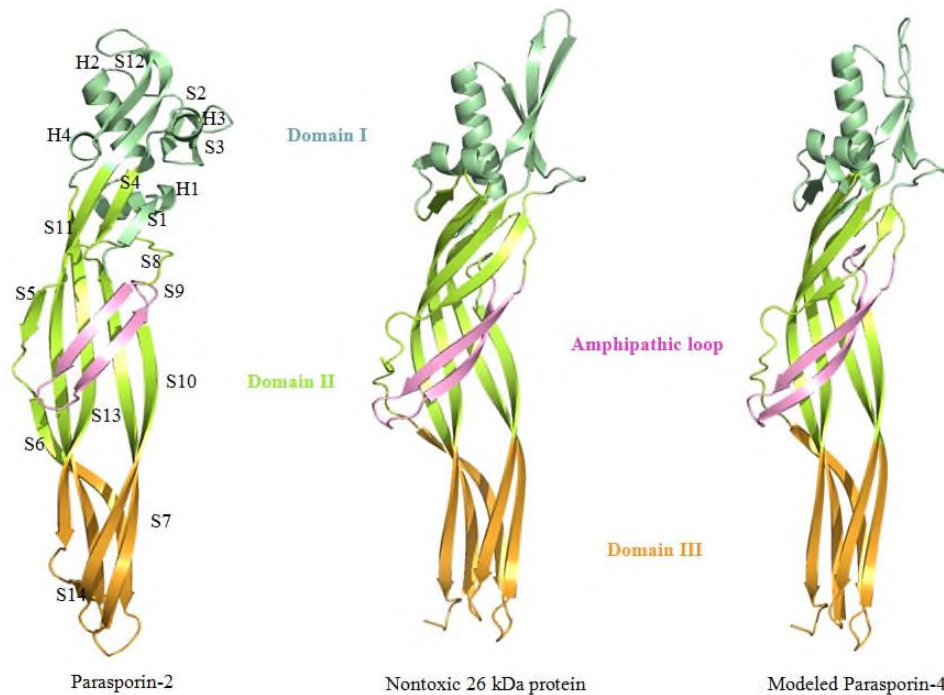


Figura 1. Comparación estructural entre la Parasporina-2, Proteína no toxica 26 kDa y la Parasporina-4 Modelada (Xu *et al.*, 2014).

La parasporina-3 (PS3Aa1 y PS3Ba1) del *B. thuringiensis* cepa A1462 fue la primera toxina reportada con una arquitectura típica de tres dominios (en la Figura 1 se muestran proteínas de 3 dominios) y es citotóxica para células cancerosas humanas, el alineamiento de secuencias indica que la PS3 tiene los mismos cinco bloques conservados que las toxinas Cry (como la Cry1, Cry3A o Cry2Aa) de tres dominios insecticidas, lo que sugiere que puede tener un modo de acción similar (Yamashita *et al.*, 2005). La PS6 muestra una similitud de secuencia del 56.4% con la toxina Cry2, lo cual sugiere que su mecanismo de acción puede estar también relacionada con las toxinas Cry de tres dominios (Nagamatsu *et al.*, 2010).

2.5 Propuestas de mecanismo de acción para las PS

El modo de acción de las PS difiere entre las familias, así como el espectro citotóxico y el nivel de actividad (Ohba *et al.*, 2008). Si bien no está claro cómo actúan, se cree que la mayoría de ellas siguen o tienen un mecanismo similar al descrito para la PS1. Es decir, provocan un aumento en el flujo de Ca^{2+} por la unión con los receptores de la proteína G o por las proteínas G heterotriméricas, induciendo, probablemente, apoptosis en las células cancerígenas (Katayama *et al.*, 2007).

El potencial apoptótico selectivo de la PS1 sobre las células cancerosas y no sobre las células sanas confirma que pertenece a una clase privilegiada de proteínas citotóxicas anticáncer (Katayama *et al.*, 2007).

Dos de las principales PS2 son producidas por las cepas *B. thuringiensis* A1547 (PS2Aa1) y A1470 (PS2Aa2). Son proteínas de 37 kDa como protoxina (Kitada *et al.*, 2006), y se diferencian entre ellas en tan solo cuatro residuos de aminoácidos. Dichas protoxinas al ser cortadas por proteasas por los extremos amino y carboxilo terminal, se genera un fragmento tóxico de aproximadamente 30 kD (Okumura *et al.*, 2013). Esta proteína ha mostrado citotoxicidad frente a líneas celulares como HepG2 (hepatocarcinoma), Jurkat (leucemia de linfocitos T) y poca actividad hacia células HeLa y hepatocitos normales (Ito *et al.*, 2004).

Las PS3 producidas por *B. thuringiensis* A1462 tienen un peso de 88 kDa y cuando está activa su tamaño es de 64 kDa. Aparentemente, estas PS son formadoras de poros, aumentando la permeabilidad de las células cancerígenas (Yamashita *et al.*, 2005).

La PS4 son de 31 kDa, pero una vez activadas son de 27 kDa presenta una fuerte actividad citotóxica contra líneas celulares humanas con cáncer como MOLT-4, CACO-2 (cáncer humano de colon), TCS (cáncer humano de cuello uterino), y células HL60 (Xu *et al.*, 2014).

La PS5 (Cry64Aa) es la única de esta clase reconocida a la fecha por el comité de nomenclatura de PS; fue aislada y purificada de la cepa A1100 de *B. thuringiensis* tohokuensis y presenta actividad citotóxica contra células T humanas con leucemia (MOLT-4). Esta PS presenta una masa molecular de 33.8 kDa en su forma de protoxina, pero cuando es cortada por proteinasa K se produce un fragmento activo de 29.8 kDa (Ekino *et al.*, 2014).

La PS6 (Cry63Aa) se aisló de las cepas de *B. thuringiensis* M109 y CP84 (Nagamatsu *et al.*, 2010). Las protoxinas de las PS1 y PS6 están relacionadas entre sí en algunos aspectos ya que comparten secuencias conservadas en sus primeros 50 residuos de aminoácidos. Además, tras ser activadas por proteólisis, dos de tres péptidos generados en el proceso, conforman un heterodímero tóxico (Lenina *et al.*, 2013) presenta mayor actividad citotóxica hacia células HepG2 y HeLa (Nagamatsu *et al.*, 2010).

Las PS tienen menos de un 25% de homología en la secuencia de aminoácidos con las toxinas Cry (Ohba *et al.*, 2008). Sin embargo, se cree que su mecanismo de acción es muy similar. Existen indicios que señalan que las PS reconocen receptores de membrana específicos en las células cancerosas para desencadenar la muerte celular (Santos *et al.*, 2022). La actividad citotóxica específica para las células cancerosas convierte a las PS en potencial agente terapéutico en la oncología clínica. Sin embargo, para que un medicamento obtenga la aprobación debe ser seguro y mostrar cierta eficacia. La comprensión de su mecanismo de acción debe ser clara de cómo funciona el fármaco. En este sentido, se sabe que algunas de las PS pueden causar citotoxicidad en las células tumorales, es a través del daño a nivel de membrana plasmática induciendo la activación del mecanismo apoptótico (Okassov *et al.*, 2014).

Por lo que surge la necesidad de identificar si las PS aisladas por el grupo de trabajo tiene como mecanismo de muerte la apoptosis.

En los primeros años de la identificación de toxinas Cry, se recurría a la caracterización de la proteína, seguida de un proceso extenso de identificación y clonación del gen, cuantos más genes fueron clonados se obtuvo más información sobre similitudes en sus secuencias, sin embargo, en los últimos años, la secuenciación de proteínas se ha convertido en una herramienta importante en la identificación de las secuencias y su homología.

2.6 Modelado de proteínas

El modelado computacional de la estructura de las proteínas ha sido objeto de un creciente interés en el campo de la bioinformática a lo largo del tiempo. Esta técnica ha demostrado ser una herramienta valiosa para complementar los esfuerzos experimentales de la biología estructural en la caracterización del universo de las proteínas (Schwede *et al.*, 2013). A través del modelado computacional, es posible predecir la estructura de una proteína, lo que puede acelerar y guiar los experimentos de biología estructural para determinar la estructura tridimensional de una proteína con mayor precisión.

Los avances en las técnicas de secuenciación de ADN han provocado un incremento en el número de secuencias, resulta evidente que no ha sido posible determinar experimentalmente la estructura de todas las proteínas de interés con las técnicas actuales a pesar del enorme progreso en la automatización de las técnicas de determinación de estructuras experimentales, la brecha en la estructura de proteínas se amplía aún más debido

a la disponibilidad de técnicas de secuenciación de próxima generación más sensibles (Schwede *et al.*, 2013).

Durante más de tres décadas, se han desarrollado y mejorado diversas técnicas computacionales para predecir la estructura de proteínas. Estas técnicas incluyen el uso de simulaciones de plegamiento basadas en funciones de energía físicas o empíricas, la creación de modelos a partir de fragmentos pequeños de estructuras ya conocidas, el enhebrado, que determina la compatibilidad entre una secuencia y un plegamiento experimental utilizando funciones de energía similares, y la modelización basada en plantilla (TBM), en la que se alinea una secuencia con una secuencia conocida de estructura mediante patrones de variación evolutiva, las estrategias que se han englobado dentro de la TBM incluyen la modelización por homología, la modelización comparativa y el reconocimiento de plegamiento (Kelley *et al.*, 2015).

Las técnicas de modelado por homología, también conocidas como modelado comparativo de estructuras de proteínas, se han desarrollado para construir modelos tridimensionales de una proteína objetivo a partir de su secuencia de aminoácidos, estos modelos se basan en el alineamiento con una proteína similar con una estructura conocida, llamada plantilla. En caso de no poder identificar una estructura de plantilla adecuada, se pueden utilizar métodos de predicción de estructura *de novo*, también conocidos como *ab initio*, para generar modelos tridimensionales de proteínas sin depender de una estructura de plantilla homóloga (Bordoli *et al.*, 2008).

El modelado por homología es uno de los métodos preferidos para construir modelos tridimensionales confiables *in silico* de una proteína cuando se pueden identificar estructuras que funcionen como plantillas. Estos modelos se utilizan ampliamente en aplicaciones como el cribado virtual, el diseño de experimentos de mutagénesis dirigida al sitio o en la racionalización de los efectos de las variaciones de secuencia (Bordoli *et al.*, 2008).

Aunque la modelización basada en plantillas es precisa, fiable y rápida para predecir estructuras, no puede utilizarse para determinar proteínas de nuevas familias o plegamientos, ya que depende de plantillas en una base de datos, este método falla si no hay plantillas adecuadas (Jisna & Jayaraj, 2021). La predicción de estructuras *ab initio* mediante estrategias basadas en el conocimiento reduce los requisitos de muestreo del espacio conformacional al reutilizar conformaciones de fragmentos de bases de datos estructurales, sin embargo, solo la modelización *ab initio* basada en la física, que no utiliza información estructural conocida, puede ofrecer información sobre los mecanismos de plegamiento reales. Para mejorar los

resultados, la mayoría de los algoritmos utilizan ahora un enfoque híbrido que combina técnicas basadas en el conocimiento y en la física (Jisna & Jayaraj, 2021).

El Deep Learning (DL) ha revolucionado la modelización de proteínas al mostrar una alta precisión en las puntuaciones de predicción. La predicción de las distancias de contacto residuo-residuo permitida por los potenciales de Deep Learning tiene éxito tanto en los objetivos basados en plantillas como en los de nuevo plegamiento. Los métodos de Deep Learning extraen automáticamente características complejas de las secuencias de proteínas y las utilizan para la predicción de estructuras, los sistemas de última generación demuestran la capacidad de la técnica de Deep Learning para reemplazar los métodos tradicionales tanto para objetivos fáciles como difíciles de nuevo plegamiento (Jisna & Jayaraj, 2021).

El rápido aumento de las puntuaciones Global Distance Test Total Score (GDTTS) en CASP14 (Critical Assessment of protein Structure Prediction) demuestra que la precisión experimental es un objetivo alcanzable para las técnicas de DL. Aunque los enfoques de Deep Learning proporcionan resultados más rápidos y fiables en la predicción de la estructura 3D, aún hay margen de mejora. Actualmente, existe un gran interés en la investigación de un sistema de extremo a extremo para la predicción de estructuras de proteínas, donde una única red neuronal profunda represente toda la tubería de predicción, sin embargo, con la disponibilidad de conjuntos de datos públicos, un aumento en la capacidad computacional y datos experimentales, se esperan más esfuerzos de los biólogos computacionales para reducir la brecha secuencia-estructura y resolver el problema en la predicción de estructuras (Jisna & Jayaraj, 2021).

2.7 Docking Molecular

El docking molecular (o acoplamiento molecular) es un procedimiento de simulación empleando software para la predicción de la formación de un complejo receptor-ligando, donde el receptor suele ser, por lo general, una proteína o una molécula de ácido nucleico (DNA o RNA) y el ligando es una proteína u otra molécula pequeña (Dias *et al.*, 2008). Es el método más común de diseño de fármacos basado en la estructura computacional (SBDD, por sus siglas en inglés) y ha sido utilizada ampliamente desde principios de la década de 1980 (Stanzione *et al.*, 2021).

La información obtenida de la orientación preferida de las moléculas unidas se puede utilizar para predecir perfiles de energía (como la energía libre de unión), la fuerza y la estabilidad (como la afinidad de unión y la constante de unión) de los complejos, esto se puede lograr

mediante una función de puntuación en el proceso de docking. En tiempos recientes, el docking molecular se ha estado utilizando comúnmente para predecir la orientación de unión de pequeñas moléculas candidatas a fármacos, a sus objetivos biomoleculares, como proteínas, carbohidratos y ácidos nucleicos, con el objetivo de determinar sus parámetros de unión tentativos (Agarwal & Mehrotra, 2016).

Los algoritmos de acoplamiento molecular realizan predicciones cuantitativas de la energía de unión y clasifican los compuestos acoplados en base a la afinidad de unión de los complejos ligando-receptor. Para identificar las conformaciones de unión más probables, se deben llevar a cabo dos pasos: en primer lugar, la exploración de un gran espacio conformacional que representa diversos modos de unión potenciales, y, en segundo lugar, la predicción precisa de la energía de interacción asociada con cada una de las conformaciones de unión predichas. Este proceso se lleva a cabo de forma cíclica mediante programas de acoplamiento molecular que evalúan la conformación del ligando utilizando funciones de puntuación específicas, y se repite recursivamente hasta alcanzar una solución de energía mínima (Ferreira *et al.*, 2015).

Durante el proceso de acoplamiento molecular la flexibilidad del receptor es crucial y puede ser regulada de dos maneras, la primera es mediante la modificación del receptor rígido estándar, mientras que la segunda implica el ingreso de múltiples receptores. Además, es necesario considerar la forma del ligando para obtener una mayor precisión en la predicción de la unión. Debido a estos factores, el acoplamiento molecular se puede dividir en tres tipos según el grado de simplificación: acoplamiento rígido, acoplamiento semiflexible y acoplamiento flexible (Tao *et al.*, 2019). El acoplamiento rígido significa que la conformación del sistema de acoplamiento (receptor y ligando) no cambia, es adecuado para sistemas grandes, como proteína-proteína y proteína-ácido nucleico, además, el acoplamiento rígido no requiere cálculos múltiples, lo que lo convierte en el método más simple.

En el acoplamiento semiflexible, se mantiene la conformación fija del receptor mientras que la conformación del ligando se modifica, por lo tanto, se puede fijar la longitud o ángulo de enlace de la parte no crítica de la estructura del ligando, este tipo de acoplamiento es adecuado para el acoplamiento de pequeñas moléculas y macromoléculas como proteínas o ácidos nucleicos con pequeñas moléculas de ligando (Tao *et al.*, 2019).

Durante el proceso de acoplamiento flexible, se permite la modificación de la conformación del sistema de acoplamiento (receptor y ligando), sin embargo, dado que el número de variables del receptor y del ligando aumenta con el número de átomos, se deben considerar

varios factores adicionales, entre ellos se incluyen la complejidad del cálculo y la dificultad inherente del proceso de acoplamiento (Tao *et al.*, 2019).

El acoplamiento de ligandos flexibles es teóricamente deseable para simular con precisión el proceso de unión en sistemas reales, pero enfrenta desafíos importantes los cuales incluyen el alto costo computacional para el muestreo, la estimación de la energía conformacional/tensión interna y la pérdida de la entropía conformacional del ligando en la unión. Por lo tanto, el acoplamiento de ligandos flexibles a menudo tiene una tasa de éxito más baja en la predicción del modo de unión en comparación con el acoplamiento de ligandos rígidos con las conformaciones unidas de complejos experimentales, además, incluso un pequeño cambio en la conformación del ligando, que es común durante el muestreo conformacional del ligando, puede conducir a una gran diferencia en las poses de acoplamiento resultantes (Huang, 2018).

2.8 Función de receptores en células normales y de cáncer

La comunicación y adaptación de las células a su entorno es esencial en la biología celular, esta capacidad se debe en gran medida a la presencia de receptores específicos en la superficie celular que son sensibles a las señales del ambiente, estas señales pueden provenir de diversas fuentes, como factores solubles (químicos, polipéptidos, proteínas, azúcares, etc.), ligandos unidos a otras células o a la matriz extracelular, estos receptores transmiten las señales a través de la membrana celular y, mediante la activación de vías de señalización intracelular, generan la respuesta funcional apropiada, además, las células han desarrollado sistemas sofisticados para integrar entradas de múltiples señales (Uings & Farrow, 2000).

Los receptores juegan roles importantes en células normales y en células cancerígenas; un proceso fisiológico fundamental tanto en el desarrollo como en la edad adulta es la angiogénesis, que requiere la acción coordinada de una variedad de factores de crecimiento y moléculas de adhesión celular en las células endoteliales y de revestimiento muscular, VEGF (Vascular endothelial growth factor)-A y sus receptores (VEGFR-1 y VEGFR-2) son la vía de señalización mejor caracterizada en la angiogénesis en desarrollo (Ferrara & Kerbel, 2005). En algunos casos, VEGFR-1 es expresado por las células tumorales y puede mediar una señal quimiotáctica, lo que potencialmente extiende el papel de este receptor en el crecimiento del cáncer, la expresión del gen VEGF-A se regula al alza por la hipoxia (Ferrara & Kerbel, 2005).

Los miembros de la familia de receptores transmembrana del factor de crecimiento epidérmico (familia ErbB) son potentes mediadores del crecimiento y desarrollo celular

normal, esta familia está compuesta por cuatro receptores tirosina quinasa tipo 1 transmembrana muy relacionados: el receptor del factor de crecimiento epidérmico (EGFR, también conocido como ERBB1), ERBB2, ERBB3 y ERBB4. La señalización a través de la familia ErbB sustenta muchas de las actividades celulares en las que dependen la supervivencia y la función celular, EGFR, ERBB2 y ERBB3 están implicados en el desarrollo y progresión del cáncer, y la heterodimerización de los receptores juega un papel crucial en su función, el papel de ERBB4 en la oncogénesis es menos claro y este receptor podría estar involucrado en la inhibición del crecimiento celular en lugar de la proliferación; la expresión o función aberrante de ERBB2 se ha relacionado con la evolución tanto del cáncer de mama como del cáncer gástrico y es evidente en otros tipos de cáncer, como los tumores ováricos y de glándulas salivales (Baselga & Swain, 2009).

2.9 Función de receptores como dianas terapéuticas.

A lo largo de los años, el tratamiento convencional del cáncer se ha mejorado enormemente con la adición de un número creciente de nuevos agentes molecularmente dirigidos, que inhiben las vías moleculares cruciales para el crecimiento, mantenimiento y metástasis del tumor, este tratamiento personalizado y dirigido contra el cáncer utiliza las características moleculares detalladas de un tumor y su microambiente para permitir terapias adaptadas que mejoren los resultados del tratamiento, reduzcan la toxicidad en los tejidos sanos y superen la resistencia a los fármacos, mejorando así el perfil beneficio/riesgo (Assaraf *et al.*, 2014).

Un receptor que se puede emplear como diana es el receptor folato (FR), en los tejidos y órganos normales, la expresión del FR está restringida a unos pocos sitios, que incluyen el riñón, el pulmón, el plexo coroideo y la placenta, donde el FR está confinado a la superficie luminal de las células epiteliales polarizadas y, por lo tanto, no está en contacto con los folatos circulantes o los conjugados de ácido fólico administrados por vía intravenosa. Numerosos estudios han demostrado que el FR está notablemente sobreexpresado en la superficie de varios tipos de tumores, incluyendo los cánceres de ovario, riñón, pulmón, cerebro, endometrio, colorrectal, pancreático, gástrico, prostático, testicular, vesical, cabeza y cuello, y de mama, así como en el cáncer de pulmón de células no pequeñas, con niveles de expresión aumentados en comparación con los tejidos normales correspondientes. Además, la evidencia sugiere que la expresión del FR aumenta con el avance del estadio de la enfermedad y que la sobreexpresión del FR es un factor pronóstico negativo para el cáncer de mama, colorrectal, ovario y endometrio (Assaraf *et al.*, 2014).

2.10 Receptores en MCF-7

En 1970, el investigador Herbert D. Soule en la Michigan Cancer Foundation intentó desarrollar una línea celular a partir de una excisión de un nódulo de la pared torácica y de un derrame pleural de la paciente Helen Marion. A pesar de que en ese momento la generación de cultivos continuos y estables de líneas celulares cancerosas presentaba dificultades técnicas, Soule utilizó un proceso estándar para desarrollar una línea celular. Pese a que los cultivos celulares derivados del nódulo de la pared torácica pronto fueron sobrecrecidos por fibroblastos y descartados, las células del derrame pleural crecieron inicialmente en suspensión y luego formaron finalmente una monocapa sobre plástico que creció como un cultivo continuo. La línea celular resultante se llamó MCF-7 en honor a la Michigan Cancer Foundation y representó el séptimo intento de Soule de generar una línea celular cancerosa (Lee *et al.*, 2015).

Se han reportado receptores expresados y/o sobreexpresados en la línea MCF-7, por ejemplo, los receptores de estrógeno (ER), estos son objetivos importantes para regular el crecimiento de las células de cáncer de mama, en los tejidos mamarios, ER α se asocia con la proliferación celular, mientras que se informa con frecuencia que el ER β desempeña un papel antiproliferativo. La presencia de ER β en los cánceres de mama se asocia con una mejor supervivencia del paciente clínicamente (Liu *et al.*, 2020). Otro receptor es el receptor folato (FR), es una proteína de glicosilfosfatidilinositol que se expresa en muchas células cancerosas humanas, el nivel de expresión de FR está relacionado con el estadio del cáncer, FR tiene una alta afinidad por el ácido fólico y sus derivados, y es el mediador en la entrega del 5-metiltetrahidrofolato como folato fisiológico a la célula. Los FR están sobreexpresados en numerosos tipos de células cancerosas humanas, incluyendo cáncer de mama, colorrectal, ovario, cerebro, epitelial y de pulmón, mientras que la expresión está limitada en células normales y, por lo tanto, podría ser utilizado como biomarcador para el cáncer (Soleymani *et al.*, 2018).

