



Universidad Nacional Autónoma de México

Facultad de Estudios Superiores Cuautitlán

Identificación de receptores celulares del PEDV
responsables de la infección mediante técnicas *in silico*.

T E S I S

QUE PARA OBTENER EL TÍTULO DE:
LICENCIADO EN QUÍMICA INDUSTRIAL

P R E S E N T A

Cárdenas Granados Luis Alfonso

ASESOR: M. en C. Manuel Alejandro Hernández Serda

COASESOR: Dr. Víctor Hugo Vázquez Valadez

Cuautitlán Izcalli, Estado de México, 2023



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



UNIVERSIDAD NACIONAL
AUTÓNOMA DE
MÉXICO

FACULTAD DE ESTUDIOS SUPERIORES CUAUTITLÁN
SECRETARÍA GENERAL
DEPARTAMENTO DE TITULACIÓN



ASUNTO: VOTO APROBATORIO

DR. DAVID QUINTANAR GUERRERO
DIRECTOR DE LA FES CUAUTITLÁN
PRESENTE

ATN: DRA. MARIA DEL CARMEN VALDERRAMA BRAVO
Jefa del Departamento de Titulación
de la FES Cuautitlán.

Con base en el Reglamento General de Exámenes, y la Dirección de la Facultad, nos permitimos comunicar a usted que revisamos el: **Trabajo de tesis y examen profesional.**

Identificación de receptores celulares del PEDV responsables de la infección mediante técnicas in silico.

Que presenta el pasante: **Luis Alfonso Cárdenas Granados**
Con número de cuenta: **419061207** para obtener el título de: **Químico Industrial**

Considerando que dicho trabajo reúne los requisitos necesarios para ser discutido en el **EXAMEN PROFESIONAL** correspondiente, otorgamos nuestro **VOTO APROBATORIO.**

ATENTAMENTE
"POR MI RAZA HABLARÁ EL ESPÍRITU"
Cuautitlán Izcalli, Méx. a 12 de Mayo de 2023.

PROFESORES QUE INTEGRAN EL JURADO

	NOMBRE	FIRMA
PRESIDENTE	Dra. Susana Elisa Mendoza Elvira	
VOCAL	Dra. Sandy María Pacheco Ortín	
SECRETARIO	M. en C. Manuel Alejandro Hernández Serda	
1er. SUPLENTE	M.C. Paola Edith Briseño Lugo	
2do. SUPLENTE	Q. F.B. David Ladislao Sánchez	

NOTA: Los sinodales suplentes están obligados a presentarse el día y hora del Examen Profesional.
En caso de que algún miembro del jurado no pueda asistir al examen profesional deberá dar aviso por anticipado al departamento.

“Somos como mariposas que vuelan durante un día pensando que lo harán por siempre”.

-Carl Sagan

“Si una idea no parece absurda de entrada, pocas esperanzas hay en ella”.

-Albert Einstein

“Podemos ser todos poéticos y perder la cabeza juntos”.

-Ellie

“Believe”.

-Ted Lasso



Dedicatorias

A mi abuela Josefina Granados Malagón por todo su cariño y apoyo, por siempre cuidar de mí y a que a pesar de todo siempre has estado presente, te estaré agradecido el resto de mi vida y siempre te llevaré en el corazón, gracias por todo "Abu".

A mi tío Luis, a pesar de que nos faltó tiempo juntos, te quiero y siempre estarás conmigo.

Agradecimientos

A mi madre Magnolia Granados Malagón por siempre confiar en mí y apoyarme en mis decisiones, por tantas cosas que hemos pasado juntos, por acompañarme, agradezco tu tiempo. Te amo

A mi padre Alfonso Cárdenas Vallejo, gracias por cambiar y por todo tu apoyo y cariño. Te quiero

A mi hermana Zoé, gracias por ser como eres, sé que lograras grandes cosas. Te amo.

A mi tía Edith Granados y mi tío Carlos Brizuela por ser como unos padres para mí, les agradezco todo lo que han hecho por mí y les debo mucho, sepan que los quiero.

A mis tías Noemi Granados, Cristina Ruiz y Delia Granados, por su cariño

Más que mis amigos a mis hermanos que llevan tantos años conmigo, gracias, Arath, Ángel (Dnu), David, Kean (Niño), Miguel (Kiwi), Santi, Sagiv y Hugo gracias por todo el tiempo en grupo y lo que hemos vivido.

A Salma por acompañarme en momentos tan difíciles. Te amo

A mis Asesores Manuel Alejandro Hernández Serda y Víctor Hugo Vázquez Valadez donde además de ser excelentes en su trabajo también los son como personas y donde creo haber encontrado unos buenos amigos.

Al Dr. Enrique Ángeles Anguiano por abrirme las puertas de su laboratorio así como permitirme ser parte de su grupo, por todos sus consejos y apoyo.

Al Dr. Levon Abrahamyan de la Universidad de Montreal en Canadá por su gran guía y cooperación para realizar este proyecto.

A mis amigos durante esta etapa Universitaria Diego, Ulises, Daira, Karina, Beto y Celeste gracias por todos esos buenos momentos.

A todos en el LQM (Rodolfo, Sabrina, Elizabeth, Omar) que poco a poco está creciendo, pero en especial al Dr. Pablo Arturo por ser una persona con tanto conocimiento y a la vez tan humilde, se te quiere Pablin, al M. C. Aldo Yoshio por sus buenos consejos y a Pablo Aguirre por siempre mostrarse dispuesto ayudar.

A los Sioux de V8 por darme los mejores momentos de mi vida.

A todos los miembros del Jurado por tomarse el tiempo de realizar tan valiosas correcciones para este trabajo.

El presente trabajo se realizó dentro de la Facultad de Estudios Superiores Cuautitlán Campo 1 en el LQM (Laboratorio de Química Medicinal), con los apoyos PAPIIT IN206023, PAPIIT IN2020, PIAPI C12201, LANCAD-UNAM-DGTIC-034, así como la secretaría de relaciones exteriores (AMEXCID), en el marco del grupo de trabajo en educación, cultura, investigación e innovación 2021-2023 México-Quebec.

Índice

1. RESUMEN	1
2. INTRODUCCIÓN	2
2.1. Enfermedades en los cerdos.....	2
3. VIRUS.....	3
3.1. Estructura Viral	4
3.2. Clasificación y tipos de virus	5
3.3. Ciclo viral.....	6
3.4. Adhesión y entrada.....	7
3.5. Receptores virales	8
3.6. Penetración a la célula	9
3.8. Patogénesis vírica.....	10
4. VIRUS DE DIARREA PORCINA EPIDEMICA (PEDV)	11
5. QUÍMICA COMPUTACIONAL	15
5.1. Dinámica molecular.....	15
5.2. Campo de Fuerza (Force Field)	16
5.3. Acoplamiento molecular	17
5.3.1. Funciones de registro.....	18
6. OBJETIVOS.....	19
7. METODOLOGÍA	20
7.1. Búsqueda de receptores.....	21
7.1.1. CD4.....	21
7.1.2. CD8.....	22
7.1.3. CD13 (Aminopeptidasa).....	22
7.1.4. CD26 (DPP4).....	22
7.1.5. CD46.....	23
7.1.6. CD54 (ICAM-1)	23
7.1.7. CD66 (CEACAM-1).....	23
7.1.8. CD209.....	24
7.1.9. CD249 (ENPEP)	24
7.1.10. ECA-2	24
7.1.11. Ocludina	25
7.1.12. Tripsina	25
7.2. Obtención de estructuras	27
7.3. Spike	29
7.4. Receptores	30
8. RESULTADOS Y ANALISIS	34
9. CONCLUSIONES.....	43
10. PERSPECTIVAS.....	44

11. REFERENCIAS	45
11. ANEXOS	49

Índice de figuras

1. Comparativa de tamaños de distintos virus	3
2. Envoltura viral.....	5
3. Clasificación de Baltimore	6
4. Membrana celular	7
5. Receptores virales	8
6. Entrada a la célula.....	9
7. Peptido de fusión	10
8. PEDV	11
9. Genoma viral.....	13
10. Dominios PEDV.....	13
11. Conformaciones Spike	14
12. Metodología general.....	20
13. Softwares utilizados.....	21
14. Estructura secundaria conformaciones Spike	30
15. Complejo Spike-Receptor	33
16. Ejemplo RMSD	34
17. Ejemplo RMSF	34
18. Resultados docking conformacion cerrada	36
19. Resultados docking conformacion abierta	37
20. Ejemplo tripsina	38
21. Distribución de interacciones para la proteína CD13 (ANPEP)	39
22. Distribución de interacciones para la proteína Tripsina.....	40
23. Distribución de interacciones para la proteína CD4.....	41
24. Aminoácidos en la región CTD.....	42

Índice de tablas

1. Residuos por receptor	26
2. Obtención de estructuras	31
3. Protocolo	32
4. Energía de acoplamiento.....	35

1. RESUMEN

El sector porcino se enfrenta a una importante amenaza derivada del virus PEDV, que se caracteriza por su elevada mortalidad y su capacidad de propagación entre granjas. La propagación de esta enfermedad ha tenido un impacto significativo en la producción porcina, uno de los alimentos más consumidos en todo el mundo, con importantes pérdidas económicas. México, en particular, es uno de los principales consumidores de carne de cerdo en el continente americano, según la Organización de las Naciones Unidas para la agricultura y la Alimentación o FAO (*Food and Agriculture Organization of the United Nations*), lo que le hace especialmente vulnerable a los efectos económicos del PEDV.

Si bien existen vacunas preventivas para la infección ocasionada por PEDV, aun no se sabe con seguridad cual o cuales son los receptores que permiten la entrada de dicho virus, por lo que este trabajo pretende mediante métodos computacionales de acoplamiento molecular proponer potenciales receptores virales para la entrada del PEDV.

2.INTRODUCCIÓN

La enfermedad se define como cualquier alteración perjudicial y alarmante que cause desviaciones o desequilibrios en el funcionamiento del organismo. Normalmente la forma de asociar enfermedades con aquello que la ocasiona suelen ser los síntomas, del estudio de las enfermedades se encarga la patología, donde abarca varios aspectos de una enfermedad, desde aquello que la causa (etiología), conocer los mecanismos de dicha enfermedad (patogénesis), así como cambios presentes en la funcionalidad de los órganos (cambios morfológicos).(Scarpelli, 2023)

2.1. Enfermedades en los cerdos

Generalmente las enfermedades de mayor importancia dentro de la industria ganadera han sido enfermedades de orígenes infecciosos o parasitarios, donde los principales causantes de enfermedades de este tipo son bacterias, parásitos o virus. Algunas de las enfermedades que perjudican a los cerdos son: trichostrongilosis, ascariasis, cisticercosis, colera, erisipela, neumonía, salmonelosis, brucelosis, virus de la diarrea epidémica porcina.(Organización de las Naciones Unidas para la Agricultura y la Alimentación (FAO), 2010)

Las enfermedades infecciosas representan una grave amenaza para la industria porcina, ya que tienen una alta capacidad de propagación entre granjas. Por lo que se deben implementar controles rigurosos de seguridad e higiene, así como estar preparados ante el posible surgimiento de nuevas infecciones. De esta manera, se podrán reducir los riesgos y proteger la salud de los animales, así como garantizar la seguridad de los productos que se obtienen de ellos.

Existe una gran cantidad de receptores celulares, cada uno con particularidades, sin embargo para la entrada del PEDV es necesario un receptor que se encuentre sobre la superficie celular, los receptores de superficie celular generalmente cuentan con tres partes principales conocidos como dominios, estos dominios son, desde la parte más externa el dominio extracelular el cual es el responsable de las uniones con ligandos, seguido de un dominio altamente hidrofóbico que se encuentra atravesando la

membrana celular, por último el dominio intracelular aquel que como su nombre lo indica se encuentra dentro del citoplasma celular.(Khan Academy, n.d.)

3. VIRUS

Los virus inicialmente fueron catalogados como aquellos agentes capaces de causar enfermedad, sin embargo desde ese entonces la forma de definir a los virus ha cambiado. Con el pasar de los años, y gracias a avances científicos y tecnológicos se han desarrollado mejores técnicas para su clasificación. Los virus en su mayoría suelen tener un tamaño pequeño, esto va desde los 20 nm para los virus más pequeños como el caso de los Picornavirus, o en el caso del virus de influenza o el VIH (Virus de Inmunodeficiencia Humana), que se encuentran cerca de los 100 nm, pero también existen virus de mucho mayor tamaño como se puede observar en la figura 1. En comparación, una célula humana promedio tiene un diámetro de 10-30 micrómetros.(Louten & Reynolds, s/f)

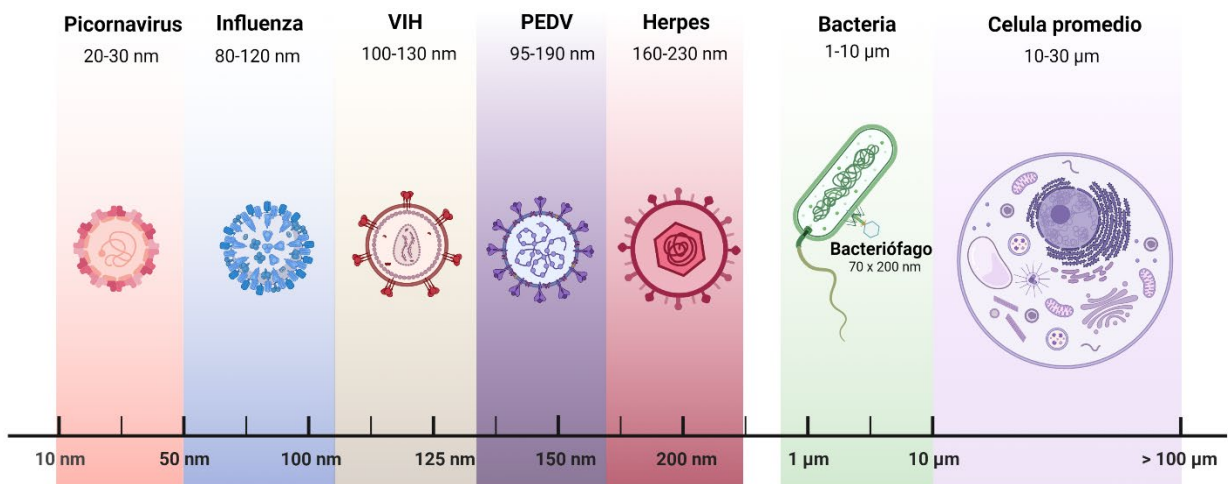


Figura 1. Comparativa de tamaños de distintos virus con otras estructuras. (Realizada con BioRender.com)

La ausencia metabólica es algo que caracteriza a los virus para su definición desde 1957 ya que el virión por sí solo no tiene la capacidad de obtener energía a través de reacción bioquímicas.

Gracias a la deficiencia metabólica, se dice que un virus no tiene vida, ya que está obligado a parasitar una célula con vida, y de esta forma crear nuevas copias infecciosas del mismo conocido como virión, un virión es aquella partícula con la capacidad de infectar.(Koonin, Dolja, et al., 2021)

A pesar de no ser clasificados como seres vivos tener, los virus comparten varias características con los organismos vivos, ya que de hecho están compuestos de sustancias muy similares, como puede ser el material genético o los aminoácidos que los componen.

