



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE QUÍMICA

**“DISEÑO *IN SILICO* Y CRIBADO VIRTUAL DE ANTAGONISTAS DE
3CLpro COMO POSIBLES TRATAMIENTOS PARA COVID-19”.**

TESIS

**QUE PARA OBTENER EL TÍTULO DE
QUÍMICO FARMACÉUTICO BIÓLOGO**

PRESENTA

BRAYAN RAZIEL CEDILLO GONZÁLEZ

CDMX

2022





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.

JURADO ASIGNADO:

PRESIDENTE: Profesor: Lira Rocha Alfonso Sebastián
VOCAL: Profesor: Pedraza Chaverri José
SECRETARIO: Profesor: Marco Antonio Velasco Velázquez
1er. SUPLENTE: Profesor: Dimitrova Dinkova Tzvetanka
2° SUPLENTE: Profesor: Jiménez Reyes Genaro

SITIO DONDE SE DESARROLLÓ EL TEMA:


Laboratorio de Farmacología Molecular, Torre de Investigación, Facultad de Medicina, UNAM.

Este proyecto fue apoyado por PAPIIT UNAM IV200121. Durante la realización del proyecto, Brayan Raziel Cedillo González recibió un estímulo económico del Sistema Nacional de Investigadores (CONACYT) como Ayudante de Investigador.

ASESOR DEL TEMA:

Dr. Marco Antonio Velasco Velázquez 

SUPERVISOR TÉCNICO:

Dr. Angel Jonathan Ruíz Moreno 

SUSTENTANTE (S):

Brayan Raziel Cedillo González 

DEDICATORIA

Lleno de regocijo, amor, vida y esperanza, dedico este proyecto a cada uno de mis seres queridos, quienes han sido mis pilares para seguir adelante. A mis padres José Francisco Cedillo González y Margarita González Arzaluz por su comprensión, aliento y confianza. A mi pareja favorita, mi hermana querida Vianey Berenice C. G., y su esposo Iván Flores, quienes me brindaron palabras de aliento, apoyo incondicional, deseos de superación, humildad, sacrificio, consejos de vida y vida misma. A María de Lourdes G. R., por siempre estar ahí de forma incondicional como mi confidente, mejor amiga y pareja. Y sin dejar atrás a toda mi familia de sangre y de vida, gracias por confiar en mi y gracias por ser parte de mi vida y por permitirme ser parte de su orgullo.

Muchas gracias a todos.

AGRADECIMIENTOS PERSONALES

Dr. Angel R. Y Dr. Marco, les agradezco de corazón su guía constante, tiempo, conocimiento, dedicación, consejos brindados y por supuesto, les agradezco por brindarme esta enorme oportunidad que me ha abierto nuevos senderos por descubrir en mi vida.

Brayan Raziel Cedillo González.

CONTENIDO

Índice de abreviaturas	3
Índice de figuras, tablas y esquemas	6
Resumen	8
I. Introducción	10
1.1 Epidemiología	10
1.2 Agente Etiológico	11
1.2.1 Genoma viral del SARS-CoV-2	12
1.2.2 Estructura del SARS-CoV-2	15
1.3 Cuadro clínico y patogénesis del SARS-CoV-2	16
1.3.1 Tratamientos disponibles	21
1.4 3CLpro como blanco terapéutico	26
1.4.1 Inhibidores de 3CLpro	30
1.5 Uso de metodologías computacionales para el desarrollo de compuestos anti-SARS-CoV-2	33
1.5.1 Biología estructural.....	33
1.5.2 Descubrimiento, diseño y desarrollo de fármacos	35
1.5.3 Estrategias para identificar compuestos líderes	40
1.5.4 Diseño de fármacos asistido por computadora	42
1.5.4.1 CADD del tipo SB: VS, HTS y acoplamiento molecular	45
1.5.5 Modelo farmacofórico	49
II. Justificación	55
III. Hipótesis	55
IV. Objetivo	55
4.1 Objetivos particulares	56
V. Metodología	56
5.1 Identificación y alineamiento estructural de 3CLpro	56
5.2 Generación de los modelos de farmacofóricos 3D de moléculas cocrystalizadas con 3CLpro	57
5.3 Generación del MFC	57
5.4 Validación del MFC	58
5.4.1 Librería de validación	58
5.4.2 Biblioteca de confórmeros.....	59
5.4.3 Validación del MFC.....	59
5.5 Búsqueda de candidatos	60
5.5.1 Submodelos farmacofóricos	61

5.6	Optimización local y filtrado	61
5.7	Acoplamiento molecular	62
5.8	Validación del Acoplamiento molecular	63
VI.	Resultados.....	64
6.1	Alineamiento de secuencia	65
6.2	Modelo farmacofórico 3D.....	65
6.3	Agrupamiento jerárquico y MFC.....	66
6.4	Validación del MFC	69
6.5	Submodelos Farmacofóricos y búsqueda de potenciales candidatos.....	73
6.6	Optimización local y filtrado	75
6.7	Validación del acoplamiento molecular.....	76
6.8	Acoplamiento molecular	77
VII.	Discusión	81
VIII.	Conclusión	96
IX.	Perspectivas	97
X.	Referencias	98

Índice de abreviaturas

(+) ssRNA	<i>Positive sense single stranded RNA</i>
3CLpro	<i>3-chimiotripsin-like protease</i>
ACE2	<i>Angiotensin-converting enzyme</i>
ADME	Proceso de absorción, distribución, metabolismo y excreción.
ADMET	Proceso de absorción, distribución, metabolismo, excreción y toxicidad.
AH	Aceptor de puente de hidrógeno
ATP	<i>Adenosín trifosfato</i>
CADD	<i>Computer-aided drug design</i>
CCDC	<i>Cambridge Crystallographic Data Center</i>
COFEPRIS	Comisión Federal para la Protección contra Riesgos Sanitarios
COVID-19	<i>Coronavirus disease – 2019</i>
CYP3A	Citocromo P450, familia 3, subfamilia A
DAMP	<i>Damage-associated molecular patterns</i>
DH	Donador de puente de hidrógeno
DNA	<i>Deoxyribonucleic acid</i>
E	Proteína de envoltura
EC ₅₀	<i>Half maximal effective concentration</i>
EK1C4	Lipopéptido derivado de EK1 el cuál es un inhibidor de fusión de pan-coronavirus dirigido al dominio HR1 de la proteína S
EPIC-HR	<i>Evaluation of Protease Inhibition for Covid-19 in High-Risk Patients</i>
ER	<i>Endoplasmic reticulum</i>
FDA	<i>Food and Drug Administration</i>
GP	Glicoproteína-P
gRNA	<i>Guide RNA</i>
GWAS	<i>Whole genome association study</i>
HIV-1	<i>Human immunodeficiency virus type 1</i>
HR2P	<i>Heptad repeat 2 domain peptide</i>
HTS	<i>High throughput screening</i>
Hyd	Hidrofóbico
IC ₅₀	Concentración inhibitoria media
ICTRP	<i>International Clinical Trials Registry Platform</i>
IFN γ	Interferón γ
IL-6	Interleucina 6
IL-1 β	Interleucina 1 β
IP-10	Proteína 10 inducible por interferón gamma
ISG	<i>Interferon-stimulated gene</i>
IUPAC	<i>International Union of Pure and Applied Chemistry</i>

KOL	<i>Key opinion leaders</i>
LB	<i>Ligand-based</i>
LBP	<i>Ligand-based pharmacophore</i>
M	<i>Proteína de membrana</i>
MCP-1	<i>Monocyte chemoattractant protein-1</i>
MERS-CoV	<i>Middle east respiratory syndrome coronavirus</i>
MFC	<i>Modelo farmacofórico consenso</i>
MMFF	<i>Merck molecular force field</i>
MolWT	<i>Molecular weight</i>
Mpro	<i>Main protease</i>
mRNA	<i>Messenger RNA</i>
MT-DTI	<i>Molecule transformer-drug target interaction</i>
N	<i>Nucleocapside</i>
N-finger	<i>Extremo N-terminal del dominio I</i>
NCBI	<i>National Center of Biotechnology Information</i>
nCOVID-19	<i>Novel coronavirus disease 2019</i>
NIH	<i>National Institutes of Health</i>
NMI	<i>Necesidades medicas insatisfechas</i>
NSP	<i>Non-structural proteins</i>
OMS	<i>Organización Mundial de la Salud</i>
ORF	<i>Open reading frames</i>
PAMP	<i>Pathogen-associated molecular patterns</i>
PD	<i>Pharmacodynamic</i>
PDB	<i>Protein Data Bank</i>
pIC ₅₀	<i>Cologaritmo de la concentración inhibitoria media</i>
PK	<i>Pharmacokinetics</i>
PLP	<i>Papain-like protease</i>
POM	<i>Proof of mechanism</i>
Pp1a	<i>Polyprotein 1a</i>
Pp1b	<i>Polyprotein 1b</i>
pro-IL-1 β	<i>Pro-interleucina 1β</i>
PRR	<i>Pattern recognition receptor</i>
QSAR	<i>Quantitative structure-activity relationship</i>
RBD	<i>Receptor-binding domain</i>
RdRp	<i>RNA-dependent RNA polymerase</i>
RMSD	<i>Root-mean-square deviation</i>
RNA	<i>Ribonucleic acid</i>
S	<i>Proteína spike</i>
SAR	<i>Structure-activity relationship</i>
SARS	<i>Severe acute respiratory syndrome</i>
SARS-CoV	<i>Severe acute respiratory syndrome coronavirus</i>
SARS-CoV-2	<i>Severe acute respiratory syndrome coronavirus type 2</i>
SB	<i>Structure-based</i>
SBP	<i>Structure-based pharmacophore</i>

SCA	Síndrome coronario agudo
SIDA	Síndrome de inmunodeficiencia adquirida
sgRNA	<i>Single guide RNA</i>
TCM	<i>Traditional Chinese medicine</i>
Th1	<i>T helper 1</i>
TLR	<i>Toll-like receptor</i>
TMPRSS2	<i>Transmembrane serine protease 2</i>
TNF- α	Factor de necrosis tumoral alfa
VHC	Virus de la hepatitis C
VS	<i>Virtual screening</i>

Índice de figuras, tablas y esquemas¹

Figura 1. <i>Organización genómica del SARS-CoV-2.</i>	15
Figura 2. <i>Procesos de ingreso y desarrollo viral dentro de la célula.</i>	18
Figura 3. <i>Potenciales blancos terapéuticos en SARS-CoV-2.</i>	25
Figura 4. <i>3CLpro de SARS-CoV-2.</i>	27
Figura 5. <i>Esquema 2D general del mecanismo catalítico de la 3CLpro de SARS-CoV-2.</i>	29
Figura 6. <i>Comparación de estructuras cristalizadas de 3CLpro de diferentes coronavirus.</i>	30
Figura 7. <i>Descripción general del proceso de descubrimiento de fármacos.</i>	36
Figura 8. <i>Etapas principales en el desarrollo clásico de fármacos.</i>	38
Figura 9. <i>Representación del proceso del CADD.</i>	45
Figura 10. <i>Alineamiento de secuencia de diversas 3CLpro de SARS-CoV-2.</i>	65
Figura 11. <i>Esquemización de la determinación de modelos farmacofóricos 3D de ligandos cocrystalizados con 3CLpro.</i>	66
Figura 12. <i>Esquema de la generación de centroides Hyd</i>	67
Figura 13. <i>Esquema de la generación de centroides AH.</i>	67
Figura 14. <i>Esquema de la generación de centroides DH.</i>	68
Figura 15. <i>Generación del MFC.</i>	69
Figura 16. <i>Distribución de los ligandos del grupo de validación con respecto a algunos de sus descriptores moleculares.</i>	70
Figura 17. <i>Resultados de validación del MFC.</i>	72
Figura 18. <i>Librerías químicas cotejadas.</i>	73
Figura 19. <i>Validación de acoplamiento molecular.</i>	77
Figura 20. <i>Resultados del acoplamiento molecular.</i>	78

¹ Las abreviaturas ubicadas en cada pie de figura se definen en el apartado **Índice de abreviaturas.**

Figura 21. <i>Mejores candidatos en el contexto de 3CLpro.</i>	80
Tabla 1. <i>Descriptores farmacofóricos y centroides del MFC.</i>	68
Tabla 2. <i>Centroides por submodelo farmacofórico.</i>	74
Tabla 3. <i>Candidatos obtenidos por Submodelo farmacofórico.</i>	74
Tabla 4. <i>Mejores candidatos derivados del acoplamiento molecular.</i> ...	79
Esquema 1. <i>Proceso de validación del MFC.</i>	71

Resumen

Hasta la fecha se han reportado a nivel mundial más de 262 millones de casos confirmados de COVID-19. El agente causal de esta enfermedad es el SARS-CoV-2, el cual apareció a finales del 2019 en Wuhan, China.

Taxonómicamente hablando, el SARS-CoV-2 pertenece a la familia de los Nidovirales, al género Coronaviridae, dentro del grupo de los β - *CoVs*.

Dentro del ciclo replicativo de este coronavirus existen varias proteínas de gran importancia, tal es el caso de la proteasa principal, 3CLpro, la cual se ha determinado como uno de los principales blancos terapéuticos para la generación de candidatos con potencial actividad inhibitoria contra SARS-CoV-2. Hasta el día de hoy se han identificado una amplia variedad de potenciales inhibidores de este virus que interactúan con diferentes blancos terapéuticos, sin embargo, la actividad de estos fármacos para prevenir o reducir la severidad de los síntomas de la COVID-19 no se ha establecido completamente, por lo que prevalece la urgencia del desarrollo e identificación de candidatos farmacológicos adicionales cuyos blancos terapéuticos en SARS-CoV-2 sean eficaces en el tratamiento de la COVID-19.

En el presente estudio se ejecutó el diseño y desarrollo de un MFC, haciendo uso de la amplia información cristalográfica generada durante la actual pandemia mundial.

El MFC logró representar el sistema proteína ligando reportado entre la díada catalítica de 3CLpro (His41 y Cys 145) y una amplia variedad de ligandos cocrystalizados, de forma tal que mediante su empleo en un cribado virtual y tras el uso de un protocolo de acoplamiento molecular, se logró la obtención de 12 moléculas como candidatos con potencial actividad inhibitoria de 3CLpro con miras al desarrollo de un agente terapéutico contra la COVID-19.

Palabras clave: SARS-CoV-2, COVID-19, farmacóforo, modelo farmacofórico consenso, inhibidores, 3CLpro.

I. Introducción

1.1 Epidemiología

Desde finales del 2019, un nuevo coronavirus hallado en Wuhan, provincia de Hubei, China, ha sido el agente etiológico responsable del brote de neumonía viral 2019-2020 por coronavirus, enfermedad conocida como COVID-19 (del inglés, *Coronavirus disease-2019*) [1-4]. Posteriormente, el 11 de marzo de 2020 la Organización Mundial de la Salud (OMS) declaró la ocurrencia de la pandemia de COVID-19, exhortando a la población mundial a tomar las medidas de higiene y distanciamiento necesarias, así como a conjuntar esfuerzos de control en la mayor emergencia en salud pública mundial de los tiempos modernos [5].

Según la OMS, para el 2 de diciembre del año 2021 la COVID-19 acumuló un total de 262.1 millones de caso confirmados y 5.2 millones de muertes alrededor del mundo, donde México aportó el 1.5% y el 5.7% de las cifras totales, respectivamente [6].

Cabe destacar que la enfermedad presenta un peor pronóstico para la población adulta de entre 30 – 79 años con una letalidad global del 2.3% (CFR) [7]. Del mismo modo, el 34º Informe Epidemiológico de la Situación de COVID-19 del gobierno de México reportó para el 25 de octubre de 2021, que la letalidad para adultos mayores de 60 años a nivel nacional fue del 29%.

1.2 Agente Etiológico

El virus del síndrome respiratorio agudo severo tipo 2 (*SARS-CoV-2*, por sus siglas en inglés) es el causante de la enfermedad COVID-19. Taxonómicamente, es un miembro del género *Betacoronavirus*, perteneciente a la subfamilia *Coronavirinae*, en la familia *Coronaviridae* dentro del orden de los Nidovirales [8-10]. Los coronavirus, llamados así por su cápsula lipoproteica de forma esférica rodeada de múltiples espículas (proteína Spike o glicoproteínas-S) que le dan aspecto de corona [11], suelen causar enfermedades respiratorias leves en humanos y animales, además se estima que producen entre el 10% y el 30% de los casos de resfriado común [5]. A pesar de lo anterior, la subfamilia *Coronaviridae*, constituida por 4 géneros (Alpha, Beta, Gamma y Delta-coronavirus) [9], presenta otras especies, principalmente dentro del género beta, que anteriormente han generado epidemias con grandes estragos en la humanidad. Tal es el caso del coronavirus responsable del síndrome agudo severo (*SARS-CoV*) en 2002-2003, con una tasa de letalidad del 9.6%, y del coronavirus responsable del síndrome agudo respiratorio del medio este (*MERS-CoV*) en 2012, con una tasa de letalidad de 34.3% [7,10].

1.2.1 Genoma viral del SARS-CoV-2

En cuanto al genoma de SARS-CoV-2, éste se encuentra constituido por ácido ribonucleico (RNA, por sus siglas en inglés) de cadena sencilla con sentido positivo [(+)ssRNA] de aproximadamente 30 kb [4] (Figura 1).

Los coronavirus cuentan con dos características particulares dentro de todos los virus de RNA: 1) cuentan con el genoma más largo y 2) son uno de los pocos virus con un mecanismo de revisión genómica que evita que acumulen mutaciones que podrían debilitarlo [12].

Cuentan con 15 marcos de lectura abiertos (ORFs, por sus siglas en inglés), capaces de formar hasta 28 proteínas, entre las cuáles encontramos 4 proteínas estructurales: la proteína Spike (S), la proteína de envoltura (E), la proteína de membrana (M) y la proteína de la nucleocápside (N); además de 16 proteínas no estructurales (NSPs, por sus siglas en inglés) y proteínas accesorias [13].

El genoma viral está listo para ser traducido por las poliproteínas replicasas, una vez dentro del hospedero, debido a que cuenta con un extremo 5' no codificante con un cap, conformado por 7-metilguanosina, y un extremo 3' con una cola de poli A, características que le permiten actuar como RNA mensajero (mRNA) [5,14].

Al enfocarnos en su RNA, las dos terceras partes codificantes en dirección al extremo 5' se encuentran ocupadas por los ORFs 1a y 1ab, siendo este último el gen más largo en SARS-CoV-2 [15]. Ambos ORFs codifican para

proteínas largas que mediante proteólisis dan lugar a una amplia variedad de NSPs de tamaño variable [15]. El ORF1ab codifica para la proteína pp1ab y 15 NSPs, mientras que el ORF1a codifica para la proteína pp1a y contiene 10 NSPs [16–18]. Dentro de las mencionadas NSPs, destacan: una RNA polimerasa dependiente de RNA (RdRp, por sus siglas en inglés), una Helicasa y dos proteasas de vital importancia, la proteasa principal (Mpro por sus siglas en inglés o 3CLpro del inglés *3-chimotripsinchymotrypsin-like protease*) y la proteasa similar a papaína (PLP por sus siglas en inglés). Estas últimas se encargan de escindir las poliproteínas en sus fragmentos funcionales [19]. La otra tercera parte del genoma hacia el extremo 3', contiene los ORFs correspondientes a las proteínas estructurales N, M, E y S, además de otras nueve proteínas pequeñas de función desconocida [5].

Entre los dos ORFs, ORF1a y ORF1ab, hay una región que mediante la replicasa de SARS-CoV-2 codifica para dos poliproteínas superpuestas – pp1a y pp1ab – las cuales son requeridas para transcripción y replicación viral. Los polipéptidos funcionales son liberados de las poliproteínas mediante un proceso proteolítico que tiene lugar gracias a PLP y, predominantemente, gracias a 3CLpro. Por otra parte, ambas proteasas bloquean la respuesta inmune innata del huésped. La 3CLpro escinde a pp1a y pp1ab en al menos 11 sitios conservados [20,21] (ver Figura 1).

Las NSPs cumplen un papel importante en los procesos de replicación, transcripción y recombinación viral durante el proceso infeccioso [10,22]. Por ejemplo, la NSP 1 es usada por el virus para evadir la respuesta inmunitaria del hospedero [23]. La NSP 2 es indispensable para la replicación viral y su función sigue sin ser muy clara, además las NSP 2 y 3 interactúan para formar la proteasa que escinde el ORF1ab [24,25]. Por otra parte, las NSP 3 y 4 interactúan con otros cofactores para inducir un rearrreglo membranal para el mecanismo de replicación viral y la pérdida de este complejo elimina este proceso [26]. La NSP 5 (3CLpro) es una proteasa de cisteína encargada de la escisión de 11 sitios en las poliproteínas para liberar NSP 4–NSP 16 durante la replicación [27]. La NSP 6 genera autofagosomas en el retículo endoplásmico y está envuelto en la autofagia [28]. La NSP 12 en complejo con las NSP 7 y 8 forman la maquinaria de replicación [29]. La NSP 9 en complejo con la NSP 8 están envueltas en la RNA replicasa y en la virulencia [30]. El complejo entre las NSP 10-NSP 16 es esencial para generar el 7-metilguanilato o cap en el mRNA viral para una eficiente traducción y ayuda a evadir la respuesta inmunitaria [31]. La NSP 14 en complejo con la NSP 10 están involucradas en la actividad de exonucleasa [32]. La NSP 13 tienen actividad RNA TPasa [33] y la NSP 14 tiene actividad exoribonucleasa [33]. Finalmente, las NSP 11 y 15 tienen actividad endorribonucleasa [22].

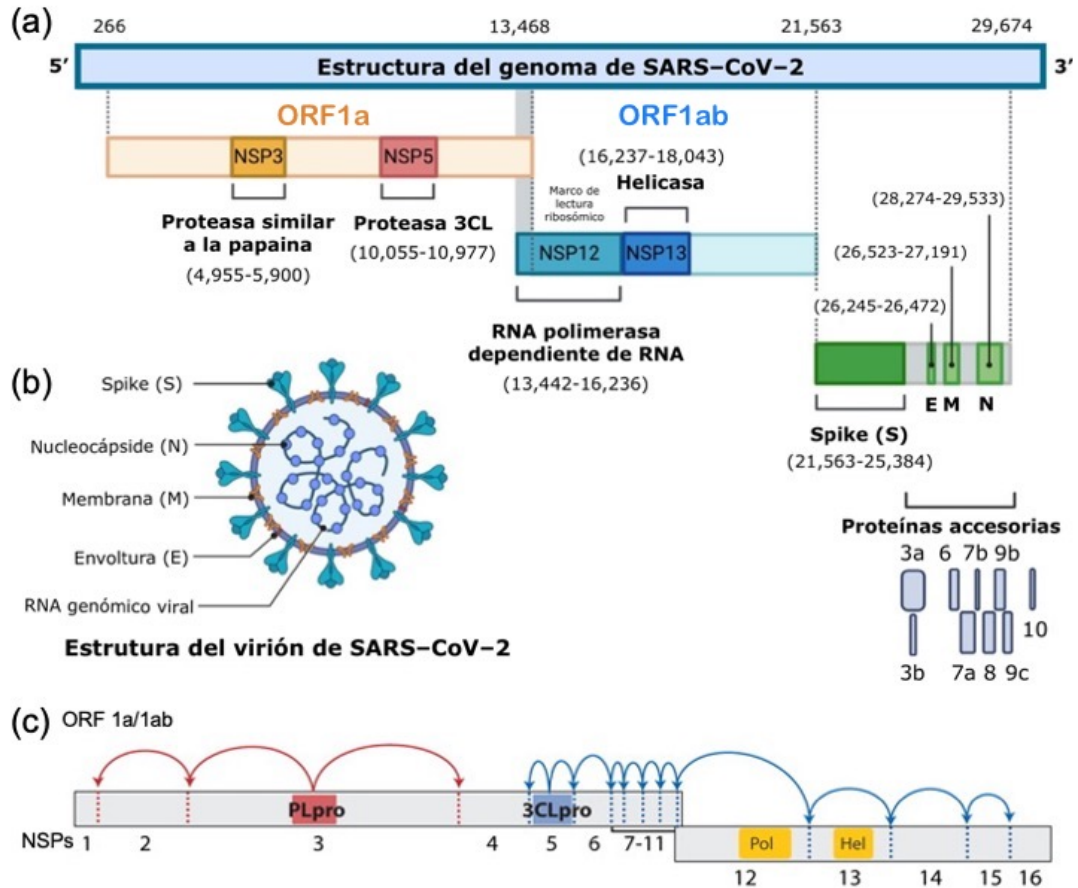


Figura 1. Organización genómica del SARS-CoV-2.

(1a) El genoma de SARS-CoV-2 codifica dos genes, ORF1a (naranja) y ORF1ab (azul). Estos codifican 16 NSPs. Las NSP 3 y NSP 5 codifican para la PLP y para la 3CL-pro, respectivamente. La NSP 12 codifica para la RdRp. La NSP 13 codifica para la RNA helicasa. Los genes que codifican las proteínas estructurales: S, E, M y N, destacan en verde. Las proteínas accesorias (gris) son únicas para SARS-CoV-2 en términos numéricos, organización genómica, secuencia y función. **(1b)** Representación estructural de una partícula viral de SARS-CoV-2. **(1c)** Sitios de corte de 3CLpro y PLpro sobre el ORF 1a y 1ab de SARS-CoV-2. Modificado de [34,35].

1.2.2 Estructura del SARS-CoV-2

Las principales características de los coronavirus son que poseen forma esférica pleomórfica, con un diámetro aproximado de 125 nm [15]. Estructuralmente está constituido por 4 proteínas estructurales incluyendo la proteína N y S, así como la glicoproteína M y E [36]; éstas

últimas tres se encuentran ancladas en la envoltura lipídica. La proteína S genera homotrómeros que se posan en la superficie del virus y le permiten el ingreso a las células hospederas mediante la identificación de sus receptores, la enzima convertidora de angiotensina II (ACE2, por sus siglas en inglés) [2] presente en las células del tracto respiratorio superior. Además, le brinda a la partícula viral la apariencia de una corona. Por otra parte, SARS-CoV-2 muestra una extensiva glucosilación en la proteína S de la superficie viral, la cual puede influenciar diversos aspectos, tales como la estabilidad de los componentes proteicos, tropismo celular, reconocimiento por mecanismos inmunitarios, y camuflaje de antígenos reconocidos por anticuerpos neutralizantes [37]. Asimismo, es el principal inductor de estos últimos [38].

La proteína N es responsable del empaquetamiento del genoma viral y la proteína M permite el ensamble de las partículas vírica [15]. Esta última determina la forma de la envoltura viral, estabiliza la proteína N y el complejo N-RNA permitiendo el ensamble completo del virus.

Finalmente, la proteína E juega un papel importante en la maduración y producción de partículas virales [5,36].

1.3 Cuadro clínico y patogénesis del SARS-CoV-2

La infección de COVID-19, inicialmente llamada *nCOVID-19* (del inglés, *novel coronavirus disease 2019*), es principalmente transmitida de

humano a humano por contacto mediante gotas de *flugge* (secreciones respiratorias $>5\mu\text{m}$) y en ocasiones también por aerosoles. Se ha reportado que puede permanecer en superficies de plástico o metal de 24 hasta 72 horas [39]. Presenta un alto índice de contagio con un número de reproducibilidad (R_0)=5, lo que indica que una persona contagiada puede infectar de 2 a 5 personas, explicando el crecimiento exponencial del número de casos [11,40]. El periodo de incubación es de 1 a 14 días con una media de entre 5 y 6 días [11].

La patogénesis de la infección por SARS-CoV-2 en humanos se manifiesta con diversos síntomas que pueden ser desde leves hasta severas insuficiencias respiratorias. Una vez que este virus se instala en células epiteliales en el tracto respiratorio, se inicia la replicación y migración a las vías respiratorias, para finalmente ingresar a las células epiteliales alveolares en los pulmones. Este proceso inicia con la proteína S y su reconocimiento por su receptor, lo cual se ha debido a un dominio de unión al receptor (RBD, por sus siglas en inglés) que media el contacto directo con la ACE2 y con un sitio de escisión polibásico S1/S2 que se escinde proteolíticamente por la catepsina L celular y la proteasa transmembranal de serina (TMPRSS2, por sus siglas en inglés). La TMPRSS2 facilita la entrada viral en la superficie de la membrana plasmática, mientras que la catepsina L activa a la proteína S en los endosomas. Una vez que el genoma viral se libera en el citosol del

huésped, los ORF1a y ORF1ab se traducen en replicasas virales, las cuales son escindidas en NSPs individuales (a través de proteasas virales – 3CLpro y P1pro – del huésped); estas replicasas forman la RdRp. A continuación, los componentes de la replicasa reorganizan el retículo endoplásmico en vesículas de doble membrana que facilitan la replicación viral de los RNA genómicos y sub genómicos; estos últimos se traducen en proteínas estructurales y accesorias que facilitan la formación de partículas virales [41]. El proceso de ingreso y desarrollo viral dentro de la célula se resume en la Figura 2.

