



Desarrollo in-silico de inhibidores de PLpro del SARS-CoV-2: Un enfoque teórico para la búsqueda de nuevos antivirales

Gianfranco Paolo Sabadini Alfaro

Tesis para optar al grado de
Doctor en ciencias mención química

UNIVERSIDAD DE VALPARAÍSO
UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA

Valparaíso, Chile
2026



CONSTANCIA DE VALIDACIÓN Y CONFIDENCIALIDAD DE MONOGRAFÍA A REPOSITORIO ACADÉMICO

1.- IDENTIFICACIÓN DEL TRABAJO ACADÉMICO

Tipo de monografía (marcar una opción): ☐ Memoria o trabajo de título ☒ Tesis de Postgrado

Título del trabajo: Desarrollo in-silico de inhibidores de PLpro del SARS-CoV-2: Un enfoque teórico para la búsqueda de nuevos antivirales

Nombre del candidato: Gianfranco Paolo Sabadini Alfaro

Carrera / Grado: Doctorado en ciencias mención química

Campus: Valparaíso **Departamento:** Química

2.- VALIDACIÓN DEL PROFESOR GUÍA/DIRECTOR DE TESIS

Yo, Jaime Alberto Mella Raipan en mi calidad de profesor guía/director del trabajo académico mencionado anteriormente

DEJO CONSTANCIA que:

- He revisado esta versión del documento y corresponde a la versión final aprobada del trabajo.
- El trabajo cumple con los requisitos académicos y de formato establecidos por la institución.

3.- EVALUACIÓN DE CONFIDENCIALIDAD POR PROPIEDAD INDUSTRIAL (marcar una opción)

☒ El trabajo **NO contiene** información que amerite confidencialidad y puede ser publicado de inmediato en repositorio con acceso abierto.

El trabajo **CONTIENE** información con potenciales implicancias de propiedad industrial o intelectual y requiere un periodo de confidencialidad (**embargo**) por (**marcar una opción**):

☐ 6 meses ☐ 12 meses ☐ 2 años ☐ 3 años ☐ 5 años ☐ 10 años

Fundamentación de la necesidad de confidencialidad (obligatorio si se solicita embargo):

4.- FIRMAS

Profesor(a) guía o director(a) de memoria o tesis:

Fecha: 30 de julio de 2026 **Firma:**

Estudiante o Candidato(a):

Fecha: 30/07/2026 **Firma:**

Este formulario debe ser insertado como página 2 de la memoria o tesis, completado y firmado por estudiante y profesor(a) antes de la entrega en portal PRISMA de Biblioteca USM.

1 Contenido

2 Introducción 26

2.1 Contexto general del SARS-CoV-2..... 26

2.2 PLpro como diana terapéutica..... 29

2.2.1 Rol funcional de PLpro en el ciclo replicativo 29

2.2.2 Función inmunomoduladora de PLpro 30

2.2.3 Características estructurales clave de PLpro 31

2.2.4 Comparación entre PLpro y Mpro..... 32

2.2.5 Justificación científica de PLpro como diana farmacológica 32

2.2.6 Inhibidores conocidos y su mecanismo de acción 32

2.2.7 PLpro: similitud estructural con deubiquitinasa humana y potencial anticancerígeno 33

2.3 Química medicinal asistida por computación en la era post-COVID..... 34

2.3.1 Técnicas computacionales clave en el diseño racional de fármacos..... 35

2.3.2 Aplicaciones prácticas, optimización de recursos y casos de éxito..... 35

2.3.3 Perspectivas futuras y desafíos éticos-colaborativos 35

2.4 Rol de las técnicas *in silico* en la química medicinal asistida por computación..... 35

2.4.1 Enfoques basados en el ligando..... 36

2.4.1.1 2D-QSAR 36

Descriptores moleculares 36

Descriptores fisicoquímicos globales 37

Descriptores de conteo estructural 40

Descriptores topológicos y geométricos 40

Descriptores basados en autocorrelación..... 40

Descriptores electrotopológicos (ETA) y electrónicos derivados de electronegatividad 41

2.4.1.2 Softwares más utilizados..... 42

2.4.1.3 Generación estadística..... 42

Elección de variables 42

Construcción modelo estadística 42

Métodos de validación 43

2.4.2 Enfoques basados en el receptor..... 45

2.4.2.1 Docking molecular 46

Tipos de docking..... 46

Algoritmo general del docking molecular..... 46

Algoritmos sistemáticos 47

Algoritmos estocásticos 47

Basadas en conocimiento 47

Algoritmo Autodock 4 48

Algoritmo *Induced fit docking* de Schrodinger 2025..... 49

2.4.2.2 Virtual Screening..... 50

Base de datos de ligandos..... 51

Convertidor SDF a PDBQT 51

Autogrid4..... 51

Autodock4 GPU 52

2.4.2.3 Dinámica molecular..... 52

2.4.2.3.1 Tiempos de simulación 53

2.4.2.3.2 Mecánica clásica 54

2.4.2.3.3 Leyes de Newton 54

2.4.2.3.4 Series de Taylor..... 54

2.4.2.3.5 Campos de fuerza..... 55

Términos enlazantes 56

Términos no enlazantes 57

Ensamblajes termodinámicos..... 59

Control de temperatura y presión..... 59

Dinámica molecular en sistemas biomoleculares..... 60

Análisis de trayectorias 60

Análisis ángulos diedros..... 62

Alcances y limitaciones 62

2.5 Administración, distribución, metabolismo, excreción y toxicidad (ADME/Tox)..... 63

2.5.1 Farmacocinética y la importancia de incluirla en estrategia de desarrollo *in silico*..... 63

2.5.2 Parámetros para evaluar absorción..... 63

2.5.3 Parámetros para evaluar distribución 63

2.5.4 Parámetros para evaluar metabolismo 64

2.5.5 Parámetros para evaluar excreción 64

2.5.6 Parámetros para evaluar toxicidad 64

2.5.7 Predicción *in silico* y uso del ADMETlab 3.0 65

2.6 Hipótesis 65

3 Objetivos 65

3.1 Objetivo general 65

3.2 Objetivos específicos 65

4 Metodología 66

4.1 Generación de ecuaciones QSAR mediante selección manual 66

4.2 Generación de la Ecuación QSAR mediante Algoritmo Genético 66

4.2.1 Base de Datos y Descriptores 66

4.2.2 Configuración del Algoritmo Genético..... 66

4.2.3 Selección y Refinamiento de Modelos..... 67

4.2.3.1 Identificación de las Mejores Ecuaciones 67

4.2.3.2	Depuración de Datos y Eliminación de Valores Atípicos	67
4.2.3.3	Control de Multicolinealidad (VIF)	67
4.2.3.4	Regresión Exploratoria y Eliminación de Valores Atípicos en residuos	67
4.2.4	Construcción y Validación del Modelo QSAR	67
4.2.5	Selección de la Ecuación Final	67
4.3	Virtual Screening	68
4.4	<i>Induced fit docking</i>	68
4.5	Dinámica molecular	69
4.5.1	Protocolos de simulación.....	69
4.5.2	Análisis de trayectorias	69
4.6	Análisis retrosintético y farmacocinético.....	71
4.7	Síntesis orgánica	72
4.7.1	Protocolo general para la síntesis de chalconas en medio básico (KOH en etanol absoluto)	72
4.7.2	Síntesis de derivados Pirimido[1,2-a]benzimidazol	72
4.7.3	Síntesis de derivados de benzimidazol sustituidos con piridinas y naftalenos piridina benzimidazoles	72
4.7.3.1	Síntesis de benzimidazoles piridina en solvente acuoso o etanólico	72
4.7.3.2	Síntesis de benzimidazoles piridina en mortero	73
4.7.3.3	Síntesis de benzimidazoles piridina sustituidos NH con acetil naftalenos	73
4.7.4	Caracterización Orgánica.....	73
5	Resultados	73
5.1	QSAR de selección manual de descriptores	73
5.1.1	Pretratado de base de datos	73
5.1.2	Validación QSAR.....	78
5.1.3	Ecuación final del modelo QSAR de selección manual de descriptores	79
5.2	QSAR de selección vía algoritmo genético	79
5.2.1	Análisis por algoritmo Generación y Selección del Modelo QSAR.....	79
5.2.2	Análisis y Depuración de Ecuaciones	80
5.2.3	Ecuaciones QSAR Seleccionadas	83
5.2.4	Validación estadística	83
5.2.5	Significado de los descriptores seleccionados.....	85
5.2.6	Predicción y diseño de nuevos compuestos	85
5.2.6.1	Compuestos derivados de chalconas y heterociclos	85
5.2.6.1.1	Distribución de actividad predicha por familia estructural	86
5.2.6.1.2	Influencia del fragmento derivado de la cetona (R^1)	87
5.2.6.1.3	Influencia del fragmento derivado del aldehído (R^2)	87
5.2.6.1.4	Comparación entre los modelos QSAR.....	87

5.2.6.2	Compuestos derivados de benzimidazoles	87
5.3	<i>Virtual screening</i>	88
5.3.1	Validación molécula referencia GRL-0617	88
5.3.2	Desempeño análisis molecular	89
5.3.3	Ranking por afinidad	90
5.3.4	Comparación de GRL-0617 poses top 1	93
5.3.5	Predicción <i>Virtual screening</i> ecuaciones QSAR.....	93
5.4	Docking inducido (<i>induced fit docking</i>).....	94
5.4.1	Validación molécula referencia GRL-0617	94
5.4.2	Docking molecular de chalconas y derivados de chalconas en PLpro.....	95
5.4.3	Docking molecular de derivados de benzimidazol.....	98
5.4.4	Docking molecular de moléculas Top 1 del <i>Virtual Screening</i>	100
5.5	Dinámica molecular	102
5.5.1	Análisis comportamiento solo proteína	102
5.5.2	Análisis GRL-0617 molécula de referencia	103
5.5.3	Análisis de la familia pirimido-benzimidazol	106
5.5.4	Top 2 mejores moléculas de cada familia	132
5.5.5	Análisis derivados de benzimidazoles	150
5.5.6	Análisis moléculas Top 1 <i>Virtual screening</i>	158
5.6	Farmacocinética	167
5.6.1	Chalconas y derivados heterocíclicos	167
5.6.1.1	Drogabilidad (<i>drugability</i>).....	167
5.6.1.2	Absorción	168
5.6.1.3	Distribución.....	170
5.6.1.4	Metabolismo	172
5.6.1.5	Excreción	174
5.6.1.6	Toxicidad	176
5.6.1.7	Solubilidad	179
5.6.2	Derivados de benzimidazol.....	181
5.6.2.1	Drogabilidad (<i>drug like</i>).....	181
5.6.2.2	Absorción	182
5.6.2.3	Distribución.....	184
5.6.2.4	Metabolismo	185
5.6.2.5	Excreción	186
5.6.2.6	Toxicidad	187
5.6.2.7	Solubilidad	188
5.6.3	Moléculas top1 <i>Virtual screening</i>	189
5.6.3.1	Drogabilidad	189

5.6.3.2	Absorción	190
5.6.3.3	Distribución	191
5.6.3.4	Metabolismo	192
5.6.3.5	Excreción	193
5.6.3.6	Toxicidad	194
5.6.3.7	Solubilidad	196
5.7	Retrosíntesis	197
5.7.1	Chalconas y heterociclos derivados.....	197
5.7.1.1	Chalconas	197
5.7.1.2	Aminopirimidinas.....	198
5.7.1.3	Dihidrobenzodiazepinas	198
5.7.1.4	Dihidropirazol	198
5.7.1.5	Benzodiazepinas	199
5.7.1.6	Pirazol	199
5.7.1.7	Pirimidina-aminobenzimidazol.....	199
5.7.1.8	Pirimido-benzimidazol.....	200
5.7.2	Derivados de Benzimidazoles	200
5.8	Avances sintéticos	201
5.8.1	Chalconas	201
5.8.1.1	Estandarización protocolo de síntesis	201
5.8.1.2	Chalcona obtenidas	202
5.8.1.3	Pirimido benzimidazol	203
5.8.2	Benzimidazoles y derivados.....	207
5.8.2.1	Benzimidazoles.....	207
5.8.2.2	Benzimidazoles acilados	208
6	Discusión	208
6.1	QSAR evolutivo y de selección manual.....	208
6.1.1	Evaluación QSAR de elección de descriptores manuales.....	208
6.1.1.1	Depuración del conjunto de datos y delimitación del dominio químico	208
6.1.1.2	Robustez estadística y validez predictiva del modelo manual	209
6.1.1.3	Significado químico de los descriptores del modelo manual	209
6.1.1.4	Utilidad del modelo manual para diseño racional	210
6.1.1.5	Limitaciones del modelo manual	210
6.1.2	Evaluación QSAR de elección de descriptores de selección evolutivo	210
6.1.2.1	Interpretabilidad frente a rendimiento predictivo	210
6.1.2.2	Convergencia del algoritmo y significado del <i>plateau</i>	211
6.1.2.3	Recurrencia de descriptores y detección de señal estructural	211
6.1.2.4	Discusión química de la ecuación evolutiva seleccionada	211

6.1.2.5	Control de multicolinealidad, refinamiento y criterio de selección final.....	212
6.1.2.6	Potencial y límites del QSAR evolutivo	212
6.1.3	Comparación crítica entre QSAR manual y QSAR evolutivo	213
6.1.3.1	Interpretación vs predicción	213
6.1.3.2	Diferencias en las predicciones y calibración del espacio de actividad	213
6.1.3.3	Complementariedad de ambos enfoques	213
6.1.3.4	Integración de los modelos QSAR manual y evolutivo	213
6.2	<i>Virtual Screening</i>	214
6.2.1	Validación del protocolo de docking mediante la molécula de referencia	214
6.2.2	Robustez y alcance del espacio químico explorado.....	215
6.2.3	Consideraciones computacionales y escalabilidad del método	215
6.2.4	Análisis de afinidad: influencia estructural y coherencia con QSAR	215
6.2.5	Análisis de poses: convergencia estructural y especificidad del sitio de unión	215
6.2.6	Implicancias para el diseño racional de inhibidores	216
6.2.7	Resultados con respecto al QSAR.....	217
6.2.8	Limitaciones del estudio	217
6.3	Docking inducido (<i>Induced fit docking</i>).....	217
6.3.1	Análisis de derivados de chalconas	217
6.3.2	Análisis de derivados de benzimidazol	219
6.3.3	Análisis de derivados de top 1 moléculas del <i>virtual screening</i>	219
6.3.4	Limitaciones asociadas al <i>induced fit docking</i>	220
6.4	Dinámica molecular	220
6.4.1	Discusión de la Dinámica molecular del sistema APO y del complejo con GRL-0617.....	220
6.4.2	Discusión de la Dinámica molecular de la familia pirimido-benzimidazol	222
6.4.3	Discusión de la Dinámica molecular del Top 2 de las mejores moléculas derivadas de chalconas.....	224
6.4.4	Discusión de la Dinámica molecular derivados de benzimidazoles	227
6.4.5	Discusión de la Dinámica molécula top 1 del <i>virtual screening</i>	228
6.4.6	Integración de moléculas favorables	228
6.5	Farmacocinética	229
6.5.1	Farmacocinética de derivados de chalconas	229
6.5.1.1	Absorción y distribución	229
6.5.1.2	Metabolismo y excreción	229
6.5.1.3	Toxicidad	230
6.5.1.4	Solubilidad	230
6.5.1.5	Integración farmacocinética con dinámica molécula.....	231
6.5.2	Farmacocinética de derivados de benzimidazoles	232
6.5.2.1	Absorción y distribución	232

6.5.2.2	Metabolismo y excreción	233
6.5.2.3	Toxicidad	233
6.5.2.4	Solubilidad	234
6.5.3	Integración de resultados de farmacodinámica predicha de benzimidazoles y derivados de chalconas.....	234
6.5.4	Farmacocinética de moléculas top 1 del <i>virtual screening</i>	234
6.5.4.1	Absorción y distribución	234
6.5.4.2	Metabolismo y excreción	235
6.5.4.3	Toxicidad	236
6.5.4.4	Solubilidad	238
6.5.4.5	Integración de moléculas top 1 del <i>virtual screening</i>	238
6.6	Retrosíntesis	240
6.6.1	Derivados de chalconas.....	240
6.6.2	Derivados de benzimidazoles	242
6.7	Avances sintéticos	242
6.7.1	Chalconas	243
6.7.2	2 Fenil 4 (4-trifluorometanofenil)Pirimido-benzimidazol	243
6.7.2.1	Punto de fusión y infrarrojo	244
6.7.2.2	Asignación ¹ H RMN	245
6.7.2.3	Asignación ¹³ C RMN	246
6.7.2.4	2D HMBC.....	247
6.7.2.5	2D COSY	248
6.7.2.6	2D NOESY	249
6.7.2.7	Asignaciones finales	250
6.7.3	Benzimidazoles	252
7	Conclusión	253
8	Proyecciones	255
9	Referencias.....	256
10	Anexo.....	264

Índice de figuras

Figura 1 Estructura y organización genómica del SARS-CoV-2.(A) Esquema estructural del virión. (B) Organización del genoma y proteínas virales codificadas, con funciones conocidas y funciones aún no determinadas⁷. 27

Figura 2 Ciclo replicativo del SARS-CoV-2. La unión de la proteína S al receptor ACE2 permite la entrada viral mediada por proteasas celulares. Tras la liberación del ARN viral, se sintetizan las poliproteínas pp1a/pp1ab, que son escindidas por Mpro y PLpro para generar las nsps que forman el complejo replicasa–transcriptasa. Este complejo dirige la replicación del genoma y la transcripción de ARN subgenómicos. Las proteínas estructurales se ensamblan con el ARN genómico en el ERGIC y los nuevos viriones son liberados por exocitosis¹³. 30

Figura 3 Rol dual de la proteasa PLpro de SARS-CoV-2 en la replicación viral y la evasión de la respuesta inmune innata. PLpro procesa la poliproteína viral para generar proteínas no estructurales maduras necesarias para la formación de un complejo replicasa funcional. Además, esta proteasa interfiere con las defensas antivirales del hospedero al eliminar ISG15 de proteínas celulares, incluida IRF3, provocando la desregulación de la señalización inmune innata¹². 31

Figura 4 a) Estructura global de PLpro (entrada PDB 7JIR) coloreada por dominios, con características estructurales relevantes indicadas. b) Interacciones entre PLpro y el inhibidor no covalente GRL-0617, destacando los residuos clave y los subsitios de unión del sustrato¹⁵. 32

Figura 5 Base molecular de la inhibición de PLpro por GRL-0617. 33

Figura 6 Comparación estructural entre PLpro de SARS-CoV-2 y la deubiquitinasa humana USP12. (A) Modelo estructural de PLpro que muestra su organización secundaria y la arquitectura tipo “thumb–palm–fingers(pulgar,palma,dedos)”, con los dominios señalados. Se destacan los residuos del sitio catalítico Cys111, His272 y Asp286, así como el ion zinc estructural.(B) Estructura de la deubiquitinasa humana USP12 (PDB: 5K1B) que exhibe un plegamiento y organización de dominios similares. Se indican los residuos catalíticos Cys48, His317 y Asp333 y el ion zinc¹⁸. 34

Figura 7 Esquema del flujo de trabajo del descubrimiento de fármacos asistido por computadora. El proceso integra bases de datos químicas, modelado QSAR y aprendizaje automático, docking molecular, simulaciones de dinámica molecular, selección de hits y validación experimental para priorizar candidatos y optimizar recursos en etapas tempranas del diseño de fármacos.(Esquema hecho con inteligencia artificial) 34

Figura 8 Representación esquemática de la relación entre la hidrofobicidad (logP) y la actividad biológica. El término lineal logP describe la tendencia general de la actividad con la lipofilia, mientras que el término cuadrático (logP)² introduce la curvatura necesaria para reflejar un comportamiento no lineal. La combinación de ambos términos genera un modelo más realista que predice la existencia de un intervalo óptimo de lipofilia asociado a máxima actividad biológica. Las regiones sombreadas ilustran zonas de hidrofobicidad desfavorable por exceso de hidrofilia o lipofilia. 38

Figura 9 Representación esquemática del Algoritmo Genético Lamarckiano (LGA) empleado en AutoDock⁵⁸. 49

Figura 10 Diagrama de flujo del algoritmo de docking inducido⁵⁹. 50

Figura 11 Etapas preparatorias y de producción en una simulación de dinámica molecular biomolecular.(Imagen generada con inteligencia artificial).	53
Figura 12 Escala de fenómenos moleculares y tiempo en que pueden observarse ⁶⁶	53
Figura 13 A) Figura en 2D de GRL-0617 B) Figura en 3D de GRL-0617.	68
Figura 14 Distancia entre los carbonos alfa de los aminoácidos C111 y H272, aminoácidos catalíticos. 70	
Figura 15 Distancias entre carbono alfa de :Tyr268-Arg166 y Tyr268-Pro299, aminoácidos que miden la movilidad del Loop BL2.....	70
Figura 16 Representación bidimensional de ángulos diedros a evaluar.	71
Figura 17 Representación tridimensional de ángulos diedros a evaluar	71
Figura 18 Gráficos de la actividad experimental frente a la actividad predicha para el conjunto de entrenamiento (A) y el conjunto de prueba (B). Gráfico de distribución de residuos (C). Los círculos negros representan el training set y los cuadrados rojos el test set. Todos los compuestos se encuentran dentro de una unidad logarítmica de residuo (residuo = $\text{pIC}_{50} \text{ exp} - \text{pIC}_{50} \text{ pred}$) ¹¹⁸	78
Figura 19 Evolución del coeficiente de determinación máximo (R^2) a lo largo de las generaciones del algoritmo genético. Se destacan los diez modelos con mayor R^2 obtenidos durante la simulación, considerados candidatos a ecuaciones predictivas finales. La estabilización del R^2 en las últimas generaciones indica la convergencia del algoritmo dentro del límite de 1.000.000 de generaciones.	80
Figura 20 Mapa de calor que muestra la frecuencia de selección de los descriptores moleculares durante la ejecución del algoritmo genético. Se evaluaron 1533 descriptores derivados de los 1875 calculados inicialmente con PaDEL. El color azul indica la presencia del descriptor en los individuos seleccionados y el rojo su ausencia a lo largo de las generaciones. El patrón observado evidencia la estabilización del proceso evolutivo y la convergencia del modelo.	80
Figura 21 Moléculas con estructuras atípicas	81
Figura 22 Distribuciones de valores de descriptores por cada molécula. Cada punto en rojo y su número representa el número de la molécula.....	82
Figura 23 Gráficos de dispersión entre la actividad biológica experimental ($\text{pIC}_{50} \text{ real}$) y la actividad predicha por los modelos QSAR seleccionados (ecuaciones 1, 3 y 4). La línea negra representa la relación ideal ($y = x$), mientras que las líneas azules y rojas delimitan los intervalos de error $\pm 1\sigma$, $\pm 1.5\sigma$ y $\pm 2\sigma$. La distribución de los puntos alrededor de la línea ideal evidencia la capacidad predictiva de los modelos y la ausencia de desviaciones sistemáticas significativas.....	83
Figura 24 Evaluación gráfica del desempeño predictivo de la Ecuación 4. (A) Relación entre los valores experimentales y predichos de pIC_{50} para el conjunto de entrenamiento, mostrando un ajuste lineal adecuado. (B) Relación entre valores experimentales y predichos para el conjunto de prueba externo, evidenciando una buena capacidad predictiva del modelo. (C) Gráfico de dispersión de los residuos en función del pIC_{50} experimental, donde se observa una distribución aleatoria alrededor de cero sin tendencias sistemáticas, lo que confirma la ausencia de sesgos y la robustez estadística del modelo. .	85
Figura 25 Familias estructurales evaluadas en el estudio. Las estructuras se generaron a partir de chalconas precursoras mediante transformaciones sintéticas que permiten acceder a diferentes sistemas heterocíclicos.	86
Figura 26 Estrategia de generación de la biblioteca molecular utilizada para las predicciones QSAR. La combinación de cinco fragmentos derivados de cetonas (R^1) y ocho fragmentos derivados de aldehídos (R^2) permitió generar 40 chalconas precursoras, las cuales	86
Figura 27 Esquema estructuras de benzimidazoles evaluados	88
Figura 28 Superposición estructural entre la conformación cristalográfica de PLpro en complejo con GRL-0617 y la pose obtenida mediante docking rígido.(A) Estructura cristalográfica del complejo PLpro–GRL-0617.(B) Pose mejor clasificada obtenida por docking rígido.(C) Superposición de ambas estructuras, donde la proteína cristalográfica se muestra en gris y la estructura obtenida por docking en verde.(D) Vista ampliada del sitio activo mostrando la coincidencia espacial entre el ligando cristalográfico (amarillo) y la pose dockeada (verde).	89

Figura 29 Núcleos heterocíclicos presentes en el análisis de virtual screening.....	89
Figura 30 Estructura del top 10 de cada familia molecular.	92
Figura 31 Distribución de afinidad de las moléculas Top 10 por familia química. La línea discontinua indica la afinidad del ligando de referencia GRL-0617.	93
Figura 32 A) Vista general del sitio de unión alostérico de PLpro mostrando la superposición de las poses Top 1 obtenidas para los compuestos seleccionados dentro del bolsillo de unión. (B) Aminopirimidinas.(C) Benzimidazoles.(D) Benzodiazepinas.(E) Benzamidas.(F) Benzotiofenos.(G) Pirimidina-aminobenzimidazoles.(H) Pirimido-benzimidazoles.(I) Ligando de referencia GRL-0617.En los paneles (B–H), cada ligando se muestra superpuesto con el ligando de referencia GRL-0617, representado en color gris translúcido, con el fin de facilitar la comparación espacial y evaluar la coincidencia en el modo de unión dentro del sitio activo.	93
Figura 33 Validación del protocolo de Induced Fit Docking utilizando el inhibidor de referencia GRL-0617. (A) Estructura del complejo PLpro–GRL-0617 obtenida a partir de la estructura cristalográfica, mostrando el ligando en el sitio de unión.(B) Pose de GRL-0617 obtenida mediante Induced Fit Docking, representada dentro del sitio de unión de PLpro. (C) Superposición entre la estructura cristalográfica (gris) y la estructura obtenida mediante Induced Fit Docking (verde), mostrando la coincidencia entre ambas conformaciones del complejo proteína–ligando.(D) Vista ampliada del sitio de unión mostrando la superposición de la pose cristalográfica y la pose obtenida mediante Induced Fit Docking para el ligando GRL-0617.	94
Figura 34 Comparación de las interacciones ligando–proteína para GRL-0617 en la estructura cristalográfica y en la pose obtenida mediante Induced Fit Docking.(A) Diagrama de interacción bidimensional correspondiente al complejo cristalográfico PLpro–GRL-0617, mostrando los residuos del sitio de unión que interactúan con el ligando.(B) Diagrama de interacción bidimensional correspondiente a la pose de GRL-0617 obtenida mediante Induced Fit Docking, donde se observan los residuos del bolsillo de unión involucrados en interacciones con el ligando.	95
Figura 35 Distribución global de chalconas y derivados de chalconas con respecto al XP GScore la línea de referencia es el ligando GRL-0617 con su puntaje de XP GScore :-5.5720.	95
Figura 36 Distribución de los valores de XP GScore obtenidos mediante docking molecular para las diferentes familias estructurales evaluadas. Cada punto representa el valor correspondiente a la mejor pose de unión de una molécula individual. La línea punteada indica el valor de XP GScore del ligando de referencia GRL-0617.	96
Figura 37 Distribución de los valores de XP GScore obtenidos mediante docking molecular en función del fragmento R ¹ derivado de la cetona utilizada en la condensación de Claisen–Schmidt. Cada punto representa el valor correspondiente a la mejor pose de unión de una molécula individual. La línea punteada indica el valor de XP GScore del ligando de referencia GRL-0617.	96
Figura 38 Distribución de los valores de XP GScore obtenidos mediante docking molecular en función del fragmento R ² derivado del aldehído utilizado en la condensación de Claisen–Schmidt. Cada punto representa el valor correspondiente a la mejor pose de unión de una molécula individual. La línea punteada indica el valor de XP GScore del ligando de referencia GRL-0617.	97
Figura 39 Mapas de calor de la mediana de XP GScore obtenida para las distintas combinaciones estructurales de la biblioteca evaluada frente al sitio alostérico de PLpro. El panel A muestra la distribución de los valores medianos de XP GScore en función de la familia estructural y del sustituyente R ² derivado del aldehído, el panel B presenta el efecto del fragmento R ¹ derivado de la cetona sobre las diferentes familias evaluadas y el panel C resume el efecto combinado de los fragmentos R ¹ y R ² independientemente del núcleo heterocíclico. Los colores representan la mediana del XP GScore calculada para cada combinación estructural. Con el fin de mejorar la legibilidad de la figura se utilizaron abreviaciones para los fragmentos aromáticos, donde Pireno corresponde al derivado de acetilpireno, 9-Antraceno al derivado de acetilantraceno, 1-Naftilo y 2-Naftilo a los derivados de 1- y 2-acetilnaftaleno respectivamente y Fenilo al derivado de acetofenona. De manera análoga, los fragmentos R ² se	

representan como p-Nitro para p-nitrofenilo, p-DMA para p-dimetilaminofenilo, Vainillina para el fragmento 4-hidroxi-3-metoxifenilo, Cinamilo para el sustituyente cinamilo, o-Hidroxifenilo para el derivado orto-hidroxifenilo, p-Isopropil para p-isopropilfenilo, p-Amino para p-aminofenilo y p-CF ₃ para p-trifluorometilfenilo.	98
Figura 40 Superposición de las poses del acoplamiento molecular de los derivados benzimidazol dentro del sitio activo de PLpro. En cada panel se muestra la pose obtenida por docking (representada en barras) superpuesta con la pose cristalográfica del ligando de referencia del complejo PLpro–GRL-0617 (PDB: 7CJM), representada en transparencia. Los paneles corresponden a los compuestos evaluados: (A) benzimidazol 1, (B) benzimidazol 2, (C) benzimidazol 3, (D) benzimidazol 4, (E) benzimidazol 5 y (F) benzimidazol 6. La proteína se muestra en representación de cintas coloreadas para cada sistema, permitiendo visualizar la orientación relativa de los ligandos dentro del bolsillo de unión.....	99
Figura 41. Los diagramas de interacción 2D muestran que todos los derivados benzimidazol se acomodan dentro del bolsillo de unión de PLpro. En general, los compuestos presentan principalmente interacciones hidrofóbicas, junto con contactos específicos de tipo π – π stacking y puentes de hidrógeno. Entre los derivados evaluados, benzimidazol 1 presenta una interacción por puente de hidrógeno, mientras que benzimidazol 5 muestra contactos compatibles con π – π stacking asociados al sistema aromático del ligando. En contraste, benzimidazol 2 y benzimidazol 3 exhiben un patrón de interacción más limitado, consistente con los valores de docking menos favorables observados para estos compuestos.....	100
Figura 42 Diagramas bidimensionales de interacción proteína–ligando para los complejos analizados. Los paneles muestran las interacciones del sitio de unión entre la proteína y los distintos compuestos: (A) aminopirimidina, (B) benzimidazol, (C) benzamida, (D) benzodiazepina, (E) benzotiofeno, (F) amino-pirimido-benzimidazol, (G) pirimido-benzimidazol y (H) GRL-0617.	101
Figura 43 Comparación estructural de los complejos proteína–ligando representativos para las distintas familias químicas analizadas. (A) Superposición global de las conformaciones proteicas en presencia de los distintos ligandos. La proteína se muestra solo el backbone y cada complejo se distingue por un color: aminopiridina (beige), benzimidazol (lila claro), benzodiazepina (celeste), benzotiofeno (salmón), GRL-0617 (blanco), amino-pirimido-benzimidazol (rosado intenso), pirimido-benzimidazol (amarillo) y benzamida (verde). (B) Complejo con aminopiridina. (C) Complejo con benzimidazol. (D) Complejo con benzamida. (E). Complejo con benzodiazepina (F) Complejo con benzotiofeno (G) Complejo con amino-pirimido-benzimidazol. (H) Complejo con pirimido-benzimidazol. (I) Complejo con GRL-0617. En todos los paneles, los ligandos se muestran en representación de sticks, con los heteroátomos coloreados según la convención estándar, también está incluido la posición del ligando de referencia.	101
Figura 44 Evolución temporal del RMSD de la proteína durante la simulación de dinámica molecular del sistema APO. Las líneas grises corresponden a las tres réplicas independientes de la simulación, mientras que la línea negra representa el promedio de las trayectorias. El eje x muestra el tiempo de simulación (ns) y el eje y el RMSD (Å) calculado respecto a la estructura inicial.	102
Figura 45 Fluctuación cuadrática media por residuo (RMSF) de la proteína durante la simulación de dinámica molecular del sistema APO. Las líneas grises corresponden a las tres réplicas independientes de la simulación, mientras que la línea negra representa el promedio de las trayectorias. El eje x indica el número de residuo, mientras que el eje y muestra el RMSF (Å), reflejando la flexibilidad local de cada región de la proteína a lo largo de la simulación.....	102
Figura 46 Evolución temporal de distancias estructurales clave durante la simulación de dinámica molecular del sistema APO. El panel izquierdo muestra la distancia catalítica entre Cys111 y His272, correspondiente a los residuos del sitio activo. Los paneles central y derecho muestran las distancias Arg166–Tyr268 y Pro299–Tyr268, respectivamente, utilizadas como descriptores del movimiento del loop BL2. Las líneas grises representan las tres réplicas independientes de la simulación, mientras que la línea negra corresponde al promedio de las trayectorias. El eje x indica el tiempo de simulación (ns) y el eje y la distancia entre los residuos (Å).....	103

Figura 47 Distribución angular de los ángulos diedros χ_1 y χ_2 del residuo Tyr268 durante la simulación de dinámica molecular del sistema APO. Los diagramas polares muestran la frecuencia de aparición de cada conformación angular a lo largo de la trayectoria. Cada barra corresponde a un intervalo angular de 5°, mientras que los círculos concéntricos representan la frecuencia relativa, marcados en incrementos de 10 unidades.....	103
Figura 48 RMSD de la proteína con respecto al frame en tiempo 0 a lo largo del tiempo de simulación 700ns	104
Figura 49 RMSD del ligando con respecto a la proteína a lo largo de la simulación 700ns.....	104
Figura 50 Fluctuación cuadrática media por residuo (RMSF) de la proteína con GRL-0617 durante la simulación. La línea azul representa las trayectorias. El eje x indica el número de residuo, mientras que el eje y muestra el RMSF (Å), reflejando la flexibilidad local de cada región de la proteína a lo largo de la simulación.....	105
Figura 51 Comparación de la fluctuación cuadrática media por residuo (RMSF) entre el sistema APO y el complejo PLpro– GRL-0617 durante la simulación de dinámica molecular. El panel izquierdo muestra el perfil completo de RMSF a lo largo de la proteína, mientras que el panel derecho presenta una ampliación de la región correspondiente al loop BL2. La línea negra representa el sistema APO (promedio de las réplicas) y la línea azul el sistema GRL-0617. El eje x indica el número de residuo y el eje y el RMSF (Å), permitiendo comparar la flexibilidad local de la proteína en ausencia y presencia del ligando.....	105
Figura 52 Evolución temporal de distancias estructurales clave durante las simulaciones de dinámica molecular para los sistemas APO y PLpro–GRL-0617. El panel izquierdo muestra la distancia catalítica entre Cys111 y His272. Los paneles central y derecho corresponden a las distancias Arg166–Tyr268 y Pro299–Tyr268, respectivamente, utilizadas como descriptores del movimiento del loop BL2. La línea negra representa el sistema APO (promedio de las réplicas) y la línea azul el complejo GRL-0617. El eje x indica el tiempo de simulación (ns) y el eje y la distancia entre residuos (Å).....	106
Figura 53 Distribución angular de los ángulos diedros χ_1 y χ_2 del residuo Tyr268 durante las simulaciones de dinámica molecular de los sistemas APO y PLpro–GRL-0617. Los diagramas polares muestran la frecuencia de las conformaciones angulares observadas a lo largo de la trayectoria. Los paneles superiores corresponden al sistema APO (promedio de las réplicas) y los inferiores al complejo GRL-0617. Cada barra representa un intervalo angular de 5°, mientras que los círculos concéntricos indican la frecuencia relativa de ocurrencia.	106
Figura 54 Evolución temporal del RMSD de la proteína durante las simulaciones de dinámica molecular para los sistemas APO, PLpro–GRL-0617 y los distintos derivados bencénicos evaluados. Cada panel corresponde a un ligando diferente (cinamilo, 4-OH-3-OMe, 4-NMe ₂ , 4-CF ₃ , 4-NH ₂ , 4-NO ₂ , 4-iPr y 2-OH). La línea negra representa el sistema APO (promedio de las réplicas), la línea azul el complejo PLpro–GRL-0617, y la línea de color corresponde al sistema con el ligando evaluado. El eje x indica el tiempo de simulación (ns) y el eje y el RMSD de la proteína (Å) respecto a la estructura inicial.	107
Figura 55 Evolución temporal del RMSD de la proteína durante las simulaciones de dinámica molecular para los sistemas APO, PLpro–GRL-0617 y los derivados 1-naftílicos evaluados.	107
Figura 56 Evolución temporal del RMSD de la proteína durante las simulaciones de dinámica molecular para los sistemas APO, PLpro–GRL-0617 y los derivados 2-naftílicos evaluados.	108
Figura 57 Evolución temporal del RMSD de la proteína durante las simulaciones de dinámica molecular para los sistemas APO, PLpro–GRL-0617 y los derivados de antraceno evaluados.	108
Figura 58 Evolución temporal del RMSD de la proteína durante las simulaciones de dinámica molecular para los sistemas APO, PLpro–GRL-0617 y los derivados de pireno evaluados.	109
Figura 59 Evolución temporal del RMSD del ligando durante las simulaciones de dinámica molecular para los distintos derivados bencénicos evaluados en el sitio de unión de PLpro. Cada panel corresponde a un ligando diferente (cinamilo, 4-OH-3-OMe, 4-NMe ₂ , 4-CF ₃ , 4-NH ₂ , 4-NO ₂ , 4-iPr y 2-OH). La línea azul representa el ligando GRL-0617 utilizado como referencia, mientras que la línea de color corresponde al ligando evaluado.	110

Figura 60 RMSD del ligando durante las simulaciones de dinámica molecular para los derivados 1 naftílicos evaluados. 110

Figura 61 RMSD del ligando durante las simulaciones de dinámica molecular para los derivados 2 naftílicos evaluados. 111

Figura 62 RMSD del ligando durante las simulaciones de dinámica molecular para los derivados de antraceno evaluados. 111

Figura 63 RMSD del ligando durante las simulaciones de dinámica molecular para los derivados de pireno evaluados. 112

Figura 64 RMSF por residuo de la proteína durante las simulaciones de Dinámica molecular de los complejos PLpro–derivados bencénicos. Cada panel muestra el perfil de flexibilidad de la proteína para un ligando diferente. La región sombreada indica el loop BL2 (residuos 255–275), asociado con el control del acceso al bolsillo de unión. 113

Figura 65 RMSF por residuo de la proteína para los derivados 1-naftílicos evaluados. La región correspondiente al loop BL2 (255–275). 113

Figura 66 RMSF por residuo de la proteína para los derivados 2-naftílicos evaluados. La región correspondiente al loop BL2 (255–275). 114

Figura 67 RMSF por residuo de la proteína para los derivados de antraceno evaluados. La región correspondiente al loop BL2 (255–275). 114

Figura 68 RMSF por residuo de la proteína para los derivados de pireno evaluados. La región correspondiente al loop BL2 (255–275). 115

Figura 69 Zoom del perfil de RMSF en el loop BL2 (residuos 255–275) durante las simulaciones de dinámica molecular de los complejos PLpro–derivados bencénicos. Se incluyen los sistemas APO, GRL-0617 y los ligandos evaluados, permitiendo comparar las variaciones de flexibilidad en esta región estructural clave del sitio de unión. 115

Figura 70 Zoom del RMSF en la región BL2 (255–275) para los derivados de 1-naftilo evaluados. 116

Figura 71 Zoom del RMSF en la región BL2 (255–275) para los derivados de 2-naftilo evaluados. 116

Figura 72 Zoom del RMSF en la región BL2 (255–275) para los derivados de antraceno evaluados. 117

Figura 73 Zoom del RMSF en la región BL2 (255–275) para los derivados de pireno evaluados. 117

Figura 74 Evolución temporal de la distancia Cys111–His272 (CA–CA) durante las simulaciones de Dinámica molecular. Cada panel muestra la trayectoria de la distancia catalítica para uno de los compuestos evaluados. La línea negra corresponde al promedio de las tres réplicas de la proteína APO, mientras que la línea azul representa el complejo con el inhibidor de referencia GRL-0617. Las líneas de colores corresponden a los distintos derivados de acetofenona evaluados con sus respectivo derivados de aldehído (Cinamilo, 4-OH-3-OMe, 4-NMe₂, 4-CF₃, 4-NH₂, 4-NO₂, 4-iPr y 2-OH). En todos los sistemas, la distancia entre los residuos catalíticos se mantiene relativamente constante a lo largo de la trayectoria de 700 ns, indicando la conservación de la geometría del sitio catalítico de PLpro durante la simulación. 118

Figura 75 Evolución temporal de la distancia Cys111–His272 (CA–CA) para los derivados con sustituyente 1-naftilo. 118

Figura 76 Evolución temporal de la distancia Cys111–His272 (CA–CA) para los derivados con sustituyente 2-naftilo. 119

Figura 77 Evolución temporal de la distancia Cys111–His272 (CA–CA) para los derivados con sustituyente antraceno. 119

Figura 78 Evolución temporal de la distancia Cys111–His272 (CA–CA) para los derivados con sustituyente pireno. 120

Figura 79 Evolución temporal de la distancia Tyr268–Arg166(CA–CA) para los derivados con sustituyente benceno. 120

Figura 80 Evolución temporal de la distancia Tyr268–Arg166(CA–CA) para los derivados con sustituyente 1-naftilo. 121

Figura 81 Evolución temporal de la distancia Tyr268-Arg166(CA–CA) para los derivados con sustituyente 2-naftilo.	121
Figura 82 Evolución temporal de la distancia Tyr268-Arg166(CA–CA) para los derivados con sustituyente antraceno.	122
Figura 83 Evolución temporal de la distancia Tyr268-Arg166(CA–CA) para los derivados con sustituyente pireno.	122
Figura 84 Evolución temporal de la distancia Tyr268-Pro299(CA–CA) para los derivados con sustituyente benceno. La definición de las líneas y condiciones de simulación es la misma que en la Figura 68.	123
Figura 85 Evolución temporal de la distancia Tyr268-Pro299(CA–CA) para los derivados con sustituyente 1-naftilo. La definición de las líneas y condiciones de simulación es la misma que en la Figura 68.	123
Figura 86 Evolución temporal de la distancia Tyr268-Pro299 (CA–CA) para los derivados con sustituyente 2-naftilo. La definición de las líneas y condiciones de simulación es la misma que en la Figura 68.	124
Figura 87 Evolución temporal de la distancia Tyr268-Pro299(CA–CA) para los derivados con sustituyente antraceno. La definición de las líneas y condiciones de simulación es la misma que en la Figura 68. ..	124
Figura 88 Evolución temporal de la distancia Tyr268-Pro299(CA–CA) para los derivados con sustituyente pireno. La definición de las líneas y condiciones de simulación es la misma que en la Figura 68.	125
Figura 89 Distribución polar del ángulo diedro χ_1 del residuo Tyr268 durante las simulaciones de Dinámica molecular para los complejos con derivados de la familia benceno. Los paneles superiores muestran la distribución para la proteína APO (promedio de tres réplicas) y para el complejo con el inhibidor de referencia GRL-0617, mientras que los paneles restantes corresponden a los distintos derivados evaluados (Cinamilo, 4-OH-3-OMe, 4-NMe ₂ , 4-CF ₃ , 4-NH ₂ , 4-NO ₂ , 4-iPr y 2-OH). La distribución angular se representa en coordenadas polares, donde el radio corresponde a la frecuencia de observación del ángulo durante la trayectoria de Dinámica molecular de 700 ns.	127
Figura 90 Distribución polar del ángulo diedro χ_2 del residuo Tyr268 durante las simulaciones de Dinámica molecular para los complejos con derivados de la familia benceno. La organización de los paneles y la definición de las curvas corresponden a las descritas en la Figura 83.	127
Figura 91 Distribución polar del ángulo diedro χ_1 del residuo Tyr268 durante las simulaciones de Dinámica molecular para los complejos con derivados de la familia 1-naftilo La organización de los paneles y la definición de las curvas corresponden a las descritas en la Figura 83.	128
Figura 92 Distribución polar del ángulo diedro χ_2 del residuo Tyr268 durante las simulaciones de Dinámica molecular para los complejos con derivados de la familia 1-naftilo. La organización de los paneles y la definición de las curvas corresponden a las descritas en la Figura 83.	128
Figura 93 Distribución polar del ángulo diedro χ_1 del residuo Tyr268 durante las simulaciones de Dinámica molecular para los complejos con derivados de la familia 2-naftilo La organización de los paneles y la definición de las curvas corresponden a las descritas en la Figura 83.	129
Figura 94 Distribución polar del ángulo diedro χ_2 del residuo Tyr268 durante las simulaciones de Dinámica molecular para los complejos con derivados de la familia 2-naftilo. La organización de los paneles y la definición de las curvas corresponden a las descritas en la Figura 83.	129
Figura 95 Distribución polar del ángulo diedro χ_1 del residuo Tyr268 durante las simulaciones de Dinámica molecular para los complejos con derivados de la familia antraceno. La organización de los paneles y la definición de las curvas corresponden a las descritas en la Figura 83.	130
Figura 96 Distribución polar del ángulo diedro χ_2 del residuo Tyr268 durante las simulaciones de Dinámica molecular para los complejos con derivados de la familia antraceno. La organización de los paneles y la definición de las curvas corresponden a las descritas en la Figura 83.	130
Figura 97 Distribución polar del ángulo diedro χ_1 del residuo Tyr268 durante las simulaciones de Dinámica molecular para los complejos con derivados de la familia pireno. La organización de los paneles y la definición de las curvas corresponden a las descritas en la Figura 83.	131

Figura 98 Distribución polar del ángulo diedro χ_2 del residuo Tyr268 durante las simulaciones de Dinámica molecular para los complejos con derivados de la familia pireno. La organización de los paneles y la definición de las curvas corresponden a las descritas en la Figura 83. 131

Figura 99 Estructuras representativas de cada familia junto con sus dos compuestos mejor posicionados según el puntaje de docking (XP GScore). 132

Figura 100 RMSD de la proteína a lo largo de las simulaciones de Dinámica molecular para los complejos con los ligandos evaluados de las cinco familias estructurales. La línea negra representa el sistema APO promedio, la línea azul el complejo con el inhibidor de referencia GRL-0617 y la línea verde el complejo con el ligando analizado en cada panel. 134

Figura 101 RMSD del ligando durante las simulaciones de Dinámica molecular para los complejos correspondientes a la familia pirazol. La línea azul representa el inhibidor de referencia GRL-0617 y la línea verde el ligando evaluado en cada panel. Los gráficos muestran la evolución temporal de la desviación estructural del ligando respecto a su conformación inicial durante la trayectoria de simulación..... 136

Figura 102 RMSF de la proteína durante las simulaciones de dinámica molecular para los complejos correspondientes a la familia pirazol. La línea negra corresponde al sistema APO promedio, la línea azul al complejo con el inhibidor de referencia GRL-0617 y la línea verde al complejo proteína-ligando correspondiente a cada derivado evaluado. Sombreado en gris correspondiente a loop BL2 (255–275). 138

Figura 103 RMSF de los residuos correspondientes al loop BL2 (255–275) durante las simulaciones de Dinámica molecular para los complejos de la familia pirazol. La línea negra representa el sistema APO promedio, la línea azul el complejo con el inhibidor de referencia GRL-0617 y la línea verde el complejo con el ligando evaluado en cada panel..... 140

Figura 104 C111–H272, Tyr268–Arg166 y Tyr268–Pro299 durante las simulaciones de dinámica molecular para los complejos proteína-ligando de cada familia. Negro: sistema APO promedio; azul: complejo con GRL-0617; verde: complejo con el ligando evaluado. 146

Figura 105 Distribución de los ángulos diedros χ_1 y χ_2 del residuo Tyr268 durante las simulaciones de Dinámica molecular. Los histogramas polares representan la frecuencia de los valores angulares observados a lo largo de la trayectoria. La escala radial corresponde al número de frames de la simulación..... 150

Figura 106 RMSD de la proteína durante las simulaciones de Dinámica molecular. Trayectorias de RMSD para la proteína en estado apo, el complejo con el ligando de referencia GRL-0617 y los complejos con los derivados benzimidazol. El panel final muestra la superposición de todas las trayectorias para facilitar la comparación entre sistemas. 151

Figura 107 RMSD de los ligandos durante la Dinámica molecular. Evolución temporal del RMSD de los ligandos respecto a su posición inicial para GRL-0617 y los derivados benzimidazol. El panel final muestra la superposición de todas las trayectorias..... 152

Figura 108 Perfil de RMSF de la proteína. Fluctuaciones cuadráticas medias por residuo para la proteína apo y los complejos proteína-ligando. La región sombreada corresponde al loop BL2..... 153

Figura 109 Superposición de RMSF y ampliación del loop BL2. Comparación directa de las fluctuaciones por residuo entre los distintos sistemas y ampliación de la región BL2 (residuos 255–275). 154

Figura 110 Distancia catalítica Cys111–His272. Evolución temporal de la distancia entre los residuos catalíticos Cys111 y His272 durante las simulaciones. 155

Figura 111 Distancia Pro299–Tyr268 asociada al movimiento del loop BL2..... 156

Figura 112 Distancia Arg166–Tyr268 asociada a la apertura del sitio alostérico. 156

Figura 113 Distribución angular del diedro χ_1 del residuo Tyr268..... 157

Figura 114 Distribución angular del diedro χ_2 del residuo Tyr268..... 158

Figura 115 RMSD de la proteína durante las simulaciones de Dinámica molecular. Trayectorias de RMSD para la proteína en estado apo, el complejo con el ligando de referencia GRL-0617 y los complejos con

las moléculas para cada familia. El panel final muestra la superposición de todas las trayectorias para facilitar la comparación entre sistemas. 159

Figura 116 RMSD del ligando durante las simulaciones de Dinámica molecular. Trayectorias de RMSD con el ligando de referencia GRL-0617 y los complejos con las moléculas para cada familia. El panel final muestra la superposición de todas las trayectorias para facilitar la comparación entre sistemas. 160

Figura 117 Perfil de RMSF de la proteína. Fluctuaciones cuadráticas medias por residuo para la proteína apo y los complejos proteína–ligando. La región sombreada corresponde al loop BL2. 161

Figura 118 Superposición de RMSF y ampliación del loop BL2. Comparación directa de las fluctuaciones por residuo entre los distintos sistemas y ampliación de la región BL2 (residuos 255–275). 162

Figura 119 Distancia catalítica Cys111–His272. Evolución temporal de la distancia entre los residuos catalíticos Cys111 y His272 durante las simulaciones. 163

Figura 120 Distancia Arg166–Tyr268 asociada al movimiento del loop BL2. 164

Figura 121 Distancia Pro299–Tyr268 asociada al movimiento del loop BL2. 165

Figura 122 Distribución angular del diedro χ_1 del residuo Tyr268. 166

Figura 123 Distribución angular del diedro χ_2 del residuo Tyr268. 167

Figura 124 Propiedades de absorción intestinal e interacción con P-glicoproteína predichas. 168

Figura 125 Distribución de la permeabilidad Caco-2 predicha para el conjunto de compuestos. El diagrama de violín representa la densidad de la distribución, mientras que los puntos individuales corresponden a cada molécula evaluada. La línea negra indica la mediana. Los compuestos con valores superiores a $-5.15 \log \text{ cm/s}$ se muestran en verde (alta permeabilidad), mientras que aquellos con valores iguales o inferiores a este umbral se muestran en rojo (baja permeabilidad). La línea roja discontinua indica el umbral empírico de permeabilidad intestinal asociado a una adecuada absorción. 169

Figura 126 Perfil de absorción intestinal promedio por familia química. Mapa de calor que resume las propiedades de absorción predichas para las distintas familias de compuestos analizadas. Las propiedades incluidas corresponden a: probabilidad de absorción intestinal humana (HIA), permeabilidad Caco-2 normalizada, probabilidad de inhibición de P-glicoproteína, y probabilidad de actuar como sustrato de P-glicoproteína. Para facilitar la interpretación visual, el valor de HIA fue transformado como $1 - \text{HIA}$, de modo que valores más altos en el mapa de calor representen un perfil de absorción más favorable. Los valores mostrados corresponden al promedio de las moléculas pertenecientes a cada familia química, permitiendo comparar tendencias globales entre los distintos scaffolds estructurales. 170

Figura 127 Distribución de las propiedades farmacocinéticas relacionadas con la distribución para el conjunto de compuestos analizados. 171

Figura 128 Mapa de calor que resume las propiedades de distribución predichas para las distintas familias estructurales analizadas. Las variables incluidas corresponden a la probabilidad de penetración a la barrera hematoencefálica (BBB), la unión a proteínas plasmáticas transformada como $1 - \text{PPB}/100$ (PPB favorable), la fracción libre en plasma (F_u) y el volumen de distribución en estado estacionario (VDss), estos últimos normalizados para facilitar la comparación visual. Los valores mostrados corresponden al promedio de cada descriptor dentro de cada familia química. La escala de color representa la magnitud relativa de cada propiedad. 172

Figura 129 Distribución de las probabilidades predichas de inhibición de isoformas del sistema Citocromo P450 (CYP1A2, CYP2C9, CYP2C19, CYP2D6 y CYP3A4) para el conjunto de moléculas analizadas. Los diagramas tipo violín representan la densidad de distribución de las probabilidades en el rango de 0 a 1, mientras que la línea central indica la mediana de cada distribución. Los puntos individuales corresponden a valores muestreados para visualizar la dispersión de los datos. 173

Figura 130 Distribución de las probabilidades predichas de comportamiento como sustrato para las principales isoformas del sistema Citocromo P450 (CYP1A2, CYP2C9, CYP2C19, CYP2D6 y CYP3A4). Los diagramas tipo violín muestran la densidad de los valores predichos en el rango de 0 a 1, mientras que la

línea central representa la mediana de cada distribución. Los puntos superpuestos ilustran la variabilidad de los datos individuales.....	173
Figura 131 Perfil metabólico promedio por familia estructural, representado como un mapa de calor de las probabilidades predichas de inhibición (inh) y comportamiento como sustrato (sub) para las principales isoformas del sistema Citocromo P450. Los valores corresponden al promedio de las probabilidades dentro de cada familia, en una escala de 0 a 1. La intensidad de color refleja la magnitud de la probabilidad promedio, permitiendo comparar las tendencias metabólicas entre familias estructurales.	174
Figura 132 Distribución global de los parámetros de excreción. Diagramas tipo violín que muestran la distribución de los valores predichos de clearance plasmático (CL) y vida media ($t_{1/2}$) para el conjunto total de compuestos. Los puntos representan valores individuales. Las líneas discontinuas indican los umbrales empíricos del modelo ADMETlab para la clasificación de clearance plasmático (5 y 15 mL/min/kg) y vida media (1, 4 y 8 h).....	175
Figura 133 Distribución de los parámetros de excreción por familia química. Diagramas tipo violín que representan la distribución del clearance plasmático (CL) y la vida media ($t_{1/2}$) agrupados según la familia estructural de los compuestos. Los puntos corresponden a valores individuales. Las líneas discontinuas indican los umbrales empíricos del modelo ADMETlab para la clasificación de clearance plasmático (5 y 15 mL/min/kg) y vida media (1, 4 y 8 h).	175
Figura 134 Perfil promedio de excreción por familia química. Mapa de calor que muestra los valores promedio de clearance plasmático (CL) y vida media ($t_{1/2}$) para cada familia química. La escala de color representa la magnitud relativa de los valores dentro de cada propiedad, mientras que los valores numéricos corresponden a los promedios calculados para cada grupo estructural.	176
Figura 135 Clasificación toxicológica por endpoint. Distribución porcentual de compuestos clasificados como positivos y negativos para los distintos endpoints toxicológicos predichos por ADMETlab 3.0. La clasificación se realizó considerando baja probabilidad (≤ 0.30), probabilidad intermedia (0.30–0.70) y alta probabilidad (> 0.70). Los endpoints incluidos corresponden a probabilidades de positividad/toxicidad reportadas por el modelo.	177
Figura 136 Perfil toxicológico promedio por familia química. Mapa de calor que representa los valores promedio de probabilidad para los distintos endpoints toxicológicos agrupados según la familia estructural de los compuestos. La escala de color muestra la magnitud relativa dentro de cada endpoint, mientras que los valores numéricos corresponden a los promedios observados en cada familia.	178
Figura 137 Distribución de la carga toxicológica global del conjunto de compuestos, definida como el número de efectos toxicológicos con probabilidad predicha superior a 0,5 por molécula. El diagrama de violín representa la densidad de la distribución, mientras que los puntos individuales corresponden a cada compuesto y la línea horizontal indica la mediana. La concentración de valores entre 13 y 16 efectos positivos sugiere una elevada acumulación de alertas toxicológicas en el conjunto analizado.	178
Figura 138 Distribución ordenada de los valores de logS predichos para la biblioteca evaluada. Los compuestos se presentan ordenados de acuerdo con su valor de logS, desde el mayor al menor. La línea horizontal punteada indica el umbral de referencia de logS = -4.	179
Figura 139 Distribución de logS por familia química. Se muestran las distribuciones de los valores de logS predichos para cada familia química mediante gráficos de violín, ordenados según la mediana de logS. La línea horizontal punteada indica el umbral de referencia de logS = -4.	180
Figura 140 Distribución de logS según el aldehído empleado. Se presentan las distribuciones de logS agrupadas de acuerdo con el aldehído precursor, mediante gráficos de violín ordenados según la mediana de cada grupo. La línea horizontal punteada indica el umbral de referencia de logS = -4.	180
Figura 141 Distribución de logS según la cetona empleada. Se muestran las distribuciones de logS agrupadas según la cetona precursora, mediante gráficos de violín ordenados de acuerdo con la mediana de cada grupo. La línea horizontal punteada indica el umbral de referencia de logS = -4.	181

Figura 142 Propiedades de absorción intestinal e interacción con P-glicoproteína predichas para la serie de benzimidazoles. (A) Clasificación de las moléculas según la probabilidad de absorción intestinal humana (HIA). Valores entre 0–0.30 corresponden a absorción favorable, entre 0.30–0.70 a absorción intermedia y mayores a 0.70 a absorción desfavorable. (B) Distribución de compuestos predichos como inhibidores y sustratos de P-glicoproteína (P-gp), empleando un umbral de probabilidad de 0.5 para la clasificación binaria. 182

Figura 143 Distribución de la permeabilidad Caco-2 predicha para la serie de benzimidazoles. El diagrama de violín representa la densidad de la distribución y los puntos individuales corresponden a cada compuesto evaluado. La línea negra indica la mediana. Los compuestos con valores superiores a $-5.15 \log \text{ cm/s}$ se consideran de alta permeabilidad, mientras que la línea roja discontinua indica el umbral empírico asociado a una adecuada absorción intestinal. 183

Figura 144 Mapa de calor que resume las propiedades de absorción predichas para el conjunto analizado. Las propiedades incluidas corresponden a 1 – HIA, permeabilidad Caco-2 normalizada, probabilidad de inhibición de P-glicoproteína y probabilidad de actuar como sustrato de P-glicoproteína. Los valores corresponden al promedio de la serie. 183

Figura 145 Propiedades de distribución predichas para la serie de benzimidazoles. (A) Clasificación de penetración a la barrera hematoencefálica (BBB) basada en un umbral de probabilidad de 0.5, mostrando el número de moléculas clasificadas como BBB+ y BBB-. (B) Distribución de la unión a proteínas plasmáticas (PPB, %). La línea negra representa la mediana y la línea discontinua roja indica el umbral de referencia del 90%. (C) Distribución de la fracción libre en plasma (F_u , %). La línea negra representa la mediana, mientras que las líneas discontinuas indican valores de referencia en 5% y 20%. (D) Distribución del volumen de distribución en estado estacionario (VD_{ss} , L/kg) en escala logarítmica. La línea negra indica la mediana y las líneas discontinuas verdes delimitan el rango de referencia (0.04–20 L/kg). Los puntos representan los valores individuales de cada compuesto. 184

Figura 146 Perfil promedio de distribución de la serie de benzimidazoles. El mapa de calor muestra los valores promedio normalizados de los parámetros de distribución: penetración BBB, proporción de PPB favorable, fracción libre en plasma (F_u) y volumen de distribución normalizado (VD_{ss}). La intensidad del color corresponde al valor relativo de cada parámetro dentro de la serie. Los valores numéricos indicados en cada celda representan los promedios calculados para cada descriptor. 185

Figura 147 Distribución de probabilidades de inhibición de isoformas del citocromo P450 para la serie de benzimidazoles. Se muestran las distribuciones de las probabilidades predichas de inhibición para las isoformas CYP1A2, CYP2B6, CYP2C19, CYP2C8, CYP2C9, CYP2D6 y CYP3A4. Los violines representan la densidad de distribución de los datos, los puntos corresponden a los valores individuales de cada compuesto y la línea horizontal negra indica la mediana. 185

Figura 148 Distribución de probabilidades de comportamiento como sustrato de isoformas del citocromo P450 para la serie de benzimidazoles. Se presentan las distribuciones de las probabilidades predichas de ser sustrato para las isoformas CYP1A2, CYP2B6, CYP2C19, CYP2C9, CYP2D6 y CYP3A4. Los violines muestran la distribución de los datos, los puntos representan los valores individuales y la línea horizontal negra indica la mediana. 186

Figura 149 Perfil metabólico promedio de la serie de benzimidazoles. El mapa de calor muestra las probabilidades promedio de inhibición (inh) y comportamiento como sustrato (sub) para distintas isoformas del citocromo P450. La intensidad del color corresponde al valor de probabilidad promedio, según la escala indicada. Los valores numéricos dentro de cada celda representan los promedios calculados para cada propiedad. 186

Figura 150 Distribución global de los parámetros de excreción predichos para la serie de benzimidazoles. Los gráficos de violín representan la densidad de los valores de clearance plasmático (mL/min/kg) y vida media (h), incluyendo la mediana (línea negra) y los valores individuales (puntos). Las líneas horizontales discontinuas indican umbrales de referencia para la clasificación de cada parámetro. 186

Figura 151 Perfil promedio de excreción de la serie de benzimidazoles. Los valores numéricos corresponden al promedio de cada propiedad, mientras que la intensidad del color representa una escala relativa intra-propiedad (0–1), permitiendo comparar tendencias entre parámetros con diferentes unidades. 187

Figura 152 Clasificación toxicológica por endpoint para la serie de benzimidazoles. Distribución porcentual de compuestos en categorías de baja, intermedia y alta probabilidad de toxicidad, según los valores predichos por ADMETlab 3.0. La clasificación se basa en umbrales de probabilidad definidos como $\leq 0,30$ (baja), $0,30\text{--}0,70$ (intermedia) y $> 0,70$ (alta). 187

Figura 153 Perfil toxicológico promedio de la serie de benzimidazoles. Mapa de calor que representa los valores promedio de probabilidad para los distintos endpoints toxicológicos. La escala de color refleja la magnitud de la probabilidad, mientras que los valores numéricos corresponden a los promedios calculados para cada efecto. 188

Figura 154 Distribución de la carga toxicológica global de la serie de benzimidazoles, definida como el número de endpoints con probabilidad superior a 0,5 por molécula. El diagrama de violín representa la densidad de la distribución, mientras que los puntos individuales corresponden a cada compuesto y la línea horizontal indica la mediana. 188

Figura 155 Distribución ordenada de los valores de logS predichos para los benzimidazoles evaluados. Los compuestos se muestran ordenados de acuerdo con su valor de logS, desde el mayor al menor. La línea vertical punteada indica el valor de referencia logS = -4. Los valores obtenidos se distribuyeron entre -4.694 y -4.181, y en todos los casos se ubicaron por debajo de dicho umbral. 189

Figura 156 Distribución ordenada de logS para los benzimidazoles evaluados. Los compuestos fueron ordenados de acuerdo con su valor de logS. La línea punteada indica el valor de referencia logS = -4. Todos los compuestos presentaron valores inferiores a este umbral. 189

Figura 157 Propiedades de absorción intestinal e interacción con P-glicoproteína predichas para el conjunto de moléculas seleccionadas mediante virtual screening. (A) Clasificación de las moléculas según la probabilidad de absorción intestinal humana (HIA). Valores entre 0–0.30 corresponden a absorción favorable, entre 0.30–0.70 a absorción intermedia y mayores a 0.70 a absorción desfavorable. (B) Distribución de compuestos predichos como inhibidores y sustratos de P-glicoproteína (P-gp), empleando un umbral de probabilidad de 0.5 para la clasificación binaria. Los puntos coloreados corresponden a las moléculas individuales identificadas en la leyenda. 191

Figura 158 Distribución de la permeabilidad Caco-2 predicha para el conjunto de moléculas seleccionadas mediante virtual screening(VS). El diagrama de violín representa la densidad de la distribución y los puntos individuales corresponden a cada molécula evaluada. La línea negra indica la mediana. Los compuestos con valores superiores a -5.15 log cm/s se consideran de alta permeabilidad intestinal, mientras que la línea roja discontinua señala el umbral empírico de referencia..... 191

Figura 159 Propiedades de distribución predichas para el conjunto de moléculas seleccionadas mediante virtual screening(VS). (A) Clasificación de penetración a la barrera hematoencefálica (BBB) basada en un umbral de probabilidad de 0.5, mostrando el número de moléculas clasificadas como BBB+ y BBB-. Los puntos coloreados indican las moléculas individuales correspondientes a cada categoría. (B) Distribución de la unión a proteínas plasmáticas (PPB, %). La línea negra representa la mediana y la línea discontinua roja indica el umbral de referencia de 90%. (C) Distribución de la fracción libre en plasma (Fu, %). La línea negra representa la mediana, mientras que las líneas discontinuas indican los valores de referencia de 5% y 20%. (D) Distribución del volumen de distribución en estado estacionario (VDss, L/kg) en escala logarítmica. La línea negra indica la mediana y las líneas discontinuas verdes delimitan el rango de referencia (0.04–20 L/kg). Los puntos representan los valores individuales de cada molécula identificada en la leyenda. 192

Figura 160 Distribución predicha de la probabilidad frente a isoformas de CYP450. En el panel de inhibición, CYP2C8 presentó la mayor tendencia inhibitoria, con una mediana cercana a 1.00, seguida de CYP2C9 con una mediana de 0.77. En cambio, CYP2B6 y CYP2D6 mostraron medianas menores,

mientras que CYP1A2, CYP2C19 y CYP3A4 exhibieron distribuciones más dispersas. En el panel de sustrato , las mayores probabilidades se observaron para CYP2C9, CYP2D6 y CYP1A2, con medianas de 0.99, 0.99 y 0.95, respectivamente, mientras que CYP2B6 y CYP3A4 presentaron una tendencia globalmente baja. En conjunto, el conjunto VS mostró un comportamiento metabólico heterogéneo, con una interacción particularmente marcada con isoformas específicas del sistema CYP450.	193
Figura 161 Distribución de probabilidades de ser sustrato de isoformas de CYP450 para el conjunto de moléculas seleccionadas mediante virtual screening. El diagrama de violín representa la densidad de la distribución para cada isoforma y los puntos coloreados corresponden a las moléculas individuales identificadas en la leyenda. La línea negra indica la mediana. Se evaluaron las isoformas CYP1A2, CYP2B6, CYP2C19, CYP2C8, CYP2C9, CYP2D6 y CYP3A4.	193
Figura 162 Propiedades de excreción predichas para las moléculas seleccionadas mediante virtual screening (VS). El clearance plasmático osciló entre 1.40 y 3.58 mL/min/kg, con una mediana de 2.97 mL/min/kg ,ubicándose todas las moléculas por debajo del valor de referencia de 5 mL/min/kg. De manera consistente, la vida media presentó valores entre 0.44 y 2.07 h, con una mediana de 0.57 h concentrándose la mayoría de los compuestos por debajo de 1 h y sin alcanzar el umbral de 4 h. En conjunto, el perfil de excreción predicho sugiere una serie con comportamiento relativamente homogéneo, dominado por bajo clearance plasmático y vida media breve.	194
Figura 163 Distribución de la probabilidad de toxicidad por efecto evaluado para el conjunto de moléculas seleccionadas mediante virtual screening. Las barras apiladas muestran el porcentaje de compuestos clasificados en baja probabilidad (≤ 0.30), probabilidad intermedia (0.30–0.70) y alta probabilidad (> 0.70) para cada endpoint toxicológico.	195
Figura 164 Perfil toxicológico promedio del conjunto de moléculas seleccionadas mediante virtual screening. El mapa de calor representa la probabilidad media predicha para cada endpoint toxicológico. La intensidad del color aumenta con la magnitud de la probabilidad promedio y los valores numéricos se indican dentro de cada celda.	195
Figura 165 Distribución global de la carga toxicológica del conjunto de moléculas seleccionadas mediante virtual screening. El diagrama de violín representa la densidad de la distribución del número de endpoints con probabilidad superior a 0.5 para cada molécula. Los puntos coloreados corresponden a los compuestos individuales identificados en la leyenda y la línea negra indica la mediana.	196
Figura 166 Distribución ordenada de logS para el conjunto de moléculas seleccionadas mediante virtual screening. Los puntos coloreados representan los valores individuales de cada compuesto, identificados en la leyenda. La línea discontinua gris indica el umbral empírico de referencia en $\log S = -4.0$, utilizado para distinguir compuestos con solubilidad relativamente más favorable de aquellos con menor solubilidad predicha.	197
Figura 167 Análisis retrosintético de las chalconas. La desconexión del sistema α,β -insaturado conduce a una cetona aromática ($R_1\text{--CO--}$) y un aldehído aromático ($R_2\text{--CHO}$), los cuales pueden ser combinados mediante condensación de Claisen–Schmidt para generar la estructura final.	197
Figura 168 Análisis retrosintético de los derivados aminopirimidínicos. La estructura heterocíclica puede ser relacionada con una chalcona como precursor carbonílico α,β -insaturado y guanidina como fragmento nitrogenado. Esta transformación corresponde a un proceso de ciclación/condensación que conduce a la formación del anillo pirimidínico.	198
Figura 169 Análisis retrosintético de los derivados de dihidrobenzodiazepina. La estructura heterocíclica puede ser relacionada con una chalcona como precursor α,β -insaturado y o-fenilendiamina como fragmento nitrogenado bifuncional, los cuales permiten la construcción del anillo benzodiazepínico.	198
Figura 170 Análisis retrosintético de los derivados de dihidropirazol. La desconexión del anillo heterocíclico conduce a una chalcona como precursor α,β -insaturado y a hidrazina como fragmento nitrogenado bifuncional.	198

Figura 171 Análisis retrosintético de los derivados de benzodiazepina. La desconexión del sistema heterocíclico conduce a una chalcona como precursor α,β -insaturado y a o-fenilendiamina como fragmento nitrogenado bifuncional.	199
Figura 172 Análisis retrosintético de los derivados de pirazol. La desconexión del anillo aromático conduce a una chalcona y a hidrazina como precursores inmediatos, evidenciando la formación de un sistema heterocíclico aromatizado a partir de estos bloques de construcción.....	199
Figura 173 Análisis retrosintético de los derivados amino-pirimidina-benzimidazol. La desconexión del sistema heterocíclico conduce a una chalcona como precursor α,β -insaturado y a una guanidina funcionalizada con benzimidazol como fragmento nitrogenado, estableciendo una relación directa entre los productos finales y sus bloques de construcción.	200
Figura 174 Análisis retrosintético de los derivados pirimido-benzimidazol. La desconexión del sistema heterocíclico fusionado conduce a una chalcona como precursor α,β -insaturado y a un 2-aminobenzimidazol como fragmento nitrogenado bifuncional, estableciendo una relación directa entre los productos finales y sus bloques de construcción.	200
Figura 175 Estrategia de funcionalización del núcleo benzimidazol mediante acilación con cloruros de ácido aromáticos. El proceso permite la incorporación de fragmentos naftílicos, generando derivados N-acilados.....	201
Figura 176 Análisis retrosintético de los derivados benzimidazol-heteroaromáticos. La estrategia se basa en la desconexión del enlace C–N clave, conduciendo a o-fenilendiamina y aldehídos heteroaromáticos como precursores, seguido de ciclación para formar el núcleo benzimidazol.	201
Figura 177 Estructura de chalconas obtenidas.	202
Figura 178 Estructura del derivado pirimido-benzimidazol.....	203
Figura 179 Espectro de infrarrojo de la molécula Pirimido benzimidazol con sustracción de solvente (acetonitrilo)	204
Figura 180 Espectro de ^1H RMN (400 MHz, DMSO- d_6) del derivado pirimido-benzimidazol. Se observan señales en la región aromática ($\delta \approx 6.5\text{--}8.5$ ppm), consistentes con los protones del sistema heteroaromático y anillos bencénicos presentes en la estructura.	204
Figura 181 Ampliación de la región aromática del espectro de ^1H RMN (400 MHz, DMSO- d_6) del derivado pirimido-benzimidazol. Se distinguen los patrones de multiplicidad e integraciones correspondientes a los protones aromáticos, evidenciando la complejidad del entorno químico.	205
Figura 182 Espectro de ^{13}C RMN (100 MHz, DMSO- d_6) del derivado pirimido-benzimidazol. Se observan señales en la región de $\delta \approx 110\text{--}160$ ppm correspondientes a carbonos aromáticos y heteroaromáticos. La señal del disolvente (DMSO- d_6) se identifica alrededor de $\delta \approx 39$ ppm.	205
Figura 183 Ampliación del espectro de ^{13}C RMN (100 MHz, DMSO- d_6) del derivado pirimido-benzimidazol, mostrando el multiplete característico en la región de $\delta \approx 125\text{--}128$ ppm, correspondiente al carbono aromático acoplado al grupo CF_3 . El patrón de cuarteto observado se atribuye al acoplamiento $^{13}\text{C}\text{--}^{19}\text{F}$	206
Figura 184 Espectro bidimensional $^1\text{H}\text{--}^1\text{H}$ COSY (400 MHz, DMSO- d_6) del derivado pirimido-benzimidazol, mostrando correlaciones entre protones acoplados a través de enlaces.	206
Figura 185 Espectro bidimensional $^1\text{H}\text{--}^{13}\text{C}$ HMBC del derivado pirimido-benzimidazol, mostrando correlaciones de largo alcance que permiten establecer la conectividad entre protones y carbonos no directamente enlazados.	207
Figura 186 Espectro bidimensional $^1\text{H}\text{--}^1\text{H}$ NOESY (400 MHz, DMSO- d_6) del derivado pirimido-benzimidazol, evidenciando correlaciones espaciales entre protones próximos.	207
Figura 187 Moléculas finales para cada familia analizada y las obtenidas por el Virtual screening.	228
Figura 188 Posibilidades de ciclación para el Pirimido-benzimidazol.....	244
Figura 189 Espectro infrarrojo con sustracción de solvente (acetonitrilo) y asignaciones.	245
Figura 190 Espectro ^1H rmn con señales etiquetadas.	246
Figura 191 Espectro de ^{13}C con señales etiquetadas.	247

Figura 192 Espectro 2D HMBC con señales etiquetadas..... 248

Figura 193 Espectro 2D COSY con señales etiquetadas..... 249

Figura 194 Espectro 2D NOESY con señales etiquetadas..... 250

Figura 195 Resumen esquemático de las correlaciones más relevantes observadas en los experimentos bidimensionales HMBC, COSY y NOESY del compuesto pirimido-benzimidazólico trifluorometilado. Las correlaciones HMBC permitieron asignar carbonos cuaternarios y establecer conectividades de largo alcance; el COSY permitió definir los sistemas de acoplamiento entre protones aromáticos vecinales; y el NOESY aportó evidencia de proximidad espacial entre protones del núcleo heterocíclico y del anillo aromático sustituido, respaldando la asignación final de la estructura y una disposición molecular plegada, consistente con la regioisomería propuesta..... 252

2 Introducción

2.1 Contexto general del SARS-CoV-2

A lo largo de los últimos años, el SARS-CoV-2 ha demostrado ser uno de los mayores desafíos para la salud global, obligando a la comunidad científica a redoblar esfuerzos en la búsqueda de terapias antivirales complementarias a la vacunación. En este contexto, la presente investigación en el diseño *in silico* de nuevos inhibidores dirigidos a la proteasa *papain-like* (PLpro) del SARS-CoV-2, como estrategia para el desarrollo de antivirales. A continuación, se presenta una exposición detallada de la relevancia del virus, la necesidad imperante de nuevas terapias, la importancia biológica y estructural de PLpro, el uso de metodologías computacionales en el diseño racional de fármacos y los avances recientes junto con las brechas existentes que sustentan la novedad y pertinencia de esta línea de investigación.

El SARS-CoV-2, agente etiológico del COVID-19, surgió a finales del 2019 y, desde entonces, ha provocado millones de infecciones y un elevado número de muertes a nivel mundial, alcanzando un total de 7.110.188 fallecimientos confirmados y 779.102.516 casos hasta Enero de 2026¹, ocasionando una crisis sanitaria . Este virus presenta una estructura característica y una organización genómica compleja (

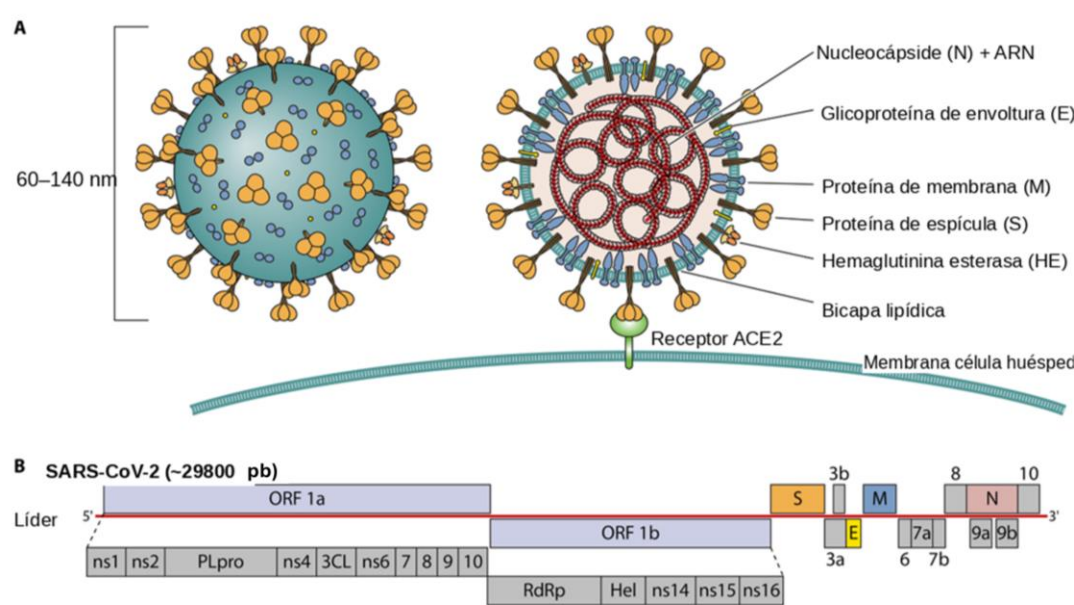


Figura 1), las cuales sustentan su capacidad replicativa y patogénica. Su alta capacidad de transmisión y la diversidad de manifestaciones clínicas, que van desde cuadros leves hasta formas extremadamente graves, han generado consecuencias socioeconómicas y sanitarias de gran magnitud, afectando incluso a países con sistemas de salud robustos². Además, la rápida propagación del virus y la aparición de variantes de preocupación han puesto en evidencia que, si bien la vacunación constituye una herramienta esencial, no es suficiente para contener completamente la propagación ni mitigar todas las complicaciones asociadas a la infección³. Esta situación refuerza la necesidad de desarrollar terapias antivirales adicionales que complementen la acción preventiva de las vacunas y ofrezcan alternativas terapéuticas para pacientes con respuestas inmunitarias insuficientes o ante la emergencia de variantes que reduzcan su eficacia^{2,3}.

A pesar de los éxitos logrados en el desarrollo de vacunas, las limitaciones inherentes a estas ,como la duración de la protección, la necesidad de dosis de refuerzo y la posibilidad de evasión inmunitaria por parte de variantes emergentes, hacen imperativo explorar nuevas estrategias terapéuticas. El desarrollo de antivirales se presenta como una herramienta complementaria crucial para reducir la carga viral, disminuir la severidad del cuadro clínico y evitar la evolución hacia estados críticos, sobre todo en poblaciones vulnerables y en escenarios donde la vacunación no alcanza la efectividad esperada^{3,4}. En este sentido, la búsqueda de nuevos fármacos que actúen sobre blancos terapéuticos esenciales para el ciclo de replicación viral es prioritaria para ampliar el arsenal terapéutico contra la COVID-19, lo que también constituye una estrategia para el manejo de futuras pandemias causadas por otros coronavirus⁴.

Entre los diferentes blancos terapéuticos que han sido estudiados, la papaina-like protease (PLpro) del SARS-CoV-2 emerge como uno de los objetivos más atractivos y estratégicos. PLpro cumple funciones fundamentales en la replicación viral (Figura1) : tras la síntesis de los polipéptidos virales pp1a y pp1ab, esta enzima es responsable de la escisión de estos precursores para liberar las proteínas no estructurales (*non-structural proteins* (nsps)) esenciales para la formación del complejo replicativo^{2,4}. Además, PLpro desempeña un papel dual al intervenir en la modulación de la respuesta inmune del huésped; ejerce actividad desubiquitinasa y desISGylasa, lo que le permite interferir en la señalización antiviral mediada por interferones y facilitar así la evasión de la respuesta inmune innata ^{2,3}. Dado que estos procesos son cruciales tanto para la replicación viral como para la supresión de la respuesta inmunitaria, la inhibición de PLpro no solo interfiere en el ensamblaje y maduración de la maquinaria replicativa del virus, sino que también puede restaurar parcialmente las defensas naturales del organismo frente a la infección ^{5,6}.

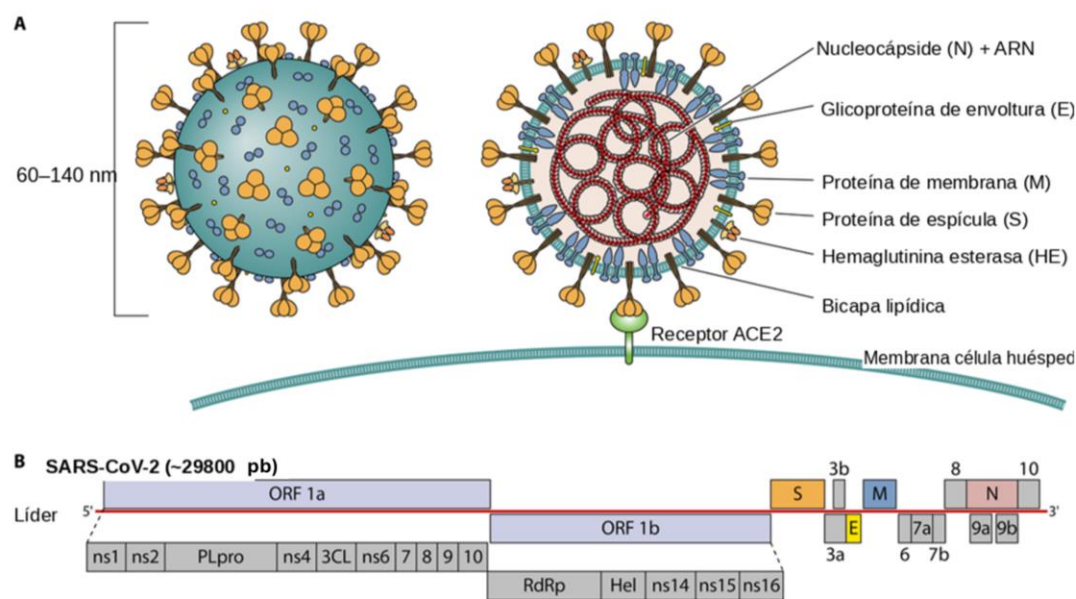


Figura 1 Estructura y organización genómica del SARS-CoV-2.(A) Esquema estructural del virión. (B) Organización del genoma y proteínas virales codificadas, con funciones conocidas y funciones aún no determinadas⁷.

Estructuralmente, PLpro presenta una arquitectura modular que facilita su funcionalidad multifacética. La enzima posee una organización de dominios que incluye un “dominio pulgar-palma-dedo” y un loopflexible conocido como BL2, el cual es fundamental para la interacción con sustratos y para la regulación de la actividad enzimática^{2,6}. Su sitio activo se compone de una triada catalítica formado por los residuos Cys111, His272 y Asp286, que es altamente conservado entre los coronavirus y, a diferencia de otras proteasas, presenta poca similitud con proteasas humanas, lo que minimiza los posibles efectos secundarios y aumenta su idoneidad como blanco terapéutico^{2,8}. Adicionalmente, la presencia de un dominio zinc-dependiente contribuye a la estabilidad estructural de PLpro, lo cual es crítico para mantener su función catalítica⁶. Esta conservación estructural entre SARS-CoV-1 y SARS-CoV-2 ha permitido reutilizar y adaptar estrategias de inhibición desarrolladas en estudios previos, lo que acelera la fase de descubrimiento y optimización de nuevos compuestos⁵.

El diseño racional de fármacos asistido por computador ha revolucionado el proceso de descubrimiento de medicamentos al permitir la evaluación teórica de la interacción fármaco-blanco antes de la síntesis y las pruebas experimentales. Herramientas como el análisis de Relación Cuantitativa Estructura-Actividad (QSAR) han permitido establecer correlaciones matemáticas entre la estructura de una molécula y su actividad biológica, facilitando la predicción y optimización de compuestos con propiedades deseables ^{4,9}. Asimismo, las técnicas de acoplamiento molecular (docking) posibilitan la simulación de la unión de pequeños compuestos a la superficie de la proteína objetivo, identificando modos de unión favorables y las interacciones clave de aminoácidos dentro del sitio activo de PLpro^{4,10}. Complementariamente, las simulaciones de dinámica molecular permiten explorar la estabilidad y la evolución temporal de los complejos inhibidor-PLpro, proporcionando información sobre la energía libre de unión y la flexibilidad del sistema, datos cruciales en la fase de optimización de los inhibidores^{4,9}.

El uso integrado de estas metodologías *in silico* resulta especialmente valioso en el contexto actual, en el que la rapidez y la precisión son esenciales para combatir emergencias sanitarias como la pandemia de COVID-19. La capacidad de realizar virtual screening a gran escala sobre bases de datos de compuestos naturales y sintéticos acelera la identificación de candidatos antivirales, reduciendo significativamente el tiempo y los costos asociados a la experimentación en laboratorio^{3,4}. Además, estos enfoques permiten explorar modificaciones estructurales y optimizaciones químicas sin la necesidad de sintetizar físicamente cada derivado, lo que facilita la generación de una biblioteca de compuestos con alta probabilidad de éxito terapéutico⁵.

En cuanto a los avances recientes en el desarrollo de inhibidores dirigidos a PLpro, diversos estudios han reportado la identificación de moléculas prometedoras a partir de estrategias de virtual screening y diseño basado en estructuras cristalinas. Por ejemplo, el compuesto GRL-0617 se ha destacado como un inhibidor no covalente capaz de bloquear eficazmente la actividad de PLpro, obstruyendo los sitios S3 y S4 del bolsillo de unión del sustrato y provocando un cierre del loopBL2, lo que impide la interacción con el motivo LXGG de los sustratos virales^{6,10}. No obstante, a pesar de la eficacia demostrada *in vitro*, muchos de estos compuestos presentan limitaciones relacionadas con la biodisponibilidad, la selectividad y la estabilidad metabólica, lo que evidencia la necesidad de continuar la optimización a nivel molecular mediante esquemas de diseño racional.

La literatura actual resalta, además, la potencialidad de estrategias que combinan inhibición de PLpro con la inhibición de la proteasa principal (Mpro), lo que podría conducir al desarrollo de antivirales de amplio espectro capaces de bloquear múltiples etapas del ciclo de replicación del SARS-CoV-2^{3,4}. Estas aproximaciones duales no solamente tienen el potencial de mejorar la eficacia antiviral, sino que también disminuyen la probabilidad de desarrollo de resistencia viral, un problema importante en el tratamiento de infecciones virales emergentes^{4,5}.

A pesar de los avances significativos en la identificación de inhibidores para PLpro, persisten diversas brechas que justifican una investigación más profunda en este campo. Una de las principales limitaciones es la discrepancia que a menudo se observa entre la actividad *in vitro* de un compuesto y su eficacia *in vivo*, lo que subraya la necesidad de optimizar no solo la afinidad y la selectividad del inhibidor, sino también sus propiedades farmacocinéticas y de toxicidad^{3,5}. Además, la variabilidad en la conformación del loopBL2 y otros elementos flexibles de PLpro representa un desafío adicional, ya que las dinámicas estructurales pueden influir significativamente en la eficacia de la unión del inhibidor^{4,9}.

La investigación en el diseño *in silico* de inhibidores de PLpro resulta, por tanto, sumamente novedosa y necesaria, ya que integra conocimientos estructurales y funcionales avanzados del SARS-CoV-2 con técnicas computacionales de última generación, permitiendo abordar las deficiencias de los compuestos existentes y proponiendo nuevas soluciones terapéuticas. La combinación de análisis QSAR, docking y simulaciones de dinámica molecular ofrece una plataforma robusta para identificar interacciones críticas y optimizar compuestos a nivel atómico, lo cual es fundamental para superar los obstáculos actuales en la optimización de fármacos antivirales^{4,9}. Además, el enfoque *in silico* facilita la modificación racional de compuestos, permitiendo iteraciones rápidas en el proceso de diseño y reduciendo la dependencia de ensayos costosos y prolongados en laboratorio^{3,5}.

El impacto potencial de esta investigación es doble. En primer lugar, permite acelerar la identificación de nuevos candidatos antivirales que puedan ser evaluados *in vitro* e *in vivo*, lo que contribuirá a la cartera de terapias disponibles para combatir no solo la COVID-19, sino potencialmente también infecciones causadas por otros coronavirus emergentes. En segundo lugar, el desarrollo de inhibidores de PLpro ofrece la posibilidad de mitigar la supresión inmune ejercida por el virus, lo que podría mejorar la respuesta inmunitaria del huésped y reducir la severidad de la enfermedad^{2,3}.

La convergencia de conocimientos provenientes de estudios estructurales, estudios de mecanismo enzimático y aplicaciones computacionales permite fundamentar sólidamente esta línea de investigación como uno de los pilares para el desarrollo de antivirales de amplio espectro. La alta conservación del sitio activo de PLpro entre diferentes coronavirus y la poca similitud con los equivalentes humanos justifican aún más su elección como blanco terapéutico. Asimismo, la integración de datos experimentales con modelados computacionales ofrece una ventana sin precedentes para comprender la dinámica de la unión del inhibidor y potenciar el diseño de compuestos con mayor probabilidad de éxito clínico^{4,10}.

En síntesis, el SARS-CoV-2 representa una amenaza continua para la salud global, y aunque las vacunas han demostrado ser eficaces en la prevención de la infección, la existencia de variantes, la demanda de dosis repetidas y la necesidad de salvaguardar a poblaciones vulnerables evidencian la urgencia de

antivirales efectivos. La proteasa PLpro se erige en un blanco terapéutico sumamente atractivo, dada su función indispensable en la replicación viral y en la evasión de la respuesta inmune, lo que le confiere doble relevancia clínica^{2,8}. El enfoque de diseño *in silico*, sustentado en metodologías computacionales avanzadas como QSAR, virtual screening, docking y dinámica molecular, permite no solo acelerar el descubrimiento de nuevos inhibidores, sino también optimizar sus propiedades biofísicas y farmacológicas, garantizando la eficiencia y seguridad de los candidatos antivirales^{4,9}.

Finalmente, la presente investigación es novedosa y de gran relevancia, ya que aborda en forma integral el reto planteado por el SARS-CoV-2 desde una perspectiva multidisciplinaria que combina química medicinal, biología estructural y ciencia computacional. La capacidad de diseñar nuevos inhibidores de PLpro mediante técnicas *in silico* no solo representa un avance metodológico considerable, sino que además abre la puerta a la implementación de terapias antivirales capaces de actuar de manera complementaria a las vacunas, contribuyendo a controlar la replicación viral y a mejorar la respuesta inmune del huésped. Este enfoque resulta especialmente pertinente en el contexto de futuras emergencias sanitarias, en las que la rapidez y eficacia de los desarrollos terapéuticos serán determinantes para contener brotes epidémicos y pandemias^{3,9}.

Tomando en cuenta lo anteriormente expuesto, el diseño *in silico* de los inhibidores de la proteasa PLpro del SARS-CoV-2 es una estrategia innovadora para el desarrollo de nuevos antivirales de amplio espectro. La integración de conocimientos sobre la biología del virus, la importancia funcional y estructural de la PLpro y la combinación de herramientas computacionales avanzadas, refuerza y resalta la importancia y necesidad de esta línea de investigación. Esta estrategia no solo promete mejorar la eficacia de los tratamientos existentes ante la COVID-19, sino que también sienta las bases para prepararse ante futuras amenazas víricas, siendo así una aportación fundamental para la salud global^{2,4,9}.

En vista de lo expuesto, la presente tesis doctoral se justifica por la imperiosa necesidad de desarrollar nuevos antivirales eficaces y seguros, aprovechando el potencial del diseño *in silico* para optimizar la interacción entre inhibidor y blanco terapéutico. La metodología propuesta integrará análisis QSAR, simulaciones de docking y dinámica molecular para identificar, evaluar y perfeccionar nuevos compuestos capaces de inhibir la actividad de PLpro, contribuyendo así al fortalecimiento del arsenal terapéutico contra la COVID-19 y futuras emergencias derivadas de otros coronavirus. Esta aproximación multidisciplinaria y tecnológicamente avanzada no solo aporta valor desde el punto de vista científico, sino que también tiene importantes implicaciones prácticas, al facilitar la transición de los hallazgos *in silico* a ensayos preclínicos y clínicos que garanticen una respuesta rápida y efectiva ante brotes virales de gran magnitud^{3,5}.

2.2 PLpro como diana terapéutica

2.2.1 Rol funcional de PLpro en el ciclo replicativo

La *papain-like protease* (PLpro) del SARS-CoV-2 es una enzima cisteína indispensable para el procesamiento de las poliproteínas virales pp1a y pp1ab, donde realiza cortes específicos en motivos de reconocimiento LXGG, liberando las proteínas no estructurales (nsp1, nsp2 y nsp3) que conforman el complejo replicativo y de transcripción del virus². Esta escisión proteolítica es fundamental para la maduración de la maquinaria replicativa viral, ya que sin la liberación ordenada de estos nsps, el virus no podría ensamblar el replicase y transcribir su genoma de ARN¹¹. La actividad de PLpro se integra en una red de eventos en la que la proteólisis inicial activa a la ARN polimerasa dependiente de ARN (RdRp), facilitando así la replicación del genoma viral¹². Además, estudios experimentales han demostrado que la inhibición de PLpro conduce a una drástica reducción de la capacidad replicativa del SARS-CoV-2, subrayando la importancia funcional de esta proteasa en el ciclo vital viral ²(Figura 2).

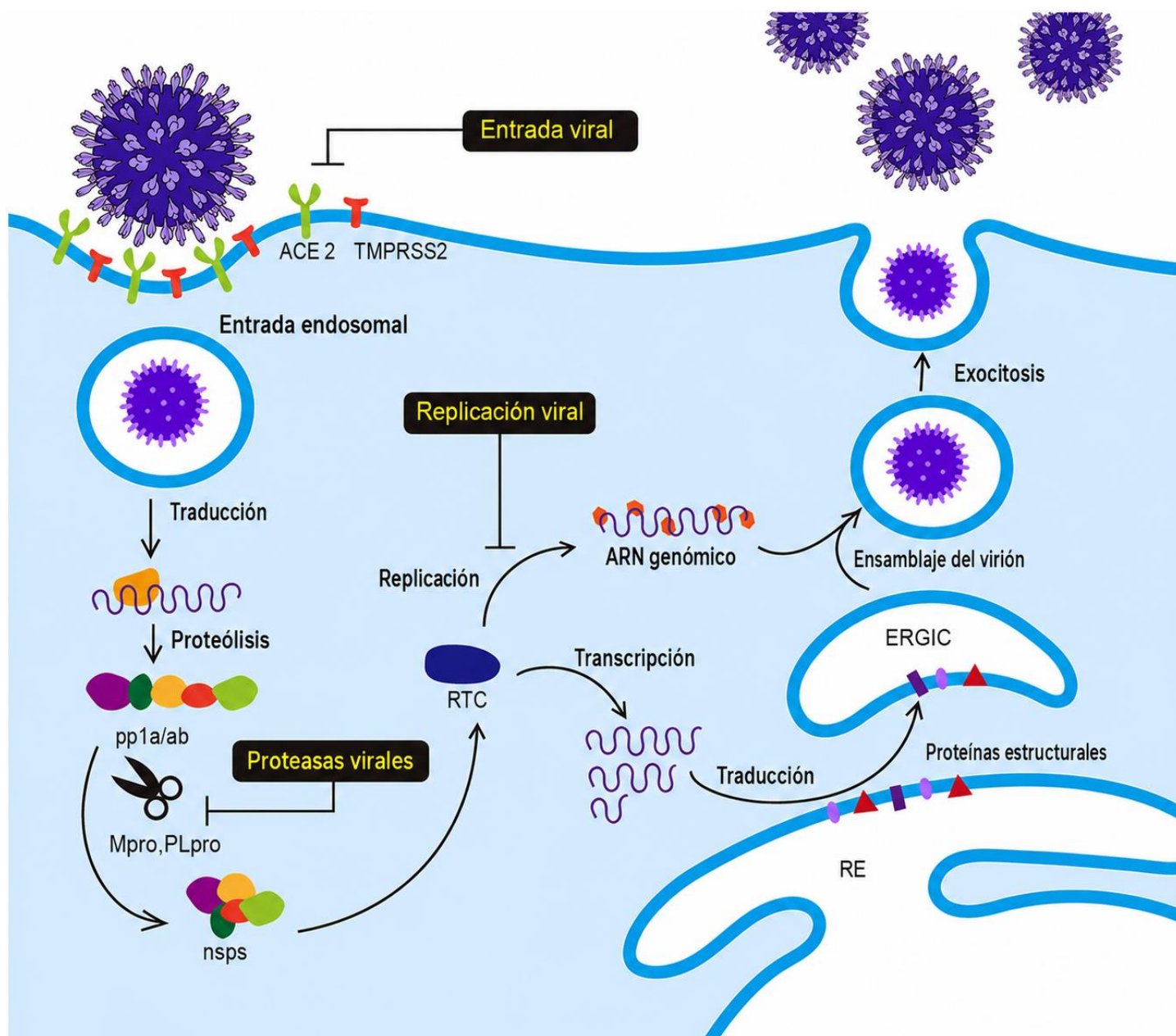


Figura 2 Ciclo replicativo del SARS-CoV-2. La unión de la proteína S al receptor ACE2 permite la entrada viral mediada por proteasas celulares. Tras la liberación del ARN viral, se sintetizan las poliproteínas pp1a/pp1ab, que son escindidas por Mpro y PLpro para generar las nsps que forman el complejo replicasa-transcriptasa. Este complejo dirige la replicación del genoma y la transcripción de ARN subgenómicos. Las proteínas estructurales se ensamblan con el ARN genómico en el ERGIC y los nuevos viriones son liberados por exocitosis¹³.

2.2.2 Función inmunomoduladora de PLpro

Más allá de su función en la proteólisis del polipéptido viral, PLpro actúa como desubiquitinasa y desISGylasa, eliminando modificaciones de ubiquitina e ISG15 en proteínas del huésped que son esenciales para la activación de la respuesta inmune innata¹⁴ (Figura 3). Al remover estas señales regulatorias, PLpro interfiere con las rutas de señalización del interferón tipo I y otros mediadores antivirales, lo que permite al virus evadir la detección inmunitaria y suprimir la respuesta antiviral temprana⁸. Esta función inmunomoduladora resulta especialmente relevante en la patogénesis del SARS-CoV-2, ya que la inhibición de la respuesta interferón favorece una replicación viral más eficiente y una mayor diseminación sistémica¹². En concordancia con ello, estudios in vitro han demostrado que la inhibición de la actividad desubiquitinasa de PLpro puede restaurar parcialmente la respuesta antiviral en

células infectadas, lo que respalda su potencial terapéutico para revertir la evasión inmune mediada por el virus¹⁴.

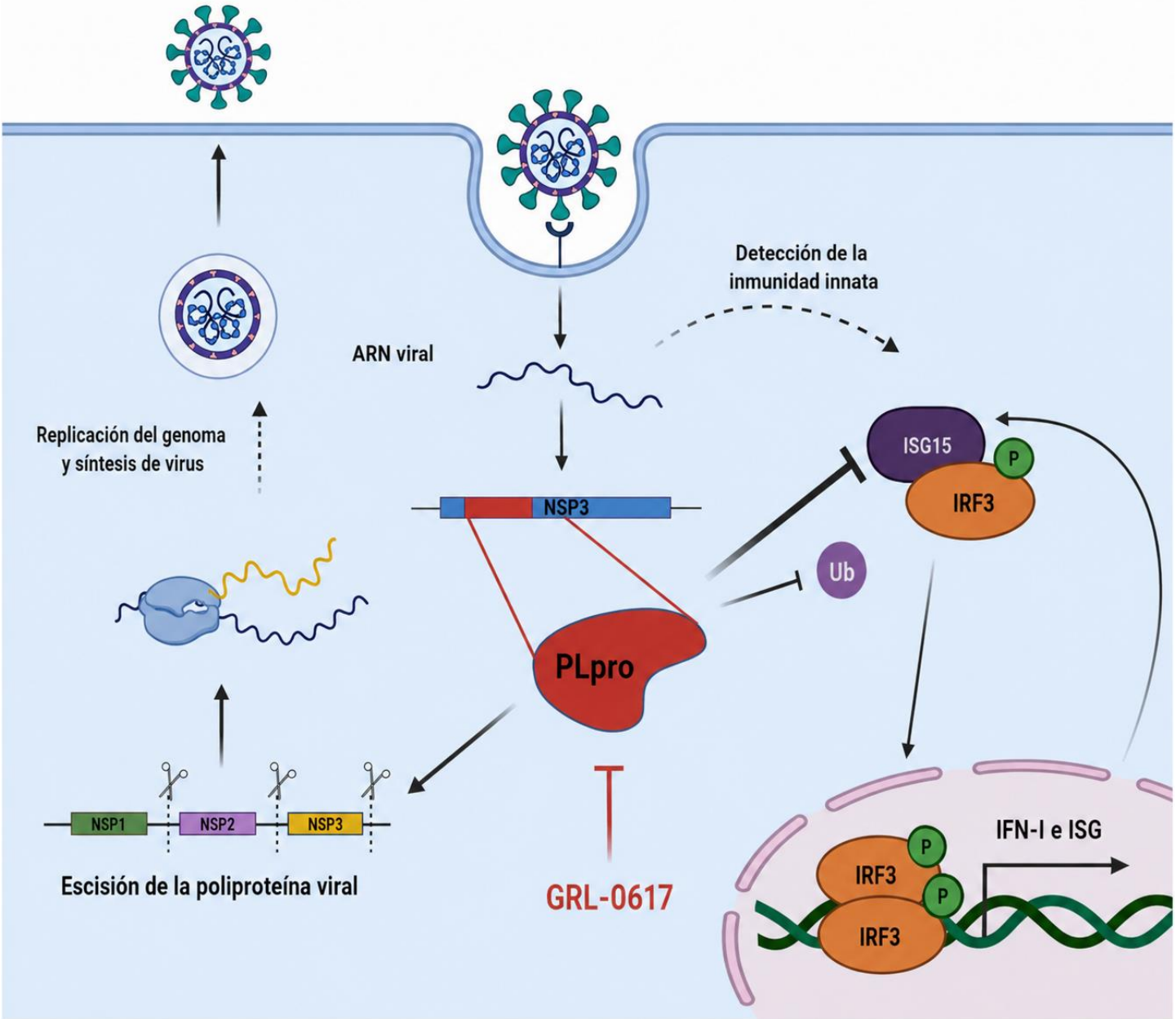


Figura 3 Rol dual de la proteasa PLpro de SARS-CoV-2 en la replicación viral y la evasión de la respuesta inmune innata. PLpro procesa la poliproteína viral para generar proteínas no estructurales maduras necesarias para la formación de un complejo replicasa funcional. Además, esta proteasa interfiere con las defensas antivirales del hospedero al eliminar ISG15 de proteínas celulares, incluida IRF3, provocando la desregulación de la señalización inmune innata¹².

2.2.3 Características estructurales clave de PLpro

Desde el punto de vista estructural, PLpro exhibe una arquitectura modular compuesta por tres dominios funcionales denominados “pulgar”, “palma” y “dedo”, conformando una estructura en forma de mano⁶(Figura 4). En el centro de la actividad catalítica se sitúa un sitio activo conformado por una tríada catalítica: Cys111, His272 y Asp286, cuyas interacciones coordinadas permiten la hidrólisis del enlace peptídico en el polipéptido viral⁶. Un componente esencial para la integridad estructural es el dominio zinc, localizado en la región del “dedo”, donde la coordinación de un ion zinc mediante residuos de cisteína estabiliza la conformación necesaria para la actividad enzimática⁸. Igualmente importante resulta el loopBL2, una estructura flexible situada en las proximidades del sitio activo, que modula el reconocimiento y la entrada del sustrato o del inhibidor, adaptándose a diferentes conformaciones en presencia o ausencia de ligandos⁶.

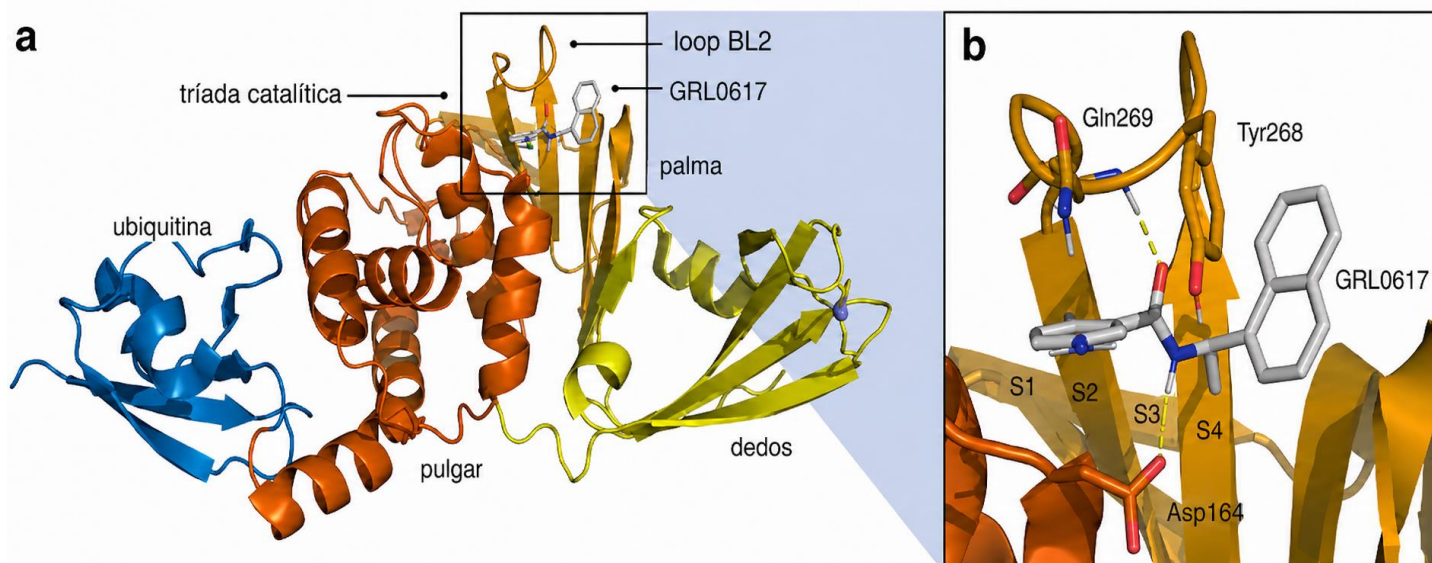


Figura 4 a) Estructura global de PLpro (entrada PDB 7JIR) coloreada por dominios, con características estructurales relevantes indicadas. b) Interacciones entre PLpro y el inhibidor no covalente GRL-0617, destacando los residuos clave y los subsitos de unión del sustrato¹⁵.

2.2.4 Comparación entre PLpro y Mpro

Aunque tanto PLpro como Mpro (proteasa principal o 3CLpro) son esenciales para el procesamiento del polipéptido viral, existen diferencias marcadas en su especificidad y roles funcionales². Mpro realiza múltiples cortes a lo largo de las poliproteínas pp1a/pp1ab, liberando proteínas no estructurales desde nsp4 hasta nsp16, mientras que PLpro se especializa en el procesamiento de la región N-terminal, liberando nsp1–3¹². A diferencia de Mpro, que cumple una función exclusivamente proteolítica, PLpro contribuye adicionalmente a la evasión inmunitaria mediante su actividad desubiquitinasa y desISGylasa, lo que le confiere una relevancia terapéutica adicional¹¹. Asimismo, aunque PLpro comparte similitud funcional con deubiquitinastas humanas, presenta características estructurales distintivas que pueden ser explotadas para el diseño de inhibidores selectivos, minimizando potenciales efectos sobre proteínas del huésped¹⁴.

2.2.5 Justificación científica de PLpro como diana farmacológica

La elección de PLpro como diana terapéutica se basa en múltiples factores. Su estructura y función conservadas entre coronavirus sustentan el potencial de desarrollar inhibidores de amplio espectro¹⁰. Además, aunque PLpro comparte similitud funcional con deubiquitinastas humanas, presenta diferencias estructurales aprovechables para el diseño de inhibidores selectivos, minimizando posibles efectos fuera de objetivo en el huésped¹⁴. En este contexto, se ha descrito que la inhibición de PLpro puede impactar tanto la replicación viral como la modulación de respuestas antivirales innatas, lo que refuerza su interés terapéutico¹². Finalmente, la disponibilidad de estructuras cristalinas de alta resolución habilita estrategias de diseño racional basadas en estructura para la optimización de nuevos compuestos⁶.

2.2.6 Inhibidores conocidos y su mecanismo de acción

Entre los inhibidores desarrollados para bloquear PLpro destaca GRL-0617, un inhibidor no covalente que se une en proximidad al loop BL2, bloqueando el acceso del sustrato natural sin interactuar covalentemente con el sitio catalítico⁸. Su unión induce una conformación cerrada del loop BL2 que impide la interacción con los sustratos virales y proteínas ubiquitina-like⁶. Este mecanismo es relevante porque el loop BL2 actúa como una compuerta conformacional que regula el acceso del sustrato al centro catalítico; por ello, su estabilización en una conformación cerrada limita el procesamiento de ubiquitina, ISG15 y sustratos virales, reduciendo la actividad catalítica de la enzima^{6,16}. Como se observa en la Figura 5, GRL-0617 se une en proximidad al loop BL2, estabilizando su conformación cerrada y bloqueando el acceso del sustrato al sitio catalítico. Además de GRL-0617, se han desarrollado otros compuestos experimentales, incluidos inhibidores covalentes dirigidos al residuo catalítico Cys111, que muestran actividad inhibitoria prometedora; sin embargo, aún requieren optimización para mejorar su potencia, selectividad y propiedades farmacológicas¹⁷. Asimismo, algunos inhibidores afectan la actividad desubiquitinasa de PLpro, lo que podría contribuir a restaurar la respuesta inmune antiviral del huésped¹⁴.

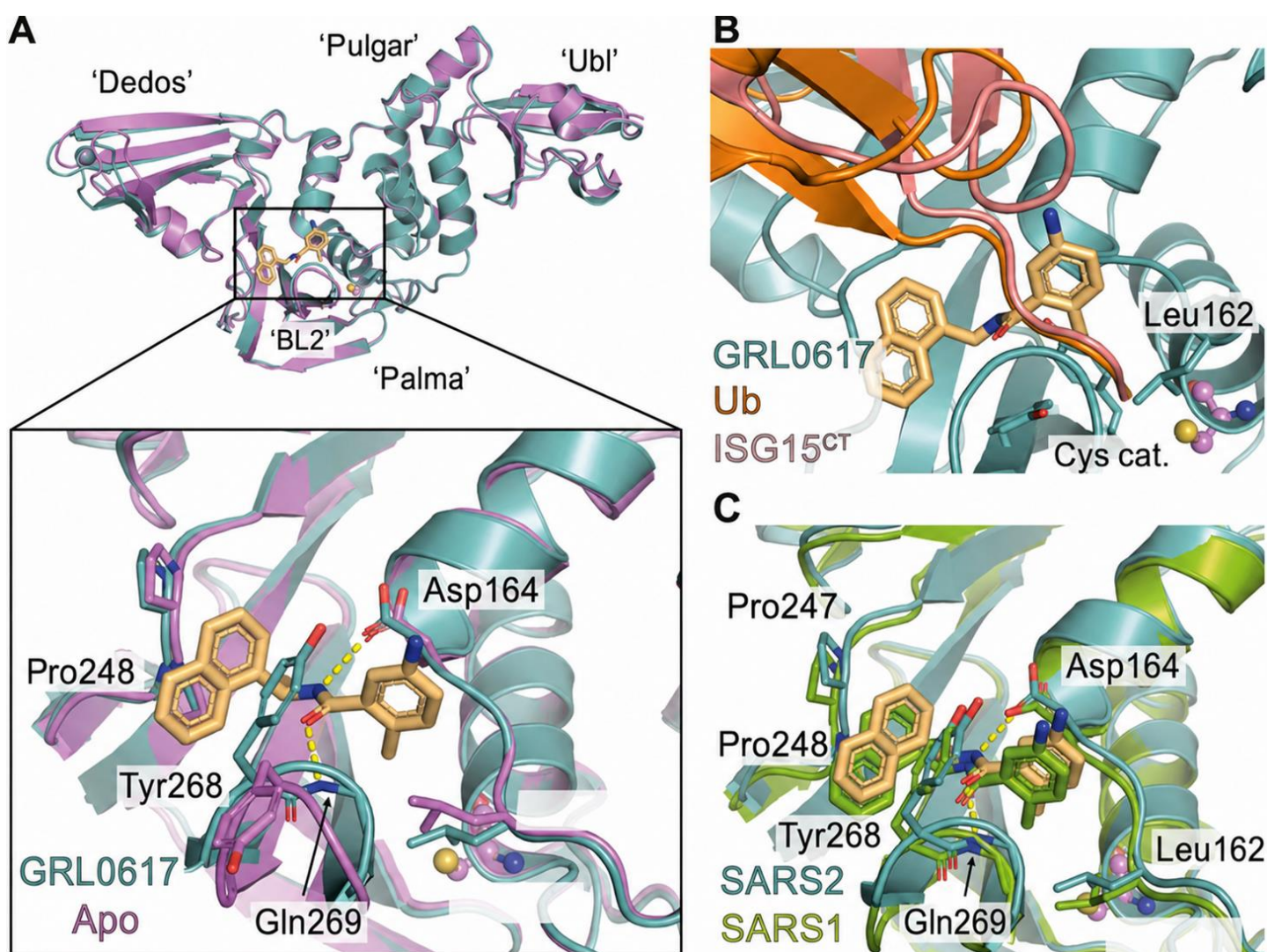


Figura 5 Base molecular de la inhibición de PLpro por GRL-0617.

(A) Estructura de PLpro unida al inhibidor no covalente GRL-0617, mostrando el sitio de unión y los residuos clave involucrados. La superposición con la forma apo indica que la unión del ligando no induce cambios conformacionales globales. (B) Comparación del sitio de unión con ubiquitina e ISG15, evidenciando que el ligando ocupa el canal de acceso del extremo C-terminal de las proteínas ubiquitina-like y bloquea el acceso al residuo catalítico Cys111. (C) Superposición estructural entre PLpro de SARS-CoV-2 y SARS-CoV que muestra la conservación de los residuos del sitio de unión, lo que explica la reactividad cruzada de los inhibidores³.

2.2.7 PLpro: similitud estructural con deubiquitinasa humana y potencial anticancerígeno

La proteasa PLpro presenta similitudes estructurales y funcionales con las deubiquitininas (DUBs) humanas (Figura 6), ya que comparte un dominio catalítico de tipo cisteína y la capacidad de remover ubiquitina e ISG15 de proteínas diana¹⁴. Esta relación refleja la estrategia evolutiva del virus para manipular rutas celulares del huésped, pero también plantea un desafío relevante en términos de selectividad farmacológica durante el diseño de inhibidores.

Las DUBs humanas desempeñan funciones esenciales en la regulación de procesos celulares críticos, incluyendo el control del ciclo celular, la apoptosis, la respuesta al daño del ADN y la señalización inmune. Debido a su papel central en la homeostasis celular, la desregulación de estas enzimas ha sido asociada con diversas patologías, particularmente el cáncer, y su inhibición está siendo explorada como estrategia terapéutica antitumoral.

En este contexto, la similitud funcional entre PLpro y DUBs humanas implica que los inhibidores dirigidos contra PLpro podrían presentar interacciones fuera de objetivo con enzimas del huésped si no se optimiza adecuadamente su selectividad. Por ello, el diseño racional de inhibidores debe considerar cuidadosamente las diferencias estructurales entre PLpro y las DUBs humanas para maximizar la actividad antiviral y minimizar efectos citotóxicos o interferencias con procesos celulares esenciales. Así, aunque la relación estructural con DUBs humanas resalta el potencial farmacológico de esta familia enzimática, también subraya la necesidad de optimización estructural y farmacodinámica para preservar el efecto antiviral y reducir posibles efectos adversos en células humanas^{14,17}.

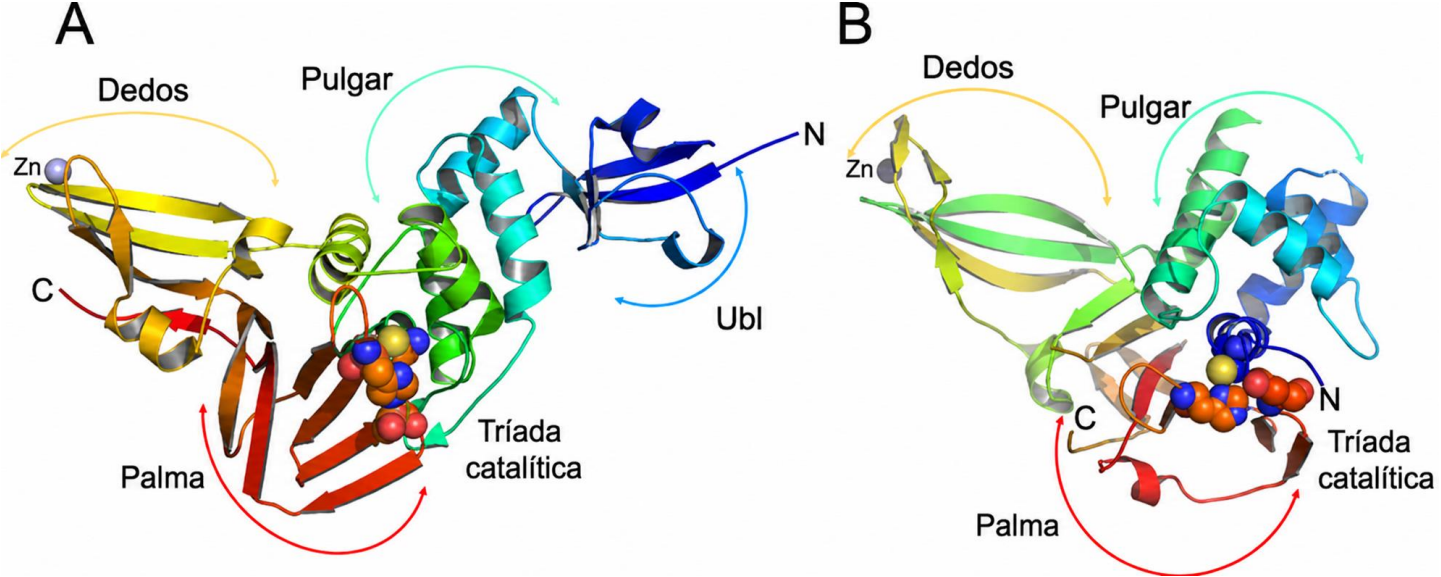


Figura 6 Comparación estructural entre PLpro de SARS-CoV-2 y la deubiquitinasa humana USP12. (A) Modelo estructural de PLpro que muestra su organización secundaria y la arquitectura tipo “thumb–palm–fingers(pulgar,palma,dedos)”, con los dominios señalados. Se destacan los residuos del sitio catalítico Cys111, His272 y Asp286, así como el ion zinc estructural. (B) Estructura de la deubiquitinasa humana USP12 (PDB: 5K1B) que exhibe un plegamiento y organización de dominios similares. Se indican los residuos catalíticos Cys48, His317 y Asp333 y el ion zinc¹⁸.

2.3 Química medicinal asistida por computación en la era post-COVID

La etapa post-pandémica ha obligado a replantear las estrategias de descubrimiento y diseño de fármacos. Aunque la crisis sanitaria aguda ha disminuido, la circulación persistente del SARS-CoV-2 y la emergencia continua de nuevas variantes requieren mantener sistemas de vigilancia epidemiológica y genómica, así como una evaluación constante del riesgo sanitario y de las estrategias de respuesta¹⁹. En este marco, la química medicinal asistida por computación integra metodologías in silico ,como QSAR, docking molecular, simulaciones de dinámica molecular (MD) e inteligencia artificial (IA),para priorizar candidatos y acelerar la generación de hipótesis mecanísticas, sin pretender sustituir la validación experimental²⁰.Estas estrategias permiten racionalizar el espacio químico, reducir iteraciones empíricas tempranas y apoyar decisiones de selección de compuestos antes de invertir en síntesis, ensayos in vitro y etapas posteriores del desarrollo^{20–22}. Estas estrategias se integran dentro de un flujo de trabajo iterativo que permite priorizar compuestos y orientar la validación experimental (Figura 7).

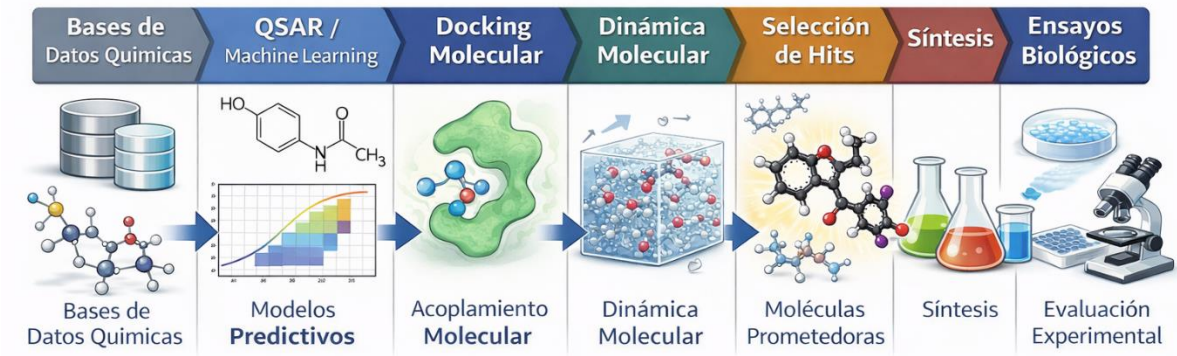


Figura 7 Esquema del flujo de trabajo del descubrimiento de fármacos asistido por computadora. El proceso integra bases de datos químicas, modelado QSAR y aprendizaje automático, docking molecular, simulaciones de dinámica molecular, selección de hits y validación experimental para priorizar candidatos y optimizar recursos en etapas tempranas del diseño de fármacos.(Esquema hecho con inteligencia artificial)

2.3.1 Técnicas computacionales clave en el diseño racional de fármacos

Las metodologías computacionales se han consolidado como herramientas centrales para proponer y evaluar candidatos antivirales en etapas tempranas. En primer lugar, los modelos QSAR (relación cuantitativa estructura-actividad), incluidos enfoques basados en aprendizaje automático, permiten construir relaciones entre descriptores moleculares y medidas de actividad (por ejemplo, datos asociados a inhibición en blancos virales) y, con ello, priorizar moléculas para evaluación posterior²³⁻²⁵.

Por su parte, el docking molecular se emplea para explorar poses de unión y estimar interacciones ligando-diana en modelos estructurales, lo que resulta útil para tareas de *virtual screening* y reposicionamiento (*repurposing*) una vez que existen estructuras tridimensionales disponibles de proteínas virales relevantes^{20,26}. En este mismo flujo de trabajo, las simulaciones de dinámica molecular permiten evaluar la estabilidad conformacional de complejos proteína-ligando a lo largo del tiempo en condiciones modeladas, complementando la foto estática del docking y aportando evidencia computacional adicional sobre persistencia de interacciones^{27,28}.

La principal fortaleza práctica de estos métodos aparece cuando se integran en un enfoque multitécnico, donde las limitaciones de cada aproximación se compensan con información complementaria: (i) QSAR aporta velocidad para filtrar grandes conjuntos; (ii) docking aporta hipótesis de modo de unión; (iii) MD aporta evaluación dinámica de estabilidad. Este tipo de integración ha sido descrita como una estrategia útil en descubrimiento asistido por computación y en revisiones de reposicionamiento y CADD aplicadas a COVID-19^{21,22,25}.

2.3.2 Aplicaciones prácticas, optimización de recursos y casos de éxito

En el contexto de COVID-19, múltiples estudios y revisiones describen el uso de repurposing computacional para identificar candidatos a partir de fármacos existentes o bibliotecas químicas, destacando su utilidad para acelerar la generación de listas priorizadas de compuestos a ensayar^{22,26}. Por ejemplo, se han reportado flujos de cribado virtual y evaluación por MD en estudios orientados a interfaces relevantes del virus, como la interacción Spike/ACE2, con el objetivo de explorar reposicionamiento^{27,28}.

En paralelo, los enfoques QSAR también se han aplicado para priorizar moléculas candidatas a inhibición en proteasas de coronavirus y, en general, para apoyar decisiones de selección temprana mediante modelos predictivos entrenados con datos existentes^{24,25}. Adicionalmente, revisiones sobre inteligencia artificial o *machine learning* en COVID-19 discuten el papel de estos métodos para analizar datos, apoyar repurposing y acelerar etapas de descubrimiento, en conjunto con herramientas clásicas de química computacional²⁹⁻³¹.

Es importante enfatizar que el objetivo típico de estos enfoques no es reemplazar el experimento, sino mejorar la priorización: generar hipótesis, reducir el espacio de búsqueda y orientar los ensayos hacia candidatos con mejor perfil computacional preliminar, para luego refinarlos con evidencia experimental y retroalimentar el ciclo de modelado^{20,22}.

2.3.3 Perspectivas futuras y desafíos éticos-colaborativos

La evolución continua del SARS-CoV-2 y la aparición esperable de nuevas variantes refuerzan la necesidad de mantener vigilancia y estrategias adaptativas¹⁹. En este escenario, los enfoques computacionales pueden contribuir a ciclos iterativos de rediseño y priorización, siempre que se disponga de datos y estructuras actualizadas y que las predicciones se integren con validación experimental^{20,21}.

A futuro, se proyecta un paradigma híbrido que combine métodos basados en datos (inteligencia artificial o *machine learning*) con herramientas fisicoquímicas (docking/MD), tanto para explorar espacios químicos amplios como para refinar hipótesis mecanísticas^{20,29}. Sin embargo, este progreso exige atender desafíos de reproducibilidad, transparencia metodológica, estandarización y colaboración multidisciplinaria para maximizar la transferibilidad y utilidad práctica de los resultados computacionales^{20,21}.

Finalmente, el uso de plataformas abiertas para explorar asociaciones diana-enfermedad puede facilitar la integración de evidencia biológica y clínica, actuando como punto de partida para priorización de blancos o racionalización de hipótesis.

2.4 Rol de las técnicas *in silico* en la química medicinal asistida por computación

El desarrollo de nuevos fármacos ha sido históricamente un proceso largo, costoso y de alta incertidumbre³². Sin embargo, con el avance de la química medicinal asistida por computación, también conocida como *computational drug design*, ha sido posible acelerar significativamente esta etapa crítica

del descubrimiento farmacéutico³³. Estas técnicas *in silico* permiten explorar y optimizar candidatos bioactivos antes de su evaluación experimental, lo que reduce el número de compuestos a sintetizar y analizar en laboratorio³⁴.

Las metodologías computacionales aplicadas en el diseño racional de fármacos pueden clasificarse, a grandes rasgos, en dos grandes enfoques: basadas en el ligando y basadas en el receptor.

2.4.1 Enfoques basados en el ligando

Las estrategias basadas en el ligando se centran en el análisis de compuestos con actividad biológica conocida, incluso en ausencia de información estructural del blanco terapéutico. Entre las herramientas más destacadas se encuentran los modelos Cuantitativos de Relación Estructura-Actividad (QSAR), y en particular los modelos 2D-QSAR y 3D-QSAR.

Estas metodologías buscan establecer una correlación matemática entre los descriptores moleculares, propiedades fisicoquímicas, topológicas, electrónicas, entre otras, y la actividad biológica experimental observada³⁵. Una vez desarrollados y validados, estos modelos permiten predecir la actividad de nuevas moléculas y priorizar aquellas con mayor probabilidad de éxito terapéutico.

QSAR ha sido aplicado exitosamente en múltiples áreas terapéuticas, incluyendo el diseño de antivirales, antitumorales, antibacterianos y antiinflamatorios³⁶. La fortaleza de estos modelos radica en su capacidad de generalización, siempre que se construyan con bases de datos representativas y siguiendo criterios estadísticos rigurosos.

2.4.1.1 2D-QSAR

El enfoque 2D-QSAR (Relación Cuantitativa Estructura-Actividad bidimensional) se basa en establecer relaciones matemáticas mediante métodos estadísticos o de aprendizaje automático entre la estructura química de una molécula y su actividad biológica. Esta actividad puede expresarse, por ejemplo, como inhibición, activación o afinidad frente a un blanco biológico.

En este tipo de modelos, la estructura molecular se representa mediante descriptores moleculares bidimensionales, tales como propiedades fisicoquímicas, topológicas, electrónicas y estéricas, entre otras. Estos descriptores cuantifican características estructurales relevantes de la molécula y se utilizan como variables independientes en modelos de regresión o clasificación.

El resultado es una ecuación matemática en la cual los descriptores estructurales explican la variación observada en la actividad biológica experimental. Una vez desarrollado y validado el modelo, esta ecuación permite predecir teóricamente la actividad biológica de nuevas moléculas a partir de su estructura química.

De manera general, la ecuación QSAR puede expresarse como: Actividad biológica:

$$Actividad = \beta_0 + \beta_1 D_1 + \beta_2 D_2 + \beta_3 D_3 + \dots + \beta_n D_n$$

En la ecuación QSAR, la actividad biológica corresponde a la variable dependiente y puede expresarse mediante parámetros experimentales como pIC_{50} , $\log(IC_{50})$ o la constante de inhibición (K_i). El término β_0 representa el intercepto del modelo, mientras que $\beta_1, \beta_2, \dots$ corresponden a los coeficientes de regresión asociados a cada variable independiente. Los términos $D_1, D_2, \dots$ representan los descriptores moleculares, es decir, variables cuantitativas que describen propiedades estructurales, fisicoquímicas o electrónicas de la molécula.

Descriptores moleculares

Para describir cuantitativamente una estructura molecular es necesario emplear parámetros que permitan representar sus características químicas de manera numérica. Estos parámetros, conocidos como descriptores moleculares, constituyen las variables independientes utilizadas en los modelos QSAR.

Los descriptores pueden derivarse de propiedades fisicoquímicas determinadas experimentalmente o calcularse mediante métodos computacionales a partir de la estructura molecular. Entre ellos se incluyen características topológicas, electrónicas, estéricas y fisicoquímicas que reflejan aspectos relevantes de la molécula.

Los descriptores moleculares establecen una conexión directa entre la estructura química y sus propiedades fisicoquímicas, permitiendo modelar y predecir la actividad biológica de los compuestos en estudios QSAR.

Descriptores fisicoquímicos globales

Los descriptores fisicoquímicos globales se fundamentan en parámetros que reflejan las propiedades físicas y químicas intrínsecas de las sustancias. Variaciones en estas propiedades pueden modificar el comportamiento de las moléculas bioactivas en sistemas biológicos, afectando procesos como su absorción, distribución, transporte, metabolismo y eliminación.

Entre los factores fisicoquímicos que influyen de manera significativa en la actividad biológica de fármacos y otras sustancias químicas destacan la hidrofobicidad, las propiedades electrónicas y los efectos estéricos de las moléculas. Estos parámetros permiten describir características globales de la estructura molecular y constituyen elementos clave para comprender y predecir su comportamiento biológico.

Hidrofobia

El coeficiente de partición octanol/agua (P) constituye una medida fundamental de la distribución de una molécula entre un medio lipofílico y uno acuoso, y se emplea de manera rutinaria como descriptor de hidrofobicidad. Matemáticamente, se expresa como:

$$P = \frac{[C]_{octanol}}{[C]_{agua}}$$

donde [C]_{octanol} y [C]_{agua} corresponden a las concentraciones de la molécula en cada fase. De manera práctica, se suele utilizar la forma logarítmica, logP, que permite correlacionar linealmente este parámetro con la actividad biológica y con otras propiedades moleculares relevantes³⁷.

Diversos estudios han demostrado que la hidrofobicidad es un determinante crítico en procesos de absorción, permeabilidad y biodisponibilidad, así como en la interacción ligando–receptor³⁸. En este contexto, valores elevados de logP reflejan una mayor afinidad lipídica, mientras que valores bajos indican preferencia por el medio acuoso.

El cálculo computacional de logP puede realizarse mediante distintos enfoques. Uno de los más empleados es el método fragmentario, en el que el valor global se obtiene como la suma de contribuciones de fragmentos estructurales y factores de corrección:

$$\log P = \sum a_i f_i + \sum b_j f_j$$

donde f_i representa la constante de cada fragmento, ponderada por su ocurrencia a_i, y f_j corresponde a factores de corrección asociados a efectos como polaridad, flexibilidad conformacional o impedimento estérico³⁹. Estos algoritmos, implementados en diversos paquetes de software de quimioinformática, permiten la estimación eficiente de este descriptor en grandes bases de datos químicas.

En modelos QSAR, además del término lineal logP, se incorpora frecuentemente su forma cuadrática (logP)². El término lineal refleja la tendencia general: si al aumentar la lipofilia la actividad tiende a incrementarse o disminuir. Sin embargo, este comportamiento no suele ser estrictamente lineal en sistemas biológicos. En la práctica, se observa que existe un intervalo óptimo de logP en el cual la actividad es máxima, mientras que valores demasiado bajos (moléculas muy hidrofílicas) o demasiado altos (moléculas excesivamente lipofílicas) resultan desfavorables.

El término cuadrático (logP)² introduce la curvatura parabólica que caracteriza esta relación. No obstante, por sí solo produciría una parábola simétrica sin dirección clara. Por ello, la combinación de logP y (logP)² permite construir un modelo más realista, en el que la pendiente inicial provista por el término lineal se complementa con la curvatura del término cuadrático. De esta forma, el modelo refleja la existencia de un punto óptimo de hidrofobicidad, donde se alcanza la máxima actividad biológica (Figura 8).

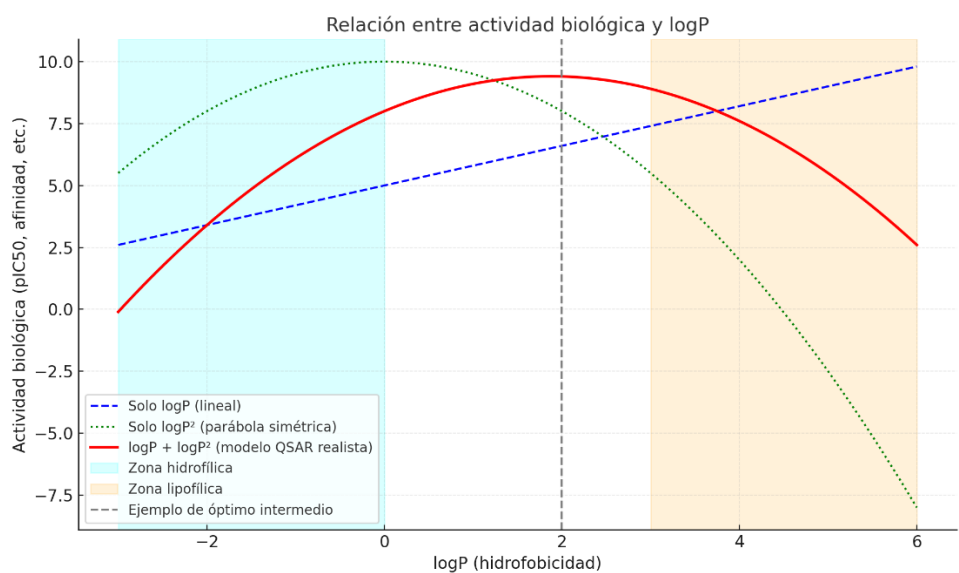


Figura 8 Representación esquemática de la relación entre la hidrofobicidad (logP) y la actividad biológica. El término lineal logP describe la tendencia general de la actividad con la lipofilia, mientras que el término cuadrático (logP)² introduce la curvatura necesaria para reflejar un comportamiento no lineal. La combinación de ambos términos genera un modelo más realista que predice la existencia de un intervalo óptimo de lipofilia asociado a máxima actividad biológica. Las regiones sombreadas ilustran zonas de hidrofobicidad desfavorable por exceso de hidrofilia o lipofilia.

Electrostáticos

La distribución eléctrica de una molécula es un aspecto fundamental en el diseño de fármacos, ya que interacciones como dipolo–dipolo o ión–dipolo dependen directamente de esta propiedad. Cada átomo o grupo funcional presente puede modificar la densidad de carga en la superficie molecular, ya sea aumentándola o disminuyéndola.

Para analizar de manera cuantitativa el aporte electrostático de un sustituyente, es habitual recurrir al estudio de la constante de disociación del ácido benzoico, empleada como patrón para estimar empíricamente el carácter aceptor o dador de electrones de un sustituyente determinado⁴⁰:

$$K_a = \frac{[A^-][H^+]}{[HA]}$$

donde H⁺ corresponde al protón y A⁻ a la base conjugada de HA. Normalmente, esta constante se expresa en su forma logarítmica:

$$pK_a = -\log(K_a).$$

El valor de pKa puede variar cuando se introducen sustituyentes en el anillo aromático. Por ejemplo, el ácido benzoico posee un pKa de 4,20. Si se añade un átomo de bromo en posición orto, el pKa disminuye a 2,85; en posición meta se modifica a 3,81; y en posición para a 3,97.

La relación entre el cambio en la constante de acidez y el efecto electrónico del sustituyente puede describirse mediante la ecuación de Hammett:

$$\log(k/k_0) = \rho\sigma$$

donde ρ es una constante propia del ácido considerado (ρ = 1 para el ácido benzoico) y σ es una constante característica de cada sustituyente. Aquí, k₀ corresponde a la constante de acidez del ácido benzoico no sustituido y k a la del derivado sustituido. En este marco, los sustituyentes que donan densidad electrónica al anillo aromático presentan valores de σ negativos, mientras que los que retiran densidad electrónica exhiben valores de σ positivos.

Polarizabilidad

En 1937 London instaure un nuevo concepto que es la capacidad de inducir dipolos entre molécula⁴¹ esa capacidad Esta capacidad depende tanto de la naturaleza del átomo como de la molécula en su conjunto, y se denomina polarizabilidad. La polarizabilidad describe en qué medida la nube electrónica de una

molécula se deforma frente a la acción de un campo eléctrico externo, generando un dipolo inducido. Cuanto mayor es la polarizabilidad, más fácil resulta distorsionar la distribución electrónica y, por lo tanto, inducir un dipolo.

Una forma de cuantificar la polarizabilidad de las moléculas es con la refractividad molar⁴²:

$$MR = \frac{4\pi N\alpha}{3}$$

donde N es el número de Avogadro y α la polarizabilidad electrónica. Desde un enfoque electrostático, si se considera una molécula esférica de radio(r), puede deducirse que $\alpha=r^3$, lo que implica que la refractividad molar se relaciona directamente con el volumen efectivo de la molécula. De este modo, la contribución atómica a la refractividad molar se interpreta como el volumen de un átomo dentro de la molécula, el cual difiere del volumen de un átomo aislado debido a la influencia de la polaridad de los enlaces y al solapamiento de las nubes electrónicas de los átomos enlazados.

Estéricos

Los componentes estéricos consisten en como el espacio de una molécula o átomo afecta propiedades químicas o biológica. En el punto de vista biológico, estos efectos desempeñan un papel crucial, ya que determinan en gran medida la capacidad de un ligando para adaptarse correctamente al sitio activo de una proteína. Un sustituyente voluminoso o mal orientado puede provocar repulsiones espaciales con residuos cercanos del receptor, alterar la conformación adoptada por el ligando o incluso impedir su unión, reduciendo así la afinidad y la actividad biológica.

Además, los efectos estéricos no solo se manifiestan como simples barreras físicas. También pueden influir indirectamente en otras propiedades moleculares, como la orientación de grupos funcionales clave, la accesibilidad a enlaces de hidrógeno otra interacción intramolecular o intermolecular. En algunos casos, una cierta contribución estérica es incluso deseable, ya que puede mejorar la selectividad al evitar interacciones no específicas o estabilizar la conformación bioactiva del ligan

El volumen que ocupa un sustituyente es un factor fundamental para lograr un acoplamiento adecuado entre el ligando y el sitio activo del receptor. Sustituyentes excesivamente voluminosos pueden generar impedimentos estéricos, dificultando el acceso o el posicionamiento correcto del ligando dentro de la cavidad de unión, lo que repercute negativamente en la interacción y afinidad global del complejo.

Para cuantificar el efecto estérico de un sustituyente se utiliza, por analogía con la ecuación de Hammett, la relación propuesta por Taft⁴³, que permite estimar la contribución estérica a partir de las constantes de hidrólisis de ésteres. En este enfoque, se compara la velocidad de hidrólisis del propanoato de metilo no sustituido con la de un éster que contenga el sustituyente de interés, mediante la siguiente expresión:

$$E_s = \log \left(\frac{k_x}{k_{CH_3}} \right)_A$$

donde k_x es la constante de hidrólisis del éster con el sustituyente X y KCH_3

la del éster no sustituido. El subíndice “A” indica que la reacción ocurre en condiciones de hidrólisis ácida. Cuanto más negativo sea el valor de E_s

E_s , mayor será el efecto estérico, es decir, mayor será la interferencia espacial generada por el sustituyente:

$$v_x = r_x - r_H$$

En esta relación, r_x corresponde al radio de Van der Waals del sustituyente⁴⁴, mientras que r_H representa el radio del átomo de hidrógeno utilizado como referencia. En el caso de sustituyentes poliatómicos, r_x se define como la distancia entre el eje principal de unión y el punto más próximo de la superficie de Van der Waals del grupo. Con esta definición, el hidrógeno presenta un valor $v = 0$, y cualquier sustituyente con un volumen mayor exhibirá valores positivos ($v > 0$).

La escala v propuesta por Charton es de carácter semiempírico, ya que resulta de la combinación de datos experimentales con consideraciones geométricas. Gracias a ello, este parámetro permite describir con mayor precisión la influencia espacial de grupos funcionales complejos, considerando no solo su tamaño absoluto, sino también su proyección tridimensional dentro del entorno molecular.

Descriptores de conteo estructural

Los descriptores de conteo estructural son parámetros que cuantifican, en forma de números enteros o fraccionarios, la presencia de determinados elementos estructurales dentro de una molécula. Estos elementos pueden corresponder a átomos individuales, tipos de enlaces, fragmentos estructurales definidos, grupos funcionales o incluso propiedades moleculares discretas susceptibles de ser contabilizadas, como el número de donadores o aceptores de puentes de hidrógeno.

La principal característica de este tipo de descriptores es que proporcionan una medida directa y objetiva de la composición estructural de la molécula, sin requerir cálculos físico-químicos complejos. Por ello, constituyen uno de los enfoques más simples y ampliamente utilizados en el análisis QSAR⁴².

Matemáticamente, un descriptor de conteo estructural D_i puede expresarse como:

$$D_i = \sum_{j=1}^n C_{ij}$$

donde: D_i es el descriptor asociado a una característica estructural específica i , C_{ij} representa la cantidad de dicha característica presente en la molécula j , n es el número total de ocurrencias contadas. Por ejemplo, si i corresponde a los donadores de puentes de hidrógeno, D_i será igual al número total de grupos capaces de donar un protón a través de un enlace de hidrógeno. De igual forma, si i representa enlaces dobles C=C, el descriptor indicará cuántas veces aparece este tipo de enlace en la molécula.

Descriptores topológicos y geométricos

Los descriptores topológicos y geométricos constituyen una categoría esencial dentro de los estudios QSAR, ya que describen de manera cuantitativa la conectividad, forma y disposición espacial de las moléculas. Estos parámetros permiten transformar la información estructural en valores numéricos que reflejan la organización interna, la ramificación y la complejidad tridimensional del sistema molecular, sin necesidad de recurrir a métodos cuánticos.

Los descriptores topológicos se derivan de la representación de la molécula como un grafo⁴², en el cual los átomos se consideran nodos y los enlaces, aristas. A partir de esta abstracción, se obtienen índices que describen la conectividad, simetría y flexibilidad estructural, como los índices de Randić, los números de Kier y Hall⁴², o los índices de distancias moleculares. Estos valores han demostrado correlacionarse eficazmente con propiedades fisicoquímicas y biológicas, proporcionando información estructural independiente de la conformación molecular.

En cambio, los descriptores geométricos incorporan información tridimensional y reflejan propiedades relacionadas con la forma molecular, la superficie accesible al solvente (ASA) y la distribución espacial de regiones polares y apolares⁴². Dentro de esta familia destacan los descriptores basados en la superficie polar accesible (PSA) y, particularmente, los DPSA (*Difference in Polar Surface Area*), que cuantifican la diferencia de área polar accesible entre distintas regiones o fragmentos de la molécula.

Los descriptores DPSA permiten evaluar la heterogeneidad en la distribución de la polaridad molecular, lo que resulta fundamental para interpretar fenómenos como la afinidad de unión ligando–receptor⁴², la permeabilidad a membranas biológicas y la biodisponibilidad. En conjunto, los descriptores topológicos y geométricos ofrecen una visión integral de la estructura global y el comportamiento espacial de las moléculas, aspectos determinantes en el reconocimiento molecular y la actividad biológica.

Descriptores basados en autocorrelación

Otra clase importante de descriptores moleculares está representada por los índices de autocorrelación, los cuales se basan en el análisis de la distribución y correlación de propiedades atómicas dentro de la molécula. Estos descriptores fueron introducidos por primera vez por Moreau y Broto (1980)⁴⁵, quienes propusieron un método para establecer relaciones entre átomos en función de su distancia topológica, permitiendo cuantificar cómo una determinada propiedad se distribuye a lo largo del grafo molecular. Posteriormente, Consonni y Todeschini⁴² ampliaron y sistematizaron este enfoque, ofreciendo una revisión exhaustiva de los diferentes tipos de descriptores de autocorrelación empleados en estudios QSAR.

En términos generales, los descriptores de autocorrelación consideran los átomos de una molécula como un conjunto discreto de puntos en el espacio, sobre los cuales se evalúa una propiedad atómica específica (por ejemplo, electronegatividad, polarizabilidad, masa atómica o potencial de ionización). A partir de estos valores, se calcula el grado de correlación de la propiedad entre pares de átomos separados por una distancia topológica k , denominada retardo (lag).

Matemáticamente, la autocorrelación puede expresarse de forma general como:

$$A_k = \sum_i \sum_j f_i f_j \text{ para todos los pares } (i, j) \text{ tales que } d_{ij} = k$$

donde f_i y f_j representan los valores de la propiedad atómica considerada en los átomos i y j , respectivamente, y d_{ij} es la distancia topológica (número de enlaces) entre ellos. Así, el descriptor A_k expresa cómo se correlacionan los valores de una propiedad a intervalos definidos por k . Estos descriptores integran simultáneamente información estructural y electrónica, combinando la topología molecular con la distribución de propiedades atómicas. Su principal utilidad radica en que permiten capturar cómo las características locales, como la electronegatividad o la polarizabilidad, influyen sobre regiones más alejadas dentro de la molécula, reflejando la coherencia estructural global del sistema. También existen variantes autocorrelacionales derivadas de los índices de Moran y de Geary, adaptadas al contexto molecular⁴⁶. En estas, los pares de átomos se consideran según una distancia topológica fija d . La autocorrelación de Moran⁴⁷ se basa en el producto de las desviaciones respecto de la media de la propiedad atómica, mientras que la autocorrelación de Geary⁴⁸ utiliza las diferencias cuadráticas entre pares de átomos; por ello, la primera refleja la covariación relativa a la media global y la segunda suele ser más sensible a discrepancias locales entre vecinos topológicos.

Descriptores electrotopológicos (ETA) y electrónicos derivados de electronegatividad

Los descriptores electrónicos y electrotopológicos constituyen una categoría avanzada de parámetros moleculares que buscan representar cuantitativamente cómo la distribución de electrones dentro de una molécula influye en su reactividad química y actividad biológica⁴². Estos descriptores permiten incorporar al análisis QSAR información que va más allá de la forma o el tamaño de la molécula, reflejando la manera en que los efectos electrónicos, como la densidad de carga, la deslocalización, la polarización o la capacidad de atraer electrones, determinan el comportamiento molecular.

En términos generales, estos descriptores expresan cómo los electrones se distribuyen e interactúan a lo largo de la estructura, considerando tanto los efectos locales (propios de un átomo o grupo funcional) como los globales (que involucran a toda la molécula). De este modo, permiten relacionar características puramente estructurales con propiedades como la afinidad de unión, la polaridad, la capacidad de formar enlaces de hidrógeno o la estabilidad electrónica del sistema.

Los descriptores electrotopológicos (ETA)⁴⁹ integran simultáneamente información topológica, es decir, cómo los átomos están conectados dentro del grafo molecular, y electrónica, relacionada con la capacidad de los átomos para influir en el entorno químico mediante sus cargas o pares de electrones. Su desarrollo se basa en el concepto de estado electrotopológico (E-state)⁵⁰, que asigna a cada átomo un valor que depende tanto de su naturaleza química (electronegatividad, tipo de hibridación, sustituyentes) como de su posición dentro de la red de enlaces.

Estos descriptores permiten visualizar cómo una perturbación electrónica en un punto de la molécula (por ejemplo, un grupo donador o atractor de electrones) puede afectar regiones más alejadas, capturando así fenómenos de resonancia, efectos inductivos o mesómeros. Un ejemplo relevante utilizado en este trabajo es el ETA_Beta_ns_d⁴⁹, que cuantifica la contribución de electrones no compartidos con capacidad de participar en resonancia dentro de sistemas aromáticos. Este parámetro permite estimar cómo los efectos de deslocalización electrónica pueden modificar la interacción con el sitio activo del receptor biológico.

Complementariamente, los descriptores electrónicos derivados de electronegatividad proporcionan una medida global de la tendencia electrónica de la molécula, al basarse en el promedio o la suma de las electronegatividades atómicas de sus componentes. Dichos descriptores reflejan la capacidad general del sistema para atraer o ceder electrones, lo cual influye directamente en la polaridad, la afinidad ligando-receptor y la estabilidad del complejo molecular. Un ejemplo es el SaasC⁵⁰, que representa la

suma de las electronegatividades de Sanderson correspondientes a los átomos de carbono de la molécula, proporcionando una visión resumida de la distribución electrónica total.

2.4.1.2 Softwares más utilizados

Entre los softwares más utilizados se encuentran Dragon, PaDEL-Descriptor, DataWarrior y librerías de código abierto implementadas en Python, como RDKit. Todos ellos permiten calcular cientos o miles de descriptores a partir de la estructura molecular (en formatos como SMILES o SDF), facilitando la construcción de bases de datos para análisis QSAR y machine learning.

Dentro de estas herramientas, destaca PaDEL-Descriptor⁵¹ por ser una alternativa de libre acceso ampliamente utilizada en la comunidad científica. Este programa, desarrollado en Java, permite calcular más de 1.800 descriptores moleculares correspondientes a 1D, 2D y huellas moleculares (*fingerprints*). Su diseño modular lo hace compatible con distintos sistemas operativos y permite una ejecución eficiente en grandes conjuntos de datos.

PaDEL combina rutinas implementadas en el paquete CDK (Chemistry Development Kit) con algoritmos propios, ofreciendo una amplia cobertura de descriptores fisicoquímicos, topológicos, autocorrelativos, geométricos y electrotopológicos. Además, su interfaz gráfica y su modo de línea de comandos facilitan la automatización del procesamiento masivo de moléculas.

Por otro lado, herramientas como Dragon y DataWarrior presentan interfaces más completas para la visualización, filtrado y análisis estadístico de los descriptores generados, mientras que RDKit constituye una alternativa flexible y programable que permite integrar el cálculo de descriptores dentro de flujos de trabajo en Python.

2.4.1.3 Generación estadística

La construcción de modelos QSAR requiere el establecimiento de una relación matemática robusta entre los descriptores moleculares seleccionados y la actividad biológica experimental. Este proceso implica la selección adecuada de variables, la construcción del modelo estadístico y la evaluación rigurosa de su capacidad predictiva.

Elección de variables

El cálculo de descriptores moleculares genera un conjunto amplio de variables que puede contener redundancias, ruido estadístico o información irrelevante. La inclusión de dichas variables puede comprometer la estabilidad y capacidad predictiva del modelo, por lo que resulta imprescindible realizar un pretratamiento adecuado de los datos.

En primer lugar, se eliminan descriptores constantes o con varianza cercana a cero, así como aquellos con valores faltantes. Posteriormente, se evalúa la colinealidad entre variables mediante matrices de correlación, eliminando descriptores altamente correlacionados (por ejemplo, $|r| > 0.85-0.95$), con el fin de reducir redundancias y evitar problemas de multicolinealidad⁵².

Una vez depurado el conjunto inicial, se procede a la selección del subconjunto óptimo de descriptores. Este proceso busca identificar variables relevantes, no redundantes y con significado fisicoquímico coherente. Entre los métodos comúnmente empleados se encuentran procedimientos análisis de correlación, regularización estadística y algoritmos evolutivos, como los algoritmos genéticos, que permiten explorar eficientemente combinaciones de variables y optimizar la calidad predictiva del modelo.

Construcción modelo estadística

Seleccionadas las variables relevantes, se establece la relación cuantitativa entre los descriptores moleculares y la actividad biológica mediante modelos estadísticos o técnicas de aprendizaje automático. En estudios QSAR, la regresión lineal múltiple (MLR) es ampliamente utilizada debido a su interpretabilidad y capacidad para cuantificar la contribución individual de cada descriptor.

De manera general, el modelo QSAR puede expresarse como:

$$\text{Actividad} = \beta_0 + \beta_1 D_1 + \beta_2 D_2 + \beta_3 D_3 + \dots + \beta_n D_n$$

D_n corresponden a los descriptores moleculares y β_n son los coeficientes de regresión que describen la contribución de cada variable al modelo.

El signo y la magnitud de los coeficientes permiten interpretar la influencia de cada descriptor sobre la actividad biológica. Para evitar el sobreajuste, es fundamental mantener una relación adecuada entre el

número de moléculas del conjunto de entrenamiento y el número de descriptores incluidos, recomendándose al menos cinco observaciones por variable⁵².

Además de la regresión lineal múltiple, pueden emplearse métodos multivariados o técnicas de aprendizaje automático cuando la relación entre estructura y actividades no lineal; sin embargo, la selección del método debe equilibrar capacidad predictiva e interpretabilidad.

Métodos de validación

La validación constituye un requisito esencial para garantizar la robustez, confiabilidad y capacidad predictiva de un modelo QSAR. El objetivo principal de esta etapa es demostrar que el modelo no solo reproduce adecuadamente los datos del conjunto de entrenamiento, sino que también es capaz de predecir con precisión la actividad biológica de compuestos nuevos. La validación es especialmente relevante en contextos regulatorios y científicos, donde se requiere evidencia de que las predicciones del modelo son confiables y reproducibles⁵³.

Los principios de validación aceptados internacionalmente establecen que un modelo QSAR debe presentar: un algoritmo definido y reproducible, medidas apropiadas de ajuste, robustez y predictividad.

Validación interna

La validación interna evalúa la estabilidad del modelo utilizando únicamente los datos del conjunto de entrenamiento (training set). El procedimiento más común es la validación cruzada, donde cada compuesto se excluye temporalmente del conjunto y su actividad es predicha por el modelo generado con los restantes datos. Entre los modelos más utilizados para validación cruzada es el *Leave one out* (LOO)⁵³.

El coeficiente de validación cruzada (Q^2) se emplea para estimar la capacidad predictiva interna del modelo:

$$Q^2 = 1 - \frac{\sum(Y_{obs(train)} - Y_{pred(train)})^2}{\sum(Y_{obs(train)} - \bar{Y}_{train})^2}$$

Donde: $Y_{obs(train)}$: Actividad observada, $Y_{pred(train)}$: Actividad predicha y $\bar{Y}_{train}$: El promedio de la actividad observada del *training set*.

El coeficiente Q^2 puede variar entre 1 y valores negativos; sin embargo, se considera que un modelo presenta capacidad predictiva interna aceptable cuando:

No obstante, este valor constituye una condición necesaria pero no suficiente para validar un modelo QSAR.

Además, se calcula el error estándar de predicción (SDEP):

$$Q^2 \geq 0.5$$

No obstante, este valor constituye una condición necesaria pero no suficiente para validar un modelo QSAR.

Además, se calcula el error estándar de predicción (RMSE):

$$RMSE_{train} = \sqrt{\frac{\sum(Y_{obs(train)} - Y_{pred(train)})^2}{n}}$$

donde n corresponde al número de compuestos del conjunto de entrenamiento. Valores bajos de RMSE indican mayor precisión predictiva.

Validación externa

La validación externa constituye el método más riguroso para evaluar la capacidad predictiva real del modelo, ya que utiliza compuestos independientes que no participaron en su construcción. Para

garantizar una evaluación representativa, el conjunto externo debe preservar la diversidad estructural del conjunto total y abarcar adecuadamente el espacio químico del sistema estudiado.
 El poder predictivo externo se evalúa mediante el coeficiente de determinación predictivo:

$$R^2_{test} = 1 - \frac{\sum(Y_{obs(test)} - Y_{pred(test)})^2}{\sum(Y_{obs(test)} - \bar{Y}_{train})^2}$$

Donde: $Y_{obs(test)}$: Actividad observada, $Y_{pred(trates)}$:Actividad predicha y $\bar{Y}_{train}$:El promedio de la actividad observada del *training set*.

Un modelo QSAR se considera predictivo cuando:

$$R^2_{test}>0.6$$

lo que indica que el modelo explica al menos el 60 % de la variabilidad biológica del conjunto externo.

$$RMSE_{test} = \sqrt{\frac{\sum(Y_{obs(test)} - Y_{pred(test)})^2}{n}}$$

donde n corresponde al número de compuestos del conjunto de evaluación (*test set*). Valores bajos de RMSE indican mayor precisión predictiva.
 Debido a esta limitación, Roy y colaboradores⁵² propusieron el parámetro r^2_m , el cual permite evaluar simultáneamente la precisión predictiva y la proximidad de las predicciones a la línea de identidad, proporcionando una medida más estricta del poder predictivo del modelo. Este parámetro penaliza desviaciones sistemáticas entre los valores observados y predichos, evitando que modelos con buena correlación pero con sesgos importantes sean considerados predictivos^{49,54}.

$$r^2_m = r^2 \left(1 - \sqrt{r^2 - r^2_0} \right)$$

donde r^2 corresponde al coeficiente de determinación entre valores observados y predichos, mientras que r^2_0 es el coeficiente de determinación obtenido al forzar la recta de regresión a pasar por el origen.
 De manera complementaria, se calcula el parámetro inverso:

$$r'^2_m = r^2 \left(1 - \sqrt{r^2 - r'^2_0} \right)$$

De manera análoga, r'^2_m se calcula intercambiando los ejes de valores observados y predichos y forzando nuevamente la regresión a pasar por el origen. La comparación entre ambos parámetros permite detectar desviaciones sistemáticas en las predicciones y evaluar la concordancia real del modelo.
 Para evaluar la confiabilidad del modelo se emplean los parámetros promedio y diferencia:

$$\overline{r^2_m} = \frac{r^2_m + r'^2_m}{2}$$

$$\Delta r^2_m = |r^2_m - r'^2_m|$$

Un modelo QSAR se considera predictivo cuando $\overline{r^2_m}>0.5$ y $\Delta r^2_m < 0.2$, criterios que indican simultáneamente buena precisión predictiva y ausencia de sesgos sistemáticos. Estos parámetros proporcionan una evaluación más rigurosa del poder predictivo del modelo, ya que integran la correlación, la exactitud y la concordancia entre valores experimentales y predichos

De manera complementaria, el error absoluto medio (MAE)⁵⁴ proporciona una estimación directa del error promedio entre los valores observados y predichos:

$$MAE_{test} = \frac{1}{n} \sum |Y_{obs(test)} - Y_{pred(test)}|$$

El MAE es menos sensible a valores atípicos que el RMSE y permite una interpretación más directa del error promedio del modelo.

Asimismo, el coeficiente de correlación concordante (CCC)^{49,54} permite evaluar simultáneamente la precisión y exactitud de las predicciones, considerando tanto la correlación entre los valores observados y predichos como su desviación respecto a la línea de identidad. Para su cálculo se emplean los valores experimentales y predichos del conjunto de validación externa:

$$CCC = \frac{2r\sigma_{obs,test}\sigma_{pred.test}}{\sigma_{obs,test} + \sigma_{pred.test} + (\mu_{obs,test} + \mu_{pred.test})^2}$$

donde *r* es el coeficiente de correlación entre los valores observados y predichos, *σ* representa las desviaciones estándar y *μ* las medias de los valores observados y predichos. Valores de CCC cercanos a la unidad indican excelente concordancia entre los valores experimentales y los predichos, mientras que valores bajos revelan desviaciones sistemáticas en las predicciones.

El uso conjunto de RMSE, MAE y CCC permite evaluar de manera integral la precisión, el error promedio y la concordancia entre los valores experimentales y predichos, proporcionando una descripción más completa del desempeño predictivo del modelo QSAR.

Prueba de aleatorización de la variable respuesta (Y-randomization)

A pesar de que un modelo QSAR pueda presentar buenos parámetros estadísticos y predictivos, existe la posibilidad de que las correlaciones obtenidas se deban al azar y no a relaciones reales entre la estructura química y la actividad biológica. Para descartar esta posibilidad, se emplea la prueba de aleatorización de la variable respuesta, conocida como Y-randomization⁵⁴.

Este procedimiento consiste en generar múltiples modelos QSAR utilizando los mismos descriptores moleculares, pero reemplazando los valores experimentales de actividad biológica por valores aleatorizados. Posteriormente, se comparan los parámetros estadísticos obtenidos para los modelos aleatorizados con los del modelo original.

Si el modelo QSAR es robusto y no producto del azar, los modelos generados con datos aleatorizados deben presentar valores significativamente menores de *Q*² en comparación con el modelo original. Por el contrario, si los modelos aleatorizados muestran parámetros estadísticos similares, ello indicaría que la correlación obtenida podría ser espuria.

La prueba de Y-randomization permite verificar que la relación encontrada entre los descriptores moleculares y la actividad biológica es real y no resultado de correlaciones fortuitas, constituyendo una herramienta fundamental para garantizar la confiabilidad del modelo QSAR.

2.4.2 Enfoques basados en el receptor

Por otro lado, las metodologías basadas en el receptor aprovechan la información estructural de la proteína blanco, generalmente obtenida por cristalografía de rayos X, RMN o predicciones de modelado por homología o IA. Dentro de este grupo destacan técnicas como el acoplamiento molecular (docking) y la dinámica molecular (MD).

El docking permite predecir la orientación y afinidad de un ligando dentro del sitio activo de la proteína, proporcionando información clave sobre posibles interacciones moleculares relevantes. Estas predicciones son fundamentales para entender el mecanismo de acción de los compuestos y optimizar su afinidad y selectividad.

Complementariamente, la dinámica molecular permite simular el comportamiento temporal del sistema ligando-receptor en condiciones fisiológicas. Gracias a esta técnica, es posible evaluar la estabilidad del complejo, la flexibilidad del sitio activo, y explorar fenómenos como la entrada o salida del ligando, reordenamientos conformacionales del receptor, o la formación de puentes de hidrógeno dinámicos. Esta

información es clave para validar y refinar los resultados del docking y obtener una visión más realista del proceso de reconocimiento molecular.

2.4.2.1 Docking molecular

El docking molecular es una técnica computacional utilizada para estudiar cómo un ligando interactúa con una proteína en un estado estacionario. Los ligandos ejercen actividad biológica al unirse a cavidades específicas del receptor, modulando su función al activarlo o inhibirlo. El objetivo central del docking es predecir la orientación (pose) y la afinidad del ligando en el sitio activo, identificando la conformación más favorable y permitiendo visualizar los aminoácidos involucrados y los tipos de interacciones intermoleculares que estabilizan el complejo.

El procedimiento se vuelve especialmente relevante en el diseño de nuevos fármacos, ya que constituye una herramienta rápida y eficiente para explorar, visualizar y predecir el potencial bioactivo de una molécula antes de pasar a etapas experimentales más complejas.

Tipos de docking

Docking rígido

Este tipo de docking tanto el ligando como el receptor se consideran estructuralmente fijos durante el proceso de acoplamiento. Bajo este enfoque, el algoritmo explora distintas orientaciones y posiciones del ligando dentro del sitio activo de la proteína, sin permitir cambios conformacionales en ninguno de los dos componentes, siguiendo el modelo clásico de llave–cerradura⁵⁵.

Este tipo de docking asume que la conformación del receptor representa adecuadamente el estado bioactivo de la proteína, lo que permite evaluar de manera eficiente la afinidad y el modo de unión del ligando. Debido a su bajo costo computacional y rapidez, el docking rígido es ampliamente utilizado en etapas iniciales del *virtual screening* , especialmente cuando se analizan grandes conjuntos de moléculas.

No obstante, este método presenta como principal limitación la ausencia de flexibilidad molecular, lo que puede afectar la precisión de las predicciones en sistemas donde la unión del ligando induce cambios estructurales relevantes en el receptor.

Docking flexible o Induced Fit

En el docking flexible o *induced fit*, tanto el ligando como el receptor pueden modificar su conformación con el fin de optimizar el acoplamiento. El ligando explora diferentes conformaciones mediante la rotación de sus enlaces, mientras que el sitio activo del receptor puede reorganizarse parcialmente para acomodar al ligando. Este enfoque representa de manera más realista el proceso de reconocimiento molecular, en el cual ambos componentes se adaptan mutuamente durante la formación del complejo ligando–receptor⁵⁶.

Para explorar las posibles orientaciones y conformaciones del ligando dentro del sitio activo, los programas de docking emplean diversos algoritmos de búsqueda conformacional, los cuales pueden clasificarse según su estrategia de exploración del espacio conformacional.

Algoritmo general del docking molecular

El docking molecular puede entenderse como un problema de optimización multidimensional, cuyo objetivo es identificar la conformación y orientación del ligando dentro del sitio activo del receptor que minimice la energía libre del complejo ligando–receptor⁵⁵. Para ello, el método de docking debe abordar simultáneamente dos desafíos fundamentales: la generación de conformaciones iniciales del ligando y la exploración eficiente del espacio conformacional con el fin de localizar mínimos energéticos relevantes. En una primera etapa, el ligando es descrito mediante un conjunto de grados de libertad traslacionales, rotacionales y torsionales, los cuales definen su espacio conformacional. A partir de estos grados de libertad, el algoritmo de docking genera un conjunto de conformaciones iniciales que sirven como puntos de partida para la búsqueda del mínimo energético.

Posteriormente, se aplica un algoritmo de búsqueda conformacional que explora el espacio de posibles orientaciones y conformaciones del ligando dentro del sitio activo del receptor. Este proceso busca identificar las poses más favorables desde el punto de vista energético, evaluadas mediante una función de puntuación (*scoring function*). Dado que el espacio conformacional crece exponencialmente con el número de enlaces rotables del ligando y la flexibilidad del receptor, la exploración exhaustiva resulta computacionalmente inviable, lo que ha motivado el desarrollo de distintos enfoques algorítmicos. En este contexto, los algoritmos de búsqueda utilizados en docking molecular pueden clasificarse, de manera general, en algoritmos sistemáticos y algoritmos estocásticos, los cuales difieren en la forma en que exploran el espacio conformacional y en el balance entre precisión y costo computacional.

Algoritmos sistemáticos

Los algoritmos sistemáticos se basan en la exploración determinística y exhaustiva del espacio conformacional del ligando dentro del sitio activo del receptor. En este enfoque, los grados de libertad del ligando, tales como rotaciones de enlaces, orientaciones espaciales y posiciones traslacionales, se muestrean de manera ordenada siguiendo intervalos predefinidos⁵⁵.

Entre los métodos sistemáticos más utilizados se encuentran la búsqueda exhaustiva (*exhaustive search*) y los algoritmos de construcción incremental, en los cuales el ligando es fragmentado a lo largo de sus enlaces rotables y ensamblado progresivamente dentro del sitio de unión. Este tipo de algoritmos permite una cobertura amplia del espacio conformacional, lo que resulta ventajoso para sistemas pequeños o ligandos con baja flexibilidad estructural.

Sin embargo, el principal inconveniente de los algoritmos sistemáticos es su alto costo computacional, el cual aumenta de forma exponencial con el número de grados de libertad del ligando. Por esta razón, su aplicación suele limitarse a etapas de refinamiento o a sistemas donde el tamaño del espacio conformacional es manejable.

Algoritmos estocásticos

Los algoritmos estocásticos constituyen uno de los enfoques más ampliamente utilizados en docking molecular, debido a su balance entre eficiencia computacional y cobertura conformacional. A diferencia de los métodos sistemáticos, estos algoritmos exploran el espacio conformacional mediante procesos aleatorios controlados, evaluando cada nueva conformación en función de criterios probabilísticos o funciones de aptitud.

Entre los algoritmos estocásticos más representativos se encuentran los algoritmos genéticos (*Genetic Algorithms*, GA)⁵⁵ y los métodos de Monte Carlo (MC). En los algoritmos genéticos, las posibles conformaciones del ligando se tratan como una población de individuos que evoluciona a través de operadores como mutación, recombinación y selección, favoreciendo aquellas poses con mejor energía de interacción. Por su parte, los métodos de Monte Carlo generan modificaciones aleatorias en la conformación del ligando, aceptando o rechazando cada cambio según criterios energéticos y probabilísticos.

Debido a la naturaleza aleatoria del muestreo, los algoritmos estocásticos suelen ejecutarse múltiples veces para asegurar la convergencia de los resultados. Este enfoque es especialmente adecuado para ligandos flexibles y sistemas complejos, y es el fundamento de programas de docking ampliamente utilizados, como AutoDock y GOLD.

Basadas en conocimiento

Los algoritmos basados en el conocimiento utilizan información derivada de estructuras experimentales previamente resueltas de complejos proteína–ligando. En este enfoque, las interacciones observadas en bases de datos estructurales se emplean para derivar potenciales estadísticos que guían la colocación y evaluación del ligando dentro del sitio activo⁵⁷.

Este tipo de algoritmos asume que las configuraciones de unión más favorables son aquellas que reproducen patrones de interacción frecuentes en complejos conocidos, tales como distancias interatómicas características, geometrías de enlaces de hidrógeno y contactos hidrofóbicos recurrentes. Como resultado, los métodos basados en el conocimiento pueden ofrecer una evaluación rápida y razonable de la afinidad de unión, especialmente cuando se dispone de información estructural abundante y de alta calidad.

No obstante, su desempeño depende fuertemente de la calidad y representatividad del conjunto de datos utilizado para construir los potenciales estadísticos, lo que puede limitar su aplicabilidad en sistemas poco caracterizados o novedosos.

Algoritmo Autodock 4

AutoDock 4 emplea una función de puntuación semiempírica diseñada para estimar la energía libre de unión del complejo ligando–receptor. Esta función se basa en parámetros previamente definidos y combina contribuciones energéticas asociadas a interacciones intermoleculares, tales como fuerzas de van der Waals, enlaces de hidrógeno, interacciones electrostáticas, efectos de desolvatación y términos entrópicos relacionados con la flexibilidad del ligando.

Para identificar la pose de unión más favorable, AutoDock 4 implementa distintos algoritmos de búsqueda conformacional. Entre ellos se incluyen métodos de búsqueda local, algoritmos genéticos (GA) y el algoritmo genético lamarckiano (Lamarckian Genetic Algorithm, LGA), el cual combina la exploración global propia de los algoritmos genéticos con procedimientos de optimización local. Esta estrategia permite mejorar la eficiencia en la exploración del espacio conformacional y aumentar la probabilidad de localizar configuraciones energéticamente favorables.

Motor de búsqueda genético Lamarckiano

AutoDock 4 utiliza el algoritmo genético lamarckiano (LGA) como uno de sus métodos principales de búsqueda, el cual está diseñado para explorar eficientemente el espacio conformacional y posicional de un ligando flexible dentro del sitio activo de un receptor rígido⁵⁸.

En este contexto, cada posible conformación del ligando se representa mediante un cromosoma, el cual codifica explícitamente los parámetros geométricos necesarios para definir su posición, orientación y conformación interna. Dicho cromosoma está compuesto, como mínimo, por siete variables reales: tres coordenadas cartesianas (x, y, z) que describen la traslación del ligando dentro del volumen delimitado por el *grid box*, y cuatro parámetros que definen un cuaternión, donde tres componentes determinan el eje de rotación y la cuarta corresponde al ángulo de giro. Adicionalmente, se incorpora un parámetro real por cada enlace rotacional del ligando, el cual describe los ángulos torsionales asociados a su flexibilidad interna.

El proceso mediante el cual el cromosoma es transformado en una estructura tridimensional explícita del ligando se denomina *mapping*, y consiste en la reconstrucción completa de las coordenadas atómicas a partir de los valores codificados en el genotipo. Una vez realizado el *mapping*, cada individuo es evaluado mediante la función de energía de AutoDock, calculada a partir de mapas de interacción previamente generados por AutoGrid, lo que permite estimar la energía libre de unión ligando–receptor. Esta energía constituye la función de fitness, donde valores más bajos indican interacciones más favorables.

El algoritmo comienza con la generación de una población inicial de cromosomas, cuyos parámetros se asignan de manera aleatoria dentro de los límites definidos por el espacio de búsqueda. A partir de esta población, el algoritmo evoluciona a lo largo de múltiples generaciones mediante un ciclo iterativo que comprende las siguientes etapas: selección, cruce, mutación y evaluación.

Durante la selección, los individuos son elegidos para reproducirse de manera proporcional a su fitness, de modo que aquellos con energías de unión más favorables tienen mayor probabilidad de contribuir a la siguiente generación. Posteriormente, se aplica un operador de crossover de dos puntos, en el cual se intercambian segmentos completos del cromosoma entre dos individuos parentales, garantizando que los cortes se realicen únicamente entre genes completos, evitando así la fragmentación de parámetros continuos.

A continuación, los cromosomas resultantes pueden experimentar mutaciones, las cuales consisten en la modificación aleatoria de uno o más genes mediante la adición de un valor perturbador. En AutoDock, estas perturbaciones siguen una distribución de Cauchy, lo que favorece principalmente cambios pequeños, pero permite ocasionalmente variaciones de mayor magnitud, promoviendo una exploración eficiente del espacio conformacional.

Una característica distintiva del LGA es la incorporación de una etapa de búsqueda local (*local search*), aplicada a una fracción de los individuos de cada generación. En esta etapa, se realiza una optimización

local de los parámetros del cromosoma con el objetivo de refinar la energía del ligando en un mínimo cercano. El carácter “lamarckiano” del algoritmo radica en que las mejoras obtenidas durante la búsqueda local son incorporadas al cromosoma, de modo que el individuo optimizado sustituye al original y transmite dicha información a generaciones posteriores. Este ciclo evolutivo se repite hasta alcanzar el número máximo de generaciones o evaluaciones de energía especificadas por el usuario. Al finalizar todas las corridas independientes del algoritmo, las soluciones obtenidas son agrupadas mediante un análisis de *clustering* conformacional, generalmente basado en el RMSD, permitiendo identificar las poses más representativas y energéticamente favorables del ligando en el sitio de unión del receptor. Con el fin de clarificar la diferencia entre el espacio genotípico y fenotípico, así como el rol de la búsqueda local en el enfoque lamarckiano, la Figura 9 presenta un esquema conceptual del funcionamiento del algoritmo.

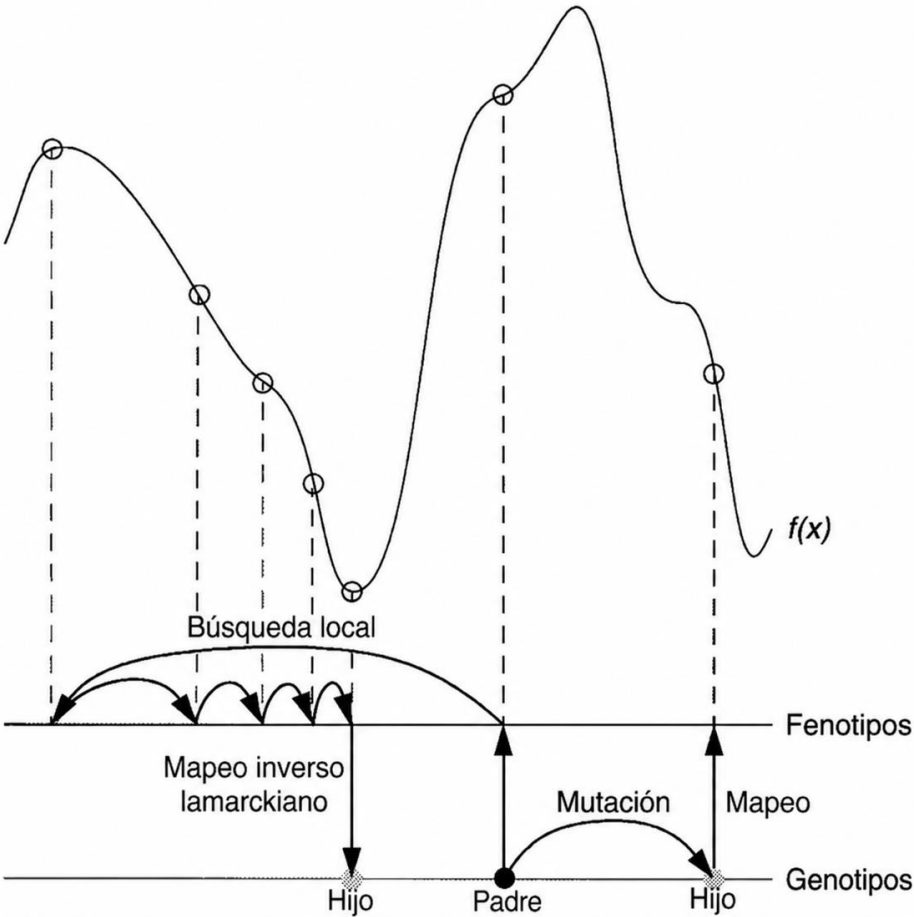


Figura 9 Representación esquemática del Algoritmo Genético Lamarckiano (LGA) empleado en AutoDock⁵⁸.

Para poder entender la diferencia entre el espacio genotípico y fenotípico, así como el rol de la búsqueda local en el enfoque lamarckiano, la Figura presenta un esquema conceptual del funcionamiento del algoritmo. En la Figura, la parte superior representa el espacio fenotípico, donde se evalúa la función objetivo $f(x)$, mientras que la parte inferior corresponde al espacio genotípico, en el cual se codifican los parámetros del individuo. El proceso de mapping transforma la información codificada en el genotipo en una solución explícita que puede ser evaluada en el espacio fenotípico. A partir de un individuo parental (parent), la mutación genera un nuevo individuo (child) en el espacio genotípico. Posteriormente, mediante local search, esta solución puede refinarse en el espacio fenotípico para aproximarse a un mínimo local de energía. En el enfoque lamarckiano, la mejora obtenida durante esta optimización local se transfiere nuevamente al genotipo mediante un proceso de inverse mapping, de modo que el individuo optimizado reemplaza al original y puede heredar esa mejora a generaciones posteriores.

Algoritmo Induced fit docking de Schrodinger 2025

Este tipo de algoritmo utilizado por Schrodinger⁵⁹ a una estrategia avanzada de acoplamiento molecular que incorpora la flexibilidad tanto del ligando como del receptor, en contraste con los enfoques de docking rígido basados en el modelo llave–cerradura, donde ambas entidades conservan su conformación durante el proceso de acoplamiento. En el docking inducido, el ligando no solo explora sus grados de libertad conformacionales, sino que además puede inducir cambios estructurales locales en el sitio de unión del receptor, permitiendo una adaptación mutua que refleja de mejor manera la naturaleza dinámica del reconocimiento molecular . Este protocolo, implementado en la suite Maestro de Schrödinger, integra los algoritmos Glide y Prime y se estructura en un flujo de trabajo secuencial que consta de cuatro etapas principales. En primer lugar, se realiza un muestreo inicial del ligando, donde se reducen los radios de Van der Waals de los átomos del ligando y del receptor, y se sustituyen temporalmente residuos cercanos al sitio activo por alaninas, con el objetivo de evitar bloqueos estéricos y generar un conjunto amplio y diverso de poses iniciales mediante acoplamiento con gradiente. En una segunda etapa, se lleva a cabo el muestreo de la proteína utilizando el algoritmo Prime, que permite la optimización conformacional de las cadenas laterales del receptor en presencia del ligando, seguido de un cálculo de la energía de solvatación del complejo mediante un modelo de solvente implícito basado en el método de Born generalizado. Posteriormente, las estructuras que presentan una energía favorable son sometidas a una tercera etapa de remuestreo del ligando, donde se realiza un nuevo docking empleando Glide en modos de mayor precisión, tales como SP (Standard Precision) o XP (Extra Precision), con el fin de refinar las poses y eliminar falsos positivos. Finalmente, la evaluación energética del complejo ligando–receptor se efectúa mediante una función de puntuación combinada, donde el puntaje final corresponde a la suma ponderada del GlideScore y el PrimeScore, asignando un mayor peso al primero para reducir la influencia del ruido estérico asociado al refinamiento proteico. Este enfoque permite obtener poses de unión más realistas y energéticamente consistentes, siendo especialmente adecuado para sistemas en los que el sitio activo no es completamente rígido o presenta una elevada plasticidad conformacional, como se ilustra en el esquema general del algoritmo de docking inducido (Figura 10).

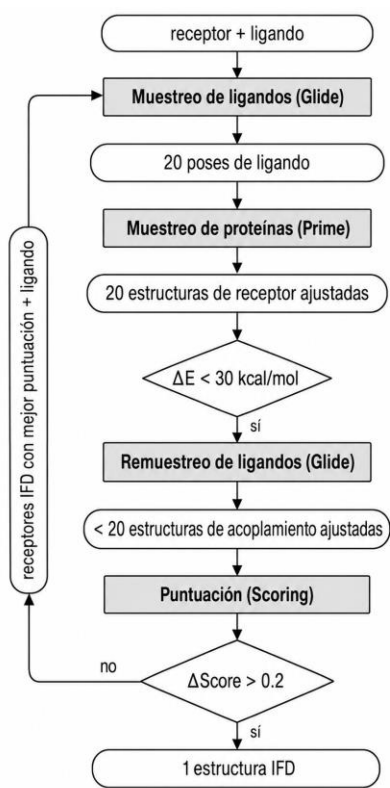


Figura 10 Diagrama de flujo del algoritmo de docking inducido⁵⁹.

2.4.2.2Virtual Screening

Esta técnica computacional se basa en el *docking* molecular, en el cual se realizan múltiples acoplamientos moleculares variando únicamente el ligando, mientras que el receptor se mantiene constante. Los ligandos evaluados pueden corresponder tanto a moléculas conocidas y previamente reportadas en la literatura como a moléculas diseñadas in silico. En este enfoque, el receptor

generalmente se considera rígido, mientras que el ligando permanece flexible, principalmente debido a las limitaciones y requerimientos computacionales asociados al cálculo.

El *virtual screening* permite evaluar de manera exhaustiva un conjunto amplio de ligandos, facilitando la selección de un subconjunto de moléculas con potencial afinidad hacia un determinado *target* biológico. Ejemplos representativos de compuestos identificados mediante esta estrategia incluyen Aggrastat® (antagonista del receptor GP IIb/IIIa), PRX-00023 (agonista del receptor 5-HT_{1A}), DMP450 (inhibidor de la proteasa del VIH-1) y Cevoglitazar (agonista dual de PPAR- α/γ)⁶⁰.

Base de datos de ligandos

Para llevar a cabo un proceso de virtual screening, es fundamental disponer de bases de datos en línea que contengan estructuras moleculares confiables y bien curadas. Existen diversas plataformas que proporcionan este tipo de información, entre las cuales destacan ZINC y PubChem.

La base de datos ZINC permite descargar bibliotecas de compuestos listas para docking, filtradas según propiedades fisicoquímicas relevantes, como lipofilicidad y peso molecular. Por su parte, PubChem ofrece opciones de filtrado más amplias, incluyendo parámetros como masa molecular, número de donadores y aceptores de enlaces de hidrógeno, entre otros. No obstante, una limitación práctica de PubChem es que las descargas masivas están restringidas a paquetes de hasta 500 000 moléculas por archivo. Estos conjuntos pueden obtenerse en formato SDF, el cual conserva la información estructural tridimensional necesaria para estudios de modelamiento molecular.

Convertidor SDF a PDBQT

Para convertir archivos en formato SDF a PDBQT es necesario emplear herramientas de conversión que preserven la información estructural tridimensional de las moléculas. Una de las herramientas más utilizadas es Open Babel⁶¹, un software de código abierto ampliamente empleado en química computacional para la interconversión de formatos químicos y el procesamiento de estructuras moleculares.

Adicionalmente, existen scripts y utilidades desarrolladas por la comunidad científica y disponibles en plataformas de intercambio de código, como GitHub, que permiten automatizar y optimizar este proceso para el manejo de grandes volúmenes de datos. Entre estas herramientas se encuentran convertidores diseñados para facilitar el procesamiento masivo de moléculas, manteniendo la integridad estructural necesaria para estudios de docking molecular⁶².

Autogrid4

AutoGrid4 es el programa encargado de generar las grillas de energía utilizadas durante el *docking* molecular, que viene en conjunto con autodock4. Este proceso calcula mapas tridimensionales que describen las posibles interacciones entre el receptor y diferentes tipos de átomos del ligando, permitiendo evaluar de manera eficiente la energía de unión durante el acoplamiento.

La generación de estas grillas es necesaria tanto en estudios de docking molecular convencionales como en protocolos de *virtual screening*. Sin embargo, cuando se trabaja con grandes bibliotecas de compuestos, es fundamental definir adecuadamente los parámetros del grid box para cubrir exhaustivamente el sitio activo y asegurar una evaluación consistente para todos los ligandos. Además, AutoGrid4 calcula mapas energéticos específicos para cada tipo atómico presente, lo que permite una estimación precisa de las interacciones ligando–receptor.

AutoGrid4 forma parte del ecosistema de AutoDock y se encuentra integrado en implementaciones aceleradas como AutoDock4GPU, facilitando la generación de mapas energéticos compatibles con ejecuciones de alto rendimiento.

Para la ejecución de un gran número de *dockings* moleculares es indispensable contar con un software lo suficientemente potente que permita realizar estos cálculos en una escala de tiempo razonable. En este contexto, AutoDock4 GPU⁶³ incorpora el uso de unidades de procesamiento gráfico (GPU) al cálculo del docking molecular, lo que permite acelerar significativamente los tiempos de simulación en comparación con implementaciones basadas exclusivamente en CPU.

El uso de AutoDock4 GPU se inicia a partir de la generación de una grilla proteína–ligando, la cual contiene la información energética asociada a la interacción del receptor con cada tipo de átomo del ligando. Esta etapa es realizada mediante el programa AutoGrid4, encargado de calcular y almacenar los mapas de energía potencial necesarios para el proceso de docking.

AutoDock4 GPU permite realizar tanto dockings con receptor rígido como dockings con receptor flexible, siendo este último computacionalmente más demandante debido al aumento en los grados de libertad del sistema y, por consiguiente, al mayor costo computacional asociado.

2.4.2.3 Dinámica molecular

La dinámica molecular es una técnica de simulación computacional que permite describir la evolución temporal de un sistema molecular mediante el modelamiento explícito de las interacciones covalentes y no covalentes entre sus átomos. Esta metodología se fundamenta en la mecánica clásica, donde el movimiento de cada partícula se calcula a partir de las fuerzas derivadas de un campo de fuerza previamente parametrizado. De este modo, esta técnica posibilita el análisis detallado del comportamiento atómico y conformacional de las moléculas a lo largo del tiempo, así como la evaluación de su estabilidad estructural y de sus interacciones bajo distintas condiciones termodinámicas.

En la práctica, una simulación de dinámica molecular no comienza directamente con la propagación de las ecuaciones de movimiento. Previamente, el sistema debe ser llevado a un estado físicamente consistente y numéricamente estable mediante una serie de etapas preparatorias, cuyo objetivo es garantizar que las trayectorias generadas sean representativas del comportamiento real del sistema y no estén dominadas por artefactos iniciales.

Estas etapas incluyen, en primer lugar, la minimización de energía, orientada a eliminar contactos estéricos desfavorables y tensiones geométricas excesivas introducidas durante la construcción inicial del modelo. Posteriormente, se asignan velocidades iniciales a las partículas, generalmente de acuerdo con una distribución de Maxwell–Boltzmann asociada a una temperatura inicial controlada. A continuación, el sistema es sometido a un proceso de calentamiento gradual, en el cual la temperatura se incrementa progresivamente hasta alcanzar el valor objetivo, evitando así inestabilidades numéricas asociadas a aumentos abruptos de energía cinética.

Una vez alcanzadas condiciones cercanas al equilibrio, se lleva a cabo una fase de equilibración termodinámica, durante la cual se aplican mecanismos de control de temperatura y presión que permiten estabilizar las variables macroscópicas del sistema. Solo después de completar estas etapas el sistema se considera adecuadamente preparado para la simulación de producción, en la cual se generan las trayectorias utilizadas para el análisis posterior (Figura 11).

Este marco general proporciona el contexto necesario para comprender los fundamentos físicos de la dinámica molecular, los modelos de interacción utilizados para describir las fuerzas interatómicas y los métodos numéricos empleados para la integración de las ecuaciones de movimiento, los cuales se describen en los apartados siguientes^{64,65}.

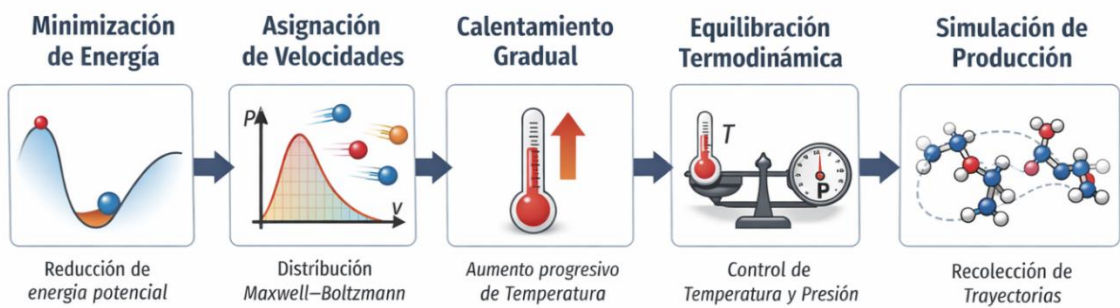


Figura 11 Etapas preparatorias y de producción en una simulación de dinámica molecular biomolecular.(Imagen generada con inteligencia artificial).

2.4.2.3.1 Tiempos de simulación

La dinámica molecular permite describir el comportamiento de sistemas biomoleculares a lo largo del tiempo, abarcando distintos niveles de movimiento según la escala temporal considerada⁶⁶(Figura 12). En la escala de femtosegundos (10^{-15} s) se observan vibraciones de enlaces y movimientos atómicos rápidos. En el rango de picosegundos a nanosegundos (10^{-12} – 10^{-9} s) predominan rotaciones de enlaces, fluctuaciones locales de cadenas laterales y ajustes iniciales del ligando dentro del sitio de unión. A escalas de nanosegundos a microsegundos (10^{-9} – 10^{-6} s) se evidencian cambios conformacionales más relevantes, como la reorganización de loops y la estabilización del complejo proteína–ligando. Finalmente, en el rango de microsegundos a milisegundos o incluso segundos, pueden ocurrir procesos de mayor magnitud, como transiciones conformacionales globales, eventos de unión/desunión o cambios alostéricos.

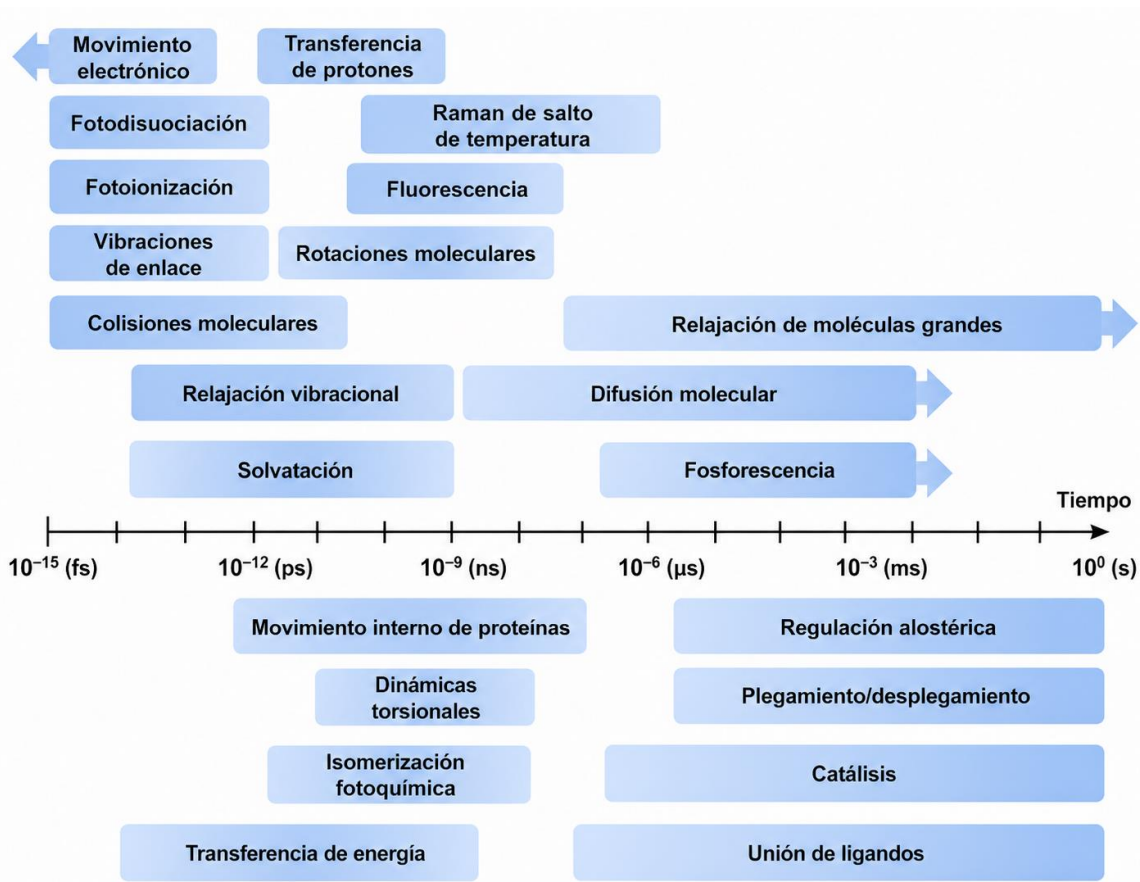


Figura 12 Escala de fenómenos moleculares y tiempo en que pueden observarse ⁶⁶.

2.4.2.3.2 Mecánica clásica

La dinámica molecular describe el comportamiento de los sistemas moleculares mediante la aplicación de los principios de la mecánica clásica, en particular a través de las leyes de Newton que gobiernan el movimiento de las partículas. En este enfoque, la evolución temporal de cada átomo se determina a partir de las fuerzas que actúan sobre él, las cuales dependen de la configuración geométrica instantánea del sistema.

A partir de estas leyes, se define una función de energía potencial que depende de las posiciones de los núcleos, dando origen a la denominada superficie de energía potencial. Dicha superficie puede conceptualizarse como un espacio multidimensional generado punto a punto, donde cada punto representa una disposición específica del sistema molecular y posee asociada una energía potencial determinada.

La superficie de energía potencial constituye el elemento central de la dinámica molecular, ya que a partir de su gradiente se obtienen las fuerzas que gobiernan el movimiento de los átomos. En consecuencia, la forma de esta superficie determina las trayectorias atómicas y la evolución global de la simulación a lo largo del tiempo.

2.4.2.3.3 Leyes de Newton

Un aspecto fundamental a considerar es que la dinámica molecular se basa en una descripción clásica del movimiento de los sistemas moleculares, por lo que su formulación se sustenta en las leyes de Newton. En particular, la segunda ley de Newton es la más relevante para la comprensión de esta técnica, ya que establece la relación entre la fuerza que actúa sobre una partícula, su masa y la aceleración resultante:

$$F = ma$$

Donde F =Fuerza representa la fuerza aplicada, m la masa de la partícula y a su aceleración. En dinámica molecular, las fuerzas que actúan sobre los átomos no se introducen de forma directa, sino que se obtienen a partir de una función de energía potencial que depende de la posición de las partículas. Esta relación se expresa mediante la derivada de la energía potencial con respecto a la posición:

$$F(x) = -\frac{dU}{dx}$$

Donde U corresponde a la energía potencial del sistema y x a la posición de la partícula. Considerando que la aceleración es la derivada temporal de la velocidad y que la velocidad corresponde a la derivada temporal de la posición, la segunda ley de Newton puede expresarse como un conjunto de ecuaciones diferenciales:

$$\begin{aligned} \frac{dx}{dt} &= v \\ \frac{dv}{dt} &= \frac{F(x)}{m} \end{aligned}$$

Estas ecuaciones son equivalentes a la forma de segundo orden:

$$-m \frac{d^2x}{dt^2} = \frac{dU}{dx}$$

En las expresiones anteriores, x representa la posición, v la velocidad, t el tiempo y m la masa de la partícula. Estas ecuaciones diferenciales describen la evolución temporal del sistema molecular y constituyen la base teórica de la dinámica molecular. En la práctica, su resolución se realiza mediante métodos de integración numérica, los cuales permiten obtener las trayectorias de posición y velocidad de las partículas a lo largo del tiempo^{64,65}.

2.4.2.3.4 Series de Taylor

En matemáticas, una serie de Taylor corresponde a una expansión que permite aproximar una función diferenciable en torno a un punto mediante una suma de potencias enteras. Esta expansión se construye

a partir de las derivadas de la función evaluadas en un punto de referencia a , y puede expresarse de forma general como:

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} (x - a)^n$$

Donde $f^{(n)}(a)$ representa la derivada de orden n de la función evaluada en el punto a , y $(x - a)^n$ corresponde al término de orden n de la expansión. En aplicaciones prácticas, la serie se trunca a un número finito de términos, obteniendo así una aproximación de la función cuya precisión depende del orden de truncamiento.

Como ejemplo ilustrativo, consideremos la función $f(x)=e^x$, cuya serie de Taylor alrededor del punto $a=0$ se expresa como:

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

Si la serie se trunca al segundo orden, se obtiene la aproximación:

$$e^x = 1 + x + \frac{x^2}{2!}$$

Esta expresión permite estimar el valor de e^x para valores pequeños de x con una precisión razonable, ilustrando cómo la serie de Taylor aproxima una función mediante un número finito de términos.

2.4.2.3.5 Campos de fuerza

Para integrar las ecuaciones de movimiento en una simulación de dinámica molecular es necesario conocer las fuerzas que actúan sobre cada átomo del sistema. Dado que dichas fuerzas se obtienen como el gradiente de la energía potencial, resulta imprescindible disponer de una expresión funcional que permita evaluar la energía del sistema en función de las posiciones atómicas en cada paso de la simulación. Para este propósito se emplean los campos de fuerza.

Un campo de fuerza corresponde a un conjunto de funciones matemáticas y parámetros empíricos que permiten aproximar la energía potencial total de un sistema molecular. Este enfoque proporciona una descripción eficiente de las interacciones atómicas, permitiendo el cálculo rápido de fuerzas y haciendo viable la simulación de sistemas de gran tamaño durante escalas de tiempo relevantes.

En el marco de los campos de fuerza clásicos, la energía potencial del sistema se descompone en contribuciones que representan distintos tipos de interacciones. De manera general, la energía potencial total puede expresarse como la suma de términos enlazantes y no enlazantes:

$$U_{Sistema} = U_{enlazante} + U_{no\ enlazante}$$

Los términos enlazantes describen las interacciones asociadas a enlaces químicos y a la geometría local de la molécula, y pueden expresarse como:

$$U_{enlazante} = U_{enlace} + U_{ángulos} + U_{Torsiones}$$

Por otra parte, los términos no enlazantes representan las interacciones de largo alcance entre átomos no conectados directamente, y se expresan como:

$$U_{no\ enlazantes} = U_{Electrostatica} + U_{Van\ Der\ Waals}$$

Esta descomposición no pretende describir de manera exacta la naturaleza cuántica de las interacciones moleculares, sino proporcionar una aproximación funcional que permita evaluar la energía potencial y las fuerzas de manera eficiente durante la simulación⁶⁴.

En aplicaciones biomoleculares, la forma funcional general del campo de fuerza se mantiene constante, mientras que los parámetros específicos varían según el conjunto desarrollado. Entre los campos de fuerza más utilizados se encuentran AMBER, CHARMM, GROMOS y OPLS, los cuales han sido parametrizados para reproducir propiedades estructurales, energéticas y termodinámicas de sistemas biológicos.

El campo de fuerza OPLS (*Optimized Potentials for Liquid Simulations*) ha sido ampliamente empleado en estudios de interacción proteína-ligando debido a su cuidadosa parametrización basada en propiedades termodinámicas de fases líquidas y en cálculos cuánticos de alta precisión⁶⁷. Las versiones más recientes, como OPLS3 y OPLS4, incorporan mejoras en la descripción de interacciones no enlazantes, torsiones y

cargas parciales, lo que permite reproducir con mayor exactitud energías de interacción, geometrías conformacionales y afinidades de unión en sistemas biomoleculares complejos. Debido a estas características, OPLS se ha consolidado como una opción robusta para estudios de dinámica molecular orientados al diseño racional de fármacos, ya que ofrece un equilibrio adecuado entre precisión fisicoquímica y eficiencia computacional.

Términos enlazantes

Estiramiento de enlaces

En los campos de fuerza clásicos, el estiramiento de enlaces covalentes se describe habitualmente mediante una aproximación armónica, la cual puede interpretarse como el truncamiento de la serie de Taylor de la energía potencial alrededor de la distancia de equilibrio del enlace. Bajo esta aproximación, la energía asociada al estiramiento de un enlace $A-B$ se expresa como:

$$U(r_{AB}) = \frac{1}{2} k_{AB} (r_{AB} - r_{AB,eq})^2$$

donde r_{AB} corresponde a la distancia interatómica, $r_{AB,eq}$ es la distancia de equilibrio y k_{AB} es la constante de fuerza del enlace.

Esta forma funcional constituye una aproximación válida para pequeñas desviaciones respecto a la geometría de equilibrio, que es el régimen de interés en la mayoría de las simulaciones de dinámica molecular clásica de sistemas biomoleculares. Aunque para grandes elongaciones del enlace esta aproximación pierde validez, su simplicidad y eficiencia computacional justifican su uso generalizado en campos de fuerza diseñados para describir estructuras estables sin ruptura de enlaces, como se discute en la literatura especializada⁶⁴.

Doblamientos de ángulos de enlace

En los campos de fuerza clásicos, las deformaciones angulares asociadas al doblamiento de ángulos de valencia se describen mediante una aproximación armónica, análoga a la empleada para el estiramiento de enlaces. Esta aproximación puede interpretarse como el truncamiento de una expansión de Taylor de la energía potencial alrededor del ángulo de equilibrio.

