



**FACULTAD DE CIENCIAS DE LA SALUD**

**Trabajo de Fin de Master Titulado:**

**“Estudio *in silico* mediante modelado QSAR e inteligencia artificial para la predicción de la actividad antimicrobiana de derivados del bencimidazol frente a *Staphylococcus aureus*”**

**Realizado por:**

**LUIS DAVID PANTOJA NARVÁEZ**

**Director del proyecto:**

**Dr. Jesús Antonio Romero L., Ph.D.**

**Como requisito para la obtención del título de:**

**MAGISTER EN BIOMEDICINA**

**QUITO, ABRIL del 2026**

## **DECLARACIÓN JURAMENTADA**

Yo, Luis David Pantoja Narváez, ecuatoriano, con Cédula de ciudadanía N° 0401780622, declaro bajo juramento que el trabajo aquí desarrollado es de mi autoría, que no ha sido presentado anteriormente para ningún grado o calificación profesional, y se basa en las referencias bibliográficas descritas en este documento.

A través de esta declaración, cedo los derechos de propiedad intelectual a la UNIVERSIDAD INTERNACIONAL SEK, según lo establecido en la Ley de Propiedad Intelectual, reglamento y normativa institucional vigente.

-----  
**BQ. F. Luis David Pantoja Narváez**

C.I.: 0401780622

### **DECLARACIÓN DEL DIRECTOR DE TESIS**

Declaro haber dirigido este trabajo a través de reuniones periódicas con el estudiante, orientando sus conocimientos y competencias para un eficiente desarrollo del tema escogido y dando cumplimiento a todas las disposiciones vigentes que regulan los Trabajos de Titulación.

-----

Dr. Jesús Antonio Romero L., Ph.D.

**LOS PROFESORES INFORMANTES:**

ING. MANUEL ANDRÉS HERRERA YELA, MSc.

DR. JOSÉ RUBÉN RAMÍREZ IGLESIAS, PhD.

Después de revisar el trabajo presentado lo han calificado como apto para su defensa  
oral ante el tribunal examinador.

---

Ing. Manuel Andrés Herrera Yela, MSc.

---

Dr. José Rubén Ramírez Iglesias, PhD.

Quito, ABRIL de 2026

## **DECLARACIÓN DE AUTORÍA DEL ESTUDIANTE**

Declaro que este trabajo es original, de mi autoría, que se han citado las fuentes correspondientes y que en su ejecución se respetaron las disposiciones legales que protegen los derechos de autor vigentes.

-----  
**BQ. F. Luis David Pantoja Narváez**

C.I.: 0401780622

## **DEDICATORIA**

Dedico este logro a mis padres, por su apoyo constante, sus valores y su ejemplo de perseverancia, que han sido fundamentales en mi formación personal y académica. A mi hermana y a su familia, por su permanente aliento y comprensión a lo largo de este proceso. Su respaldo ha sido esencial para culminar esta etapa.

## **AGRADECIMIENTO**

Agradezco profundamente a mis padres por su apoyo incondicional, sus valores y su constante motivación a lo largo de mi formación académica. A mi hermana y a su familia, por su acompañamiento, comprensión y palabras de aliento durante este proceso, que han sido fundamentales para culminar esta etapa.

Expreso también mi sincero agradecimiento a mi tutor del Trabajo de Fin de Máster (TFM), Dr. Jesús Romero Lugo, PhD., por su orientación académica, su valiosa asesoría y su disposición durante el desarrollo de este trabajo, contribuyendo de manera significativa a la culminación de esta investigación.

Artículo de tesis

David Pantoja <sup>1</sup>, Jesús Romero <sup>1</sup>

<sup>1</sup> Maestría en Biomedicina, Facultad de Ciencias de la Salud, Universidad Internacional SEK; [luis.pantoja@uisek.edu.ec](mailto:luis.pantoja@uisek.edu.ec)

<sup>1</sup> Maestría en Biomedicina, Facultad de Ciencias de la Salud, Universidad Internacional SEK; [jesus.romero@uisek.edu.ec](mailto:jesus.romero@uisek.edu.ec)

**Resumen:** La resistencia antimicrobiana representa un problema de salud pública a nivel mundial, asociado al incremento de infecciones causadas por bacterias multirresistentes y a la disminución de la eficacia de antibióticos disponibles. *Staphylococcus aureus* ha sido reportado como uno de los principales patógenos a nivel hospitalario y comunitario con alta prevalencia de cepas resistentes. En Ecuador, estudios evidencian una creciente resistencia de *S. aureus* a meticilina, implicando la necesidad de explorar nuevas terapéuticas. En este contexto, el modelado QSAR y la inteligencia artificial surgen como enfoques computacionales para estudiar la relación entre la estructura química y la actividad biológica de compuestos antimicrobianos. Este estudio tuvo como objetivo analizar, *in silico*, la relación estructura-actividad de derivados del anillo benzimidazólico seleccionado como núcleo central del estudio por su presencia recurrente en moléculas con actividad antimicrobiana frente a *Staphylococcus aureus*. A partir de bases de datos públicas y literatura científica se elaboró una matriz de 100 compuestos con el núcleo benzimidazol y valores experimentales de concentración mínima inhibitoria (MIC). Estas estructuras fueron optimizadas geométricamente y se calcularon descriptores moleculares mediante PaDEL-Descriptor, construyéndose una matriz depurada de 636 descriptores. Se elaboraron modelos QSAR empleando ElasticNetCV (Lineal), Random Forest y Support Vector Regression (No lineales), aplicando validación cruzada, análisis de residuos, evaluación del dominio de aplicabilidad y validación externa empleando compuestos independientes. Los resultados proponen que los modelos no lineales reprodujeron adecuadamente patrones descriptor-actividad, mientras que el modelo lineal identificó descriptores asociados a flexibilidad conformacional, superficie polar, hidrofilia y masa molecular, mismas que se relacionan con incrementos en los valores de MIC. El estudio evidencia que el uso integrado de modelado QSAR e inteligencia artificial permite analizar y predecir la actividad antimicrobiana de derivados del benzimidazol, aportando criterios estructurales útiles para futuras etapas de investigación computacional.

**Palabras clave:** Descriptores; Benzimidazol; Antimicrobiano; Algoritmos; Resistencias; *Staphylococcus aureus*; cepas; *in silico*; QSAR.

**Abstract:** Antimicrobial resistance is a global public health problem associated with an increase in infections caused by multidrug-resistant bacteria and a decrease in the effectiveness of available antibiotics. *Staphylococcus aureus* has been reported as one of the main pathogens in hospitals and communities, with a high prevalence of resistant strains. In Ecuador, studies show a growing resistance of *S. aureus* to methicillin, implying the need to explore new therapies. In this context, QSAR modeling and artificial intelligence emerge as computational approaches to study the relationship between the chemical structure and biological activity of antimicrobial compounds. The objective of this study was to analyze, *in silico*, the structure-activity relationship of benzimidazole ring derivatives selected as the central focus of the study due to their recurrent presence in molecules with antimicrobial activity against *Staphylococcus aureus*. A matrix of 100 compounds with the benzimidazole core and experimental minimum inhibitory concentration (MIC) values was developed from public databases and scientific literature. These structures were geometrically optimized, and molecular descriptors were calculated using PaDEL-Descriptor, constructing a refined matrix of 636 descriptors. QSAR models were developed using ElasticNetCV, Random Forest, and Support Vector Regression, applying cross-validation, residual analysis, applicability domain assessment, and external validation using independent compounds. The results suggest that the nonlinear models adequately reproduced descriptor-activity patterns, while the linear model identified descriptors associated with

conformational flexibility, polar surface area, hydrophilicity, and molecular mass, which are related to increases in MIC values. The study shows that the integrated use of QSAR modeling and artificial intelligence allows the analysis and prediction of the antimicrobial activity of benzimidazole derivatives, providing useful structural criteria for future stages of computational research.

**Keywords:** Descriptors; Benzimidazole; Antimicrobial; Algorithms; Resistance; *Staphylococcus aureus*; Strains; In silico; QSAR.

---

## 1. Introducción

La era moderna ha traído avances en medicina y tecnología, pero también grandes desafíos, uno de los más relevantes para la salud pública mundial: la resistencia antimicrobiana (RAM). Esta amenaza creciente representa una importante batalla que la humanidad está comenzando a perder, al observar cómo bacterias patógenas desarrollan mecanismos para evadir los tratamientos farmacológicos existentes, generando la pérdida de sensibilidad frente a múltiples clases o tipos de antibióticos (Organización Mundial de la Salud, 2024). Dentro de este complejo panorama, *Staphylococcus aureus* se ha posicionado como uno de los patógenos más agresivos y adaptativos, siendo capaz de causar desde infecciones cutáneas leves hasta enfermedades potencialmente letales como neumonías, endocarditis o sepsis, entre otras. La diseminación de cepas resistentes a meticilina (MRSA) preocupa, ya que han reducido drásticamente la eficacia de la terapéutica en ambientes hospitalarios y/o comunitarios (Tong et al., 2015; Vinueza et al., 2015).

En Ecuador, estudios recientes confirman una creciente y acelerada prevalencia de *S. aureus* multiresistente, especialmente en hospitales de tercer nivel, donde la tasa de aislamiento de cepas MRSA alcanza hasta un 30% en algunas regiones, empeorando la morbilidad, prolongando las estancias hospitalarias y aumentando los costos del sistema de salud pública, particularmente a nivel de Red Pública Integral de Salud (Vinueza et al., 2015). Esto genera la urgencia de identificar y desarrollar nuevos compuestos antimicrobianos capaces de contrarrestar la amenaza de cepas resistentes.

Frente a esta situación crítica, el desarrollo tradicional de antibióticos caracterizado por ser costoso, prolongado (tiempo) y con alto índice de fallos resulta insuficiente para contribuir con la terapéutica de los pacientes. La ciencia contemporánea ofrece soluciones innovadoras a través del diseño racional de fármacos (DRF), que permite predecir, optimizar y sintetizar compuestos bioactivos mediante el uso de herramientas computacionales avanzadas, todo esto con metodología *In silico* (Paul et al., 2024). Dentro de este enfoque, el modelado Relación Cuantitativa Estructura-Actividad QSAR por sus siglas en inglés Quantitative Structure-Activity

Relationship se ha posicionado como una de las metodologías esenciales al establecer relaciones matemáticas entre la estructura molecular y la actividad biológica de compuestos, permitiendo la predicción y priorización de nuevos candidatos con potencial terapéutico (Roy et al., 2015). Para ello QSAR utiliza descriptores moleculares, estos son variables numéricas que cuantifican propiedades fisicoquímicas, topológicas, geométricas, electrónicas y estructurales de una molécula. Los descriptores permiten transformar la estructura química en información analítica manejable por modelos estadísticos computacionales y algoritmos de aprendizaje automático, permitiendo facilitar la predicción de la actividad biológica de nuevos compuestos (Roy et al., 2015; Todeschini & Consonni, 2009).

La integración de QSAR con algoritmos de Inteligencia Artificial (IA) incluyendo técnicas de aprendizaje automático y aprendizaje profundo (machine learning and deep learning) ha potenciado la precisión de estos modelos, permitiendo analizar grandes cantidades de datos químicos, identificar patrones estructurales ocultos y generar nuevas moléculas con propiedades mejoradas y potenciadas (Yang et al., 2019; Vamathevan et al., 2019). Este enfoque no sólo reduce el tiempo y el costo asociados al desarrollo de nuevos fármacos, sino que también disminuye la necesidad de ensayos *in vivo* tempranos, alineándose con prácticas éticas y sostenibles con el medio ambiente y animal.

En este contexto tecnológico y biomédico, los derivados del bencimidazol surgen como una matriz estructural prometedora. Cuyo núcleo heterocíclico ha demostrado una amplia gama de actividades farmacológicas, entre ellas: antimicrobiana, antiviral, antiparasitaria y antitumoral (Zhou et al., 2013). Su capacidad para interactuar con dianas biológicas esenciales en bacterias, como la DNA girasa y la topoisomerasa IV, lo que lo convierten en una excelente base para el diseño de nuevos agentes anti *S. aureus* (Dokla et al., 2020; He et al., 2004).

Por tanto, la presente investigación propone aprovechar la potente sinergia entre la quimioinformática, el modelado QSAR y la inteligencia artificial para diseñar una ecuación que permita predecir las características fundamentales de nuevas estructuras hacia el diseño racional de nuevos derivados de bencimidazol con actividad antimicrobiana específica contra *S. aureus*. A través de la aplicación rigurosa de métodos *in silico*, modelos QSAR validados estadísticamente y algoritmos de IA de vanguardia, además este estudio utilizará compuestos reportados en bases de datos como PubChem y ChEMBL, todos ellos con valores cuantitativos de concentración mínima inhibitoria MIC, por sus siglas en inglés Minimum Inhibitory Concentration frente a *Staphylococcus aureus*

La hipótesis de este trabajo se basa en que a través del uso combinado de modelado QSAR e inteligencia artificial permitirá predecir de manera confiable la actividad antimicrobiana de

derivados del bencimidazol contra *Staphylococcus aureus* y facilitará el diseño racional de nuevas estructuras moleculares con potencial bioactivo.

Cabe señalar que a través de este enfoque *in silico*, se busca identificar los descriptores moleculares de mayor potencial, establecer un modelo predictivo confiable conceptual y orientar el diseño de nuevas moléculas con bioactividad antimicrobiana. Estos hallazgos, elaborados bajo criterios conceptuales, permiten sentar las bases fundamentales, cuyo fin permita el desarrollo futuro de alternativas farmacoterapéuticas innovadoras, al enfrentar infecciones bacterianas resistentes. Dado que los distintos algoritmos de aprendizaje pueden presentar diferencias entre su desempeño durante la validación interna y su capacidad real de generalización, en esta investigación se realizó una evaluación comparativa de modelos lineales y no lineales. Este análisis consideró no sólo su capacidad de ajuste e interpretación, sino también su desempeño predictivo al aplicarse sobre compuestos externos no incluidos en el entrenamiento.

Por ello, el presente estudio tiene como objetivo construir una ecuación QSAR que permita analizar y predecir actividad antimicrobiana de derivados del bencimidazol frente al patógeno *Staphylococcus aureus*, mediante cálculo de descriptores y construcción de modelos de aprendizaje automático.

## 2. Materiales y Métodos

Este estudio corresponde a una investigación cuantitativa, no experimental, de tipo descriptivo y predictivo, basado en modelado computacional (*in silico*), específicamente mediante técnicas QSAR, permitiendo trabajar con modelos computacionales sin la necesidad de ejecutar ensayos húmedos.

### 2.1. Obtención de datos y selección de compuestos

Se construyó una base de datos compuesta por 100 estructuras químicas derivadas del núcleo bencimidazol, las cuales presentan actividad antimicrobiana previamente reportada frente a *Staphylococcus aureus* en términos de actividad mínima inhibitoria (o su equivalente). Estos compuestos fueron seleccionados a partir de búsquedas en plataformas reconocidas principalmente PubChem (<https://pubchem.ncbi.nlm.nih.gov/>) complementándose con consultas en ChEMBL (<https://www.ebi.ac.uk/chembl/>) y artículos científicos disponibles en PubMed, plataformas

ampliamente utilizadas en estudios QSAR, por su confiabilidad y amplio rigor científico en investigaciones de modelado molecular (Gaulton et al., 2017).

### 2.1.1. Procedimiento de obtención de compuestos en PubChem

La obtención de los compuestos se realizó mediante una búsqueda estructurada en la base de datos **PubChem** (<https://pubchem.ncbi.nlm.nih.gov/>), siguiendo una secuencia operativa orientada a obtener moléculas con actividad antimicrobiana reportada.

Inicialmente, en el buscador principal de PubChem se ingresó el término “**benzimidazole**”, permitiendo acceder a los registros asociados a este núcleo químico. A continuación, se seleccionó la categoría “**Compounds**”, donde se despliega el conjunto de moléculas registradas que contienen la misma estructura.

En esta sección, se procedió a revisar individualmente los compuestos listados. Para cada registro, se accedió a su ficha descriptiva (**Compound Summary**), donde se verificó la presencia del núcleo benzimidazólico mediante el análisis de su **estructura química**, disponible en:

**-Representación bidimensional (2D)**

**-Modelo tridimensional (3D)**

Adicionalmente, desde la misma plataforma se obtuvo la representación estructural en formato **SMILES (Simplified Molecular Input Line Entry System)**, la cual fue utilizada para la estandarización y manejo computacional de las moléculas (ver Anexo 1: Matriz de 100 derivados de Benzimidazol).

Asimismo, se realizó la **descarga de las estructuras moleculares** en formatos estándar empleados en estudios quimiinformáticos, tales como:

**.SDF (Structure Data File)**

**.MOL**

Permitiendo su posterior utilización en procesos de optimización geométrica y cálculo de descriptores moleculares.

Posteriormente, dentro de cada compuesto, se exploró la sección “**BioAssay Results**”, con el objetivo de identificar ensayos biológicos asociados. En esta se revisaron los bioensayos disponibles, cuyo enfoque en aquellos relacionados con **actividad antimicrobiana**, particularmente frente a *Staphylococcus aureus*. Se estudió la información reportada en cada ensayo, incluyendo el tipo de prueba, condiciones experimentales y resultados cuantitativos.

Para confirmar y complementar los datos experimentales, se accedió a la sección “**Literature**” o “**References**” de cada compuesto, donde PubChem enlaza artículos científicos relacionados. A través de los enlaces, principalmente hacia **PubMed**, se consultaron las publicaciones originales en las que se reportan valores experimentales como **MIC (Minimum Inhibitory Concentration)**, permitiendo validar la información biológica asociada a cada molécula.

Adicionalmente, dentro de la ficha del compuesto, se recolectó la información disponible en la sección “**Chemical and Physical Properties**”, incluyendo propiedades fisicoquímicas relevantes y descriptores básicos, complementando la caracterización estructural de cada compuesto.

Cabe indicar, de manera complementaria, **algunos compuestos fueron obtenidos de otras bases de datos quimioinformáticas, como ChEMBL**, siguiendo un procedimiento análogo al descrito para PubChem, lo que permitió ampliar el conjunto de datos manteniendo la consistencia en la obtención de estructuras, propiedades y datos biológicos.

Este procedimiento se aplicó de forma sistemática hasta consolidar el conjunto final de compuestos, procurando que cada uno contara con **estructura química validada (2D / 3D), representación SMILES, archivos descargables compatibles con software especializado y datos de actividad biológica respaldados por literatura científica**.

### 2.1.2. Criterios de inclusión

- Los compuestos fueron seleccionados de acuerdo a los siguientes criterios:
- Presencia del núcleo bencimidazol con estructuras químicas definidas, completas y verificables.
- Reporte de datos obtenidos de condiciones experimentales reproducibles y correctamente reportadas.
- Exclusión de compuestos duplicados.

Los criterios mencionados han sido alineados con las buenas prácticas **QSAR (OECD)**, donde la selección y limpieza del conjunto de datos obtenido se realizó de acuerdo a los principios

establecidos por la Organización para la Cooperación y el Desarrollo Económicos (OECD, 2007), que recomiendan:

- Uso de datos experimentales confiables y reproducibles.
- Definición clara de la Concentración Mínima Inhibitoria (MIC).

### 2.2. Clasificación metodológica de la actividad antimicrobiana (MIC)

Los valores experimentales de concentración mínima inhibitoria (MIC), correspondientes a los 100 compuestos analizados, constituyeron un componente fundamental en el proceso de curación y estructuración del conjunto de datos. Dichos valores fueron organizados y clasificados de manera cuantitativa con base en su actividad antimicrobiana.

En función de los rangos de actividad reportados, los compuestos se categorizaron en: activos ( $MIC \leq 10 \mu\text{g/mL}$ ), medianamente activos ( $MIC > 10-64 \mu\text{g/mL}$ ) y poco activos ( $MIC > 64 \mu\text{g/mL}$ ). Esta estratificación permitió optimizar el análisis de relaciones estructura-actividad (SAR), facilitar la selección de descriptores moleculares relevantes y favorecer la identificación de compuestos líderes.

La clasificación se fundamentó en criterios ampliamente utilizados en estudios microbiológicos y quimioinformáticos, donde la MIC ( $\mu\text{g/mL}$ ) por debajo de 10 se emplea como parámetro cuantitativo clave para la evaluación de la potencia antimicrobiana, tal como se resume en la **Tabla 2.2.1**.

Adicionalmente, todos los valores de MIC recopilados fueron estandarizados a una unidad única ( $\mu\text{g/mL}$ ), garantizando la comparabilidad entre los compuestos previo a su análisis computacional. Este proceso proporcionó un marco de referencia cuantitativo robusto, asegurando la coherencia entre la caracterización experimental y la posterior modelización mediante enfoques QSAR.

**Tabla 2.2.1.** Categorización metodológica de la actividad antimicrobiana según valores MIC

Categoría	Rango MIC ( $\mu\text{g/mL}$ )	Número de compuestos	Descripción metodológica
Activos	$\leq 10$	28	Alta potencia antimicrobiana; considerados candidatos iniciales

			para propuestas de compuestos líderes.
<b>Medianamente activos</b>	>10 – 64	22	Actividad moderada; útiles para análisis SAR y para evaluar modificaciones estructurales.
<b>Poco activos</b>	> 64	50	Actividad baja; incluidos como referencia para análisis comparativos en los modelos QSAR.

### 2.3. Optimización molecular

Previo al cálculo de los descriptores, se realizó una optimización geométrica de todas las estructuras con el software Avogadro (v1.2.0). Este es una herramienta quimiinformática frecuentemente utilizada en estudios QSAR debido a su capacidad para obtener conformaciones confiables mediante campos de fuerza clásicos y algoritmos robustos de minimización (Hanwell et al., 2012).

Se emplearon los siguientes parámetros:

- Campo de fuerza: MMFF94 (validado para moléculas orgánicas).
- Método de minimización: algoritmo de descenso más pronunciado (Steepest Descent).
- Criterio de convergencia:  $1 \times 10^{-7}$  kcal/mol.

