Universidad Nacional Mayor de San Marcos
Universidad del Perú. Decana de América
Facultad de Ciencias Biológicas
Escuela Profesional de Genética y Biotecnología

Análisis *in silico* del dominio serinoproteasa de TMPRSS2 de *Mus musculus* y su interacción con inhibidores mediante herramientas computacionales

TESIS

Para optar el Título Profesional de Bióloga Genetista
Biotecnóloga

AUTOR

Nathalie Jackeline RAMIREZ QUISPE

ASESOR

Mg. Gustavo Adolfo Sandoval Peña

Lima, Perú

2026



Reconocimiento - No Comercial - Compartir Igual - Sin restricciones adicionales

https://creativecommons.org/licenses/by-nc-sa/4.0/

Usted puede distribuir, remezclar, retocar, y crear a partir del documento original de modo no comercial, siempre y cuando se dé crédito al autor del documento y se licencien las nuevas creaciones bajo las mismas condiciones. No se permite aplicar términos legales o medidas tecnológicas que restrinjan legalmente a otros a hacer cualquier cosa que permita esta licencia.

Referencia bibliográfica

Ramirez, N. (2026). *Análisis in silico del dominio serinoproteasa de TMPRSS2 de Mus musculus y su interacción con inhibidores mediante herramientas computacionales*. [Tesis de pregrado, Universidad Nacional Mayor de San Marcos, Facultad de Ciencias Biológicas, Escuela Profesional de Genética y Biotecnología]. Repositorio institucional Cybertesis UNMSM.

Metadatos complementarios

Datos de autor	
Nombres y apellidos	Nathalie Jackeline Ramirez Quispe
Tipo de documento de identidad	DNI
Número de documento de identidad	75966550
URL de ORCID	https://orcid.org/0000-0001-8609-0139
Datos de asesor	
Nombres y apellidos	Gustavo Adolfo Sandoval Peña
Tipo de documento de identidad	DNI
Número de documento de identidad	41020762
URL de ORCID	https://orcid.org/0000-0002-8392-9880
Datos del jurado	
Presidente del jurado	
Nombres y apellidos	Patricia Gloria Woll Toso
Tipo de documento	DNI
Número de documento de identidad	10587728
Miembro del jurado 1	
Nombres y apellidos	Wilser Andrés García Quispes
Tipo de documento	DNI
Número de documento de identidad	40480447
Miembro del jurado 2	
Nombres y apellidos	Dan Erick Vivas Ruiz
Tipo de documento	DNI
Número de documento de identidad	41951131

Datos de investigación	
Línea de investigación	Desarrollo de Productos Biológicos
Grupo de investigación	Grupo de Investigación en Bioinformática y Biología Estructural (GIBIOBES)
Agencia de financiamiento	Sin financiamiento.
Ubicación geográfica de la investigación	Edificio: Facultad de Ciencias Biológicas – Universidad Nacional Mayor de San Marcos País: Perú Departamento: Lima Provincia: Lima Distrito: Lima Calle: Germán Amezaga N° 375 Latitud: -12.059438 Longitud: -77.082136
Año o rango de años en que se realizó la investigación	Enero 2024-Junio 2026
URL de disciplinas OCDE	Bioinformática https://purl.org/pe-repo/ocde/ford#1.02.03

Universidad Nacional Mayor de San Marcos
(Universidad del Perú, Decana de América)

FACULTAD DE CIENCIAS BIOLÓGICAS

ACTA DE SESIÓN PARA OPTAR AL TÍTULO PROFESIONAL DE BIÓLOGA GENETISTA BIOTECNÓLOGA (MODALIDAD: SUSTENTACIÓN PRESENCIAL DE TESIS)

Siendo las .10:52 horas del 14 de setiembre de 2026, en el Salón de Grados de la Facultad de Ciencias Biológicas y en presencia del Jurado formado por los profesores que suscriben, se inició la sesión para optar al **Título Profesional de Bióloga Genetista Biotecnóloga** de **NATHALIE JACKELINE RAMIREZ QUISPE.**

Luego de dar lectura y conformidad al expediente **N.º UNMSM-20250137558**, la titulando expuso su Tesis: **ANÁLISIS *in silico* DEL DOMINIO SERINOPROTEASA DE TMPRSS2 DE *Mus musculus* Y SU INTERACCIÓN CON INHIBIDORES MEDIANTE HERRAMIENTAS COMPUTACIONALES** y el Jurado efectuó las preguntas del caso calificando la exposición con la nota ...19.., calificativo: ..Aprobado...con...méritos...honores

Finalmente, el expediente será enviado a la Escuela Profesional de Genética y Biotecnología y al Consejo de Facultad para que se apruebe otorgar el **Título Profesional de Bióloga Genetista Biotecnóloga** a **NATHALIE JACKELINE RAMIREZ QUISPE** y se eleve lo actuado al Rectorado para conferir el respectivo grado, conforme a ley.

Siendo las .12:16 horas se levantó la sesión.

Ciudad Universitaria, 14 de setiembre de 2026.

PATRICIA GLORIA WOLL TOSO
PRESIDENTA

GUSTAVO ADOLFO SANDOVAL PEÑA
ASESOR

WILSER ANDRES GARCIA QUISPES
MIEMBRO

DAN ERICK VIVAS RUIZ
MIEMBRO

UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS
Universidad del Perú, Decana de América

Anexo 1

CERTIFICADO DE SIMILITUD

Yo, **Gustavo Adolfo Sandoval Peña**, en mi condición de asesor acreditado con el **Dictamen N° 023-2026-EPGBT-FCB/UNMSM** de la tesis cuyo título es: **"Análisis *in silico* del dominio serinoproteasa de TMPRSS2 de *Mus musculus* y su interacción con inhibidores mediante herramientas computacionales"**, presentada por la Bach. **Nathalie Jackeline Ramirez Quispe**, para optar al Título Profesional de **Bióloga Genetista Biotecnóloga**,

CERTIFICO que se ha cumplido con lo establecido en la Directiva de Originalidad y de Similitud de Trabajos Académicos, de Investigación y Producción Intelectual. Según la revisión, análisis y evaluación mediante la plataforma PIDS, el documento evaluado cuenta con el **4%** de similitud, nivel PERMITIDO para continuar con los trámites correspondientes y para su **publicación en el repositorio institucional.**

Se emite el presente certificado en cumplimiento de lo establecido en las normas vigentes, como uno de los requisitos para la obtención del Título Profesional de Bióloga Genetista Biotecnóloga.

Lima, 20 de agosto del 2026

UNMSM
Firmado digitalmente por SANDOVAL PEÑA Gustavo Adolfo FAU 20148092282 soft
Motivo: Soy el autor del documento
Fecha: 21.08.2026 16:04:13 -05:00

Mg. Gustavo Adolfo Sandoval Peña
Profesor Principal TC 40 h
Asesor de Tesis
DNI N° 41020762

La vigencia del presente certificado corresponde a un período de noventa (90) días desde la fecha de su emisión; siempre que no haya observaciones en el texto durante el proceso de revisión por el jurado informante o surja alguna contingencia.

DEDICATORIA

A mis padres, Guisela Rocio y Raúl Martin, por apoyarme en cada paso de vida. A mis familiares, tía Tiffany y tío Gino, por todos sus consejos y apoyo incondicional. En especial a mis abuelos Carmen Enríquez y Eugenio Quispe, por siempre creer en mí y brindarme su amor inagotable. A mis hermanos Margorit y Víctor, por su cariño y por estar siempre a mi lado en este camino.

AGRADECIMIENTO

Al Dr. Gustavo Adolfo Sandoval Peña, por su invaluable apoyo y mentoría a lo largo de este proceso. Su guía en el Grupo de Investigación en Bioinformática y Estructural (GIBIOBES) me permitió desarrollar mis propias ideas y fortalecer mi formación en el campo de la biología estructural y la bioinformática.

A la Universidad Nacional Mayor de San Marcos y a todos sus docentes, quienes han sido pilares fundamentales en mi formación académica. En especial, quiero agradecer a los profesores Mónica Paredes, Wilser García, Pedro Romero y Jaime Sánchez, por su compromiso con la enseñanza, su apoyo en mi desarrollo profesional y las oportunidades de aprendizaje que me brindaron.

A mis amigas Katherine y Cristina, por los momentos compartidos, por los sueños y metas que construimos juntas, por nuestras experiencias lado a lado y, sobre todo, por su apoyo incondicional a lo largo de toda la carrera universitaria. Su amistad ha sido un pilar fundamental en este camino. A Elvira, Marina, Emilio y Natalia, por su amistad y compañía durante mi estancia en España. Su apoyo y calidez hicieron de esta experiencia un capítulo inolvidable en mi vida.

Finalmente, expreso mi agradecimiento a mis compañeros y mentores del Laboratorio de Biomoléculas del Centro de Investigación en Ciencias de la Salud por su constante apoyo, por compartir sus conocimientos y experiencias, y por sus valiosos consejos, que contribuyeron al desarrollo de esta tesis. Asimismo, agradezco la oportunidad de crecer profesionalmente en dos áreas que me apasionan, la biología molecular y la bioinformática, en un espacio donde cada día continúo aprendiendo y aportando con mis conocimientos.

ÍNDICE GENERAL

ÍNDICE DE FIGURAS vi
ÍNDICE DE TABLAS viii
ÍNDICE DE ANEXOS ix
RESUMEN xiii
ABSTRACT xiv
I. INTRODUCCIÓN 1
I.1. Introducción 1
I.2. Planteamiento del problema 2
I.2.1. Determinación del problema 2
I.2.2. Formulación del problema 6
I.3. Objetivos 6
I.3.1. Objetivo general 6
I.3.2. Objetivos específicos 6
I.4 Importancia y alcance de la investigación 6
I.5.Limitaciones de la investigación 7
II. REVISIÓN LITERARIA 8
II.1. Marco teórico 8
II.1.1. Serinoproteasas de tipo II (TMPRSS2) 8
II.1.1.1. Estructura y función de TMPRSS2 9
II.1.2. Participación de TMPRSS2 en la activación proteolítica de los virus 11
II.1.2.1. Influenza 11
II.1.2.2. Coronavirus 13
II.1.3. *Mus musculus* como modelo experimental de virulencia 15
II.1.4. Potenciales inhibidores de TMPRSS2 16
II.1.4.1. Mesilato de camostato 16
II.1.4.2. Mesilato de nafamostat 17
II.1.4.3. Lopinavir y valrubicin 18
II.2. Antecedentes del estudio 19
II.3. Bases teóricas 20
II.3.1. Modelado molecular 20
II.3.1.1. Modelado por homología 21
II.3.1.2. Modelado por *threading* 22
II.3.1.3. Modelado *ab initio* 23
II.3.1.4. Alphafold 24
II.3.2 Docking molecular 25
II.3.3. Dinámica Molecular 27
II.3.3.1. RMSD (Desviación cuadrática media) 30
II.3.3.2. RMSF (Fluctuación cuadrática media) 30
II.4 Definición de términos 30
III. HIPÓTESIS Y VARIABLES 31

III.1. Hipótesis 31
III.2. Variables 31
III.2.1. Variable independiente 31
III.2.2. Variable dependiente 31
III.3. Operacionalización de las variables 31

IV. MATERIALES Y MÉTODOS 32

IV.1. Área de estudio 32
IV.2. Diseño de investigación 32
IV.3. Población y muestra 33
IV.4. Procedimientos, Técnicas e instrumentos de recolección de información 34
IV.4.1. Recuperación de secuencia y estructura del gen 34
IV.4.2. Análisis de parámetros bioquímicos con ProtParam 34
IV.4.3. Determinación de dominios conservados 35
IV.4.4. Identificación de proteínas que interactúan con TMPRSS2 35
IV.4.5. Alineamiento de secuencias 35
IV.4.6. Predicción de estructuras secundarias 35
IV.4.7. Predicción de estructuras terciarias mediante los métodos de: 36
IV.4.7.1. Modelado por homología 36
IV.4.7.2. Modelado por *threading* 37
IV.4.7.3. Modelado *ab initio* 38
IV.4.7.4. Alphafold 38
IV.4.8. Validación de modelos 39
IV.4.8.1. PDBSum 39
IV.4.8.2. ModFOLD 39
IV.4.9. Simulación de dinámica molecular 39
IV.4.10. Acoplamiento molecular por AutoDock 41
IV.4.11. Simulación de dinámica molecular del complejo proteína - ligando 42
IV.5. Análisis estadístico 43

V. RESULTADOS 44

V.1. Presentación y análisis de los resultados 44
V.1.1. Análisis de la estructura genética 44
V.1.2. Análisis de parámetros bioquímicos 45
V.1.3. Análisis de dominios conservados. 48
V.1.4. Interacción proteína- proteína. 48
V.1.5. Alineamiento de secuencias 49
V.1.6. Predicción de estructuras secundarias: 52
V.1.7. Predicción de estructuras terciarias 52
V.1.7.1. Por Homología 52
V.1.7.2. Por *Threading* 54
V.1.7.3. *Ab initio* 55
V.1.7.4. Alphafold 56
V.1.8. Validación de modelos 57
V.1.8.1. PDBSum 57
V.1.8.2. ModFOLD 58
V.1.9. Simulación de dinámica molecular (MDS) 59

V.1.10. Análisis de la calidad estructural y conservación geométrica del sitio activo del dominio serinoproteasa de TMPRSS2 de *Mus musculus* frente a la referencia cristalográfica (PDB: 7MEQ) 64
V.1.11. Acoplamiento molecular del dominio serinoproteasa de TMPRSS2 y los 4 inhibidores por AutoDock 67
V.1.12. Simulación de dinámica molecular (MD) del dominio serinoproteasa de TMPRSS2 y 4 inhibidores 74
V.1.12.1. Análisis de distancias respecto al centro de masa (COM) y tiempo de permanencia 82
V.1.12.2. Análisis del Radio de Giro (Rg) 83
V.1.12.3. Monitoreo de distancias hacia los residuos de la tríada catalítica84
V.1.12.4. Variación de las distancias intermoleculares en la tríada catalítica 85
VI. DISCUSIÓN 86
VI.1. Caracterización *in silico* del dominio serinoproteasa de TMPRSS2: propiedades fisicoquímicas, predicción estructural y estabilidad molecular 87
VI.2. Acoplamiento molecular del dominio serinoproteasa de TMPRSS2 y los 4 inhibidores por AutoDock 95
VI.2.1. Análisis e interpretación de los inhibidores referenciales de bajo masa molecular: Nafamostat y Camostat 96
VI.2.3. Análisis e interpretación de los inhibidores exploratorios de alto masa molecular: Valrubicina y Lopinavir 99
VI.3. Simulación de dinámica molecular (MDS) del dominio serinoproteasa de TMPRSS2 y 4 inhibidores. 101
VI.3.1 Análisis dinámico y perturbación geométrica de la tríada catalítica de TMPRSS2 107
VII. CONCLUSIONES 108
VIII. RECOMENDACIONES 109
IX. REFERENCIAS BIBLIOGRÁFICAS 110
X. ANEXOS 125

ÍNDICE DE FIGURAS

Figura 1. Organización de los dominios de TMPRSS2. 10

Figura 2. El ciclo de replicación del virus de la influenza y la escisión proteolítica de las proteasas del huésped. 12

Figura 3. El ciclo de replicación del coronavirus y la escisión proteolítica de las proteasas del huésped.....14

Figura 4. Esquema de trabajo propuesto por Morris y Wilby (2008) para la metodología de Acoplamiento Molecular. 26

Figura 5. Esquema del estudio estructural y dinámico de una proteína. Adaptado de Hollingsworth y Dror, 2018. 29

Figura 6. Estructura del gen Tmprss2 de *Mus musculus.* 45

Figura 7. Gráfica de perfiles de la proteína TMPRSS2, según Prosite.....48

Figura 8. Análisis de la red de interacción proteína-proteína de la TMPRSS2 de *Mus musculus.* 49

Figura 9. Alineamiento múltiple de secuencias del dominio serinoproteasa de TMPRSS2 de *Mus musculus* con secuencias ortólogas de otros animales.....51

Figura 10. Predicción de estructuras secundarias del dominio serinoproteasa de TMPRSS2 de *Mus musculus.*.....52

Figura 11. Estructuras tridimensionales predichas del dominio serinoproteasa de TMPRSS2 en *Mus musculus* mediante modelado por homología. 54

Figura 12. Estructura 3D del dominio serinoproteasa predicha por la metodología *Threading* (plantilla c7meq) visualizada en el programa PyMOL. 55

Figura 13. Estructura 3D del dominio serinoproteasa predicha por Robetta visualizada en el programa PyMOL. 56

Figura 14. Alineamiento de los 5 modelos predichos por el algoritmo Alphafold visualizada en PyMOL.....57

Figura 15. Figura 15. Comparación estructural y organización del dominio serinoproteasa (SP) de TMPRSS2 de *Mus musculus*. 59

Figura 16. Gráficos de RMSD (nm) en función del tiempo (10 ns) para diferentes modelos de dominio serinoproteasa de TMPRSS2 obtenidos mediante distintos enfoques de modelado.. 62

Figura 17. Gráficos de fluctuación del RMSF (nm) en función del número de átomos a lo largo de la simulación de dinámica molecular (MD) de 10 ns 64

Figura 18. Superposición estructural tridimensional de los modelos del dominio serinoproteasa de TMPRSS2 de *Mus musculus* frente a la referencia cristalográfica. 65

Figura 19. Conservación geométrica de la tríada catalítica y caracterización estructural del bolsillo de especificidad S1 de TMPRSS2. 66

Figura 20. Caracterización estructural del bolsillo de especificidad S1 de TMPRSS2. 66

Figura 21. Interacción molecular entre el dominio serinoproteasa de TMPRSS2 y el ligando nafamostat obtenida mediante docking molecular en UCSF Chimera.....70

Figura 22. Interacción molecular entre el dominio serinoproteasa de TMPRSS2 y el ligando camostat obtenida mediante docking molecular en UCSF Chimera. 71

Figura 23. Interacción molecular entre el dominio serinoproteasa de TMPRSS2 y el ligando lopinavir obtenida mediante docking molecular en UCSF Chimera. 72

Figura 24. Interacción molecular entre el dominio serinoproteasa de TMPRSS2 y el ligando valrubicin obtenida mediante docking molecular en UCSF Chimera. 73

Figura 25. Desviación Media Cuadrática (RMSD) de las interacciones proteína-ligando del dominio serinoproteasa TMPRSS2 y 4 inhibidores. 78

Figura 26. Fluctuaciones residuales (RMSF) del dominio serinoproteasa de TMPRSS2 en complejos con sus inhibidores durante 50 ns de dinámica molecular. 80

Figura 27. Interacciones de puentes de hidrógeno entre el dominio serinoproteasa de TMPRSS2 y los ligandos durante la dinámica molecular.....82

ÍNDICE DE TABLAS

Tabla 1. Composición aminoacídica del dominio serinoproteasa de TMPRSS2 de *Mus musculus*.....46

Tabla 2. Características fisicoquímicas del dominio serinoproteasa de TMPRSS2 de *Mus musculus*.....47

Tabla 3. Resultados del alineamiento del dominio serinoproteasa de TMPRSS2 de *Mus musculus* con plantillas del PDB, utilizando SWISS-MODEL y BLASTp. La calidad de los modelos se evaluó mediante GMQE, QMEANDisCo Global, cobertura y porcentaje de identidad de secuencia.....53

Tabla 4. Resultados del modelamiento por *threading* del dominio serinoproteasa de TMPRSS2. La calidad de los modelos se evaluó mediante Coverage, Confidence y porcentaje de identidad.....54

Tabla 5. Resultados del acoplamiento molecular entre el dominio serinoproteasa de TMPRSS2 con sus inhibidores obtenidos en AutoDock.....69

Tabla 6. Resumen de los parámetros dinámicos de los complejos TMPRSS2–ligando durante 50 ns de simulación molecular.....82

Tabla 7. Valores promedio de distancia entre centros de masa (COM) y estimación del tiempo de permanencia intermolecular.....83

Tabla 8. Análisis cuantitativo del Radio de Giro (Rg) para los complejos del dominio serinoproteasa de TMPRSS2.....84

Tabla 9. Distancias promedio calculadas entre el centro de masa (COM) del ligando y los residuos catalíticos de TMPRSS2 durante los 50 ns de simulación.....85

Tabla 10. Valores promedio de las distancias intermoleculares (nm) y desviaciones estándar en la tríada catalítica de TMPRSS2 con nafamostat, lopinavir y valrubicina respecto a su estructura apo.....86

Tabla 11. Valores promedio de las distancias intermoleculares (nm) y desviaciones estándar en la tríada catalítica de TMPRSS2 con camostat respecto a su estructura apo.....86

ÍNDICE DE ANEXOS

Anexo 1. Perfiles y firmas de motivos determinados por alta probabilidad de ocurrencia para la proteína TMPRSS2 en *Mus musculus*, según Prosite. 125

Anexo 2. Resumen gráfico de los dominios conservados en la proteína TMPRSS2 de *Mus musculus*. LDLa, SRCR_2 y Tryp_SPc, según NCBI. También se señalan los hits específicos, no específicos y las superfamilias a las que pertenecen estos dominios. 126

Anexo 3. Lista de dominios determinados por alta probabilidad de ocurrencia para la proteína TMPRSS2 en *Mus musculus*, según NCBI. 127

Anexo 4. Resultados de los estados de predicción de estructuras secundarias por PHD de TMPRSS2 de *Mus musculus*. Codificada por colores la longitud de la secuencia y el porcentaje de cada elemento secundario. 128

Anexo 5. Estimación de calidad local para cada residuo de los modelos predichos del dominio serinoproteasaTMPRSS2 de *Mus musculus* a partir de las estructuras de referencia (A) 7MEQ, (B) 7XYD.....129

Anexo 6. QMEAN Z-score de los modelos generados a partir de las estructuras de referencia (A) 7MEQ, (B) 7XYD en función del tamaño de la proteína (en residuos) en comparación con un conjunto no redundante de estructuras del PDB. 130

Anexo 7. Resultados de predicción del dominio SP mediante el método *ab initio*. 131

Anexo 8. Cobertura e identidad de secuencia en el alineamiento múltiple de secuencias (MSA) usando ColabFold en la predicción tridimensional del dominio SP de TMPRSS2. 133

Anexo 9. Predicción de IDDT para cada posición de la proteína TMPRSS2 en cinco modelos diferentes, etiquetados como rank_1, rank_2, rank_3, rank_4 y rank_5 predichos por Alphafold. 134

Anexo 10. Gráfica del Plot de Ramachandran para validar las estructuras de los 9 modelos elegidos por las diferentes metodologías de modelamiento por el servidor de validación PDBsum Generate. Los triángulos representan los residuos de glicina (Gly). Los cuadrados representan los residuos de prolina (Pro). 135

Anexo 11. Evaluación estadística de la calidad de los modelos con ModFLOD. En base al error entre los residuos predichos y residuos reales, los valores no superaron los 4Å para la región del dominio serinoproteasa en los 9 modelos. 139

Anexo 12. Resultados del docking molecular del compuesto nafamostat con el modelo de TMPRSS2 generado con SWISS-MODEL.....142

Anexo 13. Resultados del docking molecular del compuesto camostat con el modelo de TMPRSS2 generado con SWISS-MODEL. 143

Anexo 14. Resultados del docking molecular del compuesto lopinavir con el modelo de TMPRSS2 generado con SWISS-MODEL. 144

Anexo 15. Resultados del docking molecular del compuesto valrubicin con el modelo de TMPRSS2 generado con SWISS-MODEL. 145

Anexo 16. Resultados del docking molecular del compuesto nafamostat con el modelo de TMPRSS2 generado con Alphafold (modelo 3). 146

Anexo 17. Resultados del docking molecular del compuesto camostat con el modelo de TMPRSS2 generado con Alphafold (modelo 3). 147

Anexo 18. Resultados del docking molecular del compuesto lopinavir con el modelo de TMPRSS2 generado con Alphafold (modelo 3) 148

Anexo 19. Resultados del docking molecular del control negativo valrubicin con el modelo de TMPRSS2 generado con Alphafold (modelo 3).....149

Anexo 20. Resultados del docking molecular del compuesto Adenosine triphosphate con el modelo del dominio serinoproteasaTMPRSS2. 150

Anexo 21. Interacciones de enlaces de hidrógeno entre TMPRSS2 y los ligandos nafamostat, camostat, valrubicin y lopinavir obtenidas mediante docking molecular en UCSF Chimera y PLIP. La distancia donador–aceptor (Å) representa la distancia entre el átomo donador y el aceptor. 151

Anexo 22. Interacciones significativas entre residuos del sitio activo de TMPRSS2 y los ligandos nafamostat, valrubicin, camostat y lopinavir obtenidos en UCSF Chimera y PLIP. 153

Anexo 23. Número de contactos (< 0.35 nm) del complejo TMPRSS2–Nafamostat durante 50 ns de simulación de dinámica molecular. 156

Anexo 24. Número de contactos (< 0.35 nm) del complejo TMPRSS2–Camostat durante 50 ns de simulación de dinámica molecular. 157

Anexo 25. Número de contactos (< 0.35 nm) del complejo TMPRSS2–Lopinavir durante 50 ns de simulación de dinámica molecular. 158

Anexo 26. Número de contactos (< 0.35 nm) del complejo TMPRSS2–Valrubicina durante 50 ns de simulación de dinámica molecular. 159

Anexo 27. Estabilidad estructural del dominio SP de TMPRSS2 evaluada mediante RMSD durante 50 ns de dinámica molecular en estado apo y en complejos holo con nafamostat, lopinavir y valrubicina.....160

Anexo 28. Estabilidad estructural del dominio SP de TMPRSS2 modelado con camostat, evaluada mediante RMSD en estados apo y holo durante 50 ns de dinámica molecular. 161

Anexo 29. Perfil de fluctuación residual (RMSF) del dominio SP de TMPRSS2 en estado apo y en complejos holo con nafamostat, lopinavir y valrubicina durante 50 ns de dinámica molecular. 162

Anexo 30. Perfil de fluctuación residual (RMSF) del dominio SP de TMPRSS2 en estado apo y en complejo con camostat durante 50 ns de dinámica molecular. 163

Anexo 31. Distancia entre los centros de masa de TMPRSS2 y Nafamostat durante 50 ns de simulación de dinámica molecular 164

Anexo 32. Distancia entre los centros de masa de TMPRSS2 y Camostat durante 50 ns de simulación de dinámica molecular 165

Anexo 33. Distancia entre los centros de masa de TMPRSS2 y lopinavir durante 50 ns de simulación de dinámica molecular 166

Anexo 34. Distancia entre los centros de masa de TMPRSS2 y valrubicina durante 50 ns de simulación de dinámica molecular 167

Anexo 35. Perfil del radio de giro (Rg) del dominio serinoproteasa de TMPRSS2. 168

Anexo 36. Perfil del radio de giro (Rg) del complejo TMPRSS2–Nafamostat 169

Anexo 37. Perfil del radio de giro (Rg) del complejo TMPRSS2–Camostat 170

Anexo 38. Perfil del radio de giro (Rg) del complejo TMPRSS2–Valrubicina 171

Anexo 39. Perfil del radio de giro (Rg) del complejo TMPRSS2–Lopinavir 172

Anexo 40. Distancia promedio entre nafamostat y residuos catalíticos de TMPRSS2 durante 50 ns de simulación.....173

Anexo 41. Distancia promedio entre camostat y residuos catalíticos de TMPRSS2 durante 50 ns de simulación.....174

Anexo 42. Distancia promedio entre lopinavir y residuos catalíticos de TMPRSS2 durante 50 ns de simulación.....175

Anexo 43. Distancia promedio entre valrubicina y residuos catalíticos de TMPRSS2 durante 50 ns de simulación.....176

Anexo 44. Variación de la distancia entre los residuos catalíticos His41 y Ser186 en TMPRSS2 libre (apo) y en los complejos con los ligandos durante 50 ns de dinámica molecular. 177

Anexo 45. Variación de la distancia entre los residuos catalíticos His41 y Asp90 en TMPRSS2 libre (apo) y en los complejos con los ligandos durante 50 ns de dinámica molecular 178

Anexo 46. Variación de la distancia entre los residuos catalíticos Asp90 y Ser186 en TMPRSS2 libre (apo) y en los complejos con los ligandos durante 50 ns de dinámica molecular 179

RESUMEN

En el siglo XXI, las epidemias y pandemias de influenza y coronavirus han representado un gran riesgo para la salud pública, por ello se ha demandado una serie de estudios estructurales relacionados con la proteasa de serina transmembrana tipo II (TMPRSS2) debido a su importancia en la infección viral, ya que puede servir como potencial diana farmacológica. Sin embargo, la mayoría de ellos se han visto obstaculizados por la falta de conocimiento de la estructura de TMPRSS2 en organismos modelos, como en el ratón (*Mus musculus*). Por esta razón, el presente estudio tiene como objetivo realizar un análisis *in silico* del dominio serinoproteasa de TMPRSS2 de *Mus musculus* y predecir su interacción con fármacos utilizando herramientas computacionales. El modelado se realizó por 4 estrategias: homología, *threading*, *ab initio*, y el algoritmo AlphaFold. El mejor modelo fue elegido según análisis con ModFold v8.0, PDBsum, simulación de dinámica molecular (SMD) y *docking molecular*. Se analizó las interacciones favorables entre los modelos predichos del dominio serinoproteasa de TMPRSS2 y los fármacos (nafamostat, camostat, lopinavir y valrubicina) mediante docking molecular. La evaluación de la estabilidad relativa de TMPRSS2 e inhibidores en un tiempo de 50 ns se realizó con GROMACS versión 2023.1. Se compararon los resultados del análisis de todas las estructuras predichas y se determinó que los modelos 3D de TMPRSS2 obtenidos mediante homología con SWISS-MODEL y el generado por AlphaFold mostraron estabilidad relativa, según los valores de RMSD obtenidos en la SMD. El docking molecular y dinámica molecular proteína-ligando predijeron que nafamostat y camostat presentan interacciones favorables con TMPRSS2 y un comportamiento relativo estable sin fluctuaciones abruptas respecto a los complejos lopinavir y valrubicin con TMPRSS2. En conclusión, se generó un modelo 3D de la proteína TMPRSS2 de *Mus musculus* y se predijo interacciones favorables y estabilidad relativa con inhibidores nafamostat y camostat.

ABSTRACT

In the 21st century, epidemics and pandemics of influenza and coronavirus have represented a great risk to public health, therefore a series of structural studies related to transmembrane serine protease type II (TMPRSS2) has been demanded due to its importance in viral infection, since it can serve as a potential pharmacological target. However, most of them have been hindered by the lack of knowledge of the structure of TMPRSS2 in model organisms, such as in the mouse (*Mus musculus*). For this reason, the present study aims to perform an in silico analysis of the serine protease domain of TMPRSS2 of *Mus musculus* and predict its interaction with drugs using computational tools. The modeling was performed by 4 strategies: homology, *threading*, *ab initio*, and the AlphaFold algorithm. The best model was chosen according to analysis with ModFold v8.0, PDBsum, molecular dynamics simulation (SMD) and molecular *docking*. Favorable interactions between the predicted models of the serine protease domain of TMPRSS2 and the drugs (nafamostat, camostat, lopinavir and valrubicina) were analyzed by molecular docking. The evaluation of the relative stability of TMPRSS2 and inhibitors in a time of 50 ns was performed with GROMACS versión 2023.1. The results of the analysis of all predicted structures were compared and it was determined that the 3D models of TMPRSS2 obtained by homology with SWISS-MODEL and the one generated by AlphaFold showed relative stability, according to the RMSD values obtained in the SMD. Molecular docking and protein-ligand molecular dynamics predicted that nafamostat and camostat present favorable interactions with TMPRSS2 and a relative stable behavior without abrupt fluctuations compared to the lopinavir and valrubicin complexes with TMPRSS2. In conclusion, a 3D model of the TMPRSS2 protein of *Mus musculus* was generated and favorable interactions and relative stability with inhibitors nafamostat and camostat were predicted.

I. INTRODUCCIÓN

I.1. Introducción

A principios del siglo XXI se han producido las epidemias o pandemias de virus de la influenza y del coronavirus, lo que supone un riesgo para la salud pública (Shen et al., 2017). Los virus responsables son el SARS-CoV-2, SARS-CoV, el MERS-CoV, el virus pandémico de la gripe A H1N1 del 2009, y el virus asiático de la gripe A H7N9 (Hatesuer et al., 2013; Iwata-Yoshikawa et al., 2019; Li et al., 2021; Sakai et al., 2014; Sun et al., 2020; Daodu et al., 2026). Así, en 2019 fue identificada la enfermedad COVID-19 en Wuhan, China, siendo reportada en más de 180 países, con más de 676 millones de infectados y más de 6,8 millones de muertes (Johns Hopkins University of Medicine, 2023).

Durante la infección por SARS-CoV-2, la proteína Spike (S) del coronavirus se une al receptor enzima convertidora de angiotensina 2 (ACE2) del hospedero. Sin embargo, para que ocurra la fusión de membranas y la entrada viral, la proteína S requiere un cebado proteolítico mediado por la serinoproteasa transmembrana TMPRSS2 (Matsuyama et al., 2020; Widiasta et al., 2020; Fraser et al., 2021). De forma similar, el ciclo de replicación en el virus de la influenza depende de la activación de la hemaglutinina (HA) viral mediante la escisión proteolítica por TMPRSS2 (Hatesuer et al., 2013).

Ante la emergencia sanitaria, se han explorado diversas estrategias antivirales para abordar el neumotropismo y la patogenicidad de virus como el coronavirus y la influenza. Sin embargo, hasta el momento, no existe evidencia clínica concluyente sobre la efectividad y seguridad de los antivirales dirigidos contra proteínas virales. Esto se debe a la alta tasa de mutación de estos patógenos, lo que puede generar resistencia

persistente a los tratamientos y respuestas inmunológicas no protectoras frente a nuevas variantes emergentes (Shen et al., 2017; Wang et al., 2023; Dinata et al., 2025).

En este contexto, una alternativa prometedora es el desarrollo de terapias dirigidas a proteínas del huésped esenciales para el ciclo de vida viral y altamente conservadas. Entre ellas, TMPRSS2 ha sido identificada como una diana terapéutica potencial, ya que desempeña un papel clave en la activación proteolítica de diversos virus respiratorios, incluidos el virus de la gripe A H7N9, el subtipo H1N1, el SARS-CoV-2, el SARS-CoV y el MERS-CoV (Shen et al., 2017; Wang et al., 2023)

Se incluyeron inhibidores de referencia como el nafamostat y camostat debido a que experimentalmente se ha demostrado su capacidad para inhibir la actividad de TMPRSS2 (Lee *et al.*, 2021; Shrimp *et al.*, 2020). Por otro lado, lopinavir y valrubicina fueron seleccionados como compuestos de interés exploratorio, porque estudios previos de acoplamiento molecular lo proponen como posibles inhibidores de TMPRSS2 en el contexto de estrategias de reposicionamiento farmacológico (Baby et al., 2021). La inclusión de estos dos grupos de compuestos permite comparar el comportamiento de inhibidores con evidencia experimental consolidada frente a compuestos candidatos identificados mediante aproximaciones computacionales.

I.2. Planteamiento del problema

I.2.1. Determinación del problema

El principal vacío que cubre la presente tesis es la ausencia de una caracterización estructural dinámica y molecular de TMPRSS2 de *Mus musculus* interactuando con ligandos (Brooke y Prischi, 2020), lo que conlleva a 3 limitaciones que la presente tesis resuelve. Primero, el ratón es el biomodelo preclínico estándar para evaluar terapias contra virus respiratorios, el cual asemeja en grado variable a las características clínicas y patológicas de las infecciones en humanos (van den Brand et

al., 2014; Zmora et al., 2017; Rocha et al., 2022; Färber et al., 2022). Por ello, extrapolar los datos de TMPRSS2 humano al modelo murino induce a sesgos, ya que se desconoce si las variaciones en los residuos periféricos de la secuencia de la variante murina alteran las distancias intermoleculares o el acoplamiento con ligandos. Segundo, los modelos automatizados generados por el algoritmo de AlphaFold proporcionan predicciones 3D estáticas de las proteínas, en su mayoría en su estado apo, sin embargo, no aportan información funcional en base al acoplamiento o simulación bajo condiciones fisiológicas. Tercero, la conservación evolutiva de TMPRSS2 respalda a *Mus musculus* como biomodelo de interés, aunque existe vacíos de información sobre la equivalencia dinámica frente a la proteína humana.

En específico, se ha demostrado que TMPRSS2 tiene un alto nivel de similitud entre la secuencia humana y diferentes modelos animales como los conejillos de indias, perros, gatos, ratas, conejos, hurones, ratones, hámsters y macacos. Debido a esta similitud, se sugiere que TMPRSS2 en dichos modelos animales podría escindir la glicoproteína S del SARS-CoV-2 de manera similar a la variante humana, conocida como hTMPRSS2, e inclusive se puede estudiar la infección por SARS-Cov-2 en ratones transgénicos hACE2 cuya infección se asemeja al de humano (Brooke y Prischi, 2020).

Trabajar con la variante de *Mus musculus* es importante porque el ratón es uno de los biomodelos preclínicos de referencia para reproducir las infecciones respiratorias como la influenza y coronavirus, por lo que podría proporcionar información útil sobre la eficacia, seguridad, toxicidad y efectos secundarios de los fármacos en humanos (Viera y Márquez, 2020) . En primer lugar, se ha demostrado que los ratones reproducen de manera confiable la patogénesis de la enfermedad pulmonar humana, desarrollando cuadros clínicos y descamación del epitelio bronquial muy similares a los escenarios clínicos respiratorios humanos (Menéndez et al., 2005). Adicionalmente, la localización tisular de la proteasa está altamente conservada, ya que por ejemplo en humanos, la

TMPRSS2 se expresa en el epitelio respiratorio positivo para ácido siálico, en modelos murinos el inicio de la infección ocurre también en el epitelio bronquiolar (Hatesuer et al., 2013).

En segundo lugar, la viabilidad de estos modelos de infección radica en que la entrada celular de virus como el SARS-CoV-2 y el MERS-CoV en el epitelio respiratorio no prefiere la vía endosomal, sino que depende de la fusión en la superficie celular dada por la proteasa TMPRSS2 (Li et al., 2021). Además, en el biomodelo murino se produce la coexpresión tisular coordinada de ACE2 y TMPRSS2 en las células diana del sistema respiratorio, jugando ambas proteínas un papel crucial en la activación proteolítica viral (Rocha et al., 2023). De hecho, ensayos *in vitro* e *in vivo* demostraron que TMPRSS2 de *Mus musculus* es capaz de procesar de manera homóloga los mismos sitios de escisión de virus humanos como las cepas de influenza H1N1 y H3N2, lo que respalda que el mecanismo catalítico básico está conservado (Hatesuer et al., 2013).

En tercer lugar, se ha demostrado que TMPRSS2 tiene un alto nivel de similitud entre la secuencia humana y diferentes modelos animales como los conejillos de indias, perros, gatos, ratas, conejos, hurones, ratones, hámsters y macacos. Debido a esta similitud, se sugiere que TMPRSS2 de *Mus musculus* podría escindir la glicoproteína S del SARS-CoV-2 de manera similar a la variante humana, conocida como hTMPRSS2, e inclusive se puede estudiar la infección por SARS-Cov-2 en ratones transgénicos hACE2 cuya infección se asemeja al de humano (Brooke y Prischi, 2020). Además, diversos estudios han demostrado que la eliminación de la proteasa no produce efectos adversos en ratones, demostrando que TMPRSS2 no es esencial para el funcionamiento y desarrollo normal del organismo, por ello su inhibición selectiva resulta ser una estrategia para prevenir la entrada viral (Kim et al., 2006). Asimismo, estudios preclínicos *in vivo* demuestran que la inhibición selectiva de la actividad catalítica de esta proteasa mediante compuestos como el nafamostat y el camostat reduce

drásticamente los títulos virales en el tejido pulmonar, revierte la pérdida de peso y disminuye la mortalidad en modelos murinos infectados con SARS-CoV-2 y MERS-CoV (Li et al., 2021). Debido a la eficacia de estos tratamientos, se vuelve indispensable el modelado estructural específico para optimizar el diseño de fármacos interactuando con este biomodelo animal (Li et al., 2021).

El análisis de la presente tesis aporta una validación predictiva exploratoria para el uso del ratón como modelo experimental. Al evaluar la dinámica molecular del dominio serinoproteasa de TMPRSS2 de *Mus musculus* con inhibidores de referencia se encontró interacciones favorables y con una estabilidad relativa, lo que permitiría avanzar hacia ensayos *in vivo* con este modelo animal. De esta manera, se reduciría el número de individuos sacrificados en fases preclínicas, descartando temprano los ligandos sin complementariedad estructural. Y, se optimizaría los recursos antes de la experimentación real

A pesar de la alta conservación en la región catalítica del dominio serinoproteasa de TMPRSS2 entre *Mus musculus* y *Homo sapiens*, existen sustituciones clave en su periferia estructural. Por ende, el docking molecular solo ofrece una pose estática del acoplamiento del inhibidor. Debido a esto, se requieren simulaciones de dinámica molecular para evaluar si estas diferencias periféricas en el ratón alteran la estabilidad del ligando, deforman la proteína o provocan la expulsión del fármaco a lo largo del tiempo.

El modelo obtenido complementa las estructuras ya disponibles de TMPRSS2. Por lo tanto, el objetivo de esta tesis no es reemplazar las predicciones de la versión humana, sino aportar la caracterización estructural del dominio serinoproteasa de *Mus musculus* por su relevancia como biomodelo preclínico. Además, frente a antecedentes con identidades de apenas 33.82% y 42.19% (Escalante & Ferguson, 2021;

Mahmudpour et al., 2021), el alto porcentaje de identidad alcanzado en este estudio refuerza la robustez de nuestra modelación para la variante murina.

I.2.2. Formulación del problema

¿Cómo se puede caracterizar estructuralmente el dominio serinoproteasa de TMPRSS2 de *Mus musculus* y evaluar la estimación cuantitativa computacional de su afinidad de unión y estabilidad con inhibidores mediante técnicas de modelado molecular, acoplamiento y dinámica molecular?

I.3. Objetivos

I.3.1. Objetivo general

Realizar un análisis *in silico* del dominio serinoproteasa de TMPRSS2 de *Mus musculus* y su interacción con inhibidores utilizando herramientas computacionales.

I.3.2. Objetivos específicos

- Predecir la estructura tridimensional del dominio serinoproteasa de TMPRSS2 de *Mus musculus* utilizando técnicas de modelado.
- Evaluar la afinidad entre el dominio serinoproteasa de TMPRSS2 y sus inhibidores utilizando técnicas de acoplamiento molecular.
- Evaluar la estabilidad entre el dominio serinoproteasa de TMPRSS2 e inhibidores con los parámetros RMSD, RMSF, radio de giro, número de contactos intermoleculares, así como las variaciones en las distancias interatómicas de la tríada catalítica y la proximidad de los ligandos a los residuos del sitio activo.

I.4 Importancia y alcance de la investigación

Esta investigación es importante porque TMPRSS2 es una diana terapéutica potencial contra virus como el SARS-CoV-2 y la influenza. Por ende, analizar la estructura de su dominio serinoproteasa ayuda a comprender la interacción con ligandos y aporta al diseño racional de fármacos. Debido a que esta proteína cuenta con un dominio altamente conservado que media el reconocimiento molecular (Hatesuer et al.,

2013), caracterizarla en modelos experimentales es vital. Así, este trabajo servirá de base para evaluar otros modelos como el hámster, donde el cebado de la proteína S es muy similar al humano y el nafamostat ya demostró tener un efecto quimioprofiláctico (Neary et al., 2023).

Bajo un enfoque ético, los métodos computacionales optimizan la selección del modelo animal idóneo para vacunas y tratamientos antivirales. De esta manera, aplicar el principio de las "3R" (reemplazo, reducción y refinamiento) disminuye la dependencia experimental *in vivo*, promoviendo una investigación eficiente (Ekstrand et al., 2021). Por ende, los cuestionamientos éticos hacia el uso masivo de roedores impulsan enfoques alternativos como las simulaciones *in silico*, modelos 3D y sistemas *in vitro*. Estas herramientas analizan procesos moleculares con buena precisión, reduciendo el número de animales y agilizando la identificación de inhibidores en blancos terapéuticos del hospedero (Viera y Márquez, 2020).

Finalmente, la comparación de diferentes técnicas de modelado estructural permite evaluar la confiabilidad y el comportamiento dinámico de los modelos construidos, proporcionando una base sólida para futuras investigaciones en biología estructural y diseño racional de fármacos.

I.5.Limitaciones de la investigación

Una de las principales limitaciones de la presente tesis es la dependencia de modelos teóricos para predecir la interacción del dominio serinoproteasa de TMPRSS2 con nafamostat, camostat, lopinavir y valrubicina. Por ello, al tratarse de aproximaciones *in silico*, estos resultados necesitan validación experimental mediante ensayos enzimáticos o celulares para poder corroborar su verdadero impacto en un modelo biológico real.

También, el tiempo de simulación en la dinámica molecular podría ampliarse en futuros trabajos para evaluar la estabilidad de los ligandos durante periodos más prolongados. Finalmente, debido a la disponibilidad de recursos computacionales, no se calcularon las energías libres de unión (ΔG_{bind}) mediante métodos como MM/PBSA o MM/GBSA, ni se determinó el tiempo de residencia de los compuestos.

II. REVISIÓN LITERARIA

II.1. Marco teórico

II.1.1. Serinoproteasas de tipo II (TMPRSS2)

Las serinoproteasas pertenecen a una de las familias de enzimas proteolíticas más grande y conservada conocida por sus funciones en el desarrollo, coagulación, digestión e inmunidad. Además, comparten un mecanismo catalítico común para la escisión selectiva de sustratos específicos y con frecuencia participan en cascadas proteolíticas (Antalis et al., 2011).

Dentro de las subfamilias de las serinoproteasas se encuentra las del tipo quimotripsina (S1), la cual contiene enzimas que son secretadas o almacenadas en orgánulos citoplasmáticos a la espera de una señal para su liberación regulada. También, se han identificado un grupo de enzimas ancladas a la membrana celular mediante un dominio transmembrana en el extremo C-terminal (Tipo I) o N-terminal (Tipo II) (Netzel-Arnett et al., 2003). Este último subgrupo de serinoproteasas transmembrana tipo II (TTSP II), incluye la subfamilia hepsina/TMPRSS, formada por 7 enzimas: hepsina, TMPRSS2, TMPRSS3, TMPRSS4, TMPRSS5, MSPL y enteropeptidasa, las cuales cumplen funciones de interés biomédicas como la activación de proteínas, regulación celular e interacción con patógenos (Mantzourani et al., 2022).

La proteína TMPRSS2, al igual que otras TTSP, se sintetiza como zimógeno, que se activa tras una escisión autocatalítica entre Arg255 e Ile256 produciendo una

forma de dos cadenas unidas por un puente disulfuro, donde una parte considerable de ellas pueden liberarse al medio extracelular (Antalis et al., 2011).

II.1.1.1. Estructura y función de TMPRSS2

En 1997, Paoloni-Giacobino et al. describieron la primera caracterización de la proteasa TMPRSS2 usando la técnica captura de exones. El gen *TMPRSS2* se asigna al cromosoma humano 21q22.3 (HC21) y es homólogo al gen de la enteropeptidasa humana. La proteína TMPRSS2 consta de 492 aminoácidos y se expresa principalmente en las células epiteliales de la mama, próstata, riñón, colón, conducto biliar, intestino delgado, páncreas, ovario, glándula salival, estómago y pulmón. La estructura de esta proteína se caracteriza por un dominio citoplasmático N-terminal, un dominio transmembrana, un dominio del receptor LDL de clase A (LDLRA), un dominio receptor scavenger rico en cisteína (SRCR) y un dominio de activación unido al dominio de serinoproteasa altamente conservado estabilizado por enlaces disulfuro intradominio (Figura 1). El dominio serinoproteasa contiene la triada catalítica formada por los residuos de His296, Asp345 y Ser441. Además, contiene 9 residuos de cisteína conservados formando puentes disulfuro que estabilizan la proteína (Brooke y Prischi, 2020; Mantzourani et al., 2022). Además, adopta un pliegue compuesto por dos barriles β que convergen en la hendidura catalítica donde se ubica la tríada catalítica. La selectividad biológica de esta enzima frente a sus sustratos está modulada por bucles A-E (L_A–L_E) y bucles 1-3 (L_1–L_3), los cuales delimitan los subsitios de reconocimiento (S1, S1', S2, S3 y S4) encargados de acomodar las posiciones aminoacídicas del sustrato (Figura 1). En el caso del subsitio S1 se localiza en la base de la cavidad catalítica y está definido por los residuos conservados Asp435, Ser436 y Gly464. En el caso del subsitio S1', es una región interna y estrecha que alberga cadenas laterales pequeñas e hidrofóbicas. En el subsitio S2, es una cavidad influenciada por el residuo Lys342. Mientras que los subsitios S3 y S4 son zonas expuestas y flexibles que facilitan

interacciones polares con el sustrato mediante los residuos Gln438 y Thr341, respectivamente (Fraser et al., 2022).

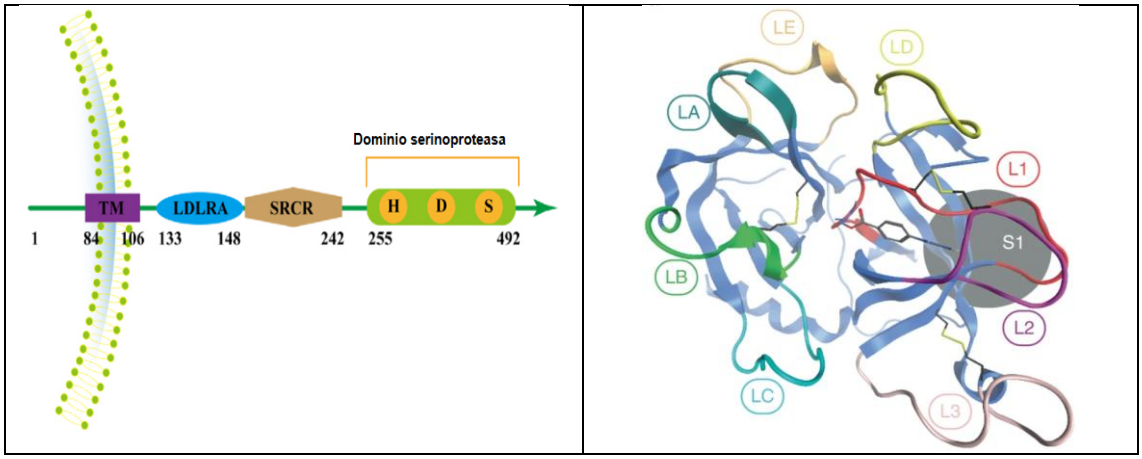


Figura 1. Organización de los dominios de TMPRSS2. Izquierda: Representación esquemática de la estructura de dominios de TMPRSS2 anclada a la membrana celular, TM: dominio transmembrana; LDLRA: dominio del receptor de lipoproteína de baja densidad clase A; SRCR: dominio rico en cisteína del receptor Scavenger; Letras H: histidina; Letras D: aspartato; Letras S: serina. Figura tomada de Shen et al., 2017. (B) Derecha: Estructura tridimensional en cintas del dominio serinoproteasa de TMPRSS2. Se resaltan los bucles superficiales variables encargados del reconocimiento del sustrato (bucles $\boldsymbol{L_A}$-$\boldsymbol{L_E}$ en tonos fríos y $\boldsymbol{L_1}$-$\boldsymbol{L_3}$ en tonos cálidos), así como la ubicación del subsitio de especificidad S1 (esfera gris). Figura tomada de Fraser et al., 2022.

Un estudio sugirió que la localización estratégica de TMPRSS2, como mediador anclado a la membrana celular, desempeña un papel en el dolor asociado al cáncer y en el dolor en general (Lam et al., 2015). Sin embargo, los ratones deficientes en TMPRSS2 se desarrollaron de forma normal, sobreviviendo hasta la edad adulta, y sin mostrar ninguna anomalía fenotípica evidente como infertilidad, enfermedad visible o muerte. Si bien la función fisiológica exacta de TMPRSS2 *in vivo* sigue siendo desconocida, se especula que TMPRSS2 puede contribuir a una función especializada pero no vital que solo se expresa en condiciones de estrés, enfermedad u alguna perturbación sistémica (Kim et al., 2006).