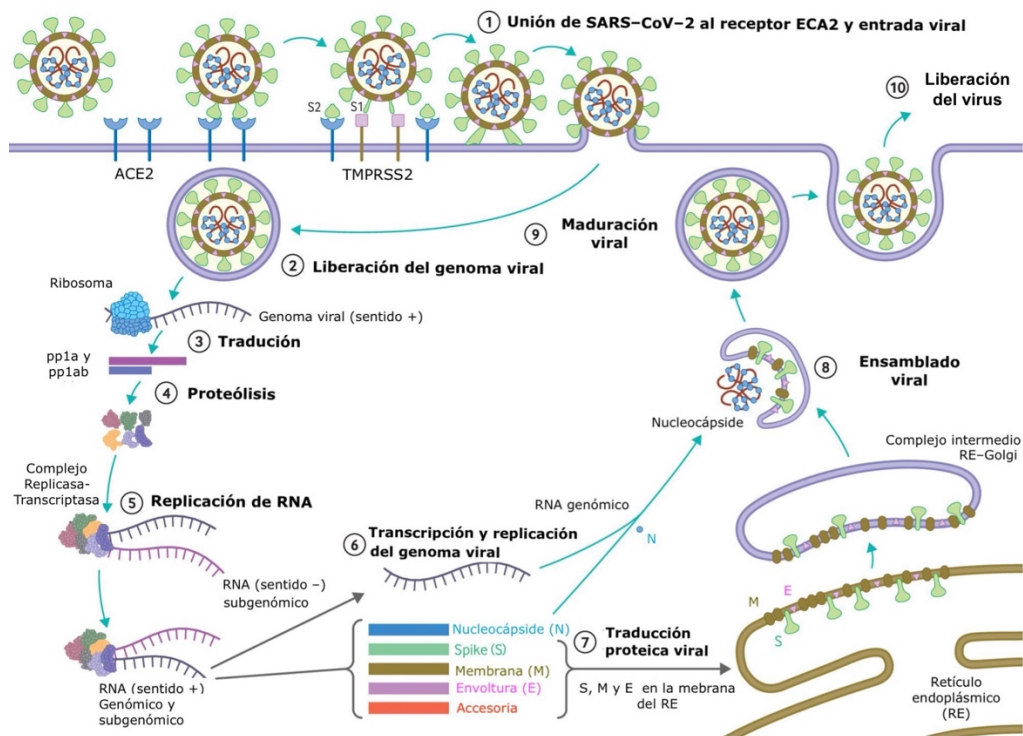


Figura 2. Procesos de ingreso y desarrollo viral dentro de la célula. (1) El virus ingresa a su blanco celular hospedero tras su unión al receptor ACE2 en sinergia con la TMPRSS2 del huésped. Esto conduce a la fusión de membranas y liberación del material genómico dentro del citoplasma del hospedero (2). Las etapas (3-7) muestran los pasos restantes de la replicación viral, que conducen al ensamblado, maduración y liberación viral. Las etapas muestran el ensamblado viral, maduración y liberación, respectivamente. Modificado de [42].

Su rápida replicación dentro del pulmón puede desencadenar una fuerte respuesta inmune. La tormenta de citosinas origina el síndrome respiratorio agudo severo (SARS, por sus siglas en inglés) e insuficiencia respiratoria, esta última se considera la principal causa de muerte por COVID-19 [43,44]. Además, múltiples grupos han reportado falla múltiple de órganos en algunos casos de COVID-19 [7,45,46].

Dentro de los signos y síntomas más comunes destacan: fiebre, fatiga y tos seca [44,45,47,48]. Los menos comunes incluyen producción de esputo, dolor de cabeza, hemoptisis, diarrea, anorexia, dolor de garganta, dolor de pecho, escalofríos, náuseas, vomito, dificultad para respirar (disnea), linfocitopenia, trombocitopenia, un aumento del dímero-D y trastornos del olfato y del gusto. Síntomas que ocasionalmente pueden derivar a complicaciones tales como neumonía, SARS e inclusive la muerte [14].

De forma general, la distribución de los receptores ACE2 en diferentes tejidos pueden ayudar a explicar los sitios de infección y los síntomas del paciente (generalmente multifactoriales) [42], por ejemplo, el receptor ACE2 se encuentra en el epitelio de órganos como el intestino y las células endoteliales del riñón y los vasos sanguíneos, lo que puede explicar los síntomas gastrointestinales [49], las complicaciones cardiovasculares (inflamación vascular, miocarditis, arritmias cardíacas [44] y el Síndrome

Coronario Agudo (SCA) [50]), trombocitopenia [51], hipercoagulabilidad [52], lesiones hepáticas [53] o lesiones renales [54]. Se ha propuesto que, fuera de la distribución y unión al receptor ACE2, estas manifestaciones están relacionadas con la inflamación inducida por citosinas, el daño citotóxico directo del virus y otras comorbilidades [55]. De forma particular, la patogenia de la neumonía inducida por SARS-CoV-2 se explica mejor en dos etapas, una fase temprana y una tardía. La primera se caracteriza por la replicación viral que resulta en daño tisular directo mediado por virus, que a su vez puede causar altos niveles de piroptosis, seguida de la fase tardía donde las células huésped infectadas desencadenan una respuesta inmune con el reclutamiento de linfocitos T, monocitos y neutrófilos. Utilizando una amplia variedad de receptores de reconocimiento de patrones (PRRs, por sus siglas en inglés), las células epiteliales alveolares y los macrófagos alveolares detectan los patrones moleculares asociado a patógenos (PAMP, por sus siglas en inglés) liberados, como el RNA viral, y los patrones moleculares asociados a daño (DAMP, por sus siglas en inglés), incluido el adenosín trifosfato (ATP) y el ácido desoxirribonucleico (DNA, por sus siglas en inglés). Lo anterior da origen a una ola de inflamación local, lo cual implica un aumento en las secreciones de citosinas proinflamatorias y quimiocinas, tales como la interleucina 6 (IL-6), el interferón γ (IFN γ), la proteína quimioatrayente de monocitos 1 (MCP-1) y la proteína 10 inducible por interferón gama

(IP-10) en la sangre de pacientes infectados. Estas citosinas, por un lado son indicadores de una respuesta de células T helper 1 (Th1); mientras que por el otro lado, atrae a las células inmunitarias, en particular a los monocitos y linfocitos T, de la sangre al sitio de infección con el fin de eliminar a las células infectadas [56].

En la COVID-19 severa, la sobre activación del sistema inmunológico da como resultado una tormenta de citosinas caracterizada por la liberación de altos niveles de citosinas (especialmente IL-6 y el factor de necrosis tumoral α o TNF- α) en circulación, lo que provoca una respuesta inflamatoria local y sistémica [57,58]. Además la unión de SARS-CoV-2 al receptor TLR (del inglés *Toll-Like Receptor*) induce la liberación de pro-interleucina 1 β (pro-IL-1 β), que se escinde en la interleucina 1 β (IL-1 β) madura y activa quien media la inflamación pulmonar hasta la fibrosis [55,59].

1.3.1 Tratamientos disponibles

Desde el inicio de la pandemia hasta el 3 de noviembre del 2021 habían cerca de 12,146 tipos de ensayos clínicos sobre la COVID-19 alrededor del mundo registrados en la plataforma Internacional de Registros de Ensayos Clínicos (*ICTRP*, por sus siglas en inglés) de la OMS [60]. Estos ensayos se enfocan en la evaluación de estrategias profilácticas, terapéuticas, o rehabilitadoras.

Respecto al desarrollo de fármacos para tratar la COVID-19, la Administración de Fármacos y Alimentos de Estados Unidos (*FDA*, por sus siglas en inglés) únicamente ha aprobado al medicamento antiviral Remdesivir (Veklury), un nucleótido análogo inhibidor de la RdRp del virus del Ébola [61], para su uso en pacientes adultos y pediátricos de 12 años o mayores, que pesen al menos 40 Kg y que requieran hospitalización [62]. Dicha aprobación requirió evidencia sustancial científica sobre su efectividad y una demostración de seguridad para su uso en humanos [62].

Por otra parte, la misma organización ha autorizado otros tratamientos farmacológicos para su uso de urgencia durante esta emergencia de salud pública moderna [63]. El pasado diciembre del 2021 se autorizó el uso de dos antivirales orales para el tratamiento de esta enfermedad, en ciertos pacientes. Estos son el Paxlovid de Pfizer [64] y el Molnupiravir de Merck [65]. El Paxlovid se compone por Nirmatrelvir, un inhibidor de 3CLpro, y Ritonavir, un inhibidor de la proteasa del virus de inmunodeficiencia humana tipo 1 (HIV-1, por sus siglas en inglés) y del citocromo P450, familia 3, subfamilia A (CYP3A) [66]. Este está autorizado para el tratamiento de COVID-19 de leve a moderado en pacientes adultos y pediátricos (mayores de 12 años y con un peso de al menos 40 Kg) con resultado positivo en la prueba directa de detección viral de SARS-CoV-2. El Molnupiravir, un profármaco de isopropil éster del análogo

nucleósido β - D - N^4 hidroxicitidina el cual es un inductor en el aumento de mutagénesis en el RNA viral por la RdRp, impidiendo de esta forma la replicación de SARS-CoV-2. Éste está autorizado para el tratamiento de COVID-19 de leve a moderado en pacientes con resultado positivo en la prueba directa de detección viral de SARS-CoV-2 [67].

Algunos procesos de inhibición viral, potenciales blancos terapéuticos en SARS-CoV-2 y principales inhibidores químicos se encuentran ilustrados en la Figura 3, de los que destacan:

- a. Anticuerpos monoclonales. Bloquean la unión de la proteína S al receptor ACE2 [68-70].
- b. Favipiravir. Un nucleósido análogo que es conocido como fármaco antiviral de amplio espectro [71].
- c. Danoprevir. Es un inhibidor de la proteasa NS3 del virus de la hepatitis C (VHC) aprobado en China para el tratamiento del genotipo 1b no cirrótico crónico de Hepatitis C [72].
- d. Darunavir. Es un inhibidor de la proteasa de HIV [73].
- e. Lopinavir y Ritonavir. Usados en conjunto para el tratamiento del Síndrome de Inmunodeficiencia Adquirida (SIDA). Ambos son inhibidores de la proteasa HIV. El Ritonavir es también un inhibidor del citocromo P450 y de la glicoproteína-P (GP) [74].

- f. Oseltamivir. Fármaco aprobado para el tratamiento de influenza A y B, debido a que inhibe la neuraminidasa viral y, en consecuencia, bloquea la liberación de las partículas virales [75].
- g. Umifenovir, también conocido como Arbidol. Fármaco aprobado en Rusia y China para el tratamiento de la infección por el virus de la influenza A y B. Está patentado su uso para el tratamiento de SARS impidiendo la interacción de la proteína S y ACE2 [76].

De cualquier forma, la actividad de estos fármacos para prevenir o reducir la severidad de los síntomas de la COVID-19 no ha sido completamente establecida [77]. Debido a la urgente necesidad de un tratamiento de prevención y control de esta enfermedad, es necesario el desarrollo e identificación de candidatos farmacológicos adicionales cuyos blancos terapéuticos en SARS-CoV-2 sean eficaces en el tratamiento de la COVID-19 [78].

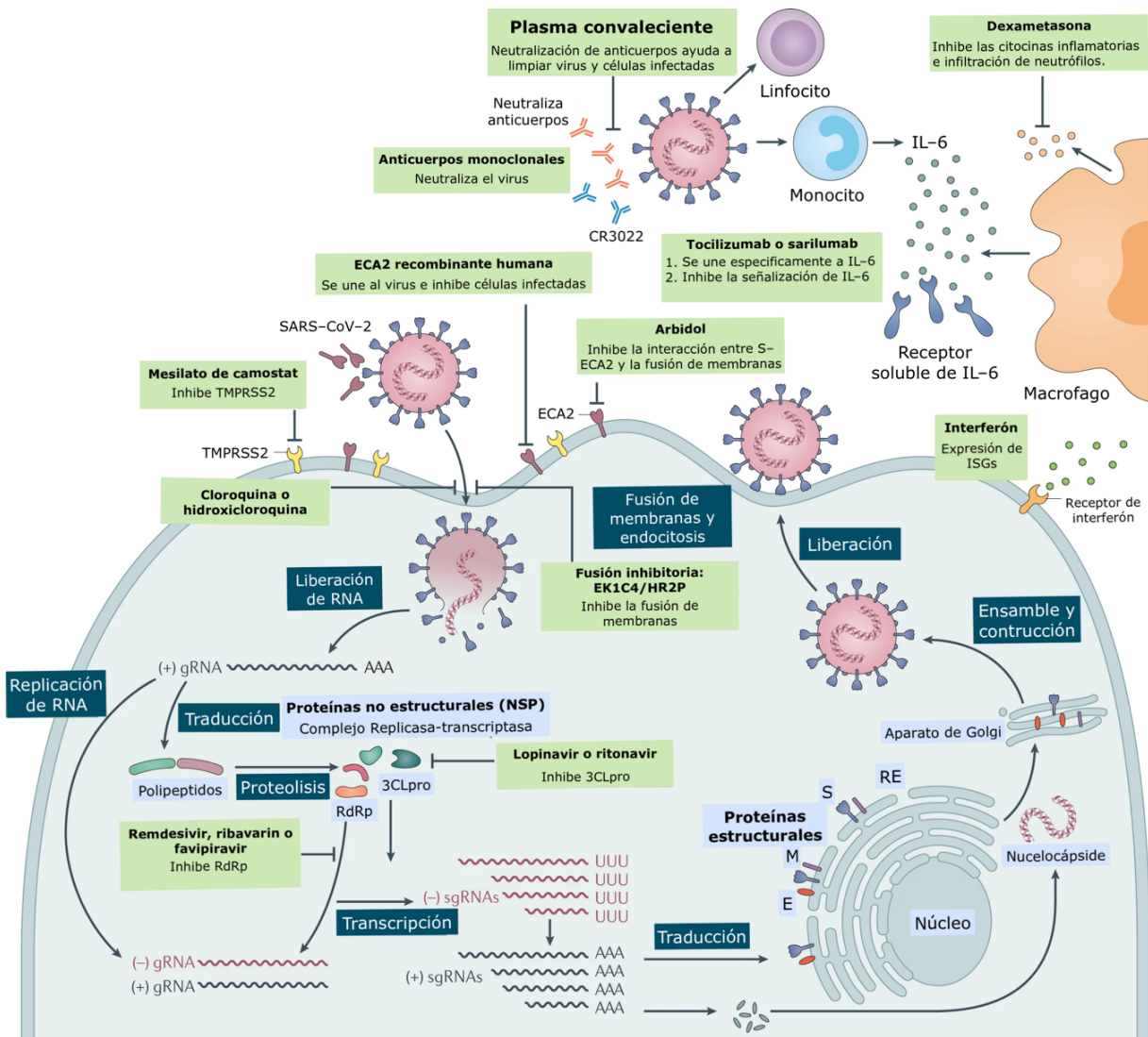


Figura 3. Potenciales blancos terapéuticos en SARS-CoV-2.

Blancos antivirales dirigidos hacia los diferentes pasos de replicación del SARS-CoV-2, que va desde la unión al receptor, entrada y la fusión para la replicación. Cabe destacar que hasta la fecha la robustez de los datos de la eficiencia clínica de la mayoría de estos tratamientos es escasa. Anticuerpo monoclonal humano anti-glicoproteína S1 de SARS-CoV-2 (conocido como CR3022); lipopéptido derivado de EK1 el cuál es un inhibidor de fusión de pan-coronavirus dirigido al dominio HR1 de la proteína S (EK1C4); retículo endoplásmico (ER); RNA guía (gRNA); peptidos derivados de la repetición 2 de la heptada de la proteína S de SARS-CoV-2 (HR2P); gen estimulador de interferón (ISG); RNA guía único (sgRNA). Modificado de [79].

1.4 3CLpro como blanco terapéutico

La 3CLpro es una proteasa de cisteína. Su forma nativa es un homodímero (Figura 4e), donde cada monómero del ensamble biológico se compone de dos dominios del tipo barriles- β -antiparalelos de 5 hebras con un dominio plegable similar a la quimiotripsina (el dominio I engloba los residuos Phe8-Tyr101; el dominio II engloba los residuos Lys102-Pro184) y un dominio de 5 α -hélices antiparalelas (dominio III, engloba los residuos Thr201-Thr303) [26] (Figura 4).

Los dominios II y III se conectan mediante un *loop* largo (181-199). Cabe destacar que el tercer dominio es importante para la dimerización y formación de una 3CLpro activa [80]. Los residuos 1-7 en el extremo N-terminal del dominio I (*N-finger*) cuentan con un rol importante en su actividad proteolítica, además ayudan a dar forma al *pocket* S1 del sitio de unión a sustrato [81,82].

El sitio activo se encuentra formado por la diada catalítica His41-Cys145, aminoácidos que forman el "agujero de oxoanión" junto con la cadena principal de amidas de Gly143 y Ser144. Esta diada catalítica se encuentra altamente conservada entre las diferentes 3CLpro de coronavirus [26]. La His41 y la Cys145 forman parte de los dominios I y II, respectivamente, y se encuentran separados por 3.6 Å, distancia óptima necesaria para la formación de interacciones del tipo puente de hidrógeno [10], la

importancia de este punto será discutida más adelante en el proceso enzimático de 3CLpro.

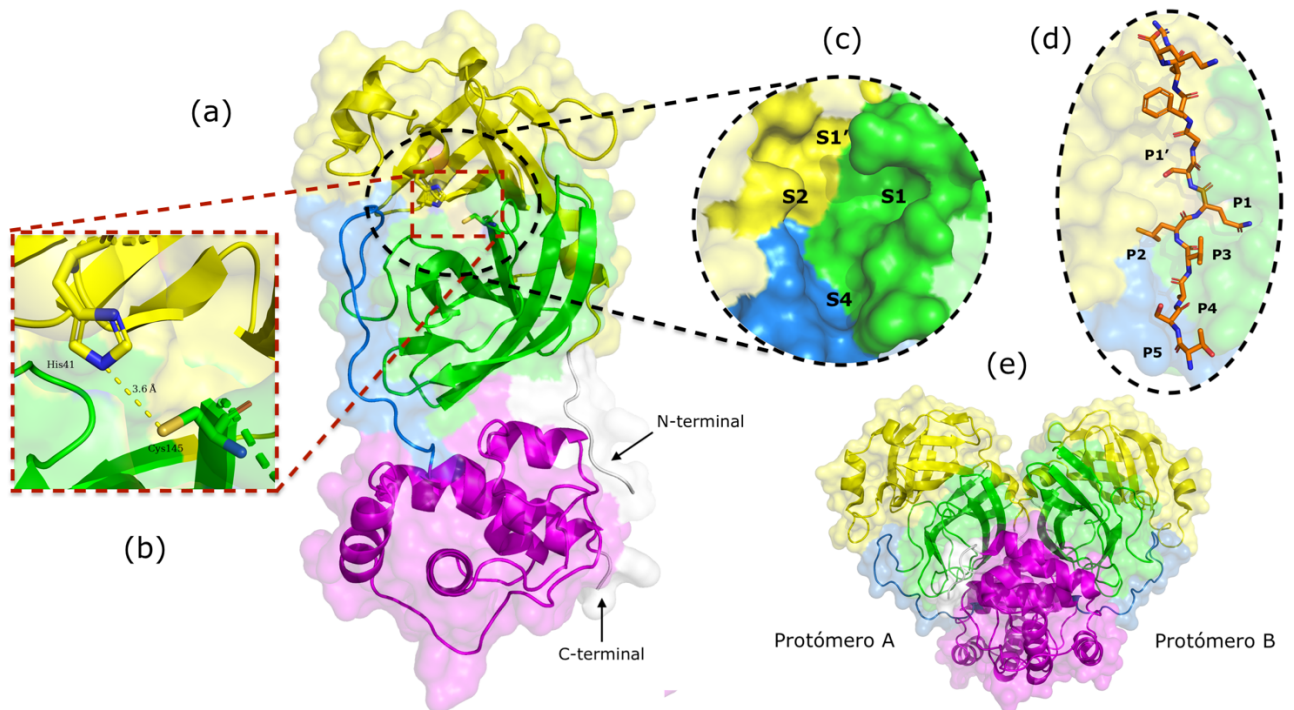


Figura 4. 3CLpro de SARS-CoV-2.

(4a) Representación en caricatura de los dominios estructurales de 3CLpro de SARS-CoV-2 (código PDB 6M2N). Los dominios I y II (residuos Phe8-Tyr101 y Lys102-Pro184) están representados de color amarillo y verde, respectivamente. El tercer dominio (residuos Thr201-Thr303), de color rosa, se une al dominio I mediante un *loop* largo, color azul. El sitio activo está localizado en la interfaz del dominio I y II. **(4b)** Diada catalítica de 3CLpro, conformada por la His41 (amarillo), la cual está a 3.6 Å de la Cys145 (verde, segundo dominio). **(4c)** Subsitios de unión al sustrato, se encuentran marcados con las leyendas S1, S2, S1' y S4. **(4d)** Posiciones del sustrato de 3CLpro (código PDB: 2Q6G) dentro de la diada catalítica, las posiciones de los residuos son marcadas con las leyendas P1, P1', P2, P3, P4 y P5. **(e)** Homodímero y protómeros de 3CLpro, los colores corresponden a los mismos marcados en el inciso **(4a)**. La figura fue preparada usando PyMol (Schrödinger LLC).

El sitio activo de la 3CLpro cuenta con diferentes cavidades o *pockets* los cuales le confieren una alta especificidad al sustrato por cada residuo en la secuencia de escisión.

El sitio de unión del sustrato está conformado por 4 subsitios denominados: S1, S1', S2 y S4 (Figura 4c), estos están seguidos de los residuos en la posición P1, P1', P2, P4, de la secuencia del sustrato (ver Figura 4d). Estudios de Modelado molecular han sugerido que la flexibilidad de las regiones 140-146 y 184-197, las cuales abarcan los *pockets* de unión S1 y S2, son cruciales para el acomodo de sustratos y análogos [83,84]. Cabe destacar que el residuo glutamina está altamente conservado en la posición P1 [85]. Otros estudios destacan que la posición P2 acomoda residuos hidrofóbicos con cadenas laterales grandes como leucina y fenilalanina, mientras que la posición P1 tolera residuos de tamaño pequeño [86].

El proceso proteolítico se ha propuesto que sigue un mecanismo de múltiples pasos. El primer paso en la reacción catalítica de 3CLpro es la desprotonación del grupo tiol de la Cys145 por la His41. El tiolato resultante funciona como nucleófilo atacando el enlace amida del sustrato, formando un tiohemiacetal, el cual colapsa en un tioéster tras la ruptura del enlace peptídico. El producto peptídico N-terminal se libera por abstracción de protones de la histidina previo a que el tioéster se hidrolice para liberar el producto C-terminal y restaurar la díada catalítica.

La formación del enlace tioéster es un paso crítico en el mecanismo catalítico de 3CLpro y su bloqueo por parte de compuestos antivirales lleva a la inactivación de la enzima [10,87]. El mecanismo catalítico se esquematiza en la Figura 5.

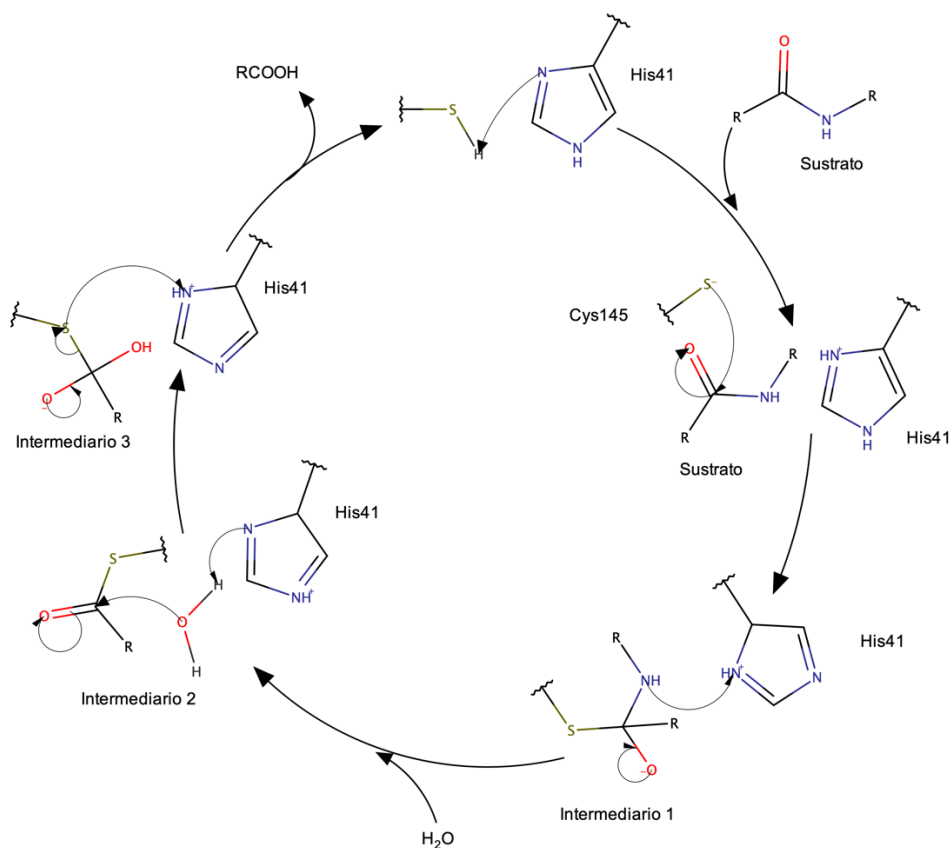


Figura 5. Esquema 2D general del mecanismo catalítico de la 3CLpro de SARS-CoV-2.
Modificado [88].

Entre la familia *Coronaviridae*, y particularmente dentro de los diferentes *Betacoronavirus*, el sitio de unión a sustrato en 3CLpro se encuentra altamente conservado [10,20,27,81,83,89,90]. La 3CLpro de SARS-CoV-2 presenta una estructura que es virtualmente idéntica a su ortólogo de SARS-CoV (96% de identidad) [91]. Ver Figura 6.

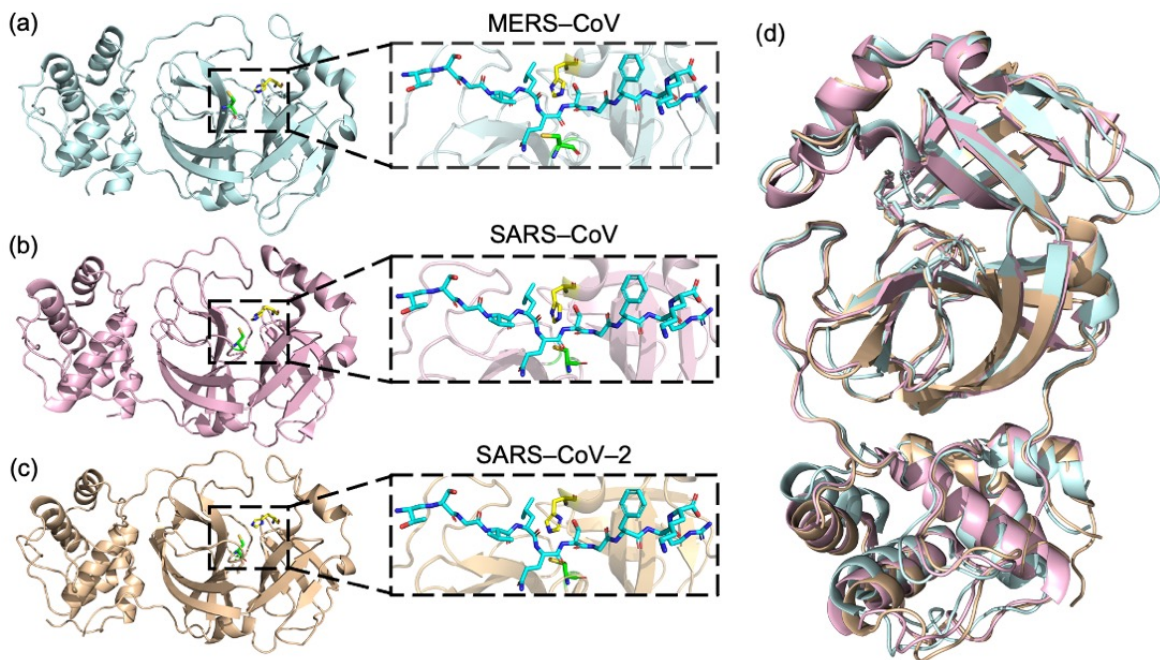


Figura 6. Comparación de estructuras cristalizadas de 3CLpro de diferentes coronavirus.

(6a) MERS-CoV (azul, código PDB 5C3N), **(6b)** SARS-CoV (rosa, código PDB 2BX4) y **(6c)** SARS-CoV-2 (naranja, código PDB 6M2N). La diada catalítica está indicada por un rectángulo a un lado de la estructura del monómero proteico. **(6d)** Alineamiento estructural de la 3CLpro de MERS-CoV, SARS-CoV-2 y SARS-CoV. La figura fue preparada usando PyMol (Schrödinger LLC).1.3.3.

1.4.1 Inhibidores de 3CLpro

Debido a que la 3CLpro de SARS-CoV-2 es crucial para el ciclo de replicación viral y se conoce a detalle su estructura, ha figurado como un blanco terapéutico atractivo, razón por la cual se han desarrollado diversos inhibidores para combatir la patología que deriva de su infección. Se han identificado una amplia variedad de potenciales inhibidores de 3CLpro que interactúan con el sitio activo tanto de forma reversible (no covalente), como irreversible (covalente). Dentro de los inhibidores reversibles reportados se encuentra el Boceprevir, un potente inhibidor

de la 3CLpro (CI_{50} 1.63 μ M) el cual exhibe una actividad antiviral dosis-dependiente significativa [92], sin embargo demostró no ser un inhibidor específico [93]. El Telaprevir, otro inhibidor reversible de 3CLpro mostró una baja inhibición moderada de la proteasa (CI_{50} >20 μ M) [94]. Dentro del grupo de los inhibidores del tipo covalente, destacan el aceptor de Michael N3 y la α -cetoamida 13b. Ambos compuestos demostraron inhibir la actividad de 3CLpro recombinante (CI_{50} 16.77 μ M y CI_{50} 4.67 μ M) [87], además el compuesto 13b demostró actividad inhibitoria contra la COVID-19 en células humanas Calu-3 infectadas con SARS-CoV-2. Cabe destacar que la actividad inhibitoria moderada de estos compuestos implica el requerimiento de inhibidores de alta afinidad contra 3CLpro [95].