Dichas partículas contienen material genético en su interior, el cual puede estar compuesto por ARN o DNA pero nunca por ambos. Por otro lado, todas las células con vida ya sean de animales, plantas o bacterianas tienen una doble cadena de ADN (dsDNA) como material genético. (Louten & Reynolds, s/f)

3.1. Estructura Viral

En el caso de los virus más simples el virión está constituido por dos partes principales, donde se tiene una parte de ácido nucleico la cual está rodeada por una cobertura proteica, la cual adquiere el nombre de cápside, ambas partes juntas conforman la nucleocápside. En virus más complejos la cápside puede estar rodeada por una envoltura lipoproteica.(Fenner et al., 1987)

Debido a las características membranales de la envoltura, los virus con envoltura deben siempre de permanecer húmedos, por lo general este tipo de virus se transmite por fluidos, cambios en temperatura o pH del medio pueden llevar a la inactivación de dicha partícula viral.(Elsevier Connect, 2021)

La cápside viral está formada por proteínas llamadas capsómeros y según su simetría se pueden clasificar en distintos tipos:

- Cápside helicoidal (presente en distintos virus).

- Cápside icosaédrica (presente típicamente en virus animales que a su vez dependiendo de los monómeros que lo componen se pueden dividir en pentones o hexones).
- Cápside compleja (presenta una cabeza o nucleocápside y una cola).

Como se muestra en la figura 2, una característica muy importante en algunos de estos es la envoltura donde algunos pueden presentarla y otros no, algo que siempre se encuentra presente es la presencia de proteínas que ayudan a la entrada, también llamadas proteínas de adhesión viral.

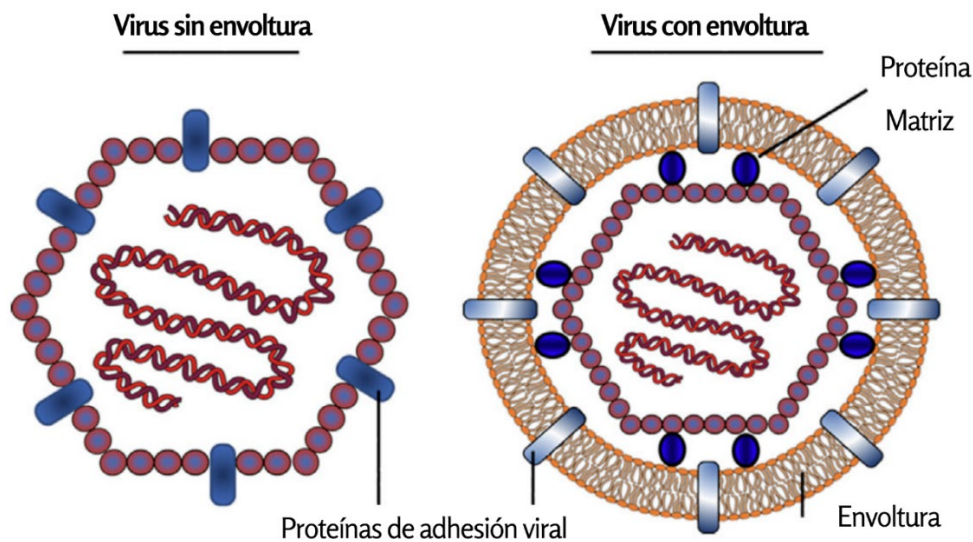


Figura 2. Ejemplo de envoltura viral y la comparación entre virus con envoltura y sin envoltura.

Obtenida de (Louten & Reynolds, n.d.)

3.2. Clasificación y tipos de virus

En 1970 se desarrolló un método de clasificación basándose en el tipo de ácido nucleico y el tipo de replicación del virus, a este método se le denominó sistema de clasificación Baltimore (*Baltimore classification system*). Este método de clasificación separa a los virus en 7 clases distintas las cuales se pueden observar en la figura 3 y son (Koonin, Krupovic, et al., 2021):

- Clase I: ds DNA virus
- Clase II: ssDNA virus
- Clase III: dsRNA virus
- Clase IV: sentido positivo ssRNA virus
- Clase V: sentido negativo ssRNA virus
- Clase VI: RNA virus de transcripción inversa.
- Clase VII: DNA virus de transcripción inversa.

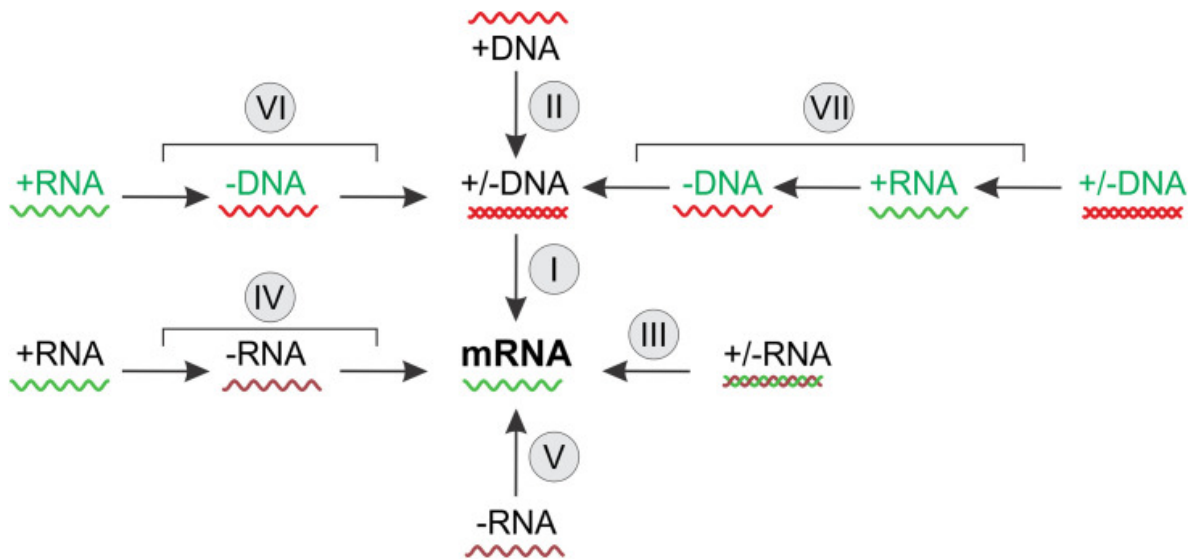


Figura 3. Clasificación de Baltimore. Obtenida de (Koonin, Krupovic, et al., 2021)

Existen otros factores para la clasificación de virus como puede ser el tamaño, la estructura de la cápside, propiedades físicas, la enfermedad que causa, entre otros. Es por eso por lo que el Comité Internacional en Taxonomía de Virus ICTV (*International Comitte on Taxonomy of Viruses*) decidió que los virus serian clasificados únicamente usando Orden, Familia, Genero y Especie. Se han categorizado 7 órdenes de virus conteniendo un total de 103 familias (Louten & Reynolds, s/f)

3.3. Ciclo viral

Los pasos tomados por el virus para la producción de nuevas partículas virales se han dividido en distintos pasos, con el fin de lograr un mayor entendimiento del ciclo viral. El cual está comprendido por la entrada de la partícula infecciosa, pasando a la

producción de mRNA (RNA mensajero) y su traducción por los organelos celulares del hospedero, la replicación del genoma viral y por último el ensamblaje y liberación de copias virales. Aquellos pasos en conjunto llevan por nombre reproducción viral.

La reproducción viral hace uso de las grandes funciones celulares, entre ellas la maquinaria de traducción viral del mRNA, funciones metabólicas entre otras. (Flint et al., 2020)

3.4. Adhesión y entrada

Debido a la necesidad de los virus de parasitar células los virus deben lograr de alguna forma entrar a dicha célula, la mayoría de los virus que infectan animales producen partículas muy grandes para cruzar a través del plasma membranar. Principalmente la infección viral comienza con la exposición del hospedero a un ambiente infeccioso. Sobre las membranas existe una gran presencia de proteínas, como se muestran en la figura 4 dentro de ellas existen glucoproteínas que atraviesan la bicapa lipídica de la membrana. (Flint et al., 2020; Louten & Reynolds, s/f)

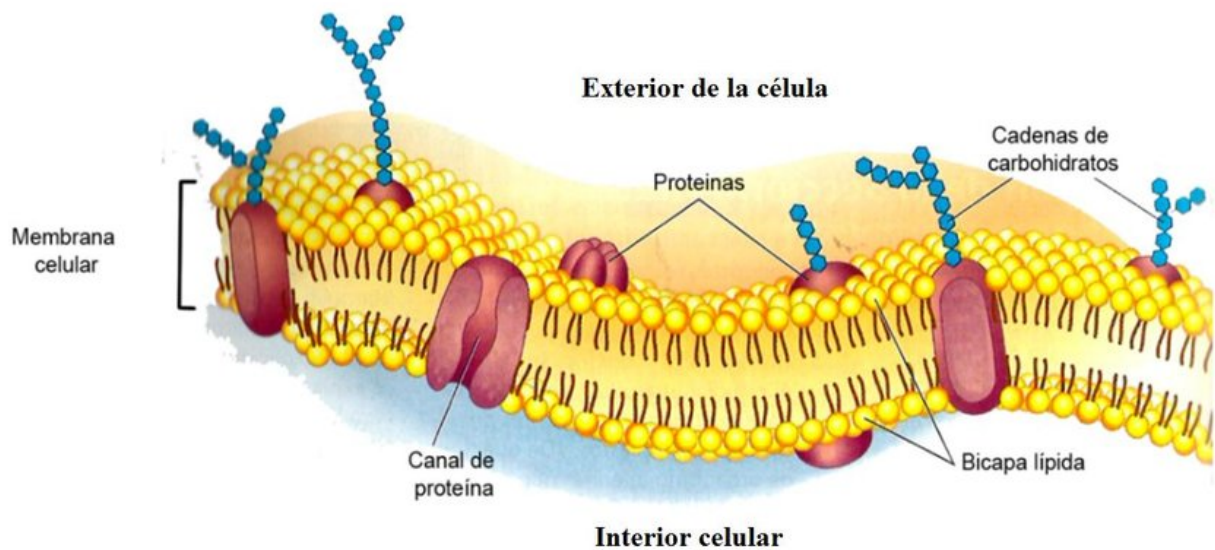


Figura 4. Esquematización de la estructura general de una membrana celular y algunos componentes comunes. (Ortega-Quintana, 2018)

3.5. Receptores virales

Aquellas proteínas que se encuentran sobre dicha membrana adquieren el nombre de proteínas integrales o *Integral proteins*, estas proteínas pueden tener distintas funciones, una de ellas ser receptores para sustancias externas.

Los virus pueden tener distintas zonas de unión en sus superficies, generalmente los virus con envoltura poseen glucoproteínas incorporadas en las cuales se encuentra dicha zona de unión. Por otro lado, aquellos virus sin envoltura carecen de estas estructuras que facilitan la unión, por lo que proteínas de la cápside funcionan como las zonas de unión. (Flint et al., 2020)

Se ha encontrado que dos diferentes virus pueden compartir un receptor celular, como es el caso del receptor para Coxsackievirus y Adenovirus donde el comparten el receptor CAR (Receptor constitutivo de androstano). Sin embargo, en el caso de los adenovirus esta unión no es suficiente para la infección y es necesaria la participación de un correceptor. (Flint et al., 2020)

Dentro de los virus que requieren un correceptor para la unión e infección se encuentra el VIH el cual en primera instancia se une al receptor CD4, sin embargo requiere de uno de dos correceptores el cual puede ser CXCR4 o CCR5, como se puede observar en la figura 5. (Louten & Reynolds, s/f)

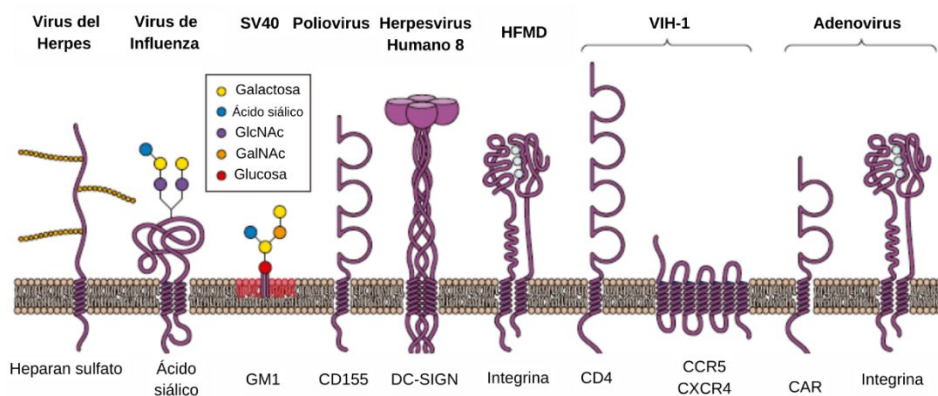


Figura 5. Ejemplos de receptores virales reportados para la entrada de distintos virus. (Obtenida de lecturas de virología Flint et al., 2020)

3.6. Penetración a la célula

Seguido de la etapa de adhesión existen distintas rutas de ingreso a la célula, y esto puede variar dependiendo de cuál es la sustancia externa que pretende cruzar la membrana, entre los mecanismos de entrada o penetración principales se encuentra la fagocitosis, y la endocitosis; donde la endocitosis es la principal ruta utilizada por los virus, en específico el tipo de endocitosis mediada por receptor. Entre estos tipos existen los mediados por clatrina, caveolina y aquellos que entran por medio de fusión en la figura 6 se esquematizan estos procesos. Este último es usado únicamente por virus con envoltura y es normalmente la proteína que se une a un receptor la misma que se encarga de la fusión. (Flint et al., 2020; Louten & Reynolds, s/f)

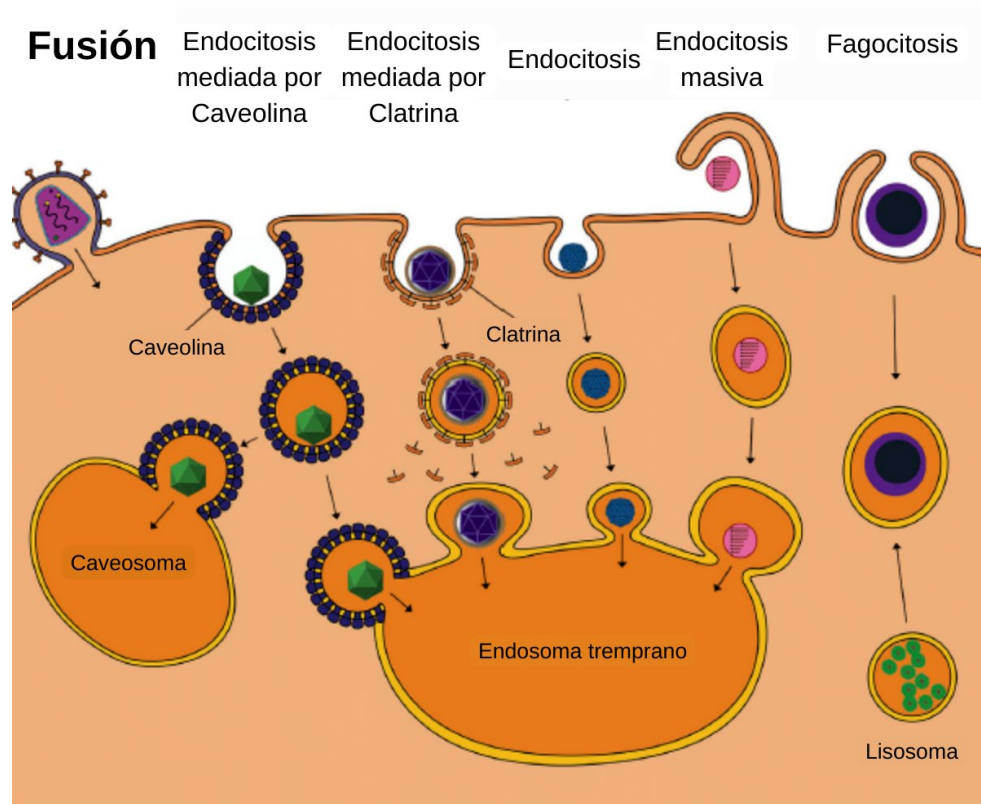


Figura 6. Esquematación de entrada de partículas externas al interior de la célula. (Obtenida de Louten & Reynolds, n.d.)

