



**UNIVERSIDAD
DEL AZUAY**

FACULTAD DE MEDICINA

**Modelo in silico con actividad contra la betalactamasa TEM-1
presente en Escherichia coli**

**Trabajo de titulación previo a la obtención de título de Médico
General**

Autoras:

**Anabel Estefanía Piña Rodríguez
Karen Melissa Torres Peñaranda**

Directora:

Dra. Miriann Mora

Cuenca-Ecuador

2023

Abstract

The bacterial resistance of Escherichia Coli is primarily determined by beta-lactamases, with TEM-1 being one of the most important ones. To minimize the limitations of traditional methods, the use of in silico methods such as QSAR has been proposed for the search for new drugs. In this study, we worked with 207 molecules collected from BindingDB, each of them having 977 descriptors calculated using Dragon 7.0. By applying genetic algorithms, the best descriptors adapted to classification methods (kNN, N3, BNN, CART) were selected, and two activity classes were created with a threshold set at the median of the IC50 values. A total of 102 molecules were classified as "highly active" and 105 molecules as "less active," resulting in models with good predictive ability. The results demonstrate that the QSAR classification models (kNN, N3, BNN, and CART) are capable of distinguishing between active and less active molecules against the TEM-1 beta-lactamase of Escherichia coli, with high predictive power both internally and externally.

Keywords: Escherichia Coli, TEM-1, beta-lactamase, QSAR, bacterial resistance.

Translated by:



Firmado electrónicamente por:
MIRIANN ALEXANDRA
MORA VERDUGO

Karen Torres

Anabel Piña



Resumen

La resistencia bacteriana de *Escherichia Coli* está dada principalmente por betalactamasas, siendo una de las más importantes TEM-1. Para minimizar los obstáculos de métodos tradicionales, se ha propuesto el uso de métodos *in silico* como QSAR para la búsqueda de nuevos medicamentos. En este estudio, se trabajó sobre 207 moléculas recolectadas de BindingDB, cada una con 977 descriptores calculados mediante Dragon 7.0. Aplicando algoritmos genéticos se seleccionaron los mejores descriptores adaptados a métodos de clasificación (kNN, N3, BNN, CART) para lo que se crearon 2 clases de actividad cuyo umbral es la mediana de los valores de IC50 y se clasificaron 102 moléculas como “muy activas” y 105 moléculas como “menos activas”, con los que se obtuvieron modelos con buena capacidad predictiva. Los resultados demuestran que los modelos QSAR de clasificación (kNN, N3, BNN y CART) son capaces de distinguir entre moléculas activas y menos activas contra la betalactamasa TEM-1 de *Escherichia coli*, con un alto poder predictivo tanto interna como externamente.

Palabras clave: *Escherichia Coli*, TEM-1, betalactamasa, QSAR, resistencia bacteriana.

Introducción

La resistencia bacteriana es la capacidad de las bacterias de evadir los mecanismos que utilizan los antibióticos en su contra (1). Este es un problema alarmante ya que el ritmo al que se están descubriendo nuevos antibióticos se ha ralentizado mientras que su mal uso y uso desmesurado continúa en aumento. Un reporte estima que, para el 2050, la resistencia antibiótica sería la causante de 10 millones de muertes anuales. Las consecuencias no solo se limitan a muerte y discapacidad sino que se convierten en una carga financiera para el individuo afectado y el estado pues requerirán mayor tiempo de hospitalización y tratamientos más costosos (2,3).

Entre las bacterias de mayor interés clínico a nivel mundial y local por su creciente resistencia antibiótica, destaca la *Escherichia coli*. En Ecuador, durante el periodo 2014-2018, el Instituto Nacional de Investigación en Salud Pública (INSPI) indicó que el microorganismo más reportado aislado a partir de los servicios hospitalarios es justamente la *Escherichia coli* (4), por lo que, esta bacteria está considerada en el Sistema de Vigilancia del país.

La *Escherichia coli* es una bacteria Gram negativa en forma de bastón perteneciente a la familia Enterobacteriaceae, es aeróbica facultativa y predominante en el tracto digestivo humano. Este microorganismo se puede transmitir de una persona a otra por medio de los alimentos contaminados o vía fecal-oral, convirtiéndose en el patógeno más comúnmente responsable de gastroenteritis e infecciones del tracto urinario (5,6). Existen seis serotipos causantes de diarrea de *Escherichia coli*: enteropatógena, enteroagregativa, enterotoxigénica, enteroinvasiva, productora de toxina Shiga y de adherencia difusa (7). Sin embargo, también pueden causar infecciones extraintestinales como síndrome hemolítico urémico, meningitis e infecciones respiratorias (8). La *Escherichia coli*, al igual que la mayoría de bacterias Gram negativas, tiene la capacidad de adquirir y propagar genes que codifican mecanismos de resistencia (9), siendo su principal instrumento la producción de betalactamasas

(10–12).

Las betalactamasas son enzimas bacterianas que se unen al antibiótico betalactámico rompiendo su enlace amida del anillo de azetidiona, agregando una molécula de agua lo que hace que el antibiótico sea incapaz de cumplir su función (1,13). Las betalactamasas capaces de hidrolizar los antibióticos de espectro extendido como penicilinas, cefalosporinas, entre otras, se les denomina betalactamasas de espectro extendido (BLEE); entre ellas destacan TEM-1 y CTX-M comúnmente encontradas en especies de Enterobacteriaceae (13,14).

TEM-1 fue descrito por primera vez en 1960, aislado de *Escherichia coli* proveniente de un cultivo sanguíneo de una paciente griega llamada Temoneira, de donde adopta su nombre (15). Sin embargo, en la misma década, fue identificada en cepas de *K. pneumoniae* resistentes a las penicilinas (16). Esta betalactamasa ha sido identificada como la principal causa de resistencia bacteriana en *Escherichia coli*, predominantemente a la ampicilina, debido a que los genes que la codifican han tenido mutaciones que los convierten en betalactamasas de espectro extendido (9,13,17).

Los fármacos betalactámicos son los antibióticos prescritos con mayor frecuencia en el ámbito hospitalario y ambulatorio debido a su seguridad, eficacia y espectro de actividad. Los principales betalactámicos son: penicilinas, cefalosporinas, carbapenémicos y monobactámicos, todos los cuales incluyen en su estructura un anillo de azetidiona de cuatro miembros (17,18). Por otro lado, los inhibidores de betalactamasas son estructuralmente similares a las penicilinas, esto les permite unirse irreversiblemente a las betalactamasas actuando como sustratos suicidas, manteniéndolas inactivas. Por lo que, una buena estrategia para superar la resistencia por betalactamasas, ha sido la unión de un fármaco betalactámico y un inhibidor de betalactamasas, puesto que, el inhibidor activará la acción de la enzima permitiendo que las penicilinas cumplan con su función (13,17).

Actualmente, los inhibidores de betalactamasas más frecuentemente utilizados son: ácido clavulánico, tazobactam y sulbactam, ya que poseen actividad inhibitoria conocida en contra de betalactamasas con altos índices de resistencia (TEM, SHV y CTX-M) (19) también llamadas betalactamasas de espectro extendido (BLEE). Estas son resistentes a antibióticos betalactámicos excepto carbapenémicos, cefamicinas, e inhibidores de las betalactamasas (20). Sin embargo, el primer caso de resistencia a un betalactámico/inhibidor de betalactamasas (amoxicilina/ácido clavulánico) en *Escherichia coli* fue en 1989, apenas 8 años tras su comercialización (5).

A lo largo de la historia se han buscado nuevos fármacos activos contra las bacterias con una resistencia creciente a antibióticos. No obstante, la baja eficacia, el alto costo de producción y los datos extensos y complejos de los ensayos clínicos e *in vitro* imponen un obstáculo que afectan el diseño y el descubrimiento de fármacos. Por tanto, el aprendizaje automático y la inteligencia artificial han modernizado esta área, integrando algoritmos que permiten la síntesis de péptidos, la detección virtual basada en estructuras, la detección virtual basada en ligandos, la predicción de toxicidad, el control y la liberación de fármacos, el modelado de farmacóforos, la relación cuantitativa estructura-actividad, el reposicionamiento de fármacos y actividad fisicoquímica (21).

Desde aproximadamente 1962, se han aplicado modelos computacionales para el descubrimiento de fármacos (22). Uno de estos modelos es la relación cuantitativa estructura-actividad (QSAR) el cual condensa estadística, bioinformática, métodos computacionales relacionados con el diseño y visualización espacial virtual de moléculas (también llamado modelo *in silico*) y descriptores (propiedades fisicoquímicas moleculares). Los modelos QSAR tienen como objetivo hacer una predicción de actividad biológica de ligandos no existentes para la búsqueda de nuevos fármacos (23). Estos avances permiten eliminar los desafíos y obstáculos del diseño y desarrollo de medicamentos tradicionales (21).

Metodología

Este es un estudio computacional que utiliza el modelado *in silico* para investigar la relación cuantitativa estructura-actividad de los inhibidores de TEM-1.

Construcción de la base de datos y depuración

Se recolectaron 316 moléculas citadas en la base de datos Binding DB (24), cada una con su actividad expresada en CI50 que representa la mitad de la concentración inhibitoria máxima. Sobre estas se hizo un control visual para eliminar moléculas repetidas y aquellas en las que la actividad no estaba expresada en nanomoles (nM), por lo tanto, se trabajó sobre 207 moléculas y se buscó sus identificadores: el número de registro CAS (Chemical Abstracts Service), InChIKey (International Chemical Identifier) y SMILES (Simplified Molecular Input Line Entry System) en la base de datos PubChem (25).

Cálculo de descriptores

Después de la depuración, se procedió a calcular los descriptores moleculares de los candidatos seleccionados; esta operación se realizó usando el programa Dragon 7.0. (26). Este cálculo abarcó hasta descriptores bidimensionales (0 D, 1 D y 2 D) y se obtuvieron aproximadamente 5000 descriptores. Posteriormente, se llevó a cabo una depuración para eliminar aquellos descriptores constantes o casi constantes. Los descriptores fueron identificados por pares con un coeficiente de correlación (R^2) superior a 0.85 y se retuvo únicamente uno de los dos descriptores en cada par. En total se obtuvo una base de datos de 207 moléculas cada una representada por 977 descriptores.

Debido a que los descriptores tienen unidades de medida diferentes para usarlos con métodos multivariantes se deben transformar en una misma unidad de medida. El escalado de datos permite homogeneizar todas las variables de estudio, sin importar sus unidades de medida (27).

En este estudio, se empleó la técnica de autoescalado, que implica centrar los datos y luego normalizarlos para obtener una varianza unitaria. Las variables

autoescaladas presentan una media de cero y una desviación estándar igual a uno, lo que sugiere que los datos son relativamente homogéneos (27).

Modelos de regresión

Para utilizar los métodos de regresión correctamente es necesario que las moléculas (objetos) sean en número mayor que las variables descriptoras (descriptores moleculares). La base de datos contaba con 207 moléculas y 977 descriptores por lo que, en la selección de las variables descriptoras, es necesario retener solamente aquellas que son verdaderamente útiles para la regresión.

En este trabajo, para realizar el método de regresión se han utilizado métodos acoplados con los algoritmos genéticos (GAs) (28). El método de algoritmos genéticos se basa en la teoría de la evolución: en un entorno biológico, los "mejores" individuos tienen una mayor probabilidad de sobrevivir y transmitir sus genes a través de la reproducción. El cruce de dos individuos "buenos" resulta en la combinación de sus genes, lo que puede dar lugar a una descendencia "mejorada". Con este método, se genera una población de cromosomas de forma aleatoria. En el contexto de la selección de características, cada cromosoma es un vector de filas que contiene tantos elementos como variables, y cada elemento se codifica con 1 si la variable correspondiente está seleccionada, y con 0 si no lo está. Esta población inicialmente aleatoria de cromosomas evoluciona a medida que se seleccionan los individuos más aptos y se produce un cruce (es decir, una combinación de sus elementos) o también conocido como crossover, junto con mutaciones ocasionales. Este proceso se repite un número determinado de veces (29).