### 2.4. Cálculo de descriptores moleculares

Los descriptores se calcularon utilizando el software PaDEL-Descriptor (versión 2.21), herramienta gratuita basada en Java ampliamente citada en la literatura QSAR y utilizada en estudios farmacéuticos y toxicológicos (Yap, 2011).

#### 2.4.1. Primera ejecución: cálculo completo

Se activaron la totalidad de casillas de PaDEL obteniendo más de 1400 descriptores por molécula. Esta ejecución inicial permitió observar la distribución de descriptores y detectar redundancias.

#### 2.4.2. Segunda ejecución: selección racional de casillas

Basándose en las directrices de la OECD y las recomendaciones de Tropsha (2010), se llevó a cabo un proceso sistemático de selección de descriptores moleculares. En primer lugar, se calcularon los descriptores utilizando herramientas computacionales en Python. Posteriormente, se realizó un filtrado inicial para eliminar variables constantes o con baja variabilidad.

A continuación, se aplicó un análisis de correlación entre descriptores, empleando el coeficiente de correlación de Pearson. Se estableció un umbral de  $|r| \geq 0.90$  para identificar pares de variables altamente correlacionadas. En estos casos, se eliminó uno de los descriptores correlacionados, priorizando la conservación de aquellos con mayor relevancia química e interpretabilidad.

Posteriormente, se seleccionaron los grupos de descriptores con mayor relevancia para el modelado cuantitativo, incluyendo:

- Descriptores topológicos 2D (Kier, Chi, Zagreb)
- Índices E-State
- Descriptores constitucionales (MW, número de heteroátomos, enlaces rotables)
- Autocorrelaciones (Moran, Geary, Broto-Moreau)
- BCUT
- Cargas PEOE
- Lipofilicidad (XlogP)
- Área superficial y volumen (TPSA, THSA, WPSA-1, Vi, Vs)
- Descriptores geométricos básicos

Por otro lado, se excluyeron del análisis:

- Huellas dactilares moleculares binarias completas (fingerprints)
- Subestructuras tipo fingerprint (FP)
- Descriptores 3D dependientes de múltiples conformaciones

Este enfoque prioriza la interpretabilidad del modelo y reduce la multicolinealidad entre variables, en concordancia con las prácticas recomendadas en estudios QSAR modernos (Cherkasov et al., 2014).

### *2.5. Preprocesamiento estadístico en Python*

El modelo QSAR fue construido en Python (versión 3.11.9.), empleando la biblioteca Scikit-learn (Pedregosa et al., 2011), Pandas y NumPy. En esta etapa se preparó la matriz de descriptores para su posterior análisis mediante los modelos de aprendizaje automático ejecutados en el estudio, incluyendo a ElasticNetCV, Random Forest y Support Vector Regression (SVR), todos estos implementados bajo un mismo esquema de preprocesamiento y validación.

Se aplicaron los siguientes procedimientos:

- Eliminación de columnas no numéricas
- Eliminación de descriptores con varianza cero
- Limpieza de valores nulo
- Reducción de multicolinealidad ( $r > 0.95$ )
- Estandarización por Z-score.

### *2.6. Selección de características*

Para identificar las variables con mayor impacto en actividad antimicrobiana se empleó diferentes métodos:

- SelectKBest (con `f_regression` y `mutual_info_regression`), (proceso de selección de descriptores que evalúa tanto relaciones lineales como no lineales)
- Importancia basada en Random Forest (Bosque Aleatorio),
- Coeficientes de ElasticNetCV

- Análisis de peso relativo mediante SVR.

Este enfoque multimodelo permite seleccionar descriptores más robustos y minimiza sesgos asociados al uso de un solo algoritmo.

### *2.7. Construcción de modelos QSAR*

Tres fueron los modelos entrenados y comparados:

- 1. ElasticNetCV**
- 2. Random Forest Regressor**
- 3. Support Vector Regression (SVR)**

Los tres modelos se implementaron con Scikit-learn, aplicando validación cruzada de 5 pliegues (C=5) siguiendo las recomendaciones de Tropsha (2010), garantizando reproducibilidad y evitar sobreajuste.

Se evaluaron tres métricas principales

- $R^2$  (coeficiente de determinación),
- RMSE (raíz del error cuadrático medio),
- MAE (Error absoluto medio).

Cada uno de los modelos fue ajustado mediante técnicas estándar de preprocesamiento y optimización de hiperparámetros.

### *2.8. Validación del modelo*

Esta validación incluyó:

- Análisis de residuos,
- Verificación de homocedasticidad, evaluación de patrones sistemáticos,

-Determinación del dominio de aplicabilidad mediante Williams Plot (Gráfico que evalúa el dominio de aplicabilidad del modelo).

Esta orientación sigue las recomendaciones internacionales para comprobar la solidez y confiabilidad de un modelo QSAR.

Como etapa final, orientada a evaluar la capacidad de generalización real de los modelos entrenados, se incorporó una etapa de validación externa utilizando nueve compuestos adicionales que no formaron parte de la matriz de entrenamiento ni de la construcción de los modelos (ver Anexo 2: Derivados de Bencimidazol – Validación externa).

### 2.9. Lineamientos racionales para la optimización estructural.

En la presente etapa se integraron los descriptores identificados como relevantes por los modelos QSAR lineales y no lineales, aquellos que mostraron mayor influencia en la variación de la MIC tales como nRotBt, nRotB, Kier1, PPSA-1, WPSA-1, THSA, MW y sumI, cuyo fin identificar las propiedades fisicoquímicas que deberían priorizarse en la modificación futura de la serie bencimidazólica.

## 3. Resultados

### 3.1. Preparación de la Base de datos de compuestos activos contra *Staphylococcus aureus*

Se construyó un base de datos (**Tabla 3.1.1.**) con 100 compuestos derivados del Bencimidazol con actividad documentada frente a *Staphylococcus aureus*. Para cada compuesto se recopiló su información estructural, fórmula molecular, nombre IUPAC y el valor experimental MIC ( $\mu\text{g/mL}$ ), permitiendo elaborar la matriz inicial para los análisis posteriores. Cuya información detallada para cada compuesto se ilustra en el **Anexo 1.**

**Tabla 3.1.1.** Número de compuesto y actividad biológica frente a *Staphylococcus aureus*.

N°	MIC $\mu\text{g/mL}$	N°	MIC $\mu\text{g/mL}$	N°	MIC $\mu\text{g/mL}$	N°	MIC $\mu\text{g/mL}$	N°	MIC $\mu\text{g/mL}$
1	65	21	0.194	41	0.328	61	64	81	256
2	65	22	0.102	42	0.328	62	4	82	256
3	65	23	0.087	43	0.328	63	64	83	6.25

4	140	24	0.092	44	0.328	64	32	84	3.12
5	140	25	0.175	45	1.764	65	64	85	3.12
6	250	26	0.183	46	1.764	66	4	86	3.12
7	65	27	0.159	47	8	67	64	87	6.25
8	65	28	625	48	32	68	4	88	3.12
9	140	29	625	49	32	69	16	89	3.12
10	32	30	200	50	32	70	16	90	3.12
11	128	31	400	51	32	71	64	91	3.12
12	64	32	400	52	32	72	64	92	25
13	64	33	400	53	16	73	32	93	25
14	64	34	12.5	54	64	74	16	94	25
15	128	35	12.5	55	64	75	64	95	6,25
16	128	36	12.5	56	16	76	16	96	12,5
17	32	37	0.104	57	64	77	32	97	12,5
18	64	38	0.076	58	8	78	256	98	25
19	128	39	0.186	59	8	79	256	99	50
20	0.206	40	0.35	60	64	80	256	100	50

### 3.2. Matriz final de descriptores moleculares y composición.

Una vez ejecutado el cálculo y la selección de los descriptores moleculares, se obtuvo una matriz final constituida por 636 descriptores numéricos para los 100 compuestos analizados. Estos descriptores corresponden al conjunto construido resultante después de la eliminación de variables redundantes, inconsistentes o altamente correlacionadas, tal cual se menciona en la metodología.

Esta matriz sirvió como base para la selección de características y construcción de los modelos QSAR presentados en secciones posteriores.

Los 636 descriptores se han agrupados de acuerdo con sus prefijos y se presenta en la **Tabla 3.2.1.**, está resume distribución y naturaleza de los descriptores utilizados en el modelado. Sin embargo, por su extensión, la matriz completa de valores se incluye en el Anexo 1.

**Tabla 3.2.1.** Clasificación de los 636 descriptores moleculares depurados según prefijo del descriptor e interpretación química.

<b>Prefijo / familia del descriptor</b>	<b>N° de descriptores</b>	<b>Interpretación química general (resumida)</b>
<b>S*</b> (SsX, SssX, SdX, etc.)	<b>119</b>	Capturan la distribución electrónica local por tipo de átomo (índices E-state).
<b>min*</b>	<b>119</b>	Representan los valores mínimos del estado electrónico para identificar centros menos activados.
<b>max*</b>	<b>119</b>	Representan los valores máximos del estado electrónico, asociados a centros más reactivos.
<b>n*/ns*/nss*</b>	<b>109</b>	Contadores estructurales que cuantifican átomos, fragmentos y entornos electrónicos específicos.
<b>L*, P*, E*, T*, A*, V*, K*, D*</b>	<b>91</b>	Descriptores 3D relacionados con forma, volumen, simetría y distribución espacial molecular (WHIM).
<b>PPSA, PNSA, WPSA, WNSA, TPSA, THSA, FPSA, FNSA, DPSA, RPSA, RPC*, RNC*, RHSA</b>	<b>29</b>	Describen superficie polar, carga superficial y accesibilidad molecular.
<b>nHB*, nwHB*</b>	<b>19</b>	Indican capacidad donora/aceptora de H-puentes y entornos de interacción polar.
<b>Fisicoquímicos globales (Kier, MW, XLogP, McGowan, RotBFrac, TopoPSA, etc.)</b>	<b>12</b>	Describen propiedades globales como tamaño, volumen, flexibilidad, polaridad y lipofilia.
<b>BCUT*</b>	<b>6</b>	Eigenvalores ponderados que reflejan masa, carga y distribución electrónica en el grafo molecular.
<b>sumI, meanI, h/g</b>	<b>6</b>	Índices globales que resumen conectividad, complejidad e información estructural.
<b>MAXDN/MAXDP/DELS</b>	<b>6</b>	Parámetros globales de donación/aceptación electrónica.
<b>LipoaffinityIndex</b>	<b>1</b>	Índice integrado de lipoafinidad y partición en medios lipofílicos.

### 3.3. Descriptores moleculares seleccionados para el modelo QSAR final

Partiendo de la matriz depurada de 636 descriptores, el modelo ElasticNetCV selecciono 10 descriptores con mayor distribución estadística a la variabilidad de MIC, frente a *S. aureus*. Estos representan el subconjunto óptimo para la construcción de la ecuación QSAR lineal interpretativa.

La **Tabla 3.3.1.** resume los descriptores seleccionados, su significado químico general y los coeficientes asignados por el modelo, permitiendo interpretar su impacto sobre la MIC en escala log10.

**Tabla 3.3.1.** Descriptores seleccionados por ElasticNetCV y su contribución al modelo QSAR final.

Descriptor	Interpretación química	Coefficiente ( $\beta_i$ )	Efecto sobre MIC
<b>nHBAcc_Lipinski</b>	Número de aceptores de enlace de hidrógeno según Lipinski	<b>+0.239590</b>	Aumenta MIC (↓ actividad)
<b>nRotBt</b>	Enlaces rotables totales	<b>+0.275819</b>	Aumenta MIC (↓ actividad)
<b>nRotB</b>	Número de enlaces rotables	<b>+0.158279</b>	Aumenta MIC (↓ actividad)
<b>THSA</b>	Área hidrofóbica total (Total Hydrophobic Surface Area)	<b>+0.325224</b>	Aumenta MIC (↓ actividad)
<b>Kier1</b>	Índice topológico-conformacional de primer orden	<b>0.000000</b>	Sin efecto significativo
<b>MW</b>	Peso molecular	<b>-0.009842</b>	Disminuye MIC (↑ actividad)
<b>sumI</b>	Índice global de conectividad/información	<b>-0.012644</b>	Disminuye MIC (↑ actividad)
<b>PPSA-1</b>	Superficie polar parcial positiva (nivel 1)	<b>-0.012309</b>	Disminuye MIC (↑ actividad)
<b>PPSA-2</b>	Superficie polar parcial positiva (nivel 2)	<b>-0.175672</b>	Disminuye MIC (↑ actividad)
<b>WPSA-1</b>	Superficie polar ponderada (Weighted PSA)	<b>-0.048229</b>	Disminuye MIC (↑ actividad)

El modelo QSAR lineal arrojó coeficientes ( $\beta_i$ ) positivos para nHBAcc\_Lipinski, nRotBt, nRotB y THSA, determinando que valores más altos en estas propiedades se asocian con incrementos en la MIC frente a *Staphylococcus aureus*. Por otro lado, los descriptores MW, sumI, PPSA-1, PPSA-2 y WPSA-1 presentaron coeficientes negativos, indicando una relación inversa con la MIC.

Estos diez descriptores conforman la base estructural y estadística de la ecuación QSAR lineal final, que se desarrolla en secciones posteriores.

**3.4. Ranking de importancia de descriptores en el modelo QSAR - 100 compuestos de entrenamiento.**

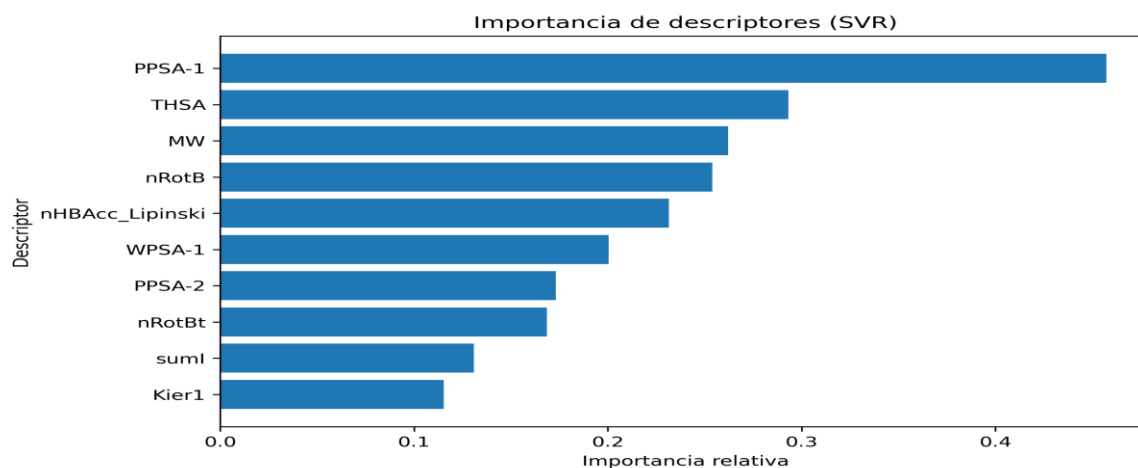
El modelo QSAR construido mediante Regresión por Vectores de Soporte (Support Vector Regression, SVR), permitió identificar el peso relativo que aporta cada uno de los diez descriptores dentro del desempeño predictivo del modelo. A partir del conjunto final de 100 compuestos, se calcularon valores de importancia promedio de cada variable, lo que permitió establecer un orden de relevancia según su influencia en la predicción de los valores de  $\log_{10}(\text{MIC})$ .

La Tabla 3.4.1. presenta el ranking de importancia de los 10 descriptores seleccionados, en orden de mayor a menor. Este análisis evidencia que no todos los descriptores contribuyen en la misma magnitud, algunos identifican un impacto considerable en la capacidad del modelo para explicar variabilidad en actividad antimicrobiana, mientras que otros presentan influencia más limitada. Estas diferencias reflejan la contribución específica de cada propiedad molecular al comportamiento global del modelo SVR.

Tabla 3.4.1. Ranking de importancia de descriptores seleccionados en el modelo QSAR (SVR).

<b>Posición</b>	<b>Descriptor</b>	<b>Importancia media</b>
<b>1</b>	<b>PPSA-1</b>	0.457318
<b>2</b>	<b>THSA</b>	0.293126
<b>3</b>	<b>MW</b>	0.262077
<b>4</b>	<b>nRotB</b>	0.253939
<b>5</b>	<b>nHBAcc_Lipinski</b>	0.231480
<b>6</b>	<b>WPSA-1</b>	0.200514
<b>7</b>	<b>PPSA-2</b>	0.173197
<b>8</b>	<b>nRotBt</b>	0.168574
<b>9</b>	<b>sumI</b>	0.130994
<b>10</b>	<b>Kier1</b>	0.115408

En la Figura 3.4.1. se aprecia que los descriptores asociados a la superficie polar parcial y el área hidrofílica (PPSA-1, THSA) expresan la mayor importancia relativa, seguidos por propiedades topológicas y de flexibilidad estructural (MW, nRotB).



**Figura 3.4.1.** Importancia relativa de los descriptores moleculares seleccionados en el modelo QSAR basado en *Support Vector Regression (SVR)*.

**Fuente:** Elaboración propia mediante Python (*Scikit-learn, Matplotlib y Pandas*).

### 3.5. Modelado QSAR y validación estadística (ecuación QSAR)

Se desarrollaron tres algoritmos de aprendizaje supervisado, cuyo fin fue la construcción y comparación de modelos predictivos, tales como: ElasticNet, Random Forest y Support Vector Regression (SVR) con núcleo RBF).

El desempeño de cada uno fue evaluado a través de validación cruzada de cinco pliegues ( $CV = 5$ ), calculando indicadores estadísticos  $R^2$ , RMSE y MAE, tanto para el conjunto de entrenamiento como en validación.

La Tabla 3.5.1. muestra los valores de desempeño obtenidos para los modelos ElasticNet, Random Forest y Support Vector Regression (SVR). Los resultados identifican diferencias en la capacidad de ajuste y predicción entre los modelos evaluados. El modelo Random Forest alcanzó el mayor  $R^2$  en el conjunto de entrenamiento, mientras que el modelo SVR mostró el mejor desempeño durante la validación interna, sin embargo, la capacidad predictiva real de cada modelo fue determinada posteriormente mediante una validación externa con compuestos adicionales.

**Tabla 3.5.1.** Desempeño comparativo de los modelos QSAR evaluados.

Modelo	R <sup>2</sup> entrenamiento	RMSE	MAE	R <sup>2</sup> CV (media)	R <sup>2</sup> CV (desv. est.)
ElasticNetCV	0,395	0,820	0,647	0,027	0,465
Random Forest	0,919	0,300	0,219	0,343	0,223
SVR (RBF)	0,761	0,516	0,335	<b>0,380</b>	<b>0,193</b>

### 3.6. Ecuación QSAR, construida (ElasticNetCV)

El modelo lineal construido permitió derivar una ecuación QSAR interpretativa a partir de los 10 descriptores seleccionados. La ecuación describe la relación cuantitativa entre las propiedades estructurales de los compuestos y valores experimentales de log<sub>10</sub>(MIC) frente a *S. aureus*. En las líneas siguientes se ilustra la expresión final del modelo.

La ecuación QSAR final, expresada en escala logarítmica:

$$\begin{aligned} \log_{10}(\text{MIC}) = & 1.2276 - (0.0126 * \text{sumI}) + (0.2396 * n\text{HB}Acc_{\text{Lipinski}}) + (0.1583 * n\text{RotB}) \\ & + (0.2758 * n\text{RotBt}) - (0.0098 * \text{MW}) - (0.0123 * \text{PPSA} - 1) - (0.1757 * \text{PPSA} - 2) \\ & - (0.0482 * \text{WPSA} - 1) + (0.3252 * \text{THSA}) \end{aligned}$$

(El descriptor **Kier1** presentó un coeficiente igual a cero y no contribuyó a la expresión lineal final.)

En unidades experimentales, se expresa como:

$$\text{MIC} = 10^{(b_0 + \sum(\beta_i \cdot X_i))}$$

Donde:

**MIC** corresponde a la concentración mínima inhibitoria expresada en µg/mL

**b<sub>0</sub>** = 1.2276 representa el intercepto del modelo

**β<sub>i</sub>** = son los coeficientes asociados a cada descriptor X<sub>i</sub>

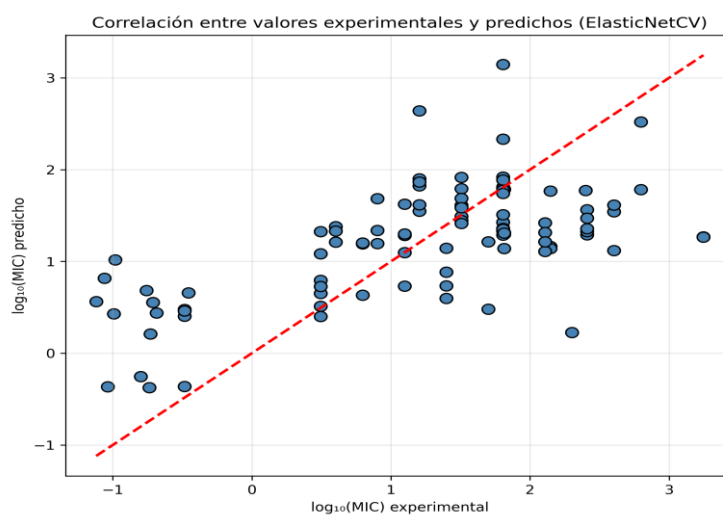
**X<sub>i</sub>** = valores de los descriptores moleculares en su escala original

A continuación, se presentan los parámetros estadísticos obtenidos para el modelo *ElasticNetCV*, estos describen el comportamiento del ajuste y la precisión predictiva alcanzada dentro del conjunto de los 100 compuestos analizados. En la Tabla 3.6.1. se resume el desempeño general del modelo lineal interpretativo.

