UNIVERSIDAD TÉCNICA DE AMBATO

FACULTAD DE CIENCIA E INGENIERÍA EN ALIMENTOS Y BIOTECNOLOGÍA

CARRERA DE BIOTECNOLOGÍA

Predicción de proteínas *de novo* mediante inteligencia artificial para la inhibición del factor NF-κB en el cáncer gástrico

Informe Final del Trabajo de Titulación, opción Proyecto de Investigación, previo a la obtención del Título de Ingeniero Biotecnólogo, otorgado por la Universidad Técnica de Ambato, a través de la Facultad de Ciencia e Ingeniería en Alimentos y Biotecnología.

Autor: Diego Martin Laverde Lomas

Tutor: Mg. Cristian Fernando Galarza Galarza

Ambato – Ecuador

Agosto - 2024

APROBACIÓN DEL TUTOR

Mg. Cristian Fernando Galarza Galarza

CERTIFICA:

Que el presente Informe Final del Trabajo de Titulación ha sido prolijamente revisado. Por lo tanto, autorizo la presentación de este Informe Final del Trabajo de Titulación, opción Proyecto de Investigación, el mismo que responde a las normas establecidas en el reglamento de Títulos y Grados de la Facultad de Ciencia e Ingeniería en Alimentos y Biotecnología.

Ambato, 27 de Junio de 2024 .

Firmado electrónicamente por:
CRISTIAN FERNANDO
GALARZA GALARZA

Mg. Cristian Fernando Galarza Galarza

1803160272

TUTOR

AUTORÍA DEL TRABAJO DE TITULACIÓN

Yo, Diego Martín Laverde Lomas, manifiesto que los resultados obtenidos en el presente Informe Final del Trabajo de Titulación, opción Proyecto de Investigación, previo a la obtención del título de Ingeniero Biotecnólogo son absolutamente originales, auténticos y personales: a excepción de las citas bibliográficas.



Diego Martin Laverde Lomas

0503853525

AUTOR

DERECHOS DE AUTOR

Autorizo a la Universidad Técnica de Ambato, para que haga de este Informe Final del Trabajo de Titulación o parte de él, un documento disponible para su lectura consulta y procesos de investigación, según las normas de la Institución .

Cedo los Derechos en línea patrimoniales de mi Informe Final del Trabajo de Titulación, con fines de difusión pública, además apruebo la reproducción de este, dentro de las regulaciones de la Universidad, siempre y cuando esta reproducción no suponga una ganancia económica y se realice respetando mis derechos de autor.



Diego Martin Laverde Lomas

0503853525

AUTOR

APROBACIÓN DEL TRIBUNAL DE GRADO

Los suscritos Profesores Calificadores, aprueban el presente Informe Final del Trabajo de Titulación, opción Proyecto de Investigación, el mismo que ha sido elaborado de conformidad con las disposiciones emitidas por la Facultad de Ciencia e Ingeniería en Alimentos y Biotecnología de la Universidad Técnica de Ambato.

Para constancia firman:

Firmado electrónicamente por:
WILLIAM RICARDO CALERO CACERES

Presidente del Tribunal

Firmado electrónicamente por:
PABLO VINICIO TUZA ALVARADO

Dr. Pablo Vinicio Tuza Alvarado

1104063241

Firmado electrónicamente por:
LORENA DE LOS ANGELES NUNEZ VILLACIS

Dra. Lorena de los Ángeles Núñez Villacís

1804256905

Ambato, 02 de Agosto de 2024

DEDICATORIA

Dedico este trabajo

A aquella persona que ya no está a mi lado,

A quienes se quedaron

Y a los que se han alejado.

Dedico este trabajo, principalmente a mi familia

Y en especial a mi madre, quien me dio la vida.

A "Zule", "Cindy" y "Lenin" que me hicieron reír día tras día,

Y con quienes compartí miles de alegrías.

Dedico este trabajo, a mis hermanas "Milagros" y "Ana Paula"

Con las que convivo semana a semana.

A mis sueños, metas y esperanzas

Que son mi razón para despertar cada mañana.

Dedico este trabajo, a mi yo del pasado,

A mi presente y mi futuro añorado.

Quiero que recuerden lo mencionado

Yo, ¡lo he logrado!

- Diego Martin Laverde Lomas -

Trabaja duro, divierte y haz historia

- Jeff Bezos -

AGRADECIMIENTO

Agradezco las buenas y malas experiencias que me permitieron apreciar la vida que me ha tocado vivir.

Agradezco al tiempo y a las oportunidades de la vida, a pesar de que nunca se detienen ni esperan a nadie, siempre llega en el momento exacto.

Agradezco a mi madre, por apoyarme en los momentos de más agobio, estrés y tristeza he sentido.

Agradezco a cada uno de los docentes que conocí durante mi vida universitaria y a los conocimientos que he adquirido de ellos.

Agradezco a mi tutor M.Sc. Cristian Galarza, por impulsarme y ayudarme durante el desarrollo de esta investigación.

Agradezco a mi computador, que ha sido mi más fiel herramienta de estudio y trabajo a lo largo de todo este proceso.

ÍNDICE GENERAL DE CONTENIDOS

APROBACIÓN DEL TUTOR....ii

AUTORÍA DEL TRABAJO DE TITULACIÓN.... iii

DERECHOS DE AUTOR....iv

APROBACIÓN DEL TRIBUNAL DE GRADO v

DEDICATORIA....vi

AGRADECIMIENTO vii

ÍNDICE DE TABLAS xi

ÍNDICE DE FIGURAS....xii

RESUMEN EJECUTIVO.... xiii

ABSTRACT....xiv

CAPITULO 1....1

MARCO TEÓRICO....1

1.1 Antecedentes Investigativos....1

1.1.1. Justificación 1

1.1.2. Cáncer Gástrico....2

1.1.3. Factor NF-κB 4

1.1.4. Proteínas *de novo* 8

1.1.5. Inteligencia artificial dentro de la Biomedicina....10

1.2. Objetivos....12

1.2.1. Objetivo General....12

1.2.2. Objetivos Específicos....12

CAPITULO 2....13

METODOLOGÍA....13

2.1. Materiales....13

2.2. Métodos....14

2.2.1. Identificación de proteínas *de novo* estables relacionadas con la inhibición del factor NF-κB en el cáncer gástrico utilizando herramientas de inteligencia artificial.....14

2.2.2. Comparación de las variables termodinámicas obtenidas de la interacción entre el factor NF-κB y las proteínas relacionadas. 17

2.2.3. Evaluación de la afinidad y selectividad de las proteínas predichas durante la inhibición del factor propuesto. 18

CAPITULO 3....19

RESULTADOS Y DISCUSIÓN 19

3.1. Análisis y discusión de los resultados....19

3.1.1. Búsqueda de Secuencias y Predicción de proteínas *de novo* 19

3.1.2. Visualización de las proteínas predichas....22

3.1.3. Análisis de estabilidad de las proteínas predichas 26

3.1.4. Comparación de las variables termodinámicas presentes en la interacción proteína-proteína....30

3.1.5. Afinidad y selectividad de las proteínas predichas 32

CAPITULO 4....35

CONCLUSIONES Y RECOMENDACIONES....35

4.1. Conclusiones....35

4.2. Recomendaciones 36

REFERENCIAS BIBLIOGRÁFICAS.....37

ANEXOS 52

ÍNDICE DE TABLAS

Tabla 1. Polimorfismos presentes en genes de NF-κB y en moléculas de señalización. 5

Tabla 2. Genes regulados por NF-κB vinculados con el desarrollo de cáncer gástrico. 7

Tabla 3. Recursos Bioinformáticos.....13

Tabla 4. Equipos.....13

Tabla 5. Genes regulados y factores de transcripción relacionados con NF- κB en el cáncer gástrico 19

Tabla 6. Proteínas generadas por las herramientas de Inteligencia Artificial en base a inhibidores del factor NF- κB.....20

Tabla 7. Composición y Estructura de las proteínas predichas en las herramientas de Inteligencia artificial.....22

Tabla 8. Datos y propiedades energéticas de las proteínas predichas por las herramientas de Inteligencia Artificial. 27

Tabla 9. Datos acerca de las variables termodinámicas obtenidas tras el proceso de Acoplamiento Molecular. 30

Tabla 10. Datos acerca de la afinidad de las proteínas predichas (ligandos).....32

Tabla 11. Identificación de la selectividad de las proteínas predichas (ligandos) 34

ÍNDICE DE FIGURAS

Figura 1. Composición y estructura de la proteína Q15653. 22

Figura 2. Composición y estructura de la proteína Q8NI38. 23

Figura 3. Composición y estructura de la proteína Q9BYH8. 23

Figura 4. Composición y estructura de la proteína A26C8. 24

Figura 5. Composición y estructura de la proteína 208E3. 24

Figura 6. Composición y estructura de la proteína 9061D. 24

Figura 7. Composición y estructura de la proteína RF715A1. 25

Figura 8. Composición y estructura de la proteína RF315AD. 25

Figura 9. Composición y estructura de la proteína RF2M31. 25

Figura 10. Gráficas RMSD de estabilidad de las proteínas predichas. 29

Figura 11. Acoplamiento obtenido de la interacción entre las proteínas control. 32

Figura 12. Acoplamiento obtenido de la interacción entre NIK y Q9BYH8. 32

Figura 13. Acoplamiento obtenido de la interacción entre la subunidad p100 y Q9BYH8. 33

Figura 14. Acoplamiento obtenido de la interacción entre IL-8 y 10FAD. 33

RESUMEN EJECUTIVO

En la presente investigación, se destaca la limitada capacidad de los actuales tratamientos contra el cáncer gástrico junto al rol de las herramientas de Inteligencia Artificial en el desarrollo de nuevas estrategias personalizadas, a través de la predicción de proteínas *de novo* destinadas a la inhibición del Factor Nuclear de las cadenas ligeras kappa de las células B activas (NF-κB) asociado con este tipo de cáncer. Se predijeron 30 proteínas con una composición similar a las que se encuentran almacenadas en bases de datos de tipo estándar y experimental. Se analizó su estabilidad y capacidad de plegamiento a partir del análisis de las energías provenientes de interacciones intra e intermoleculares mediante simulaciones de dinámica molecular. Así mismo, se efectuó un proceso de acoplamiento molecular entre varios genes o factores de transcripción regulados por NF-κB y las proteínas predichas, se calculó y comparó diferentes variables termodinámicas tales como: Energía Libre de Gibbs, Constante de disociación, Entalpía, Capacidad Calorífica y Entropía con respecto al complejo del Inhibidor Alpha de kappa B y la proteína compuesta por las subunidades p65 y p50 de NF-κB para verificar la afinidad de su interacción proteína-proteína y su estado de conformación estructural. Finalmente, se evaluó la afinidad y selectividad en las interacciones resultantes, concluyendo que las proteínas *de novo* predichas por inteligencia artificial son viables para la generación de nuevos tratamientos y el diseño de nuevos fármacos para tratar el cáncer gástrico.

Palabras clave: NF-κB, cáncer gástrico, proteínas *de novo*, inteligencia artificial, acoplamiento molecular, interacciones proteína-proteína

ABSTRACT

In this research, highlights the limited effectiveness of current treatments for gastric cancer and the role of AI tools in the development of new personalized strategies, through the prediction of *de novo* proteins aimed at inhibiting the Nuclear Factor kappa-light-chain-enhancer of activated B cells (NF-κB) associated with this type of cancer. Thirty proteins were predicted, with a similar composition to those stored in standard and experimental databases. Their stability and folding capacity were analyzed based on the energies from intra and intermolecular interactions by molecular dynamics simulations. Furthermore, a molecular docking process was performed between several genes or transcription factors regulated by NF-κB and the predicted proteins, different thermodynamic variables such as Gibbs Free Energy, Dissociation Constant, Enthalpy, Heat Capacity, and Entropy were calculated and compared with respect to the complex of Inhibitor Alpha of kappa B and the protein composed of the p65 and p50 subunits of NF-κB to verify the affinity of their protein-protein interaction and structural conformation state. Finally, the affinity and selectivity of the resulting interactions were evaluated, concluding that the *de novo* proteins predicted by artificial intelligence are viable for the generation of new treatments and the design of new drugs to treat gastric cancer.

Keywords: NF-κB, gastric cancer, *de novo* proteins, artificial intelligence, molecular docking, protein-protein interactions

CAPITULO 1

MARCO TEÓRICO

1.1 Antecedentes Investigativos

1.1.1. Justificación

A pesar de los esfuerzos de la comunidad médica, el cáncer se ha convertido en una de las principales causas de muerte a nivel global, provocando alrededor de 10 millones de defunciones en el año 2020 **(World Health Organization [WHO], 2022)**, incluso cuando su tasa de mortalidad disminuyó entre 1.4 a 1.8 % durante periodos anteriores **(Ward et al., 2019)**. El cáncer gástrico es considerado como uno de los cinco tipos más comunes de cáncer, debido a que cada año se diagnostica a más de un millón de personas con esta enfermedad y al incremento de su incidencia en poblaciones menores de 40 años dentro de países europeos, asiáticos y sudamericanos, por esta razón, se ha posicionado como el tercer tipo de cáncer con mayor mortalidad en el mundo **(Rawla & Barsouk, 2019; Wong et al., 2021).** Este cáncer registra una tasa de supervivencia del 30% durante los primeros cincos años tras su diagnóstico, no obstante, dicho porcentaje puede aumentar si los tratamientos empiezan en una etapa temprana del cáncer diagnosticado **(Eusebi et al., 2020)**.

Entre las técnicas para su tratamiento se encuentras: la radioterapia y la quimioterapia que son las más comunes durante el desarrollo de ensayos clínicos, sin embargo, este tipo de técnicas se encuentran limitadas por su incapacidad para actuar en secciones específicas **(Johnston & Beckman, 2019; Joshi & Badgwell, 2021)**, en donde, a pesar de que es habitual contrarrestar la acción tumoral de dicha enfermedad aún no se demuestran avances significativos en su eliminación definitiva **(Petryszyn et al., 2020)**. Por tal motivo, la implementación de alternativas como la terapia proteica enfocada al diseño de proteínas *de novo* se presenta como una posible estrategia para el tratamiento del cáncer debido a la facilidad para generar secuencias proteicas altamente estables y destinarlas a la interacción o unión con receptores de células cancerígenas específicas **(Kintzing et al., 2016; Kuhlman & Bradley, 2019),** todo esto añadido al uso de la inteligencia artificial como parte de la predicción y verificación del diseño de dichas estructuras, para el desarrollo de nuevas estrategias de tratamiento de cualquier tipo de cáncer y el desarrollo de nuevos fármacos dirigidos **(Huang & Li, 2023)**.

Finalmente, en la presente investigación se plantea un estudio de factibilidad acerca del uso de herramientas de inteligencia artificial para la predicción de proteínas *de novo* destinadas a la

inhibición específica del factor NF-κB presente en el cáncer gástrico con la finalidad de establecer un punto de partida en investigaciones relacionadas con la comprensión de estructuras proteicas, inteligencia artificial en la biomedicina y desarrollo de nuevos fármacos para tratar la enfermedad.

1.1.2. Cáncer Gástrico

El cáncer gástrico probablemente sea una de las enfermedades más antiguas reconocidas por la sociedad humana incluso llegando a presentar varios registros dentro de distintas civilizaciones, sin embargo, no fue considerada como tal hasta el descubrimiento de su fisiopatología (alteraciones en la mucosa gástrica y desarrollo de ulceras duodenales) durante la época del Renacimiento **(Topi et al., 2020)**, en donde, gran parte de los conocimientos actuales fueron desarrollados. Hoy en día, se la considera la neoplasia más recurrente en el estómago y el tercer tipo de cáncer con mayor mortalidad (1 de cada 12 muertes) por debajo del cáncer de pulmón y colorrectal, además, suele ser más frecuente en hombres (1 de cada 54) que en mujeres (1 de cada 126) de regiones rurales de Asia Oriental, Europa Oriental y América Latina, siendo esta última región donde se presentan la mayor parte de los pacientes jóvenes **(M. J. López et al., 2023; Smyth et al., 2020; Thrift & El-Serag, 2020)**, debido en gran parte a la gran cantidad de factores asociados a la misma, tales como: la edad, el sexo, el alto consumo de sales, las inflamaciones crónicas intestinales (gastritis), las infecciones por *Helicobacter pylori*, la virulencia bacteriana, los polimorfismos genéticos y los factores ambientales **(Lyons et al., 2019; Sepulveda, 2013; L. Yang et al., 2020)**.

1.1.2.1. Clasificación del cáncer gástrico

Los sistemas de clasificación empleados durante el último siglo no son completamente reproducibles conforme los estándares actuales, debido a que se basan principalmente en las características macroscópicas y microscópicas presentes en el cáncer gástrico, en donde, la presencia de numerosas características histológicas en una misma región dificulta su interpretación **(Agnes et al., 2020)**. Por tal razón, recientemente se ha implementado varias clasificaciones que toman en cuenta múltiples características de dicha enfermedad:

- **Clasificación de Borrmann:** Implementada y estudiada ampliamente dentro de países asiáticos y orientales, clasifica al cáncer gástrico en función de sus características macroscópicas y su patrón infiltrativo llegando a diferenciarlas en: tipo I (polipoide) caracterizada por presentar lesiones leves junto a una limitada formación del tumor,

tipo II (fungante) con aspecto similar al de un hongo tras la formación de lesiones profundas, tipo III (ulcerada) causante de lesiones ulceradas extensas dentro de la pared del estómago y tipo IV (plana/infiltrativa) relacionada con la infiltración difusa de células cancerígenas **(Agnes et al., 2020; Del Arco et al., 2021)**.

- **Clasificación de Lauren:** Conocido por ser el sistema de clasificación del cáncer gástrico más utilizado en el mundo debido al uso de subtipos histológicos (intestinal y difuso) junto a los principales tipos topográficos (cardiaco y no cardiaco) presentes en dicho cáncer, con la finalidad de identificar distintos tipos en función de sus características **(Thrift & Nguyen, 2021)**, siendo que pueden ser las siguientes: bien diferenciado (no cardiaco/intestinal) presente en adultos mayores en forma de glándulas o túbulos que se asemejan a la mucosa intestinal, poco diferenciado (cardiaco/difuso) común entre mujeres jóvenes en forma de anillo o poco cohesiva y mixto regularmente presentado por hombres **(Machlowska et al., 2020; Sexton et al., 2020)**.
- **Clasificación de la OMS:** Similar a anteriores clasificaciones, este tipo de sistema añade características de varios grupos de adenocarcinomas gástricos (subtipos histológicos) y conforme a su última edición se identifican los siguientes tipos: tubular conocido por ser el más frecuente dentro de este tipo de cáncer y la alta cohesión entre sus células, papilar que presenta una superficie luminal hacia afuera formando estructuras papilares similares a dedos, mucinoso caracterizado por ser la segunda más frecuente en aparecer y la producción de un componente gelatinoso (mucina), poco cohesivo destinados para los carcinomas de tipo no sólido junto a los de tipo difuso y mixto que a pesar de no contar con una descripción clara implica la presencia de una combinación tanto de tumores tubulares como de mucinosos **(Kushima, 2022; Machlowska et al., 2020)**.

1.1.2.2. Causas y Efectos

Generalmente, el desarrollo de cualquier tipo de cáncer suele estar influenciado por diversos factores genéticos y ambientales, se incluyen hábitos y comportamientos con un alto impacto en la generación de este tipo de diagnósticos, como lo son: la elevada ingesta de sales y alimentos ahumados o asados, consumo de alcohol, tabaquismo, las infecciones por bacterias o virus (*Helicobacter pylori* y virus de Epstein-Barr) y en menor medida los antecedentes familiares **(Machlowska et al., 2020)**. Dichos factores, comúnmente ocasiona un crecimiento anormal de células tumorales alrededor del revestimiento del estómago que puede llegar a extenderse hasta varios órganos subyacentes como es el caso del hígado o el páncreas, sin llegar a presentar ningún síntoma característico hasta etapas muy avanzadas debido a la obstrucción

de la luz perteneciente a la región gástrica **(Cleveland Clinic, 2022; Waldum & Fossmark, 2021)**.

1.1.2.3. Diagnóstico y tratamientos contra el cáncer gástrico

Recientemente, se han desarrollado diversos procedimientos (biomarcadores circulantes, microARN, entre otros) para el diagnóstico temprano del cáncer gástrico, sin embargo, aun presentan una baja sensibilidad y especificidad durante la presentación de sus resultados **(Necula et al., 2019)**. Por tal razón, procedimientos invasivos como la biopsia endoscópica siguen siendo recurrentes en el diagnóstico de este tipo de enfermedad, a pesar de que se realizan durante etapas muy avanzadas de la enfermedad, donde, la tasa de supervivencia de los pacientes es baja incluso llegando a ser del 10% **(L. Yang et al., 2020)**.

La mayoría de los tratamientos se tratan de terapias combinadas como: la cirugía tradicional y laparoscópica, la quimioterapia convencional, la terapia dirigida, la radioterapia y la inmunoterapia que ha logrado importantes avances con el control inmunológico por medio de inhibidores, de igual forma, se puede incluir al uso de diversos fármacos como **(K. Li et al., 2021; Patel & Cecchini, 2020)**:

- Anti-HER2 (receptor de tirosina quinasa)
- Trastuzumab (anticuerpo monoclonal IgG1)
- Trastuzumab emtansina (conjugado anticuerpo - fármaco)
- Pertuzumab (anticuerpo monoclonal humanizado)
- Lapatinib (inhibidor competitivo del ATP)

1.1.3. Factor NF-κB

El Factor Nuclear Kappa B (NF-κB) considerado como un factor de transcripción de proteínas, presenta un papel clave durante la supervivencia, proliferación y expresión de distintos tipos de células, además, mantiene una enorme influencia en las propiedades de proinflamación y neuro protección, en gran parte, por ser un regulador del sistema inmune innato **(Albensi, 2019)**, es responsable de la señalización biológica compleja del cuerpo, resultante de la activación de su vía de señalización (IκB quinasa o IKK) por diversos estímulos: señales de estrés, antígenos inmunes, citoquininas inflamatorias o productos microbianos que permiten al complejo IKK trasladarse hacia el núcleo, en donde, la poli ubiquitinación y fosforilación de NF-κB permite

la degradación proteasomal de dicha proteína y la liberación de sus dímeros para su traslocación **(Mulero et al., 2019; Ramadass et al., 2020)**.

1.1.3.1. Subunidades y vías de señalización del factor NF-κB

Dentro de la familia NF-κB constan cinco miembros que pueden interactuar entre sí: RelA (p65), RelB, c-Rel, NFκB1 (p105/p50) y NFκB2 (p100/p52), donde, cada subunidad presenta un dominio N-terminal (RHD) que permite la unión, dimerización e interacción del ADN con múltiples inhibidores y proteínas **(Medeiros et al., 2021)**. Además, se sabe que tanto la subunidad p50 como p52 no son capaces de activar una transcripción in vitro o in vivo y RelB solo llega a formar homo o heterodímeros con dichas subunidades **(Barnabei et al., 2021)**.

Por otra parte, el complejo quinasa IKK consta de un par de subunidades catalíticas, conocidas como: IKKα y IKKβ junto a una subunidad reguladora conocida por ser moduladora esencial de NF-κB y principal componente de la vía de señalización de NF-κB (NEMO) **(Peng et al., 2020)**; La activación de dicha vía se realiza a través de una vía canónica inducida por medio de diversas moléculas (lipopolisacáridos, interleucina-1 y TNFα) y relacionada con el desarrollo del sistema nervioso o de una vía no canónica dependiente del complejo NFκB2 (p100)/RelB de vital importancia en el desarrollo del sistema inmunológico **(Medeiros et al., 2021; Peng et al., 2020; Zinatizadeh et al., 2021)**.