Es precisamente el proceso de fusión membranar el encargado de llevar el ácido nucleico dentro de la célula, dentro de las proteínas de fusión de clase I es clave una escisión para que se puede liberar el péptido de fusión el cual debe ser una secuencia altamente hidrofóbica, para que esta pueda unirse a la membrana. Un punto de suma importancia es que dicho péptido de fusión debe encontrarse protegido hasta su llegada a la célula. Dependiendo de la ubicación del péptido de fusión, la proteína viral unida al receptor puede presentar cambios conformacionales antes o después de la unión del péptido a la membrana, como en la figura 7, donde se muestra el cambio conformacional en el caso del virus del VIH.(Flint et al., 2020)

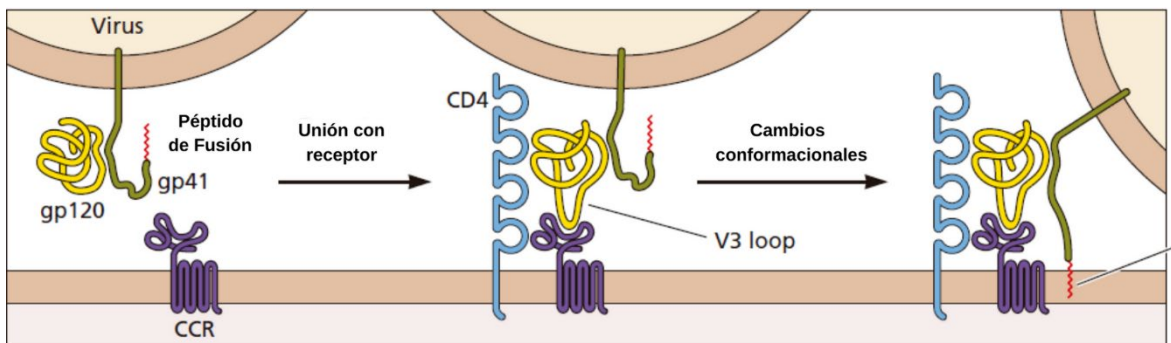


Figura 7. Ejemplo de exposición del péptido de fusión en el caso de la adhesión del VIH (Obtenida de lecturas de virología Flint et al., 2020)

3.8. Patogénesis vírica

Existen distintas formas en las que una infección vírica puede desarrollarse, existe la infección lítica la cual debido al daño dado por la replicación termina en la destrucción celular. Una infección persistente, en donde no se llega a la muerte celular, se vuelven infecciones persistentes crónicas, latentes, recurrentes y transformantes derivando en tumores. Así como una infección abortiva donde debido a las altas mutaciones víricas no se puede completar la replicación viral para la creación de nuevas partículas infectantes. (Galán-Sánchez et al., 2014)

4. VIRUS DE DIARREA PORCINA EPIDEMICA (PEDV)

El PEDV (Porcine Epidemic Diarrhea Virus), es un virus de cadena simple RNA del género *Alphacoronavirus* perteneciente a la familia *Coronaviridae* del orden de los *Nidovirales*. Se trata de un virus con envoltura, la cápside que presenta es helicoidal. Como se observa en la figura 8 donde del lado izquierdo se muestra el virus obtenido de muestras fecales mediante microscopía electrónica, al lado derecho su esquematización. (Z. Li et al., 2020)

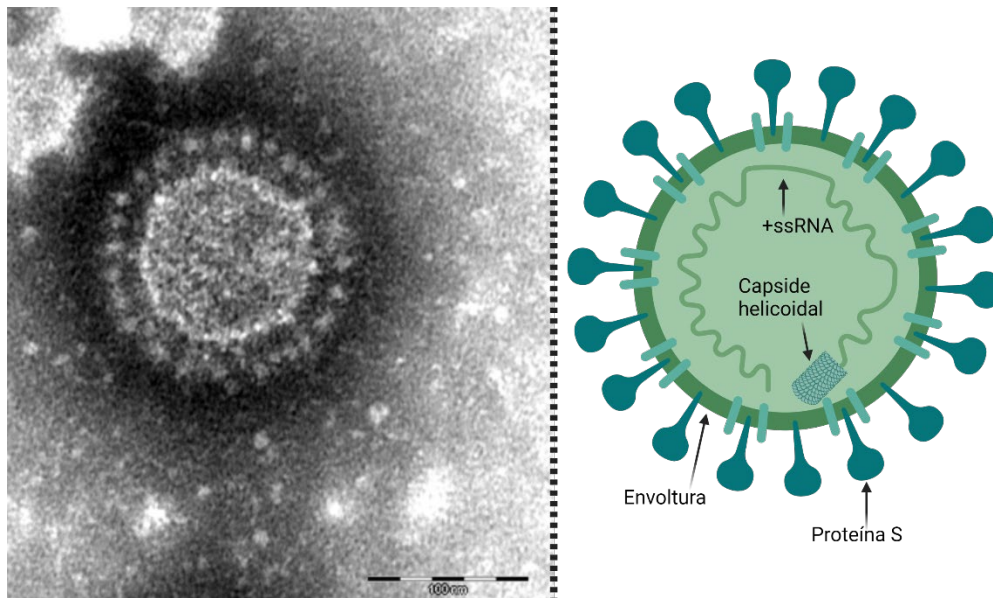


Figura 8. Virus PED obtenido por microscopía (izquierda) y esquematización (derecha).
Obtenida de (Hanke et al., 2015)

A este virus se le atribuyen efectos tales como diarrea, vomito, deshidratación, lo cual se asocia con una alta mortalidad, en especial en el caso de los lechones con una mortalidad cercana al 100%. (Jung et al., 2020)

El PEDV tuvo su primera aparición en países europeos como Inglaterra y Bélgica en los años 70's. Sin embargo este virus no se había encontrado fuera de Europa y Asia al menos hasta 2013, año en que fue encontrado en zonas de Norte América. Dicho brote se considera uno de los eventos más significativos dentro de la industria en los últimos años. A pesar de que en Europa los brotes infecciosos ocasionados por el PEDV

no han sido masivos y se han considerado como de poca importancia, en otros lugares como Asia o América han tenido gran impacto. (Davies, 2015; Diel et al., 2016; Jung et al., 2020)

La llegada de este patógeno a Estados Unidos demostró la vulnerabilidad de la mayoría de las granjas al enfrentarse al PEDV. Las formas de llegada del virus se dieron por dos tipos de contactos: Directo (Debido a la movilización de cerdos infectados) e Indirecto. Logrando llegar a otras zonas geográficas totalmente distintas a las que se consideraba endémico.(VanderWaal et al., 2018)

La transmisión del virus se puede dar a distintas edades, generalmente la primera infección se puede dar en cerdos de engorda, después de esto el virus se transmite a las cerdas donde llega a los lechones y es ahí donde debido a la alta mortalidad del virus afecta a la mayoría de los lechones provocando una gran cantidad de muertes. Durante la infección se presentan distintas rutas mediante las cuales el virus logra esparcirse, sin embargo, la ruta más común es la ruta fecal-oral, debido al contacto de material expulsado por cerdos infectados. (Zhang et al., 2022)

El virus por medio de procesos adaptativos y evolutivos presenta cambios estructurales continuamente, esto debido a cambios en su ambiente lo que puede producir la propagación de distintas cepas, debido a esto los investigadores han decidido clasificar dichas cepas en dos grupos GI (Geno grupo clásico) GII (Geno grupo de variantes).(Zhang et al., 2022)

El virus PED requiere para su entrada a las células blanco, de la presencia de la enzima tripsina, esto debido a que dicha enzima facilita la fusión membrana por activación de la proteína spike del virus, por ello se piensa que el intestino es un sitio con altas probabilidades de contar con células con el receptor del PEDV debido a su alto contenido en tripsina. (Wicht et al., 2014)

PEDV puede ser escindido gracias a distintas proteasas, especialmente la tripsina, lo cual tiene un papel muy importante para la infección, donde dicha escisión ocurre

únicamente después de que el virus se ha logrado unir a su receptor. Dicha escisión se lleva a cabo en la parte del S2 de la proteína S. (Tan et al., 2021)

Durante este trabajo la cepa utilizada fue la cepa CV777 de la proteína S la cual apareció en por primera vez en China.

El genoma del PEDV está constituido de 28kb que codifican las poliproteínas las cuales son ORF1a y ORF1b, S, ORF3, envoltura (E), membrana (M) y la nucleocápside (N) como se observa en la figura 9. (Deng et al., 2016)

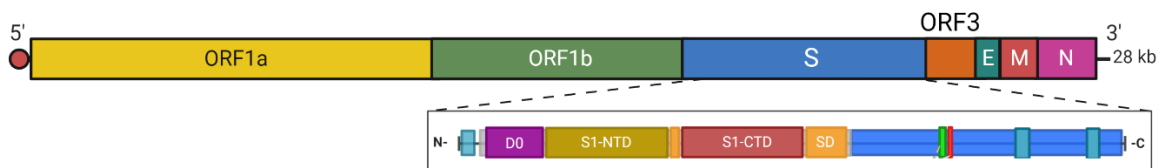


Figura 9. Genoma viral del PEDV CV777

La proteína S del PEDV es conformada por tres unidades idénticas de esta glicoproteína membranal, la cual principalmente consta de dos regiones de gran importancia las cuales son llamadas región S1 con residuos 34-764 y S2 con residuos 765-1383, donde la región S1 contiene principalmente tres dominios, el D0 (Dominio 0) con residuos 34-234, un dominio llamado NTD (Dominio N terminal) con residuos 235-474, y un dominio CTD (Dominio C terminal) con residuos 502-637, dominios esquematizados en la siguiente figura: (Huang et al., 2022; Lin et al., 2022)

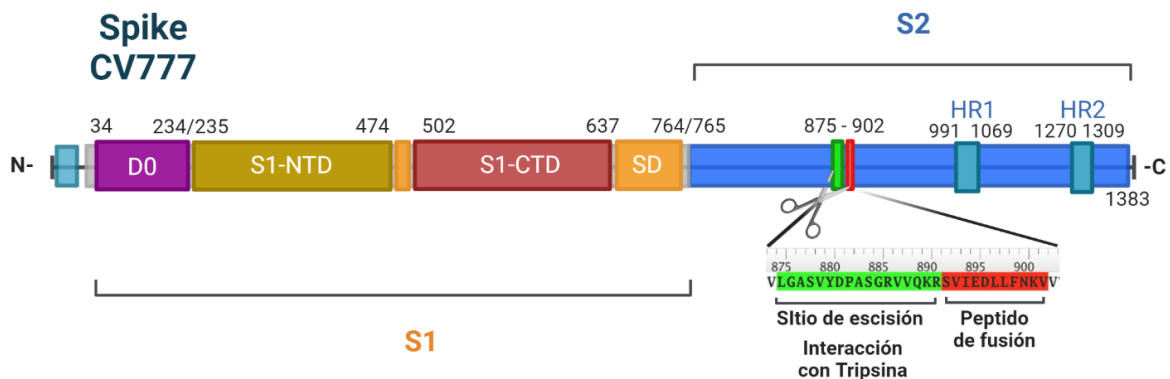


Figura 10. Dominios presentes en la proteína Spike PEDV CV777

De todos los dominios correspondientes al S1 uno de ellos adquiere gran importancia en el PEDV, este es el dominio CTD ya que dentro de este dominio se cree que existe algo que puede ser llamado RBD (*Receptor Binding Domain*), y es este RBD el cual se encarga de la unión con el posible receptor, realmente aún no se conoce con exactitud si el RBD se encuentra en el dominio NTD o CTD con total precisión, ni tampoco las funciones exactas de estos dominios. (Deng et al., 2016)

Se sabe que el dominio CTD puede tomar al menos dos distintas conformaciones como se muestra en la figura 11, una donde la proteína S se encuentra con dicho dominio con mayor proximidad al dominio NTD y el resto de la región S1, la cual se denomina conformación cerrada o abajo, por otro lado, si dicho dominio se encuentra con mayor exposición y más alejado de los dominios antes mencionados, se le conoce como conformación abierta o arriba.

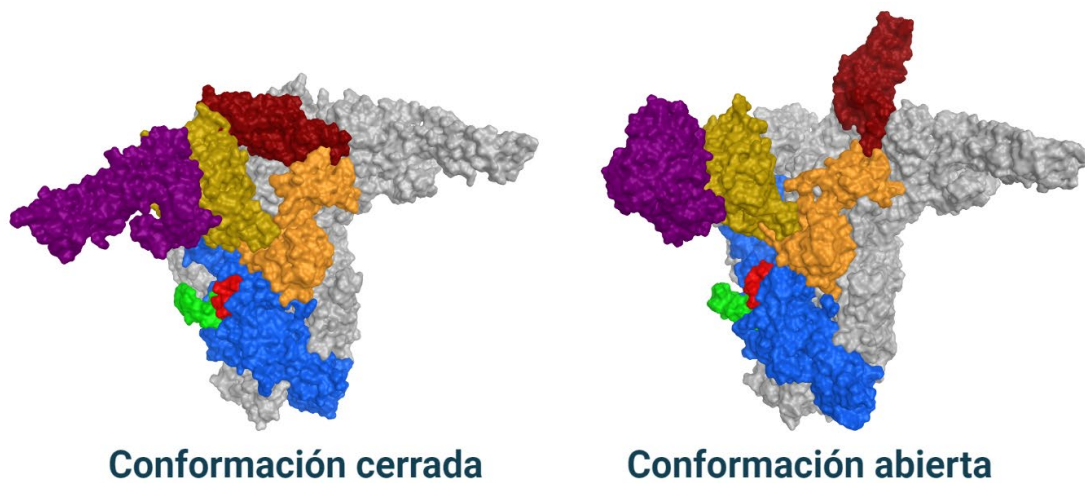


Figura 11. Representación de superficie molecular para la proteína Spike con el dominio CTD en ambas conformaciones. Imagen propia

5. QUÍMICA COMPUTACIONAL

La química computacional se encarga del estudio estructura-propiedades de átomos o moléculas, mediante la realización de cálculos utilizando métodos computacionales. Desde hace varios años la química computacional o también llamado modelado molecular ha sido de gran importancia en el desarrollo industrial, debido a que se considera económico y presenta un ahorro en los tiempos, traduciéndose en eficiencia, es por ello que en 1998 y en 2013 se entregaron premios Nobel de química a científicos que trabajan en el campo de la química computacional (Departamento de Medio Ambiente y Energía, s/f)

Dentro de la química computacional o el modelado molecular se pueden incluir distintos enfoques, entre ellos están los métodos *ab initio* y semi-empíricos usados en mecánica cuántica, mecánica molecular empírica, dinámica molecular, métodos Monte Carlo y métodos de energía libre, así como relaciones estructura actividad (SAR). (Schlick, 2010)

Se sabe que existen diferencias entre la química teórica y la mecánica molecular, donde la primera tiene sus bases en la mecánica cuántica, por otro lado la mecánica molecular se basa en la mecánica clásica.