La energía asociada a la deformación de un ángulo $A-B-C$ se expresa como:

$$U(\theta_{ABC}) = \frac{1}{2} k_{ABC} (\theta_{ABC} - \theta_{ABC,eq})^2$$

Donde θ_{ABC} corresponde al ángulo formado por los enlaces AB y BC , $\theta_{ABC,eq}$ es el valor de equilibrio del ángulo y k_{ABC} la constante de fuerza asociada al ángulo.

De esta forma funcional proporciona una descripción adecuada de las variaciones energéticas asociadas a pequeñas deformaciones angulares alrededor del equilibrio, que constituyen el régimen de interés en la mayoría de las simulaciones de dinámica molecular clásica de sistemas biomoleculares. Aunque modelos más complejos pueden incorporar términos anarmónicos o múltiples mínimos, la aproximación armónica ofrece un compromiso adecuado entre realismo físico y eficiencia computacional, razón por la cual es ampliamente utilizada en campos de fuerza biomoleculares⁶⁴.

Diedros o torsiones

En los campos de fuerza clásicos, las interacciones torsionales describen la rotación alrededor de un enlace químico, la cual se define a partir de cuatro átomos consecutivos $A-B-C-D$. El ángulo torsional, o ángulo diedro, corresponde al ángulo formado entre los planos definidos por los átomos $A-B-C$ y $B-C-D$.

Dado que la rotación alrededor de un enlace presenta una naturaleza periódica, la energía asociada a este movimiento se modela mediante funciones periódicas, usualmente expresadas como una serie de términos cosenoidales. De manera general, la energía torsional puede escribirse como:

$$U(\omega_{ABCD}) = \sum_j \frac{V_{j,ABCD}}{2} [1 + \cos j\omega_{ABCD} - \delta_j]$$

Donde ω_{ABCD} es el ángulo torsional, V_j corresponde a la amplitud del término torsional asociado a la periodicidad j , y δ_j es el desfase de fase. La inclusión de múltiples términos permite describir perfiles energéticos complejos asociados a barreras rotacionales.

Este tratamiento resulta fundamental para reproducir correctamente la conformación molecular y la flexibilidad estructural en simulaciones de dinámica molecular, especialmente en sistemas biomoleculares donde la rotación alrededor de enlaces simples juega un papel clave en la estabilidad y dinámica conformacional. Debido a su eficiencia computacional y capacidad descriptiva, este tipo de potencial torsional es ampliamente utilizado en campos de fuerza biomoleculares clásicos⁶⁴.

Términos no enlazantes

Electrostáticos

En los campos de fuerza clásicos, las interacciones electrostáticas entre átomos se describen mediante la asignación de cargas parciales a cada átomo del sistema. Estas cargas representan una aproximación del reparto electrónico promedio y permiten modelar las interacciones de largo alcance de manera eficiente.

La energía electrostática entre dos átomos A y B separados por una distancia r_{AB} se expresa mediante la ley de Coulomb:

$$U_{AB}^{elec} = \frac{q_A q_B}{4\pi\epsilon_0 \epsilon_r r_{AB}}$$

donde q_A y q_B corresponden a las cargas parciales de los átomos, ϵ_0 es la permitividad del vacío y ϵ_r es la constante dieléctrica relativa del medio.

En la mayoría de los campos de fuerza biomoleculares clásicos, las cargas parciales se consideran parámetros fijos, independientes de la geometría instantánea del sistema. Esta aproximación permite un cálculo eficiente de las interacciones electrostáticas y resulta adecuada para describir sistemas donde no se consideran efectos de polarización electrónica explícita⁶⁴.

Van Der Waals

Las interacciones de van der Waals describen fuerzas no enlazantes de corto alcance que actúan entre átomos no conectados directamente. Estas interacciones resultan de la combinación de un término atractivo asociado a fuerzas de dispersión y un término repulsivo originado en la superposición de densidades electrónicas, la cual da lugar a repulsión de Pauli a distancias cortas.

En los campos de fuerza clásicos, estas contribuciones se modelan habitualmente mediante el potencial de Lennard–Jones, cuya forma funcional se expresa como:

$$U_{AB}^{vdW}(r_{AB}) = 4\epsilon_{AB} \left[\left(\frac{\sigma_{AB}}{r_{AB}} \right)^{12} - \left(\frac{\sigma_{AB}}{r_{AB}} \right)^6 \right]$$

Donde r_{AB} es la distancia entre los átomos A y B , ϵ_{AB} representa la profundidad del pozo de energía potencial y σ_{AB} corresponde a la distancia a la cual el potencial se anula.

El término proporcional a r^{-6} describe la atracción por dispersión, mientras que el término r^{-12} representa de manera empírica la repulsión de corto alcance. Esta forma funcional proporciona una descripción eficiente y suficientemente precisa de las interacciones no enlazantes en simulaciones de dinámica molecular clásica, siendo ampliamente utilizada en campos de fuerza biomoleculares⁶⁴.

Las ecuaciones de movimiento que gobiernan la dinámica molecular corresponden a un sistema de ecuaciones diferenciales derivadas de la segunda ley de Newton que, para sistemas moleculares complejos, no admite una solución analítica. En consecuencia, la evolución temporal del sistema debe determinarse mediante métodos de integración numérica, los cuales permiten propagar las posiciones y velocidades de las partículas en intervalos discretos de tiempo Δt , generando trayectorias que describen el comportamiento dinámico del sistema.

El objetivo fundamental de la integración numérica en dinámica molecular es estimar la posición de las partículas en un tiempo posterior $t+\Delta t$ a partir de la información disponible en el tiempo t . Dado que el computador solo puede operar de forma discreta, resulta necesario aproximar la solución continua de las ecuaciones de movimiento mediante expresiones que relacionen el estado actual del sistema con su estado en un tiempo cercano.

Una estrategia ampliamente utilizada para este propósito consiste en emplear expansiones en serie de Taylor, las cuales permiten aproximar el valor de una función en un tiempo próximo a partir de su valor actual y de sus derivadas temporales. En el contexto de la dinámica molecular, esta aproximación se aplica a las coordenadas espaciales de las partículas. Así, la expansión de la posición $x(t)$ hacia un tiempo posterior puede expresarse como:

$$x(t + \Delta t) = x(t) + v(t)\Delta t + \frac{1}{2}a(t)\Delta t^2 + \theta(\Delta t^3) + \theta(\Delta t^4) + \dots \theta(\Delta t^n)$$

En la expresión anterior, el símbolo $O(\Delta t^n)$ corresponde a la notación de orden, la cual se utiliza para indicar el comportamiento de los términos de orden superior en la expansión. Estos términos representan contribuciones cuya magnitud es proporcional a Δt^n y que dependen de derivadas temporales de orden mayor. A medida que el paso de integración Δt disminuye, la contribución de estos términos se vuelve progresivamente menor, por lo que en la práctica se truncan para obtener una aproximación eficiente sin introducir nueva información física.

De manera análoga, la expansión de la posición hacia un tiempo anterior se escribe como:

$$x(t - \Delta t) = x(t) - v(t)\Delta t + \frac{1}{2}a(t)\Delta t^2 + \theta(\Delta t^3) + \theta(\Delta t^4) + \dots \theta(\Delta t^n)$$

En estas expresiones, $v(t)$ representa la velocidad y $a(t)$ la aceleración de la partícula en el instante t . La aceleración desempeña un rol central en dinámica molecular, ya que se obtiene directamente a partir de la fuerza que actúa sobre la partícula, la cual deriva de la energía potencial del sistema. Por esta razón, resulta conveniente construir esquemas de integración que dependan únicamente de posiciones y aceleraciones, evitando el uso de derivadas de orden superior y reduciendo la propagación del error numérico.

La consideración simultánea de las expansiones temporales hacia adelante y hacia atrás permite introducir simetría temporal en el esquema de integración. Al sumar ambas expresiones, los términos que dependen de la velocidad se cancelan, obteniéndose una relación que depende exclusivamente de la posición y de la aceleración:

$$x(t + \Delta t) = 2x(t) - x(t - \Delta t) + a(t)\Delta t^2 + 2\theta(\Delta t^3) + 2\theta(\Delta t^4) + \dots 2\theta(\Delta t^n)$$

Si bien el algoritmo de Verlet permite una integración robusta de las trayectorias, la formulación original no proporciona de manera explícita las velocidades en el mismo instante temporal que las posiciones. Sin embargo, las velocidades son necesarias para el cálculo de propiedades dinámicas y termodinámicas, tales como la energía cinética, la temperatura del sistema y la aplicación de termostatos.

Por esta razón, se emplean variantes del método original, como el algoritmo de *Velocity Verlet*, el cual conserva las ventajas del esquema de Verlet para la actualización de las posiciones y, adicionalmente, permite una evaluación consistente de las velocidades. En este método, las posiciones se actualizan según:

$$x(t + \Delta t) = x(t) + v(t)\Delta t + \frac{1}{2}a(t)\Delta t^2$$

y las velocidades se corrigen posteriormente considerando la aceleración en los instantes inicial y final del paso de integración:

$$v(t + \Delta t) = v(t) + \frac{1}{2}[a(t) + a(t + \Delta t)]\Delta t$$

De este modo, el método de Velocity Verlet combina una integración estable de las posiciones con una actualización coherente de las velocidades, manteniendo las propiedades numéricas deseables del algoritmo original y permitiendo el cálculo eficiente de observables dinámicos y termodinámicos⁶⁵.

Ensamblajes termodinámicos

En dinámica molecular, la evolución temporal del sistema puede interpretarse en el marco de la mecánica estadística mediante el concepto de ensamble termodinámico, el cual define el conjunto de condiciones macroscópicas bajo las cuales se describe un sistema. Un ensamble especifica qué variables termodinámicas se mantienen constantes durante la simulación y permite relacionar las trayectorias microscópicas con propiedades observables a nivel macroscópico.

El ensamble más fundamental en dinámica molecular clásica es el microcanónico (NVE), en el cual el número de partículas N , el volumen V y la energía total E permanecen constantes. En este caso, el sistema evoluciona de acuerdo con las ecuaciones de movimiento de Newton sin intercambio de energía con el entorno, representando un sistema aislado.

Con el fin de simular condiciones más cercanas a las experimentales, es habitual emplear el ensamble canónico (NVT), donde el número de partículas y el volumen se mantienen constantes, mientras que la temperatura T se controla mediante el acoplamiento del sistema a un baño térmico. Este enfoque permite regular la energía cinética promedio del sistema y mantener una temperatura definida durante la simulación.

De manera análoga, el ensamble isobárico–isotérmico (NPT) mantiene constantes el número de partículas, la presión P y la temperatura, permitiendo que el volumen del sistema fluctúe. Esto resulta especialmente relevante en simulaciones de sistemas biomoleculares en solución, donde se busca reproducir condiciones físicas cercanas a las biológicas reales.

En simulaciones de complejos ligando–receptor, el uso del ensamble NPT permite ajustar adecuadamente la densidad del solvente y aliviar tensiones internas del sistema, asegurando una distribución realista del agua y de los iones alrededor del complejo. Estas características son fundamentales para describir correctamente la estabilidad estructural y las interacciones no covalentes que gobiernan el reconocimiento molecular.

Por esta razón, el ensamble NPT se emplea habitualmente durante las etapas de equilibración de simulaciones de dinámica molecular biomolecular, previo al análisis detallado de la dinámica y estabilidad del complejo⁶⁵.

Control de temperatura y presión

En el ensamble isobárico–isotérmico (NPT), la dinámica molecular se lleva a cabo manteniendo constantes el número de partículas, la temperatura y la presión del sistema, permitiendo que el volumen fluctúe de manera controlada. Para lograr estas condiciones, es necesario incorporar mecanismos adicionales a la integración directa de las ecuaciones de movimiento de Newton, los cuales regulan la temperatura y la presión durante la simulación.

El control de la temperatura en simulaciones NPT se fundamenta en su relación con la energía cinética promedio del sistema. En la práctica, este control se implementa mediante termostatos que simulan el acoplamiento del sistema a un baño térmico externo. Durante las etapas de equilibración, es común emplear esquemas de acoplamiento suave, como el termostato de Berendsen, el cual permite ajustar progresivamente la temperatura hacia un valor objetivo y disipar energías excesivas asociadas a configuraciones iniciales poco realistas. No obstante, este método no reproduce rigurosamente el ensamble canónico, por lo que en la fase de producción se prefieren termostatos más formales, como el método de Nosé–Hoover, que generan distribuciones estadísticas apropiadas.

De manera complementaria, el control de la presión en el ensamble NPT se realiza mediante barostatos que ajustan el volumen del sistema, generalmente a través del escalamiento de las dimensiones de la celda de simulación bajo condiciones periódicas de contorno. Este procedimiento permite mantener una presión promedio constante y asegurar una densidad adecuada del solvente. De forma análoga al control térmico, durante la equilibración se emplean barostatos de relajación rápida, mientras que en etapas posteriores se utilizan métodos más precisos que garantizan una correcta descripción estadística del ensamble.

En simulaciones de sistemas biomoleculares en solución, el uso del ensamble NPT resulta fundamental durante la etapa de equilibración, ya que permite obtener una distribución realista del solvente y de los iones alrededor del complejo ligando–receptor, así como aliviar tensiones internas del sistema y estabilizar su estructura global. Una vez alcanzadas condiciones termodinámicas adecuadas, es habitual continuar la simulación en un ensamble de volumen constante para el análisis detallado de la dinámica molecular.

Por estas razones, el ensamble NPT, junto con un control adecuado de temperatura y presión, constituye un marco esencial para la preparación y equilibración de sistemas biomoleculares en simulaciones de dinámica molecular, asegurando que las trayectorias analizadas sean representativas de condiciones físicas y fisiológicas reales⁶⁵.

Dinámica molecular en sistemas biomoleculares

La dinámica molecular aplicada a sistemas biomoleculares constituye una herramienta fundamental para el estudio detallado de las interacciones proteína–ligando a nivel atómico. A diferencia del docking molecular, que proporciona una representación estática del complejo, la dinámica molecular permite describir la evolución temporal del sistema bajo condiciones termodinámicas controladas, incorporando explícitamente el efecto del solvente, la flexibilidad conformacional de la proteína y los movimientos internos del ligando.

En el contexto del diseño racional de fármacos, la dinámica molecular se utiliza para evaluar la estabilidad estructural del complejo proteína–ligando, analizar la persistencia de interacciones no covalentes (puentes de hidrógeno, interacciones hidrofóbicas, apilamientos π – π , interacciones iónicas), y explorar posibles reacomodos conformacionales inducidos por la unión del ligando. Este enfoque permite refinar hipótesis estructurales obtenidas mediante docking y detectar interacciones transitorias que no pueden ser capturadas en modelos estáticos.

Un sistema biomolecular típico en solución incluye: (i) la macromolécula (proteína), (ii) el ligando, (iii) moléculas de solvente explícitas —usualmente agua modelada mediante potenciales clásicos— y (iv) iones para neutralizar la carga total del sistema y reproducir condiciones fisiológicas. El uso de condiciones periódicas de contorno evita efectos de borde y simula un entorno continuo, mientras que el empleo de ensambles termodinámicos apropiados (frecuentemente NPT en estudios biomoleculares) garantiza una densidad y presión realistas del solvente⁶⁶.

Análisis de trayectorias

El análisis de trayectorias constituye una etapa fundamental en los estudios de dinámica molecular, ya que permite interpretar la evolución temporal del sistema simulado. Una simulación de dinámica molecular genera un conjunto de trayectorias que describen las posiciones y movimientos de todos los átomos a lo largo del tiempo, proporcionando información detallada sobre la estabilidad estructural y el comportamiento conformacional del sistema.

El examen de estas trayectorias permite evaluar cómo se comporta el sistema completo o regiones específicas bajo determinadas condiciones termodinámicas. Para ello, se emplean diversas herramientas de análisis que permiten generar representaciones gráficas y métricas cuantitativas, facilitando la interpretación de fenómenos como estabilidad estructural, flexibilidad, interacciones intermoleculares y cambios conformacionales.

Desviación cuadrática media (RMSD)

La desviación cuadrática media (*Root Mean Square Deviation*, RMSD) es uno de los parámetros más utilizados en el análisis de trayectorias de dinámica molecular. Este indicador cuantifica la desviación promedio de las posiciones atómicas de una estructura a lo largo del tiempo respecto a una configuración de referencia, permitiendo evaluar la estabilidad estructural y los cambios conformacionales del sistema durante la simulación⁶⁹.

Habitualmente, el RMSD se calcula tomando como referencia la estructura inicial de la simulación ($t = 0$), lo que permite determinar el grado de desviación del sistema a lo largo del tiempo. En sistemas proteína–ligando, también es común calcular el RMSD del ligando tras alinear previamente la proteína, con el fin de evaluar la estabilidad del ligando dentro del sitio activo. Asimismo, puede analizarse el RMSD del esqueleto proteico o de residuos específicos del sitio de unión para estudiar cambios estructurales locales relevantes.

El RMSD se determina mediante la siguiente expresión:

$$RMSD(t) = \sqrt{\frac{1}{N} \sum_{i=1}^N |r_i(t) - r_i^{ref}|^2}$$

donde N corresponde al número de átomos considerados, $r_i(t)$ representa la posición del átomo i en el tiempo t , y r_i^{ref} corresponde a la posición del átomo en la estructura de referencia.

Fluctuación cuadrática media (RMSF)

La fluctuación cuadrática media (*Root Mean Square Fluctuation*, RMSF) es un parámetro ampliamente utilizado en el análisis de trayectorias de dinámica molecular para caracterizar la flexibilidad local de una estructura⁶⁹. Este indicador describe el grado de fluctuación de las posiciones atómicas a lo largo del tiempo respecto a su posición promedio temporal, permitiendo identificar regiones con mayor movilidad estructural.

El RMSF se calcula para cada átomo o residuo mediante la siguiente expresión:

$$RMSF = \sqrt{\langle (r_i(t) - \langle r_i \rangle)^2 \rangle}$$

Donde $r_i(t)$ presenta el vector posición del átomo i en el tiempo t , $\langle r_i \rangle$ corresponde a la posición promedio temporal de dicho átomo durante la simulación, y los corchetes angulares indican el promedio temporal sobre todas las configuraciones analizadas.

Este análisis permite identificar regiones estructurales con mayor flexibilidad, tales como bucles, extremos terminales o residuos involucrados en el reconocimiento molecular. En sistemas proteína–ligando, el RMSF resulta especialmente útil para evaluar cómo la unión del ligando modula la dinámica de la proteína, pudiendo rigidizar regiones estructurales clave o, en algunos casos, aumentar su flexibilidad. Estos cambios en la movilidad local proporcionan información relevante sobre los efectos conformacionales inducidos por el ligando y su posible impacto en la estabilidad y funcionalidad del sitio activo.

Análisis de distancias

El análisis de distancias intermoleculares constituye una herramienta fundamental para evaluar interacciones estructurales y funcionales durante simulaciones de dinámica molecular. Este análisis permite monitorear la variación temporal de la distancia entre pares específicos de átomos o residuos, proporcionando información sobre la estabilidad estructural, la formación de interacciones no covalentes y los cambios conformacionales del sistema.

En sistemas proteína–ligando, el seguimiento de distancias entre el ligando y residuos clave del sitio activo permite evaluar la estabilidad del complejo y la persistencia de interacciones relevantes, tales como enlaces de hidrógeno, interacciones electrostáticas y contactos hidrofóbicos. La conservación de

distancias estables a lo largo de la simulación suele indicar una interacción favorable y una adecuada acomodación del ligando dentro del bolsillo de unión.

Asimismo, el análisis de distancias entre residuos catalíticos o estructurales permite identificar reorganizaciones conformacionales que podrían influir en la actividad biológica. Variaciones significativas en estas distancias pueden reflejar procesos de apertura o cierre del sitio activo, adaptación estructural inducida por ligandos o cambios dinámicos asociados a la función enzimática.

En la proteasa PLpro de SARS-CoV-2, este análisis adquiere especial relevancia debido al papel regulador del loop BL2, una región flexible situada próxima al sitio activo que modula el acceso de sustratos e inhibidores. El monitoreo de distancias entre residuos del loop BL2, entre los que destacan Tyr268, Gln269 y Tyr264, y el ligando o residuos del sitio activo permite evaluar su grado de apertura o cierre y, por tanto, su influencia directa sobre la actividad enzimática. La estabilización de conformaciones cerradas del loop BL2⁷⁰ suele asociarse con la inhibición efectiva de la enzima, mientras que configuraciones más abiertas pueden facilitar el acceso del sustrato y favorecer la actividad catalítica.

De este modo, el análisis temporal de estas distancias proporciona información directa sobre los mecanismos de modulación estructural inducidos por ligandos y su posible impacto en la función biológica de PLpro.

Análisis ángulos diedros

El análisis de ángulos diedros constituye una herramienta fundamental para evaluar cambios conformacionales locales en proteínas durante simulaciones de dinámica molecular. Estos ángulos describen la orientación relativa entre planos formados por cuatro átomos consecutivos y permiten caracterizar la conformación tridimensional tanto del esqueleto peptídico como de las cadenas laterales. En proteínas, los ángulos diedros del esqueleto principal, ϕ (phi) y ψ (psi), determinan la estructura secundaria y su estabilidad a lo largo del tiempo. El monitoreo de estas variables permite detectar desviaciones conformacionales, transiciones estructurales y regiones con elevada flexibilidad.

Por otra parte, los ángulos diedros de las cadenas laterales (χ) proporcionan información sobre la orientación espacial de residuos funcionales y su participación en interacciones moleculares. Variaciones en estos ángulos pueden reflejar reorganizaciones del sitio activo, adaptaciones estructurales inducidas por la unión del ligando o movimientos dinámicos asociados a la función biológica.

En conjunto, el análisis de los ángulos diedros permite identificar cambios conformacionales sutiles que no siempre son evidentes mediante parámetros estructurales globales, contribuyendo a una comprensión más profunda de la dinámica funcional del sistema biomolecular.

Alcances y limitaciones

La dinámica molecular se ha consolidado como una herramienta computacional robusta para el estudio de interacciones ligando–receptor, ya que permite describir, con resolución atómica y temporal, los fenómenos físicos que gobiernan la estabilidad y el reconocimiento molecular. A través de la simulación explícita de las fuerzas interatómicas, esta metodología permite analizar no solo las interacciones directas entre el ligando y el sitio activo, sino también los efectos estructurales inducidos en la proteína, incluyendo cambios conformacionales, movimientos de *loops* y reorganizaciones dinámicas que pueden influir en la actividad biológica.

Este enfoque posibilita la identificación de interacciones clave, como enlaces de hidrógeno, interacciones hidrofóbicas y contactos electrostáticos, así como la evaluación de la estabilidad estructural del complejo a lo largo del tiempo. Asimismo, la dinámica molecular permite inferir mecanismos estructurales asociados a procesos de activación o inhibición, aportando información relevante para comprender el modo de acción de agonistas, antagonistas o inhibidores.

Además, el análisis de trayectorias permite calcular contribuciones energéticas y estimaciones de energía libre de unión, proporcionando un marco cuantitativo para evaluar afinidad molecular. Estas capacidades convierten a la dinámica molecular en una herramienta valiosa dentro del diseño racional de fármacos, permitiendo optimizar estructuras químicas y predecir el efecto de modificaciones estructurales sobre la estabilidad y las interacciones del complejo.

No obstante, la dinámica molecular presenta limitaciones inherentes. Una de las principales es su alto costo computacional, especialmente cuando se requieren simulaciones prolongadas para explorar escalas temporales más amplias o eventos conformacionales raros. Si bien los recursos computacionales actuales permiten realizar simulaciones de decenas a cientos de nanosegundos en equipos convencionales, el acceso a escalas de microsegundos puede requerir infraestructura de alto rendimiento.

Adicionalmente, la dinámica molecular clásica se basa en aproximaciones que no describen la formación o ruptura de enlaces químicos, lo que limita su aplicabilidad en procesos reactivos. Los resultados dependen también de la calidad del campo de fuerza utilizado y de la correcta parametrización del sistema. Asimismo, el muestreo conformacional puede ser incompleto para procesos lentos, y las condiciones simuladas representan aproximaciones simplificadas del entorno biológico real.

2.5 Administración, distribución, metabolismo, excreción y toxicidad (ADME/Tox)

2.5.1 Farmacocinética y la importancia de incluirla en estrategia de desarrollo *in silico*.

En el descubrimiento de fármacos, la afinidad frente a una diana biológica constituye solo una parte del problema. Para que una molécula pueda avanzar como candidato terapéutico, no basta con que sea activa: también debe presentar propiedades adecuadas de absorción, distribución, metabolismo y excreción, junto con un perfil toxicológico aceptable. Por esta razón, la evaluación ADME/Tox se ha incorporado tempranamente al diseño racional de fármacos como una herramienta de priorización, ya que permite identificar de manera anticipada compuestos con mayor probabilidad de fracaso en etapas posteriores del desarrollo⁷¹.

Como primera aproximación a la selección de compuestos con potencial farmacológico, resulta útil considerar criterios clásicos de drogabilidad, entre los que destacan la regla de Lipinski y el criterio de Veber. La regla de Lipinski, también conocida como “regla de cinco”³⁷, establece que los problemas de absorción o permeación oral son más probables cuando una molécula presenta más de 5 donadores de puente de hidrógeno, más de 10 aceptores de puente de hidrógeno, un peso molecular superior a 500 Da y un valor de logP superior a 5. Aunque estas reglas no constituyen criterios absolutos de exclusión, siguen siendo ampliamente utilizadas como filtros tempranos para evaluar la compatibilidad de una estructura con administración oral. Posteriormente, Veber y colaboradores demostraron que la biodisponibilidad oral no depende únicamente de la masa molecular o de la lipofilidad, sino también de la flexibilidad conformacional y de la polaridad global de la molécula⁷², proponiendo como parámetros orientativos un máximo de 10 enlaces rotables y una superficie polar topológica no mayor de 140 Å². En conjunto, ambos enfoques proporcionan una base sencilla pero útil para estimar el carácter drug-like de una molécula y para priorizar compuestos antes de análisis ADME/Tox más específicos.

2.5.2 Parámetros para evaluar absorción

La absorción describe el paso del compuesto desde el sitio de administración hacia la circulación sistémica. En estudios *in silico* tempranos suele aproximarse mediante parámetros como la absorción intestinal humana (HIA), la permeabilidad en monocapas de células Caco-2 y la interacción con transportadores de eflujo como la glicoproteína-P (P-gp)^{73,74}. La HIA ofrece una estimación global de la capacidad de absorción intestinal; la permeabilidad Caco-2 se utiliza de manera extensa como aproximación experimental e *in silico* al paso a través del epitelio intestinal; y la P-gp resulta especialmente relevante porque puede expulsar compuestos desde el enterocito hacia la luz intestinal, reduciendo su absorción neta y afectando además su distribución tisular. En este sentido, conocer si una molécula actúa como sustrato o inhibidor de P-gp permite anticipar restricciones de biodisponibilidad y posibles interacciones farmacocinéticas.

2.5.3 Parámetros para evaluar distribución

La distribución se refiere al reparto del compuesto entre plasma y tejidos una vez absorbido. Dentro de este bloque, la permeación a la barrera hematoencefálica (BBB)⁷⁵ es relevante porque puede ser deseable o indeseable según el blanco terapéutico; en el caso de antivirales no dirigidos al sistema nervioso central, una alta penetración cerebral puede convertirse en una alerta adicional de seguridad. A ello se suman la unión a proteínas plasmáticas (PPB), la fracción libre en plasma (Fu)⁷⁶ y el volumen de distribución en

estado estacionario (VDss)⁷⁷. La PPB influye en la cantidad de fármaco retenido en el compartimento vascular; la fracción libre representa la porción farmacológicamente disponible para distribuirse, interactuar con su blanco, metabolizarse o eliminarse; y el VDss ofrece una medida global de la tendencia del compuesto a permanecer en plasma o repartirse ampliamente en tejidos.

2.5.4 Parámetros para evaluar metabolismo

El metabolismo comprende la biotransformación enzimática del compuesto, principalmente a nivel hepático, y constituye uno de los determinantes más importantes de estabilidad, duración del efecto e interacciones fármaco-fármaco. En humanos, el sistema citocromo P450 (CYP450) desempeña un papel central en la oxidación de xenobióticos, por lo que la predicción de comportamiento como sustrato o inhibidor de isoformas concretas resulta especialmente útil en etapas tempranas. Entre las isoformas más analizadas se encuentran CYP1A2, CYP2B6, CYP2C8, CYP2C9, CYP2C19, CYP2D6 y CYP3A4, debido a su amplia contribución al metabolismo de fármacos clínicamente utilizados⁷⁸. La inhibición de estas enzimas puede traducirse en acumulación del compuesto o en interacciones relevantes con otros medicamentos, mientras que el comportamiento como sustrato permite anticipar qué rutas metabólicas podrían condicionar su depuración.

2.5.5 Parámetros para evaluar excreción

La excreción representa la eliminación del compuesto y de sus metabolitos desde el organismo. En el contexto de predicción *in silico*, dos descriptores particularmente útiles son el *clearance* plasmático y la vida media⁷⁹. El *clearance* expresa la eficiencia con la que el organismo elimina un compuesto desde la circulación, mientras que la vida media ofrece una medida práctica del tiempo durante el cual la concentración plasmática disminuye por eliminación y distribución. Ambos parámetros son relevantes porque un compuesto con depuración excesivamente rápida puede no sostener concentraciones terapéuticas adecuadas, mientras que una vida media muy prolongada puede favorecer acumulación y aumentar el riesgo de toxicidad o de interacciones.

2.5.6 Parámetros para evaluar toxicidad

La toxicidad constituye una dimensión crítica en la selección de candidatos, ya que una molécula puede mostrar buena afinidad por la proteína blanco y, aun así, fracasar si presenta riesgos de seguridad relevantes. En este contexto, la evaluación *in silico* no reemplaza los ensayos experimentales, pero sí permite identificar alertas tempranas que orientan la priorización de compuestos. Entre estos endpoints, el ensayo de Ames^{80,81} se utiliza como una aproximación inicial a la mutagenicidad, es decir, a la capacidad de un compuesto para inducir mutaciones. La genotoxicidad corresponde a un concepto más amplio, ya que se refiere al potencial de dañar el material genético en general, mientras que la carcinogenicidad estima la probabilidad de que una exposición prolongada favorezca procesos asociados al desarrollo de cáncer.

En cuanto a la toxicidad sobre órganos y sistemas específicos, DILI (drug-induced liver injury) corresponde al riesgo de lesión hepática inducida por fármacos durante su uso^{82,83}, mientras que la hepatotoxicidad humana representa una estimación más general del potencial de daño sobre el hígado en humanos. De forma análoga, la nefrotoxicidad se refiere al posible daño renal, la neurotoxicidad al compromiso del sistema nervioso, la ototoxicidad al daño sobre estructuras auditivas, la hematotoxicidad a alteraciones tóxicas en células o funciones de la sangre, y la toxicidad respiratoria al potencial de afectar negativamente al sistema respiratorio. Estos endpoints permiten detectar si un compuesto, aun siendo activo, podría comprometer órganos especialmente sensibles durante su administración.

Dentro de la seguridad cardiovascular, el endpoint hERG evalúa la probabilidad de que una molécula bloquee el canal de potasio hERG^{84,85}, evento asociado con prolongación del intervalo QT y con riesgo de arritmias ventriculares potencialmente graves. En ADMETlab también aparece la variante hERG (10 μ M), que corresponde al mismo riesgo interpretado bajo un criterio experimental estandarizado a esa concentración, por lo que ambos descriptores apuntan a la misma preocupación de cardiotoxicidad, aunque desde condiciones de clasificación distintas.

Otros endpoints⁸⁶ informan sobre tolerancia general o daño local. La toxicidad oral aguda estima el riesgo de efectos tóxicos severos tras una exposición oral única o de corta duración. La sensibilización cutánea indica la probabilidad de que una molécula desencadene una respuesta alérgica en la piel. La irritación ocular se refiere a daño inflamatorio o molesto generalmente reversible en el ojo, mientras que la corrosión ocular implica un daño más severo, profundo y potencialmente irreversible.

Por otra parte, la citotoxicidad en líneas celulares como HEK293, A549 y RPMI-8226 ^{87,88}funciona como una medida general de daño celular en modelos biológicos. En términos simples, estos endpoints estiman si el compuesto tiende a comprometer la viabilidad de células cultivadas en laboratorio; por ello, no describen por sí solos una toxicidad clínica específica, pero sí actúan como una señal útil de agresividad biológica temprana. HEK293 corresponde a una línea derivada de riñón embrionario humano, A549 a una línea epitelial de adenocarcinoma pulmonar, y RPMI-8226 a una línea humana de mieloma múltiple. El descriptor FDAMDD (*FDA Maximum Recommended Daily Dose*) ⁸⁶se relaciona con la dosis diaria máxima recomendada y se utiliza como una aproximación práctica al umbral de exposición tolerable en humanos. En términos sencillos, no describe un órgano específico, sino que entrega una señal global sobre si un compuesto podría asociarse a limitaciones de seguridad al considerar la magnitud de exposición diaria aceptable. En conjunto, todos estos endpoints no deben interpretarse como pruebas definitivas, sino como herramientas predictivas que ayudan a distinguir entre moléculas con mejor o peor perfil de seguridad en etapas tempranas del diseño racional de fármacos

2.5.7 Predicción in silico y uso del ADMETlab 3.0

ADMETlab 3.0 ⁸⁶ es una herramienta especialmente útil para la evaluación temprana de candidatos, no solo por su amplia cobertura de propiedades fisicoquímicas, criterios de química medicinal, variables ADME, endpoints toxicológicos y reglas de toxicóforos, sino también porque, según sus autores, la plataforma mejora el desempeño y la robustez predictiva mediante una arquitectura multitarea basada en *directed message passing neural networks* (DMPNN) combinada con descriptores moleculares. Además, incorpora estimación de incertidumbre, acceso por API y funciones de apoyo a la decisión, lo que fortalece su utilidad en procesos de cribado virtual y priorización racional de compuestos. En comparación con SwissADME, ADMETlab 3.0 ofrece una evaluación más integral, particularmente por su mayor cobertura de toxicidad y herramientas complementarias para análisis y selección temprana de candidatos. Bajo esta perspectiva, la evaluación ADME/Tox no debe entenderse como un análisis accesorio, sino como una dimensión complementaria a QSAR, docking y dinámica molecular. Una molécula puede mostrar afinidad favorable o estabilidad conformacional adecuada frente a una proteína blanco y, sin embargo, ser poco atractiva como candidato si presenta absorción deficiente, distribución no deseada o múltiples alertas toxicológicas. Por ello, la incorporación de estos criterios permite una priorización más realista de las moléculas seleccionadas y orienta con mayor fundamento las etapas posteriores de optimización química, validación biológica y eventual desarrollo preclínico.

2.6 Hipótesis

Se plantea como hipótesis que: La integración de métodos QSAR, virtual screening, docking molecular, dinámica molecular y análisis ADME/Tox predictivo permite desarrollar nuevos derivados con potencial inhibitorio frente a PLpro del SARS-CoV-2.

3 Objetivos

3.1 Objetivo general

Encontrar nuevos inhibidores de la proteasa tipo papaína (PLpro) de SARS-CoV-2, preferentemente aquirales, mediante la aplicación de técnicas de modelamiento QSAR, Virtual screening, docking molecular y simulaciones de dinámica molecular.

3.2 Objetivos específicos

1. Construir y curar una base de datos robusta y actualizada de compuestos con actividad inhibitoria frente a PLpro de SARS-CoV-2, a partir de información reportada en la literatura científica.
2. Desarrollar y validar estadísticamente modelos QSAR 2D, evaluando descriptores moleculares de tipo 0D, 1D y 2D, con el fin de identificar relaciones cuantitativas entre la estructura química y la actividad inhibitoria frente a PLpro.
3. Diseñar nuevas moléculas inhibitorias de PLpro a partir de los modelos QSAR desarrollados y predecir su actividad biológica mediante las ecuaciones obtenidas.
4. Procesar y preparar la estructura cristalográfica de PLpro, así como grandes bibliotecas de compuestos orgánicos “drug-like” provenientes de bases de datos públicas, para su posterior evaluación computacional.

5. Realizar estudios de virtual screening de las bibliotecas moleculares seleccionadas mediante docking molecular rígido, con el fin de identificar compuestos con alta afinidad potencial hacia el sitio activo de PLpro.
6. Aplicar docking flexible y simulaciones de dinámica molecular para validar:
 - a) Las moléculas diseñadas a partir de los modelos QSAR.
 - b) Los compuestos mejor posicionados a partir del virtual screening.
7. Analizar las interacciones proteína–ligando y la estabilidad conformacional de los complejos PLpro–inhibidor obtenidos, mediante métricas estructurales derivadas de las simulaciones de dinámica molecular.
8. Realizar análisis retrosintético para evaluar la viabilidad sintética de las moléculas más prometedoras identificadas en el estudio computacional.
9. Evaluar propiedades de solubilidad, farmacocinética y toxicidad de las moléculas seleccionadas, utilizando algoritmos predictivos in silico, con el fin de priorizar candidatos con mejor perfil farmacológico.

4 Metodología

4.1 Generación de ecuaciones QSAR mediante selección manual

Se recopiló una serie de 142 moléculas con actividad biológica frente a PLpro a partir de la literatura^{6,18,89–98}. La actividad se reporta como IC₅₀ (Concentración de un compuesto necesario para inhibir el 50% de la actividad de proteína), medida en la enzima bajo las mismas condiciones experimentales. Los valores se convirtieron a pIC₅₀ ($pIC_{50} = -\log IC_{50}$).

Se calcularon una serie de descriptores estructurales y fisicoquímicos utilizando el programa PADEL-Descriptors⁵¹.

Se construyó una matriz con los valores de pIC₅₀ (variable dependiente) y los descriptores (variables independientes). Posteriormente, se realizó una regresión estadística PLS en Minitab, y se descartaron 29 compuestos con valores residuales elevados (residuos mayores a 0.5 unidades logarítmicas).

El modelo final se construyó con 113 moléculas, de las cuales 90 se utilizaron como conjunto de entrenamiento (79.6%) y 23 como conjunto de prueba (20.4%). La distribución de los valores de actividad biológica para los conjuntos de entrenamiento y prueba (*Training set* y *test set* respectivamente).

El modelo final se construyó utilizando los siguientes 13 descriptores: ALogP (lipofilicidad), ALogP², AMR (refractividad molar), DPSA-3 (dipolaridad/polarizabilidad, tamaño y forma, y conteo de tipos atómicos), nAr (número de anillos aromáticos), nBasic (número de átomos de nitrógeno básicos), nAcid (número de ácidos carboxílicos), nF (número de átomos de flúor), nCl (número de átomos de cloro), nHBDon (donadores de enlace de hidrógeno), nO (número de átomos de oxígeno), TopoPSA (área superficial polar topológica), VABC (volumen de van der Waals con contribuciones atómicas y de enlace)

Finalmente, se realizaron las validaciones estadísticas del modelo, calculando r^2_{test} , $r^2_{\text{rand-test}}$, q^2F1 , q^2F2 , q^2F3 , CCC, MAE y la prueba de *Y-randomization* (Prueba de actividad aleatoria).

4.2 Generación de la Ecuación QSAR mediante Algoritmo Genético

4.2.1 Base de Datos y Descriptores

Se utilizó una base de datos compuesta por 280 moléculas^{89–92,94,95,97–116} con valores experimentales de actividad biológica previamente normalizados. Para la caracterización estructural de estos compuestos, se calcularon 1875 descriptores moleculares utilizando el software *PaDEL*⁵¹. Estos descriptores incluyen propiedades fisicoquímicas, topológicas y electrónicas relevantes para la actividad biológica.

4.2.2 Configuración del Algoritmo Genético

Para la selección óptima de descriptores y la generación de la ecuación QSAR, se implementó un algoritmo genético (AG) utilizando un *Scribd* con la siguiente configuración:

- Tamaño de la población: 100 individuos por generación.
- Número total de generaciones: 1.000.000.
- Número de descriptores por individuo: 14.

- Función de aptitud (fitness): definida como el coeficiente de determinación (R^2) penalizado por un factor de colinealidad. Este criterio permitió favorecer modelos con alta capacidad explicativa y mínima redundancia entre variables.

4.2.3 Selección y Refinamiento de Modelos

4.2.3.1 Identificación de las Mejores Ecuaciones

Al finalizar la simulación del Algoritmo Genético (AG), se identificaron las 10 ecuaciones con los valores más altos de R^2 . Estas ecuaciones fueron consideradas candidatas para el refinamiento posterior.

4.2.3.2 Depuración de Datos y Eliminación de Valores Atípicos

Para asegurar la calidad del conjunto de datos, se realizó una depuración previa de los descriptores seleccionados. Se eliminaron aquellos que:

- No tenían sentido químico o eran difíciles de interpretar.
- Presentaban baja variabilidad (90% de los valores iguales).

Posteriormente, se llevó a cabo un análisis estadístico de los descriptores filtrados. Se evaluó la distribución de cada descriptor y de la variable dependiente mediante análisis exploratorio de datos. Aquellos valores que se desviaban significativamente del promedio fueron considerados atípicos y, por tanto, excluidos del análisis. Esta identificación se realizó mediante un análisis de distribución de los valores por de cada molécula en cada descriptor, estableciendo límites basados en el rango intercuartílico. Las observaciones con valores extremos fueron eliminadas para evitar sesgos en el modelo.

4.2.3.3 Control de Multicolinealidad (VIF)

Para minimizar la redundancia entre los descriptores seleccionados, se calculó el Factor de Inflación de la Varianza (VIF). Se eliminaron progresivamente aquellos descriptores con valores de VIF elevados (>5), asegurando la independencia estadística de las variables incluidas en el modelo.

4.2.3.4 Regresión Exploratoria y Eliminación de Valores Atípicos en residuos

Se realizó una regresión preliminar utilizando los descriptores filtrados y se evaluaron los residuos resultantes. Se eliminaron aquellas observaciones cuyos residuos se situaban más allá de dos desviaciones estándar respecto a la media, con el objetivo de mitigar la influencia de posibles valores aberrantes y mejorar la estabilidad del modelo.

4.2.4 Construcción y Validación del Modelo QSAR

Tras la depuración del conjunto de datos, se construyó el modelo QSAR definitivo empleando técnicas de regresión, específicamente *Partial Least Squares* (PLS). La calidad y robustez del modelo fueron evaluadas mediante el cálculo de:

- R^2 (training): Coeficiente de determinación en el conjunto de entrenamiento.
- Q^2 : Cuadrado del coeficiente de validación cruzada (LOO).
- F_1 , F_2 , F_3 : Coeficientes de la varianza explicada en la predicción externa.
- RMSE: (*Root Mean Squared Error*): Error cuadrático medio.
- MAE: (*Mean Absolute Error*): Error absoluto medio en ajuste.
- CCC: (*Concordance Correlation Coefficient*): Coeficiente de correlación de concordancia.
- $\bar{q}^2Y$ y $\bar{r}^2Y$: Promedios de Q^2 y R^2 obtenidos en la prueba de Y-randomización.
- R^2 (test): Coeficiente de regresión para el conjunto de prueba.

4.2.5 Selección de la Ecuación Final

Entre las ecuaciones que cumplieron con los criterios estadísticos de aceptabilidad, se seleccionó aquella que presentó:

- El mayor R^2 en el conjunto de entrenamiento.
- El mayor Q^2 y mejores coeficientes de varianza explicada (F_1 , F_2 , F_3).
- Los menores valores de error (RMSE y MAE).
- Un CCC elevado, garantizando la confiabilidad del modelo.

4.3 Virtual Screening

El *Virtual Screening* se llevó a cabo utilizando como estructura blanco la PLpro del SARS-CoV-2, en su conformación cristalográfica disponible en la base de datos Protein Data Bank (PDB) bajo el código 7NFV. Esta estructura fue seleccionada por su resolución y relevancia en la inhibición de la actividad enzimática viral. Para tener una comparación se usó el PDB de código 7CJM que contiene el cristal del ligando de referencia GRL-0617 y la proteína PLpro (Figura 13)

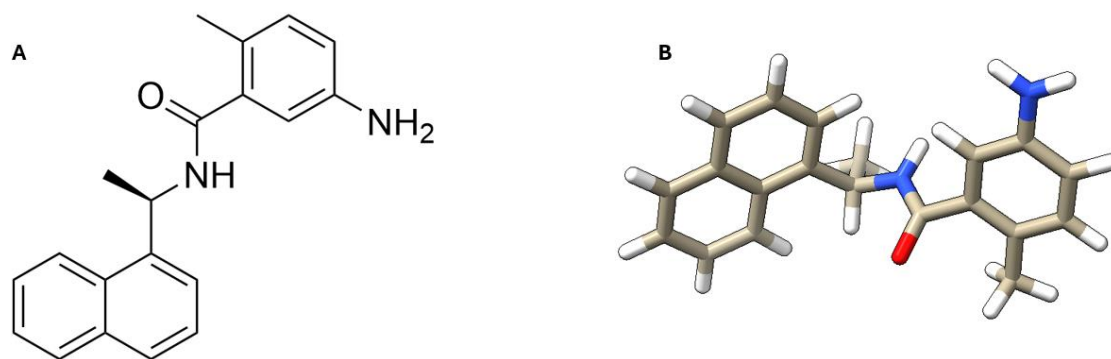


Figura 13 A) Figura en 2D de GRL-0617 B) Figura en 3D de GRL-0617.

El tratamiento inicial del archivo cristalográfico se realizó en el software UCSF Chimera versión 1.18, donde se procedió a visualizar y preparar la proteína. Posteriormente, se utilizó el software Schrodinger el módulo *protein preparation*, mediante el cual se eliminaron las moléculas de agua, ligandos co-cristalizados y otros componentes no deseados. Además, se añadieron los átomos de hidrógeno necesarios para la correcta parametrización de la estructura proteica. Finalmente, se conservó únicamente la cadena proteica correspondiente al blanco molecular.

El archivo tratado fue exportado en formato PDB e importado al programa AutoDock Tools versión 1.5.7p1. En esta plataforma, se definió una caja de búsqueda tridimensional (grid box) centrada en el átomo de carbono alfa del residuo Tirosina 268 (Y268), con un tamaño de 20 Å por lado, a fin de abarcar la cavidad activa del sitio catalítico.

Los mapas de potencial necesarios para el cálculo de las energías de acoplamiento fueron generados utilizando AutoGrid, considerando los tipos atómicos más frecuentes en los ligandos. Se incluyeron mapas para C (carbono sp³ de carácter hidrofóbico), A (carbono aromático), N (nitrógeno general), NA (nitrógeno aceptador de enlaces de hidrógeno), HD (hidrógeno donador de enlaces de hidrógeno), OA (oxígeno aceptador de enlaces de hidrógeno), SA (azufre, como el presente en residuos de cisteína), F (flúor), Cl (cloro), Br (bromo), Zn (zinc, relevante para interacciones de coordinación en sitios metálicos), así como mapas adicionales de interacción como e (electrostático general) y d (desolvatación, que representa el costo energético asociado a la eliminación de moléculas de agua durante la unión ligando-receptor).

Los ligandos fueron seleccionados a partir de la base de datos PubChem, buscando cada uno por las familias: Aminopirimidinas, Benzotiofenos, Benzodiazepinas, Benzamidas, Pirimidina aminobenzimidazol, Pirimido-Benzimidazol y Benzimidazol). Las moléculas son descargadas en formato .SDF y posteriormente convertidos a formato PDBQT mediante un script en línea de comandos utilizando el programa Open Babel (obabel), asegurando la orientación adecuada de cada molécula⁶².

Finalmente, el proceso de acoplamiento molecular se ejecutó utilizando AutoDock-GPU, versión acelerada mediante CUDA, con el fin de optimizar los tiempos de cálculo. Cada ligando fue evaluado individualmente, y los resultados de energía libre de unión fueron extraídos desde los archivos .xml generados al finalizar cada simulación.

4.4 Induced fit docking

El proceso se llevó a cabo utilizando el software Maestro en sus versiones 24.2 (Schrödinger LLC). La preparación de la proteína fue hecho en el módulo *protein preparation*, misma métodos empleados en el *virtual screening*.

Posteriormente, las moléculas candidatas fueron sometidas al protocolo de LigPrep, donde se llevó a cabo la generación de protonaciones relevantes, la estandarización de estructuras químicas y la minimización de energía bajo las condiciones predeterminadas de Force Field OPLS4. Con el conjunto de ligandos preparado, se inició el procedimiento de *Induced Fit Docking* empleando el módulo

correspondiente. Durante este proceso, se permitió la flexibilidad tanto del ligando como de residuos cercanos en el sitio activo para modelar los efectos de adaptación conformacional de la proteína.

De los resultados obtenidos, se seleccionó la pose más representativa de cada ligando de acuerdo con el IFD score generado por el programa, considerando únicamente la conformación con el valor más favorable como referencia para los análisis posteriores.

4.5 Dinámica molecular

Preparación del sistema

La proteína (PDB ID: 7NFV) preparada en la sección de *protein preparation* y los complejos proteína–ligando obtenidos en *induced fit docking* se colocaron en una caja ortorrómbica con un margen de 1.0 nm respecto a los bordes. El sistema se solvató con agua TIP3P y se añadieron iones Cl^- y Na^+ para neutralizar la carga total y ajustar la fuerza iónica a 0.15 M, reproduciendo condiciones fisiológicas. La caja fue de 100Åx64Åx60Å.

4.5.1 Protocolos de simulación

Antes de la producción, se aplicó el *Relaxation protocol* por defecto de Desmond, que incluye: Minimización de energía mediante descenso de gradiente, hasta reducir tensiones estéricas. Equilibración NVT a 310 K, con restricciones posicionales, para estabilizar la temperatura. Equilibración NPT a 310 K y 1 atm, manteniendo restricciones, para estabilizar la densidad y el volumen. Posteriormente, se llevaron a cabo las simulaciones de producción: Proteína sola: se realizaron tres réplicas independientes de 700 ns cada una bajo condiciones NPT (310 K y 1.01325 bar). Se empleó el termostato Nose–Hoover chain ($\tau = 1.0$ ps) y el barostato Martyna–Tobias–Klein con acoplamiento isotrópico ($\tau = 2.0$ ps). Complejos proteína–ligando: cada ligando seleccionado se simuló por 700 ns (una réplica) bajo las mismas condiciones NPT descritas arriba. Todas las simulaciones se ejecutaron con un paso de integración de 2 fs, registrando las trayectorias cada 700 ps, para un total de 1000 frames por corrida.

4.5.2 Análisis de trayectorias

La estabilidad estructural global se evaluó mediante el cálculo de la desviación cuadrática media (RMSD) de los átomos C α respecto a la conformación inicial. La flexibilidad local se determinó a través del cálculo del RMSF por residuo. Además, se analizaron distancias interatómicas relevantes con un intervalo de muestreo por cada frame, incluyendo:

- Cys111–His272(Figura 14)
- Arg166–Tyr268(Figura 15)
- Pro299–Tyr268()

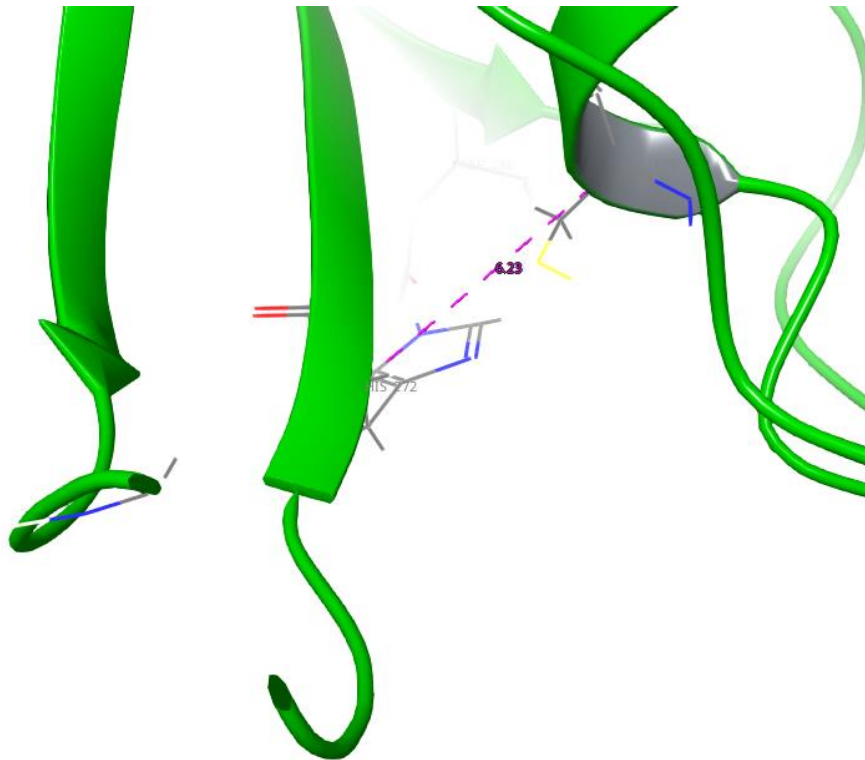


Figura 14 Distancia entre los carbonos alfa de los aminoácidos C111 y H272, aminoácidos catalíticos.

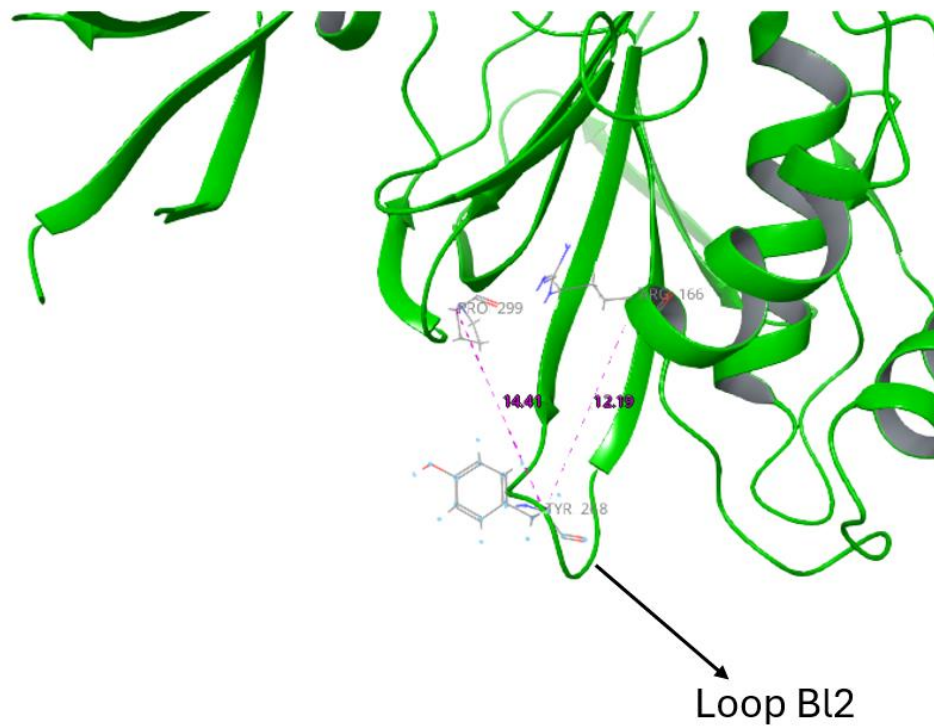


Figura 15 Distancias entre carbono alfa de :Tyr268-Arg166 y Tyr268-Pro299, aminoácidos que miden la movilidad del Loop BL2.

La dinámica conformacional de Tyr268 se evaluó mediante el seguimiento de dos ángulos diedros laterales (χ_1 y χ_2), definidos a partir de los átomos carbono alfa, carbono beta y el carbono gamma del anillo aromático (Figura 16 y Figura 17).

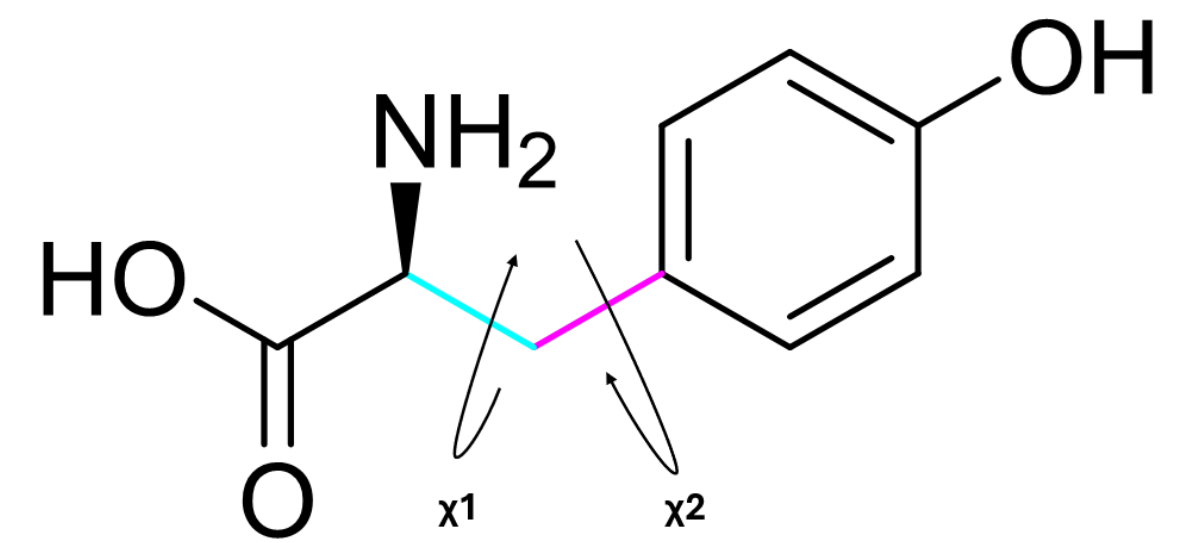


Figura 16 Representación bidimensional de ángulos diedros a evaluar.

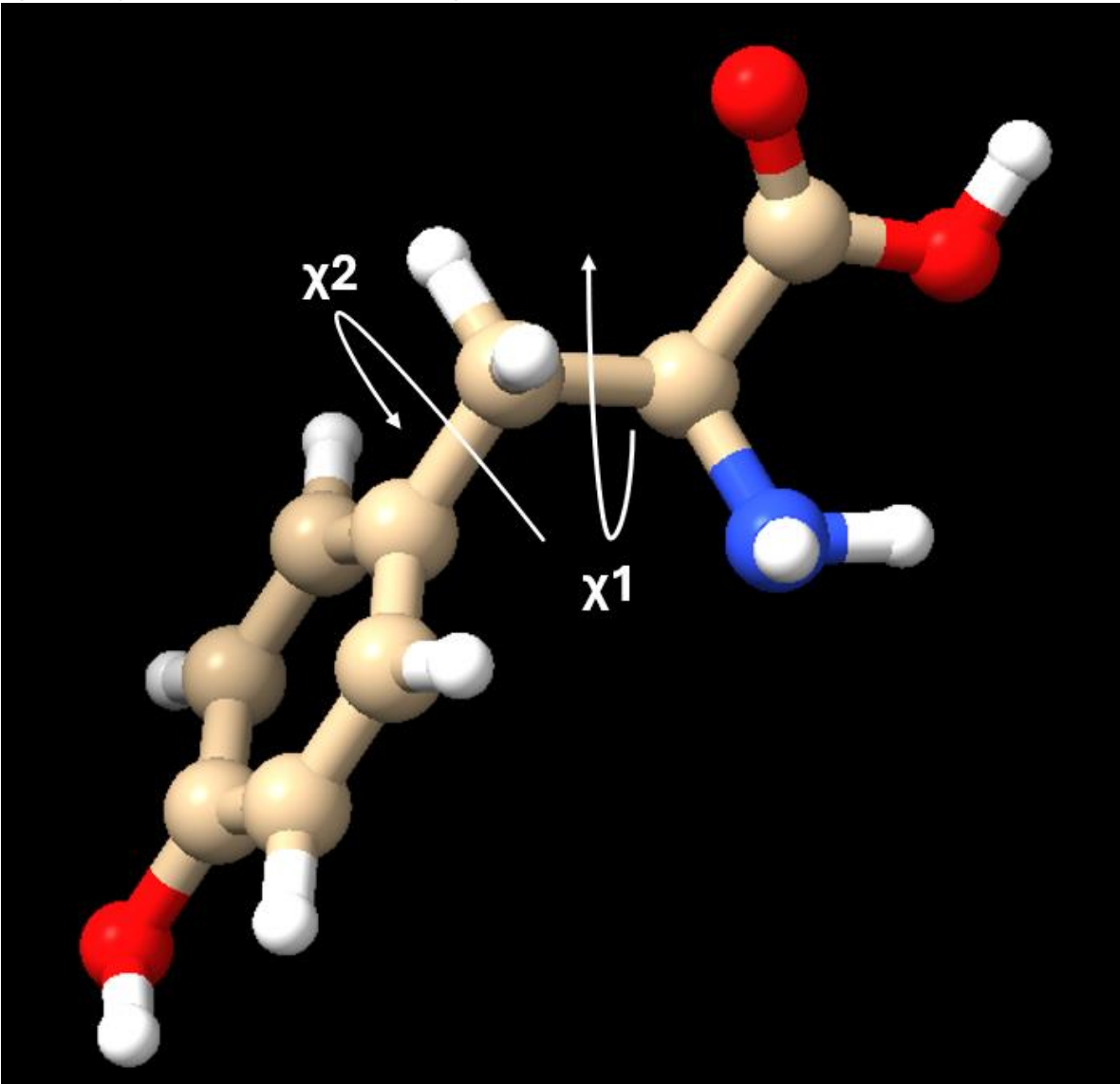


Figura 17 Representación tridimensional de ángulos diedros a evaluar

Los resultados reportados corresponden al promedio de las tres réplicas en el caso de la proteína sola. En caso de dinámicas con ligando se utiliza solo una única replica.

4.6 Análisis retrosintético y farmacocinético

Para el análisis retrosintético de los compuestos propuestos se empleó como referencia principal el texto *March's Advanced Organic Chemistry*, el cual proporciona fundamentos mecanísticos y estrategias para la identificación de fragmentos sintéticos y rutas viables de síntesis¹¹⁷. Este enfoque se complementó con

la aplicación de técnicas retrosintéticas clásicas manuales, orientadas a simplificar las estructuras objetivo en precursores accesibles y químicamente coherentes.

Paralelamente, se realizó una evaluación farmacocinética *in silico* mediante la plataforma ADMETLAB 3.0⁸⁶ con el fin de estimar propiedades fisicoquímicas, biodisponibilidad oral y parámetros relacionados con la absorción, distribución, permeabilidad molecular y solubilidad. Los valores dados en la plataforma son descargados en archivos CSV y graficados.

4.7 Síntesis orgánica

4.7.1 Protocolo general para la síntesis de chalconas en medio básico (KOH en etanol absoluto)

En un matraz de fondo redondo de volumen apropiado, equipado con barra de agitación magnética y operado a presión atmosférica, se introducen la cetona aromática (1,0 equivalente), el aldehído correspondiente (1,0 equivalente) y 10 mL etanol absoluto en cantidad suficiente para disolver completamente ambos reactivos.

La mezcla se agita a temperatura ambiente hasta lograr una disolución homogénea.

En paralelo, se prepara una solución acuosa de KOH al 40% p/p. Esta se adiciona lentamente al matraz en pequeñas alícuotas (50–100 μ L cada 2–5 min) hasta completar un total de 5 equivalentes respecto a la cetona, manteniendo la agitación continua.

Una vez completada la adición, la mezcla de reacción se mantiene bajo agitación a temperatura ambiente durante 20–22 h.

Finalizado el tiempo, el contenido se vierte sobre hielo triturado y se ajusta el pH con HCl 1 M gota a gota, hasta pH $\approx$ 7. El sólido formado se recolecta por filtración al vacío, se lava con etanol:agua frío (1:1) y se seca al aire o en estufa a baja temperatura.

Para purificación, se recrystaliza en caliente en etanol o (Acetato de etilo /hexano) (especificados posteriormente), si la precipitación es ineficiente, una extracción líquido-líquido con Acetato de etilo, concentrando el extracto en rotavapor

Nota:

1. En el caso particular del p- nitro benzaldehído y p- trifluorometano benzaldehído, debido a su mayor reactividad y tendencia a reacciones secundarias, la mezcla inicial se enfría en un baño de hielo (0 °C por 20h).

La solución de KOH 40% p/p se añade de la misma forma (alícuotas pequeñas y espaciadas), manteniendo la temperatura a 0 °C durante toda la adición y las primeras 20 h de reacción. Luego, se permite que la mezcla alcance temperatura ambiente y se mantiene en agitación durante 2 h adicionales.

4.7.2 Síntesis de derivados Pirimido[1,2-a]benzimidazol

En un tubo de reacción de 100 mL de fondo redondo equipado con barra de agitación magnética y condensador a reflujo, se introdujo la chalcona correspondiente (1.0 equivalente) junto con 2-aminopiridina (1,2 equivalentes). Ambos reactivos se disolvieron en 3 mL de etanol, cantidad mínima necesaria para asegurar la completa solubilización. Posteriormente, se adicionó 1.25 equivalente de KOH, y se calentó a reflujo por 8h.

Al completarse 8h, la solución se dejó enfriar a temperatura ambiente y se diluyó con agua fría. La fase acuosa se extrajo sucesivamente con diclorometano en dos ciclos, cada uno de 10 mL. Las fases orgánicas se reunieron, secaron sobre sulfato de sodio anhidro por 10 min y se concentraron a presión reducida mediante rotavapor. La fase orgánica es reducida hasta aproximadamente 5 mL, el concentrado es retirado y se sometió a purificación por cromatografía en columna, empleando una elución con diclorometano:metanol (95:5, v/v).

Nota: También la reacción fue hecha en dimetilformamida. En las mismas condiciones que en de etanol.

4.7.3 Síntesis de derivados de benzimidazol sustituidos con piridinas y naftalenos piridina benzimidazoles

4.7.3.1 Síntesis de benzimidazoles piridina en solvente acuoso o etanólico

La síntesis de los benzimidazoles sustituidos se llevó a cabo mediante la condensación de o-fenilendiamina con los aldehídos aromáticos correspondientes. En un tubo de reacción o balón de fondo redondo de 100 mL se combinaron o-fenilendiamina (1 equivalente) y el carbaldehído piridina (2,3 o 4) (2 equivalentes) en 50 mL de agua destilada o etanol(ambos métodos hechos por separado). La mezcla se

calentó a reflujo bajo agitación constante durante 24 horas, tiempo tras el cual se observó la formación de un sólido precipitado. Luego de enfriar a temperatura ambiente, el sólido se aisló por filtración, se lavó con agua fría y se secó al aire, posteriormente se cristaliza en una solución de etanol/agua 6:4. El producto crudo obtenido se recrystalizó en etanol para mejorar su pureza. En el caso del derivado obtenido a partir de 2-piridinacarbaldéhid, fue necesaria cromatografía en columna, empleando un gradiente de hexano a acetato de etilo en lugar de una cristalización.

4.7.3.2 Síntesis de benzimidazoles piridina en mortero

La síntesis mecanosintética de los benzimidazoles se realizó mediante molienda directa de los reactivos en medio ácido. En un mortero se combinaron el aldehído aromático correspondiente (1.1 equivalentes) y o-fenilendiamina (1 equivalente), añadiéndose posteriormente ácido acético glacial como medio de reacción (aprox. 0,5 mL por mmol de diamina). La mezcla se sometió a molienda manual durante aproximadamente 15 minutos.

Finalizada la reacción, el mortero se lavó con diclorometano y agua para transferir cuantitativamente el material. La fase acuosa se neutralizó cuidadosamente con bicarbonato de sodio hasta alcanzar pH 7–8. Posteriormente, el producto se extrajo con acetato de etilo y las fases orgánicas combinadas se secaron y concentraron. El sólido obtenido se recrystalizó a partir de una mezcla etanol/agua para su purificación. Por otra parte, el benzimidazol sintetizado con 2-piridinacarbaldéhid requirió purificación adicional mediante cromatografía en columna con hexano acetato de etilo en gradiente.

4.7.3.3 Síntesis de benzimidazoles piridina sustituidos NH con acetyl naftalenos

La obtención de los benzimidazoles N-acilados se realizó mediante la reacción de los benzimidazoles previamente sintetizados con cloruros de acilo derivados de acetyl naftalenos. En un tubo de reacción de 100mL seco se disolvió el benzimidazol correspondiente (1 equivalente) en THF anhidro, al cual se añadieron 1.2 eq de hidruro de sodio (NaH al 60% el aceite mineral) como base. Posteriormente, se añadió el cloruro de acilo (1- o 2-naftoil cloruro) (1.1 equivalentes) bajo agitación por 2h, permitiendo que la reacción transcurriera a temperatura ambiente hasta su completa conversión.

Finalizada la reacción, la mezcla se concentró y el producto crudo se purificó mediante cromatografía en columna utilizando una elución por gradiente de hexano y acetato de etilo, obteniéndose los correspondientes benzimidazoles N-acilados como sólidos purificados.

4.7.4 Caracterización Orgánica

Las moléculas tipo chalcona y los benzimidazoles finales fueron caracterizadas mediante espectroscopía de resonancia magnética nuclear de ^1H y ^{13}C empleando un equipo de 400 MHz. Por su parte, la molécula final pirimido-benzimidazol fue caracterizada adicionalmente mediante espectroscopía infrarroja en disolución de acetonitrilo y determinación de punto de fusión. La medición del punto de fusión se realizó por triplicado.

5 Resultados

5.1 QSAR de selección manual de descriptores

5.1.1 Pretratado de base de datos

El conjunto inicial estuvo compuesto por 142 moléculas con actividad inhibitoria reportada. A partir de este conjunto se generó una ecuación preliminar con el fin de evaluar la relación entre los descriptores moleculares y la actividad biológica observada.

Posteriormente, se analizaron los residuos individuales obtenidos a partir de dicha ecuación, definidos como la diferencia entre los valores experimentales y los valores predichos. Este análisis permitió identificar desviaciones respecto del comportamiento general del conjunto de datos.

Con el objetivo de obtener un modelo estadísticamente consistente y representativo, se aplicó un criterio de depuración basado en el valor absoluto del residuo, estableciéndose un umbral de $|\text{residuo}| > 0.6$ para la identificación de observaciones atípicas. La aplicación de este criterio indicó que 29 moléculas presentaban desviaciones superiores al límite establecido.

El análisis del conjunto de compuestos identificados como *outliers* (S001–S029) evidenció diferencias estructurales y fisicoquímicas respecto del dominio predominante del modelo QSAR, caracterizado por sistemas aromáticos amídicos relativamente rígidos y con basicidad moderada. Entre los compuestos excluidos se observaron estructuras con enlaces disulfuro y motivos tipo ditiocarbamato (S009, S012 y S013), así como una tioamida aromática (S010). También se identificaron sistemas polifenólicos

altamente oxigenados, como S014 y derivados fenólicos presentes entre S023 y S028. Asimismo, se registraron estructuras con elevada flexibilidad conformacional, incluyendo el compuesto peptidomimético S011 y cadenas alifáticas extensas en S018 y S019. Se observaron además diferencias electrónicas asociadas a heteroátomos y sustituyentes electronegativos, tales como heterociclos sulfurados (S003 y S020), un sustituyente trifluorometilo (S017) y sistemas heteroaromáticos nitrogenados (S004 y S029). Finalmente, algunos compuestos presentaron características asociadas a mayor lipofilia y volumen molecular, como los sistemas policíclicos en S015 y S016, así como el incremento del tamaño molecular observado en S018 y S019.

En consecuencia, estos compuestos fueron excluidos del conjunto de entrenamiento, obteniéndose un conjunto depurado de 113 moléculas, el cual fue empleado para la generación de la ecuación final del modelo QSAR.

El conjunto de datos depurado fue dividido en un conjunto de entrenamiento (training set) y un conjunto de prueba externo (test set). El training set, utilizado para la construcción del modelo, estuvo conformado por los compuestos:

1–14, 16–20, 22–27, 29–36, 38–43, 46, 48–52, 54–56, 58–62, 65–66, 68–69, 72–74, 76–77, 80–82, 84–92, 94–96, 98–101, 103–110 y 112.

Por su parte, el test set externo, empleado para evaluar la capacidad predictiva del modelo, estuvo compuesto por los compuestos:

15, 21, 28, 37, 44, 45, 47, 53, 57, 63, 64, 67, 70, 71, 75, 78, 79, 83, 93, 97, 102, 111 y 113.

Los valores de pIC_{50} experimental abarcan aproximadamente tres unidades logarítmicas, desde el compuesto 77 (pIC_{50} = 3.903), el menos activo, hasta el compuesto 43 (pIC_{50} = 6.947), el más activo.

La Tabla 1 presenta los residuos correspondientes a todas las moléculas evaluadas, junto con sus estructuras en formato SMILES, indicando los compuestos clasificados como outliers (identificados con el prefijo S0) y aquellos incluidos en el conjunto final.

Tabla 1 Valores de residuo del modelo QSAR de selección de descriptores manual y representación estructural (SMILES) de los compuestos evaluados.

ID	Residuo	Residuo	Clasificación	SMILES
1	0.522850895	0.522850895	Incluido	<chem>CC(c1cccc2ccccc12)NC(c1c(C)ccc(N)c1)=O</chem>
2	0.235785603	0.235785603	Incluido	<chem>CC(c1cccc2ccccc12)NC(c1c(C)ccc(NC(CC2)CCN2C(OC(C)(C)C)=O)c1)=O</chem>
3	-0.304126047	0.304126047	Incluido	<chem>CC(c1cccc2ccccc12)NC(c1c(C)ccc(NC(C(C2)CN2C(OC(C)(C)C)=O)=O)c1)=O</chem>
4	-0.380814179	0.380814179	Incluido	<chem>CC(c1cccc2ccccc12)NC(c1c(C)ccc(NC(C(C2)CCN2C(OC(C)(C)C)=O)=O)c1)=O</chem>
5	0.33877549	0.33877549	Incluido	<chem>CC(c1cccc2ccccc12)NC(c1c(C)ccc(NC2CNC2)c1)=O</chem>
6	0.462334116	0.462334116	Incluido	<chem>CC(c1cccc2ccccc12)NC(c1c(C)ccc(NC2CCNCC2)c1)=O</chem>
7	0.360707553	0.360707553	Incluido	<chem>CC(c1cccc2ccccc12)NC(c1c(C)ccc(NC(C2CNC2)=O)c1)=O</chem>
8	0.414317103	0.414317103	Incluido	<chem>CC(c1cccc2ccccc12)NC(c1c(C)ccc(NC(C2CCNCC2)=O)c1)=O</chem>
9	0.342613496	0.342613496	Incluido	<chem>CC(c1cccc2ccccc12)NC(c1c(C)ccc(NC2CCN(C)CC2)c1)=O</chem>
10	-0.046623442	0.046623442	Incluido	<chem>CC(c1cccc2ccccc12)NC(c1c(C)ccc(N(C)C2CN(C)C2)c1)=O</chem>
11	-0.366886215	0.366886215	Incluido	<chem>CC(c1cccc2ccccc12)NC(c1c(C)ccc(NC2CN(C)C2)c1)=O</chem>
12	-0.124794375	0.124794375	Incluido	<chem>CC(c1cccc2ccccc12)NC(c1c(C)ccc(N(C)C2CNC2)c1)=O</chem>
13	-0.013991529	0.013991529	Incluido	<chem>CC(c1cccc2ccccc12)NC(c1c(C)ccc(CNCC(O)=O)c1)=O</chem>
14	0.015553136	0.015553136	Incluido	<chem>CC(c1cccc2ccccc12)NC(c1c(C)ccc(CNC2CNC2)c1)=O</chem>
15	-0.060028029	0.060028029	Incluido	<chem>CC(c1cccc2ccccc12)NC(c1c(C)ccc(CN(C2)CC2C(O)=O)c1)=O</chem>
16	-0.405625917	0.405625917	Incluido	<chem>CC(c1cccc2ccccc12)NC(c1c(C)ccc(CN(CC2)CCC2C(O)=O)c1)=O</chem>
17	-0.026304898	0.026304898	Incluido	<chem>CC(c1cccc2ccccc12)NC(c(cc(cc1)N)c1Cl)=O</chem>
18	-0.003380091	0.003380091	Incluido	<chem>CC(c1cccc2ccccc12)NC(c(cc(cc1)N)c1Br)=O</chem>
19	-0.355763773	0.355763773	Incluido	<chem>CC(c1cccc2ccccc12)NC(c1c(C)ccc2c1ccn2CC1CNC1)=O</chem>
20	-0.279604205	0.279604205	Incluido	<chem>CC(c1cccc2ccccc12)NC(c(cc(cc1)N(C)C2CNC2)c1Cl)=O</chem>
21	-0.320231295	0.320231295	Incluido	<chem>CC(c1csc2c1cccc2)NC(c1c(C)ccc(N)c1)=O</chem>

22	-0.139754814	0.139754814	Incluido	CC(c(cc1)cc2c1scc2)NC(c1c(C)ccc(N(C)C2CNC2)c1)=O
23	-0.297399884	0.297399884	Incluido	CC(c1csc2c1cccc2)NC(c1c(C)ccc(N(C)C2CNC2)c1)=O
24	0.374159745	0.374159745	Incluido	CC(c1csc2c1cccc2)NC(c(cc(cc1)N(C)C2CNC2)c1Cl)=O
25	0.306408638	0.306408638	Incluido	CC(c1c(c(cccc2)c2[nH]2)c2ccc1)NC(c1c(C)ccc(N(C)C2CNC2)c1)=O
26	-0.040459235	0.040459235	Incluido	CC(c1csc2c1cccc2)NC(c1c(C)ccc(NC2CNC2)c1)=O
27	-0.030305617	0.030305617	Incluido	CC(c1c(c(cccc2)c2[nH]2)c2ccc1)NC(c1c(C)ccc(NC2CNC2)c1)=O
28	-0.272217122	0.272217122	Incluido	CC(c1nccc2ccccc12)NC(c1c(C)ccc(N(C)C2CNC2)c1)=O
29	-0.052510008	0.052510008	Incluido	CC(c1cc(-c2cccs2)ccc1)NC(c(cc(cc1)NC2CNC2)c1Cl)=O
30	-0.148992909	0.148992909	Incluido	CC(c1cccc(-c2cccs2)c1)NC(c1c(C)ccc(NC2CNC2)c1)=O
31	0.058437361	0.058437361	Incluido	CC(c1cccc(-c2csc2)c1)NC(c1c(C)ccc(NC2CNC2)c1)=O
32	0.188633517	0.188633517	Incluido	CC(c1cc(-c2c[nH]cc2)ccc1)NC(c1c(C)ccc(NC2CNC2)c1)=O
33	0.312122936	0.312122936	Incluido	CC(c1cccc(-c2ccc(CN3CCNCC3)s2)c1)NC(c1c(C)ccc(NC2CNC2)c1)=O
34	-0.021368565	0.021368565	Incluido	CC(c1cccc(-c2ccc(CN3CCOCC3)s2)c1)NC(c1c(C)ccc(NC2CNC2)c1)=O
35	0.084162108	0.084162108	Incluido	CC(c1cccc(-c2ccc(CNC3CCN(C)CC3)s2)c1)NC(c1c(C)ccc(NC2CNC2)c1)=O
36	-0.111684145	0.111684145	Incluido	CC(c1cccc(-c2ccc(CNCC3CCNCC3)s2)c1)NC(c1c(C)ccc(NC2CNC2)c1)=O
37	-0.148801579	0.148801579	Incluido	CC(c1cccc(-c2ccc(CNC3CCCC3)s2)c1)NC(c1c(C)ccc(NC2CNC2)c1)=O
38	0.197111553	0.197111553	Incluido	CC(c1cccc(-c2ccc(CN3CCCC3)s2)c1)NC(c1c(C)ccc(NC2CNC2)c1)=O
39	-0.204102219	0.204102219	Incluido	CC(c1cccc(-c2ccc(C)s2)c1)NC(c1c(C)ccc(NC2CNC2)c1)=O
40	-0.028556257	0.028556257	Incluido	CC(c1cccc(-c2ccc(C(O)=O)s2)c1)NC(c1c(C)ccc(NC2CNC2)c1)=O
41	0.283149219	0.283149219	Incluido	CC(c1cccc(-c2ccc(C(OC)=O)s2)c1)NC(c1c(C)ccc(NC2CNC2)c1)=O
42	-0.030361162	0.030361162	Incluido	CC(c1cccc(-c2ccc(C(NCC3OCCC3)=O)s2)c1)NC(c1c(C)ccc(NC2CNC2)c1)=O
43	0.558030032	0.558030032	Incluido	CC(c1cccc(-c2ccc(C(NCC3OCC3)=O)s2)c1)NC(c1c(C)ccc(NC2CNC2)c1)=O
44	0.203201535	0.203201535	Incluido	CC(c1cccc(-c2ccc(CN(CC3)CC3O)s2)c1)NC(c1c(C)ccc(NC2CNC2)c1)=O
45	0.133899923	0.133899923	Incluido	CC(c1cccc(-c2ccc(CN3CCC3)s2)c1)NC(c1c(C)ccc(NC2CNC2)c1)=O
46	-0.12851016	0.12851016	Incluido	CC(c1cccc(-c2ccc(CN(CC3)CC3O)s2)c1)NC(c1c(C)ccc(NC2CNC2)c1)=O
47	0.054852861	0.054852861	Incluido	CC(c1cccc(-c2ccc(CNC3COCC3)s2)c1)NC(c1c(C)ccc(NC2CNC2)c1)=O
48	0.004012449	0.004012449	Incluido	CC(c1cccc(-c2ccc(CNC(CCN3)C3=O)s2)c1)NC(c1c(C)ccc(NC2CNC2)c1)=O
49	-0.082083737	0.082083737	Incluido	CC(c1cccc(-c2ccc(CNC(C3)CNC3=O)s2)c1)NC(c1c(C)ccc(NC2CNC2)c1)=O
50	0.414876692	0.414876692	Incluido	CC(c1cccc(-c2ccc(CNC(CC3)CC3O)s2)c1)NC(c1c(C)ccc(NC2CNC2)c1)=O

51	-0.009596542	0.009596542	Incluido	CC(c1cccc(-c2ccc(CNC(CC3)CC3O)s2)c1)NC(c1c(C)ccc(NC2CN C2)c1)=O
52	0.388656322	0.388656322	Incluido	CC(c1cccc(-c2ccc(CNC(CC3)CC3O)s2)c1)NC(c1c(C)ccc(NC2CN C2)c1)=O
53	-0.147872291	0.147872291	Incluido	CC(c1cccc(-c2ccc(CNC(CC3)CC3O)s2)c1)NC(c1c(C)ccc(NC2CN C2)c1)=O
54	-0.447943309	0.447943309	Incluido	CC(c1cccc(-c2ccc(CNC(OC(C)(C)C)=O)s2)c1)NC(c1c(C)ccc(N)c1)=O
55	-0.130732783	0.130732783	Incluido	CC(c1cccc(-c2ccc(CN)s2)c1)NC(c1c(C)ccc(N)c1)=O
56	-0.274259622	0.274259622	Incluido	CC(c1cccc(-c2ccc(CNC(C3CCCC3)=O)s2)c1)NC(c1c(C)ccc(N)c1)=O
57	-0.24943269	0.24943269	Incluido	CC(c1cccc(-c2ccc(CNC(C)=O)s2)c1)NC(c1c(C)ccc(NC(C)=O)c1)=O
58	0.153552359	0.153552359	Incluido	CC(c1cccc(-c2ccc(CNC(CC3)CC3O)s2)c1)NC(c1c(C)ccc(NC(C)=O)c1)=O
59	-0.302416945	0.302416945	Incluido	CCN(CC)C(SSC(N(CC)CC)=S)=S
60	-0.095413524	0.095413524	Incluido	O=C1N(c2cccc2)[Se]c2c1cccc2
61	-0.118201954	0.118201954	Incluido	CC(c1cccc2cccc12)NC(c1c(C)cccc1)=O
62	-0.12615484	0.12615484	Incluido	CC(c1cccc2cccc12)NC(c(ccc1)c1Cl)=O
63	-0.306610225	0.306610225	Incluido	CC(c1cccc2cccc12)NC(c1c(C)ccs1)=O
64	0.411950319	0.411950319	Incluido	CC(c1cccc2cccc12)NC(c(cc1)c(C)cc1N)=O
65	0.045512112	0.045512112	Incluido	CC(c1cccc2cccc12)NC(c1c(C)ccc(N(C)C)c1)=O
66	0.071361969	0.071361969	Incluido	CC(c1cccc2cccc12)NC(c(cc1)c(C)cc1NS(C)(=O)=O)=O
67	0.028785289	0.028785289	Incluido	CC(c1cccc2cccc12)N(C1)Cc2c1cccc2
68	-0.063548183	0.063548183	Incluido	CC(c1cccc2cccc12)N(C1)Cc2c1ccc(N)c2
69	-0.192038876	0.192038876	Incluido	C=CCC(C(NC(Cc1c[nH]c2c1cccc2)C(NCc1cccc1)=O)=O)NC(CCl)=O
70	0.382810399	0.382810399	Incluido	Oc(cc1)ccc1C(Oc1c2c(O)cc(O)c1-c(cc(cc1)C(Oc3c4c(O)cc(O)c3)=CC4=O)c1O)=CC2=O
71	0.032060332	0.032060332	Incluido	COc(cc1)ccc1C(Oc1c2c(O)cc(O)c1-c(cc(cc1)C(Oc3c4c(O)cc(O)c3)=CC4=O)c1O)=CC2=O
72	0.051850485	0.051850485	Incluido	COc(cc1O)cc(OC(c(cc2)cc(-c(c(OC(c(cc3)ccc3O)=C3)c(c(O)c4)C3=O)c4O)c2OC)=C2)c1C2=O
73	-0.136561085	0.136561085	Incluido	COc(cc1)ccc1C(Oc1c2c(O)cc(O)c1-c(cc(cc1)C(Oc3c4c(O)cc(O)c3)=CC4=O)c1OC)=CC2=O
74	-0.031191553	0.031191553	Incluido	COc(cc1)ccc1C(Oc1c2c(O)cc(O)c1-c(cc(cc1)C(Oc3c4c(O)cc(OC)c3)=CC4=O)c1OC)=CC2=O
75	0.463114137	0.463114137	Incluido	Oc1ccc(C(C(c(OC(c(cc2)cc(O)c2O)=C2)c(c(O)c3)C2=O)c3O)C2=O)Oc3c2c(O)cc(O)c3cc1
76	0.381913433	0.381913433	Incluido	Oc(cc1)ccc1C(Oc(c1c2O)cc(O)c2Oc(cc2)ccc2C(Oc2 c3c(O)cc(O)c2)=CC3=O)=CC1=O
77	-0.275246253	0.275246253	Incluido	COc(cc1)ccc1C(Oc(c1c2O)cc(OC)c2Oc(cc2)ccc2C(O c2c3c(O)cc(O)c2)=CC3=O)=CC1=O
78	-0.147984281	0.147984281	Incluido	COc(ccc(C(Oc1c2c(O)cc(O)c1)=CC2=O)c1)c1Oc(cc1)ccc1C(Oc1c2c(O)cc(O)c1)=CC2=O
79	0.553518609	0.553518609	Incluido	Oc(cc1)ccc1C(Oc1c2c(O)cc(O)c1)=CC2=O
80	-0.035201717	0.035201717	Incluido	COc(cc1)ccc1C(Oc1c2c(O)cc(O)c1)=CC2=O
81	-0.173213294	0.173213294	Incluido	S=C(N1CCOCC1)SSC(N1CCOCC1)=S
82	-0.136638331	0.136638331	Incluido	CCOC(SSC(OCC)=S)=S
83	-0.040887451	0.040887451	Incluido	S=C(N1CCCC1)SSC(N1CCCC1)=S
84	0.176551068	0.176551068	Incluido	CN(C)C(SSC(N(C)C)=S)=S
85	0.23494624	0.23494624	Incluido	CN(CC1)CCN1C(SSC(N1CCN(C)CC1)=S)=S
86	-0.157345448	0.157345448	Incluido	COc(nc12)ccc2nc(cc(cc2)Cl)c2c1Nc(cc1CN2CCCC2)cc(CN2CCCC2)c1O
87	-0.09477449	0.09477449	Incluido	CN(C)Cc1cc(Nc(c2c3)c(ccc(Cl)c4)c4nc2ccc3OC)cc(CN(C)C)c1O
88	0.003516946	0.003516946	Incluido	COc(cc12)ccc2nc(cc(cc2)Cl)c2c1Nc(cc1CN2CCCC2)cc(CN2CCCC2)c1O
89	0.016652633	0.016652633	Incluido	COc(cc12)ccc2nc(cc(cc2)Cl)c2c1Nc(cc1)cc(CN2CCC C2)c1O
90	-0.273794171	0.273794171	Incluido	Oc1cc(OC(C(C2)OC(c(cc3O)cc(O)c3O)=O)c(cc3)cc(O)c3O)c2c(O)c1
91	-0.427626823	0.427626823	Incluido	CC(c1cc2cccc2cc1)NC(c1c(C)ccc(N)c1)=O

92	-0.302970214	0.302970214	Incluido	OCC(C(C(C1O)O)O)OC1Oc1cc(O)cc(/C=C/c(cc2)ccc2O)c1
93	-0.45672251	0.45672251	Incluido	Oc1ccc(/C=C/c2cc(O)cc(O)c2)cc1
94	-0.522024729	0.522024729	Incluido	Cc1[n+](Cc2ncnc2)c(C(c(cccc2)c2C2=O)=O)c2n1CCOC.[Br-]
95	0.417234524	0.417234524	Incluido	CC1C(C(C(c2c3ccc4c2CCCC4(C)C)=O)=O)=C3OC1
96	0.011557812	0.011557812	Incluido	Cc1coc(-c(ccc2c(C)cccc22)c2C2=O)c1C2=O
97	0.098287604	0.098287604	Incluido	CC(c1cccc2cccc12)NC(c1c(C)ccc(NC(NC(N)=O)=O)c1)=O
98	-0.239363001	0.239363001	Incluido	CC(c1cccc2cccc12)NC(c1c(C)ccc(NC(C=C)=O)c1)=O
99	0.365170304	0.365170304	Incluido	Nc1cc(C(Nc2cccc3cccc23)=O)c(C(F)(F)F)cc1
100	0.361796	0.361796	Incluido	CCCCNC(Nc1cc(C(NC(C)c2cccc3cccc23)=O)c(C)cc1)=O
101	0.040586388	0.040586388	Incluido	CC(c1cccc2cccc12)NC(c1c(C)ccc(NC(Oc(cc2)ccc2[N+](O-)=O)=O)c1)=O
102	-0.281343968	0.281343968	Incluido	CCCCC(Nc1cc(C(NC(C)c2cccc3cccc23)=O)c(C)cc1)=O
103	-0.107600269	0.107600269	Incluido	CC(C)c1c(CNCc2cc(C(O)=O)c(C)o2)occ1
104	-0.017898576	0.017898576	Incluido	Cc1c(C(N(CCC(O)=O)C2CCOCC2)=O)sc2c1c(F)ccc2
105	-0.097049551	0.097049551	Incluido	OC(CCN(C1CCOCC1)C(c(cc1)ccc1OC1CCCC1)=O)=O
106	0.414473462	0.414473462	Incluido	OC(CCN(C1CCOCC1)C(CCc1ncc(-c(cccc2)c2F)o1)=O)=O
107	-0.053509821	0.053509821	Incluido	Cc1nsc(NC)c1C(Nc(cc1)cc(CC2)c1N(Cc(cc1)ccc1Cl)C2=O)=O
108	-0.150281849	0.150281849	Incluido	OC(CN(Cc1cccc1)C(c1ccc(CN(Cc2cccc2)S(c(cc2)ccc2Cl)=O)=O)cc1)=O)=O
109	0.359868246	0.359868246	Incluido	OC(C(Cc1cccc1)NC(c1ccc(CN(Cc2cccc2)S(c(cc2)ccc2Cl)=O)=O)cc1)=O)=O
110	-0.127756965	0.127756965	Incluido	OC(C(c1cccc1)NC(c1ccc(CN(Cc2cccc2)S(c(cc2)ccc2Cl)=O)=O)cc1)=O)=O
111	0.425761096	0.425761096	Incluido	CC(C)(Cc(cc1)ccc1-c(cc1)cc(C(F)(F)F)c1Cl)C(O)=O
112	-0.580994506	0.58099451	Incluido	CC(c1cccc2cccc12)N(C)Cc(cc1)ccc1N(C)CCO
113	-0.078391752	0.07839175	Incluido	O=C1C2=C(N=C(C3=CN=CC=C3)O2)C(C4=CC=CC=C41)=O
S001	-0.633802264	0.633802264	Excluido (outlier)	CC(c1cccc2cccc12)NC(c(cc(cc1)N)c1Cl)=O
S002	-1.878769905	1.878769905	Excluido (outlier)	CC(c1cccc2cccc12)NC(c1c(C)ccc2c1ccn2CC1CNC1)=O
S003	-1.536680793	1.536680793	Excluido (outlier)	CC(c(cc1)cc2c1sc2)NC(c1c(C)ccc(N(C)C2CNC2)c1)=O
S004	-2.233673325	2.233673325	Excluido (outlier)	CC(c1nccc2cccc12)NC(c1c(C)ccc(N(C)C2CNC2)c1)=O
S005	-1.720168563	1.720168563	Excluido (outlier)	CC(c1cccc(-c2ccc(CN3CCNCC3)s2)c1)NC(c1c(C)ccc(NC2CNC2)c1)=O
S006	-1.813433671	1.813433671	Excluido (outlier)	CC(c1cccc(-c2ccc(CNC3CCN(C)CC3)s2)c1)NC(c1c(C)ccc(NC2CNC2)c1)=O
S007	-1.998146285	1.998146285	Excluido (outlier)	CC(c1cccc(-c2ccc(CNCC3CCNCC3)s2)c1)NC(c1c(C)ccc(NC2CNC2)c1)=O
S008	-1.600697129	1.600697129	Excluido (outlier)	CC(c1cccc(-c2ccc(CN3CCC3)s2)c1)NC(c1c(C)ccc(NC2CNC2)c1)=O
S009	-1.297006411	1.297006411	Excluido (outlier)	CCN(CC)C(SSC(N(CC)CC)=S)=S
S010	-2.012116991	2.012116991	Excluido (outlier)	CC(c1cccc2cccc12)NC(c1c(C)ccs1)=O
S011	0.66250712	0.66250712	Excluido (outlier)	C=CCC(C(NC(Cc1c[nH]c2c1cccc2)C(NC1cccc1)=O)=O)NC(CCl)=O
S012	1.120564795	1.120564795	Excluido (outlier)	S=C(N1CCOCC1)SSC(N1CCOCC1)=S
S013	-1.00036684	1.00036684	Excluido (outlier)	CN(CC1)CCN1C(SSC(N1CCN(C)CC1)=S)=S
S014	-0.658862599	0.658862599	Excluido (outlier)	Oc1cc(OC(C(C2)OC(c(cc3O)cc(O)c3O)=O)c(cc3)cc(O)c3O)c2c(O)c1
S015	-0.613457712	0.613457712	Excluido (outlier)	CC1C(C(C(c2c3ccc4c2CCCC4(C)C)=O)=O)=C3OC1
S016	-0.607811364	0.607811364	Excluido (outlier)	Cc1coc(-c(ccc2c(C)cccc22)c2C2=O)c1C2=O
S017	-1.305176242	1.305176242	Excluido (outlier)	Nc1cc(C(Nc2cccc3cccc23)=O)c(C(F)(F)F)cc1

S018	-1.719980133	1.719980133	Excluido (outlier)	CCCCNC(Nc1cc(C(NC(C)c2cccc3cccc23)=O)c(C)cc1)=O
S019	-1.458257658	1.458257658	Excluido (outlier)	CCCCC(Nc1cc(C(NC(C)c2cccc3cccc23)=O)c(C)cc1)=O
S020	-1.13570937	1.13570937	Excluido (outlier)	Cc1c(C(N(CCC(O)=O)C2CCOCC2)=O)sc2c1c(F)ccc2
S021	-2.100109499	2.100109499	Excluido (outlier)	CCn1c(ccc(NC(c2c(NC)snc2C)=O)c2)c2c2c1cccc2
S022	0.677604665	0.677604665	Excluido (outlier)	CC(c(cccc1)c1Cl)N(C)Cc(cc(cc1)O)c1Cl
S023	-0.90781259	0.90781259	Excluido (outlier)	CC(c1cccc2cccc12)N(C)Cc1c(C)ccc(O)c1
S024	-1.302688659	1.302688659	Excluido (outlier)	CC(c1cccc2cccc12)N(C)Cc1cccc(O)c1
S025	-0.70625787	0.70625787	Excluido (outlier)	CC(c1cccc2cccc12)N(C)Cc(cc1)ccc1O
S026	-1.132746977	1.132746977	Excluido (outlier)	CC(c1cccc2cccc12)N(C)Cc(cc1)c(C)cc1O
S027	-0.626653565	0.626653565	Excluido (outlier)	CC(c1cccc2cccc12)N(C)Cc(ccc(O)c1)c1Cl
S028	-0.612655503	0.612655503	Excluido (outlier)	CC(c1cccc2cccc12)N(C)Cc(ccc(O)c1)c1OC
S029	-1.526354691	1.526354691	Excluido (outlier)	CC(c1cccc2cccc12)N(C)Cc1cnc(N)nc1

5.1.2 Validación QSAR

La Figura 18 muestra la relación entre las actividades experimentales y predichas para los conjuntos de entrenamiento y prueba, así como la distribución de los residuos asociados, permitiendo visualizar el desempeño global del modelo.

Con el propósito de evaluar cuantitativamente su desempeño, se calcularon parámetros de validación interna y externa. La Tabla 2 presenta las métricas obtenidas para las ecuaciones evaluadas, incluyendo coeficientes de determinación, parámetros de validación cruzada, métricas de predicción externa y estimadores de error.

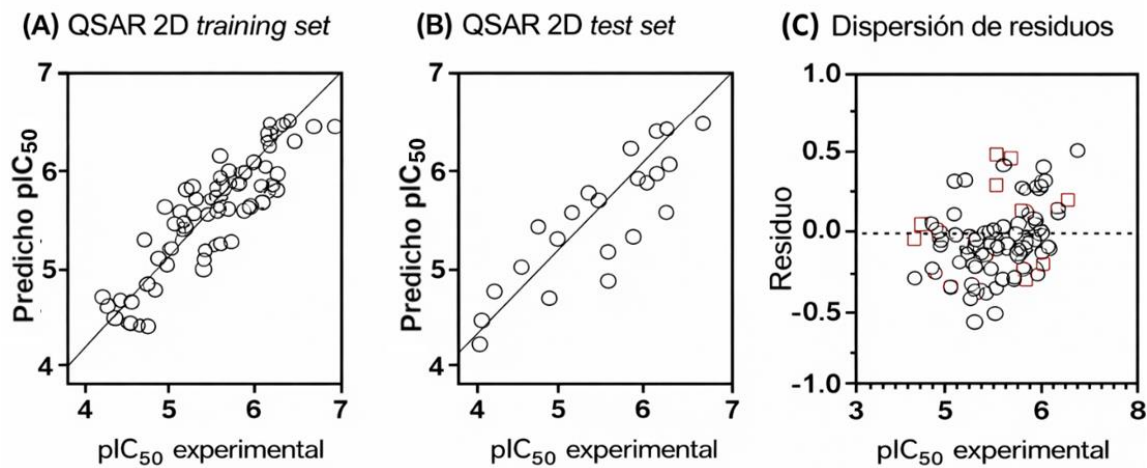


Figura 18 Gráficos de la actividad experimental frente a la actividad predicha para el conjunto de entrenamiento (A) y el conjunto de prueba (B). Gráfico de distribución de residuos (C). Los círculos negros representan el training set y los cuadrados rojos el test set. Todos los compuestos se encuentran dentro de una unidad logarítmica de residuo ($\text{residuo} = \text{pIC}_{50} \text{ exp} - \text{pIC}_{50} \text{ pred}$)¹¹⁸.

El modelo QSAR fueron evaluados mediante parámetros de validación interna, externa y aleatorización. El número óptimo de componentes PLS utilizado para cada modelo también se incluye en la Tabla 2.

Tabla 2 Parámetros estadísticos utilizados para la validación interna y externa del modelo QSAR.

Condición	Parametro	Valor tope	Resultado
1	r ²	Lo máximo	0.8333
2	q ²	>0.5	0.770
3	q ² F1	>0.7	0.817
4	q ² F2	>0.7	0.817
5	q ² F3	>0.7	0.984
6	MAE	<0.5	0.237
7	CCC	>0.85	0.905

8	q ² Y	<0.5	0.0003
9	r ² Y	<0.6	0.118
10	r ² _{test}	>0.5	0.721
11	r ² _{m training}	>0.5	0.83
12	rm test	>0.5	0.7
13	RMSEP	<0.5	0.35
14	PLS componentes	-	5

R²: coeficiente de determinación del ajuste; Q²: validación cruzada leave-one-out; Q²F1, Q²F2 y Q²F3: coeficientes de validación externa; r² test: coeficiente de determinación del conjunto externo; MAE: error absoluto medio; CCC: coeficiente de correlación concordante; q²Y y r²Y: valores promedio obtenidos por Y-randomization; rm²(training) y rm²(test): coeficientes rm² para entrenamiento y prueba; RMSEP: error cuadrático medio de predicción; PLS: número óptimo de componentes del modelo.

5.1.3 Ecuación final del modelo QSAR de selección manual de descriptores

A partir del conjunto depurado de 113 moléculas y tras la aplicación del método de regresión por mínimos cuadrados parciales (PLS), se obtuvo la ecuación final del modelo QSAR, la cual describe la relación cuantitativa entre los descriptores moleculares seleccionados y la actividad inhibitoria expresada como pIC₅₀.

A continuación, se presenta la ecuación obtenida a partir del análisis PLS utilizando cinco componentes:

$$pIC_{50} = 3.84264 - 0.23411 \, nF - 0.79971 \, nCl - 0.21961 \, nO + 0.16803 \, ALogP + 0.04589 \, ALogP^2 + 0.00796 \, AMR + 0.00119 \, DPSA-3 + 0.2659 \, nAr + 0.17442 \, nBasic + 0.22252 \, nAcid + 0.07801 \, nHBDOn + 0.00225 \, TopoPSA + 0.00145 \, VABC$$

donde:

- **ALogP** corresponde a la lipofilicidad molecular;
- **ALogP²** representa el cuadrado de la lipofilicidad;
- **AMR** es la refractividad molar;
- **DPSA-3** corresponde a la diferencia entre la superficie parcial cargada positiva (PPSA-3) y la superficie parcial cargada negativa (PNSA-3), descriptores que cuantifican la distribución de carga superficial molecular;
- **nAr** indica el número de anillos aromáticos;
- **nBasic** corresponde al número de átomos de nitrógeno básicos;
- **nAcid** indica el número de grupos ácido carboxílico;
- **nF** corresponde al número de átomos de flúor;
- **nCl** corresponde al número de átomos de cloro;
- **nHBDOn** indica el número de átomos donadores de enlaces de hidrógeno;
- **nO** corresponde al número de átomos de oxígeno;
- **TopoPSA** es el área superficial polar topológica;
- **VABC** corresponde al volumen de van der Waals calculado mediante contribuciones atómicas y de enlace.

5.2 QSAR de selección vía algoritmo genético

5.2.1 Análisis por algoritmo Generación y Selección del Modelo QSAR

Una vez finalizada la simulación mediante el algoritmo genético (AG), se registraron las 10 ecuaciones con los valores más altos de R², las cuales fueron evaluadas como candidatas a modelos predictivos finales (Figura 19). La simulación analizó un total de 1533 descriptores, seleccionados de los 1875 originalmente calculados mediante PaDEL. La presencia o ausencia de estos descriptores a lo largo de las generaciones se representa en la Figura 20, donde el color azul indica presencia y el rojo, ausencia. Ambas figuras sugieren que la simulación convergió adecuadamente dentro del límite de 1.000.000 de generaciones. Los códigos smiles de las moléculas originalmente analizadas se muestran en Anexo I.

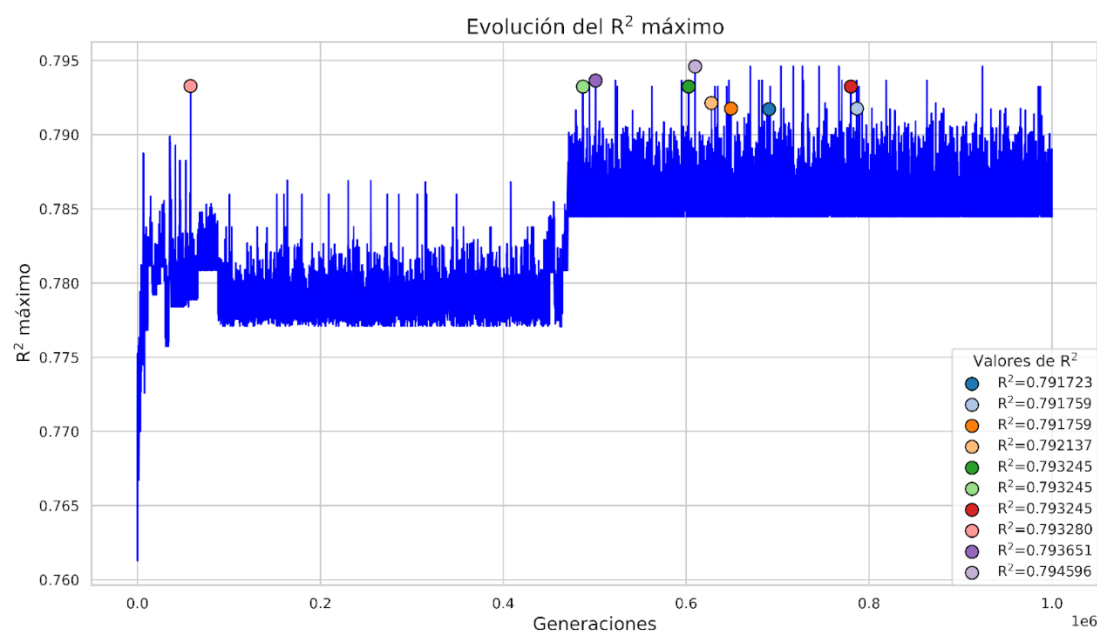


Figura 19 Evolución del coeficiente de determinación máximo (R^2) a lo largo de las generaciones del algoritmo genético. Se destacan los diez modelos con mayor R^2 obtenidos durante la simulación, considerados candidatos a ecuaciones predictivas finales. La estabilización del R^2 en las últimas generaciones indica la convergencia del algoritmo dentro del límite de 1.000.000 de generaciones.

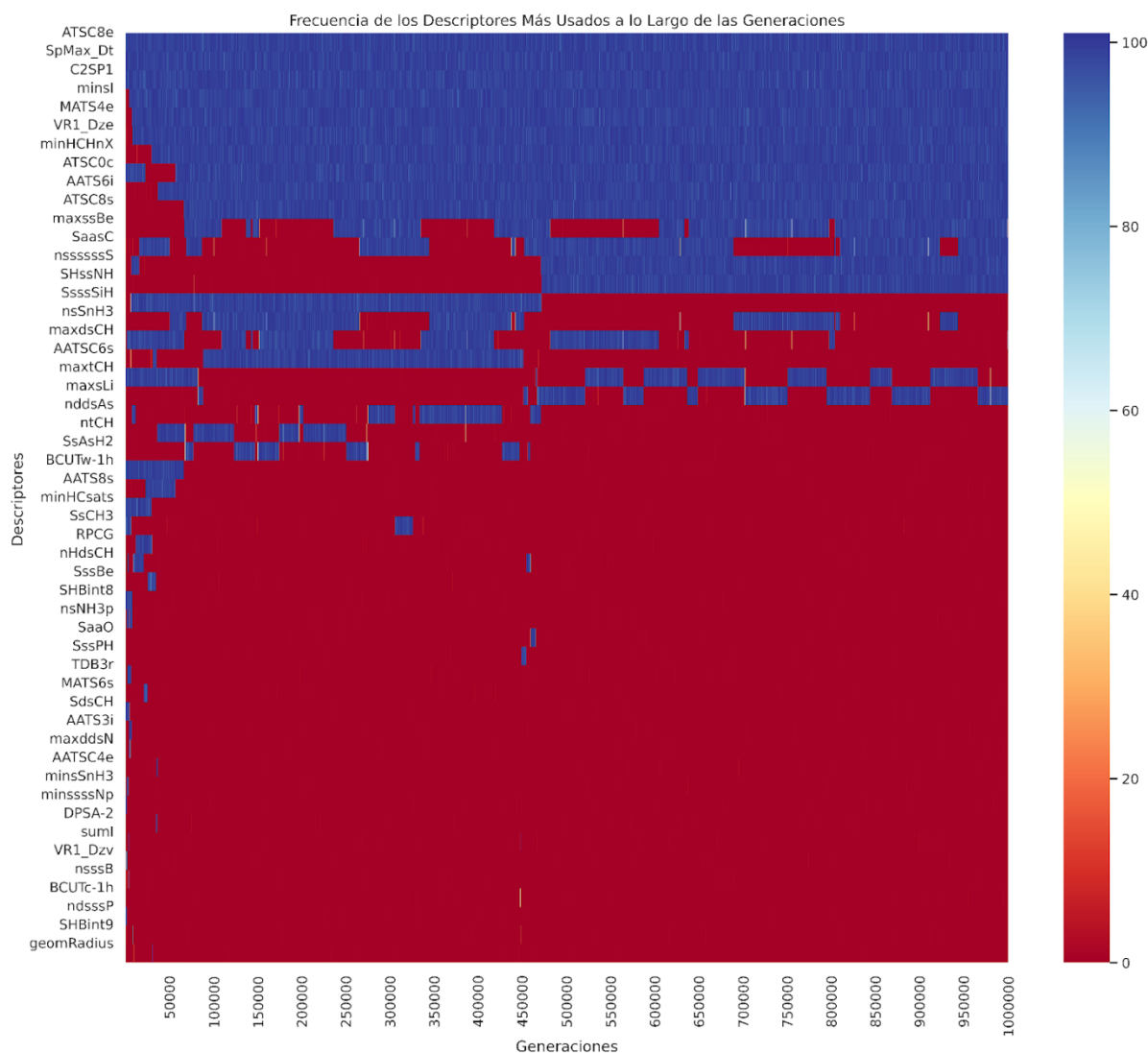


Figura 20 Mapa de calor que muestra la frecuencia de selección de los descriptores moleculares durante la ejecución del algoritmo genético. Se evaluaron 1533 descriptores derivados de los 1875 calculados inicialmente con PaDEL. El color azul indica la presencia del descriptor en los individuos seleccionados y el rojo su ausencia a lo largo de las generaciones. El patrón observado evidencia la estabilización del proceso evolutivo y la convergencia del modelo.

5.2.2 Análisis y Depuración de Ecuaciones

Durante el análisis de las 10 ecuaciones, se llevó a cabo un proceso de depuración estadística. Se eliminaron cinco valores atípicos identificados en los descriptores, utilizando diagramas de caja y criterios basados en el rango intercuartílico.

Adicionalmente, se eliminaron ciertos descriptores por motivos de carencia de fundamento químico o por baja variabilidad. En particular, se excluyeron los siguientes:

- **VR2_Dze, SpMAD_Dt, BCUTc-1l y SpMax3_Bhm:** derivados de valores propios de matrices. Si bien son útiles en contextos de inteligencia artificial, no tienen validez directa dentro del marco QSAR.
- **SddSss:** descartado por baja variabilidad, ya que el 99.7% de sus valores permanecen constantes.

Asimismo, se identificaron y eliminaron cuatro moléculas (Figura 21) cuyos perfiles descriptoriales escapaban claramente de la distribución general del conjunto. Estas moléculas fueron detectadas mediante gráficos de distribución individuales por descriptor (Figura 22), donde sus valores aparecían como atípicos extremos:

- **Molécula 106 y Molécula 110:** Presentan valores extremos en descriptores como ATSC0p y varios descriptores RDF. Ambas corresponden a estructuras de tipo **péptidos cíclicos**, lo que las aleja estructural y funcionalmente del dominio químico de las moléculas analizadas.
- **Molécula 156:** Se destacó en AATSC0c, ATSC1c y BCUTc-1l. Esta molécula posee un **grupo ácido borónico**, poco frecuente en el conjunto principal.
- **Molécula 177:** Mostró valores extremos en AATS7i y MATS5e. Se trata de una molécula **muy pequeña y clorada**, con una arquitectura química claramente distinta al núcleo aril-amida predominante.

Estas estructuras difieren sustancialmente del dominio químico general de la base de datos, el cual está compuesto principalmente por derivados de **aril-amidas, pseudo-péptidos y compuestos heterocíclicos funcionalizados**. Por tanto, su exclusión fue justificada con el objetivo de preservar la homogeneidad estructural del conjunto y mejorar la robustez estadística del modelo.

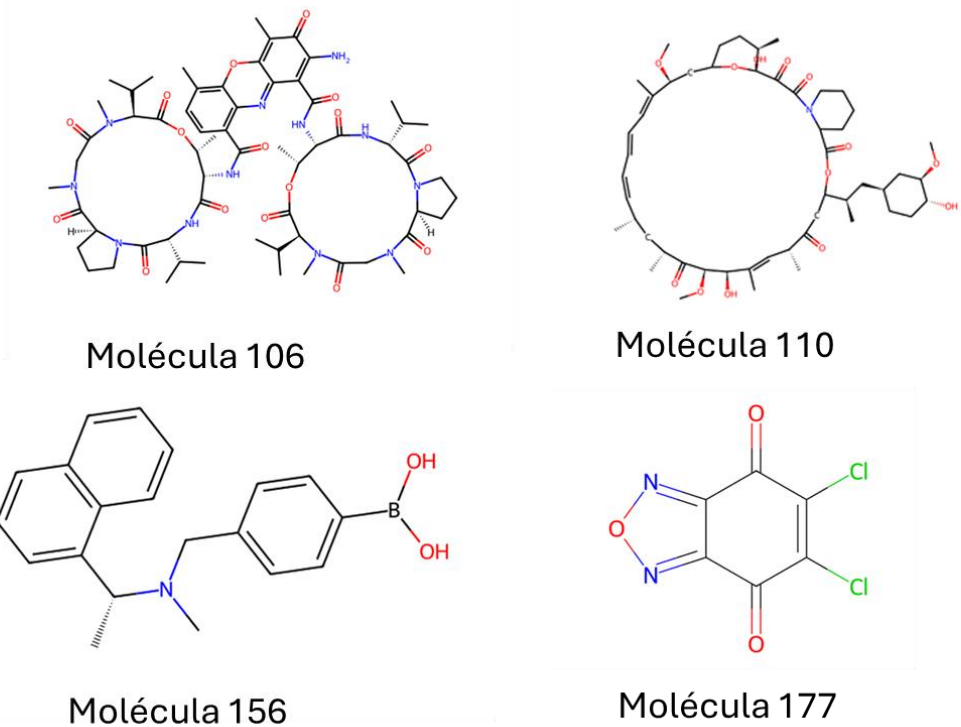


Figura 21 Moléculas con estructuras atípicas

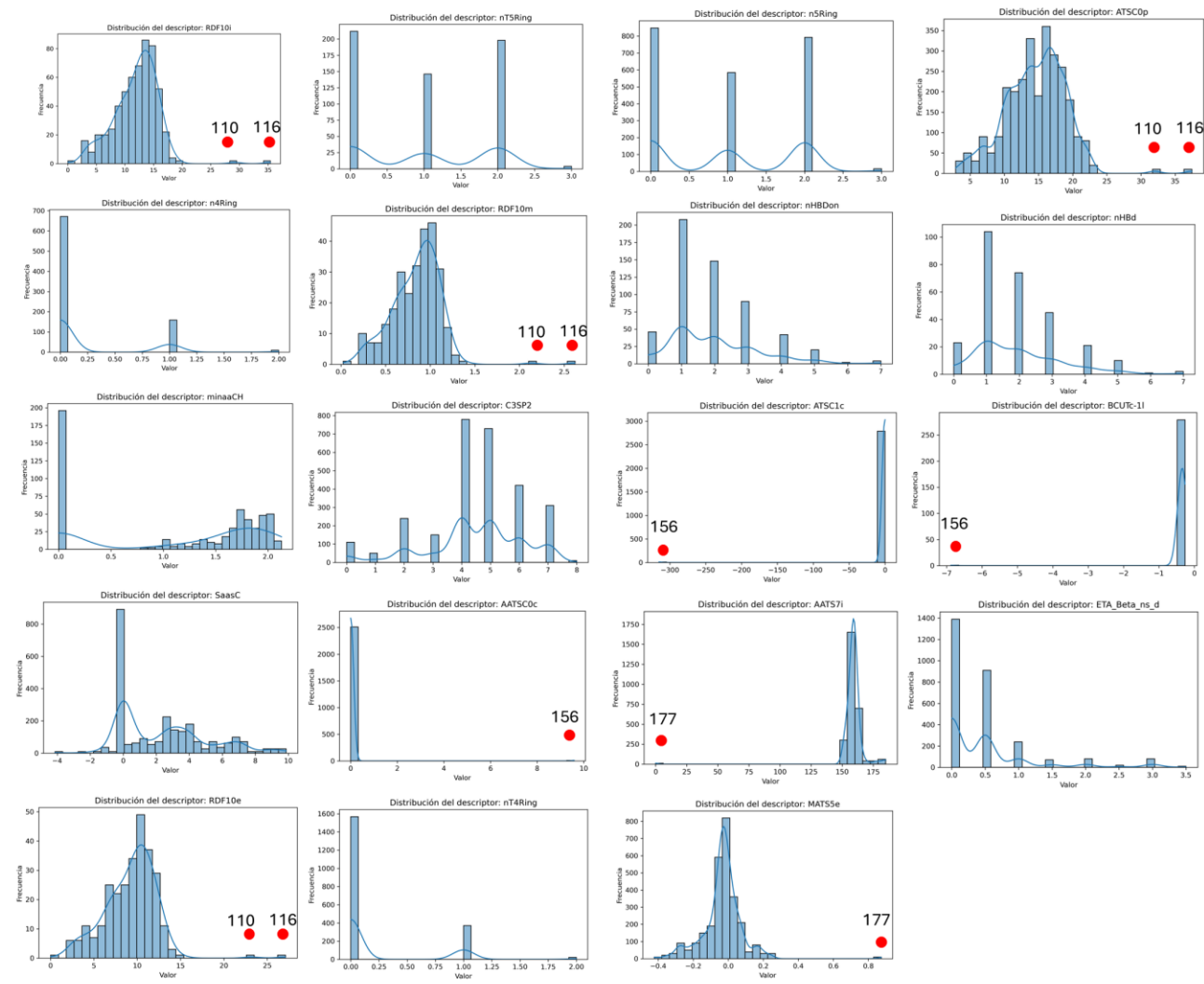


Figura 22 Distribuciones de valores de descriptores por cada molécula. Cada punto en rojo y su número representa el número de la molécula.

Posteriormente, se evaluó la multicolinealidad mediante el cálculo del **Factor de Inflación de la Varianza (VIF)**. Se eliminaron los descriptores que presentaban un VIF mayor a 5, ya que indicaban colinealidad estadísticamente significativa entre variables, lo cual podría comprometer la estabilidad del modelo (Tabla 3).

En particular, se eliminaron los siguientes descriptores:

- **minaaCH**: Eliminado en las ecuaciones 1 y 2 por presentar un VIF de 8.61.
- **ATSC1c**: Eliminado en las ecuaciones 4, 5 y 6 por un VIF de 6.65.
- **RDF10e, RDF10i, RDF10m**: Eliminados en las ecuaciones 7, 8, 10 respectivamente por presentar valores de VIF superiores a 6.

Tabla 3. Descriptores eliminados por colinealidad (VIF > 5)

Ecuación	Descriptor	VIF
1	minaaCH	8.61
2	minaaCH	8.61
4	ATSC1c	6.65
5	ATSC1c	6.65
6	ATSC1c	6.65

7	RDF10e	6.48
8	RDF10i	6.38
10	RDF10m	6.12

Se observó que las ecuaciones 1, 2, 7, 8, 9 y 10, así como 4, 5 y 6, presentaban estructuras idénticas. En consecuencia, se decidió eliminar las ecuaciones redundantes, específicamente las ecuaciones 2, 5, 6, 7, 8, 9 y 10.

Posteriormente, se realizó una regresión final con las tres ecuaciones restantes (1, 3 y 4), y se evaluó cómo se distribuían sus residuos. Se analizaron los residuos a 1σ, 1.5σ y 2σ, observándose cuántas moléculas serían eliminadas bajo distintos criterios de desviación estándar (Figura 23). Con base en estos resultados, se optó por aplicar el criterio más estricto (1 desviación estándar σ=0.434), ya que incluso con esta restricción el número de moléculas retenidas seguía siendo estadísticamente significativo: 193 moléculas para la Ecuación 1, y 194 para las Ecuaciones 3 y 4. Las ecuaciones se reportan a continuación como las finales.

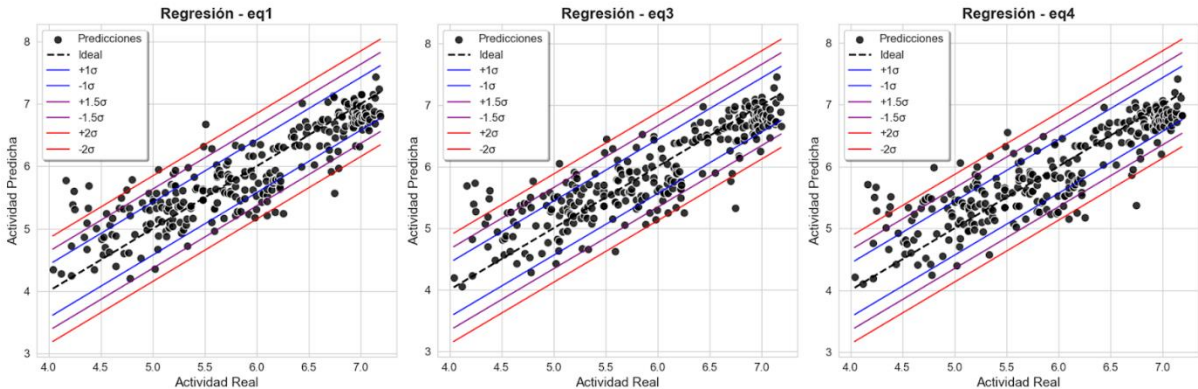


Figura 23 Gráficos de dispersión entre la actividad biológica experimental (pIC₅₀ real) y la actividad predicha por los modelos QSAR seleccionados (ecuaciones 1, 3 y 4). La línea negra representa la relación ideal (y = x), mientras que las líneas azules y rojas delimitan los intervalos de error ±1σ, ±1.5σ y ±2σ. La distribución de los puntos alrededor de la línea ideal evidencia la capacidad predictiva de los modelos y la ausencia de desviaciones sistemáticas significativas.

5.2.3 Ecuaciones QSAR Seleccionadas

Tras la evaluación por residuos y la selección basada en criterios de desviación estándar, se determinaron tres ecuaciones finales candidatas.

Ecuación 1:

$$pIC_{50} = 0.1029 + 0.0223 \text{ AATS7i} + 0.3997 \text{ ATSC1c} + 0.0823 \text{ ATSC0p} + 1.0415 \text{ MATS5e} + 0.1358 \text{ C3SP2} - 0.3847 \text{ ETA_Beta_ns_d} + 30.7332 \text{ AATSC0c} + 0.0258 \text{ SaasC} + 0.3410 \text{ n5Ring} - 0.1633 \text{ n4Ring}$$

Ecuación 3:

$$pIC_{50} = 1.7796 + 0.0149 \text{ AATS7i} + 0.4199 \text{ ATSC1c} + 0.0773 \text{ ATSC0p} + 0.7698 \text{ MATS5e} + 0.1304 \text{ C3SP2} - 0.2596 \text{ ETA_Beta_ns_d} + 0.3517 \text{ n5Ring} - 0.2245 \text{ n4Ring}$$

Ecuación 4:

$$pIC_{50} = 0.4600 + 0.0210 \text{ AATS7i} + 0.0820 \text{ ATSC0p} + 1.1181 \text{ MATS5e} + 0.1220 \text{ C3SP2} - 0.3597 \text{ ETA_Beta_ns_d} + 19.4572 \text{ AATSC0c} - 0.0303 \text{ nHBd} + 0.0316 \text{ SaasC} + 0.3549 \text{ n5Ring} - 0.2237 \text{ n4Ring}$$

Se decidió eliminar la Ecuación 3, ya que contiene un menor número de descriptores, lo que se traduce en una menor capacidad de interpretación estructural y detallada en comparación con las otras dos.

5.2.4 Validación estadística

A partir de esta decisión, se realizó un análisis estadístico comparativo entre las Ecuaciones 1 y 4. Este análisis contempló métricas de validación como R² (entrenamiento), Q² (validación cruzada leave-one-out), RMSE, MAE, CCC, R² (test), y los coeficientes F₁, F₂, F₃, así como los valores promedio q²Y y r²Y de Y-randomization. Los resultados se presentan en la Tabla 4.

Tabla 4 Validación estadística comparativa entre las Ecuaciones 1 y 4.

Métrica	Ecuación 1	Ecuación 4
r^2	0.9257	0.9351
r^2_{test}	0.9366	0.9163
$r^2_{\text{m training}}$	0.9257	0.9351
$r^2_{\text{m test}}$	0.9037	0.9137
q^2	0.9089	0.9254
q^2F1	0.9381	0.9175
q^2F2	0.9366	0.9163
q^2F3	0.9796	0.978
MAE	0.1829	0.1749
CCC	0.9676	0.9726
RMSEP	0.2274	0.2423
q^2Y	0.0633	0.0686
r^2Y	0.0054	0.0070
PLS componentes	5	5

A partir de los valores estadísticos obtenidos, se observa que ambas ecuaciones presentan un desempeño estadístico sólido. La Ecuación 1 muestra un mayor R^2_{test} (0.9366) y menor RMSEP (0.2274), lo cual indica mejor ajuste externo en términos de varianza. Sin embargo, la Ecuación 4 fue seleccionada como modelo QSAR final por cumplir de manera más equilibrada con los criterios establecidos:

- Posee un Q^2 más alto (0.9254), reflejando mejor capacidad predictiva interna.
- Tiene un MAE más bajo (0.1749), lo que implica menor desviación promedio.
- Su CCC también es más alto (0.9726), confirmando mayor concordancia entre valores experimentales y predichos.
- Mantiene consistencia en los valores de q^2F1 , q^2F2 y q^2F3 .

En consecuencia, estos parámetros respaldan la selección de la Ecuación 4 como el modelo más robusto y confiable para representar la relación estructura-actividad de las moléculas evaluadas.

Adicionalmente, la Figura 24 muestra la distribución de las moléculas correspondientes a los conjuntos de entrenamiento y prueba utilizados para validar la Ecuación 4.

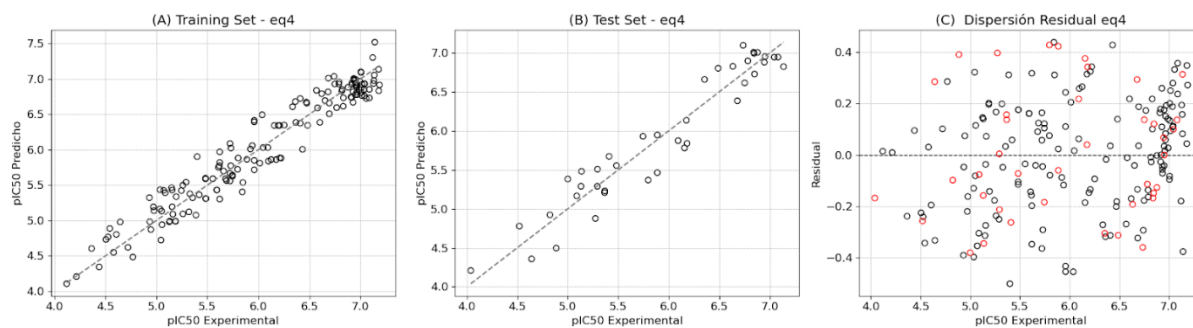


Figura 24 Evaluación gráfica del desempeño predictivo de la Ecuación 4. (A) Relación entre los valores experimentales y predichos de pIC_{50} para el conjunto de entrenamiento, mostrando un ajuste lineal adecuado. (B) Relación entre valores experimentales y predichos para el conjunto de prueba externo, evidenciando una buena capacidad predictiva del modelo. (C) Gráfico de dispersión de los residuos en función del pIC_{50} experimental, donde se observa una distribución aleatoria alrededor de cero sin tendencias sistemáticas, lo que confirma la ausencia de sesgos y la robustez estadística del modelo.

5.2.5 Significado de los descriptores seleccionados

A continuación, se presenta una descripción detallada del significado químico y matemático de cada descriptor incluido en la Ecuación 4, con el fin de facilitar su interpretación estructural dentro del modelo QSAR:

1. AATS7i: Autocorrelación de Broto-Moreau (lag 7), ponderada por potencial de ionización.
2. ATSC0p: Autocorrelación centrada de Broto-Moreau (lag 0), ponderada por polarizabilidad.
3. MATS5e: Autocorrelación de Moran (lag 5), ponderada por electronegatividad de Sanderson.
4. C3SP2: Número de carbonos sp^2 en estructuras cíclicas.
5. ETA_Beta_ns_d: Descriptor electrotopológico que mide la contribución de electrones no compartidos. Por cada átomo con capacidad directa de participar en resonancia con sistemas aromáticos, se suma 0.5 al descriptor.
6. AATSC0c: Autocorrelación centrada de Broto-Moreau (lag 0), ponderada por electronegatividad del carbono.
7. SaasC: Suma de electronegatividades de Sanderson para los átomos de carbono.
8. n5Ring: Número de anillos de 5 miembros.
9. n4Ring: Número de anillos de 4 miembros.
10. nHBd: Número de donadores de enlaces de hidrógeno.

5.2.6 Predicción y diseño de nuevos compuestos

5.2.6.1 Compuestos derivados de chalconas y heterociclos

Con el objetivo de explorar el potencial biológico de nuevas estructuras derivadas de chalconas, se aplicaron dos modelos QSAR previamente desarrollados en este trabajo: (i) un modelo basado en selección evolutiva de descriptores moleculares y (ii) un modelo basado en descriptores fisicoquímicos seleccionados manualmente. Ambos modelos se utilizaron para estimar los valores de pIC_{50} de las estructuras diseñadas y analizar la distribución de actividad predicha dentro del conjunto molecular evaluado.

La biblioteca molecular se construyó a partir de chalconas precursoras, las cuales fueron utilizadas como plataforma sintética para generar distintas familias heterocíclicas. Las familias estructurales consideradas en este estudio se muestran en la Figura 25, donde se resumen los esqueletos químicos evaluados.

Las estructuras fueron generadas mediante una estrategia combinatoria basada en la variación de dos fragmentos estructurales principales: R^1 , derivado de la cetona aromática empleada en la condensación de Claisen-Schmidt, y R^2 , derivado del aldehído aromático correspondiente.

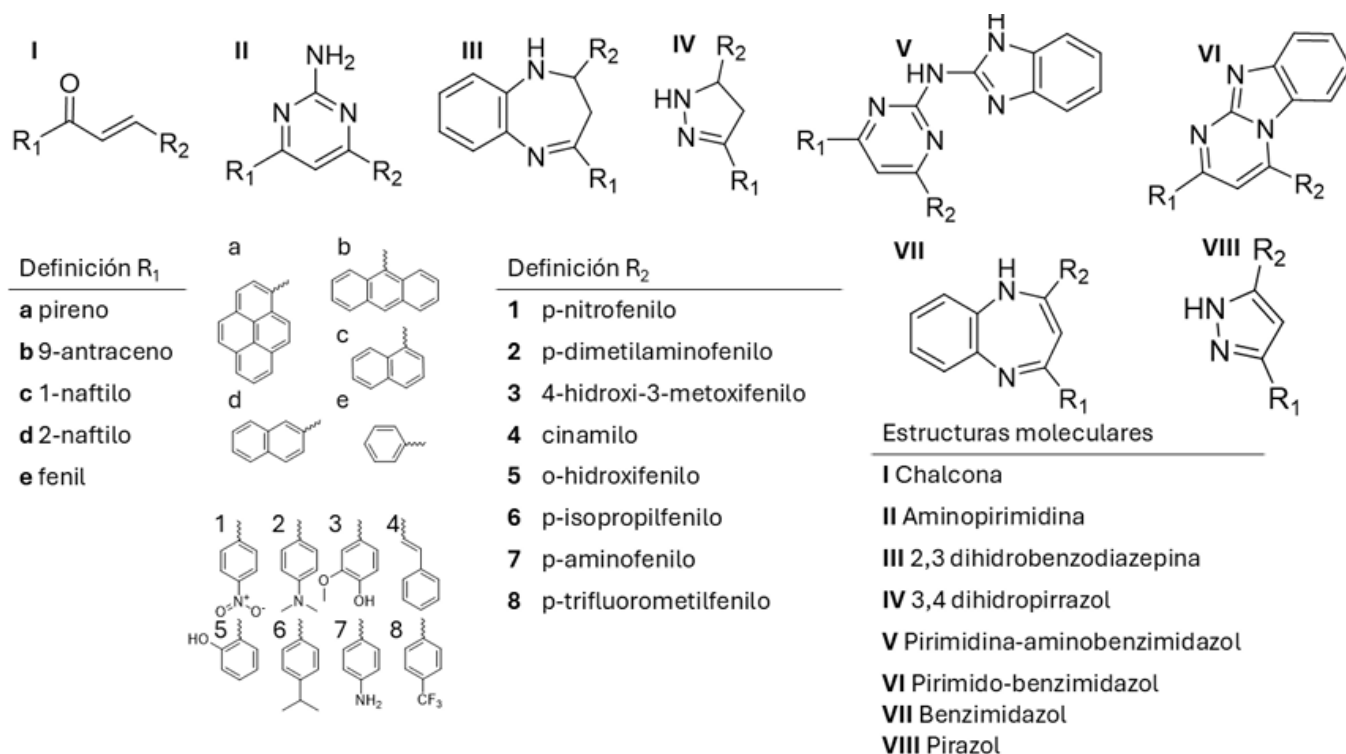


Figura 25 Familias estructurales evaluadas en el estudio. Las estructuras se generaron a partir de chalconas precursoras mediante transformaciones sintéticas que permiten acceder a diferentes sistemas heterocíclicos.

La generación de la biblioteca molecular se realizó combinando cinco fragmentos derivados de cetonas (R^1) con ocho fragmentos derivados de aldehídos (R^2). Esta combinación permitió obtener inicialmente 40 chalconas precursoras, las cuales posteriormente se transformaron conceptualmente en diferentes familias heterocíclicas. El esquema general de generación de la biblioteca molecular se muestra en la Figura 26.

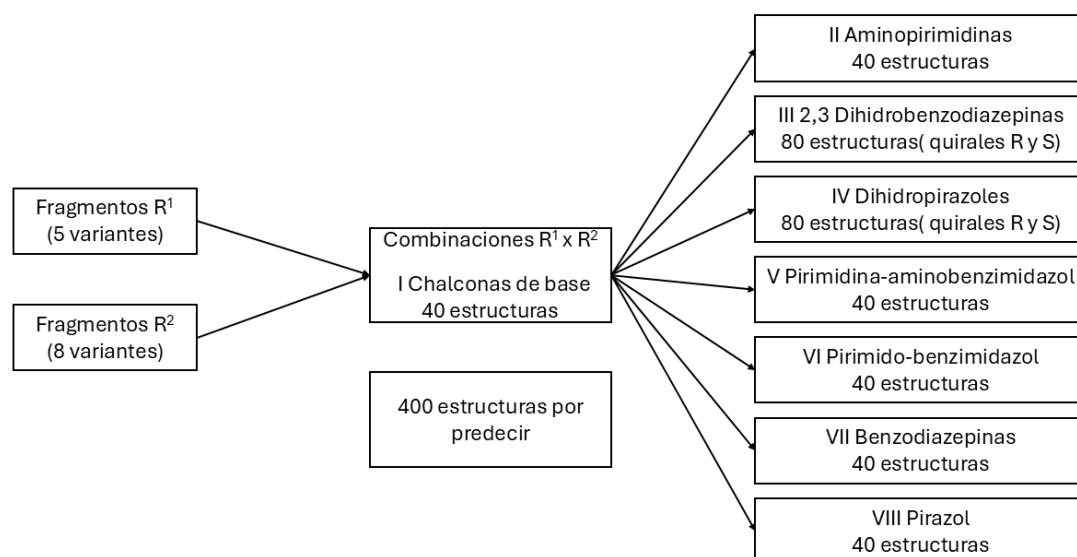


Figura 26 Estrategia de generación de la biblioteca molecular utilizada para las predicciones QSAR. La combinación de cinco fragmentos derivados de cetonas (R^1) y ocho fragmentos derivados de aldehídos (R^2) permitió generar 40 chalconas precursoras, las cuales

A partir de estas combinaciones se generó un total de 400 estructuras, las cuales fueron evaluadas mediante ambos modelos QSAR para estimar sus valores de pIC_{50} predicho.

5.2.6.1.1 Distribución de actividad predicha por familia estructural

Los valores de pIC_{50} obtenidos mediante ambos modelos QSAR mostraron diferencias sistemáticas entre las distintas familias estructurales evaluadas, cada valor en particular se muestra en Anexo II.

En términos generales, las chalconas precursoras presentaron los valores de actividad predicha más bajos dentro del conjunto analizado. En contraste, varias de las familias heterocíclicas derivadas mostraron incrementos en los valores de pIC_{50} predichos.

En el modelo QSAR evolutivo, las familias heterocíclicas derivadas de mayor complejidad estructural presentaron los valores más elevados dentro del rango de predicción del modelo. En particular, las familias pirimidina-aminobenzimidazol y pirimido-benzimidazol concentraron varios de los valores más altos del conjunto. Asimismo, las familias 3,4-dihidropirazol y 2,3-dihidrobenzodiazepina mostraron valores intermedios a altos dentro de este modelo.

En el modelo QSAR manual, los valores absolutos de pIC_{50} predicho se distribuyeron en un rango más alto en comparación con el modelo evolutivo. En este caso, las familias benzodiazepina aromática, pirimidina-aminobenzimidazol y pirimido-benzimidazol presentaron los valores más elevados dentro del conjunto evaluado.

5.2.6.1.2 Influencia del fragmento derivado de la cetona (R^1)

El análisis de los fragmentos derivados de cetona utilizados para generar las chalconas precursoras mostró que las combinaciones que incorporan sistemas aromáticos extendidos en el fragmento R^1 presentaron valores de pIC_{50} predicho más altos dentro de varias familias estructurales.

En particular, las estructuras que incorporan fragmentos derivados de 1-acetilpireno y 9-acetilantraceno aparecieron de manera recurrente entre las moléculas con mayores valores de actividad predicha en ambos modelos QSAR. Por el contrario, los fragmentos derivados de sistemas aromáticos de menor extensión conjugada presentaron valores de actividad predicha más moderados.

5.2.6.1.3 Influencia del fragmento derivado del aldehído (R^2)

El análisis del fragmento derivado del aldehído mostró que ciertas sustituciones aromáticas en R^2 se asociaron con valores de pIC_{50} predicho más elevados dentro de distintas familias estructurales.

Entre los fragmentos aldehídicos evaluados, las combinaciones que incorporan 4-isopropilbenzaldehído y cinamaldehído aparecieron con mayor frecuencia entre las moléculas con valores de actividad predicha más altos dentro del conjunto evaluado. Por otra parte el fragmento p-(trifluorometil)benzaldehído destaca en el QSAR evolutivo mas no exhibe una tendencia alta en el QSAR manual.

5.2.6.1.4 Comparación entre los modelos QSAR

Los dos modelos QSAR mostraron diferencias en la magnitud absoluta de los valores de pIC_{50} predichos, siendo el modelo manual el que presentó valores globalmente más elevados que el modelo evolutivo. Sin embargo, al analizar las tendencias estructurales dentro del conjunto evaluado, ambos modelos mostraron coincidencias en varios aspectos.

Ambos modelos identificaron una mayor actividad predicha en las familias heterocíclicas derivadas en comparación con las chalconas precursoras, así como una tendencia favorable asociada a fragmentos aromáticos extendidos en R^1 y a determinadas sustituciones aromáticas en R^2 .

Los valores completos de pIC_{50} predicho para las 400 estructuras, junto con la familia estructural correspondiente y los fragmentos R^1 y R^2 asociados, se presentan en el Anexo II.

5.2.6.2 Compuestos derivados de benzimidazoles

Otra familia evaluada correspondió a los derivados de benzimidazol, cuya estructura general se muestra en la Figura 27. Esta serie fue incluida dentro del análisis predictivo debido a que incorpora un núcleo aromático rígido y una elevada densidad de átomos de nitrógeno, características estructurales de interés dentro del diseño de heterociclos potencialmente bioactivos. Los valores de pIC_{50} predicho obtenidos para esta familia mediante ambas ecuaciones QSAR se presentan en la Tabla 5.

Estructura general de los derivados de benzimidazol

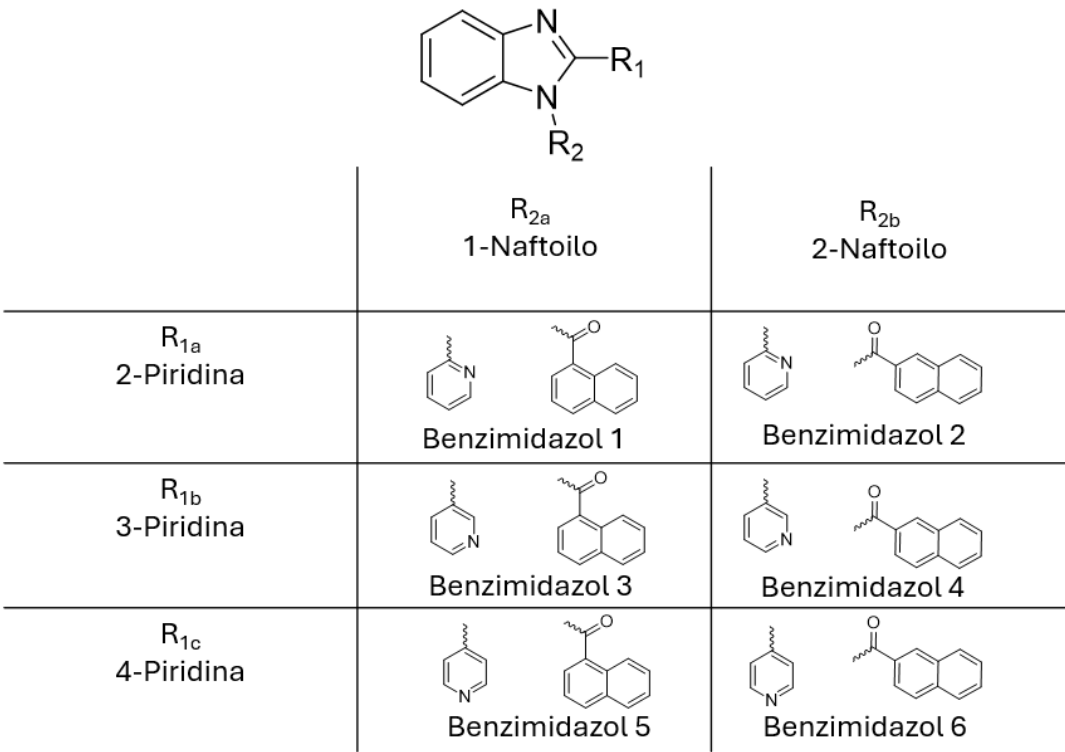


Figura 27 Esquema estructuras de benzimidazoles evaluados

Molécula	pIC ₅₀ QSAR Evolutivo	pIC ₅₀ QSAR manual
Benzimidazol 1	5.258818	6.78481339
Benzimidazol 2	5.364227	6.79515075
Benzimidazol 3	5.314317	6.77879663
Benzimidazol 4	5.413341	6.79019524
Benzimidazol 5	5.389899	6.74801064
Benzimidazol 6	5.490197	6.76076754

Tabla 5 Benzimidazole con usvalor de pIC₅₀ en ambos modelos de QSAR

En términos generales, ambas ecuaciones predijeron actividades dentro de un intervalo relativamente estrecho para los seis compuestos analizados, lo que sugiere un comportamiento homogéneo de la serie. La ecuación QSAR evolutiva entregó valores comprendidos entre 5.26 y 5.49, mientras que la ecuación QSAR manual predijo valores más altos, en un rango de 6.75 a 6.80. Por tanto, aunque ambas aproximaciones difirieron en la magnitud absoluta de la actividad estimada, coincidieron en que las variaciones estructurales dentro de esta familia producen diferencias moderadas más que cambios drásticos en la respuesta predicha.

Respecto a la priorización interna de los compuestos, ambas ecuaciones mostraron coincidencias parciales. En ambos casos, Benzimidazol 2 y Benzimidazol 4 se ubicaron entre los compuestos con mejores valores predichos de actividad, lo que sugiere una tendencia favorable asociada a estas combinaciones de sustituyentes. En contraste, Benzimidazol 1 presentó uno de los valores más bajos en ambas aproximaciones. Sin embargo, también se observaron diferencias entre los modelos, particularmente en los compuestos 5 y 6: mientras la ecuación evolutiva predijo el mejor valor para Benzimidazol 6, la ecuación manual favoreció levemente a Benzimidazol 2 y Benzimidazol 4.

5.3 Virtual screening

5.3.1 Validación molécula referencia GRL-0617

El docking rígido realizado sobre GRL-0617 generó diez poses de unión. La pose reportada por Autodock gpu presentó un puntaje de -7.83 y un RMSD de 0,43 Å respecto a la conformación cristalográfica del ligando de referencia en la estructura 7CJM, lo que indica una excelente reproducción del modo de unión experimental, además demoró 0.6s en terminar el cálculo.

La Figura 28 muestra la superposición entre la estructura cristalográfica de PLpro en complejo con GRL-0617 y la pose obtenida mediante docking rígido. La proteína cristalográfica se representa en color gris,

mientras que la estructura obtenida por docking se muestra en verde. La coincidencia espacial observada en el sitio activo evidencia una alta concordancia entre ambas conformaciones, consistente con el bajo valor de RMSD obtenido (0,43 Å).

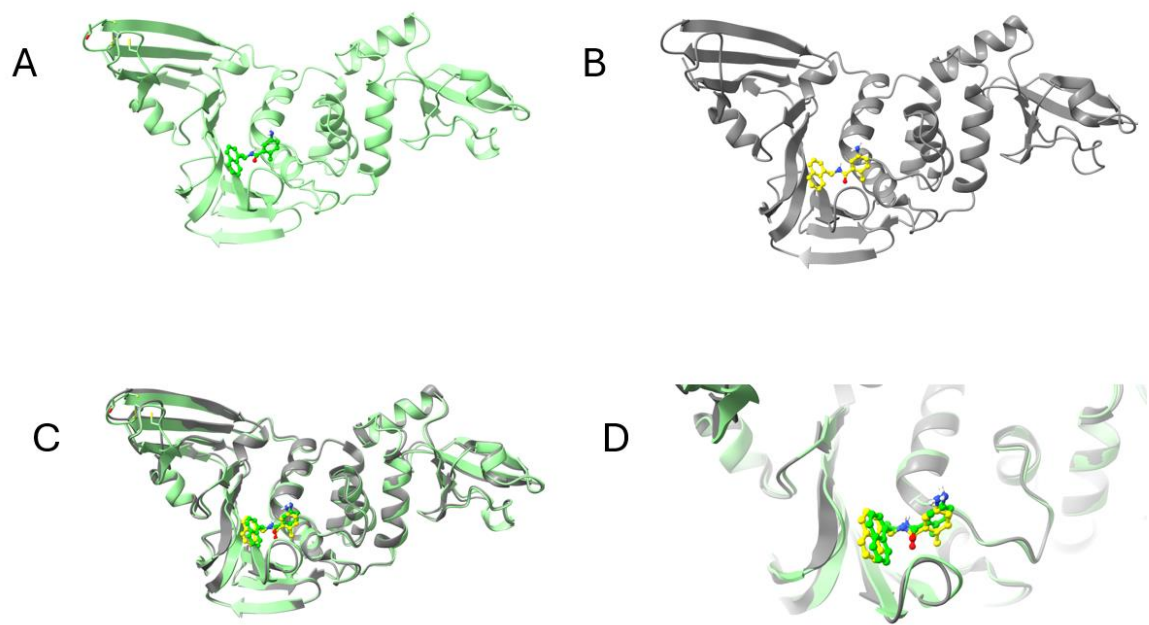


Figura 28 Superposición estructural entre la conformación cristalográfica de PLpro en complejo con GRL-0617 y la pose obtenida mediante docking rígido. (A) Estructura cristalográfica del complejo PLpro–GRL-0617. (B) Pose mejor clasificada obtenida por docking rígido. (C) Superposición de ambas estructuras, donde la proteína cristalográfica se muestra en gris y la estructura obtenida por docking en verde. (D) Vista ampliada del sitio activo mostrando la coincidencia espacial entre el ligando cristalográfico (amarillo) y la pose dockeada (verde).

5.3.2 Desempeño análisis molecular

Las familias químicas evaluadas incluyeron aminopirimidinas, benzimidazoles, benzodiazepinas, benzamidas, benzotiofenos y derivados pirimidina–benzimidazol, así como sistemas fusionados tipo pirimido–benzimidazol (Figura 29). Con el fin de estimar la disponibilidad química y la viabilidad de cribado virtual, se consultó la plataforma PubChem en mayo de 2025, la cual reportó el número de compuestos disponibles para cada familia (Tabla 6). Asimismo, se registraron los tiempos de procesamiento requeridos para el análisis de cada conjunto molecular, los cuales se presentan de forma comparativa en la Tabla 6.

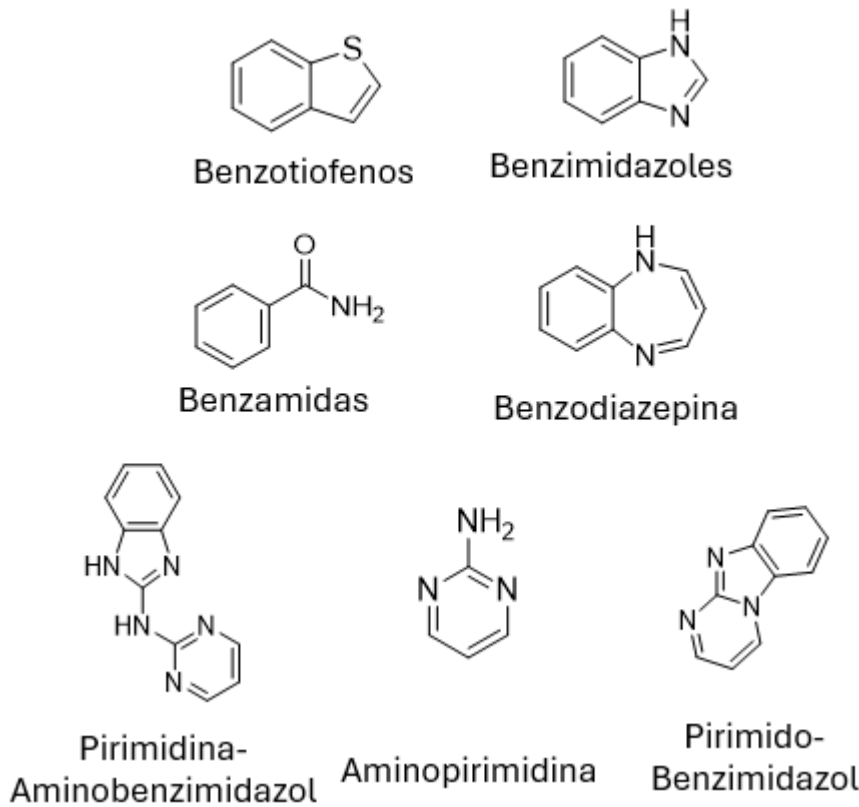


Figura 29 Núcleos heterocíclicos presentes en el análisis de virtual screening.

Tabla 6 Nombres de familia química analizada, el número de moléculas analizadas y el tiempo de computo.

Nombre de familia Química	Numero reportado por Pubchem	Tiempo de computo
Aminopirimidinas	778	14min 53s
Benzimidazoles	334974	7 días 12h 59min 54s
Benzodiazepinas	15187	6h 11min 52s
Benzamidas	2669895	34 días 8h 38min 6s
Benzotiofeno	288626	6 días 12h 21min 22s
Pirimidina–aminobenzimidazol	627	10min 39s
Pirimido-benzimidazol	2080	33min 12s

5.3.3 Ranking por afinidad

Los compuestos pertenecientes a las siete familias de heterociclos fueron analizados mediante estudios de acoplamiento molecular. Los puntajes correspondientes a las diez moléculas mejor clasificadas se resumen en la Tabla 7, mientras que sus estructuras químicas se muestran en la Figura 30.

Tabla 7 Puntaje del virtual screening obtenido.

Ranking	Aminopirimidinas	Benzimidazoles	Benzodiazepinas	Benzamidas
1	-5.29	-9.34	-14.00	-8.86
2	-5.17	-9.26	-13.46	-8.81
3	-5.12	-9.19	-12.57	-8.78
4	-5.08	-9.17	-12.53	-8.78
5	-5.08	-9.17	-12.33	-8.74
6	-5.07	-9.02	-12.28	-8.72
7	-5.06	-9.01	-12.23	-8.71
8	-5.03	-8.97	-12.23	-8.70
9	-5.02	-8.93	-12.18	-8.69
10	-4.96	-8.93	-12.16	-8.69

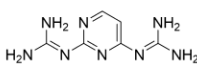
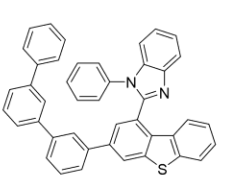
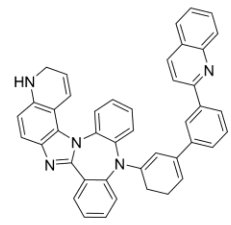
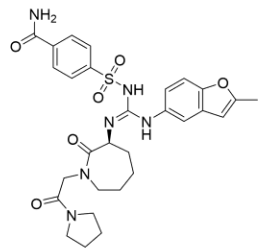
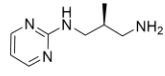
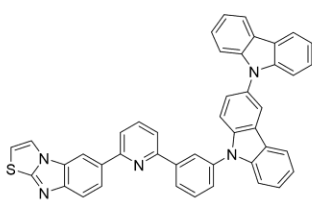
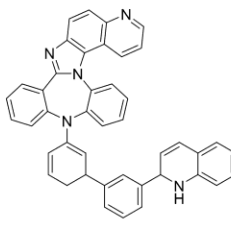
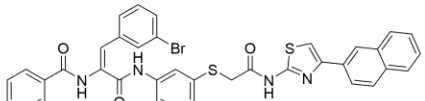
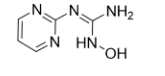
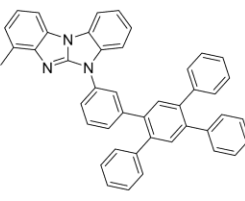
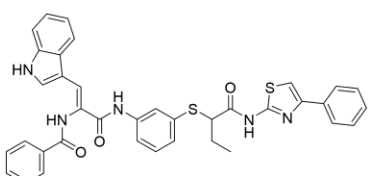
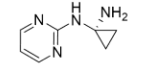
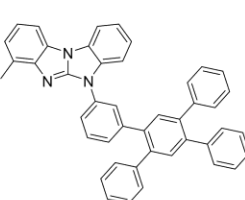
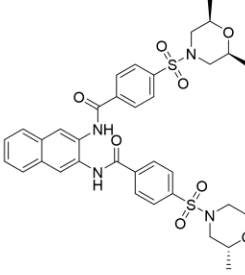
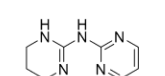
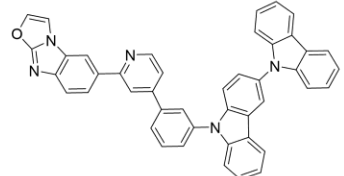
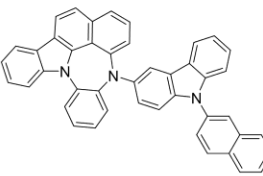
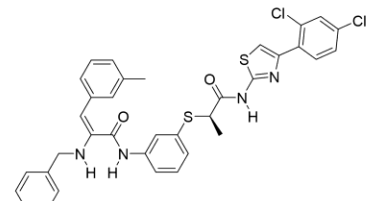
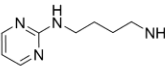
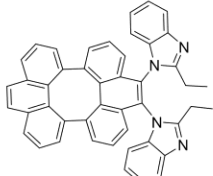
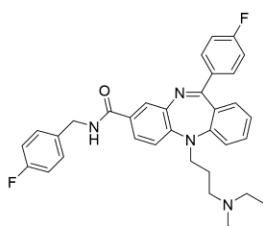
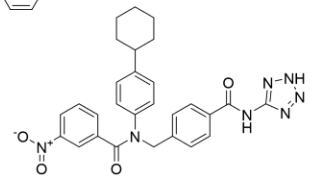
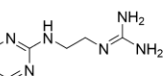
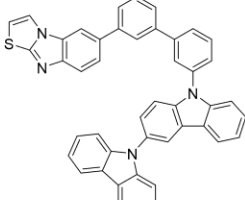
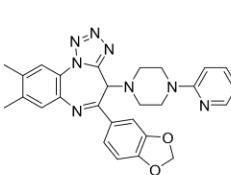
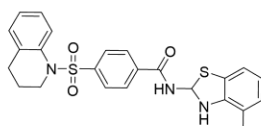
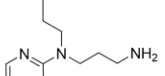
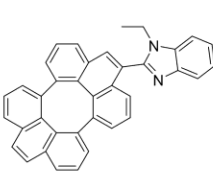
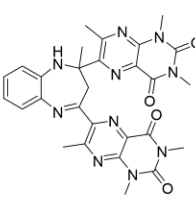
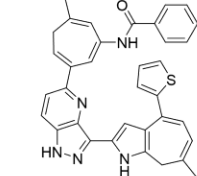
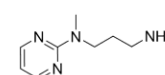
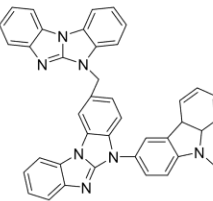
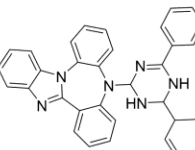
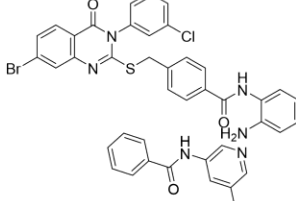
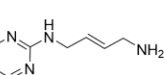
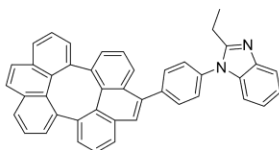
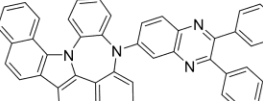
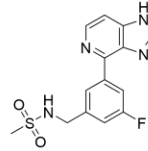
Ranking	Benzotiofeno	Pirimidina–aminobenzimidazol	Pirimido-benzimidazol
1	-9.50	-11.33	-12.58
2	-9.29	-11.27	-11.97
3	-9.16	-11.22	-11.72
4	-8.72	-11.16	-11.62
5	-8.66	-11.09	-11.56
6	-8.60	-11.06	-11.35
7	-8.59	-11.01	-11.34
8	-8.57	-11.01	-11.32
9	-8.54	-11.01	-11.13
10	-8.53	-10.98	-11.03

Aminopirimidinas

Benzimidazoles

Benzodiazepina

Benzamidas

1				
2				
3			*1	
4			*1	
5				
6				
7				
8				
9				
10				

* Misma entidad química; diferencias solo en orientación espacial.

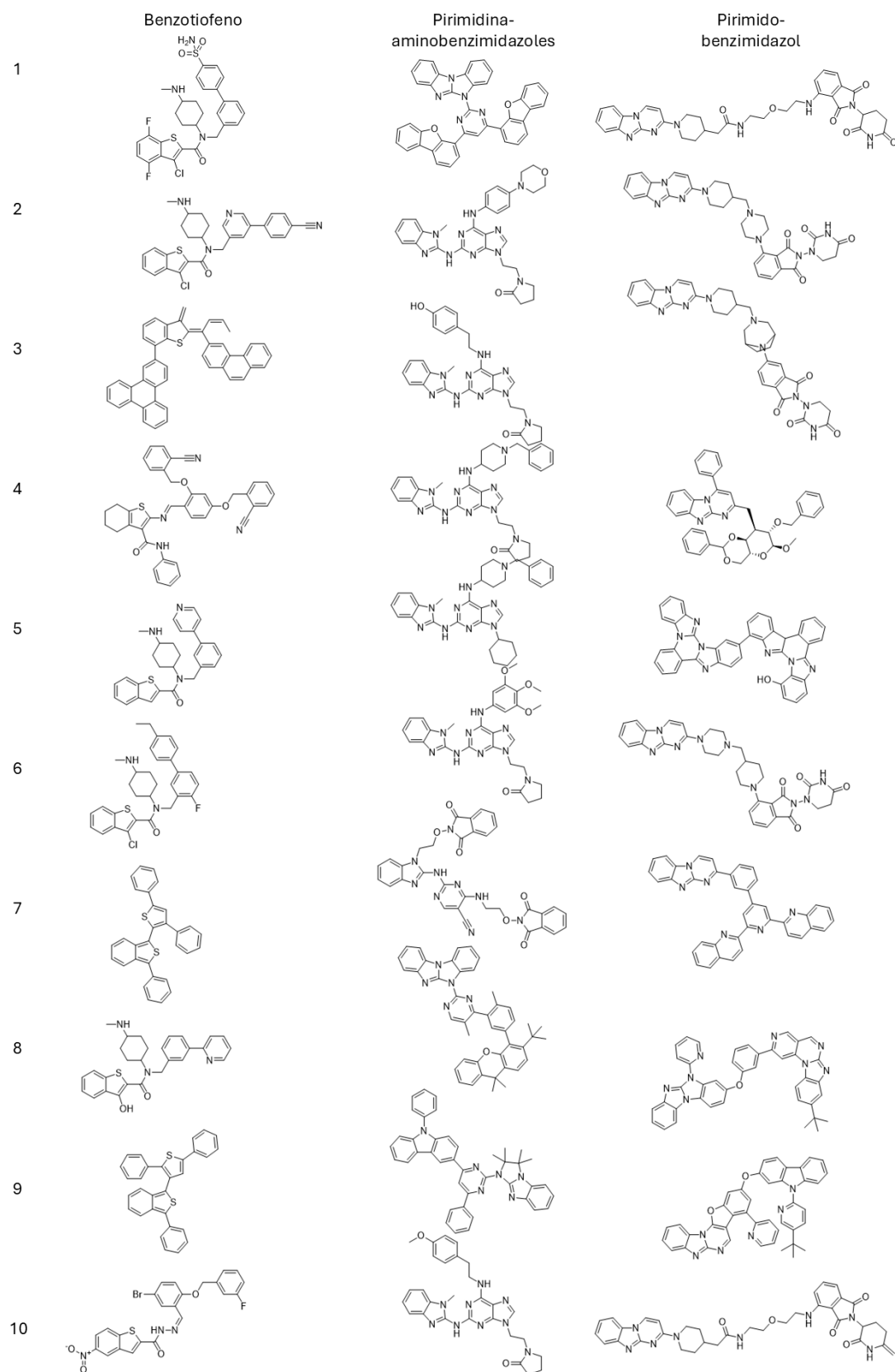


Figura 30 Estructura del top 10 de cada familia molecular.

La Figura 31 muestra la distribución de afinidad de las moléculas mejor posicionadas dentro del Top 10.

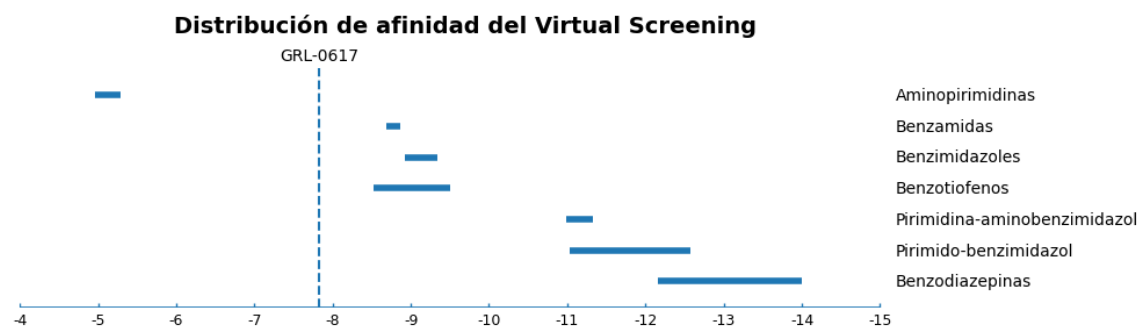


Figura 31 Distribución de afinidad de las moléculas Top 10 por familia química. La línea discontinua indica la afinidad del ligando de referencia GRL-0617.

5.3.4 Comparación de GRL-0617 poses top 1

La Figura 32A muestra la superposición de las poses top-1 obtenidas para los compuestos seleccionados dentro de la cavidad asociada al sitio alostérico cercano al aminoacido Tyr268 de PLpro. En los paneles B–H, la superposición con el ligando de referencia GRL-0617 (representado en gris translúcido) permite comparar la orientación relativa de cada familia dentro del sitio activo. Las aminopirimidinas (B) presentan una ocupación parcial del bolsillo y un solapamiento limitado con el ligando de referencia. Las benzodiazepinas (D) muestran una orientación más definida dentro de la cavidad, con sus anillos aromáticos ubicados en regiones similares a las ocupadas por GRL-0617. Los benzimidazoles (C), benzotiofenos (F) y los híbridos pirimidina–aminobenzimidazol (G) y pirimido-benzimidazol (H) ocupan extensamente el bolsillo y presentan un solapamiento parcial con el ligando de referencia. Por su parte, las benzamidas (E) adoptan orientaciones más extendidas hacia regiones periféricas del sitio.

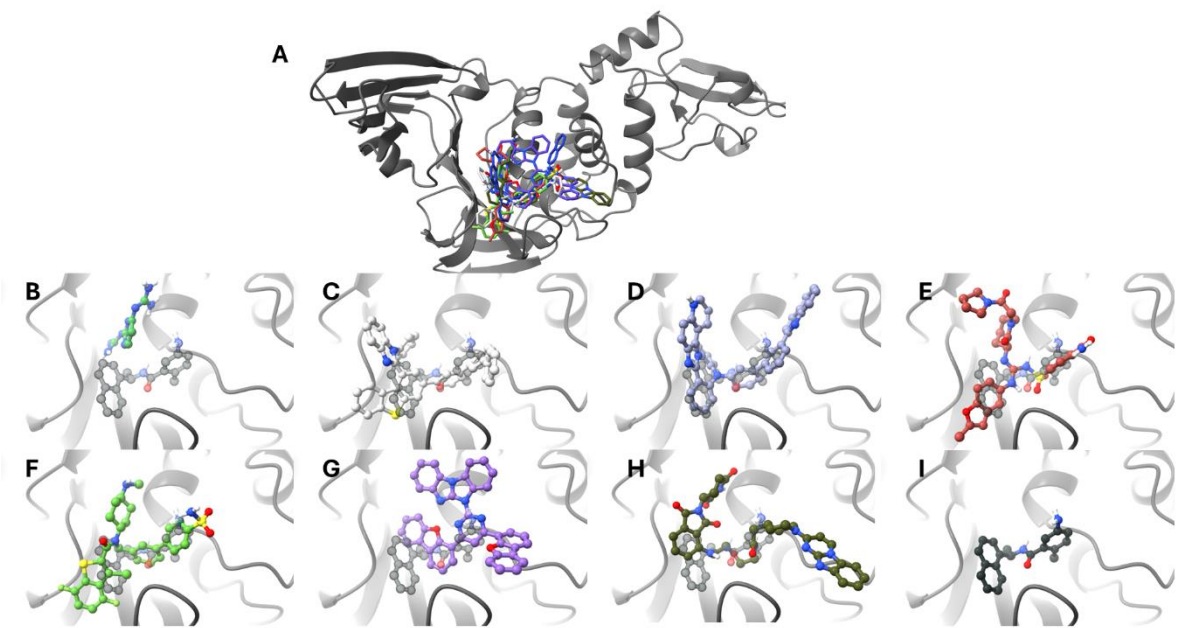


Figura 32 A) Vista general del sitio de unión alostérico de PLpro mostrando la superposición de las poses Top 1 obtenidas para los compuestos seleccionados dentro del bolsillo de unión. (B) Aminopirimidinas.(C) Benzimidazoles.(D) Benzodiazepinas.(E) Benzamidas.(F) Benzotiofenos.(G) Pirimidina-aminobenzimidazoles.(H) Pirimido-benzimidazoles.(I) Ligando de referencia GRL-0617.En los paneles (B–H), cada ligando se muestra superpuesto con el ligando de referencia GRL-0617, representado en color gris translúcido, con el fin de facilitar la comparación espacial y evaluar la coincidencia en el modo de unión dentro del sitio activo.

5.3.5 Predicción Virtual screening ecuaciones QSAR

La Tabla 8 muestra los valores de pIC₅₀ predicho para las moléculas finales calculados mediante los modelos QSAR manual y QSAR evolutivo. El modelo manual presentó valores entre 6.151850791 y 10.91787802, mientras que el modelo evolutivo mostró valores entre 4.0298087 y 7.5065760. En ambos modelos, aminopirimidina correspondió al compuesto con menor actividad predicha. Asimismo, benzimidazol se mantuvo entre los compuestos con mayor actividad predicha en ambos enfoques, mostrando valores muy próximos entre sí. Benzodiazepina y benzamida también presentaron una ubicación intermedia a alta en ambas predicciones. En contraste, las mayores diferencias entre los

modelos se observaron para pirimido-benzimidazol y pirimidina-aminobenzimidazol, que presentaron los valores más altos en el modelo manual, pero no en el modelo evolutivo.

Tabla 8 Valor predicho de pIC₅₀ para el Top 1 de las moléculas del Virtual Screening.

Molécula Top1	pIC ₅₀ QSAR manual	pIC ₅₀ QSAR evolutivo
Aminopirimidina	6.151850791	4.0298087
Benzimidazol	7.476959642	7.5065760
Benzamida	7.023729789	6.3906366
Benzodiazepina	7.747255524	6.6246253
Pirimidina-aminobenzimidazol	10.37493948	7.2381732
Pirimido-Benzimidazol	10.91787802	6.6077736

5.4 Docking inducido (*induced fit docking*)

5.4.1 Validación molécula referencia GRL-0617

Con el fin de evaluar el desempeño del procedimiento de *Induced Fit Docking*, se utilizó el inhibidor de referencia GRL-0617, cuyo modo de unión se compara con un el modelo cristalizado. La Figura 33 muestra la superposición entre la estructura cristalográfica del complejo proteína-ligando y la pose obtenida mediante *Induced Fit Docking*. En esta representación, la proteína correspondiente a la estructura cristalográfica se muestra en color gris, mientras que la estructura optimizada mediante el protocolo de *Induced Fit Docking* se representa en color verde. En la parte inferior de la figura se observa la superposición de ambas poses del ligando dentro del sitio de unión.

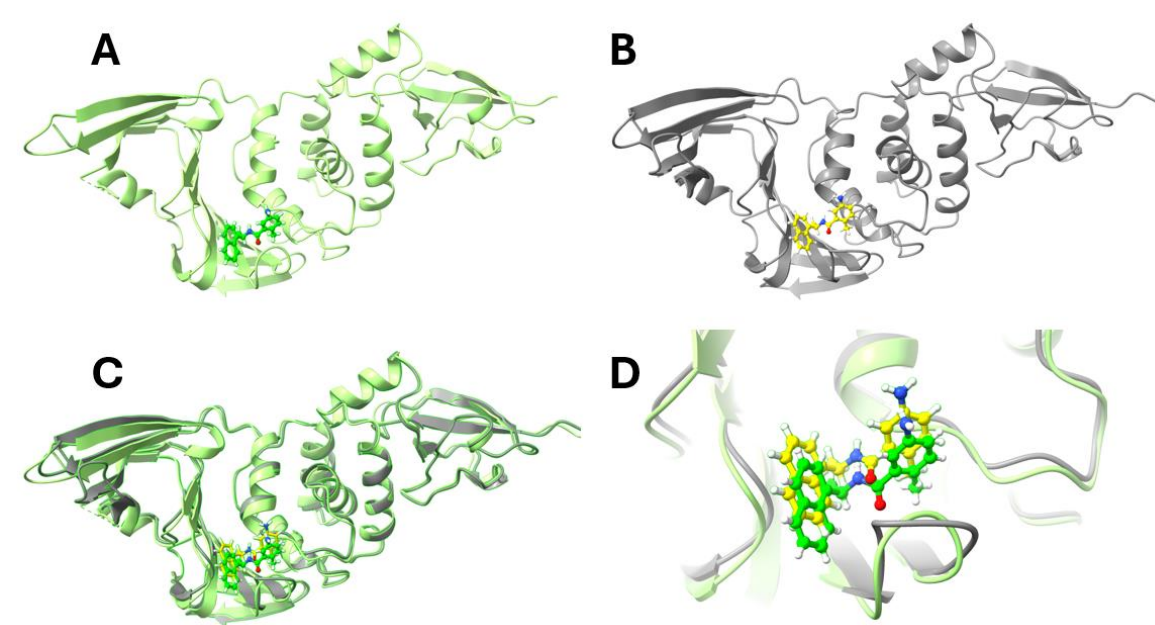


Figura 33 Validación del protocolo de *Induced Fit Docking* utilizando el inhibidor de referencia GRL-0617. (A) Estructura del complejo PLpro-GRL-0617 obtenida a partir de la estructura cristalográfica, mostrando el ligando en el sitio de unión. (B) Pose de GRL-0617 obtenida mediante *Induced Fit Docking*, representada dentro del sitio de unión de PLpro. (C) Superposición entre la estructura cristalográfica (gris) y la estructura obtenida mediante *Induced Fit Docking* (verde), mostrando la coincidencia entre ambas conformaciones del complejo proteína-ligando. (D) Vista ampliada del sitio de unión mostrando la superposición de la pose cristalográfica y la pose obtenida mediante *Induced Fit Docking* para el ligando GRL-0617.

El docking de GRL-0617 produjo un XP GScore de -5.572 y un RMSD de 1.679 Å respecto a la pose cristalográfica. Adicionalmente, se generaron diagramas de interacción bidimensional para la estructura cristalográfica y para la pose obtenida mediante *Induced Fit Docking*, los cuales se muestran en la Figura 34. Estos diagramas permiten visualizar las interacciones establecidas entre el ligando y los residuos del sitio de unión.

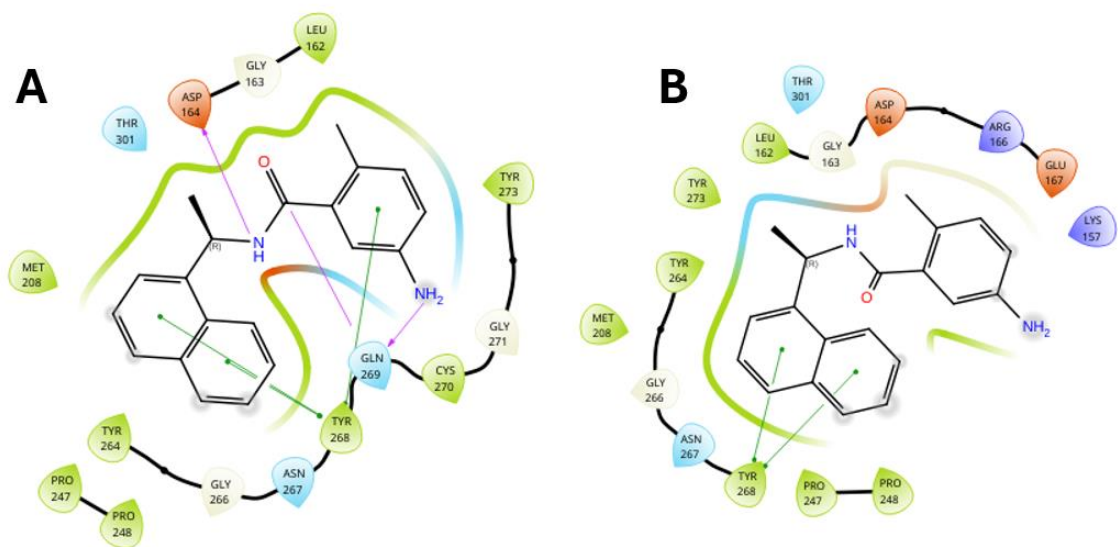


Figura 34 Comparación de las interacciones ligando–proteína para GRL-0617 en la estructura cristalográfica y en la pose obtenida mediante Induced Fit Docking. (A) Diagrama de interacción bidimensional correspondiente al complejo cristalográfico PLpro–GRL-0617, mostrando los residuos del sitio de unión que interactúan con el ligando. (B) Diagrama de interacción bidimensional correspondiente a la pose de GRL-0617 obtenida mediante Induced Fit Docking, donde se observan los residuos del bolsillo de unión involucrados en interacciones con el ligando.

5.4.2 Docking molecular de chalconas y derivados de chalconas en PLpro

Con el objetivo de evaluar el comportamiento de unión de la biblioteca molecular diseñada frente al sitio activo de PLpro, se analizaron los valores de XP GScore obtenidos mediante docking molecular para todas las estructuras generadas. La biblioteca está compuesta por derivados de chalcona y siete familias heterocíclicas derivadas de estas, cuyas estructuras generales se muestran en la Figura 25. Estas moléculas fueron generadas mediante una estrategia combinatoria basada en la variación de dos fragmentos estructurales principales: R¹, derivado de la cetona aromática utilizada en la condensación de Claisen–Schmidt, y R², derivado del aldehído aromático correspondiente.

Las 400 estructuras generadas fueron sometidas a estudios de docking molecular en el sitio alostérico adyacente al residuo Tyr268 de PLpro. Para cada compuesto se consideró el XP GScore correspondiente a la mejor pose de unión. El listado completo de las moléculas evaluadas junto con sus respectivos valores de XP GScore se presenta en el Anexo II.

La distribución global de los valores de XP GScore obtenidos para la biblioteca molecular se muestra en la Figura 35. En esta representación se observa que una proporción importante de las moléculas evaluadas presenta valores de docking más favorables que el obtenido para el ligando de referencia.

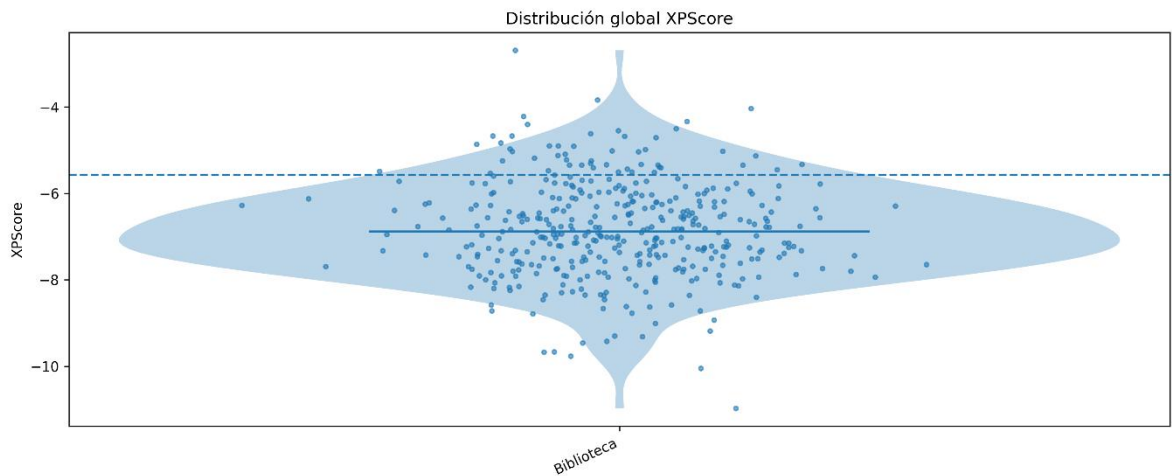


Figura 35 Distribución global de chalconas y derivados de chalconas con respecto al XP GScore la línea de referencia es el ligando GRL-0617 con su puntaje de XP GScore :-5.5720.

Al analizar la distribución de los valores de docking según las familias estructurales evaluadas, se observa que todas las familias presentan compuestos con valores de XP GScore más negativos que el del ligando

de referencia GRL-0617 (-5.572). En cada familia se identifican múltiples estructuras con energías de docking inferiores a dicho valor, como se muestra en la Figura 36.

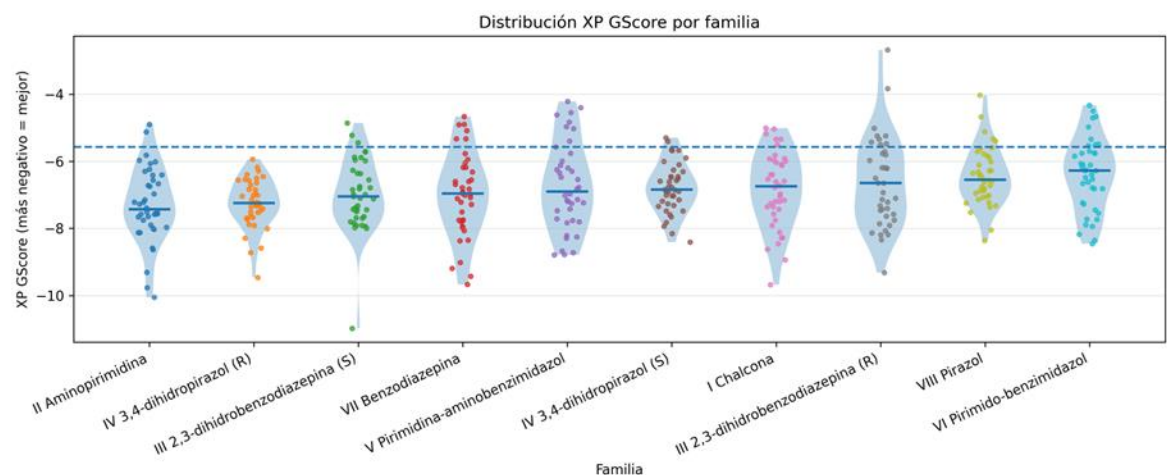


Figura 36 Distribución de los valores de XP GScore obtenidos mediante docking molecular para las diferentes familias estructurales evaluadas. Cada punto representa el valor correspondiente a la mejor pose de unión de una molécula individual. La línea punteada indica el valor de XP GScore del ligando de referencia GRL-0617.

En particular, las familias pirimidina-aminobenzimidazol, pirimido-benzimidazol y benzodiazepinas aromáticas presentan una proporción importante de compuestos con valores de XP GScore inferiores a -7 , observándose además varias estructuras con valores cercanos o inferiores a -9 dentro del conjunto analizado.

Las familias chalconas y aminopirimidinas muestran igualmente distribuciones amplias de valores de docking, con múltiples compuestos localizados por debajo del valor correspondiente al ligando de referencia. De manera similar, las familias 2,3-dihidrobenzodiazepinas y 3,4-dihidropirazoles presentan distribuciones de XP GScore que abarcan un rango amplio dentro de la biblioteca evaluada.

La distribución de los valores de XP GScore en función del fragmento R¹ muestra variaciones moderadas entre los diferentes sistemas aromáticos derivados de las cetona evaluadas, como se observa en la Figura 37. En todos los casos se identifican múltiples compuestos con valores de docking más negativos que el del ligando de referencia GRL-0617.

Los fragmentos derivados de pireno, antraceno y naftaleno presentan distribuciones de XP GScore con medianas cercanas a -7 ,mostrando además una dispersión amplia de valores dentro de cada grupo estructural. Por su parte, el fragmento derivado de fenilo presenta una distribución de valores similar, con múltiples compuestos ubicados en el rango aproximado de -6 a -8 dentro del conjunto evaluado.

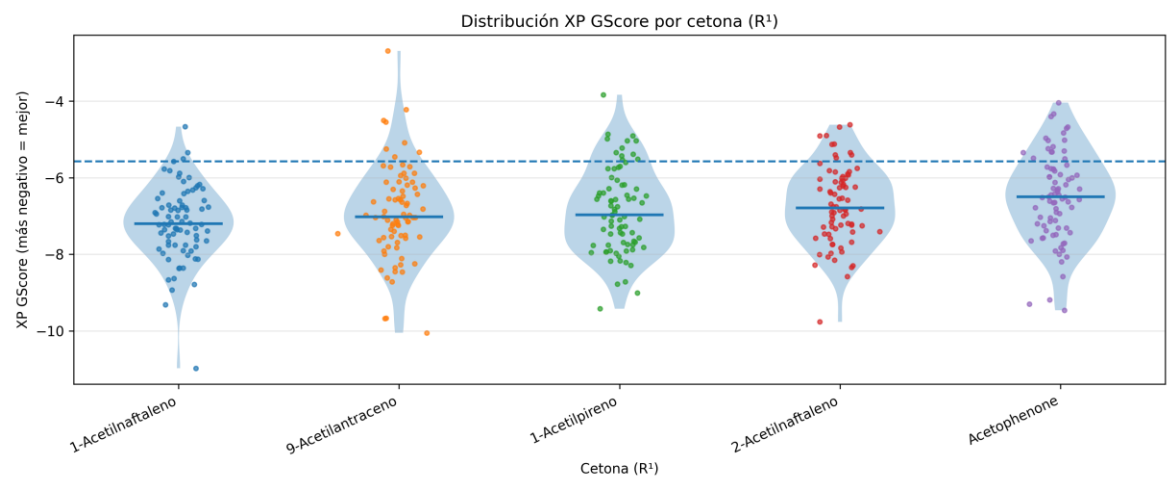


Figura 37 Distribución de los valores de XP GScore obtenidos mediante docking molecular en función del fragmento R¹ derivado de la cetona utilizada en la condensación de Claisen-Schmidt. Cada punto representa el valor correspondiente a la mejor pose de unión de una molécula individual. La línea punteada indica el valor de XP GScore del ligando de referencia GRL-0617.

La distribución de los valores de XP GScore en función del fragmento R² muestra variaciones entre los diferentes sustituyentes aromáticos derivados de los aldehídos evaluados, como se observa en la Figura 38. En todos los casos se identifican compuestos con valores de docking más negativos que el del ligando de referencia GRL-0617 .

Los sustituyentes p-nitrofenilo, cinamilo y p-isopropilfenilo presentan distribuciones de XP GScore con medianas cercanas a -6.5 a -7 , observándose además múltiples estructuras con valores inferiores a -7 dentro del conjunto evaluado.

De manera similar, los fragmentos p-dimetilaminofenilo, 4-hidroxi-3-metoxifenilo, o-hidroxifenilo, p-aminofenilo y p-trifluorometilfenilo presentan distribuciones amplias de valores de docking dentro de rangos comparables, con compuestos distribuidos aproximadamente entre -6 y -8.5.

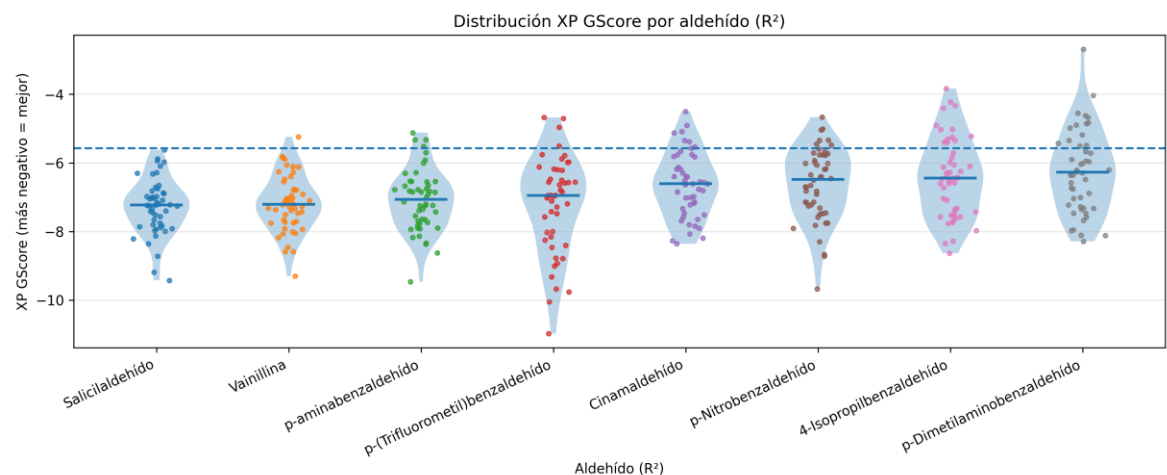


Figura 38 Distribución de los valores de XP GScore obtenidos mediante docking molecular en función del fragmento R^2 derivado del aldehído utilizado en la condensación de Claisen–Schmidt. Cada punto representa el valor correspondiente a la mejor pose de unión de una molécula individual. La línea punteada indica el valor de XP GScore del ligando de referencia GRL-0617.

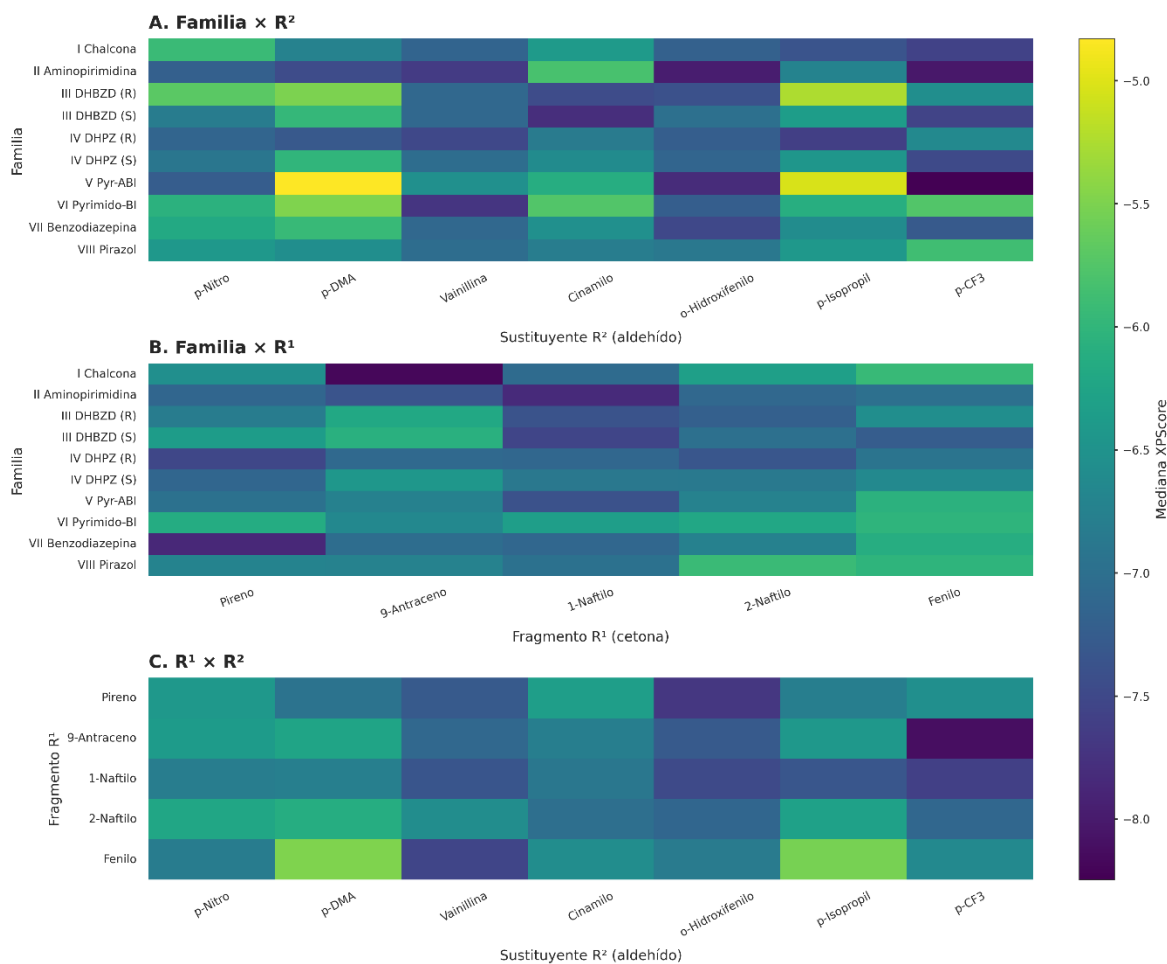
Con el fin de visualizar el efecto conjunto de la familia estructural y de los fragmentos aromáticos presentes en la biblioteca evaluada, se calcularon las medianas de XP GScore para las diferentes combinaciones estructurales, cuyos resultados se muestran en la Figura 39. El análisis de los mapas de calor revela variaciones en los valores medianos de docking dependiendo tanto de la familia heterocíclica como de los fragmentos derivados de la cetona (R^1) y del aldehído (R^2).

En relación con el efecto del sustituyente R^2 , se observan diferencias en los valores medianos de XP GScore entre las distintas familias estructurales. En general, las familias pirimidina-aminobenzimidazol, pirimido-benzimidazol y dihidrobenzodiazepina presentan valores medianos más favorables en comparación con otras familias de la biblioteca. Dentro de estas familias destacan particularmente las combinaciones que incluyen los sustituyentes p-dimetilaminofenilo (p-DMA) y p-isopropilfenilo, que muestran valores medianos más negativos en varias de las familias evaluadas. En contraste, los sustituyentes o-hidroxifenilo y p- CF_3 presentan en algunos casos valores medianos menos favorables, especialmente dentro de las familias aminopirimidina y chalcona. Por otra parte, las familias dihidropirazol y pirazol muestran una distribución más homogénea de los valores medianos de XP GScore entre los diferentes fragmentos R^2 considerados.

Por otro lado, el análisis del fragmento R^1 derivado de la cetona muestra que los diferentes sistemas aromáticos evaluados presentan valores medianos relativamente comparables entre sí en varias de las familias estructurales. Los fragmentos derivados de pireno, antraceno y naftaleno muestran distribuciones similares de XP GScore dentro de varias familias, mientras que el fragmento fenilo presenta en algunos casos valores medianos ligeramente más favorables, particularmente dentro de las familias pirimidina-aminobenzimidazol, pirimido-benzimidazol y pirazol. En contraste, dentro de la familia chalcona el fragmento antraceno muestra valores medianos menos favorables en comparación con los demás fragmentos derivados de la cetona evaluados.

Finalmente, el análisis de las combinaciones directas entre los fragmentos R^1 y R^2 evidencia variaciones moderadas en los valores medianos de XP GScore dependiendo de la combinación específica de fragmentos aromáticos. Entre las combinaciones que presentan valores medianos más favorables se encuentran aquellas que incluyen el sustituyente p-DMA o p-isopropilfenilo en R^2 combinados con fragmentos aromáticos en R^1 como fenilo o algunos sistemas policíclicos. En contraste, combinaciones que incluyen los sustituyentes p- CF_3 u o-hidroxifenilo presentan en varios casos valores medianos menos favorables, particularmente cuando se combinan con el fragmento antraceno. En general, los fragmentos

derivados de naftaleno muestran valores relativamente consistentes entre los distintos sustituyentes R^2 considerados.



estructural. Con el fin de mejorar la legibilidad de la figura se utilizaron abreviaciones para los fragmentos aromáticos, donde Pireno corresponde al derivado de acetilpireno, 9-Antraceno al derivado de acetilantraceno, 1-Naftilo y 2-Naftilo a los derivados de 1- y 2-acetilnaftaleno respectivamente y Fenilo al derivado de acetofenona. De manera análoga, los fragmentos R^2 se representan como p-Nitro para p-nitrofenilo, p-DMA para p-dimetilaminofenilo, Vainillina para el fragmento 4-hidroxi-3-metoxifenilo, Cinamilo para el sustituyente cinamilo, o-Hidroxifenilo para el derivado orto-hidroxifenilo, p-Isopropil para p-isopropilfenilo, p-Amino para p-aminofenilo y p-CF₃ para p-trifluorometilfenilo.

5.4.3 Docking molecular de derivados de benzimidazol

Los valores de afinidad obtenidos mediante docking molecular para los derivados benzimidazol se resumen en la Tabla 9, mostrando una variación moderada en los valores de XP GScore entre los compuestos evaluados. En general, los valores se distribuyen entre -4.352 y -7.259, evidenciando diferencias en la capacidad de acoplamiento dentro del sitio activo de PLpro. Entre los compuestos analizados, benzimidazol 1 presentó el valor más favorable (-7.259), seguido por benzimidazol 5 (-7.013), indicando una mayor afinidad relativa dentro de esta serie. En contraste, benzimidazol 3 mostró el valor menos favorable (-4.352), mientras que benzimidazol 2 también presentó un puntaje relativamente alto (-4.660). Los compuestos benzimidazol 4 y benzimidazol 6 exhibieron valores intermedios, con XP GScores de -6.032 y -5.541, respectivamente.

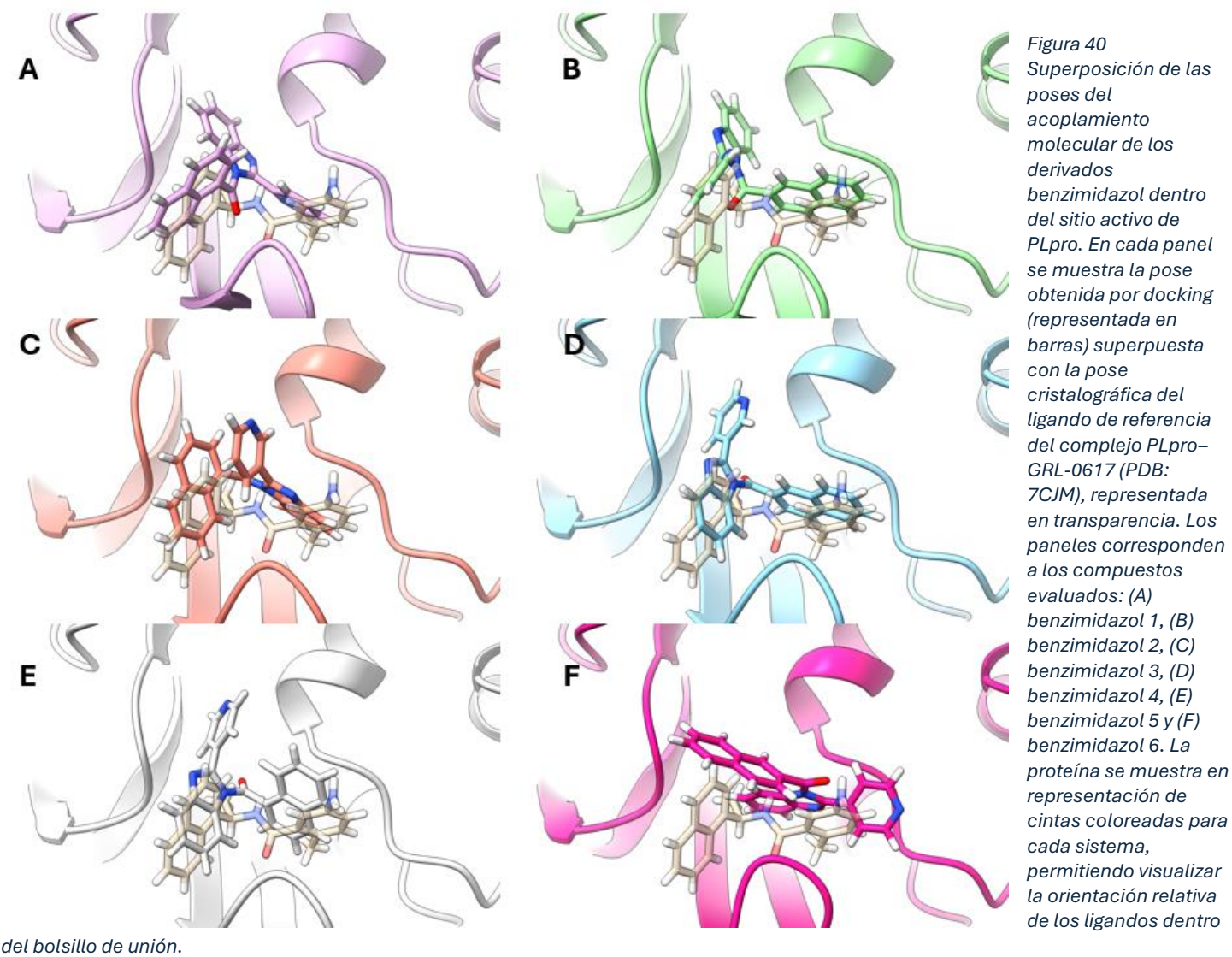
La Figura 40 muestra la superposición de las poses obtenidas por docking para los derivados benzimidazol dentro del bolsillo de unión de PLpro, comparadas con la orientación cristalográfica del ligando de referencia GRL-0617. En términos generales, los compuestos evaluados presentan una orientación consistente dentro del sitio de unión, ocupando una región similar del bolsillo adyacente al loop BL2. En los paneles A-F, correspondientes a benzimidazol 1-6, se observa que el núcleo aromático del sistema benzimidazol se orienta hacia la región hidrofóbica del bolsillo, mientras que los sustituyentes heteroaromáticos se proyectan hacia zonas más expuestas del sitio activo. En particular, los compuestos con valores de docking más favorables (benzimidazol 1 y benzimidazol 5) muestran una superposición

más cercana con la orientación del ligando cristalográfico, manteniendo una disposición espacial similar del sistema aromático principal dentro de la cavidad.

En la Figura 41 se observa que los derivados benzimidazol se acomodan dentro del bolsillo de unión de PLpro y presentan principalmente interacciones hidrofóbicas, junto con contactos puntuales de tipo puente de hidrógeno y π - π stacking. Entre los compuestos evaluados, benzimidazol 1 y benzimidazol 5 muestran una ocupación más definida del bolsillo y conservan un patrón de interacción más favorable. En particular, en benzimidazol 1 se aprecia una interacción por puente de hidrógeno, mientras que en otros derivados, como benzimidazol 5, predominan los contactos aromáticos compatibles con π - π stacking dentro del sitio de unión. En contraste, los compuestos con menor puntaje, como benzimidazol 2 y benzimidazol 3, mantienen su inserción en el bolsillo, pero exhiben un patrón de interacciones más limitado.