1.1.3.2. Factor NF-κB y su relación con el cáncer gástrico

NF- κB actúa como regulador central de la respuesta inmune e inflamatoria del organismo, por lo tanto, si existe una desregulación durante la activación de dicho factor se pueden presentar varias alteraciones (mutaciones oncogénicas, daños en los tejidos y ADN) que permiten la carcinogénesis gástrica, supresión de la apoptosis, metástasis e inestabilidad genómica que puede incluir cierto tipo de resistencia a medicamentos destinados al tratamiento del cáncer **(Chaithongyot et al., 2021; Peng et al., 2020)**. Un claro ejemplo de estas alteraciones son: los diferentes polimorfismos presentes en los genes del factor NF-κB y sus moléculas de señalización, en la Tabla 1 **(Chaithongyot et al., 2021)**.

Tabla 1.

Polimorfismos presentes en genes de NF-κB y en moléculas de señalización.

Nombre del Gen	Polimorfismos	Observaciones relacionadas con el cáncer gástrico
NFKB1	SNP_rs230521	Presente en pacientes con cáncer gástrico
	SNP_rs28362491 (−94 pulgadas/del ATTG)	Se encuentra asociado con el cáncer gástrico difuso, acelera la inflamación gástrica severa
	SNP_rs4648068	Presenta mayor riesgo de cáncer gástrico
	deleción homocigótica	Se observa atrofia gástrica en ratones relacionada a un cáncer gástrico invasivo
NFKB2	deleción homocigótica	Presenta una hiperplasia gástrica junto a una muerte postnatal temprana
		Se observa en lesiones de la mucosa gástrica de forma suprimida
NFKBIA	SNP_rs2233408 T/C homocigoto	Presenta cierta susceptibilidad al cáncer gástrico
	SNP_rs2233408 T heterocigoto	Se observa una reducción del riesgo de cáncer gástrico no cardíaco de tipo intestinal
	SNP_rs17103265	Considerado como un factor de riesgo para la carcinogénesis gástrica
	SNP_rs696	Susceptible al cáncer gástrico cardiaco
	SNP_rs2233406	Susceptible al cáncer gástrico no cardiaco
IKBKB	SNP_rs2272736 A homocigoto	Relacionado con un tiempo prolongado de supervivencia general
TNIP1	SNP_rs7708392	Asociado con el riesgo de cáncer gástrico
	SNP_rs10036748	
MYD88	deleción, mutación	Presenta daño de la mucosa gástrica y carcinogénesis
	L265P mutante	Se observa en linfomas de tejido linfoide asociado a la mucosa gástrica (MALT)
RIPK2	SNP_rs16900627	Presenta mayor riesgo de cáncer gástrico intestinal
TLR9	SNP_rs5743836 (-1237 T/C)	Asociado con el cáncer gástrico inducido por *H. pylori*

De igual forma, se puede llegar a vincular con el cáncer gástrico a los distintos genes regulados por NF-κB presentados en la Tabla 2 y que pertenecen a: citoquininas proinflamatorias, células epiteliales, macrófagos, ciclo celular, entre otros **(Chaithongyot et al., 2021)**. Es necesario establecer una inhibición selectiva de su vía de activación y señalización a través de proteínas inhibitorias tales como los microARN (miARN), compuestos farmacológicos como el salicilato de sodio, la dexametasona u otro fármaco antinflamatorio acompañado de diferentes terapias convencionales para un tratamiento más efectivo. Hay que tomar en cuenta que el uso de este tipo de inhibidores durante un tiempo prolongado puede provocar inmunodeficiencia **(Mirzaei et al., 2021; Yu et al., 2020)**.

Tabla 2.

Genes regulados por NF-κB vinculados con el desarrollo de cáncer gástrico.

Genes regulados por NF-κB	**Observaciones relacionadas con el desarrollo del cáncer gástrico**
IL-8	Se encuentra relacionado con el cáncer gástrico de tipo difuso
	Vinculado con las siguientes características: • Invasión venosa y linfática • Profundidad de la invasión • Baja tasa de supervivencia
IL-17	Vinculado positivamente con el cáncer gástrico y la migración celular
IL-1β	Promueve la displasia gástrica a cáncer gástrico
COX2, MMP9, VEGF	Se promueve la invasión y migración celular en el cáncer gástrico a través de su expresión
PD-L1	Relacionado con el cáncer gástrico de tipo intestinal en etapas temprana
	Asociado con un mal pronóstico del cáncer gástrico
NO, PGE2	Promueve un ambiente inflamatorio mediante la infiltración de macrófagos en el tejido del estómago
	Se observa un aumento de la microcirculación en la restitución celular que permite la curación de los tejidos
	Aumento en la tasa de mutagénesis del tejido inflamado por medio del recambio acelerado de las células epiteliales
iNOS , COX-2	Contribuye al progreso gradual de la carcinogénesis gástrica
STAT3	Relacionado con el desarrollo y progresión de cáncer gástrico.

Genes regulados por NF-κB	Observaciones relacionadas con el desarrollo del cáncer gástrico
c-myc , ciclinaD1	Vinculado a la alta expresión de cáncer gástrico de tipo intestinal
HNF4α	Vinculado con la inflamación prolongada (sobreexpresión)
miR-223-3p, miR-18a-3p, miR-4286	Se presenta un vínculo con la su expresión en células o tejidos y la carcinogénesis gástrica
miR-425	Promueve la proliferación del cáncer gástrico
Noxo1	Relacionado con la gastritis
Snail1	Regula negativa de E-cadherina en las células cancerígenas
hTERT	Promueve la metaplasia intestinal

1.1.4. Proteínas *de novo*

En el pasado, el diseño de proteínas *de novo* fue considerado casi imposible a causa de las limitaciones de la época, debido a la complejidad de la proteína y la inexistencia de herramientas que faciliten su diseño. Actualmente, el uso de herramientas como: las redes neuronales convolucionales o los modelos de aprendizaje profundo que permitan estabilizar la columna vertebral (*backbone*) de la estructura proteica en función de la secuencia otorgada para obtener una conformación tridimensional plegada que imite eficazmente distintas funciones de una proteína natural ha mejorado eficazmente el proceso de diseño de nuevas proteínas **(Qing et al., 2022; Wu et al., 2022)**.

Con el paso del tiempo y la evolución continua de los métodos computacionales se ha logrado diseñar varios tipos de proteínas con diversas funciones y pliegues **(Korendovych & DeGrado, 2020)**, inclusive se ha proporcionado nuevos conocimientos acerca de la estructura y plegamiento de las proteínas a partir del desarrollo de nuevas variantes proteicas que pueden o no estar relacionadas con una especie nativa (**Qing et al., 2022)**.

1.1.4.1. Enfoques generales para la predicción de proteínas de novo

En sus inicios, la construcción y predicción de proteínas empleó un enfoque basado en el ensamblaje de fragmentos de corta longitud que tomaba en cuenta datos cristalográficos de estructuras proteicas con una alta similitud entre sus secuencias, dando como resultado modelos tridimensionales aleatorios con características similares (estructuras secundarias, ángulos de torsión, entre otras) a las de sus predecesores **(Korendovych & DeGrado, 2020; J. Yang et**

al., 2020). Actualmente, gracias a la potencia informática junto con el rápido crecimiento de las bases de datos biológicas se ha logrado impulsar el desarrollo de dos nuevos enfoques generales **(Kuhlman & Bradley, 2019)**:

- **Modelado basado en plantillas:** Destinado principalmente a la construcción y ajuste de modelos en función de una plantilla estructural adecuada, en donde, se toma en cuenta las secuencias disponibles pertenecientes a cierta familia de proteínas con la finalidad de proporcionar información acerca de las mutaciones, inserciones y eliminaciones disponibles en dicha estructura.
- **Modelado sin plantillas:** Enfocado en la generación de modelos candidatos sin ninguna similitud o relación aparente con estructuras proteicas disponibles en las bases de datos, a través de, un largo proceso de construcción, refinamiento, clasificación, comparación y selección de dichos modelos conforme a las necesidades planteadas durante su elaboración.

1.1.4.2. Aplicaciones de las proteínas de novo

Su motivación principal arraigada en la comprensión del aspecto y plegamiento de las proteínas, mediante el uso de diseños computacionales enfocados en la predicción de estructuras proteicas de menor tamaño y más estables a factores de desnaturalización los cuales están, destinados a incrementar su potencial terapéutico dentro de la industria farmacéutica **(Korendovych & DeGrado, 2020; Woolfson, 2021)**, con lo que se ha logrado desarrollar terapias proteicas enfocadas en la neutralización de los efectos de las drogas por medio del diseño de enzimas estables que mejoren la acción catalítica **(Planas-Iglesias et al., 2021)**.

También, surgen otras potenciales aplicaciones como es el caso de: la generación de nanomateriales estables y funcionales o interfaces biocompatibles, el descubrimiento y diseño de nuevos fármacos a nivel molecular, predicción de las propiedades de un compuesto, entre muchos otros destinados al desarrollo y análisis de datos preclínicos **(Lin et al., 2020; Qing et al., 2022)**.

1.1.4.3. Desafíos de las proteínas de novo

El diseño y predicción de estas estructuras aún presenta grandes desafíos debido a que su construcción tradicional comienza desde el interior (cavidad proteica) hacia el exterior (cadenas

laterales polares) de la proteína con el único fin de maximizar las interacciones ligando-proteína favorables, incluso con los avances computacionales de hoy en día, dichas proteínas tienden a mantener los principios básicos de las estructuras pertenecientes al Protein Data Bank (PDB) **(Pan & Kortemme, 2021)**.

La mayoría de este tipo de proteínas no presentan una función biológica por conformación evolutiva, por tanto, destinar una función suele ser un desafío que aún no se ha logrado resolver en su totalidad junto a la generación de grandes conjuntos de interacciones con un repertorio de monómeros **(Baker, 2019; Sesterhenn et al., 2020)**. A pesar de esto, el desarrollo de proteínas con una alta afinidad y especificidad ha logrado resolver una gran problemática relacionada con la predicción de cambios en la estructura o secuencia proteica, siendo más fácil predecir dichos cambios en las conformaciones y comprender las interacciones proteína-proteína en las proteínas *de novo* que en las proteínas naturales **(Baker, 2019)**.

1.1.5. Inteligencia artificial dentro de la Biomedicina

Durante una conferencia informática en 1956, el término "Inteligencia Artificial" describió a aquellos algoritmos preliminares que permitieron reproducir la cognición humana junto a su capacidad de adaptabilidad y aprendizaje en un entorno computacional, a través del uso de redes locales para la búsqueda y recopilación de información, dando como resultado un alto rendimiento en áreas específicas: desde el reconocimiento de patrones de rostros, voz o huellas dactilares, la detección de fallas en procesos, ecuaciones y códigos hasta la toma de decisiones a partir de patrones adquiridos **(Bernstein, 2022; Hassani et al., 2020)**.

Dentro de la medicina fue presentado durante la década de 1970 como sistema complementario de las computadoras que ayudaba en la resolución de problemas médicos conforme los conocimientos de distintos especialistas **(Kulikowski, 2019)**, no obstante, solo recientemente se ha vuelto tendencia debido la inclusión de técnicas de aprendizaje profundo, aprendizaje automático y redes neuronales añadiendo a la generación de computadoras con un poder computacional eficiente que permiten cumplir con su objetivo inicial y desarrollar nuevos estándares tanto de entrada como de salida de datos e información **(Hulsen, 2021; Rong et al., 2020)**.

1.1.5.1. Potencial biomédico de la Inteligencia Artificial

El valor del uso de la Inteligencia Artificial puede llegar a ser incalculable, debido a que tiene el potencial de reducir los costos de investigación y gestión administrativa, gerencial o médica. La identificación de información cuantitativa y cualitativa en revistas lideres dentro del campo o relacionado al mismo permitirá la optimización de los procesos, destinando sus recursos en la atención médica multidisciplinaria **(Secinaro et al., 2021)**. Por tal razón, actualmente el uso de herramientas de inteligencia artificial se ha destinado en aplicaciones biomédicas, en donde, conforme a su potencial se clasifican en **(Kaul et al., 2020; Schwalbe & Wahl, 2020)**:

- De aprendizaje automático (diagnóstico)
- Procesamiento de datos (evaluación de riesgos)
- Procesamiento natural del lenguaje (Predicción y vigilancia de enfermedades)
- Planificación experta (Política y planificación)
- Robots físicos (Automatización de procesos)

Además, se ha implementado dentro de los procesos de diseño de fármacos y tratamientos contra el cáncer por medio de técnicas como: cribado virtual de estructuras y ligandos, liberación de medicamentos, evaluación de parámetros de toxicidad, actividad fisicoquímica de distintos compuestos farmacológicos y comprensión de los mecanismos moleculares **(Gupta et al., 2021; Melo et al., 2021; Nagarajan et al., 2019)**. Inclusive, se ha logrado solventar los altos costos que presenta el estudio y predicción de estructuras proteicas destinadas a mejorar el desarrollo de nuevos medicamentos mediante el refinamiento o modelado computacional de sus estructuras tridimensionales y la optimización e identificación de nuevas dianas terapéuticas para el tratamiento del cáncer **(Dhakal et al., 2022; J. L. Watson et al., 2023)**.

Por lo tanto, este tipo de herramientas representan tanto el presente como el futuro de la humanidad, especialmente dentro del desarrollo de fármacos y predicción de proteínas, siendo necesario conocer algunas herramientas de inteligencia artificial aplicadas a este enfoque, tales como: AlphaFold desarrollada por Google's Deep mind, Deepchem presentada por Googles's TensoFlow framework o Rfdiffusión elaborada por Rossetta **(Chopra et al., 2022)**.

1.2. Objetivos

1.2.1. Objetivo General

Predecir proteínas *de novo* mediante inteligencia artificial para la inhibición del factor NF-κB en el cáncer gástrico.

1.2.2. Objetivos Específicos

- Identificar proteínas *de novo* estables relacionadas con la inhibición del factor NF-κB en el cáncer gástrico utilizando herramientas de inteligencia artificial.
- Comparar las variables termodinámicas obtenidas de la interacción entre el factor NF-κB y las proteínas predichas.
- Evaluar la afinidad y selectividad de las proteínas predichas durante la inhibición del factor NF-κB.

CAPITULO 2

METODOLOGÍA

El presente trabajo de titulación fue llevado a cabo dentro de las instalaciones de la Facultad de Ciencia e Ingeniería en Alimentos y Biotecnología de la Universidad Técnica de Ambato. En la Tabla 3, se enlistan las aplicaciones, softwares y servidores web empleado durante la investigación, y en la Tabla 4 se presentan los equipos utilizados para el mismo.

2.1. Materiales

Tabla 3.

Recursos Bioinformáticos.

Recursos	Enlace
PDB (Protein Data Bank)	https://www.rcsb.org/
Uniprot	https://www.uniprot.org/
DogsiteScorer	https://proteins.plus/#dogsite
BLAST	https://blast.ncbi.nlm.nih.gov/Blast.cgi
NCBI	https://www.ncbi.nlm.nih.gov/
Jalview	https://www.jalview.org/download/
AlphaFold	https://alphafold.ebi.ac.uk/
Deepchem	https://github.com/deepchem/deepchem
RFdiffusion	https://github.com/RosettaCommons/RFdiffusion
PyMOL	https://pymol.org/
CHARMM	https://www.charmm.org/archive/charmm/showcase/news/free-charmm/
CHARMM-GUI	https://www.charmm-gui.org/
GROMACS	https://manual.gromacs.org/current/install-guide/index.html
ChimeraX	https://www.cgl.ucsf.edu/chimerax/
Protein Plus	https://proteins.plus/
AutoDock Vina	https://vina.scripps.edu/downloads/

Tabla 4.

Equipos.

Recursos	Cantidad
Computadora Portátil	1
Router para Internet	1

2.2. Métodos

2.2.1. Identificación de proteínas *de novo* estables relacionadas con la inhibición del factor NF-κB en el cáncer gástrico utilizando herramientas de inteligencia artificial.

2.2.1.1. Búsqueda de secuencias y estructuras proteicas relacionadas con el factor NF-κB presente en el cáncer gástrico

Para la predicción precisa de proteínas, se estableció la secuencia y estructura tridimensional del factor NF-κB como base para la construcción de inhibidores que presenten la capacidad de interaccionar con este, como también las diferentes mutaciones con las que se relaciona dentro del cáncer gástrico. Esto permitió identificar regiones específicas que pueden llegar a ser objetivos de inhibición, además, se verificaron otras posibles interacciones con proteínas naturales para establecer puntos de comparación en secciones futuras de esta investigación.

Fue necesario el uso de diversos repositorios de datos como: PDB (Protein Data Bank), Uniprot y DogSiteScorer para la visualización de sus estructuras proteicas cristalizadas e identificación de sus secuencias, en donde, PDB y Uniprot tuvieron como entrada principal el factor "NF-κB" o cada uno de los nombres de las mutaciones encontradas tras una revisión bibliográfica; se seleccionaron aquellas proteínas provenientes del organismo "*Homo sapiens*" **(Bittrich et al., 2023)**, mientras que para DogSiteScorer se utilizó como entrada exclusiva el código "PDB" (e.g. 2ozr) de dichos elementos y se mantuvo los parámetros de selección anteriormente mencionados **(Michel et al., 2019)**.

2.2.1.2. Predicción de proteínas de novo mediante el uso de inteligencia artificial

El uso de modelos de aprendizaje automático para la predicción de estructuras y secuencias novedosas permitió que varias proteínas *de novo* recientemente desarrolladas presenten distintos diseños **(Pan & Kortemme, 2021)**.

En el caso de construcciones proteicas de doble hélice β se empleó la metodología dispuesta por **Marcos et al., (2018)**, se especificó que la longitud de las hebras mantuviera un rango de 5 a 7 residuos y de 3 a 5 residuos para los bucles de arco β, también se permitió un rango de desplazamiento de 0 a 2 para hebras β de distinto arco junto a un desplazamiento para los arcos β de 0 a 4, mientras que para las hebras pertenecientes a un mismo arco se mantuvieron sin ningún cambio, resultando en la formación de estructuras estrechas (tipo sándwich) para aquellas proteínas con un pequeño desplazamiento y para grandes desplazamientos estructuras anchas (tipo barril); se incluyó enlaces disulfuro entre la primera y última β-cadena para

disminuir la entropía y aumentar la estabilidad termodinámica del diseño. En cambio, para construcciones que contengan los elementos bucle-hélice-bucle (LHL) su elaboración se llevó a cabo mediante el método descrito por **(Pan et al., 2020)**, que excluye todas las estructuras redundantes y comienza con un pliegue de proteína que permitió insertar bucles de 5 residuos o menos en los extremos de cada espacio y en las hélices de 2 a 5 residuos de longitud en el medio, en donde; se eliminó toda secuencia del bucle que entre en conflicto con el pliegue de la proteína inicial junto al exceso de residuos presente.

Así también, fue necesario el uso de "*templates*" con el fin de obtener una adecuada conformación proteica similar a las estructuras naturales, para esto se utilizó el programa BLAST en conjunto con el repositorio de datos NCBI los que permiten identificar diferentes "*clusters*" pertenecientes a bases de datos experimentales que presenten un alineamiento homólogo casi idéntico (mayor al 70 %) con las secuencias anteriormente investigadas. De esta manera, se generó un alineamiento "múltiple" y "por pares" entre los "*clusters*" y la "*query sequence*", utilizando los ficheros que contienen dichas secuencias en el software Jalview, esto permitió seleccionar los fragmentos con un alto índice de conservación junto a los segmentos disponibles para un mejor consenso en la predicción de las proteínas *de novo*.

En estas predicciones, se empleó diferentes programas de Inteligencia Artificial (IA) tales como: AlphaFold, Deepchem y RFdiffusion que generaron modelos 3D con cierta similitud con las proteínas naturales. Para "AlphaFold" se estableció como entrada el nombre del factor propuesto y se seleccionó solo aquellos que muestran un acceso directo a su estructura cristalizada dentro de la base de datos "Uniprot" y que pertenezcan al organismo *H. sapiens* **(Senior et al., 2020)**, en cambio, en "Deepchem" se estableció una línea de comando presentada en el Anexo 1 que permitió establecer la longitud de las secuencias proteicas con la finalidad de modelarla en colaboración de los servidores de Google Colab de "AlphaFold" **(Bellamkonda, 2023; Chang et al., 2024)**. Mientras que,"RFdifussion" ejecutó sus propias líneas de comando disponibles dentro de su propio servidor de Google Colab, permitiendo establecer los parámetros de diseño: longitud de la proteína y la generación de dímeros, oligómeros o monómeros junto al número de variantes que se quieren diseñar, estableiendo que por cada herramienta se genere 5 a 10 estructuras **(Pinzi & Rastelli, 2019; J. L. Watson & Juergens, 2023; K. Watson, 2022)**.

2.2.1.3. Visualización de las proteínas predichas

Para la visualización de estructuras moleculares tanto de proteínas como algún tipo de compuestos, se utilizó herramientas como "PyMOL", que permite una rápida y fácil exploración de la composición estructural de una proteína, por lo tanto, es fundamental su utilización en el diseño de proteínas *de novo,* desarrollo de fármacos y nuevas terapias personalizadas **(Martinez et al., 2020)**. Por tal razón, cada una de las proteínas inhibitorias predichas en la sección anterior y que presentan estructuras tridimensionales, fueron descargas en formato ".pdb" y a través de la aplicación "PyMOL" se abrieron por medio de los comandos: "File; Run script" con el propósito de visualizar de forma completa la estructura, composición, motivos y dominios estructurales, familias de proteínas, entre otras características que componen dichas moléculas **(Mooers & Brown, 2021)**.

2.2.1.4. Análisis de estabilidad de las proteínas predichas

El análisis de estabilidad de las proteínas predichas fue una de las secciones más cruciales dentro de esta investigación, ya que identificó si las secuencias proteicas establecidas son adecuadas para el proceso de interacción proteína-proteína, por lo que, se examinó su conformación y plegamiento a través de técnicas y herramientas computacionales para comprender la complejidad de dichas moléculas **(Rocklin et al., 2017)**. Se utilizó las siguientes herramientas, Chemistry at HARvard Macromolecular Mechanics conocida como CHARMM que es un programa informático con la capacidad de leer distintas estructuras proteicas con las que se llevó a cabo simulaciones (modo normal o dinámica molecular) y análisis de propiedades de equilibrio, dinámicas y estructurales presentes en las proteínas predichas conforme a lo descrito por **(Brooks et al., 1983)**.

Se empleó CHARMM apoyado del uso de su servidor web gratuito "CHRAMM-GUI" (debido a que el software "PROTDES" se encuentra descontinuado y deshabilitado), en donde, se subieron a la red las estructuras proteicas predichas en forma de archivo PDB y desde ese lugar se calcularon: energías de referencia, capacidad de plegamiento e interacciones residuo-residuo presentes en dichas estructuras **(Jo et al., 2008)**, además, de los parámetros básicos para las simulaciones de dinámica molecular que fueron ejecutados en la aplicación GROMACS **(Abraham et al., 2015)**. Finalmente, con estos resultados se llenó una tabla con los valores de los diferentes parámetros analizados y se definió aquellas proteínas que cumplen con las condiciones de estabilidad deseable.

2.2.2. Comparación de las variables termodinámicas obtenidas de la interacción entre el factor NF-κB y las proteínas relacionadas.

2.2.2.1. Acoplamiento y Docking molecular entre el factor NF-κB y las proteínas predichas

Estas se consideran técnicas destacadas en la predicción de interacciones moleculares, tienen como finalidad el análisis de las fuerzas y energías presentes en la interacción y la determinación de su conformación óptima. Por lo tanto, para iniciar este procedimiento se descargó cada una de las estructuras pertenecientes al factor NF-κB y las mutaciones relacionadas a dicho factor dentro del cáncer gástrico lo que conjuntamente con las proteínas predichas y seleccionadas como estables en secciones anteriores; se probaron para la eliminación de sitios irrelevantes en los mismos, utilizando el programa ChimeraX v1.6.1 o superior **(Haure-Mirande et al., 2022; Pettersen et al., 2021)**; luego se utilizó el servicio Protein Plus para la predicción de sitios de unión apropiados dentro de las estructuras previamente importadas y procesadas, en donde, se establecieron valores "DrugScore" entre 0.5 a 1 de significancia **(Michel et al., 2019; Schöning-Stierand et al., 2020)**.