5.1. Dinámica molecular

La mecánica molecular se basa en la física clásica (leyes de Newton), en donde la energía total está dada por el campo de fuerza. Esto está diseñado para la obtención de energías y estructuras moleculares. En el caso de la mecánica molecular en principio únicamente se involucra el movimiento de los núcleos, por lo que el modelado de los sistemas se puede asemejar a un modelo de esferas y resortes. El surgimiento de la mecánica molecular es producto de varias ecuaciones empíricas con las cuales es posible reproducir una superficie de energía potencial. Un punto para considerar es que es importante llevar a un estado de mínima energía a la molécula con el fin de llegar a un mínimo energético dentro de la superficie de energía potencial. (Arias, s/f)

MD (*Molecular dynamics*) o dinámica molecular es un tipo de simulación que hace uso de la mecánica molecular para el estudio de un sistema a través del tiempo, permitiéndonos analizar distintos parámetros de importancia, por medio de la dinámica molecular y la termodinámica estadística, se pueden obtener propiedades fisicoquímicas como energía libre, solubilidad, entalpía entre otros. Permite obtener la fuerza de interacción de un complejo dado por un receptor y un ligando a través del tiempo. (Lozano-Aponte & Scior, 2014)

Una vez realizadas las simulaciones de dinámica molecular dos principales parámetros a utilizar para analizar la dinámica son: por un lado el RMSD (*Root Mean Square Deviation*), el cual se basa en el cálculo de la diferencia entre dos posiciones atómicas, en este caso la posición inicial se compara con la posición del mismo átomo en un tiempo determinado, donde lo común es que el cambio o la diferencia de RMSD a lo largo de la etapa de producción durante una dinámica molecular no sobrepase los 2 Å, término que es absoluto, ya que depende enteramente del tamaño de la estructura estudiada, el cálculo del RMSD se obtiene mediante la siguiente expresión:

$$RMSD = \sqrt{\frac{1}{n} \sum_{i=1}^N \|v_i - w_i\|^2}$$

Ecuación 1. Obtención de RMSD de partículas atómicas.

Por otro lado, un segundo parámetro utilizado es el cálculo de RMSF (*Root Mean Square Fluctuation*), parámetro que nos permite conocer la fluctuación atómica a lo largo de la simulación, de esta forma se puede observar el movimiento por cada aminoácido a lo largo de toda la simulación. Es decir, que el punto de comparación en este caso es la posición promedio a lo largo de la dinámica.

5.2. Campo de Fuerza (Force Field)

De una forma clásica un campo de fuerza en mecánica molecular que considera a todos los átomos está compuesto de 6 términos. Un término que considera o toma en cuenta

los enlaces, así como su deformación. Otro termino que considera los ángulos formados. Después dos términos que se refieren sobre las rotaciones de ángulos diedros, que son aquellos ángulos formados entre cuatro átomos. Y como ultima contribución existen dos términos los cuales no hacen referencia a enlaces propios, estos son por un lado un término que describe los efectos de atracción y de repulsión, y aquel que considera las interacciones electrostáticas, de forma que podemos obtener la siguiente expresión: (Kmicik et al., 2016)

$$U(\vec{R}) = \sum_{\text{enlaces}} K_b(b - b_0)^2 + \sum_{\text{ángulos}} K_\theta(\theta - \theta_0)^2 + \sum_{\text{dihedros}} K_\chi(1 + \cos(n\chi - \delta)) + \sum_{\substack{\text{dihedros} \\ \text{impropios}}} K_{imp}(\varphi - \varphi_0)^2 \\ + \sum_{\text{Lennard-Jones}} \left(\epsilon_{ij} \left(\frac{R_{min_{ij}}}{r_{ij}} \right)^{12} - \left(\frac{R_{min_{ij}}}{r_{ij}} \right)^6 \right) + \sum_{\text{electrostáticas}} \frac{q_i q_j}{4\pi\epsilon_0\epsilon r_{ij}}$$

Ecuación 2. Principales componentes de un campo de Fuerza

Dentro de los campos de fuerza más utilizados se encuentran algunos para proteínas como pueden ser: AMBER-98, GROMOS98, CHARMM22, OPLS/AA. O para el caso de ácidos nucleicos: AMBER-94/98, CHARM29, BMS.

El rápido avance tecnológico ha permitido el desarrollo de nuevos campos de fuerza, por ejemplo, aquellos especializados para la simulación de membranas, entre otros.

5.3. Acoplamiento molecular

El acoplamiento molecular pretende realizar una simulación de unión o interacción entre dos moléculas, donde generalmente la de mayor tamaño es llamada receptor y la de menor tamaño ligando.

Esta técnica basada en mecánica molecular nos permite predecir energías de interacción, las formas de enlace también llamado modo de unión entre el receptor y el ligando.

El acoplamiento molecular es una forma muy interesante de abordar temas de biología molecular, así como en el diseño de fármacos.(Morris & Lim-Wilby, 2008)

5.3.1. Funciones de registro

Un estudio de acoplamiento molecular puede arrojarnos cientos o miles de resultados, por lo que debe de existir una forma de buscar los mejores resultados, a esto se le conoce como *Score function* o Funciones de puntuación o registro. Esta fase para asignar las puntuaciones a los distintos modos de unión se puede basar en el uso del campo de fuerzas utilizado, donde se consideran las interacciones entre los aminoácidos de ambas proteínas.(Sunny & Jayaraj, 2022)

Dichas funciones son usadas para entre todas las poses encontrar un modelo de unión, así como el mejor sitio para que este se encuentre, es por eso por lo que existen distintos tipos de *Scoring functions* los cuales dependerán totalmente del software utilizado para los cálculos.(J. Li et al., 2019)

6. OBJETIVOS

Objetivo general

- Realizar una búsqueda de posibles receptores virales y mediante estudios de acoplamiento molecular proponer los candidatos más favorables como receptores del PEDV.

Objetivos particulares

- Realizar una búsqueda bibliográfica de receptores implicados en la entrada viral.
- Obtener estructuras tridimensionales a partir de bases de datos de proteínas, por medio de homología o mediante uso de nuevas tecnologías como Alphafold.
- Analizar la energía libre entre los complejos Spike-receptor con el dominio CTD en conformación abierta y cerrada.
- Determinar los aminoácidos involucrados en la interacción Spike-receptor.
- Proponer los candidatos más favorables como receptores del PEDV de acuerdo con los resultados de afinidad.

7. METODOLOGÍA

La metodología utilizada de forma general se encuentra esquematizada en el siguiente diagrama:

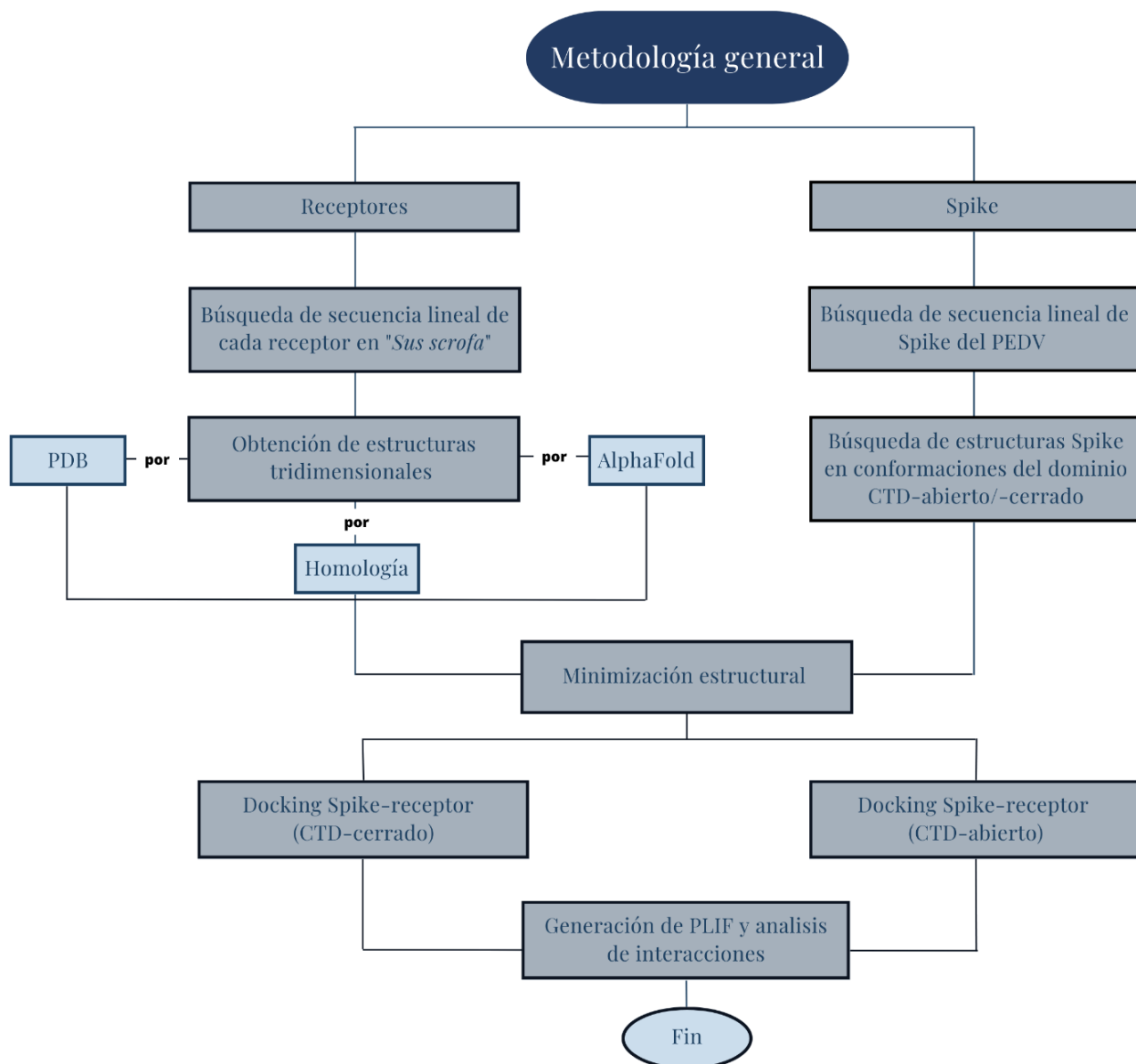


Figura 12. Diagrama de flujo de la metodología general

El uso de los distintos softwares se encuentra esquematizado en el siguiente diagrama:

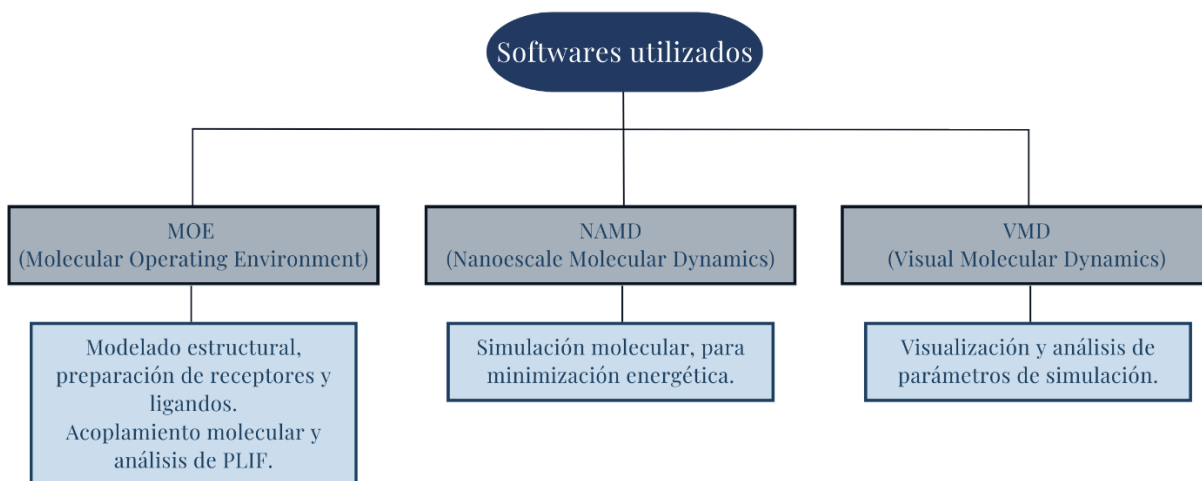


Figura 13. Resumen de los softwares utilizados

7.1. Búsqueda de receptores

Se realizó una búsqueda de algunos de los receptores celulares implicados en la entrada de ciertos virus de gran importancia, se buscaron candidatos que se encontraran ya reportados en la literatura como receptores virales, así como que existiera la posibilidad de obtener sus estructuras tridimensionales, para de esta forma mediante el modelado molecular y distintos estudios computacionales proponer de esta lista de receptores el mejor candidato como receptor del PEDV.

7.1.1. CD4

CD4 es una glucoproteína monomérica membranal de 50.9 kDa, también conocida como molécula de diferenciación 4, la cual aparece en células T, que generalmente tiene la función de T cooperador o auxiliar.

En el caso de los cerdos el CD4 tiene identificado dos dominios extracelulares Ig-like (*Inmunoglobulin-like*). En el caso de los humanos se conocen 4 dominios extracelulares, cuenta con un dominio hidrofóbico transmembranal y una parte final dentro del citoplasma.(Owen & Punt, 2016)

CD4 presenta interacciones con la glucoproteína gp120, por lo que CD4 es responsable de la entrada del VIH-1 (Virus de Inmunodeficiencia Humana 1), actuando

como receptor para un cambio conformacional que conduce a la exposición del péptido de fusión y la unión de este con la membrana celular.(Bour et al., 1995; Lama, 2003)

7.1.2. CD8

Las células T poseen en su superficie unas glicoproteínas de membrana denominadas CD8, que suelen actuar como células citotóxicas. Los CD8 pueden presentar una estructura heterodimérica $\alpha\beta$ o una homodimérica $\alpha\alpha$, y se componen de dos tipos de cadenas de pequeñas glicoproteínas, α y β , con un peso aproximado de 30-38 kDa cada una. Cada cadena está formada por un solo dominio extracelular, una región transmembrana hidrofóbica y una cola citoplasmática que incluye de 25 a 27 residuos, algunos de los cuales pueden estar fosforilados.(Owen & Punt, 2016)

7.1.3. CD13 (Aminopeptidasa)

La enzima aminopeptidasa interviene en la última etapa de la descomposición de los péptidos resultantes de la digestión de las proteínas por las proteasas gástricas y pancreáticas. Múltiples artículos la han identificado como receptor de alfa coronavirus. En artículos específicos sobre el PEDV, se menciona la aminopeptidasa como receptor del virus, así como del virus de la gastroenteritis transmisible.(Lin et al., 2022)

La aminopeptidasa se encuentra en la membrana microviliar, teniendo presencia en el intestino delgado, el riñón y otras membranas plasmáticas.