OLS

La regresión de mínimos cuadrados ordinarios (OLS) es un enfoque tradicional utilizado para crear un modelo multivariable que establece una relación entre n objetos descritos por p variables y una variable de respuesta y . Este método se fundamenta en la minimización de la siguiente fórmula.

En este caso en estudio, se debe necesariamente reducir el número de variables descriptoras por el limitado número de objetos sobre los cuales construir el modelo. Para este fin, se utilizaron algoritmos genéticos (GAs) de los descriptores más significativos (28,30).

RNN

La regresión RNN es una técnica no paramétrica que busca aproximar de forma intuitiva la relación entre variables independientes y un resultado continuo, promediando las observaciones en un vecindario determinado. El analista tiene la tarea de establecer el tamaño de dicho vecindario, o puede ser seleccionado mediante validación cruzada (abordaremos esto más adelante) para encontrar el tamaño que minimice el error cuadrático medio (30).

Validación del método

Para que el método de regresión sea aplicable es necesario que sea predictivo, por esta razón, es preciso construir el modelo sobre un subconjunto de datos (training set) y validando sobre un segundo subconjunto de datos (test set) cuyas moléculas no se utilizan para la construcción del modelo. En el training set se extrajeron 152 moléculas al azar (70%) y en el test set 62 moléculas de la misma manera (30%).

Modelos de clasificación

Debido a que el CI50 es una variable continua, se optó por crear 2 clases de actividad cuyo umbral es la mediana de los valores de CI50.

Creación de las dos clases de actividad

Se subdividió la matriz de los descriptores en dos subconjuntos en función del valor de su CI50, moléculas de alta actividad o con CI50 bajo y moléculas poco activas o con CI50 alto. La clase activa se denominó con 1 y la baja con 2. El umbral de pertenencia a una clase se fijó determinando la mediana de todos los valores de CI50 (300 nM/L) como resultado de esto se obtuvieron 102 moléculas

en clase 1 y 105 moléculas en clase 2, logrando dos clases suficientemente equilibradas.

Definición del conjunto de calibración y conjunto de validación externa

Una vez realizada la división del C150 en 2 clases, se procedió a determinar el conjunto de moléculas de calibración (training set) y el conjunto de moléculas de predicción (test set). Para este efecto, se utilizó un algoritmo implementado en Matlab en el cual se programó para que en el conjunto de calibración esté el 60% de las moléculas (125 moléculas) y en el conjunto de predicción el 40% (82 moléculas), respetando la numerosidad de las clases.

De los 977 descriptores, solo algunos de ellos serán importantes para un modelo óptimo, por lo que se debe seleccionar aquellos que permitan obtener modelos con buena predicción. En este estudio se utilizó algoritmos genéticos (GAs), el mismo que se describió previamente en los métodos de regresión.

kNN

Sobre el conjunto de calibración se procedió a calcular los modelos; se eligió el método de los k-vecinos más cercanos kNN como método base ya que es robusto y no depende de la distribución estadística de los datos. Es un enfoque no paramétrico que clasifica un objeto basándose en la información de sus k-vecinos más cercanos en un espacio multidimensional definido por los descriptores moleculares. Como prerequisites; se necesita la definición de la distancia alrededor de estas moléculas en estudio, entre la cual se consideran las clases de las moléculas vecinas. Para calcular las distancias entre los compuestos, se utilizó la métrica euclidiana después de autoescalar los descriptores moleculares. El número de moléculas que dan el valor óptimo de modelización en predicción se obtiene aplicando el método con diferentes valores de k . La optimización del valor k se logró mediante la minimización del NERcv (Non-error-rate o tasa de aciertos). El modelo está formado por las moléculas del training set, el k óptimo y el tipo de distancia utilizada (31).

N3

La técnica N-vecinos más cercanos (N3, N-Nearest Neighbours) es un método de clasificación que se basa en el enfoque del método kNN. Este enfoque se utiliza para clasificar un objeto determinado al considerar la clase a la que pertenecen los vecinos más similares, es decir, las drogas. Para clasificar un objeto específico, el clasificador N3 tiene en cuenta la contribución de todos los objetos restantes en el espacio multidimensional, ponderada según su proximidad al objeto que se está clasificando. La contribución de cada vecino se pondera en función del rango de similitud de la molécula, y este peso se modula mediante un parámetro alfa que se optimiza.

Para el método N3 se usó la distancia de Mahalanobis promedio, y el número óptimo del parámetro alfa se optimizó para lograr tasas de aciertos más altas (NER cv) con grupos de validación cruzada de cinco veces divididos en ventanas venecianas (32).

BNN

El método de los vecinos más cercanos agrupados (BNN, Binned Nearest Neighbors) utiliza un enfoque en el cual se predice la respuesta del objetivo considerando un número variable de k de vecinos y utilizando el criterio del voto mayoritario. La idea principal de este método es tomar en cuenta para la predicción a aquellos vecinos que tengan una similitud significativa y comparable con el objetivo. Para seleccionar los vecinos más similares, se definen previamente intervalos de similitud (también conocidos como contenedores) y los vecinos se distribuyen en estos intervalos en función de su similitud con el objetivo.

Todos los vecinos que caen en el contenedor con la mayor similitud se consideran para realizar la predicción. El parámetro óptimo utilizado para definir los umbrales de los contenedores se busca dentro de un rango predefinido, y se selecciona el valor que resulta en el error de clasificación más bajo utilizando un protocolo de validación (32).

Árboles de regresión y clasificación (CART)

El árbol de decisión se construye a partir del conjunto de entrenamiento mediante una búsqueda exhaustiva de la mejor división (nodo raíz) para reducir el error de clasificación. Este método se considera univariante, ya que analiza un descriptor a la vez. Una vez que se encuentra la mejor división, los datos se dividen en dos subconjuntos mutuamente excluyentes (nodos hijos), que representan las posibles trayectorias para la división de los datos. Este proceso continúa hasta que el error asociado a un nodo cae por debajo de un umbral o nivel de tolerancia establecido (31).

Validación del modelo

Para cada modelo se debe determinar su calidad global y la calidad de clasificación de cada clase, tanto en calibración como en predicción. Todos estos parámetros se calculan partiendo de una matriz de confusión (33)

Parámetros primarios

Precisión: Capacidad del modelo para evitar la inclusión de objetos de otras clases en la clase g-ésima.

Sensibilidad: Capacidad del modelo para clasificar correctamente los objetos que pertenecen a la clase g-ésima.

Parámetros globales

Tasa de aciertos (non-error rate; NER): Se calcula como el promedio de las sensibilidades de las clases.

Resultados

Modelos de regresión

Regresión mediante OLS

Este modelo no es de buena calidad en cuanto R^2 tanto en calibración como en validación interna presenta valores deficientes (Tabla 1). Mientras que el R^2 de validación externa fue de 0,862 lo cual nos indica que este resultado se debe al azar.

Tabla 1: Resultados de modelo de regresión por método OLS.

Medida	Calibración	Validación Interna	Validación Externa
RMSE	127176.232	129607.773	56370.8939
R2	0.3	0.273	0.862

Regresión mediante RNN

La aplicación del método RNN acoplado a algoritmos genéticos ha permitido construir un modelo con 19 descriptores. La calidad del modelo obtenido se presenta en la siguiente tabla (tabla 2):

Tabla 2: Resultados modelo de regresión por método RNN

Medida	Calibración	Validación Interna	Validación Externa
RMSE	756.446.836	749.310.488	288.490.140
R2	0.752	0.757	0.964

Los valores de la Tabla 2 indican que el modelo es bueno en cuanto a calibración y validación interna y externa ya que se obtuvo un R2 cercano a la unidad y lo mismo ocurre en predicción, pero al analizar la gráfica CI50 calculado vs CI50 experimental (Fig. 1), refleja la presencia de un efecto de palanca de 4 moléculas del conjunto de calibración, por lo tanto, este modelo de regresión es poco confiable y difícilmente aplicable a otros casos.

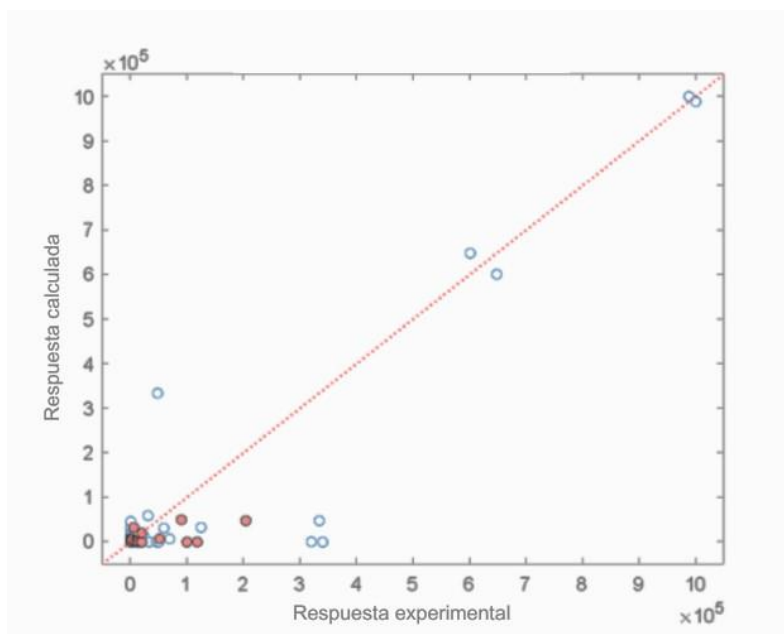


Figura 1: Comparación de respuesta experimental vs respuesta calculada.

Los modelos de regresión obtenidos en este estudio no permiten predecir con seguridad la actividad de nuevas moléculas, por esta razón se ha optado por utilizar modelos de clasificación.

Modelo de clasificación

Para realizar el modelo de clasificación se utilizaron los descriptores previamente calculados mediante Dragon 7.0 y se seleccionaron los realmente útiles mediante algoritmos genéticos (GAs) (Tabla 3):

Tabla 3: Descriptores usados en método de clasificación (34)

Descriptor	Definición
nSK	Número de átomos no hidrogenoides
Sp	Suma de las polarizabilidades atómicas (escaladas sobre carbono)
Se	Suma de electronegatividades de sanderson (escaladas sobre carbono)
GD	Densidad del grafo molecular
Mp	Promedio de las polarizabilidades atómicas (escalas sobre carbono)

Si	Suma de los primeros potenciales de ionización (escalado sobre carbono)
AMW	Peso molecular promedio

Tabla 4: Resultados de los modelos de clasificación

Método	Parámetro	Parámetros globales			Parámetro de calidad de la predicción externa			
		N.E.R _{cal}	N.E.R _{cv}	N.E.R _{pred}	Precisión Clase 1	Precisión Clase 2	Sensibilidad Clase 1	Sensibilidad Clase 2
kNN	k=8	0.752	0.792	0.820	0.755	0.909	0.925	0.714
N3	$\alpha = 1.25$	0.760	0.776	0.818	0.778	0.865	0.875	0.762
BNN	$\alpha = 1.25$	0.752	0.792	0.818	0.791	0.846	0.850	0.786
CART	-	0.769	0.712	0.790	0.770	0.820	0.820	0.760

Los resultados obtenidos con el método kNN tanto en calibración (0.752) como en predicción (0.85) son significativos en cuanto son cercanos a uno. Para este método se usaron 7 descriptores moleculares detallados en la tabla 3. Sin embargo, con este número de descriptores es problemático identificar un mecanismo de acción de las moléculas con respecto a la betalactamasa.