Para el acoplamiento molecular se empleó AutoDock Vina junto a las estructuras tanto del factor empleado como de las estructuras proteicas anteriormente cargadas, se cargaron a través de: "File: Read Molecule" y "Ligand: Input" las cuales no deben contener moléculas de hidrogeno, agua o cargas Kollman **(Desta et al., 2020; Salmaso & Moro, 2018)**. Finalmente, se estableció las coordenadas de los sitios de unión por medio de mapeos estructurales que permitieron el uso de la opción "Docking; Macromolecule; Set Rigid Filename" con los formatos ". pdbqt" de las estructuras de las proteínas que van a interaccionar **(Eberhardt et al., 2021; Tang et al., 2022)**, se estableció como parámetro de 10 predicciones de sitios de unión que se identificaron mediante la opción "DogSiteScorer" **(Pham et al., 2022)**.

2.2.2.2. Determinación y Comparación de las variables termodinámicas presentes en la interacción proteína-proteína

La comparación de las variables termodinámicas (Energía libre de Gibbs, Constante de disociación, Entalpía, Capacidad Calorífica y Entropía) suele ser parte del análisis de acoplamiento obtenido de la interacción proteína-proteína, por lo tanto, su contrastación es necesaria para la verificación de la eficacia de las proteínas en la inhibición del factor NF-kB presente en el cáncer gástrico **(Deeks et al., 2023; Holdgate & Ward, 2005)**.

Se estableció como control una de las interacciones entre proteínas naturales identificadas en la primera sección de la metodología de esta investigación, en donde, conforme a los valores registrados de las variables termodinámicas: Entalpía y Capacidad calorífica se asumió un sistema a presión y temperatura constantes (25 ºC a 1 atm) igual para cada una de las interacciones obtenidas, mientras que las variables restantes se obtuvieron a través del proceso de acoplamiento molecular de las proteínas predichas y las fórmulas de equilibrio:

$\Delta G = \Delta H - T\Delta S$ y $\Delta G = -RT \ln K_d$.

Donde:

- **ΔG** = Energía libre de Gibbs
- **ΔH** = Entalpía
- **T** = Temperatura
- **ΔS** = Entropía
- **R** = Contante de los gases
- **K_d** = Constante de disociación

2.2.3. Evaluación de la afinidad y selectividad de las proteínas predichas durante la inhibición del factor propuesto.

2.2.3.1. Determinación y evaluación de la afinidad y selectividad de las proteínas predichas durante la inhibición del factor propuesto.

Para la evaluación de la selectividad de las proteínas predichas durante la inhibición del factor propuesto, se categorizó a cada una de las interacciones proteína-proteína conforme la cantidad de genes o factores NF-kB con los que han logrado interaccionar estas proteínas **(Cao et al., 2022; Ivarsson & Jemth, 2019)**, aquellas estructuras con la menor cantidad de interacciones serán consideradas como "selectivas". Mientras que para considerar la afinidad de una proteína se tomó en cuenta la "energía libre de Gibbs" determinada en la sección anterior durante el acoplamiento molecular, donde una ΔG más negativa sugiere una interacción más fuerte y, por lo tanto, una mayor afinidad.

CAPITULO 3

RESULTADOS Y DISCUSIÓN

3.1. Análisis y discusión de los resultados

3.1.1. Búsqueda de Secuencias y Predicción de proteínas *de novo*

Si bien en estos últimos años, se ha logrado identificar y clasificar diferentes tipos de cáncer gástrico conforme a su histopatología y características macroscópicas, todavía no se refleja en su totalidad la heterogeneidad de dicha enfermedad, se ha demostrado que cada subtipo de cáncer gástrico presenta patrones moleculares específicos vinculados a alteraciones o aberraciones genómicas y vías oncogénicas **(Carino et al., 2021; Cristescu et al., 2015)**. Incluso se han asociado con la activación irregular de diversos promotores, como es el caso de los genes y factores de transcripción regulados por NF-κB disponibles en la Tabla 5.

Tabla 5.

Genes regulados y factores de transcripción relacionados con NF- κB en el cáncer gástrico.

	Factores regulados por NF- κB	Nombre del Gen	Código PDB
Subunidades de NF- κB	Quinasa inductora de NF- κB (NIK)	MAP3K14	4G3D
	Subunidad p50 (NFκB1)	NFKB1	1SVC
	Subunidad p52 (NFκB2)	NFKB2	1A3Q
	Subunidad p65 (RelA)	RELA	1NFI
	Subunidad p100 (NFκB2)	NFKB2	4OT9
	Subunidad p105 (NFκB1)	NFKB1	2DBF
Genes regulados por NF- κB	CiclinaD1	CD1	6P8H
	Ciclooxigenasa-2	COX-2	5IKV
	Interleucina-B	IL-B	9ILB
	Interleucina-8	IL-8	6N2U
	STAT3	STAT3	6NUQ

Por lo tanto, se ha planteado el uso dichos genes y factores de transcripción como objetivos terapéuticos contra del cáncer gástrico, debido a que en su mayoría se encuentran relacionados con procesos inflamatorios mediados por la vía de señalización de NF-κB o sus subunidades, por lo que fomentan la progresión de tumores asociados con la enfermedad **(Pflug & Sitcheran, 2020)**. Por ejemplo, tanto NIK como IL-8 se encuentran relacionados con el desarrollo de

linfomas y ulceras pépticas con un mayor grado de agresividad **(Marta et al., 2020; Mondragón et al., 2019)**, mientras que COX-2 al ser una enzima que se expresa en la mucosa inflamada tiene un papel clave en la tumorigénesis, así también, STAT3 está presente en etapas tempranas de la infección por *H. pylori* ayuda a la proliferación de células cancerígenas a través de la activación de las vías proinflamatorias **(Ashrafizadeh et al., 2020; Echizen et al., 2019; Soutto et al., 2021)**.

Es una prioridad el desarrollo de tratamientos con la capacidad de suprimir la expresión del factor NF-κB y de este tipo de genes. Desde este punto de vista, cada uno de los modelos contemplados en la Tabla 6 junto a las secuencias proteicas descritas en el Anexo 2 son el resultado de redes de aprendizaje altamente entrenadas para la generación de nuevas estructuras proteicas sin algún tipo de restricción o limitación en su plegamiento **(Kortemme, 2024)**, por tal motivo, la verificación de su confiabilidad y efectividad se presentan como el punto de partida para la investigación de nuevas estrategias especializadas en este contexto.

Tabla 6.

Proteínas generadas por las herramientas de Inteligencia Artificial en base a inhibidores del factor NF- κB.

N.º	**Códigos para la identificación de las proteínas predichas**		
	AlphaFold	**Deepchem**	**RFdiffusion**
1	O15226	A7886	RF21A
2	P25963	EC19F	RF415AD
3	Q15653	963D2	RF715A1
4	O15111	92B1D	RF315AD
5	O14920	A26C8	RF215A2
6	Q8NI38	208E3	RF315BD
7	O00221	10FAD	RF51B
8	Q9BYH8	9061D	RF2M31
9	Q9Y6K9	533E3	RF3M31
10	Q9UBC1	07795	RF2M32

Nota: En total, se predijeron 30 proteínas *de novo* a partir de datos disponibles en repositorios (AlphaFold) o conjuntos de datos experimentales (RFdiffusion) y estándar (Deepchem).

Uno de los problemas latentes, es la poca credibilidad que se le otorga a estos modelos, lo cual se debe a la falta de información acerca de la capacidad y precisión de las IA al momento de realizar múltiples diseños o conformaciones similares a proteínas naturales **(Del Alamo et al., 2022)**, a pesar de ello, el progreso y evolución constante de este tipo de herramientas de inteligencia artificial han permitido determinar las mejores métricas, composiciones estructurales y funciones biológicas con la única finalidad que respondan adecuadamente a las pruebas experimentales a las que serán sometidas durante las siguientes fases de desarrollo **(Schauperl & Denny, 2022; Winnifrith et al., 2024)**.

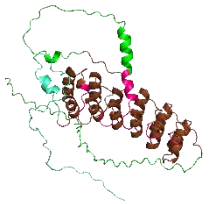
Como evidencia de este progreso, la predicción exitosa de la estructura T1008 ha demostrado que es posible generar modelos homólogos a estructuras PDB con un alto grado de precisión, pero con una clara diferencia entre los rendimientos obtenidos por una proteína *de novo* y una de origen natural **(J. Yang et al., 2020)**.

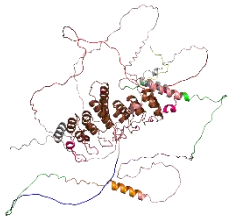
3.1.2. Visualización de las proteínas predichas

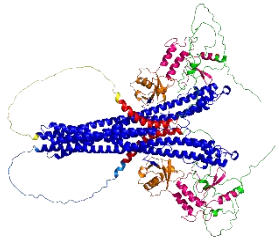
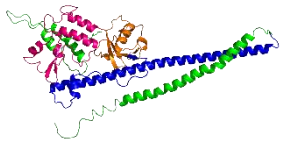
Los modelos obtenidos están basados en una amplia variedad de inhibidores y reguladores naturales de NF-κB, siendo el inhibidor NFKBIA uno de los "*templates*" más empleados en la construcción y predicción de estas estructuras, debido a que puede llegar a disminuir la inflamación causada por los complejos NF- κB/REL **(Wang et al., 2024)**. Por ende, un análisis de la estructura y composición que caracterizan a las proteínas *de novo* resultantes como lo presentado en la Tabla 7 y Anexo 3 permite mejorar la compresión de sus disposiciones tridimensionales, además, de definir las similitudes existentes con respectos a sus bases de datos **(Banach et al., 2020; Yadav et al., 2021)**.

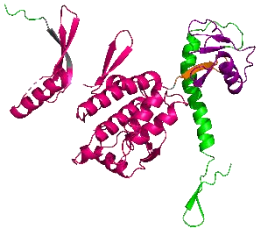
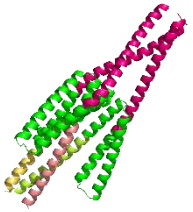
Tabla 7.

Composición y Estructura de las proteínas predichas en las herramientas de Inteligencia artificial.

IA	Código de Identificación	Composición de la proteína	Figuras
AlphaFold	Q15653	**Regiones** • Regiones desordenadas: Verde (149–193, 298–356) **Dominios** • Dominio que contiene repeticiones Ankyrin: Rosado (56-302) **Repeticiones** • ANK 1: Café (57-86) • ANK 2: Café (93-122) • ANK 3: Café (126-155) • ANK 4: Café (206-235) • ANK 5: Café (240-269) • ANK 6: Café (273-302)	 ***Figura 1.*** *Composición y estructura de la proteína Q15653.*

Q8NI38	**Dominios** • Dominio que contiene repeticiones Ankyrin: Rosado (32-303) **Repeticiones** • ANK 1: Café (48-83) • ANK 2: Café (84-113) • ANK 3: Café (117-146) • ANK 4: Café (152-201) • ANK 5: Café (206-236) • ANK 6: Café (243-276)	 ***Figura 2.*** *Composición y estructura de la proteína Q8NI38.*
Q9BYH8	**Regiones** • Requerido para la actividad transcripcional: Rojo (321-394) • Interacción con NFKB1/p50: Gris (404-718) • Regiones desordenadas: Verde (1-20, 58-108, 186-211) **Sesgos compositivos** • Residuos polares: Azul (58-82) • Residuos básicos y ácidos: Amarillo (1-15) **Dominios** • Dominio que contiene repeticiones Ankyrin: Rosado (428-692) • OCA: Naranja (108-130) **Motivos** • Señal de localización nuclear: Morado (164-179) **Repeticiones** • ANK 1: Café (443-472) • ANK 2: Café (479-508) • ANK 3: Café (512-541) • ANK 4: Café (551-580) • ANK 5: Café (582-607) • ANK 6: Café (612-641) • ANK 7: Café (648-681)	 ***Figura 3.*** *Composición y estructura de la proteína Q9BYH8.*

Deepchem	A26C8	**Regiones** • Regiones artificiales: Verde (1-60, 565-625) **Dominios** • Dominio catalítico de la serina/treonina quinasa: Rosado (59-139) • Dominio similar a la ubiquitina (Ubl): Naranja (140-220) • Dominio de dimerización del andamio IQBAL: Azul (219-493) • Dominio de unión a I-kappa-quinasa-beta NEMO: Rojo (536-565)	 ***Figura 4.*** *Composición y estructura de la proteína A26C8.*
	208E3	**Regiones** • Regiones artificiales: Verde (1-60, 565-625) **Dominios** • Dominio catalítico de la serina/treonina quinasa: Rosado (59-139) • Dominio similar a la ubiquitina (Ubl): Naranja (140-220) • Dominio de dimerización del andamio IQBAL: Azul (219-493) • Dominio de unión a I-kappa-quinasa-beta NEMO: Rojo (536-565)	 ***Figura 5.*** *Composición y estructura de la proteína 208E3.*
	9061D	**Regiones** • Regiones artificiales: Verde (1-80, 355-435) **Dominios** • Dominio catalítico de la serina/treonina quinasa: Rosado (65-159) • Dominio similar a la ubiquitina (Ubl): Naranja (160-240) • Dominio de dimerización del andamio IQBAL: Azul (239-355)	 ***Figura 6.*** *Composición y estructura de la proteína 9061D.*

RFdiffusion	RF715A1	**Regiones** • Regiones artificiales: Verde (1-5, 305-359) **Dominios** • Dominio catalítico de la serina/treonina quinasa: Rosado (13-242) • Dominio similar a la ubiquitina (Ubl): Naranja (244-302) **Familias** • Familia de función desconocida (DUF6088): Morado (255-313)	 ***Figura 7.*** *Composición y estructura de la proteína RF715A1.*
	RF315AD	**Regiones** • Regiones artificiales: Verde (113-117) **Dominios** • Dominio catalítico de la serina/treonina quinasa: Rosado (8-242) • Dominio similar a la ubiquitina (Ubl): Naranja (244-302)	 ***Figura 8.*** *Composición y estructura de la proteína RF315AD.*
	RF2M31	**Regiones** • Regiones artificiales: Verde (31-98) **Dominios** • Modulador esencial NF-kappa-B NEMO: Rosado (89-138)	 ***Figura 9.*** *Composición y estructura de la proteína RF2M31.*

Nota: Cada una de las estructuras estudiadas tanto en la presente tabla como las dispuestas en el Anexo 3 presentan una conformación Bucle-Hélice-Bucle (LHL), en donde, la variedad en sus estructuras tridimensionales se ve determinada por la Inteligencia Artificial utilizada, mientras que su composición está definida por el conjunto de datos destinado a dicho modelos.

Tal y como se observa en la Tabla 7 y en las estructuras proteicas expuestas, la mayoría de las proteínas predichas presentan composiciones similares entre sí, las que coinciden con datos cristalográficos de tipo estándar y experimental **(E. H. Li et al., 2023)**. Ante este hecho, los dominios y regiones que mejor representan la composición general de estas estructuras son: el dominio que contiene repeticiones de ankyrin visualizado en la Figura 1, Figura 2 y Figura 3 que describe a una de las familias de proteínas más comunes que puede encontrarse en todas las formas de vida, y tienen un papel clave en las interacciones proteína-proteína de iniciadores transcripcionales, transductores de señales, reguladores de desarrollo y toxinas **(Chakrabarty & Parekh, 2022)**.

De igual forma, se encuentra un dominio de dimerización similar al contenido por las subunidades RelA y p50 de NF-κB dentro de la Figura 4, Figura 5 y Figura 6, donde: su principal rol es la construcción de complejos diméricos **(Baughman et al., 2022)**, además, están el dominio similar a ubiquitina (Ubl) junto al dominio catalítico serina/treonina quinasa disponibles en la Figura 7, Figura 8 y las anteriormente mencionadas que permiten efectuar modificaciones postraduccionales en diferentes vías de señalización destinadas a la transcripción y degradación **(Dreier et al., 2022; Fulcher & Sapkota, 2020)**. El dominio de unión a NEMO visto en la Figura 9 interrumpe las interacciones con el complejo IKK sin presentar algún tipo de efecto secundario o afectar las actividades normales del factor NF-κB **(Glaeser et al., 2020)**, en cambio, en el caso de las regiones "desordenadas" y "artificiales" adquiridas por los modelos en función de la IA, se consideran como secciones únicas de cada proteína que deben ser tomadas en cuenta en las siguientes fases de desarrollo, debido a su conformación poco confiable.

3.1.3. Análisis de estabilidad de las proteínas predichas

En relación con el análisis de estabilidad efectuado a cada uno de los modelos proteicos, se conoce de forma general que, a pesar de la constante evolución de las herramientas de inteligencia artificial no se puede identificar de forma directa la capacidad de plegamiento de las estructuras predichas, de igual manera, las múltiples interacciones intra e intermoleculares consideradas como el actual foco de atención de diversos estudios físicos, químicos y biológicos por su activo papel en el rendimiento y viabilidad de las proteínas solo pueden ser calculados en un determinado punto energético **(Liu et al., 2022; Song et al., 2022)**, por lo tanto, la incorporación de simulaciones computacionales que permiten reflejar el comportamiento de estas proteínas *de novo* en entornos experimentales y con tiempos específicos, ayuda a comprender la naturaleza de este tipo de propiedades **(Pandey et al., 2023)**.

En base a estos criterios, se puede deducir que los datos relacionados con el estado de protonación y los residuos terminales, presentados tanto en la Tabla 8 como en el Anexo 4, son relevantes en el análisis de estabilidad de las proteínas *de novo* debido a que su presencia o ausencia durante el ensamblaje de la estructura tridimensional está ligado a su capacidad de plegamiento y retención. Como evidencia de ello, **D. J. López et al., (2021)** describe a los residuos N-terminal y C-terminal como regiones comunes de varias proteínas de unión al ADN con funciones reguladoras, en donde, N-terminal presenta cierta flexibilidad conformacional y capacidad de unión con moléculas de señalización, mientras C-terminal al estar conformada por residuos hidrófobos estabiliza la estructura y aumenta su afinidad con otras proteínas **(Hamann et al., 2020; Voros et al., 2018)**.

En este sentido, se puede sugerir que el estado de protonación de los residuos presentes en las proteínas predichas permite experimentar cierto grado de inestabilidad durante su plegamiento, incluso, los residuos que se encuentran asociados con dicho efecto tienden a separarse de su estructura principal con mayor frecuencia cuando se encuentran con dos o más de estos estados **(Wilson et al., 2024)**, sin embargo, no se puede tomar como prueba irrefutable o concluyente de la estabilidad de dichas estructuras, ya que en varios estudios se ha determinado su papel en la heterogeneidad de los modelos lo que resulta ambiguo debido a que puede provocar inusuales interacciones entre aquellos que lo conducen a una total desestabilización de complejos diméricos a causa del drástico aumento en la diferencia de sus energías de interacción en comparación con contrapartes monoméricas **(Dogan & Durdagi, 2023; Medina et al., 2020)**.

Por otra parte, las energías computadas se presentan como una medida de mayor confiabilidad para el análisis de estabilidad, gracias a que representan a aquellas interacciones intra e intermoleculares que mantienen unidas a este tipo de estructuras complejas **(Olivieri et al., 2023)**, en donde, las interacciones de enlace de hidrogeno asociadas con la energía de unión entre sus enlaces covalentes y de van der Waals relacionadas con energías más débiles para mantener la estabilidad tienen un papel clave en estas conformaciones **(Grassmann et al., 2024; Nunes et al., 2023; Zhang et al., 2024)**. Además, si se incluyen los valores energéticos negativos de: CMAP (Corrección de mapas de energía) y Angular relacionados con los campos de fuerza de torsión, flexibilidad y plegamiento de las proteínas **(Galano-Frutos & Sancho, 2024; Y. Li et al., 2023)**, se puede inferir que cada uno de los modelos monoméricos obtenidos presenta una buena estabilidad en su conformación estructural.

Tabla 8.

Datos y propiedades energéticas de las proteínas predichas por las herramientas de Inteligencia Artificial.

IA	Códigos de Identificación	Datos básicos de la proteína						Energía Computada			
		Cadenas	Longitud	pH	Residuos Terminal	Protonación	Ion	Enlace	CMAP	VDWaals	Angular
AlphaFold	Q15653	A	356	7,0	NTER CTER	-	NaCl	544,93	-13,17	-727,46	1147,65
	Q8NI38	A	313	7,0	NTER CTER	-	NaCl	495,22	29,63	-1088,82	948,24
	Q9BYH8	A	718	7,0	NTER CTER	-	NaCl	1461,52	-241,28	-1354,91	3124,63
Deepchem	A26C8	A B	625	7,0	NTER CTER	A (384ARG–RN1; 561ASP–ASPP; 5ASP–ASPP)	NaCl	4574,76	229,09	$0,18x10^{14}$	7053,08
	208E3	A	625	7,0	NTER CTER	-	NaCl	1460,05	-269,79	$7,45x10^{8}$	2974,29
	9061D	A	435	7,0	NTER CTER	-	NaCl	1019,61	-112,79	$5,95x10^{6}$	1856,22
RFdiffusion	RF715A1	A	433	7,0	GLYP CTER	A (229LYS–LSN; 30ASP–ASPP)	NaCl	$3,10x10^{5}$	-33,67	$7,24x10^{6}$	1452,76
	RF315BD	A B	544	7,0	NTER CTER	-	NaCl	65931,43	-62,22	$4,17x10^{10}$	4889,89
	RF2M31	A B C	138	7,0	NTER CTER	A (30CYS–CYM; 4ASP–ASPP)	NaCl	1088,81	-47,44	$0,68x10^{12}$	998,25

Nota: Dentro de la presente tabla y Anexo 4, el estado de protonación expresado en la mayoría de los modelos diméricos predichos representa la presencia o ausencia de protones (H^+) en cada uno de los residuos descritos, en donde, se incluye la forma que adopta tras dicho proceso. Por lo tanto, se puede deducir que las herramientas de inteligencia artificial empleadas aún no se encuentran familiarizadas con el desarrollo de conformaciones más complejas.