7.1.4. CD26 (DPP4)

La proteína DPP4 (Dipeptidil Peptidasa), también conocida como CD26, es una serina proteasa que se ha reportado como receptor funcional del coronavirus del Síndrome Respiratorio de Oriente Medio (hCoV-EMC). Según algunos artículos, se teoriza que virus como el SARS COV y el MERS se unen a residuos fuera del pocket, que corresponden a la numeración 336-499. Además, el MERS HKU4 reconoce la serina proteasa DPP4 en este caso. Su estructura consta de un dominio citoplasmático N-terminal de seis residuos, un segmento transmembrana de 22 aminoácidos y una parte extracelular más grande de 738 aminoácidos. La parte extracelular comprende un tallo flexible corto (aminoácidos 29-39), una región rica en glucosilación (aminoácidos 101-

350), una región rica en cisteína (aminoácidos 55-100, 351-497) y un dominio catalítico (aminoácidos 506-766).(Raj et al., 2013; Wang et al., 2013)

7.1.5. CD46

La interacción clave ocurre debido a una sección expuesta de la proteína CD46 en su dominio N-terminal. Esta sección permite una alta interacción con una zona hidrofóbica encontrada en la proteína MV-H.(Santiago et al., 2010)

Resultados indican que CD46 humano (también conocido como MCP) es un receptor del virus del sarampión (MV). Esta proteína está ampliamente distribuida en los tejidos humanos.(Dörig et al., 1993)

7.1.6. CD54 (ICAM-1)

La familia de proteínas ICAM es una de varias moléculas de adhesión celular (como integrinas, selectinas y mucinas) que contienen un número variable de dominios similares a inmunoglobulinas. La función normal de la molécula de adhesión intercelular humana-1 (ICAM-1) es proporcionar adhesión entre células endoteliales y leucocitos después de una lesión o estrés. ICAM-1 se une al antígeno asociado a la función de los leucocitos o al antígeno-1 de los macrófagos. Sin embargo, ICAM-1 también es utilizado como receptor por el grupo principal de rinovirus humanos y es un catalizador para la liberación del material viral durante la entrada a la célula.(Bella & Rossmann, 1999)

Una gran cantidad de rinovirus humanos, de los 100 serotipos identificados al menos el 80% de estos utiliza la proteína ICAM-1 como receptor sin embargo no utilizan ICAM-2 ni ICAM-3.(Bella et al., 1998)

7.1.7. CD66 (CEACAM-1)

CD66 tiene un papel como un receptor co-inhibitorio en la respuesta inmune, la acción de la insulina y funciona como activador durante la angiogénesis. (Chen et al., 2008)

CD66c se utilizó en un estudio para buscar un posible receptor para el virus de la influenza. Después de realizar el estudio, se descubrió que a pesar de que el ácido siálico actúa como receptor viral, CD66c al menos actúa como un receptor durante la infección

por influenza. Por otro lado, MHV se une a CEACAM-1 mediante interacciones proteína-proteína, donde el dominio NTD en S1 sirve como sitio de unión.(Bella et al., 1998; F. Li, 2016; Rahman et al., 2021)

7.1.8. CD209

Las proteínas CD209/DC-SIGN y CD209L/L-SIGN funcionan como receptores únicos que reconocen patógenos y facilitan la adhesión celular, permitiendo así interacciones con la célula. Estas proteínas son capaces de reconocer una variedad de patógenos, incluyendo virus como SARS-CoV-2, SARS, bacterias, hongos y parásitos. Los patógenos utilizan estas proteínas para promover la infección y evadir la detección inmunitaria. Además, CD209L y CD209 se expresan abundantemente en los órganos objetivo de SARS-CoV-2, lo que sugiere que pueden contribuir a la patogénesis y progresión de la infección.(Rahimi, 2020)

7.1.9. CD249 (ENPEP)

La familia M1 de endopeptidasas incluye la glutamil aminopeptidasa (ENPEP), que es una endopeptidasa de zinc integral de tipo II de membrana en mamíferos. Esta enzima es capaz de producir angiotensina III, un componente de la vía catabólica del sistema renina-angiotensina que juega un papel importante en el control de la presión arterial y el desarrollo de los vasos sanguíneos.(Holmes et al., 2017)

ENPEP es una proteína que ha sido identificada como un receptor de virus, incluyendo los coronavirus. Se ha descubierto que la proteína S del virus SARS-CoV-2 se adhiere a ENPEP en la superficie de las células humanas, permitiendo así la entrada y propagación del virus. En vista de esta información, se ha investigado la posibilidad de utilizar ENPEP como objetivo para el tratamiento de COVID-19.(Masre et al., 2021)

7.1.10. ECA-2

El ECA-2 es una enzima importante que forma parte del sistema renina-angiotensina-aldosterona (RAAS), una serie de proteínas que desempeñan un papel importante en los procesos fisiológicos del cuerpo. Esta enzima se encuentra y se une a la membrana plasmática de varias células del cuerpo, como el epitelio alveolar del

pulmón, el epitelio tubular renal, las células de Leydig del testículo y el tracto gastrointestinal.(Lin et al., 2022)

Se ha identificado que ACE-2 funciona como un receptor para dos tipos de coronavirus: alfa (HCoV-NL63) y beta (SARS-CoV). (F. Li, 2016)

Según estudios realizados, se ha comprobado que los coronavirus se conectan directamente a la enzima ECA-2 presente en las células huésped, utilizando para ello sus proteínas S de espiga. (Zhou et al., 2018)

7.1.11. Ocludina

La ocludina es una proteína integral de membrana tetraspanina de 59 kDa (*Sus scrofa*) que desempeña un papel fundamental en la estabilización de las uniones estrechas (TJ) y en la función óptima de la barrera. Las TJ son complejos multiproteicos intramembranosos que se encuentran en células cercanas en monocapa epiteliales y endoteliales, y que proporcionan conexiones intercelulares apicales.(Cummins, 2012)

La infección de células epiteliales por PEDV provoca la interrupción de la distribución de la ocludina en las uniones estrechas, lo que aumenta la susceptibilidad de las células a la infección. Aunque la sobreexpresión de la ocludina en las células también las hace más vulnerables a la infección, se sabe que la ocludina desempeña un papel crucial en las etapas posteriores a la unión durante la infección por PEDV.(Luo et al., 2017)

7.1.12. Tripsina

La tripsina pertenece a la familia S1 dentro de las peptidasas, La tripsina puede escindir en residuos como lisina o arginina, sin embargo, esto no ocurre si alguno de estos dos aminoácidos se encuentra seguido por un residuo de prolina.

TGEV (*Transmissible Gastroenteritis Virus*) y PRCv (*Porcine Respiratory Coronavirus*) usan como receptor en lugar de ECA-2 a la tripsina, sin embargo, en el caso del TGEV algún azúcar puede actuar como correceptor.(F. Li, 2016)

Como se muestra en la tabla 1 se buscaron los dominios reportados para cada una de las proteínas utilizadas con el objetivo de separar aquellos dominios extracelulares de los que se encuentran presentes dentro del citoplasma.

Tabla 1. Residuos correspondientes a cada dominio en cada receptor

Receptor	Dominio	Residuos
	Transmembrana	397-422
CD4	Dominio Ig	20-122
	Dominio Ig	138-196
CD8	Dominio Ig	11-118
CD13	Citoplasmático	2-8
(Aminopeptidasa)	Transmembrana	9-32
	Extracelular	33-993
CD26	Citoplasmático	1-6
(DPP4)	Transmembrana	7-27
	Extracelular	28-766
CD46	Extracelular	45-329
	Sushi 1-4	45-294
CD54	Dominio Ig	117-193
(ICAM-1)	Dominio Ig	400-478
	Transmembrana	486-508
CD66 (CEACAM-1)	Dominio Ig	117-193
	Dominio Ig	236-325
CD209	Transmembrana	38-62
	Tipo lectina	119-231
CD249	Citoplasmático	1-14
(ENPEP)	Transmembrana	15-35
	Extracelular	36-942
ECA-2	Zinc metalopeptidasa	19-611
	Transmembrana	741-761
Ocludina	Transmembrana	67-89,244-265
	OCEL	414-522
Tripsina	Peptidasa	9-229

7.2. Obtención de estructuras

Una vez preparada la lista de las proteínas como posibles receptores, fue necesario buscar las estructuras de dichas proteínas. La forma en que se buscaron cada una de las proteínas fue mediante la plataforma de UniProt (Bateman et al., 2023), la cual es una base de datos que almacena secuencias de aminoácidos de distintas proteínas. Se buscaron las proteínas seleccionadas para el organismo del cerdo o (*Sus scrofa*) y una vez obtenidas las secuencias de aminoácidos se buscaron sus representaciones estructurales.

Esta búsqueda se realizó mediante los servicios de NCBI (*National Center for Biotechnology Information*) por medio de un BLAST (*Basic Local Alignment Search Tool*) de proteína en una cola de búsqueda comparando con la secuencia lineal de la proteína comparando con estructuras depositadas en bases de datos RCSB PDB (*Research Collaboratory for Structural Bioinformatics Protein Data Bank*).

Uno de los pasos fundamentales para el proceso de acoplamiento molecular es la obtención de las estructuras tridimensionales de las proteínas o moléculas a utilizar, en el caso de moléculas pequeñas es posible realizarlas mediante el uso de softwares que permitan la elaboración de estructuras químicas, sin embargo para el caso de proteínas no es tan sencillo.

Por lo general, hay sitios web como el PDB que recopilan información sobre las estructuras de proteínas y macromoléculas biológicas derivadas de estudios científicos. Estos estudios facilitan la obtención de estructuras proteicas tridimensionales, las cuales pueden ser obtenidas mediante diferentes métodos, como el uso de rayos X o microscopía criogénica donde la elección del método o técnica depende de la resolución deseada en los experimentos. Uno de los factores más importantes para obtener estas estructuras son los datos disponibles, y cada método se utiliza según las características y la resolución que se busque alcanzar. En resumen, la elección de la técnica está

determinada por la resolución deseada y la naturaleza de la estructura que se busca obtener.

A pesar de que el PDB es el banco de estructuras más conocido y utilizado, no es el único, existen otras bases de datos disponibles que contienen información estructural de proteínas, sin embargo cada una de ellas puede tener un enfoque distinto. Al encontrar la estructura que deseemos utilizar basta con buscar dentro de la página del PDB y descargar la estructura según el formato que deseemos utilizar, el formato más utilizado para el manejo de estructuras proteicas es el *.pdb.

El número de estructuras depositadas en bases de datos ha ido en un gran aumento con el pasar de los años, hasta el punto de tener cientos de miles de estructuras accesibles para su descarga, sin embargo hay una cantidad mucho mayor de proteínas existentes en el mundo por lo que existen además de los métodos experimentales mencionados anteriormente otras formas de obtener estructuras proteicas, una de ellas es el modelado por homología.

El modelado por homología utiliza como base que proteínas homologas tienen una estructura tridimensional similar, por lo que en regiones conservadas se presenta una estructura tridimensional parecida, mediante esto es posible predecir una estructura tridimensional de una secuencia de aminoácidos sin estructura reportada usando como base una estructura conocida de su proteína homóloga conocido como modelado de proteínas.

Dicho modelado no deja de ser un modelo predictivo por lo que puede existir cierta incertidumbre en la obtención de dichas estructuras.

Cuando no existe una proteína homóloga en la cual basarse para la creación de estructuras tridimensionales, existe una forma más reciente para su obtención esa es mediante el uso de técnicas de aprendizaje profundo y redes neuronales, tal es el caso de AlphaFold, el cual es un software de inteligencia artificial que predice estructuras tridimensionales de proteínas partiendo de una secuencia de aminoácidos. La capacidad de predecir estructuras tiene gran importancia en distintas áreas de investigación.

7.3. Spike

Se comenzó con el modelado estructural de la proteína S del PEDV, se obtuvo la estructura inicial de la base de datos del PDB con código 6U7K el cual corresponde a la proteína en un estado de pre-fusión.

El archivo cristalográfico obtenido de PDB presentaba segmentos de la proteína sin modelar, es por esto se utilizó el software MOE 2022 (Chemical Computing Group ULC, 2023), con ayuda de una herramienta propia de MOE llamada *loop modeler* se modelaron estructuralmente los segmentos de faltantes siendo estos los residuos modelados:

- | | |
|------------|------------|
| a) 64-83 | e) 618-642 |
| b) 139-153 | f) 674-692 |
| c) 174-189 | g) 809-817 |
| d) 240-265 | h) 855-862 |

Una vez modelados los segmentos faltantes, dentro del mismo software se procedió a realizar una minimización estructural, dichas minimizaciones se llevaron a cabo utilizando AMBER 12: EHT como campo de fuerza, para obtener de esta forma nuestro modelo final de la proteína S en su conformación RBD-cerrado.

Para la obtención de la proteína S en conformación abierta se encontró en la base de datos de PDB que existía reportada el homotrímero de la proteína spike con la región CTD en conformación abierta con código de PDB 7Y6V.

Se preparó dicho homotrímero, con MOE para que dicha proteína constara de la misma secuencia de aminoácidos usada para la conformación cerrada correspondiente a la cepa CV777 y de esta forma modelando una cadena del homotrímero original con la región CTD expuesta se obtuvo el modelo de la proteína S en su conformación RBD-abierta, en la figura 14 se muestran ambas conformaciones.

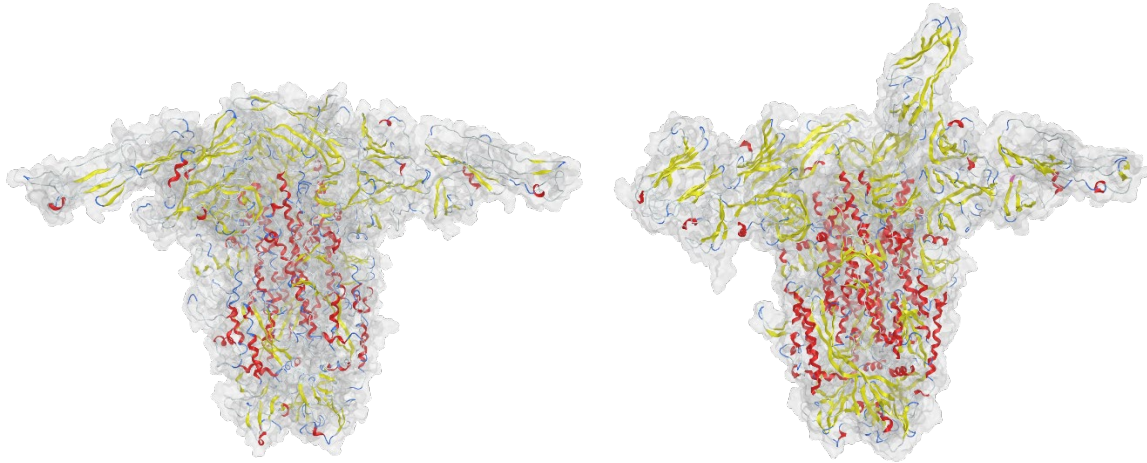


Figura 14. Proteína Spike en ambas conformaciones del dominio CTD, derecha se muestra en la conformación abierta el dominio CTD y en el lado izquierdo en la conformación cerrada.

Una vez modeladas las proteínas para ambas conformaciones se minimizaron utilizando el programa MOE.

7.4. Receptores

La obtención de estructuras de los receptores se realizó empleando distintos métodos, por una parte, directamente de la base de datos del PDB únicamente se obtuvieron las estructuras de las proteínas CD8, CD26 y Tripsina pertenecientes al organismo *Sus scrofa*.

Por medio del modelado por homología se obtuvieron las estructuras de las proteínas CD4, CD13, CD46, ICAM-1, CEACAM-1, CD249, ECA-2.

Para el modelado por homología se tomaron las secuencias de aminoácidos de las proteínas obtenidas de UniProt, una vez obtenidas dichas secuencias, dentro del programa MOE usando la herramienta de *Homology model* se prepararon las estructuras finales a usar como potenciales receptores.

Para las proteínas restantes ya que no se contaba con una plantilla modelo para modelarse, se utilizó AlphaFold, un punto a considerar es que AlphaFold muestra la confianza del modelo generado, debido a esto se buscó que los dominios extracelulares de las proteínas tuvieran un mayor puntaje de confianza.

La forma en la que se obtuvieron las estructuras así como su identificador en UniProt se muestran en la siguiente tabla:

Tabla 2. Se muestra la forma en que se obtuvieron las estructuras tridimensionales, así como su código en UniProt de cada receptor.

Receptor	Código UniProt	Estructura
CD4	Q6R3N4	Modelo de Homología
CD8	A0MNZ3	PDB: 5EDX
CD13 (Aminopeptidasa)	P15145	Modelo de Homología
CD26 (DPP4)	P22411	PDB: 5LLS
CD46	O02839	Modelo de Homología
CD54 (ICAM-1)	F1S3J9	Modelo de Homología
CD66 (CEACAM-1)	A0A8D1FU76	Modelo de Homología
CD209	B3FVQ2	AlphaFold: AF-B3FVQ2-F1
CD249 (ENPEP)	Q95334	Modelo de Homología
ECA-2	K7GLM4	Modelo de Homología
Ocludina	F1SK31	AlphaFold: AF-F1SK31-F1
Tripsina	P00761	PDB: 2A31

Una vez obtenidas las estructuras de todos los receptores, cada uno fue minimizado usando el programa MOE.