3.-JUSTIFICACIÓN

El hallazgo de inclusiones proteicas con actividad citotóxica y selectividad contra diversas líneas de células de cáncer, producidas por algunas de las cepas *B. thuringiensis*, aisladas por el grupo, las convierte en excelentes candidatos como agentes anticancerígenos, un ejemplo de estas son las proteínas A34-2 y AX-2. Si bien estas proteínas ya han sido secuenciadas, aun no se han realizado análisis a estas secuencias y por lo tanto no han sido clasificadas en alguna de las 6 familias conocidas de las PS. La importancia de conocer la secuencia y la estructura de estas proteínas está en que nos permitirá conocer y entender el mecanismo de acción que posiblemente se encuentre relacionado a su actividad antiproliferativa, así también, el análisis de las secuencias de las proteínas A34-2 y AX-2 permitirá realizar estudios *in silico* de las posibles moléculas con las que pudiera interactuar y así, desarrollar su actividad. Debido a la falta de información de la secuencia PS los estudios *in silico* son difíciles de diseñar, es este trabajo planteamos utilizar a las proteínas Cry como un modelo para diseñar una estrategia de un estudio *in silico* para utilizarlo como modelo referente a las parasporinas; como se sabe, Cry y PS tienen un alto grado de homología estructural y funcional, este trabajo tiene por objeto la generación de conocimiento sobre la hipótesis de que las parasporinas necesitan interacción con la membrana celular neoplásica mediante un receptor específico y que, de esta manera, logre los efectos citotóxicos observados en estas células.

4.- HIPÓTESIS

El modelado *in silico* empleando proteínas Cry como estrategia llevará a identificar posibles receptores diana de las parasporinas en cáncer de mama.

5.- OBJETIVOS.

5.1 Objetivo general:

Identificar *in silico* moléculas específicas de la célula cancerosa con la capacidad de interaccionar con las proteínas Cry

5.2 Objetivos particulares:

- Identificar proteínas Cry con similitudes funcionales y estructurales a las PS.
- Demostrar *in silico* la presencia de moléculas receptoras para Cry1Aa y Cry1Ac en la línea celular MCF-7.

6.- DIAGRAMA EXPERIMENTAL

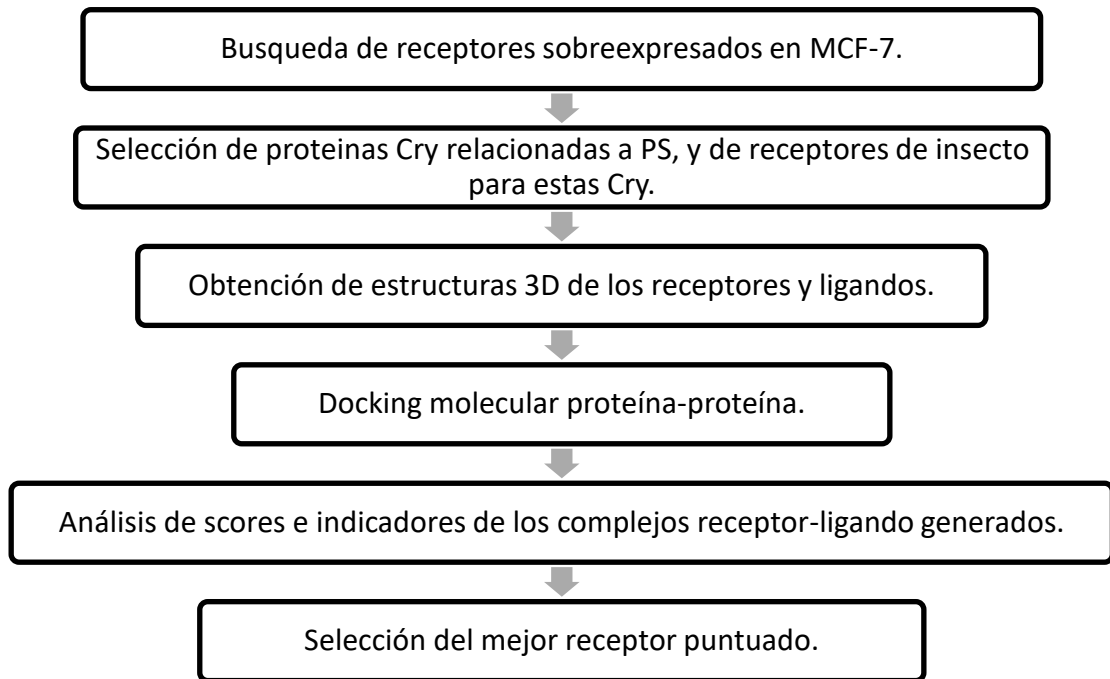


Figura 2. Diagrama experimental

7.- MATERIALES Y MÉTODOS

7.1 Búsqueda de receptores

Se realizó una búsqueda en literatura acerca de aquellos receptores que se sobreexpresaran en células de cáncer, en el grupo de trabajo, una de las líneas de cáncer de interés y con la que se ha estado trabajando es la MCF-7, línea celular de cáncer de mama, por lo tanto, los receptores elegidos fueron aquellos que estuvieran presentes en cáncer de mama, en la tabla 1, se muestran los receptores que se ocuparon en el presente trabajo.

Tabla 1. Receptores empleados en el trabajo y tipos de cáncer en los que se expresa.

RECEPTOR	CÁNCER
Folato α	Pulmón, colon, mama
Transferrina	Colon, glioma, mama
Factor de crecimiento epidérmico (EGFR)	Pulmón, ovario, mama, páncreas
Receptor de ácido hialurónico (CD44)	mama
Péptido intestinal vasoactivo (VIP)	Colon, mama, tumores endocrinos
Aminopeptidasa N	Colon

7.2 Modelos de proteínas.

Los receptores mencionados en la sección anterior fueron buscados en la base de datos Protein Data Bank (PDB), que es una fuente de datos global de estructuras de proteínas tridimensionales, la información complementaria y relevante a cada una de las estructuras se obtendrá de la base de datos The Universal Protein Resource (UniProt), así también se buscaron las estructuras e información de las proteínas Cry que se ocuparon como ligandos, es decir, Cry1Aa y Cry1Ac.

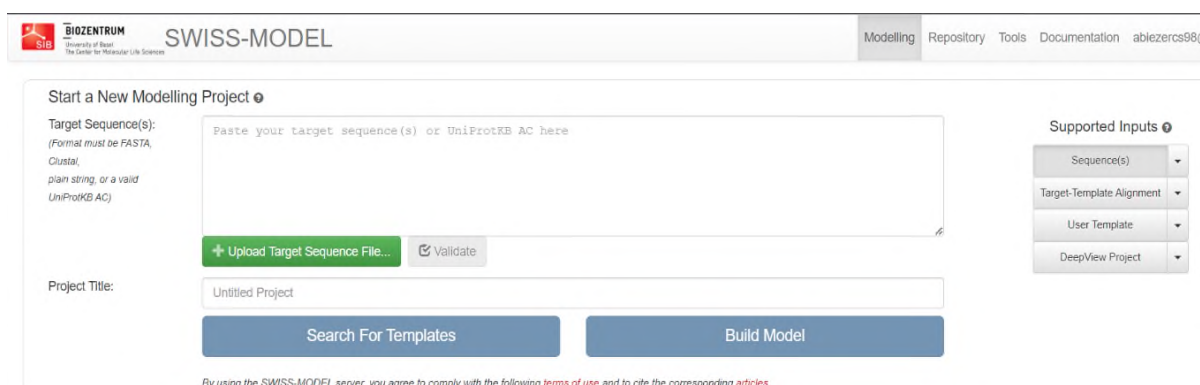
7.3 Modelado de receptores.

Los modelos de receptores de insectos no se encontraron alojados en el PDB, es decir, no están resueltas experimentalmente. La opción por la que se optó fue modelar dichas estructuras a partir de las secuencias proteicas, la información de las secuencias se encontró en UniProt; se eligieron 2 maneras para modelar los receptores faltantes: 1) mediante modelado por homología usando SWISS-MODEL, un servidor de modelado por homología de

estructuras desarrollado por el Computational Structural Biology group apoyado del Swiss Institute of Bioinformatics (SIB), 2) modelado de proteínas mediante el uso de Deep Learning usando la inteligencia artificial AlphaFold desarrollada por DeepMind, algunos detalles de éstas herramientas.

7.3.1 Modelado por homología

SWISS-MODEL fue la herramienta elegida para el modelado por homología, esto se realizó a través de su servidor en línea (Figura 3.) (<https://swissmodel.expasy.org/interactive>)



The screenshot shows the 'Start a New Modelling Project' page on the SWISS-MODEL website. At the top, there is a navigation bar with 'Modelling', 'Repository', 'Tools', 'Documentation', and 'ablezercs986'. The main content area is titled 'Start a New Modelling Project' and contains a form for entering target sequences. The form includes a text input field for 'Target Sequence(s):' with a placeholder 'Paste your target sequence(s) or UniProtKB AC here'. Below this field are two buttons: '+ Upload Target Sequence File...' and 'Validate'. To the right of the form is a 'Supported Inputs' section with a dropdown menu showing 'Sequence(s)', 'Target-Template Alignment', 'User Template', and 'DeepView Project'. Below the form is a 'Project Title:' field with the text 'Untitled Project' and two buttons: 'Search For Templates' and 'Build Model'. At the bottom of the page, there is a small disclaimer: 'By using the SWISS-MODEL server, you agree to comply with the following terms of use and to cite the corresponding articles.'

Figura 3. Interfaz Modelling del servidor en línea SWISS-MODEL.

El identificador de UniProt de cada proteína (receptores y ligandos) se introduce en el cuadro de target sequence y se nombra el proyecto, posteriormente se selecciona la búsqueda de plantados, los resultados de esta búsqueda se ordenan en función de un indicador del servidor llamado GMQE (Global Model Quality Estimate por sus siglas en ingles), este indicador puntúa del 0 a 1 la calidad de una predicción de un modelo, los plantados con mejores valores de GMQE se llevaron al paso siguiente el cual fue la construcción de modelos basados en los plantados de proteínas homologas. Los modelos resultantes son puntuados mediante el indicador QMEANDisCo que también evalúa los modelos resultantes del 0 a 1, esta función toma en cuenta criterios como la calidad de encaje entre la secuencia de aminoácidos y la estructura del plantado y la calidad de la estructura en términos de la conformación de residuos aminoacídicos. Los modelos con mejores valores de QMEANDisCo se seleccionaron para docking.

7.3.2 Modelado con Inteligencia Artificial (IA)

El modelado con inteligencia artificial se llevó a cabo con la herramienta desarrollada por DeepMind, AlphaFold, hay varias maneras de trabajar con AlphaFold, dos de ellas son mediante el uso de hardware de terceros en la nube a través de Google Colab

(<https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb>) y accediendo a la base de datos de la inteligencia artificial (<https://alphafold.ebi.ac.uk/>), para este trabajo, se optó por obtener los modelos de la base de datos directamente, los modelos alojados en esta (Figura 4.), fueron generados por el equipo de investigación del laboratorio de inteligencia artificial de DeepMind.

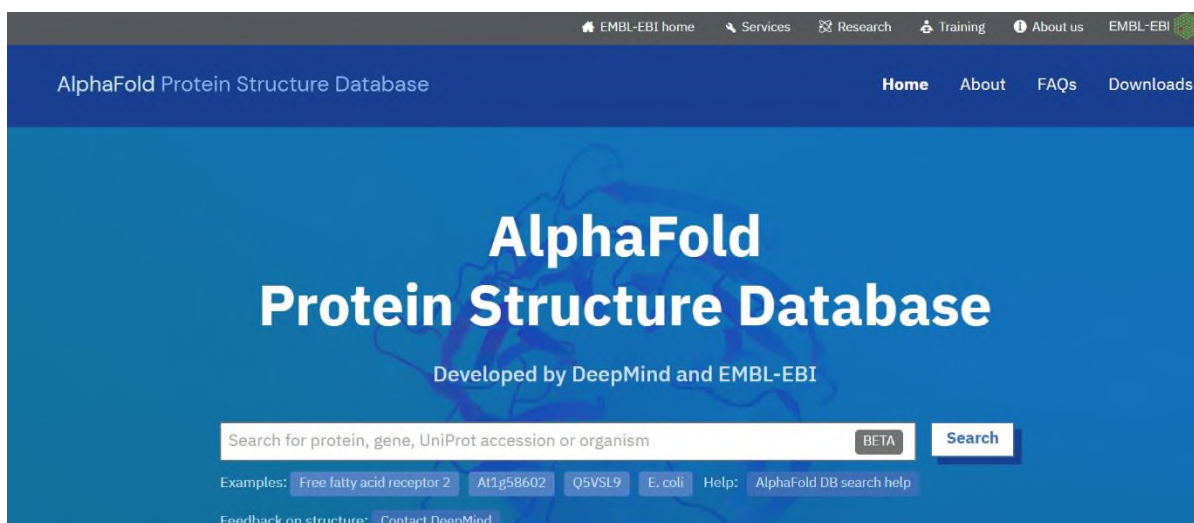


Figura 4. Interfaz de inicio de Alphafold Database.

7.4 Preparación de proteínas para docking

Las estructuras resueltas experimentalmente y provenientes de la base de datos PDB, suelen estar acompañadas de moléculas pequeñas como iones, ligandos, agua, entre otros, así también, algunos modelos pueden aparecer con múltiples copias de sí misma en un archivo debido a que técnicas de resolución de estructuras como la difracción de rayos x o técnicas de difracción en solución resuelven múltiples en la misma celda unitaria, que es el espacio más pequeño que puede contener la información necesaria para reproducir la estructura completa. Debido a lo antes mencionado, se utilizó la UCSF Chimera para eliminar las subunidades adicionales, ligandos, iones y moléculas de solvente que pueden interferir en el proceso de acoplamiento, cada estructura editada fue guardada en formato pdb.

7.5 Regiones de interés

Se realizó una búsqueda de regiones entre las proteínas que se ocuparon, tanto de los receptores como de los ligandos, las bases de datos InterPro y PROSITE fueron utilizadas para identificación y análisis de dominios de proteínas y patrones de secuencias, esto para conocer los dominios de unión a proteínas y visualizar coincidencias con los resultados del docking.

7.6 Docking Molecular

Para realizar las pruebas de acoplamiento o docking molecular, se utilizaron tres programas: ClusPro, HDock y HEX, los primeros dos cuentan con un servidor en línea el cual permite ejecutar el docking en estos mismos, el último es un software que se ejecutó de manera local. Aunque los resultados arrojados por cada una de estas herramientas no son completamente comparables entre sí, llegan a aportar una gama de resultados más variable.

7.6.1 ClusPro

ClusPro es una herramienta de docking proteína-proteína de libre acceso (<https://cluspro.bu.edu/login.php>), las pruebas se realizaron en el servidor en línea (Figura 5), se puede acceder mediante una cuenta con correo institucional o como invitado.

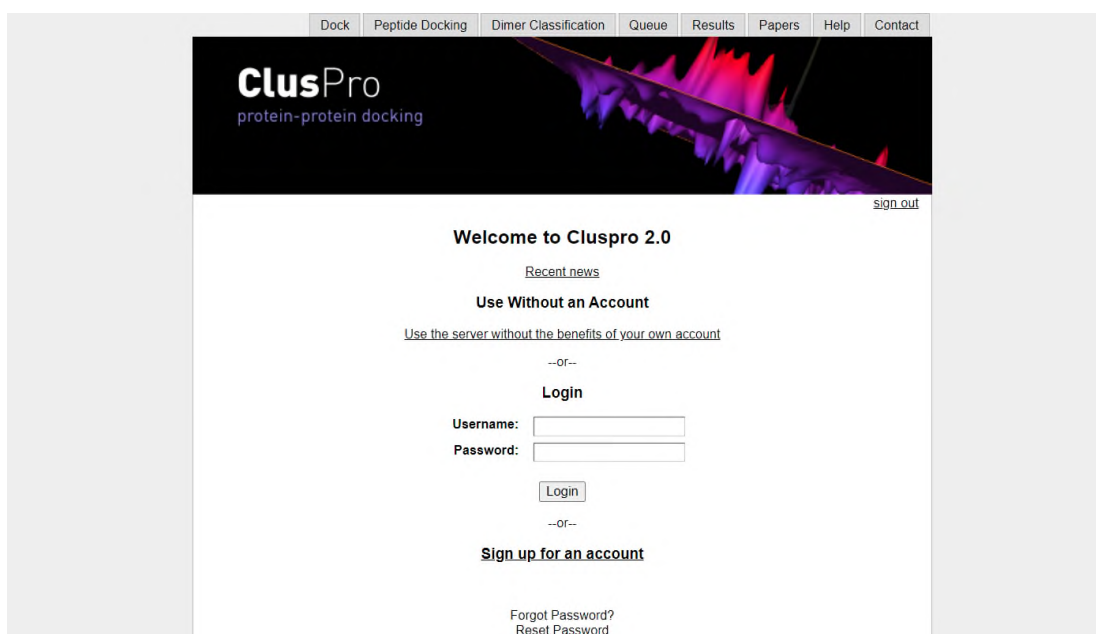


Figura 5. Interfaz principal del servidor ClusPro.

Una vez se accede al servidor, este permite nombrar el trabajo de la sesión, así como elegir si el trabajo se ejecuta en un procesador CPU o en GPU, se cargan los archivos del receptor y del ligando en formato pdb, después de esto se marca una casilla de acuerdo de uso, en la que se especifica que el servidor solo se ocupa para propósitos no comerciales y por último se presiona el botón Dock (Figura 6). El servidor carga toda la información, le asigna un identificador a la prueba y la manda a la cola de trabajo, una vez que concluye manda la prueba a la sección de resultados en donde se almacena temporalmente para poder descargar todos los archivos resultantes de la prueba.

Figura 6. Sección de carga de trabajo del servidor.

7.6.2 HDock

Este servidor permite realizar docking proteína-proteína y proteína-RNA/DNA, es de libre acceso a través de internet (Figura 7) (<http://hdock.phys.hust.edu.cn/>).

Figura 7. Interfaz principal del servidor HDock.

En la interfaz de la página se encuentran dos secciones para introducir los archivos del receptor y del ligando respectivamente, posterior a esto, en la parte inferior de la interfaz se agrega el nombre del trabajo y un correo electrónico para enviar los resultados una vez que estos estén listos, después de introducir estos datos se da click al botón submit para que la prueba se envíe a la cola.

7.6.3 HEX

HEX es un programa de libre acceso que se ejecuta localmente, está disponible su descarga en su sitio en internet (<http://hex.loria.fr/>).



Figura 8. HEX, software de docking molecular.

HEX utiliza una técnica exhaustiva para explorar todas las posibles conformaciones del receptor y el ligando, y a su vez, calcula la energía de interacción de cada conformación. La interfaz (Figura 8) de uso contiene una barra de herramientas en la sección superior, para cargar las moléculas se selecciona la pestaña File y se abren el archivo de receptor y ligando donde corresponde, posteriormente, en la pestaña Controls en la opción Dock, se ajustan los parámetros, que para este trabajo se dejaron por defecto (Figura 9). Una vez ejecutado el programa, arroja los resultados en una ventana de texto los cuales se pueden exportar a un formato de texto simple y los modelos de docking se pueden exportar al disco local en formato pdb.

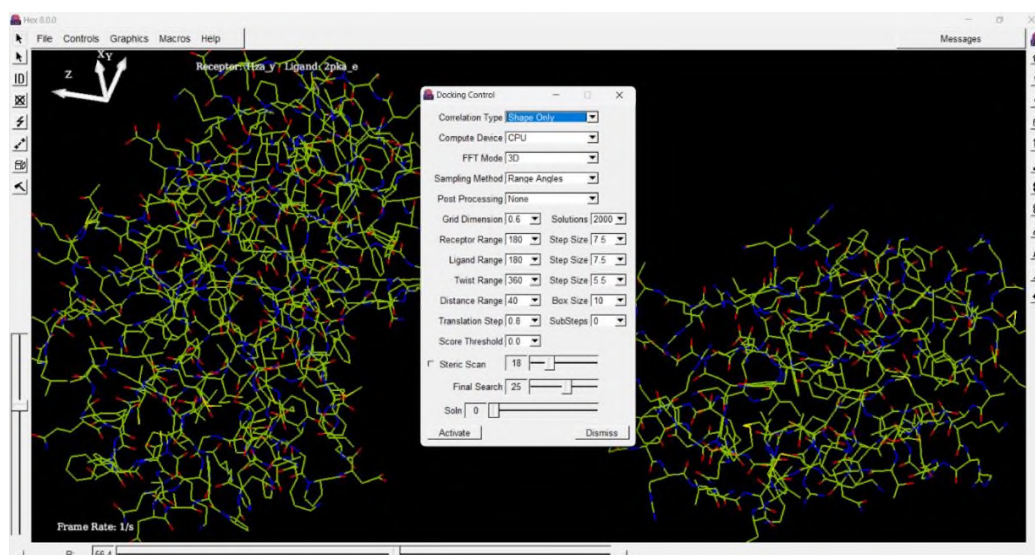


Figura 9. Pantalla de preparación de docking en HEX.

8.-RESULTADOS

En este trabajo se empleó el docking molecular para la evaluación de receptores sobreexpresados en la línea MCF-7 de cancer de mama y su interacción con las proteínas Cry (Cry1Aa y Cry1Ac), las cuales guardan homología de función, secuencia y estructura con las PS, por esto, se emplearon estas Cry como un sistema de referencia, ya que la información de parasporinas como ligandos es muy escasa.

8.1 Modelos Experimentales

La selección de los receptores usados en este trabajo se basó en una amplia revisión de la literatura científica, mediante el análisis de estudios previos, se identificaron una serie de receptores los cuales se sabe que están sobreexpresados en diversos tipos de cáncer como el de pulmón, cerebro, hígado, páncreas, próstata, ovarios, colon y mama.

A continuación, se muestra en la tabla 2, datos generales de los modelos proteicos (receptores humanos y proteínas cry) extraídos del Protein Data Bank.

Tabla 2. Información general de receptores humanos obtenida del Protein Data Bank.

Nombre	Identificador	PDB ID	UniProt ID	Longitud PDB (aa.)	Longitud UniProt (aa.)
Receptor Folato α	FOL	4KM6	P15328	208	257
Receptor Transferrina	TFRC	6OKD	P02786	670	760
Receptor EGFR	EGFR	4UV7	P00533	621	1210
Receptor CD44	CD44	1UUH	P16070	159	742
Receptor VIP	VIP	2X57	P41587	116	438
Receptor Aminopeptidasa N	AMPN	4FYQ	P15144	903	967
Proteína Cry1Aa	Cry1Aa	1CIY	P0A366	590	1176
Proteína Cry1Ac	Cry1Ac	4ARX	P05068	579	1178

PDB ID: Número de identificador en Protein Data Bank; UniProt ID: Numero de identificador en UniProt

Estos modelos experimentales de receptores y de las proteínas Cry, fueron los que posteriormente se utilizaron para realizar el docking con las diferentes herramientas, es decir, con el software local y con los servidores online.