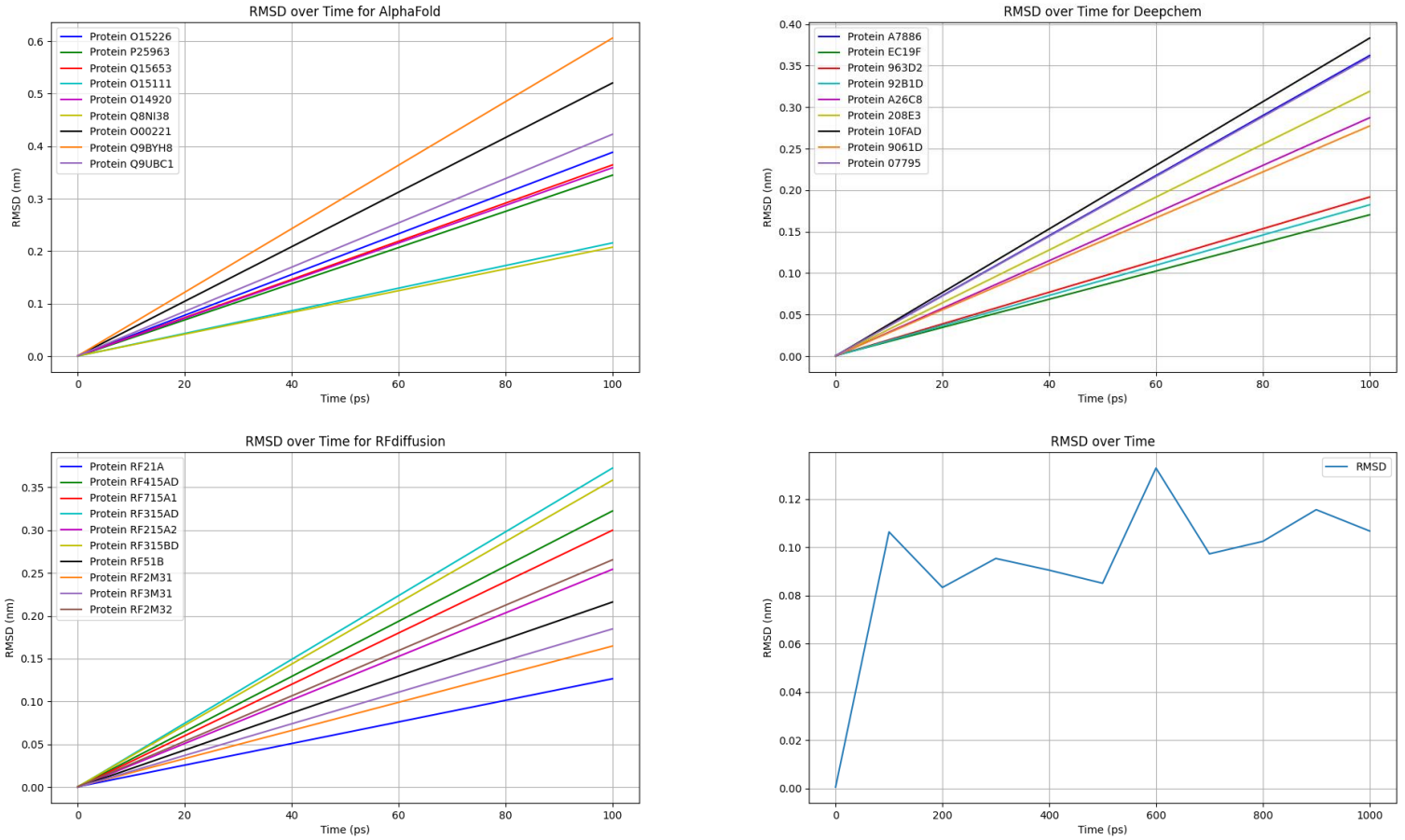


Figura 10. Gráficas RMSD de estabilidad de las proteínas predichas.

RMSD (Raíz de la desviación cuadrática media) muestra el cambio producido dentro de la conformación estructural de las proteínas predichas en un tiempo de 100 picosegundos debido a la limitada capacidad computacional. En donde, a) la simulación de los modelos predichos por AlphaFold presenta solo 9 proteínas sin considerar a Q9Y6K9 y b) la simulación de los modelos otorgados por Deepchem muestra 9 proteínas exceptuando 533E3, mientras tanto, c) la simulación de los modelos propuestos por RFdiffusion indica 10 proteínas junto a un d) análisis RMSD completo de la proteína RF21A durante una simulación de 1000 picosegundos.

Así mismo, sus contrapartes diméricas necesitan un estudio más amplio para llegar a una adecuada conclusión, por lo que, se incluyó a los valores de RMSD (desviaciones cuadráticas medias) visualizados en la Figura 10 como un análisis extra acerca de la estabilidad de las proteínas predichas, al ser considerado como un indicador útil del estado de la estructura durante su plegamiento en las simulaciones realizadas **(Irumagawa et al., 2021)**. Por esta razón, varios estudios consideran al estado estable de una proteína dentro de un rango de fluctuación de entre 0,2 a 1,0 nm tras llegar a su equilibrio a los 100 o 200 picosegundos hasta finalizar los 150 a 200 nanosegundos de dinámica molecular **(Arshia et al., 2021; Maruyama et al., 2023; Xing et al., 2023)**, por tanto, al considerar dicha información se puede concluir que todas las proteínas predichas que fluctuación no mayor a 0,6 nm son estables durante los primeros 100 picosegundos.

3.1.4. Comparación de las variables termodinámicas presentes en la interacción proteína-proteína

Luego de haber completado el acoplamiento molecular entre las proteínas predichas y los sitios de unión de los genes observados en el Anexo 5, una de las mejores maneras de identificar si dicho procedimiento fue favorable es mediante la medición de variables termodinámicas que evalúen las características estructurales obtenidas en el complejo proteína-proteína **(Singh et al., 2022)**, por lo tanto, el estudio y comparación de los valores de capacidad de interacción o unión proteína-proteína junto a los de estabilidad y flexibilidad del acoplamiento, presentados en la Tabla 9 y Anexo 6 son adecuados para confirmar este planteamiento.

Tabla 9.

Datos acerca de las variables termodinámicas obtenidas tras el proceso de Acoplamiento Molecular.

Docking y Acoplamiento Molecular		**Variables Termodinámicas obtenidas**				
Receptor	**Mejor ligando**	**ΔG (kcal/mol)**	**K_d (pM)**	**ΔH (kcal/mol)**	**C_p (cal/mol*K)**	**ΔS (cal/mol*K)**
Heterodímero NF-κB p50/p65 (Control)	Inhibidor IκBα (Control)	-15,11	8,4	-30	-119,5	50

Docking y Acoplamiento Molecular		Variables Termodinámicas obtenidas				
Receptor	**Mejor ligando**	**ΔG (kcal/mol)**	**K_d (pM)**	**ΔH (kcal/mol)**	**C_p (cal/mol*K)**	**ΔS (cal/mol*K)**
Quinasa inductora de NF- κB (NIK)	Q9BYH8	-15,59	3,7	-30	-119,5	48,3
Subunidad p100 (NFκB2)	Q9BYH8	-18,07	0,05	-30	-119,5	40,0
Interleucina-8	10FAD	-11.93	1799,9	-30	-119,5	60,6

Nota: Los valores descritos en la presente tabla y en el Anexo 6, se refieren a las siguientes variables termodinámicas: Energía libre de Gibbs (ΔG), Constante de disociación (K_d), Entalpía (ΔH), Capacidad calorífica (C_p) y Entropía (ΔS).

En primer lugar, los valores de control disponibles pertenecen al complejo altamente estudiado IκBα/NF-κB, que a través de su acoplamiento e interacción específica produce una inhibición efectiva de la actividad de NF-κB **(Huxford et al., 1998)**, por lo tanto, se presenta como punto de referencia para los nuevos prospectos a inhibidores. De este modo, se puede considerar que aquellos valores de ΔG y K_d conocidos por estar asociados con los fenómenos de interacción proteína-proteína, proteína-ligando o proteína-ADN, tienden a formar acoplamientos con mayor fuerza en complejos espontáneos con valores más negativos o menores **(Ferraz et al., 2023)**, y conforme a esta información, es factible decir que todo acoplamiento obtenido con valores de mayores a -15,11 kcal/mol de ΔG y 8,4 pM de K_d presentan interacciones proteína-proteínas menos favorables o viceversa.

Mientras tanto, aquellos valores restantes se encuentras asociados con la estabilidad y plegado de los acoplamientos, en donde, el valor de -30 kcal/mol obtenido en ΔH junto al valor de -119,5 cal/mol*K de C_p reflejan en cada uno de complejos proteicos un cierto grado de termo-estabilidad (25 ºC) durante plegamiento de su estructura (Kordes et al., 2022; Romero-Romero et al., 2021). En cambio, ΔS al representar el desorden y la espontaneidad se puede interpretar que los valores que sean mayores a 50 cal/mol*K en dicha variable presentarán un mayor desorden en su estructura y viceversa **(Seelig & Seelig, 2023)**, dando a entender que la mayoría

de los acoplamientos evaluados tienen una buena predisposición en función del complejo IκBα/NF-κB.

3.1.5. Afinidad y selectividad de las proteínas predichas

En relación con el apartado anterior, se destaca la importancia de impulsar el desarrollo de interacciones proteína-proteína con altos rangos de afinidad y especificidad, ya que, impulsan activamente la generación y descubrimiento de nuevos tratamientos o fármacos especializados en la inhibición de los múltiples factores que participan en el cáncer gástrico junto a nuevos métodos computacionales que simplifican dicho proceso **(Soleymani et al., 2022)**.

Consecuentemente, existe la necesidad de seleccionar a aquellos modelos predichos con IA con mayor afinidad y selectividad hacia los diferentes genes expuestos con antelación, por lo tanto, primero se evaluó la afinidad obtenida durante estas interacciones proteína-proteína y se visualizó a las mismas dentro de la Tabla 10 y Anexo 7.

Tabla 10.

Datos acerca de la afinidad de las proteínas predichas (ligandos).

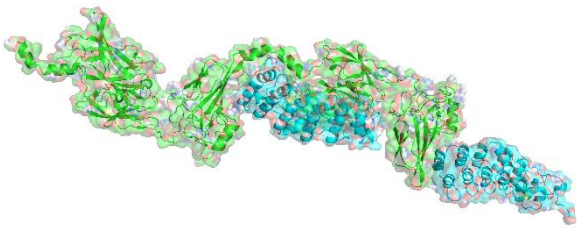
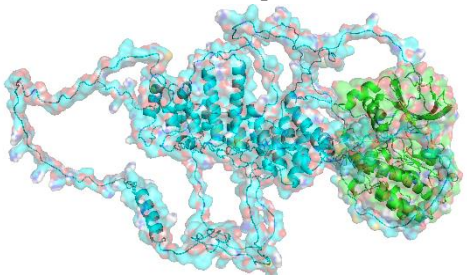
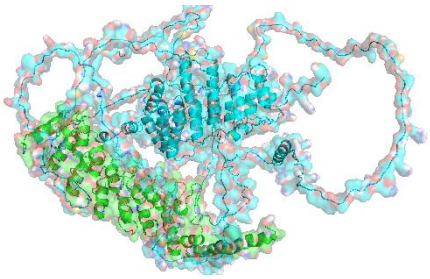
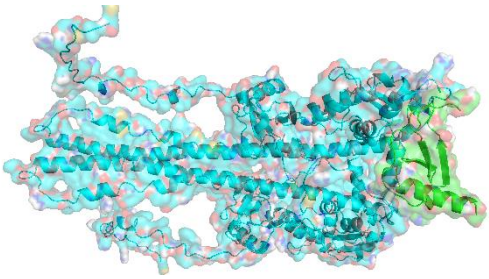
Receptor	Proteínas predichas (ligandos)	Afinidad (kcal/mol)	Figura
Heterodímero NF-κB p50/p65 (Control)	Inhibidor IκBα (Control)	-15,11	
Quinasa inductora de NF- κB (NIK)	Q9BYH8	-15,59	

Figura 11. *Acoplamiento obtenido de la interacción entre las proteínas control.*

Figura 12. *Acoplamiento obtenido de la interacción entre NIK y Q9BYH8.*

Receptor	Proteínas predichas (ligandos)	Afinidad (kcal/mol)	Figura
Subunidad p100 (NFκB2)	Q9BYH8	-18,07	 ***Figura 13.*** *Acoplamiento obtenido de la interacción entre la subunidad p100 y Q9BYH8.*
Interleucina-8	10FAD	-11,93	 ***Figura 14.*** *Acoplamiento obtenido de la interacción entre IL-8 y 10FAD.*

Nota: Las figuras observadas en la presente tabla y Anexo 7, representan la interacción proteína- proteína entre los genes regulados por el factor NF-κB (verde) y las proteínas predichas con mayor afinidad (celeste) a dichos genes.

De acuerdo con los mencionado por **(Siebenmorgen & Zacharias, 2020)**, se puede deducir que tanto la afinidad de unión como la energía libre de unión (ΔG) se encuentran asociadas a las interacciones proteína-proteína y por ende sus valores llegan a ser iguales durante una evaluación de afinidades. En consecuencia, la estructura del complejo IκBα/NF-κB vista en la Figura 11 junto a su energía libre de unión (ΔG) serán nuevamente el punto de referencia para los demás complejos obtenidos, dando como resultado que, la proteína Q9BYH8 sea altamente afín en complejos como NIK y NFκB2(p100) observados tanto en la Figura 12 como Figura 13, mientras que en complejos como el de la Figura 14 se observa que existe una menor afinidad.

Al considerar que una buena interacción proteína-proteína presenta tanto afinidad como selectividad, esta última aún permanece casi inexplorada debido a que solo se conoce su preferencia por estructuras ordenadas **(Shu et al., 2023)**, la evaluación correspondiente a la Tabla 11 se realizó de forma cualitativa. Donde, se determinó la cantidad de complejos que forma cada una de las proteínas empleadas durante el acoplamiento e interacción, infiriendo

que los modelos Q9UBC1, 963D2 y 10FAD llegan a formar complejos altamente selectivos con: NFκB1 (p50), RelA (p65) y IL-8 respectivamente.

Tabla 11.

Identificación de la selectividad de las proteínas predichas (ligandos).

Proteína predicha (Ligando)	**Receptores con los que interactúa**	**Interpretación de Selectividad**
Q9BYH8	1. Quinasa Inductora de NF-κB 2. Subunidad p52 (NFκB2) 3. Subunidad p100 (NFκB2) 4. CiclinaD1 (CD1)	Poco selectivo
A26C8	1. Subunidad p105 (NFκB1) 2. Interleucina-B (IL-B)	Levemente selectivo
O00221	1. Ciclooxigenasa-2 (COX-2) 2. STAT3	Levemente selectivo
Q9UBC1	1. Subunidad p50 (NFκB1)	Selectivo
963D2	1. Subunidad p65 (RelA)	Selectivo
10FAD	1. Interleucina-8 (IL-8)	Selectivo

En conclusión, podemos afirmar que las diferentes proteínas *de novo* predichas por las herramientas de Inteligencia Artificial como: AlphaFold, Deepchem y RFdiffusion son absolutamente viables y estables para el desarrollo de tratamientos destinados a la interacción y posterior inhibición de genes y factores de transcripción regulados por NF-κB dentro del cáncer gástrico.

CAPITULO 4

CONCLUSIONES Y RECOMENDACIONES

4.1. Conclusiones

A partir de las herramientas de inteligencia artificial AlphaFold, Deepchem y RFdiffusion se predijeron 30 proteínas *de novo* con conformación Bucle-Hélice-Bucle (LHL) y una composición similar a inhibidores y reguladores del factor NF-κB presente en varios conjuntos de bases de datos estándar y experimental, en donde, se presentaron principalmente: dominios de repeticiones de Ankyrin, dominios de Dimerización, dominios similares a ubiquitina (Ubl) y dominios catalíticos serina/treonina quinasa junto a regiones artificiales o desordenadas que permiten mejorar sus interacciones proteína-proteína. También, se analizó los valores energéticos de las interacciones moleculares y valores de RMSD, concluyendo que todas las proteínas predichas son estables durante un tiempo de 100 picosegundos.

Mediante el acoplamiento y docking molecular de las proteínas predichas con los genes regulados por el factor NF-κB presentes en el cáncer gástrico, se obtuvieron valores correspondientes a las variables termodinámicas tales como: Energía libre de Gibbs (ΔG), Constante de disociación (K_d), Entalpía (ΔH), Capacidad calorífica (C_p) y Entropía (ΔS), las cuales comparadas con los valores impuestos por el complejo IκBα/NF-κB (control), permiten inferir que la mayoría de complejos predichos tienen una mejor interacción proteína-proteína y plegamiento en comparación con el control sugerido.

En las 11 interacciones proteínas-proteína y el complejo IκBα/NF-κB (control), se evaluaron los valores de afinidad mediante el análisis de la Energía libre de Gibbs (ΔG) y la selectividad mostrada por las proteínas *de novo* durante el acoplamiento molecular. Mostrando como resultado que los modelos predichos por AlphaFold son más afines y que los modelos predichos por Deepchem son más selectivos en sus interacciones con los genes y factores de transcripción regulados por NF-κB.

4.2. Recomendaciones

Analizar los valores de RMSD durante los tiempos correspondientes a una simulación de dinámica molecular completa (150 a 200 nanosegundos) con la intención de mejorar la comprensión acerca de la capacidad de plegamiento y estabilidad de las proteínas predichas.

Plantear diferentes variables termodinámicas para contrastar y verificar los resultados obtenidos en relación con sus interacciones proteína-proteína.

Validar los valores de afinidad mediante el uso de softwares o servidores web destinados exclusivamente al acoplamiento y docking molecular entre proteínas, por ejemplo, Cluspro.

Probar la metodología expuesta dentro de un servidor con una mayor capacidad de cómputo para obtener mejores resultados con la intención de que puedan ser escalados a una etapa de prediseño.

REFERENCIAS BIBLIOGRÁFICAS

Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., & Lindah, E. (2015). Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX 1–2, 1*(2015), 19–25. https://doi.org/10.1016/j.softx.2015.06.001

Agnes, A., Biondi, A., Laurino, A., Persiani, R., & D'Ugo, D. (2020). Global updates in the treatment of gastric cancer: a systematic review. Part 1: staging, classification and surgical treatment. In *Updates in Surgery*, *72*(2), 341–353. https://doi.org/10.1007/s13304-020-00736-3

Albensi, B. C. (2019). What is nuclear factor kappa B (NF-κB) doing in and to the mitochondrion? In *Frontiers in Cell and Developmental Biology*, *7*(JULY), 1–7. https://doi.org/10.3389/fcell.2019.00154

Arshia, A. H., Shadravan, S., Solhjoo, A., Sakhteman, A., & Sami, A. (2021). De novo design of novel protease inhibitor candidates in the treatment of SARS-CoV-2 using deep learning, docking, and molecular dynamic simulations. *Computers in Biology and Medicine*, *139*(104967), 104967. https://doi.org/10.1016/j.compbiomed.2021.104967

Ashrafizadeh, M., Zarrabi, A., Orouei, S., Zarrin, V., Moghadam, E. R., Zabolian, A., Mohammadi, S., Hushmandi, K., Gharehaghajlou, Y., Makvandi, P., Najafi, M., & Mohammadinejad, R. (2020). STAT3 pathway in gastric cancer: Signaling, therapeutic targeting and future prospects. In *Biology*, *9*(6), 126. https://doi.org/10.3390/biology9060126

Baker, D. (2019). What has de novo protein design taught us about protein folding and biophysics? In *Protein Science*, *28*(4), 678–683. https://doi.org/10.1002/pro.3588

Banach, M., Fabian, P., Stapor, K., Konieczny, L., & Roterman, I. (2020). Structure of the hydrophobic core determines the 3d protein structure—verification by single mutation proteins. *Biomolecules*, *10*(5), 767. https://doi.org/10.3390/biom10050767

Barnabei, L., Laplantine, E., Mbongo, W., Rieux-Laucat, F., & Weil, R. (2021). NF-κB: At the Borders of Autoimmunity and Inflammation. In *Frontiers in Immunology*, *12*(2021), 716469. https://doi.org/10.3389/fimmu.2021.716469

Baughman, H. E. R., Narang, D., Chen, W., Villagrán Suárez, A. C., Lee, J., Bachochin, M. J., Gunther, T. R., Wolynes, P. G., & Komives, E. A. (2022). An intrinsically disordered transcription activation domain increases the DNA binding affinity and reduces the specificity of NFκB p50/RelA. *Journal of Biological Chemistry*, *298*(9), 102349. https://doi.org/10.1016/j.jbc.2022.102349

Bellamkonda, S. V. (2023). *Applications of DeepChem with Alphafold: Docking and protein-ligand interaction from protein sequence*. DeepChemXAlphafold. https://deepchem.io/tutorials/deepchemxalphafold/

Bernstein, P. (2022). What is Artificial Intelligence (AI)? In *Machine Learning*. https://doi.org/10.4324/9781003297192-3

Bittrich, S., Bhikadiya, C., Bi, C., Chao, H., Duarte, J. M., Dutta, S., Fayazi, M., Henry, J., Khokhriakov, I., Lowe, R., Piehl, D. W., Segura, J., Vallat, B., Voigt, M., Westbrook, J. D., Burley, S. K., & Rose, Y. (2023). RCSB Protein Data Bank: Efficient Searching and Simultaneous Access to One Million Computed Structure Models Alongside the PDB Structures Enabled by Architectural Advances: RCSB Protein Data Bank: Architectural Advances. *Journal of Molecular Biology*, *435*(14), 167994. https://doi.org/10.1016/j.jmb.2023.167994

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983). CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *Journal of Computational Chemistry*, *4*(2), 187–217. https://doi.org/10.1002/jcc.540040211

Cao, L., Coventry, B., Goreshnik, I., Huang, B., Sheffler, W., Park, J. S., Jude, K. M., Marković, I., Kadam, R. U., Verschueren, K. H. G., Verstraete, K., Walsh, S. T. R., Bennett, N., Phal, A., Yang, A., Kozodoy, L., DeWitt, M., Picton, L., Miller, L., … Baker, D. (2022). Design of protein-binding proteins from the target structure alone. *Nature*, *605*(7910), 551–560. https://doi.org/10.1038/s41586-022-04654-9

Carino, A., Graziosi, L., Marchianò, S., Biagioli, M., Marino, E., Sepe, V., Zampella, A., Distrutti, E., Donini, A., & Fiorucci, S. (2021). Analysis of Gastric Cancer Transcriptome Allows the Identification of Histotype Specific Molecular Signatures With Prognostic Potential. *Frontiers in Oncology*, *11*(2021), 1–17. https://doi.org/10.3389/fonc.2021.663771

Chaithongyot, S., Jantaree, P., Sokolova, O., & Naumann, M. (2021). NF-κb in gastric cancer development and therapy. *Biomedicines*, *9*(8), 870. https://doi.org/10.3390/biomedicines9080870

Chakrabarty, B., & Parekh, N. (2022). Sequence and Structure-Based Analyses of Human Ankyrin Repeats. *Molecules*, *27*(2), 423. https://doi.org/10.3390/molecules27020423

Chang, J., Chapman, B., Friedberg, I., Hamelryck, T., de Hoon, M., Cock, P., Antao, T., Talevich, E., & Wilczyński, B. (2024, January 10). *Biopython Tutorial and Cookbook*. Biopython. https://biopython.org/DIST/docs/tutorial/Tutorial.html

Chopra, H., Baig, A. A., Gautam, R. K., & Kamal, M. A. (2022). Application of Artificial Intelligence in Drug Discovery. *Current Pharmaceutical Design*, *28*(33), 2690–2703. https://doi.org/10.2174/1381612828666220608141049

Cleveland Clinic. (2022, May 17). *Stomach Cancer: Causes, Symptoms, Diagnosis & Treatment*. Cleveland Clinic. https://my.clevelandclinic.org/health/diseases/15812-stomach-cancer

Cristescu, R., Lee, J., Nebozhyn, M., Kim, K. M., Ting, J. C., Wong, S. S., Liu, J., Yue, Y. G., Wang, J., Yu, K., Ye, X. S., Do, I. G., Liu, S., Gong, L., Fu, J., Jin, J. G., Choi, M. G., Sohn, T. S., Lee, J. H., … Aggarwal, A. (2015). Molecular analysis of gastric cancer identifies subtypes associated with distinct clinical outcomes. *Nature Medicine*, *21*(5), 449–456. https://doi.org/10.1038/nm.3850

Deeks, H. M., Zinovjev, K., Barnoud, J., Mulholland, A. J., van der Kamp, M. W., & Glowacki, D. R. (2023). Free energy along drug-protein binding pathways interactively sampled in virtual reality. *Scientific Reports 2023 13:1*, *13*(1), 1–9. https://doi.org/10.1038/s41598-023-43523-x

Del Alamo, D., Sala, D., McHaourab, H. S., & Meiler, J. (2022). TITLE: Sampling alternative conformational states of transporters and receptors with AlphaFold2. *eLife*, *11*(2022), e75751. https://doi.org/10.7554/eLife.75751