Para llevar a un estado de relajación se realizaron dinámicas moleculares, para las proteínas obtenidas por homología, PDB y AlphaFold. La preparación del sistema se realizó bajo condiciones periódicas, solvatando la proteína con moléculas de agua, donde las dimensiones de la caja solvatada dependieron totalmente del tamaño de cada proteína, sin embargo se consideraron márgenes de 10 Å para la solvatación.

El programa usado para las dinámicas moleculares fue NAMD (*Nanoscale Molecular Dynamics*), software especializado en simulaciones moleculares.

La realización de las dinámicas moleculares se empleó usando el mismo protocolo para todas las estructuras el cual constó de principalmente tres etapas:

Tabla 3. Protocolo de dinámica molecular empleado.

Etapas	Tiempo (ps)	T (K)	P(kPa)
Calentamiento	1000	0-300	
Equilibrio	2000	300	
Producción	30000	300	101

Se utilizó el programa MOE para realizar el acoplamiento molecular, dicho acoplamiento utilizó la opción interna de MOE de acoplamiento proteína-proteína.

Se crearon sets para la separación de dominios de la proteína Spike, el principal set creado fue para la región S1. Por parte de los receptores de igual forma se crearon sets, estos sets principalmente se utilizaron para identificar las regiones extracelulares de cada receptor.

Para cada proteína se orientó el acoplamiento priorizando una unión de la región S1 de la proteína spike y aquellos aminoácidos que se encuentran en dominios extracelulares, utilizando los sets creados anteriormente, los cuales son los más importantes para la unión de este.

Cada receptor siguió las configuraciones predeterminadas para el acoplamiento proteína-proteína propias del programa. Esto quiere decir que en una fase de colocación previa se obtuvieron 10000 poses, seguido a esto en una etapa de colocación se mantuvieron 1000 poses y por último en una etapa de refinamiento dando 100 poses finales. El refinamiento utilizado fue de cuerpo rígido, mientras que la obtención final de energía fue mediante GB/VI.

Adicionalmente se realizó un proceso de acoplamiento para la proteína Spike aislada. Los resultados obtenidos después del proceso de acoplamiento molecular se almacenaron en bases de datos conteniendo el score final de cada pose.

Se realizaron dos acoplamientos por receptor, uno utilizando la conformación de la proteína S con la región o dominio CTD cerrada y otro usando la conformación

abierta. El acoplamiento molecular fue realizado utilizando AMBER 12: EHT como campo de fuerza, el cual se encuentra parametrizado para pequeñas moléculas y grandes moléculas. Una vez terminado el proceso de acoplamiento se tomaron como posiciones correctas aquellas que se encontraban cerca de la región S1 y tomando una ventana de energías de 30 kcal/mol.

El objetivo, como se muestra en la figura 15 es el modelado de la posible interacción entre la proteína Spike con su posible receptor viral.

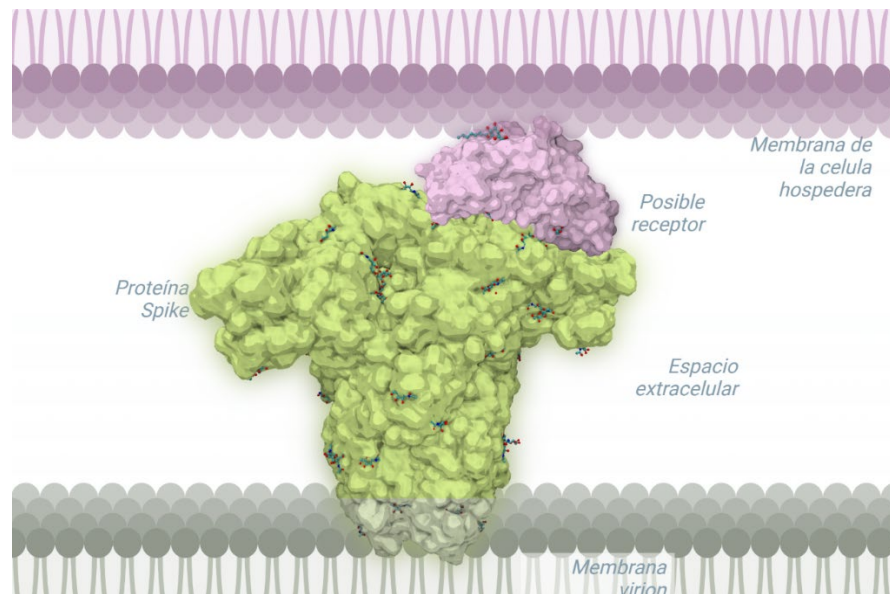


Figura 15. Esquematización de los complejos buscados entre la proteína Spike y las proteínas utilizadas como receptores. (Realizada con BioRender.com)

Se realizó un análisis PLIF (*Protein-Ligand Interaction Fingerprints*) con aquellas posiciones más favorables durante el acoplamiento molecular, esto con el objetivo de identificar las mayores zonas de interacción entre los receptores y la proteína S.

Se obtuvo una lista de interacciones por cada receptor para cada conformación para observar los cambios que se presentan ante la exposición del dominio CTD.

8. RESULTADOS Y ANALISIS

El primer parámetro que se evaluó después del proceso de dinámica molecular fue la obtención del RMSD a lo largo de la trayectoria de simulación utilizando la herramienta de VMD (Visual Molecular Dynamics). Un ejemplo de eso lo muestra la siguiente imagen, donde se puede ver el gráfico de RMSD a lo largo de la dinámica para una de las proteínas utilizadas:

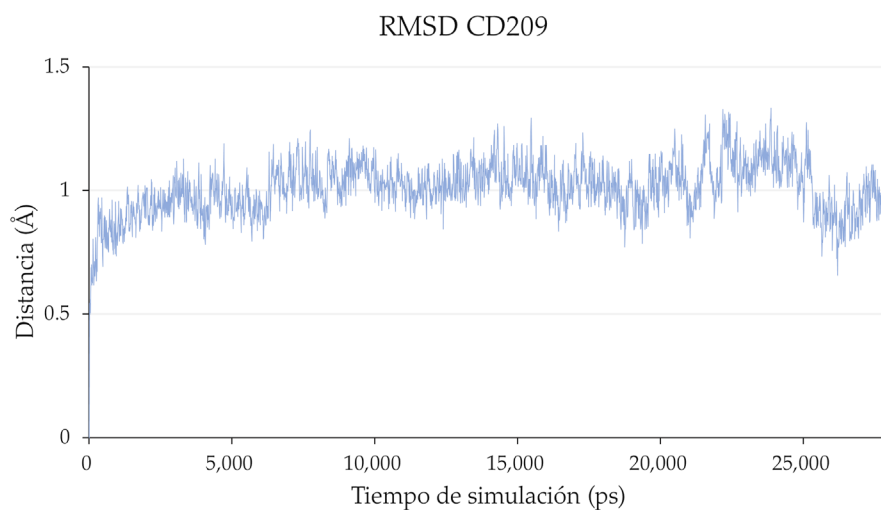


Figura 16. RMSD a lo largo de las dinamica molecular para la proteína CD209.

También se evaluó el RMSF, esto para conocer principalmente la fluctuación estructural de los carbonos alfa de la proteína, en este caso la figura 17 muestra el grafico al final de la simulación para la proteína CD209.

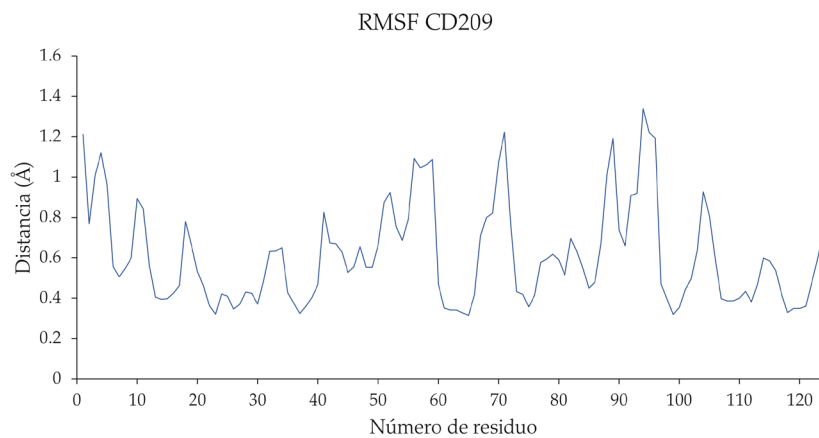


Figura 17. Ejemplo de cálculo de RMSF a lo largo de la dinámica para la proteína CD209.

Una vez que se obtuvieron las estructuras finales la energía obtenida de los resultados de acoplamiento se pueden observar en la siguiente tabla, donde se muestran los resultados para cada receptor en ambas conformaciones de la zona CTD de la proteína spike:

Tabla 4. Energía correspondiente al acoplamiento entre la proteína spike y cada receptor, donde se muestran ambas conformaciones del RBD.

Receptor	Energía de acoplamiento (kcal/mol)	
	RBD cerrado	RBD abierto
CD4	-83.55 ± 7.74	-103.84 ± 21.01
CD8	-57.24 ± 7.88	-58.89 ± 8.06
CD13 (ANPEP)	-69.13 ± 8.45	-74.41 ± 8.44
CD26 (DPP4)	-93.18 ± 8.32	-89.38 ± 9.94
CD46	-65.26 ± 10.05	-66.21 ± 6.54
CD54 (ICAM-1)	-82.20 ± 9.13	-89.61 ± 9.94
CD66 (CEACAM-1)	-63.55 ± 7.61	-80.83 ± 7.87
CD209	-59.07 ± 6.01	-70.83 ± 6.27
CD249 (ENPEP)	-79.33 ± 9.18	-77.08 ± 9.37
ECA-2	-77.57 ± 6.92	-70.26 ± 6.77
OCLIDINA	-62.46 ± 7.70	-59.16 ± 7.66
TRIPSINA	-63.49 ± 6.83	-65.55 ± 8.51

Adicionalmente se realizó un acoplamiento con las mismas características usadas para todos los receptores, pero en este caso para la proteína Spike. Debido a que la proteína Spike está compuesta de tres cadenas de aminoácidos idénticas. Se modelaron dos cadenas de la proteína para que esto pudiera ser usado como receptor y una cadena usada como ligando. De los resultados obtenidos, únicamente se eligió la posición 1, esto debido a la gran brecha energética presentada en los resultados finales. Se obtuvo que la energía final del acoplamiento de las cadenas de la proteína Spike fue de: -896.15 kcal/mol.

El resultado anterior demuestra como este complejo homotrimerico presenta una alta estabilidad gracias a la afinidad entre cadenas.

Los resultados finales de energía de la tabla 4 se pueden observar en las siguientes graficas:

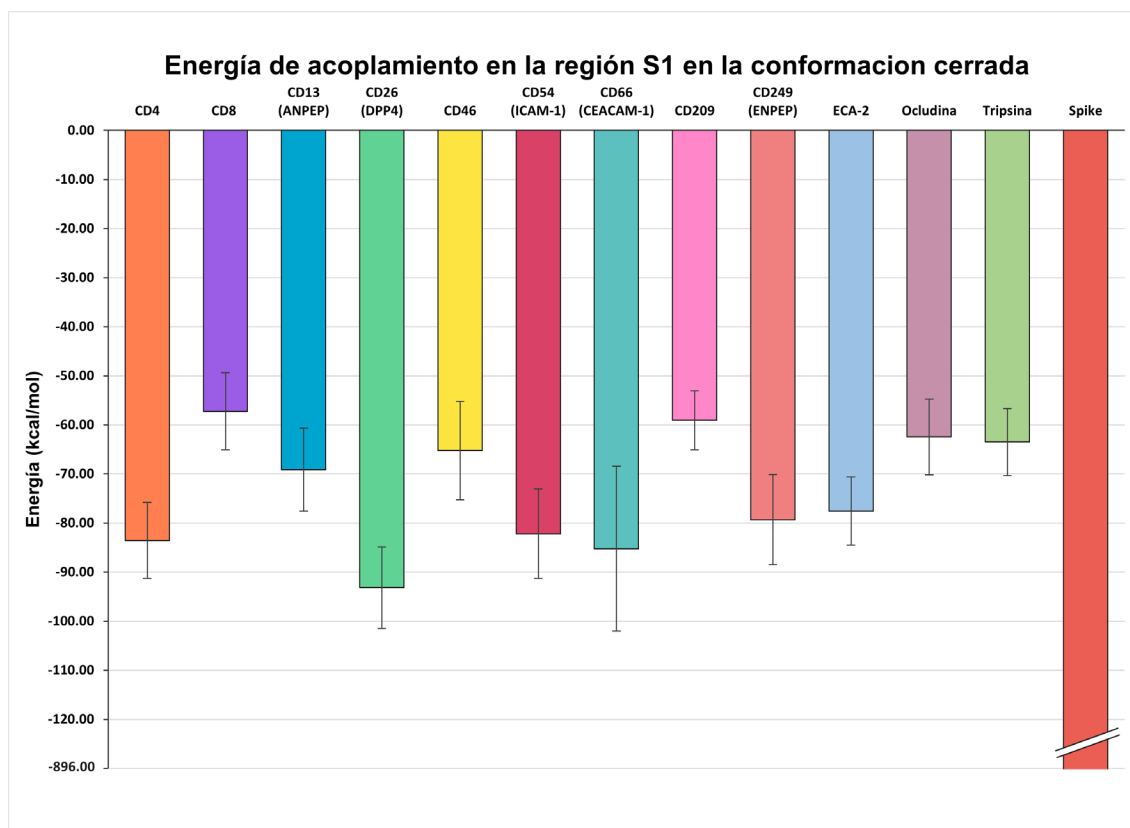


Figura 18. Comparativa entre energías finales de los resultados del acoplamiento en conformación cerrada.

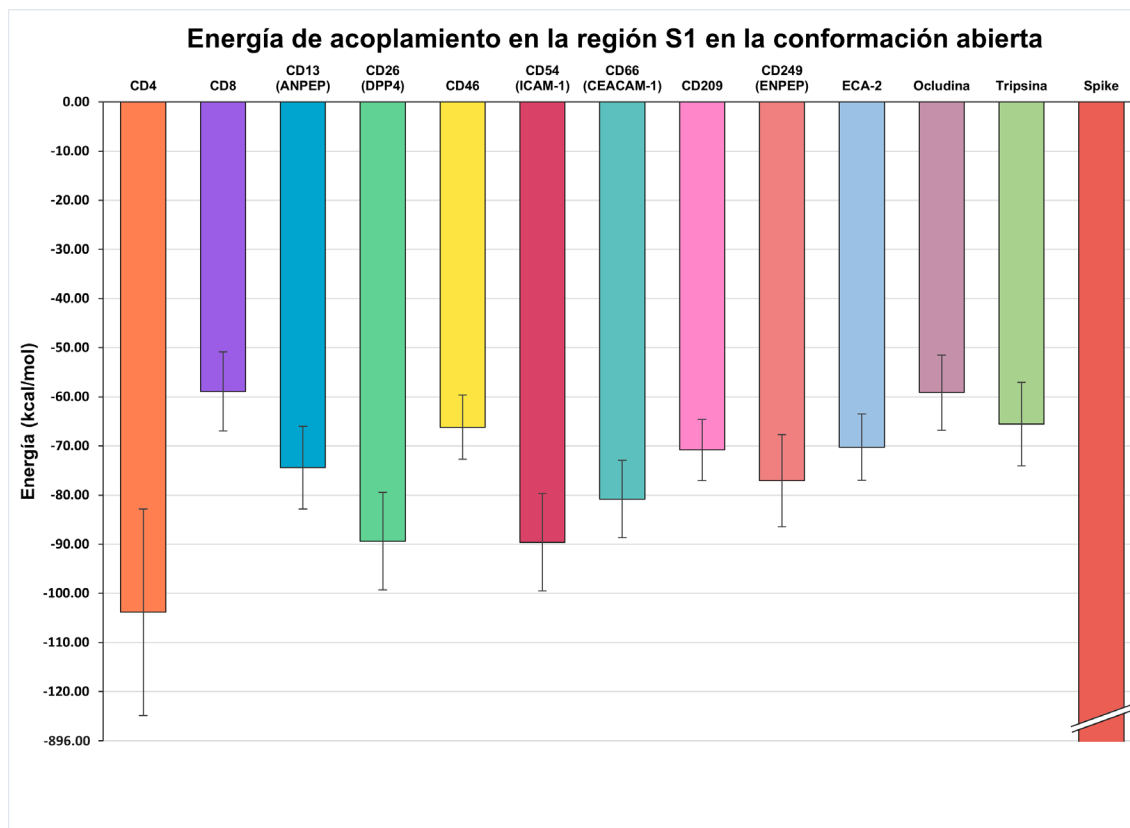


Figura 19. Comparativa entre energías finales de los resultados del acoplamiento en conformación abierta.