Utilizando los mismos descriptores con los métodos BNN y N3 se han obtenido resultados similares (Tabla 4). Asimismo, al aplicar sobre el mismo conjunto de calibración seleccionado el método CART, se obtuvieron resultados comparables con los tres métodos anteriores pero con un solo descriptor: VE1sign_B(s) definido como la suma de los coeficientes del más alto autovector de la matriz de Burden pesada por el potencial de ionización de los átomos. El

árbol de decisión de CART (Figura 2) indica que las moléculas con valores del descriptor VE1sign_B(s) menores de 0.092 y mayores de 0.173 pertenecen a la clase 1 (moléculas activas), mientras que, los que tienen valores entre 0.092 y 0.173 pertenecen a la clase 2 (menos activas).

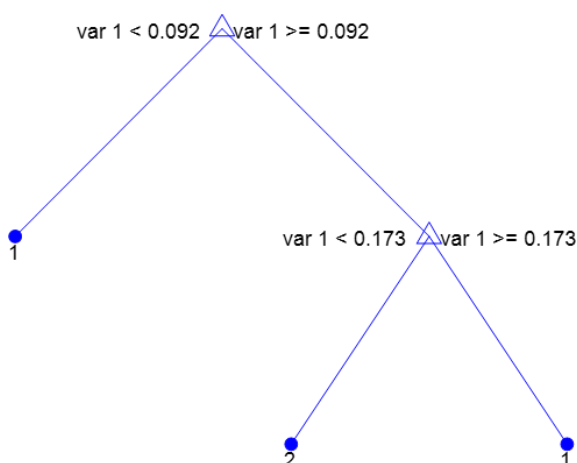


Figura 2: Árbol de decisión CART.

Este descriptor está relacionado con el potencial de ionización de los átomos de la molécula y también con la “complejidad” de la misma, se puede pensar que las moléculas de la clase 1-1 con VE1sign_B(s) menor a 0.099 (clase 1.1) tengan menos átomos conectados y menos átomos con potencial de ionización alto como se puede ver en el caso de la molécula “BDBM17817” que pertenece a la clase 1-1 cuya estructura se presenta en la figura 3:

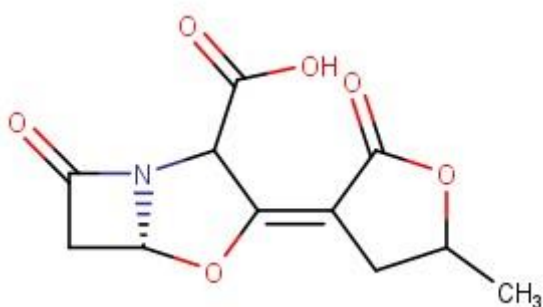


Figura 3: Estructura de la molécula “BDBM17817” (clase 1-1) valor del descriptor VE1sign_B(s): 0.07

Las moléculas de la clase 1 con el descriptor mayor que 0.173 (clase 1.2) tienen más átomos conectados y más átomos con potencial de ionización alto como en el caso de la molécula “BDBM50076794” cuya estructura se presenta en figura

4:

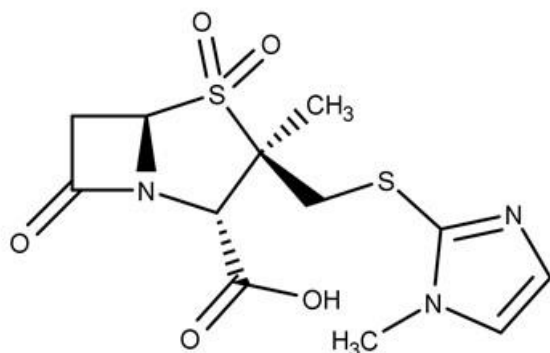


Figura 4: Estructura de la molécula “BDBM50076766” (clase 1-2) valor del descriptor $VE1sign_B(s)$: 0.185

Las moléculas menos activas tienen un nivel intermedio de este descriptor y también un nivel intermedio de átomos conectados y potencial de ionización, como en el caso de la molécula “BDBM17817” cuya estructura se presenta en figura 5:

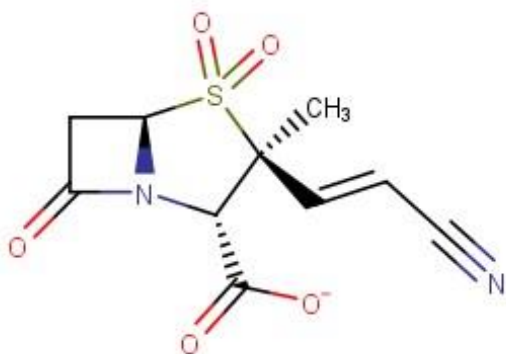


Figura 5: Estructura de la molécula “BDBM17817” (clase 2) valor del descriptor $VE1sign_B(s)$: 0.148

Discusión

El uso excesivo de betalactámico/inhibidor de betalactamasa contribuye a la resistencia y limita su eficacia. Sin embargo, no es solo la mayor resistencia contra estos fármacos lo que limita su uso terapéutico sino también su reducido espectro de acción, ya que estos inhibidores no tienen ningún efecto contra betalactamasas tipo AmpC y metalobetalactamasas (MBL) también presentes en *Escherichia coli*, por tanto, se han buscado moléculas alternativas, pero la mayoría se encuentran en fase experimental (35).

El estudio SMART (Study for Monitoring Antimicrobial Resistance Trends) en

España, demuestra la limitada actividad de antibióticos como amoxicilina/clavulanato y piperacilina/tazobactam ante organismos productores de BLEE con un aumento en las tasas de resistencia media en *Escherichia coli* causante de infección intraabdominal, desde el 17,6% hasta el 18,51% para amoxicilina/ácido clavulánico y desde el 7,7% hasta el 10% para piperacilina/tazobactam, durante los periodos 2002-2010 y 2016-2017 (36). Otro estudio realizado en Estados Unidos entre 2006 y 2016, demuestra un patrón de resistencia a piperacilina/tazobactam en *Escherichia coli* y *Klebsiella Pneumoniae* desde aproximadamente un 0,5% hasta un 1,4% (37). Igualmente, en Inglaterra, desde 2011 hasta 2015, se ha reportado un aumento en las tasas de resistencia a piperacilina/tazobactam en las bacteriemias causadas por *Escherichia coli* desde un 8,5% hasta un 11,7% y para amoxicilina/ácido clavulánico desde el 31% al 42% (5).

Se ha demostrado que *Escherichia coli* uropatógena, la cual es causante de la mayoría de infecciones urinarias a nivel mundial, que tiene la capacidad de producir enzimas BLEE evidenciaron mayor tasas de resistencia a betalactámicos, aminoglucósidos y quinolonas en comparación con las cepas no productoras de BLEE (20). Por lo tanto, la Sociedad de Enfermedades Infecciosas de América (IDSA) en 2011, difundió una recomendación para sugerir el uso de trimetoprim-sulfametoxazol (cotrimoxazol), nitrofurantoína, fosfomicina o pivmecilinam como tratamientos para infecciones del tracto urinario no complicadas, dejando de lado el uso de betalactámicos (38).

En nuestro estudio se clasificaron 102 moléculas como “muy activas” contra TEM-1 de *Escherichia coli*, las cuales podrían ser usadas para búsqueda de nuevos fármacos. Esta betalactamasa fue identificada como el principal mecanismo de resistencia de *Escherichia coli*, especialmente a la ampicilina. Inclusive el gen que codifica la betalactamasa TEM-1 (*bla*TEM-1) fue identificado como el de mayor prevalencia en las fuentes de agua de consumo humano, debido a que las fuentes de agua urbana reciben un gran número de microorganismos por agua residual proveniente de hospitales, agrícola,

industrias, entre otras fuentes (39). Sin embargo, la resistencia a *Escherichia coli* no está dada únicamente por betalactamasa TEM-1, si no que, actualmente coexiste con otras betalactamasas, entre ellas, CTX-M que se ha distribuido ampliamente entre las enterobacterias (especialmente *Escherichia coli*) (5); además que esta bacteria tiene características como fimbrias, producción de toxinas, cápsula, adhesinas, endotoxina y flagelo que dificulta la elección de un fármaco adecuado para el tratamiento de sus infecciones (20).

Además de los inhibidores de betalactamasa sintéticos, existen hierbas medicinales que contienen muchos compuestos que simulan su acción. Una especie interesante desde este punto de vista es *Chromolaena Odorata* que se encuentra en algunas zonas de África, en un estudio se usó una variante de QSAR asociada a Docking generando modelos de regresión sobre 30 moléculas (mixtas, naturales y sintéticas) y 497 descriptores, en donde señala que los compuestos de esta planta presentan propiedades farmacocinéticas tales como: absorción, distribución, metabolismo, eliminación y afinidad mejores que antibióticos estándares (avibactam, tazobactam y sulbactam) (39). Adicionalmente, en un estudio *in silico*, se buscó analizar la capacidad inhibitoria de catequinas vegetales contra la betalactamasa TEM-1 de *Escherichia coli*, en donde se demostró su actividad especialmente en aquellas encontradas en el ácido gálico (40). Por otro lado, un estudio mediante el método de detección de objetivos virtuales, demostró actividad inhibitoria del ácido rosmarínico, un polifenol presente en muchas plantas como la albahaca, orégano y romero; y otros compuestos relacionados contra metalo y serino-betalactamasas (VIM-2, NDM-1 y TEM-1) (41).

En este estudio, el modelo por el método CART es más simple que los demás métodos propuestos, dado que el valor calculado por el descriptor VE1sign_B(s) puede establecer una predicción de actividad solo controlando 2 valores del árbol de decisión (VE1sign_B(s) < 0.092 y >0.173: clase 1 o moléculas activas, VE1sign_B(s) entre 0.092 y 0.173: clase 2 o menos activas); brindando información sobre si una molécula nueva es más o menos activa contra la

betalactamasa TEM-1 de Escherichia Coli. Todos estos hallazgos podrían ser útiles tanto para el estudio como para combatir la resistencia antibiótica principalmente a betalactámicos.

Limitaciones

La limitación principal de los modelos obtenidos en este trabajo es que están contruidos sobre una base de datos de 316 moléculas, por tanto, los modelos pueden producir predicciones, pero si hubiese moléculas completamente diferentes o nuevas podrían dar predicciones equivocadas.

Otra limitación de los modelos es que están dirigidos únicamente contra la betalactamasa TEM-1 de Escherichia coli. Sin embargo, esta bacteria tiene otras betalactamasas para las cuales no es seguro que el modelo pueda funcionar. Además, existe una limitación bibliográfica sobre investigación con modelos QSAR sobre actividad de betalactamasas.

Conclusión

Se conoce que la betalactamasa TEM-1 es una enzima que tiene la capacidad de hidrolizar la penicilina y sus derivados. En este estudio se demuestra que los modelos QSAR de clasificación (kNN, N3, BNN y CART) tienen la capacidad de diferenciar las moléculas activas de las menos activas contra la betalactamasa TEM-1 presente en Escherichia coli. Además, se demuestra que los cuatro modelos tienen buena capacidad de ajuste, estabilidad interna y predictividad, contribuyendo así a proporcionar herramientas para que el estudio de estas moléculas sea más rápido.

Agradecimiento

Agradecemos a nuestras respectivas familias por su apoyo y guía incondicional, a nuestra directora y asesor metodológico por su soporte y tiempo brindado, sin ellos este trabajo no habría sido posible.