Del Arco, C. D., Ortega Medina, L., Estrada Muñoz, L., Molina Roldán, E., Cerón Nieto, M. Á., García Gómez de las Heras, S., & Fernández Aceñero, M. J. (2021). Are borrmann's types of advanced gastric cancer distinct clinicopathological and molecular entities? A western study. *Cancers*, *13*(12), 3081. https://doi.org/10.3390/cancers13123081

Desta, I. T., Porter, K. A., Xia, B., Kozakov, D., & Vajda, S. (2020). Performance and Its Limits in Rigid Body Protein-Protein Docking. *Structure*, *28*(9), 1071–1081. https://doi.org/10.1016/j.str.2020.06.006

Dhakal, A., McKay, C., Tanner, J. J., & Cheng, J. (2022). Artificial intelligence in the prediction of protein-ligand interactions: recent advances and future directions. In *Briefings in Bioinformatics, 23*(1), 1–23. https://doi.org/10.1093/bib/bbab476

Dogan, B., & Durdagi, S. (2023). Investigating the Effect of GLU283 Protonation State on the Conformational Heterogeneity of CCR5 by Molecular Dynamics Simulations. *BioRxiv*, 8(2023), 553662. https://www.biorxiv.org/content/10.1101/2023.08.17.553662v1

Dreier, J. E., Prestel, A., Martins, J. M., Brøndum, S. S., Nielsen, O., Garbers, A. E., Suga, H., Boomsma, W., Rogers, J. M., Hartmann-Petersen, R., & Kragelund, B. B. (2022). A context-

dependent and disordered ubiquitin-binding motif. *Cellular and Molecular Life Sciences*, *79*(9), 484. https://doi.org/10.1007/s00018-022-04486-w

Eberhardt, J., Santos-Martins, D., Tillack, A. F., & Forli, S. (2021). AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings. *Journal of Chemical Information and Modeling*, *61*(8), 3891–3898. https://doi.org/10.1021/acs.jcim.1c00203

Echizen, K., Horiuchi, K., Aoki, Y., Yamada, Y., Minamoto, T., Oshima, H., & Oshima, M. (2019). NF-κB-induced NOX1 activation promotes gastric tumorigenesis through the expansion of SOX2-positive epithelial cells. *Oncogene*, *38*(22), 4250–4263. https://doi.org/10.1038/s41388-019-0702-0

Eusebi, L. H., Telese, A., Marasco, G., Bazzoli, F., & Zagari, R. M. (2020). Gastric cancer prevention strategies: A global perspective. In *Journal of Gastroenterology and Hepatology (Australia), 35*(9), 1495–1502. https://doi.org/10.1111/jgh.15037

Ferraz, M. V. F., Neto, J. C. S., Lins, R. D., & Teixeira, E. S. (2023). An artificial neural network model to predict structure-based protein-protein free energy of binding from Rosetta-calculated properties. *Physical Chemistry Chemical Physics*, *25*(10), 7257–7267. https://doi.org/10.1039/d2cp05644e

Fulcher, L. J., & Sapkota, G. P. (2020). Functions and regulation of the serine/threonine protein kinase CK1 family: Moving beyond promiscuity. In *Biochemical Journal, 477*(23), 4603–4621. https://doi.org/10.1042/BCJ20200506

Galano Frutos, J. J., & Sancho, J. (2024). Energy, water, and protein folding: A molecular dynamics-based quantitative inventory of molecular interactions and forces that make proteins stable. *Protein Science*, *33*(2), e4905. https://doi.org/10.1002/pro.4905

Glaeser, J. D., Salehi, K., Kanim, L. E. A., NaPier, Z., Kropf, M. A., Cuéllar, J. M., Perry, T. G., Bae, H. W., & Sheyn, D. (2020). NF-κB inhibitor, NEMO-binding domain peptide attenuates intervertebral disc degeneration. *Spine Journal*, *20*(9), 1480–1491. https://doi.org/10.1016/j.spinee.2020.04.025

Grassmann, G., Miotto, M., Desantis, F., Di Rienzo, L., Tartaglia, G. G., Pastore, A., Ruocco, G., Monti, M., & Milanetti, E. (2024). Computational Approaches to Predict Protein–Protein Interactions in Crowded Cellular Environments. *Chemical Reviews*, *124*(7), 3932–3977. https://doi.org/10.1021/acs.chemrev.3c00550

Gupta, R., Srivastava, D., Sahu, M., Tiwari, S., Ambasta, R. K., & Kumar, P. (2021). Artificial intelligence to deep learning: machine intelligence approach for drug discovery. *Molecular Diversity*, *25*(3), 1315–1360. https://doi.org/10.1007/s11030-021-10217-3

Hamann, F., Schmitt, A., Favretto, F., Hofele, R., Neumann, P., Xiang, S. Q., Urlaub, H., Zweckstetter, M., & Ficner, R. (2020). Structural analysis of the intrinsically disordered splicing factor Spp2 and its binding to the DEAH-box ATPase Prp2. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(6), 2948–2956. https://doi.org/10.1073/pnas.1907960117

Hassani, H., Silva, E. S., Unger, S., TajMazinani, M., & Mac Feely, S. (2020). Artificial Intelligence (AI) or Intelligence Augmentation (IA): What Is the Future? *AI (Switzerland)*, *1*(2), 143–155. https://doi.org/10.3390/ai1020008

Haure Mirande, J. V., Audrain, M., Ehrlich, M. E., & Gandy, S. (2022). Microglial TYROBP/DAP12 in Alzheimer's disease: Transduction of physiological and pathological signals across TREM2. In *Molecular Neurodegeneration, 17*(1), 55 (2022). https://doi.org/10.1186/s13024-022-00552-w

Holdgate, G. A., & Ward, W. H. J. (2005). Measurements of binding thermodynamics in drug discovery. In *Drug Discovery Today, 10*(22), 1543–1550. https://doi.org/10.1016/S1359-6446(05)03610-X

Huang, T., & Li, Y. (2023). Current progress, challenges, and future perspectives of language models for protein representation and protein design. In *Innovation, 4*(4), 100446. https://doi.org/10.1016/j.xinn.2023.100446

Hulsen, T. (2021). Literature analysis of artificial intelligence in biomedicine. *Pharmacogenomics Research and Personalized Medicine*, *10*(23), 1284. https://doi.org/10.21037/prpm-21-15

Huxford, T., Huang, D. Bin, Malek, S., & Ghosh, G. (1998). The crystal structure of the IκBα/NF-κB complex reveals mechanisms of NF-κB inactivation. *Cell*, *95*(6), 759–770. https://doi.org/10.1016/S0092-8674(00)81699-2

Irumagawa, S., Kobayashi, K., Saito, Y., Miyata, T., Umetsu, M., Kameda, T., & Arai, R. (2021). Rational thermostabilisation of four-helix bundle dimeric de novo proteins. *Scientific Reports*, *11*(1), 7526 (2021). https://doi.org/10.1038/s41598-021-86952-2

Ivarsson, Y., & Jemth, P. (2019). Affinity and specificity of motif-based protein–protein interactions. In *Current Opinion in Structural Biology, 54*(1), 26–33. https://doi.org/10.1016/j.sbi.2018.09.009

Jo, S., Kim, T., Iyer, V. G., & Im, W. (2008). CHARMM-GUI: A web-based graphical user interface for CHARMM. *Journal of Computational Chemistry*, *29*(11), 1859–1865. https://doi.org/10.1002/jcc.20945

Johnston, F. M., & Beckman, M. (2019). Updates on Management of Gastric Cancer. In *Current Oncology Reports, 21*(8), 67. https://doi.org/10.1007/s11912-019-0820-4

Joshi, S. S., & Badgwell, B. D. (2021). Current treatment and recent progress in gastric cancer. *CA: A Cancer Journal for Clinicians*, *71*(3), 264–279. https://doi.org/10.3322/caac.21657

Kaul, V., Enslin, S., & Gross, S. A. (2020). History of artificial intelligence in medicine. In *Gastrointestinal Endoscopy, 92*(4), 807–812. https://doi.org/10.1016/j.gie.2020.06.040

Kintzing, J. R., Filsinger Interrante, M. V., & Cochran, J. R. (2016). Emerging Strategies for Developing Next-Generation Protein Therapeutics for Cancer Treatment. In *Trends in Pharmacological Sciences, 37*(12), 993–1008. https://doi.org/10.1016/j.tips.2016.10.005

Kordes, S., Romero-Romero, S., Lutz, L., & Höcker, B. (2022). A newly introduced salt bridge cluster improves structural and biophysical properties of de novo TIM barrels. *Protein Science*, *31*(2), 513–527. https://doi.org/10.1002/pro.4249

Korendovych, I. V., & DeGrado, W. F. (2020). De novo protein design, a retrospective. In *Quarterly Reviews of Biophysics*, 53, e3, 1–33. https://doi.org/10.1017/S0033583519000131

Kortemme, T. (2024). De novo protein design—From new structures to programmable functions. In *Cell, 187*(3), 526–544. https://doi.org/10.1016/j.cell.2023.12.028

Kuhlman, B., & Bradley, P. (2019). Advances in protein structure prediction and design. In *Nature Reviews Molecular Cell Biology, 20*(11), 681–697. https://doi.org/10.1038/s41580-019-0163-x

Kulikowski, C. A. (2019). Beginnings of Artificial Intelligence in Medicine (AIM): Computational Artifice Assisting Scientific Inquiry and Clinical Art - with Reflections on Present AIM Challenges. *Yearbook of Medical Informatics*, *28*(1), 249–256. https://doi.org/10.1055/s-0039-1677895

Kushima, R. (2022). The updated WHO classification of digestive system tumours—gastric adenocarcinoma and dysplasia. In *Pathologe, 43*(1), 8–15. https://doi.org/10.1007/s00292-021-01023-7

Li, E. H., Spaman, L. E., Tejero, R., Janet Huang, Y., Ramelot, T. A., Fraga, K. J., Prestegard, J. H., Kennedy, M. A., & Montelione, G. T. (2023). Blind assessment of monomeric AlphaFold2 protein

structure models with experimental NMR data. *Journal of Magnetic Resonance*, *352*(107481), 107481. https://doi.org/10.1016/j.jmr.2023.107481

Li, K., Zhang, A., Li, X., Zhang, H., & Zhao, L. (2021). Advances in clinical immunotherapy for gastric cancer. In *Biochimica et Biophysica Acta - Reviews on Cancer, 1876*(2), 188615. https://doi.org/10.1016/j.bbcan.2021.188615

Li, Y., Fan, Z., Rao, J., Chen, Z., Chu, Q., Zheng, M., & Li, X. (2023). An overview of recent advances and challenges in predicting compound-protein interaction (CPI). In *Medical Review, 3*(6), 465–486. https://doi.org/10.1515/mr-2023-0030

Lin, E., Lin, C. H., & Lane, H. Y. (2020). Relevant applications of generative adversarial networks in drug design and discovery: Molecular de novo design, dimensionality reduction, and de novo peptide and protein design. In *Molecules, 25*(14), 3250. https://doi.org/10.3390/molecules25143250

Liu, S., Wu, K., & Chen, C. (2022). Obtaining protein foldability information from computational models of AlphaFold2 and RoseTTAFold. *Computational and Structural Biotechnology Journal*, *20*(2022), 4481–4489. https://doi.org/10.1016/j.csbj.2022.08.034

López, D. J., Rodríguez, J. A., & Bañuelos, S. (2021). Molecular mechanisms regulating the dna repair protein ape1: A focus on its flexible n-terminal tail domain. In *International Journal of Molecular Sciences, 22*(12), 6308. https://doi.org/10.3390/ijms22126308

López, M. J., Carbajal, J., Alfaro, A. L., Saravia, L. G., Zanabria, D., Araujo, J. M., Quispe, L., Zevallos, A., Buleje, J. L., Cho, C. E., Sarmiento, M., Pinto, J. A., & Fajardo, W. (2023). Characteristics of gastric cancer around the world. In *Critical Reviews in Oncology/Hematology, 181*(JANUARY), 103841. https://doi.org/10.1016/j.critrevonc.2022.103841

Lyons, K., Le, L. C., Pham, Y. T. H., Borron, C., Park, J. Y., Tran, C. T. D., Tran, T. V., Tran, H. T. T., Vu, K. T., Do, C. D., Pelucchi, C., La Vecchia, C., Zgibor, J., Boffetta, P., & Luu, H. N. (2019). Gastric cancer: Epidemiology, biology, and prevention: A mini review. In *European Journal of Cancer Prevention, 28*(5), 397–412. https://doi.org/10.1097/CEJ.0000000000000480

Machlowska, J., Baj, J., Sitarz, M., Maciejewski, R., & Sitarz, R. (2020). Gastric cancer: Epidemiology, risk factors, classification, genomic characteristics and treatment strategies. In *International Journal of Molecular Sciences, 21*(11), 4012. https://doi.org/10.3390/ijms21114012

Marcos, E., Chidyausiku, T. M., McShan, A. C., Evangelidis, T., Nerli, S., Carter, L., Nivón, L. G., Davis, A., Oberdorfer, G., Tripsianes, K., Sgourakis, N. G., & Baker, D. (2018). De novo design

of a non-local β-sheet protein with high stability and accuracy. *Nature Structural and Molecular Biology*, *25*(11), 1028–1034. https://doi.org/10.1038/s41594-018-0141-6

Marta, Ż. N., Agnieszka, W., Jacek, P., Jeleń, A., Adrian, K., Dagmara, S. K., Sałagacka-Kubiak, A., & Balcerczak, E. (2020). NFKB2 gene expression in patients with peptic ulcer diseases and gastric cancer. *Molecular Biology Reports*, *47*(3), 2015–2021. https://doi.org/10.1007/s11033-020-05299-5

Martinez, X., Chavent, M., & Baaden, M. (2020). Visualizing protein structures - Tools and trends. In *Biochemical Society Transactions, 48*(2), 499–506. https://doi.org/10.1042/BST20190621

Maruyama, Y., Igarashi, R., Ushiku, Y., & Mitsutake, A. (2023). Analysis of Protein Folding Simulation with Moving Root Mean Square Deviation. *Journal of Chemical Information and Modeling*, *63*(5), 1529–1541. https://doi.org/10.1021/acs.jcim.2c01444

Medeiros, M., Candido, M. F., Valera, E. T., & Brassesco, M. S. (2021). The multifaceted NF-kB: are there still prospects of its inhibition for clinical intervention in pediatric central nervous system tumors? In *Cellular and Molecular Life Sciences, 78*(17–18), 6161–6200. https://doi.org/10.1007/s00018-021-03906-7

Medina, E., Villalobos, P., Coñuecar, R., Ramírez-Sarmiento, C. A., & Babul, J. (2020). Author Correction: The protonation state of an evolutionarily conserved histidine modulates domain swapping stability of FoxP1. In *Scientific Reports, 10*(1), 15986 (2020). https://doi.org/10.1038/s41598-020-70337-y

Melo, M. C. R., Maasch, J. R. M. A., & de la Fuente-Nunez, C. (2021). Accelerating antibiotic discovery through artificial intelligence. In *Communications Biology, 4*(1), 1050. https://doi.org/10.1038/s42003-021-02586-0

Michel, M., Visnes, T., Homan, E. J., Seashore-Ludlow, B., Hedenström, M., Wiita, E., Vallin, K., Paulin, C. B. J., Zhang, J., Wallner, O., Scobie, M., Schmidt, A., Jenmalm-Jensen, A., Warpman Berglund, U., & Helleday, T. (2019). Computational and Experimental Druggability Assessment of Human DNA Glycosylases. *ACS Omega*, *4*(7), 11642–11656. https://doi.org/10.1021/acsomega.9b00162

Mirzaei, S., Zarrabi, A., Hashemi, F., Zabolian, A., Saleki, H., Ranjbar, A., Seyed Saleh, S. H., Bagherian, M., Sharifzadeh, S. omid, Hushmandi, K., Liskova, A., Kubatka, P., Makvandi, P., Tergaonkar, V., Kumar, A. P., Ashrafizadeh, M., & Sethi, G. (2021). Regulation of Nuclear Factor-KappaB (NF-κB) signaling pathway by non-coding RNAs in cancer: Inhibiting or promoting

carcinogenesis? In *Cancer Letters, 509*(2021), 63–80. https://doi.org/10.1016/j.canlet.2021.03.025

Mondragón, L., Mhaidly, R., De Donatis, G. M., Tosolini, M., Dao, P., Martin, A. R., Pons, C., Chiche, J., Jacquin, M., Imbert, V., Proïcs, E., Boyer, L., Doye, A., Luciano, F., Neels, J. G., Coutant, F., Fabien, N., Sormani, L., Rubio-Patiño, C., … Verhoeyen, E. (2019). GAPDH Overexpression in the T Cell Lineage Promotes Angioimmunoblastic T Cell Lymphoma through an NF-κB-Dependent Mechanism. *Cancer Cell*, *36*(3), 268–287.e10. https://doi.org/10.1016/j.ccell.2019.07.008

Mooers, B. H. M., & Brown, M. E. (2021). Templates for writing PyMOL scripts. *Protein Science*, *30*(1), 262–269. https://doi.org/10.1002/pro.3997

Mulero, M. C., Wang, V. Y. F., Huxford, T., & Ghosh, G. (2019). Genome reading by the NF-κB transcription factors. *Nucleic Acids Research*, *47*(19), 9967–9989. https://doi.org/10.1093/NAR/GKZ739

Nagarajan, N., Yapp, E. K. Y., Le, N. Q. K., Kamaraj, B., Al-Subaie, A. M., & Yeh, H. Y. (2019). Application of Computational Biology and Artificial Intelligence Technologies in Cancer Precision Drug Discovery. In *BioMed Research International, 2019*(1), 1–15. https://doi.org/10.1155/2019/8427042

Necula, L., Matei, L., Dragu, D., Neagu, A. I., Mambet, C., Nedeianu, S., Bleotu, C., Diaconu, C. C., & Chivu-Economescu, M. (2019). Recent advances in gastric cancer early diagnosis. *World Journal of Gastroenterology*, *25*(17), 2029–2044. https://doi.org/10.3748/wjg.v25.i17.2029

Nunes, J. A., Araújo, R. S. A. de, Silva, F. N. da, Cytarska, J., Łączkowski, K. Z., Cardoso, S. H., Mendonça-Júnior, F. J. B., & Silva-Júnior, E. F. da. (2023). Coumarin-Based Compounds as Inhibitors of Tyrosinase/Tyrosine Hydroxylase: Synthesis, Kinetic Studies, and In Silico Approaches. *International Journal of Molecular Sciences*, *24*(6), 5216. https://doi.org/10.3390/ijms24065216

Olivieri, C., Walker, C., Veliparambil Subrahmanian, M., Porcelli, F., Taylor, S. S., Bernlohr, D. A., & Veglia, G. (2023). An NMR portrait of functional and dysfunctional allosteric cooperativity in cAMP-dependent protein kinase A. In *FEBS Letters, 597*(8), 1055–1072. https://doi.org/10.1002/1873-3468.14610

Pan, X., & Kortemme, T. (2021). Recent advances in de novo protein design: Principles, methods, and applications. In *Journal of Biological Chemistry, 296*(JANUARY), 100558. https://doi.org/10.1016/j.jbc.2021.100558

Pan, X., Thompson, M. C., Zhang, Y., Liu, L., Fraser, J. S., Kelly, M. J. S., & Kortemme, T. (2020). Expanding the space of protein geometries by computational design of de novo fold families. *Science*, *369*(6507), 1132–1136. https://doi.org/10.1126/science.abc0881

Pandey, P., Panday, S. K., Rimal, P., Ancona, N., & Alexov, E. (2023). Predicting the Effect of Single Mutations on Protein Stability and Binding with Respect to Types of Mutations. *International Journal of Molecular Sciences*, *24*(15), 12073. https://doi.org/10.3390/ijms241512073

Patel, T. H., & Cecchini, M. (2020). Targeted Therapies in Advanced Gastric Cancer. *Current Treatment Options in Oncology*, *21*(9), 70. https://doi.org/10.1007/s11864-020-00774-4

Peng, C., Ouyang, Y., Lu, N., & Li, N. (2020). The NF-κB Signaling Pathway, the Microbiota, and Gastrointestinal Tumorigenesis: Recent Advances. In *Frontiers in Immunology, 11*(2020), 1–13. https://doi.org/10.3389/fimmu.2020.01387

Petryszyn, P., Chapelle, N., & Matysiak-Budnik, T. (2020). Gastric Cancer: Where Are We Heading? In *Digestive Diseases, 38*(4), 280–285. https://doi.org/10.1159/000506509

Pettersen, E. F., Goddard, T. D., Huang, C. C., Meng, E. C., Couch, G. S., Croll, T. I., Morris, J. H., & Ferrin, T. E. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Science*, *30*(1), 70–82. https://doi.org/10.1002/pro.3943

Pflug, K. M., & Sitcheran, R. (2020). Targeting NF-κB-inducing kinase (NIK) in immunity, inflammation, and cancer. *International Journal of Molecular Sciences*, *21*(22), 8470. https://doi.org/10.3390/ijms21228470

Pham, T. N. H., Nguyen, T. H., Tam, N. M., Y. Vu, T., Pham, N. T., Huy, N. T., Mai, B. K., Tung, N. T., Pham, M. Q., V. Vu, V., & Ngo, S. T. (2022). Improving ligand-ranking of AutoDock Vina by changing the empirical parameters. *Journal of Computational Chemistry*, *43*(3). 160–169. https://doi.org/10.1002/jcc.26779

Pinzi, L., & Rastelli, G. (2019). Molecular docking: Shifting paradigms in drug discovery. In *International Journal of Molecular Sciences, 20*(18), 4331. https://doi.org/10.3390/ijms20184331

Planas Iglesias, J., Marques, S. M., Pinto, G. P., Musil, M., Stourac, J., Damborsky, J., & Bednar, D. (2021). Computational design of enzymes for biotechnological applications. In *Biotechnology Advances, 47*(March-April 2021), 107696. https://doi.org/10.1016/j.biotechadv.2021.107696

Qing, R., Hao, S., Smorodina, E., Jin, D., Zalevsky, A., & Zhang, S. (2022). Protein Design: From the Aspect of Water Solubility and Stability. In *Chemical Reviews, 122*(18), 14085–14179. https://doi.org/10.1021/acs.chemrev.1c00757

Ramadass, V., Vaiyapuri, T., & Tergaonkar, V. (2020). Small Molecule NF-κB Pathway Inhibitors in Clinic. *International Journal of Molecular Sciences*, *21*(14), 1–43. https://doi.org/10.3390/IJMS21145164

Rawla, P., & Barsouk, A. (2019). Epidemiology of gastric cancer: Global trends, risk factors and prevention. In *Przeglad Gastroenterologiczny, 14*(1), 26–38. https://doi.org/10.5114/pg.2018.80001

Rocklin, G. J., Chidyausiku, T. M., Goreshnik, I., Ford, A., Houliston, S., Lemak, A., Carter, L., Ravichandran, R., Mulligan, V. K., Chevalier, A., Arrowsmith, C. H., & Baker, D. (2017). Global analysis of protein folding using massively parallel design, synthesis, and testing. *Science*, *357*(6347), 168–175. https://doi.org/10.1126/science.aan0693

Romero Romero, S., Costas, M., Silva Manzano, D. A., Kordes, S., Rojas-Ortega, E., Tapia, C., Guerra, Y., Shanmugaratnam, S., Rodríguez-Romero, A., Baker, D., Höcker, B., & Fernández-Velasco, D. A. (2021). The Stability Landscape of de novo TIM Barrels Explored by a Modular Design Approach. *Journal of Molecular Biology*, *433*(18), 167153. https://doi.org/10.1016/j.jmb.2021.167153

Rong, G., Mendez, A., Bou Assi, E., Zhao, B., & Sawan, M. (2020). Artificial Intelligence in Healthcare: Review and Prediction Case Studies. In *Engineering, 6*(3), 291–301. https://doi.org/10.1016/j.eng.2019.08.015