**Tabla 3.6.1.** Métricas estadísticas del modelo QSAR lineal (*ElasticNetCV*) basado en 100 compuestos.

Parámetro	Descripción	Valor
<b>R<sup>2</sup></b>	Coficiente de determinación (entrenamiento)	0.395
<b>RMSE</b>	Raíz del error cuadrático medio ( $\log_{10}(\text{MIC})$ )	0.820
<b>MAE</b>	Error absoluto medio ( $\log_{10}(\text{MIC})$ )	0.647
<b><math>\alpha</math></b>	Parámetro de regularización óptimo	0.001
<b>L1_ratio</b>	Proporción L1–L2 en la penalización	0.10

La **Figura 3.6.1.** ilustra la relación entre los valores experimentales y los valores predichos de  $\log_{10}(\text{MIC})$  generados a partir del modelo QSAR penalizado (*ElasticNetCV*). La dispersión que se observa evidencia un ajuste moderado del modelo, con tendencia general alineada hacia la línea de identidad, indicando concordancia parcial entre los valores observados y los valores estimados por la ecuación QSAR final.



**Figura 3.6.1.** Correlación entre los valores experimentales y predichos de  $\log_{10}(\text{MIC})$  obtenidos mediante el modelo QSAR lineal penalizado (*ElasticNetCV*).

**Fuente:** Elaboración propia mediante Python (*Scikit-learn*, *Pandas* y *Matplotlib*)

### 3.7. Rol interpretativo del modelo lineal ElasticNetCV

Este modelo permitió obtener la ecuación QSAR lineal interpretativa basada en descriptores seleccionados, alcanzando un coeficiente de determinación de  $R^2 = 0.395$ , con un RMSE de 0.820 y un MAE de 0.647 en escala  $\log_{10}(\text{MIC})$ , indicando un ajuste moderado, adecuado para interpretación molecular pese a su limitación en la capacidad predictiva.

El modelo seleccionó automáticamente los parámetros de regularización  $\alpha = 0.001$  y  $L1\_ratio = 0.10$ , reflejando el uso de una penalización mixta que combina los componentes L1 y L2. e). Este ajuste contribuye a mejorar la estabilidad numérica del modelo y mitigar el sobreajuste, estimando de manera individualizada la contribución de cada descriptor con respecto a la variación de la actividad antimicrobiana.

Pese que los modelos no lineales como Random Forest y Support Vector Regression alcanzaron un desempeño predictivo superior, ElasticNetCV contribuye un elemento clave para el presente estudio: permitiendo obtener una ecuación lineal explícita que facilita la interpretación fisicoquímica de los factores que influyen en la actividad antimicrobiana frente a *S. aureus*. Esto permite entender cómo los cambios en las propiedades estructurales del núcleo Bencimidazol se traducen cuantitativamente en la respuesta biológica, permitiendo complementar así la mayor capacidad predictiva de los modelos más complejos.

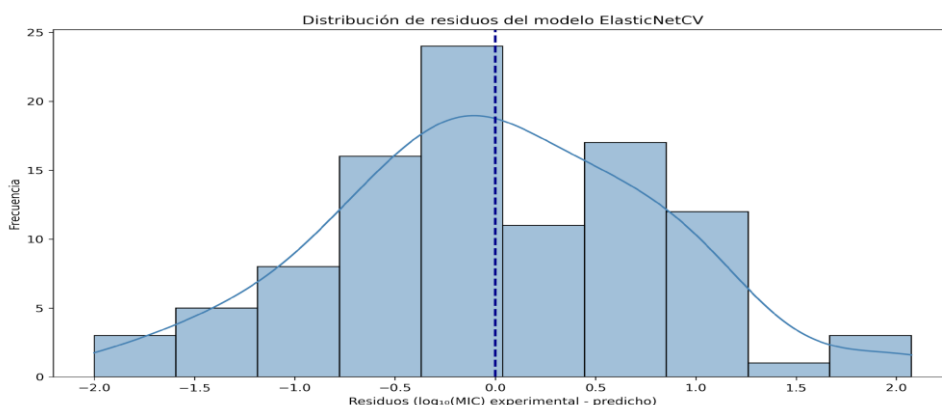
### 3.8. Validación estadística del modelo QSAR

En esta sección se comprobó la robustez, fiabilidad y coherencia interna del modelo QSAR lineal obtenido mediante ElasticNetCV, a través del análisis de residuos y validación cruzada.

#### 3.8.1. Análisis de residuos

El comportamiento de los residuos permite comprobar si se cumplen los supuestos estadísticos del modelo lineal, en especial la tendencia a la linealidad y la ausencia de sesgos sistemáticos.

En la Figura 3.8.1. Se puede observar que los residuos se mantienen principalmente alrededor de cero, sin agrupaciones inusuales que indiquen desviaciones importantes con respecto a estos supuestos. Pese a que la distribución presenta variabilidad amplia, esta no evidencia patrones que comprometan la validez interpretativa del modelo lineal dentro del conjunto analizado.



**Figura 3.8.1.** Distribución de los residuos del modelo QSAR lineal (ElasticNetCV)

*Fuente:* Elaboración propia con Python (Scikit-learn, Pandas, Matplotlib y Seaborn)

### 3.9. Validación cruzada y análisis de robustez del modelo SVR

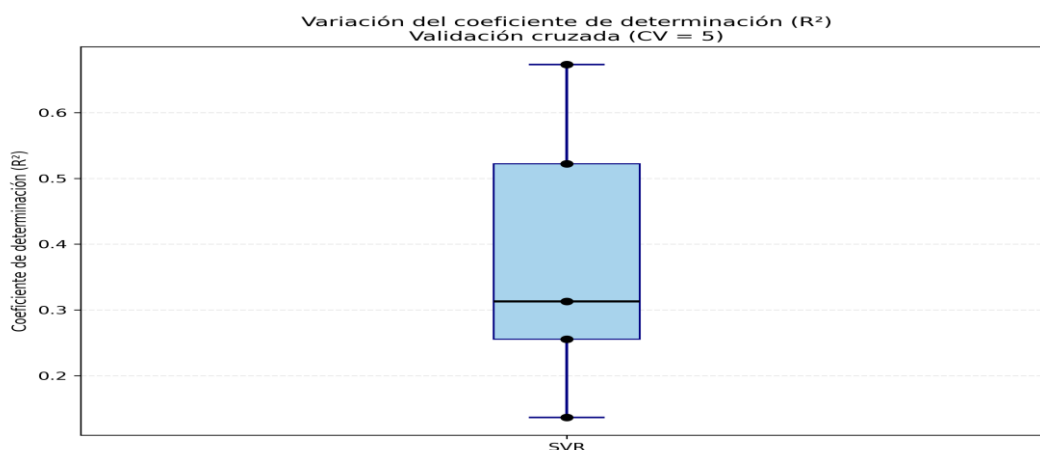
Para el modelo SVR no lineal, la validación cruzada de cinco pliegues (CV=5) generó valores  $R^2$  en un intervalo entre 0.137 y 0.673, con:

**-Media = 0.380**

**-Desviación estándar = 0.193**

Estos resultados indican una capacidad predictiva moderada, con variaciones esperadas entre pliegues debido a la naturaleza no lineal del modelo y además la diversidad estructural del conjunto evaluado. Todos los valores de  $R^2$  fueron positivos, lo que refleja que el modelo logra reproducir patrones descriptor-actividad en cada partición.

La figura 3.9.1. Representa el diagrama de caja correspondiente a los valores de validación cruzada, aquí se puede observar una dispersión moderada y ausencia de valores atípicos extremos. La presente distribución confirma la estabilidad adecuada del modelo SVR y respalda una consideración con el enfoque predictivo más consistente entre los modelos evaluados.



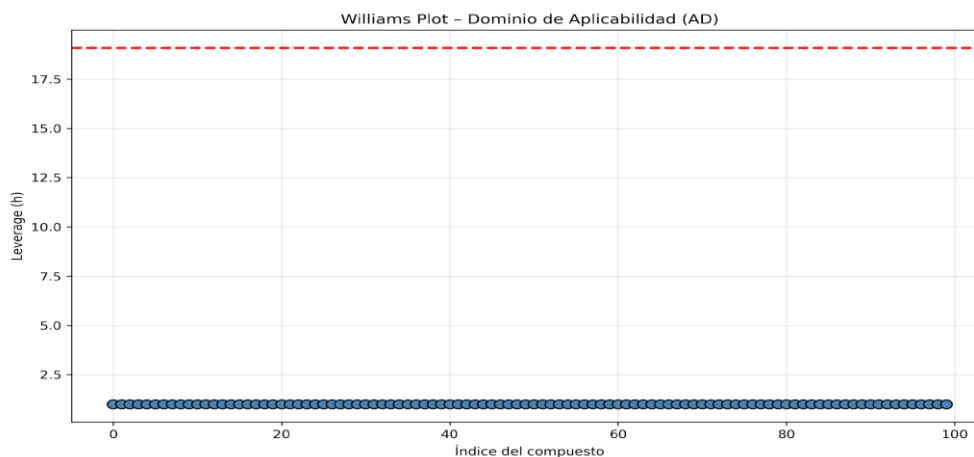
**Figura 3.9.1.** Variación del coeficiente de determinación ( $R^2$ ) en la validación cruzada ( $CV = 5$ ).

**Fuente:** Elaboración propia con Python (Scikit-learn y Matplotlib), a partir del modelo SVR validado con  $CV = 5$ .

### 3.10. Dominio de aplicabilidad (AD)

El dominio de aplicabilidad del modelo QSAR se verificó mediante el cálculo del modelo QSAR se evaluó a través del cálculo de Leverage ( $h$ ) para los 100 compuestos de la matriz final. El valor crítico obtenido fue  $h^* = 19.11$ , mientras que el leverage máximo observado fue 1.00, indicando que ningún compuesto superó el umbral.

En la Figura 3.10.1. Se aprecia como todos los puntos permanecen dentro del límite permitido, lo que permite confirmar que el modelo no realiza extrapolaciones y que opera dentro de un espacio estable del descriptor y representado por los datos de entrenamiento.



**Figura 3.10.1.** Williams Plot del dominio de aplicabilidad (AD) para los 100 compuestos, mostrando los valores de leverage ( $h$ ) y el lumbral crítico  $h^* = 19.11$

**Fuente:** elaboración propia Python.

### 3.11. Análisis Relación Estructura-Actividad (SAR)

El análisis de relaciones estructura–actividad (SAR) evidenció que los descriptores seleccionados presentan una correlación positiva global con los valores de  $\log_{10}(\text{MIC})$ , lo que indica que el incremento en dichos descriptores se asocia con una disminución de la actividad antimicrobiana. Entre los descriptores de mayor influencia se identificaron nRotBt, Kier1 y nRotB, lo que sugiere que una mayor flexibilidad conformacional y determinadas características topológicas contribuyen a la reducción de la potencia biológica.

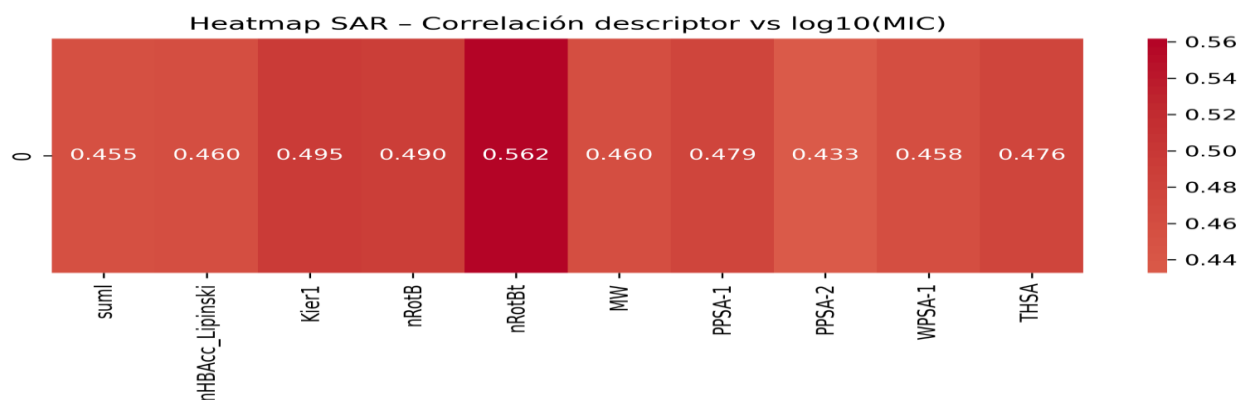
En particular, el aumento en el número de enlaces rotables (nRotB, nRotBt) refleja una mayor libertad conformacional, lo cual puede afectar negativamente la interacción específica del compuesto con su diana biológica. De manera consistente, el índice topológico de Kier (Kier1) sugiere que ciertos patrones estructurales influyen en la afinidad molecular.

Por otra parte, se observó que el incremento en la superficie polar y en las propiedades hidrofílicas se asocia con valores más elevados de MIC, lo que podría limitar la permeabilidad a través de membranas biológicas, especialmente en bacterias Gram positivas como *Staphylococcus aureus*, afectando así la eficacia antimicrobiana.

**Tabla 3.11.1.** Correlación entre los descriptores moleculares seleccionados y los valores de  $\log_{10}(\text{MIC})$  en la serie de derivados del bencimidazol.

Descriptor	Correlación vs $\log_{10}(\text{MIC})$	Interpretación química	Implicación en la actividad (SAR)
sumI	0.455	Índice de conectividad / complejidad estructural	Mayor complejidad → mayor MIC (↓ actividad)
nHBAcc_Lipinski	0.460	Aceptores de enlace de hidrógeno	Más aceptores → menor potencia antimicrobiana
Kier1	0.495	Índice topológico–conformacional	Mayor flexibilidad → mayor MIC
nRotB	0.490	Enlaces rotables	Aumento de movilidad conformacional → menor actividad
nRotBt	<b>0.562</b>	Enlaces rotables totales	Descriptor más influyente; flexibilidad elevada perjudica la actividad
MW	0.460	Masa molecular	Moléculas más pesadas → menor permeabilidad → mayor MIC

<b>PPSA-1</b>	0.479	Superficie polar parcial positiva	Aumento de polaridad → mayor MIC
<b>PPSA-2</b>	0.433	Superficie polar parcial positiva (complementaria)	Reducción de actividad al aumentar la polaridad
<b>WPSA-1</b>	0.458	Superficie polar ponderada	Superficie polar extensa → disminución de la actividad
<b>THSA</b>	0.476	Área hidrofílica total	Mayor hidrofilia → menor eficiencia inhibitoria



**Figura 3.11.1.** Heatmap (mapa de calor) de correlación entre los descriptores moleculares seleccionados y los valores de  $\log_{10}(\text{MIC})$ .

Fuente: Elaboración propia mediante Python (Pandas, Seaborn y Matplotlib)

### 3.12. Lineamientos racionales para la optimización estructural de derivados de bencimidazol.

La identificación de los descriptores con mayor influencia sobre la actividad antimicrobiana, permitió determinar los lineamientos estructurales que orientan el diseño racional de nuevos derivados del bencimidazol. Con base en los patrones del descriptor observados en los análisis QSAR y SAR, derivados tanto de los modelos lineales como de los no lineales fue posible organizar la información de manera sistemática y definir qué propiedades deben priorizarse en una optimización estructural posterior.

La **Tabla 3.12.1.** resume las propiedades críticas de los descriptores y sus implicaciones con respecto a la actividad biológica, en contraste la **Figura 3.12.1.** sintetiza la relación entre los factores que disminuyen la actividad y estrategias sugeridas para su mejora. Por su parte, la **Figura 3.12.2.** integra el proceso lógico desde los resultados obtenidos hasta la propuesta de los criterios racionales de optimización, ofreciendo una visión clara y estructurada del enfoque seguido en esta etapa.

**Tabla 3.12.1.** Lineamientos de los descriptores para el diseño racional de nuevos derivados del bencimidazol.

<b>Categoría del descriptor</b>	<b>Descriptor</b>	<b>Correlación vs <math>\log_{10}(\text{MIC})</math></b>	<b>Implicación química real</b>	<b>Recomendación de diseño racional</b>
<b>Flexibilidad conformacional</b>	<b>nRotBt</b>	<b>0.562</b>	Flexibilidad total elevada reduce actividad	Reducir enlaces rotables; usar sustituyentes rígidos o cíclicos
	<b>nRotB</b>	0.490	Mayor movilidad conformacional	Limitar cadenas alquílicas largas; preferir anillos
<b>Topología / complejidad</b>	<b>Kier1</b>	0.495	Aumento de complejidad topológica → mayor MIC	Evitar ramificaciones excesivas; promover estructuras compactas
<b>Superficie polar parcial</b>	<b>PPSA-1</b>	0.479	Polaridad positiva elevada reduce actividad	Modulación de heteroátomos; limitar grupos polarizantes
	<b>PPSA-2</b>	0.433	Complemento de PPSA-1	Evitar acumulación de cargas parciales positivas
	<b>WPSA-1</b>	0.458	Superficie polar ponderada amplia	Controlar tamaño de grupos sustituyentes polares
<b>Hidrofilia total</b>	<b>THSA</b>	0.476	Área hidrofílica grande disminuye permeabilidad	Reducir grupos muy polares (OH, NH, carbonilos múltiples)
<b>Masa y tamaño molecular</b>	<b>MW</b>	0.460	Peso molecular alto → menor penetración	Mantener MW en rango moderado (no añadir sustituyentes pesados)
<b>Densidad electrónica</b>	<b>sumI</b>	0.455	Complejidad electrónica excesiva penaliza actividad	Equilibrar grupos donadores/aceptores; evitar exceso de heteroátomos

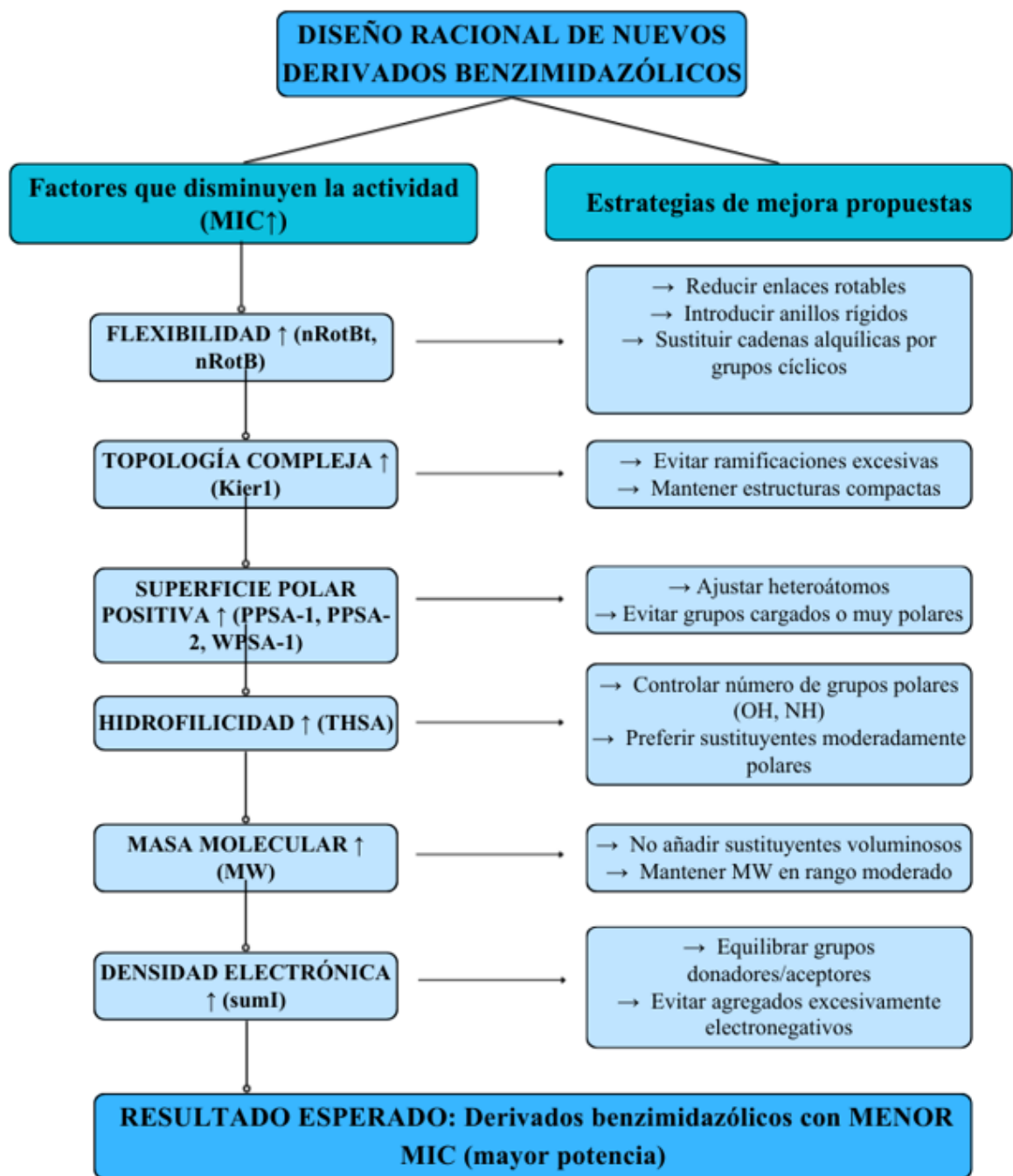
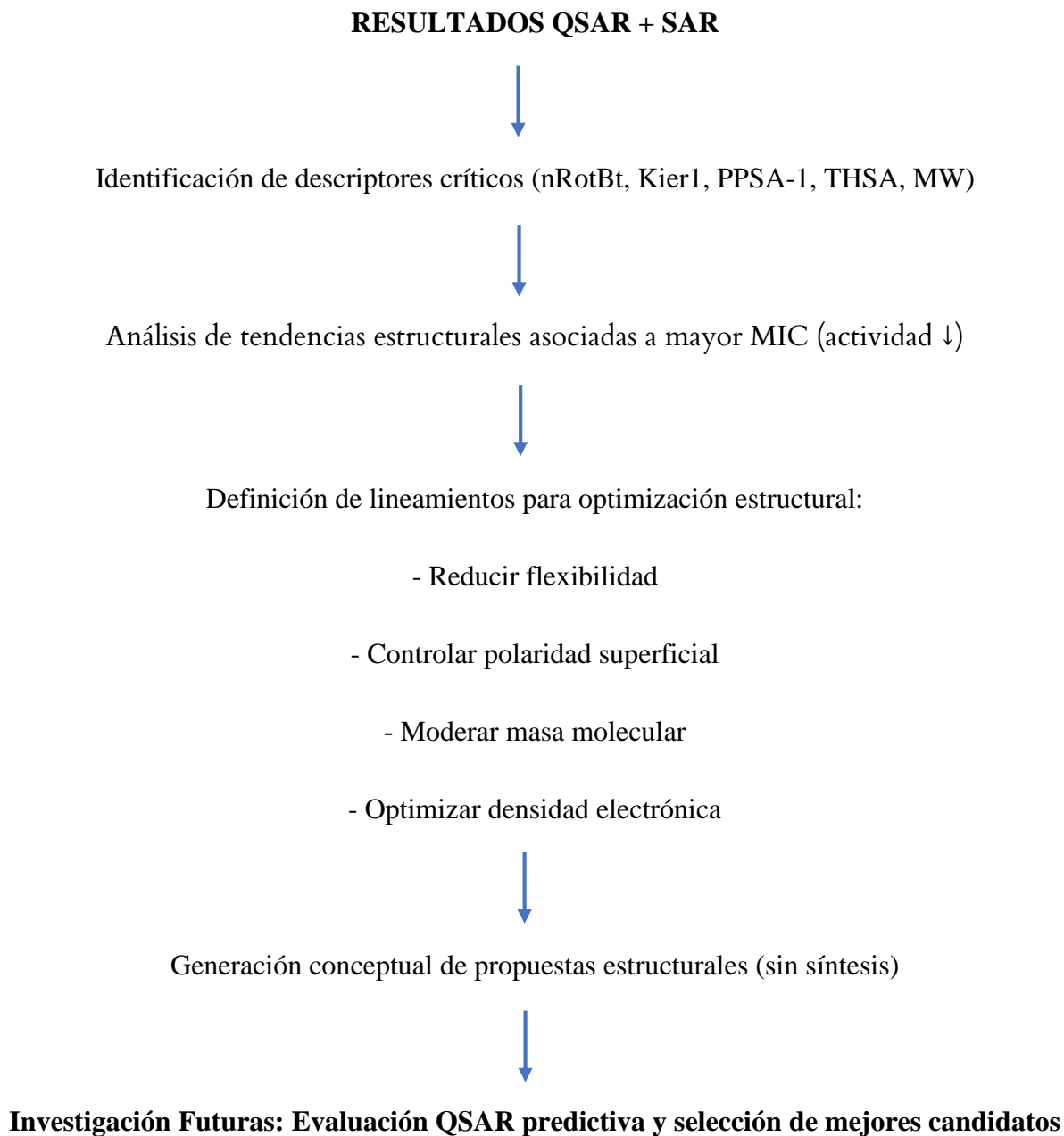


