

Análisis in silico, estructural y funcional de la proteína ACE2 en diferentes vertebrados y su relación con SARS-COV-2

In silico, structural and functional analysis of ace2 protein in different vertebrates and its relationship to SARS-COV-2

Juan J. Pedro Huaman*

Facultad de Ciencias Biológicas, Universidad Nacional de Trujillo, Av. Juan Pablo II s/n – Ciudad Universitaria, Trujillo, Perú/Facultad de Ciencias, Universidad Privada Antenor Orrego, Trujillo, PERÚ

* Autor de correspondencia: jpedro@unitru.edu.pe

 <http://orcid.org/0000-0001-6230-3670>

Loana Moya Chávez

Facultad de Ciencias Biológicas, Universidad Nacional de Trujillo, Av. Juan Pablo II s/n – Ciudad Universitaria, Trujillo, PERÚ

 <https://orcid.org/0009-0003-2741-9119>

Recibido: 22-XII-2023; aceptado: 18-I-2024; publicado online: 30-IV-2024

Resumen

A finales del 2019, se manifestó una nueva afección respiratoria de carácter infecciosa ubicada en la provincia de Hubei Wuhan-China causada por el virus SARS-CoV-2. Por otra parte, la enzima convertidora de angiotensina 2 (ACE2) se identificó rápidamente como el receptor funcional crítico para el SARS-CoV-2. El estudio de la proteína ACE2 permite determinar los posibles reservorios que podría tener en la naturaleza coronavirus como SARS-CoV, SARS-CoV-2 entre otros. La presente investigación tiene como objetivo analizar estructural y funcionalmente la proteína ACE2 en diferentes clases de vertebrados. Para ello se recuperó un total de 35 secuencias aminoacídicas, realizándose un análisis preliminar con ProtParam, SMART, InterPro y CATHdb. La estructura secundaria de las proteínas de 10 especies seleccionadas se predijo utilizando la herramienta PSIPRED. Posteriormente, se modeló la estructura terciaria con MODELLER 10.4 y se realizó la validación mediante Procheck. El refinamiento se realizó mediante las herramientas ModRefiner y GalaxyRefine, para la alineación estructural se usó CLICK. Se observó que las propiedades fisicoquímicas, dominios y funciones de las diferentes clases de vertebrados son muy similares. De igual forma, se puede observar que la estructura secundaria encontrada con mayor prevalencia en las proteínas fue la alfa hélice. Finalmente, existe una alta conservación en la estructura tridimensional al compararla con la de *Homo sapiens*. Esta conservación evolutiva denota su importancia para los seres vivos e indica la potencial susceptibilidad de las diferentes especies a los coronavirus mediante el mismo receptor ACE2.

Palabras clave: coronavirus, receptor, modelamiento por homología, dominios, alfa hélice, hospederos.

Abstract

At the end of 2019, a new infectious respiratory condition located in Hubei Wuhan-China province caused by SARS-CoV-2 virus was manifested. Moreover, angiotensin-converting enzyme 2 (ACE2) was quickly identified as the critical functional receptor for SARS-CoV-2. The study of the ACE2 protein allows us to determine the possible reservoirs that coronaviruses such as SARS-CoV, SARS-CoV-2 and others could have in nature. The present research aims to analyze structurally and functionally the ACE2 protein in different classes of vertebrates. For this purpose, a total of 35 amino acid sequences were recovered and a preliminary analysis was performed with ProtParam, SMART, InterPro and CATHdb. The secondary structure of the proteins of 10 selected species was predicted using the PSIPRED tool. Subsequently, tertiary structure was modeled with MODELLER 10.4 and validation was performed using Procheck. Refinement was performed using ModRefiner and GalaxyRefine tools, for structural alignment CLICK was used. It was observed that the physicochemical properties, domains and functions of the different vertebrate classes are very similar. Similarly, it can be observed that the secondary structure found with the highest prevalence in the proteins was the alpha helix. Finally, there is a high conservation in the three-dimensional structure when compared with that of *Homo sapiens*. This evolutionary conservation denotes its importance for living beings and indicates the potential susceptibility of different species to coronaviruses through the same ACE2 receptor.

Keywords: coronavirus, receptor, homology modeling, domains, alpha helix, hosts.

Citación: Pedro, J. & L. Moya. 2024. Análisis in silico, estructural y funcional de la proteína ace2 en diferentes vertebrados y su relación con SARS-COV-2. *Arnaldoa* 31 (1): 151-169, 2024 doi:<http://doi.org/10.22497/arnaldoa.311.31108>

Introducción

A finales de 2019, se manifestó una nueva afección respiratoria de naturaleza infecciosa ubicada en la provincia de Hubei Wuhan, China (Huang *et al.* 2020). El brote fue denominado Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), que se transmitió rápidamente alrededor del mundo lo que provocó una emergencia sanitaria global. Por otra parte, la enzima convertidora de angiotensina 2 (ACE2), se identificó como el receptor funcional crítico del SARS-CoV-2. Además, ACE2 es bien conocida como un contrarregulador del sistema renina-angiotensina (RAS) y juega un rol importante en el sistema cardiovascular (Bian & Li, 2020).

Asimismo, la regulación positiva de ACE2 también está asociada con la mejora de la patogénesis de muchos trastornos metabólicos; esto incluye afecciones como la fibrosis hepática, enfermedad renal crónica, insuficiencia cardíaca y diabetes (Gottschalk *et al.*, 2021). De la misma forma, mejora la tasa de filtración glomerular, reduce la hipertensión sistémica y atenúa la expresión de genes inflamatorios (Lo *et al.*, 2015).

La enzima ACE2 es diana de la proteína Motivo de unión al receptor (RBM) en el dominio de unión al receptor (RBD) de SARS-CoV y funciona como receptor del SARS-CoV. De igual forma ocurre en el virus SARS-CoV-2 lo que permitiría su ingreso dentro del organismo hospedero. La especificidad de la proteína ACE2 como receptor es lo que determina el rango de hospederos que puedan tener los coronavirus dentro de la naturaleza. Se presume que los primeros virus de tipo SARS tienen su origen en animales mamíferos como los murciélagos, especulándose y comprobándose a nivel de ensayos de laboratorio que su rango

de hospederos es amplio. Esto incluye a diferentes especies como gato, perro, pangolín, hámster chino y primates no humanoides como susceptibles debido a la presencia de homólogos de ACE2. El análisis de la proteína ACE2 permite determinar los posibles reservorios de coronavirus como SARS-CoV, SARS-CoV-2, entre otros. Puesto que se presume que esta proteína está ampliamente conservada a lo largo del proceso evolutivo y podría indicar la presencia de un amplio rango de hospederos. Además, para entender mejor la interacción entre la proteína viral y su receptor es necesario conocer sus estructuras tridimensionales, especialmente haciendo énfasis en los aminoácidos clave involucrados en la unión específica (Damas *et al.* 2020; Liu *et al.* 2020; Luan *et al.* 2020; Samavati & Uhal, 2020; Liu *et al.* 2021 y Wu *et al.* 2020).

El estudio enfocado a la predicción de la estructura tridimensional en proteínas tiene gran relevancia; dado que, proporciona una comprensión detallada y eficiente de sus propiedades y funciones a nivel molecular. De esta manera se puede determinar cuál es su rol biológico, al igual que las posibles interacciones que pueden darse con otras proteínas (Singh *et al.*, 2022; Mehmood *et al.*, 2020).

Actualmente, para poder obtener un análisis detallado se están utilizando recursos computacionales, lo que puede ser más rápido y menos costoso que llevar a cabo experimentos de laboratorio. Es por ello que, la bioinformática representa un campo avanzado en la biología molecular porque ofrece el empleo de métodos *In silico* para abordar cuestionamientos que resultaban desafiantes mediante enfoques moleculares convencionales. Estas metodologías han permitido desarrollar investigaciones en

diversos aspectos como la predicción de epítomos para proteínas antigénicas de *Mycobacterium tuberculosis*, la identificación de homología con genes, el modelado de proteínas, el análisis mutacional, la evolución de genes antimicrobianos y la tasa de recombinación en genes de interés (Kumar *et al.*, 2021).

Las metodologías de predicción estructural de proteínas son herramientas versátiles que permiten modelar estructuras inclusive en casos donde ya existe una estructura identificada mediante experimentación, los métodos *in silico* pueden ser utilizados para modelar los efectos de mutaciones, anticipar la localización de superficies de unión para otras macromoléculas y evaluar los impactos de moléculas pequeñas. Asimismo, estas técnicas permiten estimar las energías de enlace y prever movimientos locales y no locales en la estructura proteica (Werner *et al.* 2012).