A pesar de los grandes esfuerzos internacionales para desarrollar inhibidores específicos contra la 3CLpro de SARS-CoV-2 [96], tal y como se mencionó previamente, solo el Velkury (Remdesivir) ha sido aprobado como tratamiento contra la COVID-19. Por otra parte, se ha autorizado el uso de emergencia del peptidomimético oral de Pfizer, Paxlovid, originalmente llamado PF-07321332. Ha sido autorizado por la FDA como el primer tratamiento antiviral oral contra la COVID-19 de leve a moderada, en adultos y pacientes pediátricos (mayores a 12 años que pesen al menos 40 kg). Cabe destacar que este medicamento no está autorizado para la prevención previa o posterior a la exposición al COVID-

19 o para la iniciación del tratamiento en personas que requieren hospitalización debido a la COVID-19, grave o crítico [64].

El Paxlovid es una forma farmacéutica que se compone de Nirmatrelvir, un inhibidor irreversible de 3CLpro del SARS-CoV-2 para impedir que el virus se replique, y de Ritonavir, que desacelera la descomposición del Nirmatrelvir para ayudar a que permanezca en el organismo durante más tiempo en concentraciones más altas [64].

La evidencia para la autorización se basó en un ensayo clínico aleatorio, doble ciego y controlado con placebo, para el tratamiento de adultos mayores a 18 años, sintomáticos no hospitalizados con un diagnóstico de laboratorio confirmatorio para infección de SARS-CoV-2 (EPIC-HR, por sus siglas en inglés).

Entre los efectos secundarios se encuentra el deterioro del sentido del gusto, diarrea, presión arterial alta y dolores musculares. La FDA reporta que su uso junto con el de otro medicamento puede resultar en una interacción medicamentosa potencialmente significativa. De igual forma, su uso en pacientes con infección no controlada o sin diagnóstico de HIV puede derivar a una resistencia ante los medicamentos usados como tratamiento para esta enfermedad. El Ritonavir puede causar daño hepático. Cabe destacar que, hasta estos días no existen alternativas adecuadas, aprobadas y disponibles al Paxlovid para el tratamiento del COVID-19.

Es por todo lo anterior que es fundamental la búsqueda de nuevos y mejores candidatos con actividad inhibitoria ante 3CLpro de SARS-CoV-2 con el fin de contar con una amplia variedad de tratamientos para los casos en los que el Paxlovid no pueda ser empleado.

1.5 Uso de metodologías computacionales para el desarrollo de compuestos anti-SARS-CoV-2

Con la actual pandemia de COVID-19, la investigación relacionada con el SARS-CoV-2 ha tenido no sólo un abrupto aumento de velocidad, sino que también los científicos de todo el mundo están investigando y generando información acerca de la evolución y patogénesis del agente etiológico con la finalidad de poder controlar esta pandemia. A la luz de esta enorme cantidad de datos que se han generado, muchas preguntas fundamentales en la investigación del SARS-CoV-2 pueden ser abordadas con la ayuda de la bioinformática. El análisis adecuado de estos datos puede impulsar el descubrimiento de posibles agentes terapéuticos que combatan la COVID-19 [97]. A continuación, se abordarán más a detalle algunos temas y conceptos relevantes relacionados al descubrimiento, diseño y desarrollo de fármacos contra SARS-CoV-2.

1.5.1 Biología estructural

Cabe destacar que la estructura de 3CLpro y el cribado virtual (VS, por sus siglas en inglés) son de suma importancia, esto debido a que la

primera es usada en la identificación y selección de las dianas, mientras que el segundo es usado para la identificación de potenciales candidatos y fragmentos. Además, la biología estructural funge un rol crítico en la optimización de líderes con una mayor afinidad y selectividad [98].

Son notables los esfuerzos de la comunidad científica para visualizar y comprender la compleja biología del agente que impulsa la pandemia a través de estudios de estructura-función de diferentes proteínas del SARS-CoV-2.

Poco después de la publicación de la secuencia del genoma viral, en enero de 2020, se determinó y liberó la estructura cristalina de la proteasa principal del SARS-CoV-2 (código PDB:6LU7) [87]. A esto le siguieron varias otras estructuras de proteínas virales estructurales, NSP y accesorias [99]. Al 9 de noviembre de 2021, había 183,793 estructuras de varias proteínas del SARS-CoV-2 en forma APO (sin ligando unido) o en complejo con ligandos y otras proteínas en *Protein Data Bank* (PDB) [100], dentro de los blancos farmacológicos que destacan entre estas estructuras cocrystalizadas encontramos a la proteína M, N, S, E, 3CLpro, PLpro y RdRp. La determinación de estas estructuras hasta niveles atómicos es importante, no solo para entender las bases moleculares de la biología viral, sino también para el diseño y desarrollo de moléculas que puedan actuar como potenciales agentes terapéuticos [99]. Esta importancia fue reconocida y abordada por el portal UniProt, debido a que

generó un sub-portal llamado *COVID-19 UniProt*, el cual proporciona acceso previo a la liberación temprana de (1) secuencias de proteínas de SARS-CoV-2, (2) proteínas cercanas de SARS-CoV, (3) proteínas humanas relevantes para la biología de la infección viral (como enzimas o receptores), (4) ProtVista, herramienta que permite visualizar las características de secuencia de cada proteína, (5) enlaces de herramientas de análisis de secuencia, (6) acceso a la recopilación de publicaciones relevantes para COVID-19, aportadas por la comunidad, así como (7) enlaces a recursos relevantes [97]. El portal COVID-19 UniProt está disponible bajo la licencia de Creative Commons Attribution (CC BY 4.0) vía <https://covid-19.uniprot.org/>.

1.5.2 Descubrimiento, diseño y desarrollo de fármacos

En el descubrimiento de fármacos, la identificación de nuevos compuestos es la pieza clave para la innovación y el éxito en el posterior diseño de fármacos de moléculas pequeñas [101]. Actualmente el proceso para el descubrimiento de fármacos mediante métodos informáticos, de forma general, incluye cinco etapas: 1) la identificación de la enfermedad a ser tratada, 2) la selección del blanco molecular modulable (o *druggable*), 3) su validación, 4) el cribado de alto rendimiento (HTS, por sus siglas en inglés) de librerías de compuestos y, finalmente, 5) la optimización del compuesto con el fin de mejorar su potencia y selectividad.

Subsecuentemente, los compuestos líderes se optimizan para mejorar la eficacia y farmacocinética antes de seguir con el desarrollo de fármacos [102] (ver Figura 7).

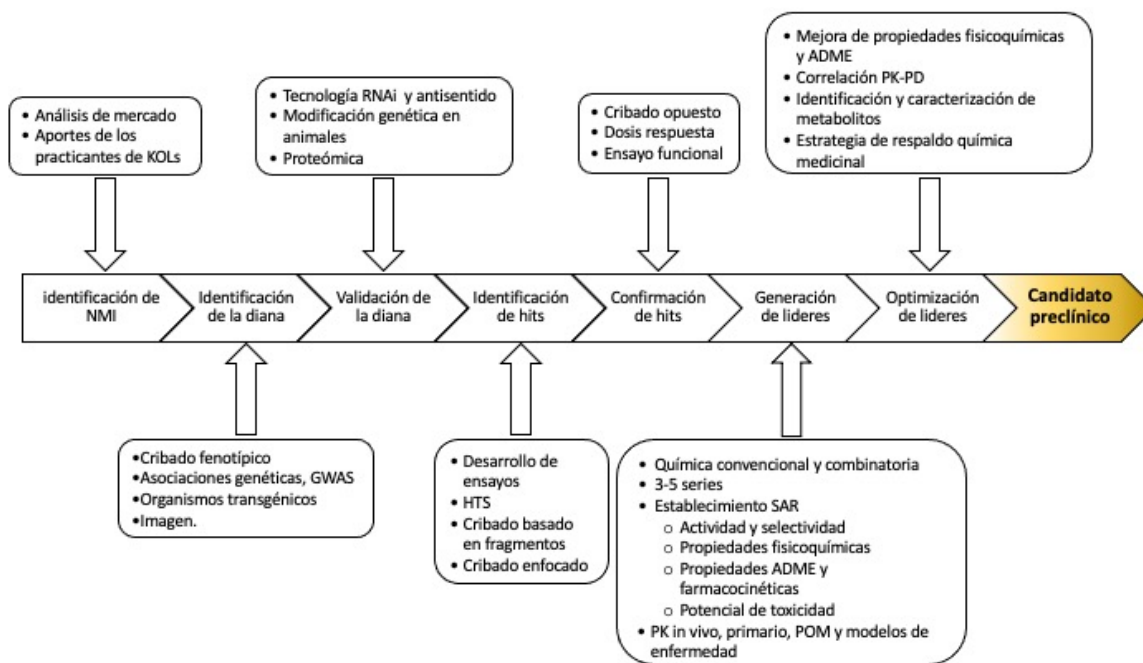


Figura 7. Descripción general del proceso de descubrimiento de fármacos.

Necesidades médicas insatisfechas (NMI); líder clave de opinión (KOL); relación estructura actividad (SAR); absorción, distribución, metabolismo y excreción (ADME); estudios de asociación de todo el genoma (GWAS); prueba de mecanismo (POM); farmacocinética (PK); farmacodinamia (PD). Modificado de [102].

El desarrollo de fármacos es un proceso complejo que puede ser dividido en dos etapas conocidas como: etapa preclínica y clínica [103], etapas que incluyen ensayos en líneas celulares, en animales y pruebas clínicas en humanos. Los compuestos que pasan satisfactoriamente por estas dos etapas son aprobados por un agente regulatorio para su uso clínico, por

ejemplo, la Comisión Federal para la Protección contra Riesgos Sanitarios (COFEPRIS) en México o la *FDA* en Estados Unidos [104].

Este proceso inicia con la identificación de compuestos que se unen a un blanco terapéutico o que muestran actividad biológica en un ensayo de tamizaje, los cuales tienen la potencia de ser candidatos a fármacos tras su evaluación. Aquellas moléculas que muestran actividad biológica reciben el nombre de *hits*. Seguido de esto se busca encontrar compuestos, típicamente a través de modificaciones químicas sobre los *hits*, cuyas propiedades farmacéuticas sean atractivas, por ejemplo: baja toxicidad, o adecuada solubilidad acuosa para administración vía oral. Los compuestos derivados de esta segunda etapa son denominados líderes (del inglés *leads*) o cabezas de serie [104].

La identificación y validación de la diana es uno de los pasos más cruciales en el desarrollo de un nuevo fármaco. Un blanco molecular o diana terapéutica es un término general que se refiere a una variedad de entidades biológicas, como proteínas, genes y RNA [105]. Un buen blanco molecular debe ser eficiente, seguro, conocido tanto clínica como comercialmente y, más importante, debe ser modulable (del inglés *druggable*) [106,107], lo cual significa que, al interactuar con un compuesto químico o una proteína, logrará disparar una respuesta biológica medible tanto *in vitro* como *in vivo* [102,105], la cual altera el

curso de la enfermedad retornando al estado normal [106]. Las etapas principales en el desarrollo de fármacos se visualizan en la Figura 8.

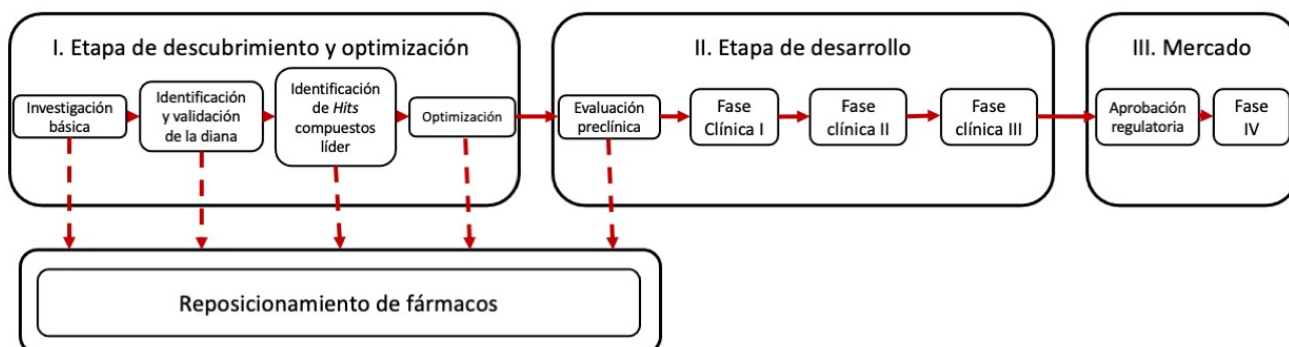


Figura 8. Etapas principales en el desarrollo clásico de fármacos.

Se consideran como etapas al 1) descubrimiento y optimización, 2) desarrollo y 3) mercado. Tomado de [104].

En el contexto de SARS-CoV-2 vale la pena señalar que, aunque se han desarrollado, probado y distribuido múltiples vacunas efectivas contra la COVID-19, con una velocidad sin precedentes, aún no se entiende por completo su eficacia a largo plazo, efectos secundarios y la cobertura que tiene ésta ante las variantes de SARS-CoV-2 que emergen rápidamente. Sin mencionar que ninguna de estas vacunas ofrece un 100% de protección a todos los candidatos vacunados, es por lo anterior que un tratamiento farmacológico es indispensable ante la actual situación mundial. Cabe destacar que las pequeñas moléculas que actúan como inhibidores y las vacunas, corresponden a enfoques totalmente complementarios, orientados a la terapia y prevención, donde ambas permitirían contener la pandemia por COVID-19 [108].

El crecimiento progresivo de la capacidad de análisis y robótica ha permitido una detección a gran escala de potenciales inhibidores contra el SARS-CoV-2 [108]. Tal es el caso de un estudio reciente que utilizó modelos basados en actividad biológica para identificar 311 sustancias químicas, donde 99 de ellas mostraron actividad *in vitro* en contra del SARS-CoV-2, de los cuales los más potentes presentaban CI_{50} en el rango nanomolar [109].

En otro estudio a gran escala se evaluaron alrededor de 12,000 compuestos en estadio clínico o aprobados por la FDA (provenientes de la biblioteca ReFRAME) en un ensayo sobre células Vero, infectadas con SARS-CoV-2, resultando en la identificación de 21 candidatos con relaciones dosis-respuesta prometedora. Dentro de los cuáles destacaron la Clofazimina (CE_{50} 0.31 μ M) y el inhibidor de la fosfatidilinositol-3-fosfato 5-quinasa Apilimod (CE_{50} 0.023 μ M) [110]. Este último entró en ensayos clínicos para COVID-19 en junio de 2020 debido a que se demostró su potencia adecuada en líneas celulares humanas (293T y Huh-7) infectadas (12 y 88 μ M, respectivamente) [111]. En julio del mismo año, la Clofazimina también avanzó a ensayos clínicos como parte de una terapia combinada [112].

Por otra parte, cabe destacar que hasta el 31 de diciembre del 2021 hubo más de 470 estudios en proceso revisados por la FDA dentro de dos categorías, etapa temprana (fase clínica 0, 1 y 1/2) y tardía (fase clínica

2, 2/3, 3 y 4). Dentro de éstos, se han autorizado 14 tratamientos farmacológicos de soporte para evitar el progreso de la COVID-19, tratamientos autorizados únicamente bajo ciertas condiciones y para su uso de emergencia durante la COVID-19. Dentro de esta lista se encuentran fármacos tales como el Molnupiravir, Paxlovid, Baricitinib-Olumiant; plasma convaleciente; anticuerpos monoclonales como Evusheld (Tixagevimab empaquetado con Cligavimab), Actemra (Tocilizumab), Sotrovimab, Bamlanivimab; agentes sedantes como Propofol -Lipuro 1% y Fresenius Kabi Propoven 2% y por último, una solución para la Terapia de Reemplazo Renal Continuo llamada REGIOCIT [66]. Finalmente, el Veklury (Remdesivir) ha sido aprobado por la FDA como tratamiento para la COVID-19 para ciertos pacientes [66,113].

1.5.3 Estrategias para identificar compuestos líderes

Existe una amplia variedad de estrategias experimentales y/o computacionales para identificar y optimizar compuestos activos o líderes [104], tales como:

- Reposicionamiento y optimización de fármacos.
- Ensayos biológicos sistemáticos de colecciones de compuestos, siendo el ejemplo más típico el HTS.
- Uso de información biológica disponible.

- El diseño y descubrimiento de fármacos asistidos por computadora (CADD, por sus siglas en inglés).

Cronológicamente hablando, con el fin de identificar y optimizar líderes, a principios de los 80's se comenzaron a usar modelos comparativos basados en homólogos [98,114]. Una década más tarde se reconoció que las estructuras cocrystalizadas 3D eran útiles para la definición de topografías de las superficies complementarias de los ligandos y sus proteínas blanco, lo que sirvió para optimizar la potencia y selectividad de los líderes [115].

En el contexto de la COVID-19, los antivirales existentes y las experiencias obtenidas de los brotes previos de SARS-CoV y MERS-CoV pueden ayudar a dilucidar el camino para la identificación de posibles fármacos empleando las técnicas antes listadas.

Tal es el caso del más reciente candidato propuesto por Pfizer, donde Owen DR y col. (2021) mediante reposicionamiento y optimización de fármacos, a partir de un *lead* previamente caracterizado como inhibidor de 3CLpro en SARS-CoV, desarrollaron el candidato oralmente biodisponible, PAXLOVID, inhibidor de 3CLpro en SARS-CoV-2 [116], el cual actualmente se encuentra bajo proceso de investigación y se considera que será el primer fármaco antiviral aprobado, desarrollado específicamente contra el SARS-CoV-2 [108,117].

1.5.4 Diseño de fármacos asistido por computadora

El CADD tiene como principio entender las relaciones estructura-actividad biológica o farmacológica de compuestos bioactivos [118]. Éste tiene la ventaja de no sólo reducir los costos asociados con el descubrimiento de fármacos, sino que también puede reducir el tiempo que le toma a un fármaco llegar al mercado de consumo, ventajas que son de suma relevancia ante la situación actual originada por la COVID-19. Además, el CADD ha sido empleado en el descubrimiento de numerosos fármacos disponibles en la actualidad, que han recibido la aprobación de la FDA y han alcanzado el mercado de consumo, tal es el caso del antiviral antes mencionado, Veklury (Remdesivir) aprobado por la FDA para el tratamiento contra la COVID-19 en casos de emergencia particulares [119].

Los métodos de CADD pueden ser clasificados en dos grupos (ver Figura 9), los basados en estructura (SB, por sus siglas en inglés) y los basados en ligando (LB, por sus siglas en inglés) [120,121]. La diferencia entre estos dos enfoques es que las herramientas de los métodos SB sólo pueden ser aplicados si se conocen las estructuras de las dianas biológicas [120,122], sin embargo, si éstas no son conocidas, la aproximación LB es usada como alternativa ya que se basa en el conocimiento de moléculas pequeñas que se unan a la diana de interés [120,123,124].

Las herramientas principales de SB son:

1. El HTS y VS.
2. Los métodos de acoplamiento molecular directo.
3. El diseño *de novo* de ligandos.

Las herramientas principales de LB son [120,125]:

1. El modelado farmacofórico [126].
2. La aproximación por similaridad molecular o *fingerprints* [127]. En este método la huella dactilar de ligandos conocidos que se unen a una diana es usada para encontrar moléculas con huellas dactilares similares a través del cribado en librerías moleculares. Los *fingerprints* son una forma de codificar la estructura de una molécula [128], son representaciones de cadenas de bits donde cada posición de bit generalmente representa la presencia o ausencia de una característica estructural, fragmento, propiedad o patrones farmacológicos [127].
3. El modelado QSAR (del inglés *quantitative structure-activity relationship*) es un método que modela la relación entre las características estructurales de los ligandos que se unen a un blanco y los correspondiente efectos sobre la actividad biológica [129].

Las herramientas principales de SB y el modelo farmacofórico serán abordados de forma particular en los siguientes apartados.

Los métodos CADD han sido exitosamente usados en el proceso de descubrimiento de fármacos contra la COVID-19. Selvaraj y col. (2021) dilucidaron la estructura tridimensional de la guanina-N7 metiltransferasa (NSP14) de SARS-CoV-2 empleando una metodología de modelado por homología, además, propusieron cinco compuestos provenientes de librerías químicas de compuestos característicos de la Medicina Tradicional China (TCM, por sus siglas en inglés), TCM 57025, TCM 3495, TCM 5376, TCM 20111 y TCM 31007 como posibles fitoquímicos antivirales basados en estudios de simulación y acoplamiento molecular [130]. En otro estudio mediante un modelo de interacción proteína-ligando preentrenado basado en *Deep learning*, llamado *Molecule Transformer-Drug Target Interaction* (MT-DTI), identificaron algunos fármacos antivirales aprobados por la FDA (Atazanavir, Remdesivir, Efavirenz, Ritonavir y Dolutegravir) que mostraron potencial actividad inhibitoria contra 3CLpro de SARS-CoV-2 [131].

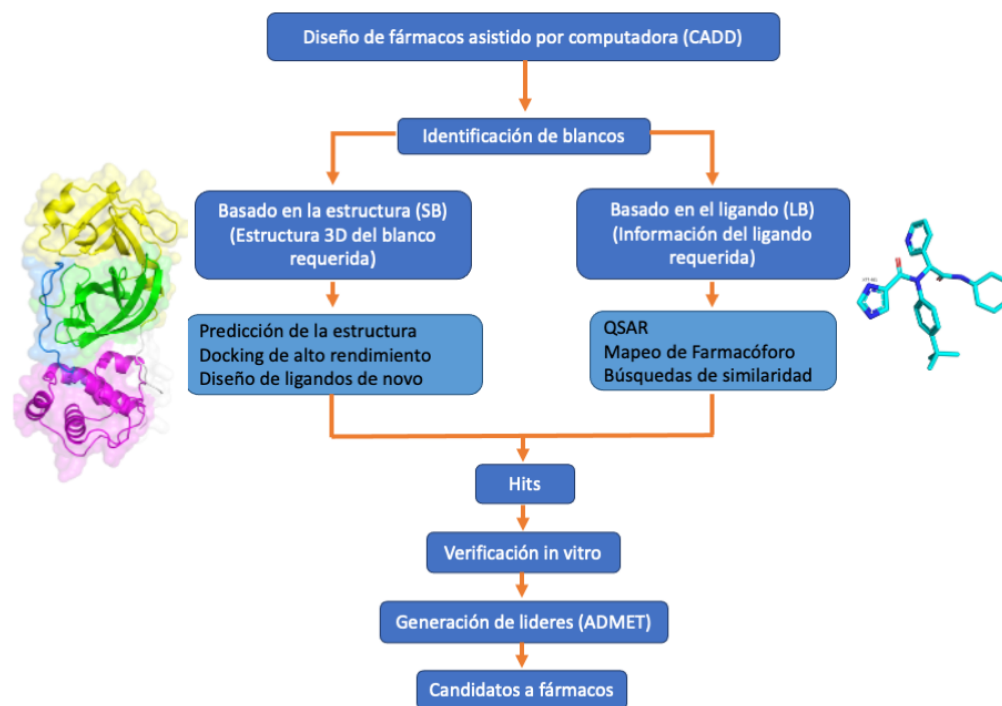


Figura 9. Representación del proceso del CADD.

Esquematación del CADD con el enfoque LB y SB. Proceso de absorción, distribución, metabolismo, excreción y toxicidad (ADMET) Modificado de [120].

1.5.4.1 CADD del tipo SB: VS, HTS y acoplamiento molecular

Como método complementario al HTS tenemos al VS, el cual es una aproximación al descubrimiento de fármacos asistido por computadora [132]. El VS, un método *in silico* del HTS, es una técnica que emplea programas computacionales con el fin de buscar potenciales *Hits* (cribado *in silico*) en librerías virtuales de fragmentos [101,133,134]. Éste consiste en colocar virtualmente (acoplamiento molecular o *molecular docking*) colecciones de millones de compuestos en un blanco biológico, seguido de una evaluación del acomodo entre éstos (puntuación o *scoring*). [135]

Lo que diferencia el HTS del VS es que el HTS es una aproximación experimental, mientras el VS es una aproximación teórica. Por un lado, el HTS requiere una alta inversión de tiempo, infraestructura y tiene un bajo índice de éxito (<5%) [135]. Mientras que el VS es rápido y rentable, sin embargo tiene una precisión relativamente baja en la predicción y una rápida acumulación de errores [101,136,137].

Los requerimientos básicos del VS son:

1. Una colección de compuestos. Estas colecciones se encuentran almacenadas y recopiladas en diversas bases de datos o librerías de compuestos, dentro de las cuales destacan [120]:
 - a. Pubchem del National Institutes of Health (NIH), un repositorio que contiene millones de moléculas pequeñas biológicamente relevantes [138].
 - b. Zinc es una librería abierta de compuestos listos para ser empleados en VS. Contiene más de 35 millones de moléculas que pueden estar disponibles comercialmente [138,139].
 - c. Drugbank, librería con más de 14,000 moléculas, dentro de los cuales más de 4200 son fármacos aprobados por la FDA [140].
2. La estructura del blanco biológico.
 - a. *PDB*, recurso global que contiene información determinada experimentalmente de macromolécula biológicas [141,142].

- b. Modelado molecular de proteínas, en caso de no contar con estructuras 3D. Técnica predictiva que permite obtener una aproximación de la estructura atómica de una cadena polipeptídica con base en la estructura conocida de proteínas relacionadas (homólogas), comúnmente llamadas moldes [143].

3. Un apropiado esquema de acoplamiento/evaluación.

Existen dos aproximaciones generales de VS, la LB y SB [133,144]. La primera puede ser empleada cuando se conoce un conjunto de moléculas ligando conocidas y hay poca o nula información disponible sobre el blanco terapéutico, tal es el caso del modelo farmacofórico y el método QSAR. El segundo depende de la información estructural del sitio de unión en la proteína blanco, siendo el acoplamiento molecular el método mayormente empleado desde los 80's [126,132,145,146].

El acoplamiento molecular automatizado es un método teórico para estudiar las interacciones y el reconocimiento entre proteínas y ligandos [147]. Consiste en la búsqueda de la conformación y posición óptima de un ligando (generalmente moléculas pequeñas) dentro del sitio activo de un blanco macromolecular [104].

Esta metodología está conformada por dos componentes: a) *docking* o proceso de búsqueda de la conformación y orientación de las moléculas

(o muestreo), y b) *scoring*, que consiste en asignar un valor o puntaje que mida la interacción entre las dos estructuras y que permita identificar el modo de unión [104,148].

Mediante el entendimiento de las interacciones que tienen lugar entre el ligando y su receptor, el acoplamiento molecular abre las puertas a la estimación de la afinidad previo a la síntesis, así como a las técnicas de optimización del ligando [134]. Existe una gran variedad de herramientas y programas para acoplamiento molecular, tales como AutoDock [149], AutoDock Vina [150], LeDock [151], USCF DOCK [152], Glide [153], Gold [154], MOE Dock [155], Surflex-Dock [156], etc., las cuales han sido desarrolladas para usos comerciales y académicos [157–159].

Lo anterior convierte a este conjunto de técnicas en estrategias de suma importancia en la búsqueda de candidatos prometedores con el potencial de dirigirse hacia algún blanco proteico en SARS-CoV-2 [21] mediante la predicción de la energía de unión proteína-ligando [160], es por lo que a lo largo de los últimos meses estas técnicas se han visto explotadas. Como ejemplos de lo anterior, Das S, y col. (2021) utilizaron un enfoque de acoplamiento molecular para identificar potenciales inhibidores en contra de 3CLpro mediante el VS de productos naturales, antifúngicos, antiprotozoarios, antivirales y antinematodos. Los potenciales inhibidores identificados en el estudio incluyen Rutosido, Ritonavir, Emetina,

Hesperidina, Lopinavir e Indinavir [161]. Por su parte, otro estudio de Khan SA, y col. (2021) reveló tres medicamentos aprobados por la FDA (Remdesivir, Saquinavir y Darunavir) y dos compuestos naturales (derivados de flavina y cumarina) como inhibidores prometedores de 3CLpro empleando un enfoque de simulación de dinámica molecular y acoplamiento molecular [162].

Se han realizado múltiples estudios con similares aproximaciones enfocados en 3CLpro, a través de los cuales se ha logrado identificar potenciales líderes [163] de diferentes tipos tales como moléculas de origen natural [164], fitoquímicos [165], análogos de estilbenoides [166], etc.

1.5.5 Modelo farmacofórico

Ehrlich, en 1909, introdujo por primera vez el concepto de farmacóforo [167], definiéndolo como una referencia que lleva las características principales responsables de la actividad biológica de un fármaco. Concepto que, a pesar de un siglo de desarrollo, permanece sin cambios. Por otra parte, la *International Union of Pure and Applied Chemistry* (IUPAC) define un modelo farmacofórico como un ensamble de características estéricas y electrónicas que son necesarias para asegurar las interacciones óptimas supramoleculares con un blanco terapéutico específico y desencadenar (o bloquear) su respuesta biológica [168]. Es

importante destacar que un farmacóforo no representa a una molécula real o a una asociación real de grupos funcionales, es un concepto abstracto que indica la capacidad de interacción molecular común de un grupo de compuestos dirigidos a un blanco farmacológico específico. Es enunciado por algunos autores como un común denominador de un conjunto de moléculas activas [104].

