



UNIVERSIDAD TÉCNICA DE AMBATO
FACULTAD DE CIENCIA E INGENIERÍA EN ALIMENTOS
Y BIOTECNOLOGÍA



CARRERA DE INGENIERÍA BIOQUÍMICA

Cribado virtual de inhibidores de la enzima M^{pro} del virus SARS-CoV-2 empleando librerías de compuestos y homología estructural con proteasas de virus patógenos humanos.

Trabajo de Titulación, Modalidad Proyecto de Investigación, previo a la obtención de título de Ingeniero Bioquímico, otorgado por la Universidad Técnica de Ambato, a través de la Facultad de Ciencia e Ingeniería en Alimentos y Biotecnología.

Este Trabajo de Titulación forma parte del Proyecto “**Diseño racional de inhibidores con potencial actividad antiviral contra el virus SARS-CoV-2 (COVID-19)**”, auspiciado por la UODIDE mediante la Convocatoria 2020 y coordinado por Mario Daniel García Solís, PhD.

Autor: Andrés Salomón Sánchez Vaca

Tutor: PhD. Mario Daniel García Solís

Ambato - Ecuador

Enero 2021

APROBACIÓN DEL TUTOR

PhD. Mario Daniel García Solís

CERTIFICA

Que el presente trabajo de titulación ha sido prolijamente revisado. Por lo tanto, autorizo la presentación de este Trabajo de Titulación bajo la modalidad de Proyecto de Investigación, el mismo que responde a las normas establecidas en el Reglamento de Títulos y Grados de la Facultad de Ciencia e Ingeniería en Alimentos y Biotecnología de la Universidad Técnica de Ambato.

Ambato, 18 de Diciembre de 2020

PhD. Mario Daniel García Solís

C.I. 110360547-1

TUTOR

DECLARACIÓN DE AUTENTICIDAD

Yo, Andrés Salomón Sánchez Vaca, manifiesto que los resultados obtenidos en el presente Trabajo de Titulación, modalidad Proyecto de Investigación, previo a la obtención del título de Ingeniero Bioquímico son absolutamente originales, auténticos y personales; a excepción de las citas bibliográficas.



Andrés Salomón Sánchez Vaca

C.I. 180437731-3

AUTOR

APROBACIÓN DE LOS MIEMBROS DE TRIBUNAL DE GRADO

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

Para constancia firman:

Presidente del Tribunal
Dr. Orestes Darío López Hernández
CI. 1754784864

Dr. William Ricardo Calero Cáceres
CI. 1714348859

Dra. Liliana Alexandra Cerda Mejía
CI. 1804148086

Ambato, 08 de Enero de 2021

DERECHOS DE AUTOR

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

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



Andrés Salomón Sánchez Vaca

C.I. 180437731-3

AUTOR

DEDICATORIA

DEDICADO A DIOS YA QUE TODO HONOR Y GLORIA SEA A ÉL POR LOS SIGLOS DE LOS SIGLOS. AMÉN

DEDICADO A MIS PADRES WALTER SÁNCHEZ Y GLORIA VACA POR TODO EL ESFUERZO QUE HICIERON A LO LARGO DE MI VIDA PARA DARMÉ TODAS LAS POSIBILIDADES PARA ESTUDIAR Y LLEGAR A ESTE PROYECTO DE TITULACIÓN, ADEMÁS DE SU APOYO Y CARIÑO A LO LARGO DEL PROCESO.

DEDICADO A MI HERMANA IBETH SÁNCHEZ POR SU AYUDA Y EXPERIENCIA QUE CONTRIBUYÓ A MI CARRERA UNIVERSITARIA DEDICADO A MI ABUELITA ELSA ARBOLEDA, POR BRINDARME SU AMOR INCONDICIONAL Y SUS CONSEJOS QUE ME PERMITIERON SALIR ADELANTE Y LLEGAR HASTA ESTE PUNTO.

DEDICADO A MIS AMIGOS ELIZABETH FALCÓN, MARITZA SANDOVAL, RAQUEL JURADO, LEONARDO ALDÁS, CHRISTIAN SALINAS CON QUIENES FORMAMOS UN VERDADERO EQUIPO, DEBIÉNDOLES PARTE DE ESTE ÉXITO ACADÉMICO.

DEDICADO A MIS AMIGOS DANIELA ACOSTA, ANDRÉS GUTIÉRREZ, PABLO ILVIS, MATEO AGUAGUIÑA, A QUIENES CONOCÍ A LO LARGO DE MI CARRERA UNIVERSITARIA, VIVIENDO EXPERIENCIAS INOLVIDABLES.

DEDICADO A MARÍA JOSÉ HERNÁNDEZ, QUIÉN DURANTE TODA LA VIDA UNIVERSITARIA ME APOYÓ, ME AYUDÓ A SER MEJOR PERSONA Y A TENER LA FUERZA PARA SEGUIR ADELANTE.

DEDICADO A SARA NARANJO, QUIEN HIZO DE MI VIDA UNIVERSITARIA UNA VERDADERA AVENTURA, SIGUIENDO JUNTOS COMO UNO SOLO Y DÁNDOME AQUELLA MOTIVACIÓN QUE NECESITABA PARA ESFORZARME CADA DÍA MÁS HASTA ALCANZAR ESTA META.

AGRADECIMIENTOS

AGRADEZCO A LA UNIVERSIDAD TÉCNICA DE AMBATO Y A LA FACULTAD DE CIENCIA E INGENIERÍA EN ALIMENTOS POR ABRIRME SUS PUERTAS Y DARME LA OPORTUNIDAD DE ESTUDIAR Y FORMARME EN TAL PRESTIGIOSA INSTITUCIÓN.

AGRADEZCO A MI TUTOR PHD. MARIO DANIEL GARCÍA SOLÍS, A QUIÉN MÁS QUE MI PROFESOR LO CONSIDERO COMO MI AMIGO, POR TODO SU APOYO Y DEDICACIÓN EN EL PROCESO DE ELABORACIÓN DE ESTE PROYECTO, ASÍ COMO EN EL AULA DE CLASE.

AGRADEZCO A LISETTE RAMOS, DAYSI ESPÍN Y CAROLINA VIZUETA, QUIENES CONFORMAMOS EL EQUIPO DE INVESTIGACIÓN PARA LA REALIZACIÓN DE ESTE PROYECTO, POR SU APOYO PARA QUE TODOS LOGREMOS ESTA META.

ÍNDICE DE CONTENIDOS

Aprobación del Tutor	ii
Declaración de Autenticidad	iii
Aprobación de los miembros del Tribunal de Grado	iv
Derechos de Autor	v
Dedicatoria.....	vi
Agradecimientos	vii
Índice de figuras.....	x
Índice de Tablas.....	xii
Resumen	xiii
Abstract	xiv
CAPÍTULO I.....	1
1 Marco Teórico	1
1.1 Antecedentes Investigativos	1
1.2 1.2 El Virus Sars Cov-2.....	2
1.2.1 Origen.....	2
1.2.2 Características Y Estructura Del Virus SARS-Cov-2.....	4
1.3 Patogenicidad Del Virus Sars-Cov-2	6
1.3.1 Proteasa Principal Del Virus SARS-Cov-2 (<i>M Pro</i>)	6
1.3.2 Mecanismo De Infección Del Virus SARS-Cov-2	7
1.4 Diseño De Inhibidores Como Tratamiento A La Infección Por Sars-Cov-2	8
1.4.1 Inhibidores.....	8
1.5 Objetivos	10
1.5.1 Objetivo General.....	10
1.5.2 Objetivos Específicos.....	10
CAPÍTULO II.....	11
2 Metodología.....	11
2.1 Materiales	11
2.2 Procedimientos Cualitativos	11
2.3 Métodos	13
2.3.1 Análisis De Homología De Proteínas	13
2.3.2 Determinación De Los Sitios Drogables De La Proteína.....	21
2.3.3 Mapeo, Determinación Y Caracterización De La Estructura De Moléculas Con Actividad Inhibitoria	23

2.3.4	Docking O Acoplamiento De La Molécula Inhibidora A La Proteína	27
2.3.5	Diseño Racional De Inhibidores	32
CAPÍTULO III	40
Resultados Y Discusión	40
3.1	Homología De La Proteasa Mpro De Sars Cov-2 Respecto A Virus Patógenos Humanos	40
3.2	Sitios Drogables De La Proteasa Mpro Del Virus Sars Cov-2 Y Coronavirus Humanos Patógenos Homólogos.....	43
3.2.1	Alineamiento Estructural De Los Coronavirus Humanos Y Sus Aminoácidos Conservados.....	46
3.2.2	Interacción De Los Inhibidores Con La Proteasa Mpro De Coronavirus Patógenos Humanos.....	47
3.3	Resíduos Flexibles De La Proteasa Mpro	59
3.4	Acoplamiento Molecular O Docking De Inhibidores En Mpro De Sars Cov-2.....	60
3.5	Diseño Racional De Inhibidores De La Enzima Mpro	63
3.5.1	Superposición De Ligandos Diseñados.....	64
3.6	Comparación De Ligandos Identificados Y Diseñados Con Ligando De Referencia N3	68
3.7	Interacciones Intermoleculares Identificadas En Los Mejores Ligandos	70
3.7.1	Ligandos Obtenidos Del Cribado Virtual.....	71
3.7.2	Ligandos Obtenidos Del Diseño Racional	73
CAPÍTULO IV	75
Conclusiones Y Recomendaciones	75
4.1	Conclusiones.....	75
4.2	Recomendaciones	76
6	Bibliografía.....	77
ANEXOS	82
Anexo 1:	Secuencias De Virus Homólogos Humanos Para Alineamiento	82
Anexo 2:	Librería De Moléculas Con Actividad Inhibitoria De Mpro Obtenidas De Protein Data Bank.....	84
Anexo 3:	Interacción De Inhibidores Del Virus SARS Cov-2 Y Homólogos Humanos Con Los Aminácidos Del Sitio Activo De La Proteasa Mpro	95
Anexo 4:	Moléculas Con Actividad Inhibitoria Diseñadas	102

ÍNDICE DE FIGURAS

Figura 1. Transmisión zoonótica de coronavirus.....	3
Figura 2. Esquema de Schechter y Berger	6
Figura 3. Proceso de infección del virus SARS-CoV-2 a la célula	8
Figura 4. Principales inhibidores de la proteasa Mpro.....	9
Figura 5. Entorno de Uniprot.....	14
Figura 6. Menú desplegable para la descarga de secuencias en formato FASTA	14
Figura 7. Herramienta Blast de Uniprot	15
Figura 8. Resultados de homología arrojados por Uniprot	15
Figura 9. Obtención de información de alineamientos individuales.....	16
Figura 10. Entorno de Protein Data Bank	17
Figura 11. Lista de resultados arrojados por Protein Data Bank	17
Figura 12. Entorno de RCSB PDB	18
Figura 13. Identificación del posible inhibidor de interés.....	18
Figura 14. Información desplegada correspondiente al ligando seleccionado	19
Figura 15. Entorno de la herramienta TCOFFEE para alineamientos de secuencias	19
Figura 16. Formatos de descarga del alineamiento realizado.....	20
Figura 17. Entorno de la herramienta destinada a dar formato al alineamiento	20
Figura 18. Botón para la descarga del alineamiento.	21
Figura 19. Menú de descarga de estructuras en formato pdb	21
Figura 20. Entorno de software Pymol versión 0.8.7.....	21
Figura 21. Visualización de la estructura del ligando	22
Figura 22. Funciones del programa Wincoot usadas en la investigación	22
Figura 23. Visualización de los aminoácidos de la proteína y su interacción con el ligando .	23
Figura 24. Entorno del software de visualización molecular Pymol	24
Figura 25. Menú desplegado del software Pymol.....	24
Figura 26. Visualización de la estructura del ligando y la proteína en 3D	25
Figura 27. Codificación correspondiente al ligando, dentro de la estructura total	25
Figura 28. Función “ray” de Pymol para obtener una imagen de la estructura.....	26
Figura 29. Entorno del programa Mercury.....	26
Figura 30. Función usada en la optimización de ligandos	27

Figura 31. Entorno de GOLD Protein Docking Ligand	27
Figura 32. Configuración de la proteína objetivo	28
Figura 33. Configuración del sitio de acoplamiento para ligandos.	28
Figura 34. Definición de programa para acoplamiento molecular.	29
Figura 35. Adición de los ligandos optimizados para el acoplamiento molecular	29
Figura 36. Elección de programa para el estudio.....	30
Figura 37. Definición de parámetros para acoplamiento molecular	30
Figura 38. Modificación de residuos flexibles de la proteína.....	31
Figura 39. Proceso de análisis y prueba de los ligandos por parte del software.....	31
Figura 40. Obtención de resultados del acoplamiento molecular	32
Figura 41. Estructura 2D de los ligandos con mayor actividad inhibitoria de Mpro.....	32
Figura 42. Posibles sitios para la incorporación de grupos funcionales al ligando MMV1580167	33
Figura 43. Transición de la molécula MMV1580167 al ligando AS1.....	33
Figura 44. Transición de la molécula MMV1580167 al ligando AS2.....	34
Figura 45. Posibles sitios para la incorporación de grupos funcionales al ligando MMV002137	34
Figura 46. Transición de la molécula MMV002137 al ligando AS6.....	35
Figura 47. Transición de la molécula MMV002137 al ligando AS7.....	35
Figura 48. Posibles sitios para la incorporación de grupos funcionales al ligando MMV002137	36
Figura 49. Transición de la molécula MMV003406 al ligando AS3.....	36
Figura 50. Transición de la molécula MMV003406 al ligando AS4.....	37
Figura 51. Transición de la molécula MMV003406 al ligando AS5.....	37
Figura 52. Entorno del dominio Arpeggio	38
Figura 53. Configuración para la obtención de interacciones intermoleculares ligando- proteína.....	39
Figura 54. Interacciones intermoleculares calculadas, visualizadas en Pymol	39
Figura 55. Proteasa Mpro de: A: Sars CoV-2; B: Sars CoV; C: Mers CoV; D: HKU1 CoV; E: NL63 CoV; F: 229E CoV.....	41
Figura 56. Alineamiento múltiple de secuencias de Mpro de diferentes virus.....	47
Figura 57. Contraste entre residuos rígidos y flexibles de la proteasa Mpro del virus SARS CoV-2	60
Figura 58. Interacciones intermoleculares ligando MMV1580167	71

Figura 59. Interacciones intermoleculares ligando MMV002137	72
Figura 60. Interacciones intermoleculares ligando MMV003406	72
Figura 61. Interacciones intermoleculares ligando AS7	73
Figura 62. Interacciones intermoleculares ligando AS1	73

ÍNDICE DE TABLAS

Tabla 1. Clasificación Taxonómica del virus SARS-CoV-2.....	4
Tabla 2. Proteínas principales del virus SARS-CoV-2	5
Tabla 3. Proteínas no estructurales del virus SARS-CoV-2.....	5
Tabla 4. Valores de E-Value.....	16
Tabla 5. Valores de Identidad.....	16
Tabla 6. Comparación estructural entre Mpro del virus SARS-CoV-2 y proteasas virales.	41
Tabla 7. Aminoácidos objetivo de inhibidores de la proteasa Mpro en coronavirus humanos44	
Tabla 8. Aminoácidos Principales del sitio Activo de la Enzima Mpro	46
Tabla 9. Imágenes de las moléculas la librería de Proteín Data Bank	48
Tabla 10. Identificación de residuos flexibles	59
Tabla 11. Resultado Docking Molecular del Cribado Virtual.....	61
Tabla 12. Acoplamiento de los mejores inhibidores identificados, dentro de la proteasa Mpro	62
Tabla 13. Resultados del docking de los ligandos diseñados en la proteasa Mpro	63
Tabla 14. Superposición de ligandos mejorados a ligandos iniciales	64
Tabla 15. Superposición de los mejores Ligandos a N3	68
Tabla 16. Identificación de interacciones Intermoleculares	71
Tabla 17. Librería de ligandos obtenidos de Pdbe Fold.....	84
Tabla 18. Aminoácidos de Mpro que interaccionan con los ligandos	95
Tabla 19. Estructura 2D de ligandos diseñados.....	102

RESUMEN

El presente proyecto de investigación fue realizado con la finalidad de encontrar y desarrollar ligandos con actividad inhibitoria contra la proteasa Mpro del virus SARS CoV-2 mediante un cribado virtual de inhibidores recientemente desarrollados contra este y aquellos que han sido probados anteriormente en virus patógenos humanos, para lo cual se recurrió a librerías de compuestos como Protein Data Bank de donde se escogió 46 moléculas con una alta afinidad de unión a Mpro a partir de un estudio de homología estructural con las proteasas de otros virus que atacan al ser humano y Covid Box, librería a la que se puede ingresar desde el siguiente dominio: <https://www.mmv.org/mmv-open/covid-box> de donde se tomó un total de 79 moléculas, los ligandos seleccionados se presentaban como inhibidores de Mpro, a partir de la formación de interacciones hidrofóbicas, polares y enlaces pi entre átomos de carbono – carbono, azufre – carbono y flúor – carbono. Para seleccionar las mejores moléculas se realizó un análisis bioinformático en software como Wincoot, Pymol y GOLD - Protein Ligand Docking Software, este último para ensayos de docking o acoplamiento molecular, estudiado a nivel conformacional por la disposición y solapamiento de átomos, nivel de incidencia de interacciones intermoleculares y posibilidad de choques estéricos de sus grupos funcionales, obteniendo como resultado que la moléculas **MMV1580167** de Covid Box presenta la mejor actividad inhibitoria dentro del cribado virtual, mientras que el ligando **AS7** diseñado, presenta la más alta actividad inhibitoria en el entorno virtual donde se realizó la investigación.

Palabras Clave: Cribado virtual, SARS CoV-2, virus patógenos humanos, bioinformática, inhibidores enzimáticos, diseño racional, actividad antiviral

ABSTRACT

The present research project was carried out in order to find and develop ligands with inhibitory activity against the Mpro protease of the SARS CoV-2 virus through a virtual screening of recently developed inhibitors against this and those that have been previously tested in human pathogenic viruses. For which, compound libraries such as Protein Data Bank were used, from which 46 molecules with a high binding affinity to Mpro were chosen from a study of structural homology with the proteases of other viruses that attack humans and Covid Box. library that can be accessed from the following domain: <https://www.mmv.org/mmv-open/covid-box> from where a total of 79 molecules were taken, the selected ligands were presented as inhibitors of Mpro, at starting from the formation of hydrophobic, polar interactions and pi bonds between carbon - carbon, sulfur - carbon and fluorine - carbon atoms. To select the best molecules, a bioinformatic analysis was carried out in software such as Wincoot, Pymol and GOLD - Protein Ligand Docking Software, the latter for docking or molecular docking tests, studied at a conformational level by the arrangement and overlapping of atoms, level of incidence of intermolecular interactions and the possibility of steric shocks of their functional groups, obtaining as a result that the **MMV1580167** molecule from Covid Box presents the best inhibitory activity within the virtual screening, while the designed **AS7** ligand presents the highest inhibitory activity in the virtual environment where the investigation was carried out.

Key Words: Virtual screening, SARS CoV-2, human pathogenic viruses, bioinformatics, enzyme inhibitors, rational design, antiviral act

CAPÍTULO I

MARCO TEÓRICO

1.1 ANTECEDENTES INVESTIGATIVOS

En los últimos días del año 2019, en la ciudad China de Wuhan se detectó casos de un nuevo tipo de infección, que con el pasar de los días derivaba en una neumonía de causas desconocidas. Investigaciones posteriores arrojaron resultados que establecían que la enfermedad era causada por un nuevo tipo de coronavirus que por primera vez infectaba a los seres humanos. Este fue bautizado como SARS-CoV-2, agente causal de la enfermedad conocida hoy en día como COVID-19 y cuyos síntomas más representativos son la tos, fiebre y dificultad respiratoria (Accinelli et al., 2020).

La capacidad del virus para infectar de manera relativamente fácil a las células humanas, y su supervivencia en varias estructuras como el vidrio, metal, plástico y otros componentes de uso común en el día a día de las personas, ha causado su expansión a todos los rincones del planeta. Nuestro país, Ecuador, no fue la excepción, causando así una crisis sanitaria generalizada. Según los datos recopilados por la Universidad Johns Hopkins, el COVID-19 alcanza, al 2 de octubre de 2020, un total de 34'289.709 casos confirmados a nivel mundial de los cuales 7'781.877 siguen aún en estado activo. El número de fallecidos asciende a 1'022.878 personas que representan el 2,98% del total de infectados (Dong, Du & Gardner, 2020). En el Ecuador, según datos oficiales, la cifra de contagios a la misma fecha alcanzó ya los 138.584, con un total de 7.621 fallecimientos confirmados y 3.812 fallecimientos probables. Esto indica que la tasa de mortalidad en nuestro país es del 5,73%, para fatalidades confirmadas por COVID-19, y 8,97% si se toma en cuenta el número total de probables fallecimientos por esta enfermedad (11.433 fallecidos) (COE Nacional, 2020). Estos valores denotan que el Ecuador presenta una elevada tasa de mortalidad con respecto a la media mundial y hacen eco de la extrema necesidad de contar con un tratamiento efectivo que pueda contribuir a la recuperación de los pacientes que padecen la enfermedad, especialmente aquellos que desarrollan cuadros agudos y están en riesgo de muerte.

Por otra parte, al tratarse de un virus que ha infectado a los seres humanos recientemente, no existe un tratamiento efectivo contra la enfermedad que produce, lo que ha derivado en las muertes antes mencionadas. Si bien desde el descubrimiento del virus varias naciones han aumentado sus esfuerzos en la síntesis de una vacuna, no es sencillo disponer de la misma en un corto tiempo, además de que esta no representa en sí una cura para la enfermedad, más bien un tratamiento preventivo. Científicos enuncian que el coronavirus SARS-CoV-2 permanecerá en el planeta indefinidamente como un virus endémico que infectará a personas a lo largo de la historia de manera similar a como lo hace la gripe. Además, teniendo en cuenta que al tratarse de un virus tiende a mutar fácilmente, las probabilidades de erradicarlo con una vacuna son reducidas (Cohen & Kupferschmidt, 2020).

1.2 1.2 EL VIRUS SARS CoV-2

1.2.1 Origen

A través de los años, un denominador común en la aparición de nuevos virus capaces de infectar a los seres humanos ha sido la zoonosis, una condición donde virus patógenos de animales adquieren la capacidad de infectar personas. Esto ocurre debido a procesos naturales de mutación y selección en los cuales los patógenos obtienen la habilidad de infectar células humanas (Paules, Marston, & Fauci, 2020).

Investigaciones sugieren que el virus SARS-CoV-2 tuvo su origen en los murciélagos, animales conocidos por ser portadores de varias especies de coronavirus, y que en ocasiones anteriores han propiciado epidemias menores como es el caso del coronavirus SARS-CoV que en el año 2003 causó la infección de 8421 personas, con el fallecimiento de 784 de ellos (Zambrano & Loaiza, 2003). Además, otra especie de coronavirus que infecta naturalmente al murciélago ocasionó la diarrea aguda porcina y la muerte de más de 24 mil lechones en China (Sun et al., 2015). Se cree que el virus causante de la pandemia actual no es la excepción, ya que mediante la secuenciación del genoma del SARS-CoV-2 se determinó una similitud de 96,2% respecto al coronavirus RaTG13 que infecta murciélagos de la especie *Rhinolophus affinis* y un 79,5% respecto al virus humano SARS-CoV (Zhou et al., 2018).

Un factor importante para el desarrollo de la capacidad infecciosa de estos virus en humanos, son los animales intermediarios. En este punto, se tiene como antecedente la infección por el MERS-CoV que en el año 2012 infectó a 2492 personas, causando la

muerte de 858 de ellas. Para este caso los intermediarios fueron los camellos, los cuales a su vez resultaron contagiados por mordeduras de murciélagos (Bratanich, 2015). No se conoce a ciencia cierta si existió o no un animal intermediario que facilitara la zoonosis del virus SARS-CoV-2, pero se maneja la hipótesis de que el pangolín o las civetas sean los animales en cuestión (Figura 1). Esta idea ha sido fundamentada con estudios genéticos de coronavirus aislados de estos animales, de los cuales se encontró una similitud altamente significativa con el SARS-CoV-2 (Zhou et al., 2020).

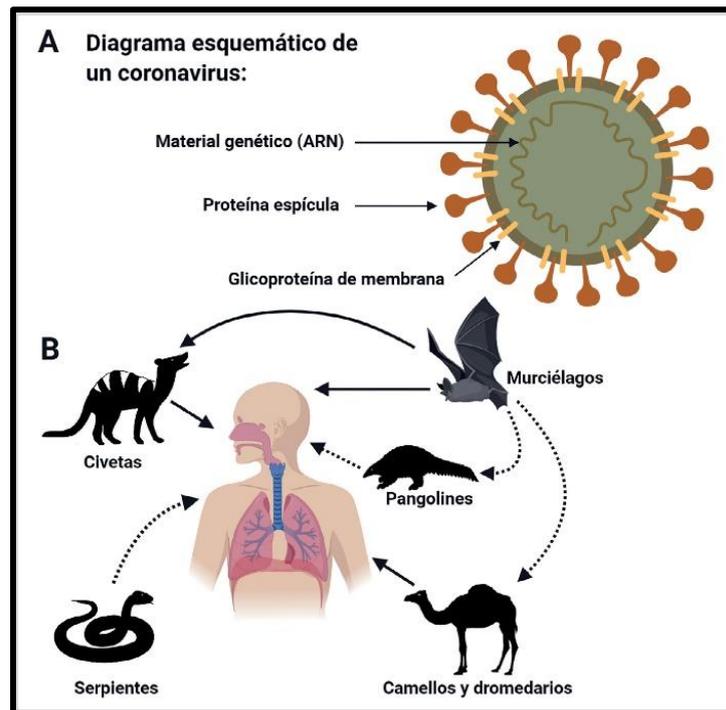


Figura 1. Transmisión zoonótica de coronavirus

Fuente: (Cortés, 2020)

El principal detonante para el traspaso del virus de animales al ser humano fue la venta ilegal de animales silvestres, entre los que se encuentran los anteriormente mencionados. Además estos animales se comercializan vivos y en condiciones insalubres en un mercado de la ciudad China de Wuhan, lugar donde se originó el contagio (Guo et al., 2020).

En contraste a lo expuesto anteriormente, estudios un tanto polémicos sugieren que el virus fue creado artificialmente en un laboratorio y diseminado entre la población por un escape accidental durante investigaciones (Andersen et al., 2020). En relación a esta hipótesis, anteriormente se ha documentado en varias ocasiones la fuga del virus SARS-CoV de laboratorios de bioseguridad nivel 2 (Christian, Poutanen, Loutfy, Muller, & Low, 2004) lo cual hace creíble la teoría de que el SARS-CoV-2 también escapó de uno

de estos laboratorios. De todas maneras, el origen del SARS-CoV-2 aún debe ser decifrado.

El virus SARS-CoV-2 se suma al total de 7 coronavirus capaces de causar una infección al ser humano. De estos, los virus HKU1, NL63, OC43 y 229E causan una sintomatología leve similar al resfriado al ingresar al organismo (Corman et al., 2018).

1.2.2 Características y Estructura del virus SARS-CoV-2

El SARS-CoV-2 es un virus catalogado como coronavirus del género *Betacoronaviridae* (Tabla 1), por su característica corona semejante a la del sol que lo rodea, observada en un microscopio electrónico.

Tabla 1. Clasificación Taxonómica del virus SARS-CoV-2

Dominio	<i>Riboviria</i>
Reino	<i>Orthornavirae</i>
Filo	<i>Pisuviricota</i>
Clase	<i>Pisoniviricetes</i>
Orden	<i>Nidovirales</i>
Suborden	<i>Cornidovirineae</i>
Familia	<i>Coronaviridae</i>
Género	<i>Betacoronaviridae</i>
Especie	Coronavirus del síndrome respiratorio agudo grave
Subespecie	Coronavirus 2 del síndrome respiratorio agudo grave

Fuente: Taxonomy browser (2020)

Su material genético consta de ARN de cadena sencilla dispuesto en sentido positivo y no segmentado, del cual sus marcos de lectura abiertos (ORFs) 1a y 1b (ORF1a y ORF1ab) se encargan de codificar proteínas esenciales para el proceso infeccioso dentro del huésped. Una de estas proteínas es glicoproteína espiga “S” que es la que se encarga del acoplamiento a las membranas celulares de las células humanas y proteínas que se encargan de la entrega del material viral y elementos complementarios que facilitarán la formación de la cápside de los nuevos virus (De Wit, Van Doremalen, Falzarano, & Munster, 2016). Estas proteínas esenciales se resumen en la tabla 2.

Tabla 2. Proteínas principales del virus SARS-CoV-2

Proteína	Función
S	Acoplamiento a células humanas
M	Entrega del material viral
N	Formación de la nucleocápside
E	Formación de la envoltura viral

Fuente: (De Wit et al., 2016).

Por otra parte, un tercio del material genético total del virus, se encarga de la formación de 16 proteínas no estructurales (NSPs) complementarias, dentro de las cuales se encuentra la proteasa principal M^{pro} y otras con funciones descritas (**Tabla 3**).

Tabla 3. Proteínas no estructurales del virus SARS-CoV-2

Proteína	Función
Nsp1	Degrada el ARNm dentro de las células
Nsp2	No definida
Nsp3	Interferencia con la respuesta inmune del hospedero
Nsp4	Formación de vesículas de doble membrana
Nsp5	Proteasa principal M^{pro} que escinde las poliproteínas
Nsp6	Restringe la expansión del auto fagosoma
Nsp7	Cofactor con nsp8 y nsp12
Nsp8	Cofactor con nsp7 y nsp12
Nsp9	Unión al ARN viral y dimerización
Nsp10	Activa enzimas replicativas
Nsp11	No definida
Nsp12	ARN polimerasa
Nsp13	ARN helicasa
Nsp14	Exorribonucleasa
Nsp15	Endorribonucleasa
Nsp16	Inhibe respuesta inmunitaria innata

Fuente: (Herrera, 2020)

1.3 PATOGENICIDAD DEL VIRUS SARS-CoV-2

1.3.1 Proteasa Principal Del Virus SARS-CoV-2 (*M^{pro}*)

M^{pro} es proteasa del tipo cisteína, llamada también tiol proteasa, que se encarga de la escisión de las poliproteínas virales pp1a y pp1ab. Para activarse, *M^{pro}* requiere auto escindirse entre las proteínas complementarias nsp4 y nsp6. Estructuralmente *M^{pro}* está conformada por un homodímero, con cada monómero organizado en tres dominios: I, II y III. El dominio I que comprende los residuos 8-101, el II abarca los residuos 102-184 y el III está compuesto de residuos entre las posiciones 201-306. Entre el dominio I y II se sitúa una diada catalítica Cys-His no canónica (Anand, Ziebuhr, Wadhwani, Mesters, & Hilgenfeld, 2003). El sitio activo se encuentra formado por una hendidura entre los dominios I y II mediante una unión previa de los protómeros que forman la proteasa mediante un dedo N-terminal (Yang et al., 2003).

La proteasa *M^{pro}* presenta una alta especificidad de sustrato, de tal manera que solo escinde en un sitio precedido por el aminoácido glutamina (Gln). La especificidad de la proteasa se encuentra dirigida hacia los residuos de sustrato en el rango de P4 a P1 según la nomenclatura de Schechter y Berger, determinándose que el sitio de escisión se encuentra entre P1 y P1', siendo P1 el residuo que requiere de la unión de la glutamina (Figura 2) (Muramatsu et al., 2016).

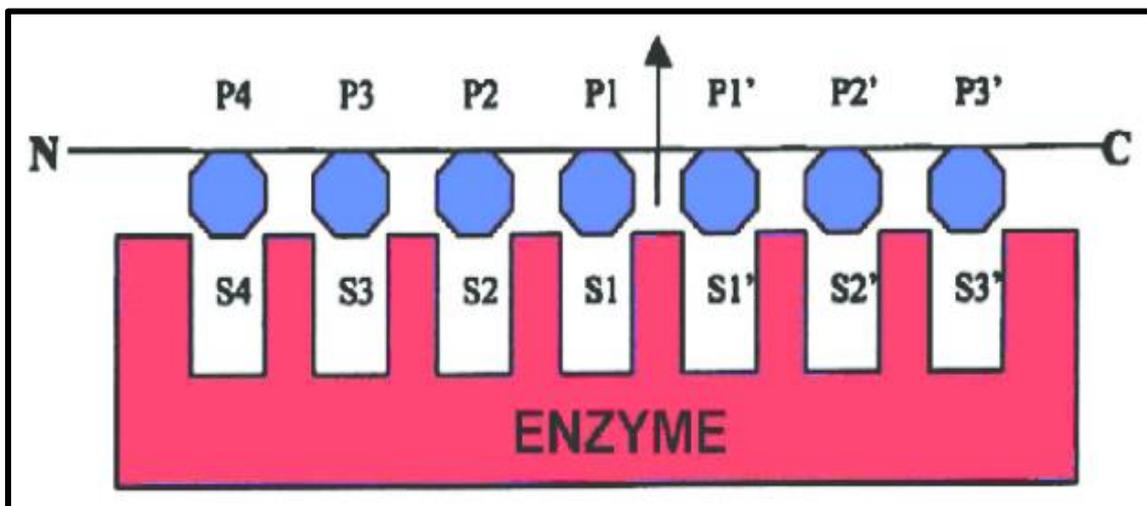


Figura 2. Esquema de Schechter y Berger

Fuente: (Rzychon et al,2004)

1.3.2 Mecanismo de Infección del Virus SARS-CoV-2

1.3.2.1 Infección de células huésped

El mecanismo general de infección por parte de un virus consiste en depositar al interior de la célula su material genético de tal manera que se obligue a esta a replicarlo. El SARS-CoV-2 lleva a cabo este proceso mediante su proteína “S” que con su forma característica de espiga se une al receptor ACE2 de las células humanas. Este último es el encargado de recibir la enzima convertidora de angiotensina 2 que participa en la regulación de la presión arterial; por lo tanto, el virus se unirá a tejidos y órganos que dispongan de este receptor en sus células componentes. El más importante de estos órganos es el pulmón, el cual alberga el receptor ACE2 en sus células alveolares. Sin embargo, el receptor también se encuentra en células del riñón, corazón, hígado, intestino, vejiga, esófago y en las células de revestimiento de la cavidad oral (Hao et al., 2020).

La infección de las células por el virus SARS-CoV-2 se da por una endocitosis final, para lo cual participan las 2 subunidades de la proteína S, S1 y S2, donde S1 es la encargada de unirse al receptor ACE2, mientras que S2 se encarga de acoplar la membrana viral a la celular. Seguidamente, la proteasa TMPRSS2 realiza un corte en 2 posiciones de S2 para separarse de S1, lo cual ocasiona la fusión de membranas que acaba llevando la carga viral hacia el interior de la célula (Ali et al., 2020).

1.3.2.2 Transcripción y traducción de las proteínas virales

Una vez que el virus ha ingresado al interior de la célula, la nucleocápside libera su genoma, el cual es confundido con el ARNm y transcrito directamente por los ribosomas. Los genes ORF1a y ORF1ab, precursores de las poliproteínas homónimas pp1a y pp1ab, respectivamente, son expresados primeramente en las fases tempranas de la infección. Finalmente, estas poliproteínas son escindidas por un coctel de proteasas entre las que se destacan la quimiotripsina viral, proteasas similares a la papaína y la proteasa principal *M^{pro}*. Luego de un exhaustivo procesamiento proteolítico se obtienen 16 NSPs virales que en conjunto conformarán el complejo replicasa transcriptasa (RTC) ensamblado en vesículas de doble membrana. El RTC tiene la función de sintetizar ARN subgenómico (ARNsg) que dará origen a las proteínas estructurales S, M, E, y N, las cuales se empaquetarán en el Aparato de Golgi de la célula dando lugar a nuevas partículas virales. Estas partículas se liberarán posteriormente en el citoplasma, de donde emergen mediante lisis celular cuando la carga viral es alta (Figura 3) (Mousavizadeh & Ghasemi, 2020).

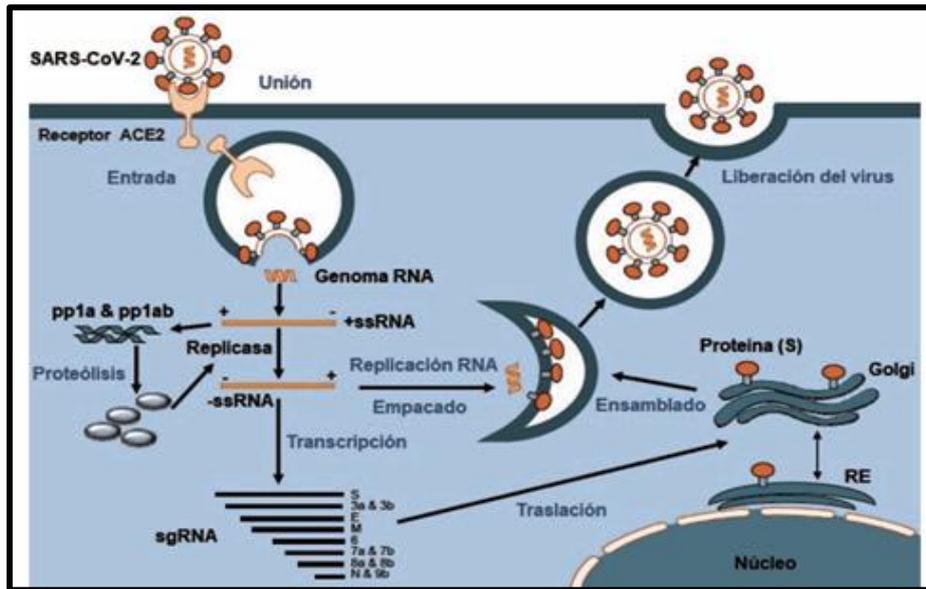


Figura 3. Proceso de infección del virus SARS-CoV-2 a la célula

Fuente: (Zhu et al., 2013)

1.4 DISEÑO DE INHIBIDORES COMO TRATAMIENTO A LA INFECCIÓN POR SARS-CoV-2

La alta similitud en los sitios de reconocimiento y escisión de poliproteínas por parte de los virus del SARS, MERS y SARS-CoV-2, sugiere que la proteasa M^{pro} es un candidato altamente atractivo para el desarrollo de fármacos. Esta hipótesis se fundamenta en el hecho de que M^{pro} cuenta con bajos niveles de mutagenicidad y que en el cuerpo humano no existe ninguna molécula homóloga a dicha proteína (Hilgenfeld, 2014).

1.4.1 Inhibidores

Inicialmente, la comunidad científica ha evaluado la efectividad de inhibidores que han sido destinados a contrarrestar los virus del SARS y MERS, mediante la anulación del residuo de cisteína catalítica, presente en este tipo de proteasas, mediante la formación de un enlace covalente entre este residuo y el inhibidor (Dai et al., 2020). Sin embargo, estos no han arrojado aún resultados totalmente convincentes. En M^{pro} también se ha considerado que el principal objetivo para el desarrollo de inhibidores es la Cisteína catalítica, Cys145. En este sentido, se han evaluado diversos tipos de moléculas entre las que destacan peptidomiméticos, que compiten con los verdaderos sustratos de la enzima,

capaces de formar un hemitioacetal con Cys145 gracias a que esta reacciona con el grupo funcional alfa-acetamida presente en el inhibidor (Zhang et al., 2020).

