



UNIVERSIDAD AUTÓNOMA DEL ESTADO DE HIDALGO

INSTITUTO DE CIENCIAS BÁSICAS E INGENIERÍA

LICENCIATURA EN QUÍMICA

TESIS

**Desarrollo de modelos QSAR de compuestos inhibidores de
alfa-glucosidasa usando herramientas computacionales**

**Para obtener el título de
Licenciado en Química**

Presenta

Alejandro Gómez Rangel

Director

Dr. Julián Cruz Borbolla

Codirector

Dr. Luis Ángel Zarate Hernández

Mineral de la Reforma, Hgo., México, agosto de 2026



Universidad Autónoma del Estado de Hidalgo

Instituto de Ciencias Básicas e Ingeniería

School of Engineering and Basic Sciences

Dirección

Office of the Director

Mineral de la Reforma, Hgo., a 27 de agosto de 2026

Número de control: ICBI-D/1223/2026

Asunto: Autorización de impresión.

DRA. OJUKY DEL ROCÍO ISLAS MALDONADO
DIRECTORA DE ADMINISTRACIÓN ESCOLAR DE LA UAEH

Con Título Quinto, Capítulo II, Capítulo V, Artículo 51 Fracción IX del Estatuto General de nuestra Institución, por este medio, le comunico que el Jurado asignado al egresado de la Licenciatura en Química **Alejandro Gómez Rangel**, quien presenta el trabajo de titulación **"Desarrollo de modelos QSAR de compuestos inhibidores de alfa-glucosidasa usando herramientas computacionales"**, ha decidido, después de revisar fundamento en lo dispuesto en el Título Tercero, Capítulo I, Artículo 18 Fracción IV; dicho trabajo en la reunión de sinodales, **autorizar la impresión del mismo**, una vez realizadas las correcciones acordadas.

A continuación, firman de conformidad los integrantes del Jurado:

Presidente: Dra. Rosa Luz Camacho Mendoza

Secretario: Dr. Carlos Zepactonal Gómez Castro

Vocal: Dr. Julián Cruz Borbolla

Suplente: Dr. Luis Ángel Zárate Hernández

Sin otro particular por el momento, reciba un cordial saludo.

Atentamente
"Amor, Orden y Progreso"

Dr. Gabriel Vergara Rodríguez
Director del ICBI

GVR/MMM

"Amor, Orden y Progreso"



Ciudad del Conocimiento, Carretera Pachuca-Tulancingo
Km. 4.5 Colonia Carboneras, Mineral de la Reforma, Hidalgo,
México. C.P 42184
Teléfono: 52 (771) 71 720 00 Ext. 40001
direccion_icbi@uaeh.edu.mx, vergara@uaeh.edu.mx

uaeh.edu.mx

Índice

Índice de Figuras	5
Índice de Tablas	5
Índice de Gráficas	6
Agradecimientos	7
Abreviaturas y Símbolos	8
Resumen	11
Abstract.....	11
1. Introducción	12
Objetivo General.....	18
Objetivos Específicos.....	18
Justificación	18
2. Metodología.	20
2.1. Selección y Preoptimización de los Compuestos Inhibidores	20
2.2. <i>Docking</i>	21
2.2.1. Preparación de la Enzima	21
2.2.2. Preparación de los Inhibidores.....	21
2.2.3. Mapas de Afinidad Atómica	21
2.2.4. Generación de Complejos Enzima-Inhibidor	22
2.3. <i>Clustering</i> Mediante <i>K-means</i>	22
2.4. Cálculos DFT	23
Cálculo de Descriptores de Dragon y RDKit	24
2.5. QSAR	25
3. Resultados	27
3.1. Estructuras y Propiedades Experimentales de los Inhibidores Seleccionados.	27
3.2. Dinámicas Moleculares	35
3.3. Resultados de <i>Docking</i>	37
3.3.1. Complejos Enzima-Inhibidor	38
3.4. Modelos QSAR	44

3.4.1.	Descriptores Estructurales del Modelo Obtenido	48
3.4.2.	Resultados Estadísticos del Modelo	48
Conclusiones		63
Referencias.....		65

Índice de Figuras

Figura 1. Representación esquemática de la reacción reversible catalizada por ambos dominios de las enzimas humanas, imagen de creación propia, inspirado en [10].	13
Figura 2. Compuestos provenientes de diferentes derivados	32
Figura 3. Estructura cristalina de la subunidad C-terminal de la maltasa-glucoamilasa humana en complejo con acarbosa (3TOP). Imagen recuperada del PDB.	37
Figura 4. Diagrama 3D y 2D de las interacciones moleculares del <i>Docking</i> de M101 sobre la enzima α -glucosidasa (3TOP).	52
Figura 5. Mapa del potencial electroestático de M101	53
Figura 6. Diagrama 3D y 2D de las interacciones moleculares del Molecular <i>Docking</i> de M118 sobre la enzima α -glucosidasa (3TOP).	56
Figura 7. Mapa del potencial electroestático de M118	57

Índice de Tablas

Tabla 1 Base de datos de los compuestos estudiados en el presente trabajo, con la información experimental reportada en la literatura.	28
Tabla 2. Intervalos donde se encuentra la mayor cantidad de complejos por cada clúster.	39
Tabla 3. <i>Complejos seleccionados a partir de los clústers que cumplían mejor con los 3 intervalos propuestos (distancia a la referencia (Å), ΔG (kcal/mol) y pIC_{50})</i>	42
Tabla 4. Modelos QSAR obtenidos.	45
Tabla 5. Descriptores resultantes del Modelo 4 para cada compuesto con su respectivo valor de pIC_{50} .	46

Índice de Gráficas

Gráfica 1. Trayectoria de la dinámica molecular para M011 , destacando el punto en rojo, siendo el confórmero final.	35
Gráfica 2. Distribución de los <i>frames</i> en 3 dimensiones de cada complejo generado respecto a los valores de pIC_{50} , ΔG (kcal/mol) y la distancia de un átomo del ligante que se encuentra más cercano a la referencia co-cristalizada (Å), agrupado por clústers.	38
Gráfica 3. Distribución en 2 dimensiones de cada complejo generado respecto a los valores de ΔG (kcal/mol) y la distancia de un átomo del ligante que se encuentra más cercano a la referencia co-cristalizada (Å), agrupado por clústers.	40
Gráfica 4. Distribución en 2 dimensiones de cada complejo generado respecto a los valores de pIC_{50} y la distancia de un átomo del ligante que se encuentra más cercano a la referencia co-cristalizada (Å), agrupado por clústers.	40
Gráfica 5. Distribución en 2 dimensiones de cada complejo generado respecto a los valores de ΔG (kcal/mol) y pIC_{50} , agrupado por clústers.	41
Gráfica 6. Modelo QSAR con pIC_{50} calculada vs. pIC_{50} experimental, donde los rombos amarillos pertenecen al conjunto de entrenamiento y los cuadrados azules al conjunto de predicción.....	50
Gráfica 7. Modelo QSAR de pIC_{50} calculada (LOO) vs pIC_{50} experimental.	59
Gráfica 8. Gráfico de Williams del modelo QSAR pIC_{50} donde las líneas punteadas horizontales representan los límites al 98.76% (-3.0 y 3.0), la línea punteada vertical corresponde al valor de h^* , los rombos amarillos pertenecen al conjunto de entrenamiento y los cuadrados azules al conjunto de predicción.	61

Agradecimientos

A mi madre, por todo su amor, por apoyarme y creer incondicionalmente en mí. Sobre todo, por no dejarme solo nunca y ser la principal guía en mi vida.

A mi hermano, por el amor y apoyo que me ha dado en todo momento. Por ser mi mejor amigo y por su infinita comprensión, incluso en mis momentos más difíciles.

A mi padre, por las grandes lecciones de vida que me ha enseñado.

A mi familia (mi abue, mis tíos y mi prima), por ser mi motor principal. Gracias por su apoyo incondicional, por creer en mí incluso en los días más pesados y por darme el aliento necesario para nunca rendirme.

A mi novia y futura esposa, por el profundo amor y respaldo que me ha brindado. Por creer en mí, motivarme a seguir adelante y acompañarme incondicionalmente ante cualquier adversidad.

A mis mascotas, Daysi y Suki, quienes me acompañaron durante toda la licenciatura, en las largas noches de desvelo haciendo tarea y durante los tiempos difíciles de la pandemia. Siempre las amaré y las llevaré en mis recuerdos por el resto de mi vida.

A mis amigos, "los bobes", con quienes compartí innumerables experiencias durante estos años. Gracias por hacer este camino más llevadero y por demostrarme el inmenso valor que tiene una verdadera amistad.

A mis asesores, Dr. Julián y Dr. Luis Ángel, por su infinita paciencia, sus valiosos consejos y por guiarme a lo largo de este proyecto. Gracias por compartir su experiencia y orientarme en los retos que implicó la investigación. Su rigor y apoyo fueron clave para llevar a buen término esta tesis.

Al Dr. Carlos, por su disposición, paciencia y valiosa ayuda en el desarrollo de los cálculos de *Docking*, importantes para el desarrollo y conclusión de este proyecto.

A la Dra. Rosa Luz, por su enorme paciencia, apoyo constante y guía en la revisión y culminación de este proyecto.

Y finalmente a SECIHTI por el apoyo con la beca de Ayudantía del SNII III, que fue muy importante para la culminación de la presente tesis.

Abreviaturas y Símbolos

2D	Dos dimensiones
3D	Tres dimensiones
3D-QSAR	Relación Cuantitativa Estructura-Actividad tridimensional
3ob	Conjunto de parámetros empleado en DFTB+
3TOP	Código PDB de la estructura cristalográfica utilizada
6-31G**	Conjunto de base usado en los cálculos DFT
A	Afinidad electrónica
AEA	Matriz de adyacencia de aristas aumentada
AEA(w)	Matriz de adyacencia de aristas aumentada ponderada
AGL	Algoritmo Genético Lamarckiano
AlogP	Coefficiente de reparto octanol-agua de Ghose-Crippen
ATAC	Autocorrelación de tipo atómico
Å	Ångström
B08[C–S]	Descriptor binario de par C–S a distancia topológica 8
B09[C–O]	Descriptor binario de par C–O a distancia topológica 9
B3LYP	Becke, 3-parámetros, Lee-Yang-Parr
BalabanJ	Índice topológico de Balaban
CACTVS	Servidor/herramienta para generación de estructuras químicas en 3D
CoMFA	<i>Comparative Molecular Field Analysis</i> /Análisis Comparativo de Campos Moleculares
CoMSIA	<i>Comparative Molecular Similarity Indices Analysis</i> /Análisis Comparativo de Índices de Similitud Molecular
DA	Dominio de aplicabilidad
DFT	<i>Density Functional Theory</i> /Teoría del Funcional de la Densidad
DFTB+	Programa basado en <i>Density Functional Tight-Binding</i> ; método de enlace fuerte basado en la Teoría del Funcional de la Densidad
DM	<i>Distance Matrix</i> /Matriz de distancias
Ds	Índice de accesibilidad total ponderado por el estado intrínseco
EA	Matriz de adyacencia de aristas
EA(w)	Matriz de adyacencia de aristas ponderada
EC	<i>Enzyme Commission</i> /Clasificación enzimática
E_{HOMO}	Energía del orbital HOMO
E_{LUMO}	Energía del orbital LUMO
ENSANUT	Encuesta Nacional de Salud y Nutrición
F07[C–N]	Descriptor de frecuencia de par C–N a distancia topológica 7
FID	Federación Internacional de Diabetes
fs	Femtosegundo
GA	<i>Genetic Algorithms</i> /Algoritmos Genéticos
GA-MLR	<i>Genetic Algorithms-Multiple Linear Regression</i> /Algoritmos genéticos-regresión lineal múltiple
GGA	<i>Generalized Gradient Approximation</i> /Aproximación del Gradiente Generalizado

<i>Hat / leverage</i>	Medida de influencia estructural de una observación en el modelo
HOMO	<i>Highest-energy Occupied Molecular Orbital</i> /Orbital molecular ocupado de mayor energía
h^*	Valor crítico de <i>leverage</i>
<i>I</i>	Potencial de ionización
IAGs	Inhibidores de α -glucosidasa
IC_{50}	Concentración inhibitoria del 50 %
K	Kelvin
kcal/mol	Kilocalorías por mol
K_i	Constante de inhibición
k	Pendiente de la regresión entre valores experimentales y predichos
k'	Pendiente de la regresión inversa entre valores predichos y experimentales
LMO	<i>Leave-Many-Out</i> ; validación cruzada por exclusión de múltiples observaciones
LOO	<i>Leave-One-Out</i> ; validación cruzada por exclusión de una observación
LRCD	Parámetro de similitud entre complejo calculado y complejo cristalográfico
LUMO	<i>Lowest-energy Unoccupied Molecular Orbital</i> /Orbital molecular desocupado de menor energía
MGAM	Maltasa-glucoamilasa
MGAM-C	Dominio C-terminal de la maltasa-glucoamilasa humana
MLR	<i>Multiple Linear Regression</i> /Regresión lineal múltiple
MPE	Mapeo del Potencial Electroestático
μ M	Micromolar
nAT	Número de átomos
nBO	Número de enlaces
nDB	Número de dobles enlaces
NaaO	Número de átomos de oxígeno en un anillo aromático
PCA	Análisis de Componentes Principales
PDB	<i>Protein Data Bank</i> /Banco de Datos de Proteínas
PDB ID	Identificador de estructura en el <i>Protein Data Bank</i>
pIC_{50}	Logaritmo negativo del valor de IC_{50} expresado en μ M
QSAR	<i>Quantitative Structure-Activity Relationship</i> /Relación Cuantitativa Estructura-Actividad
QSAR-INS	Software para desarrollo, análisis y validación de modelos QSAR
RDF	Función de Distribución Radial
RDF020i	Descriptor RDF a 2.0 Å ponderado por potencial de ionización
RDKit	Biblioteca/software de química computacional
RMSD	<i>Root Mean Square Deviation</i> /Desviación cuadrática media
R^2	Coeficiente de determinación
R^2_{adj}	Coeficiente de determinación ajustado

R_{pred}^2	Coeficiente de determinación predictivo
Q_{LOO}^2	Coeficiente de determinación de validación cruzada Leave-One-Out
Q_{LMO}^2	Coeficiente de determinación de validación cruzada Leave-Many-Out
Q_{F2}^2	Coeficiente de validación predictiva externa F2
r_m^2	Métrica de validación basada en la correlación entre valores observados y predichos
$r_m'^2$	Métrica r_m^2 calculada con los ejes invertidos
$\overline{r_m^2}$	Promedio de r_m^2 y $r_m'^2$
Δr_m^2	Diferencia absoluta entre r_m^2 y $r_m'^2$
r_0^2	Coeficiente de determinación para la regresión a través del origen
$r_0'^2$	Coeficiente de determinación para la regresión inversa a través del origen
SECIHTI	Secretaría de Ciencia, Humanidades, Tecnología e Innovación
SI	Sucrasa-isomaltasa
SlogP_VSA4	Descriptor VSA basado en contribuciones atómicas a logP
SM15_EA(bo)	Momento espectral de orden 15 de la matriz de adyacencia de aristas ponderada por orden de enlace
SMILES	Simplified Molecular Input Line Entry System
SNII	Sistema Nacional de Investigadoras e Investigadores
VSA	Área de superficie de Van der Waals
WHIM	<i>Weighted Holistic Invariant Molecular</i>
ΔG	Energía libre de unión
η	Dureza química
μ	Potencial químico
χ	Electronegatividad
S	Blandura química
ω	Electrofilicidad
ω^+	Poder electroaceptor
ω^-	Poder electrodonador
σ	Desviación estándar

Resumen

La diabetes mellitus es una enfermedad que afecta millones de personas en el mundo, desafortunadamente aún no se conoce una cura para este padecimiento. Hoy en día existen medicamentos que ayudan en el tratamiento de esta afección, un grupo de fármacos llamados inhibidores y su principal función es bloquear la actividad biológica de enzimas como la α -glucosidasa, blanco importante en el tratamiento de esta enfermedad. En el presente trabajo se desarrolló un modelo QSAR capaz de predecir la actividad inhibitoria a partir de 74 compuestos de diferentes derivados químicos, filtrándolos mediante *Docking* y posteriormente calculando descriptores topológicos y de reactividad electrónica global mediante DFT (B3LYP/6-31G**), los parámetros estructurales y electrónicos se correlacionaron con la actividad de inhibición IC_{50} mediante un modelo de regresión lineal múltiple. El modelo óptimo se estructuró con 10 descriptores y demostró una capacidad predictiva robusta y confiable, siendo validado interna y externamente con valores de $R^2 = 0.747$, $Q_{Loo}^2 = 0.621$ y $R_{pred}^2 = 0.572$. Los resultados se complementaron, analizando las interacciones de los inhibidores provenientes del *Docking* y los mapas de potencial electrostático resultantes de DFT los cuales demostraron que la integración de técnicas computacionales permite generar modelos que pueden predecir la actividad inhibitoria de diferentes compuestos.

Abstract

*Diabetes mellitus affects millions of people worldwide, and unfortunately, no cure for the condition is currently known. Medications are available to treat this ailment, including a group of drugs known as inhibitors; their primary function is to block the biological activity of enzymes such as α -glucosidase, a key target in treating the disease. In this study, a QSAR model capable of predicting inhibitory activity was developed from 74 compounds representing various chemical derivatives. These compounds were filtered using molecular docking, followed by the calculation of topological and global electronic reactivity descriptors via DFT (B3LYP/6-31G**). Structural and electronic parameters were correlated with IC_{50} inhibition activity using a multiple linear regression model. The optimal model, comprising 10 descriptors, demonstrated robust and reliable predictive capability, undergoing both internal and external validation with values of $R^2 = 0.747$, $Q_{Loo}^2 = 0.621$ y $R_{pred}^2 = 0.572$. The results were further supported by an analysis of inhibitor interactions derived from docking and electrostatic potential maps obtained via DFT, demonstrating that integrating computational techniques enables the generation of models capable of predicting the inhibitory activity of various compounds.*

1. Introducción

La diabetes mellitus, es una enfermedad crónica que se caracteriza por la persistencia de hiperglucemia como síntoma principal. Aquellas personas que la padecen no son capaces de producir o utilizar de manera correcta la insulina que proporciona su cuerpo, por consiguiente, el contenido de glucosa en sangre se eleva, ya que la insulina funciona como regulador de este monosacárido [1]. De acuerdo con el Atlas de la Diabetes por la Federación Internacional de Diabetes (FID), en 2024, aproximadamente 589 millones de personas de entre 20 y 79 años padecen diabetes mellitus diagnosticada y sin diagnosticar. Se estima que este número aumente hasta los 852.5 millones en 2050 [2]. La Encuesta Nacional de Salud y Nutrición (ENSANUT) señaló que, en el año de 2022, 14.6 millones de personas padecen diabetes mellitus en México [3]. Existen dos tipos principales de diabetes mellitus, el tipo 1 y el tipo 2. El tipo 1 es más recurrente en niños y adultos jóvenes y se asocia principalmente a la falta de producción de insulina en el cuerpo. Por otra parte en la diabetes mellitus tipo 2 es la forma más común representando entre el 90% y el 95% de los casos de diabetes mellitus y se caracteriza principalmente por la resistencia a la insulina, condición en la cual las células no responden de manera adecuada a la acción de esta hormona. Como consecuencia, el páncreas incrementa la producción de insulina con el propósito de compensar dicha deficiencia, sin embargo, con el paso del tiempo, las células beta pancreáticas pueden perder progresivamente su capacidad funcional, resultando en una producción insuficiente de insulina [2].

El presente trabajo se relaciona directamente con la diabetes mellitus tipo 2, padecimiento para el cual actualmente no existe una cura definitiva, sin embargo existen tratamientos auxiliares que controlan la enfermedad. Parte de estos tratamientos incluyen el consumo de fármacos capaces de regular los niveles de glucosa posprandial, conocidos como inhibidores de la α -glucosidasa (IAGs), los cuales inhiben enzimas encargadas de la degradación de carbohidratos complejos a glucosa [4]. La α -glucosidasa es un término utilizado para denominar un conjunto de enzimas pertenecientes a la familia de hidrolasas de glucósidos (EC 3.2.1.20) [5] cuya especificidad se dirige principalmente a la hidrólisis de enlaces (1 $\rightarrow$ 4)- α -glucosídicos en maltooligosacaridos, liberando α -D-glucosa. La α -glucosidasa participa en diferentes funciones biológicas esenciales, entre ellas la digestión de carbohidratos provenientes de la dieta, el catabolismo lisosomal de glicoconjugados y el procesamiento postraducciona l de glicoproteínas [6]. En mamíferos, estas enzimas desempeñan un papel fundamental en la digestión intestinal de carbohidratos, particularmente en el borde en cepillo de los enterocitos, donde catalizan la hidrólisis de enlaces α -glucosídicos presentes en disacáridos y oligosacáridos y depende de complejos con actividad α -glucosidasa, principalmente

sucrasa–isomaltasa (SI) y maltasa–glucoamilasa (MGAM); estos complejos, en conjunto con la acción previa de las α -amilasas, convierten oligosacáridos en monosacáridos y determinan en gran medida la velocidad de liberación de glucosa tras una comida [6], [7].

Respecto del mecanismo este se identifica del tipo catalítico debido a que puede describirse como un proceso de dos etapas principales, como se muestra en la Figura 1, en una primera etapa, el sustrato se aproxima al sitio activo, que es la región específica en la que se une el sustrato y donde ocurre la reacción, esta zona favorece la transformación de los sustratos en productos al disminuir la energía de activación de la reacción [8], orientando el enlace glucosídico entre los subsitios -1 y +1 que corresponden a las regiones de la enzima donde ocurre la ruptura del enlace glucosídico [9], posteriormente, un residuo con actividad ácido/base protona el oxígeno glucosídico, facilitando la liberación de la glucosa del subsitio +1, mientras que un residuo nucleofílico ataca el carbono anomérico obteniendo un intermediario covalente glucosil-enzima. En la segunda etapa, el residuo desprotonado actúa activando una molécula de agua que ataca el carbono anomérico del intermediario covalente. Este ataque rompe el enlace entre la glucosa y el residuo nucleofílico de la enzima, regenerando el sitio activo y liberando una segunda glucosa como producto final [10].

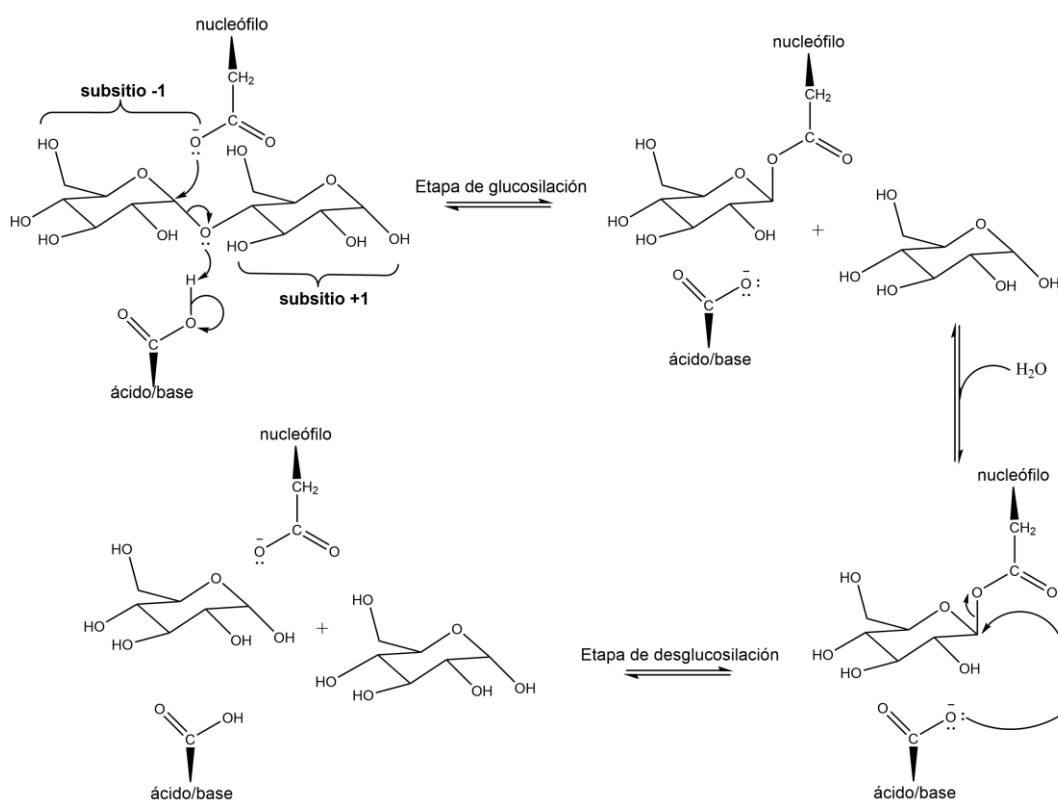


Figura 1. Representación esquemática de la reacción reversible catalizada por ambos dominios de las enzimas humanas, imagen de creación propia, inspirado en [10].

Por lo anterior, la α -glucosidasa es un blanco farmacológico importante para el control de hiperglucemia posprandial ya que al catalizar la ruptura de los enlaces glucosídicos, favorece la liberación intestinal de glucosa y posteriormente puede ser absorbida contribuyendo al incremento de su concentración en sangre, lo que puede provocar daño en los riñones, ceguera o aumentar el riesgo cardíaco en personas con diabetes. En este sentido, los IAGs son compuestos capaces de disminuir o bloquear la actividad catalítica de la enzima al unirse a ella y alterar temporal o permanentemente su capacidad de transformar sustratos en productos [11], actúan predominantemente de forma competitiva, retardando la digestión y absorción de carbohidratos, y con ello, atenuando los picos posprandiales de glucosa [12].

Entre los IAGs de uso clínico se encuentran acarbosa [13], voglibosa [14], y miglitol [15] los cuales actúan de manera independiente de la insulina al inhibir la actividad de las α -glucosidasas intestinales y disminuir la velocidad de digestión de los carbohidratos. Sin embargo, su uso puede asociarse con efectos adversos, principalmente gastrointestinales, derivados de la presencia y fermentación de carbohidratos no digeridos en el colon. Entre las manifestaciones más frecuentes se encuentran flatulencia, distensión abdominal, diarrea, dolor abdominal y náuseas. [12]. De las medidas empleadas para cuantificar la actividad inhibitoria de un compuesto IAGs es mediante el IC_{50} (*half maximal inhibitory concentration*) definida como la concentración del compuesto que produce un 50% de inhibición de una respuesta biológica o de la actividad de una enzima bajo condiciones experimentales definidas, en comparación con un control sin inhibidor [16].

En la evaluación y diseño de nuevos compuestos IAGs, las metodologías computacionales han adquirido un papel relevante, debido a que permiten relacionar la estructura molecular con la actividad biológica, así como proponer modificaciones estructurales antes de realizar ensayos experimentales. Entre estas aproximaciones destacan los modelos de Relación Cuantitativa Estructura-Actividad (QSAR, por sus siglas en inglés, *Quantitative Structure-Activity Relationship*), que buscan establecer una evaluación estadística entre un conjunto de descriptores moleculares y la actividad biológica. Una de las aproximaciones más robustas son los modelos de Relación Cuantitativa Estructura-Actividad en tres dimensiones (3D-QSAR, por sus siglas en inglés, *Three-Dimensional Quantitative Structure-Activity Relationship*). En este contexto, Zheng et al. (2017) [17] utilizaron de manera inicial una metodología de Acoplamiento Molecular (*Molecular Docking* o solo *Docking*, en inglés) para obtener las conformaciones de los ligandos con la enzima α -glucosidasas, donde el estudio reveló que las xantonas más potentes logran su alta afinidad no solo por interacciones de carga y de puentes de hidrógeno con los residuos catalíticos, sino por valiosas interacciones hidrofóbicas y un apilamiento π - π . Posteriormente estas conformaciones fueron refinadas con dinámica

molecular y con ellas emplearon los métodos CoMFA (*Comparative Molecular Field Analysis*) y CoMSIA (*Comparative Molecular Similarity Indices Analysis*) para estudiar las 54 xantonas. Concluyendo que el aumento del número de anillos aromáticos así como la adición de sustituyentes aceptores de enlaces de hidrógeno son relevantes para favorecer la bioactividad de las xantonas. A pesar de que presentan un modelo robusto con capacidad predictiva, podría complementarse con información proveniente de la Teoría del Funcional de la Densidad (DFT, por sus siglas en inglés, *Density Functional Theory*).

Otra estrategia complementaria para evaluar el potencial de los derivados como agentes antidiabéticos consiste en el desarrollo de modelos QSAR obtenidos mediante selección de variables con algoritmos genéticos (GA, por sus siglas en inglés, *Genetics Algorithms*), en esta línea, *Aminu et al.* (2022) [18] utilizaron GA para predecir la actividad de 1,4-dihidropiridinas, optimizando las estructuras mediante cálculos DFT (B3LYP/6-31G*) y a partir de las estructuras optimizadas se generaron parámetros estructurales y topológicos, con la información obtenida se generaron modelos que se utilizaron para predecir las actividades inhibitorias. El mejor modelo seleccionado y obtenido por los autores del trabajo se encontró con 5 descriptores topológicos: ATSC4v (Autocorrelación de Broto-Moreau centrada-retardo 4, ponderada por volúmenes de van der Waals), GATS5i (Autocorrelación de Geary-retardo 5, ponderada por el primer potencial de ionización), SpMAD_Dze (Desviación absoluta media espectral respecto a la matriz de Barysz, ponderada por las electronegatividades de Sanderson), VR2_Dze (Índice normalizado basado en vectores propios de tipo Randic a partir de la matriz de Barysz, ponderado por las electronegatividades de Sanderson) y SpMin5_Bhp (El valor propio absoluto más pequeño de la matriz modificada de Burden (orden de dimensión 5), ponderado por las polarizabilidades relativas de los átomos), las cuales se encontraron como variables relacionadas a la interacción de los derivados de 1,4-dihidropiridinas con la interacción del sitio activo del dominio C-terminal de la maltasa-glucoamilasa humana (MGAM-C, PDB: 3TOP). Este modelo permitió el diseño de cinco compuestos alternativos, utilizando la misma base estructural de la familia de compuestos 1,4-dihidropiridinas y se obtuvieron sus actividades teóricas mediante el modelo validado, encontrando valores cercanos de pIC_{50} a la acarbosa, posteriormente, se aplicó *Docking* para él estudió sobre las puntuaciones de unión de los compuestos alternativos contra el sitio activo mostrando buenas afinidades por este receptor y finalmente, el análisis de las propiedades de absorción, distribución, metabolismo, excreción y toxicidad (SWISSADME) de los compuestos diseñados, reveló que los compuestos poseen una buena absorción intestinal y potencial de eliminación total, además de no presentar toxicidad y mostrar un perfil adecuado respecto a su interacción con la enzima CYP3A4. En este trabajo los autores aplicaron una metodología donde utilizan DFT, *Docking* y QSAR, se

optimizaron las estructuras para posteriormente obtener descriptores topológicos y el *Docking* se utilizó como una respuesta cualitativa de cómo se unirían los compuestos diseñados con el sitio activo de la enzima, cabe resaltar que en este estudio la metodología DFT fue utilizada para la obtención de las geometrías iniciales para el cálculo de los descriptores topológicos.