Figura 3.12.1. Mapa conceptual del diseño racional basado en descriptores clave.

Fuente: Elaboración propia mediante CANVA



**Figura 3.12.2.** Flujograma del procedimiento conceptual del diseño racional

*Fuente: Elaboración propia*

### **3.13. Integración final de resultados**

Los resultados obtenidos permitieron estructurar una visión integral y coherente del comportamiento descriptor-actividad dentro del conjunto de derivados del Bencimidazol. El

proceso de modelado mostró que, aunque el algoritmo SVR (núcleo RBF) presentó un desempeño interno adecuado durante la validación cruzada, el modelo Random Forest demostró la mejor capacidad de generalización al ser evaluado con compuestos externos, consolidándose como el modelo predictivo más robusto del estudio.

En contexto el modelo lineal ElasticNetCV contribuyó estructurar una ecuación explícita que facilitó interpretar directamente la contribución individual de los descriptores con respecto a la variación de la MIC, aportando un valor fundamental para el entendimiento fisicoquímico de la serie, aun cuando su capacidad predictiva externa fue limitada. La evaluación estadística, complementada con el análisis de residuos y el dominio de aplicabilidad (AD), que al ser ejecutados demostraron que las predicciones se mantuvieron dentro de los límites del descriptor representados en el conjunto de entrenamiento, garantizando consistencia y ausencia de extrapolaciones.

La integración del análisis SAR fortificó estos hallazgos, evidenciando que la flexibilidad conformacional, la superficie polar positiva, el área hidrofílica total y el incremento de la masa molecular se relacionan sistemáticamente con valores más altos de MIC. Mediante heatmap (mapa de calor) y la síntesis organizada en la **Tabla 3.12.1** permitieron determinar con claridad los factores estructurales que disminuyen la actividad antimicrobiana frente al modelo estudiado de *S. aureus*.

Finalmente, los criterios de diseño racional derivados de los resultados QSAR aportan una guía clara para la generación futura de nuevos análogos del Bencimidazol, enfocados a reducir la flexibilidad, controlar polaridad superficial y optimizar el equilibrio electrónico sin incrementar excesivamente la masa molecular (MW). En conjunto, esta integración establece las bases para etapas posteriores de exploración y optimización computacional.

### **3.14. Validación externa del modelo QSAR**

La validación externa constituye un paso fundamental para determinar la capacidad real de generalización de los modelos QSAR construidos. Para esta evaluación se utilizaron nueve moléculas adicionales no incluidas en el entrenamiento, cada una con valores experimentales de MIC previamente reportados. A todas se les calcularon los diez descriptores oficiales seleccionados para la construcción de los modelos, empleando el mismo proceso de estandarización aplicado sobre el conjunto de entrenamiento.

Los valores predichos por los tres modelos desarrollados (ElasticNetCV, Random Forest y SVR-RBF) se compararon contra los valores experimentales, permitiendo estimar su error absoluto y su capacidad para reproducir la actividad antimicrobiana real de los compuestos externos. Los resultados detallados de esta comparación, incluyendo los valores experimentales, los valores predichos y los errores asociados para cada modelo, se presentan en la **Tabla 3.14**

**Tabla 3.14.** Validación externa del modelo QSAR

Nombre	MIC_exp_ug_m	MIC_ElasticNet_ug_m	MIC_RF_ug_m	MIC_SVR_ug_m	Error_abs_E	Error_abs_R	Error_abs_SV
	L	L	L	L	N	F	R
Compuesto_1	32	7,89	1,59	2,47	0,61	1,30	1,11
Compuesto_2	2	41,12	40,68	68,85	1,31	1,31	1,54
Compuesto_3	4	8,81	0,39	1,21	0,34	1,01	0,52
<b>Compuesto_4</b>	<b>8</b>	34,62	<b>8,18</b>	68,83	0,64	<b>0,01</b>	0,93
<b>Compuesto_5</b>	<b>16</b>	26,52	<b>14,89</b>	84,80	0,22	<b>0,03</b>	0,72
Compuesto_6	8	4,72	1,23	2,02	0,23	0,81	0,60
Compuesto_7	32	16,19	7,39	32,04	0,30	0,64	0,00
<b>Compuesto_8</b>	<b>4</b>	3,78	<b>5,41</b>	2,71	0,02	<b>0,13</b>	0,17
<b>Compuesto_9</b>	<b>8</b>	11,14	<b>4,13</b>	9,07	0,14	<b>0,29</b>	0,05

En los resultados anteriores se evidencian diferencias cuantificables en el desempeño de los 3 modelos. **ElasticNetCV** presentó errores absolutos en un rango de **0.02 a 1.31** unidades  $\log_{10}(\text{MIC})$ , con desviaciones elevadas en compuestos de baja MIC, como el compuesto **2**. **SVR-RBF** mostró un rango de error entre **0.00 y 1.54**, denotando sobreestimaciones en varios compuestos tales como **2 y 4**. En contraste, **Random Forest** presentó los menores errores absolutos en la mayoría de los casos, destacándose en los compuestos **4 (0.01), 5 (0.03), 8 (0.13) y 9 (0.29)**, indicando mayor capacidad para modelar relaciones no lineales entre descriptores moleculares y actividad biológica. Cuyos resultados posicionan a Random Forest como el modelo con mejor desempeño predictivo en la validación externa del presente estudio QSAR.

#### 4. Discusión

Los resultados obtenidos en este estudio demuestran que el modelado QSAR, apoyado en algoritmos de aprendizaje automático, es una estrategia válida para analizar la relación estructura–actividad de derivados del bencimidazol frente a *Staphylococcus aureus*. La construcción y comparación de modelos lineales y no lineales permitió abordar el problema desde dos perspectivas complementarias: tanto la interpretación fisicoquímica de los descriptores como la capacidad predictiva real del sistema, concordando con las recomendaciones metodológicas internacionales para estudios QSAR (OECD, 2007; Tropsha, 2010; Cherkasov et al., 2014).

Desde el punto de vista predictivo, Random Forest destacó en la validación externa. Este algoritmo logró reproducir con mayor precisión los valores experimentales de MIC en varios compuestos, con errores mínimos en los compuestos 4, 5, 8 y 9. Este comportamiento no es casual; los modelos basados en ensamblado, como Random Forest, están diseñados para capturar patrones complejos y no lineales que suelen caracterizar los sistemas biológicos. En el contexto QSAR, donde múltiples factores estructurales interactúan de manera simultánea, esta capacidad resulta ventajosa, tal como ha sido documentado en la literatura (Cherkasov et al., 2014; Muratov et al., 2020; Yang et al., 2019).

Un aspecto relevante para interpretar estos resultados es la estructura de la matriz de entrenamiento. En este estudio los compuestos se distribuyeron según su nivel de actividad biológica en función del valor de MIC, incluyendo compuestos activos ( $\text{MIC} \leq 10 \mu\text{g/mL}$ ), medianamente activos ( $\text{MIC}$  entre 10 y  $64 \mu\text{g/mL}$ ) y de baja actividad ( $\text{MIC} > 64 \mu\text{g/mL}$ ).

Aproximadamente la mitad de los compuestos correspondieron a moléculas de baja actividad, lo que introduce un desequilibrio en la distribución de los datos y puede dificultar la modelación de compuestos más activos.

A pesar de ello, Random Forest mostró un comportamiento consistente en el rango de MIC entre 4 y  $16 \mu\text{g/mL}$ , donde las predicciones fueron más cercanas a los valores experimentales. Esto sugiere que el modelo logra capturar con mayor eficiencia las relaciones descriptor–actividad en regiones del espacio químico con mejor representatividad estructural, aspecto clave para la

aplicabilidad de modelos QSAR en la predicción de nuevos compuestos (Tropsha, 2010). (Tropsha, 2010).

En contraste, ElasticNetCV y SVR-RBF presentaron un comportamiento menos consistente en validación externa. ElasticNetCV, al basarse en una aproximación lineal penalizada, mostró limitaciones para describir relaciones no lineales complejas, lo que se reflejó en errores más elevados en algunos compuestos con baja MIC. Sin embargo, su valor dentro del estudio no radica únicamente en la predicción, sino en su capacidad para generar una ecuación QSAR explícita, permitiendo interpretar directamente cómo cada descriptor contribuye a la actividad antimicrobiana, lo que sigue siendo un aspecto clave en la interpretación de la relación estructura-actividad (Roy et al., 2015).

Por su parte, el modelo SVR-RBF, aunque presentó un mejor desempeño durante la validación interna, mostró una menor estabilidad con compuestos externos. En varios casos se observó tendencia a sobreestimar valores de MIC, lo que indica que el modelo puede ser sensible a la distribución de los datos y a la presencia de regiones del espacio químico poco representadas en el entrenamiento. Este comportamiento concuerda con métodos basados en kernels (transformaciones), cuya capacidad de generalización depende estrictamente de la representatividad del conjunto de datos (Cherkasov et al., 2014).

Estos hallazgos refuerzan una idea clave en modelado QSAR: un buen ajuste interno no garantiza un buen desempeño predictivo en datos externos. La validación externa realizada en este estudio permitió evidenciar esta diferencia, posicionando al modelo Random Forest como el modelo más robusto para la predicción de nuevos compuestos dentro del dominio de aplicabilidad.

En consecuencia, Random Forest no solo mostró una mayor estabilidad en sus predicciones, sino también una mejor capacidad para capturar la complejidad inherente a las relaciones estructura-actividad en los derivados del bencimidazol. Convirtiéndolo en una herramienta particularmente útil para futuras aplicaciones en diseño racional de compuestos, en etapas preliminares donde la capacidad predictiva resulta crítica para la selección de candidatos con potencial actividad antimicrobiana concordando con enfoques actuales de descubrimiento de fármacos asistido por inteligencia artificial (Yang et al., 2019).

Más allá del rendimiento predictivo de los modelos QSAR, es importante destacar estos resultados en el contexto biológico de *Staphylococcus aureus*, patógeno reconocido por su gran capacidad de adaptación, elevada virulencia y, sobre todo, por su resistencia a múltiples

terapéuticas antimicrobianas. Uno de los mecanismos más relevantes es la adquisición del gen *mecA*, que codifica la proteína PBP2a, variante de las proteínas de unión a penicilina que reduce la eficacia de los antibióticos  $\beta$ -lactámicos, dando lugar a cepas resistentes como MRSA (Chambers & DeLeo, 2009).

En este escenario, radica la importancia identificar qué características estructurales de los derivados del bencimidazol están relacionadas con su actividad antimicrobiana. Estos compuestos resultan interesantes ya que pueden actuar sobre blancos bacterianos alternativos, distintos a los mecanismos tradicionales de los  $\beta$ -lactámicos. Entre estos destacan enzimas clave como la DNA girasa y la topoisomerasa IV, implicadas en la replicación del ADN, así como en rutas esenciales como la biosíntesis de ácidos grasos (FAS-II) y la síntesis de la pared celular (Bush & Bradford, 2016; Foster, 2017).

Los resultados obtenidos sugieren que la lipofilicidad, la superficie polar y la flexibilidad molecular no son simplemente variables estadísticas, sino factores directamente relacionados con la capacidad de los compuestos para atravesar la pared celular bacteriana y alcanzar sus blancos intracelulares. En *S. aureus*, la pared rica en peptidoglicano actúa como barrera importante, por lo que el equilibrio entre características hidrofílicas y la distribución de carga superficial (reflejadas en descriptores como PPSA-1, WPSA-1 y THSA), resulta determinante para la penetración del compuesto, y por ende para la actividad antimicrobiana (Foster, 2017). En este contexto, nHBAcc\_Lipinski se relaciona con la capacidad de formar enlaces de hidrógeno, fundamentales en la interacción con blancos enzimáticos.

A partir de estas relaciones descriptor – actividad identificadas, es posible establecer criterios orientados al diseño de nuevos compuestos.

Con base en los patrones de descriptores moleculares identificados, se establecieron lineamientos para orientar la optimización de futuros derivados, tales como la reducción de la flexibilidad conformacional, el control de la superficie polar parcial positiva, la moderación del área hidrofílica total, la limitación del incremento en la masa molecular (MW) y la optimización del equilibrio electrónico.

En este sentido, los resultados obtenidos constituyen un marco conceptual para el diseño y evaluación de nuevos compuestos, el cual podrá desarrollarse en etapas posteriores mediante el uso de herramientas informáticas como Avogadro y PaDEL para la construcción y cálculo de

descriptores, RDKit para la validación estructural y Python (Scikit-learn) para la ejecución y evaluación de modelos QSAR.

A esto se suma otro desafío relevante: la capacidad de *S. aureus* para construir biopelículas y expresar múltiples factores de virulencia, como toxinas, adhesinas (MSCRAMMs) y sistemas regulatorios como agr, que favorecen la persistencia de la infección y dificultan la acción de los antimicrobianos (Otto, 2014). En este contexto, los compuestos que logren penetrar estas estructuras podrían tener una ventaja terapéutica significativa.

Desde el punto de vista de la resistencia, además del gen *mecA*, este microorganismo tiene otros mecanismos que perjudican el tratamiento, como las bombas de eflujo (por ejemplo, NorA), la modificación de dianas y la producción de enzimas inactivadoras. Todo esto refuerza la necesidad de diseñar compuestos que no solo sean activos, sino que también puedan evitar estos sistemas de defensa bacterianos (Chambers & DeLeo, 2009; Otto, 2014).

En este contexto, el mejor desempeño del modelo Random Forest no solo refleja su capacidad para capturar relaciones no lineales, sino también su utilidad para detectar patrones complejos entre estructura química y actividad biológica. En el diseño racional de nuevos derivados del bencimidazol, donde el objetivo no es únicamente mejorar la actividad antimicrobiana, sino hacerlo considerando las múltiples barreras biológicas y mecanismos de resistencia propios de *Staphylococcus aureus*.

## 5. Conclusiones

Se elaboró una base de datos conformada por 100 derivados del bencimidazol con actividad antimicrobiana reportada frente a *Staphylococcus aureus*, a partir de esta se calcularon y seleccionaron diez descriptores moleculares relevantes tales como: sumI, nHBAcc\_Lipinski, Kier1, nRotB, nRotBt, MW, PPSA-1, PPSA-2, WPSA-1 y THSA. Esta selección permitió representar adecuadamente propiedades fisicoquímicas asociadas a la actividad biológica. En este sentido los descriptores relacionados con la flexibilidad molecular, superficie polar y distribución electrónica pueden vincularse con procesos clave como es la penetración en la envoltura bacteriana y la interacción con dianas biológicas, estableciendo una matriz descriptor-actividad consistente para el modelado QSAR.

Se construyeron modelos predictivos mediante ElasticNetCV, Random Forest y SVR-RBF, evaluados a través de validación interna y externa. Random Forest presentó el mejor desempeño en validación externa, con errores absolutos mínimos de 0.01, 0.03, 0.13 y 0.29 unidades

$\log_{10}(\text{MIC})$  en los compuestos 4, 5, 8 y 9, respectivamente, determinando una mayor capacidad para capturar relaciones no lineales descriptor–actividad, en particular en el rango de 4–16  $\mu\text{g}/\text{mL}$ . Dicho comportamiento es relevante en el contexto de *S. aureus*, donde la actividad antimicrobiana depende de múltiples factores estructurales y biológicos, como la interacción con blancos alternativos frente a mecanismos de resistencia como expresión de PBP2a (*mecA*).

La evaluación estadística permitió establecer limitaciones en los modelos ElasticNetCV y SVR-RBF, cuyos errores alcanzaron valores de hasta 1.31 y 1.54 unidades  $\log_{10}(\text{MIC})$ , respectivamente. Sin embargo, el análisis QSAR–SAR permitió identificar descriptores estructurales clave, tales como: la flexibilidad conformacional (*nRotBt*, *nRotB*), la superficie polar (PPSA-1, WPSA-1) y la hidrofilia (THSA), asociados a la variación de la actividad antimicrobiana. Estos resultados sugieren que estas propiedades podrían influir en la capacidad que presentan los compuestos para superar mecanismos de resistencia bacteriana como las bombas de eflujo (*NorA*) y formación de biopelículas, permitiendo determinar los lineamientos racionales para la optimización estructural de nuevos derivados del bencimidazol frente a cepas resistentes de *S. aureus*.