Referencias

1. Bush K. Epidemiology of β -Lactamase-Producing Pathogens. 2020;
2. Neill JO'. Antimicrobial Resistance: Tackling a crisis for the health and wealth of nations The Review on Antimicrobial Resistance Chaired. 2014;(December).
3. Collaborators AR. Articles Global burden of bacterial antimicrobial resistance in 2019 : a systematic analysis. 2022;399.
4. Dirección Nacional de Vigilancia Epidemiológica. Reporte de datos de resistencia a los antimicrobianos en Ecuador. Minist Salud Publica [Internet]. 2018;2(1):1-10. Disponible en: https://www.salud.gob.ec/wp-content/uploads/2019/08/gaceta_ram2018.pdf
5. Rodríguez Villodres Á, Younes S, Lepe Jiménez JA, Pachón Díaz J. Bacteriemias por Escherichia coli: análisis clínico-epidemiológico, factores de patogenicidad y mecanismos de resistencias a betalactámicos/inhibidores de bectalactamasas. 2019;1-209. Disponible en: [https://idus.us.es/bitstream/handle/11441/93377/Copia sin artículos Tesis Ángel Rodríguez Villodres.pdf?sequence=1&isAllowed=y&fbclid=IwAR1TV7yjPoO-ED6OzMzK-V3CwjW0CuhTiiLwBHIR1BQFIQ2MxIYAFhIXBjA](https://idus.us.es/bitstream/handle/11441/93377/Copia%20sin%20art%C3%ADculos%20Tesis%20Rodr%C3%ADguez%20Villodres.pdf?sequence=1&isAllowed=y&fbclid=IwAR1TV7yjPoO-ED6OzMzK-V3CwjW0CuhTiiLwBHIR1BQFIQ2MxIYAFhIXBjA)
6. Giri S, Kudva V, Shetty K, Shetty V. Prevalence and characterization of extended-spectrum β -lactamase-producing antibiotic-resistant escherichia coli and klebsiella pneumoniae in ready-to-eat street foods. Antibiotics. 1 de julio de 2021;10(7).
7. Jang J, Hur H-G, Sadowsky MJ, Byappanahalli MN, Yan T, Ishii S. Environmental Escherichia coli: ecology and public health implications-a review. 2017; Disponible en: <https://sfamjournals.onlinelibrary.wiley.com/doi/10.1111/jam.13468>
8. Yair Y, Gophna U. Pandemic bacteremic escherichia coli strains: Evolution and emergence of drug-resistant pathogens. En: Current Topics in Microbiology and Immunology. Springer Verlag; 2018. p. 163-80.
9. Bakr A, Id M, Abdullateef K, Id A. Phenotypic and genotypic detection of extended spectrum beta lactamase enzyme in Klebsiella pneumoniae.

- PLoS One [Internet]. 29 de septiembre de 2022 [citado 4 de noviembre de 2022];17(9):e0267221. Disponible en: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0267221>
10. Meini S, Tascini C, Cei M, Sozio E, Rossolini GM. AmpC β -lactamase-producing Enterobacterales: what a clinician should know. Vol. 47, *Infection*. Urban und Vogel GmbH; 2019. p. 363-75.
 11. Christaki E, Marcou M, Tofarides A. Antimicrobial Resistance in Bacteria: Mechanisms, Evolution, and Persistence. *J Mol Evol* [Internet]. 2020;88(1):26-40. Disponible en: <https://doi.org/10.1007/s00239-019-09914-3>
 12. Bonomo RA. β -Lactamases: A focus on current challenges. *Cold Spring Harb Perspect Med*. 2017;7(1):1-16.
 13. Bush K, Bradford PA. Interplay between β -lactamases and new β -lactamase inhibitors. Vol. 17, *Nature Reviews Microbiology*. Nature Publishing Group; 2019. p. 295-306.
 14. Peirano G, Pitout JDD. Extended-Spectrum β -Lactamase-Producing Enterobacteriaceae: Update on Molecular Epidemiology and Treatment Options. *Drugs* [Internet]. 2019;79(14):1529-41. Disponible en: <https://doi.org/10.1007/s40265-019-01180-3>
 15. Castanheira M, Simner PJ, Bradford PA. Extended-spectrum β -lactamases: An update on their characteristics, epidemiology and detection. *JAC-Antimicrobial Resist*. 2021;3(3).
 16. Navon-Venezia S, Kondratyeva K, Carattoli A. *Klebsiella pneumoniae*: A major worldwide source and shuttle for antibiotic resistance. Vol. 41, *FEMS Microbiology Reviews*. Oxford University Press; 2017. p. 252-75.
 17. Nakano V, A. Nishiyama S, J. Avila-Campos M. Beta-lactamasas: su importancia en la resistencia bacteriana [Internet]. [citado 3 de mayo de 2023]. Disponible en: http://www.icb.usp.br/bmm/mariojac/index.php?option=com_content&view=article&id=47&Itemid=&lang=es
 18. Biondi S, Long S, Panunzio M, Qin WL. Current Trends in-Lactam Based-Lactamases Inhibitors. Vol. 18, *Current Medicinal Chemistry*. 2011.

19. Crass RL, Pai MP. Pharmacokinetics and Pharmacodynamics of β -Lactamase Inhibitors. *Pharmacotherapy*. 2019;39(2):182-95.
20. Asadi Karam MR, Habibi M, Bouzari S. Urinary tract infection: Pathogenicity, antibiotic resistance and development of effective vaccines against Uropathogenic Escherichia coli. Vol. 108, *Molecular Immunology*. Elsevier Ltd; 2019. p. 56-67.
21. Gupta R, Srivastava D, Sahu M, Tiwari S, Ambasta RK, Kumar P. Artificial intelligence to deep learning: machine intelligence approach for drug discovery [Internet]. Vol. 25, *Molecular Diversity*. Springer International Publishing; 2021. 1315-1360 p. Disponible en: <https://doi.org/10.1007/s11030-021-10217-3>
22. Muratov EN, Bajorath J, Sheridan RP, Tetko I V., Filimonov D, Poroikov V, et al. QSAR without borders. Vol. 49, *Chemical Society Reviews*. Royal Society of Chemistry; 2020. p. 3525-64.
23. Lozano-Aponte J, Scior T. ¿Qué sabe Ud. acerca de...QSAR? *Rev Mex ciencias Farm*. 2012;43(2):82-4.
24. Binding Database Home [Internet]. [citado 24 de junio de 2023]. Disponible en: <https://www.bindingdb.org/rwd/bind/index.jsp>
25. NLM. PubChem [Internet]. Disponible en: <https://pubchem.ncbi.nlm.nih.gov/>
26. Dragon 7.0 - Kode Chemoinformatics [Internet]. [citado 24 de junio de 2023]. Disponible en: <https://chm.kode-solutions.net/pf/dragon-7-0/>
27. National University of Ireland D. Infrared Spectroscopy for Food Quality Analysis and Control. Sun D-W, editor. Dublín; 2016. 73-74 p.
28. Leardi R, Lupianez Gonzalez A, et al. Genetic algorithms applied to feature selection in PLS regression: how and when to use them. Vol. 41, *Chemometrics and Intelligent Laboratory Systems*. 1998.
29. Jouan-Rimbaud D, Massart DL, Leardi R, De Noord OE. Genetic Algorithms as a Tool for Wavelength Selection in Multivariate Calibration. *Anal Chem*. 1995;67(23):4295-301.
30. Consonni V, Baccolo G, Gosetti F, Todeschini R, Ballabio D. A MATLAB toolbox for multivariate regression coupled with variable selection.

- Chemom Intell Lab Syst. 2021;213(March).
31. Ballabio D, Consonni V. Classification tools in chemistry. Part 1: Linear models. PLS-DA. Anal Methods. 2013;5(16):3790-8.
 32. Todeschini R, Ballabio D, Cassotti M, Consonni V. N3 and BNN: Two New Similarity Based Classification Methods in Comparison with Other Classifiers. J Chem Inf Model. 2015;55(11):2365-74.
 33. Ballabio D, Grisoni F, Todeschini R. Multivariate comparison of classification performance measures. Chemom Intell Lab Syst. 15 de marzo de 2018;174:33-44.
 34. MathWorks - Makers of MATLAB and Simulink - MATLAB & Simulink [Internet]. [citado 24 de junio de 2023]. Disponible en: <https://www.mathworks.com/>
 35. Acevedo Cepeda D, Chaparro Lozano D, Serrato Ladino KN, Rincón-Riveros A, Cepeda A, Lozano C, et al. Revisión sistemática New beta-lactamase inhibitors: current status and application in practice 42. Hechos Microbiol [Internet]. 2021;12(1):pp-pp. Disponible en: <http://www.udea.edu.co/hmRecepción:10/12//2020.Aceptación:12/02/2021>
 36. Cantón R, Loza E, Aznar J, Castillo FJ, Cercenado E, Fraile-Ribot PA, et al. Monitoring the antimicrobial susceptibility of gram-negative organisms involved in intraabdominal and urinary tract infections recovered during the smart study (Spain, 2016 and 2017). Rev Esp Quimioter. 2019;32(2):145-55.
 37. Suseno M, Das S, Semel J, Richard Thomson J. 2407. Emerging Piperacillin/Tazobactam Resistance in Escherichia coli and Klebsiella sp. Open Forum Infect Dis [Internet]. 26 de noviembre de 2018 [citado 16 de junio de 2023];5(Suppl 1):S719. Disponible en: </pmc/articles/PMC6253293/>
 38. Lee DS, Lee SJ, Choe HS, Giacobbe DR. Community-Acquired Urinary Tract Infection by Escherichia coli in the Era of Antibiotic Resistance. Vol. 2018, BioMed Research International. Hindawi Limited; 2018.
 39. David TI, Okonkwo DC, Adelakun NS, Omotuyi OI. Automated QSAR and

- Docking Study of Phyto-Ligands of *Chromolaena odorata* as Inhibitor of Bacteria Beta-lactamase Enzyme. *J Biol Eng Res Rev*. 2019;6(2):32-42.
40. Académica U, Bienestar DESY, Biofarmacia CDE, De IDELABT-. UNIVERSIDAD CATÓLICA DE CUENCA Comunidad Educativa al Servicio del Pueblo “ EVALUACIÓN IN SILICO DE CATEQUINAS COMO Comunidad Educativa al Servicio del Pueblo. 2022;
41. Yu ZJ, Liu S, Zhou S, Li H, Yang F, Yang LL, et al. Virtual target screening reveals rosmarinic acid and salvianolic acid A inhibiting metallo- and serine- β -lactamases. *Bioorganic Med Chem Lett* [Internet]. 2018;28(6):1037-42. Disponible en: <https://doi.org/10.1016/j.bmcl.2018.02.025>