Un modelo farmacofórico puede ser establecido con base en el ligando (LBP, por sus siglas en inglés) y con base en la estructura (SBP, por sus siglas en inglés).

El LBP se genera mediante la superposición de un conjunto de moléculas activas (llamado generalmente grupo de entrenamiento), extrayendo las características químicas comunes, esenciales para la bioactividad, de estructuras 3D de un conjunto de ligandos representativos en el contexto de un blanco macromolecular. Este modelo se ha convertido en una estrategia computacional clave debido a que facilita el descubrimiento de fármacos en ausencia de la estructura de un blanco molecular [168].

En general, la formación de un modelo farmacofórico a partir de múltiples ligandos envuelve dos pasos principales, 1) la creación del espacio conformacional para cada ligando dentro del grupo de entrenamiento que representa la flexibilidad conformacional de los ligandos, y 2) el alineamiento de los múltiples ligandos dentro del grupo de entrenamiento

y determinando las características químicas esenciales comunes para la construcción de los modelos farmacofóricos [168].

Actualmente, se han desarrollado una amplia variedad de generadores de farmacóforos automatizados tales como: HipHop [169,170], Hypogen [170,171], DISCO [170], GASP [172], GALAHAD [173,174], PHASE [175] y MOE [176], así como múltiples programas académicos. Estos programas difieren principalmente en el algoritmo usado para manejar la flexibilidad de los ligandos y para el alineamiento molecular. Cabe destacar que en la actualidad se han establecido una serie de algoritmos avanzados para el muestreo de espacio conformacional de moléculas pequeñas, tales como la restricción de poses [177], rejillas de torsión sistemática [178], ajuste dirigido [179], algoritmos genéticos [180] y Monte Carlo [181].

Se ha demostrado que el tipo de ligando molecular, el tamaño de la base de datos y su diversidad química dentro del grupo de entrenamiento, afectan considerablemente el modelo farmacofórico generado, razón por la cual esto ha sido detectado como un problema desafiante, ya que en algunos casos modelos farmacofóricos completamente diferentes de ligandos que actúan con el mismo objetivo macromolecular podrían ser generados a partir del mismo algoritmo y programa que utiliza diferentes grupos de entrenamiento [126].

El SBP se genera mediante la exploración de los posibles puntos de interacciones entre las estructuras 3D del blanco molecular y los ligandos. El protocolo de este modelo implica un análisis de las características químicas complementarias del sitio activo y sus relaciones espaciales y un subsecuente ensamblaje del modelo farmacofórico con las características seleccionadas. Se pueden clasificar en dos categorías, los basados en complejos macromolécula-ligando y los basados en macromoléculas sin ligando. El primero es conveniente para localizar el sitio de unión del ligando de la macromolécula blanco y determinar los puntos clave de interacción entre ellos, sin embargo, la limitación de esta aproximación es la necesidad de la estructura 3D del complejo macromolécula-ligando. [168]. Este esquema se incluye y se representa de forma adecuada en LigandScout [182], Pocket [183] y GBPM [184].

Los problemas más frecuentes para el SBP son:

1. Se pueden identificar demasiadas características químicas, no prioritarias, para un sitio específico de unión a una macromolécula blanco. En el mismo sentido, Yang S-Y., y colaboradores (2010) informan que un modelo farmacofórico compuesto por más de 7 características químicas no es adecuado para aplicaciones prácticas, siendo lo ideal seleccionar entre 3 a 7 [126].
2. El modelo farmacofórico obtenido no puede reflejar la QSAR debido a que esta aproximación sólo se basa en un único complejo

macromolécula ligando. De igual forma, Yang S-Y y colaboradores (2010) sugieren el uso de un mapa completo basado en multicomplejos y en el modelo farmacofórico más frecuente [126]. Una vez que el modelo farmacofórico es generado, ya sea por la aproximación SBP o LBP, puede ser usado para 1) consultar las características químicas 3D en librerías químicas con el fin de identificar potenciales ligandos, técnica conocida como VS basado en farmacóforo, 2) diseñar ligandos *de novo*, siendo NEWLEAD [185] el primer programa en llevar a cabo este diseño. Otros softwares tales como LUDI [186], BUILDER [187] y PhDD [188] también son usados en el diseño *de novo* basado en farmacóforo [126], 3) optimización de líderes y el diseño de fármacos multiblancos [126].

El modelo del farmacóforo es un enfoque rápido y eficaz en la identificación de moléculas líderes interesantes para el descubrimiento de fármacos contra la enfermedad COVID-19 [160].

Empleando esta metodología, Daoud S, y col. (2021), usando estructuras cristalizadas de 3CLpro de SARS-CoV-2, construyeron un modelo farmacofórico y realizaron estudios de acoplamiento molecular. Lograron identificar cinco medicamentos antivirales aprobados por la FDA (Lopinavir, Remdesivir, Ritonavir, Saquinavir y Raltegravir) y los resultados del acoplamiento molecular revelaron que estos compuestos

presentan interacciones de unión específicas comparables con las presentadas por el inhibidor cocrystalizado X77 [189].

Otros múltiples estudios han sido realizados con similares aproximaciones enfocados en múltiples blancos moleculares de SARS-CoV-2, tales como 3CLpro [190–193], TMPRSS2 [194], RdRp [195], etc., a través de los cuales se han logrado identificar candidatos con potencial actividad inhibitoria contra SARS-CoV-2.

II. Justificación

Finalmente, considerando todo lo anterior, al inhibir la actividad proteolítica de 3CLpro se bloqueará la liberación de las NSPs y se inhibirá la maduración y efectividad del SARS-CoV-2 [10]. Cabe destacar que múltiples grupos de investigación han identificado a esta proteasa como uno de los blancos más atractivos en el tratamiento de infecciones por coronavirus debido a su importancia funcional en el ciclo viral, combinado con la ausencia de homólogos cercanos en humano [20].

III. Hipótesis

La identificación, por métodos computacionales, de los descriptores farmacofóricos comunes entre compuestos que se unan a 3CLpro de SARS-CoV-2 permitirá la generación de un modelo farmacofórico consenso (MFC), el cual será de utilidad en la identificación de inhibidores de esta enzima.

IV. Objetivo

Generar un MFC y aplicarlo en el cribado virtual de bibliotecas químicas para la identificación de potenciales inhibidores de 3CLpro de SARS-CoV-2.

4.1 Objetivos particulares

- Seleccionar el sitio blanco en 3CLpro con base en la estructura de la proteasa.
- Generar modelos farmacofóricos individuales basados en las interacciones de la proteasa con inhibidores cocrystalizados.
- Generar un MFC a partir de los modelos farmacofóricos previos.
- Seleccionar y analizar, mediante técnicas quimioinformáticas, moléculas como potenciales inhibidores de 3CLpro.
- Seleccionar potenciales candidatos con la capacidad teórica de unión a 3CLpro con base en el modelo farmacofórico y acoplamiento molecular.

V. Metodología

5.1 Identificación y alineamiento estructural de 3CLpro

Con la finalidad de identificar todas las estructuras cristalizadas de 3CLpro con moléculas pequeñas, se realizó una búsqueda de 3CLpro de SARS-CoV-2 en *UniProt* con el código de acceso P0DTC1 [196], hasta el 28 marzo del 2021. Las estructuras identificadas de 3CLpro, en forma APO y con moléculas pequeñas tanto covalentes, como no covalentes fueron preparadas considerando un monómero de la proteasa principal y eliminando las moléculas de agua y disolvente.

La proteína del presente estudio empleada como referencia, previamente preparada, fue la proteasa principal 3CLpro en su forma APO que lleva el código PDB 6M2Q [197]. Este cristal fue seleccionado con base en su forma APO, su resolución de 1.7 Å y por su secuencia completa sin hueco. Con la finalidad de generar una biblioteca con muchos de los inhibidores cristalográficos de 3CLpro reportados, se alinearon todas las estructuras de 3CLpro y se extrajeron los ligandos para aquellas estructuras conteniendo inhibidores cocrystalizados en el sitio catalítico reteniendo sus coordenadas 3D.

5.2 Generación de los modelos farmacofóricos 3D de moléculas cocrystalizadas con 3CLpro

Para identificar las características indispensables para que los compuestos orgánicos se unan al sitio catalítico de 3CLpro, se generó un modelo farmacofórico 3D de cada ligando de la librería generada previamente en el contexto de su proteína 3CLpro cocrystalizada empleando *Pharmer* [198]. El modelo farmacofórico generado para cada complejo ligando-3CLpro se guardó para su futuro empleo.

5.3 Generación del MFC

Con el objetivo de compilar la información presente en todos los modelos farmacofóricos obtenidos por ligando cocrystalizado, se identificaron grupos basados en la distancia entre los descriptores farmacofóricos

mediante un análisis de agrupamiento jerárquico, empleando la librería SciPy [199]. Para la separación de los grupos de descriptores, se consideró una desviación del 27% de la distancia máxima euclidiana entre cada descriptor.

Posteriormente, para cada grupo formado, se calculó el centroide como el promedio de las coordenadas x, y y z, de todos los puntos contenidos en un mismo grupo. Como resultado, el MFC se definió como el conjunto de centroides provenientes del agrupamiento jerárquico. Adicionalmente, para cada descriptor del MFC se asignó la información del radio de Van der Waals asociado al tipo de descriptor molecular identificado por *Pharmer* [198].

5.4 Validación del MFC

Para llevar a cabo la validación del MFC se generó una librería de compuestos como se describe a continuación.

5.4.1 Librería de validación

A partir de los complejos cristalográficos de moléculas pequeñas con 3CLpro, se desarrolló un grupo de validación con ligandos únicos cocrystalizados. Con la finalidad de evitar redundancia y analizar la diversidad química y estructural de los compuestos cocrystalizados las moléculas de la librería se seleccionaron empleando los siguientes filtros: 1) poseer más de un enlace rotatable, 2) contar con más de tres

descriptores farmacofóricos y 3) tener un peso molecular mayor a 200 Da.

5.4.2 Biblioteca de confórmeros

Para cada molécula en el grupo de validación, se generaron 250 confórmeros energéticamente favorables empleando el módulo AllChem.EmbedMultipleConfs perteneciente a la librería de quimioinformática RDKit [200]. Este módulo emplea la distancia geométrica para obtener un conjunto de múltiples coordenadas para una molécula. Se consideró 1) el cumplimiento de la quiralidad al tener presentes centros quirales, 2) la imposición de preferencias experimentales de ángulos de torsión y 3) la imposición de química básica como anillos planos, 4) un umbral de RMSD (del inglés *Root-Mean-Square Deviation*) de 0.5 Å, así como 5) el empleo del campo de fuerza molecular de Merck MMFF94s (del inglés *Merck molecular force field*). Con el objetivo de filtrar confórmeros redundantes estos fueron seleccionados sólo cuando había una desviación de al menos RMSD de 0,5 Å entre los diferentes confórmeros de la misma molécula. La librería de confórmeros generada fue cargada en *Pharmit* para su posterior utilización [201].

5.4.3 Validación del MFC

Con el objetivo de identificar si el MFC era capaz de recapturar las poses cristalográficas de moléculas cocrystalizadas con 3CLpro, se empleó el

MFC en contra de la librería de confórmeros generada a partir del grupo de validación usando Pharmit. Los confórmeros obtenidos por el MFC se compararon en contra de su estructura cocrystalizada homóloga mediante el cálculo de RMSD empleando *RDKit* [200]. Si uno o más confórmeros de una molécula tenían un RMSD ≤ 2.5 Å se consideró como la reproducción del modo de unión cristalográfico. Adicionalmente, se generaron los descriptores farmacofóricos para cada confórmero obtenido con el objetivo de compararlos en contra de los centroides del MFC. Finalmente, con los datos obtenidos se calculó la eficiencia del MFC como:

$$\%Eficiencia = \frac{\#M}{SV} * 100$$

Donde #M Corresponde al número de estructuras con un RMSD ≤ 2.5 Å y SV al número de ligandos dentro del grupo de validación.

5.5 Búsqueda de candidatos

Debido a que el MFC proviene de la compilación de múltiples moléculas con diferentes características químicas y modos de unión, no es posible encontrar una entidad química existente que contenga todos los descriptores identificados. Por lo tanto, se segmentó el MFC en una serie de submodelos, los cuales consideraron entre 6 y 7 puntos del MFC, con el fin de aumentar la probabilidad de identificar potenciales entidades químicas que compartieran estas características. Para identificar moléculas con un modo de unión representado por los submodelos

farmacofóricos, se buscaron confórmeros en las siguientes librerías químicas: a) ChEMBL25, b) ChemDiv, c) ChemSpace, d) MCULE, e) MCULE-ULTIMATE, f) MolPort, g) NCI Open Chemical Repository, h) LabNetwork y i) Zink.

5.5.1 Submodelos farmacofóricos

La formación de estos submodelos consideró la combinación de 6 a 7 descriptores basados en su distancia que exploraran el sitio catalítico de 3CLpro basados en una inspección visual previa y la priorización de centroides con base en el *expertis* químico, es decir, en la experiencia adquirida tras el estudio, análisis y visualización de la naturaleza de la cavidad y de sus respectivos ligandos. Consideramos algunos puntos tales como, si los ligandos cubrían la mayor parte de la cavidad, los tamaños de los ligandos, los tipos y cantidades de interacciones que los ligandos presentaban con respecto a la proteasa, por mencionar algunos. La búsqueda de potenciales candidatos se realizó a través de *Pharmit* [201].

5.6 Optimización local y filtrado

Los candidatos encontrados que cumplían con las características de los submodelos fueron sometidos a un proceso de optimización local y filtrado con la finalidad de: 1) encontrar el mejor ajuste del candidato en la cavidad de 3CLpro y 2) dar solución a los choques generados con la proteasa. El origen de estos choques yace en que los candidatos

resultantes se encontraban acomodados en el sitio de interés considerando únicamente el MFC del cual derivaban y el sitio de unión, dejando de lado el tamaño del ligando y el comportamiento de éste en el contexto de la proteasa. Con la intención de eliminar los choques de los confórmeros con la proteína y asignar conformaciones energéticamente favorables a los ligandos en el contexto de la proteasa, se realizó una optimización local empleando *Smina* [202]. La opción de *scoring* y minimización empleada fue: *minimize*, considerando 800 iteraciones y un campo de fuerza permitido de 0.1. En cuanto al proceso de filtrado, se realizó una inspección visual a través de la cual se tomó en cuenta la selección del mejor confórmero en el contexto de la cavidad del receptor, conservando el sub-modelo farmacofórico. Se descartaron ligandos repetidos y aquellos que tras dicha optimización continuaban chocando con la proteasa o dejaban de reproducir el sub-modelo farmacofórico del cuál derivaron.

5.7 Acoplamiento molecular

Se llevaron a cabo estudios de acoplamiento molecular para evaluar las afinidades de unión entre la proteína de interés y los posibles candidatos optimizados derivados de los sub-modelos farmacofóricos, así como para identificar si los candidatos encontrados tenían el potencial teórico de

unirse a 3CLpro espontáneamente y si dicho modo de unión era el más favorable y frecuente.

La proteasa empleada como referencia fue la contenida en el cristal asociado al código PDB: 6M2N, a ésta se le adicionaron cargas e hidrógenos considerando un pH de 7.4. Las moléculas seleccionadas por el modelo farmacofórico fueron empleadas como ligandos. Los ligandos fueron preparados mediante la adición de hidrógenos explícitos, de estados tautoméricos a pH 7.4, y generando coordenadas 3D con *Standardizer* v21.8.0 [203].

El acoplamiento molecular se realizó empleando el software GOLD del *Cambridge Crystallographic Data Center* (CCDC) [154]. El sitio de búsqueda se definió empleando a la Cys145 como punto céntrico y extendiendo la región de búsqueda 15 Å de radio. Para la búsqueda conformacional por ligando, se realizaron 50 ejecuciones del algoritmo genético. Cada pose fue evaluada empleando la función de PLP *Chemscore* establecida en el software GOLD. Finalmente, para cada candidato se guardaron las 50 mejores poses para su posterior análisis.

5.8 Validación del Acoplamiento molecular

Para llevar a cabo esta validación se empleó el grupo de validación descrito previamente y el *software* de acoplamiento molecular GOLD. Para ello se empleó la proteasa definida como referencia (6M2N) y los ligandos

del grupo de validación previamente tratados. De igual forma se siguió el protocolo de acoplamiento molecular tal y como se enuncia en el apartado anterior (*5.7 Acoplamiento molecular*). Para cada ligando del grupo de validación se generaron 50 confórmeros.

Tras la realización de los acoplamientos moleculares, se estudiaron las 50 poses por ligando y se determinó la pose mayoritaria con el fin de determinar si está correspondía a la pose cristalográfica. Para esto se realizaron alineamientos entre el confórmero generado del acoplamiento molecular y la pose cristalográfica encontrada en el grupo de validación. Finalmente, se consideraron aquellos confórmeros con un $\text{RMSD} \leq 3.5 \text{ \AA}$ para el establecimiento de la pose mayoritaria.

VI. Resultados

Se estudiaron un total de 181 cristales relacionados con los indicadores "3CLpro" y "R1A_SARS2" derivados de UniProt [196], de los cuales 157 correspondían a complejos proteína-ligando, de los cuales 94 de ellos contenían inhibidores del tipo covalentes y 63 del tipo no covalentes. Para fines del proyecto se emplearon únicamente 130 de los 157 complejos proteína-ligando, esto debido a que algunos ligandos se encontraban en una región diferente a la del sitio activo en 3CLpro.

6.1 Alineamiento de secuencia

Mediante un estudio de alineamiento de secuencia se determinó que el porcentaje de identidad entre los 181 cristales fue del 96-100%, resultados que pueden ser visualizados en la Figura 10.

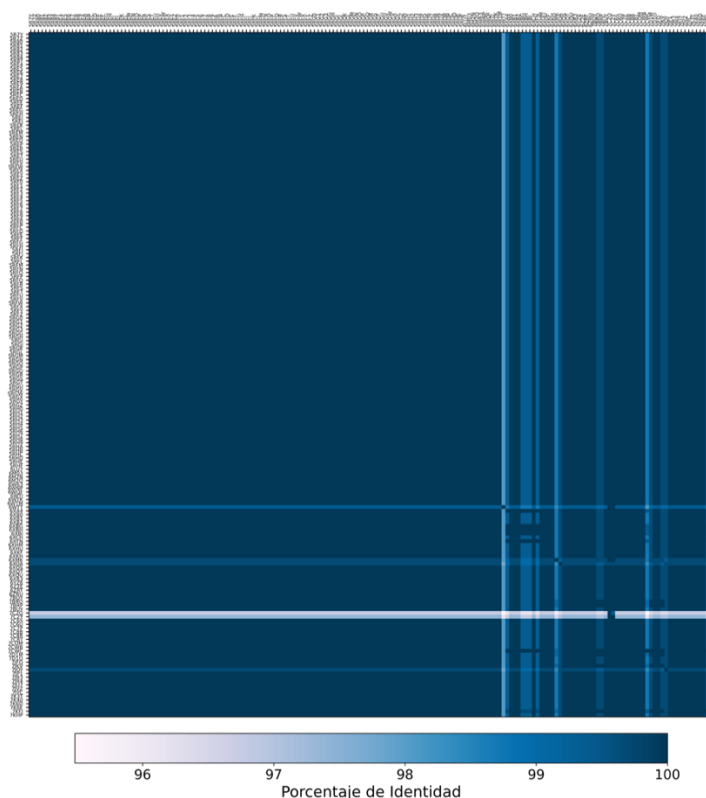


Figura 10. Alineamiento de secuencia de diversas 3CLpro de SARS-CoV-2.

Porcentaje de identidad derivado del alineamiento de secuencia entre las diversas estructuras cristalográficas disponibles de 3CLpro de SARS-CoV-2.

6.2 Modelo farmacofórico 3D

Para cada ligando cocrystalizado extraído de los complejos proteína-ligando de 3CLpro se generó un modelo farmacofórico 3D basado en el receptor. Se consideraron 130 ligandos a partir de los cuales se identificó un total de 522 descriptores farmacofóricos en el contexto de la dída

catalítica, dentro de los cuales se encontraron: a) 289 descriptores del tipo aceptor de puente de hidrógeno (AH), b) 179 descriptores del tipo donador de puente de hidrógeno (DH), y c) 54 descriptores del tipo hidrofóbico (Hyd). Ver Figura 11.

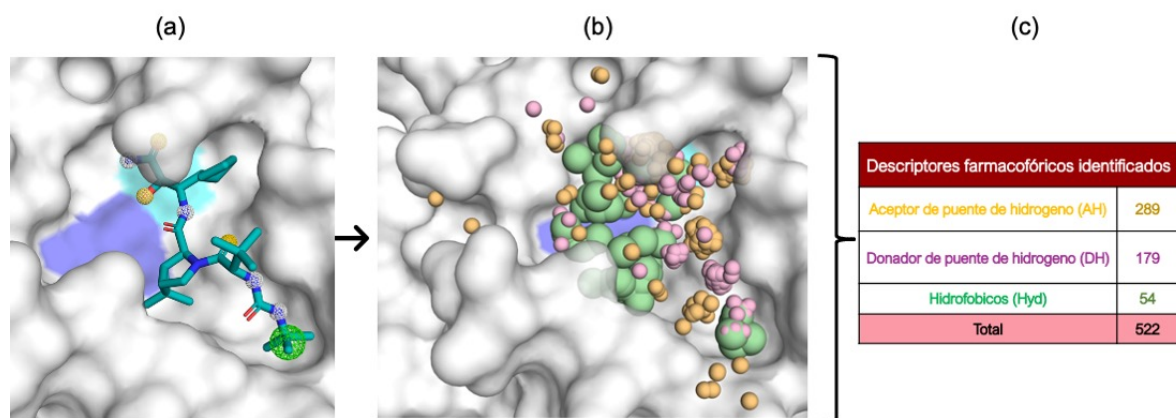


Figura 11. Esquemmatización de la determinación de modelos farmacofóricos 3D de ligandos cocrystalizados con 3CLpro.

(11a) Descriptores farmacofóricos del Boceprevir en el contexto del sitio activo de 3CLpro (PDB: 6ZRU). **(11b)** Conjunto de descriptores derivados a partir de 130 ligandos cocrystalizados con 3CLpro. **(11c)** Tabla que resume la cantidad y el tipo de descriptores determinados. El código de colores corresponde a amarillo para los descriptores AH, rosa para los descriptores DH y verde para los descriptores Hyd.

6.3 Agrupamiento jerárquico y MFC

Considerando el gran cumulo de información obtenida y la falta de una adecuada representación, se realizó un agrupamiento jerárquico basado en la distancia entre los descriptores farmacofóricos. La esquematización de la generación de los centroides por tipo de descriptor se visualiza en las Figuras 12, 13 y 14.

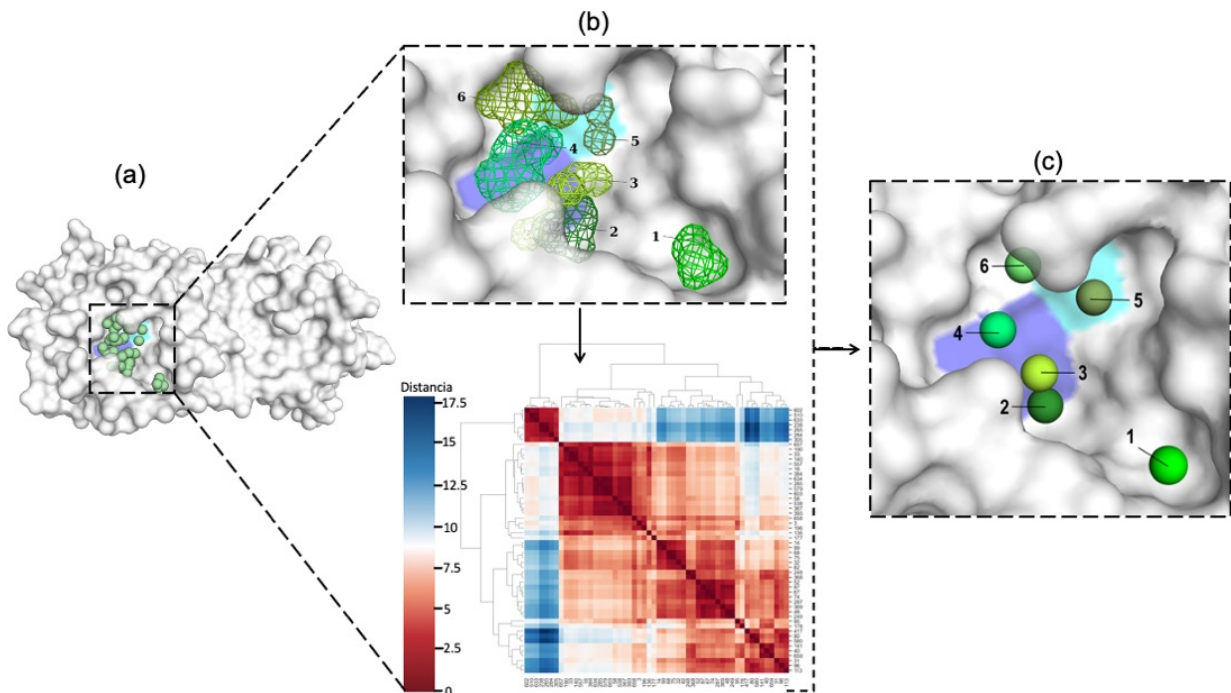


Figura 12. Esquema de la generación de centroides Hyd.

Los 54 descriptores del tipo Hyd identificados (**12a**), se estudiaron y analizaron. Considerando la estructura en red formada y el *heatmap* (**12b**) se establecieron 6 centroides (**12c**). En azul y cian se visualiza la diada catalítica conformado por la His41 y Cys145, respectivamente.

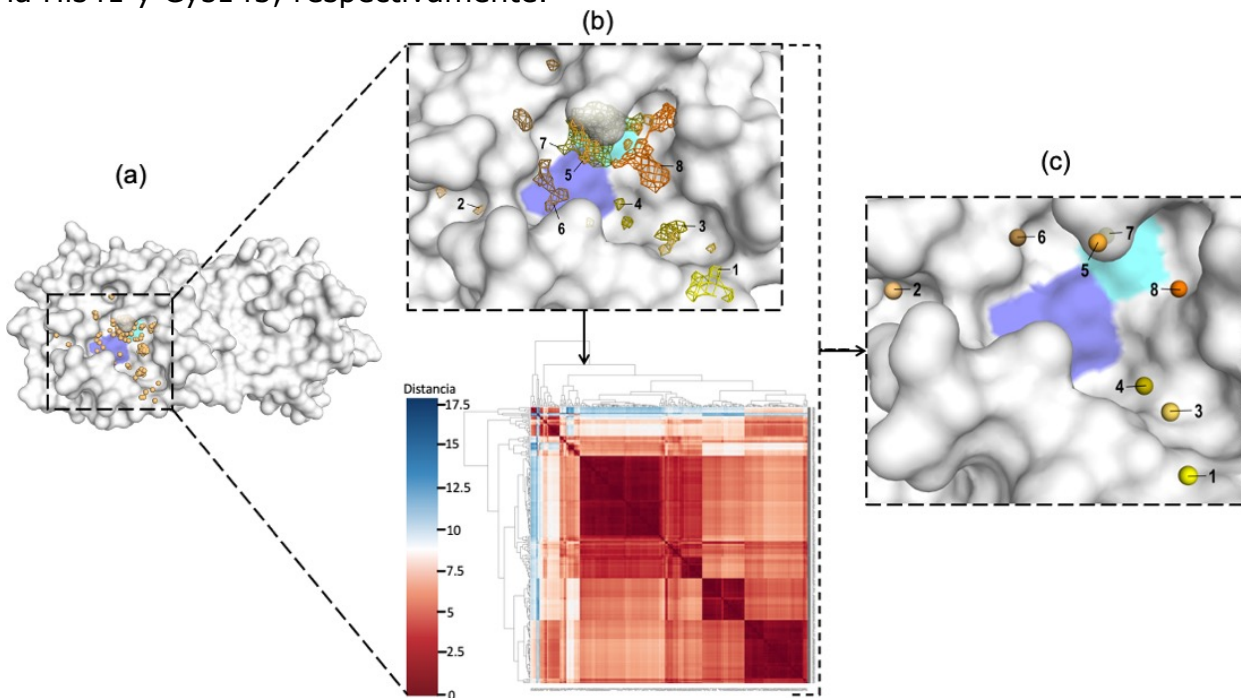


Figura 13. Esquema de la generación de centroides AH.

Los 289 descriptores del tipo AH identificados (**13a**), se estudiaron y analizaron. Considerando la estructura en red formada y el *heatmap* (**13b**) se establecieron

8 centroides (**13c**). En azul y cian se visualiza la diada catalítica conformado por la His41 y Cys145, respectivamente.