A pesar de los avances significativos en la caracterización de inhibidores mediante estas técnicas, gran parte de la literatura actual tiende a emplear el *Docking*, los cálculos DFT y el modelado QSAR de manera secuencial pero no relacionada directamente. Sin embargo, en la mayoría de los casos esta información derivada de DFT y del *Docking* no se incorpora de manera directa como parte del conjunto inicial de variables del modelo QSAR. A pesar de que los enfoques de estas herramientas han demostrado su valor individual, el análisis de la literatura revela una constante en como estas se integran, en diferentes casos la información del *Docking* o de los cálculos DFT se utiliza principalmente como validación o explicación posterior del modelo QSAR o como herramientas que permiten identificar la reactividad desde el punto de vista de dos niveles de teoría muy distintos, y no necesariamente como parte directa del conjunto inicial de variables empleadas para construir el modelo predictivo.

Aunque algunos trabajos han comenzado a cerrar esta brecha, siendo el caso de *Mora et al.* (2018) [19], al evaluar derivados de ácidos cinámicos unidos a diferentes aminoácidos como IAGs, en este estudio, los autores utilizaron descriptores químicos cuánticos obtenidos mediante DFT, así como descriptores derivados de mapas algebraicos multilineales [20] los cuales fueron ponderados por propiedades atómicas como: coeficiente de reparto octanol-agua de Ghose-Crippen (AlogP), carga (c), electronegatividad (e), dureza (η), masa (m), polarizabilidad (p), área de superficie polar (psa), refractividad (r), blandura (s) y volumen de Van der Waals (v), para construir modelos QSAR mediante regresión lineal múltiple y selección de variables por algoritmos genéticos. Si bien el modelo más robusto obtenido con los descriptores derivados de mapas algebraicos: área de superficie polar (psa), el volumen de Van der Waals (v), la dureza (h) y las cargas (c), pueden ser considerados como variables dentro del modelado QSAR y no únicamente como elementos descriptivos independientes.

Es en este contexto que en el presente trabajo se propone un estudio fundamentado en el uso de herramientas de química computacional para el desarrollo de modelos QSAR que permitan predecir la inhibición de la enzima α -glucosidasa en compuestos similares a los ya reportados en la literatura. Se plantea utilizar el *Docking* como técnica de filtrado para discernir e identificar los compuestos con interacción directa en el sitio activo de la enzima estudiada [21], de manera complementaria se realizarán cálculos de estructura electrónica mediante DFT de

los cuales se obtendrá una estructura estable y que represente un mínimo en la superficie de energía potencial, así de estos mismos cálculos se obtienen algunas propiedades electrónicas globales [22] (potencial químico, dureza [23], poder electrodonador/aceptor [24], electrofilicidad [25], afinidad electrónica y potencial de ionización) que pueden servir para dos fines, por una parte analizar la reactividad química y por otra parte como parámetros en el entrenamiento de los modelos QSAR. La tabla de descriptores para el entrenamiento de los modelos QSAR se complementara también con descriptores topológicos calculados mediante los programas Dragon [26] y RDKit [27], que finalmente se utilizaron para la construcción de los modelos con la combinación de estrategias estadísticas como algoritmos genéticos para la selección óptima de variables y regresión lineal múltiple (GA-MLR) para el cálculo de los coeficientes de los modelos candidatos, de esta manera, así como parámetros estadísticos que permitieran validar la capacidad predictiva de los candidatos y así elegir las mejores opciones de modelos predictivos. Finalmente, este trabajo muestra que la información obtenida mediante distintas etapas que involucran *Docking*, DFT y los descriptores moleculares favorecen el desarrollo de modelos QSAR predictivos y aplicados, en este caso, a conjuntos de IAGs [28], [29].

Objetivo General

Desarrollar modelos QSAR que permitan predecir la actividad inhibitoria frente a la α -glucosidasa en compuestos pertenecientes a distintas familias químicas con valores de pIC_{50} reportados en la literatura y seleccionados a partir de similitudes en las propiedades obtenidas mediante *Docking*, a fin de predecir la actividad inhibitoria de candidatos a fármacos en el tratamiento de diabetes mellitus tipo 2.

Objetivos Específicos

- Construir una base de datos de compuestos IAGs a partir de trabajos reportados en la literatura, aplicando diferentes criterios de filtrado para tener calidad en los datos.
- Pre optimizar las geometrías de la base de datos mediante dinámica molecular y realizar estudios de *Docking* entre los compuestos seleccionados y un modelo cristalográfico de la enzima α -glucosidasa con el fin de determinar interacciones, energías de unión y distancias cercanas respecto al inhibidor de referencia que se encuentre co-cristalizado.
- Calcular descriptores topológicos, propiedades electrónicas derivadas de DFT, para desarrollar modelos QSAR con validez estadística y analizar la contribución de los descriptores más significativos e identificar características estructurales y electrónicas asociadas a la actividad inhibitoria.

Justificación

Recientemente, el interés por desarrollar nuevos IAGs ha incrementado, en particular aquellos que no sean derivados de carbohidratos y que puedan ser tanto naturales como sintéticos con el objetivo de mejorar la selectividad y reducir los efectos adversos provocados por los inhibidores clínicos. En la búsqueda de compuestos novedosos, el uso de técnicas computacionales ha tomado un papel importante debido a su capacidad para disminuir costos y tiempos experimentales, ya que podrían facilitar el diseño racional y el desarrollo de candidatos con actividad inhibitoria mejorada y menor incidencia de efectos gastrointestinales, útiles para el tratamiento de pacientes con diabetes. Entre estas herramientas destacan los modelos QSAR que describen la relación matemática entre las propiedades estructurales de un compuesto y su actividad biológica experimental, estos han evolucionado desde correlaciones lineales simples hasta sofisticadas aproximaciones 3D y métodos híbridos que incorporan descriptores fisicoquímicos, topológicos y cuánticos, permitiendo predecir la actividad biológica y priorizar compuestos antes de su síntesis [30]. Sin embargo, su eficacia depende en gran medida de la calidad de los descriptores empleados y de la correcta representación

de la interacción fármaco–blanco. En este sentido, el *Docking* constituye una técnica complementaria, ya que permite modelar la orientación y el modo de unión de los ligandos en el sitio activo de la proteína, estimar afinidades de interacción y caracterizar interacciones clave [31]. Por su parte, los cálculos basados en DFT se han consolidado como el método más empleado para describir la estructura electrónica, proporcionando parámetros como energías HOMO-LUMO, parámetros de reactividad tanto globales como locales y mapeos de potencial electroestático, esenciales para interpretar la reactividad y generar descriptores confiables para QSAR [32].

2. Metodología.

2.1. Selección y Preoptimización de los Compuestos Inhibidores

Inicialmente, realizar una revisión a partir de información publicada en artículos recientes con el propósito de identificar compuestos reportados como IAGs y recopilar sus valores experimentales de IC_{50} . Posteriormente construir una base de datos inicial, considerando únicamente compuestos orgánicos, no iónicos, que no contengan en su estructura átomos pesados y excluyendo valores de IC_{50} censurados [33], asignar una etiqueta que permite identificar a cada molécula y para homogenizar la información experimental, expresar todos los valores de IC_{50} en unidades micromolares (μM) y posteriormente transformar a valores de pIC_{50} utilizando la siguiente ecuación:

$$pIC_{50} = -\log(IC_{50}) \quad \text{Ecuación 1}$$

Nótese que esta transformación permite expresar la actividad inhibitoria en una escala logarítmica, donde valores mayores de pIC_{50} corresponden a compuestos con mayor actividad inhibitoria.

El siguiente paso consiste en representar cada compuesto seleccionado mediante notación SMILES, la cual se utiliza como punto de partida para construir estructuras tridimensionales mediante la aplicación web CACTVS [34]. Una vez obtenidas las estructuras tridimensionales, someterlas a un proceso de preoptimización conformacional mediante dinámica molecular, utilizando el programa DFTB+ [35] que implementa el método *Density Functional based Tight Binding*, el cual es una aproximación basada en la teoría del funcional de la densidad que reduce el costo computacional respecto a los cálculos DFT convencionales, por lo que resulta adecuada para la exploración preliminar de geometrías moleculares y procesos de preoptimización estructural. Para dicha preoptimización se emplea el conjunto de parámetros “3ob” [36], con un paso de integración de 2 femtosegundos (fs) y un proceso simulado de enfriamiento gradual desde 600 K hasta 10 K. La temperatura inicial elevada permite explorar diferentes conformaciones moleculares, mientras que el enfriamiento progresivo favorecerá la obtención de estructuras más estables y geométricamente razonables.

2.2. Docking

El paso inicial consiste en preparar archivos de entrada y archivos de configuración tanto para las estructuras de los compuestos inhibidores como para la enzima objetivo. Realizar el estudio de interacciones moleculares empleando el programa AutoDock 4.2.6 y AutoDock Tools [37].

2.2.1. Preparación de la Enzima

Como primer paso llevar a cabo una búsqueda en el Banco de Datos de Proteínas (PDB, por sus siglas en inglés, *Protein Data Bank*) [38] a fin de identificar y seleccionar un modelo estructural tridimensional adecuado de la enzima objetivo. Para sustentar la idoneidad de esta selección y garantizar la fiabilidad de los cálculos de acoplamiento molecular, se debe evaluar rigurosamente diversos parámetros cristalográficos y biológicos. Así como priorizar aquellas estructuras que presenten una alta resolución, preferentemente inferior a 2.5 Å y que correspondan a la especie biológica de interés (origen humano) y que se encuentren co-cristalizada con un ligando o inhibidor de referencia. Posteriormente, aislar la unidad catalíticamente activa de la enzima, posteriormente adicionar átomos de hidrogeno polares, cargas parciales y parámetros de solvatación, de acuerdo con el procedimiento empleado en estudios previos [39].

2.2.2. Preparación de los Inhibidores

Los inhibidores se preparan para los cálculos de *Docking* a partir de las estructuras previamente optimizadas. La preparación se realiza mediante AutoDock Tools (ADT), considerando la definición de la estructura del ligando, la identificación de los enlaces rotables y la asignación de las regiones rígidas y flexibles de cada molécula. Asimismo, se añaden los átomos de hidrógeno polares y se calculan las cargas parciales de tipo Gasteiger [39], requeridas para los cálculos de *Docking*. Finalmente, las estructuras preparadas se guardan en el formato PDBQT, necesario para su utilización en AutoDock.

2.2.3. Mapas de Afinidad Atómica

Delimitar el espacio de búsqueda en el sitio activo definiendo cuadrículas cúbicas de 90x90x90 puntos con un espaciado de 0.375 Å [40]. Centrar estas cajas de exploración en las coordenadas cartesianas correspondientes al centro de masa del ligante co-cristalizado en la estructura original de la enzima. Finalmente, calcular los mapas de afinidad atómica utilizando el programa AutoGrid 4.2 [37]. considerando las coordenadas de la enzima procesada y especificando cada uno de los tipos de átomos presentes en los inhibidores preparados en la etapa anterior [38].

2.2.4. Generación de Complejos Enzima-Inhibidor

Aplicar el cálculo de *Docking* mediante el Algoritmo Genético Lamarkiano (AGL) implementado en el programa AutoDock 4.2.6 con las siguientes opciones de configuración: 100 corridas, 150 individuos en la población inicial, 2.500.000 evaluaciones de energía y 27.000 generaciones. Utilizar el resto de los parámetros en valores por defecto [39,40]. Los 100 modelos del complejo enzima-inhibidor que se generan para cada compuesto y proteína considerados se someten a un análisis de agrupamiento para filtrar conformaciones equivalentes entre si considerando los valores de energía libre de unión y de RMSD [41]. Adicionalmente se debe calcular el parámetro LRCD como una medida de la similitud entre los complejos calculados y el complejo cristalográfico [42].

2.3. Clustering Mediante *K-means*

Una vez teniendo los complejos generados en el *Docking*, aplicar un análisis de clústers utilizando el algoritmo *K-means* [43, 44], con la finalidad de clasificar los complejos ligando–receptor en grupos con características semejantes. Para llevar a cabo esta clasificación de los compuestos de estudio, se consideraron las variables cuantitativas más importantes, la energía libre de unión (ΔG) obtenida mediante *Docking*, la actividad inhibitoria pIC_{50} y la proximidad espacial respecto al ligando de referencia, la selección de estas variables permite integrar, dentro de un mismo análisis, parámetros relacionados entre sí, facilitando la identificación y selección del clúster que mejor cumpla con los criterios establecidos.

Dado que las variables pueden presentar escalas numéricas diferentes, inicialmente efectuar una etapa de estandarización, de manera que cada variable contribuyera de forma equivalente al proceso de partición, posteriormente, evaluar el número de clústers más adecuado mediante el método del codo, calculando la inercia para valores de K comprendidos entre 1 y 10. Con base en la forma de la curva obtenida, seleccionar el número optimo de clústers, al representar el punto a partir del cual la reducción de la variabilidad intragrupo dejó de ser sustancial. Definido este valor, ejecutar el algoritmo *K-means* para el número de clústers. Cada complejo será asignado a un clúster según su cercanía al centroide correspondiente dentro del espacio multivariado estandarizado. Finalmente, la partición obtenida se visualizará mediante una representación tridimensional construida a partir de las variables originales, en la cual los complejos se deben distinguir cromáticamente según el grupo asignado, permitiendo identificar conjuntos de complejos con propiedades semejantes.

2.4. Cálculos DFT

Llevar a cabo los cálculos de optimización geométrica y frecuencias vibracionales de las moléculas preoptimizadas y previamente seleccionadas en el clúster resultante mediante *K-means*, utilizando el programa Gaussian 09 [45] empleando el funcional B3LYP [46] y el conjunto de base 6-31G** [47, 48]. Después corroborar que todas las frecuencias sean positivas a fin de asegurar que las estructuras optimizadas corresponden a un mínimo de energía en la superficie de energía potencial, en caso de que el cálculo tenga frecuencias imaginarias, se debe corregir la estructura modificando las posiciones atómicas a una geometría más estable y volver a calcular.

Posteriormente, con las moléculas optimizadas, realizar el estudio teórico de la reactividad química de los compuestos mediante el cálculo de descriptores globales, evaluando las propiedades relacionadas con el potencial de ionización (I) y la afinidad electrónica (A) a partir de energías calculadas para distintos estados de energía: E_c^{N-1} (energía de la estructura optimizada en su forma neutra al perder un electrón), E_o^N (energía de la estructura optimizada en su forma neutra) y E_A^{N+1} (energía de la estructura optimizada en su forma neutra al ganar un electrón) [49], de modo que:

$$I = E_c^{N-1} - E_o^N \quad \text{Ecuación 2}$$

$$A = E_o^N - E_A^{N+1} \quad \text{Ecuación 3}$$

Se puede emplear el principio de máxima dureza, que se obtiene cuando se aplica la aproximación de diferencias finitas de tres puntos y expresa el potencial químico (μ), la dureza (η) y la electronegatividad (χ) a través de I y A :

$$\mu = -\chi = -\frac{1}{2}(I + A) \quad \text{Ecuación 4}$$

$$\eta = \frac{I - A}{2} \quad \text{Ecuación 5}$$

Otro descriptor cuántico es la electrofilicidad (ω) [50], que se define como la capacidad de una molécula tanto para donar como para recibir electrones y se expresa mediante la siguiente ecuación:

$$\omega = \frac{\mu^2}{2\eta} \quad \text{Ecuación 6}$$

También es posible aproximar el potencial electrodonador (ω^-) y el poder electroaceptor (ω^+) [24], mediante las ecuaciones en términos de I y A :

$$\omega^- \equiv \frac{(\mu^-)^2}{2\eta} \approx \frac{(3I + A)^2}{16(I - A)} \quad \text{Ecuación 7}$$

$$\omega^+ \equiv \frac{(\mu^+)^2}{2\eta} \approx \frac{(I + 3A)^2}{16(I - A)} \quad \text{Ecuación 8}$$

Para el análisis cualitativo de la reactividad química de los inhibidores, se usan representaciones gráficas de funciones que dependen de la estructura electrónica como el mapeo de orbitales moleculares de frontera (HOMO y LUMO) y el mapeo de potenciales electroestáticos.

Cálculo de Descriptores de Dragon y RDKit

Utilizar las estructuras optimizadas de los inhibidores obtenidas a partir de los cálculos DFT, con el fin de calcular los descriptores estructurales utilizando los programas Dragon 6.0 [51] y RDKit [27] para cada molécula. En ambos casos los descriptores moleculares, los cuales en su mayoría son definidos por Todeschini y Consonni [52], están preprogramados para calcularse de forma automatizada utilizando la estructura química en formato mol y para homogeneizar los resultados se debe verificar en todos los casos que los órdenes de enlace estén correctamente configurados en estos archivos. Posterior al procesamiento de los archivos estructurales se obtienen cantidades numéricas obtenidas mediante un procedimiento matemático que transforma la información química codificada en una representación simbólica de la molécula y está basada en diferentes estrategias tales como teoría de grafos, ponderaciones en coordenadas 3D, notaciones lineales, entre otras. Estos descriptores en su mayoría de los casos pierden interpretación física directa sin embargo se pueden encontrar tendencias y relaciones con propiedades medibles (como variables experimentales), además que son dependientes de la estructura química y su distribución espacial, por lo que son útiles como un conjunto amplio de descriptores para modelado QSAR.

2.5. QSAR

Los modelos QSAR se construyen mediante regresiones lineales múltiples, de acuerdo con la ecuación general:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n$$
Ecuación 9

Donde Y corresponde a la variable respuesta y representa al valor de la actividad inhibitoria experimental (pIC_{50}), β_0 representa el intercepto, β_1 a β_n son los coeficientes de regresión y de X_1 a X_n son las variables independientes y corresponden a los descriptores moleculares.

Se debe generar una base de datos completa que involucren los dos conjuntos de descriptores calculados, por una parte, los parámetros globales calculados mediante DFT y se agregan los descriptores estructurales de Dragon y RDKit, esto sirve para preparar la matriz de datos con dimensión $m \times n$, que estará conformada por n variables que corresponden a la variable dependiente (pIC_{50}) y las variables independientes que son los descriptores de las m moléculas seleccionadas previamente en el clúster, para fines de formato, incluir sus etiquetas numéricas en la primera columna. Posteriormente, utilizar el programa QSAR-INS [53] para importar la matriz construida y eliminar el primer descriptor de los pares que contienen un valor absoluto de su coeficiente de correlación mayor o igual al 0.95, con la finalidad de evitar una sobre estimación en los modelos, realizar la división de la matriz de datos en dos submatrices, la primera que representa el conjunto de entrenamiento se obtiene de manera aleatoria con el 85% de las filas de la matriz y es utilizado exclusivamente para la selección de variables, la estimación de los coeficientes de regresión y la validación interna y la segunda que representa el conjunto de predicción externa se obtiene con el 15% restante de las filas de la matriz y es utilizado para evaluar la capacidad predictiva de las ecuaciones frente a moléculas que no participaron en el cálculo del modelo [54]. Como paso siguiente, realizar la selección de variables independientes mediante el uso de GA implementado en QSAR-INS, este método consiste en generar modelos de distintos tamaños de manera progresiva, es decir, que en cada generación se aumenta el número de descriptores y los modelos candidatos se seleccionan mediante la evaluación de su rendimiento con la función Q_{LOO}^2 , cada solución candidata representa una combinación específica de descriptores moleculares utilizada para construir una ecuación MLR. El proceso evolutivo se llevó a cabo de manera iterativa mediante operaciones de selección, recombinación y mutación, de este modo, las nuevas generaciones exploran diferentes combinaciones de variables con el propósito de identificar subconjuntos capaces de proporcionar modelos con adecuado ajuste. QSAR-INS permite controlar la búsqueda mediante parámetros

como el tamaño de la población, la tasa de mutación y el número de generaciones, así como seleccionar la función estadística utilizada para dirigir la optimización [53]. Una vez obtenidos los modelos, analizar y validar de manera interna con los parámetros como R^2 , Q^2 , $\bar{r}_{m(LOO)}^2$, $\Delta r_{m(LOO)}^2$, y de manera externa, utilizando parámetros como R_{pred}^2 , Q_{F2}^2 , $\Delta r_{m(test)}^2$ y $\bar{r}_{m(test)}^2$ y el dominio de aplicabilidad (definidos en el apartado de **Anexos**), con la finalidad de establecer confiabilidad y predictividad al modelo obtenido que permita predecir la actividad inhibitoria en compuestos que funcionan como IAGs. Finalmente analizar la tendencia que siguen los compuestos respecto a la diagonal resultante y comparar este resultado con las interacciones obtenidas en *Docking* y los mapas de potencial electrostático (MPE), obtenidos mediante DFT.

3. Resultados

En este capítulo se presentan los resultados obtenidos en este trabajo. En la primera sección se describen las estructuras químicas y las propiedades experimentales de los compuestos inhibidores seleccionados a partir de la búsqueda realizada en la literatura. Posteriormente, se presentan los resultados de la preoptimización estructural mediante dinámica molecular de los compuestos estudiados.

A continuación, se muestran los resultados del *Docking*, incluyendo la formación de los complejos enzima-inhibidor y el filtrado a partir de estos resultados. Después, la recopilación de las propiedades candidatas para el modelado QSAR.

En la última sección de este capítulo, se incluyen los modelos QSAR generados, la selección del modelo con mejor desempeño predictivo, el análisis de los descriptores relevantes incluidos en dicho modelo y el análisis de las interacciones no covalentes y mapeos de potencial electrostático observados en dichos complejos.

3.1. Estructuras y Propiedades Experimentales de los Inhibidores Seleccionados.

Después de la búsqueda y refinamiento de datos de la literatura proveniente de tres publicaciones [55] [56] [57] se procedió con el curado de la base de datos, se recopilaron 157 compuestos con su respectiva etiqueta en formato “M###”, donde # es un contador, con sus valores de IC_{50} experimental en μM , sus transformaciones a pIC_{50} y sus respectivos nombres, los cuales se reportan en la Tabla 1. Es importante mencionar que la conversión de IC_{50} a pIC_{50} permite linealizar la relación entre estructura química y actividad biológica, facilitando el análisis y comparación de los compuestos. Esta transformación es esencial en estudios QSAR, ya que mejora la interpretación de los modelos predictivos. Los valores de pIC_{50} se obtuvieron en un rango amplio, que va desde valores cercanos a -3.28 como **M156** (el compuesto de mayor IC_{50} , $643.6 \mu\text{M}$) hasta valores positivos como 1.398 con **M079** (el compuesto con menor IC_{50} , $0.040 \mu\text{M}$), lo que indica una alta variabilidad concentraciones inhibitorias en el conjunto, deseable para construir un modelo con una buena capacidad discriminante.

Tabla 1. Base de datos de los compuestos estudiados en el presente trabajo, con la información experimental reportada en la literatura.

Compuesto	IC ₅₀ (μM)	pIC ₅₀	Nombre	Ref.
M001	50.130	-1.700	1-(6-hidroxi-2-(prop-1-en-2-il)benzofuran-5-il)etan-1-ona	[55]
M002	45.040	-1.654	(E)-1-(6-hidroxi-2-(prop-1-en-2-il)benzofuran-5-il)-3-(2-metilfenil)prop-2-en-1-ona	[55]
M003	40.270	-1.605	(E)-1-(6-hidroxi-2-(prop-1-en-2-il)benzofuran-5-il)-3-(3-metilfenil)prop-2-en-1-ona	[55]
M004	82.670	-1.917	(E)-3-(2-fluorofenil)-1-(6-hidroxi-2-(prop-1-en-2-il)benzofuran-5-il)prop-2-en-1-ona	[55]
M005	39.770	-1.600	(E)-3-(4-fluorofenil)-1-(6-hidroxi-2-(prop-1-en-2-il)benzofuran-5-il)prop-2-en-1-ona	[55]
M006	41.210	-1.615	(E)-3-(4-clorofenil)-1-(6-hidroxi-2-(prop-1-en-2-il)benzofuran-5-il)prop-2-en-1-ona	[55]
M007	57.270	-1.758	(E)-3-(4-bromofenil)-1-(6-hidroxi-2-(prop-1-en-2-il)benzofuran-5-il)prop-2-en-1-ona	[55]
M008	9.020	-0.955	(E)-1-(6-hidroxi-2-(prop-1-en-2-il)benzofuran-5-il)-3-(2-(trifluorometil)fenil)prop-2-en-1-ona	[55]
M009	42.510	-1.628	(E)-1-(6-hidroxi-2-(prop-1-en-2-il)benzofuran-5-il)-3-(3-(trifluorometil)fenil)prop-2-en-1-ona	[55]
M010	42.500	-1.628	(E)-1-(6-hidroxi-2-(prop-1-en-2-il)benzofuran-5-il)-3-(tiofen-2-il)prop-2-en-1-ona	[55]
M011	15.880	-1.201	(E)-3-(4-(difenilamino)fenil)-1-(6-hidroxi-2-(prop-1-en-2-il)benzofuran-5-il)prop-2-en-1-ona	[55]
M012	71.980	-1.857	(E)-3-([1,1'-bifenil]-4-il)-1-(6-hidroxi-2-(prop-1-en-2-il)benzofuran-5-il)prop-2-en-1-ona	[55]
M013	12.490	-1.097	(E)-3-(antracen-9-il)-1-(6-hidroxi-2-(prop-1-en-2-il)benzofuran-5-il)prop-2-en-1-ona	[55]
M014	85.320	-1.931	(E)-3-(2-bromo-4,5-dimetoxifenil)-1-(6-hidroxi-2-(prop-1-en-2-il)benzofuran-5-il)prop-2-en-1-ona	[55]
M015	96.250	-1.983	(E)-1-(6-hidroxi-2-(prop-1-en-2-il)benzofuran-5-il)-3-(2,3,4-trimetoxifenil)prop-2-en-1-ona	[55]
M016	91.320	-1.961	(E)-1-(6-hidroxi-2-(prop-1-en-2-il)benzofuran-5-il)-3-(2,4,6-trimetoxifenil)prop-2-en-1-ona	[55]
M017	4.600	-0.663	2-(5-((benzo[d]tiazol-2-iltio)metil)-1,2,4-oxadiazol-3-il)-5-fluorofenol	[56]
M018	5.600	-0.748	5-((benzo[d]tiazol-2-iltio)metil)-3-(4-(trifluorometil)fenil)-1,2,4-oxadiazol	[56]
M019	15.600	-1.193	5-((benzo[d]tiazol-2-iltio)metil)-3-(2-nitrofenil)-1,2,4-oxadiazol	[56]
M020	34.600	-1.539	5-((benzo[d]tiazol-2-iltio)metil)-3-(4-bromo-2-nitrofenil)-1,2,4-oxadiazol	[56]
M021	14.600	-1.164	5-((benzo[d]tiazol-2-iltio)metil)-3-(4-nitrofenil)-1,2,4-oxadiazol	[56]
M022	31.600	-1.500	5-((benzo[d]tiazol-2-iltio)metil)-3-(2-metilfenil)-1,2,4-oxadiazol	[56]
M023	29.600	-1.471	5-((benzo[d]tiazol-2-iltio)metil)-3-(4-metilfenil)-1,2,4-oxadiazol	[56]
M024	9.500	-0.978	5-((benzo[d]tiazol-2-iltio)metil)-3-(4-fluorofenil)-1,2,4-oxadiazol	[56]
M025	33.800	-1.529	5-((benzo[d]tiazol-2-iltio)metil)-3-(3-metilfenil)-1,2,4-oxadiazol	[56]
M026	26.500	-1.423	5-((benzo[d]tiazol-2-iltio)metil)-3-(2,5-dimetoxifenil)-1,2,4-oxadiazol	[56]
M027	31.200	-1.494	5-((benzo[d]tiazol-2-iltio)metil)-3-(2-metoxifenil)-1,2,4-oxadiazol	[56]
M028	10.500	-1.021	5-((benzo[d]tiazol-2-iltio)metil)-3-(3-fluorofenil)-1,2,4-oxadiazol	[56]
M029	36.600	-1.563	5-((benzo[d]tiazol-2-iltio)metil)-3-(3-metoxifenil)-1,2,4-oxadiazol	[56]
M030	18.600	-1.270	5-((benzo[d]tiazol-2-iltio)metil)-3-(2,3-diclorofenil)-1,2,4-oxadiazol	[56]
M031	19.500	-1.290	5-((benzo[d]tiazol-2-iltio)metil)-3-(4-cloro-2-nitrofenil)-1,2,4-oxadiazol	[56]
M032	13.400	-1.127	2-(5-((benzo[d]tiazol-2-iltio)metil)-1,2,4-oxadiazol-3-il)-3,5-diclorofenol	[56]
M033	48.400	-1.685	5-((benzo[d]tiazol-2-iltio)metil)-3-(4-bromofenil)-1,2,4-oxadiazol	[56]
M034	2.200	-0.342	2-((E)-((E)-2-(1H-benzo[d]imidazol-2-iltio)-1-(4-metoxifenil)etilideno)hidrazono)metil)-5-metoxifenol	[58]
M035	2.400	-0.380	4-((E)-((E)-2-(1H-benzo[d]imidazol-2-iltio)-1-(4-metilfenil)etilideno)hidrazono)metil)-2-metoxifenol	[58]
M036	25.200	-1.401	2-(4-((4-(1H-benzo[d]imidazol-2-il)fenoxi)metil)-1H-1,2,3-triazol-1-il)-N-(4-etilfenil)acetamida	[59]
M037	14.556	-1.163	N-(3-fluorofenil)-2-(4-((2-metil-1H-benzo[d]imidazol-1-il)metil)-1H-1,2,3-triazol-1-il)acetamida	[60]
M038	5.370	-0.730	(2R,3R,4S)-1,2,3,4-tetrahidrobenzo[4,5]imidazo[1,2-a]piridina-2,3,4-triol	[61]
M039	82.200	-1.915	(E)-N'-(4-(3-(1,3-dioxoisindolin-2-il)propoxi)bencilideno)-2-oxo-2H-cromeno-3-carbohidrazida	[62]
M040	423.300	-2.627	(E)-N'-(4-(3-(1,3-dioxoisindolin-2-il)propoxi)-3-metoxibencilideno)-2-oxo-2H-cromeno-3-carbohidrazida	[62]
M041	12.700	-1.104	2-(3-(5-oxo-2-fenoxi-[1,2,4]triazolo[1,5-a]quinazolin-4(5H)-il)propil)isindolina-1,3-diona	[63]
M042	40.500	-1.607	3-(1,3-dioxoisindolin-2-il)-N-((1-(4-fluorobencil)-1H-1,2,3-triazol-4-il)metil)benzamida	[64]