8.2 Modelado de proteínas

Para comparar el acoplamiento de los receptores humanos con las proteínas Cry, fue necesario obtener los receptores de insectos ya reportados de las proteínas Cry de interés: el receptor APN de *Manduca sexta* (especie lepidóptera conocida como gusano del tabaco), receptor APN de *Plutella xylostella* (lepidóptera conocida como polilla dorso de diamante o polilla de la col) y receptor ALP de *Heliothis virescens* (lepidóptera conocido como gusano bellotero). Estos receptores solo cuentan con la secuencia de aminoácidos alojada en la base de datos de UniProt pero no así con una estructura tridimensional resuelta. Por este motivo se recurrió a realizar el modelado de los receptores de insectos, se emplearon 2 maneras distintas para obtener los modelos de estos receptores, el modelado por homología y el modelado con Inteligencia artificial, los resultados de estas metodologías se muestran a continuación.

8.2.1 Modelado por homología

Para este tipo de modelado se utilizó el servidor de modelado por homología de estructuras proteicas SWISS-MODEL, las proteínas a las que se aplicó esta metodología fueron 3 receptores de insectos, estos han sido reportados como receptores de unión de las proteínas Cry1Aa y Cry1Ac, en la tabla 3, se muestran datos generales de estos receptores.

Tabla 3. Información general de los receptores de insecto obtenidos de UniProt.

Receptor	Organismo	UniProt ID	Masa (kDa)	Longitud (a.a.)
Aminopeptidasa N	<i>Manduca sexta</i>	P91885	106.78	942
Aminopeptidasa N	<i>Plutella xylostella</i>	P91887	106.58	946
Fosfatasa alcalina	<i>Heliothis virescens</i>	C3U1W2	59.2	539

UniProt ID: Numero de identificador en UniProt

Dentro del servidor SWISS-MODEL, el cuadro de “Target sequence” se llenó con el dato de UniProt ID de los receptores Aminopeptidasa N (APN) y Fosfatasa alcalina (ALP) para posteriormente realizar la búsqueda de templados para armar el modelo tridimensional.

Para cada receptor se utilizó solo 1 templado dado que el porcentaje de identidad es muy bajo, por lo tanto, se obtuvo 1 modelo por homología de cada receptor de insecto (tabla 4).

Dentro de los indicadores a tener en consideración está GMQE, que es una métrica de SWISS-MODEL que sirve como estimación de la calidad del modelo a generar, combina propiedades entre el alineamiento del templado-objetivo y de la estructura del templado, el valor de este indicador va de 0 a 1, siendo 1 el valor más alto (Biasini *et al.*, 2014), este valor se encuentra disponible antes de construir el modelo, por lo que es útil para seleccionar el templado adecuado, sin embargo, para el caso de los receptores de interés solo se disponía de un templado, por lo que no hubo más opción que usar el templado disponible a pesar de los valores bajos de identidad de secuencia y de GMQE. Otra de las métricas para evaluar la calidad en SWISS-MODEL es QMEANDisCo, que evalúa el ajuste de la estructura del modelo generado frente a estructuras de referencia utilizando criterios como superposición estructural, geometría de ángulos de torsión y distancias interatómicas, mientras más cercano sea este valor a 1, mayor será la similitud y la calidad estructural del modelo respecto a las estructuras de referencia (Studer *et al.*, 2020).

Tabla 4. Datos del templado para cada receptor.

Receptor	Templado				
	GMQE	PDB ID	Organismo	Identidad (%)	Cobertura (%)
APN1 (M. sexta)	0.69	4WZ9	<i>Anopheles gambiae</i>	32	93
APN2 (P. xylostella)	0.68	4WZ9	<i>Anopheles gambiae</i>	31.6	92
ALP (H. virescens)	0.69	3MK1	<i>Homo sapiens</i>	43.5	87

GMQE: (Global Model Quality Estimation por sus siglas en inglés) es una métrica de SWISS-MODEL que sirve como estimación de la calidad del modelo a generar, toma valores de 0 a 1, donde 1 es el valor más alto.

Los modelos de receptores de insecto obtenidos por SWISS-MODEL se muestran a continuación en la figura 10:

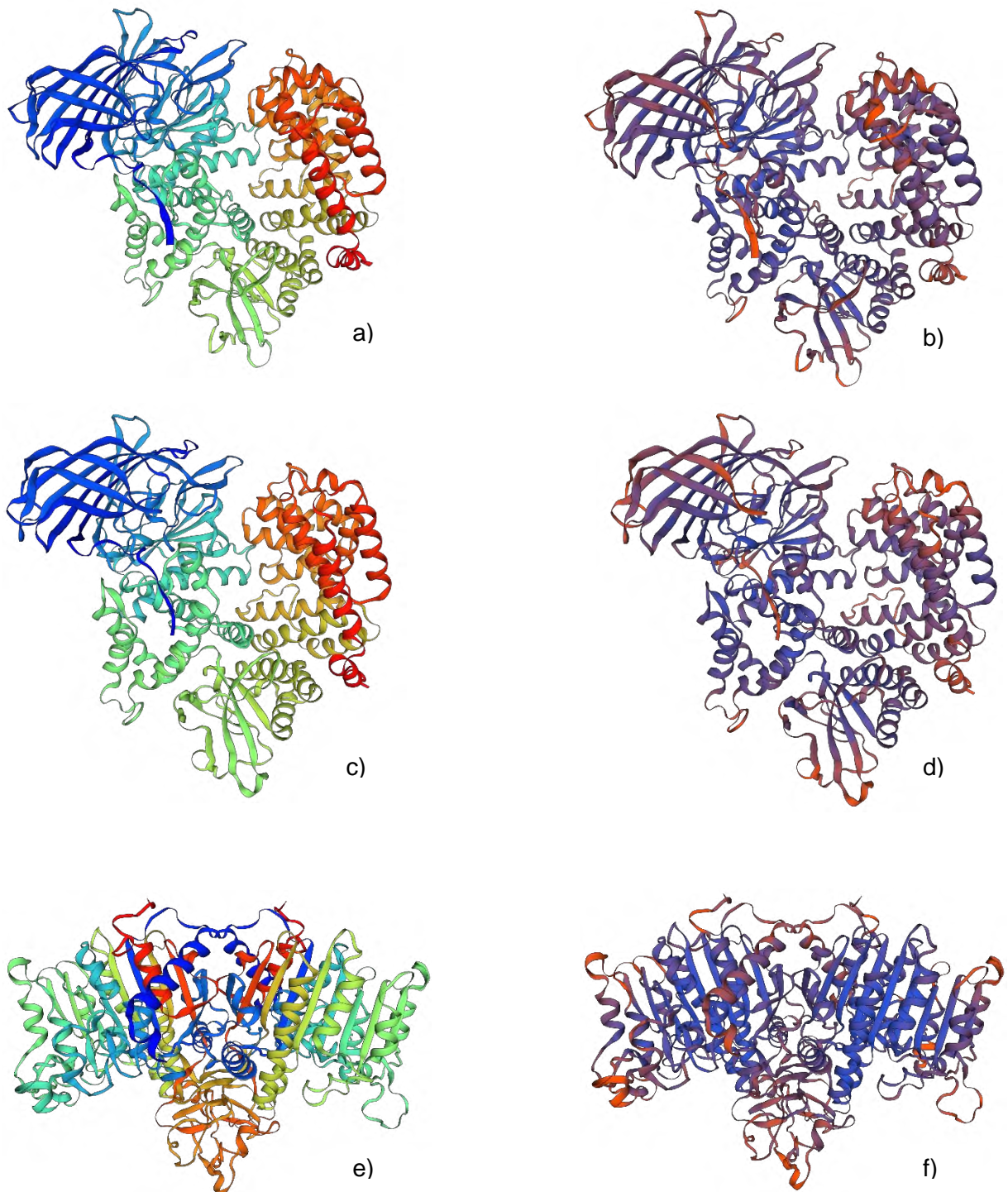


Figura 10. Modelos generados por homología mediante SWISS-MODEL de los receptores APN1 (a,b), APN2 (c,d) y ALP (e,f), paleta de color rainbow (a, c y d) y la paleta de colores de QMEANDisCo (b, d y f), donde las regiones en azul indican las regiones con un mejor valor de QMEANDisCo y las rojas regiones con valores bajos.

Los gráficos de Ramachandran muestran en que regiones se encuentran los ángulos diedros de los modelos generados, las regiones favorables están remarcadas de color verde, desde verde ligero para regiones generosamente permitidas, a verde más intenso para las regiones más favorables (figura 11).

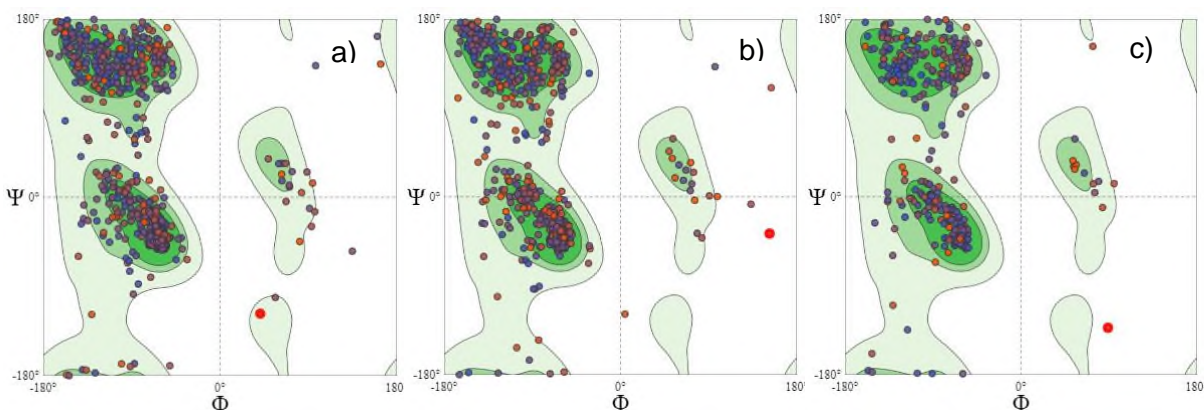


Figura 11. Gráficos de Ramachandran de los modelos APN1 (a), APN2 (b) y ALP (c), generados con SWISS-MODEL.

El modelado por homología de estos 3 receptores de insectos arrojó resultados deficientes (tabla 5), el valor QMEANDisCo global es un promedio de las puntuaciones de cada residuo que conforma a la proteína, en este caso a cada receptor, este puntaje por residuo evalúa la correlación del aminoácido del modelo con el del templado, siendo 1 el valor más alto de correlación; para el caso de los 3 receptores el valor de QMEANDisCo fue similar pero el valor en sí es bajo, por lo tanto, la calidad del modelado de aminoácidos es pobre, sin embargo, para los pasos posteriores se ocuparon estos modelos generados por homología a pesar de su baja calidad ya que no se contaban con mejores templados y para comparar el desempeño en el docking con el otro tipo de modelado.

Tabla 5. Score de diferentes indicadores estructurales para los modelos por homología.

Receptor	QMEANDisCo global	MolProbity Score	Ángulos en regiones favorables (%)
APN1 (<i>M. sexta</i>)	0.71 ± 0.05	1.80 (85%)	92.21
APN2 (<i>P. xylostella</i>)	0.70 ± 0.05	1.70 (89%)	92.70
ALP (<i>H. virescens</i>)	0.72 ± 0.05	1.92 (80%)	93.96

QMEANDisCo: Restricciones de distancia aplicadas al análisis de energía del modelo cualitativo (Distance Constraints applied on Qualitative Model Energy Analysis en inglés); MolProbity: Es un servicio web de validación de estructuras que proporciona una evaluación de amplio espectro y con base sólida de la calidad del modelo tanto a nivel global como local para proteínas y ácidos nucleicos.

Por último, respecto a los modelos de receptores de insectos en general, se cuenta con una métrica en común para comparar la calidad de la estrategia de modelado, MolProbity Score, el valor de este es una combinación ponderada en escala logarítmica del clashscore (que evalúa colisiones estéricas entre átomos adyacentes a la estructura), porcentaje de regiones Ramachandran no favorecidas y porcentaje de malos rotámeros en cadenas laterales, el valor resultante refleja la resolución cristalográfica a la que se esperarían esos parámetros, una estructura con un puntaje MolProbity numéricamente más bajo que su resolución cristalográfica real es, en términos de calidad, mejor que la estructura promedio en esa resolución, el valor de porcentaje que acompaña al puntaje MolProbity indica en qué medida la estructura evaluada se ajusta en comparación con las estructuras de referencia en la base de datos, por ejemplo, un valor de porcentaje del 90% significa que la estructura evaluada tiene una calidad igual o mejor que el 90% de las estructuras de referencia en la base de datos utilizada (Chen *et al.*, 2010).

8.2.2 Modelado con IA

El segundo método de modelado empleado para obtener la estructura de los 3 receptores de insecto se realizó con inteligencia artificial, mediante la herramienta Alphafold desarrollado por DeepMind, esta herramienta cuenta con una base de datos (AlphafoldDB) que almacena predicciones llevadas a cabo por el equipo de desarrollo, también se encuentra disponible para su uso en sistema Linux y mediante notebooks (o libretas) en Colab de Google Research que permite el uso de lenguaje Python en el navegador así como emplear hardware de mayor potencia para desarrollar estas tareas. Para el caso de los 3 receptores se consiguieron a través de la base de datos de AlphaFold, estos modelos fueron obtenidos por un grupo de investigación de DeepMind quienes prueban el rendimiento de la IA y actualizan los modelos de una proteína cuando se obtiene un mejor resultado (Varadi *et al.*, 2021), estos 3 modelos se muestran en la figura 12.

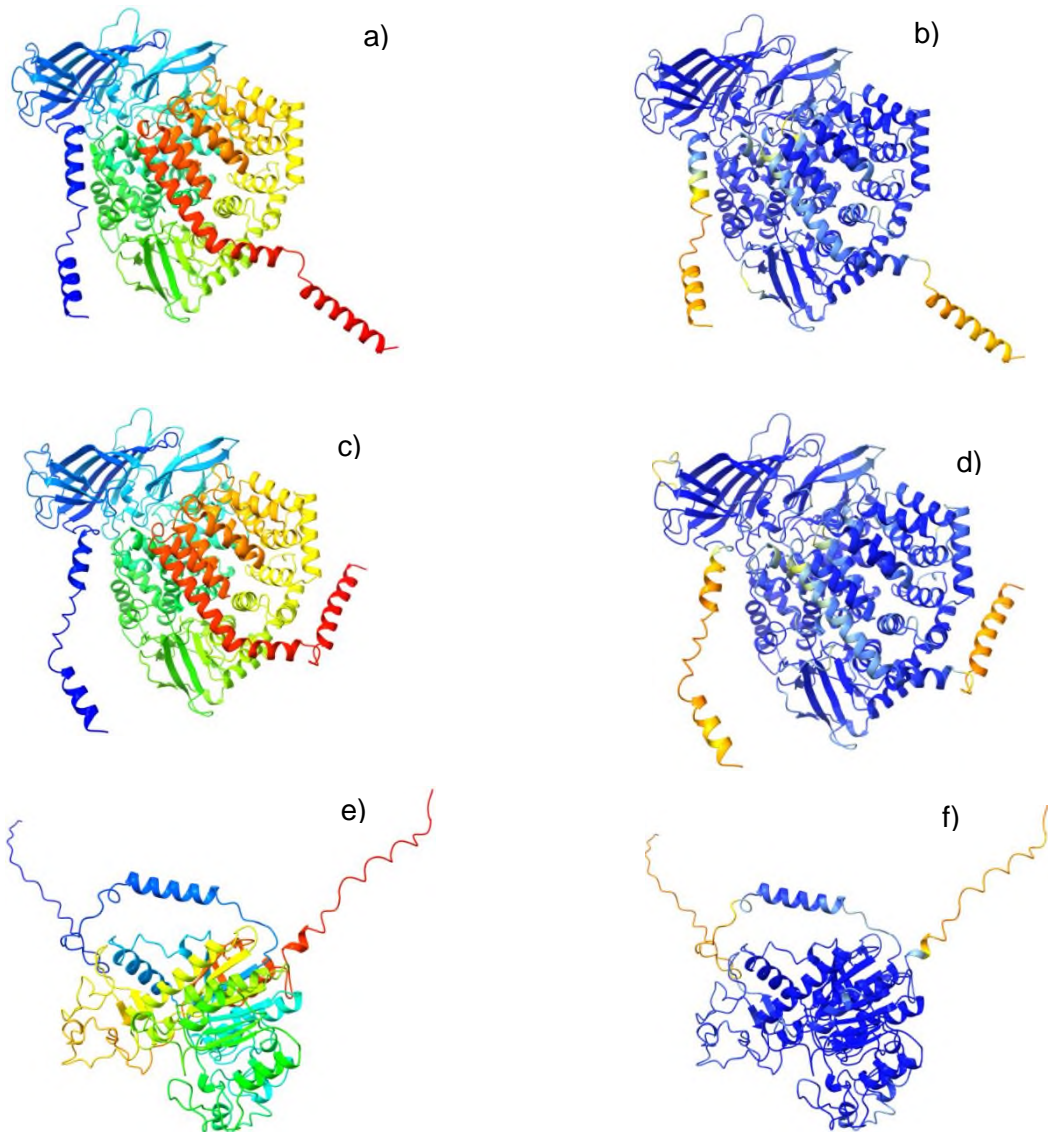


Figura 12. Modelos generados por Inteligencia Artificial mediante AlphaFold de los receptores APN1 (a,b), APN2 (c,d) y ALP (e,f), paleta de color rainbow (a, c y d) y paleta de color pLLDDT (b, d y f).

*pLLDDT: per-residue local distance difference test por sus siglas en inglés.

Los modelos de AlphaFold al ser visualizados en la base de datos y en la salida del notebook en Colab, presentan una paleta de colores de rojo a azul, esto se relaciona a un valor de puntuación llamado pLDDT (per-residue local distance difference test por sus siglas en inglés) que califica de 0-100 la confianza de la predicción de cada residuo que constituye la proteína o péptido modelado, así también, se genera una puntuación global de este valor, que es un promedio de los valores por residuo, esta puntuación global nos dice la precisión del modelado generado por IA, valores por encima de 90 tienen una precisión muy alta y son modelos confiables.

El error de alineado predicho o PAE (Predicted aligned error, por sus siglas en ingles), indica el error de distancia esperado en Angstroms de un residuo en el eje X cuando se alinea a un residuo en el eje Y, el color verde oscuro indica un error bajo mientras que el verde claro un valor de error alto, por ejemplo, tomamos el residuo al 200 del eje Y para alinearlo con el residuo 100 del eje X, si la zona en la gráfica es de color verde oscuro podemos confiar en la posición relativa del residuo 100, y si fuese verde claro no podríamos estar seguros de su posición relativa, en la figura 13 se muestran los gráficos PAE (de error alineado predicho)de los modelos con IA.

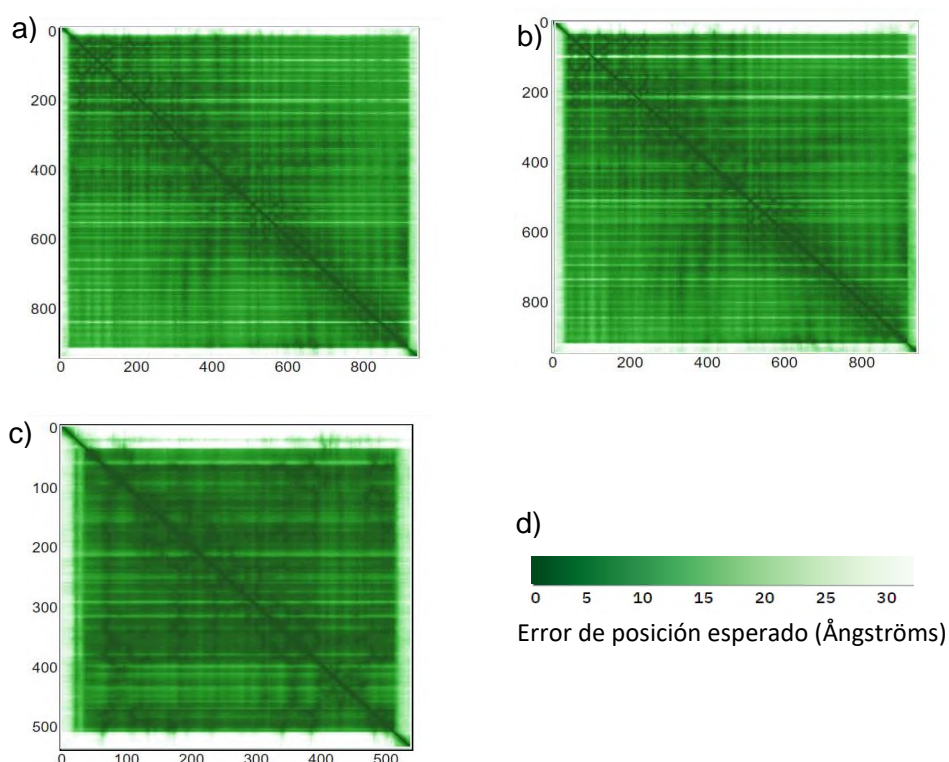


Figura 13. Gráficos PAE de los modelos APN1 (a), APN2 (b) y ALP (c), generados con AlphaFold y barra de color/error de referencia (d).

Los valores de pLDDT global para los modelos de estos 3 receptores están por encima de 90 (tabla 6), valor que alphafold maneja como muy confiable, si bien hay regiones en las que por residuo el modelado es deficiente o poco fiable, de manera global se puede decir que estos modelos son muy buenos para emplearlos en los siguientes pasos. En la figura 14 se muestran los gráficos de Ramachandran de los modelos de IA, los cuales, presentan un porcentaje de ángulos favorables por encima del 90%.

Tabla 6. Score de diferentes indicadores estructurales para los modelos por IA.

Modelo	pLDDT Global	MolProbity Score	Ángulos en regiones favorables (%)
APN1	93.32	0.84 (100%)	97.34
APN2	92.49	1.02 (100%)	96.93
ALP	91.10	1.61 (92%)	92.55

pLDDT: per-residue local distance difference test por sus siglas en inglés.

Molprobity Score: evalúa aspectos como choques entre átomos, orientaciones de la cadena lateral y geometría de la cadena principal, el puntaje refleja la resolución cristalográfica con esos parámetros (valores menores a 1.8 son deseables).