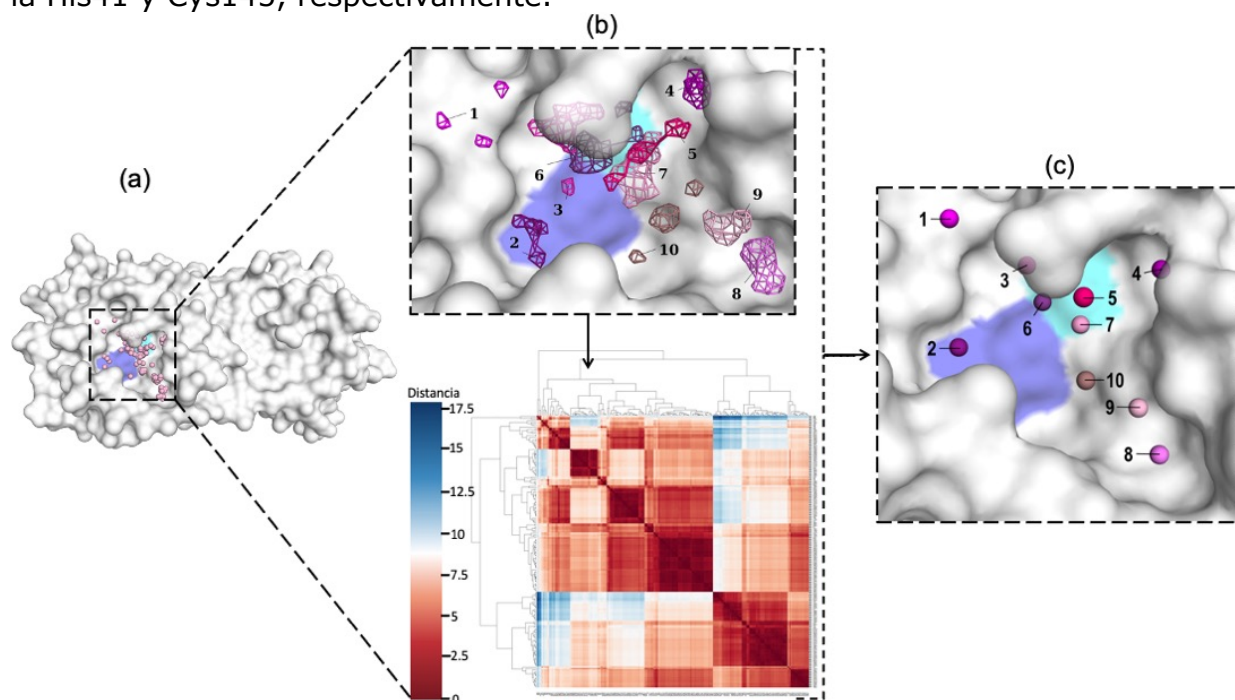


Figura 14. Esquema de la generación de centroides DH.

Los 179 descriptores del tipo DH identificados (**14a**), se estudiaron y analizaron. Considerando la estructura en red formada y el *heatmap* (**14b**) se establecieron 10 centroides (**14c**). En azul y cian se visualiza la diada catalítica conformado por la His41 y Cys145, respectivamente.

De esta manera se pasó de contar con 522 descriptores farmacofóricos a tan sólo 24 centroides representativos, resultados que se muestran en la Tabla 1.

Tabla 1. Descriptores farmacofóricos y centroides del MFC.

Descriptores		Grupos generados	Número de Centroides
Tipo	Cantidad		
Hyd	54	6	6
DH	179	10	10
AH	289	8	8
Total	522	24	24

Las abreviaciones para el tipo de descriptor corresponden a Hyd: hidrofóbico, DH: donador de puente de hidrógeno, AH: aceptor de puente de hidrógeno.

Finalmente, considerando todos los centroides antes mencionados, se generó el MFC el cual, como ya se mencionó, partió de considerar 522 descriptores derivados de 130 ligandos en complejos cocrystalizados con 3CLpro, para finalmente contemplar 24 centroides que representan la complejidad del conjunto de ligandos en el contexto del receptor y su sitio activo. La Figura 15 ilustra la generación de este MFC.

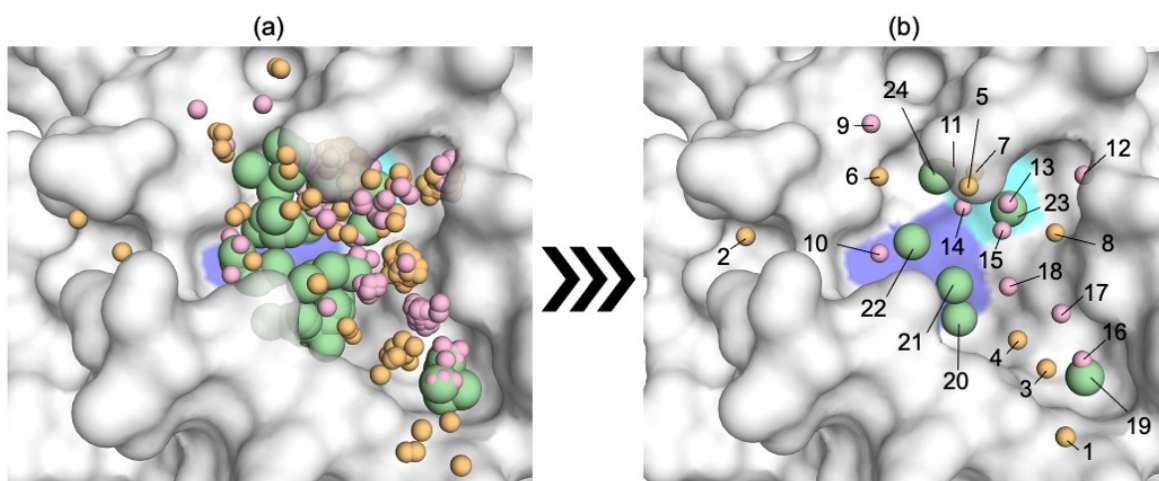


Figura 15. Generación del MFC.

Los 522 descriptores farmacofóricos identificados (**15a**) (se empleó el mismo código de color como en la Figura 10) se redujeron a 24 centroides (**15b**) de acuerdo con el agrupamiento jerárquico de su posición espacial.

6.4 Validación del MFC

Para llevar a cabo la validación del MFC se estableció un grupo de ligandos cocrystalizados. Para realizar lo anterior se empleó un proceso de filtrado partiendo de 177 complejos ligando-proteína en el contexto del sitio activo de 3CLpro. De éstos, se descartaron los ligandos duplicados dejando un total de 152 ligandos únicos. Posteriormente, se consideraron aquellos ligandos que contaban con más de un enlace rotatable, más de

tres puntos farmacofóricos y un peso molecular (MolWT, por sus siglas en inglés) mayor a 200 Da, dando lugar a un grupo de validación con 78 ligandos únicos.

La distribución de los ligandos pertenecientes al grupo de validación en relación con su MolWT, el número de enlaces rotables y número de puntos farmacofóricos se observa en la Figura 16.

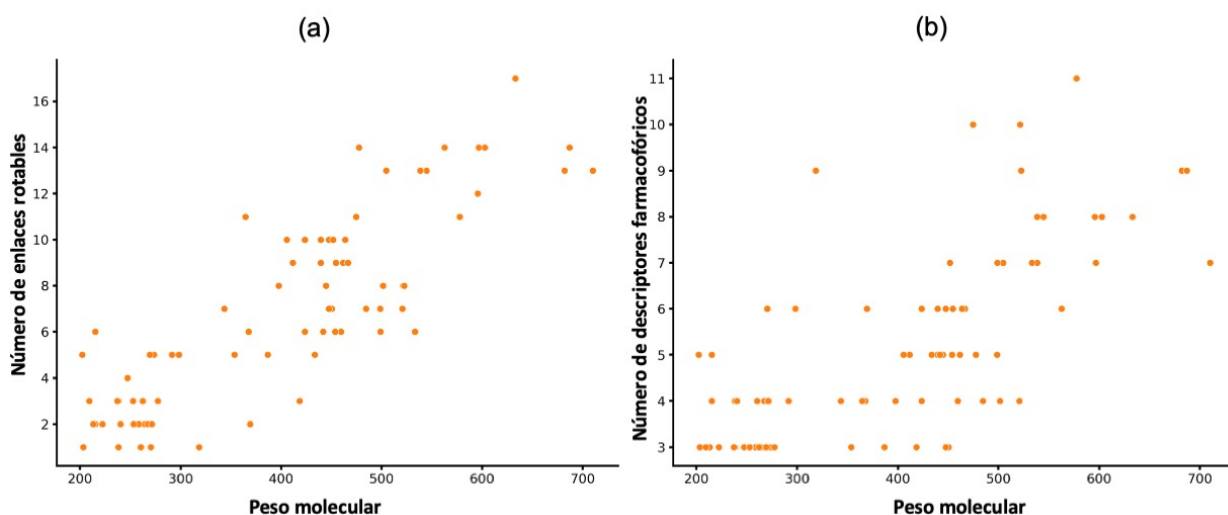


Figura 16. Distribución de los ligandos del grupo de validación con respecto a algunos de sus descriptores moleculares.

(16a) Se relacionó el MolWT con el número de enlaces rotables que poseen los ligandos del grupo de validación, **(16b)** así como el MolWT con el número de descriptores farmacofóricos que poseen los mismos.

A partir de los 78 ligandos únicos que conformaban el grupo de validación, se generó una librería de 19,500 conformeros con el fin de capturar las poses cristalográficas.

Finalmente, de un total de 78 ligandos se logró capturar la pose cristalográfica con un $\text{RMSD} \leq 2.5 \text{ \AA}$ de 31 ligandos, que mediante el empleo de la ecuación %Eficiencia, previamente definida, se determinó

que este modelo farmacofórico consenso presentó un 39.7% de eficiencia, es decir, logró captura la pose cristalográfica de aproximadamente 4 de cada 10 ligandos.

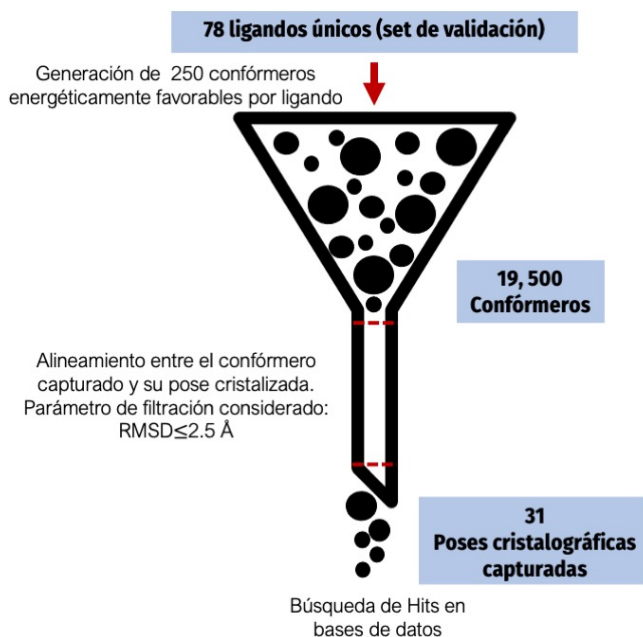
$$\%Eficiencia = \frac{\#M}{SV} * 100 \rightarrow \%Eficiencia = \frac{31}{78} * 100$$

$$\%Eficiencia = 39.7 \%$$

El proceso completo de la validación del MFC se visualiza en el Esquema

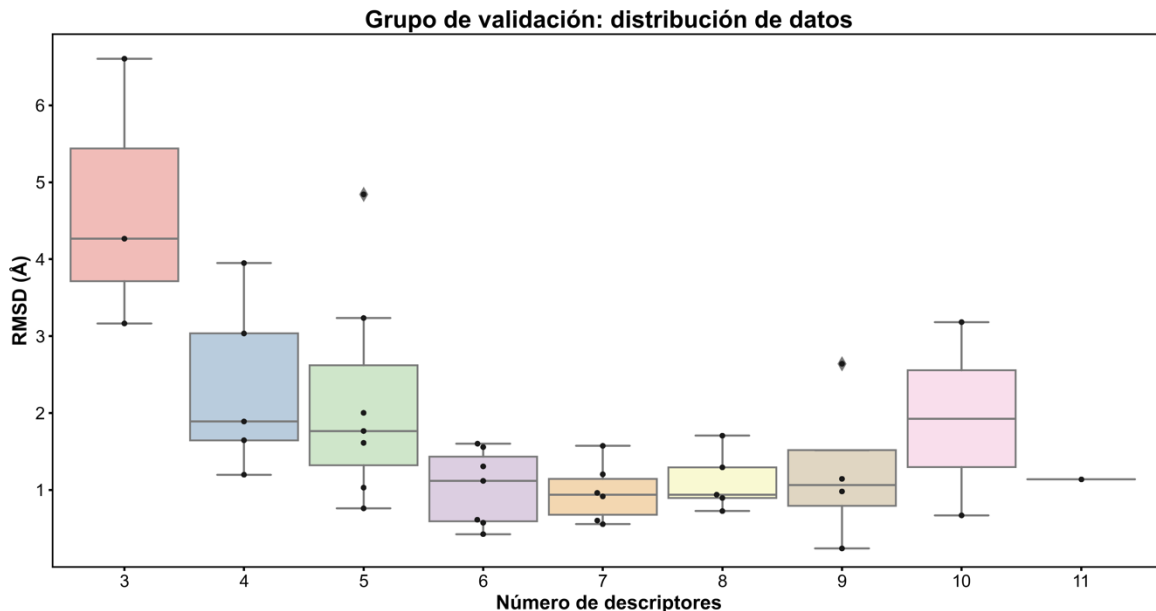
1.

Esquema 1. Proceso de validación del MFC.



Los resultados se organizaron y representaron en la Figura 17. En la Figura 17a se observa la distribución de datos de los ligandos relacionando su RMSD con respecto a la cantidad de descriptores que presentaron. En la Figura 17b se observa la relación del RMSD con cada molécula.

(a)



(b)

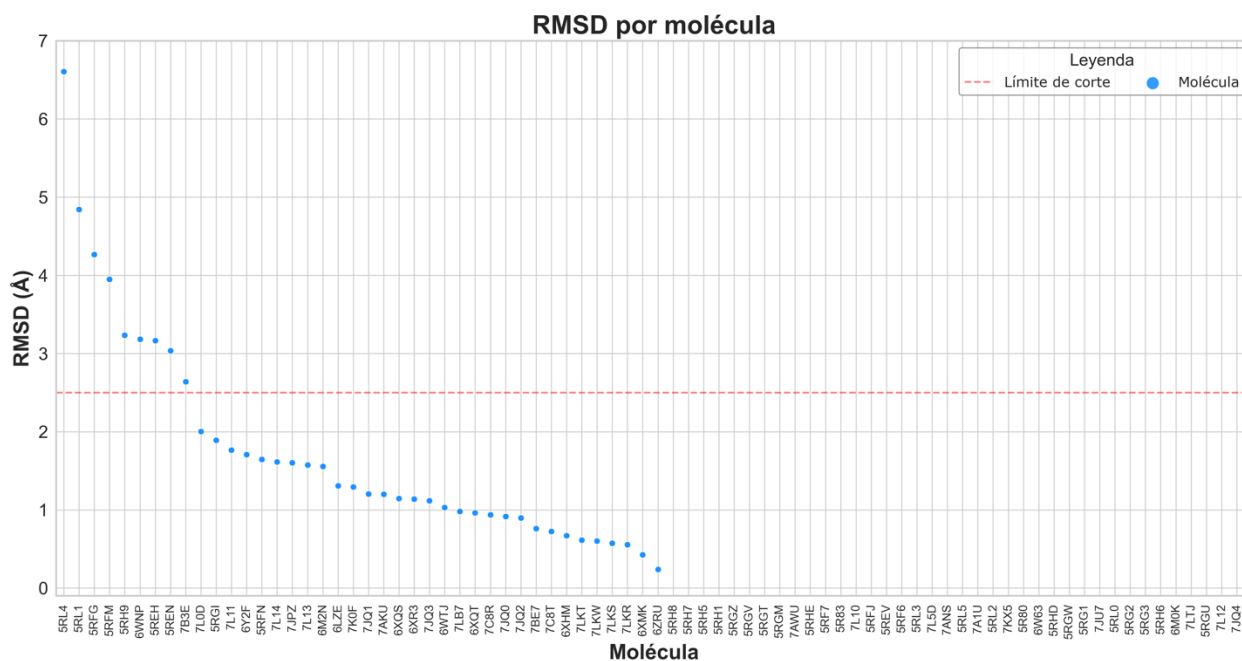


Figura 17. Resultados de validación del MFC.

Se organizaron los datos cristalográficos con respecto a la cantidad de descriptores con la que cada ligando contó **(17a)** en relación con el RMSD derivado de la comparación entre la pose capturada y la cristalográfica **(17b)**.

6.5 Submodelos Farmacofóricos y búsqueda de potenciales candidatos

Con el objetivo de identificar potenciales candidatos se cotejaron diferentes librerías químicas. En la Figura 18 se observa la composición de moléculas y conformeros que construyen cada una de las librerías químicas cotejadas.

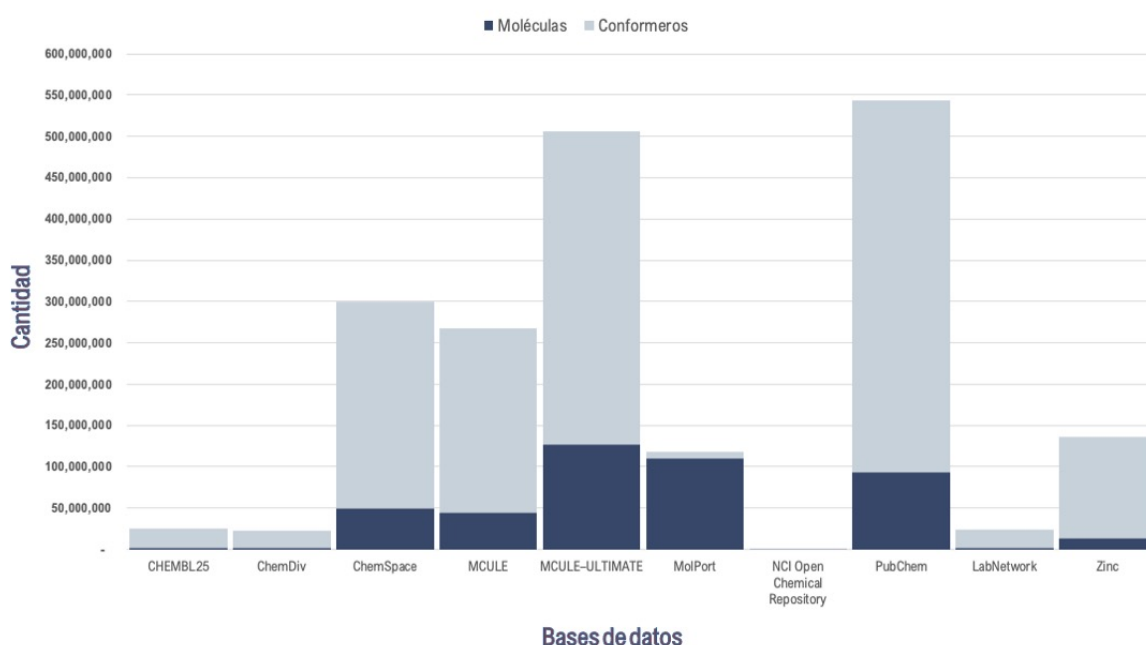


Figura 18. Librerías químicas cotejadas.

Se observan dos barras de diferente color por base de datos, por un lado, el color oscuro refleja la cantidad de moléculas mientras que, por el otro, el color claro refleja la cantidad de conformeros con los que la base de datos se encuentra constituida. Las bases cotejadas fueron CHEMBL25, ChemDiv, ChemSpace, MCULE, MCULE-ULTIMATE, MolPort, NCI Open Chemical Repository, PubChem, LabNetwork y Zinc.

Se generaron 13 submodelos farmacofóricos diferentes que exploran la cavidad y priorizan los centroides con respecto a los descriptores químicos con mayor frecuencia de aparición. Esta priorización se realizó con base

en el *expertis* químico, es decir, con base en la experiencia adquirida tras el estudio, análisis y visualización de la naturaleza de la cavidad y de sus respectivos ligandos.

Cabe destacar que estos submodelos contaban con 6 o 7 centroides. Las características de estos submodelos farmacofóricos se encuentran enlistadas en la Tabla 2.

En lo que respecta a la búsqueda de candidatos cuyos descriptores químicos coincidieron con los centroides presentados por los submodelos propuestos, se obtuvo un total de 1354 candidatos. Estos resultados se enlistan en la Tabla 3.

Tabla 2. Centroides por submodelo farmacofórico.

Centroides		Sub Modelo Farmacofórico												
Tipo	Número	B1	B2	D1	D3	D4	D5	D6	D7	D8	D9	E1	E2	E3
HA	1	•	•	•	•	•	•	•	•	•	•	•	•	•
	2	•	•	•	•	•	•	•	•	•	•	•	•	•
	3	•	•	•	•	•	•	•	•	•	•	•	•	•
	4	•	•	•	•	•	•	•	•	•	•	•	•	•
	5	•	•	•	•	•	•	•	•	•	•	•	•	•
	6	•	•	•	•	•	•	•	•	•	•	•	•	•
	7	•	•	•	•	•	•	•	•	•	•	•	•	•
	8	•	•	•	•	•	•	•	•	•	•	•	•	•
	9	•	•	•	•	•	•	•	•	•	•	•	•	•
HD	10	•	•	•	•	•	•	•	•	•	•	•	•	•
	11	•	•	•	•	•	•	•	•	•	•	•	•	•
	12	•	•	•	•	•	•	•	•	•	•	•	•	•
	13	•	•	•	•	•	•	•	•	•	•	•	•	•
	14	•	•	•	•	•	•	•	•	•	•	•	•	•
	15	•	•	•	•	•	•	•	•	•	•	•	•	•
	16	•	•	•	•	•	•	•	•	•	•	•	•	•
	17	•	•	•	•	•	•	•	•	•	•	•	•	•
	18	•	•	•	•	•	•	•	•	•	•	•	•	•
	19	•	•	•	•	•	•	•	•	•	•	•	•	•
Hyd	20	•	•	•	•	•	•	•	•	•	•	•	•	•
	21	•	•	•	•	•	•	•	•	•	•	•	•	•
	22	•	•	•	•	•	•	•	•	•	•	•	•	•
	23	•	•	•	•	•	•	•	•	•	•	•	•	•
	24	•	•	•	•	•	•	•	•	•	•	•	•	•
Centroides usados		7	6	6	6	6	7	6	6	6	7	6	7	6

Tabla 3. Candidatos obtenidos por Submodelo farmacofórico.

Base de datos	Sub modelo farmacofórico												
	B1	B2	D1	D3	D4	D5	D6	D7	D8	D9	E1	E2	E3
ChEMBL25	0	8	0	0	0	0	0	0	4	0	0	0	3
ChemDiv	0	2	0	0	0	0	0	0	0	0	0	0	1
ChemSpace	0	6	0	1	0	0	0	0	69	0	0	4	0
MCULE	0	4	0	2	0	0	0	0	46	0	0	17	1
MCULE-ULTIMATE	0	2	0	0	5	0	1	0	0	0	0	0	2
MolPort	0	5	2	3	2	0	0	0	8	0	0	1	2
NCI	0	0	0	0	0	0	0	0	0	0	0	0	0
PubChem	1	51	0	4	4	1	23	3	1004	8	1	12	22
LabNetwork	0	1	0	1	0	0	0	0	0	0	0	0	0
Zinc	0	2	1	1	0	0	1	0	6	0	0	0	4
Hits	1	81	3	12	13	1	25	3	1137	8	1	34	35
Hits totales	1354												

Los submodelos reciben sus nombres respectivos debido a que derivan de una lista extensa de submodelos farmacofóricos que contemplaban una amplia variedad de combinaciones, aquí únicamente se anexan aquellos que se

consideraron de importancia y a través de los cuales se obtuvieron potenciales candidatos. Los descriptores farmacofóricos considerados para cada submodelo se marcan con un círculo color verde.

6.6 Optimización local y filtrado

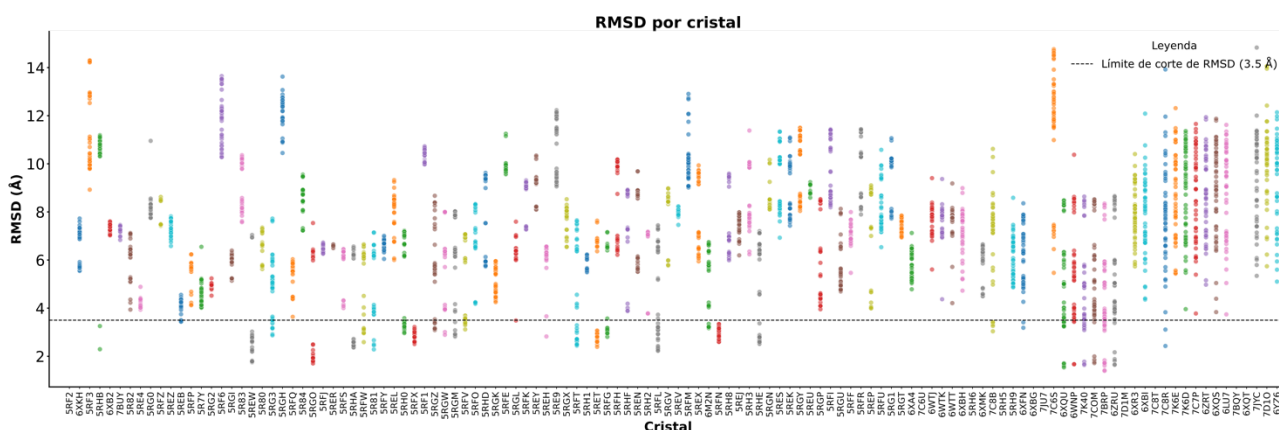
Tras la obtención de los candidatos, se llevaron a cabo los procesos de optimización local y filtrado, por medio de los cuales, de contar con 1354 candidatos, la lista finalmente se redujo a sólo 312 candidatos. El proceso de optimización local consistió en encontrar una disposición espacial de la colección de átomos que describen el *hit* en el contexto de la proteína donde la red entre la fuerza atómica en cada átomo es aceptablemente cercana a cero y la posición en la superficie de energía potencial es un punto estacionario. La intención de realizar dicha optimización es la importancia física de la estructura del *hit* obtenida y debido a que a menudo las estructuras optimizadas corresponden a las estructuras naturales [204].

El proceso de filtrado consistió en descartar aquellos ligandos que se encontraban repetidos o que, tras su optimización local, dejaban de alinearse en contra del MFC o continuaban generando choques con la 3CLpro.

6.7 Validación del acoplamiento molecular

El procedimiento de validación de acoplamiento molecular consistió en determinar la capacidad con la que 1) el protocolo empleado lograba acoplar los candidatos a 3CLpro y 2) reproducía las poses cristalográficas de los ligandos extraídos de los complejos proteína-ligando recabados con anterioridad. Para llevar a cabo lo antes descrito se usó el grupo de validación descrito previamente y el *software* de acoplamiento molecular *Gold*. Los resultados de cada pose por cristal considerando su RMSD se visualizan en la Figura 20. Finalmente, se determinó que únicamente fue posible reproducir la pose cristalográfica con alta frecuencia de 7 ligandos de un total de 78 ligandos, siendo el protocolo de acoplamiento molecular un 8.9% eficaz a la hora de reproducir una pose cristalográfica para este sistema. Cabe destacar que se consideró alta frecuencia si al menos 20 poses reproducían la pose cristalográfica (ver Figura 19).

(a)



(b)

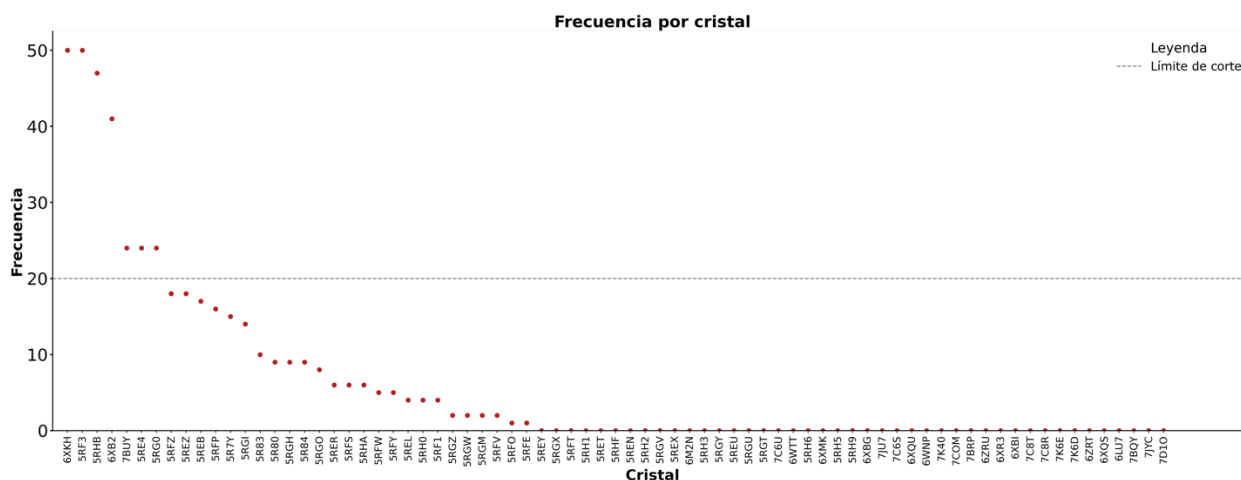


Figura 19. Validación de acoplamiento molecular.

En la Figura 19a se muestra la relación entre el RMSD por cada pose de cada cristal, considerando que los resultados están organizados de menor a mayor MolWT, yendo de izquierda a derecha. En la Figura 19b se muestra la relación entre la frecuencia con la que las poses presenta un $\text{RMSD} \leq 3.5 \text{ \AA}$ conforme a cada cristal.