Salmaso, V., & Moro, S. (2018). Bridging molecular docking to molecular dynamics in exploring ligand-protein recognition process: An overview. In *Frontiers in Pharmacology, 9*(AUG), 923–939. https://doi.org/10.3389/fphar.2018.00923

Schauperl, M., & Denny, R. A. (2022). AI-Based Protein Structure Prediction in Drug Discovery: Impacts and Challenges. In *Journal of Chemical Information and Modeling, 62*(13), 3142–3156. https://doi.org/10.1021/acs.jcim.2c00026

Schöning Stierand, K., Diedrich, K., Fährrolfes, R., Flachsenberg, F., Meyder, A., Nittinger, E., Steinegger, R., & Rarey, M. (2020). ProteinsPlus: Interactive analysis of protein–ligand binding interfaces. *Nucleic Acids Research*, *48*(W1), 48–53. https://doi.org/10.1093/NAR/GKAA235

Schwalbe, N., & Wahl, B. (2020). Artificial intelligence and the future of global health. In *The Lancet, 395*(10236), 1579–1586. https://doi.org/10.1016/S0140-6736(20)30226-9

Secinaro, S., Calandra, D., Secinaro, A., Muthurangu, V., & Biancone, P. (2021). The role of artificial intelligence in healthcare: a structured literature review. *BMC Medical Informatics and Decision Making*, *21*(1), 125 (2021). https://doi.org/10.1186/s12911-021-01488-9

Seelig, J., & Seelig, A. (2023). Protein Stability─Analysis of Heat and Cold Denaturation without and with Unfolding Models. *Journal of Physical Chemistry B*, *127*(15), 3352–3363. https://doi.org/10.1021/acs.jpcb.3c00882

Senior, A. W., Evans, R., Jumper, J., Kirkpatrick, J., Sifre, L., Green, T., Qin, C., Žídek, A., Nelson, A. W. R., Bridgland, A., Penedones, H., Petersen, S., Simonyan, K., Crossan, S., Kohli, P., Jones, D. T., Silver, D., Kavukcuoglu, K., & Hassabis, D. (2020). Improved protein structure prediction using potentials from deep learning. *Nature*, *577*(7792), 706–710. https://doi.org/10.1038/s41586-019-1923-7

Sepulveda, A. R. (2013). Helicobacter, Inflammation, and Gastric Cancer. *Current Pathobiology Reports*, *1*(1), 9–18. https://doi.org/10.1007/s40139-013-0009-8

Sesterhenn, F., Yang, C., Bonet, J., Cramer, J. T., Wen, X., Wang, Y., Chiang, C. I., Abriata, L. A., Kucharska, I., Castoro, G., Vollers, S. S., Galloux, M., Dheilly, E., Rosset, S., Corthésy, P., Georgeon, S., Villard, M., Richard, C. A., Descamps, D., … Correia, B. E. (2020). De novo protein design enables the precise induction of RSV-neutralizing antibodies. *Science*, *368*(6492), eaay5051. https://doi.org/10.1126/science.aay5051

Sexton, R. E., Al Hallak, M. N., Diab, M., & Azmi, A. S. (2020). Gastric cancer: a comprehensive review of current and future treatment strategies. In *Cancer and Metastasis Reviews, 39*(4), 1179–1203. https://doi.org/10.1007/s10555-020-09925-3

Shu, J., Li, J., Wang, S., Lin, J., Wen, L., Ye, H., & Zhou, P. (2023). Systematic analysis and comparison of peptide specificity and selectivity between their cognate receptors and noncognate decoys. *Journal of Molecular Recognition*, *36*(3), e3006. https://doi.org/10.1002/jmr.3006

Siebenmorgen, T., & Zacharias, M. (2020). Computational prediction of protein–protein binding affinities. In *Wiley Interdisciplinary Reviews: Computational Molecular Science, 10*(3), e1448. https://doi.org/10.1002/wcms.1448

Singh, S., Bani Baker, Q., & Singh, D. B. (2022). Molecular docking and molecular dynamics simulation. *Bioinformatics: Methods and Applications*, 291–304. https://doi.org/10.1016/B978-0-323-89775-4.00014-6

Smyth, E. C., Nilsson, M., Grabsch, H. I., van Grieken, N. C., & Lordick, F. (2020). Gastric cancer. *The Lancet*, *396*(10251), 635–648. https://doi.org/10.1016/S0140-6736(20)31288-5

Soleymani, F., Paquet, E., Viktor, H., Michalowski, W., & Spinello, D. (2022). Protein–protein interaction prediction with deep learning: A comprehensive review. In *Computational and Structural Biotechnology Journal, 20*(2022), 5316–5341. https://doi.org/10.1016/j.csbj.2022.08.070

Song, Q., Wu, Z., Jin, C., Yu, Z., Xu, P., & Jiang, Z. (2022). Effects of Residual Composition and Distribution on the Structural Characteristics of the Protein. *International Journal of Molecular Sciences*, *23*(22), 14263. https://doi.org/10.3390/ijms232214263

Soutto, M., Bhat, N., Khalafi, S., Zhu, S., Poveda, J., Garcia-Buitrago, M., Zaika, A., & El-Rifai, W. (2021). NF-kB-dependent activation of STAT3 by H. pylori is suppressed by TFF1. *Cancer Cell International*, *21*(1), 444 (2021). https://doi.org/10.1186/s12935-021-02140-2

Tang, S., Chen, R., Lin, M., Lin, Q., Zhu, Y., Ding, J., Hu, H., Ling, M., & Wu, J. (2022). Accelerating AutoDock Vina with GPUs. *Molecules*, *27*(9), 3041. https://doi.org/10.3390/molecules27093041

Thrift, A. P., & El-Serag, H. B. (2020). Burden of Gastric Cancer. In *Clinical Gastroenterology and Hepatology, 18*(3), 534–542. https://doi.org/10.1016/j.cgh.2019.07.045

Thrift, A. P., & Nguyen, T. H. (2021). Gastric Cancer Epidemiology. *Gastrointestinal Endoscopy Clinics of North America*, *31*(3), 425–439. https://doi.org/10.1016/j.giec.2021.03.001

Topi, S., Santacroce, L., Bottalico, L., Ballini, A., Inchingolo, A. D., Dipalma, G., Charitos, I. A., & Inchingolo, F. (2020). Gastric cancer in history: A perspective interdisciplinary study. *Cancers*, *12*(2), 264. https://doi.org/10.3390/cancers12020264

Voros, O., Szilagyi, O., Balajthy, A., Somodi, S., Panyi, G., & Hajdu, P. (2018). The C-terminal HRET sequence of Kv1.3 regulates gating rather than targeting of Kv1.3 to the plasma membrane. *Scientific Reports*, *8*(1), 5937 (2018). https://doi.org/10.1038/s41598-018-24159-8

Waldum, H., & Fossmark, R. (2021). Gastritis, gastric polyps and gastric cancer. In *International Journal of Molecular Sciences, 22*(12), 6548. https://doi.org/10.3390/ijms22126548

Wang, B., Sun, D., Li, H., & Chen, J. (2024). A bird's eye view of the potential role of NFKBIA in pan-cancer. *Heliyon*, *10*(10), e31204. https://doi.org/10.1016/J.HELIYON.2024.E31204

Ward, E. M., Sherman, R. L., Henley, S. J., Jemal, A., Siegel, D. A., Feuer, E. J., Firth, A. U., Kohler, B. A., Scott, S., Ma, J., Anderson, R. N., Benard, V., & Cronin, K. A. (2019). Annual Report to the Nation on the Status of Cancer, Featuring Cancer in Men and Women Age 20–49 Years. *Journal of the National Cancer Institute*, *111*(12), 1279–1297. https://doi.org/10.1093/JNCI/DJZ106

Watson, J. L., & Juergens, D. (2023, October 3). *RosettaCommons/RFdiffusion: Code for running RFdiffusion*. GitHub. https://github.com/RosettaCommons/RFdiffusion

Watson, J. L., Juergens, D., Bennett, N. R., Trippe, B. L., Yim, J., Eisenach, H. E., Ahern, W., Borst, A. J., Ragotte, R. J., Milles, L. F., Wicky, B. I. M., Hanikel, N., Pellock, S. J., Courbet, A., Sheffler, W., Wang, J., Venkatesh, P., Sappington, I., Torres, S. V., … Baker, D. (2023). De novo design of protein structure and function with RFdiffusion. *Nature*, *620*(7976), 1089–1100. https://doi.org/10.1038/s41586-023-06415-8

Watson, K. (2022). The Ethics of Access: Reframing the Need for Abortion Care as a Health Disparity. *Https://Doi.Org/10.1080/15265161.2022.2075976*, *22*(8), 22–30. https://doi.org/10.1080/15265161.2022.2075976

Wilson, C. J., de Groot, B. L., & Gapsys, V. (2024). Resolving coupled pH titrations using alchemical free energy calculations. *Journal of Computational Chemistry*, *45*(17), 1444–1455. https://doi.org/10.1002/jcc.27318

Winnifrith, A., Outeiral, C., & Hie, B. L. (2024). Generative artificial intelligence for de novo protein design. *Current Opinion in Structural Biology*, *86*(102794), 102794. https://doi.org/10.1016/J.SBI.2024.102794

Wong, M. C. S., Huang, J., Chan, P. S. F., Choi, P., Lao, X. Q., Chan, S. M., Teoh, A., & Liang, P. (2021). Global Incidence and Mortality of Gastric Cancer, 1980-2018. *JAMA Network Open, 4*(7), e2118457. https://doi.org/10.1001/jamanetworkopen.2021.18457

Woolfson, D. N. (2021). A Brief History of De Novo Protein Design: Minimal, Rational, and Computational: De novo protein design. In *Journal of Molecular Biology, 433*(20), 167160. https://doi.org/10.1016/j.jmb.2021.167160

World Health Organization [WHO]. (2022, February 2). *Cancer*. World Health Organization. https://www.who.int/es/news-room/fact-sheets/detail/cancer

Wu, R., Ding, F., Wang, R., Shen, R., Zhang, X., Luo, S., Su, C., Wu, Z., Xie, Q., Berger, B., Ma, J., & Peng, J. (2022). High-resolution de novo structure prediction from primary sequence. *BioRxiv, 2022*(JULY), 1–37. https://doi.org/10.1101/2022.07.21.500999

Xing, C., Chen, P., & Zhang, L. (2023). Computational insight into stability-enhanced systems of anthocyanin with protein/peptide. *Food Chemistry: Molecular Sciences*, *6*(100168), 100168. https://doi.org/10.1016/j.fochms.2023.100168

Yadav, N. S., Kumar, P., & Singh, I. (2021). Structural and functional analysis of protein. In *Bioinformatics: Methods and Applications* (pp. 189–206). https://doi.org/10.1016/B978-0-323-89775-4.00026-2

Yang, J., Anishchenko, I., Park, H., Peng, Z., Ovchinnikov, S., & Baker, D. (2020). Improved protein structure prediction using predicted interresidue orientations. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(3), 1496–1503. https://doi.org/10.1073/pnas.1914677117

Yang, L., Ying, X., Liu, S., Lyu, G., Xu, Z., Zhang, X., Li, H., Li, Q., Wang, N., & Ji, J. (2020). Gastric cancer: Epidemiology, risk factors and prevention strategies. *Chinese Journal of Cancer Research*, *32*(6), 695–704. https://doi.org/10.21147/j.issn.1000-9604.2020.06.03

Yu, H., Lin, L., Zhang, Z., Zhang, H., & Hu, H. (2020). Targeting NF-κB pathway for the therapy of diseases: mechanism and clinical study. In *Signal Transduction and Targeted Therapy, 5*(1), 209. https://doi.org/10.1038/s41392-020-00312-6

Zhang, H., Lu, C., Yao, Q., & Jiao, Q. (2024). In silico study to identify novel NEK7 inhibitors from natural sources by a combination strategy. *Molecular Diversity, 1*(2024), 1–24. https://doi.org/10.1007/s11030-024-10838-4

Zinatizadeh, M. R., Schock, B., Chalbatani, G. M., Zarandi, P. K., Jalali, S. A., & Miri, S. R. (2021). The Nuclear Factor Kappa B (NF-kB) signaling in cancer development and immune diseases. In *Genes and Diseases, 8*(3), 287–297. https://doi.org/10.1016/j.gendis.2020.06.005

ANEXOS

Anexo 1. *Línea de comando y códigos empleados para la ejecución de la Inteligencia Artificial Deepchem.*

- **Configuración:**

```
!pip install --pre deepchem
import deepchem
deepchem.__version__
!pip install biopython
import random
from Bio.Seq import Seq
from Bio.SeqUtils import seq1
from Bio.PDB import PDBParser
from Bio import BiopythonWarning
import warnings
```

- **Generación de nueva secuencia:**

```
# Suprimir las advertencias de Biopython
warnings.simplefilter('ignore', BiopythonWarning)
# 1. Establecer la secuencia proteica de entrada
query_sequence =
"HMKSSVVGSERKRGMQLEDLKQQLQQAEEALVAKQEVIDKLKEEAEQHKIVMETVPVLKA
QADIYKADFQAERQAREKLAEKKELLQEQLEQLQ" #@param {type:"string"}
#@markdown - Incorpora el fragmento mejor "conservado" dentro de `query_sequence`
correspondiente a los alineamietos realizados de los "templates" utilizadso
#@markdown `Nota:` Dicho fragmento de la secuencia no debe presentar interrepciones debido a este
comando los tomará en cuenta y durante la predicción proteica ocurrirá un error
# 2. Aumento del número de residuos (comando)
def add_random_residues(sequence, num_residues, position='end'):
    residues = 'ACDEFGHIKLMNPQRSTVWY'
    for _ in range(num_residues):
        residue = random.choice(residues)
        if position == 'end':
            sequence += residue
        elif position == 'start':
            sequence = residue + sequence
        elif position == 'both':
```

```
            sequence = residue + sequence + residue
    return sequence
# 3. Verificación del formato PDB (comando)
def validate_pdb_format(pdb_file):
    try:
        parser = PDBParser()
        structure = parser.get_structure('temp', pdb_file)
        return True
    except Exception as e:
        print("Error:", e)
        return False
# 4. Rectificación la secuencia con PDBFixer (comando)
def fix_pdb_sequence(sequence):
    fixed_sequence = sequence
    return fixed_sequence
# 5. Presentación la nueva secuencia al usuario
num_residues = 80 #@param {type:"raw"}
#@markdown - Establece el número de residuos que quieres que se agreguen a la nueva secuencia proteica
position = "both" #@param {type:"string"} ["start", "end", "both"]
#@markdown - Define la posición que tendran los residuos, siendo `start` para el inicio de la nueva secuencia, `end` para el final de la misma y `both` en el caso que se necesite añadir en ambas secciones
new_sequence = add_random_residues(query_sequence, num_residues=num_residues, position=position)
print("Nueva secuencia generada:", new_sequence)
# 6. Verificación el formato PDB de la nueva secuencia
is_valid_pdb = validate_pdb_format(new_sequence)
if is_valid_pdb:
    print("La secuencia cumple con el formato PDB.")
else:
    print("La secuencia no cumple con el formato PDB.")
# 7. Rectificación la secuencia si es necesario
fixed_sequence = fix_pdb_sequence(new_sequence)
print("Secuencia corregida:", fixed_sequence)
display(fixed_sequence)
```

Anexo 2. *Secuencias proteicas de las proteínas predichas por las herramientas de Inteligencia Artificial.*

IA	Código de Identificación	Secuencia proteica
AlphaFold	O15226	MEKILQMAEGIDIGEMPSYDLVLSKPSKGQKRHLSTCDGQNPPKKQAGSKFHARPRFEPVHFVASSSKDERQEDPYGP QTKEVNEQTHFASMPRDIYQDYTQDSFSIQDGNSQYCDSSGFILTKDQPVTANMYFDSGNPAPSTTSQQANSQSTPEPS PSQTFPESVVAEKQYFIEKLTATIWKNLSNPEMTSGSDKINYTYMLTRCIQACKTNPEYIYAPLKEIPPADIPKNKKLLT DGYACEVRCQNIYLTTGYAGSKNGSRDRATELAVKLLQKRIEVRVVRRKFKHTFGEDLVVCQIGMSSYEFPPALKPPE DLVVLGKDASGQPIFNASAKHWTNFVITENANDAIGILNNSASFNKMSIEYKYEMMPNRTWRCRVFLQDHCLAEGYG TKKTSKHAAADEALKILQKTQPTYPSVKSSQCHTGSSPRGSGKKKDIKDLVVYENSSNPVCTLNDTAQFNRMTVEYV YERMTGLRWKCKVILESEVIAEAVGVKKTVKYEAAGEAVKTLKKTQPTVINNLKKGAVEDVISRNEIQGRSAEEAYK QQIKEDNIGNQLLRKMGWTGGGLGKSGEGIREPISVKEQHKREGLGLDVERVNKIAKRDIEQIIRNYARSESHTDLTFS RELTNDERKQIHQIAQKYGLKSKSHGVGHDRYLVVGRKRRKEDLLDQLKQEGQVGHYELVMPQAN
	P25963	MFQAAERPQEWAMEGPRDGLKKERLLDDRHDSGLDSMKDEEYEQMVKELQEIRLEPQEVPRGSEPWKQQLTEDGDS FLHLAIIHEEKALTMEVIRQVKGDLAFLNFQNNLQQTPLHLAVITNQPEIAEALLGAGCDPELRDFRGNTPLHLACEQG CLASVGVLTQSCTTPHLHSILKATNYNGHTCLHLASIHGYLGIVELLVSLGADVNAQEPCNGRTALHLAVDLQNPDLV SLLLKCGADVNRVTYQGYSPYQLTWGRPSTRIQQQLGQLTLENLQMLPESEDEESYDTESEFTEFTEDELPYDDCVFG GQRLTL
	Q15653	MAGVACLGKAADADEWCDSGLGSLGPDAAAPGGPGLGAELGPGLSWAPLVFGYVTEDGDTALHLAVIHQHEPFLDF LLGFSAGTEYMDLQNDLGQTALHLAAILGETSTVEKLYAAGAGLCVAERRGHTALHLACRVGAHACARALLQPRPR RPREAPDTYLAQGPDRTPDTNHTPVALYPDSDLEKEEEESEEDWKLQLEAENYEGHTPLHVAVIHKDVEMVRLLRDA GADLDKPEPTCGRSPLHLAVEAQAADVLELLLRAGANPAARMYGGRTPLGSAMLRPNPILARLLRAHGAPEPEGEDE KSGPCSSSSDSDSGDEGDEYDDIVVHSSRSQTRLPPTPASKPLPDDPRPV
	O15111	MERPPGLRPGAGGPWEMRERLGTGGFGNVCLYQHRELDLKIAIKSCRLELSTKNRERWCHEIQIMKKLNHANVVKAC DVPEELNILIHDVPLLAMEYCSGGDLRKLLNKPENCCGLKESQILSLLSDIGSGIRYLHENKIIHRDLKPENIVLQDVGG KIIHKIIDLGYAKDVDQGSLCTSFVGTLQYLAPELFENKPYTATVDYWSFGTMVFECIAGYRPFLHHLQPFTWHEKIKK KDPKCIFACEEMSGEVRFSSHLPQPNSLCSLVVEPMENWLQLMLNWDPQQRGGPVDLTLKQPRCFVLMDHILNLKIV HILNMTSAKIISFLLPPDESLHSLQSRIERETGINTGSQELLSETGISLDPRKPASQCVLDGVRGCDSYMVYLFDKSKTVY EGPFASRSLSDCVNYIVQDSKIQLPIIQLRKVWAEAVHYVSGLKEDYSRLFQGQRAAMLSLLRYNANLTKMKNTLISA SQQLKAKLEFFHKSIQLDLERYSEQMTYGISSEKMLKAWKEMEEKAIHYAEVGVIGYLEDQIMSLHAEIMELQKSPYG

IA	Código de Identificación	Secuencia proteica
		RRQGDLMESLEQRAIDLYKQLKHRPSDHSYSDSTEMVKIIVHTVQSQDRVLKELFGHLSKLLGCKQKIIDLLPKVEVAL SNIKEADNTVMFMQGKRQKEIWHLLKIACTQSSARSLVGSSLEGAVTPQTSAWLPPTSAEHDHSLSCVVTPQDGETSA QMIEENLNCLGHLSTIIHEANEEQGNSMMNLDWSWLTE
	O14920	MSWSPSLTTQTCGAWEMKERLGTGGFGNVIRWHNQETGEQIAIKQCRQELSPRNRERWCLEIQIMRRLTHPNVVAAR DVPEGMQNLAPNDLPLLAMEYCQGGDLRKYLNQFENCCGLREGAILTLLSDIASALRYLHENRIIHRDLKPENIVLQQ GEQRLIHKIIDLGYAKELDQGSLCTSFVGTLQYLAPELLEQQKYTVTVDYWSFGTLAFECITGFRPFLPNWQPVQWHS KVRQKSEVDIVVSEDLNGTVKFSSSLPYPNNLNSVLAERLEKWLQLMLMWHPRQRGTDPTYGPNGCFKALDDILNLK LVHILNMVTGTIHTYPVTEDESLQSLKARIQQDTGIPEEDQELLQEAGLALIPDKPATQCISDGKLNEGHTLDMDLVFLF DNSKITYETQISPRPQPESVSCILQEPKRNLAFFQLRKVWGQVWHSIQTLKEDCNRLQQGQRAAMMNLLRNNSCLSK MKNSMASMSQQLKAKLDFFKTSIQIDLEKYSEQTEFGITSDKLLLAWREMEQAVELCGRENEVKLLVERMMALQTDI VDLQRSPMGRKQGGTLDDLEEQARELYRRLREKPRDQRTEGDSQEMVRLLLQAIQSFEKKVRVIYTQLSKTVVCKQK ALELLPKVEEVVSLMNEDEKTVVRLQEKRQKELWNLLKIACSKVRGPVSGSPDSMNASRLSQPGQLMSQPSTASNSLP EPAKKSEELVAEAHNLCTLLENAIQDTVREQDQSFTALDWSWLQTEEEEHSCLEQAS
	Q8NI38	MEAGPWRVSAPPSGPPQFPAVVPGPSLEVARAHMLALGPQQLLAQDEEGDTLLHLFAARGLRWAAYAAAEVLQVYR RLDIREHKGKTPLLVAAAANQPLIVEDLLNLGAEPNAADHQGRSVLHVAATYGLPGVLLAVLNSGVQVDLEARDFEG LTPLHTAILALNVAMRPSDLCPRVLSTQARDRLDCVHMLLQMGANHTSQEIKSNKTVLHLAVQAANPTLVQLLLELP RGDLRTFVNMKAHGNTALHMAAALPPGPAQEAIVRHLLAAGADPTLRNLENEQPVHLLRPGPGPEGLRQLLKRSRVA PPGLSS
	O00221	MNQRRSESRPGNHRLQAYAEPGKGDSGGAGPLSGSARRGRGGGGAIRVRRPCWSGGAGRGGGPAWAVRLPTVTAG WTWPALRTLSSLRAGPSEPHSPGRRPPRAGRPLCQADPQPGKAARRSLEPDPAQTGPRPARAAGMSEARKGPDEAEES QYDSGIESLRSLRSLPESTSAPASGPSDGSPQPCTHPPGPVKEPQEKEDADGERADSTYGSSSLTYTLSLLGGPEAEDPAP RLPLPHVGALSPQQLEALTYISEDGDTLVHLAVIHEAPAVLLCCLALLPQEVLDIQNNLYQTALHLAVHLDQPGAVRA LVLKGASRALQDRHGDTALHVACQRQHLACARCLLEGRPEPGRGTSHSLDLQLQNWQGLACLHIATLQKNQPLMEL LLRNGADIDVQEGTSGKTALHLAVETQERGLVQFLLQAGAQVDARMLNGCTPLHLAAGRGLMGISSTLCKAGADSL LRNVEDETPQDLTEESLVLLPFDDLKISGKLLLCTD
	Q9BYH8	MIVDKLLDDSRGGEGLRDAAGGCGLMTSPLNLSYFYGASPPAAAPGACDASCSVLGPSAPGSPGSDSSDFSSASSVSSC GAVESRSRGGARAERQPVEPHMGVGRQQRGPFQGVRVKNSVKELLLHIRSHKQKASGQAVDDFKTQGVNIEQFREL KNTVSYSGKRKGPDSLSDGPACKRPALLHSQFLTPPQTPTPGESMEDVHLNEPKQESSADLLQNIINIKNECSPVSLNTV