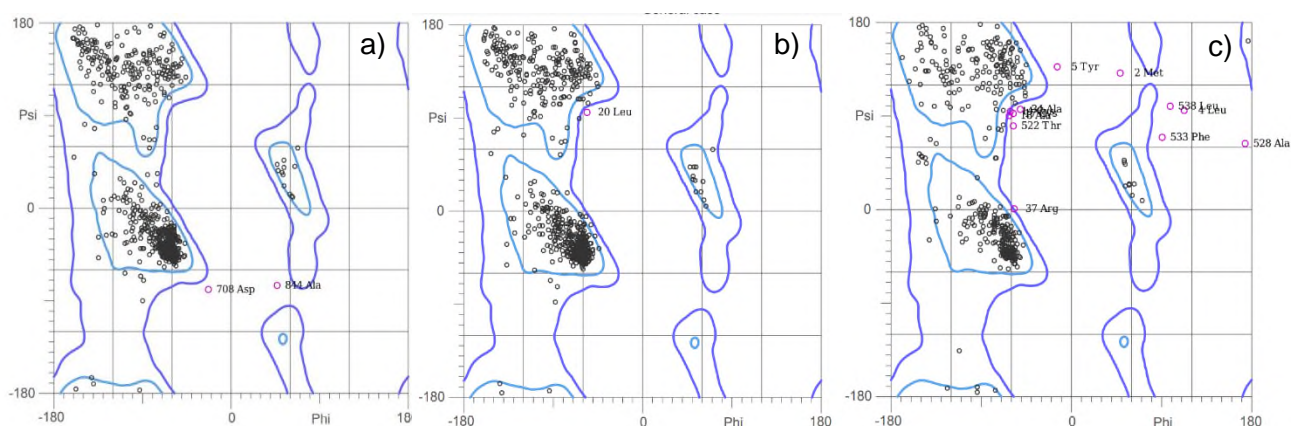


Figura 14. Gráficos de Ramachandran de los modelos APN1 (a), APN2 (b) y ALP (c), generados con AlphaFold.

8.3 Docking Molecular

Se usaron 3 herramientas para realizar el docking molecular de las proteínas, 2 de estas a través de servidores en línea y el tercero fue un software local, estos son: HDOCK, ClusPro 2.0 y HEX 8.0 respectivamente, en la siguiente sección se mostrarán los mejores acoplamientos de las proteínas Cry con los receptores de insectos y posteriormente con los receptores humanos.

8.3.1 HDOCK

HDOCK usa 2 indicadores para calificar la calidad del modelo usado para el docking, tanto para el receptor como para el ligando (tabla 7), MaxSub Y LGScore, MaxSub es una medida de la superposición entre los átomos C α de una estructura de proteína predicha y los átomos C α de una estructura determinada experimentalmente (Siew *et al.*, 2000) y LGScore es un puntaje

estadístico que se calcula a partir de la alineación de una estructura de proteína predicha con una estructura determinada experimentalmente (Cristobal *et al.*, 2001). Para el caso de los receptores humanos, todas las estructuras presentaron valores de MaxSub por encima del límite inferior de “correcto”, y en cuanto a valores de LGScore, solo el receptor VIP se encontró por debajo de “correcto” (tabla 8).

ProQ es el programa encargado de realizar esta la evaluación de los modelos, existen 3 niveles para clasificar la calidad de los modelos. En la tabla 8 se muestran los valores de cada indicador para los modelos experimentales, y en la tabla 9 los valores para cada tipo de modelado.

Tabla 7. Valores de referencia de los indicadores LGscore y MaxSub.

CORRECTO	BUENO	MUY BUENO
$1.5 < \text{LGscore}$	$3.0 \leq \text{LGscore} < 5.0$	$5.0 \leq \text{LGscore}$
$0.1 < \text{MaxSub}$	$0.5 \leq \text{MaxSub} < 0.8$	$0.8 \leq \text{MaxSub}$

LGscore: Evalúa la calidad global del modelo; MaxSub: Mide la similitud local entre el modelo y la estructura nativa de referencia. El rango de valores deseables es de 3.0 - 5.0 para LGScore y de 0.5 – 0.8 para MaxSub.

Tabla 8. Score de los indicadores LGscore y MaxSub para los modelos experimentales.

MODELO EXPERIMENTAL	INDICADOR	
	LGscore	MaxSub
Cry1Aa	5.509	0.363
Cry1Ac	5.887	0.374
Receptor Folato α	2.072	0.212
Receptor Transferrina	6.497	0.538
Receptor EGFR	3.939	0.283
Receptor CD44	2.526	0.295
Receptor VIP	1.142	0.120
Receptor N-Aminopeptidasa	5.795	0.489

Tabla 9. Score de los indicadores LGscore y MaxSub para las proteínas modeladas por homología y con IA.

PROTEÍNA MODELADA	AlphaFold		SWISS-MODEL	
	LGscore	MaxSub	LGscore	MaxSub
APN1	5.498	0.429	5.027	0.409
APN2	5.092	0.429	4.529	0.376
ALP	3.095	0.198	4.115	0.296

8.3.1.1 MODELOS ALPHAFOLD

A continuación, se muestran los 5 mejores scores de docking para los 3 mejores complejos de docking, ordenados en ranking de 1 a 5, donde 1 es el que obtuvo la mejor puntuación, en la tabla 10 se muestra los mejores complejos Cry1Aa-Receptor modelado; el RMSD es la desviación cuadrática media por sus siglas en inglés y el score de confianza se debe aproximar a 1. De igual manera, en la tabla 11 se muestran los mejores complejos Cry1Ac-Receptor modelado con los 5 mejores scores de las poses. El mejor complejo es el Cry1Ac-ALP por su bajo score de docking, con RMSD más bajo, comparado al resto de complejos (tabla 10 y 11).

Tabla 10. Scores de docking de Cry1Aa vs. Receptores de insectos generados por IA

Dock	Rank	1	2	3	4	5
Cry1Aa-APN1	Docking score	-309.30	-299.22	-294.44	-293.19	-290.68
	Score de confianza	0.9603	0.9519	0.9473	0.9460	0.9434
	RMSD	130.73	166.28	169.45	139.26	165.81
Cry1Aa-APN2	Docking score	-315.26	-298.97	-273.69	-269.09	-268.75
	Score de confianza	0.9646	0.9516	0.9223	0.9154	0.9149
	RMSD	169.17	170.97	176.19	166.49	138.19
Cry1Aa-ALP	Docking score	-309.60	-308.17	-284.05	-282.09	-278.57
	Score de confianza	0.9605	0.9594	0.9359	0.9335	0.9290
	RMSD	182.17	183.82	141.02	182.99	171.55

Tabla 11. Scores de docking de Cry1Ac vs. Receptores de insectos generados por IA

Dock	Rank	1	2	3	4	5
Cry1Ac-APN1	Docking score	-311.61	-302.21	-301.51	-296.92	-288.62
	Score de confianza	0.9620	0.9545	0.9539	0.9497	0.9412
	RMSD	63.02	70.73	66.59	62.60	58.88
Cry1Ac-APN2	Docking score	-285.80	-284.04	-283.18	-282.38	-281.65
	Score de confianza	0.9380	0.9359	0.9348	0.9339	0.9330
	RMSD	70.25	51.42	60.82	66.12	62.76
Cry1Ac-ALP	Docking score	-346.35	-331.79	-331.40	-316.03	-292.88
	Score de confianza	0.9807	0.9743	0.9741	0.9651	0.9457
	RMSD	51.77	62.23	57.59	59.71	43.20

En la figura 15, se muestran los 3 mejores modelos de complejos de docking, los cuales tienen el score más alto (tabla 10 y tabla 11), cada inciso está nombrado con abreviaturas del ligando (en amarillo y rojo) seguido del receptor (en color azul).

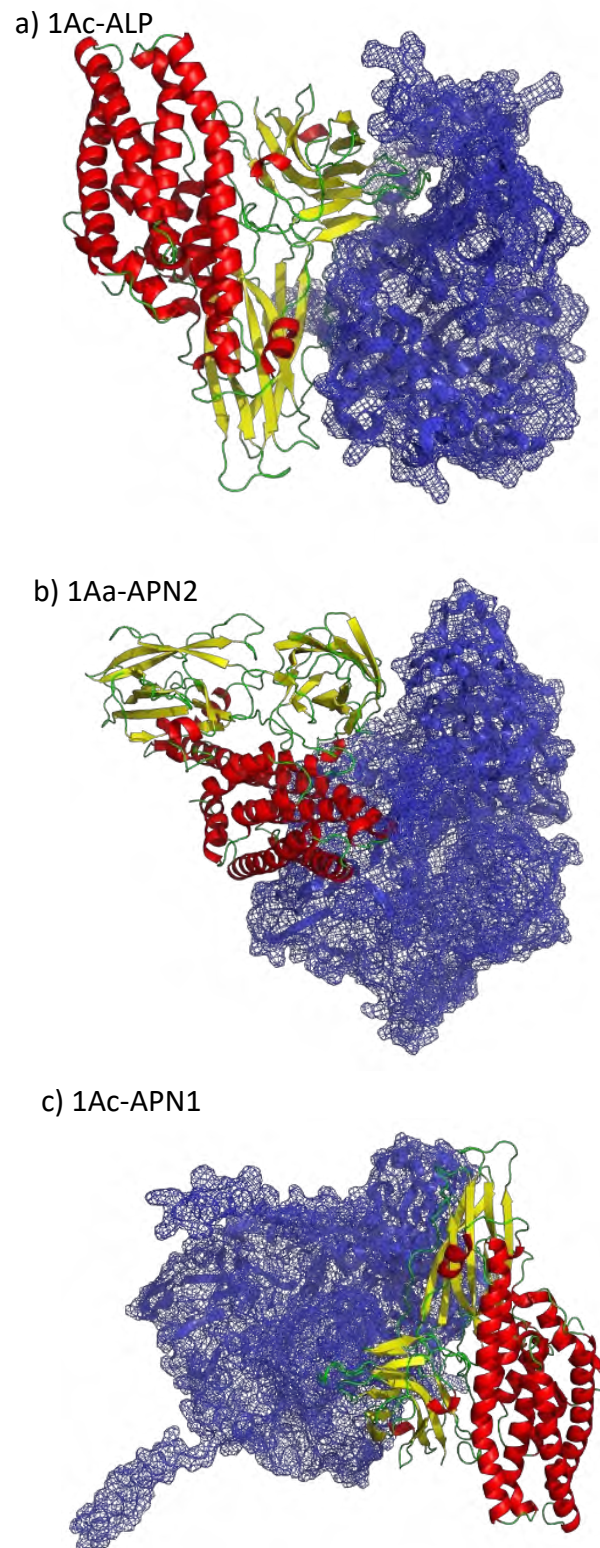


Figura 15. Top 3 complejos docking Cry-Receptores de insecto (AlphaFold) arrojados por HDOCK.

8.3.1.2 MODELOS SWISSMODEL

La tabla 12, muestra los valores de score de los mejores complejos Cry1Aa-Receptor modelado por homología, en el cual el mejor complejo se forma con el receptor APN1, aunque también el formado con ALP muestra un valor menor, destacando también en este un valor de RMSD más bajo.

Tabla 12. Scores de docking de Cry1Aa vs. Receptores de insectos generados por homología

Dock	Rank	1	2	3	4	5
Cry1Aa- APN1	Docking score	-291.41	-289.98	-281.04	-276.90	-274.95
	Score de confianza	0.9442	0.9427	0.9322	0.9268	0.9241
	RMSD	151.79	169.76	140.37	140.48	166.94
Cry1Aa- APN2	Docking score	-272.27	-269.54	-267.58	-261.27	-260.90
	Score de confianza	0.9202	0.9161	0.9131	0.9025	0.9019
	RMSD	179.64	160.17	164.47	124.95	162.50
Cry1Aa- ALP	Docking score	-276.41	-268.40	-258.54	-257.47	-257.16
	Score de confianza	0.9261	0.9144	0.8976	0.8956	0.8950
	RMSD	64.83	67.48	88.85	104.46	70.17

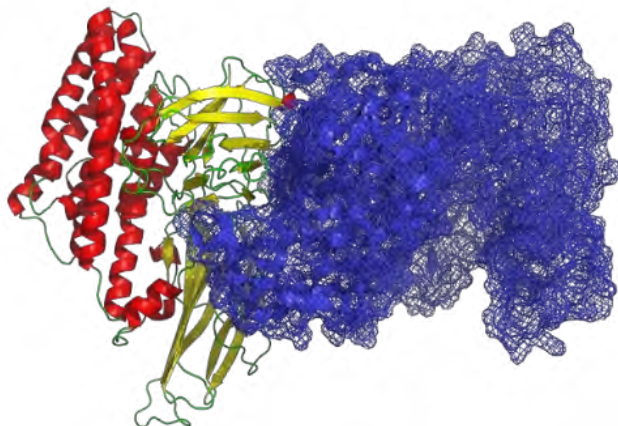
En la tabla 13, se encuentran los scores de los mejores complejos con la Cry1Ac-Receptor modelado por homología, siendo el complejo unido a APN1 el que tiene un mejor score (de -312.79), se puede destacar que con los modelos por homología el mejor receptor fue APN1.

Tabla 13. Scores de docking de Cry1Ac vs. Receptores de insectos generados por homología

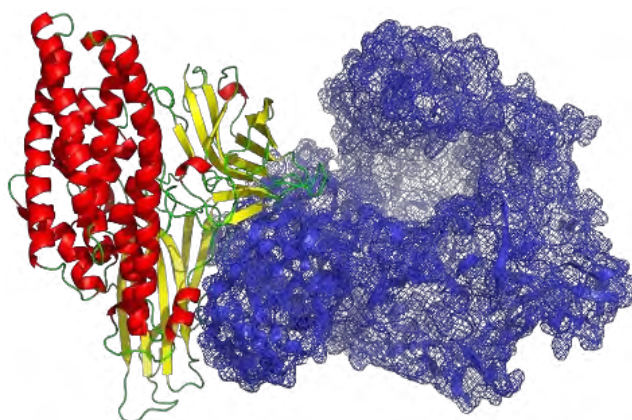
Dock	Rank	1	2	3	4	5
Cry1Ac- APN1	Docking score	-312.79	-299.84	-292.85	-292.52	-284.78
	Score de confianza	0.9629	0.9524	0.9457	0.9453	0.9368
	RMSD	53.83	52.33	72.52	39.39	73.11
Cry1Ac- APN2	Docking score	-296.22	-289.29	-283.25	-278.09	-273.55
	Score de confianza	0.9490	0.9419	0.9349	0.9284	0.9221
	RMSD	83.12	59.26	70.15	53.32	46.95
Cry1Ac- ALP	Docking score	-296.20	-294.53	-291.92	-289.58	-286.02
	Score de confianza	0.9490	0.9474	0.9447	0.9422	0.9382
	RMSD	50.88	82.93	77.46	51.78	90.56

En la figura 16 se muestran los complejos de docking mejor puntuados a partir de los datos de la tabla 12 y la tabla 13.

a) 1Ac-APN1



b) 1Ac-APN2



c) 1Ac-ALP

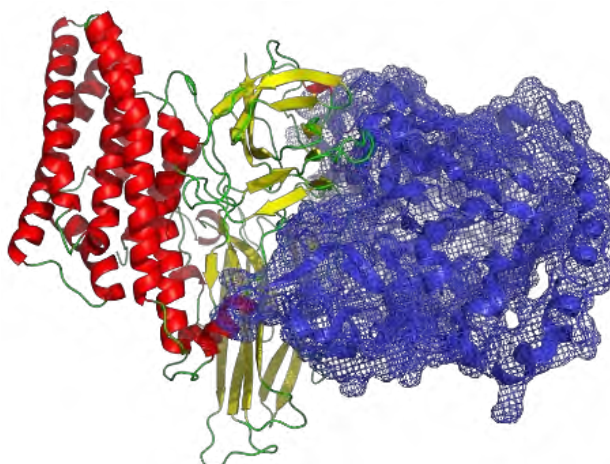


Figura 16. Top 3 complejos docking Cry-Receptores de insecto (SWISS-MODEL) arrojados por HDock.

8.3.1.3 DOCKING RECEPTORES HUMANOS

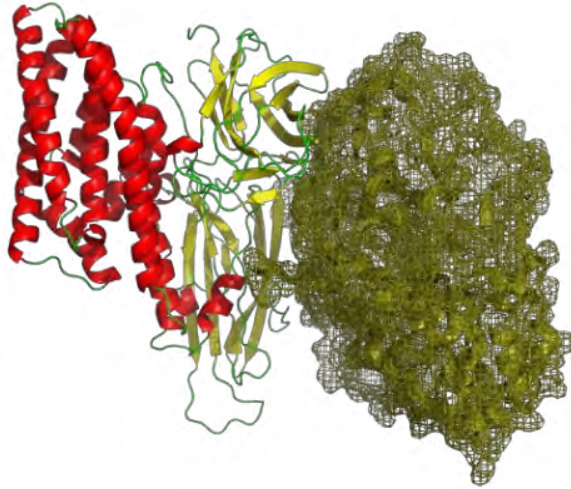
Para el docking con la Cry1Aa y receptores humanos obtenidos experimentalmente se obtuvieron los scores mostrados en la tabla 14, ordenados del 1 a 5, donde 1 es el de mejor score, el receptor TFRC obtuvo el mejor score, así como un score de confianza de 0.9839 siendo el más alto también, por otro lado, el receptor VIPR fue el peor.

Tabla 14. Scores de docking de Cry1Aa vs. Receptores humanos

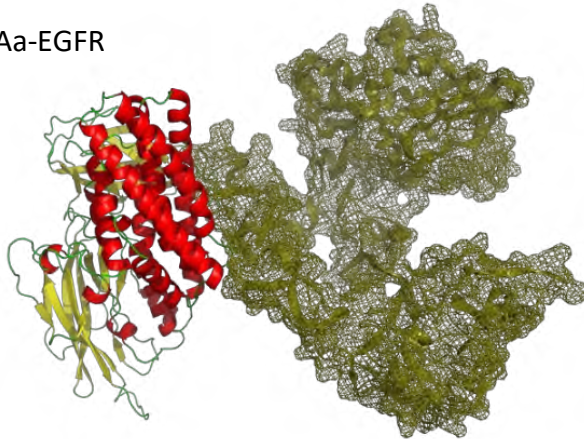
Dock	Rank	1	2	3	4	5
Cry1Aa-FOL	Docking score	-278.77	-267.91	-262.91	-257.49	-255.50
	Score de confianza	0.9293	0.9136	0.9054	0.8957	0.8919
	RMSD	53.44	93.85	99.61	80.90	98.77
Cry1Aa-TRFC	Docking score	-355.56	-312.92	-307.97	-298.05	-285.09
	Score de confianza	0.9839	0.9630	0.9593	0.9508	0.9371
	RMSD	149.42	153.51	149.42	160.89	153.54
Cry1Aa-EGFR	Docking score	-310.06	-306.82	-290.43	-280.52	-276.68
	Score de confianza	0.9609	0.9584	0.9431	0.9315	0.9265
	RMSD	155.34	171.82	154.50	139.88	132.33
Cry1Aa-CD44	Docking score	-277.26	-271.53	-258.34	-251.63	-249.78
	Score de confianza	0.9273	0.9191	0.8972	0.8842	0.8803
	RMSD	138.92	134.20	100.14	122.14	132.10
Cry1Aa-VIPR	Docking score	-277.18	-265.57	-256.88	-256.47	-238.25
	Score de confianza	0.9271	0.9098	0.8945	0.8937	0.8538
	RMSD	158.56	146.47	164.72	152.12	164.68
Cry1Aa-AMPN	Docking score	-282.16	-273.05	-272.90	-265.35	-264.00
	Score de confianza	0.9336	0.9214	0.9211	0.9095	0.9072
	RMSD	161.59	150.54	122.37	153.42	133.55

La figura 17 muestra los 3 mejores complejos de docking entre la Cry1Aa y receptores humanos, estos últimos se muestran en una tonalidad de verde para diferenciar con los receptores de insecto mostrados en color azul anterior y posterior a esta figura.

a) 1Aa-TFRC



b) 1Aa-EGFR



c) 1Aa-AMPN

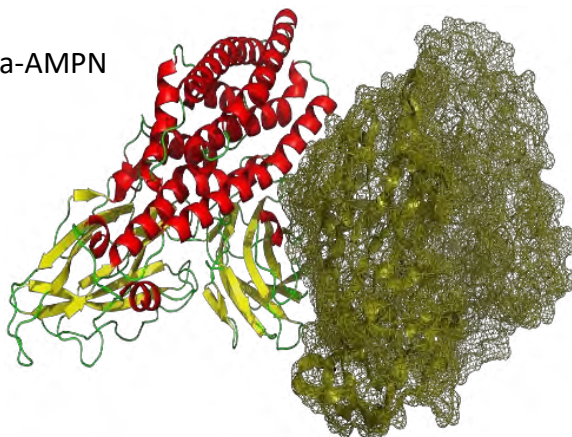


Figura 17. Top 3 complejos docking Cry1Aa-Receptores humanos arrojados por HDock.

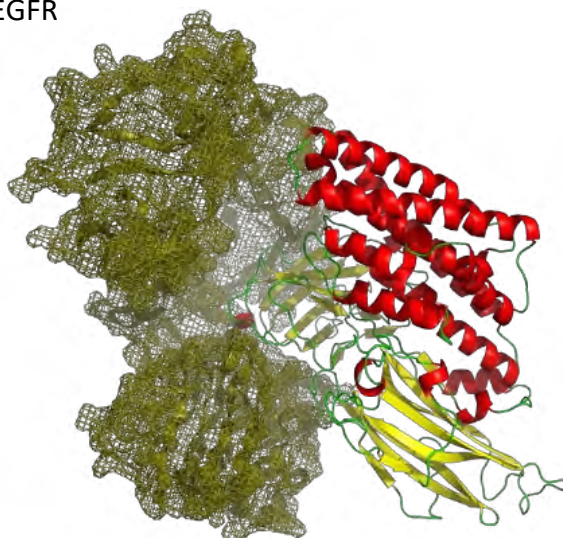
En la tabla 15 se muestran los resultados de docking entre la Cry1Ac y receptores humanos, para esta Cry, el mejor receptor fue EGFR, ligeramente por encima del TRFC, pero ambos manejan el mismo score de confianza.