6.8 Acoplamiento molecular

Los 312 candidatos obtenidos tras el filtrado fueron acoplados a 3CLpro mediante el empleo del *software* Gold. Se generaron 50 poses por cada uno de los candidatos. Los resultados de RMSD y Score por candidato se visualizan en la Figura 20.

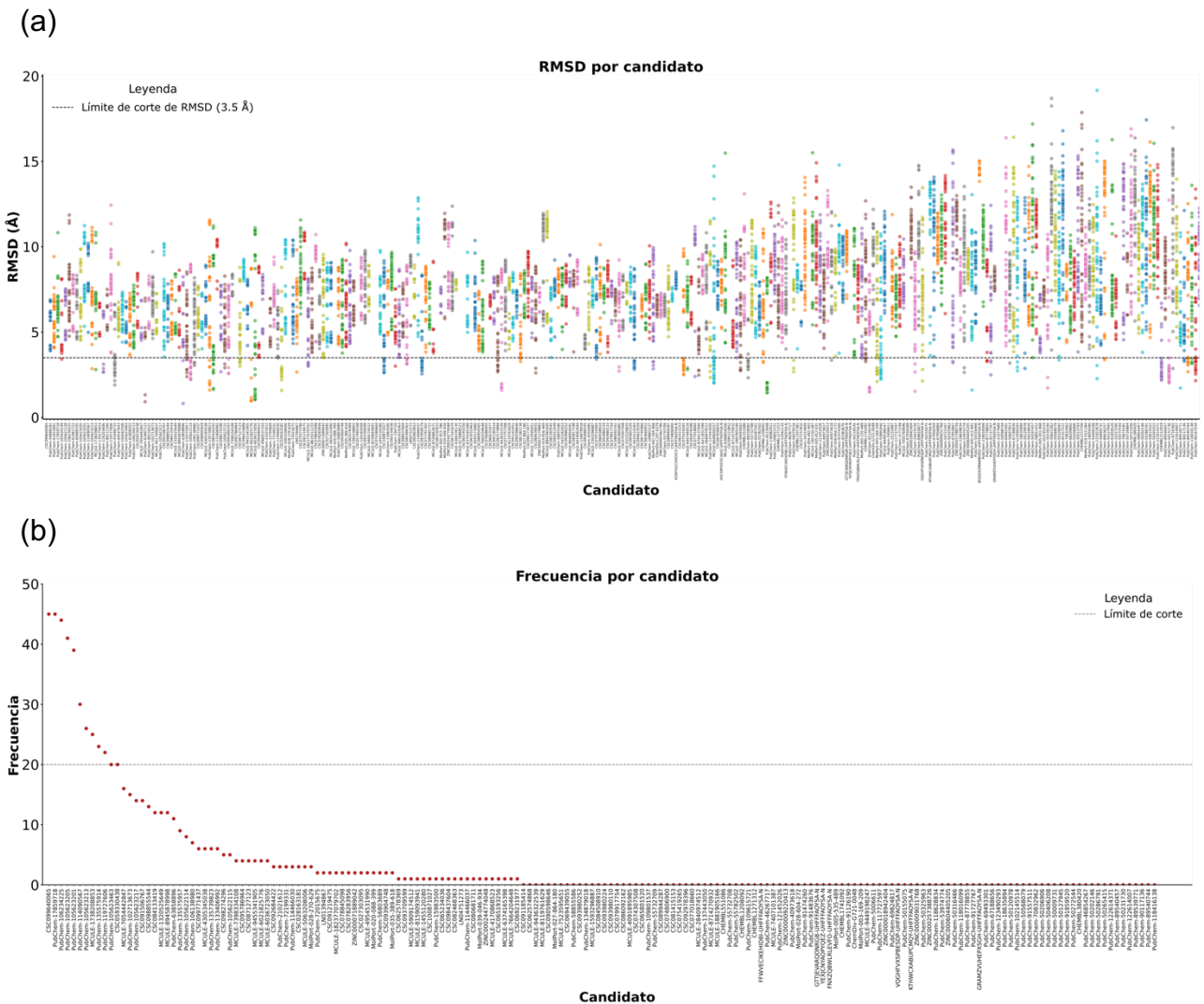


Figura 20. Resultados del acoplamiento molecular.

En la figura **20a** se muestra la relación entre el RMSD por cada pose de cada candidato, considerando que los resultados están organizados de menor a mayor MolWT, yendo de izquierda a derecha. En la figura **20b** se muestra la relación entre la frecuencia con la que las poses presentan un $\text{RMSD} \leq 3.5 \text{ \AA}$ conforme a cada candidato.

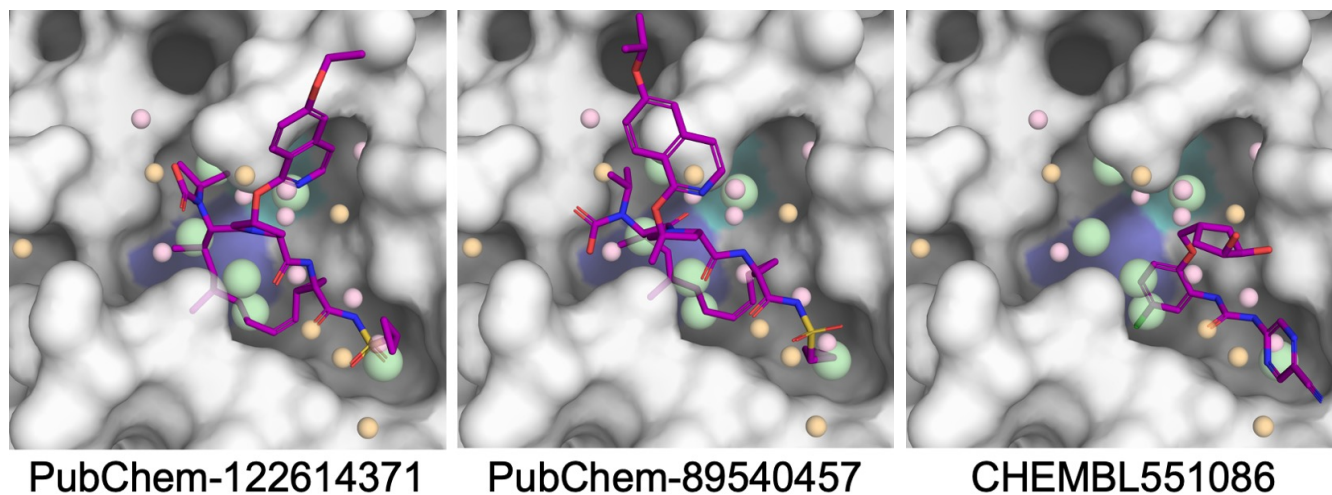
La lista de los 12 candidatos catalogados como los mejores, los cuales contaron con poses con $\text{RMSD} \leq 3.5 \text{ \AA}$, junto con sus respectivas frecuencias y scores se enlistan en la Tabla 4.

Tabla 4. Mejores candidatos derivados del acoplamiento molecular.

Software: GOLD					
ID	Candidato	Frecuencia	Pose	RMSD (Å)	Score
1	PubChem-89540457	45	24	2.02	68.29
2	PubChem-122614371	45	49	2.19	57.55
3	CHEMBL551086	44	49	2.32	58.85
4	PubChem-114466037	41	39	2.56	45.52
5	PubChem-114466030	39	12	1.58	45.53
6	PubChem-135575957	30	10	2.22	55.58
7	CSC081415153	26	7	2.49	58.39
8	PubChem-137443050	25	32	2.02	65.80
9	PubChem-119727606	23	26	1.90	43.47
10	PubChem-118416138	22	26	2.16	48.51
11	MolPort-003-169-209	20	37	2.24	51.67
12	CSC091185414	20	29	2.67	58.20

Se enlistan los mejores 12 candidatos resultantes del acoplamiento molecular con base en su frecuencia. La frecuencia hace referencia a aquella con la que aparecen poses con un $\text{RMSD} \leq 3.5 \text{ \AA}$. La columna de pose, RMSD (Å) y Score hacen referencia a la pose con un menor RMSD, ese RMSD y la calificación asignada por GOLD a dicha pose, respectivamente.

Los mejores candidatos abordados con anterioridad se visualizan en la Figura 21 y se identifican con respecto al nombre del candidato.



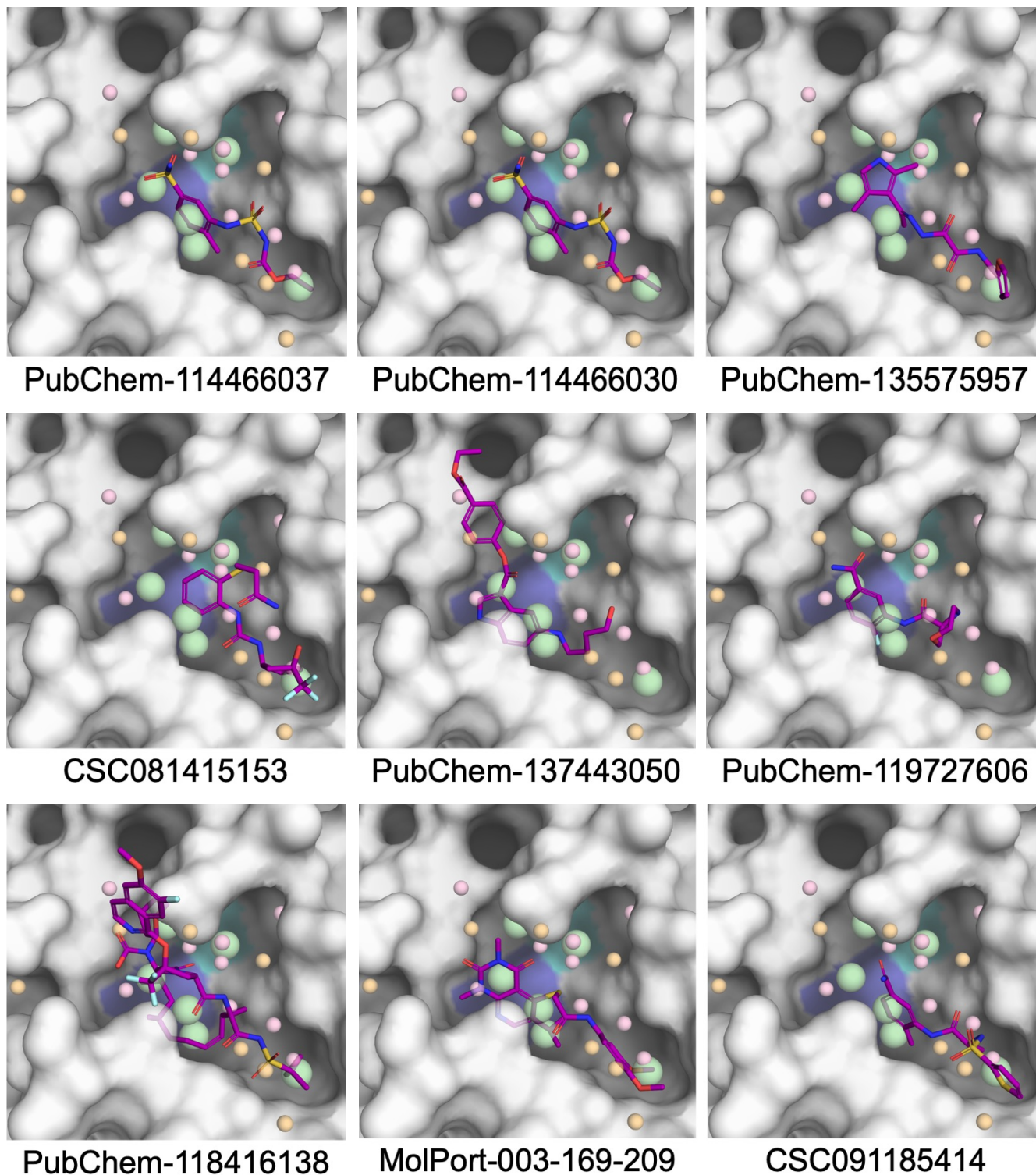


Figura 21. Mejores candidatos en el contexto de 3CLpro.

Los esqueletos de los candidatos se visualizan de color magenta, mientras que la diada catalítica, conformada por la His41 y Cys145, esta coloreada de azul y cian, respectivamente. Los descriptores farmacofóricos mantienen la misma relación de colores que ha sido marcada anteriormente.

VII. Discusión

La enfermedad conocida como COVID-19, causada por SARS-CoV-2, ha sido la responsable de más de 5.2 millones de muertes alrededor del mundo y más de 262.1 millones de casos confirmados de contagios, desde sus inicios, hasta finales del 2021 [6]. Se ha determinado a 3CLpro como uno de los principales blancos terapéuticos para la generación de candidatos con potencial actividad inhibitoria contra SARS-CoV-2 debido a su función crucial dentro del ciclo replicativo [19] (ver Figura 4 y 6).

Se han identificado una amplia variedad de potenciales inhibidores de 3CLpro que interactúan con el sitio activo, sin embargo, la actividad de estos fármacos para prevenir o reducir la severidad de los síntomas de la COVID-19 no ha sido completamente establecida [77], por lo que prevalece la urgencia del desarrollo e identificación de candidatos farmacológicos adicionales cuyos blancos terapéuticos en SARS-CoV-2 sean eficaces en el tratamiento de la COVID-19 [78].

Con la finalidad de identificar nuevos candidatos con potencial actividad inhibitoria de la 3CLpro, se llevó a cabo este proyecto para lo cual se recabó toda la información cristalográfica existente relacionada a la 3CLpro. Tras un estudio de alineamiento de secuencia se determinó que el porcentaje de identidad entre las 181 estructuras relacionadas con 3CLpro de SARS-CoV-2 (complejos proteína-ligando covalente, no

covalente y en forma APO) era del 96% al 100%. Del mismo modo se determinó que el sitio activo conformado por la His41 y Cys145 se encuentra altamente conservado. Los dos puntos previos indican que esta proteasa, así como su sitio activo se encuentran altamente conservados y, a su vez, confirma lo que otros estudios han reportado con anterioridad (ver Figura 4, 6 y 10) [10,205].

Tras los experimentos anteriores se identificaron dos puntos importantes: 1) los inhibidores del tipo covalente y no covalente en el sitio activo de 3CLpro no generan cambios conformacionales de importancia en la proteasa (ver Figura 10) y 2) el sitio con mayor fluctuación corresponde al tercer dominio, lugar donde no reside el dominio catalítico. Esto implica que el sitio activo de 3CLpro permanece constante en todos los sistemas estudiados sin importar el tipo interacciones que éstos presentan.

A continuación, se extrajeron los ligandos ubicados dentro de la diada catalítica de 3CLpro, a partir de los cuales se generaron, a su vez, modelos farmacofóricos individuales.

Un modelo farmacofórico, como ya se mencionó, es el ensamble de descriptores farmacofóricos necesarios para asegurar las interacciones entre un ligando con un blanco terapéutico [168]. Dicho modelo puede establecerse a través de dos enfoques diferentes, el LBP y el SBP [168].

El enfoque SBP se empleó en el proyecto ya que se contaba con una amplia evidencia experimental de complejos proteína-ligando. A través de

este enfoque se exploraron los posibles puntos de interacciones entre las estructuras 3D de 3CLpro y los diversos ligandos cocrystalizados [168]. Cabe destacar que el SBP se ha convertido en una estrategia computacional clave debido a que facilita el descubrimiento de fármacos considerando ensambles experimentales [168]. El empleo de la metodología SBP en 3CLpro es interesante, por la gran cantidad de información experimental existente derivada de la situación de emergencia internacional, ya que ésta nos permitió identificar puntos de suma importancia tales como la preferencia de ciertos tipos de descriptores farmacofóricos con ciertos residuos de 3CLpro, la naturaleza de los subsitios en 3CLpro y la frecuencia con la que algún tipo de interacción se presentaba.

Aplicando el enfoque SBP se generaron 130 modelos farmacofóricos a partir de los ligandos obtenidos, los cuales en conjunto se encontraban compuestos por un total de 522 descriptores farmacofóricos (ver Figura 11). Dentro de éstos, el descriptor con mayor porcentaje de aparición fue del tipo AH 55.4% (289), mientras que el 34,3% (179) y el 10.3% (54) restante, correspondían a los descriptores del tipo DH e Hyd, respectivamente. En el mismo sentido, la mayoría de los descriptores del tipo DH se presentaron con mayor frecuencia entre la cavidad S1'-S2. Estos resultados concuerdan con la especificidad de la proteasa, dado que

en esta zona (S1'-S2) sucede el ataque nucleofílico de la Cys-145 al sustrato, reacción que es catalizada por la formación de un puente de hidrógeno entre la His41 con el sustrato. Como consecuencia, la presencia de descriptores del tipo DH en dicha cavidad es de suma importancia [206].

En vista de la gran información cristalográfica obtenida a partir de los ligandos presentes en la cavidad catalítica de 3CLpro se estableció la metodología del MFC. Para este proyecto el MFC puede ser definido como el ensamble de descriptores farmacofóricos óptimos que representan eficientemente a múltiples modelos farmacofóricos en el contexto de un sitio de interés de importancia biológica. Dicha metodología tiene como objetivo el descubrimiento de potenciales fármacos a partir de información cristalográfica existente sobre un sistema proteína-ligando. Cabe destacar que el MFC tiene una serie de ventajas que ilustran su potencial frente a las estrategias convencionales de VS como el modelo farmacofórico y el acoplamiento molecular, tales como 1) se encuentra basado en una amplia evidencia experimental de una variedad de ligandos en el contexto de su receptor, 2) tiene el poder de ser representativo ya que permite considerar la frecuencia de aparición de los descriptores farmacofóricos y la frecuencia de interacción de éstos con los residuos proteicos, 3) permite identificar la naturaleza de la diada catalítica, y 4)

empleando un MFC se combate el problema más frecuentemente identificado para el SBP, ya que para este último se pueden identificar demasiadas características químicas de naturaleza subyacente [168].

Lo anterior justifica el uso del MFC, gracias a su gran potencial representativo, descriptivo e innovador. Con esta metodología se pasó de contar con 522 descriptores farmacofóricos a tan sólo 24 centroides representativos por zona (ver Figura 15).

Cuando el MFC se enfrentó al grupo de validación, se determinó que el porcentaje de eficiencia fue del 39.7% (con un $\text{RMSD} \leq 2.5 \text{ \AA}$), indicando que con este modelo es posible capturar la pose cristalográfica de aproximadamente 4 de cada 10 ligandos, sin perder las características del MFC (ver Figura 17). Aquí es pertinente hacer mención que con base en los puntos antes mencionados para el MFC y considerando que 1) se partió de información experimental y, 2) se contempló la representatividad por zona, tipo de descriptor y frecuencia, empleando lo previamente denominado como *expertise* químico, se esperaba que tras la validación se lograra capturar una mayor cantidad de poses cristalográficas (al menos más del 60%).

Comparando los resultados obtenidos y los esperados para la validación del MFC, así como realizando el análisis del flujo de trabajo, es posible que existan características importantes que no fueron consideradas para abordar una amplia representatividad por inter e intragrupo de

descriptores en el contexto de 3CLpro y que a su vez definieron el porcentaje de eficiencia obtenido, dejando de lado la modificación de la matriz de distancia (ver Perspectivas). Lo anterior con base en una idea simple: si contamos con descriptores de diferente tipo, naturaleza y frecuencia de aparición, no es sensato establecer que las interacciones óptimas generadas por todos los descriptores tengan el mismo valor para la generación del MFC.

Cabe acentuar que este modelo, al contar con una validación como base, impone un punto de inicio a partir del cual el desarrollo y la optimización pueden ser llevados a cabo para este sistema, de modo que logre representarlo de manera óptima. Por otra parte, dado que no hay metodologías similares reportadas que cuenten con una validación asociada a este sistema, se hace difícil realizar una comparación exacta. Así mismo, esto coloca a la metodología del MFC como una metodología innovadora basada en evidencia experimental.

Tras la generación y validación del MFC en el contexto de 3CLpro, con el fin de generar potenciales candidatos, se realizó un VS. Para lo anterior se determinó que el MFC seguía siendo complejo y difícil de analizar, puesto que el empleo íntegro de dicho modelo en el VS empleando *Pharmit* no generó resultado alguno. Esto debido a que en el universo químico no existe entidad química tal que cuente con 24 descriptores

farmacofóricos particulares en esa disposición espacial determinada. Con base en lo anterior, se decidió fragmentar el MFC en submodelos.

Se probó una fracción de submodelos con un número variable de centroides, dentro de los cuales destacaron 13 submodelos que exploraban la cavidad y priorizaban los centroides con respecto a los descriptores químicos con mayor frecuencia de aparición. Para los 13 submodelos farmacofóricos destacados, se consideraron entre 6 y 7 centroides (ver Tabla 2) con base en 1) el estudio que Yang S-Y y colaboradores (2010), quienes en su informe destacaron que un modelo farmacofórico compuesto por más de 7 características químicas no es adecuado para aplicaciones prácticas (recomendando de 3 a 7 descriptores) [168], y 2) se buscaron candidatos que cubrieran de forma adecuada la cavidad, potenciando la mayor cantidad de interacciones proteína-ligando. Del mismo modo, los submodelos contemplaron en promedio de 2 a 3 centroides por cada tipo de descriptor farmacofórico empleado.

Los centroides más frecuentemente utilizados en los 13 submodelos estudiados fueron el 10, 14 y 18 (DH), el 6 y 8 (AH), así como el 22 y 24 (Hyd). Estos centroides y su uso concuerdan con las características que presentan los subsitios S1', S1 y S2 de 3CLpro. Por ejemplo, el centroide 10 (DH) se localizó sobre la His41 en el subsitio S1', el cual se caracteriza por contener la díada catalítica His41-Cys145 de 3CLpro [206]. Por otra

parte, los centroides 14 y 18 (DH), así como el 8 (AH) se localizan entre S1 y S1', próximos a His163 y Glu166, residuos que han sido reportados como participantes en la formación de puentes de hidrógeno esenciales para la actividad de 3CLpro [206]. Es con lo anterior que la presencia de aceptores y/o donadores de puente de hidrógeno entre S1'-S1 son de vital importancia en el desarrollo y diseño de fármacos inhibidores de 3CLpro. Con respecto a los centroides 22 y 24 (Hyd), éstos se localizan entre S2-S1', concordando con la naturaleza de las cavidades, ya que, por un lado S2 tiene la naturaleza de ser una cavidad hidrofóbica flexible, destacando que en este subsitio generalmente se prefieren los anillos de benceno o ciclopropilmetilo cuando se diseña la posición P2 de inhibidores peptidomiméticos [206]. Por el otro lado, derivado de estudios de interacción sobre 3CLpro, se observó que la presencia de sustituyentes que forman interacciones del tipo hidrofóbico en S1' son relevantes para la unión de un candidato a 3CLpro [95].

Cabe destacar que, acorde a las observaciones de Keretsu, S. y colaboradores (2020), la presencia de sustituyentes con el potencial de formar interacciones del tipo hidrofóbico y puente de hidrógeno con los residuos localizados en S1', pueden incrementar la energía de afinidad del candidato con 3CLpro al emular el patrón de unión al sustrato [95].

Finalmente, en cuanto al centroide 6 (AH), éste se ubica entre S1-S1', sin embargo, cabe señalar que este centroide engloba una amplia red de

descriptores, dentro de los cuales la mayoría se encuentran ubicados sobre la His41, sin embargo, tras la generación del centroide, éste se localizó en el punto medio de la red formada, quedando alejado de la díada catalítica (ver Figura 13). Este punto demuestra que el MFC tiene el potencial de optimizarse e incrementar el nivel de representatividad que presenta para el sistema 3CLpro (ver Perspectivas).

Es importante señalar que tras la obtención de los 24 centroides del MFC y la consecuente generación de submodelos farmacofóricos dentro de la cavidad de 3CLpro, se determinó que existen entre 1.3×10^5 y 3.4×10^5 combinaciones diferentes de submodelos farmacofóricos que consideran entre 6 y 7 centroides, respectivamente. Lo anterior indica que, sin tomar en cuenta los 13 submodelos farmacofóricos considerados para este proyecto, contamos con un extenso universo químico por explorar con la capacidad de tener candidatos con potencial actividad inhibitoria contra 3CLpro. Cabe recordar que estos 13 submodelos farmacofóricos fueron armados y seleccionados, en su mayoría, por poseer los centroides de mayor relevancia apoyandonos en el *expertise* químico.

Una vez establecidos los submodelos farmacofóricos y tras la realización del VS, se obtuvieron 1354 entidades químicas (moléculas y conformeros) derivadas de 10 librerías químicas diferentes (ver Figura 18 y Tabla 3).

Durante el VS estos submodelos mostraron volúmenes de exclusión, los cuales corresponden a zonas prohibidas para que un ligando se ubique en

ellas. La función de estos volúmenes de exclusión es la de simular los factores estéricos del sitio activo y, por ende, evitar que ligandos con geometrías inadecuadas para el sitio activo puedan ser considerados como candidatos [207].

En lo que respecta a los submodelos, el etiquetado como B2 generó resultados en todas las librerías químicas consultadas a excepción de NCI, representando aproximadamente el 6% de los candidatos totales derivados del VS, siendo a su vez el modelo que englobó la mayor parte de los centroides más frecuentemente empleados (6, 8, 10, 14 y 22) dentro del conjunto de submodelos considerados. Por otro lado, el submodelo D8 fue el que obtuvo el 84% del total de los candidatos derivados de las librerías químicas, siendo a su vez el modelo que presentó los centroides menos comunes dentro del conjunto de submodelos considerados. En el mismo sentido, la librería química que arrojó el 74% de los candidatos de D8 fue PubChem, una base de datos operada y mantenida por el National Center of Biotechnology Information (NCBI), la cual contiene más de 111 millones de estructuras químicas [208].

Contando con las 1354 entidades químicas, se realizó un proceso de optimización local y filtrado, tras el cual se pasó a contar con sólo 312 moléculas únicas, las cuales son capaces de reproducir el MFC. Aproximadamente, el 45%, 24% y 16% de las moléculas únicas derivaron

de las librerías químicas PubChem, Chem Space y Mcule-Ultimate, respectivamente. En el mismo sentido, la mayoría de estas moléculas conservan los descriptores farmacofóricos que conforman al submodelo farmacofórico D8, tal es el caso de los candidatos PubChem-122614371, PubChem-89540475, PubChem-114466037, PubChem-114466030 y PubChem-137443050, por mencionar algunos. Este submodelo cuenta con 6 centroides, 2 del tipo AH (4 y 5), 3 del tipo DH (13, 17 y 18) y 1 del tipo Hyd (21).

El centroide 5 (AH), así como el 3 (DH) se localizan entre los subsitios S1-S1', próximos a His163 y Glu166, los cuales, como ya se mencionó, participan activamente en la formación de puentes de hidrógeno con 3CLpro. El centroide 4 (AH) se localizó sobre S4, la cual a pesar de ser reportada como una cavidad estrecha poco profunda de carácter principalmente hidrofóbico, Keretsu, S. y colaboradores (2020) reportan que S4 puede participar en interacciones adicionales a través de la formación de puentes de hidrógeno [95]. El centroide 21 (Hyd), así como el 17 y 18 (DH), se encuentran próximos a S2 potenciando las interacciones con 3CLpro tal y como varios autores han detallado sobre esta cavidad [85,95].

Considerando la gran cantidad de candidatos, se procedió a establecer como método de selección un protocolo de acoplamiento molecular, con

el fin de proponer una serie de candidatos para la determinación futura de su actividad mediante ensayos biológicos.

Previo al empleo del protocolo de acoplamiento molecular y con el objetivo de determinar qué tan efectivo sería éste para predecir adecuados modos de unión, se realizó su validación a partir de la cual resultó un porcentaje de eficiencia del 8.9% (al menos 20 poses con un $\text{RMSD} \leq 3.5 \text{ \AA}$), indicando que a través de este modelo es posible reproducir la pose cristalográfica de aproximadamente 1 de cada 10 ligandos (ver Figura 19). Justificando lo anterior, Macip G. y colaboradores (2022) reportan que de 61 estudios en los cuales se considera un paso de acoplamiento molecular dentro de un protocolo de VS para predecir posibles inhibidores de 3CLpro en SARS-CoV-2, únicamente 18 autores (29.5%) realizaron una validación del protocolo de acoplamiento molecular (*re-docking*). Por otra parte, los mismos autores mencionan que a pesar de usar dos estructuras de 3CLpro de SARS-CoV-2 y cinco *softwares* de acoplamiento molecular, ellos demostraron que 1) la correlación entre los valores del co-logaritmo de la concentración inhibitoria media (pIC_{50}) y los *scores* de acoplamiento molecular no son buenos, y 2) ni el *score* de acoplamiento molecular o el ΔG son capaces de distinguir entre compuestos con y sin actividad inhibitoria experimental frente a 3CLpro [209].

En el mismo sentido, Llanos MA, y colaboradores (2021) analizaron 168 artículos en donde se emplea el acoplamiento molecular como técnica de

VS para predecir posibles inhibidores de 3CLpro en SARS-CoV-2, durante el 2020. En éste, determinaron dos puntos principales, el primero es que sólo en el 57.7% de los casos examinaron la eficacia del sistema 3CLpro antes de realizar el VS, y de ellos, sólo el 29.9% probó tanto la predicción de la pose y calculó el rendimiento del cribado virtual. El segundo es que en sólo el 4.4% de los artículos donde se calcula la eficacia del VS, los autores utilizan para dicho cálculo, un conjunto de prueba (moléculas con actividad, sin actividad, señuelo). Adicionalmente, sólo 2 artículos (1.2%) han mencionado resultados de pruebas biológicas, siendo que en uno de ellos los resultados experimentales no son revelados [210].