Anexo 1

Modelo de Regresión					
No	Nombre	Código CAS	InChiKey	SMILES	IC50 (nM) Train/test
1	BDBM50149468		MTRYCRPKNWRHLE-CPWLGMPSA-M	C1COCC2=NC=C(N21)/C=C/3\C @ @H]4N(C3=O)C(=CS4)C(=O)[O-]	0,4 Train
2	BDBM50149466		YJSNNWLFWJHTLH-DJTUGOFGSA-M	C1CC2=CC(=NN2C1)/C=C/3\C @ @H]4N(C3=O)C(=CS4)C(=O)[O-]	1 Train
3	BDBM50149467		XFARUGBBAMTOCU-IBSXUBTNSA-M	C1CC2=C(C1)SC3=NC(=CN23)/C=C/4\C @ @H]5N(C4=O)C(=CS5)C(=O)[O-]	1 Train
4	BDBM50191390		IBSSBUWAVGFIIX-TWLAMCEISA-M	C1CS(=O)(=O)CC2=CC(=NN21)/C=C/3\C @ @H]4N(C3=O)C(=CS4)C(=O)[O-]	1 Train
5	BDBM467003		XMFBVBAWNAOLU-ZLUOBGJFSA-N	C1[C @ H](N2C[C @ H]1N(C2=O)O[C @ @H](C(=O)O)F)C(=O)N	1,3 Test
6	BDBM466965		HBJXCJGYADJSFE-ACZMJKKPSA-N	CC1=C[C @ H]2CN([C @ @H]1C(=O)N)C(=O)N2O[C @ @H](C(=O)O)F	1,35 Train
7	BDBM178215		HXKNILWLBVUKER-SUVRPQBXSAN	C1[C @ @H]2N(C1=O)C/C(=C(\C/C=C/CN(CC3=CC=CC=C3)C(=O)OCC4=C=CC=C4)/C(=O)O)/O2)C(=O)OCC5=CC=CC=C5	1,4 Train
8	BDBM466967		QYPIGCMIMCHYCS-WDSKDSINSA-N	CC1=C[C @ H]2CN([C @ @H]1C(=O)N)C(=O)N2OC(C(=O)O)(F)F	1,7 Train
9	BDBM178170		OUIWZKDIHVIQMRZ-ZKAZXBEYSAN	CC1C/C(=C\2/C(N3[C @ H](O2)CC3=O)C(=O)O)/C(=O)O1	1,8 Train
10	BDBM50339145	1192500-31-4		C1[C @ H](N2C[C @ @H]1N(C2=O)OS(=O)(=O)O)C(=O)N	1,9 Train
11	BDBM50191381		ZKKACLLQVNCVQU-VYBYTRPSA-M	COC1=CC=C(C=C1)CN2CCC3=C(C2)C=C(C3)/C=C/4\C @ @H]5N(C4=O)C(=CS5)C(=O)[O-]	2 Test
12	BDBM50284790		IKGZGJCMJFERPR-JSCIZMSRSA-M	CC(=O)/C=C/1\C @ @H]2N(C1=O)[C @ H](C(S2)(C)C)C(=O)[O-]	3 Train
13	BDBM466972		CZCMDCKIUIXEQC-XVMARJQXSA-N	CC1=C[C @ H](N2C[C @ H]1N(C2=O)O[C @ @H](C(=O)O)F)C(=O)N	3,2 Train
14	BDBM178165		VZTMHDMZOMWRDY-BNGSUFQRSAN	COC(=O)C1/C(=C\2/CC(OC2=O)CI)/O[C @ H]3N1C(=O)C3	3,5 Train
15	BDBM466974		ZEICLMMMMGQHIK-XVYDVKMFSA-N	C1CC1C2=C[C @ H]3CN([C @ @H]2C(=O)N)C(=O)N3O[C @ @H](C(=O)O)F	3,5 Test
16	BDBM50191387		MRXVYOLTKNEJHK-VVOKDJNLSA-M	CN1CCN2C=C(N=C2C1)/C=C/3\C @ @H]4N(C3=O)C(=CS4)C(=O)[O-]	4 Train
17	BDBM178171		OUIWZKDIHVIQMRZ-ZKAZXBEYSAN	CC1C/C(=C\2/C(N3[C @ H](O2)CC3=O)C(=O)O)/C(=O)O1	4,2 Train
18	BDBM466966		HBJXCJGYADJSFE-WABBHOIFSA-N	CC1=C[C @ H]2CN([C @ @H]1C(=O)N)C(=O)N2OC(C(=O)O)F	4,4 Test
19	BDBM50076683		XPRZVSWHOFCHGHP-SZVQBCOZSA-M	C[C @ @]1([C @ @H](N2[C @ H](S1(=O)=O)[C @ @H](C2=O)CO)C(=O)[O-])CN3C=CN=N3	6 Train
20	BDBM50076684		ODGTYAIFNCCHEW-JCGDXUMPSA-N	CC1([C @ @H](N2[C @ H](S1(=O)=O)[C @ @H](C2=O)CO)C(=O)O)C	8 Train
21	BDBM50076685		FVJOWYDVEXFNKI-BCMBZQKMSA-N	C1[C @ @H]2N(C1=O)C/C(=C(\C/C=C/C3=CC=CC=C3)/C(=O)O)/O2)C(=O)OCC4=CC=CC=C4	9,2 Test
22	BDBM50076686		HTRXHLGQLRRIRG-UHFFFAOYSA-N	CCC(C)CCCCCCCCCCCC1=C(C(=CC=C1)O)C(=O)O	10 Train
23	BDBM50076687		JTXLOXAHJKZYOH-ZUYGLVRRSA-M	C1CN(CC2=C1SC(=C2)/C=C/3\C @ @H]4N(C3=O)C(=CS4)C(=O)[O-])CC5=CN=CC=C5	10 Train
24	BDBM50076688		QOMLCKUMUUKMOX-XUJVJEKNSA-N	C[C @ @]1([C @ @H](N2[C @ H](S1(=O)=O)CC2=O)C(=O)O)CSC3=NC4=CC=CC=C4O3	10,5 Train
25	BDBM50076689		WUZOFVCGXRSCHI-VLCNFMPSA-N	COC(=O)C1/C(=C(\C/C=C/C2=CC=C2)/C(=O)O)O[C @ H]3N1C(=O)C3	10,6 Train
26	BDBM50076690	148529-38-8		CCCCCCCCCCCC1=C(C(=CC=C1)O)C(=O)O	12 Train
27	BDBM50076691		ZZWPJCNMJSBJDZ-UHEBKQLESA-M	CCCCO[C @ @H]1CCC[C @ @H]2C1=C(N3[C @ H]2/C(=C\)/C3=O)C(=O)[O-]	12 Train
28	BDBM50076692	26631-90-3		CC1([C @ @H](N2[C @ H](S1)[C @ @H](C2=O)Br)C(=O)O)C	14,3 Train
29	BDBM50076693		FEKJDJDOMDQNFAM-GUKKNSHRSA-N	C[C @ @]1([C @ @H](N2[C @ H](S1)[C @ @H](C2=O)Br)C(=O)O)CSC3=NC=CN3C	15,3 Train
30	BDBM50076694	32099803		C[C @ @]1([C @ @H](N2[C @ H](S1(=O)=O)CC2=O)C(=O)O)CN3C=CN=N3	17 Train
31	BDBM50076695	32099803		C[C @ @]1([C @ @H](N2[C @ H](S1(=O)=O)CC2=O)C(=O)O)CN3C=CN=N3	17,2 Train
32	BDBM50076696		XQHZBWNBBIUYN-FWVWVHSMVSA-N	C[C @ @]1([C @ @H](N2[C @ H](S1(=O)=O)CC2=O)C(=O)O)CSC3=NN=NN3C	17,3 Train
33	BDBM50076697		FUPMZGCEKTRXTR-FFKCEWRFSA-M	C/C=C/1\C @ H]2[C @ H]3CCC[C @ @H](C3=C(N2C1=O)C(=O)[O-])OCCC#N	18 Test
34	BDBM50076698		NZNHTQCXWQAKDT-QORJOZBASA-M	C[C @ @]1([C @ @H](N2[C @ H](S1(=O)=O)[C @ @H](C2=O)CO)C(=O)[O-])/C=C\C#N	19 Train
35	BDBM50076699		HBJXCJGYADJSFE-LFYFHCNISA-N	CC1=C[C @ H]2CN([C @ @H]1C(=O)N)C(=O)N2O[C @ H](C(=O)O)F	19 Train
36	BDBM50076700	32099803		C[C @ @]1([C @ @H](N2[C @ H](S1(=O)=O)CC2=O)C(=O)O)CN3C=CN=N3	20 Test
37	BDBM50076701		OVEFSUBTHUNTRK-JCKKZLPESA-M	C/C=C/1\C @ H]2[C @ H]3CCC[C @ @H](C3=C(N2C1=O)C(=O)[O-])OCCF	21 Train
38	BDBM50076702		DVGGIJCNIHENJ-ASTPYSOASA-N	C[C @ @]1([C @ @H](N2[C @ H](S1)[C @ @H](C2=O)Br)C(=O)O)CSC3=NC=NN3	21,1 Train
39	BDBM50076703		QEYQIONWPOMBON-CTJIIYQFSA-N	COC(=O)C1/C(=C(\C/C=C/C2=CC=CC=C2)/C(=O)O)O[C @ H]3N1C(=O)C3	21,7 Test
40	BDBM50076704		KYVATCVZAZHVKG-WOCUHLHXSA-M	CCO[C @ @H]1CCC[C @ @H]2C1=C(N3[C @ H]2/C(=C\)/C3=O)C(=O)[O-]	22 Test
41	BDBM50076705		MXCGPVVSKSQZRV-FYBVGQRMSA-N	C[C @ @]1([C @ @H](N2[C @ H](S1(=O)=O)CC2=O)C(=O)O)CSC3=NCCS3	22,8 Test
42	BDBM50076706		YHANUKIZKPOW-PW-WABBHOIFSA-N	CC1=C[C @ H]2CN([C @ @H]1C3=NN=C(O3)C(=O)N)C(=O)N2OC(C(=O)O)F	24 Test
43	BDBM50076707		ZEICLMMMMGQHIK-CNUIFLNQSA-N	C1CC1C2=C[C @ H]3CN([C @ @H]2C(=O)N)C(=O)N3O[C @ H](C(=O)O)F	24 Train
44	BDBM50076708		OPBQWQXRAFJUEL-GARJFASQSA-N	CC1=C[C @ H]2CN([C @ @H]1C(=O)NOCC3CC3)C(=O)N2O[C @ H](C(=O)O)F	26 Train