Compuestos con anillos de piridona son capaces de colocarse entre P2 y P3 y cuentan con la característica de poseer una alta vida media en el citoplasma y tener efectividad horas después de haber ingresado al cuerpo (Yang et al., 2005). Por otra parte, compuestos peptidomiméticos con un grupo indol en el extremo N-terminal unido a P3 se muestran como los inhibidores más efectivos conocidos hasta la fecha, gracias al aldehído unido al extremo C-terminal el cual se une covalentemente a Cys145. (Ullrich & Nitsche, 2020).

Estudios recientes a su vez han identificado 6 moléculas con una alta capacidad de inhibición de la proteasa *M^{pro}*, estas son: carmofur, ebselen, PX12, disulfiram, tideglusib, shikonin (Figura 4), de las cuales las 3 primeras se unen covalentemente a Cys145 (Jin et al, 2020).

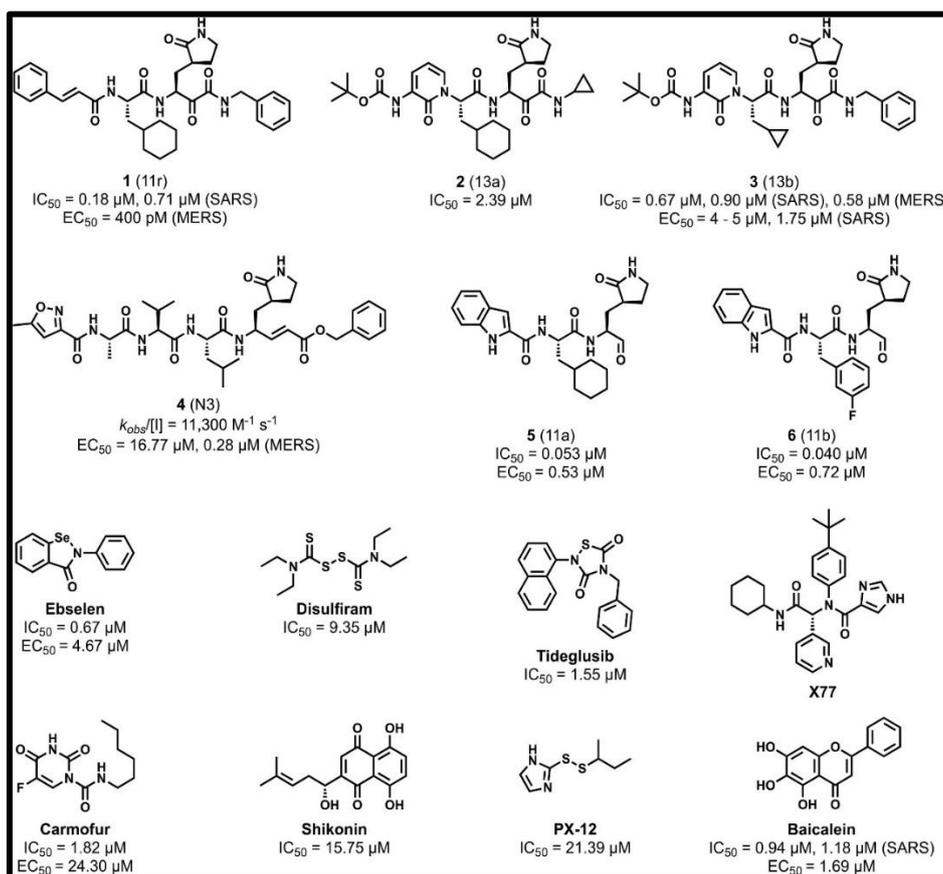


Figura 4. Principales inhibidores de la proteasa *M^{pro}*

Fuente: (Ullrich & Nitsche, 2020).

1.5 OBJETIVOS

1.5.1 OBJETIVO GENERAL

Realizar un cribado virtual de inhibidores de la enzima M^{pro} del virus SARS-CoV-2 empleando librerías de compuestos y homología estructural con proteasas de virus patógenos de humanos.

1.5.2 OBJETIVOS ESPECÍFICOS

- Analizar la homología estructural entre la enzima M^{pro} del virus SARS-CoV-2 y proteasas de otros virus que atacan al ser humano.
- Determinar sitios de posible unión de inhibidores a la enzima M^{pro} del virus SARS-CoV-2.
- Generar una librería de compuestos para análisis de acoplamiento molecular.
- Determinar los mejores inhibidores de la proteasa M^{pro} mediante el análisis de las propiedades su unión con la enzima.

CAPÍTULO II

METODOLOGÍA

1.6 MATERIALES

2.1.1 Material Virtual

- Artículos Científicos
- Base de Datos de Uniprot
- Base de Datos de PDBE Fold
- Herramientas de BLAST y Alineamientos

2.1.2 Software Bioinformático

- Wincoot versión 0.8.7
- Pymol
- GOLD - Protein Ligand Docking Software
- Autodock

2.1.3 Equipos

- ❖ Ordenador

2.1.4 Organismo estudiado

- ❖ SARS-CoV-2: Coronavirus 2 causante del Síndrome Respiratorio Agudo Grave

1.7 PROCEDIMIENTOS CUALITATIVOS

2.2.1 Obtención de la secuencia y estructura de la proteasa M^{pro}

Para la obtención de la secuencia de aminoácidos que componen la proteasa M^{pro} , se recurrió a la plataforma de base de datos de proteínas conocida como Protein Data Bank (PDB por sus siglas en inglés, <https://www.rcsb.org/>). De la base de datos del PDB se obtuvo la secuencia de aminoácidos de la enzima en su formato FASTA. Mediante una búsqueda preliminar se determinó que los datos depositados con el código PDB 6YB7 y que corresponden a la proteasa M^{pro} son aptos para esta investigación. De 6YB7 se

descargó su secuencia mediante el menú “Download files”, así como también el archivo con extensión .pdb que contiene los datos de su estructura tridimensional.

2.2.2 Análisis de homología de M^{pro} de SARS-CoV-2 en PDBeFold

PDBeFold (<https://www.ebi.ac.uk/msd-srv/ssm/>) es un recurso web que compara la estructura de una proteína objetivo con la extensa base de datos del PDB (170000 estructuras), la cual incluye diversas estructuras peptídicas y sus ligandos e inhibidores. Para esta comparación la herramienta tomó como referencia la estructura secundaria de las moléculas (SSM) y arrojó valores de similitud acorde a la disposición tridimensional de sus átomos componentes.

Mediante el uso de esta herramienta, se determinó estructuras homólogas a la proteasa M^{pro} en virus patógenos a humanos y se recopiló datos de ligandos e inhibidores de interés.

2.2.3 Análisis de homología de M^{pro} de SARS-CoV-2 en UniProt

Por otra parte, ya que la base de datos del PDB es algo limitada y proporciona únicamente información de aquellas macromoléculas de las cuales se ha determinado su estructura tridimensional, se realizó un análisis de homología de la secuencia de aminoácidos de la enzima M^{pro} mediante la base de datos Universal Protein (UniProt). UniProt es una base de datos gratuita que registra una gran cantidad de información sobre proteínas, incluyendo su secuencia de aminoácidos. Esta información proviene de secuenciaciones genómicas publicadas en revistas científicas y depositada en bases de datos como el GenBank. Para realizar el análisis de identidad de secuencia UniProt utilizó el algoritmo BLAST, herramienta de búsqueda de alineación local básica, que permite localizar regiones similares entre la secuencia objetivo con todos los archivos depositados en UniProt. De esta forma, tanto el programa como el usuario pueden inferir relaciones evolutivas y funcionales de la proteína en estudio, mediante ciertos parámetros cuantificables.

Este sitio web de UniProt (<https://www.uniprot.org/>), provee de herramientas que serán útiles en los análisis posteriores, tales como el “BLAST” que permitirá la obtención de secuencias proteica de aminoácidos homólogas y “Align” o alineamiento por su traducción, que permitirá comparar secuencias de proteasas de diversos organismos con la del SARS-CoV-2.

2.2.4 Acoplamiento Molecular

El acoplamiento molecular se llevó a cabo mediante el uso de programas bioinformáticos como GOLD - Protein Ligand Docking Software o AutoDock que evaluaron las diversas conformaciones adquiridas por una molécula en un sitio específico de la proteína objetivo, mediante un algoritmo genético, permitiendo la identificación del mejor inhibidor en el mejor sitio drogable de la proteína.

Programas bioinformáticos como Wincoot y Pymol en primera instancia y gracias a su capacidad de visualización tridimensional de estructuras proteicas, otorgaron un punto de partida en la identificación de los sitios drogables de la proteína objetivo, donde se determinó los aminoácidos afines a inhibidores y posteriormente se analizó los sitios de unión y conformación de los inhibidores de prueba.

1.8 MÉTODOS

1.8.1 ANÁLISIS DE HOMOLOGÍA DE PROTEÍNAS

Mediante el uso del navegador de un ordenador, se ingresó hacia las bases de datos de proteínas (UNIPROT, PDBe FOLD) donde se encuentran todas aquellas que han sido estudiadas hasta la fecha.

En Uniprot: Acceder al siguiente enlace: <https://www.uniprot.org/>

1. Se colocó el código de la proteína deseada (6yb7) en la barra de búsqueda de Uniprot, este desplegó entradas conforme a la molécula solicitada.



Figura 5. Entorno de Uniprot

- Se seleccionó la entrada correspondiente para desplegar el menú donde se encuentra toda la información de la proteína y descargar la secuencia fasta de la barra de herramientas, según el botón señalado (**Figura 6**).

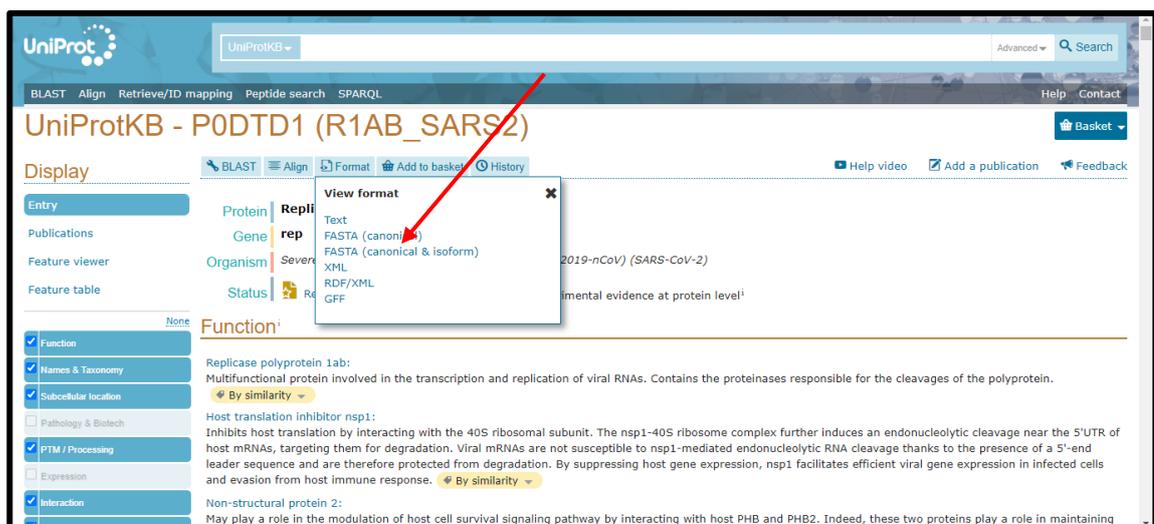


Figura 6. Menú desplegable para la descarga de secuencias en formato FASTA

- Se identificó en la secuencia FASTA los aminoácidos que permiten a la proteína tener su actividad, en este caso infecciosa, y se realizó un alineamiento (BLAST) frente a toda la base de datos, usando la propia herramienta de Uniprot, con la finalidad de encontrar estructuras homólogas, la secuencia se coloca en el espacio señalado (**Figura 7**).

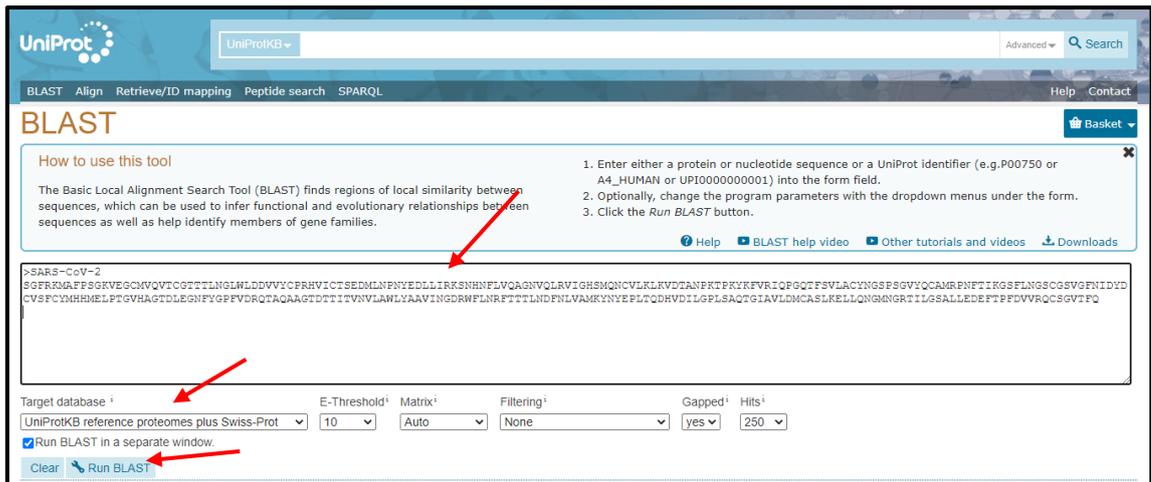


Figura 7. Herramienta Blast de Uniprot

Nota: Es importante realizar una modificación en el nombre de la proteína a continuación del símbolo “>” con la finalidad de tener una mejor identificación de la misma dentro del alineamiento.

- Del total de organismos homólogos encontrados, se hizo énfasis en los virus, que fueron separados acorde al enfoque del trabajo de investigación (virus humanos), para desplegar el alineamiento se seleccionó “View alignment” (Figura 8)

Entry	Alignment overview	Info	Status
<input type="checkbox"/> Query: SARS-CoV-2 B202009245C475328CEF75220C360D524E9D456CE0E49259			
<input type="checkbox"/> P0DTC1	R1A_SARS2 - Replicase polyprotein 1a - Severe acute res... - View alignment	E-value: 0.0 Score: 1,646 Ident.: 100.0%	
<input type="checkbox"/> P0DTD1	R1AB_SARS2 - Replicase polyprotein 1ab - Severe acute res... - View alignment	E-value: 0.0 Score: 1,646 Ident.: 100.0%	
<input type="checkbox"/> P0C6F5	R1A_BC279 - Replicase polyprotein 1a - Bat coronavirus ... - View alignment	E-value: 0.0 Score: 1,604 Ident.: 96.1%	
<input type="checkbox"/> P0C6V9	R1AB_BC279 - Replicase polyprotein 1ab - Bat coronavirus ... - View alignment	E-value: 0.0 Score: 1,604 Ident.: 96.1%	
<input type="checkbox"/> A0A0U1WHG8	A0A0U1WHG8_CVHSA - 2'-O-methyltransferase BTRF-BetaCoV/HeB2013 - View alignment	E-value: 0.0 Score: 1,601 Ident.: 95.8%	
<input type="checkbox"/> P0C6U8	R1A_CVHSA - Replicase polyprotein 1a - Human SARS coron... - View alignment	E-value: 0.0	

Figura 8. Resultados de homología arrojados por Uniprot

- Se observó los alineamientos individuales de la proteasa del coronavirus SARS-CoV-2 con sus virus homólogos, identificando también el sitio de la proteína donde se ha encontrado la similitud.

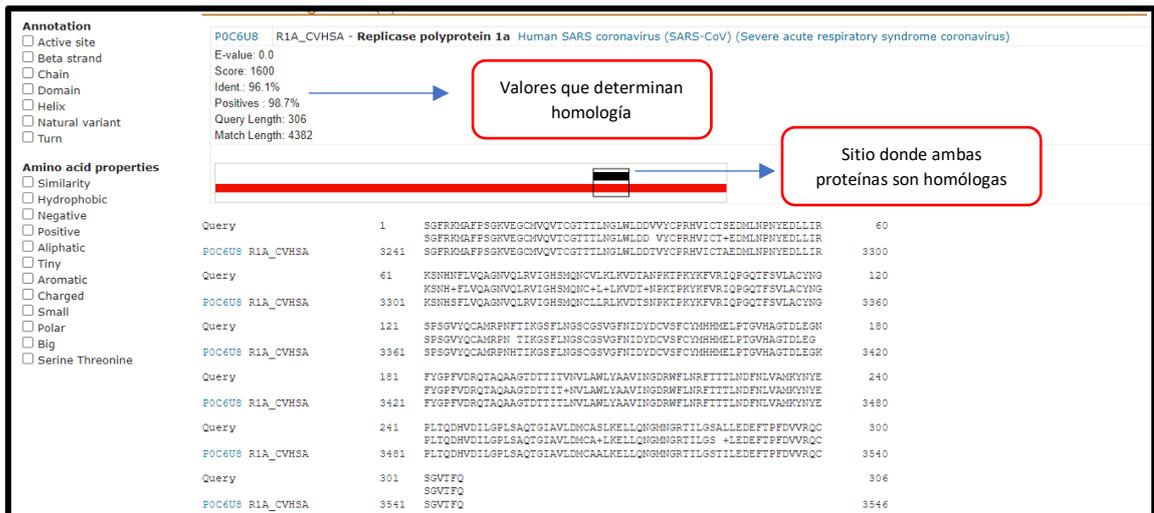


Figura 9. Obtención de información de alineamientos individuales

6. Se determinó aquellas proteínas con mayor similitud y alto valor investigativo a través de valores como el E-value y el % de Identidad por los siguientes criterios:

Tabla 4. Valores de E-Value

E-value $\leq 0,02$	Secuencias Probablemente Homólogas
$0,02 < \text{E-value} \leq 1$	Homología no descartable, corroborar con otros análisis
E-value > 1	Homología no significativa

Tabla 5. Valores de Identidad

ID $\geq 45\%$	Estructura similar y posible función similar
ID $\geq 25\%$	Similitud en el plegamiento 3D general
ID = 18-25%	Zona gris (similitud difícil de asignar)
ID $< 20\%$	No existe similitud

En Pdbe Fold: Se ingresó al siguiente link: <https://www.ebi.ac.uk/msd-srv/ssm/>

1. Se seleccionó la opción “Launch Pdbe” y en las opciones desplegadas, se colocó el código de la proteína (6yb7) y se definió el porcentaje de similitud deseado, seguidamente enviar a consulta (**Figura 10**).

Figura 10. Entorno de Protein Data Bank

2. En la lista de organismos e inhibidores desplegada, se realizó una selección previa acorde a los requerimientos de investigación (microorganismos patógenos al ser humano), tomando nota de los códigos correspondientes (**Figura 11**).

Structure Alignment Results.

Query: pdb entry 6yb7:A , 306 residues
SARS-COV-2 MAIN PROTEASE WITH UNLIGANDED ACTIVE SITE (2019-NCOV, CORONAVIRUS DISEASE 2019, COVID-19)

Examined 163448 entries, (439847 chains). Displaying Matches 1-20 of 461.

Back to query | next | last page | Sort by: Q-score | arrange by SCOP family | match 1 | jump

#	Scoring			RMSD	N _{align}	N _{seq}	%Seq	Query			Target (PDB entry)		
	Q	P	Z					%Seq	Match	%Seq	N _{seq}	Title	
1	1.00	74.2	26.0	0.01	306	0	100	100	6yb7:A	100	306	<input type="checkbox"/>	SARS-COV-2 MAIN PROTEASE WITH UNLIGANDED ACTIVE SITE (2019-NCOV, CORONAVIRUS DISEASE 2019, COVID-19)
2	1.00	66.2	24.5	0.14	306	0	100	100	7k3t:A	100	306	<input type="checkbox"/>	CRYSTAL STRUCTURE OF SARS-COV-2 MAIN PROTEASE (3CLPRO/MPRO) IN THE APO FORM AT 1.2 Å RESOLUTION
3	0.99	64.7	24.2	0.24	306	0	100	100	6ymq:A	100	306	<input type="checkbox"/>	STRUCTURE OF SARS-COV-2 MAIN PROTEASE BOUND TO 2-METHYL-1-TETRALONE.
4	0.99	58.6	23.0	0.25	306	0	100	100	6yt8:A	100	306	<input type="checkbox"/>	STRUCTURE OF SARS-COV-2 MAIN PROTEASE BOUND TO PYRITHIONE ZINC
5	0.99	74.2	26.0	0.00	304	0	100	100	6y84:A	100	304	<input type="checkbox"/>	COVID-19 MAIN PROTEASE WITH UNLIGANDED ACTIVE SITE (2019-NCOV, CORONAVIRUS DISEASE 2019, SARS-COV-2)
6	0.99	74.2	26.0	0.10	304	0	100	100	5rfe:A	100	304	<input type="checkbox"/>	PANDDAANALYSIS GROUP DEPOSITION – CRYSTAL STRUCTURE OF SARS-COV-2 MAIN PROTEASE IN COMPLEX WITH Z126932614
7	0.99	74.2	26.0	0.10	304	0	100	100	5r8t:A	100	304	<input type="checkbox"/>	PANDDAANALYSIS GROUP DEPOSITION OF GROUND-STATE MODEL OF SARS-COV-2 MAIN PROTEASE SCREENED AGAINST DSI POISED (ENAMINE), FRAGLITES AND PEPLITES (NEWCASTLE UNIVERSITY), MINI FRAGS (ASTEX), YORK 3D (YORK UNIVERSITY), ELECTROPHILE CYSTEINE COVALENT (WEIZMAN INSTITUTE) FRAGMENT LIBRARIES
8	0.99	63.2	23.9	0.26	306	0	100	95	6yvf:A	100	306	<input type="checkbox"/>	STRUCTURE OF SARS-COV-2 MAIN PROTEASE BOUND TO AZD6482
9	0.99	74.2	26.0	0.10	304	0	100	100	5rfe:A	100	304	<input type="checkbox"/>	PANDDAANALYSIS GROUP DEPOSITION – CRYSTAL STRUCTURE OF SARS-COV-2 MAIN PROTEASE IN COMPLEX WITH

Figura 11. Lista de resultados arrojados por Protein Data Bank

- Una vez obtenidos los códigos de las estructuras deseadas, se ingresó al dominio: <https://www.rcsb.org/> que va a permitir obtener información de las mismas, con la colocación del código correspondiente en la barra de búsqueda.

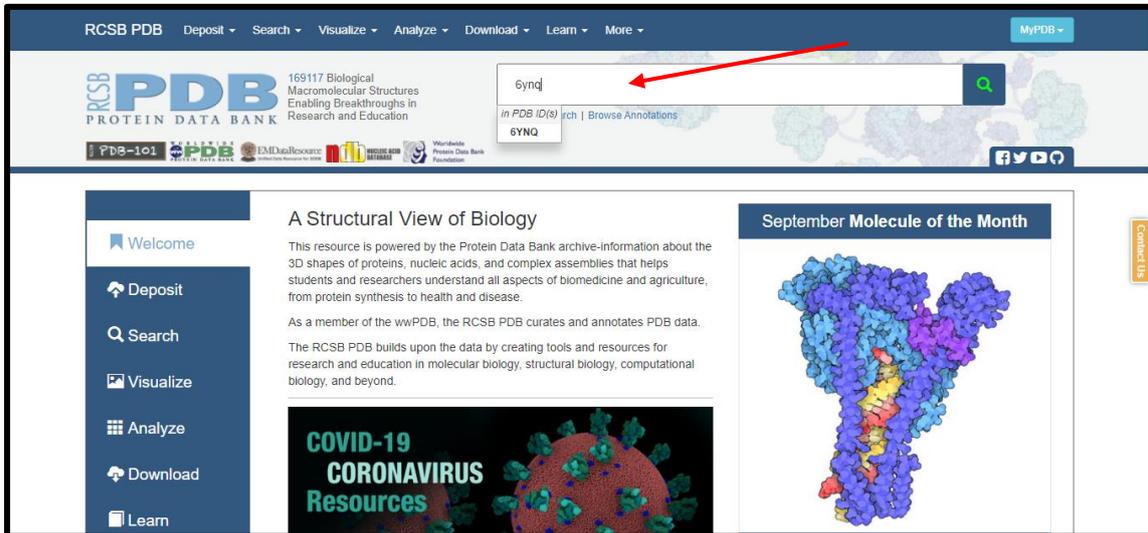


Figura 12. Entorno de RCSB PDB

- De los ligandos unidos a la proteína del virus, se seleccionó el principal (**Figura 13**).

ID	Chains	Name / Formula / InChI Key	2D Diagram	3D Interactions
P6N Query on P6N Download CCD File	A	(2-(S))-2-methyl-3,4-dihydro-2-(H)-naphthalen-1-one C ₁₁ H ₁₂ O GANIBVZSZGNMNB-QIMMGPBSA-N		Ligand Interaction
DMS Query on DMS Download CCD File	A	DIMETHYL SULFOXIDE C ₂ H ₆ O S IAZDPXIOMUYVGZ-UHFFFAOYSA-N		Ligand Interaction
CL Query on CL Download CCD File	A	CHLORIDE ION Cl VEXZGXHIMUGYJMC-UHFFFAOYSA-M		Ligand Interaction

Figura 13. Identificación del posible inhibidor de interés

5. Una vez desplegada la información correspondiente del inhibidor, se extrajo su nombre completo, estructura química en 2D y estructura isomérica smiles que fueron esenciales para la construcción de la librería y estudios posteriores (**Figura 14**).

The screenshot shows the RCSB PDB interface for ligand P6N. The top navigation bar includes 'RCSB PDB', 'Deposit', 'Search', 'Visualize', 'Analyze', 'Download', 'Learn', and 'More'. The main content area displays the 2D chemical structure on the left and a 3D ball-and-stick model on the right. To the right of the 3D model, the ligand name 'P6N' is shown, followed by its full name: '(2~{S})-2-methyl-3,4-dihydro-2~{H}-naphthalen-1-one'. Below this, there are sections for 'Find entries where:' and 'Find related ligands:'. At the bottom, there are two tables: 'Chemical Component Summary' and 'Chemical Details'. The 'Chemical Component Summary' table includes fields for Name, Identifiers, Formula, Molecular Weight, Type, Isomeric SMILES, and InChI. The 'Chemical Details' table includes Formal Charge, Atom Count, Chiral Atom Count, Bond Count, and Aromatic Bond Count. Red arrows in the original image point to the 2D structure, the 3D model, the full name, and the SMILES string in the 'Isomeric SMILES' field.

Chemical Component Summary	
Name	(2~{S})-2-methyl-3,4-dihydro-2~{H}-naphthalen-1-one
Identifiers	(2~{S})-2-methyl-3,4-dihydro-2~{H}-naphthalen-1-one
Formula	C ₁₁ H ₁₂ O
Molecular Weight	160.21
Type	NON-POLYMER
Isomeric SMILES	C[C@H]1CCc2ccccc2C1=O
InChI	InChI=1S/C11H12O/c1-8-6-7-9-4-2-3-5-10(9)/11(8)/12/h2-5,8H,6-7H2,1H3/t8-m/0/s1
	GANIBVZSZGNMNB-QMMMGPBSA-N

Chemical Details	
Formal Charge	0
Atom Count	24
Chiral Atom Count	1
Bond Count	25
Aromatic Bond Count	6

Figura 14. Información desplegada correspondiente al ligando seleccionado

1.8.1.1 Alineamiento total de estructuras

1. Se ingresó al siguiente enlace: <http://tcoffee.org.cat/apps/tcoffee/do:regular>
2. Se ingresaron las secuencias previamente obtenidas de Uniprot o PDB en formato FASTA en la sección “Entrada de secuencias” (**Figura 15**).

The screenshot shows the T-COFFEE website interface. The top navigation bar includes 'T-COFFEE', 'Casa', 'Historia', 'Tutorial', 'Referencias', 'Contactos', 'Proyectos', and 'Descargar'. The main content area displays the 'Entrada de secuencias' section, which includes a text input field for entering sequences in FASTA format. A red arrow in the original image points to the input field, which contains the following FASTA sequence data:

```
>SARS-CoV-2
SGFRKMAFPSPGKVEGCMVQVTCGTTTLNGLWLDVVYCPRHVICTSEDMLNPIYEDLLIRKSIHIFLVQAGNIVQLRVIGHSMQNCLLRLKVDTSNP
KTPKYKFRVIRIQGQTFVSLACVNGSPSCYQCAMPNFTIKGSFLNGSCGSGVFNIDYDCVFCYMHHELPDGVHAGTDLEGNFYGPVDRQTAQ
AAGTDTTITVNLAWLYAAVINGDRVFNHRFTTLLNDFNLVAMKYNYEPLTQDHVDLGLPLSAQTGIAVLDMCASLKEILLQNGMNGRTILGSALLE
DEFTPFDFVVRQCSGVTFQ

>SARS-CoV
SGFRKMAFPSPGKVEGCMVQVTCGTTTLNGLWLDVVYCPRHVICTAEDMLNPIYEDLLIRKSIH5FLVQAGNIVQLRVIGHSMQNCLLRLKVDTSNP
```

Figura 15. Entorno de la herramienta TCOFFEE para alineamientos de secuencias

3. En la nueva ventana desplegada, en la sección de “Archivos de resultados” se seleccionó el archivo fasta aln para descargar. (Figura 16).

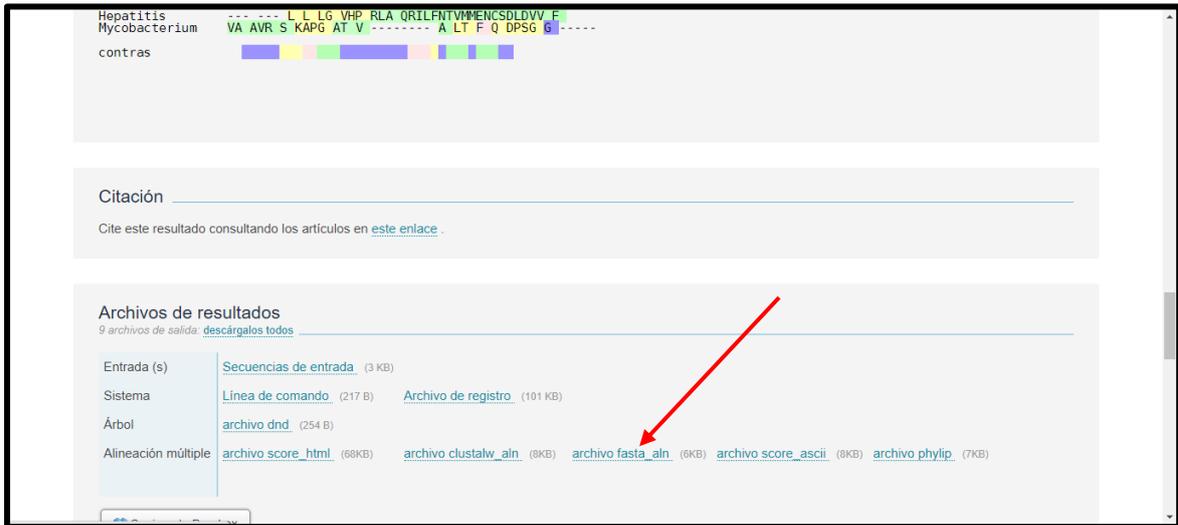


Figura 16. Formatos de descarga del alineamiento realizado.

4. Se ingresó al dominio: http://www.ch.embnet.org/software/BOX_form.html

5. Se configuró el formato de salida a “RTF new”, el formato de secuencia de entrada a “otro”, se pegó las secuencias descargadas en el formato previamente descrito y se prosiguió a “Ejecutar BOXSHADE”. (Figura 17).

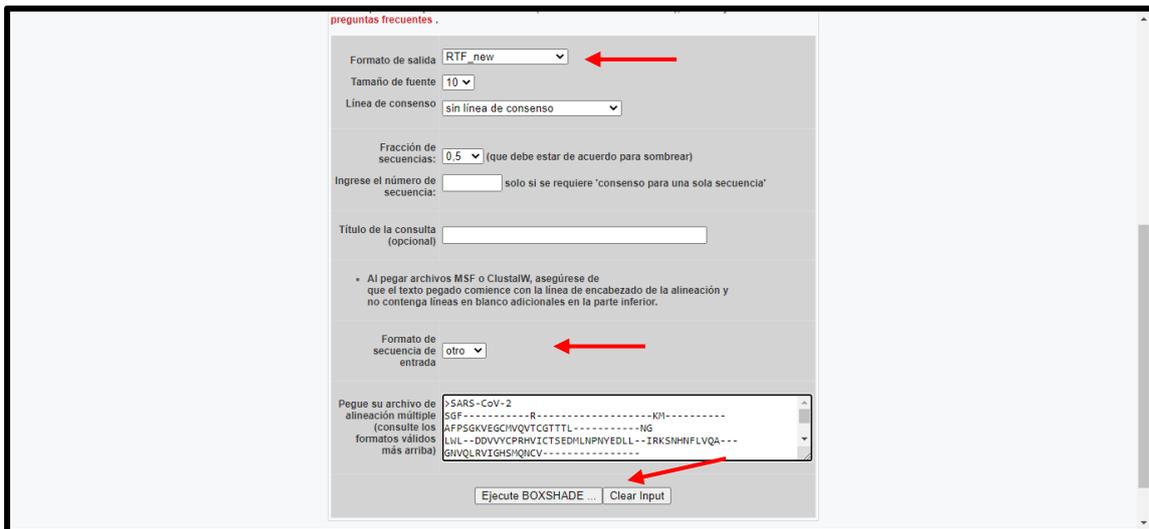


Figura 17. Entorno de la herramienta destinada a dar formato al alineamiento

6. Se descargar el alineamiento en el apartado de “Salida número 1”

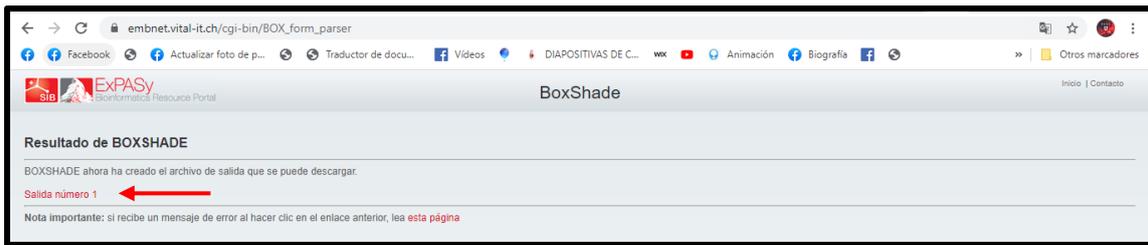


Figura 18. Botón para la descarga del alineamiento.

1.8.2 DETERMINACIÓN DE LOS SITIOS DROGABLES DE LA PROTEÍNA

1. Se descargó el formato PDB la estructura de la proteína con su inhibidor.

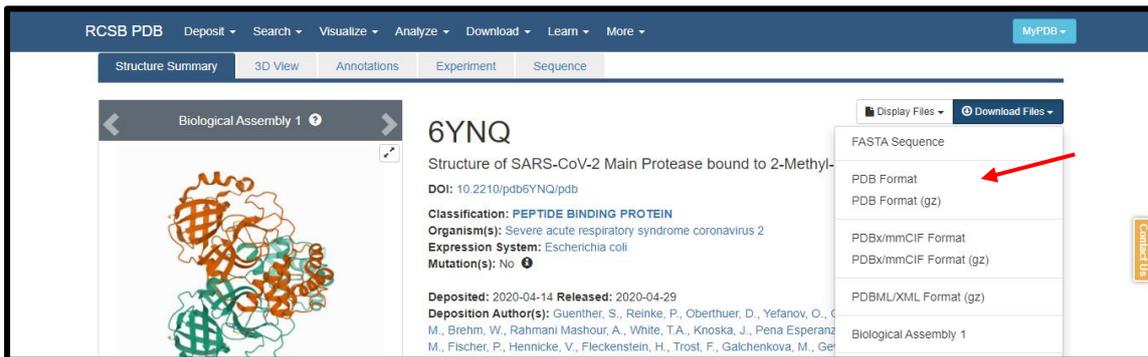


Figura 19. Menú de descarga de estructuras en formato pdb

2. Se visualizó la estructura previamente descargada en software bioinformático conocido como Wincoot (Para el proyecto se utilizó la versión 0.8.7) (**Figura 20**), para lo cual se usó en la opción “Open Coordinates” y seguidamente se seleccionó la estructura deseada en formato PDB

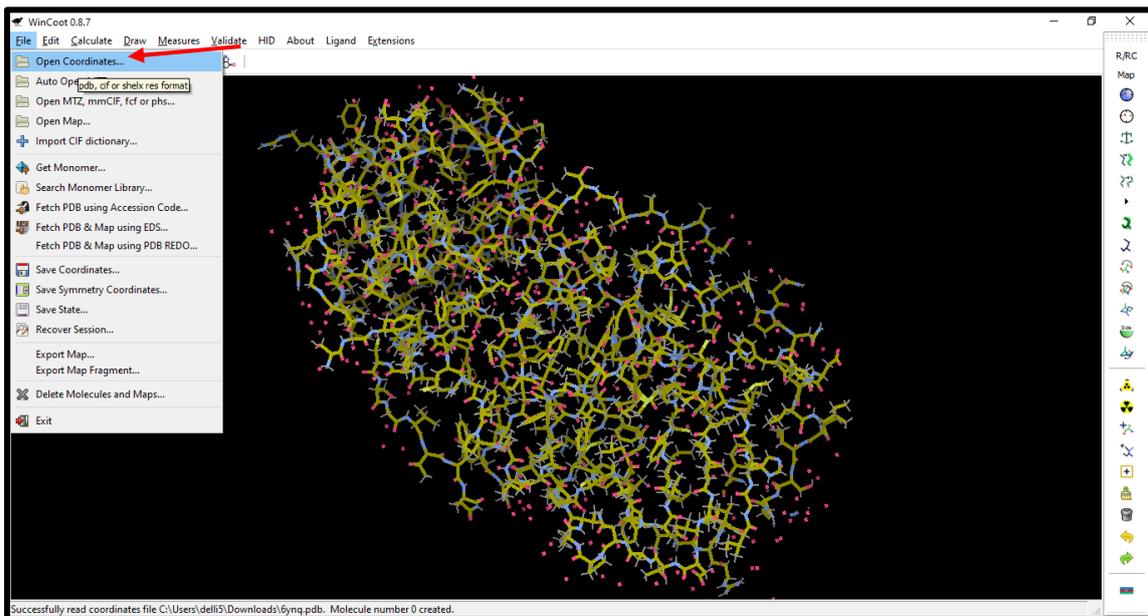


Figura 20. Entorno de software Pymol versión 0.8.7

3. En la barra de tareas superior se seleccionó la opción “Go to next Ligand” (Figura 21), para que el programa nos lleve a la sección de la proteína donde el ligando se ha unido, la estructura de la proteína principal se denota como “A” y la cadena del ligando como “B”

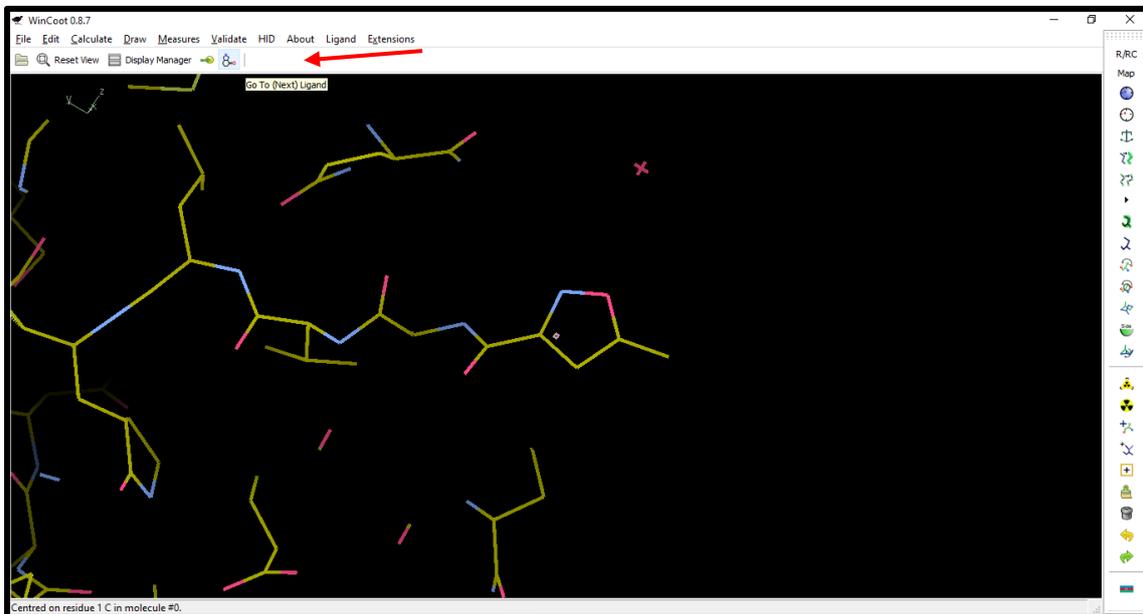


Figura 21. Visualización de la estructura del ligando

4. Se usó la opción “Measures” y seguidamente “Environment Distances” (Figura 22) con el fin de visualizar la distancia y el tipo de enlace que presenta el ligando hacia la proteína.