M043	17.900	-1.253	(R)-4-(4-fluorofenil)-1-(2-(2-metoxi-4-(naftalen-2-il)fenoxi)propil)-1H-1,2,3-triazol	[65]
M044	170.000	-2.230	(2R,3R,4R)-2-(hidroximetil)-1-(4-(4-(4-nitrofenil)-1H-1,2,3-triazol-1-il)butil)pirrolidina-3,4-diol	[66]
M045	170.000	-2.230	N-(4-((2S,3S,4S)-3,4-dihidroxi-2-(hidroximetil)pirrolidin-1-il)butil)-4-nitrobenzamida	[66]
M046	13.000	-1.114	3,3'-((4-((1-(2-clorobencil)-1H-1,2,3-triazol-4-il)metoxi)fenil)metileno)bis(4-hidroxi-2H-cromen-2-ona)	[67]
M047	16.400	-1.215	3,3'-((3-metoxi-4-((1-(2-metilbencil)-1H-1,2,3-triazol-4-il)metoxi)fenil)metileno)bis(4-hidroxi-2H-cromen-2-ona)	[67]
M048	90.400	-1.956	1-(3,5-dimetilbencil)-4-((4-(4,5-difenil-1H-imidazol-2-il)fenoxi)metil)-1H-1,2,3-triazol	[68]
M049	2.060	-0.314	N-(2,4-dimetoxifenil)-2-(4-(((1-hidroxi-6-metoxi-9-oxo-9H-xanten-3-il)oxi)metil)-1H-1,2,3-triazol-1-il)acetamida	[69]
M050	41.290	-1.616	1-(3-(4-metoxifenil)-5-(4-((1-(4-metilbencil)-1H-1,2,3-triazol-4-il)metoxi)fenil)-4,5-dihidro-1H-pirazol-1-il)etan-1-ona	[70]
M051	47.940	-1.681	1-(3-(4-bromofenil)-5-(4-((1-(4-nitrobencil)-1H-1,2,3-triazol-4-il)metoxi)fenil)-4,5-dihidro-1H-pirazol-1-il)etan-1-ona	[70]
M052	75.000	-1.875	(E)-2-(4-((1-(2,6-diclorobencil)-1H-1,2,3-triazol-5-il)oxi)-3-metoxibencilideno)-N-(4-metoxifenil)hidrazina-1-carbotioamida	[71]
M053	85.600	-1.932	1-(2-clorobencil)-4-(((4,5-difenil-1H-imidazol-2-il)tio)metil)-1H-1,2,3-triazol	[72]
M054	107.100	-2.030	(E)-N-(2-(2-(4-((1-(3-clorobencil)-1H-1,2,3-triazol-4-il)metoxi)-3-metoxibencilideno)hidrazina-1-carbonil)fenil)-4-metoxibenzamida	[73]
M055	2.090	-0.320	(E)-1-(1H-indol-2-il)-N-(4H-1,2,4-triazol-4-il)metanimina	[74]
M056	116.559	-2.067	(E)-4-((4-metoxibencilideno)amino)-5-(5-metil-1H-pirazol-3-il)-4H-1,2,4-triazol-3-tiol	[75]
M057	33.620	-1.527	4-(1-(4-bromofenil)-4,5-difenil-1H-imidazol-2-il)-3-(3-nitrofenil)-1-fenil-1H-pirazol	[76]
M058	25.190	-1.401	4-(1-(4-bromofenil)-4,5-difenil-1H-imidazol-2-il)-3-(4-clorofenil)-1-fenil-1H-pirazol	[76]
M059	12.450	-1.095	7-(4-clorofenil)-5-(3-metil-1-fenil-5-(1H-1,2,4-triazol-1-il)-1H-pirazol-4-il)-5,8-dihidro-[1,2,4]triazolo[4,3-a]pirimidina	[77]
M060	4.480	-0.651	2-(4-metoxifenilimino)-5-((3-(4-nitrofenil)-1-fenil-1H-pirazol-4-il)metileno)tiazolidin-4-ona	[78]
M061	3.260	-0.513	(Z)-3-(((E)-(4-(3-nitrofenil)-1-fenil-1H-pirazol-3-il)metileno)hidrazinilideno)indolin-2-ona	[79]
M062	5.470	-0.738	(Z)-3-(((E)-(4-(4-nitrofenil)-1-fenil-1H-pirazol-3-il)metileno)hidrazinilideno)indolin-2-ona	[79]
M063	131.815	-2.120	(E)-1-fenil-N-(piridin-2-il)metanimina	[80]
M064	0.207	0.684	(2S,3S,4S,5R)-1-(4-hidroxi-3-metoxibencil)-2-(hidroximetil)piperidina-3,4,5-triol	[81]
M065	0.276	0.559	(2S,3S,4S,5R)-1-(3-bromo-4-hidroxi-5-metoxibencil)-2-(hidroximetil)piperidina-3,4,5-triol	[81]
M066	1.332	-0.125	(Z)-etil-4-(2-(hidroximetil)-5-((6-metil-3,4-dioxo-4,5-dihidrofuro[3,4-c]piridin-1(3H)-ilideno)metil)-1H-pirrol-1-il)benzoato	[82]
M067	111.000	-2.045	2-(5-(4-clorofenil)-1-(2,4-diclorofenil)-2-metil-1H-pirrol-3-il)-2-oxo-N-(piridin-3-il)acetamida	[83]
M068	196.000	-2.292	2-(1-(2,4-diclorofenil)-2-metil-5-fenil-1H-pirrol-3-il)-2-oxo-N-(piridin-3-il)acetamida	[83]
M069	123.360	-2.091	(4-fenil-1,2,5-oxadiazol-3-il)metil-3,4-dihidroxibenzoato	[84]
M070	80.590	-1.906	(4-fenil-1,2,5-oxadiazol-3-il)metil-2-(3,4-dihidroxifenil)acetato	[84]
M071	0.300	0.523	(E)-4-(5-((4-hidroxibencilideno)amino)-1,3,4-oxadiazol-2-il)fenol	[85]
M072	52.630	-1.721	2-((5-(4-clorofenil)-1,3,4-oxadiazol-2-il)tio)-N-(3,4-dimetilfenil)propanamida	[86]
M073	52.730	-1.722	2-((5-(4-clorofenil)-1,3,4-oxadiazol-2-il)tio)-N-(3-metilfenil)propanamida	[86]
M074	0.500	0.301	4-(5-(4-(6-clorobenzo[d]tiazol-2-il)fenil)-1,3,4-oxadiazol-2-il)benceno-1,2-diol	[87]
M075	215.000	-2.332	3,4-bis(4-clorofenil)-5-fenil-4,5-dihidro-1,2,4-oxadiazol	[88]
M076	256.000	-2.408	3-(4-clorofenil)-4-(4-fluorofenil)-5-fenil-4,5-dihidro-1,2,4-oxadiazol	[88]
M077	499.600	-2.699	((4aS,6aS,6bR,10R,11R,12aR)-10,11-dihidroxi-2,2,6a,6b,9,9,12a-heptametil-1,3,4,5,6,6a,6b,7,8,8a,9,10,11,12,12a,12b,13,14b-octadecahidropicen-4a(2H)-il)(4-(2-hidroxietyl)piperazin-1-il)metanona	[89]
M078	1.800	-0.255	(1R,3aS,5aR,5bR,14aR)-5a,5b,8,8,14a-pentametil-1-(prop-1-en-2-il)-2,3,4,5,5a,5b,6,7,7a,8,9,14,14a,14b,15,16,16a,16b-octadecahidrociclopenta[7,8]fenantro[2,1-b]carbazon-3a(1H)-ácido carboxílico	[90]
M079	0.040	1.398	ácido (2-oxo-2-((1R,3aS,5aR,5bR,14aR)-5a,5b,8,8,14a-pentametil-1-(prop-1-en-2-il)-2,3,4,5,5a,5b,6,7,7a,8,9,14,14a,14b,15,16,16a,16b-octadecahidrociclopenta[7,8]fenantro[2,1-b]carbazon-3a(1H)-il)etil)carbámico	[90]
M080	364.000	-2.561	ácido N-[2α,3β-dihidroxi-urs-12-en-28-ol]-L-aspartico	[91]
M081	0.330	0.481	(6aR,6bS,14bR)-13-fluoro-4,4,6a,6b,11,11,14b-heptametil-2-((Z)-4-metilbencilideno)octadecahidro-3H,9H-12b,8a-(epoximetano)picen-3,16-diona	[92]

Tabla 1. Continuación.

M082	20.100	-1.303	(1R,3aR,5aR,5bR,11aR)-10-((Z)-4-bromobencilideno)-3a,5a,5b,8,8,11a-hexametil-1-(prop-1-en-2-il)icosahidro-9H-ciclopenta[a]crisen-9-ona	[93]
M083	19.590	-1.292	(1S,2R,4aR,6aR,6bR,12aR,14aR)-11-((Z)-2-fluorobencilideno)-1,2,4a,6b,9,9,12a-heptametiloctadecahidro-6a,14a-metanopicen-10(6bH)-ona	[94]
M084	0.465	0.333	metil-10-metoxi-2,4a,6a,6b,9,9,12a-heptametil-13-oxo-1,2,3,4,4a,5,6,6a,6b,7,8,8a,9,10,11,12,12a,12b,13,14b-icosahidropicen-2-carboxilato	[95]
M085	0.687	0.163	metil-2,4a,6a,6b,9,9,12a-heptametil-13-oxo-10-(((3R,4R,5R,6R)-3,4,5-trihidroxi-6-(hidroximetil)tetrahidro-2H-piran-2-il)oxi)-1,2,3,4,4a,5,6,6a,6b,7,8,8a,9,10,11,12,12a,12b,13,14b-icosahidropicen-2-carboxilato	[95]
M086	2.050	-0.312	N-([1,1'-bifenil]-4-il(2-hidroxinaftalen-1-il)metil)acrilamida	[96]
M087	13.020	-1.115	5-(3-(1-etil-1H-indol-3-il)-1-(naftalen-2-il)-3-oxopropil)-1,3-dimetilpirimidina-2,4,6(1H,3H,5H)-triona	[97]
M088	2.560	-0.408	(2R,3R,5S)-2-(((3,4,5-trihidroxibenzoil)oxi)metil)tetrahidro-2H-piran-3,4,5-triil tris(3,4,5-trihidroxibenzoato)	[98]
M089	1.680	-0.225	(3R,5R,6R)-2-hidroxi-6-(((3,4,5-trihidroxibenzoil)oxi)metil)tetrahidro-2H-piran-3,4,5-triil tris(3,4,5-trihidroxibenzoato)	[98]
M090	0.120	0.921	(1S,2S,4R,5S)-5-((1,3-dihidroxiopropan-2-il)amino)-1-(hidroximetil)ciclohexano-1,2,4-triol	[99]
M091	62.400	-1.795	1,4-didesoxi-1,4-imino-D-galactitol	[100]
M092	40.600	-1.609	N-bencil-1,4-didesoxi-1,4-imino-D-galactitol	[100]
M093	55.600	-1.745	2-amino-4-(4-hidroxifenil)-6-(feniltio)piridina-3,5-dicarbonitrilo	[101]
M094	3.010	-0.479	4,4'-(1,3-fenileno)bis(2-amino-5-((E)-3-(4-hidroxi-3-metoxifenil)acrilol)-6-((E)-4-hidroxi-3-metoxiestiril)-4H-pirano-3-carbonitrilo)	[101]
M095	15.200	-1.182	4-hidroxifenetil-4-cianobenzoato	[102]
M096	643.605	-2.809	4-(3-hidroxi-3-metiltriaz-1-en-1-il)bencenosulfonamida	[103]
M097	70.570	-1.849	5,7-dibromo-3-hidroxi-2-{1-[(4-metilfenil)sulfonil]-1H-indol-3-il}-4H-cromen-4-ona	[104]
M098	2.310	-0.364	etil-4-(2-bromo-4,5-dimetoxifenil)-2,7,7-trimetil-5-oxo-1,4,5,6,7,8-hexahidroquinolina-3-carboxilato	[105]
M099	12.980	-1.113	2-oxo-2H-cromen-4-il-(E)-3-(4-metilfenil)acrilato	[106]
M100	28.300	-1.452	etil-2-ciano-2-(2,5-dioxo-1-fenilpirrolidin-3-il)-2-fenilacetato	[107]
M101	0.980	0.009	etil-6-((4-((2,4-dioxotiazolidin-5-ilideno)metil)fenoxi)metil)-1,3-dimetil-2-oxo-4-fenil-1,2,3,4-tetrahidropirimidina-5-carboxilato	[107]
M102	95.000	-1.978	etil-(Z)-3-(ciclohexilimino)-2-(4-hidroxi-3-metoxifenil)-6-fenil-3H-imidazo[1,2-b]pirazol-7-carboxilato	[108]
M103	1.680	-0.225	4-acetil-2,6-dibromo-3-hidroxifenil-2-nitrobenzoato	[109]
M104	24.370	-1.387	etil-1-(dietoxifosforil)-4-((dietoxifosforil)amino)-2,6-bis(4-hidroxifenil)-1,2,5,6-tetrahidropiridina-3-carboxilato	[110]
M105	140.900	-2.149	bencil-4-((2,4-dihidroxi-6-metilbenzoil)oxi)-2-hidroxi-6-metilbenzoato	[111]
M106	203.900	-2.309	bencil-4-(2-((2,4-dihidroxi-6-metilbenzoil)oxi)etoxi)-2-hidroxi-6-metilbenzoato	[111]
M107	150.000	-2.176	3,4-dihidroxifenetil-2-hidroxi-4-[(4-metilfenil)sulfonamido]benzoato	[112]
M108	86.000	-1.934	3-(3,4-dihidroxifenil)propil-2-hidroxi-4-[(4-metilfenil)sulfonamido]benzoato	[112]
M109	320.000	-2.505	3,4-dihidroxifenetil-4-acetamido-2-hidroxibenzoato	[112]
M110	6.200	-0.792	N'-(3,4-dihidroxibencilideno)quinolina-6-carbohidrazida	[113]
M111	62.250	-1.794	(E)-N'-(2,3,4,4a-tetrahidro-1H-xanten-1-ilideno)-2-(4-(trifluorometil)fenil)acetohidrazida	[114]
M112	18.920	-1.277	(E)-4-(di(1H-indol-3-il)metil)-N'-(3,4-dihidroxibencilideno)benzohidrazida	[115]
M113	1.540	-0.188	(E)-N-(2'',4''-difluorofenil)-2-(6,6-dimetil-1-fenil-1,5,6,7-tetrahidro-4H-indazol-4-ilideno)hidrazina-1-carbotioamida	[116]
M114	13.990	-1.146	2-(5-fluoro-1H-indol-2-carbonil)-N-(2-fluorofenil)hidrazina-1-carbotioamida	[117]
M115	1.300	-0.114	2-(2-((5H-[1,2,4]triazino[5,6-b]indol-3-il)tio)-1-(4-metoxifenil)etilideno)-N-(2,3-diclorofenil)hidrazina-1-carbotioamida	[118]
M116	1.300	-0.114	2-(2-((5H-[1,2,4]triazino[5,6-b]indol-3-il)tio)-1-(3-nitrofenil)etilideno)-N-(2,3-diclorofenil)hidrazina-1-carbotioamida	[118]
M117	13.090	-1.117	(E)-3-(2-(4-clorobencilideno)hidrazinil)-5,6-difenil-1,2,4-triazina	[119]
M118	104.070	-2.017	(E)-5-(2-(2-metoxibencilideno)hidrazinil)-2-fenoxipirido[3,2-e][1,2,4]triazolo[1,5-a]pirimidina	[120]
M119	5.001	-0.699	(Z)-2-(1-(2-(6-clorobenzo[d]tiazol-2-il)hidrazinilideno)-3-metilbutil)-1H-indeno-1,3(2H)-diona	[121]
M120	0.359	0.445	2,2'-((1E,1'E)-hidrazina-1,2-diilidenobis(metaniilideno))bis(4-((E)-(4-nitrofenil)diazenil)fenol)	[122]
M121	16.100	-1.207	(4R,4aS,6aR,8S,9R,11aR,11bS)-1-((ciclohexanocarbonil)oxi)-4,8,11b-trimetiltetradecahidro-6a,9-metanociclohepta[a]naftaleno-4-ácido carboxílico	[123]

Tabla 1. Continuación.

M122	3.640	-0.561	(6'R,7'S,7a'R)-5-bromo-7'-(4-(dimetilamino)fenil)-6'-(3-(4-(dimetilamino)fenil)acrilol)-1-(4-metoxifenil)-1',6',7',7a'-tetrahidro-3'H-espiro[indolina-3,5'-pirrolo[1,2-c]tiazol]-2-ona	[124]
M123	74.000	-1.869	1-(4-bromofenil)-8-metoxi-3-fenil-6,11-dihidro-5H-imidazo[1',5':1,2]pirido[3,4-b]indol	[125]
M124	81.000	-1.908	1-(4-clorofenil)-3-fenil-6,11-dihidro-5H-imidazo[1',5':1,2]pirido[3,4-b]indol	[125]
M125	141.130	-2.150	6-metil-7H-[1,2,4]triazolo[4,3-b][1,2,4]triazepin-8(9H)-ona	[126]
M126	281.210	-2.449	6-metil-7H-[1,2,4]triazolo[4,3-b][1,2,4]triazepin-8(9H)-tione	[126]
M127	2.040	-0.310	1-(3,4-diclorofenil)-2-(5,6-dimetil-1H-benzo[d][1,2,3]triazol-1-il)etan-1-ona	[127]
M128	2.210	-0.344	1-(4-clorofenil)-2-(5,6-dimetil-1H-benzo[d][1,2,3]triazol-1-il)etan-1-ona	[127]
M129	0.400	0.398	(E)-2-fenil-3-((4aR,5S,10bR)-5-fenil-3,4,4a,5,6,10b-hexahidro-2H-pirano[3,2-c]quinolin-9-il)ácido acrílico	[128]
M130	27.620	-1.441	2-(benzo[d][1,3]dioxol-5-il)-4H-cromen-4-ona	[129]
M131	62.260	-1.794	2,4-difenil-4H-cromeno	[130]
M132	950.300	-2.978	2-(4-clorofenil)-4-fenil-4H-cromeno	[130]
M133	2.990	-0.476	dietil-4-hidroxi-6-nitro-4H-cromeno-2,3-dicarboxilato	[131]
M134	3.150	-0.498	dimetil-6-bromo-4-hidroxi-4H-cromeno-2,3-dicarboxilato	[131]
M135	25.700	-1.410	2-((1S,3aR,4S,6aR)-4-(3,4-dihidroxifenil)tetrahidro-1H,3H-furo[3,4-c]furano-1-il)-4,5-dihidroxifenil acetato	[132]
M136	171.670	-2.235	N-(3-acetil-2-metil-4-fenilquinolin-6-il)-2-nitrobenzamida	[133]
M137	35.200	-1.547	2,6-bis((E)-3-hidroxibencilideno)ciclohexan-1-ona	[134]
M138	19.400	-1.288	2-((E)-4-bromobencilideno)-6-((E)-4-hidroxibencilideno)ciclohexan-1-ona	[134]
M139	21.900	-1.340	2,6-bis((E)-4-hidroxibencilideno)ciclohexan-1-ona	[134]
M140	7161.00 0	-3.855	2,2-dimetil-5-(((3-(trifluorometil)fenil)amino)metileno)-1,3-dioxano-4,6-dione	[135]
M141	890.000	-2.949	(Z)-7-((3-(4-clorofenil)-1-(dimetilamino)-3-oxoprop-1-en-2-il)oxi)-6-metoxi-3-fenil-4H-cromen-4-ona	[136]
M142	16.930	-1.229	1-(6-bromobenzo[d]tiazol-2-il)-3-metilindeno[1,2-c]pirazol-4(1H)-ona	[137]
M143	21.533	-1.333	1-(6-clorobenzo[d]tiazol-2-il)-3-isopropilindeno[1,2-c]pirazol-4(1H)-ona	[137]
M144	1.800	-0.255	2,5-dihidroxi-3-pentadecilciclohexa-2,5-dieno-1,4-dione	[138]
M145	16.390	-1.215	7-hidroxi-3-(2-metilfenil)-2H-cromen-2-ona	[139]
M146	11.540	-1.062	6-hidroxi-3-(4-hidroxifenil)-2H-cromen-2-ona	[139]
M147	1.370	-0.137	5,7-dihidroxi-3-(4-hidroxifenil)-2H-cromen-2-ona	[139]
M148	91.670	-1.962	(1R,2S,4S,5R)-4,5-bis(hidroximetil)ciclohexano-1,2-diol	[140]
M149	527.390	-2.722	(1R,2S,4R,5S)-4,5-bis(hidroximetil)ciclohexano-1,2-diol	[140]
M150	0.280	0.553	etil-(Z)-1-(4-clorobencil)-3-(3-((4-oxo-2-tioxotiazolidin-5-ilideno)metil)fenil)-1H-indol-2-carboxilato	[141]
M151	0.610	0.215	etil-(Z)-1-bencil-3-(3-((4-oxo-2-tioxotiazolidin-5-ilideno)metil)fenil)-1H-indol-2-carboxilato	[141]
M152	15.200	-1.182	6-amino-7-(4-bromofenil)-N-fenil-2-(fenilamino)-5-(4-metilfenil)pirazolo[1,5-a]pirimidina-3-carboxamida	[142]
M153	17.600	-1.246	6-amino-7-(4-bromofenil)-N-(4-clorofenil)-2-(fenilamino)-5-(4-metilfenil)pirazolo[1,5-a]pirimidina-3-carboxamida	[142]
M154	124.240	-2.094	N,N'-bis(5-cloropiridin-2-il)-2,6-di(furan-2-il)-4-oxociclohexano-1,1-dicarboxamida	[143]
M155	148.180	-2.171	2,6-bis(2,4-diclorofenil)-4-oxo-N,N'-di(piridin-2-il)ciclohexano-1,1-dicarboxamida	[143]
M156	1.920	-0.283	5,5'-(oxibis(metileno))bis(3,4,6-tribromobenceno-1,2-diol)	[144]
M157	112.000	-2.049	(6R,7R,8R,8aS)-7,8-dihidroxi-6-(hidroximetil)-7,8,8a,9-tetrahidro-4H,6H-pirrolo[1,2-a][1,2,3]triazolo[1,5-d]pirazin-4-ona	[145]

Tabla 1. Continuación.

En la base de datos reportada en la Tabla 1 se observa la diversidad estructural del conjunto de IAGs, se obtuvieron 18 categorías, 2 basadas en similitud estructural, el primer conjunto derivados de euparin, el segundo conjunto derivados de benzotiazol basados en oxadiazol y las demás son un conjunto de categorías variadas, entre las que destacan derivados de benzimidazol, Isoindolina, 1,2,3-triazol, diazol, piridina, pirrol, oxadiazol, esteroidales, naftaleno, entre otros. Se seleccionaron los compuestos que presentaban mayor, menor y medio valor de IC_{50} resultantes en el modelo QSAR, a continuación se realiza un análisis estructural de estos:

En la Figura 2 se observa al compuesto **M011** ($pIC_{50} = -1.201$) el cual es un derivado de Euparin, que incorpora un núcleo benzofurano sustituido con un sistema de conjugación extendida tipo chalcona y un grupo difenilamino terminal, en esta estructura se observa un sistema π conjugado y un grupo carbonilo α,β -insaturado lo que le proporciona a esta familia de compuestos una alta capacidad de interacción hidrofóbica y aromática con el sitio activo enzimático.

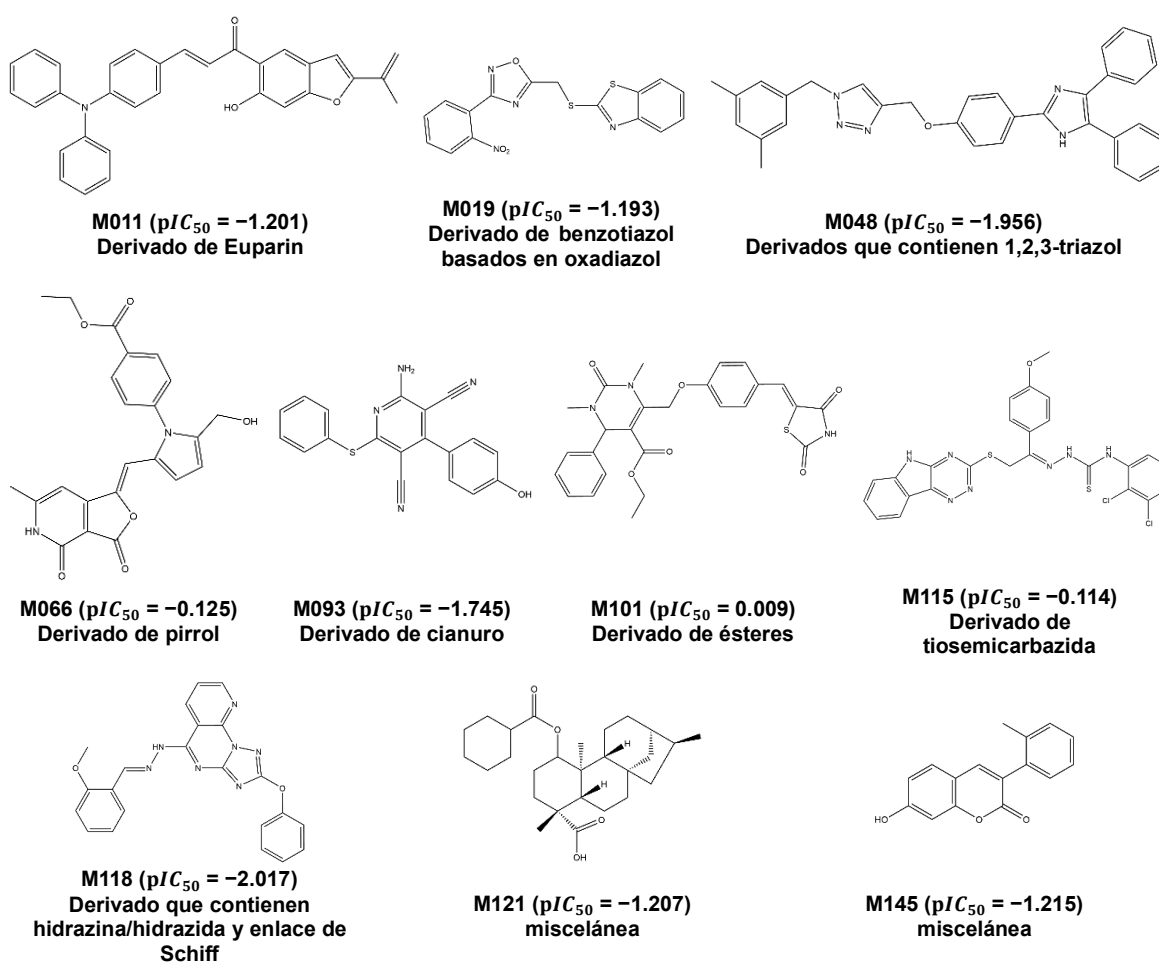


Figura 2. Compuestos provenientes de diferentes derivados

M019 ($pIC_{50} = -1.193$) pertenece a la familia de los benzotiazoles basados en oxadiazol. Su estructura combina un anillo benzotiazólico con un enlace tioéter que conecta a una subestructura de 1,2,4-oxadiazol que a su vez se enlaza a un grupo nitrofenilo. Se observan particularmente al átomo de azufre y nitrógeno del benzotiazol y los nitrógenos del oxadiazol como potenciales aceptores de puentes de hidrógeno [146], así mismo los grupos aromáticos podrían interactuar con residuos aromáticos del sitio activo, lo cual es una característica importante para su modo de unión [147].

M048 (pIC_{50} de -1.956) representa a los derivados que contienen el anillo 1,2,3-triazol. Este es un heterociclo nitrogenado de cinco miembros y bastante importante, ya que tiene la capacidad de formar puentes de hidrógeno y se caracteriza por presentar propiedades dipolares [148]. En la Figura 2 se observa que el triazol se encuentra conectado mediante un enlace éter a un sistema bifenilo-imidazol con un puente aromático de un grupo fenilo, esta estructura produce una ampliación de su superficie de contacto hidrofóbico con la enzima, por lo tanto, se espera el favorecimiento de las interacciones de tipo π - π con los residuos de fenilalanina y triptófano del sitio catalítico, debido a su naturaleza aromática [149].

El compuesto **M066** ($pIC_{50} = -0.125$) es un derivado de pirrol, presenta un anillo de benceno para sustituido por un éster etílico y un pirrol N-sustituido, unidos a un sistema bicíclico con grupos lactama y lactona. La presencia de grupos donadores y aceptores de puentes de hidrógeno distribuidos en distintas regiones de la molécula facilita interacciones polares con residuos catalíticos como Asp e His [150], mientras que el núcleo bicíclico aporta rigidez conformacional que favorece la unión en el sitio activo.

M093 ($pIC_{50} = -1.745$) representa a los derivados de cianuro, específicamente una dicianopiridina sustituida con grupos amino, tiol y un fragmento fenólico. Los grupos ciano son aceptores electrónicos potentes que modifican la distribución de densidad electrónica global de la molécula, lo que se manifiesta en un LUMO de baja energía y una brecha HOMO-LUMO relativamente estrecha, sugiriendo una mayor polarizabilidad y reactividad frente a residuos nucleófilos del sitio activo [151].

M101 ($pIC_{50} = 0.009$) es un derivado de éster con uno de los valores de pIC_{50} más altos del conjunto, lo que lo convierte en uno de los inhibidores promisorios de la base de datos. Su estructura presenta un núcleo hidantoína con sustituyentes metilo, fenilo y un éster etílico, unido de igual forma a un fragmento tiazolidinona través de un espaciador benziloxi. Esta combinación de grupos polares e hidrofóbicos distribuidos en distintos extremos de la molécula genera un perfil anfipático que favorece interacciones simultáneas con residuos polares del sitio catalítico y con el entorno hidrofóbico de la cavidad.

M115 ($pIC_{50} = -0.114$) pertenece a los derivados de tiosemicarbazida, una familia conocida por su versatilidad como inhibidores enzimáticos [152]. Su estructura incluye un fragmento tipo indol/triazina a un enlace tioéter y unido mediante un enlace imino a un grupo tiosemicarbazida que porta una cadena 2,6-diclorofenilo. El átomo de azufre del tioéter podría participar en interacciones de tipo π -azufre [153] mientras que el nitrógeno del grupo imino podría contribuir como sitio polar/aceptor de puente de hidrógeno o como parte del sistema conjugado del ligando.

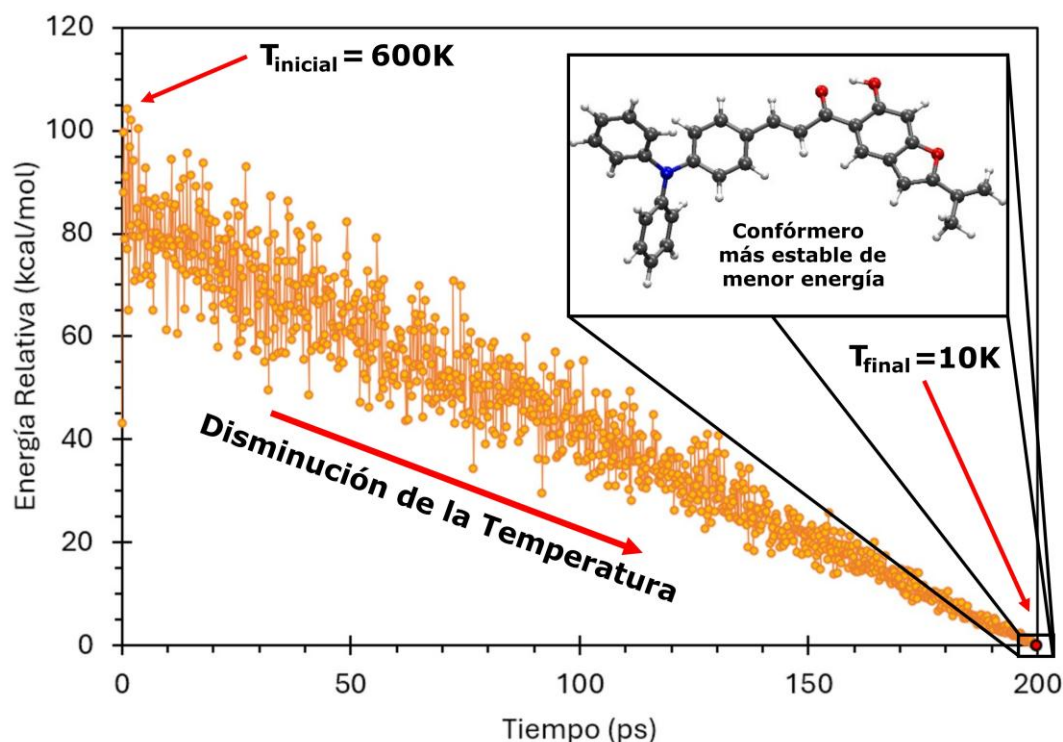
M118 ($pIC_{50} = -2.017$) es un representante de la familia de hidrazinas/hidrazidas con base de Schiff. Contiene un sistema piridina-pirimidina conectado mediante un enlace imino tipo base de Schiff a un fragmento metoxibencilo y un fragmento fenil éter unido al anillo de triazol. A pesar de presentar la mayor IC_{50} entre los representativos mostrados en la Figura 2 su inclusión es relevante desde el punto de vista estructural, ya que permite identificar los rasgos moleculares que limitan la actividad inhibitoria, como la mayor flexibilidad conformacional y la menor complementariedad estérica con el sitio activo respecto a compuestos con mejor capacidad inhibitoria.