II.1.2. Participación de TMPRSS2 en la activación proteolítica de los virus

II.1.2.1. Influenza

La propagación global de los virus de la influenza A en diversas especies de aves y mamíferos, junto a su capacidad de transmitirse entre diferentes tipos de huéspedes, crea las condiciones para el surgimiento de nuevas variantes virales, las cuáles pueden desencadenar brotes graves de enfermedades respiratorias e incluso desatar una pandemia de influenza (Tarnow et al., 2014).

En la Figura 2 se representa el ciclo de replicación del virus de la influenza. Primero, el virus se une a receptores con ácido siálico en la superficie celular y es internalizado por endocitosis. Durante la maduración del endosoma, la envoltura viral se fusiona con la membrana endosómica, liberando las ribonucleoproteicas virales (vRNP) en el citosol. Éstas se transportan al núcleo, donde ocurre la transcripción y replicación del genoma viral. Luego, los ARNm son traducidos por la maquinaria celular y los ARN virales se exportan al citosol para ensamblarse con proteínas estructurales, empaquetarse en la membrana plasmática y liberar nuevos viriones (Shen et al., 2017).

Algunos estudios han sugerido que el ciclo de replicación del virus de la influenza depende de la activación de la hemaglutinina (HA) del virus mediante la escisión proteolítica por proteasas de la célula hospedera, como TMPRSS2, HAT y otras serinoproteasas transmembrana (TTSPs), incluida TMPRSS4. Estas proteasas pueden escindir la HA de los virus de la gripe humana y aviar que presentan un residuo de arginina en el sitio de escisión (Chaipan et al., 2009; Antalis et al., 2011; Shen et al., 2017; Mantzourani et al., 2022). En la Figura 2, HAT puede escindir la HA de los virus entrantes antes de su incorporación a las células hospederas, así como la HA recién sintetizada durante la liberación de los viriones progenie. Por otro lado, TMPRSS2 únicamente escinde la HA naciente dentro de las células hospederas y no participa en la activación proteolítica de la HA de los viriones entrantes (Böttcher-Friebertshäuser

et al., 2010). En este proceso, el precursor inactivo HA0 con sitio de escisión monobásico es escindido por la proteasa específica TMPRSS2 en el aparato de Golgi del huésped dando lugar a dos subunidades funcionales, HA1 y HA2, unidas por un enlace disulfuro (Böttcher et al., 2006).

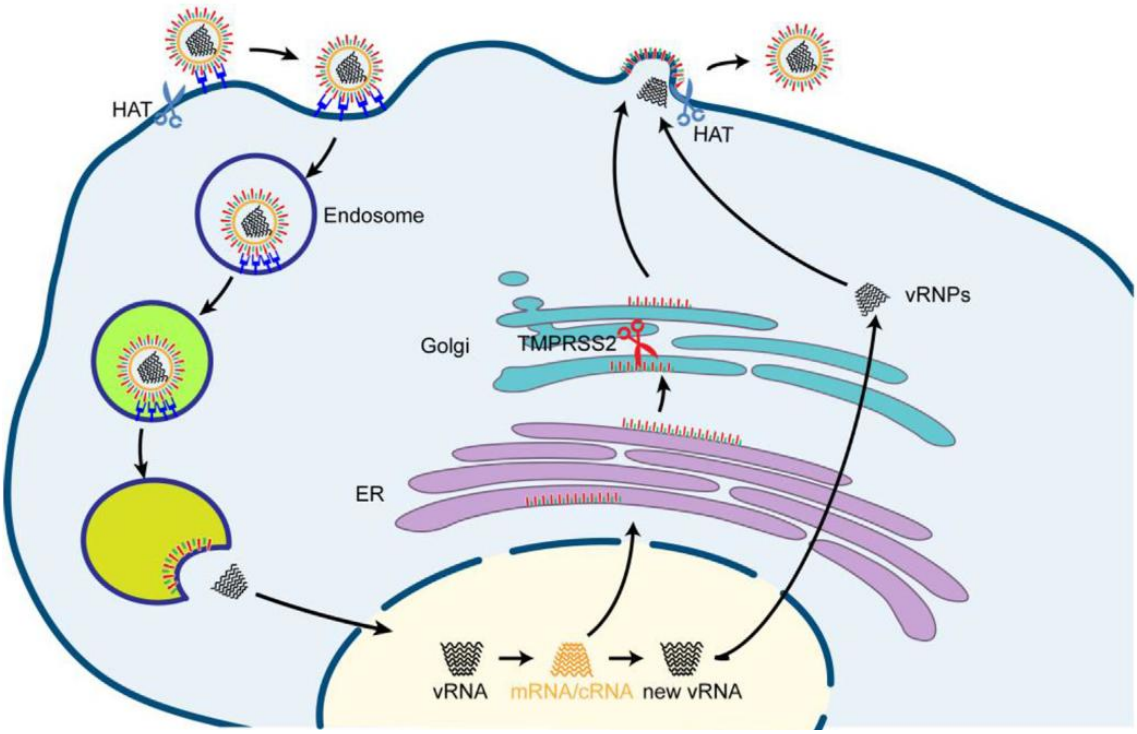


Figura 2. El ciclo de replicación del virus de la influenza y la escisión proteolítica de las proteasas del huésped. La HA se sintetiza como precursor que requiere escisión. La escisión de HA por proteasas unidas a la membrana (indicadas como tijeras) puede tener lugar en diferentes partes y en diferentes puntos de tiempo durante el ciclo de vida viral. Figura tomada de Shen et al. 2017.

En un estudio realizado en 2006 se observó que las células MDCK, las cuales no tienen proteasas endógenas capaces de activar la HA del virus de la influenza, permitieron la replicación multicíclica del virus de la influenza A tras ser transfectadas con plásmidos que codifican la proteasa TMPRSS2, lo que indica que la presencia de estas proteasas es crucial para la infección del virus proponiendo su estudio para el desarrollo de tratamientos antivirales y estrategias de prevención (Böttcher *et al.,* 2006). Posteriormente, se realizaron dos estudios que respaldaron el papel de la proteasa

TMPRSS2 en la activación y propagación de los virus de influenza A H1N1, H3N2 y H7N9 en modelos murinos. Los resultados mostraron que los ratones *knock out* de TMPRSS2 lograron sobrevivir a la infección con el virus de la influenza A, las cuales resultaron letales en los ratones de tipo salvaje (Hatesuer et al., 2013; Sakai et al., 2014).

II.1.2.2. Coronavirus

La enfermedad por coronavirus (COVID-19), causada por el síndrome respiratorio agudo severo coronavirus 2 (SARS-CoV-2), se ha propagado rápidamente desde su primera detección en Wuhan, China, a finales de 2019 (Paul et al., 2022). En este contexto, los pacientes infectados pueden experimentar desde síntomas moderados, similares a los de la gripe, hasta infecciones graves que incluyen dificultad respiratoria aguda, complicaciones cardíacas, insuficiencia renal y shock séptico, los cuales pueden conducir a la muerte o provocar complicaciones de salud a largo plazo (Alimohamadi et al., 2020).

Durante la infección por coronavirus, la entrada del virus en las células huésped depende de la interacción de su proteína Spike (S), cuya estructura y mecanismo de acción presentan similitudes con la hemaglutinina (HA) del virus de la influenza. En la Figura 3 se muestra el ciclo replicativo del coronavirus, cuya entrada celular ocurre por dos rutas. La primera está mediada por catepsina en el endosoma (ruta 1), donde el virus se une al receptor celular, es internalizado y activado por la catepsina. El descenso del pH facilita la fusión de la envoltura viral con la membrana endosómica, liberando el material genético al citosol para su transcripción, replicación y ensamblaje en el retículo endoplasmático y el Golgi. Los viriones formados se liberan mediante vesículas (Shen et al., 2017). La segunda ruta es facilitada por TMPRSS2 en la superficie celular. En esta vía, la subunidad S1 de la proteína Spike (S) se une a la enzima convertidora de angiotensina 2 (ACE2), que actúa como receptor principal para la entrada del virus.

Posteriormente, el procesamiento proteolítico por la serinoproteasa celular TMPRSS2 es fundamental para el cebado de la proteína S, un paso crucial que permite la fusión de las membranas y marca el inicio de la infección viral (Shen et al., 2017; Widiasta et al., 2020; Matsuyama et al., 2020; Fraser et al., 2021).

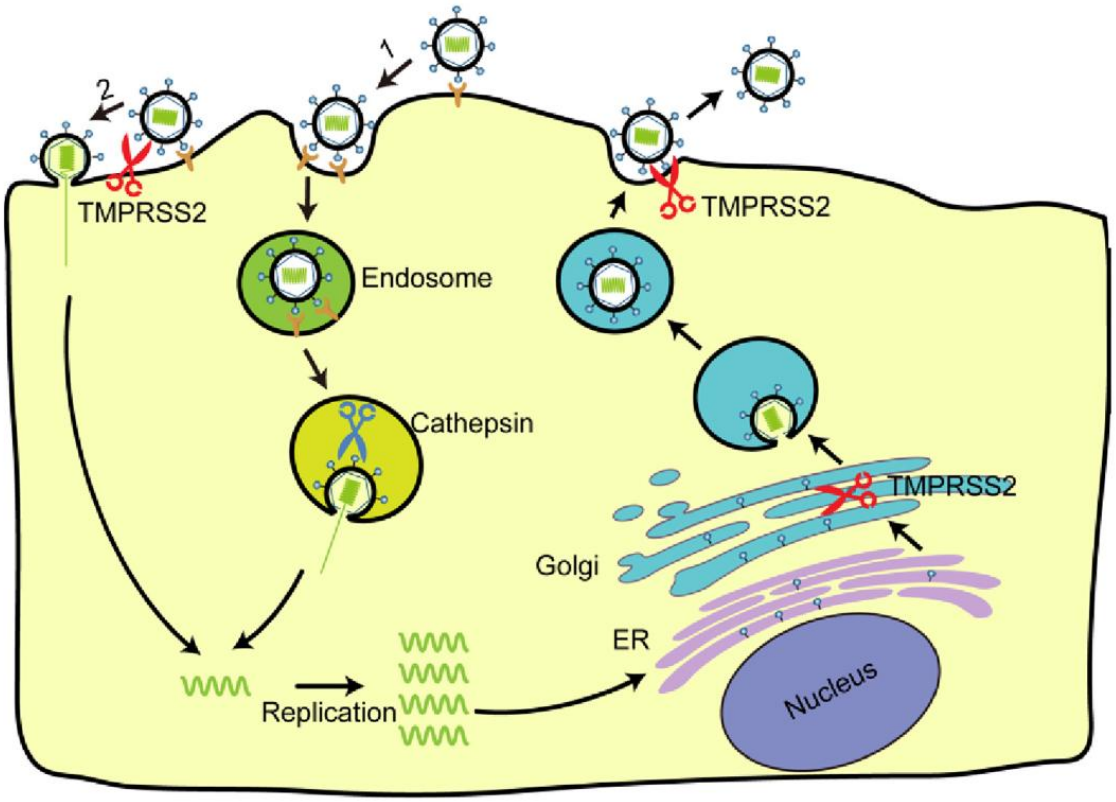


Figura 3. El ciclo de replicación del coronavirus y la escisión proteolítica de las proteasas del huésped. La proteína de la espiga se sintetiza como un precursor inactivo que requiere escisión por proteasas del huésped en distintos momentos y lugares del ciclo viral. La catepsina realiza esta escisión en el endosoma, mientras que TMPRSS2 lo hace en el aparato de Golgi o en la membrana plasmática. Figura tomada de Shen et al. 2017.

Hasta la fecha, diversos estudios realizados en modelos animales han confirmado la función de TMPRSS2 en infecciones por coronavirus *in vivo*. En un estudio reciente, se usaron ratones C57BL/6 con predisposición a respuestas Th1 para la infección por SARS-CoV, ratones transgénicos que expresaban el receptor humano MERS-CoV DPP4 (ratones hDPP4-Tg) para la infección por MERS-CoV y ratones *knock-out* para TMPRSS2 en cada experimento. Los resultados mostraron que los

ratones *knock-out* para TMPRSS2 presentaron una pérdida de peso corporal limitada y síntomas más leves en comparación con los controles, lo que reafirma que TMPRSS2 cumple un rol crucial en la propagación viral dentro de las vías respiratorias de modelos murinos infectados por MERS-CoV y SARS-CoV (Iwata-Yoshikawa et al., 2019).

II.1.3. *Mus musculus* como modelo experimental de virulencia

El modelo de *Mus musculus* podría usarse para detectar fármacos con potencial inhibitorio de la proteasa TMPRSS2. Además, los principios fundamentales de sustitución, reducción y refinamiento del uso de animales de laboratorio, 3R, podrían lograrse utilizando el ratón como alternativa a otros modelos animales de mayor complejidad biológica (Ekstrand et al., 2021).

Los ratones se han consolidado como el modelo animal de infección más utilizado en investigación biomédica debido a su rápido ciclo reproductivo, tamaño reducido y una amplia disponibilidad (Sun et al., 2020). Dentro de los estudios en ratones, Menéndez (2005) describió cómo la cepa del virus influenza tipo A infectó líneas de ratones *Balb/C*, dentro de los resultados se observó cambios clínicos y virológicos en los animales infectados; por lo que los autores concluyeron que este biomodelo es válido para el estudio de la gripe humana en ratones. Así mismo, Tarnow et al. (2014) demostraron que la eliminación de TMPRSS2 protege a los ratones *knockout* (KO) de la infección pulmonar con los virus H7N9 y H1N1 siendo letal en ratones de tipo salvaje (WT). Estos datos identifican que la TMPRSS2 es fundamental para la patogenicidad de los virus H7N9 y H1N1 en ratones, y podrían contribuir en el desarrollo potencial de inhibidores de proteasas específicos como un novedoso enfoque terapéutico contra la influenza. Diversos estudios relacionados a la infección por influenza señalan a la proteasa TMPRSS2 clave para el desarrollo de potenciales inhibidores (Hatesuer et al., 2013; Sakai et al., 2014). El beneficio de inhibir TMPRSS2 de las vías respiratorias es que, además del virus de la influenza, se podría atacar la infección de coronavirus

humanos, incluidas las variantes del SARS y MERS, en el que la escisión de TMPRSS2 activa la proteína Spike (S) en la superficie, lo que permite la entrada de la célula hospedera (Laporte y Naesens, 2017). En un estudio realizado por Zhou et al. (2015), tuvieron como resultado que el fármaco sintético camostat tiene actividad antiviral en ratones hembra *Balb/c* infectadas por SARS-CoV, demostrando que la serinaproteasa se encuentra involucrada en la infección viral *in vivo*. Los ratones, por naturaleza, no son susceptibles a la infección por SARS-CoV-2 debido a la baja afinidad del dominio de unión al receptor de la proteína S con el ACE2. Sin embargo, para superar esta limitación, se desarrolló un modelo alternativo que utiliza la transducción mediante un vector adenoviral para introducir el transgén humano ACE2 en los ratones. Esta alternativa sensibiliza a los animales, lo que permite que al infectarse con SARS-CoV-2 reproduzcan síntomas característicos de la enfermedad pulmonar COVID-19, como neumonía caracterizada por pérdida de peso, patología pulmonar grave (Sun et al., 2020). Posteriormente, un estudio usó al inhibidor nafamostat de TMPRSS2 en ratones transgénicos que expresan el gen ACE2 humanizado, debido a la baja afinidad de S por ACE2 en ratones, los resultados señalaron que dicho principio redujo la gravedad de la infección por SARS-CoV-2 (Li et al., 2021).

II.1.4. Potenciales inhibidores de TMPRSS2

II.1.4.1. Mesilato de camostato

El mesilato de camostat fue sintetizado por Ono Pharmaceutical hace 45 años. Este fármaco ha sido aprobado en Japón como tratamiento para la pancreatitis y la esofagitis por reflujo postoperatorio (Mantzourani et al., 2022). Se ha demostrado que el mesilato de camostat funcionó como inhibidor de TMPRSS2 en modelos murinos que fueron infectados con el virus de la influenza A, ya que mostraron una tasa de secreción viral del 45% en contraste con el grupo control (90%), por lo que el inhibidor fue efectivo

en reducir la replicación viral en ratones. Además, los resultados *in vitro* mostraron una eficacia contra infecciones de los virus de influenza A en células MDCK (Lee et al., 1996)

También se obtuvieron datos promisorios al evaluar la eficacia de este inhibidor de serinoproteasas en ratones *BALB/c* infectados con el virus SARS-CoV cuyo objetivo era determinar si el camostat era capaz de reducir tanto la patogénesis como la propagación del virus en condiciones *in vivo*. Los resultados mostraron que camostat protegió eficazmente a los ratones frente a la infección letal, logrando una tasa de supervivencia cercana al 60% (Zhou et al., 2015).

Según Shrimp et al. (2020), se determinaron las concentraciones necesarias del fármaco camostat para lograr una inhibición del 50% de la actividad de TMPRSS2 (IC_{50} = 6,2 nM), demostrando una mayor potencia en comparación con otros fármacos como Gabexate y bromhexina hidrocloruro (BHH), considerados menos eficaces en esta función. Por otro lado, Lee et al. (2021) reportaron que, en células Calu-3, las concentraciones de camostat requeridas para reducir la infección por tres linajes del SARS-CoV-2 mostraron una menor potencia antiviral, con valores de EC_{50} que oscilaron entre 0,102 y 0,170 µM.

II.1.4.2. Mesilato de nafamostat

El FUT-175, también conocido como nafamostat, es un inhibidor de serinoproteasas que fue sintetizado por Fujii en 1981. Este compuesto ha sido utilizado principalmente en el tratamiento de diversas condiciones como la coagulación intravascular diseminada, la pancreatitis aguda y el síndrome de respuesta inflamatoria sistémica, dado que suprime la actividad de diversas enzimas, entre ellas la trombina, la plasmina, la calicreína y la tripsina (Mantzourani et al., 2022).

Un estudio realizado por Li et al. (2021) empleó nafamostat como inhibidor de TMPRSS2 en un modelo murino transgénico que expresaba el gen ACE2 humanizado. Este enfoque se adoptó debido a la baja afinidad de la proteína S del SARS-CoV-2 por

el ACE2 de los ratones. De manera significativa, los resultados señalaron que nafamostat logró reducir la gravedad de la infección por SARS-CoV-2 en estos animales, lo que sugiere su potencial terapéutico en el contexto de la COVID-19.

Sin embargo, aunque nafamostat es ampliamente utilizado en Japón y Corea del Sur para tratar la pancreatitis aguda y la coagulación intravascular diseminada, aún persiste una limitación significativa, la escasez de estudios clínicos que respalden y validen de manera concluyente los resultados obtenidos en investigaciones *in vitro* (Hernández-Mitre et al., 2022).

Según Shrimp *et al.* (2020), se determinaron las concentraciones necesarias del fármaco nafamostat para lograr una inhibición del 50% de TMPRSS2 (IC_{50} = 0,27 nM), destacándose como el más efectivo en comparación con el camostat (IC_{50}= 6,2 nM). Por otro lado, Lee et al. (2021) reportaron que, en células Calu-3, los valores de EC_{50} de nafamostat para reducir la infección por SARS-CoV-2 oscilaron entre 0,016 y 0,024 µM. Estos resultados indican que nafamostat es altamente efectivo contra el SARS-CoV-2, ya que requiere concentraciones muy bajas para lograr un efecto antiviral significativo.

II.1.4.3. Lopinavir y valrubicin

Lopinavir es un inhibidor de proteasa aprobado por la FDA para el tratamiento de la infección por el virus de la inmunodeficiencia humana (VIH-1), en combinación con el fármaco ritonavir (Choy et al., 2020). Por otro lado, valrubicin es un análogo semisintético de la doxorubicina, un agente quimioterapéutico que interfiere con la síntesis de ADN y ARN, lo cual puede impactar la replicación viral (Baby *et al.,* 2021).

Un estudio de acoplamiento molecular demostró que tanto lopinavir como valrubicin poseen un potencial de inhibición dual, al prevenir la entrada del SARS-CoV-2 en células huésped. Específicamente, se identificaron interacciones tridimensionales de ambos compuestos con residuos clave de aminoácidos de la proteasa TMPRSS2. Sin embargo, para avanzar en la reutilización de estos fármacos, se requieren estudios

adicionales, tanto *in vitro* como *in vivo*, que permitan explorar su viabilidad clínica. Asimismo, los resultados derivados de las variaciones de desviación cuadrática media (RMSD) mostraron que las interacciones entre el ligando y la proteína permanecieron estables, con desviaciones dentro de un rango de 2,0 Å. Por lo tanto, los autores concluyen que estos inhibidores podrían actuar como posibles agentes duales contra TMPRSS2. Además, enfatiza la necesidad de realizar investigaciones más profundas mediante modelos experimentales *in vitro* e *in vivo*, con el fin de validar las predicciones realizadas *in silico* (Baby et al., 2021).

II.2. Antecedentes del estudio

Según Kishk et al. (2020), a partir de las plantillas 5CE1_A, 1Z8G_A, 6KD5_B y 6O1G_A se construyó la estructura tridimensional del TMPRSS2 humano mediante SWISS-MODEL. Posteriormente, se realizaron evaluaciones de validación mediante los gráficos de Ramachandran, Verify 3D y análisis estadístico de proteínas (ProSA). Así mismo, se emplearon simulaciones de dinámica molecular para investigar la estabilidad del modelo construido. Además, se realizó el acoplamiento de inhibidores de TMPRSS2, camostat, nafamostat, gabexato y sivelestat, en el modelo construido y los complejos proteína-ligando se sometieron a simulaciones MD y cálculos computacionales de afinidad de unión. Los resultados indican que nafamostat tendría una mejor afinidad de unión, lo que respalda las observaciones de RMSD proteína-ligando.

De acuerdo con Sonawane et al. (2021), el modelado de homología de TMPRSS2 se realizó usando la plantilla calicreína plasmática humana (5TJX) y la hepsina (5CE1). El modelo se validó usando diferentes servidores como PROSA, PROCHECK y PDBsum. El modelo de homología verificado de TMPRSS2 que tuvo buena calidad se sometió además a minimización de energía y, finalmente, simulación de dinámica molecular (MD) de 500 ns. Así mismo, se realizó el acoplamiento molecular con inhibidores experimentalmente conocidos como mesilato de camostat, nafamostat

y clorhidrato de bromhexina (BHH) utilizando técnicas de modelado molecular. Los resultados indicaron que el mesilato de camostat y su análogo estructural nafamostat interactúan fuertemente con los residuos His296 y Ser441 presentes en la tríada catalítica de TMPRSS2, mientras que BHH se une a Ala386 junto con otros residuos.

La presente tesis tuvo como objetivo evaluar la afinidad entre potenciales inhibidores y el dominio serinoproteasa de TMPRSS2 de *Mus musculus,* similar a un estudio realizado por Ramakrishnan et al. (2021), quienes generaron la estructura tridimensional de la proteína TMPRSS2 humana a partir de la secuencia de aminoácidos del dominio catalítico, usando como plantilla la hepsina (PDB: 5CE1). La estructura modelada fue optimizada y validada con el gráfico de Ramachandran. Además, se realizó simulación de acoplamiento y dinámica molecular (MD) de mesilato de camostat, nafamostat y leupeptina con el TMPRSS2. El estudio de acoplamiento molecular reveló que la leupeptina forma fuertes interacciones con los aminoácidos clave Ser186, His41 y Asp180 de la tríada catalítica presentes en el sitio activo de TMPRSS2, al igual que las otras dos moléculas que carecen de interacciones Ser186 e His41, mientras que en las simulaciones de dinámica molecular de estos tres complejos, la leupeptina es altamente estable (RMSD y RMSF) y esta molécula también forma interacciones estables con los aminoácidos clave Ser186, His41 y Asp186 de TMPRSS2 durante las simulaciones MD de 100 ns.

II.3. Bases teóricas

II.3.1. Modelado molecular

El modelado molecular de proteínas es una herramienta computacional clave en bioinformática y biología estructural, ya que permite predecir la estructura tridimensional de una proteína objetivo a partir de su secuencia de aminoácidos. Esta metodología es fundamental para comprender tanto su funcionamiento como las interacciones de éstas en los organismos vivos. Además, desempeña un papel crucial en la aceleración del

descubrimiento de nuevos fármacos y terapias, lo que resulta particularmente relevante en la industria farmacéuticas, ya que contribuye a disminuir significativamente los costos y el tiempo necesarios en los procesos de investigación y desarrollo (Adelusi et al., 2022).

Diversos estudios previos han modelado la estructura tridimensional de proteínas de ratón, pez cebra y lombriz de tierra empleando como plantillas estructuras cristalinas de proteínas humanas obtenidas por difracción de rayos X. Estos trabajos se han enfocado en la identificación de posibles dianas farmacológicas asociadas a enfermedades como el cáncer, la psoriasis, la cardiotoxicidad, y diversos trastornos neurodegenerativos (Bonifasius et al., 2019; Tripathi et al., 2019; Subaraja et al., 2019; Pandey-Rai et al., 2022).

II.3.1.1. Modelado por homología

Lesk (2014) explica en su libro *Introduction to Bioinformatics* que el modelado por homología predice la estructura tridimensional de una proteína objetivo a partir de su secuencia conocida (plantilla candidata), siempre que esté relacionada con al menos una proteína cuya secuencia y estructura hayan sido previamente determinadas. Por otro lado, cuando las proteínas presentan una relación estrecha, las estructuras conocidas pueden emplearse como base para construir el modelo de la proteína objetivo. En consecuencia, la calidad del modelo depende directamente del grado de similitud entre las secuencias; en este sentido, una identidad superior al 40-50% generalmente garantiza una precisión adecuada para diversas aplicaciones biológicas.

SWISS-MODEL es un servidor bioinformático desarrollado por T. Schwede, M.C. Peitsch y N. Guex, diseñado para facilitar el modelado por homología. Este servidor dispone de una biblioteca de plantillas que se actualiza constantemente según las estructuras almacenadas en el Protein Data Bank (PDB). A través de SWISS-MODEL,

se puede ingresar la secuencia de aminoácidos de una proteína desconocida (proteína blanco), identificar proteínas homólogas con estructuras conocidas en el PDB y, como resultado, obtener las coordenadas necesarias para predecir su estructura tridimensional (Lesk, 2014).

Durante el proceso de modelado, se proporcionan estimaciones de calidad, como el porcentaje de identidad y cobertura entre la secuencia a modelar y las diferentes plantillas. Posteriormente, los modelos generados son evaluados mediante la métrica QMEAN, una función de puntuación basada en potenciales estadísticos que compara las características energéticas y geométricas del modelo con estructuras proteicas experimentales. Esta métrica permite valorar la calidad global y local del modelo, donde valores del QMEAN entre 0 y –4.0 se consideran aceptables. Una de las principales ventajas de este servidor es su capacidad para generar modelos tridimensionales en pocos minutos, ofreciendo además gráficos de alta calidad para su visualización (Arreola-Barroso et al., 2022).

II.3.1.2. Modelado por *threading*

Según Lesk (2014), el *threading* es un método para predecir la estructura tridimensional de proteínas basado en el reconocimiento de plegamientos. Este enfoque se aplica cuando la similitud de la secuencia entre la proteína objetivo, cuya estructura es desconocida, y las estructuras conocidas (contenidas en bibliotecas como el PDB) es menor al 40%.

El principal objetivo del *threading* es generar múltiples modelos aproximados de la proteína objetivo. Esto se logra al comparar las estructuras conocidas con la secuencia de la proteína objetivo, utilizando diversos alineamientos entre ambas. Cabe destacar que, a diferencia del modelado por homología, el *threading* se caracteriza por explorar numerosos alineamientos y evaluar modelos aproximados, mientras que el

modelado por homología se centra en un único alineamiento óptimo con el fin de construir un modelo más detallado (Lesk, 2014).

Phyre2, desarrollado por el Grupo de Bioinformática Estructural del Imperial College London, es una herramienta que predice estructuras proteicas a partir de secuencias de aminoácidos mediante comparación de secuencias y estructuras conocidas. En primer lugar, al recibir una secuencia de proteína, Phyre2 genera un perfil evolutivo que identifica relaciones con otras proteínas. Este perfil permite buscar plantillas adecuadas en una base de datos de estructuras, facilitando así el modelado. Posteriormente, la calidad de los modelos se evalúa utilizando parámetros como el "confidence score", que mide la fiabilidad de la predicción, y el porcentaje de identidad, que refleja la similitud entre la proteína objetivo y las plantillas empleadas (Kelley et al., 2015).

II.3.1.3. Modelado *ab initio*

Según Kihara (2014), el modelado *ab initio* es un enfoque clave para predecir la estructura terciaria de proteínas en ausencia de plantillas estructurales disponibles en el PDB. Este método utiliza únicamente la información de la secuencia aminoacídica de la proteína y recurre a algoritmos basados en principios físicos y datos experimentales. El proceso implica la generación de múltiples conformaciones preliminares, conocidas como decoys, mediante una función de energía diseñada. Posteriormente, se selecciona el modelo final evaluando las conformaciones según criterios como la energía libre de Gibbs y su compatibilidad con datos experimentales. Este enfoque resulta fundamental para construir modelos tridimensionales en situaciones donde no existen estructuras homologadas, contribuyendo significativamente al avance de la bioinformática estructural.

El servidor Robetta ofrece herramientas totalmente automatizadas para la predicción y análisis de la estructura de proteínas. Las secuencias que se suben al

servidor son analizadas para identificar dominios putativos, y a partir de esta información, se generan modelos estructurales combinando diferentes métodos de predicción, como el modelado *de novo* y el modelado comparativo (basado en homología o *threading*). En el caso de que se identifiquen proteínas homólogas con estructuras ya conocidas, estas se emplean como plantillas para el modelado comparativo. Sin embargo, si no se encuentra una homología, el servidor recurre al método de predicción de estructura *de novo* de Rosetta, el cual permite explorar el espacio conformacional mediante la inserción de fragmentos, lo que produce un gran conjunto de estructuras preliminares (Chivian et al., 2003; Kim et al., 2004).

II.3.1.4. Alphafold

AlphaFold es una herramienta desarrollada por DeepMind para la predicción de estructuras tridimensionales de proteínas a partir de sus secuencias de aminoácidos, incluso sin una estructura similar conocida. En este sentido, utiliza redes neuronales profundas y un enfoque basado en aprendizaje automático para inferir dichas estructuras. Para lograrlo, incorpora conocimiento físico y biológico sobre la estructura proteica; aprovechando alineamientos multisecuencia (MSA) (Jumper et al., 2021; Mirdita et al., 2022).

AlphaFold se beneficia de un extenso conjunto de datos de entrenamiento basado en proteínas resueltas experimentalmente. Además, fue entrenado y evaluado en el contexto del conjunto de datos del CASP (Critical Assessment of Structure Prediction), un estándar en la evaluación de predicción de estructuras de proteínas. Cabe destacar que, en dichas pruebas, obtuvo resultados significativamente superiores en comparación con otros enfoques de predicción de estructuras de proteínas (Jumper *et al.,* 2021).

Asimismo, ColaFold es una herramienta derivada de AlphaFold que está disponible en Google Colaboratory (Colab), un servicio en línea gratuito para la

ejecución de código y análisis de datos mediante Jupyter Notebook. A través de esta plataforma, los usuarios pueden subir secuencias de proteínas y obtener predicciones de su estructura tridimensional utilizando el algoritmo AlphaFold (Mirdita et al., 2022).

II.3.2 Docking molecular

El docking molecular es una técnica computacional consolidada como una herramienta primordial en la investigación biomédica y el diseño de fármacos. El principal objetivo es predecir y calcular la pose más favorable de interacción entre un ligando, como un fármaco, y un blanco molecular, que suele ser una proteína con una estructura tridimensional conocida. Para lograr este objetivo, los métodos de docking efectivos analizan de manera exhaustiva todas las configuraciones y orientaciones posibles que el ligando puede asumir al interactuar con la proteína. Posteriormente, las configuraciones se evalúan mediante funciones de puntuación que asignan valores de energía libre de Gibbs (ΔG) a cada configuración, en el cual valores bajos de ΔG indican mayor estabilidad del sistema y, por ende, una interacción más favorable entre el ligando y la proteína (Morris y Lim-Wilby, 2008).

Según Morris y Wilby (2008), la mayoría de los protocolos de acoplamiento molecular siguen un esquema de trabajo común, como se ilustra en la Figura 4. En primer lugar, se seleccionan las estructuras tridimensionales tanto del ligando como de la proteína. Posteriormente, se realiza la preparación de ambos componentes para asegurar su acoplamiento en la disposición óptima. Este proceso incluye etapas como la adición de hidrógenos, la eliminación de moléculas de agua no esenciales y la protonación de grupos funcionales, entre otras modificaciones necesarias. Luego del docking, los resultados son analizados y se selecciona las poses con mejores valores de energía.

La teoría del candado y llave, propuesta por Fischer en 1984, explica la unión entre un ligando y su receptor como un encaje específico entre estructuras rígidas, sirviendo de base para los primeros métodos de acoplamiento molecular. Sin embargo, este modelo no logró describir adecuadamente muchos eventos moleculares observados, lo que llevó al desarrollo de un enfoque dinámico. Este nuevo modelo plantea que el sitio activo de la proteína se adapta al ligando durante la interacción, reconociendo la flexibilidad de ambas moléculas. Este avance mejoró la precisión en la descripción de las interacciones moleculares, destacando la importancia de la flexibilidad en el acoplamiento molecular (Meng et al., 2011).

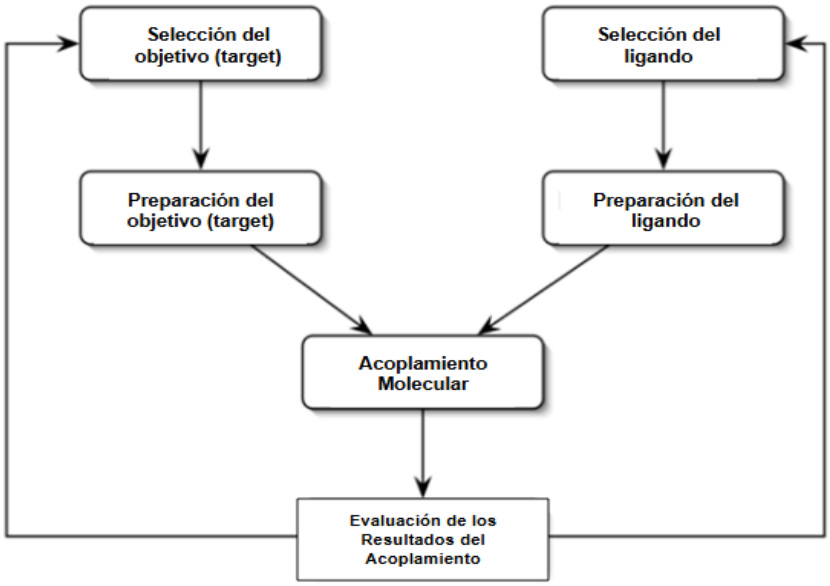


Figura 4. Esquema de trabajo propuesto por Morris y Wilby (2008) para la metodología de Acoplamiento Molecular.

AutoDock es una herramienta de código abierto y gratuito para el acoplamiento proteína-ligando, la cual permite predecir cómo las proteínas interactúan con moléculas pequeñas, como fármacos o sustratos. Para lograr este objetivo, AutoDock ajusta la posición del ligando dentro del sitio activo de la proteína y estima con precisión la energía libre de unión entre ambos. Para ello, emplea el algoritmo genético Lamarckiano

(LGA), el cual combina dos estrategias de búsqueda complementarias. Primero la búsqueda global que se lleva a cabo mediante un algoritmo genético que imita el proceso de selección natural, permitiendo encontrar la mejor pose del ligando en la proteína. Luego, la búsqueda local se realiza a través del algoritmo de Solis & Wets, el cual ajusta la precisión y orientación de las mejores poses con el fin de mejorar su energía de enlace. Además de estas ventajas, AutoDock ofrece una funcionalidad clave para definir el área de la proteína donde se espera que ocurran las interacciones, a través de la opción GridBox. De esta manera, no solo permite enfocar el estudio en una región de interés, sino que también mejora la calidad de los resultados al limitar el espacio de búsqueda a un área específica, lo que incrementa la relevancia de las interacciones predichas. Finalmente, el análisis en AutoDock genera como resultado las diez mejores conformaciones, el número de enlaces involucrados en la interacción y la constante de inhibición (Ki) (Ravi, 2016).

II.3.3. Dinámica Molecular

Karplus y McCammon (2002) señalan que la simulación de Dinámica Molecular (DM) permite predecir y estudiar el comportamiento de cada átomo dentro de las macromoléculas biológicas, como las proteínas, a lo largo del tiempo. En este sentido, este método ha proporcionado numerosos conocimientos sobre los movimientos internos de estos sistemas. De esta manera, se ha podido revelar que las proteínas no son estructuras rígidas, sino sistemas dinámicos cuyas interacciones y movimientos son fundamentales para su función biológica. Además, según Hollingsworth y Dror (2018), dichas simulaciones son capaces de captar una amplia gama de procesos biomoleculares, como cambios conformacionales, la unión de ligandos y el plegamiento de proteínas. Además, revelan la posición de cada átomo en intervalos de tiempo tan breves como femtosegundos (fs) o nanosegundos (ns), lo que permite registrar eventos ultrarrápidos con alta precisión y realizar un análisis detallado de la dinámica molecular.

Asimismo, es importante destacar que estas simulaciones también pueden predecir cómo las biomoléculas responden, a nivel atómico, ante perturbaciones como, fosforilación, mutaciones, protonación, o la eliminación o adición de un ligando.

Hace 25 años, se publicó la primera simulación de dinámica molecular de una macromolécula biológica. Esta simulación se centró en el inhibidor de tripsina pancreática bovina (BPTI), que es un modelo ideal para estudiar las dinámicas proteicas, gracias a su pequeño tamaño, alta estabilidad y estructura precisa obtenida por cristalografía de rayos X en 1975. A pesar de que la simulación se realizó en condiciones de vacío, con un modelo rudimentario de mecánica molecular y solo duró 9,2 ps, los resultados fueron cruciales para cambiar la visión de las proteínas, pasando de ser vistas como estructuras rígidas a ser consideradas como sistemas dinámicos, cuyas interacciones internas cumplen una función esencial (Karplus y McCammon, 2002).

La idea básica de las simulaciones de dinámica molecular (MD) radica en calcular las posiciones de los átomos dentro de un sistema biomolecular, como una proteína rodeada de agua o una bicapa lipídica, considerando las fuerzas que ejercen entre sí, tales como los enlaces químicos, las fuerzas de Van der Waals y las interacciones electrostáticas (Hollingsworth y Dror, 2018). Para ello, se utilizan las leyes del movimiento de Newton, lo que permite predecir el desplazamiento de cada átomo a lo largo del tiempo. El proceso se lleva a cabo mediante dos pasos iterativos: en primer lugar, se calcula la fuerza neta que actúa sobre cada átomo, sumando las contribuciones de las fuerzas ejercidas por los átomos cercanos; en segundo lugar, con esta información, se actualizan las posiciones y velocidades de los átomos para el siguiente instante de tiempo. A través de estas iteraciones, se puede simular cómo los átomos interactúan y se reorganizan a lo largo del tiempo, generando una representación tridimensional que ilustra cómo cambian las posiciones atómicas durante un intervalo

de tiempo específico. Este enfoque permite observar procesos clave, como el plegamiento de proteínas, la unión de ligandos o los cambios conformacionales en respuesta a diferentes estímulos (Hollingsworth y Dror, 2018).

La Figura 5 describe el proceso de análisis estructural y dinámico de proteínas, dividido en dos etapas: exploración conformacional y refinamiento del modelo. En la primera etapa, se parte de una estructura inicial obtenida experimentalmente o por modelado, la cual se somete a simulaciones de dinámica molecular para explorar su flexibilidad y las posibles trayectorias conformacionales. En la segunda etapa, el modelo inicial, que puede estar en un estado de alta energía, se somete a un proceso de refinamiento, incorporando restricciones experimentales o ajustes computacionales para reducir su energía. Como resultado, se obtiene un modelo refinado, más estable y representativo de la conformación funcional de la proteína. Este enfoque es clave para yqwentender la dinámica y estabilidad de las proteínas, con aplicaciones en diseño de fármacos y bioingeniería (Hollingsworth y Dror, 2018).

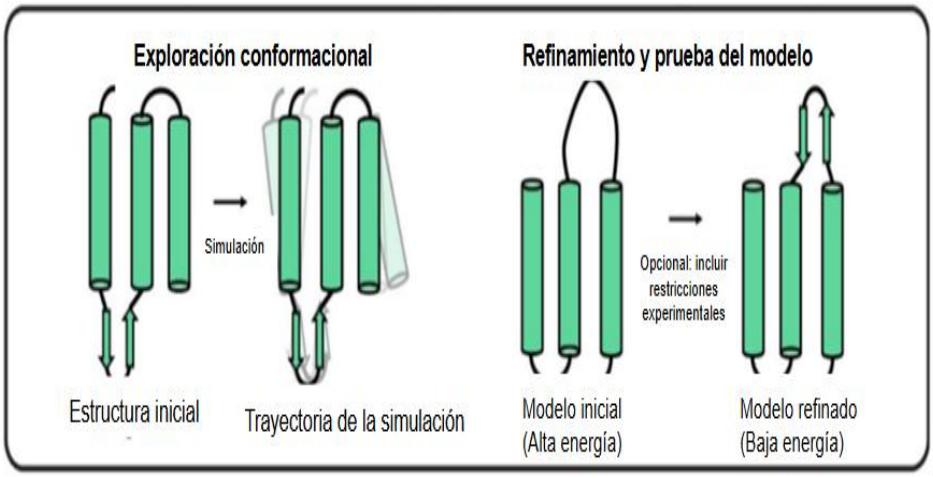


Figura 5. Esquema del estudio estructural y dinámico de una proteína. Adaptado de Hollingsworth y Dror, 2018.

II.3.3.1. RMSD (Desviación cuadrática media)

Según Arcón (2018), la RMSD es un parámetro utilizado para medir cambios en la posición relativa de los átomos entre dos conformaciones distintas de la misma molécula. La RMSD representa el desplazamiento promedio de los átomos con respecto a una posición inicial o de referencia en un instante dado y permite observar cómo evolucionan las coordenadas de una molécula respecto al punto de partida.

II.3.3.2. RMSF (Fluctuación cuadrática media)

Según Sepúlveda Durán (2007), la RMSF mide la flexibilidad individual de los residuos de una estructura, es decir, cómo se mueve un residuo particular durante la simulación. Son las fluctuaciones atómicas en función de una estructura de referencia, lo que ayuda a identificar cuáles regiones de una proteína son más estables o flexibles durante la simulación.

II.4 Definición de términos

TMPRSS2:	Transmembrane Protease Serine 2
SP	Serinoproteasa
ACE2:	Angiotensin-Converting Enzyme 2
HA:	Hemaglutinina
S:	Spike protein
PDB:	Protein Data Bank
MDS:	Simulación de dinámica molecular
RMSD:	Desviación cuadrática media
RMSF:	Fluctuación cuadrática media
ΔG:	Energía libre de Gibbs

III. HIPÓTESIS Y VARIABLES

III.1. Hipótesis

El modelamiento tridimensional del dominio serinoproteasa de TMPRSS2 de *Mus musculus* presenta afinidad con los inhibidores mediante herramientas computacionales.

III.2. Variables

III.2.1. Variable independiente

Modelos *in silico* de TMPRSS2 de *Mus musculus*.

III.2.2. Variable dependiente

Afinidad de unión y estabilidad del complejo proteína–ligando

III.3. Operacionalización de las variables

Variable	Definición conceptual	Dimensiones	Indicadores
Variable independiente: Conformación estructural del modelo del dominio serinoproteasa de TMPRSS2.	Representación tridimensional del dominio serinoproteasa de TMPRSS2 de *Mus musculus*, obtenida mediante diferentes metodologías de modelado estructural.	-Variabilidad según metodología de modelado.	-Homología (SWISS-MODEL), *Threading* (Phyre2), *Ab initio* (Robetta), AlphaFold
		-Calidad del modelo.	- % regiones permitidas (Ramachandran) y puntuación ModFOLD
		-Estabilidad estructural	- RMSD, RMSF
Variable dependiente: Afinidad de unión y estabilidad del complejo proteína–ligando.	Capacidad de los inhibidores para unirse al dominio serinoproteasa de TMPRSS2 y formar complejos estables.	-Afinidad de unión	- Energía de unión (kcal/mol).
		-Interacción proteína-ligando.	- Puentes de hidrógeno, número de contactos intermoleculares, así como las variaciones en las distancias interatómicas de la tríada catalítica y la proximidad de los

			ligandos a los residuos del sitio activo.
		-Estabilidad del complejo.	- RMSD y RMSF, radio de giro (Rg),

IV. MATERIALES Y MÉTODOS

IV.1. Área de estudio

La presente tesis se realizó en un entorno computacional, mediante el uso de herramientas bioinformáticas y plataformas especializadas en análisis estructural, acoplamiento molecular y simulación de dinámica molecular. El procesamiento y análisis de datos se realizaron en una estación de trabajo (PC Workstation HP Z4) con sistema operativo Linux Ubuntu 24.04, equipada con un procesador Intel Xeon® W-2145 de 3.7 GHz (8 núcleos), 64 GB de memoria RAM y una tarjeta gráfica Nvidia Quadro P4000 de 8 GB.

Asimismo, se emplearon diversos recursos bioinformáticos de libre acceso, incluyendo programas, bases de datos y servidores web especializados, tales como UniProt, SWISS-MODEL, Phyre2, Robetta, ColabFold, AutoDock y GROMACS, los cuales permitieron la ejecución de las diferentes etapas del estudio. Todas las simulaciones y análisis computacionales fueron realizados en un entorno de sistema operativo Linux (Ubuntu).

IV.2. Diseño de investigación

Tipo de estudio: Computacional, analítico.

Enfoque metodológico: Cuantitativo, con alcance exploratorio-descriptivo.

Diseño específico: Modelado molecular, acoplamiento molecular y simulación de dinámica molecular *in silico*.

Nivel de investigación: Aplicado.

IV.3. Población y muestra

La población está formada por todos los posibles modelos estructurales tridimensionales del dominio serinoproteasa de TMPRSS2 de *Mus musculus*, generados mediante diversas técnicas de modelado molecular. Mientras que la muestra está conformada por los modelos estructurales de TMPRSS2 seleccionados que cumplieron con los criterios de filtro estereoquímico inicial (PDBsum y ModFOLD), filtro de estabilidad térmica (Dinámica molecular de equilibrio), y criterio funcional final (*Docking Molecular*).

Criterios de selección

✓ Inclusión:

- Modelos con una calidad geométrica global óptima, con un porcentaje mayor al 95% de residuos en regiones favorables y permitidas en el gráfico de Ramachandran.
- Modelos que muestran estabilidad durante la simulación de dinámica molecular de equilibrio de 10 ns, en el que previamente el sistema haya convergido en sus fases previas de minimización de energía, equilibrio térmico y equilibrio barométrico, y que evidencien una pronta convergencia de los valores RMSD del dominio serinoproteasa con fluctuaciones menores a 1.5 Å.
- Modelos que presentan afinidad termodinámica y complementariedad geométrica con los inhibidores de referencia y exploratorios evaluados mediante *docking molecular,* priorizando los complejos que presentan menores energías libres de unión.

✓ Exclusión:

- Modelos estructurales que presentan perfiles de error por residuo superiores a 2 Å en la zona del dominio catalítico en la validación local de ModFOLD.

- Modelos que no logran llegar a la convergencia durante la simulación de equilibrio en solvente explícito en los 10 ns.
- Complejos del dominio serinoproteasa TMPRSS2-ligandos resultantes del docking molecular que presenten energías de unión desfavorables o que carezcan de interacciones con la tríada catalítica.

IV.4. Procedimientos, Técnicas e instrumentos de recolección de información

IV.4.1. Recuperación de secuencia y estructura del gen

La secuencia aminoacídica de TMPRSS2 de *Mus musculus* (UniProt: Q9JIQ8) fue recuperada en formato FASTA desde la base de datos UniProt (https://www.uniprot.org/). Para el presente estudio se seleccionó exclusivamente el dominio serinoproteasa (SP) correspondiente a los residuos 254–487 de la proteína completa.

Los residuos catalíticos corresponden a His294, Asp343 y Ser439 respecto a la secuencia completa de TMPRSS2, son equivalentes a His41, Asp90 y Ser186 en la numeración correspondiente del dominio serinoproteasa modelado.

El FGENESH+ disponible por el servidor web Softberry (http://www.softberry.com/) (Solovyev, 2007) se utilizó para determinar la estructura del gen.

IV.4.2. Análisis de parámetros bioquímicos con ProtParam

El análisis de parámetros bioquímicos se realizó utilizando ProtParam (https://web.expasy.org/protparam/). Se ingresó el número de acceso de una entrada de Swiss-Prot/TrEMBL (Q9JIQ8), seleccionándose el dominio serinoproteasa. A partir de esta región se determinaron parámetros fisicoquímicos como la masa molecular, el punto isoeléctrico teórico (pI), la composición de aminoácidos, el índice de inestabilidad, el índice alifático y el valor GRAVY.

IV.4.3. Determinación de dominios conservados

La identificación y anotación de los dominios conservados se realizó utilizando las herramientas Prosite (https://prosite.expasy.org/) y NCBI Conversed Domains (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) (Lu et al., 2020; Sigrist et al., 2013).

IV.4.4. Identificación de proteínas que interactúan con TMPRSS2

Las interacciones proteína-proteína se predijeron utilizando el software STRING (https://string-db.org/). Dicha herramienta recopila, clasifica e integra todas las fuentes disponibles públicamente de información de interacción proteína-proteína, y permite complementarlas con predicciones computacionales (Szklarczyk et al., 2019).

IV.4.5. Alineamiento de secuencias

La secuencia aminoacídica correspondiente al dominio serinoproteasa de TMPRSS2 de *Mus musculus* y de otros modelos experimentales como cobayo, hámster, rata, conejo, humano, macaco, gato, perro y hurón fueron obtenidas a partir de NCBI (https://www.ncbi.nlm.nih.gov/) y alineadas mediante Clustal Omega (https://www.ebi.ac.uk/jdispatcher/msa/clustalo). Posteriormente, el alineamiento fue analizado para resaltar las regiones conservadas y similares, así como para superponer información de estructuras secundarias, hélices alfa y láminas beta, del dominio catalítico usando la herramienta en línea ESPript3 (https://espript.ibcp.fr/ESPript/ESPript/) (Robert y Gouet, 2014). También, mediante BLASTp se determinó el porcentaje de identidad entre las secuencias proteicas del humano y ratón.

IV.4.6. Predicción de estructuras secundarias

La secuencia del dominio serinoproteasa de TMPRSS2 de *Mus musculus* se utilizó para predecir su estructura secundaria utilizando la herramienta PHD secondary

structure prediction method (https://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_phd.html), el cual se basa en el alineamiento de múltiples secuencias de proteínas relacionadas e identificación de patrones conservados en PSI-BLAST (Rost et al., 1994).

Así mismo, la misma secuencia se subió al servidor PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred/) en formato de código de aminoácidos de una sola letra y luego se seleccionó el método de predicción PSIPRED, el cual se basa en el método de aprendizaje automático y la matriz de puntuación de posiciones específicas (PSSM) generadas por la búsqueda de secuencias homólogas en PSI-BLAST (Buchan y Jones, 2019).

IV.4.7. Predicción de estructuras terciarias mediante los métodos de:

Para la construcción de las estructuras tridimensionales del dominio serinoproteasa de TMPRSS2 de *Mus musculus*, se realizó un enfoque metodológico de validación cruzada basado en cuatro estrategias de modelamiento independiente, con el fin de disminiur los sesgos algorítmicos propios de las plataformas de predicción estructural. De tal forma, se pueda garantizar la selección de un modelo óptimo que pueda superar los filtros del docking molecular y de mantener la estabilidad termodinámica de los complejos proteína-ligando durante la dinámica molecular.