Tabla 15. Scores de docking de Cry1Ac vs. Receptores humanos

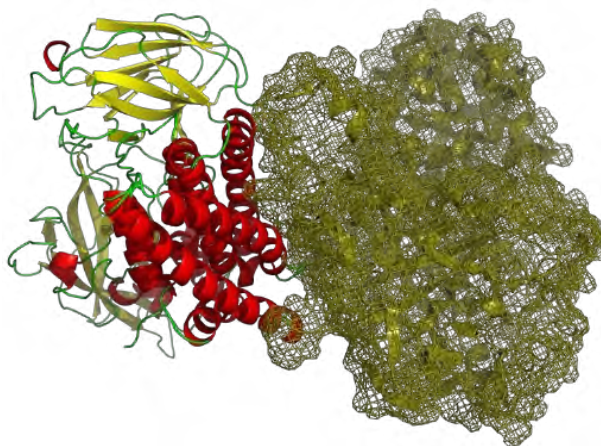
Dock	Rank	1	2	3	4	5
Cry1Ac-FOL	Docking score	-289.10	-280.60	-272.31	-271.59	-270.46
	Score de confianza	0.9417	0.9316	0.9203	0.9192	0.9175
	RMSD	94.35	84.73	75.61	88.37	84.20
Cry1Ac-TRFC	Docking score	-327.33	-319.23	-311.57	-308.75	-292.55
	Score de confianza	0.9720	0.9672	0.9620	0.9599	0.9454
	RMSD	147.74	138.18	148.65	116.39	120.03
Cry1Ac-EGFR	Docking score	-327.41	-307.59	-300.58	-300.44	-299.98
	Score de confianza	0.9720	0.9590	0.9531	0.9530	0.9526
	RMSD	69.34	73.01	72.32	56.86	73.03
Cry1Ac-CD44	Docking score	-273.52	-263.40	-260.77	-260.02	-258.13
	Score de confianza	0.9220	0.9062	0.9016	0.9003	0.8968
	RMSD	50.42	52.30	67.77	64.94	36.29
Cry1Ac-VIPR	Docking score	-277.77	-268.10	-264.56	-257.78	-255.19
	Score de confianza	0.9279	0.9139	0.9081	0.8962	0.8913
	RMSD	48.27	48.46	51.01	51.84	49.26
Cry1Ac-AMPN	Docking score	-304.02	-290.31	-282.17	-281.60	-274.79
	Score de confianza	0.9561	0.9430	0.9336	0.9329	0.9238
	RMSD	128.44	84.54	128.95	165.63	152.39

La figura 18 contiene los 3 mejores complejos de docking Cry1Ac-Receptor humano, tomados a partir de los mejores resultados arrojados en la tabla 15.

a) 1Ac-EGFR



b) 1Ac-TFRC



c) 1Ac-AMPN

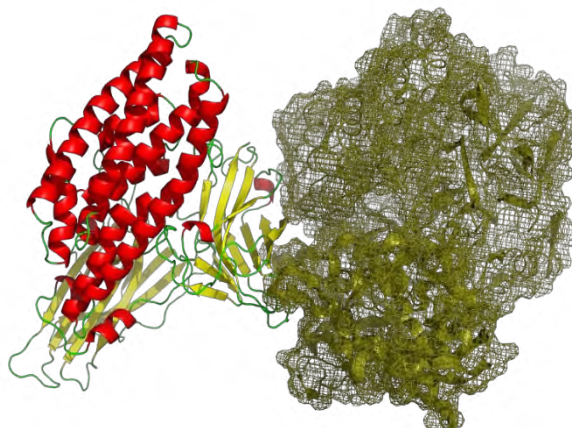


Figura 18. Top 3 complejos docking Cry1Ac-Receptores humanos arrojados por HDock.

8.3.2 ClusPro

En el servidor en línea de ClusPro se cargaron los archivos correspondientes a el ligando y el receptor respectivamente, dejando la opción de Server en su configuración por defecto, es decir, usando CPU como el hardware a utilizar durante el proceso, los resultados arrojados por el servidor son los siguientes (tabla 17 a la 21).

8.3.2.1 MODELOS ALPHAFOLD

Los resultados de docking para los modelos construidos con Alphafold se muestran en la tabla 16 para el caso de la Cry1Aa-Receptor de insecto, y en la tabla 17 los de Cry1Ac-Receptor de insecto. El valor de energía es adimensional y la energía del centro es la que se ocupará de referencia, los mejores receptores fueron ALP y APN1 para Cry1Aa y Cry1Ac respectivamente.

Tabla 16. Scores de docking de Cry1Aa vs. Receptores de insectos generados por IA

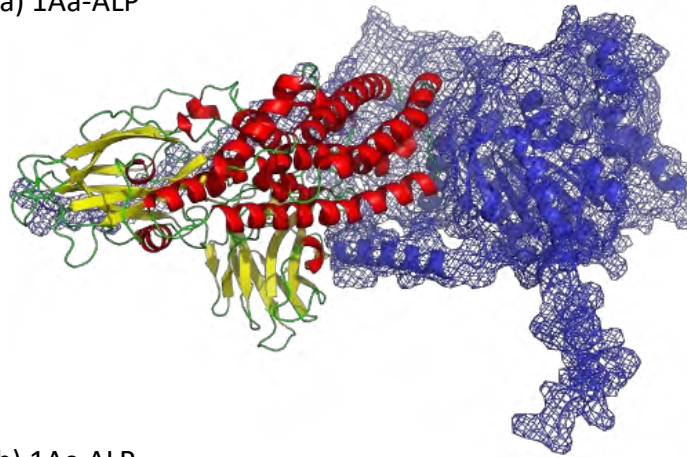
Dock	Rank	1	2	3	4	5
Cry1Aa-APN1	Energía más baja	-1121	-1066.3	-1066.7	-1064	-1011.4
	Energía del centro	-1092.2	-1066.3	-887.5	-953.3	-1011.4
	Tamaño del clúster	30	30	28	27	26
Cry1Aa-APN2	Energía más baja	-1067.5	-1071.1	-1035.8	-1119.3	-994.8
	Energía del centro	-859.7	-978.7	-810	-950.1	-968.6
	Tamaño del clúster	43	32	32	31	29
Cry1Aa-ALP	Energía más baja	-1432.9	-1061.2	-1042.4	-1093.1	-1120.4
	Energía del centro	-1308.3	-1061.2	-1042.4	-944.4	-1120.4
	Tamaño del clúster	38	37	31	27	26

Tabla 17. Scores de docking de Cry1Ac vs. Receptores de insectos generados por IA

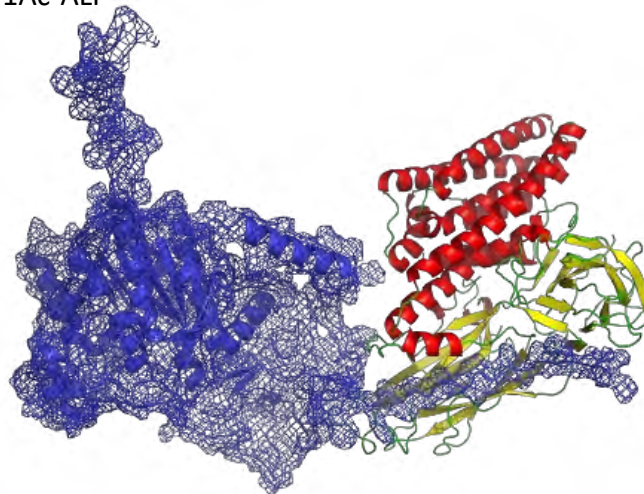
Dock	Rank	1	2	3	4	5
Cry1Ac-APN1	Energía más baja	-1040.9	-1017.2	-929.7	-1011.6	-970.8
	Energía del centro	-1040.9	-846.3	-842.9	-862.9	-917.2
	Tamaño del clúster	48	44	37	28	26
Cry1Ac-APN2	Energía más baja	-978.1	-1036.4	-1009.5	-880.7	-971.3
	Energía del centro	-831	-903	-922	-880.7	-884.2
	Tamaño del clúster	68	41	30	28	26
Cry1Ac-ALP	Energía más baja	-1222	-1133	-1077.8	-1163.8	-1008.8
	Energía del centro	-974	-1014	-986.6	-1122.2	-1008.8
	Tamaño del clúster	47	39	37	29	28

Los mejores complejos de docking entre la Cry1Aa (tabla 16) y Cry1Ac (tabla 17), se muestran en la figura 19, las Cry en amarillo y rojo, y los modelos de AlphaFold en color azul.

a) 1Aa-ALP



b) 1Ac-ALP



c) 1Aa-APN1

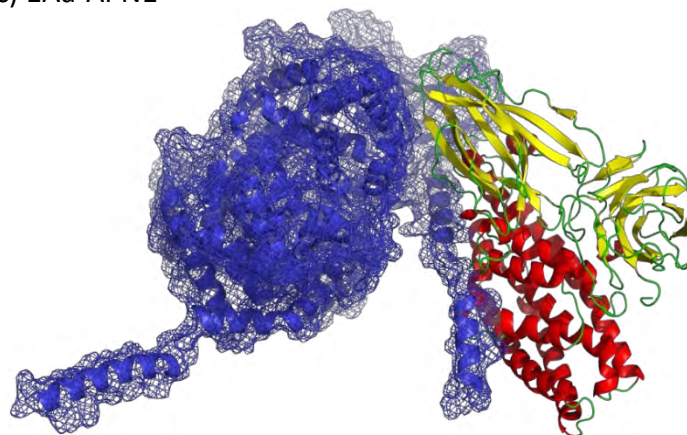


Figura 19. Top 3 complejos docking Cry-Receptores de insecto (AlphaFold) arrojados por ClusPro.

8.3.2.2 MODELOS SWISS MODEL

En esta sección se muestran los resultados de docking entre las Cry y los receptores de insecto modelados por homología, en la tabla 18 para la Cry1Aa y en la tabla 19 para la Cry1Ac. Los mejores receptores coinciden con los de la sección anterior (tabla 16 y 17) siendo ALP y APN1.

Tabla 18. Scores de docking de Cry1Aa vs. Receptores de insectos generados por homología

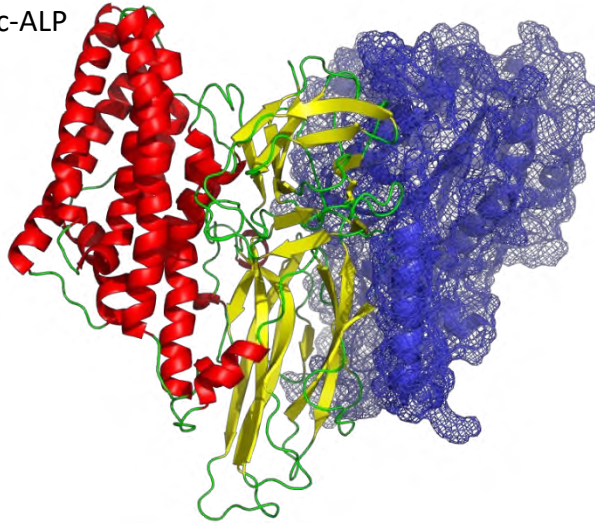
Dock	Rank	1	2	3	4	5
Cry1Aa-APN1	Energía más baja	-864.7	-875.9	-1050.7	-983	-1073.2
	Energía del centro	-838.4	-847.6	-1050.7	-809.4	-1073.2
	Tamaño del clúster	39	33	32	29	28
Cry1Aa-APN2	Energía más baja	-822.4	-936.5	-870.8	-905.1	-819.4
	Energía del centro	-729.4	-761.4	-772.1	-905.1	-800.7
	Tamaño del clúster	55	52	47	41	30
Cry1Aa-ALP	Energía más baja	-1047.1	-962.8	-980	-1004.5	-875
	Energía del centro	-916.4	-839.6	-882	-855.3	-812.6
	Tamaño del clúster	41	33	32	31	30

Tabla 19. Scores de docking de Cry1Ac vs. Receptores de insectos generados por homología

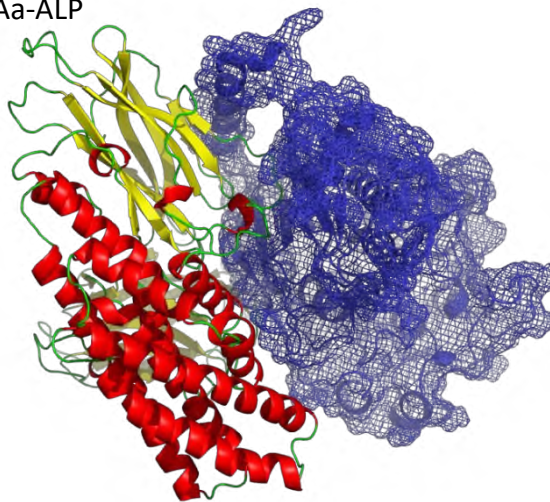
Dock	Rank	1	2	3	4	5
Cry1Ac-APN1	Energía más baja	-959.3	-1097.8	-1021.3	-973.5	-1025.9
	Energía del centro	-935.8	-1097.8	-1021.3	-814.7	-946.2
	Tamaño del clúster	45	31	30	29	27
Cry1Ac-APN2	Energía más baja	-938.6	-905.6	-973.3	-965.4	-766.8
	Energía del centro	-721.9	-719.3	-973.3	-815.7	-741.4
	Tamaño del clúster	58	37	29	28	26
Cry1Ac-ALP	Energía más baja	-1073.4	-980.6	-1068.2	-904.1	-974.6
	Energía del centro	-836.7	-833.4	-948.6	-904.1	-960.8
	Tamaño del clúster	54	49	40	33	24

A continuación, en la figura 20, se muestran los mejores complejos entre las Cry (ambas) y los receptores de insecto modelados por homología (SWISSModel).

a) 1Ac-ALP



b) 1Aa-ALP



c) 1Ac-APN1

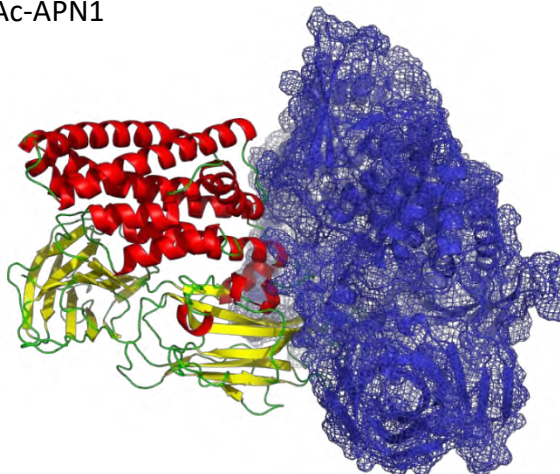


Figura 20. Top 3 complejos docking Cry-Receptores de insecto (SWISS-MODEL) arrojados por ClusPro.

8.3.2.3 DOCKING RECEPTORES HUMANOS

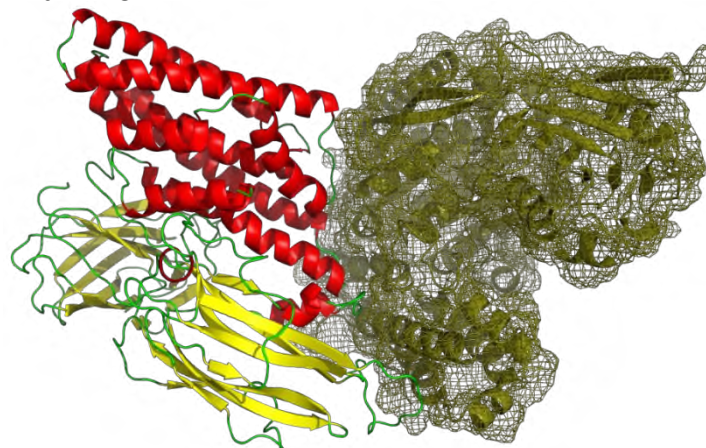
En la tabla 20, están los scores de los complejos de docking Cry1Aa-receptor humano, siendo el receptor TFRC el de mejor resultado para esta Cry, mismo receptor al que se llegó con el programa anterior (HDOCK, tabla 14).

Tabla 20. Scores de docking de Cry1Aa vs. Receptores humanos

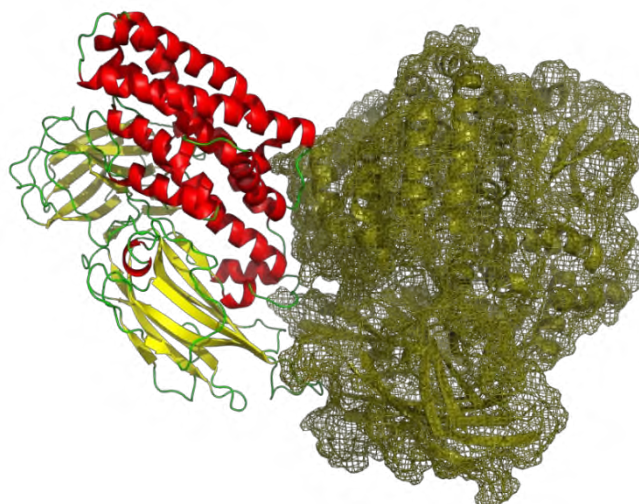
Dock	Rank	1	2	3	4	5
Cry1Aa-FOL	Energía más baja	-745.7	-811.3	-852.7	-922	-786
	Energía del centro	-722.4	-811.3	-709.1	-714.7	-700.1
	Tamaño del clúster	92	78	75	57	54
Cry1Aa-TRFC	Energía más baja	-1124.7	-910.8	-1004.4	-1023	-1030.9
	Energía del centro	-1014.5	-910.8	-1004.4	-1023	-958.8
	Tamaño del clúster	85	43	38	38	35
Cry1Aa-EGFR	Energía más baja	-830.2	-890.8	-1010.2	-882.4	-867.3
	Energía del centro	-810	-890.8	-859.7	-882.4	-759.1
	Tamaño del clúster	86	57	42	33	32
Cry1Aa-CD44	Energía más baja	-865.5	-937.1	-812.4	-769.7	-801.3
	Energía del centro	-706.4	-937.1	-712	-711.2	-715.9
	Tamaño del clúster	71	69	49	46	42
Cry1Aa-VIPR	Energía más baja	-808.1	-801.1	-814.1	-924.5	-871.5
	Energía del centro	-808.1	-695.4	-764.8	-924.5	-871.5
	Tamaño del clúster	75	47	44	43	41
Cry1Aa-AMPN	Energía más baja	-875.5	-754.6	-803.6	-766.7	-775.9
	Energía del centro	-875.5	-724.9	-782	-752	-706.9
	Tamaño del clúster	57	53	38	30	26

Los 3 mejores modelos de los complejos de docking puntuados en la tabla 20, se muestran la figura 21, correspondientes a la Cry1Aa-Receptor humano.

a) 1Aa-TFRC



b) 1Aa-AMPN



c) 1Aa-CD44

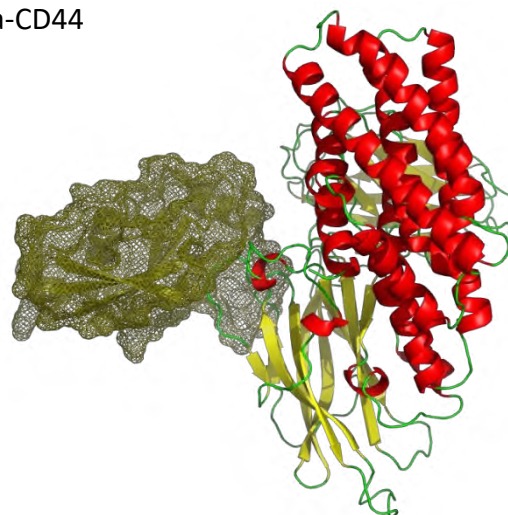


Figura 21. Top 3 complejos docking Cry1Aa-Receptores humanos arrojados por ClusPro.

Para en el docking de Cry1Ac con receptores humanos, los resultados están alojados en la tabla 21, los receptores de mejor desempeño fueron el VIPR seguido del TRFC, este último también fue el segundo mejor complejo con el programa HDock (tabla 15).

Tabla 21. Scores de docking de Cry1Ac vs. Receptores humanos

Dock	Rank	1	2	3	4	5
Cry1Ac-FOL	Energía más baja	-850.9	-860.6	-910.8	-827.4	-782.7
	Energía del centro	-751.7	-750.8	-825.1	-666.6	-701
	Tamaño del clúster	147	132	113	43	37
Cry1Ac-TRFC	Energía más baja	-1056.7	-989.8	-1006.9	-946.9	-900.3
	Energía del centro	-911.5	-982.4	-1006.9	-871	-900.3
	Tamaño del clúster	54	47	45	41	35
Cry1Ac-EGFR	Energía más baja	-1063.5	-945.8	-926.5	-940.9	-857.9
	Energía del centro	-772.1	-807	-832.4	-741	-857.9
	Tamaño del clúster	55	39	36	36	25
Cry1Ac-CD44	Energía más baja	-965.3	-807.6	-845	-881.1	-874
	Energía del centro	-752.2	-746.1	-781.4	-881.1	-691.6
	Tamaño del clúster	121	84	52	50	45
Cry1Ac-VIPR	Energía más baja	-970.5	-883.1	-995.3	-849.4	-868.3
	Energía del centro	-970.5	-883.1	-846.7	-773.2	-787.5
	Tamaño del clúster	45	45	44	44	38
Cry1Ac-AMPN	Energía más baja	-997.9	-916	-951	-856.6	-787
	Energía del centro	-865.6	-681	-715.3	-696.6	-777.2
	Tamaño del clúster	67	55	48	45	41

La figura 22, contiene los mejores modelos obtenidos del docking Cry1Ac-Receptor humano, tomados de los datos de la tabla 21.