Finalmente, el propósito del presente estudio fue analizar estructural y funcionalmente la proteína ACE2 en diferentes especies de vertebrados. Esto puede servir para futuras investigaciones que busquen determinar el rango de posibles hospederos del virus SARS-CoV-2 basándose en la homología de la proteína receptora ACE2.

Materiales y métodos

Recuperación de las secuencias aminoacídicas de la proteína ACE2

Las secuencias de la proteína ACE2 de las diferentes especies de vertebrados se recuperó de la base de datos National Center for Biotechnology Information NCBI (<https://www.ncbi.nlm.nih.gov/>). Se recuperaron un total de 35 secuencias de la sección de ortólogos del gen ACE2, se tomaron como datos la especie, nombre común, tamaño de secuencia (número de aminoácidos) y los códigos de NCBI como puede evidenciarse en la Tabla N°1. Las secuencias fueron elegidas considerando especies que pertenezcan a las diferentes clases del Subfilo Vertebrata, incluyendo a la clase Mammalia, Ave, Reptilia, Amphibia, Sarcopterygii, Actinopterygii, Chondrichthyes y Myxini. Además, se procuró que el tamaño de las proteínas no difiera mucho y que de preferencia sean determinadas con evidencia experimental. Se descargaron las secuencias en archivos de formato FASTA y fueron almacenados para posteriores análisis (Bhattacharya *et al.*, 2018; Dutta *et al.*, 2018; Lorraine *et al.*, 2023; Mohanty *et al.*, 2022; Yasin *et al.*, 2020).

Tabla 1. Secuencias de ACE2 en especies diferentes de Vertebrados

Número	Especie	Nombre común	Tamaño de la secuencia (aa)	Código en NCBI
1	<i>Homo sapiens</i>	Humano	805	NP_068576
2	<i>Petromyzon marinus</i>	Lamprea marina	849	XP_032835032.1
3	<i>Latimeria chalumnae</i>	Celacanto	859	XP_005997915.2
4	<i>Protopterus annectens</i>	Pez pulmonado	837	XP_043928726.1
5	<i>Amblyraja radiata</i>	Raya radiante	773	XP_032888812.1
6	<i>Callorhinchus milii</i>	Tiburón elefante	836	XP_007889845.1
7	<i>Scomber japonicus</i>	Pez caballa	815	XP_053176828.1
8	<i>Oreochromis niloticus</i>	Tilapia del nilo	821	XP_003445853.2
9	<i>Danio rerio</i>	Pez cebra	785	NP_001007298.1
10	<i>Microcaecilia unicolor</i>	Cecilia	811	XP_030058174.1
11	<i>Xenopus tropicalis</i>	Xenopus	862	XP_002938293.2
12	<i>Bufo bufo</i>	Sapo común	833	XP_040279250.1
13	<i>Gekko japonicus</i>	Gecko	816	XP_015273067.1
14	<i>Alligator sinensis</i>	Aligátor chino	803	XP_025066628.1
15	<i>Crocodylus porosus</i>	Cocodrilo marino	803	XP_019384826.1
16	<i>Chelonia mydas</i>	Tortuga verde marina	828	XP_037756773.2
17	<i>Varanus komodoensis</i>	Dragón de Komodo	707	XP_044305433.1
18	<i>Columbia livia</i>	Paloma	799	XP_021154486.1
19	<i>Meleagris gallopavo</i>	Pavo	861	XP_019467554.1
20	<i>Athene cunicularia</i>	Lechuza de los arenales	834	XP_026705725.1
21	<i>Falco peregrinus</i>	Cernícalo	813	XP_005231984.3
22	<i>Egretta garzetta</i>	Garceta común	806	XP_009638257.1
23	<i>Capra aegagrus hircus</i>	Cabra	804	NP_001277036.1
24	<i>Vicugna pacos</i>	Alpaca	805	XP_006212709.1
25	<i>Ornithorhynchus anatinus</i>	Ornitorrinco	806	XP_001515597.2
26	<i>Trichechus manatus latirostris</i>	Manatí del caribe	800	XP_004386381.1
27	<i>Sus scrofa</i>	Jabalí	805	NP_001116542.1
28	<i>Canis lupus familiaris</i>	Perro	804	NP_001158732.1
29	<i>Oryctolagus cuniculus</i>	Conejo	805	XP_002719891.1
30	<i>Felis catus</i>	Gato	805	NP_001034545.1
31	<i>Ovis aries</i>	Oveja	804	XP_011961657.1
32	<i>Bos taurus</i>	Toro	804	NP_001019673.2
33	<i>Macaca mulatta</i>	Mono Rhesus	805	NP_001129168.1
34	<i>Gorilla gorilla</i>	Gorila	805	XP_018874749.1
35	<i>Pan troglodytes</i>	Chimpancé	805	XP_016798468.1

Análisis de parámetros físico-químicos de la proteína ACE2

Las propiedades físicas y químicas fueron determinadas mediante análisis computacionales utilizando el servidor ProtParam (<https://web.expasy.org/protparam/>). Las propiedades evaluadas fueron Residuos, Masa molecular, Punto Isoeléctrico (pI), Índice de Inestabilidad (II), Estabilidad (tt), Índice alifático, Valor GRAVY y Coeficiente de Extinción de la proteína ACE2 de las 35 secuencias de vertebrados elegidas (Amobonye *et al.*, 2022; Premachandran & Srinivasan, 2022; Satyanarayana *et al.*, 2018; Singh *et al.*, 2022 y Bhattacharya *et al.*, 2019).

Análisis preliminar de secuencias de la proteína ACE2

El análisis preliminar de la estructura primaria busca determinar la presencia de motivos, dominios conservados, clasificación en familias de proteínas e inferir su función. En el análisis preliminar de secuencias se utilizó la herramienta bioinformática online InterPro (<https://www.ebi.ac.uk/interpro/>) para determinar los dominios y clasificarlas en familias. Además, la anotación funcional y de dominios se complementa con la herramienta SMART (<http://smart.embl-heidelberg.de/>). Asimismo, en la búsqueda de información adicional sobre la función, clasificación estructural y de superfamilias se utilizó la herramienta CATH (<https://www.cathdb.info/>) (Flores-Castañón *et al.*, 2022; Goh *et al.*, 2022; Lenin *et al.*, 2023; Mukherjee *et al.*, 2023 y Singh *et al.*, 2021).

Predicción de estructura secundaria de la proteína ACE2

La predicción de la estructura secundaria de la proteína ACE2 de *Homo sapiens* y sus ortólogos se realizó con 10

secuencias seleccionadas de las 35 especies de vertebrados previamente elegidas. De las distintas clases del Subfilo Vertebrata, se consideró *Petromyzon marinus* de la clase Myxini, *Amblyraja radiata* de Chondrichthyes, *Scomber japonicus* de Actinopterygii, *Bufo bufo* de Amphibia, *Chelonia mydas* de Reptilia, *Columbia livia* de Ave y de la clase Mammalia se incluyó *Canis lupus familiaris*, *Felis catus*, *Pan troglodytes* y *Homo sapiens*.

Para su realización se utilizó el servidor PSIPRED 4.0 (<http://bioinf.cs.ucl.ac.uk/psipred/>) con las opciones por defecto obteniéndose la visualización esquemática de la estructura secundaria considerando la alfa hélice, beta hoja plegada y coils. Dicho servidor online es simplificado, confiable y permite llevar a cabo la predicción de estructuras secundarias empleando redes neuronales para analizar las secuencias mediante PSI-BLAST (Kumar *et al.*, 2021; Shrivastava *et al.*, 2020 y Xu *et al.*, 2021).

Predicción de estructura terciaria de la proteína ACE2

La estructura terciaria de las proteínas fue predicha utilizando el enfoque por homología mediante el uso del software MODELLER 10.4. La búsqueda de los templados se llevó a cabo usando BLAST contra la base de datos Protein Data Bank PDB (<https://www.rcsb.org/>) utilizando como query a las secuencias FASTA de las diferentes especies, considerando el porcentaje de identidad, query coverly y eligiendo solo modelos obtenidos de análisis experimentales. Los modelos estructurales de los templados en formato PDB fueron editados utilizando PyMOL 2.5.5 para remover las cadenas adicionales como la de la proteína Spike de SARS-CoV-2 y átomos adicionales antes de su utilización. Los archivos fueron preparados

adecuadamente y se generaron 5 modelos aleatorizados predichos utilizando MODELLER, almacenados en formato PDB y seleccionando el que tiene mejor valor DOPE como nuestro top-model para los siguientes pasos (Kumar *et al.*, 2020 y Rout & Mahapatra, 2019).