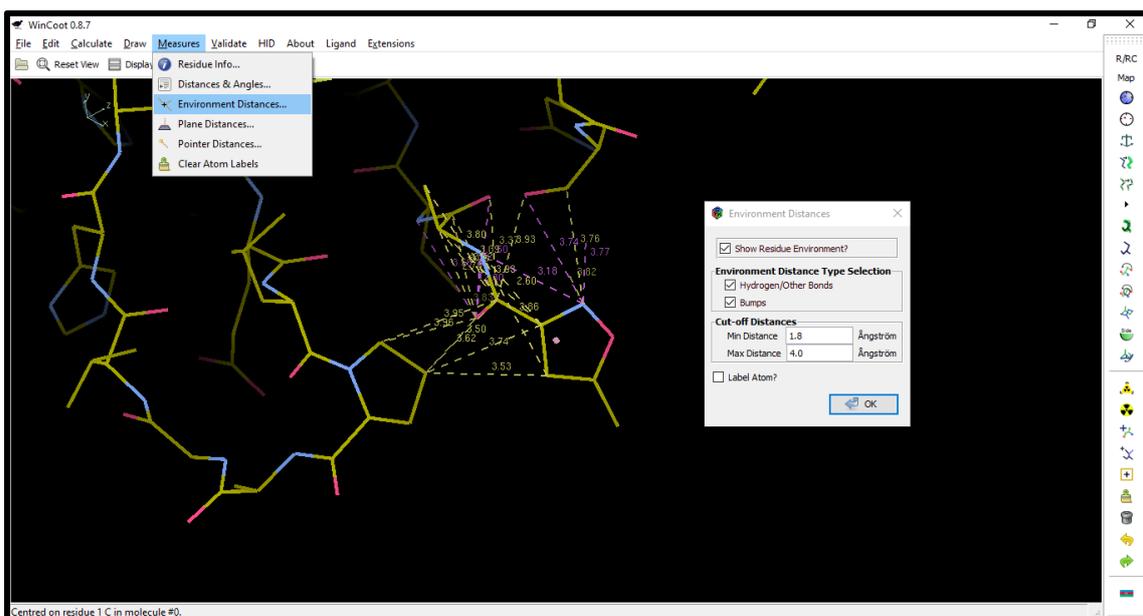


Figura 22. Funciones del programa Wincoot usadas en la investigación

2.3.3.1 Obtención de imágenes de inhibidores dentro del sitio activo de la proteasa

Mpro

1. Se dispuso del programa de análisis molecular Pymol
2. En el entorno de Pymol ase abrió el archivo PDB obtenido previamente

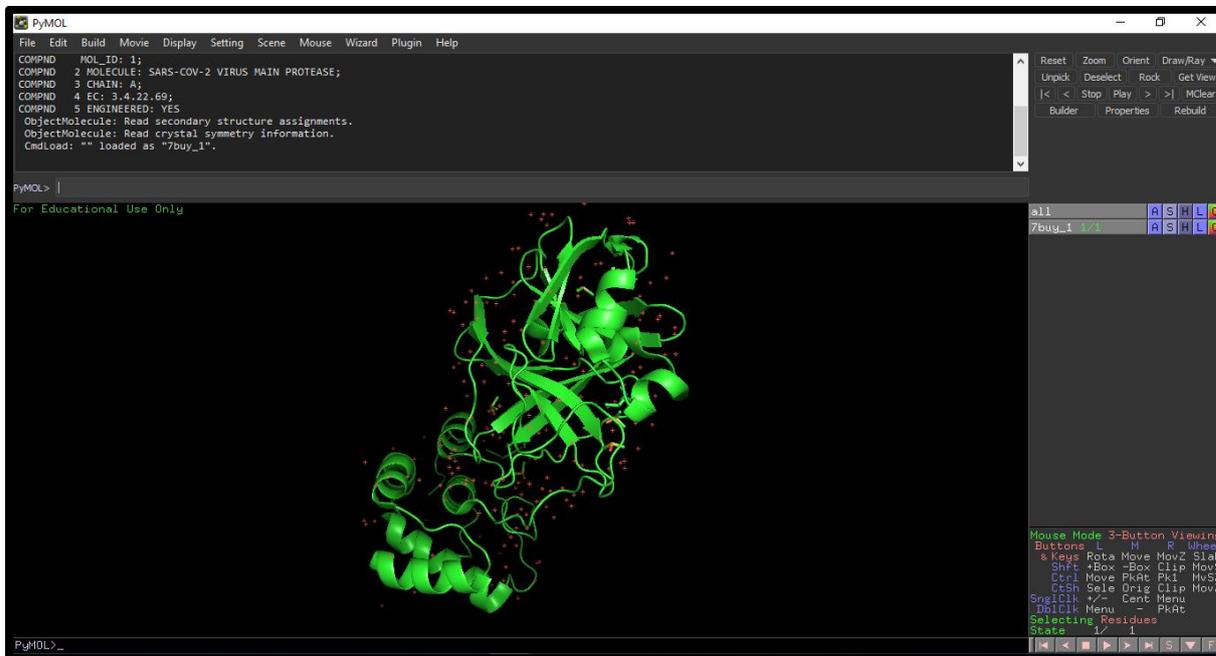


Figura 24. Entorno del software de visualización molecular Pymol

3. En el menú dispuesto a la derecha, se eliminó las aguas remanentes en “Hide” y se convirtió la estructura a modo superficie en “Show” (**Figura 25**).

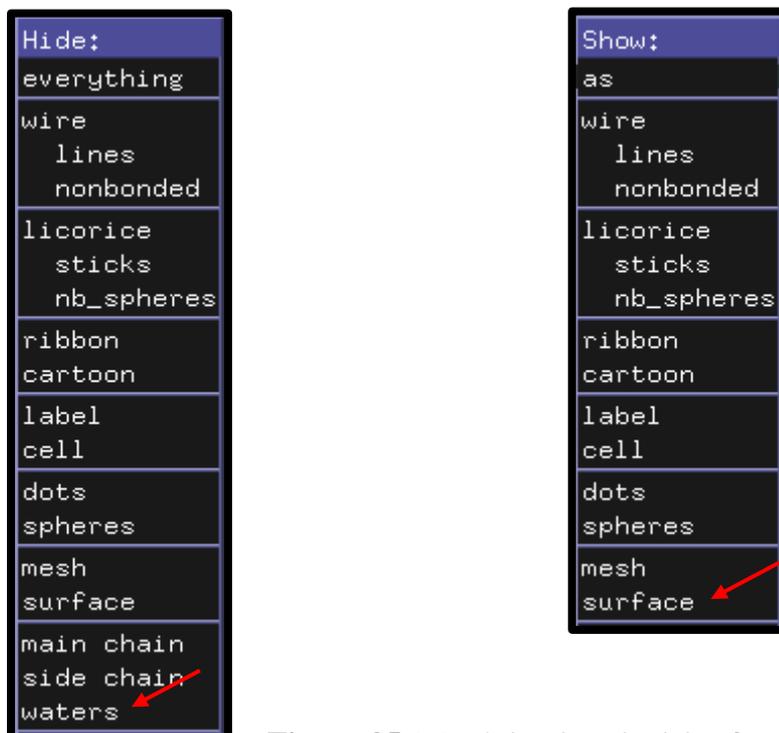


Figura 25. Menú desplegado del software Pymol

- Se seleccionó en la secuencia, la estructura correspondiente al inhibidor y fue visualizado a manera de palillos y esferas en el menú “Show” correspondiente al ligando seleccionado.

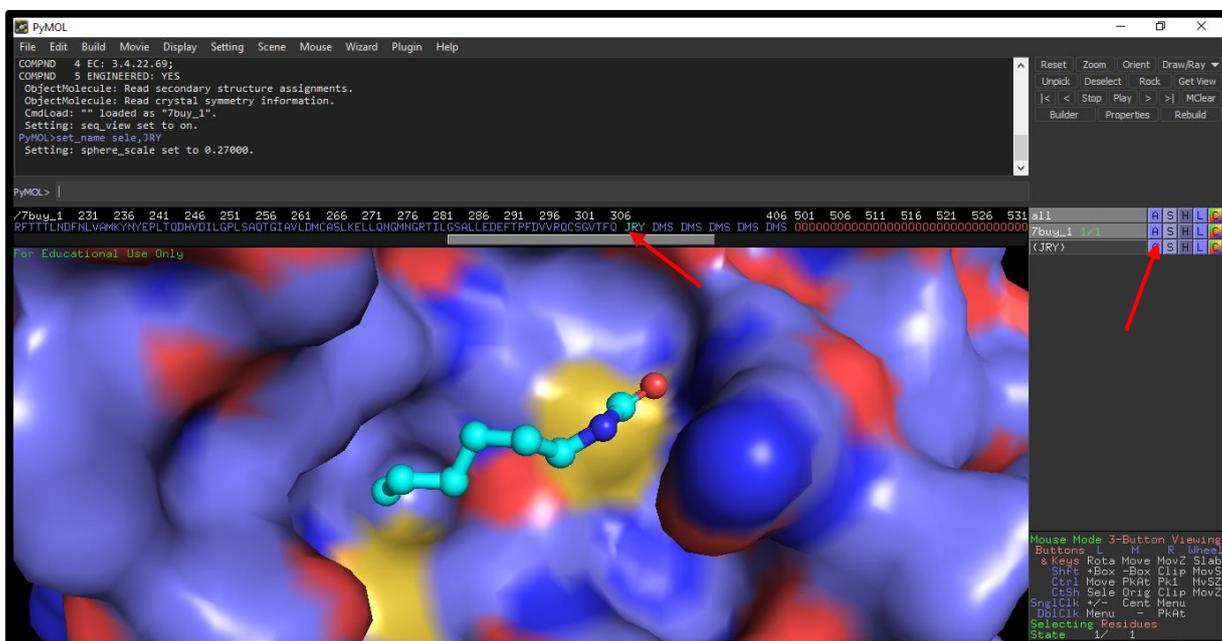


Figura 26. Visualización de la estructura del ligando y la proteína en 3D

Nota: En algunos casos los átomos de los ligandos formaron parte de la estructura en superficie de la proteína y en ese caso se requirió modificar el archivo en formato PDB valiéndose de la herramienta block de notas, donde se eliminó el inhibidor seleccionado en (**Figura 27**) y la proteína vacía se acopló al ligando obtenido de la estructura inicial

Archivo	Edición	Formato	Ver	Ayuda
ATOM	2365	CD	GLN A 306	-45.493 -1.109 30.383 1.00 79.05 C
ATOM	2366	OE1	GLN A 306	-45.900 0.004 30.065 1.00 83.70 O
ATOM	2367	NE2	GLN A 306	-46.086 -2.221 29.992 1.00 58.61 N
TER	2368		GLN A 306	
HETATM	2369	C07	JRY A 401	-11.876 17.360 65.601 0.84 34.61 C
HETATM	2370	C09	JRY A 401	-11.911 15.645 67.452 0.84 41.74 C
HETATM	2371	C10	JRY A 401	-11.004 15.546 68.685 0.84 45.76 C
HETATM	2372	C11	JRY A 401	-11.699 15.770 70.038 0.84 47.49 C
HETATM	2373	C12	JRY A 401	-12.665 14.669 70.456 0.84 46.09 C
HETATM	2374	C13	JRY A 401	-12.996 14.683 71.948 0.84 48.54 C
HETATM	2375	C14	JRY A 401	-14.494 14.792 72.230 0.84 43.21 C
HETATM	2376	N08	JRY A 401	-11.204 16.308 66.355 0.84 40.21 N
HETATM	2377	O15	JRY A 401	-11.338 17.944 64.721 0.84 32.05 O
HETATM	2378	S	DMS A 402	-12.502 12.986 63.022 1.00 45.05 S
HETATM	2379	O	DMS A 402	-13.779 12.314 63.355 1.00 52.52 O
HETATM	2380	C1	DMS A 402	-12.001 13.803 64.499 1.00 42.19 C
HETATM	2381	C2	DMS A 402	-11.271 11.721 63.013 1.00 32.15 C
HETATM	2382	S	DMS A 403	-16.809 33.064 54.309 1.00 82.45 S
HETATM	2383	O	DMS A 403	-17.651 32.808 53.108 1.00 63.18 O
HETATM	2384	C1	DMS A 403	-15.342 31.998 54.227 1.00 58.63 C
HETATM	2385	C2	DMS A 403	-17.683 32.412 55.765 1.00 44.20 C
HETATM	2386	S	DMS A 404	-26.023 3.704 38.278 1.00105.76 S

Figura 27. Codificación correspondiente al ligando, dentro de la estructura total

5. Finalmente se escribió escribir en la barra de comandos “ray 5000” para obtener una imagen de alta resolución en formato PNG

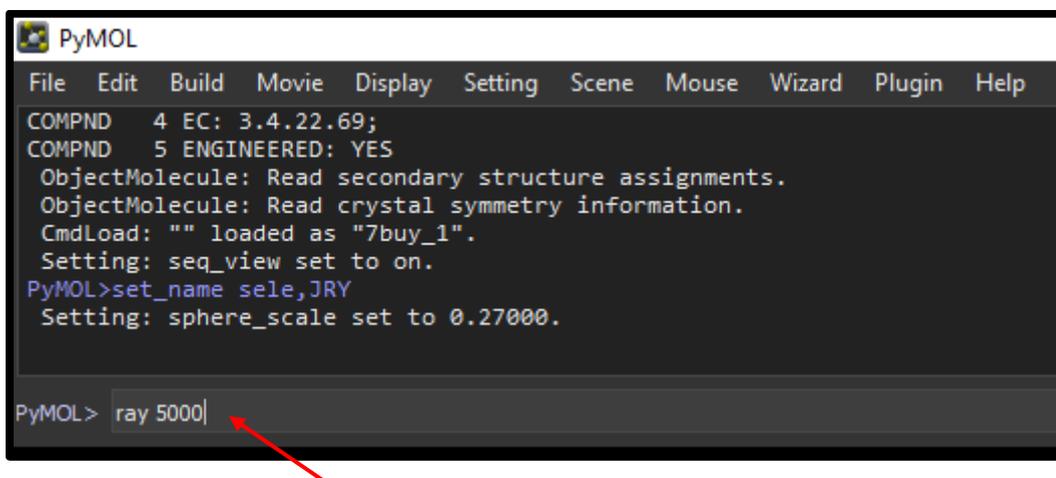


Figura 28. Función “ray” de Pymol para obtener una imagen de la estructura

2.3.3.2 Optimización de Ligandos

1. Se dispuso del programa de análisis molecular Mercury
2. En el entorno de Mercury se cargó la proteína deseada

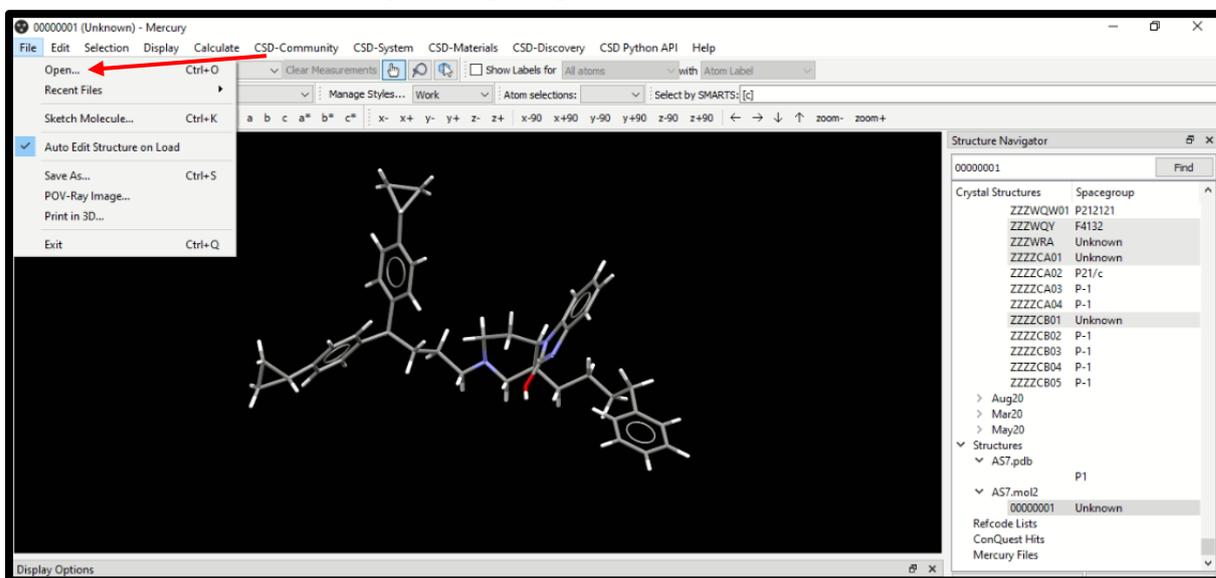


Figura 29. Entorno del programa Mercury

3. En el menú Edit. se seleccionó la opción “Auto-Edit Structure” que desplegó un menú por el que el programa añadirá átomos faltantes y estandarizará los ligandos presentes en la molécula. (Figura 30).

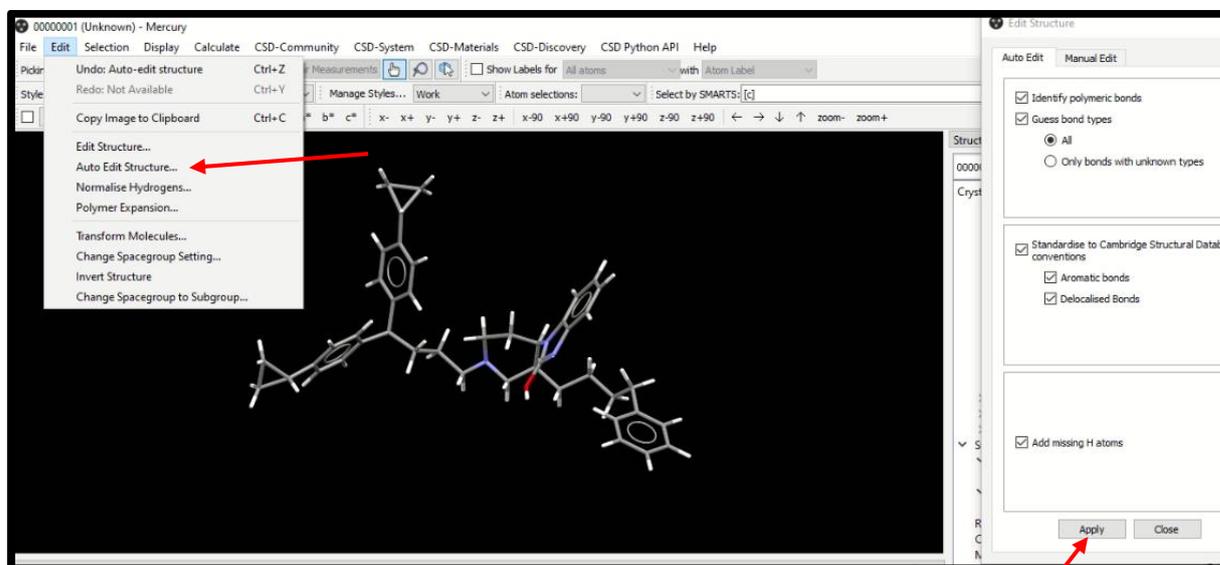


Figura 30. Función usada en la optimización de ligandos

4. El archivo fue guardado en formato “MOL.2”

1.8.4 DOCKING O ACOPLAMIENTO DE LA MOLÉCULA INHIBIDORA A LA PROTEÍNA

1. Se dispuso del software de acoplamiento molecular GOLD
2. Se definió la proteína objetivo para el acoplamiento

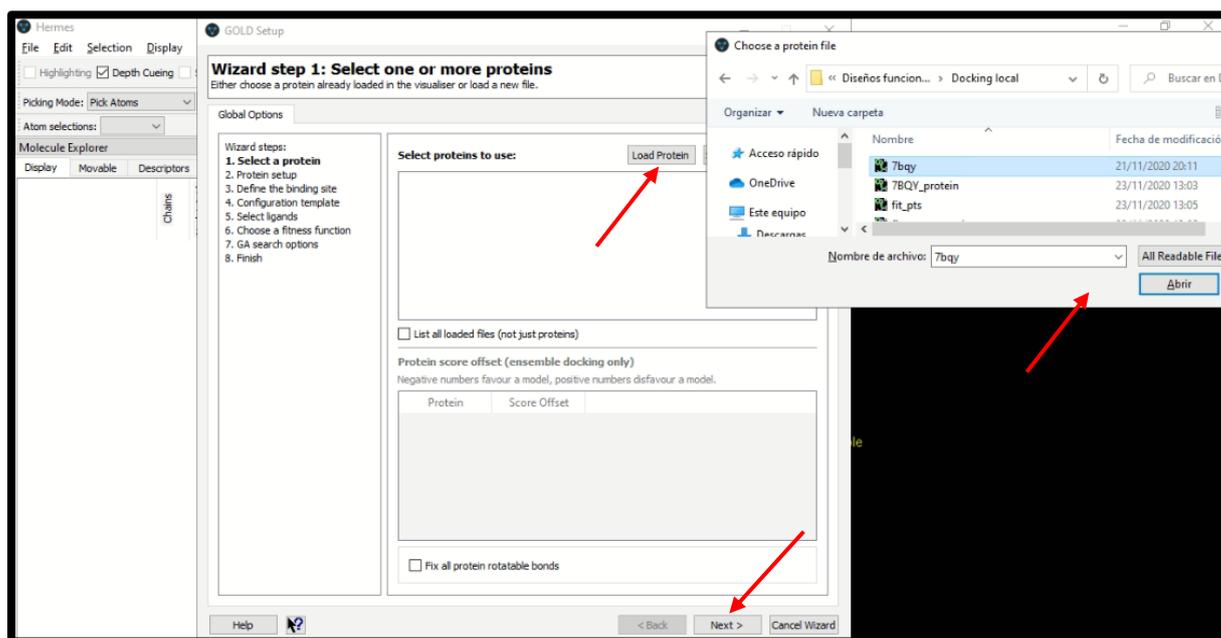


Figura 31. Entorno de GOLD Protein Docking Ligand

3. Se configuró la proteína de tal manera que se añadan los hidrógenos faltantes, se eliminen las moléculas de agua sobrantes y se extraiga el ligando de referencia, funciones señaladas en (Figura 32)

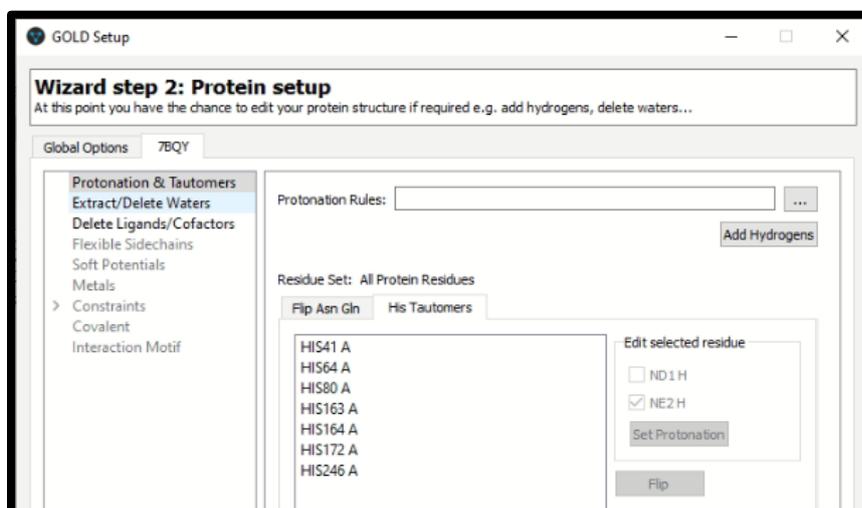


Figura 32. Configuración de la proteína objetivo

4. Se definió el sitio de unión de los inhibidores a la proteína, para lo cual se tomó como referencia el ligando que originalmente se encontraba unido a la enzima.

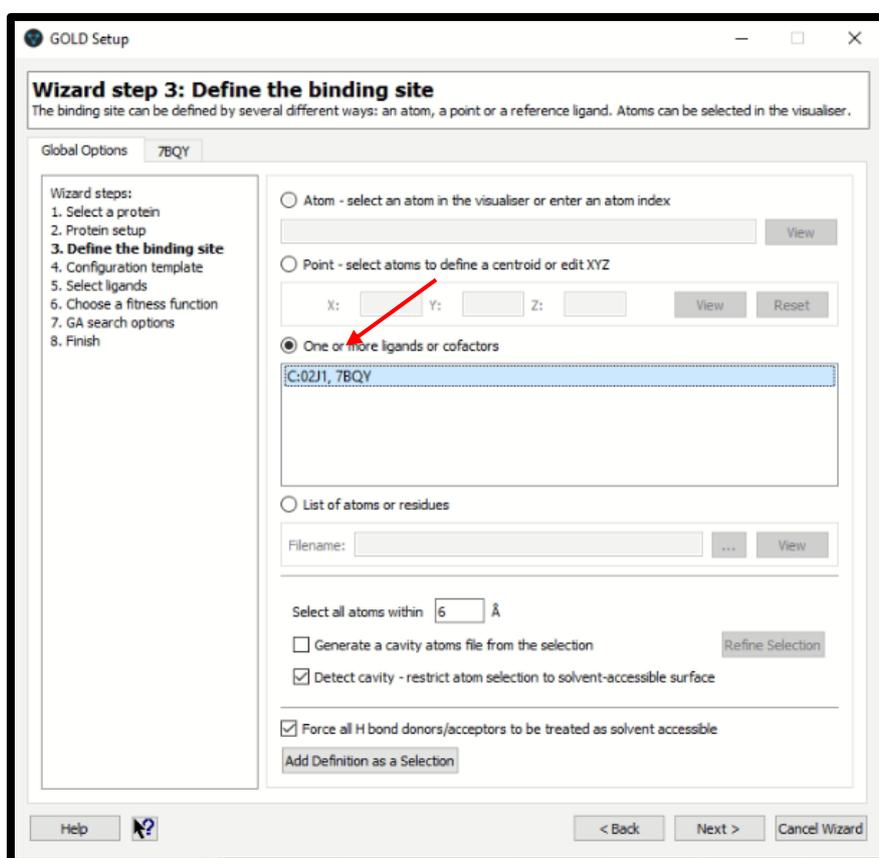


Figura 33. Configuración del sitio de acoplamiento para ligandos.

5. Se seleccionó el programa de acuerdo al tipo de enzima estudiado, en este caso una serine proteasa.

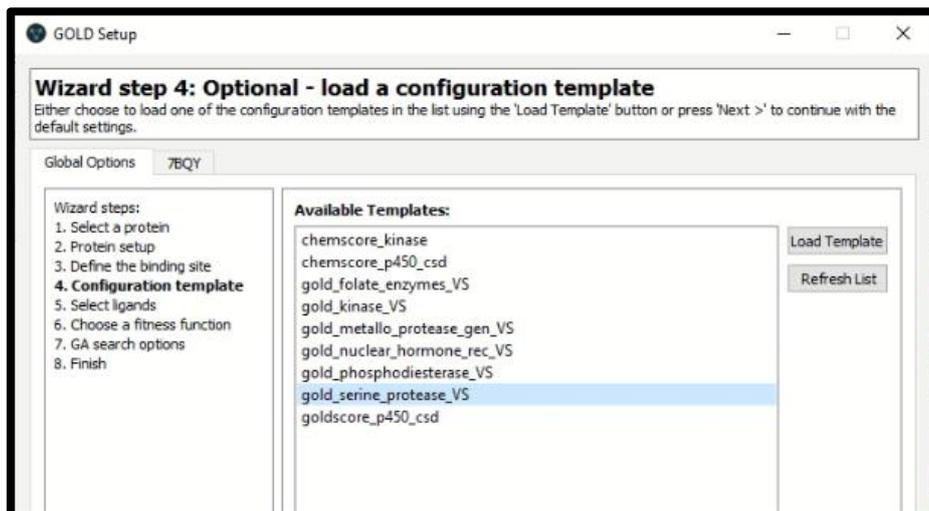


Figura 34. Definición de programa para acoplamiento molecular.

6. Se añadieron los ligandos obtenidos de PDBe Fold, Covid Box y aquellos que fueron diseñados.

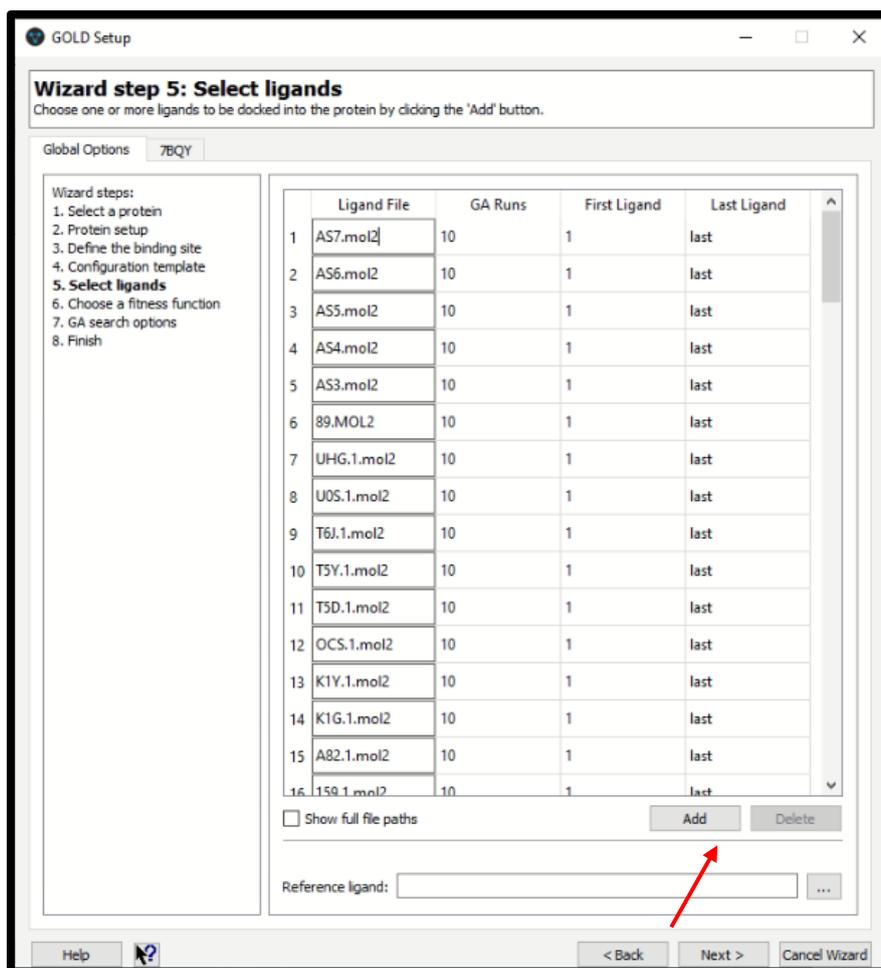


Figura 35. Adición de los ligandos optimizados para el acoplamiento molecular

7. Se escogió el programa CHEMPL y se modificó las opciones de terminación rápida a un número entero que sería 1 para lograr mayor afinidad en el acoplamiento.

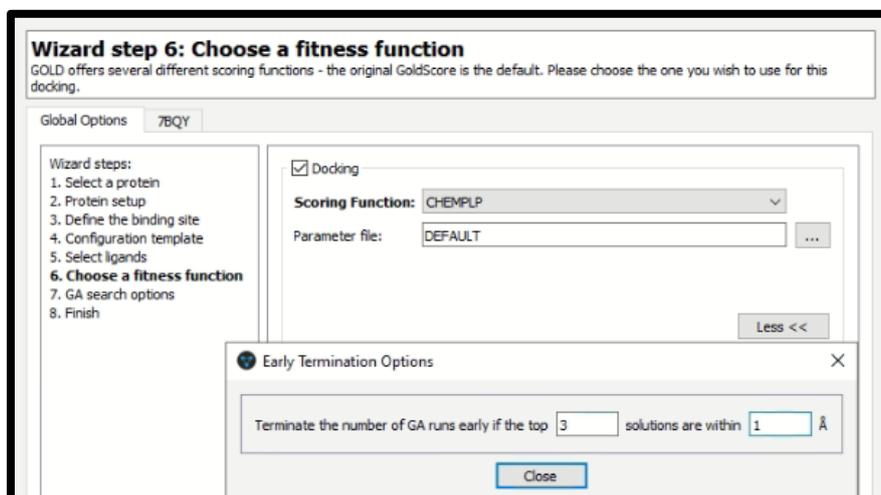


Figura 36. Elección de programa para el estudio

8. Se selecciono velocidad de análisis lenta para asegurar la exactitud de los resultados posteriores

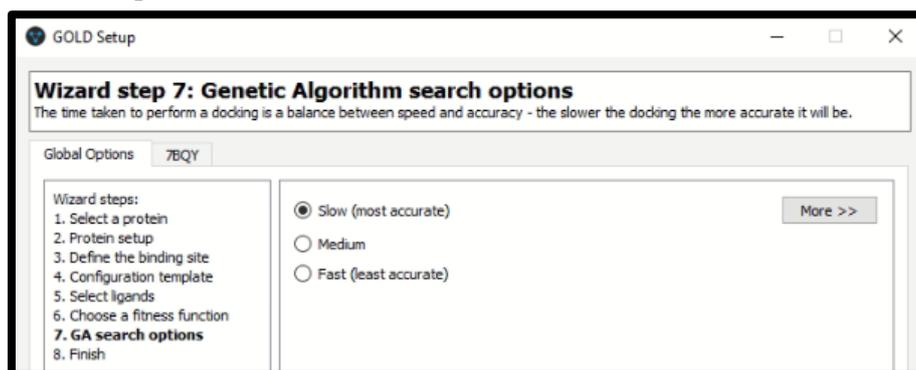


Figura 37. Definición de parámetros para acoplamiento molecular

9. Antes de finalizar, en la sección “Advanced” (Figura 38) se modificó aquellos residuos identificados como flexibles, incluyendo sus Rotámeros en el análisis, seguidamente se mandó a correr el programa mediante el botón “Run Gold”

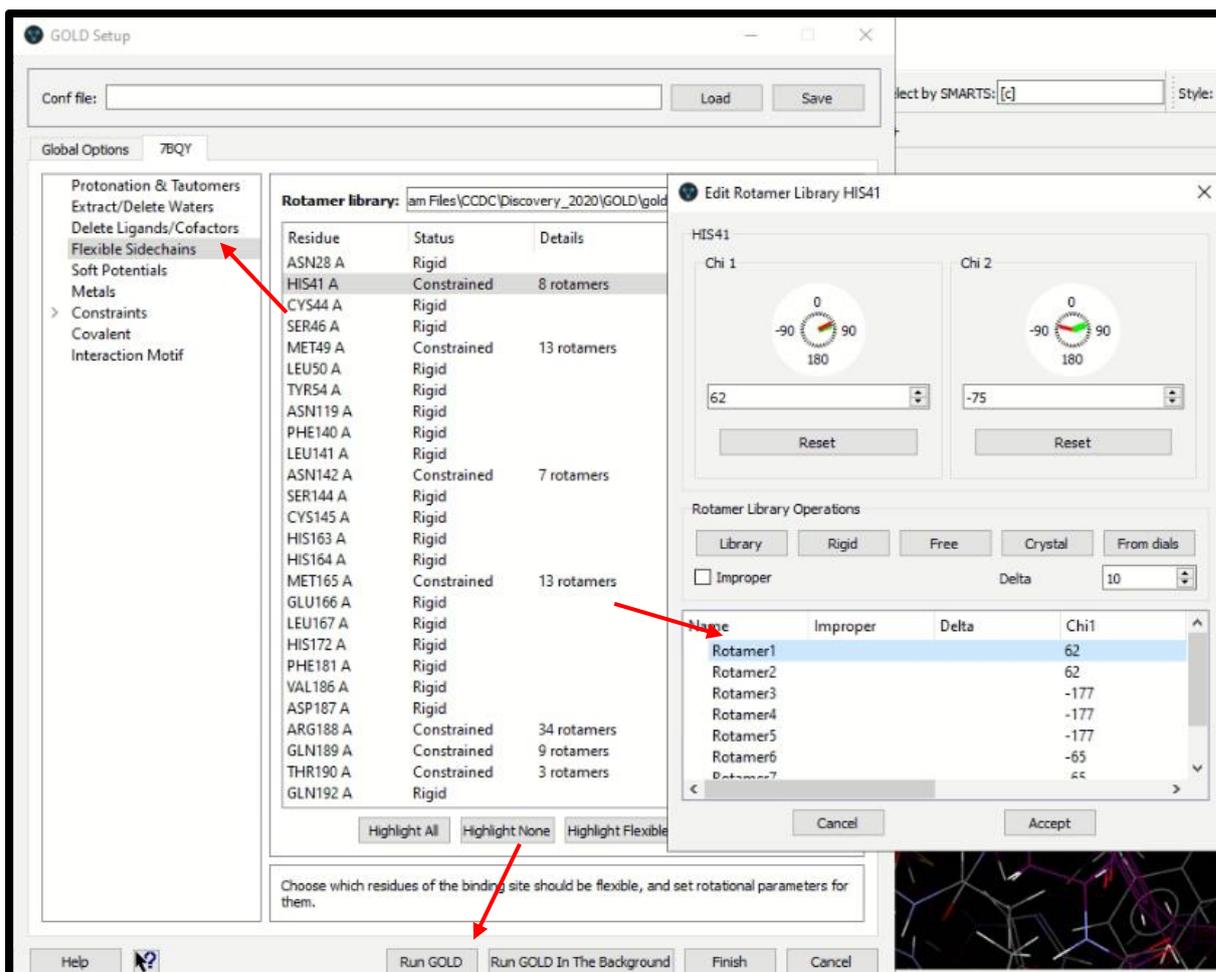


Figura 38. Modificación de residuos flexibles de la proteína

10. Se esperó hasta que el programa pruebe los ligandos uno a uno

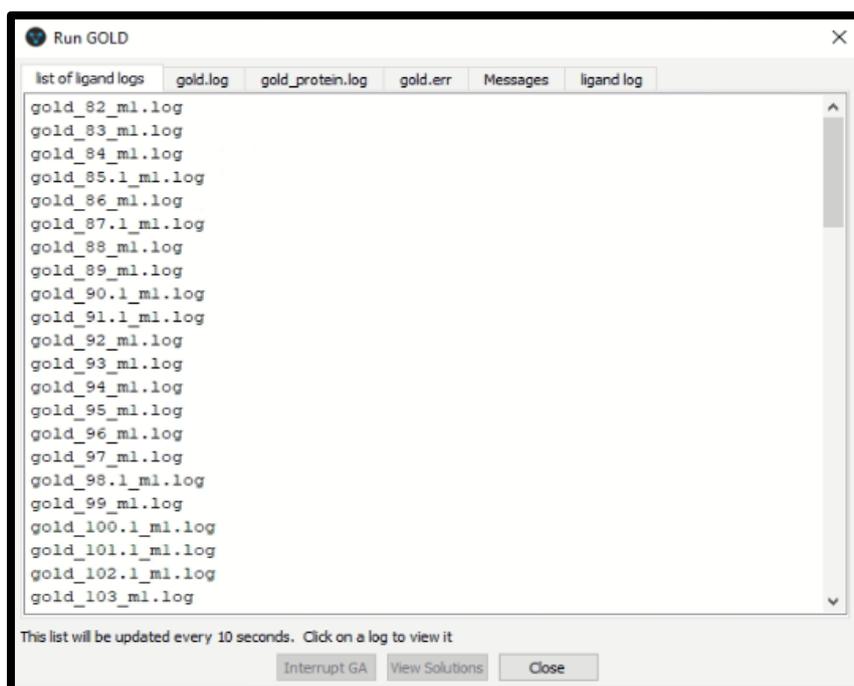


Figura 39. Proceso de análisis y prueba de los ligandos por parte del software

11. Se ordenó las soluciones en función de su PLP Score y se extrajo los 3 primeros resultados de los ligandos de covid box y todos los correspondientes a los ligandos diseñados.