M121 ($pIC_{50} = -1.207$) y **M145** ($pIC_{50} = -1.215$) se agrupan bajo la categoría de miscelánea de familias variadas. **M121** corresponde a un esteroide bicíclico de gran tamaño molecular con grupos éster y ácido carboxílico, mientras que **M145** presenta un núcleo de cumarina con un sustituyente hidroxilo en posición C-6 y un sustituyente tolueno. Ambos compuestos ilustran la heterogeneidad estructural de la base de datos y representan casos de inhibición moderada cuya inclusión garantiza que el modelo QSAR final no esté sesgado hacia un solo grupo, ampliando así su dominio de aplicabilidad.

En conjunto, los diez compuestos reflejan la amplia diversidad estructural y química de la base de datos construida en este trabajo, dicha variedad no es trivial desde el punto de vista del modelado QSAR: la coexistencia de compuestos con distintos perfiles electrónicos desde sistemas ricos en electrones como los derivados de euparina y pirrol, hasta sistemas electrón-deficientes como los derivados de cianuro y tiosemicarbazida y con rangos de pIC_{50} que van desde -2.017 hasta 0.009 , garantiza que el modelo matemático resultante capture relaciones estructura-actividad genuinas y generalizables, minimizando el riesgo de sobreajuste a una sola familia química. Asimismo, la presencia de inhibidores eficaces como **M101** junto con compuestos de actividad moderada o baja como **M118** y **M093** dota al conjunto de una cobertura adecuada del espacio químico relevante, condición indispensable para que el dominio de aplicabilidad del modelo sea suficientemente amplio y con utilidad predictiva real en el diseño racional de nuevos IAGs.

3.2. Dinámicas Moleculares

Las geometrías obtenidas de los SMILES fueron sometidas a un proceso de dinámica molecular, usando DFTB+ [35] con el conjunto de parámetros “3ob”; la simulación se ejecutó durante 100,000 pasos de 2 fs y un enfriamiento gradual de 600 K a 10 K, estas temperaturas fueron seleccionadas para explorar la superficie de energía potencial de cada inhibidor desde una temperatura suficientemente alta como para proporcionar la energía cinética suficiente para generar distintos rotámeros. El proceso antes descrito se realizó para cada una de las moléculas seleccionadas de nuestra base de datos.



Gráfica 1. Trayectoria de la dinámica molecular para **M011**, destacando el punto en rojo, siendo el conformero final.

La Gráfica 1 muestra la trayectoria de energía relativa en función del tiempo de simulación obtenida mediante dinámica molecular con el método DFTB/3ob para el compuesto **M011** (derivado de euparin), empleada como ejemplo representativo del comportamiento general observado para los 157 compuestos de la base de datos, ya que todos presentaron trayectorias cualitativamente similares.

Al inicio de la simulación, con una temperatura de $T = 600\text{ K}$, la energía relativa del sistema se encuentra en valores elevados, del orden de 80–105 kcal/mol, y exhibe fluctuaciones amplias y de alta frecuencia, características de un sistema con suficiente energía térmica para superar barreras conformacionales y explorar de

manera extensa la superficie de energía potencial. Este comportamiento es precisamente el objetivo de la etapa de alta temperatura: evitar que la molécula quede atrapada prematuramente en un mínimo local poco representativo de la geometría real del inhibidor.

A medida que la temperatura disminuye de forma gradual y controlada hacia $T = 10$ K, la energía relativa desciende de manera progresiva y continua, siguiendo una tendencia monótona decreciente con fluctuaciones que se amortiguan conforme avanza la simulación. Este patrón de disipación suave de la energía térmica indica que el esquema de enfriamiento adoptado —100,000 pasos con un tamaño de paso de 2 fs— es numéricamente estable y permite que el sistema relaje hacia conformaciones de menor energía sin que se produzcan saltos abruptos ni discontinuidades en la trayectoria, descartando inestabilidades numéricas o rupturas de enlaces durante la dinámica.

Al final de la trayectoria, con la temperatura próxima a 10 K, la energía relativa converge hacia un valor mínimo cercano a cero en la escala relativa empleada. El punto marcado en rojo corresponde al confórmero seleccionado, que representa una aproximación al confórmero de mínima energía accesible bajo el protocolo empleado. Es importante señalar que, dado el carácter estocástico de la dinámica molecular y la naturaleza del enfriamiento gradual, no se garantiza que la geometría obtenida corresponda al mínimo global absoluto de la superficie de energía potencial, sino a una geometría de baja energía que constituye una aproximación razonable y reproducible al confórmero más estable. Estas geometrías resultantes se emplearon como punto de partida para las etapas posteriores de *Docking* y cálculos DFT, lo cual es metodológicamente aceptable dado que AutoDock 4.2.6 realiza su propia exploración conformacional durante el *Docking* mediante el Algoritmo Genético Lamarckiano.

El recuadro insertado en la gráfica muestra la estructura tridimensional del confórmero final de **M011**, en la que se aprecia la geometría del sistema π extendido del núcleo benzofurano-chalcona y la disposición espacial del grupo difenilamino, confirmando que la estructura obtenida es químicamente razonable y consistente con lo esperado para esta familia de inhibidores.

3.3. Resultados de *Docking*

Las estructuras resultantes de la búsqueda conformacional para los compuestos estudiados, descritas en el apartado 4.2, se utilizaron como ligandos para realizar los estudios de *Docking* proteína-ligando. La estructura del receptor (proteína) utilizada en el presente trabajo (Figura 3) consiste en un modelo estructural de la enzima α -glucosidasa (PDB ID: 3TOP). Este modelo, PDB corresponde al dominio C-terminal de la maltasa-glucoamilasa intestinal humana (MGAM-C) [154]. La elección de 3TOP se sustenta en que se trata de una estructura determinada por difracción de rayos X con resolución adecuada de 2.88 Å y que se encuentra co-cristalizada con acarbosa, un inhibidor de referencia utilizado ya como medicamento contra la diabetes mellitus tipo 2, lo cual permite definir el sitio activo de la enzima.

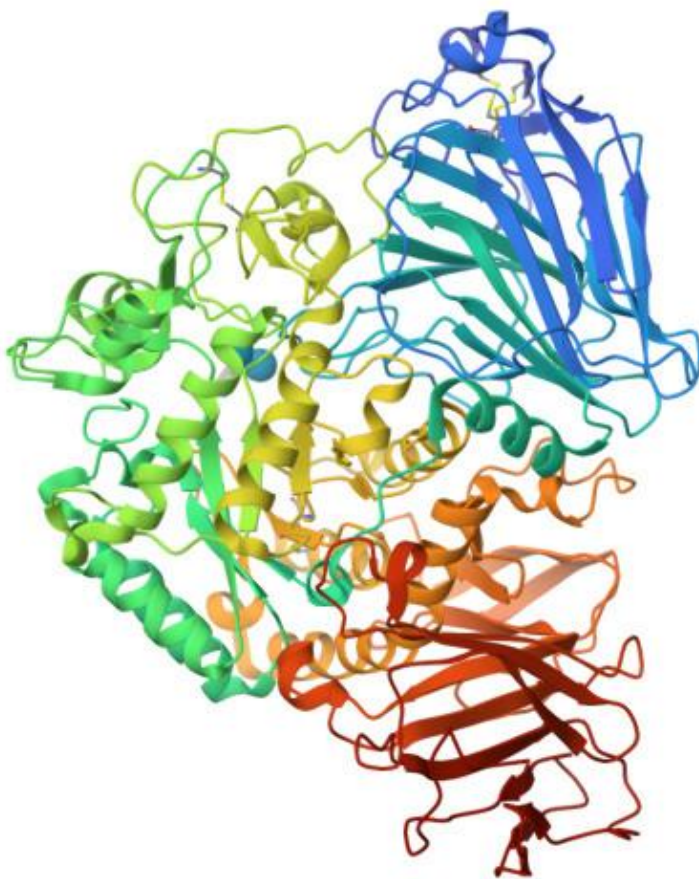
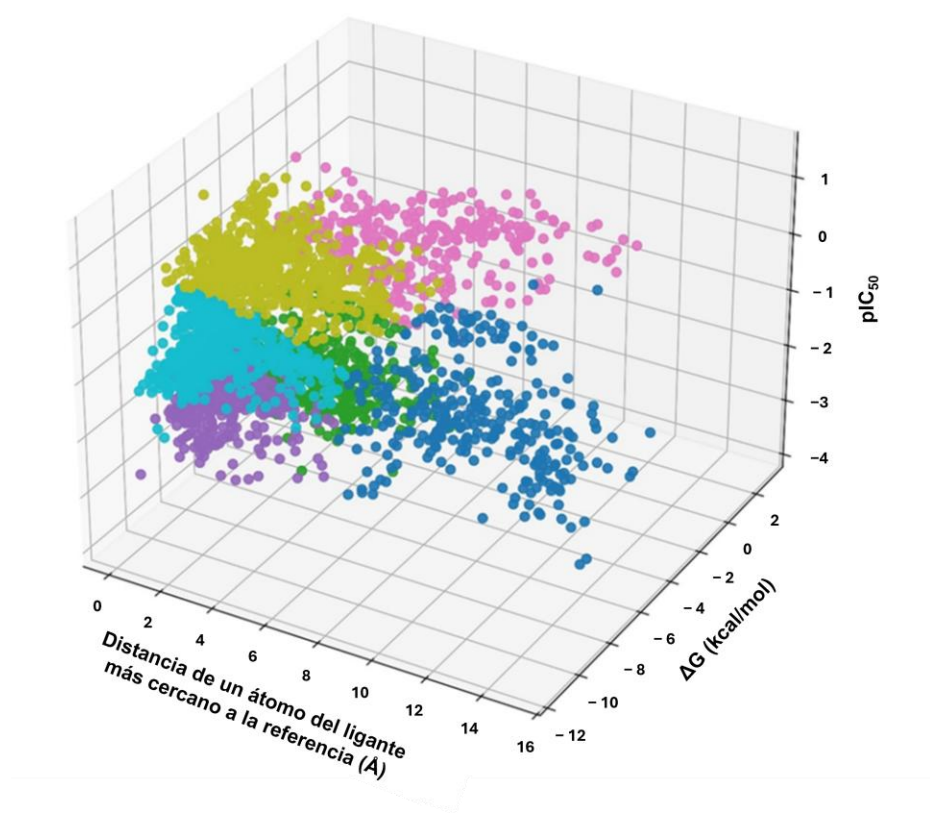


Figura 3. Estructura cristalina de la subunidad C-terminal de la maltasa-glucoamilasa humana en complejo con acarbosa (3TOP). Imagen recuperada del PDB.

3.3.1. Complejos Enzima-Inhibidor

Cada uno de los 100 modelos generados en el cálculo de *Docking* para cada complejo, se sometió a un análisis de agrupamiento para posteriormente filtrar conformaciones equivalentes entre si considerando un valor de corte de RMSD de 2 Å. Finalmente se obtuvieron un total de 3,556 complejos formados entre los 157 inhibidores y la enzima 3TOP.

Para reducir el número de inhibidores utilizados en la generación de modelos QSAR, se hizo un segundo filtrado a partir de los complejos remanentes del paso anterior, el cual consistió en agrupar los complejos con ayuda de *K-means* (explicado anteriormente en el apartado 2.3). A partir de los 3,556 complejos obtenidos se realizó un gráfico para analizar la distribución que estos tendrían, tomando como variables el valor de pIC_{50} , la distancia más cercana entre cualquier átomo del ligante y al centro de masa de la referencia (acarbose) y la energía libre de unión (ΔG), obteniendo como resultado la Gráfica 2. En esta distribución se pueden apreciar 6 cúmulos que permiten diferenciar los complejos que presentan mayor afinidad y menor distancia al sitio de unión de la referencia.



Gráfica 2. Distribución de los frames en 3 dimensiones de cada complejo generado respecto a los valores de pIC_{50} , ΔG (kcal/mol) y la distancia de un átomo del ligante que se encuentra más cercano a la referencia co-cristalizada (Å), agrupado por clústers.

Para la reducción del número de inhibidores, se optó por elegir un solo clúster, por lo que se definieron intervalos para delimitar cada uno de los grupos como se muestra en la Tabla 2.

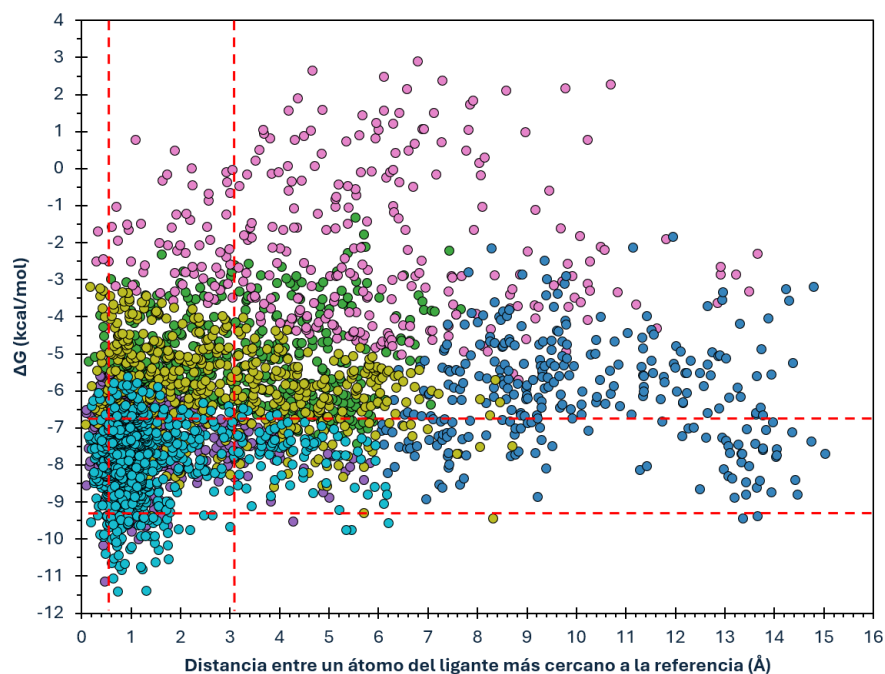
Tabla 2. Intervalos donde se encuentra la mayor cantidad de complejos por cada clúster.

Clúster	Color	Intervalo		
		Distancia a la referencia (Å)	ΔG (kcal/mol)	pIC_{50}
0	Morado	0.489 a 2.166	-8.400 a -6.230	-2.505 a -1.615
1	Azul Claro	0.549 a 3.094	-9.270 a -6.740	-1.500 a -0.651
2	Amarillo	0.642 a 5.322	-7.821 a -4.647	-0.479 a 0.523
3	Verde	1.063 a 5.561	-6.493 a -3.490	-2.505 a -1.387
4	Rosa	2.077 a 9.267	-4.382 a 0.792	-0.479 a 0.684
5	Azul Oscuro	7.211 a 13.438	-7.769 a -4.153	-2.561 a -0.283
		Rango de 0.489 a 13.438	Rango de -9.270 a 0.792	Rango de -2.505 a 0.684

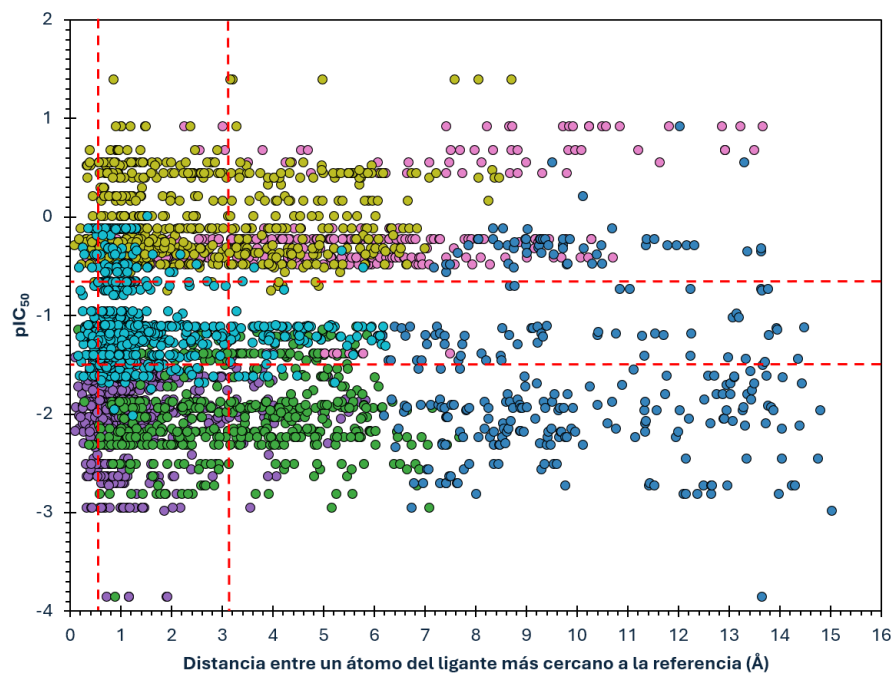
Como podemos observar en los intervalos, el clúster que tiene mayor cercanía con la referencia es el número 0 o de color morado, por otro lado el clúster que presenta una energía libre de unión más favorable es el número 1 o de color azul claro y finalmente el que presenta una mejor pIC_{50} es el clúster número 4 o de color rosa.

Para la elección del clúster y con los intervalos establecidos, el conjunto que cumplía de mejor manera con los 3 intervalos era el clúster 1, ya que a pesar de que el clúster 0 tenía mejor cercanía con la referencia y energía libre de unión favorable, sus valores de pIC_{50} eran bastante bajos. El clúster número 4 también fue descartado ya que a pesar de contar con valores más altos de pIC_{50} y cercanía con la referencia, el intervalo de su energía libre de unión no era favorable, en comparación con el clúster 1.

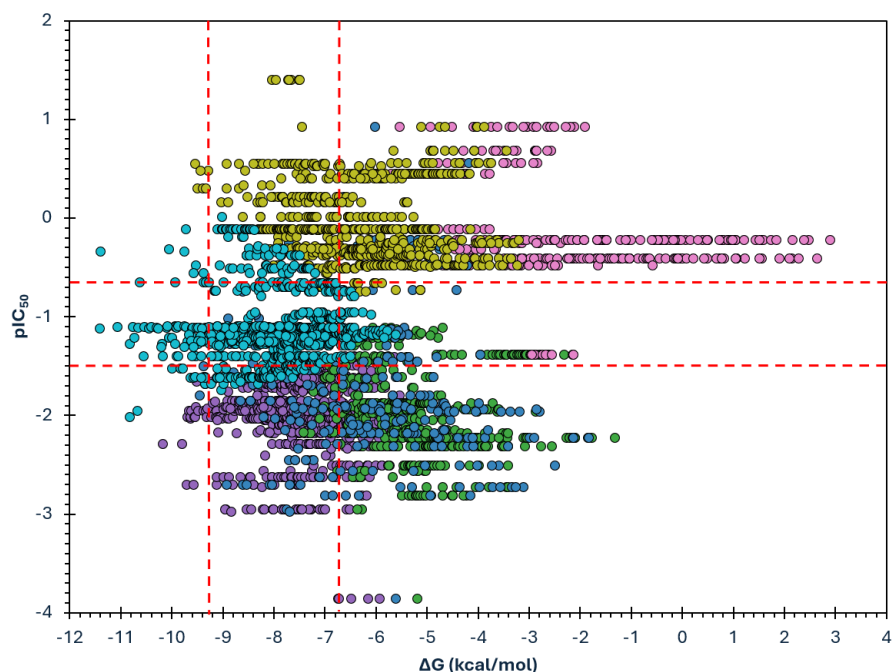
En las gráficas 3, 4 y 5 se presenta la distribución en 2 dimensiones de los complejos agrupados por clúster, destacando los intervalos del grupo seleccionado para el cálculo de los modelos QSAR.



Gráfica 3. Distribución en 2 dimensiones de cada complejo generado respecto a los valores de ΔG (kcal/mol) y la distancia de un átomo del ligante que se encuentra más cercano a la referencia co-cristalizada (Å), agrupado por clústers.



Gráfica 4. Distribución en 2 dimensiones de cada complejo generado respecto a los valores de pIC_{50} y la distancia de un átomo del ligante que se encuentra más cercano a la referencia co-cristalizada (Å), agrupado por clústers.



Gráfica 5. Distribución en 2 dimensiones de cada complejo generado respecto a los valores de ΔG (kcal/mol) y pIC_{50} , agrupado por clústers.

En la Tabla 3 se integró un conjunto de 74 complejos provenientes del clúster 1, para los cuales se emplearon los ligantes como base para el desarrollo de los cálculos para modelos QSAR, destacando los que tienen alto (color verde), medio (color amarillo) y bajo (color rojo) valor de pIC_{50} .

Tabla 3. Complejos seleccionados a partir de los clústers que cumplían mejor con los 3 intervalos propuestos (distancia a la referencia (Å), ΔG (kcal/mol) y pIC_{50})

Complejo	pIC_{50}	Distancia a la referencia (Å)	LRCD	ΔG (kcal/mol)	Complejo	pIC_{50}	Distancia a la referencia (Å)	LRCD	ΔG (kcal/mol)
M002_3TOPC-AC13_0	-1.654	0.611	0.696	-8.650	M060_3TOPC-AC13_0	-0.651	1.490	0.788	-10.630
M003_3TOPC-AC13_0	-1.605	1.290	0.706	-7.980	M061_3TOPC-AC13_0	-0.513	0.699	0.757	-8.860
M008_3TOPC-AC13_0	-0.955	0.831	0.682	-8.370	M062_3TOPC-AC13_0	-0.738	0.972	0.853	-9.160
M011_3TOPC-AC13_0	-1.201	0.591	0.677	-9.430	M066_3TOPC-AC13_0	-0.125	0.904	0.811	-8.390
M013_3TOPC-AC13_0	-1.097	0.703	0.687	-10.080	M082_3TOPC-AC13_0	-1.303	1.017	0.652	-10.210
M017_3TOPC-AC13_0	-0.663	0.938	0.726	-7.860	M083_3TOPC-AC13_0	-1.292	0.526	0.650	-9.120
M018_3TOPC-AC13_1	-0.748	1.211	0.818	-7.570	M086_3TOPC-AC13_0	-0.312	0.247	0.721	-8.490
M019_3TOPC-AC13_0	-1.193	0.588	0.853	-7.960	M087_3TOPC-AC13_0	-1.115	1.478	0.705	-10.100
M020_3TOPC-AC13_0	-1.539	0.656	0.883	-8.570	M093_3TOPC-AC13_0	-1.745	1.596	0.769	-9.030
M021_3TOPC-AC13_0	-1.164	0.946	0.749	-8.400	M095_3TOPC-AC13_0	-1.182	1.166	0.780	-6.900
M022_3TOPC-AC13_1	-1.500	0.833	0.715	-8.990	M099_3TOPC-AC13_0	-1.113	1.532	0.766	-8.500
M023_3TOPC-AC13_1	-1.471	1.062	0.717	-8.020	M100_3TOPC-AC13_0	-1.452	0.747	0.775	-7.150
M024_3TOPC-AC13_0	-0.978	1.199	0.705	-8.450	M101_3TOPC-AC13_0	0.009	1.524	0.793	-9.010
M025_3TOPC-AC13_0	-1.529	0.902	0.734	-8.190	M110_3TOPC-AC13_0	-0.792	0.252	0.804	-8.250
M026_3TOPC-AC13_0	-1.423	0.935	0.850	-8.120	M112_3TOPC-AC13_0	-1.277	0.807	0.711	-10.830
M027_3TOPC-AC13_0	-1.494	0.436	0.772	-8.220	M113_3TOPC-AC13_0	-0.188	0.633	0.697	-8.900
M028_3TOPC-AC13_1	-1.021	0.688	0.706	-8.500	M114_3TOPC-AC13_0	-1.146	0.411	0.755	-7.510
M029_3TOPC-AC13_1	-1.563	0.655	0.860	-8.550	M115_3TOPC-AC13_0	-0.114	0.814	0.854	-9.720
M030_3TOPC-AC13_0	-1.270	1.229	0.707	-8.510	M117_3TOPC-AC13_0	-1.117	0.734	0.706	-11.410
M031_3TOPC-AC13_0	-1.290	0.499	0.815	-8.310	M118_3TOPC-AC13_0	-2.017	1.230	0.748	-10.820
M032_3TOPC-AC13_1	-1.127	0.635	0.786	-8.530	M119_3TOPC-AC13_0	-0.699	0.546	0.856	-8.660
M034_3TOPC-AC13_0	-0.342	1.314	0.624	-11.390	M121_3TOPC-AC13_0	-1.207	0.261	0.752	-8.430
M035_3TOPC-AC13_1	-0.380	1.238	0.811	-8.530	M122_3TOPC-AC13_0	-0.561	0.932	0.667	-9.580
M036_3TOPC-AC13_0	-1.401	0.790	0.769	-9.180	M127_3TOPC-AC13_0	-0.310	0.767	0.773	-8.170
M037_3TOPC-AC13_0	-1.163	1.481	0.715	-8.160	M128_3TOPC-AC13_0	-0.344	0.906	0.793	-7.950
M038_3TOPC-AC13_0	-0.730	0.835	0.854	-6.800	M130_3TOPC-AC13_1	-1.441	0.573	0.792	-7.220
M041_3TOPC-AC13_0	-1.104	0.650	0.793	-11.060	M135_3TOPC-AC13_0	-1.410	0.889	0.900	-8.350
M042_3TOPC-AC13_0	-1.607	0.192	0.787	-9.540	M137_3TOPC-AC13_0	-1.547	1.020	0.756	-8.160

M043_3TOPC-AC13_0	-1.253	0.549	0.702	-9.690	M138_3TOPC-AC13_0	-1.288	0.583	0.719	-7.310
M046_3TOPC-AC13_0	-1.114	0.606	0.736	-10.650	M139_3TOPC-AC13_0	-1.340	0.665	0.713	-7.830
M047_3TOPC-AC13_0	-1.215	0.771	0.803	-10.730	M142_3TOPC-AC13_0	-1.229	1.027	0.745	-7.860
M048_3TOPC-AC13_0	-1.956	0.870	0.731	-10.670	M143_3TOPC-AC13_0	-1.333	1.185	0.735	-8.010
M050_3TOPC-AC13_0	-1.616	1.150	0.748	-9.630	M145_3TOPC-AC13_0	-1.215	0.521	0.755	-7.580
M051_3TOPC-AC13_0	-1.681	1.074	0.746	-9.350	M146_3TOPC-AC13_0	-1.062	0.991	0.786	-7.170
M057_3TOPC-AC13_0	-1.527	0.820	0.708	-9.990	M152_3TOPC-AC13_0	-1.182	3.011	0.688	-9.670
M058_3TOPC-AC13_0	-1.401	1.761	0.650	-10.55	M153_3TOPC-AC13_0	-1.246	1.198	0.686	-9.620
M059_3TOPC-AC13_0	-1.095	1.383	0.731	-9.930	M156_3TOPC-AC13_0	-3.283	0.691	0.761	-8.310

Tabla 3. Continuación

3.4. Modelos QSAR

Se calculó un total de 3,596 descriptores, entre los que se incluyeron los parámetros derivados de los cálculos de DFT, así como descriptores topológicos; de manera particular índices de conectividad y autocorrelación; descriptores basados en propiedades atómicas, cargas, superficie y volumen molecular; conteos de fragmentos estructurales y descriptores tridimensionales relacionados con la forma y distribución espacial de los átomos. Se importó la matriz construida en el programa QSAR-INS, donde se excluyeron un total de 2006 descriptores que presentaban una correlación mayor o igual a 0.95 con otro descriptor, teniendo una matriz final de 1590 descriptores y utilizando como variable de respuesta los valores de pIC_{50} , el conjunto final evaluado incluyó 74 moléculas que presentan propiedades como inhibidores, se distribuyeron en un conjunto de 60 compuestos de entrenamiento y otro conjunto de 14 compuestos de predicción externa. Para los GA se utilizó una población de 300 modelos, una tasa de mutación de 40% y un total de 3000 generaciones, se seleccionó Q_{Loo}^2 como la función de evaluación [155] y se estableció como máximo 10 descriptores por ecuación.

Una vez concluido el cálculo de los modelos, se buscó aquel que cumplía con los parámetros estadísticos de validación interna y externa, por ejemplo, en la Tabla 4 se observa que al aumentar el número de los descriptores seleccionados, mejora la capacidad de ajuste y predicción estadística. Los primeros 2 modelos presentaron valores de R^2 y Q_{Loo}^2 relativamente bajos, indicando que las variables incluidas no eran suficientes para describir adecuadamente la relación estructura-actividad, en contraste, el modelo 3 presentó una $R^2 = 0.672$, que es considerada aceptable, pero su valor de $Q_{Loo}^2 = 0.466$ evidenciaba que el modelo no cumplía con los parámetros establecidos. El Modelo 4 presentó los mejores parámetros estadísticos, alcanzando un valor de $R^2 = 0.747$ y $Q_{Loo}^2 = 0.621$, lo cual indica una correlación adecuada entre los valores experimentales y calculados, este modelo final se construyó utilizando diez descriptores moleculares seleccionados automáticamente mediante algoritmos genéticos, la selección indica que la actividad inhibitoria depende de múltiples factores estructurales, en lugar de una sola propiedad molecular.

La comparación entre los parámetros de validación de los cuatro modelos obtenidos demuestra que la predicción de la pIC_{50} experimental se puede realizar únicamente con descriptores topológicos, esto debido a que no se encontró un modelo validado con descriptores provenientes de DFT.

Tabla 4. Modelos QSAR obtenidos.

Modelo	Descriptores	R^2	R^2_{adj}	Q^2_{LOO}	Q^2_{LMO}	$R^2 - Q^2_{LOO}$	r^2_m	r'^2_m	$\Delta r^2_{m(LOO)}$	$\bar{r}^2_{m(LOO)}$	¿Cumple?
Modelo 1	SMR_VSA6, fr_nitrile, GATS2e, P_VSA_m_5, F09[O-O]	0.474	0.423	0.410	0.352	0.064	0.392	0.106	0.285	0.249	No
Modelo 2	Chi3v, SpMax_L, VE1_H2, nArOR, B02[N-S], B03[N-O]	0.591	0.543	0.407	0.370	0.184	0.365	0.208	0.158	0.286	No
Modelo 3	SlogP_VSA4, SpPosA_Dz(m), VE1_B(e), MATS1m, P_VSA_LogP_2, RDF100m, HATS3s F07[O-S]	0.672	0.618	0.466	0.417	0.206	0.416	0.288	0.129	0.352	No
Modelo 4	BalabanJ, SlogP_VSA4, nDB, SM15_EA(bo), RDF020i, Ds, NaaO, B08[C-S], B09[C-O], F07[C-N]	0.747	0.694	0.621	0.574	0.126	0.578	0.425	0.153	0.501	Sí

Tabla 5. Descriptores resultantes del Modelo 4 para cada compuesto con su respectivo valor de pIC_{50} .