IV.4.7.1. Modelado por homología

Para el modelado de homología, se utilizó el servidor automatizado SWISS-MODEL accesible a través del servidor web ExPASy (https://swissmodel.expasy.org/). Debido al grado de conservación, se subió la secuencia aminoacídica del dominio serinoproteasa (SP) de la proteína TMPRSS2 (Q9JIQ8) como consulta para buscar estructuras de proteínas relacionadas evolutivamente en la biblioteca de plantillas SMTL de SWISS-MODEL. SWISS-MODEL realizó esta tarea utilizando dos métodos de

búsqueda en bases de datos como BLAST y HHblits (Waterhouse et al., 2018). Este método de modelamiento por homología es útil debido a que la secuencia objetico cuenta con plantillas estructurales evolutivamente cercanas depositadas en PDB database, las cuáles han sido cristalizados juntos a inhibidores, preservando la arquitectura molecular necesaria para el reconocimiento de inhibidores.

Se seleccionaron en total 2 plantillas y para cada una se generó automáticamente un modelo de proteína 3D. Para estimar la calidad esperada de los modelos antes de su construcción, en base a la calidad de la plantilla seleccionada y el alineamiento de secuencias entre la plantilla y la secuencia TMPRSS2, se usó la puntuación GMQE. La calidad del modelo construido se evaluó mediante varias puntuaciones: QMEANDisCo Global, QMEANDisCo Local y el QMEAN Z-score. El primero, analiza la calidad general del modelo en relación con estructuras experimentales de alta resolución del PDB. El segundo evalúa la calidad local de cada residuo, identificando áreas de mayor y menor confianza. El tercero contrasta la calidad del modelo con un conjunto de estructuras no redundantes del PDB (Waterhouse et al., 2018). Además, se evaluó la cobertura del alineamiento y el porcentaje de identidad entre las secuencias. Finalmente, se seleccionaron y descargaron los archivos PDB de 2 plantillas según la metodología y resolución del modelado.

IV.4.7.2. Modelado por *threading*

En el modelado por *threading* del dominio serinoproteasa de TMPRSS2 se utilizó el servidor web Phyre2 (http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index). Phyre2 alinea la secuencia de consulta con proteínas anotadas en base de datos como SCOP y PDB para predecir el modelo tridimensional de la proteína de interés. A diferencia del modelado por homología, *threading* funcionaría como un control de calidad cruzado y validación para demostrar la robustez en la selección del dominio serinoproteasa de TMPRSS2.

La selección del modelo estructural basado en 10 plantillas con mejores coincidencias se realizó considerando la confidencialidad y el porcentaje de identidad con las plantillas utilizadas (Kelley et al., 2015).

IV.4.7.3. Modelado *ab initio*

Se cargó la secuencia aminoacídica del dominio serinoproteasa de TMPRSS2 al servidor ROBETTA (http://robetta.bakerlab.org/), el cual recurre al modelado *ab initio* para la predicción de la estructura tridimensional. Se generaron 5 modelos preliminares, de los cuales se seleccionó el mejor modelo basado en la estimación de la posición de los residuos. El modelo seleccionado se descargó en formato PDB para su posterior análisis (Raman et al., 2009), éste modelo sirve como un indicador físico independiente, construido sin algún modelo previo basado.

IV.4.7.4. Alphafold

Finalmente, la secuencia anteriormente indicada se subió a ColabFold (https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb), un servicio en línea gratuito que permite cargar secuencias de proteínas y obtener la predicción de la estructura 3D a partir del alineamiento múltiple de secuencias (MSA) que se usa para entrenar el modelo de Alphafold, este método se incorporó como la referencia metodológica de vanguardia y de alta confianza biológica. ColabFold predijo 5 posibles modelos estructurales de TMPRSS2, se evaluaron los modelos basándose en la cobertura de secuencia, identidad de secuencia, e IDDT (Interface-Driven Docking Topology); que mide la confianza de los 5 modelos.

IV.4.8. Validación de modelos

Los modelos generados mediante las distintas técnicas computacionales fueron evaluados a través de un flujo secuencial de filtros de calidad, el cual incluyó la validación estereoquímica inicial, análisis de estabilidad por dinámica molecular y el criterio de compatibilidad funcional mediante docking molecular.

A continuación, se describe el primer filtro estereoquímico, en el cual se evalúa la calidad geométrica global de los modelos:

IV.4.8.1. PDBSum

La evaluación de la calidad de los modelos predichos se realizó utilizando los servidores PDBsum Generate (https://www.ebi.ac.uk/thornton-srv/databases/pdbsum/) y ModFOLD (https://www.reading.ac.uk/bioinf/ModFOLD/). Para validar la geometría de las estructuras predichas, se empleó el servidor web PDBsum Generate. Todos los modelos seleccionados se subieron en formato PDB. El análisis de Ramachandran se utilizó para estimar la calidad general del modelo (>95%) basándose en la posición de los residuos. Además, se considerará el factor G (>-0.5) como un parámetro adicional para evaluar la calidad del modelo (Laskowski et al., 2018).

IV.4.8.2. ModFOLD

El servidor web ModFOLD, por su parte, proporcionó una evaluación estadística de la calidad de cada modelo, incluyendo gráficos de errores locales e imágenes anotadas de los modelos 3D. Los resultados se presentaron en una tabla que incluía la clasificación y el ID del modelo, puntuaciones globales, una puntuación de confianza y un valor P, junto con miniaturas de los resultados gráficos (McGuffin et al., 2021).

IV.4.9. Simulación de dinámica molecular

El segundo criterio de evaluación de los modelos predichos fue el filtro de estabilidad termodinámica, el cual consistió en obtener modelos estables para el

procedimiento de acoplamiento molecular. Se ejecutó una simulación de dinámica molecular en los archivos PDB de todos los modelos predichos del dominio serinoproteasa de TMPRSS2 de *Mus musculus.* No obstante, los archivos obtenidos de *ab initio* y AlphaFold fueron optimizados con el servidor en línea MolProbity de Duke Biochemistry, EE. UU (http://molprobity.biochem.duke.edu/) mediante la adición de hidrógenos (Williams et al., 2018). Las simulaciones se realizaron en entorno Linux utilizando GROMACS versión 2023.1 y el campo de fuerza OPLS-AA para generar la topología del receptor y evaluar su estabilidad en un entorno acuoso (Kaminski et al., 2001). Se generó una caja cúbica definiendo una distancia mínima de 1.0 nm entre la superficie de la proteína y las paredes de la caja, según lo indicado por Lemkul (2018). Para analizar la estabilidad del dominio serinoproteasa de TMPRSS2 en un entorno acuoso se utilizó un modelo de solvente de carga puntual única (SPC/E), seguido de la adición de iones cloruro para neutralizar el sistema. La minimización de energía se realizó mediante el algoritmo de descenso más pronunciado con un máximo de 50 000 pasos.

Para iniciar la dinámica molecular real, el sistema minimizado se sometió a un paso de equilibrio de dos estapas. En la primera etapa, se realizó una simulación NVT (número de partículas, volumen y temperatura constantes) a 310 K utilizando el termostato V-rescale. Posteriormente, se efectuó una simulación NPT (número de partículas, presión y temperatura constantes) a 310 K y 1 bar empleando el barostato Parrinello-Rahman. Finalmente, se realizaron simulaciones de producción de 10 ns y 50 ns. La estabilidad estructural del dominio serinoproteasa de TMPRSS2 se evaluó mediante la desviación cuadrática media (RMSD) y la fluctuación cuadrática media (RMSF) respecto a la conformación inicial del sistema. El trazado de gráficos y la visualización de las trayectorias se realizaron utilizando Xmgrace y PyMOL, respectivamente.

IV.4.10. Acoplamiento molecular por AutoDock

El tercer filtro para la selección del modelo consistió en el acoplamiento molecular cruzado. En esta fase de selección, se evaluó los dos mejores modelos predichos del dominio serinoproteasa de TMPRSS2 con la finalidad de examinar su afinidad y complementariedad con inhibidores de referencia (nafamostat y camostat) e inhibidores exploratorios (lopinavir y valrubicina). Asimismo, se incluyó el nucleótido Adenosín Trifosfato (ATP) como control negativo de selectividad, ya que es una molécula que no interactúa fisiológicamente con el sitio activo de serinoproteasas, permitiendo de esta formar validar la especificidad del algoritmo de búsqueda y descartar acoplamientos que podrían ser inespecíficos. Las estructuras tridimensionales de estos ligandos se descargaron de la base de datos ZINC Database (https://zinc.docking.org/). El acoplamiento molecular con el sitio activo del modelo del dominio serinoproteasa de TMPRSS2 previamente minimizado se realizó utilizando el programa AutoDock v4.2.6 (Morris et al., 1998).

Antes de la simulación de acoplamiento, los inhibidores y el control negativo fueron preparados mediante la adición de átomos de hidrógeno, la fusión de hidrógenos no polares y la asignación de cargas parciales de Gasteiger. De manera homóloga, a la estructura de la serinoproteasa TMPRSS2 se le añadieron cargas de Kollman, se fusionaron los hidrógenos no polares y se mantuvieron únicamente los hidrógenos polares en AutoDockTools (ADT).

Las afinidades de unión descritas para compuestos como el nafamostat en el sitio activo de TMPRSS2 corresponden a los residuos de la tríada catalítica His, Asp y Ser (Ramakrishnan et al., 2021). En el presente estudio, estos residuos clave corresponden a His41 (His294), Asp90 (Asp343) y Ser186 (Ser439), según la numeración asignada al fragmento modelado del dominio y a la secuencia completa de TMPRSS2 de *Mus musculus*, respectivamente. Con base en esto, se ejecutó un

acoplamiento molecular dirigido específicamente a dicha cavidad catalítica. El espacio de búsqueda (Grid Box) se centró de forma exacta en las coordenadas cartesianas: x = 37.772, y = 38.107 y z = 37.898. Las dimensiones de la caja de cuadrícula se establecieron en 72, 64 y 66 puntos para los ejes X, Y y Z, respectivamente, empleando un espaciado (spacing) estándar de 0.375 Å, de esta forma se corrobora la cobertura óptima del bolsillo catalítico y el sitio de especificidad S1.

La exploración conformacional se llevó a cabo utilizando el Algoritmo Genético Lamarckiano (LGA). Para cada uno de los ligandos evaluados y el control negativo se generó un total de 10 corridas o conformaciones independientes. La selección de las mejores poses se realizó considerando la interacción específica con los residuos del sitio catalítico, la formación y persistencia de puentes de hidrógeno, así como la menor energía libre de unión (kcal/mol), priorizando la coherencia estructural y la estabilidad del complejo proteína–ligando. Las interacciones ligando-receptor y las representaciones visuales tridimensionales se analizarán mediante AutoDock, UCSF Chimera (Sonawane et al., 2021), PyMOL y el software PLIP (Schake et al., 2025).

IV.4.11. Simulación de dinámica molecular del complejo proteína - ligando

Para entender el efecto de la unión en la estabilidad estructural y la flexibilidad conformacional de los modelos predichos del dominio serinoproteasa de TMPRSS2 de *Mus musculus* seleccionados con mesilato de camostat, nafamostat, lopinavir y valrubicin se realizó la simulación de dinámica molecular del complejo. La topología del dominio serinoproteasa de TMPRSS2 y las moléculas de ligando se generaron por separado de los complejos respectivos.

La topología del dominio serinoproteasa de TMPRSS2 se generó utilizando GROMACS versión 2023.1 empleando el campo de fuerza OPLS-AA. Los ligandos fueron parametrizados mediante el servidor LigParGen, obteniéndose parámetros

compatibles con el mismo campo de fuerza (Dodda et al., 2017). Los complejos proteína-ligando se colocaron en una caja cúbica con una distancia mínima de 1,0 nm respecto a los bordes bajo condiciones periódicas de contorno y se solvataron con moléculas de agua explícitas utilizando el modelo SPC/E. La neutralización del sistema se realizó mediante la adición de iones Na+ y Cl− cuando fue necesario. La minimización de energía se llevó a cabo mediante el algoritmo de descenso más pronunciado con un máximo de 50 000 pasos. El sistema se equilibró en dos etapas consecutivas, en la primera etapa, se realizó una simulación NVT (número de partículas, volumen y temperatura constantes) a 310 K utilizando el termostato V-rescale. Posteriormente, se efectuó una simulación NPT (número de partículas, presión y temperatura constantes) a 310 K y 1 bar utilizando el barostato de Berendsen. Finalmente, se realizó una simulación de producción de 50 ns a 310 K y 1 bar. Cabe mencionar que los estados apo y holo se simularon y compararon bajo las mismas condiciones termodinámicas, físicas y de campo de fuerza.

El análisis estructural de los complejos dominio serinoproteasa-ligando se realizó a partir de las trayectorias de simulación mediante parámetros de la desviación cuadrática media (RMSD) y la fluctuación cuadrática media (RMSF). Además, se analizaron el radio de giro, el número de contactos, la distancia entre los centros de masa de TMPRSS2 y los ligandos, la variación de la distancia entre los residuos catalíticos en TMPRSS2 apo y en los complejos holo, así como la distancia entre los ligandos y los residuos catalíticos del dominio serinoproteasa de TMPRSS2.

IV.5. Análisis estadístico

El análisis de los datos se realizó mediante estadística descriptiva para resumir las variables cuantitativas obtenidas *in silico*. Los parámetros evaluados fueron la energía de afinidad (kcal/mol), RMSD, RMSF, puentes de hidrógeno, radio de giro (R_g),

distancias intermoleculares y contactos. Con ello, se calcularon promedios, desviaciones estándar y valores extremos de las trayectorias.

La recopilación de los datos se extrajeron comandos de GROMACS 2023.1 (*gmx rms*, *gmx rmsf*, *gmx hbond*, *gmx gyrate* y *gmx mindist*), mientras que el cálculo estadístico se ejecutó directamente con el comando *gmx analyze*.

Por último, se compararon los complejos proteína-ligando para explorar de manera preliminar su estabilidad estructural e interacciones. Todas las gráficas temporales se generaron y exportaron utilizando el comando xmgrace.

V. RESULTADOS

V.1. Presentación y análisis de los resultados

V.1.1. Análisis de la estructura genética

El gen *Tmprss2* de ratón (GenBank ID: NC_000082.7) está situado en el brazo largo del cromosoma 16 entre 97365882 y 97460741 pares de bases. El gen *Tmprss2* de 3175 bp consta de trece exones (min 12 bp/max 120 bp) interrumpidos por doce intrones de diferente tamaño (min 391 bp/max 12164 bp), como se muestra en la Figura 6. La región codificante (CDS) de 1473 nt no comprende el primer exón, dado que inicia en la posición 231 y termina en 1703.

A partir de la región codificante de *Tmprss2* se obtuvo la secuencia aminoacídica completa de TMPRSS2 de *Mus musculus* (490 aminoácidos). Sin embargo, el presente estudio se realizó exclusivamente sobre el dominio serinoproteasa (SP) correspondiente a los residuos 254–487 de la proteína completa.

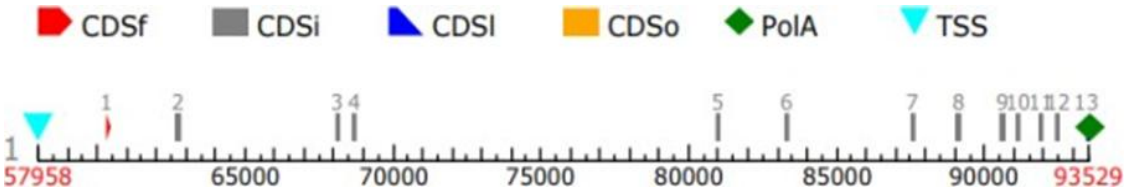


Figura 6. Estructura del gen Tmprss2 de *Mus musculus.* Se identifican trece exones señalados por números cardinales. La región codificante (CDS) no comprende el primer exón (flecha roja), dado que se encuentra en la posición 231-1703 (1473 bp). El sitio de inicio de la transcripción (TSS) y el sitio de poliadenilación (polA) son representados por un triángulo celeste y un rombo verde, respectivamente.

V.1.2. Análisis de parámetros bioquímicos

La composición de aminoácidos (Tabla 1) y las propiedades fisicoquímicas (Tabla 2) del dominio serinoproteasa de TMPRSS2 de *Mus musculus* se evaluó mediante ExPASy. La proteína TMPRSS2 consta de 234 aminoácidos y pesa 25,92 kDa. La proteasa tiene un pI < 7, lo que indica que tiene carácter ácido. Este dato es relevante para la optimización de sistemas de tampones en procesos de purificación, como la focalización isoeléctrica, donde el pI es un parámetro clave para la separación de proteínas. Aunque ProtParam de Expasy calcula el coeficiente de extinción para un rango de longitudes de onda (276, 278, 279, 280 y 282 nm), 280 nm es preferido porque las proteínas absorben fuertemente allí mientras que otras sustancias comúnmente presentes en soluciones proteicas no lo hacen. El coeficiente de extinción de la proteasa a 280 nm es 67420 M^{-1} cm^{-1} con respecto a la concentración de Cys, Trp y Tyr.

La proteasa presenta una vida media de 20 horas en reticulocitos de mamíferos *in vitro*, 30 minutos en levaduras *in vivo* y más de 10 horas en *Escherichia coli in vivo*. Según el índice de inestabilidad, el programa ProtParam de Expasy clasifica el dominio serinoproteasa como estable, ya que su índice de inestabilidad es inferior a 40. El índice alifático (AI) es 85,95. El índice de hidrofobicidad promedio (GRAVY) del dominio catalítico es -0,032. El índice GRAVY muy bajo de la proteasa sugiere que estas enzimas podrían tener una mejor interacción con el agua.

Tabla 1. Composición aminoacídica del dominio serinoproteasa de TMPRSS2 de *Mus musculus*

N°	AA	Cantidad	Porcentaje
1	Ala (A)	15	6.3%
2	Arg (R)	5	2.1%
3	Asn (N)	13	5.5%
4	Asp (D)	11	4.6%
5	Cys (C)	8	3.4%
6	Gln (Q)	10	4.2%
7	Glu (E)	7	3.0%
8	Gly (G)	23	9.7%
9	His (H)	5	2.1%
10	Ile (I)	15	6.3%
11	Leu (L)	20	8.4%
12	Lys (K)	11	4.6%
13	Met (M)	7	3.0%
14	Phe (F)	5	2.1%
15	Pro (P)	15	6.3%

16	Ser (S)	19	8.0%
17	Thr (T)	12	5.1%
18	Trp (W)	10	4.2%
19	Tyr (Y)	8	3.4%
20	Val (V)	18	7.6%

Tabla 2. Características fisicoquímicas del dominio serinoproteasa de TMPRSS2 de *Mus musculus*.

Características fisicoquímicas	**Valores**
Número de aminoácidos	234
Masa molecular (Da)	25923.73
pI teórico	6.21
Número total de residuos cargados negativamente (Asp + Glu)	18
Número total de residuos cargados positivamente (Arg + Lys)	16
Número total de átomos	3607
Coeficiente de extinción	67420 M^{-1} cm^{-1}
Índice de inestabilidad (II)	33.79

Índice alifático	85.95
Gran promedio de hidropaticidad (GRAVY)	-0.032

V.1.3. Análisis de dominios conservados.

Los resultados obtenidos mediante Prosite (Figura 7, Anexo 1) y NCBI Conserved Domains (Anexos 2 y 3) predijeron tres dominios altamente conservados en TMPRSS2 de *Mus musculus*. El dominio LDLRA (Dominio tipo A del receptor de lipoproteínas de baja densidad) corresponde a una región N-terminal rica en cisteínas que presenta siete repeticiones sucesivas esenciales para la estabilidad estructural. El dominio SRCR (Dominio rico en cisteínas del receptor scavenger) localizado contiguo a la región catalítica con alto contenido de cisteínas. Finalmente, el dominio Tryp_SPc, correspondiente a la serina proteasa tipo tripsina, se sintetiza como un zimógeno inactivo que requiere una proteólisis para convertirse en su forma catalíticamente activa.

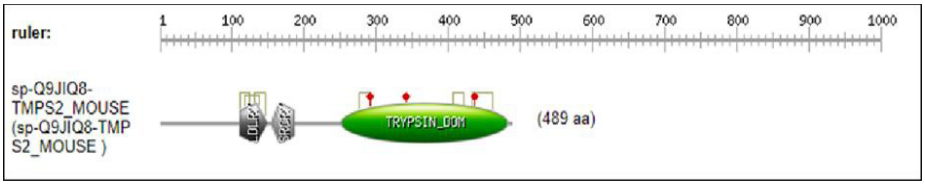


Figura 7. Gráfica de perfiles de la proteína TMPRSS2, según Prosite. Análisis de dominio conservados de la proteína TMPRSS2 que muestra la ubicación de los dominios LPL, SRCR y peptidasa S1. Los sitios activos están representados por figuras de color rojo y los puentes disulfuro por las líneas de color gris.

V.1.4. Interacción proteína- proteína.

El análisis de interacción proteína-proteína *in silico* mostró que la TMPRSS2 interactúa con varias proteínas; reguladores de la transcripción, enzimas y receptores, como se revela en la Figura 8. La proteína TMPRSS2 interactuó principalmente con Ar (proteína receptora de andrógenos), Prkdc (proteína kinasa), TMPRSS4 (proteína de transmembrana serina proteasa 4), Etv1 (proteína ETS de translocación variante 1),

Ace2 (enzima convertidora de angiostensia I) y Nkx3-1 (Proteína homeobox Nk-3.1), con puntajes de interacción significativos (>8).

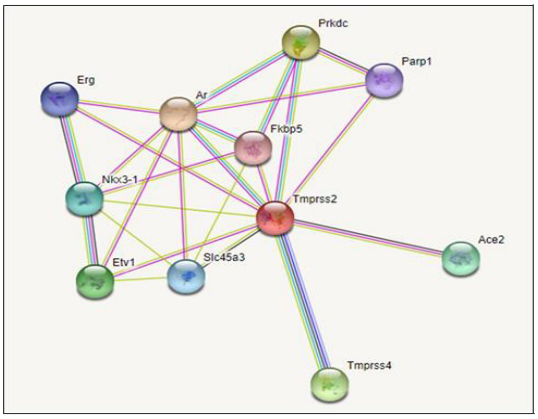


Figura 8. Análisis de la red de interacción proteína-proteína de la TMPRSS2 de *Mus musculus.* Interacciones conocidas: de bases de datos seleccionadas (línea celeste) y determinado experimentalmente (línea rosa). Proximidad génica (línea verde) y coocurrencia de genes (línea azul).

V.1.5. Alineamiento de secuencias

El alineamiento múltiple de la secuencia de aminoácidos del dominio serinoproteasa (SP) de TMPRSS2 de *Mus musculus* con las secuencias ortólogas de *Homo sapiens*, *Cavia porcellus*, *Mesocricetus auratus*, *Rattus norvegicus*, *Oryctolagus cuniculus*, *Macaca mulatta*, *Felis catus*, *Canis lupus familiaris* y *Mustela putorius furo*, realizado mediante Clustal Omega y visualizado con ESPript 3.0, mostró un elevado grado de conservación a nivel de secuencia (Figura 9).

En específico, el análisis comparativo directo entre *Mus musculus* y *Homo sapiens* cuantificó un 84.48% de identidad de secuencia y un 90.94% de similitud fisicoquímica a lo largo del dominio catalítico. El análisis de las regiones funcionales mostró la conservación completa de la tríada catalítica (His296, Asp345 y Ser441, señalados con triángulos rojos) entre ambas especies. Asimismo, los residuos que conforman los subsitios de especificidad S1, S1', S2 y S4 permanecieron conservados. En el subsitio S1 se observó la conservación del residuo Asp433, localizado en el fondo del bolsillo catalítico y flanqueado por Gly460 y Gly462. En el subsitio S1' se conservaron

los residuos comprendidos entre las posiciones 278–296, incluyendo los motivos VCG y HCV. De igual manera, el subsitio S2 mantuvo invariable el motivo 332-HPNYDS-337, mientras que el subsitio S4 conservó el bloque aminoacídico 337-SKTK-340 (Figura 9).

El alineamiento estructural evidenció además la conservación del Bucle A (LA), región donde se localizan los subsitios S2 y S4, el cual se encuentra delimitada por las láminas β6 y β7. En contraste, el subsitio S3, localizado en el contorno del Bucle 1 (L1), justo antes de la lámina β11, presentó la única diferencia observada en las proximidades del sitio activo. Mientras que en *Homo sapiens* se identificó el motivo 427-LQGN-430, en *Mus musculus* se observó la secuencia 427-LQGS-430, correspondiente a la sustitución de una asparagina por una serina en la posición 430 (Figura 9).

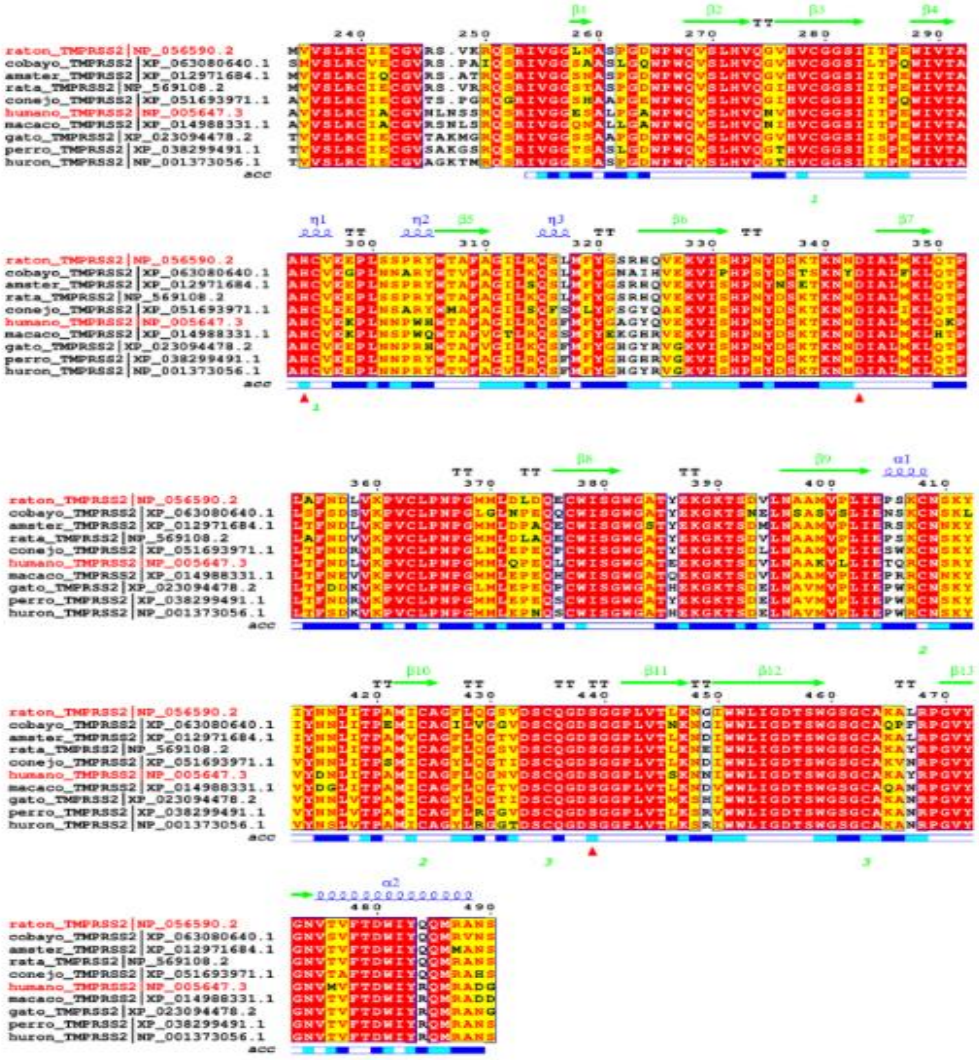


Figura 9. Alineamiento múltiple de secuencias del dominio serinoproteasa de TMPRSS2 de *Mus musculus* con secuencias ortólogas de otros animales. Se representan la tríada catalítica (flechas rojas), en la parte superior se predijo estructuras secundarias divididas en hélices eta (η1 a η3), hélices alfa (α1 y α2) y láminas betas (β11 a β13)

V.1.6. Predicción de estructuras secundarias:

El análisis de la secuencia del dominio serinoproteasa de TMPRSS2 en PHD predijo que su estructura secundaria está constituida por 22.04% de láminas beta y 10% de hélices alfa. Se observó una alta probabilidad de formación de estructuras o plegamientos azarosos o "random coil" (Anexo 4). Específicamente en el dominio serinoproteasa, PSIPRED predijo las estructuras secundarias en forma de láminas beta (12) y alpha hélices (4) separados por bucles (Figura 10).

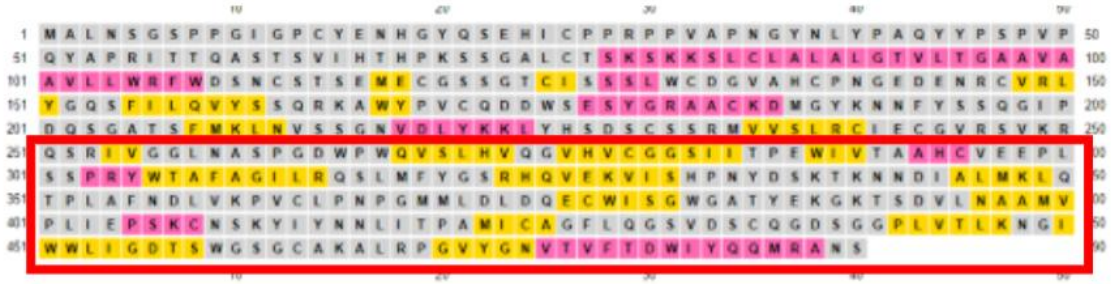


Figura 10. Predicción de estructuras secundarias del dominio serinoproteasa de TMPRSS2 de *Mus musculus.* Formado por hélices α coloreadas en rosa, láminas β coloreadas en amarillo y bucles coloreados en gris. La secuencia aminoacídica del dominio serinoproteasa de TMPRSS2 está sombreado en recuadro rojo.

V.1.7. Predicción de estructuras terciarias

V.1.7.1. Por Homología

Los modelos predichos por la metodología homología para el dominio serinoproteasa de TMPRSS2 de *Mus musculus* se generaron utilizando SWISS-MODEL a partir de las 2 plantillas identificadas, y fueron observados y modificados mediante el programa PyMOL. La detección inicial de plantillas se llevó a cabo mediante análisis BLAST contra la base de datos de Protein Data Bank (PDB) utilizando NCBI BlastP y SWISS-MODEL. El análisis SWISS-MODEL y BlastP revelaron que 7meq.1, y 7xyd.1.B resultaron plantillas óptimas para el modelado, con al menos un 40% de similitud (Tabla 3). Se observó que ambos modelos predijeron el dominio serinopeptidasa. Una diferencia notable entre ellos es que el Modelo 1 predijo 4 hélices alfa y 12 láminas beta,

mientras que el Modelo 2 predijo 5 hélices alfa y 12 láminas beta. En específico, al utilizar la plantilla 7XYD, el Modelo 2 incluyó la tercera hélice alfa (QSL) (Figura 11).

Los parámetros QMEANDisCo Global (estimación mínima de 0.90 ± 0.05) y GMQE (estimación mínima de 0.92) evidenciaron la alta calidad global de los modelos generados. Esto fue corroborado por la estimación de la calidad local de cada residuo en los modelos generados de la proteína TMPRSS2 de *Mus musculus* en comparación con las estructuras de referencia 7MEQ (A) y 7XYD (B). Los resultados indican que, en general, los modelos muestran una calidad aceptable en la mayoría de sus residuos especialmente a nivel del dominio serinoproteasa, con ciertas regiones con similitud local predicha baja (<0.8) (Anexo 5). La puntuación QMEAN Z-score (indicada por la estrella roja) indican que ambos modelos generados a partir de las estructuras de referencia, **(A)** 7MEQ y **(B)** 7XYD, tienen un Z-score cercano a 0 (regiones de puntos negros). Esto indica que los modelos presentan una buena calidad estructural en comparación con las estructuras experimentales de tamaño similar (Anexo 6).

Tabla 3. Resultados del alineamiento del dominio serinoproteasa de TMPRSS2 de *Mus musculus* con plantillas del PDB, utilizando SWISS-MODEL y BLASTp. La calidad de los modelos se evaluó mediante GMQE, QMEANDisCo Global, cobertura y porcentaje de identidad de secuencia.

Modelo	**Oligo-State**	**Método**	**GMQE**	**QMEANDisCo Global**	**Coverage**	**Seq Identity**
Modelo 01 (plantilla_7meq.1)	Monómero	X-ray, 1.95 Å	0.93	0.91 ± 0.05	100%	84.32%
Modelo 02 (plantilla_7xyd.1.B)	Heterodímero	X-ray, 2.6Å	0.92	0.90 ± 0.05	100%	84.32%

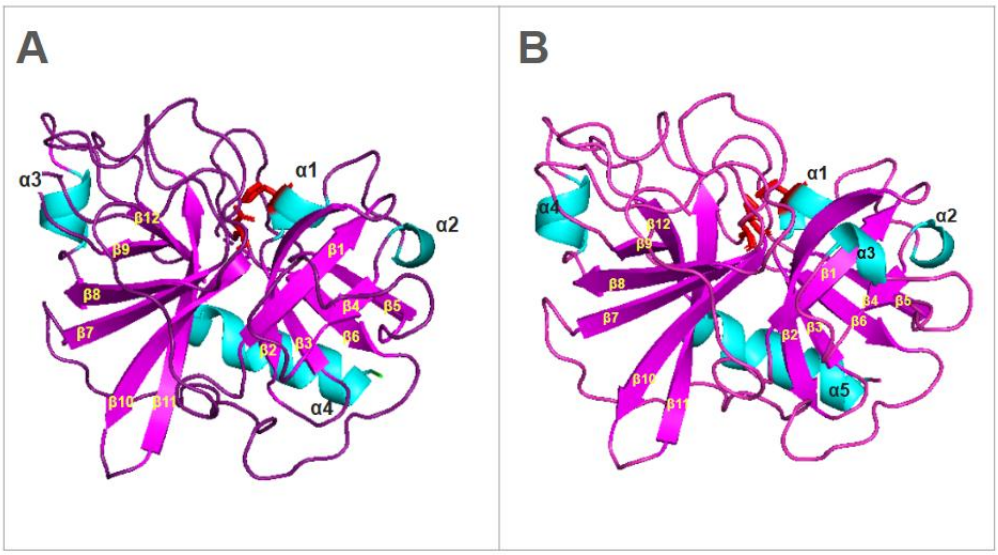


Figura 11. Estructuras tridimensionales predichas del dominio serinoproteasa de TMPRSS2 en *Mus musculus* mediante modelado por homología. Los modelos fueron visualizados en PyMOL, donde las hélices α (cian) y las láminas β (magenta) destacan la organización secundaria. El sitio catalítico se resalta en rojo. (A) Modelo 01 de predicción por homología. (B) Modelo 02 de predicción por homología.

V.1.7.2. Por *Threading*

En el modelado predicho por la metodología *threading*, se seleccionó el modelo con mejor resultado (c7meqA) debido a que presentó un porcentaje de confidencia del 100% y un porcentaje de identidad de 84% (Tabla 4). Se observó que el modelo predijo 4 hélices Alpha y 12 láminas beta del dominio SP (Figura 12).

Tabla 4. Resultados del modelamiento por *threading* del dominio serinoproteasa de TMPRSS2. La calidad de los modelos se evaluó mediante Coverage, Confidence y porcentaje de identidad.

Accesión	Coverage	Confidence %	%i.d
c7meqA	99%	100	84
c1z8gA_	99%	100	43
d1z8ga1	99%	100	43

c4durA	99%	100	43
c6o1gA_	99%	100	42

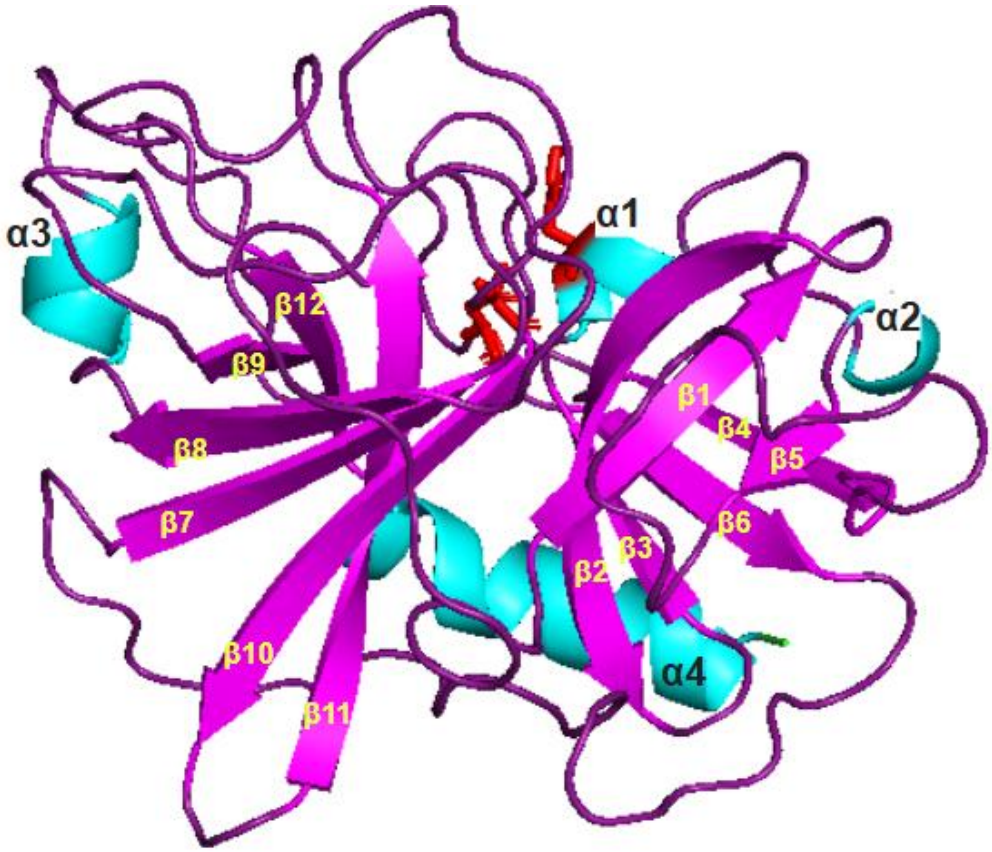


Figura 12. Estructura 3D del dominio serinoproteasa predicha por la metodología *Threading* (plantilla c7meq) visualizada en el programa PyMOL. Las hélices alfa (en cian) y las láminas beta (en magenta). El sitio catalítico se encuentra resaltado en rojo.

V.1.7.3. *Ab initio*

También se predijo el dominio serinoproteasa de TMPRSS2 utilizando el enfoque de modelo *ab initio* del servidor Robetta. El servidor generó cinco estructuras modelo (Anexo 7); sin embargo, se seleccionó el modelo 1 (Figura 13) que presentaba un error estimado menor a 1Å en la mayoría de los residuos, indicando una alta confianza en la predicción para los residuos del dominio SP. El modelo predicho presentó 6 hélices alfa

y 12 láminas beta. En particular, la sexta hélice alfa (CQGDS) no fue identificada por PHD, que únicamente predijo 5 hélices alfa.

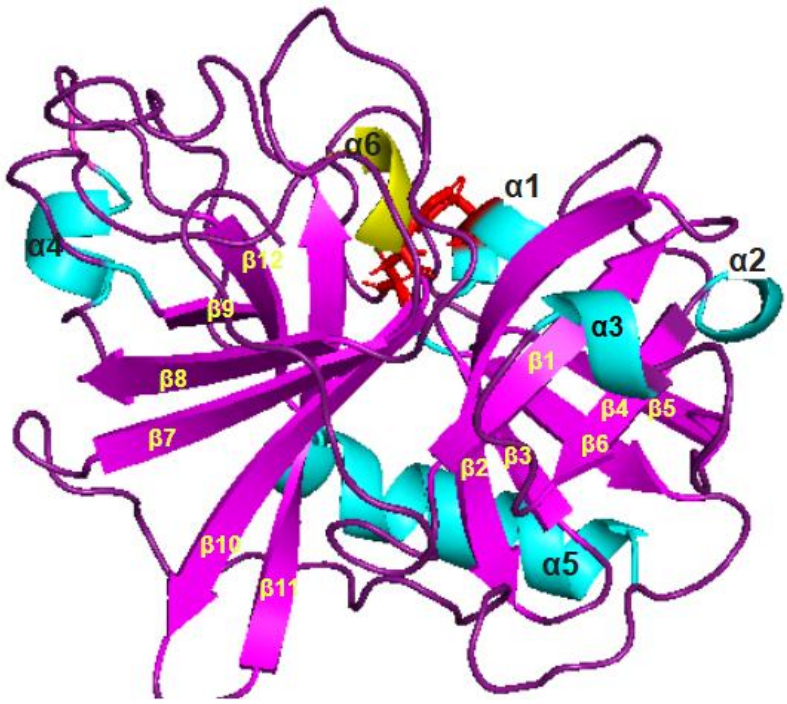


Figura 13. Estructura 3D del dominio serinoproteasa predicha por Robetta visualizada en el programa PyMOL. Las hélices alfa (en cian) y las láminas beta (en magenta). El sitio catalítico se encuentra resaltado en rojo.

V.1.7.4. Alphafold

También, el dominio serinoproteasa fue predicho por el algoritmo Alphafold y se obtuvieron 5 plantillas, las cuales fueron alineadas. Se observó que generó 5 hélices alfa y 12 láminas beta. A diferencia del modelo 2 generado por Swiss Model, la tercera hélice predicha estaba conformada por un mayor número de aminoácidos (QSLMFY) (Figura 14). Los resultados señalan una alta cobertura de secuencia a lo largo de la mayoría de las posiciones de la secuencia del dominio serinoproteasa deTMPRSS2, con una alta identidad de secuencia, lo que indica una fuerte conservación y una alta confianza en las predicciones generadas (Anexo 8). Además, los valores del IDDT son bastantes altos (>70) en la mayoría de las posiciones de los residuos del dominio SP de TMPRSS2, sugiriendo una buena confianza en la predicción (Anexo 9).

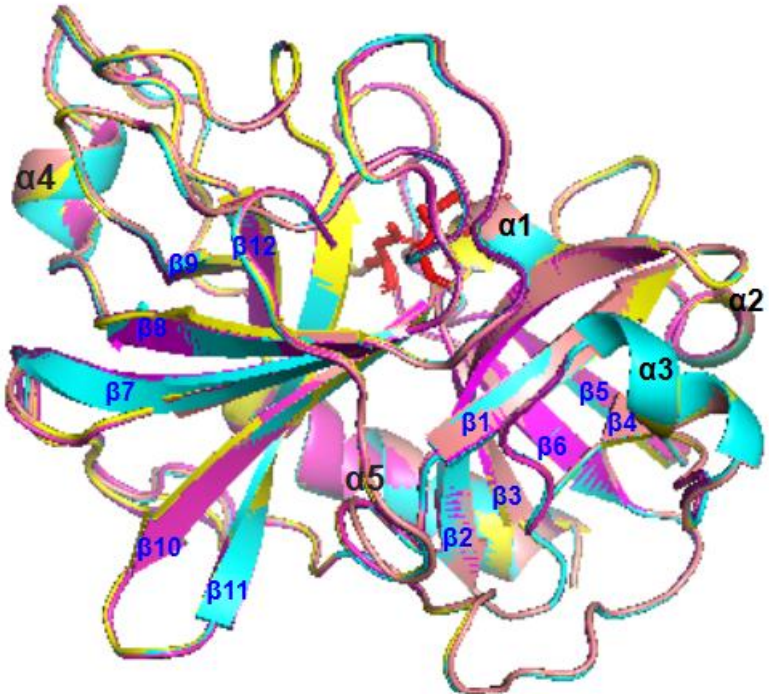


Figura 14. Alineamiento de los 5 modelos predichos por el algoritmo Alphafold visualizada en PyMOL. Las hélices alfa (en cian) y las láminas beta (en magenta). El sitio catalítico se encuentra resaltado en rojo.

V.1.8. Validación de modelos

V.1.8.1. PDBSum

La calidad de los 9 modelos obtenidos por las diferentes estrategias de modelamiento (homología, *threading* y *ab initio*), así como del algoritmo Alphafold, se realizó utilizando el servidor web PDBsum Generate. Se evaluó la sumatoria (%) del número de residuos de la estructura ubicados en las "regiones favorecidas" y "adicionalmente permitidas" del Plot de Ramachandran (Anexo 10). Los resultados mostraron que los 9 modelos elegidos obtuvieron una sumatoria superior al 95% de las "regiones favorecidas" y "adicionalmente permitidas" y muy pocos residuos caen en las regiones desfavorables, y presentan un valor del parámetro G-factor mayor a -0.5. Se eligió el modelo 03 generado mediante Alphafold como potencial modelo para la modelización del dominio serinoproteasa de TMPRSS2 en ratones, ya que presentaron

los porcentajes de sumatoria más altos de todos los modelos, con un 99.5% (G-factor: 0.57).

V.1.8.2. ModFOLD

La calidad de los 9 modelos obtenidos por homología, *threading*, *ab initio* y Alphafold se evaluó utilizando el servidor web ModFOLD. Todos los modelos obtuvieron un puntaje global de calidad superior a 0.4, lo que indicó la complejidad de estos. Estos resultados indicaron que los modelos fueron confiables y muy similares a la estructura nativa. Además, los modelos obtuvieron buenos resultados en la estimación de error de residuos, con valores inferiores a < 2 Å en el dominio serinoproteasa, donde se encuentran el sitio activo de la proteína TMPRSS2 (Anexo 11). Entre los modelos evaluados, el análisis sugirió que el mejor modelo fue el obtenido mediante Alphafold (Modelo 3). Este modelo presentó una confianza media, un p-valor de 2.958E-5 y un puntaje global de calidad de 0.9455, destacando como el modelo más preciso y confiable para la predicción de la estructura de TMPRSS2, por esta razón, se seleccionó para análisis posteriores en comparación con los demás modelos generados mediante AlphaFold.

Posteriormente, se alineó el modelo 3 generado por Alphafold junto a los modelos generados por Swiss Model, Phyre2 y Robetta (Figura 15A). Aunque todos los modelos predichos generaron las 12 láminas beta del dominio SP; las hélices alfa variaron entre 4 a 6, además se observó una ligera variación en los bucles fuera del sitio catalítico. La estructura 3D del modelo de homología se muestra en la Figura 15B, se señala la matriz extracelular (área superficial de Van der Waals) y los dominios SRCR, LDL-Receptor clase A y el dominio transmembrana, luego se señala los residuos de la triada catalítica en la Figura 15C.

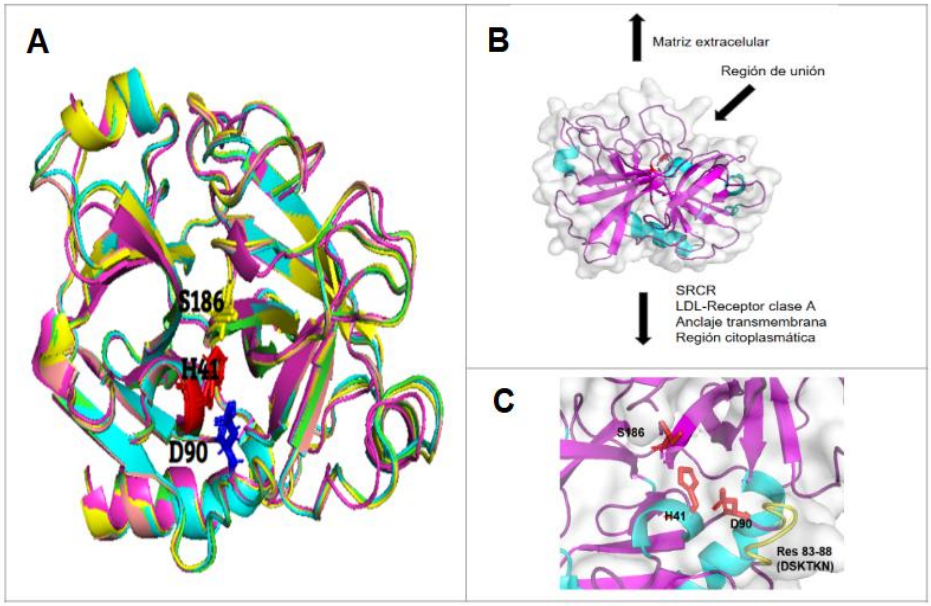
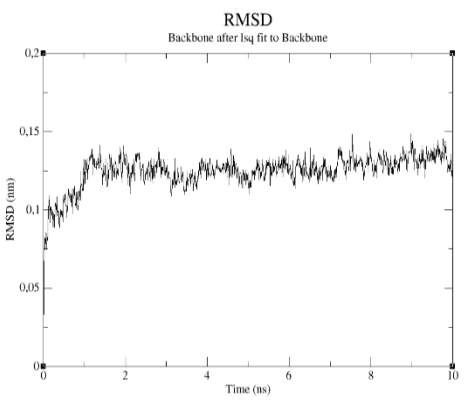
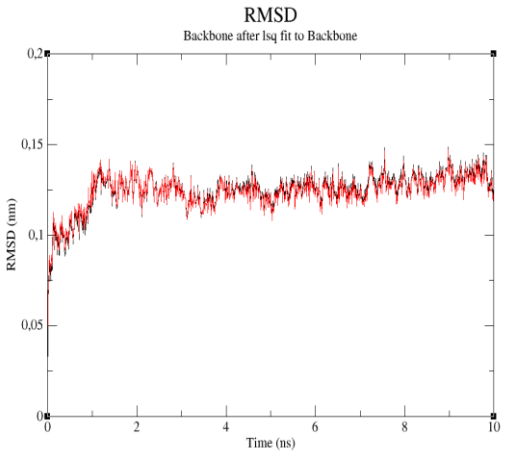


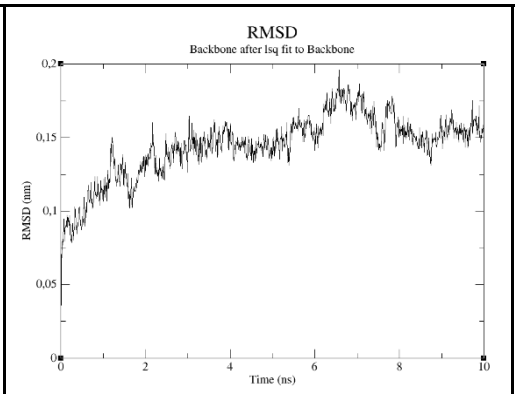
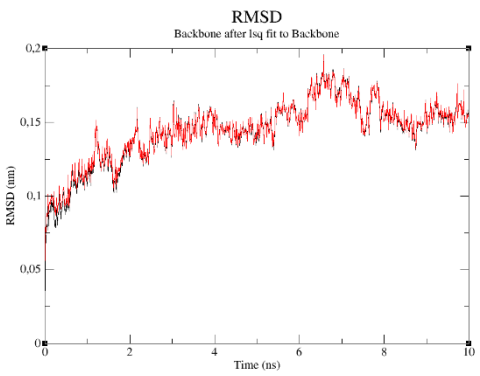
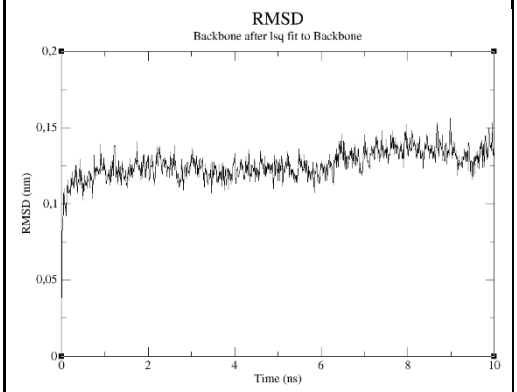
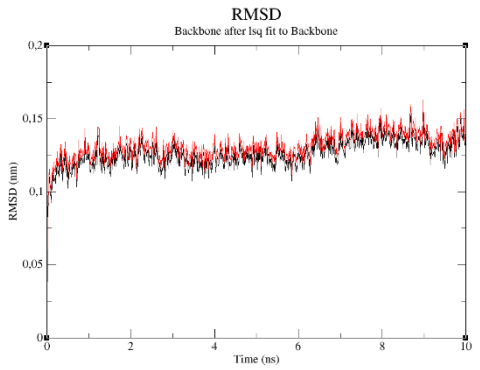
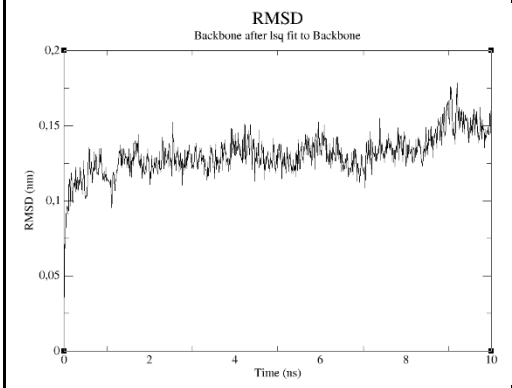
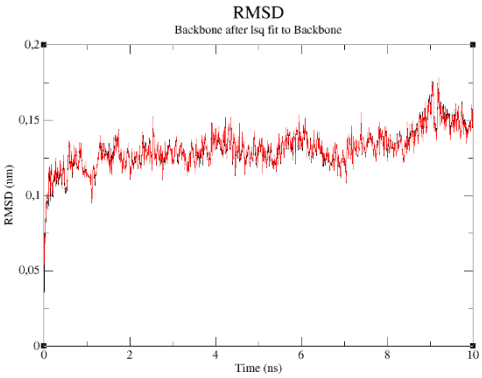
Figura 15. Figura 15. Comparación estructural y organización del dominio serinoproteasa (SP) de TMPRSS2 de *Mus musculus*. (A) Superposición estructural de los modelos tridimensionales generados, destacando la conservación del sitio activo conformado por la tríada catalítica (His41, Asp90 y Ser186). Se detallan los modelos según su origen metodológico: SWISS-MODEL (celeste), Phyre2 (verde), Robetta (rosado) y AlphaFold (amarillo). (B) Orientación espacial de la estructura terciaria del dominio SP de TMPRSS2. La flecha superior indica la proyección hacia la matriz extracelular (región de unión), mientras que la flecha inferior señala la orientación hacia los dominios adyacentes (SRCR, LDL-Receptor clase A, anclaje transmembrana y región citoplasmática). La superficie semitransparente (gris) representa el área de Van der Waals; los elementos de estructura secundaria se colorean en malva (bucles), cian (hélices alfa) y magenta (láminas beta). (C) Vista detallada de la tríada catalítica (His41, Asp90 y Ser186) y de los residuos adyacentes 83-88 (DSKTKN) en el modelo tridimensional de ratón.