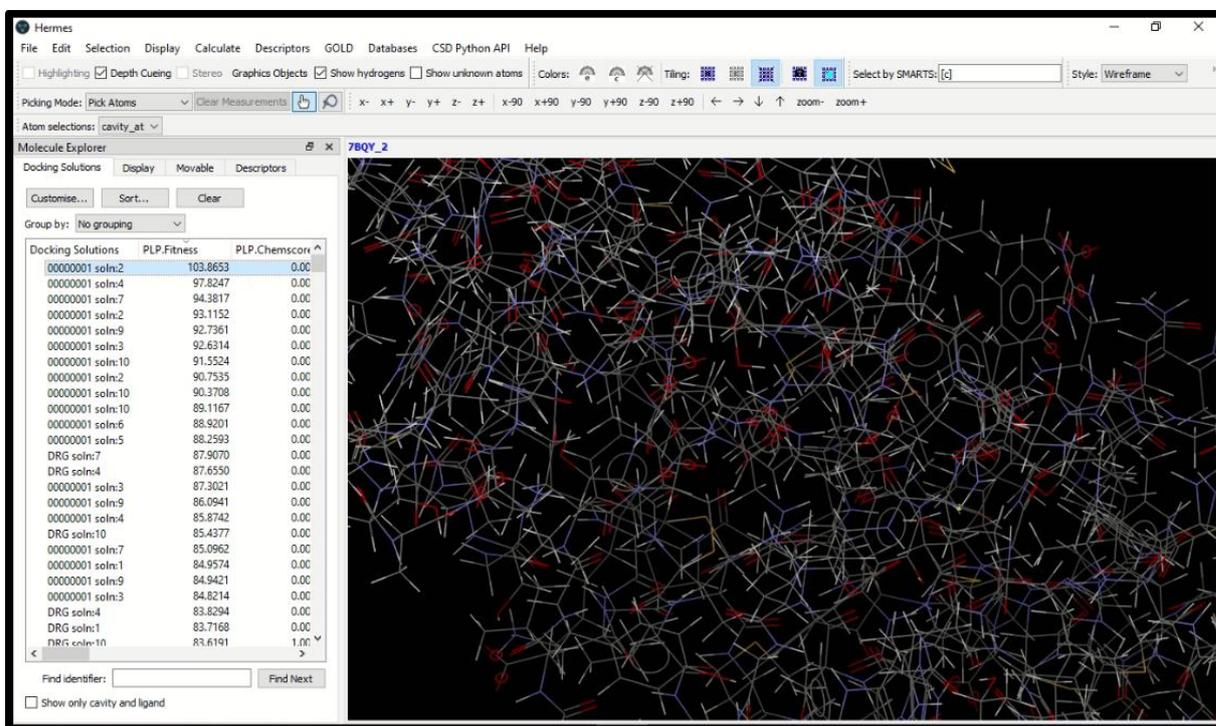


Figura 40. Obtención de resultados del acoplamiento molecular

1.8.5 DISEÑO RACIONAL DE INHIBIDORES

1. Se tomó como punto de partida los 3 inhibidores con mayor score obtenidos del cribado virtual y luego de ser analizados mediante el estudio de acoplamiento molecular.

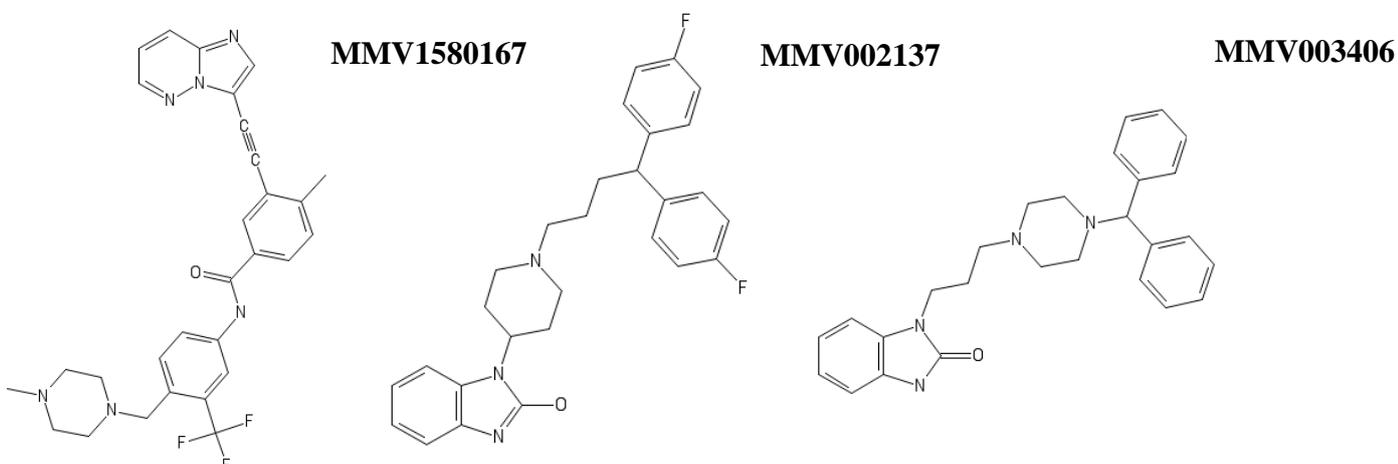


Figura 41. Estructura 2D de los ligandos con mayor actividad inhibitoria de Mpro

2. Se identificó en las imágenes de las mejores soluciones, los sitios donde existe interacción de inhibidor – proteína como también aquellos donde se podría aumentar grupos funcionales para elevar el score de interacción.

❖ **Ligando MMV1580167**

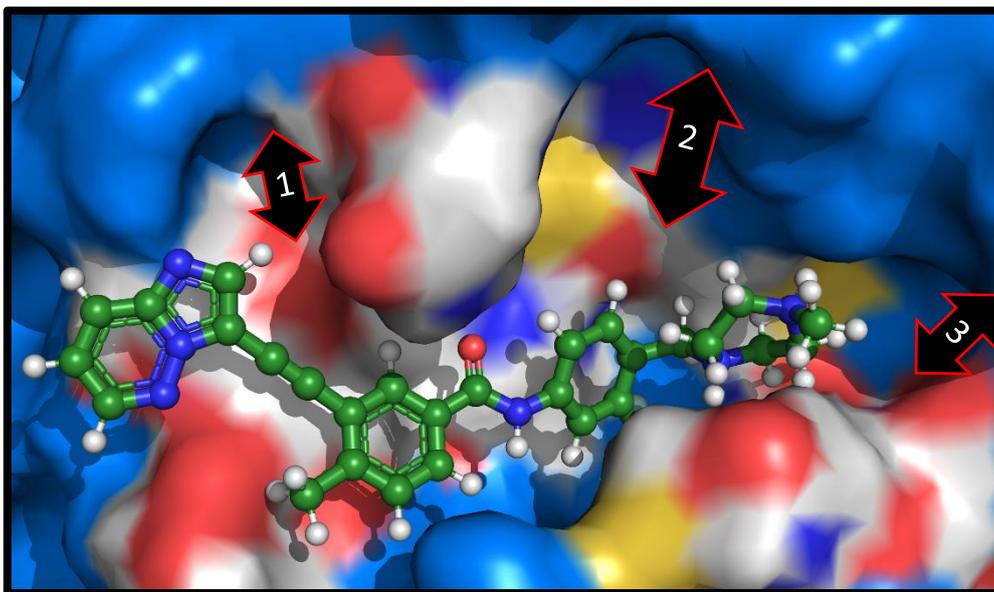


Figura 42. Posibles sitios para la incorporación de grupos funcionales al ligando MMV1580167

3. Una vez identificados los sitios sin interacción dentro del sitio activo de la proteasa, se procedió al diseño de grupos funcionales que puedan llegar esos espacios, tomando en cuenta la distancia entre enlaces Carbono – Carbono que es de 154 Å°

❖ **Diseño del Ligando AS1**

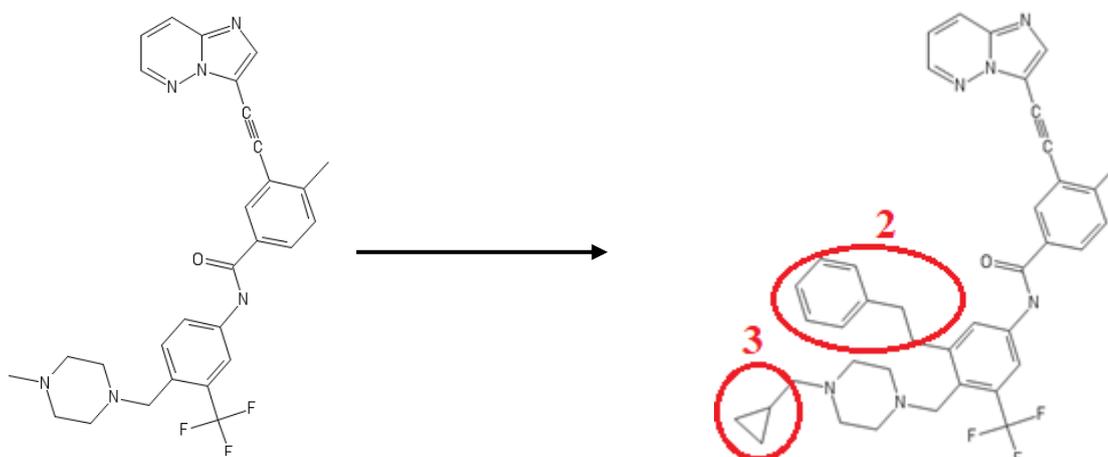


Figura 43. Transición de la molécula MMV1580167 al ligando AS1.

El ligando se modificó añadiendo un grupo ciclopentano en la posición 3 y un grupo fenilo precedido por 3 enlaces carbono-carbono para acoplarse a la posición 2, en este primer ligando la posición 1 no fue tomada en cuenta.

❖ Diseño del Ligando AS2

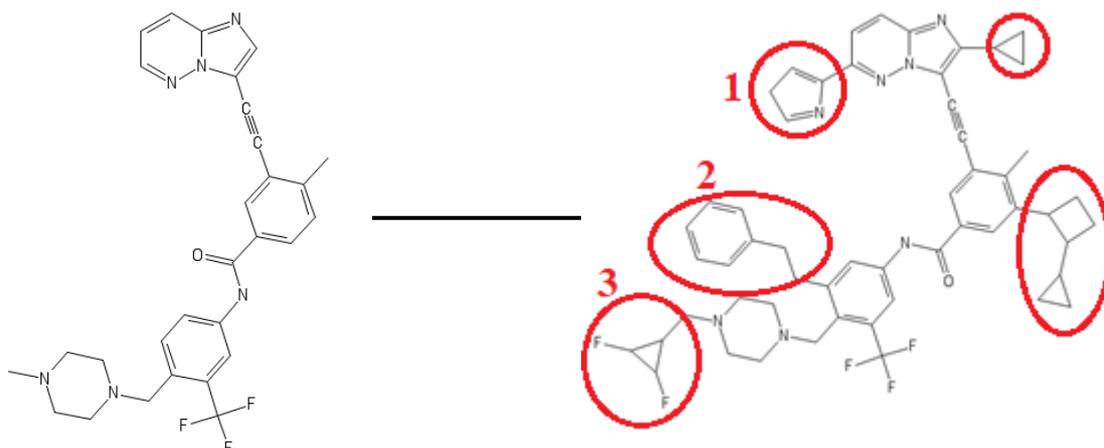


Figura 44. Transición de la molécula MMV1580167 al ligando AS2

En este ligando se incluyeron una gran cantidad de grupos funcionales, un anillo de ciclo pentano para cubrir la posición 1, un anillo de ciclopropano para otorgarle estabilidad, una combinación de un ciclo butano y ciclopropano para mejorar la interacción hacia la diada catalítica, un grupo químico similar al mencionado anteriormente para el sitio 2 y un anillo de ciclopropano terminado con átomos de flúor para el sitio 3.

❖ Ligando MMV002137

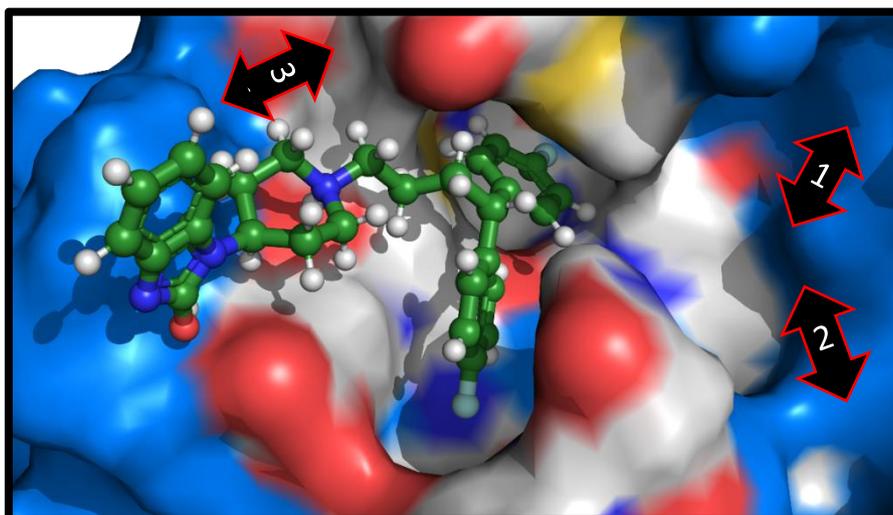


Figura 45. Posibles sitios para la incorporación de grupos funcionales al ligando MMV002137

❖ Diseño del Ligando AS6

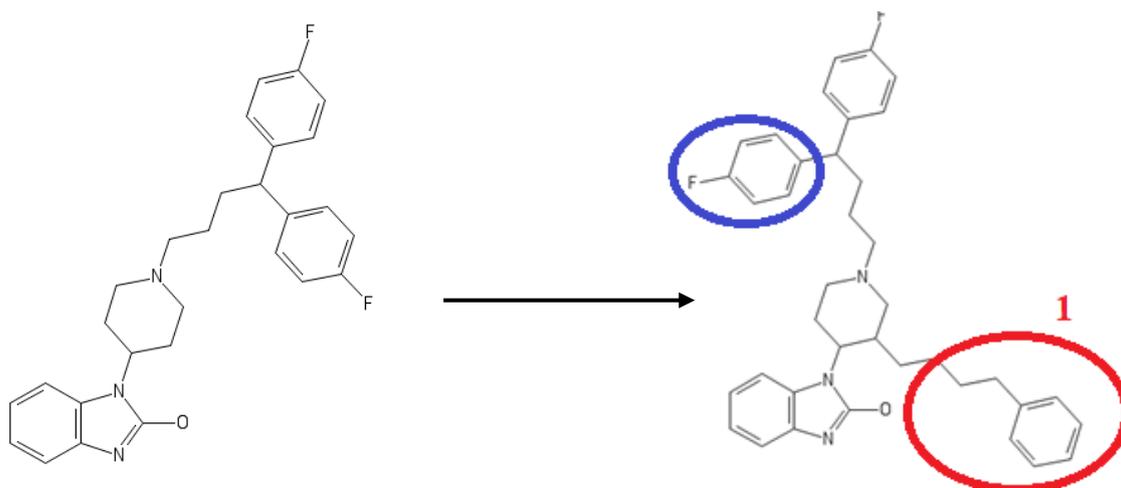


Figura 46. Transición de la molécula MMV002137 al ligando AS6

Para el diseño de este ligando, fue necesario cambiar la conformación espacial de uno de sus grupos funcionales (azul) para evitar la colocación de muchos grupos funcionales en un mismo sitio de la molécula, que generaría en choques estéricos entre los átomos que afectarían a la unión y afinidad del inhibidor, y se colocó un grupo formado por 4 enlaces carbono-carbono y un fenilo.

❖ Diseño del Ligando AS7

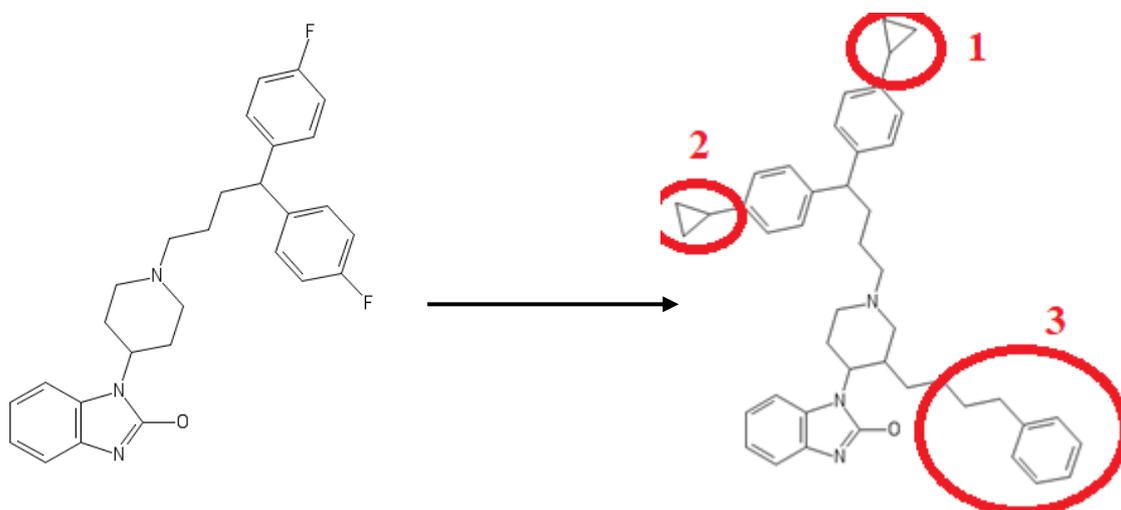


Figura 47. Transición de la molécula MMV002137 al ligando AS7

Para este ligando se evaluó la colocación de anillos de ciclopropano en sustitución de los átomos de Flúor, para acoplarlos a los espacios 1, 2 y a su vez se varió la conformación

espacial del grupo funcional similar al caso anterior, por otra parte, el grupo conformado por enlaces carbono-carbono y grupo fenilo fue evaluado para el acoplamiento al sitio 3.

❖ Ligando MMV003406

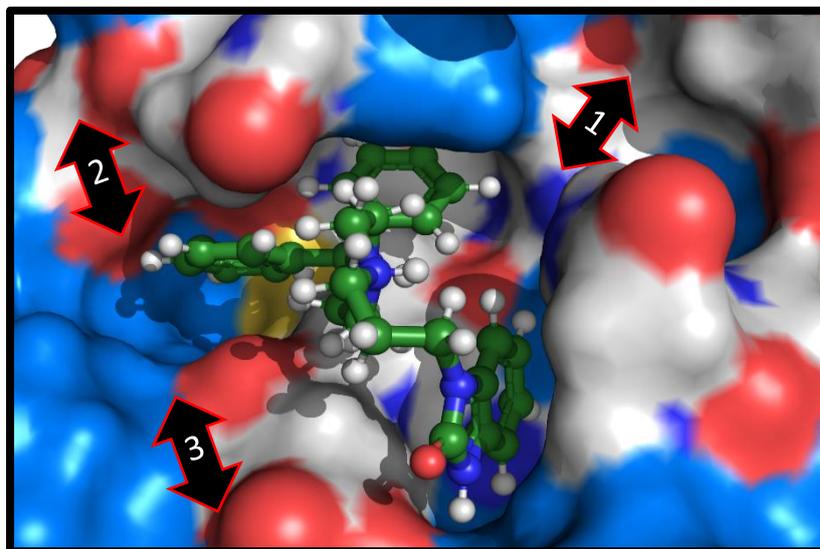


Figura 48. Posibles sitios para la incorporación de grupos funcionales al ligando MMV002137

❖ Diseño del Ligando AS3

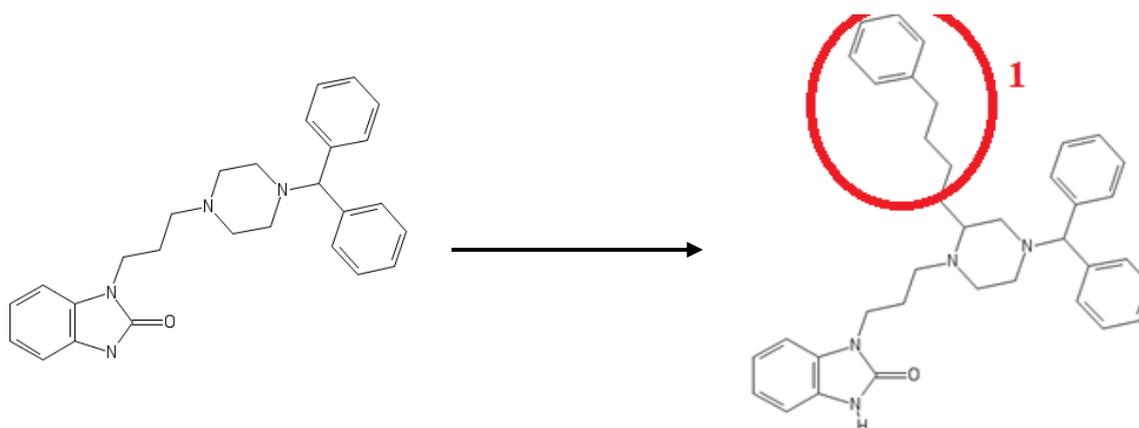


Figura 49. Transición de la molécula MMV003406 al ligando AS3

Para mejorar el acoplamiento de este ligando a la proteína, se optó por añadir un grupo funcional conformado por enlaces carbono-carbono terminados en un fenilo de tal manera que se disponga en la posición 1.

❖ Diseño del Ligando AS4

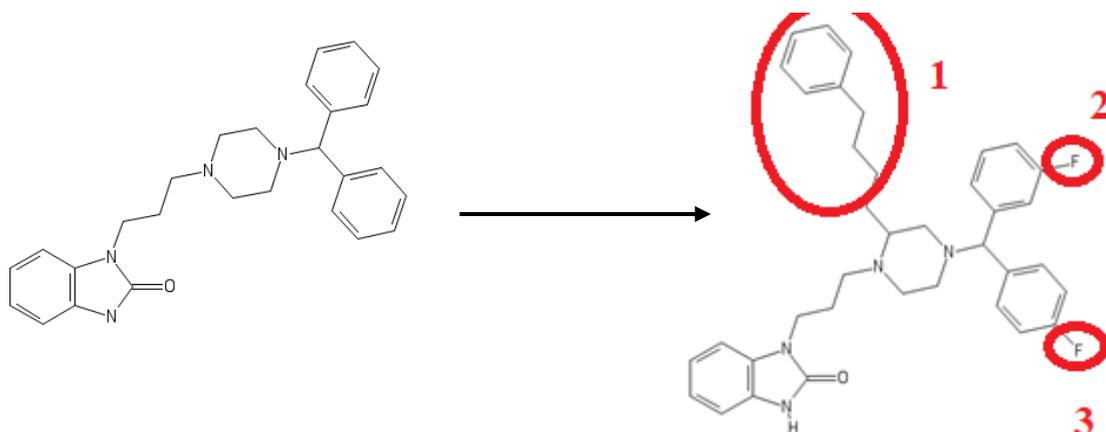


Figura 50. Transición de la molécula MMV003406 al ligando AS4

Este ligando tiene una estructura similar a AS3, pero con la inclusión de átomos de fósforo en los extremos moleculares, buscando que estos orienten a la misma y se acoplen a los sitios 2 y 3 y permitan a la molécula extenderse por todo el sitio activo.

❖ Diseño del Ligando AS5

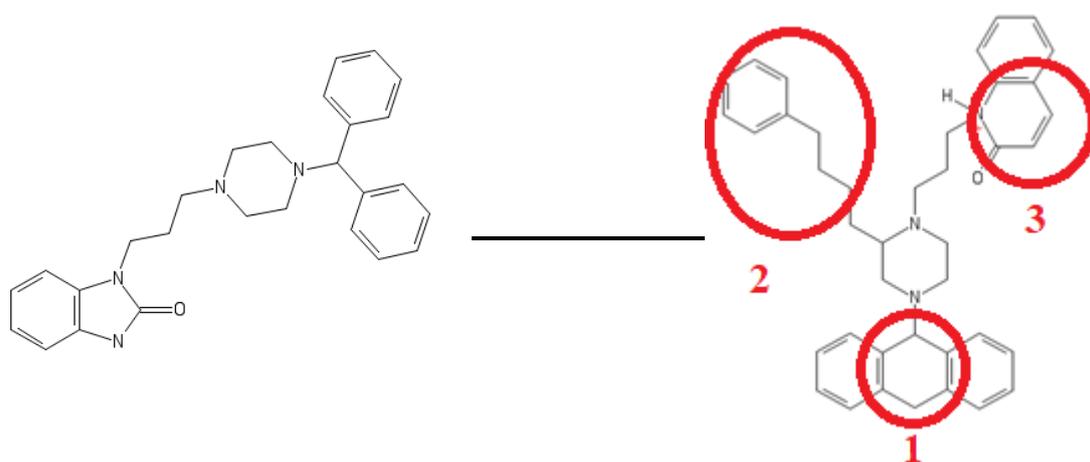


Figura 51. Transición de la molécula MMV003406 al ligando AS5

Para la preparación de este ligando se optó por la construcción de grupos funcionales grandes que sean capaces de introducirse completamente en las cavidades del sitio activo y permitan una inhibición más completa, se introdujo un anillo de ciclohexano entre los 2 grupos benceno del inicio de la molécula y se cambió el ciclopentano del ligando original por un ciclohexano.

2.3.4.1 Determinación de Interacciones Intermoleculares Ligando – Proteína

1. Se ingresó desde el navegador al siguiente dominio:
<http://biosig.unimelb.edu.au/arpeggioweb/about/>
2. Se ingresó al menú “Calculate”
3. Se cargó el archivo en formato PDB

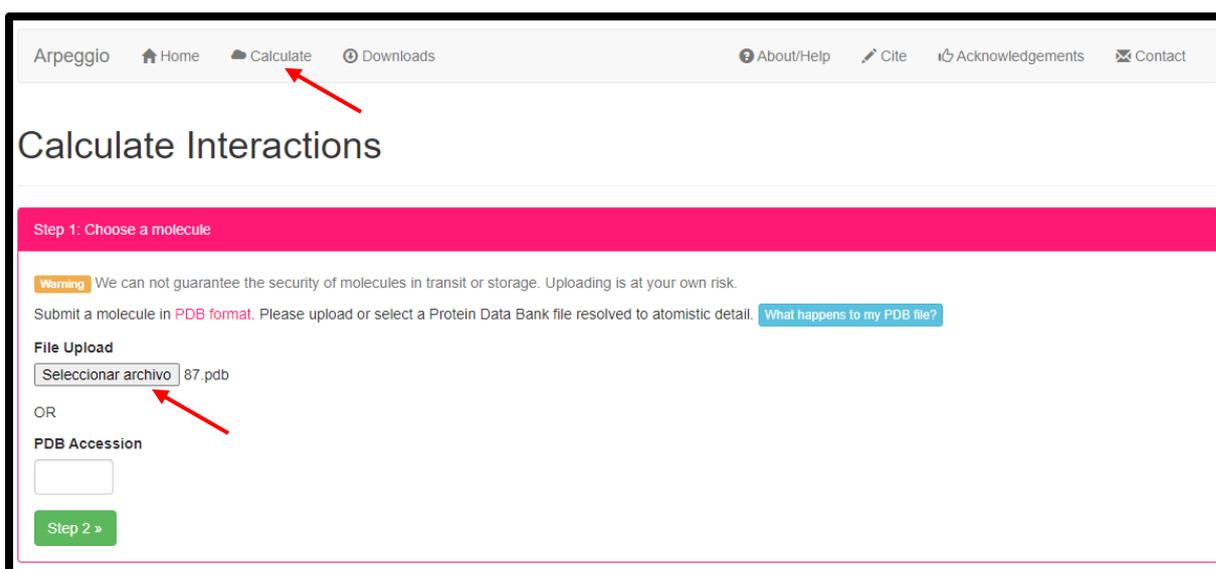


Figura 52. Entorno del dominio Arpeggio

4. Una vez seleccionado el archivo, el software permitió escoger la sección de la proteína en la que se desea calcular las interacciones, una vez seleccionadas, se envió a calcular.

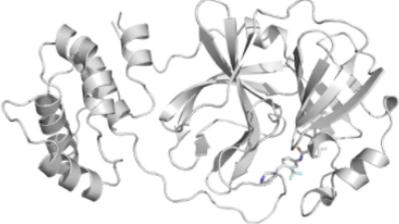
Entities to calculate contacts for

Heteroatom Groups
 Chain / Residue 1 (***)

Chain-Everything Interactions
 Interfaces between a chain and everything else in the structure.
 Chain
 Chain A

Selection
 Separate each selection with a new line. [How do I make a custom selection?](#)
 Leave the selection blank to calculate all contacts.

Calculate interactions »

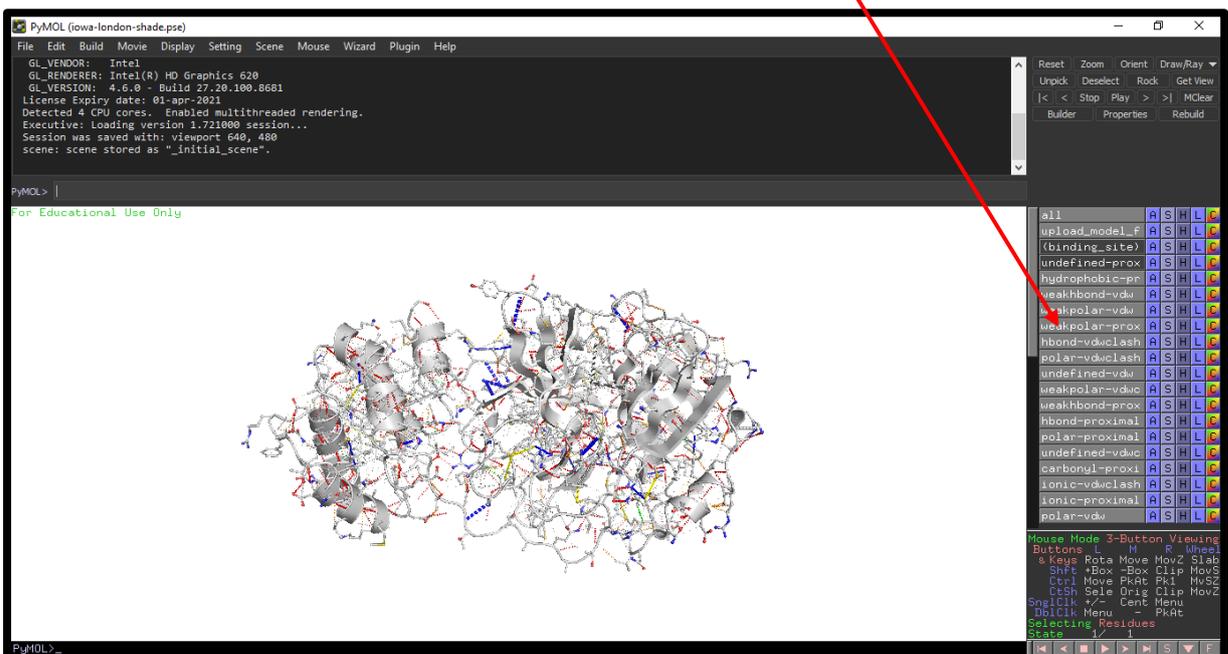


87.pdb

This is a preview of your structure following preprocessing. Please let us know if something doesn't look right at this point, quoting iowa-london-shade.

Figura 53. Configuración para la obtención de interacciones intermoleculares ligando-proteína.

5. El programa arrojó el resultado del cálculo en un archivo en formato PDB, el cual fue analizado en Pymol donde únicamente se tomó en cuenta las interacciones polares, hidrofóbicas, carbono Pi y azufre Pi, en el menú ubicado a la derecha (**Figura 54**).



PyMOL (iowa-london-shade.pse)

File Edit Build Movie Display Setting Scene Mouse Wizard Plugin Help

GL_RENDERER: Intel
 GL_RENDERER: Intel(R) HD Graphics 620
 GL_VERSION: 4.6.0 - Build 27.20.100.8681
 License Expiry date: 01-apr-2021
 Detected 4 CPU cores. Enabled multithreaded rendering.
 Executive: loading version 1.721800 session...
 Session was saved with: viewport 640, 480
 scene: scene stored as "_initial_scene".

PyMOL > |

For Educational Use Only

all	A	S	H	L	E
upload_model_f	A	S	H	L	E
(binding_site)	A	S	H	L	E
undefined-prox	A	S	H	L	E
hydrophobic-pr	A	S	H	L	E
weakbond-vdw	A	S	H	L	E
polar-vdw	A	S	H	L	E
weakpolar-prox	A	S	H	L	E
hbond-vdwlash	A	S	H	L	E
polar-vdwlash	A	S	H	L	E
undefined-vdw	A	S	H	L	E
weakpolar-vduc	A	S	H	L	E
weakbond-prox	A	S	H	L	E
hbond-proximal	A	S	H	L	E
polar-proximal	A	S	H	L	E
undefined-vduc	A	S	H	L	E
carbonyl-proxi	A	S	H	L	E
ionic-vdwlash	A	S	H	L	E
ionic-proximal	A	S	H	L	E
polar-vdw	A	S	H	L	E

Reset Zoom Orient Draw/Ray
 Unpick Deselect Rock Get View
 < > Stop Play > > | MClear
 Builder Properties Rebuild

Mouse Mode 3-Button Viewing
 Buttons L M R Wheel
 8 Keys Rotate Move Move2 Slab
 Shift +Box -Box Clip MoveS
 Ctrl Move PKAC PKI MoveZ
 Ctrl Sele Drag Clip MoveZ
 SingleClick Ctrl Menu
 DoubleClick Menu - PKAt
 Selecting Residues
 State 1 / 1

Figura 54. Interacciones intermoleculares calculadas, visualizadas en Pymol.

CAPÍTULO III

RESULTADOS Y DISCUSIÓN

3.1 HOMOLOGÍA DE LA PROTEASA Mpro de SARS CoV-2 RESPECTO A VIRUS PATÓGENOS HUMANOS

La investigación utilizó un enfoque que plantea el uso de moléculas biológicamente activas capaces de anular la actividad de la enzima Mpro la cual es un objetivo por excelencia al ser una enzima imprescindible para la replicación del virus, estructurada por 2 subunidades que forman un sitio activo conformado por la diada catalítica de His-41 y Cys 145 (Pérez & Arce, 2008) y a los que necesariamente la molécula inhibitoria debe unirse, junto con la mayor cantidad posible de aminoácidos que conforman el sitio activo, (Tabla 12) y (Anexo 3) resumen las posiciones de estas interacciones.