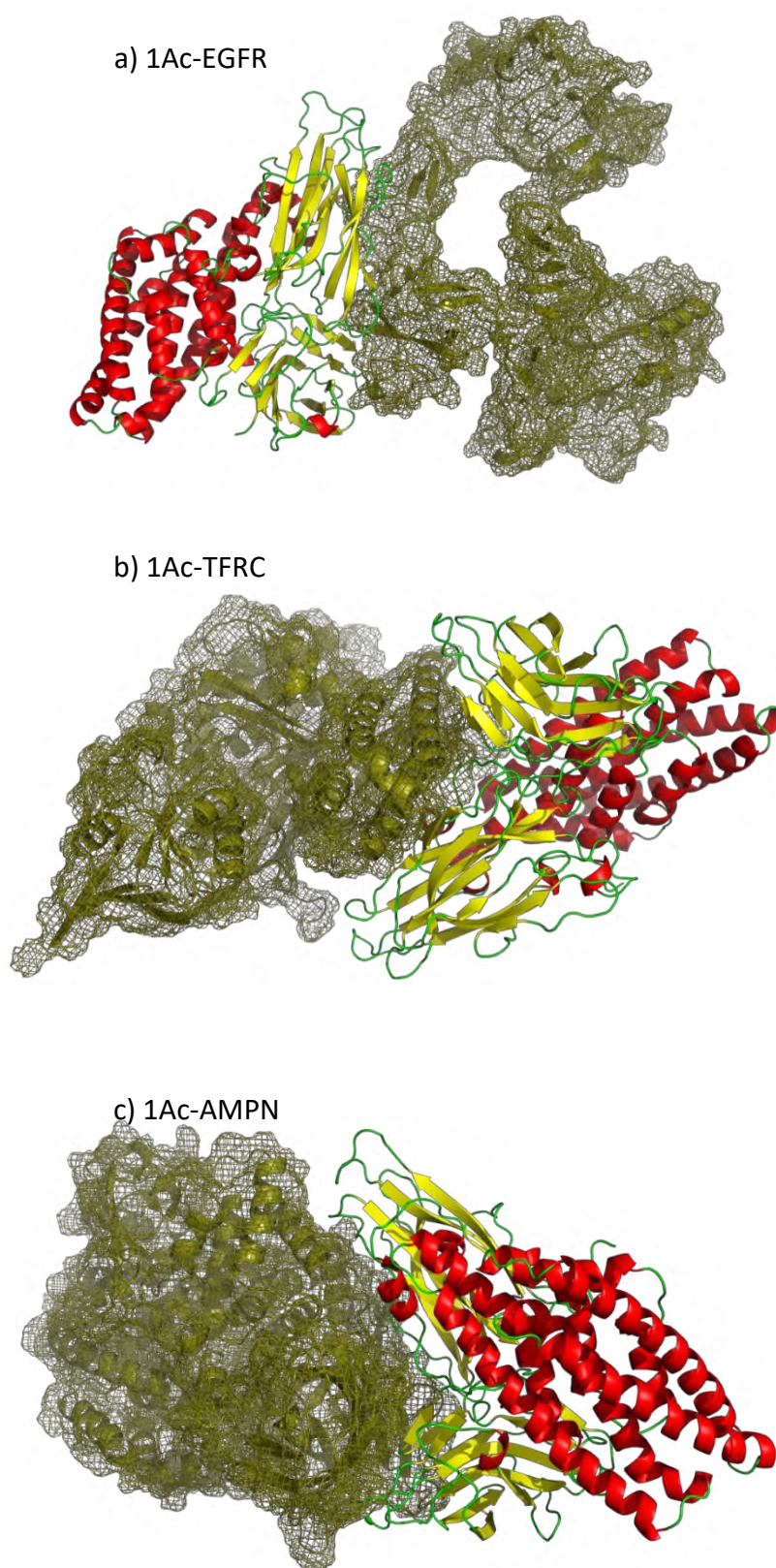


Figura 22. Top 3 complejos docking Cry1Ac-Receptores humanos arrojados por ClusPro.

8.3.3 HEX

En el programa ejecutado localmente, las pruebas con los modelos generados por IA no tuvieron éxito a excepción de una, que es la que se muestra en la tabla 22.

8.3.3.1 MODELOS ALPHAFOLD

En la tabla 22, se muestra el único resultado obtenido en el docking Cry-Receptor de insecto, ya que el programa HEX no tuvo un buen desempeño con los receptores modelados con Alphafold, así también, en la figura 23 se muestra el único modelo de complejo de docking obtenido por HEX, por la compatibilidad con el visualizador de modelos, las estructuras tridimensionales no conservan el formato de visualización con los anteriores modelos mostrados (en las figuras anteriores), pero sí los colores que los representan, en rojo y amarillo los ligandos (Cry), y en azul los receptores, esto, para todos los modelos obtenidos con HEX.

Tabla 22. Scores de docking de Cry1Aa vs. Receptor APN generado por IA

Dock	Rank	1	2	3	4	5
Cry1Aa-APN2	Energía total	-590.8	-589.0	-570.6	-568.3	-562.8

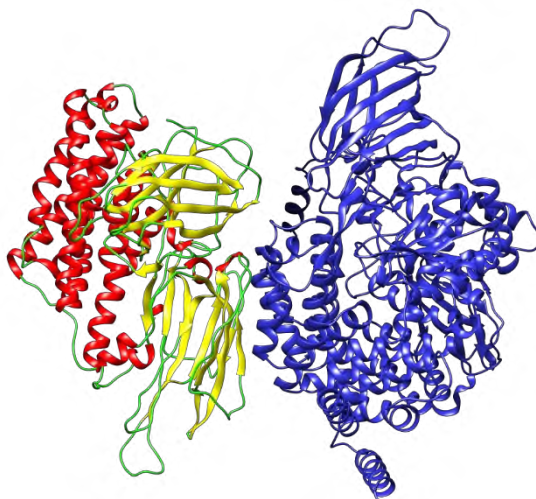


Figura 23. Complejo docking Cry1Aa-Receptor APN2 (SWISS-MODEL) arrojado por HEX.

8.3.3.2 MODELOS SWISS MODEL

Para los resultados usando los receptores modelados con SWISSModel no se tuvieron problemas con el docking, en la tabla 23 se muestran los datos para la Cry1Aa y en la tabla 24 para el docking con la Cry1Ac. En el caso de este programa, el mejor receptor para ambas Cry modeladas por homología, fue el ALP, que se encuentra en los mejores para HDOCK y Cluspro también.

Tabla 23. Scores de docking de Cry1Aa vs. Receptores de insectos generados por homología

Dock	Rank	1	2	3	4	5
Cry1Aa-APN1	Energía total	-673.6	-660.5	-659.5	-644.6	-629.6
Cry1Aa-APN2	Energía total	-892.2	-788.7	-657.0	-654.0	-616.0
Cry1Aa-ALP	Energía total	-902.2	-888.1	-873.2	-830.6	-825.5

Energía en kcal/mol.

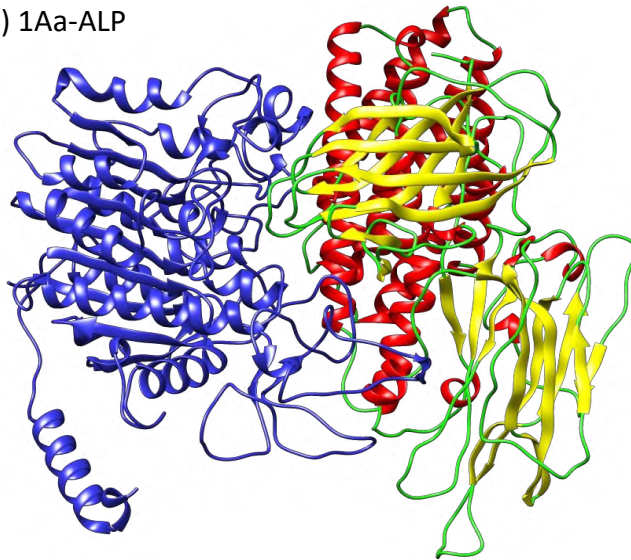
Tabla 24. Scores de docking de Cry1Ac vs. Receptores de insectos generados por homología

Dock	Rank	1	2	3	4	5
Cry1Ac-APN1	Energía total	-690.5	-687.2	-682.7	-676.0	-673.6
Cry1Ac-APN2	Energía total	-702.0	-696.7	-679.0	-664.9	-648.2
Cry1Ac-ALP	Energía total	-871.1	-855.0	-844.2	-813.7	-796.1

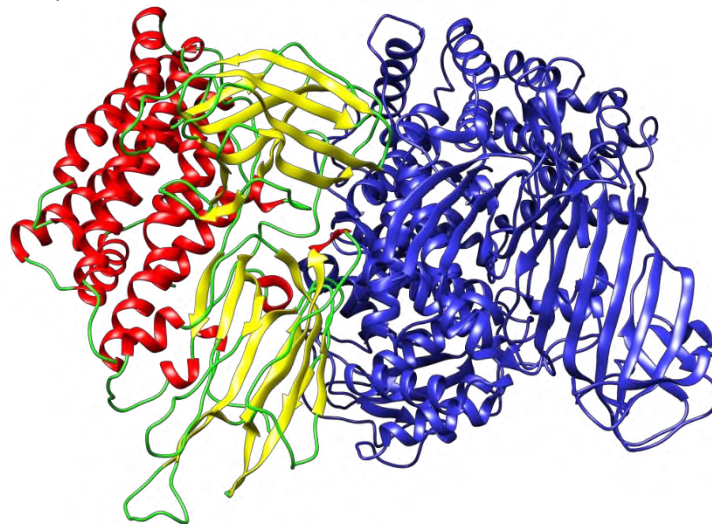
Energía en kcal/mol.

Los mejores modelos de las tablas 23 y 24, se muestran en la figura 24, al igual que con la figura de los modelos con Alphafold, se presenta en un formato visual distinto, pero en este caso se conservan los 3 mejores complejos según datos de la tabla 23 y 24, al igual que como se venía mostrando en las figuras de los dos programas anteriores.

a) 1Aa-ALP



b) 1Aa-APN2



c) 1Ac-ALP

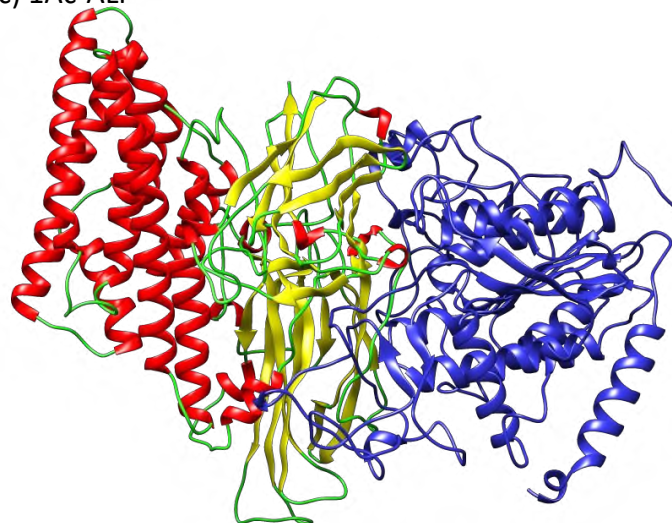


Figura 24. Top 3 complejos docking Cry-Receptores de insecto (SWISS-MODEL) arrojados por HEX.

8.3.3.3 DOCKING RECEPTORES HUMANOS

Con los receptores humanos, HEX, no presentó problemas, en la tabla 25 se muestran los datos de docking con la Cry1Aa y receptores humanos, y en la figura 25, los 3 mejores modelos mejor puntuados. De manera similar con HDOCK y ClusPro, el receptor de mejor desempeño para la Cry1Aa fue el TFRC (tabla 25).

Tabla 25. Scores de docking de Cry1Aa vs. Receptores humanos

Dock	Rank	1	2	3	4	5
Cry1Aa-FOL	Energía total	-247.0	-238.4	-229.3	-221.6	-219.6
Cry1Aa-TRFC	Energía total	-856.7	-753.1	-748.6	-733.5	-722.5
Cry1Aa-EGFR	Energía total	-589.5	-550.2	-521.6	-507.7	-498.4
Cry1Aa-CD44	Energía total	-701.0	-684.3	-668.0	-662.7	-658.9
Cry1Aa-VIPR	Energía total	-839.1	-742.1	-661.2	-633.7	-629.0
Cry1Aa-AMPN	Energía total	-699.8	-607.0	-592.2	-556.5	-550.9

Energía en kcal/mol.

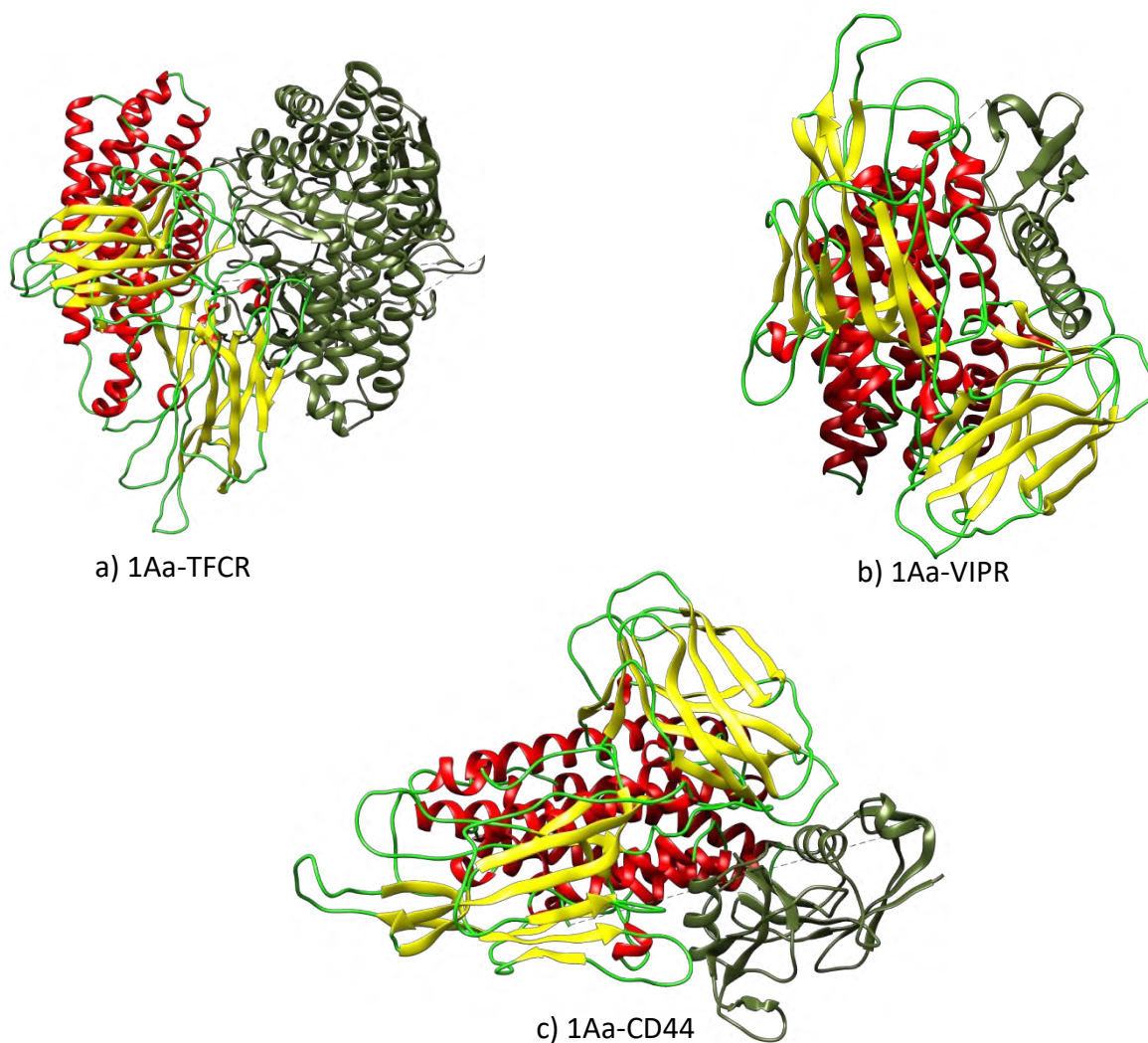


Figura 25. Top 3 complejos docking Cry1Aa-Receptores humanos arrojados por HEX.

La tabla 26 contiene los datos para docking Cry1Ac con receptores humanos, y en la figura 26 los respectivos mejor puntuados. También para esta Cry, el receptor con mejor score fue el TFRC, comparable con los complejos de esta misma Cry en los programas HDOCK y ClusPro.

Tabla 26. Scores de docking de Cry1Ac vs. Receptores humanos

Dock	Rank	1	2	3	4	5
Cry1Ac-FOL	Energía total	-752.5	-695.0	-662.7	-660.8	-644.5
Cry1Ac-TRFC	Energía total	-820.9	-744.8	-733.2	-709.2	-706.2
Cry1Ac-EGFR	Energía total	-696.7	-628.6	-576.3	-515.0	-511.5
Cry1Ac-CD44	Energía total	-747.7	-740.0	-690.0	-678.5	-669.0
Cry1Ac-VIPR	Energía total	-633.0	-622.7	-620.7	-617.2	-615.2
Cry1Ac-AMPN	Energía total	-721.9	-670.8	-621.5	-567.7	-559.6

Energía en kcal/mol.

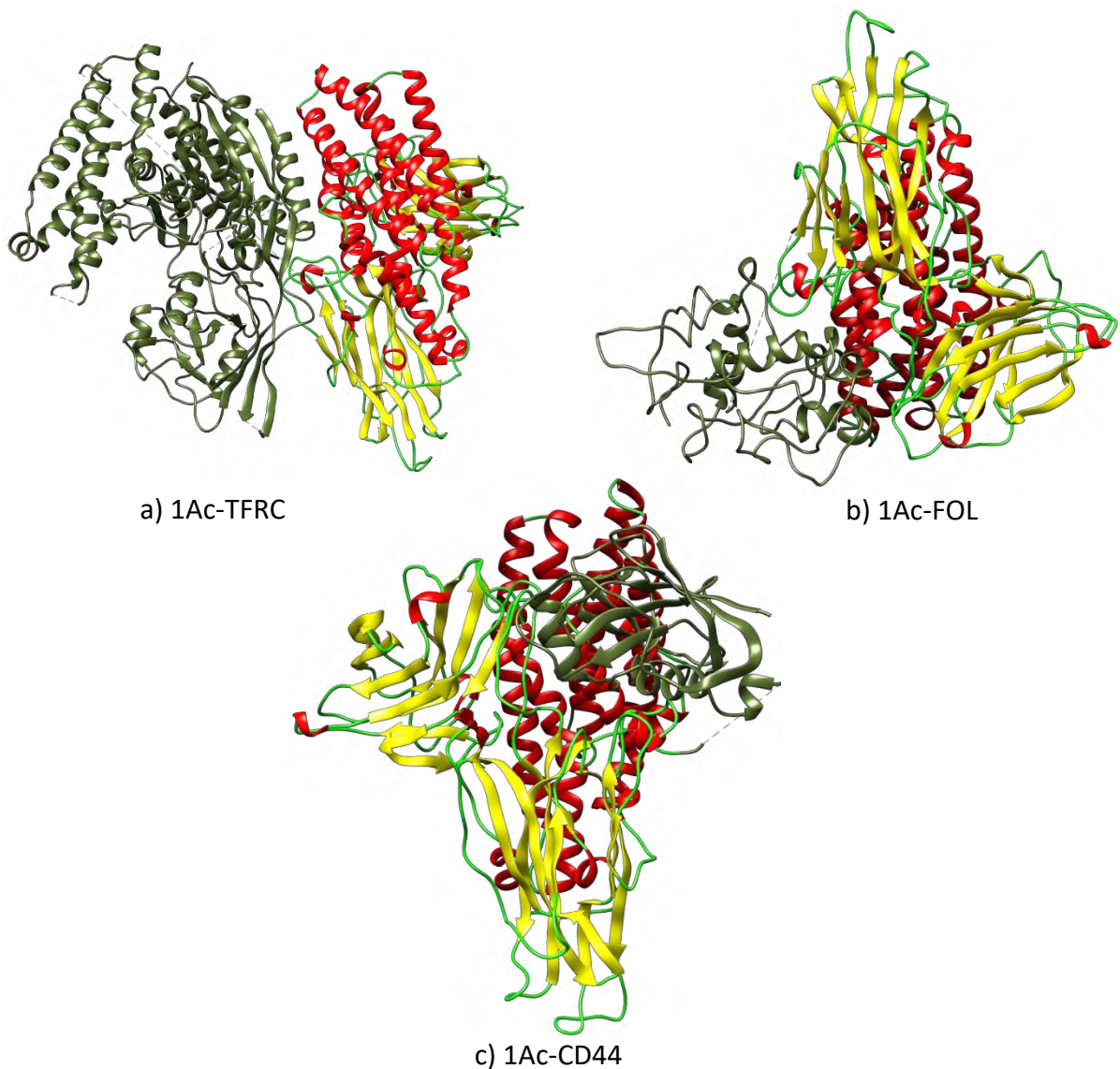


Figura 26. Top 3 complejos docking Cry1Ac-Receptores humanos arrojados por HEX.

8.3.4 Top 3 receptores para proteínas Cry.

A continuación, se resumen los mejores scores de complejos docking de las Cry, con los receptores de insecto y los receptores humanos, se presentan en la tabla 27 los 3 mejores complejos (con los mejores receptores) con el programa HDock para ambas Cry y con ambos tipos de modelado (IA y homología), se obtuvo un mejor desempeño utilizando los modelos generados con AlphaFold, ya que los valores de score son más bajos. En la tabla 28, se encuentran los mejores complejos con los receptores humanos que tuvieron mejor desempeño en el docking para las 2 Cry, en ambas tablas se ordenan de arriba a abajo las puntuaciones, según su valor de score. Para las Cry modeladas por homología el mejor receptor fue APN1, y en el caso de los modelos de IA, los dos mejores fueron ALP y APN2 (tabla 27). Para los receptores humanos HDock arrojó que el receptor de mejor desempeño fue el TFRC para ambas Cry, aunque para el caso de Cry1Ac esté ligeramente por debajo de EGFR, el score de confianza es idéntico (tabla 28).

Tabla 27. Tres mejores receptores de insecto para Cry1Aa y Cry1Ac según HDock

	HDock					
	SWISS-MODEL			ALPHAFOLD		
	Receptor	Confidence	Score	Receptor	Confidence	Score
Cry1Aa	APN1	0.9442	-291.41	APN2	0.9646	-315.26
	ALP	0.9261	-276.41	ALP	0.9605	-309.6
	APN2	0.9202	-272.27	APN1	0.9603	-309.3
Cry1Ac	APN1	0.9629	-312.79	ALP	0.9807	-346.35
	APN2	0.949	-296.22	APN1	0.962	-311.61
	ALP	0.949	-296.2	APN2	0.938	-285.8

Tabla 28. Tres mejores receptores humanos para Cry1Aa y Cry1Ac según HDock

	HDock		
	Receptor	Confidence	Score
Cry1Aa	TFRC	0.9839	-355.56
	EGFR	0.9609	-310.06
	AMPN	0.9336	-282.16
Cry1Ac	EGFR	0.972	-327.41
	TFRC	0.972	-327.33
	AMPN	0.9561	-304.02

Con el programa ClusPro se ordenaron de igual forma que las dos tablas anteriores, en la tabla 29, se encuentra los 3 de receptores con mejor desempeño para las dos Cry, con los dos tipos de modelado, de igual manera que con los datos para HDock, en ClusPro se obtuvieron mejores resultados con los modelos de AlphaFold.

La tabla 30 tiene los datos de los complejos para ambas Cry con los receptores humanos; para los datos de ClusPro, más adelante se usa los valores de energía para comparar con los otros dos programas, aunque en el programa el tamaño de los clusters es importante. Con ClusPro, el receptor ALP fue el de mejor resultado para la unión con ambas Cry, tanto en homología como con IA (tabla 29). Por otro lado, de los receptores humanos, TFRC y EGFR son los de mejor desempeño ya que son los de menor score.

Tabla 29. Tres mejores receptores de insecto para Cry1Aa y Cry1Ac según ClusPro

	CLUSPRO					
	SWISS-MODEL			ALPHAFOLD		
	Receptor	Cluster	Score (Energía)	Receptor	Cluster	Score (Energía)
Cry1Aa	ALP	41	-1047.1	ALP	38	-1432.9
	APN1	39	-864.7	APN1	30	-1121
	APN2	55	-822.4	APN2	43	-1067.5
Cry1Ac	ALP	54	-1073.4	ALP	47	-1222
	APN1	45	-959.3	APN1	48	-1040.9
	APN2	58	-938.6	APN2	68	-978.1

*La energía del score esta correlacionada con la energía de libre de Gibbs, sin embargo, es adimensional.