Tabla 9 XP GScore Benzimidazoles

Molécula	XP GScore
Benzimidazol 1	-7.259
Benzimidazol 2	-4.660
Benzimidazol 3	-4.352
Benzimidazol 4	-6.032
Benzimidazol 5	-7.013
Benzimidazol 6	-5.541



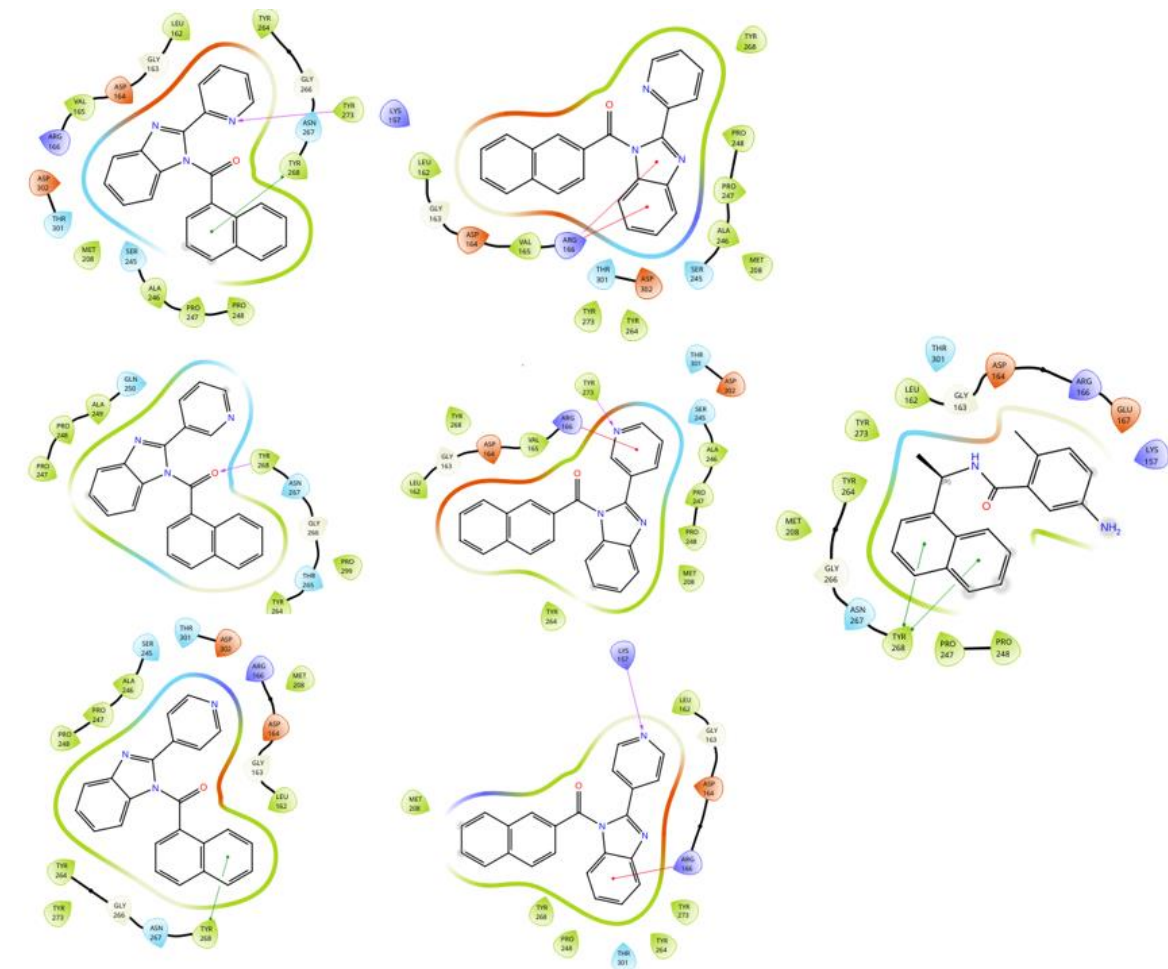


Figura 41. Los diagramas de interacción 2D muestran que todos los derivados benzimidazol se acomodan dentro del bolsillo de unión de PLpro. En general, los compuestos presentan principalmente interacciones hidrofóbicas, junto con contactos específicos de tipo π - π stacking y puentes de hidrógeno. Entre los derivados evaluados, benzimidazol 1 presenta una interacción por puente de hidrógeno, mientras que benzimidazol 5 muestra contactos compatibles con π - π stacking asociados al sistema aromático del ligando. En contraste, benzimidazol 2 y benzimidazol 3 exhiben un patrón de interacción más limitado, consistente con los valores de docking menos favorables observados para estos compuestos

5.4.4 Docking molecular de moléculas Top 1 del Virtual Screening

En el análisis de docking flexible de las moléculas top 1 provenientes del virtual screening, obteniendo diferentes valores de puntaje de interacción mostrados en la Tabla 10, la familia benzimidazol presentó el valor más favorable de XP GScore (-6.858), seguida muy de cerca por aminopirimidina (-5.291). En un segundo nivel se ubicaron benzotiofeno y pirimidina aminobenzimidazol, ambos con un valor de -6.572, mientras que benzodiazepina mostró un puntaje intermedio de -6.386. Los valores menos favorables dentro de este conjunto correspondieron a pirimido benzimidazol (-5.937) y benzamida (-5.895).

Los diagramas bidimensionales (Figura 42) de interacción evidenciaron un patrón recurrente de reconocimiento molecular en torno a residuos del bolsillo, destacando contactos frecuentes con Leu162, Gly163, Asp164, Val165, Arg166, Tyr264, Tyr273, Asn267 y Gln269.

La superposición tridimensional de los complejos (Figura 43) mostró que las distintas familias químicas se localizaron en una región común del sitio de unión, con una conservación general del acomodo global de la proteína y variaciones principalmente en la orientación fina de los ligandos dentro del bolsillo. En todos los casos, los compuestos ocuparon una cavidad comparable a la observada para el ligando de referencia GRL-0617, aunque con diferencias en la extensión espacial de sus sustituyentes y en la manera en que cada andamiaje se proyectó hacia los bordes del sitio de unión.

Tabla 10 valores de Xp Gscore del top1 de cada familia molecular de Virtual Screening.

Molécula top 1 Virtual screening	Xp Gscore
Aminopirimidina	-5.291
Benzimidazol	-6.858

Benzodiazepina	-6.386
Benzamida	-5.895
Benzotiofeno	-6.572
Pirimidina aminobenzimidazol	-6.572
Pirimido benzimidazol	-5.937

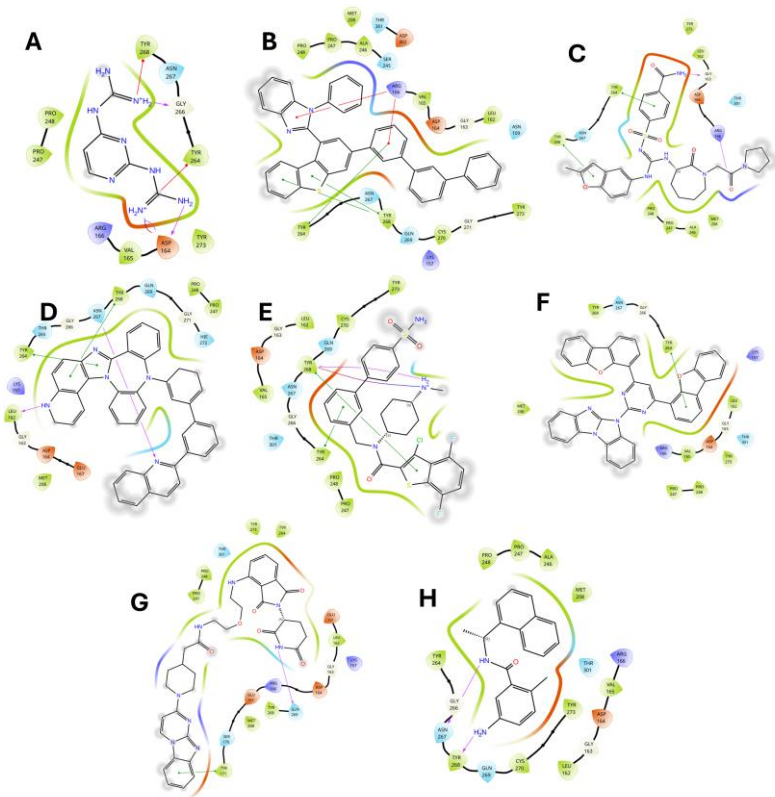


Figura 42 Diagramas bidimensionales de interacción proteína–ligando para los complejos analizados.Los paneles muestran las interacciones del sitio de unión entre la proteína y los distintos compuestos: (A) aminopirimidina, (B) benzimidazol, (C) benzamida, (D) benzodiazepina, (E) benzotiofeno, (F) amino-pirimido-benzimidazol, (G) pirimido-benzimidazol y (H) GRL-0617.

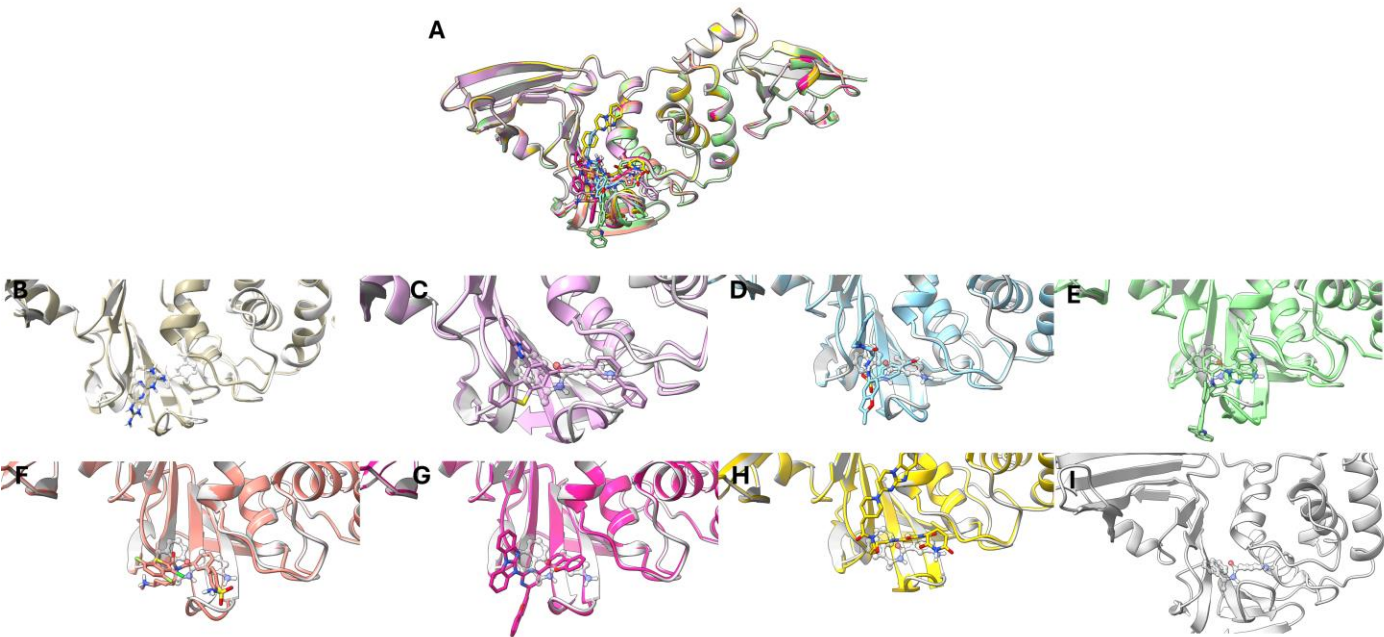


Figura 43 Comparación estructural de los complejos proteína–ligando representativos para las distintas familias químicas analizadas. (A) Superposición global de las conformaciones proteicas en presencia de los distintos ligandos. La proteína se muestra solo el backbone y cada complejo se distingue por un color: aminopiridina (beige), benzimidazol (lila claro), benzodiazepina (celeste), benzotiofeno (salmón), GRL-0617 (blanco), amino-pirimido-benzimidazol (rosado intenso), pirimido-benzimidazol (amarillo) y benzamida (verde). (B) Complejo con aminopiridina. (C) Complejo con benzimidazol. (D) Complejo con benzamida. (E). Complejo con benzodiazepina (F) Complejo con benzotiofeno (G) Complejo con amino-pirimido-benzimidazol. (H) Complejo con pirimido-benzimidazol. (I) Complejo con GRL-0617. En todos los paneles, los ligandos se muestran en representación de sticks, con los heteroátomos coloreados según la convención estándar, también está incluido la posición del ligando de referencia.

5.5 Dinámica molecular

5.5.1 Análisis comportamiento solo proteína

Las simulaciones de Dinámica molecular mostraron una convergencia consistente entre las réplicas, evidenciada por la estabilización del RMSD a lo largo del tiempo de simulación (Figura 44). El análisis de RMSF identificó una región de mayor movilidad en el loop BL2, donde el residuo Tyr268 presentó la fluctuación más elevada, un valor promedio de 5.48 Å, indicando una marcada flexibilidad local (Figura 45).

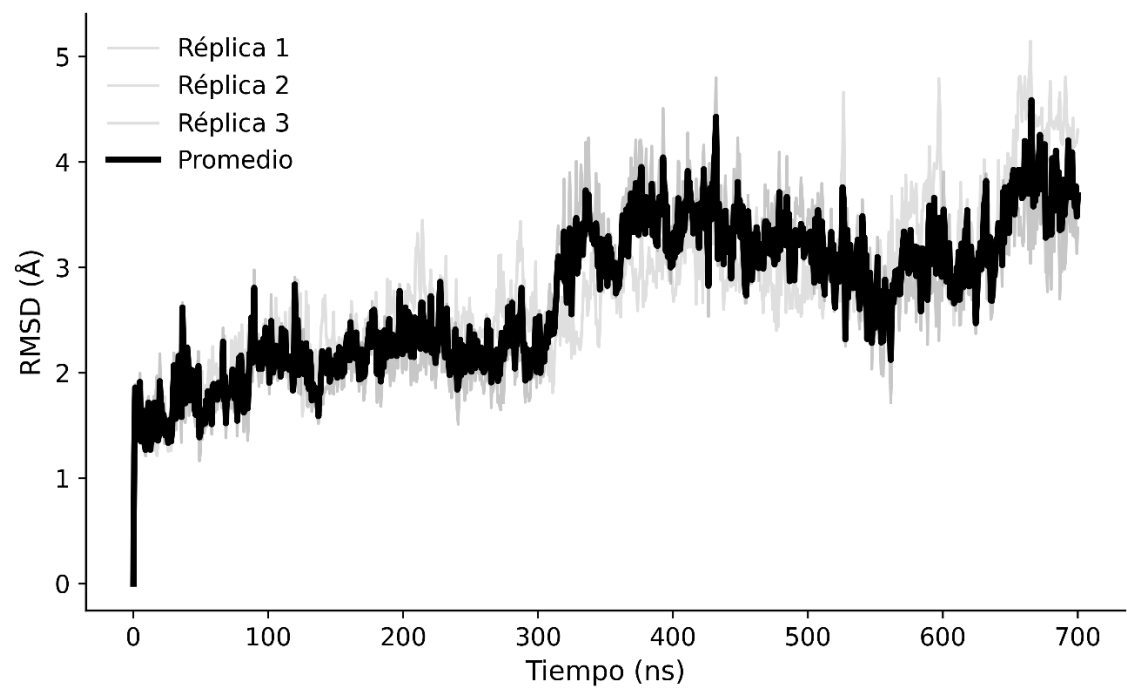


Figura 44 Evolución temporal del RMSD de la proteína durante la simulación de dinámica molecular del sistema APO. Las líneas grises corresponden a las tres réplicas independientes de la simulación, mientras que la línea negra representa el promedio de las trayectorias. El eje x muestra el tiempo de simulación (ns) y el eje y el RMSD (Å) calculado respecto a la estructura inicial.

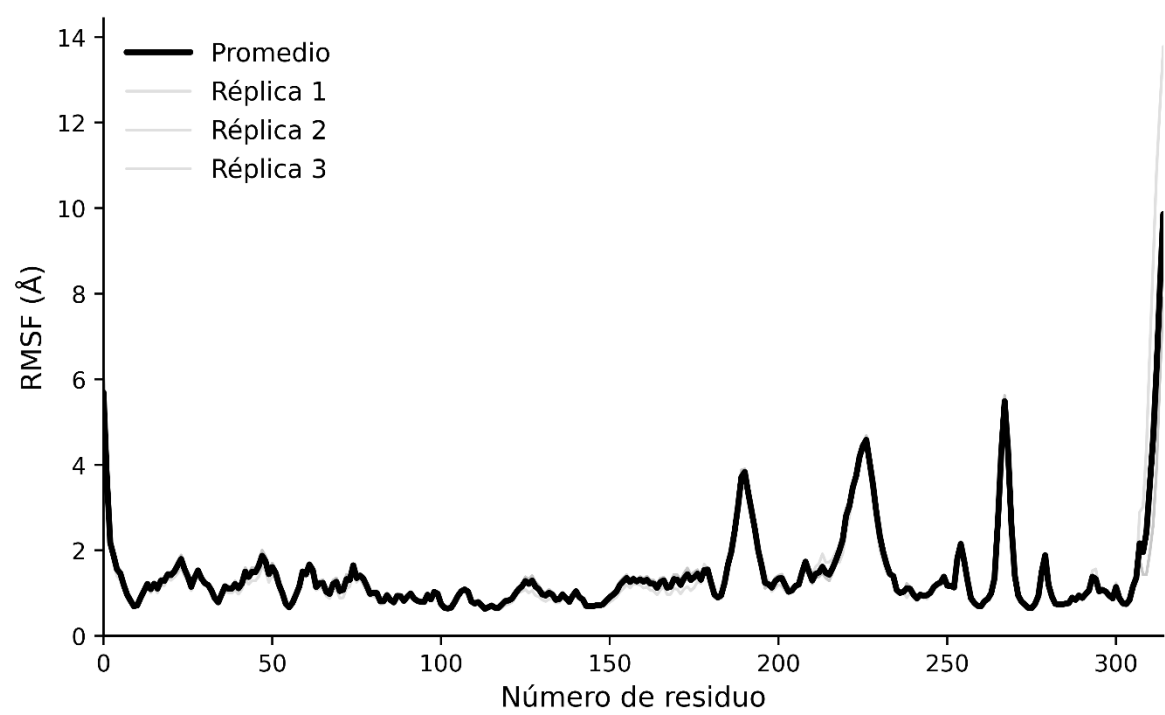


Figura 45 Fluctuación cuadrática media por residuo (RMSF) de la proteína durante la simulación de dinámica molecular del sistema APO. Las líneas grises corresponden a las tres réplicas independientes de la simulación, mientras que la línea negra representa el promedio de las trayectorias. El eje x indica el número de residuo, mientras que el eje y muestra el RMSF (Å), reflejando la flexibilidad local de cada región de la proteína a lo largo de la simulación.

La distancia catalítica entre Cys111 y His272 se mantuvo estable durante toda la trayectoria, lo que sugiere la conservación de la arquitectura del sitio catalítico. Por el contrario, las distancias Tyr268–Arg166

y Tyr268–Pro299, empleadas como indicadores del movimiento del loop BL2, mostraron variaciones continuas a lo largo de la simulación, reflejando la dinámica conformacional de esta región (Figura 46).

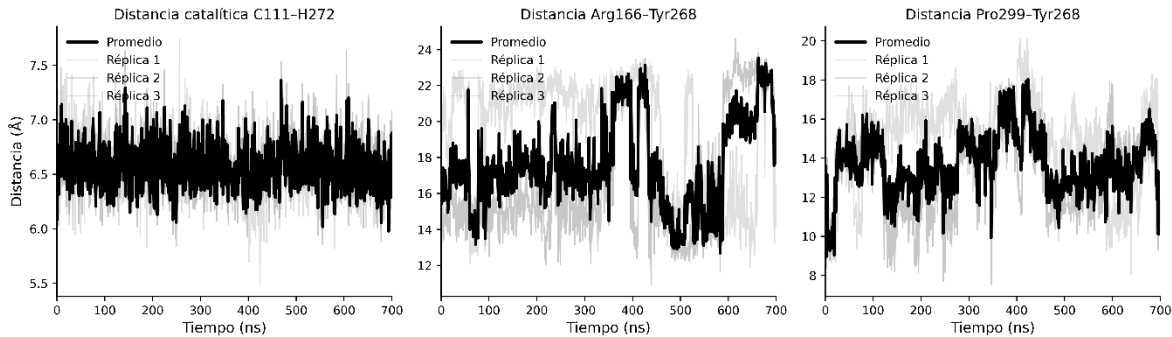


Figura 46 Evolución temporal de distancias estructurales clave durante la simulación de dinámica molecular del sistema APO. El panel izquierdo muestra la distancia catalítica entre Cys111 y His272, correspondiente a los residuos del sitio activo. Los paneles central y derecho muestran las distancias Arg166–Tyr268 y Pro299–Tyr268, respectivamente, utilizadas como descriptores del movimiento del loop BL2. Las líneas grises representan las tres réplicas independientes de la simulación, mientras que la línea negra corresponde al promedio de las trayectorias. El eje x indica el tiempo de simulación (ns) y el eje y la distancia entre los residuos (Å).

El análisis de los ángulos diedros de Tyr268 reveló la presencia de poblaciones conformacionales preferenciales para los ángulos χ_1 y χ_2 (Figura 47) indicando la adopción recurrente de orientaciones específicas de la cadena lateral durante la simulación.

Distribución angular Tyr268 - sistema APO

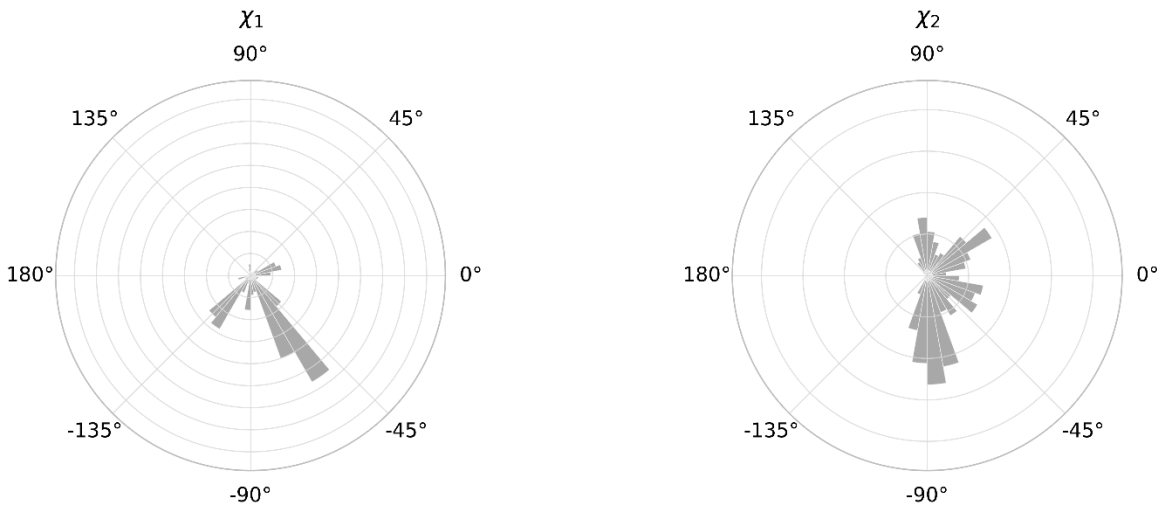


Figura 47 Distribución angular de los ángulos diedros χ_1 y χ_2 del residuo Tyr268 durante la simulación de dinámica molecular del sistema APO. Los diagramas polares muestran la frecuencia de aparición de cada conformación angular a lo largo de la trayectoria. Cada barra corresponde a un intervalo angular de 5°, mientras que los círculos concéntricos representan la frecuencia relativa, marcados en incrementos de 10 unidades.

5.5.2 Análisis GRL-0617 molécula de referencia

La simulación de Dinámica molecular del complejo proteína–GRL-0617 mostró una estabilización progresiva del RMSD de la proteína a lo largo de la trayectoria, alcanzando valores prácticamente constantes después de los primeros nanosegundos de simulación (Figura 48). Este comportamiento indica que la estructura global de la proteína se mantiene estable en presencia del ligando. Al comparar estos resultados con el RMSD de la proteína en ausencia de ligando (Figura 44), se observa que el sistema complejo presenta una menor fluctuación estructural, sugiriendo que la unión de GRL-0617 contribuye a

estabilizar la conformación de la proteína.

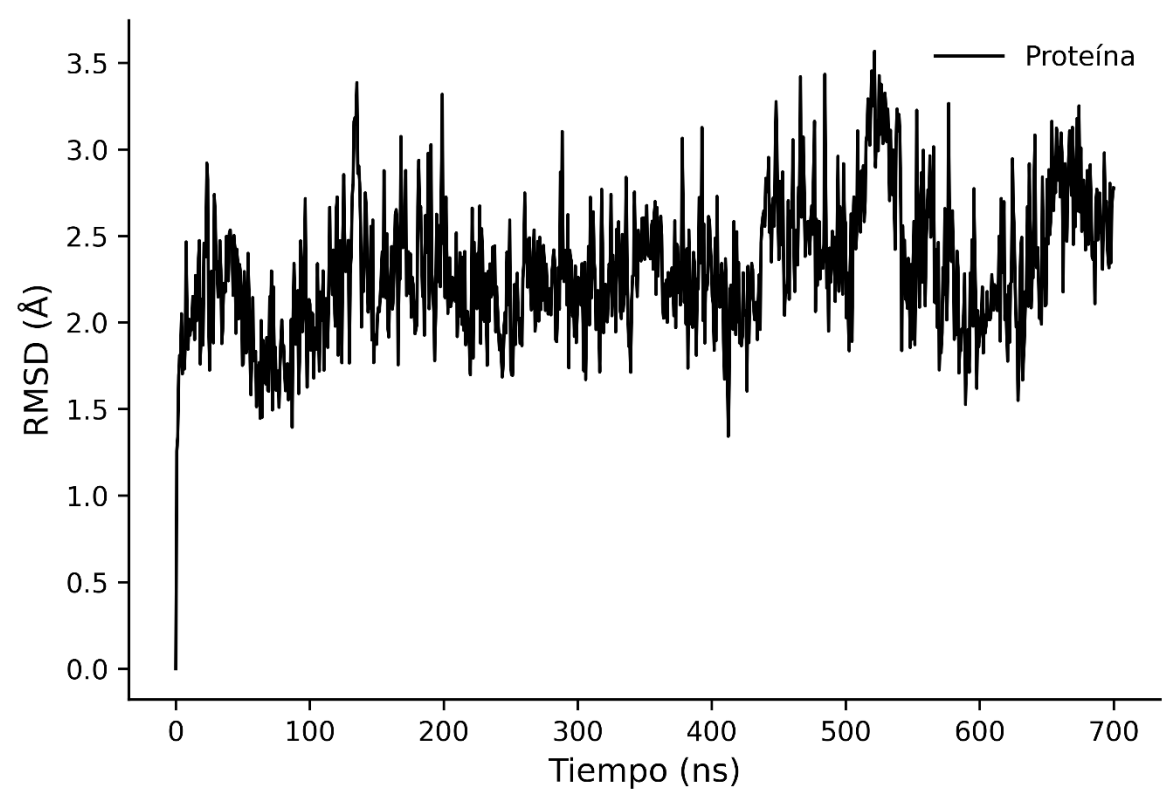


Figura 48 RMSD de la proteína con respecto al frame en tiempo 0 a lo largo del tiempo de simulación 700ns

De manera complementaria, el RMSD del ligando GRL-0617 se mantuvo dentro de un rango reducido durante toda la simulación (Figura 49), lo que sugiere que el compuesto permanece firmemente anclado en el sitio de unión sin experimentar desplazamientos significativos ni procesos de disociación.

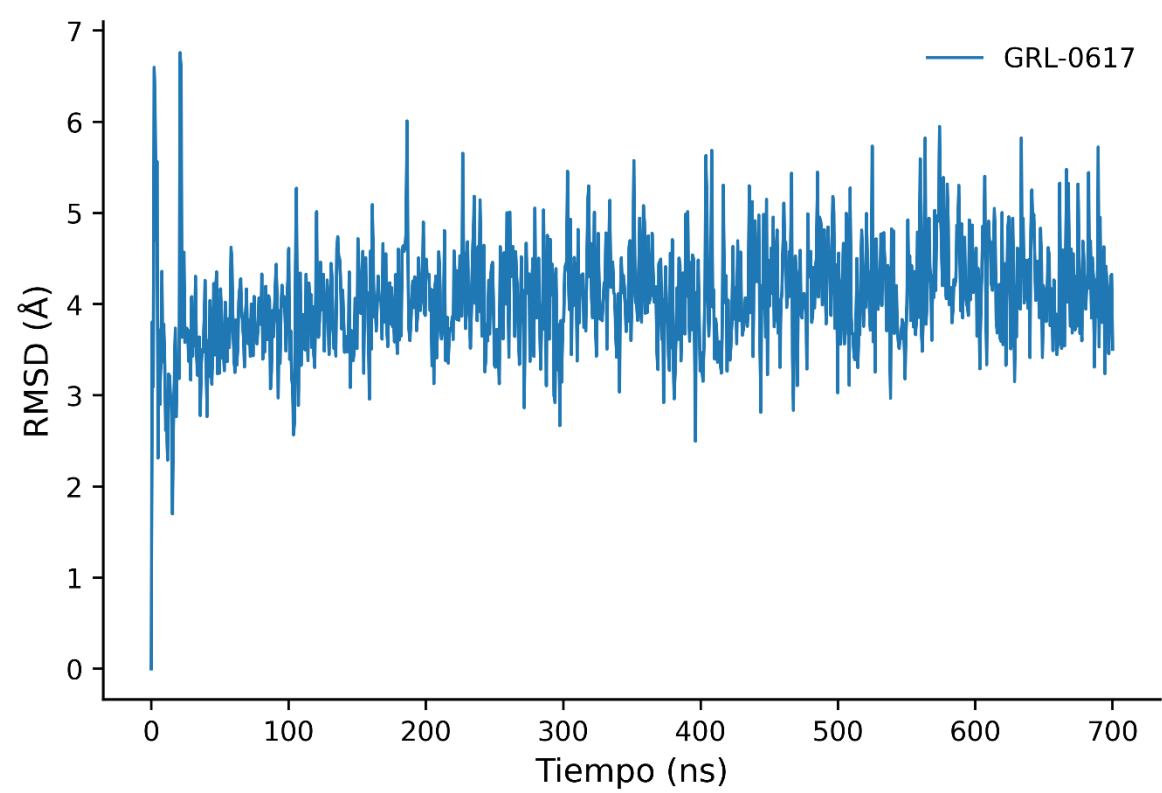


Figura 49 RMSD del ligando con respecto a la proteína a lo largo de la simulación 700ns.

El análisis de RMSF de la proteína en el complejo reveló una mayor movilidad en la región correspondiente al loop BL2, donde el residuo Tyr268 presentó la mayor fluctuación conformacional (Figura 50). Este

comportamiento es consistente con el papel dinámico de esta región en la regulación del acceso al sitio de unión.

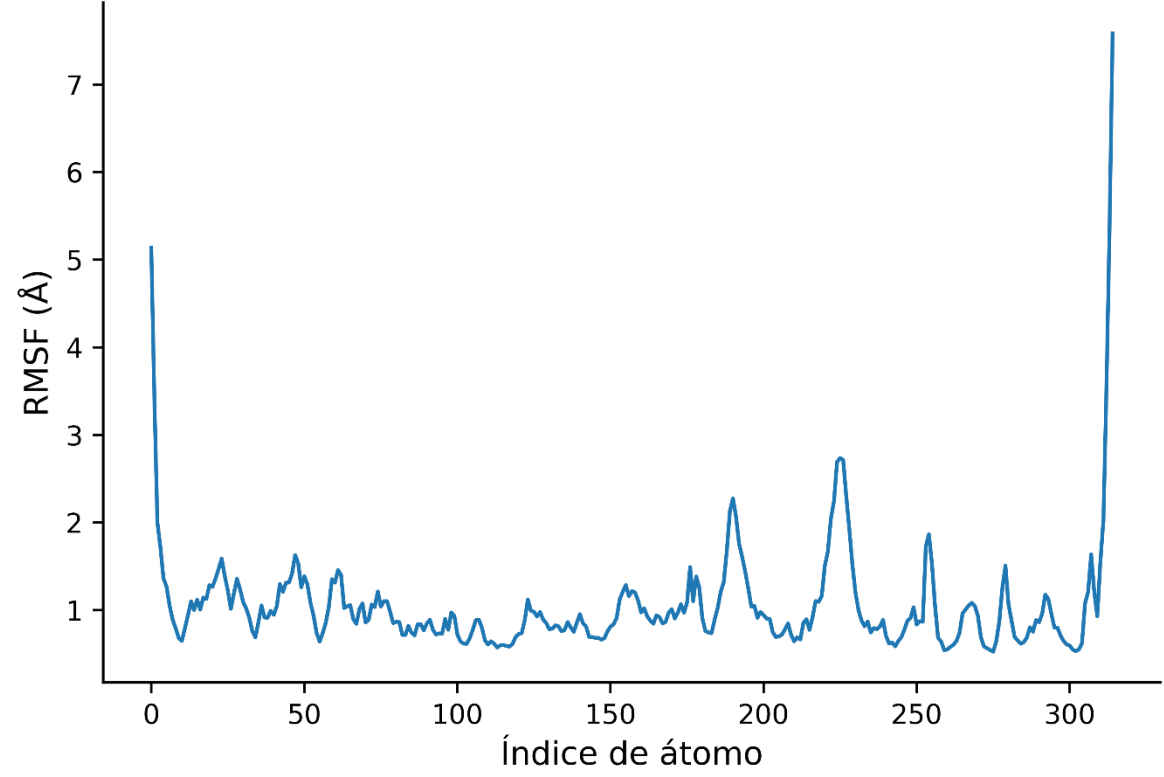


Figura 50 Fluctuación cuadrática media por residuo (RMSF) de la proteína con GRL-0617 durante la simulación. La línea azul representa las trayectorias. El eje x indica el número de residuo, mientras que el eje y muestra el RMSF (Å), reflejando la flexibilidad local de cada región de la proteína a lo largo de la simulación.

En la Figura 51 se compara el perfil de RMSF de la PLpro en estado apo con el obtenido para el complejo con el ligando GRL-0617. En esta comparación se observa una disminución de la flexibilidad en la región del loop BL2 en presencia del ligando, indicando una estabilización local de esta zona de la proteína.

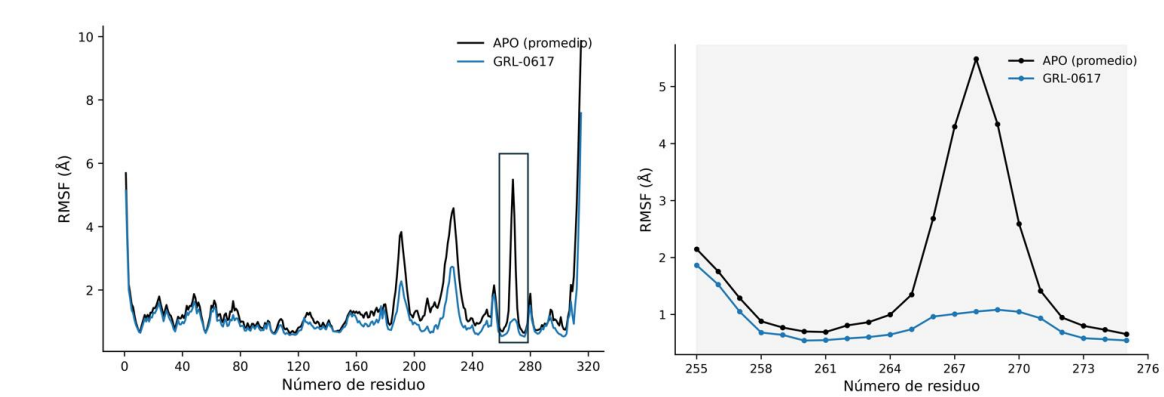


Figura 51 Comparación de la fluctuación cuadrática media por residuo (RMSF) entre el sistema APO y el complejo PLpro– GRL-0617 durante la simulación de dinámica molecular. El panel izquierdo muestra el perfil completo de RMSF a lo largo de la proteína, mientras que el panel derecho presenta una ampliación de la región correspondiente al loop BL2. La línea negra representa el sistema APO (promedio de las réplicas) y la línea azul el sistema GRL-0617. El eje x indica el número de residuo y el eje y el RMSF (Å), permitiendo comparar la flexibilidad local de la proteína en ausencia y presencia del ligando.

El análisis de las distancias estructurales clave mostró que la distancia catalítica entre Cys111 y His272 se mantuvo estable durante toda la simulación, indicando que la arquitectura del sitio catalítico se conserva en presencia del ligando. En contraste, las distancias Tyr268–Arg166 y Tyr268–Pro299

mostraron variaciones a lo largo del tiempo (Figura 52), reflejando un cambio en el movimiento dinámico del loop BL2.

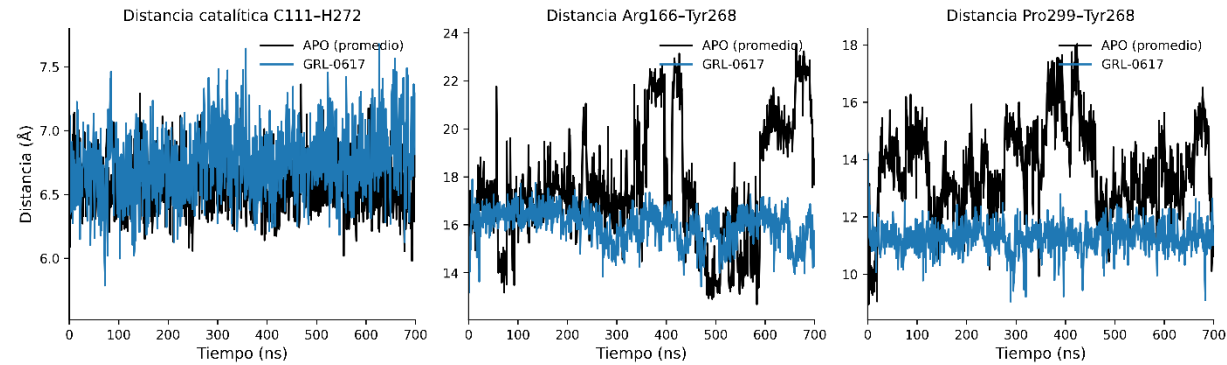


Figura 52 Evolución temporal de distancias estructurales clave durante las simulaciones de dinámica molecular para los sistemas APO y PLpro-GRL-0617. El panel izquierdo muestra la distancia catalítica entre Cys111 y His272. Los paneles central y derecho corresponden a las distancias Arg166-Tyr268 y Pro299-Tyr268, respectivamente, utilizadas como descriptores del movimiento del loop BL2. La línea negra representa el sistema APO (promedio de las réplicas) y la línea azul el complejo GRL-0617. El eje x indica el tiempo de simulación (ns) y el eje y la distancia entre residuos (Å).

Finalmente, el análisis de los ángulos diedros de Tyr268 mostró poblaciones conformacionales preferenciales para los ángulos χ_1 y χ_2 (Figura 47), indicando que la cadena lateral del residuo adopta orientaciones recurrentes durante la simulación del complejo proteína-ligando.

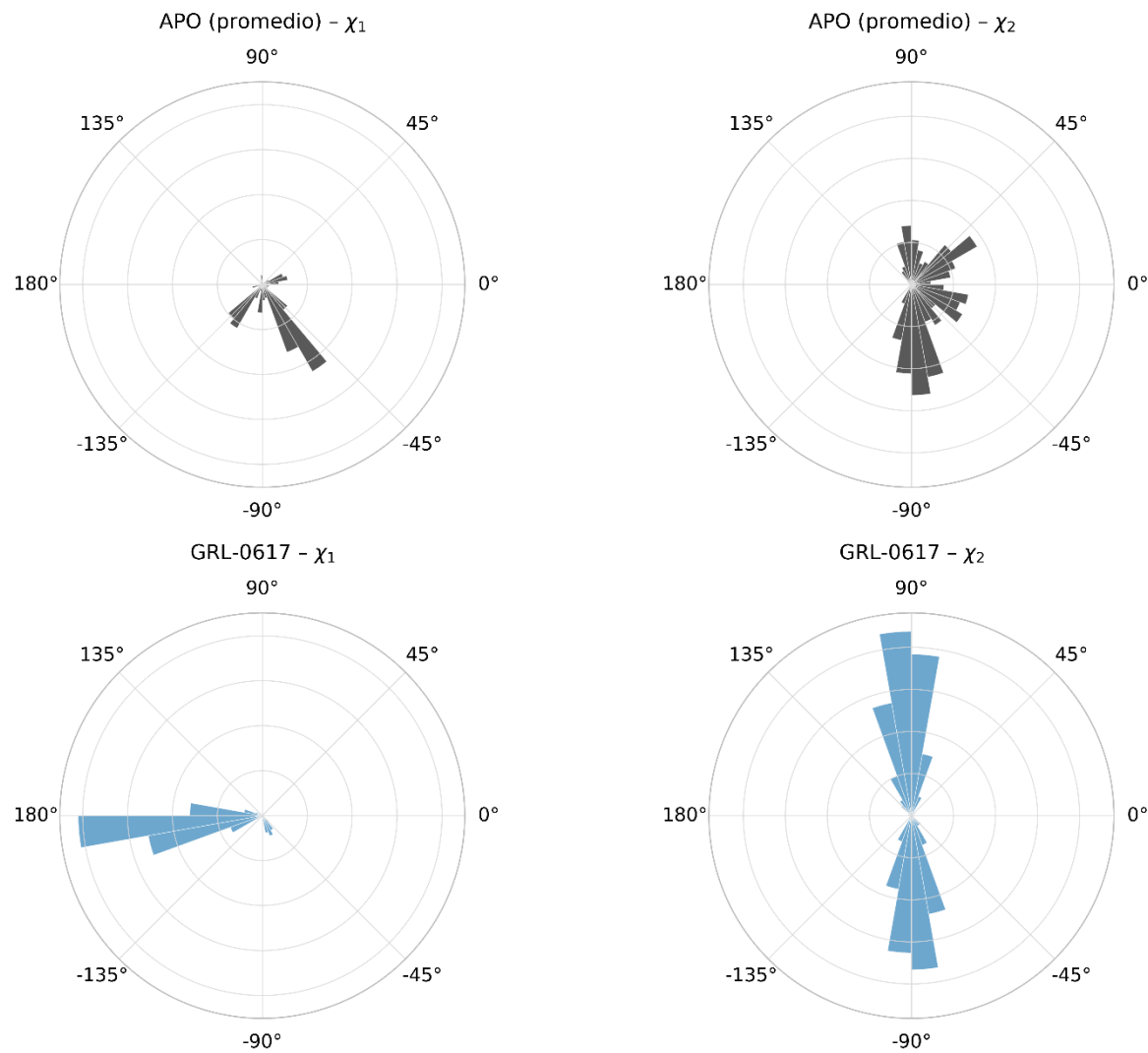


Figura 53 Distribución angular de los ángulos diedros χ_1 y χ_2 del residuo Tyr268 durante las simulaciones de dinámica molecular de los sistemas APO y PLpro-GRL-0617. Los diagramas polares muestran la frecuencia de las conformaciones angulares observadas a lo largo de la trayectoria. Los paneles superiores corresponden al sistema APO (promedio de las réplicas) y los inferiores al complejo GRL-0617. Cada barra representa un intervalo angular de 5°, mientras que los círculos concéntricos indican la frecuencia relativa de ocurrencia.

5.5.3 Análisis de la familia pirimido-benzimidazol

Dentro de las familias benceno, 1-naftilo, 2-naftilo, antraceno y pireno, se observa que el comportamiento dinámico de la proteína en presencia de la mayoría de los ligandos es similar al observado para el inhibidor

de referencia GRL-0617 (Figura 54-Figura 58). En particular, el sustituyente 4-NMe₂ muestra el comportamiento más consistentemente cercano al del inhibidor de referencia, manteniendo valores de RMSD contenidos a lo largo de todas las series analizadas. Un comportamiento igualmente similar al control se observa para 4-OH-3-OMe, especialmente en las familias benceno, 2-naftilo, antraceno y pireno, aunque en la familia 1-naftilo este derivado presenta una desviación transitoria más marcada durante la simulación.

En contraste, los mayores incrementos del RMSD de la proteína se detectan para el sustituyente 2-OH, particularmente en la serie 1-naftilo, donde la trayectoria alcanza valores cercanos o incluso superiores a los observados para la proteína en estado apo. Este mismo sustituyente también genera aumentos notorios del RMSD en las familias benceno y pireno, mientras que en la familia 2-naftilo permanece relativamente estable.

Por su parte, el derivado 4-NO₂ muestra en general un comportamiento similar al del inhibidor de referencia, a excepción de la familia 1-naftilo, donde se observa un incremento más evidente del RMSD de la proteína.

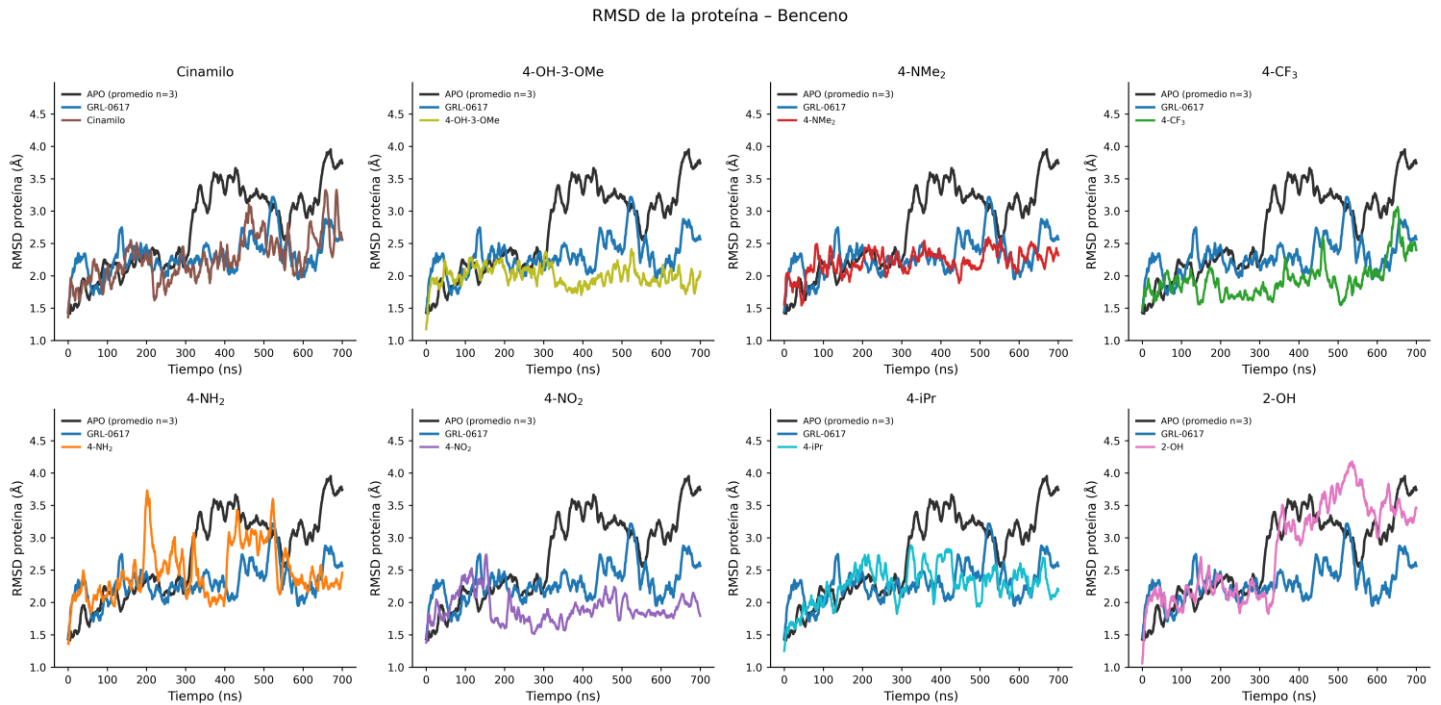


Figura 54 Evolución temporal del RMSD de la proteína durante las simulaciones de dinámica molecular para los sistemas APO, PLpro–GRL-0617 y los distintos derivados bencénicos evaluados. Cada panel corresponde a un ligando diferente (cinamilo, 4-OH-3-OMe, 4-NMe₂, 4-CF₃, 4-NH₂, 4-NO₂, 4-iPr y 2-OH). La línea negra representa el sistema APO (promedio de las réplicas), la línea azul el complejo PLpro–GRL-0617, y la línea de color corresponde al sistema con el ligando evaluado. El eje x indica el tiempo de simulación (ns) y el eje y el RMSD de la proteína (Å) respecto a la estructura inicial.

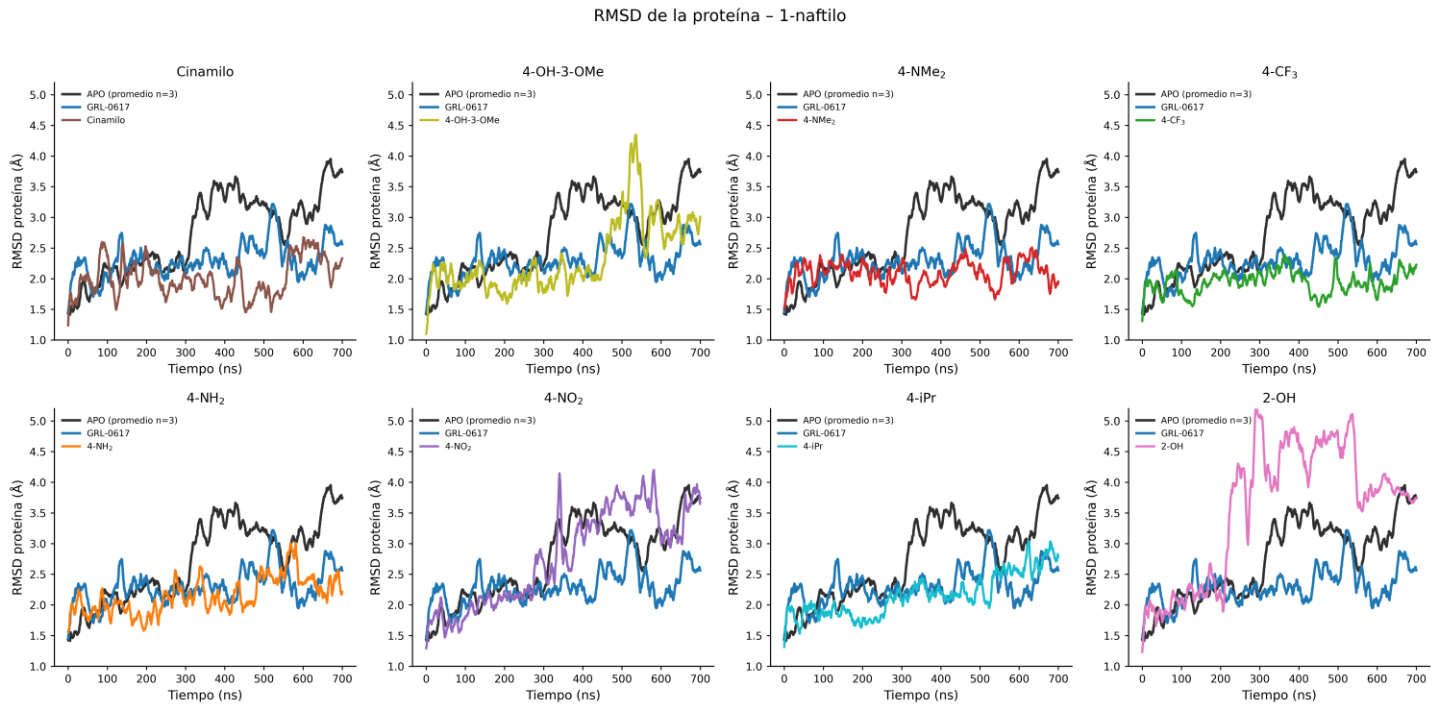


Figura 55 Evolución temporal del RMSD de la proteína durante las simulaciones de dinámica molecular para los sistemas APO, PLpro–GRL-0617 y los derivados 1-naftílicos evaluados.

RMSD de la proteína - 2-naftilo

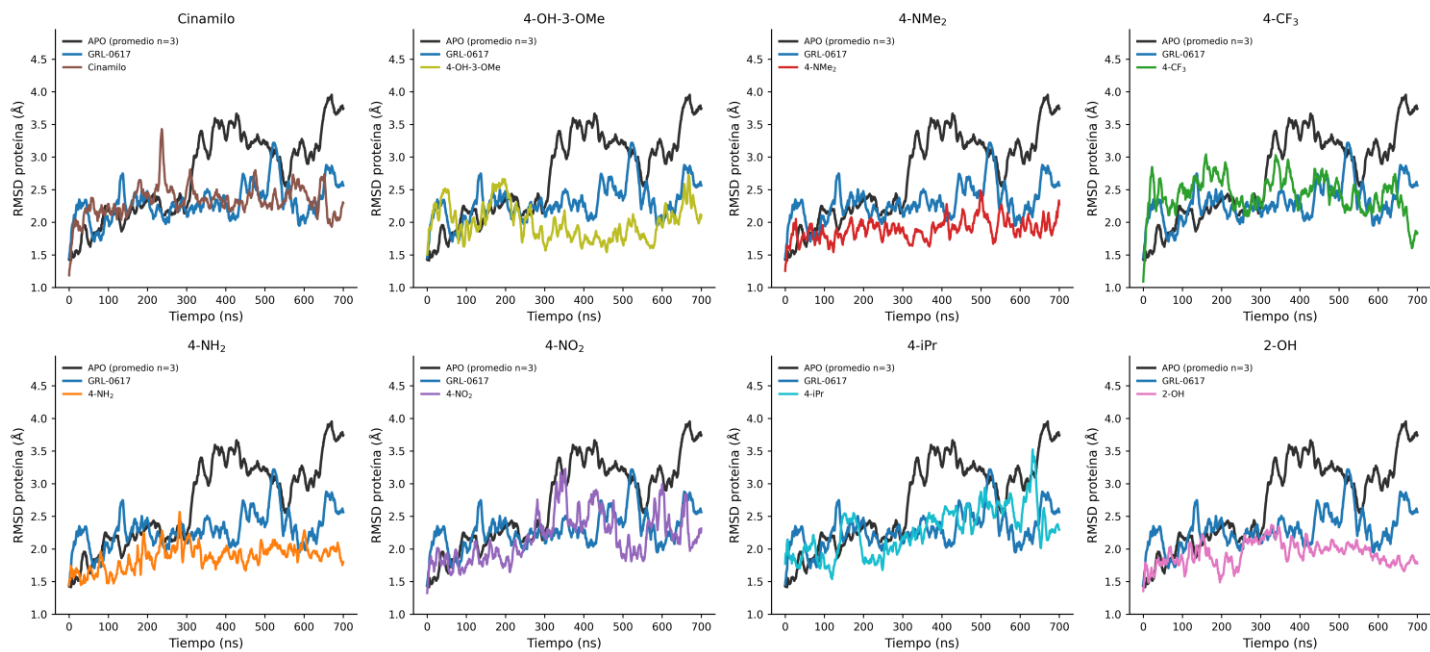


Figura 56 Evolución temporal del RMSD de la proteína durante las simulaciones de dinámica molecular para los sistemas APO, PLpro-GRL-0617 y los derivados 2-naftílicos evaluados.

RMSD de la proteína - Antraceno

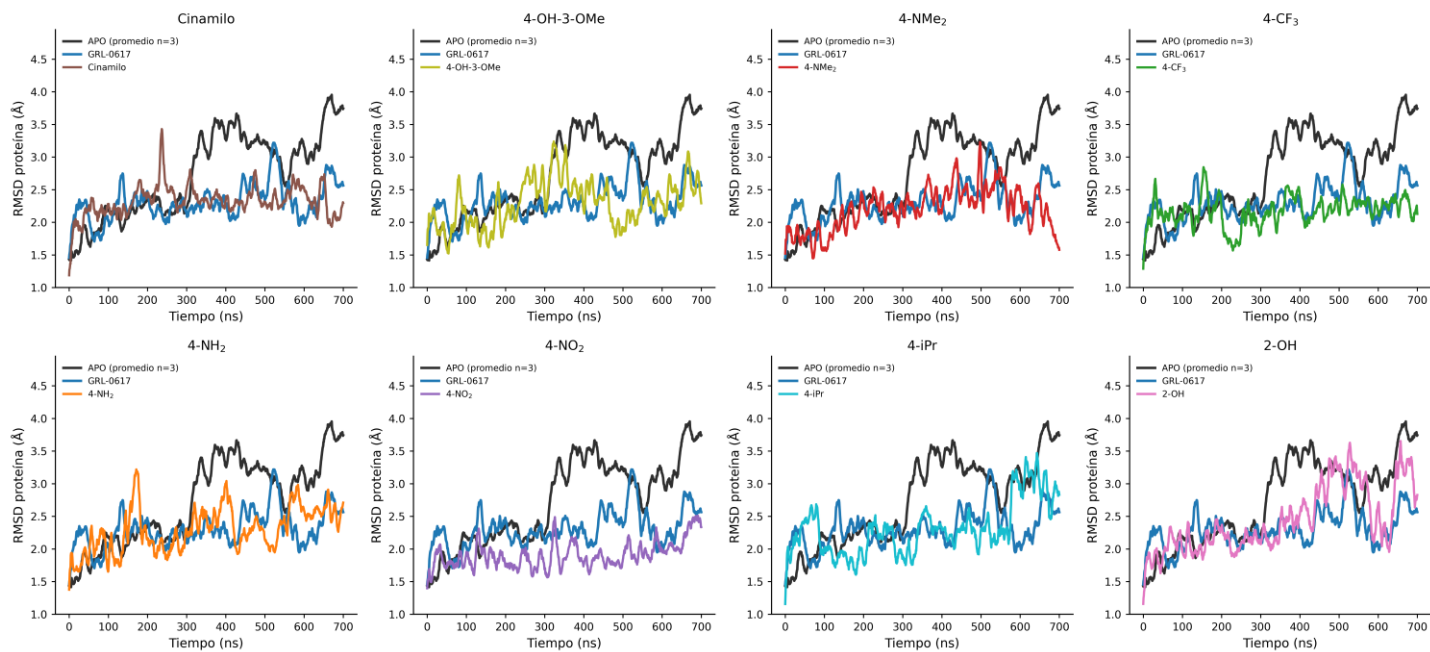


Figura 57 Evolución temporal del RMSD de la proteína durante las simulaciones de dinámica molecular para los sistemas APO, PLpro-GRL-0617 y los derivados de antraceno evaluados.

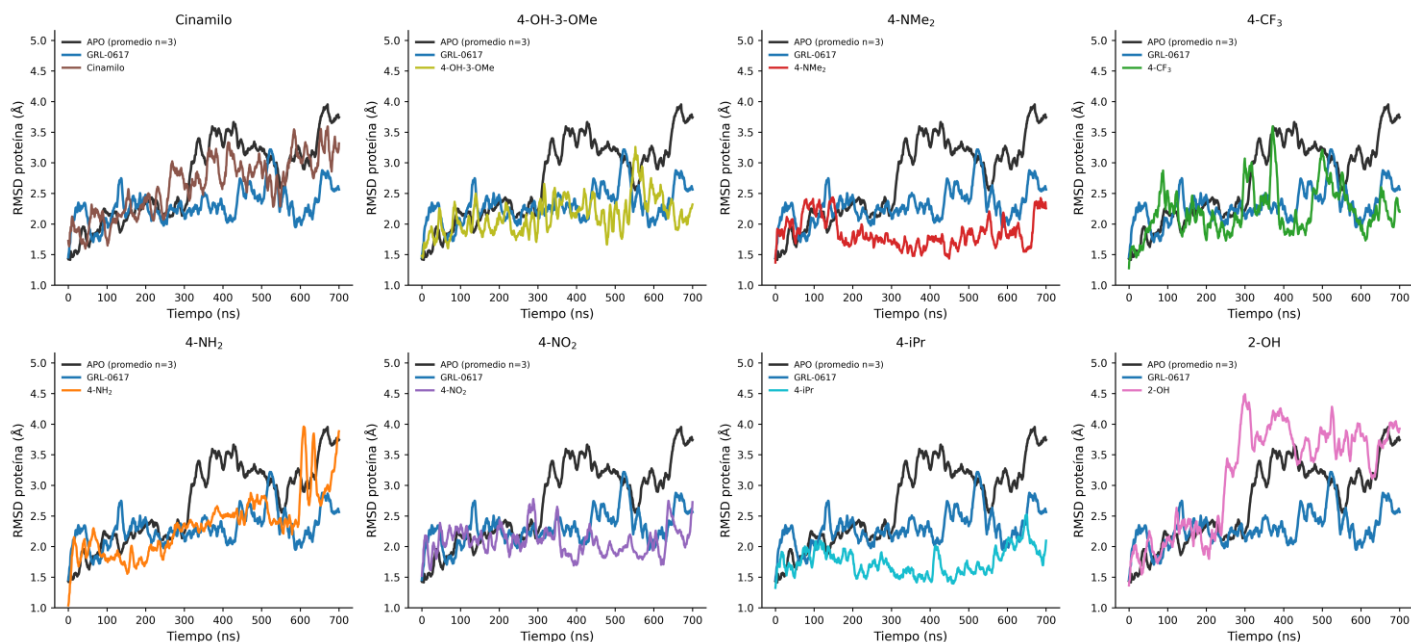


Figura 58 Evolución temporal del RMSD de la proteína durante las simulaciones de dinámica molecular para los sistemas APO, PLpro–GRL-0617 y los derivados de pireno evaluados.

El análisis del RMSD del ligando permite evaluar la estabilidad de cada compuesto dentro del sitio activo a lo largo de la simulación (Figura 59–Figura 63). Como referencia, el inhibidor cristalográfico GRL-0617 mantiene un RMSD cercano a ~ 4 Å.

En la familia benceno, ninguno de los derivados evaluados muestra un comportamiento comparable al del inhibidor de referencia. La mayoría de los ligandos presenta incrementos pronunciados del RMSD, alcanzando valores superiores a 20–40 Å, lo que indica que abandonan la posición inicial dentro del sitio activo durante la simulación. Algunos derivados muestran mesetas relativamente definidas tras el aumento inicial del RMSD, lo que sugiere que se estabilizan en otras regiones de la proteína. Este comportamiento se observa, por ejemplo, en 4-OH-3-OMe, 4-CF₃ y 4-iPr, cuyos valores de RMSD se mantienen relativamente constantes después de un desplazamiento inicial. Otros ligandos presentan desviaciones aún más marcadas; el derivado 4-NMe₂ muestra fluctuaciones amplias con incrementos que superan los 50 Å, mientras que cinamilo y 4-NH₂ exhiben incrementos progresivos del RMSD a lo largo de la simulación. En conjunto, estos resultados indican que los derivados de esta familia no permanecen en el sitio activo y probablemente exploran otras regiones de la proteína.

El comportamiento cambia al analizar los derivados 1-naftilo. En esta serie se observan varios ligandos que, aunque abandonan el sitio activo, muestran trayectorias con mesetas relativamente definidas en valores superiores a 10 Å. Este patrón se aprecia particularmente en cinamilo y 4-NO₂, cuyos RMSD se estabilizan en torno a ~ 15 –20 Å durante gran parte de la simulación, lo que sugiere una posible estabilización en cavidades cercanas o regiones alostéricas de la proteína. En contraste, el sustituyente 2-OH presenta un comportamiento mucho más cercano al del inhibidor de referencia, manteniendo valores relativamente bajos de RMSD durante la simulación. Los demás derivados, incluyendo 4-NMe₂, 4-CF₃, 4-NH₂ y 4-iPr, muestran incrementos más pronunciados del RMSD, lo que indica desplazamientos importantes respecto a la posición inicial del ligando.

En la serie 2-naftilo se observan nuevamente desplazamientos importantes para varios compuestos. Los derivados cinamilo, 4-NH₂ y 2-OH presentan incrementos significativos del RMSD que alcanzan valores cercanos a 20–45 Å, lo que sugiere que abandonan el sitio activo y se estabilizan en otras regiones de la proteína. El derivado 4-CF₃ también muestra un desplazamiento temprano seguido de una estabilización en valores cercanos a ~ 17 Å. Sin embargo, en esta familia aparece uno de los comportamientos más similares al inhibidor de referencia: el derivado 4-OH-3-OMe, cuyo RMSD permanece cercano a 3–4 Å durante toda la trayectoria. Los sustituyentes 4-NMe₂ y 4-iPr muestran valores intermedios de RMSD, generalmente entre 7 y 12 Å, lo que sugiere desplazamientos moderados dentro de la proteína.

Los derivados de la familia antraceno presentan comportamientos más heterogéneos. Algunos ligandos muestran incrementos importantes del RMSD desde etapas tempranas de la simulación, como ocurre con cinamilo y 4-NO₂, cuyos valores alcanzan aproximadamente 35–40 Å, indicando desplazamientos significativos del sitio inicial. Otros derivados presentan trayectorias más contenidas, con estabilización

moderada en valores cercanos a 8–10 Å, como se observa para 4- NH₂, 4-OH-3-OMe y 4-CF₃, lo que sugiere una posible permanencia en cavidades cercanas al sitio activo. En contraste, los ligandos 4-iPr y 2-OH muestran valores de RMSD más cercanos al inhibidor de referencia, manteniéndose alrededor de 4–6 Å durante gran parte de la trayectoria. Finalmente, en la familia pireno se observa un comportamiento más irregular para la mayoría de los compuestos. Varios derivados presentan incrementos tempranos del RMSD seguidos de fluctuaciones amplias, indicando desplazamientos importantes del sitio inicial de unión. Este comportamiento se observa, por ejemplo, en cinamilo y 4-NMe₂, cuyos valores alcanzan aproximadamente 15–20 Å. Algunos ligandos muestran trayectorias relativamente más contenidas, como 4-CF₃ y 4-iPr, cuyos RMSD permanecen entre 7 y 12 Å durante gran parte de la simulación. Dentro de esta serie, el derivado 4-OH-3-OMe presenta el comportamiento más cercano al inhibidor de referencia, aunque aún con desviaciones mayores que las observadas para GRL-0617.

Los resultados muestran que la mayoría de los ligandos presentan desplazamientos importantes respecto a la posición inicial de unión, mientras que solo algunos derivados específicos muestran valores de RMSD comparables al inhibidor de referencia, sugiriendo una mayor estabilidad dentro del bolsillo proteico.

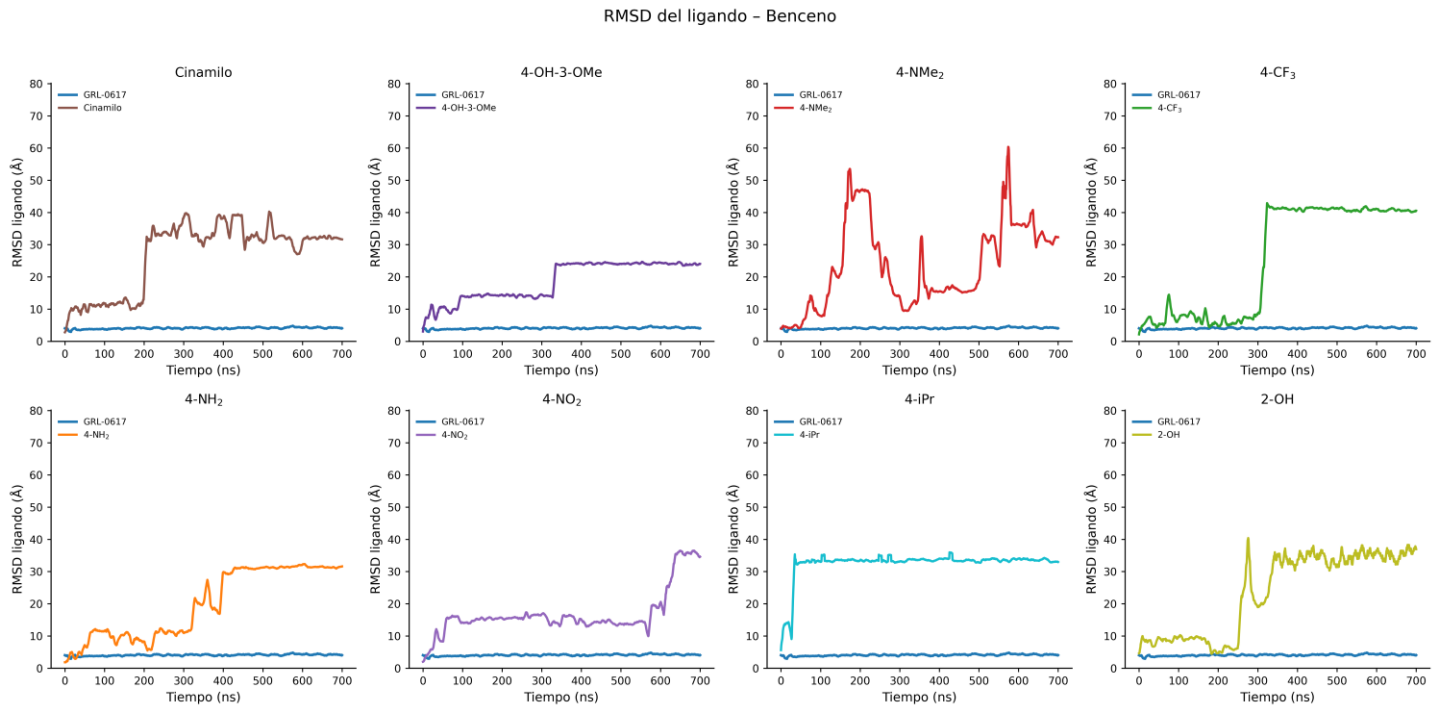


Figura 59 Evolución temporal del RMSD del ligando durante las simulaciones de dinámica molecular para los distintos derivados bencénicos evaluados en el sitio de unión de PLpro. Cada panel corresponde a un ligando diferente (cinamilo, 4-OH-3-OMe, 4-NMe₂, 4-CF₃, 4-NH₂, 4-NO₂, 4-iPr y 2-OH). La línea azul representa el ligando GRL-0617 utilizado como referencia, mientras que la línea de color corresponde al ligando evaluado.

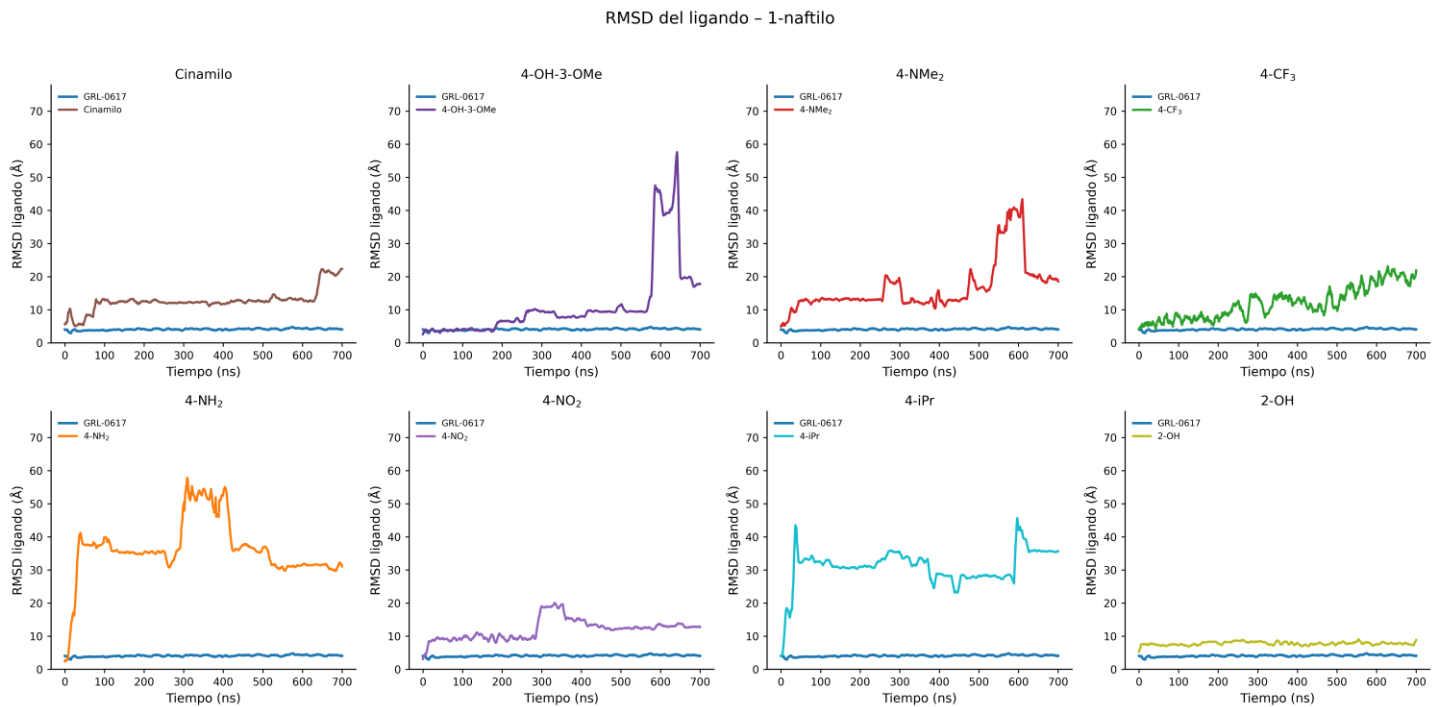


Figura 60 RMSD del ligando durante las simulaciones de dinámica molecular para los derivados 1 naftílicos evaluados.

RMSD del ligando - 2-naftilo

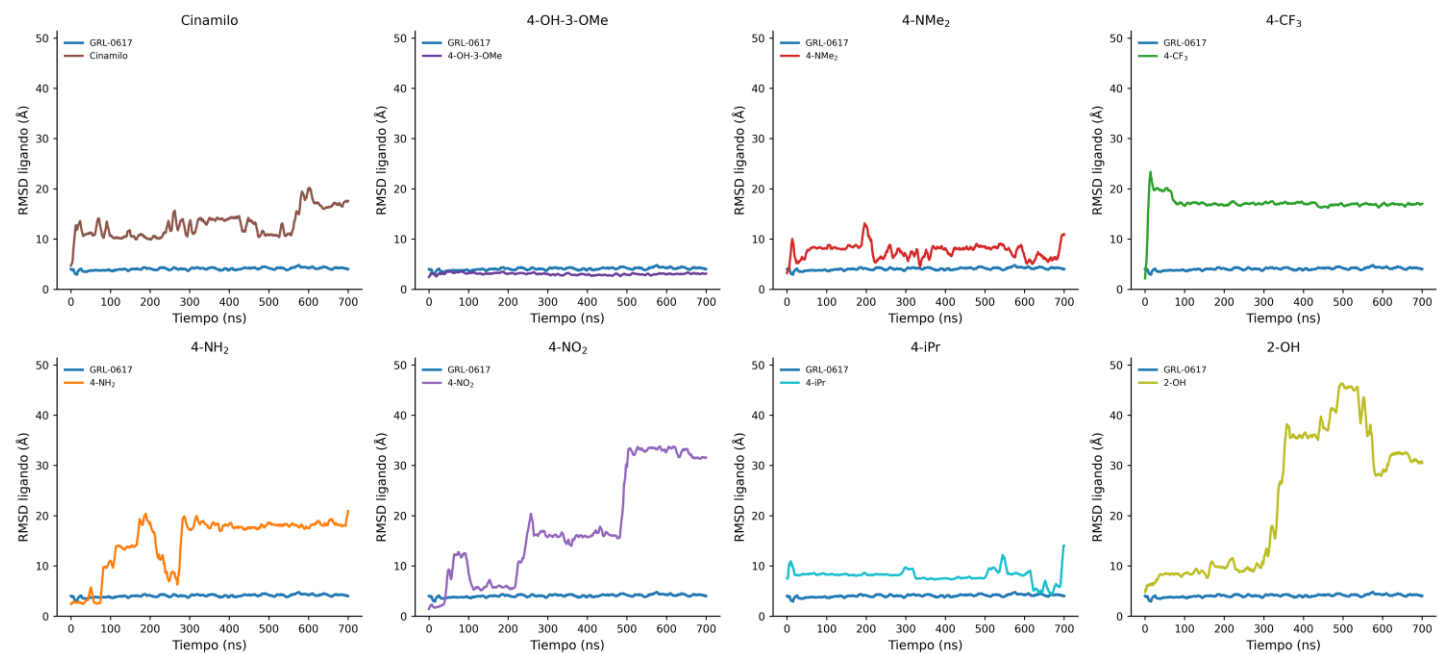


Figura 61 RMSD del ligando durante las simulaciones de dinámica molecular para los derivados 2 naftílicos evaluados.

RMSD del ligando - Antraceno

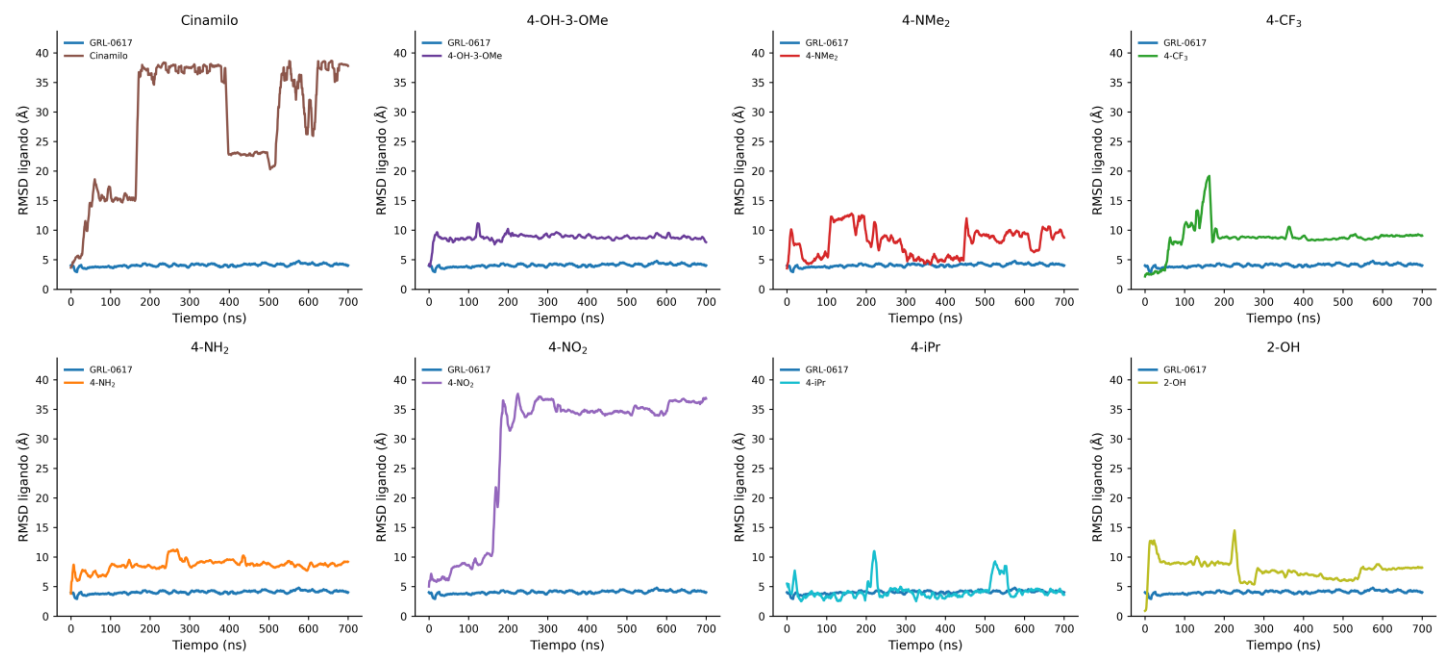


Figura 62 RMSD del ligando durante las simulaciones de dinámica molecular para los derivados de antraceno evaluados.

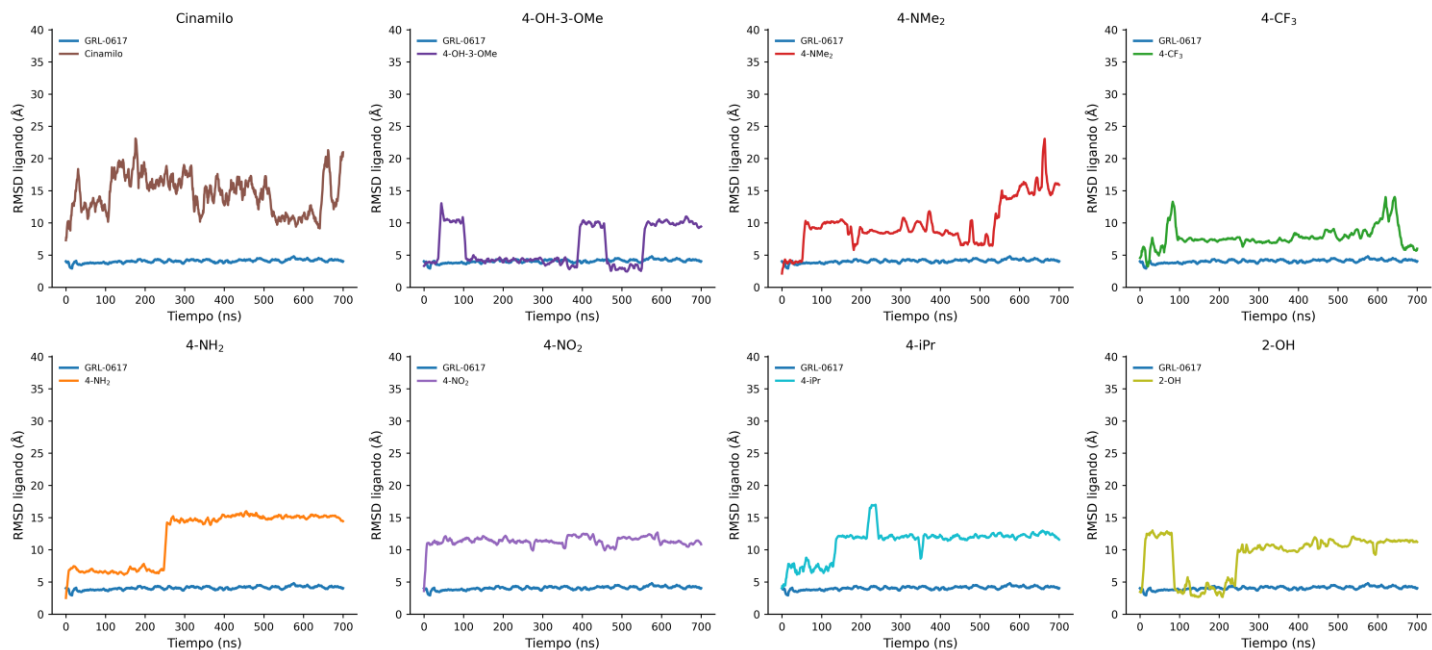


Figura 63 RMSD del ligando durante las simulaciones de dinámica molecular para los derivados de pireno evaluados.

El análisis de las fluctuaciones residuo a residuo (RMSF) permitió evaluar el impacto de los distintos ligandos sobre la dinámica del loop BL2 (residuos 255–275), una región clave en la regulación conformacional del sitio activo de la proteína. En el sistema apo, esta región presenta las mayores fluctuaciones estructurales, con un máximo centrado aproximadamente en el residuo 267, lo que refleja la naturaleza altamente flexible del loop en ausencia de ligando. En contraste, el inhibidor de referencia GRL-0617 reduce significativamente estas fluctuaciones, indicando una clara estabilización de esta región estructural. Los perfiles completos de RMSF para cada sistema se muestran en las Figura 64-Figura 68, mientras que un acercamiento específico a la región BL2 se presenta en las Figura 69-Figura 73.

En la familia de los bencenos, la mayoría de los derivados produce un efecto intermedio de rigidización del loop BL2 en comparación con el estado apo. En general, todos los sustituyentes reducen parcialmente la flexibilidad del loop, aunque sin alcanzar el nivel de estabilización observado para el inhibidor de referencia. Entre ellos, el derivado 4-CF₃ destaca por inducir la mayor rigidización dentro de esta familia, mostrando una disminución más marcada en las fluctuaciones del BL2. Otros sustituyentes, como 4-OH-3-OMe, 4-NMe₂ y 4-iPr, presentan un comportamiento intermedio, reduciendo la movilidad del loop de forma moderada.

En la serie 1-naftilo se observa nuevamente un comportamiento predominantemente intermedio. Sin embargo, algunos compuestos muestran una capacidad mayor de estabilización del loop BL2. En particular, el derivado 2-OH presenta una de las reducciones más notables de la flexibilidad en esta región. En contraste, el derivado 4-NO₂ muestra un perfil de fluctuación más similar al observado en la proteína en estado apo, lo que sugiere una menor capacidad para rigidizar el loop BL2. Por su parte, el sustituyente 4-OH-3-OMe mantiene un comportamiento relativamente estable, con fluctuaciones moderadas dentro de esta región.

Para la familia 2-naftilo, se observa una modulación más clara de la dinámica del loop BL2. Dentro de esta serie, los derivados que inducen la mayor rigidización corresponden a 4-OH-3-OMe (derivado tipo vainilina) y 4-iPr, los cuales muestran una reducción más marcada en las fluctuaciones del centro del BL2. Otros compuestos presentan un efecto estabilizador intermedio, aunque sin alcanzar el nivel observado para estos dos sustituyentes. En contraste, los derivados 4-CF₃ y 4-NO₂ muestran perfiles de fluctuación más elevados en esta región, con valores que se aproximan al comportamiento observado para la proteína en estado apo, lo que sugiere una menor capacidad para restringir la movilidad del loop BL2.

Para la familia antraceno, la mayoría de los compuestos tiende a rigidizar el loop BL2 de forma más clara que en las familias anteriores. Varios derivados muestran una disminución significativa de las fluctuaciones en esta región, indicando una mayor restricción conformacional del loop. No obstante, una excepción notable corresponde al derivado 4-NO₂, el cual incrementa la flexibilidad del BL2 incluso por

encima del comportamiento observado en el estado apo, sugiriendo que este sustituyente no favorece la estabilización de esta región estructural.

Finalmente, en la serie pireno se observa un comportamiento más limitado en cuanto a la rigidización del loop BL2. En general, la mayoría de los compuestos mantiene fluctuaciones relativamente cercanas al estado apo, indicando un efecto menor sobre la dinámica de esta región. Dentro de esta familia, el único derivado que muestra una reducción clara de la movilidad del BL2 es 4-NMe₂, cuyo perfil de RMSF evidencia una estabilización más marcada del loop en comparación con los demás compuestos.

El RMSF destaca distintos ligandos modula la dinámica del loop BL2 en distintos grados dependiendo tanto del núcleo aromático como del sustituyente presente. Algunos compuestos como 4-CF₃ en la serie benceno, 4-OH-3-OMe y 4-iPr en la familia 2-naftilo, y 4-NMe₂ en la serie pireno, destacaron por disminuir la rigidez de la zona.

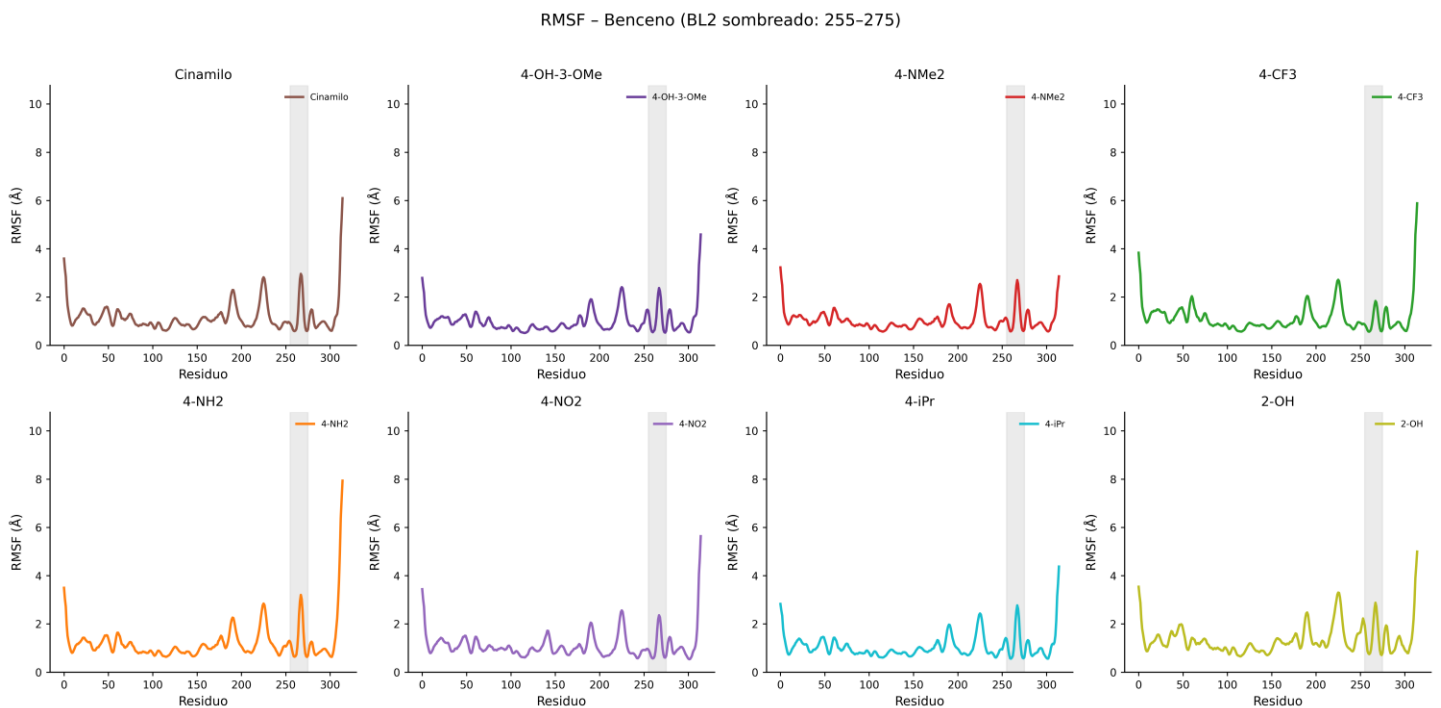


Figura 64 RMSF por residuo de la proteína durante las simulaciones de Dinámica molecular de los complejos PLpro–derivados bencénicos. Cada panel muestra el perfil de flexibilidad de la proteína para un ligando diferente. La región sombreada indica el loop BL2 (residuos 255–275), asociado con el control del acceso al bolsillo de unión.

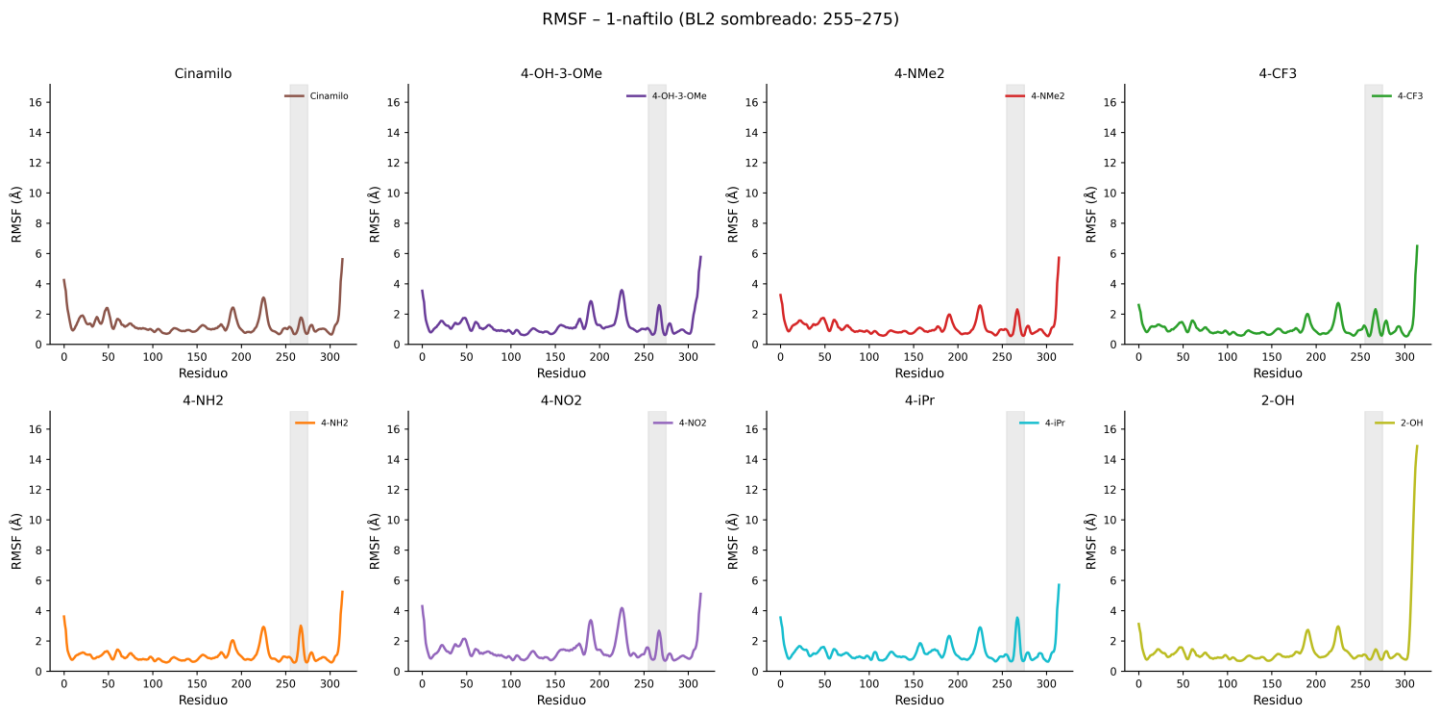


Figura 65 RMSF por residuo de la proteína para los derivados 1-naftílicos evaluados. La región correspondiente al loop BL2 (255–275).

RMSF - 2-naftilo (BL2 sombreado: 255-275)

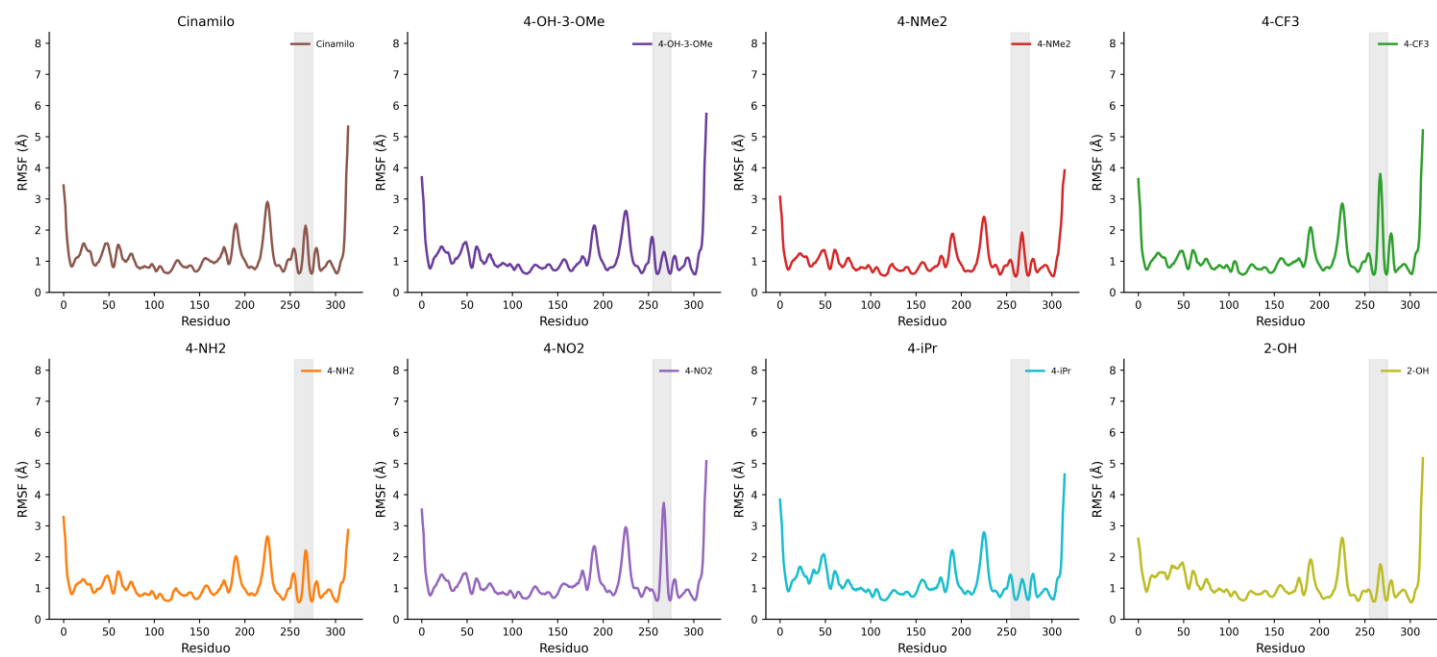


Figura 66 RMSF por residuo de la proteína para los derivados 2-naftílicos evaluados. La región correspondiente al loop BL2 (255–275).

RMSF - Antraceno (BL2 sombreado: 255-275)

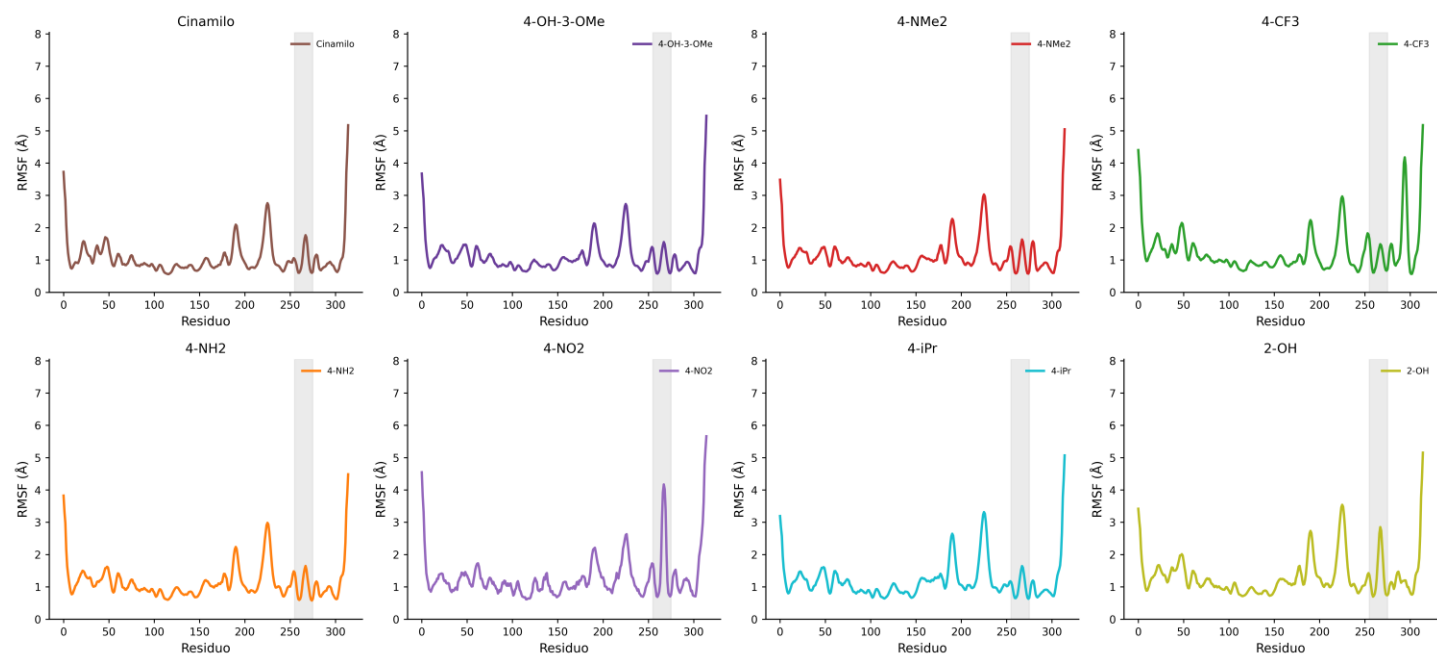


Figura 67 RMSF por residuo de la proteína para los derivados de antraceno evaluados. La región correspondiente al loop BL2 (255–275).

RMSF - Pireno (BL2 sombreado: 255-275)

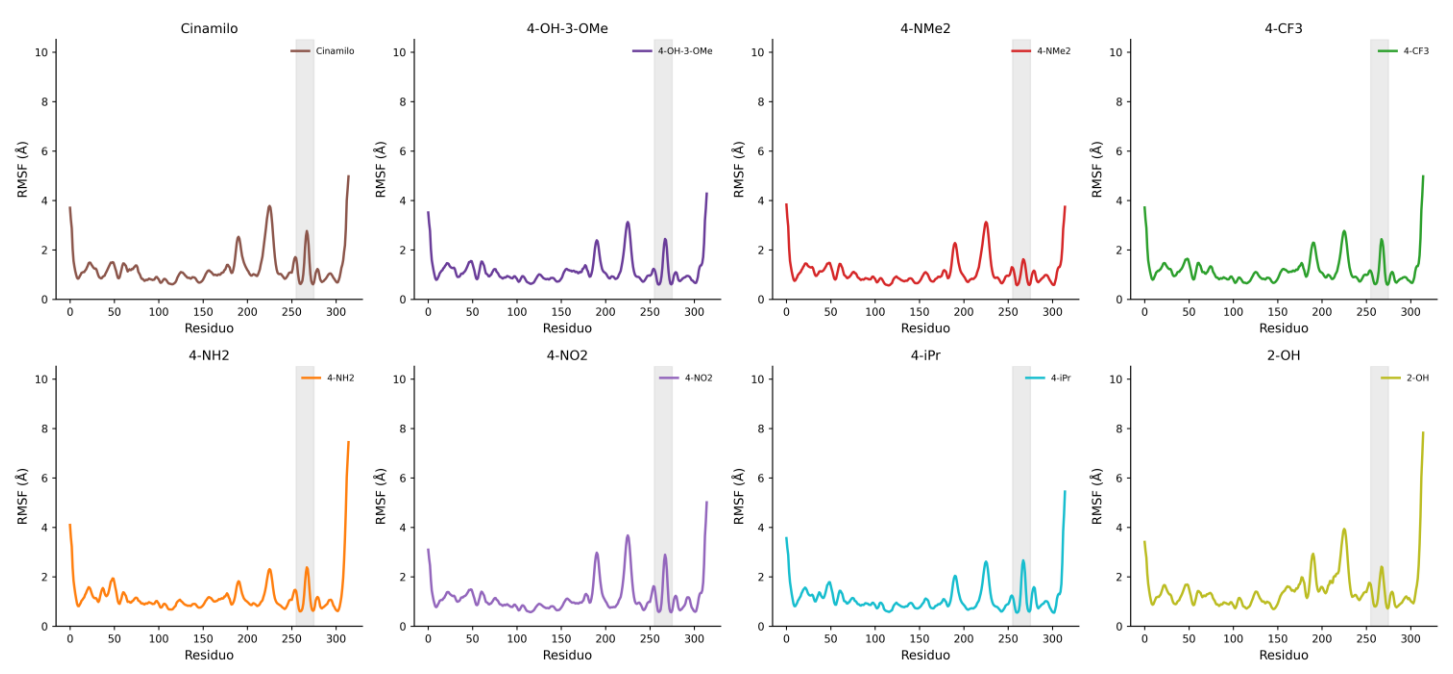


Figura 68 RMSF por residuo de la proteína para los derivados de pireno evaluados. La región correspondiente al loop BL2 (255-275).

Zoom del RMSF en la región BL2 - Benceno (residuos 255-275)

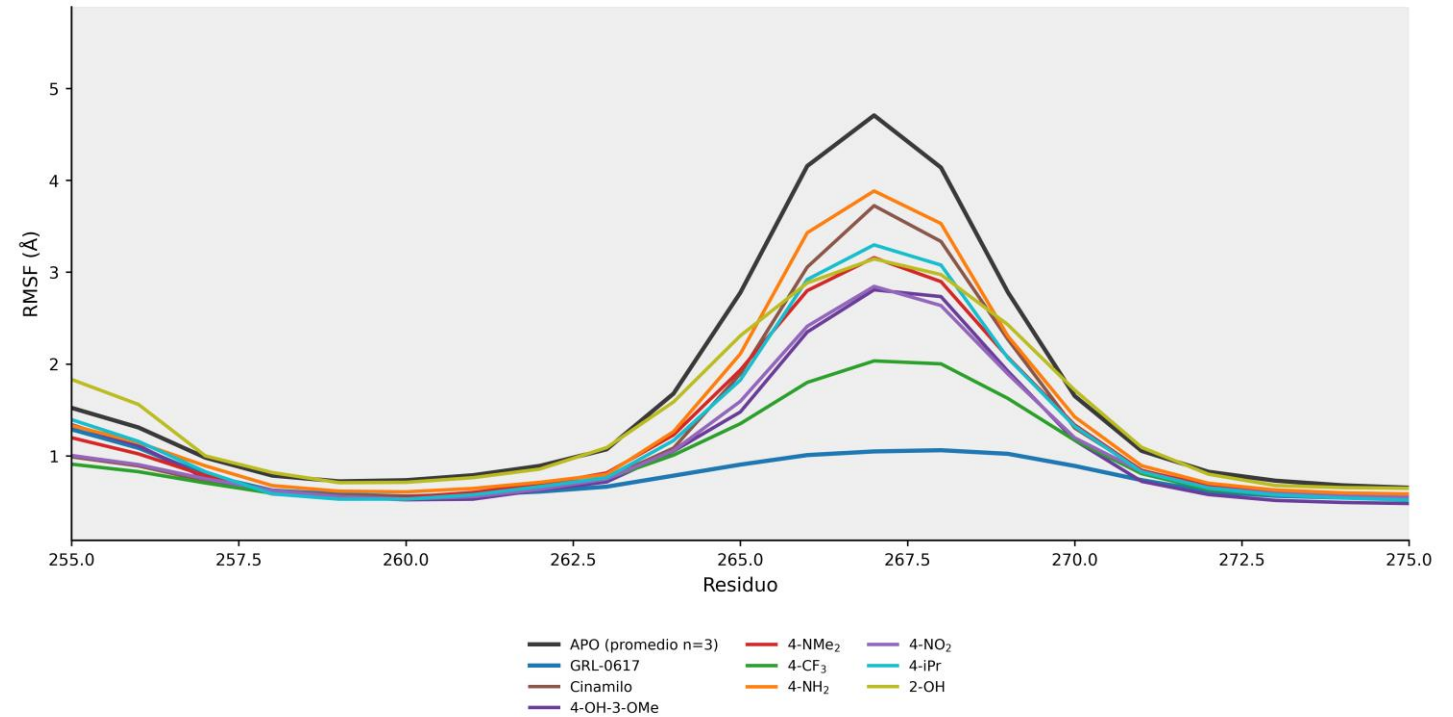


Figura 69 Zoom del perfil de RMSF en el loop BL2 (residuos 255-275) durante las simulaciones de dinámica molecular de los complejos PLpro-derivados bencénicos. Se incluyen los sistemas APO, GRL-0617 y los ligandos evaluados, permitiendo comparar las variaciones de flexibilidad en esta región estructural clave del sitio de unión.

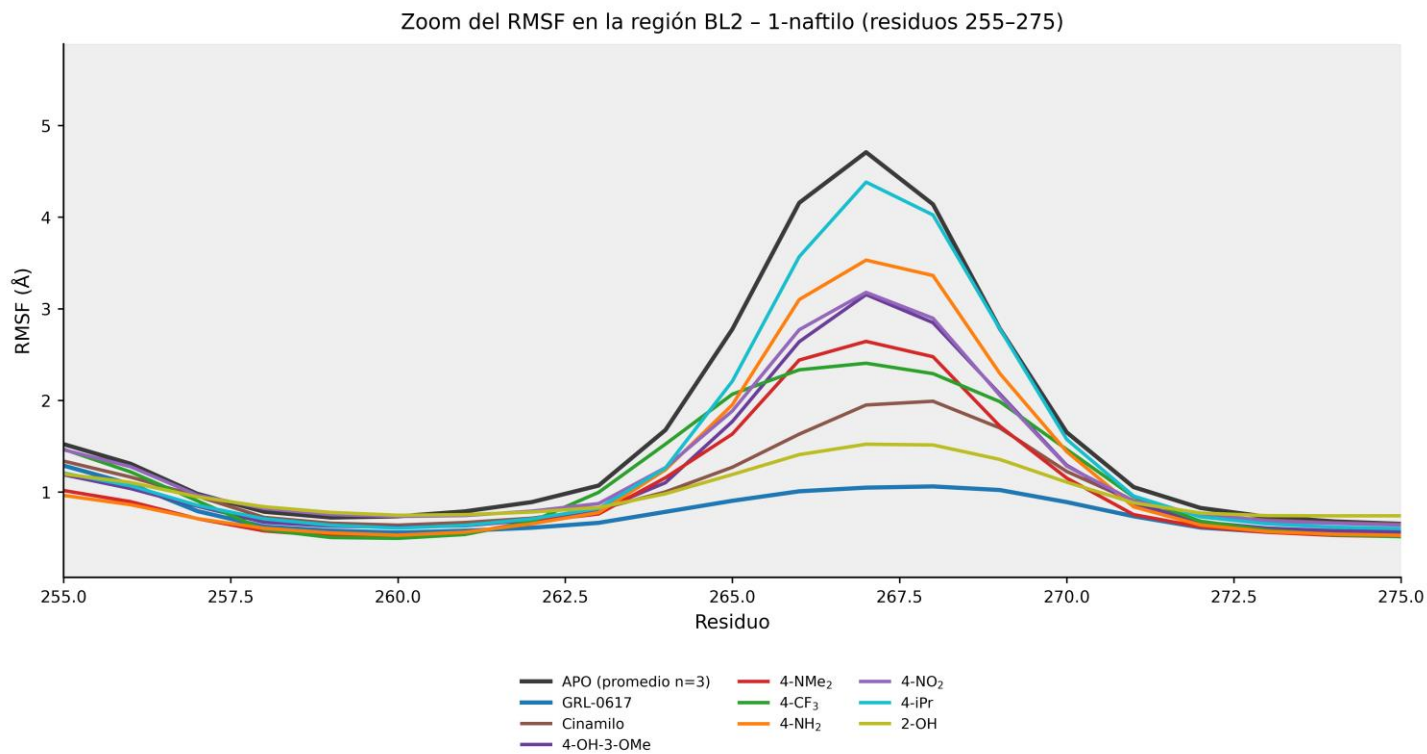


Figura 70 Zoom del RMSF en la región BL2 (255-275) para los derivados de 1-naftilo evaluados.

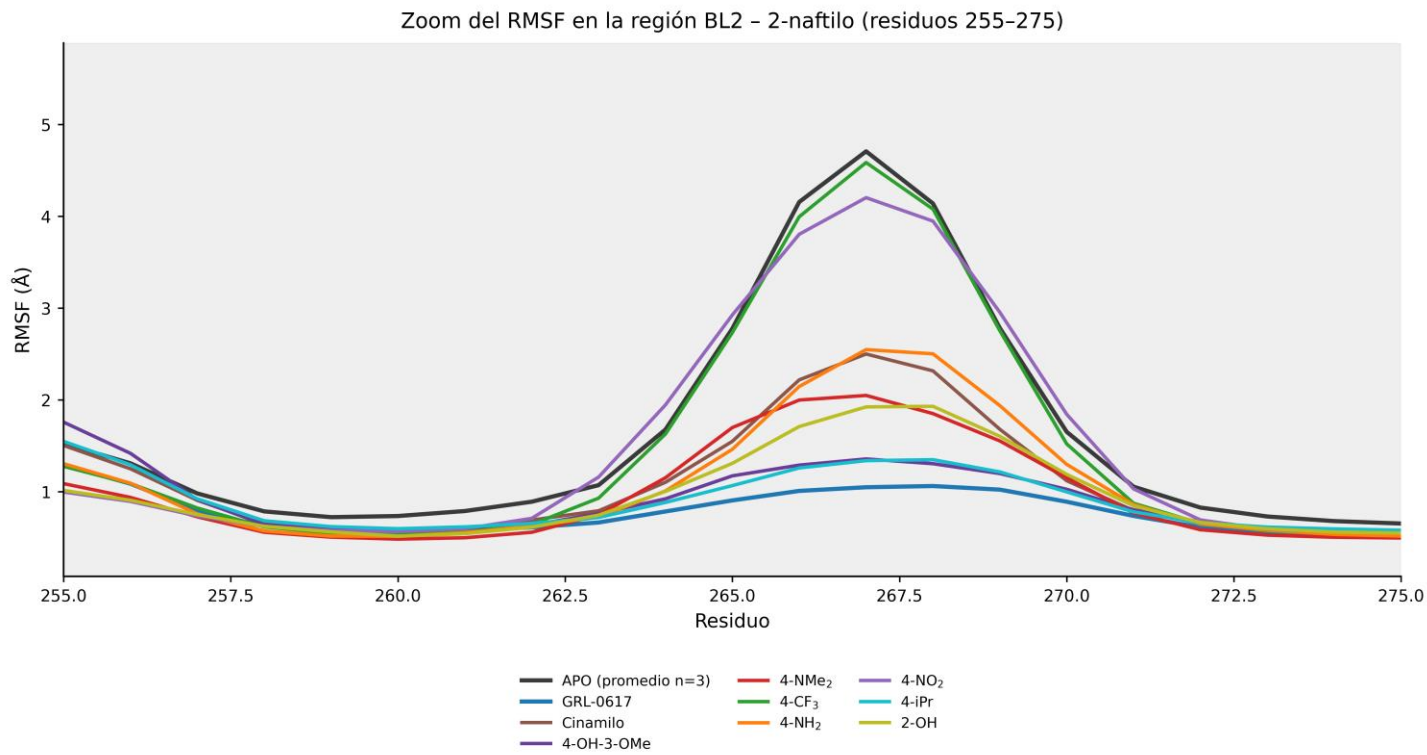


Figura 71 Zoom del RMSF en la región BL2 (255-275) para los derivados de 2-naftilo evaluados.

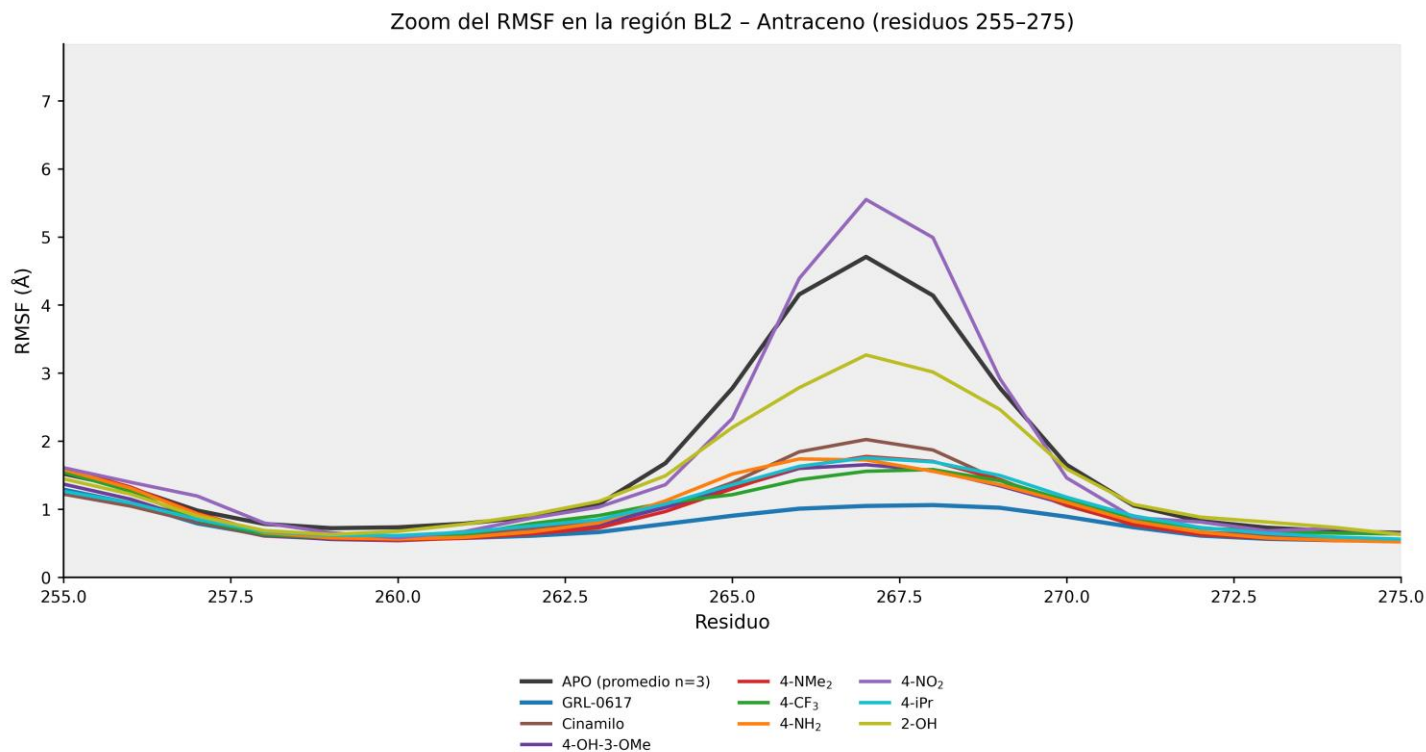


Figura 72 Zoom del RMSF en la región BL2 (255–275) para los derivados de antraceno evaluados.

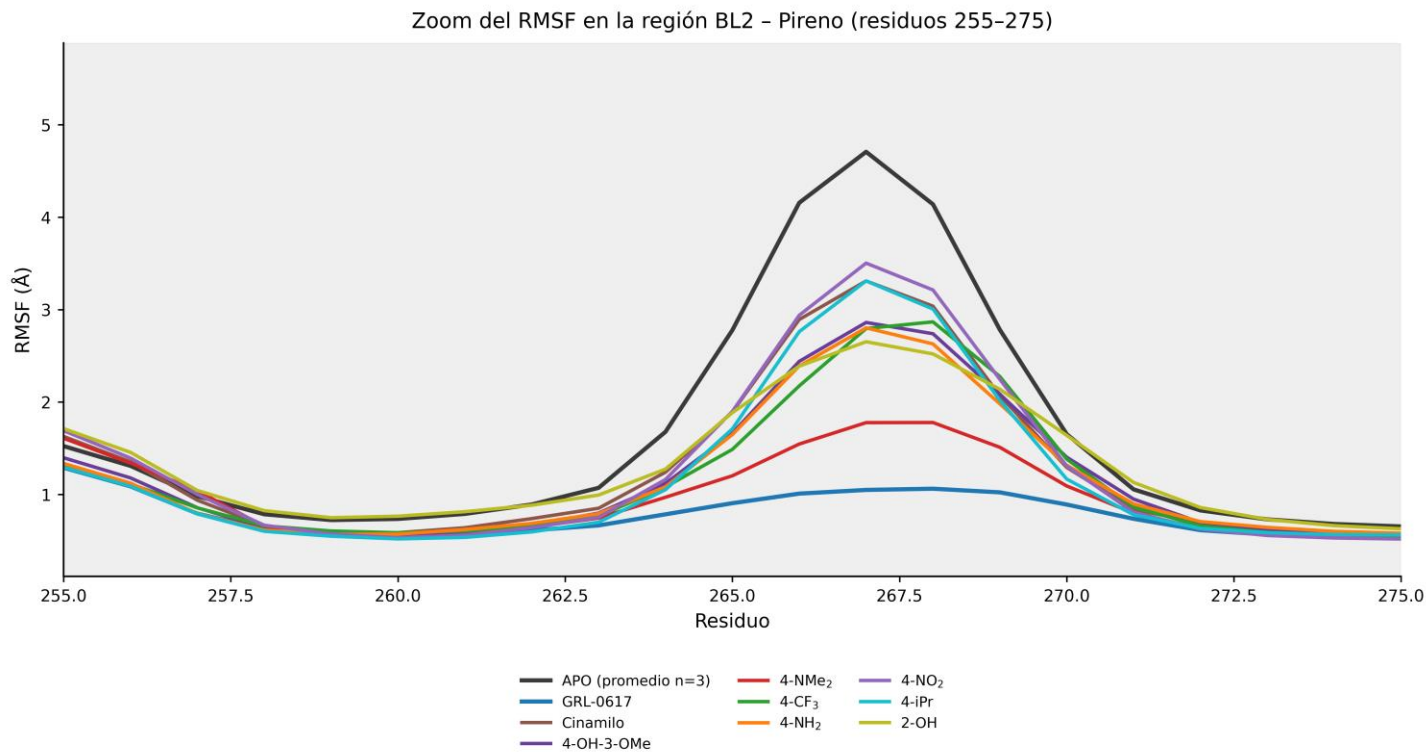


Figura 73 Zoom del RMSF en la región BL2 (255–275) para los derivados de pireno evaluados.

Los resultados de distancias Cys111–His272 muestran que la mayoría de los sistemas mantiene valores de distancia muy similares a los observados tanto para la proteína en estado apo como para el complejo con el inhibidor de referencia GRL-0617(Figura 74-Figura 78), con fluctuaciones centradas aproximadamente entre 6.4 y 6.8 Å a lo largo de las trayectorias de Dinámica molecular. Esto indica que la presencia de los distintos ligandos no genera perturbaciones significativas en la geometría del sitio catalítico.

La única desviación relativamente más evidente se observa en el derivado 4-OH-3-OMe del fragmento 2-naftilo, el cual presenta una tendencia a valores ligeramente menores de distancia en comparación con los demás sistemas. Sin embargo, incluso en este caso las variaciones permanecen dentro de un rango estrecho y no representan un cambio estructural importante en la organización de la díada catalítica.

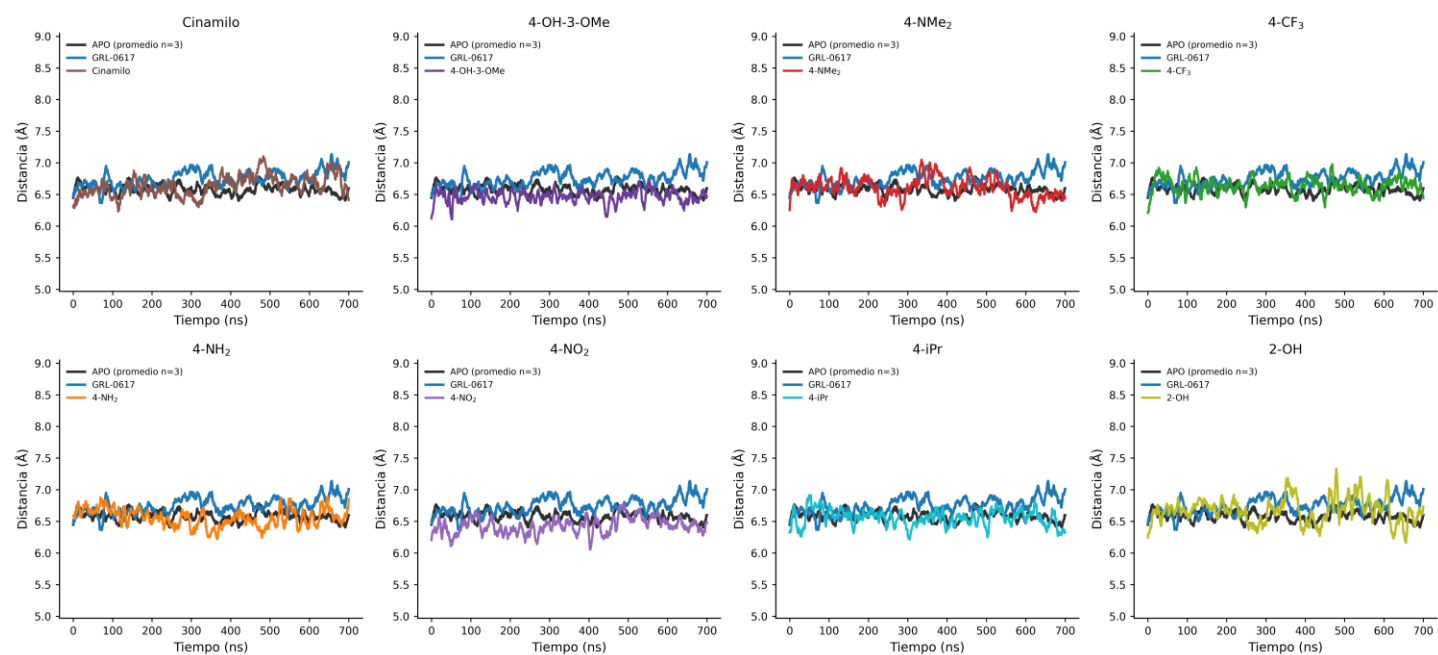


Figura 74 Evolución temporal de la distancia Cys111–His272 (CA–CA) durante las simulaciones de Dinámica molecular. Cada panel muestra la trayectoria de la distancia catalítica para uno de los compuestos evaluados. La línea negra corresponde al promedio de las tres réplicas de la proteína APO, mientras que la línea azul representa el complejo con el inhibidor de referencia GRL-0617. Las líneas de colores corresponden a los distintos derivados de acetofenona evaluados con sus respectivo derivados de aldehído (Cinamilo, 4-OH-3-OMe, 4-NMe₂, 4-CF₃, 4-NH₂, 4-NO₂, 4-iPr y 2-OH). En todos los sistemas, la distancia entre los residuos catalíticos se mantiene relativamente constante a lo largo de la trayectoria de 700 ns, indicando la conservación de la geometría del sitio catalítico de PLpro durante la simulación.

Distancia C111-H272 - 1-naftilo

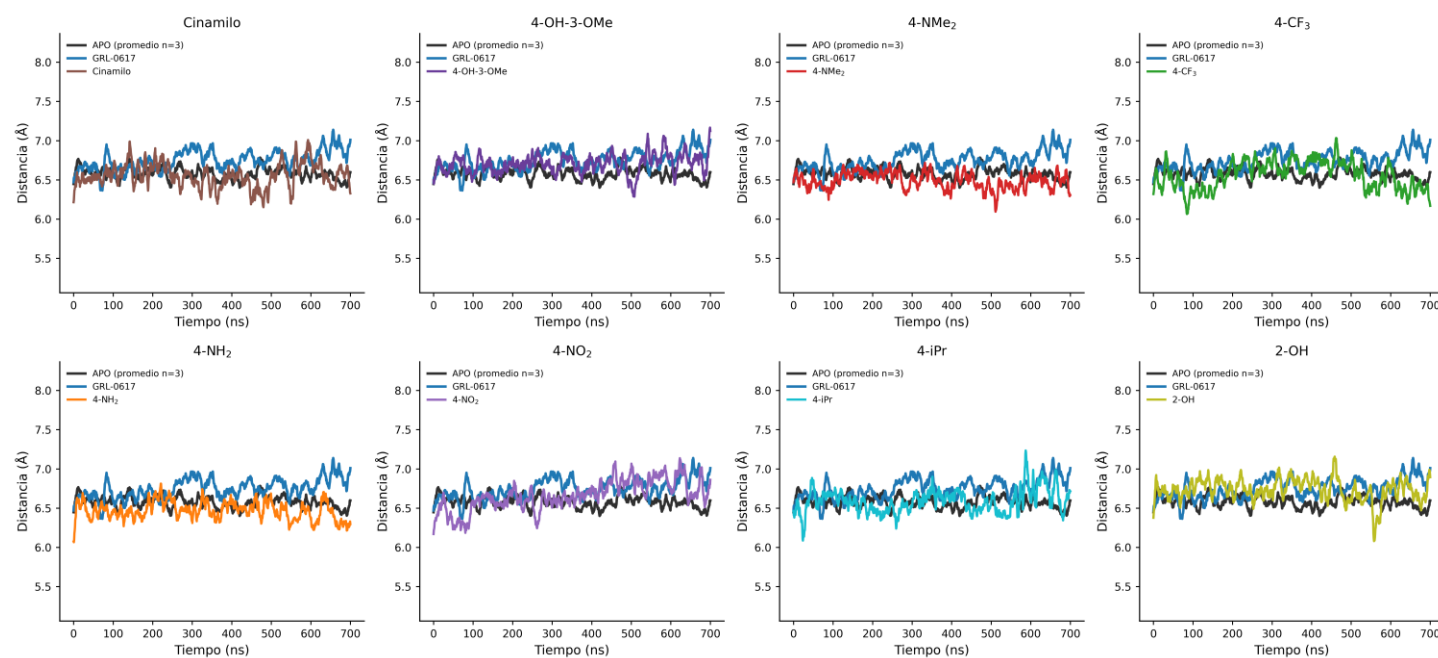


Figura 75 Evolución temporal de la distancia Cys111–His272 (CA–CA) para los derivados con sustituyente 1-naftilo.

Distancia C111-H272 – 2-naftilo

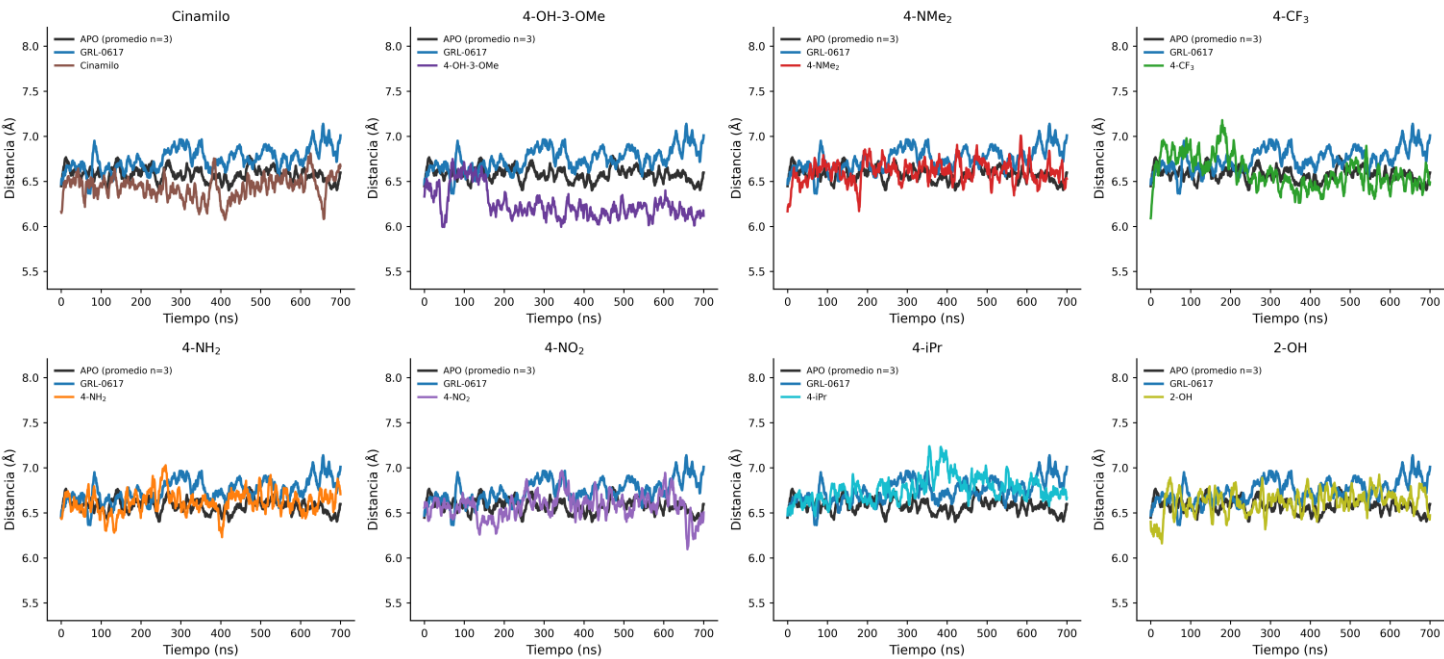


Figura 76 Evolución temporal de la distancia Cys111-His272 (CA-CA) para los derivados con sustituyente 2-naftilo.

Distancia C111-H272 – Antraceno

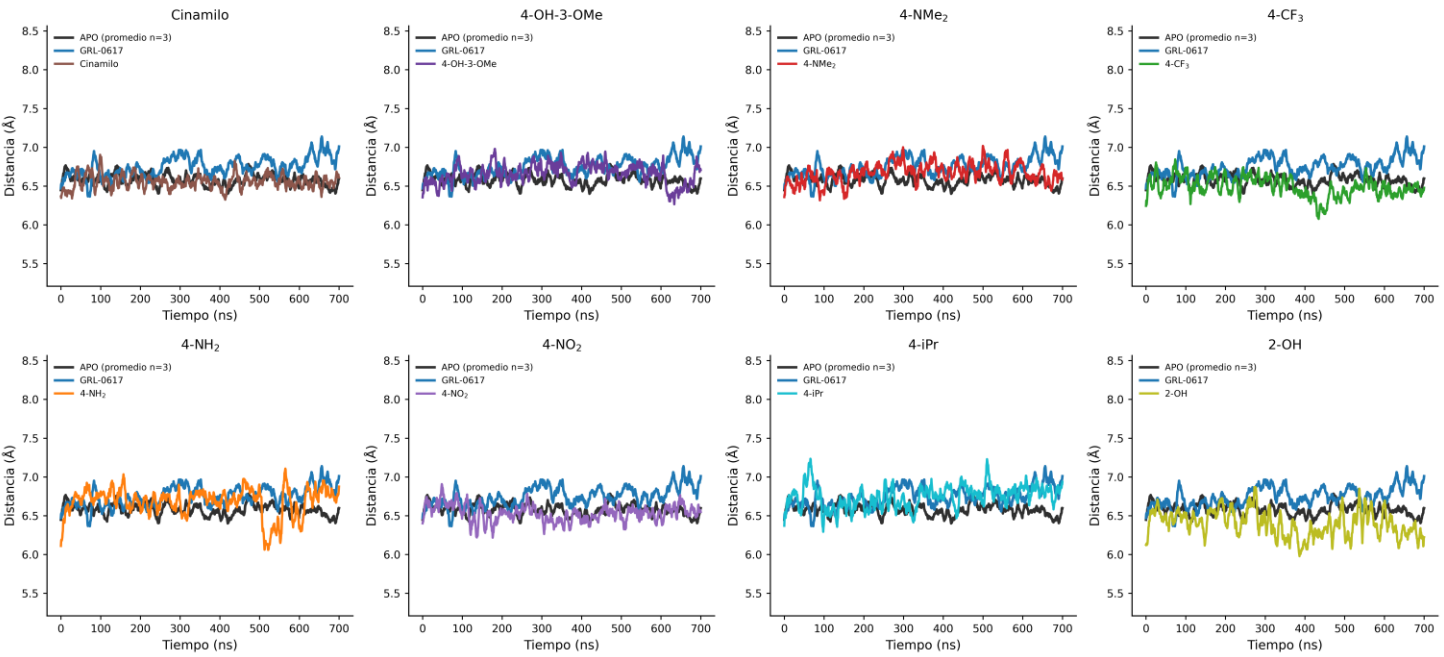


Figura 77 Evolución temporal de la distancia Cys111-His272 (CA-CA) para los derivados con sustituyente antraceno.

Distancia C111-H272 – Pireno

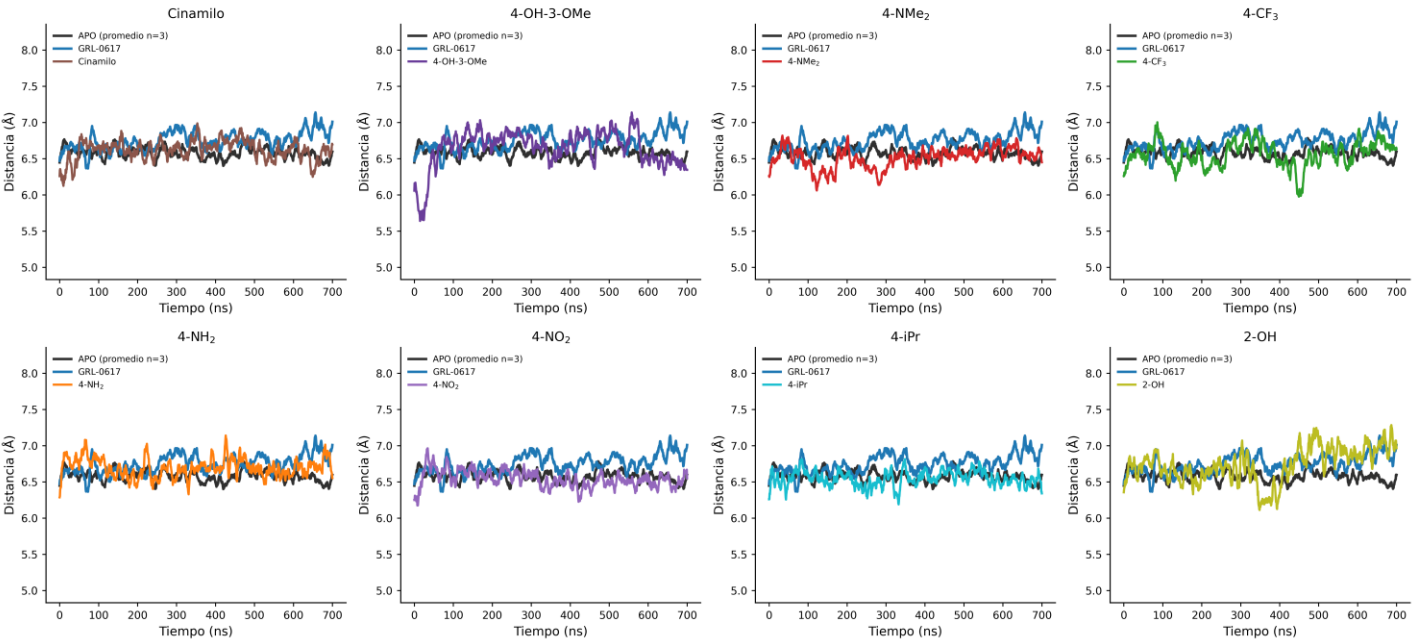


Figura 78 Evolución temporal de la distancia Cys111–His272 (CA–CA) para los derivados con sustituyente pireno.

Con respecto a las distancias Tyr268–Arg166 se vieron diferentes comportamientos con respecto a cada familia(Figura 79-Figura 83).

En la familia benceno, varios derivados mostraron una tendencia a disminuir la distancia Tyr268–Arg166 respecto al estado apo, acercándose al comportamiento del inhibidor de referencia. Entre ellos, 4-CF₃ y 2-OH presentaron los perfiles más claramente asociados a distancias menores, mientras que otros sistemas, como 4-OH-3-OMe, 4-NMe₂, 4-NH₂ y 4-iPr, mostraron trayectorias más erráticas o transiciones entre estados.

En la serie 1-naftilo, las trayectorias mostraron comportamientos más variables entre sistemas. El derivado 4-NMe₂ presentó valores de distancia más elevados durante gran parte de la trayectoria. Por otro lado, 4-CF₃ y 2-OH mostraron perfiles con valores cercanos a los observados para GRL-0617. Otros ligandos, como 4-NH₂, 4-NO₂ y 4-iPr, alternaron entre intervalos de distancia menores y mayores a lo largo de la simulación.

En la familia 2-naftilo, se observaron diferencias claras entre los sistemas. El derivado cinamilo mantuvo los valores de distancia más altos incluso que la propia proteína nativa durante gran parte de la trayectoria. En contraste, 4-NMe₂, 4-iPr y 2-OH mostraron valores de distancia menores , relativamente estables y similares a GRL-0617. El derivado 4-OH-3-OMe presentó un perfil relativamente constante, con valores ligeramente superiores a los del inhibidor de referencia. Por su parte, 4-CF₃ y 4-NH₂ se mantuvieron en un rango intermedio de valores.

En la serie antraceno, varios sistemas mostraron valores de distancia Tyr268–Arg166 menores que los observados para el estado nativo. Este comportamiento se observó particularmente en los derivados 4-NMe₂, 4-CF₃, 4-NH₂ y 4-iPr, cuyos perfiles se mantuvieron cercanos o por debajo del rango observado para GRL-0617. En contraste, el derivado cinamilo presentó valores más elevados durante la trayectoria, mientras que 2-OH mostró fluctuaciones más pronunciadas con cambios hacia valores menores en la parte final de la simulación.

Para la familia pireno también se observaron diferencias entre sistemas. El derivado 4-NMe₂ presentó valores de distancia menores durante gran parte de la trayectoria. Otros compuestos, como 2-OH, también alcanzaron intervalos de distancia baja, aunque con mayor variabilidad. En contraste, cinamilo y 4-NO₂ mostraron valores de distancia altos como la proteína en estado nativo.

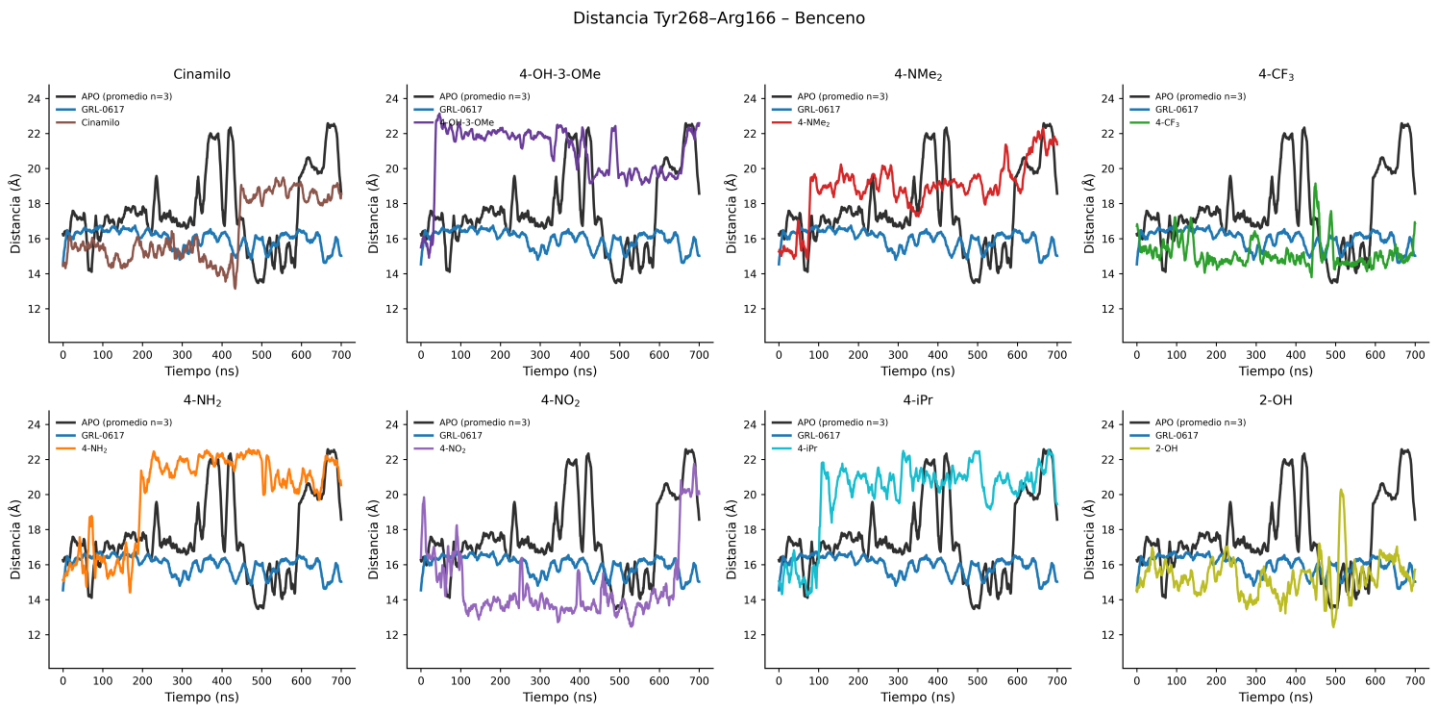


Figura 79 Evolución temporal de la distancia Tyr268–Arg166(CA–CA) para los derivados con sustituyente benceno.

Distancia Tyr268-Arg166 – 1-naftilo

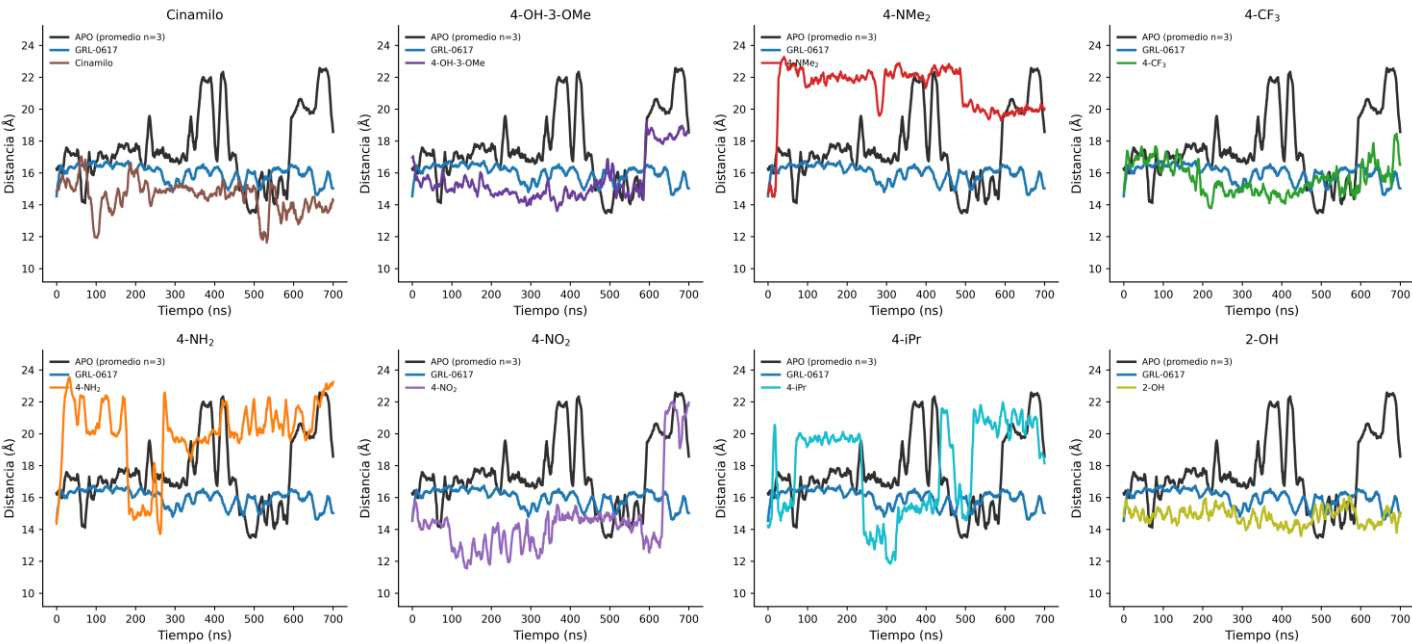


Figura 80 Evolución temporal de la distancia Tyr268-Arg166(CA-CA) para los derivados con sustituyente 1-naftilo.

Distancia Tyr268-Arg166 – 2-naftilo

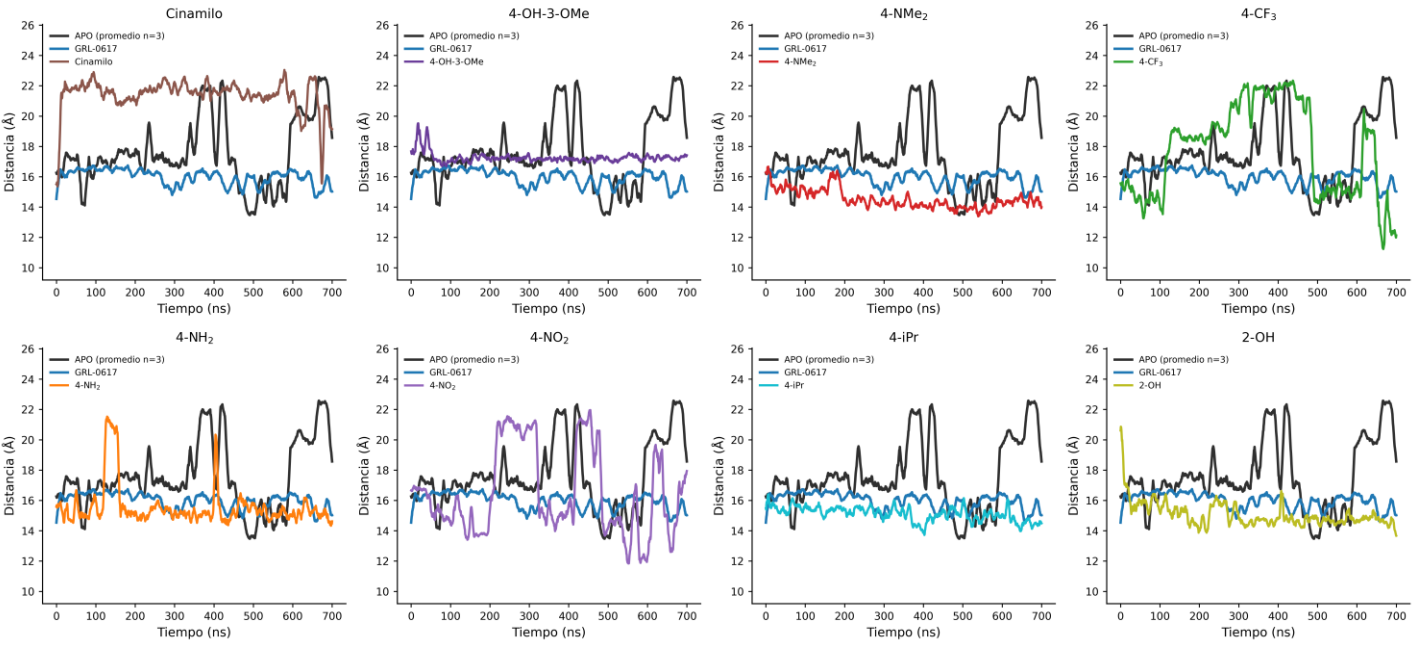


Figura 81 Evolución temporal de la distancia Tyr268-Arg166(CA-CA) para los derivados con sustituyente 2-naftilo.

Distancia Tyr268-Arg166 – Antraceno

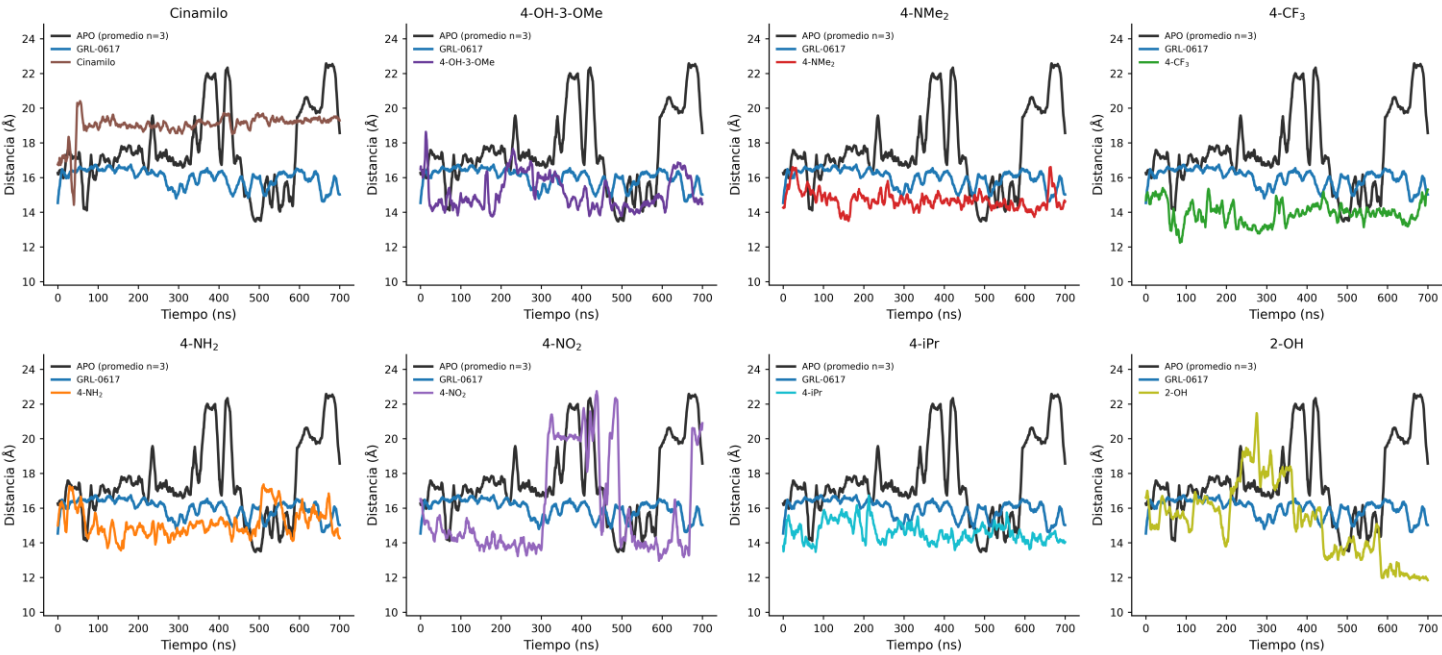


Figura 82 Evolución temporal de la distancia Tyr268-Arg166(CA-CA) para los derivados con sustituyente antraceno.

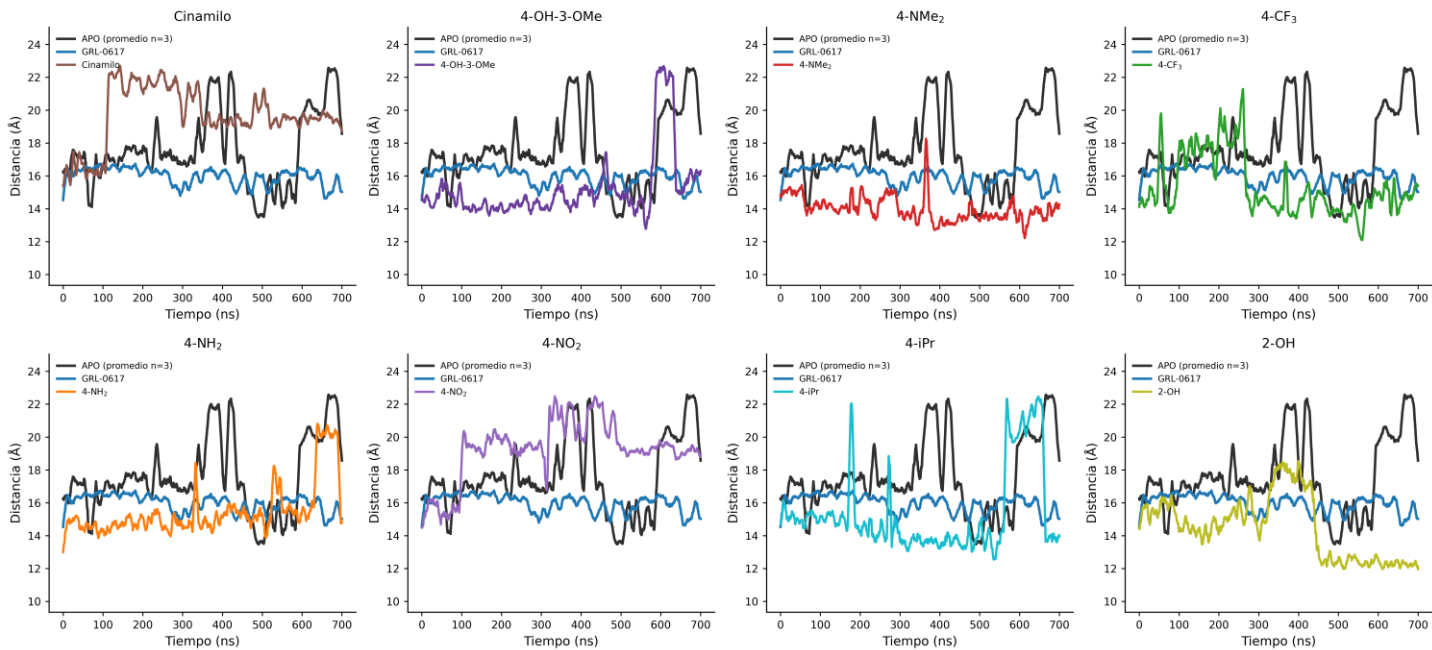


Figura 83 Evolución temporal de la distancia Tyr268-Arg166(CA-CA) para los derivados con sustituyente pireno.

Para la distancia Tyr268–Pro299, se observaron diferencias entre sistemas en comparación con el estado apo y con el complejo GRL-0617 (Figura 84-Figura 88).

En la familia benceno, el derivado cinamilo mostró inicialmente valores estables cercanos al rango observado para GRL-0617, aunque hacia el final de la trayectoria se registró un incremento de la distancia. El derivado 4-OH-3-OMe presentó valores comparables a los del estado apo durante gran parte de la simulación. De manera similar, 4-NMe₂ mostró valores cercanos a apo, con fluctuaciones moderadas. En contraste, 4-CF₃, 2-OH y 4-NO₂ mostraron intervalos en los que la distancia disminuyó por debajo de los valores observados para GRL-0617, aunque con trayectorias más variables. Otros sistemas dentro de esta serie mostraron fluctuaciones amplias sin una tendencia constante.

En la serie 1-naftilo, se observaron comportamientos más heterogéneos. El derivado cinamilo presentó valores de distancia intermedios entre apo y GRL-0617. El sistema 4-OH-3-OMe mostró inicialmente distancias menores, seguidas de un aumento progresivo durante la trayectoria. El derivado 4-NMe₂ alcanzó valores de distancia superiores tanto a apo como a GRL-0617 durante una parte significativa de la simulación. En contraste, 4-CF₃ presentó intervalos con distancias menores, cercanas o inferiores a las observadas para el inhibidor de referencia, aunque con fluctuaciones. Los derivados 4-NH₂ y 4-NO₂ mostraron trayectorias más variables, alternando entre valores mayores y menores respecto a GRL-0617. El sistema 4-iPr mostró valores cercanos al inhibidor de referencia, mientras que 2-OH presentó valores próximos a GRL-0617 durante gran parte de la trayectoria.

En la familia 2-naftilo, el derivado cinamilo presentó valores de distancia superiores a los observados para GRL-0617. El sistema 4-OH-3-OMe mostró un perfil prácticamente superpuesto al del inhibidor de referencia. El derivado 4-NMe₂ presentó valores de distancia menores que GRL-0617 durante gran parte de la trayectoria. Por su parte, 4-CF₃ mostró valores más elevados y fluctuantes. Los sistemas 4-NH₂ y 4-NO₂ presentaron trayectorias erráticas con transiciones entre valores mayores y menores respecto al inhibidor de referencia. El derivado 4-iPr mostró valores cercanos a GRL-0617, mientras que 2-OH presentó fluctuaciones moderadas con valores ligeramente menores.

En la serie antraceno, los derivados 4-NMe₂ y 4-OH-3-OMe mostraron valores de distancia comparables a los del estado apo, con fluctuaciones a lo largo de la trayectoria. Un comportamiento similar se observó para 4-NH₂. Otros sistemas presentaron intervalos con valores menores o mayores respecto a GRL-0617, sin un patrón constante.

Finalmente, en la familia pireno, el derivado cinamilo mostró valores de distancia superiores a los observados para GRL-0617 durante gran parte de la trayectoria. El sistema 4-OH-3-OMe presentó valores cercanos a los del estado apo, con fluctuaciones moderadas. El derivado 4-NMe₂ mostró valores próximos al inhibidor de referencia. En algunos sistemas, como 2-OH, se observó una disminución progresiva de la distancia hacia el final de la simulación.

Distancia Tyr268-Pro299 – Benceno

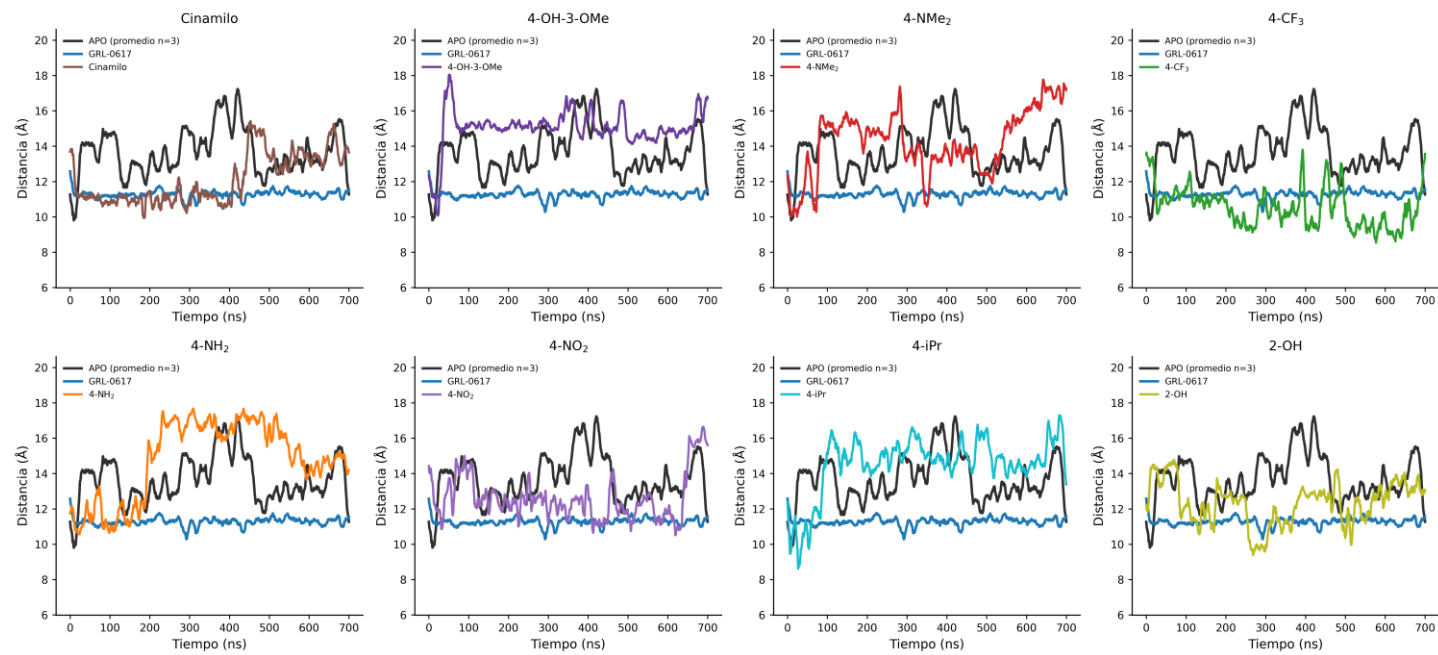


Figura 84 Evolución temporal de la distancia Tyr268-Pro299(CA-CA) para los derivados con sustituyente benceno. La definición de las líneas y condiciones de simulación es la misma que en la Figura 68.

Distancia Tyr268-Pro299 – 1-naftilo

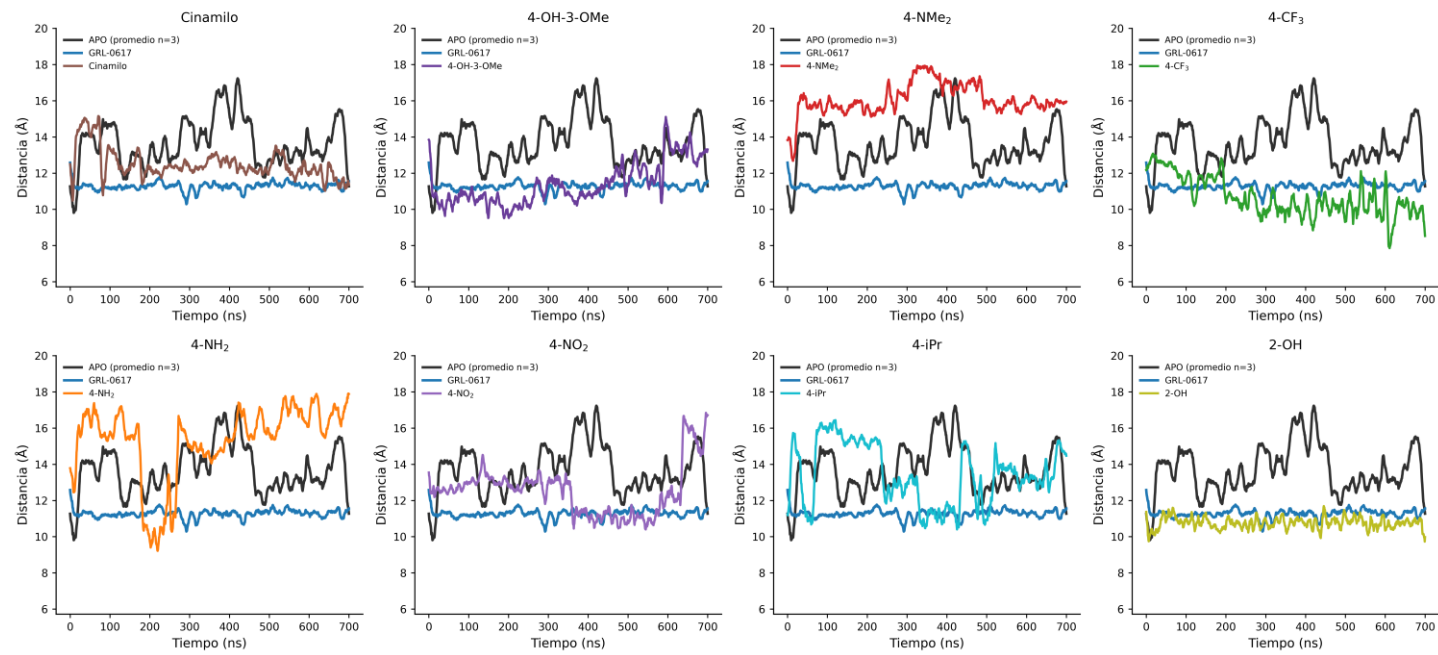


Figura 85 Evolución temporal de la distancia Tyr268-Pro299(CA-CA) para los derivados con sustituyente 1-naftilo. La definición de las líneas y condiciones de simulación es la misma que en la Figura 68.

Distancia Tyr268-Pro299 – 2-naftilo

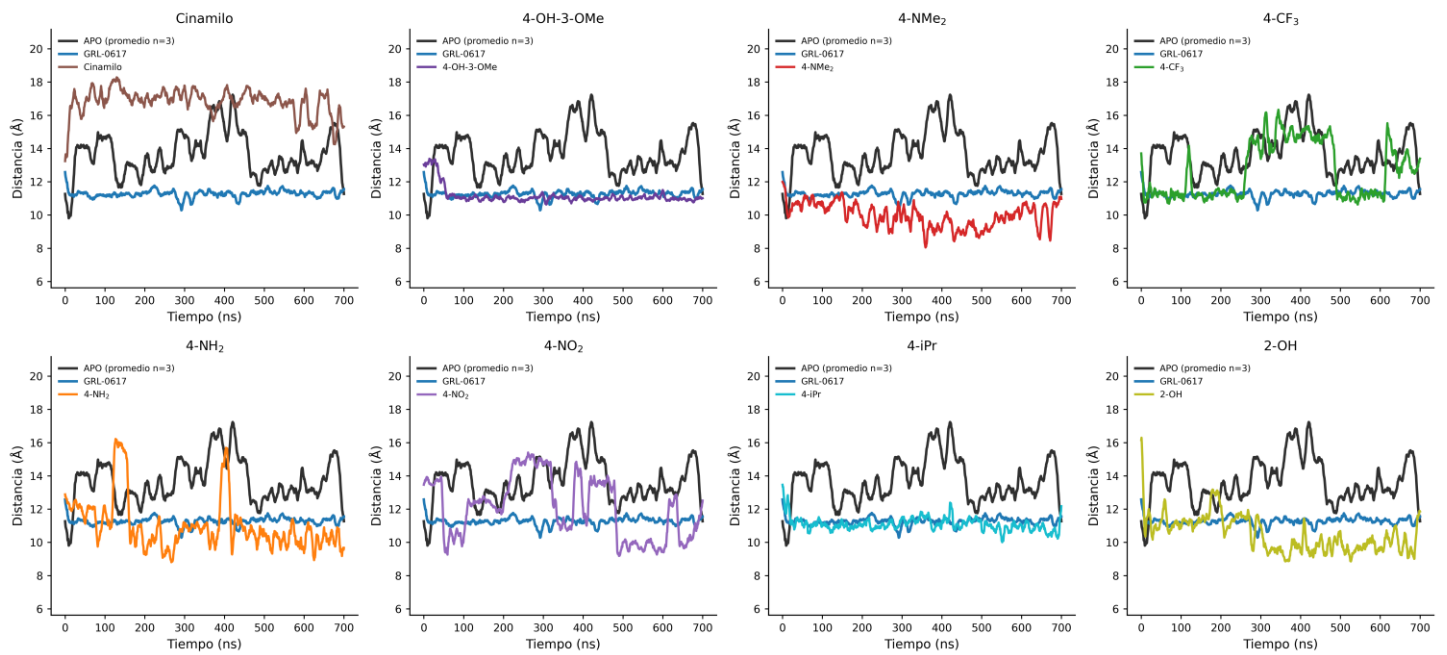


Figura 86 Evolución temporal de la distancia Tyr268-Pro299 (CA-CA) para los derivados con sustituyente 2-naftilo. La definición de las líneas y condiciones de simulación es la misma que en la Figura 68.

Distancia Tyr268-Pro299 – Antraceno

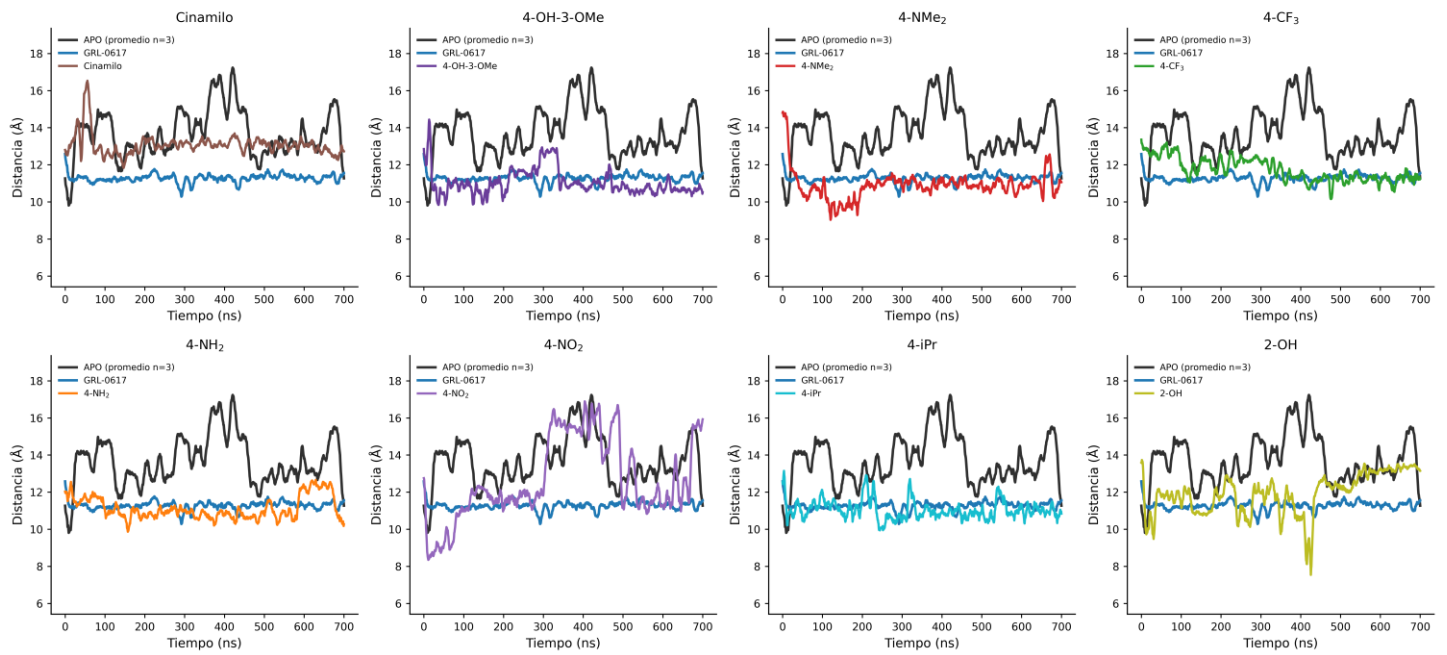


Figura 87 Evolución temporal de la distancia Tyr268-Pro299(CA-CA) para los derivados con sustituyente antraceno. La definición de las líneas y condiciones de simulación es la misma que en la Figura 68.

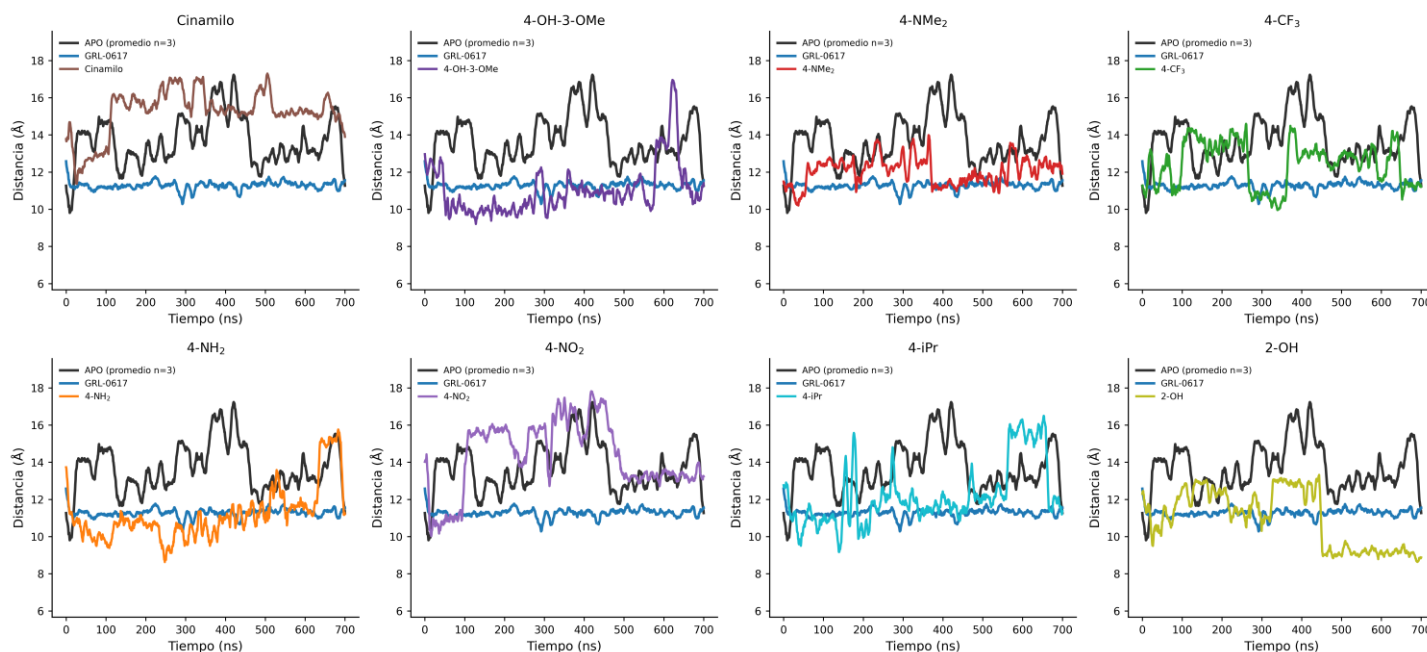


Figura 88 Evolución temporal de la distancia Tyr268-Pro299(CA-CA) para los derivados con sustituyente pireno. La definición de las líneas y condiciones de simulación es la misma que en la Figura 68.

Para los ángulos de diedro del aminoácido Y268 se observaron diferentes tendencias(Figura 89-Figura 98) : En la familia benceno, la distribución del ángulo χ_1 mostró que la mayoría de los derivados presentan poblaciones predominantes alrededor de -45° , similares a las observadas para varios sistemas con ligando. En contraste, los derivados 4-NO₂ y 2-OH mostraron poblaciones dominantes cercanas a 180° , coincidiendo con la orientación predominante observada en el inhibidor de referencia GRL-0617. El resto de los sustituyentes, incluyendo cinamilo, 4-OH-3-OMe, 4-NMe₂, 4-CF₃ y 4-iPr, presentaron distribuciones centradas principalmente alrededor de -45° , con menor población en otros intervalos angulares. Para el ángulo χ_2 en esta misma familia, la mayoría de los sistemas presentaron distribuciones verticales alrededor de $\pm 90^\circ$, patrón también observado para GRL-0617. Este comportamiento fue más pronunciado en los derivados 4-OH-3-OMe, 4-iPr, 2-OH y 4-NH₂. En contraste, el derivado 4-NO₂ mostró una distribución más dispersa con poblaciones adicionales fuera de este eje principal. En la serie 1-naftilo, el ángulo χ_1 mostró que el derivado cinamilo presenta una distribución cercana a 180° , similar a la observada para GRL-0617. Un patrón comparable también se observó para 4-OH-3-OMe. Por el contrario, otros sistemas como 4-CF₃, 4-iPr y 2-OH mostraron poblaciones predominantes alrededor de -45° , mientras que 4-NO₂ presentó nuevamente una distribución centrada en 180° . Para el ángulo χ_2 en 1-naftilo, se observó un patrón diferente para cinamilo, cuya población principal se desplazó hacia valores cercanos a 45° , alejándose del comportamiento observado para GRL-0617 y el sistema apo. Un comportamiento similar también se observó para 4-NO₂. En contraste, los demás sistemas presentaron distribuciones dominantes alrededor de $\pm 90^\circ$, comparables a las observadas para el inhibidor de referencia. En la familia 2-naftilo, el ángulo χ_1 mostró que cinamilo presenta poblaciones predominantes cercanas a -45° , similares al patrón observado para el sistema apo. En cambio, el derivado 4-OH-3-OMe mostró una distribución prácticamente superpuesta con la del inhibidor GRL-0617, centrada en 180° . Otros derivados como 4-CF₃, 4-NH₂ y 4-iPr presentaron poblaciones predominantes cercanas a -45° , mientras que 4-NO₂ mostró nuevamente una distribución centrada en 180° . Para el ángulo χ_2 en esta serie, el derivado 4-OH-3-OMe mostró nuevamente una distribución prácticamente idéntica a la observada para GRL-0617, con poblaciones dominantes alrededor de $\pm 90^\circ$. El resto de los sistemas también presentaron poblaciones en este intervalo, aunque con mayor dispersión. En la familia antraceno, el ángulo χ_1 mostró que cinamilo presenta una distribución centrada alrededor de 180° , similar a GRL-0617. Un patrón comparable también se observó para 4-CF₃ y 2-OH, cuyos ángulos se agrupan en torno a esta región. En contraste, derivados como 4-OH-3-OMe, 4-NMe₂ y 4-iPr mostraron poblaciones predominantes alrededor de -45° . Para el ángulo χ_2 en antraceno, ninguno de los sistemas reprodujo exactamente la distribución observada para el inhibidor de referencia. La mayoría de los derivados presentaron poblaciones alrededor de $\pm 90^\circ$,

aunque con distribuciones más amplias. El derivado cinamilo mostró una distribución particularmente desplazada hacia -45° .

En la familia pireno, el ángulo χ_1 mostró que la mayoría de los derivados presentan poblaciones alrededor de -45° , mientras que el derivado 4-NMe₂ mostró una distribución más concentrada en torno a 180° , similar al patrón observado para GRL-0617. En cuanto al ángulo χ_2 , la mayoría de los sistemas mostraron distribuciones dispersas sin una orientación predominante claramente definida, aunque el derivado 4-iPr presentó una distribución más marcada alrededor de $\pm 90^\circ$, cercana al patrón observado para el inhibidor de referencia.

Distribución polar de χ_1 – Benceno

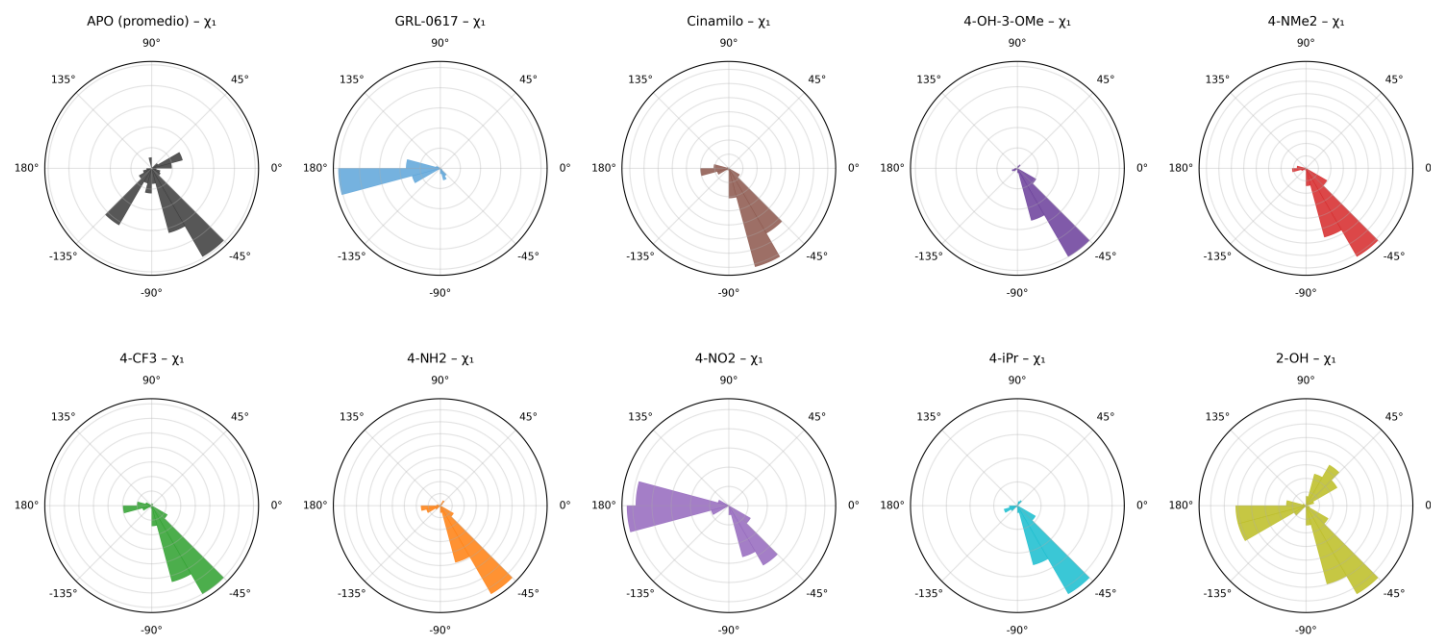


Figura 89 Distribución polar del ángulo diedro χ_1 del residuo Tyr268 durante las simulaciones de Dinámica molecular para los complejos con derivados de la familia benceno. Los paneles superiores muestran la distribución para la proteína APO (promedio de tres réplicas) y para el complejo con el inhibidor de referencia GRL-0617, mientras que los paneles restantes corresponden a los distintos derivados evaluados (Cinamilo, 4-OH-3-OMe, 4-NMe₂, 4-CF₃, 4-NH₂, 4-NO₂, 4-iPr y 2-OH). La distribución angular se representa en coordenadas polares, donde el radio corresponde a la frecuencia de observación del ángulo durante la trayectoria de Dinámica molecular de 700 ns.

Distribución polar de χ_2 – Benceno

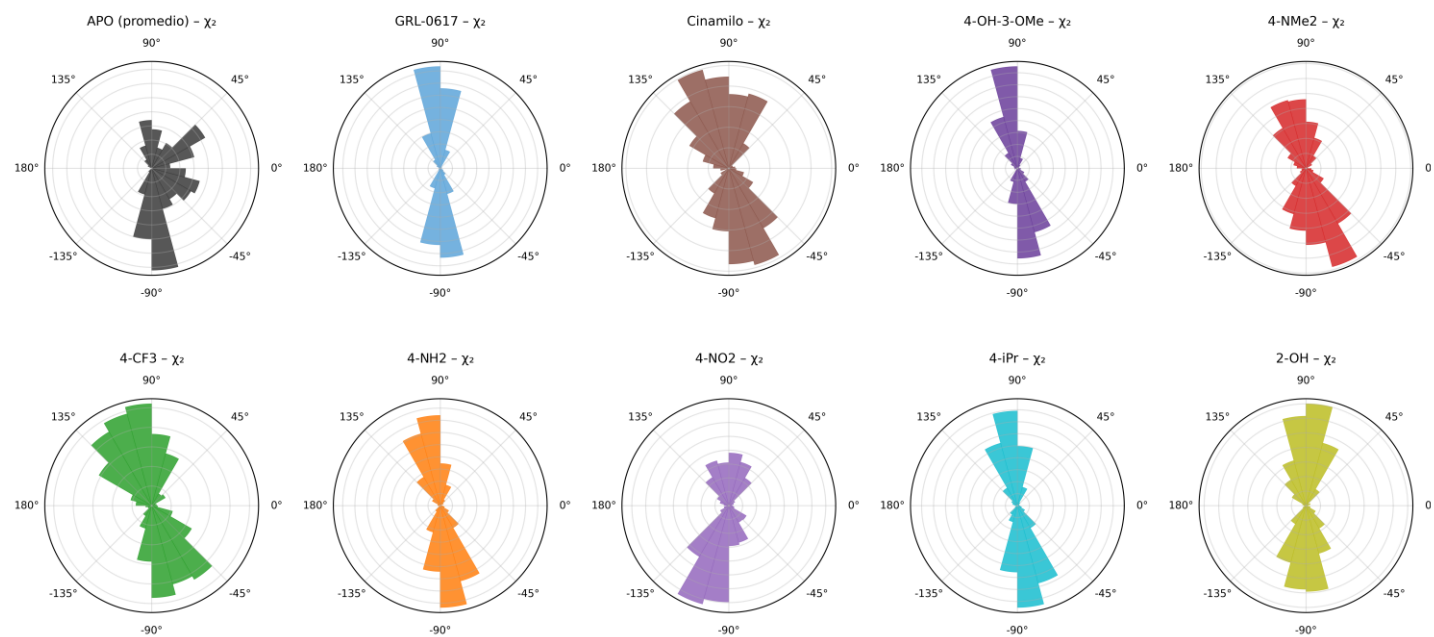


Figura 90 Distribución polar del ángulo diedro χ_2 del residuo Tyr268 durante las simulaciones de Dinámica molecular para los complejos con derivados de la familia benceno. La organización de los paneles y la definición de las curvas corresponden a las descritas en la Figura 83.

Distribución polar de χ_1 - 1-naftilo

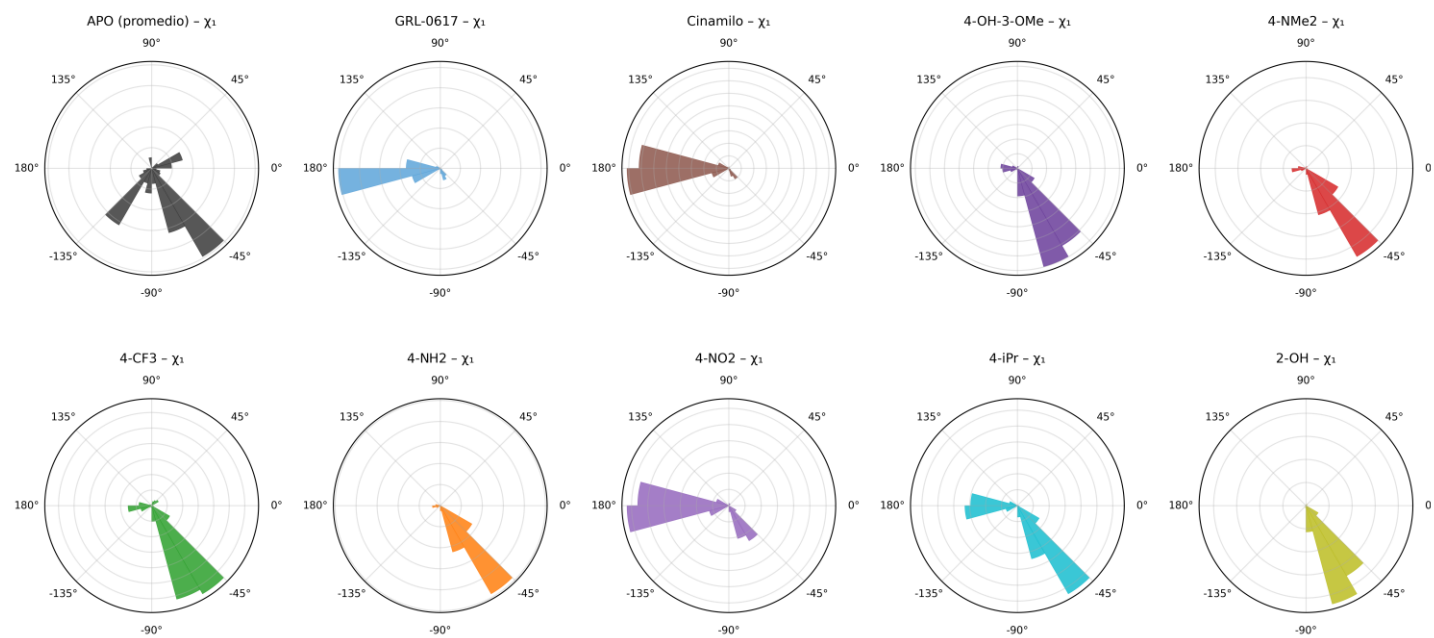


Figura 91 Distribución polar del ángulo diedro χ_1 del residuo Tyr268 durante las simulaciones de Dinámica molecular para los complejos con derivados de la familia 1-naftilo. La organización de los paneles y la definición de las curvas corresponden a las descritas en la Figura 83.

Distribución polar de χ_2 - 1-naftilo

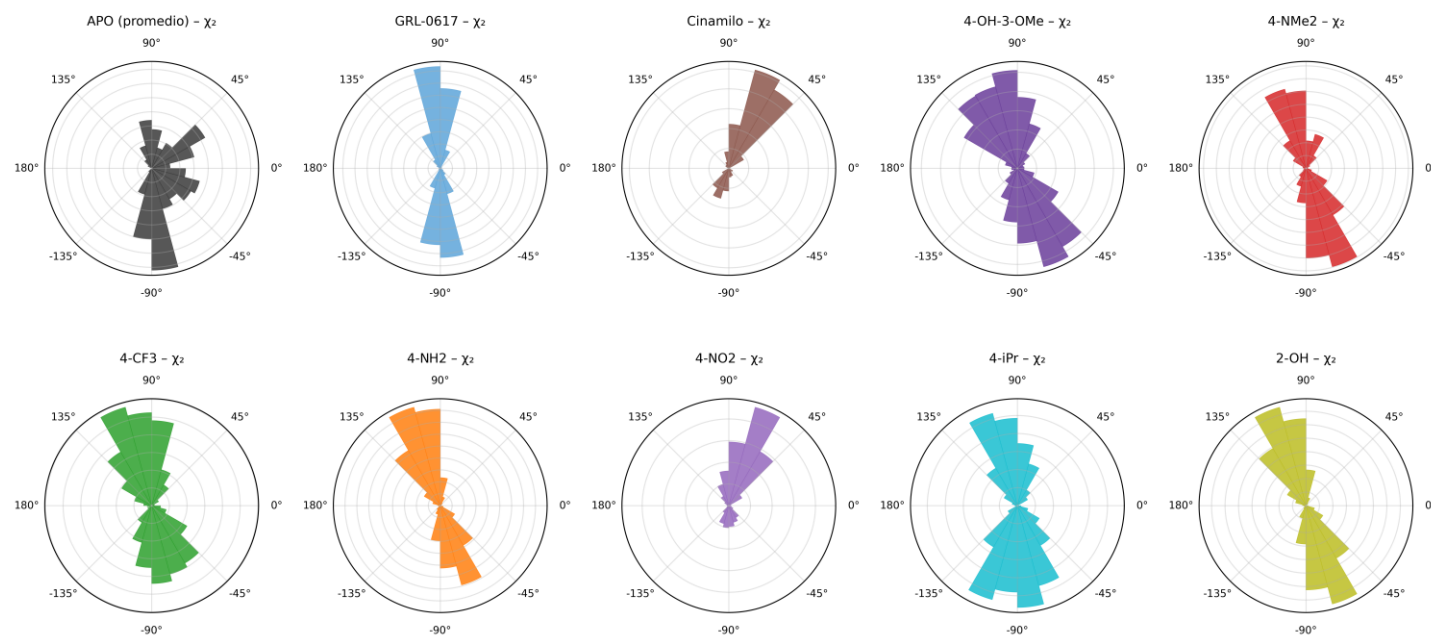


Figura 92 Distribución polar del ángulo diedro χ_2 del residuo Tyr268 durante las simulaciones de Dinámica molecular para los complejos con derivados de la familia 1-naftilo. La organización de los paneles y la definición de las curvas corresponden a las descritas en la Figura 83.

Distribución polar de χ_1 - 2-naftilo

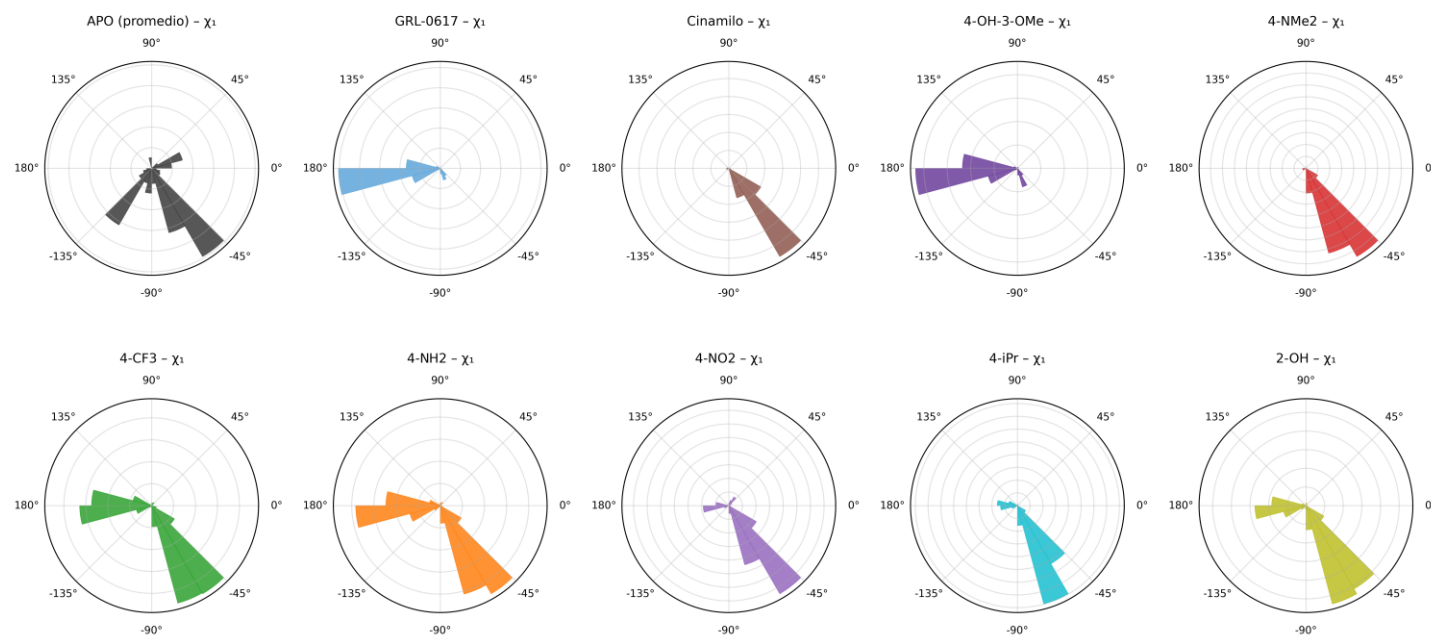


Figura 93 Distribución polar del ángulo diedro χ_1 del residuo Tyr268 durante las simulaciones de Dinámica molecular para los complejos con derivados de la familia 2-naftilo. La organización de los paneles y la definición de las curvas corresponden a las descritas en la Figura 83.

Distribución polar de χ_2 - 2-naftilo

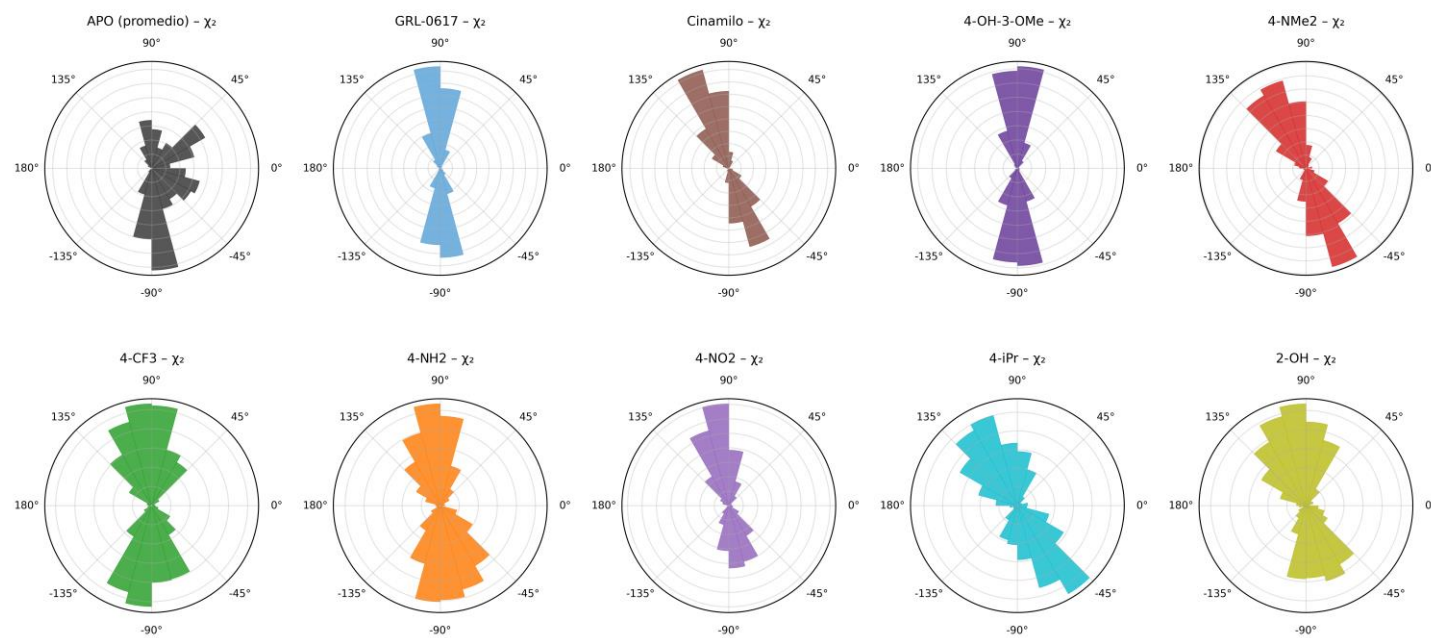


Figura 94 Distribución polar del ángulo diedro χ_2 del residuo Tyr268 durante las simulaciones de Dinámica molecular para los complejos con derivados de la familia 2-naftilo. La organización de los paneles y la definición de las curvas corresponden a las descritas en la Figura 83.

Distribución polar de χ_1 – Antraceno

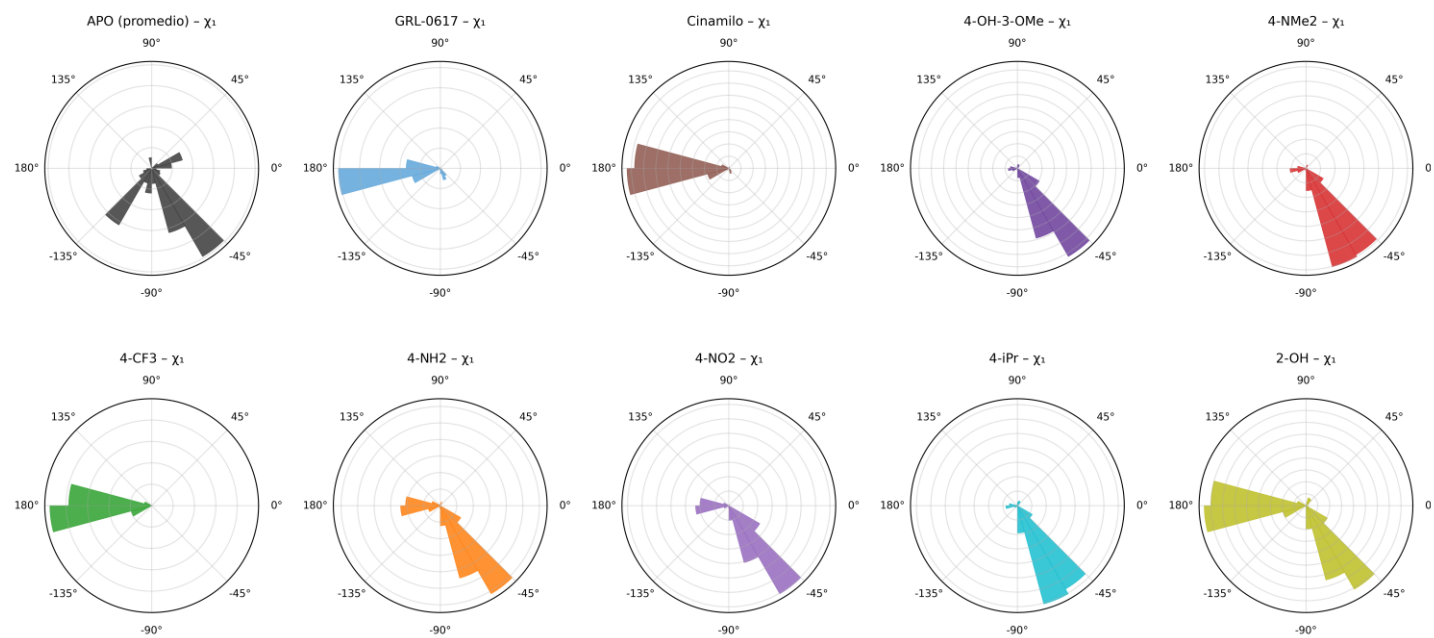


Figura 95 Distribución polar del ángulo diedro χ_1 del residuo Tyr268 durante las simulaciones de Dinámica molecular para los complejos con derivados de la familia antraceno. La organización de los paneles y la definición de las curvas corresponden a las descritas en la Figura 83.

Distribución polar de χ_2 – Antraceno



Figura 96 Distribución polar del ángulo diedro χ_2 del residuo Tyr268 durante las simulaciones de Dinámica molecular para los complejos con derivados de la familia antraceno. La organización de los paneles y la definición de las curvas corresponden a las descritas en la Figura 83.

Distribución polar de χ_1 – Pireno

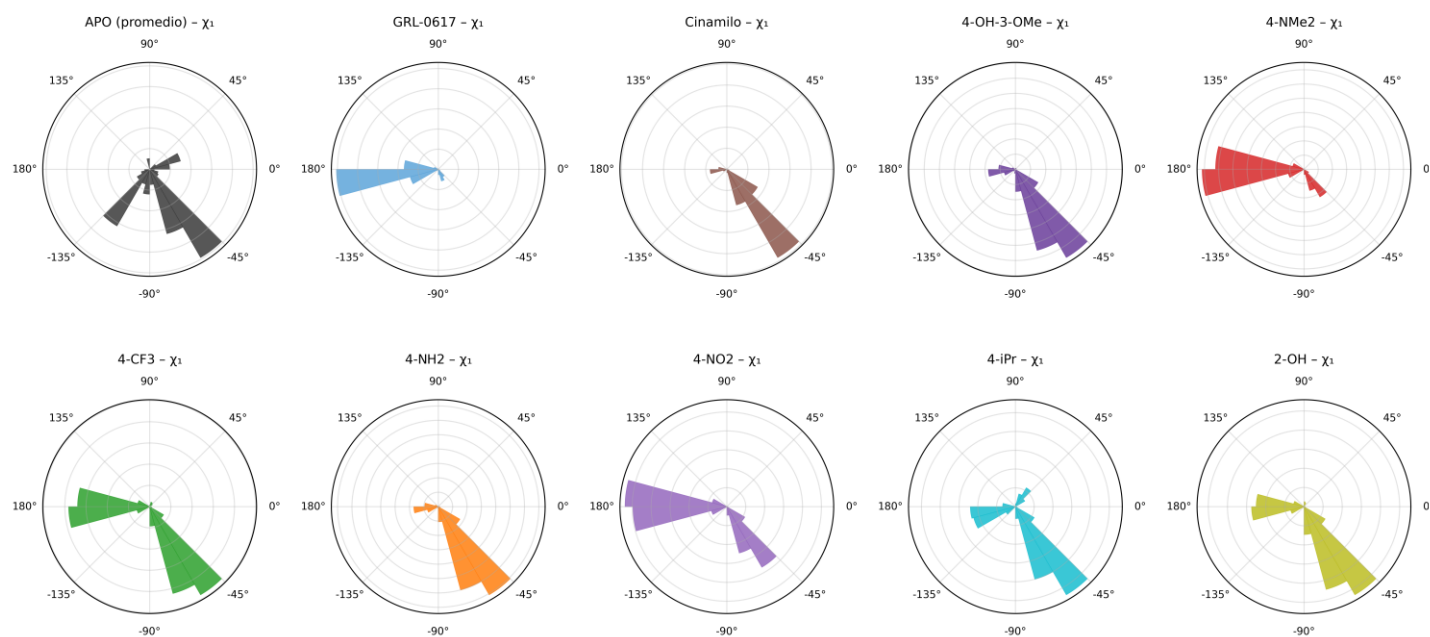


Figura 97 Distribución polar del ángulo diedro χ_1 del residuo Tyr268 durante las simulaciones de Dinámica molecular para los complejos con derivados de la familia pireno. La organización de los paneles y la definición de las curvas corresponden a las descritas en la Figura 83.