Los modelos predichos deben ser evaluados y verificados para conocer su calidad y mejorarlos, para dicha labor se utilizaron las herramientas bioinformáticas disponibles en la web de UCLA-DOE LAB – SAVES v6.0 (<https://saves.mbi.ucla.edu/>), donde se evaluaron los parámetros que se contemplan en la herramienta PROCHECK, enfocándose principalmente en el Ramachandran Plot para el control de calidad de la estabilidad estructural (Lorraine *et al.*, 2023; Mohanty *et al.*, 2022 y Premachandran & Srinivasan, 2022). Aquellos modelos que no cumplieron con los parámetros evaluados, procedieron a ser mejorados mediante el refinamiento con los servidores bioinformáticos online ModRefiner (<https://zhanggroup.org/ModRefiner/>) y GalaxyRefine (<https://galaxy.seoklab.org/cgi-bin/submit.cgi?type=REFINE>). Posteriormente, se volvió a evaluar la calidad de los modelos resultantes en PROCHECK. Este proceso se realizó iterativamente hasta lograr un modelo adecuado (Amobonye *et al.*, 2022; Desai & Chauhan, 2018). Finalmente, los modelos fueron visualizados utilizando Rasmol y PyMOL para obtener la ilustración de las estructuras más óptimas correspondientes a la estructura terciaria de cada especie.

Comparación estructural de la proteína ACE2 de diferentes especies de vertebrados

Con el objetivo de poder comparar adecuadamente los modelos estructurales, se realizó superposición de los mismos

mediante un alineamiento estructural átomo a átomo empleando el servidor web CLICK (<http://cospi.iiserpune.ac.in/click/>) el cual se encarga de establecer una comparación independiente de topología de la estructura tridimensional de las proteínas, llevándose a cabo con los ajustes por defecto y mejorando la visualización de las mismas para determinar la similitud estructural (Nguyen *et al.* 2011).

Resultados

Parámetros físico-químicos de la proteína ACE2

Se calcularon los parámetros físico-químicos como Residuos, Masa molecular, pI, Índice de Inestabilidad, Estabilidad, Índice alifático, Valor GRAVY y Coeficiente de Extinción como se puede apreciar en la Tabla N°2. El valor de pI para *Homo sapiens* es de 5.36 lo que es considerado una carga positiva en soluciones ácidas. Las demás especies analizadas presentan valores entre 5 y 6 aproximadamente. La especie con valor más bajo de pI es *Sus scrofa* con 4.95 mientras que el valor más alto fue de 6.03 con una especie de *Meleagris gallopavo*, por lo que este parámetro está altamente conservado a lo largo del proceso evolutivo. Los valores de pI ilustran cómo la composición de aminoácidos de las proteínas influye en su comportamiento en diferentes entornos de pH.

Uno de los parámetros físico-químicos que tiene gran relevancia es el GRAVY (Gran Promedio de Hidrofobicidad) el cual proporciona datos sobre la capacidad de solubilidad de una proteína. La especie que obtuvo el valor más alto fue *Meleagris gallopavo* (-0.257) mientras que el valor más bajo perteneció a la especie *Latimeria chalumnae* (-0.484), una especie bastante antigua y poco evolucionada de celacanto.

La presente investigación muestra valores de Índice de Inestabilidad en un rango de 33.03 a 44.68. Para *Homo sapiens*, fue de 40.09 lo que sugiere que es una proteína levemente inestable en solución y se tienen que tomar consideraciones para su almacenamiento. De la totalidad de organismos, 18 serían inestables como por ejemplo *Egretta garzetta* y *Athene cunicularia*, mientras que 17 serían

estables como *Microcaecilia unicolor*, *Alligator sinensis* entre otras.

Finalmente, el rango de valores para el Índice Alifático fue de 73.46 a 82.37, siendo 80.55 para *Homo sapiens* lo que indica una alta termoestabilidad. La más alta termoestabilidad corresponde a *Meleagris gallopavo* mientras que la menos termoestable sería *Latimeria chalumnae*.

Tabla N°2. Parámetros físico-químicos de ACE2 en especies diferentes de Vertebrados

Especie	Residuos	Mr	pI	Ins. Index (II)	Stability (tt)	Aliph. Index	GRA-VY	Ex. Coeff.	Ex. Coeff. Red.
<i>Homo sapiens</i>	805	92463.04	5.36	40.09	Inestable	80.55	-0.375	176170	175670
<i>Petromyzon marinus</i>	849	95803.40	5.48	39.72	Estable	73.65	-0.447	179150	178650
<i>Latimeria chalumnae</i>	859	99210.65	5.74	34.22	Estable	73.46	-0.484	190165	189540
<i>Protopterus annectens</i>	837	96265.01	5.06	36.91	Estable	77.07	-0.418	189455	189080
<i>Amblyraja radiata</i>	773	89622.38	5.61	34.53	Estable	78.36	-0.354	175140	174640
<i>Callorhinchus milii</i>	836	96238.15	5.57	36.34	Estable	75.05	-0.482	181670	181170
<i>Scomber japonicus</i>	815	94153.56	4.99	36.24	Estable	77.53	-0.424	171715	171090
<i>Oreochromis niloticus</i>	821	94992.45	5.10	41.00	Inestable	75.91	-0.476	190040	189540
<i>Danio rerio</i>	785	90862.23	5.09	38.25	Estable	75.95	-0.435	179735	179110
<i>Microcaecilia unicolor</i>	811	93632.20	5.73	33.03	Estable	76.10	-0.393	182590	182090
<i>Xenopus tropicalis</i>	862	98849.94	5.10	42.08	Inestable	76.52	-0.426	185110	184610
<i>Bufo bufo</i>	833	95983.76	5.24	41.34	Inestable	80.43	-0.402	198060	197560
<i>Gekko japonicus</i>	816	94044.58	5.43	40.42	Inestable	78.55	-0.399	200705	200080
<i>Alligator sinensis</i>	803	92212.27	5.97	33.44	Estable	77.85	-0.359	185235	184610
<i>Crocodylus porosus</i>	803	92327.14	5.76	34.05	Estable	77.10	-0.371	181350	180600
<i>Chelonia mydas</i>	828	95040.59	5.31	38.54	Estable	76.81	-0.418	185110	184610
<i>Varanus komodoensis</i>	707	81876.10	5.52	36.10	Estable	74.67	-0.420	175140	174640
<i>Columba livia</i>	799	undefined	-	37.11	Estable	80.56	-0.312	170350	169600
<i>Meleagris gallopavo</i>	861	97758.31	6.03	41.86	Inestable	82.37	-0.257	189160	188160

Especie	Resi- duos	Mr	pl	Ins. Index (II)	Stability (tt)	Aliph. Index	GRA- VY	Ex. Coeff.	Ex. Coe- ff. Red.
<i>Athene cunicularia</i>	834	96057.97	5.36	44.32	Inestable	77.06	-0.420	186850	186100
<i>Falco peregrinus</i>	813	93928.70	5.28	42.76	Inestable	78.23	-0.381	186725	186100
<i>Egretta garzetta</i>	806	93492.63	5.45	44.68	Inestable	76.36	-0.448	196695	196070
<i>Capra aegagrus hircus</i>	804	93282.69	5.20	38.56	Estable	76.78	-0.462	190610	190110
<i>Vicugna pacos</i>	805	92726.33	5.11	40.51	Inestable	79.11	-0.377	193130	192630
<i>Ornithorhynchus anatinus</i>	806	92853.16	5.46	41.64	Inestable	76.13	-0.478	179610	179110
<i>Trichechus manatus latirostris</i>	800	91527.57	5.15	41.01	Inestable	77.15	-0.387	176170	175670
<i>Sus scrofa</i>	805	92161.59	4.95	43.41	Inestable	80.21	-0.346	186140	185640
<i>Canis lupus familiaris</i>	804	92618.77	5.16	42.68	Inestable	76.88	-0.431	194620	194120
<i>Oryctolagus cuniculus</i>	805	92653.75	4.97	44.33	Inestable	75.95	-0.422	177660	177160
<i>Felis catus</i>	805	92708.08	5.64	42.67	Inestable	76.43	-0.436	188660	188160
<i>Ovis aries</i>	804	93283.71	5.20	37.83	Estable	76.65	-0.465	190610	190110
<i>Bos taurus</i>	804	93101.39	5.19	39.91	Estable	76.90	-0.447	189120	188620
<i>Macaca mulatta</i>	805	92478.78	5.18	40.05	Inestable	80.55	-0.397	177660	177160
<i>Gorilla gorilla</i>	805	92476.08	5.33	40.32	Inestable	80.67	-0.374	176170	175670
<i>Pan troglodytes</i>	805	92445.91	5.33	39.98	Estable	80.06	-0.389	176170	175670

Nota: Mr: Masa molecular, pl: Punto isoeléctrico, Ins. Index (II): Índice de Inestabilidad, Stability (tt): Estabilidad, Aliph. Index: Índice alifático, GRAVY: Gran promedio de Hidrofobicidad, Ex. Coeff.: Coeficiente de extinción y Ex. Coeff. Red.: Coeficiente de extinción reducido.