IA	Código de Identificación	Secuencia proteica
		QVSWLNPVVVPQSSPAEQCQDFHGGQVFSPPQKCQPFQVRGSQQMIDQASLYQYSPQNQHVEQQPHYTHKPTLEYSP FPIPPQSPAYEPNLFDGPESQFCPNQSLVSLLGDQRESENIANPMQTSSSVQQQNDAHLHSFSMMPSSACEAMVGHEM ASDSSNTSLPFSNMGNPMNTTQLGKSLFQWQVEQEESKLANISQDQFLSKDADGDTFLHIAVAQGRRALSYVLARKM NALHMLDIKEHNGQSAFQVAVAANQHLIVQDLVNIGAQVNTTDCWGRTPLHVCAEKGHSQVLQAIQKGAVGSNQFV DLEATNYDGLTPLHCAVIAHNAVVHELQRNQQPHSPEVQELLLKNKSLVDTIKCLIQMGAAVEAKDRKSGRTALHLA AEEANLELIRLFLELPSCLSFVNAKAYNGNTALHVAASLQYRLTQLDAVRLLMRKGADPSTRNLENEQPVHLVPDGP VGEQIRRILKGKSIQQRAPPY
	Q9Y6K9	MNRHLWKSQLCEMVQPSGGPAADQDVLGEESPLGKPAMLHLPSEQGAPETLQRCLEENQELRDAIRQSNQILRERCE ELLHFQASQREEKEFLMCKFQEARKLVERLGLEKLDLKRQKEQALREVEHLKRCQQQMAEDKASVKAQVTSLLGEL QESQSRLEAATKECQALEGRARAASEQARQLESEREALQQQHSVQVDQLRMQGQSVEAALRMERQAASEEKRKLAQ LQVAYHQLFQEYDNHIKSSVVGSERKRGMQLEDLKQQLQQAEEALVAKQEVIDKLKEEAEQHKIVMETVPVLKAQA DIYKADFQAERQAREKLAEKKELLQEQLEQLQREYSKLKASCQESARIEDMRKRHVEVSQAPLPPAPAYLSSPLALPS QRRSPPEEPPDFCCPKCQYQAPDMDTLQIHVMECIE
	Q9UBC1	MSNPSPQVPEEEASTSVCRPKSSMASTSRRQRRERRFRRYLSAGRLVRAQALLQRHPGLDVDAGQPPPLHRACARHD APALCLLLRLGADPAHQDRHGDTALHAAARQGPDAYTDFFLPLLSRCPSAMGIKNKDGETPGQILGWGPPWDSAEEE EEDDASKEREWRQKLQGELEDEWQEVMGRFEGDASHETQEPESFSAWSDRLAREHAQKCQQQQREAEGSRRPPRAE GSSQSWRQQEEEQRLFRERARAKEEELRESRARRAQEALGDREPKPTRAGPREEHPRGAGRGSLWRFGDVPWPCPGG GDPEAMAAALVARGPPLEEQGALRRYLRVQQVRWHPDRFLQRFRSQIETWELGRVMGAVTALSQALNRHAEALK
Deepchem	A7886	DTTTNVFVKSYLEFREVGVSQCESPKFMHLTLVMWNVLRVAHFRAQNVQPYYRVPLYYVPQIWVYPLHRMGRHEEL KWYEFPMPHFYTRSMEVIRQVKGDLAFLNFQNNLQQTPLHLAVITNQPEIAEALLGAGCDPELRDFRGNTPLHLACEQ GCLASVGVLTQSCTTPHLHSILKATNYNGHTCLHLASIHGYLGIVELLVSLGADVNAQEPCNGRTALHLAVDLQNPDL VSLLLKCGADVNRVTYQGYSPYQLTWGRPSTRIQQQLGQLTLENLQMLPE
	EC19F	DTTTNVFVKSYLEFREVGVSQCESPKFMHLTLVMWNVLRVAHFRAQNVQPYYRVPLYYVPQIWVYPLHRMGRHEEL KWYEFPMPHFYTRSMEVIRQVKGDLAFLNFQNNLQQTPLHLAVITNQPEIAEALLGAGCDPELRDFRGNTPLHLACEQ GCLASVGVLTQSCTTPHLHSILKATNYNGHTCLHLASIHGYLGIVELLVSLGADVNAQEPCNGRTALHLAVDLQNPDL VSLLLKCGADVNRVTYQGYSPYQLTWGRPSTRIQQQLGQLTLENLQMLPE
	963D2	AYMYFQQYNEEWMLIRFPWSKQNPIHMRIEFCGMQKNHCSYQHRELDLKIAIKSCRLELSTKNRERWCHEIQIMKKL NHANVVKACDVPEELNILIHDVPLLAMEYCSGGDLRKLLNKPENCCGLKESQILSLLSDIGSGIRYLHENKIIHRDLKPE

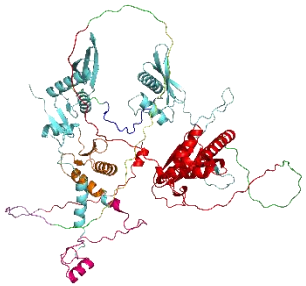
IA	Código de Identificación	Secuencia proteica
		NIVLQDVGGKIIHKIIDLGYAKDVDQGSLCTSFVGTLQYLAPELFENKPYTATVDYWSFGTMVFECIAGYRPFLHHLQP FTWHEKIKKKDPKCIFACEEMSGEVRFSSHLPQPNSLCSLIVEPMENWLQLMLNWDPQQRGGPVDLTLKQPRCFVLM DHILNLKIVHILNMTSAKIISFLLPPDESLHSLQSRIERETGINTGSQELLSETGISLDPRKPASQCVLDGVSCHNKQMGCF EIRMHIPNQKSWPFRILMWEENYQQFYMYA
	92B1D	AYMYFQQYNEEWMLIRFPWSKQNPIHMRIEFCGMQKNHCSYQHRELDLKIAIKSCRLELSTKNRERWCHEIQIMKKL NHANVVKACDVPEELNILIHDVPLLAMEYCSGGDLRKLLNKPENCCGLKESQILSLLSDIGSGIRYLHENKIIHRDLKPE NIVLQDVGGKIIHKIIDLGYAKDVDQGSLCTSFVGTLQYLAPELFENKPYTATVDYWSFGTMVFECIAGYRPFLHHLQP FTWHEKIKKKDPKCIFACEEMSGEVRFSSHLPQPNSLCSLIVEPMENWLQLMLNWDPQQRGGPVDLTLKQPRCFVLM DHILNLKIVHILNMTSAKIISFLLPPDESLHSLQSRIERETGINTGSQELLSETGISLDPRKPASQCVLDGVSCHNKQMGCF EIRMHIPNQKSWPFRILMWEENYQQFYMYA
	A26C8	IMWQDAVFHFYFVCCFGELKEMWVGPPQIAQNEISHVKKKEIFCNVETYHSDNQTGNQMKWHSKVRQKSEVDIVVS EDLNGTVKFSSSLPYPNNLNSVLAERLEKWLQLMLMWHPRQRGTDPTYGPNGCFKALDDILNLKLVHILNMVTGTIH TYPVTEDESLQSLKARIQQDTGIPEEDQELLQEAGLALIPDKPATQCISDGKLNEGHTLDMDLVFLFDNSKITYETQISP RPQPESVSCILQEPKRNLAFFQLRKVWGQVWHSIQTLKEDCNRLQQGQRAAMMNLLRNNSCLSKMKNSMASMSQQL KAKLDFFKTSIQIDLEKYSEQTEFGITSDKLLLAWREMEQAVELCGRENEVKLLVERMMALQTDIVDLQRSPMGRKQ GGTLDDLEEQARELYRRLREKPRDQRTEGDSQEMVRLLLQAIQSFEKKVRVIYTQLSKTVVCKQKALELLPKVEEVVS LMNEDEKTVVRLQEKRQKELWNLLKIACSKVRGPVSGSPDSMNASRLSQPGQLMSQPSTASNSLPEPAKKSEELVAE AHNLCTLLENAIQDTVREQDQSFTKMQNGTQNDSHYTEVNCFIEKKKVHSIENQAIQPPGVWMEKLEGFCCVFYFHF VADQWMI
	208E3	IMWQDAVFHFYFVCCFGELKEMWVGPPQIAQNEISHVKKKEIFCNVETYHSDNQTGNQMKWHSKVRQKSEVDIVVS EDLNGTVKFSSSLPYPNNLNSVLAERLEKWLQLMLMWHPRQRGTDPTYGPNGCFKALDDILNLKLVHILNMVTGTIH TYPVTEDESLQSLKARIQQDTGIPEEDQELLQEAGLALIPDKPATQCISDGKLNEGHTLDMDLVFLFDNSKITYETQISP RPQPESVSCILQEPKRNLAFFQLRKVWGQVWHSIQTLKEDCNRLQQGQRAAMMNLLRNNSCLSKMKNSMASMSQQL KAKLDFFKTSIQIDLEKYSEQTEFGITSDKLLLAWREMEQAVELCGRENEVKLLVERMMALQTDIVDLQRSPMGRKQ GGTLDDLEEQARELYRRLREKPRDQRTEGDSQEMVRLLLQAIQSFEKKVRVIYTQLSKTVVCKQKALELLPKVEEVVS LMNEDEKTVVRLQEKRQKELWNLLKIACSKVRGPVSGSPDSMNASRLSQPGQLMSQPSTASNSLPEPAKKSEELVAE AHNLCTLLENAIQDTVREQDQSFTKMQNGTQNDSHYTEVNCFIEKKKVHSIENQAIQPPGVWMEKLEGFCCVFYFHF VADQWMI

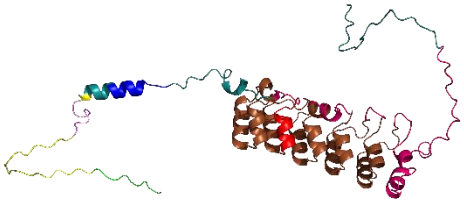
IA	Código de Identificación	Secuencia proteica
	10FAD	MLCGEGKPRKAKNMRKHETENWSNEILINIASIWVKGEWTNHQQRRFNSDSLEWYGPSQSIYIYAYGCMRNLPPWA WRENWHSKVRQKSEVDIVVSEDLNGTVKFSSSLPYPNNLNSVLAERLEKWLQLMLMWHPRQRGTDPTYGPNGCFKA LDDILNLKLVHILNMVTGTIHTYPVTEDESLQSLKARIQQDTGIPEEDQELLQEAGLALIPDKPATQCISDGKLNEGHTL DMDLVFLFDNSKITYETQISPRPQPESVSCILQEPKRNLAFFQLRKVWGQVWHSIQTLKEDCNRLQQGQRAAMMNLLR NNSCLSKMKNSMASMSQQLKAKLDFFKTSIQIDLEKYSEQTEFGINERWAWPPLNRMCGYAYIYISQSPGYWELSDSN FRRQQHNTWEGKVWISAINILIENSWNETEHKRMNKAKRPKGEGCLM
	9061D	MLCGEGKPRKAKNMRKHETENWSNEILINIASIWVKGEWTNHQQRRFNSDSLEWYGPSQSIYIYAYGCMRNLPPWA WRENWHSKVRQKSEVDIVVSEDLNGTVKFSSSLPYPNNLNSVLAERLEKWLQLMLMWHPRQRGTDPTYGPNGCFKA LDDILNLKLVHILNMVTGTIHTYPVTEDESLQSLKARIQQDTGIPEEDQELLQEAGLALIPDKPATQCISDGKLNEGHTL DMDLVFLFDNSKITYETQISPRPQPESVSCILQEPKRNLAFFQLRKVWGQVWHSIQTLKEDCNRLQQGQRAAMMNLLR NNSCLSKMKNSMASMSQQLKAKLDFFKTSIQIDLEKYSEQTEFGINERWAWPPLNRMCGYAYIYISQSPGYWELSDSN FRRQQHNTWEGKVWISAINILIENSWNETEHKRMNKAKRPKGEGCLM
	533E3	NPLMDKYYCDKFKMNKWIGIRLKKMTVDKGYPNAGRQCYQAHHWSLWGCAQVNHHQCYWDRNSFQLDSVFYVSI HSMIYSHMKSSVVGSERKRGMQLEDLKQQLQQAEEALVAKQEVIDKLKEEAEQHKIVMETVPVLKAQADIYKADFQ AERQAREKLAEKKELLQEQLEQLQSYIMSHISVYFVSDLQFSNRDWYCQHHNVQACGWLSWHHAQYCQRGANPYG KDVTMKKLRIGIWKNMKFKDCYYKDMLPN
	07795	GIYVFAESCDYFCHQRDSGQLYRKWQQAPRRKQYWPWEVFMHSRNLPFRDDKIWHWIGQQAVYSSDSSPAYSDNG RFSLQADIYKADFQAERQAREKLAEKKELLQEQLEQLQREYSKLKASCQESARIEDMRKRHVEVSQAPLPPAPAYLSS PLALPSQRRSPPEEPPDFCCPKCQYQAPDMDTLQIHVMECIE
RFdiffusion	RF21A	MLAEILALIDSLEPEALARALELFRAALALPDEELQALFLAVMLRLERQIKASSPEQAALLRQLRIEMTRLLIETRPEVA RALCPLLAAAGLSQETLATMLHLAVITNQPEIAEALLGAGCDPELRDFRGNTPLHLACEQGCLASVGVLTQSCTTPHL HSILKATNYNGHTCLHLASIHGYLGIVELLVSLGADVNAQEKNGNEEVYKKVQEAVASL
	RF415AD	LLTPEELEELPWEMRERLGTGGFGNVCLYQHRELDLKIAIKSLSTKNRERWCHEIQIMKKLRKLLNKPENCCGLKESQI LSLLSDIGSGIRYLHENKIIHRDLKPENIVLQDVGGKIIHKIILQYLAPELFENKPYTATVDYWSFGTMVFECIAGYRPFL HHLQPFTWHEKIKKKDPKCIFACEEMSGEVRFSSHLPQPNSLCSLIVEPMENWLQLMLNWDPQQRGGPVDLTLKQPR CFVLMDHILNLKIVHILNMTSAKIISFLLPPDESLHSLQSRIERETGINTGSQELLSETGISLDPRKPMVYLITKSDTSFRTR LEIENLKKEHPEAAAMIGPISDEVSVEDVRRLYDVVVELL

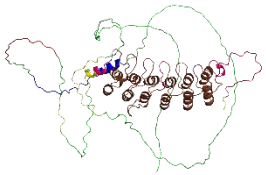
IA	Código de Identificación	Secuencia proteica
	RF715A1	GPVIIPWEMRERLGTGGFGNVCLYQHRELDLKIAIKSLSTKNRERWCHEIQIMKKLRKLLNKPENCCGLKESQILSLLS DIGSGIRYLHENKIIHRDLKPENIVLQDVGGKIIHKIILQYLAPELFENKPYTATVDYWSFGTMVFECIAGYRPFLHHLQP FTWHEKIKKKDPKCIFACEEMSGEVRFSSHLPQPNSLCSLIVEPMENWLQLMLNWDPQQRGGPVDLTLKQPRCFVLM DHILNLKIVHILNMTSAKIISFLLPPDESLHSLQSRIERETGINTGSQELLSETGISLDPRKPMVYLSDGKPREIKPEDKEKF EENYVEEHEKTKKAKKEVEKKEKVEFVKIGDKYYILRVL
	RF315AD	PWEMRERLGTGGFGNVCLYQHRELDLKIAIKSLSTKNRERWCHEIQIMKKLRKLLNKPENCCGLKESQILSLLSDIGSG IRYLHENKIIHRDLKPENIVLQDVGGKIIHKIIPRPLDLQYLAPELFENKPYTATVDYWSFGTMVFECIAGYRPFLHHLQP FTWHEKIKKKDPKCIFACEEMSGEVRFSSHLPQPNSLCSLIVEPMENWLQLMLNWDPQQRGGPVDLTLKQPRCFVLM DHILNLKIVHILNMTSAKIISFLLPPDESLHSLQSRIERETGINTGSQELLSETGISLDPRKPMVYL
	RF215A2	PWEMRERLGTGGFGNVCLYQHRELDLKIAIKSLSTKNRERWCHEIQIMKKLRKLLNKPENCCGLKESQILSLLSDIGSG IRYLHENKIIHRDLKPENIVLQDVGGKIIHKIIPYPLSSKESRLQYLAPELFENKPYTATVDYWSFGTMVFECIAGYRPFL HHLQPFTWHEKIKKKDPKCIFACEEMSGEVRFSSHLPQPNSLCSLIVEPMENWLQLMLNWDPQQRGGPVDLTLKQPR CFVLMDHILNLKIVHILNMTSAKIISFLLPPDESLHSLQSRIERETGINTGSQELLSETGISLDPRKPMVYL
	RF315BD	MVRHEVLVLERGEIAPRSKEKVIKLVTEVLEMRSLPVEKQSEWILEYALKKITLRRVTVDVLYVVSPELLARLEAFAAE HAPLFLKAAAELRIYGELGNLIEACSRKPVAMTTLLDDHQREIADKAIKKGGKGIVVSEDLNGTVKFSSSLPYPNNLNS VLAERLEKWLQLMLMWHPRQRGTDPTYGPNGCFKALDDILNLKLVHILNMVTGTIHTYPVTEDESLQSLKARIQQDT GIPEEDQELLQEAGLALIPDKPATQCISLNEGHTLDMDLVFLFDNSKITYETQRPQPESVSCILQEPKRNLAFFQLRKVW GQVWHSIQTLKEDCNRLQQGQRAAMMNLLRNNSCLSKMKNSMASMSQQLKAKLDFFKTSIQIDLEKYSEQTDKLLL AWREMLKKLPPLSPEAEEFYKKKLEEIEKQRAELEKYTAEVTAKYQAEIGKLITTLCTTLIAQAKAFAEEYAKHAADN PDVLLHFLFVIGEALAGAAKFDPALVAATLAEIEEIFAAHGAHAELTALAMEYCRLYVLAEAA
	RF51B	VIIERTPGLIIAYYPGALIVYCSGPSCTPEKIKEIIDLLKKRKYKNQLIICISTDLSKEEIIEIAAAIVEAAKESGSPEIWIAISS GRLLEKGHHTVFLHFVLPLSPEDVRIEVEKIIAAKEKHFEKTKGAIVVSEDLNGTVKFSSSLPYPNNLNSVLAERLEKW LQLMLMWHPRQRGTDPTYGPNGCFKALDDILNLKLVHILNMVTGTIHTYPVTEDESLQSLKARIQQDTGIPEEDQELL QEAGLALIPDKPATQCISLNEGHTLDMDLVFLFDNSKITYETQRPQPESVSCILQEPKRNLAFFQLRKVWGQVWHSIQT LKEDCNRLQQGQRAAMMNLLRNNSCLSKMKNSMASMSQQLKAKLDFFKTSIQIDLEKYSEQTDKLLLAWREMLEKL KEVEKKIEELEKQIEELDKKEEEAEKRLESLNKAVLKYILEKLKQATKLLKAVLELLKQAPLEALPALFPIVEQMLKQV KELIKLAKEQLKQVKTLTEEEKKEIEKEIKEVEKEMKELEKAIEEARKAVAEA

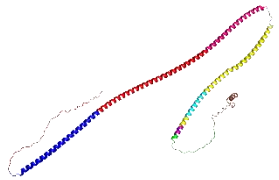
IA	Código de Identificación	Secuencia proteica
	RF2M31	ELEDKNEELLSEIAHLKNEVARLKKLLQRCLGGLLGLEMSRLRDAGDEAAVAAIEAALAKEREEVKKELKKLKEEGA PEEDIENKEALLEALEESAETLAANQELRDAIRQSNQILRERAEELLHFQASQREEKEFLM
	RF3M31	ELEDKNEELLSEIAHLKNEVARLKKLLQRCGEPGRRLVEELRDAPLEERLAVLAEREAELQAILAALEALRAKETLEEK KAIIDELRAALLAANQELRDAIRQSNQILRERAEELLHFQASQREEKEFLM
	RF2M32	HHMMTVARFDGMTLEDLELLNKALKTGELSPSMSEKVKAVLEAVKKGVDVILITSVGADPEKAWKAFLELAKLFKE KGIALLIVHANTKEDADAWADKLEPIVKKYKSNIVAIVEKHPDEPTKVKVVCDTEESAQELLELLKDLPGLSMKVEDG WYIIYTSEGPADFCCPKCQYQAPDMDTLQIHVMECIEEEKHQKLMKEKE

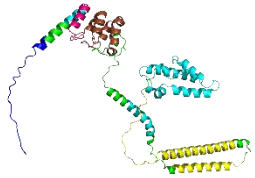
Anexo 3. *Estructuras y composición de las proteínas complementarias.*

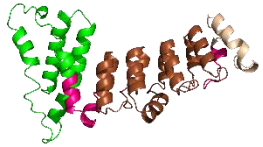
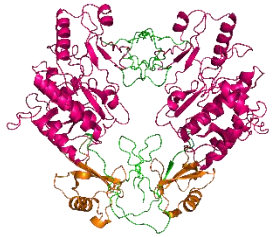
IA	Código de Identificación	Composición de la proteína	Figuras
AlphaFold	O15226	**Regiones** • Dominio de Represión Activa: Rojo (1-296) • Regiones desordenadas: Verde (27–87, 132–163, 414–437) **Sesgos compositivos** • Residuos polares: Azul (29–43, 414–430) • Residuos básicos y ácidos: Amarillo (51–78) **Dominios** • Parche G: Rosado (551-596) • R3H: Naranja (600-664) **Motivos** • Señal de localización nuclear: Violeta (25-45)	 Composición y estructura de la proteína O15226

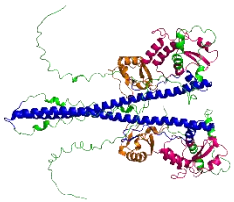
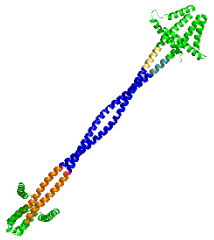
IA	Código de Identificación	Composición de la proteína	Figuras
	P25963	**Regiones** • Regiones desordenadas: Verde (1–39) **Sesgos compositivos** • Residuos básicos y ácidos: Amarillo (11–39) **Dominios** • Dominio que contiene repeticiones Ankyrin: Rosado (73-293) **Motivos** • Motivo de destrucción: Violeta (30-36) • Señal de exportación nuclear: Azul (45-54) • Señal de importación nuclear: Rojo (110-120) **Repeticiones** • ANK1: Café (73-103) • ANK2: Café (110-139) • ANK3: Café (143-172) • ANK4: Café (182-211) • ANK5: Café (216-245)	 Composición y estructura de la proteína P25963