## Referencias citadas

ACD/Labs. (2024). ACD/ChemSketch Freeware (Version 2024) [Software]. <https://www.acdlabs.com/>

Bansal, S., Gaur, R., Bhardwaj, H., Kumar, N., Sharma, G. K., & Mishra, S. S. (2024). Synthesis, biological evaluation and molecular docking studies of benzimidazole derivatives substitutes as new antimicrobial agents. *Journal of Drug Design and Medicinal Chemistry*, 10(2), 54–66. <https://doi.org/10.11648/j.jddmc.20241002.12>

Bush, K., & Bradford, P. A. (2016).  $\beta$ -Lactams and  $\beta$ -Lactamase Inhibitors: An Overview. *Cold Spring Harbor Perspectives In Medicine*, 6(8), a025247. <https://doi.org/10.1101/cshperspect.a025247>

Chambers, H. F., & DeLeo, F. R. (2009). Waves of resistance: *Staphylococcus aureus* in the antibiotic era. *Nature Reviews Microbiology*, 7(9), 629–641. <https://doi.org/10.1038/nrmicro2200>

Chen, H., Liu, H., Feng, H., Fu, H., Cai, W., Shao, X., & Chipot, C. (2021). MLCV: Bridging machine-learning-based dimensionality reduction and free-energy calculation. *Journal of Chemical Information and Modeling*, 62(1). <https://pubs.acs.org/doi/10.1021/acs.jcim.1c01010>

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). <https://doi.org/10.1145/2939672.2939785>

Cherkasov, A., Muratov, E. N., Fourches, D., Varnek, A., Baskin, I., Cronin, M., ... & Tropsha, A. (2014). QSAR modeling: Where have you been? Where are you going to? *Journal of Medicinal Chemistry*, 57(12), 497–5010. <https://doi.org/10.1021/jm4004285>

Daina, A., Michielin, O., & Zoete, V. (2017). SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Scientific Reports*, 7, 42717. <https://doi.org/10.1038/srep42717>

Da Rocha, M. N., De Sousa, D. S., Da Silva Mendes, F. R., Santos, H. S. D., Marinho, G. S., Marinho, M. M., & Marinho, E. S. (2024b). Ligand and structure-based virtual screening approaches in drug discovery: minireview. *Molecular Diversity*, 29(3), 2799–2809. <https://doi.org/10.1007/s11030-024-10979-6>

Dokla, E. M. E., Abutaleb, N. S., Milik, S. N., Li, D., El-Baz, K., Shalaby, M.-A. W., Al-Karaki, R., Nasr, M., Klein, C. D., Abouzid, K. A. M., & Seleem, M. N. (2020). *Development of benzimidazole-based derivatives as antimicrobial agents and their synergistic effect with colistin against gram-negative bacteria*. *European Journal of Medicinal Chemistry*, 186, 111850. <https://doi.org/10.1016/j.ejmech.2019.111850>

Foster, T. J. (2017). Antibiotic resistance in *Staphylococcus aureus*. Current status and future prospects. *FEMS Microbiology Reviews*, 41(3), 430–449. <https://doi.org/10.1093/femsre/fux007>

Gaulton, A., Hersey, A., Nowotka, M., Bento, A. P., Chambers, J., Mendez, D., ... & Leach, A. R. (2017). The ChEMBL database in 2017. *Nucleic Acids Research*, 45(D1), D945–D954. <https://doi.org/10.1093/nar/gkw1074>

Hanwell, M. D., Curtis, D. E., Lonie, D. C., Vandermeersch, T., Zurek, E., & Hutchison, G. R. (2012). Avogadro: An advanced semantic chemical editor, visualization, and analysis platform. *Journal of Cheminformatics*, 4, 17. <https://doi.org/10.1186/1758-2946-4-17>

He, Y., Yang, J., Wu, B., Risen, L., & Swayze, E. E. (2004). Synthesis and biological evaluations of novel benzimidazoles as potential antibacterial agents. *Bioorganic & Medicinal Chemistry Letters*, 14(5), 1217-1220. <https://doi.org/10.1016/j.bmcl.2003.12.051>

Kovalishyn, V. V., Aires-de-Sousa, J., Ventura, C., & Leitao, R. (2011). QSAR modeling of antitubercular activity of diverse organic compounds. *Chemometrics and Intelligent Laboratory Systems* 107(1):69-74. DOI:10.1016/j.chemolab.2011.01.011

Küçükbay, H., Uçkun, M., Apohan, E., & Yeşilada, Ö. (2021). Cytotoxic and antimicrobial potential of bencimidazole derivatives. *Archiv der Pharmazie*, 354(4), e2100076. <https://doi.org/10.1002/ardp.202100076>

Landrum, G. (s. f.). RDKit: Open-source cheminformatics [Software]. <https://www.rdkit.org>

Lipinski, C. A. (2004). Lead- and drug-like compounds: The rule-of-five revolution. *Drug Discovery Today: Technologies*, 1(4), 337–341. <https://doi.org/10.1016/j.ddtec.2004.11.007>

Li, Q., Cheng, T., Wang, Y., & Bryant, S. H. (2010). PubChem as a public resource for drug discovery. *Drug Discovery Today*, 15(23–24), 1052–1057. <https://doi.org/10.1016/j.drudis.2010.10.003>

Marvin - Product Introduction 2024. (2025, 20 octubre). [Vídeo]. <https://chemaxon.com/products/marvin>

Malasala, S., Ahmad, M. N., Akunuri, R., Shukla, M., Kaul, G., Dasgupta, A., Madhavi, Y. V., Chopra, S., & Nanduri, S. (2020). Synthesis and evaluation of new quinazoline-bencimidazole hybrids as potent anti-microbial agents against multidrug resistant *Staphylococcus aureus* and *Mycobacterium tuberculosis*. *European Journal of Medicinal Chemistry*, 212, 112996. <https://doi.org/10.1016/j.ejmech.2020.112996>

Marinescu, M. (2023). Bencimidazole-triazole hybrids as antimicrobial and antiviral agents: A systematic review. *Antibiotics*, 12(7), 1220. <https://doi.org/10.3390/antibiotics12071220>

Mishra, V. R., Ghanavatkar, C. W., Mali, S. N., Qureshi, S. I., Chaudhari, H. K., & Sekar, N. (2019). Design, synthesis, antimicrobial activity and computational studies of novel azo linked substituted bencimidazole, benzoxazole and benzothiazole derivatives. *Computational Biology and Chemistry*, 78, 330–337. <https://doi.org/10.1016/j.compbiolchem.2019.01.003>

Muratov, E. N., Bajorath, J., Sheridan, R. P., Tetko, I. V., Filimonov, D., Poroikov, V., ... & Tropsha, A. (2020). QSAR without borders. *Chemical Society Reviews*, 49(11), 3525–3564. <https://doi.org/10.1039/D0CS00098A>

Narasimhan, B., Sharma, D., Kumar, P., Yogeewari, P., & Sriram, D. (2011). Synthesis, antimicrobial and antimycobacterial evaluation of [2-(substituted phenyl)-imidazol-1-yl]-pyridin-3-yl-methanones. *Journal of Enzyme Inhibition and Medicinal Chemistry*, 26(5), 720–727. <https://doi.org/10.3109/14756366.2010.548331>

Organisation for Economic Co-operation and Development (OECD). (2007). Guidance document on the validation of (quantitative) structure–activity relationships [(Q)SAR] models (ENV/JM/MONO(2007)2). [https://one.oecd.org/document/ENV/JM/MONO\(2007\)2/en/pdf](https://one.oecd.org/document/ENV/JM/MONO(2007)2/en/pdf)

Organización Mundial de la Salud. (2024). La crisis de resistencia a los antimicrobianos “es ya un hecho”. Noticias ONU. <https://news.un.org/es/story/2024/11/1534346>

Otto, M. (2014). Staphylococcus aureus toxins. *Current Opinion In Microbiology*, 17, 32-37. <https://doi.org/10.1016/j.mib.2013.11.004>

Padalkar, V. S., Borse, B. N., Gupta, V. D., Phatangare, K. R., Patil, V. S., Umape, P. G., & Sekar, N. (2011). Synthesis and antimicrobial activity of novel 2-substituted benzimidazole, benzoxazole and benzothiazole derivatives. *Arabian Journal Of Chemistry*, 9, S1125-S1130. <https://doi.org/10.1016/j.arabjc.2011.12.006>

Paul, D., Sanap, G., Shenoy, S., Kalyane, D., Kalia, K., & Tekade, R. K. (2020). Artificial intelligence in drug discovery and development. *Drug Discovery Today*, 26(1), 80-93. <https://doi.org/10.1016/j.drudis.2020.10.010>

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. <https://jmlr.csail.mit.edu/papers/v12/pedregosa11a.html>

Pires, D. E. V., Blundell, T. L., & Ascher, D. B. (2015). pkCSM: predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. *Journal of Medicinal Chemistry*, 58(9), 4066–4072. <https://doi.org/10.1021/acs.jmedchem.5b00104>

Pham, E. C., Le, T. V. T., & Truong, T. N. (2022). Design, synthesis, bio-evaluation, and in silico studies of some N-substituted 6-(chloro/nitro)-1H-benzimidazole derivatives as antimicrobial and anticancer agents. *RSC Advances*, 12(33), 21621-21646. <https://doi.org/10.1039/d2ra03491c>

Roy, K., Kar, S., & Das, R. N. (2015). Understanding the basics of QSAR for applications in pharmaceutical sciences and risk assessment. Academic Press. [https://www.researchgate.net/publication/273348080\\_Understanding\\_the\\_Basics\\_of\\_QSAR\\_for\\_Applications\\_in\\_Pharmaceutical\\_Sciences\\_and\\_Risk\\_Assessment](https://www.researchgate.net/publication/273348080_Understanding_the_Basics_of_QSAR_for_Applications_in_Pharmaceutical_Sciences_and_Risk_Assessment). ISBN: 9780128016336 [https://www.researchgate.net/publication/273348080\\_Understanding\\_the\\_Basics\\_of\\_QSAR\\_for\\_Applications\\_in\\_Pharmaceutical\\_Sciences\\_and\\_Risk\\_Assessment](https://www.researchgate.net/publication/273348080_Understanding_the_Basics_of_QSAR_for_Applications_in_Pharmaceutical_Sciences_and_Risk_Assessment).

Roy, K., Kar, S., & Das, R. N. (2015a). A Primer on QSAR/QSPR Modeling. En *Springer briefs in molecular science*. <https://doi.org/10.1007/978-3-319-17281-1>

Sanner, M. F. (1999). Python: A programming language for software integration and development. PubMed. <https://pubmed.ncbi.nlm.nih.gov/10660911/>

Schrödinger, LLC. (2026). PyMOL Molecular Graphics System, Version 3.1. <https://pymol.org/2/>

Sheridan, R. P. (2013). Time-split cross-validation as a method for estimating the goodness of prospective prediction. *Journal of Chemical Information and Modeling*, 53(4), 783–790. <https://doi.org/10.1021/ci400084k>

Salahuddin, Shaharyar, M., & Mazumder, A. (2012). Benzimidazoles: A biologically active compounds. *Arabian Journal Of Chemistry*, 10, S157-S173. <https://doi.org/10.1016/j.arabjc.2012.07.017>

Shalaby, M., Dokla, E., Serya, R., & Abouzid, K. (2019). Identification of novel pyrazole and benzimidazole based derivatives as PBP2a inhibitors: Design, synthesis, and biological evaluation. *Archives Of Pharmaceutical Sciences Ain Shams University*, 3(2), 228-245.

<https://doi.org/10.21608/aps.2019.16625.1010>. chrome-extension://efaidnbmninnibpcjpcglclefindmkaj/https://pdfs.semanticscholar.org/0347/6acd3aff360c86ab49a73dd1ea8f41e8f11b.pdf

Sharma, D., Narasimhan, B., Kumar, P., & Jalbout, A. (2008). Synthesis and QSAR evaluation of 2-(substituted phenyl)-1H-benzimidazoles and [2-(substituted phenyl)-benzimidazol-1-yl]-pyridin-3-yl-methanones. *European Journal Of Medicinal Chemistry*, 44(3), 1119-1127. <https://doi.org/10.1016/j.ejmech.2008.06.009>

Singh, R., Lin, Y.-W., Zhao, J., & Li, J. (2026). Artificial intelligence and machine learning in antimicrobial discovery, resistance prediction, and precision therapy. *International Journal of Antimicrobial Agents*, 67(3), 107715. <https://doi.org/10.1016/j.ijantimicag.2026.107715>

Tiwari, D., Narang, R., & Khokra, S. L. (2021b). A review on microwave assisted synthesis, mechanism of action and structure activity relationship of 1, 3, 4-oxadiazoleA review on microwave assisted synthesis, mechanism of action and structure activity relationship of 1, 3, 4-oxadiazole derivatives as anticancer agent derivatives as anticancer agent. *World Journal Of Advanced Research And Reviews*, 9(1), 086-096. <https://doi.org/10.30574/wjarr.2021.9.1.0472>

Tunçbilek, M., Kiper, T., & Altanlar, N. (2008). Synthesis and in vitro antimicrobial activity of some novel substituted benzimidazole derivatives having potent activity against MRSA. *European Journal Of Medicinal Chemistry*, 44(3), 1024-1033. <https://doi.org/10.1016/j.ejmech.2008.06.026>

Tong, S. y. C., Davis, J. S., Eichenberger, E., Holland, T. L., & Fowler, V. G. (2015). Staphylococcus aureus Infections: Epidemiology, Pathophysiology, Clinical Manifestations, and Management. *Clinical Microbiology Reviews*, 28(3), 603-661. <https://doi.org/10.1128/cmr.00134-14>

Tropsha, A. (2010). Best practices for QSAR model development, validation, and exploitation. *Molecular Informatics*, 29(6-7), 476-488. <https://doi.org/10.1002/minf.201000061>

Trott, O., & Olson, A. J. (2009). AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal Of Computational Chemistry*, 31(2), 455-461. <https://doi.org/10.1002/jcc.21334>

Todeschini, R., & Consonni, V. (2009). Molecular Descriptors for Chemoinformatics. En *Methods and principles in medicinal chemistry*. <https://doi.org/10.1002/9783527628766>. <https://www.scirp.org/reference/referencespapers?referenceid=2596048>

Vamathevan, J., Clark, D., Czodrowski, P., Dunham, I., Ferran, E., Lee, G., Li, B., Madabhushi, A., Shah, P., Spitzer, M., & Zhao, S. (2019). Applications of machine learning in drug discovery and development. *Nature Reviews Drug Discovery*, 18(6), 463-477. <https://doi.org/10.1038/s41573-019-0024-5>

Vinueza, H., Valeria, K. (2015). Prevalencia de Staphylococcus aureus resistente a meticilina (SARM) en pacientes atendidos en el Hospital General Enrique Garcés durante el periodo enero 2013-diciembre 2013. *Repositorio Digital UCE*. Quito. <https://www.dspace.uce.edu.ec/entities/publication/0a33afaf-6acb-46b7-926e-d34df937edce>

Wang, Y., Bheemanaboina, R. R. Y., Cai, G., & Zhou, C. (2018). Novel purine benzimidazoles as antimicrobial agents by regulating ROS generation and targeting clinically resistant Staphylococcus aureus DNA groove. *Bioorganic & Medicinal Chemistry Letters*, 28(9), 1621-1628. <https://doi.org/10.1016/j.bmcl.2018.03.046>

Wu, Z., Ramsundar, B., Feinberg, E. N., Gomes, J., Geniesse, C., Pappu, A. S., Leswing, K., & Pande, V. (2017). MoleculeNet: a benchmark for molecular machine learning. *Chemical Science*, 9(2), 513-530. <https://doi.org/10.1039/c7sc02664a>

Yang, X., Wang, Y., Byrne, R., Schneider, G., & Yang, S. (2019). Concepts of Artificial Intelligence for Computer-Assisted Drug Discovery. *Chemical Reviews*, 119(18), 10520-10594. <https://doi.org/10.1021/acs.chemrev.8b00728>

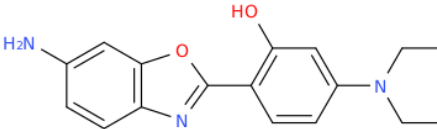
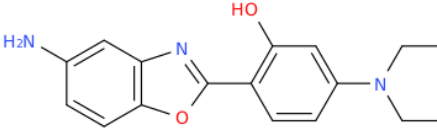
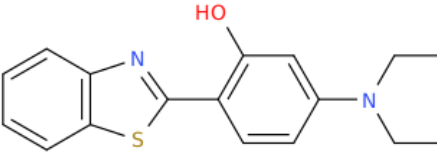
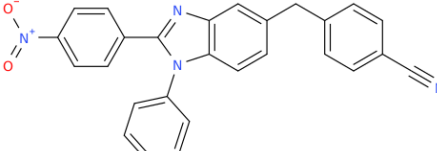
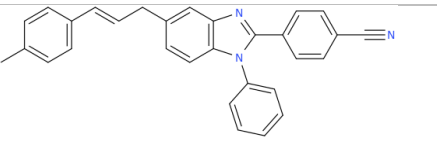
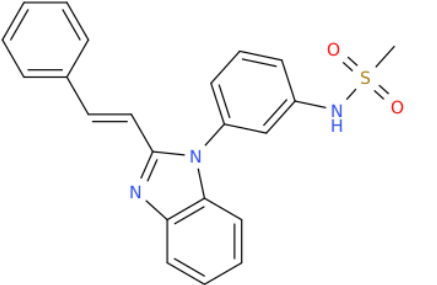
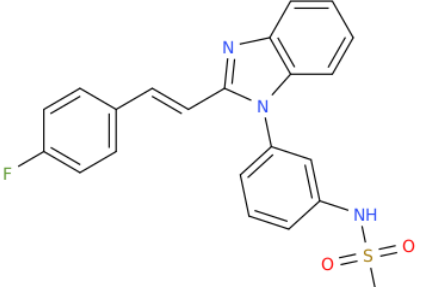
Yap, C. W. (2010). PaDEL-descriptor: An open source software to calculate molecular descriptors and fingerprints. *Journal Of Computational Chemistry*, 32(7), 1466-1474. <https://doi.org/10.1002/jcc.21707>

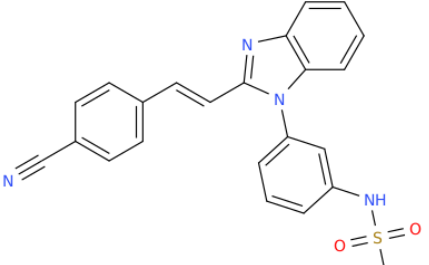
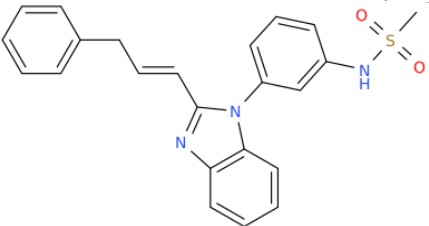
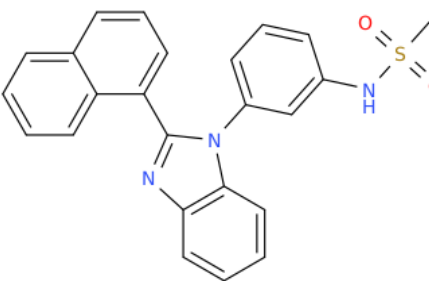
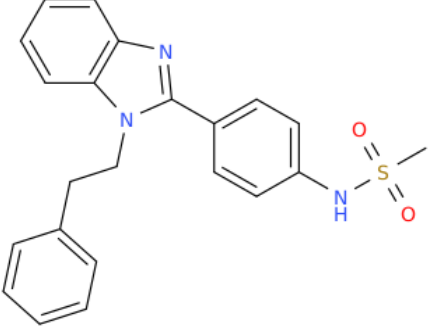
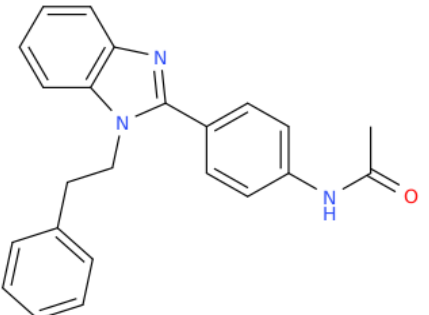
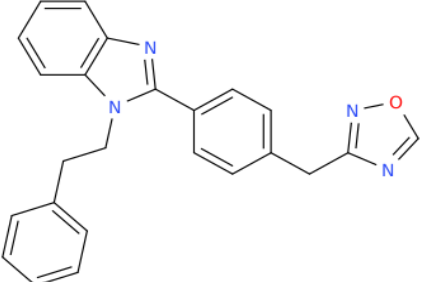
Zhou, B., Li, B., Yi, W., Bu, X., & Ma, L. (2013). Synthesis, antioxidant, and antimicrobial evaluation of some 2-arylbenzimidazole derivatives. *Bioorganic & Medicinal Chemistry Letters*, 23(13), 3759-3763. <https://doi.org/10.1016/j.bmcl.2013.05.004>

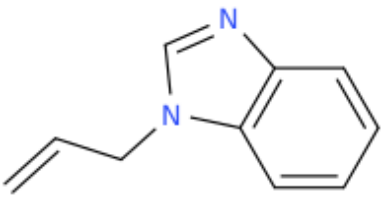
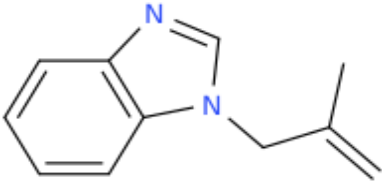
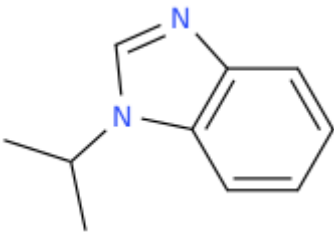
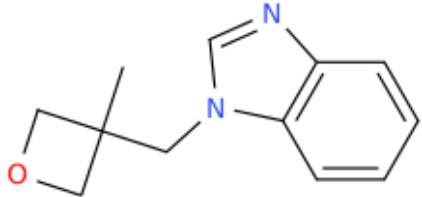
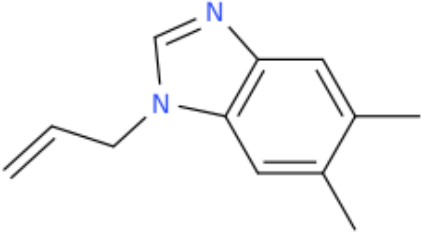
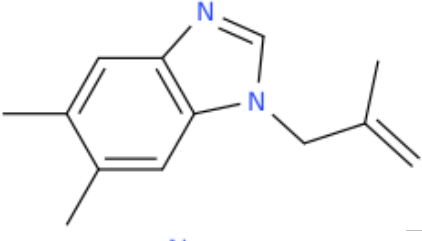
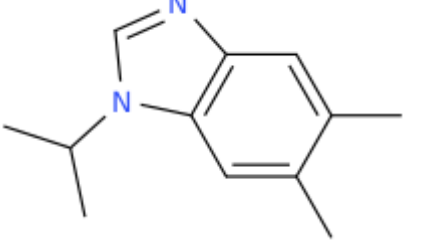
## Anexos

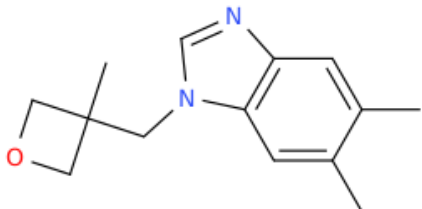
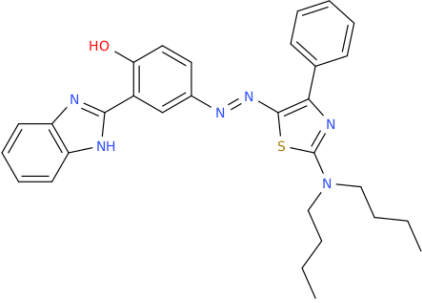
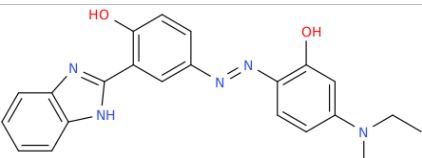
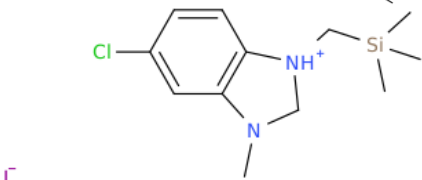
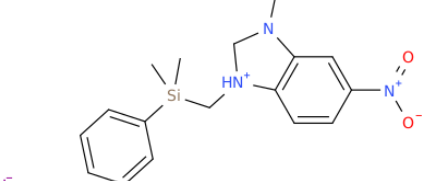
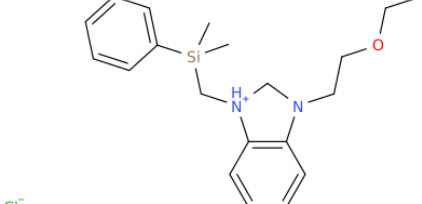
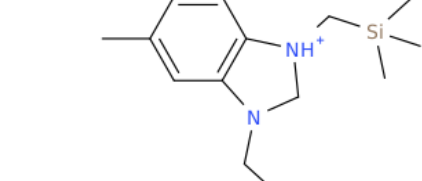
### Anexo 1: Matriz de 100 derivados de Bencimidazol obtenidos de las plataformas PubChem, ChemBL y Papers indexados en PubMed.