V.1.9. Simulación de dinámica molecular (MDS)

Se evaluó la estabilidad de los modelos generados para el dominio serinoproteasa de TMPRSS2 de ratón mediante una simulación de dinámica molecular empleando el campo de fuerza OPLS-AA para todos los átomos en GROMACS versión 2023.1. Previo al análisis a la fase de producción, los modelos predichos convergieron hacia su minimización energética, posteriormente la temperatura del sistema se estabilizó alrededor de los 310 K y la presión alcanzó 1 bar, de esta forma las condiciones permanecieron homogéneas antes de la fase de producción. La simulación

inicial de 10 ns se realizó con el objetivo de que los modelos predichos del dominio serinoproteasa de TMPRSS2 se encuentren dinámicamente equilibrados, minimizados y libre de artefactos estructurales al llegar a la convergencia estructural. Los resultados muestran que el RMSD de los modelos generados por el método de homología con plantilla 7meq y el modelo generado por AlphaFold llegaron a una rápida convergencia estructural entre 0.1 y 0.15 nm a los 2 ns de los 10 ns de simulación (Figura 16) respecto al resto de los modelos predictores. No se observaron variaciones significativas entre el RMSD de la estructura presente en el sistema equilibrado minimizado y la estructura cristalizada. Asimismo, el RMSF de ambos modelos mostraron fluctuaciones más consistentes y valores más bajos en general (en el rango 0.05 a 0.2 y 0.05 a 0.25 nm) (Figura 17). En consecuencia, los modelos 3D obtenidos por homología y AlphaFold del dominio serinoproteasa de TMPRSS2 de ratón fueron elegidos para hacer el análisis de docking molecular.

Modelo	**RMSD (backbone respecto al estado inicial)**	**RMSD (referencia cristalizada)**
1		

2	RMSD Backbone after lsq fit to Backbone RMSD (nm) Time (ns)	RMSD Backbone after lsq fit to Backbone RMSD (nm) Time (ns)
3	RMSD Backbone after lsq fit to Backbone RMSD (nm) Time (ns)	RMSD Backbone after lsq fit to Backbone RMSD (nm) Time (ns)
4	RMSD Backbone after lsq fit to Backbone RMSD (nm) Time (ns)	RMSD Backbone after lsq fit to Backbone RMSD (nm) Time (ns)

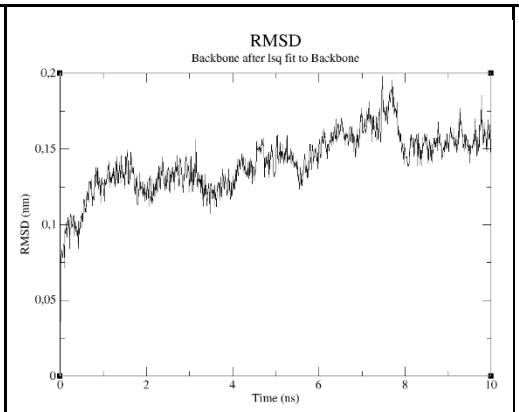
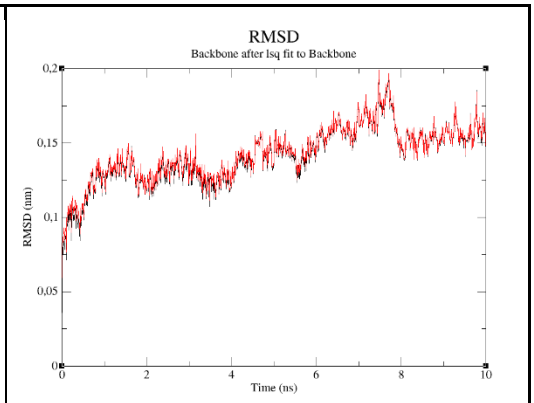
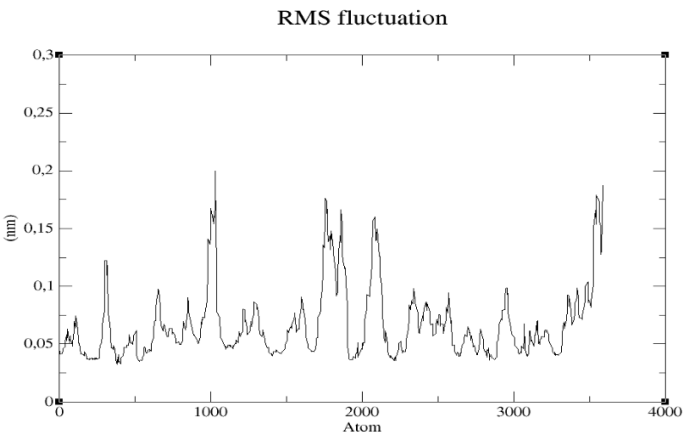
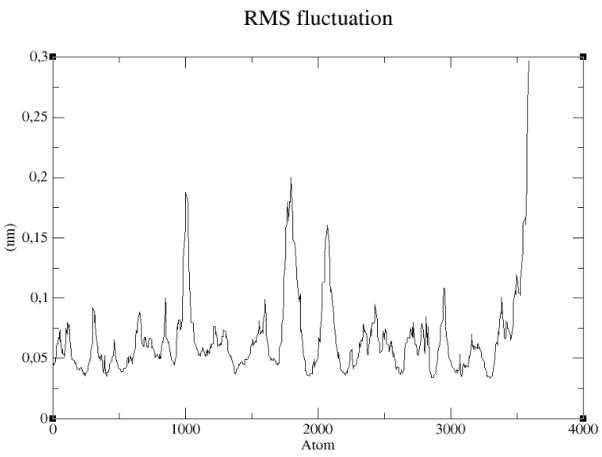
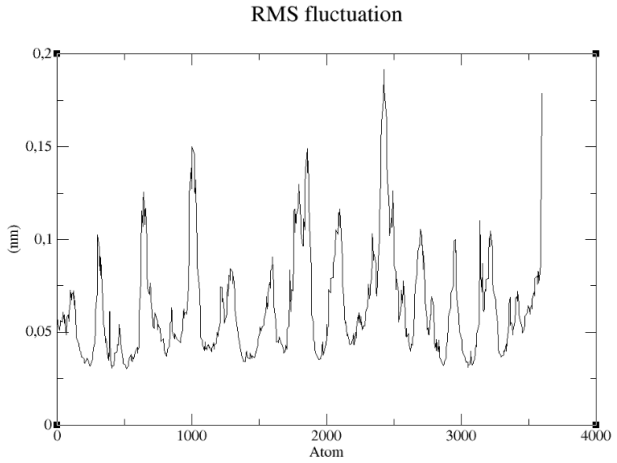
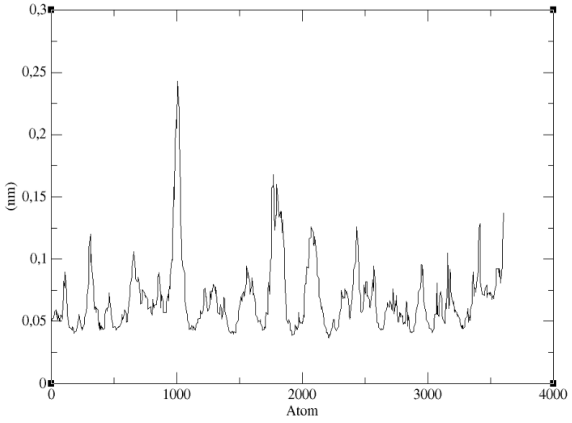
5		

Figura 16. Gráficos de RMSD (nm) en función del tiempo (10 ns) para diferentes modelos de dominio serinoproteasa de TMPRSS2 obtenidos mediante distintos enfoques de modelado. Se presentan dos análisis de RMSD: (izquierda) variación del RMSD del backbone con respecto a su estado inicial y (derecha) variación del RMSD con respecto a la estructura cristalizada de referencia. Modelo 1: plantilla 7meq (Swiss Model), Modelo 2: plantilla 7xyd (Swiss Model), Modelo 3: modelo generado por Robetta, Modelo 4: modelo generado por AlphaFold, Modelo 5: 7 meq generado por Phyre2.

Modelo	RMSF
1	

2	RMS fluctuation 0,3 0,25 0,2 0,15 0,1 0,05 0 (nm) 0 1000 2000 3000 4000 Atom
3	RMS fluctuation 0,2 0,15 0,1 0,05 0 (nm) 0 1000 2000 3000 4000 Atom
4	RMS fluctuation 0,3 0,25 0,2 0,15 0,1 0,05 0 (nm) 0 1000 2000 3000 4000 Atom

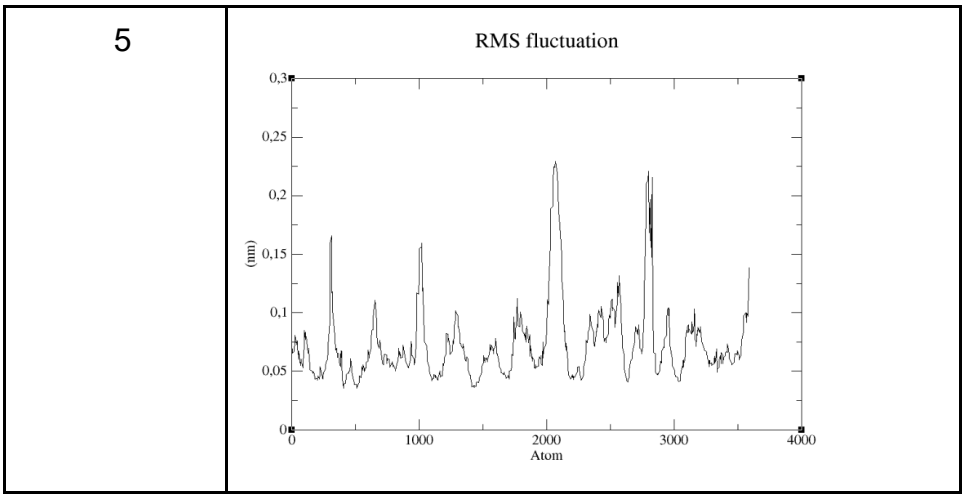


Figura 17. Gráficos de fluctuación del RMSF (nm) en función del número de átomos a lo largo de la simulación de dinámica molecular (MD) de 10 ns para distintos modelos del dominio serinoproteasa de TMPRSS2 obtenidos mediante diferentes enfoques de modelado. Modelo 1: plantilla 7meq (Swiss Model), Modelo 2: plantilla 7xyd (Swiss Model), Modelo 3: modelo generado por Robetta, Modelo 4: modelo generado por AlphaFold, Modelo 5: 7 meq generado por Phyre2.

V.1.10. Análisis de la calidad estructural y conservación geométrica del sitio activo del dominio serinoproteasa de TMPRSS2 de *Mus musculus* frente a la referencia cristalográfica (PDB: 7MEQ)

La superposición estructural tridimensional permitió confirmar la alta fidelidad de los modelos del dominio serinoproteasa de TMPRSS2 predichos por Swiss Model y Alphafold.

El alineamiento del modelo obtenido por homología mediante SWISS-MODEL frente al cristal 7MEQ mostró valores de RMSD de 0.285 Å. Por otro lado, el alineamiento de la estructura predicha por AlphaFold frente al mismo molde experimental dio como resultado un RMSD de 1.367 Å (Figura 18). Asimismo, para verificar la integridad estructural de la zona catalítica, se realizó la superposición molecular del sitio activo de los modelos tridimensionales obtenidos frente a la estructura cristalográfica experimental del receptor (PDB: 7MEQ). Las distancias interresiduales medidas entre los entre los carbonos alfa ($C\alpha$) de la tríada catalítica (His41, Asp90 y Ser186) oscilan entre 6.4 Å y 10.0 Å (His41-Asp90: 6.4 Å; His41-Ser186: 8.0 Å; Asp90-Ser186: 10.0 Å).

Se observó una alta conservación geométrica y espacial al comparar el modelo por homología de SWISS-MODEL (rojo) y la predicción de AlphaFold (azul) con la estructura cristalográfica de referencia 7MEQ (verde), evidenciando desviaciones estructurales mínimas (Figura 19). Finalmente, se caracterizó la cavidad de unión de la serinoproteasa de *Mus musculus*, donde la triada catalítica His41, Asp90 y Ser186 (color magenta) se encuentra adyacente al subsitio de especificidad S1, la cavidad profunda del dominio serinoproteasa cuya base se encuentra Asp178 y flanqueada por los residuos Gly207 y Gly209 (naranja), los cuales se encuentran en las paredes de entrada del bolsillo. Además, se muestra el inhibidor nafamostat cerca de la triada como al bolsillo s1 del dominio serinoproteasa de TMPRSS2 (Figura 20).

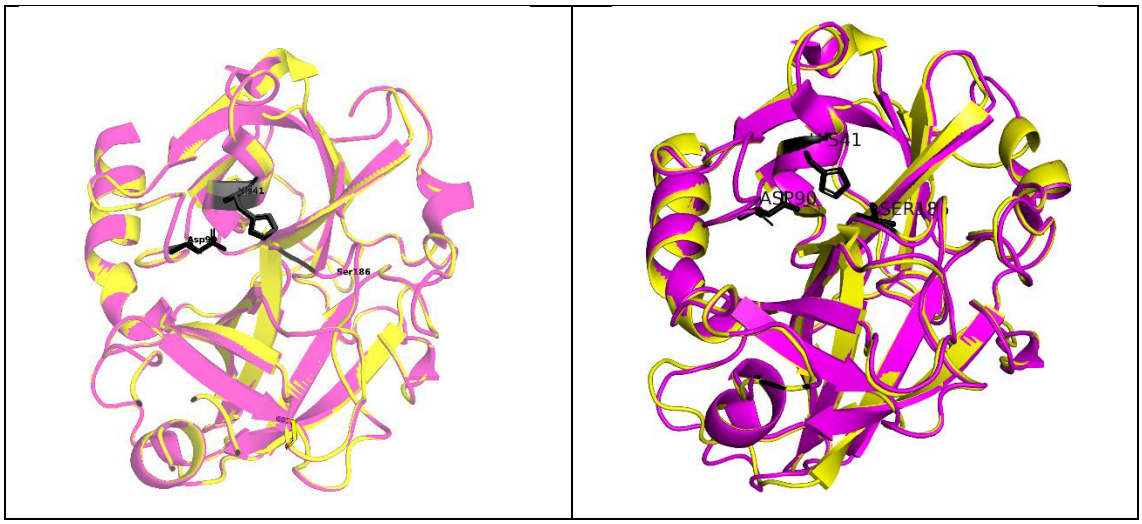


Figura 18. Superposición estructural tridimensional de los modelos del dominio serinoproteasa de TMPRSS2 de *Mus musculus* frente a la referencia cristalográfica. (Izquierda) Alineamiento estructural del modelo obtenido por homología mediante SWISS-MODEL frente a la estructura experimental del cristal de referencia (PDB: 7MEQ, RMSD = 0.285 Å). (Derecha) Alineamiento tridimensional del modelo predicho por AlphaFold frente al mismo molde experimental (PDB: 7MEQ, RMSD = 1.367 Å). En ambos paneles se destacan en representación de bastones (*sticks*) los residuos aminoacídicos altamente conservados de la tríada catalítica (His41, Asp90 y Ser186).

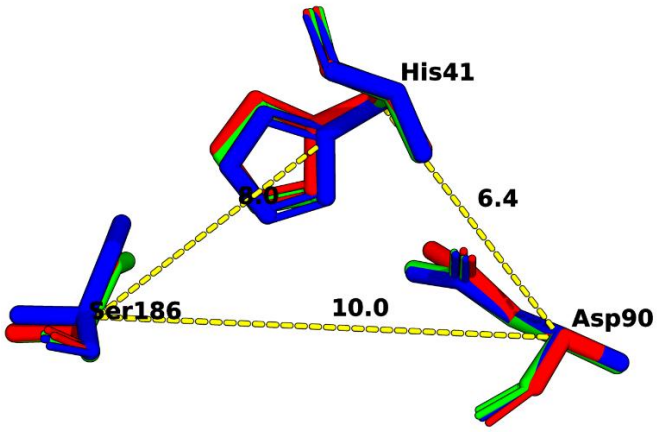


Figura 19. Conservación geométrica de la tríada catalítica y caracterización estructural del bolsillo de especificidad S1 de TMPRSS2. Superposición tridimensional del sitio activo conformado por los residuos His41, Asp90 y Ser186 del modelo por homología de SWISS-MODEL (rojo), la predicción de AlphaFold (azul) y la estructura cristalográfica de referencia (PDB: 7MEQ, verde). Las líneas discontinuas amarillas indican las distancias interresiduales medidas en Angstroms (Å) entre sus carbonos alfa (Cα)

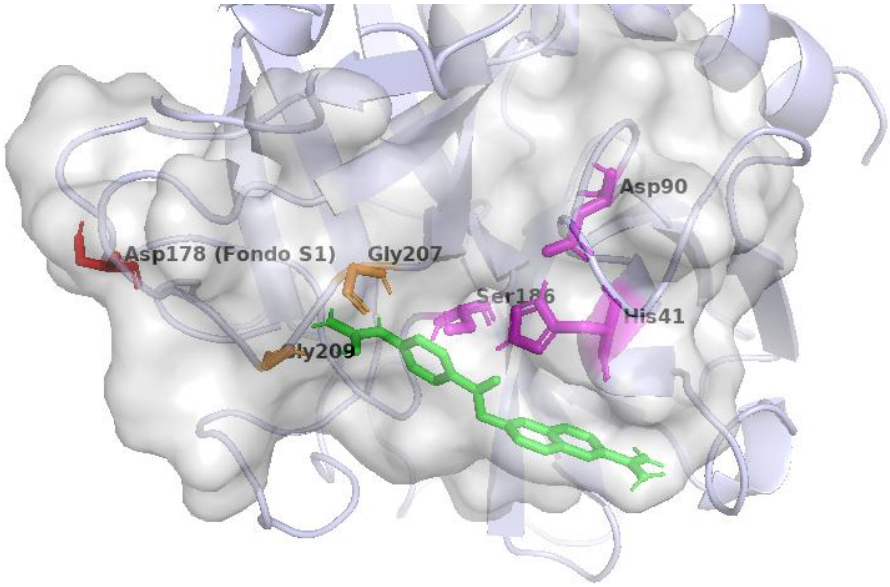


Figura 20. Caracterización estructural del bolsillo de especificidad S1 de TMPRSS2. Representación de la cavidad de unión del dominio serinoproteasa de *Mus musculus*. Se muestra la tríada catalítica (magenta) adyacente al bolsillo de especificidad S1, destacando a Asp178 (rojo) en la base de la cavidad, los residuos Gly207 y Gly209 (naranja) flanqueando las paredes de entrada, y el acoplamiento espacial de un inhibidor de referencia (verde) en el sitio activo.

V.1.11. Acoplamiento molecular del dominio serinoproteasa de TMPRSS2 y los 4 inhibidores por AutoDock

Se utilizaron el modelo de homología y AlphaFold simulados de dinámica molecular (MD) del dominio serinoproteasa de TMPRSS2 para estudiar su interacción con 4 inhibidores experimentalmente probados, mesilato de camostat, nafamostat, lopinavir y valrubicina. Asimismo, se incorporó el Adenosín Trifosfato (ATP) como control negativo del ensayo para evaluar la selectividad posicional y energética del algoritmo de búsqueda.

En el análisis del acoplamiento molecular realizado por AutoDock se obtuvieron 10 poses para cada complejo, de los cuáles se eligió a la mejor pose en base a los valores de energía de unión, interacción con puentes de hidrógeno y residuos catalíticos, por lo que al final se obtuvieron 4 poses potenciales entre el dominio serinoproteasa TMPRSS2 y los ligandos (Figura 21-24). Los valores de energía de enlace, la constante de inhibición (Ki), número de enlaces de hidrógenos y los residuos de TMPRSS2 involucrados en dichas interacciones se presentan en el Anexo 12-20, el cual se encuentran ordenados en base al número de run del cual provinieron.

El análisis estructural del acoplamiento molecular con los inhibidores de bajo masa molecular, nafamostat y camostat, mostró interacciones esenciales en el bolsillo de especificidad primaria S1, el bolsillo secundario S2 y el entorno de entrada de TMPRSS2. Mientras que los inhibidores exploratorios de alto masa molecular, lopinavir y valrubicina, se ubicaron en las regiones periféricas externas. Sin embargo, el control negativo (ATP) exhibió una ausencia total de reconocimiento específico.

En la tríada catalítica, nafamostat mostró una interacción de tipo puente salino y contactos atómicos estrechos con la His41 a distancias de entre 3.28 Å y 4.51 Å, además de una proximidad con la Ser186 catalítica a 3.28 Å. Y camostat mostró interacciones puente de hidrógeno con Ser186 (2.65 Å) y Asp90 (2.5Å y 2.7 Å), además de

interacciones hidrofóbicas estables y contactos cercanos con el anillo imidazólico de la His41 (3.19 Å y 3.23 Å) (Anexo 21 y Anexo 22).

En el bolsillo de especificidad primaria S1, nafamostat mostró interacciones tipo puente de hidrógeno donde destaca un acoplamiento bidentado con la Gly209 (2.79 Å y 3.01 Å) y la Gly207 (3.47 Å), complementado por interacciones hacia Glu45 (2.87 Å), Ser181 (3.10 Å) y Cys42 (3.53 Å). Mientras que camostat presentó una orientación que abarca múltiples subsitios; estabilizó su anclaje en el motivo conservado de S1 a través de puentes de hidrógeno con la Gly184 (2.07 Å y 2.20 Å), pero extendió su estructura hacia el bolsillo secundario S2, en el cual destacaron interacciones puentes de hidrógeno, π-Catión y contactos cercanos con Lys87, Tyr82 y Ile91. El acoplamiento de ambos inhibidores de referencia se complementa en la boca de la cavidad mediante residuos del entorno de entrada y periferia; nafamostat interactuó con los residuos Glu44, Gln183, Leu47 y Pro46, mientras que el camostat consolidó contactos con Val25, Gln183, Glu44 (Anexo 21 y Anexo 22).

A diferencia de los inhibidores de referencia, valrubicina y lopinavir no lograron internalizar en la cavidad estrecha S1, por lo que adoptaron una ubicación superficial en los lazos periféricos. Valrubicina se ubica en la entrada del sitio activo, bloqueando el acceso al subsitio S1. Entre las interacciones destaca los puentes de hidrógeno entre el ligando y los residuos como Thr138 a una distancia corta de 2.35 Å, Gln183 (3.28 Å) y Val25 (3.22 Å) pertenecientes al bucle de entrada N-terminal S1'. Se evidenció interacciones tipo puente salino entre el grupo carboxilato de la valrubicina y el anillo imidazólico de la His41 (5.01 Å y 5.18 Å), los cuales resultan demasiados lejanas en comparación a los ligandos de referencia, lo que indica que el fármaco no se introduce a la cavidad catalítica, sino se posiciona transversalmente sobre ella. De la misma manera, lopinavir tuvo una interacción tipo puente de hidrógeno con el residuo distintivo del lazo de reconocimiento Asp140, la cual estuvo reforzada por interacciones con Gln183 (2.90 Å) y Thr138 (3.22 Å). Además, la estabilidad exógena de lopinavir fue dada

por Val23 (3.24 Å) y Val25 (3.62 Å) mediante contactos hidrofóbicos (Anexo 21 y Anexo 22).

El análisis estructural y geométrico del control negativo confirmó una ausencia total de reconocimiento molecular y especificidad posicional. El ATP no consolidó ningún enlace de hidrógeno ni interacciones estables con ninguno de los residuos de la tríada catalítica (His41, Asp90 y Ser186) en ninguna de las 10 conformaciones generadas. Las poses resultantes mostraron una elevada dispersión geométrica, quedando el ligando confinado fuera de la cavidad catalítica.

Por otro lado, el análisis de la energía de unión de los 4 complejos evaluados, la interacción entre nafamostat y el dominio serinoproteasa de TMPRSS2 mostró la pose con la energía de enlace más baja (-5.52 kcal/mol), con 6 enlaces de hidrógeno (Tabla 5). Por otro lado, la interacción entre lopinavir y el dominio serinoproteasa de TMPRSS2 presentó una energía más alta (-2.62 kcal/mol), con 4 enlaces de hidrógeno (Tabla 5). En complejo acoplado del dominio serinoproteasa de TMPRSS2-camostat tuvo una energía de -3.19 kcal/mol y formó 15 enlaces de hidrógeno, mientras que valrubicin-TMPRSS2 tuvo una energía de -4.02 kcal/mol y 4 enlaces puente de hidrógeno (Tabla 5). Finalmente, el control negativo (ATP) evidenció un perfil energético marcadamente desfavorable e inespecífico en comparación con los inhibidores diana. En las corridas se obtuvieron valores de energía de unión netamente positivos (alcanzando hasta $+1.23$ kcal/mol).

Tabla 5. Resultados del acoplamiento molecular entre el dominio serinoproteasa de TMPRSS2 con sus inhibidores obtenidos en AutoDock.

Ligando	Modelo	Pose	Energía de Unión (Kcal/mol)	H-Bonds	Residuos Clave
nafamostat	SwissModel	pose 1	-5.52	6	His41
valrubicin	SwissModel	pose 1	-4.02	4	His41
camostat	Alphafold	pose 1	-3.19	15	His41, Ser186

lopinavir	SwissModel	pose 1	-2.62	4	His41

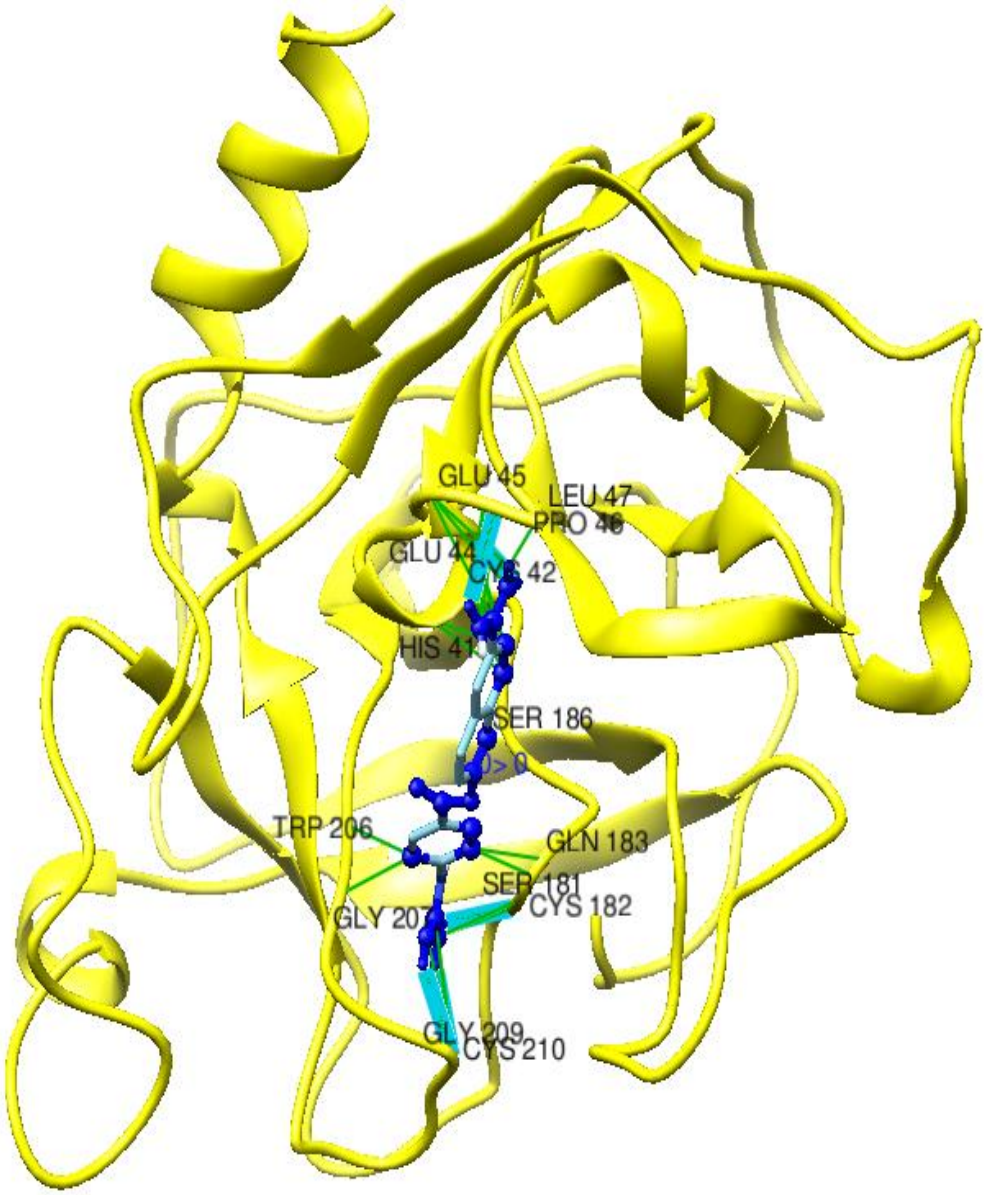


Figura 21. Interacción molecular entre el dominio serinoproteasa de TMPRSS2 y el ligando nafamostat obtenida mediante docking molecular en UCSF Chimera. La proteína se muestra en amarillo y el ligando en azul. Los residuos clave de interacción (Gly 209, Glu 45, Gln 183, His 41) están etiquetados en negro. Los enlaces de hidrógeno se indican en celeste, y otras interacciones relevantes en verde.

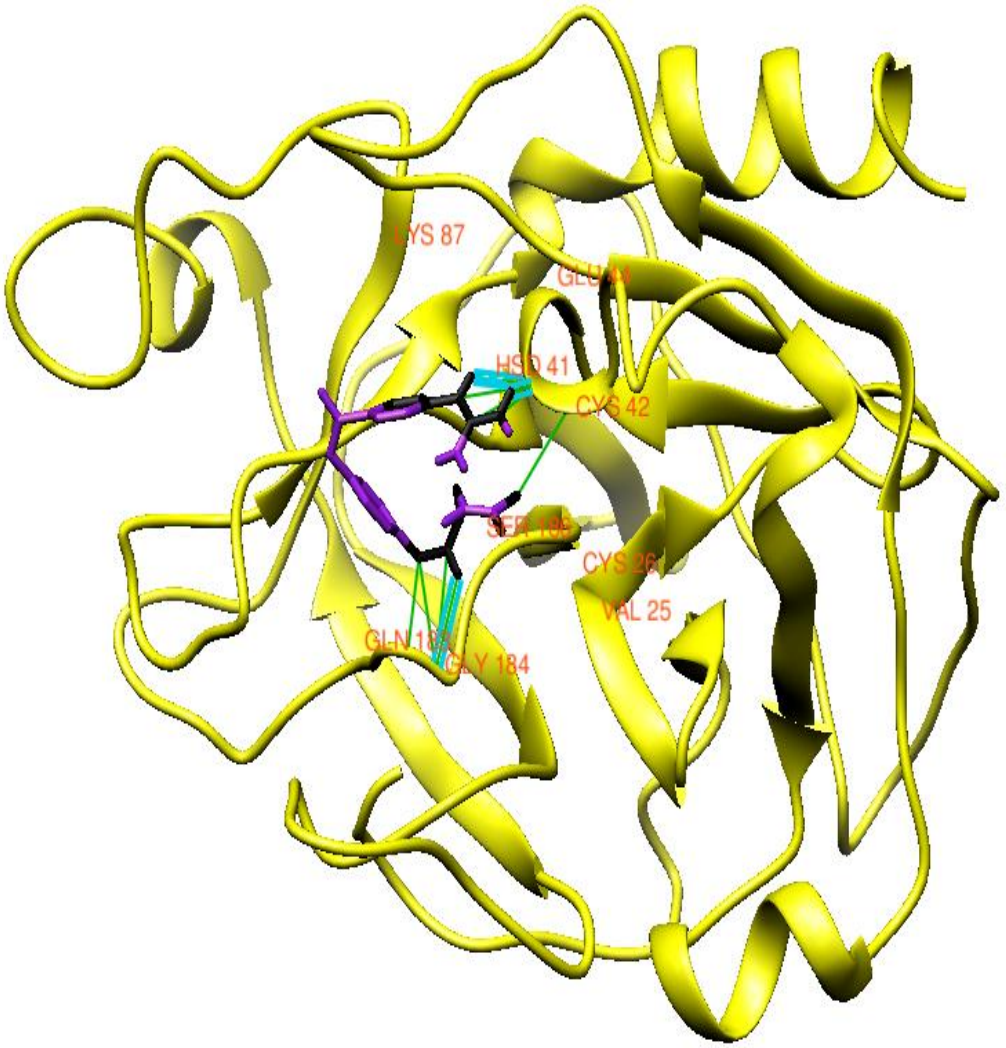


Figura 22. Interacción molecular entre el dominio serinoproteasa de TMPRSS2 y el ligando camostat obtenida mediante docking molecular en UCSF Chimera. La proteína se muestra en amarillo y el ligando en morado. Los residuos clave de interacción (Ser 186, His 41, Glu 44, Cys 26, Gln 183, Gly 184) están etiquetados en rojo. Los enlaces de hidrógeno se indican en celeste, y otras interacciones relevantes en verde.

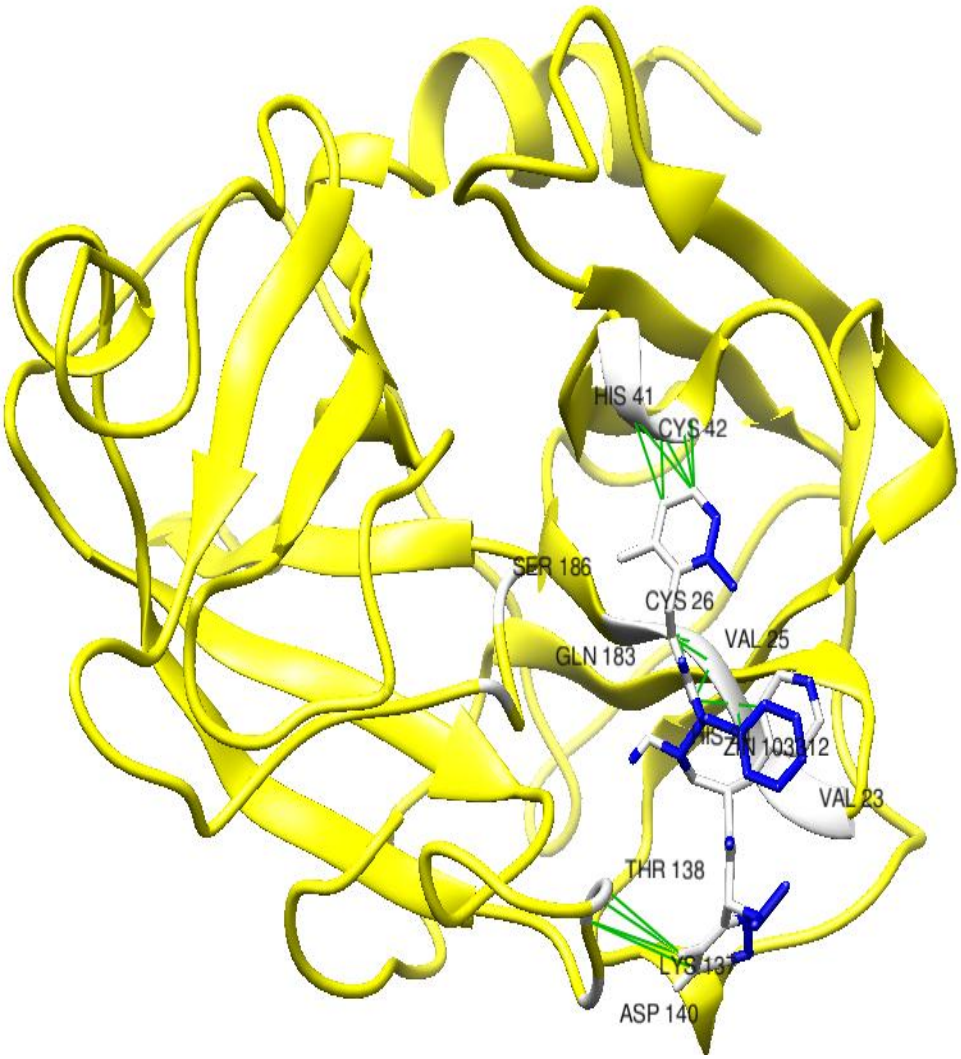


Figura 23. Interacción molecular entre el dominio serinoproteasa de TMPRSS2 y el ligando lopinavir obtenida mediante docking molecular en UCSF Chimera. La proteína se muestra en amarillo y el ligando en azul. Los residuos clave de interacción (His 41, Cys 42, Cys 26, Val 25, Thr 138, Lys 137) están etiquetados en negro.

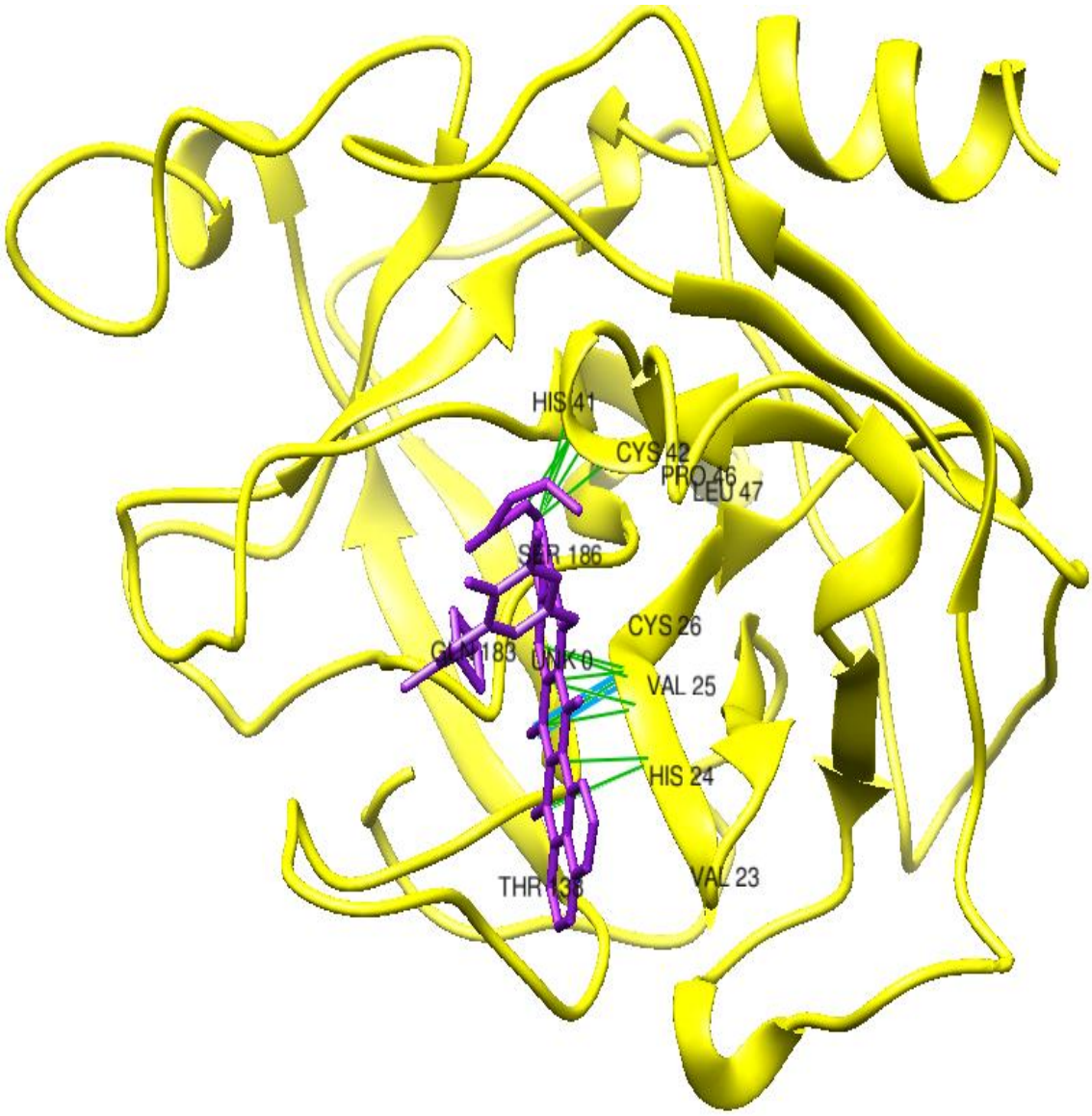


Figura 24. Interacción molecular entre el dominio serinoproteasa de TMPRSS2 y el ligando valrubicin obtenida mediante docking molecular en UCSF Chimera. La proteína se muestra en amarillo y el ligando en morado. Los residuos clave de interacción (His 41, Cys 42, Ser 186, Val 25, His 24, Cys 26) están etiquetados en negro. Los enlaces de hidrógeno se indican en celeste, y otras interacciones relevantes en verde.

V.1.12. Simulación de dinámica molecular (MD) del dominio serinoproteasa de TMPRSS2 y 4 inhibidores

Luego de la selección de las poses óptimas realizadas durante el acoplamiento molecular estático, se procedió con la simulación de dinámica molecular de 50 ns en solvente explícito para cada complejo proteína-ligando, bajo las mismas condiciones de campo de fuerza (OPLS-AA), temperatura (310 K) y presión (1 bar) empleadas para el dominio serinoproteasa libre en forma apo.

Previo al análisis de las variaciones en la desviación cuadrática media (RMSD) y la fluctuación media cuadrática de los residuos (RMSF) del dominio serinoproteasa de TMPRSS2 y RMSD del ligando ajustado a la proteína durante la trayectoria de producción, se confirmó el equilibrio termodinámico de todos los sistemas antes del inicio de la fase de producción.

Los resultados de la presente tesis se basan en un alcance exploratorio de las interacciones iniciales proteína-ligando, por lo que los tiempos de simulación ejecutados no son suficientes para demostrar una estabilidad concluyente a largo plazo. Sin embargo, nos permite identificar de manera comparativa el comportamiento dinámico inicial y la flexibilidad relativa de los ligandos en la triada catalítica del dominio serinoproteasa de TMPRSS2 de *Mus musculus.*

El complejo del dominio serinoproteasa de TMPRSS2-camostat mostró una pronta convergencia y estabilidad a los inicios de la simulación de 50 ns. El dominio serinoproteasa presentó un RMSD promedio de 0.205 ± 0.020 nm, mientras que el ligando mostró RMSD promedio de 0.245 ± 0.031 nm, sin cambios drásticos en la conformación del complejo. El complejo del dominio serinoproteasa de TMPRSS2-nafamostat mostró estabilidad durante los primeros 40 ns. El dominio serinoproteasa presentó un RMSD promedio de 0.212 ± 0.0175 nm, mientras que el ligando mostró

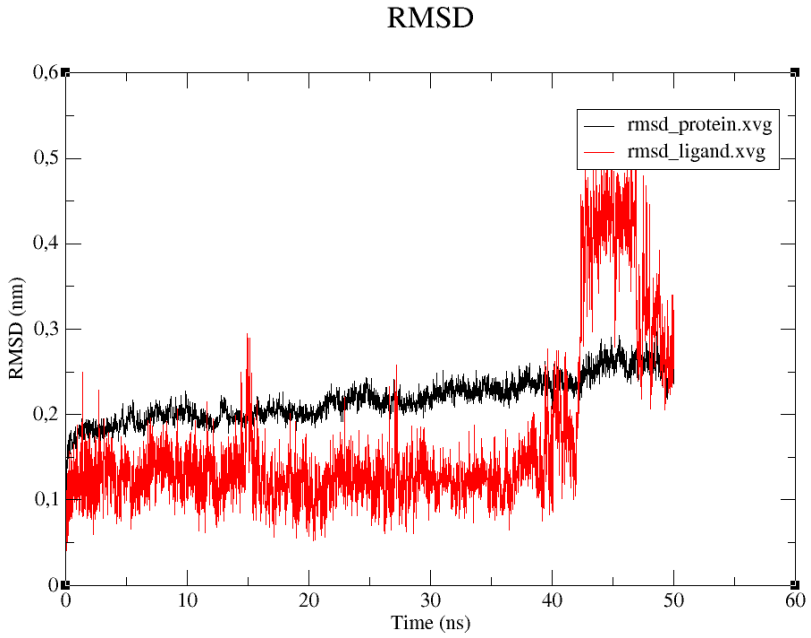
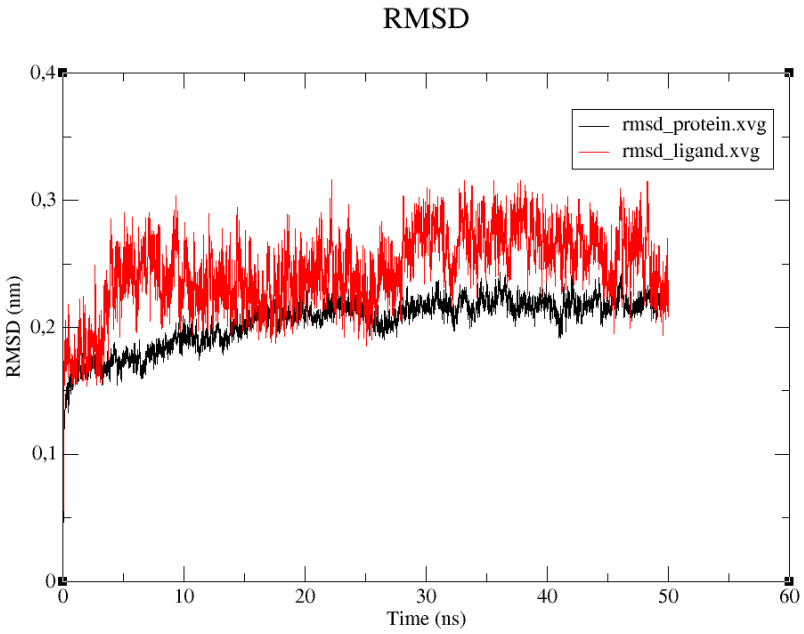
RMSD promedio de 0.126 nm ± 0. 028 nm. No obstante, tras los 40 ns, el RMSD del ligando experimentó un aumento abrupto, alcanzando aproximadamente 0.5 nm. Por otro lado, el complejo del dominio serinoproteasa deTMPRSS2-lopinavir mostró que el dominio serinoproteasa mantiene su estabilidad estructural durante la simulación con un RMSD promedio de 0.235 ± 0.020 nm, sin variaciones que indiquen cambios conformacionales bruscos. Sin embargo, lopinavir sufre fluctuaciones constantes con un RMSD promedio de 0.308 ± 0.047 nm. Finalmente, el RMSD del complejo del dominio serinoproteasa TMPRSS2-valrubicin muestra que el dominio serinoproteasa mantiene su estabilidad estructural a lo largo de la simulación con un RMSD promedio de 0.2101 ± 0.0161 nm, mientras que el ligando mostró RMSD promedio de 0.310 ± 0.080 nm evidenciando un comportamiento más fluctuante que el resto de los inhibidores analizados (Figura 25). De la misma forma, los valores de RMSF del dominio serinoproteasa de TMPRSS2 con los 4 ligandos evaluados indicaron que los residuos del sitio catalítico His41, Asp90 y Ser186 presentan fluctuaciones menores a 0.1 nm (Tabla 6 y Figura 26)

El número de puentes de hidrógeno en los complejos del dominio serinoproteasa de *Mus musculus* con los inhibidores de referencia, camostat y nafamostat, se mantuvo estable entre 2 a 4 y 3 a 5 interacciones durante toda la simulación. En contraste, con los ligandos exploratorios, el complejo con lopinavir mostró un leve incremento a partir de los 30 ns, manteniéndose por debajo de las 3 interacciones; por su parte, el complejo con valrubicina exhibió marcadas fluctuaciones de 0 a 7 interacciones a lo largo de toda la trayectoria (Tabla 6 y Figura 27).

La estabilidad del complejo con nafamostat se corroboró mediante el análisis de contactos atómicos, registrando un promedio de 132.78 ± 20.63. Al nanosegundo 15, un reajuste conformacional del ligando permitió que el sistema convergiera de manera estable entre 100 y 130 contactos hasta el final de la simulación. Por su parte, el

complejo con camostat promedió 96.18 ± 18.77 contactos, manteniéndose constante alrededor de los 100 contactos en el tramo final, lo que confirma su convergencia. En contraste, la inestabilidad del complejo con lopinavir se reflejó en un promedio de 104.31 ± 42.84 contactos; el ligando inició con interacciones bajas (30 a 90 contactos) los primeros 26 ns, para luego registrar un incremento abrupto y estabilizarse entre 120 y 180 contactos hacia los 50 ns. Finalmente, la inestabilidad con valrubicina fue más crítica, con un promedio de 146.51 ± 316.76 contactos, donde la ausencia de interacción durante la mayor parte del tiempo y la presencia de picos delgados e intermitentes evidencian una disociación temprana del ligando respecto al dominio SP (Anexo 23-Anexo 26).

Adicionalmente, la comparación de los complejos del dominio serinoproteasa TMPRSS2-inhibidor con su estado apo mostró un aumento moderado en los valores RMSD durante la simulación de 50 ns (Anexo 27 y Anexo 28), indicando cambios conformacionales asociados a la presencia del ligando, pero conservando su estabilidad estructural global. También, el análisis de RMSF mostró variaciones locales en la flexibilidad residual respecto al estado apo (Anexo 29 y Anexo 30) observándose en los residuos catalíticos una disminución de la movilidad conformacional, lo que sugiere que existe una estabilización local del sitio activo tras la interacción del ligando a la proteína. Los resultados de RMSF fueron consistentes con las distancias promedio entre los ligandos y los residuos catalíticos (His41, Asp90 y Ser186). Ya que, los inhibidores de referencia, nafamostat y camostat, presentaron menores distancias promedio hacia la tríada catalítica que los inhibidores exploratorios, lopinavir y valrubicina (Tabla 7).