Como se puede observar en las figuras anteriores existe un cambio en la afinidad de los receptores según la exposición del dominio RBD.

La tripsina es una de las proteínas que presenta menor afinidad a la región del RBD, estos resultados presentan similitud con lo predicho experimentalmente, ya que se sabe que la tripsina es importante durante para la replicación del PEDV sin embargo no tiene acción como receptor, si no que ayuda catalizando por medio de hidrolisis ciertos aminoácidos que se encuentran en la región S2 de la proteína y de esta forma ayudando a la exposición del péptido de fusión, péptido que posteriormente se une a la membrana.

Los resultados de acoplamiento para la tripsina presentan una gran tendencia a la región S2, cerca del péptido de fusión, esquematizado en la figura 20.

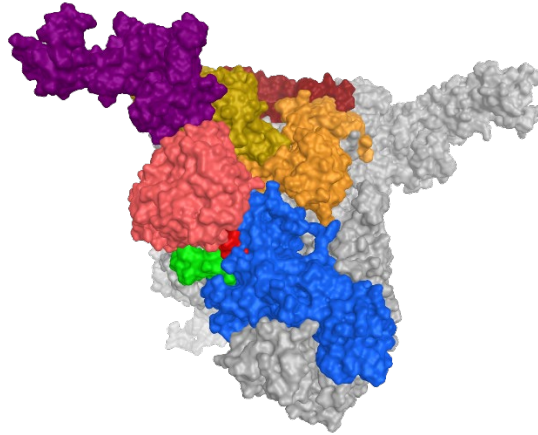


Figura 20. Lugar de acoplamiento favorecido en el caso de la tripsina(rosa), donde se observa la cercanía al péptido de fusión.

Por medio del análisis de interacciones entre proteínas es posible determinar las zonas con mayor interacción en la proteína Spike, se analizó el PLIF únicamente de las poses consideradas, posteriormente se buscaron los aminoácidos presentes en dichas interacciones para cada complejo Spike-receptor y a su vez se dividieron por dominio dichas interacciones.

A continuación, se presentarán dichos resultados a forma de comparación entre las dos conformaciones del dominio CTD, donde cada color corresponde a un dominio de la proteína Spike. En este caso se muestra la tendencia de interacción con aminoácidos de la proteína Spike por cada receptor, así como el cambio de distribución de estas interacciones a la exposición del dominio CTD.

Se presentan a continuación tres proteínas, CD4 debido a que según los resultados de acoplamiento presenta una mejor afinidad a la proteína Spike, CD13 o aminopeptidasa debido a que se consideraba como el receptor del PEDV y la tripsina, ya que no tiene efecto como receptor del PEDV y como se muestra en la figura 17, presenta una tendencia a la interacción con el péptido de fusión, si no como catalizador para la exposición del péptido de fusión.

Dominio Cero ■ Dominio NTD ■ Dominio CTD ■ Dominio SD ■ Dominio S2 ■

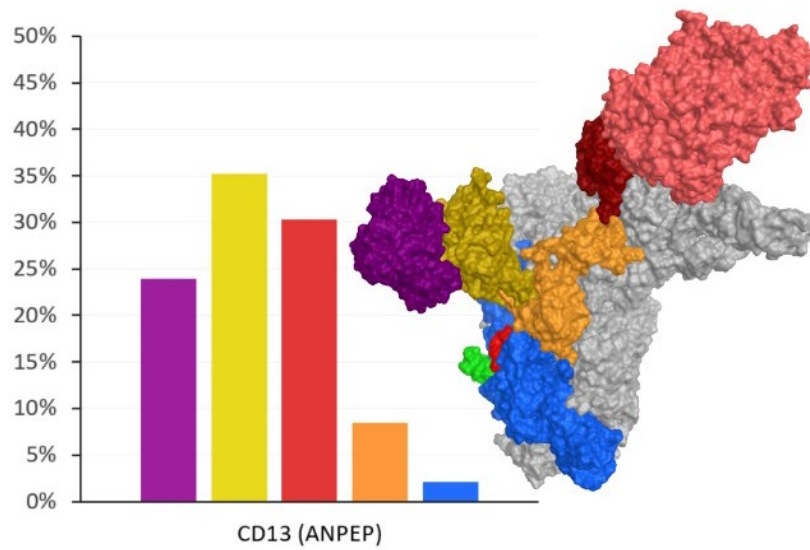
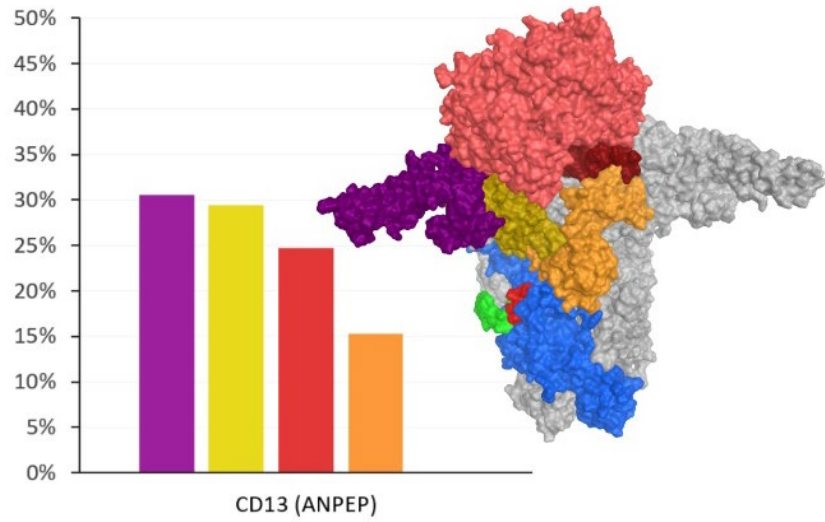


Figura 21. Distribución de interacciones para la proteína CD13 ANPEP (rosa) con la proteína Spike, donde en la parte inferior se muestra la exposición del dominio CTD.

Dominio Cero ■ Dominio NTD ■ Dominio CTD ■ Dominio SD ■ Dominio S2 ■

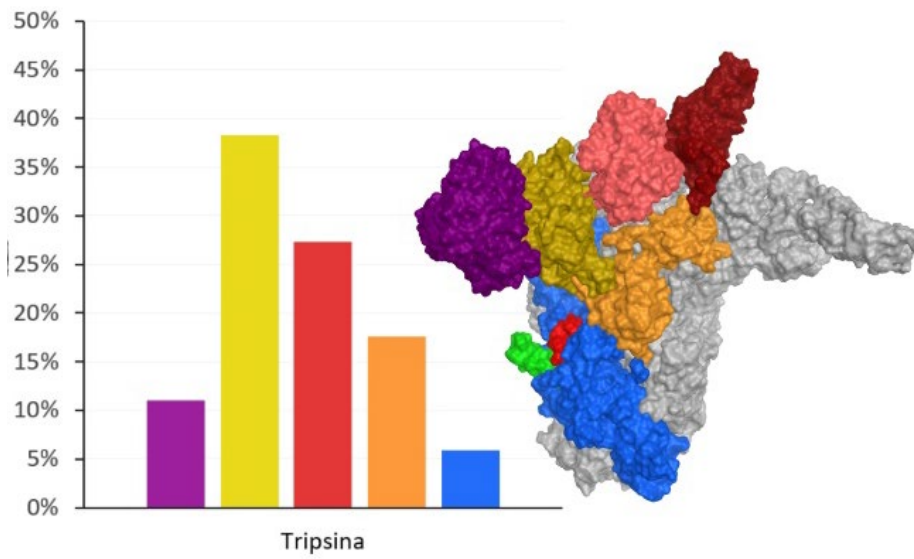
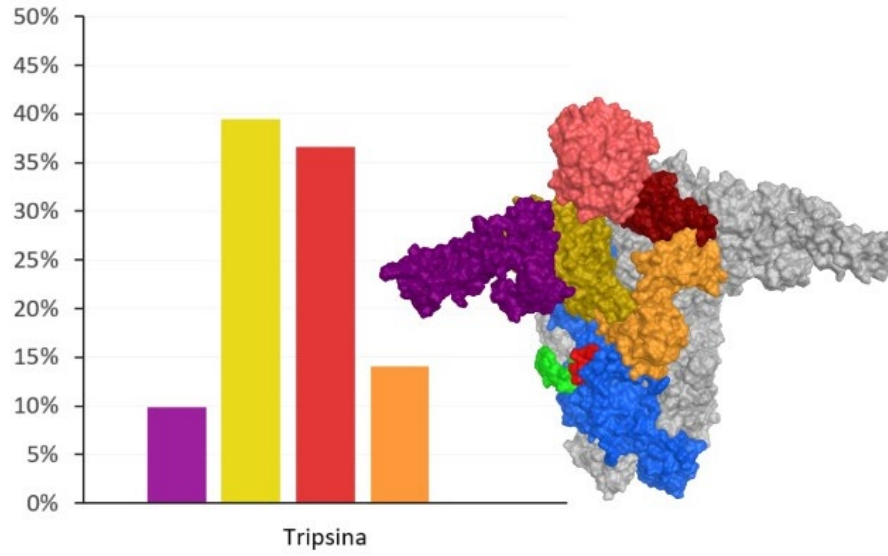


Figura 22. Distribución de interacciones para la proteína Tripsina (rosa) con la proteína Spike, donde en la parte inferior se muestra la exposición del dominio CTD.

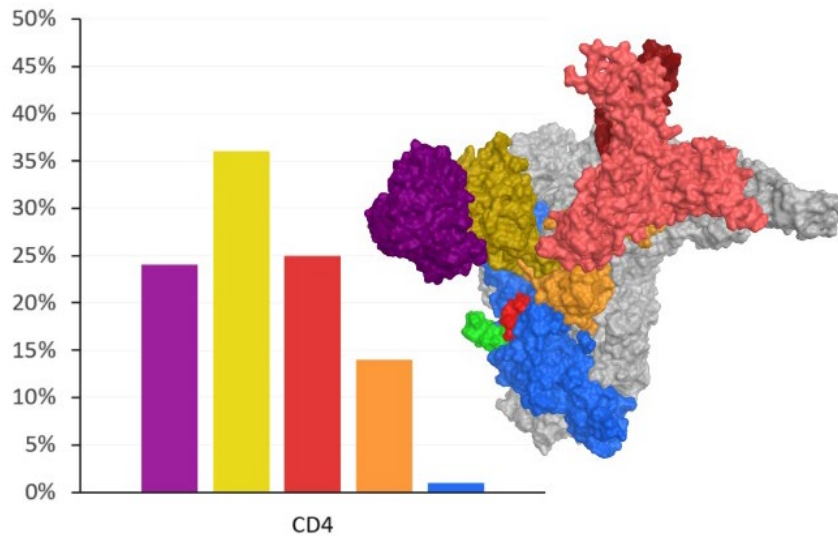
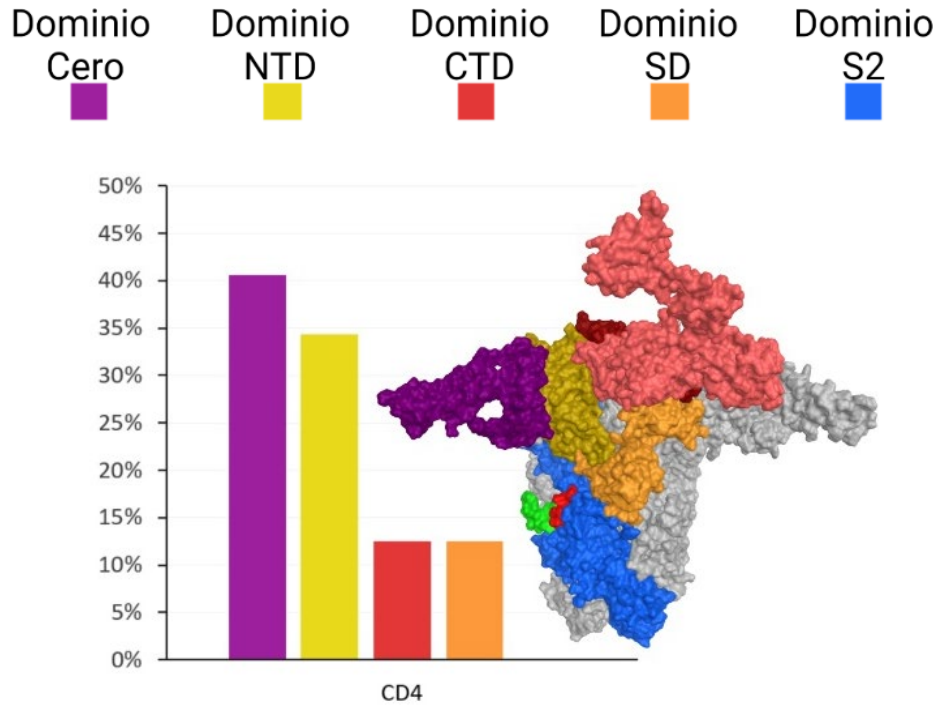


Figura 23. Distribución de interacciones para la proteína CD4 (rosa) con la proteína Spike, donde en la parte inferior se muestra la exposición del dominio CTD.

Como se pueden observar la exposición del dominio CTD ocasiona en la mayoría de los casos, que las interacciones de los receptores con la proteína Spike cambien, lo que se busca es que principalmente exista una buena energía de afinidad entre la

proteína Spike y el receptor, por otro lado que las interacciones aumenten hacia el dominio RBD que se encuentra dentro del dominio CTD de la proteína Spike.