Distribución polar de χ_2 – Pireno

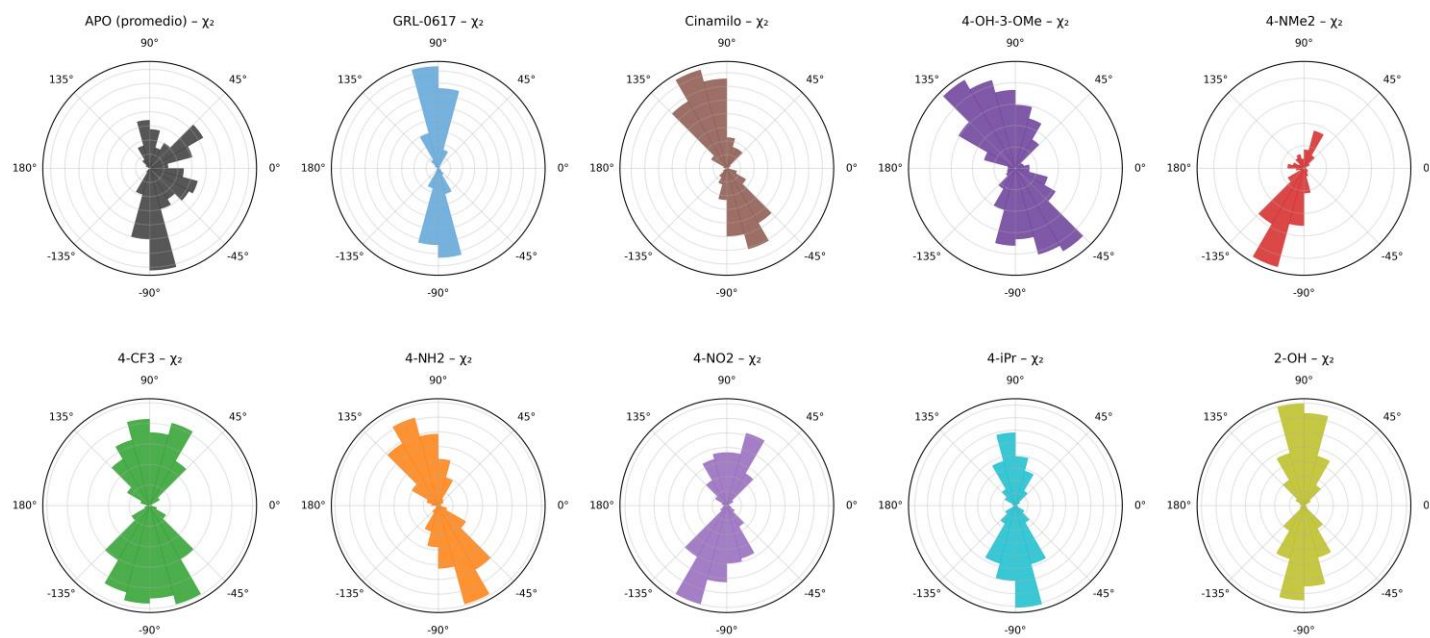


Figura 98 Distribución polar del ángulo diedro χ_2 del residuo Tyr268 durante las simulaciones de Dinámica molecular para los complejos con derivados de la familia pireno. La organización de los paneles y la definición de las curvas corresponden a las descritas en la Figura 83.

5.5.4 Top 2 mejores moléculas de cada familia

Las dos moléculas con mejor puntaje de docking dentro de cada familia (Figura 99) fueron seleccionadas para los análisis posteriores, con excepción de las dihidrobenzodiazepinas y dihidropirazoles, cuya exclusión se discute en el apartado de discusión de esta sección.

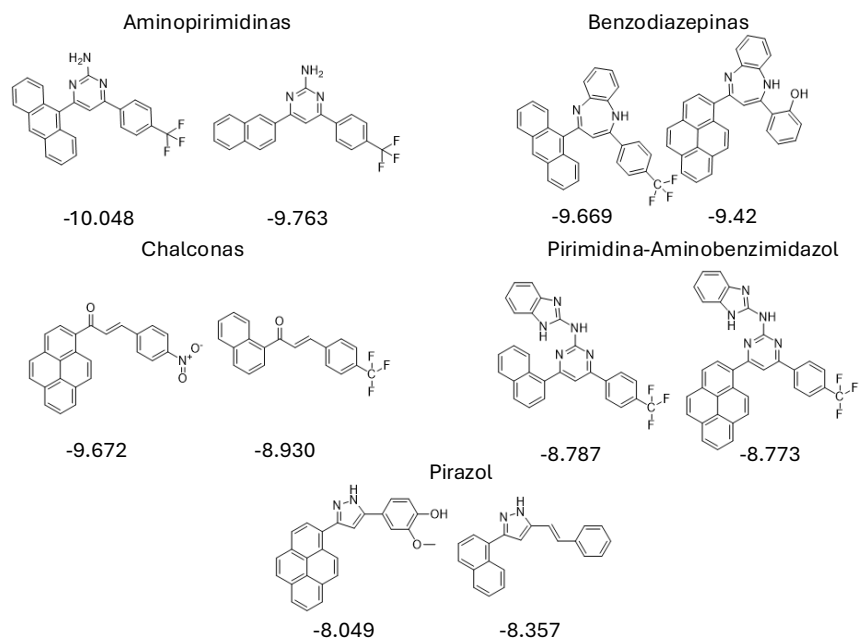
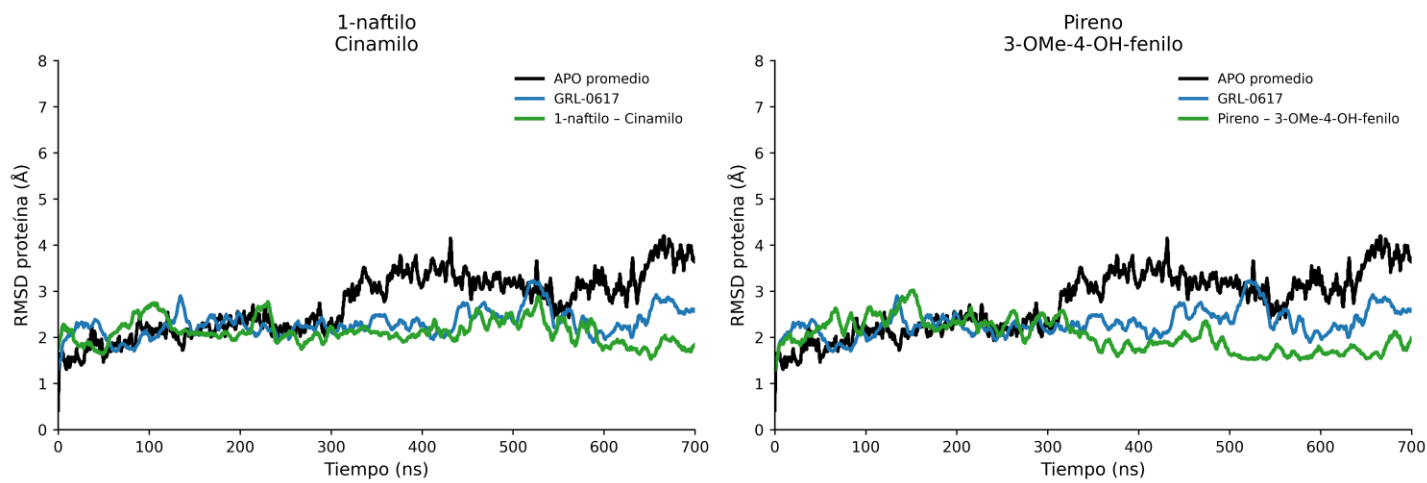


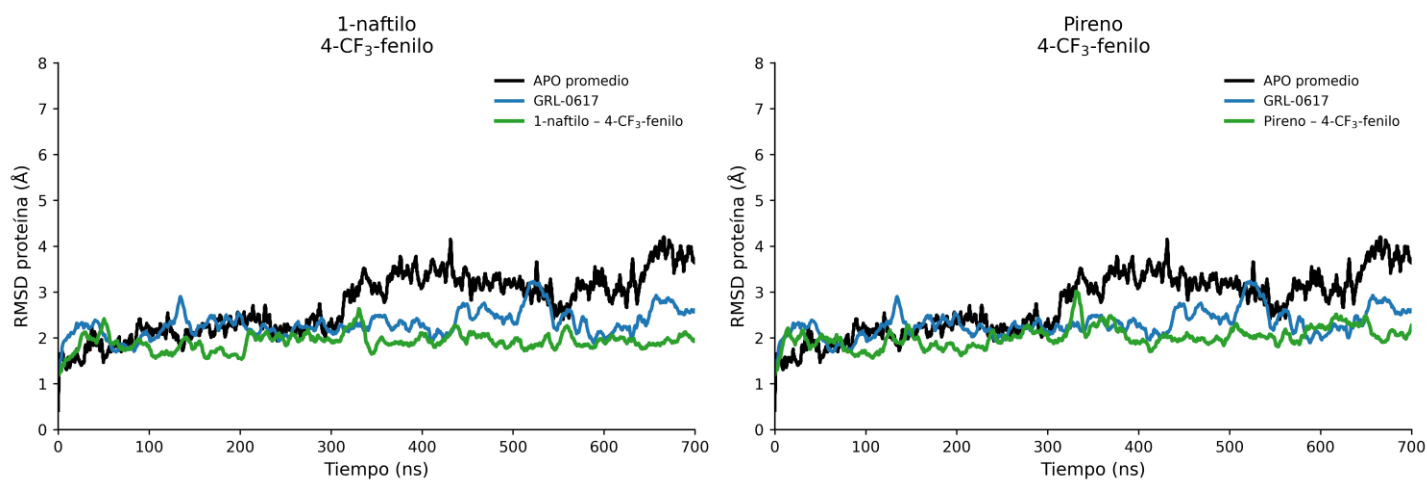
Figura 99 Estructuras representativas de cada familia junto con sus dos compuestos mejor posicionados según el puntaje de docking (XP GScore).

Los resultado de su RMSD de proteínas se ven en la Figura 100 . De manera consistente, todos los complejos proteína-ligando evaluados presentan valores de RMSD comparables o inferiores a los observados para GRL-0617, manteniéndose predominantemente en el rango de 2–3 Å durante la mayor parte de la simulación.

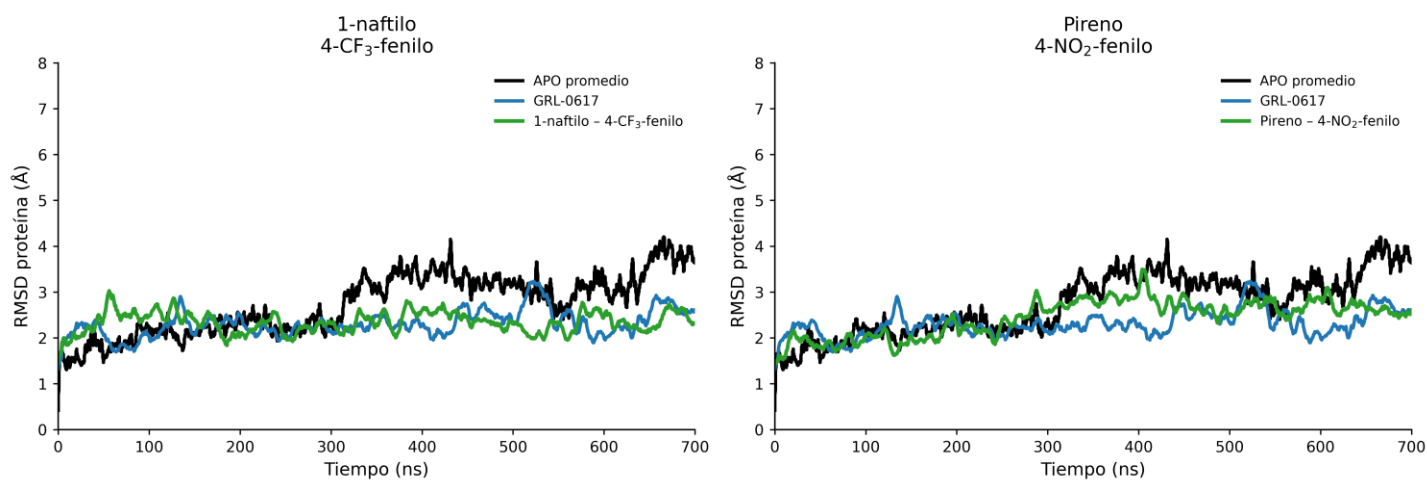
RMSD de la proteína – Pirazol



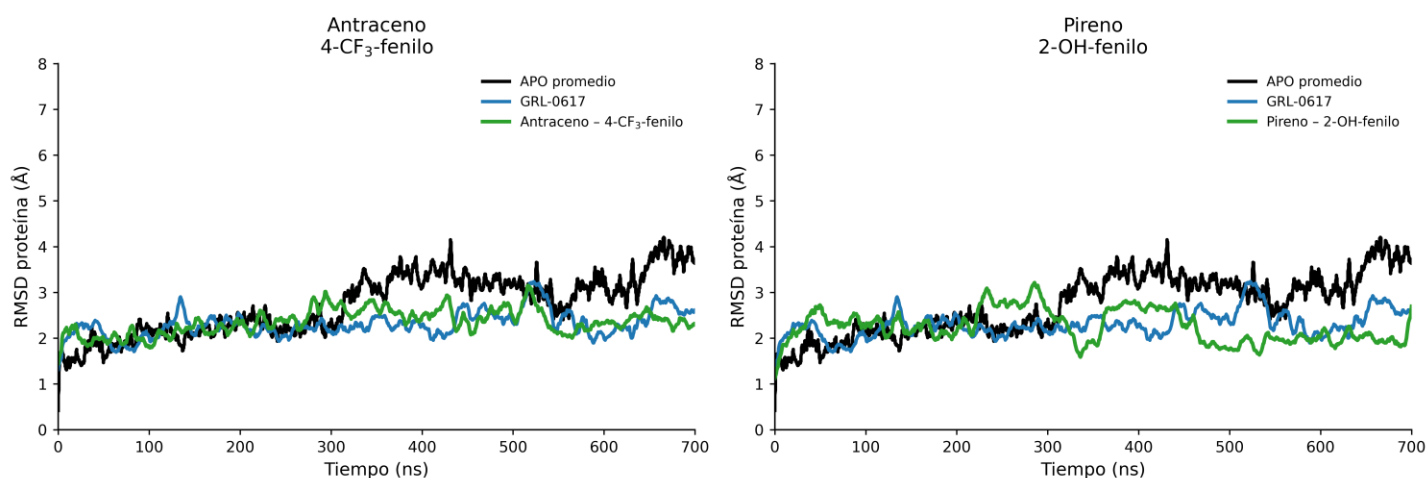
RMSD de la proteína – Pirimidina aminobenzimidazol



RMSD de la proteína – Chalcona



RMSD de la proteína – Benzodiazepinas



RMSD de la proteína – Aminopirimidina

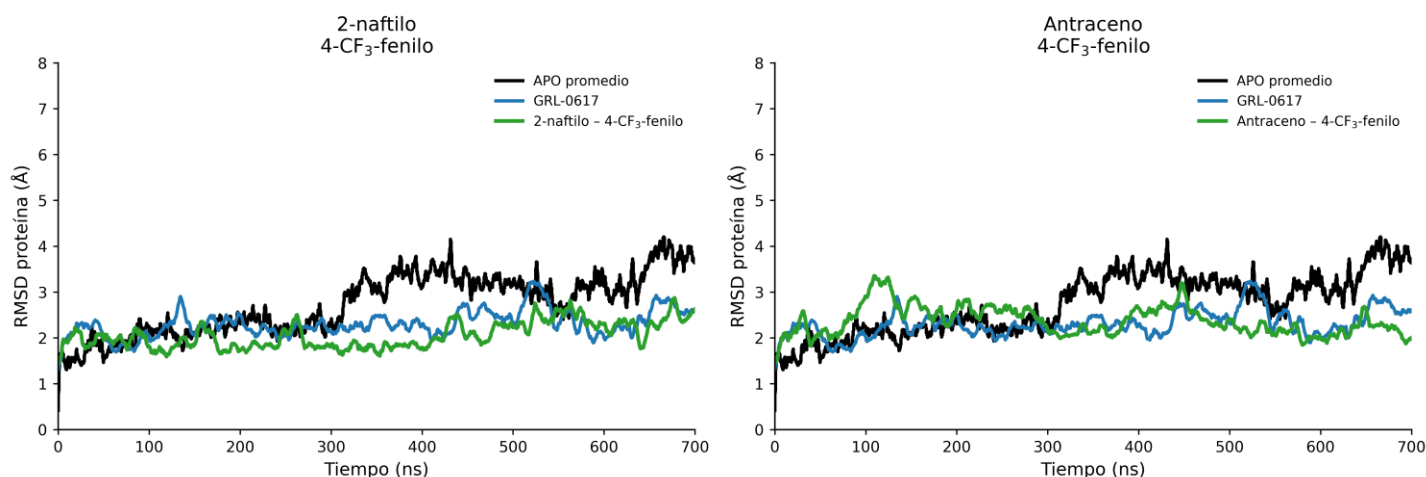


Figura 100 RMSD de la proteína a lo largo de las simulaciones de Dinámica molecular para los complejos con los ligandos evaluados de las cinco familias estructurales. La línea negra representa el sistema APO promedio, la línea azul el complejo con el inhibidor de referencia GRL-0617 y la línea verde el complejo con el ligando analizado en cada panel.

Los perfiles de RMSD del ligando durante las simulaciones de Dinámica molecular se muestran en la Figura 101. En el caso de las aminopirimidinas, el derivado 2-naftilo-4- CF_3 -fenilo se mantiene relativamente estable durante aproximadamente los primeros 400 ns, tras lo cual se observa un aumento abrupto en el RMSD que indica una migración del ligando fuera de la posición inicial. En contraste, el derivado antraceno-4- CF_3 -fenilo presenta una alta variabilidad conformacional durante la simulación, con fluctuaciones amplias del RMSD.

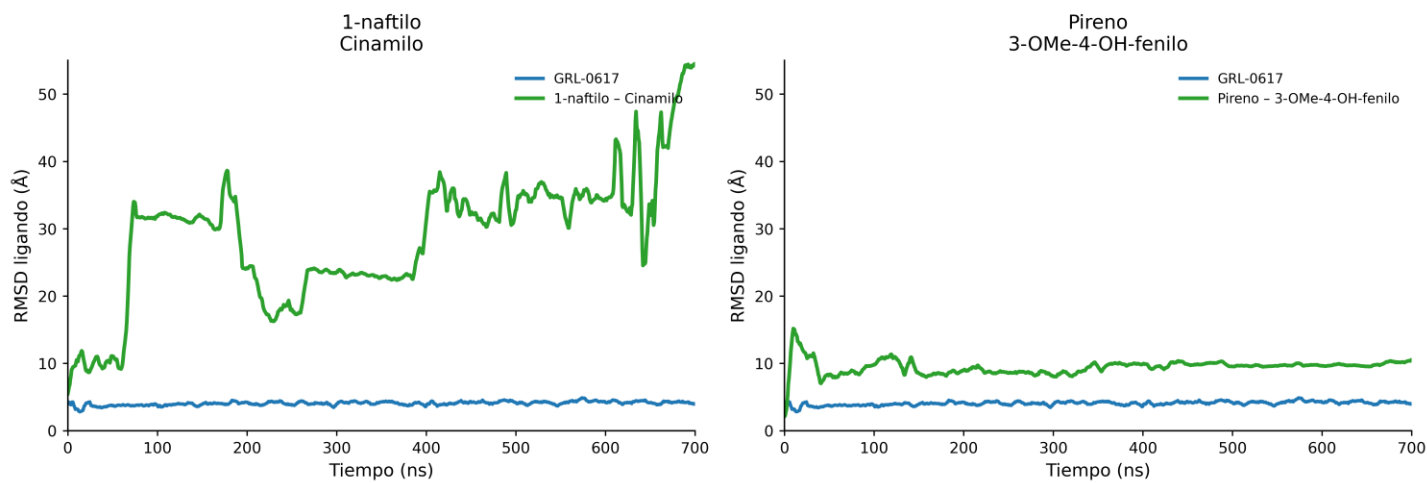
Para la familia de benzodiazepinas, ambos ligandos muestran un comportamiento relativamente estable a lo largo de la simulación. En particular, el derivado pireno-2-OH-fenilo presenta un perfil de RMSD muy similar al observado para el ligando de referencia GRL-0617, manteniéndose en valores cercanos a $\sim 4\text{--}5$ Å durante toda la trayectoria.

En el caso de las chalconas, se observa en general una mayor desviación estructural respecto a la conformación inicial. El derivado 1-naftilo-4- CF_3 -fenilo presenta un incremento abrupto del RMSD seguido de una estabilización en valores elevados (~ 30 Å), mientras que el derivado pireno-4- NO_2 -fenilo muestra fluctuaciones moderadas con valores cercanos a $15\text{--}18$ Å durante la mayor parte de la simulación.

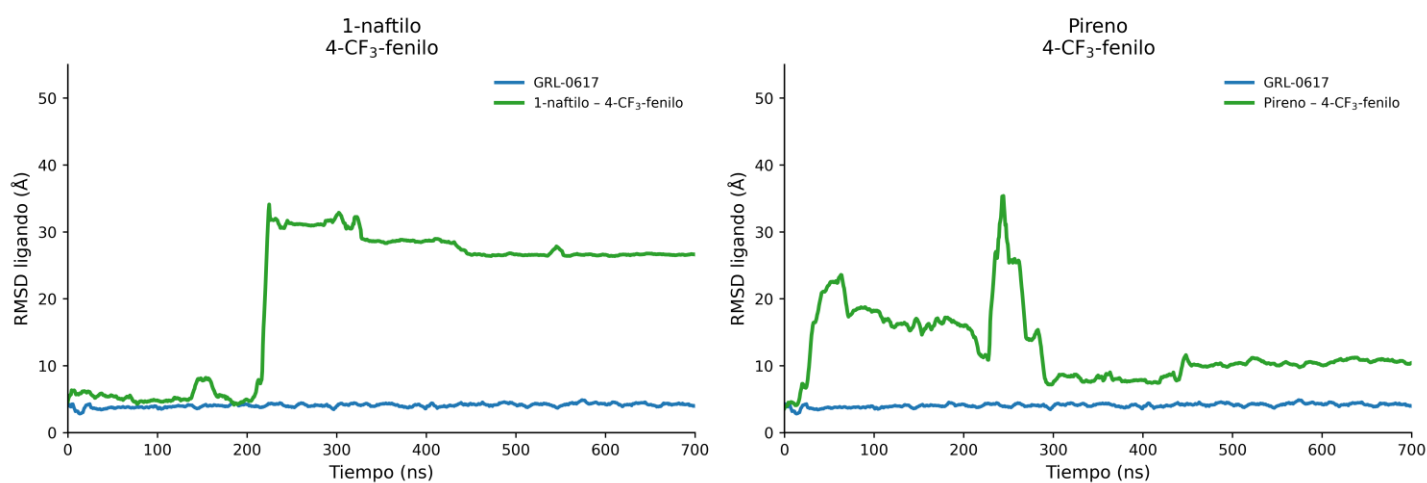
Para la familia pirimidina-aminobenzimidazol, el derivado 1-naftilo-4- CF_3 -fenilo se mantiene estable durante aproximadamente los primeros 200 ns, tras lo cual el RMSD aumenta bruscamente y se estabiliza posteriormente en valores cercanos a $27\text{--}30$ Å. Por su parte, el derivado pireno-4- CF_3 -fenilo presenta una alta variabilidad en el RMSD durante la simulación, con episodios de aumento abrupto seguidos de una posterior estabilización.

Finalmente, en la familia pirazol, el derivado 1-naftilo-cinamilo muestra una elevada desviación estructural y una marcada variabilidad durante toda la simulación, alcanzando valores superiores a 40 Å. En contraste, el derivado pireno-3-OMe-4-OH-fenilo presenta un comportamiento más estable, manteniendo valores de RMSD cercanos a $8\text{--}10$ Å a lo largo de la trayectoria.

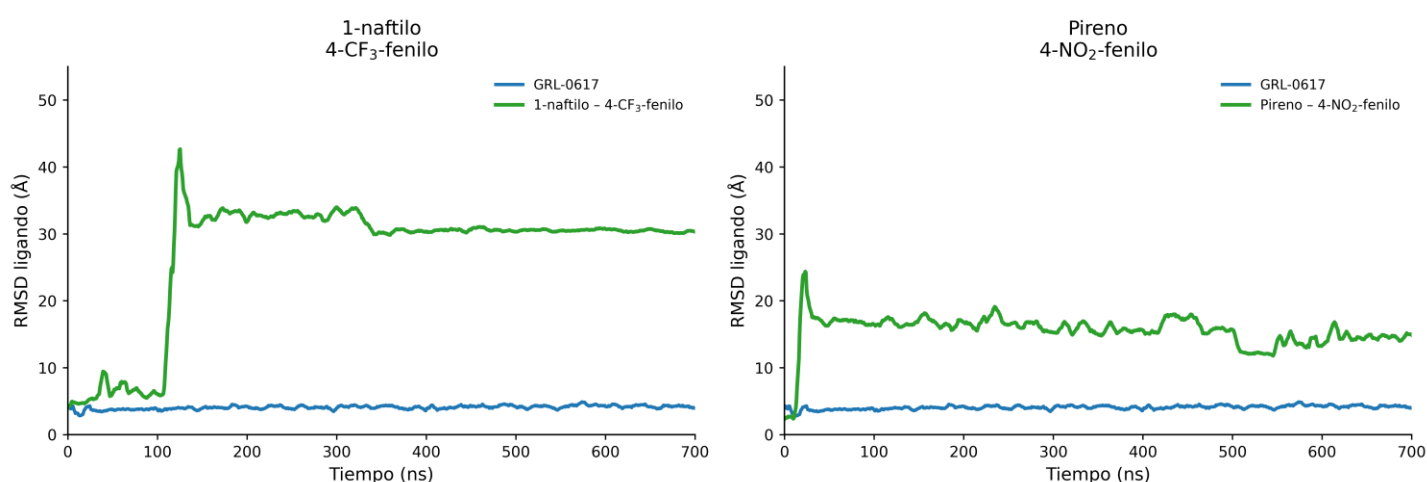
RMSD del ligando - Pirazol



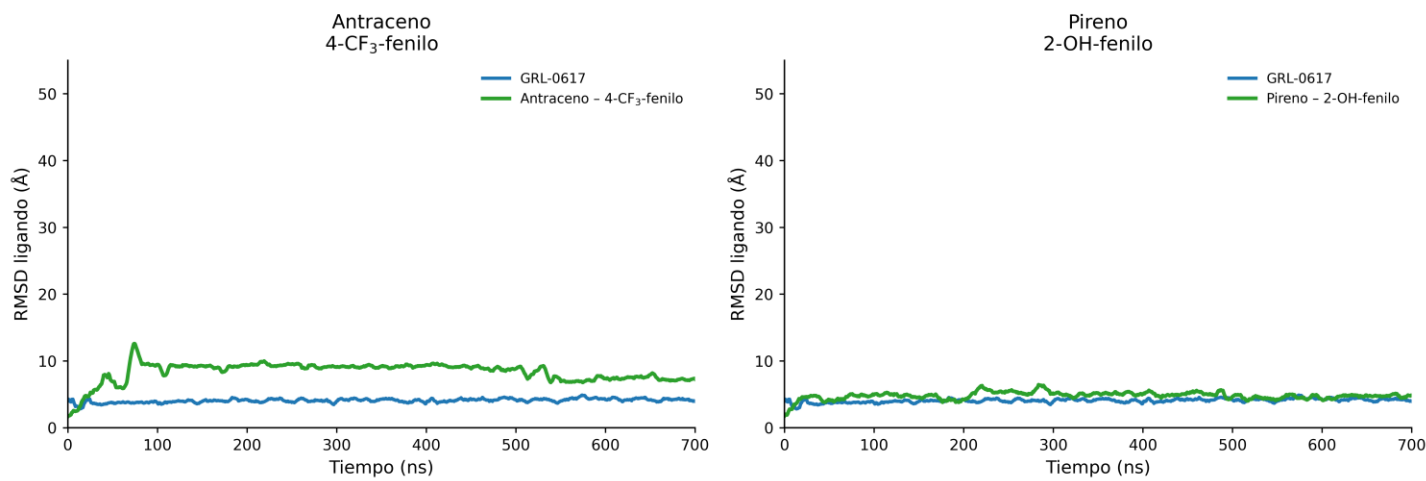
RMSD del ligando - Pirimidina aminobenzimidazol



RMSD del ligando - Chalcona



RMSD del ligando – Benzodiazepinas



RMSD del ligando – Aminopirimidina

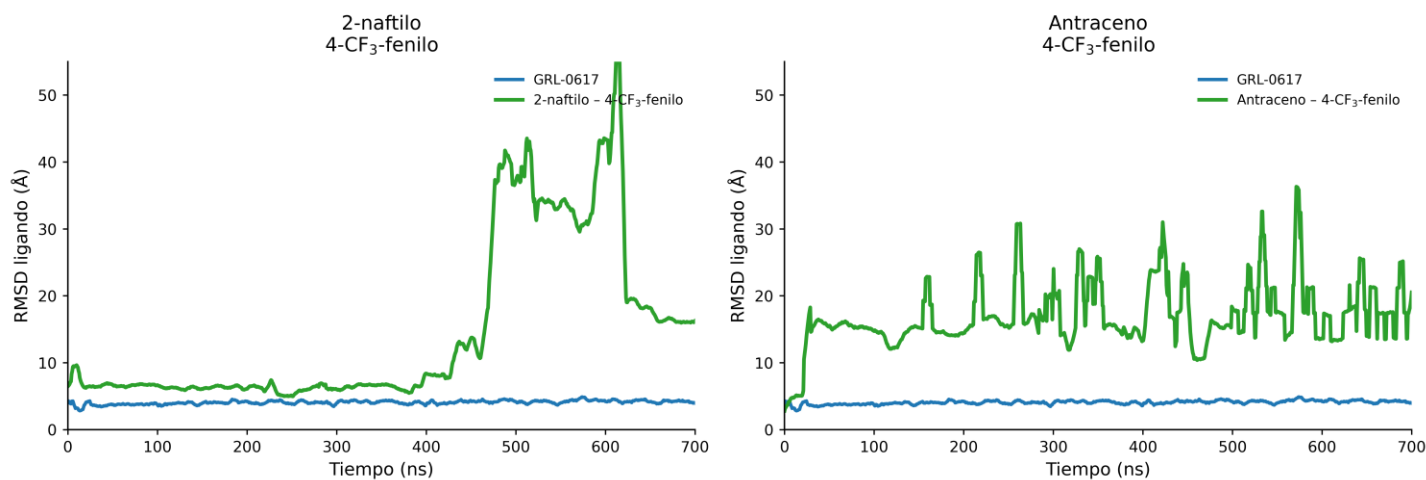


Figura 101 RMSD del ligando durante las simulaciones de Dinámica molecular para los complejos correspondientes a la familia pirazol. La línea azul representa el inhibidor de referencia GRL-0617 y la línea verde el ligando evaluado en cada panel. Los gráficos muestran la evolución temporal de la desviación estructural del ligando respecto a su conformación inicial durante la trayectoria de simulación. Los perfiles de RMSF de la proteína durante las simulaciones de Dinámica molecular se muestran en la Figura 102, mientras que un acercamiento a la región del loop BL2 (residuos 255–275) se presenta en la Figura 103. En todas las simulaciones, el sistema APO presenta las mayores fluctuaciones en la región

correspondiente al loop BL2, con un máximo alrededor del residuo Tyr268, mientras que el complejo con el inhibidor de referencia GRL-0617 muestra valores de RMSF menores en esta región.

En la familia de aminopirimidinas, el derivado 2-naftilo-4-CF₃-fenilo presenta un perfil de RMSF similar al sistema APO en la región BL2, mientras que el derivado antraceno-4-CF₃-fenilo muestra valores ligeramente menores de RMSF en comparación con APO.

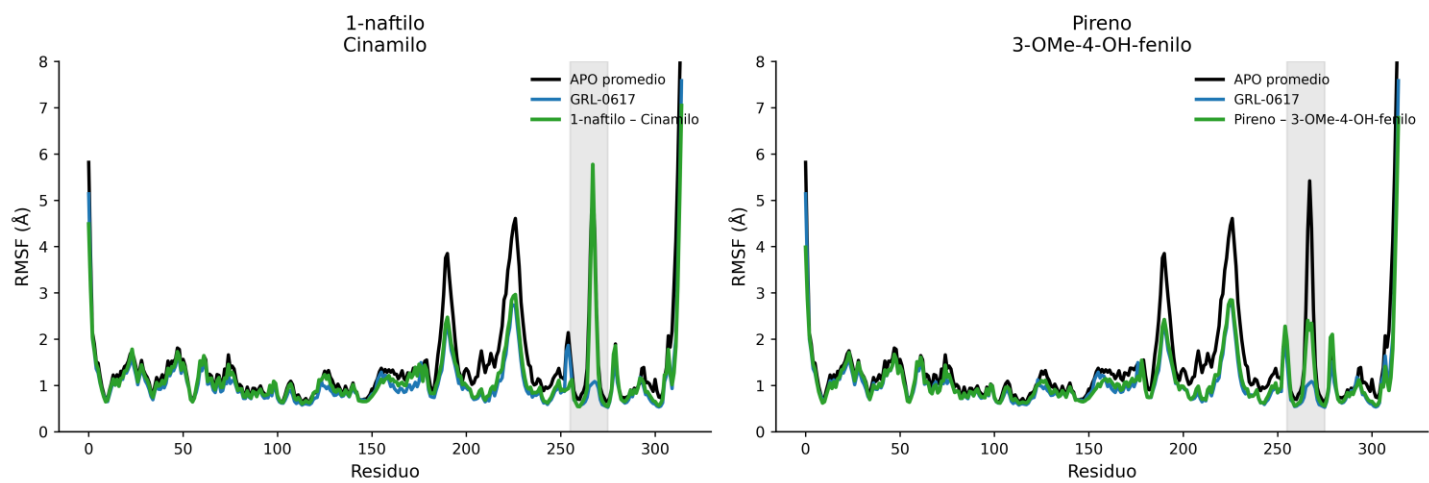
Para la familia de benzodiazepinas, ambos derivados evaluados (antraceno-4-CF₃-fenilo y pireno-2-OH-fenilo) presentan valores de RMSF menores que el sistema APO en la región BL2. En ambos casos, los valores observados son cercanos a los del complejo con GRL-0617.

En el caso de las chalconas, los derivados 1-naftilo-4-CF₃-fenilo y pireno-4-NO₂-fenilo muestran perfiles de RMSF en la región BL2 similares a los observados para el sistema APO, con diferencias mínimas en la amplitud de las fluctuaciones.

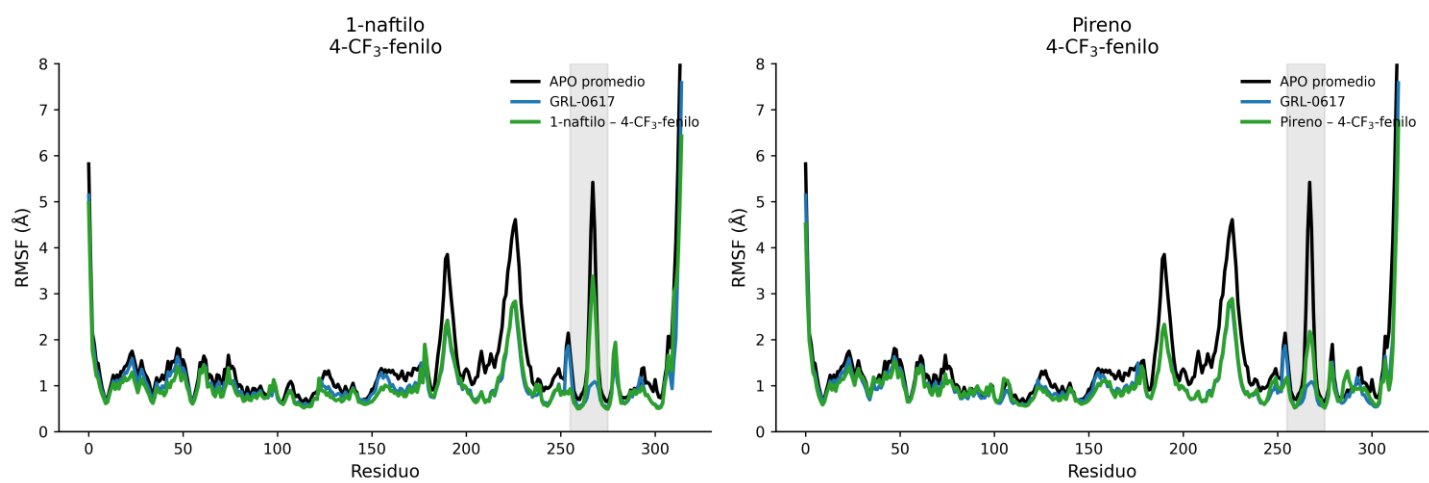
Para la familia pirimidina-aminobenzimidazol, el derivado 1-naftilo-4-CF₃-fenilo presenta una disminución parcial del RMSF en la región BL2 en comparación con APO, mientras que el derivado pireno-4-CF₃-fenilo muestra valores de RMSF menores en esta región.

Finalmente, en la familia de pirazoles, el derivado 1-naftilo-cinamilo presenta valores de RMSF similares a los del sistema APO en el loop BL2, mientras que el derivado pireno-3-OMe-4-OH-fenilo muestra valores menores de RMSF en esta región.

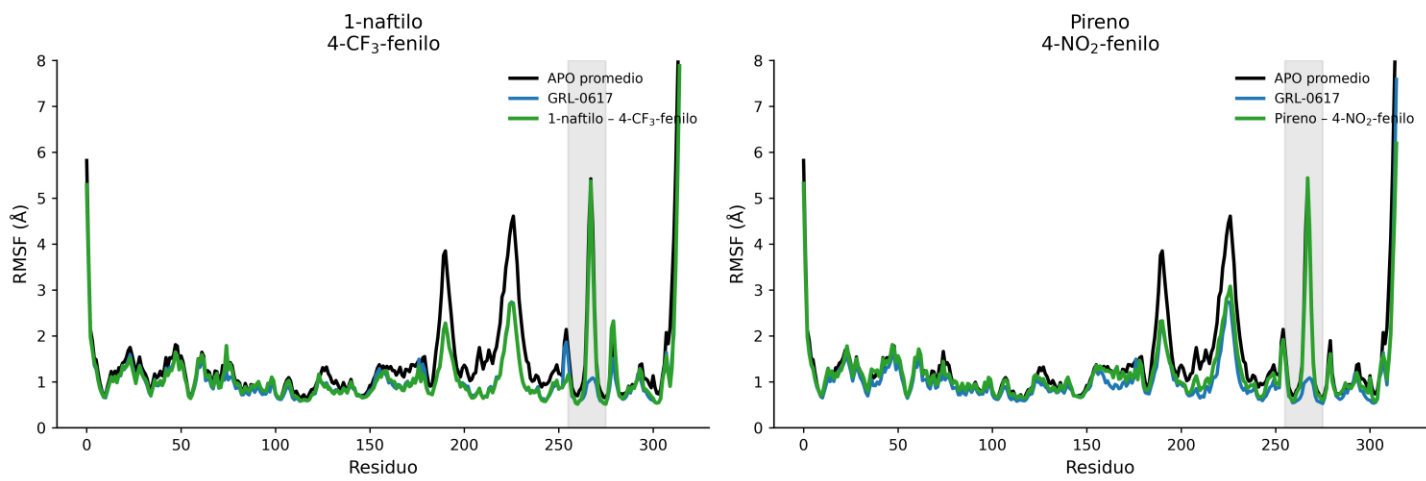
RMSF completo – Pirazol



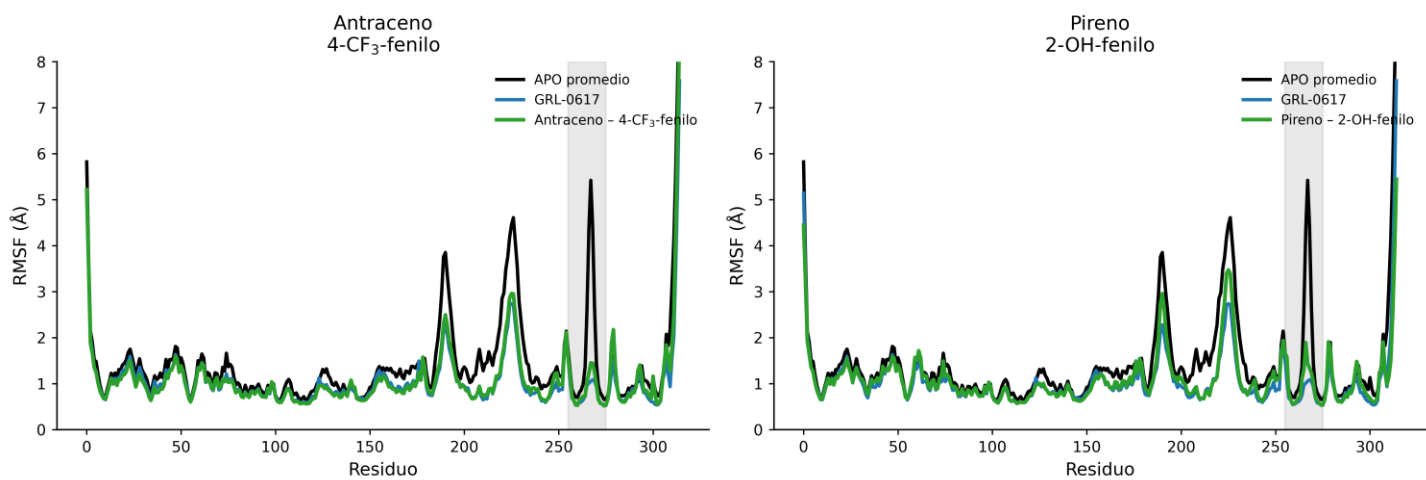
RMSF completo – Pirimidina aminobenzimidazol



RMSF completo – Chalcona



RMSF completo – Benzodiazepinas



RMSF completo – Aminopirimidina

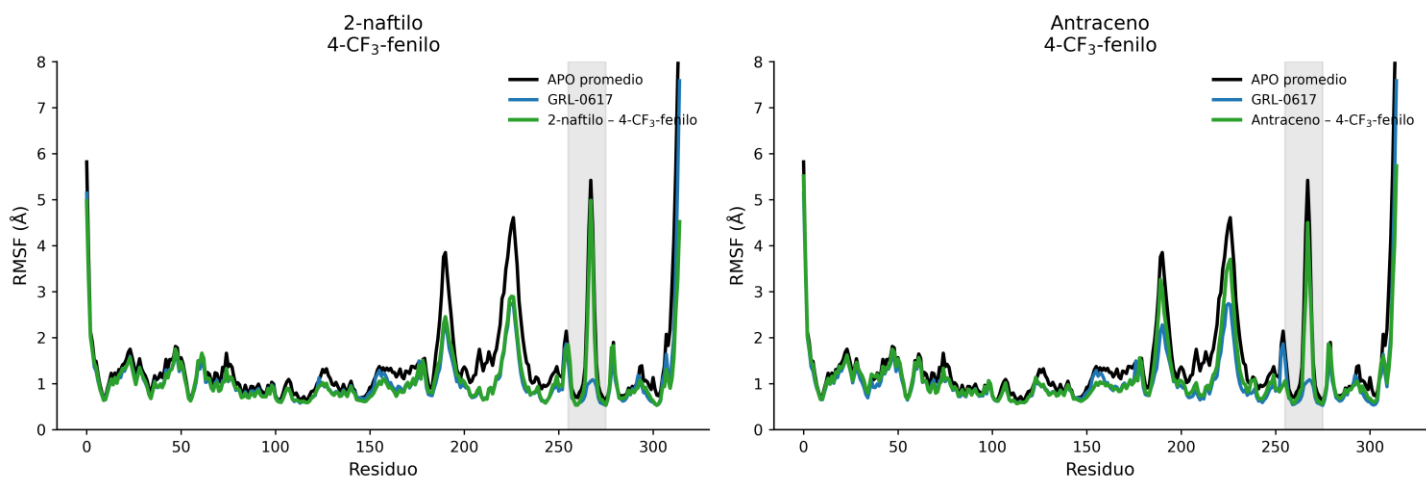
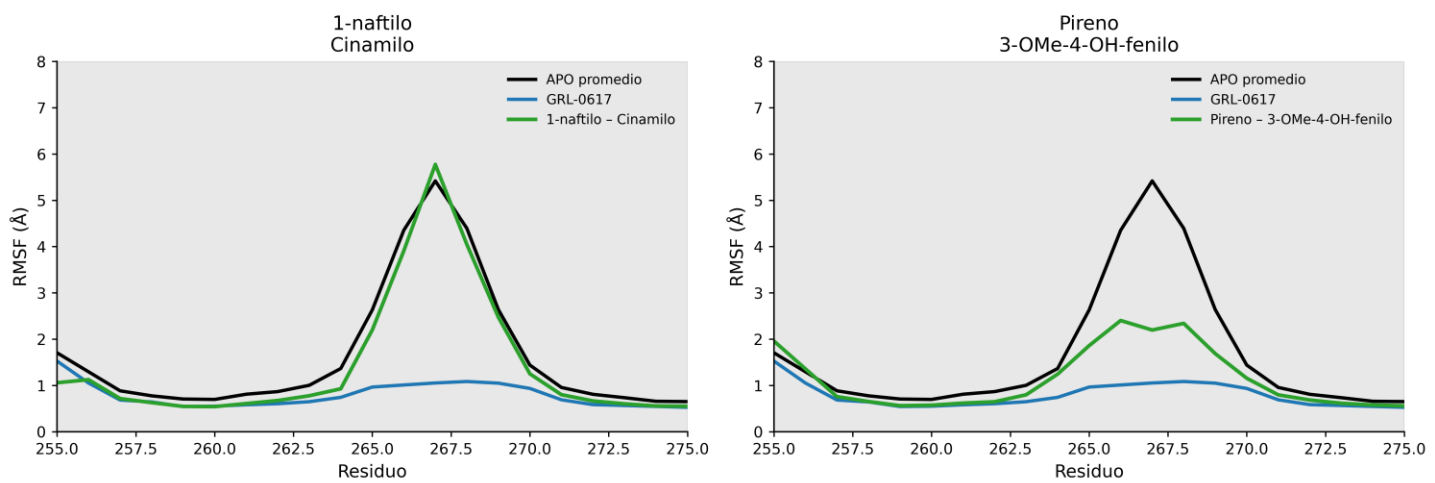
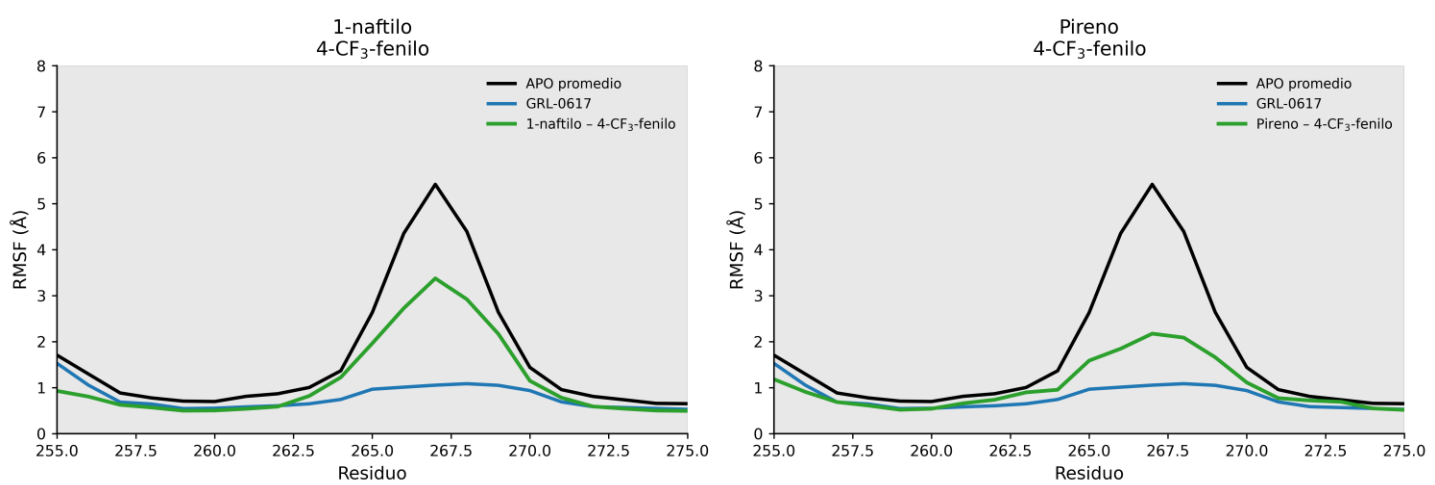


Figura 102 RMSF de la proteína durante las simulaciones de dinámica molecular para los complejos correspondientes a la familia pirazol. La línea negra corresponde al sistema APO promedio, la línea azul al complejo con el inhibidor de referencia GRL-0617 y la línea verde al complejo proteína-ligando correspondiente a cada derivado evaluado. Sombreado en gris correspondiente a loop BL2 (255–275).

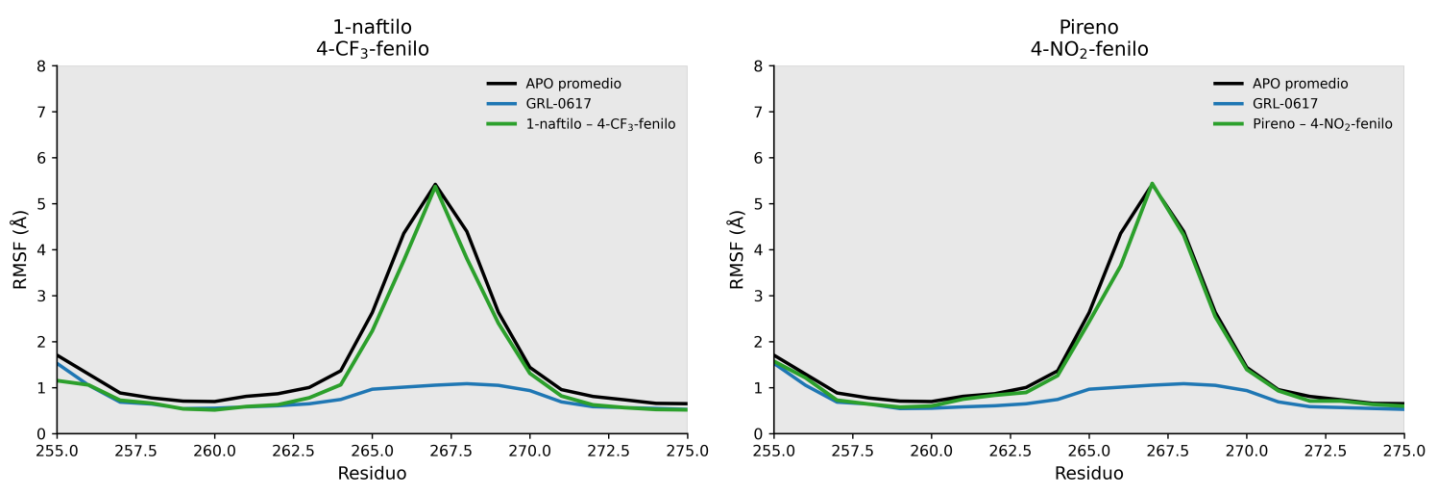
RMSF BL2 (255-275) - Pirazol



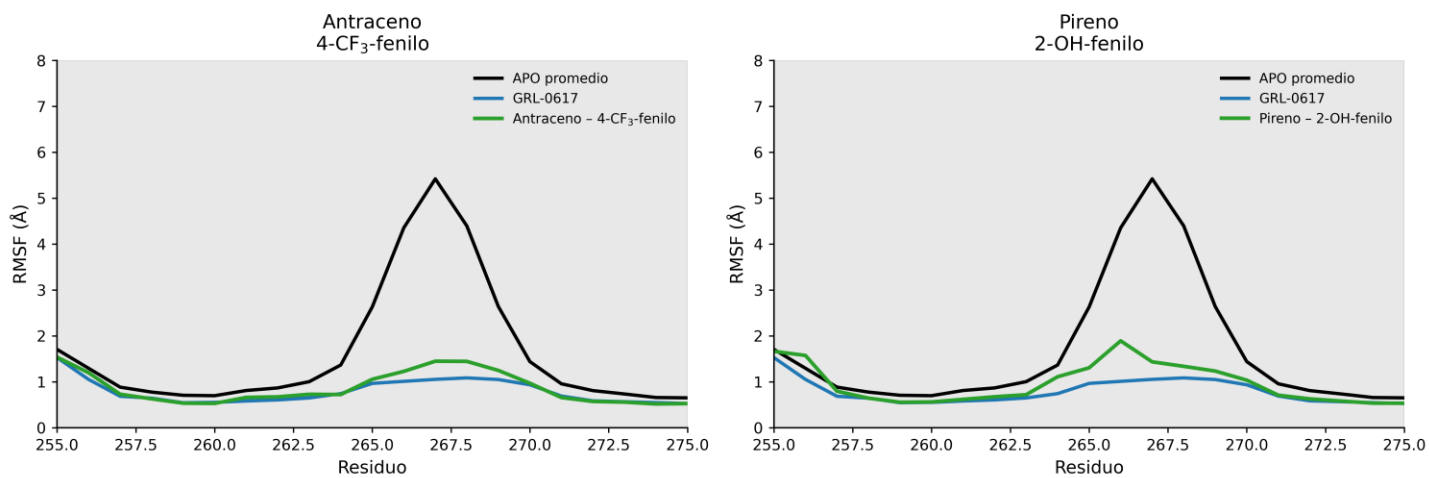
RMSF BL2 (255-275) - Pirimidina aminobenzimidazol



RMSF BL2 (255-275) - Chalcona



RMSF BL2 (255–275) – Benzodiazepinas



RMSF BL2 (255–275) – Aminopirimidina

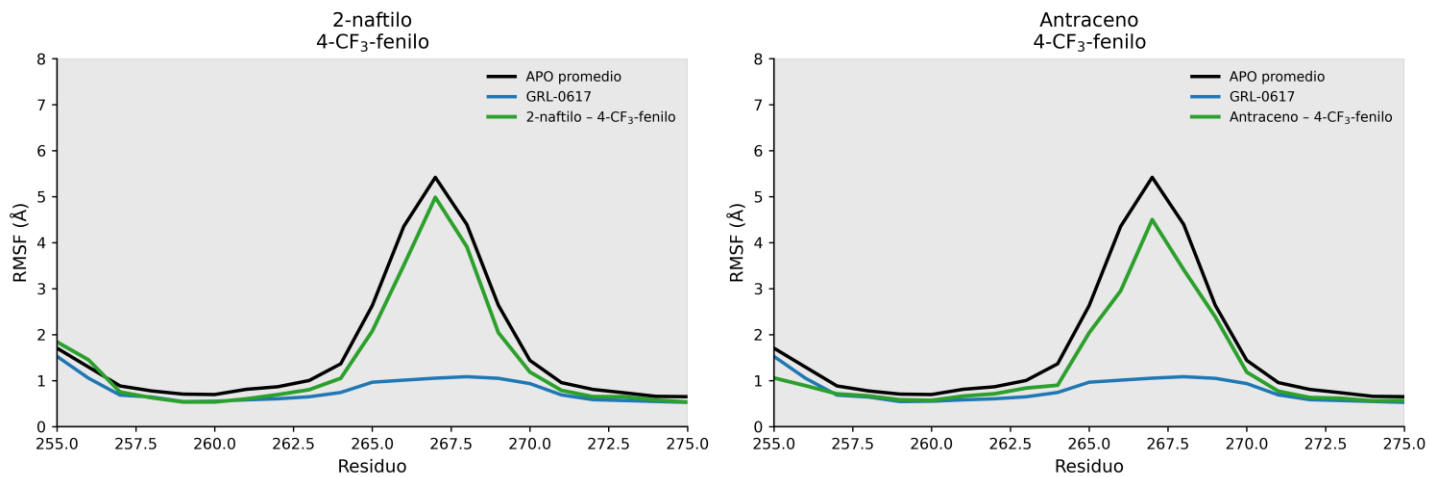


Figura 103 RMSF de los residuos correspondientes al loop BL2 (255–275) durante las simulaciones de Dinámica molecular para los complejos de la familia pirazol. La línea negra representa el sistema APO promedio, la línea azul el complejo con el inhibidor de referencia GRL-0617 y la línea verde el complejo con el ligando evaluado en cada panel.

Las distancias estructurales monitoreadas durante las simulaciones de Dinámica molecular se muestran en la Figura 104. En todos los sistemas analizados, la distancia catalítica C111–H272 se mantiene estable a lo largo de la simulación, fluctuando principalmente alrededor de 6–7 Å, sin diferencias significativas entre los complejos proteína-ligando y los sistemas de referencia (APO y GRL-0617).

Para la familia aminopirimidina, ambos derivados evaluados mantienen valores de la distancia catalítica comparables a los sistemas de referencia durante toda la trayectoria. En cuanto a las distancias asociadas al movimiento del loop BL2, el derivado antraceno-4-CF₃-fenilo muestra valores similares a los observados en el sistema APO, mientras que el derivado 2-naftilo-4-CF₃-fenilo presenta variaciones moderadas en la distancia Tyr268–Arg166, manteniéndose dentro del rango observado para los sistemas de referencia.

En el caso de las benzodiazepinas, la distancia catalítica permanece estable en ambos complejos durante toda la simulación. Las distancias relacionadas con el loop BL2 muestran fluctuaciones comparables a las observadas en el sistema APO, sin alteraciones relevantes respecto al ligando de referencia.

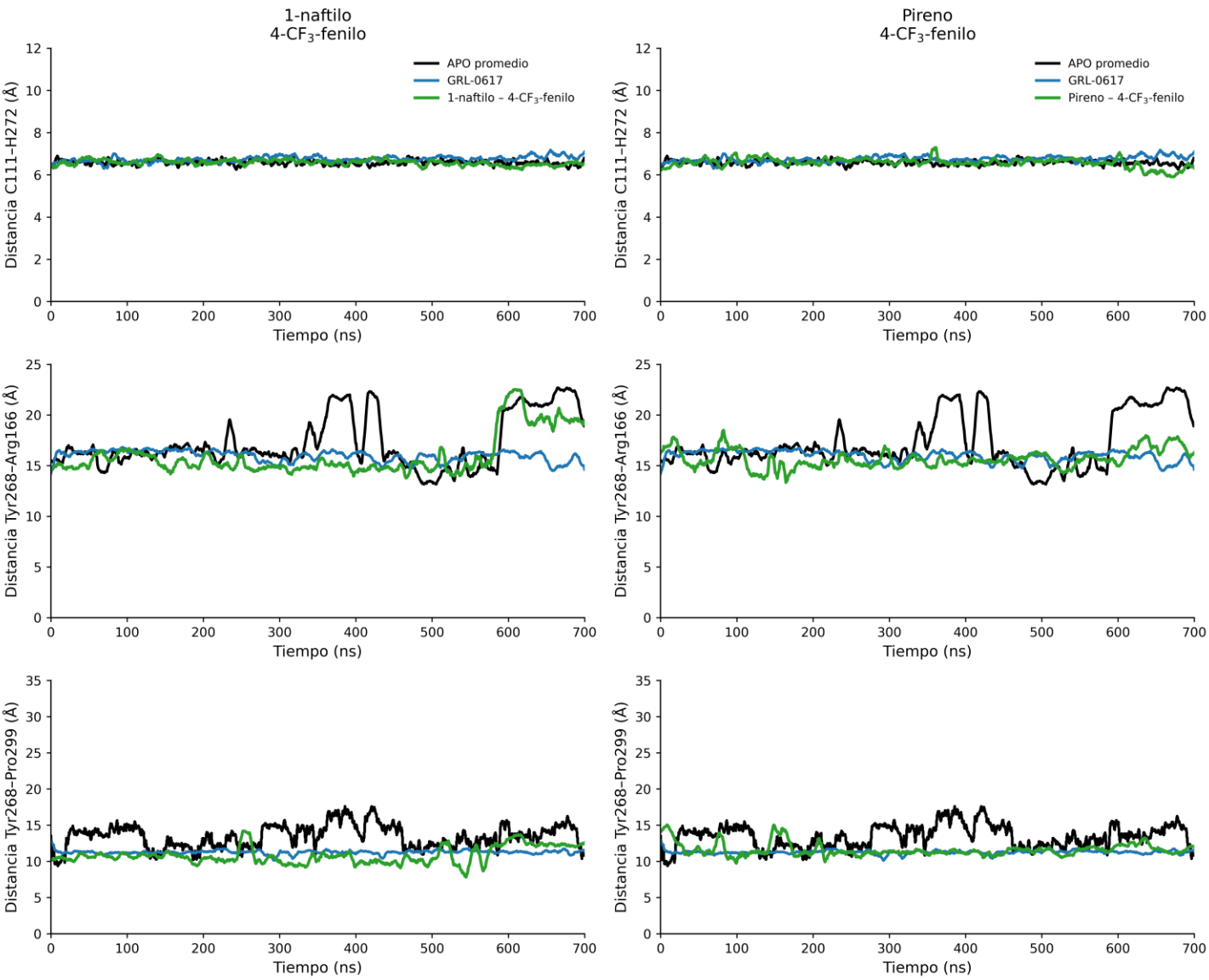
Para la familia chalcona, la distancia catalítica C111–H272 se mantiene igualmente constante durante toda la trayectoria. Sin embargo, se observan variaciones más amplias en las distancias Tyr268–Arg166 y Tyr268–Pro299, particularmente para el derivado pireno-4-NO₂-fenilo, que muestra valores superiores a los observados para el complejo con GRL-0617 durante ciertos intervalos de la simulación.

En la familia pirimidina-aminobenzimidazol, ambos ligandos mantienen la distancia catalítica estable durante toda la trayectoria. Las distancias asociadas al loop BL2 presentan fluctuaciones moderadas, especialmente en el derivado 1-naftilo-4-CF₃-fenilo, que muestra aumentos transitorios en la distancia Tyr268–Arg166 hacia la segunda mitad de la simulación.

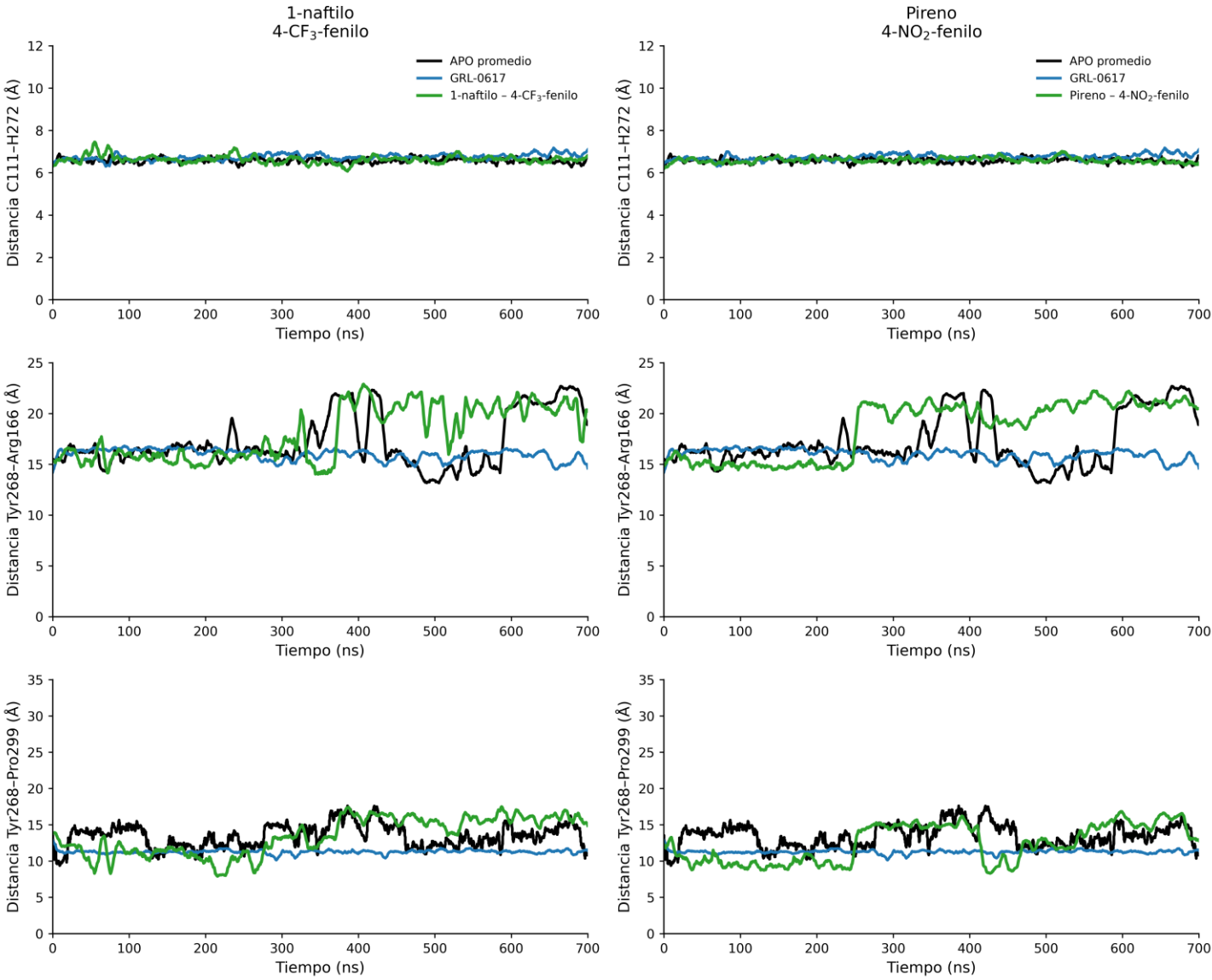
Finalmente, en la familia pirazol, la distancia catalítica permanece estable durante toda la simulación para ambos complejos. El derivado 1-naftilo-cinamilo presenta variaciones moderadas en las distancias

relacionadas con el loop BL2, mientras que el derivado pireno-3-OMe-4-OH-fenilo mantiene valores más cercanos a los observados para el ligando de referencia durante la mayor parte de la trayectoria.

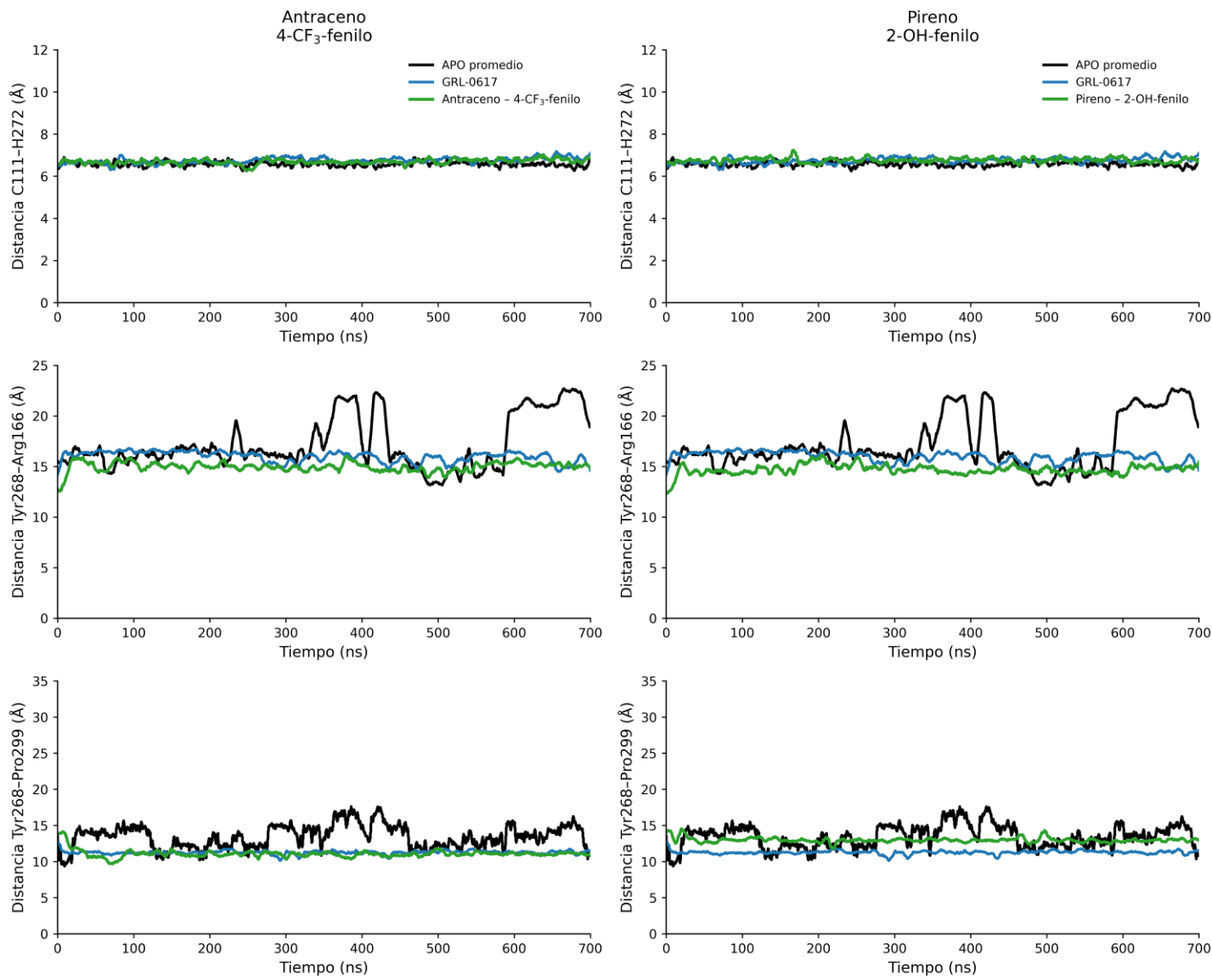
Distancias - Pirimidina aminobenzimidazol



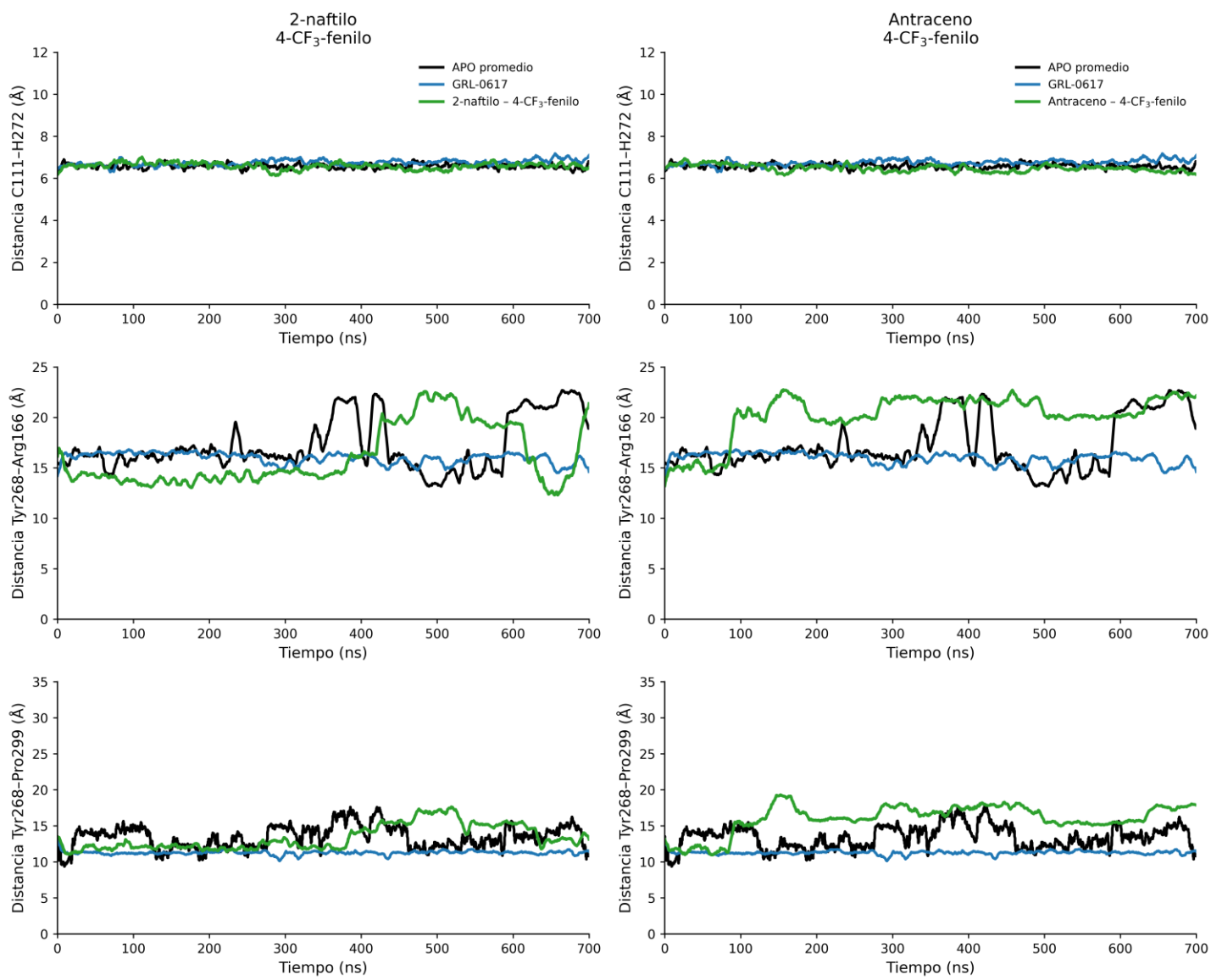
Distancias – Chalcona



Distancias - Benzodiazepinas



Distancias - Aminopirimidina



Distancias – Pirazol

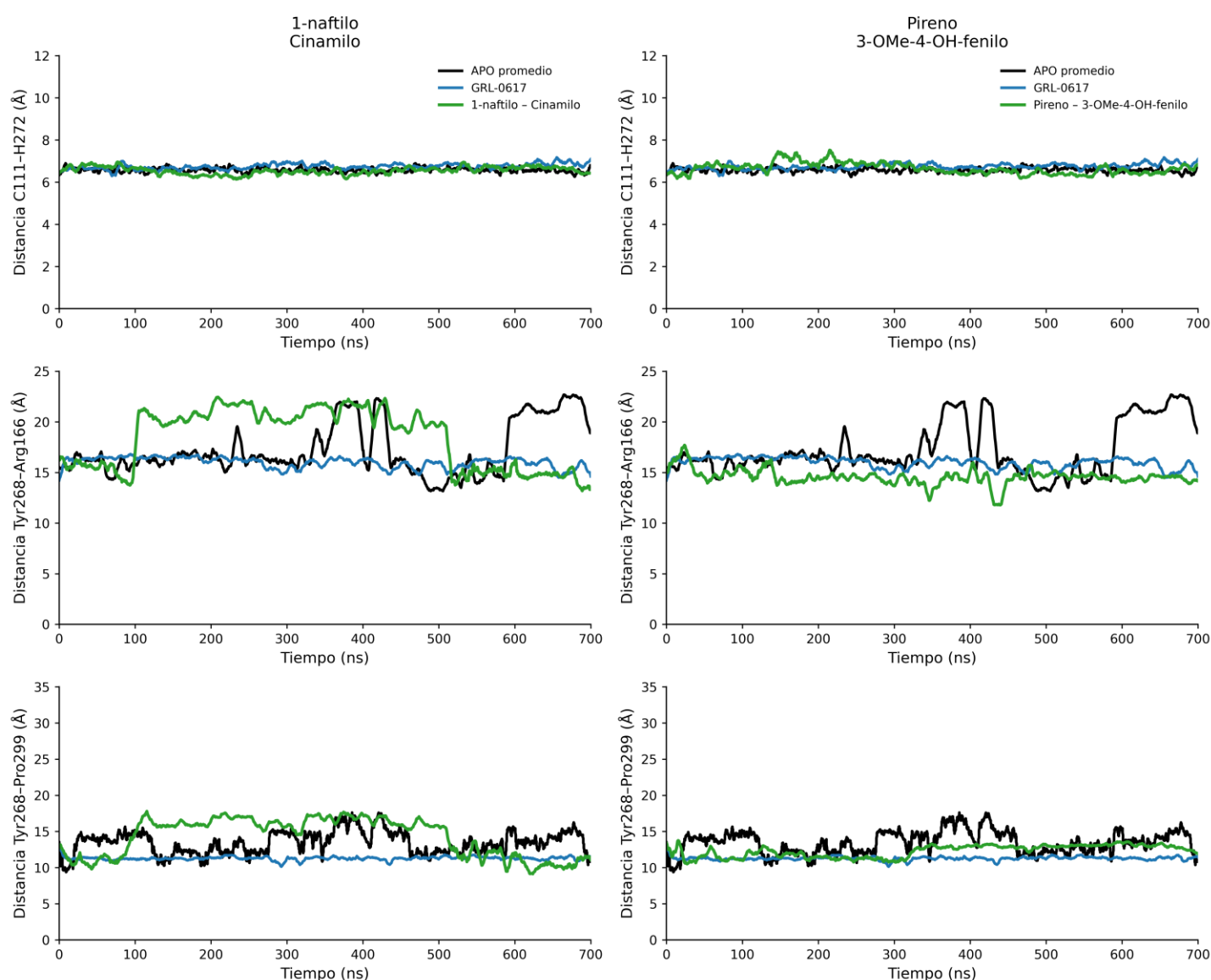


Figura 104 C111-H272, Tyr268-Arg166 y Tyr268-Pro299 durante las simulaciones de dinámica molecular para los complejos proteína-ligando de cada familia. Negro: sistema APO promedio; azul: complejo con GRL-0617; verde: complejo con el ligando evaluado.

El análisis de la distribución angular de los diedros χ_1 y χ_2 del residuo Tyr268 revela patrones conformacionales diferenciados entre las distintas familias evaluadas (Figura 105). En general, la conformación dominante del ángulo χ_1 se concentra alrededor de valores cercanos a 180° , consistentes con la orientación observada en los sistemas de referencia, lo que sugiere que la mayoría de los ligandos no inducen cambios conformacionales significativos en esta rotación lateral del residuo.

En la familia aminopirimidina, los compuestos analizados muestran distribuciones angulares de χ_1 comparables al sistema de referencia, con un máximo predominante cercano a 180° . En contraste, el ángulo χ_2 presenta una distribución más amplia, evidenciando poblaciones adicionales alrededor de -90° a -120° , lo que indica una mayor flexibilidad rotacional del anillo fenólico de Tyr268 en esta dimensión conformacional.

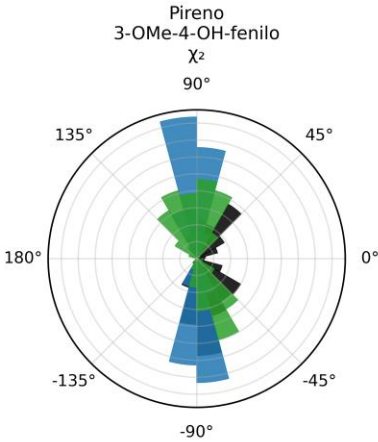
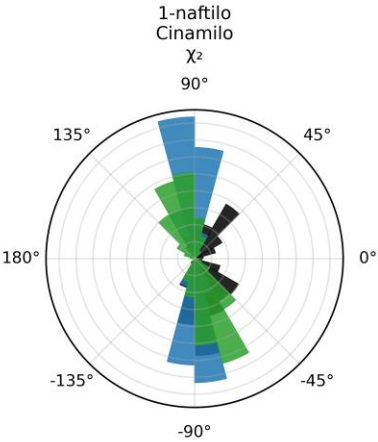
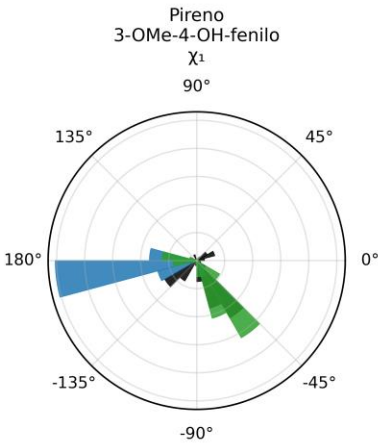
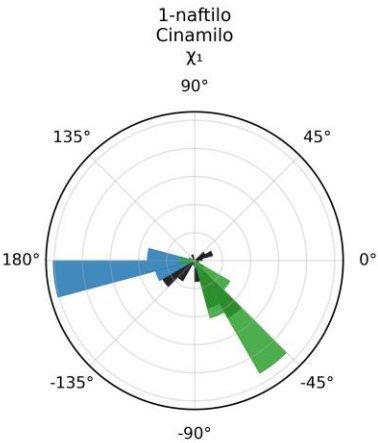
En los sistemas derivados de benzodiazepina, la distribución de χ_1 también permanece centrada alrededor de 180° , manteniendo una conformación similar a la observada en el sistema de referencia. Sin embargo, el análisis de χ_2 muestra poblaciones bien definidas en regiones angulares cercanas a $\pm 120^\circ$, sugiriendo la coexistencia de múltiples estados conformacionales accesibles durante la simulación.

Las chalconas presentan patrones conformacionales altamente similares a los sistemas de referencia tanto para χ_1 como para χ_2 . En particular, χ_1 permanece fuertemente concentrado alrededor de 180° , mientras que χ_2 muestra poblaciones dominantes alrededor de 90° y -90° , indicando que la dinámica del residuo Tyr268 permanece mayormente conservada en presencia de estos ligandos.

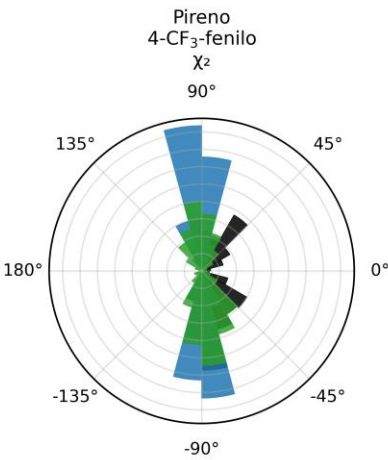
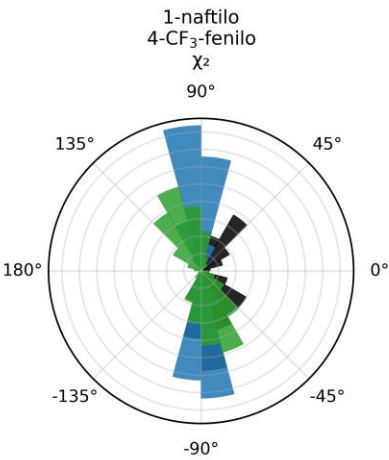
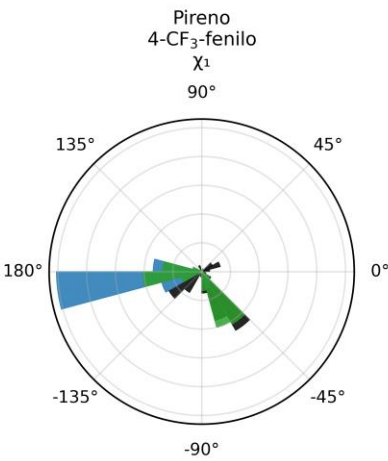
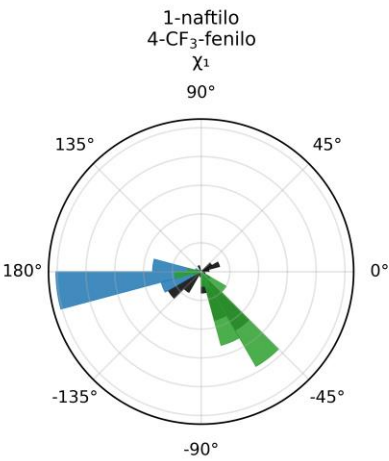
En la familia pirimidina-aminobenzimidazol, la distribución angular de χ_1 continúa mostrando un máximo predominante en 180° , con escasa dispersión hacia otras regiones angulares. Para χ_2 se observa

nuevamente una distribución bimodal alrededor de $\pm 120^\circ$, similar a la observada en otras familias, lo que sugiere un comportamiento conformacional consistente del residuo dentro del bolsillo de unión. Los derivados pirazol muestran distribuciones de χ_1 comparables a las demás familias, manteniendo un máximo alrededor de 180° . No obstante, en χ_2 se observa una distribución más dispersa, con poblaciones extendidas entre -120° y -60° , indicando una mayor variabilidad conformacional del residuo Tyr268 en presencia de estos ligandos.

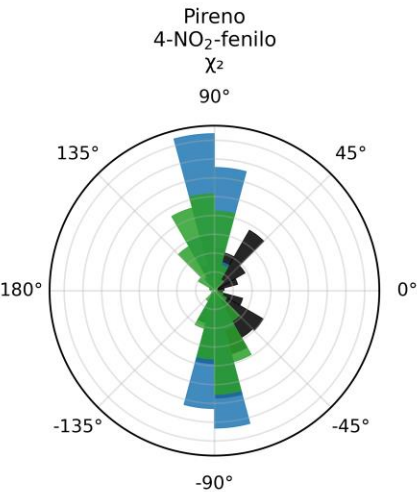
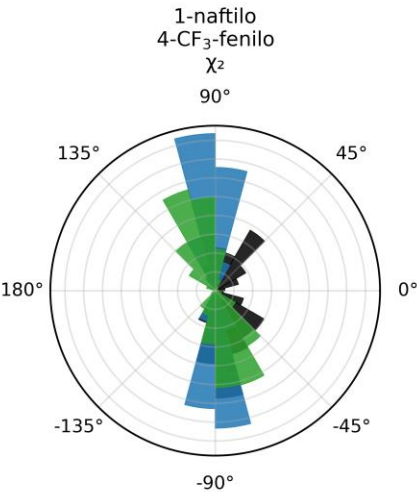
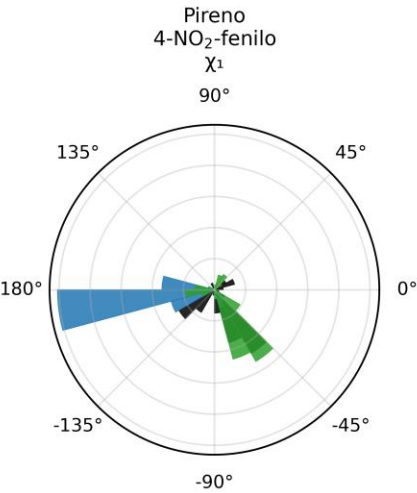
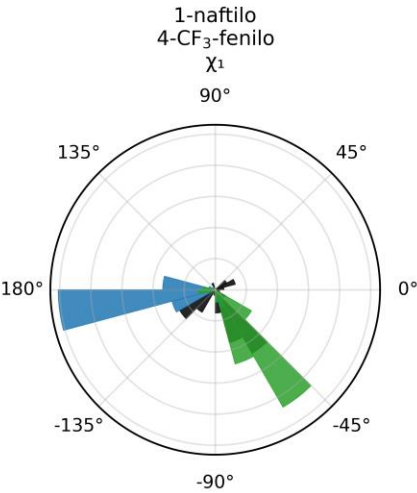
Diedros Tyr268 – Pirazol



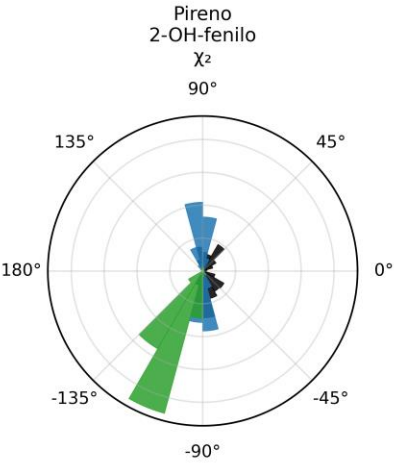
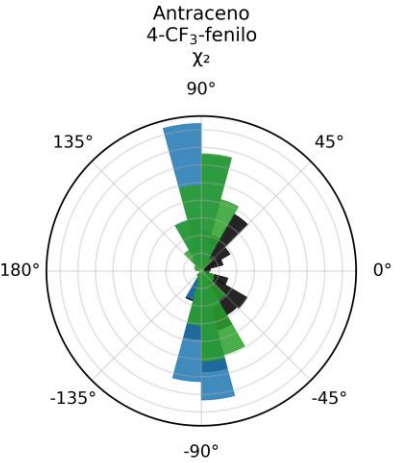
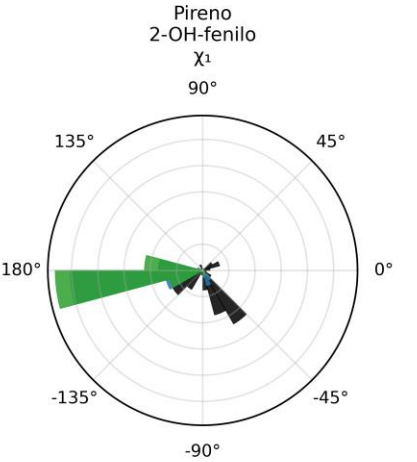
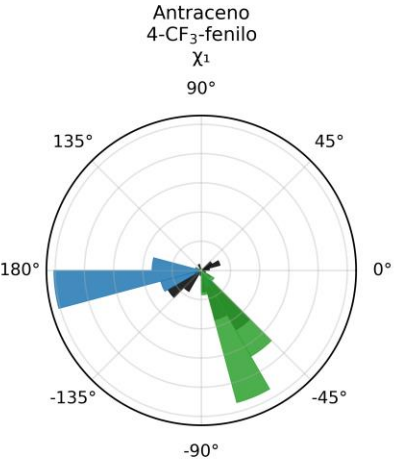
Diedros Tyr268 – Pirimidina aminobenzimidazol



Diedros Tyr268 – Chalcona



Diedros Tyr268 – Benzodiazepinas



Diedros Tyr268 – Aminopirimidina

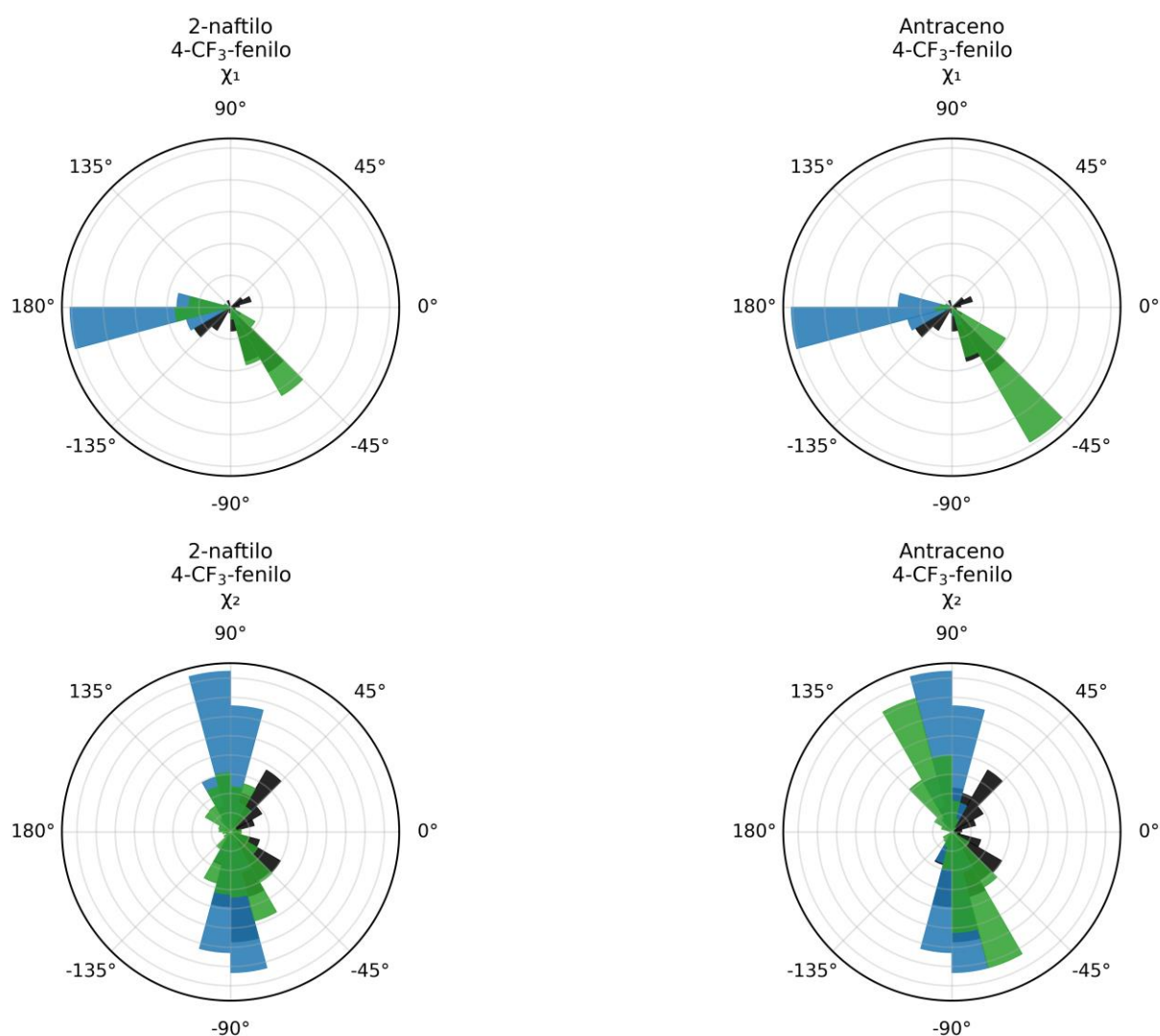


Figura 105 Distribución de los ángulos diedros χ_1 y χ_2 del residuo Tyr268 durante las simulaciones de Dinámica molecular. Los histogramas polares representan la frecuencia de los valores angulares observados a lo largo de la trayectoria. La escala radial corresponde al número de frames de la simulación.

5.5.5 Análisis derivados de benzimidazoles

El análisis de las trayectorias de Dinámica molecular permitió evaluar la estabilidad estructural del sistema proteína–ligando y los efectos estructurales inducidos por los derivados benzimidazol. La estabilidad global de la proteína se evaluó mediante el RMSD de la proteína a lo largo de la simulación (Figura 106).

En la proteína en estado apo se observa un incremento progresivo del RMSD que alcanza valores cercanos a 3.5–4.0 Å, indicando una mayor desviación estructural a lo largo de la trayectoria. En contraste, varios de los complejos proteína–ligando presentan trayectorias más estables, con valores de RMSD menores y fluctuaciones más contenidas. En particular, los complejos con benzimidazol 1, benzimidazol 3 y benzimidazol 6 muestran trayectorias relativamente estables, manteniendo valores cercanos a ~2 Å durante gran parte de la simulación. El complejo con GRL-0617 presenta también un comportamiento estable, con fluctuaciones moderadas alrededor de ~2.0–2.5 Å. La superposición de todas las trayectorias confirma que varios complejos con benzimidazoles presentan desviaciones estructurales menores que la proteína en estado apo.

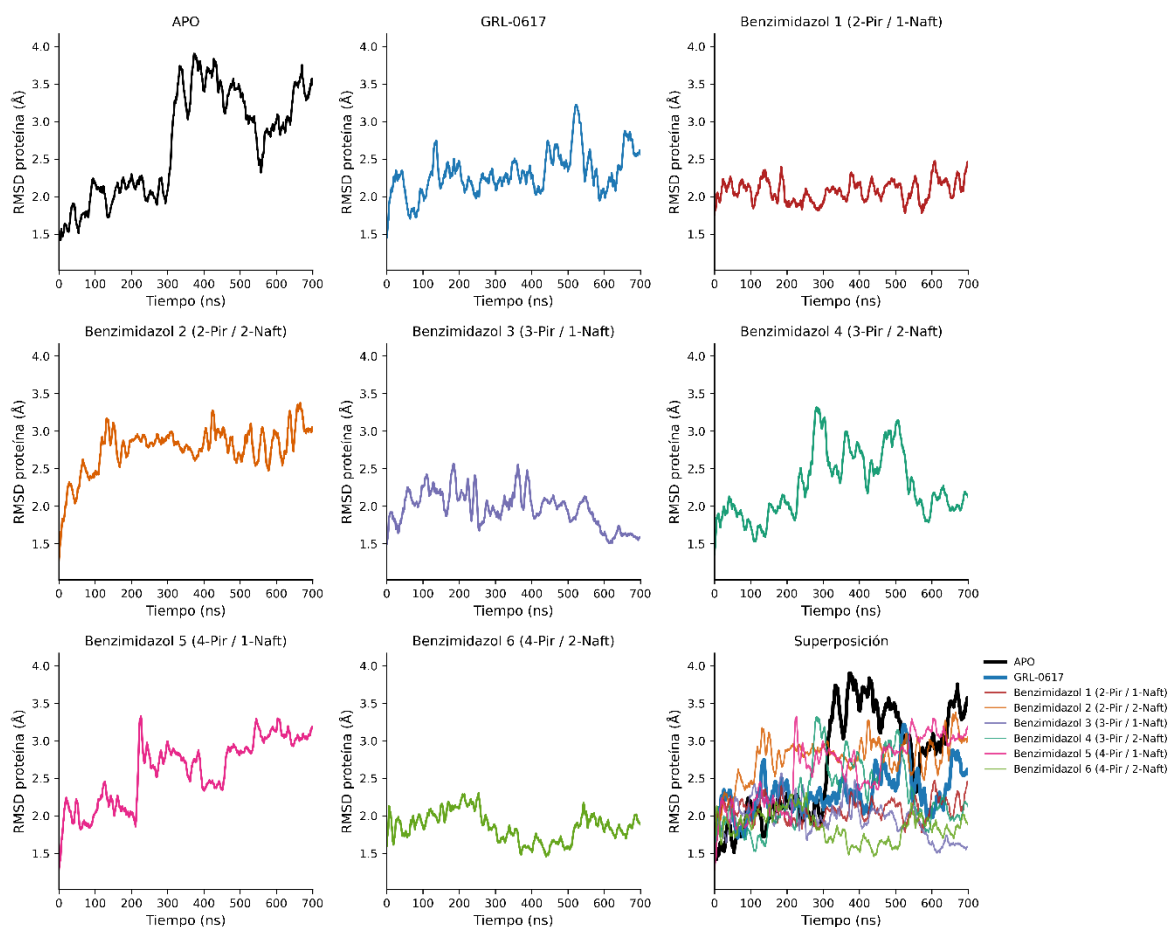


Figura 106 RMSD de la proteína durante las simulaciones de Dinámica molecular. Trayectorias de RMSD para la proteína en estado apo, el complejo con el ligando de referencia GRL-0617 y los complejos con los derivados benzimidazol. El panel final muestra la superposición de todas las trayectorias para facilitar la comparación entre sistemas.

El comportamiento dinámico de los ligandos se evaluó mediante el RMSD del ligando respecto a su posición inicial (Figura 107). El ligando de referencia GRL-0617 mantiene valores bajos de RMSD (~3–5 Å), indicando una posición estable dentro del sitio alostérico durante toda la simulación. Entre los derivados evaluados, benzimidazol 3 y benzimidazol 5 muestran valores relativamente estables de RMSD tras los primeros nanosegundos de simulación, lo que indica que permanecen localizados en la región de unión durante la trayectoria. En contraste, otros derivados presentan incrementos marcados del RMSD, alcanzando valores superiores a 20–40 Å, lo que sugiere desplazamientos del ligando respecto a su posición inicial. En algunos casos estos desplazamientos convergen posteriormente hacia valores constantes, indicando que el ligando puede estabilizarse en otras regiones de la proteína.

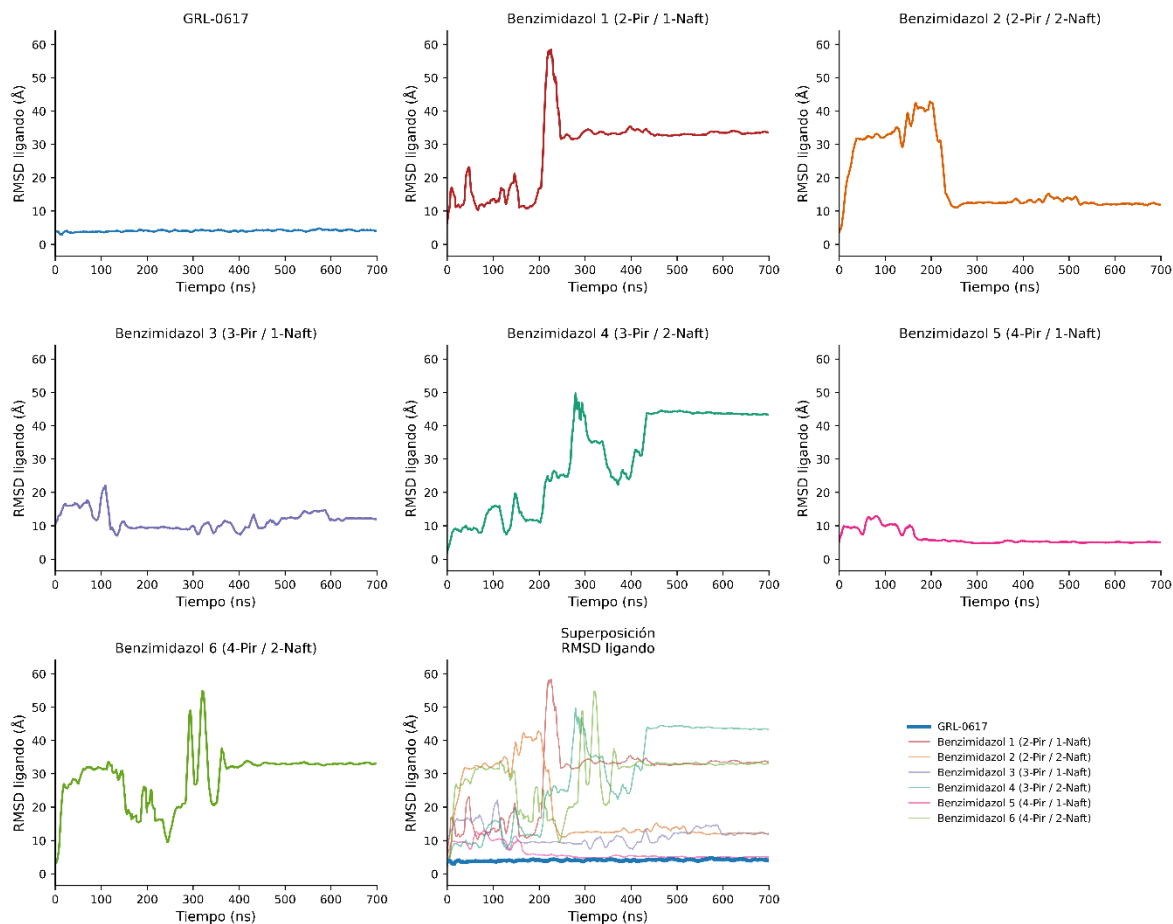


Figura 107 RMSD de los ligandos durante la Dinámica molecular. Evolución temporal del RMSD de los ligandos respecto a su posición inicial para GRL-0617 y los derivados benzimidazol. El panel final muestra la superposición de todas las trayectorias.

La flexibilidad local de la proteína se analizó mediante el RMSF por residuo (Figura 108). En general, la mayor parte de la estructura presenta fluctuaciones inferiores a $\sim 2 \text{ \AA}$, mientras que regiones específicas muestran picos de mayor movilidad. Entre estas destaca la región correspondiente al loop BL2 destacado en gris la Figura 108. El análisis comparativo muestra que algunos complejos presentan una reducción en la amplitud de las fluctuaciones en esta región. En particular, los complejos con benzimidazol 2 y benzimidazol 3 muestran valores menores de RMSF en el loop BL2 en comparación con la proteína apo. La superposición de los perfiles RMSF confirma esta tendencia, observándose una disminución de la amplitud del máximo localizado alrededor de los residuos 265–270 (Figura 109).

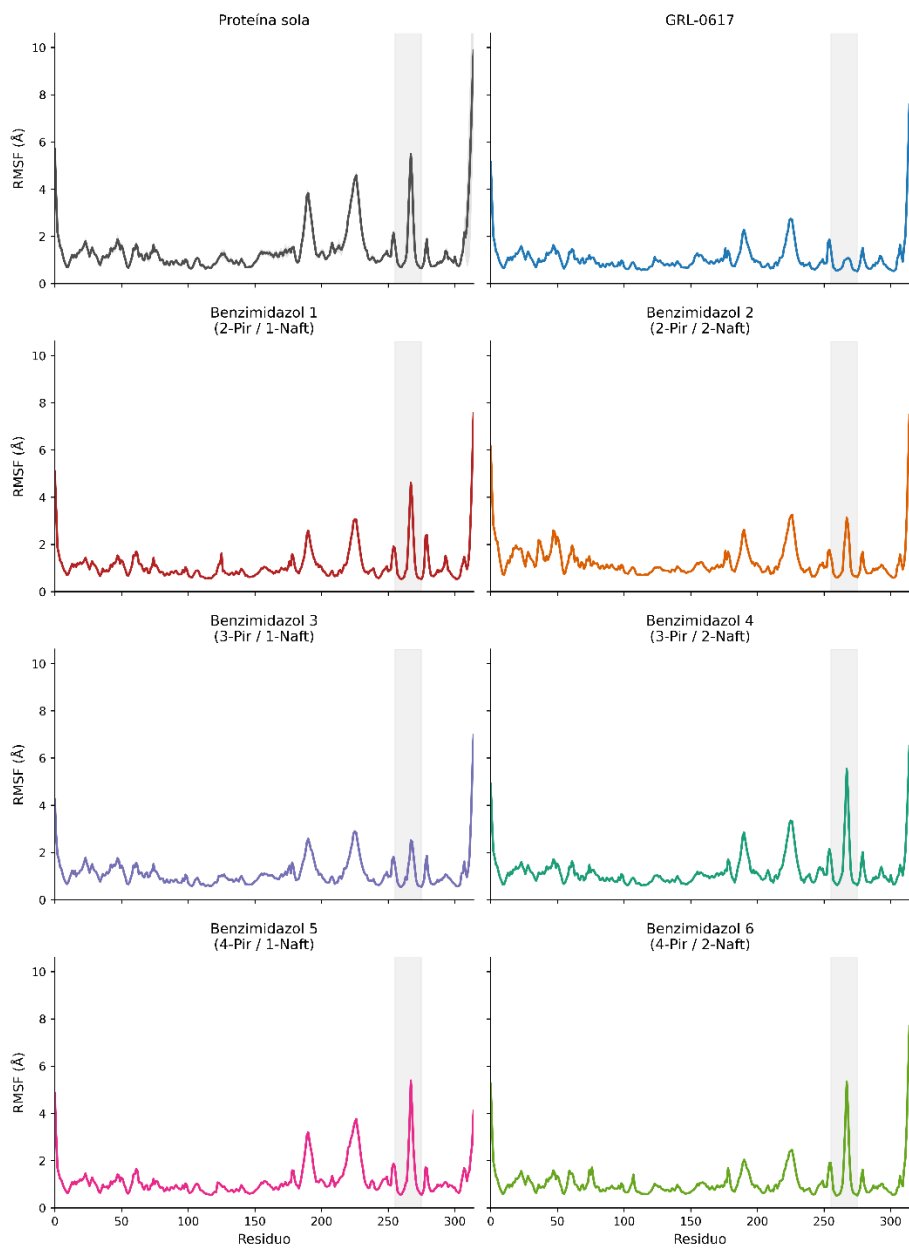


Figura 108 Perfil de RMSF de la proteína. Fluctuaciones cuadráticas medias por residuo para la proteína apo y los complejos proteína–ligando. La región sombreada corresponde al loop BL2.

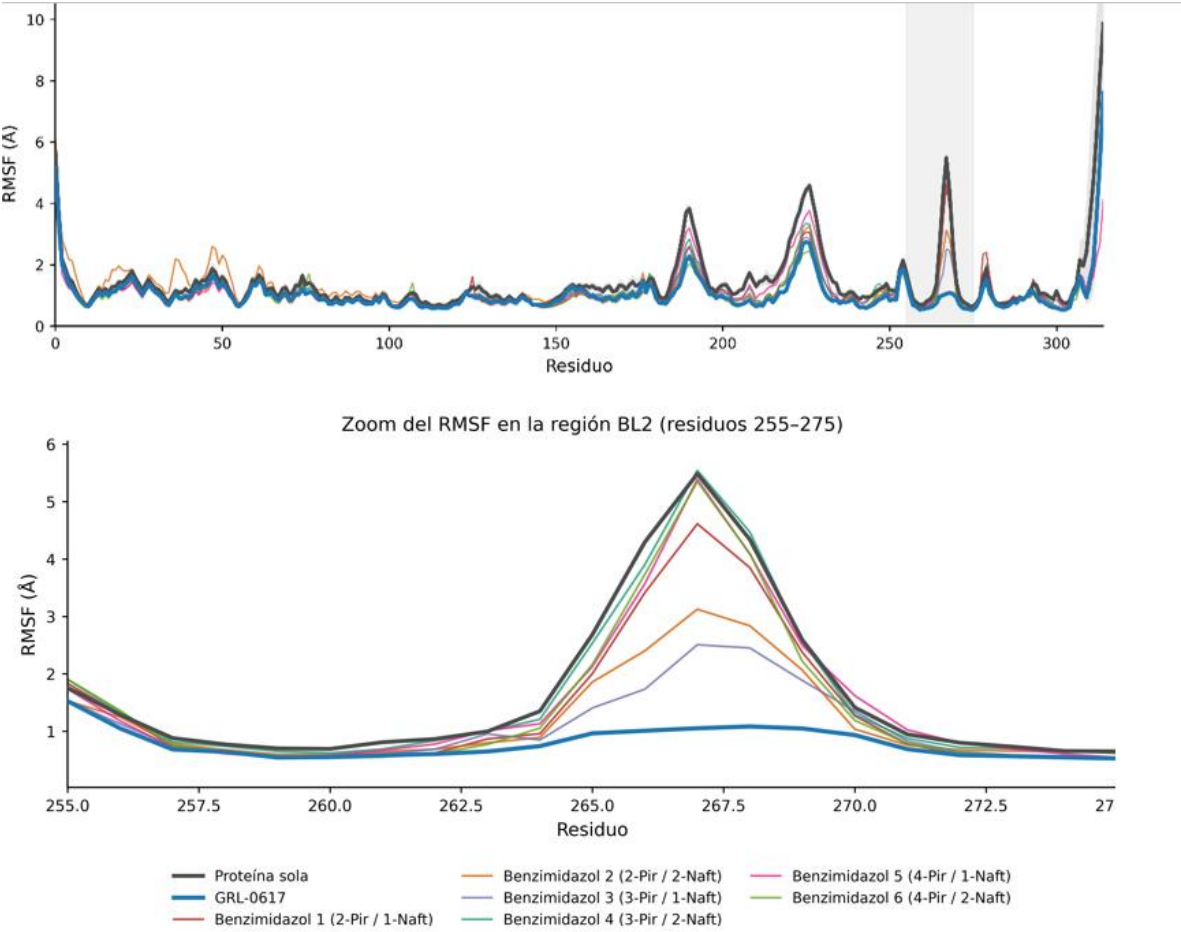


Figura 109 Superposición de RMSF y ampliación del loop BL2. Comparación directa de las fluctuaciones por residuo entre los distintos sistemas y ampliación de la región BL2 (residuos 255–275).

Con el fin de caracterizar posibles cambios estructurales en el entorno del sitio alostérico se analizaron distancias entre residuos clave a lo largo de la simulación. La distancia catalítica entre Cys111 y His272 se mantiene aproximadamente constante en todos los sistemas evaluados (Figura 110), con valores cercanos a ~6–7 Å, lo que indica que la geometría del sitio catalítico no presenta cambios significativos durante las trayectorias.

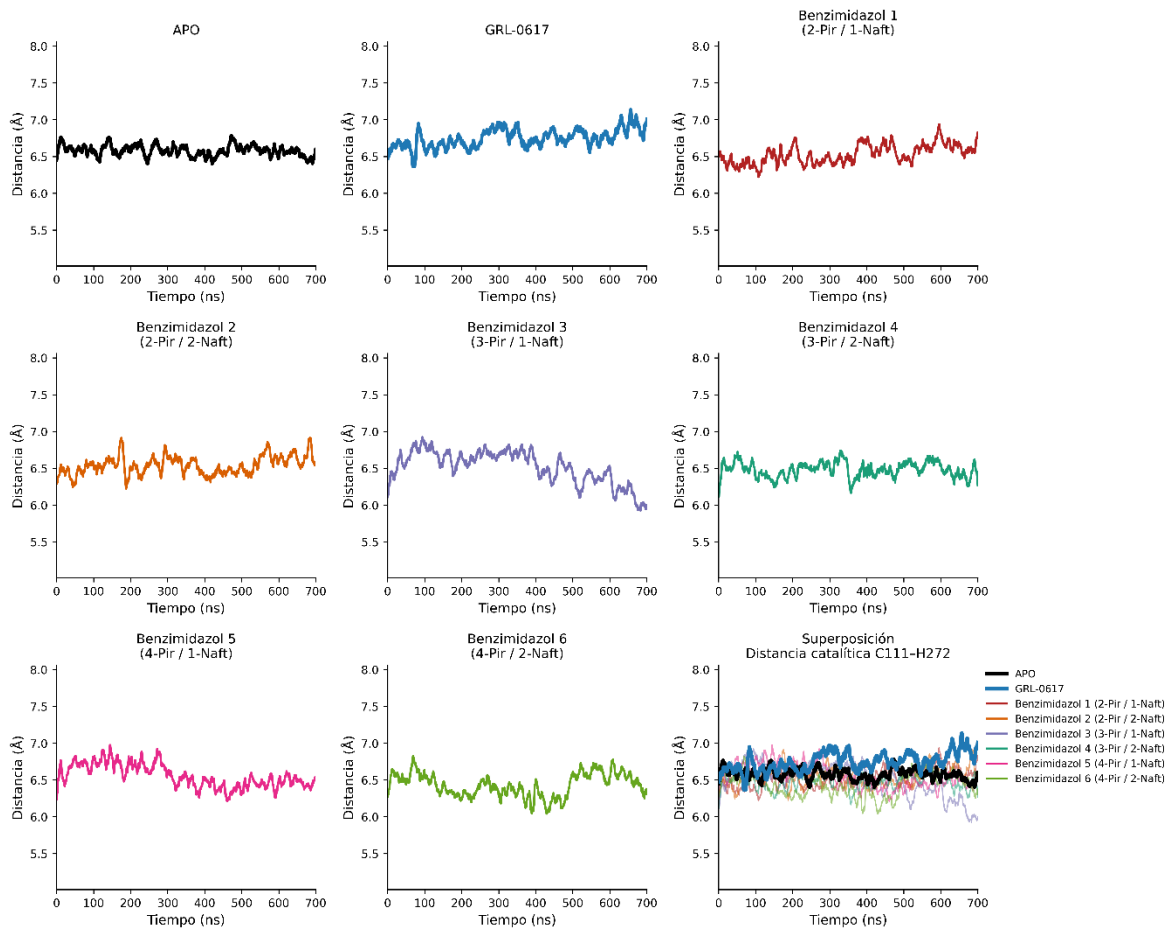


Figura 110 Distancia catalítica Cys111–His272. Evolución temporal de la distancia entre los residuos catalíticos Cys111 y His272 durante las simulaciones.

Se analizaron además distancias asociadas al movimiento del loop BL2. La distancia entre Pro299 y Tyr268 presenta variaciones dependientes del ligando (Figura 111). El complejo con GRL-0617 mantiene valores cercanos a ~11–12 Å durante la simulación. Entre los derivados evaluados, el complejo con benzimidazol 3 muestra un comportamiento similar, con valores comparables al ligando de referencia. Otros derivados presentan desplazamientos más amplios, con valores que oscilan entre ~10 y ~18 Å, indicando diferentes grados de apertura del loop BL2.

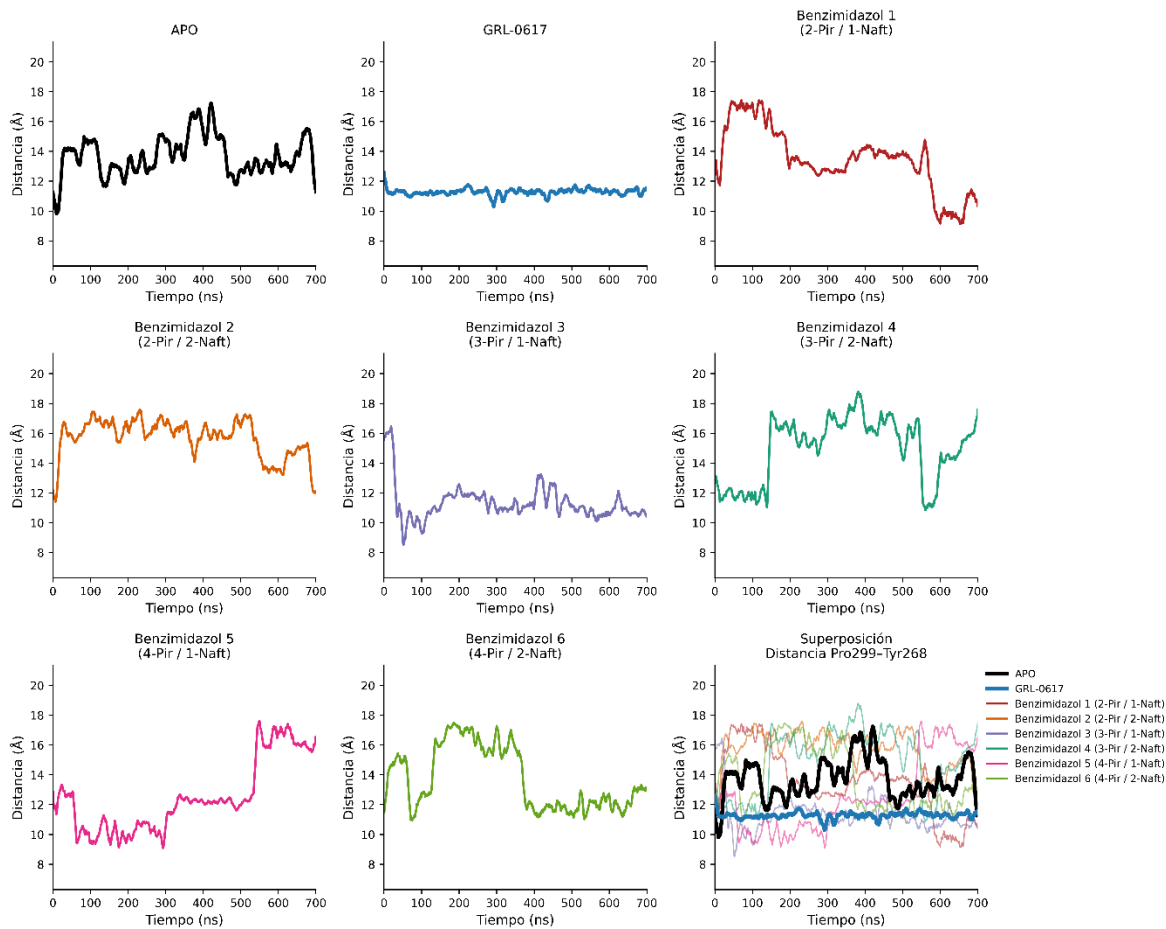


Figura 111 Distancia Pro299–Tyr268 asociada al movimiento del loop BL2.

De manera similar, la distancia entre Arg166 y Tyr268 también muestra diferencias entre los sistemas (Figura 112). Mientras que el ligando de referencia mantiene valores relativamente constantes cercanos a ~16 Å, algunos derivados presentan cambios más pronunciados en esta distancia a lo largo de la trayectoria, reflejando variaciones en la posición relativa del loop BL2, destacando nuevamente el benzimidazol 3 por su estabilidad en la distancia entre ambos aminoácidos por toda la simulación.

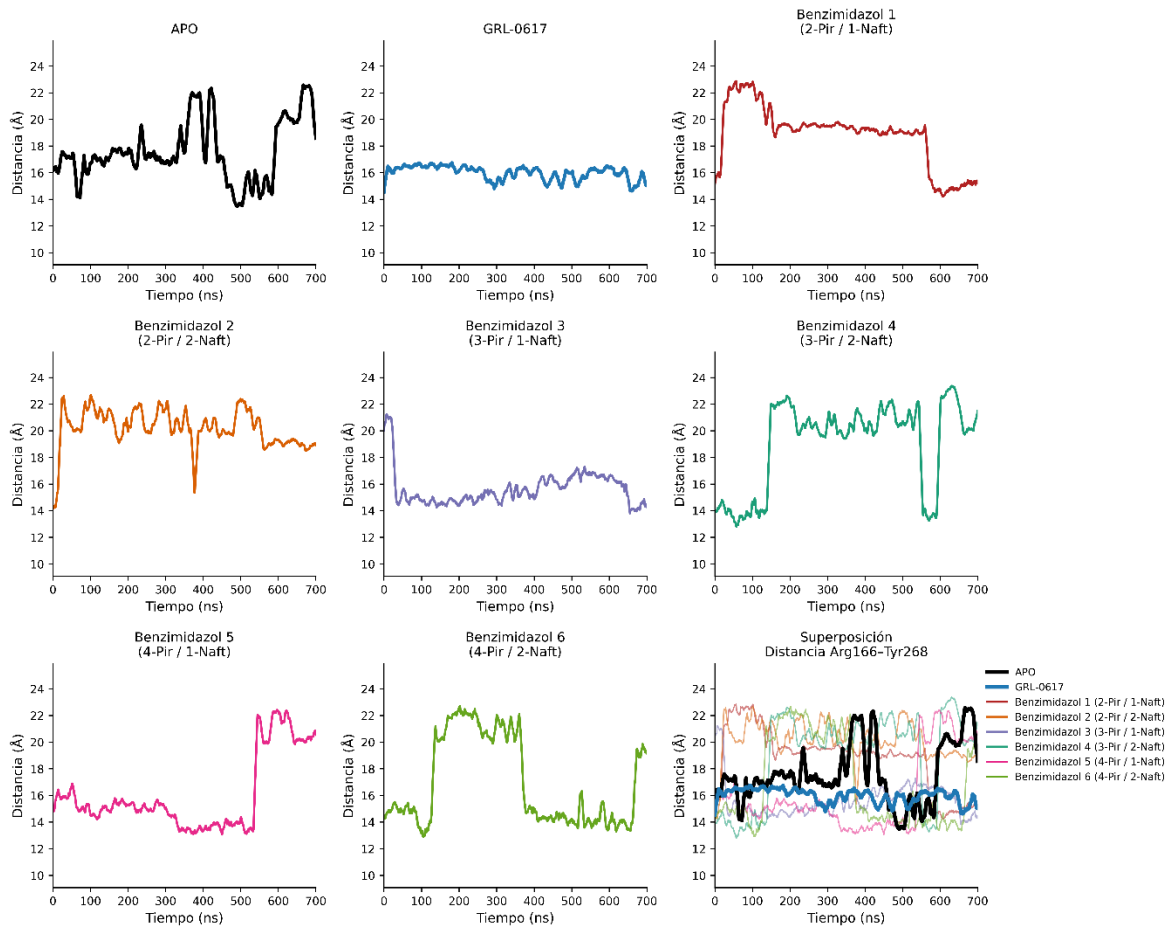


Figura 112 Distancia Arg166–Tyr268 asociada a la apertura del sitio alostérico.

Finalmente, se analizaron las distribuciones de los ángulos diedros χ_1 y χ_2 del residuo Tyr268 (Figura 113 y Figura 114). Las distribuciones angulares muestran que diferentes ligandos inducen poblaciones conformacionales distintas de este residuo. En el caso de χ_1 , los complejos con benzimidazol 1, 3, 5 y 6 muestran distribuciones similares, con poblaciones predominantes cercanas a $\pm 180^\circ$. En contraste, otros sistemas presentan distribuciones más dispersas. Para el ángulo χ_2 , varios complejos muestran poblaciones concentradas alrededor de valores cercanos a $\pm 90^\circ$, mientras que otros presentan distribuciones más amplias. En particular, benzimidazol 3 presenta una distribución más repartida entre diferentes regiones angulares, mientras que benzimidazol 2 y benzimidazol 4 muestran poblaciones concentradas alrededor de $\sim 100^\circ$ y $\sim -60^\circ$.

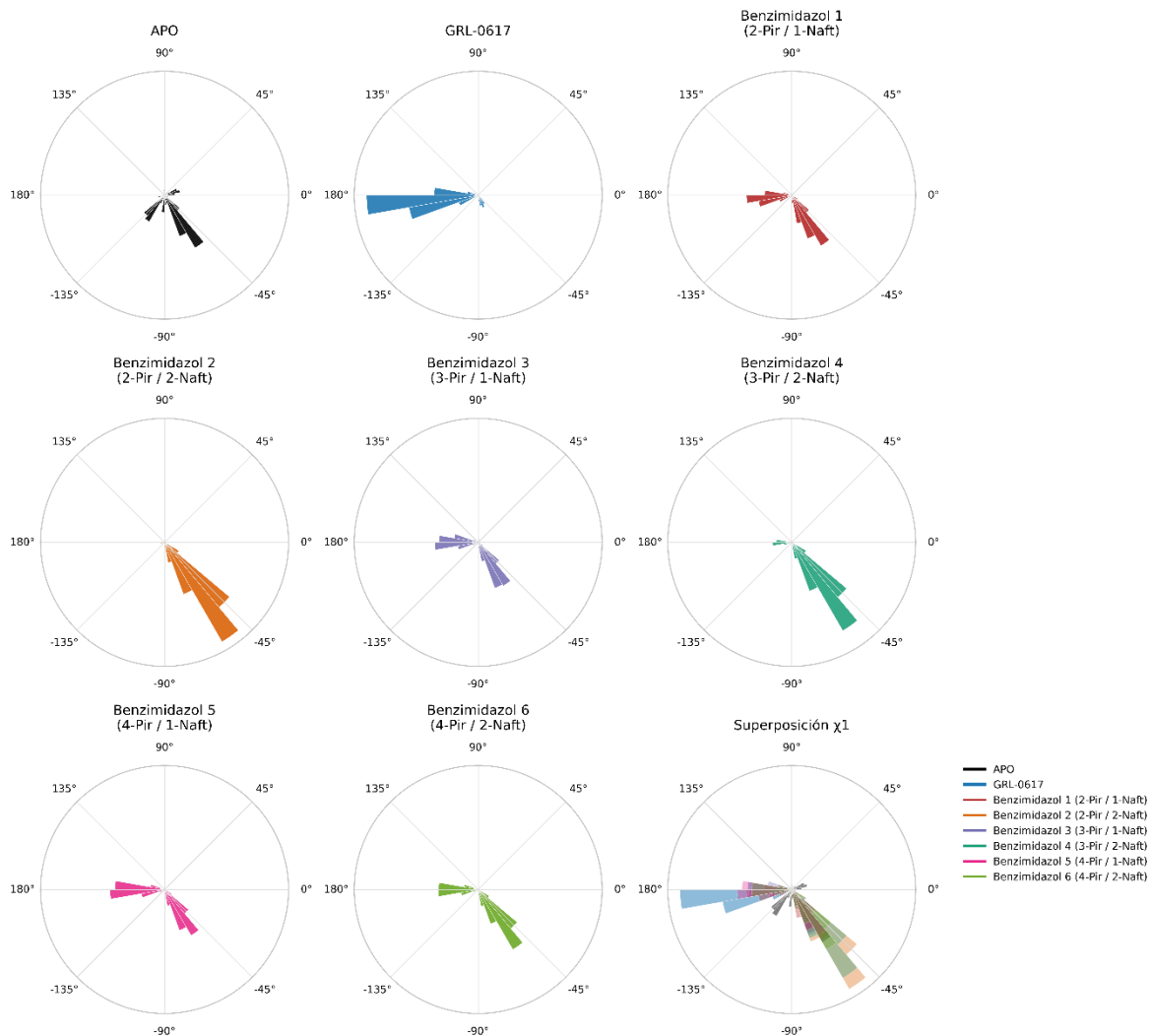


Figura 113 Distribución angular del diedro χ_1 del residuo Tyr268.



Figura 114 Distribución angular del diedro χ_2 del residuo Tyr268.

5.5.6 Análisis moléculas Top 1 Virtual screening

En los resultados de la evolución del RMSD de la proteína durante las simulaciones mostrado en Figura 115. La proteína en estado apo mostró una tendencia progresiva al aumento del RMSD, alcanzando los valores más altos hacia la segunda mitad de la trayectoria. En cambio, el complejo con GRL-0617 y la mayoría de los complejos proteína–ligando presentaron fluctuaciones más acotadas, con valores generalmente intermedios respecto al sistema apo. Entre los complejos evaluados, la aminopirimidina mostró el comportamiento más cercano al sistema apo, con un aumento más marcado del RMSD proteico durante la simulación. El resto de las familias mantuvo trayectorias relativamente similares entre sí y más próximas al comportamiento observado para el complejo con el ligando de referencia.

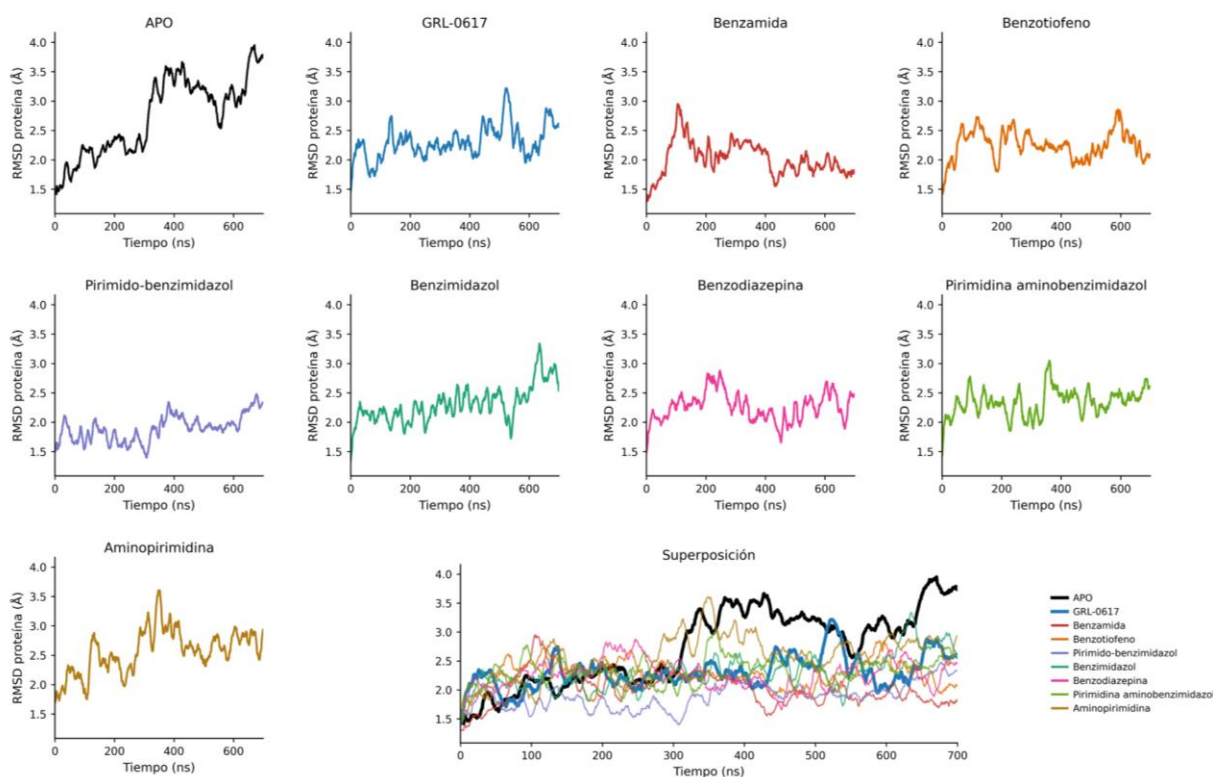


Figura 115 RMSD de la proteína durante las simulaciones de Dinámica molecular. Trayectorias de RMSD para la proteína en estado apo, el complejo con el ligando de referencia GRL-0617 y los complejos con las moléculas para cada familia. El panel final muestra la superposición de todas las trayectorias para facilitar la comparación entre sistemas.

El RMSD del ligando para el complejo con GRL-0617 y para las moléculas representativas de cada familia mostrado en la Figura 116. El ligando de referencia presentó una trayectoria estable y de baja magnitud durante toda la simulación. Entre las familias evaluadas, benzamida, benzodiazepina y pirimidina aminobenzimidazol mostraron perfiles relativamente estables desde etapas tempranas de la dinámica, aunque con diferentes magnitudes absolutas de RMSD. El benzimidazol presentó un comportamiento intermedio, con un aumento progresivo del RMSD durante la trayectoria. En contraste, benzotiofeno, pirimido-benzimidazol y aminopirimidina presentaron los valores más altos de RMSD del ligando, con

cambios abruptos y trayectorias más variables, lo que las diferencia claramente de los sistemas más estables.

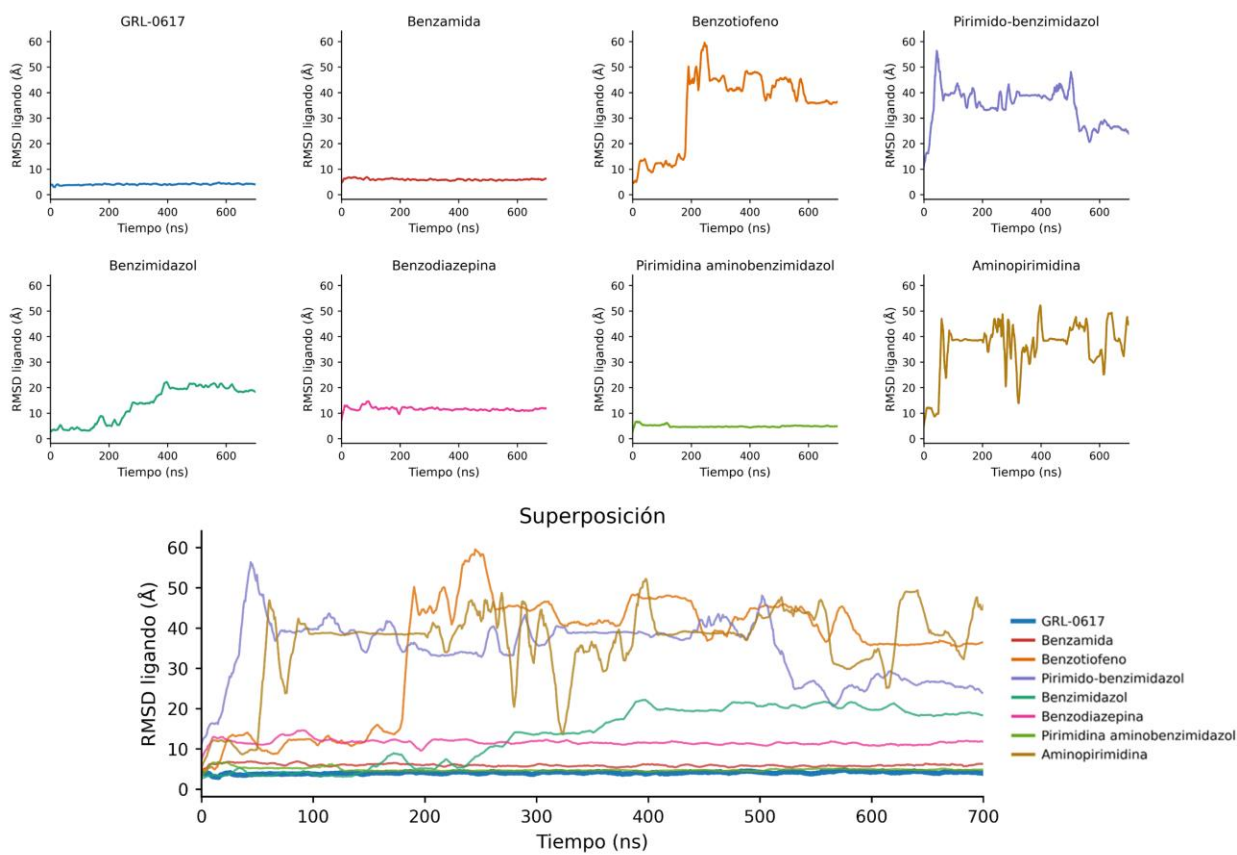


Figura 116 RMSD del ligando durante las simulaciones de Dinámica molecular. Trayectorias de RMSD con el ligando de referencia GRL-0617 y los complejos con las moléculas para cada familia. El panel final muestra la superposición de todas las trayectorias para facilitar la comparación entre sistemas.

En la Figura 117 se presentan los perfiles de RMSF por residuo para la proteína apo, el complejo con GRL-0617 y los complejos por familia. En todos los sistemas se observaron fluctuaciones localizadas en regiones específicas de la proteína, incluyendo los extremos de la secuencia y la región correspondiente al loop BL2. La proteína apo presentó una fluctuación marcada en la región BL2, especialmente alrededor de los residuos cercanos a Tyr268. El complejo con GRL-0617 mostró una reducción notoria de la fluctuación en esta región. Entre las familias evaluadas, benzamida y pirimidina aminobenzimidazol fueron las que presentaron una menor fluctuación del loop BL2 en comparación con el sistema apo. Las demás familias mostraron perfiles intermedios o con mayor flexibilidad local en esta región.

Perfiles de RMSF por residuo para la proteína sola, GRL-0617 y los complejos por familia

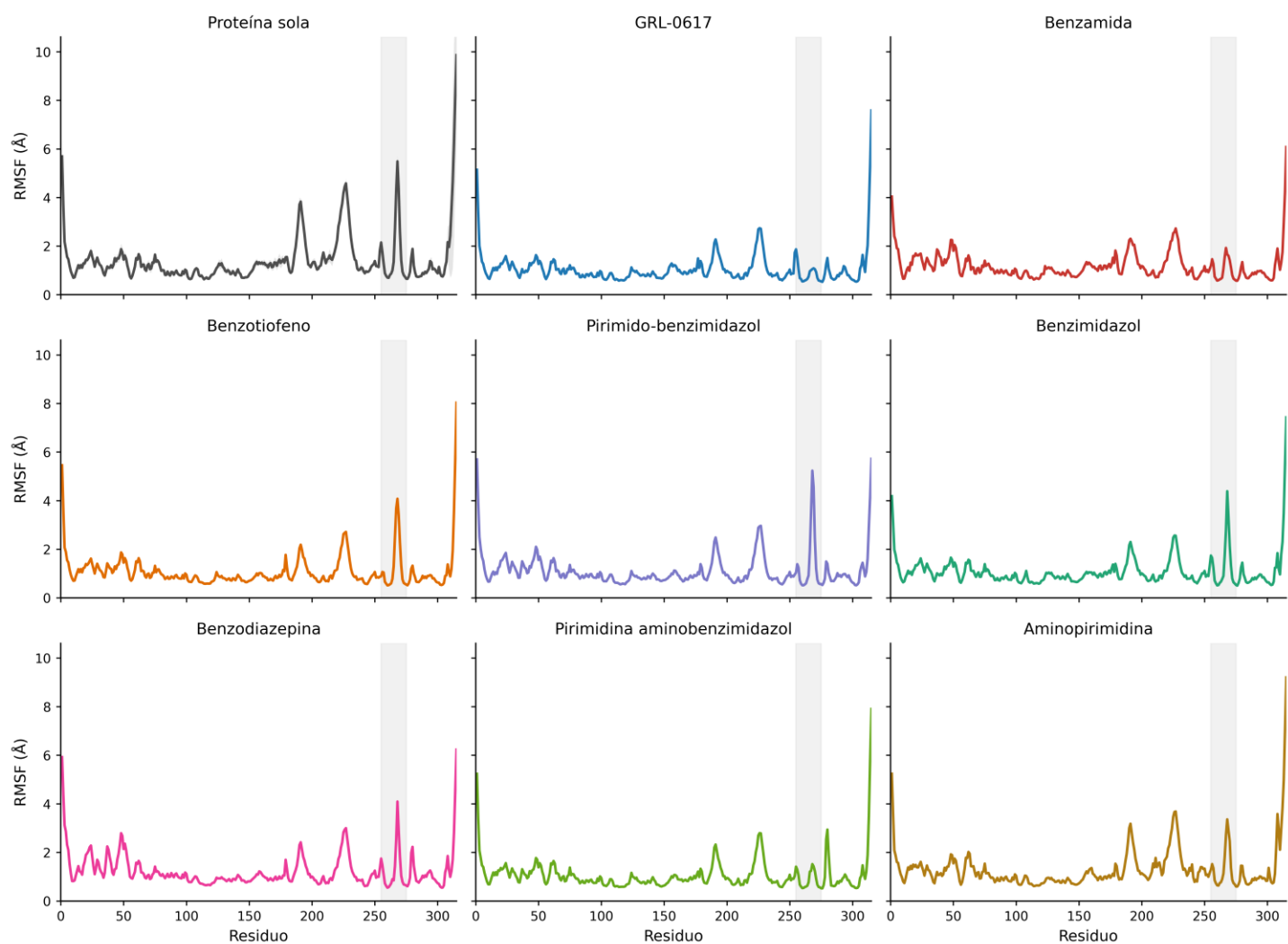


Figura 117 Perfil de RMSF de la proteína. Fluctuaciones cuadráticas medias por residuo para la proteína apo y los complejos proteína–ligando. La región sombreada corresponde al loop BL2.

El efecto de reducción de la flexibilidad del loop bl2 se ve con mayor claridad en la Figura 118.

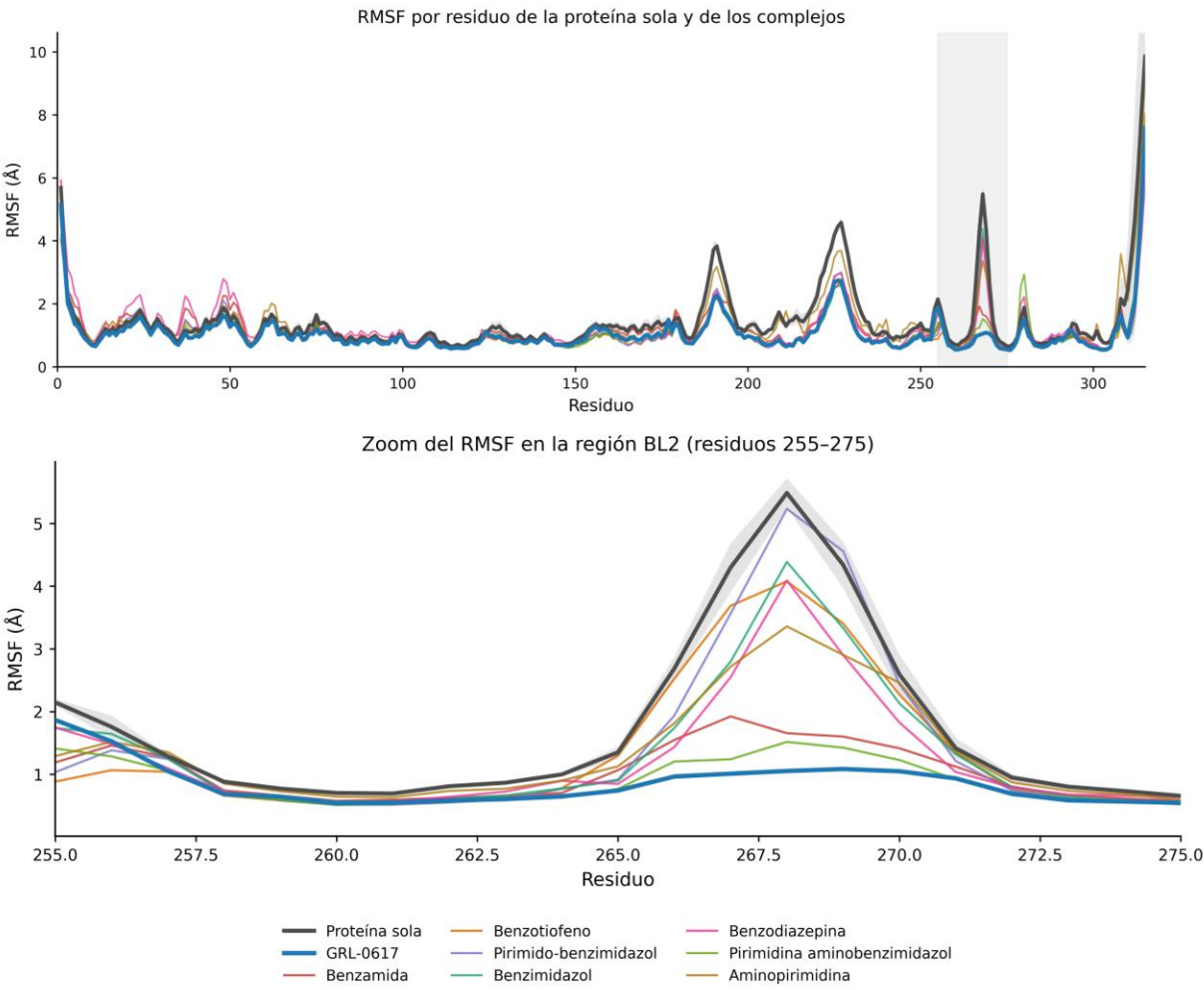


Figura 118 Superposición de RMSF y ampliación del loop BL2. Comparación directa de las fluctuaciones por residuo entre los distintos sistemas y ampliación de la región BL2 (residuos 255–275).

En la Figura 119 se muestra la evolución temporal de la distancia entre Cys111 e His272. En general, todos los sistemas mantuvieron esta distancia dentro de un rango relativamente estrecho durante la simulación, sin separaciones abruptas ni cambios sostenidos de gran magnitud. El complejo con GRL-0617 tendió a ubicarse en valores ligeramente superiores hacia la segunda mitad de la trayectoria. Benzimidazol presentó una mayor variabilidad local, mientras que benzodiazepina y aminopirimidina mostraron algunos intervalos con distancias algo menores. Sin embargo, en términos generales, la distancia catalítica Cys111–His272 no evidenció diferencias tan marcadas como las observadas en el RMSD del ligando o en las distancias asociadas al loop BL2.

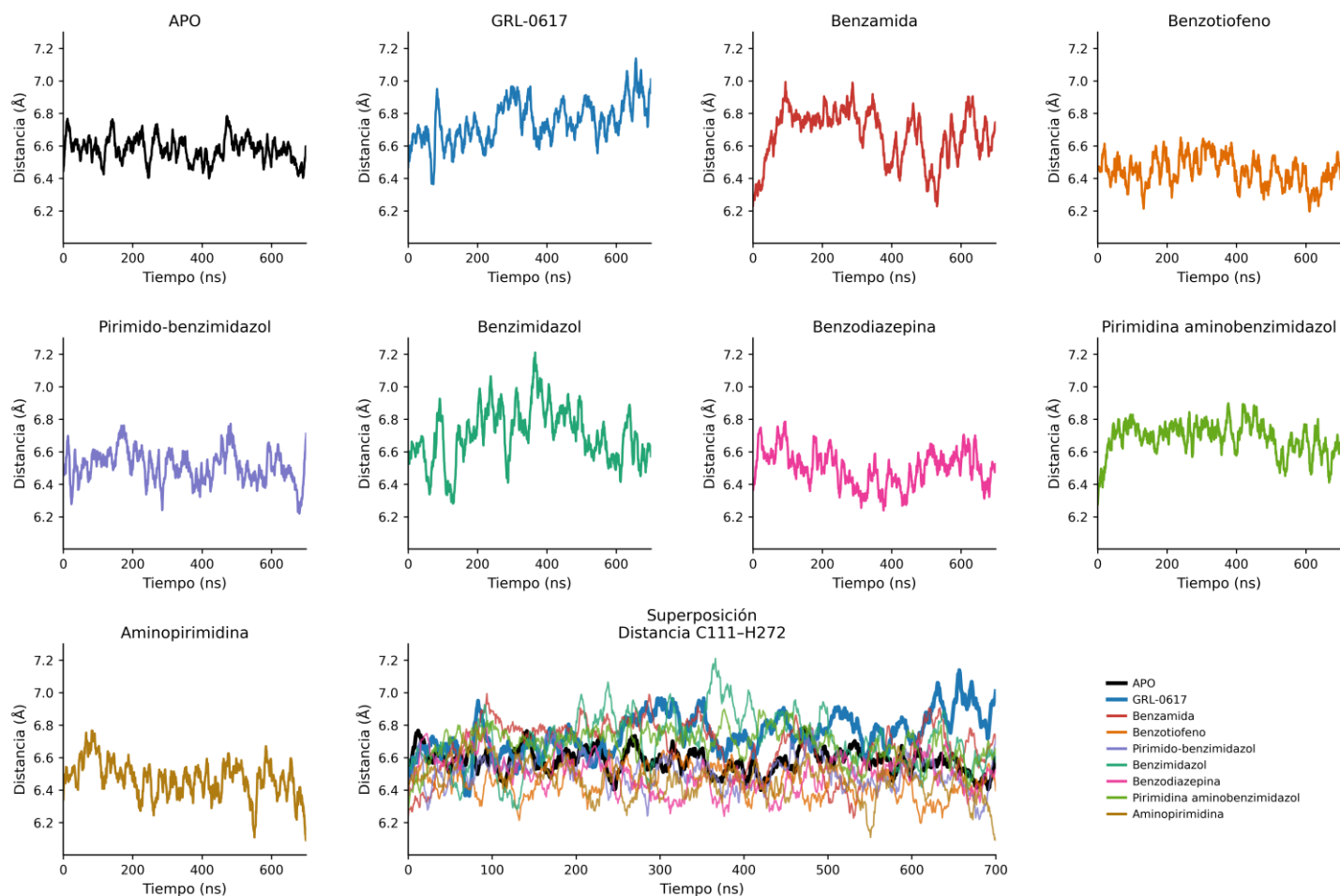


Figura 119 Distancia catalítica Cys111–His272. Evolución temporal de la distancia entre los residuos catalíticos Cys111 y His272 durante las simulaciones.

En las distancias asociadas al movimiento del loop BL2, se observaron diferencias más evidentes entre sistemas(Figura 120). Para la distancia Tyr268–Arg166, el complejo con GRL-0617 mantuvo un perfil relativamente estable. Benzamida y pirimidina aminobenzimidazol presentaron comportamientos cercanos al del ligando de referencia, con distancias moderadas y sin grandes transiciones. En cambio, benzotiofeno y benzodiazepina mostraron valores más altos durante gran parte de la simulación. Pirimido-benzimidazol presentó un cambio abrupto cercano a la mitad de la trayectoria, mientras que benzimidazol mostró un aumento marcado hacia los últimos nanosegundos. Aminopirimidina también presentó una transición temprana hacia valores mayores de distancia.

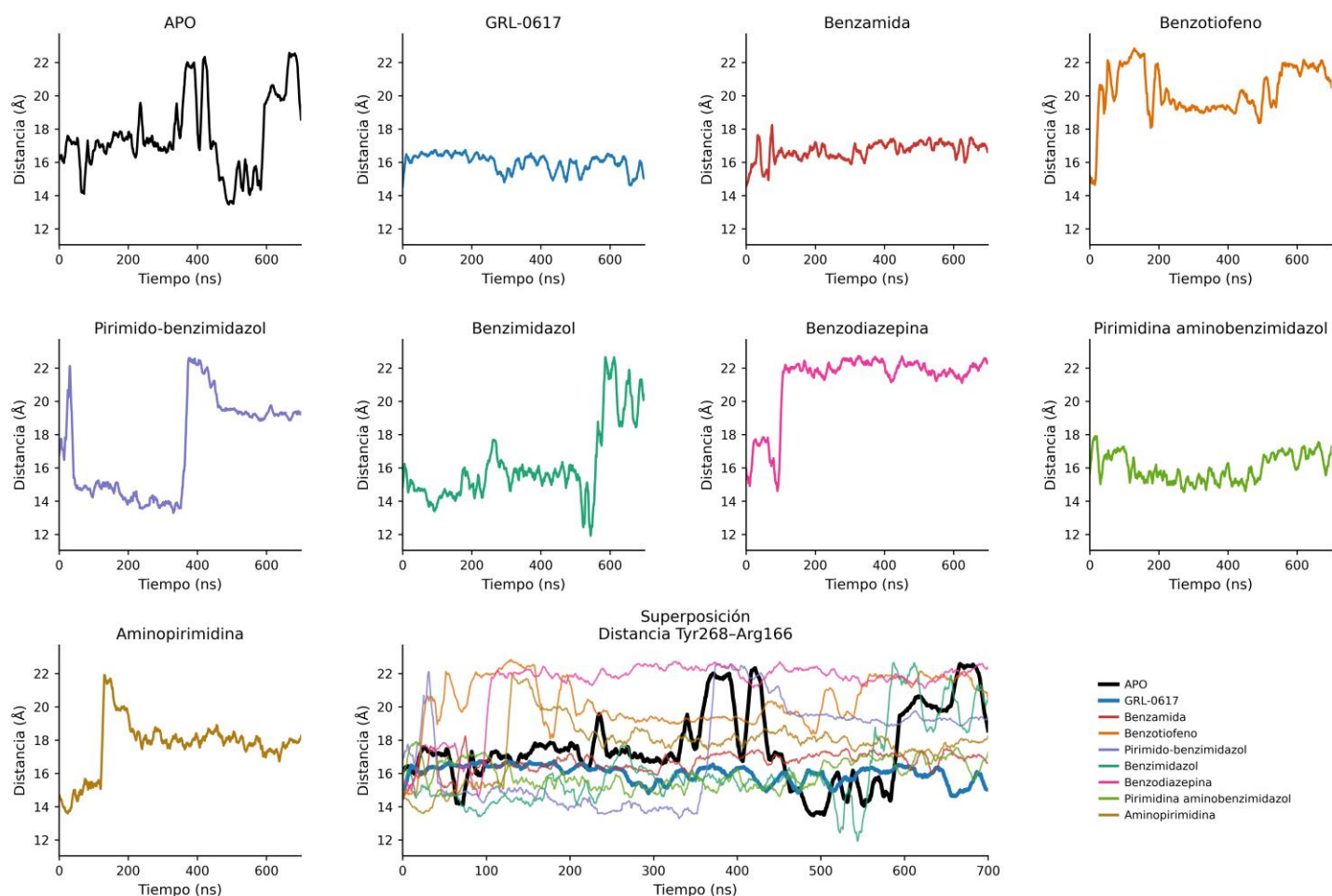


Figura 120 Distancia Arg166–Tyr268 asociada al movimiento del loop BL2.

Para la distancia Tyr268–Pro299, se observó un patrón similar. El complejo con GRL-0617 mantuvo una distancia baja y estable durante la simulación (Figura 121). Benzamida y pirimidina aminobenzimidazol conservaron valores cercanos a este comportamiento. En contraste, benzodiazepina y benzotiofeno mostraron distancias más elevadas, mientras que pirimido-benzimidazol y benzimidazol presentaron transiciones claras durante la trayectoria. Aminopirimidina mostró un aumento temprano de la distancia y luego una estabilización en valores superiores a los observados para GRL-0617.

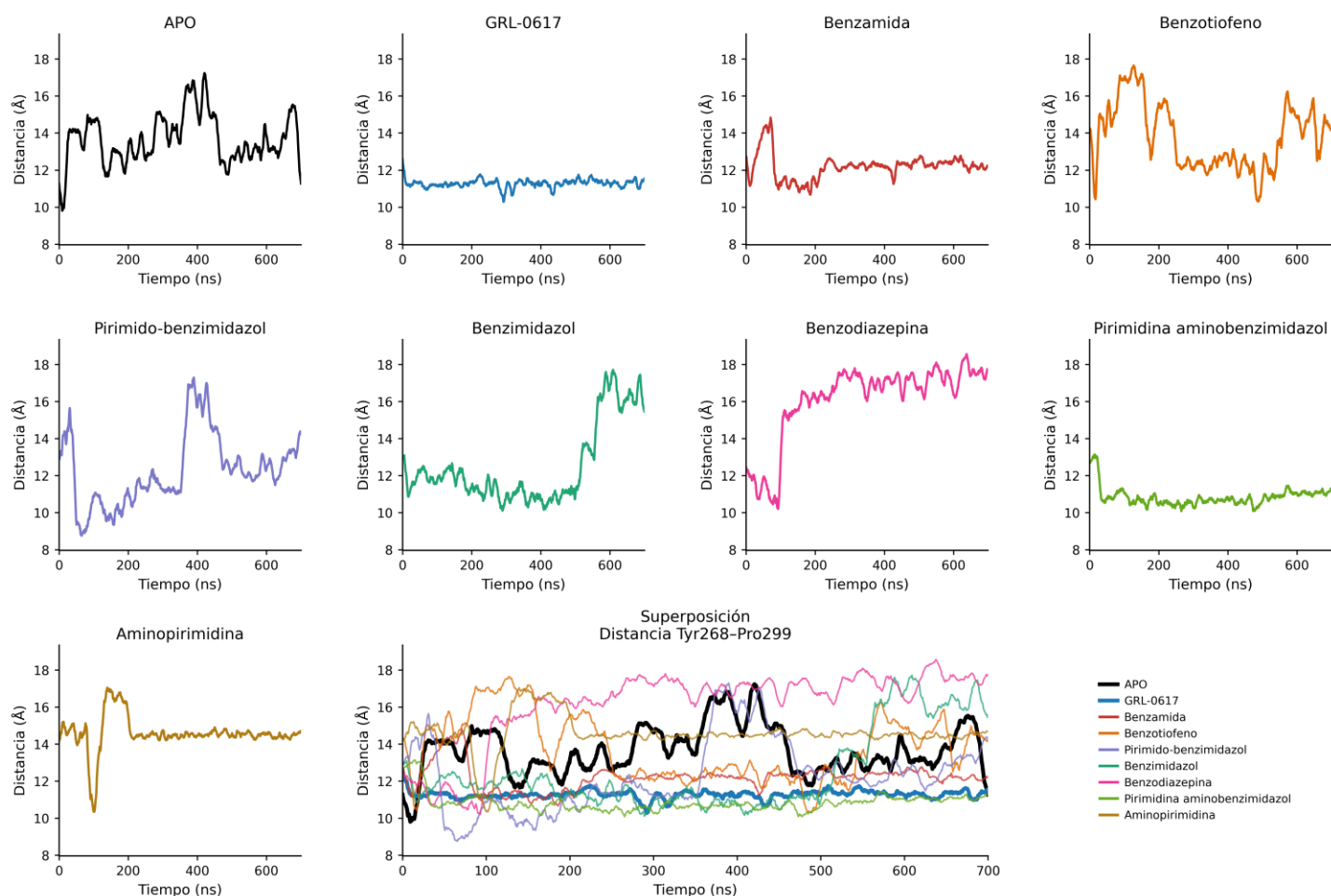


Figura 121 Distancia Pro299–Tyr268 asociada al movimiento del loop BL2.

En la Figura 122 se presenta la distribución angular del diedro χ_1 de Tyr268. La proteína apo mostró una distribución concentrada principalmente en ángulos negativos, mientras que GRL-0617 presentó una población angular desplazada hacia valores cercanos a $\pm 180^\circ$. Las familias evaluadas mostraron distribuciones diferentes entre sí. Benzamida destacó por presentar una orientación angular distinta tanto del sistema apo como del complejo con GRL-0617, con una población predominante en valores positivos. Benzotiofeno, pirimido-benzimidazol y benzimidazol mostraron distribuciones parcialmente superpuestas con regiones ocupadas por el sistema apo o por el ligando de referencia. Por otro lado, benzodiazepina, pirimidina aminobenzimidazol y aminopirimidina presentaron poblaciones más concentradas en sectores angulares negativos, con cierta similitud entre ellas.

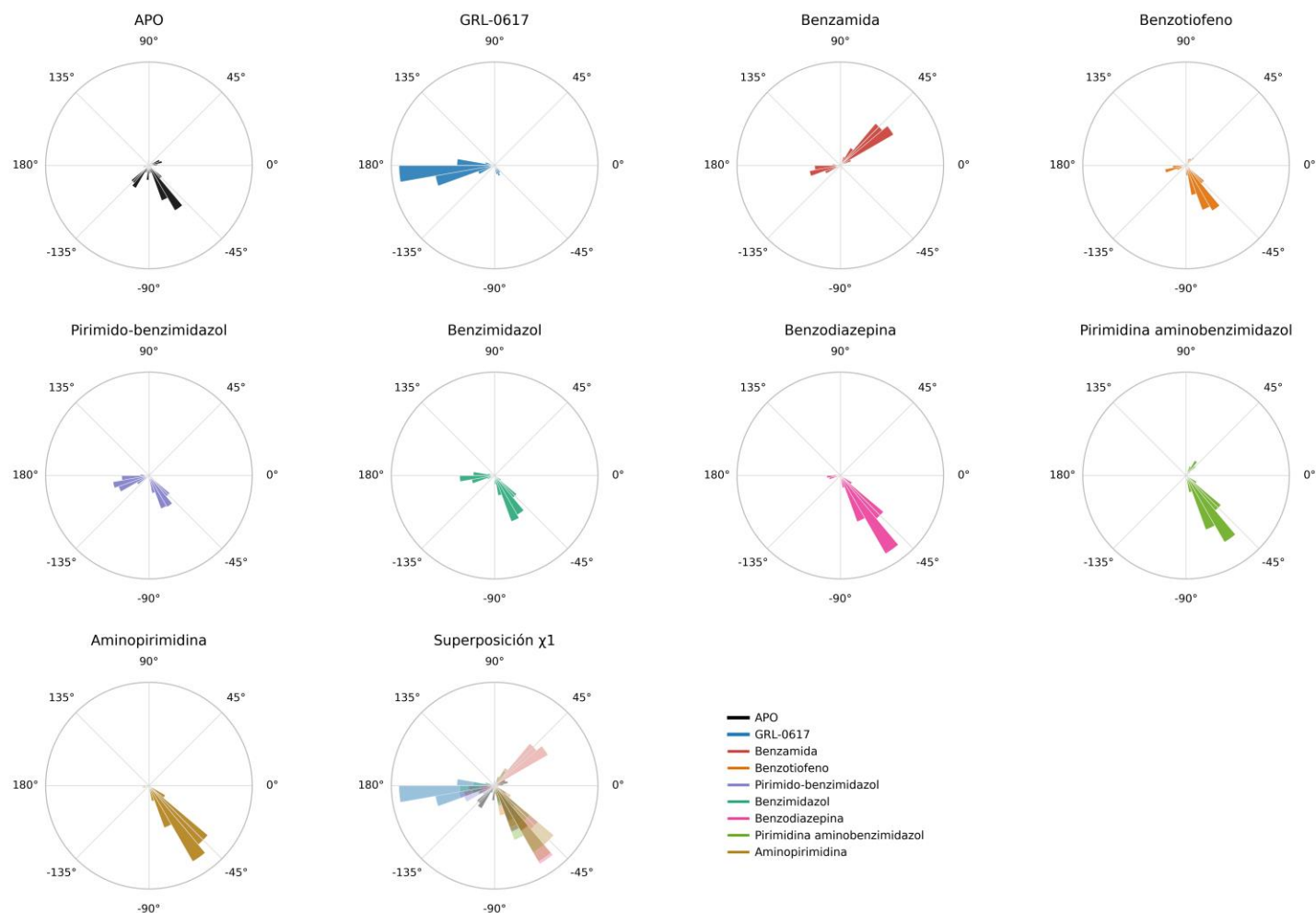


Figura 122 Distribución angular del diedro χ_1 del residuo Tyr268.

En la Figura 123 se muestra la distribución angular del diedro χ_2 de Tyr268. El sistema apo presentó una distribución más dispersa, mientras que GRL-0617 mostró poblaciones angulares más definidas en torno a orientaciones cercanas a $\pm 90^\circ$. Entre las familias evaluadas, benzotiofeno y benzamida fueron las que mostraron distribuciones más próximas al patrón observado para GRL-0617, aunque con desplazamientos en la orientación predominante. Pirimido-benzimidazol, benzimidazol, benzodiazepina y pirimidina aminobenzimidazol también presentaron poblaciones preferentes, pero con mayor variabilidad o desplazamiento respecto al ligando de referencia. Aminopirimidina mostró una distribución más diferenciada, con una orientación angular desplazada respecto a los demás sistemas.

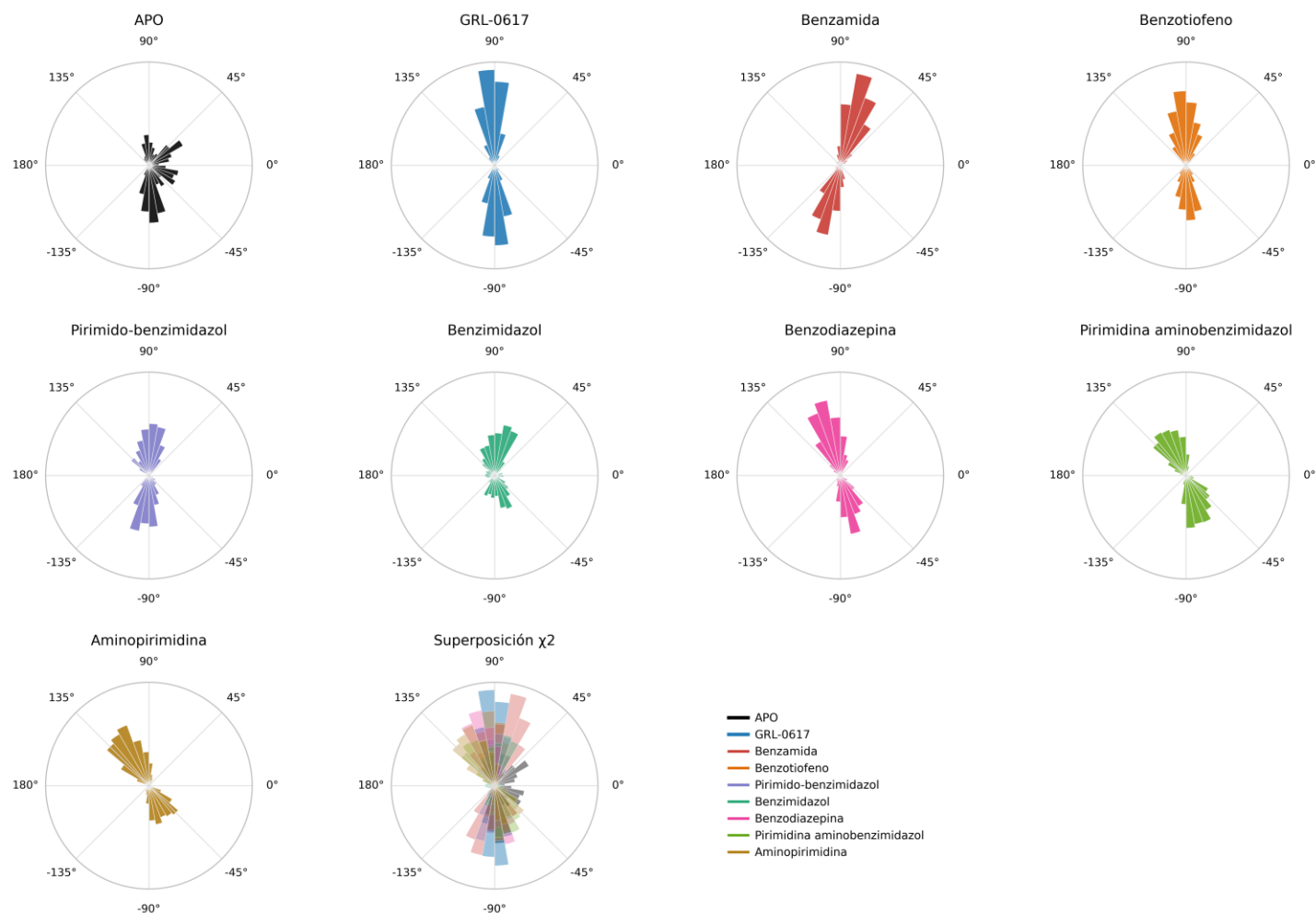


Figura 123 Distribución angular del diedro χ_2 del residuo Tyr268.

5.6 Farmacocinética

5.6.1 Chalconas y derivados heterocíclicos

5.6.1.1 Drogabilidad (*drugability*)

Se evaluó la drogabilidad de los compuestos de los compuestos mediante la regla de Lipinski y el criterio de Veber. La primera considera peso molecular, lipofilicidad y capacidad de formación de enlaces de hidrógeno, mientras que la regla de Veber relaciona la biodisponibilidad oral con la superficie polar topológica y la flexibilidad molecular .

La regla de Lipinski establece que los compuestos con mayor probabilidad de presentar adecuada biodisponibilidad oral suelen cumplir los siguientes criterios: peso molecular ≤ 500 Da, $\log P \leq 5$, no más de 5 donadores de enlace de hidrógeno (HBD) y no más de 10 aceptores de enlace de hidrógeno (HBA). Por su parte, la regla de Veber relaciona la biodisponibilidad oral con la superficie polar topológica (TPSA) y la flexibilidad molecular, estableciendo como criterios $TPSA \leq 140 \text{ \AA}^2$ y ≤ 10 enlaces rotables.

Los resultados obtenidos para las propiedades fisicoquímicas de los compuestos analizados se presentan en la Tabla IV del Anexo. En total, se evaluaron 400 moléculas pertenecientes a las distintas familias estructurales consideradas en este estudio.

El análisis del cumplimiento de la regla de Lipinski mostró que 386 compuestos no presentan violaciones a este criterio, mientras que 14 moléculas presentan al menos una violación. En relación con el criterio de Veber, todas las moléculas analizadas (400 compuestos) cumplen los parámetros establecidos para la superficie polar topológica y el número de enlaces rotables.

Los valores individuales de peso molecular, lipofilicidad ($\log P$), número de donadores y aceptores de enlace de hidrógeno, superficie polar topológica y número de enlaces rotables para cada uno de los compuestos analizados se presentan en Anexo IV.

5.6.1.2 Absorción

Las propiedades relacionadas con la absorción intestinal de las moléculas fueron estimadas mediante modelos de predicción ADMET. Los resultados se resumen en las Figura 124 y Figura 126.

La clasificación empírica de absorción intestinal humana (HIA) indica que la totalidad de las moléculas analizadas se ubican dentro de la categoría favorable, de acuerdo con los umbrales establecidos por el modelo predictivo (Figura 124). En consecuencia, ninguna molécula fue clasificada dentro de las categorías intermedia o desfavorable, observándose una distribución completamente concentrada en la región correspondiente a absorción favorable.

La predicción de interacción con la glicoproteína P (P-gp) muestra que 235 moléculas fueron clasificadas como potenciales inhibidores de P-gp, mientras que 165 compuestos no presentaron esta característica (Figura 124). En contraste, el análisis de sustrato de P-gp indica que únicamente 5 moléculas fueron predichas como posibles sustratos, mientras que 395 compuestos no presentan probabilidad significativa de ser transportados por esta proteína.

La distribución de los valores de permeabilidad predicha en el modelo Caco-2 se presenta en la Figura 125. El gráfico de violín muestra una distribución relativamente compacta de los valores de permeabilidad, con una mediana aproximada cercana a $-4.9 \log \text{ cm/s}$. Considerando el umbral empírico de permeabilidad ($-5.15 \log \text{ cm/s}$), se observa que la mayoría de los compuestos se sitúa en la región asociada a alta permeabilidad, mientras que un total de 14 moléculas presenta valores iguales o inferiores a dicho límite.

Con el fin de comparar estas propiedades entre las diferentes familias estructurales estudiadas, se calculó el promedio de cada descriptor de absorción y se representó mediante un mapa de calor (Figura 126). En este análisis, la probabilidad de absorción intestinal favorable (expresada como $1 - \text{HIA}$)

($1 - \text{HIA}$) presenta valores cercanos a 1.0 en todas las familias, indicando una tendencia uniforme en este parámetro. Por su parte, la permeabilidad Caco-2 promedio varía entre las distintas familias químicas, con valores que oscilan aproximadamente entre 0.16 y 0.72 en la escala normalizada utilizada.

Respecto a la interacción con P-gp, se observan diferencias entre las familias estructurales. Algunas familias presentan probabilidades promedio más altas de inhibición de P-gp, mientras que otras muestran valores considerablemente menores. En contraste, la probabilidad de actuar como sustrato de P-gp se mantiene baja en todas las familias analizadas.

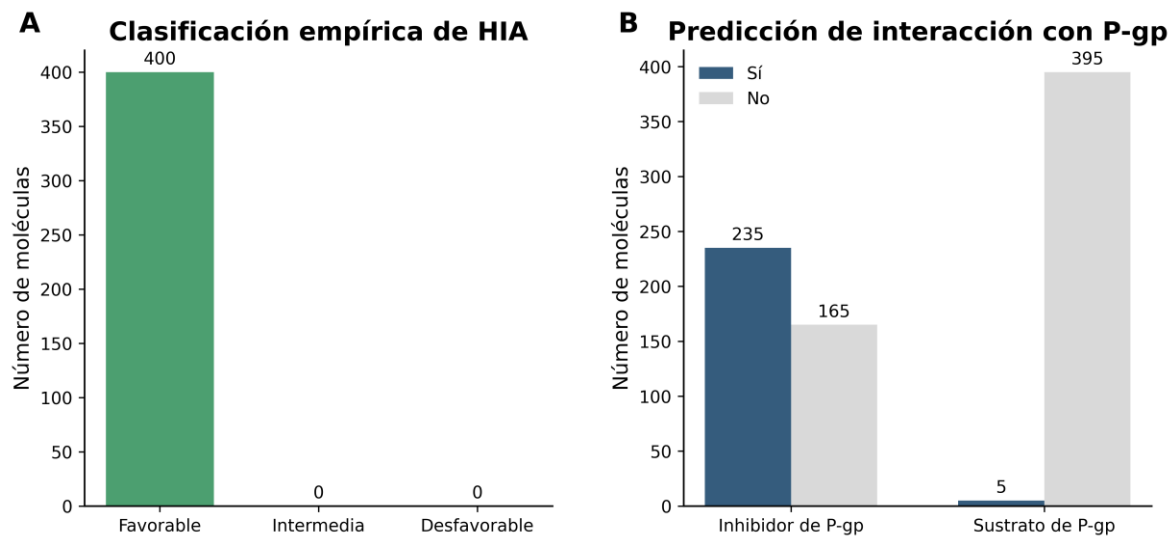


Figura 124 Propiedades de absorción intestinal e interacción con P-glicoproteína predichas. (A) Clasificación de las moléculas según la probabilidad de absorción intestinal humana (HIA). De acuerdo con la escala empírica del modelo, valores entre 0–0.3 corresponden a compuestos con absorción intestinal favorable, valores entre 0.3–0.7 indican absorción intermedia, mientras que valores entre 0.7–1.0 se asocian a una absorción desfavorable. (B) Distribución de moléculas predichas como inhibidores de P-glicoproteína (P-gp) y sustratos de P-gp, utilizando un umbral de probabilidad de 0.5 para la clasificación binaria. Los valores indican el número total de compuestos en cada categoría.

Distribución de la permeabilidad Caco-2

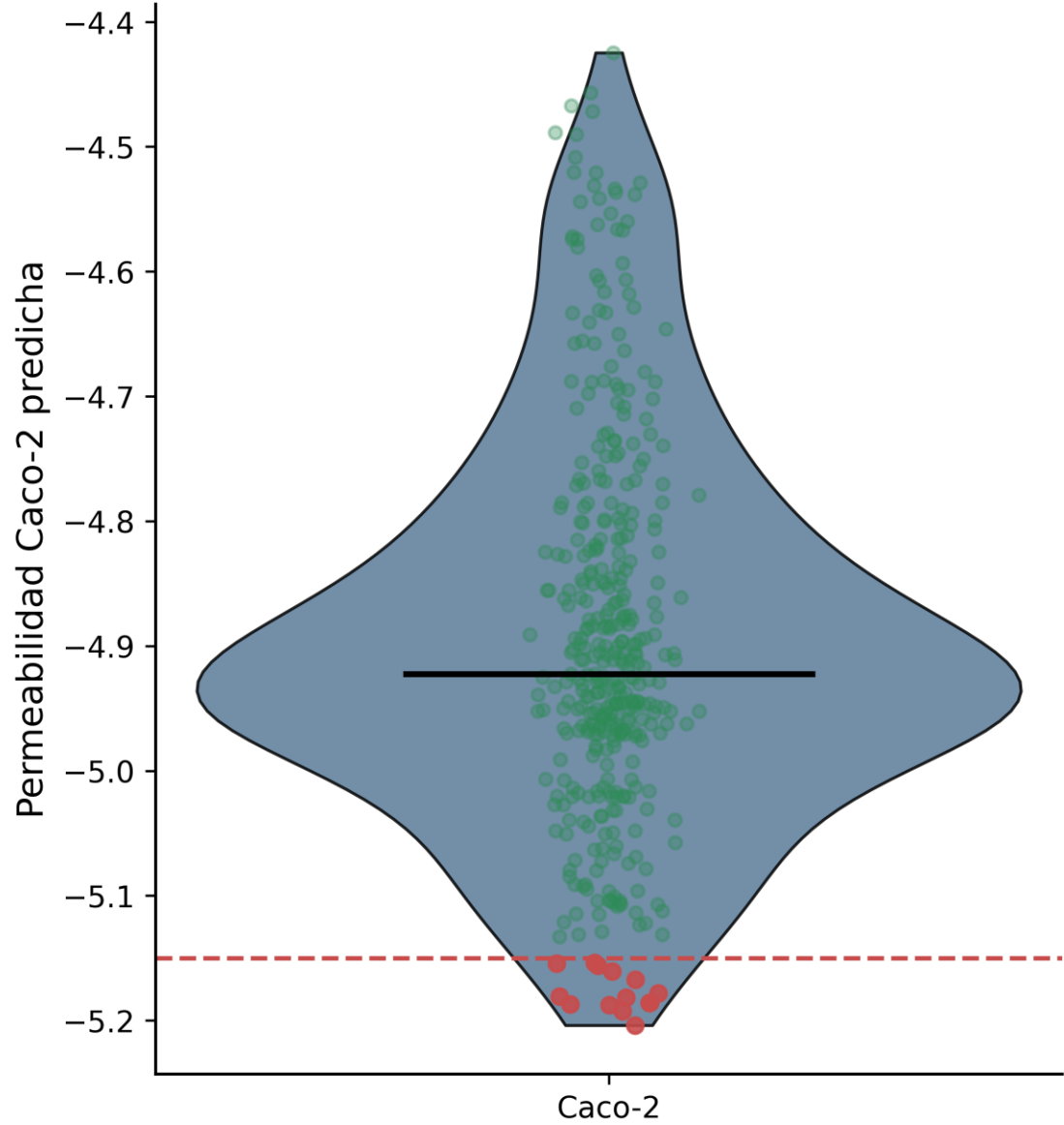


Figura 125 Distribución de la permeabilidad Caco-2 predicha para el conjunto de compuestos. El diagrama de violín representa la densidad de la distribución, mientras que los puntos individuales corresponden a cada molécula evaluada. La línea negra indica la mediana. Los compuestos con valores superiores a $-5.15 \log \text{ cm/s}$ se muestran en verde (alta permeabilidad), mientras que aquellos con valores iguales o inferiores a este umbral se muestran en rojo (baja permeabilidad). La línea roja discontinua indica el umbral empírico de permeabilidad intestinal asociado a una adecuada absorción.

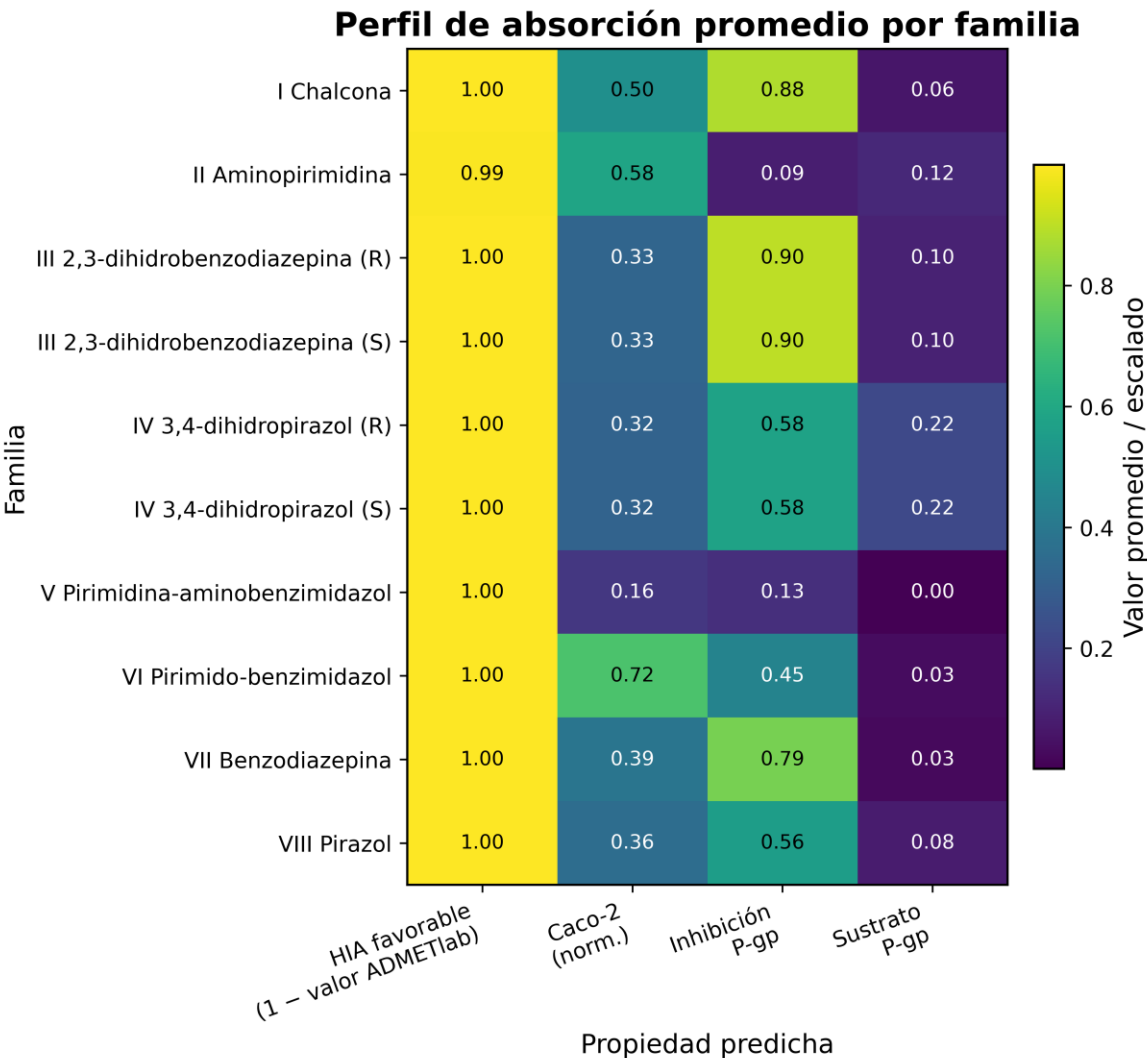


Figura 126 Perfil de absorción intestinal promedio por familia química. Mapa de calor que resume las propiedades de absorción predichas para las distintas familias de compuestos analizadas. Las propiedades incluidas corresponden a: probabilidad de absorción intestinal humana (HIA), permeabilidad Caco-2 normalizada, probabilidad de inhibición de P-glicoproteína, y probabilidad de actuar como sustrato de P-glicoproteína. Para facilitar la interpretación visual, el valor de HIA fue transformado como 1 – HIA, de modo que valores más altos en el mapa de calor representen un perfil de absorción más favorable. Los valores mostrados corresponden al promedio de las moléculas pertenecientes a cada familia química, permitiendo comparar tendencias globales entre los distintos scaffolds estructurales.

5.6.1.3Distribución

Las propiedades de distribución de las moléculas fueron evaluadas mediante los descriptores de permeabilidad a la barrera hematoencefálica (BBB), unión a proteínas plasmáticas (PPB), fracción libre en plasma (Fu) y volumen de distribución en estado estacionario (VDss), utilizando los modelos predictivos. La clasificación de la permeabilidad a la barrera hematoencefálica muestra que el conjunto de moléculas se distribuye entre compuestos con capacidad de penetración cerebral (BBB+) y aquellos sin dicha capacidad (BBB-) (Figura 127A). En términos cuantitativos, se observa una distribución relativamente equilibrada entre ambas categorías, lo que evidencia una variabilidad en este parámetro dentro del conjunto analizado.

La distribución de los valores de unión a proteínas plasmáticas (PPB) se presenta en la Figura 127B. El gráfico de violín muestra que la mayoría de las moléculas exhibe valores elevados de PPB, concentrándose en un rango estrecho cercano a los valores máximos del modelo predictivo. Esta tendencia se refleja en una distribución sesgada hacia altos porcentajes de unión, con una dispersión limitada de los datos.

En concordancia con lo anterior, la distribución de la fracción libre en plasma (Fu) en la Figura 127(C) muestra valores predominantemente bajos para la mayoría de las moléculas analizadas. La distribución

se encuentra concentrada en la región inferior del rango evaluado, indicando una baja proporción de compuesto libre en plasma en términos generales. Los valores de VDss de la Figura 127D muestran que todas la moléculas analizadas se encuentran dentro del rango óptimo.

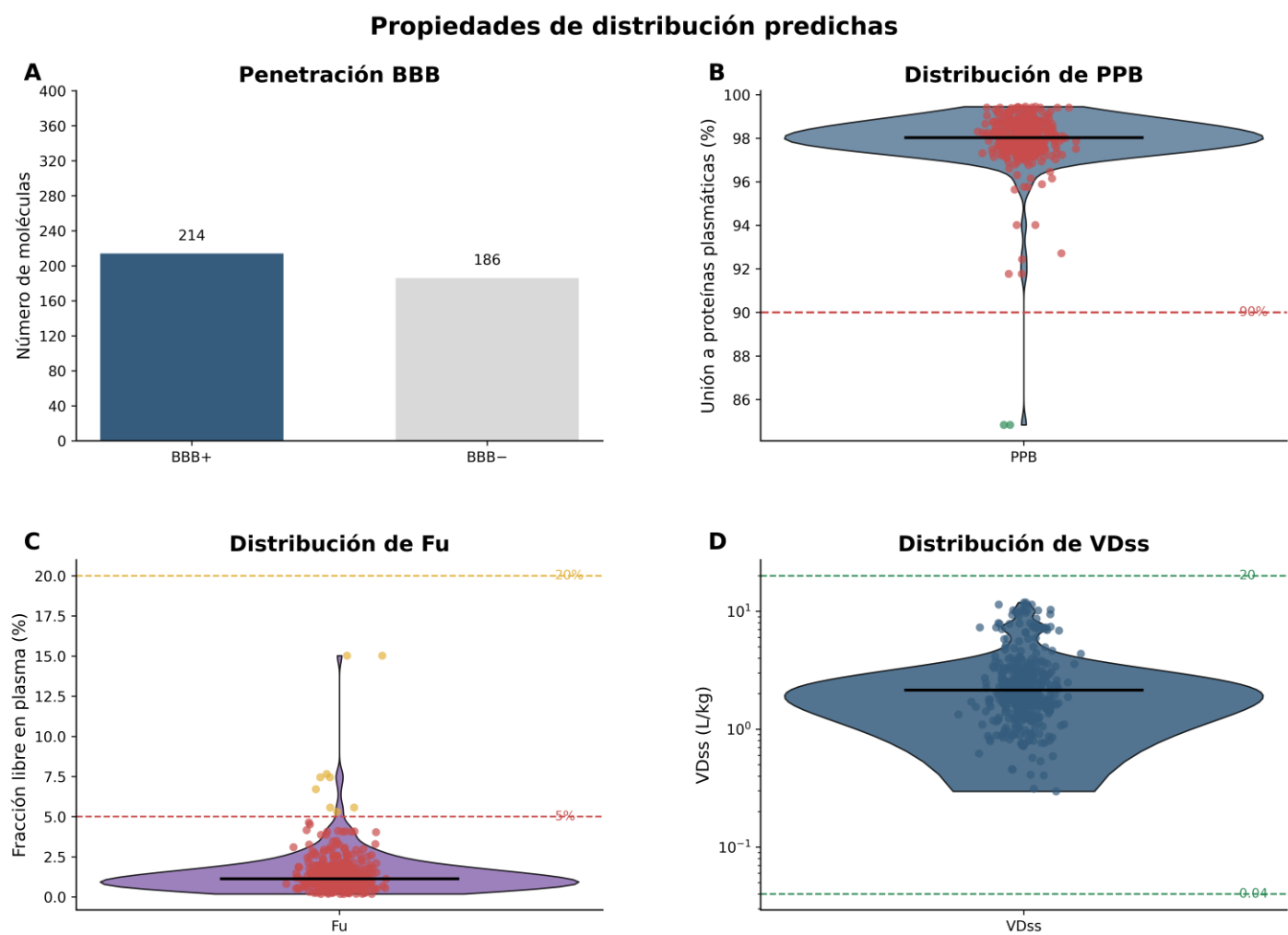


Figura 127 Distribución de las propiedades farmacocinéticas relacionadas con la distribución para el conjunto de compuestos analizados. (A) Clasificación de la permeabilidad a la barrera hematoencefálica (BBB), donde valores de probabilidad superiores a 0.5 se consideran indicativos de compuestos con capacidad de penetración al sistema nervioso central (BBB+), mientras que valores inferiores corresponden a compuestos con baja probabilidad de penetración (BBB-). (B) Distribución de la unión a proteínas plasmáticas (PPB), expresada como porcentaje. El diagrama de violín representa la densidad de los datos, mientras que los puntos individuales corresponden a cada molécula. La línea central indica la mediana de la distribución. La línea discontinua representa el umbral empírico de alta unión a proteínas plasmáticas (>90%). (C) Distribución de la fracción libre en plasma (Fu), expresada como porcentaje. El gráfico de violín muestra la dispersión de los valores predichos, donde la línea central corresponde a la mediana. Las líneas discontinuas indican valores de referencia empíricos para fracción libre baja (~5%) y moderada (~20%). (D) Distribución del volumen de distribución en estado estacionario (VDss), expresado en L/kg y representado en escala logarítmica. De acuerdo con el criterio empírico del modelo ADMETlab 3.0, valores entre 0.04 y 20 L/kg se consideran favorables (rango en verde), mientras que valores fuera de este rango se consideran desfavorables.

Con el fin de comparar las propiedades de distribución entre las distintas familias estructurales, se calculó el valor promedio de cada descriptor y se representó mediante un mapa de calor (Figura 128). En este análisis, la probabilidad de penetración BBB muestra variaciones entre familias, evidenciando diferencias en la capacidad de acceso al sistema nervioso central. Por su parte, los valores asociados a PPB favorable y fracción libre se mantienen bajos de manera consistente entre las distintas familias químicas. En contraste, el volumen de distribución (VDss) presenta variaciones moderadas entre familias, indicando diferencias en la tendencia de los compuestos a distribuirse en compartimentos tisulares, pero estos siempre dentro del rango óptimo.

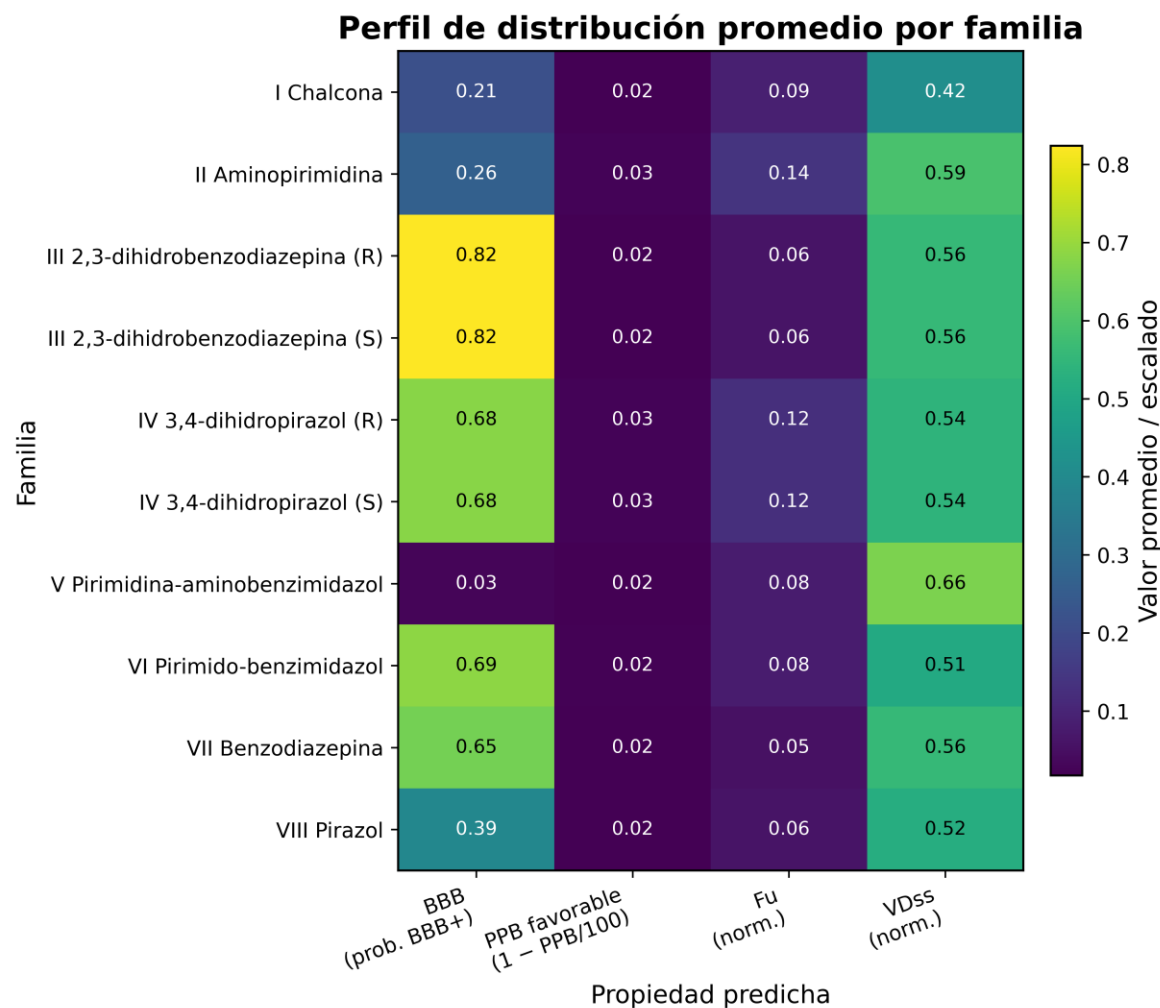


Figura 128 Mapa de calor que resume las propiedades de distribución predichas para las distintas familias estructurales analizadas. Las variables incluidas corresponden a la probabilidad de penetración a la barrera hematoencefálica (BBB), la unión a proteínas plasmáticas transformada como 1-PPB/100 (PPB favorable), la fracción libre en plasma (Fu) y el volumen de distribución en estado estacionario (VDss), estos últimos normalizados para facilitar la comparación visual. Los valores mostrados corresponden al promedio de cada descriptor dentro de cada familia química. La escala de color representa la magnitud relativa de cada propiedad.

5.6.1.4 Metabolismo

La interacción de las moléculas con enzimas del sistema citocromo p450 fue evaluada mediante modelos predictivos, considerando tanto la probabilidad de inhibición como de sustrato para las principales isoformas involucradas en el metabolismo de fármacos (CYP1A2, CYP2C9, CYP2C19, CYP2D6 y CYP3A4). La distribución de probabilidades de inhibición para cada isoforma se presenta en la Figura 129. En general, se observa una variabilidad en los valores predichos entre las distintas isoformas, evidenciando perfiles diferenciados de interacción en la serie de compuestos analizados. Algunas isoformas presentan distribuciones desplazadas hacia valores elevados de probabilidad, mientras que otras muestran una mayor dispersión o tendencia hacia valores intermedios.

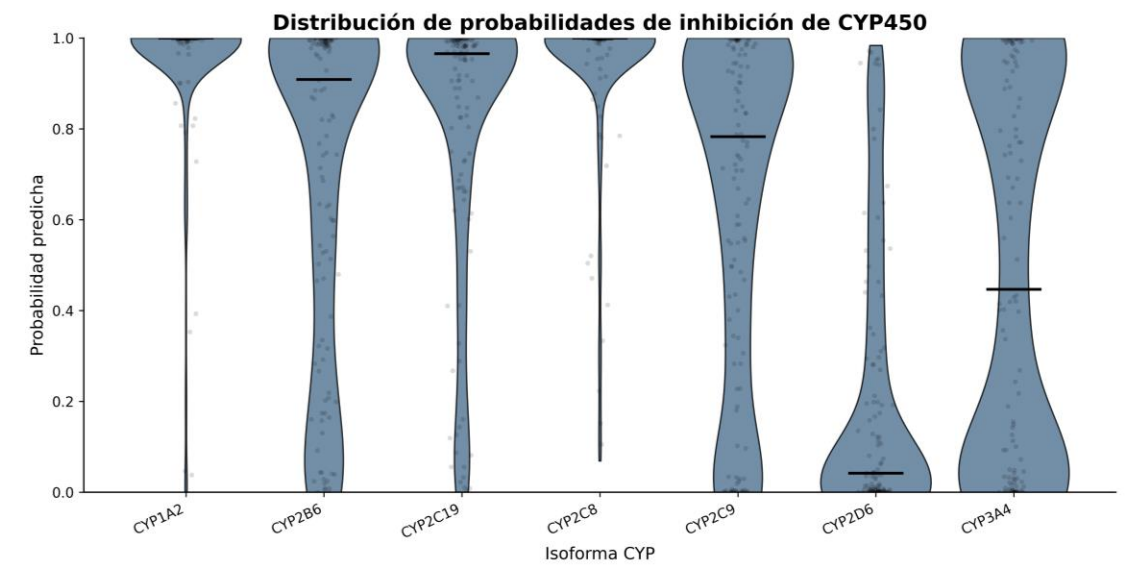


Figura 129 Distribución de las probabilidades predichas de inhibición de isoformas del sistema Citocromo P450 (CYP1A2, CYP2C9, CYP2C19, CYP2D6 y CYP3A4) para el conjunto de moléculas analizadas. Los diagramas tipo violín representan la densidad de distribución de las probabilidades en el rango de 0 a 1, mientras que la línea central indica la mediana de cada distribución. Los puntos individuales corresponden a valores muestreados para visualizar la dispersión de los datos.

En la Figura 130 se muestra la distribución de probabilidades asociadas al comportamiento como sustrato de las mismas isoformas. De manera análoga a lo observado para la inhibición, se identifican diferencias entre isoformas, con distribuciones que abarcan desde valores bajos hasta altos, indicando una heterogeneidad en la propensión de las moléculas a ser metabolizadas por estas enzimas.

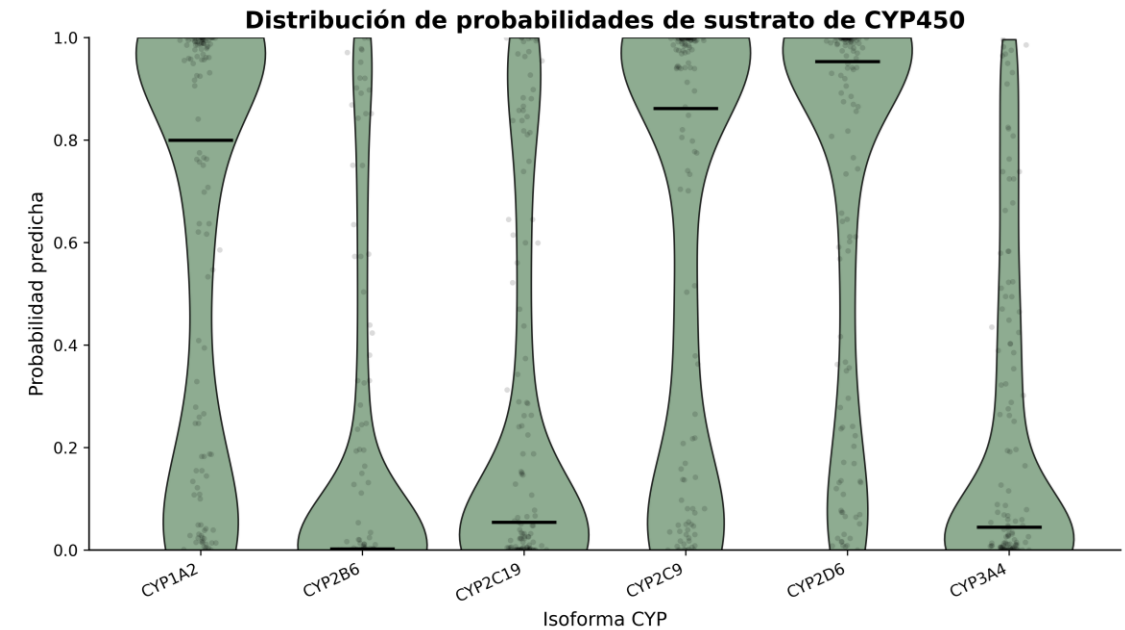


Figura 130 Distribución de las probabilidades predichas de comportamiento como sustrato para las principales isoformas del sistema Citocromo P450 (CYP1A2, CYP2C9, CYP2C19, CYP2D6 y CYP3A4). Los diagramas tipo violín muestran la densidad de los valores predichos en el rango de 0 a 1, mientras que la línea central representa la mediana de cada distribución. Los puntos superpuestos ilustran la variabilidad de los datos individuales.

Por otra parte, el perfil metabólico promedio agrupado por familia estructural se resume en la Figura 131. Este análisis permite visualizar tendencias globales en la interacción con las distintas isoformas de CYP450, observándose variaciones en los valores promedio de probabilidad tanto para inhibición como para sustrato entre las distintas familias.

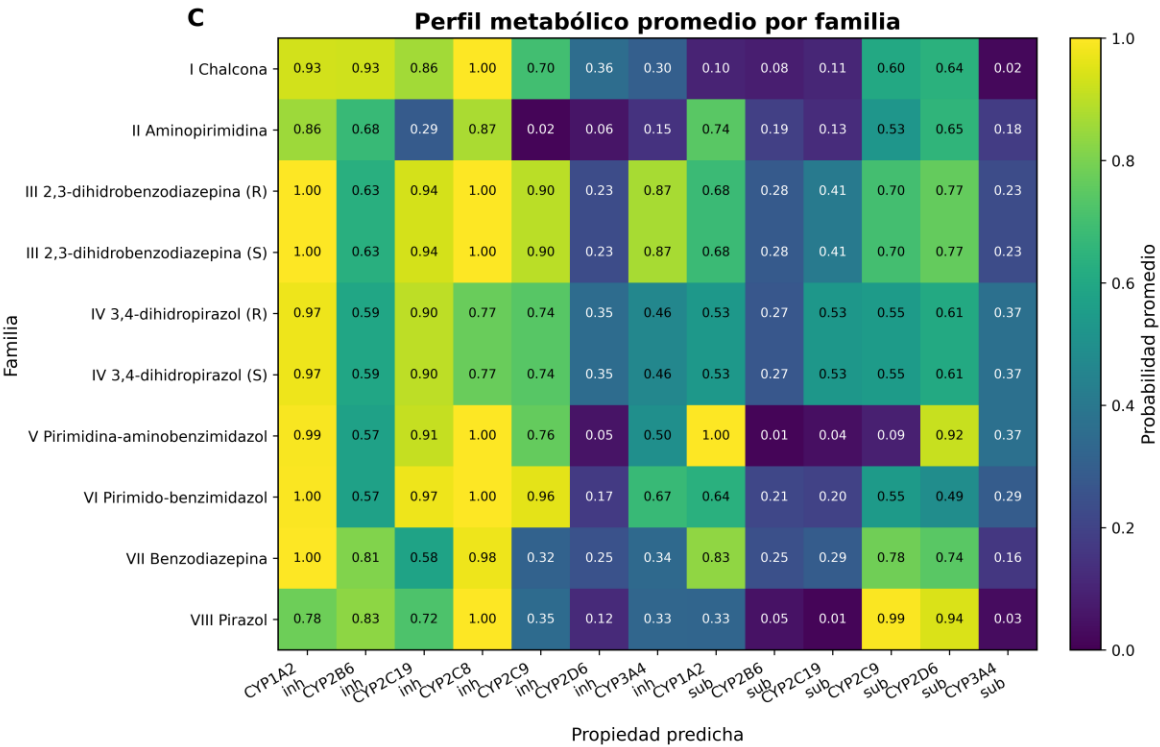


Figura 131 Perfil metabólico promedio por familia estructural, representado como un mapa de calor de las probabilidades predichas de inhibición (inh) y comportamiento como sustrato (sub) para las principales isoformas del sistema Citocromo P450. Los valores corresponden al promedio de las probabilidades dentro de cada familia, en una escala de 0 a 1. La intensidad de color refleja la magnitud de la probabilidad promedio, permitiendo comparar las tendencias metabólicas entre familias estructurales.

5.6.1.5 Excreción

Las propiedades relacionadas con la excreción de los compuestos fueron evaluadas mediante predicciones ADMET, considerando el clearance plasmático (CL) y la vida media ($t_{1/2}$) como parámetros principales.

La distribución global de estos parámetros se presenta en la Figura 132. En el caso del clearance plasmático, los valores predichos se distribuyen principalmente en un rango intermedio, con una dispersión que abarca desde valores bajos hasta cercanos al umbral superior del modelo. La mediana del conjunto se sitúa en torno al límite inferior de la categoría intermedia. Por su parte, la vida media presenta una distribución más concentrada, con valores predominantemente bajos y una menor dispersión en comparación con el clearance plasmático.