Compuesto	pIC_{50}	BalabanJ	SlogP_VSA4	nDB	SM15_EA(bo)	RDF020i	Ds	NaaO	B08[C-S]	B09[C-O]	F07[C-N]
M002	-1.654	2.445	6.852	3	22.883	5.383	0.374	1	0	1	0
M003	-1.605	2.417	6.852	3	22.801	5.359	0.366	1	0	1	0
M008	-0.955	2.445	0.000	3	23.000	5.407	0.519	1	0	1	0
M011	-1.201	1.846	0.000	3	22.977	7.464	0.419	1	0	1	1
M013	-1.097	1.960	0.000	3	23.771	6.963	0.432	1	0	1	0
M017	-0.663	1.740	5.817	0	22.620	3.565	0.393	1	1	0	5
M018	-0.748	1.736	0.000	0	22.548	3.609	0.496	1	1	0	5
M019	-1.193	1.765	10.114	2	23.076	3.459	0.466	1	1	1	5
M020	-1.539	1.765	10.114	2	22.766	3.517	0.317	1	1	1	5
M021	-1.164	1.734	10.114	2	22.642	3.357	0.458	1	1	1	6
M022	-1.500	1.779	6.852	0	22.786	3.343	0.304	1	1	0	5
M023	-1.471	1.736	6.852	0	22.541	3.404	0.414	1	1	0	5
M024	-0.978	1.726	5.817	0	22.504	3.113	0.393	1	1	0	5
M025	-1.529	1.757	6.852	0	22.447	3.323	0.374	1	1	0	5
M026	-1.423	1.801	0.000	0	22.674	8.297	0.325	1	1	1	5
M027	-1.494	1.759	0.000	0	22.640	5.853	0.336	1	1	0	5
M028	-1.021	1.726	5.817	0	22.447	3.074	0.432	1	1	0	5
M029	-1.563	1.733	0.000	0	22.529	5.813	0.353	1	1	1	5
M030	-1.270	1.726	0.000	0	22.573	2.987	0.369	1	1	0	5
M031	-1.290	1.765	10.114	2	22.858	3.103	0.494	1	1	1	5
M032	-1.127	1.740	0.000	0	23.142	3.179	0.410	1	1	0	5
M034	-0.342	2.043	0.000	2	22.402	17.728	0.473	0	0	1	0
M035	-0.380	2.063	6.852	2	22.712	9.044	0.442	0	1	1	12
M036	-1.401	1.518	0.000	1	22.640	12.229	0.504	0	0	1	13
M037	-1.163	1.907	12.669	1	23.099	9.002	0.447	0	0	1	11
M038	-0.730	2.641	0.000	0	23.038	8.283	0.426	0	0	0	0
M041	-1.104	1.638	0.000	3	22.853	8.952	0.342	0	0	1	12
M042	-1.607	1.610	5.817	3	22.611	8.306	0.353	0	0	1	12
M043	-1.253	1.899	5.817	0	22.882	9.255	0.401	0	0	1	7
M046	-1.114	1.546	0.000	4	23.077	13.956	0.414	0	0	1	6
M047	-1.215	1.605	6.852	4	23.047	16.438	0.352	0	0	1	7
M048	-1.956	1.566	13.704	0	22.483	8.082	0.427	0	0	1	7
M050	-1.616	1.674	6.852	2	22.508	12.548	0.364	0	0	1	9
M051	-1.681	1.691	10.114	4	22.204	7.618	0.476	0	0	1	7
M057	-1.527	1.903	10.114	2	23.031	6.234	0.490	0	0	1	20
M058	-1.401	1.899	0.000	0	22.987	6.170	0.413	0	0	0	18
M059	-1.095	1.901	6.852	1	23.253	8.754	0.455	0	0	0	25
M060	-0.651	1.818	10.114	5	22.845	8.626	0.398	0	1	1	6
M061	-0.513	1.878	10.114	5	23.177	5.522	0.469	0	0	1	13
M062	-0.738	1.866	10.114	5	22.891	4.947	0.421	0	0	1	13
M066	-0.125	2.183	6.852	6	23.059	9.657	0.487	0	0	1	3
M082	-1.303	2.399	56.545	3	21.835	20.144	0.474	0	0	1	0
M083	-1.292	2.122	67.777	2	22.716	16.989	0.506	0	0	1	0

M086	-0.312	2.390	0.000	2	23.001	9.578	0.360	0	0	1	2
M087	-1.115	2.141	5.894	4	23.332	11.355	0.472	0	0	1	12
M093	-1.745	2.594	22.662	0	22.778	4.720	0.420	0	0	1	8
M095	-1.182	2.637	11.331	1	21.146	5.176	0.458	0	0	1	0
M099	-1.113	2.235	6.852	4	21.473	3.611	0.486	0	0	1	0
M100	-1.452	2.610	17.225	3	22.170	5.836	0.353	0	0	1	4
M101	0.009	2.268	0.000	6	22.070	15.351	0.462	0	1	1	5
M110	-0.792	2.182	0.000	2	22.588	7.393	0.428	0	0	1	2
M112	-1.277	1.727	0.000	2	23.637	9.531	0.439	0	0	1	9
M113	-0.188	2.173	17.049	2	22.780	10.945	0.370	0	1	0	8
M114	-1.146	2.049	11.634	2	22.424	10.049	0.415	0	1	1	6
M115	-0.114	1.707	0.000	2	23.612	11.618	0.443	0	1	0	16
M117	-1.117	2.038	0.000	1	22.091	5.996	0.341	0	0	0	14
M118	-2.017	1.718	0.000	1	23.692	11.953	0.242	0	0	1	12
M119	-0.699	2.115	11.788	3	23.224	5.753	0.413	0	1	1	6
M121	-1.207	2.410	45.715	2	21.015	11.366	0.495	0	0	1	0
M122	-0.561	1.832	5.894	3	23.080	19.287	0.445	0	1	1	10
M127	-0.310	2.409	13.704	1	23.038	3.936	0.526	0	0	0	3
M128	-0.344	2.409	13.704	1	22.851	4.076	0.557	0	0	0	3
M130	-1.441	2.138	0.000	2	21.894	5.255	0.353	0	0	1	0
M135	-1.410	2.208	11.788	1	22.302	10.811	0.457	0	0	1	0
M137	-1.547	2.582	0.000	3	21.008	5.954	0.374	0	0	1	0
M138	-1.288	2.565	0.000	3	21.144	5.602	0.463	0	0	1	0
M139	-1.340	2.565	0.000	3	20.485	6.190	0.397	0	0	1	0
M142	-1.229	1.928	6.852	1	23.525	2.662	0.296	0	0	1	1
M143	-1.333	1.942	0.000	1	23.840	2.915	0.327	0	0	1	2
M145	-1.215	2.637	6.852	2	21.888	2.953	0.420	0	0	1	0
M146	-1.062	2.563	0.000	2	21.700	3.263	0.479	0	0	1	0
M152	-1.182	1.956	6.852	1	23.851	11.869	0.345	0	0	1	27
M153	-1.246	1.956	6.852	1	23.851	13.147	0.398	0	0	1	27
M156	-3.283	2.737	0.000	0	23.239	11.093	0.530	0	0	1	0

Tabla 5. Continuación

*La definición e interpretación de cada descriptor obtenido en el Modelo 4 se presentan en el apartado de **Anexos**.

3.4.1. Descriptores Estructurales del Modelo Obtenido

En la Tabla 5 muestra los valores individuales de cada descriptor para las moléculas estudiadas, puede observarse que existe una importante diversidad estructural entre los compuestos, lo cual favorece la construcción de modelos QSAR más robustos debido a la variabilidad química del conjunto.

La exclusión de la molécula **M156** fue necesaria ya que se realizó una detección preliminar de valores atípicos en la variable experimental, estableciendo en los límites superior e inferior ± 3 desviaciones estándar (σ). Las observaciones fuera de este intervalo se consideraron atípicas y se evaluó su exclusión del conjunto final.

3.4.2. Resultados Estadísticos del Modelo

La ecuación para el Modelo 4 es la siguiente:

$$\begin{aligned} pIC_{50} = & -9.3854 + 0.3581 \text{ BalabanJ} - 0.0128 \text{ SlogP_VSA4} + 0.1141 \text{ nDB} \\ & + 0.2855 \text{ SM15_EA(bo)} + 0.0242 \text{ RDF020i} + 2.8277 \text{ Ds} \\ & - 0.4664 \text{ NaaO} + 0.4929 \text{ B08[C - S]} - 0.4157 \text{ B09[C - O]} \\ & - 0.0169 \text{ F07[C - N]} \end{aligned} \quad \text{Ecuación 10}$$

La ecuación QSAR obtenida describe matemáticamente cómo cada descriptor molecular contribuye al valor final de pIC_{50} . Los coeficientes positivos indican que un aumento en el descriptor favorece la actividad inhibitoria, mientras que los coeficientes negativos sugieren una disminución de la actividad.

La combinación simultánea de descriptores topológicos, geométricos y electrónicos demuestra que la inhibición de α -glucosidasa es un fenómeno multifactorial. Esto significa que la afinidad inhibitoria no depende únicamente de la presencia de ciertos grupos funcionales, sino también de la forma molecular, distribución electrónica y propiedades fisicoquímicas globales.

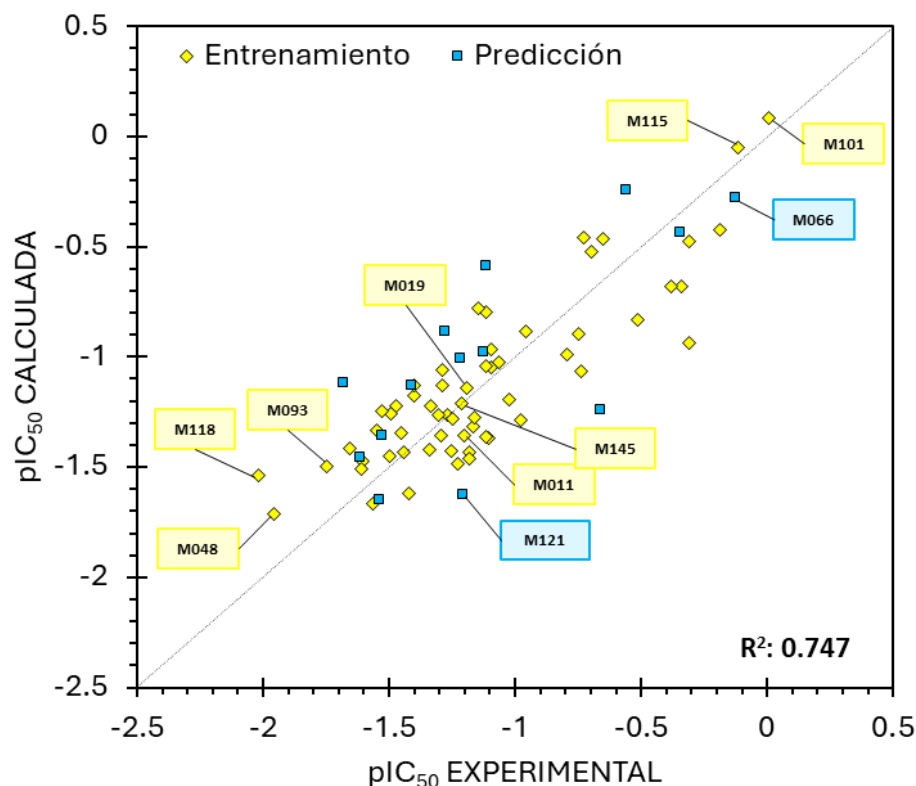
En la Gráfica 6 se presenta la correlación entre los valores experimentales de pIC_{50} y los valores calculados por el modelo QSAR. La línea diagonal representa el ajuste ideal, es decir, el escenario en el cual los valores predichos coinciden exactamente con los experimentales. El método estadístico para determinar cuantitativamente la cercanía que tienen estos datos a la recta, es el conocido coeficiente de determinación (R^2), el cual alcanza un valor de 1.0 si hay un ajuste perfecto, mientras que valores cercanos a cero nos indican que el modelo no tiene correlación con los datos experimentales. El modelo fue seleccionado debido a que supera el valor de R^2 de 0.747, el cual de acuerdo con *Goldbraik et al.* [156], es aceptable para modelos con capacidad predictiva.

Los rombos amarillos corresponden al conjunto de entrenamiento, mientras que los cuadrados azules representan el conjunto de validación externa, la distribución de

los puntos alrededor de la diagonal permite evaluar visualmente la calidad predictiva del modelo, se observa que la mayoría de los compuestos se agrupan relativamente cerca de la línea diagonal, indicando que el modelo reproduce adecuadamente la tendencia experimental, la proximidad de los puntos respecto a la diagonal implica errores de predicción relativamente pequeños. Los compuestos alejados de la diagonal representan moléculas cuya actividad fue subestimada o sobreestimada por el modelo, sin embargo, incluso estos compuestos mantienen una tendencia general coherente, lo que indica que el modelo conserva estabilidad estadística. La dispersión moderada observada es esperable en sistemas biológicos complejos, debido a que la actividad inhibitoria puede verse influenciada por factores experimentales adicionales que no son completamente descritos mediante los descriptores moleculares utilizados.

Entre las moléculas señaladas en la Gráfica 6, la molécula **M011** presenta una desviación moderada entre el valor experimental y el calculado. La posición observada sugiere que el modelo describe razonablemente bien su actividad, aunque con cierta limitación asociada posiblemente a propiedades geométricas o electrónicas particulares.

Las moléculas **M019**, y **M115** destacan por encontrarse muy cercana a la diagonal ideal. Esto indica que el modelo reproduce adecuadamente su valor experimental de pIC_{50} , evidenciando que las propiedades estructurales responsables de su actividad se encuentran correctamente representadas por los descriptores seleccionados.



Gráfica 6. Modelo QSAR con pIC_{50} calculada vs. pIC_{50} experimental, donde los rombos amarillos pertenecen al conjunto de entrenamiento y los cuadrados azules al conjunto de predicción.

Por otro lado, la molécula **M066**, perteneciente al conjunto de predicción externa, presenta una desviación moderada respecto a la diagonal ideal. Aunque el modelo no reproduce exactamente su valor experimental, la proximidad relativa a la línea de correlación demuestra que conserva capacidad predictiva sobre compuestos externos. Esta desviación puede atribuirse a diferencias estructurales particulares dentro del conjunto de entrenamiento.

La molécula **M121** corresponde a uno de los compuestos con mayor desviación respecto a la diagonal ideal dentro del conjunto de predicción externa. Este resultado evidencia que el modelo tiene dificultades para describir completamente su comportamiento biológico. La discrepancia observada puede atribuirse a una combinación de factores, como propiedades electrónicas poco representadas, elevada flexibilidad conformacional o patrones estructurales únicos dentro del conjunto estudiado. Sin embargo, aunque presenta uno de los mayores errores de predicción, la molécula continúa siguiendo la tendencia general de correlación del modelo.

M145 presenta una adecuada concordancia entre el valor experimental y el valor calculado de pIC_{50} , observándose relativamente cercana a la diagonal ideal de correlación. Este comportamiento indica que el modelo QSAR reproduce

satisfactoriamente su actividad inhibitoria y que las propiedades estructurales responsables de su comportamiento biológico se encuentran correctamente representadas mediante los descriptores seleccionados.

La molécula **M048** y **M093** son de las más alejadas de la diagonal ideal en la región inferior izquierda de la gráfica. Este comportamiento indica un error de predicción considerable, donde el modelo no logra reproducir adecuadamente su actividad inhibitoria experimental. La discrepancia observada podría deberse a características estructurales poco representadas dentro del conjunto de moléculas.

Con el propósito de integrar la información experimental, predictiva y estructural, se compararon los compuestos ubicados en los extremos de actividad biológica. El compuesto **M101** ($pIC_{50} = 0.009$, $\Delta G = -9.010 \text{ kcal/mol}$) fue identificado como uno de los inhibidores con mayor actividad dentro del conjunto analizado, al presentar el valor de pIC_{50} más alto. Este comportamiento fue reproducido adecuadamente por el modelo QSAR, ya que el valor predicho fue muy cercano al experimental. Como se describió previamente, **M101** corresponde a un derivado de éster, cuyo modo de unión en el sitio activo de MGAM-C, ilustrado en la Figura 4, representa uno de los casos más informativos, debido a que combina una energía libre de unión altamente favorable con la mayor actividad experimental observada en el conjunto estudiado.

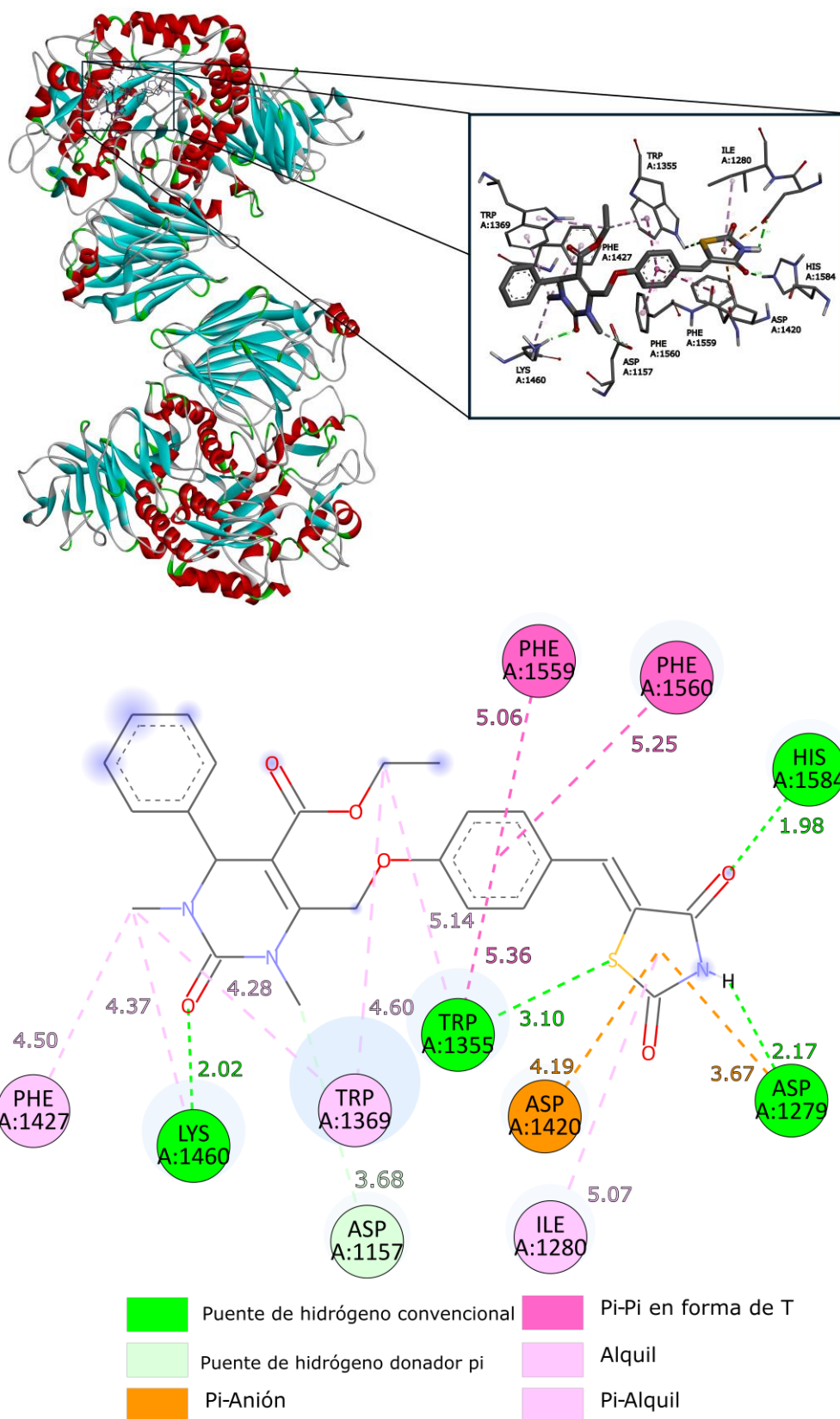


Figura 4. Diagrama 3D y 2D de las interacciones moleculares del Docking de **M101** sobre la enzima α - glucosidasa (3TOP).

El análisis del *Docking* muestra que **M101** tiene interacciones de puentes de hidrógeno convencionales, entre ellos destaca la interacción con **His1584** a 1.98 Å, una distancia corta que sugiere una contribución importante a la estabilidad del complejo, donde el grupo carbonilo de la tiazolidinona actúa como aceptor frente al imidazol de la histidina. Asimismo, se observa una interacción con **Asp1279**, donde el grupo N–H de la tiazolidinona puede actuar como donador hacia el grupo carboxilato del aspartato. Esta interacción resulta particularmente relevante, ya que, como se observa en la Figura 5, el MPE muestra una región electropositiva asociada al hidrógeno unido al nitrógeno, representada por zonas azules, lo que favorece la interacción por residuos cargados negativamente como los aspartatos.

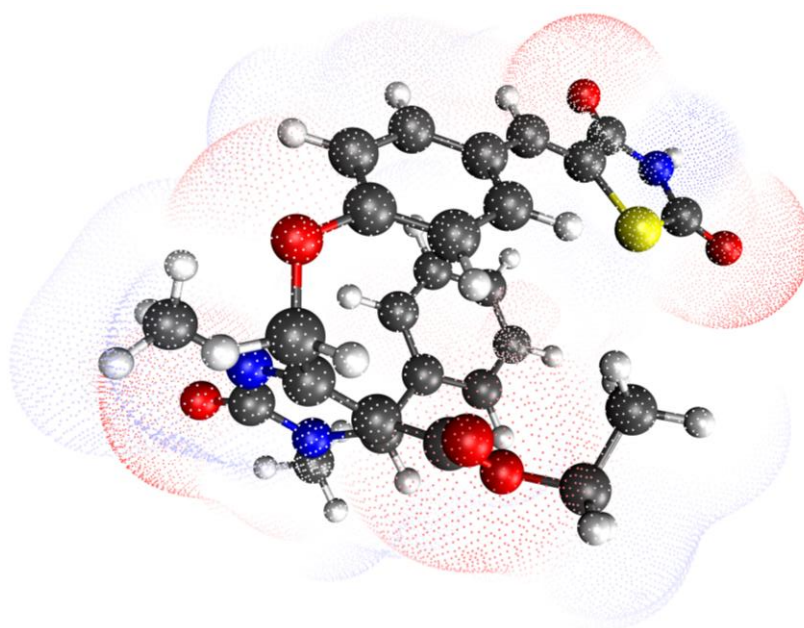


Figura 5. Mapa del potencial electrostático de **M101**

De manera complementaria, **M101** establece una interacción con **Trp1355** a 3.10 Å, atribuible a un contacto N–H⋯S, en el cual el grupo N–H del anillo indólico del triptófano actúa como donador y el átomo de azufre de la tiazolidinona como aceptor débil. Desde el punto de vista del potencial electrostático, la participación del átomo de azufre puede explicarse por la polarizabilidad de su nube electrónica y por la presencia de densidad electrónica disponible asociada a sus pares libres, características que le permiten actuar como un centro aceptor débil y participar en interacciones no covalentes [157]. También se observa un puente de hidrógeno con **Lys1460** a 2.02 Å, en el que uno de los oxígenos del carbonilo del ligando actúa

como aceptor frente a la cadena lateral de la lisina. Esta interacción concuerda con el MPE, donde los átomos de oxígeno de los grupos carbonilo aparecen como regiones de potencial negativo, representadas en rojo con capacidad aceptora de puentes de hidrógeno.

Los contactos electrostáticos aromáticos se amplían mediante interacciones π -anión con **Asp1420** a 4.19 Å y con **Asp1279** a 3.67 Å. Esta última es particularmente importante, ya que como se mencionó antes **Asp1279** también participa en un puente de hidrógeno directo con el ligando y mediante una interacción π -anión con el sistema aromático. La interpretación del MPE permite justificar este comportamiento, ya que los sistemas aromáticos de **M101** presentan una distribución electrostática parcialmente polarizada; aunque predominan regiones hidrofóbicas y de baja polaridad, la presencia de heteroátomos cercanos modifica la densidad electrónica del sistema π , generando zonas capaces de interactuar favorablemente con residuos aniónicos como **Asp1279** y **Asp1420**.

Las interacciones hidrofóbicas y aromáticas de **M101** se extiende mediante contactos no covalentes, incluyendo interacciones π - π , π -alquil y contactos hidrofóbicos. En este sentido, destacan las interacciones π - π tipo conformación T con **Phe1559** a 5.06 Å, **Phe1560** a 5.25 Å y **Trp1355** a 5.36 Å, así como una interacción π -alquil con **Trp1369** a 4.60 Å. Estas interacciones reflejan el acomodo del sistema aromático del ligando dentro de una región rica en residuos aromáticos, particularmente triptófano y fenilalanina, cuyas cadenas laterales pueden establecer interacciones π - π y contactos hidrofóbicos que contribuyen a la estabilización de fragmentos aromáticos en complejos proteína-ligando [149] , [158]. En el MPE estas regiones aromáticas muestran una superficie menos polar en comparación con los heteroátomos del ligando, lo que favorece su inserción en zonas hidrofóbicas del sitio activo.

Asimismo, se observa un contacto alquil con **Lys1460** a 4.37 Å, que complementa el puente de hidrógeno previamente descrito con este residuo, añadiendo un componente hidrofóbico adicional a la interacción molecular. La estabilización periférica se amplía mediante interacciones π -alquil con **Phe1427** a 4.50 Å, **Ile1280** a 5.07 Å y residuos de triptófano a 5.14 y 4.28 Å. Estas distancias se encuentran dentro de los intervalos empleados comúnmente en análisis de *Docking* para identificar contactos π - π y π -alquil. Desde el MPE, estas interacciones pueden asociarse con las superficies aromáticas y alifáticas de menor polaridad de **M101**, las cuales favorecen el contacto hidrofóbico dentro del sitio activo.

Un aspecto estructuralmente relevante de **M101** es que no establece una interacción directa con **Asp1526**, residuo catalítico ácido-base principal del subsitio -1 de MGAM-C [159]. Este hallazgo sugiere que **M101** adopta un modo de unión

desplazado hacia subsitios más externos o hacia una región lateral del sitio activo. Sin embargo, la ausencia de contacto con **Asp1526** parece ser compensada por un conjunto de interacciones con **His1584**, **Asp1279**, **Lys1460**, **Trp1355**, **Trp1369** y otros residuos hidrofóbicos periféricos. Esta complementariedad entre la distribución electrostática del ligando y el entorno químico del sitio activo contribuye a explicar su energía libre de unión favorable ($\Delta G = -9.010 \text{ kcal/mol}$) y su alta actividad experimental, reflejada en el mayor valor de pIC_{50} dentro del conjunto analizado.

En contraste, el compuesto **M118** ($pIC_{50} = -2.017$, $\Delta G = -10.630 \text{ kcal/mol}$), correspondiente al valor de pIC_{50} más bajo del conjunto analizado, presentando una menor actividad inhibitoria experimental. Este compuesto se aleja considerablemente de la diagonal ideal entre los valores experimentales y predichos, lo que indica que el modelo subestima parcialmente su actividad biológica y que, experimentalmente, la molécula muestra un comportamiento distinto al esperado con base en sus propiedades estructurales calculadas. **M118** corresponde a un derivado que contiene un sistema hidrazida con enlace de Schiff, cuyo modo de unión en el sitio activo de MGAM-C, ilustrado en la Figura 6, representa uno de los casos más interesantes del conjunto analizado: aunque presenta una de las energías libres de unión más favorables, también corresponde al compuesto con el pIC_{50} más bajo del conjunto, es decir, al inhibidor menos activo.

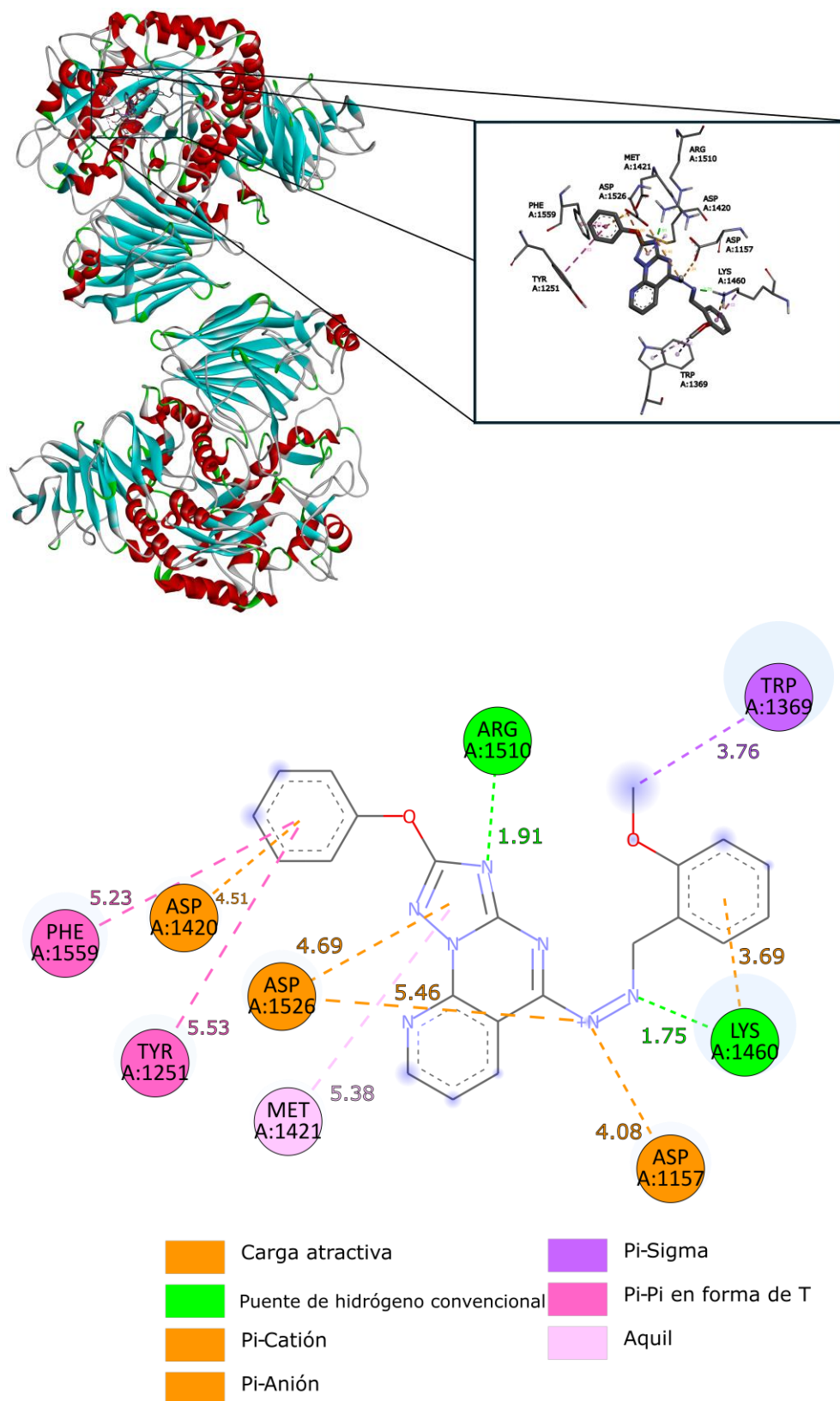


Figura 6. Diagrama 3D y 2D de las interacciones moleculares del Molecular Docking de **M118** sobre la enzima α -glucosidasa (3TOP).

La interpretación del MPE mostrado en la Figura 7 permite comprender por qué el *Docking* favorece tanto el modo de unión de **M118**, se observa una distribución electrostática altamente polarizada, con regiones de potencial negativo, representadas en rojo, localizadas principalmente sobre los átomos electronegativos del ligando, como oxígenos y nitrógenos del sistema hidrazida-base de Schiff y del núcleo heterocíclico. Estas zonas indican regiones ricas en densidad electrónica, capaces de actuar como aceptores de puentes de hidrógeno. De manera complementaria, las zonas de potencial positivo, representadas en azul, se distribuyen alrededor de hidrógenos unidos a nitrógenos y regiones parcialmente protonadas del sistema quinoxalina-triazol, lo que favorece interacciones con residuos ácidos del sitio activo.