Con ayuda del análisis de interacciones por aminoácido, se buscaron aquellos aminoácidos que más aparecieron en los resultados de acoplamiento, al tomar el dominio CDT y analizar aquellos aminoácidos con más repetición se obtuvo que estos fueron los siguientes:

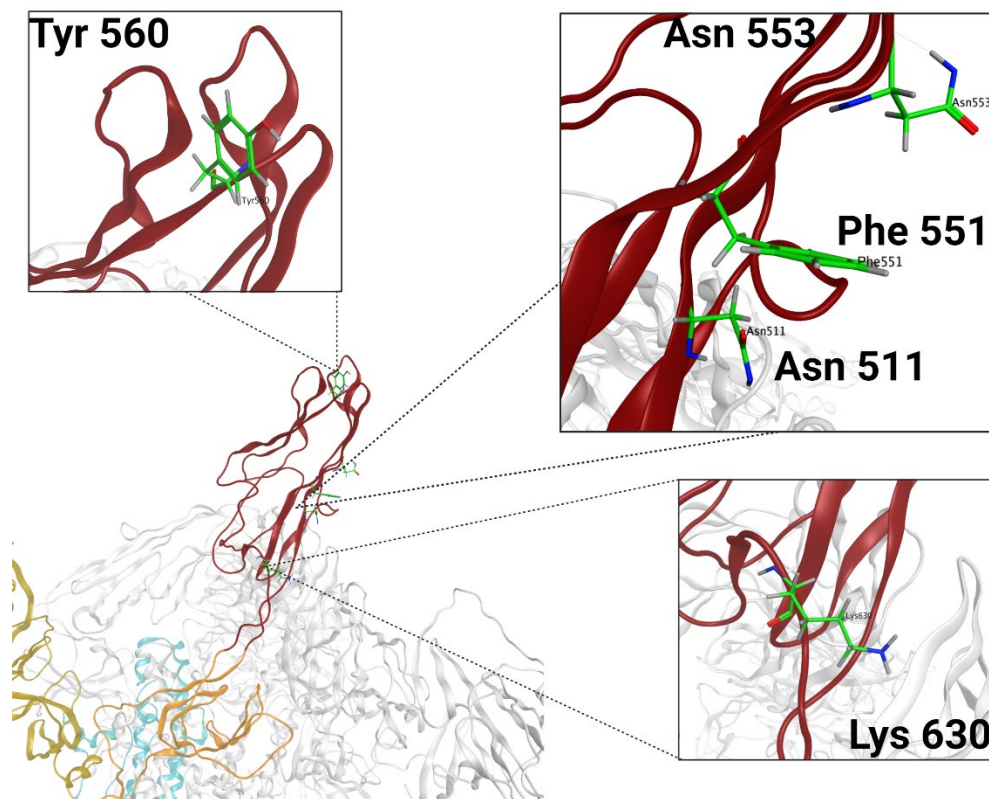


Figura 24. Aminoácidos con más interacciones en el dominio CTD con los receptores estudiados.

Como se puede observar en la Figura 21 los aminoácidos con más interacciones se encuentran dispersos a través del dominio CTD, únicamente los aminoácidos Asn 511, Phe 551 y Asn 553 se encuentran cerca uno de otro, donde se presentan glucosilaciones en los residuos 511 y 553, por otro lado los aminoácidos Tyr 560 y Lys 630 se distribuidos en otras partes del dominio.

9. CONCLUSIONES

Se obtuvieron las secuencias de aminoácidos que componen las distintas proteínas utilizadas así como sus dominios extracelulares.

Gracias a distintos métodos computacionales como la homología o el uso de herramientas basadas en Deep learning como Alphafold, y a partir de una secuencia proteica lineal se logró obtener la estructura tridimensional.

Se encontró por métodos computacionales que la exposición del dominio CTD de la proteína Spike afecta en la energía de afinidad que se presenta entre el complejo Spike-receptor así como la distribución de interacciones entre los dominios de la proteína Spike.

Se encontró que los aminoácidos con más interacciones en el Spike dentro del dominio CTD con todos los receptores fueron: Asn 511, Phe 551, Asn 553, Tyr 560 y Lys 630

Mediante el uso de estudios de acoplamiento molecular, se determinó que ante la exposición del dominio CTD en la proteína Spike el mejor candidato a receptor del PEDV fue el CD4.

10. PERSPECTIVAS

Si bien el estudio muestra resultados energéticos favorables para la unión de la partícula viral con un posible receptor, esto no deja de ser un estudio predictivo, es por esto por lo que existe la necesidad de comprobar dichas predicciones, mediante estudios *invitro* o alguno otro tipo de estudio que permita comprobar los resultados obtenidos.

La búsqueda de receptores virales puede tener un gran avance utilizando técnicas de *Deep learning* para hacer de esto un trabajo más automatizado, esto puede presentar grandes ventajas, ya que ante la presencia de algún nuevo virus, la posibilidad de encontrar aquel receptor que es necesario para la entrada se puede ahorrar largos tiempos de investigación.

Diseñar fármacos que puedan inhibir las interacciones por los aminoácidos más representativos en el dominio CTD con el receptor, y observar su efecto.

11. REFERENCIAS

1. Arias, S. (s/f). *Mecánica molecular*. Departamento de Química Orgánica.- Universidad de Alcalá. . Recuperado el 2 de marzo de 2023, de https://bibliotecavirtual.ranf.com/es/catalogo_imagenes/grupo.do?path=6026700
2. Bateman, A., Martin, M.-J., Orchard, S., Magrane, M., Ahmad, S., Alpi, E., Bowler-Barnett, E. H., Britto, R., Bye-A-Jee, H., Cukura, A., Denny, P., Dogan, T., Ebenezer, T., Fan, J., Garmiri, P., da Costa Gonzales, L. J., Hatton-Ellis, E., Hussein, A., Ignatchenko, A., ... Zhang, J. (2023). UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Research*, *51*(D1), 523–531. <https://doi.org/10.1093/nar/gkac1052>
3. Bella, J., Kolatkar, P. R., Marlor, C. W., Greve, J. M., & Rossmann, M. G. (1998). The structure of the two amino-terminal domains of human ICAM-1 suggests how it functions as a rhinovirus receptor and as an LFA-1 integrin ligand. *Proceedings of the National Academy of Sciences*, *95*(8), 4140–4145. <https://doi.org/10.1073/pnas.95.8.4140>
4. Bella, J., & Rossmann, M. G. (1999). Review: Rhinoviruses and Their ICAM Receptors. *Journal of Structural Biology*, *128*(1), 69–74. <https://doi.org/10.1006/jsbi.1999.4143>
5. Bour, S., Geleziunas, R., & Wainberg, M. A. (1995). The human immunodeficiency virus type 1 (HIV-1) CD4 receptor and its central role in promotion of HIV-1 infection. *Microbiological Reviews*, *59*(1), 63–93. <https://doi.org/10.1128/mr.59.1.63-93.1995>
6. Chemical Computing Group ULC. (2023). *Molecular Operating Environment (MOE)* (2022.02).
7. Chen, Z., Chen, L., Qiao, S.-W., Nagaishi, T., & Blumberg, R. S. (2008). Carcinoembryonic Antigen-Related Cell Adhesion Molecule 1 Inhibits Proximal TCR Signaling by Targeting ZAP-70. *The Journal of Immunology*, *180*(9), 6085–6093. <https://doi.org/10.4049/jimmunol.180.9.6085>
8. Cummins, P. M. (2012). Occludin: One Protein, Many Forms. *Molecular and Cellular Biology*, *32*(2), 242–250. <https://doi.org/10.1128/MCB.06029-11>
9. Davies, P. R. (2015). The dilemma of rare events: Porcine epidemic diarrhea virus in North America. *Preventive Veterinary Medicine*, *122*(1–2), 235–241. <https://doi.org/10.1016/j.prevetmed.2015.08.006>
10. Deng, F., Ye, G., Liu, Q., Navid, M., Zhong, X., Li, Y., Wan, C., Xiao, S., He, Q., Fu, Z., & Peng, G. (2016). Identification and Comparison of Receptor Binding Characteristics of the Spike Protein of Two Porcine Epidemic Diarrhea Virus Strains. *Viruses*, *8*(3), 55. <https://doi.org/10.3390/v8030055>
11. Departamento de Medio Ambiente y Energía. (s/f). *Nanoquímica Medicinal Computacional*. Recuperado el 1 de marzo de 2023, de <https://cimav.edu.mx/investigacion/medio-ambiente-y-energia/area-quimica-computacional/>
12. Diel, D. G., Lawson, S., Okda, F., Singrey, A., Clement, T., Fernandes, M. H. V., Christopher-Hennings, J., & Nelson, E. A. (2016). Porcine epidemic diarrhea virus: An overview of current virological and serological diagnostic methods. *Virus Research*, *226*, 60–70. <https://doi.org/10.1016/j.virusres.2016.05.013>
13. Dörig, R. E., Marcil, A., Chopra, A., & Richardson, C. D. (1993). The human CD46 molecule is a receptor for measles virus (Edmonston strain). *Cell*, *75*(2), 295–305. [https://doi.org/10.1016/0092-8674\(93\)80071-L](https://doi.org/10.1016/0092-8674(93)80071-L)
14. Elsevier Connect. (2021, abril 21). *Virología. Estructura del virión, la partícula vírica*. <https://www.elsevier.com/es-es/connect/medicina/edu-virologia-estructura-del->

28. Li, J., Fu, A., & Zhang, L. (2019). An Overview of Scoring Functions Used for Protein–Ligand Interactions in Molecular Docking. *Interdisciplinary Sciences: Computational Life Sciences*, 11(2), 320–328. <https://doi.org/10.1007/s12539-019-00327-w>
29. Li, Z., Ma, Z., Li, Y., Gao, S., & Xiao, S. (2020). Porcine epidemic diarrhea virus: Molecular mechanisms of attenuation and vaccines. *Microbial Pathogenesis*, 149, 104553. <https://doi.org/10.1016/j.micpath.2020.104553>
30. Lin, F., Zhang, H., Li, L., Yang, Y., Zou, X., Chen, J., & Tang, X. (2022). PEDV: Insights and Advances into Types, Function, Structure, and Receptor Recognition. *Viruses*, 14(8), 1744. <https://doi.org/10.3390/v14081744>
31. Louten, J., & Reynolds, N. (s/f). *Essential human virology*. Academic Press.
32. Lozano-Aponte, J., & Scior, T. (2014). ¿Qué sabe Ud. Acerca de... Dinámica Molecular? *Revista mexicana de ciencias farmacéuticas*, 45(1), 86–88.
33. Luo, X., Guo, L., Zhang, J., Xu, Y., Gu, W., Feng, L., & Wang, Y. (2017). Tight Junction Protein Occludin Is a Porcine Epidemic Diarrhea Virus Entry Factor. *Journal of Virology*, 91(10). <https://doi.org/10.1128/JVI.00202-17>
34. Masre, S. F., Jufri, N. F., Ibrahim, F. W., & Abdul Raub, S. H. (2021). Classical and alternative receptors for SARS-CoV-2 therapeutic strategy. *Reviews in Medical Virology*, 31(5), 1–9. <https://doi.org/10.1002/rmv.2207>
35. Morris, G. M., & Lim-Wilby, M. (2008). *Molecular Docking*. 365–382. https://doi.org/10.1007/978-1-59745-177-2_19
36. Organización de las Naciones Unidas para la Agricultura y la Alimentación (FAO). (2010). *Principales Enfermedades de los Cerdos*. <https://www.fao.org/3/as540s/as540s.pdf>
37. Ortega-Quintana, F. (2018). *Modelado de Bioprocesos: Relaciones Matemáticas entre el Medioambiente y el Material Biológico*. <https://doi.org/10.13140/RG.2.2.33014.09286>
38. Owen, J., & Punt, J. (2016). *KUBY Inmunología* (7e ed.). McGraw Hill.
39. Rahimi, N. (2020). C-type Lectin CD209L/L-SIGN and CD209/DC-SIGN: Cell Adhesion Molecules Turned to Pathogen Recognition Receptors. *Biology*, 10(1), 1. <https://doi.org/10.3390/biology10010001>
40. Rahman, S., Ansari, M., Gaur, P., Ahmad, I., Chakravarty, C., Verma, D., Sharma, A., Chhibber, S., Nehal, N., Wirth, D., & Lal, S. (2021). The Immunomodulatory CEA Cell Adhesion Molecule 6 (CEACAM6/CD66c) Is a Protein Receptor for the Influenza A Virus. *Viruses*, 13(5), 726. <https://doi.org/10.3390/v13050726>
41. Raj, V. S., Mou, H., Smits, S. L., Dekkers, D. H. W., Müller, M. A., Dijkman, R., Muth, D., Demmers, J. A. A., Zaki, A., Fouchier, R. A. M., Thiel, V., Drosten, C., Rottier, P. J. M., Osterhaus, A. D. M. E., Bosch, B. J., & Haagmans, B. L. (2013). Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC. *Nature*, 495(7440), 251–254. <https://doi.org/10.1038/nature12005>
42. Santiago, C., Celma, M. L., Stehle, T., & Casasnovas, J. M. (2010). Structure of the measles virus hemagglutinin bound to the CD46 receptor. *Nature Structural & Molecular Biology*, 17(1), 124–129. <https://doi.org/10.1038/nsmb.1726>
43. Scarpelli, D. G. and B. William. (2023, abril 4). *disease*. Encyclopedia Britannica. <https://www.britannica.com/science/disease>.
44. Schlick, T. (2010). *Molecular Modeling and Simulation: An Interdisciplinary Guide* (A. S.S, Ed.; 2nd ed., Vol. 21). Springer New York. <https://doi.org/10.1007/978-1-4419-6351-2>
45. Sunny, S., & Jayaraj, P. B. (2022). Protein–Protein Docking: Past, Present, and Future. *The Protein Journal*, 41(1), 1–26. <https://doi.org/10.1007/s10930-021-10031-8>

46. Tan, Y., Sun, L., Wang, G., Shi, Y., Dong, W., Fu, Y., Fu, Z., Chen, H., & Peng, G. (2021). Trypsin-Enhanced Infection with Porcine Epidemic Diarrhea Virus Is Determined by the S2 Subunit of the Spike Glycoprotein. *Journal of Virology*, *95*(11). <https://doi.org/10.1128/JVI.02453-20>
47. VanderWaal, K., Perez, A., Torremorrell, M., Morrison, R. M., & Craft, M. (2018). Role of animal movement and indirect contact among farms in transmission of porcine epidemic diarrhea virus. *Epidemics*, *24*, 67–75. <https://doi.org/10.1016/j.epidem.2018.04.001>
48. Wang, N., Shi, X., Jiang, L., Zhang, S., Wang, D., Tong, P., Guo, D., Fu, L., Cui, Y., Liu, X., Arledge, K. C., Chen, Y.-H., Zhang, L., & Wang, X. (2013). Structure of MERS-CoV spike receptor-binding domain complexed with human receptor DPP4. *Cell Research*, *23*(8), 986–993. <https://doi.org/10.1038/cr.2013.92>
49. Wicht, O., Li, W., Willems, L., Meuleman, T. J., Wubbolts, R. W., van Kuppeveld, F. J. M., Rottier, P. J. M., & Bosch, B. J. (2014). Proteolytic Activation of the Porcine Epidemic Diarrhea Coronavirus Spike Fusion Protein by Trypsin in Cell Culture. *Journal of Virology*, *88*(14), 7952–7961. <https://doi.org/10.1128/JVI.00297-14>
50. Zhang, Y., Chen, Y., Zhou, J., Wang, X., Ma, L., Li, J., Yang, L., Yuan, H., Pang, D., & Ouyang, H. (2022). Porcine Epidemic Diarrhea Virus: An Updated Overview of Virus Epidemiology, Virulence Variation Patterns and Virus–Host Interactions. *Viruses*, *14*(11), 2434. <https://doi.org/10.3390/v14112434>
51. Zhou, P., Fan, H., Lan, T., Yang, X.-L., Shi, W.-F., Zhang, W., Zhu, Y., Zhang, Y.-W., Xie, Q.-M., Mani, S., Zheng, X.-S., Li, B., Li, J.-M., Guo, H., Pei, G.-Q., An, X.-P., Chen, J.-W., Zhou, L., Mai, K.-J., ... Ma, J.-Y. (2018). Fatal swine acute diarrhoea syndrome caused by an HKU2-related coronavirus of bat origin. *Nature*, *556*(7700), 255–258. <https://doi.org/10.1038/s41586-018-0010-9>

11. ANEXOS

Resto de proteínas con su distribución de interacciones en los dominios de la proteína Spike, donde en el lado izquierdo se muestran las interacciones con el dominio CTD no expuesto, y en el lado derecho con la exposición del dominio.