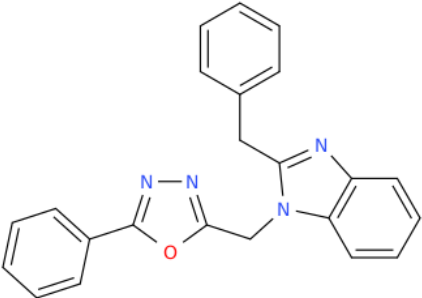
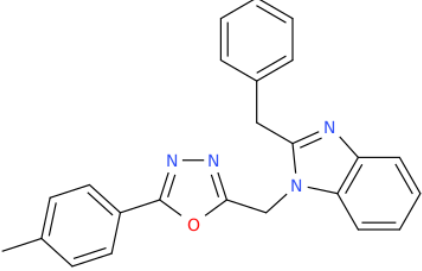
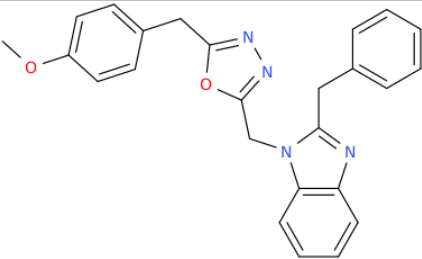
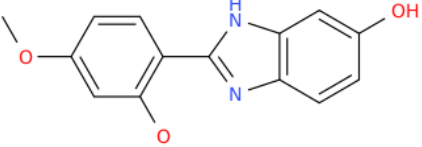
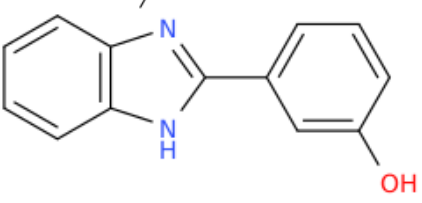
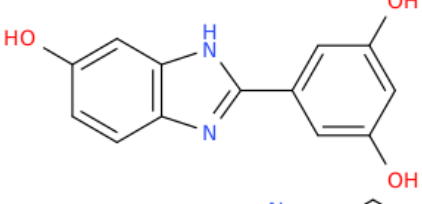
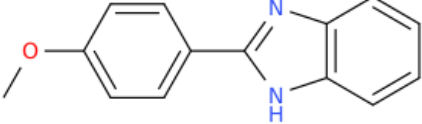
N°	FÓRMULA MOLECULAR	ESCRITURA IUPAC	MIC $\mu\text{g/mL}$	ESTRUCTURA 2D	CID/DOI/SMILES
1	$\text{C}_{15}\text{H}_{17}\text{N}_3\text{O}$	5-(N,N-dietilamino)-2-(1H-benzimidazol-2-il)fenol	65		136649796 <chem>C(C)N(CC)C=1C=CC(=C(C1)O)C1=NC2=C(N1)C=CC=C2</chem>
2	$\text{C}_{15}\text{H}_{16}\text{N}_4\text{O}_3$	5-(dietilamino)-2-(5,6-dinitro-1H-benzimidazol-2-il)fenol	65		154638964 <chem>C(C)N(C=1C=CC(=C(C1)O)C1=NC2=C(N1)C=C(C(=C2)[N+](=O)[O-])[N+](=O)[O-])CC</chem>
3	$\text{C}_{15}\text{H}_{18}\text{N}_4\text{O}$	2-(5-amino-1H-benzimidazol-2-il)-5-(N,N-dietilamino)fenol	65		<a href="https://doi.org/10.1016/j.arabjc.2011.12.006">https://doi.org/10.1016/j.arabjc.2011.12.006</a> DOI: 10.1016/j.heliyon.2020.e04126 1372934-38-7 <chem>NC1=CC2=C(NC(=N2)C2=C(C=C(C(=C2)N(CC)CC)O)C=C1</chem>
4	$\text{C}_{15}\text{H}_{16}\text{N}_2\text{O}_2$	2-(1,3-benzoxazol-2-il)-5-(N,N-dietilamino)fenol	140		136827403 <chem>O1C(=NC2=C1C=CC=C2)C2=C(C=C(C=C2)N(CC)CC)O</chem>
5	$\text{C}_{15}\text{H}_{15}\text{N}_3\text{O}_4$	5-(N,N-dietilamino)-2-(6-nitro-1,3-benzoxazol-2-il)fenol	140		136835551 <chem>C(C)N(CC)C=1C=CC(=C(C1)O)C=1OC2=C(N1)C=CC(=C2)[N+](=O)[O-]</chem>
6	$\text{C}_{15}\text{H}_{15}\text{N}_3\text{O}_4$	5-(N,N-dietilamino)-2-(5-nitro-1,3-benzoxazol-2-il)fenol	250		136275176 <chem>C(C)N(CC)C=1C=CC(=C(C1)O)C=1OC2=C(N1)C=C(C(=C2)[N+](=O)[O-])</chem>

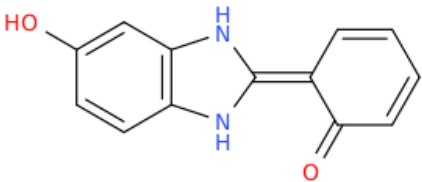
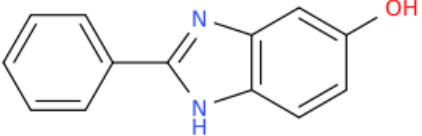
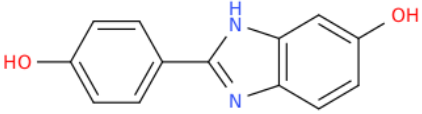
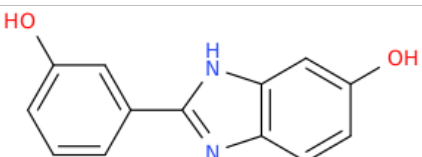
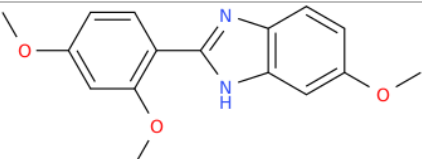
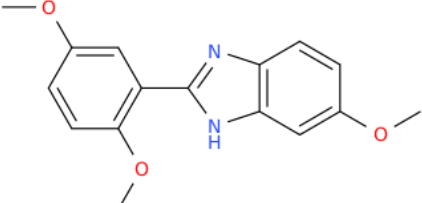
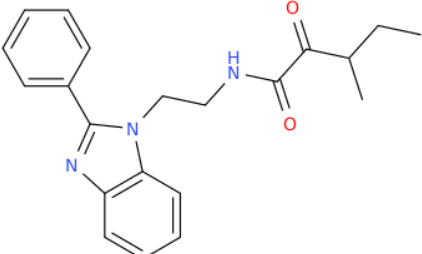
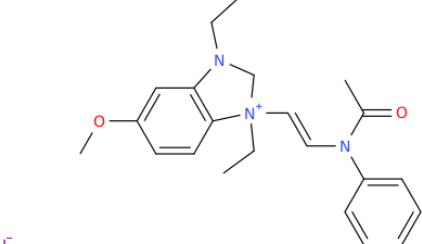
7	<b>C<sub>15</sub>H<sub>17</sub>N<sub>3</sub>O<sub>2</sub></b>	2-(6-amino-1,3-benzoxazol-2-il)-5-(N,N-dietilamino)fenol	65		136827405 <chem>NC1=CC2=C(N=C(O2)C2=C(C=C(C=C2)N(CC)CC)O)C=C1</chem>
8	<b>C<sub>15</sub>H<sub>17</sub>N<sub>3</sub>O<sub>2</sub></b>	2-(5-amino-1,3-benzoxazol-2-il)-5-(N,N-dietilamino)fenol (isómero)	65		136827404 <chem>CCN(CC)C1=CC(=C(C=C1)C2=NC3=C(O2)C=CC(=C3)N)O</chem>
9	<b>C<sub>15</sub>H<sub>16</sub>N<sub>2</sub>OS</b>	2-(1,3-benzotiazol-2-il)-5-(N,N-dietilamino)fenol	140		193939 <chem>S1C(=NC2=C1C=CC=C2)C2=C(C=C(C=C2)N(CC)CC)O</chem>
10	<b>C<sub>27</sub>H<sub>21</sub>N<sub>4</sub>O<sub>2</sub></b>	2-(4-nitrofenil)-5-(4-cianobencil)-1-fenil-1H-bencimidazol	32		<a href="https://doi.org/10.1016/j.bmc.1.2018.03.046">https://doi.org/10.1016/j.bmc.1.2018.03.046</a> <chem>[N+](=O)([O-])C1=CC=C(C=C1)C2=NC3=C(N1C1=CC=CC=C1)C=C(C=C2)CC2=CC=C(C=C2)C#N</chem>
11	<b>C<sub>32</sub>H<sub>25</sub>N<sub>3</sub></b>	1-fenil-2-(4-cianofenil)-5-[3-(4-metilfenil)prop-2-en-1-il]-1H-bencimidazol	128		<a href="https://doi.org/10.1016/j.bmc.1.2018.03.046">https://doi.org/10.1016/j.bmc.1.2018.03.046</a> <chem>C1(=CC=CC=C1)N1C(=NC2=C1C=CC(=C2)CC=CC2=C(C=C2)C)C2=CC=C(C=C2)C#N</chem>
12	<b>C<sub>22</sub>H<sub>19</sub>N<sub>3</sub>O<sub>2</sub>S</b>	(E)-N-(3-(2-estiril-1H-bencimidazol-1-il)fenil)metanosulfonamida	64		155817108 <chem>C(=C(C1=CC=CC=C1)/C1=NC2=C(N1C=1C=C(C=CC1)NS(=O)(=O)C)C=CC=C2</chem>
13	<b>C<sub>22</sub>H<sub>18</sub>FN<sub>3</sub>O<sub>2</sub>S</b>	(E)-N-(3-(2-[(4-fluorofenil)eten-1-il]-1H-bencimidazol-1-il)fenil)metanosulfonamida	64		155817108 DOI: <a href="https://doi.org/10.3390/antibiotics12071220">10.3390/antibiotics12071220</a> <chem>FC1=CC=C(C=C1)/C=C/C1=NC2=C(N1C=1C=C(C=CC1)NS(=O)(=O)C)C=CC=C2</chem>

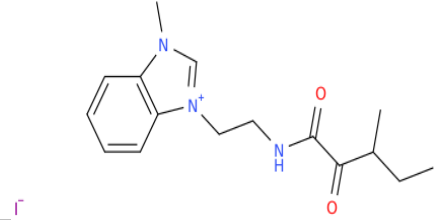
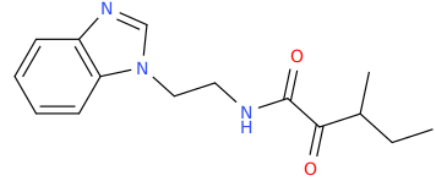
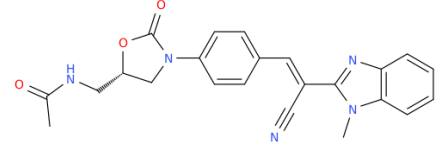
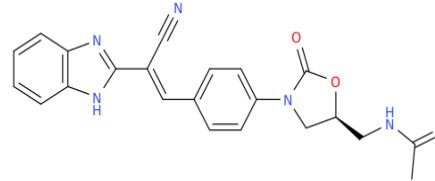
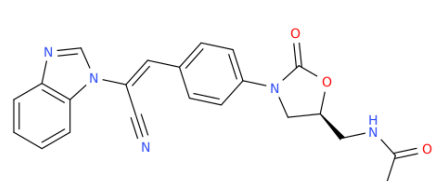
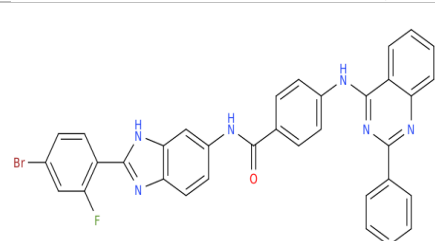
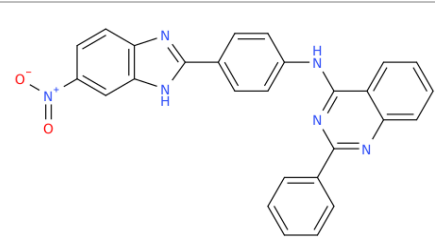
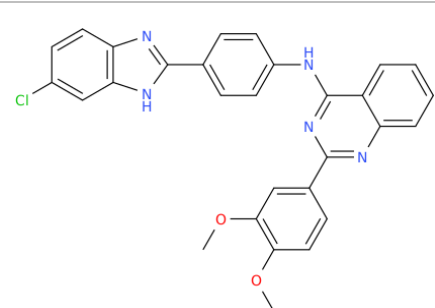
14	<b>C23H18N4O2S</b>	(E)-N-(3-(2-[(4-cianofenil)eten-1-il]-1H-bencimidazol-1-il)fenil)metanosulfonamida	64		DOI: 10.3390/antibiotics12071220  <chem>C(#N)C1=CC=C(C=C1)/C=C/C1=NC2=C(N1C=1C=C(C=CC1)NS(=O)(=O)C)C=C2</chem>
15	<b>C23H21N3O2S</b>	(E)-N-(3-(2-(3-fenilprop-1-en-1-il)-1H-bencimidazol-1-il)fenil)metanosulfonamida	128		DOI: 10.3390/antibiotics12071220  <chem>C1(=CC=CC=C1)C/C=C/C1=NC2=C(N1C=1C=C(C=CC1)NS(=O)(=O)C)C=CC=C2</chem>
16	<b>C24H19N3O2S</b>	N-(3-(2-(naftalen-1-il)-1H-bencimidazol-1-il)fenil)metanosulfonamida	128		162666053  <a href="https://doi.org/10.1016/j.bmc.1.2018.03.046">https://doi.org/10.1016/j.bmc.1.2018.03.046</a>  <chem>C1(=CC=CC2=CC=CC=C12)C1=NC2=C(N1C=1C=C(C=CC1)NS(=O)(=O)C)C=CC=C2</chem>
17	<b>C22H21N3O2S</b>	N-(4-(1-(2-feniletil)-1H-bencimidazol-2-il)fenil)metanosulfonamida	32		Doi: 10.1016/j.bmc.2008.02.020  <chem>C1(=CC=CC=C1)CCN1C(=NC2=C1C=CC=C2)C2=CC=C(C=C2)NS(=O)(=O)C</chem>
18	<b>C23H21N3O</b>	N-(4-(1-(2-feniletil)-1H-bencimidazol-2-il)fenil)acetamida	64		Doi: 10.1016/j.bmc.2008.02.020.  <chem>C1(=CC=CC=C1)CCN1C(=NC2=C1C=CC=C2)C2=CC=C(C=C2)NC(C)=O</chem>
19	<b>C24H19N4O</b>	1-(1-(2-feniletil)-1H-bencimidazol-2-il)-4-[(1,2,4-oxadiazol-3-il)metil]benceno	128		<a href="https://doi.org/10.1016/j.bmc.1.2018.03.046">https://doi.org/10.1016/j.bmc.1.2018.03.046</a>  <chem>C1(=CC=CC=C1)CCN1C(=NC2=C1C=CC=C2)C2=CC=C(C=C2)CC2=NOC=N2</chem>

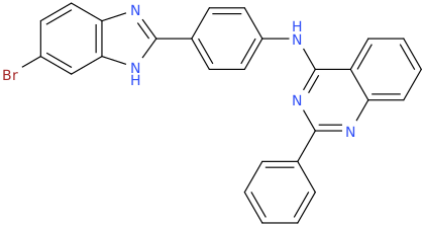
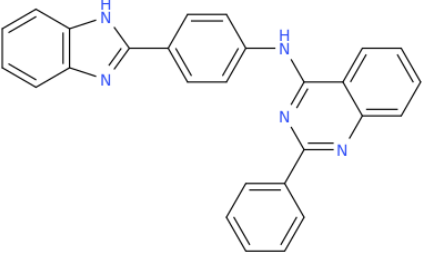
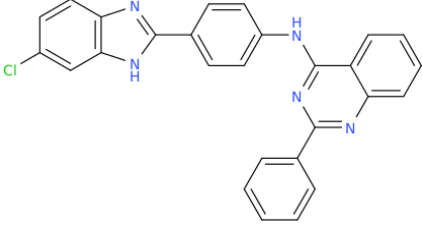
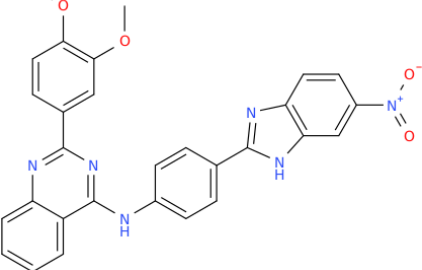
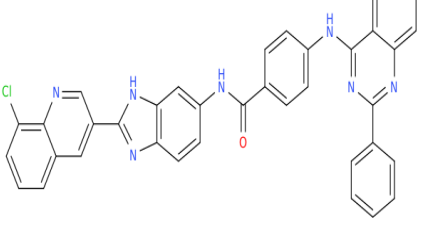
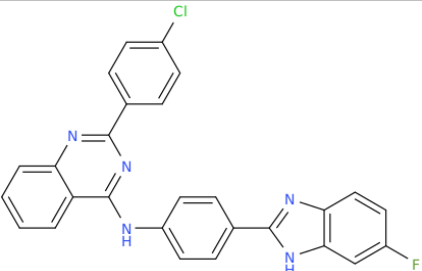
20	<b>C10H10N2</b>	1-(prop-2-en-1-il)-1H-bencimidazol	0.206		594117 <chem>C(C=C)N1C=NC2=C1C=CC=C2</chem>
21	<b>C11H12N2</b>	1-(2-metilprop-2-en-1-il)-1H-bencimidazol	0.194		847268 <chem>CC(CN1C=NC2=C1C=CC=C2)=C</chem>
22	<b>C10H12N2</b>	1-isopropil-1H-bencimidazol	0.102		914516 <chem>C(C)(C)N1C=NC2=C1C=C</chem> <chem>C=C2</chem>
23	<b>C13H16N2O</b>	1-(3-metiloxetan-3-il)metil-1H-bencimidazol	0.087		139208561 <chem>CC1(COC1)CN1C=NC2=C1C=CC=C2</chem>
24	<b>C12H14N2</b>	1-alil-5,6-dimetil-1H-bencimidazol	0.092		962377 <chem>C(C=C)N1C=NC2=C1C=C(C</chem> <chem>C(=C2)C)C</chem>
25	<b>C13H16N2</b>	1-(2-metilalil)-5,6-dimetil-1H-bencimidazol	0.175		962379 <chem>CC(CN1C=NC2=C1C=C(C</chem> <chem>(=C2)C)C)=C</chem>
26	<b>C12H16N2</b>	1-isopropil-5,6-dimetil-1H-bencimidazol	0.183		20453556 <chem>C(C)(C)N1C=NC2=C1C=C(C</chem> <chem>C(=C2)C)C</chem>