Tabla 30. Tres mejores receptores humanos para Cry1Aa y Cry1Ac según ClusPro

	CLUSPRO		
	Receptor	Cluster	Score (Energía)
Cry1Aa	TFRC	85	-1124.7
	AMPN	57	-875.5
	CD44	71	-865.5
Cry1Ac	EGFR	55	-1063.5
	TFRC	54	-1056.7
	AMPN	67	-997.9

La energía del score esta correlacionada con la energía de libre de Gibbs, sin embargo, es adimensional.

Los datos de los mejores receptores para el programa HEX están truncados debido al desempeño del docking con los receptores modelados con AlphaFold, con los modelos de insectos hechos por SWISSModel el docking se llevó a cabo de manera correcta y se seleccionaron los 3 mejores receptores, a diferencia de Alphafold, en el que solo se reportó 1 solo receptor en la tabla 31 para hacer la comparación, ALP es el mejor receptor (tabla 31) para ambas Cry (receptor modelado por homología).

En la tabla 32 se encuentran los mejores receptores humanos obtenidos con el docking en el programa HEX para las dos proteínas Cry ordenadas de menor a mayor energía, de nuevo, el receptor mejor puntuado es el TFRC, siendo constante en los tres programas usados en este trabajo.

Tabla 31. Tres mejores receptores de insecto para Cry1Aa y Cry1Ac según HEX

	HEX			
	SWISS-MODEL		ALPHAFOLD	
	Receptor	Energía*	Receptor	Energía*
Cry1Aa	ALP	-902.2	-	-
	APN2	-892.2	APN2	-590.8
	APN1	-673.6	-	-
Cry1Ac	ALP	-871.1	-	-
	APN2	-702	-	-
	APN1	-690.5	-	-

*Energía de enlace (kcal/mol).

Tabla 32. Tres mejores receptores humanos para Cry1Aa y Cry1Ac según HEX

	HEX	
	Receptor	Energía*
Cry1Aa	TFRC	-856.7
	VIPR	-839.1
	CD44	-701
Cry1Ac	TFRC	-820.9
	FOL	-752.5
	CD44	-747.7

*Energía de enlace (kcal/mol).

8.3.5 Regiones de interacción.

A continuación, se muestra gráficamente las zonas de interacción entre los dos mejores receptores (TFRC en azul y EGFR en naranja) y las proteínas Cry1Aa y Cry1Ac, destacando los aminoácidos de los receptores en la tabla 33, los cuales interactúan con los residuos de los aminoácidos del ligando.

La figura 27 muestra la interacción entre la Cry1Aa en color amarillo y verde, con el receptor TFRC en azul, hay al menos 10 aminoácidos del receptor en contacto, en comparación, la Cry1Ac presenta una interacción con 13 aminoácidos del mismo receptor.

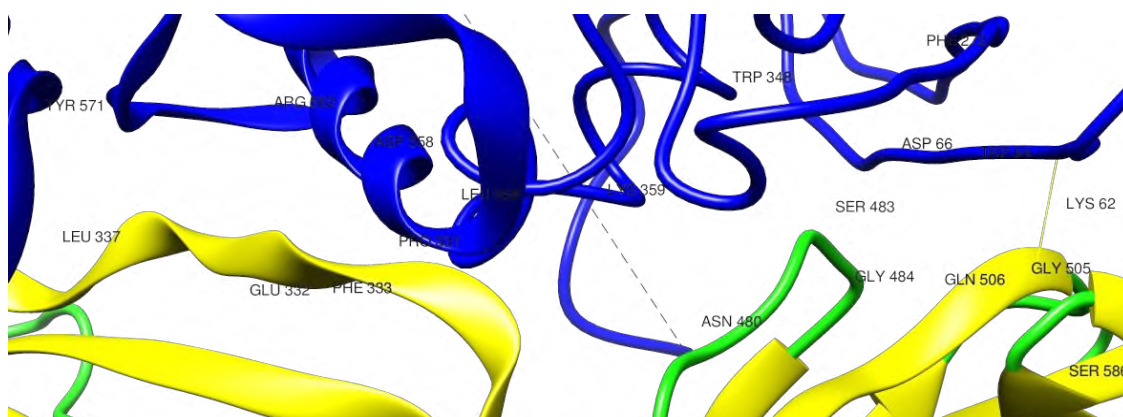


Figura 27. Región de interacción entre Cry1Aa y receptor TFRC (docking con HDock).

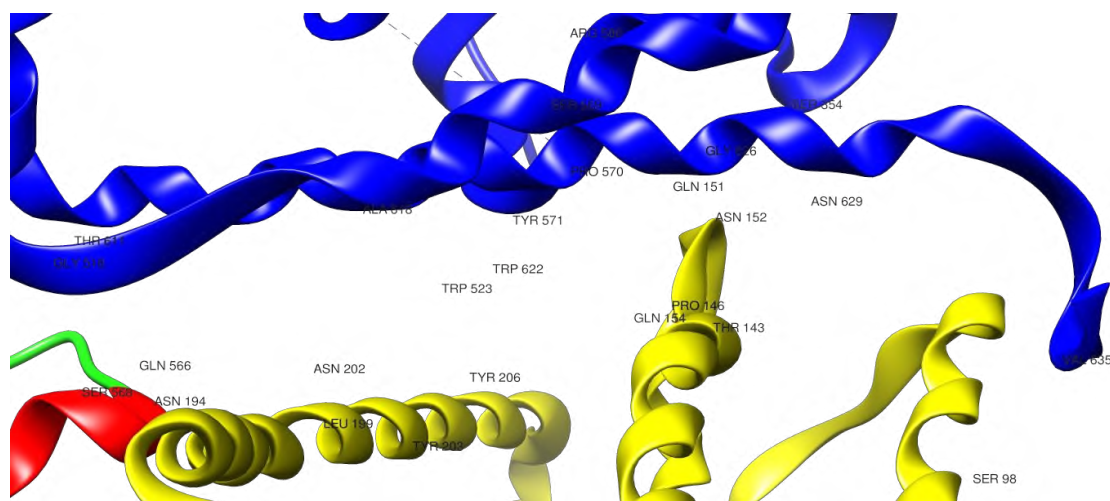


Figura 28. Región de interacción entre Cry1Ac y receptor TFRC (docking con HDock).

Los modelos obtenidos con ClusPro, mostraron mayores interacciones entre los ligandos y receptores, teniendo la Cry1Aa interacción con 23 aminoácidos del receptor TFRC (figura 29), por su parte, la Cry1Ac interaccionó con 21 aminoácidos, mostrados en la figura 30.

Los residuos del receptor TFRC que prevalecieron mayormente entre las pruebas con dos programas (HDOCK y ClusPro) fueron TRP622, TRP636, ARG614, TYR571, ALA618, VAL635. Los aminoácidos en HEX no tuvieron coincidencia con los otros dos programas, tampoco entre las dos Cry en el propio HEX (tabla 33).

Los residuos del receptor EGFR que prevalecieron mayormente entre las pruebas con HDOCK y ClusPro fueron THR 605, PRO607, PRO613, GLU610, ASP588. El EGFR no figuró entre los mejores receptores según HEX (tabla 33).

Tabla 33. Residuos aminoacídicos de los receptores involucrados en el docking.

COMPLEJO DOCKING	RESIDUOS DEL RECEPTOR EN CONTACTO
<i>1Aα-TFRC (HDOCK)</i>	LYS62, TRP64, ASP66, PHE272, TRP348, LEU358, LYS359, ASP558, ARG562, TYR571
<i>1Ac-TFRC (HDOCK)</i>	SER354, GLY518, TRP523, SER569, PRO570, TYR571, ARG580, THR611, ALA618, TRP622, GLY626, ASN629, VAL635
<i>1Aα-TFRC (ClusPro)</i>	GLY351, LEU353, SER354, SER355, LEU358, LYS359, LEU519, SER520, TRP523, TYR565, HIS566, LEU568, PRO570, TYR571, ARG614, ALA618, LEU619, TRP622, GLN625, GLY626, ASN629, VAL635, TRP636
<i>1Ac-TFRC (ClusPro)</i>	MET517, GLY518, LEU519, GLN522, TRP523, SER526, ARG533, ARG537, TYR571, THR611, ARG614, ASN615, ALA618, LEU619, TRP622, GLN625, GLY626, ASP634, VAL635, TRP636, ASP637
<i>1Aα-TFRC (HEX)</i>	LYS123, LYS143, THR168, LYS169, ASP238, LYS240, SER374, VAL399, GLN402, TYR405, GLN406, SER408, ASN409, SER412, LYS413, VAL414, GLU415, LYS416, ASP442, PRO463, ASP506, GLN509, TRP584, PRO592, LEU595, GLU596, LYS599
<i>1Ac-TFRC (HEX)</i>	VAL35, PRO36, ARG37, GLU38, SER41, GLN42, LYS43, GLU45, ASN46, ARG65, ASP76, VAL262, LEU263, LYS264, LYS267, TRP294, THR451, MET452,
<i>1Aα-EGFR (HDOCK)</i>	TYR251, GLN541, ALA542, ASN544, CYS555, ILE556, GLN557, CYS558, ILE562, GLY564, HIS566, CYS567, TYR586, ASP588, ALA589, GLY590, HIS591, VAL592, CYS593, HIS594, HIS597, PRO598, ASN599, THR605, GLY606, PRO607, GLY608, LEU609, GLU610
<i>1Ac-EGFR (HDOCK)</i>	TYR88, LEU120, HIS121, GLY122, ARG141, SER145, SER146, LYS188, ILE189, GLN193, ASN210, GLN211, ARG228, LYS229, ASP238, TYR561, LYS569, THR570, CYS571, ALA573, TRP584, TYR602, CYS604, THR605, CYS612, PRO613
<i>1Ac-EGFR (ClusPro)</i>	TYR251, LEU517, VAL526, ASN544, ILE545, THR548, ASP553, ASN554, CYS555, ILE556, GLN557, CYS558, ILE562, GLY564, PRO565, TYR586, ALA587, ASP588, ALA589, VAL592, HIS594, LEU595, THR605, GLY606, PRO607, LEU609, GLU610, PRO613

Para las proteínas Cry, el dominio II, el cual es una β -prisma compuesto por tres láminas β antiparalelas que se empaquetan entre sí, es sugerido como el dominio de unión a receptores (Xu *et al.*, 2014). Por el lado de los receptores, TFRC, posee tres dominios, el similar a proteasa, el dominio apical y el dominio helicoidal (Aisen, 2004), de los cuales, el dominio helicoidal se describe como el que comprende todo el sitio de unión (Kleven *et al.*, 2018), abarcando los residuos 607-760 (Aisen, 2004).

El receptor EGFR tiene como sitio de unión dominios I y III, conocidos también como L1 y L2 (Ogiso *et al.*, 2002), esto se puede corroborar en la base de datos InterPro (<https://www.ebi.ac.uk/interpro/protein/UniProt/P02786/entry/InterPro/>). En la tabla 34 se muestra si tanto las Cry como los receptores, se unieron en regiones reportadas en literatura como regiones de unión a otras proteínas.

Tabla 34. Dominios de unión del ligando y receptor involucrados en el docking.

COMPLEJO DOCKING	Dominio de unión Cry involucrado	Dominio(s) de unión receptor involucrado
1A α -TFRC (HDOCK)	Sí	No
1Ac-TFRC (HDOCK)	No	Sí
1A α -TFRC (ClusPro)	Sí	Sí
1Ac-TFRC (ClusPro)	Sí	Sí
1A α -TFRC (HEX)	Sí	No
1Ac-TFRC (HEX)	Sí	No
1A α -EGFR (HDOCK)	No	Sí
1Ac-EGFR (HDOCK)	Sí	Sí
1Ac-EGFR (ClusPro)	Sí	No

9. DISCUSIÓN

Para la selección de receptores humanos, entre los autores revisados, Kunjiappan *et al.* (2021) hacen mención que a sobreexpresión de los receptores folato y transferrina en la línea MCF-7 de cáncer de mama, así también, para cáncer de mama, Large *et al.* (2019) reportaron sobreexpresados los receptores transferrina, folato, EGFR y CD44, además, Hoppenz *et al.* (2020) mencionan a los receptores EGFR y VIP como sobreexpresados en cáncer de mama también. Por otro lado, Suarez-Barrera *et al.* (2022) utilizan el receptor APN (aminopeptidasa) para observar su interacción con la PS2Aa1, ya que se ha sugerido como receptor en el mecanismo de acción de esa parasporina y posee una fuerte actividad frente a la línea MCF-7 aunque esta carece del APN, sin embargo, se sabe que PS2Aa1 está relacionada con las proteínas Cry que utilizan APN como receptor en el intestino medio de insectos (Suarez-Barrera *et al.*, 2022). En la tabla 2 se enlistan los receptores empleados.

Se eligió un receptor que se uniera a una Cry, uno que se uniera a la otra y un tercero el cual no se uniera a ninguna de las Cry seleccionadas (tabla 3), siendo los siguientes:

APN2 (*P. xylostella*) es una aminopeptidasa de clase 2, un estudio indicó que este receptor posee una región putativa de unión a las proteínas Cry1A, entre ellas Cry1Aa y Cry1Ab (Nakanishi *et al.*, 2002). APN1 (*M. sexta*) es también una aminopeptidasa clase 2, sin embargo, ha demostrado afinidad a la Cry1Ab (Soberón *et al.*, 2010) y no a las proteínas Cry usadas en este trabajo. ALP (*H. virescens*) es una glucoproteína de membrana GPI-anclada la cual se ha descrito que presenta unión a la Cry1Ac, la cual se demostró mediante análisis ligand blot de las microvellosidades del intestino medio del insecto (Pigott & Ellar, 2007).

El modelado por homología se llevó a cabo a través de SWISS-MODEL mediante su servidor en línea, los modelos obtenidos de los tres receptores de insectos fueron deficientes, esto se debe al nivel de homología de las secuencias de interés y las de los templados, si bien el porcentaje de cobertura de secuencias fueron altas, no así lo fueron los porcentajes de identidad, 32% para APN1, 31.6% para APN2 y 43.5% para ALP (tabla 4), esto se puede justificar por la divergencia evolutiva entre organismos, los modelos para APN1 y APN2 se generaron utilizando un templado de *Anopheles gambiae*, el cual es un complejo de especies de mosquitos del orden díptero (Lanzaro & Lee, 2013), y para ALP se empleó un templado de *Homo sapiens*; los valores de GMQE para cada modelo fueron, para APN1 y ALP de 0.69 y para APN2 de 0.68, en este sentido, los valores de QMEANDisCo de los 3 modelos generados se

encuentran alrededor de 0.7 (tabla 5), siendo estos de una calidad media, a pesar de esto, se procedió utilizar los modelos para las pruebas de docking.

En el caso de los modelos obtenidos con Alphafold (modelado con IA), se tienen 2 indicadores de calidad, siendo el pLLDT y gráficos PAE, el pLLDT para los 3 modelos fueron: 93.2, 92.49 y 91.1 para APN1, APN2 y ALP respectivamente (tabla 6), siendo el rango de este valor de 0-100, se consideran modelos de alta calidad por encima de 90, por otro lado, los gráficos PAE (figura 14) arrojan un color verde intenso en la mayoría de área de los ejes X y Y, delimitado por una zona en blanco, esto indica que las posiciones relativas de los extremos de las proteínas no se han modelado correctamente, esto se puede apreciar también en la coloración amarillo-naranja de los extremos de los modelos con la paleta de colores pLLDT de AlphaFold (figura 12).

Por otra parte, para comparar ambos tipos de modelado se usó la puntuación de MolProbity, los valores MolProbity para los modelos generados con IA (tabla 6) son mejores que los modelos generados por homología (tabla 5), de igual manera, el porcentaje de ángulos en regiones favorables es mayor para los modelos de IA (tabla 6), tomando en cuenta que los templados para el modelado por homología eran de baja calidad se podía esperar que los modelos generados a partir de estos fueran de relativamente de baja calidad, entonces, englobando todos los indicadores de calidad, los modelos por IA fueron mejores que los de homología.

De HDock, tanto los modelos generados por homología e IA los valores de MaxSub y LGScore se encuentran por encima de "correcto"; se puede observar valores más altos de estos 2 indicadores para los modelos de AlphaFold en APN1 y APN2 (tabla 9), lo que se podría justificar con la calidad del modelado puesto que los templados para esos modelos generados por homología eran de baja calidad, en cambio en AlphaFold son los modelos con un valor de pLLDT más alto, sin embargo, para el caso de ALP pasa lo contrario, el modelo por homología presenta valores más altos de MaxSub y LGScore que el modelo por IA a pesar de que el pLLDT indica que es un modelo de calidad "alta" mientras que GMQE clasifica el modelo por homología como medianamente bueno, aunque, considerando el porcentaje de ángulos favorables, el ALP de homología, presenta un porcentaje mayor que el ALP de IA, esto también sucede con los modelos APN1 y APN2 pero con los modelos de IA con un porcentaje mayor que los de homología.

Los resultados de docking, para cada programa, se subdividen en 3 partes de acuerdo con el origen de los receptores: receptores por homología, receptores por IA y receptores humanos.

En HDOCK, los mejores receptores dado el score proporcionado por el propio programa fueron: Para Cry1Aa, el receptor APN1 de SWISS-MODEL con score -291.41 y valor de confidence de 0.9442 (tabla 27), el receptor APN2 de AlphaFold con score -315.26 y confidence 0.9646 (tabla 27), y el receptor humano TFRC con score -355.56 y confidence 0.9839.

Para Cry1Ac, el receptor APN1 de SWISS-MODEL con score -312.79 y confidence de 0.9629, el receptor ALP de AlphaFold con score -346.35 y confidence de 0.9807 y el receptor humano EGFR con score -327.41 y confidence de 0.972, para este último hay un segundo receptor humano con un score cercano al de EGFR, siendo este el receptor TFRC con score -327.33 y el mismo valor de confidence. Para Cry1Aa, el valor de score para el docking con receptor de insecto fue más bajo con el receptor de AlphaFold, es muy probable que se deba a que el modelo de IA esté mejor resuelto y por la calidad según el pLLDT, además, el programa arroja al receptor APN1 modelado con AlphaFold como el mejor para unirse a Cry1Aa, el cual en la literatura se describe como uno de los receptores a los que se une esta proteína Cry. Con Cry1Ac sucede algo similar, el modelo de AlphaFold obtiene un score de docking menor al modelo de SWISS-MODEL, y además este receptor, el ALP, también está reportado como uno de los cuales se une Cry1Ac. Para el caso del receptor humano, EGFR y TFRC son los 2 mejores para ambas Cry, sus valores de score son muy cercanos y para el caso de Cry1Ac, esos valores son prácticamente los mismos, así que, se puede tener cierto grado de confianza dado el acierto con los receptores de insecto, y en el caso de Cry1Aa el mejor receptor fue TFRC.

En ClusPro, los mejores receptores son los siguientes:

Para Cry1Aa, el mejor receptor de insectos fue ALP, tanto de SWISS-MODEL como de AlphaFold, con score -1047.1 y tamaño de cluster de 41, y score de -1432.9 y cluster de 38, respectivamente (tabla 29), en cuanto al mejor receptor humano para esta Cry, fue el TFRC con score -1124.7 y cluster de 85 (tabla 30).

Para Cry1Ac, también el receptor ALP obtuvo mejor desempeño en docking, tanto el modelo de SWISS-MODEL como el de AlphaFold con scores de, -1073.4 y cluster de 54, y -1222 y cluster de 47, respectivamente (tabla 29), para el caso del receptor humano, el receptor con el mejor score fue EGFR, el valor de score fue -1063.5 y el cluster de 55 (tabla 30), y de igual forma a HDOCK, el segundo receptor humano es el TFRC con valores cercanos, score -1056.7 y cluster de 54. Con Cry1Aa el mejor receptor de insecto en este caso, es el receptor ALP, tanto del obtenido por homología como el de IA, sin embargo, los valores de score más bajos se dieron

con los modelos de IA (tabla 29), y en este caso no logra acertar con el receptor de literatura, pero sí al tener mejor desempeño con los modelos de IA los cuales tienen mejor calidad. Con Cry1Ac de igual manera el ALP es el mejor receptor para los 2 tipos de modelado, y en este caso, acierta al ser ALP el receptor reportado. Con receptores humanos existe cierta consistencia respecto al programa HDOCK, dado que se repiten ambos receptores (EGFR y TFRC) para Cry1Ac y también el receptor TFRC para Cry1Aa, existe consistencia en estos resultados incluso siendo algoritmos distintos.

Por último, con el software HEX los mejores receptores fueron los siguientes:

Para Cry1Aa, de los modelos por homología, el receptor ALP fue el de mejor resultado con una energía de -902.2 (Tabla 31) y en cuanto el modelo de AlphaFold, el único receptor que se pudo correr en el programa fue el APN2 con -590.8 de energía (tabla 31), más adelante se detallan los problemas relacionados; en el caso del receptor humano, el complejo con menos energía fue el que tenía a TFRC como receptor, con una energía de -856.7 (Tabla 32).

Para Cry1Ac, solo se obtuvo resultado del receptor de insecto obtenido por homología, siendo ALP con -871.1 de energía (Tabla 31), y para el caso del receptor humano, el mejor fue TFRC con -820.9 de energía (Tabla 32). El docking en general con este software, tuvo un bajo rendimiento en comparación con los otros programas, dado que este no ha recibido soporte en varios años, y por lo tanto, se encuentra un poco obsoleto en el sentido de que no se ha actualizado. Sin embargo, en su momento era una buena opción para realizar docking y con el equipo de cómputo utilizado el programa se ejecutaba en cuestión de segundos, no obstante, la ejecución del programa mostró errores con los modelos de AlphaFold, ya que corrompía el archivo PDB del complejo de docking resultante, dificultando así su visualización, por eso, los modelos mostrados en la sección de resultados son diferentes para HEX, debido a que no se podían visualizar con el mismo software; el único resultado del modelo de AlphaFold fue deficiente en comparación con la tendencia a que con estos modelos se mejoraba el score de docking con los otros dos programas, pero este único receptor fue el APN2, el cual está reportado que se une a Cry1Aa, para el caso del receptor por homología, el que el programa arrojó como el mejor para unirse a esta Cry fue el ALP. para el caso de Cry1Ac, el mejor receptor, el obtenido por homología, fue ALP, también reportado para unión a Cry1Ac. En cuanto a los receptores humanos, HEX mostró similitud en estos respecto a los otros 2 programas, teniendo al receptor TFRC como el mejor para ambas Cry, con energía de -856.7 para Cry1Aa y -820.9 para Cry1Ac. De manera general, se observó una mejora en el docking

de receptores de insecto usando los modelos de AlphaFold, los cuales, en comparación con los de SWISS-MODEL, son de mejor calidad, esto respaldado por las métricas de los dos tipos de modelado y también por los scores de docking de ambos tipos de modelado, ya que en HDOCK y ClusPro los scores son mejores (más bajos) para los modelos de AlphaFold, aunque, los modelos de homología obtuvieron buenos scores, ya que en HDOCK reciben un valor de confidence por encima de 0.9, lo cual indica que los complejos receptor-ligando son de una calidad alta, y por el lado de ClusPro, el tamaño de los cluster es similar entre receptores de homología y receptores de IA, lo que indica que, a pesar de que el desempeño mejora con el uso de los receptores de IA, los receptores de homología también arrojan buenos resultados. Un aspecto que destacar es, en los resultados, casi la mitad de los casos, el mejor complejo incluye al receptor que se une *in vivo* al organismo correspondiente, lo cual es una evidencia prometedora de la capacidad de predicción y precisión de los métodos de docking molecular, además, esto sugiere que los enfoques en docking pueden proporcionar buenas inferencias de la interacción receptor-ligando, sin olvidar que los algoritmos guían a que el complejo adquiera una conformación estable y favorable, lo cual no refleja la complejidad de las interacciones biomoleculares.