El análisis por familia química (Figura 133) evidencia diferencias en la distribución de ambos parámetros entre los distintos grupos estructurales. Para el clearance plasmático, las familias muestran rangos de valores diferenciados, con distribuciones que varían en amplitud y posición central. En particular, algunas familias presentan distribuciones centradas en valores intermedios, mientras que otras exhiben una mayor dispersión hacia valores superiores. En el caso de la vida media, las distribuciones se mantienen mayoritariamente en rangos bajos, con una variabilidad limitada entre familias.

El análisis de los valores promedio por familia química (Figura 134) permite observar tendencias globales en el comportamiento de los parámetros de excreción. El clearance plasmático presenta valores promedio diferenciados entre familias, destacando variaciones en la magnitud de este parámetro entre los distintos grupos estructurales. Por su parte, la vida media muestra valores promedio relativamente homogéneos entre familias, con diferencias de menor magnitud en comparación con el clearance plasmático.

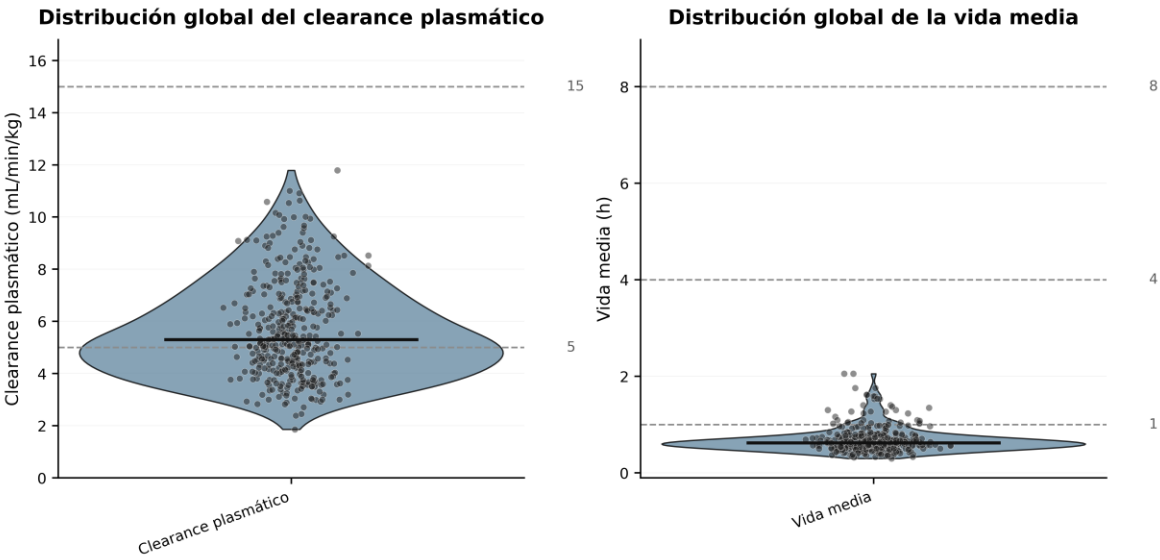


Figura 132 Distribución global de los parámetros de excreción. Diagramas tipo violín que muestran la distribución de los valores predichos de clearance plasmático (CL) y vida media (t_{1/2}) para el conjunto total de compuestos. Los puntos representan valores individuales. Las líneas discontinuas indican los umbrales empíricos del modelo ADMETlab para la clasificación de clearance plasmático (5 y 15 mL/min/kg) y vida media (1, 4 y 8 h).

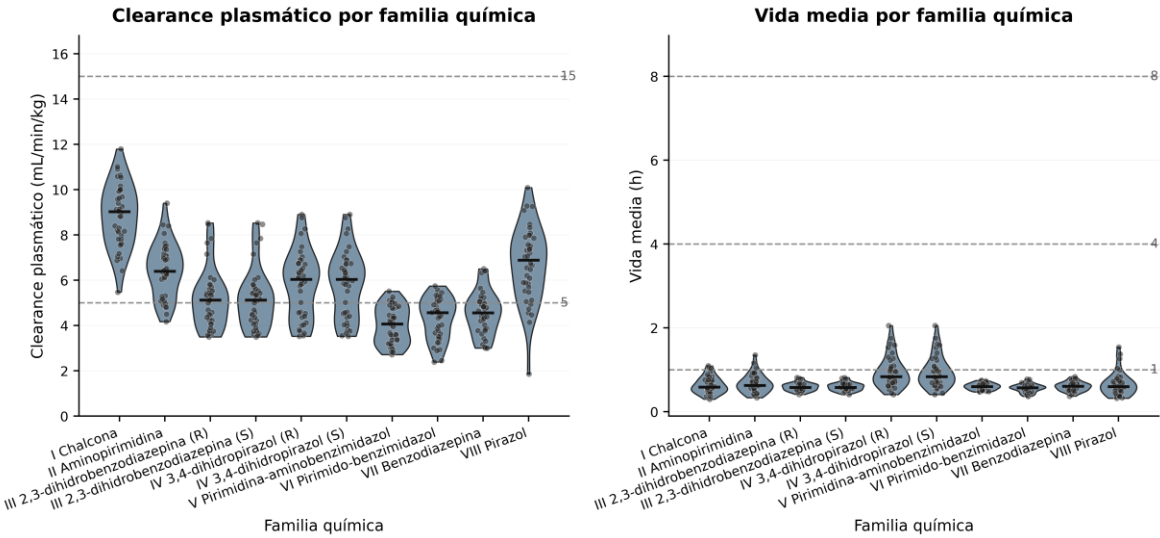


Figura 133 Distribución de los parámetros de excreción por familia química. Diagramas tipo violín que representan la distribución del clearance plasmático (CL) y la vida media (t_{1/2}) agrupados según la familia estructural de los compuestos. Los puntos corresponden a valores individuales. Las líneas discontinuas indican los umbrales empíricos del modelo ADMETlab para la clasificación de clearance plasmático (5 y 15 mL/min/kg) y vida media (1, 4 y 8 h).

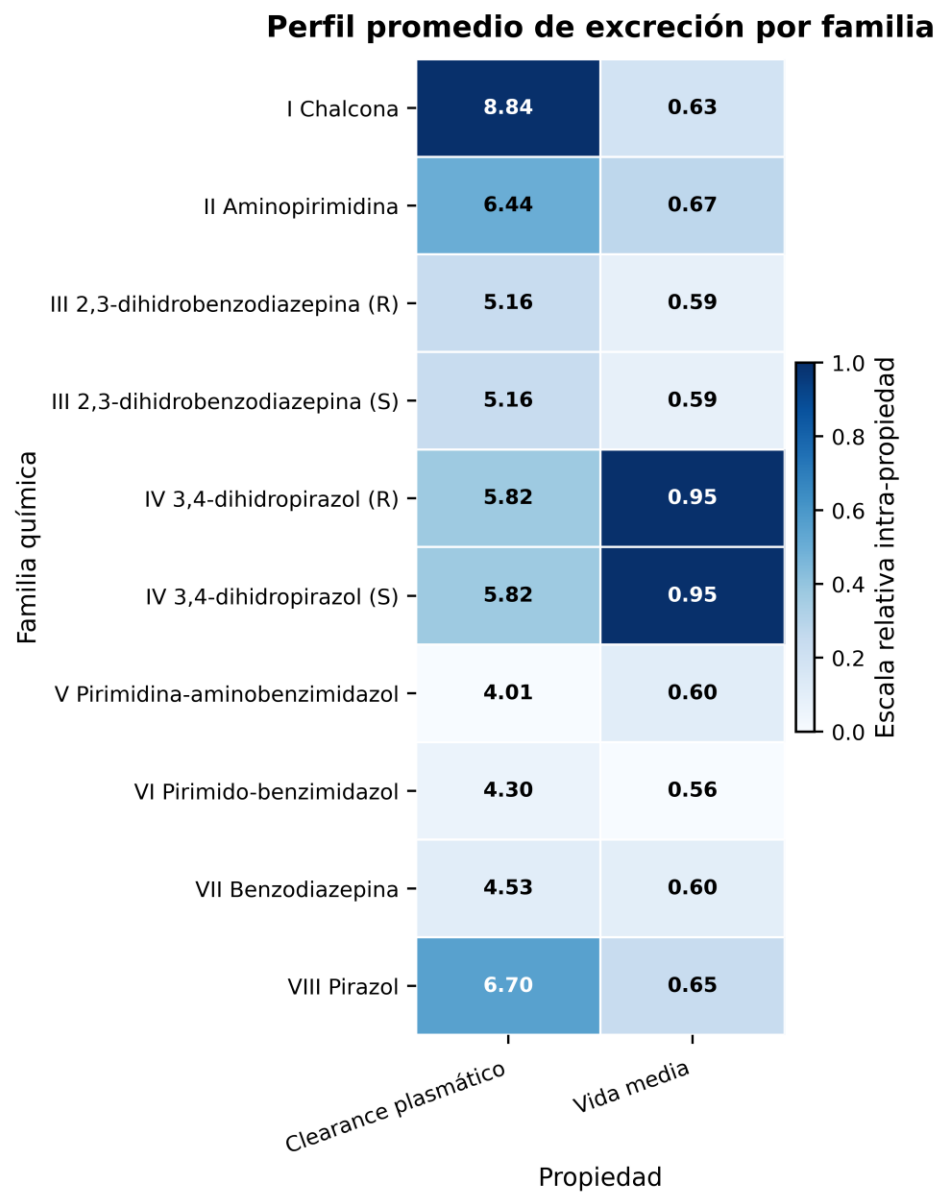


Figura 134 Perfil promedio de excreción por familia química. Mapa de calor que muestra los valores promedio de clearance plasmático (CL) y vida media ($t_{1/2}$) para cada familia química. La escala de color representa la magnitud relativa de los valores dentro de cada propiedad, mientras que los valores numéricos corresponden a los promedios calculados para cada grupo estructural.

5.6.1.6 Toxicidad

Los resultados predichos revelaron una alta prevalencia de eventos clasificados como positivos en múltiples endpoints toxicológicos, entendidos como efectos adversos específicos evaluados de manera independiente por el modelo (Figura 135). . Con el fin de evitar una clasificación estrictamente binaria, las probabilidades predichas se agruparon en tres categorías interpretativas: baja ($\leq 0,30$), intermedia ($0,30-0,70$) y alta ($> 0,70$) probabilidad de toxicidad. La distribución de estas categorías (Figura 135) revela que una proporción significativa del conjunto se concentra en el rango de alta probabilidad para varios efectos críticos. En particular, la hepatotoxicidad humana, la toxicidad respiratoria y el daño hepático inducido por fármacos (DILI) presentan las mayores fracciones de compuestos en la categoría de alto riesgo. De manera consistente, efectos como la

genotoxicidad, la inhibición del canal hERG y FDAMDD también muestran una clara dominancia de valores elevados, indicando un perfil toxicológico globalmente desfavorable. En contraste, algunos efectos como la corrosión ocular y la citotoxicidad en líneas celulares específicas presentan una mayor proporción de compuestos en las categorías de baja o intermedia probabilidad, lo que sugiere una menor contribución relativa de estos mecanismos al riesgo global del conjunto.

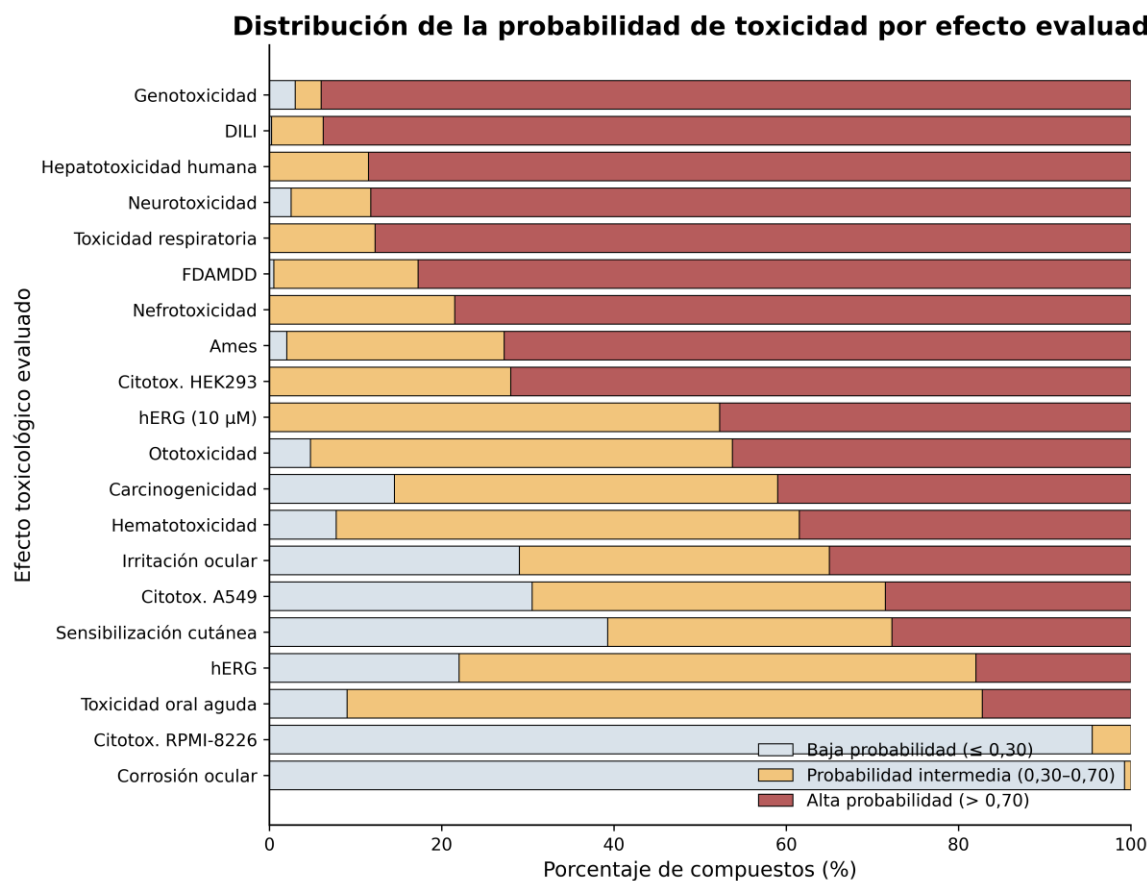


Figura 135 Clasificación toxicológica por endpoint. Distribución porcentual de compuestos clasificados como positivos y negativos para los distintos endpoints toxicológicos predichos por ADMETlab 3.0. La clasificación se realizó considerando baja probabilidad (≤ 0.30), probabilidad intermedia (0.30–0.70) y alta probabilidad (> 0.70). Los endpoints incluidos corresponden a probabilidades de positividad/toxicidad reportadas por el modelo.

El análisis del perfil promedio por familia química (Figura 136) revela diferencias estructurales marcadas en la propensión a la toxicidad. La familia pirimidina–aminobenzimidazol destaca por presentar valores consistentemente elevados en múltiples efectos evaluados, indicando un perfil de riesgo pronunciado y generalizado. Por el contrario, las chalconas exhiben valores relativamente más moderados en varios

parámetros, aunque aún mantienen niveles intermedios a altos en efectos críticos como la inhibición de hERG y la carcinogenicidad, lo que sugiere que ninguna familia está completamente exenta de riesgo.

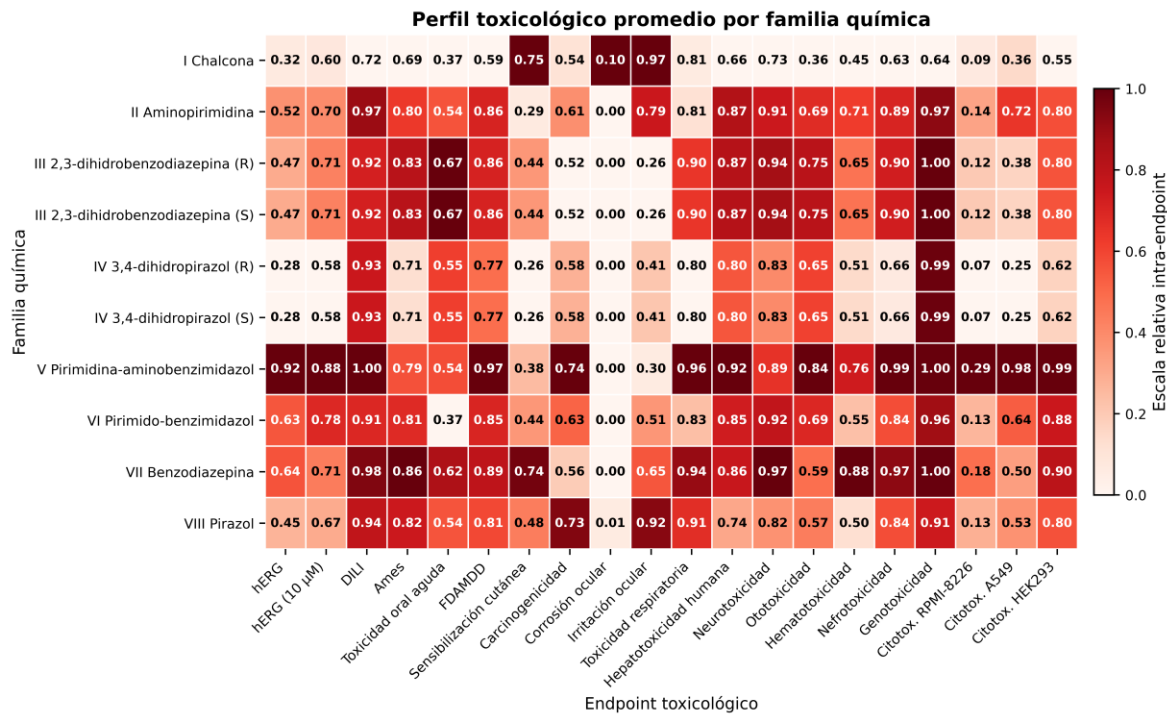


Figura 136 Perfil toxicológico promedio por familia química. Mapa de calor que representa los valores promedio de probabilidad para los distintos endpoints toxicológicos agrupados según la familia estructural de los compuestos. La escala de color muestra la magnitud relativa dentro de cada endpoint, mientras que los valores numéricos corresponden a los promedios observados en cada familia.

La distribución global de la carga toxicológica (Figura 137) definida como el número de efectos con probabilidad superior a 0,5 por molécula, muestra que la mayoría de los compuestos se agrupa entre 13 y 16 efectos positivos, con una mediana cercana a 14. Esta acumulación de señales adversas indica que las moléculas evaluadas presentan un perfil de riesgo multifactorial, caracterizado por la coexistencia de múltiples alertas toxicológicas simultáneas.

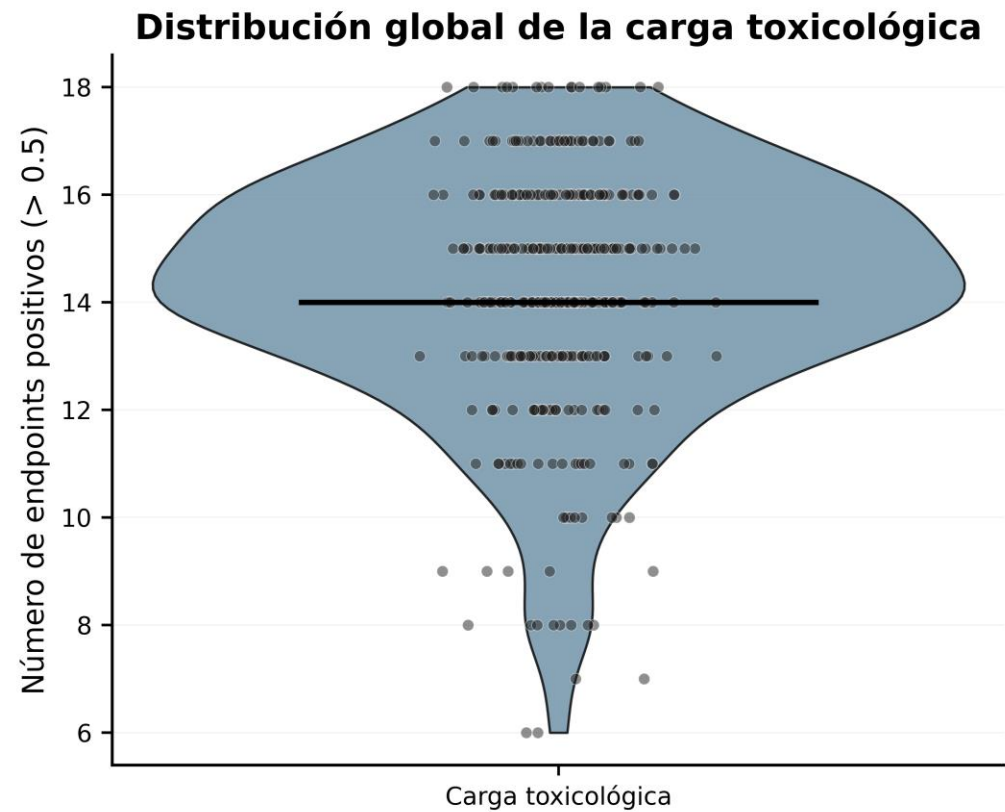


Figura 137 Distribución de la carga toxicológica global del conjunto de compuestos, definida como el número de efectos toxicológicos con probabilidad predicha superior a 0,5 por molécula. El diagrama de violín representa la densidad de la distribución, mientras que los puntos individuales corresponden a cada compuesto y la línea horizontal indica la mediana. La concentración de valores entre 13 y 16 efectos positivos sugiere una elevada acumulación de alertas toxicológicas en el conjunto analizado.

Los principales resultados evidencian que el perfil toxicológico del conjunto está dominado por una alta probabilidad de eventos adversos en múltiples sistemas biológicos, destacando especialmente aquellos

asociados a cardiotoxicidad, hepatotoxicidad y genotoxicidad. Estas tendencias sugieren que, si bien las moléculas presentan potencial farmacológico, la optimización estructural orientada a reducir su perfil toxicológico resulta necesaria para su eventual desarrollo. estructural será necesaria para reducir el riesgo toxicológico global.

5.6.1.7Solubilidad

La muestra el panorama general de los valores de logS (definido como el logaritmo decimal de la solubilidad molar en agua. En este parámetro, valores más negativos indican menor solubilidad acuosa, mientras que valores menos negativos reflejan una mayor solubilidad) predichos para la biblioteca completa(Figura 138). Los valores se distribuyeron entre -9.415 y -2.170, con una mediana global de -5.726. En conjunto, la mayor parte de los compuestos se ubicó por debajo del umbral de referencia de logS = -4, y solo 57 de 400 moléculas (14.3%) presentaron valores iguales o superiores a dicho valor. La distribución ordenada evidenció además un descenso progresivo de logS a lo largo del conjunto, sin agrupamientos discretos marcados en el panorama global.

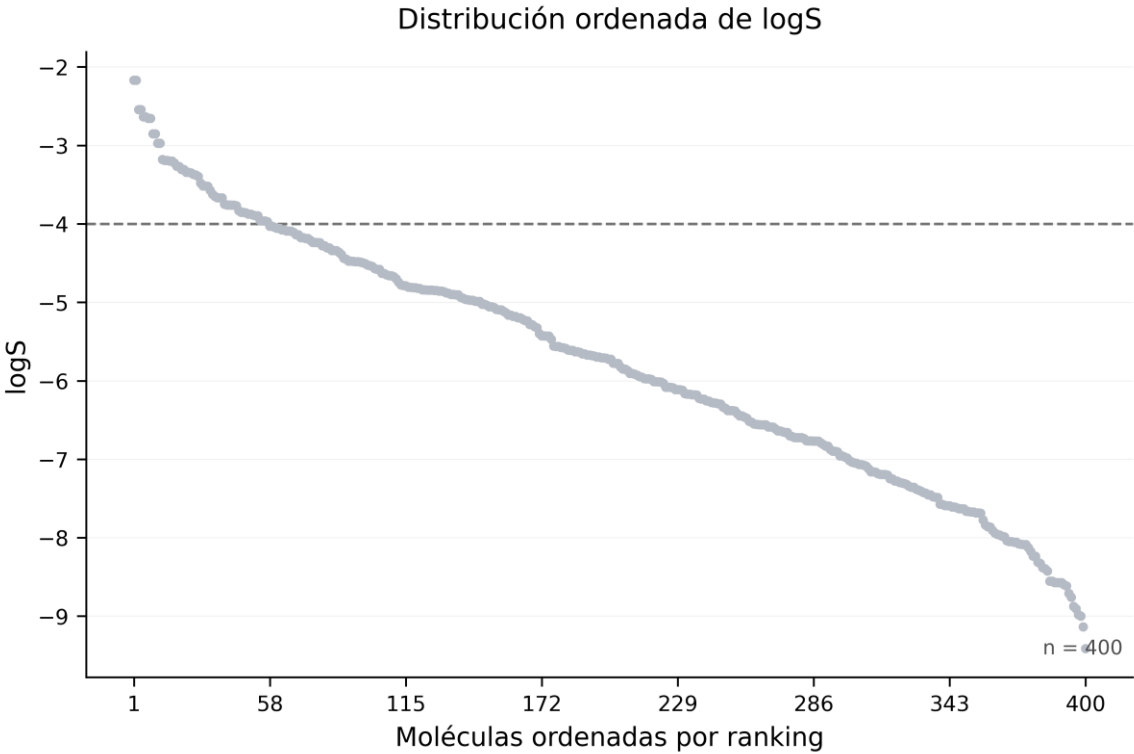


Figura 138 Distribución ordenada de los valores de logS predichos para la biblioteca evaluada. Los compuestos se presentan ordenados de acuerdo con su valor de logS, desde el mayor al menor. La línea horizontal punteada indica el umbral de referencia de logS = -4.

La distribución de logS por familia se presenta en la Figura 139. Las medianas menos negativas correspondieron a las familias IV 3,4-dihidropirazol (R) y IV 3,4-dihidropirazol (S), ambas con una mediana de -5.000. A continuación se ubicaron II Aminopirimidina (-5.515) y las familias III 2,3-dihidrobenzodiazepina (R) y (S) (-5.588 en ambos casos). Las familias con medianas más negativas fueron VI Pirimido-benzimidazol (-6.535) y V Pirimidina-aminobenzimidazol (-6.856). En general, las distribuciones mostraron desplazamientos diferenciales entre familias, con una tendencia hacia valores más negativos en los sistemas benzimidazólicos fusionados.

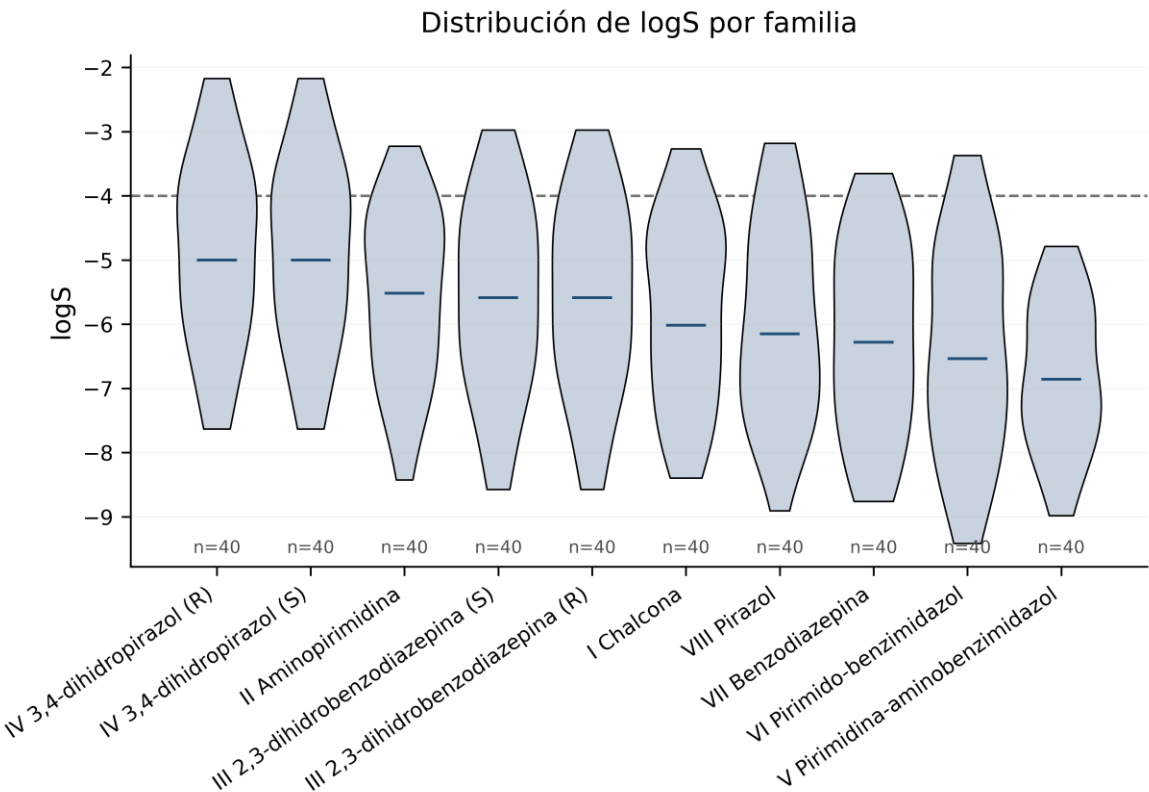


Figura 139 Distribución de logS por familia química. Se muestran las distribuciones de los valores de logS predichos para cada familia química mediante gráficos de violín, ordenados según la mediana de logS. La línea horizontal punteada indica el umbral de referencia de logS = -4.

La Figura 140 muestra la distribución de logS según el aldehído empleado. Las medianas menos negativas se observaron para los derivados de p-aminobenzaldehído (-4.780), salicilaldehído (-4.835) y vainillina (-4.893). En una zona intermedia se ubicaron los compuestos derivados de p-dimetilaminobenzaldehído (-5.157) y cinamalaldehído (-5.672). Las medianas más negativas correspondieron a p-nitrobenzaldehído (-6.207), 4-isopropilbenzaldehído (-6.322) y, especialmente, p-(trifluorometil)benzaldehído (-7.057). En conjunto, el cambio del aldehído se asoció con desplazamientos notorios en la distribución de logS.

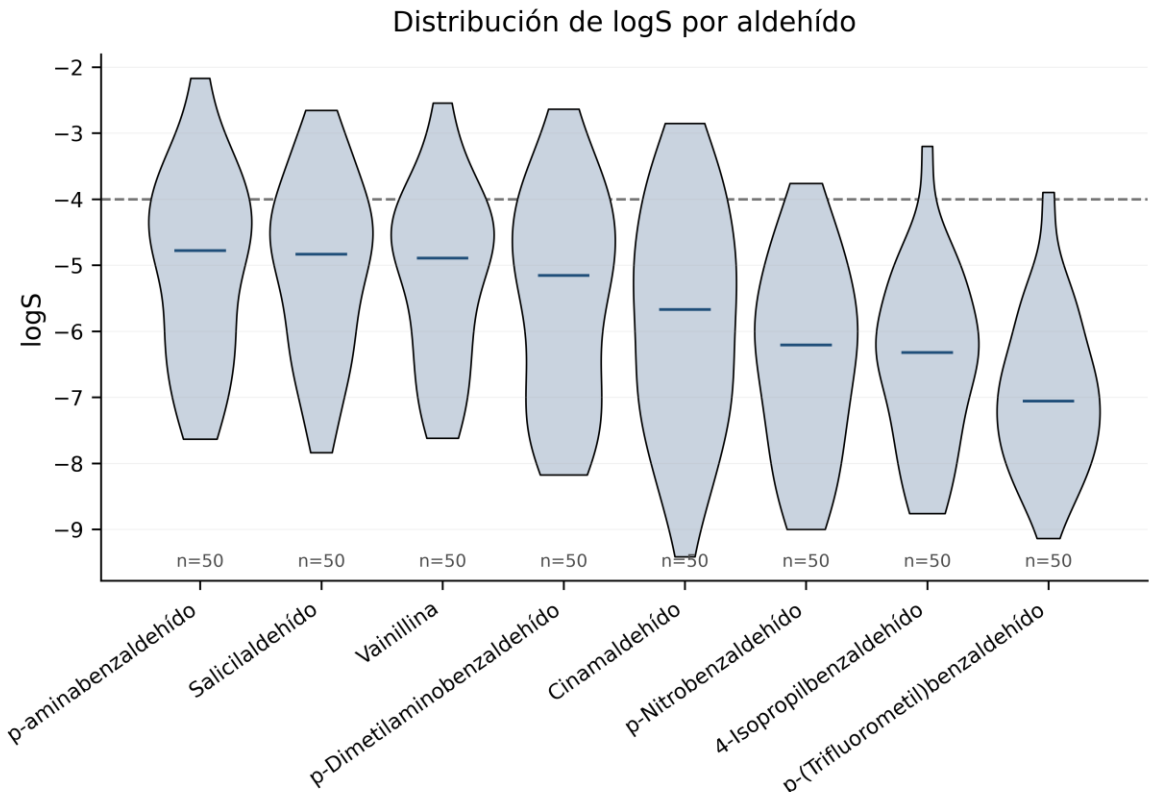


Figura 140 Distribución de logS según el aldehído empleado. Se presentan las distribuciones de logS agrupadas de acuerdo con el aldehído precursor, mediante gráficos de violín ordenados según la mediana de cada grupo. La línea horizontal punteada indica el umbral de referencia de logS = -4.

La distribución de logS por cetona se presenta en la Figura 141. Los valores menos negativos se observaron para los derivados de acetofenona, con una mediana de -3.802, siendo además la única cetona cuya mediana se ubicó por encima del umbral de -4. Le siguieron 1-acetilnaftaleno (-5.026) y 2-acetilnaftaleno (-5.140). Las medianas más negativas correspondieron a 9-acetilantraceno (-6.575) y 1-acetilpireno (-7.614). Así, la variación del fragmento cetónico produjo un gradiente claro en los valores de logS, desde sistemas con medianas próximas o superiores al umbral hasta series desplazadas hacia valores marcadamente más negativos.

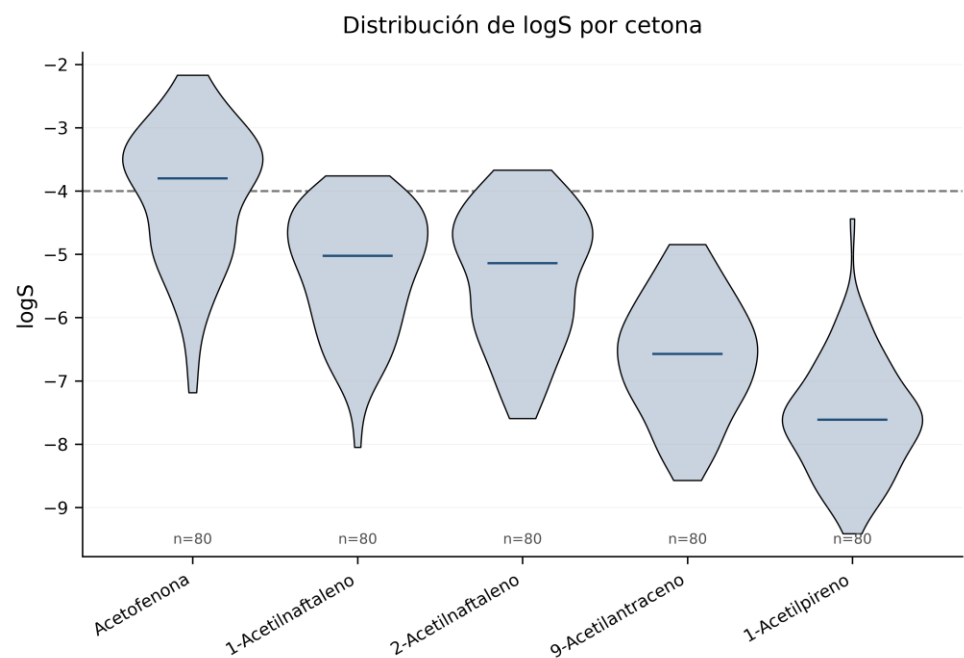


Figura 141 Distribución de logS según la cetona empleada. Se muestran las distribuciones de logS agrupadas según la cetona precursora, mediante gráficos de violín ordenados de acuerdo con la mediana de cada grupo. La línea horizontal punteada indica el umbral de referencia de logS = -4.

5.6.2 Derivados de benzimidazol

5.6.2.1 Drogabilidad (drug like)

Se evaluó el carácter drug-like de los seis benzimidazoles mediante la regla de Lipinski y el criterio de Veber. Los valores individuales de peso molecular, lipofilicidad, donadores y aceptores de enlace de hidrógeno, superficie polar topológica y enlaces rotables (Tabla 11). Los seis compuestos cumplen completamente la regla de Lipinski, ya que ninguno presenta violaciones. De manera consistente, todos exhiben un peso molecular de 349.12 Da, valores de logP entre 3.38 y 3.67, 0 donadores y 4 aceptores de enlace de hidrógeno. Estos valores se mantienen dentro de los umbrales clásicos asociados a una adecuada biodisponibilidad oral. En cuanto al criterio de Veber, la totalidad de la serie también cumple con los parámetros establecidos. Todos los compuestos presentan una TPSA de 47.78 Å² y 3 enlaces rotables, valores claramente por debajo de los límites empíricos de 140 Å² y 10 enlaces rotables, respectivamente. En conjunto, estos resultados muestran una alta homogeneidad fisicoquímica dentro de la serie y un perfil globalmente favorable de drogabilidad.

Tabla 11 Descripciones de benzimidazoles para reglas de Lipinsky y Veber(Sí=cumple con las reglas, No= No cumple con las reglas).

Molécula	MM	logP	HBD	HBA	Violaciones de Lipinsky	TPSA	nRot	Veber
Benzimidazol 1	349.12	3.48773818	0	4	0	47.78	3	Sí
Benzimidazol 2	349.12	3.63685682	0	4	0	47.78	3	Sí
Benzimidazol 3	349.12	3.3769211	0	4	0	47.78	3	Sí
Benzimidazol 4	349.12	3.52631881	0	4	0	47.78	3	Sí
Benzimidazol 5	349.12	3.50366277	0	4	0	47.78	3	Sí

Benzimidazol 6	349.12	3.66982531	0	4	0	47.78	3	Sí
-------------------	--------	------------	---	---	---	-------	---	----

5.6.2.2 Absorción

La clasificación empírica de absorción intestinal humana predicha (HIA) indica que los seis compuestos se ubican en la categoría favorable. Los valores predichos de HIA se encuentran entre 0.0002 y 0.0029, muy por debajo del umbral de 0.30 asociado a absorción intestinal favorable(Figura 142A).

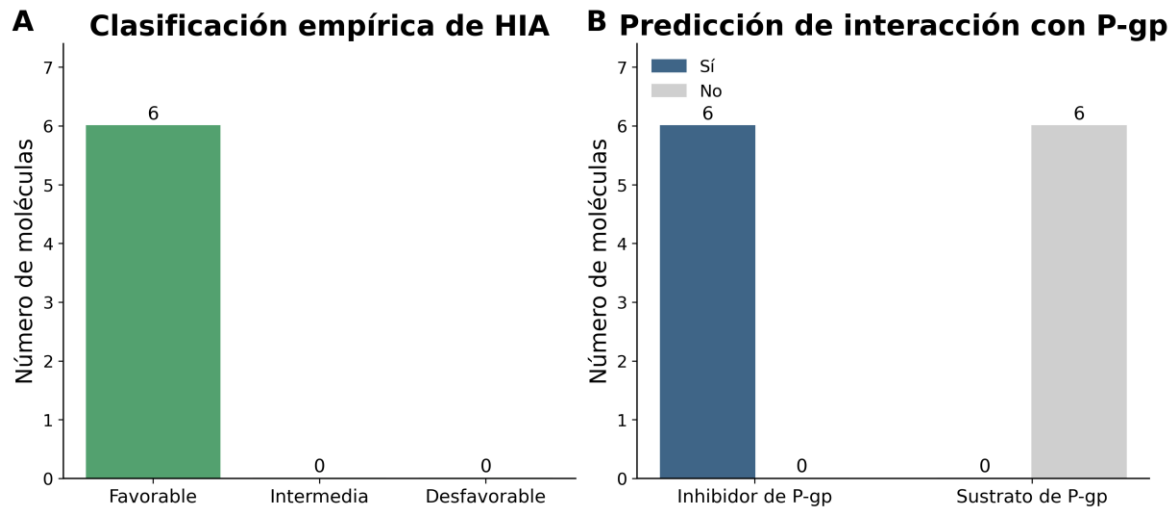


Figura 142 Propiedades de absorción intestinal e interacción con P-glicoproteína predichas para la serie de benzimidazoles. (A) Clasificación de las moléculas según la probabilidad de absorción intestinal humana (HIA). Valores entre 0–0.30 corresponden a absorción favorable, entre 0.30–0.70 a absorción intermedia y mayores a 0.70 a absorción desfavorable. (B) Distribución de compuestos predichos como inhibidores y sustratos de P-glicoproteína (P-gp), empleando un umbral de probabilidad de 0.5 para la clasificación binaria.

En relación con la interacción con P-glicoproteína, todos los compuestos fueron clasificados como potenciales inhibidores de P-gp, ya que presentaron probabilidades superiores a 0.5, con un promedio de 0.95. En contraste, ninguno fue predicho como sustrato de esta proteína, con probabilidades bajas y homogéneas (promedio 0.011) (Figura 142 B)

La distribución de la permeabilidad predicha en el modelo Caco-2 (Figura 143)muestra valores comprendidos entre –4.55 y –4.14 log cm/s, con una mediana de –4.32 log cm/s. Considerando el umbral empírico de –5.15 log cm/s, todos los benzimidazoles se sitúan en la región asociada a alta permeabilidad intestinal, sin registrarse compuestos en la categoría de baja permeabilidad.

Distribución de la permeabilidad Caco-2

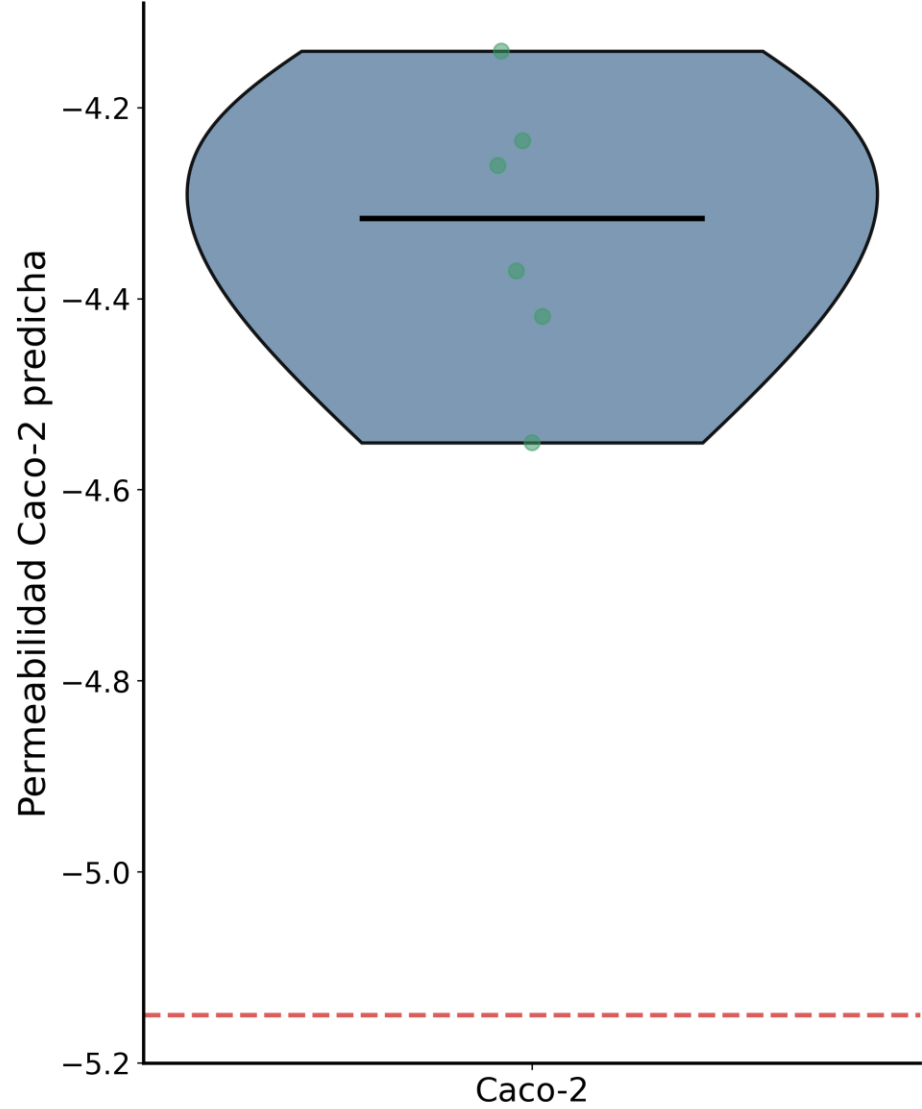


Figura 143 Distribución de la permeabilidad Caco-2 predicha para la serie de benzimidazoles. El diagrama de violín representa la densidad de la distribución y los puntos individuales corresponden a cada compuesto evaluado. La línea negra indica la mediana. Los compuestos con valores superiores a $-5.15 \log \text{ cm/s}$ se consideran de alta permeabilidad, mientras que la línea roja discontinua indica el umbral empírico asociado a una adecuada absorción intestinal.

El perfil promedio (Figura 144) de absorción de la serie evidencia un comportamiento muy favorable. La transformación $1 - \text{HIA}$ presenta un valor cercano a 1.00, lo que refleja la consistencia de la absorción intestinal predicha. Del mismo modo, la probabilidad promedio de inhibición de P-gp es alta, mientras que la probabilidad de sustrato de P-gp se mantiene cercana a cero. En conjunto, estos resultados indican que la serie posee un perfil de absorción uniforme y favorable, con escasa dispersión entre compuestos.

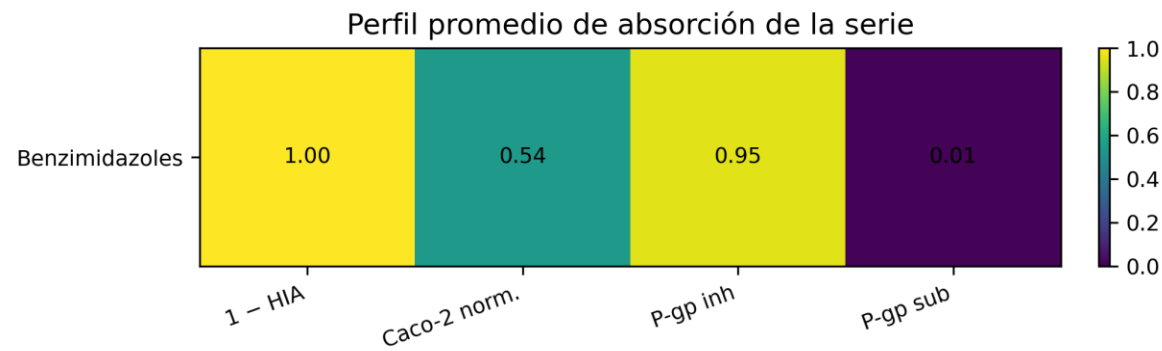


Figura 144 Mapa de calor que resume las propiedades de absorción predichas para el conjunto analizado. Las propiedades incluidas corresponden a $1 - \text{HIA}$, permeabilidad Caco-2 normalizada, probabilidad de inhibición de P-glicoproteína y probabilidad de actuar como sustrato de P-glicoproteína. Los valores corresponden al promedio de la serie.

5.6.2.3Distribución

La caracterización de las propiedades de distribución de la serie de benzimidazoles se presenta en la Figura 145. En relación con la penetración a la barrera hematoencefálica (BBB), la serie muestra una distribución equilibrada, con tres compuestos clasificados como BBB+ y tres como BBB-, utilizando un umbral de 0.5 (Figura 145A). Los valores individuales se encuentran en el rango de 0.36 a 0.60. La unión a proteínas plasmáticas (PPB) presenta valores consistentemente altos en toda la serie, comprendidos entre 97.08 y 97.94%, con una mediana de 97.53% (Figura 145B). En concordancia, la fracción libre en plasma (Fu) se mantiene baja, con valores entre 1.92 y 2.78% y una mediana de 2.14% (Figura 145C). Por su parte, el volumen de distribución en estado estacionario (VDss) muestra valores de logVDss entre -0.28 y 0.12, con un promedio de -0.09 (Figura 145D).

Propiedades de distribución predichas para la serie de benzimidazoles

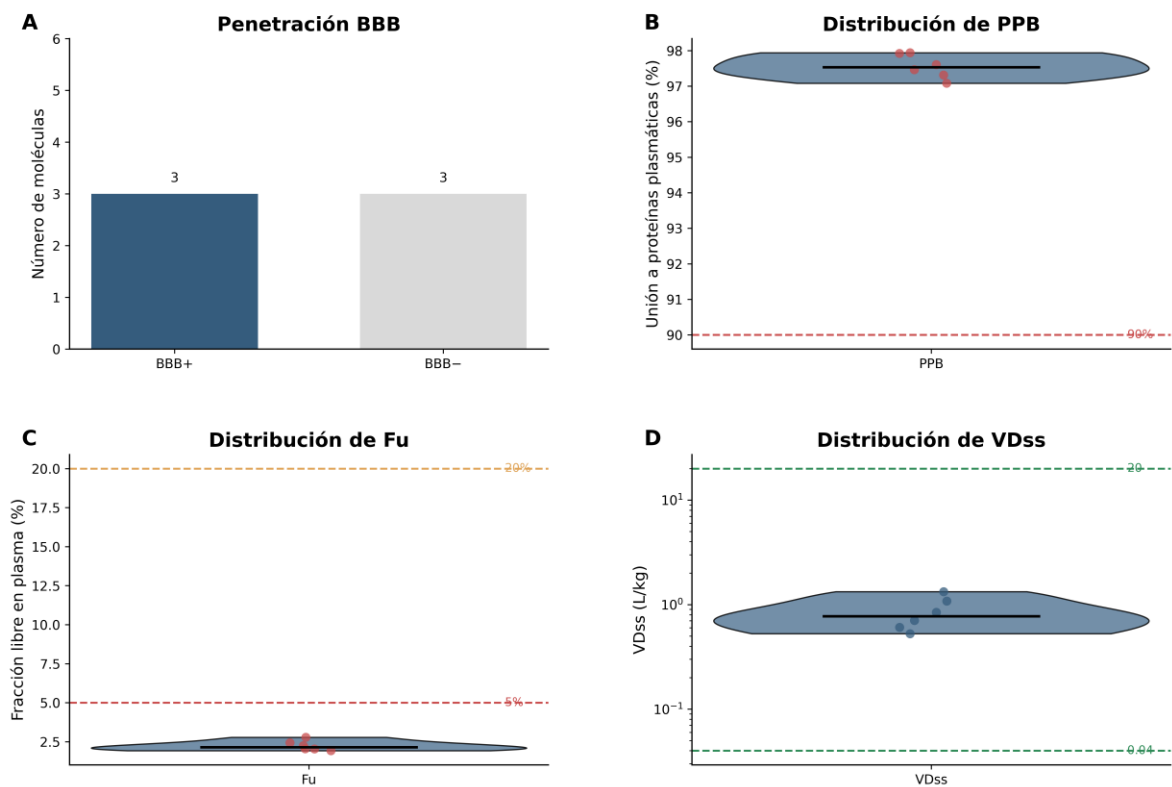


Figura 145 Propiedades de distribución predichas para la serie de benzimidazoles.(A) Clasificación de penetración a la barrera hematoencefálica (BBB) basada en un umbral de probabilidad de 0.5, mostrando el número de moléculas clasificadas como BBB+ y BBB-.(B) Distribución de la unión a proteínas plasmáticas (PPB, %). La línea negra representa la mediana y la línea discontinua roja indica el umbral de referencia del 90%.(C) Distribución de la fracción libre en plasma (Fu, %). La línea negra representa la mediana, mientras que las líneas discontinuas indican valores de referencia en 5% y 20%.(D) Distribución del volumen de distribución en estado estacionario (VDss, L/kg) en escala logarítmica. La línea negra indica la mediana y las líneas discontinuas verdes delimitan el rango de referencia (0.04–20 L/kg). Los puntos representan los valores individuales de cada compuesto.

El perfil promedio de distribución de la serie (Figura 146) resume estos parámetros, evidenciando valores relativos de 0.48 para BBB, 0.02 para PPB favorable, 2.24 para Fu y 0.40 para VDss normalizado.

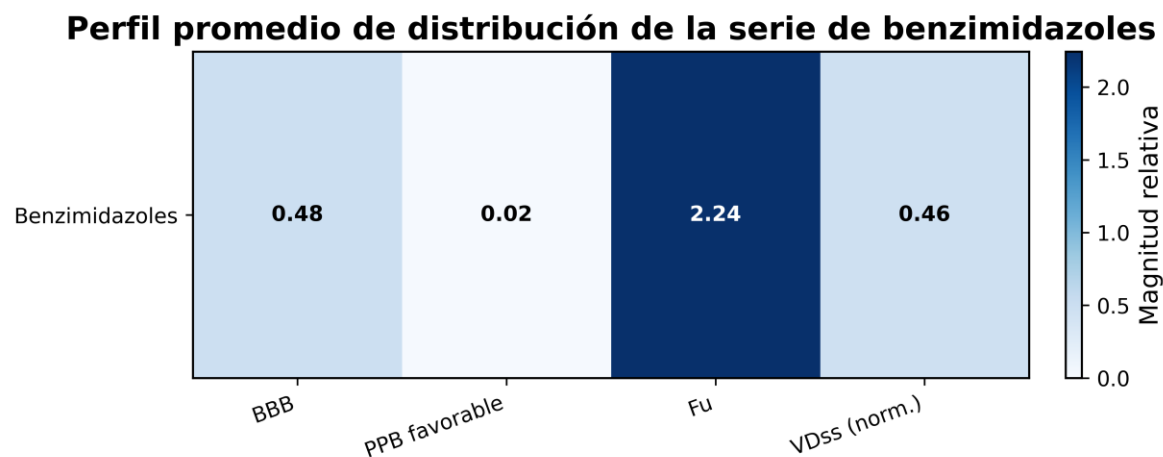


Figura 146 Perfil promedio de distribución de la serie de benzimidazoles. El mapa de calor muestra los valores promedio normalizados de los parámetros de distribución: penetración BBB, proporción de PPB favorable, fracción libre en plasma (Fu) y volumen de distribución normalizado (VDss). La intensidad del color corresponde al valor relativo de cada parámetro dentro de la serie. Los valores numéricos indicados en cada celda representan los promedios calculados para cada descriptor.

5.6.2.4 Metabolismo

El análisis del potencial de inhibición de isoformas del citocromo P450 muestra valores elevados para CYP1A2, CYP2C19 y CYP2C8, con probabilidades promedio cercanas a 1.00, mientras que CYP2C9 presenta valores altos (0.92) y CYP3A4 moderados (0.78) (Figura 147). En contraste, CYP2B6 y CYP2D6 exhiben probabilidades de inhibición considerablemente menores, con valores promedio de 0.24 y 0.15, respectivamente.

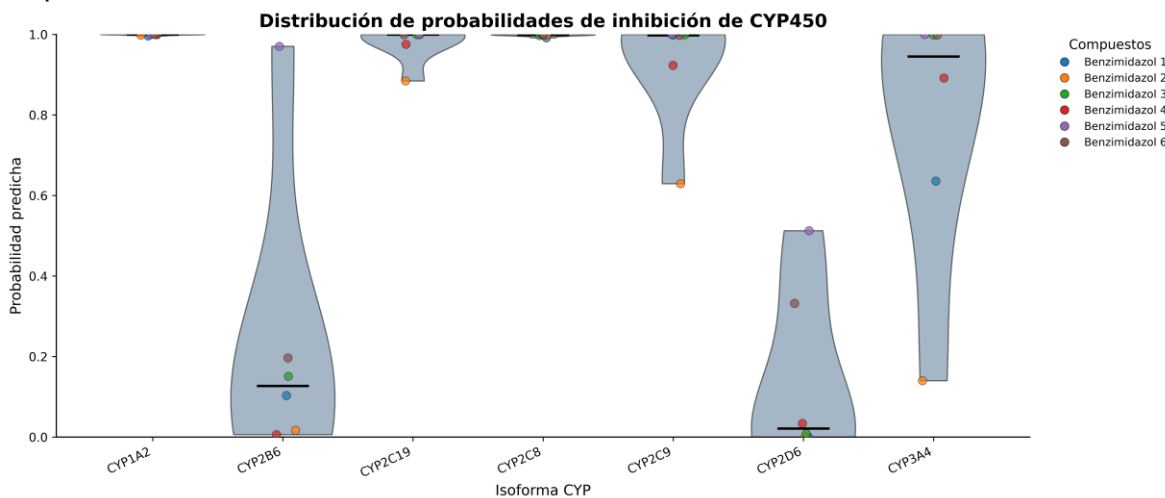


Figura 147 Distribución de probabilidades de inhibición de isoformas del citocromo P450 para la serie de benzimidazoles. Se muestran las distribuciones de las probabilidades predichas de inhibición para las isoformas CYP1A2, CYP2B6, CYP2C19, CYP2C8, CYP2C9, CYP2D6 y CYP3A4. Los violines representan la densidad de distribución de los datos, los puntos corresponden a los valores individuales de cada compuesto y la línea horizontal negra indica la mediana.

En cuanto al comportamiento como sustrato, se observa una mayor variabilidad entre isoformas (Figura 148). CYP2C9 presenta las probabilidades más altas, con valores entre aproximadamente 0.40 y 0.63 y una mediana cercana a 0.43. CYP2D6 muestra valores intermedios, distribuidos entre 0.03 y 0.62, con una mediana cercana a 0.18. Por su parte, CYP1A2 presenta valores bajos a moderados (0.01–0.39; mediana ~0.24), mientras que CYP2B6, CYP2C19 y CYP3A4 presentan probabilidades cercanas a cero en toda la serie.

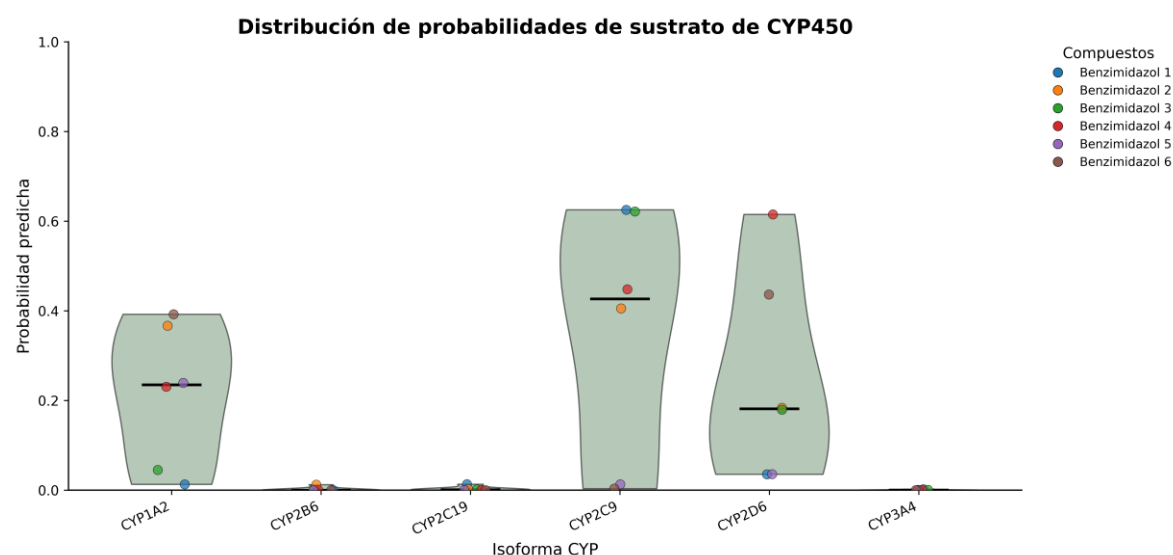


Figura 148 Distribución de probabilidades de comportamiento como sustrato de isoformas del citocromo P450 para la serie de benzimidazoles. Se presentan las distribuciones de las probabilidades predichas de ser sustrato para las isoformas CYP1A2, CYP2B6, CYP2C19, CYP2C9, CYP2D6 y CYP3A4. Los violines muestran la distribución de los datos, los puntos representan los valores individuales y la línea horizontal negra indica la mediana.

El perfil promedio integrado (Figura 149) resume estos resultados, evidenciando una alta probabilidad de inhibición para múltiples isoformas del sistema CYP450, particularmente CYP1A2, CYP2C19 y CYP2C8, junto con una contribución predominante de CYP2C9 y, en menor medida, CYP2D6 en el comportamiento como sustrato.

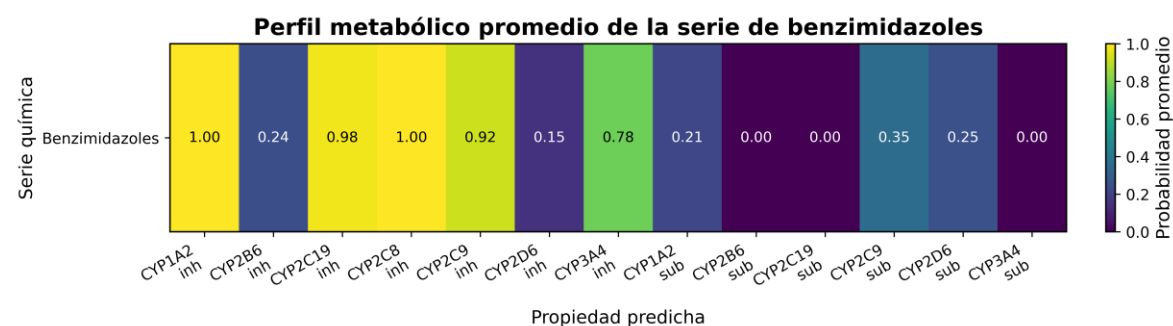


Figura 149 Perfil metabólico promedio de la serie de benzimidazoles. El mapa de calor muestra las probabilidades promedio de inhibición (inh) y comportamiento como sustrato (sub) para distintas isoformas del citocromo P450. La intensidad del color corresponde al valor de probabilidad promedio, según la escala indicada. Los valores numéricos dentro de cada celda representan los promedios calculados para cada propiedad.

5.6.2.5 Excreción

La distribución global de estos parámetros se presenta en la Figura 150. En el caso del clearance plasmático, los valores predichos se distribuyen en un rango bajo. La mediana del conjunto se centra es ente mismo rango. Por su parte, la vida media presenta una distribución más concentrada, con valores bajos y vidas medias menor a 1h.

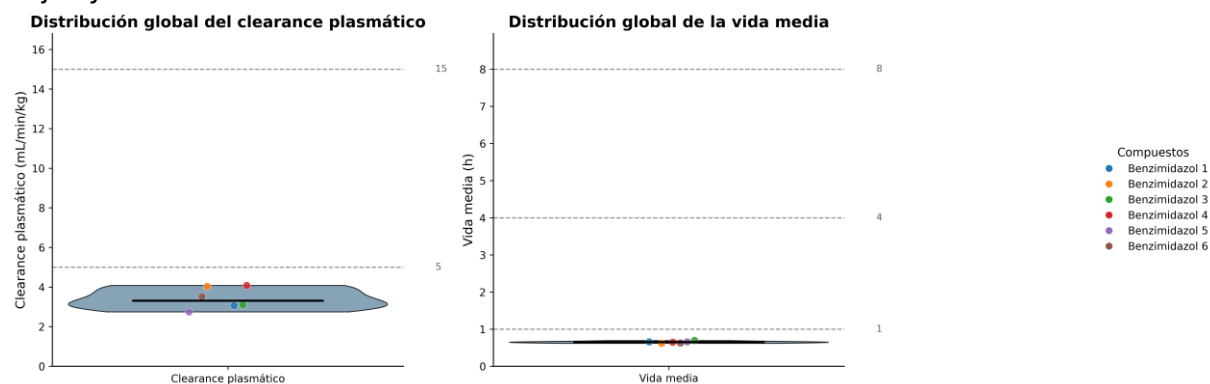


Figura 150 Distribución global de los parámetros de excreción predichos para la serie de benzimidazoles. Los gráficos de violín representan la densidad de los valores de clearance plasmático (mL/min/kg) y vida media (h), incluyendo la mediana (línea negra) y los valores individuales (puntos). Las líneas horizontales discontinuas indican umbrales de referencia para la clasificación de cada parámetro.

El perfil promedio se ve en la Figura 151, donde se observa que la tendencia previamente descrita se mantiene a nivel global, evidenciando un comportamiento homogéneo en la serie analizada.

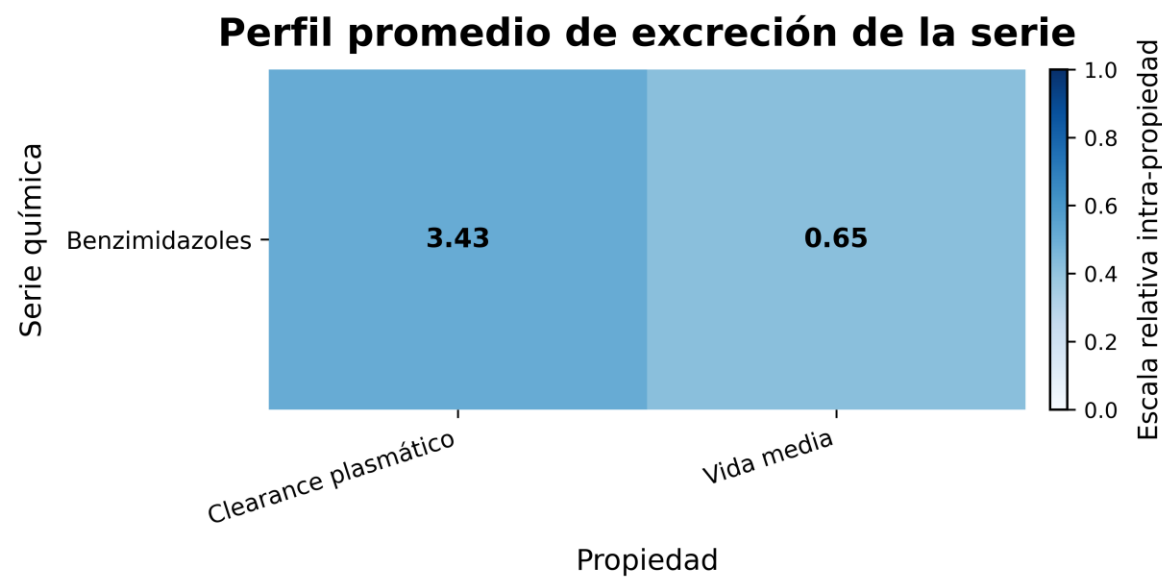


Figura 151 Perfil promedio de excreción de la serie de benzimidazoles. Los valores numéricos corresponden al promedio de cada propiedad, mientras que la intensidad del color representa una escala relativa intra-propiedad (0–1), permitiendo comparar tendencias entre parámetros con diferentes unidades.

5.6.2.6 Toxicidad

La distribución de estas categorías (Figura 152) evidencia que una fracción considerable del conjunto se concentra en el rango de alta probabilidad para varios efectos críticos. En particular, la genotoxicidad, la toxicidad respiratoria, la neurotoxicidad, la nefrotoxicidad, la hepatotoxicidad humana y el daño hepático inducido por fármacos (DILI) presentan una dominancia clara de valores elevados, con la totalidad o la mayoría de los compuestos clasificados en la categoría de alto riesgo.

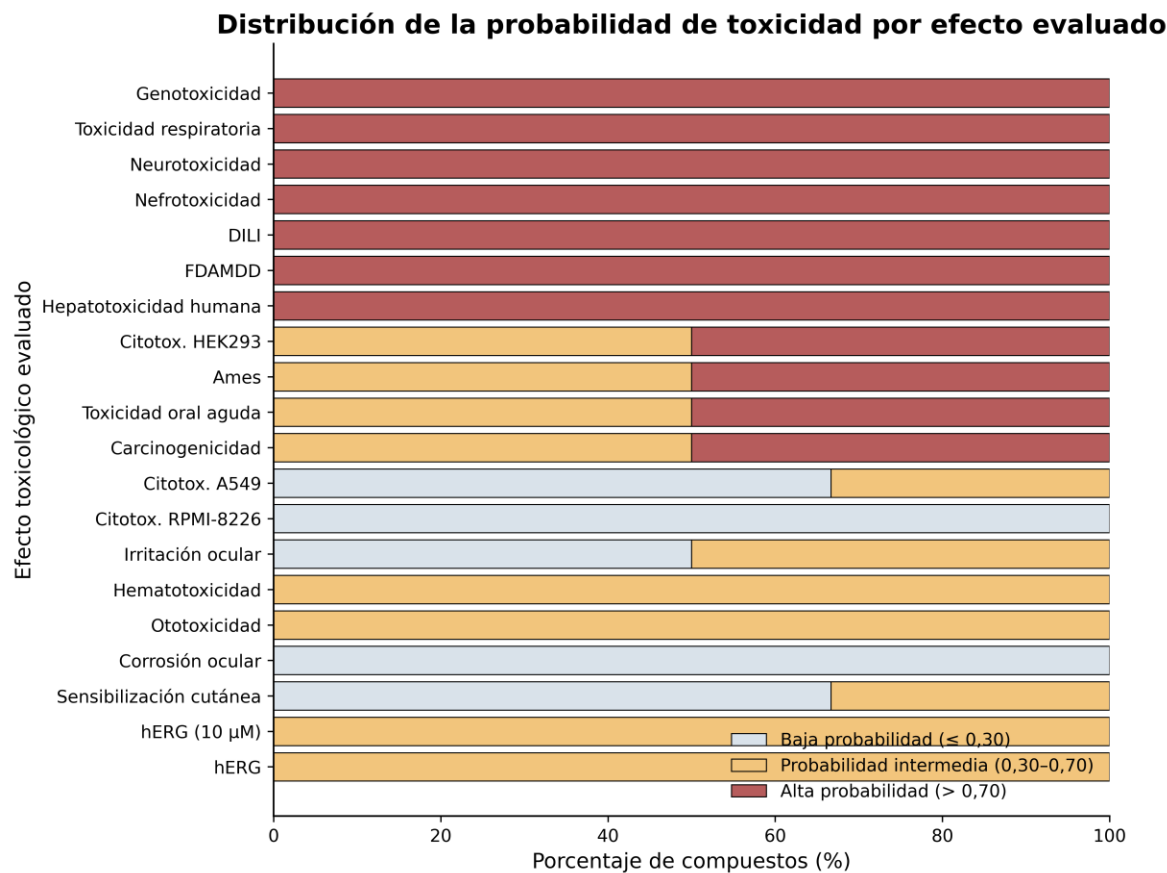


Figura 152 Clasificación toxicológica por endpoint para la serie de benzimidazoles. Distribución porcentual de compuestos en categorías de baja, intermedia y alta probabilidad de toxicidad, según los valores predichos por ADMETlab 3.0. La clasificación se basa en umbrales de probabilidad definidos como $\leq 0,30$ (baja), $0,30\text{--}0,70$ (intermedia) y $> 0,70$ (alta).

De manera consistente, endpoints como FDAMDD, la carcinogenicidad, la toxicidad oral aguda y el ensayo de Ames muestran una distribución sesgada hacia probabilidades intermedias a altas, indicando la

persistencia de señales toxicológicas relevantes en el conjunto. Asimismo, la inhibición del canal hERG (tanto en condiciones estándar como a 10 μ M) presenta predominantemente valores intermedios, lo que sugiere un riesgo moderado de cardiotoxicidad.

En contraste, algunos efectos como la corrosión ocular, la citotoxicidad en líneas celulares específicas (RPMI-8226) y, en menor medida, la irritación ocular y la sensibilización cutánea, muestran una mayor proporción de compuestos en la categoría de baja probabilidad, indicando una contribución menos significativa de estos mecanismos al perfil toxicológico global.

El análisis del perfil promedio (Figura 153) confirma estas tendencias, evidenciando valores elevados de probabilidad en múltiples endpoints, destacando especialmente DILI (0,96), FDAMDD (0,92), neurotoxicidad (0,97), genotoxicidad (0,91) y nefrotoxicidad (0,84). En contraste, endpoints como corrosión ocular (0,00), sensibilización cutánea (0,28) y citotoxicidad en líneas específicas presentan valores considerablemente menores.

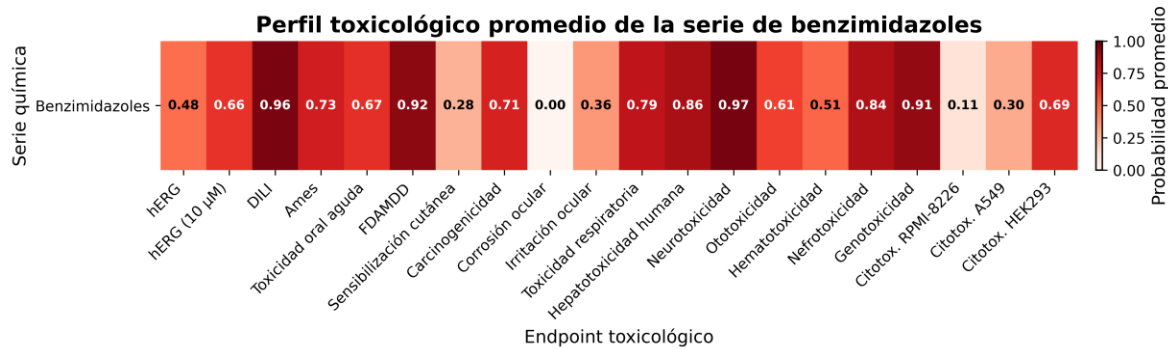


Figura 153 Perfil toxicológico promedio de la serie de benzimidazoles. Mapa de calor que representa los valores promedio de probabilidad para los distintos endpoints toxicológicos. La escala de color refleja la magnitud de la probabilidad, mientras que los valores numéricos corresponden a los promedios calculados para cada efecto.

La distribución global de la carga toxicológica (Figura 154), definida como el número de efectos con probabilidad superior a 0,5 por molécula, muestra que los compuestos se agrupan principalmente entre 13 y 15 efectos positivos, con una mediana cercana a 14. Esta concentración indica una elevada acumulación de señales adversas, reflejando un perfil toxicológico multifactorial caracterizado por la coexistencia simultánea de múltiples riesgos.

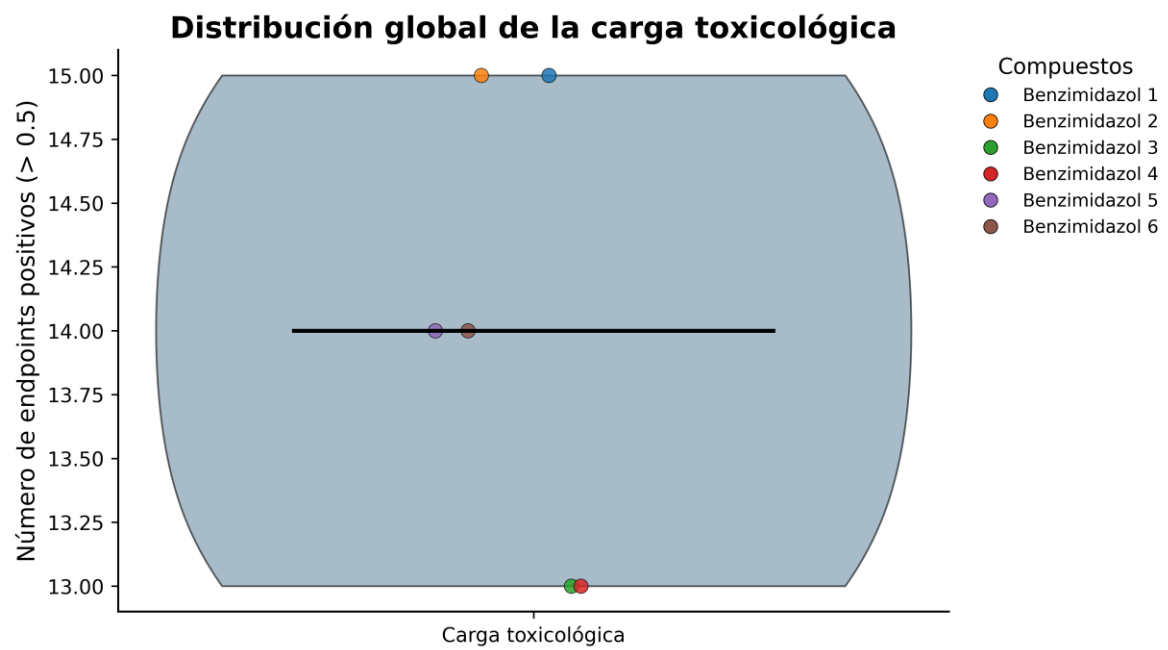


Figura 154 Distribución de la carga toxicológica global de la serie de benzimidazoles, definida como el número de endpoints con probabilidad superior a 0,5 por molécula. El diagrama de violín representa la densidad de la distribución, mientras que los puntos individuales corresponden a cada compuesto y la línea horizontal indica la mediana.

5.6.2.7 Solubilidad

La solubilidad acuosa predicha de los seis benzimidazoles(Figura 155 y Figura 156), expresada como logS, se distribuyó en un rango entre -4.694 y -4.181. El compuesto benzimidazol 5 presentó el valor menos

negativo (-4.181), seguido de benzimidazol 3 (-4.188) y benzimidazol 4 (-4.389). En valores intermedios se ubicó benzimidazol 6 (-4.420), mientras que los valores más negativos correspondieron a benzimidazol 1 (-4.566) y benzimidazol 2 (-4.694). En conjunto, todos los compuestos presentaron valores de logS inferiores a -4.

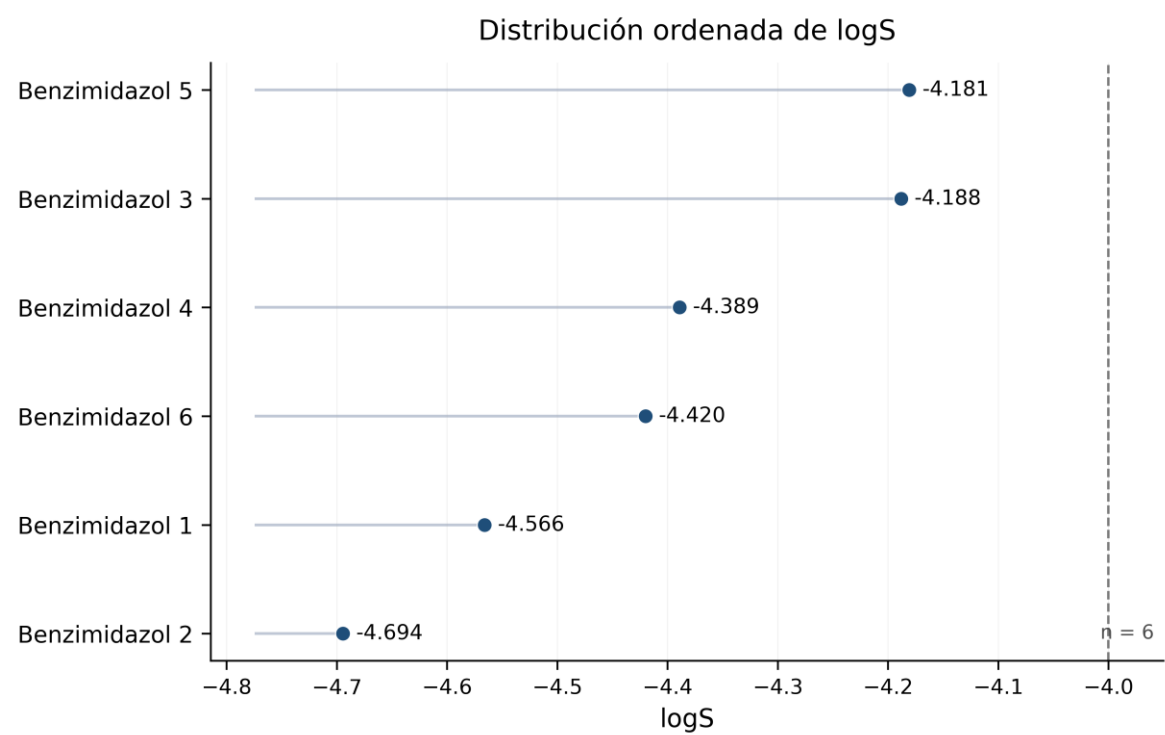


Figura 155 Distribución ordenada de los valores de logS predichos para los benzimidazoles evaluados. Los compuestos se muestran ordenados de acuerdo con su valor de logS, desde el mayor al menor. La línea vertical punteada indica el valor de referencia logS = -4. Los valores obtenidos se distribuyeron entre -4.694 y -4.181, y en todos los casos se ubicaron por debajo de dicho umbral.

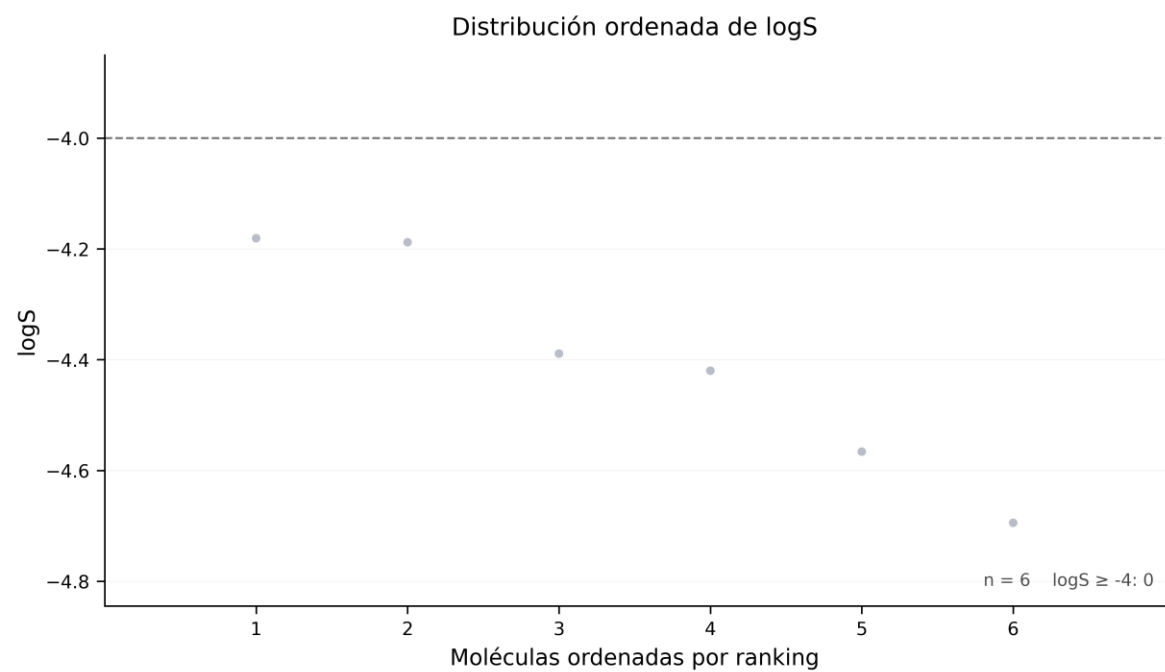


Figura 156 Distribución ordenada de logS para los benzimidazoles evaluados. Los compuestos fueron ordenados de acuerdo con su valor de logS. La línea punteada indica el valor de referencia logS = -4. Todos los compuestos presentaron valores inferiores a este umbral.

5.6.3 Moléculas top1 Virtual screening

5.6.3.1 Drogabilidad

Se evaluó el carácter drug-like de las moléculas representativas de seis familias químicas mediante la regla de Lipinski y el criterio de Veber (Tabla 12). Los parámetros analizados incluyeron peso molecular, lipofilidad (logP), número de donadores y aceptores de enlace de hidrógeno, superficie polar topológica (TPSA) y número de enlaces rotables.

A diferencia de lo observado en series más homogéneas, en este caso se evidenció una marcada variabilidad fisicoquímica entre las familias. Ninguna de las moléculas cumplió completamente la regla de Lipinski, ya que todas presentaron al menos una violación. Las familias aminopirimidina y benzotiofeno mostraron una única violación, mientras que benzimidazol, benzodiazepina y benzamida presentaron dos, y la familia pirimidina aminobenzimidazol registró tres violaciones, de acuerdo con lo consignado en la tabla. En general, las desviaciones se asociaron principalmente a valores elevados de peso molecular y/o lipofilidad, como se observó en benzimidazol, benzodiazepina y pirimidina aminobenzimidazol, mientras que aminopirimidina y benzamida destacaron por presentar propiedades más polares.

En cuanto al criterio de Veber, cuatro de las seis familias cumplieron con los parámetros establecidos: benzimidazol, benzodiazepina, benzotiofeno y pirimidina aminobenzimidazol. Por el contrario, aminopirimidina y benzamida no cumplieron este criterio, principalmente debido a sus altos valores de TPSA (154.58 y 167.41 Å², respectivamente), superiores al límite empírico de 140 Å². Estos resultados indican que la serie presenta una drogabilidad variable entre familias, con perfiles más favorables en aquellas que, pese a presentar violaciones de Lipinski, conservaron valores compatibles con el criterio de Veber.

Tabla 12 Descripciones de las moléculas top 1 del virtual screening para reglas de Lipinsky y Veber(Sí=cumple con las reglas, No= No cumple con las reglas).

Familia Molecular	MM	logP	HBD	HBA	Violaciones de Lipinsky	TPSA	nRot	Veber
Aminopirimidina	194.1	0.1184	8	8	1	154.58	2	No
Benzimidazol	604.2	10.667	0	2	2	17.82	5	Sí
Benzodiazepina	617.26	7.8306	1	5	2	45.98	3	Sí
Benzamida	594.23	1.2709	4	12	2	167.41	10	No
Benzotiofeno	603.12	3.7980	3	6	1	92.5	8	Sí
Pirimidina Aminobenzimidazol	617.19	8.2154	0	7	3	74.29	3	Sí

5.6.3.2Absorción

La clasificación empírica de absorción intestinal humana predicha (HIA) indica que las siete moléculas del conjunto de virtual screening se ubican en la categoría favorable. Los valores predichos de HIA se encuentran entre 0.0000 y 0.00010, todos muy por debajo del umbral de 0.30 asociado a absorción intestinal favorable (Figura 157).

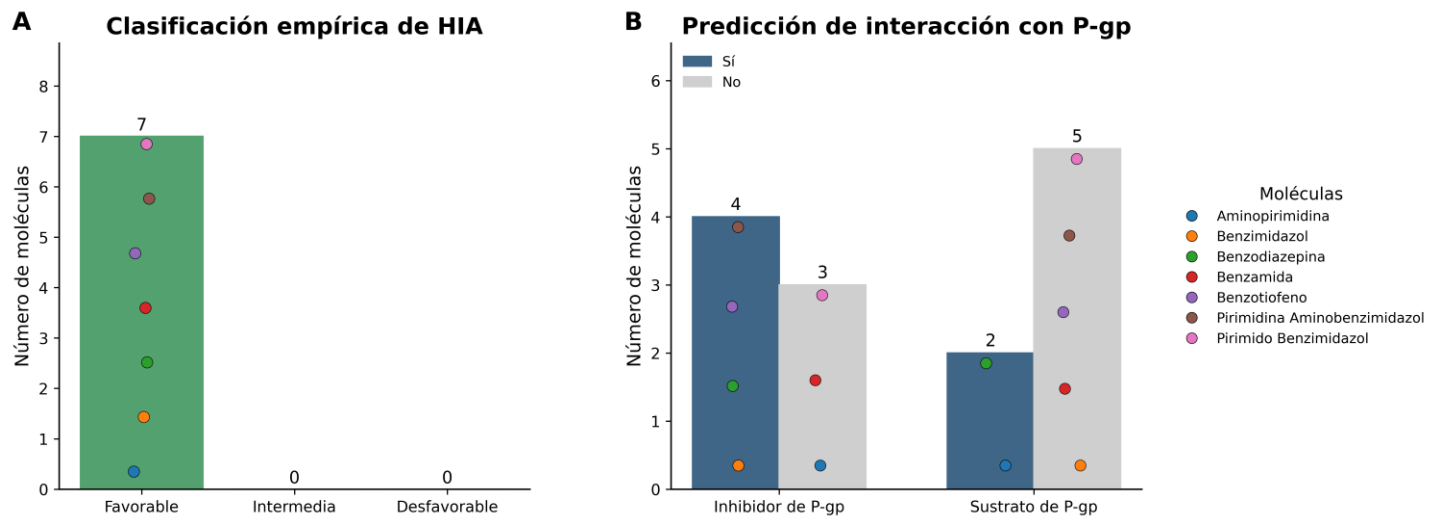


Figura 157 Propiedades de absorción intestinal e interacción con P-glicoproteína predichas para el conjunto de moléculas seleccionadas mediante virtual screening. (A) Clasificación de las moléculas según la probabilidad de absorción intestinal humana (HIA). Valores entre 0–0.30 corresponden a absorción favorable, entre 0.30–0.70 a absorción intermedia y mayores a 0.70 a absorción desfavorable. (B) Distribución de compuestos predichos como inhibidores y sustratos de P-glicoproteína (P-gp), empleando un umbral de probabilidad de 0.5 para la clasificación binaria. Los puntos coloreados corresponden a las moléculas individuales identificadas en la leyenda.

En relación con la interacción con P-glicoproteína (P-gp), el conjunto presenta un comportamiento más heterogéneo. Cuatro moléculas fueron clasificadas como potenciales inhibidores de P-gp al presentar probabilidades superiores a 0.5, mientras que tres se ubicaron bajo ese umbral. En contraste, solo dos compuestos fueron predichos como sustratos de P-gp, mientras que los cinco restantes presentaron probabilidades inferiores a 0. Los resultados muestran que, aunque la absorción intestinal predicha por HIA es uniformemente favorable, la interacción con P-gp no es homogénea entre las moléculas.

La distribución de la permeabilidad predicha en el modelo Caco-2 muestra valores comprendidos entre -5.90 y -4.82 log cm/s, con una mediana de -5.27 log cm/s (Figura 158). Considerando el umbral empírico de -5.15 log cm/s, tres compuestos se sitúan en la región asociada a alta permeabilidad intestinal, mientras que cuatro se ubican por debajo de ese valor, en la zona de menor permeabilidad relativa. Por tanto, a diferencia de la clasificación HIA, la permeabilidad Caco-2 evidencia una dispersión más marcada dentro del conjunto evaluado.

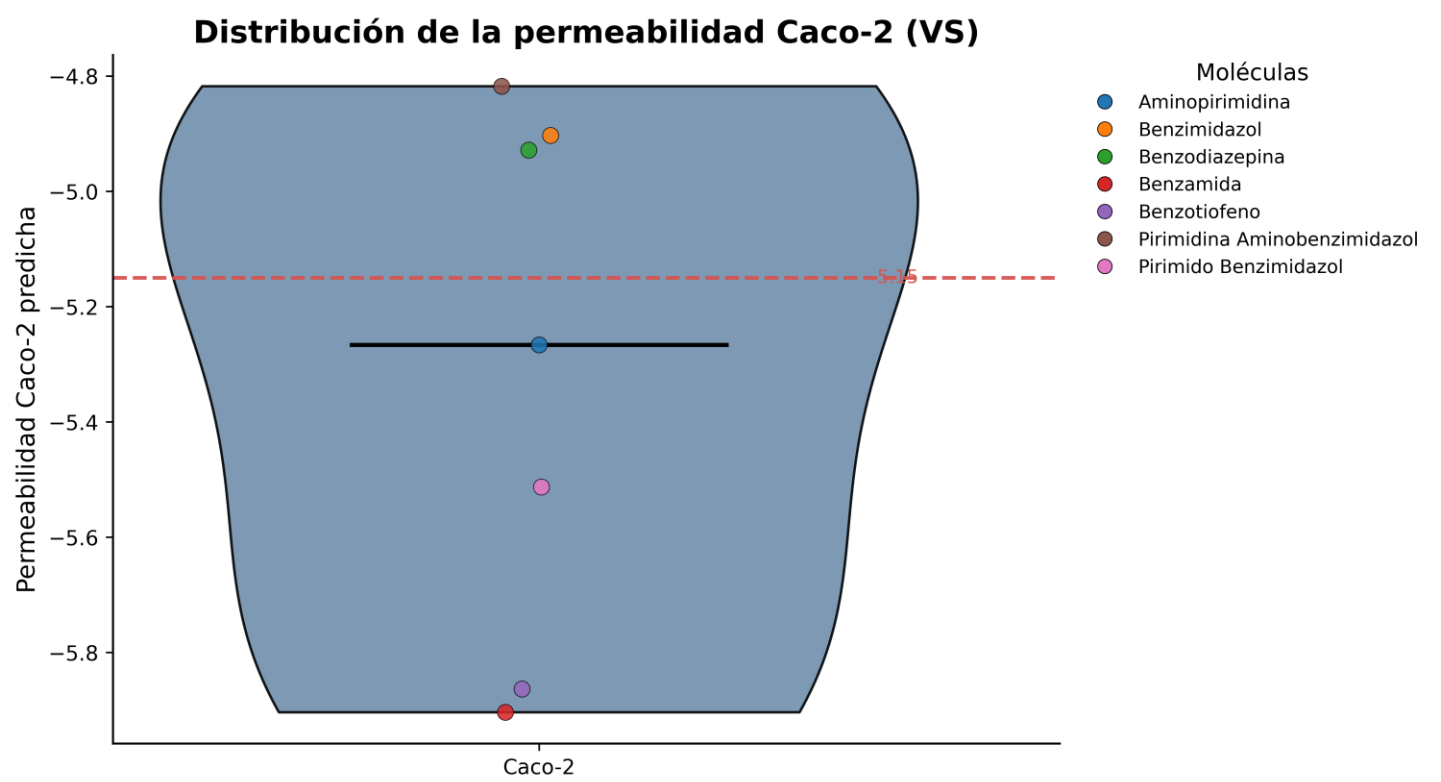


Figura 158 Distribución de la permeabilidad Caco-2 predicha para el conjunto de moléculas seleccionadas mediante virtual screening(VS). El diagrama de violín representa la densidad de la distribución y los puntos individuales corresponden a cada molécula evaluada. La línea negra indica la mediana. Los compuestos con valores superiores a -5.15 log cm/s se consideran de alta permeabilidad intestinal, mientras que la línea roja discontinua señala el umbral empírico de referencia.

5.6.3.3Distribución

La caracterización de las propiedades de distribución del conjunto de virtual screening se resume en la Figura 159. La probabilidad de penetración a la barrera hematoencefálica (BBB) indicó que cinco moléculas fueron clasificadas como BBB+ y dos como BBB-, utilizando un umbral de 0.5, con valores comprendidos entre 1.09×10^{-8} y 0.998. La PPB presentó una distribución heterogénea, con valores entre -0.44 y 98.53% y una mediana de 95.06% , reflejando que la mayoría de los compuestos exhibe una elevada unión a proteínas plasmáticas. En correspondencia, la Fu osciló entre 0.23 y 99.02%, con una mediana de 5.03% , mostrando que la serie se concentra principalmente en valores bajos, aunque con un compuesto claramente desplazado hacia valores altos. Finalmente, el VDss varió entre 0.72 y 3.91 L/kg, con una mediana de 1.89 L/kg , ubicándose todos los compuestos dentro del rango de referencia indicado en la figura.

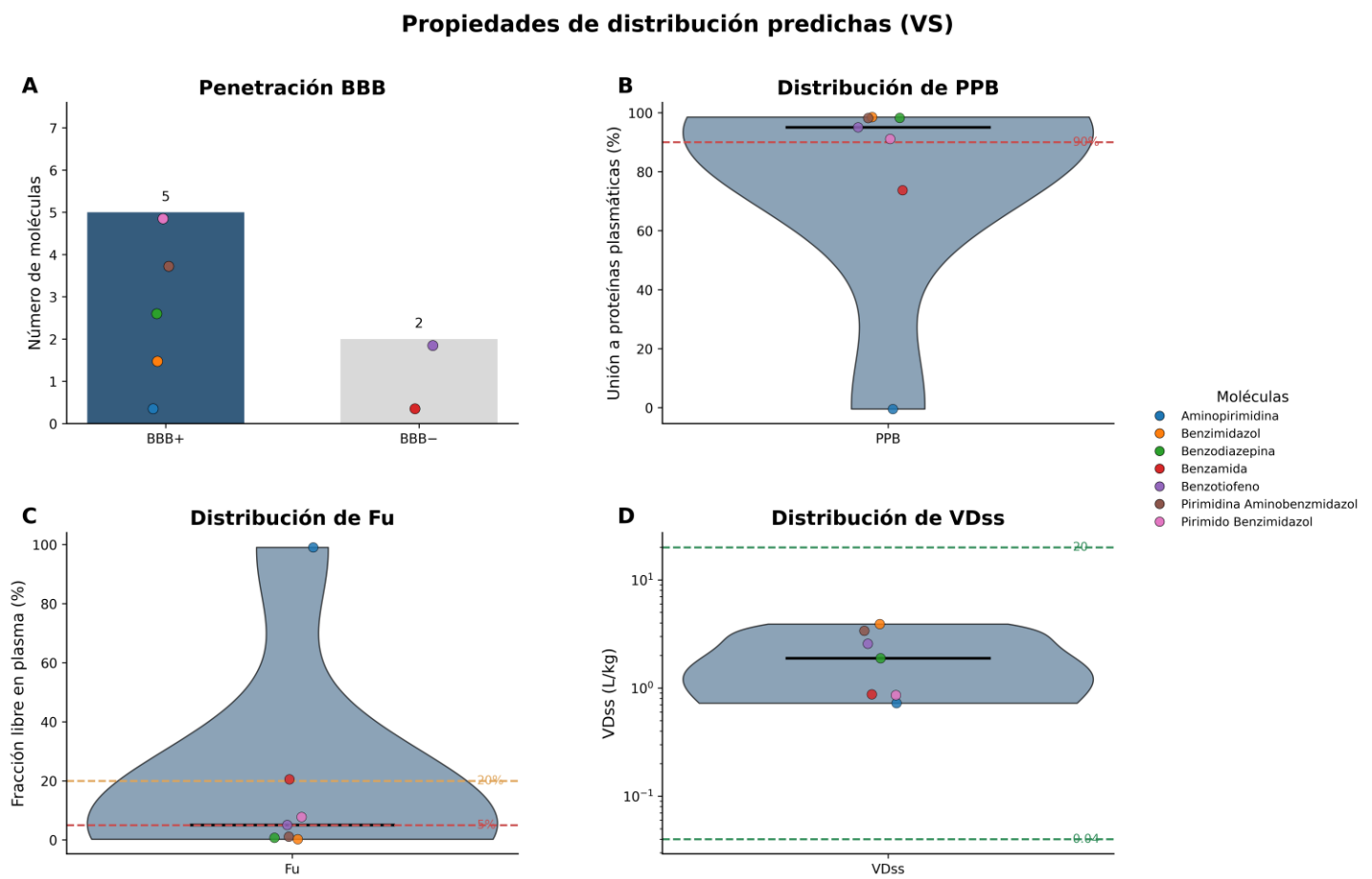


Figura 159 Propiedades de distribución predichas para el conjunto de moléculas seleccionadas mediante virtual screening(VS). (A) Clasificación de penetración a la barrera hematoencefálica (BBB) basada en un umbral de probabilidad de 0.5, mostrando el número de moléculas clasificadas como BBB+ y BBB-. Los puntos coloreados indican las moléculas individuales correspondientes a cada categoría. (B) Distribución de la unión a proteínas plasmáticas (PPB, %). La línea negra representa la mediana y la línea discontinua roja indica el umbral de referencia de 90%. (C) Distribución de la fracción libre en plasma (Fu, %). La línea negra representa la mediana, mientras que las líneas discontinuas indican los valores de referencia de 5% y 20%. (D) Distribución del volumen de distribución en estado estacionario (VDss, L/kg) en escala logarítmica. La línea negra indica la mediana y las líneas discontinuas verdes delimitan el rango de referencia (0.04–20 L/kg). Los puntos representan los valores individuales de cada molécula identificada en la leyenda.

5.6.3.4Metabolismo

La caracterización del perfil metabólico del conjunto de moléculas seleccionadas mediante virtual screening se presenta en la Figura 160 y Figura 161. En términos generales, las probabilidades predichas de inhibición y sustrato frente a isoformas de CYP450 muestran un comportamiento heterogéneo, con distribuciones amplias y, en varios casos, polarizadas hacia valores muy bajos o altos.

En el análisis de inhibición de CYP450 (Figura 160), la isoforma que mostró la mayor tendencia a inhibición fue CYP2C8, con probabilidades elevadas en casi todo el conjunto y una mediana cercana a 1.00. También se observó una propensión importante hacia la inhibición de CYP2C9, con una mediana de 0.77. En contraste, CYP2B6 y CYP2D6 presentaron medianas más bajas, de 0.27 y 0.12,

respectivamente, lo que sugiere un comportamiento menos uniforme frente a estas isoformas. Para CYP1A2, CYP2C19 y CYP3A4 se observó una dispersión marcada, con compuestos distribuidos tanto en la zona de baja como de alta probabilidad de inhibición.

El análisis de probabilidad de sustrato de CYP450 (Figura 161) también reveló un perfil diferencial entre isoformas. Las mayores probabilidades de comportamiento como sustrato se observaron para CYP2C9, CYP2D6 y CYP1A2, con medianas de 0.99, 0.99 y 0.95, respectivamente. CYP2C19 presentó una mediana intermedia de 0.55, lo que sugiere una tendencia moderada a comportarse como sustrato. En cambio, CYP2B6 y CYP3A4 mostraron medianas bajas, de 0.0008 y 0.05, respectivamente, indicando una menor propensión global a ser reconocidos como sustratos de estas isoformas.

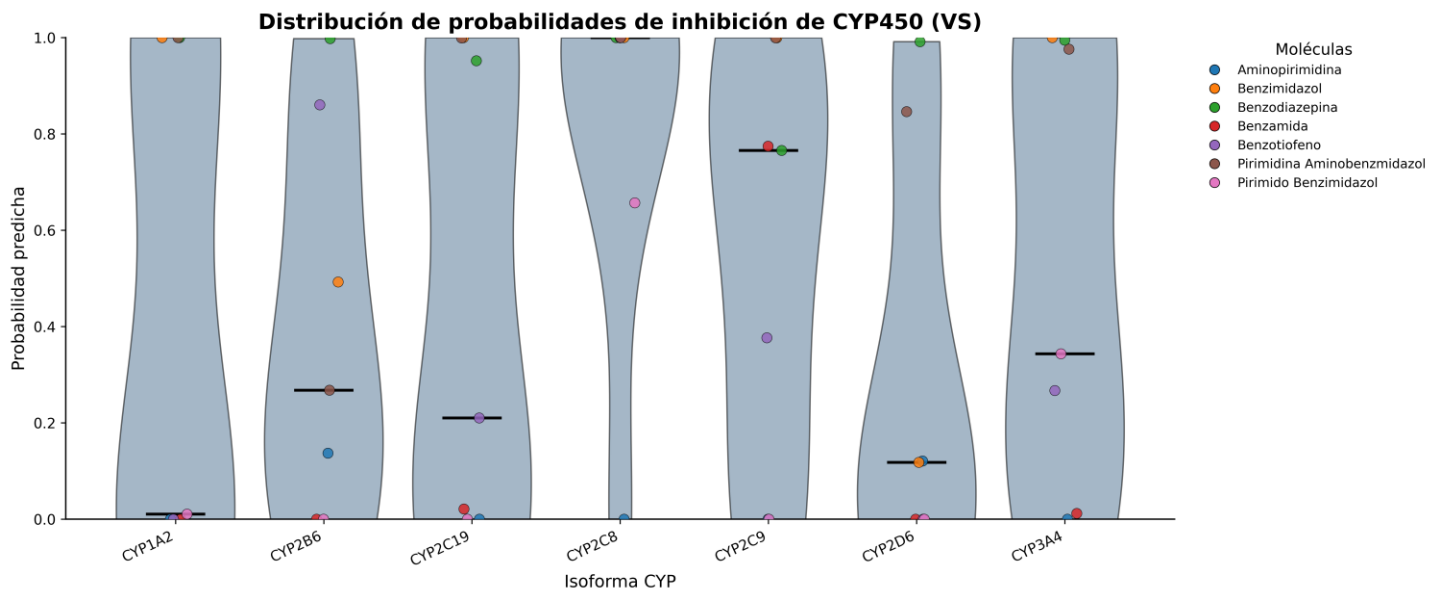


Figura 160 Distribución predicha de la probabilidad frente a isoformas de CYP450. En el panel de inhibición, CYP2C8 presentó la mayor tendencia inhibitoria, con una mediana cercana a 1.00, seguida de CYP2C9 con una mediana de 0.77. En cambio, CYP2B6 y CYP2D6 mostraron medianas menores, mientras que CYP1A2, CYP2C19 y CYP3A4 exhibieron distribuciones más dispersas. En el panel de sustrato, las mayores probabilidades se observaron para CYP2C9, CYP2D6 y CYP1A2, con medianas de 0.99, 0.99 y 0.95, respectivamente, mientras que CYP2B6 y CYP3A4 presentaron una tendencia globalmente baja. En conjunto, el conjunto VS mostró un comportamiento metabólico heterogéneo, con una interacción particularmente marcada con isoformas específicas del sistema CYP450.

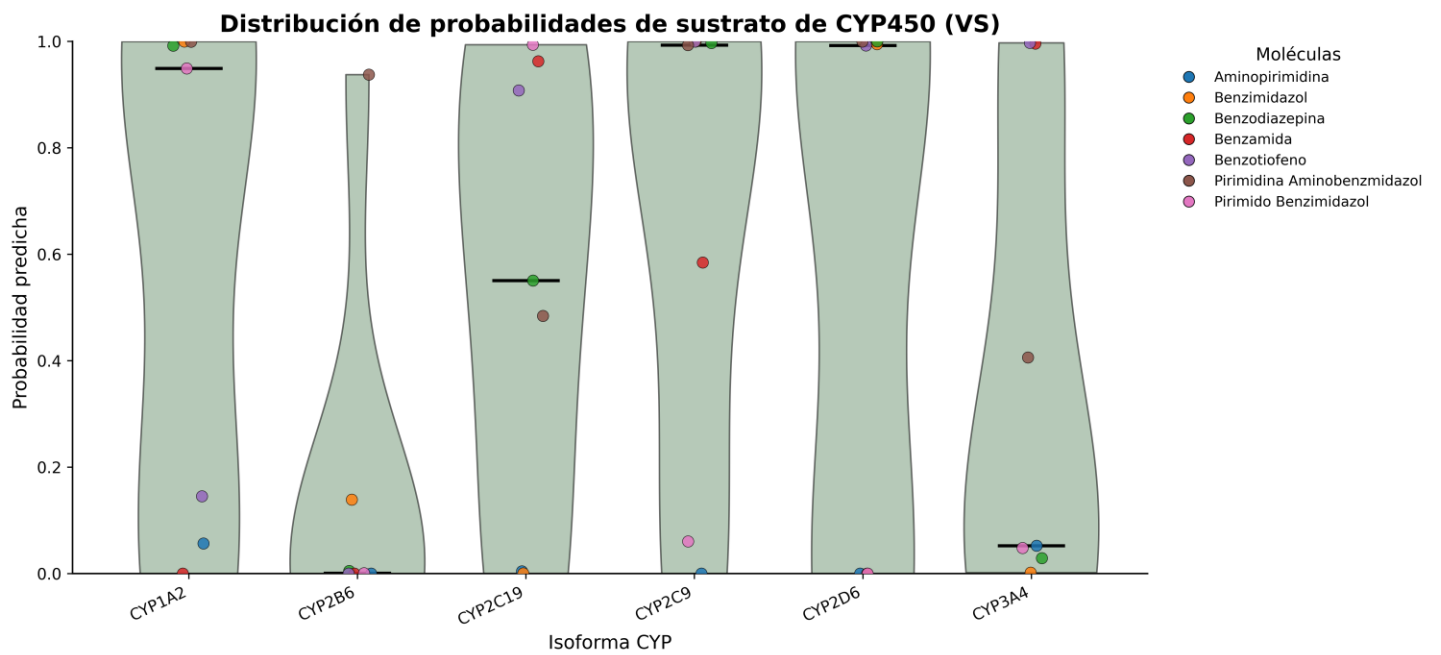


Figura 161 Distribución de probabilidades de ser sustrato de isoformas de CYP450 para el conjunto de moléculas seleccionadas mediante virtual screening. El diagrama de violín representa la densidad de la distribución para cada isoforma y los puntos coloreados corresponden a las moléculas individuales identificadas en la leyenda. La línea negra indica la mediana. Se evaluaron las isoformas CYP1A2, CYP2B6, CYP2C19, CYP2C8, CYP2C9, CYP2D6 y CYP3A4.

5.6.3.5Excreción

La caracterización de las propiedades de excreción del conjunto de moléculas seleccionadas mediante virtual screening se presenta en la Figura 162. El clearance plasmático predicho mostró valores comprendidos entre 1.40 y 3.58 mL/min/kg, con una mediana de 2.97 mL/min/kg. Todos los compuestos se ubicaron por debajo del valor de referencia de 5 mL/min/kg, lo que indica un comportamiento homogéneo dentro del rango de bajo clearance plasmático.

Por su parte, la vida media predicha presentó valores entre 0.44 y 2.07 h, con una mediana de 0.57 h. La mayor parte de las moléculas se concentró por debajo de 1 h, mientras que solo un compuesto mostró un valor cercano a 2 h. En todos los casos, los valores permanecieron por debajo del umbral de 4 h, lo que sugiere una tendencia general hacia una vida media corta en el conjunto evaluado.

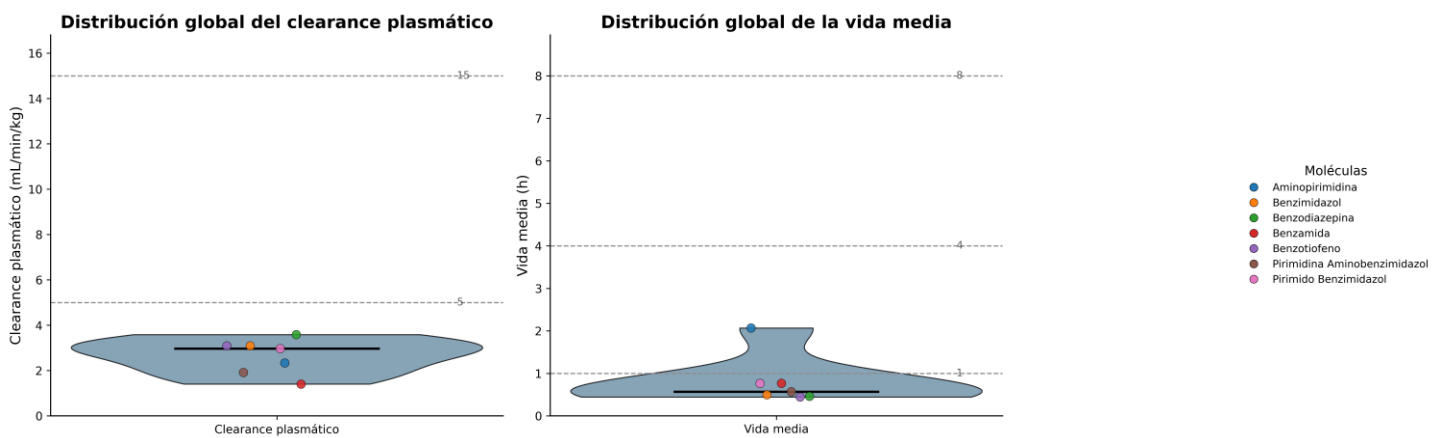


Figura 162 Propiedades de excreción predichas para las moléculas seleccionadas mediante virtual screening (VS). El clearance plasmático osciló entre 1.40 y 3.58 mL/min/kg, con una mediana de 2.97 mL/min/kg, ubicándose todas las moléculas por debajo del valor de referencia de 5 mL/min/kg. De manera consistente, la vida media presentó valores entre 0.44 y 2.07 h, con una mediana de 0.57 h concentrándose la mayoría de los compuestos por debajo de 1 h y sin alcanzar el umbral de 4 h. En conjunto, el perfil de excreción predicho sugiere una serie con comportamiento relativamente homogéneo, dominado por bajo clearance plasmático y vida media breve.

5.6.3.6Toxicidad

La caracterización del perfil toxicológico predicho del conjunto de moléculas seleccionadas mediante virtual screening se presenta en la . La distribución de probabilidades por efecto toxicológico muestra un patrón marcadamente heterogéneo, con endpoints dominados por probabilidades altas y otros concentrados en probabilidades bajas (Figura 163)

Los efectos con mayor señal de riesgo fueron genotoxicidad, DILI y hepatotoxicidad humana, ya que en los tres casos el 100% de los compuestos se ubicó en la categoría de alta probabilidad. También se observó una predominancia de probabilidades altas para Ames y ototoxicidad, con 71.4% de los compuestos en la categoría alta, así como para neurotoxicidad, hematotoxicidad, nefrotoxicidad y citotoxicidad en HEK293, donde 57.1% de las moléculas se situó en la franja de alta probabilidad. En contraste, corrosión ocular e irritación ocular mostraron un perfil claramente favorable, con el 100% de los compuestos en la categoría de baja probabilidad. De forma similar, sensibilización cutánea y citotoxicidad en RPMI-8226 concentraron 71.4% de los compuestos en la categoría baja .

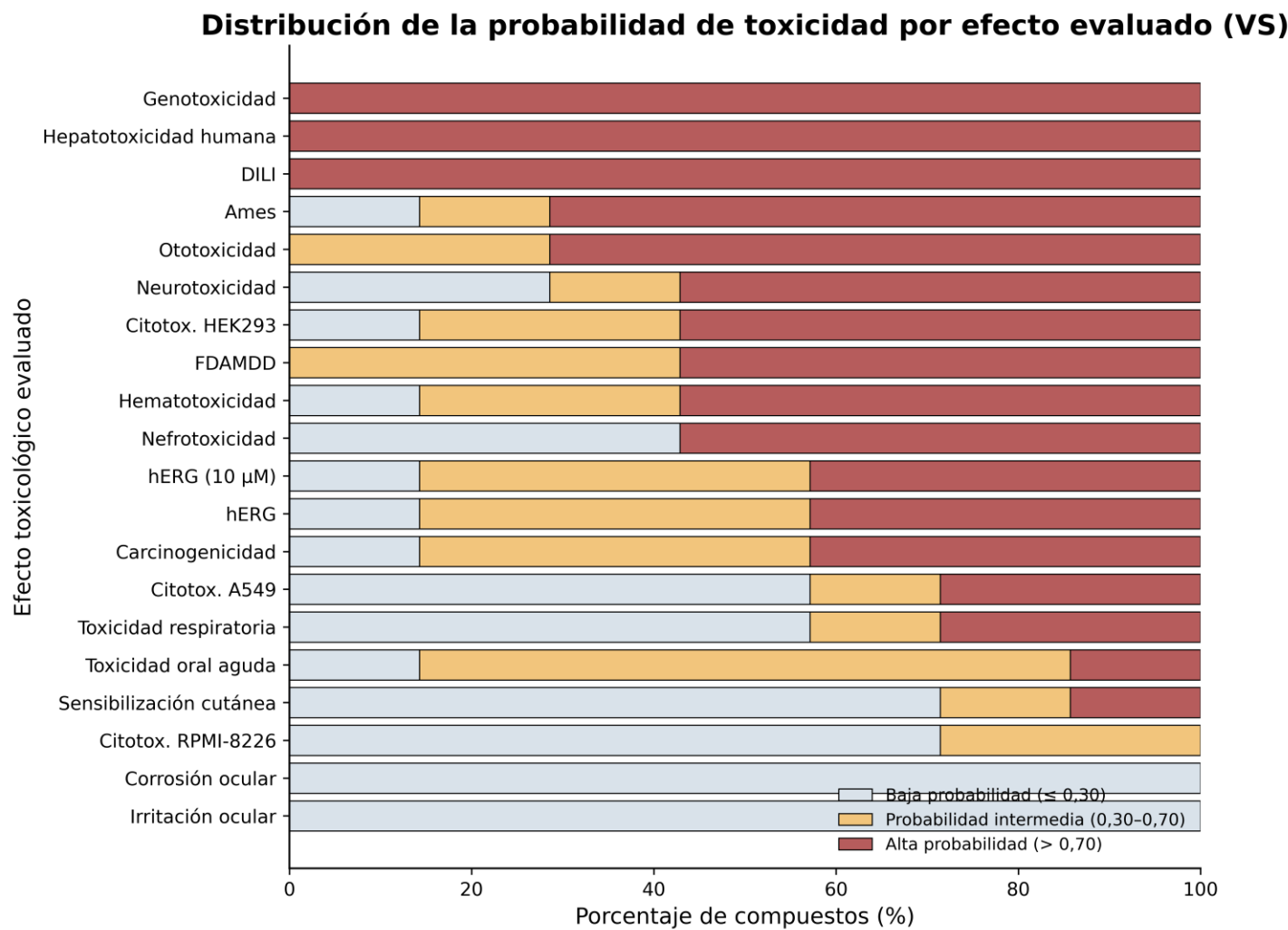


Figura 163 Distribución de la probabilidad de toxicidad por efecto evaluado para el conjunto de moléculas seleccionadas mediante virtual screening. Las barras apiladas muestran el porcentaje de compuestos clasificados en baja probabilidad (≤ 0.30), probabilidad intermedia (0.30–0.70) y alta probabilidad (> 0.70) para cada endpoint toxicológico.

El perfil toxicológico promedio (Figura 164) refuerza estas tendencias. Las mayores probabilidades medias se observaron para genotoxicidad (1.00), DILI (0.96) y hepatotoxicidad humana (0.93). También destacaron valores promedio elevados para FDAMDD (0.80), ototoxicidad (0.78), Ames (0.71), citotoxicidad en HEK293 (0.70) y neurotoxicidad (0.70). En un rango intermedio se ubicaron hERG (0.68), nefrotoxicidad (0.63), hERG a 10 µM (0.61), carcinogenicidad (0.61) y hematotoxicidad (0.60). Por el contrario, los valores promedio más bajos correspondieron a corrosión ocular (0.00), irritación ocular (0.03), citotoxicidad en RPMI-8226 (0.18) y sensibilización cutánea (0.24).

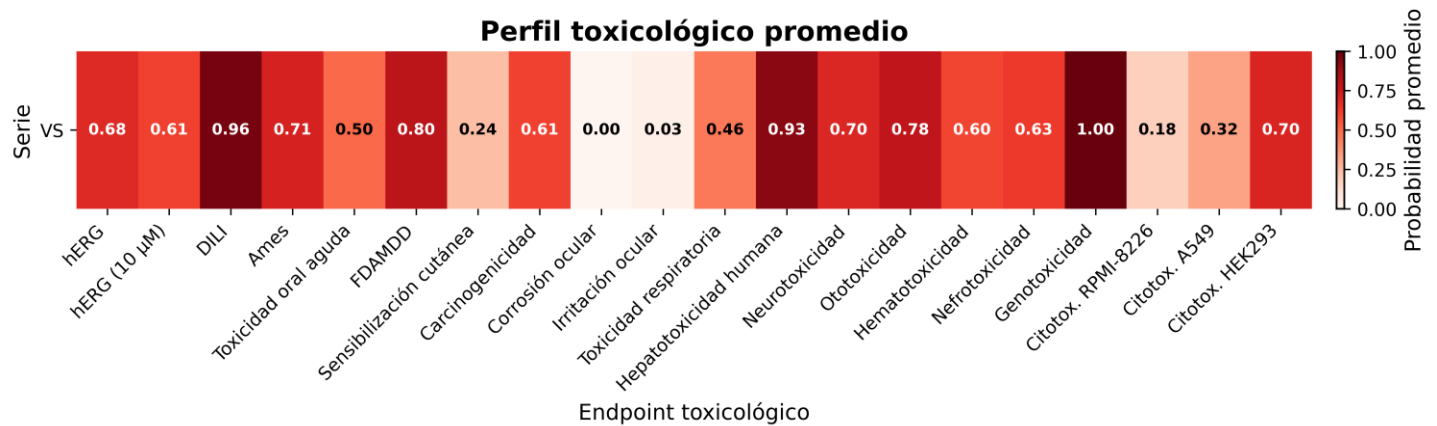


Figura 164 Perfil toxicológico promedio del conjunto de moléculas seleccionadas mediante virtual screening. El mapa de calor representa la probabilidad media predicha para cada endpoint toxicológico. La intensidad del color aumenta con la magnitud de la probabilidad promedio y los valores numéricos se indican dentro de cada celda.

La distribución global de la carga toxicológica, definida como el número de endpoints con probabilidad superior a 0.5, mostró valores entre 4 y 16, con una mediana de 11 (Figura 165). La mayor parte de las moléculas se agrupó en torno a valores intermedios-altos, lo que indica que el conjunto presenta una acumulación relevante de señales toxicológicas positivas a través de múltiples endpoints, aunque con una variabilidad apreciable entre compuestos.

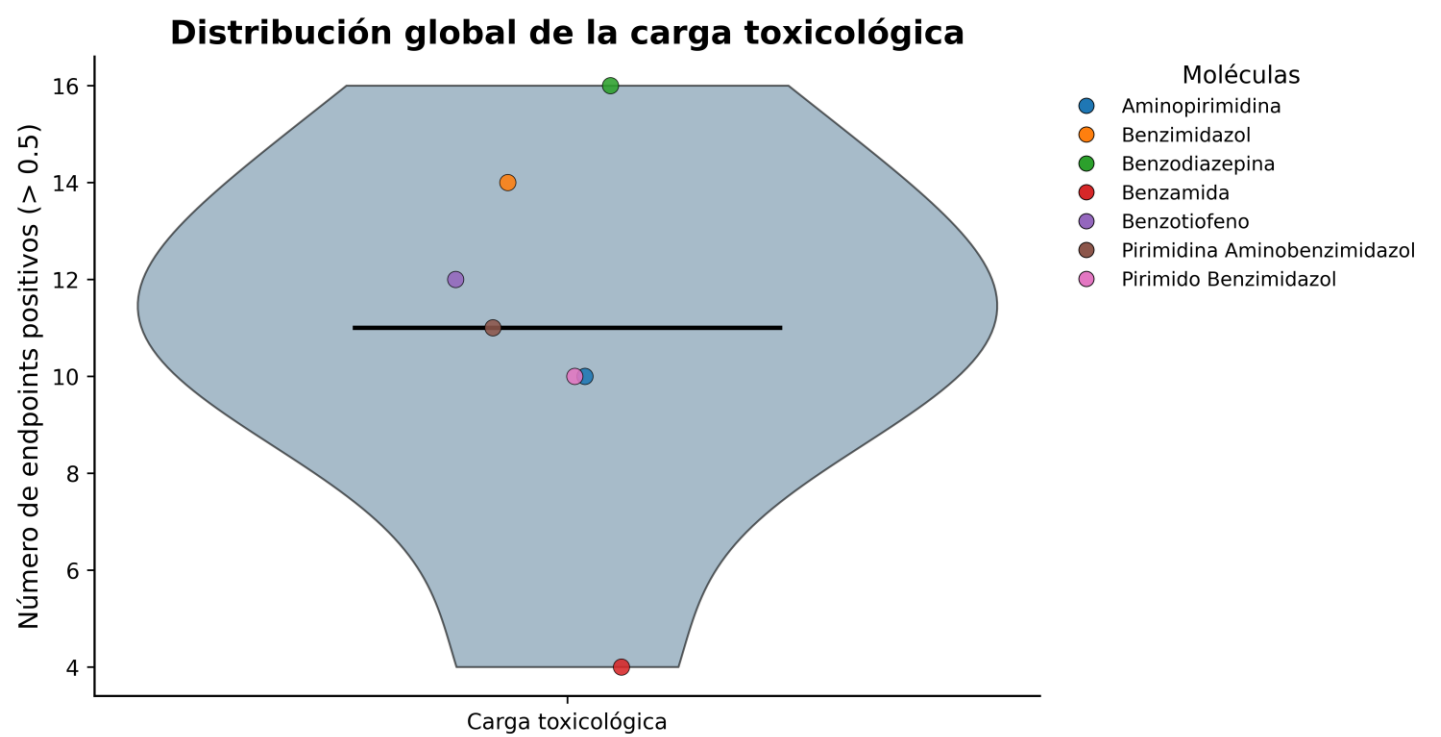


Figura 165 Distribución global de la carga toxicológica del conjunto de moléculas seleccionadas mediante virtual screening. El diagrama de violín representa la densidad de la distribución del número de endpoints con probabilidad superior a 0.5 para cada molécula. Los puntos coloreados corresponden a los compuestos individuales identificados en la leyenda y la línea negra indica la mediana.

5.6.3.7Solubilidad

La distribución ordenada de logS para las moléculas seleccionadas mediante virtual screening se presenta en laFigura 166. Los valores predichos abarcan un rango amplio, desde -12.44 hasta -1.88, lo que evidencia una variabilidad importante en la solubilidad acuosa del conjunto. La molécula con mayor solubilidad predicha fue Aminopirimidina (logS = -1.88), seguida de Benzamida (-3.20) y Pirimido Benzimidazol (-3.23). En contraste, las menores solubilidades se observaron para Benzimidazol (-12.44) y Pirimidina aminobenzimidazol (-10.59).

Considerando el umbral empírico de logS = -4.0, tres moléculas se ubicaron por encima de este valor, lo que indica una solubilidad relativamente más favorable, mientras que las cuatro restantes quedaron por debajo del umbral, sugiriendo limitaciones de solubilidad más marcadas. En este contexto, Benzotiofeno presentó un valor intermedio cercano al punto de corte (-4.44), mientras que Benzodiazepina (-7.48), Pirimidina aminobenzimidazol (-10.59) y Benzimidazol (-12.44) se desplazaron claramente hacia la región de menor solubilidad.

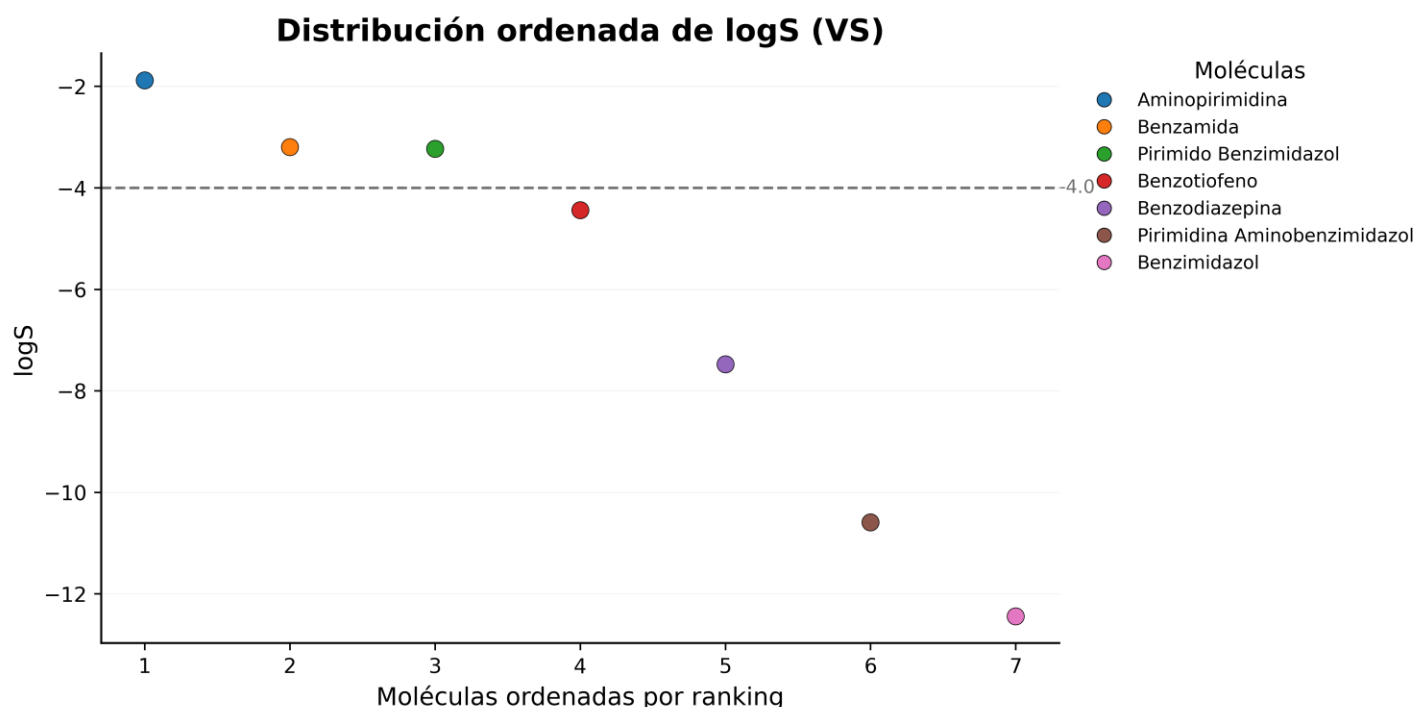


Figura 166 Distribución ordenada de logS para el conjunto de moléculas seleccionadas mediante virtual screening. Los puntos coloreados representan los valores individuales de cada compuesto, identificados en la leyenda. La línea discontinua gris indica el umbral empírico de referencia en $\log S = -4.0$, utilizado para distinguir compuestos con solubilidad relativamente más favorable de aquellos con menor solubilidad predicha.

5.7 Retrosíntesis

5.7.1 Chalconas y heterociclos derivados

5.7.1.1 Chalconas

Las chalconas constituyen el núcleo estructural central de la serie de compuestos estudiada, estando presentes tanto como productos finales como intermediarios en la obtención de derivados heterocíclicos. Desde un punto de vista retrosintético, estas estructuras pueden ser desconectadas a nivel del doble enlace α,β -insaturado, generando como fragmentos una cetona aromática y un aldehído aromático (Figura 167)

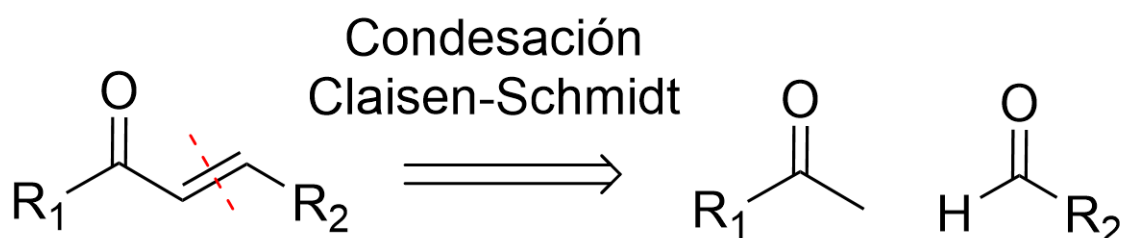


Figura 167 Análisis retrosintético de las chalconas. La desconexión del sistema α,β -insaturado conduce a una cetona aromática ($R_1\text{-CO-}$) y un aldehído aromático ($R_2\text{-CHO}$), los cuales pueden ser combinados mediante condensación de Claisen-Schmidt para generar la estructura final.

La desconexión identificada corresponde a una transformación tipo condensación de Claisen-Schmidt, en la cual una cetona con hidrógenos α reacciona con un aldehído aromático para dar lugar al sistema α,β -insaturado característico de las chalconas.

5.7.1.2Aminopirimidinas

Por su parte, los derivados aminopirimidínicos pueden ser relacionados retrosintéticamente con una chalcona y guanidina como precursores inmediatos. En este caso, la formación del anillo heterocíclico se asocia a un proceso de ciclación/condensación entre el sistema α,β-insaturado de la chalcona y el fragmento nitrogenado aportado por la guanidina, permitiendo establecer una correspondencia directa entre las estructuras finales y sus bloques d,e construcción(Figura 168).

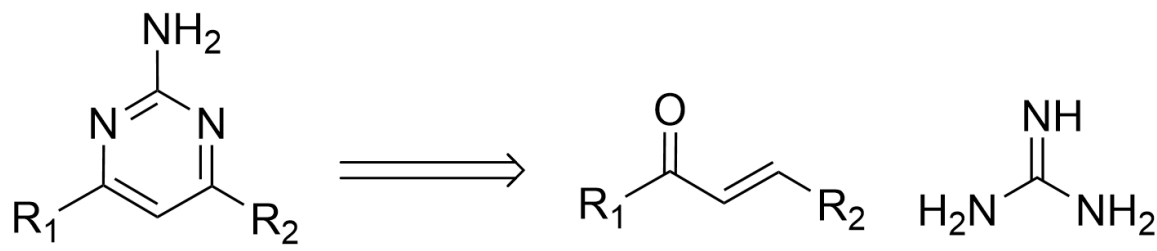


Figura 168 Análisis retrosintético de los derivados aminopirimidínicos. La estructura heterocíclica puede ser relacionada con una chalcona como precursor carbonílico α,β-insaturado y guanidina como fragmento nitrogenado. Esta transformación corresponde a un proceso de ciclación/condensación que conduce a la formación del anillo pirimidínico.

5.7.1.3 Dihidrobenzodiazepinas

Los derivados de dihidrobenzodiazepina pueden ser relacionados retrosintéticamente con una chalcona y o-fenilendiamina como precursores inmediatos (Figura 169). En este contexto, la chalcona aporta el sistema α,β-insaturado y el grupo carbonilo, mientras que la o-fenilendiamina constituye el fragmento bifuncional nitrogenado necesario para la formación del anillo heterocíclico. Este patrón de desconexión permite establecer una correspondencia directa entre las estructuras finales y sus bloques de construcción, evidenciando una estrategia común dentro de la serie analizada.

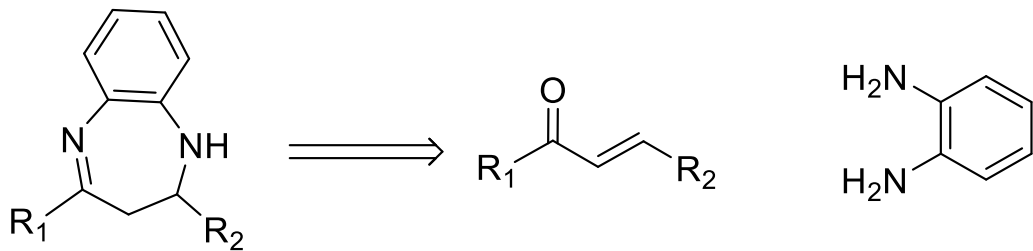


Figura 169 Análisis retrosintético de los derivados de dihidrobenzodiazepina. La estructura heterocíclica puede ser relacionada con una chalcona como precursor α,β-insaturado y o-fenilendiamina como fragmento nitrogenado bifuncional, los cuales permiten la construcción del anillo benzodiazepínico.

5.7.1.4 Dihidropirazol

Los derivados de dihidropirazol pueden ser relacionados retrosintéticamente con una chalcona y una hidrazina como precursores inmediatos (Figura 170). En este contexto, la chalcona aporta el sistema α,β-insaturado necesario para la formación del anillo heterocíclico, mientras que la hidrazina constituye el fragmento nitrogenado bifuncional incorporado en la estructura final.

Este patrón de desconexión permite establecer una relación directa entre los productos obtenidos y sus bloques de construcción, manteniendo una estrategia común dentro de la serie derivada de chalconas.

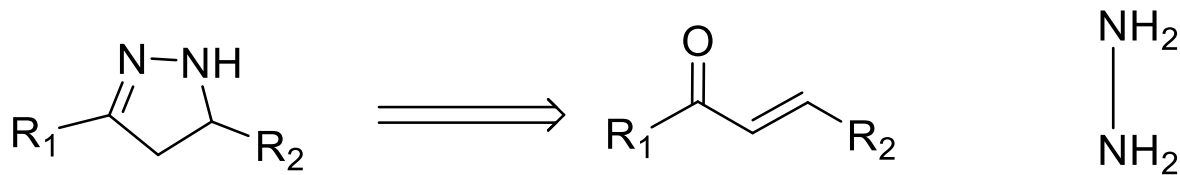


Figura 170 Análisis retrosintético de los derivados de dihidropirazol. La desconexión del anillo heterocíclico conduce a una chalcona como precursor α,β-insaturado y a hidrazina como fragmento nitrogenado bifuncional.

5.7.1.5 Benzodiazepinas

Los derivados de benzodiazepina pueden ser relacionados retrosintéticamente con una chalcona y o-fenilendiamina como precursores inmediatos (Figura 171). En este esquema, la chalcona actúa como precursor α,β-insaturado, mientras que la o-fenilendiamina aporta el fragmento nitrogenado bifuncional necesario para la formación del sistema heterocíclico.

A diferencia de los derivados dihidro, estos compuestos presentan un sistema aromático, lo cual implica una etapa adicional de eliminación que conduce a la aromatización del anillo, permitiendo la obtención de la estructura final más estabilizada.

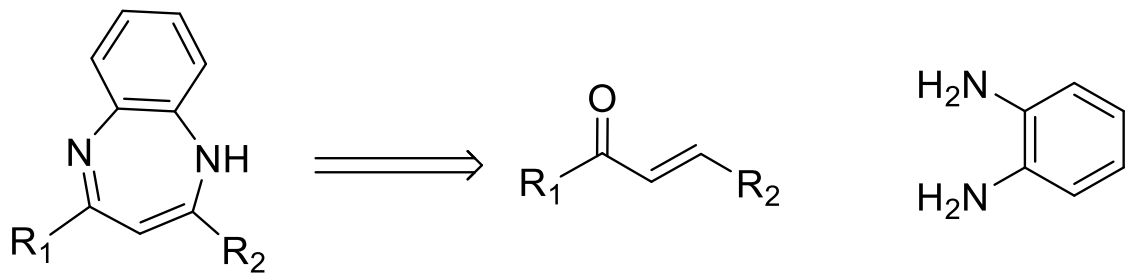


Figura 171 Análisis retrosintético de los derivados de benzodiazepina. La desconexión del sistema heterocíclico conduce a una chalcona como precursor α,β-insaturado y a o-fenilendiamina como fragmento nitrogenado bifuncional.

5.7.1.6 Pirazol

De manera análoga, los derivados de pirazol pueden ser desconectados retrosintéticamente en una chalcona y una hidrazina como precursores inmediatos (Figura 172). En este caso, la chalcona proporciona el sistema α,β-insaturado, mientras que la hidrazina actúa como fragmento nitrogenado bifuncional incorporado en la estructura final.

En contraste con los dihidropirazoles, la formación de pirazoles aromáticos involucra una etapa de eliminación adicional que conduce a la aromatización del anillo, generando estructuras más estables desde el punto de vista electrónico.

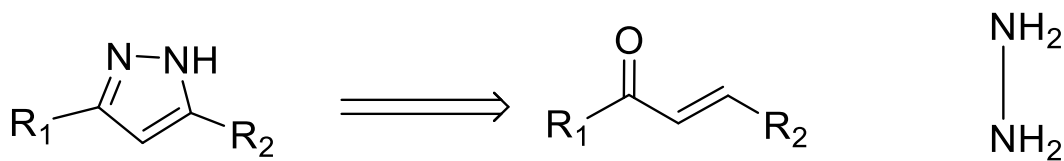


Figura 172 Análisis retrosintético de los derivados de pirazol. La desconexión del anillo aromático conduce a una chalcona y a hidrazina como precursores inmediatos, evidenciando la formación de un sistema heterocíclico aromático a partir de estos bloques de construcción.

5.7.1.7 Pirimidina-aminobenzimidazol

Los derivados amino-pirimidina-benzimidazol pueden ser relacionados retrosintéticamente con una chalcona y una guanidina funcionalizada con benzimidazol como precursores inmediatos. En la Figura 173, la chalcona actúa como sistema α,β-insaturado, mientras que la guanidina-benzimidazol constituye el fragmento nitrogenado responsable de la formación del sistema heterocíclico fusionado.

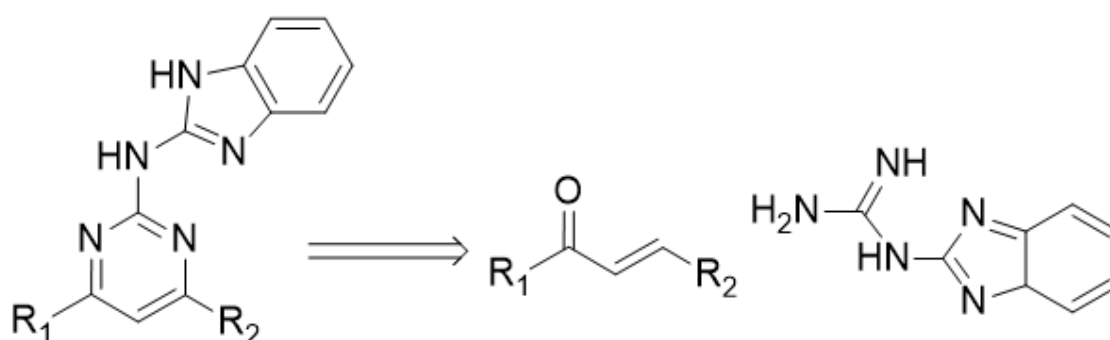


Figura 173 Análisis retrosintético de los derivados amino-pirimidina-benzimidazol. La desconexión del sistema heterocíclico conduce a una chalcona como precursor α,β -insaturado y a una guanidina funcionalizada con benzimidazol como fragmento nitrogenado, estableciendo una relación directa entre los productos finales y sus bloques de construcción.

5.7.1.8 Pirimido-benzimidazol

Los derivados pirimido-benzimidazol pueden ser relacionados retrosintéticamente con una chalcona y un 2-aminobenzimidazol como precursores inmediatos. En este esquema (Figura 174) la chalcona actúa como sistema α,β -insaturado, mientras que el 2-aminobenzimidazol constituye el fragmento nitrogenado que participa en la formación del sistema heterocíclico fusionado. Este patrón de desconexión permite establecer una correspondencia directa entre las estructuras finales y sus bloques de construcción, manteniendo una estrategia sintética común dentro de la serie derivada de chalconas.

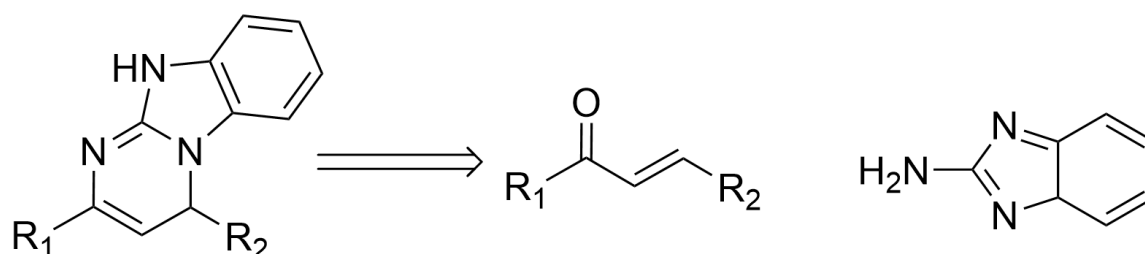


Figura 174 Análisis retrosintético de los derivados pirimido-benzimidazol. La desconexión del sistema heterocíclico fusionado conduce a una chalcona como precursor α,β -insaturado y a un 2-aminobenzimidazol como fragmento nitrogenado bifuncional, estableciendo una relación directa entre los productos finales y sus bloques de construcción.

5.7.2 Derivados de Benzimidazoles

Se estableció una estrategia retrosintética para los derivados benzimidazol-heteroaromáticos basada en la desconexión del enlace C–N que conecta el núcleo benzimidazol con el anillo heteroaromático (Figura 175). Esta desconexión conduce a dos fragmentos sintéticos principales: o-fenilendiamina y un aldehído heteroaromático correspondiente.

A partir de esta aproximación, se propone la formación del núcleo benzimidazol mediante una reacción de condensación entre o-fenilendiamina y el aldehído heteroaromático, generando el sistema heterocíclico sustituido. Posteriormente, los compuestos obtenidos fueron funcionalizados mediante una etapa de acilación, empleando cloruros de ácido aromáticos como agentes acilantes (Figura 176).

Los esquemas muestran la obtención de derivados N-acilados del benzimidazol, en los cuales se incorporan fragmentos aromáticos adicionales, manteniendo la conectividad entre el sistema benzimidazol y el heterociclo. La variación estructural se introduce tanto en el aldehído heteroaromático como en el agente acilante, permitiendo generar una serie de compuestos análogos.

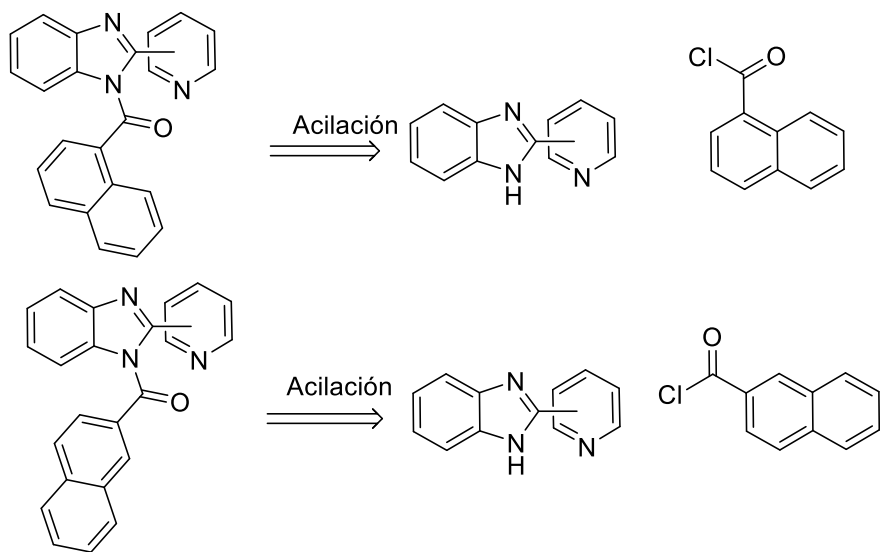


Figura 175 Estrategia de funcionalización del núcleo benzimidazol mediante acilación con cloruros de ácido aromáticos. El proceso permite la incorporación de fragmentos naftílicos, generando derivados N-acilados.

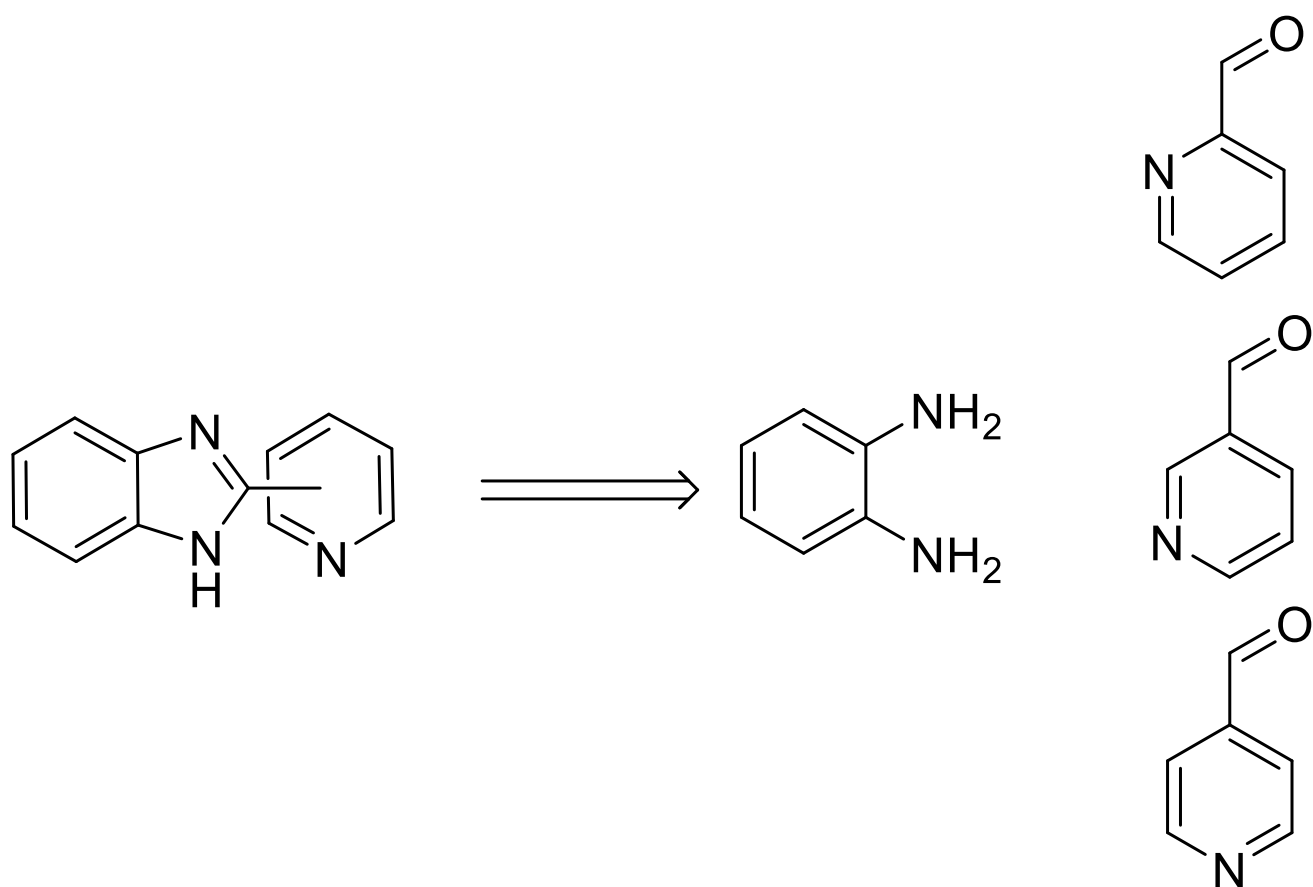


Figura 176 Análisis retrosintético de los derivados benzimidazol–heteroaromáticos. La estrategia se basa en la desconexión del enlace C–N clave, conduciendo a o-fenilendiamina y aldehídos heteroaromáticos como precursores, seguido de ciclación para formar el núcleo benzimidazol.

5.8 Avances sintéticos

5.8.1 Chalconas

5.8.1.1 Estandarización protocolo de síntesis

Para la reacción de tipo Claisen–Schmidt, se evaluaron distintas condiciones variando los equivalentes de KOH (Tabla 13). Se observó que, al emplear 5 equivalentes de base, se alcanza el mayor rendimiento en la síntesis de chalconas derivadas de acetilpireno y nitrobenceno.

Tabla 13 Rendimientos obtenidos con diferentes equivalentes de KOH y temperaturas durante 24 h.

Equivalentes de KOH	Rendimiento a RT 24 h	Rendimiento a 0 °C 24 h
0.1	trazas	40
0.5	trazas	46
1	0	70
2	0	87
5	0	91
10	0	87
20	0	83

5.8.1.2Chalcona obtenidas

Las chalconas obtenidas, derivadas de pireno, antraceno y acetofenona, se muestran en la Figura 177 Estructura de chalconas obtenidas., mientras que sus rendimientos se resumen en la Tabla 14. Los compuestos fueron caracterizados mediante espectroscopía de RMN de ¹H y ¹³CRMN(Anexo V), todos los compuestos sintetizados presentaron una adecuada concordancia entre las integraciones y las señales correspondientes en los espectros . No obstante, se observó una excepción en la molécula derivada de acetilpireno y p-dimetilaminobenzaldehído (chalcona pireno-DMA, ver Anexo VI), la cual presentó contaminación residual con dimetilamina (DMA). Esta impureza interfirió en la correcta integración de las señales, dificultando su interpretación espectral precisa.

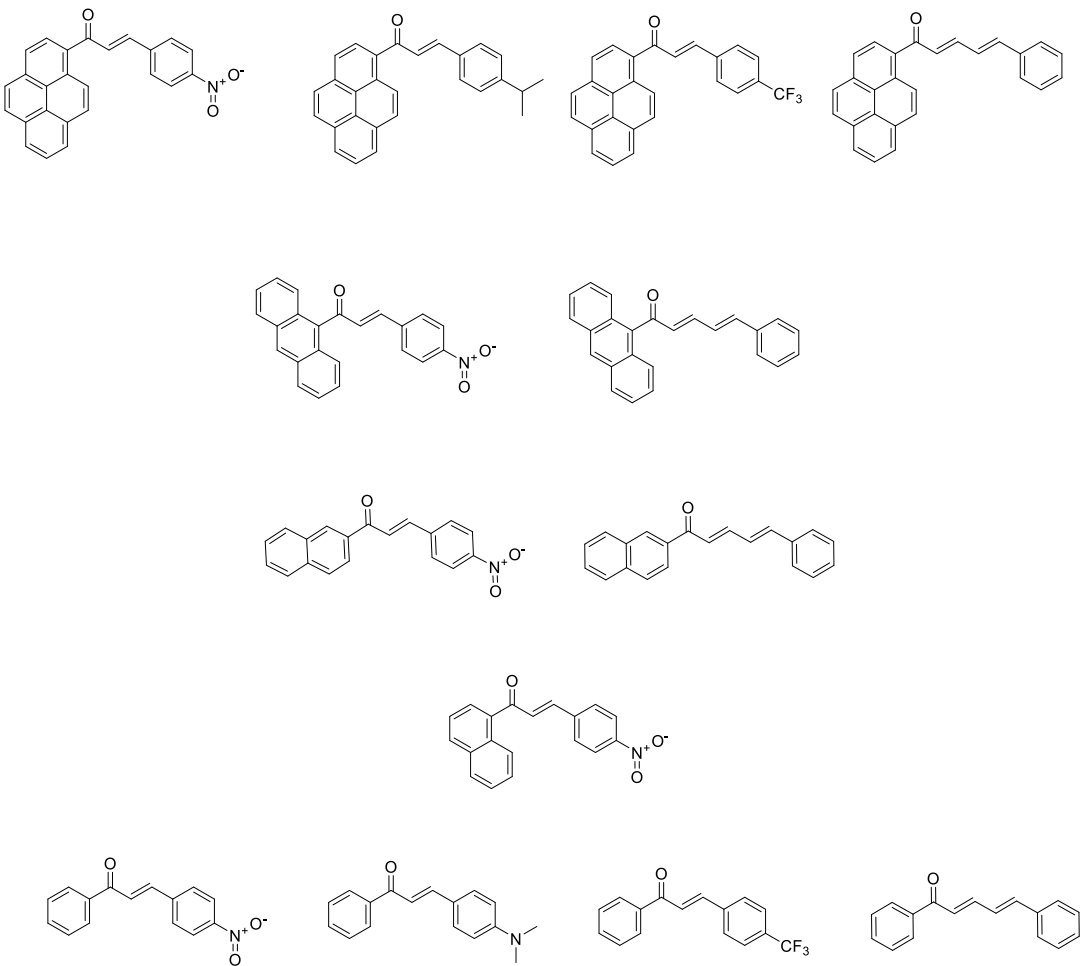


Figura 177 Estructura de chalconas obtenidas.

Tabla 14 Chalconas ordenadas según el fragmento acetilado y el aldehído empleado, con sus respectivos rendimientos.

Acetil	Aldehído	Rendimiento
Pireno	P-dimetilaminobenzaldehído (contaminado)	61.5%
	P-isopropilbenzaldehído	62.2%
	P-nitrobenzaldehído	91.0%
	Cinamaldehído	82.0%
Antraceno	P-nitrobenzaldehído	82.5%
	Cinamaldehído	81.8%
2 Naftoil	P-nitrobenzaldehído	84.3%
	Cinamaldehído	70.4%
1 Naftoil	P-nitrobenzaldehído	81.6%
Acetofenona	P-dimetilaminobenzaldehído	66.6%
	P-nitrobenzaldehído	83.5%
	P-trifluorometanobenzaldehído	85.3%
	Cinamaldehído	60.1%

5.8.1.3 Pirimido benzimidazol

Se obtuvo el derivado pirimido-benzimidazol correspondiente a la estructura mostrada en la Figura 178, con un rendimiento de 45%.

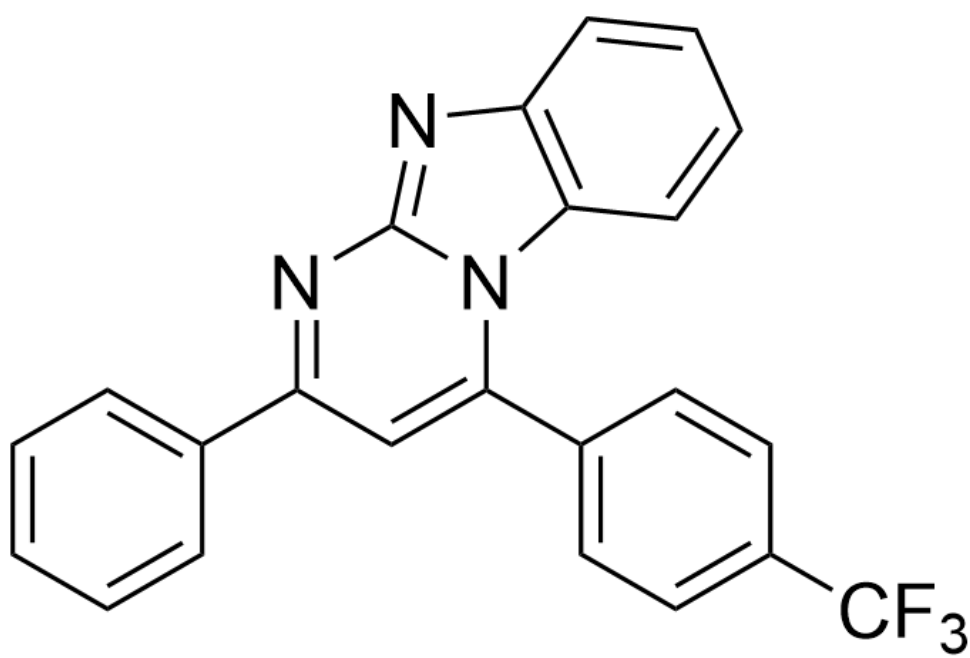


Figura 178 Estructura del derivado pirimido-benzimidazol

Los rangos obtenidos para punto de fusión, correspondieron a punto de descomposición : (249.5-254.0), (250.0-254.0) y (249.5-253.8). El infrarrojo se puede visualizar en la Figura 179. Para su tratamiento, se utilizó el software OMNIC, mediante sustracción de las señales correspondientes al solvente.

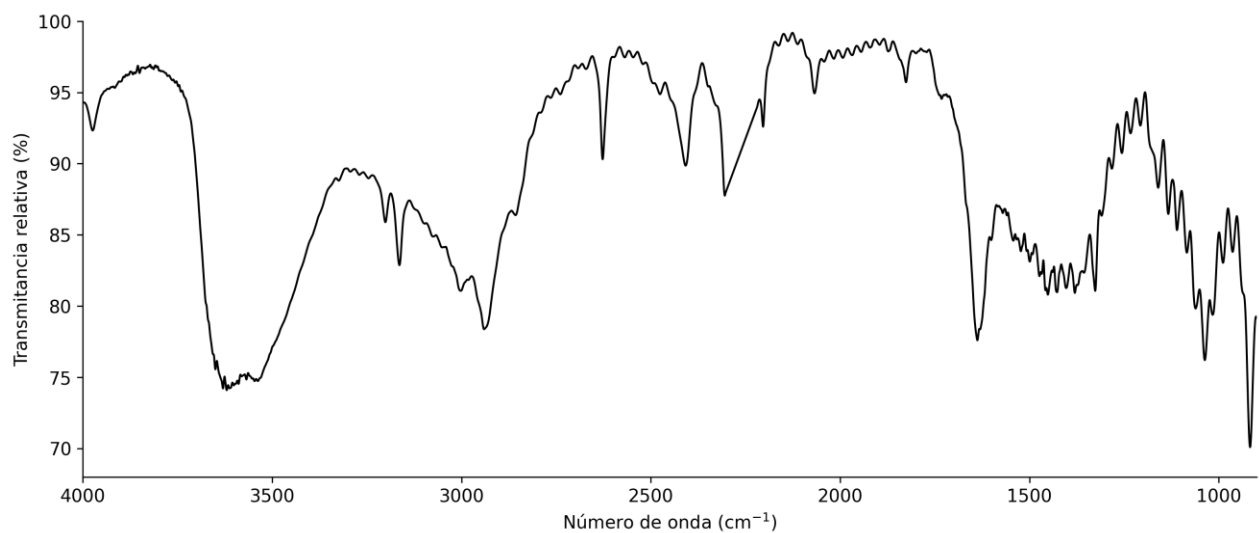


Figura 179 Espectro de infrarrojo de la molécula Pirimido benzimidazol con sustracción de solvente (acetonitrilo)

El espectro de ^1H RMN (Figura 180) muestra señales en la región aromática comprendida entre $\delta \approx 6.5\text{--}8.5$ ppm, consistentes con la presencia de múltiples sistemas aromáticos en la estructura. Las integraciones observadas son coherentes con el número esperado de protones, evidenciando un buen ajuste con la estructura propuesta.

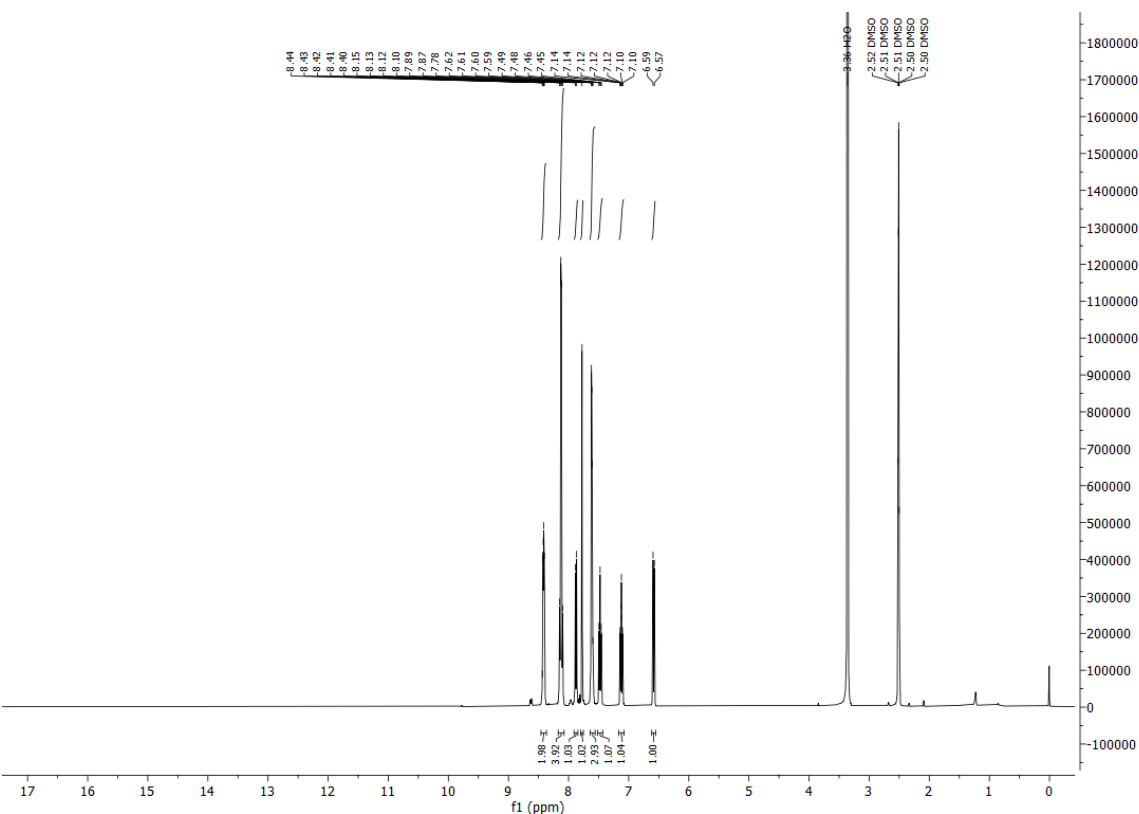


Figura 180 Espectro de ^1H RMN (400 MHz, DMSO-d_6) del derivado pirimido-benzimidazol. Se observan señales en la región aromática ($\delta \approx 6.5\text{--}8.5$ ppm), consistentes con los protones del sistema heteroaromático y anillos bencénicos presentes en la estructura.

En particular, se distinguen señales bien definidas en la zona de $\delta \approx 8.4\text{--}8.0$ ppm, atribuibles a protones aromáticos desprotegidos, junto con múltiples señales entre $\delta \approx 7.8\text{--}7.0$ ppm correspondientes al resto del sistema aromático.

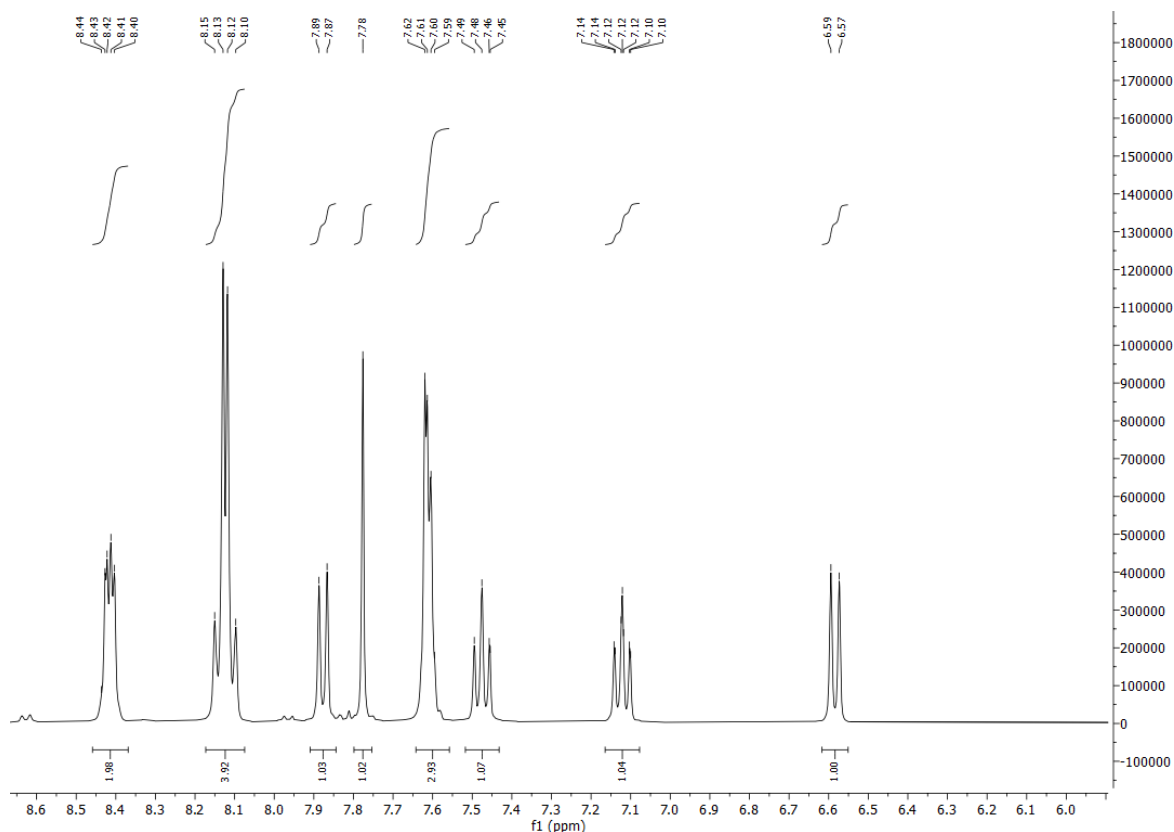


Figura 181 Ampliación de la región aromática del espectro de ^1H RMN (400 MHz, DMSO-d_6) del derivado pirimido-benzimidazol. Se distinguen los patrones de multiplicidad e integraciones correspondientes a los protones aromáticos, evidenciando la complejidad del entorno químico.

Por su parte, el espectro de ^{13}C RMN (Figura 182) presenta señales en la región de $\delta \approx 110\text{--}160$ ppm, correspondientes a carbonos aromáticos y heteroaromáticos. Se observa además la señal característica del carbono del grupo CF_3 en $\delta \approx 126.7$ (Figura 183), así como la señal del disolvente (DMSO) en la región de $\delta \approx 39$ ppm.

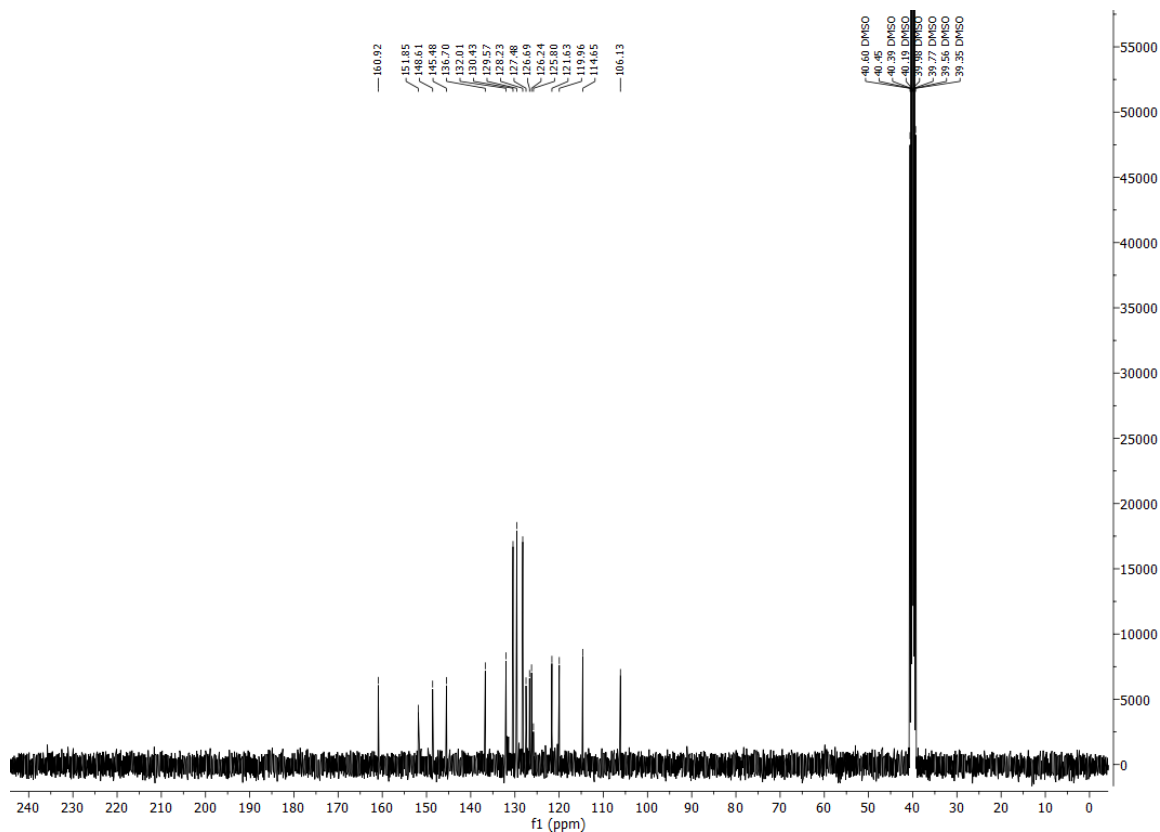


Figura 182 Espectro de ^{13}C RMN (100 MHz, DMSO-d_6) del derivado pirimido-benzimidazol. Se observan señales en la región de $\delta \approx 110\text{--}160$ ppm correspondientes a carbonos aromáticos y heteroaromáticos. La señal del disolvente (DMSO-d_6) se identifica alrededor de $\delta \approx 39$ ppm.

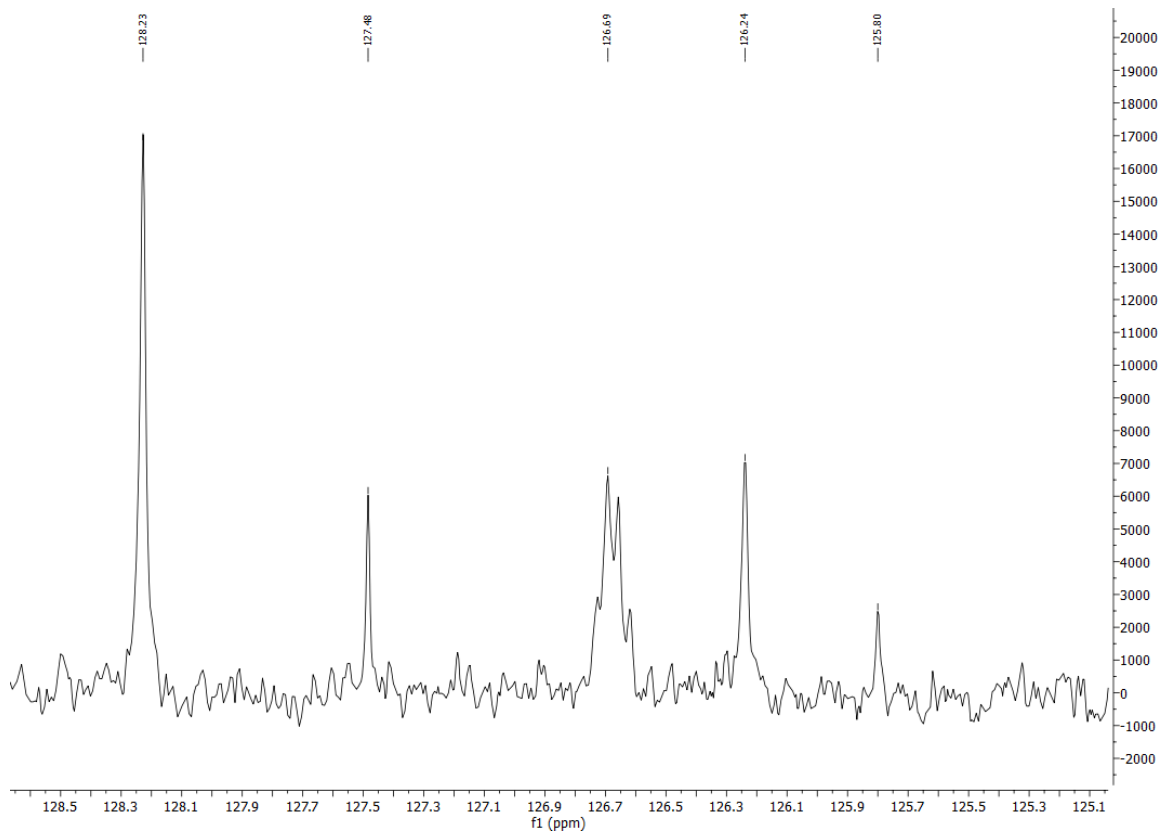


Figura 183 Ampliación del espectro de ^{13}C RMN (100 MHz, DMSO-d_6) del derivado pirimido-benzimidazol, mostrando el multiplete característico en la región de $\delta \approx 125\text{--}128$ ppm, correspondiente al carbono aromático acoplado al grupo CF_3 . El patrón de cuarteto observado se atribuye al acoplamiento $^{13}\text{C}\text{--}^{19}\text{F}$.

Los experimentos bidimensionales (COSY, HMBC y NOESY) evidencian correlaciones protón–protón y protón–carbono que respaldan la conectividad estructural del sistema.

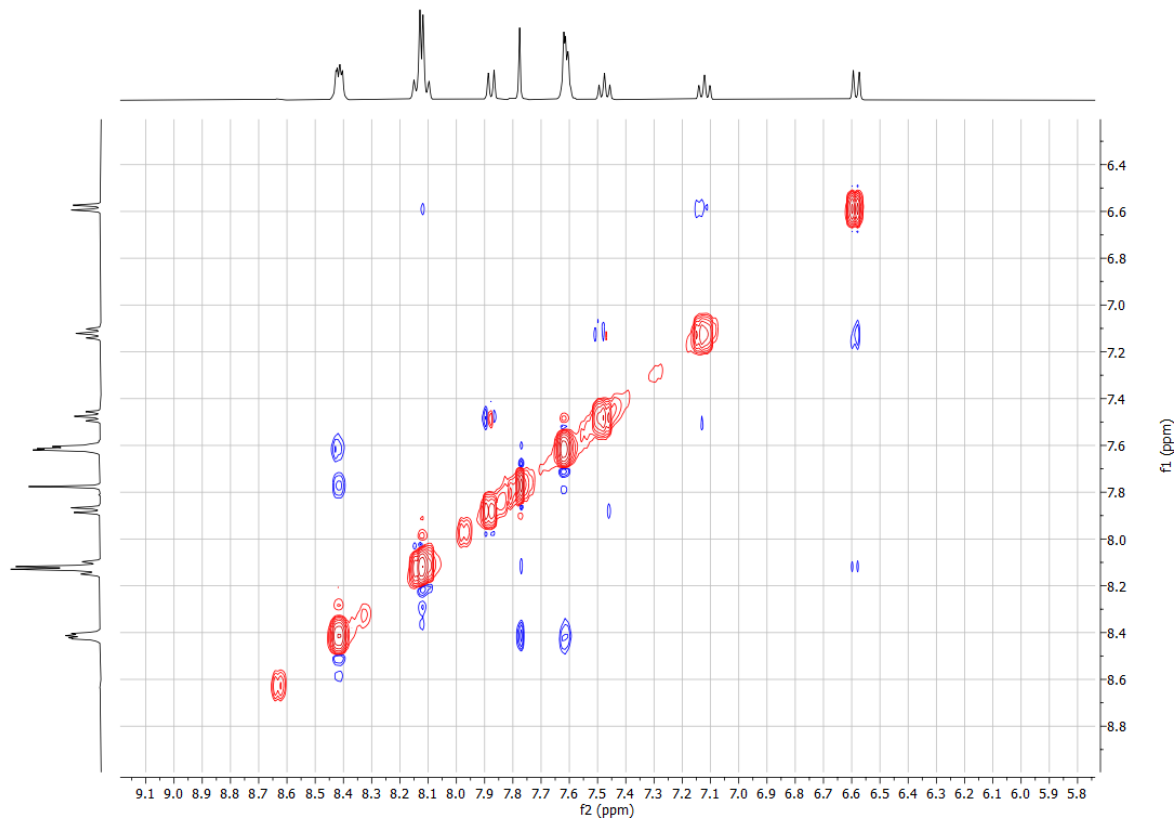


Figura 184 Espectro bidimensional $^1\text{H}\text{--}^1\text{H}$ COSY (400 MHz, DMSO-d_6) del derivado pirimido-benzimidazol, mostrando correlaciones entre protones acoplados a través de enlaces.

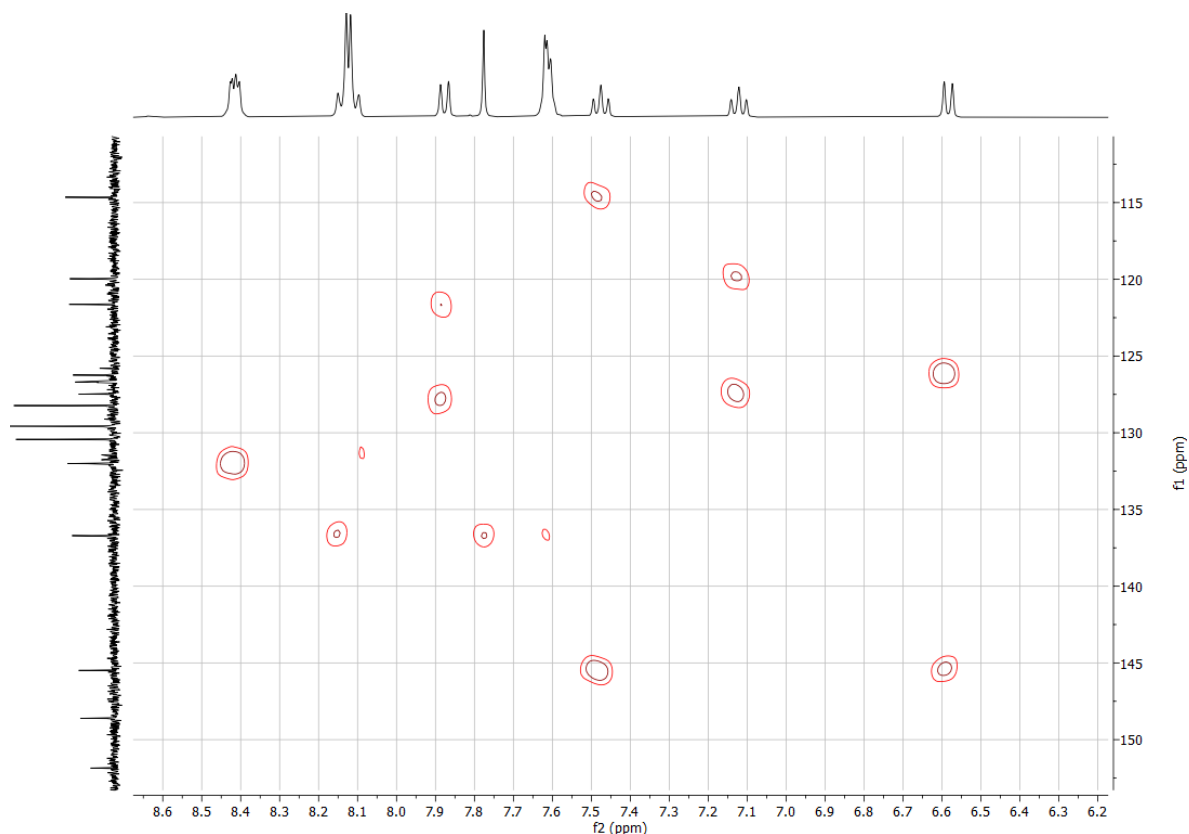


Figura 185 Espectro bidimensional ^1H - ^{13}C HMBC del derivado pirimido-benzimidazol, mostrando correlaciones de largo alcance que permiten establecer la conectividad entre protones y carbonos no directamente enlazados.

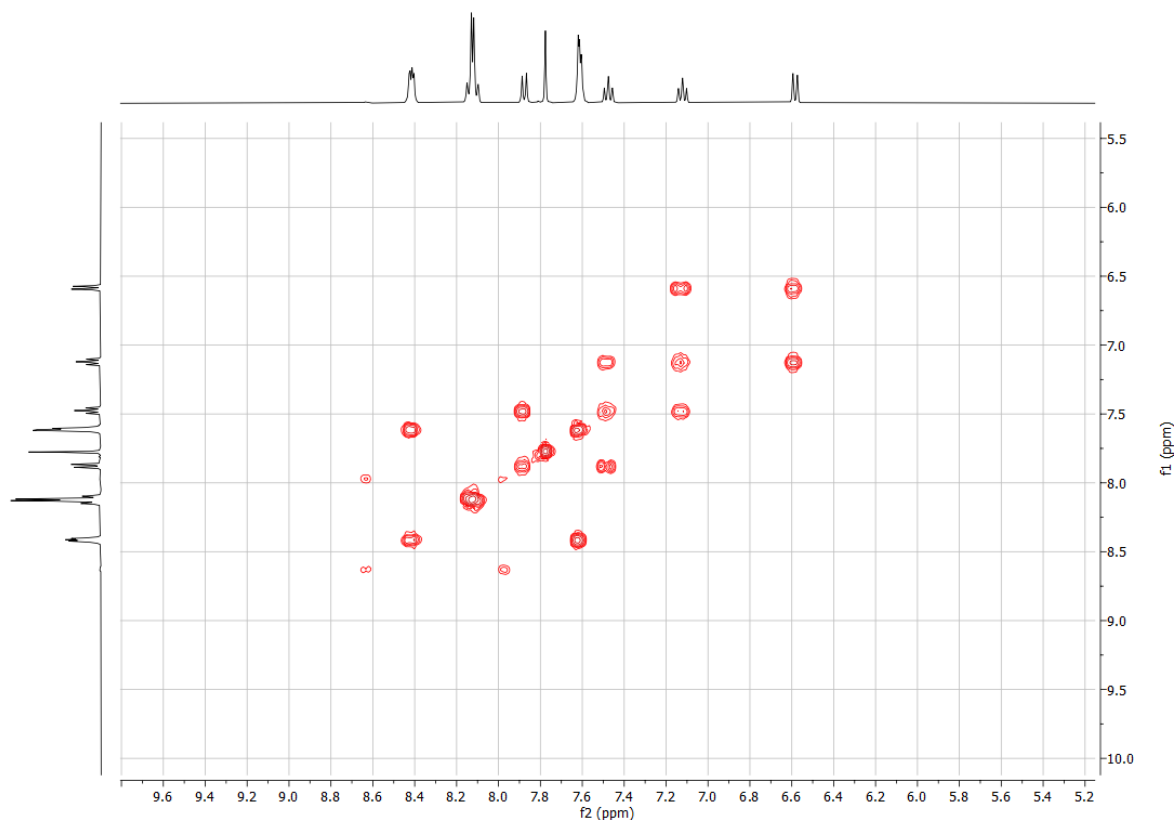


Figura 186 Espectro bidimensional ^1H - ^1H NOESY (400 MHz, DMSO-d_6) del derivado pirimido-benzimidazol, evidenciando correlaciones espaciales entre protones próximos.

Los datos espectroscópicos obtenidos son consistentes con la estructura propuesta. La asignación completa de protones y carbonos, basada en el análisis conjunto de los experimentos unidimensionales y bidimensionales de RMN, se presenta en el apartado de Discusión.

5.8.2 Benzimidazoles y derivados

5.8.2.1 Benzimidazoles

Los benzimidazoles fueron obtenidos por los tres métodos (Etanol en reflujo, agua en reflujo y mortero y pistilo en ácido acético), los rendimientos obtenidos para los 3 benzimidazoles se muestran en la Tabla 15.

Tabla 15 Rendimientos de los tres métodos de síntesis evaluados. Los compuestos fueron caracterizados vía ¹H y ¹³C(anexo V)

Producto	Método en agua reflujo	Mortero y pistilo en ácido acético	Etanol reflujo
Benzimidazol 2 piridina	24%	34%	44%
Benzimidazol 3 piridina	63%	54%	64%
Benzimidazol 4 piridina	91%	82%	76%

Los compuestos fueron caracterizados vía ¹H y ¹³C RMN Anexo V.

5.8.2.2 Benzimidazoles acilados

Los rendimientos obtenidos en cada acilación se resumen en la Tabla 16.

Tabla 16 Rendimientos obtenidos para cada acilación

Producto	cloruro de 1 naftoil	cloruro de 2 naftoil
Benzimidazol 2 piridina	62%	58%
Benzimidazol 3 piridina	-	63%
Benzimidazol 4 piridina	71%	-

Los espectros de ¹H RMN y ¹³C RMN obtenidos muestran una adecuada concordancia con la estructura propuesta.

6 Discusión

6.1 QSAR evolutivo y de selección manual

6.1.1 Evaluación QSAR de elección de descriptores manuales

6.1.1.1 Depuración del conjunto de datos y delimitación del dominio químico

La construcción del modelo QSAR manual descansa, en primer lugar, en una premisa fundamental de toda modelación estructura/actividad: la calidad del modelo no depende únicamente del algoritmo de regresión empleado, sino de la coherencia química del conjunto de datos que lo alimenta. En este sentido, el pretratamiento y depuración de la base de datos no debe interpretarse como una corrección cosmética destinada a mejorar artificialmente los parámetros estadísticos, sino como una etapa de definición del dominio químico sobre el cual la relación estructura/actividad puede considerarse válida. El modelo manual se desarrolla a partir de una base depurada y posteriormente validada con métricas internas y externas, lo que confirma que la ecuación final no surge de una correlación fortuita, sino de una señal estructura/actividad suficientemente estable.

Desde esta perspectiva de la exclusión de compuestos con residuos elevados cobra pleno sentido químico. Cuando una molécula se aparta sistemáticamente del comportamiento general de la serie, ello

puede reflejar que contiene determinantes estructurales no representados de manera adecuada por el resto del conjunto, o bien que responde a un mecanismo de interacción parcialmente distinto al predominante en la serie. Esto es especialmente relevante en familias químicas heterogéneas, donde ciertos derivados pueden incorporar heteroátomos, patrones de conjugación o arquitecturas rígidas que alteran de manera desproporcionada la relación entre propiedades globales y actividad biológica. En consecuencia, la eliminación de estos compuestos no solo mejora la estabilidad de la regresión, sino que evita mezclar en una misma ecuación señales químicas incompatibles entre sí.

6.1.1.2 Robustez estadística y validez predictiva del modelo manual

Una vez delimitado el dominio químico, el siguiente punto central es la validez estadística del modelo manual. Este modelo fue sometido a validación interna, validación externa y pruebas de aleatorización de la variable respuesta, observándose un ajuste lineal adecuado para entrenamiento, capacidad predictiva razonable sobre el conjunto de prueba y una distribución aleatoria de residuos en torno a cero, sin tendencias sistemáticas. Todo ello es consistente con un modelo robusto y no sesgado.

La coherencia entre R^2 , Q^2 y validación externa indica que la ecuación no solo reproduce el comportamiento del conjunto de entrenamiento, sino que conserva señal predictiva frente a compuestos no usados en el ajuste. Eso es crítico, porque en QSAR la principal amenaza no es construir una ecuación que describa bien los datos conocidos, sino construir una ecuación que parezca buena por memorizar ruido estructural. En ese sentido, la ausencia de comportamiento anómalo en el gráfico de residuos y el resultado favorable de Y-randomization sostienen que la relación encontrada tiene contenido informativo real.

Aun así, desde un enfoque prudente: ningún modelo lineal multivariado, por robusto que sea, demuestra causalidad estricta. Lo que demuestra es que, dentro del dominio químico estudiado, existe una asociación cuantificable y consistente entre ciertas propiedades moleculares y la actividad biológica. Esa distinción es importante, porque evita sobreinterpretar la ecuación como si describiera exhaustivamente el mecanismo molécula de inhibición. El QSAR no reemplaza el mecanismo; lo aproxima desde el espacio de descriptores.

La fortaleza del modelo manual reside en un doble equilibrio: por un lado, presenta respaldo estadístico suficiente para ser considerado predictivo; por otro, mantiene una complejidad moderada, compatible con una lectura química racional. Esa combinación es justamente la que lo vuelve valioso como herramienta de diseño.

6.1.1.3 Significado químico de los descriptores del modelo manual

La principal ventaja del modelo manual frente a aproximaciones más automatizadas es que sus descriptores poseen un significado químico directo. En otras palabras, la ecuación no solo predice, sino que también “habla” en un lenguaje químico comprensible. Esto constituye una fortaleza sustancial en química medicinal, donde el interés no recae únicamente en clasificar moléculas como más o menos activas, sino en comprender qué modificaciones estructurales tienen potencial para mejorar la actividad. De acuerdo con la interpretación preliminar, dentro del modelo manual destacan contribuciones ligadas a lipofilia, conteo o naturaleza de heteroátomos y otros términos de lectura fisicoquímica relativamente directa. Esa clase de descriptores es particularmente útil porque permite traducir la ecuación en hipótesis de diseño. Si la lipofilia contribuye positivamente, ello sugiere que una superficie molécula más hidrofóbica o aromática favorece el reconocimiento del blanco, probablemente al mejorar el acoplamiento en regiones apolares del sitio de unión. Si, en cambio, ciertos heteroátomos como oxígeno o flúor aparecen con coeficientes negativos, ello sugiere que no toda incorporación de polaridad o modulación electrónica resulta favorable, y que el efecto de estos átomos depende del contexto estructural en el que se encuentren.

Este punto merece una discusión cuidadosa. La contribución negativa de oxígeno o flúor no debe interpretarse como una regla absoluta del tipo “estos átomos disminuyen la actividad”, sino como una observación contextual: dentro de la serie analizada, su presencia parece asociarse a configuraciones que desplazan la molécula fuera del balance óptimo entre hidrofobicidad, distribución electrónica y geometría de reconocimiento. En términos de química medicinal, esto sugiere que el sistema no recompensa

simplemente el aumento de polaridad o la introducción de sustituyentes fuertemente moduladores, sino la sintonía fina entre capacidad de interacción y compatibilidad con el bolsillo.

Ese balance es especialmente verosímil en PLpro, considerando que el reconocimiento de inhibidores no covalentes se produce en una zona donde la topología del sitio, la presencia del loop BL2 y la organización de subsitios próximos al canal de entrada imponen restricciones combinadas de tamaño, aromaticidad y polaridad. En este contexto si una molécula es demasiado polar, puede perder afinidad porque le cuesta salir del agua y entrar bien al sitio de unión. En cambio, si su estructura no está bien orientada o es muy flexible, puede no acomodarse de la forma correcta y no aprovechar bien las interacciones hidrofóbicas y aromáticas

El modelo manual sugiere una interpretación importante: la actividad no depende de un único atributo dominante, sino del equilibrio entre contribuciones hidrofóbicas, electrónicas y estructurales. Esa lectura es mucho más valiosa que una simple lista de coeficientes, porque permite transformar la ecuación en criterio químico.

6.1.1.4 Utilidad del modelo manual para diseño racional

La verdadera prueba de relevancia de un modelo QSAR no es solo si alcanza buenos indicadores estadísticos, sino si genera conocimiento accionable. Bajo ese criterio, el modelo manual destaca por su utilidad para diseño racional. Al emplear descriptores interpretables, este modelo permite inferir qué tipos de modificaciones podrían desplazar la actividad en una dirección favorable y cuáles podrían deteriorarla. Esto lo convierte en una herramienta especialmente útil en etapas tempranas de optimización, donde el investigador necesita justificar por qué selecciona un nuevo derivado, por qué descarta otro y qué tendencia química desea explorar. A diferencia de modelos excesivamente abstractos, el QSAR manual puede integrarse con intuición química, literatura previa y conocimiento estructural del blanco. En otras palabras, sirve como puente entre estadística y química medicinal.

Además, este tipo de modelo tiene una virtud adicional: obliga a mantener una disciplina interpretativa. Como cada descriptor debe ser defendible, el modelo reduce el riesgo de caer en correlaciones de alto ajuste pero escasa plausibilidad química. Desde esta perspectiva, el valor del QSAR manual no radica únicamente en predecir actividades absolutas, sino en aportar una gramática estructural para pensar nuevas moléculas.

6.1.1.5 Limitaciones del modelo manual

Precisamente por su énfasis en interpretabilidad, el QSAR manual también presenta limitaciones. La primera es que la selección racional de descriptores, aunque químicamente sensata, puede dejar fuera combinaciones no obvias de variables que capturen patrones complejos de actividad. Dicho de otro modo, al privilegiar descriptores legibles, el modelo puede sacrificar parte de la sensibilidad estadística frente a relaciones no lineales o distribuciones más sutiles de propiedades a lo largo de la molécula.

La segunda limitación es que la linealidad impone una simplificación inevitable. En sistemas biológicos como PLpro, donde la unión depende de un entorno conformacional flexible y de una arquitectura de reconocimiento parcialmente inducida, es improbable que toda la variabilidad de la actividad pueda resumirse en unos pocos términos lineales. El mérito del modelo manual no es capturarlo todo, sino capturar una fracción significativa y químicamente inteligible del fenómeno.

Esta observación no invalida el modelo; más bien ayuda a ubicarlo: es un modelo particularmente potente como herramienta de interpretación y priorización inicial, más que como árbitro único del valor absoluto de pIC50.

6.1.2 Evaluación QSAR de elección de descriptores de selección evolutivo

6.1.2.1 Interpretabilidad frente a rendimiento predictivo

El empleo de un algoritmo genético para construir una segunda ecuación QSAR representa una ampliación metodológica importante y novedosa respecto del enfoque manual. Mientras el modelo manual explora un conjunto acotado de hipótesis químicamente intuitivas, el enfoque evolutivo permite

rastrear de manera sistemática un espacio descriptorial mucho más amplio, integrando variables topológicas, autocorrelacionales, electrotopológicas y de conteo estructural que serían difíciles de evaluar exhaustivamente de forma manual. La tesis explicita esta lógica al separar metodológicamente la generación de la ecuación mediante algoritmo genético, el refinamiento de modelos, el control de VIF y la selección de la ecuación final.

Este punto es crucial para la discusión: el enfoque evolutivo no debe presentarse como un sustituto “más moderno” del manual, sino como una estrategia complementaria para explorar patrones estructura-actividad que escapan a la intuición inmediata. En ese sentido, el valor del algoritmo genético no radica solo en encontrar una ecuación con mejor ajuste, sino en revelar qué regiones del espacio de descriptores contienen señal predictiva relevante.

6.1.2.2 Convergencia del algoritmo y significado del *plateau*

El análisis señalaría que el algoritmo muestra una tendencia de convergencia hacia un *plateau* de R^2 , con una dinámica donde el rendimiento crece, se estabiliza y luego alcanza una nueva meseta a generaciones más avanzadas. Esto permite argumentar que el algoritmo no se comportó de manera errática, sino que efectuó una exploración progresiva del espacio de soluciones hasta alcanzar regiones relativamente estables de desempeño.

El *plateau*, sin embargo, debe interpretarse con cuidado. No significa necesariamente que el algoritmo haya encontrado “la verdad” del sistema, sino que encontró una región del espacio “descriptorial” donde mejoras adicionales son marginales bajo la configuración utilizada. Esto es compatible con la noción de un óptimo local robusto. En modelación evolutiva, alcanzar una meseta puede considerarse una señal favorable de estabilización, siempre que el proceso no haya quedado atrapado prematuramente en una solución pobre. En este caso, la recurrencia de ciertos descriptores y la selección final de ecuaciones top apuntan a que la señal encontrada no fue arbitraria.

6.1.2.3 Recurrencia de descriptores y detección de señal estructural

Uno de los hallazgos más relevantes del enfoque evolutivo es la repetida aparición de determinados descriptores a lo largo de múltiples generaciones. Se detallan descriptores como AATS7i, ATSC0p, MATS5e, C3SP2, ETA_Beta_ns_d, AATSC0c, SaasC, n5Ring, n4Ring y nHBd, y se explica su significado químico y matemático.

La recurrencia de estas variables tiene implicancias fuertes. En primer lugar, indica que el algoritmo no está seleccionando rasgos de manera caprichosa, sino detectando atributos persistentes en la explicación de la actividad. En segundo lugar, sugiere que la actividad en esta serie no depende solo de propiedades globales simples, sino también de cómo las propiedades electrónicas y topológicas se distribuyen internamente a través del esqueleto molécula.

Una de las diferencias más relevantes respecto del QSAR manual radica en que una mayor lipofilia favorece la actividad o que ciertos heteroátomos la reducen en promedio, el modelo evolutivo parece sugerir algo más fino: no importa solo cuánto carácter electrónico o aromático tenga una molécula, sino cómo ese carácter se propaga a determinada distancia topológica, cómo se distribuye su polarizabilidad a lo largo de la red de átomos y qué arquitectura cíclica acompaña esa distribución.

Ese nivel de detalle es químicamente menos intuitivo, pero teóricamente muy potente. En sistemas aromáticos y heteroaromáticos, donde la conjugación, la resonancia y el patrón de sustitución alteran de forma marcada la densidad electrónica, tiene sentido que descriptores autocorrelacionales y electrotopológicos capturen mejor ciertas diferencias finas de actividad que no quedarían reflejadas en descriptores globales sencillos.

6.1.2.4 Discusión química de la ecuación evolutiva seleccionada

La ecuación final seleccionada por el algoritmo genético incorpora descriptores que pueden organizarse en varios bloques conceptuales. Un primer bloque corresponde a descriptores de autocorrelación ponderados por ionización, polarizabilidad o electronegatividad, como AATS7i, ATSC0p, MATS5e y

AATSC0c. Estos términos codifican la distribución espacial de propiedades electrónicas a distintas distancias topológicas.

Un segundo bloque corresponde a descriptores estructurales de conteo, como C3SP2, n5Ring, n4Ring y nHBd. Estos términos reflejan rasgos más directamente conectables con arquitectura molécula: grado de insaturación cíclica, presencia de anillos de distinto tamaño y capacidad de donar enlaces de hidrógeno. Un tercer bloque corresponde a descriptores de carácter electrotopológico y electrónico global, como ETA_Beta_ns_d y SaasC, que resumen cómo la disponibilidad electrónica y la electronegatividad del sistema se organizan en la molécula. En este contexto ETA_Beta_ns_d cuantifica la contribución de electrones no compartidos con capacidad de participar en resonancia, mientras que SaasC resume la electronegatividad de los carbonos del sistema.

Desde el punto de vista químico, este resultado es relevante, porque Sugiere que el algoritmo percibe como relevantes tres dimensiones simultáneas: la arquitectura aromática e insaturada del esqueleto, la capacidad del sistema para distribuir y modular densidad electrónica, y la presencia de grupos capaces de establecer interacciones dirigidas. Dicho de otra manera, el modelo evolutivo no premia simplemente moléculas “más grandes” o “más lipofílicas”, sino moléculas cuya organización electrónica y topológica parece especialmente adecuada para interactuar con el blanco.

En el contexto de PLpro, esto es razonable. El reconocimiento de inhibidores no covalentes alrededor del loopBL2 y de los subsitios próximos al canal de acceso no depende solamente de una interacción puntual, sino de cómo el ligando presenta simultáneamente superficie aromática, compatibilidad geométrica y un patrón electrónico que favorezca contactos estabilizantes sin incurrir en penalizaciones excesivas. La aparición de descriptores relacionados con carbonos sp^2 , resonancia y autocorrelación electrónica es congruente con la idea de que la planicidad parcial, la conjugación extendida y la distribución de polarizabilidad son factores centrales en la serie.

6.1.2.5 Control de multicolinealidad, refinamiento y criterio de selección final

Un aspecto que fortalece notablemente el enfoque evolutivo es que no se aceptó de manera automática la ecuación con mejor ajuste, sino que se aplicaron filtros adicionales de multicolinealidad y criterio químico. En la metodología y resultados se incluye explícitamente el control de VIF y la depuración de ecuaciones antes de elegir la final.

Este paso es esencial. En un entorno de alta dimensionalidad descriptorial, la correlación entre variables puede inflar artificialmente la estabilidad aparente de la regresión. El uso de VIF permite reducir esta redundancia y obliga al modelo a sostener su poder predictivo con variables menos solapadas entre sí. Esto argumentaría que el algoritmo genético no fue tratado como una caja negra, sino que fue sometido a un proceso de refinamiento racional.

Igualmente importante es la decisión de descartar ecuaciones con descriptores excesivamente abstractos o poco defendibles químicamente, aunque pudieran mostrar buen desempeño estadístico. Lo que se intenta hacer es que la mejor ecuación no es la que maximiza ciegamente una métrica, sino la que mantiene un equilibrio entre robustez matemática y capacidad de interpretación.

6.1.2.6 Potencial y límites del QSAR evolutivo

El modelo evolutivo tiene un poder indudable para capturar relaciones complejas dentro del conjunto de datos. Su principal fortaleza es que puede integrar simultáneamente variables que representan efectos electrónicos distribuidos, topología molécula y rasgos de conteo estructural, generando una lectura más rica del espacio químico que la ofrecida por un modelo manual restringido a pocas hipótesis. Además, parece producir predicciones más conservadoras en las familias nuevas, lo que puede interpretarse como una calibración más restrictiva del rango de actividad.

No obstante, esta misma fortaleza constituye también su principal límite. A medida que la ecuación incorpora descriptores más abstractos, disminuye la facilidad con que un químico puede traducirla en decisiones sintéticas concretas. Saber que una determinada combinación de autocorrelaciones ponderadas mejora el ajuste es valioso desde la perspectiva predictiva, pero menos inmediato desde la perspectiva de diseño racional.

Es por esto que el modelo evolutivo debe entenderse como una herramienta de alta resolución para exploración del espacio estructura/actividad, no necesariamente como un manual directo de síntesis. Su

utilidad máxima aparece cuando se integra con el modelo manual y con otras aproximaciones estructurales.

6.1.3 Comparación crítica entre QSAR manual y QSAR evolutivo

6.1.3.1 Interpretación vs predicción

La comparación entre ambos modelos revela una tensión metodológica clásica en QSAR: el compromiso entre interpretabilidad y rendimiento. El modelo manual ofrece descriptores más cercanos al lenguaje químico tradicional; el evolutivo, en cambio, explora relaciones de mayor complejidad y parece capturar matices estructurales menos evidentes. Ambos, por tanto, responden a necesidades distintas.

Desde un punto de vista estrictamente aplicado, el modelo manual es superior cuando el objetivo es orientar el diseño racional de análogos. Su lectura permite formular hipótesis claras: aumentar o modular aromaticidad, ajustar lipofilia, evitar sobrecargar de heteroátomos ciertas regiones, preservar un determinado balance estructural. En cambio, el modelo evolutivo es superior cuando el objetivo es detectar patrones de actividad complejos en una biblioteca amplia y filtrar compuestos con mayor exigencia estadística.