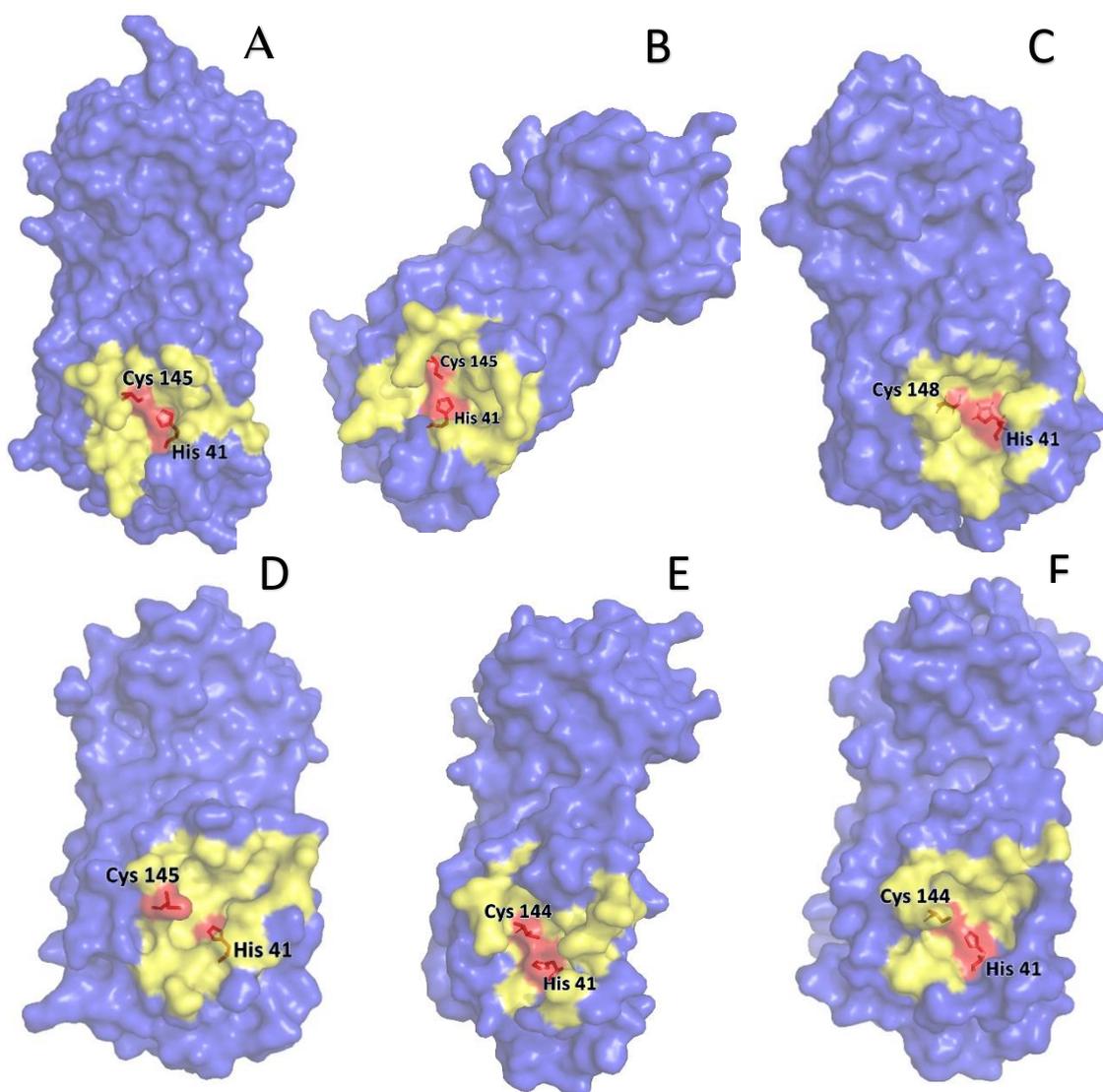


Figura 55. Proteasa Mpro de: A: Sars CoV-2; B: Sars CoV; C: Mers CoV; D: HKU1 CoV; E: NL63 CoV; F: 229E CoV, en amarillo se muestra el sitio activo y en rojo la diada catalítica formada por Histidina y Cisteína.

El punto de partida de esta investigación fue el análisis de homología de la cadena de péptidos que conforman el sitio activo de Mpro de SARS CoV-2 con respecto a otros coronavirus patógenos al ser humano, esto proporcionó información valiosa sobre los sitios de unión de las moléculas inhibidoras con respecto a la enzima gracias a los ligandos que habían sido desarrollados previamente para Mpro de virus similares. (**Tabla 6**). La selección de los virus homólogos se realizó mediante el uso de la herramienta BLAST de Uniprot, la cual tiene la función de comparar secuencias de 2 o más estructuras mediante el solapamiento de su cadena de aminoácidos con el fin de encontrar regiones altamente conservadas, el nivel de similitud entre las cadenas se conoce como % de identidad y mide el grado de correlación estructural entre las proteínas donde valores cercanos a 100 se pueden considerar homólogos. Por otra parte, el servidor arroja un valor conocido como “E-value” el cual mide la significancia estadística de la homología mediante el análisis de la probabilidad de un alineamiento al azar, siendo más confiable el resultado cuando este valor se acerca a cero (**Alonso, 2017**). Siguiendo estas directrices, se tomaron como válidos para el estudio los coronavirus SARS, MERS, HKU1, NL63 y 229 E ya que, si bien solo el primero contiene un % de identidad alto (96,1%), el valor de E-value determina su homología la cual está dada por la función de la cadena de aminoácidos del sitio activo de la enzima más que por su conformación estructural (**Figura 55**), donde se muestra el sitio activo de cada una de las proteasas de los diferentes coronavirus.

Tabla 6. Comparación estructural entre Mpro del virus SARS-CoV-2 y proteasas virales.

Organismos Similares	Uniprot			PDBe fold		
	Valor E	Identidad (%)	Identificador	P-score	rmsd	Código PDB
Coronavirus del Síndrome respiratorio agudo grave (SARS)	0,00	96,1%	P0C6X7	20,8	0,722	1zli

Coronavirus del Síndrome respiratorio de Oriente Medio (MERS)	8.1e-95	50,6%	A0A023YA54	38,22	1,297	4wmd
Coronavirus Humano (HKU1)	1.4e-89	49,0%	P0C6U5	33,06	1,333	3d23
Coronavirus humano (NL63)	1.5e-72	44.3%	P0C6X5	32,92	1,562	3tlo
Coronavirus humano (229E)	9.6e-68	41.0%	P0C6U2	36,62	1,203	2zu2
Rhinovirus Humano HRV-14	-	-	P03303	6,7	2,35	6kzo
Virus de la Hepatitis A	-	-	P08617	3,835	2,522	1hav

Valor E: Secuencias probablemente homólogas (0.02 y 1.0: homología no descartable; >1: Homología no significativa).

Identidad: Homología estructural (>45%: Estructura similar y posible función similar; >25% Similitud en el plegamiento 3D general; <20% No existe similitud).

rmsd: Raíz del error cuadrático medio de la superposición de los C α del polipéptido.

P-score: Representa logaritmo negativo del valor P y representa la probabilidad de encontrar un emparejamiento de calidad al elegir las estructuras de una base de datos al azar, es decir, es una medida de la calidad del emparejamiento que toma en cuenta el RMSD, el número de aminoácidos alineados, el número de brechas, la cantidad de elementos de estructura secundaria coincidentes y la puntuación de coincidencia de elementos de estructura secundaria.

El valor de homología reportado en la base de datos para SARS CoV (96,1%) difiere del porcentaje estipulado en la publicación de (Zhu et al, 2020) donde es del 86,9% obtenido a partir de ensayos experimentales de extracción y secuenciación del genoma viral a partir de muestras de lavado bronco alveolar obtenidos de pacientes con la infección, sin embargo esta variación puede explicarse dado que el estudio fue realizado usando el material genético completo de los virus, mientras que este estudio se basó únicamente en la secuencia proteica de la proteasa Mpro. El estudio realizado por (Liu & Wang) en cambio si corrobora el % de identidad determinado en este proyecto de investigación, reportando un valor de 96% de homología entre las enzimas Mpro del SARS CoV y SARS CoV-2.

Para identificar posibles inhibidores de Mpro, se utilizó la base de datos de Protein Data Bank, el cual arrojó resultados basados en otro tipo de valores, los que han sido tomados en cuenta en esta investigación son P-score, que es significativo cuanto mayor sea el valor, y el RMSD que ofrece un resultado confiable cuanto menor sea el valor. (Mas Benavente, 2001). Bajo estos parámetros se descartó los virus HRV-14 y el causante de la hepatitis A ya que no cuentan con una similitud estructural significativa respecto al SARS CoV-2 (Tabla 6).

3.2 SITIOS DROGABLES DE LA PROTEASA MPRO DEL VIRUS SARS COV-2 Y CORONAVIRUS HUMANOS PATÓGENOS HOMÓLOGOS

A continuación, se determinó las posiciones de los aminoácidos que interactúan con los diferentes inhibidores dentro de la proteasa Mpro, y se presentó aquellos que se muestran conservados en al menos dos coronavirus humanos (Tabla 7).

Para el análisis de los sitios drogables de Mpro se utilizó la base de datos de inhibidores de SARS CoV-2 y los virus homólogos seleccionados otorgada por Protein Data Bank, donde además se encontraban inhibidores recientemente desarrollados con la Covid 19 (Anexo 2) donde se encuentra los datos obtenidos del análisis de los sitios de unión de los diversos inhibidores a la proteasa. Se realizó un segundo filtrado del total de resultados arrojados por la base de datos y se escogió aquellos que presentaban mayores valores de similitud. La identificación de los aminoácidos objetivo se realizó a través del programa Wincoot. Mediante este análisis se pudo identificar dos tipos de ligandos, aquellos que

únicamente forman interacciones polares, apolares y puentes de hidrógeno y aquellos que además forman un enlace covalente con el residuo catalítico Cys145. Todos los inhibidores analizados se unen al sitio activo de la enzima, con lo cual se concluye que Mpro contiene solo un sitio drogable.

Tabla 7. Aminoácidos objetivo de inhibidores de la proteasa Mpro en coronavirus humanos

Aminoácidos	ORGANISMOS/POSICIÓN					
	SARS CoV-2	SARS CoV	MERS CoV	HKU1	NL63	229E
Ala	-	46	46	-	-	-
	-	-	-	-	143	143
	191	191	-	-	-	-
Arg	188	188	188	-	-	-
Asn	-	-	-	-	141	141
	142	142	-	-	-	-
Asp	187	187	-	-	187	-
Cys	44	44	44	-	-	-
	-	-	-	-	144	144
	145	145	145	145	-	-
Gln	-	-	164	164	164	-
	189	189	-	189	-	-
	192	192	192	192	192	-
Glu	166	166	-	166	166	-
Gly	143	143	-	143	-	-
His	41	41	41	41	41	-
	163	163	-	163	163	-
	164	164	-	-	-	-

	172	172	-	172	172	-
Ile	-	-	-	-	140	140
Leu	27	27	-	27	-	-
	141	141	-	141	-	-
Met	49	49	-	-	-	-
	165	165	-	-	-	-
Phe			-	-	139	139
	140	140	-	140	-	-
Pro	168	168	-		-	-
Ser	144	144	-	144	-	-
Thr	25	25	-		-	-
	26	26	-	26	-	-
	45	45	-	-	-	-
	190	190	-	-	-	-
Tyr	-	54	54	-	-	-

Los aminoácidos identificados en el presente proyecto de investigación fueron comparados con aquellos reportados bibliográficamente por (**Khaerunnisa et al, 2020**) donde identifica como drogables los residuos Thr 24, Thr 26, Phe 140, Asn 142, Gly 143, Cys 145, His 163, His 164, Glu 166, His 172, tomando como base de estudio únicamente el inhibidor N3 para SARS CoV-2 y Lys 5, Ala 7, Thr 25, His 41, Met 49, Tyr 54, Val 125, Tyr 126, Gly 127, Phe 140, Leu 141, Asn 142, Gly 143, Ser 144, Cys 145, His 163, His 164, Met 165, Glu 166, Leu 167, Pro 168, His 172, Asp 187, Arg 188, Gln 189, Gln 192, Ala 198, Lys 236, Tyr 237, Gln 273 para el SARS CoV tomando como base el inhibidor AZP, los cuales en términos generales son similares a los obtenidos en el presente proyecto de investigación (**Tabla 12**), (**Anexo 3**), lo que demuestra que este se encuentra correctamente encaminado, al haber coincidido en el reporte de residuos objetivo, con la inclusión de otros derivados del estudio de una librería de ligandos que ofrecen variadas interacciones.

En la literatura se estipula que la Cisteína catalítica es el principal objetivo de los inhibidores de Mpro, especialmente de aquellos que cuentan con un grupo indol capaz de

unirse covalentemente al azufre terminal de la misma (Ullrich & Nitsche, 2020). Sin embargo, en la presente investigación no se ha tomado en cuenta el diseño de inhibidores capaces de formar este tipo de enlaces con la proteasa Mpro ya que el acoplamiento molecular final fue realizado únicamente bajo condiciones de fuerzas intermoleculares no covalentes sigma. El docking con aquellos inhibidores que forman enlaces covalentes se incluirá en investigaciones posteriores.

Tabla 8. Aminoácidos Principales del sitio Activo de la Enzima Mpro

Ala	Arg	Asn	Asp	Cys	Gln	Glu	Gly	His	Ile	Leu	Lys	Met	Phe	Pro	Ser	Thr	Val
94	105	95	33	44	189	166	79	41	78	27	12	49	103	96	46	23	35
191	188	142	34	145	192	178	138	80		141	88	165	140	99	81	24	104
		228	145				143	163			90			168	144	25	171
		231	176				170	164			97			241		26	
			187					172			137					45	
																98	
																190	

En verde: Residuos del sitio activo de Mpro de SARS CoV-2 que típicamente interactúan con los ligandos

Los aminoácidos que conforman el sitio drogable de Mpro de SARS CoV-2 (Tabla 8) y aquellos conservados en todos los coronavirus homólogos (Tabla 7), proporcionan información importante, ya que el mejor ligando identificado en el cribado virtual y aquellos diseñados posteriormente, pueden potencialmente tener actividad inhibitoria en todos los coronavirus patógenos humanos, suponiendo un tratamiento para todas aquellas enfermedades en conjunto.

3.2.1 ALINEAMIENTO ESTRUCTURAL DE LOS CORONAVIRUS HUMANOS Y SUS AMINOÁCIDOS CONSERVADOS

El estudio de homología de SARS CoV-2 y virus homólogos se complementó con un alineamiento múltiple de sus secuencias (Figura 56) que dio como resultado la

identificación de los aminoácidos conservados del sitio activo en cada una de las proteínas y además permitió la identificación de la diada catalítica en cada una de ellas. La realización de este análisis permitió también observar que la posición de los aminoácidos que componen la diada catalítica no es la misma en todas las enzimas, en SARS CoV-2, SARS CoV y HKU1 se conserva en los residuos His 41 y Cys 45, en MERS CoV se presenta en His 41 y Cys 148, mientras que en NL63 CoV y 229E CoV en His 41 y Cys 148.

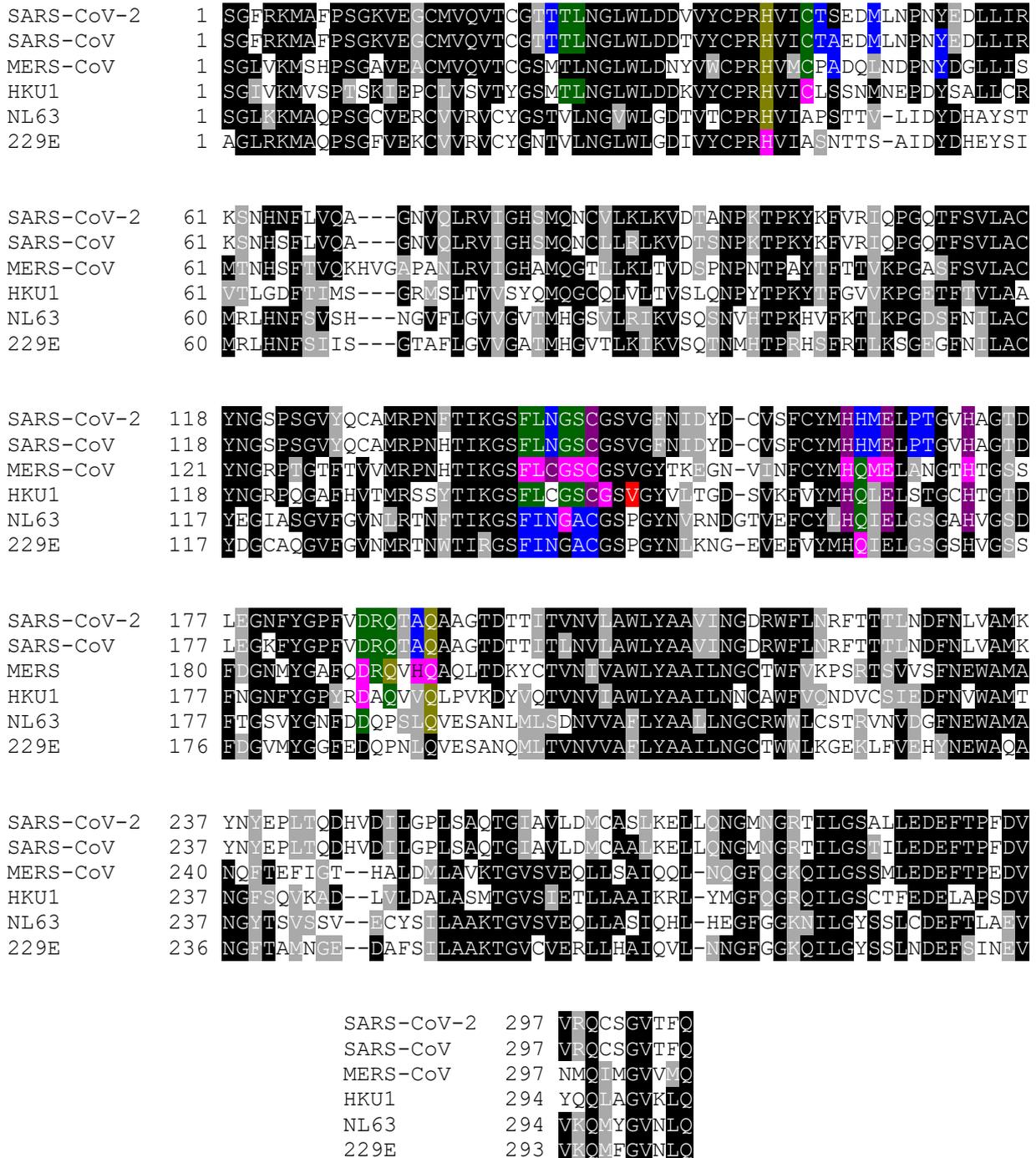
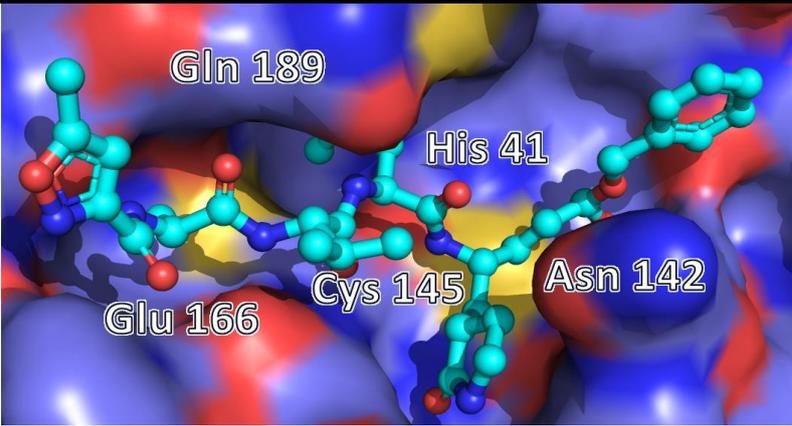
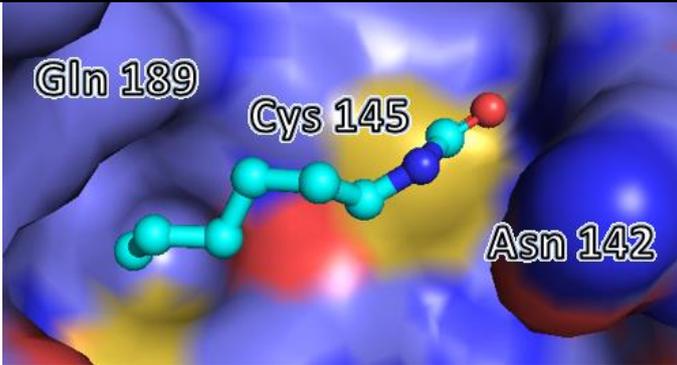
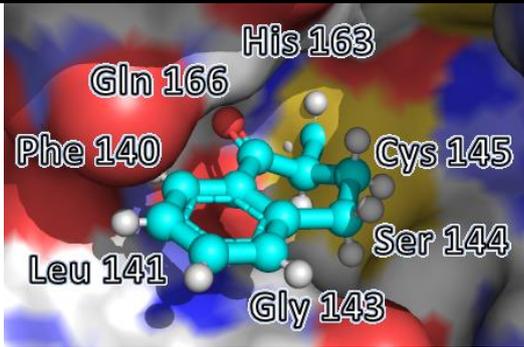
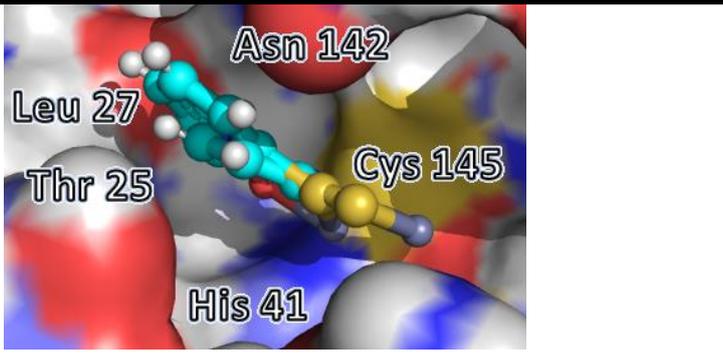
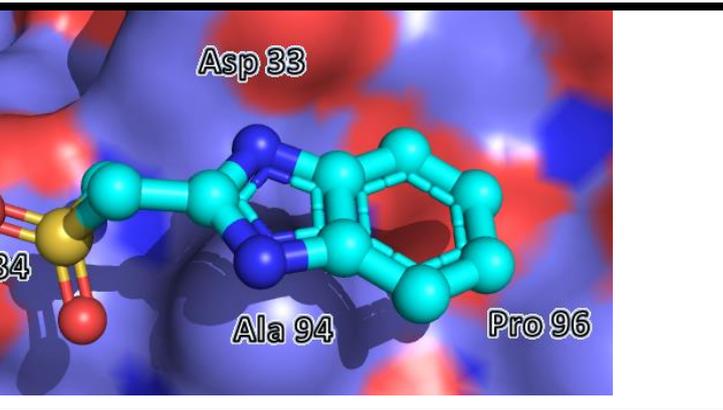
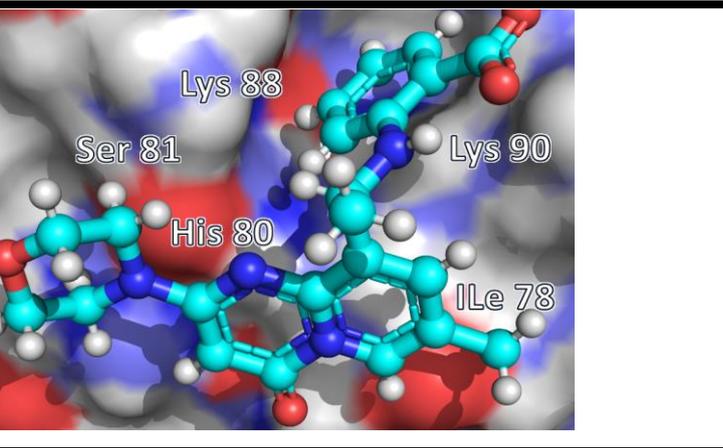
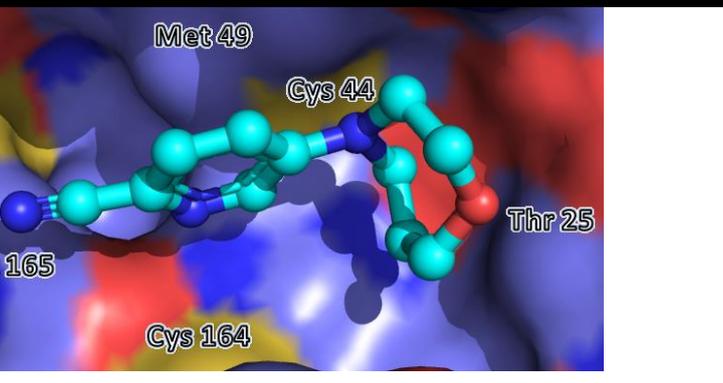


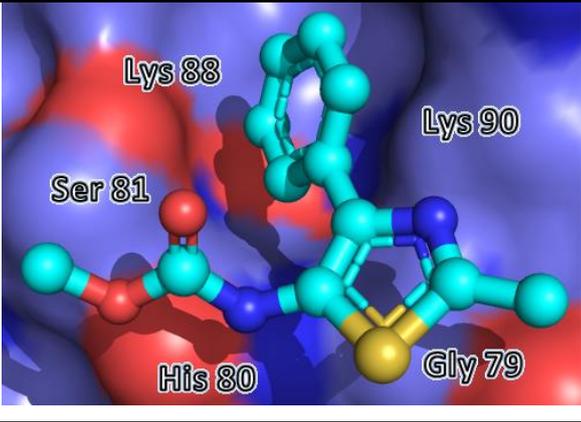
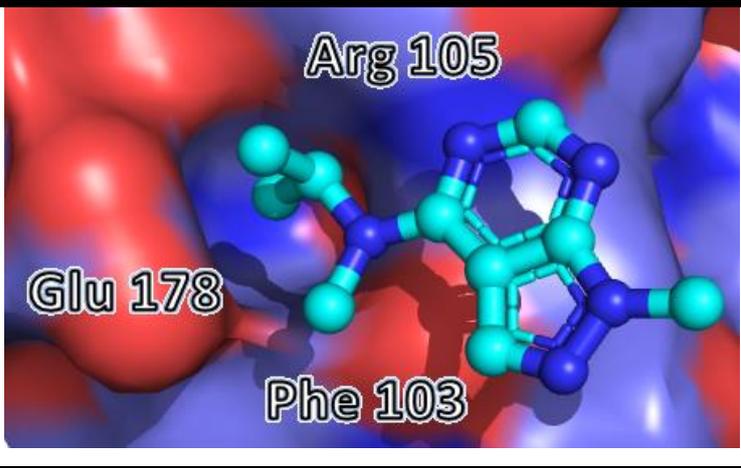
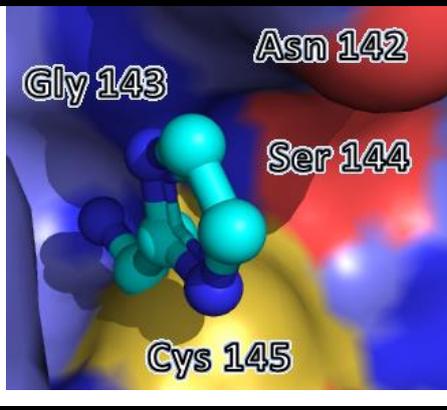
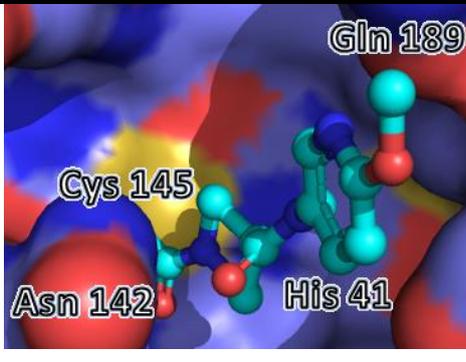
Figura 56. Alineamiento múltiple de secuencias de Mpro de diferentes virus.

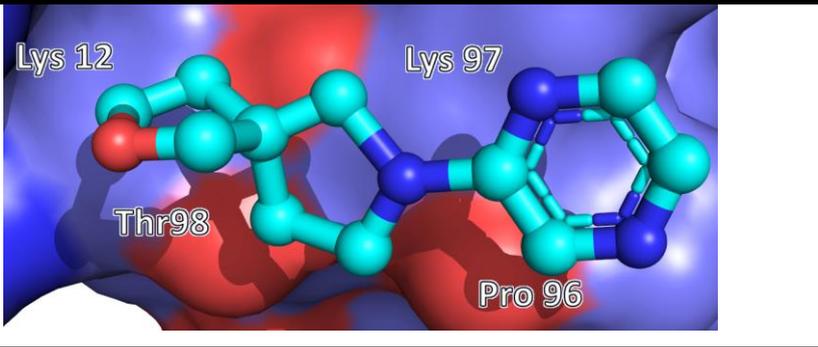
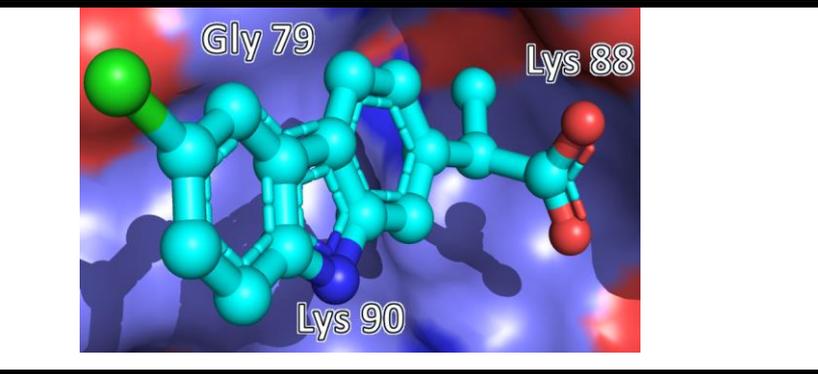
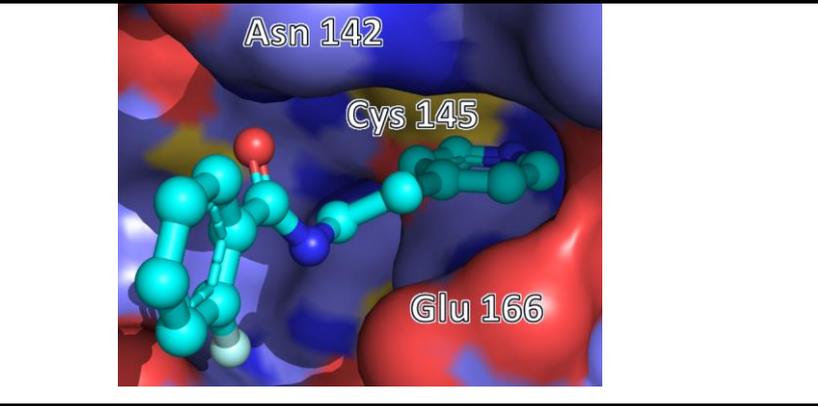
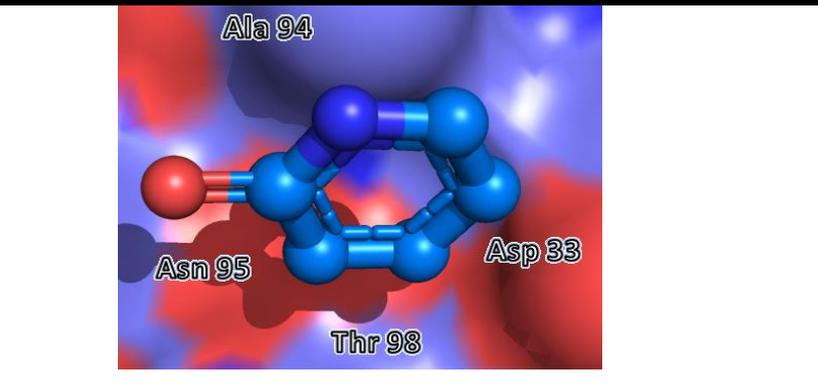
3.2.2 INTERACCIÓN DE LOS INHIBIDORES CON LA PROTEASA Mpro DE CORONAVIRUS PATÓGENOS HUMANOS

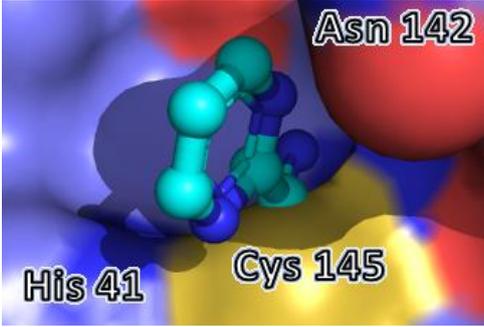
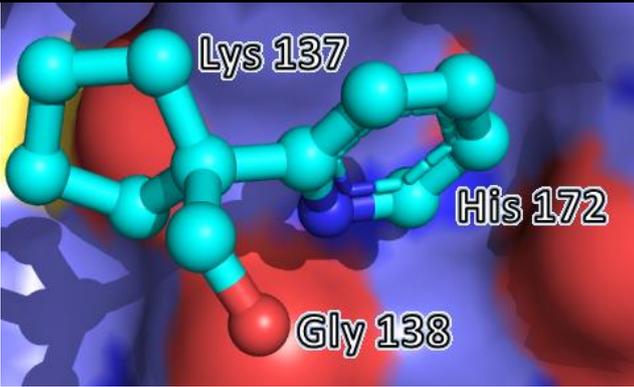
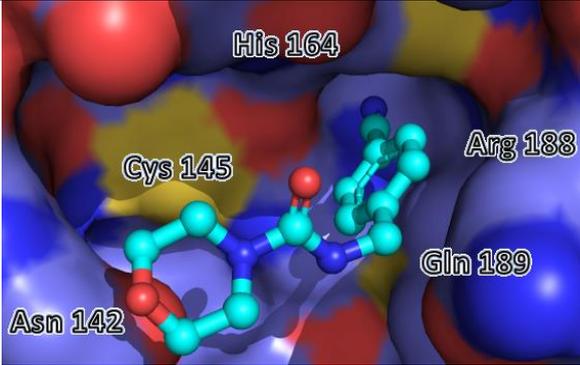
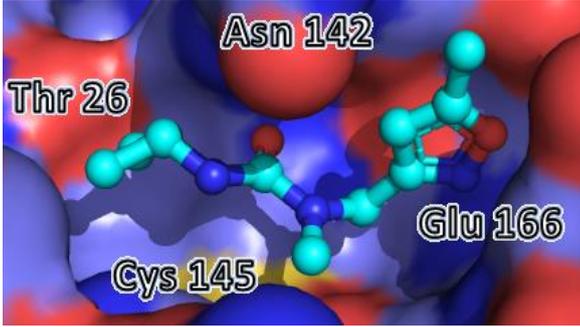
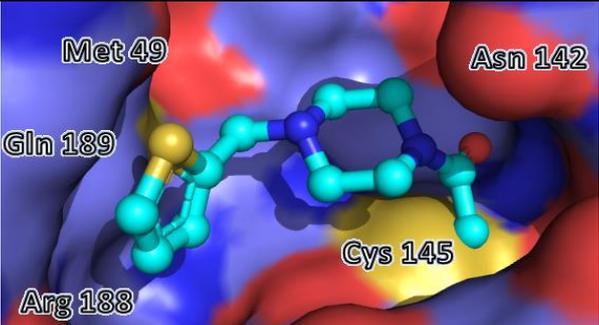
Tabla 9. Imágenes de las moléculas la librería de Proteín Data Bank

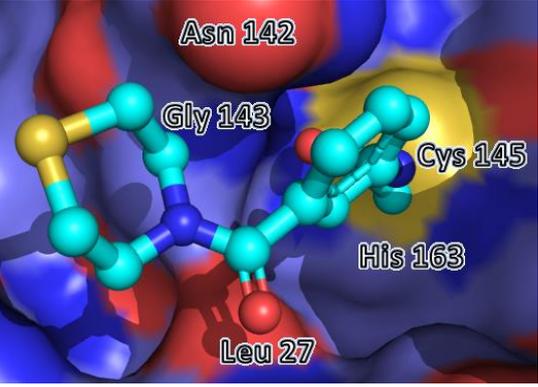
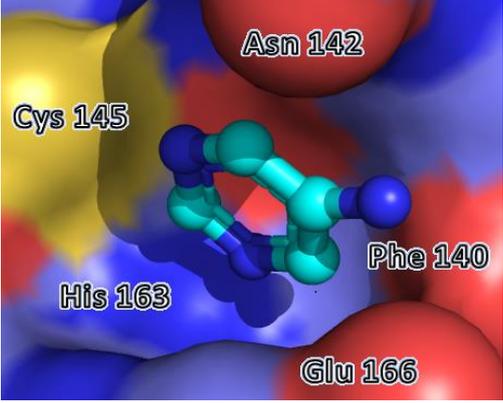
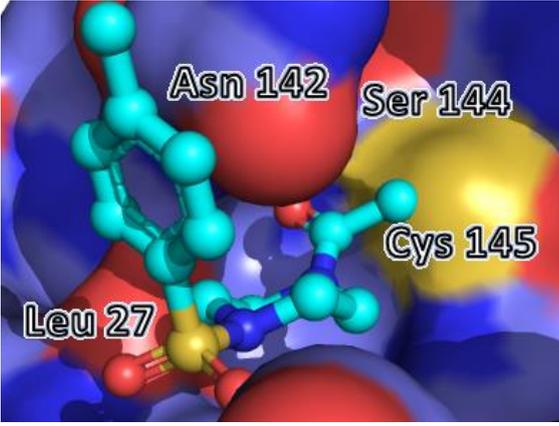
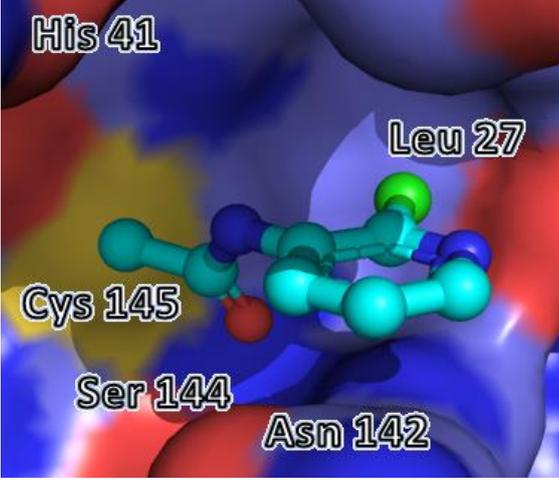
Proteasa	Inhibidor	Imagen
SARS CoV-2 P0TD1	N3 (covalente)	
SARS CoV-2 P0TD1	Carmofur (covalente)	
SARS CoV-2 P0TD1	P6N (covalente)	

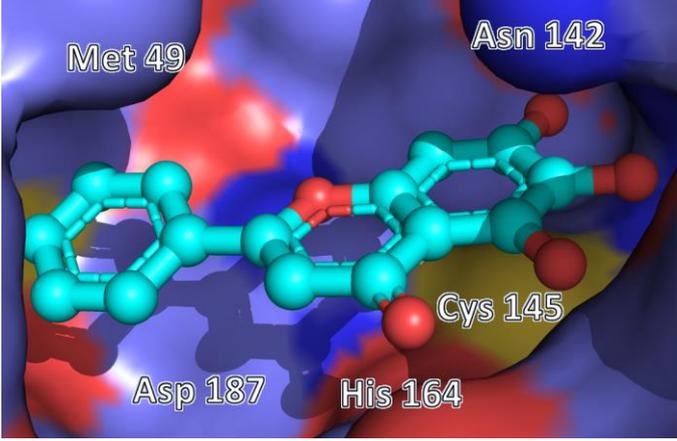
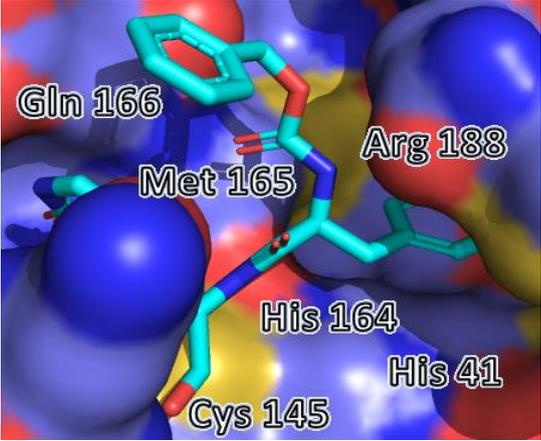
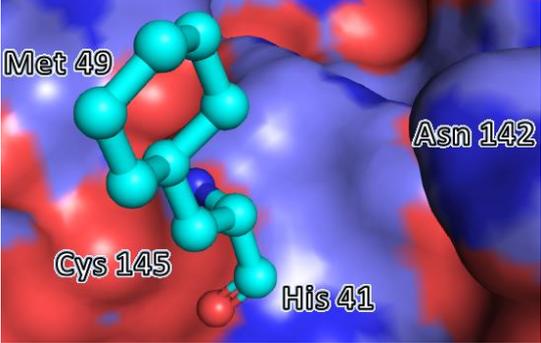
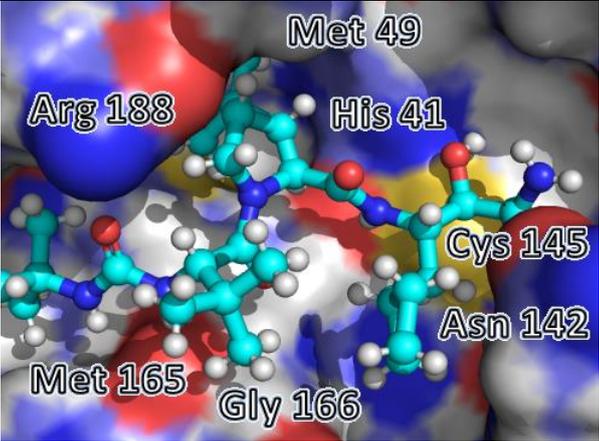
<p>SARS CoV-2 P0TD1</p>	<p>PK8 (covalente)</p>	
<p>SARS CoV-2 P0TD1</p>	<p>T6J (no covalente)</p>	
<p>SARS CoV-2 P0TD1</p>	<p>A82 (no covalente)</p>	
<p>SARS CoV-2 P0TD1</p>	<p>NTG (no covalente)</p>	