Complejo	RMSD
nafamostat-TMPRSS2	
camostat-TMPRSS2	

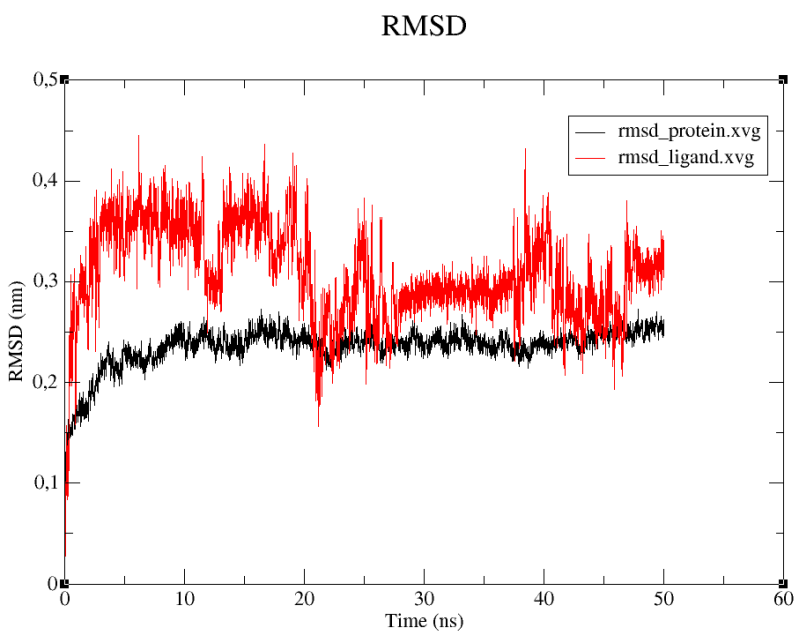
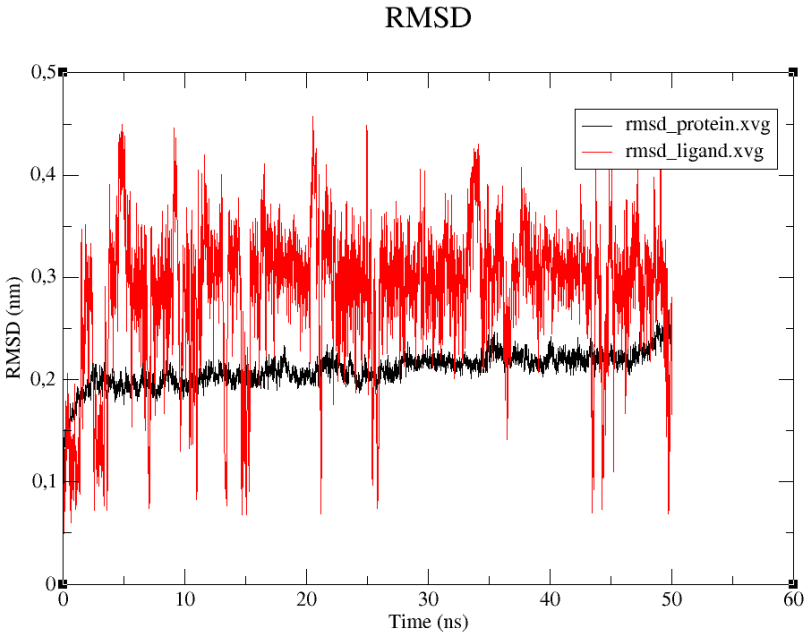
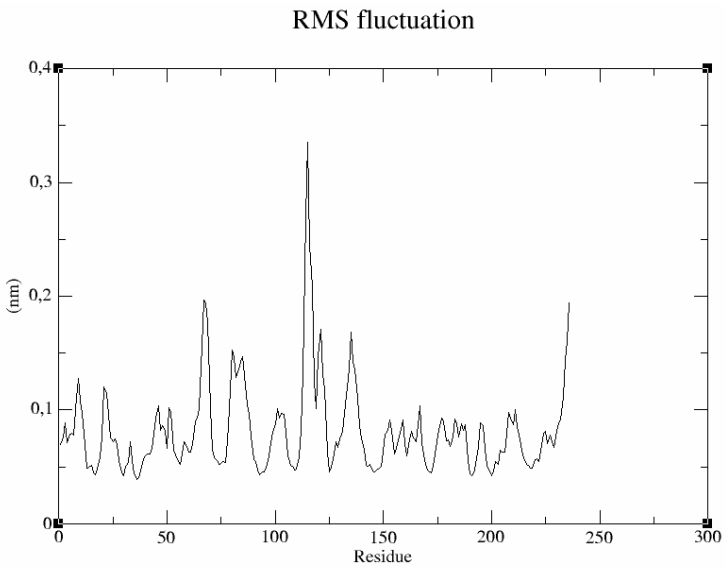
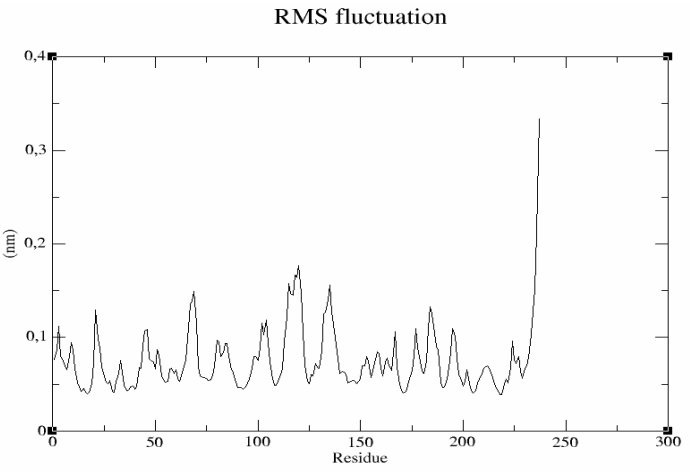
lopinavir - TMPRSS2	
valrubicin-TMPRSS2	

Figura 25. Desviación Media Cuadrática (RMSD) de las interacciones proteína-ligando del dominio serinoproteasa TMPRSS2 y 4 inhibidores. La simulación DM se ejecutó durante 50ns. El color negro representa la RMSD de la proteína principal y el violeta, la del ligando. La interacción del complejo es altamente estable si las fluctuaciones no superan los 1,5 Å (0.15 nm).

COMPLEJO	RMSF
lopinavir-TMPRSS2	
camostat-TMPRSS2	

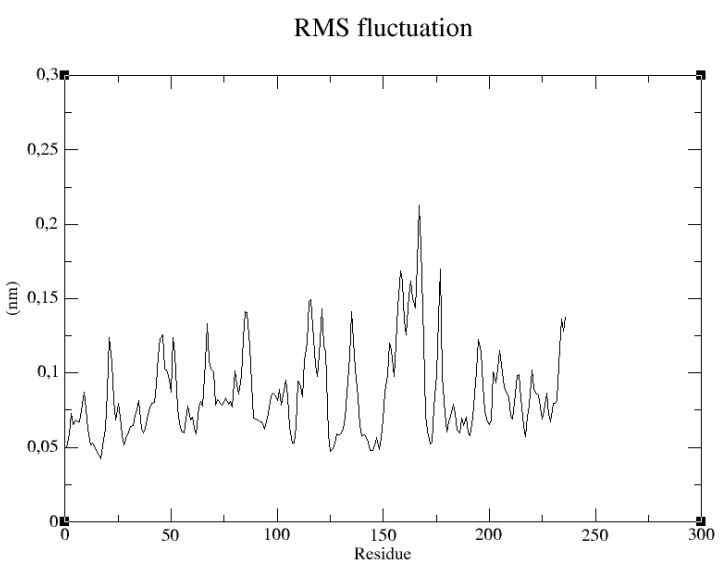
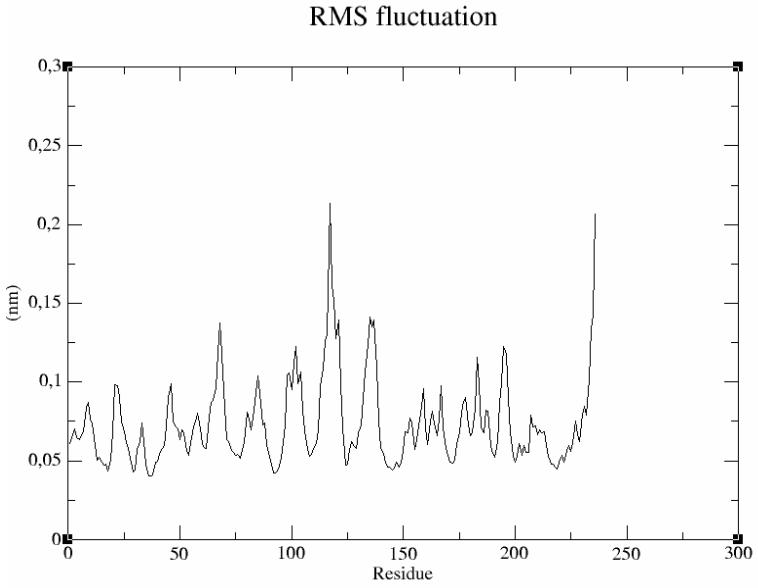
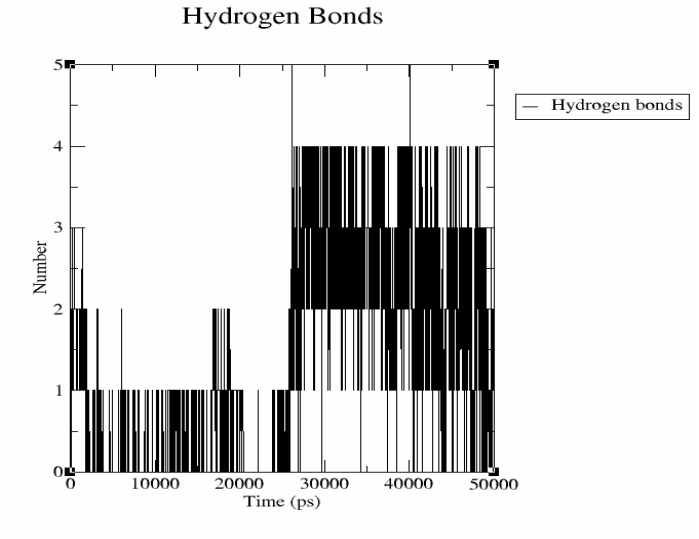
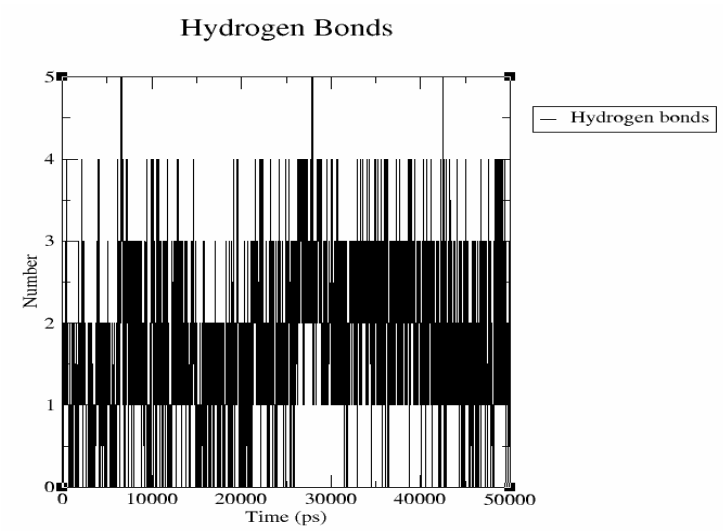
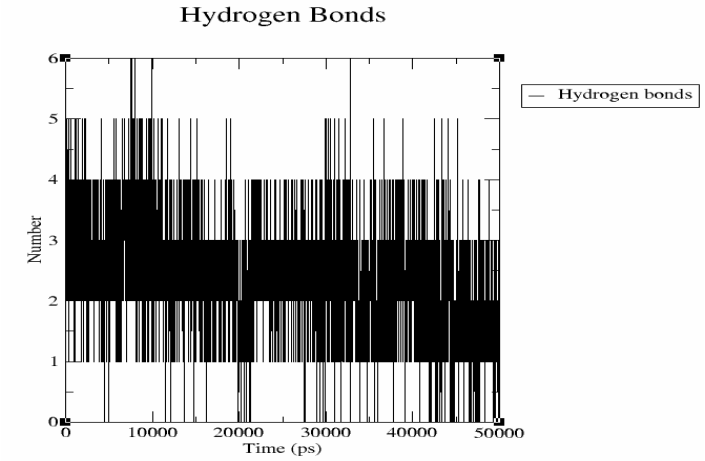
nafamostat-TMPRSS2	
valrubicin-TMPRSS2	

Figura 26. Fluctuaciones residuales (RMSF) del dominio serinoproteasa de TMPRSS2 en complejos con sus inhibidores durante 50 ns de dinámica molecular.

COMPLEJO	ENLACE PUENTES DE HIDRÓGENO
lopinavir-TMPRSS2	
camostat-TMPRSS2	
nafamostat-TMPRSS2	

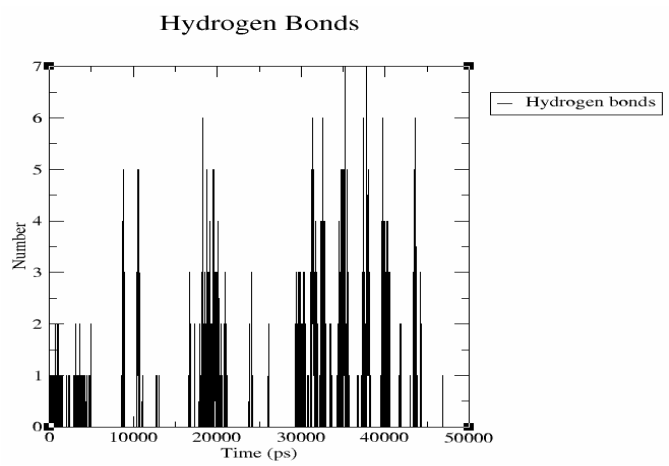
valrubicin-TMPRSS2	

Figura 27. Interacciones de puentes de hidrógeno entre el dominio serinoproteasa de TMPRSS2 y los ligandos durante la dinámica molecular.

Tabla 6. Resumen de los parámetros dinámicos de los complejos TMPRSS2–ligando durante 50 ns de simulación molecular.

Parámetro	Apo (Swiss Model)	Apo (AlphaFold)	Nafamostat	Camostat	Lopinavir	Valrubicina
RMSD proteína (nm)	0.132 ± 0.016	0.131 ± 0.015	0.212 ± 0.017	0.205 ± 0.020	0.235 ± 0.020	0.2101 ± 0.0161
RMSD ligando (nm)	-	-	0.126 ± 0. 028	0.245 ± 0.031	0.308 ± 0.047	0.310 ± 0.08
Puentes de hidrógeno	-	-	3 a 5	2 a 4	<3	0-7

V.1.12.1. Análisis de distancias respecto al centro de masa (COM) y tiempo de permanencia

Para determinar la estabilidad posicional y el tiempo de permanencia de los ligandos en el dominio serinoproteasa de TMPRSS2, se evaluó la distancia entre el centro de masa (COM) de la proteína y el COM de cada compuesto a lo largo de los 50000 ps (50 ns) de simulación. Los valores promedio, las desviaciones estándar y la estimación del tiempo de permanencia se muestran en la Tabla 7.

Tabla 7. Valores promedio de distancia entre centros de masa (COM) y estimación del tiempo de permanencia intermolecular.

Complejo	Distancia Promedio Proteína-Ligando	Tiempo de permanencia estimado
TMPRSS2-Nafamostat	*1.07 ± 0.04*	50 ns
TMPRSS2-Camostat	1.46 ± 0.06	50 ns
TMPRSS2-Valrubicina	*3.30 ± 0.84*	50 ns (reacomodo a los 22 ns)
TMPRSS2-Lopinavir	*2.20 ± 0.32*	< 5 ns (Disociación temprana)

Los inhibidores de referencia mostraron las distancias más cortas y estables durante la simulación de 50 ns. Por un lado, nafamostat mostró un promedio de 1.07 ± 0.04 nm, manteniéndose estable con un leve reajuste cerca del nanosegundo 42. Por otro lado, camostat fluctuó con un promedio de 1.46 ± 0.06 nm. De esta forma, ambos ligandos lograron un tiempo de permanencia del 100% en la simulación. Sin embargo, el complejo con lopinavir promedió 2.20 ± 0.32 nm; su trayectoria inició con valores altos de hasta 3.0 nm, pero cayó a los 22 ns para luego estabilizarse a los 2.0 nm. Por otro lado, valrubicina se disoció antes de los 5 ns. La distancia del ligando aumentó desde 1.5 nm hasta picos de 5.0 nm. Es por ello que, el complejo mostró un promedio más alto (3.30 nm) y la mayor desviación estándar (± 0.84 nm) (Anexo 31-Anexo 34).

V.1.12.2. Análisis del Radio de Giro (R_g)

Para evaluar la estabilidad macromolecular y el grado de compactación del dominio serinoproteasa de TMPRSS2 bajo condiciones dinámicas, se calculó el Radio de Giro (R_g). Los valores promedio generales y desviación estándar se muestran en la Tabla 8.

Tabla 8. Análisis cuantitativo del Radio de Giro (R_g) para los complejos del dominio serinoproteasa de TMPRSS2.

Complejo	Promedio de Radio de Giro (nm)
TMPRSS2 - Camostat	1.714± 0.006
TMPRSS2 - Nafamostat	1.730± 0.013
TMPRSS2 - Lopinavir	1.656± 0.008
TMPRSS2 - Valrubicina	1.657± 0.006
TMPRSS2	1.714± 0.009

Los inhibidores de referencia exhibieron los valores de $R_g$más elevados. Camostat se mantuvo estable con un promedio de 1.714 ± 0.006 nm y fluctuaciones mínimas. También, nafamostat mostró una media de 1.730 ± 0.013 nm, mostrando un incremento desde 1.71 nm hasta estabilizarse en 1.75 nm al final de la trayectoria. Mientras que los inhibidores exploratorios mostraron una mayor compactación en la proteína. El complejo con lopinavir promedió 1.656 ± 0.008 nm; su perfil descendió los primeros 5 ns hasta un mínimo de 1.63 nm, pero luego se equilibró en 1.66 nm. El complejo con valrubicina estuvo estable durante los 50 ns, sin fluctuaciones (Anexo 35-Anexo 39).

V.1.12.3. Monitoreo de distancias hacia los residuos de la tríada catalítica

Para determinar la cercanía y estabilidad de los ligandos frente al sitio activo de TMPRSS2, se midieron las distancias intermoleculares respecto a los residuos catalíticos His41, Asp90 y Ser186. Los valores promedio y sus desviaciones estándar se presentan en la Tabla 9.

Tabla 9. Distancias promedio calculadas entre el centro de masa (COM) del ligando y los residuos catalíticos de TMPRSS2 durante los 50 ns de simulación. Valores menores indican mayor proximidad y posible interacción estable.

Residuo catalítico	Nafamostat (nm)	Lopinavir (nm)	Valrubicina (nm)	Camostat (nm)
His41	0.66045 ± 0.04663 nm	2.02820 ± 0.34912 nm	3.13838 ± 0.91520 nm	0.94326 **±** 0.06985 nm
Asp90	1.02484 ± 0.04886 nm	2.53923 ± 0.30822 nm	3.15951 ± 0.90107 nm	1.28957 **±** 0.06836 nm
Ser186	0.54325 ± 0.05614 nm	1.49683± 0.32317 nm	3.13505 ± 0.91189 nm	0.99416 **±** 0.11494 nm

Los inhibidores de referencia presentaron la mayor cercanía al sitio activo durante la simulación. Nafamostat mostró las menores distancias respecto a His41 y Ser186, manteniéndose estable a lo largo de la trayectoria, mientras que camostat mostró un acoplamiento sostenido con fluctuaciones mínimas. En contraste, los inhibidores exploratorios presentaron mayores distancias. Lopinavir inició alejado del núcleo catalítico, pero tras un reacomodo alrededor de los 22 ns redujo su separación y permaneció más próximo a Ser186 hacia el final de la simulación. Por su parte, valrubicina exhibió amplias fluctuaciones durante los 50 ns y registró las mayores distancias respecto a la tríada catalítica, desplazándose fuera del bolsillo de unión (Anexo 40- Anexo 43).

V.1.12.4. Variación de las distancias intermoleculares en la tríada catalítica

Se midieron las distancias intermoleculares promedio entre los residuos de la tríada catalítica (His41, Asp90 y Ser186) a lo largo de los 50000 ps de simulación. Los valores promedio y sus respectivas desviaciones estándar para el estado libre (*apo*) y los complejos (*holo*) se resumen en la Tabla 10 y Tabla 11.

Tabla 10. Valores promedio de las distancias intermoleculares (nm) y desviaciones estándar en la tríada catalítica de TMPRSS2 con nafamostat, lopinavir y valrubicina respecto a su estructura apo.

Distancia	Apo_Swiss	TMPRSS2-Nafamostat	TMPRSS2-Lopinavir	TMPRSS2-Valrubicina
His 41-Asp90	0.568 ± 0.0163	0.568 ± 0.01511	0.5930 ± 0.05111	0.5956 ± 0.0126
His 41-Ser186	0.8968 ± 0.1323	0,736 ± 0.0359	0.7793 ± 0.072558	0.8846 ± 0.0544
Asp 90-Ser186	1.0814 ± 0.1389	0,987 ± 0.0416	1.1618 ± 0.073825	1.2220 ± 0.0587

Tabla 11. Valores promedio de las distancias intermoleculares (nm) y desviaciones estándar en la tríada catalítica de TMPRSS2 con camostat respecto a su estructura apo.

Distancia	Apo_Alphafold	TMPRSS2-Camostat
His 41-Asp90	0.569 ± 0.0165	0.5599 ± 0.01293
His 41-Ser186	0.8970 ± 0.1302	0.8541 ± 0.08558
Asp 90-Ser186	1.0818 ± 0.1380	1,2479 ± 0.097379

Las distancias entre los residuos de la tríada catalítica mostraron diferencias según el ligando unido. En general, nafamostat y camostat mantuvieron los perfiles más estables, con una menor separación entre His41 y Ser186, mientras que lopinavir presentó una mayor variabilidad, especialmente en His41–Asp90. En cambio, valrubicina mostró un comportamiento similar al estado apo, lo que sugiere una menor influencia sobre la conformación del sitio catalítico (Anexo 44 – Anexo 46).

VI. DISCUSIÓN

A lo largo de las primeras décadas del siglo XXI, han surgido epidemias y pandemias causadas por virus de la influenza y coronavirus, lo que representa un riesgo significativo para la salud pública. Ante estas crisis sanitarias, se han explorado

estrategias antivirales para mitigar la patogenicidad de ambos virus. Sin embargo, la evidencia clínica sobre la eficacia y seguridad de los fármacos dirigidos a proteínas virales sigue siendo inconclusa, debido principalmente a la capacidad de los virus para mutar (Shen et al., 2017). Además, otro de los desafíos para el desarrollo de vacunas y tratamientos contra estos virus es la falta de modelos de animales apropiados (Brooke y Prischi, 2020). Debido a la conservación de TMPRSS2 en diversas especies, el ratón albino (*Mus musculus*) ha sido propuesto como modelo en investigación biomédica debido a su susceptibilidad a infecciones como la influenza y SARS-CoV (Brooke y Prischi, 2020; Menéndez, 2005; van den Brand et al., 2014). TMPRSS2 de ratón constituye una proteína de interés en estudios orientados a la evaluación de compuestos con potencial capacidad de interacción con proteasas, debido a su participación en la activación de virus como influenza A H7N9, H1N1, SARS-CoV-2, SARS-CoV y MERS-CoV (Hatesuer et al., 2013; Sakai et al., 2014; Tarnow et al., 2014; Shen et al., 2017; Iwata-Yoshikawa et al., 2019). Sin embargo, ante la escasez de estudios estructurales sobre TMPRSS2 murina, los hallazgos de esta tesis aportan información sobre su conformación tridimensional y sobre las interacciones moleculares predichas con inhibidores de referencia y exploratorios.

VI.1. Caracterización *in silico* del dominio serinoproteasa de TMPRSS2: propiedades fisicoquímicas, predicción estructural y estabilidad molecular

El modelado molecular de proteínas es una herramienta computacional crucial en la bioinformática y biología estructural, ya que permite predecir la estructura tridimensional de una proteína objetivo a partir de su secuencia de aminoácidos permitiendo acelerar el descubrimiento de nuevos fármacos o reutilización de los mismos, con el fin de reducir significativamente los costos y el tiempo necesario en los procesos de desarrollo e investigación (Adelusi et al., 2022). En la presente tesis se utilizaron cuatro métodos de predicción de la estructura tridimensional del dominio

serinoproteasa de TMPRSS2 de ratón que son el modelo basado en plantillas (modelo de homología y *threading*), modelado sin plantillas (*ab initio*) y el algoritmo Alphafold, lo que permitió comparar la consistencia estructural del dominio serinoproteasa de TMPRSS2 de *Mus musculus* bajo distintos algoritmos. Es importante resaltar que el uso de métodos computacionales permite evaluar la robustez del modelo estructural obtenido, lo cual resulta relevante para estudios comparativos entre especies y para la exploración de posibles interacciones moleculares con compuestos reportados en la literatura como inhibidores de proteasas. Además, es importante el uso de las “3R” en investigaciones con animales de laboratorio, la cual se basa en principios fundamentales de reemplazo, reducción y refinamiento, ya que esto puede minimizar significativamente la dependencia de modelos animales en investigaciones farmacéuticas, permitiendo una investigación más ética y eficiente (Ekstrand et al., 2021).

La evaluación de la estructura del gen *Tmprss2* de ratón (GenBank ID: NC_000082.7) reveló su localización en el cromosoma 16, con una organización de 13 exones y 12 intrones. Su región codificante (CDS) inicia en la posición 231 del segundo exón y se extiende hasta la posición 1703 (Figura 6). Un estudio ha demostrado que los genes *ACE2* y *Tmprss2* se expresan ampliamente en los pulmones de múltiples mamíferos, incluidos *Rattus norvegicus*, *Mus musculus*, *Oryctolagus cuniculus*, *Macaca mulatta* y *Canis lupus familiaris* (Brooke y Prischi, 2020), lo que respalda su uso como modelo comparativo en estudios estructurales.

La evaluación de la estructura primaria del dominio serinoproteasa de TMPRSS2 de ratón permite analizar los residuos que lo conforman y sus características fisicoquímicas. La herramienta ProtParam calcula diferentes parámetros en función de la secuencia analizada, proporcionando información relevante sobre el comportamiento de la proteína. El punto isoeléctrico (pI) calculado sugiere que la proteína es ácida, lo que influye en su separación mediante técnicas basadas en carga, como la electroforesis isoeléctrica, en la que las proteínas migran hasta alcanzar la región de pH

correspondiente a su pI. Además, la combinación de la masa molecular (25.923 kDa) y el pI (6.21) puede ser utilizada en la electroforesis en gel bidimensional para su caracterización. El coeficiente de extinción (EC) muestra la luz absorbida por una proteína en relación con su composición a una longitud de onda específica. El Índice de Inestabilidad (II), inferior a 40, sugiere que TMPRSS2 es estable *in vitro*. Según el valor GRAVY indica que el dominio evaluado de TMPRSS2 es hidrofílico (Gasteiger et al., 2005). Esta hidrofilicidad es biológicamente coherente, debido a que el dominio serinoproteasa de TMPRSS2 se encuentra expuesto al microambiente acuoso del espacio extracelular y alveolar, el cual requiere una alta interacción con el solvente para mantener su conformación activa.

Al comparar el pI estimado de TMPRSS2 (6.21) con el de otras serinoproteasas extracelulares, como la de *Cryptococcus neoformans* (pI: 4.77), se observa una diferencia atribuible a la distancia evolutiva entre ambas especies. Sin embargo, esta discrepancia no parece afectar de manera determinante su funcionalidad, ya que en ambos casos se ha reportado actividad proteolítica en condiciones cercanas al pH fisiológico. En el caso de *C. neoformans*, dicha actividad ha sido asociada con procesos de virulencia y evasión del sistema inmune (Yoo et al., 2004). Estos hallazgos sugieren que la actividad proteolítica de las serinoproteasas no está completamente restringida por su pI, ya que la enzima sigue siendo funcional catalíticamente en un rango de pH fisiológico del organismo huésped.

De la misma forma en otro estudio se caracterizó la presencia y actividad enzimática de TMPRSS2 en vesículas extracelulares derivadas del líquido seminal humano (SF_Evs) y se evaluó potenciales inhibidores de esta proteasa mediante ensayos bioquímicos. La presencia física y el tipo de fragmentos proteicos presentes en el lisado de SF_Evs mediante la técnica de SDS-PAGE y detección por Western blot detectó un patrón de bandas que coinciden con la masa teórica del dominio peptidasa S1(~ 26 kDa) (Verhulst et al., 2025). Esta concordancia molecular demuestra la

conservación estructural de la región catalítica entre mamíferos y confirma que el fragmento analizado corresponde al dominio maduro remanente tras el proceso de autoprocesamiento proteolítico característico de las serinoproteasas.

El alineamiento de secuencias múltiples de TMPRSS2 en distintos modelos experimentales revela una alta conservación del dominio serinoproteasa, que contiene la tríada catalítica (Figura 9). Además, los residuos que conforman el bolsillo de unión de TMPRSS2 en humanos son idénticos en ratón, cobayo, hámster, rata, conejo, macaco, gato, perro y hurón. No se observaron diferencias significativas en comparación con la secuencia humana, lo que sugiere una elevada conservación de los residuos asociados al reconocimiento del sustrato. Estos hallazgos son consistentes con estudios previos que han propuesto que TMPRSS2 de distintas especies podría presentar funciones similares en los mecanismos de activación proteolítica viral (Brooke y Prischi, 2020).

El alineamiento global por pares mediante BLASTp del dominio serinoproteasa (Peptidasa S1) mostró un porcentaje de identidad de secuencia del 84.48 % y una similitud del 90.94 % entre *Mus musculus* y *Homo sapiens*. Esta elevada homología primaria podría justificar el uso del modelo murino para la extrapolación de datos estructurales hacia la proteína humana. El análisis comparativo en base a la estructura cristalina resuelta por Fraser et al. (2022) ayuda a interpretar las coincidencias y divergencias entre ambas proteínas. Los autores mencionan que el dominio adopta el pliegue canónico de quimotripsina/tripsina. En esta conformación, la especificidad del sustrato y el acceso a la tríada catalítica (His294, Asp345 y Ser441) dependen de bucles superficiales altamente variables (Bucles A–E y Bucles 1–3). Al contrastar nuestros resultados, se evidenció una coincidencia estructural absoluta (100 % de identidad) en el núcleo del bolsillo de unión primario (subsitio $S1$). En específico, los residuos Asp433, Gly460 y Gly462 (en numeración murina) se encuentran conservados en ambas secuencias (Figura 9). Con base en los hallazgos encontrados por Fraser et al. (2022)

y Hoffmann et al. (2020), esta conservación es crucial, ya que el residuo Asp433 interactúa directamente mediante puentes salinos con las argininas del sitio de escisión polibásico de la proteína Spike del SARS-CoV-2. Por lo tanto, la identidad en el subsitio $S1$ sustenta la hipótesis de que el mecanismo fundamental de reconocimiento y procesamiento proteolítico del virus se mantiene conservado en el modelo murino.

Por el contrario, se encontró discrepancias estructurales en las regiones de lazos flexibles o bucles periféricos que flanquean los subsitios secundarios $S2$, $S3$ y $S4$. Específicamente, en el Bucle A que rodea el acceso al sitio activo, se identificó la sustitución del triplete humano 337-SKN-339 por la secuencia murina 337-SKT-339, adyacente al residuo modulador Thr341. También, se observan variaciones en el bucle externo 429-LQQ-431. Como describe el estudio estructural de Fraser et al. (2022), estos bucles superficiales son altamente flexibles y determinan la topografía periférica de la cavidad. Debido a esto, un mismo inhibidor competitivo puede mostrar variaciones en sus poses de acoplamiento o ligeras fluctuaciones en su energía de unión al comparar las dos especies durante las simulaciones. Esto ocurre por sutiles modificaciones en el volumen accesible al solvente y nuevos contactos polares secundarios en la entrada del pocket.

La comparación realizada demuestra que *Mus musculus* comparte el mismo núcleo catalítico y bolsillo de especificidad $S1$ que el humano, de esta manera sería un modelo ideal para un cribado inicial de compuestos, cuyas diferencias en los bucles periféricos marcan límites estructurales normales que deberán ser validados mediante ensayos funcionales biológicos posteriores.

El modelado de serinoproteasa de TMPRSS2 de ratón a través de los servidores AlphaFold2, Swiss-Model, Phyre2 y Robetta permitió obtener su estructura tridimensional. Los modelos predichos por Swiss Model y Phyre2 utilizaron como plantillas las estructuras 7MEQ.1 (84.32%) y 7XYD.1. B (84.32%) para Swiss-Model, mientras que Phyre2 empleó la plantilla c7meqA (84%). Dado que los modelos

generados presentaron una identidad superior al 40-50%, se considera que alcanzan una precisión adecuada para diversas aplicaciones biológicas (Lesk, 2014). Todas las plantillas fueron obtenidas mediante cristalografía de rayos X. Las estructuras de referencia 7meq (Fraser et al., 2022) y 7xyd (Wang et al., 2023) corresponden a la proteasa transmembrana de serina humana y muestran una identidad del 84.32% con TMPRSS2 de ratón. La diferencia en la predicción de hélices alfa en el modelado por homología se debe a la resolución de los cristales usados como molde. Por un lado, 7MEQ se resolvió a una resolución superior de 1.95 Å, mientras que 7XYD presenta una resolución menor de 2.58 Å. Conforme a esto, la mayor nitidez de los datos en 7MEQ reduce el ruido en la densidad electrónica. Por lo tanto, el algoritmo logra delimitar con alta precisión los bucles periféricos más flexibles.

Luego, se seleccionó un modelo generado por el servidor Robetta y los 5 modelos generados por el algoritmo AlphaFold. En total, los 9 modelos fueron evaluados mediante PDBsum y ModFold para determinar su calidad estructural, todos los modelos se encontraron dentro de los estándares de aceptación en los 2 métodos de validación (Anexo 10 -11). Entre ellos, el modelo 3 de los cinco generados por AlphaFold destacó por su calidad estructural, este modelo representa la conformación más estable de la enzima en su estado apo libre de tensiones conformacionales inducidas por la presencia de un ligando.

Antes del modelado tridimensional de TMPRSS2 de ratón, se predijo su estructura secundaria. Los modelos generados a través de Swiss-Model (modelo 1) y por *threading* (Figura 11A, Figura 12) coincidieron en la formación de cuatro hélices alfa y doce láminas beta, en concordancia con la predicción de PSIPRED. Por otro lado, el modelo 2 de Swiss-Model y el modelo 3 de AlphaFold (Figura 11B, Figura 14) mostraron cinco hélices alfa y doce láminas beta, lo que se alineó con los resultados obtenidos por ESPript3, con la diferencia en el tamaño de la tercera hélice alfa. En contraste, el modelo generado por Robetta presentó seis hélices alfa y doce láminas beta, aunque la sexta

hélice no coincidió con las predicciones de ningún software de estructura secundaria (Figura 13). La sexta hélice ($\alpha 6$) predicha por Robetta podría ser un artefacto matemático de la minimización energética. Al modelar sin plantilla, los algoritmos *ab initio* fuerzan la creación de puentes de hidrógeno para estabilizar artificialmente los bucles y bajar la energía libre. Por el contrario, la fluctuación entre 4 y 5 hélices en los otros modelos responde a la flexibilidad de esa región. En el caso de AlphaFold predijo un consenso evolutivo para la secuencia de ratón que estabiliza la hélice $\alpha 3$, mientras que la variante humana de 7MEQ favorece un bucle desordenado.

El alineamiento de los cinco modelos generados mediante homología, *threading*, *ab initio* y el modelo 3 de AlphaFold (Figura 15A) mostró que las hélices alfa y láminas beta se mantienen estructuralmente estables. Sin embargo, se observaron ligeras variaciones en los bucles, lo cual era esperable, dado que estas regiones suelen ser flexibles, debido a que pueden contener sitios de unión para ligando (Teilum et al., 2011). En términos generales, no se encontraron diferencias significativas entre los modelos, lo que concuerda con estudios previos que indican que TMPRSS2 de *Mus musculus* comparte aproximadamente el 85% de similitud con su homólogo humano (Hatesuer et al., 2013).

La ubicación de la tríada catalítica es clave para la funcionalidad de TMPRSS2, ya que debe ser accesible al sustrato para permitir el clivaje eficiente. En la Figura 15B, se observa que la región de unión al sustrato se encuentra en un área superficial delimitada por interacciones de Van der Waals, lo que favorece su acoplamiento tanto con sustratos naturales como con inhibidores potenciales. Además, se confirmó la conservación de la tríada catalítica de TMPRSS2 en ratón, compuesta por His41, Ser186 y Asp90. Esta tríada cumple un papel fundamental en la actividad enzimática, funcionando como un relé de carga, donde His41 estabiliza la Ser186 y facilita el ataque nucleofílico al sustrato, mientras que Asp90 estabiliza a His41 y contribuye al reconocimiento del sustrato (Escalante & Ferguson, 2021).

Por otro lado, un estudio realizado por Rensi et al. (2020), en el que se alineó el dominio serinoproteasa de TMPRSS2 humano utilizando como plantillas la calicreína plasmática humana, el factor XIa y la hepsina, reportó variaciones estructurales notorias en los bucles fuera del sitio catalítico. En particular, identificaron una variabilidad significativa en la región comprendida entre los residuos 83-88 (NKTKSD), sugiriendo que esta zona podría influir en el control del tamaño de los ligandos que pueden acceder a la cavidad catalítica (Figura 15C).

Tras el modelado molecular y la validación estructural de los cinco modelos predichos, estos fueron sometidos a dinámica molecular con el objetivo de optimizar su estructura, minimizar la energía y evaluar su estabilidad en un periodo corto de 10 ns en solución, utilizando GROMACS. El análisis de los valores de RMSD y RMSF permitió seleccionar dos modelos que alcanzaron la estabilidad sin presentar fluctuaciones bruscas durante la simulación: el modelo generado por homología (plantilla 7meq) y el modelo 3 de AlphaFold. El modelado por homología con Swiss-Model ha sido ampliamente utilizado para predecir la estructura de TMPRSS2 en humanos, logrando valores de identidad entre 34.9% y 43.6%, y en algunos casos 33.82%, cuando se emplearon como plantillas proteínas estructuralmente similares, como prostasina, calicreína plasmática humana y hepsina (Escalante & Ferguson, 2021; Mahmudpour et al., 2021). Hasta la fecha, esta técnica sigue siendo una herramienta valiosa en la predicción de nuevas estructuras proteicas.

Por otro lado, el modelado mediante AlphaFold utiliza aprendizaje automático e incorpora conocimientos físicos y biológicos sobre la estructura proteica para generar modelos estructurales de alta precisión. Un estudio realizado por Frumenzio et al., (2023) modeló con alta precisión la estructura tridimensional de TMPRSS2 humano utilizando la plantilla 7meq y completó su estructura con AlphaFold, incluyendo el dominio transmembrana. En este estudio, se destacó que la inclusión del dominio transmembrana reveló la presencia de una cavidad entre los dominios LDLRA/SRCR y

serinoproteasa, lo que sugiere una mayor flexibilidad estructural y abre nuevas oportunidades para el diseño de inhibidores más eficientes.

VI.2. Acoplamiento molecular del dominio serinoproteasa de TMPRSS2 y los 4 inhibidores por AutoDock

En el presente estudio, se identificó que el dominio serinoproteasa de TMPRSS2 en el modelo de ratón está conformado por la tríada catalítica Ser186, His41 y Asp90, la cual es esencial para su actividad enzimática. TMPRSS2 puede reconocer y escindir las secuencias en la glicoproteína S del SARS-CoV-2 al igual que la hemaglutinina de la influenza permitiendo la entrada viral a la célula huésped (Böttcher-Friebertshäuser et al., 2010; Brooke y Prischi, 2020). Para bloquear la actividad de la triada catalítica, los inhibidores de proteasas deben establecer interacciones específicas con estos residuos clave que forman parte del relé de carga (Escalante & Ferguson, 2021).

Los cuatro ligandos evaluados fueron clasificados como inhibidores reconocidos de referencia (nafamostat y camostat) e inhibidores de interés exploratorio para reposicionamiento farmacológico (lopinavir y valrubicina). La inclusión de los dos últimos permite contrastar si la afinidad teórica predicha por el docking molecular se sostiene al ser evaluados en dinámica molecular.

El análisis comparativo del acoplamiento molecular muestra que los inhibidores de bajo masa molecular, nafamostat y camostat, se ubican en el interior del bolsillo S1 de TMPRSS2, a diferencia de compuestos de mayor masa molecular, valrubicina y lopinavir, los cuales quedan limitados a interactuar con las regiones periféricas externas debido a restricciones de espacio.

Estas simulaciones se realizaron utilizando el algoritmo genético lamarckiano (LGA) como método de búsqueda, basado en el principio de que una menor energía de unión indica una interacción más favorable entre la proteína y el ligando (Ravi, 2016).

VI.2.1. Análisis e interpretación de los inhibidores referenciales de bajo masa molecular: Nafamostat y Camostat

Nafamostat es un inhibidor de la serina proteasa utilizado principalmente en Japón y Corea para el tratamiento de la pancreatitis aguda y la coagulación intravascular diseminada (Hernández-Mitre et al., 2022; Mantzourani et al., 2022). Al igual que camostat, nafamostat se seleccionó como inhibidor de referencia porque ha demostrado experimentalmente actividad inhibitoria frente a TMPRSS2. Mientras que, camostat es un fármaco usado como tratamiento para la pancreatitis y la esofagitis por reflujo postoperatorio en Japón (Mantzourani et al., 2022). Además, se seleccionó debido a que ha demostrado experimentalmente bloquear la actividad proteolítica de esta serinoproteasa. Por ello, fue considerado un inhibidor de referencia.

Al evaluar los inhibidores de referencia, nafamostat presentó la pose con la energía libre de unión de enlace más baja y favorable de todo el estudio (-5.52 kcal/mol). Mientras que, camostat mostró una menor afinidad teórica inicial, con una energía de enlace de -3.19 kcal/mol. Al contrastar estos hallazgos con estudios previos, se observó una clara coincidencia en la jerarquía de los valores energéticos. La superioridad termodinámica predicha de nafamostat sobre camostat mostrada guarda relación con lo reportado por Kishk et al. (2020) y Sonawane et al. (2021), en la que los cálculos en AutoDock alcanzaron -7.20 kcal/mol para nafamostat y -6.23 kcal/mol para camostat, volviendo a nafamostat como el ligando con mayor afinidad teórica en la parte estática.

El análisis del acoplamiento molecular permitió interpretar estas diferencias energéticas desde un punto de vista estructural. La arquitectura lineal de nafamostat parece favorecer un anclaje más profunda en el bolsillo de especificidad primaria S1, estableciendo interacciones tipo puentes de hidrógeno teóricos con residuos clave como Gly209 (2.79 Å y 3.01 Å) y Gly207 (3.47 Å), además de interacciones con Glu45 (2.87 Å), Ser181 (3.10 Å) y Cys42 (3.53 Å). Además, se identificaron un posible puente salino y contactos atómicos cercanos con residuos de la triada catalítica como His41 (3.28–

4.51 Å), así el grupo éster central de nafamostat presentó una proximidad de 3.28 Å con Ser186. Todas estas interacciones indican una orientación favorable de nafamostat dentro del sitio activo. Desde un punto de vista fisicoquímico, este comportamiento es consistente con las propuestas de Ahuja (2024), Salleh y Deris (2022) y Fraser et al. (2022), quienes plantearon que el grupo guanidino favorece el posicionamiento del grupo éster frente a Ser186, lo que constituye un paso previo necesario para la posterior formación del enlace covalente. Esto se relaciona con el acoplamiento bidentado sobre Gly209 y Gly207 observado en nuestro modelo estructural, el cual se encuentra orientando de forma óptima la molécula para inmovilizar electrostáticamente el relé de carga. Esta disposición estructural coincide con la descrita por Kishk et al. (2020) y con la estructura cristalográfica de Fraser et al. (2022), quienes reportaron la ocupación del bolsillo S1 por nafamostat y su interacción con residuos catalíticos de TMPRSS2.

No obstante, se evidencia una discrepancia metodológica crítica con el estudio de Ramakrishnan et al. (2021). Mediante un algoritmo de acoplamiento por ajuste inducido (IFD) en la plataforma Schrödinger, dichos autores reportaron que nafamostat carecía de interacciones directas con los residuos catalíticos His41 y Ser186, quedando limitado únicamente a un contacto con la Asp180. Esta divergencia puede explicarse por la naturaleza de los algoritmos de búsqueda conformacional. El modelado de flexibilidad inducida permite el movimiento de las cadenas laterales del receptor, lo que en el estudio de Ramakrishnan et al. (2021) provocó un ensanchamiento de la cavidad que desplazó al nafamostat hacia el fondo del bolsillo. En contraste, el algoritmo estático con receptor rígido empleado en la presente tesis captura el estado de afinidad geométrica basal de la forma compacta de la proteasa. Bajo este escenario, la linealidad del nafamostat le otorga una ventaja estérica para aproximarse de forma directa y estrecha al núcleo catalítico, estableciendo los contactos descritos con la His41 y Ser186.

En contraste, camostat adoptó una orientación más extendida, ocupando tanto el bolsillo S1 como regiones del bolsillo secundario S2. El ligando estableció puentes de hidrógeno con Gly184 (2.07 Å y 2.20 Å) y presentó interacciones con Lys87, Tyr82 e Ile91, además de contactos con Val25, Gln183 y Glu44. Sin embargo, en el sitio catalítico únicamente se observaron puentes de hidrógeno con Ser186 (2.65 Å) y Asp90 (2.5 Å y 2.7 Å), así como contactos hidrofóbicos con His41 (3.19 Å y 3.23 Å). Este patrón extendido coincide con el mapeo reportado por Ramakrishnan et al. (2021), confirmando que la topología ramificada del camostat posee una tendencia intrínseca a expandir sus contactos hacia subsitios periféricos adyacentes como S2.

Estas diferencias en el patrón de interacción, sumadas a la naturaleza estática del receptor rígido en el algoritmo, explican la menor afinidad de unión predicha para camostat (−3.19 kcal/mol) respecto a nafamostat. Al mantener la proteína fija, AutoDock penaliza matemáticamente la pérdida de entropía conformacional del ligando. Por lo tanto, la estructura más larga y flexible de camostat recibe una mayor penalización energética al restringir los grados de libertad de sus enlaces rotatorios para ocupar múltiples subsitios simultáneamente, a diferencia del nafamostat cuya estructura es más compacta (Huey et al., 2009).

Por otro lado, estudios bioquímicos de Shrimp et al. (2020) y Wang et al. (2023) demostraron experimentalmente que nafamostat posee una potencia inhibitoria superior a la de camostat, los resultados de la presente tesis únicamente muestran una interacción molecular predicha compatible con dichos hallazgos y no constituyen una demostración directa de la inhibición enzimática.

Finalmente, el uso de los 4 modelos generados por las diferentes técnicas fue primordial, ya que se evitó el sesgo de confirmación algorítmica. El modelo de SWISS-MODEL heredó la conformación *holo* de la plantilla 7MEQ, lo que optimiza el acoplamiento molecular. Al alinearlos contra este cristal, el modelo por homología logró un RMSD de 0.285 Å, superando al de AlphaFold (1.367 Å). Por ende, el primero define

con mayor exactitud los rotómeros de la tríada catalítica. De hecho, este valor es más competitivo que el RMSD de 0.33 Å reportado previamente con la plantilla 1Z8G (Hussain et al., 2020). A nivel global, la alta similitud estructural coincide con el 85% de identidad entre los homólogos de ratón y humano (Hatesuer et al., 2013). No obstante, la presencia de la hélice alfa 3 (α3) en el bucle periférico contiguo al sitio activo es crítica para la unión del ligando. Así que, aunque AlphaFold (5 hélices) obtuvo mejor calidad global en estado libre (*apo*), el modelo por homología (4 hélices) mostró energías de *docking* mucho más favorables. Este comportamiento se explica por el ajuste inducido. Como 7MEQ se cristalizó con nafamostat, su bolsillo ya estaba pre-organizado en estado activo (*holo*). En cambio, AlphaFold predijo un bolsillo cerrado. Debido a esto, se propone un enfoque conformacional dual para la TMPRSS2 de *Mus musculus*: AlphaFold sirve como referente anatómico del estado libre, mientras que el modelo por homología es la opción funcional para el diseño de fármacos.

VI.2.3. Análisis e interpretación de los inhibidores exploratorios de alto masa molecular: Valrubicina y Lopinavir

Lopinavir es un inhibidor de proteasa aprobado por la FDA para el tratamiento de la infección por el virus de la inmunodeficiencia humana (VIH-1) (Choy et al., 2020). Por su parte, valrubicin es un análogo semisintético de la doxorubicina, un agente quimioterapéutico que interfiere con la síntesis de ADN y ARN, lo que podría afectar la replicación viral (Baby et al., 2021). Ambos fueron seleccionados porque estudios previos de reposicionamiento farmacológico, basado en acoplamiento y dinámica moleculares, los identificaron como candidatos prometedores de inhibir TMPRSS2 y ACE2 (Baby et al., 2021).

. Al evaluar los compuestos exploratorios mediante acoplamiento molecular, la Tabla 5 muestra que la valrubicina presentó una energía de unión de −4.02 kcal/mol,

superior a la obtenida para lopinavir (−2.62 kcal/mol) y ligeramente más favorable que la registrada para el inhibidor de referencia camostat (−3.19 kcal/mol).

La valrubicina estableció puentes de hidrógeno con Thr138 (1.95 Å y 2.34 Å) y Gln183 (2.32 Å), además de contactos hidrofóbicos con Val25 (3.07 Å), Cys26 (3.44 Å) y Val23 (3.33 Å). Asimismo, presentó interacciones de tipo puente salino con His41 a distancias de 5.01 Å y 5.18 Å. Aunque estas interacciones contribuyen a la estabilidad del complejo predicho, la ausencia de enlaces de hidrógeno directos con los residuos del núcleo catalítico y las mayores distancias respecto a His41 sugieren una interacción periférica, diferente a la observada para nafamostat y camostat.

De manera similar, el lopinavir permaneció localizado en la región externa del sitio activo, estableciendo dos puentes de hidrógeno con Asp140 (2.08 Å y 2.93 Å), además de una interacción con Gln183 (2.90 Å) y contactos hidrofóbicos con Val23 y Val25. Las distancias observadas respecto a Ser186 (4.88 Å) e His41 (5.62 Å) indican que el ligando permanece alejado del núcleo catalítico durante la conformación obtenida por docking, lo que refuerza la diferencia en su modo de unión respecto a los inhibidores de referencia. Estos resultados concuerdan con lo reportado previamente por Baby et al. (2021), tanto la valrubicina como el lopinavir permanecieron localizados en regiones periféricas del dominio serinoproteasa y no lograron entrar en el bolsillo de especificidad primaria S1. En nuestro modelo, ambos ligandos adoptaron una orientación superficial sobre los lazos externos de la entrada del sitio activo, en contraste con la ubicación observada para los inhibidores de referencia.

Con el objetivo de validar la especificidad del espacio de búsqueda y descartar interacciones espurias por sesgo del algoritmo matemático, se incorporó un acoplamiento control negativo utilizando Adenosine triphosphate. Este ligando de control registró una energía libre de unión menos favorable y más alta de 0.5 kcal/mol, situándose en el umbral inferior de todo el espectro evaluado. El análisis geométrico evidenció que el control negativo se ubicó de forma errática en la periferia externa del

dominio, careciendo por completo de la capacidad estérica para establecer puentes de hidrógeno, contactos hidrofóbicos estables o puentes salinos con la tríada catalítica (His41, Asp90 y Ser186) o con el bolsillo de especificidad S1.