Análisis preliminar de la proteína ACE2 de *Homo sapiens* y sus ortólogos

Se ha realizado un análisis detallado utilizando herramientas bioinformáticas como INTERPRO, SMART y CATH. Donde se evidencia que pertenecen a la familia Peptidase M2 del dominio Collectrin (617-770, 590-742, 620-773, 630-782, 623-776, 637-790, 607-760, 617-770, 616-769, 617-770) correspondientes a las especies *Homo sapiens*, *Petromyzon marinus*, *Amblyraja radiata*, *Scomber japonicus*, *Bufo bufo*, *Chelonia mydas*, *Columba livia*, *Canis lupus familiaris*, *Felis catus* y *Pan troglodytes* respectivamente.

A su vez, también destaca por presentar como función molecular actividad metalopeptidasa y actividad peptidil-dipeptidasa en la membrana plasmática y el espacio extracelular. Sin embargo, al analizar las secuencias con la herramienta CATH, no se lograron encontrar coincidencias.

Predicción de estructura secundaria de la proteína ACE2 de *Homo sapiens* y sus ortólogos

La predicción de estructuras secundarias se puede observar en la Tabla N°3. En dicha tabla se puede apreciar el porcentaje de Alfa Hélice, Beta Hoja Plegada y Coils

de cada una de las secuencias. Para *Homo sapiens*, la alfa hélice es la más frecuente en la conformación estructural con 57.64% mientras que la estructura menos frecuente para dicha proteína fue la beta hoja plegada con 2.73. El rango de alfa hélice varía entre 52.88 para *Petromyzon marinus* y 60.67 para *Amblyraja radiata*, lo que nos indica que es independiente del nivel evolutivo. Para la estructura beta hoja plegada los valores varían entre 2.48 para *Felis catus* y 3.12 para *Bufo bufo*, nuevamente no se puede inferir de manera directa una relación con el nivel evolutivo. Finalmente, el rango de valores para la estructura Coil fue de 36.35 para *Amblyraja radiata* y 44.53 de *Petromyzon marinus*, por lo cual no se puede inferir una asociación entre el grado evolutivo y la variación entre las diferentes estructuras secundarias de la proteína ACE2.

Tabla N°3. Estructura secundaria de la proteína ACE2 en especies diferentes de Vertebrados

	Alfa Hélice (%)	Beta Hoja Plegada (%)	Coils (%)
<i>Homo sapiens</i>	57.64	2.73	39.63
<i>Petromyzon marinus</i>	52.88	2.59	44.53
<i>Amblyraja radiata</i>	60.67	2.98	36.35
<i>Scomber japonicus</i>	58.16	2.58	39.26
<i>Bufo bufo</i>	56.06	3.12	40.82
<i>Chelonia mydas</i>	57.97	2.90	39.13
<i>Columba livia</i>	56.44	3.00	40.67
<i>Canis lupus familiaris</i>	58.21	2.98	38.81
<i>Felis catus</i>	57.39	2.48	40.13
<i>Pan troglodytes</i>	58.21	2.98	38.81

Predicción de estructura terciaria de la proteína ACE2.

Como se puede visualizar en la figura, la estructura predominante para las 10 especies estudiadas fue la de alfa hélice; seguida de la estructura coil y la estructura beta hoja plegada que es la que se aprecia en una menor cantidad.

Un aspecto a destacar, es la divergencia observada en los modelos estructurales expuestos, teniendo una mayor pronunciación en el modelo 2 (*Petromyzon marinus*), con respecto a las demás especies.

Alineamiento estructural de los modelos estructurales de la proteína ACE2

Se estableció una comparación estructural entre los modelos previamente predichos para la proteína ACE2, de *Homo sapiens* y sus ortólogos. Esto gracias a la herramienta bioinformática CLICK, como se puede evidenciar en la figura 2. De esta manera fue posible evaluar el grado de similitud que existe entre la estructura tridimensional target con respecto a las otras especies.

Tomando en cuenta los resultados obtenidos, podemos evidenciar que existe una total correspondencia estructural entre *Homo sapiens* y *Pan troglodytes* con un porcentaje de superposición del 100%.

Por otra parte, la semejanza entre la estructura de *Homo sapiens* y las demás especies, también obtuvo altos porcentajes. Sin embargo, se evidenció que la alineación entre *Homo sapiens* y *Petromyzon marinus* dio como resultado un 27.30%, siendo este el menor valor obtenido.

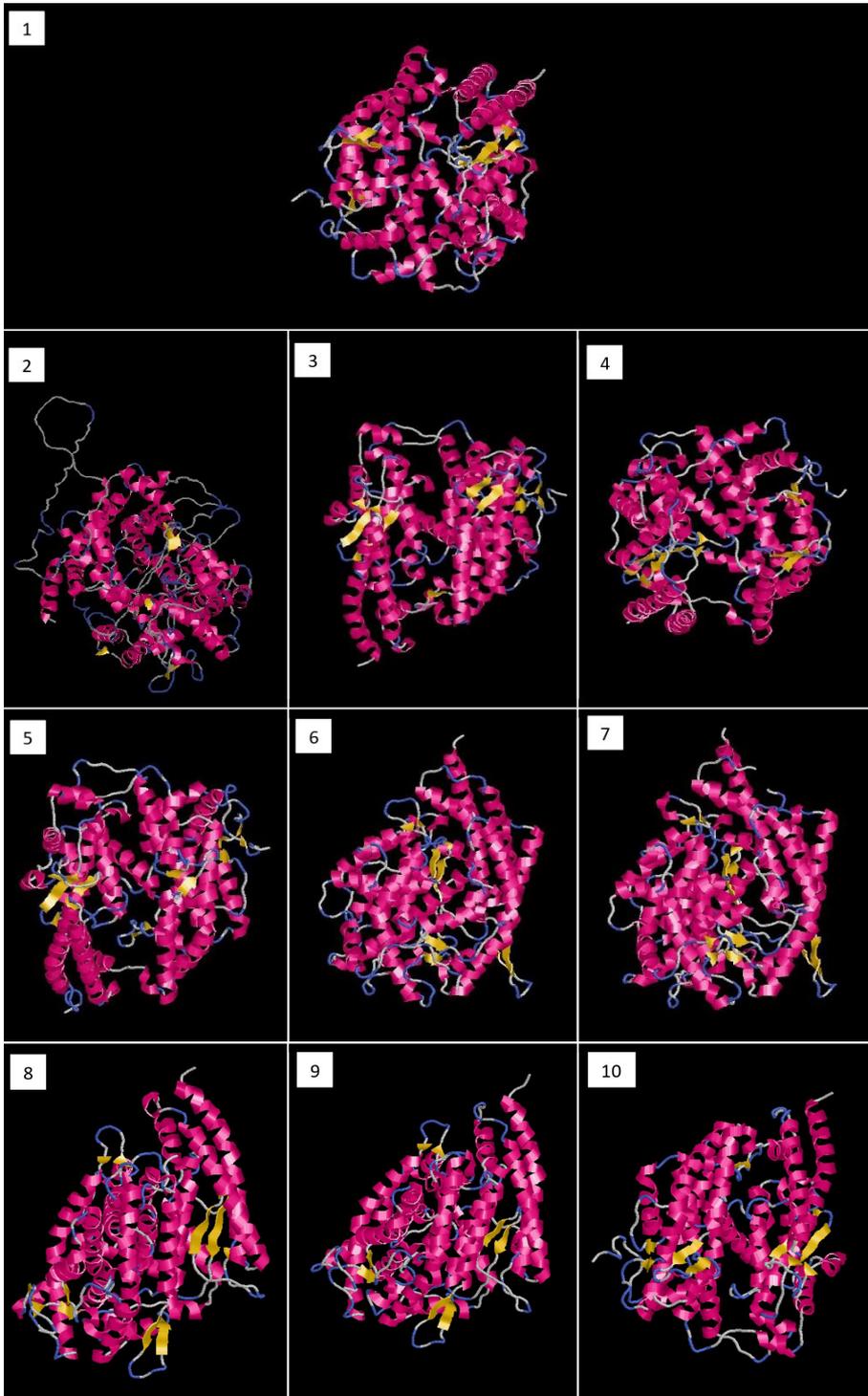


Figura N° 1. Modelos estructurales iniciales predichos de la proteína ACE2 de diferentes especies.