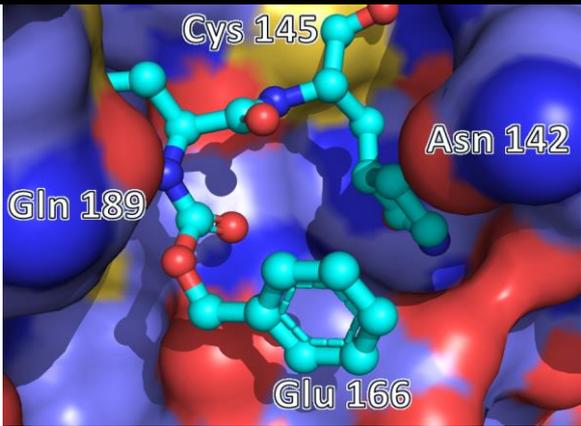
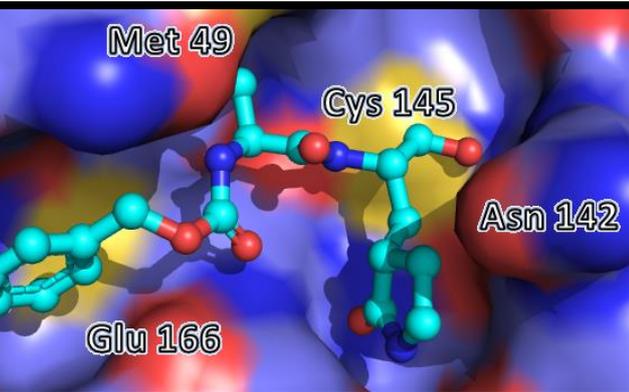
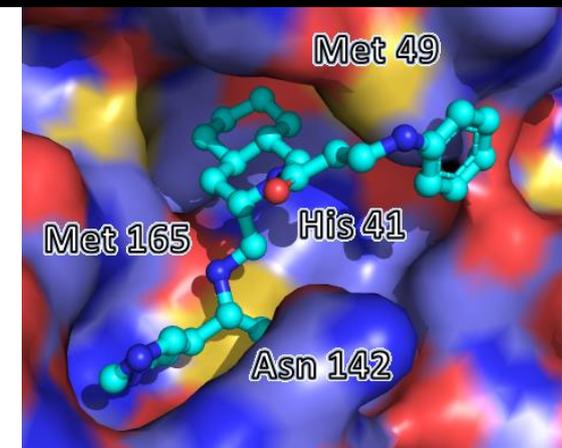
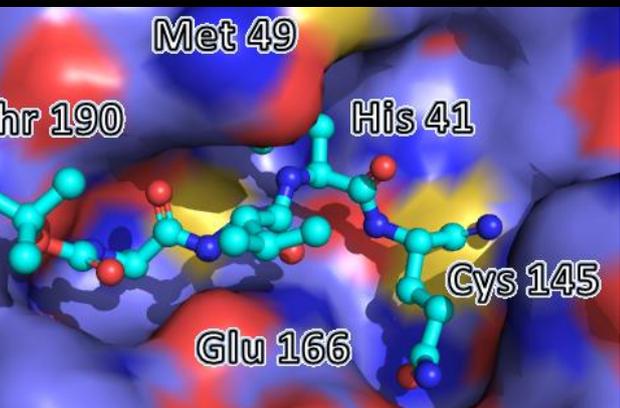
<p>SARS CoV-2 P0TD1</p>	<p>K1Y (no covalente)</p>	
<p>SARS CoV-2 P0TD1</p>	<p>K1G (no covalente)</p>	
<p>SARS CoV-2 P0TD1</p>	<p>USA (covalente)</p>	
<p>SARS CoV-2 P0TD1</p>	<p>UPD (covalente)</p>	

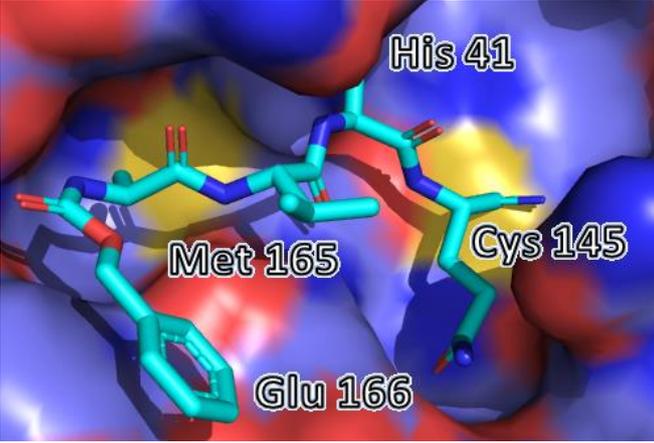
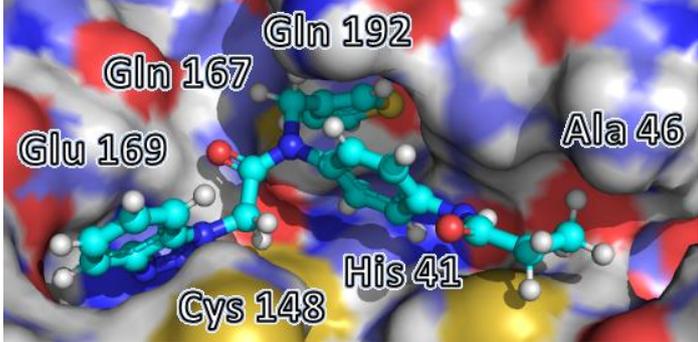
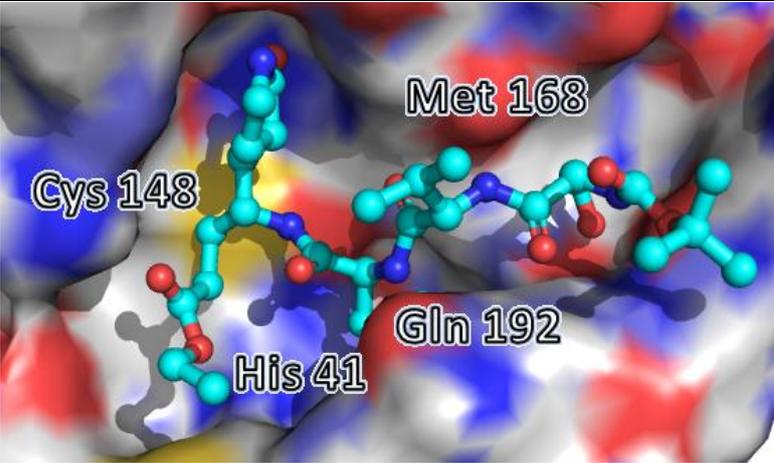
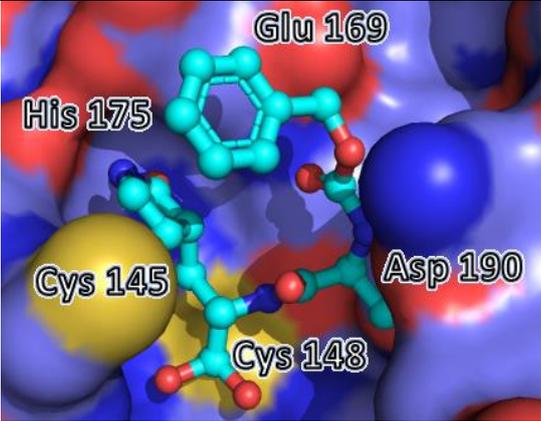
<p>SARS CoV-2 P0TD1U</p>	<p>U0S (no covalente)</p>	 <p>Lys 12 Thr98 Lys 97 Pro 96</p>
<p>SARS CoV-2 P0TD1U</p>	<p>UHG (no covalente)</p>	 <p>Gly 79 Lys 88 Lys 90</p>
<p>SARS CoV-2 P0TD1U</p>	<p>U0V (covalente)</p>	 <p>Asn 142 Cys 145 Glu 166</p>
<p>SARS CoV-2 P0TD1</p>	<p>T5Y (no covalente)</p>	 <p>Ala 94 Asn 95 Asp 33 Thr 98</p>

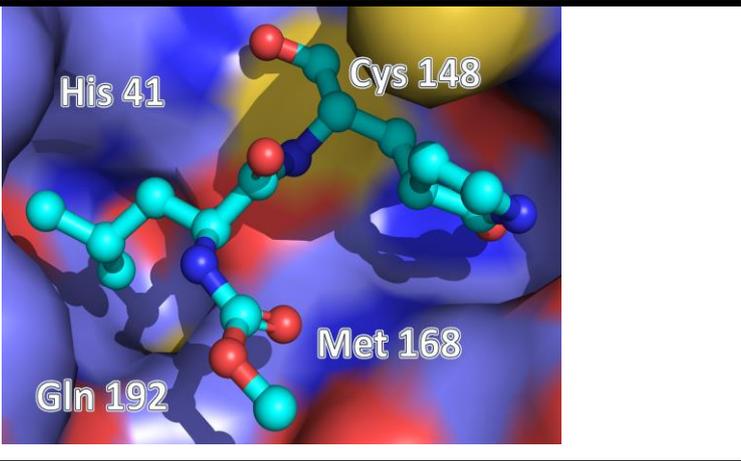
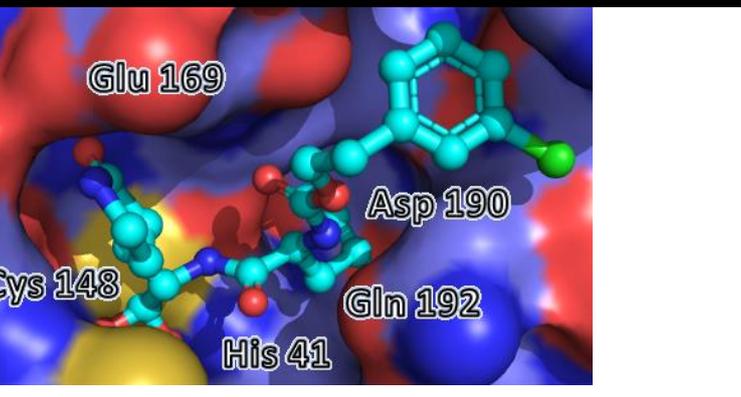
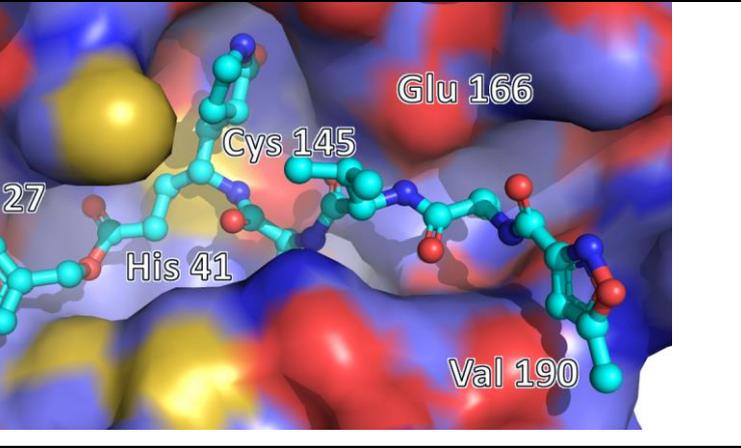
<p>SARS CoV-2 P0TD1</p>	<p>USD (covalente)</p>	
<p>SARS CoV-2 P0TD1</p>	<p>T5D (no covalente)</p>	
<p>SARS CoV-2 P0TD1</p>	<p>JGG (no covalente)</p>	
<p>SARS CoV-2 P0TD1</p>	<p>U0P (covalente)</p>	
<p>SARS CoV-2 P0TD1</p>	<p>T8M (covalente)</p>	

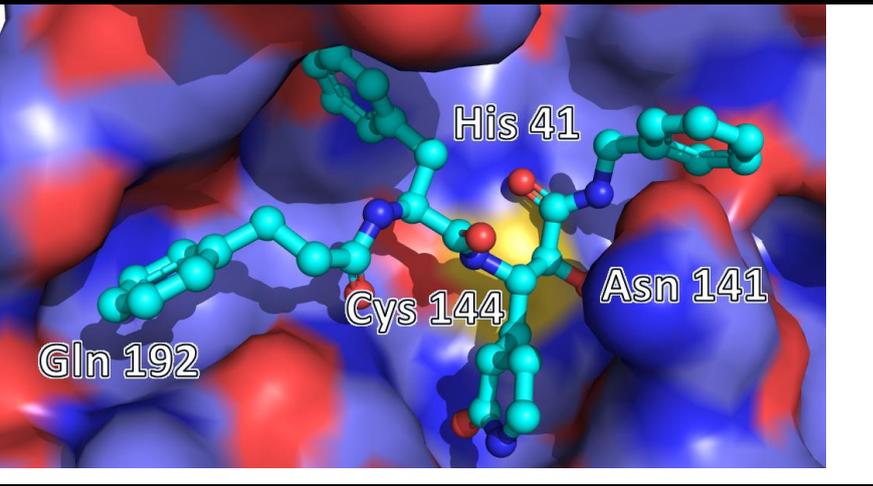
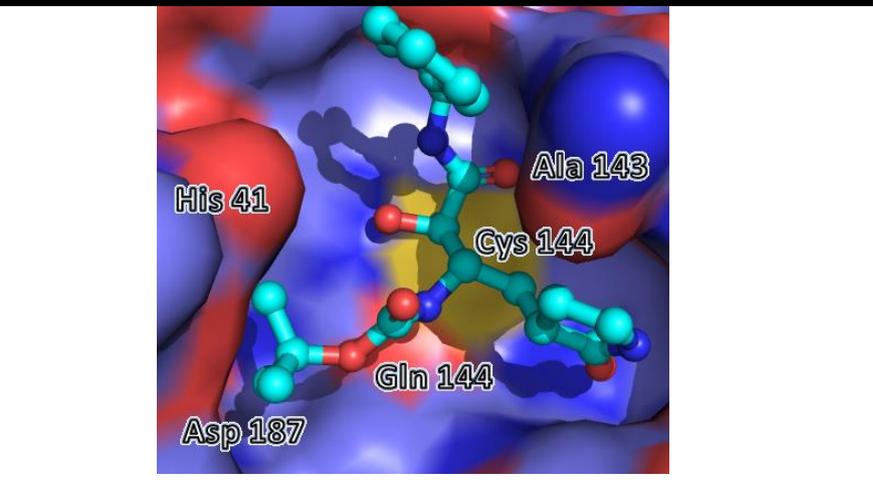
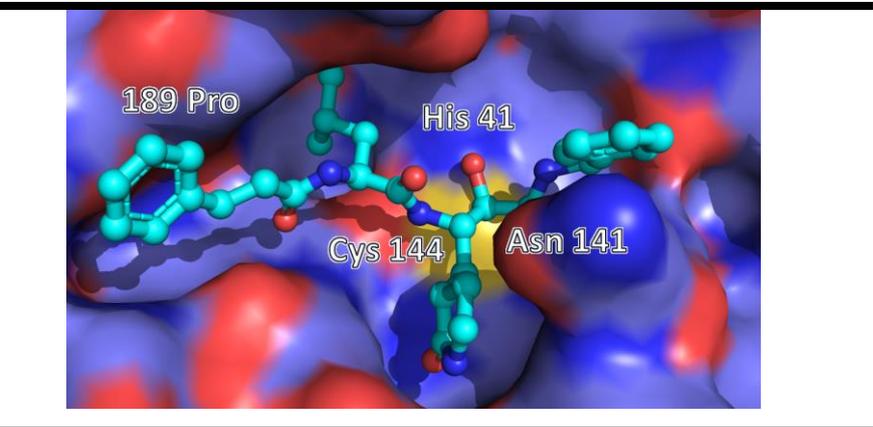
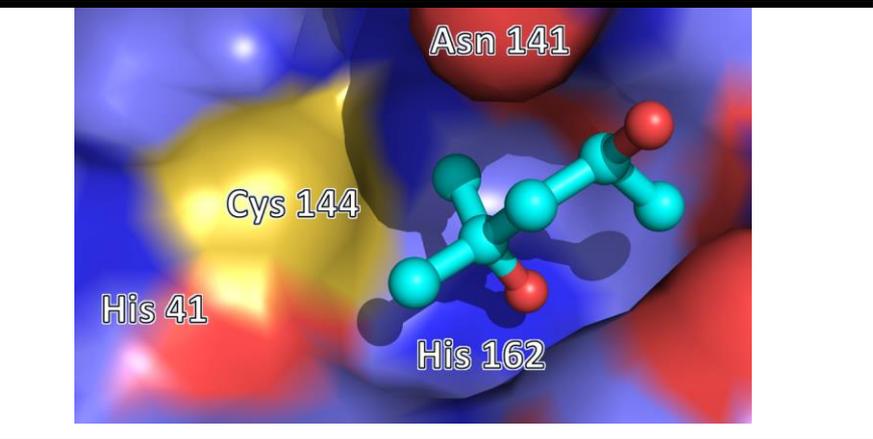
<p>SARS CoV-2 P0TD1</p>	<p>T4J (covalente)</p>	
<p>SARS CoV-2 P0TD1</p>	<p>T5V (no covalente)</p>	
<p>SARS CoV-2 P0TD1</p>	<p>U1A (covalente)</p>	
<p>SARS CoV-2 P0TD1</p>	<p>T8V (covalente)</p>	

<p>SARS CoV-2 P0TD1</p>	<p>3WL (no covalente)</p>	
<p>SARS CoV-2 P0TD1</p>	<p>UAW247 (covalente)</p>	
<p>SARS CoV: P0C6X7</p>	<p>ALC (covalente)</p>	
<p>SARS CoV: P0C6X7</p>	<p>U5G (covalente)</p>	

<p>SARS CoV: P0C6X7</p>	<p>UAW248 (covalente)</p>	
<p>SARS CoV: P0C6X7</p>	<p>K36 (covalente)</p>	
<p>SARS CoV: P0C6X7</p>	<p>SDJ (covalente)</p>	
<p>SARS CoV: P0C6X7</p>	<p>PRD_000910 (covalente)</p>	

<p>SARS CoV: P0C6X7</p>	<p>PRD_000911 (covalente)</p>	
<p>MERS CoV: A0A023YA54</p>	<p>R30 (No covalente)</p>	
<p>MERS CoV: A0A023YA54</p>	<p>PRD_002174 (covalente)</p>	
<p>MERS CoV: A0A023YA54</p>	<p>GC376 (covalente)</p>	

<p>MERS CoV: A0A023YA54</p>	<p>B6Y (covalente)</p>	
<p>MERS CoV: A0A023YA54</p>	<p>GC813 (covalente)</p>	
<p>HKU1: P0C6U5</p>	<p>PRD_002214 (covalente)</p>	

<p>NL63: P0C6U5</p>	<p>D03 (covalente)</p>	
<p>NL63: P0C6U5</p>	<p>8X8 (covalente)</p>	
<p>NL63: P0C6U5</p>	<p>E8E (covalente)</p>	
<p>229E: P0C6U2</p>	<p>MPD (no covalente)</p>	

La **Tabla 9** muestra imágenes del tipo de unión de los ligandos obtenidos de Protein Data Bank a la superficie de Mpro, permitiendo identificar aquellos que lo hacen a partir de un enlace covalente y los que solo propician interacciones intermoleculares no covalentes.

3.3 RESÍDUOS FLEXIBLES DE LA PROTEASA Mpro

Antes de someter las moléculas estudiadas a acoplamiento molecular, se tomó en cuenta un parámetro muy importante dentro de la proteómica, que es la flexibilidad de los residuos, identificados en (**Figura 57**), la superposición de la estructura tridimensional de Mpro unida a diferentes ligandos permitió la identificación de estos aminoácidos que cambian su posición de acuerdo a la estructura química de cada inhibidor (**Tabla 10**). Como menciona (**Naganathan & Orozco, 2011**) las diferentes conformaciones que adquieren los residuos en una proteína son esenciales para asegurar su correcta funcionalidad y lo más importante de todo es que esta movilidad se da sin ninguna alteración de la energía libre, razón por la cual no existe ningún problema con ensayar ligandos que necesiten la rotación de un residuo para su interacción.

Tabla 10. Identificación de residuos flexibles

Aminoácido	Posición	Rotámeros
Histidina	41	8
Metionina	49	13
Asparagina	142	7
Metionina	165	13
Arginina	188	34
Glutamina	189	9
Treonina	190	3

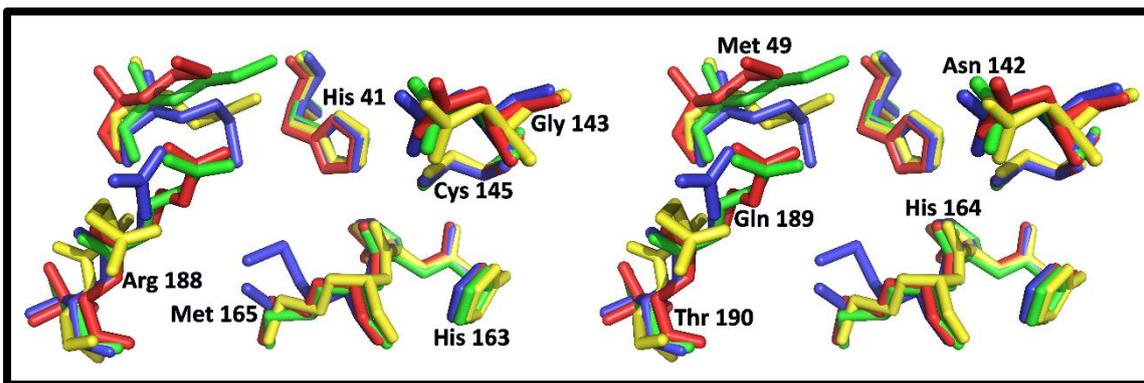


Figura 57. Contraste entre residuos rígidos y flexibles de la proteasa Mpro del virus SARS CoV-2

3.4 ACOPLAMIENTO MOLECULAR O DOCKING DE INHIBIDORES EN Mpro de SARS CoV-2

Para el estudio de docking o acoplamiento molecular se utilizó el software de GOLD - Protein Ligand Docking Software, donde una vez procesados todos los ligandos obtenidos de las diferentes librerías, el programa arrojó resultados (**Tabla 11**), en función del Piecewise Linear Potential (PLP) que es un valor establecido como resultado de la aptitud de unión del ligando a la proteína dado por la configuración espacial de átomos y grupos funcionales químicos complementarios en ambos, a los que sumados la incidencia de interacciones moleculares por puente de hidrógeno y metálicas, la rotación y posibles choques estéricos, forman el Fitness PLP que es el principal parámetro por el que se identificó los mejores ligandos en el estudio, siendo mejor cuando mayor sea el valor del mismo. Así mismo otros parámetros utilizados en la discriminación de ligandos fue el Clash Score que otorga una puntuación a nivel únicamente del ligando, sobre la posibilidad del choque de grupos funcionales grandes, y el CHOScore que es una medida de la presencia de enlaces débiles de carbono, hidrógeno y oxígeno. (**Cambridge Crystallographic Data Centre, 2019**).

Respecto a los ligandos y sus conformaciones que mayor afinidad de unión presentan hacia la proteasa Mpro del virus SARS CoV-2, se han seleccionado aquellos que cuentan con un PLP mayor a 80, constando tanto inhibidores de covid Box como de Protein Data bank, señalándose en verde oscuro los que mejor score total presentaron (**Tabla 10**).

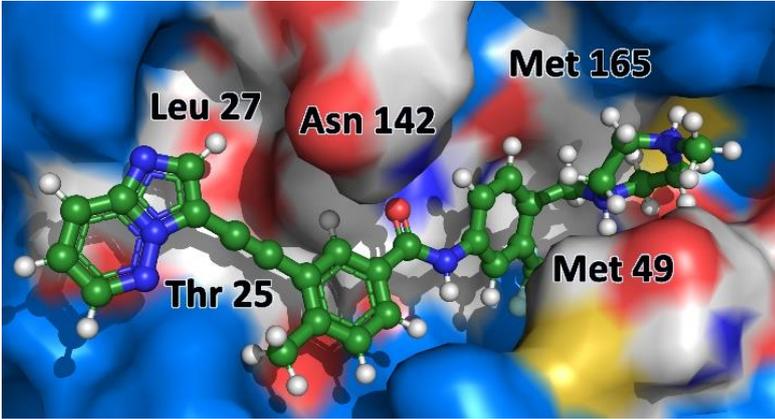
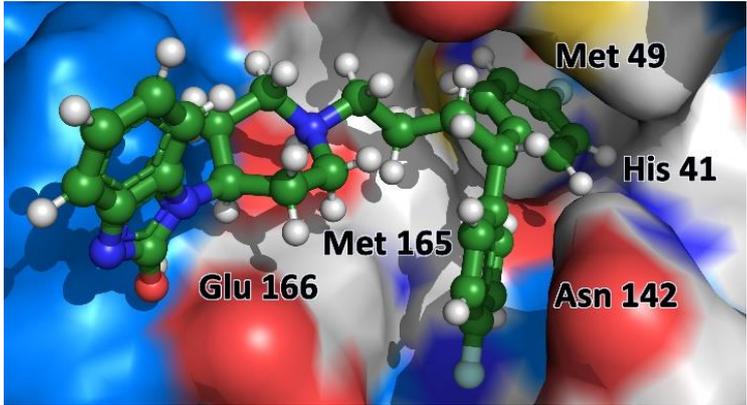
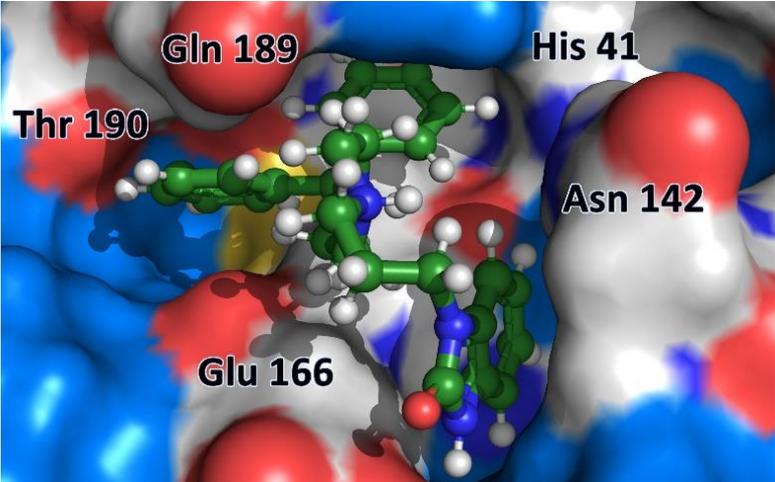
Tabla 11. Resultado Docking Molecular del Cribado Virtual

Ligando	Score	S (PLP)	S (Hbond)	S (Cho)	DE (clash)
MMV1580167	87,9070	-85,90	1,80	0,00	0,00
MMV002137	87,6550	-84,50	2,00	0,00	0,00
MMV003406	83,8294	-84,09	1,17	0,00	0,01
MMV000017	83,5532	-83,98	0,93	0,00	0,00
MMV1804312	83,7168	-85,35	1,33	0,00	0,00
MMV1804479	83,017	-86,65	1,00	0,00	0,14
MMV1804478	82,5887	-84,71	1,00	0,00	0,00
MMV1804620	82,0974	-82,53	1,07	0,00	1,40
MMV690733	80,7325	-83,47	1,00	1,00	0,06
MMV009415	80,3618	-73,18	0,99	2,59	0,0029

Según menciona (**Vivas-Reyes, Ruiz & Caraballo, 2009**) es sus estudios sobre inhibidores para la proteasa 3CL de SARS CoV, el diseño de inhibidores debe enfocarse en anular la entrada del sustrato al sitio activo, para lo cual es necesario construir ligandos capaces de interactuar con los residuos situados a los alrededores del sitio activo, en el estudio el autor utiliza un radio de 6,5 Å, lo que difiere minúsculamente de los 6°Å utilizados en la presente investigación, dentro de este radio se encuentran los aminoácidos que interactuarán con el inhibidor de entre los que destacan las Treoninas 21, 24, 25, 26 y Serinas 139, 144, 147 (**Bacha et al., 2004**).

Por otra parte, un estudio realizado por (**Marinho et al, 2020**) utilizó como sitio de referencia al interior de la proteína las coordenadas $x = -26.734$, $center_y = 13.009$, $center_z = 56.185$, $size_x = 94$, $size_y = 112$, $size_z = 108$ tomando como referencia el complejo de Mpro unido al ligando N3, de manera similar a cómo fue realizado en la presente investigación.

Tabla 12. Acoplamiento de los mejores inhibidores identificados, dentro de la proteasa Mpro

Ligando	Imagen
MMV1580167	
MMV002137	
MMV003406	

Se obtuvo imágenes de los tres mejores resultados para ligandos con actividad inhibitoria de la proteasa Mpro de SARS CoV-2 (**Tabla 12**) donde el de mayor score fue el ligando **MMV1580167**, esto significa que existió un acoplamiento tridimensional adecuado entre sus átomos y los de la proteína, no existen choques estéricos ya los grupos funcionales que lo conforman no son grandes y además cuenta con una alta incidencia de enlaces de puente de hidrógeno circunstantes que hacen la interacción más estable (**Figura 58**). Respecto al ligando **MMV002137**, presenta un score con una variación mínima respecto al primero, que de igual manera acomoda su estructura química correctamente a la proteína, pero, sin embargo, al tratarse de una molécula más pequeña, el número de interacciones con los aminoácidos del sitio activo disminuyen y por consiguiente su afinidad. El ligando **MMV003406** presenta un score más alejado con respecto a las otras dos moléculas debido en parte al tamaño pequeño de la molécula, a la menor presencia de enlaces de puentes de hidrógeno que deriva en una interacción débil con los átomos de la enzima y a la pequeña posibilidad de choques estéricos entre sus grupos funcionales. El score fue el parámetro seleccionado para la identificación de mejor ligando (**Tabla 10**).

3.5 DISEÑO RACIONAL DE INHIBIDORES DE LA ENZIMA Mpro

Una vez identificados los mejores ligandos a partir del cribado virtual, se procedió a mejorarlos a través de un diseño racional (*ver sección 2.3.4.1 de Metodología*), obteniéndose dos ligandos a partir de **MMV1580167**, dos a partir de **MMV002137** y tres a partir de **MMV003406**, esto debido a que este último presentaba una conformación molecular pequeña, siendo posible la evaluación de varios grupos funcionales.

Tabla 13. Resultados del docking de los ligandos diseñados en la proteasa Mpro

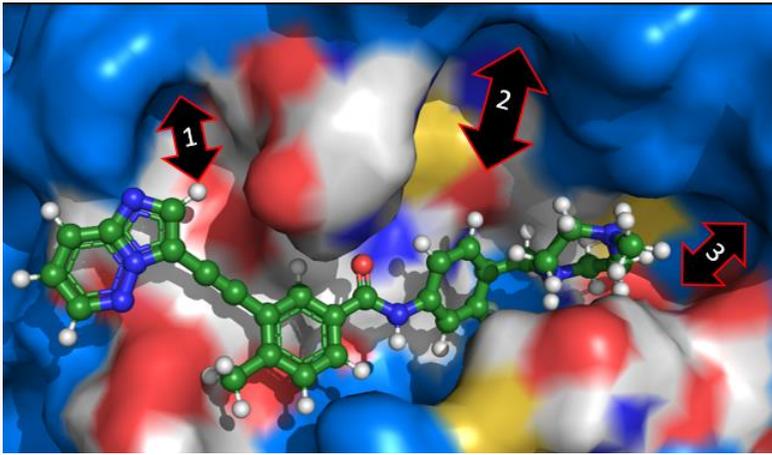
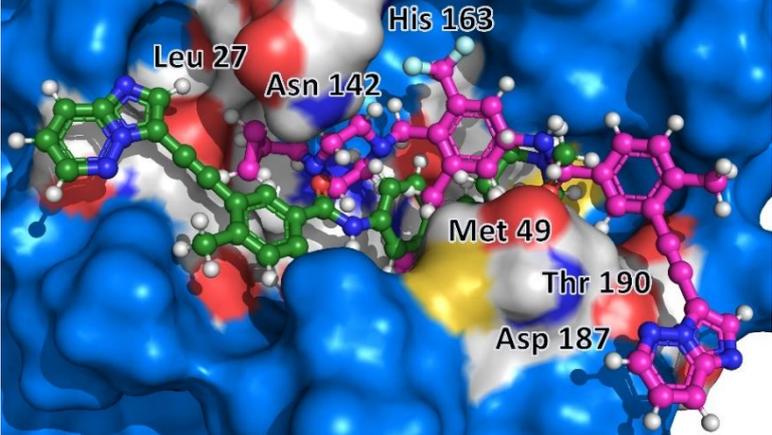
Ligando	Score	S(PLP)	S(Hbond)	PLP CHO	DE (Clash)
AS7	103,8653	-112,63	0,00	0,00	4,26
AS1	102,9805	-114,36	0,00	0,00	8,25
AS2	98,8814	-103,62	0,00	0,00	4,30
AS4	98,3817	-99,53	0,00	0,00	3,55
AS6	97,8247	-103,93	1,00	0,00	11,55
AS3	93,1152	-98,82	0,13	0,00	10,06
AS5	90,7535	-94,29	1,00	0,00	10,45

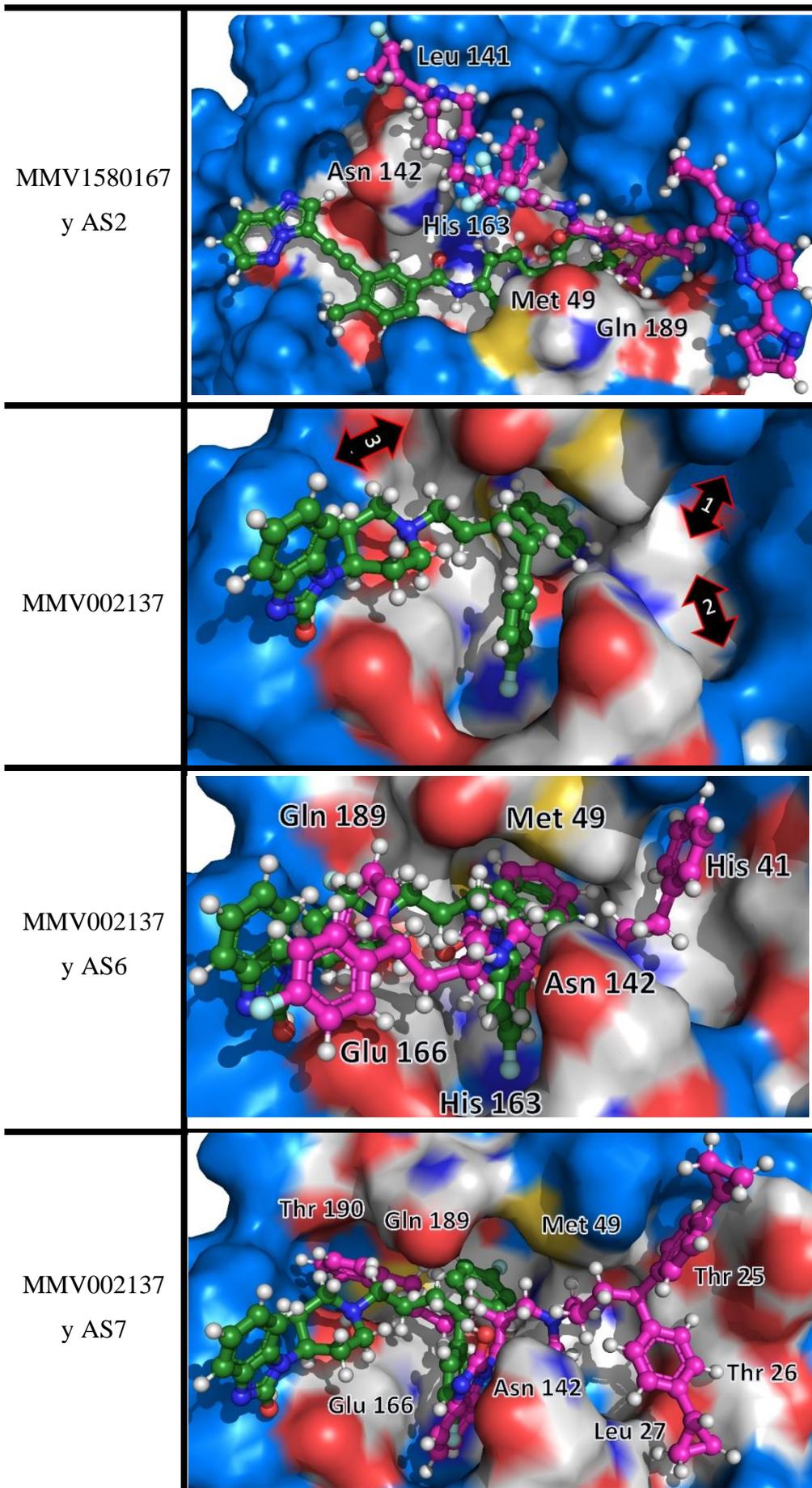
Una vez diseñados los ligandos, fueron sometidos a estudios de acoplamiento molecular al igual que aquellos obtenidos del cribado, identificándose al ligando AS7 diseñado a partir de MMV002137 (Figura 47) como el mejor, con un nivel de acoplamiento superior a los 100 puntos de score. Los valores de score fueron superiores a los determinados para sus compuestos parentales, lo cual sugiere que estos presentan una mayor afinidad por la enzima Mpro (Tabla 13).