El modo de unión de **M118** está sostenido por dos puentes de hidrógeno convencionales: uno con **Arg1510** a 1.91 Å, en el que un nitrógeno del sistema imino/hidrazida puede actuar como aceptor frente al grupo guanidina de la arginina; y otro con **Lys1460** a 1.75 Å, el más corto de los observados en el conjunto representativo, asociado a la interacción entre el hidrógeno del grupo amino de la lisina y el nitrógeno del sistema hidrazida. Estas interacciones concuerdan con el MPE, ya que las regiones de potencial negativo localizadas sobre los heteroátomos de **M118** favorecen los contacto con residuos básicos como arginina y lisina. Por ello, la presencia de estos contactos cortos contribuye de manera importante a la energía de unión favorable calculada por *Docking*.

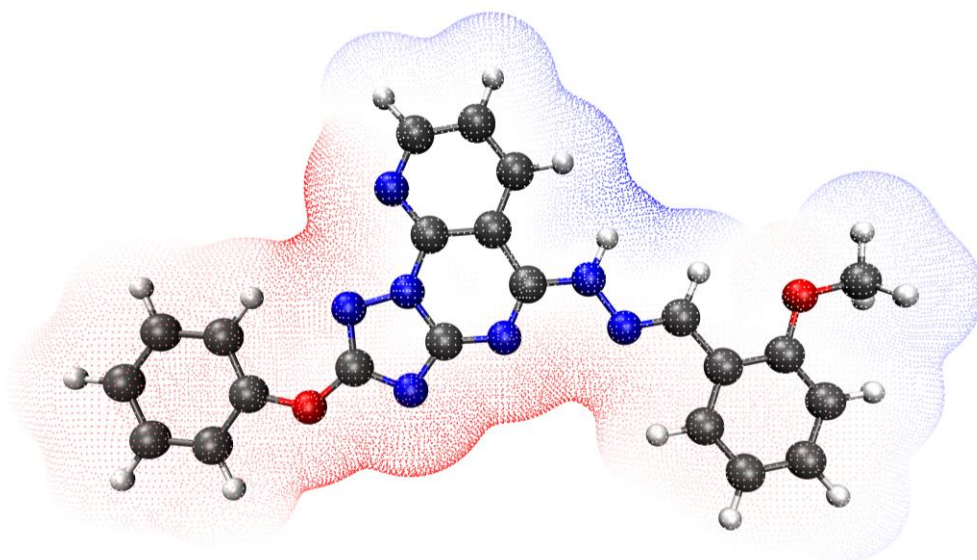


Figura 7. Mapa del potencial electrostático de **M118**

Los contactos electrostáticos se amplían mediante tres interacciones de carga atractiva, que constituyen el rasgo más distintivo de **M118** en comparación con **M101**: una con **Asp1526** a **4.69 Å**, una con **Asp1420** a **4.51 Å** y una con **Asp1157** a **4.08 Å**. Estas interacciones pueden explicarse directamente a partir del MPE, debido a que las regiones electropositivas del ligando, asociadas principalmente a los nitrógenos del sistema quinoxalina-triazol e hidrazida, presentan complementariedad con los grupos carboxilato cargados negativamente de los residuos aspartato. En particular, la interacción con **Asp1526** es estructuralmente significativa, ya que este residuo corresponde al ácido/base catalítico principal del subsitio -1, por lo que su participación sugiere que **M118** logra posicionarse en una región funcionalmente importante del sitio activo.

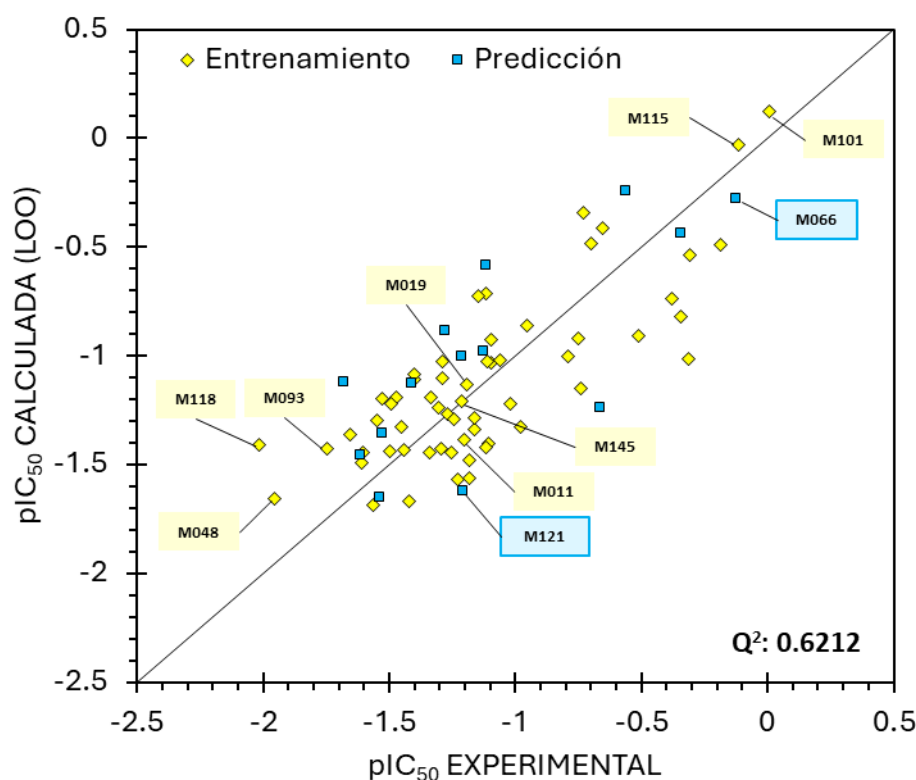
Adicionalmente, se observa una interacción π -catión con **Lys1460** a 3.69 Å, que añade un componente electrostático aromático al contacto polar ya establecido con este residuo. Esta interacción también puede relacionarse con el MPE, ya que los sistemas aromáticos y heteroaromáticos de **M118** presentan una distribución electrónica polarizada por la presencia de múltiples átomos de nitrógeno. Esta polarización favorece la interacción entre la nube π del ligando y residuos cargados positivamente como lisina, reforzando la unión del compuesto dentro del sitio activo.

Los contactos aromáticos e hidrofóbicos se establecen mediante interacciones π - π en conformación de T con **Tyr1251** a **5.53 Å** y **Phe1559** a **5.23 Å**, las cuales reflejan el acomodo del sistema aromático del núcleo quinoxalina dentro de una región rica en residuos aromáticos del sitio activo. También se observa una interacción π -sigma con **Trp1369** a **3.76 Å**, que sugiere un contacto favorable entre el sistema π del triptófano y enlaces σ del ligando, así como una interacción alquil con **Met1421** a 5.38 Å. En el MPE, estas regiones aromáticas muestran zonas menos intensamente polarizadas que los heteroátomos, lo que permite que los fragmentos aromáticos de **M118** participen en contactos hidrofóbicos y π con residuos como Trp, Phe y Tyr.

Sin embargo, aunque el MPE ayuda a explicar la alta afinidad aparente de **M118** en el *Docking*, también permite entender por qué esta afinidad computacional no necesariamente se traduce en mayor actividad experimental. La molécula presenta una distribución electrostática muy marcada, con múltiples regiones polares y centros capaces de establecer interacciones de carga o puentes de hidrógeno. La molécula presenta una distribución electrostática muy marcada, con múltiples regiones polares y centros capaces de establecer interacciones electrostáticas o puentes de hidrógeno. Esta característica puede favorecer la energía de unión estimada, ya que la función de *scoring* de AutoDock4 incluye explícitamente contribuciones asociadas con las interacciones electrostáticas y los puentes de hidrógeno [160]. No obstante, en un entorno biológico, dicha polaridad puede

implicar una penalización importante por desolvación, ya que las regiones cargadas o altamente polares deben perder parte de sus interacciones con el solvente antes de unirse al receptor [161].

Por lo tanto, el caso de **M118** muestra que una energía libre de unión muy favorable no necesariamente implica una mayor actividad inhibitoria experimental. Su MPE revela una molécula altamente polarizada, capaz de establecer múltiples interacciones electrostáticas con **Asp1526**, **Asp1420**, **Asp1157**, **Arg1510** y **Lys1460**, lo que explica su ΔG favorable en *Docking*. Sin embargo, esa misma polaridad puede generar efectos desfavorables en términos de desolvación, flexibilidad conformacional y posible biodisponibilidad.

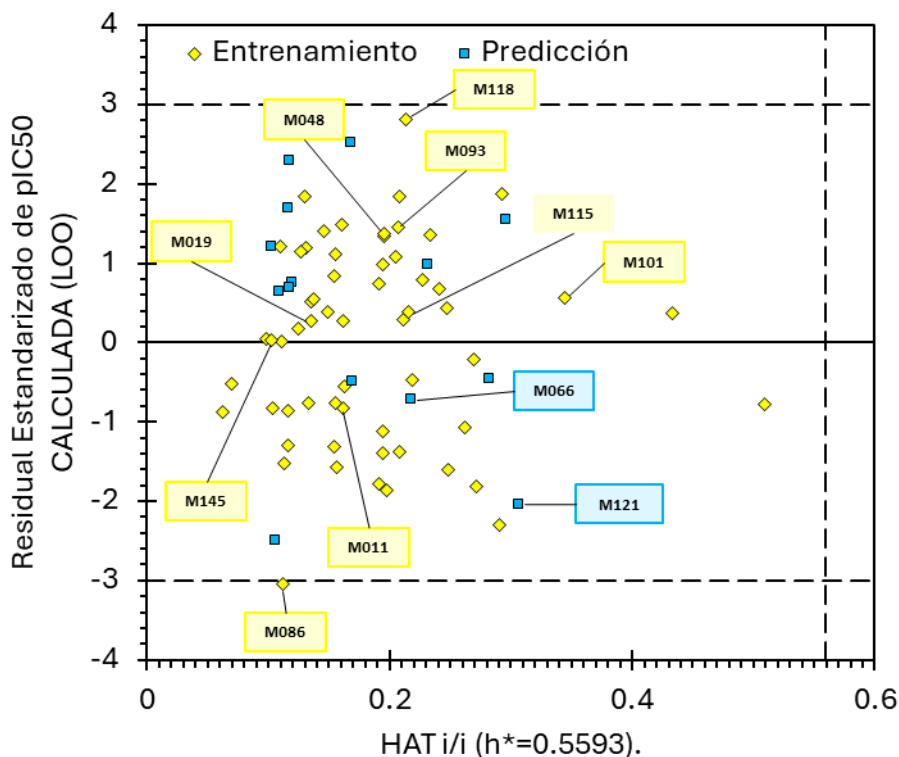


Gráfica 7. Modelo QSAR de pIC_{50} calculada (LOO) vs pIC_{50} experimental.

El coeficiente R^2 no es el único parámetro que puede dar una estimación de su capacidad predictiva, también el valor de Q^2_{LOO} el cual cuenta con un valor de 0.621, siendo aceptable de acuerdo con *Goldbraik et al.* [156]. Este parámetro estima el valor de pIC_{50} para cada compuesto del set de entrenamiento, excluyéndolo del modelo, utilizando así las mismas variables, pero recalculando los coeficientes. En la Gráfica 7 corresponde a la validación cruzada *Leave-One-Out*. Visualmente, la gráfica mantiene una distribución muy similar a la observada en la Gráfica 7, lo cual

es un indicador importante de estabilidad interna del modelo. Si el modelo estuviera sobreajustado, la dispersión aumentaría considerablemente durante la validación LOO. La cercanía entre los valores experimentales y los valores calculados mediante LOO demuestra que la ecuación QSAR conserva capacidad predictiva incluso cuando se elimina información del conjunto de entrenamiento. Además, la diferencia relativamente pequeña entre $R^2 - Q_{LOO}^2$ siendo 0.126, que es válido de acuerdo con *Ayoub L. et al.* [162], ya que no excede el valor límite de 0.3, confirma que el modelo no depende excesivamente de moléculas específicas para mantener su desempeño estadístico. La validación LOO también permite detectar compuestos potencialmente problemáticos. Los compuestos que se alejan considerablemente de la diagonal pueden interpretarse como estructuras cuya actividad inhibitoria depende de características que no están completamente representadas por los descriptores seleccionados.

En la validación externa, los parámetros nos permiten comprobar la capacidad predictiva del Modelo. Estos resultados indican que el modelo se considera predictivo, ya que el valor de R_{pred}^2 es de 0.572, siendo este mayor a 0.5, este parámetro es particularmente importante porque demuestra que la ecuación QSAR puede utilizarse para estimar la actividad de nuevos compuestos. El parámetro $Q_{F2}^2 = 0.489$ también respalda la estabilidad predictiva del modelo, ya que permanece por debajo de R_{pred}^2 , cumpliendo con los criterios establecidos para modelos QSAR válidos. Los valores de k y k' se encuentran en el intervalo de 0.85 a 1.15, lo que indica que no existe un sesgo importante en la pendiente de regresión entre valores experimentales y calculados. Esto significa que el modelo no tiende sistemáticamente a sobrestimar ni subestimar la actividad biológica. También cumplió con el criterio de $\frac{R_{pred}^2 - r_0'^2}{R_{pred}^2}$ con valor de 0.059 siendo menor que 0.1, $|R_{pred}^2 - r_0'^2|$ con un valor de 0.034 siendo menor que 0.3, $\Delta r_{m(pred)}^2$ con valor de 0.111 siendo menor que 0.2 y finalmente $\bar{r}_{m(pred)}^2$ con valor de 0.522 cumpliendo con la condición de ser mayor a 0.5.



Gráfica 8. Gráfico de Williams del modelo QSAR pIC_{50} donde las líneas punteadas horizontales representan los límites al 98.76% (-3.0 y 3.0), la línea punteada vertical corresponde al valor de h^* , los rombos amarillos pertenecen al conjunto de entrenamiento y los cuadrados azules al conjunto de predicción.

La Gráfica 8 corresponde al gráfico de Williams, una herramienta ampliamente utilizada para evaluar simultáneamente el dominio de aplicabilidad y la presencia de valores atípicos. En el eje horizontal se representan los valores de *leverage* o *Hat*, que indican la influencia estructural de cada compuesto sobre el modelo. Valores elevados de *Hat* significan que la molécula posee características estructurales poco comunes respecto al resto del conjunto. La línea vertical punteada representa el valor crítico h^* . Las moléculas localizadas a la derecha de esta línea son consideradas compuestos influyentes, ya que poseen gran impacto sobre la ecuación QSAR. En el eje vertical se muestran los residuales estandarizados, los cuales representan el error entre el valor experimental y el valor calculado para cada compuesto. Las líneas horizontales en +3 y -3 delimitan el intervalo aceptable de error estadístico. Los compuestos ubicados fuera de estos límites presentan errores de predicción elevados y pueden considerarse valores atípicos. La mayoría de las moléculas del conjunto de entrenamiento y predicción se encuentran dentro de la región delimitada por h^* y los residuales, indicando que el modelo posee un dominio de aplicabilidad bien definido.

El compuesto **M086** presentó un residual superior al límite establecido de 3 [163], lo que indica un error de predicción relativamente elevado. Sin embargo, debido a que permanece dentro del dominio de aplicabilidad estructural (*Hat* menor a h^*), no se considera una molécula externa al modelo. Este comportamiento sugiere que la desviación observada podría deberse a factores experimentales, conformacionales

o electrónicos no descritos completamente por los descriptores seleccionados. En términos generales, el gráfico de Williams confirma que el modelo QSAR desarrollado presenta estabilidad estadística y capacidad predictiva confiable dentro del espacio químico definido por las moléculas analizadas.

Conclusiones

A partir de los resultados obtenidos en el presente trabajo, fue posible desarrollar un modelo QSAR dentro de una estrategia computacional apoyada por estudios de *Docking* y cálculos DFT, con el propósito de analizar y predecir la actividad inhibitoria de diferentes compuestos frente a la enzima α -glucosidasa. La integración de estas herramientas permitió estudiar la interacción molecular y establecer relaciones entre la estructura química de los compuestos y su actividad biológica.

La recopilación de información experimental permitió construir una base de datos químicamente diversa, integrada por 157 IAGs, distribuidos en 18 categorías estructurales y caracterizados por un amplio intervalo de actividades biológicas. Esta diversidad proporcionó una base adecuada para investigar las relaciones estructura-actividad, sin restringir el análisis a un solo conjunto de derivados químicos.

La preoptimización conformacional mediante dinámica molecular permitió explorar diferentes conformaciones y obtener geometrías de baja energía, químicamente razonables y adecuadas para los estudios posteriores de *Docking* y el cálculo de descriptores moleculares. De esta manera, fue posible disponer de estructuras tridimensionales representativas para las siguientes etapas del flujo de trabajo computacional.

Los estudios de *Docking* permitieron generar 3,556 complejos con diferentes modos de unión, posteriormente, mediante el algoritmo *K-means*, estos complejos se agruparon en seis clústers, lo que permitió seleccionar los inhibidores representativos considerando simultáneamente tres criterios: la cercanía al sitio activo, una energía de unión favorecida y buena actividad biológica experimental mediante pIC_{50} . El clúster 1 presentó el mejor conjunto de estructuras con las características requeridas para ser incluidas en el QSAR cumpliendo los tres criterios mencionados, por lo que se seleccionaron los 74 compuestos pertenecientes a este conjunto de complejos.

El análisis QSAR permitió identificar un modelo estadísticamente válido, robusto y con capacidad predictiva, siendo el Modelo 4 el que cumplió de manera más adecuada con los criterios de validación establecidos, al presentar valores de $R^2 = 0.747$, $Q_{LOO}^2 = 0.621$ y $R_{pred}^2 = 0.572$, estos resultados indican que los descriptores incluidos en la ecuación representan adecuadamente características moleculares relacionadas con la inhibición de la α -glucosidasa y permiten predecir la actividad inhibitoria de los compuestos pertenecientes al modelo. Sin embargo, la construcción del Modelo 4 demostró que también es posible obtener un modelo estadísticamente válido, robusto y predictivo sin incorporar variables derivadas de los cálculos DFT en su ecuación final, por lo tanto, aunque los descriptores electrónicos pueden aportar información relevante para la interpretación molecular,

su inclusión no necesariamente garantiza una mejora en el desempeño estadístico del modelo QSAR.

Finalmente, la combinación de las herramientas computacionales empleadas permitió desarrollar una aproximación adecuada para relacionar la estructura molecular con la actividad inhibitoria frente a la α -glucosidasa. El modelo seleccionado presentó estabilidad estadística y capacidad predictiva, por lo que puede utilizarse como una herramienta para orientar la predicción *in silico* y la priorización de nuevos compuestos con posible actividad inhibitoria, no obstante, es importante tener presente que estos resultados siguen limitados al dominio de aplicabilidad además de que su uso debe complementarse con estudios experimentales antes de proponer candidatos definitivos para el desarrollo de nuevos IAGs.

Referencias

- [1] R. Goyal, M. Singhal e I. Jialal, "Type 2 Diabetes," en *StatPearls [Internet]*. Treasure Island, FL, EE. UU.: StatPearls Publishing, 2026. Actualizado: 23 jun. 2023. [En línea]. Disponible en: <https://www.ncbi.nlm.nih.gov/books/NBK513253/>
- [2] International Diabetes Federation, *IDF Diabetes Atlas*, 11.^a ed. Bruselas, Bélgica: International Diabetes Federation, 2025. [En línea]. Disponible en: <https://diabetesatlas.org/resources/idf-diabetes-atlas-2025/>
- [3] A. Basto-Abreu et al., "Prevalencia de prediabetes y diabetes en México: Ensanut 2022," *Salud Publica Mex*, vol. 65, pp. s163–s168, jun. 2023, doi: 10.21149/14832.
- [4] S. R. Joshi, E. Standl, N. Tong, P. Shah, S. Kalra, y R. Rathod, "Therapeutic potential of α -glucosidase inhibitors in type 2 diabetes mellitus: an evidence-based review", *Expert Opinion on Pharmacotherapy*, vol. 16, no. 13, pp. 1959–1981, ago. 2015, doi: 10.1517/14656566.2015.1070827.
- [5] National Center for Biotechnology Information, "PubChem Enzyme Summary for Enzyme 3.2.1.20, Alpha-glucosidase (EC 3.2.1.20)," *PubChem*, 2026. [Accedido: 18 jun. 2026]. [En línea]. Disponible en: <https://pubchem.ncbi.nlm.nih.gov/protein/EC:3.2.1.20>
- [6] B. L. Nichols, S. Avery, P. Sen, D. M. Swallow, D. Hahn, y E. Sterchi, "The maltase-glucoamylase gene: Common ancestry to sucrase-isomaltase with complementary starch digestion activities", *Proc. Natl. Acad. Sci. U.S.A.*, vol. 100, no. 3, pp. 1432–1437, ene. 2003, doi: 10.1073/pnas.0237170100.
- [7] J. Mitra, A. Sajeewan, S. Sujith, S. Paramasivam, K. Shanmugam y A. P. Solomon, " α -Glucosidase-driven metabolism as a potential therapeutic vulnerability in *Candida albicans*," *Front. Microbiol.*, vol. 17, abr. 2026, Art. no. 1812644, doi: 10.3389/fmicb.2026.1812644.
- [8] ScienceDirect Topics, "Enzyme Active Site," *ScienceDirect*. [Accedido: 19 jun. 2026]. [En línea]. Disponible en: <https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/enzyme-active-site>
- [9] G. J. Davies, K. S. Wilson, y B. Henrissat, "Nomenclature for sugar-binding subsites in glycosyl hydrolases", *Biochemical Journal*, vol. 321, no. 2, pp. 557–559, ene. 1997, doi: 10.1042/bj3210557.
- [10] N. F. Brás, D. Santos-Martins, P. A. Fernandes, y M. J. Ramos, "Mechanistic Pathway on Human α -Glucosidase Maltase-Glucoamylase Unveiled by QM/MM Calculations", *J. Phys. Chem. B*, vol. 122, no. 14, pp. 3889–3899, mar. 2018, doi: 10.1021/acs.jpcc.8b01321.
- [11] E. Feduchi Canosa, C. Romero Magdalena, E. Yáñez Conde, y C. García-Hoz Jiménez, *Bioquímica: Conceptos esenciales*, 3.^a ed. Madrid, España: Editorial Médica Panamericana, 2021.
- [12] M. Akmal, P. Patel y R. Wadhwa, "Alpha Glucosidase Inhibitors," en *StatPearls [Internet]*. Treasure Island, FL, EE. UU.: StatPearls Publishing, 2026. Actualizado: 28 feb. 2024. [Accedido: 3 jun. 2026]. [En línea]. Disponible en: <https://www.ncbi.nlm.nih.gov/books/NBK557848/>

- [13] L. A. McIver, C. V. Preuss y J. Tripp, "Acarbose," en *StatPearls [Internet]*. Treasure Island, FL, EE. UU.: StatPearls Publishing, 2026. Actualizado: 12 feb. 2024. [En línea]. Disponible en: <https://www.ncbi.nlm.nih.gov/books/NBK493214/>
- [14] "Miglitol," en *LiverTox: Clinical and Research Information on Drug-Induced Liver Injury [Internet]*. Bethesda, MD, EE. UU.: National Institute of Diabetes and Digestive and Kidney Diseases, 2012—. Actualizado: 12 ene. 2021. [En línea]. Disponible en: <https://www.ncbi.nlm.nih.gov/books/NBK547878/>
- [15] A. S. Dabhi, N. R. Bhatt y M. J. Shah, "Voglibose: an alpha glucosidase inhibitor," *J. Clin. Diagn. Res.*, vol. 7, no. 12, pp. 3023–3027, dic. 2013, doi: 10.7860/JCDR/2013/6373.3838.
- [16] International Union of Pure and Applied Chemistry (IUPAC), "Inhibitory concentration," en *Compendium of Chemical Terminology (the Gold Book)*, 2014, doi: 10.1351/goldbook.i03036.
- [17] X. Zheng, S. Zhou, C. Zhang, D. Wu, H.-B. Luo y Y. Wu, "Docking-assisted 3D-QSAR studies on xanthenes as α -glucosidase inhibitors," *J. Mol. Model.*, vol. 23, no. 9, sep. 2017, Art. no. 272, doi: 10.1007/s00894-017-3438-1.
- [18] K. S. Aminu, A. Uzairu, S. E. Abechi, G. A. Shallangwa y A. B. Umar, "Ligand-based drug design, molecular docking and pharmacokinetic studies of some series of 1,4 - dihydropyridines derivatives as human intestinal maltase-glucoamylase inhibitor," *Chemical Data Collections*, vol. 41, Art. no. 100911, oct. 2022, doi: 10.1016/j.cdc.2022.100911.
- [19] J. R. Mora, E. A. Márquez, y L. Calle, "Computational molecular modelling of N-cinnamoyl and hydroxycinnamoyl amides as potential α -glucosidase inhibitors," *Med Chem Res*, vol. 27, no. 9, pp. 2214–2223, ago. 2018, doi: 10.1007/s00044-018-2229-2.
- [20] C. R. García-Jacas, Y. Marrero-Ponce, S. J. Barigye, T. Hernández-Ortega, L. Cabrera-Leyva, y A. Fernández-Castillo, "N-tuple topological/geometric cutoffs for 3D N-linear algebraic molecular codifications: variability, linear independence and QSAR analysis", *SAR and QSAR in Environmental Research*, vol. 27, no. 12, pp. 949–975, oct. 2016, doi: 10.1080/1062936x.2016.1231714.
- [21] D. B. Kitchen, H. Decornez, J. R. Furr, y J. Bajorath, "Docking and scoring in virtual screening for drug discovery: methods and applications", *Nat Rev Drug Discov*, vol. 3, no. 11, pp. 935–949, nov. 2004, doi: 10.1038/nrd1549.
- [22] F. De Proft, P. Geerlings, F. Heidar-Zadeh, y P. W. Ayers, "Conceptual Density Functional Theory", *Comprehensive Computational Chemistry*. Elsevier, pp. 306–321, 2024. doi: 10.1016/b978-0-12-821978-2.00025-8.
- [23] F. De Proft, P. Geerlings, F. Heidar-Zadeh y P. W. Ayers, "Conceptual Density Functional Theory," en *Comprehensive Computational Chemistry*, M. Yáñez y R. J. Boyd, Eds., 1.^a ed. Oxford, Reino Unido: Elsevier, 2024, pp. 306–321, doi: 10.1016/B978-0-12-821978-2.00025-8.
- [24] J. L. Gázquez, A. Cedillo, y A. Vela, "Electrodonating and Electroaccepting Powers", *J. Phys. Chem. A*, vol. 111, no. 10, pp. 1966–1970, feb. 2007, doi: 10.1021/jp065459f.
- [25] P. K. Chattaraj, U. Sarkar, y D. R. Roy, "Electrophilicity Index", *Chem. Rev.*, vol. 106, no. 6, pp. 2065–2091, may 2006, doi: 10.1021/cr040109f.

- [26] A. Mauri, V. Consonni, M. Pavan, y R. Todeschini, "DRAGON software: An easy approach to molecular descriptor calculations," *MATCH Communications in Mathematical and in Computer Chemistry*, vol. 56, pp. 237–248, 2006.
- [27] G. Landrum et al., *rdkit/rdkit: 2025_09_6 (Q3 2025) Release*. Zenodo, 2026, doi: 10.5281/ZENODO.591637.
- [28] L. Naanaai et al., "Design of Novel Antidiabetic Agents Using 3D-QSAR, Molecular Docking, ADMET Analysis, Molecular Dynamics, Ligand Transport, and Retrosynthesis," *ChemistrySelect*, vol. 10, no. 38, oct. 2025, doi: 10.1002/slct.202502280.
- [29] L. Naanaai et al., "Design of Novel Antidiabetic Agents Using 3D-QSAR, Molecular Docking, ADMET Analysis, Molecular Dynamics, Ligand Transport, and Retrosynthesis," *ChemistrySelect*, vol. 10, no. 38, oct. 2025, Art. no. e02280, doi: 10.1002/slct.202502280.
- [30] A. Cherkasov et al., "QSAR Modeling: Where Have You Been? Where Are You Going To?," *J. Med. Chem.*, vol. 57, no. 12, pp. 4977–5010, ene. 2014, doi: 10.1021/jm4004285.
- [31] X.-Y. Meng, H.-X. Zhang, M. Mezei, y M. Cui, "Molecular Docking: A Powerful Approach for Structure-Based Drug Discovery," *CAD*, vol. 7, no. 2, pp. 146–157, jun. 2011, doi: 10.2174/157340911795677602.
- [32] N. Mardirossian y M. Head-Gordon, "Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals," *Molecular Physics*, vol. 115, no. 19, pp. 2315–2372, jun. 2017, doi: 10.1080/00268976.2017.1333644.
- [33] I. Gijbels, "Censored data", *WIREs Computational Stats*, vol. 2, no. 2, pp. 178–188, mar. 2010, doi: 10.1002/wics.80.
- [34] National Cancer Institute (NCI/CADD Group), "Online SMILES Translator and Structure File Generator (CACTVS Translate Service)," 2024. [Accedido: 2 feb. 2026]. [En línea]. Disponible en: <https://cactus.nci.nih.gov/translate/>
- [35] B. Hourahine et al., "Recent Developments in DFTB+, a Software Package for Efficient Atomistic Quantum Mechanical Simulations", *J. Phys. Chem. A*, vol. 129, no. 24, pp. 5373–5390, Jun. 2025, doi: 10.1021/acs.jpca.5c01146.
- [36] M. Gaus, X. Lu, M. Elstner, y Q. Cui, "Parameterization of DFTB3/3OB for Sulfur and Phosphorus for Chemical and Biological Applications", *J. Chem. Theory Comput.*, vol. 10, no. 4, pp. 1518–1537, mar. 2014, doi: 10.1021/ct401002w.
- [37] G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell, y A. J. Olson, "AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility," *Journal of Computational Chemistry*, vol. 30, no. 16, pp. 2785–2791, 2009.
- [38] RCSB Protein Data Bank (RCSB PDB), "RCSB Protein Data Bank," s. f. [Accedido: 30 ene. 2025]. [En línea]. Disponible en: <https://www.rcsb.org/>
- [39] C. Z. Gómez Castro, "Síntesis, caracterización y afinidad por docking de derivados del 1-fenilbencimidazol por la ciclooxigenasa," tesis de maestría, Sección de Estudios de Posgrado e Investigación, Escuela Superior de Medicina, Instituto Politécnico Nacional, Ciudad de México, México, 2006.
- [40] E. H. Acosta-Jiménez et al., "QSTR Modeling to Find Relevant DFT Descriptors Related to the Toxicity of Carbamates", *Molecules*, vol. 27, no. 17, Art. no. 5530, ago. 2022, doi: 10.3390/molecules27175530.