Nota: 1. *Homo sapiens* 2. *Petromyzon marinus*, 3. *Amblyraja radiata*, 4. *Scomber japonicus*, 5. *Bufo bufo*, 6. *Chelonia mydas*, 7. *Columba livia*, 8. *Canis lupus familiaris*, 9. *Felis catus* y 10. *Pan troglodytes*

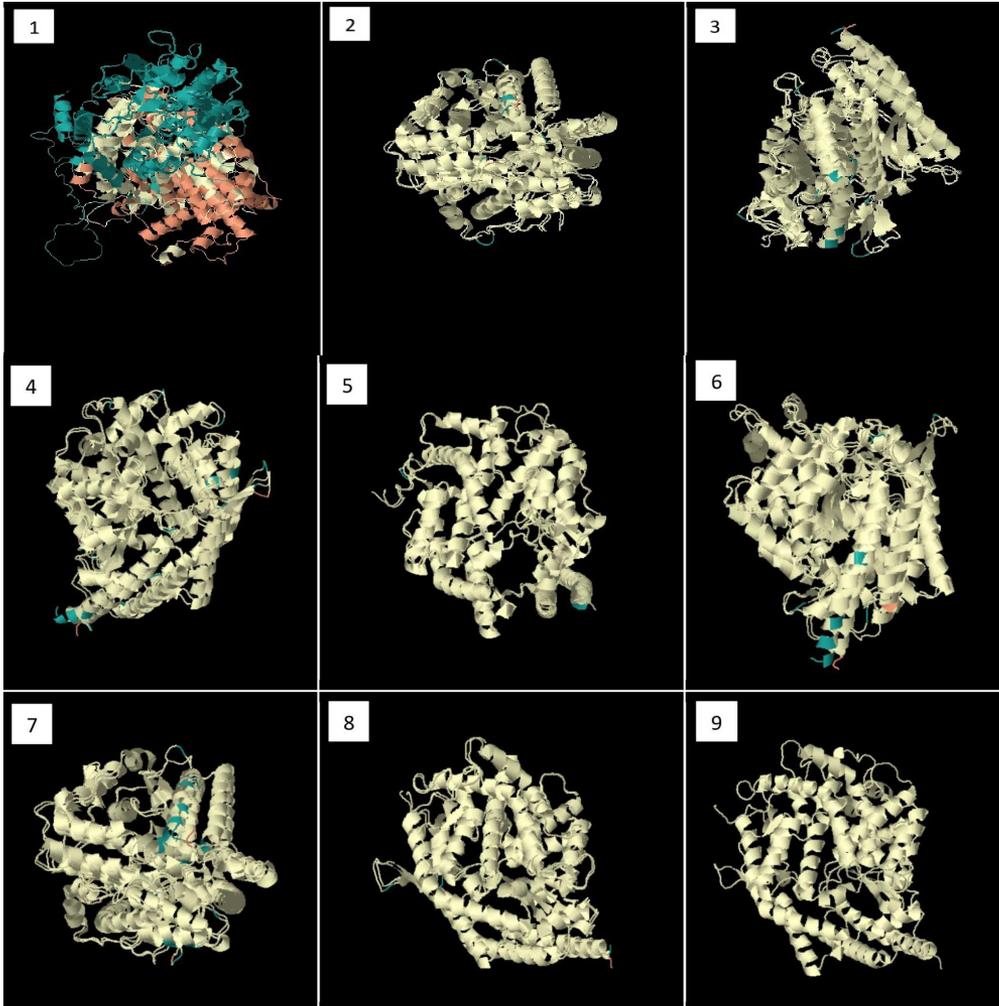


Figura N° 2. Alineamiento estructural de los modelos predichos de la proteína ACE2 de diferentes especies con *Homo sapiens*.

Nota: 1. *Petromyzon marinus*, 2. *Amblyraja radiata*, 3. *Scomber japonicus*, 4. *Bufo bufo*, 5. *Chelonia mydas*, 6. *Columba livia*, 7. *Canis lupus familiaris*, 8. *Felis catus* y 9. *Pan troglodytes*.

Discusión

La proteína ACE2 desempeña un papel fundamental en las diferentes formas de vida, sobre todo en el ser humano, puesto que participa en la regulación osmótica como parte del eje Sistema Renina-Angiotensina-Aldosterona (SRA), clivando la Angiotensina II en Angiotensina 1-7 que tiene efectos importantes sobre la regulación del Sistema Cardiovascular. Además, está involucrada en la entrada del virus SARS-Cov-2 en las células hospederas, mediante su interacción con la proteína RBD del mismo.

Este virus provoca una respuesta inflamatoria, inmunitaria e infecciosa, así como también genera alteraciones en el SRA, lo que conlleva a lesiones en el pulmón y otros órganos (Bejoy *et al.*, 2023; de Alcántara-Santos *et al.*, 2021 y Waryah *et al.*, 2023).

Los parámetros físico-químicos resultan determinantes de propiedades de interés en las proteínas. El número de Residuos, Masa molecular, Punto Isoeléctrico, Índice de Inestabilidad, Estabilidad, Índice Alifático, Valor GRAVY y Coeficiente de Extinción fueron determinados para las 35 secuencias de diferentes especies de vertebrados. Donde se observó que en su mayoría presentan un número de residuos parecido y cercano a la media de 805 aminoácidos, lo que estaría asociado con el grado de conservación de la estructura de la proteína. En relación al Punto Isoeléctrico (pI), se entiende como el valor de pH en el cual las proteínas alcanzan un equilibrio entre sus cargas (generalmente entre 5 a 6), por lo que no existen fuerzas electromagnéticas de repulsión y se precipitan de la manera más adecuada (Nolsoe & Undeland, 2009). Las proteínas de ACE2 estudiadas tienen un valor de pI que oscila entre 5 y 6 lo que

indica que necesitan un pH ácido para alcanzar un equilibrio entre sus cargas.

Un valor de GRAVY negativo indica que la proteína es no polar y por naturaleza su interacción con el agua es de carácter hidrofílico (Ahmad *et al.* 2023). En *Homo sapiens*, como principal secuencia objetivo, tuvo un GRAVY de -0.375 y para las demás especies evaluadas, se encontraron valores negativos en este parámetro, sugiriendo que tienen un comportamiento hidrofílico debido a la presencia de regiones aminoácidas de este tipo. Esto indicaría que su naturaleza hidrofílica se ha conservado desde la especie que ha experimentado una menor evolución hasta la más evolucionada.

Por su parte, los Índices de Inestabilidad brindan información sobre cuál sería el mejor solvente para el almacenaje de las proteínas. Los valores inferiores a 40 indican que la proteína es estable, si los valores son más bajos presentan una mayor estabilidad mientras que valores superiores a 40 indican una inestabilidad de la proteína en solución (Grasso *et al.*, 2016). Con respecto a este índice, se observó un comportamiento diferenciado y presentando variaciones de la estabilidad dentro de las distintas clases de Vertebrados. Esto permite inferir si las proteínas serían fáciles o difícilmente conservadas en solución y no está relacionado con un comportamiento evolutivo conservado.

El coeficiente de extinción señala la cantidad de luz que puede absorber una proteína a una longitud de onda de 280 nm, lo que alude de manera implícita la cantidad de proteína presente en una solución y permite la realización de estudios cuantitativos de proteínas (Satyanarayana *et al.*, 2018 y Dutta *et al.* 2018). Para el caso de *Gekko japonicus*, se obtuvo el valor más

alto indicando una mayor presencia de proteínas en la solución, mientras que, el valor más bajo se encontró para *Columba livia* por lo cual no sería posible relacionarlo con el nivel evolutivo de manera directa. A su vez, el índice alifático es un parámetro físico-químico que sugiere el volumen relativo que ocupan las cadenas alifáticas laterales de aminoácidos como la valina, alanina, leucina e isoleucina en proteínas sometidas a altas temperaturas, dejando entrever su termoestabilidad.