La comparación entre los resultados del acoplamiento molecular y las simulaciones de dinámica molecular de 50 ns permitió complementar la interpretación de estas interacciones. Mientras que los complejos formados con nafamostat y camostat mostraron perfiles de RMSD y RMSF relativamente estables durante el periodo de simulación, los complejos con valrubicina y lopinavir presentaron mayores fluctuaciones estructurales, indicando una menor estabilidad conformacional inicial.

Ambos compuestos han sido estudiados *in silico* por su potencial para inhibir la entrada del SARS-CoV-2 en las células del huésped. No obstante, para validar estos hallazgos y evaluar su viabilidad como tratamientos antivirales, es necesario llevar a cabo estudios experimentales (Baby et al., 2021).

VI.3. Simulación de dinámica molecular (MDS) del dominio serinoproteasa de TMPRSS2 y 4 inhibidores.

Se realizaron simulaciones de dinámica molecular para determinar si la interacción entre el TMPRSS2 con inhibidores de referencia (nafamostat y camostat) y TMPRSS2 con inhibidores exploratorios (lopinavir y valrubicina) se mantienen estables en condiciones cercanas a la fisiológicas en 50 ns. Además, la estabilidad y flexibilidad de los complejos en un entorno solventado se evaluaron mediante RMSD y RMSF. La interacción es altamente estable si las fluctuaciones no superan los 1.5 Å (Baby et al., 2021). También, el análisis se complementó mediante la evaluación del radio de giro, número de contactos entre proteína y ligando, número de puentes de hidrógeno y las distancias intermoleculares entre los ligandos y los residuos de la triada catalítica.

Los análisis de RMSD mostraron que los complejos con los inhibidores de referencia alcanzaron una estabilidad basal de forma temprana. Por un lado, el complejo

con camostat mostró una convergencia gradual hacia los 50 ns, registrando un valor promedio de 0.245 ± 0.031 nm y con un importante número de contactos 96.18 ± 18.77. Además, se mantuvo constante el número de puentes de hidrógeno durante la trayectoria, lo que sugiere que camostat permanecería anclado en el sitio activo del dominio serinoproteasa de TMPRSS2.

Por otra parte, el complejo con nafamostat mantuvo un comportamiento inicial muy estable. Sus fluctuaciones fueron mínimas (0.126 nm ± 0.028 nm) durante la mayor parte de la trayectoria. De tal forma, esta tendencia inicial se corroboró con el elevado número de contactos basales y la persistencia de entre 4 y 5 puentes de hidrógeno. No obstante, el ligando nafamostat experimentó un incremento abrupto en su RMSD durante los últimos 10 ns de la simulación, alcanzando valores cercanos a los 0.5 nm. Debido a esto, se plantea que nafamostat sufrió un reacomodo conformacional dentro de la cavidad catalítica. Este comportamiento se relaciona con la reducción de los contactos intermoleculares a partir de los 40 ns. Conforme avanzó la trayectoria, estos contactos se redujeron de un promedio de 150 a aproximadamente 100 hacia el final de la trayectoria. A pesar de este reajuste geométrico, los puentes de hidrógeno se mantuvieron continuos en esta última ventana temporal. Además, las curvas de RMSD retornaron hacia valores menores al cierre de los 50 ns. Por lo tanto, estos indicadores sugieren que el ligando no sufrió una disociación completa, sino una optimización conformacional de su acoplamiento en el sitio activo.

En un estudio realizado por Sonawane et al. (2021), quienes analizaron computacionalmente el dominio serinoproteasa de TMPRSS2 en complejo con camostat y nafamostat mediante simulaciones de 500 ns, observaron valores de RMSD estables (0.15–0.30 nm; <1.5 Å) sin fluctuaciones significativas. A pesar de que la trayectoria fue más extensa y favorece la estabilización del sistema, los resultados coinciden en que ambos inhibidores mantienen interacciones estables con TMPRSS2. Adicionalmente, Kishk et al. (2020) reportaron que nafamostat presenta una estabilidad ligeramente

superior y alcanza el equilibrio más rápido que camostat en simulaciones de 200 ns, tendencia que también se refleja en la comparación dinámica observada en este estudio.

De la misma manera, al analizar la distancia entre el centro de masa global del dominio serinoproteasa de TMPRSS2-nafamostat y TMPRSS2-camostat se confirmó que los complejos no sufrieron eventos de disociación, manteniendo una permanencia del 100% durante los 50 ns de simulación. Al quedar retenidos muy cerca de la tríada catalítica, ambos inhibidores restringen la flexibilidad termodinámica de este sitio activo, mostrando una estabilización diferencial en el análisis del radio de giro. Con nafamostat se observó un incremento sutil del radio de giro (1.730 ± 0.013 nm), lo que sugiere una conformación ligeramente más extendida, mientras que camostat conservó exactamente el volumen de compactación del estado libre Apo (1.714 ± 0.006 nm). Esta fuerte sujeción en la cavidad activa explica por qué los valores de RMSF locales de la tríada catalítica disminuyeron respecto al Apo. Esta tendencia dinámica se alinea con lo reportado por Sonawane et al. (2021) para la variante humana. En dicho estudio, el complejo con nafamostat mostró el radio de giro más alto (2.280 ± 0.02 nm) en comparación con el estado libre (2.264 ± 0.02 nm), mientras que camostat mantuvo un valor casi idéntico al Apo (2.268 ± 0.02 nm). Aunque los valores absolutos de Sonawane son mayores debido al uso de la proteína completa (*full-length*), el comportamiento relativo es idéntico en nuestro modelo del dominio de *Mus musculus*. De este modo, ambos fármacos estabilizan el volumen de la proteasa sin inducir un colapso estructural, lo que sugiere un comportamiento dinámico conservado entre ambas especies

Al comparar la trayectoria apo frente a estas formas holo, se evidenció que la unión de estos inhibidores de referencia ejerce un efecto restrictivo sobre la triada catalítica; ya que el perfil de RMSF mostró una reducción de la flexibilidad en esa región, específicamente en el residuo Ser186. En el complejo TMPRSS2–nafamostat, el RMSF de Ser186 disminuyó de 0.1686 nm en la forma apo a 0.0594 nm, indicando una reducción notoria de la movilidad residual. De manera similar, camostat redujo las

fluctuaciones de His41 y Asp90 respecto a su forma apo. Por otro lado, aunque lopinavir y valrubicina también mostraron disminuciones parciales de RMSF en algunos residuos (Anexo 29 y Anexo 30), el análisis de distancias promedio entre el ligando y los residuos catalíticos evidenció que estos compuestos permanecieron alejados del sitio activo durante gran parte de la simulación. Lopinavir presentó distancias promedio entre 1.49 y 2.54 nm respecto a los residuos catalíticos, mientras que valrubicina mostró distancias aún mayores (~3.1 nm). Estos valores sugieren ausencia de interacción estable con la tríada catalítica. En contraste, nafamostat presentó las menores distancias promedio, especialmente con Ser186 (0.543 ± 0.056 nm), indicando proximidad sostenida al sitio catalítico. De igual forma, camostat también permaneció relativamente cercano a los residuos catalíticos, aunque con distancias mayores que nafamostat (Tabla 9).

A diferencia de los inhibidores de referencias, los inhibidores exploratorios (lopinavir y valrubicina) mostraron una marcada inestabilidad estructural temprana. Durante los 50 ns, sus trayectorias de RMSD presentaron fuertes fluctuaciones y reajustes constantes en la cavidad catalítica. Esto ocurrió debido a que el solvente explícito interrumpió progresivamente las interacciones periféricas débiles que se habían establecido en el docking estático. Esto se corrobora, ya que el complejo con lopinavir mostró un salto abrupto a mitad de la simulación, donde los contactos subieron de 30 a 120 y los puentes de hidrógeno de 2 a 4. Este comportamiento indica que el ligando sufrió un reajuste estructural para buscar una mejor posición en el sitio activo. Por otro lado, la valrubicina mantuvo valores de cero contactos y puentes de hidrógeno casi toda la simulación, lo que evidencia su disociación de la proteasa alrededor de los 5 ns.

Las distancias de centro de masa de los inhibidores exploratorios confirmaron una inestabilidad temprana que contradice la afinidad obtenida en el docking estático. En primer lugar, el lopinavir registró una distancia promedio elevada (2.20 ± 0.32 nm) debido a que se reacomodó a mitad de la trayectoria hacia la periferia externa. Al no

poder ingresar al bolsillo S1, la proteína conservó podría haber experimentado un empaquetamiento alostérico, manteniendo un radio de giro bajo (1.656 ± 0.008 nm). En segundo lugar, la valrubicina mostró un aumento drástico en su distancia intermolecular después de los 5 ns, alcanzando un máximo de 5.0 nm. Aunque el fármaco inició cerca del bolsillo S1, se alejó por completo conforme avanzó la simulación. Debido a esta separación prematura de valrubicina, el sitio activo experimentó una desestabilización inicial. Esto provocó que los bucles superficiales de la proteína se reordenaran abruptamente hacia el interior para compensar el vacío del bolsillo, disminuyendo el radio de giro promedio a 1.657 ± 0.006 nm (Tabla 8 y Tabla 9).

Al evaluar la dinámica global, se observó que el RMSD de los complejos holo TMPRSS2-nafamostat y TMPRSS2-camostat fue ligeramente superior al de su forma apo. Además, los valores de RMSF de ambos sistemas disminuyeron notablemente respecto a la proteína libre (Anexo 27-Anexo 30). Esta primera tendencia molecular se contrapone parcialmente a lo reportado por Sonawane et al. (2021), quienes describieron una sutil reducción en las curvas de RMSD de las formas holo en comparación con el estado apo, aunque sí coincidieron en registrar una caída en las fluctuaciones del RMSF tras el acoplamiento. De tal forma, las discrepancias estructurales encontradas con la presente tesis se explican por criterios metodológicos en el modelado por homología. En primer lugar, los autores construyeron un modelo bidominio que incluía el dominio serinoproteasa junto al dominio scavenger receptor cysteine-rich (SRCR); por lo tanto, debido a esto, emplearon una plantilla que apenas alcanzaba un 42% de identidad de secuencia, lo cual se tradujo en una menor predicción de estructuras secundarias. En contraste, en el presente estudio se utilizó una plantilla estructural con un 84,32% de similitud de secuencia, proporcionando una conformación inicial más robusta. Por ende, el incremento en el RMSD de los complejos holo en la presente tesis se relaciona con los ajustes conformacionales locales necesarios para

optimizar las interacciones en el sitio activo, conservando intacta la estabilidad global de la enzima.

Al comparar nuestros resultados con simulaciones de 200 ns (Kishk et al., 2020; Ramakrishnan et al., 2020), el camostat y nafamostat de esta tesis convergieron a valores de RMSD más bajos. En dichos estudios el RMSD se estabilizó cerca de los 0.4 nm, mientras que aquí alcanzó una meseta constante alrededor de los 0.2 nm durante los 50 ns. Específicamente, el nafamostat mostró la menor fluctuación, sugiriendo un anclaje óptimo que coincide con los reportes previos sobre sus fuertes interacciones en el sitio activo de TMPRSS2. Esta variación numérica no implica falta de coherencia biológica, ya que ambos casos lograron convergencia estructural. La diferencia se explica por tres factores metodológicos: la elección del PDB 7MEQ como estructura inicial, la menor duración de nuestra simulación (donde 50 ns representa una estabilización temprana en solvente explícito) y las fases previas de minimización. De tal forma, a pesar de las escalas distintas, ambos trabajos validan mutuamente la estabilidad del complejo.

Estos resultados demuestran que, dentro del alcance y los límites temporales evaluados en este estudio, lopinavir y valrubicina no presentan un comportamiento estructural comparable con camostat y nafamostat como inhibidores inmediatos de TMPRSS2, ya que carecen de la complementariedad necesaria para sostener una unión rígida en condiciones fisiológicas simuladas durante 50 ns. Mientras que los fármacos de referencia poseen un mecanismo de acoplamiento rápido, directo y protector del sitio activo, los compuestos exploratorios muestran una naturaleza altamente dinámica, requiriendo de transiciones conformacionales de la proteína a mayor escala temporal para estabilizar eficazmente sus modos de unión.

VI.3.1. Análisis dinámico y perturbación geométrica de la tríada catalítica de TMPRSS2

Las distancias interresiduales iniciales medidas entre los carbonos alfa ($C\alpha$) de la tríada catalítica (His41, Asp90 y Ser186) oscilaron entre 6.4 Å y 10.0 Å (His41-Asp90: 6.4 Å; His41-Ser186: 8.0 Å; Asp90-Ser186: 10.0 Å) (Figura 19). Se observó una alta conservación geométrica y espacial al comparar el modelo por homología de SWISS-MODEL y la predicción de AlphaFold con la estructura cristalográfica de referencia 7MEQ. Este posicionamiento estático inicial coincide con lo reportado por Hussain (2020), cuyas distancias interresiduales en la tríada de TMPRSS2 fluctuaron entre 6.49 Å y 9.75 Å, un rango espacialmente equivalente a serinoproteasas homólogas como la hepsina (PDB: 1Z8G) y la calicreína plasmática (PDB: 5TJX). Esto confirma que, antes de la simulación, ambos modelos de partida presentan una estructura catalítica correcta.

Las trayectorias de 50 ns mostraron cambios estructurales clave. En estado libre (*Apo_Swiss*), el eje His41–Ser186 se abrió notablemente (0.8968 ± 0.1323 nm), coincidiendo con la flexibilidad reportada por Sonawane (2021) de 10.33 Å en el apo humano. Por el contrario, nafamostat se contrajó a 0.736 ± 0.0359 nm, alineándose con la compactación descrita por dicho autor para nafamostat (5.44 Å) y camostat (5.50 Å). Sin embargo, en la distancia entre His41–Asp90, el modelo de *Mus musculus* se mantuvo rígida (0.5599 nm para camostat y 0.568 nm para nafamostat), a diferencia del incremento reportado en humanos, sugiriendo un anclaje más estricto en el ratón (Tabla 10 y Tabla 11).

Por otro lado, los ligandos sin afinidad directa distorsionaron el sitio activo. Lopinavir, al unirse en la periferia, desestabilizó la tríada a los 40 ns, elevando el eje His41–Asp90 hasta 0.72 nm (0.5930 ± 0.05111 nm). Esta perturbación es similar a la descrita para la Bromhexina (8.02 Å) por Sonawane (2021), evidenciando un efecto alostérico. Asimismo, la valrubicina fluctuó de forma errática e idéntica al estado libre (0.8846 ± 0.0544 nm) al no lograr acoplarse a la proteína (Tabla 10 y Tabla 11).

Por lo tanto, tanto la contracción inducida por nafamostat y camostat como la distorsión alostérica asociada a lopinavir sugieren una alteración en la geometría basal de la tríada. A nivel computacional, esta desorganización espacial podría comprometer la funcionalidad catalítica de TMPRSS2, y su función proteolítica.

Es importante destacar que los valores de energía de unión obtenidos en el acoplamiento molecular fueron bajos, lo que sugiere una adecuada selección de los modelos de homología y AlphaFold en comparación con los enfoques de *threading* y *ab initio.* Además, la interacción intermolecular de los cuatro complejos fue evaluada mediante dinámica molecular, lo que permitió analizar su estabilidad a lo largo del tiempo. Sin embargo, una de las principales limitaciones de este estudio fue el tiempo de simulación empleado en la dinámica molecular. Aunque los complejos nafamostat y camostat demostraron estabilidad estructural dentro de los 50 ns de simulación, lopinavir y valrubicin no mostraron fluctuaciones estables durante ese periodo de tiempo, lo que sugiere que tiempos de simulación más prolongados podrían ser necesarios para evaluar con mayor precisión su comportamiento estructural.

Esta investigación haciendo uso de recursos computacionales sienta las bases de la interacción intermolecular entre inhibidores de proteasas y el dominio serinoproteasa de TMPRSS2 de ratón corroborando su papel como modelo experimental para futuras investigaciones en desarrollo de fármacos y tratamientos contra influenza y coronavirus. Ante la falta de modelos animales apropiados se espera que se proponga el estudio computacional en modelos como rata, cobaya, conejo, perro, hurón, macaco y hámster (Brooke y Prischi, 2020).

VII. CONCLUSIONES

Se generó y evaluó el modelo estructural del dominio serinoproteasa de TMPRSS2 en ratón mediante Swiss-Model y AlphaFold, los cuales cumplieron con

criterios de calidad y estabilidad estructural, presentando 12 láminas β y entre 4 y 5 hélices α.

El acoplamiento molecular indica que nafamostat y camostat muestran interacciones favorables específicos en la triada catalítica del dominio serinoproteasa de TMPRSS2 con energías de unión favorables (-5.52 kcal/mol y -3.19 kcal/mol) y estabilidad relativa en las condiciones de dinámica molecular de 50 ns.

El acoplamiento molecular indica que lopinavir y valrubicina se localizan en las regiones periféricas del dominio serinopoteasa de TMPRSS2 con energía de unión predichas de -4,02 kcal/mol y -2 kcal/mol. Adicionalmente, las simulaciones de dinámica molecular confirmaron la inestabilidad con fluctuaciones considerables y falta de convergencia durante los 50 ns.

Se destaca la necesidad de una validación experimental complementaria mediante ensayos enzimáticos, celulares o biofísicos para corroborar la eficacia inhibitoria y el comportamiento dinámico predictivo de los fármacos evaluados frente a TMPRSS2.

VIII. RECOMENDACIONES

Se recomienda la validación experimental mediante estudios como ensayos de inhibición enzimática o celulares y comparación con TMPRSS2 humana, y así se pueda corroborar los hallazgos obtenidos en esta tesis, determinando su relevancia en un contexto biológico real.

Además, se propone la evaluación de estos inhibidores en dinámica molecular con membranas en un periodo de tiempo mayor con réplicas independientes, debido a que TMPRSS2 es una proteína transmembrana y el entorno lipídico podría influir en la formación de interacciones relevantes para el diseño de fármacos dirigidos a este tipo de proteasas.

IX. REFERENCIAS BIBLIOGRÁFICAS

Adelusi, T. I., Oyedele, A.-Q. K., Boyenle, I. D., Ogunlana, A. T., Adeyemi, R. O., Ukachi, C. D., Idris, M. O., Olaoba, O. T., Adedotun, I. O., Kolawole, O. E., Xiaoxing, Y., & Abdul-Hammed, M. (2022). Molecular modeling in drug discovery. *Informatics in Medicine Unlocked*, *29*, 100880. https://doi.org/10.1016/j.imu.2022.100880

Ahuja, S. (2024). Molecular docking and molecular dynamics simulation study of interactions of TMPRSS2 inhibitors camostat, nafamostat, and 4-guanidinobenzoic acid with TMPRSS2. *European Biophysics Journal, 53*, 27–46. https://doi.org/10.1007/s00249-023-01695-4

Alimohamadi, Y., Sepandi, M., Taghdir, M., & Hosamirudsari, H. (2020). Determine the most common clinical symptoms in COVID-19 patients: A systematic review and meta-analysis. *Journal of Preventive Medicine and Hygiene*, *61*(3), Article 3. https://doi.org/10.15167/2421-4248/jpmh2020.61.3.1530

Antalis, T. M., Bugge, T. H., & Wu, Q. (2011). Membrane-Anchored Serine Proteases in Health and Disease. *Progress in Molecular Biology and Translational Science*, *99*, 1. https://doi.org/10.1016/B978-0-12-385504-6.00001-4

Arcón, J. P. (2018). Determinantes moleculares de la interacción droga-proteína: uso de cosolventes como moléculas de prueba (Tesis doctoral, Universidad de Buenos Aires).

Arreola-Barroso, R., Quintero-Hernández, V., Muñoz-Rojas, J., Gaytán, P., Rivera-Urbalejo, P., Rosete-Enríquez, M., Martínez-Martínez, L., & Juárez-González, V. R. (2022). SWISS-MODEL es un generador de modelos estructurales de proteínas cuyas estructuras aún no están depositadas en el PDB. *Alianzas y Tendencias BUAP*, *7*(27), 117-130. https://doi.org/10.5281/zenodo.7106150

Baby, K., Maity, S., Mehta, C. H., Suresh, A., Nayak, U. Y., & Nayak, Y. (2021). SARS-CoV-2 entry inhibitors by dual targeting TMPRSS2 and ACE2: An *in silico* drug

repurposing study. *European Journal of Pharmacology*, *896*, 173922. https://doi.org/10.1016/j.ejphar.2021.173922

Böttcher, E., Matrosovich, T., Beyerle, M., Klenk, H.-D., Garten, W., & Matrosovich, M. (2006). Proteolytic Activation of Influenza Viruses by Serine Proteases TMPRSS2 and HAT from Human Airway Epithelium. *Journal of Virology*, *80*(19), 9896. https://doi.org/10.1128/JVI.01118-06

Böttcher-Friebertshäuser, E., Freuer, C., Sielaff, F., Schmidt, S., Eickmann, M., Uhlendorff, J., Steinmetzer, T., Klenk, H.-D., & Garten, W. (2010). Cleavage of Influenza Virus Hemagglutinin by Airway Proteases TMPRSS2 and HAT Differs in Subcellular Localization and Susceptibility to Protease Inhibitors. *Journal of Virology*, *84*(11), 5605-5614. https://doi.org/10.1128/JVI.00140-10

Bonifasius, S., Santoso, F., Lee, J.-H., Yu, W.-H., Wu, C.-C., Audira, G., Stevhen Juniardi, Chen, J.-R., Lin, Y.-T., & Hsiao, C.-D. (2019). Cardiac Rhythm and Molecular Docking Studies of Ion Channel Ligands with Cardiotoxicity in Zebrafish. *Cells*, *8*(6), 566–566. https://doi.org/10.3390/cells8060566

Brooke, G. N., & Prischi, F. (2020). Structural and functional modelling of SARS-CoV-2 entry in animal models. *Scientific Reports*, *10*, 15917. https://doi.org/10.1038/s41598-020-72528-z

Buchan, D. W. A., & Jones, D. T. (2019). The PSIPRED Protein Analysis Workbench: 20 years on. *Nucleic Acids Research*, *47*(W1), W402-W407. https://doi.org/10.1093/nar/gkz297

Chaipan, C., Kobasa, D., Bertram, S., Glowacka, I., Steffen, I., Solomon Tsegaye, T., Takeda, M., Bugge, T. H., Kim, S., Park, Y., Marzi, A., & Pöhlmann, S. (2009). Proteolytic Activation of the 1918 Influenza Virus Hemagglutinin. *Journal of Virology*, *83*(7), 3200-3211. https://doi.org/10.1128/JVI.02205-08

Chivian, D., Kim, D. E., Malmström, L., Bradley, P., Robertson, T., Murphy, P., Strauss, C. E. M., Bonneau, R., Rohl, C. A., & Baker, D. (2003). Automated prediction of

CASP-5 structures using the Robetta server. *Proteins: Structure, Function, and Bioinformatics*, *53*(S6), 524-533. https://doi.org/10.1002/prot.10529

Choy, K.-T., Wong, A. Y.-L., Kaewpreedee, P., Sia, S. F., Chen, D., Hui, K. P. Y., Chu, D. K. W., Chan, M. C. W., Cheung, P. P.-H., Huang, X., Peiris, M., & Yen, H.-L. (2020). Remdesivir, lopinavir, emetine, and homoharringtonine inhibit SARS-CoV-2 replication in vitro. *Antiviral Research*, *178*, 104786. https://doi.org/10.1016/j.antiviral.2020.104786

Daodu, L. P., Yusuf, O. O., & Okouzi, M. (2026). Distinción entre la pandemia de COVID-19 y las pandemias de influenza (A/H1N1, A/H2N2, A/H3N2): una revisión narrativa. *Bull Natl Res Cent*, *50*, Artículo 18. https://doi.org/10.1186/s42269-026-01409-8

Dinata, R., Baindara, P., & Mandal, S. M. (2025). Evolución de la resistencia a los fármacos antivirales en el SARS-CoV-2. *Viruses*, *17*(5), Artículo 722. https://doi.org/10.3390/v17050722

Dodda, L. S., Cabeza de Vaca, I., Tirado-Rives, J., & Jorgensen, W. L. (2017). LigParGen web server: An automatic OPLS-AA parameter generator for organic ligands. *Nucleic Acids Research*, *45*(Web Server issue), W331-W336. https://doi.org/10.1093/nar/gkx312

Ekstrand, K., Flanagan, A. J., Lin, I. E., Vejseli, B., Cole, A., Lally, A. P., Morris, R. L., & Morgan, K. N. (2021). Animal Transmission of SARS-CoV-2 and the Welfare of Animals during the COVID-19 Pandemic. *Animals : an Open Access Journal from MDPI*, *11*(7), 2044. https://doi.org/10.3390/ani11072044

Escalante, D. E., & Ferguson, D. M. (2021). Structural modeling and analysis of the SARS-CoV-2 cell entry inhibitor camostat bound to the trypsin-like protease TMPRSS2. *Medicinal Chemistry Research*, *30*(2), 399-409. https://doi.org/10.1007/s00044-021-02708-7

Färber, I., Krüger, J., Rocha, C., Armando, F., von Köckritz-Blickwede, M., Pöhlmann, S., Braun, A., Baumgärtner, W., Runft, S., & Krüger, N. (2022). Investigations on SARS-CoV-2 Susceptibility of Domestic and Wild Animals Using Primary Cell Culture Models Derived from the Upper and Lower Respiratory Tract. *Viruses*, *14*(4), Article 4. https://doi.org/10.3390/v14040828

Fraser, B. J., Beldar, S., Seitova, A., Hutchinson, A., Mannar, D., Li, Y., Kwon, D., Tan, R., Wilson, R. P., Leopold, K., Subramaniam, S., Halabelian, L., Arrowsmith, C. H., & Bénard, F. (2021). *Structure, activity and inhibition of human TMPRSS2, a protease implicated in SARS-CoV-2 activation* (p. 2021.06.23.449282). bioRxiv. https://doi.org/10.1101/2021.06.23.449282

Fraser, B. J., Beldar, S., Seitova, A., Hutchinson, A., Mannar, D., Li, Y., Kwon, D., Tan, R., Wilson, R. P., Leopold, K., Subramaniam, S., Halabelian, L., Arrowsmith, C. H., & Bénard, F. (2022). Structure and activity of human TMPRSS2 protease implicated in SARS-CoV-2 activation. *Nature Chemical Biology*, *18*(9), Article 9. https://doi.org/10.1038/s41589-022-01059-7

Frumenzio, G., Chandramouli, B., Besker, N., Grottesi, A., Talarico, C., Frigerio, F., Emerson, A., & Musiani, F. (2023). Conformational response to ligand binding of TMPRSS2, a protease involved in SARS-CoV-2 infection: Insights through computational modeling. *Proteins: Structure, Function, and Bioinformatics*, *91*(9), 1288-1297. https://doi.org/10.1002/prot.26548

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). Protein Identification and Analysis Tools on the ExPASy Server. En J. M. Walker (Ed.), *The Proteomics Protocols Handbook* (pp. 571-607). Humana Press. https://doi.org/10.1385/1-59259-890-0:571

Hatesuer, B., Bertram, S., Mehnert, N., Bahgat, M. M., Nelson, P. S., Pöhlmann, S., & Schughart, K. (2013). Tmprss2 is essential for influenza H1N1 virus pathogenesis

in mice. *PLoS Pathogens*, *9*(12), e1003774. https://doi.org/10.1371/journal.ppat.1003774

Hernández-Mitre, M. P., Tong, S. Y. C., Denholm, J. T., Dore, G. J., Bowen, A. C., Lewin, S. R., Venkatesh, B., Hills, T. E., McQuilten, Z., Paterson, D. L., Morpeth, S. C., & Roberts, J. A. (2022). Nafamostat Mesylate for Treatment of COVID-19 in Hospitalised Patients: A Structured, Narrative Review. *Clinical Pharmacokinetics*, *61*(10), 1331-1343. https://doi.org/10.1007/s40262-022-01170-x

Hollingsworth, S. A., & Dror, R. O. (2018). Molecular dynamics simulation for all. *Neuron*, *99*(6), 1129-1143. https://doi.org/10.1016/j.neuron.2018.08.011

Huey, R., Morris, G. M., Olson, A. J., & Goodsell, D. S. (2007). A semiempirical free energy force field with charge-based desolvation. *Journal of Computational Chemistry, 28*(6), 1145–1152. https://doi.org/10.1002/jcc.20634

Hussain, M., Jabeen, N., Amanullah, A., Baig, A. A., Aziz, B., Shabbir, S., Raza, F., & Uddin, N. (2020). Molecular docking between human TMPRSS2 and SARS-CoV-2 spike protein: Conformation and intermolecular interactions. *AIMS Microbiology*, *6*(3), 350–360. https://doi.org/10.3934/microbiol.2020021

Iwata-Yoshikawa, N., Okamura, T., Shimizu, Y., Hasegawa, H., Takeda, M., & Nagata, N. (2019). TMPRSS2 Contributes to Virus Spread and Immunopathology in the Airways of Murine Models after Coronavirus Infection. *Journal of Virology*, *93*(6), e01815. https://doi.org/10.1128/JVI.01815-18

Johns Hopkins University of Medicine. (2023). COVID-19 map - Johns Hopkins Coronavirus Resource Center. Recuperado el 10 de marzo de 2023, de https://coronavirus.jhu.edu/map.html

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer,

C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., … Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, *596*(7873), 583-589. https://doi.org/10.1038/s41586-021-03819-2

Kaminski, G. A., Friesner, R. A., Tirado-Rives, J., & Jorgensen, W. L. (2001). Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides. *The Journal of Physical Chemistry B*, *105*(28), 6474-6487. https://doi.org/10.1021/jp003919d

Karplus, M., & McCammon, J. A. (2002). Molecular dynamics simulations of biomolecules. *Nature Structural Biology*, *9*(9), 646-652. https://doi.org/10.1038/nsb0902-646

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., & Sternberg, M. J. E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols*, *10*(6), 845-858. https://doi.org/10.1038/nprot.2015.053

Kihara, D (2014). Protein structure prediction (3rd ed.). Humana Press

Kim, D. E., Chivian, D., & Baker, D. (2004). Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Research*, *32*(Web Server issue), W526-W531. https://doi.org/10.1093/nar/gkh468

Kim, T. S., Heinlein, C., Hackman, R. C., & Nelson, P. S. (2006). Phenotypic Analysis of Mice Lacking the Tmprss2-Encoded Protease. *Molecular and Cellular Biology*, *26*(3), 965-975. https://doi.org/10.1128/MCB.26.3.965-975.2006

Kishk, S. M., Kishk, R. M., Yassen, A. S. A., Nafie, M. S., Nemr, N. A., ElMasry, G., Al-Rejaie, S., & Simons, C. (2020). Molecular Insights into Human Transmembrane Protease Serine-2 (TMPS2) Inhibitors against SARS-CoV2: Homology Modelling, Molecular Dynamics, and Docking Studies. *Molecules*, *25*(21), 5007. https://doi.org/10.3390/molecules25215007

Lam, D. K., Dang, D., Flynn, A. N., Hardt, M., & Schmidt, B. L. (2015). TMPRSS2, a novel membrane-anchored mediator in cancer pain. *PAIN*, *156*(5), 923. https://doi.org/10.1097/j.pain.0000000000000130

Laporte, M., & Naesens, L. (2017). Airway proteases: An emerging drug target for influenza and other respiratory virus infections. *Current Opinion in Virology*, *24*, 16-24. https://doi.org/10.1016/j.coviro.2017.03.018

Laskowski, R. A., Jabłońska, J., Pravda, L., Vařeková, R. S., & Thornton, J. M. (2018). PDBsum: Structural summaries of PDB entries. *Protein Science: A Publication of the Protein Society*, *27*(1), 129-134. https://doi.org/10.1002/pro.3289

Lee, J., Lee, J., Kim, H. J., Ko, M., Jee, Y., & Kim, S. (2021). TMPRSS2 and RNA-Dependent RNA Polymerase Are Effective Targets of Therapeutic Intervention for Treatment of COVID-19 Caused by SARS-CoV-2 Variants (B.1.1.7 and B.1.351). *Microbiology Spectrum*, *9*(1), 10.1128/spectrum.00472-21. https://doi.org/10.1128/spectrum.00472-21

Lee, M. G., Kim, K. H., Park, K. Y., & Kim, J. S. (1996). Evaluation of anti-influenza effects of camostat in mice infected with non-adapted human influenza viruses. *Archives of Virology*, *141*(10), 1979-1989. https://doi.org/10.1007/BF01718208

Lemkul, J. A. (2018). From Proteins to Perturbed Hamiltonians: A Suite of Tutorials for the GROMACS-2018 Molecular Simulation Package [Article v1.0]. *Living Journal of Computational Molecular Science*, *1*(1), Article 1. https://doi.org/10.33011/livecoms.1.1.5068

Lesk, A (2014).Introduction to Bioinformatics (4ª ed). Oxford University Press.

Li, K., Meyerholz, D. K., Bartlett, J. A., & Paul B. McCray, J. (2021). The TMPRSS2 Inhibitor Nafamostat Reduces SARS-CoV-2 Pulmonary Infection in Mouse Models of COVID-19. *mBio*, *12*(4). https://doi.org/10.1128/mBio.00970-21

Lu, S., Wang, J., Chitsaz, F., Derbyshire, M. K., Geer, R. C., Gonzales, N. R., Gwadz, M., Hurwitz, D. I., Marchler, G. H., Song, J. S., Thanki, N., Yamashita, R. A.,

Yang, M., Zhang, D., Zheng, C., Lanczycki, C. J., & Marchler-Bauer, A. (2020). CDD/SPARCLE: The conserved domain database in 2020. *Nucleic Acids Research*, *48*(D1), D265-D268. https://doi.org/10.1093/nar/gkz991

Mahmudpour, M., Nabipour, I., Keshavarz, M., & Farrokhnia, M. (2021). Virtual Screening on Marine Natural Products for Discovering TMPRSS2 Inhibitors. *Frontiers in Chemistry*, *9*. https://doi.org/10.3389/fchem.2021.722633

Mantzourani, C., Vasilakaki, S., Gerogianni, V.-E., & Kokotos, G. (2022). The discovery and development of transmembrane serine protease 2 (TMPRSS2) inhibitors as candidate drugs for the treatment of COVID-19. *Expert Opinion on Drug Discovery*, 1. https://doi.org/10.1080/17460441.2022.2029843

Matsuyama, S., Nao, N., Shirato, K., Kawase, M., Saito, S., Takayama, I., Nagata, N., Sekizuka, T., Katoh, H., Kato, F., Sakata, M., Tahara, M., Kutsuna, S., Ohmagari, N., Kuroda, M., Suzuki, T., Kageyama, T., & Takeda, M. (2020). Enhanced isolation of SARS-CoV-2 by TMPRSS2-expressing cells. *Proceedings of the National Academy of Sciences*, *117*(13), 7001-7003. https://doi.org/10.1073/pnas.2002589117

McGuffin, L. J., Aldowsari, F. M. F., Alharbi, S. M. A., & Adiyaman, R. (2021). ModFOLD8: Accurate global and local quality estimates for 3D protein models. *Nucleic Acids Research*, *49*(W1), W425-W430. https://doi.org/10.1093/nar/gkab321

Menéndez, A. B. (2005). Aspectos Clínicos y Virológicos útiles en la Implementación de un Modelo de Influenza A en líneas de ratones Balb/C. *acta farmacéutica bonaerense*, *24*.

Meng, X.-Y., Zhang, H.-X., Mezei, M., & Cui, M. (2011). Molecular Docking: A powerful approach for structure-based drug discovery. *Current computer-aided drug design*, *7*(2), 146-157.

Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., & Steinegger, M. (2022). ColabFold: Making protein folding accessible to all. *Nature Methods*, *19*(6), 679-682. https://doi.org/10.1038/s41592-022-01488-1

Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K., & Olson, A. J. (1998). Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *Journal of Computational Chemistry*, *19*(14), 1639-1662. https://doi.org/10.1002/(SICI)1096-987X(19981115)19:14<1639::AID-JCC10>3.0.CO;2-B

Morris, G. M., & Lim-Wilby, M. (2008). Molecular docking. *Methods in Molecular Biology (Clifton, N.J.)*, *443*, 365-382. https://doi.org/10.1007/978-1-59745-177-2_19

Neary, M., Sharp, J., Gallardo-Toledo, E., Herriott, J., Kijak, E., Bramwell, C., Cox, H., Tatham, L., Box, H., Curley, P., Arshad, U., Rajoli, R. K. R., Pertinez, H., Valentijn, A., Dhaliwal, K., Mc Caughan, F., Hobson, J., Rannard, S., Kipar, A., … Owen, A. (2023). Evaluation of Nafamostat as Chemoprophylaxis for SARS-CoV-2 Infection in Hamsters. *Viruses*, *15*(8), 1744. https://doi.org/10.3390/v15081744

Netzel-Arnett, S., Hooper, J. D., Szabo, R., Madison, E. L., Quigley, J. P., Bugge, T. H., & Antalis, T. M. (2003). Membrane anchored serine proteases: A rapidly expanding group of cell surface proteolytic enzymes with potential roles in cancer. *Cancer Metastasis Reviews*, *22*(2-3), 237-258. https://doi.org/10.1023/a:1023003616848

Pandey-Rai, S., Rai, K. K., Pandey, N., Tripathi, D., Apoorva, Singh, V. K., & Rai, S. K. (2022). In-vivo Studies and Molecular Docking of Modeled Mus musculas 8S Lipoxygenase Protein Using Some Natural Bioactive Compounds. *Proceedings of the National Academy of Sciences, India Section B: Biological Sciences*, *92*(2), 359–370. https://doi.org/10.1007/s40011-021-01284-5

Paoloni-Giacobino, A., Chen, H., Peitsch, M. C., Rossier, C., & Antonarakis, S. E. (1997). Cloning of the TMPRSS2 Gene, Which Encodes a Novel Serine Protease with Transmembrane, LDLRA, and SRCR Domains and Maps to 21q22.3. *Genomics*, *44*(3), 309-320. https://doi.org/10.1006/geno.1997.4845

Paul, S., Bravo Vázquez, L. A., Reyes-Pérez, P. R., Estrada-Meza, C., Aponte Alburquerque, R. A., Pathak, S., Banerjee, A., Bandyopadhyay, A., Chakraborty, S., & Srivastava, A. (2022). The role of microRNAs in solving COVID-19 puzzle from infection to therapeutics: A mini-review. *Virus Research*, *308*, 198631. https://doi.org/10.1016/j.virusres.2021.198631

Ramakrishnan, J., Kandasamy, S., Iruthayaraj, A., Magudeeswaran, S., Chinnasamy, K., & Poomani, K. (2021). Strong Binding of Leupeptin with TMPRSS2 Protease May Be an Alternative to Camostat and Nafamostat for SARS-CoV-2 Repurposed Drug: Evaluation from Molecular Docking and Molecular Dynamics Simulations. *Applied Biochemistry and Biotechnology*, *193*(6), 1909-1923. https://doi.org/10.1007/s12010-020-03475-8

Raman, S., Vernon, R., Thompson, J., Tyka, M., Sadreyev, R., Pei, J., Kim, D., Kellogg, E., DiMaio, F., Lange, O., Kinch, L., Sheffler, W., Kim, B.-H., Das, R., Grishin, N. V., & Baker, D. (2009). Structure prediction for CASP8 with all-atom refinement using Rosetta. *Proteins: Structure, Function, and Bioinformatics*, *77*(S9), 89-99. https://doi.org/10.1002/prot.22540

Ravi, L., & K, K. (2016). A Handbook on Protein-Ligand Docking Tool: AutoDock 4. *Innovare Journal of Medical Sciences*, 28-33.

Rensi, S., Altman, R. B., Liu, T., Lo, Y.-C., McInnes, G., Derry, A., & Keys, A. (2020). *Homology Modeling of TMPRSS2 Yields Candidate Drugs That May Inhibit Entry of SARS-CoV-2 into Human Cells*. ChemRxiv. https://doi.org/10.26434/chemrxiv.12009582.v1

Robert, X., & Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research*, *42*(W1), W320-W324. https://doi.org/10.1093/nar/gku316

Rocha, S. M., Fagre, A. C., Latham, A. S., Cummings, J. E., Aboellail, T. A., Reigan, P., Aldaz, D. A., McDermott, C. P., Popichak, K. A., Kading, R. C., Schountz, T., Theise, N. D., Slayden, R. A., & Tjalkens, R. B. (2022). A Novel Glucocorticoid and Androgen Receptor Modulator Reduces Viral Entry and Innate Immune Inflammatory Responses in the Syrian Hamster Model of SARS-CoV-2 Infection. *Frontiers in Immunology*, *13*, 811430. https://doi.org/10.3389/fimmu.2022.811430

Rost, B., Sander, C., & Schneider, R. (1994). PHD--an automatic mail server for protein secondary structure prediction. *Computer Applications in the Biosciences: CABIOS*, *10*(1), 53-60. https://doi.org/10.1093/bioinformatics/10.1.53

Sakai, K., Ami, Y., Tahara, M., Kubota, T., Anraku, M., Abe, M., Nakajima, N., Sekizuka, T., Shirato, K., Suzaki, Y., Ainai, A., Nakatsu, Y., Kanou, K., Nakamura, K., Suzuki, T., Komase, K., Nobusawa, E., Maenaka, K., Kuroda, M., … Takeda, M. (2014). The host protease TMPRSS2 plays a major role in *in vivo* replication of emerging H7N9 and seasonal influenza viruses. *Journal of Virology*, *88*(10), 5608-5616. https://doi.org/10.1128/JVI.03677-13

Salleh, M. Z., & Deris, Z. Z. (2022). In silico molecular characterization of human TMPRSS2 protease polymorphic variants and associated SARS-CoV-2 susceptibility. *Life, 12*(2), 231. https://doi.org/10.3390/life12020231

Schake, P., Bolz, S. N., Linnemann, K., & Schroeder, M. (2025). PLIP 2025: introducing protein–protein interactions to the protein–ligand interaction profiler. *Nucleic Acids Research*, *53*(W1), W463–W465. https://doi.org/10.1093/nar/gkaf361

Sepúlveda Durán, L. A. (2007). *Estudio de dinámica molecular con solvente implícito de la influencia de la interfase de interacción entre los dominios de la proteína FtsZ*

en la estabilidad y plegamiento (Tesis de licenciatura). Universidad de Chile, Facultad de Ciencias Químicas y Farmacéuticas, Departamento de Bioquímica y Biología Molecular.

Shen, L. W., Mao, H. J., Wu, Y. L., Tanaka, Y., & Zhang, W. (2017). TMPRSS2: A potential target for treatment of influenza virus and coronavirus infections. *Biochimie*, *142*, 1-10. https://doi.org/10.1016/j.biochi.2017.07.016

Shrimp, J. H., Kales, S. C., Sanderson, P. E., Simeonov, A., Shen, M., & Hall, M. D. (2020). An Enzymatic TMPRSS2 Assay for Assessment of Clinical Candidates and Discovery of Inhibitors as Potential Treatment of COVID-19. *ACS Pharmacology & Translational Science*, *3*(5), 997-1007. https://doi.org/10.1021/acsptsci.0c00106

Sigrist, C. J. A., de Castro, E., Cerutti, L., Cuche, B. A., Hulo, N., Bridge, A., Bougueleret, L., & Xenarios, I. (2013). New and continuing developments at PROSITE. *Nucleic Acids Research*, *41*(Database issue), D344-347. https://doi.org/10.1093/nar/gks1067

Solovyev, V. (2007). Statistical Approaches in Eukaryotic Gene Prediction. En D. J. Balding, M. Bishop, & C. Cannings (Eds.), *Handbook of Statistical Genetics* (1.ª ed., pp. 97-159). Wiley. https://doi.org/10.1002/9780470061619.ch4

Sonawane, K. D., Barale, S. S., Dhanavade, M. J., Waghmare, S. R., Nadaf, N. H., Kamble, S. A., Mohammed, A. A., Makandar, A. M., Fandilolu, P. M., Dound, A. S., Naik, N. M., & More, V. B. (2021). Structural insights and inhibition mechanism of TMPRSS2 by experimentally known inhibitors Camostat mesylate, Nafamostat and Bromhexine hydrochloride to control SARS-coronavirus-2: A molecular modeling approach. *Informatics in Medicine Unlocked*, *24*, 100597. https://doi.org/10.1016/j.imu.2021.100597

Subaraja, M., Kulandaisamy, A., Shanmugam, N. R. S., & Vanisree, A. J. (2019). Homology modeling identified for purported drug targets to the neuroprotective

effects of levodopa and asiaticoside-D in degenerated cerebral ganglions of Lumbricus terrestris. *Indian Journal of Pharmacology*, *51*(1), 31-39. https://doi.org/10.4103/ijp.IJP_600_18

Sun, J., Zhuang, Z., Zheng, J., Li, K., Wong, R. L.-Y., Liu, D., Huang, J., He, J., Zhu, A., Zhao, J., Li, X., Xi, Y., Chen, R., Alshukairi, A. N., Chen, Z., Zhang, Z., Chen, C., Huang, X., Li, F., … Zhao, J. (2020). Generation of a Broadly Useful Model for COVID-19 Pathogenesis, Vaccination, and Treatment. *Cell*, *182*(3), 734-743.e5. https://doi.org/10.1016/j.cell.2020.06.010

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N. T., Morris, J. H., Bork, P., Jensen, L. J., & Mering, C. von. (2019). STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Research*, *47*(D1), D607-D613. https://doi.org/10.1093/nar/gky1131

Tarnow, C., Engels, G., Arendt, A., Schwalm, F., Sediri, H., Preuss, A., Nelson, P. S., Garten, W., Klenk, H.-D., Gabriel, G., & Böttcher-Friebertshäuser, E. (2014). TMPRSS2 Is a Host Factor That Is Essential for Pneumotropism and Pathogenicity of H7N9 Influenza A Virus in Mice. *Journal of Virology*, *88*(9), 4744-4751. https://doi.org/10.1128/jvi.03799-13

Teilum, K., Olsen, J. G., & Kragelund, B. B. (2011). Protein stability, flexibility and function. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, *1814*(8), 969-976. https://doi.org/10.1016/j.bbapap.2010.11.005

Tripathi, A., Kriti Shrinet, Singh, V. K., & Kumar, A. (2019). Molecular modelling and docking of Mus musculus HMGB1 inflammatory protein with CGA. Bioinformation, 15(7), 467–473. https://doi.org/10.6026/97320630015467

van den Brand, J. M. A., Haagmans, B. L., van Riel, D., Osterhaus, A. D. M. E., & Kuiken, T. (2014). The Pathology and Pathogenesis of Experimental Severe Acute

Respiratory Syndrome and Influenza in Animal Models. *Journal of Comparative Pathology*, *151*(1), 83-112. https://doi.org/10.1016/j.jcpa.2014.01.004

Verhulst, E., Bruyn, M. D., Pascale Berckmans, Sim, Y., Koen Augustyns, Pintelon, I., Berg, M., Pieter Van Wielendaele, Anne-Marie Lambeir, Yann G.-J. Sterckx, Nelissen, I., & Meester, I. D. (2025). Human Transmembrane Serine Protease 2 (TMPRSS2) on Human Seminal Fluid Extracellular Vesicles Is Proteolytically Active. *Journal of Extracellular Vesicles*, *14*(3). https://doi.org/10.1002/jev2.70061

Viera, K. R. M., & Márquez, P. L. (2020). Utilización de animales en la investigación biomédica y médica: Un estudio preliminar. *Revista Iberoamericana de Bioética*, *12*, Article 12. https://doi.org/10.14422/rib.i12.y2020.011

Wang, H., Yang, Q., Liu, X., Xu, Z., Shao, M., Li, D., Duan, Y., Tang, J., Yu, X., Zhang, Y., Hao, A., Wang, Y., Chen, J., Zhu, C., Guddat, L., Chen, H., Zhang, L., Chen, X., Jiang, B., … Yang, H. (2023). Structure-based discovery of dual pathway inhibitors for SARS-CoV-2 entry. *Nature Communications*, *14*(1), 7574. https://doi.org/10.1038/s41467-023-42527-5

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T. A. P., Rempfer, C., Bordoli, L., Lepore, R., & Schwede, T. (2018). SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Research*, *46*(W1), W296-W303. https://doi.org/10.1093/nar/gky427

Widiasta, A., Sribudiani, Y., Nugrahapraja, H., Hilmanto, D., Sekarwana, N., & Rachmadi, D. (2020). Potential role of ACE2-related microRNAs in COVID-19-associated nephropathy. *Non-coding RNA Research*, *5*(4), 153-166. https://doi.org/10.1016/j.ncrna.2020.09.001

Williams, C. J., Headd, J. J., Moriarty, N. W., Prisant, M. G., Videau, L. L., Deis, L. N., Verma, V., Keedy, D. A., Hintze, B. J., Chen, V. B., Jain, S., Lewis, S. M.,

Arendall, W. B., Snoeyink, J., Adams, P. D., Lovell, S. C., Richardson, J. S., & Richardson, D. C. (2018). MolProbity: More and better reference data for improved all-atom structure validation. *Protein Science: A Publication of the Protein Society*, *27*(1), 293-315. https://doi.org/10.1002/pro.3330

Wu, C.-Y., Lin, Y.-S., Yang, Y.-H., Shu, L.-H., Cheng, Y.-C., & Liu, H. T. (2020). Potential Simultaneous Inhibitors of Angiotensin-Converting Enzyme 2 and Transmembrane Protease, Serine 2. *Frontiers in Pharmacology*, *11*, 584158. https://doi.org/10.3389/fphar.2020.584158

Yoo, J. il, Lee, Y. S., Song, C.-Y., & Kim, B. S. (2004). Purification and Characterization of a 43-Kilodalton Extracellular Serine Proteinase from Cryptococcus neoformans. *Journal of Clinical Microbiology*, *42*(2), 722-726. https://doi.org/10.1128/jcm.42.2.722-726.2004

Zhou, Y., Vedantham, P., Lu, K., Agudelo, J., Carrion, R., Nunneley, J. W., Barnard, D., Pöhlmann, S., McKerrow, J. H., Renslo, A. R., & Simmons, G. (2015). Protease inhibitors targeting coronavirus and filovirus entry. *Antiviral Research*, *116*, 76-84. https://doi.org/10.1016/j.antiviral.2015.01.011

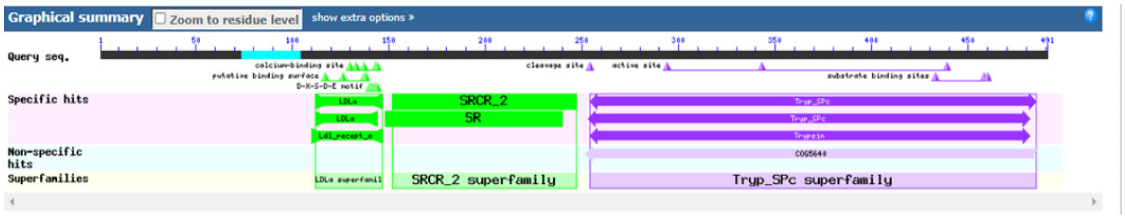
Zmora, P., Molau-Blazejewska, P., Bertram, S., Walendy-Gnirß, K., Nehlmeier, I., Hartleib, A., Moldenhauer, A.-S., Konzok, S., Dehmel, S., Sewald, K., Brinkmann, C., Curths, C., Knauf, S., Gruber, J., Mätz-Rensing, K., Dahlmann, F., Braun, A., & Pöhlmann, S. (2017). Non-human primate orthologues of TMPRSS2 cleave and activate the influenza virus hemagglutinin. *PloS One*, *12*(5), e0176597. https://doi.org/10.1371/journal.pone.0176597

X. ANEXOS

Anexo 1. Perfiles y firmas de motivos determinados por alta probabilidad de ocurrencia para la proteína TMPRSS2 en *Mus musculus*, según Prosite.

Accesión	Predicción de enlaces disulfuro	Descripción	Intervalo	Score
PS50240: Dominio de serinoproteasa tipo tripsina	Cys 279 a Cys 295 Cys 408 a Cys 424 Cys 435 a Cys 463	Su actividad catalítica proporciona un sistema de relé de carga que implica un residuo de D372 unido por hidrógeno a una H294, que a su vez está unido por hidrógeno a una S439.	254 - 487	36.662
PS50287: Dominio rico en cisteína del receptor scavenger (SRCR)	Cys 171 a Cys 188 Cys 184 a Cys 188 Cys 188 a Cys 188	Dominio antiguo y altamente conservado en aprox. 110 residuos de diversas proteínas secretadas y de la superficie celular.	157 - 188	8.501
PS50068: Dominio del receptor de lipoproteínas de baja densidad de clase A (LDLRA)	Cys 112 a Cys 125 Cys 119 a Cys 138 Cys 132 a Cys 147	Son las principales lipoproteínas transportadoras de colesterol. Presenta repeticiones sucesivas de cisteína de aprox. 40 aminoácidos en el terminal N de esta proteína membrana multidominio.	111-148	11.600

Anexo 2. Resumen gráfico de los dominios conservados en la proteína TMPRSS2 de *Mus musculus*. LDLa, SRCR_2 y Tryp_SPc, según NCBI. También se señalan los hits específicos, no específicos y las superfamilias a las que pertenecen estos dominios.