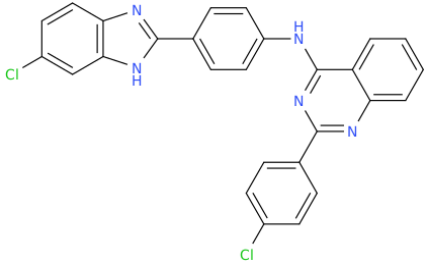
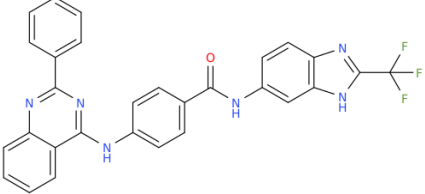
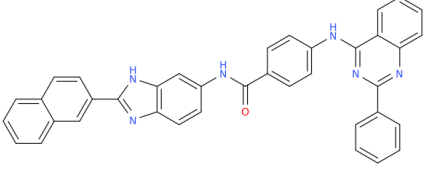
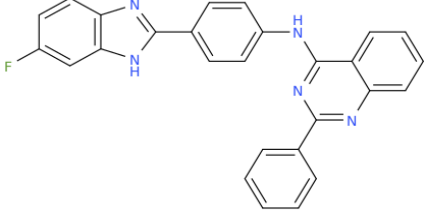
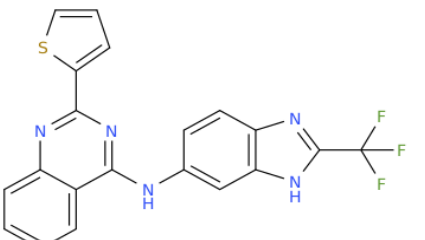
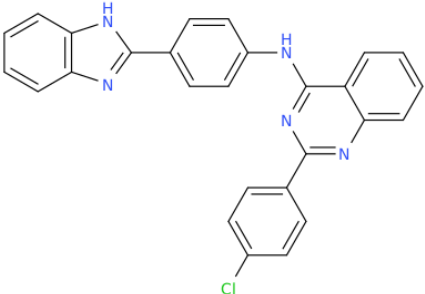
27	<b>C15H20N2O</b>	1-(3-metiloxetan-3-il)metil-5,6-dimetil-1H-bencimidazol	0.159		20453556 <chem>CC1(COC1)CN1C=NC2=C1C=C(C(=C2)C)C</chem>
28	<b>C30H32N6OS</b>	2-(1H-bencimidazol-2-il)-4-[(2-(dibutilamino)-4-feniltiazol-5-il)diazenil]fenol	625		<a href="https://doi.org/10.1016/j.pbiolchem.2019.01.003">https://doi.org/10.1016/j.pbiolchem.2019.01.003</a> <chem>N1C(=NC2=C1C=CC=C2)C2=C(C=CC(=C2)N=NC2=C(N=C(S2)N(CCCC)CCCC)C2=CC=CC=C2)O</chem>
29	<b>C23H23N5O2</b>	2-(1H-bencimidazol-2-il)-4-[(4-(dietilamino)-2-hidroxifenil)diazenil]fenol	625		<a href="https://doi.org/10.1016/j.pbiolchem.2019.01.003">https://doi.org/10.1016/j.pbiolchem.2019.01.003</a> <chem>N1C(=NC2=C1C=CC=C2)C2=C(C=CC(=C2)N=NC2=C(C=C(C(=C2)N(CC)CC)O)O</chem>
30	<b>C12H18N2ClISi</b>	yoduro de 5-cloro-3-metil-1-[(trimetilsilil)metil]-2,3-dihidro-1H-bencimidazolio	200		<a href="https://doi.org/10.1016/j.bmc.1.2013.05.004">https://doi.org/10.1016/j.bmc.1.2013.05.004</a> [I-].C1C1=CC2=C([NH+](CN2)C[Si](C)(C)C)C=C1
31	<b>C17H20N3O2SiI</b>	yoduro de 1-[(dimetil(fenil)silil)metil]-3-metil-5-nitro-1H-bencimidazolio	400		46919760 [I-].C[Si](C1=CC=CC=C1)(C)C[NH+]1CN(C2=C1C=CC(=C2)[N+](=O)[O-])C
32	<b>C20H27N2OSiCl</b>	cloruro de 1-[(dimetil(fenil)silil)metil]-3-(2-etoxietil)-1H-bencimidazolio	400		46919760 [Cl-].C[Si](C1=CC=CC=C1)(C)C[N+]1=CN(C2=C1C=CC(=C2)CCOCC
33	<b>C14H23N2SiI</b>	yoduro de 3-etil-5-metil-1-[(trimetilsilil)metil]-1H-bencimidazolio	400		46919843 [I-].C(C)N1C=[N+](C2=C1C=C(C(=C2)C)C[Si](C)(C)C

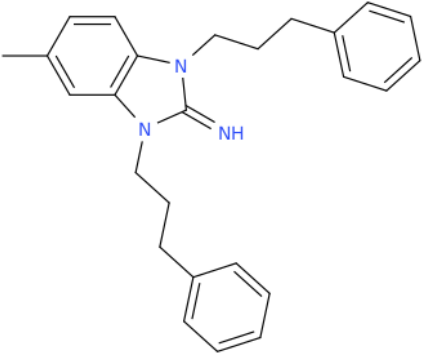
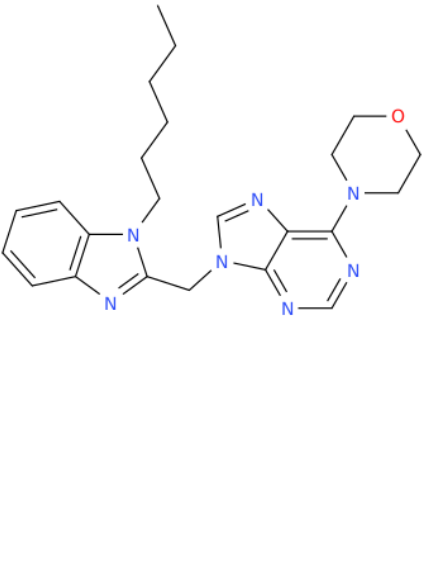
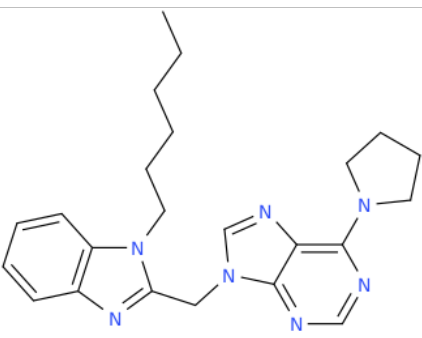
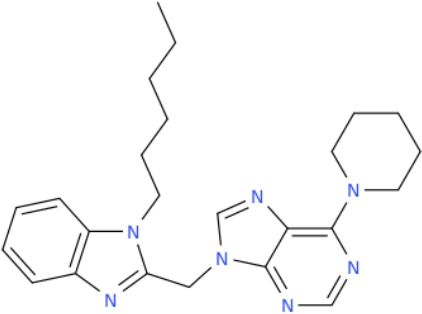
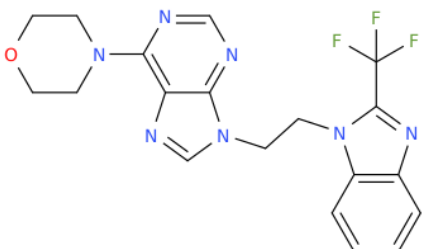
34	<b>C23H18N4O</b>	2-bencil-1-[(5-fenil-1,3,4-oxadiazol-2-il)metil]-1H-bencimidazol	12.5		46919760 <chem>C(C1=CC=CC=C1)C1=NC2=C(N1CC=1OC(=NN1)C1=CC=CC=C1)C=CC=C2</chem>
35	<b>C24H20N4O</b>	2-bencil-1-[[5-(p-tolil)-1,3,4-oxadiazol-2-il]metil]-1H-bencimidazol	12.5		46919842 <chem>C(C1=CC=CC=C1)C1=NC2=C(N1CC=1OC(=NN1)C1=CC=C(C=C1)C)C=CC=C2</chem>
36	<b>C24H20N4O2</b>	2-bencil-1-[[5-(p-anisil)-1,3,4-oxadiazol-2-il]metil]-1H-bencimidazol	12.5		46919843 <chem>C(C1=CC=CC=C1)C1=NC2=C(N1CC=1OC(=NN1)C1=CC=C(C=C1)OC)C=CC=C2</chem>
37	<b>C15H14N2O3</b>	2-(2,4-dimetoxifenil)-3H-bencimidazol-5-ol	0.104		71681388 <chem>COC1=C(C=CC(=C1)OC)C=1NC2=C(N1)C=CC(=C2)O</chem>
38	<b>C13H10N2O</b>	3-(1H-bencimidazol-2-il)fenol	0.076		763204 <chem>N1C(=NC2=C1C=CC=C2)C=2C=C(C=CC2)O</chem>
39	<b>C13H10N2O3</b>	5-(6-hidroxi-1H-bencimidazol-2-il)benceno-1,3-diol	0.186		71681237 <chem>OC=1C=CC2=C(NC(=N2)C=2C=C(C=C(C2)O)O)C1</chem>
40	<b>C14H12N2O</b>	2-(4-metoxifenil)-1H-bencimidazol	0.35		345686 <chem>COC1=CC=C(C=C1)C1=NC2=C(N1)C=CC=C2</chem>

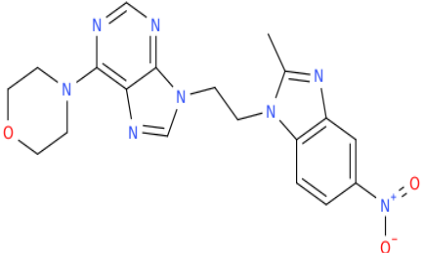
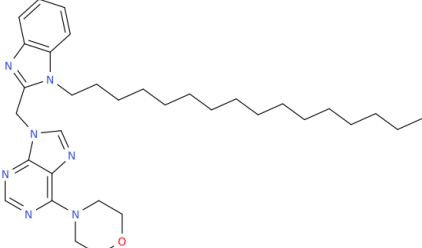
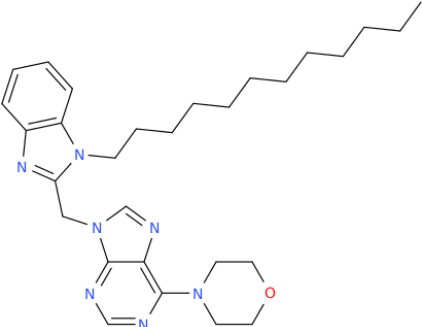
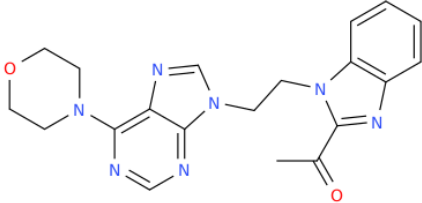
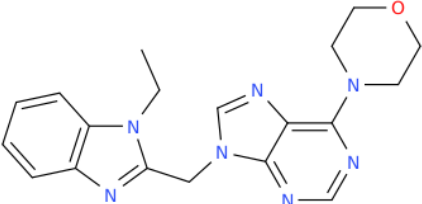
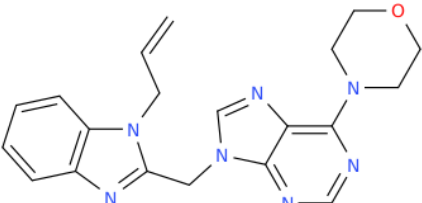
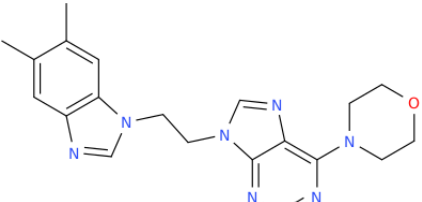
41	<b>C13H10N2O2</b>	6-(5-hidroxi-1,3-dihidro-2H-bencimidazol-2-ilideno)ciclohexa-2,4-dien-1-ona	0.328		136260506 <chem>OC1=CC2=C(NC(N2)=C2C=CC=CC2=O)C=C1</chem>
42	<b>C13H10N2O</b>	2-fenil-1H-bencimidazol-5-ol	0.328		802112 <chem>C1(=CC=CC=C1)C1=NC2=C(N1)C=CC(=C2)O</chem>
43	<b>C13H10N2O2</b>	2-(4-hidroxifenil)-3H-bencimidazol-5-ol	0.328		136244456 <chem>OC1=CC=C(C=C1)C=1NC2=C(N1)C=CC(=C2)O</chem>
44	<b>C13H10N2O2</b>	2-(3-hidroxifenil)-3H-bencimidazol-5-ol	0.328		71681235 <chem>OC=1C=C(C=CC1)C=1NC2=C(N1)C=CC(=C2)O</chem>
45	<b>C16H16N2O3</b>	2-(2,4-dimetoxifenil)-6-metoxi-1H-bencimidazol	1.764		65017829 <chem>COC1=C(C=CC(=C1)OC)C1=NC2=C(N1)C=C(C=C2)OC</chem>
46	<b>C16H16N2O3</b>	2-(2,5-dimetoxifenil)-6-metoxi-1H-bencimidazol	1.764		65018060 <chem>COC1=C(C=C(C=C1)OC)C1=NC2=C(N1)C=C(C=C2)OC</chem>
47	<b>C21H23N3O2</b>	3-metil-2-oxo-N-[2-(2-fenilbencimidazol-1-il)etil]pentanamida	8		44294077 <chem>CC(C(C(=O)NCCC1C(=NC2=C1C=CC=C2)C2=CC=CC=C2)=O)CC</chem>
48	<b>C22H26IN3O2</b>	Yoduro de N-[(E)-2-(1,3-dietil-5-metoxi-bencimidazol-1-ilio)etenil]-N-fenilacetamida	32		44294329 <chem>[I-].[C(C)[N+](I)(CN(C2=C1C=C(C=C2)OC)CC)/C=C/N(C(C)=O)C2=CC=CC=C2</chem>

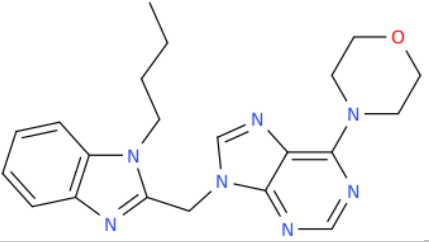
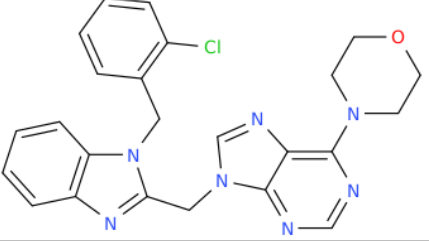
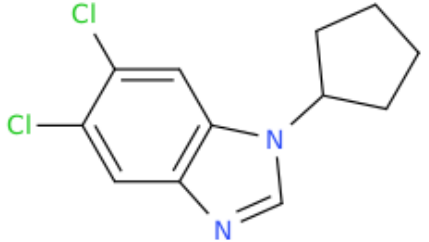
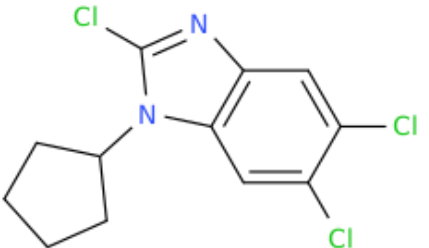
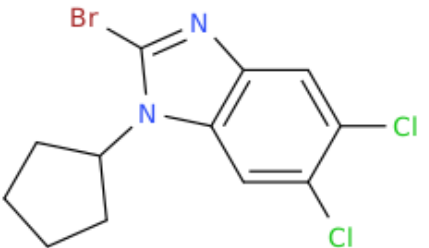
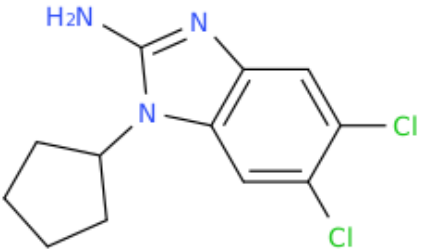
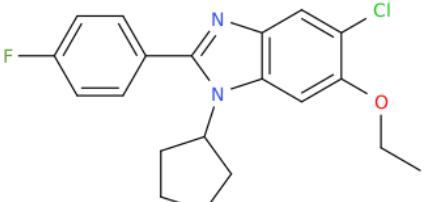
49	<b>C16H22IN3 O2</b>	Yoduro de 3-metil-N-[2-(3-metil-bencimidazol-1-ilio)etil]-2-oxopentanamida	32		44294329 <chem>[I-].CC(C(C(=O)NCC[N+]1=CNC(C=C1)C=CC=C2)C)=O)C</chem>
50	<b>C15H19N3O 2</b>	N-(2-(bencimidazol-1-il)etil)-3-metil-2-oxopentanamida	32		44294331 <chem>N1(C=NC2=C1C=CC=C2)CNC(C(C(C)C)C)=O=O</chem>
51	<b>C23H21N5O 3</b>	N-(((5S)-3-(4-((E)-2-ciano-2-(1-metilbencimidazol-2-il)etenil)fenil)-2-oxo-1,3-oxazolidin-5-il)metil)acetamida	32		44388745 <chem>C(#N)C(=C/C1=CC=C(C=C1)N1C(O[C@H](C1)CNC(C)=O)=O)C1=NC2=C(N1)C=CC=C2</chem>
52	<b>C22H19N5O 3</b>	N-(((5S)-3-(4-((E)-2-(1H-bencimidazol-2-il)-2-cianoetenil)fenil)-2-oxo-1,3-oxazolidin-5-il)metil)acetamida	32		44391733 <chem>N1(C=NC2=C1C=CC=C2)/C(=C/C2=CC=C(C=C2)N2C(O[C@H](C2)CNC(C)=O)=O)/C#N</chem>
Anexo 1 (cont.)					
53	<b>C22H19N5O 3</b>	N-(((5S)-3-(4-((E)-2-(bencimidazol-1-il)-2-cianoetenil)fenil)-2-oxo-1,3-oxazolidin-5-il)metil)acetamida – isomero	16		44391733 <chem>N1(C=NC2=C1C=CC=C2)/C(=C/C2=CC=C(C=C2)N2C(O[C@H](C2)CNC(C)=O)=O)/C#N</chem>
54	<b>C34H22BrF N6O</b>	N-(2-(4-bromo-2-fluorofenil)-3H-bencimidazol-5-il)-4-((2-fenilquinazolin-4-il)amino)bencenamida	64		162661601 <a href="https://doi.org/10.1016/j.ejmech.2020.112996">https://doi.org/10.1016/j.ejmech.2020.112996</a> <chem>BrC1=CC(=C(C=C1)C=1NC2=C(N1)C=CC(=C2)NC(C2=CC=C(C=C2)NC2=NC(=N1)C=CC=C(C2)C2=CC=C(C=C2)=O)F</chem>
55	<b>C27H18N6O 2</b>	N-(4-(6-nitro-1H-bencimidazol-2-il)fenil)-2-fenilquinazolin-4-amina	64		162675376 <a href="https://doi.org/10.1016/j.ejmech.2020.112996">https://doi.org/10.1016/j.ejmech.2020.112996</a> <chem>[N+](=O)([O-])C=1C=CC2=C(NC(=N2)C2=CC=C(C=C2)NC2=NC(=NC3=CC=CC=C3)C2=CC=CC=C2)C1</chem>
56	<b>C29H22ClN5 O2</b>	N-(4-(6-cloro-1H-bencimidazol-2-il)fenil)-2-(3,4-dimetoxifenil)quinazolin-4-amina	16		162673101 <a href="https://doi.org/10.1016/j.ejmech.2020.112996">https://doi.org/10.1016/j.ejmech.2020.112996</a> <chem>ClC1=CC=CC2=C(NC(=N2)C2=CC=C(C=C2)NC2=NC(=NC3=CC(OC)OC=C3)C2=CC=CC=C2)OC1</chem>

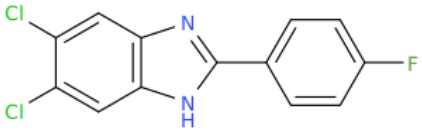
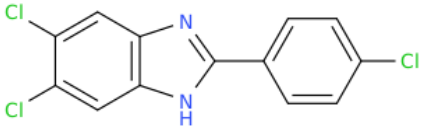
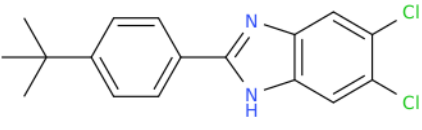
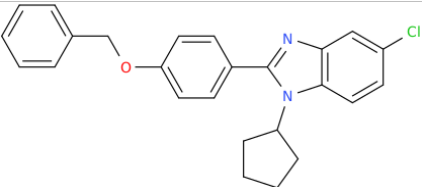
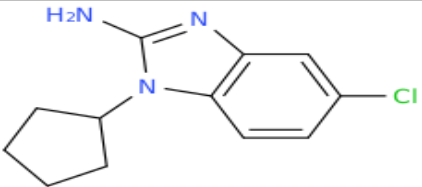
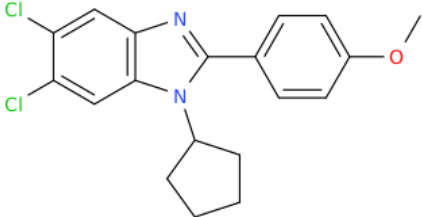
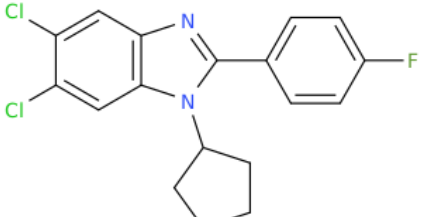
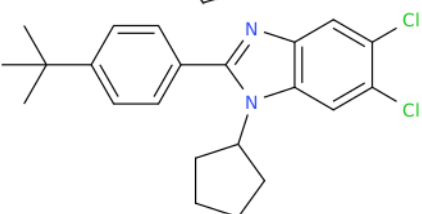
57	<b>C27H18BrN5</b> <b>5</b>	N-(4-(6-bromo-1H-benzimidazol-2-il)fenil)-2-fenilquinazolin-4-amina	64		162672240  <a href="https://doi.org/10.1016/j.ejmech.2020.112996">https://doi.org/10.1016/j.ejmech.2020.112996</a>  <chem>BrC1=CC=C(C=C1)N2=CN=C(C2)N3=NC(=N2)C2=CC=C(C=C2)NC2=NC(=NC3=CC=CC=C3)C2=CC=CC=C2)C1</chem>
58	<b>C27H19N5</b>	N-(4-(1H-benzimidazol-2-il)fenil)-2-fenilquinazolin-4-amina	8		162670354  <a href="https://doi.org/10.1016/j.ejmech.2020.112996">https://doi.org/10.1016/j.ejmech.2020.112996</a>  <chem>N1C(=NC2=C1C=CC=C2)C2=CC=C(C=C2)NC2=NC(=NC1=CC=CC=C1)C2=CC=CC=C2</chem>
59	<b>C27H18ClN5</b>	N-(4-(6-cloro-1H-benzimidazol-2-il)fenil)-2-fenilquinazolin-4-amina	8		162670345  <a href="https://doi.org/10.1016/j.ejmech.2020.112996">https://doi.org/10.1016/j.ejmech.2020.112996</a>  <chem>ClC1=CC=C(C=C1)N2=CN=C(C2)N3=NC(=N2)C2=CC=C(C=C2)NC2=NC(=NC3=CC=CC=C3)C2=CC=CC=C2)C1</chem>
60	<b>C29H22N6O4</b>	2-(3,4-dimetoxifenil)-N-(4-(6-nitro-1H-benzimidazol-2-il)fenil)quinazolin-4-amina	64		162667102  <a href="https://doi.org/10.1016/j.ejmech.2020.112996">https://doi.org/10.1016/j.ejmech.2020.112996</a>  <chem>COC1=CC=C(C=C1OC)C1=NC2=CC=CC=C2C(=N1)NC1=CC=C(C=C1)C1=NC2=C(N1)C=C(C=C2)[N+](=O)[O-]</chem>
61	<b>C37H24ClN7O</b>	N-(2-(8-cloroquinolin-3-il)-3H-benzimidazol-5-il)-4-((2-fenilquinazolin-4-il)amino)bencenamida	64		162662355  <a href="https://doi.org/10.1016/j.ejmech.2020.112996">https://doi.org/10.1016/j.ejmech.2020.112996</a>  <chem>ClC1=CC=C(C=C1)N2=CN=C(C2)N3=NC(=N2)C2=CC=C(C=C2)NC(=N3)C(=O)N4=CC=C(C=C4)N5=NC(=N5)C6=CC=C(C=C6)N5)C2=CC=CC=C2)O</chem>
62	<b>C27H17ClFN5</b>	2-(4-clorofenil)-N-(4-(6-fluoro-1H-benzimidazol-2-il)fenil)quinazolin-4-amina	4		162660639  <a href="https://doi.org/10.1016/j.ejmech.2020.112996">https://doi.org/10.1016/j.ejmech.2020.112996</a>  <chem>ClC1=CC=C(C=C1)C1=NC2=CC=CC=C2)NC(=N1)NC1=C(C=C1)C1=NC2=C(N1)C=C(C=C2)F</chem>