En lo que respecta a los dominios proteicos, son importantes puesto que permiten comprender la función molecular de las proteínas y clasificarlas en familias. Para el caso de las diferentes especies de vertebrados se observó un alto grado de conservación de estos dominios ya que todas presentan el dominio colectrina y peptidasa M2, lo que las coloca en la familia de proteínas Peptidasa M2 y su función de clivaje de proteínas (Dutta *et al.*, 2018; Lorraine *et al.*, 2023 y Sharma *et al.*, 2023).

Asimismo, la estructura secundaria de las proteínas resulta ser de vital importancia en el análisis de sus propiedades teniendo como base la formación de puentes de hidrógeno entre aminoácidos no contiguos. Haciendo un enfoque en las 10 especies de vertebrados seleccionadas, se observaron diferencias porcentuales entre las estructuras predichas con PSIPRED 4. La principal especie de estudio fue *Homo sapiens*, la cual obtuvo el porcentaje más alto de alfa hélice. que al mismo tiempo se encuentra relacionado con su presencia como proteína integral transmembrana (Feng *et al.* 2020 y Sang *et al.* 2019). Sin embargo, analizando las diferentes especies de vertebrados, se puede inferir que la composición de alfa hélice es independiente del nivel evolutivo, puesto que *Amblyraja radiata* tiene el valor más alto a pesar de su bajo nivel evolutivo,

mientras que *Petromyzon marinus* presentó el valor más bajo y el menor nivel evolutivo dentro de las especies estudiadas.

Por su parte, la estructura tipo coil o giro fue la segunda más abundante que se encontró como resultado de la predicción estructural secundaria de las secuencias de proteínas. Las estructuras tipo coils resultan importantes en la estabilidad y flexibilidad así como su interacción molecular con otras proteínas (De Brevern, 2022; Hoang *et al.*, 2019; Xu & Kennedy, 2020; Zhong *et al.*, 2020). Al igual que con la estructura alfa hélice, no se puede inferir una asociación entre el nivel evolutivo y la composición porcentual de coils. La estructura tipo beta hoja plegada fue la menos frecuente entre todas las especies de vertebrados estudiadas. Del mismo modo, estas estructuras también son importantes por su rol en la estabilidad de las proteínas (Dhanjal *et al.*, 2019).

Para la predicción de modelos estructurales fueron utilizadas diferentes herramientas bioinformáticas. La bioinformática estructural trata de completar el vacío de información que existe respecto a la estructura tridimensional, puesto que técnicas como la cristalografía de Rayos X, Resonancia Nuclear y Magnética y la Microscopía electrónica no se pueden utilizar tan fácilmente, ni a gran escala como para estudiar todas las secuencias de proteínas disponibles debido a su elevado costo y condiciones técnicas necesarias. Existe una alta diversidad de métodos de modelaje computacional de proteínas, algunos requieren el uso previo de bases de datos, el reconocimiento de pliegues, determinación de moldes o plantados y el alineamiento comparativo de proteínas. El modelado de proteínas se fundamenta en que existe una alta similitud entre las secuencias durante el proceso evolutivo y la homología de las diferentes proteínas, especialmente las

proteínas más cercanas resultan ser las más conservadas. Entre los diferentes modelos predichos, debemos de considerar que la estructura más predominante fue la alfa hélice, seguida de los coils, lo que coincide con lo observado a nivel de la estructura secundaria predicha y resultan bastante familiares. Para comprobar la calidad de los modelos predichos se utilizaron servidores de validación que indicaron que los modelos podrían mejorarse, se utilizaron programas online de refinamiento y se volvió a verificar la calidad hasta conseguir un modelo adecuado o al menos aceptable (Dorn *et al.*, 2014; Huang *et al.*, 2023; Werner *et al.*, 2012).

Para confirmar la homología estructural entre las diferentes proteínas, se realizó un alineamiento estructural con superposición de las estructuras similares. Se comprobó que la identidad entre las estructuras predichas superó inclusive el 99% para todas las comparaciones entre las diferentes especies con *Homo sapiens*. Solamente se encontró una marcada diferencia con *Petromyzon marinus*, en cuyo caso solamente hubo una identidad de 27%, lo que demuestra que esta proteína ha diferido mucho en su proceso evolutivo con todas las demás. Sin embargo, también fue el modelo predicho que presentó el mayor margen de error en su evaluación y la posible diferencia podría deberse a dificultades no superadas durante el modelaje y refinamiento de la estructura tridimensional de la proteína. Este hallazgo resulta significativo debido a que se podría deducir que existe reconocimiento grupal de ligandos basado en receptores homólogos (Kumar *et al.*, 2021; Pollet *et al.*, 2022; Sanchez-Pulido & Ponting, 2021).

Conclusiones

Se encontró que los parámetros físico-químicos de la proteína ACE2 entre las distintas clases de Vertebrados son muy variables. Además, se identificó que el dominio colectrina y peptidasa M2 se encontraba en todas las especies estudiadas, lo que indica su alto grado de conservación a lo largo del proceso evolutivo. También, se estableció que la estructura más predominante en los modelos tridimensionales fue la alfa hélice relacionada con su ubicación de proteína transmembrana y dicha característica está ampliamente conservada entre las diferentes clases de vertebrados.

Agradecimientos

A la Universidad Nacional de Trujillo y la Universidad Privada Antenor Orrego por su generosa colaboración y por proporcionar las facilidades de uso de infraestructura y equipamiento esenciales para llevar a cabo la presente investigación. Su apoyo ha sido fundamental para el éxito de este proyecto por lo que es debido hacer presente el profundo agradecimiento por la invaluable contribución de ambas instituciones.

Contribución de los autores

J.P. y L.M.: Concepción, diseño, ejecución y análisis de resultados, interpretación y discusión de los hallazgos, revisión exhaustiva del artículo y aprobación de la versión definitiva. Cada uno de los autores ha revisado el manuscrito y dado su visto bueno a la edición final del artículo.

Conflicto de intereses

Los autores a cargo de esta investigación declaran que no existe conflicto de interés.