Esto no significa que uno sea “mejor” y el otro “peor”. Significa que cada uno ocupa un lugar distinto en la cadena de descubrimiento. El error sería juzgarlos con un único criterio. Si el único criterio fuera la legibilidad, el manual ganaría siempre. Si el único criterio fuera la sofisticación descriptorial, el evolutivo sería preferible. La tesis gana fuerza justamente porque no cae en esa simplificación.

6.1.3.2 Diferencias en las predicciones y calibración del espacio de actividad

Un punto particularmente interesante de los resultados es que ambos modelos parecen conservar tendencias generales semejantes, pero difieren en la magnitud absoluta de las actividades predichas, con un manual más alto y un evolutivo más conservador. Esa diferencia debe discutirse críticamente, no como una contradicción, sino como una señal de que ambos modelos calibran de manera diferente el espacio de actividad.

El modelo manual probablemente responde con mayor sensibilidad a descriptores globales que empujan la molécula hacia perfiles considerados favorables, lo que puede hacer que sus predicciones se expandan más. El modelo evolutivo, en cambio, al incluir términos que penalizan o modulan con más fineza la organización electrónica y topológica, parece actuar con mayor restricción. Esto podría indicar que el modelo evolutivo es más severo al momento de extrapolar combinaciones estructurales nuevas.

Desde un punto de vista práctico, esta diferencia es valiosa. Un modelo más expansivo puede ser útil para no perder candidatos prometedores en una etapa inicial; uno más conservador puede servir como segundo filtro para evitar optimismo excesivo. Así, en lugar de elegir uno u otro, la estrategia más sólida es utilizarlos secuencial o complementariamente.

6.1.3.3 Complementariedad de ambos enfoques

El modelo manual y el evolutivo no compiten por describir exactamente la misma realidad, sino que la abordan desde escalas distintas. El modelo manual captura la capa más visible y químicamente interpretable de la relación estructura/actividad. El evolutivo, por su parte, captura una capa más distribuida y abstracta, donde intervienen patrones de autocorrelación, resonancia y organización topológica menos evidentes.

Por eso, la lectura más madura no es escoger uno como “correcto” y descartar el otro, sino reconocer que ambos expresan dimensiones diferentes del mismo fenómeno. El manual responde mejor a la pregunta “¿qué rasgos químicos parecen favorecer la actividad?”. El evolutivo responde mejor a la pregunta “¿qué combinaciones cuantificables de propiedades discriminan mejor la actividad dentro del conjunto?”. En conjunto, ambos ofrecen una imagen más completa que cualquiera de los dos por separado.

6.1.3.4 Integración de los modelos QSAR manual y evolutivo

La evaluación conjunta de los modelos QSAR manual y evolutivo permite establecer una diferenciación clara en cuanto a su utilidad científica y operacional, evidenciando que ambos enfoques abordan la relación estructura/actividad desde perspectivas complementarias. En términos de entendimiento químico del sistema, el modelo manual presenta una ventaja significativa, ya que permite una interpretación directa de los determinantes estructurales de la actividad, facilitando la identificación de tendencias asociadas a propiedades fisicoquímicas reconocibles. En contraste, el modelo evolutivo destaca por su capacidad para explorar con mayor profundidad el espacio “descriptorial”, integrando combinaciones de variables que no resultan evidentes a partir de un análisis racional convencional.

Esta diferencia conceptual se traduce también en su aplicabilidad práctica. El modelo manual constituye una herramienta particularmente útil para orientar modificaciones estructurales concretas, permitiendo vincular de manera directa los cambios en la estructura molécula con su efecto esperado en la actividad. Por otro lado, el modelo evolutivo adquiere mayor relevancia como herramienta de filtrado computacional, al proporcionar una discriminación más estricta dentro del dominio químico definido, lo que resulta especialmente útil en etapas de priorización de compuestos.

Desde una perspectiva doctoral, esta distinción resulta fundamental, ya que evita una interpretación reduccionista basada exclusivamente en indicadores estadísticos. La calidad de un modelo QSAR no se limita a su capacidad de ajuste, sino que también depende de su contribución al entendimiento del sistema y de su utilidad en la toma de decisiones químicas. Bajo este criterio, el modelo manual presenta una mayor centralidad conceptual, mientras que el modelo evolutivo introduce una capa adicional de validación y refinamiento que fortalece la robustez del análisis.

En conjunto, ambos modelos configuran una plataforma QSAR complementaria que permite abordar la relación estructura/actividad desde distintos niveles de complejidad. El enfoque manual proporciona una lectura químicamente inteligible, a partir de la cual es posible establecer lineamientos claros para el diseño racional de nuevas moléculas. En contraste, el enfoque evolutivo amplía la capacidad de exploración del espacio estructural, identificando patrones complejos y aportando un criterio predictivo más exigente.

La integración de ambos enfoques permite reconocer que la actividad frente a PLpro no depende únicamente de propiedades globales simples, sino del equilibrio entre hidrofobicidad, organización aromática, distribución electrónica y arquitectura topológica de la molécula. En este sentido, la complementariedad entre interpretabilidad y optimización estadística no constituye una limitación, sino una fortaleza metodológica.

Por lo tanto, no resulta adecuado plantear una elección excluyente entre ambos modelos, sino más bien considerar su uso conjunto como una estrategia integradora. El modelo manual emerge como la herramienta principal para la interpretación química y el diseño racional, mientras que el modelo evolutivo actúa como un instrumento de exploración profunda y refinamiento del espacio predictivo. Esta aproximación evita tanto el reduccionismo químico como el estadístico, y refuerza la solidez global del análisis QSAR en el contexto del diseño de nuevos inhibidores.

6.2 Virtual Screening

6.2.1 Validación del protocolo de docking mediante la molécula de referencia

Uno de los aspectos críticos en cualquier estudio de virtual screening es la validación del protocolo de acoplamiento, particularmente en términos de su capacidad para reproducir modos de unión experimentalmente conocidos. En este contexto, el *redocking* del inhibidor de referencia GRL-0617 (Figura 28) mostró un valor de RMSD bajo, lo que indica una alta concordancia entre la pose predicha y la conformación experimental reportada.

Este resultado no solo valida la parametrización del sistema (función de *scoring*, caja de búsqueda y condiciones del docking), sino que además sugiere que el protocolo es capaz de capturar adecuadamente las interacciones clave dentro del sitio S3/S4 de PLpro, incluyendo la estabilización del loopBL2. La correcta reproducción de esta interacción es particularmente relevante, dado que la inhibición de PLpro depende fuertemente de la ocupación estérica de este bolsillo más que de una interacción directa con la tríada catalítica.

El bajo RMSD obtenido respalda la aplicabilidad del protocolo para el muestreo masivo de compuestos, proporcionando confianza en la interpretación de los rankings de afinidad obtenidos posteriormente.

6.2.2 Robustez y alcance del espacio químico explorado

El virtual screening realizado se caracteriza por un enfoque exhaustivo, abarcando múltiples familias estructurales relevantes en química medicinal, incluyendo benzotiofenos, benzamidas, benzimidazoles, benzodiazepinas, aminopirimidinas, pirimidinas, aminobenzimidazoles y derivados tipo pirimido benzimidazol.

La selección de estas familias no es arbitraria, sino que responde a dos criterios fundamentales: Relevancia farmacológica conocida, Accesibilidad sintética y versatilidad estructural.

Núcleos como benzamidas, benzotiofenos y benzimidazoles están ampliamente representados en fármacos aprobados, lo que sugiere propiedades fisicoquímicas favorables para interacción proteína-ligando. También, sistemas derivados de chalconas y heterociclos condensados (como pirimido-benzimidazoles) permiten una rápida diversificación estructural, lo que los convierte en candidatos ideales dentro de una línea de investigación QSAR-*Docking*-Dinámica molecular-síntesis.

Adicionalmente, la magnitud del cribado, que incluyó bases de datos extensas (alcanzando incluso millones de compuestos en algunas familias como benzamidas), refuerza la representatividad del espacio químico explorado. Este nivel de exhaustividad permite considerar que los resultados obtenidos no corresponden a un muestreo limitado, sino a una exploración amplia y sistemática del espacio de posibles inhibidores.

6.2.3 Consideraciones computacionales y escalabilidad del método

Un aspecto frecuentemente subestimado en estudios de virtual screening es el costo computacional asociado. En este trabajo, el uso de AutoDockGPU permitió reducir significativamente los tiempos de cálculo, haciendo viable el análisis de bibliotecas de gran tamaño en un entorno académico.

No obstante, es importante destacar que el tiempo total del proceso (mostrado en la Tabla 6) no está dominado exclusivamente por el docking, sino que también por las etapas previas de preparación molécula, particularmente la conversión de archivos desde formatos como SDF a PDBQT. Este paso, especialmente en bibliotecas masivas (por ejemplo, benzamidas), puede extenderse hasta uno o dos días, constituyendo un cuello de botella relevante.

El flujo de trabajo completo demuestra ser escalable y reproducible a nivel de laboratorio, lo cual representa una ventaja significativa frente a enfoques que requieren infraestructura de alto rendimiento exclusiva. Esto posiciona la metodología empleada como una herramienta accesible para el descubrimiento temprano de fármacos.

6.2.4 Análisis de afinidad: influencia estructural y coherencia con QSAR

El análisis de los rankings de afinidad revela tendencias estructurales claras entre las diferentes familias evaluadas. En particular, las benzodiazepinas destacan como la familia con mejores puntajes de unión, alcanzando valores cercanos a -14, seguidas por derivados pirimido-benzimidazol.

Esta superioridad en afinidad puede explicarse principalmente por: Mayor extensión aromática, mayor superficie de interacción π - π , capacidad de ocupar eficientemente el bolsillo hidrofóbico.

Estas observaciones son altamente consistentes con el modelo QSAR desarrollado previamente, especialmente el modelo manual, el cual indica que un aumento en la contribución de descriptores asociados a aromaticidad y volumen molécula correlaciona positivamente con la actividad biológica.

El hecho de que ambos enfoques (QSAR y docking) converjan en la misma interpretación estructural constituye un resultado de alto valor científico, ya que: Refuerza la validez de los modelos predictivos, reduce la probabilidad de falsos positivos y proporciona una base racional para futuras modificaciones estructurales.

También se observa que los compuestos top dentro de distintas familias (benzamidas, benzotiofenos, etc.) comparten una característica común: alta densidad de anillos aromáticos, lo que sugiere que este rasgo estructural es un determinante clave de afinidad en PLpro.

6.2.5 Análisis de poses: convergencia estructural y especificidad del sitio de unión

El análisis cualitativo de las poses revela que no solo la energía de unión es relevante, sino también la localización del sitio de interacción.

Se identifican dos comportamientos principales:

a) Compuestos con alta afinidad y correcta localización.

Moléculas como la benzodiazepina (D) presentan una superposición significativa con la pose del inhibidor de referencia (GRL-0617), ocupando el bolsillo de inhibición. Este comportamiento es coherente con sus altos puntajes de afinidad y sugiere un mecanismo de inhibición análogo.

b) Compuestos con baja afinidad y sitio alternativo.

En contraste, compuestos como la aminopiridina (B) muestran poses desplazadas hacia regiones distintas del sitio activo, lo que explica sus menores energías de unión. Esto indica que no solo la estructura química, sino también la orientación espacial, determina la efectividad del acoplamiento.

Este análisis pone en evidencia que: La afinidad energética debe interpretarse en conjunto con la pose, la ocupación del sitio correcto es determinante para la inhibición. Existen falsos positivos potenciales si solo se considera el score.

6.2.6 Implicancias para el diseño racional de inhibidores

Los resultados obtenidos a partir del virtual screening permiten establecer directrices estructurales claras para el diseño racional de nuevos inhibidores de PLpro, las cuales emergen de la integración entre el análisis de afinidad, la evaluación de las poses de unión y la consistencia con los modelos QSAR previamente desarrollados.

Uno de los hallazgos más consistentes del estudio es la relevancia de la aromaticidad extendida como determinante principal de la afinidad. Los compuestos con mejores puntajes presentan múltiples sistemas aromáticos conjugados, lo que sugiere que la interacción con la cavidad de unión de PLpro está dominada por contribuciones de tipo hidrofóbico y π - π . Esta observación es coherente con la naturaleza del entorno proteico, el cual favorece la estabilización de ligandos con superficies aromáticas amplias capaces de establecer interacciones dispersivas y de apilamiento con residuos clave.

En este contexto, la planaridad molécula emerge como un factor crítico, ya que permite una mejor complementariedad geométrica con la cavidad de unión. Las estructuras planas facilitan un acoplamiento más eficiente al reducir la penalización conformacional y maximizar el área de contacto con la proteína. Esto explica por qué familias como las benzodiazepinas y los sistemas fusionados tipo pirimido-benzimidazol muestran desempeños superiores en comparación con estructuras más simples o menos conjugadas.

El análisis de las poses refuerza esta interpretación, evidenciando que los compuestos con mayor afinidad no solo presentan mejores energías de unión, sino que además adoptan orientaciones consistentes dentro de la cavidad funcional de la proteína. En particular, aquellos compuestos que logran reproducir la región de interacción del inhibidor de referencia tienden a mostrar un comportamiento más favorable, lo que sugiere que la correcta ocupación espacial de esta zona es un requisito indispensable para una inhibición efectiva. En contraste, se observa que existen moléculas que, a pesar de presentar características químicas globales similares, como un alto grado de aromaticidad o tamaño comparable, no logran adoptar una orientación espacial adecuada dentro de la cavidad de unión, lo que se traduce en una disminución significativa de su afinidad. Este comportamiento evidencia que la afinidad no depende exclusivamente de la composición química del ligando, sino también de su capacidad para establecer una disposición geométrica compatible con el entorno proteico. En consecuencia, propiedades como la conectividad entre fragmentos, la distribución tridimensional de los sistemas aromáticos y la flexibilidad conformacional adquieren un rol determinante en la eficiencia del acoplamiento.

Otro aspecto relevante es la naturaleza de las interacciones predominantes. Los resultados indican que la afinidad está principalmente gobernada por interacciones hidrofóbicas extensas complementadas por contactos π - π , más que por interacciones polares específicas. Esto implica que, en el diseño de nuevos inhibidores, debe priorizarse la optimización de la superficie hidrofóbica y la distribución electrónica del sistema aromático, en lugar de una sobrecarga de grupos funcionales polares que podrían no contribuir significativamente a la estabilidad del complejo.

Por otro lado, se observa que las estructuras con alta flexibilidad conformacional o bajo peso molecular tienden a presentar un desempeño inferior. Estas moléculas muestran una mayor dispersión en sus poses y, en muchos casos, no logran establecer interacciones estables dentro de la cavidad de unión. Esto sugiere que una rigidez estructural moderada es esencial para el anclaje efectivo, ya que permite

mantener una orientación definida y reducir la penalización entrópica asociada a la unión. En consecuencia, el diseño de nuevos compuestos debe evitar tanto la excesiva flexibilidad como la simplicidad estructural, favoreciendo arquitecturas que combinen rigidez y extensión aromática.

Un elemento particularmente robusto del estudio es la convergencia entre los resultados de docking y el modelo QSAR, especialmente en relación con descriptores asociados a aromaticidad, volumen molécula y topología. Esta coincidencia entre enfoques independientes refuerza la validez de las conclusiones obtenidas, ya que indica que las tendencias observadas no son artefactos metodológicos, sino reflejan propiedades intrínsecas del sistema proteína-ligando.

En este sentido el flujo de trabajo implementado (QSAR , virtual screening , docking y luego Dinámica molecular) demostraría ser conceptualmente sólido y altamente eficiente para la exploración del espacio químico. Este enfoque permite una optimización iterativa de los compuestos, donde las características estructurales identificadas como favorables pueden ser refinadas progresivamente en cada etapa del proceso. Así, el virtual screening no solo cumple una función de filtrado, sino que se convierte en una herramienta generadora de conocimiento estructural aplicable al diseño racional.

6.2.7 Resultados con respecto al QSAR

La coincidencia observada para aminopirimidina y benzimidazol sugiere que ambos modelos reconocen de manera consistente ciertos rasgos estructurales asociados a baja y alta actividad relativa, respectivamente. En contraste, las diferencias observadas para pirimido-benzimidazol y pirimidina-aminobenzimidazol indican que el modelo manual y el evolutivo no ponderan del mismo modo los atributos vinculados al extremo alto de actividad. En este sentido, el modelo manual parece favorecer una separación más marcada entre los compuestos mejor posicionados, mientras que el modelo evolutivo muestra una predicción más conservadora. En conjunto, esto respalda el uso complementario de ambos enfoques durante la priorización de compuestos.

6.2.8 Limitaciones del estudio

A pesar de la robustez del enfoque empleado, es importante reconocer las limitaciones inherentes a las metodologías utilizadas. En primer lugar, el docking molecular considera una representación simplificada del sistema proteico, en la cual la flexibilidad conformacional de la proteína es limitada. Este aspecto es particularmente relevante en proteínas como PLpro, donde regiones móviles pueden influir significativamente en la interacción con ligandos.

Asimismo, las funciones de scoring utilizadas presentan limitaciones en la estimación cuantitativa de las energías de unión. Estas funciones están diseñadas principalmente para establecer comparaciones relativas entre compuestos, por lo que las diferencias absolutas en energía deben interpretarse con cautela. En este sentido, pequeñas variaciones en el score no necesariamente reflejan diferencias reales en afinidad biológica.

El docking no considera explícitamente factores farmacocinéticos ni biológicos complejos, tales como solubilidad, permeabilidad, metabolismo o toxicidad. Esto implica que una alta afinidad teórica no garantiza necesariamente una actividad biológica efectiva en sistemas experimentales.

También considerar que las moléculas evaluados son reportadas en literatura, por lo que se limita a un espacio químico de molécula sintetizada, las moléculas teóricas son excluidas de este análisis.

6.3 Docking inducido (*Induced fit docking*)

6.3.1 Análisis de derivados de chalconas

El induced fit docking constituyó una etapa de refinamiento estructural indispensable dentro del flujo de trabajo empleado, ya que permitió reevaluar la compatibilidad de los ligandos seleccionados bajo un esquema más realista que el docking rígido previo. A diferencia del virtual screening inicial, en este caso

no solo se permitió la flexibilidad del ligando, sino también la reorganización local de residuos cercanos al sitio de unión, lo que resulta particularmente pertinente en PLpro dada la plasticidad conformacional de regiones adyacentes al bolsillo de reconocimiento, en especial el loop BL2. Desde este punto de vista, el induced fit docking no debe entenderse simplemente como una repetición de docking con mayor costo computacional, sino como una aproximación de mayor exigencia estructural, donde la calidad del acoplamiento depende de la capacidad del ligando para sostener interacciones favorables incluso después de que la proteína se adapta a su presencia.

La validación con GRL-0617 respalda esta interpretación. El hecho de que la pose inducida reproduzca razonablemente la orientación cristalográfica con un RMSD cercano a 1.6 Å indica que el protocolo conserva el modo general de reconocimiento del inhibidor de referencia aun cuando se permite la relajación conformacional del entorno proteico. Este punto es relevante porque, en un contexto de flexibilidad proteína-ligando, una desviación moderada respecto de la estructura cristalográfica no implica necesariamente una pérdida de validez del método; por el contrario, puede reflejar la exploración de un estado conformacional alternativo pero físicamente plausible del complejo. En consecuencia, la validación obtenida no solo confirma que el protocolo fue correctamente implementado, sino que además legitima el uso de estas conformaciones refinadas como base para la comparación entre familias químicas. Más aún, si el patrón esencial de reconocimiento del ligando de referencia se mantiene tras el ajuste inducido, entonces los compuestos que logran aproximarse a esa lógica espacial adquieren mayor relevancia mecanística que aquellos que solo obtienen un buen score en una proteína tratada como rígida.

En términos globales, los resultados muestran que el refinamiento inducido no benefició por igual a todas las series, lo que en sí mismo es un hallazgo importante. Si una familia mantiene o incluso mejora su comportamiento al pasar desde una evaluación inicial de acoplamiento a una etapa donde la proteína puede reorganizarse localmente, ello sugiere que su afinidad no depende de una coincidencia geométrica accidental con una conformación fija, sino de una compatibilidad estructural más profunda con el entorno de PLpro. Esta es precisamente una de las mayores fortalezas del induced fit docking dentro del diseño racional: permite distinguir entre ligandos que “caben” en el bolsillo y ligandos que realmente son tolerados y estabilizados por la proteína tras un proceso de reconocimiento mutuo.

En el caso de las chalconas y sus derivados heterocíclicos, la discusión de los resultados sugiere que el mejor desempeño de determinadas familias no puede atribuirse únicamente al incremento inespecífico de aromaticidad, aunque ese factor ciertamente contribuye. Lo que emerge con mayor claridad es que las familias más favorecidas son aquellas que combinan extensión aromática, rigidez relativa y una distribución espacial capaz de adaptarse con eficiencia al bolsillo de unión sin sacrificar complementariedad tras el ajuste inducido. En este sentido, el destacado comportamiento de la familia pirimidina-aminobenzimidazol resulta especialmente significativo, ya que una mediana superior al resto de las familias indica no solo la presencia de compuestos extremos muy favorables, sino una robustez de andamiaje a nivel poblacional. En términos de química medicinal, esto es más valioso que un solo top hit aislado, porque sugiere que el núcleo central de esa familia es intrínsecamente compatible con la arquitectura de unión de PLpro y que pequeñas variaciones sobre ese esqueleto tienden a conservar una afinidad razonable. Ese tipo de comportamiento es precisamente el que interesa para etapas posteriores de optimización racional.

La influencia del fragmento derivado de la cetona también ofrece una lectura estructural convincente. El mejor comportamiento de los derivados de 1-acetilnaftaleno es coherente con la hipótesis de que ciertos sistemas aromáticos fusionados presentan una geometría particularmente apta para ser estabilizada dentro del bolsillo de PLpro una vez que la proteína se reorganiza localmente. Este resultado no debe reducirse a la afirmación simple de que “más anillos aromáticos implican mejor score”, porque no todos los sistemas extendidos se comportan igual. Lo que parece favorecer al motivo 1-naftilo es una combinación más específica de superficie hidrofóbica, rigidez conformacional y direccionalidad espacial. Su semejanza conceptual con el fragmento naftalénico presente en GRL-0617 refuerza además la idea de que el bolsillo de PLpro reconoce favorablemente este tipo de arquitectura aromática. No se trata de afirmar una equivalencia farmacofórica estricta entre ambos sistemas, pero sí de reconocer que la topología del 1-acetilnaftaleno parece alinearse de manera particularmente eficiente con los requerimientos geométricos del sitio de unión. Desde una perspectiva de diseño, este resultado lo posiciona como un fragmento privilegiado para desarrollos posteriores.

La influencia del fragmento derivado del aldehído introduce un matiz aún más interesante, porque revela que la tendencia central de una serie y sus mejores compuestos individuales no necesariamente coinciden. El mejor comportamiento global de los salicílicos, reflejado en una mediana más favorable,

sugiere que este sustituyente confiere una ventaja más uniforme a lo largo de la serie. Esa uniformidad probablemente refleja una combinación equilibrada entre aromaticidad, polaridad moderada y capacidad de orientar la molécula dentro del bolsillo sin imponer penalizaciones estéricas excesivas. En contraste, el hecho de que los derivados trifluorometilados no lideren la distribución global, pero sí generen algunos de los compuestos individuales con mejores puntajes, indica un comportamiento más dependiente del contexto molécula. Es decir, el grupo trifluorometilo no actúa como un beneficio generalizado, sino como un modulador que puede producir un efecto altamente favorable solo cuando se incorpora sobre una arquitectura que ya posee una orientación y una distribución electrónica adecuadas. Este hallazgo es especialmente valioso en discusión, porque obliga a abandonar una lógica simplista de “sustituyente bueno” o “sustituyente malo” y reemplazarla por una visión más realista, donde el rendimiento depende de sinergias estructurales entre fragmentos. En otras palabras, el induced fit docking pone en evidencia que el comportamiento del ligando no es aditivo, sino emergente.

Esto se vuelve todavía más claro cuando se consideran las combinaciones R^1/R^2 . Los resultados sugieren que la afinidad final no depende de la contribución aislada de cada fragmento, sino de la manera en que ambos modulan conjuntamente la orientación global, la rigidez del sistema y la capacidad de estabilización dentro del bolsillo. Así, la aparición de compuestos trifluorometilados entre los mejores puntajes individuales, pese a una mediana menos destacada, puede entenderse como evidencia de combinaciones estructurales privilegiadas más que de un efecto universal del grupo. Desde el punto de vista metodológico, este es precisamente uno de los mayores aportes del induced fit docking frente a enfoques más simples: permite detectar relaciones no lineales entre fragmentos, mostrando que ciertas arquitecturas moleculares adquieren valor no por sus partes en abstracto, sino por la forma en que esas partes cooperan una vez enfrentadas a una proteína flexible.

6.3.2 Análisis de derivados de benzimidazol

Los resultados obtenidos en los derivados benzimidazol refuerzan esta misma lógica a una escala menor y más controlada. Dado que la serie es más corta, el análisis se vuelve menos estadístico y más comparativo entre compuestos puntuales. Aun así, la dispersión de XP GScores entre aproximadamente -4.35 y -7.26 demuestra con claridad que el núcleo benzimidazol, por sí solo, no garantiza una afinidad homogénea. Los mejores valores observados para benzimidazol 1 y benzimidazol 5, junto con su superposición más cercana respecto al ligando de referencia en la Figura 40 y su patrón de interacción más favorable en la Figura 41, indican que estos dos compuestos no solo poseen mejor desempeño energético, sino también una lógica espacial de reconocimiento más convincente. En ellos, la ocupación del bolsillo es más definida y la orientación del sistema aromático principal reproduce de forma más fiel la disposición que la proteína parece estabilizar preferentemente. En cambio, benzimidazol 2 y benzimidazol 3, aun cuando logran insertarse en la cavidad, muestran una red de interacciones más limitada y una adaptación menos eficiente, lo que explica coherentemente sus scores más modestos. Los compuestos intermedios, benzimidazol 4 y benzimidazol 6, ilustran una situación mecanísticamente útil: es posible lograr una inserción aceptable y una afinidad moderada sin alcanzar la geometría óptima de los mejores candidatos, lo que sugiere modos de unión permisivos pero no ideales.

6.3.3 Análisis de derivados de top 1 moléculas del *virtual screening*

Los resultados del virtual screening refinado mediante docking flexible mostraron puntajes menos negativos que los obtenidos previamente con AutoDock rígido, lo que no representa una contradicción metodológica, sino una consecuencia esperable de las diferencias entre ambos enfoques. En primer lugar, los valores de XP GScore de Glide y las energías estimadas por AutoDock provienen de funciones de puntuación distintas, por lo que no son directamente comparables en términos absolutos. En segundo lugar, el docking flexible incorpora un tratamiento más realista de la acomodación del ligando dentro del sitio de unión, e incluso de ajustes locales del entorno proteico, mientras que el docking rígido evalúa el reconocimiento molecular bajo una geometría más restringida. En este contexto, es razonable que algunas moléculas presenten puntajes menos favorables en el esquema flexible, ya que este procedimiento no solo recompensa contactos favorables, sino que también penaliza con mayor sensibilidad orientaciones subóptimas, tensiones conformacionales o pérdidas de complementariedad fina durante el ajuste en el bolsillo.

El docking flexible permitió observar un comportamiento relevante desde el punto de vista estructural. Todas las familias evaluadas conservaron su capacidad de ocupar una región común del sitio de unión, compatible con el bolsillo donde se acomoda el ligando de referencia GRL-0617, aunque con diferencias en la orientación, extensión y distribución de sus sustituyentes. Esto sugiere que, más allá de las variaciones en score con la metodología del docking rígido hecho por Autodock, los compuestos seleccionados por virtual screening mantienen un reconocimiento espacial coherente con el sitio alostérico analizado. En particular, familias como benzimidazol, aminopirimidina y benzotiofeno mostraron un desempeño relativamente consistente, mientras que otras, como el pirimido-benzimidazol, evidencian una mayor dependencia del protocolo empleado. Este comportamiento indica que la respuesta de cada andamiaje químico no depende únicamente de su afinidad teórica, sino también de su capacidad de adaptarse geométrica y electrónicamente a un entorno proteico dinámico.

Desde una perspectiva química, estas diferencias son informativas, ya que muestran que la priorización de candidatos puede variar cuando se pasa desde una evaluación inicial más simplificada a un refinamiento estructural más exigente. Por ello, los resultados del docking flexible no deben interpretarse como una invalidación de los obtenidos por AutoDock, sino como una etapa complementaria de discriminación, capaz de revelar qué familias conservan interacciones favorables bajo condiciones más cercanas a un reconocimiento molecular realista. En este sentido, los compuestos que mantuvieron puntajes competitivos y modos de unión coherentes tras el refinamiento flexible adquieren mayor solidez como candidatos para etapas posteriores de análisis.

6.3.4 Limitaciones asociadas al *induced fit docking*

A pesar del mayor nivel de realismo estructural que introduce el *induced fit docking* en comparación con enfoques de *docking* rígido, es fundamental reconocer que los resultados obtenidos siguen estando sujetos a limitaciones inherentes al modelo computacional empleado, las cuales deben ser consideradas al momento de interpretar la afinidad y el modo de unión de los compuestos evaluados.

En primer lugar si bien el *induced fit docking* incorpora flexibilidad local de la proteína, esta flexibilidad es parcial y restringida, limitada a ciertos residuos cercanos al sitio de unión. En consecuencia, no se captura completamente la dinámica global del sistema proteico ni los posibles cambios conformacionales de mayor escala que pueden ocurrir en condiciones fisiológicas. Esto es particularmente relevante en proteínas como PLpro, donde regiones móviles pueden influir significativamente en el reconocimiento molécula. Por tanto, las conformaciones obtenidas deben interpretarse como aproximaciones locales de estados energéticamente favorables, más que como representaciones completas del comportamiento dinámico del sistema.

En segundo lugar, las funciones de *scoring* utilizadas, incluyendo el XP GScore, presentan limitaciones conocidas en la estimación cuantitativa de las energías de unión. Estas funciones están diseñadas principalmente para establecer comparaciones relativas entre compuestos, y no para predecir valores absolutos de afinidad con precisión termodinámica. En este sentido, diferencias pequeñas en los scores no deben interpretarse como variaciones significativas en afinidad real, especialmente cuando se comparan compuestos estructuralmente similares.

Adicionalmente, el modelo no contempla si la interacción ligando-proteína se mantiene en el tiempo, además de cómo se comporta esa molécula en contexto acuoso. Por lo que hace importante y fundamental complementar la técnica con Dinámica molecular.

6.4 Dinámica molecular

6.4.1 Discusión de la Dinámica molecular del sistema APO y del complejo con GRL-0617

El análisis de Dinámica molecular del sistema APO y del complejo con GRL-0617 permite establecer una base estructural y dinámica fundamental para interpretar el mecanismo de inhibición de PLpro. Más que constituir un simple complemento del docking, esta etapa permite distinguir entre una proteína

intrínsecamente flexible y un estado conformacional estabilizado por la unión del ligando, lo que resulta especialmente relevante en una enzima como PLpro, cuya actividad depende no solo de la integridad del centro catalítico, sino también del comportamiento conformacional de regiones móviles como el loop BL2. La simulación del sistema sin ligando y la del complejo con el inhibidor de referencia, por tanto, no deben leerse como análisis independientes, sino como dos estados comparativos del mismo paisaje conformacional.

En el sistema APO, el comportamiento global de la proteína muestra que PLpro conserva estabilidad estructural general, pero no permanece rígida durante la trayectoria. Aunque las réplicas convergen de forma consistente, el RMSD aumenta progresivamente hasta valores cercanos a 4 Å, lo que sugiere que la proteína explora conformaciones alejadas de la estructura inicial antes de alcanzar un régimen relativamente estable. Este punto es importante, porque evita una interpretación simplista del RMSD como si toda desviación implicara inestabilidad. En este caso, más que una desestructuración global, lo que se observa es la relajación de la proteína hacia un conjunto de conformaciones accesibles en ausencia de ligando. En otras palabras, el sistema APO no parece colapsar ni desorganizarse, sino desplazarse dentro de su propio espacio conformacional, lo que es coherente con una proteína que posee elementos móviles funcionalmente relevantes.

El análisis de RMSF refuerza esta interpretación y muestra que esa movilidad no está distribuida de manera homogénea. La mayor fluctuación local se concentra en regiones específicas, destacando particularmente el loop BL2, donde Tyr268 presenta el valor de RMSF más alto del sistema, con un promedio reportado de 5.48 Å. Este resultado es altamente significativo desde el punto de vista funcional, porque identifica al BL2 no como una región periférica pasivamente flexible, sino como el principal foco de plasticidad conformacional de la proteína en ausencia de ligando. En palabras más simples estado APO de PLpro está caracterizado por una movilidad acentuada precisamente en el elemento estructural que regula el acceso al sitio de reconocimiento. Esto apoya fuertemente la idea de que el BL2 constituye una compuerta conformacional dinámica cuya modulación puede ser explotada farmacológicamente.

Las distancias interresiduo analizadas en el sistema APO ayudan a precisar aún más esta lectura. La distancia catalítica entre Cys111 y His272 se mantiene esencialmente constante entre réplicas, lo que indica que el centro catalítico conserva su organización geométrica a lo largo de la simulación. Este hallazgo es relevante porque demuestra que, aun cuando la proteína explora estados conformacionales diversos, la maquinaria catalítica no se desorganiza. Por tanto, la flexibilidad observada no corresponde a una pérdida general de integridad estructural, sino a una movilidad localizada en regiones regulatorias del reconocimiento molécula. Esta distinción es crucial: la proteína puede ser dinámicamente flexible y, al mismo tiempo, conservar la arquitectura de su núcleo funcional.

En contraste, las distancias entre Tyr268 y Arg166, y entre Tyr268 y Pro299, muestran una variabilidad mucho mayor, con cambios abruptos a lo largo del tiempo. Este comportamiento es coherente con movimientos de apertura y cierre del loop BL2 respecto del resto de la proteína. La amplitud de estas variaciones sugiere que el BL2 no oscila alrededor de una única conformación promedio, sino que transita entre múltiples estados estructuralmente diferenciados. En términos mecanísticos, esto refuerza la noción de que el estado APO de PLpro está asociado a una alta plasticidad del acceso al sitio de unión, condición que probablemente facilita el reconocimiento de sustratos naturales como ubiquitina e ISG15. Así, el sistema sin ligando puede describirse como conformacionalmente permisivo: mantiene la tríada catalítica intacta, pero conserva un entorno de entrada dinámico y adaptable.

El análisis angular de Tyr268 ofrece una dimensión adicional a esta discusión. En el sistema APO, la distribución de los diedros χ_1 y χ_2 muestra poblaciones predominantes bien definidas, con una fuerte representación de ciertos estados conformacionales y la coexistencia de otras orientaciones secundarias. Desde una perspectiva estructural, esto indica que Tyr268 no rota aleatoriamente, sino que ocupa un conjunto restringido pero múltiple de conformaciones preferentes. La existencia de estos estados discretos refuerza la idea de que el BL2 alterna entre geometrías funcionalmente diferenciadas y que Tyr268 puede actuar como un marcador conformacional útil para cuantificar dichos estados. Este punto es especialmente valioso para la tesis, porque transforma una observación geométrica en una posible métrica mecanística para comparar ligandos posteriores.

Cuando se compara este comportamiento con el del complejo proteína–GRL-0617, emerge con claridad el efecto conformacional del inhibidor. La simulación del complejo muestra una estabilización progresiva del RMSD de la proteína, alcanzando valores prácticamente constantes tras los primeros nanosegundos. A diferencia del sistema APO, donde la proteína continúa desplazándose hacia conformaciones progresivamente alejadas de la estructura inicial antes de estabilizarse, en presencia de GRL-0617 la fluctuación estructural global es menor. Esto sugiere que la unión del ligando restringe parte del espacio

conformacional accesible de PLpro y favorece un estado estructural más definido. En consecuencia, GRL-0617 no solo se une a la proteína, sino que parece actuar como un estabilizador conformacional.

El RMSD del propio ligando aporta un elemento importante a esta interpretación. Aunque GRL-0617 experimenta un ajuste inicial dentro del bolsillo, posteriormente mantiene un RMSD cercano a 3.5–4 Å durante la simulación, lo que indica permanencia en el sitio de unión sin expulsión ni reorientaciones drásticas incompatibles con una unión estable. Este comportamiento es particularmente relevante porque muestra que la estabilización de la proteína no depende de un ligando absolutamente inmóvil, sino de un ligando que, aun conservando cierta libertad local, permanece anclado dentro de una región funcionalmente competente. En otras palabras, la estabilidad del complejo no exige rigidez absoluta del ligando, sino persistencia espacial dentro del sitio alostérico compatible con el mantenimiento del estado conformacional inhibido.

El efecto más importante del GRL-0617, sin embargo, no se aprecia solo en el RMSD, sino en el patrón de flexibilidad local. La comparación del RMSF entre el sistema APO y el complejo muestra una disminución clara de la movilidad en la región del BL2, con una rigidización particularmente evidente alrededor de Tyr268. Este hallazgo tiene gran peso mecanístico porque reproduce precisamente el fenómeno esperado para un inhibidor no covalente de referencia: no altera de forma relevante la organización de la tríada catalítica, pero sí modula el paisaje conformacional del loop que regula el acceso al sitio. En este sentido, la simulación no solo valida el protocolo computacional, sino que también reproduce el modelo de inhibición descrito en la literatura para GRL-0617, donde la estabilización de una conformación cerrada del BL2 constituye el evento funcional clave.

Las distancias interresiduo apoyan directamente esta interpretación. Tal como observaste, la distancia catalítica entre Cys111 y His272 no muestra un efecto diferencial importante al comparar APO y GRL-0617, lo que sugiere que la inhibición no depende de una perturbación directa del núcleo catalítico. En cambio, las distancias que involucran a Tyr268 sí revelan un cambio mucho más claro. En presencia de GRL-0617, las separaciones entre Tyr268 y los otros puntos de referencia del BL2 tienden a valores más cortos y menos dispersos que en el sistema APO, lo que es consistente con una conformación más cerrada y menos fluctuante del loop. Esta observación tiene implicancias funcionales directas: si el BL2 permanece más próximo y restringido, el canal de acceso para sustratos voluminosos como ubiquitina o ISG15 se vuelve menos permisivo, lo que ofrece una explicación estructural plausible para la inhibición observada.

Los diedros de Tyr268 completan esta lectura al mostrar un cambio conformacional mucho más definido en el complejo con GRL-0617 que en el sistema APO. Mientras que en ausencia de ligando Tyr268 explora varias orientaciones preferentes, en presencia del inhibidor las distribuciones angulares se concentran de manera más marcada en regiones específicas del espacio conformacional. Esto sugiere que la unión del ligando no solo reduce la amplitud del movimiento del loop, sino que selecciona estados rotacionales concretos de la cadena lateral de Tyr268. Desde el punto de vista conceptual, este es un resultado especialmente potente, porque indica que los diedros de Tyr268 pueden funcionar como cuantificadores estructurales del estado inhibido. No se trata solo de un residuo flexible, sino de un sensor conformacional cuya orientación distingue entre un entorno abierto y uno rigidizado por ligando.

En conjunto, la comparación entre APO y GRL-0617 permite extraer una conclusión mecanística robusta. El sistema APO describe una PLpro conformacionalmente competente y altamente móvil en la región del BL2, donde el acceso al sitio de reconocimiento permanece dinámico y probablemente funcional. En contraste, la unión de GRL-0617 no altera de forma importante la geometría del centro catalítico, pero sí restringe de manera clara la movilidad del BL2, disminuye la fluctuación global de la proteína, acorta y estabiliza distancias asociadas al movimiento del loop y concentra la distribución angular de Tyr268 en estados más definidos. Este conjunto de evidencias apoya un modelo de inhibición basado en estabilización conformacional más que en perturbación catalítica directa. Dicho de otro modo, GRL-0617 parece inhibir a PLpro no porque rompa su arquitectura enzimática, sino porque la fija en un estado menos permisivo para el reconocimiento de sustratos.

6.4.2 Discusión de la Dinámica molecular de la familia pirimido-benzimidazol

Si se comienza por el RMSD de la proteína, se aprecia que esta familia no genera una respuesta uniforme sobre la estabilidad global de PLpro. En términos generales, los sustituyentes 4-NMe₂ y 4-OH-3-OMe son los que tienden a mostrar trayectorias más cercanas al complejo con GRL-0617 en varias subseries, lo que podría sugerir una modulación global más compatible con el estado de referencia. En particular, 4-OH-3-OMe mantiene un comportamiento favorable en las series benceno, 2-naftilo, antraceno y pireno,

aunque en 1-naftilo presenta una desviación transitoria más marcada. En contraste, el sustituyente 2-OH es el que más frecuentemente se asocia con aumentos notorios del RMSD proteico, especialmente en 1-naftilo, donde la trayectoria llega a aproximarse e incluso a superar al sistema APO; también en benceno y pireno este mismo fragmento induce una desviación mayor, mientras que en 2-naftilo su comportamiento es comparativamente más estable. El derivado 4-NO₂ ocupa una posición intermedia: en varias series se aproxima al patrón del inhibidor de referencia, pero en 1-naftilo vuelve a aparecer una desviación más evidente. En este primer nivel de análisis, por tanto, no parecería correcto proponer a 2-OH como el mejor comportamiento global, porque su efecto sobre la proteína es más inestable y dependiente del núcleo aromático.

El RMSD del ligando introduce un criterio más exigente, porque aquí ya no se evalúa solo la respuesta de la proteína, sino la capacidad del compuesto de conservar su ubicación inicial en el bolsillo. En la serie benceno, el patrón dominante es desfavorable: la mayoría de los ligandos abandona claramente la pose inicial y varios parecen estabilizarse en otras regiones de la proteína. En este grupo, 4-iPr y 2-OH son los menos desfavorables, mientras que 4-OH-3-OMe, 4-CF₃ y especialmente 4-NMe₂ muestran desplazamientos más notorios. En 1-naftilo, la situación mejora parcialmente: varios sistemas todavía abandonan el sitio, pero 2-OH es el que más se aproxima al comportamiento del ligando de referencia, mientras que cinamilo y 4-NO₂ tienden a estabilizarse en mesetas alrededor de 15–20 Å y 4-OH-3-OMe exhibe un reacomodo tardío importante. En 2-naftilo, en cambio, aparece el patrón más convincente de toda la familia: 4-OH-3-OMe mantiene un RMSD muy próximo al de GRL-0617 durante prácticamente toda la trayectoria, mientras que 4-NMe₂ y 4-iPr quedan en una zona intermedia, y cinamilo, 4-NH₂, 2-OH y 4-CF₃ muestran desplazamientos claramente mayores. En antraceno, los mejores perfiles corresponden a 4-iPr y 2-OH, con trayectorias cercanas a 4–6 Å, mientras que 4-OH-3-OMe, 4-NH₂ y 4-CF₃ permanecen en valores moderados, y cinamilo y 4-NO₂ son los más inestables. Finalmente, en pireno predomina otra vez un comportamiento irregular, aunque 4-OH-3-OMe es el sistema más cercano al control dentro de esa subserie, mientras que 4-CF₃ y 4-iPr muestran perfiles intermedios y cinamilo y 4-NMe₂ vuelven a ser más erráticos. En conjunto, el RMSD del ligando sugiere que los mejores desempeños no se concentran en un único sustituyente para todas las series, pero sí coloca a 2-naftilo–4-OH-3-OMe como el caso más robusto cuando se prioriza permanencia espacial dentro del bolsillo.

Cuando se pasa al RMSF, la lectura cambia desde la estabilidad espacial hacia la capacidad de rigidizar el loop BL2. Aquí se vuelve evidente que no siempre el compuesto que mejor conserva la pose es el que más rigidiza esta región. En la serie benceno, varios ligandos muestran un efecto solo moderado, y los más claros en reducción de la movilidad son 4-CF₃ y, en menor medida, 2-OH, mientras que 4-OH-3-OMe, 4-NMe₂ y 4-iPr ocupan un comportamiento intermedio. En 1-naftilo, el derivado que más claramente destaca es 2-OH, que presenta una de las reducciones más notables de la flexibilidad del BL2; 4-OH-3-OMe mantiene una estabilización moderada, mientras que 4-NO₂ se aproxima más al patrón del APO y parecería menos apto para restringir el loop. En la serie 2-naftilo, la modulación es más clara y aquí sí sobresalen 4-OH-3-OMe y 4-iPr, que muestran la mayor rigidización del centro del BL2; por el contrario, 4-CF₃ y 4-NO₂ conservan perfiles más altos y próximos al estado apo. En antraceno, la mayoría de los compuestos tiende a rigidizar con bastante claridad, salvo 4-NO₂, que incluso incrementa la flexibilidad por encima del APO. En pireno, en cambio, el efecto general es más débil y el único sustituyente que realmente sobresale es 4-NMe₂. Por tanto, si el criterio es estrictamente la reducción del RMSF en BL2, los compuestos que más destacan serían 1-naftilo–2-OH, 2-naftilo–4-OH-3-OMe, 2-naftilo–4-iPr, varios antracénicos y pireno–4-NMe₂.

Las distancias permiten afinar todavía más la discusión. La distancia catalítica Cys111–His272 se mantiene relativamente constante en toda la familia y, por lo tanto, no parece ser el descriptor que mejor discrimina entre complejos favorables y desfavorables; más bien sugiere que la geometría del núcleo catalítico se conserva aun cuando el resto del entorno dinámico varíe. La información útil está, en cambio, en las distancias Tyr268–Arg166 y Tyr268–Pro299, que sí capturan mejor el comportamiento del BL2. Para Tyr268–Arg166, en la serie benceno destacan 4-CF₃ y 2-OH por mostrar valores menores y más cercanos al control, mientras que 4-OH-3-OMe, 4-NMe₂, 4-NH₂ y 4-iPr muestran trayectorias más erráticas. En 1-naftilo, los perfiles más cercanos a GRL-0617 corresponden otra vez a 4-CF₃ y 2-OH, mientras que 4-NMe₂ es claramente el más desfavorable. En 2-naftilo, sin embargo, el panorama es más equilibrado: 4-NMe₂, 4-iPr y 2-OH muestran distancias menores y relativamente estables, mientras que 4-OH-3-OMe mantiene

un perfil constante, aunque ligeramente por encima del control. Esto es importante porque sugiere que 4-OH-3-OMe no domina esta distancia en forma extrema, pero tampoco incurre en una trayectoria errática.

La distancia Tyr268–Pro299 resulta incluso más reveladora. En la serie benceno, cinamilo se acerca inicialmente a GRL-0617 pero luego se abre, 4-OH-3-OMe y 4-NMe₂ permanecen más cerca del APO, mientras que 4-CF₃, 2-OH y 4-NO₂ muestran intervalos de cierre más marcados, aunque con trayectorias variables. En 1-naftilo, el patrón vuelve a favorecer a 2-OH, que permanece próximo a GRL-0617 durante gran parte de la trayectoria, mientras que 4-OH-3-OMe empieza con distancias menores pero luego se abre progresivamente; 4-iPr también mantiene valores bastante cercanos al control. En 2-naftilo, en cambio, el resultado más fuerte corresponde a 4-OH-3-OMe, cuyo perfil se describe como prácticamente superpuesto al del inhibidor de referencia; 4-NMe₂ se mantiene incluso por debajo de GRL-0617 gran parte del tiempo, mientras que 4-CF₃, 4-NH₂ y 4-NO₂ son más fluctuantes o más altos. Así, si Tyr268–Pro299 se interpreta como el descriptor más directo de un estado cerrado y sostenido del BL2, el complejo 2-naftilo–4-OH-3-OMe sería el que mejor reproduce el patrón del control.

Para los diedros χ_1 y χ_2 de Tyr268 ayudan a distinguir si esa modulación se acompaña o no de una firma conformacional similar a la observada con GRL-0617. En la serie benceno, el patrón más parecido al control para χ_1 corresponde a 4-NO₂ y 2-OH, mientras que 4-OH-3-OMe se concentra principalmente alrededor de -45° ; para χ_2 , en cambio, 4-OH-3-OMe, 4-iPr, 2-OH y 4-NH₂ muestran el patrón vertical alrededor de $\pm 90^\circ$ compatible con GRL-0617. En 1-naftilo, cinamilo y 4-OH-3-OMe son los que más se aproximan al control en χ_1 , pero el resto de la subserie se dispersa más. En 2-naftilo es donde aparece el resultado más convincente: 4-OH-3-OMe reproduce simultáneamente la distribución dominante de χ_1 cercana a 180° y la de χ_2 alrededor de $\pm 90^\circ$, es decir, la firma conformacional más semejante al ligando de referencia dentro de toda la familia. Desde este punto de vista, la serie 2-naftilo no solo aporta un ligando estable, sino también el residuo Tyr268 más claramente “GRL-like”.

Visto así cada gráfico privilegia sistemas distintos. El RMSD de proteína favorece sobre todo a 4-NMe₂, 4-OH-3-OMe y en algunos casos 4-NO₂; el RMSD del ligando destaca con fuerza a 1-naftilo–2-OH, 2-naftilo–4-OH-3-OMe, antraceno–4-iPr, antraceno–2-OH y, más parcialmente, pireno–4-OH-3-OMe; el RMSF del BL2 favorece a 1-naftilo–2-OH, 2-naftilo–4-OH-3-OMe, 2-naftilo–4-iPr, varios antracénicos y pireno–4-NMe₂; las distancias Tyr268–Arg166 privilegian sobre todo a 4-CF₃ y 2-OH en benceno/1-naftilo, y a 4-NMe₂, 4-iPr y 2-OH en 2-naftilo; mientras que Tyr268–Pro299 y los diedros favorecen de manera más clara a 2-naftilo–4-OH-3-OMe. Esto muestra que la familia no responde a una regla simple de “mejor sustituyente universal”, sino a una interacción entre núcleo aromático y sustituyente, donde un mismo grupo puede ser favorable en una métrica y desfavorable en otra.

Al integrar todos los análisis el panorama general sugiere dos comportamientos principales. Por un lado, existen ligandos que pueden reducir la flexibilidad del BL2 o acortar alguna distancia relevante, pero lo hacen a costa de una peor permanencia espacial o de una desviación mayor de la proteína; ahí entrarían varios sistemas con 2-OH, en especial 1-naftilo–2-OH. Por otro lado, hay complejos que no necesariamente son los mejores en cada métrica aislada, pero sí muestran la mayor coherencia global entre estabilidad del ligando, modulación del BL2 y firma conformacional de Tyr268. En esa segunda categoría, el compuesto que más claramente sobresale es 2-naftilo–4-OH-3-OMe: no solo mantiene un RMSD del ligando prácticamente superpuesto al del control, sino que además rigidiza el BL2, reproduce muy bien la distancia Tyr268–Pro299 y muestra una distribución de χ_1/χ_2 muy semejante a la de GRL-0617. Bajo una lectura integrada, podría proponerse como el candidato más equilibrado y mecanísticamente más convincente de la familia pirimido-benzimidazol.

6.4.3 Discusión de la Dinámica molecular del Top 2 de las mejores moléculas derivadas de chalconas

La selección de los dos compuestos mejor posicionados de cada familia permite comparar, dentro de un mismo marco dinámico, si los candidatos mejor rankeados convergen realmente hacia un comportamiento semejante al del inhibidor de referencia o si, por el contrario, su buen puntaje de docking responde únicamente a una complementariedad inicial que no se sostiene a lo largo de la simulación. En este análisis no se incluyeron las dihidrobenzodiazepinas ni los dihidropirazoles. En el caso de los dihidropirazoles, esta exclusión puede justificarse con mayor solidez, ya que las pirazolininas han sido descritas como sistemas susceptibles a aromatización oxidativa hacia pirazoles aromáticos incluso en presencia de atmósfera ambiente, lo que hace razonable priorizar la forma aromatizada como especie más pertinente para el análisis estructural. Para las dihidrobenzodiazepinas, en cambio, se adoptó una

decisión más conservadora de priorización analítica, favoreciendo las formas aromáticas o más oxidadas, al considerarlas potencialmente más persistentes dentro del marco comparativo del presente estudio²³. Si se observa primero el RMSD de la proteína, el resultado más general es que este descriptor discrimina poco entre familias, ya que todos los complejos evaluados mantienen valores comparables o incluso inferiores a los observados para GRL-0617, predominantemente entre 2 y 3 Å durante la mayor parte de la trayectoria. Esto indica que, en términos globales, ninguna de las series induce una desorganización estructural amplia de PLpro. Sin embargo, precisamente por esa homogeneidad relativa, el RMSD proteico debe interpretarse más como un criterio de compatibilidad global que como un marcador directo de inhibición. Bajo esta lectura, las aminopirimidinas, benzodiazepinas, pirimidina-aminobenzimidazol, pirazoles y chalconas comparten una estabilidad general aceptable del esqueleto proteico, por lo que la diferencia real entre familias debe buscarse en descriptores más finos, asociados a la permanencia del ligando y a la modulación del BL2.

El RMSD del ligando sí introduce una separación mucho más clara entre familias. En las aminopirimidinas, ninguno de los dos compuestos parece sostener una permanencia convincente durante toda la simulación. El derivado 2-naftilo-4-CF₃-fenilo se mantiene relativamente bien solo durante los primeros ~400 ns y luego presenta un aumento abrupto, compatible con migración fuera de la posición inicial, mientras que antraceno-4-CF₃-fenilo muestra una variabilidad amplia desde etapas tempranas. Esto sugiere que, aunque el acoplamiento inicial pudiera haber sido favorable, la retención dinámica del ligando no se consolida en esta familia. En las benzodiazepinas, en cambio, ambos complejos muestran un perfil mucho más robusto: el derivado pireno-2-OH-fenilo se mantiene prácticamente superpuesto al comportamiento de GRL-0617, con valores del RMSD del ligando cercanos a 4–5 Å, mientras que antraceno-4-CF₃-fenilo también conserva una estabilidad local razonable. Aquí, a diferencia de las aminopirimidinas, sí se observa una permanencia espacial compatible con un modo de unión sostenido. En las chalconas, el panorama es notoriamente distinto: 1-naftilo-4-CF₃-fenilo sufre un desplazamiento abrupto y termina estabilizándose alrededor de 30 Å, y pireno-4-NO₂-fenilo permanece la mayor parte del tiempo en torno a 15–18 Å. Este comportamiento sugiere que, bajo un esquema no covalente convencional, la familia no logra conservar una pose estable comparable al inhibidor de referencia. En pirimidina-aminobenzimidazol, el derivado 1-naftilo-4-CF₃-fenilo permanece estable solo durante aproximadamente 200 ns antes de desplazarse hasta ~27–30 Å, mientras que pireno-4-CF₃-fenilo muestra una trayectoria muy variable con reacomodos bruscos. Se trata, por tanto, de una familia con efecto parcial, pero sin una permanencia tan limpia como las benzodiazepinas. En los pirazoles, el contraste interno es fuerte: 1-naftilo-cinamilo se comporta de forma claramente desfavorable, con valores superiores a 40 Å y gran variabilidad, mientras que pireno-3-OMe-4-OH-fenilo mantiene valores cercanos a 8–10 Å durante la trayectoria, lo que lo convierte en el mejor representante de esta familia, aunque todavía más dinámico que GRL-0617.

Ese mal desempeño de las chalconas en RMSD del ligando abre, de hecho, una lectura mecanística distinta. Como su motivo α,β-insaturado constituye un aceptor de Michael, sería razonable plantear que su actividad potencial frente a una proteasa cisteína como PLpro podría no expresarse adecuadamente en un esquema puramente no covalente, sino eventualmente mediante interacción covalente con residuos nucleofílicos, particularmente cisteína. La literatura sobre chalconas describe justamente su carbonilo α,β-insaturado como un grupo capaz de reaccionar con tioles de cisteína, y en el contexto de proteasas cisteína de coronavirus se ha planteado que el scaffold chalcona podría aportar inhibición irreversible por interacción tipo Michael. Por ello, el pobre desempeño dinámico observado aquí no necesariamente invalida a la familia, pero sí sugiere que el modelo usado probablemente no captura su mecanismo más plausible de acción.

Al pasar al RMSF, la comparación ya no se centra en si el ligando permanece en el sitio, sino en si logra reducir la movilidad del BL2. A partir de este análisis se observa un patrón definido. En las aminopirimidinas, el derivado 2-naftilo-4-CF₃-fenilo mantiene un perfil prácticamente similar al APO en la región BL2, mientras que antraceno-4-CF₃-fenilo solo produce una reducción leve. Esto sugiere que, aun cuando la proteína globalmente permanezca estable, la familia aminopirimidina no ejerce una rigidización marcada del loop funcional. En las benzodiazepinas, en cambio, ambos derivados —antraceno-4-CF₃-fenilo y pireno-2-OH-fenilo— muestran valores menores de RMSF en BL2 y cercanos a los del sistema con GRL-0617. Esta coincidencia entre buena permanencia del ligando y reducción clara de la movilidad del BL2 vuelve a la familia benzodiazepina particularmente convincente. En las chalconas, los dos compuestos seleccionados mantienen perfiles muy parecidos al APO, con diferencias mínimas en la amplitud de las fluctuaciones; es decir, no se observa un efecto real de rigidización del BL2. En pirimidina-

aminobenzimidazol, el efecto es intermedio: 1-naftilo-4-CF₃-fenilo muestra una reducción parcial del RMSF, mientras que pireno-4-CF₃-fenilo presenta valores menores de RMSF y, por lo tanto, parecería más prometedor desde esta métrica. Finalmente, en los pirazoles, 1-naftilo-cinamilo se comporta casi como APO, mientras que pireno-3-OMe-4-OH-fenilo sí reduce la movilidad del BL2 de manera apreciable. En consecuencia, si se jerarquiza solo por capacidad de rigidizar el BL2, las benzodiazepinas destacan con claridad, seguidas por el pirazol pireno-3-OMe-4-OH-fenilo y, de forma más parcial, por pireno-4-CF₃-fenilo en la serie pirimidina-aminobenzimidazol.

Las distancias confirman esa separación entre familias. La distancia catalítica Cys111–His272 se mantiene relativamente constante en casi todos los sistemas, lo que indica que el núcleo catalítico no es el descriptor más útil para discriminar el efecto de estos ligandos. Las diferencias relevantes emergen en las distancias que involucran a Tyr268, es decir, aquellas asociadas al movimiento del BL2. Las diferencias observadas son consistentes con una lectura diferenciada por familias. En las aminopirimidinas, las trayectorias de Tyr268–Arg166 y Tyr268–Pro299 son más erráticas y no se alinean de forma convincente con GRL-0617; incluso el antraceno-4-CF₃-fenilo podría favorecer un aumento de la distancia respecto del sistema de referencia, lo que sería más compatible con un efecto contrario al cierre del BL2 que con una estabilización del estado inhibido. En las benzodiazepinas, en cambio, ambas distancias son claramente más favorables: los perfiles son similares a GRL-0617 y, en el caso de antraceno-4-CF₃-fenilo, incluso hay una superposición casi completa o valores menores para una de estas distancias. Esa observación es importante, porque sugiere que esta familia no solo rigidiza BL2, sino que también lo mantiene en una geometría espacial muy próxima a la del control. En las chalconas, por el contrario, no se aprecia un efecto claro: las trayectorias permanecen más próximas al APO y no sugieren un cierre sostenido del loop. En pirimidina-aminobenzimidazol, sí se advierte una modulación parcial, aunque menos limpia y menos consistente que en benzodiazepinas. En los pirazoles, el sistema pireno-3-OMe-4-OH-fenilo parece ser el único que se aleja de la tendencia más errática y podría inducir un efecto parcial de aproximación al patrón del control. Bajo esta métrica, nuevamente las benzodiazepinas serían la familia más fuerte, mientras que las chalconas quedarían claramente rezagadas.

El análisis de los diedros χ_1 y χ_2 de Tyr268 añade un último nivel de resolución. Para el análisis Top 2, en general, la conformación dominante de χ_1 se concentra alrededor de 180°, similar a los sistemas de referencia, y que las diferencias más relevantes aparecen en χ_2 , donde algunas familias muestran distribuciones más amplias y poblaciones adicionales. En las aminopirimidinas, χ_1 se mantiene cercano al patrón de referencia, pero χ_2 revela una mayor flexibilidad rotacional, lo que sugiere que la orientación fina del anillo fenólico de Tyr268 no queda tan restringida como en GRL-0617. En las benzodiazepinas, especialmente el derivado con pireno-2-OH-fenilo, reproducen de manera bastante convincente el patrón diedral del control, reforzando la idea de que esta familia sí induce un estado conformacional similar al del inhibidor de referencia. En las chalconas, en cambio, no se aprecia una reorganización diedral particularmente coherente con un estado inhibido sostenido. Para pirimidina-aminobenzimidazol y pirazol, la lectura es más parcial: hay señales de efecto, pero con mayor dispersión y menor nitidez que en benzodiazepinas. En este punto, es importante destacar que los diedros no contradicen la lectura de RMSF y distancias, sino que la refinan. Cuando existe evidencia de rigidización del loop BL2, las distribuciones diedrales tienden a acompañar esta tendencia; en cambio, en los sistemas con efectos parciales, la firma angular se vuelve más dispersa.

La integración de los descriptores dinámicos permite establecer una jerarquía general entre las familias evaluadas. Las aminopirimidinas muestran una estabilidad proteica global aceptable, pero fallan en permanencia del ligando y no inducen una rigidización clara ni una reorganización espacial convincente del BL2; por ello, su efecto global parecería débil o incluso parcialmente contraproducente. Las benzodiazepinas emergen como la familia más sólida, porque combinan permanencia espacial del ligando, reducción clara del RMSF en BL2, distancias muy próximas a GRL-0617 y una firma diedral compatible con el control; dentro de esta familia, pireno-2-OH-fenilo parece el más consistente en RMSD del ligando, mientras que antraceno-4-CF₃-fenilo sería especialmente fuerte en las distancias del BL2. Las chalconas muestran el peor comportamiento conjunto en el marco no covalente: sus ligandos migran, no rigidizan BL2 y no reproducen el patrón espacial del control, por lo que, las chalconas podrían discutirse más apropiadamente como posibles estructuras reactivas con potencial covalente, más que como inhibidores no covalentes clásicos. La familia pirimidina-aminobenzimidazol presenta un comportamiento intermedio: algunos indicios de rigidización y modulación del BL2 están presentes, pero la permanencia del ligando sigue siendo insuficiente para colocarla al nivel de las benzodiazepinas. Finalmente, los pirazoles muestran una dicotomía interna marcada: 1-naftilo-cinamilo tiene un

comportamiento dinámico desfavorable, mientras que pireno-3-OMe-4-OH-fenilo conserva una señal parcial de inhibición y sería el único representante de esa familia que destaca.

6.4.4 Discusión de la Dinámica molecular derivados de benzimidazoles

El comportamiento del RMSD de la proteína sugiere que varios complejos conservan una estabilidad global aceptable a lo largo de la trayectoria, particularmente los sistemas con benzimidazol 1, 3 y 6, los cuales tienden a estabilizarse tras la etapa inicial de relajación. Sin embargo, esta métrica, por sí sola, no permite discriminar con suficiente precisión cuáles de estos derivados inducen un estado funcionalmente más cercano al del inhibidor de referencia, ya que una proteína globalmente estable no implica necesariamente una modulación efectiva de las regiones estructuralmente relevantes para la inhibición. La diferencia entre los compuestos se vuelve más evidente al considerar el RMSD del ligando. Bajo este descriptor, benzimidazol 3 y benzimidazol 5 emergen como los casos más favorables, ya que ambos muestran una permanencia relativamente estable dentro de la región de unión tras los primeros nanosegundos de simulación, mientras que los demás derivados presentan desplazamientos más marcados o trayectorias menos consistentes. Este resultado sugiere que, dentro de la serie, solo una fracción de los compuestos logra sostener una interacción espacial suficientemente persistente como para justificar una interpretación mecanística más robusta. En ese sentido, benzimidazol 3 y benzimidazol 5 no solo destacan por estabilidad local, sino porque constituyen los dos sistemas en los que la pose inicial parece mantenerse de manera más razonable durante la trayectoria.

El análisis del RMSF aporta un criterio adicional de diferenciación y permite evaluar de forma más específica el efecto de cada ligando sobre la movilidad del loop BL2. En este punto, benzimidazol 3 adquiere particular relevancia, ya que sería el derivado que muestra la reducción más marcada de la fluctuación en esta región. Esto resulta importante porque sugiere que su efecto no se limita a permanecer dentro del bolsillo, sino que además podría restringir con mayor eficacia la movilidad local del entorno estructural asociado a Tyr268. Aunque en los resultados generales también se describe una disminución de la flexibilidad para benzimidazol 2, la lectura conjunta de los gráficos indicaría que benzimidazol 3 es el que logra el efecto más claro y consistente sobre el BL2, lo que le confiere mayor peso interpretativo dentro de la serie.

Las distancias interresiduo parecen reforzar esta interpretación. Aunque el efecto sobre la distancia catalítica no muestra una tendencia particularmente marcada, en benzimidazol 3 se observaría un acortamiento más evidente de las distancias asociadas al loop BL2, especialmente aquellas que involucran Tyr268 con Pro299 y Arg166. Esto podría interpretarse como una aproximación del loop hacia una conformación más contenida o menos abierta, en concordancia con la reducción del RMSF observada para este mismo compuesto. Así, benzimidazol 3 no solo tendería a rigidizar localmente la región, sino que además parecería desplazarla hacia una geometría más próxima a un estado conformacional restringido. Desde esta perspectiva, su comportamiento sería más convincente como modulador estructural del BL2 que el de otros derivados de la misma serie.

El análisis de los diedros de Tyr268 introduce, sin embargo, un matiz importante. Según la lectura comparativa de las distribuciones angulares, benzimidazol 5, y en menor medida benzimidazol 6, mostrarían una mayor semejanza con el ligando de referencia en términos de la orientación conformacional de Tyr268. Esto sugiere que, aunque benzimidazol 3 destaque por rigidización del BL2 y por el comportamiento de las distancias, benzimidazol 5 podría ser el sistema que mejor reproduce la firma angular característica del estado inducido por GRL-0617. En otras palabras, benzimidazol 5 no parecería ser el modulador más fuerte del loop, pero sí podría constituir el derivado que más se aproxima al patrón conformacional fino del inhibidor de referencia.

La serie benzimidazol apunta al benzimidazol 3, este destacaría por su capacidad de reducir la movilidad del BL2 y por inducir un acortamiento más claro de las distancias asociadas a esta región, lo que lo posicionaría como el mejor modulador estructural del loop. Por otra, benzimidazol 5 sobresaldría por la estabilidad del ligando y por una firma diedral más cercana a la del ligando de referencia. Bajo esta lectura integrada, ambos compuestos serían los más relevantes de la serie, aunque representando comportamientos distintos: benzimidazol 3 como el sistema con mayor impacto sobre la dinámica del BL2 y benzimidazol 5 como el sistema con mayor estabilidad en el sitio alostérico.

6.4.5 Discusión de la Dinámica molécula top 1 del *virtual screening*

Los análisis dinámicos sugieren que las benzamidas y las pirimidina aminobenzimidazol presentan el comportamiento más favorable entre las familias evaluadas. Ambas mantienen una estabilidad proteica comparable al complejo de referencia, presentan menores alteraciones en la región BL2 y conservan distancias Tyr268–Arg166 y Tyr268–Pro299 más cercanas al patrón observado para GRL-0617. Este comportamiento es relevante, ya que el loop BL2 participa directamente en la conformación del sitio alostérico de PLpro y su estabilización puede asociarse con una mejor permanencia del ligando en el bolsillo de unión.

El benzimidazol, en cambio, mostró un comportamiento intermedio. Aunque no produjo una desestabilización global evidente de la proteína, el RMSD del ligando aumentó progresivamente y las distancias asociadas a Tyr268 mostraron cambios hacia el final de la trayectoria. Por tanto, esta familia puede considerarse parcialmente estable, pero menos robusta que benzamida y pirimidina aminobenzimidazol.

Las familias benzotiofeno, pirimido-benzimidazol y aminopirimidina mostraron un comportamiento menos favorable, principalmente por los valores elevados de RMSD del ligando y por transiciones marcadas en las distancias asociadas al loop BL2. Esto sugiere una menor conservación de la pose inicial y una mayor reorganización local del sitio de unión durante la simulación.

La benzodiazepina mostró un comportamiento particular, caracterizado por distancias Tyr268–Arg166 y Tyr268–Pro299 persistentemente elevadas, junto con una distribución angular de Tyr268 distinta al patrón observado para GRL-0617. Este perfil sugiere que el complejo podría favorecer o estabilizar una conformación más abierta del loop BL2. Sin embargo, esta interpretación debe considerarse como una hipótesis conformacional, ya que los análisis realizados no permiten asignar directamente un efecto funcional activador o inhibidor.

6.4.6 Integración de moléculas favorables

Las moléculas resultantes de este análisis se muestran en la Figura 187 . Donde se destaca por una parte la presencia de anillos aromáticos para cada una, esto abre la posibilidad de integrar nueva familia de moléculas fuera de las conocidas benzamidas con un carbono quiral. Esto abriría la oportunidad a una nueva familia molecular con propiedades antivirales.

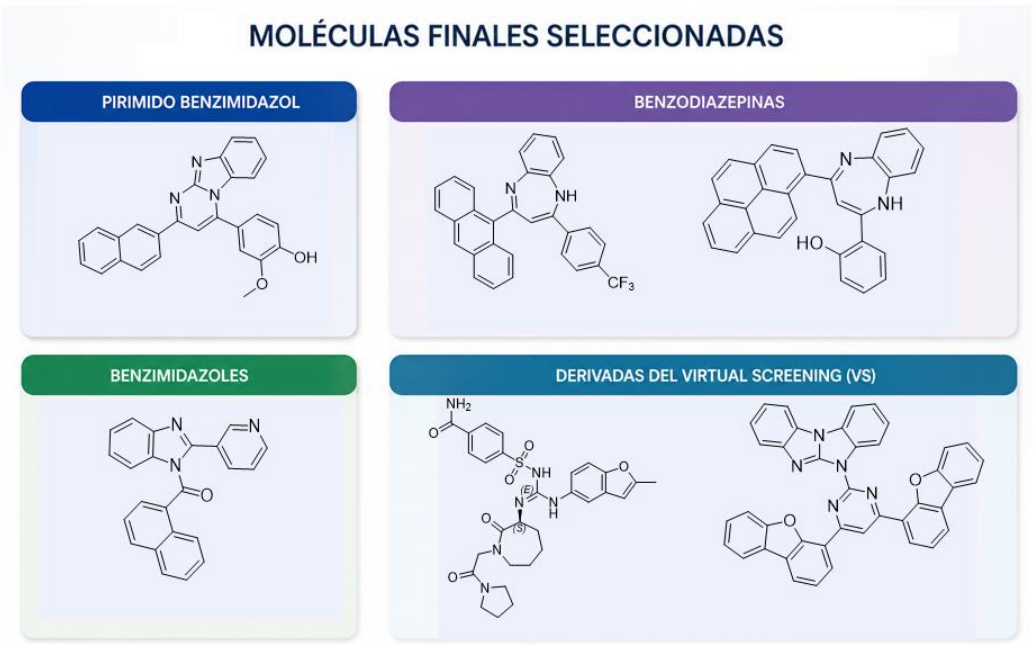


Figura 187 Moléculas finales para cada familia analizada y las obtenidas por el Virtual screening.

6.5 Farmacocinética

6.5.1 Farmacocinética de derivados de chalconas

6.5.1.1 Absorción y distribución

En términos generales, los compuestos evaluados presentaron una predicción favorable de absorción intestinal, lo que sugiere que la mayoría de las estructuras conservan una probabilidad adecuada de ser absorbidas a nivel gastrointestinal. Sin embargo, al examinar descriptores más específicos, como la permeabilidad en Caco-2 y el comportamiento frente a P-gp, emergen diferencias relevantes entre familias.

En particular, algunas familias mostraron una mayor tendencia a comportarse como inhibidores de P-gp, lo que podría modificar su absorción y distribución, aunque este efecto no necesariamente implica una ventaja farmacocinética neta, ya que también podría aumentar el riesgo de interacciones. En este contexto, las chalconas y varios derivados dihidroaromáticos presentaron valores más altos de inhibición de P-gp, mientras que otras familias, como los pirimido-benzimidazoles y algunos pirazoles, mostraron perfiles más moderados.

Respecto a la barrera hematoencefálica, una parte importante de las moléculas presentó probabilidad de penetración, especialmente dentro de las familias benzodiazepina y pirimido-benzimidazol. Este comportamiento puede interpretarse como una posible desventaja desde el punto de vista de seguridad, ya que una mayor distribución al sistema nervioso central podría favorecer efectos adversos no deseados. Esta observación resulta particularmente en las benzodiazepinas, que sugiere que estos compuestos podrían distribuirse al sistema nervioso central, lo que tiene lógica, ya que se sabe que esta familia es una famosa familia molécula de fármacos contra diferentes alteraciones o patologías neuronales o psiquiátricas. Aunque este comportamiento puede ser favorable en fármacos dirigidos a dianas cerebrales, en el contexto de inhibidores de PLpro constituiría una desventaja potencial, al aumentar el riesgo de efectos adversos centrales sin aportar necesariamente un beneficio terapéutico directo.

Los valores elevados de unión a proteínas plasmáticas y la baja fracción libre observada de manera general indican que, aunque varias moléculas presenten una absorción teórica favorable, su disponibilidad libre en plasma sería limitada. Así, la absorción inicial no debería interpretarse de forma aislada como un indicador suficiente para el desarrollo del potencial antiviral, ya que una alta unión plasmática puede restringir la fracción realmente disponible para interactuar con la diana biológica.

6.5.1.2 Metabolismo y excreción

El análisis metabólico sugiere que gran parte de las moléculas evaluadas presenta probabilidad de interacción con múltiples isoformas del citocromo P450, tanto como inhibidores como sustratos. Esta tendencia indica que, en varios casos, el metabolismo podría ser complejo y potencialmente susceptible a interacciones farmacológicas.

Dentro de este panorama, CYP2D6 parecería mantener una posición particular, ya que en varios compuestos la probabilidad de inhibición fue menor que la observada para otras isoformas, mientras que la condición de sustrato se mantuvo relativamente alta. Esto podría sugerir que, al menos para una fracción de las moléculas, esta isoforma participe de manera relevante en su metabolismo. No obstante, esta tendencia no es completamente uniforme entre familias, por lo que conviene interpretarla como una posibilidad predominante más que como una regla absoluta.

En relación con la excreción, la mayoría de los compuestos mostró valores de clearance bajos a moderados y vidas medias predichas relativamente cortas. Este patrón indica que, aun cuando algunas moléculas tengan un comportamiento estructural favorable frente a PLpro, su permanencia sistémica podría ser limitada, lo que constituye una desventaja para desarrollo oral convencional. Desde un punto de vista farmacológico, este resultado sugiere que varios candidatos podrían requerir optimización estructural adicional o estrategias de formulación si se quisiera mejorar su persistencia en circulación. Por otra parte, desde una perspectiva clínica, esta limitación podría mitigarse mediante vías de administración parenterales, como la infusión intravenosa.

6.5.1.3 Toxicidad

El análisis toxicológico muestra que ninguna de las familias evaluadas puede considerarse completamente libre de alertas, ya que en prácticamente todas se observaron probabilidades elevadas para uno o más endpoints de toxicidad, entendidos como categorías específicas de riesgo toxicológico predicho, tales como genotoxicidad, neurotoxicidad, hepatotoxicidad o toxicidad respiratoria.

La genotoxicidad predicha relativamente alta en varias estructuras podría relacionarse con la abundancia de sistemas aromáticos extendidos, los cuales favorecen interacciones π y podrían aumentar la probabilidad de interacciones no deseadas con biomacromoléculas. Del mismo modo, la neurotoxicidad cobra especial relevancia en aquellas familias que además muestran alta probabilidad de penetración de barrera hematoencefálica, como las benzodiazepinas, ya que ambos factores podrían potenciarse entre sí desde el punto de vista de seguridad.

La toxicidad se transforma en un criterio decisivo de priorización, ya que impide considerar como candidatos equivalentes a compuestos que, aunque dinámicamente sean favorables, presentan un perfil toxicológico comparativamente más problemático. Se tiene que considerar que el fármaco final no sería un medicamento de uso prologando, sino al ser un antiviral sería de un uso temporal. Se tiene que considerar las ventajas en comparación a los posibles efectos adversos.

6.5.1.4 Solubilidad

La distribución de los valores de logS mostró que la solubilidad acuosa predicha constituye una limitación transversal en la biblioteca evaluada, ya que la mayoría de los compuestos se ubicó por debajo del umbral de referencia de -4, con una mediana global de -5.726. Este comportamiento sugiere que, en términos generales, la colección analizada tiende a presentar un carácter marcadamente hidrofóbico, coherente con la presencia predominante de sistemas aromáticos extendidos, fragmentos fusionados y motivos estructurales poco ionizables. En este contexto, aunque algunos compuestos alcanzaron valores menos negativos, el panorama general indica que la solubilidad podría transformarse en un factor limitante durante etapas posteriores de optimización, particularmente si se busca avanzar hacia perfiles farmacocinéticos más equilibrados.

El análisis por familia química mostró que las series 3,4-dihidropirazol presentaron las medianas menos negativas de logS, lo que podría relacionarse con una menor planaridad efectiva y con la interrupción parcial de la conjugación respecto de otras familias más aromáticas. En contraste, las familias pirimidina-aminobenzimidazol y pirimido-benzimidazol exhibieron los valores más desfavorables, probablemente como consecuencia de una mayor rigidez estructural y de una superficie aromática más extensa. Aunque estos sistemas incorporan heteroátomos, ello no necesariamente se traduce en una mayor solubilidad, ya que el efecto favorable de la polaridad puede verse compensado por el aumento de hidrofobicidad asociado al núcleo policíclico y a la elevada planitud molécula. En este sentido, los resultados sugieren que la incorporación de anillos fusionados y sistemas heteroaromáticos condensados mejora ciertas características estructurales de interés, pero al costo de deteriorar la solubilidad acuosa.

La comparación según el aldehído precursor también reveló tendencias consistentes. Los derivados de p-aminobenzaldehído, salicilaldehído y vainillina mostraron valores menos negativos de logS, lo que puede explicarse por la introducción de grupos funcionales capaces de aumentar la polaridad global de la molécula, como amino, hidroxilo y metoxi. Estos sustituyentes favorecen interacciones con el medio acuoso y, por tanto, tienden a mejorar la solubilidad relativa dentro de la serie. En cambio, los compuestos derivados de p-(trifluorometil)benzalaldehído, 4-isopropilbenzalaldehído y p-nitrobenzalaldehído presentaron distribuciones desplazadas hacia valores más negativos. En el caso del grupo trifluorometilo y del sustituyente isopropilo, este resultado es coherente con un aumento del carácter lipofílico. Para el grupo nitro, aunque aporta polaridad, su carácter no ionizable y su contribución a sistemas aromáticos altamente conjugados no parece suficiente para compensar la menor afinidad global por el medio acuoso. Así, el tipo de sustituyente introducido desde el aldehído emerge como un factor importante en la modulación de la solubilidad.

El efecto de la cetona precursora fue particularmente marcado. Los derivados de acetofenona mostraron los valores menos negativos e incluso constituyeron el único grupo cuya mediana se ubicó por encima de -4, lo que sugiere que esqueletos aromáticos más simples y menos extendidos favorecen un perfil de

solubilidad relativamente mejor. A medida que aumenta el tamaño y la complejidad del sistema aromático, como ocurre en 1-acetilnaftaleno, 2-acetilnaftaleno, 9-acetilantraceno y especialmente 1-acetilpireno, la solubilidad disminuye de forma progresiva. Esta tendencia es químicamente razonable, ya que el incremento de la superficie π conjugada suele asociarse con un aumento de la hidrofobicidad y con una menor compatibilidad con el entorno acuoso. En particular, los derivados de 1-acetilpireno y 9-acetilantraceno evidencian que la expansión del núcleo policíclico aromático ejerce un efecto claramente desfavorable sobre logS.

Los resultados indican que la solubilidad acuosa predicha está fuertemente gobernada por el balance entre polaridad funcional y extensión del sistema aromático. La introducción de grupos polares en el aldehído puede mejorar parcialmente este parámetro, pero dicho efecto puede verse anulado cuando la molécula incorpora núcleos altamente fusionados o fragmentos aromáticos extensos. Desde el punto de vista del diseño racional, esto sugiere que futuras estrategias de optimización deberían considerar no solo la afinidad por el blanco molecula, sino también modificaciones orientadas a reducir la hidrofobicidad global o a introducir elementos que mejoren la solubilidad sin comprometer de manera importante la actividad biológica. En consecuencia, dentro de la biblioteca estudiada, los compuestos derivados de esqueletos menos extendidos, como los provenientes de acetofenona, y aquellos que contienen sustituyentes más polares en el anillo aromático, aparecen como puntos de partida más favorables desde la perspectiva de solubilidad.

6.5.1.5Integración farmacocinética con dinámica molécula

Al integrar los resultados ADME/Toxicidad con los análisis de dinámica molécula, se observa que los compuestos con mejor desempeño conformacional no son necesariamente los que presentan el perfil farmacocinético y toxicológico más favorable. Esta discrepancia es particularmente evidente en la familia benzodiazepina, cuyos representantes destacaron claramente en dinámica molecular por su estabilidad del ligando, rigidización del loop BL2 y comportamiento similar al de GRL-0617, pero mostraron desventajas importantes en distribución y seguridad, especialmente por su mayor probabilidad de penetración al sistema nervioso central y sus alertas toxicológicas.

En contraste, dentro de la familia pirimido-benzimidazol, el derivado 2-naftilo–4-OH-3-OMe surge como uno de los mejores compromisos entre comportamiento dinámico y perfil ADME/Tox relativo(Tabla 17). Aunque no puede considerarse exento de limitaciones, mostró una combinación más equilibrada entre estabilidad conformacional, permanencia del ligando, modulación del BL2 y propiedades farmacocinéticas más aceptables que las observadas en las benzodiazepinas.

Tabla 17 Valores de farmacocinética del pirimido benzimidazol 2-naftilo–4-OH-3-Ome.

Bloque	Parámetro	Valor	Lectura
Fisicoquímico	logS	-4.96	solubilidad baja
Absorción	Caco-2	-4.865	permeabilidad favorable
Absorción	HIA	0.00003	favorable
Absorción	F20 / F30	0.894 / 0.634	favorable / favorable
Absorción	P-gp inh / sub	No / No	sin señal relevante de P-gp
Distribución	BBB	0.272	BBB–
Distribución	PPB (%)	98.11	alta unión
Distribución	Fu (%)	1.41	baja fracción libre
Distribución	logVDss	-0.032	distribución moderada
Metabolismo	CYP1A2 inh / sub	Sí / Sí	probable interacción / probable sustrato

Metabolismo	CYP2C19 inh / sub	Sí / No	probable interacción / sustrato poco probable
Metabolismo	CYP2C9 inh / sub	Sí / No	probable interacción / sustrato poco probable
Metabolismo	CYP2D6 inh / sub	No / No	baja probabilidad / sustrato poco probable
Metabolismo	CYP3A4 inh / sub	Sí / No	probable interacción / sustrato poco probable
Excreción	CL-plasma	4.573	clearance baja- moderada
Toxicidad	hERG / hERG-10 μM	0.609 / 0.786	alerta
Toxicidad	DILI	0.885	alerta
Toxicidad	Ames	0.829	alerta

Por otra parte, algunos compuestos con desempeño dinámico más parcial, como ciertos pirazoles, podrían conservar interés como puntos de partida para optimización, precisamente porque su perfil de seguridad relativa parece menos desfavorable que el de los compuestos dinámicamente más potentes. Esto refuerza la idea de que la priorización final no debe realizarse únicamente en función de la estabilidad del complejo proteína–ligando, sino a partir de una evaluación integrada que considere simultáneamente afinidad estructural, comportamiento dinámico, farmacocinética y toxicidad. Estos resultados indican que los compuestos más prometedores del estudio no son necesariamente aquellos que maximizan una sola dimensión, sino los que logran el mejor equilibrio entre eficacia estructural y viabilidad farmacológica. Bajo esta perspectiva, los derivados benzodiazepínicos pueden considerarse los mejores hits mecanísticos, mientras que compuestos como 2-naftilo–4-OH-3-OMe representarían candidatos más razonables como leads para futuras etapas de optimización.

6.5.2 Farmacocinética de derivados de benzimidazoles

6.5.2.1 Absorción y distribución

Desde el punto de vista de absorción, la serie de benzimidazoles presenta un comportamiento globalmente favorable. La homogeneidad fisicoquímica observada en la serie, junto con el cumplimiento completo de las reglas de Lipinski y Veber, sugiere que estos compuestos conservan propiedades compatibles con biodisponibilidad oral adecuada. En este sentido, no se trata de una serie penalizada por peso molecula excesivo, polaridad elevada o flexibilidad estructural desmedida, sino de un conjunto relativamente bien balanceado desde el punto de vista de drogabilidad. Esta apreciación se ve reforzada por los descriptores de absorción predicha, donde los seis derivados muestran valores compatibles con una absorción intestinal favorable y una permeabilidad epitelial relativamente alta. En otras palabras, la serie no parece presentar una barrera importante en la etapa inicial de ingreso al organismo. Sin embargo, este perfil debe interpretarse con cierto cuidado, ya que todos los compuestos exhiben una alta probabilidad de comportarse como inhibidores de P-gp, mientras que la probabilidad de actuar como sustratos es baja. Esto podría ser ventajoso en términos de evitar un eflujo excesivo, pero también podría traducirse en un mayor riesgo de interacciones farmacológicas si estas moléculas se evaluaran en combinación con otros compuestos dependientes de transportadores. En distribución, el panorama es menos uniforme y resulta más discriminante. La serie se divide de manera relativamente equilibrada entre compuestos con mayor y menor probabilidad de atravesar la barrera hematoencefálica. Esto es importante, porque en el contexto de inhibidores de PLpro una alta penetración al sistema nervioso central no constituiría necesariamente una ventaja, y podría incluso convertirse en una limitación de seguridad. Bajo esta lectura, benzimidazol 3 y benzimidazol 4 serían relativamente más favorables que otros miembros de la serie, ya que presentan una menor probabilidad de distribución central. Por otra parte, todos los derivados muestran una unión a proteínas plasmáticas muy alta y una fracción libre muy baja. Esta combinación indica que, aunque la absorción inicial pueda ser favorable, una porción

importante del compuesto permanecería secuestrada por proteínas plasmáticas, reduciendo la fracción realmente disponible para alcanzar la diana. Para la serie parecería favorable en absorción, pero menos ventajosa en términos de disponibilidad libre sistémica. Dentro de este contexto, benzimidazol 3 y benzimidazol 4 mostrarían una leve ventaja relativa, al presentar una fracción libre algo mayor y una unión plasmática ligeramente menor que el resto, aunque la diferencia no es suficientemente amplia como para modificar por completo la interpretación global.

6.5.2.2 Metabolismo y excreción

El perfil metabólico de la serie constituye uno de sus principales puntos de atención. En general, los benzimidazoles presentan una alta probabilidad de inhibición frente a varias isoformas del citocromo P450, particularmente CYP1A2, CYP2C19 y CYP2C8, y en varios casos también frente a CYP2C9. Esto sugiere que la serie podría mostrar una tendencia relativamente marcada a interacciones metabólicas, lo que reduce su atractivo como candidatos directos sin optimización adicional. En contraste, la inhibición de CYP2D6 parece mucho menor en comparación con otras isoformas, lo que abre la posibilidad de que esta ruta conserve mayor participación metabólica relativa en algunos derivados, aunque no de forma exclusiva.

Dentro de la serie, benzimidazol 2 parecería ser el menos problemático respecto a CYP3A4, mientras que benzimidazol 5 y benzimidazol 6 mostrarían una inhibición mucho más marcada de esta isoforma. En ese sentido, benzimidazol 3 vuelve a posicionarse como un caso interesante, ya que aunque conserva la tendencia general de la serie a inhibir múltiples isoformas, su perfil no parece tan desfavorable como el de varios de sus análogos más extremos.

En cuanto a excreción, la serie vuelve a mostrar un comportamiento bastante homogéneo. El clearance plasmático predicho se mantiene en rangos bajos a moderados, mientras que la vida media predicha es corta en todos los compuestos. Esto sugiere que, aun si la absorción inicial es favorable, la permanencia sistémica podría ser limitada, lo que sería una desventaja para desarrollo oral convencional. No obstante, las diferencias internas permiten nuevamente matizar la serie: benzimidazol 5 presenta el clearance más bajo, mientras que benzimidazol 3 se ubica entre los compuestos con vida media relativamente más alta. Así, aunque ninguno de los derivados destaque claramente por un perfil ideal de excreción, benzimidazol 3 parecería mantener un compromiso algo más favorable que otros miembros de la serie.

6.5.2.3 Toxicidad

La toxicidad constituye la principal limitación de los benzimidazoles. Los resultados del análisis muestran que la serie no puede considerarse limpia desde el punto de vista toxicológico, ya que múltiples endpoints presentan probabilidades elevadas de respuesta positiva. Entre los más preocupantes destacan genotoxicidad, toxicidad respiratoria, neurotoxicidad, nefrotoxicidad y hepatotoxicidad inducida por fármacos. Este patrón no corresponde a una alerta aislada, sino a una carga toxicológica múltiple y relativamente consistente dentro de toda la serie.

A pesar de ello, la comparación interna sugiere que no todos los compuestos son igualmente desfavorables. Benzimidazol 3 parecería ser el menos problemático del conjunto, ya que presenta valores relativamente menores en varios endpoints críticos, incluyendo Ames, neurotoxicidad, hERG y nefrotoxicidad, aunque todavía mantiene probabilidades altas en DILI y genotoxicidad. En contraste, benzimidazol 4 y benzimidazol 6 muestran perfiles más comprometidos en varios de estos mismos parámetros, lo que los vuelve menos atractivos desde una perspectiva de priorización farmacológica.

Esta lectura resulta particularmente relevante al cruzarse con la dinámica molecular. En la discusión previa, benzimidazol 3 ya había emergido como uno de los compuestos más interesantes por su capacidad de modular el BL2 y mantener una permanencia razonable en la región de unión, mientras que benzimidazol 5 se había destacado por su semejanza conformacional con el estado inducido por GRL-0617. Al integrar ahora el perfil ADME/Tox, la balanza tendería a inclinarse más claramente hacia benzimidazol 3, ya que conserva interés dinámico y además presenta el perfil farmacocinético y toxicológico relativamente menos desfavorable de la serie. Por ello, dentro de los benzimidazoles, benzimidazol 3 podría proponerse como el mejor compromiso global entre comportamiento dinámico y viabilidad farmacológica, mientras que benzimidazol 5 conservaría valor principalmente como hit mecanístico, pero con un perfil ADME/Toxicidad menos atractivo para priorización directa.

6.5.2.4 Solubilidad

En la serie de benzimidazoles, los valores de logS se distribuyeron entre -4.694 y -4.181, ubicándose en todos los casos por debajo del umbral de referencia de -4. Esto indica que, dentro de este subconjunto, la solubilidad acuosa predicha es uniformemente baja, aunque con diferencias moderadas entre compuestos. En términos comparativos, Benzimidazol 5 y Benzimidazol 3 presentaron los valores menos negativos, mientras que Benzimidazol 2 y Benzimidazol 1 mostraron los valores más desfavorables. Este comportamiento sugiere que, aun cuando los benzimidazoles incorporan heteroátomos nitrogenados capaces de aportar polaridad, ello no resulta suficiente para compensar el efecto hidrofóbico asociado a los fragmentos aromáticos presentes en la serie. Por tanto, en estos compuestos la solubilidad probablemente constituye una limitación a considerar en etapas posteriores de optimización, especialmente si se busca mejorar su perfil farmacocinético sin comprometer la actividad biológica.

6.5.3 Integración de resultados de farmacodinámica predicha de benzimidazoles y derivados de chalconas

La integración de los resultados de dinámica molecular con el análisis farmacocinético y toxicológico muestra que los compuestos con mejor desempeño estructural frente a PLpro no coinciden necesariamente con aquellos que presentan el perfil ADME/Toxicidad más favorable. Dentro de las familias derivadas de chalconas y heterociclos, las benzodiazepinas emergieron como los sistemas más sólidos desde el punto de vista dinámico, al reproducir de manera más consistente el patrón conformacional asociado al inhibidor de referencia; sin embargo, su perfil de distribución y toxicidad las vuelve menos atractivas como candidatos directos. En contraste, la familia pirimido-benzimidazol presentó un equilibrio más favorable entre comportamiento dinámico y desarrollabilidad, destacando particularmente el derivado 2-naftilo-4-OH-3-OMe como uno de los compuestos más consistentes al integrar estabilidad del ligando, modulación del BL2 y un perfil farmacocinético relativamente menos desfavorable. A su vez, en la serie independiente de benzimidazoles, benzimidazol 3 se posicionó como el mejor compromiso global, al combinar un efecto dinámico relevante con el perfil ADME/Tox comparativamente más aceptable dentro de su serie. En este sentido, los resultados no solo permiten identificar hits mecanísticos, sino también diferenciar aquellos compuestos que podrían constituir puntos de partida más realistas para optimización futura, orientando el diseño posterior hacia estructuras que conserven la capacidad de modulación conformacional de PLpro, pero con menores restricciones farmacocinéticas y toxicológicas.

6.5.4 Farmacocinética de moléculas top 1 del *virtual screening*

6.5.4.1 Absorción y distribución

Aunque todas las moléculas seleccionadas mediante virtual screening fueron clasificadas como favorables según HIA, el análisis conjunto con P-gp y Caco-2 muestra que el comportamiento absortivo no es equivalente entre compuestos. En este contexto, Benzimidazol, Benzodiazepina y Pirimidina aminobenzimidazol presentaron los valores más favorables de permeabilidad predicha en Caco-2, ubicándose por sobre el umbral asociado a alta permeabilidad intestinal. En cambio, Aminopirimidina, Pirimido benzimidazol, Benzamida y Benzotiofeno quedaron por debajo de dicho umbral, lo que sugiere una permeación intestinal más limitada pese a la clasificación favorable observada en HIA. Esto indica que la señal global de absorción intestinal debe interpretarse con cautela, ya que la permeabilidad efectiva parece ser claramente dependiente de la familia química.

La interacción con P-glicoproteína refuerza esta diferenciación. Benzimidazol, Benzotiofeno y Pirimidina aminobenzimidazol mostraron altas probabilidades de comportarse como inhibidores de P-gp, pero bajas probabilidades de ser sustratos, lo que sugiere que podrían no verse fuertemente penalizados por eflujo directo y, al mismo tiempo, tener potencial para interferir con este transportador. Benzodiazepina presentó un comportamiento más complejo, ya que mostró alta probabilidad tanto de inhibición como de sustrato, lo que sugiere una interacción dual con P-gp: por una parte podría ser reconocida y expulsada por el transportador y, por otra, competir con él o modular su actividad. En términos prácticos, esto vuelve más difícil anticipar su absorción neta y su exposición real. En contraste, Aminopirimidina mostró una probabilidad prácticamente nula de inhibición, pero una probabilidad muy alta de comportarse como sustrato, lo que sugiere que su absorción efectiva podría verse particularmente afectada por mecanismos de eflujo. Benzamida y Pirimido benzimidazol, en cambio, mostraron bajas probabilidades tanto de inhibición como de comportamiento como sustrato, lo que sugiere una interacción menor con P-gp y un perfil más neutro desde este punto de vista.

Desde la perspectiva de la distribución, la mayoría de las moléculas fue clasificada como BBB+, incluyendo Aminopirimidina, Benzimidazol, Benzodiazepina, Pirimidina aminobenzimidazol y Pirimido benzimidazol, mientras que Benzamida y Benzotiofeno fueron clasificadas como BBB-. Este resultado sugiere que la mayor parte del conjunto presenta una propensión a atravesar la barrera hematoencefálica, lo que podría traducirse en exposición al sistema nervioso central. Esto podría sugerir mayor riesgo de efectos secundarios centrales. En ese escenario, Benzamida y Benzotiofeno tendrían un perfil más conservador, al mostrar una menor probabilidad de penetración cerebral.

La unión a proteínas plasmáticas (PPB) fue alta para la mayoría de los compuestos, especialmente en Benzimidazol, Benzodiazepina y Pirimidina aminobenzimidazol, todos con valores cercanos al 98%, lo que fue consistente con fracciones libres muy bajas. Benzotiofeno y Pirimido benzimidazol también mostraron una unión plasmática elevada, aunque menos extrema. Benzamida presentó una PPB claramente menor y, en consecuencia, una fracción libre (F_u) más alta. Aminopirimidina constituyó un caso atípico dentro del conjunto, ya que el modelo predijo una PPB prácticamente nula y una F_u cercana al 100%; este comportamiento sugiere una disponibilidad libre muy alta, aunque por tratarse de un valor extremo conviene interpretarlo con cautela. En términos comparativos, Benzamida, Benzotiofeno y Pirimido benzimidazol son los compuestos que presentan las fracciones libres más intermedias o amplias dentro del conjunto, mientras que Benzimidazol, Benzodiazepina y Pirimidina aminobenzimidazol se mantuvieron más restringidos en su fracción libre circulante.

Por otra parte, el VDss se mantuvo dentro de un rango globalmente adecuado para todas las moléculas, lo que indica que ningún compuesto se desplazó hacia comportamientos extremos de distribución tisular. No obstante, se observaron diferencias relativas: Benzimidazol y Pirimidina aminobenzimidazol presentaron los valores más altos, seguidos de Benzotiofeno y Benzodiazepina, lo que sugiere una distribución tisular comparativamente mayor. En cambio, Aminopirimidina, Benzamida y Pirimido benzimidazol mostraron valores más bajos, aunque todavía dentro del rango considerado aceptable.

Estos resultados indican que los compuestos con el perfil más consistente en absorción y distribución fueron Benzimidazol y Pirimidina aminobenzimidazol, ya que combinaron buena permeabilidad en Caco-2, alta probabilidad de inhibición de P-gp pero baja tendencia a comportarse como sustratos, penetración BBB favorable y parámetros de distribución compatibles con una adecuada exposición sistémica. Benzodiazepina también mostró buena permeabilidad, aunque su interacción dual con P-gp introduce mayor incertidumbre. Por su parte la Aminopirimidina y Benzamida presentaron perfiles más particulares: la primera por su alta probabilidad de ser sustrato de P-gp y su comportamiento plasmático atípico, y la segunda por su baja permeabilidad y su menor penetración BBB. Benzotiofeno y Pirimido benzimidazol ocuparon posiciones intermedias, con ventajas en algunos parámetros pero sin un desempeño tan equilibrado como el observado para los compuestos más favorables.

6.5.4.2 Metabolismo y excreción

El análisis del metabolismo predicho mostró diferencias importantes entre las moléculas seleccionadas mediante virtual screening, lo que confirma que no comparten un comportamiento metabólico uniforme. En este contexto, Aminopirimidina fue el compuesto con el perfil más favorable, ya que no presentó probabilidades altas ni de inhibición ni de comportamiento como sustrato frente a las isoformas de

CYP450 evaluadas. Esto sugiere una menor dependencia de estas rutas metabólicas y, al mismo tiempo, una menor probabilidad de interferir con ellas, lo que representa una ventaja farmacocinética relativa dentro del conjunto.

En el extremo opuesto, Benzodiazepina presentó el perfil más comprometido, con alta probabilidad de inhibición sobre CYP1A2, CYP2C19, CYP2C9, CYP3A4, CYP2B6 y CYP2C8, y además alta probabilidad de comportarse como sustrato de CYP1A2, CYP2C19, CYP2C9 y CYP2D6. Este patrón sugiere una interacción metabólica amplia y poco favorable, ya que combina una fuerte dependencia de varias rutas oxidativas con una elevada probabilidad de interferir con ellas, lo que podría traducirse en un mayor riesgo de interacciones metabólicas y en una farmacocinética menos predecible.

Benzimidazol y Pirimidina aminobenzimidazol mostraron un comportamiento similarmente desfavorable. Benzimidazol presentó alta probabilidad de inhibición sobre CYP1A2, CYP2C19, CYP2C9, CYP3A4 y CYP2C8, además de probabilidad alta de ser sustrato de CYP1A2, CYP2C9 y CYP2D6. Por su parte, Pirimidina aminobenzimidazol mostró inhibición probable de CYP1A2, CYP2C19, CYP2C9, CYP2D6, CYP3A4 y CYP2C8, junto con comportamiento como sustrato de CYP1A2, CYP2C9, CYP2D6 y CYP2B6. En ambos casos, la combinación de múltiples isoformas inhibidas y varias rutas potenciales de biotransformación sugiere un perfil metabólico complejo, con riesgo elevado de variabilidad en depuración e interacciones farmacológicas.

En una posición intermedia se ubicaron Benzamida, Benzotiofeno y Pirimido Benzimidazol. Benzamida mostró inhibición probable de CYP2C9 y CYP2C8, y comportamiento como sustrato de CYP2C19, CYP2C9 y CYP3A4. Benzotiofeno presentó inhibición de CYP2B6 y CYP2C8, además de probabilidad alta de ser sustrato de CYP2C19, CYP2C9, CYP2D6 y CYP3A4. En cambio, Pirimido Benzimidazol fue más acotado, ya que solo mostró inhibición relevante sobre CYP2C8 y comportamiento como sustrato de CYP1A2 y CYP2C19. Por lo tanto, aunque estos tres compuestos no presentan un perfil tan problemático como Benzodiazepina, Benzimidazol o Pirimidina aminobenzimidazol, sí mantienen una dependencia metabólica apreciable de isoformas específicas.

Desde el punto de vista de la excreción, el comportamiento fue más homogéneo. Todas las moléculas presentaron clearance plasmático bajo, con valores inferiores al umbral de referencia de 5 mL/min/kg. Dentro de este rango, Benzamida presentó el valor más bajo, seguida de Pirimidina aminobenzimidazol y Aminopirimidina, mientras que Benzodiazepina, Benzotiofeno, Benzimidazol y Pirimido Benzimidazol mostraron valores algo mayores, aunque todavía dentro de un intervalo bajo. Esto indica que, en términos de clearance, ninguna molécula se asocia a una eliminación particularmente rápida.

Sin embargo, la vida media sí permitió distinguir mejor a los compuestos. Aminopirimidina destacó claramente al presentar la vida media más prolongada del conjunto (2.07 h), mientras que el resto de las moléculas se concentró en valores considerablemente más bajos, entre aproximadamente 0.44 y 0.77 h. Los valores más breves correspondieron a Benzotiofeno, Benzodiazepina y Benzimidazol, mientras que Benzamida, Pirimido Benzimidazol y Pirimidina aminobenzimidazol mostraron valores intermedios, aunque todavía cortos. Esto sugiere que, aunque el clearance plasmático es bajo en todo el conjunto, la permanencia sistémica efectiva sigue siendo reducida para la mayoría de las moléculas.

Los resultados indican que Aminopirimidina presenta el perfil más favorable en el eje metabolismo–excreción, al combinar una baja interacción predicha con CYP450 con la vida media más prolongada del conjunto. Por otra parte la Benzodiazepina, Benzimidazol y Pirimidina aminobenzimidazol son las moléculas más comprometidas desde el punto de vista metabólico, debido a su alta probabilidad simultánea de inhibición y comportamiento como sustrato frente a múltiples isoformas. Benzamida, Benzotiofeno y Pirimido Benzimidazol ocupan una posición intermedia, con perfiles menos extremos pero todavía dependientes de rutas metabólicas específicas. Aunque la vida media corta observada en la mayoría del conjunto no excluye necesariamente su interés como candidatos, sí sugiere que, en etapas posteriores de desarrollo, podrían requerirse ajustes en el esquema de administración o en la estrategia de formulación para sostener exposiciones adecuadas.

6.5.4.3 Toxicidad

El perfil toxicológico predicho mostró una diferenciación marcada entre las moléculas seleccionadas mediante virtual screening, lo que indica que la carga de riesgo no se distribuye de manera uniforme entre las distintas familias químicas. En términos generales, los endpoints que concentraron las señales más consistentes y recurrentes en el conjunto fueron genotoxicidad, DILI y hepatotoxicidad humana, ya que todos o casi todos los compuestos presentaron probabilidades elevadas en estas categorías. En contraste, los endpoints asociados a corrosión e irritación oculares mostraron valores muy bajos en todas las moléculas, lo que sugiere un comportamiento comparativamente favorable en esta dimensión.

Dentro del conjunto, Benzodiazepina fue el compuesto con el perfil toxicológico más comprometido. Presentó la mayor carga toxicológica global, con 16 endpoints por sobre el umbral de 0.5, y mostró probabilidades altas en una amplia gama de parámetros, incluyendo hERG, hERG a 10 μ M, DILI, Ames, FDAMDD, carcinogenicidad, toxicidad respiratoria, hepatotoxicidad humana, neurotoxicidad, ototoxicidad, hematotoxicidad, nefrotoxicidad, genotoxicidad, así como citotoxicidad en A549 y HEK293. Este patrón sugiere que Benzodiazepina es el compuesto de menor atractivo toxicológico dentro del conjunto, no por un único endpoint dominante, sino por la acumulación simultánea de múltiples señales adversas.

Benzimidazol también presentó un perfil claramente desfavorable, con 14 endpoints por sobre 0.5. Mostró probabilidades altas de hERG, hERG a 10 μ M, DILI, Ames, FDAMDD, carcinogenicidad, hepatotoxicidad humana, neurotoxicidad, ototoxicidad, nefrotoxicidad, genotoxicidad, y además una señal marcada de citotoxicidad en A549 y HEK293. En conjunto, su perfil sugiere una toxicidad potencial amplia, especialmente en dimensiones genotóxicas, hepáticas y citotóxicas.

Un patrón también comprometido se observó para Benzotiofeno, que presentó 12 endpoints positivos. En este caso destacaron hERG, DILI, toxicidad oral aguda (ROA), FDAMDD, toxicidad respiratoria, hepatotoxicidad humana, neurotoxicidad, ototoxicidad, hematotoxicidad, nefrotoxicidad, genotoxicidad y citotoxicidad en HEK293. A diferencia de Benzimidazol y Benzodiazepina, Benzotiofeno no mostró una señal importante en Ames ni en carcinogenicidad, pero sí mantuvo una combinación de alertas que lo sitúan entre los compuestos más problemáticos del grupo.

La Pirimidina aminobenzimidazol presentó una carga toxicológica intermedia-alta, con 11 endpoints por sobre el umbral. Mostró probabilidades elevadas en hERG, hERG a 10 μ M, DILI, Ames, FDAMDD, carcinogenicidad, toxicidad respiratoria, hepatotoxicidad humana, neurotoxicidad, genotoxicidad y citotoxicidad en HEK293. Aunque no alcanzó la amplitud de señales observada en Benzodiazepina o Benzimidazol, sí conserva una combinación relevante de alertas, especialmente en el eje hepático-genotóxico.

La Aminopirimidina y Pirimido Benzimidazol mostraron una carga toxicológica similar, con 10 endpoints positivos cada uno, aunque con perfiles cualitativamente distintos. Aminopirimidina destacó por probabilidades altas de DILI, Ames, sensibilización cutánea, carcinogenicidad, hepatotoxicidad humana, ototoxicidad, hematotoxicidad y genotoxicidad, pero mostró valores bajos en hERG, toxicidad respiratoria, nefrotoxicidad y citotoxicidad celular. Esto sugiere que, aunque no sea uno de los compuestos con mayor burden global, sí presenta alertas específicas importantes, particularmente en genotoxicidad y toxicidad hepática. La Pirimido Benzimidazol mostró un patrón más dirigido hacia DILI, Ames, hepatotoxicidad humana, neurotoxicidad, ototoxicidad, hematotoxicidad, nefrotoxicidad, genotoxicidad y HEK293, además de una señal positiva en hERG. Por tanto, aunque su carga global es similar a la de Aminopirimidina, su perfil está más desplazado hacia toxicidad neurológica, renal y cardíaca.

El caso más favorable fue la Benzamida, que presentó solo 4 endpoints con probabilidad superior a 0.5: DILI, FDAMDD, hepatotoxicidad humana y genotoxicidad. El resto de los parámetros se mantuvo en niveles bajos o moderados, sin señales relevantes en hERG, Ames, carcinogenicidad, toxicidad respiratoria, nefrotoxicidad ni en citotoxicidad celular. Esto posiciona a la Benzamida como el compuesto con el perfil toxicológico más limpio del conjunto, aunque las señales persistentes de DILI, hepatotoxicidad humana y genotoxicidad indican que tampoco puede considerarse exento de alertas.

Estos resultados indicarían que la Benzamida presenta el perfil toxicológico más favorable dentro del grupo, seguida a distancia por la Aminopirimidina y Pirimido Benzimidazol, aunque ambos conservan advertencias importantes en endpoints específicos. En contraste, Benzodiazepina, Benzimidazol y Benzotiofeno emergen como los compuestos más comprometidos, debido a la acumulación de señales altas en múltiples dimensiones de toxicidad. Pirimidina aminobenzimidazol ocupa una posición

intermedia, con una carga relevante pero menos extrema. Por lo tanto, desde la perspectiva toxicológica, la priorización de candidatos no debería basarse solo en el número total de endpoints positivos, sino también en la naturaleza de los riesgos asociados, particularmente aquellos vinculados a genotoxicidad, hepatotoxicidad y cardiotoxicidad.

6.5.4.4 Solubilidad

La solubilidad predicha constituyó uno de los parámetros más discriminantes entre las moléculas seleccionadas mediante virtual screening, evidenciando diferencias marcadas entre familias químicas. Los valores de logS se distribuyeron entre -12.44 y -1.88, con una mediana de -4.44, lo que indica que el conjunto no presenta un comportamiento homogéneo desde el punto de vista fisicoquímico. Considerando el umbral empírico de logS = -4.0, solo Aminopirimidina, Benzamida y Pirimido Benzimidazol se ubicaron por sobre ese valor, mientras que Benzotiofeno, Benzodiazepina, Pirimidina Aminobenzimidazol y Benzimidazol quedaron por debajo, sugiriendo limitaciones crecientes de solubilidad.

Dentro del conjunto, Aminopirimidina presentó el perfil más favorable, con un logS de -1.88, claramente superior al resto de las moléculas. Este resultado sugiere una solubilidad acuosa comparativamente alta y la posiciona como el compuesto más ventajoso desde esta propiedad. También Benzamida (-3.20) y Pirimido Benzimidazol (-3.23) mostraron valores compatibles con una solubilidad relativamente aceptable, ambos por encima del umbral de referencia. Aunque estos dos compuestos no alcanzan el comportamiento de Aminopirimidina, sí se distinguen del resto por presentar un perfil más equilibrado y potencialmente menos limitante para el desarrollo.

En una posición intermedia se ubicó Benzotiofeno, con un logS de -4.44. Aunque este valor se encuentra por debajo del umbral considerado favorable, su cercanía al punto de corte sugiere una situación menos desfavorable que la observada para otros compuestos del conjunto. Por lo tanto, Benzotiofeno podría interpretarse como una molécula con solubilidad moderadamente limitada, pero no tan comprometida como aquellas con valores claramente más bajos.

Los perfiles menos favorables correspondieron a Benzodiazepina, Pirimidina Aminobenzimidazol y Benzimidazol. Benzodiazepina presentó un logS de -7.48, lo que ya sugiere una solubilidad marcadamente reducida. Sin embargo, los casos más críticos fueron Pirimidina Aminobenzimidazol (-10.59) y, sobre todo, Benzimidazol (-12.44), que mostraron los valores más bajos de todo el conjunto. Estos resultados indican que, aun cuando algunos de estos compuestos puedan presentar ventajas en otros parámetros ADMET o en el estudio de afinidad, su desarrollo podría verse fuertemente condicionado por una solubilidad acuosa muy deficiente.

La solubilidad permite establecer una diferenciación clara dentro del grupo. Aminopirimidina emerge como el compuesto más favorable desde esta propiedad, seguida por Benzamida y Pirimido Benzimidazol, que también mantienen valores aceptables. En contraste, Benzimidazol y Pirimidina Aminobenzimidazol representan los casos más problemáticos, mientras que Benzodiazepina también muestra una limitación importante. Por ello, la solubilidad aparece como un criterio de priorización particularmente relevante dentro del conjunto VS, ya que puede transformarse en una restricción práctica importante incluso en moléculas con desempeño favorable en otras dimensiones.

6.5.4.5 Integración de moléculas top 1 del *virtual screening*

Las moléculas seleccionadas mediante virtual screening mostraron un perfil ADMET claramente heterogéneo, lo que confirma que no pueden discutirse como un grupo uniforme. Desde el punto de vista de la absorción, todas presentaron una predicción favorable según HIA, pero este resultado se matiza al considerar Caco-2 y P-gp. Benzimidazol, Benzodiazepina y Pirimidina aminobenzimidazol fueron las que mostraron mejor permeabilidad predicha en Caco-2, mientras que Aminopirimidina, Benzamida, Benzotiofeno y Pirimido Benzimidazol quedaron por debajo del umbral de alta permeabilidad. Además, la

interacción con P-gp introdujo una segunda fuente de diferenciación: Benzimidazol, Benzotiofeno y Pirimidina aminobenzimidazol se asociaron principalmente a un perfil de inhibición, Aminopirimidina destacó como sustrato claro, y Benzodiazepina presentó simultáneamente alta probabilidad de inhibición y de sustrato, lo que sugiere un comportamiento absortivo menos predecible. Por lo tanto, aunque la absorción basal del conjunto parece favorable, su comportamiento real probablemente dependerá de manera importante de la familia química y del peso relativo de la permeabilidad y el eflujo.

Las propiedades de distribución también permitieron diferenciar a los compuestos. La mayoría fue clasificada como BBB+, excepto Benzamida y Benzotiofeno, mientras que la unión a proteínas plasmáticas fue alta para casi todo el conjunto, especialmente en Benzimidazol, Benzodiazepina y Pirimidina aminobenzimidazol. En cambio, Benzamida, Benzotiofeno y Pirimido Benzimidazol mostraron fracciones libres más intermedias, mientras que Aminopirimidina se apartó del resto por una predicción extrema de PPB muy baja y Fu muy alta, que debe interpretarse con cautela. El VDss se mantuvo dentro de un rango globalmente adecuado para todas las moléculas, aunque Benzimidazol y Pirimidina aminobenzimidazol presentaron los valores más altos, lo que sugiere una distribución tisular comparativamente mayor. En conjunto, absorción y distribución favorecen especialmente a Benzimidazol y Pirimidina aminobenzimidazol en términos de permeabilidad y partición sistémica, mientras que Benzamida y Pirimido Benzimidazol muestran perfiles más conservadores, y Aminopirimidina conserva un comportamiento particular por su condición de sustrato de P-gp y por sus parámetros plasmáticos atípicos.

El metabolismo fue uno de los dominios más discriminantes. Aminopirimidina presentó el perfil más limpio, ya que no mostró probabilidades altas ni de inhibición ni de comportamiento como sustrato frente a las isoformas evaluadas, lo que sugiere menor dependencia del sistema CYP450. En el extremo opuesto, Benzodiazepina fue la molécula metabólicamente más comprometida, al combinar alta probabilidad de inhibición sobre múltiples isoformas con alta probabilidad de comportarse como sustrato de varias de ellas; este patrón sugiere una interacción metabólica amplia, compleja y menos deseable. Benzimidazol y Pirimidina aminobenzimidazol mostraron un comportamiento similar, también con una superposición importante entre rutas inhibidas y rutas de biotransformación probables, lo que dificulta anticipar una vía metabólica dominante simple. En cambio, algunos compuestos sí permiten inferir rutas más definidas: Benzotiofeno mostró probabilidad de comportarse como sustrato de CYP2C19, CYP2C9, CYP2D6 y CYP3A4 sin una señal paralela importante de inhibición sobre esas mismas isoformas, por lo que su biotransformación podría depender preferentemente de esas vías; Pirimido Benzimidazol mostró un patrón relativamente acotado como sustrato de CYP1A2 y CYP2C19; y Benzamida sugiere una dependencia principal de CYP2C19 y CYP3A4, con una participación adicional de CYP2C9, aunque en esta última isoforma también aparece señal de inhibición. Así, el análisis metabólico permite distinguir compuestos con rutas más previsibles de otros con mayor riesgo de complejidad farmacocinética e interacciones.

La excreción fue comparativamente más homogénea. Todas las moléculas presentaron clearance plasmático bajo, y la principal diferencia se observó en la vida media, donde Aminopirimidina destacó al conservar el valor más alto del conjunto, mientras que el resto mostró vidas medias cortas. Esto sugiere que, aunque ninguna molécula se asocia a una depuración particularmente rápida, la permanencia sistémica efectiva podría ser limitada en la mayoría de los casos. Desde una perspectiva de desarrollo, este comportamiento no invalida necesariamente a los compuestos, pero sí indica que podría requerirse optimización posterior del esquema de administración o de la formulación para sostener exposiciones adecuadas.

La toxicidad y la solubilidad terminan de refinar la priorización. En toxicidad, Benzamida emergió como el compuesto con el perfil más limpio, mientras que Benzodiazepina, Benzimidazol y Benzotiofeno concentraron la mayor carga toxicológica. De manera consistente, los endpoints más preocupantes del conjunto fueron genotoxicidad, DILI y hepatotoxicidad humana, mientras que los parámetros de toxicidad ocular fueron comparativamente favorables. En solubilidad, Aminopirimidina mostró el mejor comportamiento con amplia diferencia, seguida por Benzamida y Pirimido Benzimidazol, mientras que Benzimidazol y Pirimidina aminobenzimidazol presentaron los valores más desfavorables, y Benzodiazepina también mostró una limitación importante. Integrando todas las dimensiones, Benzamida y Pirimido Benzimidazol aparecen como los compuestos más equilibrados desde el punto de vista de desarrollabilidad global, aunque con limitaciones de permeabilidad que no deben ignorarse.

Aminopirimidina destaca por su solubilidad y por su perfil metabólico más limpio, pero su comportamiento como sustrato de P-gp, su patrón plasmático atípico y varias alertas toxicológicas obligan a interpretarla con cautela. Por su parte el Benzimidazol y Pirimidina aminobenzimidazol presentan ventajas en absorción y distribución, pero quedan penalizados por un perfil metabólico más complejo, mayor carga toxicológica y, en algunos casos, solubilidad desfavorable. Por ello, el principal valor del análisis integrado no radica en validar uniformemente al conjunto, sino en discriminar cuáles compuestos mantienen un balance más razonable entre propiedades favorables y limitaciones críticas.

6.6 Retrosíntesis

6.6.1 Derivados de chalconas

Las familias de los diferentes heterociclos pueden racionalizarse retrosintéticamente a partir de sus chalconas precursoras correspondientes, tal como se muestra en los esquemas de retrosíntesis. Esta desconexión resulta especialmente útil porque conduce a un núcleo sintético común, desde el cual es posible acceder a varias familias heterocíclicas mediante transformaciones posteriores relativamente directas. En este sentido, la chalcona no solo actúa como un intermedio estructural clave, sino también como una plataforma convergente que permite variar de manera independiente el fragmento derivado del aldehído, el proveniente de la cetona y, posteriormente, el nucleófilo empleado en la ciclación¹¹⁹.

La obtención de estas chalconas puede plantearse mediante diferentes metodologías descritas en la literatura; sin embargo, una de las más comunes y versátiles corresponde a la condensación de Claisen-Schmidt entre aldehídos aromáticos y acetiles aromáticos, generalmente en medio básico. Esta transformación ha sido ampliamente utilizada debido a la accesibilidad de los precursores, la simplicidad operacional y la posibilidad de generar bibliotecas de chalconas con distintos patrones de sustitución. Además, también se han descrito variantes asistidas por microondas y otras metodologías más modernas orientadas a reducir tiempos de reacción y mejorar rendimientos¹²⁰.

Una vez obtenidas las chalconas, las distintas familias heterocíclicas pueden derivarse mediante ciclocondensaciones con nucleófilos o binucleófilos apropiados. Así, según la familia considerada, pueden emplearse hidrazinas, guanidinas, diaminas aromáticas u otros socios nucleofílicos para transformar el sistema α,β -insaturado en el heterociclo final. Estas transformaciones también han sido descritas bajo múltiples condiciones, incluyendo medios básicos, medios ácidos suaves, calentamiento convencional y activación por microondas, dependiendo de la reactividad del nucleófilo y del tipo de anillo que se desea formar. En consecuencia, la estrategia sintética global puede considerarse modular, convergente y relativamente corta, ya que permite acceder a múltiples familias estructurales a partir de una misma lógica de construcción molecular.

Las rutas propuestas convergen en una lógica mecanística común: la formación inicial de una chalcona como intermediario clave, seguida de transformaciones de adición nucleofílica y ciclación que conducen a los distintos heterociclos obtenidos. Esta estrategia permite racionalizar de manera unificada familias estructuralmente diversas, ya que los derivados de una misma serie comparten secuencias mecanísticas análogas y difieren principalmente en la identidad de sus sustituyentes. En consecuencia, para evitar repeticiones innecesarias, esta tesis se limita a discutir los principios mecanísticos generales (Anexo VI), mientras que los mecanismos completos y representativos de cada transformación se presentan en el. Las chalconas corresponden a sistemas α,β -insaturados del tipo 1,3-diaril-2-propen-1-ona, y justamente esa combinación entre carbonilo conjugado y doble enlace explica su valor como plataforma sintética y como precursor versátil de múltiples heterociclos.

En particular, la generación de la chalcona procede mediante la formación del enolato del compuesto carbonílico, su adición al aldehído y la obtención posterior del correspondiente β -hidroxicarbonilo. Aunque esta conversión suele resumirse de manera global como una “deshidratación”, la evidencia experimental disponible apoya que, en medio básico, dicha etapa se describe mejor mediante una vía E1cB que como una E2 concertada. En este escenario, la base abstrae primero el protón en α para generar un intermedio aniónico estabilizado por resonancia, y posteriormente ocurre la eliminación de hidróxido con formación del doble enlace conjugado. Esta formulación resulta mecanísticamente más razonable, ya que el grupo hidroxilo constituye un mal grupo saliente bajo condiciones básicas, mientras que la formación previa del enolato y la estabilización del producto α,β -insaturado favorecen una eliminación

escalonada. De hecho, Perrin y Chang¹²¹ concluyeron que, en la condensación aldólica base-catalizada que conduce a chalconas, la etapa determinante corresponde a la pérdida final de hidróxido y formación del enlace C=C; posteriormente, Perrin y Kim observaron también en agua resultados compatibles con que la etapa lenta es la eliminación final desde el intermedio enolato. Por ello, este proceso se representa en el Anexo VI como una eliminación de tipo E1cB.

De manera complementaria, también podría proponerse una ruta alternativa en ciertos casos mediante la conversión del sistema enona en un cetoalquino (ynona), ya que las ynonas son intermedios sintéticamente valiosos y reconocidos como plataformas reactivas en síntesis heterocíclica. Desde un punto de vista estratégico, este acceso alternativo podría concebirse a partir de una funcionalización adicional de la chalcona, seguida de una eliminación adecuada para generar el sistema acetilénico conjugado. Aunque esta posibilidad se presenta aquí como una alternativa mecanística y sintética plausible más que como la ruta principal del trabajo, su consideración resulta útil porque amplía el marco de transformación disponible a partir del mismo precursor común. Las ynonas, de hecho, son reconocidas como bloques de construcción relevantes en síntesis orgánica avanzada¹²².

Partiendo de una plataforma común(chalconas) , las transformaciones posteriores pueden entenderse como ciclocondensaciones sobre el sistema α,β-insaturado de la chalcona. En términos generales, el carácter biselectrofílico de este motivo permite la adición de nucleófilos o binucleófilos sobre el carbonilo conjugado y/o sobre el carbono β, seguida por cierre intramolecular de anillo y, en aquellas familias en que corresponde, por etapas adicionales de deshidratación, tautomerización y/o aromatización que conducen al heterociclo final. Por esta razón, aunque cada serie presenta particularidades estructurales, todas responden a una lógica global de adición nucleofílica/ciclación a partir de la chalcona. Los detalles mecanísticos específicos de estas ciclocondensaciones, así como los esquemas generales representativos para cada familia, se incluyen en el Anexo VI, dado que los distintos derivados siguen secuencias análogas y difieren principalmente en la naturaleza de sus sustituyentes y en la identidad del nucleófilo empleado. La versatilidad de las chalconas como precursores de pirroles, benzodiazepinas y pirimidinas está ampliamente documentada¹²³.

Un criterio adicional para racionalizar estas ciclaciones corresponde a las reglas de Baldwin¹²⁴, empleadas aquí solo como una guía sencilla de conveniencia estereo-electrónica para el paso de cierre de anillo. En términos generales, los cierres exo-trig suelen ser los más favorables. Bajo este marco, la ciclación de las hidrazinas sobre la chalcona puede describirse como un cierre 5-exo-trig, por lo que resulta favorable; la ciclación con guanidina hacia núcleos pirimidínicos puede interpretarse como un cierre 6-exo-trig, igualmente favorable; la reacción con o-fenilendiamina hacia benzodiazepinas corresponde a un cierre 7-exo-trig, viable aunque menos directo que los de cinco y seis miembros; y la ciclación con 2-aminobenzimidazol hacia sistemas pirimido[1,2-a]benzimidazólicos puede entenderse como un cierre 6-exo-trig, también favorable. En todos estos casos, la clasificación según Baldwin no reemplaza el mecanismo detallado, pero sí permite justificar de forma simple que el paso de ciclación intramolecular es geométricamente razonable. Esta lectura es consistente con la amplia precedencia sintética de pirazolinas a partir de chalconas e hidrazinas, de pirimidinas a partir de chalconas y guanidina, de benzodiazepinas a partir de chalconas y o-fenilendiamina, y de pirimido[1,2-a]benzimidazoles a partir de chalconas y 2-aminobenzimidazol(Tabla 18).

Tabla 18 Resumen de diferentes heterociclos y su evaluación

Nucleófilo	Heterociclo esperado	Tipo de cierre	Evaluación
Hidrazina	Pirazol	5 exo-trig	Favorable
Guanidina	Aminopirimidina	6- exo-trig	Favorable
o-Fenilendiamina	Benzodiazepina	7-exo-trig	Favorable
2-Aminobenzimidazol	Pirimido[1,2-a]benzimidazol	6- exo-trig	Favorable
Guanidino-benzimidazol	Pirimidina aminobenzimidazol	No aplica, cierre catalizado de Zn(L-prolina) ₂	Caso aparte; probable formación de producto vía Zn(L-prolina) ₂

Distinto es el caso del guanidino-benzimidazol. Aquí el problema principal no es que el cierre sea favorable por la regla de Baldwin, sino que en medio básico probablemente no conduciría de manera selectiva a la pirimidina-aminobenzimidazol deseado, ya que la presencia de varios centros nitrogenados con distinta basicidad y nucleofilia puede desviar el ataque hacia rutas alternativas y generar productos distintos a los contemplados en este estudio. Por ello, este sistema conviene discutirlo como un caso aparte, y no como una simple extensión del mecanismo general anterior. En términos prácticos, su formación requeriría condiciones más controladas o una activación catalítica específica; como precedente análogo, la ciclación de guanidinas con chalconas hacia pirimidinas ha sido descrita bajo catálisis de $\text{Zn}(\text{L-prolina})_2$ en medio acuoso¹²⁵. La proposición anterior para el guanidino-benzimidazol debe entenderse como una inferencia mecanística razonable Anexo VI, no como una demostración experimental directa para ese sustrato concreto.

Un aspecto relevante de estas ciclaciones es que, en algunos casos, la transformación puede transcurrir a través de un intermedio no completamente aromatizado, el cual puede presentar estabilidad suficiente como para persistir bajo determinadas condiciones de reacción. Dado que la aromatización implica una ganancia adicional de estabilización electrónica, la formación del heterociclo aromático final resultaría termodinámicamente favorecida; sin embargo, la conversión completa puede requerir energía adicional o condiciones oxidantes apropiadas. En consecuencia, no puede descartarse que, en ciertas reacciones, el intermedio dihidro o parcialmente saturado actúe como intermedio persistente o incluso como contaminante principal si la aromatización no progresa de manera eficiente. Este punto tiene además una consecuencia práctica importante en purificación, ya que la especie no aromatizada y el producto aromático final pueden exhibir polaridades y comportamientos cromatográficos próximos, dificultando su separación por TLC o por cromatografía en columna. Desde una perspectiva sintética, esta situación justifica considerar condiciones que favorezcan explícitamente la etapa de oxidación/aromatización, ya sea mediante tiempos de reacción más prolongados o mediante la incorporación controlada de agentes oxidantes.

La retrosíntesis planteada no solo resulta químicamente plausible, sino también coherente con una estrategia de síntesis racional orientada a generar diversidad estructural a partir de bloques accesibles y transformaciones conocidas. La combinación de chalconas fácilmente obtenibles con ciclaciones posteriores mediadas por nucleófilos apropiados ofrece una ruta sintética flexible para el acceso a las familias estudiadas, mientras que la consideración de posibles intermediarios no aromatizados y de sus implicancias prácticas en la purificación permite anticipar dificultades experimentales razonables y plantear eventuales ajustes en las condiciones de reacción.

6.6.2 Derivados de benzimidazoles

Desde una perspectiva retrosintética, los derivados de benzimidazol pueden racionalizarse a partir de dos desconexiones principales. La primera corresponde a la ruptura del enlace N-C(O) , lo que conduce directamente al benzimidazol libre y al correspondiente cloruro de ácido aromático como agente acilante. Esta etapa puede interpretarse como una N-acilación clásica, en la que una base fuerte, como hidruro de sodio, genera el anión benzimidazolato, el cual posteriormente reacciona con el cloruro de ácido para formar el derivado N-acilado. En términos prácticos, se trata de una transformación directa y sintéticamente sencilla, aunque normalmente requiere condiciones anhidras y atmósfera inerte, dada la sensibilidad de los reactivos involucrados¹²⁶. El mecanismo general propuesto para esta etapa de acilación se presenta en el Anexo VI.

La segunda desconexión corresponde a la formación del propio núcleo benzimidazólico, que puede trazarse retrosintéticamente hasta o-fenilendiamina y el correspondiente aldehído heteroaromático. Esta aproximación coincide con una de las rutas más clásicas y versátiles para la síntesis de benzimidazoles, basada en una condensación seguida de ciclación y aromatización. La literatura describe numerosas variantes para esta transformación, incluyendo procedimientos promovidos por ácido acético, protocolos en etanol y otras metodologías térmicas o catalíticas¹²⁶⁻¹²⁸, lo que confirma que se trata de una vía directa, accesible y ampliamente utilizada para la construcción de este heterociclo. El mecanismo general correspondiente a la formación del núcleo benzimidazólico también se incluye en el Anexo VI.

6.7 Avances sintéticos

6.7.1 Chalconas

Con respecto a los protocolos de síntesis, la etapa de estandarización se realizó inicialmente utilizando la chalcona derivada de 1-acetilpireno y p-nitrobenzaldehído. La selección de este par de sustratos no fue arbitraria, sino que respondió a criterios prácticos de factibilidad sintética. Por una parte, el p-nitrobenzaldehído corresponde a un aldehído de bajo costo y fácil acceso. Por otra, el fragmento acetilpireno había mostrado, en ensayos preliminares, mayores dificultades experimentales que otros acetilos aromáticos, manifestadas en menores rendimientos y en una manipulación más compleja. En consecuencia, su elección como sistema modelo permitió establecer condiciones suficientemente exigentes y, al mismo tiempo, representativas para luego extrapolarlas al resto de la serie. Aunque desde un punto de vista electrónico podría esperarse que el grupo nitro en posición para aumentara la electrofilia del aldehído y facilitara la condensación, en la práctica se observaron con frecuencia mezclas oscuras, formación de amalgamas negras y escasa obtención del producto, lo que hizo necesaria una estandarización cuidadosa de las condiciones de reacción.

A partir de ello, se determinó que la condensación en medio básico con hidróxido de sodio en etanol constituía una condición adecuada para la obtención de las chalconas, en concordancia con la utilidad ampliamente reconocida de la condensación de Claisen–Schmidt como ruta general para la síntesis de este tipo de sistemas α,β -insaturados. En algunos casos, especialmente con derivados de pireno, la solubilidad del sustrato en etanol frío resultó insuficiente, por lo que fue necesario incorporar hasta un 30% de acetonitrilo como cosolvente, lo que permitió mejorar de manera importante la homogeneidad del medio y favorecer el desarrollo de la reacción. Bajo las condiciones optimizadas, fue posible alcanzar rendimientos de hasta 91% para la formación de las chalconas.

Dado que estas chalconas constituyen intermedios sintéticos clave para las ciclaciones posteriores, fueron caracterizadas mediante RMN de ^1H y ^{13}C . En los espectros de ^1H RMN, una de las señales más diagnósticas correspondió a los protones vinílicos del sistema α,β -insaturado, observados como dobletes con una constante de acoplamiento vicinal elevada, característica de la configuración trans (E) del doble enlace de la chalcona. En los espectros de ^{13}C RMN, la señal del carbonilo conjugado constituyó otro elemento confirmatorio de la formación de estos intermediarios. En conjunto, la integración, el patrón de multiplicidad y la presencia del sistema enona conjugado confirmaron la obtención de las chalconas esperadas, dejando esta etapa resuelta para su posterior transformación hacia los heterociclos de interés.

6.7.2 2 Fenil 4 (4-trifluorometanofenil)Pirimido-benzimidazol

Con respecto a la síntesis del ciclado final hacia los derivados pirimido-benzimidazol, se seleccionó en primera instancia un sistema modelo que presentara ventajas tanto desde el punto de vista sintético como práctico. Esta elección respondió, por una parte, a la mayor accesibilidad de los precursores involucrados y, por otra, a la expectativa de una purificación más sencilla del producto final en comparación con otros sustratos evaluados preliminarmente. En particular, se priorizó una chalcona con un extremo fenólico/aromático simple y un sustituyente trifluorometilo, ya que esta combinación permitía explorar un sistema electrónicamente interesante, aunque al mismo tiempo representaba un desafío sintético relevante.

Sobre esta base, se realizaron ensayos de ciclación utilizando dicha chalcona bajo dos condiciones principales: en etanol a reflujo durante 24 h, y en dimetilformamida a 80 °C. Aunque ambas condiciones permitieron avanzar en la transformación, se observaron diferencias importantes en el comportamiento del sistema. En DMF, la reacción presentó mayores dificultades de trabajo posterior, principalmente debido a la formación de subproductos y a la complejidad de la extracción, lo que dificultó el aislamiento limpio del compuesto de interés. En cambio, en etanol fue posible obtener un material más manejable desde el punto de vista experimental, permitiendo su aislamiento mediante cristalización y, cuando fue necesario, mediante cromatografía. En este sentido, el uso de etanol surgió como la alternativa más conveniente para continuar con la optimización de esta familia, aun cuando el rendimiento del derivado trifluorometilado todavía requiere mejoras.

Estos resultados indican que la metodología en etanol constituye una base prometedora para la síntesis de esta serie, particularmente porque permitió acceder al producto ciclado bajo condiciones relativamente simples y con un trabajo experimental más limpio que en medios más polares como DMF. No obstante, la reacción aún requiere una etapa de optimización más fina, especialmente en lo relativo al

tiempo de reacción, rendimiento aislado y condiciones exactas de cristalización, con el fin de mejorar la reproducibilidad y extender el protocolo a otros sustratos de la misma familia.

Con respecto a la molécula final y su caracterización, los espectros de RMN de ^1H y ^{13}C permitieron corroborar la formación del sistema heterocíclico esperado, cuyos desplazamientos y asignaciones se resumen en la Tabla 19 y Tabla 20. Desde el punto de vista estructural, este análisis resultó especialmente relevante debido a la posibilidad de formación de regioisómeros, tal como se esquematiza en la Figura 188, dependiendo de cuál centro electrofílico de la chalcona participe en la etapa inicial de adición y en la ciclación posterior. Sin embargo, considerando la discusión mecanística desarrollada previamente en el apartado de retrosíntesis, la ruta más razonable corresponde a una adición conjugada de tipo 1,4, seguida de ciclación intramolecular, en lugar de una adición 1,2 directa sobre el carbonilo. En consecuencia, la estructura finalmente asignada se interpreta como la más coherente tanto con la lógica mecanística propuesta como con la evidencia espectroscópica obtenida.

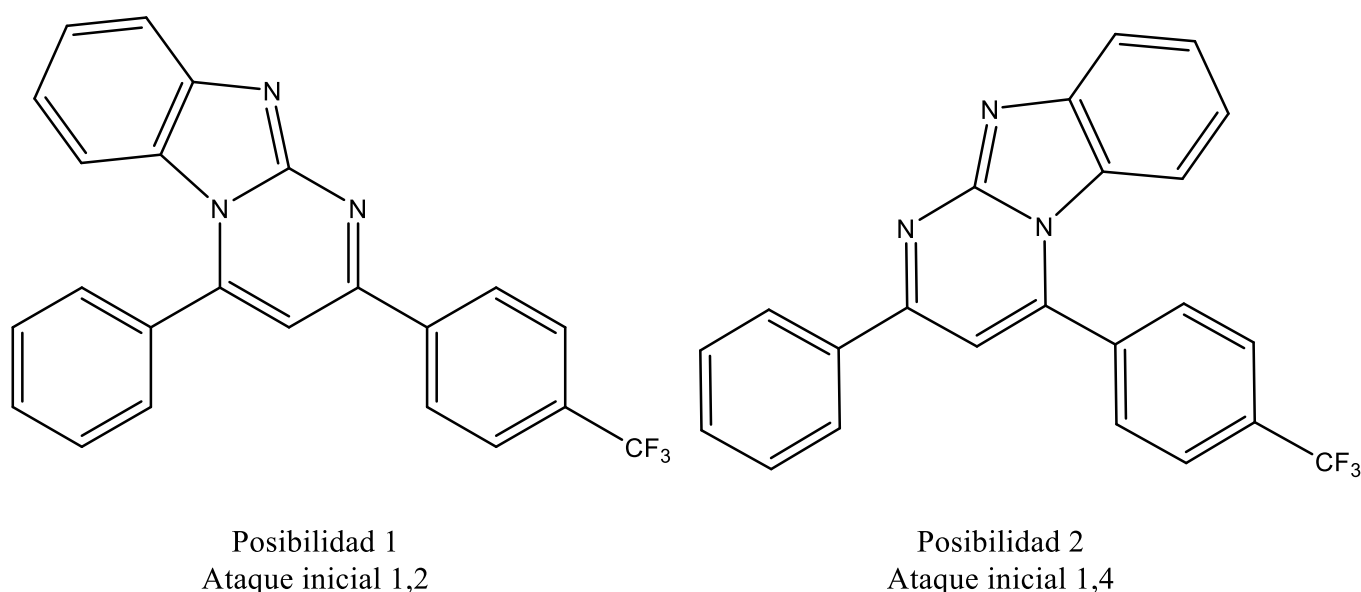


Figura 188 Posibilidades de ciclación para el Pirimido-benzimidazol.

6.7.2.1 Punto de fusión y infrarrojo

Los rangos obtenidos en la determinación del punto de fusión correspondieron en realidad a un proceso de descomposición, registrándose valores de 249.5–254.0 °C, 250.0–254.0 °C y 249.5–253.8 °C. La alta reproducibilidad entre las tres mediciones sugiere que el compuesto presenta un comportamiento térmico consistente, con inicio de descomposición cercano a 250 °C. Más que un punto de fusión nítido, estos resultados indican que la molécula sufre degradación al calentarse, lo que puede asociarse a la naturaleza rígida, altamente conjugada y heteroaromática de la estructura. En este contexto, el punto de fusión no debe interpretarse como una transición física simple, sino como una evidencia de estabilidad térmica hasta temperaturas relativamente elevadas, seguida de descomposición. Por otra parte, el espectro infrarrojo, tratado mediante sustracción de las señales del solvente, mostró bandas concordantes con la estructura propuesta. La señal observada a 1638 cm^{-1} se atribuyó a vibraciones $\text{C}=\text{N}/\text{C}=\text{C}$ conjugadas del sistema heteroaromático, mientras que las bandas a 1499 y 1454 cm^{-1} se asociaron a vibraciones $\text{C}=\text{C}$ aromáticas. Asimismo, las señales a 1379 , 1327 , 1133 y 1037 cm^{-1} fueron consistentes con deformaciones y estiramientos $\text{C}-\text{F}$ del grupo trifluorometilo. Las vibraciones de estiramiento $\text{C}-\text{H}$ aromático pueden contribuir en la región de $3100\text{--}3000\text{ cm}^{-1}$, aunque en este caso dicha zona no se consideró diagnóstica debido a posibles interferencias de línea base y del medio de análisis. En cambio, las señales remanentes en la región de $2250\text{--}2300\text{ cm}^{-1}$ se atribuyeron principalmente a acetonitrilo residual, mientras que las bandas anchas en la zona de $3200\text{--}3600\text{ cm}^{-1}$ probablemente se debieron a humedad o trazas de agua. En conjunto, la ausencia de señales atribuibles a grupos $\text{O}-\text{H}$, $\text{N}-\text{H}$ o $\text{C}=\text{O}$, junto con la presencia de las bandas características del sistema aromático conjugado y del grupo CF_3 , respalda la formación del compuesto esperado.

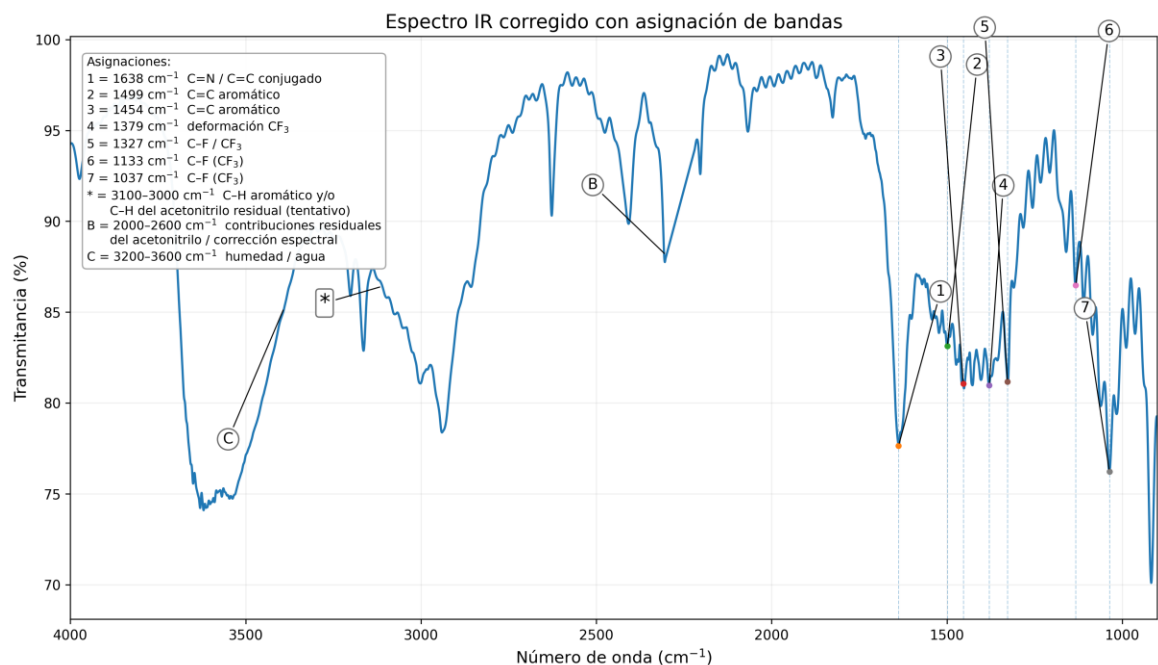


Figura 189 Espectro infrarrojo con sustracción de solvente (acetonitrilo) y asignaciones.

6.7.2.2Asignación ¹H RMN

Con respecto a la asignación de las señales observadas en el espectro de ¹H RMN(Figura 190), la señal que puede asignarse con mayor certeza es la **d**, correspondiente a un singlete, la cual se atribuye al protón del sistema heterocíclico condensado. Esta señal resulta diagnóstica debido a su patrón aislado y a que no presenta acoplamiento apreciable con protones vecinos.

Adicionalmente, las señales **c** y **h**, ambas con patrón de doblete e integración para ¹H, pueden asociarse tentativamente a protones aromáticos del fragmento pirimido-benzimidazol. Esta propuesta se sustenta en que dichos protones corresponden a posiciones no equivalentes dentro del sistema fusionado y, por tanto, pueden dar señales individuales de un protón con acoplamiento vecinal. En cambio, los protones de los anillos fenilo sustituyentes no se esperarían, en principio, como dobletes aislados de integración unitaria, sino como conjuntos de señales con integración mayor o con patrones más complejos debido al solapamiento.

Asimismo, en el espectro se observan señales con patrón de triplete aparente, cuyas constantes de acoplamiento son similares a las de los dobletes anteriormente descritos, por lo que también podrían corresponder a protones del mismo sistema aromático fusionado. Sin embargo, no puede descartarse completamente que alguna de estas señales provenga de protones del anillo fenilo unido al heterociclo, especialmente de posiciones sometidas a acoplamientos similares dentro de un entorno aromático conjugado

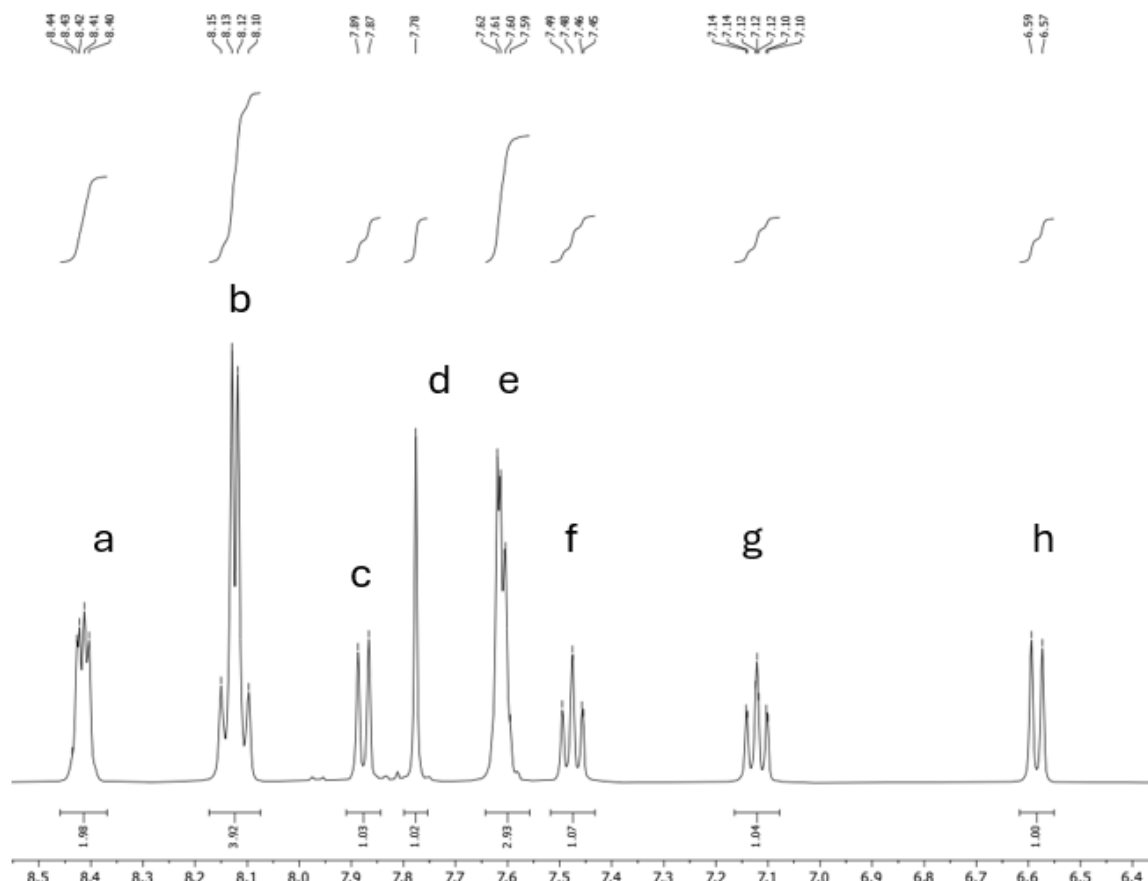


Figura 190 Espectro ^1H rmn con señales etiquetadas.

6.7.2.3Asignación ^{13}C RMN

Para la asignación de las señales de carbono también permite identificar ciertos elementos estructurales característicos, aunque el solapamiento de señales en la región aromática dificultó la asignación. La señal más fácilmente reconocible corresponde al carbono del grupo $-\text{CF}_3$, el cual aparece como un cuarteto debido al acoplamiento del núcleo de ^{13}C con los tres átomos equivalentes de flúor (^{19}F), constituyendo una evidencia clara de la incorporación de dicho sustituyente en la molécula. Además, las señales observadas a campo más bajo, aproximadamente entre 145 y 161 ppm, son consistentes con carbonos sp^2 del sistema heteroaromático enlazados a nitrógeno, compatibles con carbonos de tipo $\text{C}=\text{N}$ o con carbonos cuaternarios en un entorno electrónicamente deficiente. En cambio, la región comprendida aproximadamente entre 125 y 137 ppm concentra la mayor parte de las señales correspondientes a los carbonos aromáticos de los anillos fenilo y del sistema fusionado, lo que dificulta una asignación individual inequívoca a partir del espectro unidimensional. Se puede identificar rápidamente el carbono perteneciente al $-\text{CF}_3$, debido a que es un cuadruplete en la señal 13, más detalle en la Figura 183, esto ya que el flúor acopla con las señales de carbono.

Ya que todas las señales están dentro de la zona aromática, debido al solapamiento de señales en la región aromática, la asignación individual de cada señal no es trivial, por lo que fue necesario complementar el análisis con experimentos bidimensionales de RMN. Por lo que es necesario el uso de espectro 2D.

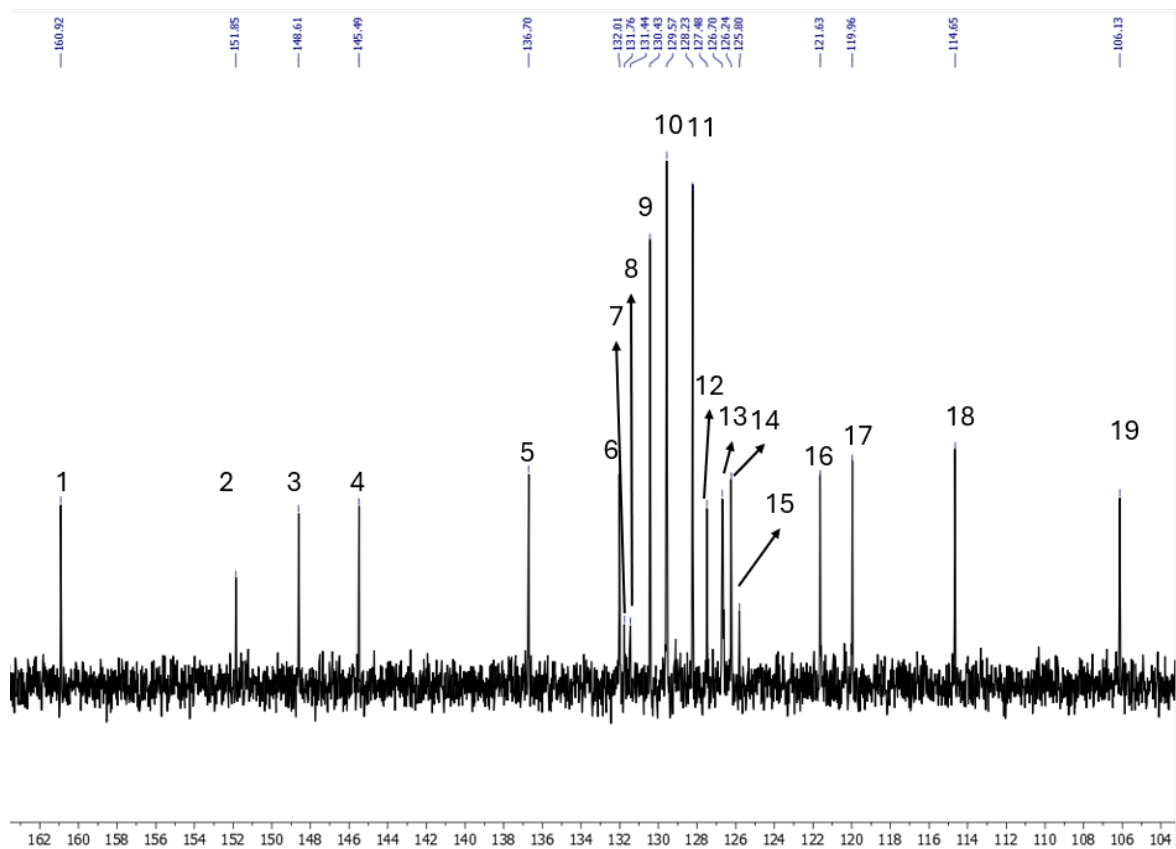


Figura 191 Espectro de ^{13}C con señales etiquetadas.

6.7.2.42D HMBC

En particular, las correlaciones a dos enlaces , se observadas para las señales c y g con el carbono 12(), así como las correlaciones de f y h con el carbono 4, sugieren que estos protones pertenecen a un mismo subfragmento aromático dentro del sistema pirimido-benzimidazol. El hecho de que dos pares de protones distintos muestren correlación hacia un mismo carbono resulta consistente con un sistema aromático fusionado, en el que varios protones pueden presentar correlaciones de dos o tres enlaces hacia los carbonos de fusión o hacia carbonos no protonados del heterociclo.

De manera especialmente relevante, la señal d, previamente propuesta como el singlete atribuible al protón del núcleo heterocíclico, presenta correlaciones con los carbonos 3 y 5. Este comportamiento es coherente con su ubicación dentro del sistema pirimido-benzimidazol, ya que desde esa posición se esperan correlaciones de largo alcance hacia los carbonos cuaternarios del anillo central, particularmente hacia aquellos involucrados en la unión con los anillos aromáticos vecinos. En este contexto, la correlación con el carbono 3, más desplazado a campo bajo, resulta razonable si dicho carbono corresponde a un centro sp^2 en un entorno empobrecido electrónicamente por la unión directa a un nitrógeno. De forma análoga, la correlación con el carbono 5 apoya que este protón se encuentra unido al anillo de benceno que contiene al CF_3 .

Las correlaciones adicionales, tales como a/6, c/16, g/17, h/14 y f/18, son también compatibles con la conectividad general del sistema aromático conjugado, aunque por sí solas no permiten asignar de manera inequívoca cada posición individual.

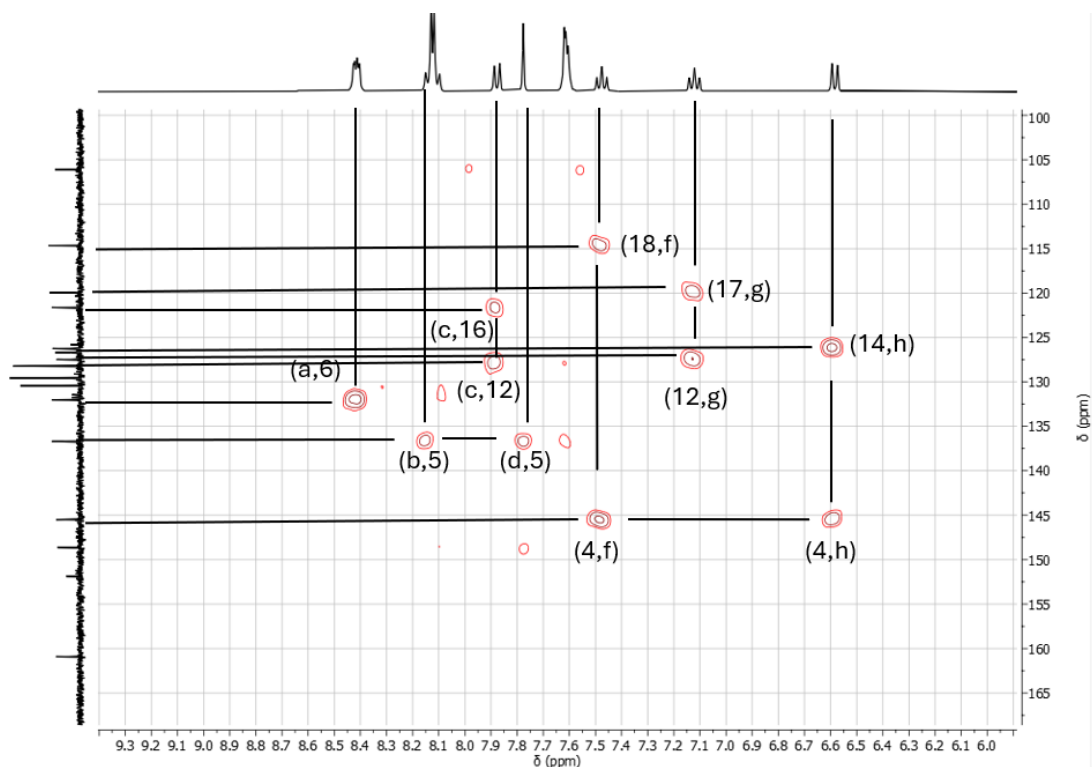


Figura 192 Espectro 2D HMBC con señales etiquetadas.

6.7.2.52D COSY

El espectro COSY (Figura 193) permitió reconocer dos sistemas principales de acoplamiento en la región aromática. Por una parte, las correlaciones sucesivas c–f, f–g y g–h indican la presencia de un conjunto de protones aromáticos vecinales que forman una secuencia continua de acoplamiento, compatible con los cuatro hidrógenos del anillo bencénico fusionado del fragmento benzimidazólico. Este comportamiento respalda la asignación de estas señales al sistema aromático condensado del núcleo heterocíclico. Por otra parte, la correlación a–e sugiere un segundo sistema de acoplamiento independiente, consistente con un anillo independiente, dado que integra para dos hidrógenos la señal a solo queda la opción que sea un doblete deformado. En contraste, la señal d no participa en esta red de acoplamientos vecinales, lo que resulta coherente con su asignación previa como el protón aislado del núcleo pirimidínico.

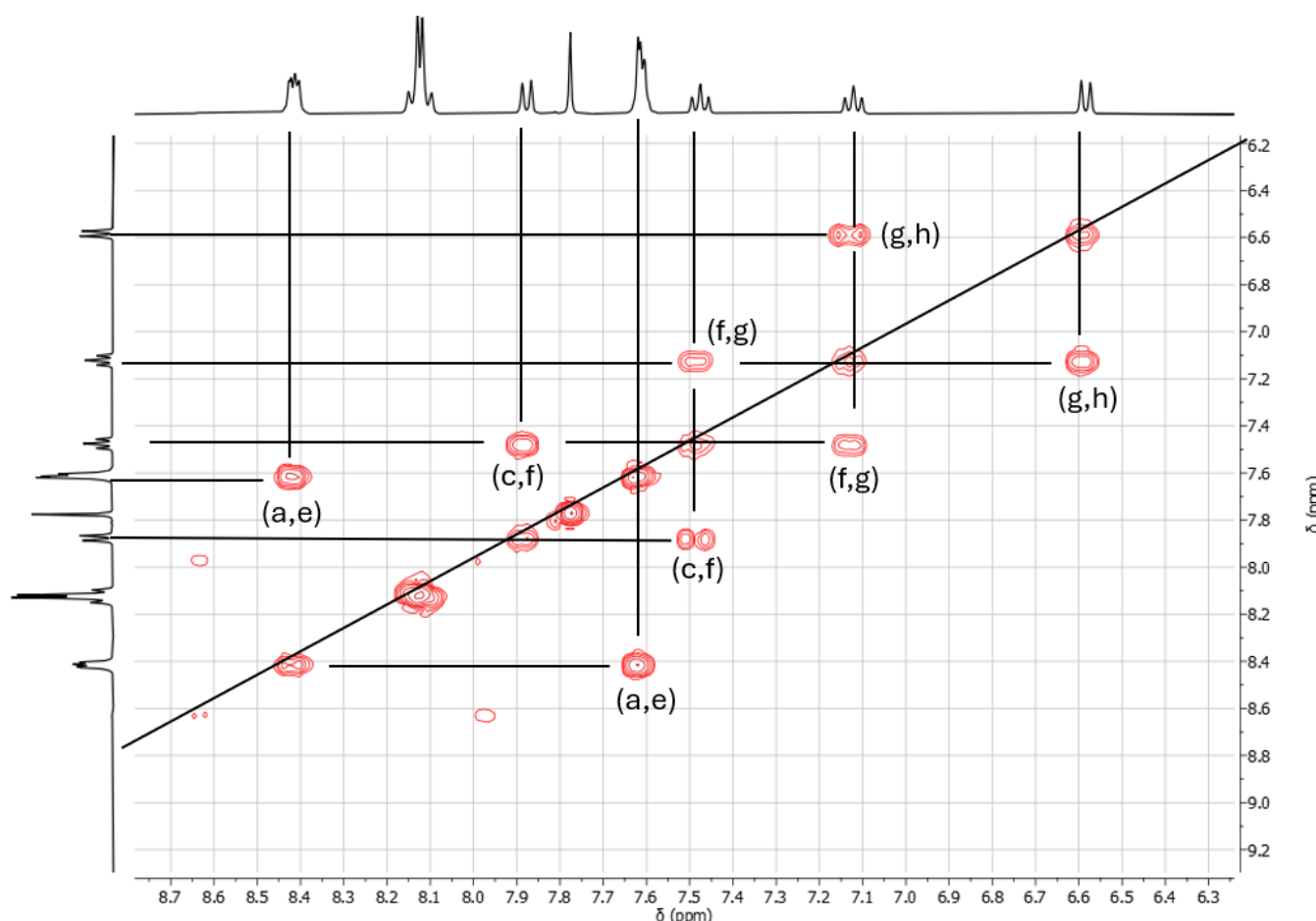


Figura 193 Espectro 2D COSY con señales etiquetadas.

6.7.2.62D NOESY

El espectro NOESY (Figura 194) aportó información decisiva para el análisis. En primer lugar, la señal d, observada como un singlete, se mantiene como la resonancia más diagnóstica del sistema, ya que su patrón indica que corresponde a un protón aislado dentro del núcleo heterocíclico pirimido-benzimidazol, sin protones vecinales directos que generen acoplamiento escalar apreciable. A partir de este punto, el COSY permite ordenar el resto de las señales aromáticas en distintos sistemas de acoplamiento. Por una parte, la secuencia c–f–g–h define claramente un conjunto continuo de protones vecinales, consistente con los cuatro hidrógenos del anillo bencénico fusionado del fragmento benzimidazólico. Esto permite separar ese subgrupo del resto de las resonancias aromáticas y concentrar la discusión de las señales a, b y e en los dos anillos fenilo sustituyentes.

Dentro de ese conjunto restante, la correlación a–e resulta especialmente informativa. Dado que la señal a integra para dos protones y se comporta como una señal de carácter doblete, esta puede asignarse razonablemente a los protones orto del anillo fenilo unido directamente al núcleo heterocíclico, es decir, a los hidrógenos más próximos al sistema pirimido-benzimidazol. Si a correlaciona con e, entonces e no puede interpretarse como una señal simple aislada, sino como una resonancia compuesta que contiene al menos una de las señales correspondientes a los protones siguientes de ese mismo anillo. En otras palabras, la correlación a–e indica que dentro de e se encuentra la señal de un protón adyacente al doblete a, lo que explica que esta región aparezca deformada o solapada y no como un patrón limpio y trivial. Esta observación permite concluir que la señal e contiene, al menos en parte, protones del anillo fenilo no sustituido o monosustituido unido al heterociclo, y no puede tratarse exclusivamente de un triplete puro separado del resto del sistema.

Una vez establecido eso, el sistema restante que integra globalmente para cuatro protones se vuelve mucho más restringido en su interpretación. Por descarte estructural y espectral, dicho conjunto corresponde de manera más razonable al anillo para-trifluorometilfenilo, el cual debería originar dos pares de protones equivalentes. En este contexto, la señal b puede entenderse como una señal compuesta correspondiente a esos protones aromáticos del anillo con CF₃, cuyos dobletes aparecen muy cercanos o parcialmente colapsados. Esta asignación no solo es coherente con la integración observada, sino también con la simetría propia de un anillo para-sustituido.

El elemento que termina de fortalecer esta interpretación proviene del experimento NOESY, ya que este no informa conectividad por enlace, sino proximidad espacial. La observación de correlaciones que involucran a la señal d con señales asignables a los anillos aromáticos periféricos demuestra que el protón singlete del núcleo heterocíclico se encuentra espacialmente cercano a esos fragmentos. En particular, la proximidad entre la señal del núcleo y la del anillo para-trifluorometilfenilo constituye una evidencia muy importante, porque indica que ambos fragmentos no se encuentran dispuestos en una geometría extendida y alejada, sino en una conformación en la que se aproximan en el espacio. Dicho de otro modo, el NOESY respalda una disposición plegada de la molécula. Esta proximidad espacial es altamente consistente con la regioisomería propuesta en Figura 188, posibilidad 1, y por lo tanto favorece la interpretación de que el proceso de ciclación ocurrió siguiendo la ruta que conduce a ese arreglo estructural, es decir, la asociada a una adición inicial 1,4, más que a la alternativa 1,2.

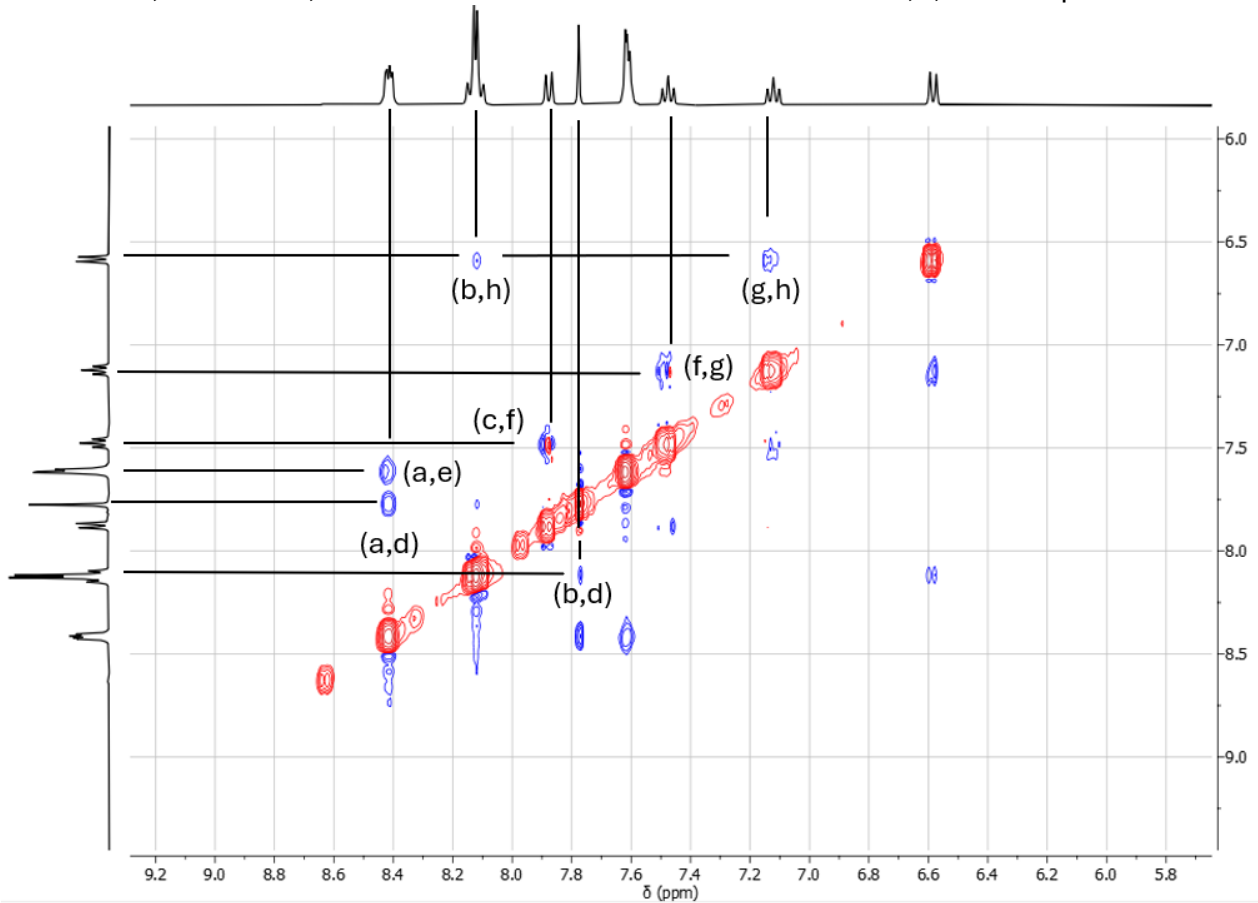


Figura 194 Espectro 2D NOESY con señales etiquetadas.

6.7.2.7Asignaciones finales

En la asignación final de cada señal de los espectros de 1H RMN y 13C RMN resulta consistente con la estructura propuesta, tal como se resume en la Tabla 19, la Tabla 20 y Figura 195, ya que las resonancias observadas calzan de manera coherente con la integración, la multiplicidad y las correlaciones obtenidas en los experimentos bidimensionales. En el caso del 1H RMN, los distintos sistemas aromáticos pudieron diferenciarse a partir del análisis combinado de los espectros 1D, COSY y NOESY, permitiendo sostener una asignación convergente de los protones del núcleo pirimido-benzimidazólico y de los anillos aromáticos sustituyentes, cuya distribución final se presenta en la Tabla 19. Por su parte, en el espectro de 13C RMN se observaron 19 señales, en concordancia con el número esperado de carbonos no equivalentes para la estructura propuesta, cuya asignación final se muestra en la Tabla 20. Aunque las señales 7 y 8 presentan menor intensidad y, en una primera inspección, podrían considerarse dudosas, su presencia resulta razonable al analizar el conjunto completo de la evidencia espectroscópica, especialmente por su compatibilidad con la distribución de desplazamientos químicos y con las correlaciones observadas en los experimentos 2D, por lo que es más apropiado considerarlas señales reales y no ruido.

Dentro de los experimentos bidimensionales, resumidos esquemáticamente en la Figura 195, el NOESY fue particularmente relevante, ya que no solo respaldó la disposición relativa de los protones del sistema aromático, sino que además entregó la evidencia más importante para la interpretación estructural global. En especial, la correlación b,h resultó decisiva, debido a que pone en evidencia una proximidad espacial entre un protón del sistema pirimido-benzimidazólico y un protón del anillo para-trifluorometilfenilo. Esta

interacción solo es compatible con una disposición en la que ambos fragmentos se encuentran orientados hacia el mismo sector del espacio, es decir, con una conformación plegada hacia ese lado de la molécula. En este sentido, dicha correlación respalda fuertemente la regioisomería propuesta y resulta altamente consistente con un producto originado por una adición inicial de tipo 1,4, más que con la alternativa 1,2. Por lo tanto, el conjunto de la evidencia espectroscópica, resumida en las Tabla 19 y Tabla 20, y en la Figura 195, no solo permite asignar de manera confiable las señales de protón y carbono, sino que además apoya de forma sólida la conectividad y la disposición espacial propuestas para el compuesto obtenido.

Tabla 19 Asignación ¹H RMN de 2 Fenil 4 (4-trifluorometanofenil)Pirimido-benzimidazol

Señal	Desplazamiento químico δ (ppm)	Integración	Multiplicidad	Constante de acoplamiento (Hz)	Asignación
a	8.46 – 8.37	2	Multiplete	-	H _a H _{a'}
b	8.17 – 8.07	4	Multiplete	-	H _{b1} H _{b1'} H _{b2} H _{b2'}
c	7.88	1	Doblete	8.2	H _c
d	7.78	1	Singlete	-	H _d
e	7.64 – 7.56	3	Multiplete	-	H _e H _{e'} H _{e2}
f	7.48	1	Triplete	7.5	H _f
g	7.12	1	Triplete	7.8	H _g
h	6.58	1	Doblete	8.4	H _h
		Σ H = 14			

Tabla 20 Asignación ^{13}C de 2 Fenil 4 (4-trifluorometanofenil)Pirimido-benzimidazol

Señal	Desplazamiento químico δ (ppm)	Asignación
1	160.92	Asignado por desplazamiento químico
2	151.85	-
3	148.61	Asignado por 2D HMBC
4	145.49	Asignado por 2D HMBC
5	136.70	Asignado por 2D HMBC
6	132.01	Asignado por 2D HMBC
7	131.44	-
8	131.76	-
9	131.44	-
10	130.43	-
11	129.57	-
12	128.23	Asignado por 2D HMBC
13	127.48	-
14	126.24	Asignado por 2D HMBC
15	125.80	-
16	121.63	Asignado por 2D HMBC
17	119.96	Asignado por 2D HMBC
18	114.65	Asignado por 2D HMBC
19	106.13	-

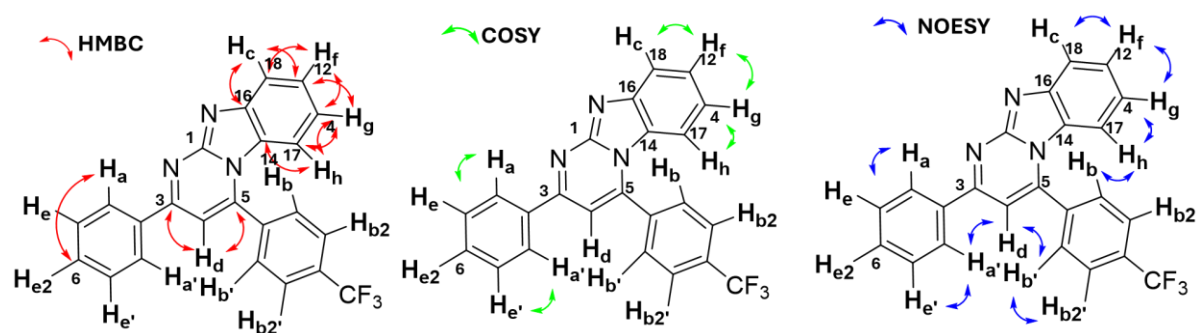


Figura 195 Resumen esquemático de las correlaciones más relevantes observadas en los experimentos bidimensionales HMBC, COSY y NOESY del compuesto pirimido-benzimidazólico trifluorometilado. Las correlaciones HMBC permitieron asignar carbonos cuaternarios y establecer conectividades de largo alcance; el COSY permitió definir los sistemas de acoplamiento entre protones aromáticos vecinales; y el NOESY aportó evidencia de proximidad espacial entre protones del núcleo heterocíclico y del anillo aromático sustituido, respaldando la asignación final de la estructura y una disposición molecular plegada, consistente con la regioisomería propuesta.

6.7.3 Benzimidazoles

Los resultados obtenidos permitieron confirmar preliminarmente la formación de tres benzimidazoles, cuyos espectros de RMN presentaron integraciones y señales de carbono consistentes con las estructuras propuestas. Una señal decisiva fue la observada alrededor de 13.0 ppm en los espectros de ^1H RMN, atribuida al protón N–H del anillo benzimidazólico. La presencia de esta señal permitió distinguir la

formación del núcleo benzimidazol y constituyó una evidencia diagnóstica relevante para esta etapa sintética.

Posteriormente, se llevó a cabo la reacción de acilación a partir de los benzimidazoles obtenidos. Si bien solo se dispuso de tres precursores benzimidazólicos, este resultado representó un avance importante, ya que demostró que las condiciones reportadas en la bibliografía utilizando hidruro de sodio fueron reproducibles bajo las condiciones empleadas. En los productos acilados, las señales observadas en los espectros de ^1H RMN concordaron con las integraciones esperadas para las estructuras propuestas. Además, la desaparición de la señal característica del N-H alrededor de 13 ppm sugirió que la acilación ocurrió sobre el nitrógeno del anillo benzimidazólico.

Esta asignación fue reforzada por los espectros de ^{13}C RMN, en los cuales se observó una señal alrededor de 169 ppm, atribuida al carbono carbonílico de la amida formada. La presencia de esta señal en las moléculas aciladas analizadas entregó evidencia adicional de la formación del enlace amida y, por tanto, de la obtención de los productos finales.

También se observó una tendencia importante en las reacciones de formación de benzimidazoles según la posición del nitrógeno en el anillo de piridina del aldehído utilizado. En general, las reacciones realizadas con 4-piridinacarboxaldehído entregaron los rendimientos más altos. De hecho, en la reacción efectuada en agua se observó la formación directa de cristales al enfriar la mezcla de reacción, lo que sugiere una conversión eficiente y una baja formación de subproductos. Este comportamiento podría atribuirse al efecto electrón-atractor del nitrógeno piridínico en posición 4, el cual favorece la electrofilia del grupo aldehído. Como consecuencia, se facilitaría el ataque nucleofílico de la o-fenilendiamina y la posterior formación del núcleo benzimidazólico.

Por otra parte, los derivados obtenidos a partir de 3-piridinacarboxaldehído presentaron rendimientos intermedios. En este caso, el nitrógeno piridínico aún puede ejercer un efecto electrónico sobre el sistema, aunque probablemente de manera menos eficiente que en el isómero 4-sustituido. Esto podría explicar que la reacción ocurriera de forma favorable, pero con una menor eficiencia global.

En contraste, las reacciones con 2-piridinacarboxaldehído entregaron los rendimientos más bajos. Además, estas reacciones tendieron a generar un residuo oscuro de aspecto aceitoso después de la extracción de la fase orgánica y la evaporación del solvente. Este resultado sugiere la formación de subproductos o mezclas complejas de reacción. Una posible explicación es que la proximidad del nitrógeno piridínico al grupo aldehído genere efectos estéricos o electrónicos desfavorables, dificultando el ataque nucleofílico de la o-fenilendiamina. Posiblemente el par de electrones no enlazante del nitrógeno piridínico podría participar en interacciones intramoleculares o estabilización de intermediarios no productivos, favoreciendo así reacciones colaterales.

Con respecto a los diferentes medios probados el reflujo en agua pareció favorecer la reacción en los aldehídos derivados de 3- y 4-piridinacarboxaldehído, probablemente debido al aumento de la temperatura del medio, lo que facilitaría la formación de la imina inicial y la posterior ciclización hacia el benzimidazol. Sin embargo, este mismo efecto fue desfavorable en el caso del 2-piridinacarboxaldehído, donde el aumento de temperatura podría haber promovido en mayor medida las rutas secundarias o procesos de degradación, en lugar de favorecer selectivamente la formación del producto deseado.

7 Conclusión

En la presente tesis doctoral se logró cumplir el objetivo general de identificar nuevos candidatos inhibidores de la proteasa tipo papaína (PLpro) de SARS-CoV-2 mediante una estrategia integrada de modelamiento QSAR, virtual screening, docking molecular y simulaciones de dinámica molecular. En una primera etapa, se construyó y depuró una base de datos de compuestos con actividad inhibitoria frente a PLpro, lo que permitió desarrollar y validar estadísticamente dos modelos QSAR 2D complementarios. En este contexto, el modelo manual aportó mayor interpretabilidad química para orientar el diseño racional de nuevas estructuras, mientras que el modelo evolutivo proporcionó un criterio predictivo más estricto para el filtrado computacional, demostrando que ambos enfoques no son excluyentes, sino metodológicamente complementarios.

A partir de estos modelos fue posible diseñar nuevas moléculas y explorar de manera amplia el espacio químico asociado a PLpro. El virtual screening abarcó un total de 3.312.167 moléculas distribuidas en siete familias químicas, confirmando la escalabilidad del flujo de trabajo y permitiendo identificar tendencias estructurales consistentes. Entre ellas, las benzodiazepinas destacaron como la familia con mejores puntajes de afinidad, seguidas por los derivados pirimido-benzimidazol. Esta convergencia entre los resultados de docking y los modelos QSAR refuerza la validez del enfoque empleado y su utilidad para priorizar compuestos dentro de un espacio químico amplio y diverso.

El refinamiento mediante induced fit docking y los análisis de dinámica molecular permitieron avanzar desde una priorización basada en afinidad hacia una evaluación mecanística más profunda de la estabilidad proteína-ligando y de la modulación conformacional de regiones funcionales de PLpro, particularmente del loop BL2. En este nivel de análisis, las benzodiazepinas sobresalieron como los sistemas con mejor desempeño mecanístico, al mostrar buena permanencia en el sitio de unión y una rigidización del BL2 comparable a la observada para GRL-0617. Sin embargo, la integración con los resultados farmacocinéticos y toxicológicos mostró que estas mismas familias presentan limitaciones importantes de distribución y seguridad, lo que impide priorizarlas únicamente en función de su desempeño dinámico. En contraste, dentro de la familia pirimido-benzimidazol, el derivado 2-naftilo-4-OH-3-OMe emergió como uno de los candidatos más equilibrados del estudio, al combinar estabilidad conformacional, modulación estructural del sitio de unión y un perfil ADME/Tox relativamente más favorable. En la serie benzimidazol, los compuestos 3 y 5 destacaron como los representantes más relevantes, aunque por atributos distintos: el primero por su mayor efecto sobre la dinámica del BL2 y el segundo por su mayor estabilidad en el sitio alostérico y semejanza conformacional con el ligando de referencia. Con respecto a las moléculas top 1 del virtual screening, las familias de benzamidas y pirimidina aminobenzimidazol destacaron como las más favorables. Sin embargo, al integrar los resultados ADME/Tox, las benzamidas presentan un perfil más equilibrado y desarrollable, principalmente por su menor carga toxicológica y un comportamiento metabólico más predecible. En contraste, la pirimidina aminobenzimidazol, aunque mantiene buenas propiedades de absorción y distribución, se ve limitada por una mayor complejidad metabólica, menor solubilidad y una mayor carga toxicológica. En paralelo el análisis retrosintético permitió establecer rutas de desconexión coherentes para las familias derivadas de chalcona y para los benzimidazoles heteroaromáticos, confirmando la viabilidad conceptual de los esqueletos propuestos. Experimentalmente, se demostró la factibilidad sintética del trabajo desarrollado, logrando la obtención de chalconas bajo condiciones optimizadas con rendimientos de hasta 91%, y la síntesis de un derivado representativo de tipo pirimido-benzimidazol trifluorometilado, cuya caracterización espectroscópica permitió respaldar la estructura propuesta y apoyar una ciclación consistente con una adición inicial 1,4. En conjunto, estos resultados confirman que la estrategia computacional no quedó restringida a una etapa teórica, sino que pudo vincularse con resultados sintéticos concretos. También se lograron sintetizar cuatro estructuras benzimidazólicas 2,N-disustituídas, obtenidas inicialmente a partir de o-fenilendiamina y distintos piridinacarboxaldehídos, y posteriormente funcionalizadas mediante acilación del N-H benzimidazólico con cloruros de acilo.

Esta tesis demuestra que la integración de QSAR, virtual screening, docking flexible, dinámica molecular, análisis ADME/Tox, retrosíntesis y avances sintéticos constituye una plataforma robusta para la identificación y priorización de nuevos candidatos frente a PLpro de SARS-CoV-2. Más que señalar una única familia “ganadora”, los resultados permiten distinguir entre compuestos con mejor desempeño mecanístico y compuestos con mejor equilibrio global entre afinidad, estabilidad y viabilidad farmacológica. De este modo, el trabajo no solo permitió proponer nuevas moléculas promisorias, sino también establecer criterios racionales de priorización para futuras etapas de optimización química y validación biológica.

8 Proyecciones

Los resultados obtenidos en esta tesis abren diversas líneas de continuación tanto en el ámbito sintético como en la validación biológica y en la expansión racional del diseño molecular. En primer lugar, una proyección directa de este trabajo corresponde a la síntesis de los compuestos mejor priorizados mediante el análisis integrado de QSAR, virtual screening, induced fit docking, dinámica molecular y evaluación farmacocinética/toxicológica. En este contexto, resulta especialmente relevante avanzar en la obtención de los derivados más prometedores, en particular aquellos que presentaron un perfil global más equilibrado entre afinidad teórica, estabilidad dinámica y propiedades ADME/Tox favorables. La síntesis de estas moléculas permitiría trasladar los hallazgos computacionales a una etapa experimental concreta y, con ello, validar de manera más robusta la estrategia de diseño implementada.

De manera complementaria, una segunda proyección importante consiste en profundizar la optimización de las rutas sintéticas desarrolladas, especialmente en la serie de benzimidazoles y en los sistemas piridina-benzimidazol. Si bien en esta tesis se logró demostrar la factibilidad de obtención de varios compuestos de interés, aún resulta necesario estandarizar las condiciones de reacción con el fin de mejorar la selectividad, reducir la formación de subproductos y maximizar los rendimientos. Esta optimización metodológica sería clave para consolidar un protocolo reproducible y escalable, capaz de facilitar la preparación de bibliotecas más amplias de derivados estructuralmente relacionados.

Asimismo, a partir de las tendencias observadas en los compuestos estudiados, surge como proyección natural la expansión racional de la serie de benzimidazoles mediante la incorporación de nuevos cloruros de acilo y otros fragmentos estructurales en posiciones estratégicas. Esta aproximación permitiría explorar con mayor profundidad el efecto de variaciones electrónicas, estéricas e hidrofóbicas sobre la actividad biológica predicha y sobre las propiedades farmacocinéticas de la serie. En este sentido, los resultados obtenidos en la presente tesis no solo permiten identificar compuestos promisorios, sino también extraer criterios de diseño útiles para orientar una segunda generación de análogos con perfiles más balanceados.

Una etapa posterior fundamental será la validación biológica experimental de los compuestos sintetizados frente a la proteasa PLpro de SARS-CoV-2. En este marco, el hecho de contar en el laboratorio con el kit EnzoLyte® 520 SARS-CoV-2 Papain-like Protease/Deubiquitinase Activity Assay Kit (Fluorimetric) constituye una ventaja significativa, ya que permitirá evaluar de manera directa la capacidad inhibitoria de las moléculas finales sobre el enzima blanco. Esta fase resulta indispensable para determinar si las tendencias observadas mediante las herramientas in silico se traducen efectivamente en inhibición enzimática real, y para establecer una primera correlación entre la priorización computacional y la respuesta experimental.

Una vez verificada la actividad inhibitoria sobre PLpro, será posible proyectar estudios de mayor complejidad biológica, incluyendo ensayos de citotoxicidad, evaluación de selectividad y, eventualmente, estudios celulares orientados a determinar el efecto de los compuestos sobre sistemas infectados o

modelos biológicos más cercanos al contexto fisiológico. Si bien estas etapas exceden el alcance de la presente tesis doctoral, representan una continuación lógica y necesaria para avanzar desde la identificación de hits hacia la validación de candidatos con potencial terapéutico.

Una proyección especialmente interesante deriva de la relación funcional y estructural existente entre PLpro y deubiquitininas humanas. Considerando que algunas de estas enzimas participan en procesos asociados a proliferación celular desregulada y cáncer, los compuestos identificados en esta investigación podrían constituir un punto de partida para explorar una nueva línea de trabajo orientada al desarrollo de inhibidores de deubiquitininas humanas. De este modo, las moléculas activas frente a PLpro no solo podrían ser optimizadas como antivirales, sino también servir como esqueletos químicos iniciales para una eventual especialización hacia aplicaciones oncológicas, ampliando de manera significativa el impacto potencial de esta investigación.

Las proyecciones muestran que la presente tesis no solo permitió identificar compuestos promisorios y establecer una metodología de priorización racional, sino también sentó bases concretas para futuras etapas de síntesis, validación experimental y expansión del diseño medicinal asistido por computación.

9 Referencias

- (1) COVID-19 Deaths | WHO COVID-19 Dashboard. <https://data.who.int/dashboards/covid19/cases>
- (2) Amin, Sk. A.; Banerjee, S.; Ghosh, K.; Gayen, S.; Jha, T. Protease Targeted COVID-19 Drug Discovery and Its Challenges: Insight into Viral Main Protease (Mpro) and Papain-like Protease (PLpro) Inhibitors. *Bioorg. Med. Chem.* 2021, 29, 115860. <https://doi.org/10.1016/j.bmc.2020.115860>.
- (3) Calleja, D. J.; Lessene, G.; Komander, D. Inhibitors of SARS-CoV-2 PLpro. *Front. Chem.* 2022, 10. <https://doi.org/10.3389/fchem.2022.876212>.
- (4) Advances in the Search for SARS-CoV-2 Mpro and PLpro Inhibitors. <https://www.mdpi.com/2076-0817/13/10/825> (accessed 2026-04-05).
- (5) Capasso, C.; Nocentini, A.; Supuran, C. T. Protease Inhibitors Targeting the Main Protease and Papain-like Protease of Coronaviruses. *Expert Opin. Ther. Pat.* 2021, 31 (4), 309–324. <https://doi.org/10.1080/13543776.2021.1857726>.
- (6) Gao, X.; Qin, B.; Chen, P.; Zhu, K.; Hou, P.; Wojdyla, J. A.; Wang, M.; Cui, S. Crystal Structure of SARS-CoV-2 Papain-like Protease. *Acta Pharm. Sin. B* 2021, 11 (1), 237–245. <https://doi.org/10.1016/j.apsb.2020.08.014>.
- (7) Safiabadi Tali, S. H.; LeBlanc, J. J.; Sadiq, Z.; Oyewunmi, O. D.; Camargo, C.; Nikpour, B.; Armanfard, N.; Sagan, S. M.; Jahanshahi-Anbuhi, S. Tools and Techniques for Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2)/COVID-19 Detection. *Clin. Microbiol. Rev.* 2021, 34 (3), 10.1128/cmr.00228-20. <https://doi.org/10.1128/cmr.00228-20>.
- (8) Mechanism and Inhibition of the Papain-like Protease, PLpro, of SARS-CoV-2 | The EMBO Journal | Springer Nature Link.
- (9) Magwaza, N. N.; Mushebenge, A. G.-A.; Ugbaja, S. C.; Mbatha, N. A.; Khan, R. B.; Kumalo, H. M. Mechanistic Insights into Targeting SARS-CoV-2 Papain-like Protease in the Evolution and Management of COVID-19. *BioChem* 2024, 4 (3), 268–299. <https://doi.org/10.3390/biochem4030014>.
- (10) The Complex Structure of GRL0617 and SARS-CoV-2 PLpro Reveals a Hot Spot for Antiviral Drug Discovery | Nature Communications.
- (11) Coronavirus Biology and Replication: Implications for SARS-CoV-2 | Nature Reviews Microbiology.
- (12) McClain, C. B.; Vabret, N. SARS-CoV-2: The Many Pros of Targeting PLpro. *Signal Transduct. Target. Ther.* 2020, 5, 223. <https://doi.org/10.1038/s41392-020-00335-z>.

- (13) Jeong, G. U.; Song, H.; Yoon, G. Y.; Kim, D.; Kwon, Y.-C. Therapeutic Strategies Against COVID-19 and Structural Characterization of SARS-CoV-2: A Review. *Front. Microbiol.* 2020, 11. <https://doi.org/10.3389/fmicb.2020.01723>.
- (14) Clemente, V.; D'Arcy, P.; Bazzaro, M. Deubiquitinating Enzymes in Coronaviruses and Possible Therapeutic Opportunities for COVID-19. *Int. J. Mol. Sci.* 2020, 21 (10), 3492. <https://doi.org/10.3390/ijms21103492>.
- (15) Sanders, B. C.; Pokhrel, S.; Labbe, A. D.; Mathews, I. I.; Cooper, C. J.; Davidson, R. B.; Phillips, G.; Weiss, K. L.; Zhang, Q.; O'Neill, H.; Kaur, M.; Schmidt, J. G.; Reichard, W.; Surendranathan, S.; Parvathareddy, J.; Phillips, L.; Rainville, C.; Sterner, D. E.; Kumaran, D.; Andi, B.; Babnigg, G.; Moriarty, N. W.; Adams, P. D.; Joachimiak, A.; Hurst, B. L.; Kumar, S.; Butt, T. R.; Jonsson, C. B.; Ferrins, L.; Wakatsuki, S.; Galanie, S.; Head, M. S.; Parks, J. M. Potent and Selective Covalent Inhibition of the Papain-like Protease from SARS-CoV-2. *Nat. Commun.* 2023, 14 (1), 1733. <https://doi.org/10.1038/s41467-023-37254-w>.
- (16) Identification of Novel Allosteric Sites of SARS-CoV-2 Papain-like Protease (PLpro) for the Development of COVID-19 Antivirals - Journal of Biological Chemistry.
- (17) Ying Huang, H.; Pinus, S.; Zhang, X. C.; Wang, G.; Broquière, M.; Boulon, R.; Rueda, A. M.; Souaibou, Y.; Huck, S.; Huot, M.; Vlaho, D.; Pottel, J.; Venegas, F.; Lu, Z.; Hennecker, C.; Stille, J.; Tjutrins, J.; Miron, C.; Labarre, A.; Plescia, J.; Burai-Patrascu, M.; Laplante, S.; Chatel-Chaix, L.; Mittermaier, A.; Moitessier, N. Integration of Computational and Experimental Techniques for the Discovery of SARS-CoV-2 PLpro Covalent Inhibitors. *ChemRxiv* 2025 (1223). <https://doi.org/10.26434/chemrxiv-2023-r0v7t-v2>.
- (18) Osipiuk, J.; Azizi, S.-A.; Dvorkin, S.; Endres, M.; Jedrzejczak, R.; Jones, K. A.; Kang, S.; Kathayat, R. S.; Kim, Y.; Lisnyak, V. G.; Maki, S. L.; Nicolaescu, V.; Taylor, C. A.; Tesar, C.; Zhang, Y.-A.; Zhou, Z.; Randall, G.; Michalska, K.; Snyder, S. A.; Dickinson, B. C.; Joachimiak, A. Structure of Papain-like Protease from SARS-CoV-2 and Its Complexes with Non-Covalent Inhibitors. *Nat. Commun.* 2021, 12 (1), 743. <https://doi.org/10.1038/s41467-021-21060-3>.
- (19) World Health Organization. WHO Policy Brief: COVID-19 Surveillance, 2024. <https://www.who.int/publications/m/item/who-policy-brief-covid-19-surveillance> (accessed 2026-04-05).
- (20) Simone Brogi. In Silico Methods for Drug Design and Discovery.
- (21) Villoutreix, B. O. Post-Pandemic Drug Discovery and Development: Facing Present and Future Challenges. *Front. Drug Discov.* 2021, 1. <https://doi.org/10.3389/fddsv.2021.728469>.
- (22) Mohamed, K.; Yazdanpanah, N.; Saghadzadeh, A.; Rezaei, N. Computational Drug Discovery and Repurposing for the Treatment of COVID-19: A Systematic Review. *Bioorganic Chem.* 2021, 106, 104490. <https://doi.org/10.1016/j.bioorg.2020.104490>.
- (23) Multi-Condition QSAR Model for the Virtual Design of Chemicals with Dual Pan-Antiviral and Anti-Cytokine Storm Profiles | ACS Omega.
- (24) QSAR Modeling of SARS-CoV Mpro Inhibitors Identifies Sufugolix, Cenicriviroc, Proglumetacin, and Other Drugs as Candidates for Repurposing against SARS-CoV-2 - Alves - 2021 - Molecular Informatics - Wiley Online Library.
- (25) Madani, A.; Benkortbi, O.; Laidi, M. In Silico Prediction of the Inhibition of New Molecules on SARS-CoV-2 3CL Protease by Using QSAR: PSOSVR Approach. *Braz. J. Chem. Eng.* 2023, 1–16. <https://doi.org/10.1007/s43153-023-00332-z>.
- (26) Wang, J. Fast Identification of Possible Drug Treatment of Coronavirus Disease-19 (COVID-19) through Computational Drug Repurposing Study. *J. Chem. Inf. Model.* 2020, 60 (6), 3277–3286. <https://doi.org/10.1021/acs.jcim.0c00179>.

- (27) Pandey, P.; Rane, J. S.; Chatterjee, A.; Kumar, A.; Khan, R.; Prakash, A.; Ray, S. Targeting SARS-CoV-2 Spike Protein of COVID-19 with Naturally Occurring Phytochemicals: An in Silico Study for Drug Development. *J. Biomol. Struct. Dyn.* 2021, 39 (16), 6306–6316. <https://doi.org/10.1080/07391102.2020.1796811>.
- (28) Pirololi, D.; Righino, B.; Camponeschi, C.; Ria, F.; Di Sante, G.; De Rosa, M. C. Virtual Screening and Molecular Dynamics Simulations Provide Insight into Repurposing Drugs against SARS-CoV-2 Variants Spike Protein/ACE2 Interface. *Sci. Rep.* 2023, 13, 1494. <https://doi.org/10.1038/s41598-023-28716-8>.
- (29) Application of Artificial Intelligence and Machine Learning for COVID-19 Drug Discovery and Vaccine Design - PubMed.
- (30) Arora, G.; Joshi, J.; Mandal, R. S.; Shrivastava, N.; Virmani, R.; Sethi, T. Artificial Intelligence in Surveillance, Diagnosis, Drug Discovery and Vaccine Development against COVID-19. *Pathogens* 2021, 10 (8), 1048. <https://doi.org/10.3390/pathogens10081048>.
- (31) Frontiers | Artificial Intelligence for COVID-19 Drug Discovery and Vaccine Development.
- (32) DiMasi, J. A.; Grabowski, H. G.; Hansen, R. W. Innovation in the Pharmaceutical Industry: New Estimates of R&D Costs. *J. Health Econ.* 2016, 47, 20–33. <https://doi.org/10.1016/j.jhealeco.2016.01.012>.
- (33) Lionta, E.; Spyrou, G.; Vassilatis, D. K.; Cournia, Z. Structure-Based Virtual Screening for Drug Discovery: Principles, Applications and Recent Advances. *Curr. Top. Med. Chem.* 2014, 14 (16), 1923–1938. <https://doi.org/10.2174/1568026614666140929124445>.
- (34) Schneider, G. Automating Drug Discovery. *Nat. Rev. Drug Discov.* 2018, 17 (2), 97–113. <https://doi.org/10.1038/nrd.2017.232>.
- (35) Tropsha, A. Best Practices for QSAR Model Development, Validation, and Exploitation. *Mol. Inform.* 2010. <https://doi.org/10.1002/minf.201000061>.
- (36) Cherkasov, A.; Muratov, E. N.; Fourches, D.; Varnek, A.; Baskin, I. I.; Cronin, M.; Dearden, J.; Gramatica, P.; Martin, Y. C.; Todeschini, R.; Consonni, V.; Kuz'min, V. E.; Cramer, R.; Benigni, R.; Yang, C.; Rathman, J.; Terfloth, L.; Gasteiger, J.; Richard, A.; Tropsha, A. QSAR Modeling: Where Have You Been? Where Are You Going To? *J. Med. Chem.* 2014, 57 (12), 4977–5010. <https://doi.org/10.1021/jm4004285>.
- (37) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings1. *Adv. Drug Deliv. Rev.* 2001, 46 (1), 3–26. [https://doi.org/10.1016/S0169-409X\(00\)00129-0](https://doi.org/10.1016/S0169-409X(00)00129-0).
- (38) Leeson, P. D.; Springthorpe, B. The Influence of Drug-like Concepts on Decision-Making in Medicinal Chemistry. *Nat. Rev. Drug Discov.* 2007, 6 (11), 881–890. <https://doi.org/10.1038/nrd2445>.
- (39) Hansch, C.; Rockwell, S. D.; Jow, P. Y. C.; Leo, A.; Steller, E. E. Substituent Constants for Correlation Analysis. *J. Med. Chem.* 1977, 20 (2), 304–306. <https://doi.org/10.1021/jm00212a024>.
- (40) Hansch, Corwin.; Leo, A.; Taft, R. W. A Survey of Hammett Substituent Constants and Resonance and Field Parameters. *Chem. Rev.* 1991, 91 (2), 165–195. <https://doi.org/10.1021/cr00002a004>.
- (41) London, F. The General Theory of Molecular Forces. *Trans. Faraday Soc.* 1937, 33 (0), 8b–26. <https://doi.org/10.1039/TF937330008B>.
- (42) Todeschini, R.; Consonni, V. *Molecular Descriptors for Chemoinformatics, Volumes I & II*.
- (43) Taft, R. W. Jr. Linear Free Energy Relationships from Rates of Esterification and Hydrolysis of Aliphatic and Ortho-Substituted Benzoate Esters. *J. Am. Chem. Soc.* 1952, 74 (11), 2729–2732. <https://doi.org/10.1021/ja01131a010>.
- (44) Charton, M. Steric Effects. I. Esterification and Acid-Catalyzed Hydrolysis of Esters. *J. Am. Chem. Soc.* 1975, 97 (6), 1552–1556. <https://doi.org/10.1021/ja00839a047>.

- (45) MOREAU G; BROTO P, G. THE AUTOCORRELATION OF A TOPOLOGICAL STRUCTURE: A NEW MOLECULAR DESCRIPTOR. *AUTOCORRELATION Topol. Struct. NEW Mol. Descr.* 1980.
- (46) U.S. Environmental Protection Agency. Molecular Descriptor Guide Version 1.0.2, 2012.
- (47) Moran, P. A. P. Notes on Continuous Stochastic Phenomena. *Biometrika* 1950, 37 (1/2), 17. <https://doi.org/10.2307/2332142>.
- (48) Geary, R. C. The Contiguity Ratio and Statistical Mapping. *Inc. Stat.* 1954, 5 (3), 115. <https://doi.org/10.2307/2986645>.
- (49) Roy, K.; Ghosh, G. QSTR with Extended Topochemical Atom Indices. 2. Fish Toxicity of Substituted Benzenes. *J. Chem. Inf. Comput. Sci.* 2004. <https://doi.org/10.1021/ci0342066>.
- (50) Hall, L. H.; Kier, L. B. Electrotological State Indices for Atom Types: A Novel Combination of Electronic, Topological, and Valence State Information. *J. Chem. Inf. Comput. Sci.* 1995, 35 (6), 1039–1045. <https://doi.org/10.1021/ci00028a014>.
- (51) Yap, C. W. PaDEL-Descriptor: An Open Source Software to Calculate Molecular Descriptors and Fingerprints. *J. Comput. Chem.* 2011, 32 (7), 1466–1474. <https://doi.org/10.1002/jcc.21707>.
- (52) *Understanding the Basics of QSAR for Applications in Pharmaceutical Sciences and Risk Assessment*; 2015.
- (53) Gramatica, P. Principles of QSAR Models Validation: Internal and External. *QSAR Comb. Sci.* 2007, 26 (5), 694–701. <https://doi.org/10.1002/qsar.200610151>.
- (54) Dastmalchi, S.; Hamzeh-Mivehroud, M.; Sokouti, B. *Quantitative Structure - Activity Relationship: A Practical Approach*; CRC Press, Taylor & Francis Group: Boca Raton, 2018.
- (55) Meng, X.-Y.; Zhang, H.-X.; Mezei, M.; Cui, M. Molecular Docking: A Powerful Approach for Structure-Based Drug Discovery. *Curr. Comput. Aided Drug Des.* 2011, 7 (2), 146–157. <https://doi.org/10.2174/157340911795677602>.
- (56) Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T. Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein–Ligand Complexes. *J. Med. Chem.* 2006, 49 (21), 6177–6196. <https://doi.org/10.1021/jm051256o>.
- (57) Gohlke, H.; Klebe, G. Approaches to the Description and Prediction of the Binding Affinity of Small-Molecule Ligands to Macromolecular Receptors. *Angew. Chem. Int. Ed.* 2002, 41 (15), 2644–2676. [https://doi.org/10.1002/1521-3773\(20020802\)41:15%3C2644::AID-ANIE2644%3E3.0.CO;2-O](https://doi.org/10.1002/1521-3773(20020802)41:15%3C2644::AID-ANIE2644%3E3.0.CO;2-O).
- (58) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function. *J. Comput. Chem.* 1998, 19 (14), 1639–1662. [https://doi.org/10.1002/\(SICI\)1096-987X\(19981115\)19:14%3C1639::AID-JCC10%3E3.0.CO;2-B](https://doi.org/10.1002/(SICI)1096-987X(19981115)19:14%3C1639::AID-JCC10%3E3.0.CO;2-B).
- (59) Sherman, W.; Day, T.; Jacobson, M. P.; Friesner, R. A.; Farid, R. Novel Procedure for Modeling Ligand/Receptor Induced Fit Effects. *J. Med. Chem.* 2006, 49 (2), 534–553. <https://doi.org/10.1021/jm050540c>.
- (60) Clark, D. E. What Has Virtual Screening Ever Done for Drug Discovery? *Expert Opin. Drug Discov.* 2008, 3 (8), 841–851. <https://doi.org/10.1517/17460441.3.8.841>.
- (61) O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: An Open Chemical Toolbox. *J. Cheminformatics* 2011, 3 (1), 33. <https://doi.org/10.1186/1758-2946-3-33>.
- (62) avinashkarkada. Avinashkarkada/Pdbqt_maker, 2024. https://github.com/avinashkarkada/pdbqt_maker (accessed 2026-04-05).
- (63) Santos-Martins, D.; Solis-Vasquez, L.; Tillack, A. F.; Sanner, M. F.; Koch, A.; Forli, S. Accelerating AutoDock4 with GPUs and Gradient-Based Local Search. *J. Chem. Theory Comput.* 2021, 17 (2), 1060–1073. <https://doi.org/10.1021/acs.jctc.0c01006>.