- [41] G. M. Morris et al., *AutoDock 4.2 User Guide*. La Jolla, CA, EE. UU.: The Scripps Research Institute, 2011. [En línea]. Disponible en: <http://autodock.scripps.edu/faqs-help/manual>
- [42] C. Z. Gómez-Castro, M. López-Martínez, J. Hernández-Pineda, J. G. Trujillo-Ferrara e I. I. Padilla-Martínez, "Profiling the interaction of 1-phenylbenzimidazoles to cyclooxygenases," *J. Mol. Recognit.*, vol. 32, no. 11, nov. 2019, Art. no. e2801, doi: 10.1002/jmr.2801.
- [43] J. MacQueen, "Some Methods for Classification and Analysis of Multivariate Observations," in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, L. M. Le Cam y J. Neyman, Eds. Berkeley, CA, USA: University of California Press, 1967, pp. 281–297.
- [44] S. Lloyd, "Least squares quantization in PCM", *IEEE Trans. Inform. Theory*, vol. 28, no. 2, pp. 129–137, mar. 1982, doi: 10.1109/tit.1982.1056489.
- [45] M. J. Frisch et al., *Gaussian 09*, Rev. D.01. Wallingford, CT, USA: Gaussian, Inc., 2009.
- [46] W. J. Hehre, R. Ditchfield, y J. A. Pople, "Self—Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian—Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules", *The Journal of Chemical Physics*, vol. 56, no. 5, pp. 2257–2261, mar. 1972, doi: 10.1063/1.1677527.
- [47] C. Lee, W. Yang, y R. G. Parr, "Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density", *Phys. Rev. B*, vol. 37, no. 2, pp. 785–789, ene. 1988, doi: 10.1103/physrevb.37.785.
- [48] P. C. Hariharan y J. A. Pople, "The influence of polarization functions on molecular orbital hydrogenation energies", *Theoret. Chim. Acta*, vol. 28, no. 3, pp. 213–222, 1973, doi: 10.1007/bf00533485.
- [49] R. Vijayaraj, V. Subramanian, y P. K. Chattaraj, "Comparison of Global Reactivity Descriptors Calculated Using Various Density Functionals: A QSAR Perspective", *J. Chem. Theory Comput.*, vol. 5, no. 10, pp. 2744–2753, Sep. 2009, doi: 10.1021/ct900347f.
- [50] P. Sharma, P. Ranjan y T. Chakraborty, "Applications of conceptual density functional theory in reference to quantitative structure–activity/property relationship," *Mol. Phys.*, vol. 122, no. 23, 2024, Art. no. e2331620, doi: 10.1080/00268976.2024.2331620.
- [51] Talete srl, *DRAGON (Software for Molecular Descriptor Calculation)*, ver. 6.0. Milano, Italia: Talete srl, 2012. [En línea]. Disponible en: <http://www.talete.mi.it/>
- [52] R. Todeschini y V. Consonni, *Handbook of Molecular Descriptors*. Weinheim, Alemania: Wiley-VCH, 2000. doi: 10.1002/9783527613106.
- [53] P. Gramatica, N. Chirico, E. Papa, S. Cassani, y S. Kovarich, "QSARINS: A new software for the development, analysis, and validation of QSAR MLR models", *J Comput Chem*, vol. 34, no. 24, pp. 2121–2132, Jun. 2013, doi: 10.1002/jcc.23361.
- [54] H. Zhang, J. Zheng, G. Cao, H. Lin, y B. Jia, "Real-Time Machine Learning-Based Position Recognition in Laser Nanofabrication with Sub-Half-Wavelength Precision", *Engineering*, vol. 57, pp. 200–209, feb. 2026, doi: 10.1016/j.eng.2025.03.037.
- [55] J. Wang et al., "Synthesis and inhibitory activity of euparin derivatives as potential dual inhibitors against α -glucosidase and protein tyrosine phosphatase 1B

- (PTP1B)", *Fitoterapia*, vol. 169, Art. no. 105596, Sep. 2023, doi: 10.1016/j.fitote.2023.105596.
- [56] Y. Khan et al., "Identification of novel oxadiazole-based benzothiazole derivatives as potent inhibitors of α -glucosidase and urease: Synthesis, in vitro bio-evaluation and their in silico molecular docking study", *Journal of Saudi Chemical Society*, vol. 27, no. 4, Art. no. 101682, Jul. 2023, doi: 10.1016/j.jscs.2023.101682.
- [57] A. Singh et al., "Recent developments in synthetic α -glucosidase inhibitors: A comprehensive review with structural and molecular insight", *Journal of Molecular Structure*, vol. 1281, p. 135115, Jun. 2023, doi: 10.1016/j.molstruc.2023.135115.
- [58] F. Rahim et al., "Synthesis, in vitro alpha-glucosidase inhibitory potential of benzimidazole bearing bis-Schiff bases and their molecular docking study", *Bioorganic Chemistry*, vol. 94, Art. no. 103394, ene. 2020, doi: 10.1016/j.bioorg.2019.103394.
- [59] N. Asemanipoor et al., "Synthesis and biological evaluation of new benzimidazole-1,2,3-triazole hybrids as potential α -glucosidase inhibitors", *Bioorganic Chemistry*, vol. 95, Art. no. 103482, ene. 2020, doi: 10.1016/j.bioorg.2019.103482.
- [60] L. Deswal et al., "Synthesis and antidiabetic evaluation of benzimidazole-tethered 1,2,3-triazoles," *Arch. Pharm.*, vol. 353, no. 9, sep. 2020, Art. no. e2000090, doi: 10.1002/ardp.202000090.
- [61] L. Yan et al., "Synthesis of tricyclic benzimidazole-iminosugars as potential glycosidase inhibitors via a Mitsunobu reaction", *Carbohydrate Research*, vol. 485, Art. no. 107807, nov. 2019, doi: 10.1016/j.carres.2019.107807.
- [62] M. Sherafati et al., "Design, synthesis and biological evaluation of novel phthalimide-Schiff base-coumarin hybrids as potent α -glucosidase inhibitors," *Chem. Pap.*, vol. 74, no. 12, pp. 4379–4388, 2020, doi: 10.1007/s11696-020-01246-7.
- [63] H. A. Abuelizz, E. H. Anouar, R. Ahmad, N. I. I. N. Azman, M. Marzouk, y R. Al-Salahi, "Triazoloquinazolines as a new class of potent α -glucosidase inhibitors: in vitro evaluation and docking study", *PLoS ONE*, vol. 14, no. 8, Art. no. e0220379, ago. 2019, doi: 10.1371/journal.pone.0220379.
- [64] S. E. Sadat-Ebrahimi et al., "New phthalimide-benzamide-1,2,3-triazole hybrids; design, synthesis, α -glucosidase inhibition assay, and docking study", *Med Chem Res*, vol. 29, no. 5, pp. 868–876, mar. 2020, doi: 10.1007/s00044-020-02522-7.
- [65] S. K. Avula et al., "Synthesis of novel (R)-4-fluorophenyl-1H-1,2,3-triazoles: A new class of α -glucosidase inhibitors", *Bioorganic Chemistry*, vol. 91, Art. no. 103182, oct. 2019, doi: 10.1016/j.bioorg.2019.103182.
- [66] X. Ferhati et al., "Dual targeting of PTP1B and glucosidases with new bifunctional iminosugar inhibitors to address type 2 diabetes", *Bioorganic Chemistry*, vol. 87, pp. 534–549, Jun. 2019, doi: 10.1016/j.bioorg.2019.03.053.
- [67] M. S. Asgari et al., "Biscoumarin-1,2,3-triazole hybrids as novel anti-diabetic agents: Design, synthesis, in vitro α -glucosidase inhibition, kinetic, and docking studies", *Bioorganic Chemistry*, vol. 92, Art. no. 103206, nov. 2019, doi: 10.1016/j.bioorg.2019.103206.
- [68] M. Saeedi et al., "Design, synthesis, in vitro, and in silico studies of novel diarylimidazole-1,2,3-triazole hybrids as potent α -glucosidase inhibitors", *Bioorganic & Medicinal Chemistry*, vol. 27, no. 23, Art. no. 115148, dic. 2019, doi: 10.1016/j.bmc.2019.115148.

- [69] G.-J. Ye et al., "Design and synthesis of novel xanthone-triazole derivatives as potential antidiabetic agents: α -Glucosidase inhibition and glucose uptake promotion", *European Journal of Medicinal Chemistry*, vol. 177, pp. 362–373, Sep. 2019, doi: 10.1016/j.ejmech.2019.05.045.
- [70] L. Kumar, K. Lal, P. Yadav, A. Kumar, y A. K. Paul, "Synthesis, characterization, α -glucosidase inhibition and molecular modeling studies of some pyrazoline-1H-1,2,3-triazole hybrids", *Journal of Molecular Structure*, vol. 1216, Art. no. 128253, Sep. 2020, doi: 10.1016/j.molstruc.2020.128253.
- [71] Z. Bakherad et al., "New thiosemicarbazide-1,2,3-triazole hybrids as potent α -glucosidase inhibitors: Design, synthesis, and biological evaluation", *Journal of Molecular Structure*, vol. 1192, pp. 192–200, Sep. 2019, doi: 10.1016/j.molstruc.2019.04.082.
- [72] M. S. Asgari et al., "Design and synthesis of 4,5-diphenyl-imidazol-1,2,3-triazole hybrids as new anti-diabetic agents: in vitro α -glucosidase inhibition, kinetic and docking studies", *Mol Divers*, vol. 25, no. 2, pp. 877–888, mar. 2020, doi: 10.1007/s11030-020-10072-8.
- [73] E. Nasli-Esfahani et al., "A new series of Schiff base derivatives bearing 1,2,3-triazole: Design, synthesis, molecular docking, and α -glucosidase inhibition," *Arch. Pharm.*, vol. 352, no. 8, ago. 2019, Art. no. e1900034, doi: 10.1002/ardp.201900034.
- [74] E. O. Yeye et al., "Syntheses, in vitro α -amylase and α -glucosidase dual inhibitory activities of 4-amino-1,2,4-triazole derivatives their molecular docking and kinetic studies", *Bioorganic & Medicinal Chemistry*, vol. 28, no. 11, Art. no. 115467, Jun. 2020, doi: 10.1016/j.bmc.2020.115467.
- [75] R. R. Pillai et al., "Synthesis, spectroscopic characterization, reactive properties by DFT calculations, molecular dynamics simulations and biological evaluation of Schiff bases tethered 1,2,4-triazole and pyrazole rings", *Journal of Molecular Structure*, vol. 1177, pp. 47–54, feb. 2019, doi: 10.1016/j.molstruc.2018.09.037.
- [76] F. Chaudhry et al., "Imidazole-pyrazole hybrids: Synthesis, characterization and in-vitro bioevaluation against α -glucosidase enzyme with molecular docking studies", *Bioorganic Chemistry*, vol. 82, pp. 267–273, feb. 2019, doi: 10.1016/j.bioorg.2018.10.047.
- [77] V. Pogaku, K. Gangarapu, S. Basavoju, K. K. Tatapudi, y S. B. Katragadda, "Design, synthesis, molecular modelling, ADME prediction and anti-hyperglycemic evaluation of new pyrazole-triazolopyrimidine hybrids as potent α -glucosidase inhibitors", *Bioorganic Chemistry*, vol. 93, Art. no. 103307, dic. 2019, doi: 10.1016/j.bioorg.2019.103307.
- [78] P. Kumar, M. Duhan, J. Sindhu, K. Kadyan, S. Saini, y N. Panihar, "Thiazolidine-4-one clubbed pyrazoles hybrids: Potent α -amylase and α -glucosidase inhibitors with NLO properties", *Journal of Heterocyclic Chem*, vol. 57, no. 4, pp. 1573–1587, dic. 2019, doi: 10.1002/jhet.3882.
- [79] R. Kaur, K. Palta, y M. Kumar, "Hybrids of Isatin-Pyrazole as Potential α -Glucosidase Inhibitors: Synthesis, Biological Evaluations and Molecular Docking Studies", *ChemistrySelect*, vol. 4, no. 45, pp. 13219–13227, dic. 2019, doi: 10.1002/slct.201903418.
- [80] A. Aispuro-Pérez et al., "Synthesis and molecular docking studies of imines as α -glucosidase and α -amylase inhibitors", *Bioorganic Chemistry*, vol. 94, Art. no. 103491, ene. 2020, doi: 10.1016/j.bioorg.2019.103491.

- [81] F. Zeng et al., "Design, Synthesis, and Activity Evaluation of Novel N-benzyl Deoxynojirimycin Derivatives for Use as α -Glucosidase Inhibitors", *Molecules*, vol. 24, no. 18, Art. no. 3309, Sep. 2019, doi: 10.3390/molecules24183309.
- [82] T. Luthra et al., "Discovery of novel pyrido-pyrrolidine hybrid compounds as alpha-glucosidase inhibitors and alternative agent for control of type 1 diabetes", *European Journal of Medicinal Chemistry*, vol. 188, Art. no. 112034, feb. 2020, doi: 10.1016/j.ejmech.2020.112034.
- [83] T. B. Tafesse et al., "Synthesis and biological evaluation of 2-(2-methyl-1H-pyrrol-3-yl)-2-oxo-N-(pyridine-3-yl) acetamide derivatives: in vitro α -glucosidase inhibition, and kinetic and molecular docking study", *Chem. Pap.*, vol. 74, no. 5, pp. 1583–1596, nov. 2019, doi: 10.1007/s11696-019-00999-0.
- [84] Y.-D. Xie et al., "Design, synthesis and evaluation of phenylfuroxan nitric oxide-donor phenols as potential anti-diabetic agents", *Bioorganic Chemistry*, vol. 89, Art. no. 103000, ago. 2019, doi: 10.1016/j.bioorg.2019.103000.
- [85] H. Ullah et al., "Aryl-oxadiazole Schiff bases: Synthesis, α -glucosidase in vitro inhibitory activity and their in silico studies", *Arabian Journal of Chemistry*, vol. 13, no. 4, pp. 4904–4915, abr. 2020, doi: 10.1016/j.arabjc.2020.01.005.
- [86] M. Iftikhar et al., "A novel five-step synthetic route to 1,3,4-oxadiazole derivatives with potent α -glucosidase inhibitory potential and their in silico studies," *Arch. Pharm.*, vol. 352, no. 12, dic. 2019, Art. no. e1900095, doi: 10.1002/ardp.201900095.
- [87] M. Gollapalli et al., "Synthesis of benzothiazole derivatives as a potent α -glucosidase inhibitor", *Bioorganic Chemistry*, vol. 85, pp. 33–48, abr. 2019, doi: 10.1016/j.bioorg.2018.12.021.
- [88] A. Khosravi, G. Vaezi, V. Hojati, y K. Abdi, "Study on the interaction of triaryl-dihydro-1,2,4-oxadiazoles with α -glucosidase", *DARU J Pharm Sci*, vol. 28, no. 1, pp. 109–117, ene. 2020, doi: 10.1007/s40199-019-00322-y.
- [89] X. Liu et al., "Semi-synthesis of C28-modified triterpene acid derivatives from maslinic acid or corosolic acid as potential α -glucosidase inhibitors", *Bioorganic Chemistry*, vol. 97, Art. no. 103694, abr. 2020, doi: 10.1016/j.bioorg.2020.103694.
- [90] E. F. Khusnutdinova et al., "Structural modifications of 2,3-indolobetulinic acid: Design and synthesis of highly potent α -glucosidase inhibitors", *Bioorganic Chemistry*, vol. 88, Art. no. 102957, Jul. 2019, doi: 10.1016/j.bioorg.2019.102957.
- [91] Z. Zeng, X. Yin, X. Wang, W. Yang, X. Liu, y Y. Hong, "Synthesis of water soluble pentacyclic dihydroxyterpene carboxylic acid derivatives coupled amino acids and their inhibition activities on α -glucosidase", *Bioorganic Chemistry*, vol. 86, pp. 277–287, may 2019, doi: 10.1016/j.bioorg.2019.02.001.
- [92] Y.-Y. Zhong et al., "Synthesis and biological evaluation of novel oleanolic acid analogues as potential α -glucosidase inhibitors", *European Journal of Medicinal Chemistry*, vol. 164, pp. 706–716, feb. 2019, doi: 10.1016/j.ejmech.2018.12.046.
- [93] H.-V.-T. Phan et al., "Design and synthesis of new lupeol derivatives and their α -glucosidase inhibitory and cytotoxic activities", *Natural Product Research*, vol. 36, no. 1, pp. 1–7, may 2020, doi: 10.1080/14786419.2020.1758095.
- [94] N.-H. Nguyen et al., "Design, modification of phyllanthone derivatives as anti-diabetic and cytotoxic agents", *Natural Product Research*, vol. 36, no. 1, pp. 371–378, Jul. 2020, doi: 10.1080/14786419.2020.1788023.

- [95] W. Zhang, H.-Y. Wang, H.-X. Wang, y Z.-Y. Zhu, "Synthesis and inhibition of α -glucosidase of methyl glycyrrhetinate glycosides", *Natural Product Research*, vol. 35, no. 11, pp. 1874–1880, Jul. 2019, doi: 10.1080/14786419.2019.1639181.
- [96] K. Boudebous, H. Boulebd, C. Bensouici, D. Harakat, R. Boulcina, y A. Debache, "Synthesis, Docking Study and Biological Activities Evaluation of 1-Amidoalkyl-2-naphthol Derivatives as Dual Inhibitors of Cholinesterase and α -Glucosidase", *ChemistrySelect*, vol. 5, no. 19, pp. 5515–5520, may 2020, doi: 10.1002/slct.202000558.
- [97] F. A. Badria et al., "Synthesis and Inhibitory Effect of Some Indole-Pyrimidine Based Hybrid Heterocycles on α -Glucosidase and α -Amylase as Potential Hypoglycemic Agents", *ChemistryOpen*, vol. 8, no. 10, pp. 1288–1297, oct. 2019, doi: 10.1002/open.201900240.
- [98] S. Machida, S. Mukai, R. Kono, M. Funato, H. Saito, y T. Uchiyama, "Synthesis and Comparative Structure–Activity Study of Carbohydrate-Based Phenolic Compounds as α -Glucosidase Inhibitors and Antioxidants", *Molecules*, vol. 24, no. 23, Art. no. 4340, nov. 2019, doi: 10.3390/molecules24234340.
- [99] C. Jiang et al., "Selective synthesis of 3-deoxy-5-hydroxy-1-amino-carbasugars as potential α -glucosidase inhibitors", *Org. Biomol. Chem.*, vol. 17, no. 21, pp. 5381–5391, 2019, doi: 10.1039/c9ob00762h.
- [100] L. Guazzelli et al., "Synthesis and investigation of polyhydroxylated pyrrolidine derivatives as novel chemotypes showing dual activity as glucosidase and aldose reductase inhibitors", *Bioorganic Chemistry*, vol. 92, Art. no. 103298, nov. 2019, doi: 10.1016/j.bioorg.2019.103298.
- [101] Z. Tavaf, S. K. Dangolani, R. Yousefi, F. Panahi, M. B. Shahsavani, y A. Khalafi-Nezhad, "Synthesis of new curcumin derivatives as influential antidiabetic α -glucosidase and α -amylase inhibitors with anti-oxidant activity", *Carbohydrate Research*, vol. 494, Art. no. 108069, ago. 2020, doi: 10.1016/j.carres.2020.108069.
- [102] L. Zhang, Q. Xu, J. Zhu, G. Xia, y H. Zang, "Synthesis, α -Glucosidase inhibition and molecular docking studies of tyrosol derivatives", *Natural Product Research*, vol. 35, no. 10, pp. 1596–1604, Jun. 2019, doi: 10.1080/14786419.2019.1628750.
- [103] P. Sharma et al., "Synthesis of sulpha drug based hydroxytriazene derivatives: Anti-diabetic, antioxidant, anti-inflammatory activity and their molecular docking studies", *Bioorganic Chemistry*, vol. 96, Art. no. 103642, mar. 2020, doi: 10.1016/j.bioorg.2020.103642.
- [104] J. Ashraf et al., "Design and synthesis of new flavonols as dual α -amylase and α -glucosidase inhibitors: Structure-activity relationship, drug-likeness, in vitro and in silico studies", *Journal of Molecular Structure*, vol. 1218, Art. no. 128458, oct. 2020, doi: 10.1016/j.molstruc.2020.128458.
- [105] H. Yousuf et al., "Dihydropyridines as potential α -amylase and α -glucosidase inhibitors: Synthesis, in vitro and in silico studies", *Bioorganic Chemistry*, vol. 96, Art. no. 103581, mar. 2020, doi: 10.1016/j.bioorg.2020.103581.
- [106] X.-T. Xu et al., "Synthesis and biological evaluation of coumarin derivatives as α -glucosidase inhibitors", *European Journal of Medicinal Chemistry*, vol. 189, Art. no. 112013, mar. 2020, doi: 10.1016/j.ejmech.2019.112013.
- [107] F. Hussain et al., "Synthesis, in-vitro α -glucosidase inhibition, antioxidant, in-vivo antidiabetic and molecular docking studies of pyrrolidine-2,5-dione and

- thiazolidine-2,4-dione derivatives”, *Bioorganic Chemistry*, vol. 91, Art. no. 103128, oct. 2019, doi: 10.1016/j.bioorg.2019.103128.
- [108] F. Peytam et al., “Design and synthesis of new imidazo[1,2-b]pyrazole derivatives, in vitro α -glucosidase inhibition, kinetic and docking studies”, *Mol Divers*, vol. 24, no. 1, pp. 69–80, mar. 2019, doi: 10.1007/s11030-019-09925-8.
- [109] W.-J. Dan, Q. Zhang, F. Zhang, W.-W. Wang, y J.-M. Gao, “Benzonate derivatives of acetophenone as potent α -glucosidase inhibitors: synthesis, structure–activity relationship and mechanism”, *Journal of Enzyme Inhibition and Medicinal Chemistry*, vol. 34, no. 1, pp. 937–945, ene. 2019, doi: 10.1080/14756366.2019.1604519.
- [110] K. Madhu Kumar Reddy et al., “Ceric ammonium nitrate (CAN) catalyzed synthesis and α -glucosidase activity of some novel tetrahydropyridine phosphonate derivatives”, *Phosphorus, Sulfur, and Silicon and the Related Elements*, vol. 194, no. 8, pp. 812–819, ene. 2019, doi: 10.1080/10426507.2018.1550641.
- [111] B. Rama Krishna, S. Ramakrishna, S. Rajendra, K. Madhusudana, y U. V. Mallavadhani, “Synthesis of some novel orsellinates and lecanoric acid related depsides as α -glucosidase inhibitors”, *Journal of Asian Natural Products Research*, vol. 21, no. 10, pp. 1013–1027, Jul. 2018, doi: 10.1080/10286020.2018.1490274.
- [112] J. Chen, W. Lu, H. Chen, X. Bian, y G. Yang, “A New Series of Salicylic Acid Derivatives as Non-saccharide α -Glucosidase Inhibitors and Antioxidants”, *Biological & Pharmaceutical Bulletin*, vol. 42, no. 2, pp. 231–246, feb. 2019, doi: 10.1248/bpb.b18-00661.
- [113] M. Taha et al., “Synthesis of quinoline derivatives as diabetic II inhibitors and molecular docking studies”, *Bioorganic & Medicinal Chemistry*, vol. 27, no. 18, pp. 4081–4088, Sep. 2019, doi: 10.1016/j.bmc.2019.07.035.
- [114] Q.-N. Tariq et al., “Xanthenone-based hydrazones as potent α -glucosidase inhibitors: Synthesis, solid state self-assembly and in silico studies”, *Bioorganic Chemistry*, vol. 84, pp. 372–383, mar. 2019, doi: 10.1016/j.bioorg.2018.11.053.
- [115] I. Uddin et al., “Synthesis, in vitro alpha glucosidase, urease activities and molecular docking study of bis-indole bearing Schiff base analogs”, *Chemical Data Collections*, vol. 28, Art. no. 100396, ago. 2020, doi: 10.1016/j.cdc.2020.100396.
- [116] R. Rafique et al., “Synthesis of new indazole based dual inhibitors of α -glucosidase and α -amylase enzymes, their in vitro, in silico and kinetics studies”, *Bioorganic Chemistry*, vol. 94, Art. no. 103195, ene. 2020, doi: 10.1016/j.bioorg.2019.103195.
- [117] A.-N. Kawde et al., “Exploring efficacy of indole-based dual inhibitors for α -glucosidase and α -amylase enzymes: In silico, biochemical and kinetic studies”, *International Journal of Biological Macromolecules*, vol. 154, pp. 217–232, Jul. 2020, doi: 10.1016/j.ijbiomac.2020.03.090.
- [118] M. Taha et al., “Synthesis of Novel Triazinoindole-Based Thiourea Hybrid: A Study on α -Glucosidase Inhibitors and Their Molecular Docking”, *Molecules*, vol. 24, no. 21, Art. no. 3819, oct. 2019, doi: 10.3390/molecules24213819.
- [119] S. Shamim et al., “Synthesis and screening of (E)-3-(2-benzylidenehydrazinyl)-5,6-diphenyl-1,2,4-triazine analogs as novel dual inhibitors of α -amylase and α -glucosidase”, *Bioorganic Chemistry*, vol. 101, Art. no. 103979, ago. 2020, doi: 10.1016/j.bioorg.2020.103979.

- [120] H. A. Abuelizz, N. A. N. I. Iwana, R. Ahmad, E.-H. Anouar, M. Marzouk y R. Al-Salahi, "Synthesis, biological activity and molecular docking of new tricyclic series as α -glucosidase inhibitors," *BMC Chem.*, vol. 13, no. 1, abr. 2019, Art. no. 52, doi: 10.1186/s13065-019-0560-4.
- [121] S. Mor, S. Sindhu, M. Khatri, N. Singh, N. Vasudeva, y N. Panihar, "Synthesis, Type II Diabetes Inhibitory Activity, and Antimicrobial Tests of Benzothiazole Derivatives Bridged with Indenedione by Methylenehydrazone", *Russ J Gen Chem*, vol. 89, no. 9, pp. 1867–1873, Sep. 2019, doi: 10.1134/s1070363219090226.
- [122] D. Shahzad et al., "Novel C-2 Symmetric Molecules as α -Glucosidase and α -Amylase Inhibitors: Design, Synthesis, Kinetic Evaluation, Molecular Docking and Pharmacokinetics", *Molecules*, vol. 24, no. 8, Art. no. 1511, abr. 2019, doi: 10.3390/molecules24081511.
- [123] X.-T. Yang, C.-A. Geng, T.-Z. Li, Z.-T. Deng, y J.-J. Chen, "Synthesis and biological evaluation of chepraecoxin A derivatives as α -glucosidase inhibitors", *Bioorganic & Medicinal Chemistry Letters*, vol. 30, no. 8, Art. no. 127020, abr. 2020, doi: 10.1016/j.bmcl.2020.127020.
- [124] L. Hao et al., "Synthesis of tetracyclic oxindoles and evaluation of their α -glucosidase inhibitory and glucose consumption-promoting activity", *Bioorganic & Medicinal Chemistry Letters*, vol. 30, no. 14, Art. no. 127264, Jul. 2020, doi: 10.1016/j.bmcl.2020.127264.
- [125] M. Adib et al., "Design and synthesis of new fused carbazole-imidazole derivatives as anti-diabetic agents: In vitro α -glucosidase inhibition, kinetic, and in silico studies", *Bioorganic & Medicinal Chemistry Letters*, vol. 29, no. 5, pp. 713–718, mar. 2019, doi: 10.1016/j.bmcl.2019.01.012.
- [126] Y. El Bakri et al., "Synthesis, biological activity and molecular modeling of a new series of condensed 1,2,4-triazoles", *Bioorganic Chemistry*, vol. 92, Art. no. 103193, nov. 2019, doi: 10.1016/j.bioorg.2019.103193.
- [127] S. Hameed et al., "Synthesis of benzotriazoles derivatives and their dual potential as α -amylase and α -glucosidase inhibitors in vitro: Structure-activity relationship, molecular docking, and kinetic studies", *European Journal of Medicinal Chemistry*, vol. 183, Art. no. 111677, dic. 2019, doi: 10.1016/j.ejmech.2019.111677.
- [128] G. Lavanya, K. Venkatapathy, C. J. Magesh, M. Ramanathan, y R. Jayasudha, "The first target specific, highly diastereoselective synthesis, design and characterization of pyranoquinoliny acrylic acid diastereomers as potential α -glucosidase inhibitors", *Bioorganic Chemistry*, vol. 84, pp. 125–136, mar. 2019, doi: 10.1016/j.bioorg.2018.11.026.
- [129] S. N. Meena, U. kumar, M. M. Naik, S. C. Ghadi, y S. G. Tilve, " α -Glucosidase inhibition activity and in silico study of 2-(benzo[d][1,3]dioxol-5-yl)-4H-chromen-4-one, a synthetic derivative of flavone", *Bioorganic & Medicinal Chemistry*, vol. 27, no. 12, pp. 2340–2344, Jun. 2019, doi: 10.1016/j.bmc.2018.12.021.
- [130] A. A. Spasov et al., "Synthesis, in vitro and in vivo evaluation of 2-aryl-4H-chromene and 3-aryl-1H-benzo[f]chromene derivatives as novel α -glucosidase inhibitors", *Bioorganic & Medicinal Chemistry Letters*, vol. 29, no. 1, pp. 119–123, ene. 2019, doi: 10.1016/j.bmcl.2018.10.018.
- [131] L. Dinparast et al., "An efficient, catalyst-free, one-pot synthesis of 4H-chromene derivatives and investigating their biological activities and mode of

- interactions using molecular docking studies”, *Journal of Molecular Structure*, vol. 1203, Art. no. 127426, mar. 2020, doi: 10.1016/j.molstruc.2019.127426.
- [132] W. Worawalai, T. Doungwichitkul, W. Rangubpit, P. Taweecat, P. Sompornpisut, y P. Phuwapraisirisan, “Furofuran lignans as a new series of antidiabetic agents exerting α -glucosidase inhibition and radical scavenging: Semisynthesis, kinetic study and molecular modeling”, *Bioorganic Chemistry*, vol. 87, pp. 783–793, Jun. 2019, doi: 10.1016/j.bioorg.2019.03.077.
- [133] L. J. Kumar, Y. Suresh, R. Rajasekaran, S. R. Reddy, y V. Vijayakumar, “Synthesis and exploration of in-silico and in-vitro α -glucosidase and α -amylase inhibitory activities of N-(3-acetyl-2-methyl-4-phenylquinolin-6-yl)arylamides”, *J IRAN CHEM SOC*, vol. 16, no. 5, pp. 1071–1080, ene. 2019, doi: 10.1007/s13738-018-01580-4.
- [134] S. W. Leong, T. Awini, S. M. Mohd Faudzi, M. Maulidiani, K. Shaari, y F. Abas, “Synthesis and biological evaluation of asymmetrical diarylpentanoids as antiinflammatory, anti- α -glucosidase, and antioxidant agents”, *Med Chem Res*, vol. 28, no. 11, pp. 2002–2009, ago. 2019, doi: 10.1007/s00044-019-02430-5.
- [135] H. Mehfooz et al., “Facile one-pot synthesis, butyrylcholinesterase and α -glucosidase inhibitory activities, structure–activity relationship, molecular docking and DNA–drug binding analysis of Meldrum’s acid derivatives”, *Res Chem Intermed*, vol. 46, no. 5, pp. 2437–2456, feb. 2020, doi: 10.1007/s11164-020-04100-3.
- [136] A. Mendieta-Moctezuma et al., “Synthesis and biological evaluation of novel chromonyl enamines as α -glucosidase inhibitors”, *Med Chem Res*, vol. 28, no. 6, pp. 831–848, abr. 2019, doi: 10.1007/s00044-019-02320-w.
- [137] S. Mor y S. Sindhu, “Synthesis, Type II diabetes inhibitory activity, antimicrobial evaluation and docking studies of indeno[1,2-c]pyrazol-4(1H)-ones”, *Med Chem Res*, vol. 29, no. 1, pp. 46–62, oct. 2019, doi: 10.1007/s00044-019-02457-8.
- [138] X. Chen et al., “Design, synthesis and α -glucosidase inhibition study of novel embelin derivatives”, *Journal of Enzyme Inhibition and Medicinal Chemistry*, vol. 35, no. 1, pp. 565–573, ene. 2020, doi: 10.1080/14756366.2020.1715386.
- [139] Y. Hu, B. Wang, J. Yang, T. Liu, J. Sun, y X. Wang, “Synthesis and biological evaluation of 3-aryl coumarin derivatives as potential anti-diabetic agents”, *Journal of Enzyme Inhibition and Medicinal Chemistry*, vol. 34, no. 1, pp. 15–30, oct. 2018, doi: 10.1080/14756366.2018.1518958.
- [140] E. Karakılıç, Ş. Baran, H. Ögütçü, A. Akdemir y A. Baran, “rac- and meso-Cyclohexanoids: Their α -, β -glycosidases, antibacterial, antifungal activities, and molecular docking studies,” *Arch. Pharm.*, vol. 353, no. 3, mar. 2020, Art. no. e1900267, doi: 10.1002/ardp.201900267.
- [141] R. Mall, A. Singh, G. Singh, V. Singh, y R. K. Verma, “Indolyl Linked Meta-Substituted Benzylidenes as Novel Ligands: Synthesis, Biological Evaluation, and Molecular Docking Studies”, *Journal of Heterocyclic Chem*, vol. 56, no. 5, pp. 1542–1552, mar. 2019, doi: 10.1002/jhet.3529.
- [142] F. Peytam et al., “An efficient and targeted synthetic approach towards new highly substituted 6-amino-pyrazolo[1,5-a]pyrimidines with α -glucosidase inhibitory activity,” *Sci. Rep.*, vol. 10, no. 1, feb. 2020, Art. no. 2595, doi: 10.1038/s41598-020-59079-z.