Entonces, por el lado de los receptores humanos, se tuvo una consistencia en los resultados a pesar de que se emplearon tres programas, con tres algoritmos y tres métodos distintos de realizar docking molecular, en todos los casos, el principal receptor implicado fue el TFRC, para el caso de Cry1Aa fue el top 1 con los tres programas, y con Cry1Ac quedó segundo en HDOCK y ClusPro por una mínima diferencia en el score con el receptor EGFR, así que, se puede decir que de manera general, que el receptor TFRC, tuvo el mejor desempeño en el docking con las dos proteínas Cry empleadas en este trabajo, seguido del receptor EGFR y AMPN. Se puede sugerir que estos resultados son confiables dado el rendimiento usando los receptores de insecto los cuales ya tienen descrito mecanismo de acción y por lo tanto, de unión, además en la tabla 34 se menciona las coincidencias entre los sitios de unión descritos en literatura para el ligando y los receptores, de los cuales, 3 complejos están unidos en estos sitios descritos; dada la similitud estructural y funcional de las proteínas Cry con las parasporinas, los receptores humanos sobreexpresados en MCF-7 que se usaron y, obtuvieron el mejor rendimiento en el docking, se pueden proponer como receptores para las parasporinas, incluyendo las que se emplean en el grupo de trabajo, además, con el análisis de las regiones de receptores que interactúan con Cry, se podría esperar que, al replicar estas pruebas con parasporinas como ligando, los residuos prevalentes obtenidos en este trabajo, pudieran estar

involucrados en la unión a PS. Aunque, por supuesto, el siguiente paso sería realizar las pruebas *in silico* con dichas parasporinas y posterior a eso, realizar estudios *in vivo* para validar las pruebas de docking molecular.

10. CONCLUSIONES

En este trabajo se emplearon dos modelos de proteínas Cry (Cry1Aa y Cry1Ac) de tres dominios cuyas funciones son similares de las parasporinas las cuales poseen tres dominios también. Para estas proteínas Cry, se encontró interacción receptor-ligando usando los tres programas, con dos receptores sobreexpresados en la línea MCF-7, siendo estos los receptores TFRC y EGFR, también se obtuvo una lista de residuos de estos receptores que están implicados en estas interacciones y que se repiten en las distintas pruebas de docking, lo cual sugiere que existen regiones de aminoácidos específicas y/o favorables para la interacción con parasporinas.

Como perspectivas de este trabajo, se deben de realizar pruebas *in vivo*, como ensayos de unión que permitan cuantificar la afinidad y cinética de unión entre las moléculas, una estrategia puede ser el bloqueo de los receptores para evaluar cómo afectaría la unión con los ligandos.

11. BIBLIOGRAFÍA

1. Agarwal, S., & Mehrotra, R. (2016). An overview of Molecular Docking. *JSM Chemistry*, 4(2), 1024. <https://www.jscimedcentral.com/public/assets/articles/chemistry-4-1024.pdf>
2. Aisen, P. (2004). Transferrin receptor 1. *The International Journal of Biochemistry & Cell Biology*, 36(11), 2137-2143. <https://doi.org/10.1016/j.biocel.2004.02.007>
3. Akiba, T., Abe, Y., Kitada, S., Kusaka, Y., Ito, A., Ichimatsu, T., Katayama, H., Akao, T., Higuchi, K., Mizuki, E., Ohba, M., Kanai, R., & Harata, K. (2009). Crystal Structure of the Parasporin-2 *Bacillus thuringiensis* Toxin That Recognizes Cancer Cells. *Journal of Molecular Biology*, 386(1), 121-133. <https://doi.org/10.1016/j.jmb.2008.12.002>
4. Ammons, D. R., Short, J. D., Bailey, J., Hinojosa, G., Tavarez, L., Salazar, M., & Rampersad, J. N. (2015). Anti-cancer Parasporin Toxins are Associated with Different Environments: Discovery of Two Novel Parasporin 5-like Genes. *Current Microbiology*, 72(2), 184-189. <https://doi.org/10.1007/s00284-015-0934-3>
5. Assaraf, Y. G., Leamon, C. P., & Reddy, J. A. (2014). The folate receptor as a rational therapeutic target for personalized cancer treatment. *Drug Resistance Updates*, 17(4-6), 89-95. <https://doi.org/10.1016/j.drug.2014.10.002>
6. Baselga, J., & Swain, S. M. (2009). Novel anticancer targets: revisiting ERBB2 and discovering ERBB3. *Nature Reviews Cancer*, 9(7), 463-475. <https://doi.org/10.1038/nrc2656>
7. Biasini, M., Bienert, S., Waterhouse, A. L., Arnold, K., Studer, G., Schmidt, T., Kiefer, F. W., Cassarino, T. G., Bertoni, M., Bordoli, L., & Schwede, T. (2014). SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Research*, 42(W1), W252-W258. <https://doi.org/10.1093/nar/gku340>
8. Bordoli, L., Kiefer, F., Arnold, K., Benkert, P., Battey, J., & Schwede, T. (2008). Protein structure homology modeling using SWISS-MODEL workspace. *Nature Protocols*, 4(1), 1-13. <https://doi.org/10.1038/nprot.2008.197>
9. Chen, V. C., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R., Kapral, G. J., Murray, L. L., Richardson, J. S., & Richardson, D. J. (2010). *MolProbity*: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D-biological Crystallography*, 66(1), 12-21. <https://doi.org/10.1107/s0907444909042073>

10. Cristobal, S., Zemla, A., Fischer, D. A., Rychlewski, L., & Elofsson, A. (2001). A study of quality measures for protein threading models. *BMC Bioinformatics*, 2(1), 5. <https://doi.org/10.1186/1471-2105-2-5>
11. de Maagd, R. (2001). How *Bacillus thuringiensis* has evolved specific toxins to colonize the insect world. *Trends in Genetics*, 17(4), 193-199. [https://doi.org/10.1016/s0168-9525\(01\)02237-5](https://doi.org/10.1016/s0168-9525(01)02237-5)
12. Dias, R., & de Azevedo Jr., W. (2008). Molecular Docking Algorithms. *Current Drug Targets*, 9(12), 1040-1047. <https://doi.org/10.2174/138945008786949432>
13. Ekino, K., Okumura, S., Ishikawa, T., Kitada, S., Saitoh, H., Akao, T., Oka, T., Nomura, Y., Ohba, M., Shin, T., & Mizuki, E. (2014). Cloning and Characterization of a Unique Cytotoxic Protein Parasporin-5 Produced by *Bacillus thuringiensis* A1100 Strain. *Toxins*, 6(6), 1882-1895. <https://doi.org/10.3390/toxins6061882>
14. Ferrara, N., & Kerbel, R. S. (2005). Angiogenesis as a therapeutic target. *Nature*, 438(7070), 967-974. <https://doi.org/10.1038/nature04483>
15. Ferreira, L., dos Santos, R., Oliva, G., & Andricopulo, A. (2015). Molecular Docking and Structure-Based Drug Design Strategies. *Molecules*, 20(7), 13384-13421. <https://doi.org/10.3390/molecules200713384>
16. Hoppenz, P., Els-Heindl, S., & Beck-Sickinger, A. G. (2020). Peptide-Drug Conjugates and Their Targets in Advanced Cancer Therapies. *Frontiers in Chemistry*, 8. <https://doi.org/10.3389/fchem.2020.00571>
17. Huang, S. (2018). Comprehensive assessment of flexible-ligand docking algorithms: current effectiveness and challenges. *Briefings in Bioinformatics*, 19(5), 982-994. <https://doi.org/10.1093/bib/bbx030>
18. Ito, A., Sasaguri, Y., Kitada, S., Kusaka, Y., Kuwano, K., Masutomi, K., Mizuki, E., Akao, T., & Ohba, M. (2004). A *Bacillus thuringiensis* Crystal Protein with Selective Cytocidal Action to Human Cells. *Journal of Biological Chemistry*, 279(20), 21282-21286. <https://doi.org/10.1074/jbc.m401881200>
19. Jisna, V. A., & Jayaraj, P. B. (2021). Protein Structure Prediction: Conventional and Deep Learning Perspectives. *The Protein Journal*, 40(4), 522-544. <https://doi.org/10.1007/s10930-021-10003-y>
20. Katayama, H., Kusaka, Y., Yokota, H., Akao, T., Kojima, M., Nakamura, O., Mekada, E., & Mizuki, E. (2007). Parasporin-1, a Novel Cytotoxic Protein from *Bacillus thuringiensis*, Induces Ca²⁺ Influx and a Sustained Elevation of the Cytoplasmic Ca²⁺ Concentration

- in Toxin-sensitive Cells. *Journal of Biological Chemistry*, 282(10), 7742-7752.
<https://doi.org/10.1074/jbc.m611382200>
21. Kelley, L. A., Mezulis, S., Yates, C. J., Wass, M. N., & Sternberg, M. J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols*, 10(6), 845-858. <https://doi.org/10.1038/nprot.2015.053>
22. Kitada, S., Abe, Y., Shimada, H., Kusaka, Y., Matsuo, Y., Katayama, H., Okumura, S., Akao, T., Mizuki, E., Kuge, O., Sasaguri, Y., Ohba, M., & Ito, A. (2006). Cytocidal Actions of Parasporin-2, an Anti-tumor Crystal Toxin from *Bacillus thuringiensis*. *Journal of Biological Chemistry*, 281(36), 26350-26360. <https://doi.org/10.1074/jbc.m602589200>
23. Kleven, M. S., Jue, S., & Enns, C. A. (2018). Transferrin Receptors TfR1 and TfR2 Bind Transferrin through Differing Mechanisms. *Biochemistry*, 57(9), 1552-1559. <https://doi.org/10.1021/acs.biochem.8b00006>
24. Kunjiappan, S., Parasuraman, P., Vellaichamy, S., Pandian, S. R. K., Ravishankar, V., Palanisamy, P., Govindaraj, S., Srinivasan, G., Premanand, A., Murugesan, S., & Panneerselvam, T. (2021). Surface receptor-mediated targeted drug delivery systems for enhanced cancer treatment: A state-of-the-art review. *Drug Development Research*, 82(3), 309-340. <https://doi.org/10.1002/ddr.21758>
25. Lanzaro, G. C., & Lee, Y. (2013). Speciation in *Anopheles gambiae* — The Distribution of Genetic Polymorphism and Patterns of Reproductive Isolation Among Natural Populations. En *InTech eBooks*. <https://doi.org/10.5772/56232>
26. Large, D. E., Soucy, J. R., Hebert, J. M., & Auguste, D. T. (2019). Advances in Receptor-Mediated, Tumor-Targeted Drug Delivery. *Advanced therapeutics*, 2(1), 1800091. <https://doi.org/10.1002/adtp.201800091>
27. Lee, A. V., Oesterreich, S., & Davidson, N. E. (2015). MCF-7 Cells--Changing the Course of Breast Cancer Research and Care for 45 Years. *Journal of the National Cancer Institute*, 107(7), djv073. <https://doi.org/10.1093/jnci/djv073>
28. Lenina, N. K., Naveenkumar, A., Sozhavendan, A. E., Balakrishnan, N., Balasubramani, V., & Udayasuriyan, V. (2013). Characterization of parasporin gene harboring Indian isolates of *Bacillus thuringiensis*. *3 Biotech*, 4(5), 545-551. <https://doi.org/10.1007/s13205-013-0190-9>
29. Liu, W., Ji, Y., Sun, Y., Si, L., Fu, J., Hayashi, T., Onodera, S., & Ikejima, T. (2020). Estrogen receptors participate in silibinin-caused nuclear translocation of apoptosis-inducing

- factor in human breast cancer MCF-7 cells. *Archives of Biochemistry and Biophysics*, 689, 108458. <https://doi.org/10.1016/j.abb.2020.108458>
30. Melo, A. L. D. A., Soccol, V. T., & Soccol, C. R. (2014). *Bacillus thuringiensis*: mechanism of action, resistance, and new applications: a review. *Critical Reviews in Biotechnology*, 36(2), 317-326. <https://doi.org/10.3109/07388551.2014.960793>
 31. Melo, A. L., & Kitada, S. (2020). Selection of the *Bacillus thuringiensis* Berliner strain to produce a parasporin with cytotoxic activity against MCF-7 breast cancer cells. *Breast Disease*, 39(1), 37-42. <https://doi.org/10.3233/bd-190405>
 32. Mendoza-Almanza, G., Esparza-Ibarra, E. L., Ayala-Lujan, J. L., Mercado-Reyes, M., Godina-González, S., Hernández-Barrales, M., & Olmos-Soto, J. (2020). The Cytocidal Spectrum of *Bacillus thuringiensis* Toxins: From Insects to Human Cancer Cells. *Toxins*, 12(5), 301. <https://doi.org/10.3390/toxins12050301>
 33. Michio Ohba, Eiichi Mizuki, & Akiko Uemori. (2008). Parasporin, a new anticancer protein group from *Bacillus thuringiensis*. *Anticancer Research*, 29(1), 427-433.
 34. Mizuki, E., Ohba, M., Akao, T., Yamashita, S., Saitoh, H., & Park, Y. S. (1999). Unique activity associated with non-insecticidal *Bacillus thuringiensis* parasporal inclusions: in vitro cell-killing action on human cancer cells. *Journal of Applied Microbiology*, 86(3), 477-486. <https://doi.org/10.1046/j.1365-2672.1999.00692.x>
 35. Nagamatsu, Y., Okamura, S., Saitou, H., Akao, T., & Mizuki, E. (2010). Three Cry Toxins in Two Types from *Bacillus thuringiensis* Strain M019 Preferentially Kill Human Hepatocyte Cancer and Uterus Cervix Cancer Cells. *Bioscience, Biotechnology, and Biochemistry*, 74(3), 494-498. <https://doi.org/10.1271/bbb.90615>
 36. Nakanishi, K., Yaoi, K., Nagino, Y., Hara, H., Kitami, M., Atsumi, S., Miura, N., & Sato, R. (2002). Aminopeptidase N isoforms from the midgut of *Bombyx mori* and *Plutella xylostella*- their classification and the factors that determine their binding specificity to *Bacillus thuringiensis* Cry1A toxin. *FEBS Letters*, 519(1-3), 215-220. [https://doi.org/10.1016/s0014-5793\(02\)02708-4](https://doi.org/10.1016/s0014-5793(02)02708-4)
 37. Ogiso, H., Ishitani, R., Nureki, O., Fukai, S., Yamanaka, M., Kim, J. N., Saito, K., Sakamoto, A., Inoue, M., Shirouzu, M., & Yokoyama, S. (2002). Crystal Structure of the Complex of Human Epidermal Growth Factor and Receptor Extracellular Domains. *Cell*, 110(6), 775-787. [https://doi.org/10.1016/s0092-8674\(02\)00963-7](https://doi.org/10.1016/s0092-8674(02)00963-7)

38. Okassov, A., Nersesyan, A., Kitada, S., & Ilin, A. (2014). Parasporins as new natural anticancer agents: a review. *Journal of B.U.ON. : official journal of the Balkan Union of Oncology*, 20(1), 5-16.
39. Okumura, S., Ishikawa, T., Saitoh, H., Akao, T., & Mizuki, E. (2013). Identification of a second cytotoxic protein produced by *Bacillus thuringiensis* A1470. *Biotechnology Letters*, 35(11), 1889-1894. <https://doi.org/10.1007/s10529-013-1275-6>
40. Okumura, S., Saitoh, H., Ishikawa, T., Mizuki, E., & Inouye, K. (2008). Identification and characterization of a novel cytotoxic protein, parasporin-4, produced by *Bacillus thuringiensis* A1470 strain. *Biotechnology Annual Review*, 225-252. [https://doi.org/10.1016/s1387-2656\(08\)00009-4](https://doi.org/10.1016/s1387-2656(08)00009-4)
41. Palma, L., Muñoz, D., Berry, C., Murillo, J., & Caballero, P. (2014). *Bacillus thuringiensis* Toxins: An Overview of Their Biocidal Activity. *Toxins*, 6(12), 3296-3325. <https://doi.org/10.3390/toxins6123296>
42. Pigott, C. R., & Ellar, D. J. (2007). Role of Receptors in *Bacillus thuringiensis* Crystal Toxin Activity. *Microbiology and Molecular Biology Reviews*, 71(2), 255-281. <https://doi.org/10.1128/membr.00034-06>
43. Santos, E. N., Menezes, L. O., Dolabella, S. S., Souto, E. B., Severino, P., Capasso, R., Zielińska, A., Souto, E. B., & Jain, S. (2022). *Bacillus thuringiensis*: From biopesticides to anticancer agents. *Biochimie*, 192, 83-90. <https://doi.org/10.1016/j.biochi.2021.10.003>
44. Schwede, T. (2013). Protein Modeling: What Happened to the “Protein Structure Gap”? *Structure*, 21(9), 1531-1540. <https://doi.org/10.1016/j.str.2013.08.007>
45. Siew, N., Elofsson, A., Rychlewski, L., & Fischer, D. A. (2000). MaxSub: an automated measure for the assessment of protein structure prediction quality. *Bioinformatics*, 16(9), 776-785. <https://doi.org/10.1093/bioinformatics/16.9.776>
46. Soberón, M., Pardo, L., Muñoz-Garay, C., Sánchez, J. L., Gómez, I. M., Porta, H., & Bravo, A. (2010). Pore formation by Cry toxins. En *Advances in Experimental Medicine and Biology* (pp. 127-142). Springer Nature. https://doi.org/10.1007/978-1-4419-6327-7_11
47. Soleymani, J., Hasanzadeh, M., Somi, M. H., Shadjou, N., & Jouyban, A. (2018). Probing the specific binding of folic acid to folate receptor using amino-functionalized mesoporous silica nanoparticles for differentiation of MCF 7 tumoral cells from MCF

- 10A. *Biosensors and Bioelectronics*, *115*, 61-69.
<https://doi.org/10.1016/j.bios.2018.05.025>
48. Stanzione, F., Giangreco, I., & Cole, J. C. (2021). Use of molecular docking computational tools in drug discovery. *Progress in Medicinal Chemistry*, 273-343.
<https://doi.org/10.1016/bs.pmch.2021.01.004>
49. Studer, G., Rempfer, C., Waterhouse, A. L., Gumienny, R., Haas, J., & Schwede, T. (2020). QMEANDisCo—distance constraints applied on model quality estimation. *Bioinformatics*, *36*(6), 1765-1771. <https://doi.org/10.1093/bioinformatics/btz828>
50. Suarez-Barrera, M. O., Visser, L., Pinzon-Reyes, E., Rondón-Villarreal, P., Alarcón-Aldana, J. S., & Rueda-Forero, N. J. (2022). Site-Directed Mutants of Parasporin PS2Aa1 with Enhanced Cytotoxic Activity in Colorectal Cancer Cell Lines. *Molecules*, *27*(21), 7262. <https://doi.org/10.3390/molecules27217262>
51. Tao, X., Huang, Y., Wang, C., Chen, F., Yang, L., Ling, L., Che, Z., & Chen, X. (2019). Recent developments in molecular docking technology applied in food science: a review. *International Journal of Food Science & Technology*, *55*(1), 33-45.
<https://doi.org/10.1111/ijfs.14325>
52. Uings, I., & Farrow, S. N. (2000). Cell receptors and cell signalling. *Journal of Clinical Pathology-molecular Pathology*, *53*(6), 295-299. <https://doi.org/10.1136/mp.53.6.295>
53. Vakser, I. A. (2014). Protein-Protein Docking: From Interaction to Interactome. *Biophysical Journal*, *107*(8), 1785-1793. <https://doi.org/10.1016/j.bpj.2014.08.033>
54. Varadi, M., Anyango, S., Deshpande, M. S., Nair, S., Natassia, C., Yordanova, G., Yuan, D. C., Stroe, O., Wood, G., Laydon, A., Žídek, A., Green, T. J., Tunyasuvunakool, K., Petersen, S., Jumper, J. M., Clancy, E., Green, R., Vora, A., Lutfi, M., . . . Velankar, S. (2021). AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Research*, *50*(D1), D439-D444. <https://doi.org/10.1093/nar/gkab1061>
55. Velásquez Cardona, L. F., Rojas Torres, D. S., & Cerón Salamanca, J. (2018). TOXINAS DE *Bacillus thuringiensis* CON ACTIVIDAD ANTICANCERÍGENA: PARASPORINAS. *Revista Colombiana de Biotecnología*, *20*(2), 89-100.
<https://doi.org/10.15446/rev.colomb.biote.v20n2.73668>
56. Wong, R. S. Y. (2010). *Bacillus Thuringiensis* Parasporal Proteins And Their Effect On Human Cancer Cells : An Overview. *International e-Journal of Science, Medicine and Education*, *4*(2), 3-9. <https://doi.org/10.56026/imu.4.2.3>

57. Wong, R. S. Y., Mohamed, S. M., Nadarajah, V. D., & Tengku, I. A. (2010). Characterisation of the Binding Properties of *Bacillus Thuringiensis* 18 Toxin on Leukaemic Cells. *Journal of Experimental & Clinical Cancer Research*, 29(1). <https://doi.org/10.1186/1756-9966-29-86>
58. Xu, C., Wang, B. C., Yu, Z., & Sun, M. (2014). Structural Insights into *Bacillus thuringiensis* Cry, Cyt and Parasporin Toxins. *Toxins*, 6(9), 2732-2770. <https://doi.org/10.3390/toxins6092732>
59. Yamashita, S., Katayama, H., Saitoh, H., Akao, T., Park, Y. S., Mizuki, E., Ohba, M., & Ito, A. (2005). Typical Three-Domain Cry Proteins of *Bacillus thuringiensis* Strain A1462 Exhibit Cytocidal Activity on Limited Human Cancer Cells. *The Journal of Biochemistry*, 138(6), 663-672. <https://doi.org/10.1093/jb/mvi177>