- (64) Cramer, C. J. *Essentials of Computational Chemistry: Theories and Models*, 2nd ed.; Wiley: Chichester, West Sussex, England ; Hoboken, NJ, 2004.
- (65) Frenkel, D.; Smit, B. *Understanding Molecular Simulation : From Algorithms to Applications. 2nd Ed*; 1996; Vol. 50. <https://doi.org/10.1063/1.881812>.
- (66) Nam, K. H. Molecular Dynamics—From Small Molecules to Macromolecules. *Int. J. Mol. Sci.* 2021, 22 (7), 3761. <https://doi.org/10.3390/ijms22073761>.
- (67) Lu, C.; Wu, C.; Ghoreishi, D.; Chen, W.; Wang, L.; Damm, W.; Ross, G. A.; Dahlgren, M. K.; Russell, E.; Von Bargen, C. D.; Abel, R.; Friesner, R. A.; Harder, E. D. OPLS4: Improving Force Field Accuracy on Challenging Regimes of Chemical Space. *J. Chem. Theory Comput.* 2021, 17 (7), 4291–4300. <https://doi.org/10.1021/acs.jctc.1c00302>.
- (68) Ballante, F. *Protein-Ligand Interactions and Drug Design*; Methods in molecular biology, 2266; Humana Press: New York, NY, 2021.
- (69) Posansee, K.; Khunrae, P.; Sutthibutpong, T. A Molecular Dynamics Study to Assess the Positive Ion Distribution and the Effects of Protonation on the N-Terminus Region of a Xylanase. *J. Phys. Conf. Ser.* 2018, 1144 (1), 012011. <https://doi.org/10.1088/1742-6596/1144/1/012011>.
- (70) Ferreira, G. M.; Pillaiyar, T.; Hirata, M. H.; Poso, A.; Kronenberger, T. Inhibitor Induced Conformational Changes in SARS-COV-2 Papain-like Protease. *Sci. Rep.* 2022, 12 (1), 11585. <https://doi.org/10.1038/s41598-022-15181-y>.
- (71) Ferreira, L. L. G.; Andricopulo, A. D. ADMET Modeling Approaches in Drug Discovery. *Drug Discov. Today* 2019, 24 (5), 1157–1165. <https://doi.org/10.1016/j.drudis.2019.03.015>.
- (72) Veber, D. F.; Johnson, S. R.; Cheng, H.-Y.; Smith, B. R.; Ward, K. W.; Kopple, K. D. Molecular Properties That Influence the Oral Bioavailability of Drug Candidates. *J. Med. Chem.* 2002, 45 (12), 2615–2623. <https://doi.org/10.1021/jm020017n>.
- (73) Miyake, T.; Tsutsui, H.; Haraya, K.; Tachibana, T.; Morimoto, K.; Takehara, S.; Ayabe, M.; Kobayashi, K.; Kazuki, Y. Quantitative Prediction of P-Glycoprotein-Mediated Drug–Drug Interactions and Intestinal Absorption Using Humanized Mice. *Br. J. Pharmacol.* 2021, 178 (21), 4335–4351. <https://doi.org/10.1111/bph.15612>.
- (74) Korjamo, T.; Honkakoski, P.; Toppinen, M.-R.; Niva, S.; Reinisalo, M.; Palmgrén, J. J.; Mönkkönen, J. Absorption Properties and P-Glycoprotein Activity of Modified Caco-2 Cell Lines. *Eur. J. Pharm. Sci. Off. J. Eur. Fed. Pharm. Sci.* 2005, 26 (3–4), 266–279. <https://doi.org/10.1016/j.ejps.2005.06.004>.
- (75) Neumaier, F.; Zlatopolskiy, B. D.; Neumaier, B. Drug Penetration into the Central Nervous System: Pharmacokinetic Concepts and In Vitro Model Systems. *Pharmaceutics* 2021, 13 (10), 1542. <https://doi.org/10.3390/pharmaceutics13101542>.
- (76) Pore, S.; Roy, K. Insights into Pharmacokinetic Properties for Exposure Chemicals: Predictive Modelling of Human Plasma Fraction Unbound (f_u) and Hepatocyte Intrinsic Clearance (Cl_{int}) Data Using Machine Learning. *Digit. Discov.* 2024, 3 (9), 1852–1877. <https://doi.org/10.1039/D4DD00082J>.
- (77) Limaye, P. B.; Parikh, K. Distribution - In Vitro Test: Protein Binding. In *Drug Discovery and Evaluation: Safety and Pharmacokinetic Assays*; Springer, Cham, 2024; pp 1–14. https://doi.org/10.1007/978-3-030-73317-9_101-1.
- (78) Zhao, M.; Ma, J.; Li, M.; Zhang, Y.; Jiang, B.; Zhao, X.; Huai, C.; Shen, L.; Zhang, N.; He, L.; Qin, S. Cytochrome P450 Enzymes and Drug Metabolism in Humans. *Int. J. Mol. Sci.* 2021, 22 (23), 12808. <https://doi.org/10.3390/ijms222312808>.
- (79) Korzekwa, K.; Nagar, S. Process and System Clearances in Pharmacokinetic Models: Our Basic Clearance Concepts Are Correct. *Drug Metab. Dispos.* 2023, 51 (4), 532–542. <https://doi.org/10.1124/dmd.122.001060>.

- (80) Vijay, U.; Gupta, S.; Mathur, P.; Suravajhala, P.; Bhatnagar, P. Microbial Mutagenicity Assay: Ames Test. *Bio-Protoc.* 2018, 8 (6), e2763. <https://doi.org/10.21769/BioProtoc.2763>.
- (81) Mortelmans, K.; Zeiger, E. The Ames Salmonella/Microsome Mutagenicity Assay. *Mutat. Res.* 2000, 455 (1–2), 29–60. [https://doi.org/10.1016/s0027-5107\(00\)00064-6](https://doi.org/10.1016/s0027-5107(00)00064-6).
- (82) Wang, Y.; Xie, W. Drug-Induced Liver Injury: An Overview and Update. *Gastroenterol. Endosc.* 2023, 1 (2), 102–109. <https://doi.org/10.1016/j.gande.2022.11.005>.
- (83) Hosack, T.; Damry, D.; Biswas, S. Drug-Induced Liver Injury: A Comprehensive Review. *Ther. Adv. Gastroenterol.* 2023, 16, 17562848231163410. <https://doi.org/10.1177/17562848231163410>.
- (84) Lagrutta, A. A.; Trepakova, E. S.; Salata, J. J. The hERG Channel and Risk of Drug-Acquired Cardiac Arrhythmia: An Overview. *Curr. Top. Med. Chem.* 2008, 8 (13), 1102–1112. <https://doi.org/10.2174/156802608785700016>.
- (85) Recanatini, M.; Poluzzi, E.; Masetti, M.; Cavalli, A.; De Ponti, F. QT Prolongation through hERG K(+) Channel Blockade: Current Knowledge and Strategies for the Early Prediction during Drug Development. *Med. Res. Rev.* 2005, 25 (2), 133–166. <https://doi.org/10.1002/med.20019>.
- (86) Fu, L.; Shi, S.; Yi, J.; Wang, N.; He, Y.; Wu, Z.; Peng, J.; Deng, Y.; Wang, W.; Wu, C.; Lyu, A.; Zeng, X.; Zhao, W.; Hou, T.; Cao, D. ADMETlab 3.0: An Updated Comprehensive Online ADMET Prediction Platform Enhanced with Broader Coverage, Improved Performance, API Functionality and Decision Support. *Nucleic Acids Res.* 2024, 52 (W1), W422–W431. <https://doi.org/10.1093/nar/gkae236>.
- (87) *RPMI-8226 | Cell lines | DSMZCellDive*. <https://celldive.dsmz.de/celllines/RPMI-8226> (accessed 2026-04-06).
- (88) Ling, S. C. W.; Lau, E. K. K.; Al-Shabeeb, A.; Nikolic, A.; Catalano, A.; Iland, H.; Horvath, N.; Ho, P. J.; Harrison, S.; Fleming, S.; Joshua, D. E.; Allen, J. D. Response of Myeloma to the Proteasome Inhibitor Bortezomib Is Correlated with the Unfolded Protein Response Regulator XBP-1. *Haematologica* 2012, 97 (1), 64–72. <https://doi.org/10.3324/haematol.2011.043331>.
- (89) Shen, Z.; Ratia, K.; Cooper, L.; Kong, D.; Lee, H.; Kwon, Y.; Li, Y.; Alqarni, S.; Huang, F.; Dubrovskiy, O.; Rong, L.; Thatcher, G. R. J.; Xiong, R. Design of SARS-CoV-2 PLpro Inhibitors for COVID-19 Antiviral Therapy Leveraging Binding Cooperativity. *J. Med. Chem.* 2022, 65 (4), 2940–2955. <https://doi.org/10.1021/acs.jmedchem.1c01307>.
- (90) Ma, C.; Sacco, M. D.; Xia, Z.; Lambrinidis, G.; Townsend, J. A.; Hu, Y.; Meng, X.; Szeto, T.; Ba, M.; Zhang, X.; Gongora, M.; Zhang, F.; Marty, M. T.; Xiang, Y.; Kolocouris, A.; Chen, Y.; Wang, J. Discovery of SARS-CoV-2 Papain-like Protease Inhibitors through a Combination of High-Throughput Screening and a FlipGFP-Based Reporter Assay. *ACS Cent. Sci.* 2021, 7 (7), 1245–1260. <https://doi.org/10.1021/acscentsci.1c00519>.
- (91) Meewan, I.; Kattoula, J.; Kattoula, J. Y.; Skinner, D.; Fajtová, P.; Giardini, M. A.; Woodworth, B.; McKerrow, J. H.; Lage de Siqueira-Neto, J.; O'Donoghue, A. J.; Abagyan, R. Discovery of Triple Inhibitors of Both SARS-CoV-2 Proteases and Human Cathepsin L. *Pharmaceuticals* 2022, 15 (6), 744. <https://doi.org/10.3390/ph15060744>.
- (92) Xu, H.; Li, J.; Song, S.; Xiao, Z.; Chen, X.; Huang, B.; Sun, M.; Su, G.; Zhou, D.; Wang, G.; Hao, R.; Wang, N. Effective Inhibition of Coronavirus Replication by Polygonum Cuspidatum. *Front. Biosci. Landmark Ed.* 2021, 26 (10), 789–798. <https://doi.org/10.52586/4988>.
- (93) Zhao, Y.; Du, X.; Duan, Y.; Pan, X.; Sun, Y.; You, T.; Han, L.; Jin, Z.; Shang, W.; Yu, J.; Guo, H.; Liu, Q.; Wu, Y.; Peng, C.; Wang, J.; Zhu, C.; Yang, X.; Yang, K.; Lei, Y.; Guddat, L. W.; Xu, W.; Xiao, G.; Sun, L.; Zhang, L.; Rao, Z.; Yang, H. High-Throughput Screening Identifies Established Drugs as SARS-CoV-2 PLpro Inhibitors. *Protein Cell* 2021, 12 (11), 877–888. <https://doi.org/10.1007/s13238-021-00836-9>.
- (94) Di Sarno, V.; Lauro, G.; Musella, S.; Ciaglia, T.; Vestuto, V.; Sala, M.; Scala, M. C.; Smaldone, G.; Di Matteo, F.; Novi, S.; Tecce, M. F.; Moltedo, O.; Bifulco, G.; Campiglia, P.; Gomez-Monterrey, I. M.;

- Snoeck, R.; Andrei, G.; Ostacolo, C.; Bertamino, A. Identification of a Dual Acting SARS-CoV-2 Proteases Inhibitor through in Silico Design and Step-by-Step Biological Characterization. *Eur. J. Med. Chem.* 2021, 226, 113863. <https://doi.org/10.1016/j.ejmech.2021.113863>.
- (95) Li, L.; Ma, L.; Hu, Y.; Li, X.; Yu, M.; Shang, H.; Zou, Z. Natural Biflavones Are Potent Inhibitors against SARS-CoV-2 Papain-like Protease. *Phytochemistry* 2022, 193, 112984. <https://doi.org/10.1016/j.phytochem.2021.112984>.
- (96) Zhou, Q. A.; Kato-Weinstein, J.; Li, Y.; Deng, Y.; Granet, R.; Garner, L.; Liu, C.; Polshakov, D.; Gessner, C.; Watkins, S. Potential Therapeutic Agents and Associated Bioassay Data for COVID-19 and Related Human Coronavirus Infections. *ACS Pharmacol. Transl. Sci.* 2020, 3 (5), 813–834. <https://doi.org/10.1021/acsptsci.0c00074>.
- (97) Puhl, A. C.; Gomes, G. F.; Damasceno, S.; Godoy, A. S.; Noske, G. D.; Nakamura, A. M.; Gawriljuk, V. O.; Fernandes, R. S.; Monakhova, N.; Riabova, O.; Lane, T. R.; Makarov, V.; Veras, F. P.; Batah, S. S.; Fabro, A. T.; Oliva, G.; Cunha, F. Q.; Alves-Filho, J. C.; Cunha, T. M.; Ekins, S. Pyronaridine Protects against SARS-CoV-2 Infection in Mouse. *ACS Infect. Dis.* 2022, 8 (6), 1147–1160. <https://doi.org/10.1021/acsinfecdis.2c00091>.
- (98) Welker, A.; Kersten, C.; Müller, C.; Madhugiri, R.; Zimmer, C.; Müller, P.; Zimmermann, R.; Hammerschmidt, S.; Maus, H.; Ziebuhr, J.; Sottriffer, C.; Schirmeister, T. Structure-Activity Relationships of Benzamides and Isoindolines Designed as SARS-CoV Protease Inhibitors Effective against SARS-CoV-2. *ChemMedChem* 2021, 16 (2), 340–354. <https://doi.org/10.1002/cmdc.202000548>.
- (99) Yan, H.; Liu, Z.; Yan, G.; Liu, X.; Liu, X.; Wang, Y.; Chen, Y. A Robust High-Throughput Fluorescence Polarization Assay for Rapid Screening of SARS-CoV-2 Papain-like Protease Inhibitors. *Virology* 2022, 574, 18–24. <https://doi.org/10.1016/j.virol.2022.07.006>.
- (100) Fuzo, C. A.; Martins, R. B.; Fraga-Silva, T. F. C.; Amstalden, M. K.; Canassa De Leo, T.; Souza, J. P.; Lima, T. M.; Faccioli, L. H.; Okamoto, D. N.; Juliano, M. A.; França, S. C.; Juliano, L.; Bonato, V. L. D.; Arruda, E.; Dias-Baruffi, M. Celastrol: A Lead Compound That Inhibits SARS-CoV-2 Replication, the Activity of Viral and Human Cysteine Proteases, and Virus-Induced IL-6 Secretion. *Drug Dev. Res.* 2022, 83 (7), 1623–1640. <https://doi.org/10.1002/ddr.21982>.
- (101) Liu, X.; Zheng, M.; Zhang, H.; Feng, B.; Li, J.; Zhang, Y.; Zhang, J.; Zhao, N.; Li, C.; Song, N.; Song, B.; Yang, D.; Chen, J.; Qi, A.; Zhao, L.; Luo, C.; Zang, Y.; Liu, H.; Li, J.; Zhang, B.; Zhou, Y.; Zheng, J. Characterization and Noncovalent Inhibition of the K63-Deubiquitinase Activity of SARS-Cov-2 PLpro. *Antiviral Res.* 2024, 228, 105944. <https://doi.org/10.1016/j.antiviral.2024.105944>.
- (102) Hasan, A.; Zhang, M.; Shang, Z.-P.; Yi, Y.; Kuang, Y.; Yu, R.; Fan, J.-J.; Huang, Y.-X.; Nijat, D.; Qiao, X.; Ye, M. Bioactive Prenylated Phenolic Compounds from the Aerial Parts of *Glycyrrhiza Uralensis*. *Phytochemistry* 2022, 201, 113284. <https://doi.org/10.1016/j.phytochem.2022.113284>.
- (103) Yi, Y.; Yu, R.; Xue, H.; Jin, Z.; Zhang, M.; Bao, Y.; Wang, Z.; Wei, H.; Qiao, X.; Yang, H. Chrysin 7-O- β -D-Glucuronide, a Dual Inhibitor of SARS-CoV-2 3CLpro and PLpro, for the Prevention and Treatment of COVID-19. *Int. J. Antimicrob. Agents* 2024, 63 (1), 107039. <https://doi.org/10.1016/j.ijantimicag.2023.107039>.
- (104) Tan, B.; Zhang, X.; Ansari, A.; Jadhav, P.; Tan, H.; Li, K.; Chopra, A.; Ford, A.; Chi, X.; Ruiz, F. X.; Arnold, E.; Deng, X.; Wang, J. Design of a SARS-CoV-2 Papain-like Protease Inhibitor with Antiviral Efficacy in a Mouse Model. *Science* 2024, 383 (6690), 1434–1440. <https://doi.org/10.1126/science.adm9724>.
- (105) Tian, X.; Zhao, Q.; Chen, X.; Peng, Z.; Tan, X.; Wang, Q.; Chen, L.; Yang, Y. Discovery of Novel and Highly Potent Inhibitors of SARS CoV-2 Papain-Like Protease Through Structure-Based Pharmacophore Modeling, Virtual Screening, Molecular Docking, Molecular Dynamics

- Simulations, and Biological Evaluation. *Front. Pharmacol.* 2022, 13. <https://doi.org/10.3389/fphar.2022.817715>.
- (106) Chen, X.; Chen, K.; Zhang, Z.; Wei, P.; Zhang, L.; Xu, Y.; Lun, Q.; Ma, Y.; Wu, F.; Zhang, Y.; Wang, Y.; Zhao, J.; Zhou, Y.; Zhan, J.; Xu, W. Investigating Derivatives of Tanshinone IIA Sulfonate Sodium and Chloroxine for Their Inhibition Activities against the SARS-CoV-2 Papain-like Protease. *ACS Omega* 2022, 7 (51), 48416–48426. <https://doi.org/10.1021/acsomega.2c06675>.
- (107) Wasilewicz, A.; Kirchweiger, B.; Bojkova, D.; Abi Saad, M. J.; Langeder, J.; Bütikofer, M.; Adelsberger, S.; Grienke, U.; Cinatl Jr., J.; Petermann, O.; Scapozza, L.; Orts, J.; Kirchmair, J.; Rabenau, H. F.; Rollinger, J. M. Identification of Natural Products Inhibiting SARS-CoV-2 by Targeting Viral Proteases: A Combined in Silico and in Vitro Approach. *J. Nat. Prod.* 2023, 86 (2), 264–275. <https://doi.org/10.1021/acs.jnatprod.2c00843>.
- (108) Zhao, G.; Liu, X.; Wang, S.; Bai, Z.; Zhang, S.; Wang, Y.; Yu, H.; Xu, X. Hydrogen Bonding Penalty Used for Virtual Screening to Discover Potent Inhibitors for Papain-Like Cysteine Proteases of SARS-CoV-2. *Chem. Biol. Drug Des.* 2022, 100 (4), 502–514. <https://doi.org/10.1111/cbdd.14115>.
- (109) Zang, Y.; Su, M.; Wang, Q.; Cheng, X.; Zhang, W.; Zhao, Y.; Chen, T.; Jiang, Y.; Shen, Q.; Du, J.; Tan, Q.; Wang, P.; Gao, L.; Jin, Z.; Zhang, M.; Li, C.; Zhu, Y.; Feng, B.; Tang, B.; Xie, H.; Wang, M.-W.; Zheng, M.; Pan, X.; Yang, H.; Xu, Y.; Wu, B.; Zhang, L.; Rao, Z.; Yang, X.; Jiang, H.; Xiao, G.; Zhao, Q.; Li, J. High-Throughput Screening of SARS-CoV-2 Main and Papain-like Protease Inhibitors. *Protein Cell* 2023, 14 (1), 17–27. <https://doi.org/10.1093/procel/pwac016>.
- (110) Cho, C.-C.; Li, S. G.; Lalonde, T. J.; Yang, K. S.; Yu, G.; Qiao, Y.; Xu, S.; Ray Liu, W. Drug Repurposing for the SARS-CoV-2 Papain-Like Protease. *ChemMedChem* 2022, 17 (1), e202100455. <https://doi.org/10.1002/cmdc.202100455>.
- (111) Yu, P.-C.; Huang, C.-H.; Kuo, C.-J.; Liang, P.-H.; Wang, L. H.-C.; Pan, M. Y.-C.; Chang, S.-Y.; Chao, T.-L.; leong, S.-M.; Fang, J.-T.; Huang, H.-C.; Juan, H.-F. Drug Repurposing for the Identification of Compounds with Anti-SARS-CoV-2 Capability via Multiple Targets. *Pharmaceutics* 2022, 14 (1), 176. <https://doi.org/10.3390/pharmaceutics14010176>.
- (112) Puhl, A. C.; Gomes, G. F.; Damasceno, S.; Godoy, A. S.; Noske, G. D.; Nakamura, A. M.; Gawriljuk, V. O.; Fernandes, R. S.; Monakhova, N.; Riabova, O.; Lane, T. R.; Makarov, V.; Veras, F. P.; Batah, S. S.; Fabro, A. T.; Oliva, G.; Cunha, F. Q.; Alves-Filho, J. C.; Cunha, T. M.; Ekins, S. Pyronaridine Protects Against SARS-CoV-2 in Mouse. *bioRxiv* September 30, 2021, p 2021.09.30.462449. <https://doi.org/10.1101/2021.09.30.462449>.
- (113) Osipiuk, J.; Azizi, S.-A.; Dvorkin, S.; Endres, M.; Jedrzejczak, R.; Jones, K. A.; Kang, S.; Kathayat, R. S.; Kim, Y.; Lisnyak, V. G.; Maki, S. L.; Nicolaescu, V.; Taylor, C. A.; Tesar, C.; Zhang, Y.-A.; Zhou, Z.; Randall, G.; Michalska, K.; Snyder, S. A.; Dickinson, B. C.; Joachimiak, A. Structure of Papain-like Protease from SARS-CoV-2 and Its Complexes with Non-Covalent Inhibitors. *bioRxiv* August 7, 2020, p 2020.08.06.240192. <https://doi.org/10.1101/2020.08.06.240192>.
- (114) Santos, L.; Kronenberger, T.; Almeida, R.; Silva, E.; Oliveira Rocha, R.; Oliveira, J.; Barreto, L.; Skinner, D.; Fajtová, P.; Giardini, M.; Woodworth, B.; Bardine, C.; Lourenço, A.; Craik, C.; Poso, A.; Podust, L.; Mckerrow, J.; Siqueira-Neto, J.; O'Donoghue, A.; Ferreira, R. *Structure-Based Identification of Naphthoquinones and Derivatives as Novel Inhibitors of Main Protease Mpro and Papain-like Protease PLpro of SARS-CoV-2*; 2022. <https://doi.org/10.1101/2022.01.05.475095>.
- (115) Yuan, S.; Gao, X.; Tang, K.; Cai, J.-P.; Hu, M.; Luo, P.; Wen, L.; Ye, Z.-W.; Luo, C.; Tsang, J. O.-L.; Chan, C. C.-Y.; Huang, Y.; Cao, J.; Liang, R.; Qin, Z.; Qin, B.; Yin, F.; Chu, H.; Jin, D.-Y.; Sun, R.; Chan, J. F.-W.; Cui, S.; Yuen, K.-Y. Targeting Papain-like Protease for Broad-Spectrum Coronavirus Inhibition. *Protein Cell* 2022, 13 (12), 940–953. <https://doi.org/10.1007/s13238-022-00909-3>.
- (116) Ma, C.; Wang, J. Validation and Invalidation of SARS-CoV-2 Papain-like Protease Inhibitors. *ACS Pharmacol. Transl. Sci.* 2022, 5 (2), 102–109. <https://doi.org/10.1021/acsptsci.1c00240>.

- (117) Smith, M. B. *March's Advanced Organic Chemistry: Reactions, Mechanisms, and Structure*; John Wiley & Sons, 2013.
- (118) Sabadini, G.; Mellado, M.; Morales, C.; Mella, J. Arylamines QSAR-Based Design and Molecular Dynamics of New Phenylthiophene and Benzimidazole Derivatives with Affinity for the C111, Y268, and H73 Sites of SARS-CoV-2 PLpro Enzyme. *Pharmaceuticals* 2024, 17 (5), 606. <https://doi.org/10.3390/ph17050606>.
- (119) Marotta, L.; Rossi, S.; Ibba, R.; Brogi, S.; Calderone, V.; Butini, S.; Campiani, G.; Gemma, S. The Green Chemistry of Chalcones: Valuable Sources of Privileged Core Structures for Drug Discovery. *Front. Chem.* 2022, 10, 988376. <https://doi.org/10.3389/fchem.2022.988376>.
- (120) Karrouchi, K.; Radi, S.; Ramli, Y.; Taoufik, J.; Mabkhot, Y. N.; Al-aizari, F. A.; Ansar, M. Synthesis and Pharmacological Activities of Pyrazole Derivatives: A Review. *Molecules* 2018, 23 (1), 134. <https://doi.org/10.3390/molecules23010134>.
- (121) Perrin, C. L.; Chang, K.-L. The Complete Mechanism of an Aldol Condensation. *J. Org. Chem.* 2016, 81 (13), 5631–5635. <https://doi.org/10.1021/acs.joc.6b00959>.
- (122) Nájera, C.; Sydnes, L. K.; Yus, M. Conjugated Ynones in Organic Synthesis. *Chem. Rev.* 2019, 119 (20), 11110–11244. <https://doi.org/10.1021/acs.chemrev.9b00277>.
- (123) Bhathiwal, A. S.; Bendi, A.; Tiwari, A. A Study on Synthesis of Benzodiazepine Scaffolds Using Biologically Active Chalcones as Precursors. *J. Mol. Struct.* 2022, 1258, 132649. <https://doi.org/10.1016/j.molstruc.2022.132649>.
- (124) Baldwin, J. E. Rules for Ring Closure. *J. Chem. Soc. Chem. Commun.* 1976, No. 18, 734–736. <https://doi.org/10.1039/C39760000734>.
- (125) Ahmed, E. A.; Soliman, A. M. M.; Ali, A. M.; Ali El-Remaily, M. A. E. A. Boosting the Catalytic Performance of Zinc Linked Amino Acid Complex as an Eco-Friendly for Synthesis of Novel Pyrimidines in Aqueous Medium. *Appl. Organomet. Chem.* 2021, 35 (5), e6197. <https://doi.org/10.1002/aoc.6197>.
- (126) Mella-Raipán, J. A.; Lagos, C. F.; Recabarren-Gajardo, G.; Espinosa-Bustos, C.; Romero-Parra, J.; Pessoa-Mahana, H.; Iturriaga-Vásquez, P.; Pessoa-Mahana, C. D. Design, Synthesis, Binding and Docking-Based 3D-QSAR Studies of 2-Pyridylbenzimidazoles—A New Family of High Affinity CB1 Cannabinoid Ligands. *Molecules* 2013, 18 (4), 3972–4001. <https://doi.org/10.3390/molecules18043972>.
- (127) Chung, N. T.; Dung, V. C.; Duc, D. X. Recent Achievements in the Synthesis of Benzimidazole Derivatives. *RSC Adv.* 13 (46), 32734–32771. <https://doi.org/10.1039/d3ra05960j>.
- (128) Zhang, P.; Liu, C.; Yu, L.; Hou, H.; Sun, W.; Ke, F. Synthesis of Benzimidazole by Mortar–Pestle Grinding Method. *Green Chem. Lett. Rev.* 2021, 14 (4), 612–619. <https://doi.org/10.1080/17518253.2021.1991483>.

10 Anexo

Anexo adjunto en formato PDF

Anexo

Anexo I

Nombre	pIC50 experimental	SMILES	Residuo
molécula 001	5.793174124	<chem>O=C(N[C@@H](C1=CC2=C(C=C1)C=CC=C2)C)C3=CC(N)=CC=C3C</chem>	0.519
molécula 002	4.614393726	<chem>O=C(C([C@@H](C/C=C(C)/C)OC(C)=O)=CC1=O)C2=C1C(O)=CC=C2O</chem>	-0.589
molécula 003	4.165579296	<chem>CC1=C(C=CC(C(O[C[C@@H]2C)=C2C3=O)=C4C3=O)C4=CC=C1</chem>	-1.658
molécula 004	5.616184634	<chem>CC([C@H]1O)(C)OC2=C(O)C=C3C4=C2[C@@]1([H])OC5=CC(O)=C(C/C=C(C)/C)C(CC3)=C45</chem>	-0.163
molécula 005	5.728158393	<chem>CC([C@@H]1O)(C)OC2=C(O)C=C3C4=C2[C@]1([H])OC5=CC(O)=C(C/C=C(C(C)/C)C(CC3)=C45</chem>	-0.051
molécula 006	5.050609993	<chem>CC1=C(C(C=C2C1=CC=C3[C@]2(CC[C@@]4([C@@]3(CC[C@@]5([C@H]4C[C@](C([O-])=O)(CC5)C)C)C)C)=O)O</chem>	-0.927
molécula 007	5.721246399	<chem>O=C(C1=C(C2=C3C(N4C=C(C5=CC=CC=C5)N=N4)C(C)</chem>	-0.37

		(C)O2)C=CC=C1)C3=O	
molécula 008	5.488116639	O=C(C1=C(C2=C3C(N4C=C(CNC5=CC(C(C=CC=C6)=C6C5=O)=O)N=N4)C(C)(C)O2)C=CC=C1)C3=O	-0.99
molécula 009	5.508638306	O=C(C1=C(C2=C3C(N4C=C(CNC5=C(Br)C(C(C=CC=C6)=C6C5=O)=O)N=N4)C(C)(C)O2)C=CC=C1)C3=O	-1.163
molécula 010	4.219826756	ClC1=CC=C(C=C1)N2CCN(CC2)CC(NC3=CC=C(C=C3)OC4=CC=CC=C4)=O	-0.514
molécula 011	4.22848601	ClC1=CC=C(/C=C2C(NC3=C(C=CC(C(NCCCN(CC)CC)=O)=C3)O/2)=O)C=C1	-1.225
molécula 012	5.847711656	COC1=CC2=C(C=C1)N=C/C2=C\N/N=C(N)/NCCCCC	0.305
molécula 013	4.975925013	CCCCCCCCCCC1=C(C(O)=O)C(O)=CC=C1	-0.168
molécula 014	5.730487056	COC1=NC2=C(C3=C(C=C(C=C3)Cl)N=C2C=C1)NC4=CC(CN5CCCC5)=C(C(CN6CCCC6)=C4)O	-0.297
molécula 015	5.484126156	COC1=CC2=C(C3=C(C=C(C=C3)Cl)N=C2C=C1)NC4=CC(CN(C)C)=C(C(CN(C)C)=C4)O	0.17
molécula 016	5.612610174	COC1=CC2=C(C3=C(C=C(C=C3)Cl)N=C2C=C1)NC4=	-0.588

		<chem>CC(CN5CCCC5)=C(C(CN6CCCC6)=C4)O</chem>	
molécula 017	5.591760035	<chem>COC1=CC2=C(C3=C(C=C(C=C3)Cl)N=C2C=C1)NC4=CC(CN5CCCC5)=C(C=C4)O</chem>	0.156
molécula 018	5.420216403	<chem>CCN(C(SSC(N(CC)CC)=S)=S)CC</chem>	0.052
molécula 019	6.698970004	<chem>S=C(SSC(N1CCOC1)=S)N2CCOCC2</chem>	0.997
molécula 020	6.22184875	<chem>S=C(SSC(OCC)=S)OCC</chem>	0.776
molécula 021*	6.15490196	<chem>S=C(SSC(N1CCCC1)=S)N2CCCC2</chem>	0.37
molécula 022	6.22184875	<chem>S=C(SSC(N(C)C)=S)N(C)C</chem>	0.781
molécula 023	5.721246399	<chem>S=C(SSC(N1CCN(C)CC1)=S)N2CCN(C)CC2</chem>	0.051
molécula 024	6.142667504	<chem>O=C(C1CCN([C@H](C2=CC=CC3=CC=CC=C32)C)C1)NCC4=CC(OC)=NC=C4</chem>	0.514
molécula 025	5.991399828	<chem>O=C(C1CCN([C@H](C2=CC=CC3=CC=CC=C32)C)C1)NCC4=CC(F)=CC=C4</chem>	0.508
molécula 026	5.13076828	<chem>O=C(C(SC1=S)=C(N2)N1C3=C(C=CC=C3)C2=O)NC4=C(Cl)C=CC=C4</chem>	0.135
molécula 027	5.086186148	<chem>O=C(C(SC1=S)=C(N2)N1C3=C(C=C(OC)C(OC)=C3)C2=O)NC4=C(C)C=CC(Cl)=C4</chem>	-0.311
molécula 028	4.801342913	<chem>O=C(C(SC1=S)=C(N2)N1C3=C(C=C4C(OCO4)=C3)C2=</chem>	-1.297

		<chem>O)NC5=CC=CC(C)=C5</chem>	
molécula 029	5.216096421	<chem>C[C@@H](NC(C1=C(C)C=CC(NC2CN(C2)C(OC(C)(C)C)=O)=C1)=O)C3=C(C=CC4=C3C=CC=C4</chem>	-0.331
molécula 030	5.958607315	<chem>C[C@H](C1=CC=C(C2=C1C=CC=C2)NC(C3=C(C=CC(NC4CCN(C(OC(C)(C)C)=O)CC4)=C3)C)=O</chem>	0.237
molécula 031	5.259637311	<chem>C[C@@H](NC(C1=C(C)C=CC(NC(C2CN(C2)C(OC(C)(C)C)=O)=O)=C1)=O)C3=CC=CC4=C3C=CC=C4</chem>	-0.243
molécula 032	5.22184875	<chem>C[C@@H](NC(C1=C(C)C=CC(NC(C2CCN(CC2)C(OC(C)(C)C)=O)=O)=C1)=O)C3=CC=CC4=C3C=CC=C4</chem>	-0.525
molécula 033	5.995678626	<chem>C[C@@H](NC(C1=C(C)C=CC(NC2CN(C2)=C1)=O)C3=C(C=CC4=C3C=CC=C4</chem>	0.783
molécula 034	6.22184875	<chem>C[C@@H](NC(C1=C(C)C=CC(NC2CCNCC2)=C1)=O)C3=CC=CC4=C3C=CC=C4</chem>	0.668
molécula 035	5.920818754	<chem>C[C@@H](NC(C1=C(C)C=CC(NC(C2CNC2)=O)=C1)=O)C3=CC=CC4=C3C=CC=C4</chem>	0.745
molécula 036	6.096910013	<chem>C[C@@H](NC(C1=C(C)C=CC(NC(C2CCNCC2)=O)=C1)</chem>	0.575

		<chem>=O)C3=CC=CC4=C3C=CC=C4</chem>	
molécula 037	6.15490196	<chem>C[C@@H](NC(C1=C(C)C=CC(NC2CCN(C)CC2)=C1)=O)C3=CC=CC4=C3C=CC=C4</chem>	0.482
molécula 038*	5.795880017	<chem>C[C@@H](NC(C1=C(C)C=CC(N(C)C2CN(C)C2)=C1)=O)C3=CC=CC4=C3C=CC=C4</chem>	0.422
molécula 039*	5.366531544	<chem>C[C@@H](NC(C1=C(C)C=CC(NC2CN(C)C2)=C1)=O)C3=CC=CC4=C3C=CC=C4</chem>	0.131
molécula 040	5.619788758	<chem>C[C@@H](NC(C1=C(C)C=CC(N(C)C2CNC2)=C1)=O)C3=CC=CC4=C3C=CC=C4</chem>	0.312
molécula 041*	5.408935393	<chem>C[C@@H](NC(C1=C(C)C=CC(CNCC(O)=O)=C1)=O)C2=CC=CC3=CC=CC=C32</chem>	-0.269
molécula 042*	5.886056648	<chem>C[C@@H](NC(C1=C(C)C=CC(CNC2CNC2)=C1)=O)C3=CC=CC4=CC=CC=C43</chem>	0.417
molécula 043	5.48148606	<chem>C[C@@H](NC(C1=C(C)C=CC(CN2CC(C(O)=O)C2)=C1)=O)C3=CC=CC4=C3C=CC=C4</chem>	-0.111
molécula 044	5.214670165	<chem>C[C@@H](NC(C1=C(C)C=CC(CN2CC(C(C(O)=O)CC2)=C1)=O)C3=CC=CC4=C3C=CC=C4</chem>	-0.73
molécula 045	5.318758763	<chem>C[C@@H](NC(C1=C(Cl)C=CC(N)=C1)</chem>	0.193

		<chem>=O)C2=CC=CC3=C2C=CC=C3</chem>	
molécula 046	4.677780705	<chem>C[C@@H](NC(C1=C(Br)C=CC(N)=C1)=O)C2=CC=CC3=C2C=CC=C3</chem>	-0.625
molécula 047	5.15490196	<chem>C[C@@H](NC(C1=C(Br)C=NC(N)=C1)=O)C2=CC=CC3=C2C=CC=C3</chem>	-0.28
molécula 048	6.096910013	<chem>C[C@@H](NC(C1=CC=C(C=CN2)C2=C1)=O)C3=CC=CC4=C3C=CC=C4</chem>	0.253
molécula 049	5.200659451	<chem>C[C@@H](NC(C1=C(C)C=CC2=C1C=CN2)=O)C3=CC=C4=C3C=CC=C4</chem>	-0.642
molécula 050	5.193820026	<chem>C[C@@H](NC(C1=C(C)C=CC2=C1C=CN2CC3CNC3)=O)C4=CC=CC5=C4C=CC=C5</chem>	-0.697
molécula 051*	5.13076828	<chem>C[C@@H](NC(C1=C(C)C=NC(NC2CNC2)=C1)=O)C3=CC=CC4=CC=CC=C43</chem>	-0.164
molécula 052	5.387216143	<chem>C[C@@H](NC(C1=C(Cl)C=CC(NC2CNC2)=C1)=O)C3=CC=CC4=CC=CC=C43</chem>	0.306
molécula 053	4.970616222	<chem>C[C@@H](NC(C1=C(Cl)C=CC(N(C)C2CNC2)=C1)=O)C3=CC=CC4=CC=C4C=C43</chem>	-0.226
molécula 054	5.48148606	<chem>C[C@@H](NC(C1=C(C)C=CC(N)=C1)=O)C2=CSC3=C2C=CC=C3</chem>	-0.121

molécula 055	5.327902142	<chem>CC(NC(C1=C(C)C=CC(N)=C1)=O)C2=CC=CC3=C2NC=C3</chem>	-0.054
molécula 056	5.244125144	<chem>O=C(C1=C(C)C=C(C(N(C2CNC2)C)=C1)NC(C)C3=CC=CC4=C3NC=C4</chem>	-0.183
molécula 057	5.173925197	<chem>CC1=C(C(N[C@H](C)C2=CC3=C(SC=C3)C=C2)=O)C=C(N(C)C4CNC4)C=C1</chem>	-0.516
molécula 058*	5.48148606	<chem>C[C@@H](NC(C1=C(C)C=CC(N(C)C2CNC2)=C1)=O)C3=CC=CC4=C3C=CN4CC5CCC5</chem>	-0.079
molécula 059	5.619788758	<chem>C[C@@H](NC(C1=C(C)C=CC(N(C2CNC2)C)=C1)=O)C3=CSC4=C3C=CC=C4</chem>	-0.204
molécula 060	5.070581074	<chem>C[C@@H](NC(C1=C(Cl)C=CC(NC2CNC2)=C1)=O)C3=CSC4=C3C=CC=C4</chem>	-0.36
molécula 061	4.962573502	<chem>C[C@@H](NC(C1=C(Cl)C=CC(N(C2CNC2)C)=C1)=O)C3=CSC4=C3C=CC=C4</chem>	-0.603
molécula 062	5.795880017	<chem>C[C@@H](NC(C1=C(C)C=CC(N(C)C2CNC2)=C1)=O)C3=C4C(NC5=C4C=CC=C5)=CC=C3</chem>	0.069
molécula 063	5.721246399	<chem>C[C@@H](NC(C1=C(C)C=CC(NC2CNC2)=C1)=O)C3=CS4=C3C=CC=C4</chem>	0.158
molécula 064	5.744727495	<chem>C[C@@H](NC(C1=C(C)C=CC(NC2CNC2)=C1)=O)C3=CC=CC4=C3C=CC=C4</chem>	0.116

		<chem>C2)=C1)=O)C3=C4 C(NC5=C4C=CC= C5)=CC=C3</chem>	
molécula 065	4.251811973	<chem>C[C@@H](NC(C1= C(C)C=CC(N(C2C NC2)C)=C1)=O)C3 =NC=CC4=CC=CC =C43</chem>	-0.972
molécula 066	5.552841969	<chem>C[C@@H](NC(C1= C(Cl)C=CC(NC2C NC2)=C1)=O)C3= CC(C4=CC=CS4)= CC=C3</chem>	-0.041
molécula 067	6.229147988	<chem>C[C@@H](NC(C1= C(C)C=CC(NC2CN C2)=C1)=O)C3=C C(C4=CC=CS4)=C C=C3</chem>	0.489
molécula 068	5.886056648	<chem>C[C@@H](NC(C1= C(C)C=CC(NC2CN C2)=C1)=O)C3=C C(C4=CSC=C4)=C C=C3</chem>	0.018
molécula 069	5.744727495	<chem>C[C@@H](NC(C1= C(C)C=CC(NC2CN C2)=C1)=O)C3=C C(C4=CNC=C4)=C C=C3</chem>	0.099
molécula 070	5.920818754	<chem>C[C@@H](NC(C1= C(C)C=CC(NC2CN C2)=C1)=O)C3=C C(C4=CC=C(CN5 CCNCC5)S4)=CC= C3</chem>	-0.094
molécula 071	6.045757491	<chem>C[C@@H](NC(C1= C(C)C=CC(NC2CN C2)=C1)=O)C3=C C(C4=CC=C(CN5 CCOCC5)S4)=CC= C3</chem>	0.01
molécula 072	5.795880017	<chem>C[C@H](C1=CC(C 2=CC=C(S2)CNC3 CCN(CC3)C)=CC= C1)NC(C4=C(C=C</chem>	-0.357

		<chem>C(NC5CNC5)=C4)C)=O</chem>	
molécula 073	5.568636236	<chem>C[C@@H](NC(C1=C(C)C=CC(NC2CN(C2)=C1)=O)C3=C(C(C4=CC=C(CNC5CCNCC5)S4)=CC=C3</chem>	-0.539
molécula 074	6.408935393	<chem>C[C@@H](NC(C1=C(C)C=CC(NC2CN(C2)=C1)=O)C3=C(C(C4=CC=C(CNC5CCCC5)S4)=CC=C3</chem>	0.029
molécula 075	6.251811973	<chem>C[C@@H](NC(C1=C(C)C=CC(NC2CN(C2)=C1)=O)C3=C(C(C4=CC=C(CN5CCCC5)S4)=CC=C3</chem>	-0.1
molécula 076	6.124938737	<chem>C[C@@H](NC(C1=C(C)C=CC(NC2CN(C2)=C1)=O)C3=C(C(C4=CC=C(C)S4)=CC=C3</chem>	0.26
molécula 077	6.013228266	<chem>C[C@@H](NC(C1=C(C)C=CC(NC2CN(C2)=C1)=O)C3=C(C(C4=CC=C(C(O)=O)S4)=CC=C3</chem>	0.199
molécula 078*	6.091514981	<chem>C[C@@H](NC(C1=C(C)C=CC(NC2CN(C2)=C1)=O)C3=C(C(C4=CC=C(C(OC)=O)S4)=CC=C3</chem>	0.212
molécula 079	5.958607315	<chem>C[C@@H](NC(C1=C(C)C=CC(NC2CN(C2)=C1)=O)C3=C(C(C4=CC=C(C(NC)=O)S4)=CC=C3</chem>	0.104
molécula 080	6.716020716	<chem>FC(C=C1)=CC=C1CSC2=NC3=CC=CN(C(C(NC4=CC(NC5CCNCC5)=CC=</chem>	0.886

		C4C)=O)CC)C3=N 2	
molécula 081	6.193820026	C[C@@H](NC(C1=C(C)C=CC(NC2CN C2)=C1)=O)C3=C C(C4=CC=C(CN5 CC[C@@H](O)C5) S4)=CC=C3	-0.153
molécula 082	5.958607315	C[C@@H](NC(C1=C(C)C=CC(NC2CN C2)=C1)=O)C3=C C(C4=CC=C(CN5 CC[C@@H](C)C5) S4)=CC=C3	-0.439
molécula 083	5.657577319	C[C@@H](NC(C1=C(C)C=CC(NC2CN C2)=C1)=O)C3=C C(C4=CC=C(CN5 CCC5)S4)=CC=C3	-0.047
molécula 084	6.15490196	C[C@@H](NC(C1=C(C)C=CC(NC2CN C2)=C1)=O)C3=C C(C4=CC=C(CN5 CC[C@H](O)C5)S4)=CC=C3	-0.192
molécula 085	5.187086643	C[C@@H](NC(C1=C(C)C=CC(NC2CN C2)=C1)=O)C3=C C(Br)=CC=C3	0.191
molécula 086	6.48148606	C[C@@H](NC(C1=C(C)C=CC(NC2CN C2)=C1)=O)C3=C C(C4=CC=C(CN[C @@H]5CCOC5)S4)=CC=C3	0.133
molécula 087	6.207608311	C[C@@H](NC(C1=C(C)C=CC(NC2CN C2)=C1)=O)C3=C C(C4=CC=C(CN[C @H]5CCOC5)S4)= CC=C3	-0.141
molécula 088	6.431798276	C[C@@H](NC(C1=C(C)C=CC(NC2CN C2)=C1)=O)C3=C	0.422

		<chem>C(C4=CC=C(CNC(C)=O)S4)=CC=C3</chem>	
molécula 089*	5.273272791	<chem>O=C1C2=C(O)C(O)=C(O[C@H]3O[C@H](CO)[C@@H](O)[C@H](O)[C@H]3O)C=C2OC(C4=CC=CC=C4)=C1</chem>	0.391
molécula 090	5.595166283	<chem>O=C1C2=C(O)C=C(O[C@H]3O[C@H](C(O)=O)[C@@H](O)[C@H](O)[C@H]3O)C=C2OC(C4=CC=CC=C4)=C1</chem>	0.631
molécula 091*	6.677780705	<chem>C[C@@H](NC(C1=C(C)C=CC(NC2CN(C2)=C1)=O)C3=CC(C4=CC=C(CN[C@@H]5CC[C@H](O)C5)S4)=CC=C3</chem>	0.288
molécula 092	6.366531544	<chem>C[C@@H](NC(C1=C(C)C=CC(NC2CN(C2)=C1)=O)C3=CC(C4=CC=C(CN[C@@H]5CC[C@@H](O)C5)S4)=CC=C3</chem>	-0.024
molécula 093	6.946921557	<chem>C[C@@H](NC(C1=C(C)C=CC(NC2CN(C2)=C1)=O)C3=CC(C4=CC=C(CN[C@H]5CC[C@@H](O)C5)S4)=CC=C3</chem>	0.557
molécula 094	6.602059991	<chem>C[C@@H](NC(C1=C(C)C=CC(NC2CN(C2)=C1)=O)C3=CC(C4=CC=C(CN[C@H]5CC[C@H](O)C5)S4)=CC=C3</chem>	0.212
molécula 095	6.091514981	<chem>C[C@@H](NC(C1=C(C)C=CC(N)=C1)=O)C2=CC(C3=CC=C(CNC(OC(C)(C)C)=O)S3)=CC=C2</chem>	0.057

molécula 096*	5.744727495	<chem>C[C@@H](NC(C1=C(C)C=CC(N)=C1)=O)C2=CC(C3=CC=C(CN)S3)=CC=C2</chem>	-0.19
molécula 097	5.958607315	<chem>C[C@@H](NC(C1=C(C)C=CC(N)=C1)=O)C2=CC(C3=CC=C(CNC(C4CCCC4)=O)S3)=CC=C2</chem>	-0.46
molécula 098	5.638272164	<chem>C[C@@H](NC(C1=C(C)C=CC(NC(C)=O)=C1)=O)C2=CC(C3=CC=C(CNC(C)=O)S3)=CC=C2</chem>	-0.477
molécula 099	5.853871964	<chem>C[C@@H](NC(C1=C(C)C=CC(NC(C)=O)=C1)=O)C2=CC(C3=CC=C(CN[C@H]4CC[C@@H](O)C4)S3)=CC=C2</chem>	-0.682
molécula 100	5.619788758	<chem>C[C@]([H])(C1=CC=CC2=C1C3=CC=CC=C3N2)NC(C4=CC(NC5CCNCC5)=CC=C4C)=O</chem>	-0.354
molécula 101	6.22184875	<chem>C[C@]([H])(C1=CC(C2=CC=C(S2)C(OCC)=O)=CC=C1)NC(C3=CC(NC4CN(C4)=CC=C3C)=O</chem>	0.337
molécula 102	5.677780705	<chem>CC1=C(C(N[C@@](C)([H])C2=CSC3=C2C=CC=C3)=O)C=C(NC4COC4)C=C1</chem>	0.084
molécula 103	6.096910013	<chem>C[C@](C1=C2C=CC=CC2=CC=C1)(NC(C3=CC(NC4CCN(CC4)C(F)(F)F)=CC=C3C)=O)[H]</chem>	0.482
molécula 104	4.695509472	<chem>O=C1C2=C(N(CCOC)C=[N+]2CC3=</chem>	-0.817

		<chem>NC=CN=C3)C(C4=CC=CC=C41)=O</chem>	
molécula 105	4.341035157	<chem>CN(C/C=C/C(C(NC1=C(C=C2N=CN=C(C2=C1)NC3=CC(C1)=C(C=C3)F)O[C@H]4CCOC4)=O)C</chem>	-0.731
molécula 106	4.8326826	<chem>[H][C@@]12CCCN1C([C@@H](C(C)C)NC([C@H]([C@H](OC([C@@H](N(C(CN(C2=O)C)=O)C)C(C)C)=O)C)NC(C3=C4N=C5C(OC4=C(C=C3)C)=C(C(C(N)=C5C(N[C@H]6[C@H](OC([C@@H](N(C(CN(C([C@@]7(CCCN7C([C@@H](C(C)C)NC6=O)=O)[H])=O)C)=O)C)C(C)C)=O)C)=O)=O)C)=O)=O</chem>	-
molécula 107	5.337242168	<chem>C[C@H]1[C@@H](O)[C@@H](N)C[C@H](O[C@H]2C[C@@])(CC3=C2C(O)=C4C(C(C5=C(C(O)C)=CC=C5)C4=O)=O)=C3O)(C(CO)=O)O)O1</chem>	-0.715
molécula 108	4.935542011	<chem>CCOC1=C(NC(/C=C/C/CN(C)C)=O)C=C2C(N=CC(C#N)=C2NC3=CC(Cl)=C(OCC4=CC=CC=N4)C=C3)=C1</chem>	-0.053
molécula 109	4.779891912	<chem>O=[N+](O-)]C1=CC(Cl)=C(NC(C2=C(O)C=CC(Cl)=C2)=O)C=C1</chem>	0.524
molécula 110	4.058638	<chem>C[C@@H]1CCC2C[C@@H](/C(=C/C=C/C=C\C[C@H](</chem>	-

		<chem>C[C@H](C(=O)[C@@H]([C@@H](/C(=C/[C@H](C(=O)CC(OC(=O)C3CCC(CN3C(=O)C(=O)[C@@]1(O2)O)[C@H](C)C[C@@H]4CC[C@H]([C@@H](C4)OC)O)C)/C)O)OC)C)/C)OC</chem>	
molécula 111	4.387216143	<chem>CN(CCOC1=CC=C(/C(C2=CC=CC=C2)=C(CC)\C3=CC=CC=C3)C=C1)C</chem>	-1.101
molécula 112	4.12090412	<chem>O=C1C=C(C2=CC=C(O)C=C2)OC3=C1C(O)=CC(O)=C3</chem>	0.009
molécula 113*	4.040005162	<chem>O=C1C=C(C2=CC=C(OC)C=C2)OC3=C1C(O)=CC(O)=C3</chem>	-0.173
molécula 114	4.555955204	<chem>C/C(C)=C/CC1=CC(C(OC2=C3C=CC(O)=C2)=C3C(O4)=O)=C4C=C1O</chem>	-0.879
molécula 115*	6.173925197	<chem>O=C(NCC1=CC=CC=C1)[C@@H](NC([C@@H](NC(CCl)=O)CC=C)=O)CC2=CNC3=CC=CC=C32</chem>	0.034
molécula 116*	5.088842391	<chem>O=C1C2=C(N=C(C3=CN=CC=C3)O2)C(C4=CC=CC=C41)=O</chem>	-0.082
molécula 117	4.215668052	<chem>OC1=CC(O)=CC(/C=C/C2=CC=C(O)C=C2)=C1</chem>	0.003
molécula 118	6.036212173	<chem>C[C@@H](NC(C1=CC(C)C=CC(NC2CN(C2)=C1)=O)C3=CC(C4=CC=C(C(NC[C@H]5CCCO5)=O)S4)=CC=C3</chem>	-0.461

molécula 119	5.161150909	<chem>CC1=C(C(N[C@@H](C2=C(C=CC=C3)C3=CC=C2)C)=O)C=CC=C1</chem>	0.066
molécula 120*	4.823908741	<chem>ClC1=C(C(N[C@@H](C2=C(C=CC=C3)C3=CC=C2)C)=O)C=CC=C1</chem>	-0.104
molécula 121	4.37675071	<chem>CC1=C(C(N[C@@H](C2=C(C=CC=C3)C3=CC=C2)C)=O)SC=C1</chem>	-1.046
molécula 122	5.124938737	<chem>O=C(N[C@@H](C1=C(C=CC=C2)C2=CC=C1)C)C3=C(C)C=C(N)C=C3</chem>	0.14
molécula 123	5	<chem>O=C(N[C@@H](C1=C(C=CC=C2)C2=CC=C1)C)C3=C(C)C=CC(N(C)C)=C3</chem>	-0.46
molécula 124	5	<chem>O=C(N[C@@H](C1=C(C=CC=C2)C2=CC=C1)C)C3=C(C)C=C(NS(C)(=O)=O)C=C3</chem>	-0.387
molécula 125	5.119186408	<chem>C[C@@H](C1=C2C=CC=CC2=CC=C1)N3CC4=C(C3)C=CC=C4</chem>	-0.62
molécula 126	4.244125144	<chem>C[C@H](C1=C2C=CC=CC2=CC=C1)N3CC4=C(C3)C=C(N)C=C4</chem>	-1.537
molécula 127*	4.886056648	<chem>OC(C=C1)=C(C2=C(O)C=C(O)C3=C2OC(C4=CC=C(O)C=C4)=CC3=O)C=C1C5=CC(C6=C(O)C=C(O)C=C6O5)=O</chem>	0.385
molécula 128	4.364516253	<chem>OC(C=C1)=C(C2=C(O)C=C(O)C3=C2OC(C4=CC=C(OC)C=C4)=CC3=O)C=</chem>	-0.244

		<chem>C1C5=CC(C6=C(O)C=C(O)C=C6O5)=O</chem>	
molécula 129*	4.525783736	<chem>OC(C=C1)=CC=C1C(O2)=CC(C3=C2C(C4=C(OC)C=CC(C5=CC(C6=C(O)C=C(OC)C=C6O5)=O)=C4)=C(O)C=C3O)=O</chem>	-0.247
molécula 130	4.505845406	<chem>OC1=CC(O)=C2C(OC(C3=CC(C4=C(O)C=C(O)C5=C4OC(C6=CC=C(OC)C=C6)=CC5=O)=C(OC)C=C3)=CC2=O)=C1</chem>	-0.231
molécula 131	4.458420756	<chem>OC1=C2C(OC(C3=CC(C4=C(O)C=C(O)C5=C4OC(C6=C=C(OC)C=C6)=C5=O)=C(OC)C=C3)=CC2=O)=CC(O)=C1</chem>	-0.417
molécula 132	4.438898616	<chem>OC(C=C1)=CC=C1[C@H]2OC3=C(C([C@H]2C4=C(O)C=C(O)C5=C4OC(C6=CC(O)=C(O)C=C6)=CC5=O)=O)C(O)=CC(O)=C3</chem>	0.089
molécula 133	5.022276395	<chem>OC1=CC2=C(C(C=C(C3=CC=C(OC4=C(O)C(C(C=C(C5=CC=C(O)C=C5)O6)=O)=C6C=C4O)C=C3)O2)=O)C(O)=C1</chem>	0.665
molécula 134	4.580044252	<chem>OC1=CC2=C(C(C=C(C3=CC=C(OC4=C(O)C(C(C=C(C5=CC=C(OC)C=C5)O6)=O)=C6C=C4OC</chem>	0.025

		<chem>)C=C3)O2)=O)C(O)=C1</chem>	
molécula 135*	4.642065153	<chem>OC1=CC2=C(C(C=C(C3=CC(OC4=C(C=C(C(OC5=C6C(O)=CC(O)=C5)=CC6=O)C=C4)=C(OC)C=C3)O2)=O)C(O)=C1</chem>	0.279
molécula 136	5.484126156	<chem>ClC1=CC=C2C(N=C3C=CC(OC)=CC3=C2NC4=CC(CN(C)C)=C(O)C(CN(C)C)=C4)=C1</chem>	0.17
molécula 137	5.591760035	<chem>ClC1=CC=C2C(N=C3C=CC(OC)=CC3=C2NC4=CC=C(O)C(CN5CCCC5)=C4)=C1</chem>	0.156
molécula 138*	5.292429824	<chem>CC1=CC=C(NC(NC(N)=O)=O)C=C1C(N[C@H](C)C2=CC=CC3=C2C=CC=C3)=O</chem>	-0.002
molécula 139	5.193820026	<chem>CC1=CC=C(NC(C=C)=O)C=C1C(N[C@H](C)C2=CC=CC3=C2C=CC=C3)=O</chem>	-0.148
molécula 140	4.364516253	<chem>O=C(C1=C(C(F)(F)F)C=CC(N)=C1)NC2=CC=CC3=C2C=CC=C3</chem>	-0.874
molécula 141	4.774690718	<chem>CC1=CC=C(NC(NCCCC)=O)C=C1C(N[C@H](C)C2=CC=CC3=C2C=CC=C3)=O</chem>	-0.675
molécula 142	5.15490196	<chem>CC1=CC=C(NC(OC2=CC=C([N+])([O-])=O)C=C2)=O)C=C1C(N[C@H](C)C3=CC=CC4=C3C=CC=C4)=O</chem>	-0.043

molécula 143	4.896196279	<chem>CC1=CC=C(NC(CCCC)=O)C=C1C(N[C@H](C)C2=CC=CC3=C2C=CC=C3)=O</chem>	-0.543
molécula 144	4.716020716	<chem>C[C@H](C1=CC=C2C2=C1C=CC=C2)N(C)CC3=CNC(NC3=O)=O</chem>	-0.756
molécula 145	5.148741651	<chem>CN([C@@H](C1=CC=CC2=C1C=CC=C2)C)CC3=CC=C3N3</chem>	-0.32
molécula 146*	4.931443105	<chem>CN([C@@H](C1=CC=CC2=C1C=CC=C2)C)CC3=CC=C(OC)N=C3</chem>	-0.397
molécula 147	5.289882635	<chem>CN([C@@H](C1=CC=CC2=C1C=CC=C2)C)CC3=CC4=C(N=CC=C4)C=C3</chem>	-0.256
molécula 148	6.207608311	<chem>CN([C@@H](C1=CC=CC2=C1C=CC=C2)C)CC3=CC=CC4=C3C=CN4</chem>	0.319
molécula 149	5.801342913	<chem>CN([C@@H](C1=CC=CC2=C1C=CC=C2)C)CC3=CC4=C(NC=C4)C=C3</chem>	-0.085
molécula 150*	5.886056648	<chem>CN([C@@H](C1=CC=CC2=C1C=CC=C2)C)CC3=CC4=C(C=CN4)C=C3</chem>	-0.066
molécula 151	5.607303047	<chem>CN([C@@H](C1=CC=CC2=C1C=CC=C2)C)CC3=CC4=C(NN=N4)C=C3</chem>	-0.15
molécula 152*	6.180456064	<chem>CN([C@@H](C1=CC=CC2=C1C=CC=C2)C)CC3=CC4=C(NN=C4)C=C3</chem>	0.336
molécula 153	6.173925197	<chem>CN(CC1=CNC2=C1C=CC=C2)[C@H]</chem>	0.309

		(C)C3=CC=CC4=C3C=CC=C4	
molécula 154	5.93930216	CN(CC1=CN(C2=C1C=CC=C2)CC(N)=O)[C@H](C)C3=CC=CC4=C3C=C=C4	-0.124
molécula 155	5.399027104	CN([C@@H](C1=CC=CC2=C1C=CC=C2)C)CC3=NNC4=C3C=C(O)C=C4	-0.507
molécula 156	5.886056	CN([C@@H](C1=CC=CC2=C1C=CC=C2)C)CC3=CC=C(B(O)O)C=C3	-
molécula 157	6.060480747	CN([C@@H](C1=CC=CC2=C1C=CC=C2)C)CC3=C(C)C=CC(O)=C3	0.534
molécula 158	5.812479279	CN([C@@H](C1=CC=CC2=C1C=CC=C2)C)CC3=CC=CC(O)=C3	0.552
molécula 159	6.173925197	CN([C@@H](C1=CC=CC2=C1C=CC=C2)C)CC3=CC=C(O)C=C3	0.921
molécula 160	5.187755303	CN([C@@H](C1=CC=CC2=C1C=CC=C2)C)CC3=CC(F)=C(O)C(F)=C3	0.196
molécula 161	6.045757491	CN([C@@H](C1=CC=CC2=C1C=CC=C2)C)CC3=C(C)C=C(O)C=C3	0.674
molécula 162	6.096910013	CN([C@@H](C1=CC=CC2=C1C=CC=C2)C)CC3=C(Cl)C=C(O)C=C3	0.929
molécula 163	6.251811973	CN([C@@H](C1=CC=CC2=C1C=CC=C2)C)CC3=C(OC)C=C(O)C=C3	1.074

molécula 164	5.051098239	<chem>CN([C@@H](C1=CC=CC2=C1C=CC=C2)C)CC3=CC=C(O)C(O)=C3</chem>	-0.087
molécula 165	5.565431096	<chem>CN([C@@H](C1=CC=CC2=C1C=CC=C2)C)CC3=CC(O)=CC(O)=C3</chem>	0.487
molécula 166	5.259637311	<chem>CN([C@@H](C1=CC=CC2=C1C=CC=C2)C)CC3=NC=C(O)C=C3</chem>	0.161
molécula 167	5.213248578	<chem>CN([C@@H](C1=CC=CC2=C1C=CC=C2)C)CC3=NC=CC(S(=O)(N)=O)=C3</chem>	-0.452
molécula 168	6.050609993	<chem>CN([C@@H](C1=CC=CC2=C1C=CC=C2)C)CC3=C(C)C=CC(N)=C3</chem>	0.516
molécula 169	4.780153614	<chem>CN([C@@H](C1=CC=CC2=C1C=CC=C2)C)CC3=CN=C(N)N=C3</chem>	-0.563
molécula 170	5.354577731	<chem>CN([C@@H](C1=CC=CC2=C1C=CC=C2)C)CC3=CC=C(NC(C)=O)C=C3</chem>	0.027
molécula 171	5.841637508	<chem>CN([C@@H](C1=CC=CC2=C1C=CC=C2)C)CC3=C(F)C=C(N)C=C3F</chem>	0.433
molécula 172	5.035269079	<chem>CN([C@@H](C1=CC=CC2=C1C=CC=C2)C)CC3=CC=C(N(CCO)C)C=C3</chem>	-0.404
molécula 173	6.744727495	<chem>BrC1=CC=C2C(C(C(C)=O)=C(C)N3)=C3C(C2=O)=O)=C1</chem>	1.261
molécula 174	6.552841969	<chem>O=C(C1=CC=CC=C1C2=C3N(CC4=CC=CC=C4)C(CN</chem>	0.486

		<chem>C(C)C=C2C(OCC)=O)C3=O</chem>	
molécula 175	5.982966661	<chem>O=C(C(O)=C1C2(C)C)C3=CC=CC=C3C1=NC2=O</chem>	0.633
molécula 176	5.886056648	<chem>O=C1C(NC(C)=O)=CC(C2=NON=C21)=O</chem>	0.426
molécula 177	4.9531148	<chem>ClC1=C(Cl)C(C2=NON=C2C1=O)=O</chem>	-
molécula 178	4.94577009	<chem>O=C(C1=C(C2=C3C=CC=N2)C=CC=C1)C3=O</chem>	0.068
molécula 179	4.747389659	<chem>O=C(OCC)C1=C(C)N(C)C(C2=O)=C1C3=CC(Br)=CC=C3C2=O</chem>	-0.845
molécula 180	4.769551079	<chem>OC1=C2C(C=CC=C2)=CN=C1Br</chem>	0.279
molécula 181	4.721246399	<chem>CC1=CC(N)=C(O)C2=C1C=CC=N2</chem>	0.095
molécula 182	5.68613278	<chem>O=C(OC)C1=CC(C2=CC=CC=C2)=NC(C1C3=O)C(N4CCCC4)=CC3=O</chem>	0.118
molécula 183	5.681936665	<chem>O=C(C1=NON=C12)C(NC3=CC=CC=C3)=C(NC(C)=O)C2=O</chem>	0.435
molécula 184	4.644931794	<chem>CN(CN1C)C2=C(C1=CC=C3)C3=CC=C2</chem>	-0.339
molécula 185	4.607303047	<chem>O=C1C(Cl)=C(N2CCCC2)C(C3=C1C=CC=C3)=O</chem>	-0.201
molécula 186	4.589560214	<chem>NC1=CC2=NSC=C2C=C1</chem>	-0.722
molécula 187	4.543026986	<chem>NC1=CN(CC2CCOCC2)N=C1</chem>	-0.35
molécula 188	4.520280765	<chem>CCN1C(C=C(O)N=C2O)=C2C(N)=N1</chem>	-0.264

molécula 189*	5.364516253	<chem>O=C(C1=O)C=C(N2CCOCC2)C3=C1C=CC=C3</chem>	0.15
molécula 190	5.333482019	<chem>OC1=C(Br)C2=CC=CC=C2C=N1</chem>	0.68
molécula 191*	5.294136288	<chem>O=C(C1=CC=CC=C1C2=C3N(C)C(C)=C2C(OCC)=O)C3=O</chem>	-0.219
molécula 192*	5.136082623	<chem>O=[N+](O-)[C1=CC(C)=C(N(C)C(C2=O)=O)C2=C1</chem>	-0.351
molécula 193	5.055024092	<chem>O=C1NC(C2=CC=CC=C2C1)=O</chem>	0.108
molécula 194	5.044793462	<chem>NC(C(C(N1CC2=C(C=CC=C2)=O)=C(C)C=C1O)=O</chem>	0.316
molécula 195	5.033858267	<chem>O=C(NC1=O)C2=CC=CC=C2C1=O</chem>	-0.137
molécula 196*	6.750312572	<chem>C[C@H](C1=CC(C2=CC=CS2)=CC(C3=CC=CS3)=C1)NC(C4=C(C)C=CC(N)=C4)=O</chem>	0.131
molécula 197	6.76649624	<chem>C[C@H](C1=CC(C2=CC=CS2)=CC(C3=CC=CS3)=C1)NC(C4=C(C)C=CC(NC5CNC5)=C4)=O</chem>	0.167
molécula 198	6.468776625	<chem>C[C@H](C1=CC(C2=CC=CS2)=CC(C3=CC=CS3)=C1)NC(C4=C(C)C=CC(NCC5CNC5)=C4)=O</chem>	-0.207
molécula 199*	6.734946211	<chem>C[C@H](C1=CC(C2=CC=CS2)=CC(C3=CC=CS3)=C1)NC(C4=C(C)C=CC(NC5CCN(CC5)C)=C4)=O</chem>	-0.366

molécula 200	6.93330145	<chem>C[C@H](C1=CC(C2=CC=CS2)=CC(C3=CC=CS3)=C1)N(C(C4=C(C)C=CC(NCCN5CCCCC5)=C4)=O</chem>	-0.063
molécula 201	6.83209219	<chem>C[C@H](C1=CC(C2=CC=CS2)=CC(C3=CC=CS3)=C1)N(C(C4=C(C)C=CC(NC(CN)=O)=C4)=O</chem>	0.095
molécula 202	6.927382523	<chem>C[C@H](C1=CC(C2=CC=CS2)=CC(C3=CC=CS3)=C1)N(C(C4=C(C)C=CC(NC(CN(C)C)=O)=C4)=O</chem>	0.253
molécula 203	6.788345599	<chem>C[C@H](C1=CC(C2=CC=CS2)=CC(C3=CC=CS3)=C1)N(C(C4=C(C)C=CC(NC(C5CCNCC5)=O)=C4)=O</chem>	-0.143
molécula 204	7.186419011	<chem>C[C@H](C1=CC(C2=CC=CS2)=CC(C3=CC=CS3)=C1)N(C(C4=C(C)C=CC(OCCN)=C4)=O</chem>	0.266
molécula 205*	7.179142011	<chem>C[C@H](C1=CC(C2=CC=CS2)=CC(C3=CC=CS3)=C1)N(C(C4=C(C)C=CC(OCCN(C)C)=C4)=O</chem>	0.342
molécula 206	7.119758224	<chem>C[C@H](C1=CC(C2=CC=CS2)=CC(C3=CC=CS3)=C1)N(C(C4=C(C)C=CC(OCCN5CCCCC5)=C4)=O</chem>	-0.186
molécula 207	6.435215615	<chem>C[C@H](C1=CC(C2=CC=CS2)=CC(C3=CC=CS3)=C1)N</chem>	-0.532

		<chem>C(C4=C(C)C=CC(OCCN5CCOCC5)=C4)=O</chem>	
molécula 208	6.877456476	<chem>C[C@@H](NC(C1=C(C=CC(OCC[N+](C)(C)C)=C1)C)=O)C2=CC(C3=CC=CS3)=CC(C4=CC=CS4)=C2</chem>	-0.008
molécula 209	7.003050752	<chem>C[C@H](C1=CC(C2=CSC=C2)=CC(C3=CSC=C3)=C1)NC(C4=C(C)C=CC(OCCN(C)C)=C4)=O</chem>	-0.102
molécula 210*	6.844056982	<chem>C[C@H](C1=CC(C2=CC=C(S2)C)=CC(C3=CC=C(S3)C)=C1)NC(C4=C(C)C=CC(OCCN(C)C)=C4)=O</chem>	-0.155
molécula 211	6.694218849	<chem>C[C@H](C1=CC(C2=CC=CO2)=CC(C3=CC=CO3)=C1)NC(C4=C(C)C=CC(OCCN(C)C)=C4)=O</chem>	0.182
molécula 212	7.077274542	<chem>C[C@H](C1=CC(C2=COC=C2)=CC(C3=COC=C3)=C1)NC(C4=C(C)C=CC(OCCN(C)C)=C4)=O</chem>	0.334
molécula 213	6.962573502	<chem>C[C@H](C1=CC(C2=CN=CS2)=CC(C3=CN=CS3)=C1)NC(C4=C(C)C=CC(OCCN(C)C)=C4)=O</chem>	0.053
molécula 214	6.987162775	<chem>C[C@H](C1=CC(C2=CC=CC=C2)=CC(C3=CC=CC=C3)=C1)NC(C4=C(C)C</chem>	0.881

		<chem>=CC(OCCN(C)C)=C4)=O</chem>	
molécula 215	6.776504059	<chem>C[C@H](C1=CC(C2=CC=C(C=C2)O)=CC(C3=CC=C(C=C3)O)=C1)NC(C4=C(C)C=CC(OCCN(C)C)=C4)=O</chem>	0.647
molécula 216	5.709965389	<chem>C[C@H](C1=CC(C2=CN=C(N=C2)OC)=CC(C3=CN=C(N=C3)OC)=C1)NC(C4=C(C)C=CC(OCCN(C)C)=C4)=O</chem>	-0.498
molécula 217	6.820735536	<chem>C[C@H](C1=CC(C2=CN(N=C2)C)=CC(C3=CN(N=C3)C)=C1)NC(C4=C(C)C=CC(OCCN(C)C)=C4)=O</chem>	-0.033
molécula 218	6.74836178	<chem>C[C@H](C1=CC(C2=CN(N=C2)C(F)F)=CC(C3=CN(N=C3)C(F)F)=C1)NC(C4=C(C)C=CC(OCCN(C)C)=C4)=O</chem>	-0.184
molécula 219	6.408601449	<chem>C[C@H](C1=CC(C2=C(ON=C2C)C)=CC(C3=C(ON=C3)C)=C1)NC(C4=C(C)C=CC(OCCN(C)C)=C4)=O</chem>	-0.319
molécula 220	6.740645073	<chem>C[C@H](C1=CC(C2=CN(N=C2)CCO)=CC(C3=CN(N=C3)CCOC)=C1)NC(C4=C(C)C=CC(OCCN(C)C)=C4)=O</chem>	-0.299
molécula 221*	6.781727146	<chem>C[C@H](C1=CC(C2=CN(N=C2)CCO)=CC(C3=CN(N=C3)CCO)=C1)NC(C4=C(C)C=CC(OCCN(C)C)=C4)=O</chem>	-0.12

molécula 222*	6.353204311	<chem>C[C@H](C1=CC(C2=CC=NN2CC)=C(C3=CC=NN3CC)=C1)NC(C4=C(C)C=CC(OCCN(C)C)=C4)=O</chem>	-0.312
molécula 223	6.333856573	<chem>C[C@H](C1=CC(C2=CC=NN2C)=CC(C3=CC=NN3C)=C1)NC(C4=C(C)C=CC(OCCN(C)C)=C4)=O</chem>	-0.278
molécula 224	6.816160963	<chem>O=C(C1=C(C)C=CC(OCCN(C)C)=C1)N[C@@H](C2=CC(C3=CC=CS3)=CC(C4=CN(N=C4)C)=C2)C</chem>	-0.016
molécula 225	7.086716098	<chem>O=C(C1=C(C)C=CC(OCCN(C)C)=C1)N[C@@H](C2=CC(C3=CC=CS3)=CC(C4=CN(N=C4)=C2)C</chem>	0.351
molécula 226	7.126679398	<chem>O=C(C1=C(C)C=CC(OCCN(C)C)=C1)N[C@@H](C2=CC(C3=CC=CS3)=CC(C4=CN(N=C4)C(F)F)=C2)C</chem>	0.188
molécula 227	6.93404702	<chem>O=C(C1=C(C)C=CC(OCCN(C)C)=C1)N[C@@H](C2=CC(C3=CC=CS3)=CC(C4=CN(N=C4)CCN(C)C)=C2)C</chem>	-0.078
molécula 228*	6.951558196	<chem>O=C(C1=C(C)C=CC(OCCN(C)C)=C1)N[C@@H](C2=CC(C3=CC=CS3)=CC(C4=CN(N=C4)CCOC)=C2)C</chem>	-0.006
molécula 229*	6.945004138	<chem>O=C(C1=C(C)C=CC(OCCN(C)C)=C1)</chem>	0.061

		<chem>N[C@@H](C2=CC(C3=CC=CS3)=CC(C4=CN(N=C4)CCO)=C2)C</chem>	
molécula 230	7.086716098	<chem>O=C(C1=C(C)C=C(C(OCCN(C)C)=C1)N[C@@H](C2=CC(C3=CC=CS3)=CC(C4=CN(N=C4)CCN)=O)=C2)C</chem>	0.231
molécula 231	7.131943638	<chem>O=C(C1=C(C)C=C(C(OCCN(C)C)=C1)N[C@@H](C2=CC(C3=CC=CS3)=CC(C4=CN(N=C4)CC(F)(F)F)=C2)C</chem>	0.151
molécula 232*	7.041436117	<chem>O=C(C1=C(C)C=C(C(OCCN(C)C)=C1)N[C@@H](C2=CC(C3=CC=CS3)=CC(C4=CC=C(S4)CN)=C2)C</chem>	0.106
molécula 233	7.177831921	<chem>O=C(C1=C(C)C=C(C(OCCN(C)C)=C1)N[C@@H](C2=CC(C3=CC=CS3)=CC(C4=CN(N=C4)CCN5CCOCC5)=C2)C</chem>	0.038
molécula 234	7.038578906	<chem>O=C(C1=C(C)C=C(C(OCCN(C)C)=C1)N[C@@H](C2=CC(C3=CC=CS3)=CC(C4=CN=C(C=C4)N5CCOCC5)=C2)C</chem>	0.273
molécula 235	6.964570262	<chem>O=C(C1=C(C)C=C(C(OCCN(C)C)=C1)N[C@@H](C2=CC(C3=CN(C)N=C3)=CC(C4=CN=CS4)=C2)C</chem>	0.144
molécula 236	6.920095532	<chem>O=C(N[C@H](C)C1=CC(C2=CN(N=C2)C)=CC(C3=CC=</chem>	-0.006

		<chem>C(O3)CN4CCOCC4)=C1)C5=C(C)C=CC(OCCN(C)C)=C5</chem>	
molécula 237	7.135488919	<chem>O=C(N[C@H](C)C1=CC(C2=CN(N=C2)C)=CC(C3=CN(N=C3)CC(N)=O)=C1)C4=C(C)C=CC(OCCN(C)C)=C4</chem>	0.308
molécula 238	6.927382523	<chem>O=C(N[C@H](C)C1=CC(C2=CN(N=C2)C)=CC(C3=CN(N=C3)C(C)=O)=C1)C4=C(C)C=CC(OC CN(C)C)=C4</chem>	0.103
molécula 239	7.135133086	<chem>O=C(N[C@H](C)C1=CC(C2=CN(N=C2)C)=CC(C3=CN(N=C3)CC4CCOCC4)=C1)C5=C(C)C=C C(OCCN(C)C)=C5</chem>	0.076
molécula 240	7.005243055	<chem>CN1N=CC(C2=CC(C3=CN(N=C3)CC N4CCOCC4)=CC([C@H](NC(C5=C(C)C=CC(OCCN(C)C)=C5)=O)C)=C2)=C1</chem>	-0.089
molécula 241	6.963371105	<chem>CN1N=CC(C2=CC(C3=CN(N=C3)C(C)C)=CC([C@H](NC(C4=C(C)C=CC(OCCN(C)C)=C4)=O)C)=C2)=C1</chem>	0.077
molécula 242	6.988429556	<chem>CN1N=CC(C2=CC(C3=CN(N=C3)C4CC4)=CC([C@H](NC(C5=C(C)C=CC(OCCN(C)C)=C5)=O)C)=C2)=C1</chem>	0.134
molécula 243	6.950007143	<chem>CN1N=CC(C2=CC(C3=CN(N=C3)CC(F)(F)F)=CC([C@H](</chem>	0.028

		<chem>NC(C4=C(C)C=CC(OCCN(C)C)=C4)=O)C=C2)=C1</chem>	
molécula 244*	6.877784122	<chem>CN1N=CC(C2=CC(C3=CC=C(C4=C3)C=NN4C)=CC([C@H](NC(C5=C(C)C=CC(OCCN(C)C)=C5)=O)C)=C2)=C1</chem>	-0.133
molécula 245	7.055024092	<chem>CN1N=CC(C2=CC(C3=CN=C(N4CCOCC4)C=C3)=CC([C@@H](C)NC(C5=C(C=CC(OCCN(C)C)=C5)C)=O)=C2)=C1</chem>	0.327
molécula 246	7.007446482	<chem>CN1N=CC(C2=CC(C3=CN(C(F)F)N=C3)=CC([C@@H](C)NC(C4=C(C=CC(OCCN(C)C)=C4)C)=O)=C2)=C1</chem>	0.128
molécula 247	7.038104526	<chem>CC(C=CC(OCCN(C)C)=C1)=C1C(N[C@H](C)C2=CC(C3=CN(N=C3)CCO)=CC(C4=CN(C(F)F)N=C4)=C2)=O</chem>	0.103
molécula 248	6.78994915	<chem>CC(C=CC(OCCN(C)C)=C1)=C1C(N[C@H](C)C2=CC(C3=CN(N=C3)C(C)C)=CC(C4=CN(C(F)F)N=C4)=C2)=O</chem>	-0.17
molécula 249	6.951169913	<chem>CC(C=CC(OCCN(C)C)=C1)=C1C(N[C@H](C)C2=CC(C3=CN(N=C3)CCO)=CC(C4=CN(C(F)F)N=C4)=C2)=O</chem>	-0.048
molécula 250	6.915423722	<chem>CC(C=CC(OCCN(C)C)=C1)=C1C(N[C@H](C)C2=CC(C3=CN(N=C3)COC)</chem>	-0.033

		<chem>=CC(C4=CN(C(F)F)N=C4)=C2)=O</chem>	
molécula 251	7.015022874	<chem>CC(C=CC(OCCN(C)C)=C1)=C1C(N[C@H](C)C2=CC(C3=CN(N=C3)C4CC4)=CC(C5=CN(C(F)F)N=C5)=C2)=O</chem>	0.087
molécula 252*	6.838631998	<chem>CC(C=CC(OCCN(C)C)=C1)=C1C(N[C@H](C)C2=CC(C3=CN(N=C3)CCN(C)C)=CC(C4=CN(CCO)N=C4)=C2)=O</chem>	-0.174
molécula 253	7.140261434	<chem>CC(C=CC(OCCN(C)C)=C1)=C1C(N[C@H](C)C2=CC(C3=CC=C(CN4CCC4)S3)=CC(C5=C(N(CCO)N=C5)=C2)=O</chem>	-0.382
molécula 254	6.991399828	<chem>CC(C=CC(OCCN(C)C)=C1)=C1C(N[C@H](C)C2=CC(C3=CN(N=C3)CCO)C)=CC(C4=CN(CCO)N=C4)=C2)=O</chem>	0.021
molécula 255	7.034328029	<chem>CN1N=CC(C2=CC(C3=CN(CCO)N=C3)=CC([C@@H](C)NC(C4=C(C=CC(OCCN(C)C)=C4)C)=O)=C2)=C1</chem>	0.183
molécula 256*	6.847711656	<chem>CC(C=CC(OCCN(C)C)=C1)=C1C(N[C@H](C)C2=CC(C3=CC=NN3C)=CC(C4=CN(CCO)N=C4)=C2)=O</chem>	0.115
molécula 257	7.04349541	<chem>CC(C=CC(OCCN(C)C)=C1)=C1C(N[C@H](C)C2=CC(C3=CN(CCN)=O)N=</chem>	0.094

		<chem>C3)=CC(C4=CN(CCOC)N=C4)=C2)=O</chem>	
molécula 258	6.957424488	<chem>CC(C=CC(OCCN(C)C)=C1)=C1C(N[C@H](C)C2=CC(C3=CN=CS3)=CC(C4=CN(CCOC)N=C4)=C2)=O</chem>	0.011
molécula 259	6.684659523	<chem>CC(C=CC(OCCN(C)C)=C1)=C1C(N[C@H](C)C2=CC(C3=CN(CC(F)F)N=C3)=CC(C4=CN(CCOC)N=C4)=C2)=O</chem>	-0.33
molécula 260	6.830325566	<chem>CC(C=CC(OCCN(C)C)=C1)=C1C(N[C@H](C)C2=CC(C3=NN(C)C=C3)=CC(C4=NN(C)C=C4)=C2)=O</chem>	0.212
molécula 261	7.007004902	<chem>CC(C=CC(OCCN(C)C)=C1)=C1C(N[C@H](C)C2=CC(C3=CN(CCOC)N=C3)=CC(C4=NN(C)C=C4)=C2)=O</chem>	0.194
molécula 262	7.014124643	<chem>CC(C=CC(OCCN(C)C)=C1)=C1C(N[C@H](C)C2=CC(C3=CN(C(F)F)N=C3)=CC(C4=NN(C)C=C4)=C2)=O</chem>	0.244
molécula 263	6.949233689	<chem>CC(C=CC(OCCN(C)C)=C1)=C1C(N[C@H](C)C2=CC(C3=CN(C(C)C)N=C3)=CC(C4=NN(C)C=C4)=C2)=O</chem>	0.169
molécula 264	6.906928694	<chem>CC(C=CC(OCCN(C)C)=C1)=C1C(N[C@H](C)C2=CC(C3=CN(C4CC4)N=C</chem>	0.158

		<chem>3)=CC(C5=NN(C)C=C5)=C2)=O</chem>	
molécula 265	6.932929144	<chem>CC1=C(C(N[C@@H](C2=CC(C3=CN(CCN4CCOCC4)N=C3)=CC(C5=NN(C=C5)C)=C2)C)=O)C=C(OCCN(C)C)C=C1</chem>	-0.052
molécula 266*	7.080398976	<chem>CC1=C(C(N[C@@H](C2=CC(C3=CN(CC4CCOCC4)N=C3)=CC(C5=NN(C=C5)C)=C2)C)=O)C=C(OCCN(C)C)C=C1</chem>	0.13
molécula 267	7.061980903	<chem>CC1=C(C(N[C@@H](C2=CC(C3=CN(CF)N=C3)=CC(C4=NN(C=C4)C)=C2)C)=O)C=C(OCCN(C)C)C=C1</chem>	0.335
molécula 268	6.913996294	<chem>CC1=C(C(N[C@@H](C2=CC(C3=CN(C)N=C3)=CC(C4=NN(C=C4)C)=C2)C)=O)C=C(OCCN(C)C)C=C1</chem>	0.172
molécula 269*	6.488250289	<chem>CC1=C(C(N[C@@H](C2=CC(C3=CN(C(F)F)N=C3)=CC(C4=NN(C=C4)CC)=C2)C)=O)C=C(OCCN(C)C)C=C1</chem>	-0.319
molécula 270	6.693574972	<chem>CC1=C(C(N[C@@H](C2=CC(C3=NN(CC)C=C3)=CC(C4=NN(C=C4)CC)=C2)C)=O)C=C(OCCN(C)C)C=C1</chem>	0.03
molécula 271	6.682145511	<chem>CC1=C(C(N[C@@H](C2=CC(C3=CN(C(C)C)N=C3)=CC(C4=NN(C=C4)CC)</chem>	-0.14

		<chem>=C2)C)=O)C=C(OC CN(C)C)C=C1</chem>	
molécula 272	6.578725209	<chem>CC1=C(C(N[C@@ H])(C2=CC(C3=CN(C4COC4)N=C3)=C C(C5=NN(C=C5)C C)=C2)C)=O)C=C(OCCN(C)C)C=C1</chem>	-0.017
molécula 273*	6.633203617	<chem>CC1=C(C(N[C@@ H])(C2=CC(C3=CN(CC(F)F)N=C3)=CC (C4=NN(C=C4)CC) =C2)C)=O)C=C(OC CN(C)C)C=C1</chem>	-0.199
molécula 274	6.473531488	<chem>CC1=C(C(N[C@@ H])(C2=CC(C3=NN(C)C=C3)=CC(C4= NN(C=C4)CC)=C2) C)=O)C=C(OCCN(C)C)C=C1</chem>	-0.187
molécula 275	6.36081244	<chem>CC1=C(C(N[C@@ H])(C2=CC(C3=NN(C(F)F)C=C3)=CC(C4=NN(C=C4)CC) =C2)C)=O)C=C(OC CN(C)C)C=C1</chem>	-0.325
molécula 276	6.652279783	<chem>CC1=C(C(N[C@@ H])(C2=CC(C3=NN(CC(F)F)C=C3)=CC (C4=NN(C=C4)CC) =C2)C)=O)C=C(OC CN(C)C)C=C1</chem>	-0.047
molécula 277	6.567993313	<chem>CC1=C(C(N[C@@ H])(C2=CC(C3=CN(CC(F)(F)F)N=C3)= CC(C4=NN(C=C4) CC)=C2)C)=O)C=C (OCCN(C)C)C=C1</chem>	-0.275
molécula 278	6.971428747	<chem>CC1=C(C(N[C@@ H])(C2=CC(C3=CN(C)N=C3)=CC(C4= NN(C=C4)CC)=C2) C)=O)C=C(OCCN(C)C)C=C1</chem>	0.187

molécula 279	6.619426997	<chem>CC1=C(C(N[C@@H](C2=CC(C3=CN(COC)N=C3)=CC(C4=NN(C=C4)CC)=C2)C)=O)C=C(OCN(C)C)C=C1</chem>	-0.176
molécula 280	6.783570169	<chem>CC1=C(C(N[C@@H](C2=CC(C3=CN(COC)N=C3)=CC(C4=NN(C=C4)CC)=C2)C)=O)C=C(OCN(C)C)C=C1</chem>	0.702

Anexo II

Molécula	Familia	Aldehído	Cetona	SMILES	pIC50 QSAR Evolutivo	pIC50 QSAR manual
mol_001	I Chalcona	p- Nitrobenz aldehído	1- Acetilpire no	<chem>O=C(/C=C/C/C1=CC=C(C=C1)[N+](=[O-])=O)C2=C3C=CC4=CC=C5=CC=C(C=C2)C3=C45</chem>	5.561311 136	7.510226 199
mol_002	II Aminopiri midina	p- Nitrobenz aldehído	1- Acetilpire no	<chem>[O-][N+](C(C=C1)=CC=C1C2=NC(N)=NC(C3=C(C=CC4=CC=C5)C6=C4C5=C=C6C=C3)=C2)=O</chem>	5.574015 118	7.287869 561
mol_003R	III 2,3- dihidrobe nzodiaze pina (R)	p- Nitrobenz aldehído	1- Acetilpire no	<chem>[O-][N+](C1=CC=C(C2CC(C3=C</chem>	5.877305 379	7.586449 125

				<chem>C=C4C5=C3C=CC6=C5C(C=C4)=CC=C6)=NC7=C(N2)C=CC=C7)C=C1)=O</chem>		
mol_003S	III 2,3-dihidrobenzodiazepina (S)	p-Nitrobenzaldehído	1-Acetilpireno	<chem>[O-][N+](C1=CC=C(C2CC(C3=C=C4C5=C3C=CC6=C5C(C=C4)=CC=C6)=NC7=C(N2)C=CC=C7)C=C1)=O</chem>	5.877305379	7.586461284
mol_004R	IV 3,4-dihidropirazol (R)	p-Nitrobenzaldehído	1-Acetilpireno	<chem>[O-][N+](C1=CC=C(C2NN=C(C2)C3=C4C=CC5=C=C6C=CC=C(C=C3)C4=C56)C=C1)=O</chem>	6.025335866	6.999739482
mol_004S	IV 3,4-dihidropirazol (S)	p-Nitrobenzaldehído	1-Acetilpireno	<chem>[O-][N+](C1=CC=C(C2NN=C(C2)C3=C4C=CC5=C=C6C=CC=C(C=C3)C4=C56)C=C1)=O</chem>	6.025335866	6.997424173

mol_005	V Pirimidina - aminobenzimidazol	p- Nitrobenzaldehído	1- Acetilpireno	<chem>[O-][N+](C1=CC=C(C2=NC(NC3=NC4=C(N3)C=CC=C4)=NC(C5=C6C=CC7=CC=CC8=C(C=C(C=C5)C6=C78)=C2)C=C1)=O</chem>	6.067876 693	7.569442 099
mol_006	VI Pirimidobenzimidazol	p- Nitrobenzaldehído	1- Acetilpireno	<chem>O=[N+](C1=CC=C(C2=CC(C3=C4C=CC5=C6C(C=CC(C=C3)=C46)=CC=C5)=NC7=NC8=CC=CC=C8N27)C=C1)[O-]</chem>	6.102668 705	7.979342 618
mol_007	I Chalcona	p- Nitrobenzaldehído	9- Acetilantraceno	<chem>O=C(C1=C2C(C=C(C=C2)=CC3=C1C=CC=C3)/C=C/C4=CC=C([N+])([O-])=O)C=C4</chem>	5.304011 325	7.128897 374
mol_008	II Aminopirimidina	p- Nitrobenzaldehído	9- Acetilantraceno	<chem>[O-][N+](C(C=C1)=CC=C1C2=NC(N)=NC(C3=C4C(C=CC=C3)=C4</chem>	5.294559 61	6.919304 713

				4)=CC5=C3C=CC=C5)=C2)=O		
mol_009R	III 2,3-dihidrobenzodiazepina (R)	p-Nitrobenzaldehído	9-Acetilantraceno	[O-][N+](C1=CC=C(C2CC(C3=C4C(C=CC=C4)=CC5=C3C=C(C=C5)=NC6=C(N2)C=CC=C6)C=C1)=O	5.603013333	7.472569206
mol_009S	III 2,3-dihidrobenzodiazepina (S)	p-Nitrobenzaldehído	9-Acetilantraceno	[O-][N+](C1=CC=C(C2CC(C3=C4C(C=CC=C4)=CC5=C3C=C(C=C5)=NC6=C(N2)C=CC=C6)C=C1)=O	5.603013333	7.472569206
mol_010R	IV 3,4-dihidropirazol (R)	p-Nitrobenzaldehído	9-Acetilantraceno	[O-][N+](C1=CC=C(C2NN=C(C2)C3=C4C(C=CC=C4)=CC5=C3C=CC=C5)C=C1)=O	5.814867713	6.621150243
mol_010S	IV 3,4-dihidropirazol (S)	p-Nitrobenzaldehído	9-Acetilantraceno	[O-][N+](C1=CC=C(C2NN=C(C2)C3=C4C(C=CC=C4)=CC5=C3C=CC=C5)C=C1)=O	5.814867713	6.618772634

				4)=CC5=C3C=CC=C5)C=C1)=O		
mol_011	V Pirimidina - aminobenzimidazol	p- Nitrobenzaldehído	9- Acetilantraceno	[O-][N+](C1=CC=C(C2=NC(NC3=NC4=C(N3)C=CC=C4)=NC(C5=C6C(C=CC=C6)=CC7=C5C=CC=C7)=C2)C=C1)=O	5.788831558	8.255850468
mol_012	VI Pirimidobenzimidazol	p- Nitrobenzaldehído	9- Acetilantraceno	O=[N+](C1=CC=C(C2=CC(C3=C4C=CC=CC4=CC5=C3C=CC=C5)=NC6=NC7=CC=CC=C7N26)C=C1)[O-]	5.837872049	7.598362706
mol_013	I Chalcona	p- Nitrobenzaldehído	1- Acetilnaftaleno	O=C(C1=C2C(C=C(C=C2))=C(C=C1)/C=C/C3=CC=C([N+])([O-])=O)C=C3	4.938657264	6.391564246
mol_014	II Aminopirimidina	p- Nitrobenzaldehído	1- Acetilnaftaleno	[O-][N+](C(C=C1)=CC=C1C2=NC(N)=NC(C3=CC=	4.984653055	6.489056429

				<chem>CC4=C3C=CC=C4)=O</chem>		
mol_015R	III 2,3-dihidrobenzodiazepina (R)	p-Nitrobenzaldehído	1-Acetilnaftaleno	<chem>[O-][N+](C1=CC=C(C2CC(C3=C4C(C=CC=C4)=CC=C3)=NC5=C(N2)C=CC=C5)C=C1)=O</chem>	5.275016852	6.736605877
mol_015S	III 2,3-dihidrobenzodiazepina (S)	p-Nitrobenzaldehído	1-Acetilnaftaleno	<chem>[O-][N+](C1=CC=C(C2CC(C3=C4C(C=CC=C4)=CC=C3)=NC5=C(N2)C=CC=C5)C=C1)=O</chem>	5.275016852	6.736605877
mol_016R	IV 3,4-dihidropirazol (R)	p-Nitrobenzaldehído	1-Acetilnaftaleno	<chem>[O-][N+](C1=CC=C(C2NN=C(C2)C3=CC=CC4=C3C=CC=C4)C=C1)=O</chem>	5.467943623	6.977782224
mol_016S	IV 3,4-dihidropirazol (S)	p-Nitrobenzaldehído	1-Acetilnaftaleno	<chem>[O-][N+](C1=CC=C(C2NN=C(C2)C3=CC=CC4=C3C=CC=C4)C=C1)=O</chem>	5.467943623	6.971686093
mol_017	V Pirimidina -	p-Nitrobenzaldehído	1-Acetilnaftaleno	<chem>[O-][N+](C1=CC=C(C2</chem>	5.444114446	7.530496346

	aminobenzimidazol			<chem>=NC(NC3=NC4=C(N3)C=CC=C4)=NC(C5=CC=CC6=C5C=CC=C6)=C2)C=C1)=O</chem>		
mol_018	VI Pirimido-benzimidazol	p-Nitrobenzaldehído	1-Acetilnaftaleno	<chem>O=[N+](C1=CC=C(C2=CC(C3=C4C=CC=CC4=CC=C3)=NC5=NC6=CC=C(C=C6N25)C=C1)[O-]</chem>	5.473129 952	6.877461 233
mol_019	I Chalcona	p-Nitrobenzaldehído	2-Acetilnaftaleno	<chem>O=C(C1=CC=C(C=CC=C2)C2=C1)/C=C/C3=CC=C([N+])([O-])=O)C=C3</chem>	4.956094 292	6.392761 425
mol_020	II Aminopirimidina	p-Nitrobenzaldehído	2-Acetilnaftaleno	<chem>[O-][N+](C(C=C1)=CC=C1C2=NC(N)=NC(C3=CC(C=CC=C4)=C4C=C3)=C2)=O</chem>	4.996789 588	6.219898 832
mol_021R	III 2,3-dihidrobenzodiazepina (R)	p-Nitrobenzaldehído	2-Acetilnaftaleno	<chem>[O-][N+](C1=CC=C(C2CC(C3=CC=C(C=C3)C=C4)C4</chem>	5.296713 437	6.474538 064

				<chem>=C3)=NC5=C(N2)C=CC=C5)C=C1)=O</chem>		
mol_021S	III 2,3-dihidrobenzodiazepina (S)	p-Nitrobenzaldehído	2-Acetilnaftaleno	<chem>[O-][N+](C1=CC=C(C2CC(C3=CC=C(C=C(C=C4)C4=C3)=NC5=C(N2)C=CC=C5)C=C1)=O</chem>	5.296713437	6.474524629
mol_022R	IV 3,4-dihidropirazol (R)	p-Nitrobenzaldehído	2-Acetilnaftaleno	<chem>[O-][N+](C1=CC=C(C2NN=C(C2)C3=CC4=C(C=C3)C=CC=C4)C=C1)=O</chem>	5.423881891	6.707942723
mol_022S	IV 3,4-dihidropirazol (S)	p-Nitrobenzaldehído	2-Acetilnaftaleno	<chem>[O-][N+](C1=CC=C(C2NN=C(C2)C3=CC4=C(C=C3)C=CC=C4)C=C1)=O</chem>	5.423881891	6.712306642
mol_023	V Pirimidina-aminobenzimidazol	p-Nitrobenzaldehído	2-Acetilnaftaleno	<chem>[O-][N+](C1=CC=C(C2=NC(NC3=NC4=C(N3)C=CC=C4)=NC(C5=CC6=C(C=C5)C=CC=C6)=C2)C=C1)=O</chem>	5.451506616	7.26704882

mol_024	VI Pirimido- benzimid azol	p- Nitrobenz aldehído	2- Acetilnaft aleno	<chem>O=[N+](C1=CC=C(C2=CC(C3=CC=C4C(C=CC=C4)=C3)=NC5=NC6=CC=C(C=C6N25)C=C1)[O-]</chem>	5.474855 499	6.881845 075
mol_025	I Chalcona	p- Nitrobenz aldehído	Acetophe none	<chem>O=C(C1=CC=CC=C1)/C=C/C2=CC=C([N+](O-))=O)C=C2</chem>	4.531064 451	5.681488 109
mol_026	II Aminopiri midina	p- Nitrobenz aldehído	Acetophe none	<chem>[O-][N+](C(C=C1)=CC=C1C2=NC(N)=NC(C3=CC=CC=C3)=C2)=O</chem>	4.648788 601	6.616011 461
mol_027R	III 2,3- dihidrobe nzodiaze pina (R)	p- Nitrobenz aldehído	Acetophe none	<chem>[O-][N+](C1=CC=C(C2CC(C3=CC=CC=C3)=NC4=C(N2)C=CC=C4)C=C1)=O</chem>	4.951048 832	6.830309 755
mol_027S	III 2,3- dihidrobe nzodiaze pina (S)	p- Nitrobenz aldehído	Acetophe none	<chem>[O-][N+](C1=CC=C(C2CC(C3=CC=CC=C3)=NC4=C(N2)C=</chem>	4.951048 832	6.830238 324

				<chem>CC=C4)C=C1)=O</chem>		
mol_028R	IV 3,4-dihidropirazol (R)	p-Nitrobenzaldehído	Acetophenone	<chem>[O-][N+](C1=CC=C(C2=NN=C(C2)C3=CC=CC=C3)C=C1)=O</chem>	5.067018795	6.028811421
mol_028S	IV 3,4-dihidropirazol (S)	p-Nitrobenzaldehído	Acetophenone	<chem>[O-][N+](C1=CC=C(C2=NN=C(C2)C3=CC=CC=C3)C=C1)=O</chem>	5.067018795	6.028030756
mol_029	V Pirimidina-aminobenzimidazol	p-Nitrobenzaldehído	Acetophenone	<chem>[O-][N+](C1=CC=C(C2=NC(NC3=NC4=C(N3)C=CC=C4)=NC(C5=CC=CC=C5)=C2)C=C1)=O</chem>	5.06977716	6.834744563
mol_030	VI Pirimido-benzimidazol	p-Nitrobenzaldehído	Acetophenone	<chem>O=[N+](C1=CC=C(C2=CC(C3=CC=CC=C3)=NC4=NC5=CC=CC=C5N24)C=C1)[O-]</chem>	5.087923109	6.185452311
mol_031	I Chalcona	p-Dimetilaminobenzaldehído	1-Acetilpireno	<chem>CN(C)C1=CC=C(/C=C/C(C2=C(C=C3=CC=C4)C5=C3C4=CC=</chem>	5.660165753	7.738779864

				C5C=C2) =O)C=C1		
mol_032	II Aminopiri midina	p- Dimetila minobenz aldehído	1- Acetilpire no	CN(C)C(C=C1)=C C=C1C2= NC(N)=N C(C3=C4 C(C(C(C= C4)=CC= C5)=C5C =C6)=C6 C=C3)=C 2	5.679547 654	7.562932 602
mol_033R	III 2,3- dihidrobe nzodiaze pina (R)	p- Dimetila minobenz aldehído	1- Acetilpire no	CN(C)C1 =CC=C(C 2CC(C3= CC=C4C 5=C3C=C C6=C5C(C=C4)=C C=C6)=N C7=C(N2) C=CC=C 7)C=C1	5.933612 417	7.819872 505
mol_033S	III 2,3- dihidrobe nzodiaze pina (S)	p- Dimetila minobenz aldehído	1- Acetilpire no	CN(C)C1 =CC=C(C 2CC(C3= CC=C4C 5=C3C=C C6=C5C(C=C4)=C C=C6)=N C7=C(N2) C=CC=C 7)C=C1	5.933612 417	7.823195 231
mol_034R	IV 3,4- dihidropir azol (R)	p- Dimetila minobenz aldehído	1- Acetilpire no	CN(C)C1 =CC=C(C 2NN=C(C 2)C3=C4 C=CC5= CC=CC6 =CC=C(C =C3)C4=	6.059954 999	7.252631 627

				C56)C=C 1		
mol_034S	IV 3,4- dihidropir azol (S)	p- Dimetila minobenz aldehído	1- Acetilpire no	CN(C)C1 =CC=C(C 2NN=C(C 2)C3=C4 C=CC5= CC=CC6 =CC=C(C =C3)C4= C56)C=C 1	6.059954 999	7.253044 217
mol_035	V Pirimidina - aminobe nzimidaz ol	p- Dimetila minobenz aldehído	1- Acetilpire no	CN(C)C1 =CC=C(C 2=NC(NC 3=NC4=C (N3)C=C C=C4)=N C(C5=C6 C=CC7= CC=CC8 =CC=C(C =C5)C6= C78)=C2) C=C1	6.188068 708	7.809471 575
mol_036	VI Pirimido- benzimid azol	p- Dimetila minobenz aldehído	1- Acetilpire no	CN(C)C1 =CC=C(C =C1)C2= CC(C3=C 4C=CC5= C6C(C=C C(C=C3)= C46)=CC =C5)=NC 7=NC8=C C=CC=C 8N27	6.206640 907	8.217081 06
mol_037	I Chalcona	p- Dimetila minobenz aldehído	9- Acetilantr aceno	CN(C)C1 =CC=C(/ C=C/C(C 2=C3C(C =CC=C3) =CC4=C2 C=CC=C	5.450237 852	7.625885 566

				4)=O)C=C 1		
mol_038	II Aminopiri midina	p- Dimetila minobenz aldehído	9- Acetilantr aceno	CN(C)C(C=C1)=C C=C1C2= NC(N)=N C(C3=C4 C(C=CC= C4)=CC5 =C3C=C C=C5)=C 2	5.433039 947	7.455990 127
mol_039R	III 2,3- dihidrobe nzodiaze pina (R)	p- Dimetila minobenz aldehído	9- Acetilantr aceno	CN(C)C1 =CC=C(C 2CC(C3= C4C(C=C C=C4)=C C5=C3C= CC=C5)= NC6=C(N 2)C=CC= C6)C=C1	5.685873 184	7.707935 412
mol_039S	III 2,3- dihidrobe nzodiaze pina (S)	p- Dimetila minobenz aldehído	9- Acetilantr aceno	CN(C)C1 =CC=C(C 2CC(C3= C4C(C=C C=C4)=C C5=C3C= CC=C5)= NC6=C(N 2)C=CC= C6)C=C1	5.685873 184	7.709247 742
mol_040R	IV 3,4- dihidropir azol (R)	p- Dimetila minobenz aldehído	9- Acetilantr aceno	CN(C)C1 =CC=C(C 2NN=C(C 2)C3=C4 C(C=CC= C4)=CC5 =C3C=C C=C5)C= C1	5.899687 431	7.145082 943
mol_040S	IV 3,4- dihidropir azol (S)	p- Dimetila	9- Acetilantr aceno	CN(C)C1 =CC=C(C 2NN=C(C	5.899687 431	7.142277 469

		minobenz aldehído		2)C3=C4 C(C=CC= C4)=CC5 =C3C=C C=C5)C= C1		
mol_041	V Pirimidina - aminobe nzimidaz ol	p- Dimetila minobenz aldehído	9- Acetilantr aceno	CN(C)C1 =CC=C(C 2=NC(NC 3=NC4=C (N3)C=C C=C4)=N C(C5=C6 C(C=CC= C6)=CC7 =C5C=C C=C7)=C 2)C=C1	5.931152 517	8.764722 237
mol_042	VI Pirimido- benzimid azol	p- Dimetila minobenz aldehído	9- Acetilantr aceno	CN(C)C1 =CC=C(C =C1)C2= CC(C3=C 4C=CC= CC4=CC 5=C3C=C C=C5)=N C6=NC7= CC=CC= C7N26	5.968103 171	8.108738 847
mol_043	I Chalcona	p- Dimetila minobenz aldehído	1- Acetilnaft aleno	CN(C)C1 =CC=C(/ C=C/C(C 2=CC=C C3=C2C= CC=C3)= O)C=C1	5.065845 583	6.909491 808
mol_044	II Aminopiri midina	p- Dimetila minobenz aldehído	1- Acetilnaft aleno	CN(C)C(C=C1)=C C=C1C2= NC(N)=N C(C3=CC =CC4=C3 C=CC=C 4)=C2	5.114140 912	7.045638 846

mol_045R	III 2,3-dihidrobenzodiazepina (R)	p-Dimetilaminobenzaldehído	1-Acetilnaftaleno	<chem>CN(C)C1=CC=C(C2CC(C3=CC=CC4=C3C=CC=C4)=NC5=C(N2)C=CC=C5)C=C1</chem>	5.336689167	7.259601303
mol_045S	III 2,3-dihidrobenzodiazepina (S)	p-Dimetilaminobenzaldehído	1-Acetilnaftaleno	<chem>CN(C)C1=CC=C(C2CC(C3=CC=CC4=C3C=CC=C4)=NC5=C(N2)C=CC=C5)C=C1</chem>	5.336689167	7.26330152
mol_046R	IV 3,4-dihidropirazol (R)	p-Dimetilaminobenzaldehído	1-Acetilnaftaleno	<chem>CN(C)C1=CC=C(C2NN=C(C2)C3=CC=CC4=C3C=CC=C4)C=C1</chem>	5.503882502	7.521714922
mol_046S	IV 3,4-dihidropirazol (S)	p-Dimetilaminobenzaldehído	1-Acetilnaftaleno	<chem>CN(C)C1=CC=C(C2NN=C(C2)C3=CC=CC4=C3C=CC=C4)C=C1</chem>	5.503882502	7.519244247
mol_047	V Pirimidina-aminobenzimidazol	p-Dimetilaminobenzaldehído	1-Acetilnaftaleno	<chem>CN(C)C1=CC=C(C2=NC(NC3=NC4=C(N3)C=CC(C5=CC=CC6=C5C=CC=C6)=C2)C=C1</chem>	5.590960803	8.059624249

mol_048	VI Pirimido- benzimid azol	p- Dimetila minobenz aldehído	1- Acetilnaft aleno	<chem>CN(C)C1=CC=C(C=C1)C2=CC(C3=CC=CC4=C3C=CC=C4)=NC5=NC6=C(C=CC=C6)N25</chem>	5.600471 06	7.409275 643
mol_049	I Chalcona	p- Dimetila minobenz aldehído	2- Acetilnaft aleno	<chem>CN(C)C1=CC=C(/C=C/C(C2=CC(C=CC=C3)=C3C=C2)=O)C=C1</chem>	5.086391 755	6.907739 399
mol_050	II Aminopiri midina	p- Dimetila minobenz aldehído	2- Acetilnaft aleno	<chem>CN(C)C(C=C1)=CC=C1C2=NC(N)=NC(C3=CC(C=CC=C4)=C4C=C3)=C2</chem>	5.136382 36	6.779172 018
mol_051R	III 2,3- dihidrobe nzodiaze pina (R)	p- Dimetila minobenz aldehído	2- Acetilnaft aleno	<chem>CN(C)C1=CC=C(C2CC(C3=CC(C=CC=C4)=C4C=C3)=NC5=C(N2)C=CC=C5)C=C1</chem>	5.366309 772	8.064971 134
mol_051S	III 2,3- dihidrobe nzodiaze pina (S)	p- Dimetila minobenz aldehído	2- Acetilnaft aleno	<chem>CN(C)C1=CC=C(C2CC(C3=CC(C=CC=C4)=C4C=C3)=NC5=C(N2)C=CC=C5)C=C1</chem>	5.366309 772	8.064940 861

mol_052R	IV 3,4-dihidropirazol (R)	p-Dimetilaminobenzaldehído	2-Acetilnaftaleno	<chem>CN(C)C1=CC=C(C2NN=C(C2)C3=CC4=C(C=C3)C=CC=C4)C=C1</chem>	5.453709681	7.253420804
mol_052S	IV 3,4-dihidropirazol (S)	p-Dimetilaminobenzaldehído	2-Acetilnaftaleno	<chem>CN(C)C1=CC=C(C2NN=C(C2)C3=CC4=C(C=C3)C=CC=C4)C=C1</chem>	5.453709681	7.257246324
mol_053	V Pirimidina-aminobenzimidazol	p-Dimetilaminobenzaldehído	2-Acetilnaftaleno	<chem>CN(C)C1=CC=C(C2=NC(NC3=NC4=C(N3)C=C(C=C4)=NC(C5=CC6=C(C=C5)C=CC=C6)=C2)C=C1</chem>	5.607217235	7.793257906
mol_054	VI Pirimido-benzimidazol	p-Dimetilaminobenzaldehído	2-Acetilnaftaleno	<chem>CN(C)C1=CC=C(C=C1)C2=CC(C3=C(C(C=CC=C4)=C4C=C3)=NC5=NC6=C(C=CC=C6N25</chem>	5.613343765	7.407309893
mol_055	I Chalcona	p-Dimetilaminobenzaldehído	Acetophenone	<chem>CN(C)C1=CC=C(/C=C/C(C2=CC=C(C=C2)=O)C=C1</chem>	4.678262735	6.215146156
mol_056	II Aminopirimidina	p-Dimetilaminobenzaldehído	Acetophenone	<chem>CN(C)C(C=C1)=CC=C1C2=</chem>	4.796213289	7.194801653

		minobenz aldehído		<chem>NC(N)=NC(C3=CC=CC=C3)=C2</chem>		
mol_057R	III 2,3- dihidrobe nzodiaze pina (R)	p- Dimetila minobenz aldehído	Acetophe none	<chem>CN(C)C1=CC=C(C2CC(C3=CC=CC=C3)=NC4=C(N2)C=CC=C4)C=C1</chem>	5.000561 752	7.378131 583
mol_057S	III 2,3- dihidrobe nzodiaze pina (S)	p- Dimetila minobenz aldehído	Acetophe none	<chem>CN(C)C1=CC=C(C2CC(C3=CC=CC=C3)=NC4=C(N2)C=CC=C4)C=C1</chem>	5.000561 752	7.377878 304
mol_058R	IV 3,4- dihidropir azol (R)	p- Dimetila minobenz aldehído	Acetophe none	<chem>CN(C)C1=CC=C(C2NN=C(C2)C3=CC=CC=C3)C=C1</chem>	5.068240 216	6.594602 406
mol_058S	IV 3,4- dihidropir azol (S)	p- Dimetila minobenz aldehído	Acetophe none	<chem>CN(C)C1=CC=C(C2NN=C(C2)C3=CC=CC=C3)C=C1</chem>	5.068240 216	6.592626 231
mol_059	V Pirimidina - aminobe nzimidaz ol	p- Dimetila minobenz aldehído	Acetophe none	<chem>CN(C)C1=CC=C(C2=NC(NC3=NC4=C(N3)C=C(C4)=NC(C5=CC=CC=C5)=C2)C=C1</chem>	5.241506 537	7.383607 624
mol_060	VI Pirimido-	p- Dimetila	Acetophe none	<chem>CN(C)C1=CC=C(C</chem>	5.230620 016	6.738796 971

	benzimidazol	minobenzaldehído		<chem>=C1)C2=CC(C3=CC=CC=C3)=NC4=NC5=CC=CC=C5N24</chem>		
mol_061	I Chalcona	Vainillina	1-Acetilpireno	<chem>COC1=C(O)C=CC(/C=C/C(C2=C(C=C3=CC=C4)C5=C3C4=CC=C5C=C2)=O)=C1</chem>	5.398274007	7.293939243
mol_062	II Aminopirimidina	Vainillina	1-Acetilpireno	<chem>OC(C(OC)=C1)=CC=C1C2=NC(N)=NC(C3=C4C(C(C=C4)=CC=C5)=C5C=C6C=C3)=C2</chem>	5.48084487	7.185426577
mol_063R	III 2,3-dihidrobenzodiazepina (R)	Vainillina	1-Acetilpireno	<chem>OC1=CC=C(C2CC(C3=C(C=CC4=C5C(C=C6)=CC=C4)C5=C6C=C3)=NC7=C(N2)C=CC=C7)C=C1OC</chem>	5.743312317	7.397251063
mol_063S	III 2,3-dihidrobenzodiazepina (S)	Vainillina	1-Acetilpireno	<chem>OC1=CC=C(C2CC(C3=C(C=CC4=C5C(C=C6)=CC=C4)C5=C6C=C3)=NC7=C(N2)C=CC=C7)C=C1OC</chem>	5.743312317	7.394365435

				3)=NC7=C(N2)C=CC=C7)C=C1OC		
mol_064R	IV 3,4-dihidropirazol (R)	Vainillina	1-Acetilpireno	OC1=CC=C(C2NN=C(C2)C3=C4C=C5=CC=CC6=CC=C(C=C3)C4=C56)C=C1OC	5.91830844	6.862635902
mol_064S	IV 3,4-dihidropirazol (S)	Vainillina	1-Acetilpireno	OC1=CC=C(C2NN=C(C2)C3=C4C=C5=CC=CC6=CC=C(C=C3)C4=C56)C=C1OC	5.91830844	6.861551133
mol_065	V Pirimidina-aminobenzimidazol	Vainillina	1-Acetilpireno	OC1=C(OC)C=C(C2=NC(NC3=NC4=C(N3)C=C(C=C4)=NC(C5=C6C=CC7=CC=CC(C=C5)C6=C78)=C2)C=C1	5.991878648	7.400132653
mol_066	VI Pirimido-benzimidazol	Vainillina	1-Acetilpireno	OC1=C(OC)C=C(C=C1)C2=CC(C3=C4C=CC5=C6C(C=C(C(C=C3)=C46)=CC=C5)=NC	6.005044706	7.813617455

				7=NC8=C C=CC=C 8N27		
mol_067	I Chalcona	Vainillina	9- Acetilantr aceno	<chem>COC1=C(O)C=CC(=C/C=C/C(C2=C3C(C=CC=C3)=CC4=C2C=CC=C4)=O)=C1</chem>	5.149354 061	6.925079 772
mol_068	II Aminopiri midina	Vainillina	9- Acetilantr aceno	<chem>OC(C(OC)=C1)=CC=C1C2=NC(N)=NC(C3=C4C(C=CC=C4)=CC5=C3C=CC=C5)=C2</chem>	5.213888 892	6.823123 886
mol_069R	III 2,3- dihidrobe nzodiaze pina (R)	Vainillina	9- Acetilantr aceno	<chem>OC1=CC=C(C2CC(C3=C(C=CC=C4)C4=CC5=C3C=CC=C5)=NC6=C(N2)C=CC=C6)C=C1OC</chem>	5.474363 732	7.286267 134
mol_069S	III 2,3- dihidrobe nzodiaze pina (S)	Vainillina	9- Acetilantr aceno	<chem>OC1=CC=C(C2CC(C3=C(C=CC=C4)C4=CC5=C3C=CC=C5)=NC6=C(N2)C=CC=C6)C=C1OC</chem>	5.474363 732	7.286267 134
mol_070R	IV 3,4- dihidropir azol (R)	Vainillina	9- Acetilantr aceno	<chem>OC1=CC=C(C2NN=C(C2)C3=C4C(C=</chem>	5.709168 915	6.492363 844

				<chem>CC=C4)=CC5=C3C=CC=C5)C=C1O</chem>		
mol_070S	IV 3,4-dihidropirazol (S)	Vainillina	9-Acetilantraceno	<chem>OC1=CC=C(C2NN=C(C2)C3=C4C(C=CC=C4)=CC5=C3C=CC=C5)C=C1O</chem>	5.709168915	6.492371122
mol_071	V Pirimidina - aminobenzimidazol	Vainillina	9-Acetilantraceno	<chem>OC1=C(O)C=C(C2=NC(NC3=NC4=C(N3)C=C(C=C4)=NC(C5=C6C(C=CC=C6)=CC7=C5C=C(C=C7)=C2)C=C1</chem>	5.689893211	8.09181147
mol_072	VI Pirimido-benzimidazol	Vainillina	9-Acetilantraceno	<chem>OC1=CC=C(C=C1OC)C2=C(C3=C4C=CC=C4=CC5=C3C=CC=C5)=NC6=NC7=C(C=CC=C7)N26</chem>	5.749111927	7.438971213
mol_073	I Chalcona	Vainillina	1-Acetilnaftaleno	<chem>COC1=C(O)C=CC(/C=C/C(C2=CC=C(C3=C2C=CC=C3)=O)=C1</chem>	4.809928151	6.246017046

mol_074	II Aminopiri midina	Vainillina	1- Acetilnaft aleno	<chem>OC(C(OC)=C1)=CC=C1C2=NC(N)=NC(C3=C4C(C=CC=C4)=CC=C3)=C2</chem>	4.923126 226	6.451557 778
mol_075R	III 2,3- dihidrobe nzodiaze pina (R)	Vainillina	1- Acetilnaft aleno	<chem>OC1=CC=C(C2CC(C3=CC=C(C(C=CC=C4)C4=C3)=NC5=C(N2)C=CC=C5)C=C1OC</chem>	5.175727 329	6.613218 767
mol_075S	III 2,3- dihidrobe nzodiaze pina (S)	Vainillina	1- Acetilnaft aleno	<chem>OC1=CC=C(C2CC(C3=CC=C(C(C=CC=C4)C4=C3)=NC5=C(N2)C=CC=C5)C=C1OC</chem>	5.175727 329	6.616474 282
mol_076R	IV 3,4- dihidropir azol (R)	Vainillina	1- Acetilnaft aleno	<chem>OC1=CC=C(C2NN=C(C2)C3=CC=CC4=C3C=C(C=C4)C=C1OC</chem>	5.378105 769	6.911660 16
mol_076S	IV 3,4- dihidropir azol (S)	Vainillina	1- Acetilnaft aleno	<chem>OC1=CC=C(C2NN=C(C2)C3=CC=CC4=C3C=C(C=C4)C=C1OC</chem>	5.378105 769	6.908655 621
mol_077	V Pirimidina - aminobe	Vainillina	1- Acetilnaft aleno	<chem>OC1=CC=C(C2=NC(NC3=N(C4=C(N3)C</chem>	5.363373 218	7.698385 979

	nzimidazol			<chem>C=CC=C4)=NC(C5=CC=CC=C5C=C6)=C2)C=C1O</chem>		
mol_078	VI Pirimido- benzimidazol	Vainillina	1-Acetilnaftaleno	<chem>OC1=CC=C(C=C1OC)C2=CC(C3=CC=CC4=C3C=CC=C4)=NC5=NC6=CC=CC=C6N25</chem>	5.403231329	7.048305137
mol_079	I Chalcona	Vainillina	2-Acetilnaftaleno	<chem>COC1=C(O)C=CC(=C/C=C/C(C2=CC(C=CC=C3)=C3C=C2)=O)=C1</chem>	4.82188353	6.247722541
mol_080	II Aminopirimidina	Vainillina	2-Acetilnaftaleno	<chem>OC(C(OC)=C1)=CC=C1C2=NC(N)=NC(C3=CC=C(C=CC=C4)C4=C3)=C2</chem>	4.936818006	6.451424936
mol_081R	III 2,3-dihidrobenzodiazepina (R)	Vainillina	2-Acetilnaftaleno	<chem>OC1=CC=C(C2CC(C3=CC=CC4=C3C=CC=C4)=NC5=C(N2)C=CC=C5)C=C1OC</chem>	5.154426204	7.424760189
mol_081S	III 2,3-dihidrobe	Vainillina	2-Acetilnaftaleno	<chem>OC1=CC=C(C2CC(C3=CC=</chem>	5.154426204	7.424760189

	nzodiaze pina (S)			<chem>CC4=C3C=CC=C4)=NC5=C(N2)C=CC=C5)C=C1OC</chem>		
mol_082R	IV 3,4- dihidropir azol (R)	Vainillina	2- Acetilnaft aleno	<chem>OC1=C(OC)C=C(C2)NN=C(C2)C3=CC4=C(C=C3)C=CC=C4)C=C1</chem>	5.344306 761	6.651351 318
mol_082S	IV 3,4- dihidropir azol (S)	Vainillina	2- Acetilnaft aleno	<chem>OC1=C(OC)C=C(C2)NN=C(C2)C3=CC4=C(C=C3)C=CC=C4)C=C1</chem>	5.344306 761	6.651034 656
mol_083	V Pirimidina - aminobe nzimidaz ol	Vainillina	2- Acetilnaft aleno	<chem>OC1=C(OC)C=C(C2=NC(NC3=NC4=C(N3)C=C(C=C4)=NC(C5=CC6=C(C=C5)C=CC=C6)=C2)C=C1</chem>	5.372666 009	7.166842 699
mol_084	VI Pirimido- benzimid azol	Vainillina	2- Acetilnaft aleno	<chem>OC1=CC=C(C=C1OC)C2=CC(C3=CC(C=CC=C4)=C4C=C3)=NC5=NC6=CC=CC=C6N25</chem>	5.407249 331	6.782303 484
mol_085	I Chalcona	Vainillina	Acetophe none	<chem>COC1=C(O)C=CC(/C=C/C(C</chem>	4.451626 117	5.593131 477

				<chem>O=C1C=CC(=O)C=C1</chem>		
mol_086	II Aminopiri midina	Vainillina	Acetophe none	<chem>OC(C(OC)=C1)=CC=C1C2=NC(N)=NC(C3=CC=CC=C3)=C2</chem>	4.624971 549	6.643446 691
mol_087R	III 2,3- dihidrobe nzodiaze pina (R)	Vainillina	Acetophe none	<chem>OC1=CC=C(C2CC(C3=CC=CC=C3)=NC4=C(N2)C=CC=C4)C=C1OC</chem>	4.841597 208	6.768515 102
mol_087S	III 2,3- dihidrobe nzodiaze pina (S)	Vainillina	Acetophe none	<chem>OC1=CC=C(C2CC(C3=CC=CC=C3)=NC4=C(N2)C=CC=C4)C=C1OC</chem>	4.841597 208	6.768515 102
mol_088R	IV 3,4- dihidropir azol (R)	Vainillina	Acetophe none	<chem>OC1=CC=C(C2NN=C(C2)C3=CC=CC=C3)C=C1OC</chem>	5.014047 467	6.026873 853
mol_088S	IV 3,4- dihidropir azol (S)	Vainillina	Acetophe none	<chem>OC1=CC=C(C2NN=C(C2)C3=CC=CC=C3)C=C1OC</chem>	5.014047 467	6.025642 668
mol_089	V Pirimidina - aminobe nzimidaz ol	Vainillina	Acetophe none	<chem>OC1=C(OC)C=C(C2=NC(NC3=NC4=C(N3)C=C(C4)=N</chem>	5.020779 889	6.798269 679

				<chem>C(C5=CC=CC=C5)=C2)C=C1</chem>		
mol_090	VI Pirimido- benzimidazol	Vainillina	Acetophe none	<chem>OC1=C(OC)C=C(C=C1)C2=CC(C3=CC=CC=C3)=NC4=NC5=CC=CC=C5N24</chem>	5.049386138	6.150853182
mol_091	I Chalcona	Cinamald ehído	1-Acetilpire no	<chem>O=C(C=C/C=C/C1=CC=CC=C1)C2=C(C=CC3=CC=C4)C5=C3C4=CC=C5C=C2</chem>	5.651653192	8.319704876
mol_092	II Aminopiri midina	Cinamald ehído	1-Acetilpire no	<chem>NC1=NC(C2=C3C4=C5C(C=C3)=CC=CC5=CC=C4C=C2)=CC(/C=C/C6=CC=CC=C6)=N1</chem>	5.671910699	8.022646072
mol_093R	III 2,3-dihidrobe nzodiaze pina (R)	Cinamald ehído	1-Acetilpire no	<chem>C1(C2=C3C(C4=C5C=C3)=C(C=CC4=CC=C5)C=C2)=NC6=CC=CC=C6N(C1)/C=C/C7=CC=CC=C7</chem>	5.97373416	8.387901486

mol_093S	III 2,3-dihidrobenzodiazepina (S)	Cinamaldehído	1-Acetilpireno	<chem>C1(C2=C3C(C4=C5C=C3)=C(C=CC4=CC=C5)C=C2)=NC6=CC=CC=C6N(C(C1)/C=C/C7=CC=CC=C7</chem>	5.97373416	8.38746301
mol_094R	IV 3,4-dihidropirazol (R)	Cinamaldehído	1-Acetilpireno	<chem>C1(/C=C/C2NN=C(C2)C3=C4C=CC5=CC=CC6=CC=C(C=C3)C4=C56)=CC=CC=C1</chem>	6.19006201	7.790573144
mol_094S	IV 3,4-dihidropirazol (S)	Cinamaldehído	1-Acetilpireno	<chem>C1(/C=C/C2NN=C(C2)C3=C4C=CC5=CC=CC6=CC=C(C=C3)C4=C56)=CC=CC=C1</chem>	6.19006201	7.788567621
mol_095	V Pirimidina-aminobenzimidazol	Cinamaldehído	1-Acetilpireno	<chem>C1(/C=C/C2=NC(NC3=NC4=CC=C4)=NC(C5=C6C=CC7=CC=CC8=CC=C(C=C5)C6=C78)=C2)=CC=CC=C1</chem>	6.196533448	8.367055279

mol_096	VI Pirimido- benzimid azol	Cinamald ehído	1- Acetilpire no	C1(/C=C/ C2=CC(C 3=C4C=C C5=C6C(C=CC(C= C3)=C46) =CC=C5) =NC7=N C8=CC= CC=C8N 27)=CC= CC=C1	6.176393 972	8.769588 899
mol_097	I Chalcona	Cinamald ehído	9- Acetilantr aceno	O=C(C=C /C=C/C1= CC=CC= C1)C2=C 3C(C=CC =C3)=CC 4=C2C=C C=C4	5.453994 718	7.933514 219
mol_098	II Aminopiri midina	Cinamald ehído	9- Acetilantr aceno	NC1=NC(C2=C(C= CC=C3)C 3=CC4=C 2C=CC= C4)=CC(/ C=C/C5= CC=CC= C5)=N1	5.437293 752	7.648527 347
mol_099R	III 2,3- dihidrobe nzodiaze pina (R)	Cinamald ehído	9- Acetilantr aceno	C1(C2=C 3C(C=CC =C3)=CC 4=C2C=C C=C4)=N C5=CC= CC=C5N C(C1)/C= C/C6=CC =CC=C6	5.738745 357	8.275030 47
mol_099S	III 2,3- dihidrobe nzodiaze pina (S)	Cinamald ehído	9- Acetilantr aceno	C1(C2=C 3C(C=CC =C3)=CC 4=C2C=C C=C4)=N	5.738745 357	8.275016 58

				<chem>C5=CC=CC=C5N</chem> <chem>C(C1)/C=C/C6=CC=CC=C6</chem>		
mol_100R	IV 3,4-dihidropirazol (R)	Cinamaldéhid	9-Acetilantraceno	<chem>C1/C=C/C2=NN=C(C2)C3=C4C(C=CC=C4)=CC5=C3C=C(C=C5)=CC=C1</chem>	6.051621074	7.411268581
mol_100S	IV 3,4-dihidropirazol (S)	Cinamaldéhid	9-Acetilantraceno	<chem>C1/C=C/C2=NN=C(C2)C3=C4C(C=CC=C4)=CC5=C3C=C(C=C5)=CC=C1</chem>	6.051621074	7.407234965
mol_101	V Pirimidina-aminobenzimidazol	Cinamaldéhid	9-Acetilantraceno	<chem>C1/C=C/C2=NC(NC3=NC4=CC=C4)C=C(NC(C5=C6C(C=CC=C6)=CC7=C5C=C(C=C7)=C2)=CC=C1</chem>	5.941962963	9.049639575
mol_102	VI Pirimido-benzimidazol	Cinamaldéhid	9-Acetilantraceno	<chem>C1/C=C/C2=CC(C3=C4C=C(C=CC4=CC5=C3C=CC=C5)=NC6=NC7=CC=CC=C7</chem>	5.947012652	8.386695818

				N26)=CC =CC=C1		
mol_103	I Chalcona	Cinamald ehído	1- Acetilnaft aleno	O=C(C=C /C=C/C1= CC=CC= C1)C2=C 3C(C=CC =C3)=CC =C2	5.068685 286	7.182551 211
mol_104	II Aminopiri midina	Cinamald ehído	1- Acetilnaft aleno	NC1=NC(C2=C(C= CC=C3)C 3=CC=C2)=CC(/C= C/C4=CC =CC=C4) =N1	5.109880 961	7.208218 069
mol_105R	III 2,3- dihidrobe nzodiaze pina (R)	Cinamald ehído	1- Acetilnaft aleno	C1(C2=C 3C(C=CC =C3)=CC =C2)=NC 4=CC=C C=C4NC(C1)/C=C/ C5=CC= CC=C5	5.392463 985	7.523945 5
mol_105S	III 2,3- dihidrobe nzodiaze pina (S)	Cinamald ehído	1- Acetilnaft aleno	C1(C2=C 3C(C=CC =C3)=CC =C2)=NC 4=CC=C C=C4NC(C1)/C=C/ C5=CC= CC=C5	5.392463 985	7.523284 979
mol_106R	IV 3,4- dihidropir azol (R)	Cinamald ehído	1- Acetilnaft aleno	C1(/C=C/ C2NN=C(C3=CC= CC4=C3 C=CC=C 4)C2)=CC =CC=C1	5.668379 973	7.751171 546

mol_106S	IV 3,4-dihidropirazol (S)	Cinamaldéhidó	1-Acetilnaftaleno	<chem>C1(/C=C/C2NN=C(C3=CC=CC4=C3C=CC=C4)C2)=CC=CC=C1</chem>	5.668379973	7.7490449
mol_107	V Pirimidina-aminobenzimidazol	Cinamaldéhidó	1-Acetilnaftaleno	<chem>C1(/C=C/C2=NC(NC3=NC4=C(N3)C=CC=C4)=NC(C5=C(C=CC6=C5C=CC=C6)=C2)=CC=CC=C1</chem>	5.597378206	8.309774736
mol_108	VI Pirimido-benzimidazol	Cinamaldéhidó	1-Acetilnaftaleno	<chem>C1(/C=C/C2=CC(C3=CC=C4=C3C=CC=C4)=NC5=NC6=CC=C(C=C6N25)=CC=CC=C1</chem>	5.567352447	7.653585713
mol_109	I Chalcona	Cinamaldéhidó	2-Acetilnaftaleno	<chem>O=C(C1=CC=C(C=CC=C2)C2=C1)C=C/C=C/C3=CC=C(C=C3</chem>	5.105358444	7.190752876
mol_110	II Aminopirimidina	Cinamaldéhidó	2-Acetilnaftaleno	<chem>NC1=NC(C2=CC=C(C(C=CC=C3)C3=C2)=CC(/C=C/C4=C(C=CC=C4)=N1</chem>	5.15763984	6.796936172

mol_111R	III 2,3-dihidrobenzodiazepina (R)	Cinamaldehído	2-Acetilnaftaleno	<chem>C1(C2=CC=CC3C(C=CC=C3)=C2)=NC4=CC=C(C=C4NC(C1)/C=C/C5=CC=CC=C5)</chem>	5.440185767	8.329066282
mol_111S	III 2,3-dihidrobenzodiazepina (S)	Cinamaldehído	2-Acetilnaftaleno	<chem>C1(C2=CC=CC3C(C=CC=C3)=C2)=NC4=CC=C(C=C4NC(C1)/C=C/C5=CC=CC=C5)</chem>	5.440185767	8.328835274
mol_112R	IV 3,4-dihidropirazol (R)	Cinamaldehído	2-Acetilnaftaleno	<chem>C1(/C=C/C2NN=C(C2)C3=CC4=C(C=C3)C=CC=C4)=CC=CC=C1</chem>	5.639540649	7.481141193
mol_112S	IV 3,4-dihidropirazol (S)	Cinamaldehído	2-Acetilnaftaleno	<chem>C1(/C=C/C2NN=C(C2)C3=CC4=C(C=C3)C=CC=C4)=CC=CC=C1</chem>	5.639540649	7.483826943
mol_113	V Pirimidina-aminobenzimidazol	Cinamaldehído	2-Acetilnaftaleno	<chem>C1(/C=C/C2=NC(NC3=NC4=CC(N3)C=CC=C4)=NC(C5=CC6=C(C=C5)C=CC=C6)=C2)=CC=CC=C1</chem>	5.622716601	8.043592779

mol_114	VI Pirimido- benzimid azol	Cinamald ehído	2- Acetilnaft aleno	C1(/C=C/ C2=CC(C 3=CC(C= CC=C4)= C4C=C3) =NC5=N C6=CC= CC=C6N 25)=CC= CC=C1	5.591666 968	7.652480 985
mol_115	I Chalcona	Cinamald ehído	Acetophe none	O=C(C1= CC=CC= C1)C=C/ C=C/C2= CC=CC= C2	4.665125 886	6.461608 422
mol_116	II Aminopiri midina	Cinamald ehído	Acetophe none	NC1=NC(C2=CC= CC=C2)= CC(/C=C/ C3=CC= CC=C3)= N1	4.789589 403	7.203345 563
mol_117R	III 2,3- dihidrobe nzodiaze pina (R)	Cinamald ehído	Acetophe none	C1(C2=C C=CC=C 2)=NC3= CC=CC= C3NC(C1)C=C/C4 =CC=CC =C4	5.061603 102	7.873221 977
mol_117S	III 2,3- dihidrobe nzodiaze pina (S)	Cinamald ehído	Acetophe none	C1(C2=C C=CC=C 2)=NC3= CC=CC= C3NC(C1)C=C/C4 =CC=CC =C4	5.061603 102	7.866479 22
mol_118R	IV 3,4- dihidropir azol (R)	Cinamald ehído	Acetophe none	C1(/C=C/ C2NN=C(C3=CC= CC=C3)C	5.250928 333	7.053970 178

				2)=CC=C C=C1		
mol_118S	IV 3,4-dihidropirazol (S)	Cinamaldéhid	Acetophenone	C1(/C=C/C2NN=C(C3=CC=CC=C3)C2)=CC=C C=C1	5.250928333	7.053993857
mol_119	V Pirimidina-aminobenzimidazol	Cinamaldéhid	Acetophenone	C1(C2=NC(NC3=NC4=C(N3)C=CC=C4)=NC(/C=C/C5=CC=CC=C5)=C2)=CC=C1	5.204010974	7.407396445
mol_120	VI Pirimido-benzimidazol	Cinamaldéhid	Acetophenone	C1(/C=C/C2=CC(C3=CC=C(C=C3)=NC4=NC5=CC=CC=C5N24)=CC=CC=C1	5.179210391	7.211327663
mol_121	I Chalcona	Salicilaldéhid	1-Acetilpireno	OC1=CC=CC=C1C=CC(C2=C(C=CC3=CC=C4)C5C3C4=CC=C5C=C2)=O	5.427152279	7.346772372
mol_122	II Aminopirimidina	Salicilaldéhid	1-Acetilpireno	NC1=NC(C2=C3C4=C5C(C=C3)=CC=CC5=CC=C4C=C2)=CC(C6=C(O)C=C	5.418870136	7.186759221

				C=C6)=N 1		
mol_123R	III 2,3-dihidrobe nzodiaze pina (R)	Salicilald ehído	1-Acetilpire no	OC1=CC =CC=C1 C2CC(C3 =C(C=CC 4=C5C(C =C6)=CC =C4)C5= C6C=C3) =NC7=C(N2)C=CC =C7	5.635989 118	7.428448 68
mol_123S	III 2,3-dihidrobe nzodiaze pina (S)	Salicilald ehído	1-Acetilpire no	OC1=CC =CC=C1 C2CC(C3 =C(C=CC 4=C5C(C =C6)=CC =C4)C5= C6C=C3) =NC7=C(N2)C=CC =C7	5.635989 118	7.428448 68
mol_124R	IV 3,4-dihidropir azol (R)	Salicilald ehído	1-Acetilpire no	OC1=CC =CC=C1 C2NN=C(C2)C3=C 4C=CC5= CC=CC6 =CC=C(C =C3)C4= C56	5.868247 015	6.883757 284
mol_124S	IV 3,4-dihidropir azol (S)	Salicilald ehído	1-Acetilpire no	OC1=CC =CC=C1 C2NN=C(C2)C3=C 4C=CC5= CC=CC6 =CC=C(C =C3)C4= C56	5.868247 015	6.880806 623

mol_125	V Pirimidina - aminobe nzimidaz ol	Salicilald ehído	1- Acetilpire no	<chem>OC1=CC=CC=C1C2=NC(NC3=NC4=C(N3)C=CC=C4)NC(C5=C6C=CC(=CC=C(C=C5)C6)=C78)=C2</chem>	5.907175 485	7.431916 004
mol_126	VI Pirimido- benzimid azol	Salicilald ehído	1- Acetilpire no	<chem>OC(C=CC=C1)=C1C2=CC(C3=C4C=C5=C6C(C=CC(C=C3)=C46)=CC=C5)=NC7=NC8=CC=C8N27</chem>	5.934399 55	7.840174 09
mol_127	I Chalcona	Salicilald ehído	9- Acetilantr aceno	<chem>OC1=CC=CC=C1C=CC(C2=C3C(C=CC=C3)=CC4=C2C=CC=C4)=O</chem>	5.200048 065	6.978959 07
mol_128	II Aminopiri midina	Salicilald ehído	9- Acetilantr aceno	<chem>NC1=NC(C2=C3C(C=CC=C3)=CC4=C2C=CC=C4)=CC(C5=C(O)C=CC=C5)=N1</chem>	5.156938 305	6.818221 308
mol_129R	III 2,3- dihidrobe	Salicilald ehído	9- Acetilantr aceno	<chem>OC1=CC=CC=C1C2CC(C3</chem>	5.376918 894	7.322418 747

	nzodiaze pina (R)			<chem>=C(C=CC=C4)C4=CC5=C3C=CC=C5)=NC6=C(N2)C=CC=C6</chem>		
mol_129S	III 2,3- dihidrobe nzodiaze pina (S)	Salicilald ehído	9- Acetilantr aceno	<chem>OC1=CC=CC=C1C2CC(C3=C(C=CC=C4)C4=CC5=C3C=CC=C5)=NC6=C(N2)C=CC=C6</chem>	5.376918 894	7.322455 508
mol_130R	IV 3,4- dihidropir azol (R)	Salicilald ehído	9- Acetilantr aceno	<chem>OC1=C(C2NN=C(C2)C3=C4C(C=CC=C4)=CC5=C3C=C(C=C5)C=CC=C1</chem>	5.696402 2	6.508297 481
mol_130S	IV 3,4- dihidropir azol (S)	Salicilald ehído	9- Acetilantr aceno	<chem>OC1=C(C2NN=C(C2)C3=C4C(C=CC=C4)=CC5=C3C=C(C=C5)C=CC=C1</chem>	5.696402 2	6.508987 625
mol_131	V Pirimidina - aminobe nzimidaz ol	Salicilald ehído	9- Acetilantr aceno	<chem>OC1=C(C2=NC(NC3=NC4=C(N3)C=C(C4)=NC(C5=C6C(C=CC=C6)=CC7=C5C=C(C=C7)=C</chem>	5.639666 863	8.120920 185

				2)C=CC=C1		
mol_132	VI Pirimido- benzimidazol	Salicilald ehído	9- Acetilantraceno	OC1=CC=CC=C1C2=CC(C3=C4C=CC=CC4=CC5=C3C=CC=C5)=NC6=NC7=CC=CC=C7N26	5.684089805	7.458088973
mol_133	I Chalcona	Salicilald ehído	1- Acetilnaftaleno	OC1=CC=CC=C1C=CC(C2=CC=CC3=C2C=CC=C3)=O	4.837161089	6.267254047
mol_134	II Aminopirimidina	Salicilald ehído	1- Acetilnaftaleno	NC1=NC(C2=C3C(C=CC=C3)=CC=C2)=CC(C4=C(O)C=CC=C4)=N1	4.851669211	6.424205496
mol_135R	III 2,3-dihidrobenzodiazepina (R)	Salicilald ehído	1- Acetilnaftaleno	OC1=CC=CC=C1C2CC(C3=C(C=CC=C4)C4=CC=C3)=NC5=C(N2)C=CC=C5	5.037228301	6.623750881
mol_135S	III 2,3-dihidrobenzodiazepina (S)	Salicilald ehído	1- Acetilnaftaleno	OC1=CC=CC=C1C2CC(C3=C(C=CC=C4)C4=CC=C3)=NC5=C(N	5.037228301	6.623750881

				2)C=CC=C5		
mol_136R	IV 3,4-dihidropirazol (R)	Salicilaldehído	1-Acetilnftaleno	OC1=CC=CC=C1C2NN=C(C2)C3=CC=CC4=C3C=CC=C4	5.326432034	6.901831714
mol_136S	IV 3,4-dihidropirazol (S)	Salicilaldehído	1-Acetilnftaleno	OC1=CC=CC=C1C2NN=C(C2)C3=CC=CC4=C3C=CC=C4	5.326432034	6.898695639
mol_137	V Pirimidina-aminobenzimidazol	Salicilaldehído	1-Acetilnftaleno	OC1=CC=CC=C1C2=NC(NC3=NC4=CC(N3)=CC=C4)=NC(C5=CC=CC6=C5C=CC=C6)=C2	5.298269812	7.432148336
mol_138	VI Pirimido-benzimidazol	Salicilaldehído	1-Acetilnftaleno	OC1=CC=CC=C1C2=CC(C3=CC=CC4=C3C=CC=C4)=NC5=NC6=CC=C6C=C5N2	5.318552084	6.780685416
mol_139	I Chalcona	Salicilaldehído	2-Acetilnftaleno	OC1=CC=CC=C1C=CC(C2=CC(C=C3C=C3)=CC=C2)=O	4.856889854	6.269814259

mol_140	II Aminopiri midina	Salicilald ehído	2- Acetilnaft aleno	<chem>NC1=NC(C2=CC=C(C(C=CC=C3)C3=C2)=CC(C4=C(O)C=CC=C4)=N1</chem>	4.868818 228	6.157442 983
mol_141R	III 2,3- dihidrobe nzodiaze pina (R)	Salicilald ehído	2- Acetilnaft aleno	<chem>OC1=CC=CC=C1C2CC(C3=CC(C=C4C=C3)=NC5=C(N2)C=CC=C5</chem>	5.058775 53	7.423152 704
mol_141S	III 2,3- dihidrobe nzodiaze pina (S)	Salicilald ehído	2- Acetilnaft aleno	<chem>OC1=CC=CC=C1C2CC(C3=CC(C=C4C=C3)=NC5=C(N2)C=CC=C5</chem>	5.058775 53	7.415927 276
mol_142R	IV 3,4- dihidropir azol (R)	Salicilald ehído	2- Acetilnaft aleno	<chem>OC1=C(C2NN=C(C2)C3=CC4=C(C=C3)C=CC=C4)C=CC=C1</chem>	5.269710 008	6.630362 623
mol_142S	IV 3,4- dihidropir azol (S)	Salicilald ehído	2- Acetilnaft aleno	<chem>OC1=C(C2NN=C(C2)C3=CC4=C(C=C3)C=CC=C4)C=CC=C1</chem>	5.269710 008	6.632570 421
mol_143	V Pirimidina - aminobe	Salicilald ehído	2- Acetilnaft aleno	<chem>OC1=C(C2=NC(NC3=NC4=C(N3)C=C</chem>	5.309923 263	7.169303 776

	nzimidazol			<chem>C=C4)=N C(C5=CC 6=C(C=C 5)C=CC= C6)=C2)C =CC=C1</chem>		
mol_144	VI Pirimido- benzimidazol	Salicilaldehído	2-Acetilnaftaleno	<chem>OC(C=CC =C1)=C1 C2=CC(C 3=CC(C= CC=C4)= C4C=C3) =NC5=N C6=CC= CC=C6N 25</chem>	5.32508652	6.783401218
mol_145	I Chalcona	Salicilaldehído	Acetophenone	<chem>OC1=CC =CC=C1 C=CC(C2 =CC=CC =C2)=O</chem>	4.453996118	5.592494445
mol_146	II Aminopirimidina	Salicilaldehído	Acetophenone	<chem>NC1=NC(C2=CC= CC=C2)= CC(C3=C (O)C=CC =C3)=N1</chem>	4.532149468	6.590130858
mol_147R	III 2,3-dihidrobenzodiazepina (R)	Salicilaldehído	Acetophenone	<chem>OC1=CC =CC=C1 C2CC(C3 =CC=CC =C3)=NC 4=C(N2)C =CC=C4</chem>	4.705614133	6.75392292
mol_147S	III 2,3-dihidrobenzodiazepina (S)	Salicilaldehído	Acetophenone	<chem>OC1=CC =CC=C1 C2CC(C3 =CC=CC =C3)=NC 4=C(N2)C =CC=C4</chem>	4.705614133	6.753912702
mol_148R	IV 3,4-dihidropirazol (R)	Salicilaldehído	Acetophenone	<chem>OC1=C(C 2NN=C(C 2)C3=CC</chem>	4.89877441	5.994029004

				<chem>=CC=C3)C=CC=C1</chem>		
mol_148S	IV 3,4-dihidropirazol (S)	Salicilaldehído	Acetophenone	<chem>OC1=C(C2NN=C(C2)C3=CC=CC=C3)C=CC=C1</chem>	4.89877441	5.986883532
mol_149	V Pirimidina - aminobenzimidazol	Salicilaldehído	Acetophenone	<chem>OC1=C(C2=NC(NC3=NC4=C(N3)C=C(C=C4)=NC(C5=CC=CC=C5)=C2)C=C(C=C1</chem>	4.933111577	6.775280937
mol_150	VI Pirimido-benzimidazol	Salicilaldehído	Acetophenone	<chem>OC(C=CC=C1)=C1C2=CC(C3=CC=C(C=C3)=NC4=NC5=CC=CC=C5N24</chem>	4.938270211	6.123924439
mol_151	I Chalcona	4-Isopropilbenzaldehído	1-Acetilpireno	<chem>CC(C)C(C=C1)=CC=C1/C=C/C(C2=C(C=CC3=CC=C4)C5=C3C4=CC=C5C=C2)=O</chem>	5.970567937	8.044844667
mol_152	II Aminopirimidina	4-Isopropilbenzaldehído	1-Acetilpireno	<chem>NC1=NC(C2=CC=C(C=C3)C4=C2C=CC5=CC=CC3=C45)=CC(C6=CC=C(C</chem>	5.988006113	7.810692146

				(C)C)C=C 6)=N1		
mol_153R	III 2,3- dihidrobe nzodiaze pina (R)	4- Isopropilb enzaldehí do	1- Acetilpire no	CC(C)C1 =CC=C(C 2CC(C3= CC=C4C 5=C3C=C C6=CC= CC(C=C4)=C65)=N C7=C(N2) C=CC=C 7)C=C1	6.236498 442	8.124258 733
mol_153S	III 2,3- dihidrobe nzodiaze pina (S)	4- Isopropilb enzaldehí do	1- Acetilpire no	CC(C)C1 =CC=C(C 2CC(C3= CC=C4C 5=C3C=C C6=CC= CC(C=C4)=C65)=N C7=C(N2) C=CC=C 7)C=C1	6.236498 442	8.124575 773
mol_154R	IV 3,4- dihidropir azol (R)	4- Isopropilb enzaldehí do	1- Acetilpire no	CC(C)C1 =CC=C(C 2NN=C(C 2)C3=C4 C=CC5= CC=CC6 =CC=C(C =C3)C4= C56)C=C 1	6.338618 052	7.524669 545
mol_154S	IV 3,4- dihidropir azol (S)	4- Isopropilb enzaldehí do	1- Acetilpire no	CC(C)C1 =CC=C(C 2NN=C(C 2)C3=C4 C=CC5= CC=CC6 =CC=C(C =C3)C4= C56)C=C 1	6.338618 052	7.521959 369

mol_155	V Pirimidina - aminobe nzimidaz ol	4- Isopropilb enzaldehí do	1- Acetilpire no	<chem>CC(C)C1=CC=C(C2=NC(NC3=NC4=C(N3)C=C(C=C4)=NC(C5=C6C=CC7=CC=CC8=CC=C(C=C5)C6=C78)=C2)C=C1</chem>	6.509195 865	8.097133 539
mol_156	VI Pirimido- benzimid azol	4- Isopropilb enzaldehí do	1- Acetilpire no	<chem>CC(C)C1=CC=C(C=C1)C2=CC(C3=C4C=CC5=C6C(C=C(C(C=3)=C46)=CC=C5)=NC7=NC8=C(C=CC=C8N27</chem>	6.516997 839	8.503549 099
mol_157	I Chalcona	4- Isopropilb enzaldehí do	9- Acetilantr aceno	<chem>CC(C)C(C=C1)=CC=C1/C=C/C(C2=C3C(C=CC=C3)=CC4=C2C=C(C=C4)=O</chem>	5.777794 878	7.926799 288
mol_158	II Aminopiri midina	4- Isopropilb enzaldehí do	9- Acetilantr aceno	<chem>NC1=NC(C2=C(C=CC=C3)C3=CC4=C2C=CC=C4)=CC(C5=CC=C(C(C(C)C)C=C5)=N1</chem>	5.754846 598	7.699783 025

mol_159R	III 2,3-dihidrobe nzodiaze pina (R)	4- Isopropilb enzaldehí do	9- Acetilantr aceno	<chem>CC(C)C1=CC=C(C2CC(C3=C(C=CC=C4)C4=C(C5=C3C=CC=C5)=NC6=C(N2)C=CC=C6)C=C1</chem>	6.003752 941	8.002152 086
mol_159S	III 2,3-dihidrobe nzodiaze pina (S)	4- Isopropilb enzaldehí do	9- Acetilantr aceno	<chem>CC(C)C1=CC=C(C2CC(C3=C(C=CC=C4)C4=C(C5=C3C=CC=C5)=NC6=C(N2)C=CC=C6)C=C1</chem>	6.003752 941	8.001802 726
mol_160R	IV 3,4-dihidropir azol (R)	4- Isopropilb enzaldehí do	9- Acetilantr aceno	<chem>CC(C)C1=CC=C(C2NN=C(C2)C3=C4C(C=CC=C4)=CC5=C3C=CC=C5)C=C1</chem>	6.204759 567	7.412835 85
mol_160S	IV 3,4-dihidropir azol (S)	4- Isopropilb enzaldehí do	9- Acetilantr aceno	<chem>CC(C)C1=CC=C(C2NN=C(C2)C3=C4C(C=CC=C4)=CC5=C3C=CC=C5)C=C1</chem>	6.204759 567	7.403490 877
mol_161	V Pirimidina - aminobe nzimidaz ol	4- Isopropilb enzaldehí do	9- Acetilantr aceno	<chem>CC(C)C1=CC=C(C2=NC(NC3=NC4=C(N3)C=C(C=C4)=N</chem>	6.259505 072	9.754464 053

				<chem>C(C5=C6C(C=CC=C6)=CC7=C5C=C(C=C7)=C2)C=C1</chem>		
mol_162	VI Pirimido- benzimidazol	4- Isopropilb enzaldehí do	9- Acetilnitr aceno	<chem>CC(C)C1=CC=C(C=C1)C2=CC(C3=C4C=CC=CC4=CC5=C3C=C(C=C5)=NC6=NC7=CC=CC=C7N26</chem>	6.290165 686	9.093006 738
mol_163	I Chalcona	4- Isopropilb enzaldehí do	1- Acetilnaft aleno	<chem>CC(C)C(C=C1)=CC=C1/C=C/C(C2=C3C(C=CC=C3)=CC=C2)=O</chem>	5.370507 944	7.886584 126
mol_164	II Aminopiri midina	4- Isopropilb enzaldehí do	1- Acetilnaft aleno	<chem>NC1=NC(C2=C(C=CC=C3)C3=CC=C2)=CC(C4=CC=C(C(C)C)C=C4)=N1</chem>	5.418821 692	7.966176 699
mol_165R	III 2,3- dihidrobe nzodiaze pina (R)	4- Isopropilb enzaldehí do	1- Acetilnaft aleno	<chem>CC(C)C1=CC=C(C2CC(C3=C(C=CC=C4)C4=CC=C3)=NC5=C(N2)C=CC=C5)C=C1</chem>	5.637309 021	8.235140 486
mol_165S	III 2,3- dihidrobe	4- Isopropilb	1- Acetilnaft aleno	<chem>CC(C)C1=CC=C(C2CC(C3=</chem>	5.637309 021	8.234693 555

	nzodiaze pina (S)	enzaldehí do		<chem>C(C=CC=C4)C4=C(C=C3)=NC5=C(N2)C=CC=C5)C=C1</chem>		
mol_166R	IV 3,4- dihidropir azol (R)	4- Isopropilb enzaldehí do	1- Acetilnaft aleno	<chem>CC(C)C1=CC=C(C2NN=C(C2)C3=CC=CC4=C3C=CC=C4)C=C1</chem>	5.770068 432	8.464051 344
mol_166S	IV 3,4- dihidropir azol (S)	4- Isopropilb enzaldehí do	1- Acetilnaft aleno	<chem>CC(C)C1=CC=C(C2NN=C(C2)C3=CC=CC4=C3C=CC=C4)C=C1</chem>	5.770068 432	8.459841 482
mol_167	V Pirimidina - aminobe nzimidaz ol	4- Isopropilb enzaldehí do	1- Acetilnaft aleno	<chem>CC(C)C1=CC=C(C2=NC(NC3=NC4=C(N3)C=C(C4)=NC(C5=CC=CC6=C5C=CC=C6)=C2)C=C1</chem>	5.911756 599	9.018771 737
mol_168	VI Pirimido- benzimid azol	4- Isopropilb enzaldehí do	1- Acetilnaft aleno	<chem>CC(C)C1=CC=C(C=C1)C2=CC(C3=C(C=CC4=C3C=CC=C4)=NC5=NC6=C(C=CC=C6N25</chem>	5.910209 763	8.365467 148
mol_169	I Chalcona	4- Isopropilb	2- Acetilnaft aleno	<chem>CC(C)C(C=C1)=CC=C1/C=C/</chem>	5.387333 735	7.884149 62

		enzaldehído		<chem>C(C2=CC=C(C=CC=C3)C3=C2)=O</chem>		
mol_170	II Aminopirimidina	4-Isopropilbenzaldehído	2-Acetilnaftaleno	<chem>NC1=NC(C2=CC=C3C(C=C(C=C3)=C2)=CC(C4=CC=C(C(C)C)C=C4)=N1</chem>	5.440296534	7.700856922
mol_171R	III 2,3-dihidrobenzodiazepina (R)	4-Isopropilbenzaldehído	2-Acetilnaftaleno	<chem>CC(C)C1=CC=C(C2CC(C3=CC=C4C(C=CC=C4)=C3)=NC5=C(N2)C=CC=C5)C=C1</chem>	5.666778716	9.030471993
mol_171S	III 2,3-dihidrobenzodiazepina (S)	4-Isopropilbenzaldehído	2-Acetilnaftaleno	<chem>CC(C)C1=CC=C(C2CC(C3=CC=C4C(C=CC=C4)=C3)=NC5=C(N2)C=CC=C5)C=C1</chem>	5.666778716	9.030145959
mol_172R	IV 3,4-dihidropirazol (R)	4-Isopropilbenzaldehído	2-Acetilnaftaleno	<chem>CC(C)C1=CC=C(C2NN=C(C2)C3=CC4=C(C=C3)C=CC=C4)C=C1</chem>	5.710558418	8.199430762
mol_172S	IV 3,4-dihidropirazol (S)	4-Isopropilbenzaldehído	2-Acetilnaftaleno	<chem>CC(C)C1=CC=C(C2NN=C(C2)C3=CC4=C(C=C3)C=CC=C4)C=C1</chem>	5.710558418	8.193199774

mol_173	V Pirimidina - aminobe nzimidaz ol	4- Isopropilb enzaldehí do	2- Acetilnaft aleno	<chem>CC(C)C1=CC=C(C2=NC(NC3=NC4=C(N3)C=C(C=C4)=NC(C5=CC6=C(C=C5)C=CC=C6)=C2)C=C1</chem>	5.927910 3	8.753209 513
mol_174	VI Pirimido- benzimid azol	4- Isopropilb enzaldehí do	2- Acetilnaft aleno	<chem>CC(C)C1=CC=C(C=C1)C2=CC(C3=C(C(C=CC=C4)=C4C=C3)=NC5=NC6=C(C=CC=C6N25</chem>	5.923455 292	8.364683 565
mol_175	I Chalcona	4- Isopropilb enzaldehí do	Acetophe none	<chem>CC(C)C(C=C1)=CC=C1/C=C/C(C2=CC=CC=C2)=O</chem>	4.960931 56	7.161003 447
mol_176	II Aminopiri midina	4- Isopropilb enzaldehí do	Acetophe none	<chem>NC1=NC(C2=CC=CC=C2)=CC(C3=C(C=C(C(C)C)C=C3)=N1</chem>	5.086491 537	8.086148 792
mol_177R	III 2,3- dihidrobe nzodiaze pina (R)	4- Isopropilb enzaldehí do	Acetophe none	<chem>CC(C)C1=CC=C(C2CC(C3=CC=CC=C3)=NC4=C(N2)C=CC=C4)C=C1</chem>	5.286970 367	8.317963 986
mol_177S	III 2,3- dihidrobe	4- Isopropilb	Acetophe none	<chem>CC(C)C1=CC=C(C</chem>	5.286970 367	8.315714 527

	nzodiaze pina (S)	enzaldehí do		<chem>2CC(C3=CC=CC=C3)=NC4=C(N2)C=CC=C4)C=C1</chem>		
mol_178R	IV 3,4- dihidropir azol (R)	4- Isopropilb enzaldehí do	Acetophe none	<chem>CC(C)C1=CC=C(C2NN=C(C2)C3=CC=CC=C3)C=C1</chem>	5.295315 65	7.508853 835
mol_178S	IV 3,4- dihidropir azol (S)	4- Isopropilb enzaldehí do	Acetophe none	<chem>CC(C)C1=CC=C(C2NN=C(C2)C3=CC=CC=C3)C=C1</chem>	5.295315 65	7.501057 927
mol_179	V Pirimidina - aminobe nzimidaz ol	4- Isopropilb enzaldehí do	Acetophe none	<chem>CC(C)C1=CC=C(C2=NC(NC3=NC4=C(N3)C=C(C4)=NC(C5=CC=CC=C5)=C2)C=C1</chem>	5.557738 873	8.313699 265
mol_180	VI Pirimido- benzimid azol	4- Isopropilb enzaldehí do	Acetophe none	<chem>CC(C)C1=CC=C(C=C1)C2=CC(C3=CC=CC=C3)=NC4=NC5=CC=CC=C5N24</chem>	5.531861 088	7.662770 986
mol_181	I Chalcona	p- aminaben zaldehído	1- Acetilpire no	<chem>NC1=CC=C(/C=C/C(=O)C2=CC=C3C=CC4=C(C=CC5=CC=C2C</chem>	5.524093 566	8.293072 47

				3=C45)C=C1		
mol_182	II Aminopirimidina	p-aminabenzaldehído	1-Acetilpireno	<chem>NC1=CC=C(C2=CC(C3=CC=C4C=C5=CC=CC6=CC=C3C4=C56)=NC(N)=N2)C=C1</chem>	5.53065466	8.191489683
mol_183R	III 2,3-dihidrobenzodiazepina (R)	p-aminabenzaldehído	1-Acetilpireno	<chem>NC1=CC=C(C2CC(C3=CC=C4C=CC5=CC=CC6=CC=C3C4=C56)=NC3=CC=CC=C3N2)C=C1</chem>	5.787734986	8.386036453
mol_183S	III 2,3-dihidrobenzodiazepina (S)	p-aminabenzaldehído	1-Acetilpireno	<chem>NC1=CC=C(C2CC(C3=CC=C4C=CC5=CC=CC6=CC=C3C4=C56)=NC3=CC=CC=C3N2)C=C1</chem>	5.787734986	8.386036453
mol_184R	IV 3,4-dihidropirazol (R)	p-aminabenzaldehído	1-Acetilpireno	<chem>NC1=CC=C(C2CC(C3=CC=C4C=CC5=CC=CC6=CC=C3C4=C56)=NN2)C=C1</chem>	5.935916126	7.861439421

mol_184S	IV 3,4-dihidropirazol (S)	p-aminabenzaldehído	1-Acetilpireno	<chem>NC1=CC=C(C2CC(C3=CC=C4C=CC5=CC=C6=CC=C3C4=C56)=NN2)C=C1</chem>	5.935916126	7.861907944
mol_185	V Pirimidina-aminobenzimidazol	p-aminabenzaldehído	1-Acetilpireno	<chem>NC1=CC=C(C2=C(C3=CC=C4C=C(C5=CC=CC6=CC=C3C4=C56)=NC(NC3=NC4=C(N3)C=CC=C4)=N2)C=C1</chem>	6.007520825	8.398261837
mol_186	VI Pirimido-benzimidazol	p-aminabenzaldehído	1-Acetilpireno	<chem>NC1=CC=C(C=C1)C2=CC(C3=C4C=C(C5=C6C(C=CC(C=C3)=C46)=CC=C5)=NC7=NC8=CC=C8N27</chem>	6.021764452	8.81633265
mol_187	I Chalcona	p-aminabenzaldehído	9-Acetilantraceno	<chem>NC1=CC=C(C(=C/C(=O)C2=C3C=CC=CC3=C(C3=CC=CC=C32)C=C1</chem>	5.314613805	7.926495062
mol_188	II Aminopirimidina	p-aminabenzaldehído	9-Acetilantraceno	<chem>NC1=CC=C(C2=C(C3=C4</chem>	5.280036604	7.822999565

				<chem>C=CC=C C4=CC4= CC=CC= C43)=NC(N)=N2)C= C1</chem>		
mol_189R	III 2,3- dihidrobe nzodiaze pina (R)	p- aminaben zaldehído	9- Acetilantr aceno	<chem>NC1=CC =C(C2CC (C3=C4C =CC=CC 4=CC4=C C=CC=C 43)=NC3 =CC=CC =C3N2)C =C1</chem>	5.538482 861	8.281873 698
mol_189S	III 2,3- dihidrobe nzodiaze pina (S)	p- aminaben zaldehído	9- Acetilantr aceno	<chem>NC1=CC =C(C2CC (C3=C4C =CC=CC 4=CC4=C C=CC=C 43)=NC3 =CC=CC =C3N2)C =C1</chem>	5.538482 861	8.281878 814
mol_190R	IV 3,4- dihidropir azol (R)	p- aminaben zaldehído	9- Acetilantr aceno	<chem>NC1=CC =C(C2CC (C3=C4C =CC=CC 4=CC4=C C=CC=C 43)=NN2) C=C1</chem>	5.775423 92	7.503441 46
mol_190S	IV 3,4- dihidropir azol (S)	p- aminaben zaldehído	9- Acetilantr aceno	<chem>NC1=CC =C(C2CC (C3=C4C =CC=CC 4=CC4=C C=CC=C 43)=NN2) C=C1</chem>	5.775423 92	7.492146 498

mol_191	V Pirimidina - aminobenzimidazol	p-aminobenzalaldehído	9-Acetilnitraceno	<chem>NC1=CC=C(C2=C(C(C3=C4C=CC(=C4)CC(=C3)NC(=NC4=C(N3)C=CC=C4)=N2)C=C1</chem>	5.746161667	9.090404616
mol_192	VI Pirimido-benzimidazol	p-aminobenzalaldehído	9-Acetilnitraceno	<chem>NC1=CC=C(C=C1)C2=CC(C3=C4C=CC(=C4)CC5=C3C=CC=C5)=NC6=NC7=CC=CC=C7N26</chem>	5.78062798	8.444702263
mol_193	I Chalcona	p-aminobenzalaldehído	1-Acetilnaftaleno	<chem>NC1=CC=C(C(=O)C2=CC=CC=C2C(=O)C1</chem>	4.959155919	7.251970764
mol_194	II Aminopirimidina	p-aminobenzalaldehído	1-Acetilnaftaleno	<chem>NC1=CC=C(C2=C(C(C3=CC(=C4C=CC(=C4)NC(N)=N2)C=C1</chem>	4.983512249	7.465022956
mol_195R	III 2,3-dihidrobenzodiazepina (R)	p-aminobenzalaldehído	1-Acetilnaftaleno	<chem>NC1=CC=C(C2CC(C3=CC(=CC4=CC=CC=C43)=NC3=C</chem>	5.206432241	7.617354892

				<chem>C=CC=C3N2)C=C1</chem>		
mol_195S	III 2,3-dihidrobenzodiazepina (S)	p-aminabenzaldehído	1-Acetilnaftaleno	<chem>NC1=CC=C(C2CC(C3=CC=CC4=CC=CC=C43)=NC3=C(C=CC=C3N2)C=C1</chem>	5.206432241	7.617354892
mol_196R	IV 3,4-dihidropirazol (R)	p-aminabenzaldehído	1-Acetilnaftaleno	<chem>NC1=CC=C(C2CC(C3=CC=CC4=CC=CC=C43)=NN2)C=C1</chem>	5.412567539	7.91952117
mol_196S	IV 3,4-dihidropirazol (S)	p-aminabenzaldehído	1-Acetilnaftaleno	<chem>NC1=CC=C(C2CC(C3=CC=CC4=CC=CC=C43)=NN2)C=C1</chem>	5.412567539	7.917869942
mol_197	V Pirimidina-aminobenzimidazol	p-aminabenzaldehído	1-Acetilnaftaleno	<chem>NC1=CC=C(C2=C(C(C3=CC=CC4=C(C=CC=C43)=NC(NC3=NC4=C(N3)C=CC=C4)=N2)C=C1</chem>	5.411220292	8.701165349
mol_198	VI Pirimido-benzimidazol	p-aminabenzaldehído	1-Acetilnaftaleno	<chem>NC1=CC=C(C(C=C1)C2=CC(C3=CC=C(C4=C3C=CC=C4)=NC5=NC</chem>	5.419144599	8.060171531

				<chem>6=CC=C C=C6N25</chem>		
mol_199	I Chalcona	p- aminaben zaldehído	2- Acetilnaft aleno	<chem>NC1=CC =C(/C=C/ C(=O)C2= CC=C3C =CC=CC 3=C2)C= C1</chem>	4.998753 985	7.258174 205
mol_200	II Aminopiri midina	p- aminaben zaldehído	2- Acetilnaft aleno	<chem>NC1=CC =C(C2=C C(C3=CC =C4C=C C=CC4= C3)=NC(N)=N2)C= C1</chem>	5.019490 723	7.463750 118
mol_201R	III 2,3- dihidrobe nzodiaze pina (R)	p- aminaben zaldehído	2- Acetilnaft aleno	<chem>NC1=CC =C(C2CC (C3=CC= C4C=CC =CC4=C3)=NC3=C C=CC=C 3N2)C=C 1</chem>	5.246523 045	8.589259 152
mol_201S	III 2,3- dihidrobe nzodiaze pina (S)	p- aminaben zaldehído	2- Acetilnaft aleno	<chem>NC1=CC =C(C2CC (C3=CC= C4C=CC =CC4=C3)=NC3=C C=CC=C 3N2)C=C 1</chem>	5.246523 045	8.589304 809
mol_202R	IV 3,4- dihidropir azol (R)	p- aminaben zaldehído	2- Acetilnaft aleno	<chem>NC1=CC =C(C2CC (C3=CC= C4C=CC =CC4=C3)=NN2)C= C1</chem>	5.371754 517	7.826904 387

mol_202S	IV 3,4-dihidropirazol (S)	p-aminabenzaldehído	2-Acetilnftaleno	<chem>NC1=CC=C(C2CC(C3=CC=C4C=CC=CC4=C3)=NN2)C=C1</chem>	5.371754517	7.828347815
mol_203	V Pirimidina-aminobenzimidazol	p-aminabenzaldehído	2-Acetilnftaleno	<chem>NC1=CC=C(C2=C(C(C3=CC=C4C=C(C=CC4=C3)=NC(NC3=NC4=C(N3)C=CC=C4)=N2)C=C1</chem>	5.434212574	8.345538599
mol_204	VI Pirimido-benzimidazol	p-aminabenzaldehído	2-Acetilnftaleno	<chem>NC1=CC=C(C=C1)C2=CC(C3=CC(C=CC=C4)=C4C=C3)=NC5=NCC6=CC=CC=C6N25</chem>	5.439815709	7.964461894
mol_205	I Chalcona	p-aminabenzaldehído	Acetophenone	<chem>NC1=CC=C(/C=C/C(=O)C2=CC=CC=C2)C=C1</chem>	4.597386143	6.782561873
mol_206	II Aminopirimidina	p-aminabenzaldehído	Acetophenone	<chem>NC1=CC=C(C2=C(C(C3=CC=CC=C3)=NC(N)=N2)C=C1</chem>	4.683171309	7.835995603
mol_207R	III 2,3-dihidrobenzodiazepina (R)	p-aminabenzaldehído	Acetophenone	<chem>NC1=CC=C(C2CC(C3=CC=CC=C3)=NC3=CC</chem>	4.888023362	7.95178923

				<chem>=CC=C3N2)C=C1</chem>		
mol_207S	III 2,3-dihidrobenzodiazepina (S)	p-aminobenzaldehído	Acetophenone	<chem>NC1=CC=C(C2CC(C3=CC=CC=C3)=NC3=CC=CC=C3N2)C=C1</chem>	4.888023362	7.95178923
mol_208R	IV 3,4-dihidropirazol (R)	p-aminobenzaldehído	Acetophenone	<chem>NC1=CC=C(C2CC(C3=CC=CC=C3)=NN2)C=C1</chem>	5.008762188	7.214442455
mol_208S	IV 3,4-dihidropirazol (S)	p-aminobenzaldehído	Acetophenone	<chem>NC1=CC=C(C2CC(C3=CC=CC=C3)=NN2)C=C1</chem>	5.008762188	7.216562811
mol_209	V Pirimidina-aminobenzimidazol	p-aminobenzaldehído	Acetophenone	<chem>NC1=CC=C(C2=CC(C3=CC=CC=C3)=NC(NC3=NC4=C(N3)C=CC=C4)=N2)C=C1</chem>	5.056901749	7.984013374
mol_210	VI Pirimido-benzimidazol	p-aminobenzaldehído	Acetophenone	<chem>NC1=CC=C(C=C1)C2=CC(C3=CC=CC=C3)=NC4=NC5=CC=CC=C5N24</chem>	5.047299842	7.338533111
mol_211	I Chalcona	p-(Trifluorometil)benzaldehído	1-Acetilpireno	<chem>O=C(/C=C/C/C1=CC=C(C(F)(F)F)C=C1)C2=C3C4=C(C(C=C2)C=C3)C=C4</chem>	5.889492257	8.721893097

				C3)=CC=C5)C5=C C=C4C=C2		
mol_212	II Aminopirimidina	p-(Trifluorometil)benzaldehído	1-Acetilpireno	FC(F)(F)C1=CC=C(C=C1)C2=NC(N)=NC(C3=C4C5=C(C(C=C4)=CC=C6)C6=CC=C5C=C3)=C2	5.871656525	8.465626959
mol_213R	III 2,3-dihidrobenzodiazepina (R)	p-(Trifluorometil)benzaldehído	1-Acetilpireno	FC(F)(F)C1=CC=C(C=C1)C(N2)CC(C3=C4C5=C(C(C=C4)=CC=C6)C6=CC=C5C=C3)=NC7=C2C=CC=C7	6.171832661	8.818297592
mol_213S	III 2,3-dihidrobenzodiazepina (S)	p-(Trifluorometil)benzaldehído	1-Acetilpireno	FC(F)(F)C1=CC=C(C=C1)C(N2)CC(C3=C4C5=C(C(C=C4)=CC=C6)C6=CC=C5C=C3)=NC7=C2C=CC=C7	6.171832661	8.809887937
mol_214R	IV 3,4-dihidropirazol (R)	p-(Trifluorometil)benzaldehído	1-Acetilpireno	FC(F)(F)C1=CC=C(C=C1)C2CC(C3=C4C5=C(C(C=C4)=CC=C6)C6=CC=C5C=C3)=NC7=C2C=CC=C7	6.368052903	8.190469684

				<chem>C=C4)=C C=C6)C6 =CC=C5 C=C3)=N N2</chem>		
mol_214S	IV 3,4- dihidropir azol (S)	p- (Trifluoro metil)ben zaldehído	1- Acetilpire no	<chem>FC(F)(F)C 1=CC=C(C=C1)C2 CC(C3=C 4C5=C(C(C=C4)=C C=C6)C6 =CC=C5 C=C3)=N N2</chem>	6.368052 903	8.194480 715
mol_215	V Pirimidina - aminobe nzimidaz ol	p- (Trifluoro metil)ben zaldehído	1- Acetilpire no	<chem>FC(F)(F)C 1=CC=C(C2=NC(N C3=NC4= C(N3)C= CC=C4)= NC(C5=C 6C7=C(C(C=C6)=C C=C8)C8 =CC=C7 C=C5)=C 2)C=C1</chem>	6.349999 009	8.773744 474
mol_216	VI Pirimido- benzimid azol	p- (Trifluoro metil)ben zaldehído	1- Acetilpire no	<chem>FC(F)(F)C 1=CC=C(C2=CC(C 3=C4C5= C(C(C=C 4)=CC=C 6)C6=CC =C5C=C3)=NC7=N C8=C(C= CC=C8)N 27)C=C1</chem>	6.399556 526	9.162977 109
mol_217	I Chalcona	p- (Trifluoro metil)ben zaldehído	9- Acetilantr aceno	<chem>O=C(C1= C2C(C=C C=C2)=C C=C1)/C=</chem>	5.624341 01	8.333239 665

				<chem>C/C3=CC=C(C(F)(F)F)C=C3</chem>		
mol_218	II Aminopiridina	p-(Trifluorometil)benzaldehído	9-Acetilantraceno	<chem>FC(F)(F)C(C=C1)=CC=C1C2=NC(N)=NC(C3=CC=CC4=C3C=CC=C4)=C2</chem>	5.590047326	8.085618082
mol_219R	III 2,3-dihidrobenzodiazepina (R)	p-(Trifluorometil)benzaldehído	9-Acetilantraceno	<chem>FC(F)(F)C1=CC=C(C2CC(C3=C4C(C=CC=C4)=CC=C3)=NC5=C(N2)C=CC=C5)C=C1</chem>	6.005783285	8.732123161
mol_219S	III 2,3-dihidrobenzodiazepina (S)	p-(Trifluorometil)benzaldehído	9-Acetilantraceno	<chem>FC(F)(F)C1=CC=C(C2CC(C3=C4C(C=CC=C4)=CC=C3)=NC5=C(N2)C=CC=C5)C=C1</chem>	6.005783285	8.730128193
mol_220R	IV 3,4-dihidropirazol (R)	p-(Trifluorometil)benzaldehído	9-Acetilantraceno	<chem>FC(F)(F)C1=CC=C(C2NN=C(C2)C3=CC=CC4=C3C=CC=C4)C=C1</chem>	6.142490973	7.803626277
mol_220S	IV 3,4-dihidropirazol (S)	p-(Trifluorometil)benzaldehído	9-Acetilantraceno	<chem>FC(F)(F)C1=CC=C(C2NN=C(C2)C3=CC=CC4=C3C=CC=C4)C=C1</chem>	6.142490973	7.796509917

				=C4)C=C 1		
mol_221	V Pirimidina - aminobe nzimidaz ol	p- (Trifluoro metil)ben zaldehído	9- Acetilantr aceno	FC(F)(F)C 1=CC=C(C2=NC(N C3=NC4= C(N3)C= CC=C4)= NC(C5=C C=CC6= C5C=CC =C6)=C2) C=C1	6.070727 682	9.435624 782
mol_222	VI Pirimido- benzimid azol	p- (Trifluoro metil)ben zaldehído	9- Acetilantr aceno	FC(F)(F)C 1=CC=C(C2=CC(C 3=C4C=C C=CC4= CC=C3)= NC5=NC 6=CC=C C=C6N25)C=C1	6.133921 684	8.765559 024
mol_223	I Chalcona	p- (Trifluoro metil)ben zaldehído	1- Acetilnaft aleno	O=C(C1= CC=C(C= CC=C2)C 2=C1)/C= C/C3=CC =C(C(F)(F)F)C=C3	5.301142 696	7.562789 677
mol_224	II Aminopiri midina	p- (Trifluoro metil)ben zaldehído	1- Acetilnaft aleno	FC(F)(F)C (C=C1)=C C=C1C2= NC(N)=N C(C3=CC (C=CC=C 4)=C4C= C3)=C2	5.303258 909	7.621084 61
mol_225R	III 2,3- dihidrobe nzodiaze pina (R)	p- (Trifluoro metil)ben zaldehído	1- Acetilnaft aleno	FC(F)(F)C 1=CC=C(C2CC(C3 =CC=C(C =CC=C4) C4=C3)=	5.581144 967	7.901516 033

				NC5=C(N2)C=CC=C5)C=C1		
mol_225S	III 2,3-dihidrobenzodiazepina (S)	p-(Trifluorometil)benzaldehído	1-Acetilnaftaleno	FC(F)(F)C1=CC=C(C2CC(C3=CC=C(C4=CC=C4)C4=C3)=NC5=C(N2)C=CC=C5)C=C1	5.581144967	7.916767269
mol_226R	IV 3,4-dihidropirazol (R)	p-(Trifluorometil)benzaldehído	1-Acetilnaftaleno	FC(F)(F)C1=CC=C(C2NN=C(C2)C3=C(C4=C(C=C3)C=CC=C4)C=C1	5.839554697	8.110776564
mol_226S	IV 3,4-dihidropirazol (S)	p-(Trifluorometil)benzaldehído	1-Acetilnaftaleno	FC(F)(F)C1=CC=C(C2NN=C(C2)C3=C(C4=C(C=C3)C=CC=C4)C=C1	5.839554697	8.119896351
mol_227	V Pirimidina-aminobenzimidazol	p-(Trifluorometil)benzaldehído	1-Acetilnaftaleno	FC(F)(F)C1=CC=C(C2=NC(NC3=NC4=C(N3)C=CC=C4)=NC(C5=C(C6=C(C=C5)C=CC=C6)=C2)C=C1	5.742374352	8.666504404
mol_228	VI Pirimido-benzimidazol	p-(Trifluorometil)benzaldehído	1-Acetilnaftaleno	FC(F)(F)C1=CC=C(C2=CC(C3=CC=C4C(C=CC=	5.78649227	8.044049709

				C4)=C3)= NC5=NC 6=CC=C C=C6N25)C=C1		
mol_229	I Chalcona	p- (Trifluoro metil)ben zaldehído	2- Acetilnaft aleno	O=C(/C= C/C1=CC =C(C=C1) C(F)(F)F) C2=CC= CC=C2	5.301986 98	7.559065 907
mol_230	II Aminopiri midina	p- (Trifluoro metil)ben zaldehído	2- Acetilnaft aleno	FC(F)(C1 =CC=C(C 2=NC(N)= NC(C3=C C=CC=C 3)=C2)C= C1)F	5.304835 092	7.350526 38
mol_231R	III 2,3- dihidrobe nzodiaze pina (R)	p- (Trifluoro metil)ben zaldehído	2- Acetilnaft aleno	FC(F)(C1 =CC=C(C =C1)C2C C(C3=CC =CC=C3) =NC4=C(C=CC=C 4)N2)F	5.596044 374	8.718237 494
mol_231S	III 2,3- dihidrobe nzodiaze pina (S)	p- (Trifluoro metil)ben zaldehído	2- Acetilnaft aleno	FC(F)(C1 =CC=C(C =C1)C2C C(C3=CC =CC=C3) =NC4=C(C=CC=C 4)N2)F	5.596044 374	8.718232 438
mol_232R	IV 3,4- dihidropir azol (R)	p- (Trifluoro metil)ben zaldehído	2- Acetilnaft aleno	CC(F)(C1 =CC=C(C =C1)C2N N=C(C3= CC=CC= C3)C2)F	5.798741 724	7.842716 099
mol_232S	IV 3,4- dihidropir azol (S)	p- (Trifluoro metil)ben zaldehído	2- Acetilnaft aleno	CC(F)(C1 =CC=C(C =C1)C2N N=C(C3= CC=CC= C3)C2)F	5.798741 724	7.864227 432

				<chem>CC=CC=C3C2)F</chem>		
mol_233	V Pirimidina - aminobenzimidazol	p- (Trifluorometil)benzaldehído	2-Acetilnaftaleno	<chem>FC(F)(C1=CC=C(C=C1)C2=NC(NC3=NC4=C(C=CC=C4)N3)=NC(C5=CC=CC=C5)=C2)F</chem>	5.741099245	8.404708
mol_234	VI Pirimido-benzimidazol	p- (Trifluorometil)benzaldehído	2-Acetilnaftaleno	<chem>FC(F)(C1=CC=C(C=C1)C2=CC(C3=CC=CC3)=NC4=NC5=CC=CC=C5N24)F</chem>	5.779084495	8.016562558
mol_235	I Chalcona	p- (Trifluorometil)benzaldehído	Acetophenone	<chem>O=C(C1=CC2(C=C(C=C2)=CC=C1)/C=C/C3=CC=C(C(F)(F)F)C=C3</chem>	4.938313753	6.812796203
mol_236	II Aminopirimidina	p- (Trifluorometil)benzaldehído	Acetophenone	<chem>FC(F)(F)C(C=C1)=CC=C1C2=NC(N)=NC(C3=CC=CC4=C3C=CC=C4)=C2</chem>	5.010863191	7.737065071
mol_237R	III 2,3-dihidrobenzodiazepina (R)	p- (Trifluorometil)benzaldehído	Acetophenone	<chem>FC(F)(F)C1=CC=C(C2CC(C3=C4C(C=CC=C4)=CC=C3)=NC5=C(N</chem>	5.275015518	8.24075011

				2)C=CC=C5)C=C1		
mol_237S	III 2,3-dihidrobenzodiazepina (S)	p-(Trifluorometil)benzaldehído	Acetophenone	FC(F)(F)C1=CC=C(C2CC(C3=C4C(C=CC=C4)=CC=C3)=NC5=C(N2)C=CC=C5)C=C1	5.275015518	8.244057897
mol_238R	IV 3,4-dihidropirazol (R)	p-(Trifluorometil)benzaldehído	Acetophenone	FC(F)(F)C1=CC=C(C2NN=C(C2)C3=CC=CC4=C3C=CC=C4)C=C1	5.500413424	7.409385632
mol_238S	IV 3,4-dihidropirazol (S)	p-(Trifluorometil)benzaldehído	Acetophenone	FC(F)(F)C1=CC=C(C2NN=C(C2)C3=CC=CC4=C3C=CC=C4)C=C1	5.500413424	7.406223815
mol_239	V Pirimidina-aminobenzimidazol	p-(Trifluorometil)benzaldehído	Acetophenone	FC(F)(F)C1=CC=C(C2=NC(NC3=NC4=CC(N3)C=CC=C4)=NC(C5=CC=CC6=C5C=CC=C6)=C2)C=C1	5.389720337	7.937695466
mol_240	VI Pirimido-benzimidazol	p-(Trifluorometil)benzaldehído	Acetophenone	FC(F)(F)C1=CC=C(C2=CC(C3=C4C=CC=CC4=C3)=	5.423697888	7.562718138

				NC5=NC 6=CC=C C=C6N25)C=C1		
BZD_001	VII Benzodia zepina	p- Nitrobenz aldehído	1- Acetilpire no	[O-][N+](C1= CC=C(C2 =CC(C3= CC=C4C 5=C3C=C C6=C5C(C=C4)=C C=C6)=N C7=C(C= CC=C7)N 2)C=C1)= O	5.741311 79	8.360086 006
BZD_002	VII Benzodia zepina	p- Nitrobenz aldehído	9- Acetilntr aceno	[O-][N+](C1= CC=C(C2 =CC(C3= C4C(C=C C=C4)=C C5=C3C= CC=C5)= NC6=C(C =CC=C6) N2)C=C1) =O	5.456521 261	7.981963 638
BZD_003	VII Benzodia zepina	p- Nitrobenz aldehído	1- Acetilnaft aleno	[O-][N+](C1= CC=C(C2 =CC(C3= C4C(C=C C=C4)=C C=C3)=N C5=C(C= CC=C5)N 2)C=C1)= O	5.117699 185	7.248175 772
BZD_004	VII Benzodia zepina	p- Nitrobenz aldehído	2- Acetilnaft aleno	[O-][N+](C1= CC=C(C2 =CC(C3= C4C(C=C C=C4)=C C=C3)=N C5=C(C= CC=C5)N 2)C=C1)= O	5.151879 097	7.250225 46

				<chem>CC=C(C=CC=C4)C4=C3)=NC5=C(C=CC=C5)N2)C=C1)=O</chem>		
BZD_005	VII Benzodia zepina	p-Nitrobenz aldehído	Acetophe none	<chem>[O-][N+](C1=CC=C(C2=CC(C3=CC=CC=C3)=NC4=C(C=CC=C4)N2)C=C1)=O</chem>	4.782820 11	6.547662 377
BZD_006	VII Benzodia zepina	p-Dimetila minobenz aldehído	1-Acetilpire no	<chem>CN(C)C1=CC=C(C2=CC(C3=CC=C4C5=C3C=CC6=C5C(C=C4)=CC=C6)=NC7=C(C=CC=C7)N2)C=C1</chem>	5.829097 586	8.593737 942
BZD_007	VII Benzodia zepina	p-Dimetila minobenz aldehído	9-Acetilntr aceno	<chem>CN(C)C1=CC=C(C2=CC(C3=C4C(C=CC=C4)=CC5=C3C=CC=C5)=NC6=C(C=CC=C6)N2)C=C1</chem>	5.576360 144	8.216948 174
BZD_008	VII Benzodia zepina	p-Dimetila minobenz aldehído	1-Acetilnaft aleno	<chem>CN(C)C1=CC=C(C2=CC(C3=CC=CC4=C3C=CC=C4)=N</chem>	5.222948 15	7.507552 98

				C5=C(C=CC=C5)N2)C=C1		
BZD_009	VII Benzodia zepina	p-Dimetila minobenz aldehído	2-Acetilnaft aleno	CN(C)C1=CC=C(C2=CC(C3=CC(C=C C=C4)=C4C=C3)=NC5=C(C=CC=C5)N2)C=C1	5.26603398	7.509920694
BZD_010	VII Benzodia zepina	p-Dimetila minobenz aldehído	Acetophe none	CN(C)C1=CC=C(C2=CC(C3=CC=CC=C3)=NC4=C(C=C C=C4)N2)C=C1	4.88746868	6.827780006
BZD_011	VII Benzodia zepina	Vainillina	1-Acetilpire no	OC1=CC=C(C2=C C(C3=C(C=CC4=C5C(C=C6)=CC=C4)C5=C6C=C3)=NC7=C(C=CC=C7)N2)C=C1O C	5.642982959	7.910684875
BZD_012	VII Benzodia zepina	Vainillina	9-Acetilant raceno	OC1=CC=C(C2=C C(C3=C(C=CC=C4)C4=CC5=C3C=C C=C5)=NC6=C(C=CC=C6)N2)C=C1O C	5.367463658	7.535262222

BZD_013	VII Benzodia zepina	Vainillina	1- Acetilnaft aleno	<chem>OC1=CC=C(C2=CC(C3=CC=C(C4=CC3)=NC5=C(C=CC=C5)N2)C=C1OC</chem>	5.077484 048	6.870470 018
BZD_014	VII Benzodia zepina	Vainillina	2- Acetilnaft aleno	<chem>OC1=CC=C(C2=CC(C3=CC=CC4=CC=C4)=NC5=C(C=CC=C5)N2)C=C1OC</chem>	5.043991 793	6.867886 671
BZD_015	VII Benzodia zepina	Vainillina	Acetophe none	<chem>OC1=CC=C(C2=CC(C3=CC=C3)=NC4=C(C=CC=C4)N2)C=C1OC</chem>	4.731733 337	6.227528 703
BZD_016	VII Benzodia zepina	Cinamald ehído	1- Acetilpire no	<chem>C1(C2=CC3C4=C(C=C2)C=CC5=CC=CC(C=C3)=C45)=NC6=CC=CC=C6N1C/C=C/C7=CC=C(C=C7)=C</chem>	5.866954 185	8.891965 394
BZD_017	VII Benzodia zepina	Cinamald ehído	9- Acetilantr aceno	<chem>C1(C2=CC3C(C=CC=C3)=CC4=C2C=CC=C4)=NC5=CC=</chem>	5.630943 205	8.510959 12

				CC=C5N C(/C=C/C 6=CC=C C=C6)=C 1		
BZD_018	VII Benzodia zepina	Cinamald ehído	1- Acetilnaft aleno	C1(C2=C 3C(C=CC =C3)=CC =C2)=NC 4=CC=C C=C4NC(/C=C/C5= CC=CC= C5)=C1	5.276014 479	7.766058 916
BZD_019	VII Benzodia zepina	Cinamald ehído	2- Acetilnaft aleno	C1(C2=C C=C3C(C =CC=C3) =C2)=NC 4=CC=C C=C4NC(/C=C/C5= CC=CC= C5)=C1	5.337044 471	7.767893 797
BZD_020	VII Benzodia zepina	Cinamald ehído	Acetophe none	C1(C2=C C=CC=C 2)=NC3= CC=CC= C3NC(/C =C/C4=C C=CC=C 4)=C1	4.939386 894	7.050814 149
BZD_021	VII Benzodia zepina	Salicilald ehído	1- Acetilpire no	OC1=CC =CC=C1 C2=CC(C 3=C(C=C C4=C5C(C=C6)=C C=C4)C5 =C6C=C3)=NC7=C(C=CC=C 7)N2	5.531975 74	8.203072 527

BZD_022	VII Benzodia zepina	Salicilald ehído	9- Acetilantr aceno	<chem>OC1=CC=CC=C1C2=CC(C3=C(C=CC=C4)C4=CC5=C3C=CC=C5)=NC6=C(C=CC=C6)N2</chem>	5.263776 085	7.834285 136
BZD_023	VII Benzodia zepina	Salicilald ehído	1- Acetilnaft aleno	<chem>OC1=CC=CC=C1C2=CC(C3=C(C=CC=C4)C4=CC=C3)=NC5=C(C=CC=C5)N2</chem>	4.916934 672	7.143686 932
BZD_024	VII Benzodia zepina	Salicilald ehído	2- Acetilnaft aleno	<chem>OC1=CC=CC=C1C2=CC(C3=CC(C=CC=C4)=C4C=C3)=NC5=C(C=CC=C5)N2</chem>	4.953157 625	7.133414 941
BZD_025	VII Benzodia zepina	Salicilald ehído	Acetophe none	<chem>OC1=CC=CC=C1C2=CC(C3=CC=CC=C3)=NC4=C(C=CC=C4)N2</chem>	4.579233 718	6.470859 384
BZD_026	VII Benzodia zepina	4- Isopropilb enzaldehído	1- Acetilpire no	<chem>CC(C)C1=CC=C(C2=CC(C3=CC=C4C5=C3C=CC6=CC=CC(C=C4)=C65)=</chem>	6.134116 977	8.889710 843

				<chem>NC7=C(C=CC=C7)N2C=C1</chem>		
BZD_027	VII Benzodia zepina	4- Isopropilb enzaldehí do	9- Acetilntr aceno	<chem>CC(C)C1=CC=C(C2=CC(C3=C(C=CC=C4)C4=CC5=C3C=CC=C5)=NC6=C(C=CC=C6)N2)C=C1</chem>	5.898532 543	8.509019 719
BZD_028	VII Benzodia zepina	4- Isopropilb enzaldehí do	1- Acetilnaft aleno	<chem>CC(C)C1=CC=C(C2=CC(C3=C(C=CC=C4)C4=CC=C3)=NC5=C(C=CC=C5)N2)C=C1</chem>	5.527141 918	7.768983 484
BZD_029	VII Benzodia zepina	4- Isopropilb enzaldehí do	2- Acetilnaft aleno	<chem>CC(C)C1=CC=C(C2=CC(C3=CC=C4C(C=CC=C4)=C3)=NC5=C(C=CC=C5)N2)C=C1</chem>	5.569850 565	7.767359 163
BZD_030	VII Benzodia zepina	4- Isopropilb enzaldehí do	Acetophe none	<chem>CC(C)C1=CC=C(C2=CC(C3=CC=CC=C3)=NC4=C(C=CC=C4)N2)C=C1</chem>	5.177412 466	7.057523 58
BZD_031	VII Benzodia zepina	p- aminaben zaldehído	1- Acetilpire no	<chem>NC1=CC=C(C2=CC(C3=CC=C4C=C</chem>	5.662741 267	8.196131 782

				<chem>C5=CC=CC6=CC=C3C4=C56)=NC3=CC=CC=C3N2)C=C1</chem>		
BZD_032	VII Benzodia zepina	p- aminaben zaldehído	9- Acetilantr aceno	<chem>NC1=CC=C(C2=CC(C3=C4C=CC=C4=CC=CC=C43)=NC3=CC=C(C=C3N2)C=C1</chem>	5.409143 865	7.825163 259
BZD_033	VII Benzodia zepina	p- aminaben zaldehído	1- Acetilnaft aleno	<chem>NC1=CC=C(C2=CC(C3=CC=CC4=C(C=CC=C43)=NC3=CC=CC=C3N2)C=C1</chem>	5.067866 377	7.164071 904
BZD_034	VII Benzodia zepina	p- aminaben zaldehído	2- Acetilnaft aleno	<chem>NC1=CC=C(C2=CC(C3=CC=C4C=CC4=C3)=NC3=CC=CC=C3N2)C=C1</chem>	5.122023 147	7.161675 979
BZD_035	VII Benzodia zepina	p- aminaben zaldehído	Acetophe none	<chem>NC1=CC=C(C2=CC(C3=CC=C3)=NC3=C(C=CC=C3N2)C=C1</chem>	4.742444 035	6.532361 857

BZD_036	VII Benzodia zepina	p- (Trifluoro metil)ben zaldehído	1- Acetilpire no	<chem>FC(F)(C1=CC=C(C2=CC(C3=C4C5=C6C(C=C4)=CC=CC6=CC=C5C=C3)=NC7=C(C=CC=C7)N2)C=C1)F</chem>	6.027756 832	8.441451 182
BZD_037	VII Benzodia zepina	p- (Trifluoro metil)ben zaldehído	9- Acetilantr aceno	<chem>FC(F)(C1=CC=C(C=C1)C2=CC(C3=C4C5=C3C=CC6=C5C(C=C4)=CC=C6)=NC7=C(N2)C=CC=C7)F</chem>	5.740027 039	8.312307 959
BZD_038	VII Benzodia zepina	p- (Trifluoro metil)ben zaldehído	1- Acetilnaft aleno	<chem>FC(F)(C1=CC=C(C=C1)C2=CC(C3=C4C(C=CC=C4)=CC5=C3C=CC=C5)=NC6=C(N2)C=CC=C6)F</chem>	5.415784 079	7.535082 598
BZD_039	VII Benzodia zepina	p- (Trifluoro metil)ben zaldehído	2- Acetilnaft aleno	<chem>FC(F)(C1=CC=C(C=C1)C2=CC(C3=C4C(C=CC=C4)=CC=C3)=NC5=C(N2)C=CC=C5)F</chem>	5.440582 218	7.279884 085

BZD_040	VII Benzodia zepina	p- (Trifluoro metil)ben zaldehído	Acetophe none	<chem>FC(F)(C1=CC=C(C=C1)C2=CC(C3=CC=C4C=CC=CC4=C3)=NC5=C(N2)C=CC=C5)F</chem>	5.097854 76	6.808249 362
PYZ_001	VIII Pirazol	p- Nitrobenz aldehído	1- Acetilpire no	<chem>[O-][N+](C1=CC=C(C2=CC(C3=C4C=CC5=CC=C(C6=CC=C(C(C=C3)C4=C56)=NN2)C=C1)=O</chem>	5.943380 472	7.757914 584
PYZ_002	VIII Pirazol	p- Nitrobenz aldehído	9- Acetilnitr aceno	<chem>[O-][N+](C1=CC=C(C2=CC(C3=C4C(C=C(C=C4)=C5=C3C=CC=C5)=NN2)C=C1)=O</chem>	5.726388 976	7.382688 772
PYZ_003	VIII Pirazol	p- Nitrobenz aldehído	1- Acetilnaft aleno	<chem>[O-][N+](C1=CC=C(C2=CC(C3=CC=CC4=C3C=CC=C4)=N2)C=C1)=O</chem>	5.349228 032	6.680213 143
PYZ_004	VIII Pirazol	p- Nitrobenz aldehído	2- Acetilnaft aleno	<chem>[O-][N+](C1=CC=C(C2=CC(C3=</chem>	5.315744 228	6.675567 146

				<chem>CC4=C(C=C3)C=C(C=C4)=NN2)C=C1)=O</chem>		
PYZ_005	VIII Pirazol	p- Nitrobenz aldehído	Acetophe none	<chem>[O-][N+](C1=CC=C(C2=CC(C3=CC=CC=C3)=NN2)C=C1)=O</chem>	4.891455 24	5.997617 365
PYZ_006	VIII Pirazol	p- Dimetila minobenz aldehído	1- Acetilpire no	<chem>CN(C)C1=CC=C(C2=CC(C3=C4C=C(C5=CC=CC6=CC=C(C=C3)C4=C56)=NN2)C=C1</chem>	6.024110 359	8.017101 202
PYZ_007	VIII Pirazol	p- Dimetila minobenz aldehído	9- Acetilantr aceno	<chem>CN(C)C1=CC=C(C2=CC(C3=C4C(C=CC=C4)=CC5=C3C=CC=C5)=NN2)C=C1</chem>	5.863273 348	7.642126 5
PYZ_008	VIII Pirazol	p- Dimetila minobenz aldehído	1- Acetilnaft aleno	<chem>CN(C)C1=CC=C(C2=CC(C3=CC=CC4=C3C=C(C=C4)=NN2)C=C1</chem>	5.452034 198	6.960192 728
PYZ_009	VIII Pirazol	p- Dimetila minobenz aldehído	2- Acetilnaft aleno	<chem>CN(C)C1=CC=C(C2=CC(C3=CC4=C(C=C3)C=CC=C4)=</chem>	5.415585 987	6.959852 767

				NN2)C=C 1		
PYZ_010	VIII Pirazol	p- Dimetila minobenz aldehído	Acetophe none	CN(C)C1 =CC=C(C 2=CC(C3 =CC=CC =C3)=NN 2)C=C1	4.997632 957	6.305321 854
PYZ_011	VIII Pirazol	Vainillina	1- Acetilpire no	OC1=CC =C(C2=C C(C3=C4 C=CC5= CC=CC6 =CC=C(C =C3)C4= C56)=NN 2)C=C1O C	5.883197 616	7.366891 771
PYZ_012	VIII Pirazol	Vainillina	9- Acetilantr aceno	OC1=CC =C(C2=C C(C3=C4 C(C=CC= C4)=CC5 =C3C=C C=C5)=N N2)C=C1 OC	5.673694 336	7.002820 742
PYZ_013	VIII Pirazol	Vainillina	1- Acetilnaft aleno	OC1=CC =C(C2=C C(C3=CC =CC4=C3 C=CC=C 4)=NN2)C =C1OC	5.330338 873	6.354813 006
PYZ_014	VIII Pirazol	Vainillina	2- Acetilnaft aleno	OC1=C(O C)C=C(C 2=CC(C3 =CC4=C(C=C3)C= CC=C4)= NN2)C=C 1	5.306443 303	6.362035 088
PYZ_015	VIII Pirazol	Vainillina	Acetophe none	OC1=CC =C(C2=C	4.946018 817	5.745980 652

				<chem>C(C3=CC=CC=C3)=NN2)C=C1OC</chem>		
PYZ_016	VIII Pirazol	Cinamald ehído	1-Acetilpireno	<chem>C1(/C=C/C2=CC(C3=C4C=C5=CC=CC6=CC=C(C=C3)C4=C56)=NN2)=CC=CC=C1</chem>	6.160918459	8.282072288
PYZ_017	VIII Pirazol	Cinamald ehído	9-Acetilntraceno	<chem>C1(/C=C/C2=CC(C3=C4C(C=CC=C4)=CC5=C3C=CC=C5)=NN2)=CC=CC=C1</chem>	6.030545275	7.900858976
PYZ_018	VIII Pirazol	Cinamald ehído	1-Acetilnaftaleno	<chem>C1(/C=C/C2=CC(C3=CC=C4=C3C=CC=C4)=NN2)=CC=CC=C1</chem>	5.668379973	7.210546084
PYZ_019	VIII Pirazol	Cinamald ehído	2-Acetilnaftaleno	<chem>C1(/C=C/C2=CC(C3=CC4=C(C=C3)C=CC=C4)=NN2)=CC=CC=C1</chem>	5.611314795	7.188226448
PYZ_020	VIII Pirazol	Cinamald ehído	Acetophenone	<chem>C1(/C=C/C2=CC(C3=CC=C(C=C3)=NN2)=CC=CC=C1</chem>	5.250928333	6.518237055

				<chem>C4=C56)=NN2)C=C1</chem>		
PYZ_027	VIII Pirazol	4-Isopropilbenzaldehído	9-Acetilnitraceno	<chem>CC(C)C1=CC=C(C2=CC(C3=C4C(C=CC=C4)=CC5=C3C=CC=C5)=NN2)C=C1</chem>	6.170818653	7.906542922
PYZ_028	VIII Pirazol	4-Isopropilbenzaldehído	1-Acetilnaftaleno	<chem>CC(C)C1=CC=C(C2=CC(C3=CC=CC4=C3C=CC=C4)=NN2)C=C1</chem>	5.718054566	7.191875551
PYZ_029	VIII Pirazol	4-Isopropilbenzaldehído	2-Acetilnaftaleno	<chem>CC(C)C1=CC=C(C2=CC(C3=CC4=C(C=C3)C=CC=C4)=NN2)C=C1</chem>	5.671346959	7.191773943
PYZ_030	VIII Pirazol	4-Isopropilbenzaldehído	Acetophenone	<chem>CC(C)C1=CC=C(C2=CC(C3=CC=CC=C3)=NN2)C=C1</chem>	5.223573165	6.504113501
PYZ_031	VIII Pirazol	p-aminobenzaldehído	1-Acetilpireno	<chem>NC1=CC=C(C2=C(C3=CC=C4C=C5C=CC=CC6=CC=C3C4=C56)=NN2)C=C1</chem>	5.869592232	7.659924205

PYZ_032	VIII Pirazol	p- aminaben zaldehído	9- Acetilantr aceno	<chem>NC1=CC=C(C2=C(C(C3=C4C=CC=C4=CC=CC(=C43)=NN2)C=C1</chem>	5.711368 225	7.293441 52
PYZ_033	VIII Pirazol	p- aminaben zaldehído	1- Acetilnaft aleno	<chem>NC1=CC=C(C2=C(C(C3=CC=CC4=C(C=CC=C43)=NN2)C=C1</chem>	5.320813 259	6.659002 571
PYZ_034	VIII Pirazol	p- aminaben zaldehído	2- Acetilnaft aleno	<chem>NC1=CC=C(C2=C(C(C3=CC=C4C=C(C=CC4=C3)=NN2)C=C1</chem>	5.290573 556	6.659911 231
PYZ_035	VIII Pirazol	p- aminaben zaldehído	Acetophe none	<chem>NC1=CC=C(C2=C(C(C3=CC=CC=C3)=NN2)C=C1</chem>	4.868173 938	6.052351 047
PYZ_036	VIII Pirazol	p- (Trifluoro metil)ben zaldehído	1- Acetilpire no	<chem>FC(F)(C1=CC=C(C2=NN=C(C3=C4C5=C6C(C=C4)=CC=CC6=CC=C5C=C3)C2)C=C1)F</chem>	6.265067 537	7.535922 17
PYZ_037	VIII Pirazol	p- (Trifluoro metil)ben zaldehído	9- Acetilantr aceno	<chem>FC(C1=CC=C(C2=CC(C3=CC4=CC=CC5=C3C=C</chem>	6.037904 886	7.424458 595

				<chem>C=C5)=N N2)C=C1) (F)F</chem>		
PYZ_038	VIII Pirazol	p- (Trifluoro metil)ben zaldehído	1- Acetilnaft aleno	<chem>FC(C1=C C=C(C2= CC(C3=C 4C=CC= CC4=CC =C3)=NN 2)C=C1)(F)F</chem>	5.700931 706	6.680740 482
PYZ_039	VIII Pirazol	p- (Trifluoro metil)ben zaldehído	2- Acetilnaft aleno	<chem>FC(C1=C C=C(C2= CC(C3=C C(C=CC= C4)=C4C =C3)=NN 2)C=C1)(F)F</chem>	5.664913 269	6.681053 066
PYZ_040	VIII Pirazol	p- (Trifluoro metil)ben zaldehído	Acetophe none	<chem>FC(C1=C C=C(C2= CC(C3=C C=CC=C 3)=NN2)C =C1)(F)F</chem>	5.301299 789	5.965991 117

Anexo III

Molécula	XP GScore
BZD_001	-6.184
BZD_002	-5.335
BZD_003	-7.755
BZD_004	-6.686
BZD_005	-5.769
BZD_006	-7.953
BZD_007	-5.944
BZD_008	-6.317
BZD_009	-4.897
BZD_010	-4.67
BZD_011	-7.738
BZD_012	-7.106
BZD_013	-6.947
BZD_014	-6.911

BZD_015	-8.066
BZD_016	-6.604
BZD_017	-5.088
BZD_018	-7.215
BZD_019	-6.55
BZD_020	-6.182
BZD_021	-9.42
BZD_022	-6.978
BZD_023	-7.51
BZD_024	-7.078
BZD_025	-9.185
BZD_026	-7.971
BZD_027	-8.348
BZD_028	-6.596
BZD_029	-4.908

BZD_030	-6.046
BZD_031	-7.765
BZD_032	-7.823
BZD_033	-8.366
BZD_034	-6.827
BZD_035	-5.327
BZD_036	-9.007
BZD_037	-9.669
BZD_038	-7.011
BZD_039	-7.285
BZD_040	-6.783
mol_001	-5.037
mol_002	-7.754
mol_003R	-5.34
mol_003S	-5.726

mol_004R	-8.716
mol_004S	-7.817
mol_005	-6.821
mol_006	-6.654
mol_007	-9.672
mol_008	-7.145
mol_009R	-5.708
mol_009S	-5.449
mol_010R	-7.142
mol_010S	-5.68
mol_011	-6.345
mol_012	-7.454
mol_013	-5.994
mol_014	-7.215
mol_015R	-5.768
mol_015S	-7.904
mol_016R	-6.176
mol_016S	-6.906
mol_017	-8.665
mol_018	-4.666
mol_019	-5.927
mol_020	-6.011
mol_021R	-5.475
mol_021S	-6.818
mol_022R	-6.392
mol_022S	-7.477
mol_023	-8.293
mol_024	-6.063
mol_025	-5.013
mol_026	-7.579
mol_027R	-6.503
mol_027S	-7.496
mol_028R	-7.311
mol_028S	-6.453
mol_029	-7.252
mol_030	-5.762
mol_031	-6.726
mol_032	-7.274
mol_033R	-5.431
mol_033S	-4.86
mol_034R	-8.285

mol_034S	-7.01
mol_035	-7.464
mol_036	-4.982
mol_037	-8.11
mol_038	-7.576
mol_039R	-2.689
mol_039S	-5.714
mol_040R	-7.036
mol_040S	-6.008
mol_041	-4.545
mol_042	-6.482
mol_043	-7.353
mol_044	-8.118
mol_045R	-7.321
mol_045S	-7.974
mol_046R	-7.223
mol_046S	-6.254
mol_047	-6.202
mol_048	-5.345
mol_049	-6.035
mol_050	-7.467
mol_051R	-6.212
mol_051S	-5.968
mol_052R	-7.677
mol_052S	-5.897
mol_053	-4.616
mol_054	-6.254
mol_055	-5.183
mol_056	-6.294
mol_057R	-5.503
mol_057S	-6.358
mol_058R	-7.311
mol_058S	-5.67
mol_059	-4.828
mol_060	-5.486
mol_061	-6.396
mol_062	-6.061
mol_063R	-7.911
mol_063S	-7.098
mol_064R	-7.322
mol_064S	-7.264

mol_065	-6.472
mol_066	-8.169
mol_067	-7.152
mol_068	-7.217
mol_069R	-7.1
mol_069S	-5.885
mol_070R	-7.5
mol_070S	-7.034
mol_071	-6.546
mol_072	-8.453
mol_073	-7.755
mol_074	-7.658
mol_075R	-7.426
mol_075S	-6.784
mol_076R	-7.466
mol_076S	-7.316
mol_077	-7.022
mol_078	-6.789
mol_079	-6.12
mol_080	-8.582
mol_081R	-5.811
mol_081S	-7.933
mol_082R	-8.007
mol_082S	-6.103
mol_083	-6.249
mol_084	-7.274
mol_085	-7.196
mol_086	-9.3
mol_087R	-5.24
mol_087S	-7.382
mol_088R	-8.58
mol_088S	-6.769
mol_089	-7.708
mol_090	-7.723
mol_091	-6.398
mol_092	-4.902
mol_093R	-7.462
mol_093S	-7.683
mol_094R	-6.289
mol_094S	-6.609
mol_095	-5.759

mol_096	-5.532
mol_097	-8.272
mol_098	-7.509
mol_099R	-6.186
mol_099S	-7.802
mol_100R	-7.008
mol_100S	-5.654
mol_101	-6.975
mol_102	-4.505
mol_103	-5.816
mol_104	-6.408
mol_105R	-7.141
mol_105S	-7.645
mol_106R	-7.019
mol_106S	-6.763
mol_107	-5.57
mol_108	-6.398
mol_109	-5.341
mol_110	-5.126
mol_111R	-8.063
mol_111S	-7.835
mol_112R	-6.843
mol_112S	-7.583
mol_113	-7.167
mol_114	-5.754
mol_115	-6.607
mol_116	-5.821
mol_117R	-8.193
mol_117S	-7.908
mol_118R	-6.564
mol_118S	-6.516
mol_119	-6.123
mol_120	-6.623
mol_121	-7.258
mol_122	-6.975
mol_123R	-7.751
mol_123S	-6.751
mol_124R	-7.874
mol_124S	-7.654
mol_125	-8.21
mol_126	-7.937

mol_127	-7.911
mol_128	-7.557
mol_129R	-7.396
mol_129S	-5.879
mol_130R	-7.243
mol_130S	-7.345
mol_131	-8.714
mol_132	-6.814
mol_133	-6.093
mol_134	-8.124
mol_135R	-7.856
mol_135S	-7.442
mol_136R	-6.711
mol_136S	-7.146
mol_137	-7.8
mol_138	-8.352
mol_139	-7.194
mol_140	-7.969
mol_141R	-7.212
mol_141S	-6.992
mol_142R	-7.407
mol_142S	-6.891
mol_143	-6.295
mol_144	-7.236
mol_145	-5.971
mol_146	-7.995
mol_147R	-6.653
mol_147S	-7.049
mol_148R	-5.936
mol_148S	-7.03
mol_149	-7.826
mol_150	-6.28
mol_151	-7.428
mol_152	-7.566
mol_153R	-3.833
mol_153S	-5.221
mol_154R	-7.689
mol_154S	-6.454
mol_155	-5.029
mol_156	-6.529
mol_157	-7.036

mol_158	-6.72
mol_159R	-5.246
mol_159S	-6.269
mol_160R	-7.532
mol_160S	-6.574
mol_161	-4.22
mol_162	-6.102
mol_163	-7.354
mol_164	-8.628
mol_165R	-5.979
mol_165S	-7.32
mol_166R	-7.616
mol_166S	-6.591
mol_167	-7.758
mol_168	-6.28
mol_169	-8.281
mol_170	-6.241
mol_171R	-7.573
mol_171S	-6.361
mol_172R	-7.736
mol_172S	-5.405
mol_173	-5.403
mol_174	-5.846
mol_175	-5.339
mol_176	-6.419
mol_177R	-5.02
mol_177S	-7.435
mol_178R	-6.45
mol_178S	-5.301
mol_179	-4.403
mol_180	-4.331
mol_181	-5.511
mol_182	-7.898
mol_183R	-8.177
mol_183S	-7.427
mol_184R	-7.349
mol_184S	-7.934
mol_185	-7.134
mol_186	-5.697
mol_187	-8.617
mol_188	-6.668

mol_189R	-7.635
mol_189S	-7.69
mol_190R	-6.54
mol_190S	-6.303
mol_191	-7.199
mol_192	-7.542
mol_193	-6.773
mol_194	-7.391
mol_195R	-8.134
mol_195S	-6.843
mol_196R	-7.91
mol_196S	-6.851
mol_197	-6.777
mol_198	-6.816
mol_199	-6.53
mol_200	-6.751
mol_201R	-8.339
mol_201S	-6.993
mol_202R	-7.247
mol_202S	-6.845
mol_203	-7.232
mol_204	-7.737
mol_205	-5.915
mol_206	-7.648
mol_207R	-6.644
mol_207S	-6.548
mol_208R	-9.46
mol_208S	-6.9
mol_209	-5.999
mol_210	-7.893
mol_211	-6.966
mol_212	-6.56
mol_213R	-6.192
mol_213S	-5.988
mol_214R	-6.567
mol_214S	-6.47
mol_215	-8.773

mol_216	-5.757
mol_217	-8.456
mol_218	-10.048
mol_219R	-6.21
mol_219S	-7.998
mol_220R	-6.636
mol_220S	-8.406
mol_221	-8.246
mol_222	-6.116
mol_223	-8.93
mol_224	-8.021
mol_225R	-9.317
mol_225S	-10.973
mol_226R	-6.838
mol_226S	-7.184
mol_227	-8.787
mol_228	-5.506
mol_229	-6.586
mol_230	-9.763
mol_231	-7.627
mol_232R	-6.942
mol_232S	-8.153
mol_233	-7.418
mol_234	-6.174
mol_235	-7.571
mol_236	-5.973
mol_237R	-6.935
mol_237S	-7.113
mol_238R	-6.498
mol_238S	-7.479
mol_239	-4.965
mol_240	-4.706
PYZ_001	-5.595
PYZ_002	-6.426
PYZ_003	-6.713
PYZ_004	-6.013
PYZ_005	-7.14

PYZ_006	-6.885
PYZ_007	-6.577
PYZ_008	-6.352
PYZ_009	-6.755
PYZ_010	-4.037
PYZ_011	-8.049
PYZ_012	-7.044
PYZ_013	-7.361
PYZ_014	-5.838
PYZ_015	-6.273
PYZ_016	-5.385
PYZ_017	-6.623
PYZ_018	-8.357
PYZ_019	-7.327
PYZ_020	-6.81
PYZ_021	-7.021
PYZ_022	-6.877
PYZ_023	-7.222
PYZ_024	-5.62
PYZ_025	-6.32
PYZ_026	-7.079
PYZ_027	-6.115
PYZ_028	-7.518
PYZ_029	-6.428
PYZ_030	-5.716
PYZ_031	-6.281
PYZ_032	-7.241
PYZ_033	-6.537
PYZ_034	-5.119
PYZ_035	-5.326
PYZ_036	-6.558
PYZ_037	-6.949
PYZ_038	-5.879
PYZ_039	-4.677
PYZ_040	-5.78

Anexo IV

Propiedades fisicoquímicas asociadas al carácter drug-like de los compuestos analizados, calculadas mediante la plataforma ADMETlab 3.0. La tabla incluye el identificador de cada molécula, peso molecular (MW), lipofilicidad (logP), número de donadores (HBD) y aceptores (HBA) de enlace de hidrógeno, número de violaciones a la regla de Lipinski, superficie polar topológica (TPSA) y número de enlaces rotables (nRot).