- [143] A. M. Al-Majid et al., "Synthesis of Pyridine-Dicarboxamide-Cyclohexanone Derivatives: Anticancer and α -Glucosidase Inhibitory Activities and In Silico Study", *Molecules*, vol. 24, no. 7, Art. no. 1332, abr. 2019, doi: 10.3390/molecules24071332.
- [144] P. Paudel, S. H. Seong, H. J. Park, H. A. Jung, y J. S. Choi, "Anti-Diabetic Activity of 2,3,6-Tribromo-4,5-Dihydroxybenzyl Derivatives from *Symphycyclia latiuscula* through PTP1B Downregulation and α -Glucosidase Inhibition", *Marine Drugs*, vol. 17, no. 3, Art. no. 166, mar. 2019, doi: 10.3390/md17030166.
- [145] V. Santhanam, P. Pant, B. Jayaram, y N. G. Ramesh, "Design, synthesis and glycosidase inhibition studies of novel triazole fused iminocyclitol- δ -lactams", *Org. Biomol. Chem.*, vol. 17, no. 5, pp. 1130–1140, 2019, doi: 10.1039/c8ob03084g.
- [146] J. Graton, J.-Y. Le Questel, P. Maxwell, y P. Popelier, "Hydrogen-Bond Accepting Properties of New Heteroaromatic Ring Chemical Motifs: A Theoretical Study", *J. Chem. Inf. Model.*, vol. 56, no. 2, pp. 322–334, feb. 2016, doi: 10.1021/acs.jcim.5b00574.
- [147] E. A. Meyer, R. K. Castellano, y F. Diederich, "Interactions with Aromatic Rings in Chemical and Biological Recognition", *Angew Chem Int Ed*, vol. 42, no. 11, pp. 1210–1250, mar. 2003, doi: 10.1002/anie.200390319.
- [148] Q. Guan et al., "Triazoles in Medicinal Chemistry: Physicochemical Properties, Bioisosterism, and Application", *J. Med. Chem.*, vol. 67, no. 10, pp. 7788–7824, may 2024, doi: 10.1021/acs.jmedchem.4c00652.
- [149] J. Shao et al., "The Role of Tryptophan in π Interactions in Proteins: An Experimental Approach", *J. Am. Chem. Soc.*, vol. 144, no. 30, pp. 13815–13822, Jul. 2022, doi: 10.1021/jacs.2c04986.
- [150] A. W. E. Chan, R. A. Laskowski, y D. L. Selwood, "Chemical Fragments that Hydrogen Bond to Asp, Glu, Arg, and His Side Chains in Protein Binding Sites", *J. Med. Chem.*, vol. 53, no. 8, pp. 3086–3094, mar. 2010, doi: 10.1021/jm901696w.
- [151] S. W. Lee et al., "Effect of electron-withdrawing fluorine and cyano substituents on photovoltaic properties of two-dimensional quinoxaline-based polymers," *Sci. Rep.*, vol. 11, no. 1, dic. 2021, Art. no. 24381, doi: 10.1038/s41598-021-03763-1.
- [152] A. Gumus et al., "Thiosemicarbazide-Substituted Coumarins as Selective Inhibitors of the Tumor Associated Human Carbonic Anhydrases IX and XII", *Molecules*, vol. 27, no. 14, Art. no. 4610, Jul. 2022, doi: 10.3390/molecules27144610.
- [153] Y.-H. Peng et al., "Unique Sulfur–Aromatic Interactions Contribute to the Binding of Potent Imidazothiazole Indoleamine 2,3-Dioxygenase Inhibitors", *J. Med. Chem.*, vol. 63, no. 4, pp. 1642–1659, ene. 2020, doi: 10.1021/acs.jmedchem.9b01549.
- [154] Y. Shen, X. H. Qin, y L. M. Ren, "Crystal Structure of the C-terminal Subunit of Human Maltase-Glucoamylase in Complex with Acarbose". *Worldwide Protein Data Bank*, nov. 23, 2011. doi: 10.2210/pdb3top/pdb.
- [155] P. Y. Piñero, L. Arco, M. M. García y L. Acevedo, "Algoritmos genéticos en la construcción de funciones de pertenencia borrosas," *Inteligencia Artificial. Revista Iberoamericana de Inteligencia Artificial*, vol. 7, no. 18, pp. 25–35, 2003. [En línea]. Disponible en: <https://www.redalyc.org/articulo.oa?id=92501803>
- [156] A. Golbraikh y A. Tropsha, "Beware of q^2 !", *Journal of Molecular Graphics and Modelling*, vol. 20, no. 4, pp. 269–276, ene. 2002, doi: 10.1016/s1093-3263(01)00123-1.

- [157] B. M. Francuski, S. B. Novaković y G. A. Bogdanović, "Electronic features and hydrogen bonding capacity of the sulfur acceptor in thioureido-based compounds. Experimental charge density study of 4-methyl-3-thiosemicarbazide," *CrystEngComm*, vol. 13, no. 10, pp. 3580–3591, 2011, doi: 10.1039/c0ce00760a.
- [158] G. B. McGaughey, M. Gagné, y A. K. Rappé, "π-Stacking Interactions", *Journal of Biological Chemistry*, vol. 273, no. 25, pp. 15458–15463, Jun. 1998, doi: 10.1074/jbc.273.25.15458.
- [159] L. Ren et al., "Structural insight into substrate specificity of human intestinal maltase-glucoamylase", *Protein Cell*, vol. 2, no. 10, pp. 827–836, oct. 2011, doi: 10.1007/s13238-011-1105-3.
- [160] R. Huey, G. M. Morris, A. J. Olson, y D. S. Goodsell, "A semiempirical free energy force field with charge-based desolvation", *J Comput Chem*, vol. 28, no. 6, pp. 1145–1152, feb. 2007, doi: 10.1002/jcc.20634.
- [161] M. M. Mysinger y B. K. Shoichet, "Rapid Context-Dependent Ligand Desolvation in Molecular Docking", *J. Chem. Inf. Model.*, vol. 50, no. 9, pp. 1561–1573, ago. 2010, doi: 10.1021/ci100214a.
- [162] L. Ayoub, E. Aissam, K. Yassine, E. Said, E. M. Mohammed y A. Souad, "A specific QSAR model for proteasome inhibitors from *Olea europaea* and *Ficus carica*," *Bioinformation*, vol. 14, no. 7, pp. 384–392, jul. 2018, doi: 10.6026/97320630014384.
- [163] P. Gramatica, "Principles of QSAR models validation: internal and external", *QSAR Comb. Sci.*, vol. 26, no. 5, pp. 694–701, may 2007, doi: 10.1002/qsar.200610151.
- [164] K. Roy, S. Kar, y R. N. Das, *A Primer on QSAR/QSPR Modeling*. Springer International Publishing, 2015. doi: 10.1007/978-3-319-17281-1.
- [165] P. K. Ojha y K. Roy, "Comparative QSARs for antimalarial endochins: Importance of descriptor-thinning and noise reduction prior to feature selection", *Chemometrics and Intelligent Laboratory Systems*, vol. 109, núm. 2, pp. 146–161, dic. 2011, doi: 10.1016/j.chemolab.2011.08.007.
- [166] G. Schüürmann, R.-U. Ebert, J. Chen, B. Wang, y R. Kühne, "External Validation and Prediction Employing the Predictive Squared Correlation Coefficient-Test Set Activity Mean vs Training Set Activity Mean", *J. Chem. Inf. Model.*, vol. 48, núm. 11, pp. 2140–2145, oct. 2008, doi: 10.1021/ci800253u.
- [167] E. W. Weisstein, "Balaban Index," *MathWorld—A Wolfram Web Resource*, Wolfram Research, Inc. [En línea]. Disponible en: <https://mathworld.wolfram.com/BalabanIndex.html>. [Accedido: 26-ago-2026].
- [168] S. A. Wildman y G. M. Crippen, "Prediction of Physicochemical Parameters by Atomic Contributions", *J. Chem. Inf. Comput. Sci.*, vol. 39, núm. 5, pp. 868–873, ago. 1999, doi: 10.1021/ci990307l.
- [169] P. Labute, "A widely applicable set of descriptors," *J. Mol. Graph. Model.*, vol. 18, núms. 4–5, pp. 464–477, 2000, doi: 10.1016/S1093-3263(00)00068-1.
- [170] M. Ghorbanzadeh, J. Zhang y P. L. Andersson, "Binary classification model to predict developmental toxicity of industrial chemicals in zebrafish," *J. Chemometrics*, vol. 30, núm. 6, pp. 298–307, 2016, doi: 10.1002/cem.2791.

- [171] Alvascience, alvaDesc, ver. 3.0.16, 2026. [En línea]. Disponible en: <https://www.alvascience.com/alvadesc/> [Accedido: 26-ago-2026].
- [172] Y.-T. Lin, "A tandem regression-outlier analysis of a ligand cellular system for key structural modifications around ligand binding," J. Cheminform., vol. 5, núm. 1, art. núm. 21, abr. 2013, doi: 10.1186/1758-2946-5-21.
- [173] R. E. Carhart, D. H. Smith y R. Venkataraghavan, "Atom pairs as molecular features in structure-activity studies: Definition and applications," J. Chem. Inf. Comput. Sci., vol. 25, núm. 2, pp. 64–73, may. 1985, doi: 10.1021/ci00046a002.

Anexos

Parámetros de validación interna

La validación interna de un modelo QSAR se realiza por medio de las moléculas utilizadas para el desarrollo de dicho modelo y también implica la predicción de la inhibición contra la enzima de las moléculas estudiadas, esto mediante la estimación de parámetros que ayudan a detectar la precisión de las predicciones, es decir, considera que el intervalo de confianza de la muestra sea aceptable [164]

Coeficiente de Determinación R^2

Permite calcular el grado de correlación entre los datos experimentales y los datos predichos del conjunto de datos utilizados para el desarrollo del modelo [164].

$$R^2 = 1 - \frac{\sum (y_{obs} - y_{calc})^2}{\sum (y_{obs} - \bar{y}_{obs})^2} \quad \text{Ecuación 11}$$

Donde:

y_{obs} : El valor de nuestro dato experimental.

y_{calc} : El valor calculado.

$\bar{y}_{obs}$: Promedio del valor experimental.

Validación cruzada de exclusión (Q^2_{LOO})

La validación cruzada Leave-One-Out (LOO, por sus siglas en inglés), consiste en dejar fuera una molécula del modelo QSAR el cual posteriormente se reconstruye a partir de las moléculas restantes del conjunto, utilizando la combinación de descriptores originalmente seleccionados, y la propiedad de la molécula que se dejó fuera se calcula en función de la ecuación QSAR resultante. Este proceso se repite hasta que todas las moléculas del conjunto hayan sido dejadas fuera una vez [164].

$$Q^2 = 1 - \frac{\sum (y_{obs(train)} - y_{pred(train)})^2}{\sum (y_{obs(train)} - \bar{y}_{training})^2} \quad \text{Ecuación 12}$$

Donde:

$y_{obs(train)}$: Valor experimental del conjunto de moléculas basado en la técnica LOO

$y_{pred(train)}$: Valor calculado del conjunto de moléculas basado en la técnica de LOO.

$\bar{y}_{training}$: Promedio del valor experimental del conjunto de moléculas basado en la técnica de LOO.

Si $Q^2_{LOO} > 0.5$ entonces el modelo es aceptable.

La métrica r_m^2

Un valor aceptable de Q^2 no indica necesariamente que los datos de actividad previstos sean muy próximos a los observados, aunque puede existir una buena correlación general entre los valores. Por lo tanto, para obviar este problema y para indicar mejor la predictibilidad del modelo, se realizan validaciones con las métricas r_m^2 pueden calcularse mediante las siguientes ecuaciones [164]:

$$\bar{r}_m^2 = \frac{(r_m^2 + r_m'^2)}{2} \quad \text{Ecuación 13}$$

$$\Delta r_m^2 = |r_m^2 - r_m'^2| \quad \text{Ecuación 14}$$

Donde:

$$r_m^2 = r^2 \times \left(1 - \sqrt{(r^2 - r_0^2)}\right) \quad \text{Ecuación 15}$$

$$r_m'^2 = r^2 \times \left(1 - \sqrt{(r^2 - r_0'^2)}\right) \quad \text{Ecuación 16}$$

Los parámetros r^2 y r_0^2 son los coeficientes de correlación entre los valores experimentales y los valores calculados con y sin intersección respectivamente. El parámetro $r_0'^2$ tienen el mismo significado de r_0^2 , pero los ejes se invierten.

$\bar{r}_m^2$ es el valor promedio de r_m^2 y $r_m'^2$, y Δr_m^2 es el valor absoluto de la diferencia entre $r_m^2 - r_m'^2$. En el caso de la validación interna del set de entrenamiento $r_{m(L00)}^2$ y $\Delta r_{m(L00)}^2$ pueden ser empleados y se ha demostrado que el valor de $\Delta r_{m(L00)}^2 < 0.2$ y $r_{m(L00)}^2 > 0.5$ indicando que el modelo es predictivo [165].

Parámetros de validación externa

Para determinar la predictibilidad de un modelo QSAR no basta con solo la validación interna, esto ya que solo considera a los compuestos que pertenecen al mismo conjunto utilizado para el desarrollo del modelo, por lo que es importante realizar una validación externa, donde se emplearan compuestos que no pertenezcan al conjunto utilizado previamente[164].

Coeficiente de determinación predictivo (R_{pred}^2 o $Q_{(F1)}^2$)

Este coeficiente refleja el nivel de correlación entre los datos experimentales y los predichos en el conjunto de predicción.

$$R_{pred}^2 = 1 - \frac{\sum (y_{obs(test)} - y_{pred(test)})^2}{\sum (y_{obs(test)} - \bar{y}_{training})^2} \quad \text{Ecuación 17}$$

Donde:

$y_{obs(test)}$: pIC_{50} experimental del conjunto de predicción

$y_{pred(test)}$: pIC_{50} calculado del conjunto de predicción

$\bar{y}_{training}$: Promedio del pIC_{50} experimental del conjunto de entrenamiento con el que se desarrolló el modelo.

Los modelos que con valores de $R_{pred}^2 > 0.5$ se consideran predictivos [164]

Variante del coeficiente de determinación predictivo Q^2 basada en el conjunto de prueba ($Q_{(F2)}^2$)

$$Q_{(F2)}^2 = 1 - \frac{\sum (y_{obs(test)} - y_{pred(test)})^2}{\sum (y_{obs(test)} - \bar{y}_{test})^2} \quad \text{Ecuación 18}$$

Donde:

$y_{obs(test)}$ pIC_{50} experimental del conjunto de predicción

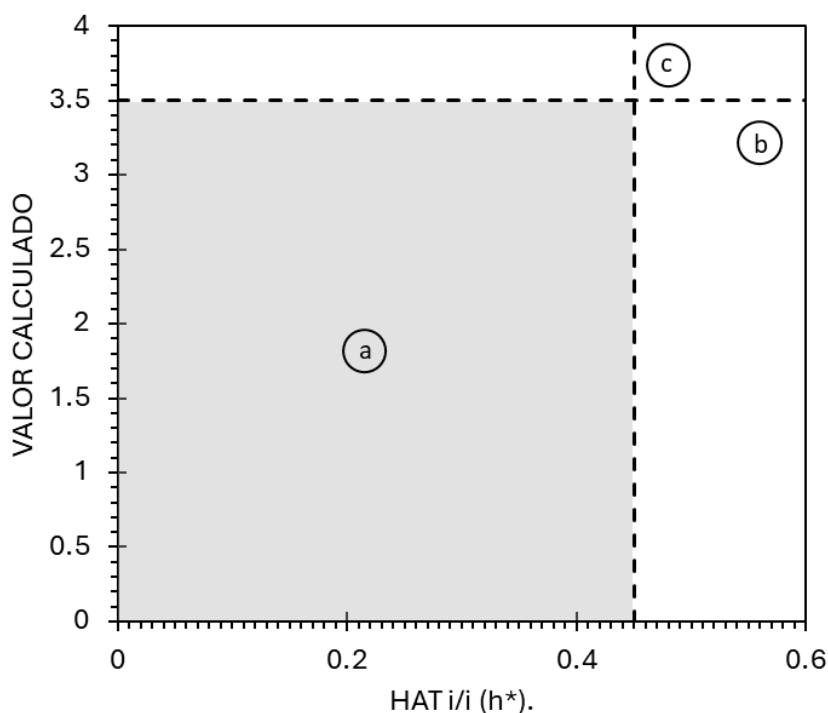
$y_{pred(test)}$: pIC_{50} calculado del conjunto de predicción

$\bar{y}_{test}$: Promedio del pIC_{50} experimental del conjunto de predicción

Para poder validar el modelo se debe cumplir con la condición de $Q_{(F2)}^2 < Q_{(F1)}^2$ [164], [166].

Dominio de Aplicabilidad

El dominio de aplicabilidad (DA) se define como la región teórica determinada por los descriptores del modelo y la respuesta modelada y, por consiguiente, de la naturaleza de los compuestos que participan en dicho modelado, en función de las variables del modelo [164].



Gráfica. Ejemplo de un dominio de aplicabilidad.

En la Gráfica el apartado **a** es la zona de interpolación y representa la región donde la respuesta calculada para las moléculas es confiable, el apartado **b** (línea punteada horizontal) representa el valor máximo calculado de los compuestos del conjunto de entrenamiento y el punto **c** (línea punteada vertical) representa la influencia crítica h^* (hat crítico) y se calcula de la siguiente manera:

$$\frac{3p'}{n} \quad \text{Ecuación 19}$$

Donde:

p' : Número de descriptores que contiene el modelo más uno.

n : Número de compuestos del conjunto base de entrenamiento.

HAT i/i representa la influencia de las moléculas y está definido como:

$$HAT\ i/i = x_i^T (X^T X)^{-1} x_i \quad (i = 1, \dots, n) \quad \text{Ecuación 20}$$

Donde:

i : Compuesto

x_i : Descriptor fila-vector de cada compuesto

X : Matriz $n \times k-1$ de k valores de los descriptores del modelo para n compuestos de conjunto base.

El superíndice T hace referencia a la transpuesta de la matriz o vector

Definición de los descriptores del Modelo 4

BalabanJ

El J de Balaban [167] (denominado formalmente como conectividad de suma de distancias promediadas) es un índice topológico basado en distancias diseñado para traducir la constitución química de una molécula en valores numéricos útiles para correlacionar propiedades físicas o relaciones biológicas de estructura-actividad. Su característica principal y el motivo de su creación es su muy baja degeneración, lo que lo convierte en un índice altamente discriminatorio con una capacidad superior para diferenciar entre isómeros constitucionales en comparación con otros índices topológicos propuestos anteriormente.

Matemáticamente, el índice J se basa en una fórmula del tipo Randić, pero reemplazando los grados de los vértices por sumas de distancias promediadas. La fórmula general corregida para anillos se define de la siguiente manera:

$$J = \frac{q}{\mu + 1} \sum_{adjac. i, j} (s_i s_j) \quad \text{Ecuación 21}$$

Los componentes de esta fórmula representan lo siguiente:

q : Es el número de adyacencias de los vértices, lo que equivale al número de enlaces en grafos simples de moléculas saturadas.

μ : Es el número ciclomático, introducido para permitir una mejor comparación entre grafos cíclicos y acíclicos. Se calcula como $\mu = q - n + 1$, donde n representa el número total de vértices (átomos distintos de hidrógeno) en el grafo.

s_i y s_j : Son las sumas de distancias para los vértices adyacentes i y j . Estas sumas se calculan sumando las entradas en la fila o columna correspondientes a cada vértice dentro de la matriz de distancias (DM) de la molécula.

Para moléculas acíclicas saturadas (grafos simples sin anillos), la fórmula se simplifica a:

$$J_{acyclic} = q \sum_{adjac. i, j} (s_i s_j)^{-\frac{1}{2}} \quad \text{Ecuación 22}$$

Por otro lado, cuando la definición de J se aplica a moléculas insaturadas o aromáticas (multigrafos), se emplean números fraccionarios en la matriz de distancias que dependen del orden del enlace (se usa $\frac{1}{b}$, donde b es el orden del enlace) al momento de calcular las sumas de distancias s_i

Contribuciones individuales al área de superficie de Van der Waals (VSA) que caen en un rango del logP.

Estos descriptores pertenecen a un conjunto de 10 descriptores moleculares, diseñados para capturar los efectos hidrofóbicos e hidrofílicos de una molécula. Para entender cómo se define específicamente SlogP_VSA4, es necesario comprender la fórmula base de los descriptores de tipo VSA:

Se calcula sumando las contribuciones individuales al área de superficie de van der Waals (VSA) de cada átomo en una molécula, siempre y cuando el valor de una propiedad numérica específica de ese átomo (P_i) caiga dentro de un rango determinado. En este caso, la propiedad evaluada es la contribución atómica al logP (el coeficiente de partición octanol/agua), utilizando el método de cálculo de Wildman y Crippen [168].

El autor define estadísticamente 10 intervalos (para crear 10 descriptores SlogP) basándose en los siguientes límites. Si seguimos la fórmula P_VSA_k donde k es el número del descriptor, podemos deducir los rangos:

- SlogP_VSA1: $-\infty$ a -0.4
- SlogP_VSA2: -0.4 a -0.2
- SlogP_VSA3: -0.2 a 0.0
- SlogP_VSA4: 0.0 a 0.1
- SlogP_VSA5: 0.1 a 0.15
- SlogP_VSA6: 0.15 a 0.2
- SlogP_VSA7: 0.2 a 0.25
- SlogP_VSA8: 0.25 a 0.3
- SlogP_VSA9: 0.3 a 0.4
- SlogP_VSA10: 0.4 a ∞

Por lo tanto, **SlogP_VSA4** se define exactamente como la suma del área de superficie de Van der Waals de todos los átomos de una molécula cuya contribución individual al logP se encuentra en el rango de 0 a 0.1 [168], [169].

nDB

El descriptor nDB se define como el número de dobles enlaces (por sus siglas en inglés, *number of Double Bonds*) de un compuesto no aromático. Es un descriptor de tipo constitucional, que refleja la composición molecular de un compuesto químico sin proporcionar ninguna información sobre su geometría molecular [171].

Índices de Adyacencia de Aristas

Los índices de adyacencia de aristas [172] son índices topológicos basados en la matriz de adyacencia de aristas (EA), las matrices de adyacencia de aristas aumentadas ($AEA(w)$) y las matrices de adyacencia de aristas ponderadas asimétricas ($EA(w)$) de un grafo molecular con deficiencia de hidrogeno.

La matriz de adyacencia de aristas (EA) es una matriz cuadrada simétrica de dimensión nBO, donde nBO es el número de enlaces en un grafo molecular con deficiencia de hidrógeno. Las entradas de la matriz son iguales a uno si los enlaces considerados son adyacentes y cero en caso contrario.

$$[EA]_{ij} = \begin{cases} 1 & \text{si } i, j \text{ son enlaces adyacentes, en caso contrario} \end{cases} \quad \text{Ecuación 23}$$

La matriz de adyacencia de aristas aumentada ($AEA(W)$) se obtiene a partir de la matriz de adyacencia de aristas reemplazando los elementos diagonales cero de la matriz de adyacencia de aristas con propiedades de enlace específicas.

$$[AEA(w)]_{ij} = \begin{cases} 1 & \text{si } i, j \text{ son enlaces adyacentes, si } i = j, \text{ en caso contrario} \\ w_i & \end{cases} \quad \text{Ecuación 24}$$

La matriz de adyacencia de aristas ponderada ($EA(w)$) se obtiene a partir de la matriz de adyacencia de aristas reemplazando los elementos correspondientes a los enlaces adyacentes con propiedades de enlace específicas:

$$[EA(w)]_{ij} = \begin{cases} w_j & \text{si } i, j \text{ son enlaces adyacentes} \\ 0 & \end{cases} \quad \text{Ecuación 25}$$

Si dos aristas i y j son adyacentes, la entrada i, j se define por el peso w_j de la arista j , y la entrada simétrica j, i se define por el peso w_i de la arista i . Todos los demás elementos se establecen en 0.

Las matrices de adyacencia de aristas aumentadas y ponderadas se definen utilizando los siguientes pesos:

- Grado de arista (ed)
- Orden de enlace (bo)

- Momentos dipolares (dm)
- Integrales de resonancia (ri)

El bloque de índices de adyacencia de aristas proporciona cuatro grupos lógicos de descriptores:

- Índices espectrales
- Índices de conectividad
- Momentos espectrales
- Autovalores

Momentos Espectrales

Se calcula la matriz de adyacencia de aristas ponderada ($EA(w)$) con la siguiente ecuación:

$$SMk_M(w) = \ln \left(1 + \sum_{i=1}^{nBO} \lambda_i^k \right) \quad \text{Ecuación 26}$$

donde nBO es el número de enlaces en el grafo molecular empobrecido en H, M es la matriz seleccionada, w es el peso seleccionado, λ es el valor propio y k es el orden (de 2 a 15). Por lo que nuestro descriptor es el Momento Espectral de orden 15 de la matriz de adyacencia de aristas ponderada asimétrica ponderada por el orden de enlace:

$$SM15_M(bo) = \ln \left(1 + \sum_{i=1}^{nBO} \lambda_i^{15} \right) \quad \text{Ecuación 27}$$

Descriptores RDF

Los descriptores RDF [172] toman su nombre de la función de distribución radial. La función de distribución radial de un conjunto de n átomos puede interpretarse como la distribución de probabilidad de encontrar un átomo en un volumen esférico dado.

Se calculan utilizando la siguiente fórmula:

$$RDFRw = \sum_{i=1}^{nAT-1} \sum_{j=i+1}^{nAT} w_i \cdot w_j \cdot e^{-B(R-r_{ij})^2} \quad \text{Ecuación 28}$$

Donde:

nAT : es el número de átomos.

w : es uno de los siguientes esquemas de ponderación:

- u : sin ponderación; en este caso, w_i y w_j son iguales a 1.
- m : masa atómica.
- v : volumen atómico de van der Waals.
- e : electronegatividad de Sanderson.
- p : polarizabilidad atómica.
- i : potencial de ionización atómica.
- s : estado intrínseco.
- carga: carga parcial.

R : radio del volumen esférico.

r_{ij} : es la distancia entre los átomos i y j .

B : es el parámetro de suavizado que define la distribución de probabilidad de las distancias individuales.

Por lo que nuestro descriptor se definiría como: La Función de la Distribución Radial a 2.0 Å, ponderada por el potencial de ionización

$$RDF020i = \sum_{i=1}^{nAT-1} \sum_{j=i+1}^{nAT} w_i \cdot w_j \cdot e^{-B(R-r_{ij})^2} \quad \text{Ecuación 29}$$

Invariante Molecular Holístico Ponderado

Los descriptores WHIM [172] (Por sus siglas en inglés, *Weighted Holistic Invariant Molecular*) contienen información sobre el tamaño, la forma, la simetría y la distribución atómica de la molécula. Debido a su método de cálculo, son invariantes a la traslación y la rotación y, por lo tanto, evitan cualquier problema de alineación molecular.

Se basan en un algoritmo que implica el cálculo de la matriz de covarianza ponderada (S) que es una matriz de tres por tres definida de la siguiente manera:

$$[S]_{ij} = \frac{\sum_{a=1}^{nAT} w_a (q_{ai} - \bar{q}_i) (q_{aj} - \bar{q}_j)}{\sum_{a=1}^{nAT} w_a} \quad i = 1, 2, 3 \text{ and } j = 1, 2, 3 \quad \text{Ecuación 30}$$

donde i y j representan las tres coordenadas posibles (x, y, z), nAT es el número de átomos, w_a es el esquema de ponderación, q_{ai} y q_{aj} son las coordenadas centradas i y j del átomo a , y $\bar{q}_i$ y $\bar{q}_j$ son el valor promedio de las coordenadas i y j , respectivamente. Posteriormente, se realiza un Análisis de Componentes Principales (PCA, por sus siglas en inglés, *Principal Component Analysis*) sobre S ,

obteniendo los autovalores y autovectores correspondientes a los tres ejes de componentes principales.

Los descriptores WHIM se dividen en dos grupos:

- Descriptores direccionales: calculados mediante las proyecciones de los átomos a lo largo de cada eje de componente principal.
- Descriptores globales: calculados como una combinación de los ejes de componentes principales, proporcionando información global sobre la molécula.

Entonces nuestro descriptor se define como el índice de accesibilidad total con esquema ponderado por el estado intrínseco (Ds)

$$Dw = E1w + 2w + E3w \quad \text{Ecuación 31}$$

NaaO

Número de átomos de oxígeno en un anillo aromático [173].

2D pares de átomos

Los pares de átomos binarios/de frecuencia [172], constituyen un caso particular de descriptores de autocorrelación, los descriptores de autocorrelación de tipo atómico (ATAC). Los descriptores de autocorrelación de tipo atómico son descriptores discretos que indican la presencia/ausencia de pares de átomos ("Binarios") o cuentan su frecuencia ("Frecuencia"), donde los pares de átomos se definen como cualquier par de átomos con características particulares a una distancia topológica predefinida k_1 .

Se pueden definir de la siguiente manera:

$$ATAC_k(u, v) = \frac{1}{2} \cdot \sum_{i=1}^{|V|} \sum_{j=1}^{|V|} \delta(i; u) \cdot \delta(j; v) \cdot \delta(d_{ij}; k) \quad \text{Ecuación 32}$$

donde u y v representan dos tipos de átomos diferentes, y $\delta(i; u)$, $\delta(j; v)$ y $\delta(d_{ij}; k)$ son tres funciones delta de Kronecker iguales a uno si el átomo i es de tipo u , el átomo j es de tipo v y la distancia topológica d_{ij} es igual a k , respectivamente; cero en caso contrario.

Por ejemplo en el caso de nuestros descriptores:

B08[C-S]

Tiene un valor de uno si la molécula tiene al menos un caso de 8 enlaces entre átomos de carbono y azufre (C-S) en otras palabras, una distancia topológica de 8 y el mismo descriptor obtiene un valor de cero en caso de no presentar alguno.

B09[C-O]

Tiene un valor de uno si la molécula tiene al menos un caso de distancia topológica de 9 entre átomos de carbono y oxígeno (C-O) y cero en caso contrario.

F07[C-N]

Cuenta todos los pares de átomos de carbono-nitrógeno (C-N) a una distancia topológica de 7 contenidos en la molécula, tomando en cuenta el camino más corto desde un átomo a otro [174].