45	BDBM50076709	CZCMDCKIUIXEQC-RRKCRQDMSA-N	CC1=C[C@H](N2C[C@H]1N(C2=O)O[C@H](C(=O)O)F)C(=O)N	26	Test
46	BDBM50076710	XQZUNXLRBIFXDD-XRWXPBKOSA-N	CC(C)(C)C(=O)OCOC(=O)C1/C(=C\C/C=C/C2=CC=CC=C2)/C(=O)/O[C@H]3N1C(=O)C3	27,1	Test
47	BDBM50076711	RMNUWMLNSDZJHP-QRHSGQBVSA-N	C1[C@H](N2C[C@H]1N(C2=O)OC(C(=O)O)F)C(=O)NC3CCNCC3	30	Test
48	BDBM50076712	XJAMHNGKVYJQKN-HMDURAKOSA-N	C[C@@]1([C@@H](N2[C@H](S1)[C@@H](C2=O)Br)C(=O)O)CSC3=NC4=CC=CC=C4O3	30,3	Train
49	BDBM50076713	RBWUPNXZJVLGSA-OPGGXJNDSA-N	C=CC/C(=C/1\C(N2[C@H](O1)CC2=O)C(=O)OCC3=CC=CC=C3)/C(=O)O	30,6	Train
50	BDBM50076714	ZZWPJCNMJSBJDZ-GNURROEBSA-M	CCCCO[C@H]1CCC[C@H]2C1=C(N3[C@H]2/C(=C\C)/C3=O)C(=O)[O-]	32	Test
51	BDBM50076715	FCBUHCWXTMAQJG-BIIVOSGPSA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NCNS(=O)(=O)N)C(=O)N2O[C@H](C(=O)O)F	33	Train
52	BDBM50076716	BVKNUAKQEQUMSY-RNJXMRFFSA-N	CC1=C[C@H](N2C[C@H]1N(C2=O)O[C@H](C(=O)O)F)C#N	35	Train
53	BDBM50076717	HTXPDRHSGPNWKA-VRGHXPGPSA-N	C[C@@H]1[C@@H]2[C@H](C(=O)N2C(=C1S[C@@H]3CCO[C@@H]3CN)C(=O)O)[C@@H](C)O	35	Train
54	BDBM50076718	WOFAABWEQYFPBT-HMDURAKOSA-N	C[C@@]1([C@@H](N2[C@H](S1)[C@@H](C2=O)Br)C(=O)O)CSC3=NC4=CC=CC=C4S3	35,9	Train
55	BDBM50076719	IOZBBLWTTHEQCC-CPWLGJMPSA-M	CC(=O)OCC1=C(N2[C@@H](C(=C\C(=O)OC)/C2=O)S(=O)(=O)C1)C(=O)[O-]	37	Train
56	BDBM50076720	VDZDNZXKYQEEM-UIXYVRDFSA-M	C/C=C/1\C@H]2[C@H]3CCC[C@@H](C3=C(N2C1=O)C(=O)[O-])OC(C)C	39	Train
57	BDBM50076721	SIVNUZUWNGEGKW-AGNDVQSCSA-N	C[C@@]1([C@@H](N2[C@H](S1)[C@@H](C2=O)Br)C(=O)O)CSC3=NCCS3	40,3	Train
58	BDBM50076722	JYGFXUBGZBBAC-RCRIAUEHSA-M	C[C@@]1([C@@H](N2[C@H](S1(=O)=O)CC2=O)C(=O)[O-])C=C/C#N	42	Train
59	BDBM50076723	ZXZLCZGIXSRHMF-PJTZRWAQSA-N	CC(C)(C)C(=O)OCOC(=O)C1/C(=C\C/C=C/C2=CC=CC=C2)/C(=O)O/[C@H]3N1C(=O)C3	44,2	Train
60	BDBM50076724	BLOSWWBHNLGJRP-JVIMKECRSA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NNC(=O)C)C(=O)N2OC(C(=O)O)F	47	Train
61	BDBM50076725	DRBXTSRICBKWDR-GPLJDEGSA-M	C[C@@]1([C@@H](C2C(S1(=O)=O)CC2=O)C(=O)[O-])CN3C=C(N=3)C(=O)OC	49	Train
62	BDBM50076726	FXRKUVQNAIPQC-LPEHRKFASA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NCNC(=O)C)C(=O)N2O[C@H](C(=O)O)F	54	Train
63	BDBM50076727	SUJDDWICKSOENB-PTNRUYDSSA-M	CCOC(=O)C1=CN(N=N1)C[C@]2([C@@H](C3C(S2(=O)=O)CC3=O)C(=O)[O-])C	54	Test
64	BDBM50076728	ACQFJWSLWQLLIA-PLWQYUFKSA-M	C/C=C/1\C@H]2[C@H]3CCC[C@@H](C3=C(N2C1=O)C(=O)[O-])OC	55	Test
65	BDBM50076729	YFHXSIOXRFREU-XHNCKOQMSA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NCCS(=O)(=O)N)C(=O)N2O[C@H](C(=O)O)F	57	Test
66	BDBM50076730	KYVATC VZAZHVKG-SCMGXACRSA-M	CCO[C@H]1CCC[C@H]2C1=C(N3[C@H]2/C(=C\C)/C3=O)C(=O)[O-]	57	Train
67	BDBM50076731	HZZVJAQRINQKSD-PBFISSAISA-M	C1[C@H]2N(C1=O)[C@H](C(=C/CO)O2)C(=O)[O-]	60	Train
68	BDBM50076732	32099803	C[C@@]1([C@@H](N2[C@H](S1(=O)=O)CC2=O)C(=O)O)CN3C=CN=N3	60	Train
69	BDBM50076733	FQSPFOVCHGZCIK-YOMQDVOZSA-M	C[C@@]1([C@@H](N2[C@H](S1(=O)=O)C[C@@H](C2=O)CO)C(=O)[O-])C=N/OC	64	Train
70	BDBM50076734	VRAFEQHIWHNGLC-LPEHRKFASA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NCCNS(=O)(=O)C)C(=O)N2O[C@H](C(=O)O)F	64	Test
71	BDBM50076735	HLBYYVYPQDVZIQ-XETXAFAXSA-N	COC(=O)C1/C(=C\CC=C)/C(=O)O/[C@H]2N1C(=O)C2	65,2	Train
72	BDBM50076736	LSWIFYHUYIAXPOB-VLCNFMPYSA-N	C=CC/C(=C/1\C(N2[C@H](O1)CC2=O)C(=O)OCC3=CC=CC=C3)/C=O	66,6	Test
73	BDBM50076737	QDJWMEGFBPTSBD-FYBVGQRMSA-N	C[C@@]1([C@@H](N2[C@H](S1(=O)=O)CC2=O)C(=O)O)CSC3=NN=CN3C	72,2	Train
74	BDBM50076738	XSDSQTYLBBGLJR-OPGGXJNDSA-N	COC(=O)C1/C(=C\CCCC2=CC=CC=C2)/C(=O)O/[C@H]3N1C(=O)C3	72,8	Train
75	BDBM50076739	NZNHTQCXWOAKDT-SURRTWTPSA-M	C[C@@]1([C@@H](N2[C@H](S1(=O)=O)C[C@@H](C2=O)CO)C(=O)[O-])C=C/C#N	74	Test
76	BDBM50076740	FCBPOMQVJRBNEN-FSOTQKASA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NOC[C@@H]3CCC(=O)N3)C(=O)N2O[C@H](C(=O)O)F	77	Test
77	BDBM50076741	PIPGAOFSHNWTAH-QGNKKJBSA-N	C1[C@H]2N(C1=O)C/C(=C\C/C=C/CO)/C(=O)O2)C(=O)OCC3=CC=CC=C3	78,4	Test
78	BDBM50076742	SFPTWFMBZJRMKD-NDQFZYFBSA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NCC3=NC=CN=C3)C(=O)N2OC(C(=O)O)F	81	Test
79	BDBM50076743	ABPLMHATNPQHNC-XHNCKOQMSA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NCCNS(=O)(=O)N)C(=O)N2O[C@H](C(=O)O)F	83	Train
80	BDBM50076744	OOBLDYJFUBJOPF-XHNCKOQMSA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NCC#N)C(=O)N2O[C@H](C(=O)O)F	85	Train
81	BDBM50076745	NMQXSANBHIOUJJ-TYJSTWMOSA-N	C1[C@H]2N(C1=O)C/C(=C\C/C=C/C3=CC=CC=C3)/C(=O)O2)C(=O)OCC4=CC=CC=C4	89	Test
82	BDBM50076746	32099803	C[C@@]1([C@@H](N2[C@H](S1(=O)=O)CC2=O)C(=O)O)CN3C=CN=N3	90	Train
83	BDBM50076747	STUCJEURXSAYOH-OOHKFKTASA-N	C[C@@H]1C2[C@H](C(=O)N2C(=C1S[C@@H]3CCO[C@@H]3CN[C@@H](C(C)C)C(=O)O)C(=O)O)[C@@H](C)O	91	Train
84	BDBM50076748	WYRIPYFYOUAORN-XHNCKOQMSA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NOCNS(=O)(=O)N)C(=O)N2O[C@H](C(=O)O)F	93	Train
85	BDBM50076749	ALXADSUETYDFDB-RFFQBLADSA-N	CCC/C(=C/1\C(N2[C@H](O1)CC2=O)C(=O)OC)/C(=O)O	97	Test
86	BDBM50076750	GCUVODSDAPCLNV-ASTPYSOASA-N	C[C@@]1([C@@H](N2[C@H](S1)[C@@H](C2=O)Br)C(=O)O)CSC3=NN=NN3C	101	Train
87	BDBM50076751	WCRPDOYIEXPPSD-UJNFCWOMSA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NCCC(=O)N)C(=O)N2O[C@H](C(=O)O)F	110	Train
88	BDBM50076752	QEEMJZXDEIFDDR-XWLWVQCSSA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NOC[C@@H]3CC(CN3)(F)F)C(=O)N2O[C@H](C(=O)O)F	110	Train
89	BDBM50076753	BFVDGHQWCMHZCR-YRNGGQEVSA-L	CC1([C@@H](N2[C@H](S1(=O)=O)C(=C\C(=O)[O-])/C2=O)C(=O)[O-])C	120	Test
90	BDBM50076754	DMUDPEVSVHCUNB-RFFQBLADSA-N	COC(=O)C1/C(=C\CC=C)/C(=O)O/[C@H]2N1C(=O)C2	122	Train
91	BDBM50076755	XUUNBOUUAUWVDD-RKEPMNIXSA-M	C[C@@H]1C2[C@H](C(=O)N2C(=C1OC)C(=O)[O-])[C@@H](C)O	130	Train