Todo lo anterior sugiere que el protocolo de acoplamiento molecular no ha logrado representar de forma efectiva el sistema conformado por la diada catalítica de 3CLpro de SARS-CoV-2 y sus respectivos ligandos cocrystalizados, razones por las que el uso del acoplamiento molecular como técnica única para la realización de VS para predecir posibles inhibidores de 3CLpro en SARS-CoV-2, no es viable. Es aquí donde se refuerza la idea que señala que el empleo de metodologías tales como el modelo farmacofórico y el MFC ganan importancia y tienen la posibilidad de arrojar potenciales candidatos basados en resultados experimentales.

A pesar de los resultados obtenidos para la validación del protocolo de acoplamiento molecular, teorizamos que aquellos ligandos con alta

frecuencia de poses cercanas al MFC podrían ser mejores candidatos para estudios posteriores.

El estudio de acoplamiento molecular determinó que sólo 12 de los 312 candidatos presentaron un $\text{RMSD} \leq 3.5 \text{ \AA}$ en al menos 20 poses, lo cual indica que esos 12 candidatos mantenían las características farmacofóricas derivadas del MFC con alta frecuencia (ver Figura 20).

Por ejemplo, el ligando PubChem-89540475, que cuenta con la mayor frecuencia presentada de poses con un $\text{RMSD} \leq 3.5 \text{ \AA}$ mantiene las características derivadas del MFC con alta frecuencia y, a su vez, cuenta con un score de 68.29, lo cual sugiere que tendrían una buena energía de unión con 3CLpro. A nivel molecular este candidato cuenta con 3 características importantes a destacar dentro de los subsitios (ver Figura 21). La primera de ellas es que la estructura de este candidato es capaz de abarcar la mayor parte de la cavidad. La segunda es que sobre S2 se encuentran dos componentes importantes, por un lado se localiza parte del macrociclo y por el otro, un grupo donador de puente de hidrógeno, componentes que potencian las interacciones con 3CLpro tal y como Keretsu, S. y colaboradores (2020), así como otros autores [85,95,206], han detallado en sus estudios [95]. La tercera es que en S4 se encuentra un grupo sulfonamida el cual presenta dos puntos a destacar. Por un lado, tiene la capacidad de actuar como donador y aceptor de puente de hidrógeno, y por el otro presenta un valor de la constante de lipofilia de

Hansch $p=-1.82$, lo que marca una hidrofilia importante [211]. En el contexto de la diada catalítica de 3CLpro un grupo de esta índole dentro de S4 concuerda con la naturaleza de dicho subsitio, el cual se ha reportado como una cavidad hidrofóbica flexible [206].

Los otros 11 candidatos catalogados como los mejores tras el empleo del MFC y del protocolo de acoplamiento molecular, también tienen el potencial de ser inhibidores de 3CLpro, ya que, es necesario remarcar, dichos candidatos mantienen características fisicoquímicas, estéricas y espaciales extraídas de los submodelos farmacofóricos derivados del MFC, el cual a su vez deriva de toda la información cristalográfica considerada.

Finalmente, con el fin de probar la actividad biológica de los 12 candidatos con potencial inhibitorio ante 3CLpro de SARS-CoV-2, es necesario la realización de ensayos biológicos. Lo anterior con miras a establecer dos puntos, 1) un potencial tratamiento que haga frente ante la actual situación mundial causada por la COVID-19 y, 2) que se permita el establecimiento de la funcionalidad, y posterior optimización, de la metodología innovadora del MFC. Adicionalmente, no hay que olvidar que, al tratarse de una metodología automatizada para este sistema, ésta es capaz de actualizarse con frecuencia, permitiendo recabar toda la información novedosa y procesarla en cortos lapsos de tiempo.

VIII. Conclusión

Con base en su estructura y en la investigación bibliográfica previa, se seleccionó a la diada catalítica de 3CLpro, es decir, al contexto de los residuos His41 y Cys145, como el sitio de interés para el presente proyecto. A partir de ésta diada catalítica y de toda la información cristalográfica existente se generó un conjunto de modelos farmacofóricos individuales (por ligando cocrystalizado con 3CLpro) el cual se empleó en la generación de un MFC.

Se seleccionaron, analizaron e identificaron potenciales candidatos inhibidores de 3CLpro derivados de un VS mediante el empleo de un MFC y un protocolo de acoplamiento molecular, sin embargo, se determinó que la segunda metodología no representa de forma adecuada al sitio activo de 3CLpro, ya que sólo se logró reproducir la pose cristalográfica en un 8.9% del total de los casos, corroborando lo reportado en relación con el uso de protocolos de acoplamiento molecular para este sistema. Finalmente, se determinó que MFC representa de forma adecuada a los inhibidores cocrystalizados de 3CLpro y logró capturar un 39.7% de las poses cristalizadas, destacando que este modelo tiene el potencial de ser optimizado.

IX. Perspectivas

Durante el análisis del MFC se visualizó una limitante, los centroides no parecían representar de la mejor forma al gran conjunto de descriptores farmacofóricos considerados para este estudio. Esto con base en que dichos centroides fueron calculados y generados considerando únicamente las coordenadas promedio por grupo de descriptores farmacofóricos, dejando de lado 1) la frecuencia de aparición de los descriptores de forma particular, por tipo de descriptor, y de forma general, así como 2) la dispersión que presentó cada descriptor dentro de cada grupo. De esta manera el MFC tiene el potencial de ser optimizado al considerar estos dos últimos puntos. En lo que respecta a la jerarquización y selección de potenciales candidatos con actividad inhibitoria contra la 3CLpro de SARS-CoV-2 se sugiere, en vez del empleo de protocolos de acoplamiento molecular, considerar parámetros tales como: 1) el RMSD tras optimización, 2) las reglas de Lipinski, 3) el score tras minimización, 4) la eficiencia de ligando y 5) la inspección visual.

X. Referencias

1. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *New England Journal of Medicine*. 2020 [cited 17 Jun 2021]. doi:10.1056/NEJMoa2001017
2. Zhou P, Yang X-L, Wang X-G, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. 2020;579: 270–273. doi:10.1038/s41586-020-2012-7
3. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. *New England Journal of Medicine*. 2020 [cited 17 Jun 2021]. doi:10.1056/NEJMoa2001316
4. Wu F, Zhao S, Yu B, Chen Y-M, Wang W, Song Z-G, et al. A new coronavirus associated with human respiratory disease in China. *Nature*. 2020;579: 265–269. doi:10.1038/s41586-020-2008-3
5. Díaz-Castrillón FJ, Toro-Montoya AI. SARS-CoV-2/COVID-19: el virus, la enfermedad y la pandemia. *Med Lab*. 2020;24: 183–205. doi:10.36384/01232576.268
6. WHO Coronavirus (COVID-19) Dashboard. [cited 17 Jun 2021]. Available: <https://covid19.who.int>
7. Wu Z, McGoogan JM. Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72 314 Cases From the Chinese Center for Disease Control and Prevention. *JAMA*. 2020;323: 1239. doi:10.1001/jama.2020.2648
8. Gorbalenya AE, Baker SC, Baric RS, de Groot RJ, Drosten C, Gulyaeva AA, et al. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nat Microbiol*. 2020;5: 536–544. doi:10.1038/s41564-020-0695-z
9. Chan JF-W, Kok K-H, Zhu Z, Chu H, To KK-W, Yuan S, et al. Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. *Emerging Microbes & Infections*. 2020;9: 221–236. doi:10.1080/22221751.2020.1719902
10. Ferreira JC, Rabeh WM. Biochemical and biophysical characterization of the main protease, 3-chymotrypsin-like protease (3CLpro) from the novel coronavirus SARS-CoV 2. *Sci Rep*. 2020;10: 22200. doi:10.1038/s41598-020-79357-0
11. Escudero X, Guarner J, Galindo-Fraga A, Escudero-Salamanca M, Alcocer-Gamba MA, Río CD, et al. La pandemia de Coronavirus SARS-CoV-2 (COVID-19): Situación actual e implicaciones para México. *Archivos de cardiología de México*. 2020;90: 7–14.

- doi:10.24875/acm.m20000064
12. Profile of a killer: the complex biology powering the coronavirus pandemic. [cited 17 Jun 2021]. Available: <https://www.nature.com/articles/d41586-020-01315-7>
 13. Cui J, Li F, Shi Z-L. Origin and evolution of pathogenic coronaviruses. *Nat Rev Microbiol.* 2019;17: 181–192. doi:10.1038/s41579-018-0118-9
 14. Pal M, Berhanu G, Desalegn C, Kandi V. Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2): An Update. *Cureus.* 2020;12. doi:10.7759/cureus.7423
 15. Fehr AR, Perlman S. Coronaviruses: An Overview of Their Replication and Pathogenesis. *Coronaviruses.* 2015;1282: 1–23. doi:10.1007/978-1-4939-2438-7_1
 16. Wu A, Peng Y, Huang B, Ding X, Wang X, Niu P, et al. Genome Composition and Divergence of the Novel Coronavirus (2019-nCoV) Originating in China. *Cell Host & Microbe.* 2020;27: 325–328. doi:10.1016/j.chom.2020.02.001
 17. Lu R, Zhao X, Li J, Niu P, Yang B, Wu H, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *The Lancet.* 2020;395: 565–574. doi:10.1016/S0140-6736(20)30251-8
 18. Chen Y, Liu Q, Guo D. Emerging coronaviruses: Genome structure, replication, and pathogenesis. *J Med Virol.* 2020;92: 418–423. doi:10.1002/jmv.25681
 19. Raj R. Analysis of non-structural proteins, NSPs of SARS-CoV-2 as targets for computational drug designing. *Biochemistry and Biophysics Reports.* 2021;25: 100847. doi:10.1016/j.bbrep.2020.100847
 20. Jin Z, Du X, Xu Y, Deng Y, Liu M, Zhao Y, et al. Structure of M pro from SARS-CoV-2 and discovery of its inhibitors. *Nature.* 2020;582: 289–293. doi:10.1038/s41586-020-2223-y
 21. Hosseini M, Chen W, Xiao D, Wang C. Computational molecular docking and virtual screening revealed promising SARS-CoV-2 drugs. *Precision Clinical Medicine.* 2021;4: 1–16. doi:10.1093/pcmmedi/pbab001
 22. Deng X, Baker SC. An “Old” protein with a new story: Coronavirus endoribonuclease is important for evading host antiviral defenses. *Virology.* 2018;517: 157–163. doi:10.1016/j.virol.2017.12.024
 23. Jauregui AR, Savalia D, Lowry VK, Farrell CM, Wathelet MG. Identification of Residues of SARS-CoV nsp1 That Differentially Affect Inhibition of Gene Expression and Antiviral Signaling. Li K, editor. *PLoS ONE.* 2013;8: e62416. doi:10.1371/journal.pone.0062416
 24. Snijder EJ, Bredenbeek PJ, Dobbe JC, Thiel V, Ziebuhr J, Poon LLM, et al. Unique and Conserved Features of Genome and Proteome of SARS-coronavirus, an Early Split-off From the Coronavirus Group 2

- Lineage. *Journal of Molecular Biology*. 2003;331: 991–1004. doi:10.1016/S0022-2836(03)00865-9
25. Graham RL, Sims AC, Baric RS, Denison MR. The NSP2 Proteins of Mouse Hepatitis Virus and Sars Coronavirus are Dispensable for Viral Replication. In: Perlman S, Holmes KV, editors. *The Nidoviruses*. Boston, MA: Springer US; 2006. pp. 67–72. doi:10.1007/978-0-387-33012-9_10
 26. Sakai Y, Kawachi K, Terada Y, Omori H, Matsuura Y, Kamitani W. Two-amino acids change in the nsp4 of SARS coronavirus abolishes viral replication. *Virology*. 2017;510: 165–174. doi:10.1016/j.virol.2017.07.019
 27. Anand K, Ziebuhr J, Wadhwani P, Mesters JR, Hilgenfeld R. Coronavirus Main Proteinase (3CLpro) Structure: Basis for Design of Anti-SARS Drugs. *Science*. 2003;300: 1763–1767. doi:10.1126/science.1085658
 28. Benvenuto D, Angeletti S, Giovanetti M, Bianchi M, Pascarella S, Cauda R, et al. Evolutionary analysis of SARS-CoV-2: how mutation of Non-Structural Protein 6 (NSP6) could affect viral autophagy. *Journal of Infection*. 2020;81: e24–e27. doi:10.1016/j.jinf.2020.03.058
 29. te Velthuis AJW, van den Worm SHE, Snijder EJ. The SARS-coronavirus nsp7+nsp8 complex is a unique multimeric RNA polymerase capable of both de novo initiation and primer extension. *Nucleic Acids Research*. 2012;40: 1737–1747. doi:10.1093/nar/gkr893
 30. Littler DR, Gully BS, Colson RN, Rossjohn J. Crystal Structure of the SARS-CoV-2 Non-structural Protein 9, Nsp9. *iScience*. 2020;23: 101258. doi:10.1016/j.isci.2020.101258
 31. Rosas-Lemus M, Minasov G, Shuvalova L, Inniss NL, Kiryukhina O, Wiersum G, et al. The crystal structure of nsp10-nsp16 heterodimer from SARS-CoV-2 in complex with S-adenosylmethionine. *bioRxiv*. 2020; 2020.04.17.047498. doi:10.1101/2020.04.17.047498
 32. Ma Y, Wu L, Shaw N, Gao Y, Wang J, Sun Y, et al. Structural basis and functional analysis of the SARS coronavirus nsp14–nsp10 complex. *Proceedings of the National Academy of Sciences of the United States of America*. 2015 [cited 28 Oct 2021]. Available: <https://doi.org/10.1073/pnas.1508686112>
 33. Chen P, Jiang M, Hu T, Liu Q, Chen XS, Guo D. Biochemical characterization of exoribonuclease encoded by SARS coronavirus. *J Biochem Mol Biol*. 2007;40: 649–655. doi:10.5483/bmbrep.2007.40.5.649
 34. Alanagreh L, Alzoughool F, Atoum M. The Human Coronavirus Disease COVID-19: Its Origin, Characteristics, and Insights into Potential Drugs and Its Mechanisms. *Pathogens*. 2020;9: 331. doi:10.3390/pathogens9050331

35. Ratia K, Pegan S, Takayama J, Sleeman K, Coughlin M, Baliji S, et al. A noncovalent class of papain-like protease/deubiquitinase inhibitors blocks SARS virus replication. *PNAS*. 2008;105: 16119–16124. doi:10.1073/pnas.0805240105
36. Verma D, Mitra D, Paul M, Chaudhary P, Kamboj A, Thatoi H, et al. Potential inhibitors of SARS-CoV-2 (COVID 19) proteases PLpro and Mpro/ 3CLpro: molecular docking and simulation studies of three pertinent medicinal plant natural components. *Current Research in Pharmacology and Drug Discovery*. 2021;2: 100038. doi:10.1016/j.crphar.2021.100038
37. Boechat JL, Chora I, Morais A, Delgado L. The immune response to SARS-CoV-2 and COVID-19 immunopathology – Current perspectives. *Pulmonology*. 2021;27: 423–437. doi:10.1016/j.pulmoe.2021.03.008
38. Mousavizadeh L, Ghasemi S. Genotype and phenotype of COVID-19: Their roles in pathogenesis. *Journal of Microbiology, Immunology and Infection*. 2021;54: 159–163. doi:10.1016/j.jmii.2020.03.022
39. Science brief: SARS-CoV-2 and surface (fomite) transmission for indoor community environments. [cited 12 Nov 2021]. Available: <https://stacks.cdc.gov/view/cdc/104762>
40. Sanche S, Lin YT, Xu C, Romero-Severson E, Hengartner N, Ke R. High Contagiousness and Rapid Spread of Severe Acute Respiratory Syndrome Coronavirus 2 - Volume 26, Number 7—July 2020 - *Emerging Infectious Diseases journal - CDC*. [cited 20 Jul 2021]. doi:10.3201/eid2607.200282
41. Harrison AG, Lin T, Wang P. Mechanisms of SARS-CoV-2 Transmission and Pathogenesis. *Trends Immunol*. 2020;41: 1100–1115. doi:10.1016/j.it.2020.10.004
42. Cevik M, Kuppalli K, Kindrachuk J, Peiris M. Virology, transmission, and pathogenesis of SARS-CoV-2. *BMJ*. 2020;371: m3862. doi:10.1136/bmj.m3862
43. Mehta P, McAuley DF, Brown M, Sanchez E, Tattersall RS, Manson JJ. COVID-19: consider cytokine storm syndromes and immunosuppression. *The Lancet*. 2020;395: 1033–1034. doi:10.1016/S0140-6736(20)30628-0
44. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet*. 2020;395: 497–506. doi:10.1016/S0140-6736(20)30183-5
45. Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *The Lancet*. 2020;395: 507–513. doi:10.1016/S0140-6736(20)30211-7
46. Yao XH, Li TY, He ZC, Ping YF, Liu HW, Yu SC, et al. [A pathological report of three COVID-19 cases by minimal invasive autopsies]. *Zhonghua Bing Li Xue Za Zhi*. 2020;49: 411–417.

- doi:10.3760/cma.j.cn112151-20200312-00193
47. Wang D, Hu B, Hu C, Zhu F, Liu X, Zhang J, et al. Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus–Infected Pneumonia in Wuhan, China. *JAMA*. 2020;323: 1061–1069. doi:10.1001/jama.2020.1585
 48. Guan W, Ni Z, Hu Y, Liang W, Ou C, He J, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. *New England Journal of Medicine*. 2020;382: 1708–1720. doi:10.1056/NEJMoa2002032
 49. Patel KP, Patel PA, Vunnam RR, Hewlett AT, Jain R, Jing R, et al. Gastrointestinal, hepatobiliary, and pancreatic manifestations of COVID-19. *J Clin Virol*. 2020;128: 104386. doi:10.1016/j.jcv.2020.104386
 50. Hua A, O’Gallagher K, Sado D, Byrne J. Life-threatening cardiac tamponade complicating myo-pericarditis in COVID-19. *Eur Heart J*. 2020; ehaa253. doi:10.1093/eurheartj/ehaa253
 51. Tan L, Wang Q, Zhang D, Ding J, Huang Q, Tang Y-Q, et al. Lymphopenia predicts disease severity of COVID-19: a descriptive and predictive study. *Signal Transduct Target Ther*. 2020;5: 33. doi:10.1038/s41392-020-0148-4
 52. Abou-Ismaïl MY, Diamond A, Kapoor S, Arafah Y, Nayak L. The hypercoagulable state in COVID-19: Incidence, pathophysiology, and management. *Thromb Res*. 2020;194: 101–115. doi:10.1016/j.thromres.2020.06.029
 53. Aleem A, Shah H. Gastrointestinal And Hepatic Manifestations Of Coronavirus (COVID-19). StatPearls. Treasure Island (FL): StatPearls Publishing; 2021. Available: <http://www.ncbi.nlm.nih.gov/books/NBK570562/>
 54. Gabarre P, Dumas G, Dupont T, Darmon M, Azoulay E, Zafrani L. Acute kidney injury in critically ill patients with COVID-19. *Intensive Care Med*. 2020; 1–10. doi:10.1007/s00134-020-06153-9
 55. Cascella M, Rajnik M, Aleem A, Dulebohn SC, Di Napoli R. Features, Evaluation, and Treatment of Coronavirus (COVID-19). StatPearls. Treasure Island (FL): StatPearls Publishing; 2021. Available: <http://www.ncbi.nlm.nih.gov/books/NBK554776/>
 56. Tay MZ, Poh CM, Rénia L, MacAry PA, Ng LFP. The trinity of COVID-19: immunity, inflammation and intervention. *Nat Rev Immunol*. 2020;20: 363–374. doi:10.1038/s41577-020-0311-8
 57. Wang J, Jiang M, Chen X, Montaner LJ. Cytokine storm and leukocyte changes in mild versus severe SARS-CoV-2 infection: Review of 3939 COVID-19 patients in China and emerging pathogenesis and therapy concepts. *J Leukoc Biol*. 2020; 10.1002/JLB.3COVR0520-272R. doi:10.1002/JLB.3COVR0520-272R
 58. Azkur AK, Akdis M, Azkur D, Sokolowska M, van de Veen W,

- Brüggen M, et al. Immune response to SARS-CoV-2 and mechanisms of immunopathological changes in COVID-19. *Allergy*. 2020; 10.1111/all.14364. doi:10.1111/all.14364
59. Conti P, Ronconi G, Caraffa A, Gallenga C, Ross R, Frydas I, et al. Induction of pro-inflammatory cytokines (IL-1 and IL-6) and lung inflammation by Coronavirus-19 (COVI-19 or SARS-CoV-2): anti-inflammatory strategies. *J Biol Regul Homeost Agents*. 2020;34: 327–331. doi:10.23812/CONTI-E
 60. COVID19-web.csv. [cited 3 Nov 2021]. Available: https://worldhealthorg-my.sharepoint.com/:x:/g/personal/karamg_who_int/EReItiDVLS5HuSNOVB21VRAByz3zAv-jxPoNS3cgtTywKA?e=5NWl6s
 61. Dyer O. Two Ebola treatments halve deaths in trial in DRC outbreak. *BMJ*. 2019; l5140. doi:10.1136/bmj.l5140
 62. Commissioner O of the. FDA Approves First Treatment for COVID-19. In: FDA [Internet]. FDA; 22 Oct 2020 [cited 16 Aug 2021]. Available: <https://www.fda.gov/news-events/press-announcements/fda-approves-first-treatment-covid-19>
 63. Research C for DE and. Know Your Treatment Options for COVID-19. FDA. 2021 [cited 16 Aug 2021]. Available: <https://www.fda.gov/consumers/consumer-updates/know-your-treatment-options-covid-19>
 64. Commissioner O of the. Coronavirus (COVID-19) Update: FDA Authorizes First Oral Antiviral for Treatment of COVID-19. In: FDA [Internet]. FDA; 22 Dec 2021 [cited 21 Jan 2022]. Available: <https://www.fda.gov/news-events/press-announcements/coronavirus-covid-19-update-fda-authorizes-first-oral-antiviral-treatment-covid-19>
 65. Commissioner O of the. Coronavirus (COVID-19) Update: FDA Authorizes Additional Oral Antiviral for Treatment of COVID-19 in Certain Adults. In: FDA [Internet]. FDA; 23 Dec 2021 [cited 21 Jan 2022]. Available: <https://www.fda.gov/news-events/press-announcements/coronavirus-covid-19-update-fda-authorizes-additional-oral-antiviral-treatment-covid-19-certain>
 66. Commissioner O of the. Emergency Use Authorization. FDA. 2021 [cited 9 Nov 2021]. Available: <https://www.fda.gov/emergency-preparedness-and-response/mcm-legal-regulatory-and-policy-framework/emergency-use-authorization>
 67. Commissioner O of the. Coronavirus (COVID-19) Update: December 23, 2021. In: FDA [Internet]. FDA; 23 Dec 2021 [cited 21 Jan 2022]. Available: <https://www.fda.gov/news-events/press-announcements/coronavirus-covid-19-update-december-23-2021>
 68. Monteil V, Kwon H, Prado P, Hagelkrüys A, Wimmer RA, Stahl M, et al. Inhibition of SARS-CoV-2 Infections in Engineered Human Tissues

- Using Clinical-Grade Soluble Human ACE2. *Cell*. 2020;181: 905-913.e7. doi:10.1016/j.cell.2020.04.004
69. Tian X, Li C, Huang A, Xia S, Lu S, Shi Z, et al. Potent binding of 2019 novel coronavirus spike protein by a SARS coronavirus-specific human monoclonal antibody. *Emerg Microbes Infect*. 2020;9: 382-385. doi:10.1080/22221751.2020.1729069
 70. Xia S, Liu M, Wang C, Xu W, Lan Q, Feng S, et al. Inhibition of SARS-CoV-2 (previously 2019-nCoV) infection by a highly potent pan-coronavirus fusion inhibitor targeting its spike protein that harbors a high capacity to mediate membrane fusion. *Cell Res*. 2020;30: 343-355. doi:10.1038/s41422-020-0305-x
 71. Wang M, Cao R, Zhang L, Yang X, Liu J, Xu M, et al. Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. *Cell Res*. 2020;30: 269-271. doi:10.1038/s41422-020-0282-0
 72. Markham A, Keam SJ. Danoprevir: First Global Approval. *Drugs*. 2018;78: 1271-1276. doi:10.1007/s40265-018-0960-0
 73. Santos JR, Curran A, Navarro-Mercade J, Ampuero MF, Pelaez P, Pérez-Alvarez N, et al. Simplification of Antiretroviral Treatment from Darunavir/Ritonavir Monotherapy to Darunavir/Cobicistat Monotherapy: Effectiveness and Safety in Routine Clinical Practice. *AIDS Research and Human Retroviruses*. 2019;35: 513-518. doi:10.1089/aid.2018.0178
 74. Lopinavir-Ritonavir Combination Induces Endoplasmic Reticulum Stress and Kills Urological Cancer Cells | *Anticancer Research*. [cited 16 Aug 2021]. Available: <https://ar.iijournals.org/content/39/11/5891.long>
 75. Uyeki TM. Oseltamivir Treatment of Influenza in Children. *Clin Infect Dis*. 2018;66: 1501-1503. doi:10.1093/cid/cix1150
 76. Blaising J, Polyak SJ, Pécheur E-I. Arbidol as a broad-spectrum antiviral: An update. *Antiviral Research*. 2014;107: 84-94. doi:10.1016/j.antiviral.2014.04.006
 77. Mody V, Ho J, Wills S, Mawri A, Lawson L, Ebert MCCJC, et al. Identification of 3-chymotrypsin like protease (3CLPro) inhibitors as potential anti-SARS-CoV-2 agents. *Commun Biol*. 2021;4: 1-10. doi:10.1038/s42003-020-01577-x
 78. Kifle Z. Current Therapeutic Drugs Treating COVID-19. 2020.
 79. Hu B, Guo H, Zhou P, Shi Z-L. Characteristics of SARS-CoV-2 and COVID-19. *Nat Rev Microbiol*. 2021;19: 141-154. doi:10.1038/s41579-020-00459-7
 80. Zhang L, Lin D, Kusov Y, Nian Y, Ma Q, Wang J, et al. α -Ketoamides as Broad-Spectrum Inhibitors of Coronavirus and Enterovirus Replication: Structure-Based Design, Synthesis, and Activity Assessment. *J Med Chem*. 2020;63: 4562-4578.

- doi:10.1021/acs.jmedchem.9b01828
81. Zhang L, Lin D, Sun X, Curth U, Drosten C, Sauerhering L, et al. Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α -ketoamide inhibitors. *Science*. 2020;368: 409–412. doi:10.1126/science.abb3405
 82. Zhong N, Zhang S, Zou P, Chen J, Kang X, Li Z, et al. Without Its N-Finger, the Main Protease of Severe Acute Respiratory Syndrome Coronavirus Can Form a Novel Dimer through Its C-Terminal Domain. *Journal of Virology*. 2008;82: 4227–4234. doi:10.1128/JVI.02612-07
 83. Gioia M, Ciaccio C, Calligari P, De Simone G, Sbardella D, Tundo G, et al. Role of proteolytic enzymes in the COVID-19 infection and promising therapeutic approaches. *Biochemical Pharmacology*. 2020;182: 114225. doi:10.1016/j.bcp.2020.114225
 84. Komatsu TS, Okimoto N, Koyama YM, Hirano Y, Morimoto G, Ohno Y, et al. Drug Binding Dynamics of the Dimeric SARS-CoV-2 Main Protease, Determined by Molecular Dynamics Simulation. 2020 [cited 19 Jul 2021]. doi:10.26434/chemrxiv.12332678.v1
 85. Rut W, Groborz K, Zhang L, Sun X, Zmudzinski M, Hilgenfeld R, et al. Substrate specificity profiling of SARS-CoV-2 Mpro protease provides basis for anti-COVID-19 drug design. 2020 Mar p. 2020.03.07.981928. doi:10.1101/2020.03.07.981928
 86. Chuck C-P, Chong L-T, Chen C, Chow H-F, Wan DC-C, Wong K-B. Profiling of Substrate Specificity of SARS-CoV 3CLpro. *PLOS ONE*. 2010;5: e13197. doi:10.1371/journal.pone.0013197
 87. Jin Z, Du X, Xu Y, Deng Y, Liu M, Zhao Y, et al. Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. *Nature*. 2020;582: 289–293. doi:10.1038/s41586-020-2223-y
 88. Wang H, He S, Deng W, Zhang Y, Li G, Sun J, et al. Comprehensive Insights into the Catalytic Mechanism of Middle East Respiratory Syndrome 3C-Like Protease and Severe Acute Respiratory Syndrome 3C-Like Protease. *ACS Catal*. 2020;10: 5871–5890. doi:10.1021/acscatal.0c00110
 89. Hilgenfeld R. From SARS to MERS: crystallographic studies on coronaviral proteases enable antiviral drug design. *The FEBS Journal*. 2014;281: 4085–4096. doi:10.1111/febs.12936
 90. Xue X, Yang H, Shen W, Zhao Q, Li J, Yang K, et al. Production of Authentic SARS-CoV Mpro with Enhanced Activity: Application as a Novel Tag-cleavage Endopeptidase for Protein Overproduction. *Journal of Molecular Biology*. 2007;366: 965–975. doi:10.1016/j.jmb.2006.11.073
 91. Ramos-Guzmán CA, Ruiz-Pernía JJ, Tuñón I. Unraveling the SARS-CoV-2 Main Protease Mechanism Using Multiscale Methods. *ACS Catal*. 2020;10: 12544–12554. doi:10.1021/acscatal.0c03420
 92. Pathak N, Chen Y-T, Hsu Y-C, Hsu N-Y, Kuo C-J, Tsai HP, et al.