Anexo 3. Lista de dominios determinados por alta probabilidad de ocurrencia para la proteína TMPRSS2 en *Mus musculus*, según NCBI.

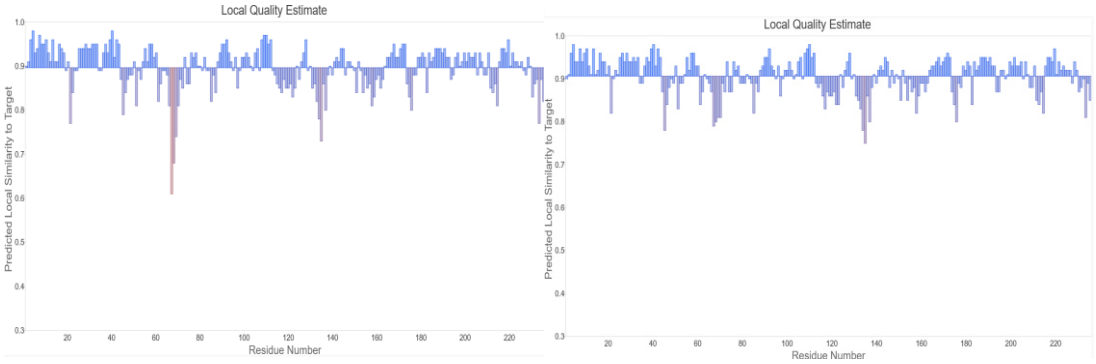
Nombre	Accesión	Descripción	Intervalo	E-Value
Tryp_SPc	cd00190	La proteasa serina, similar a la tripsina, se sintetiza como zimógenos precursores inactivos que se escinden durante una proteólisis limitada para producir sus formas activas.	254-485	3.06e-99
SRCR_2	pfam15494	Familia de dominios ricos en cisteína ubicado aguas arriba de la serina proteasa transmembrana	152-247	2.17e-42
LDLa	cd00112	Región rica en cisteínas, fundamental para el metabolismo del colesterol en los mamíferos. En el extremo N-terminal de esta proteína de membrana multidominio, hay 7 repeticiones sucesivas ricas en cisteína, cruciales para la función de unión de LDL y la estabilidad estructural del receptor.	112-147	4.27e-08

Anexo 4. Resultados de los estados de predicción de estructuras secundarias por PHD de TMPRSS2 de *Mus musculus*. Codificada por colores la longitud de la secuencia y el porcentaje de cada elemento secundario.



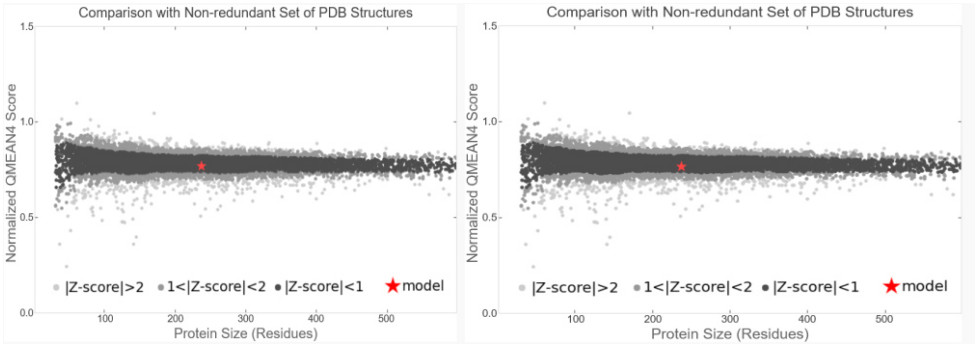
Anexo 5. Estimación de calidad local para cada residuo de los modelos predichos del dominio serinoproteasaTMPRSS2 de *Mus musculus* a partir de las estructuras de referencia (A) 7MEQ, (B) 7XYD.

A B

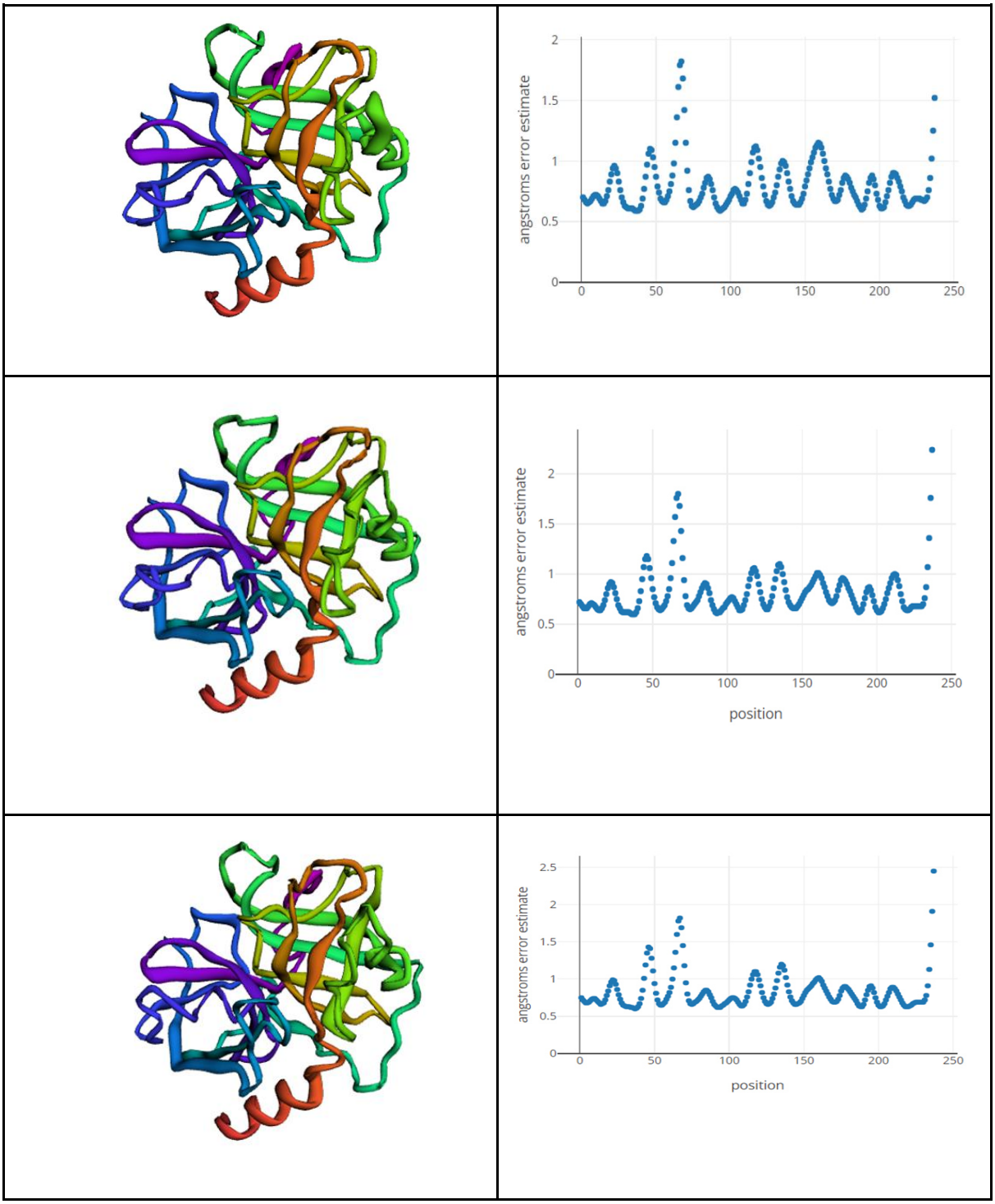


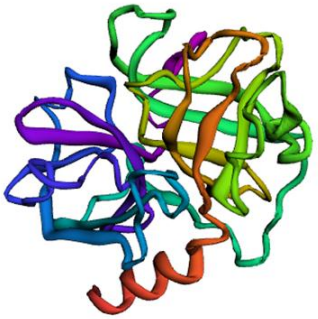
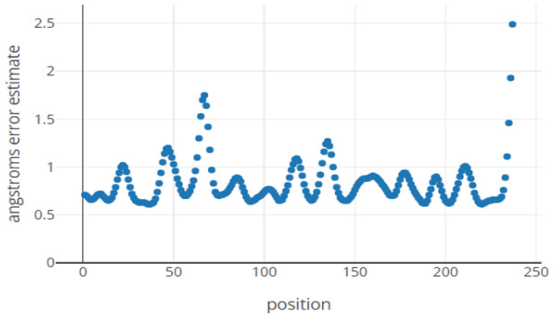
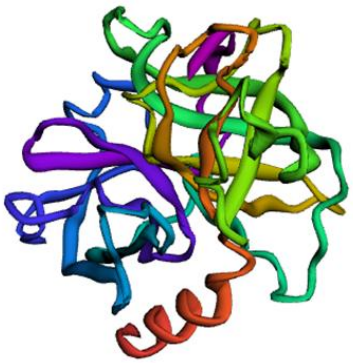
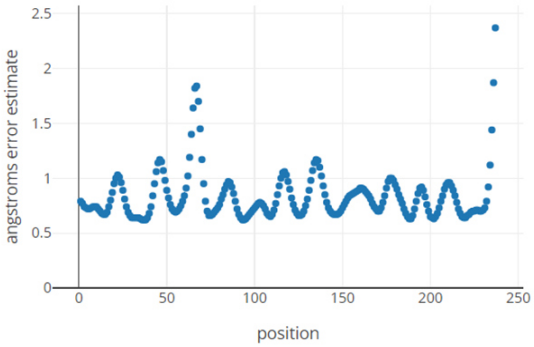
Anexo 6. QMEAN Z-score de los modelos generados a partir de las estructuras de referencia (A) 7MEQ, (B) 7XYD en función del tamaño de la proteína (en residuos) en comparación con un conjunto no redundante de estructuras del PDB.

A B

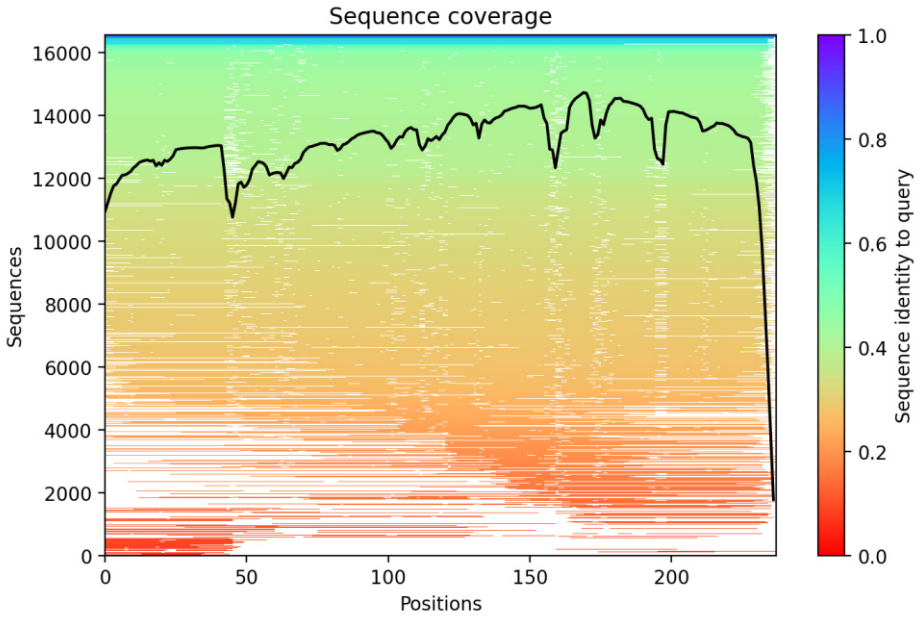


Anexo 7. Resultados de predicción del dominio SP mediante el método *ab initio*.

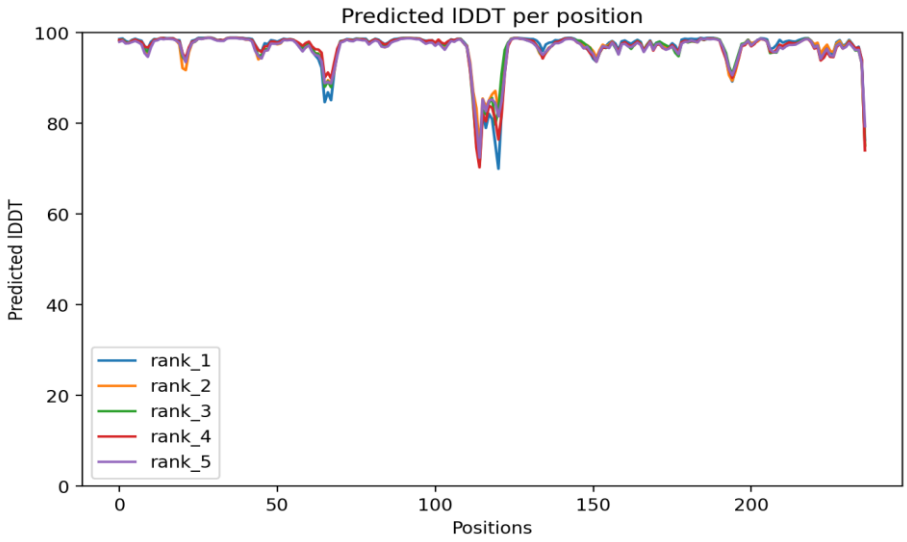


	angstroms error estimate 2.5 2 1.5 1 0.5 0 0 50 100 150 200 250 position
	angstroms error estimate 2.5 2 1.5 1 0.5 0 0 50 100 150 200 250 position

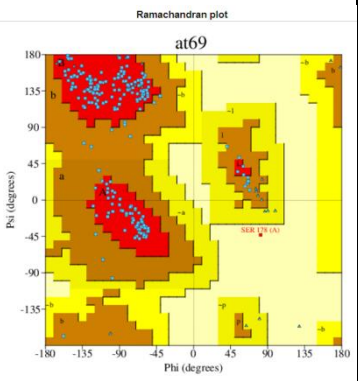
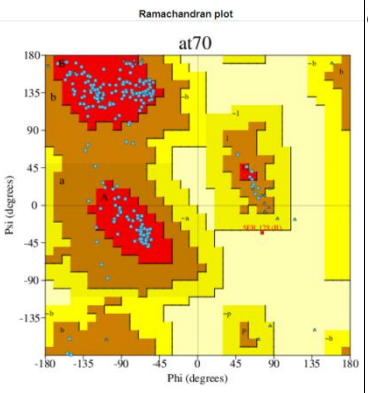
Anexo 8. Cobertura e identidad de secuencia en el alineamiento múltiple de secuencias (MSA) usando ColabFold en la predicción tridimensional del dominio SP de TMPRSS2.

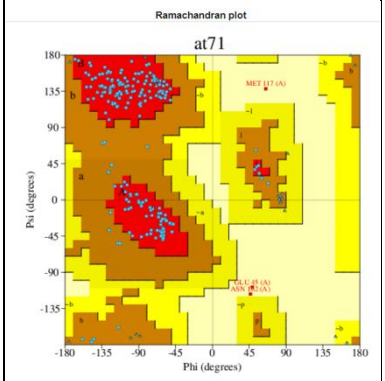
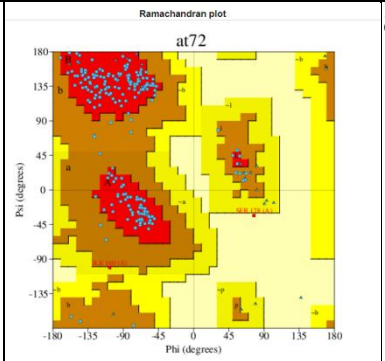
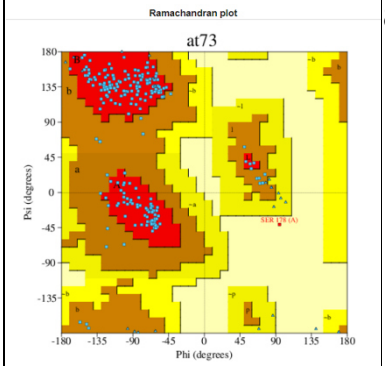


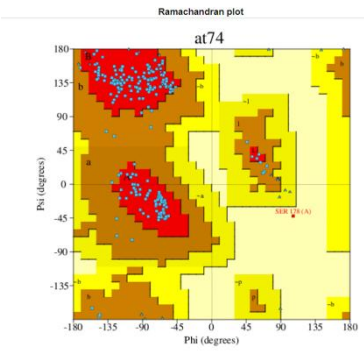
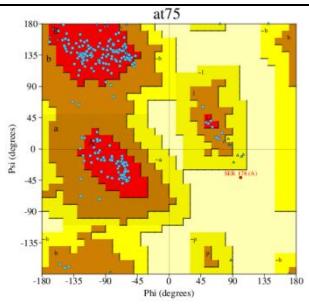
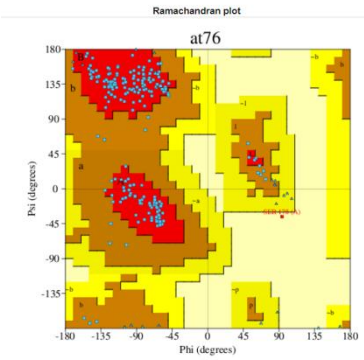
Anexo 9. Predicción de lDDT para cada posición de la proteína TMPRSS2 en cinco modelos diferentes, etiquetados como rank_1, rank_2, rank_3, rank_4 y rank_5 predichos por Alphafold.

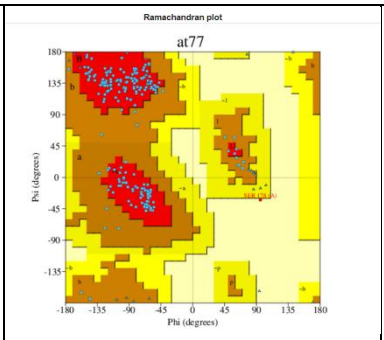


Anexo 10. Gráfica del Plot de Ramachandran para validar las estructuras de los 9 modelos elegidos por las diferentes metodologías de modelamiento por el servidor de validación PDBsum Generate. Los triángulos representan los residuos de glicina (Gly). Los cuadrados representan los residuos de prolina (Pro).

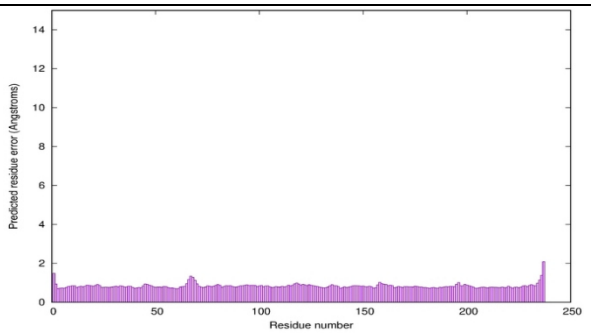
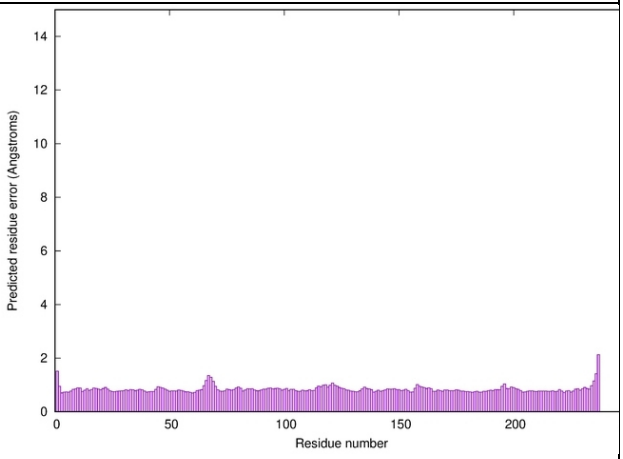
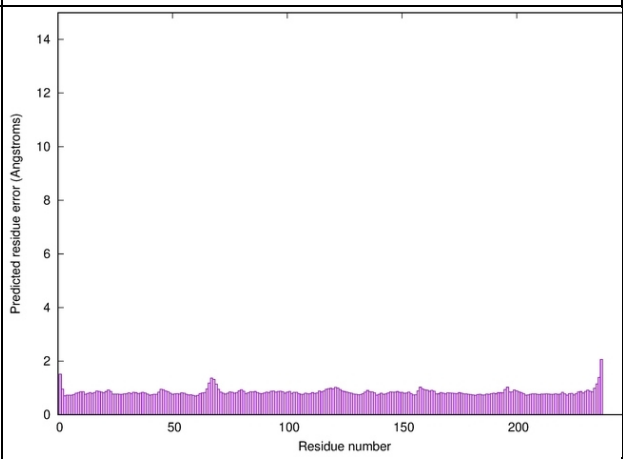
Modelo	Ramachandran plot	Regiones más favorecidas	Regiones adicionales permitidas	Sumatoria	G-Factors
Homología 01 (plantilla_7meq.1)		87.2%	12.2%	99.4%	-0.10
Homología 02 (plantilla_7xyd.1)		87.2%	12.2%	99.4%	-0.09

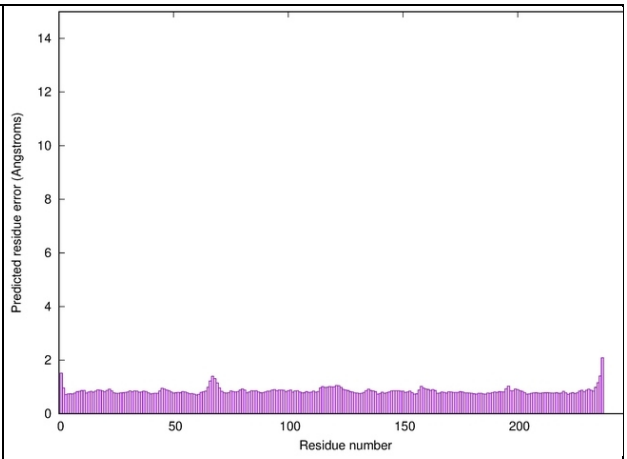
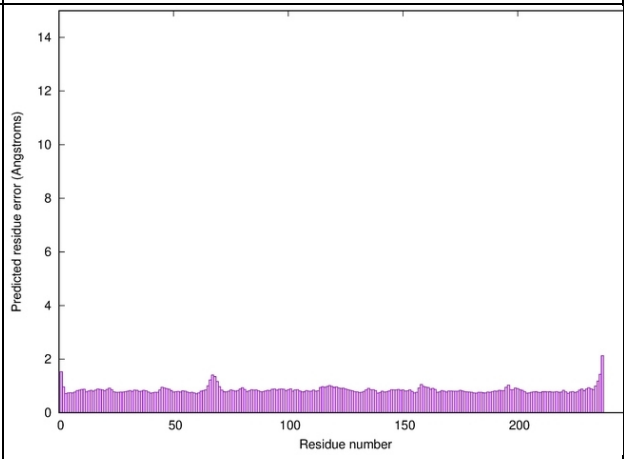
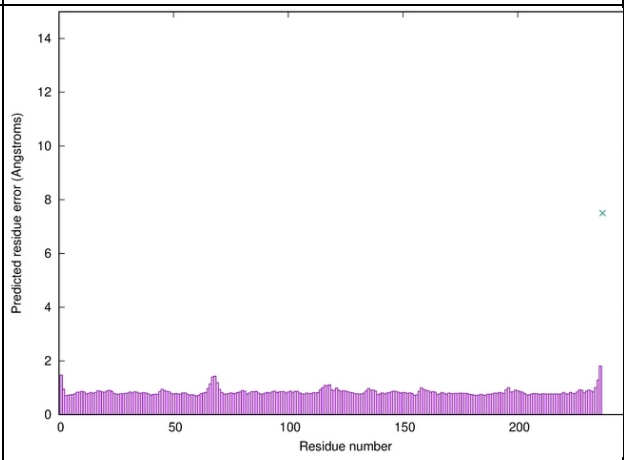
ab initio		89.8%	8.6%	98.4%	0.19
phyre2 (plantilla_c7meqA)		87.8%	11.2%	99%	0.14
alphafold 01		89.3%	10.2%	99.5%	-0.07

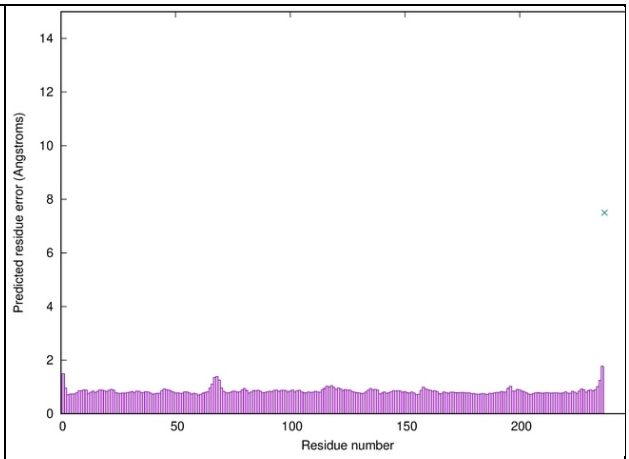
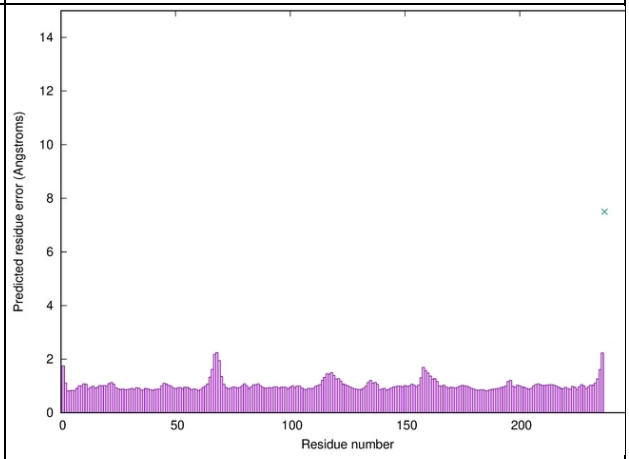
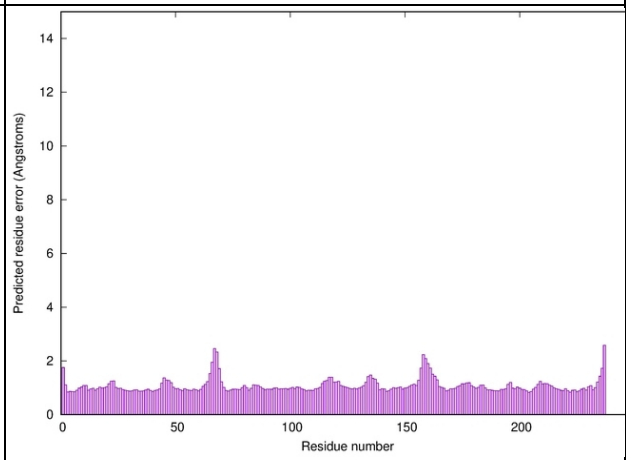
alphafold 02		89.3%	10.2%	99.5%	-0.07
alphafold 03		89.3%	10.2%	99.5%	0.57
alphafold 04		89.3%	10.2%	99.5%	-0.04

alphafold 05		87.8%	11.7%	99.5%	-0.08

Anexo 11. Evaluación estadística de la calidad de los modelos con ModFLOD. En base al error entre los residuos predichos y residuos reales, los valores no superaron los 4Å para la región del dominio serinoproteasa en los 9 modelos.

Estrategia de modelamiento	Confidencia & P-value	Score global de calidad	Plot de error de residuos
alphafold (Modelo 3)	2.958E-5	0.9455	
alphafold (Modelo 5)	2.973E-5	0.9451	
alphafold (Modelo 4)	2.978E-5	0.9449	

alphafold (Modelo 2)	3.01E-5	0.944 0	
alphafold (Modelo 1)	3.013E-5	0.943 9	
Swiss Model (Plantilla_7xyd.1)	3.106E-5	0.941 2	

Swiss model (Plantilla_**7meq.1**)	3.106E-5	0.9412	
Phyre 2 (Plantilla_c7meqA)	4.119E-5	0.9164	
ab initio Modelo 1	4.315E-5	0.9123	

Anexo 12. Resultados del docking molecular del compuesto nafamostat con el modelo de TMPRSS2 generado con SWISS-MODEL.

RUN	Energía de unión (Kcal/mol)	Ki (uM)	No. enlaces de Hidrógeno	Residuos clave del sitio catalítico
1	-3.99	1.2	ninguno	-
2	-3.63	2.19	1	-
3	-5.52	90.26	4	His 41
4	-4.58	435.83	Ninguno	-
5	-4.01	1.14	2	-
6	-4.88	266.6	1	-
7	-4.3	707.81	1	-
8	-4.37	625.68	1	-
9	-3.92	1.34	None	-
10	-4.97	225.78	2	-

Anexo 13. Resultados del docking molecular del compuesto camostat con el modelo de TMPRSS2 generado con SWISS-MODEL.

RUN	Energía de unión (Kcal/mol)	Ki (uM)	No. enlaces de Hidrógeno	Residuos clave del sitio catalítico
1	-3.32	3.66	Ninguno	-
2	-3.09	5.41	Ninguno	His41 Ser186
3	-3.74	1.82	1	-
4	-3.14	5.03	ninguno	-
5	-3.18	4.7	1	His41
6	-2.11	28.49	1	-
7	-3.47	2.84	1	-
8	-3.53	2.57	2	-
9	-3.31	3.72	ninguno	-
10	1.67	59.25	1	-

Anexo 14. Resultados del docking molecular del compuesto lopinavir con el modelo de TMPRSS2 generado con SWISS-MODEL.

RUN	Energía de unión (Kcal/mol)	Ki (uM)	No. enlaces de Hidrógeno	Residuos clave del sitio catalítico
1	-1.59	67.8	ninguno	-
2	-0.37	532.9	ninguno	-
3	-0.93	209.14	ninguno	-
4	-1.59	68.09	ninguno	-
5	-1.7	56.71	ninguno	-
6	-0.9	219.75	ninguno	His41
7	-1.69	57.68	ninguno	-
8	-2.19	24.76	1	His41
9	-0.55	397.38	Ninguno	-
10	-2.62	11.93	2	His41

Anexo 15. Resultados del docking molecular del compuesto valrubicin con el modelo de TMPRSS2 generado con SWISS-MODEL.

RUN	Energía de unión (Kcal/mol)	Ki (uM)	No. enlaces de Hidrógeno	Residuos clave del sitio catalítico
1	-4.58	441.49	ninguno	His41
2	-3.24	4.18	1	His41
3	-3.66	2.09	ninguno	His41
4	-3.58	2.37	1	
5	-4.02	1.13	ninguno	His41
6	-2.79	9.01	ninguno	
7	-4.52	484.73	1	His41
8	-3.38	3.31	ninguno	His41
9	-4.5	499.05	ninguno	His41
10	-4.48	517.53	2	His41

Anexo 16. Resultados del docking molecular del compuesto nafamostat con el modelo de TMPRSS2 generado con Alphafold (modelo 3).

RUN	Energía de unión (Kcal/mol)	Ki (uM)	No. enlaces de hidrógeno	Residuos clave del sitio catalítico
1	-4.15	911.39	1	-
2	-4.5	503	0	Ser186
3	-4.07	1.05	1	-
4	-3.92	1.34	ninguno	-
5	-3.41	3.14	ninguno	-
6	-3.02	6.16	1	-
7	-3.38	3.34	ninguno	-
8	-3.25	4.12	1	-
9	-3.78	1.69	1	-
10	-4.88	980.01	ninguno	-

Anexo 17. Resultados del docking molecular del compuesto camostat con el modelo de TMPRSS2 generado con Alphafold (modelo 3).

RUN	Energía de unión (Kcal/mol)	Ki (uM)	No. enlaces de hidrógeno	Residuos clave del sitio catalítico
1	-3.52	2.62	ninguno	Ser186
2	-4.28	727.05	ninguno	His41
3	-3.9	1.38	2	His41
4	-4.27	740.54	1	Ser186 His41
5	-3.51	2.66	ninguno	Ser186 His41
6	-3.19	4.61	2	Ser186 His41
7	-4.16	895.17	1	His41
8	-3.2	4.51	1	Ser186 His41
9	-3.26	4.09	ninguno	Ser186 His41
10	-4.39	604.87	3	Ser186 His41

Anexo 18. Resultados del docking molecular del compuesto lopinavir con el modelo de TMPRSS2 generado con Alphafold (modelo 3)

RUN	Energía de unión (Kcal/mol)	Ki (uM)	No. enlaces de hidrógeno	Residuos clave del sitio catalítico
1	-0.75	283.75	ninguno	-
2	-3.06	5.68	ninguno	-
3	-0.99	187.76	ninguno	-
4	-1.57	70.37	ninguno	His41 Ser186
5	-0.94	203.19	ninguno	His41
6	-0.83	245.97	ninguno	-
7	-0.2	715.32	ninguno	-
8	-1.11	154.01	ninguno	-
9	-1.08	162.68	ninguno	-
10	-2.17	25.49	ninguno	His41

Anexo 19. Resultados del docking molecular del control negativo valrubicin con el modelo de TMPRSS2 generado con Alphafold (modelo 3).

RUN	Energía de unión (Kcal/mol)	Ki (uM)	No. enlaces de hidrógeno	Residuos clave del sitio catalítico
1	-4.18	870.16	ninguno	-
2	-5.65	71.82	2	-
3	-5.56	84.51	1	-
4	-3.55	2.51	ninguno	-
5	-5.62	76.42	2	-
6	-2.9	7.46	1	-
7	-3.96	1.26	3	-
8	-3.37	3.36	2	-
9	-3.05	5.83	ninguno	-
10	-4.74	334.64	3	Ser 186

Anexo 20. Resultados del docking molecular del compuesto Adenosine triphosphate con el modelo del dominio serinoproteasaTMPRSS2.

RUN	Energía de unión (Kcal/mol)	Ki (uM)	No. enlaces de hidrógeno	Residuos clave del sitio catalítico
1	-1.04	173.22	ninguno	-
2	0.57	-	ninguno	-
3	1.23	-		-
4	0.44	-	ninguno	-
5	-1.03	175.25	ninguno	-
6	0.57	-	ninguno	-
7	-0.0	-	ninguno	-
8	-0.25	659.79	ninguno	-
9	-0.49	434.49	ninguno	-
10	-0.32	579.86	ninguno	-

Anexo 21. Interacciones de enlaces de hidrógeno entre TMPRSS2 y los ligandos nafamostat, camostat, valrubicin y lopinavir obtenidas mediante docking molecular en UCSF Chimera y PLIP. La distancia donador–aceptor (Å) representa la distancia entre el átomo donador y el aceptor.

Interacciones entre residuos del sitio activo de TMPRSS2 con nafamostat	**Distancia donador–aceptor (Å)**
Glu 45 (O) – H38 Naf	2.870
Cys 42 (O) – H21 Naf	3.532
Ser 181 (O) – H31 Naf	3.096
Gly 209 (O) – H30 Naf	2.789
Gly 209 (O) – H29 Naf	3.010
Gly 207 (O) – H29 Naf	3.470
Interacciones entre residuos del sitio activo de TMPRSS2 con valrubicin	**Distancia donador–aceptor (Å)**
Cys 26 (SG) – O Val	3.284
Gln 183 (NE2) – O Val	3.281
Val 25 -- (O) Val	**3.215**
O Val – Thr 138 (OG1) Val	2.348
Interacciones entre residuos del sitio activo de TMPRSS2 con camostat	**Distancia donador–aceptor (Å)**
His 41 (NE2) – O7 Cam	3.167
His 41 (O) – H46 Cam	3.203
His 41 (O) — H27 Cam	2.911
SER 186 (OG) — O17 Cam	2.650
ASP 90 (O-) — Cam	2.570
ASP 90 (O3) — Cam	2.630
SER 17 (O3) — Cam	2.700
Gly 184 (N) – O9 Cam	3.415
Glu 44 (OE1) – H46 Cam	2.776
THR 54 (O2) — Cam	2.930
THR 54 (O2) — Cam	3.230
ALA 55 (O2) — Cam	2.901

ALA 55 (O2) — Cam TYR 82 (O2) — Cam ILE 91 (O2) — Cam	3.040 3.190 2.840
Interacciones entre residuos del sitio activo de TMPRSS2 con lopinavir	**Distancia donador–aceptor (Å)**
Gln 183 (NE2) – O9 Lop Asp 140 (OD1) – N48 Lop	2.902 2.931

Anexo 22. Interacciones significativas entre residuos del sitio activo de TMPRSS2 y los ligandos nafamostat, valrubicin, camostat y lopinavir obtenidos en UCSF Chimera y PLIP. Se presentan los residuos clave involucrados en la interacción, el solapamiento (Overlap) y la distancia de interacción en Ångstroms (Å). Las distancias corresponden a la métrica Donador-Aceptor para interacciones polares y puentes de hidrógeno, y a distancias interatómicas mínimas para contactos estéricos

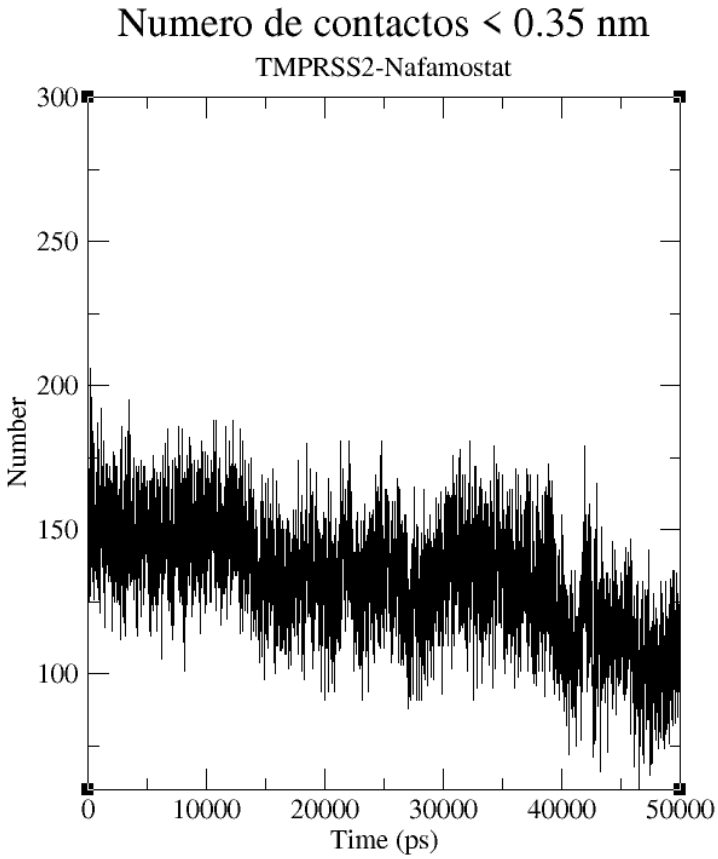
Región estructural	TMPRSS2 con Nafamostat	Tipo de Interacción	Overlap (Å)	Distancia (Å)
Tríada Catalítica	Ser 186 (OG) — C28 Naf	Contacto cercano	-0.022	3.282
	His 41 (O) — C16 Naf	Contacto cercano	-0.040	3.280
	His 41 (CD2) — O10 Naf	Contacto cercano	-0.098	3.338
	His41--- COO Naf	Puente salino	-	4,51
Bolsillo de Especificidad S1	Gly 209 (O) — H30 Naf	PH	0.573	1.907
	Ser 181 (O) — H31 Naf	PH	0.184	2.296
	Gly 209 (O) — H29 Naf	PH	0.269	2.211
	Gly 209 (O) — N3 Naf	PH	0.316	2.789
	Gly 207 (O) – H29 Naf	PH	-	3.470
Entorno de Estabilización y Entrada	Glu 45 (O) — H38 Naf	PH	0.545	1.935
	Gln 183 (2HE2) — O11 Naf	Contacto cercano	0.443	2.070
	Glu 44 (O) — N20 Naf	Contacto cercano	0.279	2.826
	Gln 183 (OE1) — C27 Naf	Contacto cercano	0.263	2.977
	Leu 47 (CD1) — N20 Naf	Contacto cercano	0.194	3.311
	Pro 46 (CD) — N19 Naf	Contacto cercano	0.084	3.421
	TMPRSS2 con valrubicin	**Tipo de Interacción**	**Overlap (Å)**	**Distancia (Å)**

	Thr 138 (HG1) — O Val	PH	0.501	1.959
Tríada	His 41— C Val	Puente salino	-	5.01
Catalítica y	His 41 -- Val	Puente salino	-	5.18
Subsitio S1'	Thr 138 (OG1) — O Val	PH	0.612	2.348
	Ser 186 (CB) — O Val	Contacto cercano	-0.061	3.401
Bucle de	Cys 42 (CA) — C Val	Contacto cercano	0.374	3.386
Entrada N-				
terminal S1'	Val 25 (O) — C Val	Contacto Hidrofóbico	0.283	3.077
	Cys 26 (SG) — C Val	Contacto Hidrofóbico	0.202	3.448
	Val 23 (CG2) — C Val	Contacto Hidrofóbico	0.152	3.338
	His 24 (O) — C Val	Contacto Van der Waals	0.010	3.080
Lazo de				
Regulación	Gln 183 (1HE2) — C Val	Contacto Van der Waals	0.129	2.751
Periférica	Gln 183 (2HE2) — O Val	PH	0.100	2.320
	Gln 183 (NE2) — F Val	Contacto cercano	-0.187	3.092
	Pro 46 (CD) — C Val	Contacto Hidrofóbico	0.169	3.591

	TMPRSS2 con camostat	**Tipo de Interacción**	**Overlap (Å)**	**Distancia (Å)**
Tríada	O17 Cam — SER 186 (OG)	PH	0.330	2.650
Catalítica y	O17 Cam — SER 186 (HN)	Contacto cercano	0.342	2.138
Subsitio S1'	C16 Cam — SER 186 (OG)	Contacto cercano	0.270	2.930
	C31 Cam — HSD 41 (CD2)	Contacto cercano	0.293	3.227
	C31 Cam — HSD 41 (NE2)	Contacto cercano	0.196	3.189
	Cam –-- Asp 90 (O-)	PH	-	2.570
	Cam –-- Asp 90 (O3-)	PH	-	2.700
Bolsillo de	C18 Cam — GLY 184 (HN)	Contacto cercano	0.498	2.202
Especificidad	O17 Cam — GLY 184 (HN)	PH	0.408	2.072
Primario S1	C16 Cam — GLY 184 (HN)	Contacto cercano	0.418	2.282
Bolsillo de	C8 Cam — LYS 87 (HZ2)	Contacto cercano	0.371	2.329
Especificidad	O9 Cam — LYS 87 (HZ2)	Contacto cercano	0.052	2.428
Secundario	C8 Cam — LYS 87 (NZ)	Contacto cercano	0.127	3.198
S2	Cam -- LYS 87 (Anillo)	π-Catión	-	3.950

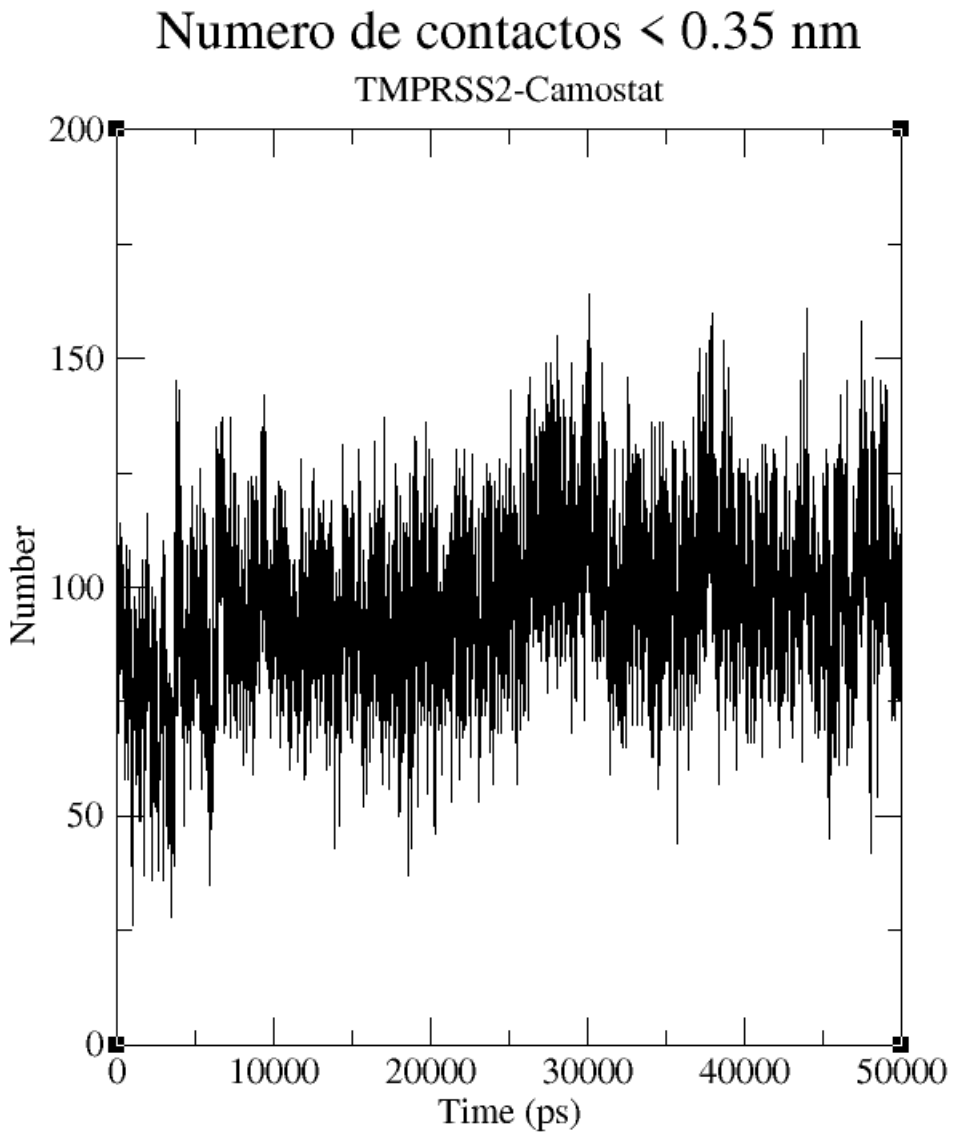
	Cam -- TYR 82 (O2)	PH	-	3.190
	Cam -- ILE 91 (O2)	PH	-	2.840
	Cam – ILE 91	I.Hidrofóbica	-	3.760
Bucle de Entrada y Periferia	H48 Cam — HSD 24 (O)	Contacto cercano	0.405	2.075
	C28 Cam — VAL 25 (O)	Contacto cercano	0.537	2.703
	N24 Cam — VAL 25 (CA)	Contacto cercano	0.509	2.996
	C16 Cam — GLN 183 (CA)	Contacto cercano	0.281	3.299
	Cam --- GLU 44 (Guanidine)	Puente salino	-	4.520
	TMPRSS2 con lopinavir	**Tipo de interacción**	**Overlap (Å)**	**Distancia (Å)**
Tríada Catalítica y Subsitio S1'	Thr 138 (OG1) — C26 Lop	Contacto Hidrofóbico	0.487	2.893
	Thr 138 (OG1) — C37 Lop	Contacto Hidrofóbico	0.258	2.942
	Thr 138 (OG1) — O38 Lop	PH	0.215	3.220
	Ser 186 (CB) — C1 Lop	Contacto débil	-0.125	4.885
	Cys 42 (CA) — C3 Lop	Contacto débil	0.269	3.371
	His 41 (NE2) — C1 Lop	Contacto débil	-0.122	5.627
Lazo de Reconocimiento Asp140	Asp 140 (OD1) — H94 Lop	PH	0.394	2.086
	Asp 140 (OD1) — N48 Lop	PH	0.174	2.931
Bucle de Entrada N-terminal	Val 25 (O) — C10 Lop	Contacto Hidrofóbico	0.528	2.832
	Cys 26 (SG) — C1 Lop	Contacto débil	0.023	3.627
	Val 23 (CG2) — C31 Lop	Contacto Hidrofóbico	0.400	3.240
Lazo de Regulación Periférica	Gln 183 (1HE2) — O9 Lop	Contacto cercano	0.347	2.153
	Gln 183 (NE2) — O9 Lop	PH	0.223	2.902

Anexo 23. Número de contactos (< 0.35 nm) del complejo TMPRSS2–Nafamostat durante 50 ns de simulación de dinámica molecular.