92	BDBM50076756	OWOUVVRZHCNSFW-RWTFEVOBSA-M	CC1([C@@H](N2[C@H](S1(=O)=O)/C(=C\C(=O)OC(C)(C)C)/C2=O)C(=O)[O-])C	136 Train
93	BDBM50076757	OUNQKFHIKDLFI-FTVIDMELSA-N	CC(=O)OC/C=C/C/C(=C/1\C(N2[C@H](O1)CC2=O)C(=O)OCC3=CC=CC=C3)/C=O	140 Test
94	BDBM50076758	DIRLPWJRIWGPCS-BIIVOSGPSA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NCO)C(=O)N2O[C@H](C(=O)O)F	150 Test
95	BDBM50076759	TYGMQYXAGAWAHU-NKWVEPMBSA-N	CC1=C[C@H](N2C[C@H]1N(C2=O)OCC(=O)O)C(=O)N	150 Train
96	BDBM50076760	SFPTWFMBJRMKD-SDDRHHMPSA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NCC3=NC=CN=C3)C(=O)N2O[C@H](C(=O)O)F	180 Train
97	BDBM50076761	SWIHLNXGOJYRM-FIQHERPVSA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NOC[C@@H]3CCCN3)C(=O)N2O[C@H](C(=O)O)F	210 Test
98	BDBM50076762	XUUNBOUUAUVVDD-CUWOBHIPSA-M	C[C@H]1C2[C@H](C(=O)N2C(=C1OC)C(=O)[O-])[C@@H](C)O	230 Train
99	BDBM50076763	SOQURWNDDXDSTI-YKFGJVJNSA-N	CC1=C[C@H]2CN([C@@H]1C(=O)N)C(=O)N2OC(C(=O)O)F	230 Train
100	BDBM50076764	QAABSWVJCGWPMP-ITHRYJQJSA-N	CCC/C(=C/1\C(N2[C@H](O1)CC2=O)C(=O)O)/C(=O)O	254 Test
101	BDBM50076765	KMIMTPSCZSJCPM-LPEHRKFASA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NCCCS(=O)(=O)N)C(=O)N2O[C@H](C(=O)O)F	280 Test
102	BDBM50076766	RROHCWIBSPGHX-PTRXPTGYSA-N	C[C@@]1([C@@H](N2[C@H](S1(=O)=O)CC2=O)C(=O)O)CSC3=NC=CN3C	289 Train
103	BDBM50076767	MWZMTCLEBFOZKXV-GPCLJDEGSA-M	C[C@@]1([C@@H](C2C(S1(=O)=O)CC2=O)C(=O)[O-])CN3C(=CN=N3)C(=O)OC	300 Train
104	BDBM50076768	CJHINZCBCHTSOE-PRFJZKWSA-M	C[C@@]1([C@@H](C2C(S1(=O)=O)CC2=O)C(=O)[O-])CN3C(=C(N=N3)C(=O)OC)C(=O)OC	300 Test
105	BDBM50076769	LTSHHIPSCFMNJ-LZUCFRIBSA-M	CC1([C@@H](N2[C@H](S1(=O)=O)/C(=C\C(=O)OC)/C2=O)C(=O)[O-])C	308 Train
106	BDBM50076770	CBBUHGXUJWGHXR-INTQDDNPSA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NCC3=NC=CO3)C(=O)N2O[C@H](C(=O)O)F	310 Test
107	BDBM50076771	AGRPURZKBDMSSM-PTNRUYDSSA-M	CCOC(=O)C1=CN=NN1C[C@]2([C@@H](C3C(S2(=O)=O)CC3=O)C(=O)[O-])C	340 Test
108	BDBM50076772	OWSYJEORQIBAC-IGVIJBNSA-L	CC(=O)OCC1=C(N2[C@@H](/C(=C\C(=O)[O-])/C2=O)S(=O)(=O)C1)C(=O)[O-]	400 Train
109	BDBM50076773	FJEAJGJSMLOQFME-MJSMECKXSA-N	COC(=O)C1(/C(=C/C(=O)O)/O[C@H]2N1C(=O)C2)/C=C/C3=CC=CC=C3	408 Train
110	BDBM50076774	TXPJIXCYUFODBL-WWTCDSDASA-N	C1[C@H]2N(C1=O)C(/C(=C/C(=O)O)/O2)(C/C=C/C3=CC=CC=C3)C(=O)OCC4=CC=CC=C4	426 Train
111	BDBM50076775	RUDGJWAKUPBPKU-XPUUQOCRSA-N	CC1=C[C@H]2CN([C@@H]1C(=O)N)C(=O)N2OCC(=O)O	430 Test
112	BDBM50076776	FKENQMMABCRJMK-RITPCOANSA-N	CC1([C@@H](N2[C@H](S1(=O)=O)CC2=O)C(=O)O)C	557 Train
113	BDBM50076777	GIFHLGGIPIAVNV-RMVUORRISA-L	C[C@@]1([C@@H](C2C(S1(=O)=O)CC2=O)C(=O)[O-])CN3C=C(N=N3)C(=O)[O-]	600 Train
114	BDBM50076778	QEYUTORLTFVDOI-VNHYZAJKSA-N	CC1([C@@H](N2[C@H](S1(=O)=O)[C@@H](C2=O)CC3=CN(N=N3)C4=CC=C(C=C4)N)C(=O)O)C	600 Test
115	BDBM50076779	CRIIDFPLOBDCDK-IESOIDBDSA-N	COC(=O)C1/C(=C(/C=C/C2=CC=CC=C2)\C=O)/O[C@H]3N1C(=O)C3	639 Train
116	BDBM50076780	ODGTYAIFNCCHEW-JCGDXUMPSA-N	CC1([C@@H](N2[C@H](S1(=O)=O)[C@@H](C2=O)CO)C(=O)O)	650 Train
117	BDBM50076781	XXDOEORVWMNHY-LPEHRKFASA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NC3COC3)C(=O)N2O[C@H](C(=O)O)F	660 Train
118	BDBM50076782	AVPVJZPIBNHZQN-CAUGXEKISA-M	C[C@@]1([C@@H](C2C(S1(=O)=O)CC2=O)C(=O)[O-])CN3C=CN=N3	690 Test
119	BDBM50076783	KPAZRDMJBFRIBA-VNHYZAJKSA-N	CC1([C@@H](N2[C@H](S1(=O)=O)[C@@H](C2=O)CC3=CN(N=N3)C4=CC=C(C=C4)O)C(=O)O)C	700 Train
120	BDBM50076784	LDDMNOXUAMVHDZ-WMUAKXSPSA-N	C[C@@]1([C@@H](C2C(S1(=O)=O)CC2=O)C(=O)O)CN3C=C(N=N3)N	700 Train
121	BDBM50076785	XSJFKEVFSNPNJZ-XRRVDJEJSA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NCC[C@@H]3CCC(=O)N3)C(=O)N2O[C@H](C(=O)O)F	730 Test
122	BDBM50076787	ITFVZKLBKASFHP-NYXMJBRJSA-M	CC(=O)OCC1=C(N2[C@@H](/C(=C\C3=CC=CC=N3)/C2=O)S(=O)(=O)C1)C(=O)[O-]	800 Train
123	BDBM50076788	SUKFPPMYJISQEB-SNVBAGLBSA-M	CC(=O)OCC1=C(N2[C@@H](C(=C)C2=O)S(=O)(=O)C1)C(=O)[O-]	800 Train
124	BDBM50076789	PODACQWVDNEGKW-WMYCVDERSA-N	C1[C@H]2N(C1=O)C(/C(=C/C(=O)O)/O2)(C/C=C\C)C(=O)OCC3=CC=CC=C3	828 Test
125	BDBM50076790	RTZMBJMWCEYEE-HUCLUGSFSA-N	CC(C)C(=O)OCOC(=O)C1(/C(=C/C(=O)O)/O[C@H]2N1C(=O)C2)/C=C/C3=CC=CC=C3	871 Train
126	BDBM50076791	NHCGBXGSGNHZFN-HBNTYKESA-N	CC1=C[C@H]2CN([C@@H]1CNC(=O)OC(C)(C)C)C(=O)N2O[C@H](C(=O)O)F	940 Train
127	BDBM50076792	SPTMRZXEWVYFQ-LGJWYTRLSA-N	C=CCC1(/C(=C/C(=O)O)/O[C@H]2N1C(=O)C2)C(=O)OCC3=CC=CC=C3	945 Train
128	BDBM50076793	JLLMLHDL0MWYNZ-VNHYZAJKSA-N	CC1([C@@H](N2[C@H](S1(=O)=O)[C@@H](C2=O)CC3=CN(N=N3)C4=CC=CC=C4)C(=O)O)C	1000 Test
129	BDBM50076794	JYGFUXUBGZBBAC-RCRIAUHESA-M	C[C@@]1([C@@H](N2[C@H](S1(=O)=O)CC2=O)C(=O)[O-])C=C/C#N	1070 Test
130	BDBM50076795	FJEAJGJSMLOQFME-MJSMECKXSA-N	COC(=O)C1(/C(=C/C(=O)O)/O[C@H]2N1C(=O)C2)/C=C/C3=CC=CC=C3	1560 Train
131	BDBM50076796	QSGHUBRFRC SOPB-VNHYZAJKSA-N	CC1([C@@H](N2[C@H](S1(=O)=O)[C@@H](C2=O)CC3=CN(N=N3)C4=CC=C(C=C4)F)C(=O)O)C	1600 Test
132	BDBM50076797	DZCHPZBOQLTTII-RMVUORRISA-L	C[C@@]1([C@@H](C2C(S1(=O)=O)CC2=O)C(=O)[O-])CN3C(=CN=N3)C(=O)[O-]	1700 Train
133	BDBM50076798	ZUIFBYULLQUXTF-VLCNFMPSA-N	C=CC/C(=C/1\C(N2[C@H](O1)CC2=O)C(=O)OCC3=CC=CC=C3)/CO	1950 Test
134	BDBM50076799	KBUJSCWFZMYQLR-QQKZXBOMSA-N	C1[C@H]2N(C1=O)C(/C(=C/C(=O)O)/O2)(C/C=C/C3=CC=CC=C3)C(=O)OCC4=CC=CC=C4	2.79E+03 Test
135	BDBM50076800	WHIIGHHQWOKEP-KWLOKCHISA-N	CC1([C@@H](N2[C@H](S1(=O)=O)C(C2=O)C/C=C/C(=O)OC)C(=O)O)C	2800 Train
136	BDBM50076801	SURSZFGYJRFQEX-IVZWLZJFSA-N	CC1=C[C@H]2CN([C@@H]1CNC(=O)C)C(=O)N2O[C@H](C(=O)O)F	3600 Train
137	BDBM50076802	FAKBCYAIPOVYTK-TYJSTWMOSA-N	C1[C@H]2N(C1=O)C(/C(=C/C(=O)O)/O2)(C/C=C/C3=CC=CC=C3)C(=O)OCC4=CC=CC=C4	4120 Train
138	BDBM50076803	WKUUKIISZDMTQV-XLPZGREQSA-N	CC1=C[C@H]2CN([C@@H]1CNC(=O)C(F)(F)F)C(=O)N2O[C@H](C(=O)O)F	4200 Train