- Uncovering Flexible Active Site Conformations of SARS-CoV-2 3CL Proteases through Protease Pharmacophore Clusters and COVID-19 Drug Repurposing. *ACS Nano*. 2021;15: 857–872. doi:10.1021/acsnano.0c07383
93. Ma C, Sacco MD, Hurst B, Townsend JA, Hu Y, Szeto T, et al. Boceprevir, GC-376, and calpain inhibitors II, XII inhibit SARS-CoV-2 viral replication by targeting the viral main protease. *bioRxiv*. 2020; 2020.04.20.051581. doi:10.1101/2020.04.20.051581
 94. Xia Z, Sacco M, Hu Y, Ma C, Meng X, Zhang F, et al. Rational Design of Hybrid SARS-CoV-2 Main Protease Inhibitors Guided by the Superimposed Cocrystal Structures with the Peptidomimetic Inhibitors GC-376, Telaprevir, and Boceprevir. *ACS Pharmacol Transl Sci*. 2021;4: 1408–1421. doi:10.1021/acspstsci.1c00099
 95. Keretsu S, Bhujbal SP, Cho SJ. Rational approach toward COVID-19 main protease inhibitors via molecular docking, molecular dynamics simulation and free energy calculation. *Sci Rep*. 2020;10: 17716. doi:10.1038/s41598-020-74468-0
 96. Vankadara S, Wong YX, Liu B, See YY, Tan LH, Tan QW, et al. A head-to-head comparison of the inhibitory activities of 15 peptidomimetic SARS-CoV-2 3CLpro inhibitors. *Bioorg Med Chem Lett*. 2021;48: 128263. doi:10.1016/j.bmcl.2021.128263
 97. Hufsky F, Lamkiewicz K, Almeida A, Aouacheria A, Arighi C, Bateman A, et al. Computational strategies to combat COVID-19: useful tools to accelerate SARS-CoV-2 and coronavirus research. *Briefings in Bioinformatics*. 2021;22: 642–663. doi:10.1093/bib/bbaa232
 98. Congreve M, Murray CW, Blundell TL. Keynote review: Structural biology and drug discovery. *Drug Discovery Today*. 2005;10: 895–907. doi:10.1016/S1359-6446(05)03484-7
 99. Arya R, Kumari S, Pandey B, Mistry H, Bihani SC, Das A, et al. Structural insights into SARS-CoV-2 proteins. *J Mol Biol*. 2021;433: 166725. doi:10.1016/j.jmb.2020.11.024
 100. Bank RPD. RCSB PDB: Homepage. [cited 15 Jul 2021]. Available: <https://www.rcsb.org/>
 101. Chen H, Zhou X, Gao Y, Chen H, Zhou J. Fragment-Based Drug Design: Strategic Advances and Lessons Learned. *Comprehensive Medicinal Chemistry III*. Elsevier; 2017. pp. 212–232. doi:10.1016/B978-0-12-409547-2.12319-4
 102. Hughes J, Rees S, Kalindjian S, Philpott K. Principles of early drug discovery: Principles of early drug discovery. *British Journal of Pharmacology*. 2011;162: 1239–1249. doi:10.1111/j.1476-5381.2010.01127.x
 103. Sinha S, Vohora D. *Drug Discovery and Development. Pharmaceutical Medicine and Translational Clinical Research*. Elsevier;

2018. pp. 19–32. doi:10.1016/B978-0-12-802103-3.00002-X
104. Saldívar-González F, Prieto-Martínez FD, Medina-Franco JL. Descubrimiento y desarrollo de fármacos: un enfoque computacional. *Educación Química*. 2017;28: 51–58. doi:10.1016/j.eq.2016.06.002
 105. Santos R, Ursu O, Gaulton A, Bento AP, Donadi RS, Bologa CG, et al. A comprehensive map of molecular drug targets. *Nat Rev Drug Discov*. 2017;16: 19–34. doi:10.1038/nrd.2016.230
 106. Hernández Cuellar CE, Castrillón-Martínez E, Gallego-Gómez JC. Descubrimiento de fármacos basado en imagenología de células vivas. *Rev Colomb Cienc Quím Farm*. 2020;49. doi:10.15446/rcciquifa.v49n1.87026
 107. Structure-Based Target Druggability Assessment | SpringerLink. [cited 19 Oct 2021]. Available: https://link.springer.com/protocol/10.1007%2F978-1-62703-311-4_10
 108. Muratov EN, Amaro R, Andrade CH, Brown N, Ekins S, Fourches D, et al. A critical overview of computational approaches employed for COVID-19 drug discovery. *Chem Soc Rev*. 2021;50: 9121–9151. doi:10.1039/D0CS01065K
 109. Huang R, Xu M, Zhu H, Chen CZ, Zhu W, Lee EM, et al. Biological activity-based modeling identifies antiviral leads against SARS-CoV-2. *Nat Biotechnol*. 2021;39: 747–753. doi:10.1038/s41587-021-00839-1
 110. Riva L, Yuan S, Yin X, Martin-Sancho L, Matsunaga N, Pache L, et al. Discovery of SARS-CoV-2 antiviral drugs through large-scale compound repurposing. *Nature*. 2020;586: 113–119. doi:10.1038/s41586-020-2577-1
 111. AI Therapeutics, Inc. A Phase II Randomized, Double-Blind, Placebo-Controlled Study of LAM-002A for the Prevention of Progression of COVID-19. *clinicaltrials.gov*; 2020 Nov. Report No.: NCT04446377. Available: <https://clinicaltrials.gov/ct2/show/NCT04446377>
 112. MD IFH. An Open-label Randomised Controlled Trial on Dual Therapy With Interferon Beta-1b and Clofazimine Combination, as Treatment for COVID-19 Infection. *clinicaltrials.gov*; 2020 Jul. Report No.: NCT04465695. Available: <https://clinicaltrials.gov/ct2/show/NCT04465695>
 113. Research C for DE and. Coronavirus Treatment Acceleration Program (CTAP). FDA. 2022 [cited 22 Jan 2022]. Available: <https://www.fda.gov/drugs/coronavirus-covid-19-drugs/coronavirus-treatment-acceleration-program-ctap>
 114. Blundell TL. Structure-based drug design. *Nature*. 1996;384: 23–26. doi:10.1038/384023a0
 115. Campbell SF. Science, art and drug discovery: a personal

- perspective. *Clin Sci (Lond)*. 2000;99: 255–260.
116. Owen DR, Allerton CMN, Anderson AS, Aschenbrenner L, Avery M, Berritt S, et al. An oral SARS-CoV-2 Mpro inhibitor clinical candidate for the treatment of COVID-19. *Science*. 0: eabl4784. doi:10.1126/science.abl4784
117. Pfizer. AN INTERVENTIONAL EFFICACY AND SAFETY, PHASE 2/3, DOUBLE-BLIND, 2-ARM STUDY TO INVESTIGATE ORALLY ADMINISTERED PF-07321332/RITONAVIR COMPARED WITH PLACEBO IN NONHOSPITALIZED SYMPTOMATIC ADULT PARTICIPANTS WITH COVID-19 WHO ARE AT INCREASED RISK OF PROGRESSING TO SEVERE ILLNESS. *clinicaltrials.gov*; 2021 Oct. Report No.: NCT04960202. Available: <https://clinicaltrials.gov/ct2/show/NCT04960202>
118. Medina-Franco J, López-Vallejo F, Castillo R. Diseño de fármacos asistido por computadora. *Educación Química*. 2006;17: 452–457. doi:10.22201/fq.18708404e.2006.4.66027
119. Commissioner O of the. Know Your Treatment Options for COVID-19. FDA. 2021 [cited 20 Oct 2021]. Available: <https://www.fda.gov/consumers/consumer-updates/know-your-treatment-options-covid-19>
120. Leelananda SP, Lindert S. Computational methods in drug discovery. *Beilstein J Org Chem*. 2016;12: 2694–2718. doi:10.3762/bjoc.12.267
121. Tanaji TT, Santosh AK, Alan CR. Successful Applications of Computer Aided Drug Discovery: Moving Drugs from Concept to the Clinic. *Current Topics in Medicinal Chemistry*. 2009;10: 127–141.
122. Kühlbrandt W. The Resolution Revolution. *Science*. 2014 [cited 1 Sep 2021]. Available: <https://www.science.org/doi/abs/10.1126/science.1251652>
123. Loew GH, Villar HO, Alkorta I. Strategies for Indirect Computer-Aided Drug Design. *Pharm Res*. 1993;10: 475–486. doi:10.1023/A:1018977414572
124. Mason J, Good A, Martin E. 3-D Pharmacophores in Drug Discovery. *CPD*. 2001;7: 567–597. doi:10.2174/1381612013397843
125. Acharya C, Coop A, E. Polli J, D. MacKerell A. Recent Advances in Ligand-Based Drug Design: Relevance and Utility of the Conformationally Sampled Pharmacophore Approach. *CAD*. 2011;7: 10–22. doi:10.2174/157340911793743547
126. Yang S-Y. Pharmacophore modeling and applications in drug discovery: challenges and recent advances. *Drug Discov Today*. 2010;15: 444–450. doi:10.1016/j.drudis.2010.03.013
127. Bajorath J, editor. *Chemoinformatics and Computational Chemical Biology*. Totowa, NJ: Humana Press; 2011. doi:10.1007/978-1-60761-839-3

128. Molecular fingerprints and similarity searching — Open Babel v2.3.1 documentation. [cited 20 Oct 2021]. Available: <https://openbabel.org/docs/dev/Fingerprints/intro.html>
129. Verma J, Khedkar V, Coutinho E. 3D-QSAR in Drug Design - A Review. *CTMC*. 2010;10: 95–115. doi:10.2174/156802610790232260
130. Selvaraj C, Dinesh DC, Panwar U, Abhirami R, Boura E, Singh SK. Structure-based virtual screening and molecular dynamics simulation of SARS-CoV-2 Guanine-N7 methyltransferase (nsp14) for identifying antiviral inhibitors against COVID-19. *Journal of Biomolecular Structure and Dynamics*. 2021;39: 4582–4593. doi:10.1080/07391102.2020.1778535
131. Beck BR, Shin B, Choi Y, Park S, Kang K. Predicting commercially available antiviral drugs that may act on the novel coronavirus (SARS-CoV-2) through a drug-target interaction deep learning model. *Computational and Structural Biotechnology Journal*. 2020;18: 784–790. doi:10.1016/j.csbj.2020.03.025
132. Liu Z, Chen H, Wold EA, Zhou J. Small-Molecule Inhibitors of Protein–Protein Interactions. *Comprehensive Medicinal Chemistry III*. Elsevier; 2017. pp. 329–353. doi:10.1016/B978-0-12-409547-2.12326-1
133. McInnes C. Virtual screening strategies in drug discovery. *Current Opinion in Chemical Biology*. 2007;11: 494–502. doi:10.1016/j.cbpa.2007.08.033
134. Zoete V, Grosdidier A, Michielin O. Docking, virtual high throughput screening and in silico fragment-based drug design. *J Cell Mol Med*. 2009;13: 238–248. doi:10.1111/j.1582-4934.2008.00665.x
135. Kitchen DB, Decornez H, Furr JR, Bajorath J. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat Rev Drug Discov*. 2004;3: 935–949. doi:10.1038/nrd1549
136. Rabal O, Urbano-Cuadrado M, Oyarzabal J. Computational medicinal chemistry in fragment-based drug discovery: what, how and when. *Future Medicinal Chemistry*. 2011;3: 95–134. doi:10.4155/fmc.10.277
137. Tao W, Mian-Bin W, Zheng-Jie C, Hua C, Jian-Ping L, Li-Rong Y. Fragment-Based Drug Discovery and Molecular Docking in Drug Design. *Current Pharmaceutical Biotechnology*. 2014;16: 11–25.
138. Kim S, Thiessen PA, Bolton EE, Chen J, Fu G, Gindulyte A, et al. PubChem Substance and Compound databases. *Nucleic Acids Res*. 2016;44: D1202–D1213. doi:10.1093/nar/gkv951
139. Irwin JJ, Sterling T, Mysinger MM, Bolstad ES, Coleman RG. ZINC: A Free Tool to Discover Chemistry for Biology. *J Chem Inf Model*. 2012;52: 1757–1768. doi:10.1021/ci3001277
140. Wishart DS, Knox C, Guo AC, Shrivastava S, Hassanali M, Stothard

- P, et al. DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Research*. 2006;34: D668–D672. doi:10.1093/nar/gkj067
141. Bernstein FC, Koetzle TF, Williams GJB, Meyer EF, Brice MD, Rodgers JR, et al. The Protein Data Bank. *European Journal of Biochemistry*. 1977;80: 319–324. doi:10.1111/j.1432-1033.1977.tb11885.x
 142. Berman HM, Kleywegt GJ, Nakamura H, Markley JL. The Protein Data Bank at 40: Reflecting on the Past to Prepare for the Future. *Structure*. 2012;20: 391–396. doi:10.1016/j.str.2012.01.010
 143. Díaz Caballero A, Martínez Serrano E, Vivas Reyes R, Puerta Llerena L, Méndez Cuadro D, Cabrales Salgado R, et al. Modelación por homología de la proteína LuxS de *Porphyromonas gingivalis* cepa W83. *Revista Clínica de Periodoncia, Implantología y Rehabilitación Oral*. 2012;5: 105–113. doi:10.1016/S0718-5391(12)70102-7
 144. Sheng C, Zhang W. Fragment Informatics and Computational Fragment-Based Drug Design: An Overview and Update. *Medicinal Research Reviews*. 2013;33: 554–598. doi:10.1002/med.21255
 145. Zheng X, Gan L, Wang E, Wang J. Pocket-Based Drug Design: Exploring Pocket Space. *AAPS J*. 2013;15: 228–241. doi:10.1208/s12248-012-9426-6
 146. Meng X-Y, Zhang H-X, Mezei M, Cui M. Molecular Docking: A powerful approach for structure-based drug discovery. *Curr Comput Aided Drug Des*. 2011;7: 146–157.
 147. Chen G, Seukep AJ, Guo M. Recent Advances in Molecular Docking for the Research and Discovery of Potential Marine Drugs. *Mar Drugs*. 2020;18: 545. doi:10.3390/md18110545
 148. Sousa SF, Fernandes PA, Ramos MJ. Protein-ligand docking: Current status and future challenges. *Proteins*. 2006;65: 15–26. doi:10.1002/prot.21082
 149. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, et al. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem*. 2009;30: 2785–2791. doi:10.1002/jcc.21256
 150. Trott O, Olson AJ. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem*. 2009; NA-NA. doi:10.1002/jcc.21334
 151. Zhao H, Caflisch A. Discovery of ZAP70 inhibitors by high-throughput docking into a conformation of its kinase domain generated by molecular dynamics. *Bioorganic & Medicinal Chemistry Letters*. 2013;23: 5721–5726. doi:10.1016/j.bmcl.2013.08.009
 152. Allen WJ, Balius TE, Mukherjee S, Brozell SR, Moustakas DT, Lang PT, et al. DOCK 6: Impact of new features and current docking

- performance. *J Comput Chem.* 2015;36: 1132–1156. doi:10.1002/jcc.23905
153. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, et al. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J Med Chem.* 2004;47: 1739–1749. doi:10.1021/jm0306430
 154. Jones G, Willett P, Glen RC, Leach AR, Taylor R. Development and validation of a genetic algorithm for flexible docking 1 Edited by F. E. Cohen. *Journal of Molecular Biology.* 1997;267: 727–748. doi:10.1006/jmbi.1996.0897
 155. Corbeil CR, Williams CI, Labute P. Variability in docking success rates due to dataset preparation. *J Comput Aided Mol Des.* 2012;26: 775–786. doi:10.1007/s10822-012-9570-1
 156. Spitzer R, Jain AN. Surflex-Dock: Docking benchmarks and real-world application. *J Comput Aided Mol Des.* 2012;26: 687–699. doi:10.1007/s10822-011-9533-y
 157. Chen Y-C. Beware of docking! *Trends in Pharmacological Sciences.* 2015;36: 78–95. doi:10.1016/j.tips.2014.12.001
 158. Bello M, Martínez-Archundia M, Correa-Basurto J. Automated docking for novel drug discovery. *Expert Opinion on Drug Discovery.* 2013;8: 821–834. doi:10.1517/17460441.2013.794780
 159. S. F. S, A. J.m. R, J. T.s. C, R. P.p. N, S. A. M, N. S.h.n. M, et al. Protein-Ligand Docking in the New Millennium – A Retrospective of 10 Years in the Field. *Current Medicinal Chemistry.* 2013;20: 2296–2314.
 160. Gurung AB, Ali MA, Lee J, Farah MA, Al-Anazi KM. An Updated Review of Computer-Aided Drug Design and Its Application to COVID-19. *BioMed Research International.* 2021;2021: e8853056. doi:10.1155/2021/8853056
 161. Das S, Sarmah S, Lyndem S, Singha Roy A. An investigation into the identification of potential inhibitors of SARS-CoV-2 main protease using molecular docking study. *Journal of Biomolecular Structure and Dynamics.* 2021;39: 3347–3357. doi:10.1080/07391102.2020.1763201
 162. Khan SA, Zia K, Ashraf S, Uddin R, Ul-Haq Z. Identification of chymotrypsin-like protease inhibitors of SARS-CoV-2 via integrated computational approach. *Journal of Biomolecular Structure and Dynamics.* 2021;39: 2607–2616. doi:10.1080/07391102.2020.1751298
 163. Elmezayen AD, Yelekçi K. Homology modeling and in silico design of novel and potential dual-acting inhibitors of human histone deacetylases HDAC5 and HDAC9 isozymes. *Journal of Biomolecular Structure and Dynamics.* 2021;39: 6396–6414. doi:10.1080/07391102.2020.1798812
 164. Joshi RS, Jagdale SS, Bansode SB, Shankar SS, Tellis MB, Pandya

- VK, et al. Discovery of potential multi-target-directed ligands by targeting host-specific SARS-CoV-2 structurally conserved main protease. *Journal of Biomolecular Structure and Dynamics*. 2021;39: 3099–3114. doi:10.1080/07391102.2020.1760137
165. Gurung AB, Ali MA, Lee J, Farah MA, Al-Anazi KM. Unravelling lead antiviral phytochemicals for the inhibition of SARS-CoV-2 Mpro enzyme through in silico approach. *Life Sciences*. 2020;255: 117831. doi:10.1016/j.lfs.2020.117831
166. Wahedi HM, Ahmad S, Abbasi SW. Stilbene-based natural compounds as promising drug candidates against COVID-19. *Journal of Biomolecular Structure and Dynamics*. 2021;39: 3225–3234. doi:10.1080/07391102.2020.1762743
167. Ehrlich P. Über den jetzigen Stand der Chemotherapie. *Berichte der deutschen chemischen Gesellschaft*. 1909;42: 17–47. doi:10.1002/cber.19090420105
168. Yang S-Y. Pharmacophore modeling and applications in drug discovery: challenges and recent advances. *Drug Discovery Today*. 2010;15: 444–450. doi:10.1016/j.drudis.2010.03.013
169. Barnum D, Greene J, Smellie A, Sprague P. Identification of common functional configurations among molecules. *J Chem Inf Comput Sci*. 1996;36: 563–571. doi:10.1021/ci950273r
170. Güner OF. Pharmacophore Perception, Development, and Use in Drug Design. *Internat'l University Line*; 2000.
171. Kurogi Y, Guner OF. Pharmacophore Modeling and Three-dimensional Database Searching for Drug Design Using Catalyst. *Current Medicinal Chemistry*. 8: 1035–1055.
172. Jones: GASP: genetic algorithm superimposition program - Google Académico. [cited 31 Jan 2022]. Available: https://scholar.google.com/scholar_lookup?title=Chapter+GASP:Genetic+Algorithm+Superposition+Program&author=G+Jones&author=P+Willett&author=R+Glen&publication_year=2000&
173. Richmond NJ, Abrams CA, Wolohan PRN, Abrahamian E, Willett P, Clark RD. GALAHAD: 1. Pharmacophore identification by hypermolecular alignment of ligands in 3D. *J Comput Aided Mol Des*. 2006;20: 567–587. doi:10.1007/s10822-006-9082-y
174. Shepphird JK, Clark RD. A marriage made in torsional space: using GALAHAD models to drive pharmacophore multiplet searches. *J Comput Aided Mol Des*. 2006;20: 763–771. doi:10.1007/s10822-006-9070-2
175. Dixon SL, Smondryev AM, Knoll EH, Rao SN, Shaw DE, Friesner RA. PHASE: a new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1. Methodology and preliminary results. *J Comput Aided Mol Des*. 2006;20: 647–671. doi:10.1007/s10822-006-9087-6

176. Vilar S, Cozza G, Moro S. Medicinal Chemistry and the Molecular Operating Environment (MOE): Application of QSAR and Molecular Docking to Drug Discovery. *Current Topics in Medicinal Chemistry*. 8: 1555–1572.
177. Poling: Promoting conformational variation - Smellie - 1995 - *Journal of Computational Chemistry* - Wiley Online Library. [cited 31 Jan 2022]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.540160205>
178. Gippert GP, Wright PE, Case DA. Distributed torsion angle grid search in high dimensions: a systematic approach to NMR structure determination. *J Biomol NMR*. 1998;11: 241–263. doi:10.1023/a:1008209806860
179. Hurst T. Flexible 3D searching: The directed tweak technique. *J Chem Inf Comput Sci*. 1994. doi:10.1021/ci00017a025
180. Jones G, Willett P, Glen RC. A genetic algorithm for flexible molecular overlay and pharmacophore elucidation. *J Comput Aided Mol Des*. 1995;9: 532–549. doi:10.1007/BF00124324
181. Li Z, Scheraga HA. Monte Carlo-minimization approach to the multiple-minima problem in protein folding. *Proc Natl Acad Sci U S A*. 1987;84: 6611–6615.
182. LigandScout: 3-D Pharmacophores Derived from Protein-Bound Ligands and Their Use as Virtual Screening Filters | *Journal of Chemical Information and Modeling*. [cited 2 Feb 2022]. Available: <https://pubs.acs.org/doi/10.1021/ci049885e>
183. Chen J, Lai L. Pocket v.2: further developments on receptor-based pharmacophore modeling. *J Chem Inf Model*. 2006;46: 2684–2691. doi:10.1021/ci600246s
184. Ortuso F, Langer T, Alcaro S. GBPM: GRID-based pharmacophore model: concept and application studies to protein–protein recognition. *Bioinformatics*. 2006;22: 1449–1455. doi:10.1093/bioinformatics/btl115
185. Tschinke V, Cohen NC. The NEWLEAD program: a new method for the design of candidate structures from pharmacophoric hypotheses. *J Med Chem*. 1993;36: 3863–3870. doi:10.1021/jm00076a016
186. Böhm HJ. The computer program LUDI: a new method for the de novo design of enzyme inhibitors. *J Comput Aided Mol Des*. 1992;6: 61–78. doi:10.1007/BF00124387
187. Roe DC, Kuntz ID. BUILDER v.2: improving the chemistry of a de novo design strategy. *J Comput Aided Mol Des*. 1995;9: 269–282. doi:10.1007/BF00124457
188. Huang Q, Li L-L, Yang S-Y. PhDD: A new pharmacophore-based de novo design method of drug-like molecules combined with assessment of synthetic accessibility. *Journal of Molecular Graphics and Modelling*. 2010;28: 775–787. doi:10.1016/j.jmgm.2010.02.002

189. Daoud S, Alabed SJ, Dahabiyeh LA. Identification of potential COVID-19 main protease inhibitors using structure-based pharmacophore approach, molecular docking and repurposing studies. *Acta Pharmaceutica*. 2021;71: 163–174. doi:10.2478/acph-2021-0016
190. Beura S, Chetti P. In-silico strategies for probing chloroquine based inhibitors against SARS-CoV-2. *Journal of Biomolecular Structure and Dynamics*. 2021;39: 3747–3759. doi:10.1080/07391102.2020.1772111
191. Arun KG, Sharanya CS, Abhithaj J, Francis D, Sadasivan C. Drug repurposing against SARS-CoV-2 using E-pharmacophore based virtual screening, molecular docking and molecular dynamics with main protease as the target. *J Biomol Struct Dyn*. 2020; 1–12. doi:10.1080/07391102.2020.1779819
192. Gentile D, Patamia V, Scala A, Sciortino MT, Piperno A, Rescifina A. Putative Inhibitors of SARS-CoV-2 Main Protease from A Library of Marine Natural Products: A Virtual Screening and Molecular Modeling Study. *Marine Drugs*. 2020;18: 225. doi:10.3390/md18040225
193. Arun KG, Sharanya CS, Abhithaj J, Francis D, Sadasivan C. Drug repurposing against SARS-CoV-2 using E-pharmacophore based virtual screening, molecular docking and molecular dynamics with main protease as the target. *Journal of Biomolecular Structure and Dynamics*. 2021;39: 4647–4658. doi:10.1080/07391102.2020.1779819
194. Alzain AA, Elbadwi FA. Identification of novel TMPRSS2 inhibitors for COVID-19 using e-pharmacophore modelling, molecular docking, molecular dynamics and quantum mechanics studies. *Informatics in Medicine Unlocked*. 2021;26: 100758. doi:10.1016/j.imu.2021.100758
195. Pundir H, Joshi T, Pant M, Bhat S, Pandey J, Chandra S, et al. Identification of SARS-CoV-2 RNA dependent RNA polymerase inhibitors using pharmacophore modelling, molecular docking and molecular dynamics simulation approaches. *Journal of Biomolecular Structure and Dynamics*. 2021;0: 1–12. doi:10.1080/07391102.2021.1987329
196. The UniProt Consortium. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Research*. 2021;49: D480–D489. doi:10.1093/nar/gkaa1100
197. Bank RPD. RCSB PDB - 6M2Q: SARS-CoV-2 3CL protease (3CL pro) apo structure (space group C21). [cited 2 Feb 2022]. Available: <https://www.rcsb.org/structure/6m2q>
198. Koes DR, Camacho CJ. Pharmer: Efficient and Exact Pharmacophore Search. *J Chem Inf Model*. 2011;51: 1307–1314. doi:10.1021/ci200097m

199. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods*. 2020;17: 261–272. doi:10.1038/s41592-019-0686-2
200. RDKit. [cited 30 Jun 2022]. Available: <http://www.rdkit.org/>
201. Sunseri J, Koes DR. Pharmit: interactive exploration of chemical space. *Nucleic Acids Research*. 2016;44: W442–W448. doi:10.1093/nar/gkw287
202. smina download | SourceForge.net. [cited 2 Mar 2022]. Available: <https://sourceforge.net/projects/smina/>
203. Chemaxon - Software Solutions and Services for Chemistry & Biology. [cited 30 Jun 2022]. Available: <https://chemaxon.com/>
204. Using redundant internal coordinates to optimize equilibrium geometries and transition states - Peng - 1996 - *Journal of Computational Chemistry* - Wiley Online Library. [cited 25 Mar 2022]. Available: <https://onlinelibrary.wiley.com/doi/10.1002/%28SICI%291096-987X%2819960115%2917%3A1%3C49%3A%3AAID-JCC5%3E3.0.CO%3B2-0>
205. Tahir ul Qamar M, Alqahtani SM, Alamri MA, Chen L-L. Structural basis of SARS-CoV-2 3CLpro and anti-COVID-19 drug discovery from medicinal plants. *Journal of Pharmaceutical Analysis*. 2020;10: 313–319. doi:10.1016/j.jpha.2020.03.009
206. Zhu J, Zhang H, Lin Q, Lyu J, Lu L, Chen H, et al. Progress on SARS-CoV-2 3CLpro Inhibitors: Inspiration from SARS-CoV 3CLpro Peptidomimetics and Small-Molecule Anti-Inflammatory Compounds. *DDDT*. 2022;16: 1067–1082. doi:10.2147/DDDT.S359009
207. Schaller D, Šribar D, Noonan T, Deng L, Nguyen TN, Pach S, et al. Next generation 3D pharmacophore modeling. *WIREs Computational Molecular Science*. 2020;10: e1468. doi:10.1002/wcms.1468
208. PubChem Statistics. [cited 27 May 2022]. Available: <https://pubchemdocs.ncbi.nlm.nih.gov/statistics>
209. Macip G, Garcia-Segura P, Mestres-Truyol J, Saldivar-Espinoza B, Ojeda-Montes MJ, Gimeno A, et al. Haste makes waste: A critical review of docking-based virtual screening in drug repurposing for SARS-CoV-2 main protease (M-pro) inhibition. *Medicinal Research Reviews*. 2022;42: 744–769. doi:10.1002/med.21862
210. Llanos MA, Gantner ME, Rodriguez S, Alberca LN, Bellera CL, Talevi A, et al. Strengths and Weaknesses of Docking Simulations in the SARS-CoV-2 Era: the Main Protease (Mpro) Case Study. *J Chem Inf Model*. 2021;61: 3758–3770. doi:10.1021/acs.jcim.1c00404
211. Braña M, Cacho M, Guisado C. Sulfonamida: el grupo mágico. *Anales de la Real Academia Nacional de Farmacia*, ISSN 0034-0618, Nº 4 2, 2006, pags 317-341. 2009.