Literatura citada

- Amobonye, A.; S. Singh; K. Mukherjee; C. Jobichen; I. Qureshi & S. Pillai.** 2022. Structural and functional insights into fungal glutaminase using a computational approach. *Process Biochemistry*, 117: 76–89. <https://doi.org/10.1016/j.procbio.2022.03.019>
- Ahmad, I.; A. K. Singh & S. K. Katari.** 2023. In silico insight into structural and functional attributes of arsenic resistance proteins from *Rhizobium radiobacter* strain F4. *Journal of Hazardous Materials Advances*, 12(100329): 100329. <https://doi.org/10.1016/j.hazadv.2023.100329>
- Bejoy, J.; C. I. Williams; H. J. Cole; S. Manzoor; P. Davoodi; J. I. Battaile; A. Kaushik; S. I. Nikolaienko; T. I. Brelidze; S. G. Gychka & Y. J. Suzuki.** 2023. Effects of spike proteins on angiotensin converting enzyme 2 (ACE2). *Archives of Biochemistry and Biophysics*, 748(109769): 109769. <https://doi.org/10.1016/j.abb.2023.109769>
- Bhattacharya, M.; A. Hota; A. Kar; D. Sankar; R. Chandra-Malick; B. Chandra-Patra & B. Kumar-Das.** 2018. In silico structural and functional modelling of Antifreeze protein (AFP) sequences of Ocean pout (*Zoarces americanus*, Bloch & Schneider 1801). *Journal, Genetic Engineering & Biotechnology*, 16(2): 721–730. <https://doi.org/10.1016/j.jgeb.2018.08.004>
- Bhattacharya, S.; S. Dhar; A. Banerjee & S. Ray.** 2019. Structural, functional, and evolutionary analysis of late embryogenesis abundant proteins (LEA) in *Triticum aestivum*: A detailed molecular level biochemistry using in silico approach. *Computational Biology and Chemistry*, 82: 9–24. <https://doi.org/10.1016/j.compbiolchem.2019.06.005>
- Bian, J. & Z. Li.** 2021. Angiotensin-converting enzyme 2 (ACE2): SARS-CoV-2 receptor and RAS modulator. *Acta Pharmaceutica Sinica B*, 11: 1-12. <https://doi.org/10.1016/j.apsb.2020.10.006>
- Damas, J.; G. M. Hughes; K. C. Keough; C. A. Painter; N. S. Persky; M. Corbo; M. Hiller; K. Koepfli; A. R. Pfening; H. Zhao; D. P. Genereux; R. Swofford; K. S. Pollard; O. A. Ryder; M. T. Nweeia; K. Lindblad-Toh; E. C. Teeling; E. K. Karlsson & H. A. Lewin.** 2020. Broad host range of SARS-CoV-2 predicted by comparative and structural analysis of ACE2 in vertebrates. *Proceedings of the National Academy of Sciences of the United States of America*, 117(36): 22311–22322. <https://doi.org/10.1073/pnas.2010146117>
- De Alcantara-Santos, R.; V. Guzzoni; K. A. S. Silva; D. S. Aragão; R. de Paula Vieira; N. Bertoncello; N. Schor; F. Aimbire; D. E. Casarini & T. S. Cunha.** 2021. Resistance exercise shifts the balance of renin-angiotensin system toward ACE2/Ang 1–7 axis and reduces inflammation in the kidney of diabetic rats. *Life Sciences*, 287(120058): 120058. <https://doi.org/10.1016/j.lfs.2021.120058>
- De Brevern, A. G.** 2022. A perspective on the (rise and fall of) protein β -turns. *International Journal of Molecular Sciences*, 23(20): 12314. <https://doi.org/10.3390/ijms232012314>
- Desai, M. & J. B. Chauhan.** 2018. Computational analysis for the determination of deleterious nsSNPs in human MTHFR gene. *Computational biology and chemistry*, 74: 20–30. <https://doi.org/10.1016/j.compbiolchem.2018.02.022>
- Dhanjal, J. K.; V. Malik; N. Radhakrishnan; M. Sigar; A. Kumari & D. Sundar.** 2019. Computational protein engineering approaches for effective design of new molecules. En S. Ranganathan, M. Gribskov, K. Nakai, & C. Schönbach (Eds.), *Encyclopedia of Bioinformatics and Computational Biology*: 631–643. Elsevier.
- Dorn, M., M. B. Silva; L. S. Buriol & L. C. Lamb.** 2014. Three-dimensional protein structure prediction: Methods and computational strategies. *Computational Biology and Chemistry*, 53: 251–276. <https://doi.org/10.1016/j.compbiolchem.2014.10.001>
- Dutta, B., A. Banerjee; P. Chakraborty & R. Bandopadhyay.** 2018. In silico studies on bacterial xylanase enzyme: Structural and functional insight. *Journal, Genetic Engineering & Biotechnology*, 16(2): 749–756. <https://doi.org/10.1016/j.jgeb.2018.05.003>
- Flores-Castañón, N., S. Sarkar & A. Banerjee.** 2022. Structural, functional, and molecular docking analyses of microbial cutinase enzymes against polyurethane monomers. *Journal of Hazardous Materials Letters*, 3(100063): 100063. <https://doi.org/10.1016/j.hazl.2022.100063>
- Feng, S.; W. Zhang; J. Yang; Y. Yang & H. Shen.** 2020. Topology prediction improvement of α -helical transmembrane proteins through helix-tail modeling and multiscale deep learning fusion. *Journal of Molecular Biology*, 432(4): 1279–1296. <https://doi.org/10.1016/j.jmb.2019.12.007>
- Goh, N., M. F. Mohamad-Razif; Y. Yap; C. L. Ng & S. Fung.** 2022. In silico analysis and characterization of medicinal mushroom cystathionine beta-synthase as an angiotensin converting enzyme (ACE) inhibitory protein. *Computational Biology*

- and Chemistry, 96(107620): 107620. <https://doi.org/10.1016/j.compbiolchem.2021.107620>
- Gottschalk, G.; K. Knox & A. Roy.** 2021. ACE2: At the crossroad of COVID-19 and lung cancer. *Gene Reports*, 23(101007): 101007. <https://doi.org/10.1016/j.genrep.2021.101077>
- Grasso, E. J.; A. E. Sottile & C. E. Coronel.** 2016. Structural Prediction and In Silico Physicochemical Characterization for Mouse Caltrin I and Bovine Caltrin Proteins, 10: 225-236. doi: 10.4137/BBI.S38191
- Hoang, H. N.; T. A. Hill; G. Ruiz-Gómez; F. Diness; J. M. Mason; C. Wu; G. Abbenante; N. E. Shepherd & D. P. Fairlie.** 2019. Twists or turns: stabilising alpha vs. beta turns in tetrapeptides. *Chemical Science (Royal Society of Chemistry: 2010)*, 10(45): 10595–10600. <https://doi.org/10.1039/c9sc04153b>
- Huang, B.; L. Kong; C. Wang; F. Ju; Q. Zhang; J. Zhu; T. Gong; H. Zhang; C. Yu; W. Zheng & D. B.** 2023. Protein structure prediction: Challenges, advances, and the shift of research paradigms. *Genomics, Proteomics & Bioinformatics*. <https://doi.org/10.1016/j.gpb.2022.11.014>
- Huang, C.; Y. Wang; X. Li; L. Ren; J. Zhao; Y. Hu; L. Zhang; G. Fan; J. Xu; X. Gu; Z. Cheng; T. Yu; J. Xia; Y. Wei; W. Wu; X. Xie; W. Yin; H. Li; M. Liu; Y. Xiao; H. Gao; L. Guo; J. Xie; G. Wang; R. Jiang; Z. Gao; Q. Jin; J. Wang & B. Cao.** 2020. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet*, 395(10223): 497-506. doi: F10.1016/S0140-6736(20)30183-5.
- Kumar, A.; Agarwal, P.; Shivangi & L. S. Meena.** 2021. Structural and functional investigation of mycobacterial Hfx protein and its mutational hotspots annotation by in silico approach. *Gene Reports*, 25(101418): 101418. <https://doi.org/10.1016/j.genrep.2021.101418>
- Kumar, V.; N. Sushma-Sri; N. Tripathi; V. K. Sharma; P. V. Bharatam; P. Garg & S. Singh.** 2020. Structural exploration of glutamine synthetase from *Leishmania donovani*: Insights from in silico and in vitro analysis. *International Journal of Biological Macromolecules*, 146: 860–874. <https://doi.org/10.1016/j.ijbiomac.2019.09.209>
- Lenin, K. L. D.; A. Raveendran & S. P. Antony.** 2023. Identification and in silico structural and functional analysis of a histone H2A-derived antimicrobial peptide from teleost fishes. *Animal Gene*, 28(200148): 200148. <https://doi.org/10.1016/j.angen.2023.200148>
- Liu, Y.; G. Hu; Y. Wang; W. Ren; X. Zhao; F. Ji; Y. Zhu; F. Feng; M. Gong; X. Ju; Y. Zhu; X. Cai; J. Lan; J. Guo; M. Xie; L. Dong; Z. Zhu; J. Na; J. Wu; ... & Q. Ding.** 2021. Functional and genetic analysis of viral receptor ACE2 orthologs reveals a broad potential host range of SARS-CoV-2. *Proceedings of the National Academy of Sciences of the United States of America*, 118(12). <https://doi.org/10.1073/pnas.2025373118>
- Liu, Z., X. Xiao; X. Wei; J. Li; J. Yang; H. Tan; J. Zhu; Q. Zhang; J. Wu & L. Liu.** 2020. Composition and divergence of coronavirus spike proteins and host ACE2 receptors predict potential intermediate hosts of SARS-CoV-2. *Journal of Medical Virology*, 92(6): 595–601. <https://doi.org/10.1002/jmv.25726>
- Lo, C.S.; Y. Shi; S. Y. Chang; S. Abdo; I. Chenier; J.G. Filep; J.R. Ingelfinger; S.L. Zhang & D.J.S. Chan.** 2015. Overexpression of heterogeneous nuclear ribonucleoprotein F stimulates renal Ace-2 gene expression and prevents TGF- β 1-induced kidney injury in a mouse model of diabetes. *Diabetologia*, 58: 2443–2454. <https://doi.org/10.1007/s00125-015-3700-y>
- Lorraine, O. E.; R. N. Z. R. A. Rahman; T. Joo-Shun; A. B. Salleh & S. N. Oslan.** 2023. In silico structural exploration of serine protease from a CTG-clade yeast *Meyerozyma guilliermondii* strain SO. *Analytical Biochemistry*, 668(115092), 115092. <https://doi.org/10.1016/j.ab.2023.115092>
- Luan, J., Y. Lu; X. Jin & L. Zhang.** 2020. Spike protein recognition of mammalian ACE2 predicts the host range and an optimized ACE2 for SARS-CoV-2 infection. *Biochemical and Biophysical Research Communications*, 526(1): 165–169. <https://doi.org/10.1016/j.bbrc.2020.03.047>
- Mehmood S.; R. Dad; A. Ahmad; M. I. Ullah; P. John; A. Ali; C. A. Hubner; A. Mohyuddin & M.J. Hassan.** 2020. Structural and functional annotation of PR/SET Domain (PRDM) protein family: In-silico study elaborating role of PRDM12 mutation in congenital insensitivity to pain. *Computational Biology and Chemistry*, 89(107382): 107382. <https://doi.org/10.1016/j.compbiolchem.2020.107382>
- Mohanty, R.; M. Manoswini; A. K. Dhal & N. Ganguly.** 2022. In silico analysis of a novel protein in CAR T- cell therapy for the treatment of hematologic cancer through molecular modelling, docking, and dynamics approach. *Computers in Biology and Medicine*, 151(106285), 106285. <https://doi.org/10.1016/j.compbiomed.2022.106285>