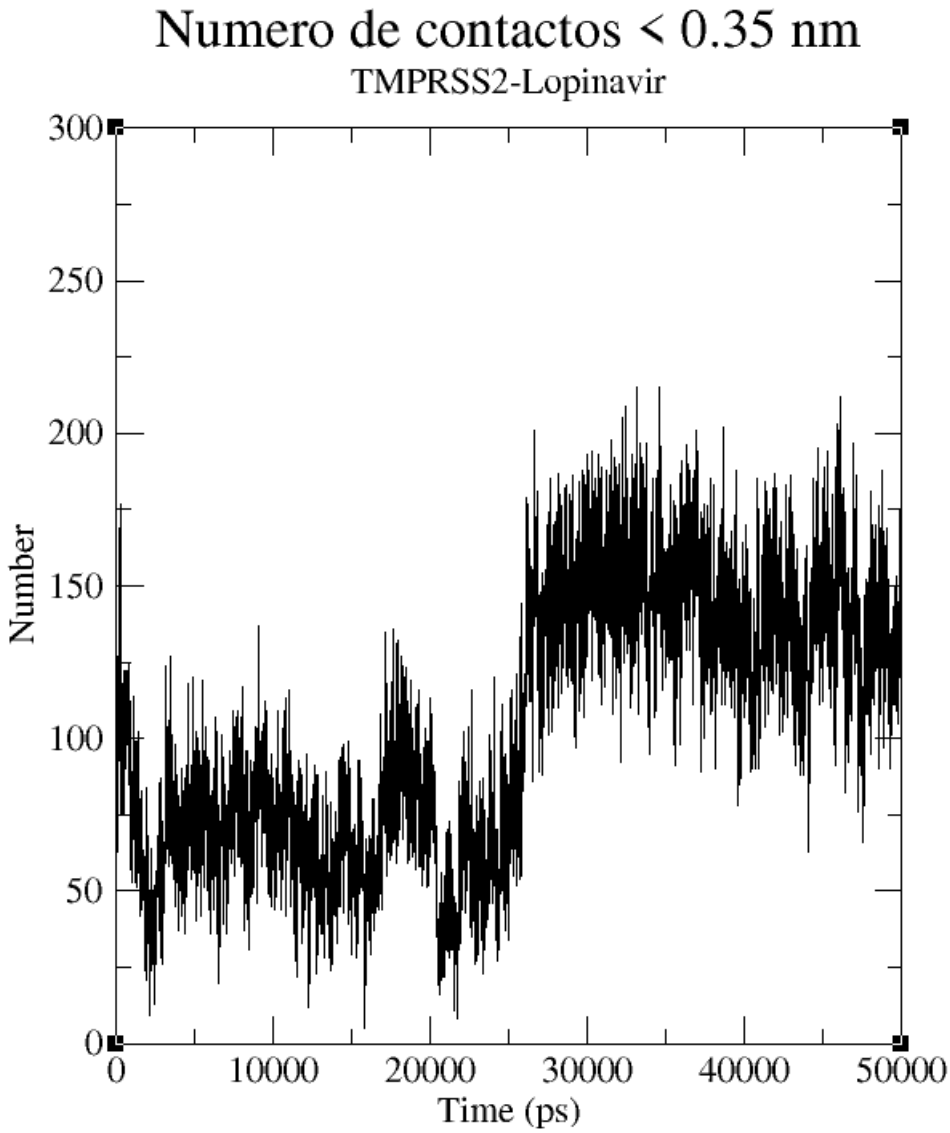


"

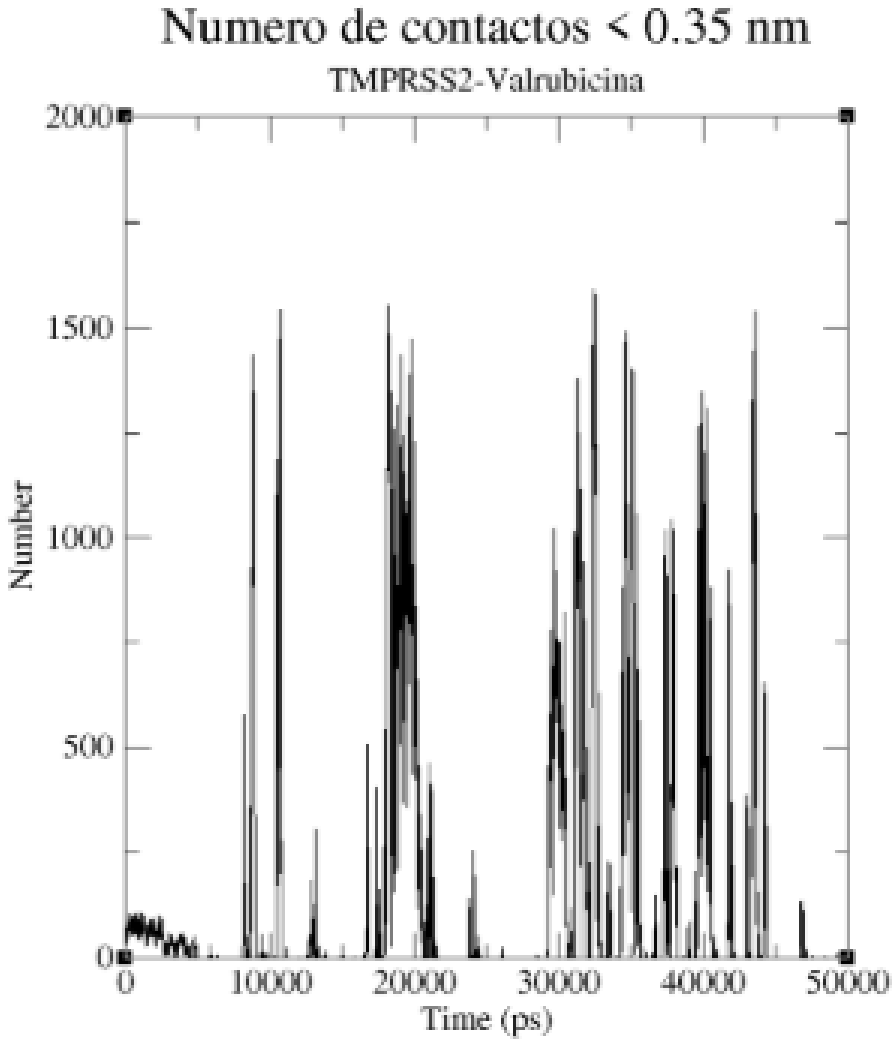
Anexo 24. Número de contactos (< 0.35 nm) del complejo TMPRSS2–Camostat durante 50 ns de simulación de dinámica molecular.



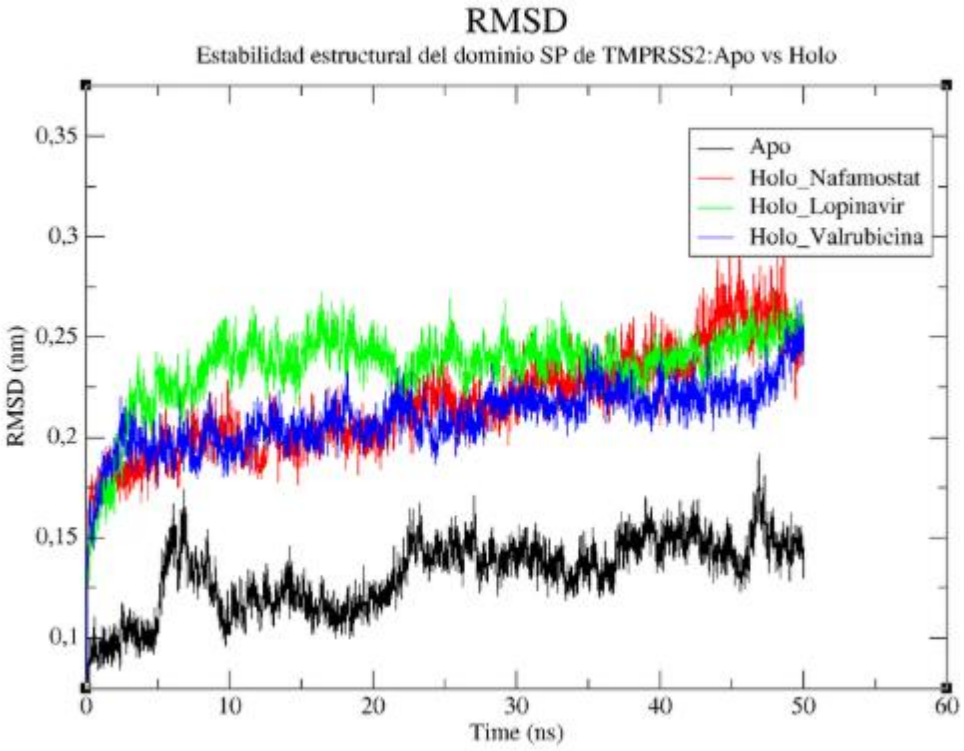
Anexo 25. Número de contactos (< 0.35 nm) del complejo TMPRSS2–Lopinavir durante 50 ns de simulación de dinámica molecular.



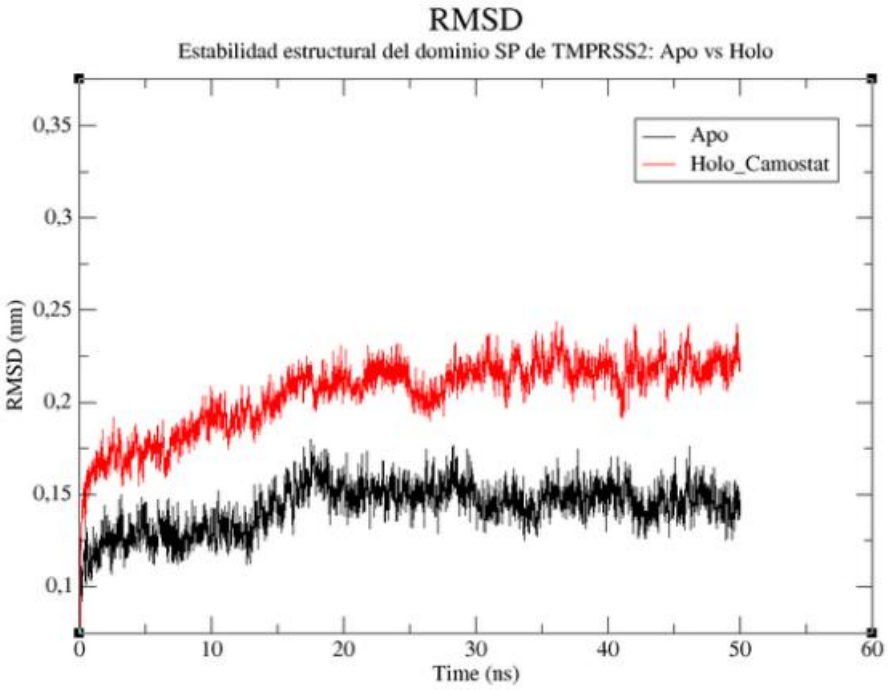
Anexo 26. Número de contactos (< 0.35 nm) del complejo TMPRSS2–Valrubicina durante 50 ns de simulación de dinámica molecular.



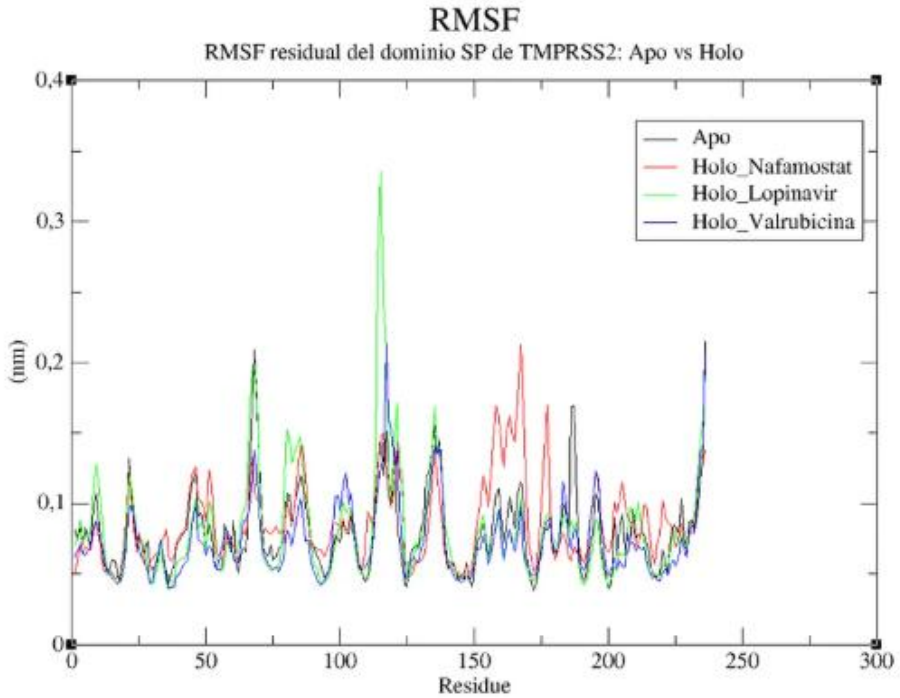
Anexo 27. Estabilidad estructural del dominio SP de TMPRSS2 evaluada mediante RMSD durante 50 ns de dinámica molecular en estado apo y en complejos holo con nafamostat, lopinavir y valrubicina.



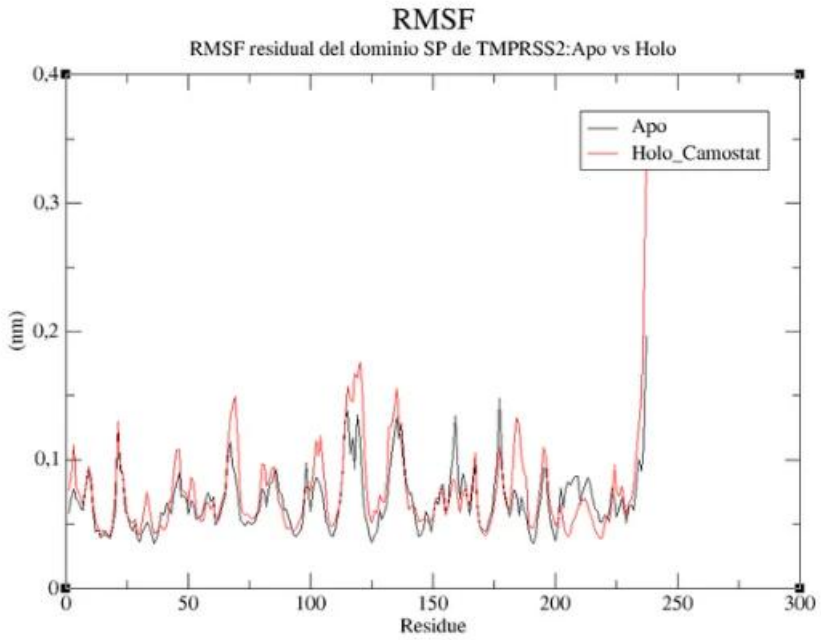
Anexo 28. Estabilidad estructural del dominio SP de TMPRSS2 modelado con camostat, evaluada mediante RMSD en estados apo y holo durante 50 ns de dinámica molecular.



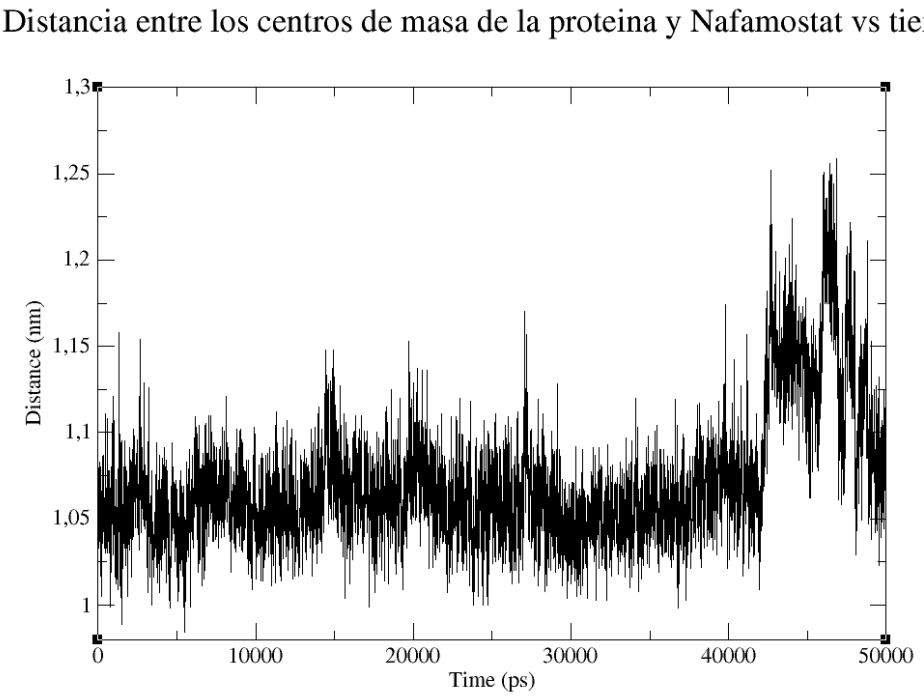
Anexo 29. Perfil de fluctuación residual (RMSF) del dominio SP de TMPRSS2 en estado apo y en complejos holo con nafamostat, lopinavir y valrubicina durante 50 ns de dinámica molecular.



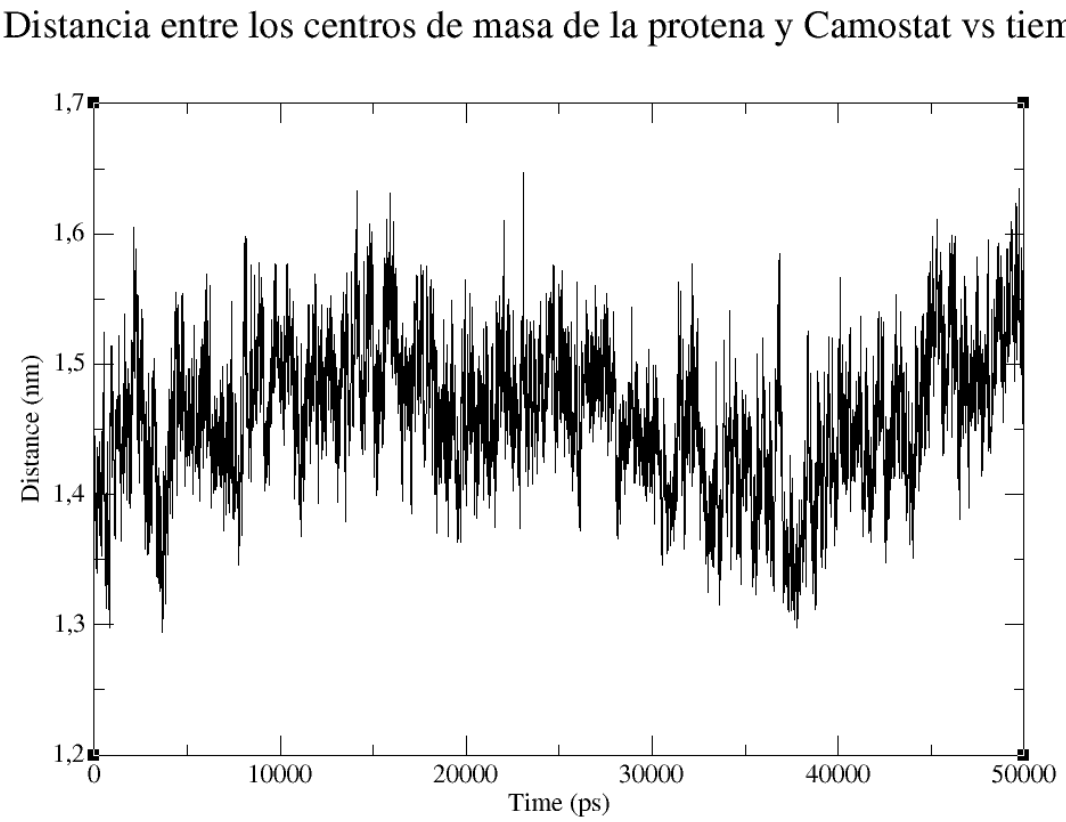
Anexo 30. Perfil de fluctuación residual (RMSF) del dominio SP de TMPRSS2 en estado apo y en complejo con camostat durante 50 ns de dinámica molecular.



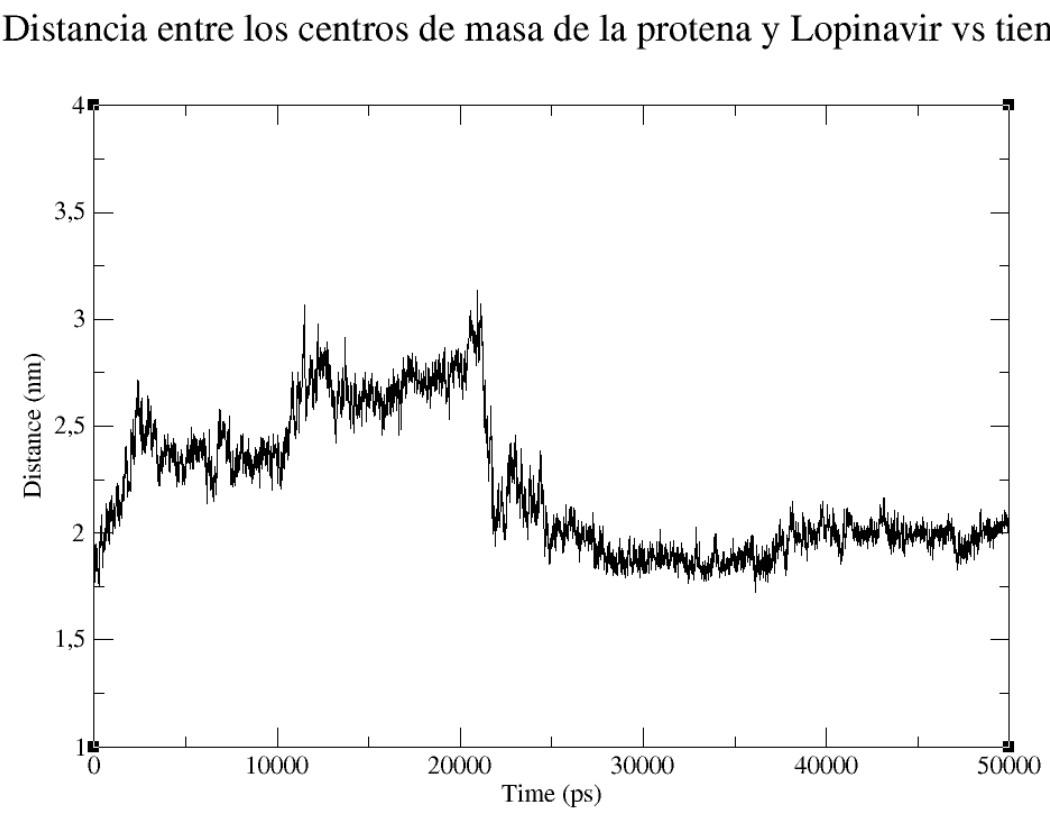
Anexo 31. Distancia entre los centros de masa de TMPRSS2 y Nafamostat durante 50 ns de simulación de dinámica molecular



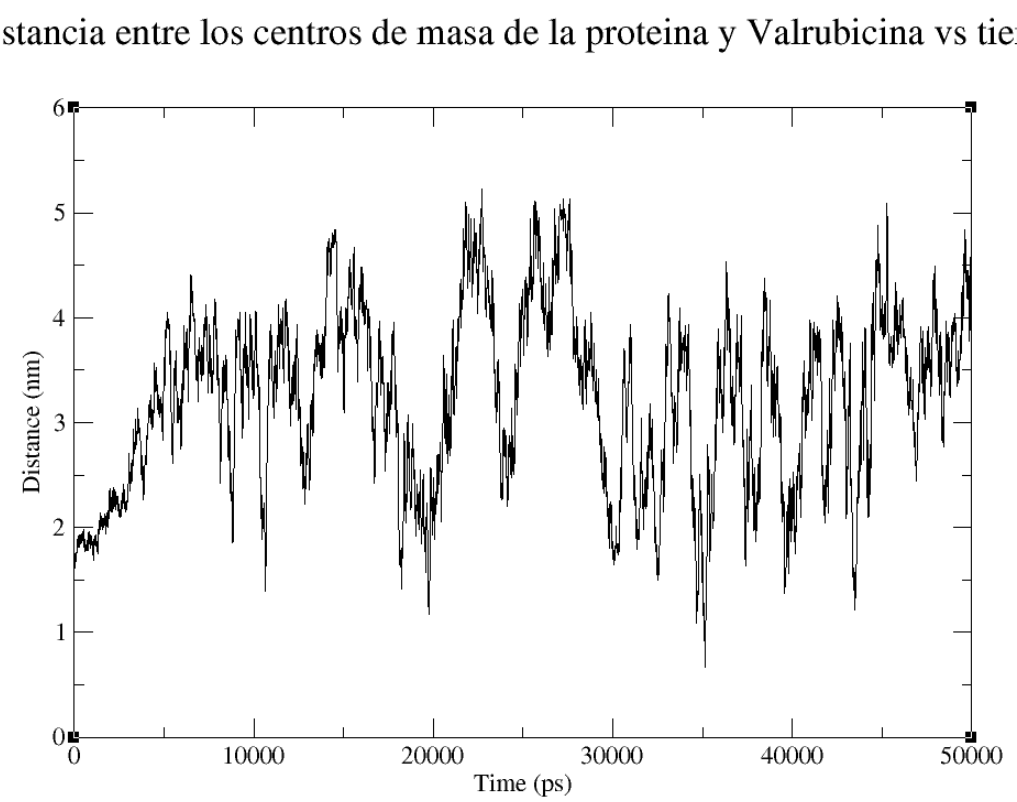
Anexo 32. Distancia entre los centros de masa de TMPRSS2 y Camostat durante 50 ns de simulación de dinámica molecular



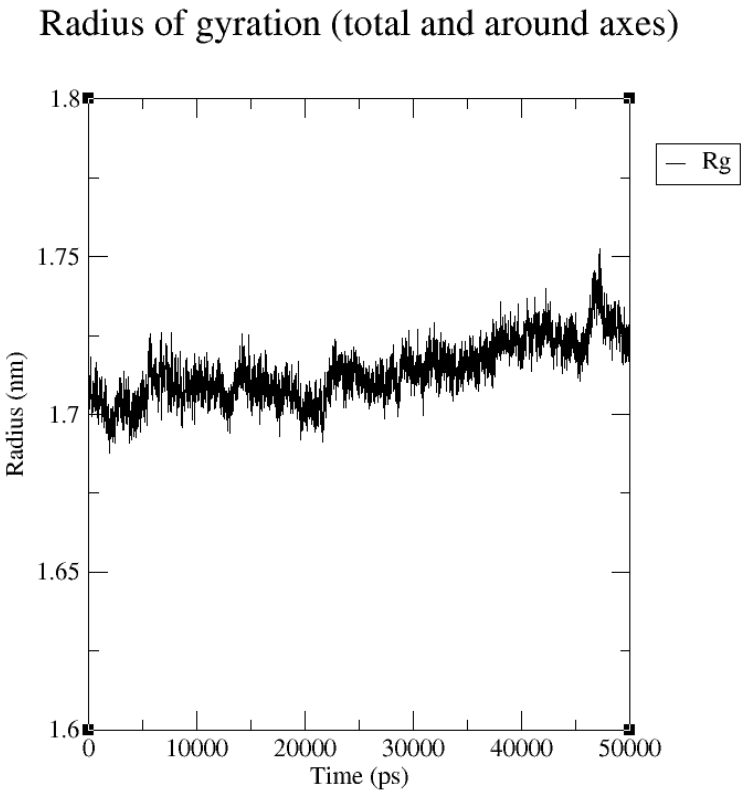
Anexo 33. Distancia entre los centros de masa de TMPRSS2 y lopinavir durante 50 ns de simulación de dinámica molecular



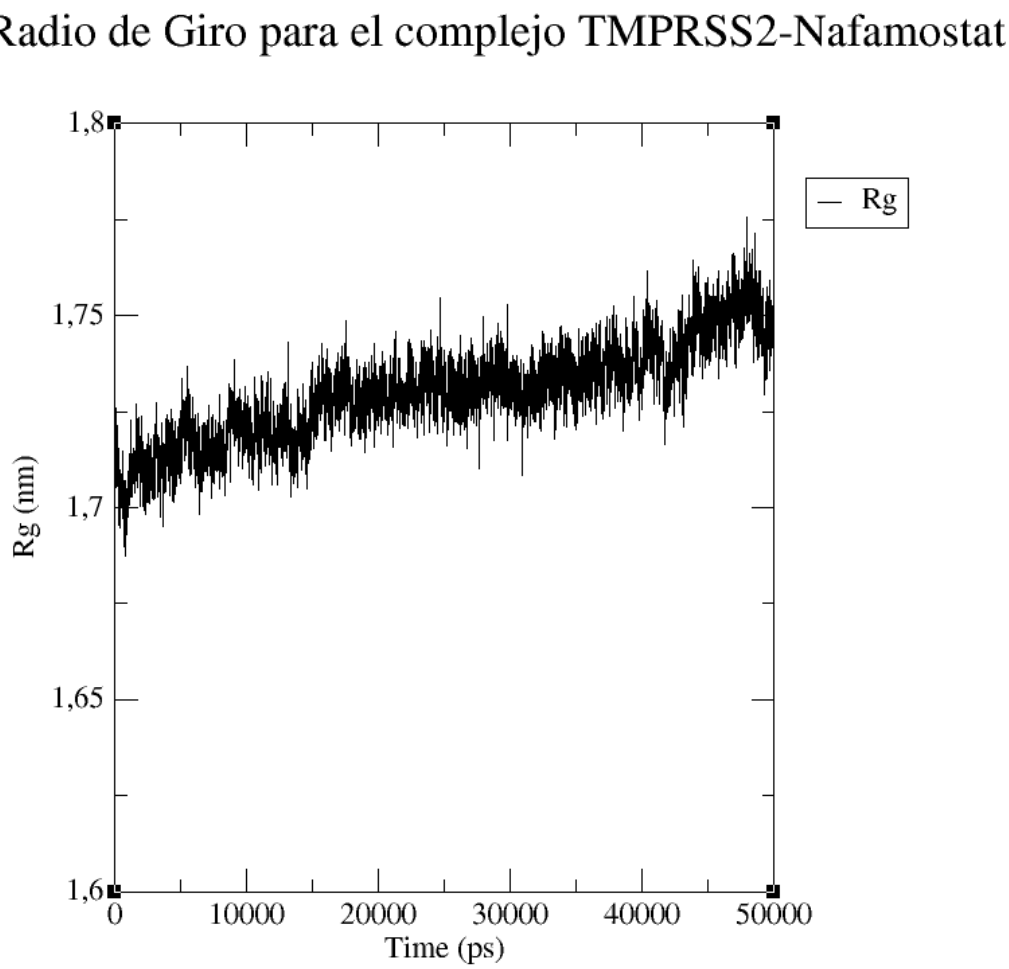
Anexo 34. Distancia entre los centros de masa de TMPRSS2 y Valrubicina durante 50 ns de simulación de dinámica molecular



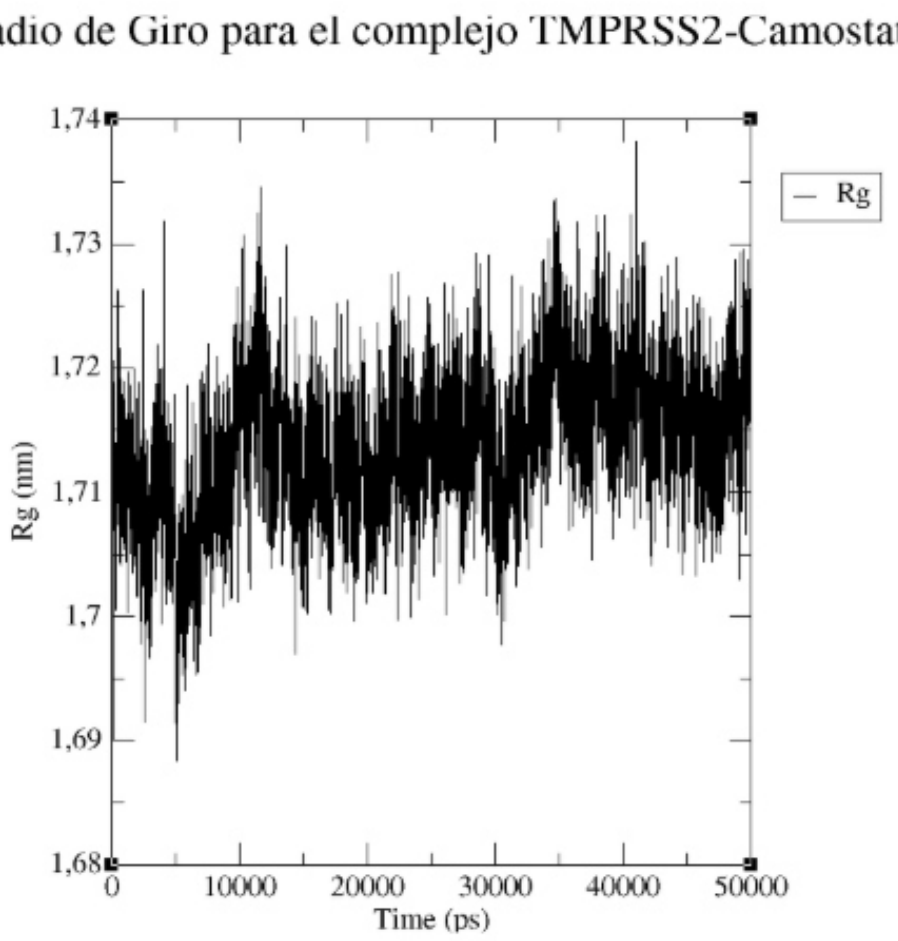
Anexo 35. Perfil del radio de giro (Rg) del dominio serinoproteasa de TMPRSS2.



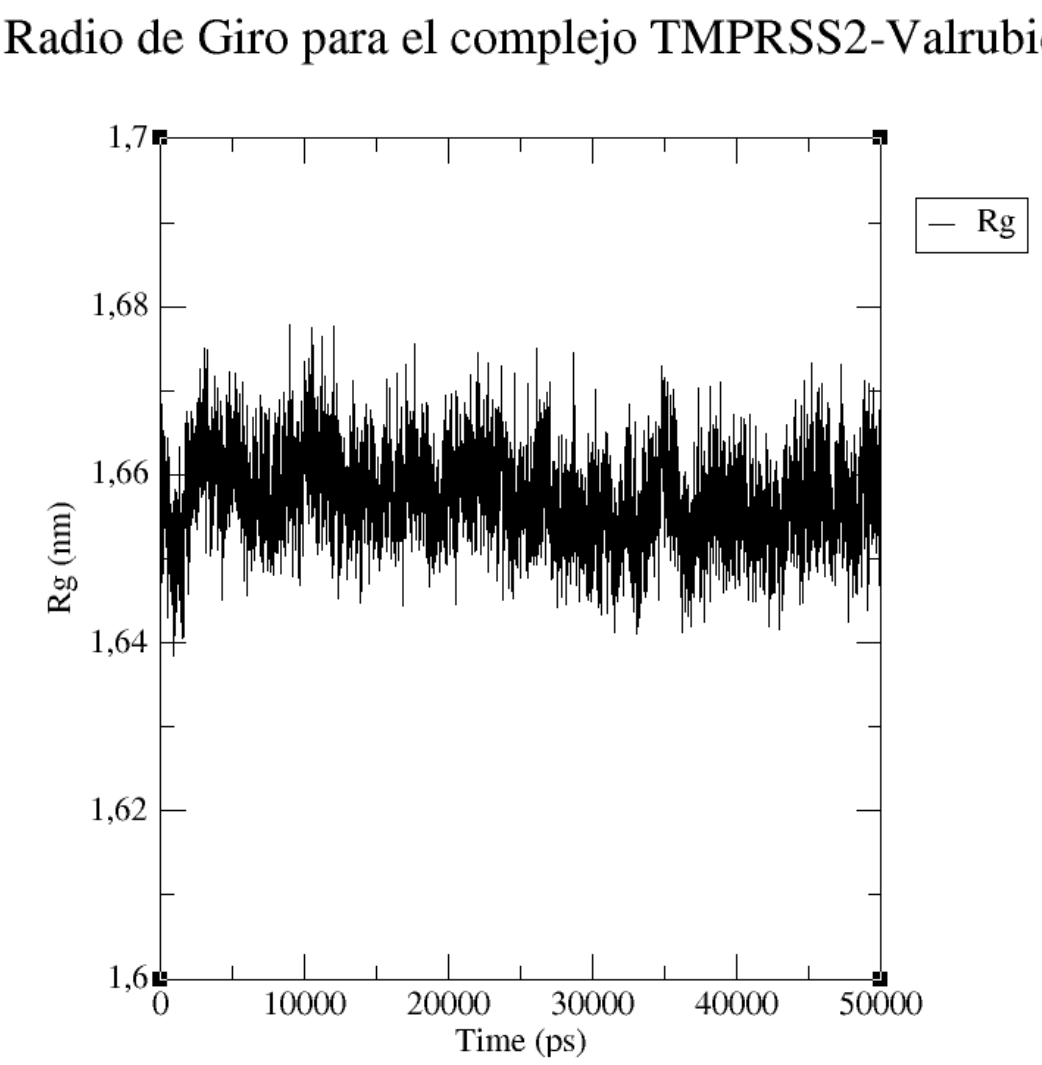
Anexo 36. Perfil del radio de giro (Rg) del complejo TMPRSS2–Nafamostat



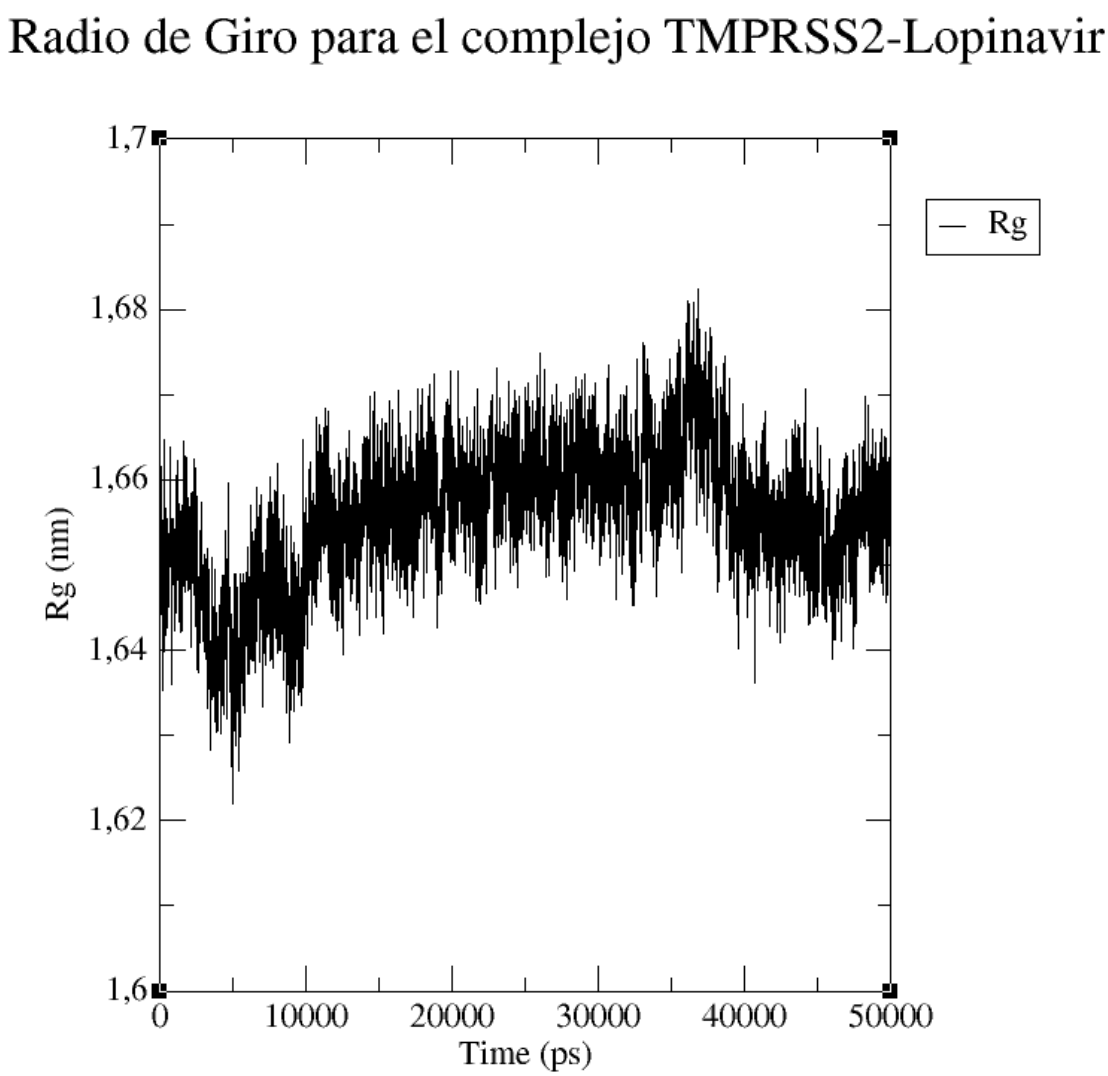
Anexo 37. Perfil del radio de giro (Rg) del complejo TMPRSS2–Camostat



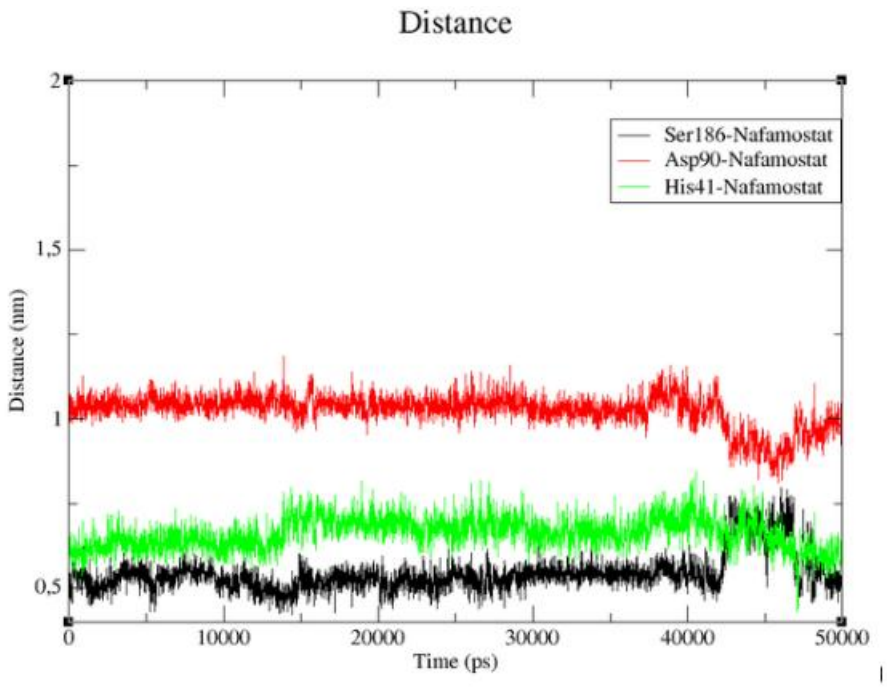
Anexo 38. Perfil del radio de giro (Rg) del complejo TMPRSS2–Valrubicina



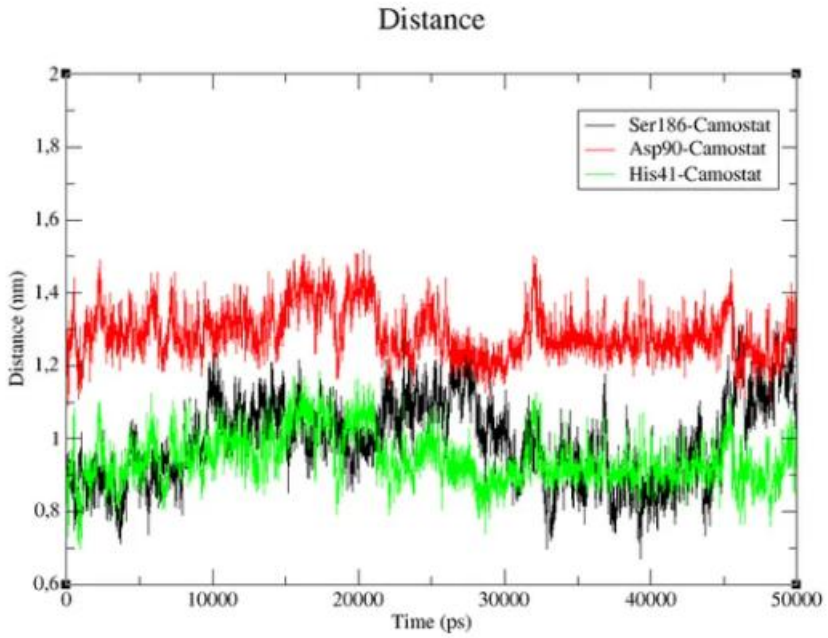
Anexo 39. Perfil del radio de giro (Rg) del complejo TMPRSS2–Lopinavir



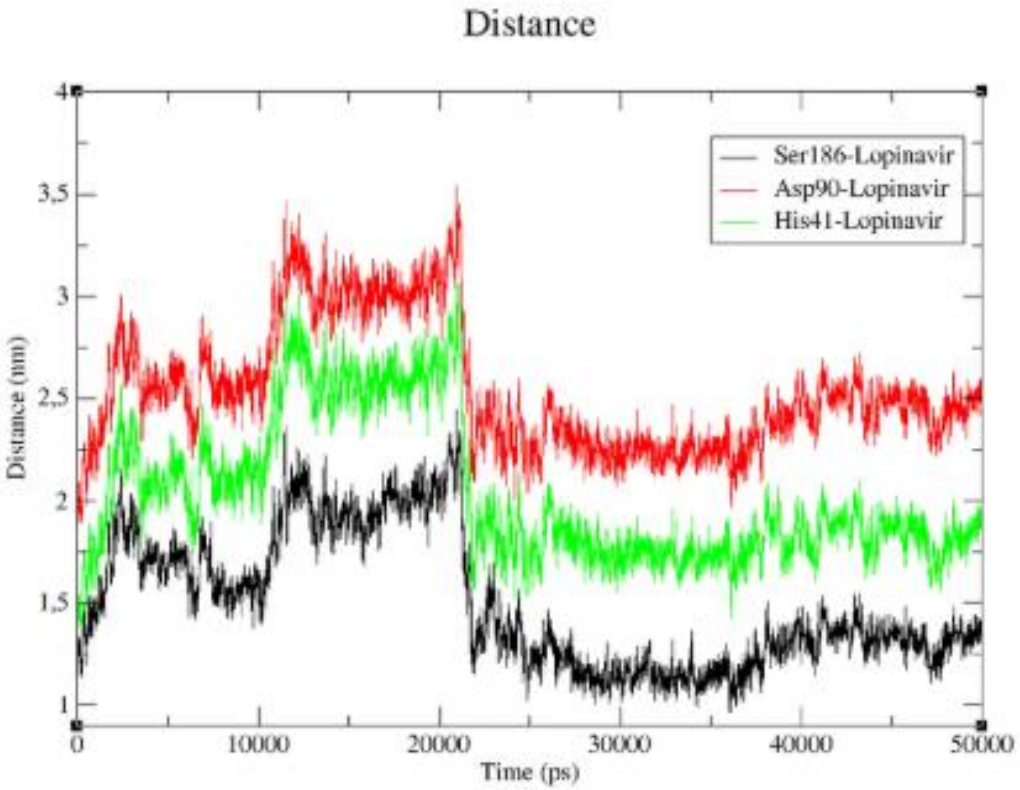
Anexo 40. Distancia promedio entre nafamostat y residuos catalíticos de TMPRSS2 durante 50 ns de simulación



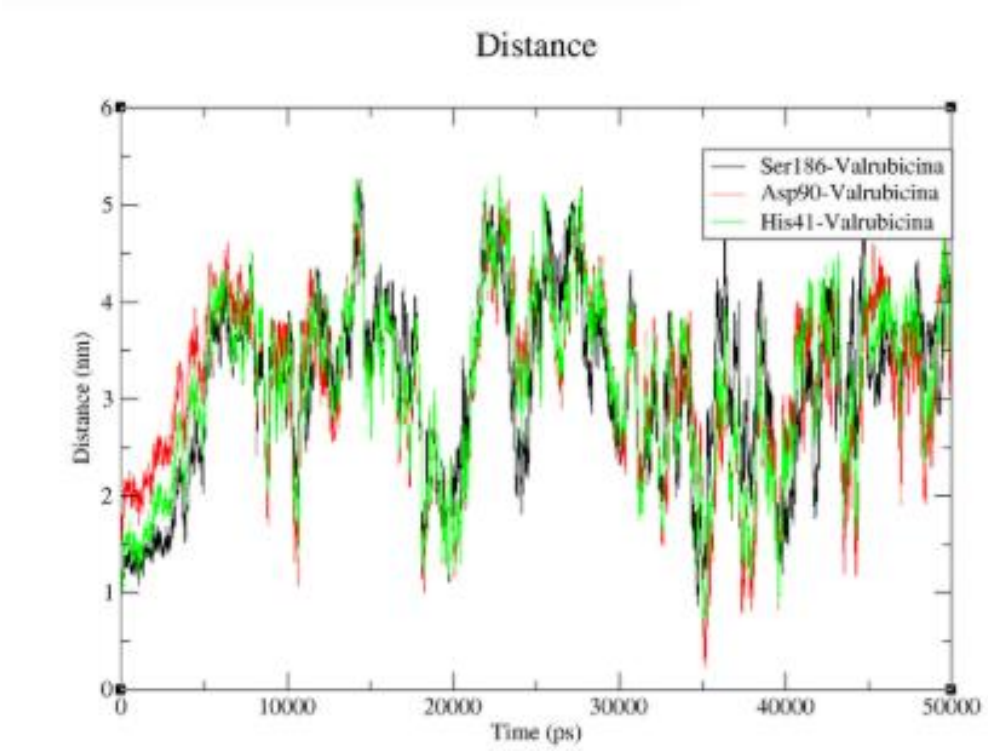
Anexo 41. Distancia promedio entre camostat y residuos catalíticos de TMPRSS2 durante 50 ns de simulación.



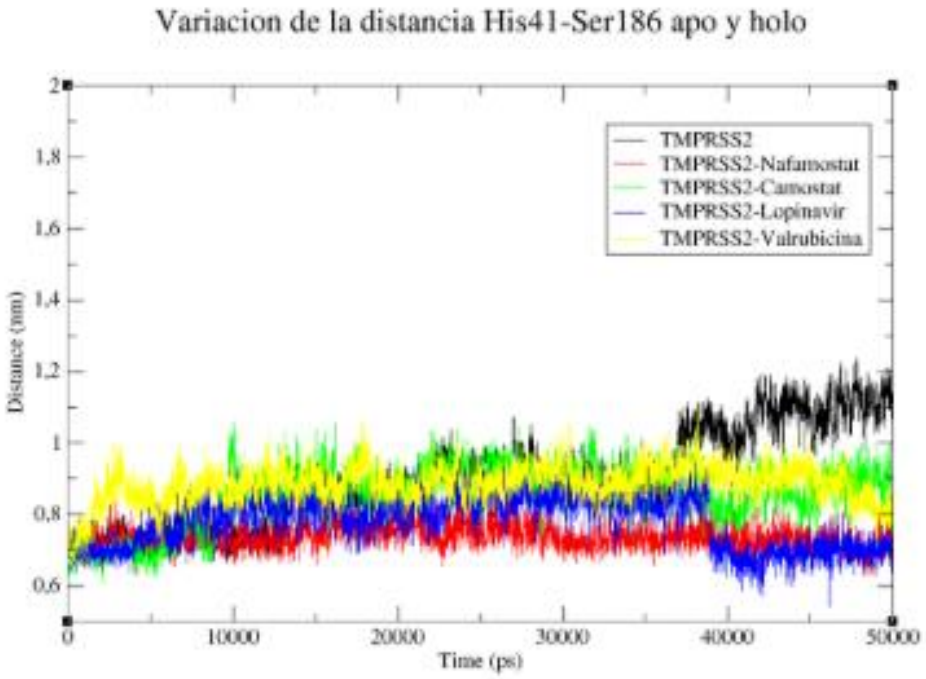
Anexo 42. Distancia promedio entre lopinavir y residuos catalíticos de TMPRSS2 durante 50 ns de simulación



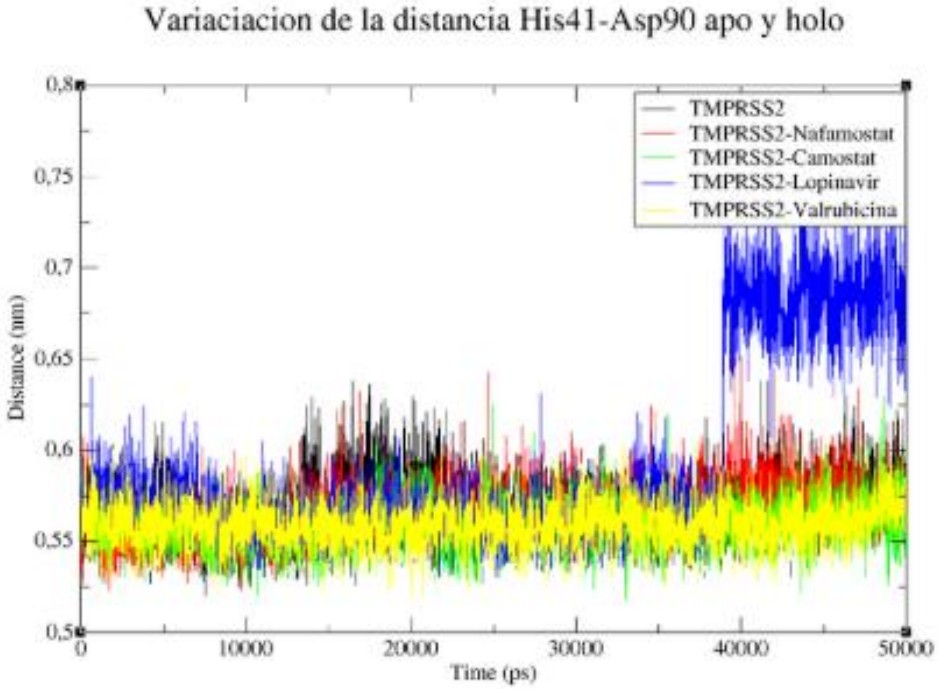
Anexo 43. Distancia promedio entre valrubicina y residuos catalíticos de TMPRSS2 durante 50 ns de simulación



Anexo 44. Variación de la distancia entre los residuos catalíticos His41 y Ser186 en TMPRSS2 libre (apo) y en los complejos con los ligandos durante 50 ns de dinámica molecular.



Anexo 45. Variación de la distancia entre los residuos catalíticos His41 y Asp90 en TMPRSS2 libre (apo) y en los complejos con los ligandos durante 50 ns de dinámica molecular



Anexo 46. Variación de la distancia entre los residuos catalíticos Asp90 y Ser186 en TMPRSS2 libre (apo) y en los complejos con los ligandos durante 50 ns de dinámica molecular