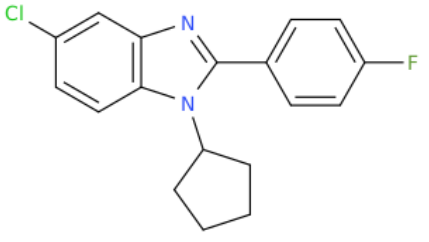
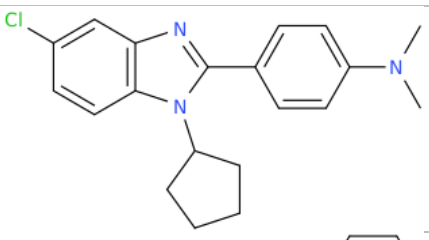
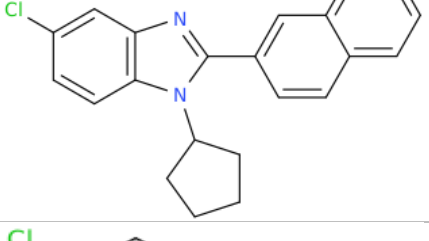
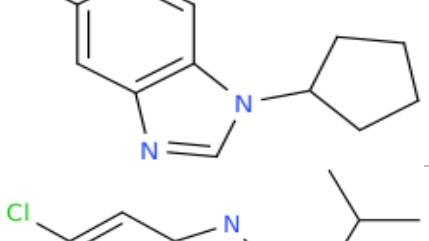
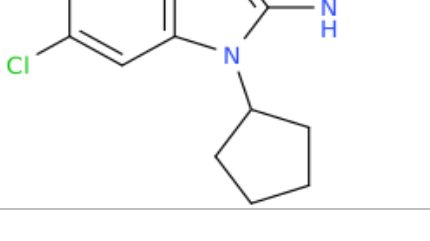
63	<b>C27H17Cl2N5</b>	N-(4-(6-cloro-1H-bencimidazol-2-il)fenil)-2-(4-clorofenil)quinazolin-4-amina	64		<p>162660424</p> <p><a href="https://doi.org/10.1016/j.ejmech.2020.112996">https://doi.org/10.1016/j.ejmech.2020.112996</a></p> <p><chem>C1=C(C=CC2=C(NC(=N2)C2=CC=C(C=C2)NC2=NC(=NC3=CC=CC=C3)C2=CC=C(C=C2)Cl)C1</chem></p>
64	<b>C29H19F3N6O</b>	4-((2-fenilquinazolin-4-il)amino)-N-(2-(trifluorometil)-3H-bencimidazol-5-il)bencenamida	32		<p>162655892</p> <p><a href="https://doi.org/10.1016/j.ejmech.2020.112996">https://doi.org/10.1016/j.ejmech.2020.112996</a></p> <p><chem>C1=(CC=CC=C1)C1=NC2=CC=CC=C2C(=N1)NC1=CC=C(C(=O)NC2=CC3=C(N=C(N3)C(F)(F)F)C=C2)C=C1</chem></p>
65	<b>C38H26N6O</b>	N-(2-(naftalen-2-il)-3H-bencimidazol-5-il)-4-((2-fenilquinazolin-4-il)amino)bencenamida	64		<p>162654802</p> <p><a href="https://doi.org/10.1016/j.ejmech.2020.112996">https://doi.org/10.1016/j.ejmech.2020.112996</a></p> <p><chem>C1=C(C=CC2=CC=CC=C2)C(=N1)NC2=CC(=C(N)C=C(C(=O)NC2=CC3=C(N=C(N3)C(F)(F)F)C=C2)C=C1</chem></p>
66	<b>C27H18FN5</b>	N-(4-(6-fluoro-1H-bencimidazol-2-il)fenil)-2-fenilquinazolin-4-amina	4		<p>162652052</p> <p><a href="https://doi.org/10.1016/j.ejmech.2020.112996">https://doi.org/10.1016/j.ejmech.2020.112996</a></p> <p><chem>FC=C1=CC2=C(NC(=N2)C2=CC=C(C=C2)NC2=NC(=NC3=CC=CC=C3)C2=CC=C(C=C2)C1</chem></p>
67	<b>C20H12F3N5S</b>	2-(tiofen-2-il)-N-(2-(trifluorometil)-3H-bencimidazol-5-il)quinazolin-4-amina	64		<p>162651931</p> <p><a href="https://doi.org/10.1016/j.ejmech.2020.112996">https://doi.org/10.1016/j.ejmech.2020.112996</a></p> <p><chem>S1C(=CC=C1)C1=NC2=CC=CC=C2C(=N1)NC1=CC=C(C(N=C(N1)C(F)(F)F)C=C2)C=C1</chem></p>
68	<b>C27H18ClN5</b>	N-(4-(1H-bencimidazol-2-il)fenil)-2-(4-clorofenil)quinazolin-4-amina - isomero	4		<p>162644165</p> <p><a href="https://doi.org/10.1016/j.ejmech.2020.112996">https://doi.org/10.1016/j.ejmech.2020.112996</a></p> <p><chem>N1C(=NC2=C1C=CC=C2)C2=CC=C(C=C2)NC2=NC(=NC1=CC=CC=C1)C2=CC=C(C=C2)Cl</chem></p>

69	<b>C26H29N3</b>	5-metil-1,3-bis(3-fenilpropil) bencimidazol-2-imina	16		122190542 <chem>CC1=CC2=C(N(C(N2CCCC2=CC=CC=C2)=N)CCCC=CC=CC=C2)C=C1</chem>
70	<b>C23H29N7O</b>	4-(9-((1-hexilbencimidazol-2-il)metil)purin-6-il)morfolina	16		145969885 <chem>C(CCCCC)N1C(=NC2=C1C=CC=C2)CN2C1=NC=NC(=C1N=C2)N2CCOCC2</chem>
71	<b>C23H29N7</b>	9-((1-hexilbencimidazol-2-il)metil)-6-(pirrolidin-1-il)purina	64		145969501 <a href="https://doi.org/10.1016/j.bmc.1.2018.03.046">https://doi.org/10.1016/j.bmc.1.2018.03.046</a> <chem>C(CCCCC)N1C(=NC2=C1C=CC=C2)CN2C1=NC=NC(=C1N=C2)N2CCCC2</chem>
72	<b>C24H31N7</b>	9-((1-hexilbencimidazol-2-il)metil)-6-(piperidin-1-il)purina	64		145969490 <a href="https://doi.org/10.1016/j.bmc.1.2018.03.046">https://doi.org/10.1016/j.bmc.1.2018.03.046</a> <chem>C(CCCCC)N1C(=NC2=C1C=CC=C2)CN2C1=NC=NC(=C1N=C2)N2CCCCC2</chem>
73	<b>C19H18F3N7O</b>	4-(9-(2-(2-(trifluorometil) bencimidazol-1-il)etil)purin-6-il)morfollina	32		145969043 <chem>FC(C1=NC2=C(N1CCN1C3=NC=NC(=C3N=C1)N1CCOCC1)C=CC=C2)(F)F</chem>

74	<b>C19H20N8O3</b>	4-(9-(2-(2-metil-5-nitrobencimidazol-1-il)etil)purin-6-il)morfollina	16		145970414  <chem>CC1=NC2=C(N1CCN1C3=NC=NC(=C3N=C1)N1CCOCC1)C=CC(=C2)[N+](=O)[O-]</chem>
75	<b>C33H49N7O</b>	4-(9-((1-hexadecilbencimidazol-2-il)metil)purin-6-il)morfollina	64		145969457  <a href="https://doi.org/10.1016/j.bmc.1.2018.03.046">https://doi.org/10.1016/j.bmc.1.2018.03.046</a>  <chem>C(CCCCCCCCCCCCCC)N1C(=NC2=C1C=CC=C2)CN2C1=NC=NC(=C1N=C2)N2CCOCC2</chem>
76	<b>C29H41N7O</b>	4-(9-((1-dodecilbencimidazol-2-il)metil)purin-6-il)morfollina	16		145968566  <chem>C(CCCCCCCCCCCC)N1C(=NC2=C1C=CC=C2)CN2C1=NC=NC(=C1N=C2)N2CCOCC2</chem>
77	<b>C20H21N7O2</b>	1-(1-(2-(6-(morfolin-4-il)purin-9-il)etil)bencimidazol-2-il)etanona	32		145968819  <chem>N1(CCOC1)C1=C2N=CN(C2=NC=N1)CCN1C(=NC2=C1C=CC=C2)C(=O)</chem>
78	<b>C19H21N7O</b>	4-(9-((1-etilbencimidazol-2-il)metil)purin-6 il)morfollina	256		145967654  <a href="https://doi.org/10.1016/j.bmc.1.2018.03.046">https://doi.org/10.1016/j.bmc.1.2018.03.046</a>  <chem>C(C)N1C(=NC2=C1C=CC=C2)CN2C1=NC=NC(=C1N=C2)N2CCOCC2</chem>
79	<b>C20H21N7O</b>	4-(9-((1-(prop-2-en-1-il)bencimidazol-2-il)metil)purin-6-il)morfollina	256		145968105  <a href="https://doi.org/10.1016/j.bmc.1.2018.03.046">https://doi.org/10.1016/j.bmc.1.2018.03.046</a>  <chem>C(C=C)N1C(=NC2=C1C=CC=C2)CN2C1=NC=NC(=C1N=C2)N2CCOCC2</chem>
80	<b>C20H23N7O</b>	4-(9-(2-(5,6----ilbencimidazol-1-il)etil)purin-6-il)morfollina	256		145970229  <a href="https://doi.org/10.1016/j.bmc.1.2018.03.046">https://doi.org/10.1016/j.bmc.1.2018.03.046</a>  <chem>N1(C=NC2=C1C=CC=C2)CN2C1=NC=NC(=C1N=C2)N2CCOCC2</chem>

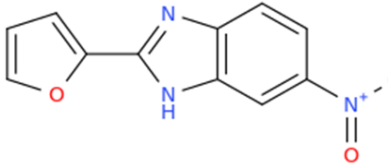
81	<b>C21H25N7O</b>	4-(9-((1-butilbencimidazol-2-il)metil)purin-6-il)morfollina	256		145968809 <a href="https://doi.org/10.1016/j.bmc.1.2018.03.046">https://doi.org/10.1016/j.bmc.1.2018.03.046</a> <chem>C(CCC)N1C(=NC2=C1C=C(C=C2)CN2C1=NC=NC(=C1N=C2)N2CCOCC2</chem>
82	<b>C24H22ClN7O</b>	4-(9-((1-((2-clorofenil)metil)bencimidazol-2-il)metil)purin-6-il)morfollina	256		145968837 <a href="https://doi.org/10.1016/j.bmc.1.2018.03.046">https://doi.org/10.1016/j.bmc.1.2018.03.046</a> <chem>C1C=C(C=CC=C1)CN1C(=NC2=C1C=CC=C2)CN2C1=NC=NC(=C1N=C2)N2CCOCC2</chem>
83	<b>C12H12Cl2N2</b>	5,6-dicloro-1-ciclopentil-1H-bencimidazol	6.25		42600434 10.1016/j.ejmech.2008.06.026 <chem>C1C=CC2=C(N(C=N2)C2CCC2)C=C1Cl</chem>
84	<b>C12H11Cl3N2</b>	2,5,6-tricloro-1-ciclopentil-1H-bencimidazol	3.12		42597043 10.1016/j.ejmech.2008.06.026 <chem>C1C=NC2=C(N1C1CCCC1)C=C(C(=C2)Cl)Cl</chem>
85	<b>C12H11BrCl2N2</b>	2-bromo-5,6-dicloro-1-ciclopentil-1H-bencimidazol	3.12		42600436 10.1016/j.ejmech.2008.06.026 <chem>BrC1=NC2=C(N1C1CCCC1)C=C(C(=C2)Cl)Cl</chem>
86	<b>C12H13Cl2N3</b>	2-amino-5,6-dicloro-1-ciclopentil-1H-bencimidazol	3.12		42597044 10.1016/j.ejmech.2008.06.026 <chem>NC1=NC2=C(N1C1CCCC1)C=C(C(=C2)Cl)Cl</chem>
87	<b>C20H20ClF N2O</b>	5-cloro-1-ciclopentil-6-etoxi-2-(4-fluorofenil) bencimidazol	6.25		42597045 10.1016/j.ejmech.2008.06.026 <chem>C1C=CC2=C(N(C(=N2)C2=CC=C(C(=C2)F)C2CCOCC2)C=C1OCC</chem>

88	<b>C13H7Cl2FN2</b>	5,6-dicloro-2-(4-fluorofenil)-1H-benzimidazol	3.12		4915231 10.1016/j.ejmech.2008.06.026 <chem>C1C=CC2=C(NC(=N2)C2=CC=C(C=C2)F)C=C1Cl</chem>
89	<b>C13H7Cl3N2</b>	5,6-dicloro-2-(4-clorofenil)-1H-benzimidazol	3.12		39158450 10.1016/j.ejmech.2008.06.026 <chem>C1C=CC2=C(NC(=N2)C2=CC=C(C=C2)Cl)C=C1Cl</chem>
90	<b>C17H16Cl2N2</b>	2-(4-terc-butilfenil)-5,6-dicloro-1H-benzimidazol	3.12		39158439 10.1016/j.ejmech.2008.06.026 <chem>C(C)(C)(C)C1=CC=C(C=C1)C2=CC=C(C=C2)N1C=NC2=C(N1)C=C(C=C2)Cl</chem>
91	<b>C25H23ClN2O</b>	2-(4-(benciloxi)fenil)-5-cloro-1-ciclopentil-1H-benzimidazol	3.12		42597291 10.1016/j.ejmech.2008.06.026 <chem>C(C1=CC=CC=C1)OC1=CC=C(C=C1)N1C=NC2=C(N1C1CCCC1)C=CC(=C2)Cl</chem>
92	<b>C12H14ClN3</b>	2-amino-5-cloro-1-ciclopentil-1H-benzimidazol	25		42600437 <chem>NC1=NC2=C(N1C1CCCC1)C=CC(=C2)Cl</chem>
93	<b>C19H18Cl2N2O</b>	5,6-dicloro-1-ciclopentil-2-(4-metoxifenil)-1H-benzimidazol	25		42597287 <chem>C1C=CC2=C(N(C(=N2)C2=CC=C(C=C2)OC)C2CCCC2)C=C1Cl</chem>
94	<b>C18H15Cl2FN2</b>	5,6-dicloro-1-ciclopentil-2-(4-fluorofenil) benzimidazol	25		42597288 <chem>C1C=CC2=C(N(C(=N2)C2=CC=C(C=C2)F)C2CCCC2)C=C1Cl</chem>
95	<b>C22H24Cl2N2</b>	2-(4-terc-butilfenil)-5,6-dicloro-1-ciclopentil-1H-benzimidazol	6,25		42597289 <chem>C(C)(C)(C)C1=CC=C(C=C1)C2=CC=C(C=C2)N1C1CCCC1C=C(C=C2)Cl</chem>

96	<b>C18H16ClF N2</b>	5-cloro-1-ciclopentil-2-(4-fluorofenil) bencimidazol	12,5		42597290 <chem>C1C=CC2=C(N(C(=N2)C2=CC=C(C=C2)F)C2CCCC2)C=C1</chem>
97	<b>C20H22ClN3</b>	5-cloro-1-ciclopentil-2-(4-(N,N-dimetilamino)fenil)-1H-bencimidazol	12,5		42597040 <chem>C1C=CC2=C(N(C(=N2)C2=CC=C(C=C2)N(C)C)C2CC(C)C)C=C1</chem>
98	<b>C22H19ClN2</b>	5-cloro-1-ciclopentil-2-(naftalen-2-il) bencimidazol	25		42597041 <chem>C1C=CC2=C(N(C(=N2)C2=CC3=CC=CC=C3C=C2)C2CCCC2)C=C1</chem>
99	<b>C12H13ClN2</b>	5-cloro-1-ciclopentil-1H-bencimidazol	50		42597042 <chem>C1C=CC2=C(N(C(=N2)C2C(C)C)C=C1</chem>
100	<b>C15H19 Cl2N3</b>	5,6-dicloro-1-ciclopentil-2-(isopropilamino)-1H-bencimidazol	50		42600435 <chem>C1C=CC2=C(N(C(=N2)NC(C)C)C2CCCC2)C=C1Cl</chem>

**Anexo 2: Derivados de Bencimidazol – Validación externa**

Nº	FÓRMULA MOLECULAR	ESCRITURA IUPAC	MIC µg/ml	ESTRUCTURA 2D	CID/DOI/SMILES
1	C16H13Cl2N2	1-alil-6-cloro-2-(2 clorofenil)-1H-bencimidazol	32		<p><a href="https://doi.org/10.1039/d2ra03491c">https://doi.org/10.1039/d2ra03491c</a></p> <chem>C(C=C)N1C(=NC2=C1C=C(C=C2)Cl)C2=C(C=CC=C2)Cl</chem>
2	C22H21ClN4O2	N,N-dimetil-4-[1-(4 clorobencil)-6-nitro-1H-bencimidazol-2-il]anilina	2		<p><a href="https://doi.org/10.1039/d2ra03491c">https://doi.org/10.1039/d2ra03491c</a></p> <chem>CN(C)C1=CC=C(C=C1)CN(C2=CC=C(C=C2)Cl)C(=C(C=C2)[N+](=O)[O-])C</chem>
3	C13H8Cl3N2	6-cloro-2-(3,4-diclorofenil)-1H-bencimidazol	4		<p>59018277</p> <p><a href="https://doi.org/10.1039/d2ra03491c">https://doi.org/10.1039/d2ra03491c</a></p> <chem>ClC1=CC=C(C=C1)C2=CC(=NC(=N2)C3=CC=C(C=C3)Cl)Cl</chem>
4	C13H6Cl2N3O2	2-(3,4-diclorofenil)-6-nitro-1H-bencimidazol	8		<p>23085583</p> <p><a href="https://doi.org/10.1039/d2ra03491c">https://doi.org/10.1039/d2ra03491c</a></p> <chem>ClC1=CC=C(C=C1)C2=CC(=NC(=N2)C3=CC=C(C=C3)Cl)C(=C(C=C2)[N+](=O)[O-])</chem>
5	C20H15ClN3O2	1-bencil-6-cloro-2-(4-nitrofenil)-1H-bencimidazol	16		<p><a href="https://doi.org/10.1039/d2ra03491c">https://doi.org/10.1039/d2ra03491c</a></p> <chem>C(C1=CC=C(C=C1)CN2=CC=C(C=C2)Cl)C(=C(C=C2)[N+](=O)[O-])C3=CC=C(C=C3)</chem>
6	C13H9Cl2N2	6-cloro-2-(4-clorofenil)-1H-bencimidazol	8		<p>376348</p> <p><a href="https://doi.org/10.1039/d2ra03491c">https://doi.org/10.1039/d2ra03491c</a></p> <chem>ClC1=CC=C(C=C1)C2=CC(=NC(=N2)C3=CC=C(C=C3)Cl)Cl</chem>
7	C13H9ClN3O2	5-cloro-2-(4-nitrofenil)-1H-bencimidazol	32		<p>376352</p> <p><a href="https://doi.org/10.1039/d2ra03491c">https://doi.org/10.1039/d2ra03491c</a></p> <chem>ClC1=CC=C(C=C1)C2=CC(=NC(=N2)C3=CC=C(C=C3)[N+](=O)[O-])C=C1</chem>
8	C11H8ClN2O	6-cloro-2-(furan-2-il)-1H-bencimidazol	4		<p>5056728</p> <p><a href="https://doi.org/10.1039/d2ra03491c">https://doi.org/10.1039/d2ra03491c</a></p> <chem>ClC1=CC=C(C=C1)C2=CC(=NC(=N2)C3=CC=CO3)C=C1</chem>

9	<b>C11H6N3 O3</b>	2-(furan-2-il)-6-nitro-1H-bencimidazol	8		<p>615240</p> <p><a href="https://doi.org/10.1039/d2ra03491c">https://doi.org/10.1039/d2ra03491c</a></p> <p><chem>O1C(=CC=C1)C1=NC2=C(N1)C=C(C=C2)[N+](=O)[O-]</chem></p>
---	-----------------------	--	---	--	---