Asimismo, se indica el cumplimiento del criterio de Veber, el cual considera valores de $TPSA \leq 140 \text{ \AA}^2$ y ≤ 10 enlaces rotables, parámetros asociados a una adecuada biodisponibilidad oral.

Molécula	MM	logP	HBD	HBA	Violaciones de Lipinsky	TPSA	nRot	Veber
mol_001	377.11	5.462588	0	4	0	60.21	4	sí
mol_002	416.13	5.400002	2	6	0	94.94	3	sí
mol_003S	467.16	5.963416	1	5	0	67.53	3	sí
mol_003R	467.16	5.963416	1	5	0	67.53	3	sí
mol_004R	391.13	5.121396	1	5	0	67.53	3	sí
mol_004S	391.13	5.121396	1	5	0	67.53	3	sí
mol_005	532.16	6.696017	2	8	1	109.63	5	sí
mol_006	490.14	6.323592	0	6	0	73.33	3	sí
mol_007	353.11	5.230761	0	4	0	60.21	4	sí
mol_008	392.13	5.001425	2	6	0	94.94	3	sí
mol_009R	443.16	5.211498	1	5	0	67.53	3	sí
mol_009S	443.16	5.211498	1	5	0	67.53	3	sí
mol_010R	367.13	4.518197	1	5	0	67.53	3	sí
mol_010S	367.13	4.518197	1	5	0	67.53	3	sí
mol_011	508.16	6.177677	2	8	1	109.63	5	sí
mol_012	466.14	5.987644	0	6	0	73.33	3	sí
mol_013	303.09	4.077532	0	4	0	60.21	4	sí
mol_014	342.11	3.941191	2	6	0	94.94	3	sí
mol_015S	393.15	4.28501	1	5	0	67.53	3	sí
mol_015R	393.15	4.28501	1	5	0	67.53	3	sí
mol_016S	317.12	3.615884	1	5	0	67.53	3	sí
mol_016R	317.12	3.615884	1	5	0	67.53	3	sí
mol_017	458.15	5.138977	2	8	0	109.63	5	sí
mol_018	416.13	4.9825	0	6	0	73.33	3	sí
mol_019	303.09	3.814819	0	4	0	60.21	4	sí
mol_020	342.11	3.872639	2	6	0	94.94	3	sí
mol_021S	393.15	4.224708	1	5	0	67.53	3	sí
mol_021R	393.15	4.224708	1	5	0	67.53	3	sí
mol_022S	317.12	3.52723	1	5	0	67.53	3	sí

mol_022R	317.12	3.52723	1	5	0	67.53	3	sí
mol_023	458.15	5.163808	2	8	0	109.63	5	sí
mol_024	416.13	4.792204	0	6	0	73.33	3	sí
mol_025	253.07	3.153394	0	4	0	60.21	4	sí
mol_026	292.1	3.150072	2	6	0	94.94	3	sí
mol_027S	343.13	3.356397	1	5	0	67.53	3	sí
mol_027R	343.13	3.356397	1	5	0	67.53	3	sí
mol_028R	267.1	2.702512	1	5	0	67.53	3	sí
mol_028S	267.1	2.702512	1	5	0	67.53	3	sí
mol_029	408.13	4.332745	2	8	0	109.63	5	sí
mol_030	366.11	4.08337	0	6	0	73.33	3	sí
mol_031	375.16	5.582977	0	2	0	20.31	4	sí
mol_032	414.18	5.40561	2	4	0	55.04	3	sí
mol_033R	465.22	6.166201	1	3	0	27.63	3	sí
mol_033S	465.22	6.166201	1	3	0	27.63	3	sí
mol_034R	389.19	5.421849	1	3	0	27.63	3	sí
mol_034S	389.19	5.421849	1	3	0	27.63	3	sí
mol_035	530.22	6.763238	2	6	1	69.73	5	sí
mol_036	488.2	6.74825	0	4	0	33.43	3	sí
mol_037	351.16	5.207333	0	2	0	20.31	4	sí
mol_038	390.18	4.948923	2	4	0	55.04	3	sí
mol_039S	441.22	5.406712	1	3	0	27.63	3	sí
mol_039R	441.22	5.406712	1	3	0	27.63	3	sí
mol_040R	365.19	4.695447	1	3	0	27.63	3	sí
mol_040S	365.19	4.695447	1	3	0	27.63	3	sí
mol_041	506.22	6.547128	2	6	1	69.73	5	sí
mol_042	464.2	6.351662	0	4	0	33.43	3	sí
mol_043	301.15	4.238604	0	2	0	20.31	4	sí
mol_044	340.17	3.859383	2	4	0	55.04	3	sí
mol_045S	391.2	4.602653	1	3	0	27.63	3	sí
mol_045R	391.2	4.602653	1	3	0	27.63	3	sí
mol_046R	315.17	3.744005	1	3	0	27.63	3	sí
mol_046S	315.17	3.744005	1	3	0	27.63	3	sí
mol_047	456.21	5.542509	2	6	0	69.73	5	sí
mol_048	414.18	5.101959	0	4	0	33.43	3	sí
mol_049	301.15	4.153122	0	2	0	20.31	4	sí
mol_050	340.17	3.843454	2	4	0	55.04	3	sí
mol_051R	391.2	4.360615	1	3	0	27.63	3	sí
mol_051S	391.2	4.360615	1	3	0	27.63	3	sí
mol_052R	315.17	3.727169	1	3	0	27.63	3	sí
mol_052S	315.17	3.727169	1	3	0	27.63	3	sí

mol_053	456.21	5.469561	2	6	0	69.73	5	sí
mol_054	414.18	5.112394	0	4	0	33.43	3	sí
mol_055	251.13	3.315058	0	2	0	20.31	4	sí
mol_056	290.15	3.209343	2	4	0	55.04	3	sí
mol_057S	341.19	3.414975	1	3	0	27.63	3	sí
mol_057R	341.19	3.414975	1	3	0	27.63	3	sí
mol_058R	265.16	2.746006	1	3	0	27.63	3	sí
mol_058S	265.16	2.746006	1	3	0	27.63	3	sí
mol_059	406.19	4.783384	2	6	0	69.73	5	sí
mol_060	364.17	4.299092	0	4	0	33.43	3	sí
mol_061	378.13	4.998035	1	3	0	46.53	4	sí
mol_062	417.15	4.703287	3	5	0	81.26	3	sí
mol_063R	468.18	5.455076	2	4	0	53.85	3	sí
mol_063S	468.18	5.455076	2	4	0	53.85	3	sí
mol_064R	392.15	4.635906	2	4	0	53.85	3	sí
mol_064S	392.15	4.635906	2	4	0	53.85	3	sí
mol_065	533.19	5.952647	3	7	1	95.95	5	sí
mol_066	491.16	5.802257	1	5	0	59.65	3	sí
mol_067	354.13	4.616485	1	3	0	46.53	4	sí
mol_068	393.15	4.477564	3	5	0	81.26	3	sí
mol_069R	444.18	4.7069	2	4	0	53.85	3	sí
mol_069S	444.18	4.7069	2	4	0	53.85	3	sí
mol_070R	368.15	4.039364	2	4	0	53.85	3	sí
mol_070S	368.15	4.039364	2	4	0	53.85	3	sí
mol_071	509.19	5.692689	3	7	1	95.95	5	sí
mol_072	467.16	5.345022	1	5	0	59.65	3	sí
mol_073	304.11	3.65521	1	3	0	46.53	4	sí
mol_074	343.13	3.446652	3	5	0	81.26	3	sí
mol_075R	394.17	3.684317	2	4	0	53.85	3	sí
mol_075S	394.17	3.684317	2	4	0	53.85	3	sí
mol_076R	318.14	3.174563	2	4	0	53.85	3	sí
mol_076S	318.14	3.174563	2	4	0	53.85	3	sí
mol_077	459.17	4.7642	3	7	0	95.95	5	sí
mol_078	417.15	4.262027	1	5	0	59.65	3	sí
mol_079	304.11	3.545606	1	3	0	46.53	4	sí
mol_080	343.13	3.374183	3	5	0	81.26	3	sí
mol_081S	394.17	3.9034	2	4	0	53.85	3	sí
mol_081R	394.17	3.9034	2	4	0	53.85	3	sí
mol_082R	318.14	3.098436	2	4	0	53.85	3	sí
mol_082S	318.14	3.098436	2	4	0	53.85	3	sí
mol_083	459.17	4.734218	3	7	0	95.95	5	sí

mol_084	417.15	4.193557	1	5	0	59.65	3	sí
mol_085	254.09	3.01128	1	3	0	46.53	4	sí
mol_086	293.12	2.68821	3	5	0	81.26	3	sí
mol_087S	344.15	2.720231	2	4	0	53.85	3	sí
mol_087R	344.15	2.720231	2	4	0	53.85	3	sí
mol_088R	268.12	2.158639	2	4	0	53.85	3	sí
mol_088S	268.12	2.158639	2	4	0	53.85	3	sí
mol_089	409.15	3.924618	3	7	0	95.95	5	sí
mol_090	367.13	3.575188	1	5	0	59.65	3	sí
mol_091	358.14	6.473861	0	1	0	17.07	4	sí
mol_092	397.16	6.219105	2	3	0	51.8	3	sí
mol_093S	448.19	5.999181	1	2	0	24.39	3	sí
mol_093R	448.19	5.999181	1	2	0	24.39	3	sí
mol_094S	372.16	5.568888	1	2	0	24.39	3	sí
mol_094R	372.16	5.568888	1	2	0	24.39	3	sí
mol_095	513.2	7.385144	2	5	1	66.49	5	sí
mol_096	471.17	7.464186	0	3	0	30.19	3	sí
mol_097	334.14	6.20714	0	1	0	17.07	4	sí
mol_098	373.16	5.676127	2	3	0	51.8	3	sí
mol_099S	424.19	5.109571	1	2	0	24.39	3	sí
mol_099R	424.19	5.109571	1	2	0	24.39	3	sí
mol_100R	348.16	4.837844	1	2	0	24.39	3	sí
mol_100S	348.16	4.837844	1	2	0	24.39	3	sí
mol_101	489.2	6.940699	2	5	0	66.49	5	sí
mol_102	447.17	7.056642	0	3	0	30.19	3	sí
mol_103	284.12	5.107	0	1	0	17.07	4	sí
mol_104	323.14	4.627462	2	3	0	51.8	3	sí
mol_105S	374.18	4.366065	1	2	0	24.39	3	sí
mol_105R	374.18	4.366065	1	2	0	24.39	3	sí
mol_106R	298.15	3.918581	1	2	0	24.39	3	sí
mol_106S	298.15	3.918581	1	2	0	24.39	3	sí
mol_107	439.18	5.869715	2	5	0	66.49	5	sí
mol_108	397.16	5.832225	0	3	0	30.19	3	sí
mol_109	284.12	5.060595	0	1	0	17.07	4	sí
mol_110	323.14	4.59717	2	3	0	51.8	3	sí
mol_111R	374.18	4.286911	1	2	0	24.39	3	sí
mol_111S	374.18	4.286911	1	2	0	24.39	3	sí
mol_112S	298.15	3.832745	1	2	0	24.39	3	sí
mol_112R	298.15	3.832745	1	2	0	24.39	3	sí
mol_113	439.18	5.819748	2	5	0	66.49	5	sí
mol_114	397.16	5.86118	0	3	0	30.19	3	sí

mol_115	234.1	4.122652	0	1	0	17.07	4	sí
mol_116	273.13	3.728416	2	3	0	51.8	3	sí
mol_117R	324.16	3.499912	1	2	0	24.39	3	sí
mol_117S	324.16	3.499912	1	2	0	24.39	3	sí
mol_118R	248.13	2.969092	1	2	0	24.39	3	sí
mol_118S	248.13	2.969092	1	2	0	24.39	3	sí
mol_119	389.16	5.11803	2	5	0	66.49	5	sí
mol_120	347.14	4.98078	0	3	0	30.19	3	sí
mol_121	350.13	4.526994	1	2	0	37.3	3	sí
mol_122	387.14	5.132399	3	4	0	72.03	2	sí
mol_123R	438.17	5.689491	2	3	0	44.62	2	sí
mol_123S	438.17	5.689491	2	3	0	44.62	2	sí
mol_124R	362.14	4.894479	2	3	0	44.62	2	sí
mol_124S	362.14	4.894479	2	3	0	44.62	2	sí
mol_125	503.17	6.346548	3	6	1	86.72	4	sí
mol_126	461.15	6.138387	1	4	0	50.42	2	sí
mol_127	324.12	5.30913	1	2	0	37.3	3	sí
mol_128	363.14	4.596734	3	4	0	72.03	2	sí
mol_129R	414.17	4.912592	2	3	0	44.62	2	sí
mol_129S	414.17	4.912592	2	3	0	44.62	2	sí
mol_130S	338.14	4.286872	2	3	0	44.62	2	sí
mol_130R	338.14	4.286872	2	3	0	44.62	2	sí
mol_131	479.17	5.871303	3	6	0	86.72	4	sí
mol_132	437.15	5.589455	1	4	0	50.42	2	sí
mol_133	274.1	4.302758	1	2	0	37.3	3	sí
mol_134	313.12	3.608872	3	4	0	72.03	2	sí
mol_135R	364.16	4.06101	2	3	0	44.62	2	sí
mol_135S	364.16	4.06101	2	3	0	44.62	2	sí
mol_136S	288.13	3.421107	2	3	0	44.62	2	sí
mol_136R	288.13	3.421107	2	3	0	44.62	2	sí
mol_137	429.16	4.932319	3	6	0	86.72	4	sí
mol_138	387.14	4.524839	1	4	0	50.42	2	sí
mol_139	274.1	4.147111	1	2	0	37.3	3	sí
mol_140	313.12	3.662756	3	4	0	72.03	2	sí
mol_141R	364.16	4.059021	2	3	0	44.62	2	sí
mol_141S	364.16	4.059021	2	3	0	44.62	2	sí
mol_142R	288.13	3.443712	2	3	0	44.62	2	sí
mol_142S	288.13	3.443712	2	3	0	44.62	2	sí
mol_143	429.16	4.853338	3	6	0	86.72	4	sí
mol_144	387.14	4.623509	1	4	0	50.42	2	sí
mol_145	224.08	3.377909	1	2	0	37.3	3	sí

mol_146	263.11	2.798826	3	4	0	72.03	2	sí
mol_147S	314.14	3.043849	2	3	0	44.62	2	sí
mol_147R	314.14	3.043849	2	3	0	44.62	2	sí
mol_148S	238.11	2.485656	2	3	0	44.62	2	sí
mol_148R	238.11	2.485656	2	3	0	44.62	2	sí
mol_149	379.14	3.959332	3	6	0	86.72	4	sí
mol_150	337.12	3.757239	1	4	0	50.42	2	sí
mol_151	374.17	6.699996	0	1	0	17.07	4	sí
mol_152	413.19	6.240409	2	3	0	51.8	3	sí
mol_153S	464.23	7.35352	1	2	0	24.39	3	sí
mol_153R	464.23	7.35352	1	2	0	24.39	3	sí
mol_154R	388.19	6.466205	1	2	0	24.39	3	sí
mol_154S	388.19	6.466205	1	2	0	24.39	3	sí
mol_155	529.23	7.673796	2	5	1	66.49	5	sí
mol_156	487.2	7.683191	0	3	0	30.19	3	sí
mol_157	350.17	6.385403	0	1	0	17.07	4	sí
mol_158	389.19	6.152174	2	3	0	51.8	3	sí
mol_159S	440.23	6.602773	1	2	0	24.39	3	sí
mol_159R	440.23	6.602773	1	2	0	24.39	3	sí
mol_160R	364.19	5.874802	1	2	0	24.39	3	sí
mol_160S	364.19	5.874802	1	2	0	24.39	3	sí
mol_161	505.23	7.329709	2	5	1	66.49	5	sí
mol_162	463.2	7.37086	0	3	0	30.19	3	sí
mol_163	300.15	5.352818	0	1	0	17.07	4	sí
mol_164	339.17	5.261924	2	3	0	51.8	3	sí
mol_165S	390.21	5.840871	1	2	0	24.39	3	sí
mol_165R	390.21	5.840871	1	2	0	24.39	3	sí
mol_166R	314.18	4.935832	1	2	0	24.39	3	sí
mol_166S	314.18	4.935832	1	2	0	24.39	3	sí
mol_167	455.21	6.260587	2	5	0	66.49	5	sí
mol_168	413.19	6.178052	0	3	0	30.19	3	sí
mol_169	300.15	5.352061	0	1	0	17.07	4	sí
mol_170	339.17	5.200683	2	3	0	51.8	3	sí
mol_171R	390.21	5.708015	1	2	0	24.39	3	sí
mol_171S	390.21	5.708015	1	2	0	24.39	3	sí
mol_172R	314.18	4.966457	1	2	0	24.39	3	sí
mol_172S	314.18	4.966457	1	2	0	24.39	3	sí
mol_173	455.21	6.450499	2	5	0	66.49	5	sí
mol_174	413.19	6.103869	0	3	0	30.19	3	sí
mol_175	250.14	4.164834	0	1	0	17.07	4	sí
mol_176	289.16	4.433146	2	3	0	51.8	3	sí

mol_177S	340.19	4.882288	1	2	0	24.39	3	sí
mol_177R	340.19	4.882288	1	2	0	24.39	3	sí
mol_178R	264.16	3.89324	1	2	0	24.39	3	sí
mol_178S	264.16	3.89324	1	2	0	24.39	3	sí
mol_179	405.2	5.535758	2	5	0	66.49	5	sí
mol_180	363.17	5.239106	0	3	0	30.19	3	sí
mol_181	347.13	4.792729	2	2	0	43.09	3	sí
mol_182	386.15	4.896097	4	4	0	77.82	2	sí
mol_183R	437.19	5.253049	3	3	0	50.41	2	sí
mol_183S	437.19	5.253049	3	3	0	50.41	2	sí
mol_184S	361.16	4.537302	3	3	0	50.41	2	sí
mol_184R	361.16	4.537302	3	3	0	50.41	2	sí
mol_185	502.19	5.800654	4	6	1	92.51	4	sí
mol_186	460.17	5.813654	2	4	0	56.21	2	sí
mol_187	323.13	4.201265	2	2	0	43.09	3	sí
mol_188	362.15	4.51924	4	4	0	77.82	2	sí
mol_189S	413.19	4.519446	3	3	0	50.41	2	sí
mol_189R	413.19	4.519446	3	3	0	50.41	2	sí
mol_190R	337.16	3.855047	3	3	0	50.41	2	sí
mol_190S	337.16	3.855047	3	3	0	50.41	2	sí
mol_191	478.19	5.564659	4	6	0	92.51	4	sí
mol_192	436.17	5.339275	2	4	0	56.21	2	sí
mol_193	273.12	3.455986	2	2	0	43.09	3	sí
mol_194	312.14	3.390687	4	4	0	77.82	2	sí
mol_195R	363.17	3.568352	3	3	0	50.41	2	sí
mol_195S	363.17	3.568352	3	3	0	50.41	2	sí
mol_196R	287.14	2.81199	3	3	0	50.41	2	sí
mol_196S	287.14	2.81199	3	3	0	50.41	2	sí
mol_197	428.17	4.341362	4	6	0	92.51	4	sí
mol_198	386.15	4.182087	2	4	0	56.21	2	sí
mol_199	273.12	3.225479	2	2	0	43.09	3	sí
mol_200	312.14	3.354753	4	4	0	77.82	2	sí
mol_201R	363.17	3.488526	3	3	0	50.41	2	sí
mol_201S	363.17	3.488526	3	3	0	50.41	2	sí
mol_202R	287.14	2.695588	3	3	0	50.41	2	sí
mol_202S	287.14	2.695588	3	3	0	50.41	2	sí
mol_203	428.17	4.321802	4	6	0	92.51	4	sí
mol_204	386.15	4.209909	2	4	0	56.21	2	sí
mol_205	223.1	2.761135	2	2	0	43.09	3	sí
mol_206	262.12	2.602042	4	4	0	77.82	2	sí
mol_207R	313.16	2.536034	3	3	0	50.41	2	sí

mol_207S	313.16	2.536034	3	3	0	50.41	2	sí
mol_208R	237.13	1.946386	3	3	0	50.41	2	sí
mol_208S	237.13	1.946386	3	3	0	50.41	2	sí
mol_209	378.16	3.587596	4	6	0	92.51	4	sí
mol_210	336.14	3.416914	2	4	0	56.21	2	sí
mol_211	400.11	6.499122	0	1	0	17.07	4	sí
mol_212	439.13	6.245844	2	3	0	51.8	3	sí
mol_213R	490.17	7.096185	1	2	0	24.39	3	sí
mol_213S	490.17	7.096185	1	2	0	24.39	3	sí
mol_214R	414.13	6.191365	1	2	0	24.39	3	sí
mol_214S	414.13	6.191365	1	2	0	24.39	3	sí
mol_215	555.17	7.225262	2	5	1	66.49	5	sí
mol_216	513.15	7.300122	0	3	1	30.19	3	sí
mol_217	376.11	5.98366	0	1	0	17.07	4	sí
mol_218	415.13	5.849574	2	3	0	51.8	3	sí
mol_219S	466.17	6.615369	1	2	0	24.39	3	sí
mol_219R	466.17	6.615369	1	2	0	24.39	3	sí
mol_220S	390.13	5.643922	1	2	0	24.39	3	sí
mol_220R	390.13	5.643922	1	2	0	24.39	3	sí
mol_221	531.17	7.050601	2	5	1	66.49	5	sí
mol_222	489.15	6.868117	0	3	0	30.19	3	sí
mol_223	326.09	5.061863	0	1	0	17.07	4	sí
mol_224	365.11	4.895397	2	3	0	51.8	3	sí
mol_225S	416.15	5.767878	1	2	0	24.39	3	sí
mol_225R	416.15	5.767878	1	2	0	24.39	3	sí
mol_226S	340.12	4.708726	1	2	0	24.39	3	sí
mol_226R	340.12	4.708726	1	2	0	24.39	3	sí
mol_227	481.15	6.237141	2	5	0	66.49	5	sí
mol_228	439.13	5.828275	0	3	0	30.19	3	sí
mol_229	326.09	5.001898	0	1	0	17.07	4	sí
mol_230	365.11	4.924361	2	3	0	51.8	3	sí
mol_231R	416.15	5.688045	1	2	0	24.39	3	sí
mol_231S	416.15	5.688045	1	2	0	24.39	3	sí
mol_232S	340.12	4.768653	1	2	0	24.39	3	sí
mol_232R	340.12	4.768653	1	2	0	24.39	3	sí
mol_233	481.15	6.100554	2	5	0	66.49	5	sí
mol_234	439.13	5.831207	0	3	0	30.19	3	sí
mol_235	276.08	4.222243	0	1	0	17.07	4	sí
mol_236	315.1	4.057943	2	3	0	51.8	3	sí
mol_237S	366.13	4.756235	1	2	0	24.39	3	sí
mol_237R	366.13	4.756235	1	2	0	24.39	3	sí

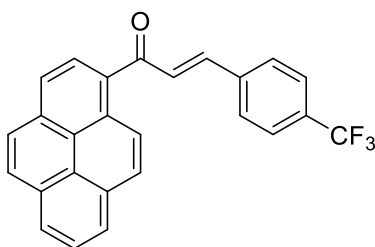
mol_238S	286.13	3.503865	1	2	0	24.39	3	sí
mol_238R	286.13	3.503865	1	2	0	24.39	3	sí
mol_239	431.14	5.685931	2	5	0	66.49	5	sí
mol_240	389.11	5.011809	0	3	0	30.19	3	sí
BZD_001	465.15	6.58419	1	5	0	67.53	3	sí
BZD_002	441.15	5.869276	1	5	0	67.53	3	sí
BZD_003	391.13	5.011221	1	5	0	67.53	3	sí
BZD_004	391.13	5.111449	1	5	0	67.53	3	sí
BZD_005	341.12	4.236793	1	5	0	67.53	3	sí
BZD_006	463.2	6.798177	1	3	0	27.63	3	sí
BZD_007	439.2	6.254962	1	3	0	27.63	3	sí
BZD_008	389.19	5.267936	1	3	0	27.63	3	sí
BZD_009	389.19	5.36421	1	3	0	27.63	3	sí
BZD_010	339.17	4.313008	1	3	0	27.63	3	sí
BZD_011	466.17	5.903633	2	4	0	53.85	3	sí
BZD_012	442.17	5.462046	2	4	0	53.85	3	sí
BZD_013	392.15	4.588891	2	4	0	53.85	3	sí
BZD_014	392.15	4.643611	2	4	0	53.85	3	sí
BZD_015	342.14	3.854398	2	4	0	53.85	3	sí
BZD_016	446.18	7.367328	1	2	0	24.39	3	sí
BZD_017	422.18	6.872385	1	2	0	24.39	3	sí
BZD_018	372.16	5.846875	1	2	0	24.39	3	sí
BZD_019	372.16	5.923435	1	2	0	24.39	3	sí
BZD_020	322.15	4.908738	1	2	0	24.39	3	sí
BZD_021	436.16	6.28774	2	3	0	44.62	2	sí
BZD_022	412.16	5.557749	2	3	0	44.62	2	sí
BZD_023	362.14	4.645437	2	3	0	44.62	2	sí
BZD_024	362.14	4.75769	2	3	0	44.62	2	sí
BZD_025	312.13	3.84363	2	3	0	44.62	2	sí
BZD_026	462.21	7.74346	1	2	0	24.39	3	sí
BZD_027	438.21	7.243425	1	2	0	24.39	3	sí
BZD_028	388.19	6.265563	1	2	0	24.39	3	sí
BZD_029	388.19	6.223028	1	2	0	24.39	3	sí
BZD_030	338.18	5.430071	1	2	0	24.39	3	sí
BZD_031	435.17	5.830454	3	3	0	50.41	2	sí
BZD_032	411.17	5.176316	3	3	0	50.41	2	sí
BZD_033	361.16	4.13532	3	3	0	50.41	2	sí
BZD_034	361.16	4.215167	3	3	0	50.41	2	sí
BZD_035	311.14	3.381032	3	3	0	50.41	2	sí
BZD_037	488.15	7.446261	1	2	0	24.39	3	sí
BZD_036	488.15	7.446261	1	2	0	24.39	3	sí

BZD_038	464.15	6.95833	1	2	0	24.39	3	sí
BZD_039	414.13	6.054144	1	2	0	24.39	3	sí
BZD_040	414.13	6.069486	1	2	0	24.39	3	sí
PYZ_001	389.12	5.830814	1	5	0	71.82	3	sí
PYZ_002	365.12	5.28928	1	5	0	71.82	3	sí
PYZ_003	315.1	4.442112	1	5	0	71.82	3	sí
PYZ_004	315.1	4.53203	1	5	0	71.82	3	sí
PYZ_005	265.09	3.688992	1	5	0	71.82	3	sí
PYZ_006	387.17	6.045176	1	3	0	31.92	3	sí
PYZ_007	363.17	5.801407	1	3	0	31.92	3	sí
PYZ_008	313.16	4.710723	1	3	0	31.92	3	sí
PYZ_009	313.16	4.800116	1	3	0	31.92	3	sí
PYZ_010	263.14	3.975187	1	3	0	31.92	3	sí
PYZ_011	390.14	5.209741	2	4	0	58.14	3	sí
PYZ_012	366.14	4.752519	2	4	0	58.14	3	sí
PYZ_013	316.12	3.901655	2	4	0	58.14	3	sí
PYZ_014	316.12	3.783086	2	4	0	58.14	3	sí
PYZ_015	266.11	3.239478	2	4	0	58.14	3	sí
PYZ_016	370.15	6.742321	1	2	0	28.68	3	sí
PYZ_017	346.15	6.126141	1	2	0	28.68	3	sí
PYZ_018	296.13	5.099466	1	2	0	28.68	3	sí
PYZ_019	296.13	5.191927	1	2	0	28.68	3	sí
PYZ_020	246.12	4.33996	1	2	0	28.68	3	sí
PYZ_021	360.13	5.663718	2	3	0	48.91	2	sí
PYZ_022	336.13	5.017045	2	3	0	48.91	2	sí
PYZ_023	286.11	4.198972	2	3	0	48.91	2	sí
PYZ_024	286.11	4.156327	2	3	0	48.91	2	sí
PYZ_025	236.09	3.462885	2	3	0	48.91	2	sí
PYZ_026	386.18	7.038371	1	2	0	28.68	3	sí
PYZ_027	362.18	6.74713	1	2	0	28.68	3	sí
PYZ_028	312.16	5.810082	1	2	0	28.68	3	sí
PYZ_029	312.16	5.922351	1	2	0	28.68	3	sí
PYZ_030	262.15	5.011027	1	2	0	28.68	3	sí
PYZ_031	359.14	5.334721	3	3	0	54.7	2	sí
PYZ_032	335.14	4.825225	3	3	0	54.7	2	sí
PYZ_033	285.13	3.76802	3	3	0	54.7	2	sí
PYZ_034	285.13	3.833489	3	3	0	54.7	2	sí
PYZ_035	235.11	3.092663	3	3	0	54.7	2	sí
PYZ_036	412.12	6.114848	0	2	0	24.72	3	sí
PYZ_037	388.12	6.406038	1	2	0	28.68	3	sí
PYZ_038	338.1	5.778353	1	2	0	28.68	3	sí

PYZ_039	338.1	5.692806	1	2	0	28.68	3	sí
PYZ_040	288.09	4.842889	1	2	0	28.68	3	sí

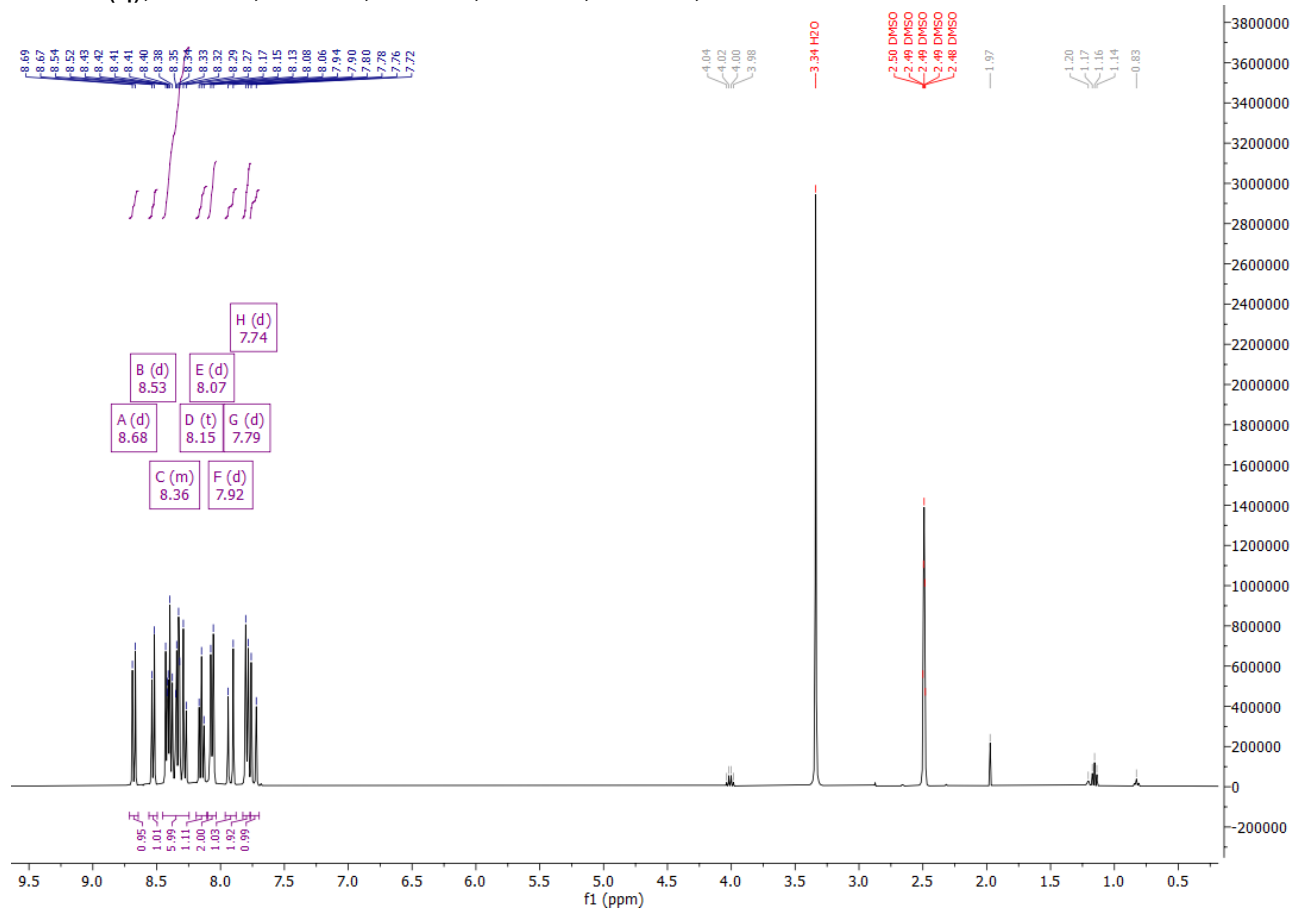
Anexo V

1)Chalcona pireno CF₃

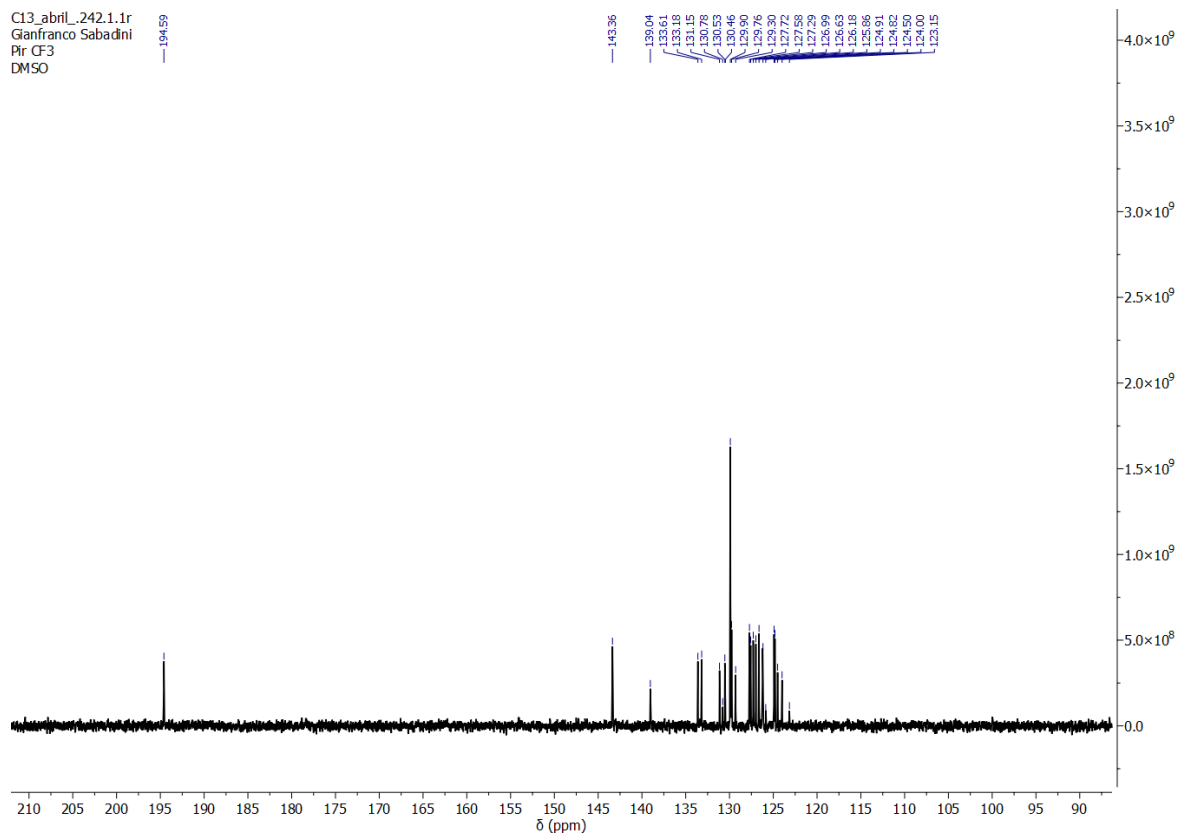


¹H NMR (400 MHz, DMSO) δ 8.68 (d, *J* = 9.3 Hz, 1H), 8.53 (d, *J* = 8.0 Hz, 1H), 8.45 – 8.25 (m, 6H), 8.15 (t, *J* = 7.7 Hz, 1H), 8.07 (d, *J* = 8.1 Hz, 2H), 7.92 (d, *J* = 16.0 Hz, 1H), 7.79 (d, *J* = 8.2 Hz, 2H), 7.74 (d, *J* = 15.9 Hz, 1H). ¹³C NMR (101 MHz, DMSO) δ 194.59, 143.36, 139.04, 133.61, 133.18, 131.15, 130.78, 130.53, 130.46, 129.90, 129.76, 129.30, 127.72, 127.58, 127.29, 126.99, 126.63,

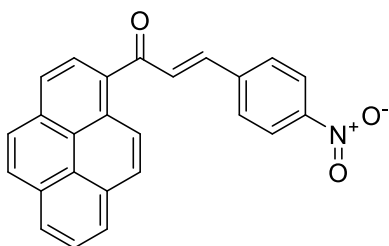
126.18(q), 125.86, 124.91, 124.82, 124.50, 124.00, 123.15.



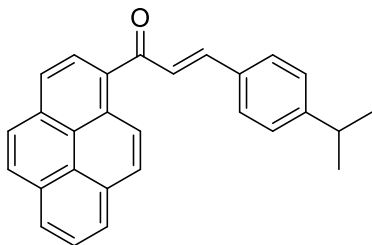
C13_abril_242.1.1r
Gianfranco Sabadini
Pir CF3
DMSO



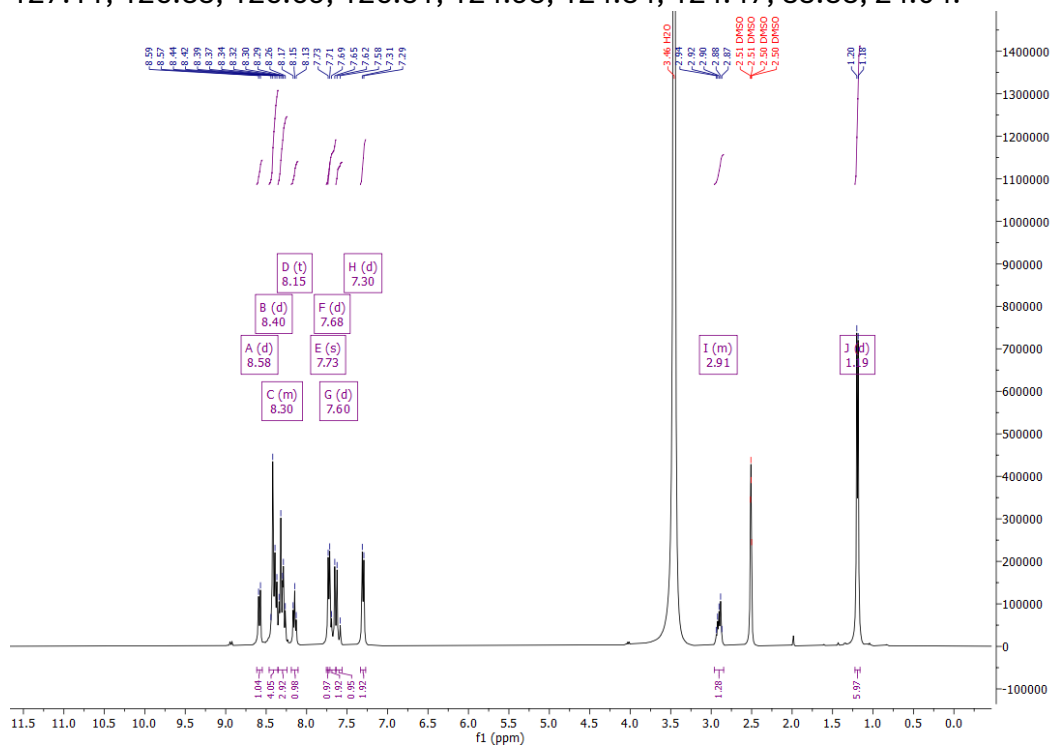
1)Chalcona pireno NO2

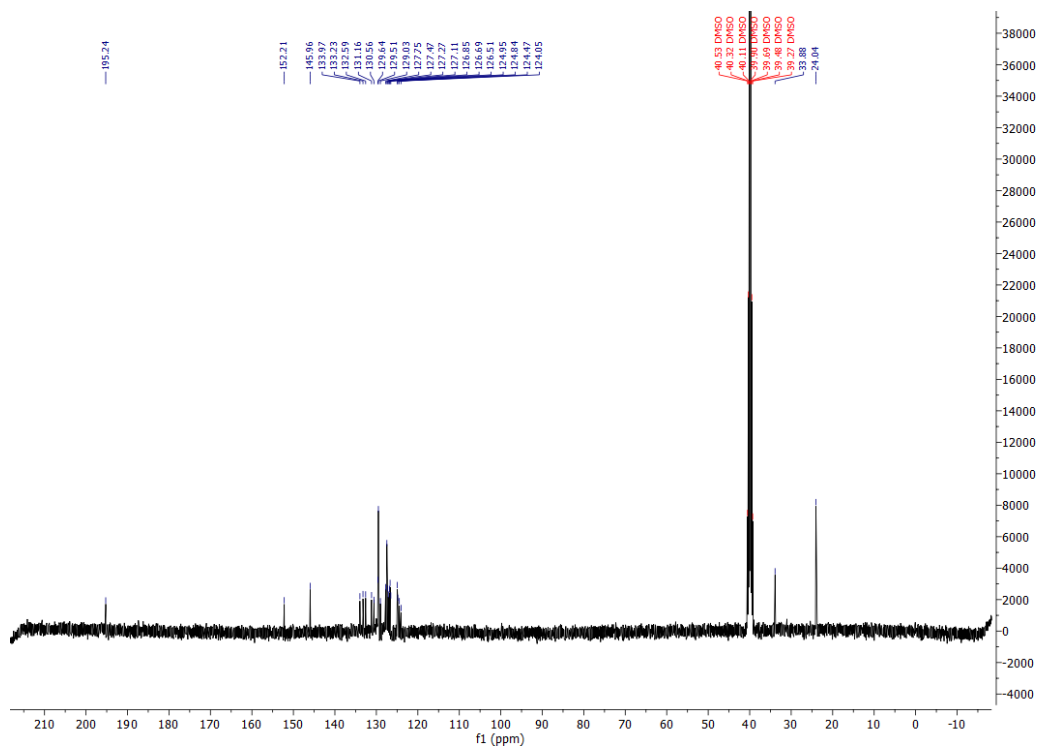


^1H NMR (400 MHz, DMSO) δ 8.68 (d, $J = 9.3$ Hz, 1H), 8.51 (d, $J = 8.0$ Hz, 1H), 8.43 – 8.35 (m, 3H), 8.32 (d, $J = 8.9$ Hz, 2H), 8.29 – 8.21 (m, 2H), 8.20 (d, $J = 2.0$ Hz, 1H), 8.13 (t, $J = 7.6$ Hz, 1H), 8.09 – 8.00 (m, 2H), 7.93 (d, $J = 15.9$ Hz, 1H), 7.72 (d, $J = 16.0$ Hz, 1H). ^{13}C NMR (101 MHz, DMSO) δ 194.01, 148.20, 142.08, 141.07, 133.39, 132.53, 130.76, 130.61, 130.15, 129.94, 129.67, 129.53, 129.02, 127.37, 126.99, 126.74, 126.36, 124.58, 124.42, 124.04.

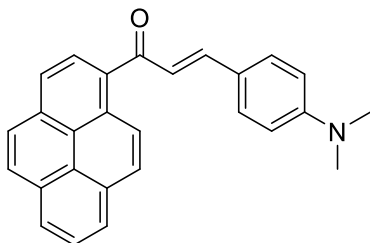


^1H NMR (400 MHz, DMSO) δ 8.58 (d, J = 9.3 Hz, 1H), 8.40 (d, J = 11.1 Hz, 4H), 8.35 – 8.24 (m, 3H), 8.15 (t, J = 7.6 Hz, 1H), 7.73 (s, 1H), 7.68 (d, J = 24.9 Hz, 2H), 7.60 (d, J = 15.9 Hz, 1H), 7.30 (d, J = 7.9 Hz, 2H), 2.91 (dd, J = 14.4, 7.4 Hz, 1H), 1.19 (d, J = 6.9 Hz, 6H). ^{13}C NMR (101 MHz, DMSO) δ 152.21, 145.96, 133.97, 133.23, 132.59, 131.16, 130.56, 129.65, 129.51, 129.03, 127.75, 127.47, 127.27, 127.11, 126.85, 126.69, 126.51, 124.95, 124.84, 124.47, 33.88, 24.04.



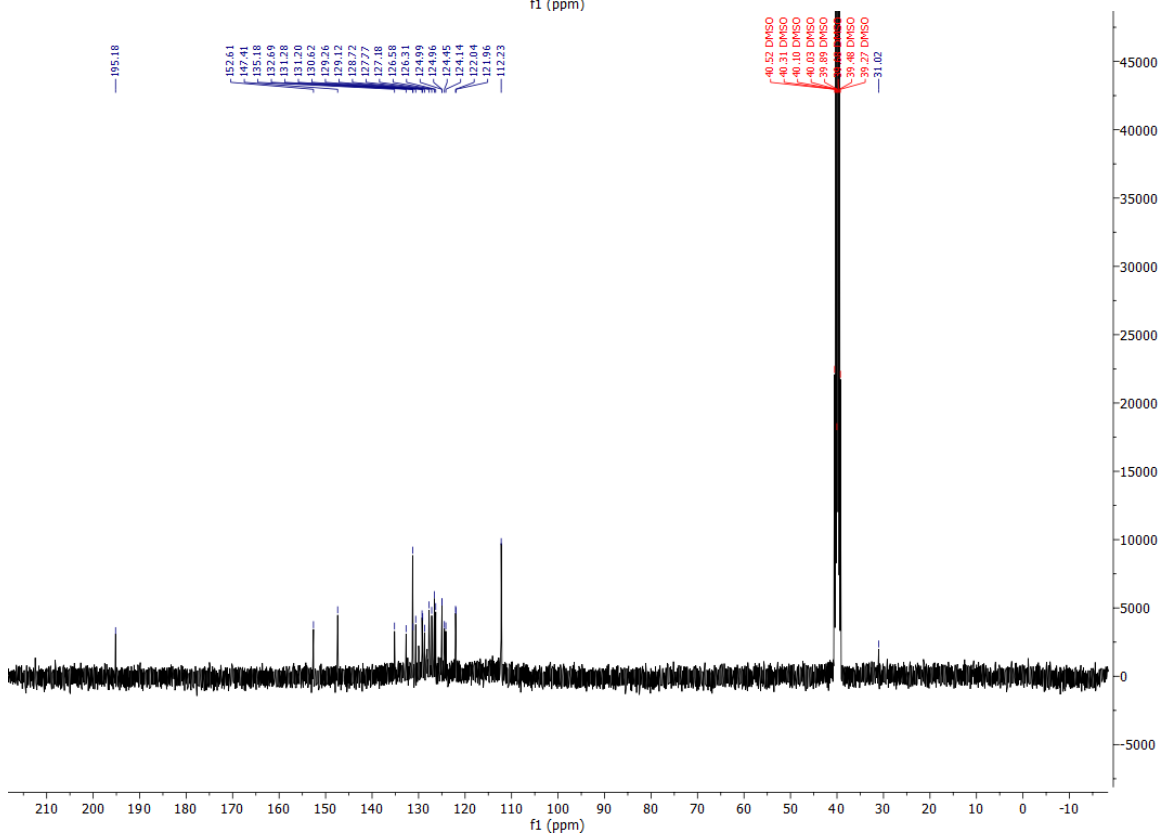
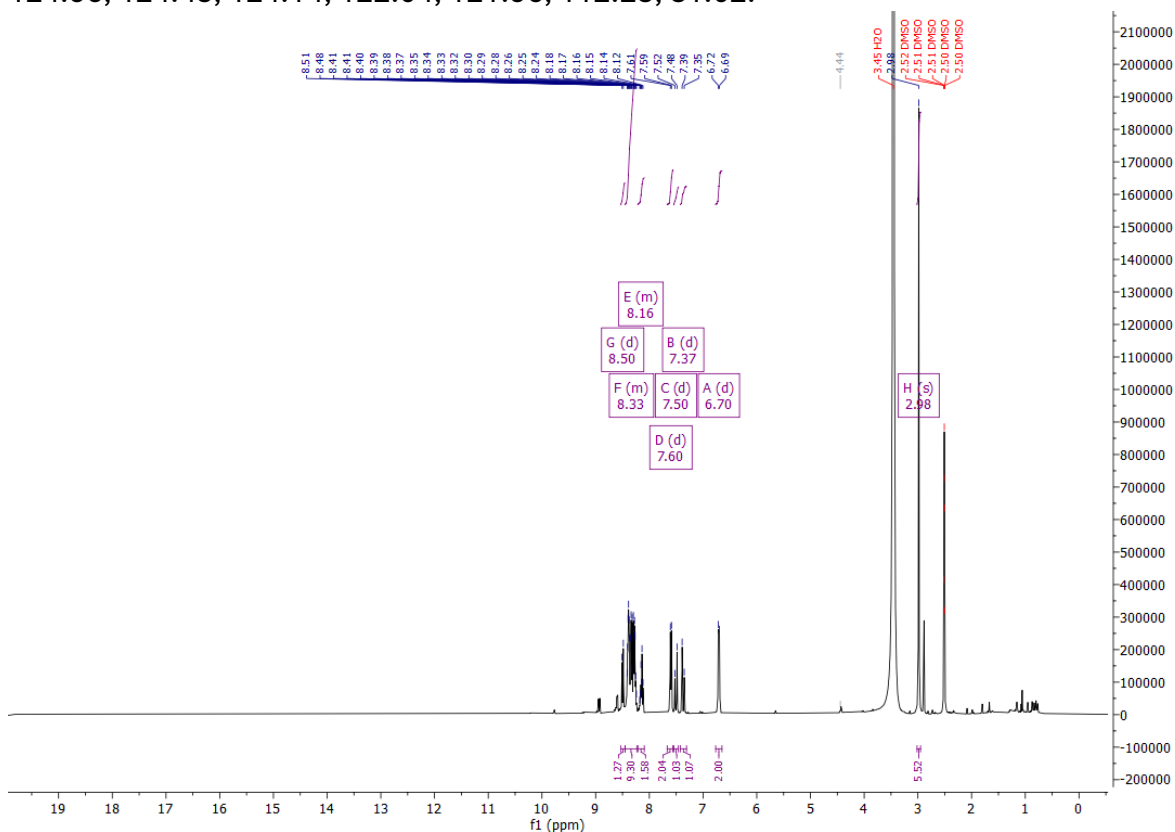


4) Chalcona pireno DMA

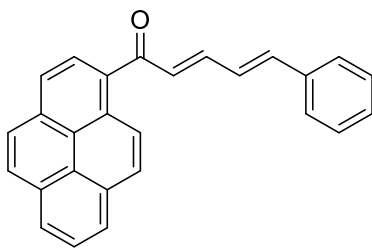


^1H NMR (400 MHz, DMSO (contaminado)) δ 8.50 (d, $J = 9.3$ Hz, 1H), 8.33 (m, $J = 31.4, 14.1, 8.2, 3.9$ Hz, 9H), 8.21 – 8.09 (m, 2H), 7.60 (d, $J = 8.6$ Hz, 2H), 7.50 (d, $J = 15.7$ Hz, 1H), 7.37 (d, $J = 15.8$ Hz, 1H), 6.70 (d, $J = 8.6$ Hz, 2H), 2.98 (s, 6H). ^{13}C NMR (101 MHz, DMSO) δ 195.18, 152.61, 147.41, 135.18, 132.69, 131.28, 131.20, 130.62, 129.26, 129.12, 128.72, 127.77, 127.18, 126.58, 126.31, 124.99,

124.96, 124.45, 124.14, 122.04, 121.96, 112.23, 31.02.

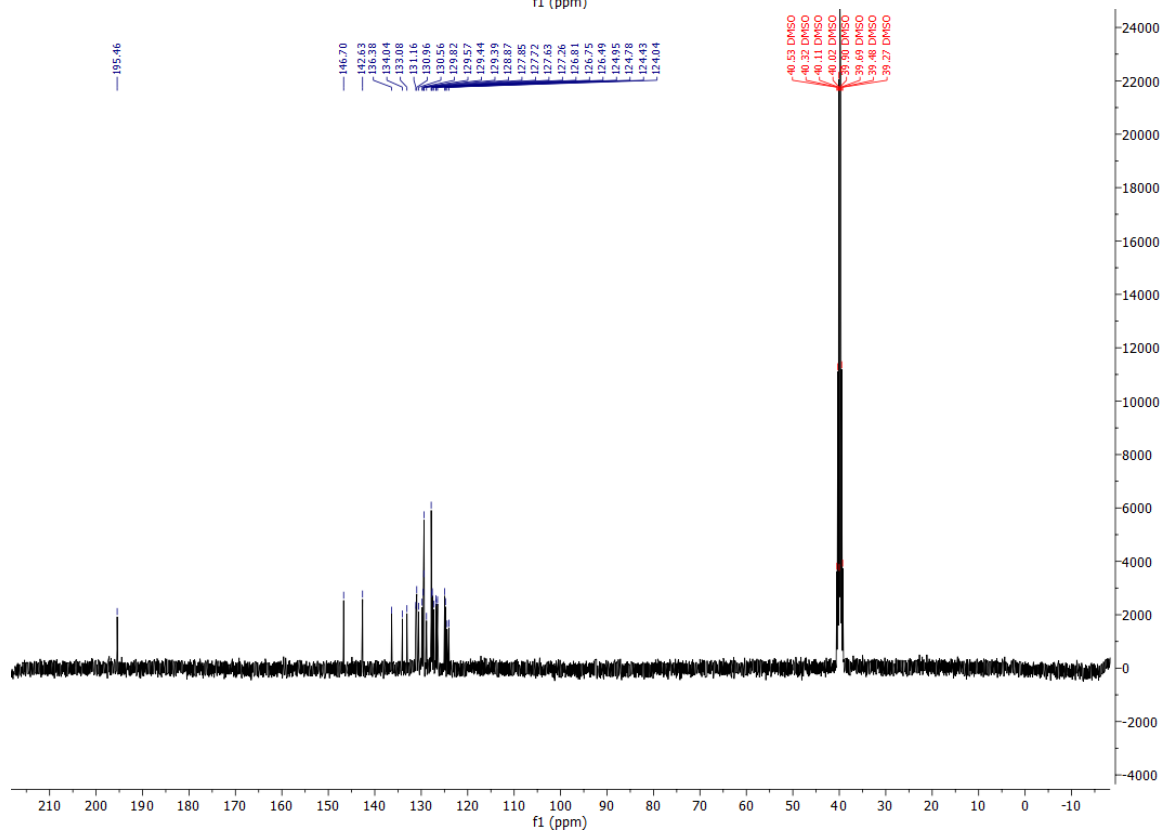
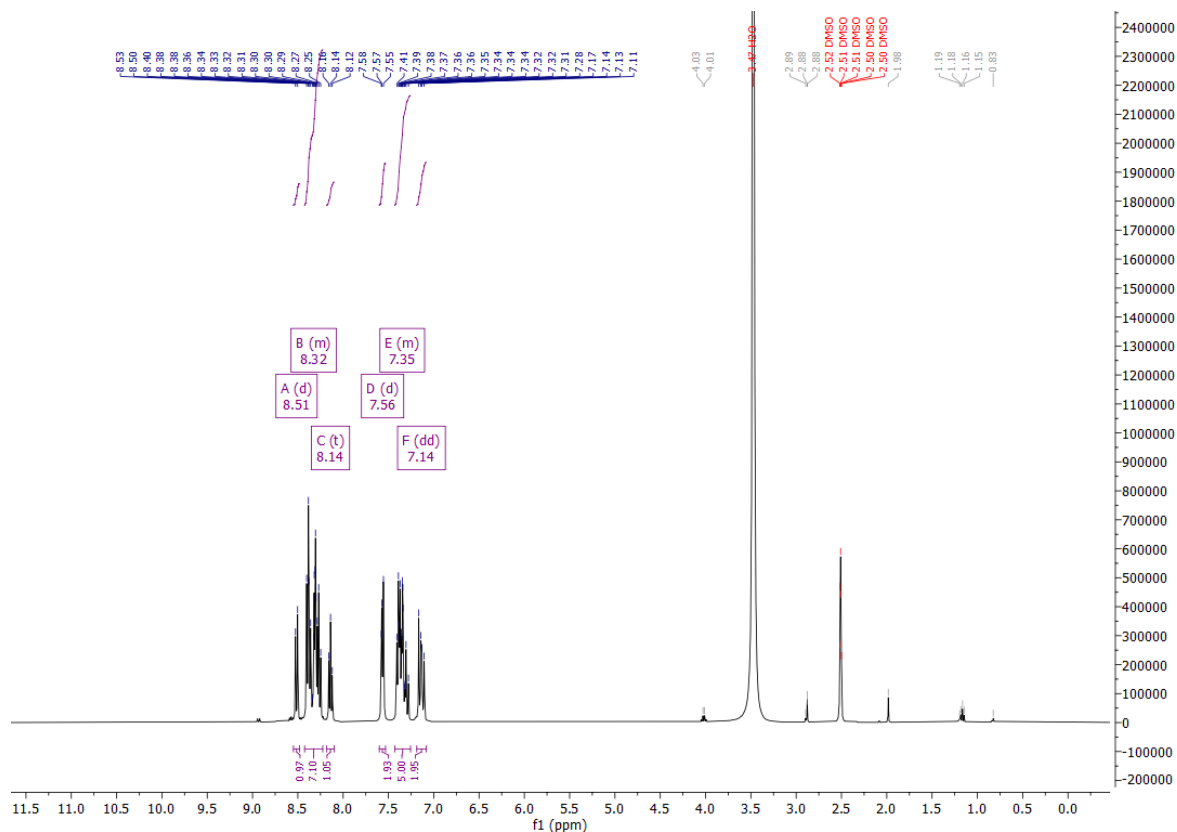


5) Chalcona pireno cinamaldehído

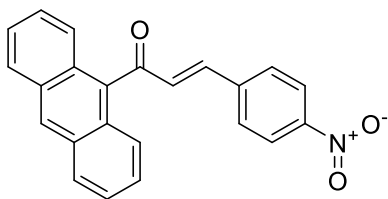


124.43, 124.04.

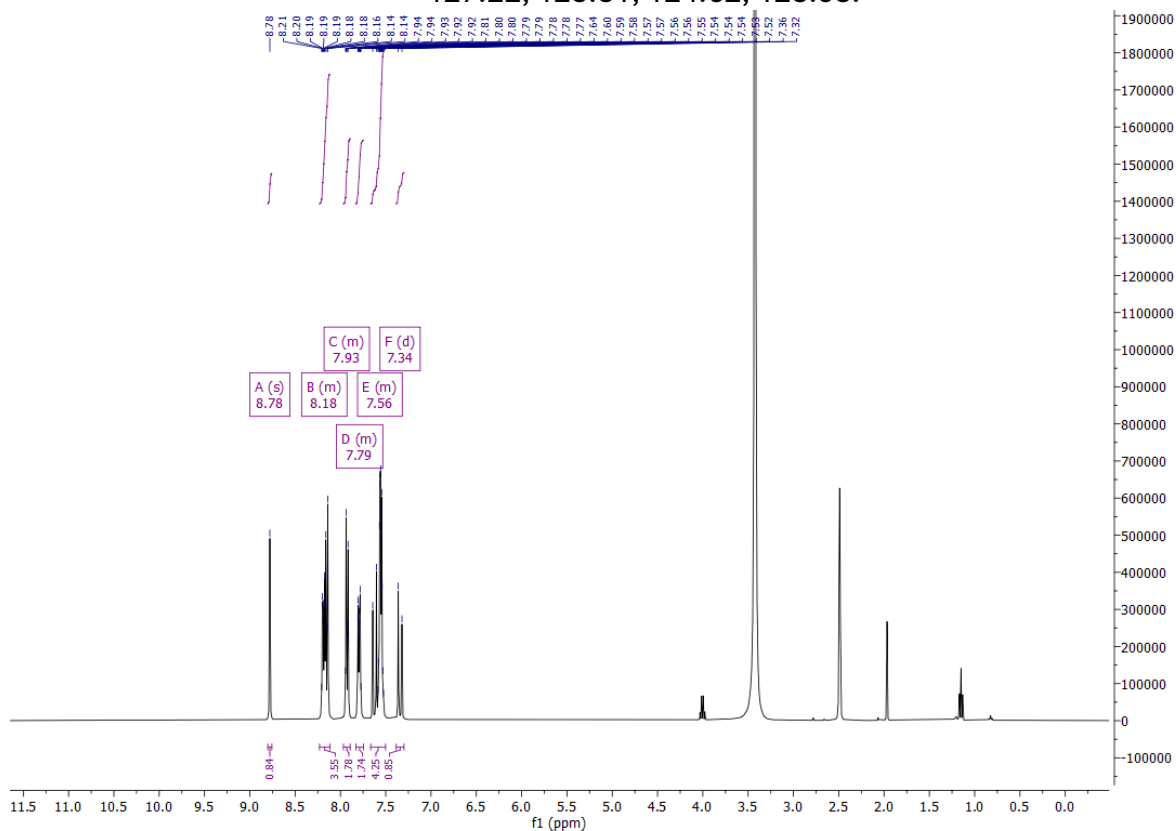
^1H NMR (400 MHz, DMSO) δ 8.51 (d, J = 9.3 Hz, 1H), 8.43 – 8.22 (m, 7H), 8.14 (t, J = 7.6 Hz, 1H), 7.56 (d, J = 7.6 Hz, 2H), 7.43 – 7.26 (m, 5H), 7.14 (dd, J = 14.4, 9.4 Hz, 2H). ^{13}C NMR (101 MHz, DMSO) δ 195.46, 146.70, 142.64, 136.38, 134.04, 133.08, 131.16, 130.97, 130.56, 129.82, 129.57, 129.44, 129.39, 128.87, 127.86, 127.72, 127.63, 127.26, 126.81, 126.75, 126.49, 124.95, 124.78,

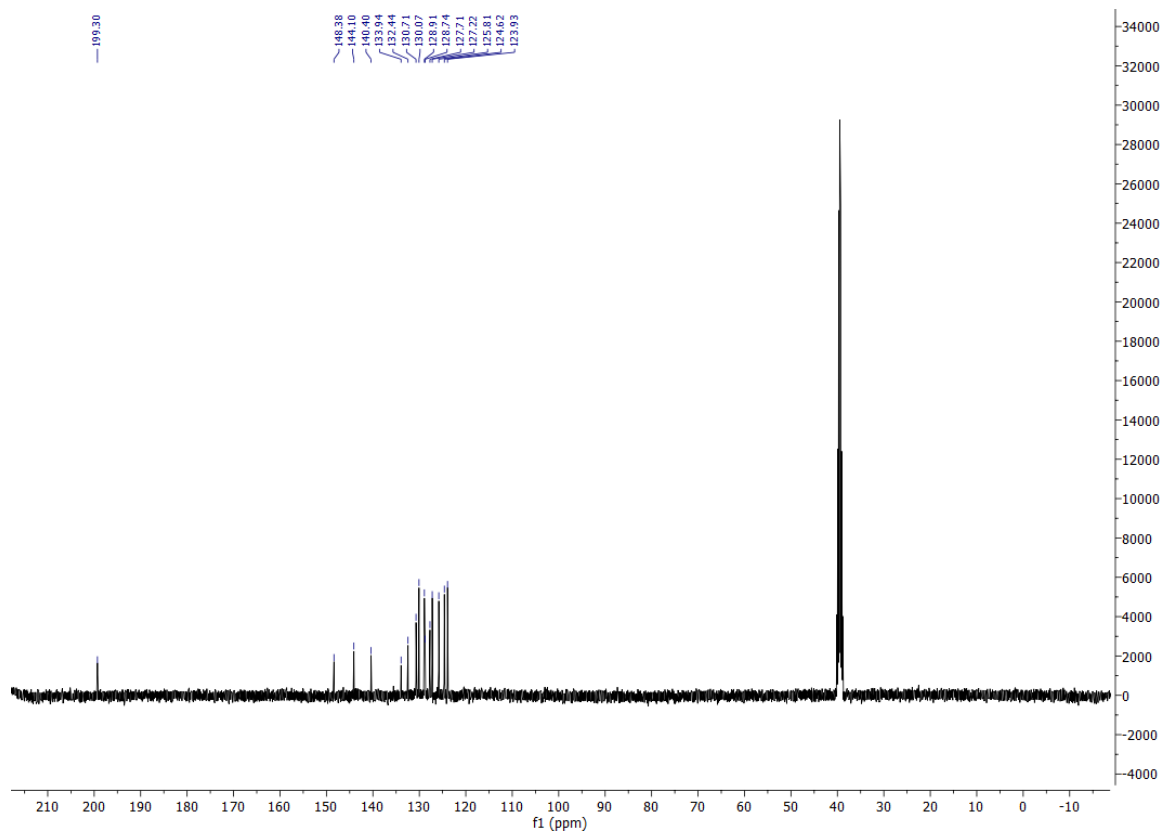


6) Chalcona antraceno NO2

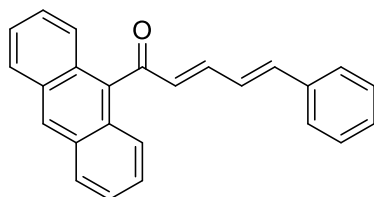


^1H NMR (400 MHz, DMSO) δ 8.78 (s, 1H), 8.23 – 8.12 (m, 4H), 7.97 – 7.89 (m, 2H), 7.83 – 7.75 (m, 2H), 7.67 – 7.50 (m, 4H), 7.34 (d, $J = 16.3$ Hz, 1H). ^{13}C NMR (101 MHz, DMSO) δ 199.30, 148.38, 144.10, 140.40, 133.94, 132.44, 130.71, 130.07, 128.91, 128.74, 127.71, 127.22, 125.81, 124.62, 123.93.

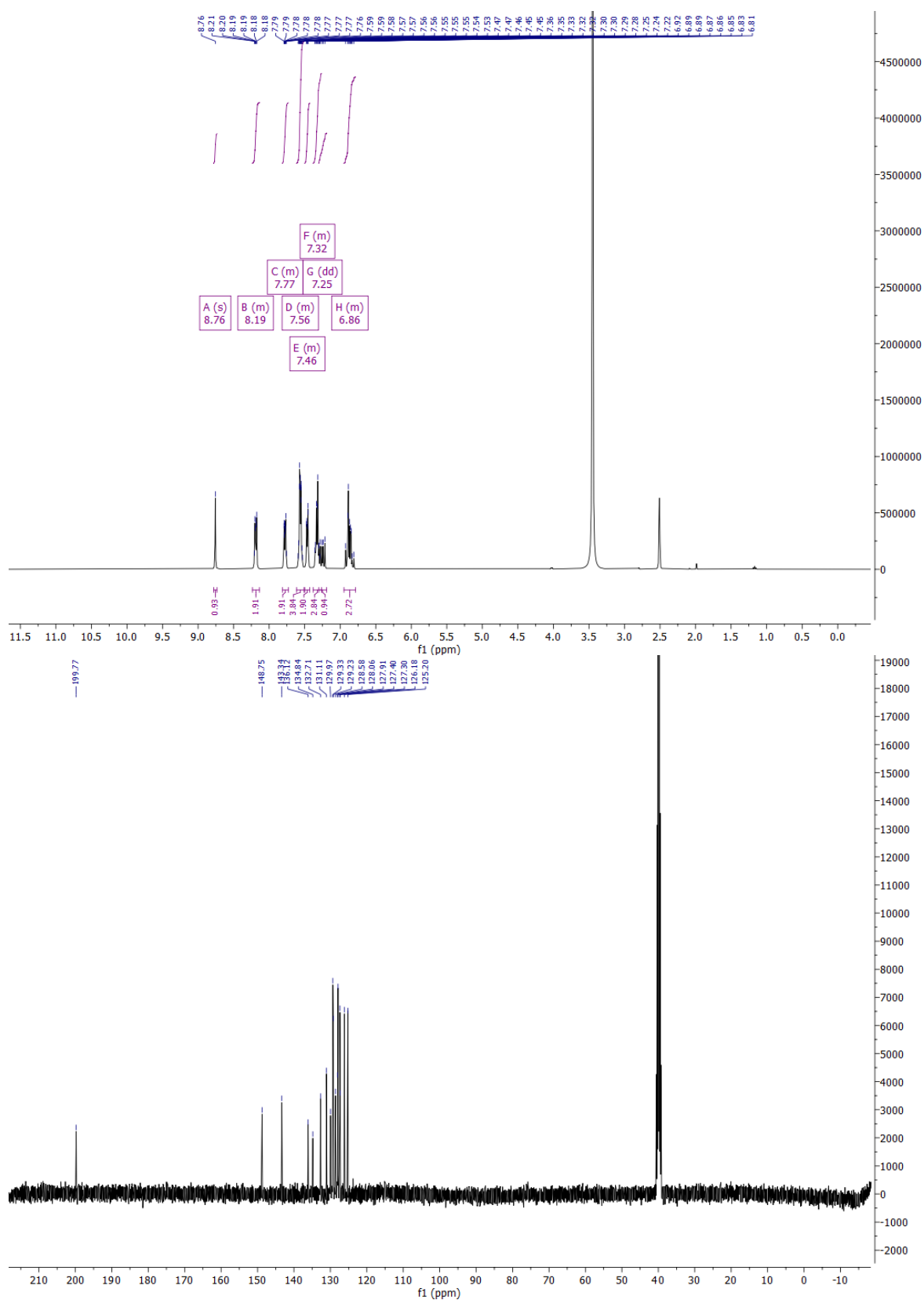




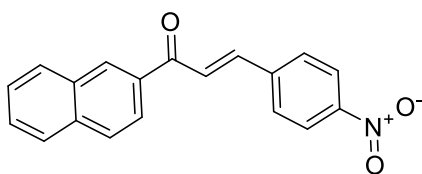
7) Chalcona antraceno cinamaldehído



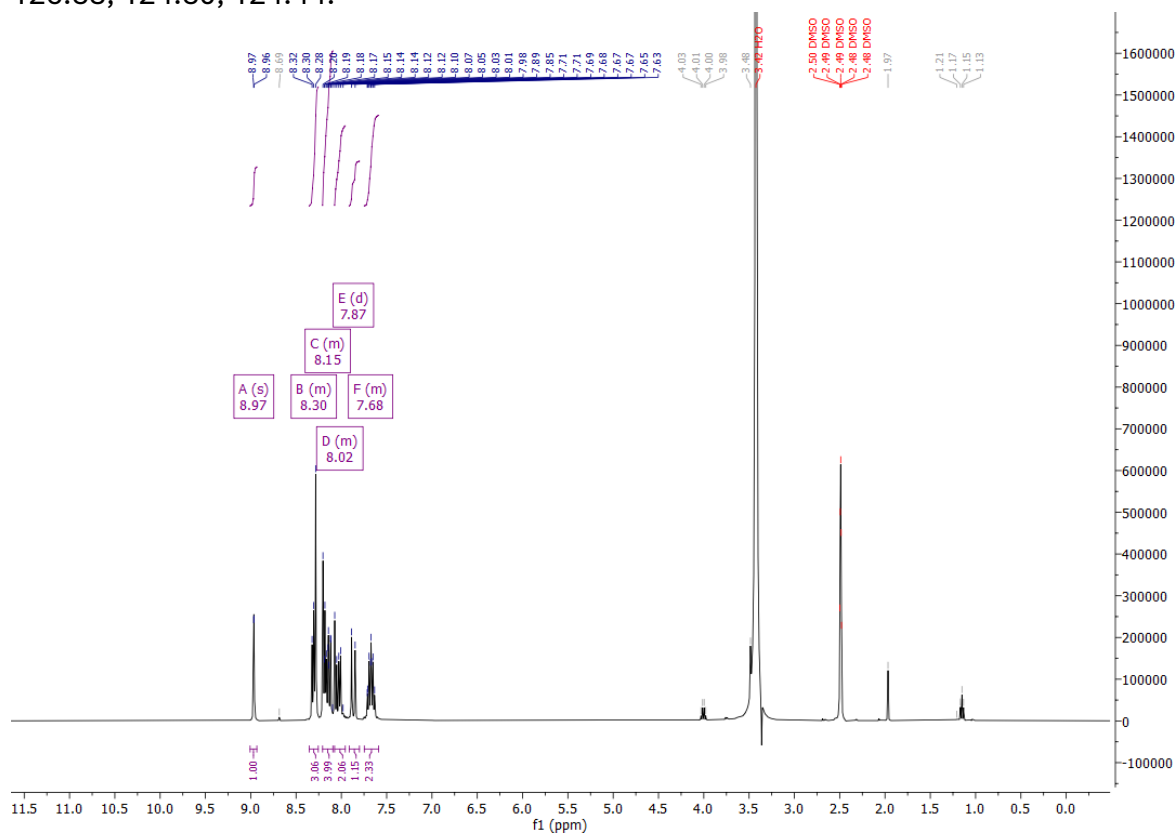
^1H NMR (400 MHz, DMSO) δ 8.76 (s, 1H), 8.24 – 8.14 (m, 2H), 7.81 – 7.73 (m, 2H), 7.61 – 7.51 (m, 4H), 7.50 – 7.43 (m, 2H), 7.38 – 7.26 (m, 3H), 7.25 (dd, J = 15.6, 9.9 Hz, 1H), 6.95 – 6.78 (m, 3H). ^{13}C NMR (101 MHz, DMSO) δ 199.77, 148.75, 143.34, 136.12, 134.84, 132.71, 131.11, 129.97, 129.33, 129.23, 128.58, 128.06, 127.91, 127.40, 127.30, 126.18, 125.20.

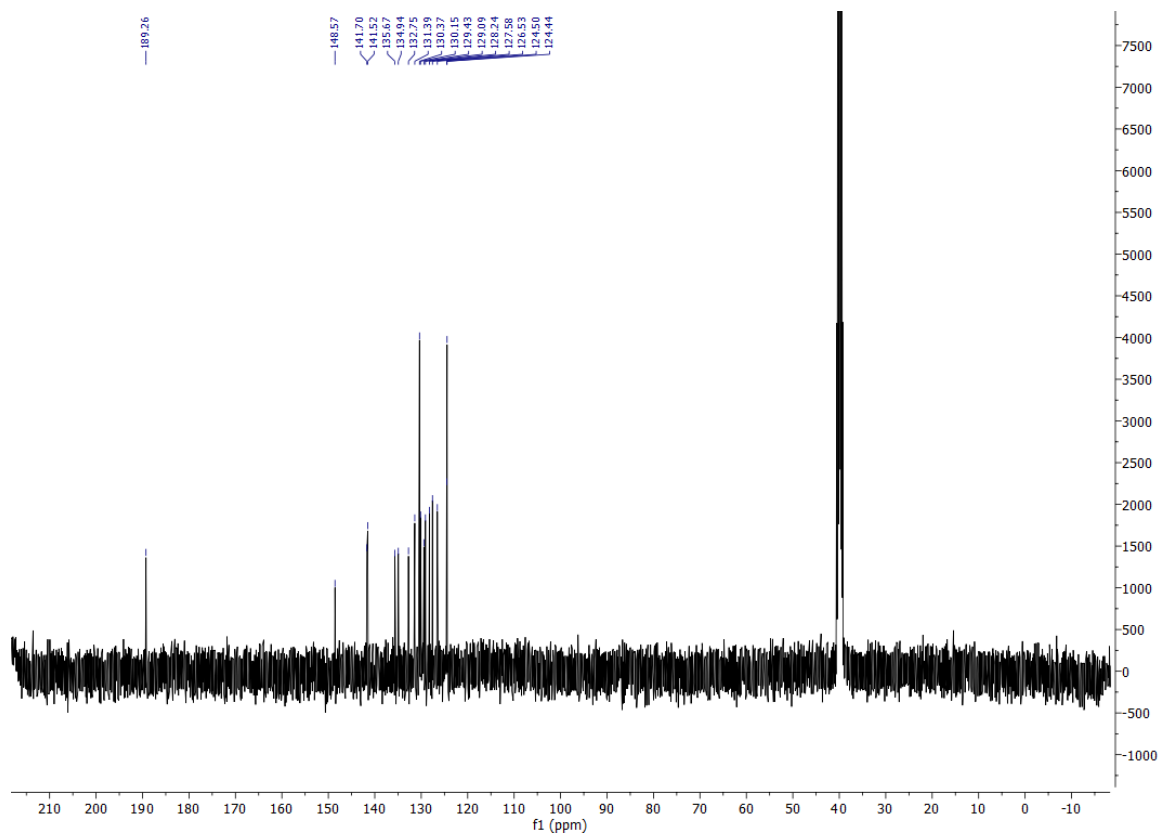


8) Chalcona 2 naftoil NO₂

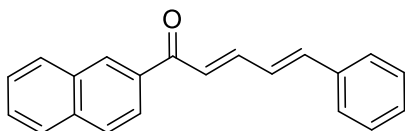


^1H NMR (400 MHz, DMSO) δ 8.97 (s, 1H), 8.35 – 8.25 (m, 3H), 8.21 – 8.09 (m, 4H), 8.07 – 7.96 (m, 2H), 7.87 (d, J = 15.7 Hz, 1H), 7.75 – 7.59 (m, 2H). ^{13}C NMR (101 MHz, DMSO) δ 189.26, 148.57, 141.70, 141.52, 135.67, 134.94, 132.75, 131.39, 130.37, 130.15, 129.43, 129.09, 128.24, 127.58, 126.53, 124.50, 124.44.

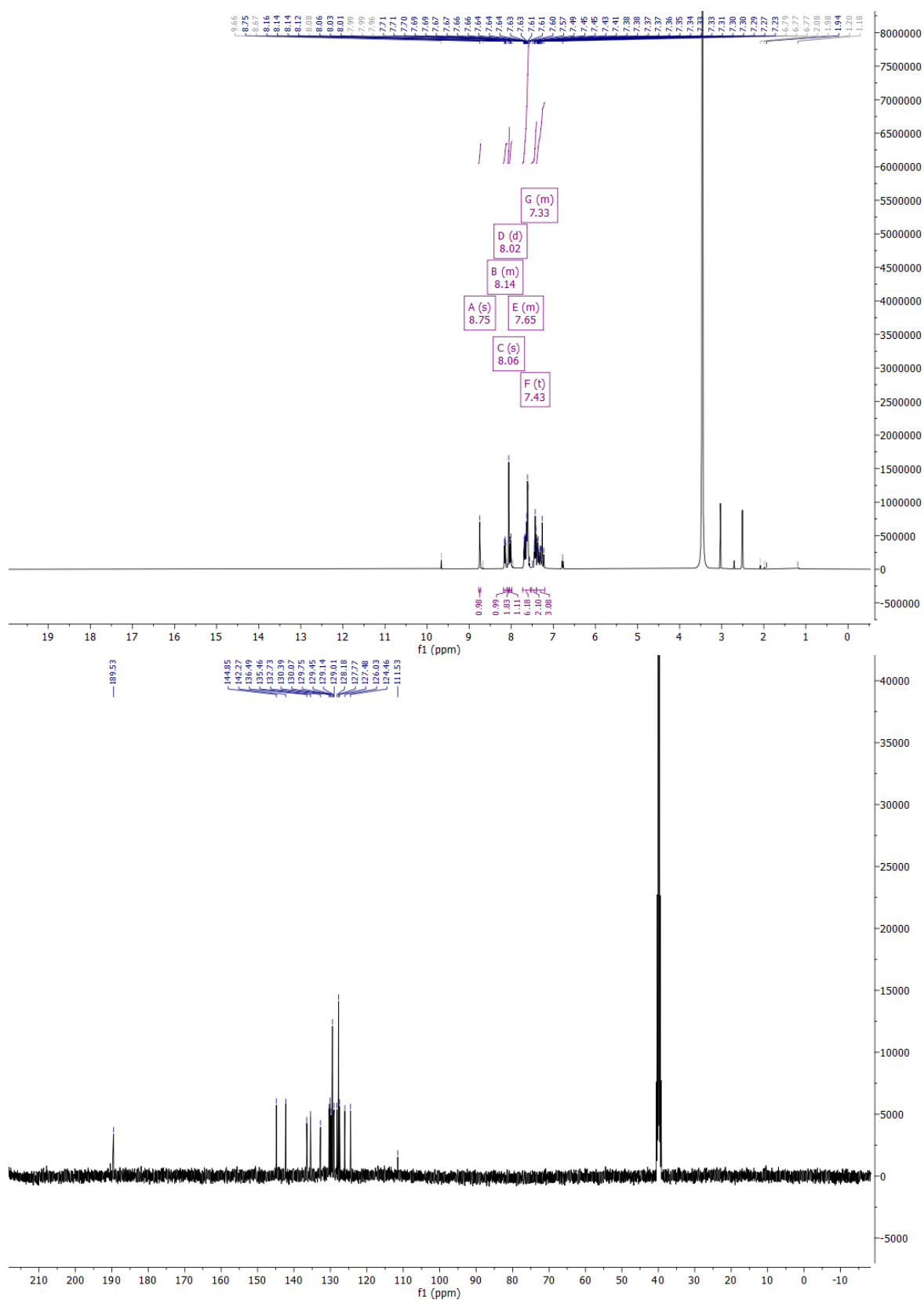




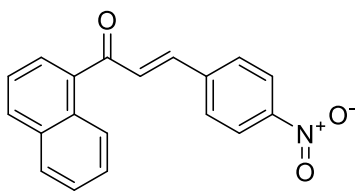
9) Chalcona 2 naftoil cinamaldehydo



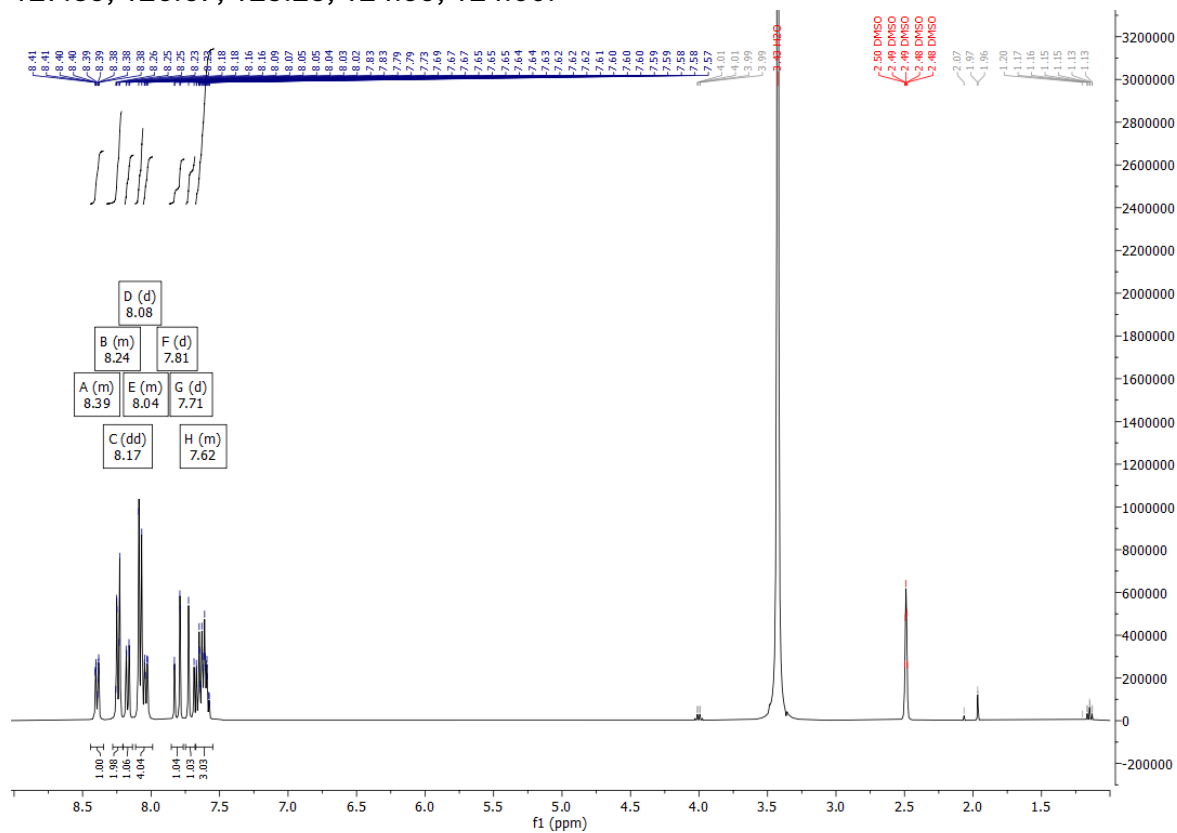
^1H NMR (400 MHz, DMSO) δ 8.75 (s, 1H), 8.18 – 8.10 (m, 1H), 8.06 (s, 2H), 8.02 (d, J = 8.0 Hz, 1H), 7.73 – 7.55 (m, 6H), 7.43 (t, J = 7.3 Hz, 2H), 7.40 – 7.20 (m, 3H). ^{13}C NMR (101 MHz, DMSO) δ 189.53, 144.85, 142.27, 136.49, 135.46, 132.73, 130.39, 130.07, 129.75, 129.45, 129.14, 129.01, 128.18, 127.77, 127.48, 126.03, 124.46, 111.53.

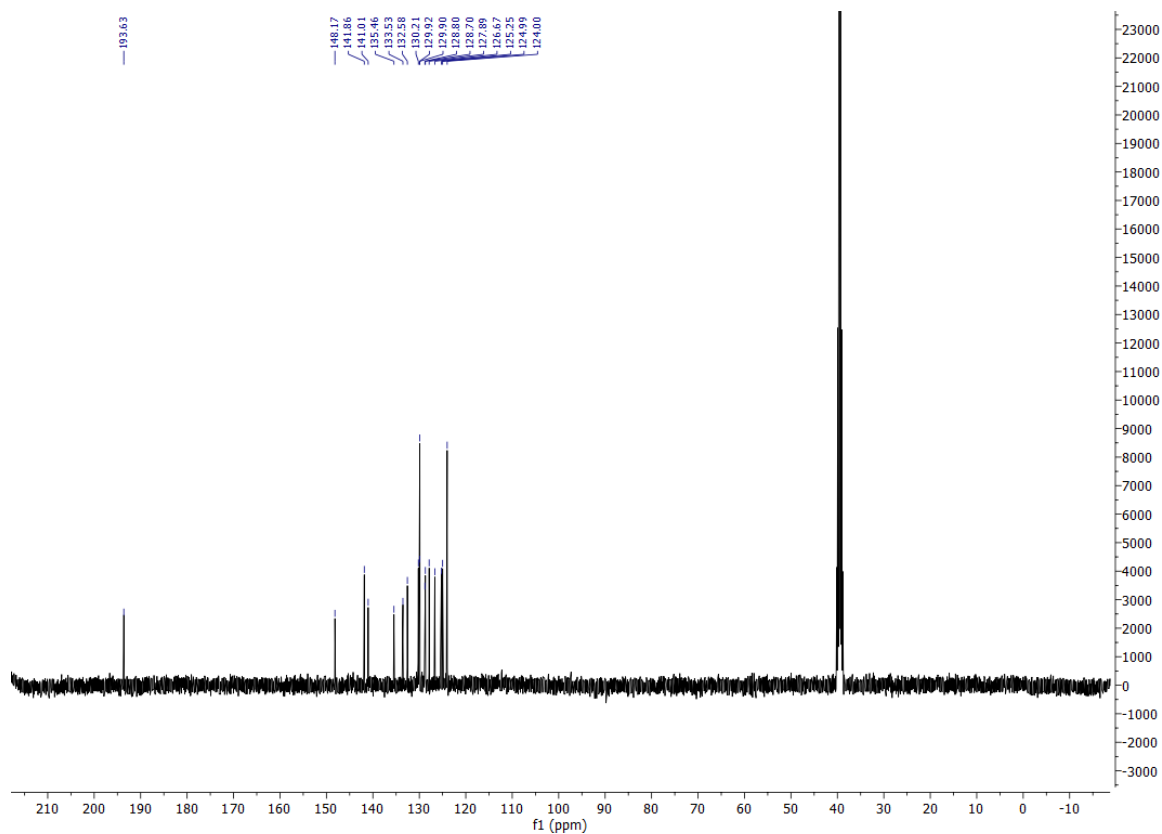


9) Chalcona 2 naftoil cinamaldehydo

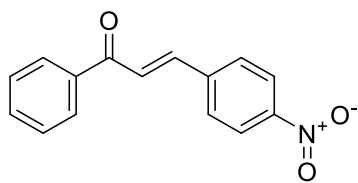


^1H NMR (400 MHz, DMSO) δ 8.39 (d, J = 8.5 Hz, 1H), 8.24 (d, J = 9.2 Hz, 5H), 8.17 (d, J = 8.3 Hz, 3H), 8.08 (d, J = 9.0 Hz, 7H), 8.07 – 7.99 (m, 4H), 7.81 (d, J = 16.0 Hz, 1H), 7.75 – 7.55 (m, 18H). ^{13}C NMR (101 MHz, DMSO) δ 193.63, 148.17, 141.86, 141.01, 135.46, 133.53, 132.58, 130.21, 129.92, 129.90, 128.80, 128.70, 127.89, 126.67, 125.25, 124.99, 124.00.

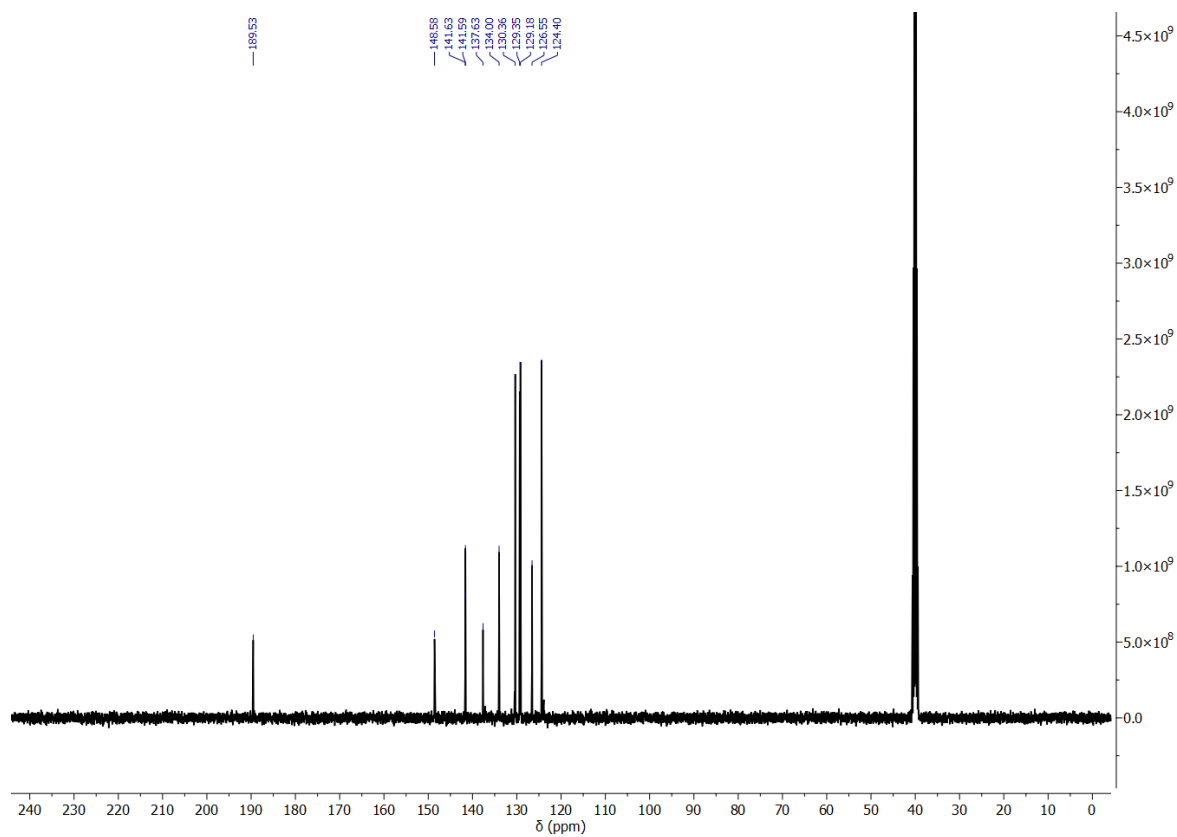
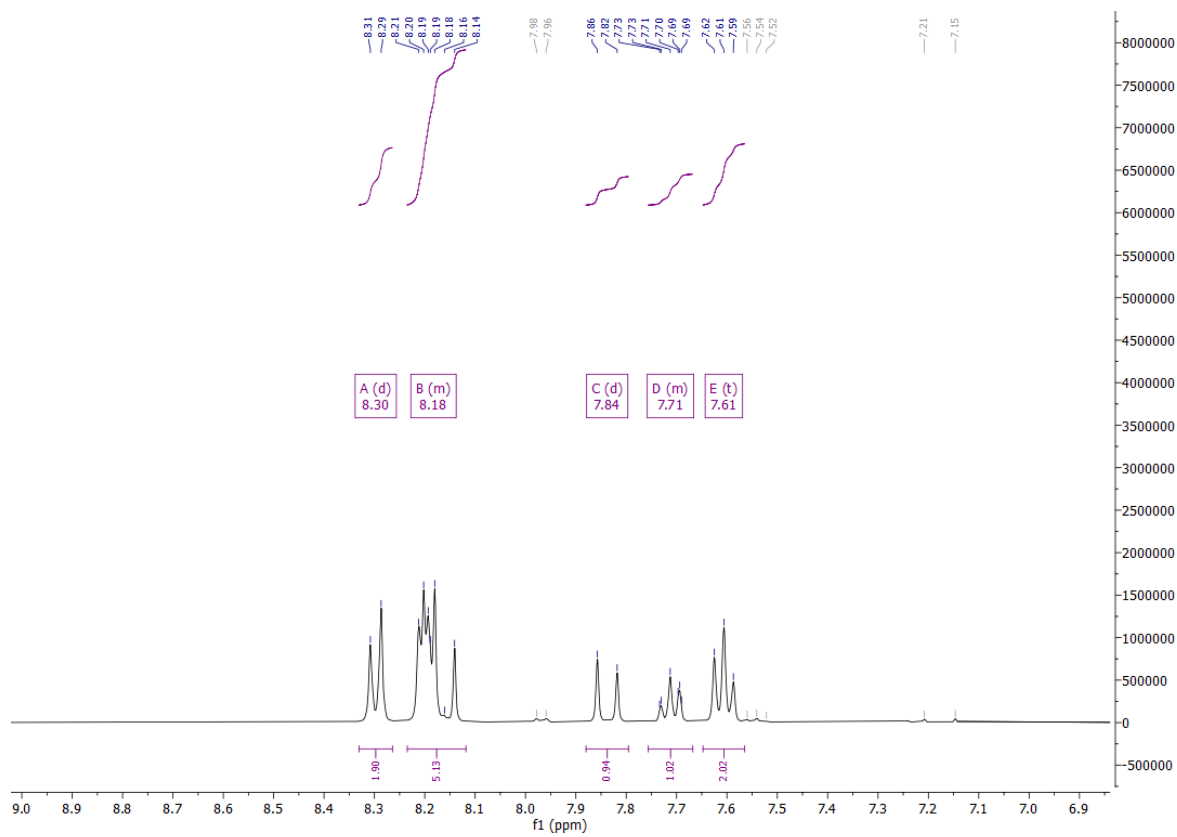




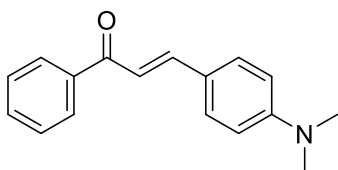
10) Chalcona acetofenona NO₂



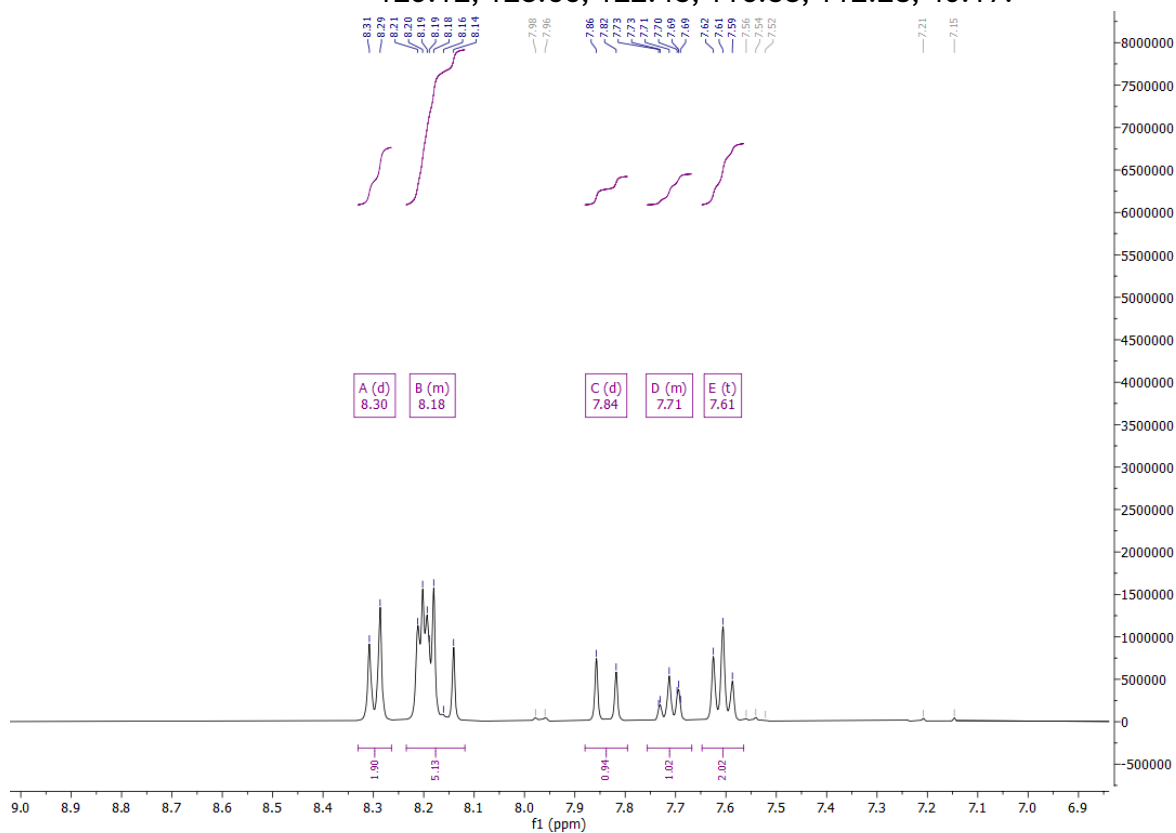
¹H NMR (400 MHz, DMSO) δ 8.30 (d, *J* = 8.6 Hz, 2H), 8.23 – 8.12 (m, 5H), 7.84 (d, *J* = 15.7 Hz, 1H), 7.76 – 7.67 (m, 1H), 7.61 (t, *J* = 7.6 Hz, 2H). ¹³C NMR (101 MHz, DMSO) δ 189.53, 148.58, 141.63, 141.59, 137.63, 134.00, 130.36, 129.35, 129.18, 126.55, 124.40.



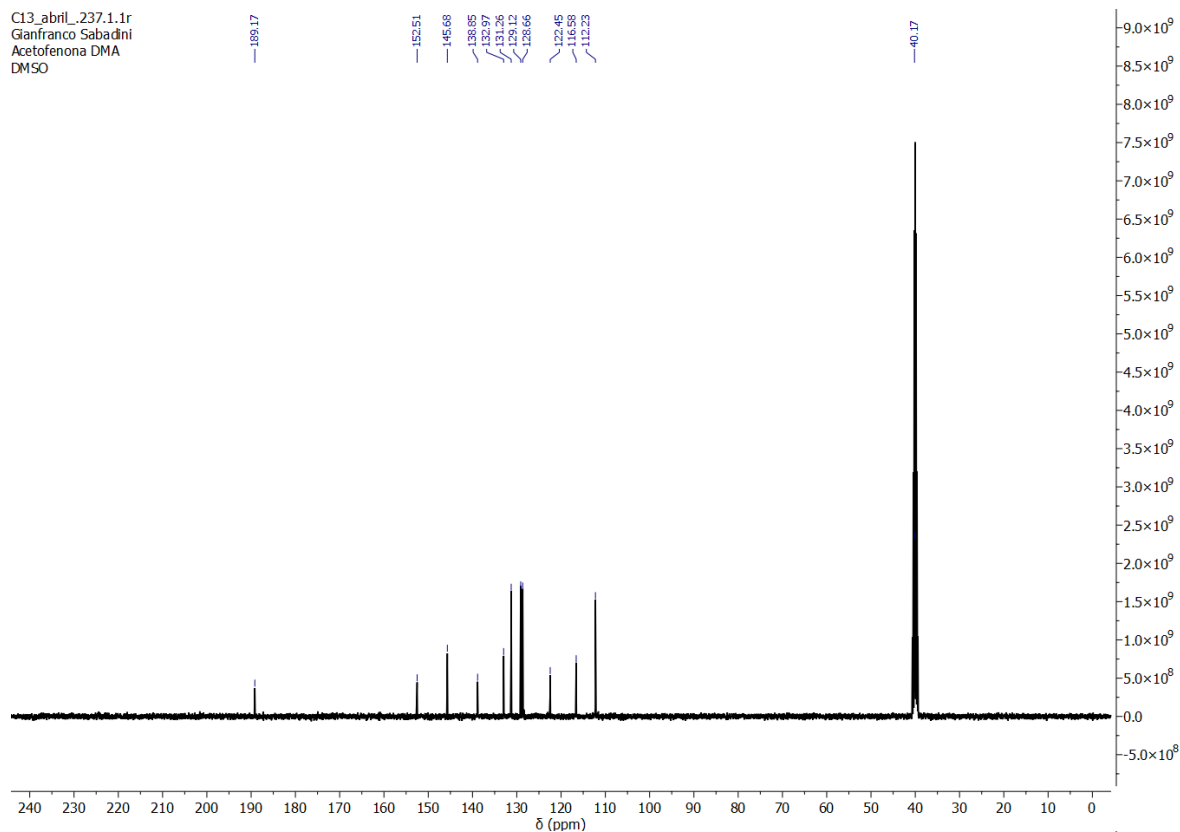
11) Chalcona acetofenona DMA



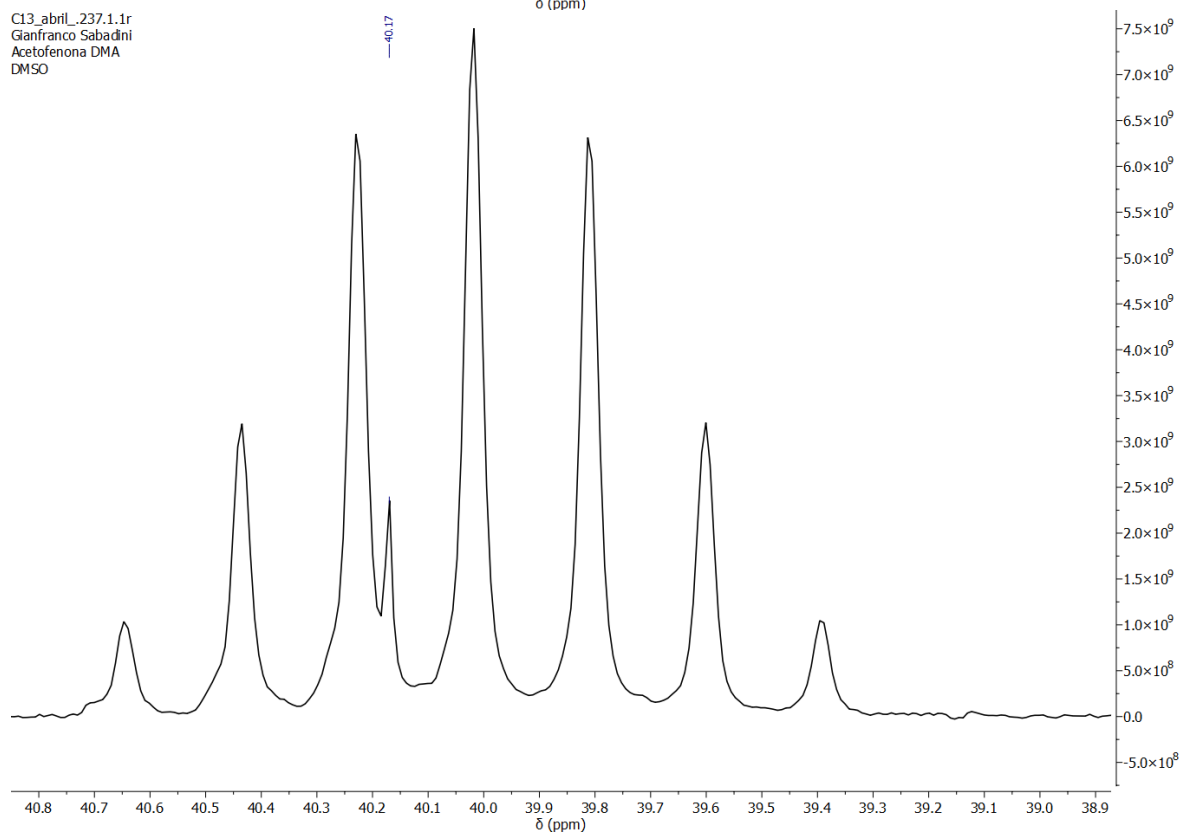
^1H NMR (400 MHz, DMSO) δ 8.15 – 8.08 (m, 2H), 7.75 – 7.67 (m, 3H), 7.67 – 7.59 (m, 2H), 7.56 (dd, J = 8.1, 6.7 Hz, 2H), 6.75 (d, J = 8.5 Hz, 2H), 3.02 (s, 6H). ^{13}C NMR (101 MHz, DMSO) δ 189.17, 152.51, 145.68, 138.85, 132.97, 131.26, 129.12, 128.66, 122.45, 116.58, 112.23, 40.17.



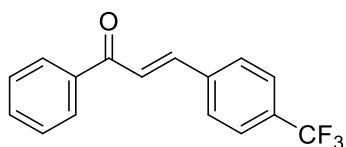
C13_abril_237.1.1r
 Gianfranco Sabadini
 Acetofenona DMA
 DMSO



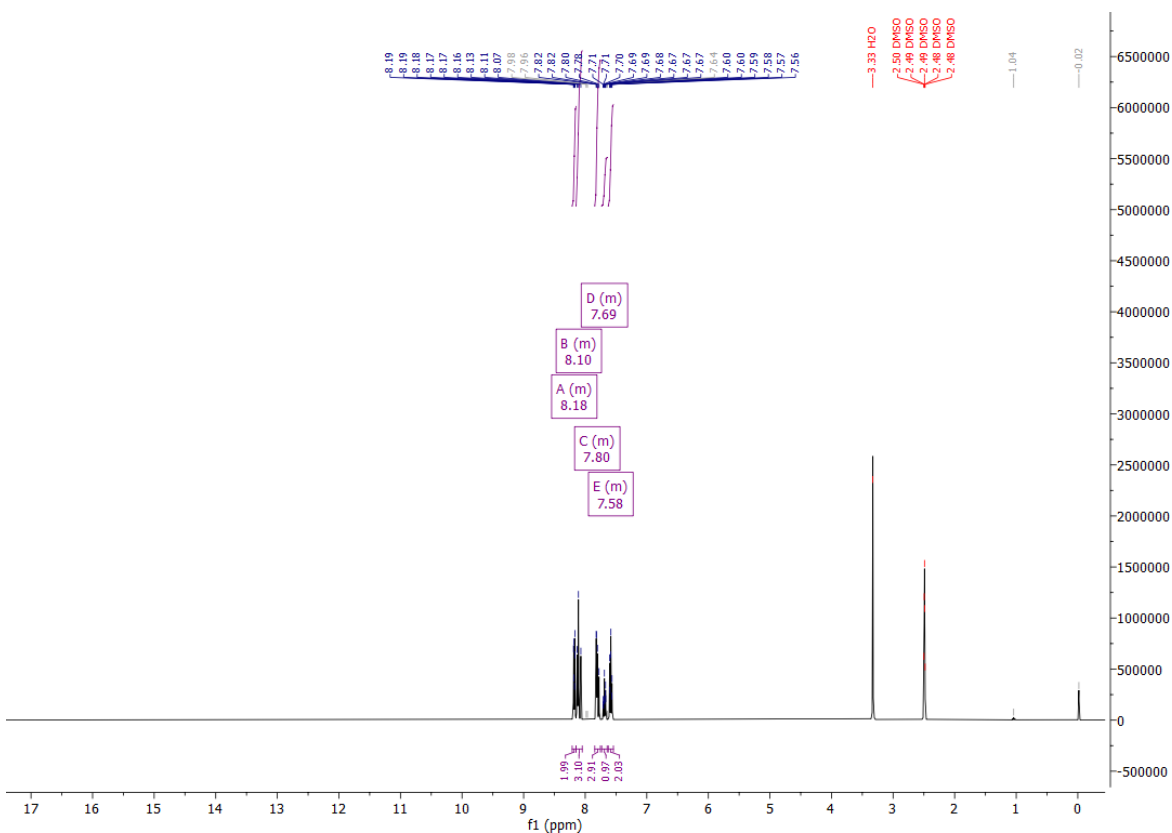
C13_abril_237.1.1r
 Gianfranco Sabadini
 Acetofenona DMA
 DMSO



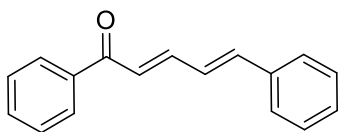
12) Chalcona acetofenona CF₃



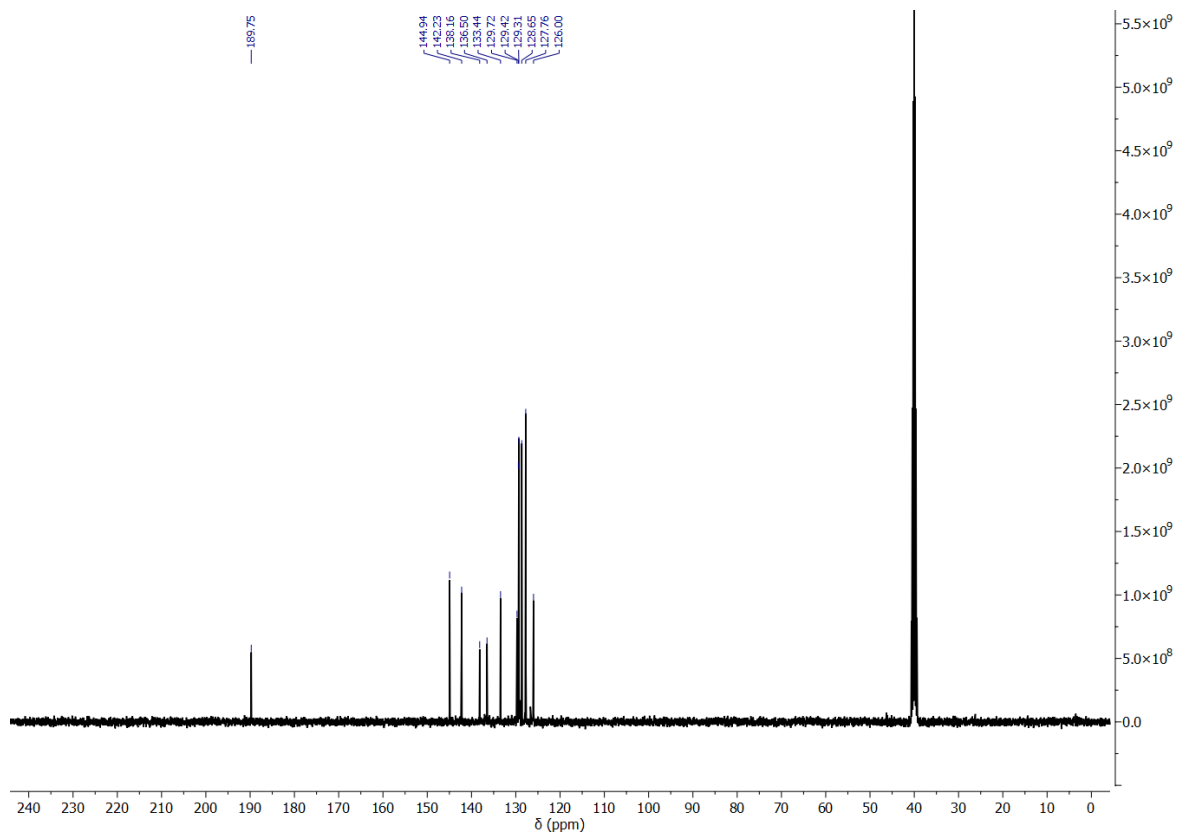
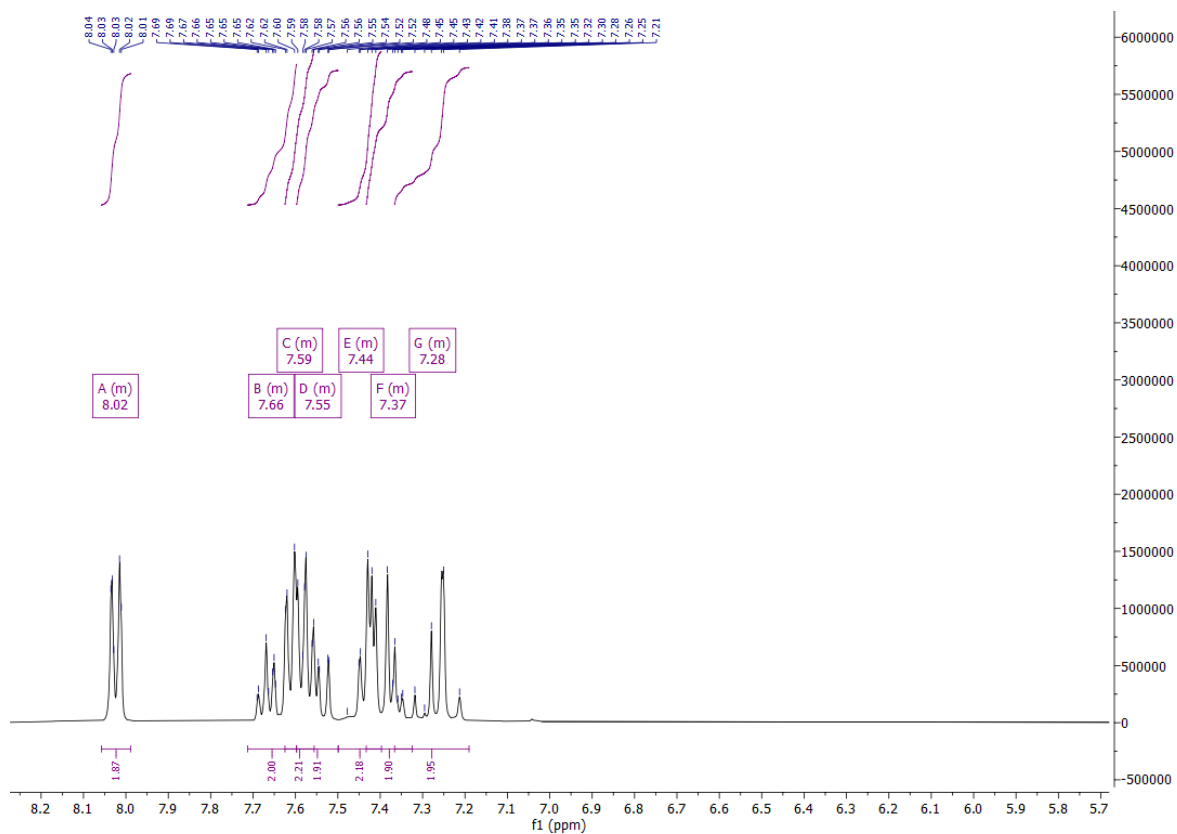
¹H NMR (400 MHz, DMSO) δ 8.21 – 8.15 (m, 2H), 8.15 – 8.05 (m, 3H), 7.84 – 7.75 (m, 3H), 7.73 – 7.64 (m, 1H), 7.63 – 7.54 (m, 2H). ¹³C NMR (101 MHz, DMSO) δ 189.61, 142.44, 139.16, 137.73, 133.89, 131.01, 130.69, 130.37, 129.95, 129.33, 129.15, 128.59, δ 126.16(q), 125.88, 125.23, 123.18, 120.48)



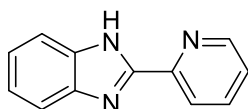
11) Chalcona acetofenona cinamaldehydo



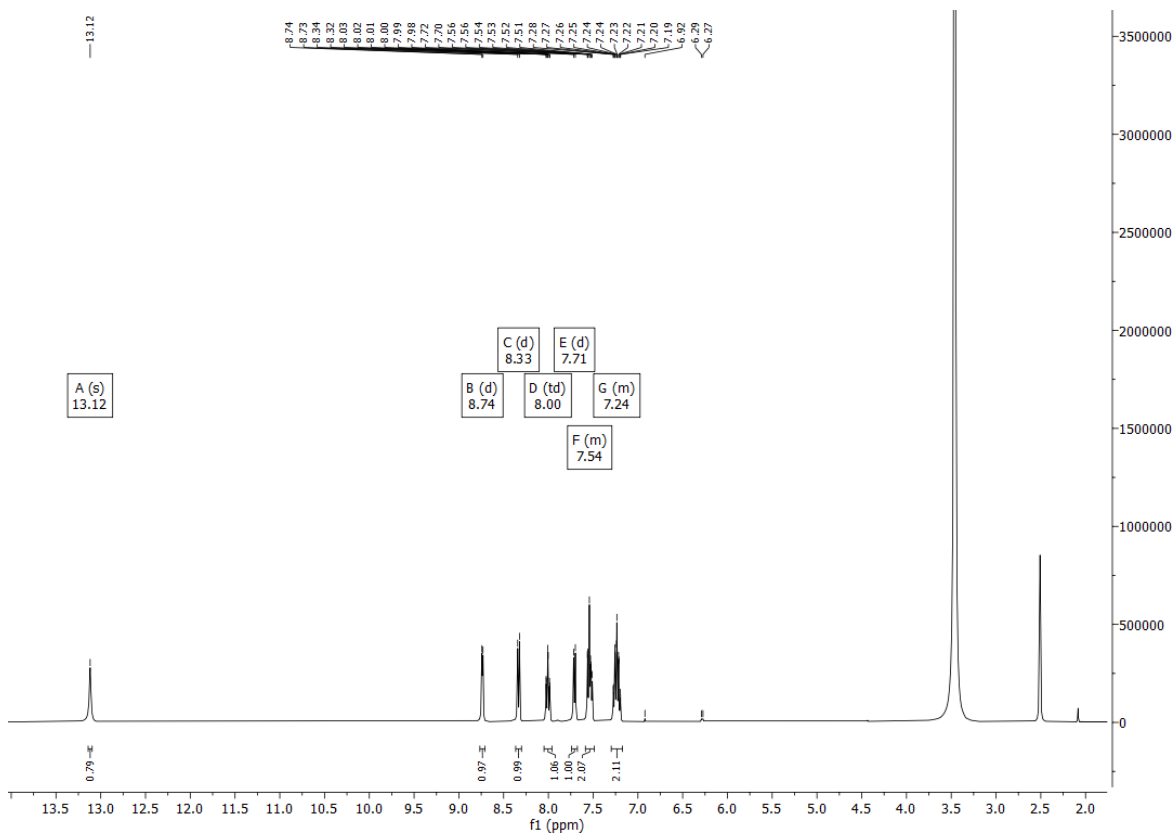
¹H NMR (400 MHz, DMSO) δ 8.06 – 7.99 (m, 2H), 7.71 – 7.60 (m, 2H), 7.62 – 7.56 (m, 2H), 7.60 – 7.50 (m, 2H), 7.50 – 7.40 (m, 2H), 7.43 – 7.32 (m, 2H), 7.37 – 7.19 (m, 2H). ¹³C NMR (101 MHz, DMSO) δ 189.75, 144.94, 142.23, 138.16, 136.50, 133.44, 129.72, 129.42, 129.31, 128.65, 127.76, 126.00.

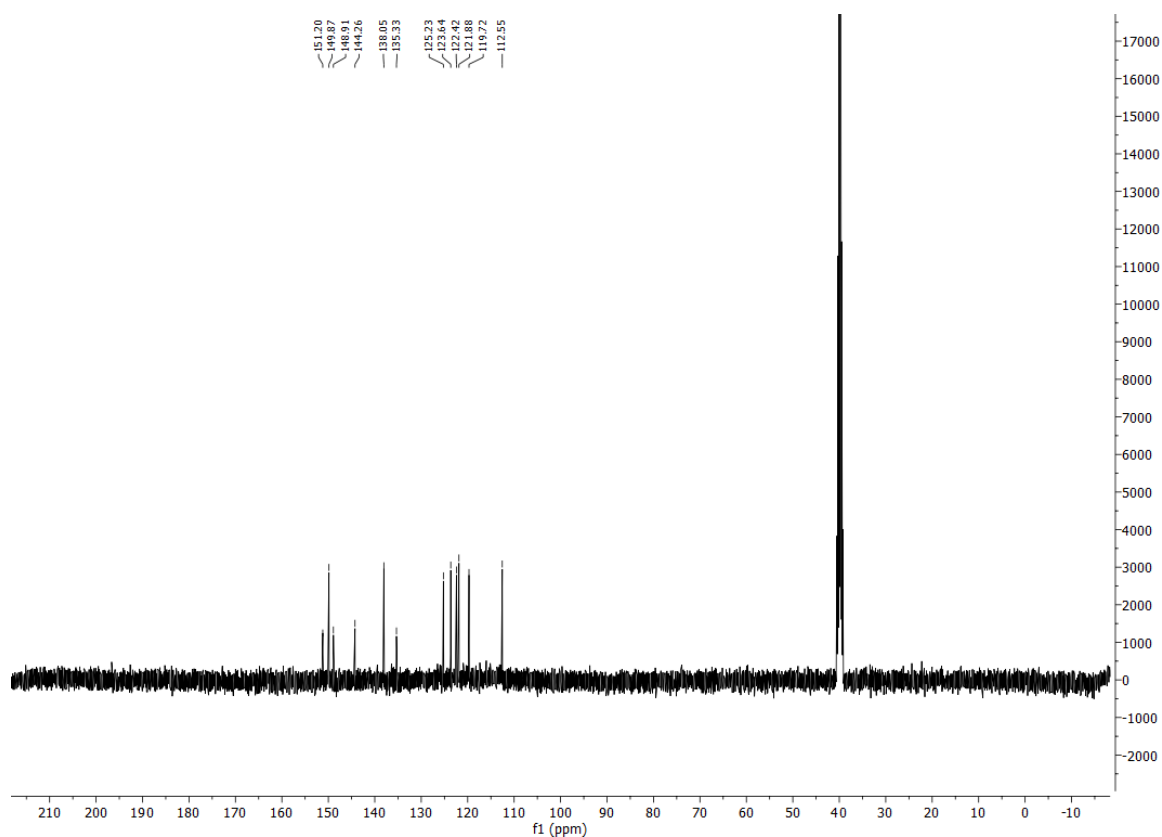


12) Benzimidazol 2 (2 piridina)

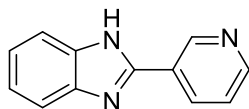


^1H NMR (400 MHz, DMSO) δ 13.12 (s, 1H), 8.74 (d, J = 4.9 Hz, 1H), 8.33 (d, J = 7.9 Hz, 1H), 8.00 (td, J = 7.7, 1.8 Hz, 1H), 7.71 (d, J = 7.6 Hz, 1H), 7.59 – 7.49 (m, 2H), 7.30 – 7.17 (m, 2H). ^{13}C NMR (101 MHz, DMSO) δ 151.20, 149.87, 148.91, 144.26, 138.05, 135.33, 125.23, 123.64, 122.42, 121.88, 119.72, 112.55.

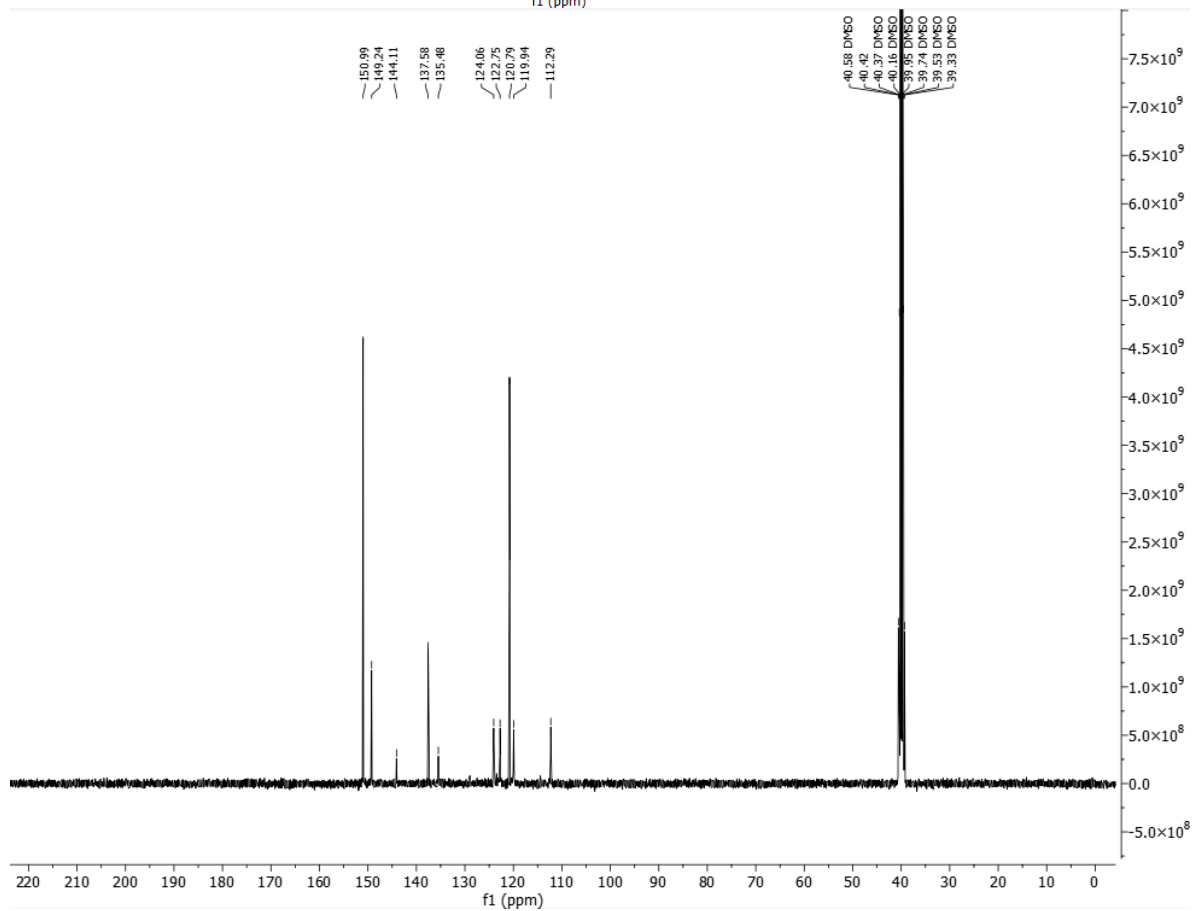
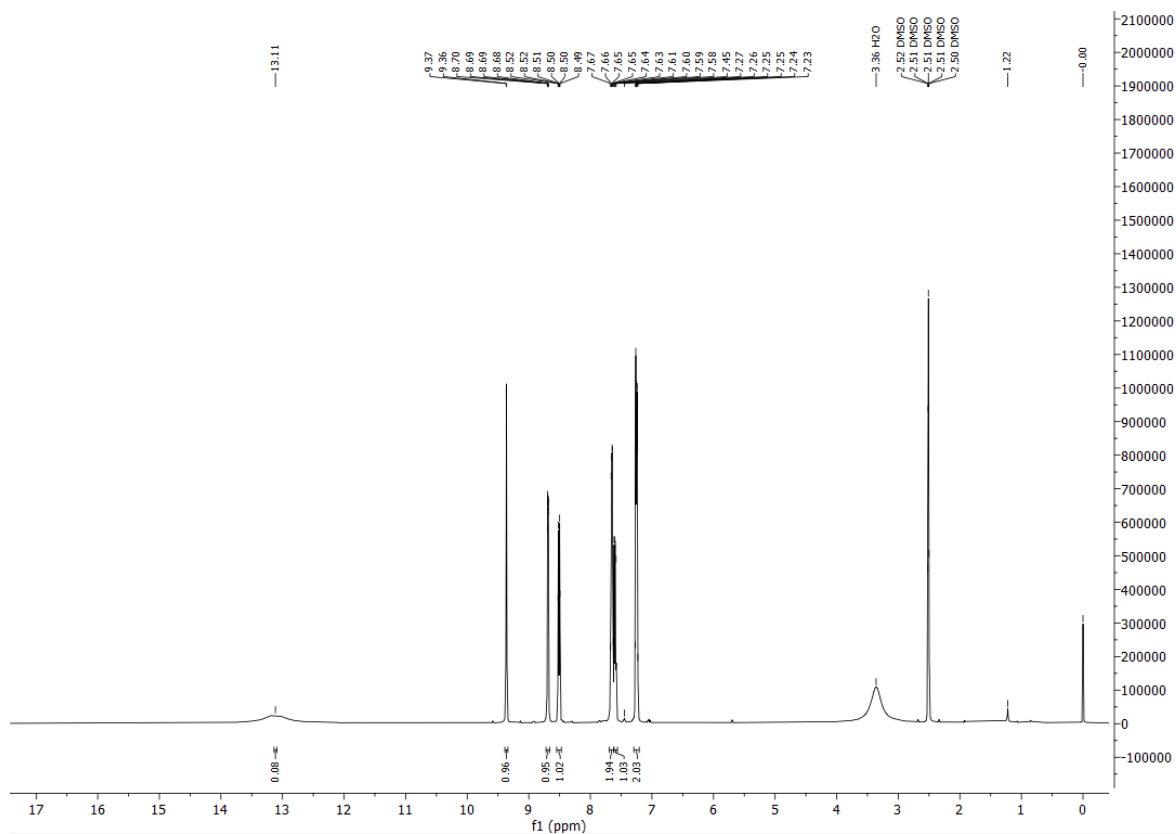




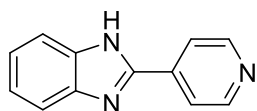
13) Benzimidazol 2 (3 piridina)



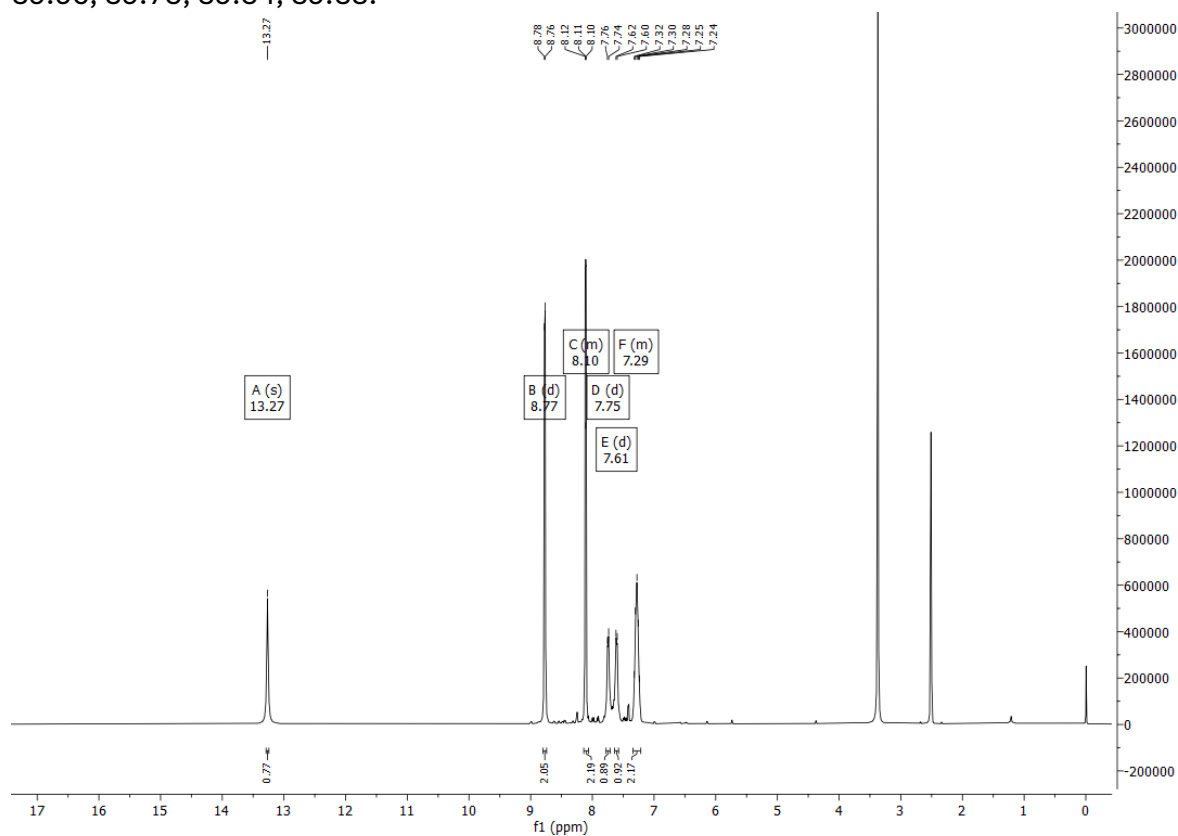
^1H NMR (400 MHz, DMSO) δ 13.11, 9.37, 9.36, 8.70, 8.69, 8.69, 8.68, 8.52, 8.52, 8.51, 8.50, 8.50, 8.49, 7.67, 7.66, 7.65, 7.65, 7.64, 7.63, 7.61, 7.60, 7.59, 7.58, 7.45, 7.27, 7.26, 7.25, 7.25, 7.24, 7.23. ^{13}C NMR (101 MHz, DMSO) δ 150.99, 149.24, 144.11, 137.58, 135.48, 124.06, 122.75, 120.79, 119.94, 112.29.

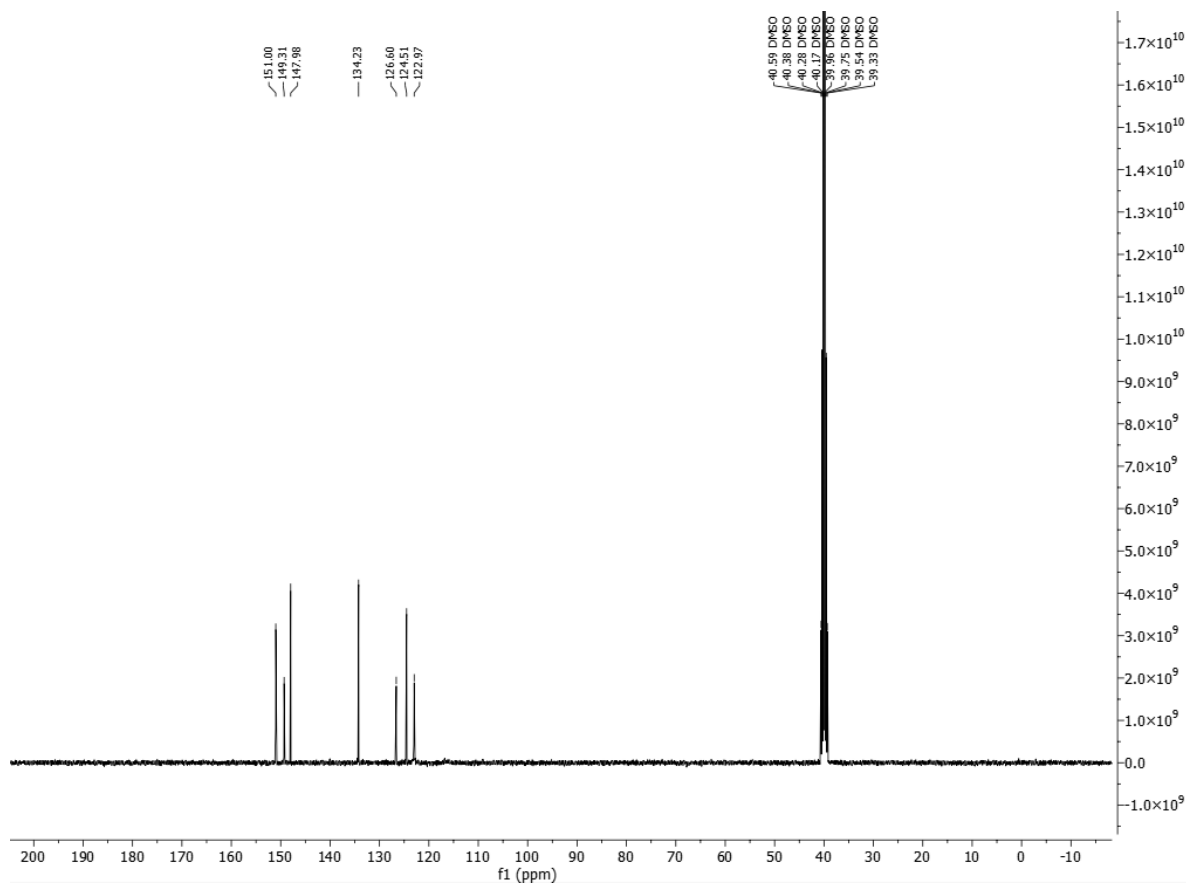


14)Benzimidazol 2 (4 piridina)

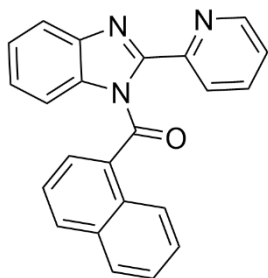


^1H NMR (400 MHz, DMSO) δ 13.27, 8.78, 8.76, 8.12, 8.11, 8.10, 7.76, 7.74, 7.62, 7.60, 7.32, 7.30, 7.28, 7.25, 7.24. ^{13}C NMR (101 MHz, DMSO) δ 151.00, 149.31, 147.98, 134.23, 126.60, 124.51, 122.97, 40.59, 40.38, 40.28, 40.17, 39.96, 39.75, 39.54, 39.33.

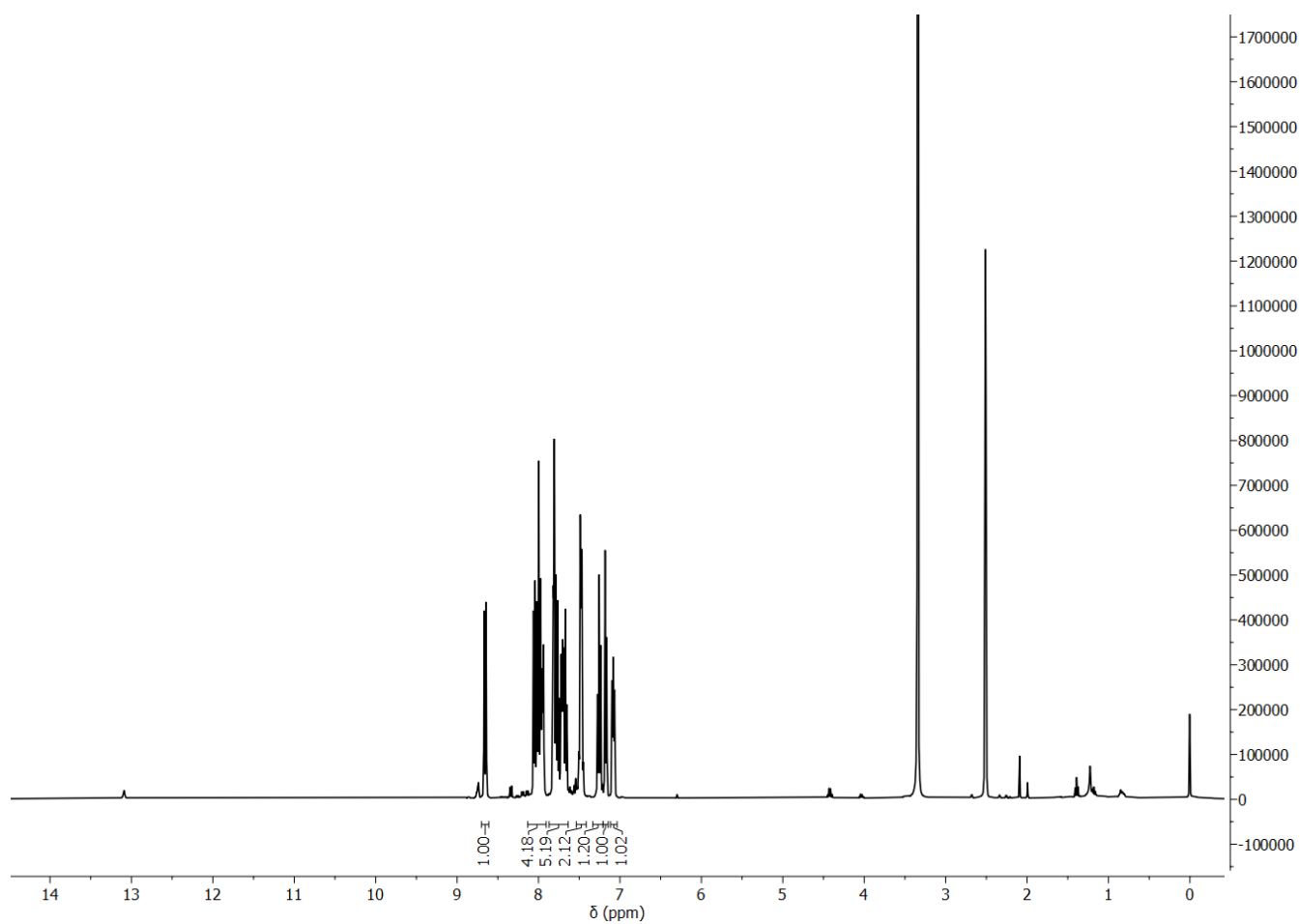


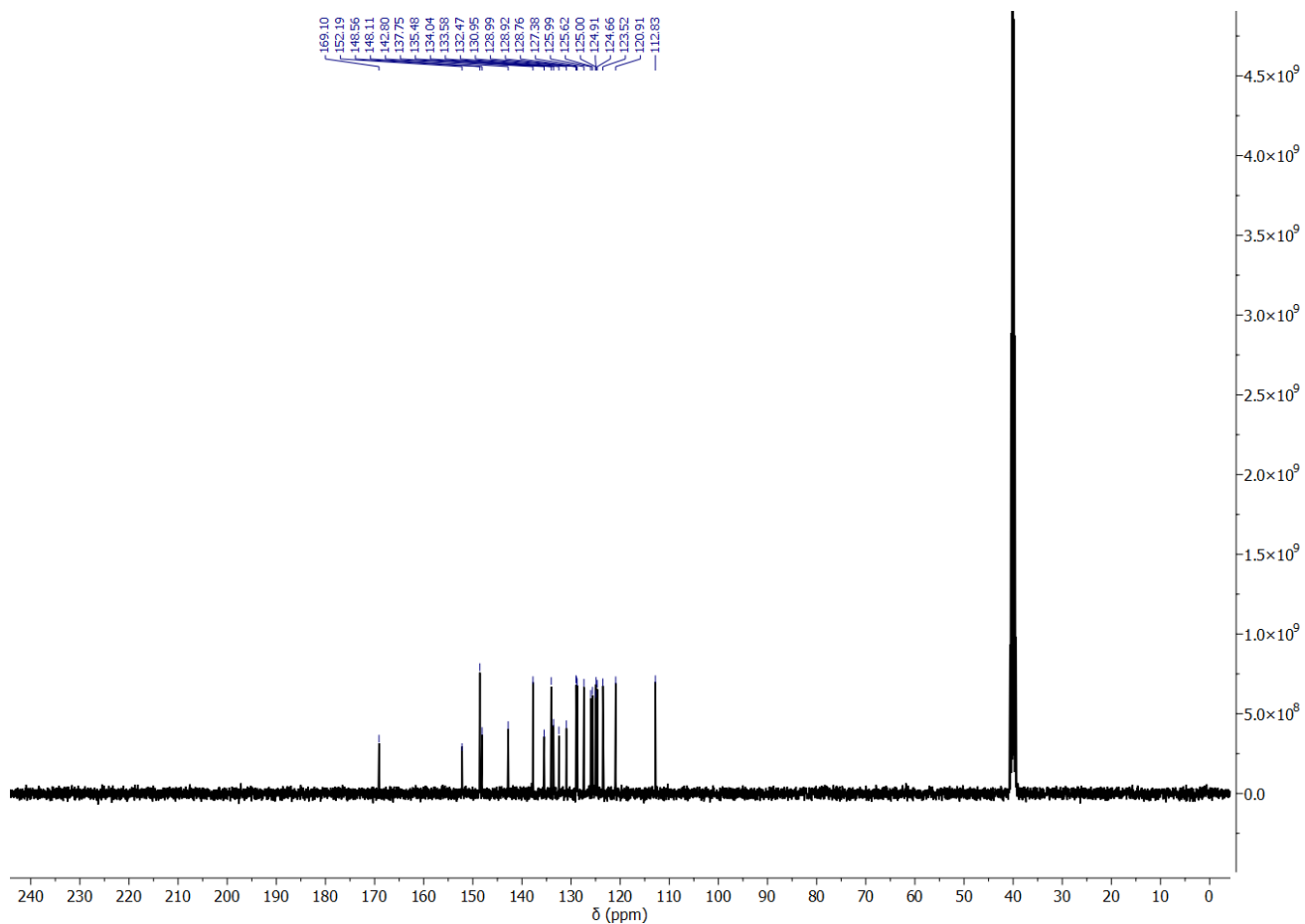


15) Benzimidazol 1 (1-(naftalen-1-carbonil)-2-(piridin-2-il)-benzimidazol)

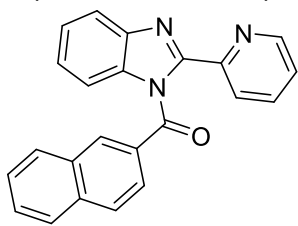


^{13}C NMR (101 MHz, DMSO) δ 169.10, 152.19, 148.56, 148.11, 142.80, 137.75, 135.48, 134.04, 133.58, 132.47, 130.95, 128.99, 128.92, 128.76, 127.38, 125.99, 125.62, 125.00, 124.91, 124.66, 123.52, 120.91, 112.83. ^1H NMR (400 MHz, DMSO) δ 8.66 (d, J = 8.6 Hz, 1H), 8.00 (m, 4H), 7.83–7.59 (m, 5H), 7.73–7.59 (m, 2H), 7.34–7.23 (m, 1H), 7.18–7.16 (m, 1H), 7.10 (m, 1H). ^{13}C NMR (101 MHz, DMSO) δ 169.10, 152.19, 148.56, 148.11, 142.80, 137.75, 135.48, 134.04, 133.58, 132.47, 130.95, 128.99, 128.92, 128.76, 127.38, 125.99, 125.62, 125.00, 124.91, 124.66, 123.52, 120.91, 112.83.

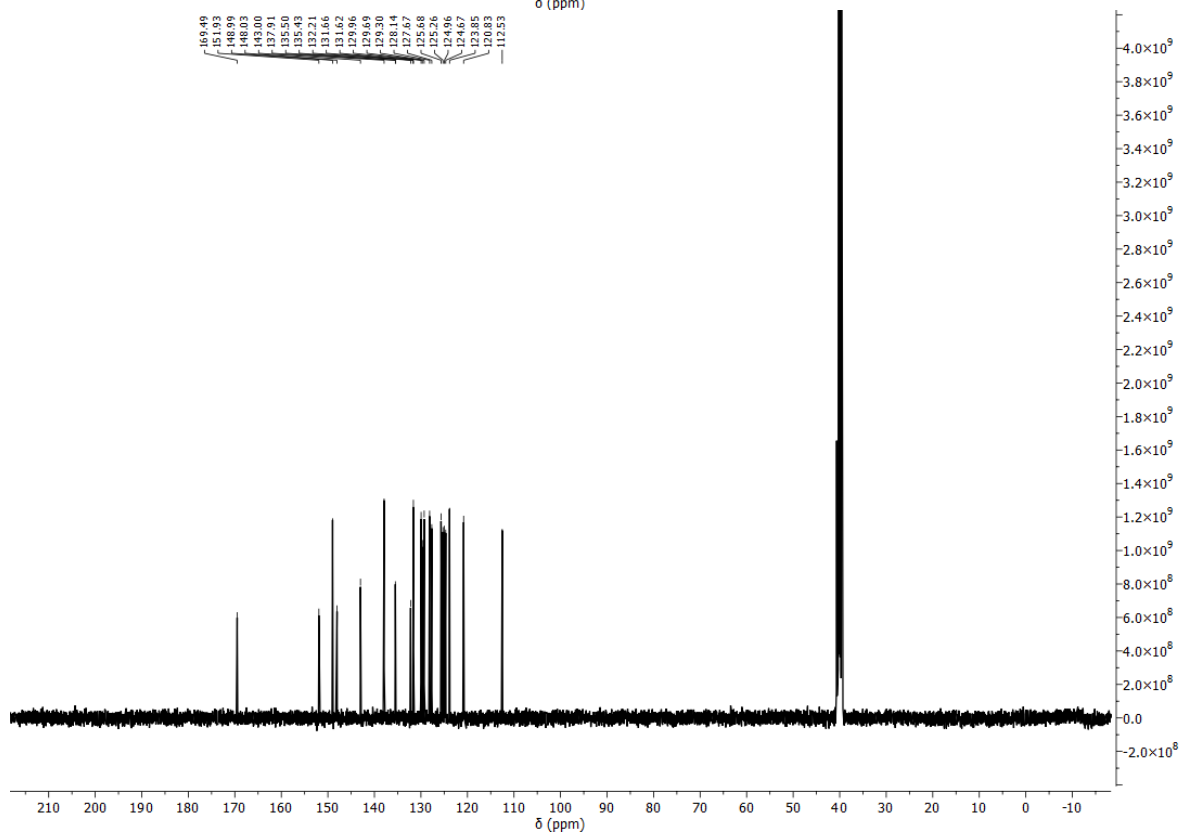
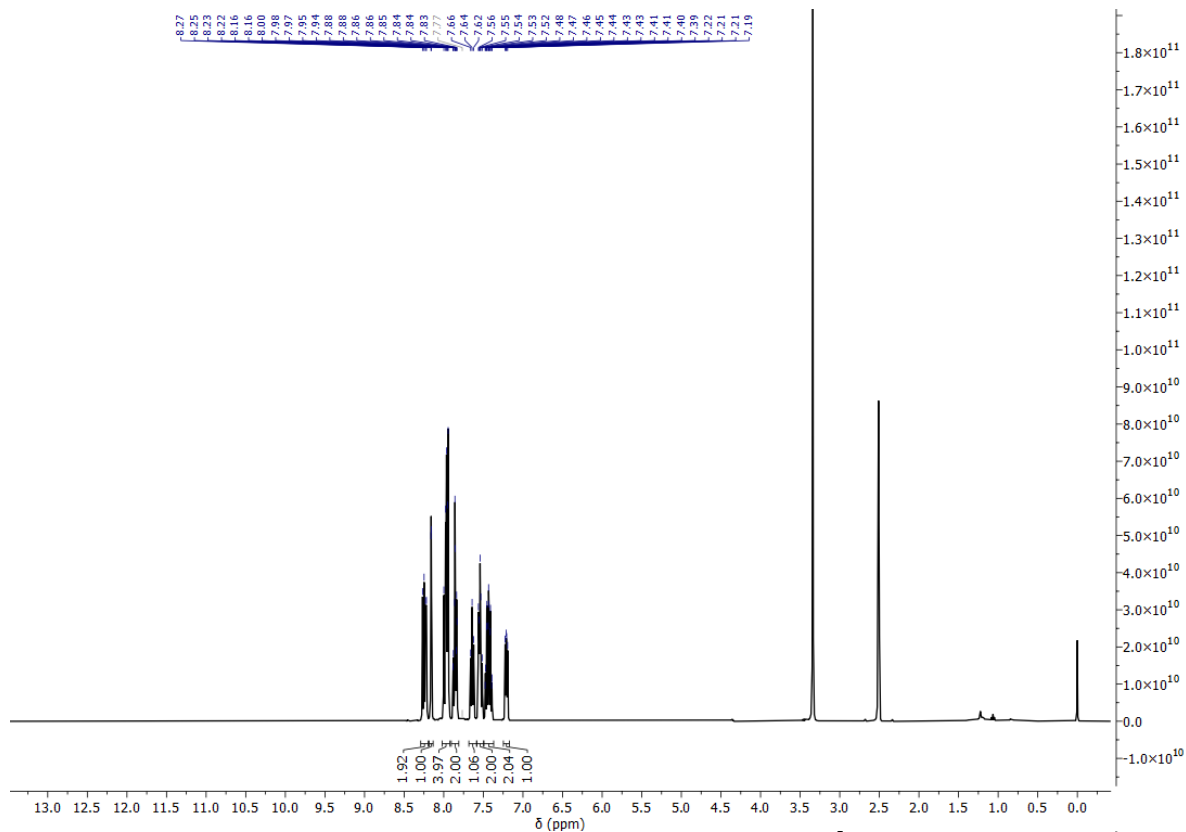




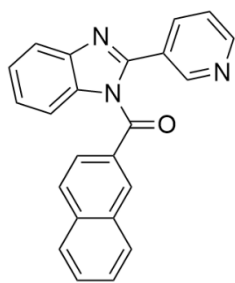
15)Benzimidazol 2 (1-(naftalen-2-carbonil)-2-(piridin-2-il)-)benzimidazol



^1H NMR (400 MHz, DMSO) δ 8.24 (dd, J = 13.3, 6.4 Hz, 2H), 8.16 (d, J = 1.8 Hz, 1H), 8.02 – 7.92 (m, 4H), 7.90 – 7.81 (m, 2H), 7.64 (t, J = 7.6 Hz, 1H), 7.58 – 7.49 (m, 2H), 7.49 – 7.37 (m, 2H), 7.21 (dd, J = 7.6, 4.9 Hz, 1H). ^{13}C NMR (101 MHz, DMSO) δ 169.50, 151.93, 148.99, 148.03, 143.00, 137.92, 135.50, 135.43, 132.21, 131.66, 131.62, 129.96, 129.69, 129.30, 128.14, 127.67, 125.68, 125.26, 124.96, 124.67, 123.85, 120.83, 112.53.

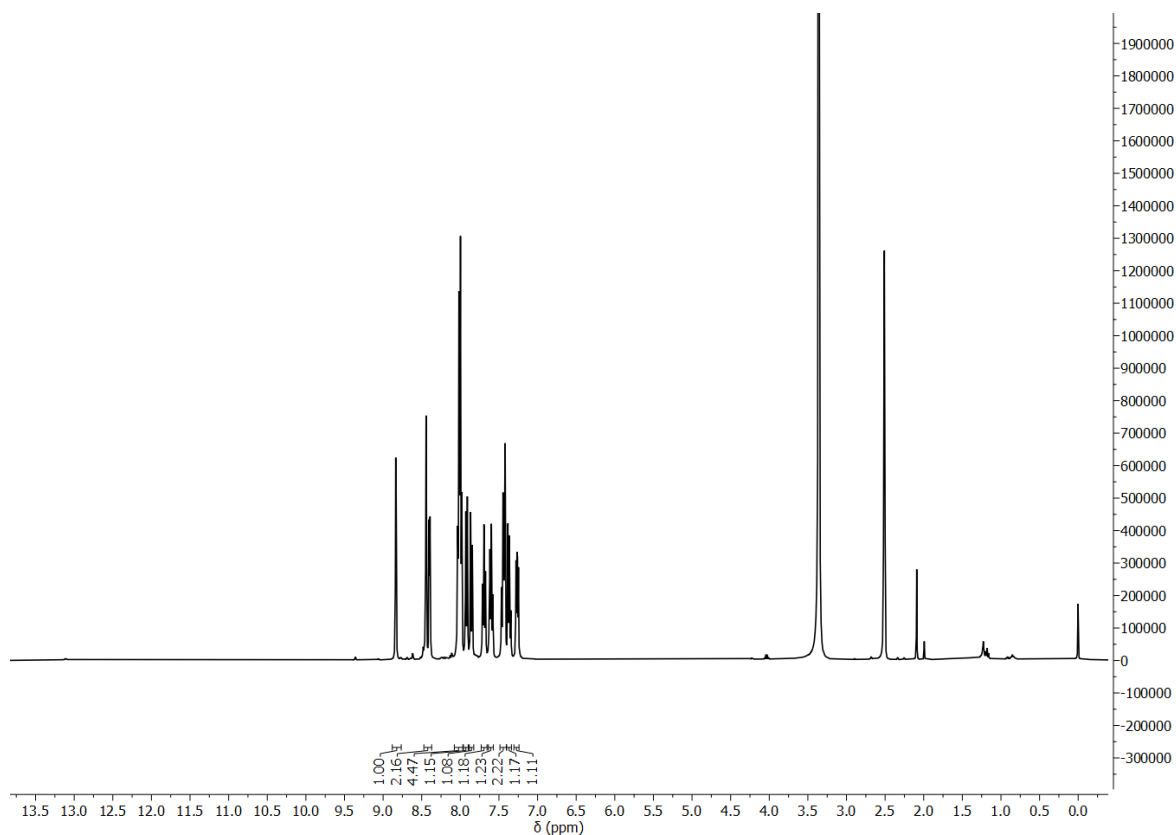


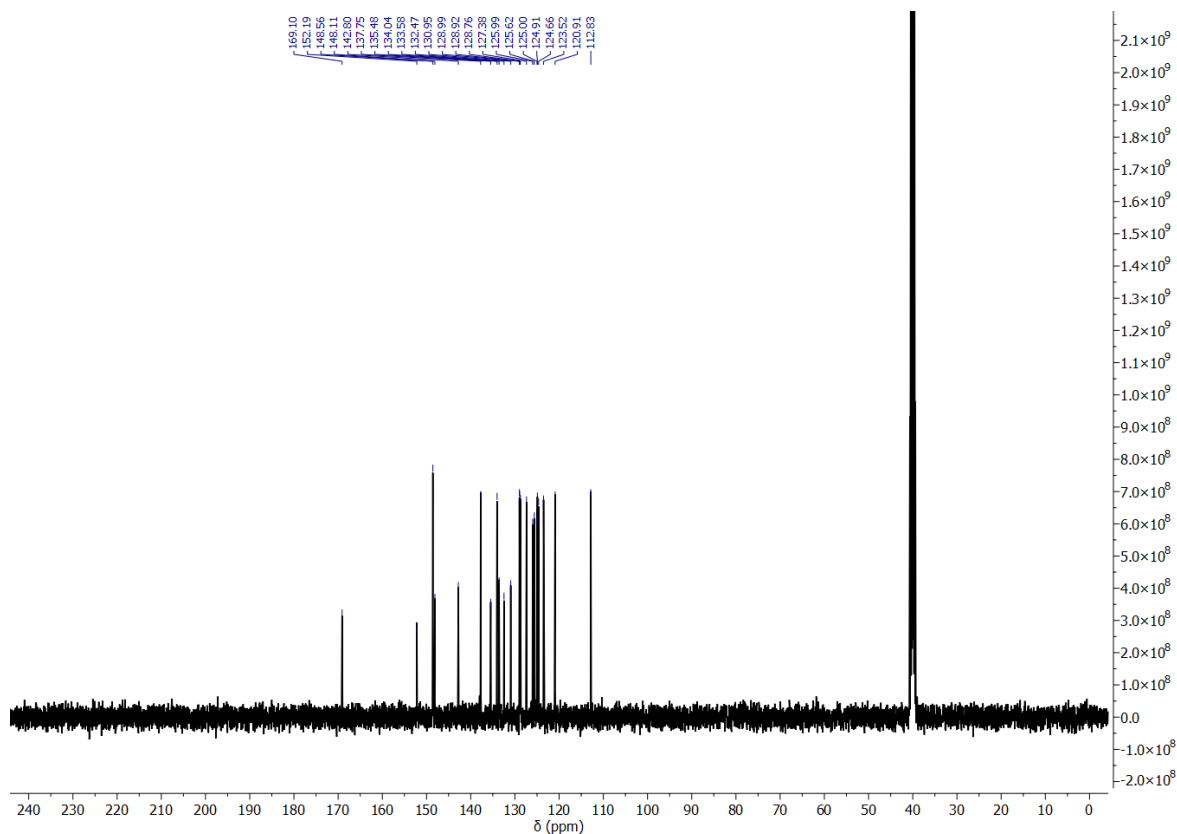
16) Benzimidazol 4 (Benzimidazol 2 (1-(naftalen-2-carbonil)-2-(piridin-3-il)-)benzimidazol)



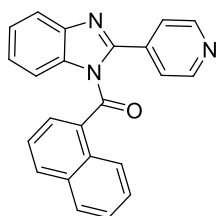
^1H NMR (400 MHz, DMSO) δ 8.83 (d, J = 2.3 Hz, 1H), 8.47 – 8.37 (m, 2H), 8.08 – 7.97 (m, 4H), 7.92 (d, J = 7.9 Hz, 1H), 7.85 (d, J = 1.8 Hz, 1H), 7.69 (t, J = 7.5 Hz, 1H), 7.64 – 7.57 (m, 1H), 7.49 – 7.40 (m, 2H), 7.40 – 7.34 (m, 1H), 7.26 (d, J = 4.8 Hz, 1H). ^{13}C NMR (101 MHz, DMSO) δ 168.94, 151.57, 150.60, 149.77, 143.06, 137.02, 135.64, 135.09, 133.38, 132.12, 130.13, 129.95, 129.93, 129.31, 128.22, 127.81, 127.33, 125.68, 125.57,

125.07, 123.49, 120.61, 113.85.

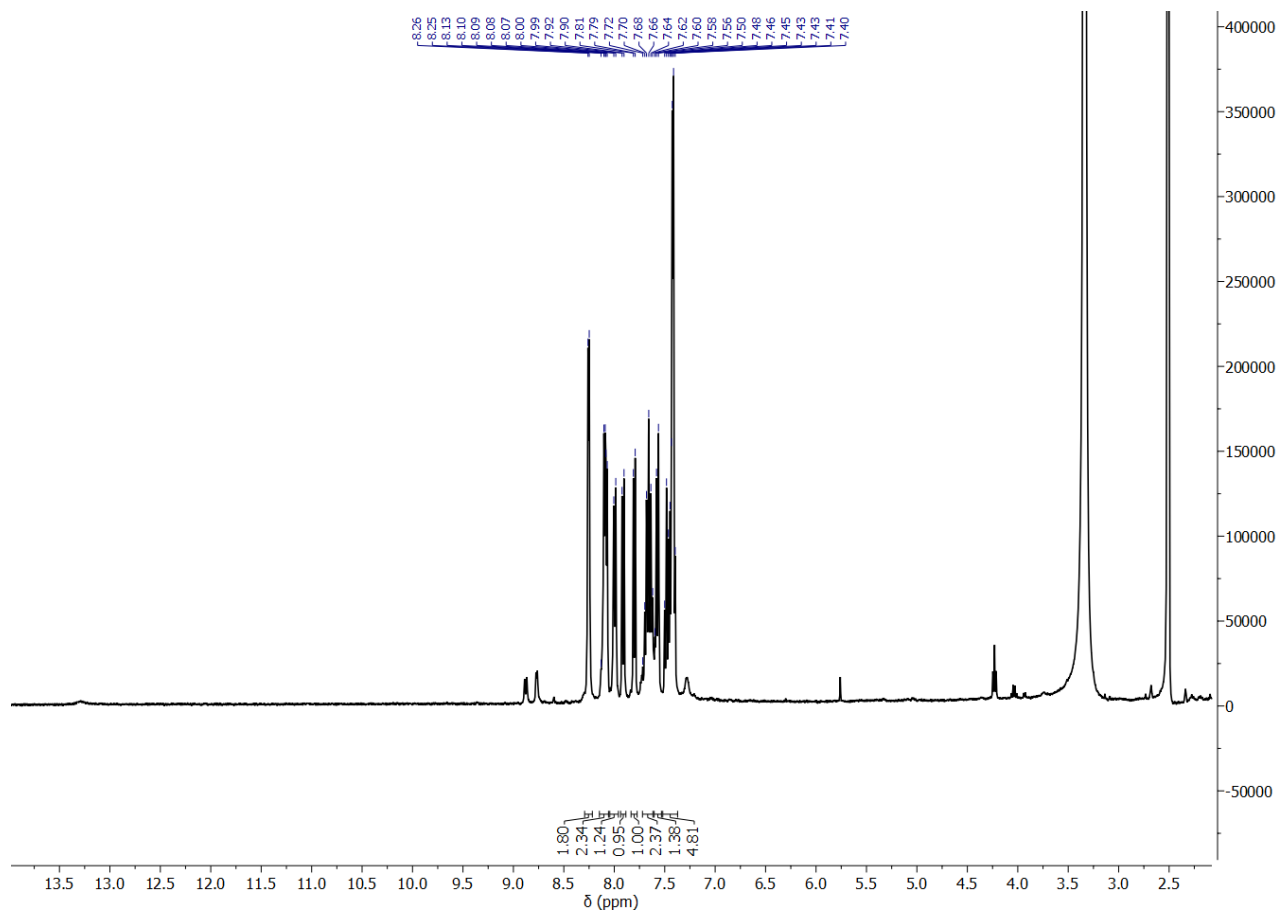




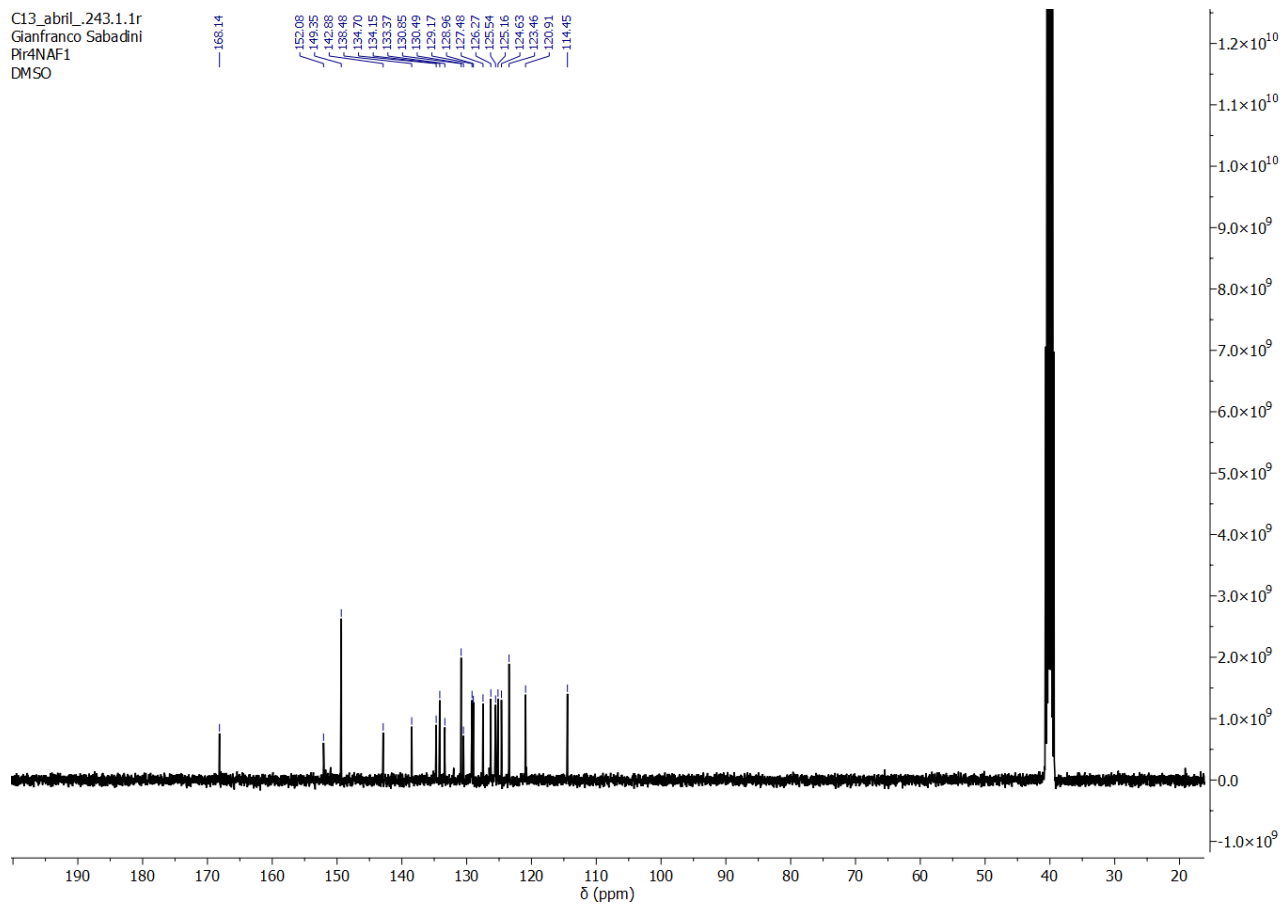
17)Benzimidazol 5 (Benzimidazol 2 (1-(naftalen-1-carbonil)-2-(piridin-4-il)-benzimidazol)



^1H NMR (400 MHz, DMSO) δ 8.25 (d, J = 5.0 Hz, 2H), 8.13 – 8.06 (m, 2H), 7.99 (d, J = 7.8 Hz, 1H), 7.91 (d, J = 7.8 Hz, 1H), 7.80 (d, J = 7.2 Hz, 1H), 7.70 – 7.62 (m, 2H), 7.59 – 7.53 (m, 1H), 7.52 – 7.37 (m, 5H). ^{13}C NMR (101 MHz, DMSO) δ 168.14, 152.08, 149.35, 142.88, 138.48, 134.70, 134.15, 133.37, 130.85, 130.49, 129.17, 128.96, 127.48, 126.27, 125.54, 125.16, 124.63, 123.46, 120.91, 114.45.

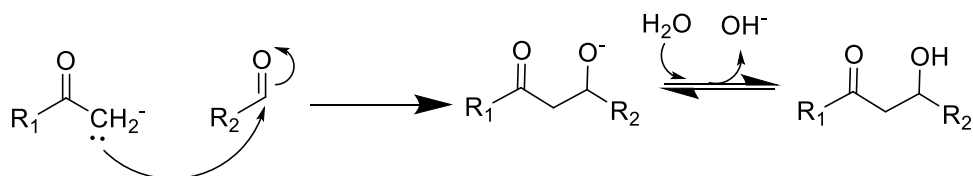
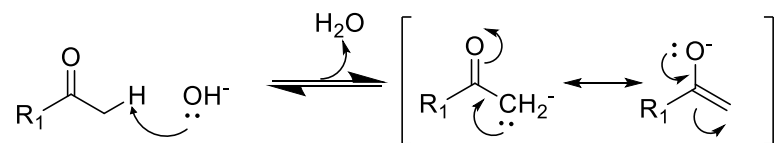


C13_abril_243.1.1r
Gianfranco Sabadini
Pir4NAF1
DMSO

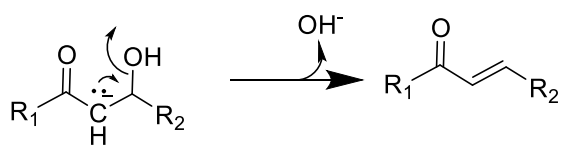
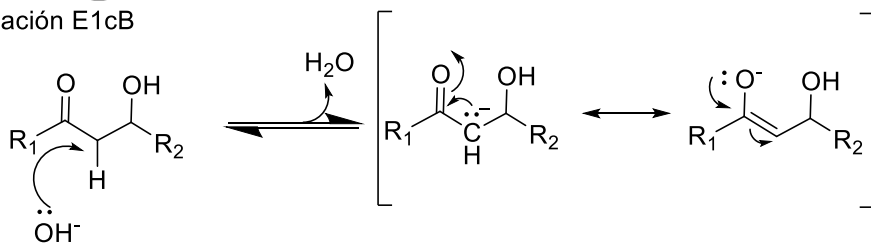


Anexo VI

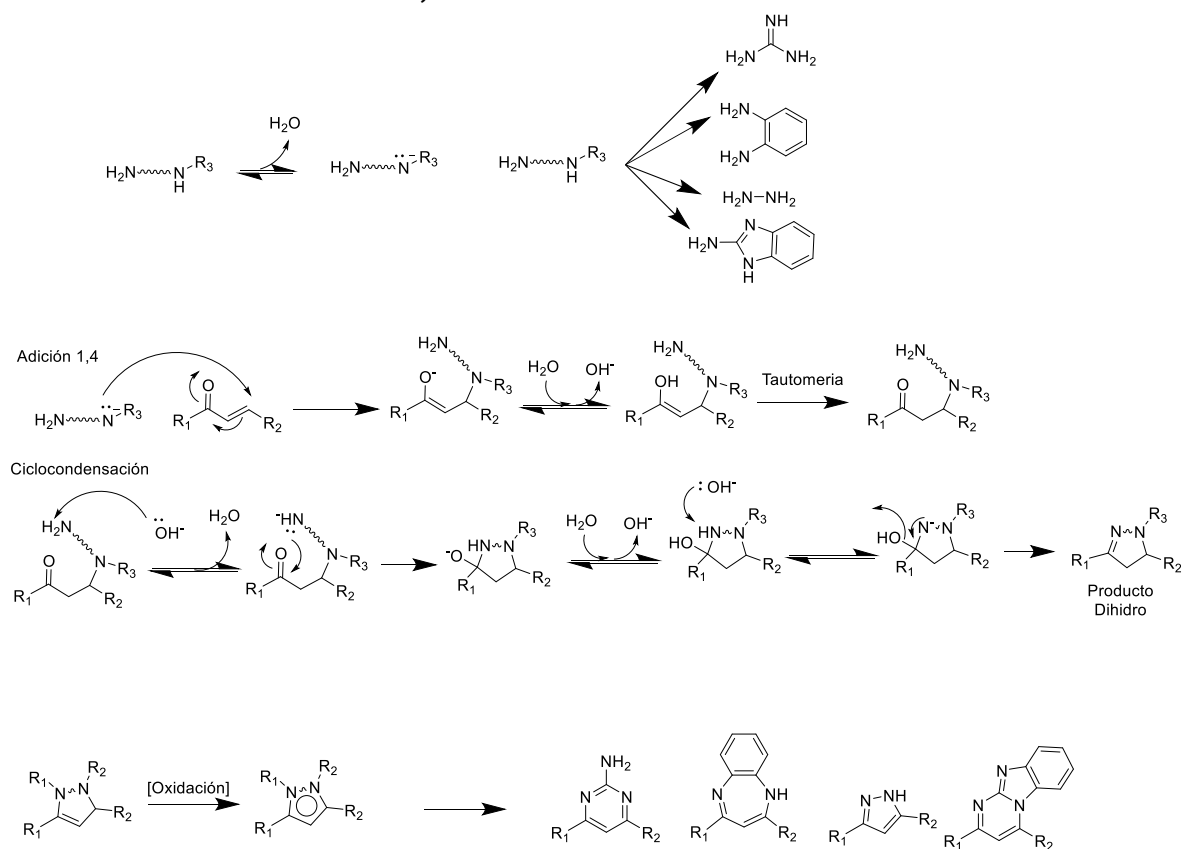
Mecanismo de síntesis de chalconas en medio básico con hidróxido



Eliminación E1cB



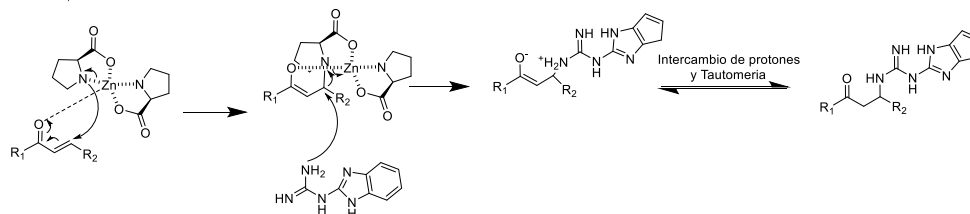
Mecanismo de síntesis de heterociclos con guanidina,2 aminobenzimidazol, hidrazina u orto-fenilendiamina



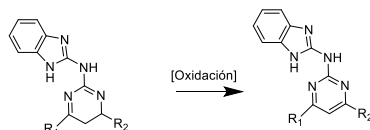
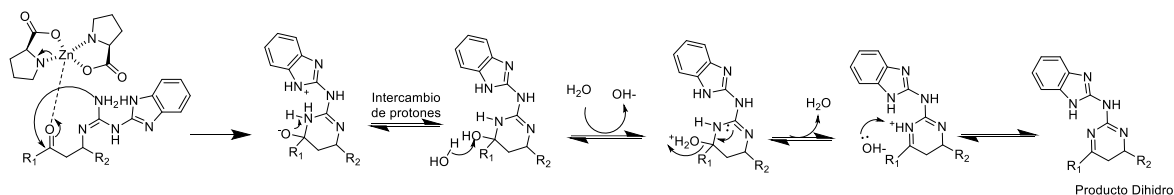
En estas condiciones, se asume la formación de un nucleófilo aniónico. Pero en medio proteicos como etanol la reacción puede transcurrir mediante una vía neutra o parcialmente anionica, en la cual el nucleófilo adiciona al sistema α,β -insaturado de la chalcona, seguido de transferencias de protón, ciclación intramolecular y posterior deshidratación, tautomerización u oxidación/aromatización, según la familia heterocíclica formada. La posible formación parcial de especies desprotonadas o zwitteriónicas puede aumentar la nucleofilia en ciertas etapas, pero no modifica la conectividad general propuesta.

Mecanismo de síntesis de pirimidina amino benzimidazol con Zinc (L-prolinato)₂

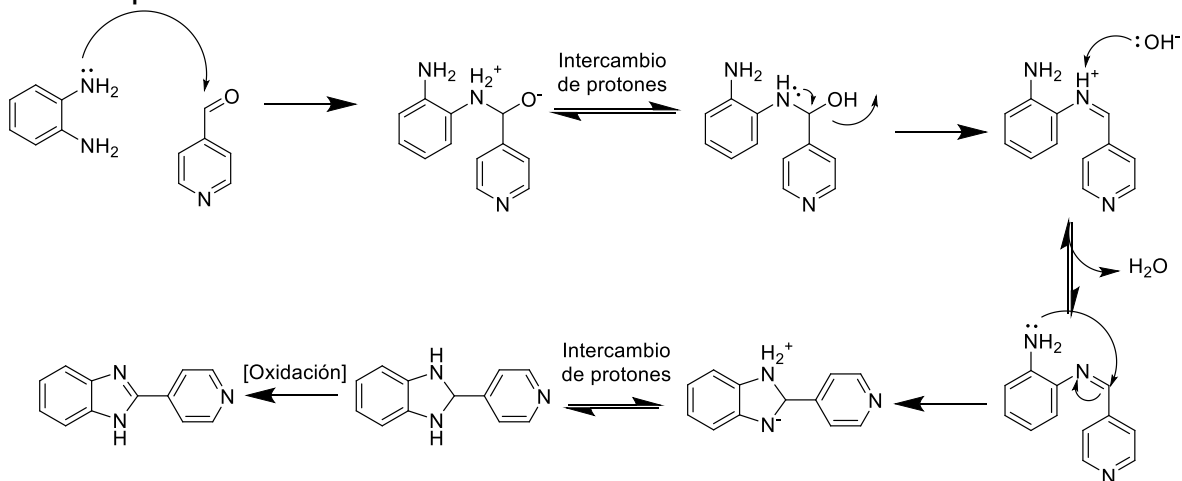
Adición 1,4



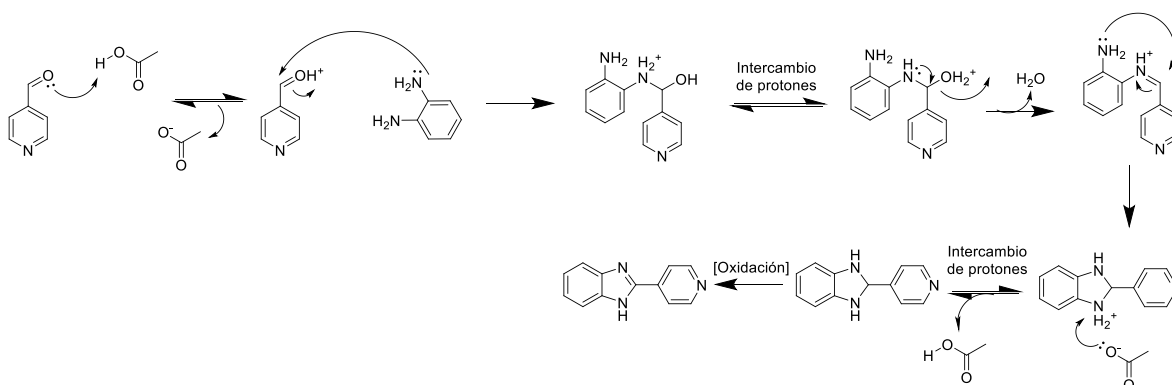
Ciclocondensación



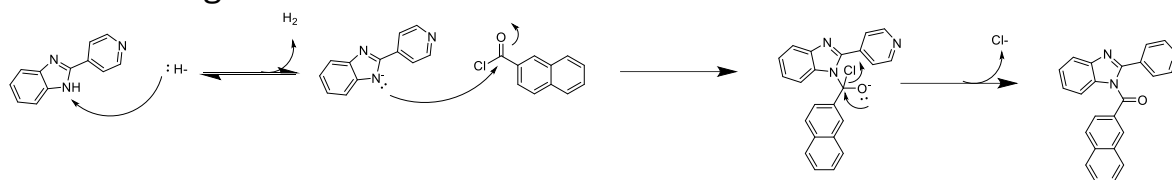
Mecanismo síntesis de benzimidazol, tomando como punto de partida el 4 piridina carbaldehído en medio neutro



Mecanismo síntesis de benzimidazol, tomando como punto de partida el 4 piridina carbaldehído en medio ácido



Mecanismo general de acilación de benzimidazol



Código de algoritmo genético QSAR evolutivo

El archivo de entrada debe estar en formato Excel (.xlsx), organizado en forma tabular, con una molécula por fila. La primera columna debe corresponder a un identificador (por ejemplo, ID), las siguientes 1444 columnas deben contener los descriptores numéricos, y debe existir una columna final llamada exactamente pIC50. La primera fila debe contener los nombres de las columnas y no debe haber celdas vacías en los descriptores ni en pIC50.

```
-*- coding: utf-8 -*-
```

```
"""
```

```
Created on Sat Dec 30 19:50:33 2023
```

```
@author: Gianfranco_S
```

```
"""
```

```
import pandas as pd
```

```
import numpy as np
```

```
x=datos.drop(["pIC50"],axis=1)
```

```
from sklearn.model_selection import train_test_split
```

```
from sklearn.linear_model import LinearRegression
```

```
from sklearn.metrics import r2_score
```

```
from time import time
```

```
start_time=time()
```

```
#datos
```

```
datos=pd.read_excel(r"/home/supermicro2/Documentos/Gianfranco_S/Plpro/Entrenamiento_algoritmo/Entrenamiento_algoritmo/")
```

```
x=datos.drop(["plC50"],axis=1)
```

```
y=datos["plC50"]
```

```
#Generar lo cromosomas de cada individuo inicial
```

```
#Aqui tenemos que poner cuantos individuos hay inicialmente
```

```
#descriptores_iniciales = Son el numero de descriptores que  
queremos usar(En esta simulacion , no variaran)
```

```
#descriptores totales ha evaluar = Es el universo de descriptores  
que se evaluan , si es en paddle habitualmente +1400
```

```
def
```

```
Poblacion_inicial(num_individuos,descriptores_iniciales,descriptores_totales_ha_evaluar):
```

```
array=np.zeros((num_individuos,descriptores_totales_ha_evaluar),dtype=int)
```

```
for i in range(num_individuos):
```

```
    unos = np.random.choice(descriptores_totales_ha_evaluar,  
descriptores_iniciales, replace=False)
```

```
    array[i, unos] = 1
```

return array

```
"""def poblacion_inicial(num_individuals):
```

```
    Inicializa una población de individuos.
```

```
    return np.random.randint(2, size=(num_individuals, 1444))"""
```

```
#Poblacion = debe ser definida como Poblacion_inicial(argumentos)
```

```
def gen(Poblacion,numero_individuo):
```

```
    return Poblacion[numero_individuo]
```

```
#gen=genotipo de la poblacion
```

```
#mutaciones
```

```
def
```

```
mutaciones(Poblacion,numero_individuo,descriptores_totales_ha_e  
valuar):
```

```
    i=np.random.randint(1,101)
```

```
    #5% de mutar sustitucion
```

```
    if i>=95:
```

```
Posicion_mutacion_s=np.random.randint(0,descriptores_totales_ha  
_evaluar)
```

```
mut=gen(Poblacion,numero_individuo)[Posicion_mutacion_s]
```

```
Descriptores_presentes=np.where(gen(Poblacion,numero_individuo)  
==1)[0]
```

```
if mut==0:
```

```
    mut=1
```

```
    gen(Poblacion,numero_individuo)[Posicion_mutacion_s] = mut
```

```
    #para concervar 14 descriptores presentes, se borra otro  
    descriptor
```

```
    gen_cambiar=np.random.choice(Descriptores_presentes)
```

```
    gen(Poblacion,numero_individuo)[gen_cambiar]=0
```

```
else:
```

```
    mut=0
```

```
    gen(Poblacion,numero_individuo)[Posicion_mutacion_s] =mut
```

```
    gen_cambiar=np.random.choice(Descriptores_presentes)
```

```
    gen(Poblacion,numero_individuo)[gen_cambiar]=1
```

```
#5% Crossover
```

```
p=np.random.randint(1,101)
```

```
if p>=95:
```

```
Posicion_crossover=np.random.randint(descriptores_totales_ha_ev  
aluar)
```



```
fragmento1=gen(Poblacion,numero_individuo)[:Posicion_crossover]
```

```
fragmento2=gen(Poblacion,numero_individuo)[Posicion_crossover:]
```

```
#se elige un individuo aleatorio
```

```
individuoX=gen(Poblacion, np.random.choice(len(Poblacion)))
```

```
fragmento3=individuoX[:Posicion_crossover]
```

```
fragmento4=individuoX[Posicion_crossover:]
```

```
Hijo_1=np.append(fragmento1,fragmento4)
```

```
Hijo_2=np.append(fragmento3,fragmento2)
```

```
#para mentener la cantidad de descriptores se restringe aquellos  
crossovers que generen + de 14 descriptores
```

```
if (Hijo_1.sum() <= Poblacion[0].sum()) or (Hijo_2.sum() <=  
Poblacion[0].sum()):
```

```
Poblacion=np.append(Poblacion,[Hijo_1],axis=0)
```

```
Poblacion=np.append(Poblacion,[Hijo_2],axis=0)
```

```
#reproduccion
```

```
def
```

```
reproduccion(Poblacion,num_individuos,descriptores_totales_ha_ev  
aluar):
```

```
p=len(Poblacion)
```

```

for i in range(0,len(Poblacion)):
    mutaciones(Poblacion,i,descriptores_totales_ha_evaluar)
while p <= num_individuos:
    #Clonacion
    indice_clonacion=np.random.randint(len(Poblacion))
    Poblacion=np.append(Poblacion,[Poblacion[indice_clonacion]],
axis=0)
    p=len(Poblacion)
return Poblacion

#generacion
def
simulacion(num_individuos,descriptores_iniciales,descriptores_totales_ha_evaluar,generaciones):
    Poblacion=Poblacion_inicial(num_individuos,
descriptores_iniciales, descriptores_totales_ha_evaluar)
    Base_genetica=[]
    R2=[]
    for i in range(1,generaciones+1) :
        #Fitnes de la nueva poblacion
        fitness=[]
        for i in range(0,len(Poblacion)):

```

```
model = LinearRegression()
```

```
x_subset=x[x.columns[1:][Poblacion[i]==1]]
```

```
model.fit(x_subset, y)
```

```
# Realizar predicciones en el conjunto de entrenamiento
```

```
predictions = model.predict(x_subset)
```

```
# Calcular el R cuadrado
```

```
r2 = r2_score(y, predictions)
```

```
# Guardar la combinación y su R cuadrado en las listas  
correspondientes
```

```
fitness.append(r2)
```

```
#para analisis más al grano , no es necesario guardar todo, para  
eso eliminar.append y poner R2=fitness y base genetica  
=Poblacion.copy()
```

```
R2.append(fitness)
```

```
Base_genetica.append(Poblacion.copy())
```

```
#Selección
```

```
posicion_Mejores= np.where(fitness >= np.percentile(fitness,  
30))[0]
```

```
Mejores=Poblacion[posicion_Mejores]
```

Poblacion=Mejores

Poblacion=reproduccion(Poblacion,num_individuos,descriptores_tales_ha_evaluar)

if i % 100 == 0:

np.save(f'Base_genetica_{i}.npy', Base_genetica[i])

np.save(f'R2_{i}.npy', R2[i])

Limpiar memoria

del Base_genetica

del R2

Base_genetica = []

R2 = []

#el resultado no da R2 para la generacion 1 dado que es inicial

#para leer

genoma,R2=simulacion(100, 10, 1444, 4000)

diferencia=time()-start_time

print(diferencia)

""" #para cargar =

```
#en este caso solo queria la generacion 220000
```

```
a=np.load("R2_22000.npz")
```

```
b=np.load("Base_genetica_22000.npz")
```

```
x_set=a["arr_1999"]
```

```
p=b["Base_genetica_1999"]
```

```
model = LinearRegression()
```

```
x_test=x[x.columns[1:][p[30]==1]]
```

```
model.fit(x_test, y)
```

```
# Realizar predicciones en el conjunto de entrenamiento
```

```
prediction = model.predict(x_test)
```

```
# Calcular el R cuadrado
```

```
r2 = r2_score(y, prediction)
```

```
coeficientes=list(np.append(model.coef_,model.intercept_))
```

```
descriptores=list(x_test.columns)
```

```
r2,coeficientes,descriptores
```

```
"""
```

```
#Para varias generaciones
```

```
from matplotlib

datos_R2 = []

# Cargar todos los archivos R2_100.npz a R2_500000.npz
for i in range(100, 500001, 100):

    nombre_archivo = f"R2_{i}.npz"

    datos = np.load(nombre_archivo)

    datos_R2.append(np.max(datos)) # Reemplaza 'arr_0' con la
    clave correcta si es diferente

# Convertir la lista de datos en un arreglo numpy
datos_R2 = np.array(datos_R2)
```