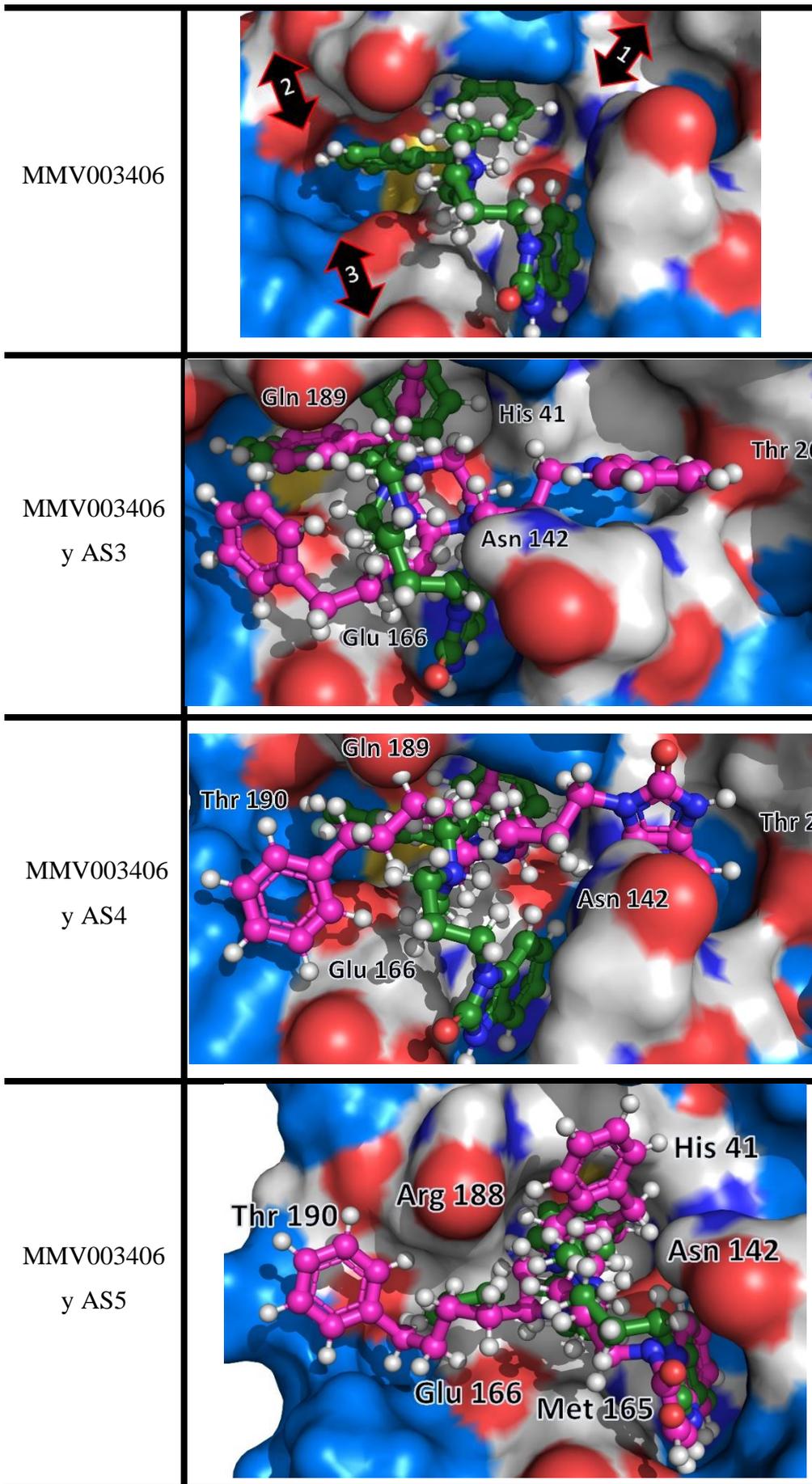
3.5.1 SUPERPOSICIÓN DE LIGANDOS DISEÑADOS

La superposición de ligandos diseñados a sus ligandos parentales arroja una idea sobre la conservación de la actividad biológica de la molécula y su efectividad de acoplarse a los sitios vacíos identificados.

Tabla 14. Superposición de ligandos mejorados a ligandos iniciales

Ligando	Imagen
MMV1580167	
MMV1580167 y AS1	





El ligando AS7 ocupa los tres espacios identificados dentro del sitio activo de la enzima (**Figura 8**), observándose en la **Tabla 14** que la conformación general del ligando diseñado se dispone de manera similar a **MMV002137**, pero con los grupos funcionales dispuestos a aproximadamente 2 Å de su posición original, tomando como referencia en la medición la distancia de un enlace carbono-carbono, esta disposición se debe al aumento de componentes moleculares que hicieron que el ligando se distribuya por todo el sitio activo de la proteína, con una torsión del ciclohexano central que permitió un mejor acoplamiento del grupo funcional conformado por la unión de un ciclopentano y un benceno. El único defecto de este ligando es la tendencia al choque estérico de sus grupos funcionales, dado por el valor de “Clash” igual a 4,26 (**Tabla 13**), efecto causado principalmente por el anillo de ciclopropano unido al sitio identificado uno, que sin embargo con una ligera torsión se ha acomodado correctamente.

El ligando identificado como AS1 fue desarrollado a partir del mejor ligando obtenido del cribado virtual, sin embargo, su actividad biológica pudo haberse anulado al colocarse en posición opuesta al ligando original (**Tabla 14**) y además si bien llena los sitios identificados (**Figura 5**), deja sin interacción a otros como aquel identificado en la posición de Leu 27, siendo la explicación de su alto score, la gran cantidad de enlaces intermoleculares creados a su alrededor.

El ligando AS2 por otra parte, muestra un desacoplamiento del sitio activo de Mpro debido a que se han colocado una cantidad exagerada de grupo funcionales que en la práctica no contribuyeron al mejoramiento de la actividad inhibitoria, denotándose que el alto score obtenido es gracias al gran tamaño de la molécula que es capaz de interactuar con otras regiones de la enzima (**Tabla 14**).

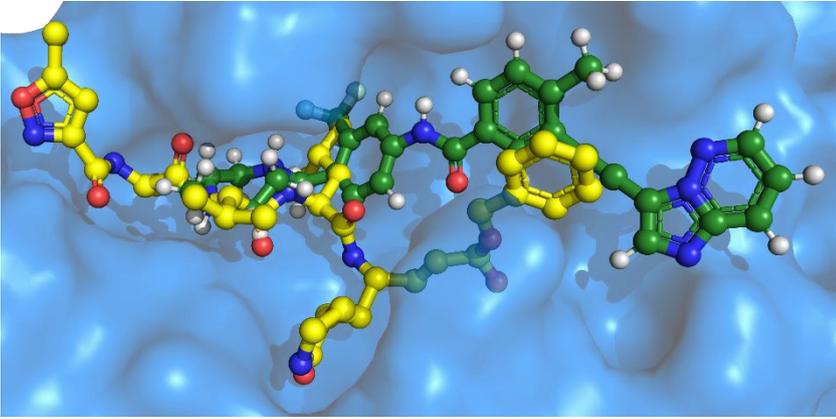
En el diseño de los ligandos AS4 y AS6 se incluyó además de los grupos funcionales clásicos, átomos de flúor al extremo de la molécula, tal como se observa en las **Figuras 13 y 9** respectivamente, con la finalidad de incluir un enlace iónico fuerte entre este átomo y el carbono gracias a la alta electronegatividad del flúor de 4,0 (**O'Hagan, 2008**) por lo cual es necesario que estos átomos se acerquen lo más posible a la estructura de la proteasa ya que el enlace se da a una distancia corta. El resultado presentado en (**Tabla 13**) permite corroborar esta información al tener el ligando AS4 un mayor score de afinidad con respecto a su homólogo AS3 que no contiene los átomos de flúor.

Finalmente, el ligando AS5 fue el que menor variación presentó en su disposición dentro del sitio activo con respecto al ligando original MMV003406, solventando todos los sitios de acoplamiento identificados en la **Figura 48**, gracias a la adición de grupos funcionales (**Tabla 14**) sin embargo la adición de un anillo de ciclohexano entre sus grupos fenílicos (**Figura 51**) dio como resultado un alto grado de choque estérico que disminuyó su score de afinidad hacia Mpro, a pesar de esto, el valor de score obtenido aún sigue siendo alto respecto a los obtenidos mediante el cribado virtual, en todos los casos solo una evaluación en vivo del inhibidor frente a Mpro para asegurar su efectividad.

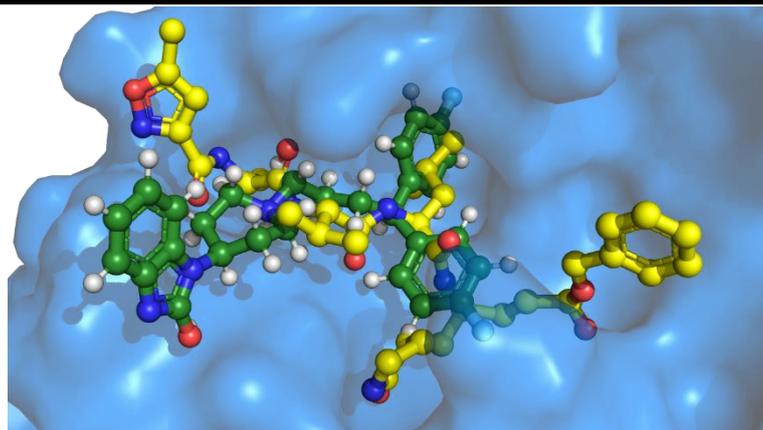
3.6 COMPARACIÓN DE LIGANDOS IDENTIFICADOS Y DISEÑADOS CON LIGANDO DE REFERENCIA N3

El ligando N3 fue la molécula de referencia que se utilizó para el estudio de docking o acoplamiento molecular, ya que es una molécula capaz de introducirse completamente dentro del sitio activo de Mpro, donde estudios de actividad biológica demostraron que este se acopla inhibiendo la enzima misma de manera irreversible, uniéndose directamente al bolsillo de unión del sustrato con una dosis letal media de 133uM (**Jin et al, 2020**).

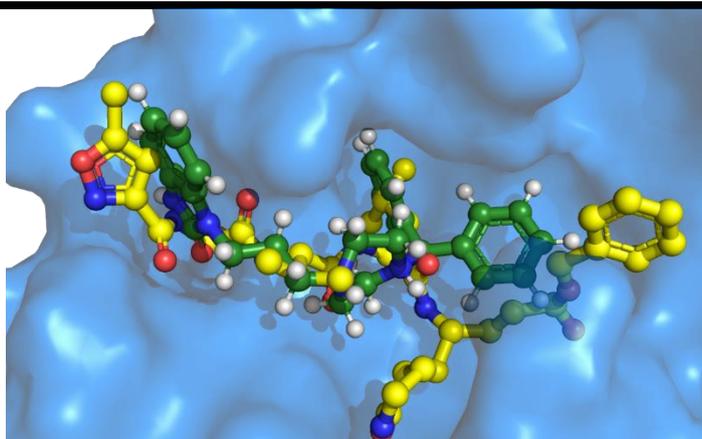
Tabla 15. Superposición de los mejores Ligandos a N3

Ligando	Imagen
MMV1580167 (verde) y N3 (amarillo)	

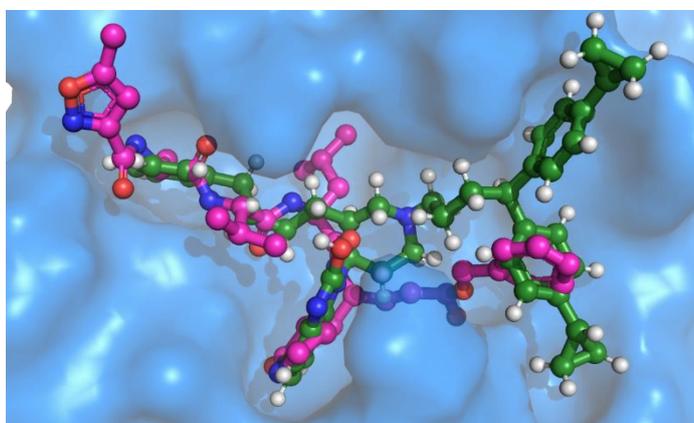
MMV002137
(verde) y N3
(amarillo)



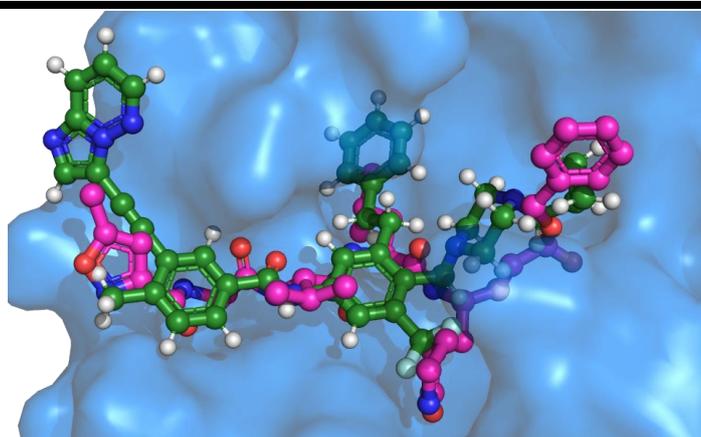
MMV003406
(verde) y N3
(amarillo)



AS7
(verde) y N3
(magenta)



AS1
(verde) y N3
(magenta)



El análisis de comparación de estructuras entre los ligandos identificados y N3 permite tener una idea del grado de actividad biológica de los primeros, al ser N3 una molécula que ha sido probada eficientemente en ensayos de laboratorio, en el caso de **MMV1580167**, se observa que la disposición espacial del ligando se encuentra un poco salida del sitio activo de la enzima, a comparación de N3 que se encuentra incrustado plenamente, sin embargo ofrece nuevas interacciones en sitios no alcanzados por este último (**Tabla 15**). El ligando **MMV002137** presenta regiones de incrustación más profundas a comparación de N3, sin embargo, la configuración espacial de su grupo funcional indol se encuentra alejado del sitio activo respecto a la molécula de referencia, la cual a su vez es más grande que el ligando (**Tabla 15**). Respecto a **MMV003406**, su disposición espacial es similar a N3, por lo que se puede asegurar que presenta actividad biológica de inhibición, al encontrarse incrustado justo en el bolsillo de unión del sustrato, sin embargo, su pequeño tamaño causa que el número de interacciones intermoleculares con la proteasa Mpro sea menor (**Tabla 15**).

En el caso de los ligandos diseñados se puede observar que **AS7** se encuentra incrustado dentro del bolsillo de unión del sustrato de la proteína de la misma manera que lo hace N3 (**Tabla 15**), y como sus grupos funcionales se encuentran en mayor número, se acopla con mayor firmeza a la proteína por lo que se puede decir que **AS7** es una molécula potencial. Por otra parte ligando **AS1** se une también en términos conformacionales de manera similar a N3, pero su grupo indol se encuentra interaccionando en las afueras del sitio activo (**Tabla 15**), lo que lo que podría reducir su capacidad inhibitoria, en el caso de estas dos moléculas el espacio que ocupan es mayor con respecto a N3, por lo que podría aumentar la afinidad de unión de estos le ganas si utiliza una molécula de referencia que equipare eso tamaño, incluyendo también parámetros de flexibilidad en los residuos periféricos del sitio activo

3.7 INTERACCIONES INTERMOLECULARES IDENTIFICADAS EN LOS MEJORES LIGANDOS

Los enlaces carbono pi se presentan cuando se da la hibridación de dos orbitales P, generando un enlace del tipo no covalente entre átomos de carbono, sin embargo, son más débiles que lo enlaces covalentes de tipo sigma ya que los primeros requieren de una mayor energía al encontrar su nube de electrones cargada negativamente a una distancia más extendida del núcleo atómico (**Jemmis et al., 2006**).

Respecto a los enlaces azufre pi, el fundamento es el mismo descrito anteriormente, con la diferencia de que en esta ocasión los átomos de S son los que forman el enlace pi con el carbono

Por otra parte, las interacciones hidrofóbicas se encuentran presentes de manera general en todo sistema ligando - proteína, debido a la presencia de grupos funcionales hidrofóbicos superficiales en la misma (Grigera, 1989), en Mpro están dados principalmente por los aminoácidos Leucina, Metionina y Fenilalanina, que son objetivos recurrentes de los inhibidores de esta proteasa.

Tabla 16. Identificación de interacciones Intermoleculares

Tipo de Interacción	Identificación
Polar	Enlaces rojos
Hidrofóbica	Puntos verdes
Carbono Pi	Enlaces magenta
Azufre Pi	Enlaces amarillos

3.7.1 LIGANDOS OBTENIDOS DEL CRIBADO VIRTUAL

❖ Ligando MMV1580167

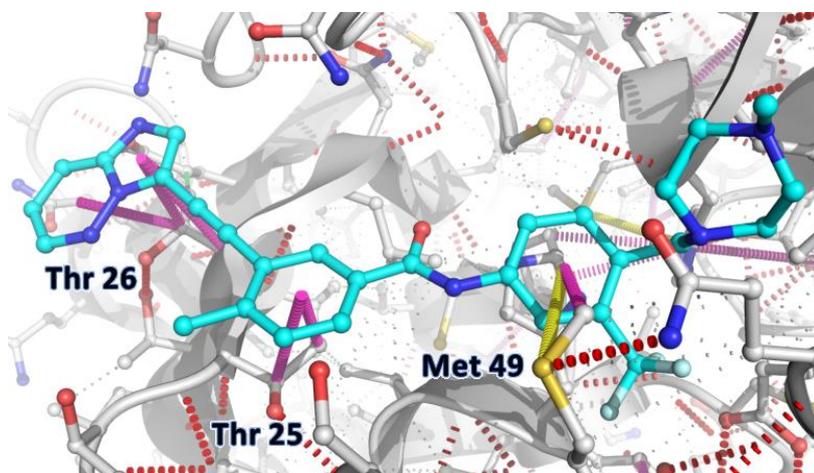


Figura 58. Interacciones intermoleculares ligando MMV1580167

❖ **Ligando MMV002137**

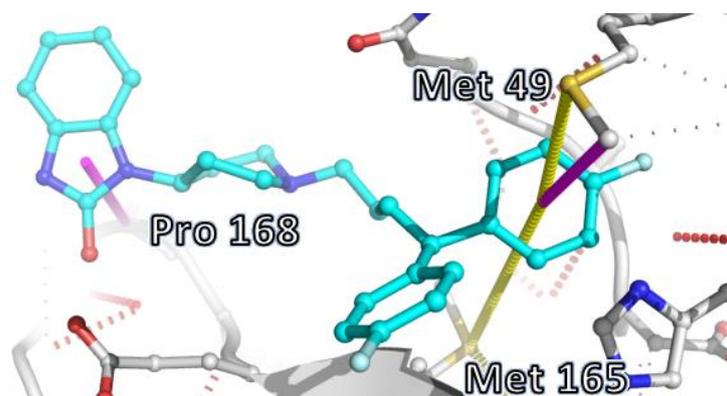


Figura 59. Interacciones intermoleculares ligando MMV002137

❖ **Ligando MMV003406**

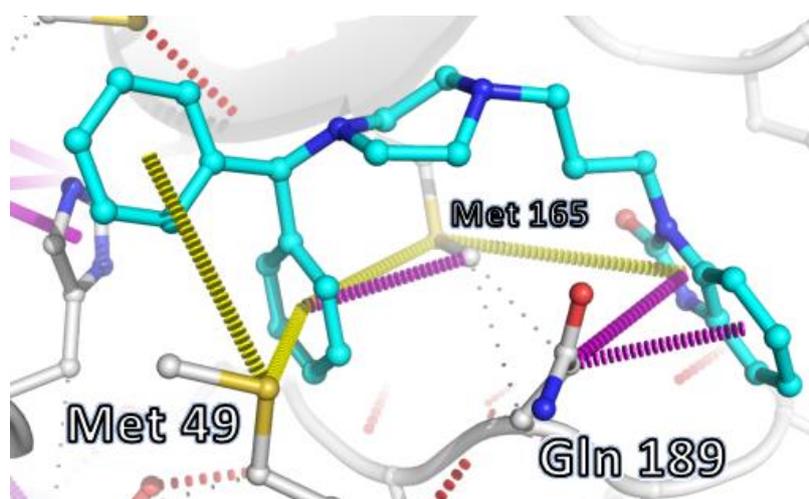


Figura 60. Interacciones intermoleculares ligando MMV003406

La presencia de enlaces carbono-carbono le otorga al ligando una mejor afinidad de unión al sitio activo de la proteína, en las **Figuras 57, 58 y 59** respectivamente, se observa que el ligando **MMV1580167** forma este tipo de enlaces con los aminoácidos Thr 25, Thr 26 y Met 49, respecto a **MMV002137** que solo lo hace con Met 49 y Pro 168 y **MMV003406** con un enlace de este tipo con Met 49 y Gln 189, el criterio de presencia de estos enlaces a su vez corrobora el resultado de docking molecular, donde **MMV1580167** es el mejor ligando.

En las **Figuras 57, 58 y 59** respectivamente, se observa también que los ligandos **MMV1580167** y **MMV002137** forman un único enlace pi de este tipo con Met 49 la cual contiene azufre en su estructura, mientras que **MMV003406** lo hace también con Met 169, siendo por este parámetro, superior a los otros dos ligandos.

3.7.2 LIGANDOS OBTENIDOS DEL DISEÑO RACIONAL

❖ Ligando AS7

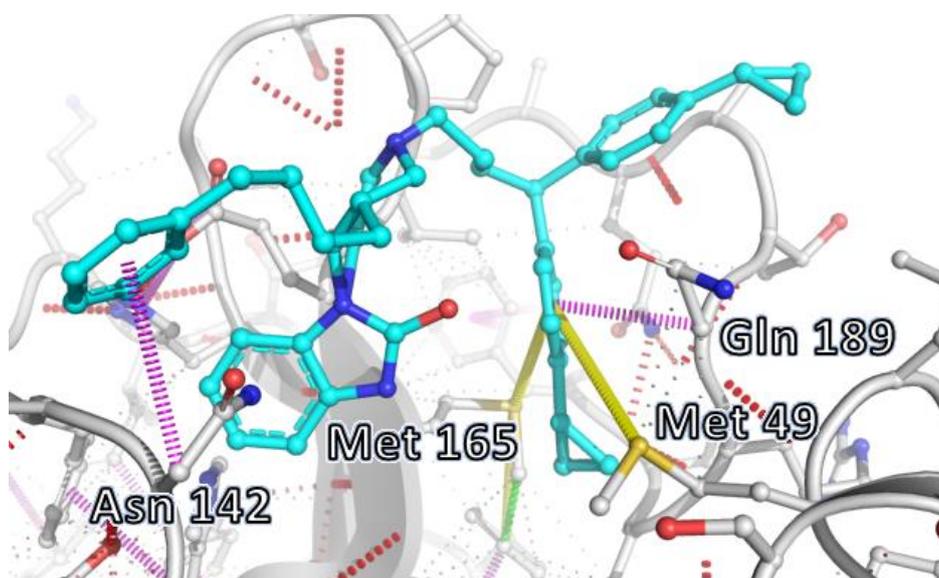


Figura 61. Interacciones intermoleculares ligando AS7

❖ Ligando AS1

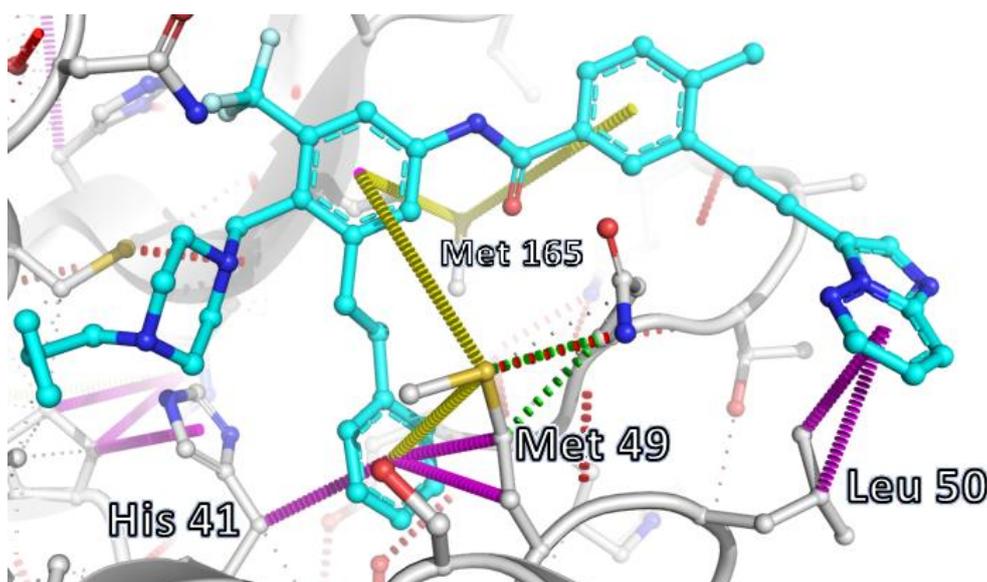


Figura 62. Interacciones intermoleculares ligando AS1

Para el ligando AS7, se determinó que la intromisión de grupos funcionales puede aumentar la inhibición, para este propósito se ha utilizado un grupo fenilo en el mejoramiento ya que su alta densidad atómica deriva en un sistema rico en electrones propicios para la formación de enlaces pi carbono – carbono (**Burgos, 1975**) siendo clave en el aumento de su afinidad con Mpro (**Figura 61**) donde se observa el enlace formado por el grupo funcional introducido, además de otro unido a Gln 189, compensando la ausencia de interacciones polares como tal. Además, el ligando forma dos enlaces de azufre pi con Met 45 y Met 165 a partir de un solo grupo funcional, ayudando a su estabilidad y facilitando su permanencia en el tiempo dentro del ligando.

Por otra parte, el ligando AS1 presenta múltiples enlaces pi entre carbonos y carbono-azufre, e interacciones hidrofóbicas (**Figura 62**) presentando el mayor número de enlaces intermoleculares que sus similares, los cuales se disponen a manera de una cadena que mantiene unido al inhibidor con el sitio activo de la enzima, lo que aparentemente se traduce como una gran actividad biológica del mismo a pesar de la variación de su conformación espacial respecto al ligando base.

CAPÍTULO IV

CONCLUSIONES Y RECOMENDACIONES

4.1 CONCLUSIONES

El estudio general del coronavirus SARS CoV-2 permitió entender la importancia de su enzima Mpro en los procesos de replicación viral, siendo el principal objetivo para el desarrollo de inhibidores con actividad biológica que permitan el tratamiento de la enfermedad COVID-19 provocada por la infección del virus, por otra parte la alta conservación de los aminoácidos que conforman el sitio activo de esta enzima en todos los coronavirus patógenos al ser humano hace posible el desarrollo de inhibidores capaces de combatir a todos estos organismos, presentándose como una opción terapéutica total.

El presente proyecto permitió la identificación de los mejores sitios drogables de la proteasa Mpro, identificados como los residuos Thr 25, Thr 26, Asn 142, Met 49, Met 165, Glu 166 y Gln 189, donde al colocarse un ligando que interaccione con los mismos se evitaría la intromisión del sustrato viral a la diada catalítica formada por His 41 y Cys 145, evitando su replicación.

Se ha establecido una librería de compuestos con actividad inhibitoria obtenida a partir de Proteín Data Bank, de la cual han sido ensayados únicamente aquellas moléculas que no forman un enlace covalente con Cys 145, sin embargo, las estructuras obtenidas de este cribado virtual pueden servir como punto de partida para futuras investigaciones.

Como resultado principal de este proyecto, se ha identificado a los ligandos **MMV1580167** obtenido a partir del cribado virtual y a **AS7** desarrollado mediante un diseño racional, como los mejores inhibidores virtuales de la proteasa Mpro del virus SARS CoV-2, siendo **AS7** el que presentó el más alto nivel de interacción con la enzima tanto en el acoplamiento molecular de sus grupos funcionales en todos los espacios identificados del sitio activo, así como en interacciones intermoleculares hidrofóbicas y enlaces pi múltiples con los aminoácidos identificados como drogables que conforman el mismo, presentando también una conformación espacial similar a N3 dentro de la enzima,

por lo que teóricamente se la acepta como una molécula biológicamente activa para el tratamiento de la infección del virus SARS CoV-2.

4.2 RECOMENDACIONES

- Realizar un estudio de docking o acoplamiento molecular tomando en cuenta los ligandos de la librería de Protein Data Bank que forman un enlace covalente con Cys 145.
- Incluir en el diseño racional de inhibidores, grupos funcionales capaces de unirse covalentemente a Cys 145 para comprobar si el tipo de enlace covalente permite obtener un score de afinidad aún mayor.
- Realizar un ensayo de laboratorio con el desarrollo real de los ligandos **MMV1580167**, **AS7** y **AS1** para ser probados in vivo frente a la enzima Mpro y corroborar su capacidad de inhibición de la misma.

BIBLIOGRAFÍA

- Ali, R.; Shamsah, A.; Shafiul, H.; Ranjit, S.; Ruchi, T.; Yashpal, M.; Kuldeep, D.; Iqbal, Y.; Bonilla-Aldana, K. & Rodríguez-Morales, A (2020). SARS-CoV-2, SARS-CoV, and MERS-CoV: a comparative overview. *Infez. Med.*, 2:174-84
- Alonso Ravelo, F. R. (2017). Estudio bioinformático de los secretomas de *Botrytis cinerea*.
- Andersen, KG, Rambaut, A., Lipkin, WI, Holmes, EC y Garry, RF (2020). El origen proximal del SARS-CoV-2. *Medicina natural* , 26 (4), 450-452.
- Bacha, U.; Barrila, J.; Velázquez-Campoy, A.; Leavitt, S. A.; Freire, E. (2004) Identification of Novel Inhibitors of the SARS Coronavirus Main Protease 3CLpro. *Biochemistry*. (17): 4906-4912.
- Bratanich, A. (2015). MERS-CoV: transmisión y el papel de nuevas especies hospederas. *Revista Argentina de Microbiología*, 47(4), 279.
- Cambridge Crystallographic Data Centre (2019) GOLD User Guide. *A Component of the CSD-Discovery Suite* Registered Charity No 800579
- Centro de Operaciones de Emergencia (COE) Nacional (2020). *Situación Nacional Por Covid-19 INFOGRAFÍA N°217*. Recuperado el 01 de octubre de 2020 de: <https://www.gestionderiesgos.gob.ec/wp-content/uploads/2020/10/INFOGRAFIA-NACIONALCOVID19-COE-NACIONAL-08h00-01102020.pdf>
- Cohen, J., & Kupferschmidt, K. (2020). Strategies shift as coronavirus pandemic looms. *Science*, 367(6481), 962–963. <https://doi.org/10.1126/science.367.6481.962>
- Corman, VM, Muth, D., Niemeyer, D. y Drosten, C. (2018). Hospedadores y fuentes de coronavirus humanos endémicos. En *Avances en la investigación de virus* (Vol. 100, págs. 163-188). Prensa académica.
- Cortés, M (2020) Diagrama Esquemático de un Coronavirus. Recuperado el 30 de septiembre de 2020 de: https://scielo.conicyt.cl/scielo.php?script=sci_arttext&pid=S003498872020000100124

- Burgos, E. M. (1975). Estudio sobre dinámica reticular y fuerzas intermoleculares: Vibraciones y mezcla de modos en el cristal de bifenilo e intensidades Raman de modos reticulares (Doctoral dissertation, Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires).
- Dai, W., Zhang, B., Jiang, X. M., Su, H., Li, J., Zhao, Y., ... & Li, C. (2020). Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease. *Science*, 368(6497), 1331-1335.
- De Wit, E., Van Doremalen, N., Falzarano, D., & Munster, V. J. (2016). SARS and MERS: Recent insights into emerging coronaviruses. *Nature Reviews Microbiology*, 14(8), 523–534. <https://doi.org/10.1038/nrmicro.2016.81>
- Dong E, Du H, Gardner L. (2020) An interactive web-based dashboard to track COVID-19 in real time. *Lancet Inf Dis*. 20(5):533-534. doi: 10.1016/S1473-3099(20)30120-1
- Anand, K., Ziebuhr, J., Wadhwani, P., Mesters, J. R., & Hilgenfeld, R. (2003). Coronavirus main proteinase (3CLpro) Structure: Basis for design of anti-SARS drugs. *Science*, 300(5626), 1763–1767. <https://doi.org/10.1126/science.1085658>
- Christian, M. D., Poutanen, S. M., Loutfy, M. R., Muller, M. P., & Low, D. E. (2004). Severe acute respiratory syndrome. *Clinical Infectious Diseases*, 38(10), 1420–1427. <https://doi.org/10.1086/420743>
- Cohen, J., & Kupferschmidt, K. (2020). Strategies shift as coronavirus pandemic looms. *Science*, 367(6481), 962–963. <https://doi.org/10.1126/science.367.6481.962>
- De Wit, E., Van Doremalen, N., Falzarano, D., & Munster, V. J. (2016). SARS and MERS: Recent insights into emerging coronaviruses. *Nature Reviews Microbiology*, 14(8), 523–534. <https://doi.org/10.1038/nrmicro.2016.81>
- Grigera, J. R. (1989). La Interacción Hidrofóbica en Sistemas de Interés Biológico. 11. Su Evaluación y algunas Aplicaciones. *Acta Farm. Bonaerense*, 8(2), 107-24.
- Hao, X.; Liang, Z.; Jiabin, D.; Jiakuan, P.; Hongxia, D.; Xin, Z.; Taiwen, L. & Qianming, Chen. (2020) High expression of ACE2 receptor of 2019nCoV on the epithelial cells of oral mucosa. *Int. J. Oral Sci.*, 12:8
- Hilgenfeld, R. (2014). From SARS to MERS: crystallographic studies on coronaviral proteases enable antiviral drug design. *The FEBS journal*, 281(18), 4085-4096.

- Jin, Z., Du, X., Xu, Y., Deng, Y., Liu, M., Zhao, Y., ... & Duan, Y. (2020). Structure of M pro from SARS-CoV-2 and discovery of its inhibitors. *Nature*, 1-5.
- Jemmis, E. D., Pathak, B., King, R. B., & Schaefer III, H. F. (2006). Bond length and bond multiplicity: σ -bond prevents short π -bonds. *Chemical communications*, (20), 2164-2166.
- Jin, Z., Du, X., Xu, Y., Deng, Y., Liu, M., Zhao, Y., ... & Duan, Y. (2020). Structure of M pro from SARS-CoV-2 and discovery of its inhibitors. *Nature*, 1-5.
- Khaerunnisa, S., Kurniawan, H., Awaluddin, R., Suhartati, S., & Soetjipto, S. (2020). Potential inhibitor of COVID-19 main protease (Mpro) from several medicinal plant compounds by molecular docking study. *Prepr. doi10, 20944*, 1-14.
- Marinho, EM, de Andrade Neto, JB, Silva, J., da Silva, CR, Cavalcanti, BC, Marinho, ES, & Júnior, HVN (2020). Cribado virtual basado en el acoplamiento molecular de posibles inhibidores de la proteasa principal Covid-19. *Patogenia microbiana* , 148 , 104365.
- Mas Benavente, J. M. (2001). Comparación computacional de estructuras de proteínas aplicación al estudio de un inhibidor de carboxipeptidasa como agente antitumoral. *Universitat Autònoma de Barcelona*,.
- Mousavizadeh, L., & Ghasemi, S. (2020). Genotype and phenotype of COVID-19: Their roles in pathogenesis. *Journal of Microbiology, Immunology and Infection*.
- Muramatsu, T., Takemoto, C., Kim, YT, Wang, H., Nishii, W., Terada, T., ... y Yokoyama, S. (2016). La proteasa SARS-CoV 3CL escinde su sitio de autoprosesamiento C-terminal mediante una nueva cooperatividad del subsitio. *Actas de la Academia Nacional de Ciencias* , 113 (46), 12997-13002.
- Naganathan, A. N., & Orozco, M. (2011). The native ensemble and folding of a protein molten-globule: functional consequence of downhill folding. *Journal of the American Chemical Society*, 133(31), 12154-12161.
- O'Hagan, D. (2008). Understanding organofluorine chemistry. An introduction to the C–F bond. *Chemical Society Reviews*, 37(2), 308-319.
- Paules, C. I., Marston, H. D., & Fauci, A. S. (2020). Coronavirus infections—more than just the common cold. *Jama*, 323(8), 707-708.