- Mukherjee, A.; P. Dandapat; M. Z. Haque; S. Mandal; P. S. Jana; S. Samanta; S. Pal; A. K. Das; P. K. Nanda; S. Bandyopadhyay & C. Guha.** 2023. Computational analysis of hypothetical proteins from *Mycobacterium orygis* identifies proteins with therapeutic and diagnostic potentials. *Animal Gene*, 29(200154), 200154. <https://doi.org/10.1016/j.angen.2023.200154>
- Nguyen, M. N., K. P. Tan & M. S. Madhusudhan.** 2011. CLICK—topology-independent comparison of biomolecular 3D structures. *Nucleic Acids Research*, 39(suppl_2): W24–W28. <https://doi.org/10.1093/nar/gkr393>
- Nolsoe, H. & I. Undeland.** 2009. The Acid and Alkaline Solubilization Process for the Isolation of Muscle Proteins: State of the Art. *Food Bioprocess Technology* 2: 1-27 <https://link.springer.com/article/10.1007/s11947-008-0088-4>
- Pollet, L.; L. Lambourne & Y. Xia.** 2022. Structural determinants of yeast protein-protein interaction interface evolution at the residue level. *Journal of Molecular Biology*, 434(19), 167750. <https://doi.org/10.1016/j.jmb.2022.167750>
- Premachandran, K., & T. S. Srinivasan.** 2022. In silico modelling and interactive profiling of BPH resistance NBS-LRR proteins with salivary specific proteins of rice planthoppers. *Gene Reports*, 28(101648): 101648. <https://doi.org/10.1016/j.genrep.2022.101648>
- Rout, S. & R. K. Mahapatra.** 2019. In silico analysis of *Plasmodium falciparum* CDPK5 protein through molecular modeling, docking and dynamics. *Journal of Theoretical Biology*, 461: 254–267. <https://doi.org/10.1016/j.jtbi.2018.10.045>
- Samavati, L. & B. D. Uhal.** 2020. ACE2, much more than just a receptor for SARS-COV-2. *Frontiers in cellular and infection microbiology*, 10. <https://doi.org/10.3389/fcimb.2020.00317>
- Sanchez-Pulido, L. & C. P. Ponting.** 2021. Extending the horizon of homology detection with coevolution-based structure prediction. *Journal of Molecular Biology*, 433(20): 167106. <https://doi.org/10.1016/j.jmb.2021.167106>
- Sang, P.; M. Zhang; Y. Shi; C. Li; S. Abdulkadir; Q. Li; H. Ji & J. Cai.** 2019. Inhibition of β -catenin/B cell lymphoma 9 protein–protein interaction using α -helix–mimicking sulfono- γ -AApeptide inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*, 116(22): 10757–10762. <https://doi.org/10.1073/pnas.1819663116>
- Satyanarayana, S. D. V.; M. S. R. Krishna; P. Pavan-Kumar & S. Jeerreddy.** 2018. In silico structural homology modeling of nif A protein of rhizobial strains in selective legume plants. *Journal, Genetic Engineering & Biotechnology*, 16(2): 731–737. <https://doi.org/10.1016/j.jgeb.2018.06.006>
- Sharma, L.; B. Maurya; S. Pandey; K. K. Rai & S. Pandey-Rai.** 2023. Structural and functional prediction of WsMYB34 transcription factor in *Withania somnifera* (L.) Dunal by deciphering its role in NaCl-induced secondary metabolism. *Industrial Crops and Products*, 206(117682): 117682. <https://doi.org/10.1016/j.indcrop.2023.117682>
- Shrivastava, N.; A. Verma & P. K. Dash.** 2020. Identification of functional epitopes of structural proteins and in-silico designing of dual acting multi-epitope anti-tick vaccine against emerging Crimean-Congo hemorrhagic fever virus. *European Journal of Pharmaceutical Sciences: Official Journal of the European Federation for Pharmaceutical Sciences*, 151(105396): 105396. <https://doi.org/10.1016/j.ejps.2020.105396>
- Singh, S.; S. Biswas; A. Srivastava; Y. Mishra & T. P. Chaturvedi.** 2021. In silico characterization and structural modeling of a homeobox protein MSX1 from *Homo sapiens*. *Informatics in Medicine Unlocked*, 22(100497): 100497. <https://doi.org/10.1016/j.imu.2020.100497>
- Singh, V.; P. Dhankhar; V. Dalal; S. Tomar & P. Kumar.** 2022. In-silico functional and structural annotation of hypothetical protein from *Klebsiella pneumoniae*: A potential drug target. *Journal of Molecular Graphics & Modelling*, 116(108262): 108262. <https://doi.org/10.1016/j.jmgm.2022.108262>
- Waryah, Y. M.; F. F. Khidri; R. Nigar; T. Devrajani; A. R. Rajput; A. M. Waryah & I. D. Ujjan.** 2023. Impact of ACE2 gene variations on COVID-19 pathogenicity in Pakistani patients. *Saudi Journal of Biological Sciences*, 30(10): 103813. <https://doi.org/10.1016/j.sjbs.2023.103813>
- Werner, T.; M. B. Morris; S. Dastmalchi & W. B. Church.** 2012. Structural modelling and dynamics of proteins for insights into drug interactions. *Advanced Drug Delivery Reviews*, 64(4), 323–343. <https://doi.org/10.1016/j.addr.2011.11.011>
- Wu, L.; Q. Chen; K. Liu; J. Wang; P. Han; Y. Zhang; Y. Hu; Y. Meng; X. Pan; C. Qiao; S. Tian; P. Du; H. Song; W. Shi; J. Qi; H. Wang; J. Yan; G. F. Gao & Q. Wang.** 2020. Broad host range of SARS-CoV-2 and the molecular basis for SARS-CoV-2 binding to cat ACE2. *Cell Discovery*, 6(1): 1–12. <https://doi.org/10.1038/s41421-020-00210-9>

- Xu, J.; Q.Li; J. Zhang; X. Li & T. Sun.** 2021. In silico structural and functional analysis of cold shock proteins in *Pseudomonas fluorescens* PF08 from marine fish. *Journal of Food Protection*, 84(8), 1446–1454. <https://doi.org/10.4315/jfp-21-044>
- Xu, S. & M. A. Kennedy.** 2020. Structural dynamics of pentapeptide repeat proteins. *Proteins*, 88(11): 1493–1512. <https://doi.org/10.1002/prot.25969>
- Yasin, N.; H. S. Laxmanappa; U. M. Muddapur; J. Cheruvathur; S. M. U. Prakash & H. V. Thulasiram.** 2020. Structural, molecular, functional and immunological characterization of *Wuchereria bancrofti-galectin*. *International Journal of Biological Macromolecules*, 150: 206–217. <https://doi.org/10.1016/j.ijbiomac.2020.02.024>
- Zhong, Y.; Q. Tang; D. P. Miller; E. Zurek; R. Liu; Z. Lu & B. Gong.** 2020. Major factors for the persistent folding of hybrid α , β , γ -hybrid peptides into hairpins. *Frontiers in chemistry*, 8. <https://doi.org/10.3389/fchem.2020.530083>