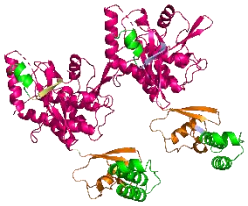
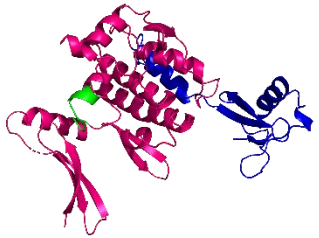
IA	Código de Identificación	Composición de la proteína	Figuras
	O15111	**Regiones** • Cremallera de leucina: Rojo (455-476) • Encuadernación NEMO: Azul (738–743) **Dominios** • Dominio catalítico de la serina/treonina quinasa: Rosado (21-309) • Dominio similar a la ubiquitina (Ubl): Amarillo (311-387) • Dominio de dimerización de andamio: Naranja (386-659)	 Composición y estructura de la proteína O15111
	O14920	**Regiones** • Cremallera de leucina: Rojo (458-479) • Región desordenada: Verde (667–706) • Encuadernación NEMO: Naranja (737-742) **Sesgos compositivos** • Residuos polares: Azul (670-700) **Dominios** • Proteína quinasa: Rosado (15-300)	 Composición y estructura de la proteína O14920
	O00221	**Regiones** • Regiones desordenadas: Verde (1–66, 84–215, 222–241) **Sesgos compositivos** • Residuos polares: Azul (157–183) • Residuos básicos y ácidos: Rojo (194–208) **Dominios**	

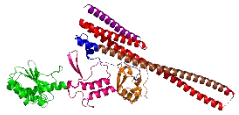
IA	Código de Identificación	Composición de la proteína	Figuras
		• Dominio que contiene repeticiones Ankyrin: Rosado (249-496) **Repeticiones** • ANK 1: Café (258-291) • ANK 2: Café (293-322) • ANK 3: Café (326-355) • ANK 4: Café (369-398) • ANK 5: Café (403-432) • ANK 6: Café (436-465)	Composición y estructura de la proteína O00221
	Q9Y6K9	**Regiones** • Requerido para la interacción y ubiquitinación por MARCHF2: Rojo (1-197) • Interacción con CHUK/IKBKB: Azul (44-111) • Interacción con TANK: Rosado (150-257) • Unión de ubiquitina (UBAN): Amarillo (242-350) • Auto asociación: Morado (246-365) • Requerido para la interacción con TNFAIP3: Gris (251-419) • Cremallera de leucina: Cyan (322-343) • Regiones desordenadas: Verde (358-395) • Interacción con CYLD: Olivo oscuro (382-419) **ZNF** • Dedo de Zinc: Café (389-419)	 Composición y estructura de la proteína Q9Y6K9

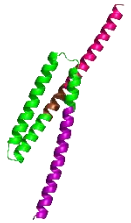
IA	Código de Identificación	Composición de la proteína	Figuras
	Q9UBC1	**Regiones** • Regiones desordenadas: Verde (1-34, 129-167, 186-294) **Sesgos compositivos** • Residuos polares: Azul (1-25) • Residuos básicos y ácidos: Amarillo (201-226, 235-290) **Dominios** • Dominio que contiene repeticiones Ankyrin: Rosado (43-138) **Repeticiones** • ANK 1: Café (64-93) • ANK 2: Café (97-130)	 Composición y estructura de la proteína Q9UBC1
Deepchem	A7886	**Regiones** • Regiones artificiales: Verde (1-90) **Dominios** • Dominio que contiene repeticiones Ankyrin: Rosado (91-256) **Repeticiones** • ANK 1: Café (110-141) • ANK 2: Café (143-180) • ANK 3: Café (182-213) • ANK 4: Café (216-246)	 Composición y estructura de la proteína A7886

IA	Código de Identificación	Composición de la proteína	Figuras
	EC19F	**Regiones** • Regiones artificiales: Verde (1-90) **Dominios** • Dominio que contiene repeticiones Ankyrin: Rosado (91-256) **Repeticiones** • ANK 1: Café (110-141) • ANK 2: Café (143-180) • ANK 3: Café (182-213) • ANK 4: Café (216-246)	 Composición y estructura de la proteína EC19F
	963D2	**Regiones** • Regiones artificiales: Verde (1-40, 385-425) **Dominios** • Dominio catalítico de la serina/treonina quinasa: Rosado (37-318) • Dominio similar a la ubiquitina (Ubl): Naranja (320-382)	 Composición y estructura de la proteína 963D2
	92B1D	**Regiones** • Regiones artificiales: Verde (1-40, 385-425) **Dominios** • Dominio catalítico de la serina/treonina quinasa: Rosado (37-318) • Dominio similar a la ubiquitina (Ubl): Naranja (320-382)	 Composición y estructura de la proteína 92B1D

IA	Código de Identificación	Composición de la proteína	Figuras
	10FAD	**Regiones** • Regiones artificiales: Verde (1-80, 355-435) **Dominios** • Dominio catalítico de la serina/treonina quinasa: Rosado (65-159) • Dominio similar a la ubiquitina (Ubl): Naranja (160-240) • Dominio de dimerización del andamio IQBAL: Azul (239-355)	 Composición y estructura de la proteína 10FAD
	533E3	**Regiones** • Regiones artificiales: Verde (1-80, 174-254) • Cremallera de leucina: Azul (96-160) **Dominios** • Dominio de unión a poliubiquitina de NEMO y proteínas relacionadas (UBAN): Rosado (95-160) • Dominio en espiral N-terminal no caracterizado del peptidoglicano hidrolasa CwlO: Naranja (62-158)	 Composición y estructura de la proteína 533E3
	07795	**Regiones** • Regiones artificiales: Verde (1-80) • Cremallera de leucina: Azul (80-120) **Dominios** • Dominio de unión a poliubiquitina de NEMO y proteínas relacionadas (UBAN): Rosado (80-120) **ZNF** • Dedos de zinc tipo C2H2: Café (169-194)	 Composición y estructura de la proteína 07795

IA	Código de Identificación	Composición de la proteína	Figuras
RFdiffusion	RF21A	**Regiones** • Regiones artificiales: Verde (1-100, 199-217) **Dominios** • Dominio que contiene repeticiones Ankyrin: Rosado (101-198) **Repeticiones** • ANK 1: Café (110-127) • ANK 2: Café (129-166) • ANK 3: Café (168-193)	 Composición y estructura de la proteína RF21A
	RF415AD	**Regiones** • Regiones artificiales: Verde (1-10, 310-360) **Dominios** • Dominio catalítico de la serina/treonina quinasa: Rosado (18-247) • Dominio similar a la ubiquitina (Ubl): Naranja (249-307)	 Composición y estructura de la proteína RF415AD
	RF215A2	**Regiones** • Regiones artificiales: Verde (113-122) **Dominios** • Serina/treonina proteína quinasa: Rosado (1-225)	 Composición y estructura de la proteína RF215A2

IA	Código de Identificación	Composición de la proteína	Figuras
	RF315BD	**Regiones** • Regiones artificiales: Verde (1-132, 397-532) **Dominios** • Dominio catalítico de la serina/treonina quinasa: Rosado (125-200) • Dominio similar a la ubiquitina (Ubl): Naranja (201-278) • Dominio de dimerización del andamio IQBAL: Azul (277-396)	 Composición y estructura de la proteína RF315BD
	RF51B	**Regiones** • Regiones artificiales: Verde (1-131, 395-530) **Dominios** • Dominio catalítico de la serina/treonina quinasa: Rosado (126-198) • Dominio similar a la ubiquitina (Ubl): Naranja (199-276) • Dominio de dimerización del andamio IQBAL: Azul (275-415) **Proteínas** • Proteína de segregación cromosómica SMC: Café (315-492) • Proteína conservada YhaN no caracterizada: Morado (382-530) **Motivo** • Reparación de roturas de doble cadena de ADN ATPasa Rad50 (PRK03918): Rojo (377-494)	 Composición y estructura de la proteína RF51B

IA	Código de Identificación	Composición de la proteína	Figuras
	RF3M31	**Regiones** • Regiones artificiales: Verde (32-90) **Dominios** • Modulador esencial NF-kappa-B NEMO: Rosado (90-130) **Proteínas** • Enzima de restricción tipo I (hsdR): Café (3-100) • Posible nucleasa del pliegue RNasa H (COG2433): Morado (1-90)	 Composición y estructura de la proteína RF3M31
	RF3M32	**Regiones** • Regiones artificiales: Verde (1-165) **Dominios** • Dominio de unión a FMN similar a la dihidrouridina sintasa (tipo DUS): Rosado (35-91) **Proteínas** • Carbamoilfosfato sintasa, subunidad grande: Café (12-139) **ZFN** • Dedos de zinc tipo C2H2: Azul (167-190)	 Composición y estructura de la proteína RF2M32

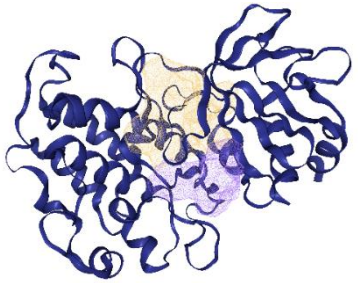
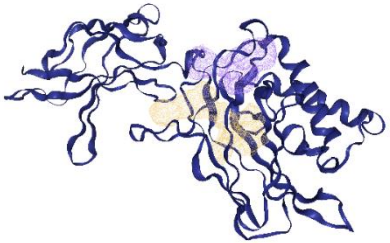
Anexo 4. *Datos y propiedades energéticas de las proteínas complementarias.*

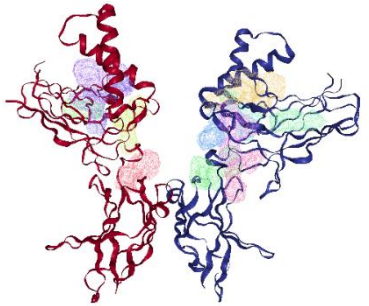
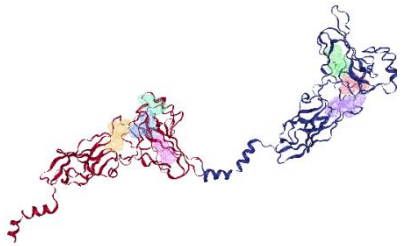
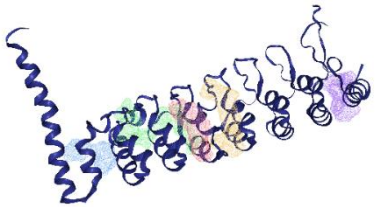
IA	Códigos de Identificación	Datos básicos de la proteína						Energía Computada			
		Cadenas	Longitud	pH	Residuos Terminal	Protonación	Ion	Enlace	CMAP	VDWaals	Angular
AlphaFold	O15226	A	690	7,0	NTER CTER	-	NaCl	1307,79	-147,50	-1621,88	2280,32
	P25963	A	317	7,0	NTER CTER	-	NaCl	503,45	-77,77	-817,69	940,27
	O15111	A	745	7,0	NTER CTER	-	NaCl	1194,87	-149,25	-2627,20	2103,24
	O14920	A	756	7,0	NTER CTER	-	NaCl	1322,26	-154,52	-2773,19	2340,87
	O00221	A	500	7,0	NTER CTER	-	NaCl	984,23	-53,19	-998,98	1986,16
	Q9Y6K9	A	419	7,0	NTER CTER	-	NaCl	706,91	-80,30	-927,44	1289,95
	Q9UBC1	A	381	7,0	NTER CTER	-	NaCl	732,13	21,92	-862,41	1236,55
Deepchem	A7886	A B	282	7,0	NTER CTER	A (39ARG – RN1; 1ASP - ASPP) B (89ARG – RN1)	NaCl	1598,84	15,78	$0,21x10^{17}$	2194,07
	EC19F	A	282	7,0	NTER CTER	-	NaCl	596,67	-24,61	$1,40x10^{8}$	939,08
	963D2	A B	425	7,0	NTER CTER	A (111ASP–ASPP; 114LYS–LSN; 47ASP–ASPP) B (114LYS–LSN; 415GLU– GLUP)	NaCl	2785,85	52,55	$0,74x10^{12}$	3691,54

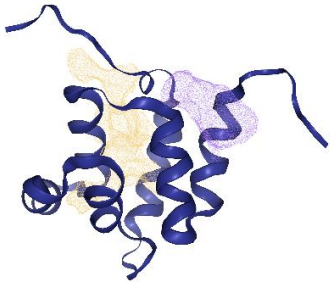
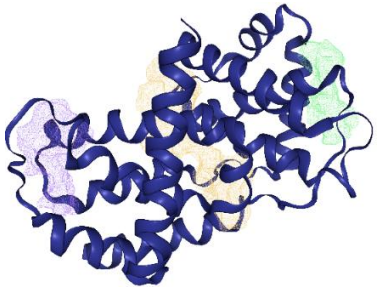
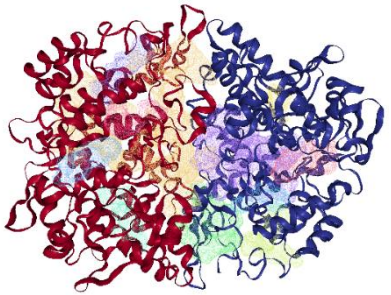
IA	Códigos de Identificación	Datos básicos de la proteína						Energía Computada			
		Cadenas	Longitud	pH	Residuos Terminal	Protonación	Ion	Enlace	CMAP	VDWaals	Angular
	92B1D	A	425	7,0	NTER CTER	-	NaCl	861,35	-134,29	$1,14x10^{6}$	1708,99
	10FAD	A B	435	7,0	NTER CTER	-	NaCl	2678,92	177,09	$0,88x10^{12}$	5111,01
	533E3	A B	254	7,0	NTER CTER	-	NaCl	1229,68	42,48	$0,10x10^{17}$	1680,55
	07795	A	195	7,0	GLYP CTER	-	NaCl	518,72	-43,51	$1,55x10^{6}$	989,84
RFdiffusion	RF21A	A	217	7,0	NTER CTER	-	NaCl	439,73	-24,16	3517,64	481,06
	RF415AD	A B	434	7,0	NTER CTER	-	NaCl	$6,68x10^{5}$	-56,24	$0,13x10^{16}$	2797,02
	RF315AD	A B	358	7,0	PROP CTER	A (124GLU–GLUP; 209LYS–LSN; 264ASP–ASPP; 268LYS–LSN; 25ASP–ASPP) B (204GLU–GLUP; 205LYS–LSN; 209LYS–LSN; 218GLU–GLUP)	NaCl	$5,07x10^{5}$	-68,28	$0,34x10^{13}$	2544,19
	RF215A2	A	363	7,0	PROP CTER	A (214LYS–LSN; 25ASP–ASPP)	NaCl	$2,18x10^{5}$	-41,88	$1,69x10^{5}$	960,61
	RF51B	A	542	7,0	NTER CTER	-	NaCl	38152,88	-33,95	$0,36x10^{13}$	1867,60

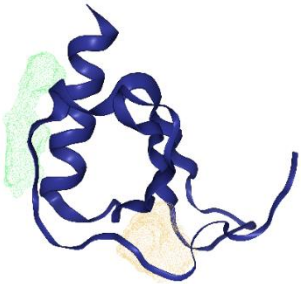
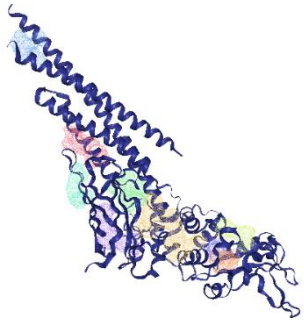
IA	Códigos de Identificación	Datos básicos de la proteína						Energía Computada			
		Cadenas	Longitud	pH	Residuos Terminal	Protonación	Ion	Enlace	CMAP	VDWaals	Angular
	RF3M31	A	130	7,0	NTER CTER	-	NaCl	336,06	-8,86	60017,24	330,47
	RF3M32	A	203	7,0	NTER CTER	-	NaCl	426,31	-35,35	$1,24x10^6$	628,71

Anexo 5. *Sitios de unión de los genes y factores de transcripción relacionados con NF- κB en el cáncer para el proceso de Acoplamiento Molecular.*

Genes y Factores relacionados con NF- κB	Sitios de Unión
Quinasa inductora de NF- κB (NIK)	 Sitios de Unión de la Quinasa inductora NF- κB (NIK)
Subunidad p50 (NFκB1)	 Sitios de Unión de la Subunidad p50 (NFκB1)

Genes y Factores relacionados con NF- κB	Sitios de Unión
Subunidad p52 (NFκB2)	 Sitios de Unión de la Subunidad p52 (NFκB2)
Subunidad p65 (RelA)	 Sitios de Unión de la Subunidad p65 (RelA)
Subunidad p100 (NFκB2)	 Sitios de Unión de la Subunidad p100 (NFκB2)

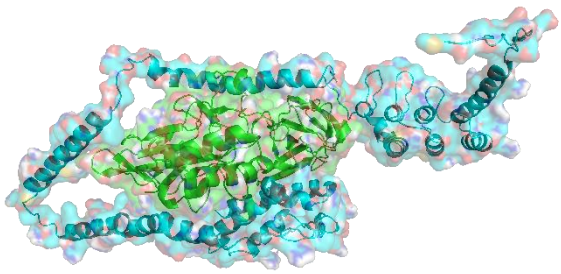
Genes y Factores relacionados con NF- κB	Sitios de Unión
Subunidad p105 (NFκB1)	 Sitios de Unión de la Subunidad p105 (NFκB1)
CiclinaD1	 Sitios de Unión de la CiclinaD1
Ciclooxigenasa-2	 Sitios de Unión de la Ciclooxigenasa-2

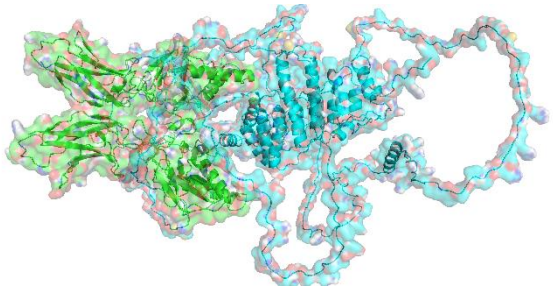
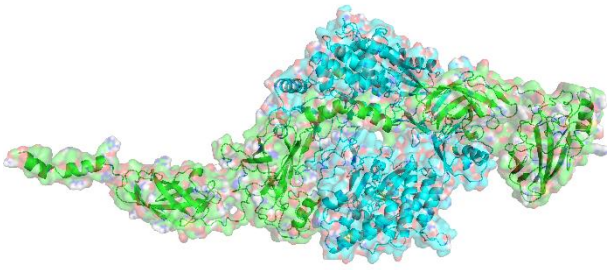
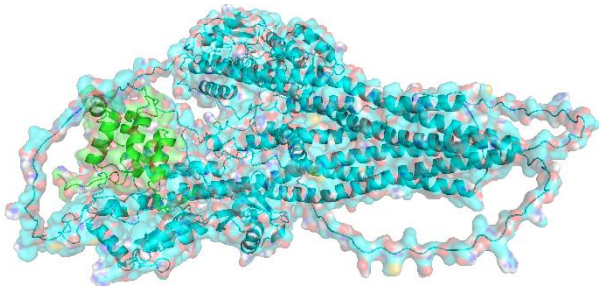
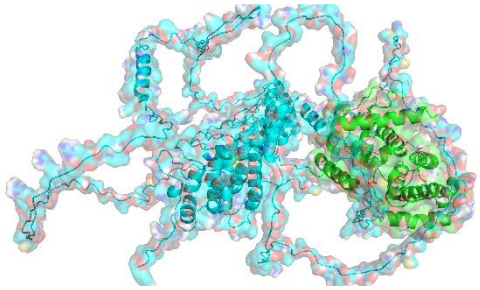
Genes y Factores relacionados con NF- κB	Sitios de Unión
Interleucina-B	 Sitios de Unión de la Interleucina-B
Interleucina-8	 Sitios de Unión de la Interleucina-8
STAT3	 Sitios de Unión del STAT3

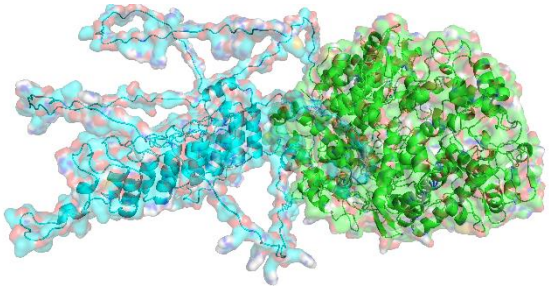
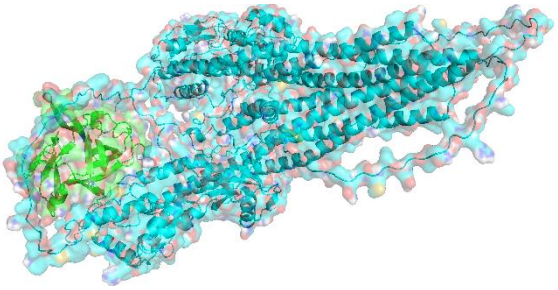
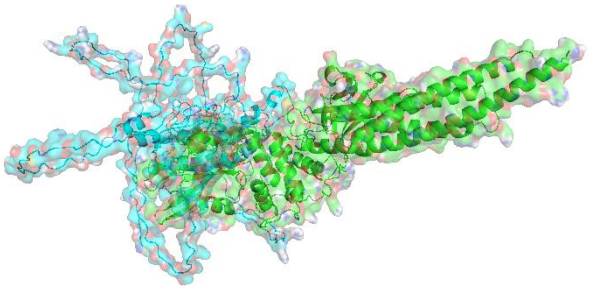
Anexo 6. *Datos complementarios acerca de las variables termodinámicas obtenidas tras el proceso de Acoplamiento Molecular.*

Acoplamiento Molecular		**Variables termodinámicas**				
Receptor	**Mejor Ligando**	**ΔG (kcal/mol)**	**K_d (pM)**	**ΔH (kcal/mol)**	**C_p (cal/mol*K)**	**ΔS (cal/mol*K)**
Subunidad p50 (NFκB1)	Q9UBC1	-15,30	6,1	-30	-119.5	49,3
Subunidad p52 (NFκB2)	Q9BYH8	-17,96	0,07	-30	-119.5	40,4
Subunidad p65 (RelA)	963D2	-18,01	0,06	-30	-119.5	40,2
Subunidad p105 (NFκB1)	A26C8	-12,25	1048.8	-30	-119.5	59,5
CiclinaD1	Q9BYH8	-14,86	12,8	-30	-119.5	50,8
Ciclooxigenasa-2	O00221	-18,27	0,04	-30	-119.5	39,3
Interleucina-B	A26C8	-12,77	436,1	-30	-119.5	57,8
STAT3	O00221	-13,90	64.8	-30	-119.5	54,0

Anexo 7. *Datos complementarios acerca de la afinidad de las proteínas predichas (ligandos).*

Receptor	**Proteínas predichas (ligandos)**	**Afinidad (kcal/mol)**	**Figura**
Subunidad p50 (NFκB1)	Q9UBC1	-15,30	 Acoplamiento obtenido de la interacción entre la subunidad p50 y Q9UBC1

Receptor	Proteínas predichas (ligandos)	Afinidad (kcal/mol)	Figura
Subunidad p52 (NFκB2)	Q9BYH8	-17,96	 Acoplamiento obtenido de la interacción entre la subunidad p52 y Q9BYH8
Subunidad p65 (RelA)	963D2	-18,01	 Acoplamiento obtenido de la interacción entre la subunidad p65 y 963D2
Subunidad p105 (NFκB1)	A26C8	-12,25	 Acoplamiento obtenido de la interacción entre la subunidad p105 y A26C8
CiclinaD1	Q9BYH8	-14,89	 Acoplamiento obtenido de la interacción entre CD1 y Q9BYH8

Receptor	Proteínas predichas (ligandos)	Afinidad (kcal/mol)	Figura
Ciclooxigenasa-2	O00221	-18,27	 Acoplamiento obtenido de la interacción entre COX-2 y O00221
Interleucina-B	A26C8	-12,77	 Acoplamiento obtenido de la interacción entre IL-B y A26C8
STAT3	O00221	-13,90	 Acoplamiento obtenido de la interacción entre STAT3 y O00221