- Pérez, L. J., & de Arce, H. D. (2008). Optimization Of Scoring Function Docking For Antiviral Drug Design Against Coronavirus Genus. *Rev. Salud Anim.* Vol, 30(3), 152-159.
- Rzychon, Malgorzata & Chmiel, Dorota & stec-niemczyk, Justyna. (2004). Modes of inhibition of cysteine proteases Circlred white star. *Acta biochimica Polonica.* 51. 861-73
- Sun, D., Wang, X., Wei, S., Chen, J. y Feng, L. (2015). Epidemiología y vacuna del virus de la diarrea epidémica porcina en China: una mini revisión. *Revista de ciencia médica veterinaria* , 15-0446.
- Ullrich, S., & Nitsche, C. (2020). The SARS-CoV-2 main protease as drug target. *Bioorganic & Medicinal Chemistry Letters*, 127377.
- Vivas-Reyes, R., Ruiz, J. L., & Caraballo, R. (2009). Estudio teórico del virus del síndrome respiratorio agudo severo (sars) a través del uso de métodos de acoplamiento molecular. *Revista Colombiana de Química*, 38(3), 327-339.
- Yang, H., Yang, M., Ding, Y., Liu, Y., Lou, Z., Zhou, Z., ... & Gao, G. F. (2003). The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor. *Proceedings of the National Academy of Sciences*, 100(23), 13190-13195.
- Yang, H., Xie, W., Xue, X., Yang, K., Ma, J., Liang, W., ... & Hilgenfeld, R. (2005). Design of wide-spectrum inhibitors targeting coronavirus main proteases. *PLoS Biol*, 3(10), e324.
- Zambrano, S. R., & Loaiza, E. J. (2003). Neumonía atípica-síndrome respiratorio agudo severo (SARS). *Medicina*, 8(3), 236-240.
- X. Liu and X.-J. Wang (2020) “Potential inhibitors against 2019-nCoV coronavirus M protease from clinically approved medicines,” *J. Genet. Genomics*, doi: 10.1016/j.jgg.2020.02.001.
- Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 March 2020
doi:10.20944/preprints202003.0226.v1

- approved medicines,” *J. Genet. Genomics*, 2020, doi: 10.1016/j.jgg.2020.02.001.
- Zhang, L., Lin, D., Kusov, Y., Nian, Y., Ma, Q., Wang, J., ... y De Wilde, A. (2020). α -cetoamidas como inhibidores de amplio espectro de la replicación de coronavirus y enterovirus: evaluación de la actividad, síntesis y diseño basado en estructuras. *Revista de química médica* .
- Zhou, P., Fan, H., Lan, T., Yang, X. Lou, Shi, W. F., Zhang, W., ... Ma, J. Y. (2018). Fatal swine acute diarrhoea syndrome caused by an HKU2-related coronavirus of bat origin. *Nature*, 556(7700), 255–259. <https://doi.org/10.1038/s41586-018-0010-9>
- Zhou, P., Yang, X. Lou, Wang, X. G., Hu, B., Zhang, L., Zhang, W., ... Shi, Z. L. (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, 579(7798), 270–273. <https://doi.org/10.1038/s41586-020-2012-7>
- Zhu, X.; Liu, Q.; Du, L. & Jiang, S. (2013) Receptor-binding domain as a target for developing severe acute respiratory syndrome vaccines. *J. Thorac. Dis.*, 5(S2):S142

ANEXOS

ANEXO 1: SECUENCIAS DE VIRUS HOMÓLOGOS HUMANOS PARA ALINEAMIENTO

>SARS-CoV-2

SGFRKMAFPSPGKVEGCMVQVTCGTTTLNGLWLDVVVYCPRHVICTSEDMLNPNYEDLLIRKSNH
NFLVQAGNVQLRVI GHSMQNCVLKLVDTANPKTPKYKFVRIQPGQTFVSLACYNGSPSGVYQC
AMRPNFTTIKGSFLNGSCGSGVGFNIDYDCVSFCYMHMELPTGVHAGTDLEGNFYGFVDRQTAQ
AAGTDTTITVNVLAWLYAAVINGDRWFLNRFTTTLNDFNLVAMKYNYEPLTQDHVDILGPLSAQ
TGI AVLDMCASLKELLQNGMNGRTILGSALLEDEFTPFVVRQCSGVTFQ

>SARS-CoV

SGFRKMAFPSPGKVEGCMVQVTCGTTTLNGLWLDVTVYCPRHVICTAEDMLNPNYEDLLIRKSNH
SFLVQAGNVQLRVI GHSMQNCVLLRKLKVDTSNPKTPKYKFVRIQPGQTFVSLACYNGSPSGVYQC
AMRPNHTTIKGSFLNGSCGSGVGFNIDYDCVSFCYMHMELPTGVHAGTDLEGKFYGFVDRQTAQ
AAGTDTTITLNVLAWLYAAVINGDRWFLNRFTTTLNDFNLVAMKYNYEPLTQDHVDILGPLSAQ
TGI AVLDMCAALKELLQNGMNGRTILGSTILEDEFTPFVVRQCSGVTFQ

>MERS-CoV

SGLVKMSHPSGAVEACMVQVTCGSMTLNGLWLDNYVWCPRHVMCPADQLNDPNYDGLLI SMTNH
SFTVQKHVGAPANLRVIGHAMQGTLLKLTVDSPNPNTPAYTFTTVKPGASFVSLACYNGRPTGT
FTVVMRPNHTIKGSFLCGSCGSGVGTKEGNVINFCYMHQMELANGHTTGSSFDGNMYGAFQDRQ
VHQAQLTDKYCTVNIVAWLYAAI LNNGCTWFVKPSRTSVVSFNEWAMANQFTEFIGTHALD
MLAVKTGVSVEQLLSAIQQLNQGFQKQILGSSMLEDEFTPEDVNMQIMGVVMQ

>HKU1

SGIVKMSVPTSKIEPCLVSVTYGSMTLNGLWLDKVVYCPRHVICLSSNMNEPDYSALLCRVTLG
DFTIMSGRMSLTVVSYQMQGCQLVLTVSLQNPYTPKYTFGVVKPGETFTVLAAYNGRPQGAFHV
TMRSSYTIKGSFLCGSCGSGVGYVLTGDSVKFVYMHQLELSTGCHTGTDFNGNFYGPYRDAQVVQ
LPVKDYVQTVNVIAWLYAAI LNCAWFVQNDVCSIEDFNWAMTNGFSQVKADLVLDALASMTG
VSIETLLAAIKRLYMGFQGRQILGSCTFEDELAPSDVYQQLAGVKLQ

>NL63

SGLKKMAQPSGCVRECVVRVYCYGSTVLNGVWLGDTVTCPRHVIAPSTTVLIDYDHAYSTMRLHN
FSVSHNGVFLGVVGTMHGSVLRIVKVSQSNVHTPKHVFKTLKPGDSFNILACYEGIASGVFGVN
LRTNFTTIKGSFINGACGSPGYNVRNDGTVEFCYLHQIELGSGAHVGSDFGTGSVYGNFDDQPSLQ
VESANLMLSDNVVAFLYAALLNGCRWWLCSTRVNVDFNEWAMANGYTSVSSVECYSILAATG
VSVEQLLASIQHLHEGFGGKNILGYSSLCDEFTLAEVVKQMYGVNLQ

> 229E

AGLRKMAQPSGFVEKCVVRVCYGNLTVLNGLWLGDIVYCPRHVIASNTTSAIDYDHEYSIMRLHN
FSIISGTAFLGVGATMHGVTLKI KVSQTNMHTPRHSFRTLKSGEGFNILACYDGCAQGVFGVN
MRTNWTIRGSFINGACGSPGYNLKNGEVEFVYMHQIELGSGSHVGS SFDGVMYGGFEDQPNLQV
ESANQMLTVNVVAFLYAAI LNGCTWWLKGEKLFVEHYNEWAQANGFTAMNGEDAFS ILAAKTGV
CVERLLHAIQVLNNGFGGKQILGYSSLNDEFSINEVVKQMFVNLQ

>HRV-14

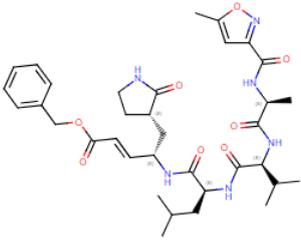
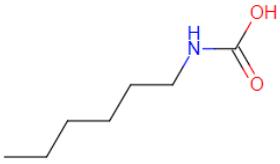
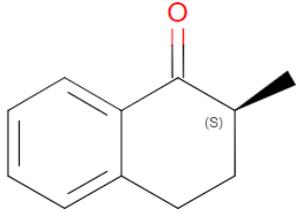
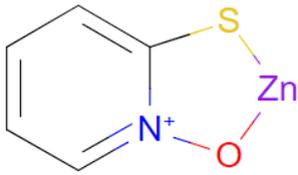
GLGPRYGGIYTSNVKIMNYHLMTPEDHHNLIAPYPNRDLAIVSTGGHGAETIPHCNCTSGVYYS
TYRKYYP IICEKPTNIWIEGNPYYP SRFQAGVMKGVGPAEPGDCGILRCIHGPIGLLTAGGS
GYVCFADIRQLECIAEEQGLSDYITGLGRAFGVGF TDQISTKVTELQEVAKDFLTTKVLSKVVK
MVSALV IICRNHDDLVTVTATLALLGCDGSPWRFLKMYISKHFQVPYIERQANDGWRKFN DAC
NAAKLEWIANKISKLIEWIKNKVLPQAKEKLEFC SKLKQLDILERQITTMHISNPTQEKREQL
FNNVLWLEQMSQKFAPLYAVESK

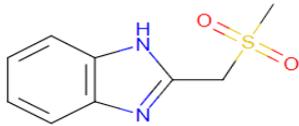
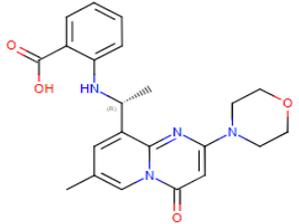
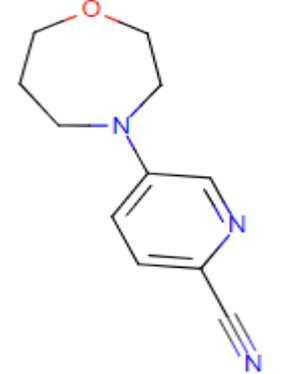
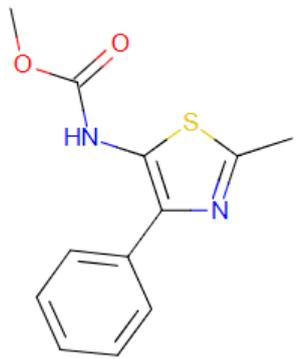
>Hepatitis A

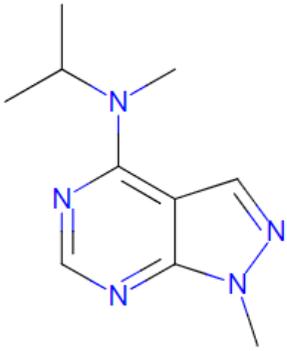
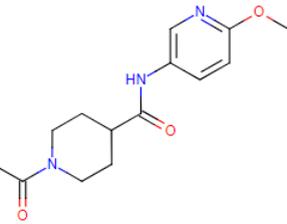
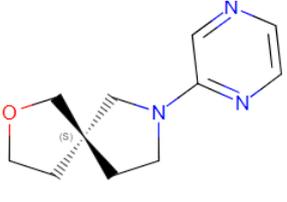
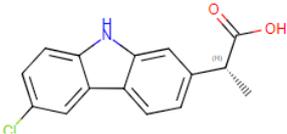
GLVRKNLVQFGVGEKNGCVRWVMNALGVKDDWLLVPSHAYKFEKDYEMMEFYFNRRGGTYYSISA
GNVVIQSLDVGFQDVVLMKVPTIPKFRDITQHFIKKGDVPRALNRLATLVTTVNGTPMLISEGP
LKMEEKATYVHKNDGTTVDLTVDQAWRGKGEGLPGMCGALVSSNQSIQNAILGIHVAGGNSI
LVAKLVTQEMFQNI DKKIESQRIMKVEFTQCSMNVVSKTLFRKSPIYHHIDKTMINFPAAMPFS
KAEIDPMAVMLSKYSLPIVEEPEDYKEASIF YQNKIVGKTQLVDDFLDLDMAITGAPGIDAINM
DSSPGFPYVQEKLTRDLIWLDENGLLLGVHPRLAQRI LFNTVMMENCSDL DVVF

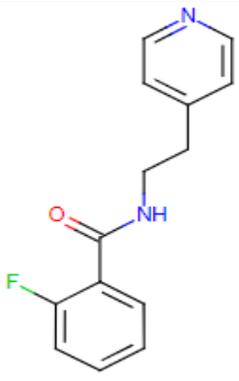
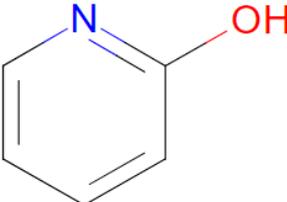
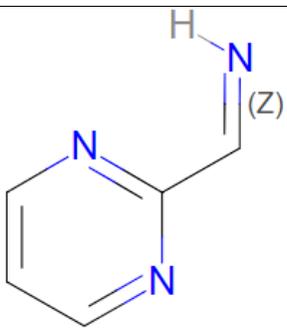
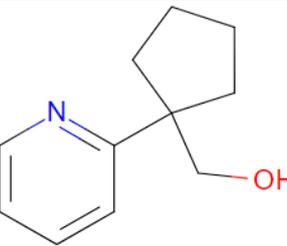
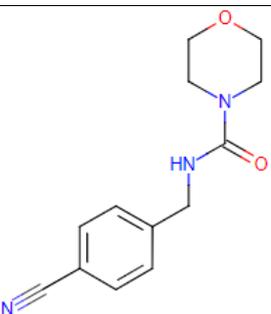
**ANEXO 2: LIBRERÍA DE MOLÉCULAS CON ACTIVIDAD
INHIBITORIA DE M_{pro} OBTENIDAS DE PROTEIN DATA BANK**

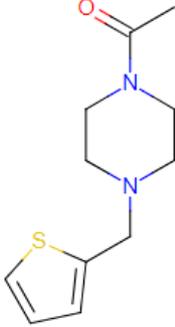
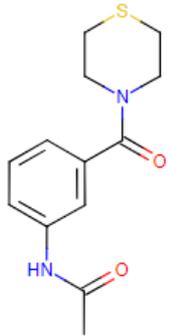
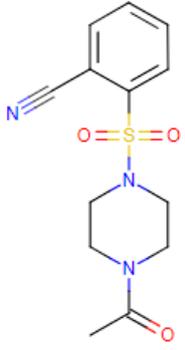
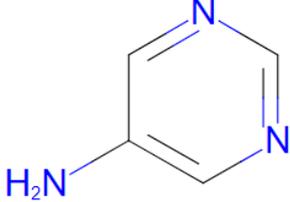
Tabla 17. Librería de ligandos obtenidos de Pdbe Fold

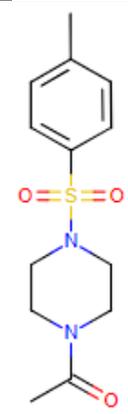
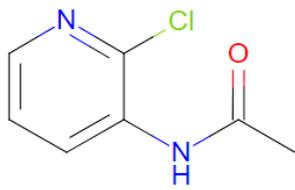
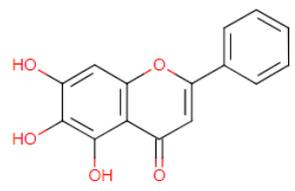
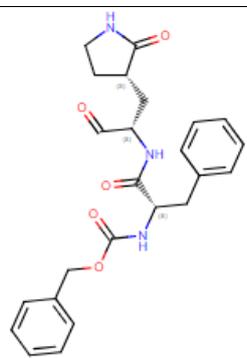
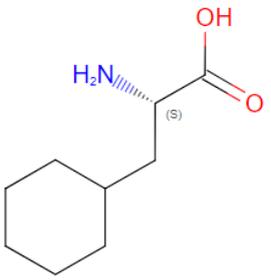
Organismo	Código PDB	Ligando	Código PDB	Nomenclatura SMILES	Estructura química
SARS CoV-2	7BQY	n-[(5-methylisoxazol-3-yl)carbonyl]alanyl-l-valyl-n-1-((1r,2z)-4-(benzyloxy)-4-oxo-1-[(3r)-2-oxopyrrolidin-3-yl]methyl}but-2-enyl)-l-leucinamide	N3	<chem>CC(C)C[C@H](NC(=O)[C@@H](NC(=O)[C@H](C)NC(=O)c1cc(C)on1)C(C)C(=O)N[C@@H](C[C@@H]2CCNC2=O)\C=C/C(=O)OCc3ccccc3</chem>	
	7BUY	hexylcarbamic acid (Carmofur)	JRY	<chem>CCCCCCNC(=O)O</chem>	
	6YNQ	(2~{S})-2-metil-3,4-dihidro-2~{H}-naftalen-1-ona	P6N	<chem>C[C@H]1CCc2ccccc2C1=O</chem>	
	6YT8	9-oxa-7-tia-1-azonia-8 \$ 1 ^ {2} -zincabicyclo [4.3.0] nona-1,3,5-trieno	PK8	<chem>O1 [Zn] Sc2cccc [n +] 12</chem>	

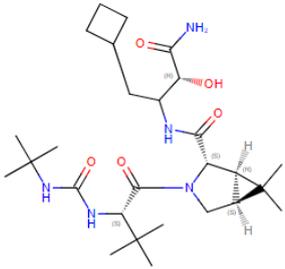
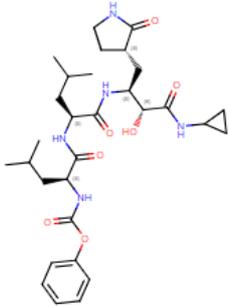
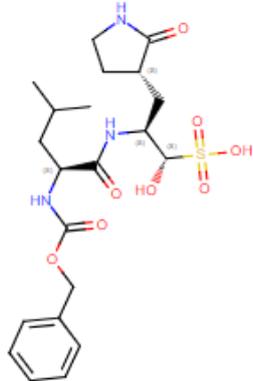
	5RFD	2 - [(metilsulfonil) metil] -1H- bencimidazol	T6J	C [S] (= O) (= O) Cc1 [nH] c2ccccc2n1	
	6YVF	AZD-6482	A82	C [C @@ H] (Nc1ccccc1C (O) = O) C2 = CC (= CN3C (= O) C = C (N = C23) N4CCOCC4) C	
SARS CoV-2	5RF6	5- (1,4-oxazepan- 4-il) piridina-2- carbonitrilo	NTG	N # Cc1ccc (cn1) N2CCCCOCC2	
	5RFC	metil (2-metil-4- fenil-1,3-tiazol-5- il) carbamato	K1Y	COC (= O) Nc1sc (C) nc1c2ccccc2	

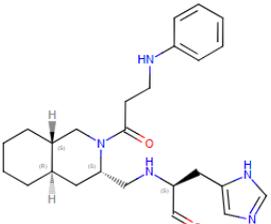
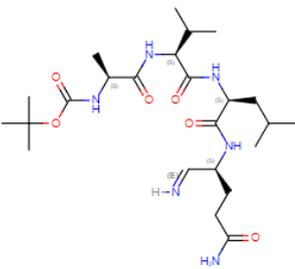
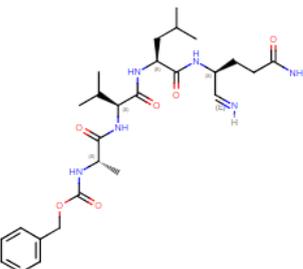
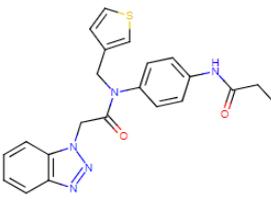
	5RGR	N, 1-dimetil-N-(propan-2-il) -1H-pirazolo [3,4-d] pirimidin-4-amina	K1G	CC (C) N (C) c1ncnc2n (C) ncc12	
	5RHC	(E) -1- (1H-imidazol-2-il) metanimina	USA	N = Cc1 [nH] ccn1	
	5RHE	1-acetil-N- (6-metoxipiridin-3-il) piperidina-4-carboxamida	UPD	COc1ccc (NC (=O) C2CCN (CC2) C (C) = O) cn1	
	5RGJ	(5S) -7- (pirazin-2-il) -2-oxa-7-azaspiro [4.4] nonano	U0S	C1C [C @] 2 (CCN (C2) c3cncn3) CO1	
SARS CoV-2	5RH4	Ácido (2R) -2- (6-cloro-9H-carbazol-2-il) propanoico	UHG	C [C @@ H] (C (O) = O) c1ccc2c ([nH] c3ccc (Cl) cc23) c1	

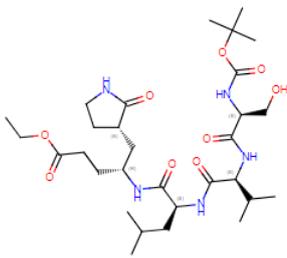
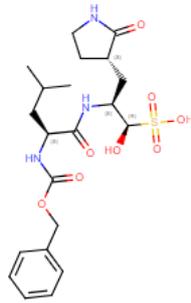
	5RGK	2-fluoro-N- [2-(piridin-4-il) etil] benzamida	U0V	<chem>Fc1ccccc1C(=O)NCCc2ccncc2</chem>	
	5RF4	piridin-2-ol	T5Y	<chem>Oc1ccccn1</chem>	
	5RHH BVB	(E) -1- (pirimidin-2-il) metanimina	USD	<chem>N=Cc1ncccn1</chem>	
	5RFO	[1- (piridin-2-il) ciclopentil] metanol	T5D	<chem>OCC1(CCCC1)c2ccccn2</chem>	
SARS CoV-2	5RFE	N - [(4-cianofenil) metil] morfolina-4-carboxamida	JGG	<chem>O=C(NCc1ccc(C#N)cc1)N2CCOCC2</chem>	
	5RGI	N'-ciclopropil-N-metil-N - [(5-metil-1,2-oxazol-3-il) metil] urea	U0P	<chem>CN(Cc1cc(C)nn1)C(=O)NC2CC2</chem>	

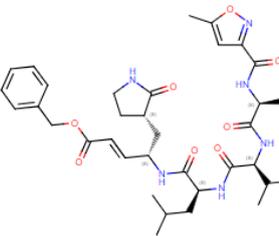
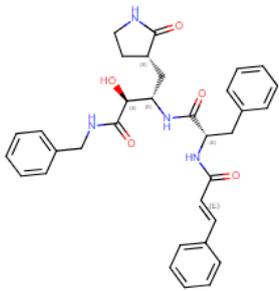
				<chem>on1) C (= O) NC2CC2</chem>	
	5RHA	1- {4 - [(tiofen-2-il) metil] piperazin-1-il} etan-1-ona	T8M	<chem>CC (= O) N1CCN (CC1) Cc2sccc2</chem>	
	5REV	N- [3- (tiomorfolina-4-carbonil) fenil] acetamida	T4J	<chem>CC (= O) Nc1cccc (c1) C (= O) N2CCSCC2</chem>	
	5REU	2 - [(4-acetilpiperazin-1-il) sulfonil] benzonitrilo	T4D	<chem>CC (= O) N1CCN (CC1) [S] (= O) (= O) c2ccccc2C # N</chem>	
SARS CoV-2	5RF3	pirimidin-5-amina	T5V	<chem>Nc1cncnc1</chem>	

	5RGN	1- {4 - [(4- metilfenil) sulfonil] piperazin-1-il} etan-1-ona	U1A	CC (= O) N1CCN (CC1) [S] (= O) (= O) c2ccc (C) cc2	
	5RFZ	N- (2- cloropiridin-3-il) acetamida	T8V	CC (= O) Nc1cccnc1Cl	
	6M2N	5,6,7-trihidroxi-2- fenil-4H-cromen- 4-ona	3WL	Oc1cc2OC (= CC (= O) c2c (O) c1O) c3ccccc3	
	6XBN	PRD_002393	UAW 247	O=C[C@H](C[C @@H]1CCNC1= O)NC(=O)[C@H] (Cc2ccccc2)NC(= O)OCc3ccccc3	
SARS CoV	3ATW	ácido 2-amino-3- ciclohexil- propionico	ALC	N [C @@ H] (CC1CCCCC1) C (O) = O	

SARS CoV	6WNP	Boceprevir	U5G	<chem>CC(C)(C)NC(=O)N[C@H](C(=O)N1C[C@H]2[C@@H]([C@H]1)C(=O)N[C@@H](CC3CCCC3)[C@@H](O)C(N)=O)C2(C)C(C)(C)C</chem>	
	6XBI	PRD_002394	UAW 248	<chem>CC(C)C[C@H](NC(=O)Oc1cccc1)C(=O)N[C@@H](CC(C)C)C(=O)N[C@@H](C[C@@H]2CCNC2=O)[C@@H](O)C(=O)NC3CC3</chem>	
	6WTT	(1S, 2S) -2 - ({N - [(benciloxi) carbonil] -L-leucil} amino) -1-hidroxi-3 - [(3S) - 2-oxopirrolidin-3-il] propano-1-sulfónico ácido	K36	<chem>CC(C)C[C@H](NC(=O)OCc1cccc1)C(=O)N[C@@H](C[C@@H]2CCNC2=O)[C@@H](O)[S](=O)(=O)O</chem>	

SARS CoV	5C50	(2S)-3-(1H-imidazol-5-yl)-2-((3S,4aR,8aS)-2-(N-phenyl-beta-alanyl)decahydroisoquinolin-3-yl)methylamino)propanal	SDJ	<chem>O=C[C@H](Cc1[nH]cnc1)NC[C@@H]2C[C@H]3CCCC[C@@H]3CN2C(=O)CCNc4cccc4</chem>	
	3VB4	Inhibidor de B4Z	PRD_000910	<chem>CC(C)C[C@H](NC(=O)[C@@H](NC(=O)[C@H](C)NC(=O)OC(C)(C)C(C)C(=O)N[C@@H](CCC(N)=O)C=N</chem>	
	3VB5	Inhibidor de C4Z	PRD_000911	<chem>CC(C)C[C@H](NC(=O)[C@@H](NC(=O)[C@H](C)NC(=O)OCc1cccc1)C(C)C(=O)N[C@@H](CCC(N)=O)C=N</chem>	
MERS CoV	4YLU	N-{4-[(1H-benzotriazol-1-yl)acetyl](thiophen-3-yl)methyl}amino]p	R30	<chem>CCC(=O)Nc1ccc(cc1)N(Cc2csc2)C(=O)Cn3nnc4ccccc34</chem>	

		henyl}propanamide			
	4RSP	N-(tert-butoxycarbonyl)-L-seryl-L-valyl-N-{(2S)-5-ethoxy-5-oxo-1-[(3S)-2-oxopyrrolidin-3-yl]pentan-2-yl}-L-leucinamide	PRD_002174	<chem>CCOC(=O)CC[C@@H](C[C@@H]1CCNC1=O)NC(=O)[C@H](CC(C)C)NC(=O)[C@@H](NC(=O)[C@H](CO)NC(=O)OC(C)(C)C)C(C)C</chem>	
MERS CoV	5WKJ	(1R,2S)-2-({N-[(benzyloxy)carbonyl]-L-leucyl}amino)-1-hydroxy-3-[(3S)-2-oxopyrrolidin-3-yl]propane-1-sulfonic acid	B1S	<chem>CC(C)C[C@H](NC(=O)OCc1ccccc1)C(=O)N[C@@H](C[C@@H]2CNC2=O)[C@H](O)[S](O)(=O)=O</chem>	
	5WKM	(1R,2S)-2-({[N-({[1-(tert-butoxycarbonyl)-4-ethylpiperidin-4-yl]oxy}carbonyl)-L-leucyl]amino}-1-hydroxy-3-[(3S)-2-oxopyrrolidin-3-	B6Y	<chem>CCC1(CCN(CC1)C(=O)OC(C)(C)C)OC(=O)N[C@@H](CC(C)C)C(=O)N[C@@H](C[C@@H]2CCNC2=O)[C@H](O)[S](O)(=O)=O</chem>	92

		yl]propane-1-sulfonic acid			
	5WKK	(1S,2S)-2-[(N- {[2-(3-chlorophenyl)ethoxy]carbonyl}-L-leucyl)amino]-1-hydroxy-3-[(3S)-2-oxopyrrolidin-3-yl]propane-1-sulfonic acid	AW4	<chem>CC(C)C[C@H](NC(=O)OCCc1cccc(Cl)c1)C(=O)N[C@@H](C[C@@H]2CCNC2=O)[C@@H](O)[S](O)(=O)=O</chem>	
HKU1	3D23	n-[(5-methylisoxazol-3-yl)carbonyl]alanyl-L-valyl-n-1-((1r,2z)-4-(benzyloxy)-4-oxo-1-[(3r)-2-oxopyrrolidin-3-yl]methyl}but-2-enyl)-L-leucinamide	PRD_0 02214	<chem>CC(C)C[C@H](NC(=O)[C@@H](NC(=O)[C@H](C)NC(=O)c1cc(C)on1)C(C)C(=O)N[C@@H](C[C@@H]2CCNC2=O)\C=C/C(=O)OCc3ccccc3</chem>	
NL63	6FV2	(S)-N-benzyl-3-((S)-2-cinnamamido-3-phenylpropanamido)-2-oxo-4-((S)-2-oxopyrrolidin-3-yl)butanamide	D03	<chem>O[C@@H]([C@H](C[C@@H]1CNC1=O)NC(=O)[C@H](Cc2ccccc2)NC(=O)/C=C/c3ccccc3)C(=O)NCc4ccccc4</chem>	

	5NHO	<p>~{tert}-butyl ~{N}- [(2~{S},3~{R})- 3-oxidanyl-4- oxidanylidene-1- [(3~{S})-2- oxidanylidene-pyrro- lidin-3-yl]-4- [(phenylmethyl)a- mino]butan-2- yl]carbamate</p>	8X8	<chem>CC(C)(C)OC(=O)N[C@@H](C[C@@H]1CCNC1=O)[C@@H](O)C(=O)NCC2=CC=CC=C2</chem>	
	6FV1	<p>(2~{S})-4- methyl-~{N}- [(2~{S},3~{R})- 3-oxidanyl-4- oxidanylidene-1- [(3~{S})-2- oxidanylidene-pyrro- lidin-3-yl]-4- [(phenylmethyl)a- mino]butan-2-yl]- 2-[[(~{E})-3- phenylprop-2- enoyl]amino]pent- anamide</p>	E8E	<p>(2~{S})-4- methyl-~{N}- [(2~{S},3~{R})- 3-oxidanyl-4- oxidanylidene-1- [(3~{S})-2- oxidanylidene-pyrro- lidin-3-yl]-4- [(phenylmethyl)a- mino]butan-2-yl]- 2-[[(~{E})-3- phenylprop-2- enoyl]amino]pent- anamide</p>	
229E	2ZU2	<p>(4s)-2-methyl- 2,4-pentanediol</p>	MPD	<chem>C[C@H](O)CC(C)(C)O</chem>	

ANEXO 3: INTERACCIÓN DE INHIBIDORES DEL VIRUS SARS CoV-2 Y HOMÓLOGOS HUMANOS CON LOS AMINÁCIDOS DEL SITIO ACTIVO DE LA PROTEASA Mpro

Tabla 18. Aminoácidos de Mpro que interaccionan con los ligandos

		Aminoácidos / Posición																				
Proteasa	Inhibidor	Ala	Arg	Asn	Asp	Cys	Gln	Glu	Gly	His	Ile	Leu	Lys	Met	Phe	Pro	Ser	Thr	Trp	Tyr	Val	
SARS CoV-2 P0TD1	N3	191		142	187	145	189 192	166	143	41 163 164 172		27 141		49 165	140	168	144	23 190				
	Carmofur					145			143	41 164				49 165			144					
	P6N			142		145		166	143	163 164		141			140		144					
	PK8			142		145			143	41 164		27		165				25 26				
	T6J	94		95	33 34												96					
	A82								79	80	78		88 90					81				35
	NTG						44		166		41				49			46	45			

									164				165				25			
	K1Y							79	80	78		88 90				81				
	K1G		105	228 231	176			178						103	241					104
	USA			142		145			143	41		27				144				
	UPD			142		145	189		143	41						46 144				
	U0S											12 97			96 99		98			
	UHG								79	80			88 90			81				
	U0V			142		145		166		163		141			140					
	T5Y	94		95	33												98			
	USD			142		145			143	41		27				144				
	T5D								138 170	172			13 7							171
	JGG		188	142	187		189		41 164				49 165							

SARS CoV-2 P0TD1	U0P			142		145		166	143	163 164		141			140		144	25 26			
	T8M		188	142		145	189		143	41 163 164				49 165			144				
	T4J			142		145			143	41		27					144	24 25 26			
	T4D			142		44 145			143	41				49			46 144	25 45			
	T5V			142		145		166		163		141			140		144				
	U1A			142		145			143	41		27					144	25 26			
	T8V			142		145			143	41 164		27					144	26			
	3WL			142	187	44 145		166	143	41 164		141		49 165							
	UAW247		188	142	187	145	189	166	143	41 163		141		49 165	140		144				

										164											
										172											
SARS CoV: P0C6X7	ALC y HSV			142	187	145		166	143	41 163 164		141		49 165	140		144			54	
	U5G		188	142	187	145	189 192	166	143	41 164		27 141 167		49 165		168	144	26 190		54	
	UAW248	191		142		145	189 192	166	143	41 163 164 172				165	140	168	144	190			
	K36	191		142	187	145	189	166 192	143	41 163 172		141		49 165	140	168	144	190			
	SDJ	46		142	187	44 145	189	166	143	41 163 164	188	141		49 165	140		144	25 45			

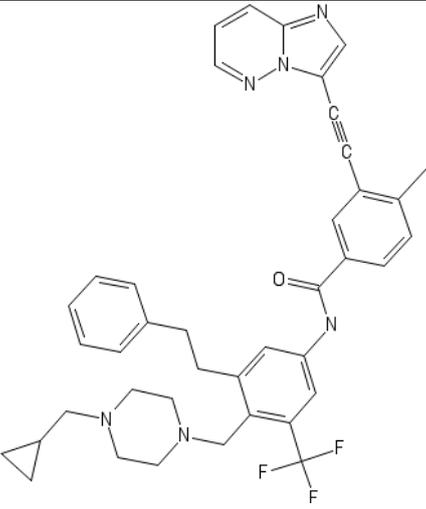
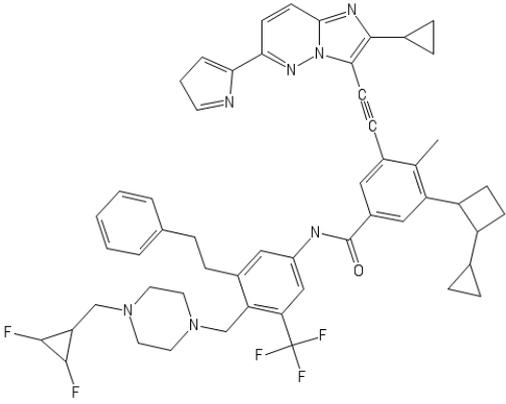
	PRD_000 911	191				145	192	166	143	163 164 172		141 167		165	140	168		190			
	PRD_000 910	191				145		166	143	41 163 164 172		27 141		49 165	140	168	144	190			
MERS: A0A023 YA54	R30	46			190	44 145 148	167 192	169		41 166		49 144	191	25 168	143	45	24 147			54	
	PRD_002 174					145 148	167 192 195	169	146	41 166 175 194		144	191	168	143		147				193
	GC376				190	145 148	167 192	169	146	41 166 175		49 144	191	168	143		147				
	B6Y				190	145 148	167	169	146	41 166		49 144	191	168	143		147				

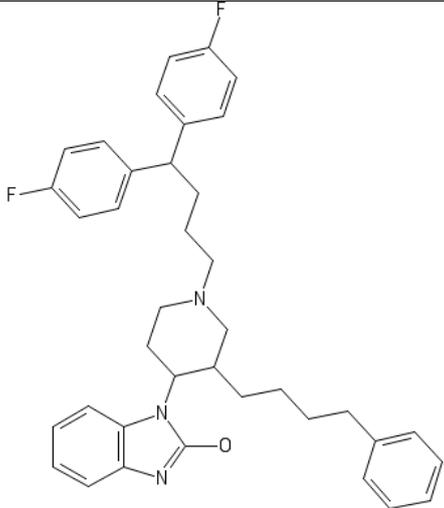
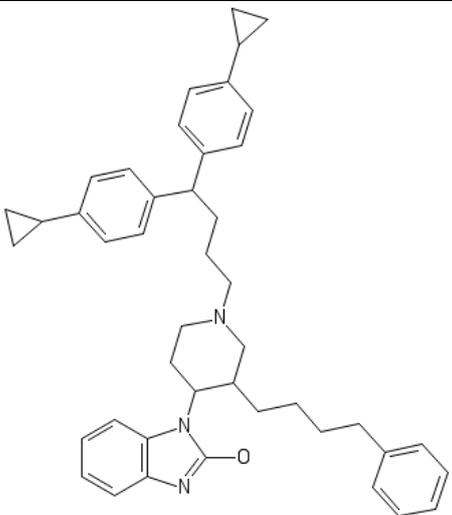
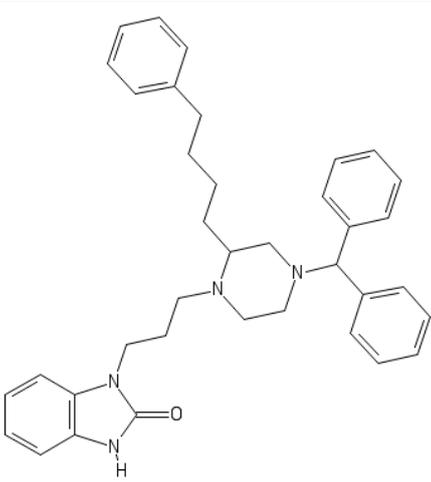
										175											
	GC813	171			190	145 148	167 192	169	146	41 166 175 194		49 144		168	143		147				193
HKU1: P0C6U5	PRD_002 214					142 145	164 189 192	166	143	41 163 172		27 141 167		25 73	140	96	24 144	26		97	190 191
NL63: P0C6U5	D03	143		141	187	144	164 188 192		142	41 163 172	140 165				139	189	190				26
	8X8	143		141	187	144	164 188	166	142 168	41 163 172	165				139	189		47			26
	ESE	143		141	187	144	164	166	142 168	41 163 172	51 140 165	191			139	189	190				26

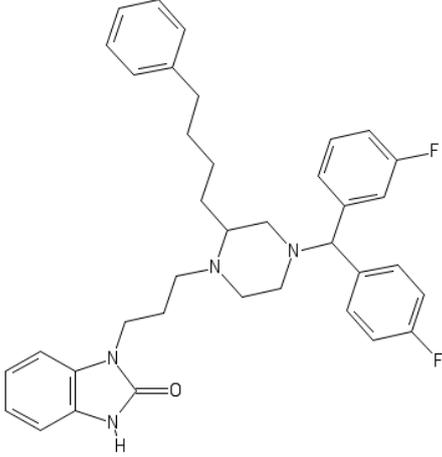
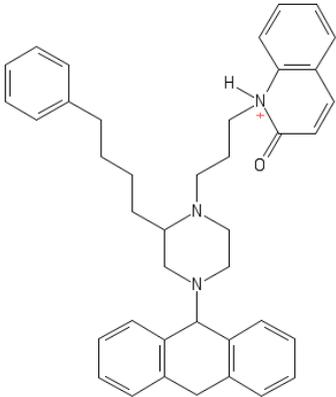
229E: POC6U2																					
MPD		143		141		144	163	165		162	140 164				139						

**ANEXO 4: MOLÉCULAS CON ACTIVIDAD INHIBITORIA
DISEÑADAS**

Tabla 19. Estructura 2D de ligandos diseñados

Nombre	Código de Sonrisas Isoméricas	Imagen
AS1	<chem>Cc5ccc(C(=O)Nc4cc(CCC1c cccc1c(CN3CCN(CC2CC2)CC3) c (C(F)(F)F)c4)cc5C#Cc6nc7cc cnn67</chem>	
AS2	<chem>Cc9c(C#Cc3c(C1CC1)nc4ccc (/C2=C/C/C=N2) nn34)cc(C(=O)Nc8cc(CCc5ccc cc5)c(CN7CCN(CC6C(F) C6F)CC7)c (C(F)(F)F)c8)cc9C%10CCC% 10C%11CC%11</chem>	

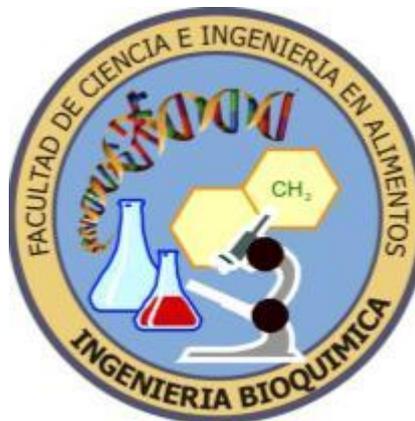
<p>AS6</p>	<p>Oc2nc1cccc1n2C5CCN (CCCC(c3ccc(F)cc3)c4ccc (F)cc4)CC5CCCCc6cccc6</p>	
<p>AS7</p>	<p>Oc2nc1cccc1n2C7CCN (CCCC(c4ccc(C3CC3)cc4) c6ccc(C5CC5)cc6)CC7CCCC c8cccc8</p>	
<p>AS3</p>	<p>O=c2[nH]c1cccc1n2CCCN5 CCN(C(c3ccccc3)c4ccccc4) CC5CCCCc6cccc6</p>	

<p>AS4</p>	<chem>O=c2[nH]c1cccc1n2CCCN5CCN(C(c3ccc(F)cc3)c4cccc(F)c4)CC5CCCCc6cccc6</chem>	
<p>AS5</p>	<chem>O=C2/C=C\c1cccc1[NH+]2CCN6CCN(C5c3cccc3Cc4cccc45)CC6CCCCc7cccc7</chem>	

UNIVERSIDAD TÉCNICA DE AMBATO



FACULTAD DE CIENCIA E INGENIERÍA EN ALIMENTOS Y BIOTECNOLOGÍA



CARRERA: INGENIERÍA BIOQUÍMICA

AMBATO – ECUADOR

2021