139	BDBM50076804	GWGNBTHOPHGFDX-WXRZDVUGUSA-N	C1[C@@H]2N(C1=O)C/C(=C\C3/CC(OC3)C1)/O2)C(=O)OCC4=CC=CC=C4	4390 Train
140	BDBM50076805	AHWXHWAYVLZNSQ-QHCPKHFHSA-N	C1[C@@H]2N(C1=O)C/C(=C/C=O)/O2)(C/C=C/C3=CC=CC=C3)C(=O)OCC4=CC=CC=C4	4920 Train
141	BDBM50076806	DRYKFHDMYZDMRX-UHFFFAOYSA-N	C1=CC(=C(C(=C1)C)C)C2=CC(=C(C=C2)C(=O)O)CS	5500 Test
142	BDBM50076807	PODACQWVDNEGKW-WMYCVDERSA-N	C1[C@@H]2N(C1=O)C/C(=C/C=O)/O2)(C/C=C\CO)C(=O)OCC3=CC=CC=C3	5640 Train
143	BDBM50076808	BGQRDDSRQMAZIR-YDJAWIAZSA-N	CC(=O)OC/C=C\C1(C(=C/C=O)/O[C@H]2N1C(=O)C2)C(=O)OCC3=CC=CC=C3	5900 Train
144	BDBM50076809	KKSAGQLHTXTMBB-AOPMUKPMSA-N	CC(=O)C/C=C/C1[C@H]2N(C1=O)[C@H](C(S2(=O)=O)(C)C)C(=O)O	6500 Train
145	BDBM50076810	MQNFPHMTRMTJB-JCGDXUMPSA-N	CC1([C@@H](N2[C@H](S1(=O)=O)[C@@H](C2=O)CS)C(=O)O)C	6800 Train
146	BDBM50076811	AVPVJZPIBNHZQN-CAUGXEKISA-M	CC[C@@]1([C@@H](C2C(S1(=O)=O)CC2=O)C(=O)[O-])CN3C=CN=N3	7000 Train
147	BDBM50076813	BGQRDDSRQMAZIR-YDJAWIAZSA-N	CC(=O)OC/C=C\C1(C(=C/C=O)/O[C@H]2N1C(=O)C2)C(=O)OCC3=CC=CC=C3	7310 Train
148	BDBM50076814	ANCXGKPUGCZNAI-OOPHUOBASA-N	CC1([C@@H](N2[C@H](S1(=O)=O)[C@@H](C2=O)CC3=CN(N=N3)CC4CCCN4)C(=O)O)C	7900 Test
149	BDBM50076815	ODGTYAIFNCCEHW-KZLJYQGOSA-N	CC1([C@@H](N2[C@H](S1(=O)=O)[C@H](C2=O)CO)C(=O)O)C	8000 Train
150	BDBM50076816	ZUIFBYULLQUXTF-VLCNFMPSA-N	C=CC/C(=C/1\C(N2[C@H](O1)CC2=O)C(=O)OCC3=CC=CC=C3)/CO	8020 Train
151	BDBM50076817	ZLZSKFBDQKTSB-SNVBAGLBSA-M	CC(=O)OCC1=C(N2[C@@H](C(=C(C)C)C2=O)S(=O)(=O)C1)C(=O)[O-]	8300 Train
152	BDBM50076818	RNEZBAWBDZGLDN-CGLKDYMESA-N	C=CC/C(=C/1\C(N2[C@H](O1)CC2=O)(CC=C)C(=O)OCC3=CC=CC=C3)/C=O	8390 Train
153	BDBM50076819	KKSAGQLHTXTMBB-XFDVOCVDSA-N	CC(=O)C/C=C/C1[C@H]2N(C1=O)[C@H](C(S2(=O)=O)(C)C)C(=O)O	8400 Train
154	BDBM50076820	WHIHHHQWOXKEP-ASQRRCTNSA-N	CC1([C@@H](N2[C@H](S1(=O)=O)C(C2=O)C/C=C(=O)OC)C(=O)O)C	8600 Train
155	BDBM50076822	ZCVZNXFILRURF-GOZTUDAPSA-M	C[C@H]([C@@H]1[C@@H]2N(C1=O)C(=C(S2(=O)=O)C(C)C)C(=O)[O-])O	10000 Train
156	BDBM50076823	NODZXNWC PAUXPG-IDLZWDITSA-N	CC1([C@@H](N2[C@H](S1=O)C(C2=O)C/C=C/C(=O)N(C)OC)C(=O)O)C	12000 Test
157	BDBM50076824	APLLQHCDOOGSFW-OQAFJANISA-N	CC1([C@@H](N2[C@H](S1(=O)=O)C(C2=O)C/C=C/C(=O)C(=O)O)C	12000 Train
158	BDBM50076825	WEJZNOFAZVDYFB-FOXUPHSNSA-N	COC(=O)C1/C(=C\C=O)/O[C@H]2N1C(=O)C2)CC=C	13300 Test
159	BDBM50076826	83173-24-4	CC/C=C\C/C=C\C=C\C1CC1CC1=C(C(=CC=C1)O)C(=O)O	13400 Test
160	BDBM50076827	ODGTYAIFNCCEHW-KZLJYQGOSA-N	CC1([C@@H](N2[C@H](S1(=O)=O)[C@H](C2=O)CO)C(=O)O)C	14600 Train
161	BDBM50076828	NRHINWWSKCNJN-NZHQHIMKSA-N	CC1([C@@H](N2[C@H](S1(=O)=O)[C@@H](C2=O)CC3=CN(N=N3)CC(C4=CC=C(C=C4)O)N)C(=O)O)C	15000 Train
162	BDBM50076829	FPJUOUJTKUOLCV-ISDAKNICSA-M	CCCC1/C(=C\C=O)/O[C@H]2N1C(=O)C2)C(=O)[O-]	17800 Test
163	BDBM50076830	VW SIAIDFEKNQNE-SRNKDFPPSA-N	COC(=O)C1/C(=C\C=O)/O[C@H]2N1C(=O)C2)C/C=C/C3=CC=CC=C3	19600 Train
164	BDBM50076831	YOCVKSISBJBEES-ROCHPLJESA-M	CC(=O)OCC1=C(N2[C@@H](/C(=C/C3=CC=CC=C3)/C2=O)S(=O)(=O)C1)C(=O)[O-]	20000 Train
165	BDBM50076832	YOCVKSISBJBEES-VXXVJXPCSA-M	CC(=O)OCC1=C(N2[C@@H](/C(=C\C3=CC=CC=C3)/C2=O)S(=O)(=O)C1)C(=O)[O-]	20000 Train
166	BDBM50076833	ZSPQKSEVIUMSR-KNVNGTOESA-M	CC(=O)/C=C/1\C@@H]2N(C1=O)C(=C(CS2)COC(=O)C)C(=O)[O-]	20000 Train
167	BDBM50076834	HFICAXNAWUOCGS-XJQCFMKWSA-M	CC(=O)OCC1=C(N2[C@@H](C(=C=CC(C)(C)C)C2=O)S(=O)(=O)C1)C(=O)[O-]	20000 Train
168	BDBM50076835	GGHHCDXUXTXKBM-OMPYUXFZSA-M	CC(=O)OCC1=C(N2[C@@H](/C(=C\C(C)C)C)C2=O)S(=O)(=O)C1)C(=O)[O-]	20000 Train
169	BDBM50076836	KKHROROKABOCS-XTXQVDMPSA-M	CC(=O)OCC1=C(N2[C@@H](/C(=C\C(=O)OC(C)C)C)C2=O)SC1)C(=O)[O-]	20000 Train
170	BDBM50076837	BIUVASJLHVNNPE-GPWUBDCCSA-M	CC(=O)OCC1=C(N2[C@@H](/C(=C\C=O)/C2=O)S(=O)(=O)C1)C(=O)[O-]	20000 Train
171	BDBM50076838	IAYUXXWVOBMOKN-SNVBAGLBSA-M	CC(=O)OCC1=C(N2[C@@H](C(=C(Br)Br)C2=O)SC1)C(=O)[O-]	20000 Train
172	BDBM50076839	VGEKZJBPIFSJW-VXXVJXPCSA-M	CC(=O)OCC1=C(N2[C@@H](/C(=C\C3=CC=CC=C3)/C2=O)SC1)C(=O)[O-]	20000 Train
173	BDBM50076840	MDAINRPXGSCIOC-IGVIJNBESA-L	CC(=O)OCC1=C(N2[C@@H](/C(=C\C(=O)[O-])/C2=O)SC1)C(=O)[O-]	20000 Train
174	BDBM50076841	BQYSW XZGFUSYTK-RWTFEVOBSA-M	CC1([C@@H](N2[C@H](S1)/C(=C\C(=O)OC(C)C)C2=O)C(=O)[O-])C	20000 Train
175	BDBM50076842	ZLYQUOKXGNQRIC-LLVKDONJSA-M	CC(=O)OCC1=C(N2[C@@H](C(=C=C)C2=O)S(=O)(=O)C1)C(=O)[O-]	20000 Train
176	BDBM50076843	WMEHQPKGOVJLPV-LZUCFRIBSA-M	CC1([C@@H](N2[C@H](S1)/C(=C\C(=O)OC)/C2=O)C(=O)[O-])C	20000 Train
177	BDBM50076844	AOZHCHLXLZFCF-CPWLGJMPMSA-M	CC(=O)OCC1=C(N2[C@@H](/C(=C\C(=O)OC)/C2=O)SC1)C(=O)[O-]	20000 Train
178	BDBM50076845	BJXIZXYMGFDKBP-YRNGGQVESA-L	CC1([C@@H](N2[C@H](S1)/C(=C\C(=O)[O-])/C2=O)C(=O)[O-])C	20000 Test
179	BDBM50076846	KKHROROKABOCS-XTXQVDMPSA-M	CC(=O)OCC1=C(N2[C@@H](/C(=C\C(=O)OC(C)C)C)C2=O)SC1)C(=O)[O-]	20000 Test
180	BDBM50076847	IWSSJMHCIOGHBY-CYZOPGGISA-M	CC(=O)OCC1=C(N2[C@@H](/C(=C/Br)/C2=O)S(=O)(=O)C1)C(=O)[O-]	20000 Test
181	BDBM50076848	PXOKVALSRAGFAH-SNVBAGLBSA-M	CC(=O)OCC1=C(N2[C@@H](C(=C(Br)Br)C2=O)S(=O)(=O)C1)C(=O)[O-]	20000 Test
182	BDBM50076849	SPOYTMW GOMJDIX-SNVBAGLBSA-M	CC(=O)OCC1=C(N2[C@@H](C(=C(C)C)C2=O)SC1)C(=O)[O-]	20000 Test
183	BDBM50076850	LEADWAGQWAFRFE-CYZOPGGISA-M	CC(=O)OCC1=C(N2[C@@H](/C(=C/Br)/C2=O)SC1)C(=O)[O-]	20000 Test
184	BDBM50076851	GVESJW VLYJUUMT-OMPYUXFZSA-M	CC(=O)OCC1=C(N2[C@@H](/C(=C\C(C)C)C)C2=O)SC1)C(=O)[O-]	20000 Test
185	BDBM50076852	ADMWSBHYYMUDFV-LGJW TYRLSA-N	C=CCC1/C(=C\C=O)/O[C@H]2N1C(=O)C2)C(=O)OCC3=CC=CC=C3	23600 Train

186	BDBM50076853	HXCNDHVLZREMRV-QYJMCNQRSA-N	CC1([C@@H](N2[C@@H](S1(=O)=O)C(C2=O)C/C=C/C#N)C(=O)O)C	24000 Train
187	BDBM50076854	DJIMYYWAZIOWRD-JCGDXUMPSA-M	CC1([C@@H](N2[C@@H](S1)[C@@H](C2=O)CO)C(=O)[O-])C	31000 Train
188	BDBM50076855	KYTCAZSHGDLTNM-UHFFFAOYSA-N	C1=CC(=CC=C1C2=CC(=C(C=C2)C(=O)O)CS)F	32800 Train
189	BDBM50076856	TXPJXCXUFODBL-WWTCCDDSASA-N	C1[C@@H]2N(C1=O)C(C(=O)O)/O2(C/C=C/C3=CC=CC=C3)C(=O)OCC4=CC=CC=C4	45700 Train
190	BDBM50076857	JWBKZXUHEFRKRV-FZGKIIMS-M	C[C@@H]([C@@H]1[C@@H]2N(C1=O)[C@@H](C(S2(=O)=O)(C)C)C(=O)[O-])O	47900 Train
191	BDBM50076858	YJWBOGUQTYCPBN-GOZTUDAPSA-M	C[C@@H]([C@@H]1[C@@H]2N(C1=O)C(=C(S2)C(C)C)C(=O)[O-])O	50000 Train
192	BDBM50076859	MQNFPHMTROMTJB-KZLJYQGOSA-N	CC1([C@@H](N2[C@@H](S1(=O)=O)[C@@H](C2=O)CS)C(=O)O)C	51700 Test
193	BDBM50076860	DJIMYYWAZIOWRD-KZLJYQGOSA-M	CC1([C@@H](N2[C@@H](S1)[C@@H](C2=O)CO)C(=O)[O-])C	59000 Train
194	BDBM50076861	AFPLSUQTYBHSDG-RVJQKOHUSA-N	CC1([C@@H](N2[C@@H](S1(=O)=O)[C@@H](C2=O)S)C(=O)O)C	69200 Train
195	BDBM50076862	GSVZWBPOOFMGAJ-PISQCCPLSA-M	C[C@@H]([C@@H]1[C@@H]2N(C1=O)C(=C(S2=O)C(C)C)C(=O)[O-])O	90000 Test
196	BDBM50076863	VNFWGXIPXW XBS-XLPZGREQSA-N	CC1=C[C@@H]2CN([C@@H]1CN)C(=O)N2O[C@@H](C(=O)O)F	100000 Test
197	BDBM50076864	MMKATEMYLYLBTG-RFFQBLADSA-M	CCC/C(=C/1\C(N2[C@@H](O1)CC2=O)C(=O)[O-])/CO	118000 Test
198	BDBM50076865	22916-47-8	C1=CC(=C(C=C1C1)C1)COC(CN2C=CN=C2)C3=C(C=C(C=C3)C1)C1	125000 Train
199	BDBM50076866	JWBKZXUHEFRKRV-SKHQTKALSA-M	C[C@@H]([C@@H]1[C@@H]2N(C1=O)[C@@H](C(S2(=O)=O)(C)C)C(=O)[O-])O	204000 Test
200	BDBM50076867	DJIMYYWAZIOWRD-KZLJYQGOSA-N	CC1([C@@H](N2[C@@H](S1)[C@@H](C2=O)CO)C(=O)O)C	275000 Train
201	BDBM50076869	JWBKZXUHEFRKRV-VUBPFVNUSA-M	C[C@@H]([C@@H]1[C@@H]2N(C1=O)[C@@H](C(S2(=O)=O)(C)C)C(=O)[O-])O	334000 Train
202	BDBM50076870	JQVSNQBQXNNKIW-WYDQCIBASA-M	C[C@@H]([C@@H]1[C@@H]2CC(=C(N2C1=O)C(=O)[O-])OC)O	340000 Train
203	BDBM50076871	RHBQBGRVEOOJQL-JCGDXUMPSA-N	CC1([C@@H](N2[C@@H](S1)[C@@H](C2=O)CS)C(=O)O)C	601000 Train
204	BDBM50076872	RHBQBGRVEOOJQL-KZLJYQGOSA-N	CC1([C@@H](N2[C@@H](S1)[C@@H](C2=O)CS)C(=O)O)C	648000 Train
205	BDBM50076873	DJIMYYWAZIOWRD-JCGDXUMPSA-N	CC1([C@@H](N2[C@@H](S1)[C@@H](C2=O)CO)C(=O)O)C	752000 Train
206	BDBM50076874	WOYLBXUCZPJYOV-RVJQKOHUSA-N	CC1([C@@H](N2[C@@H](S1)[C@@H](C2=O)S)C(=O)O)C	988000 Train
207	BDBM50076875	WOYLBXUCZPJYOV-ALEPSDHESA-N	CC1([C@@H](N2[C@@H](S1)[C@@H](C2=O)S)C(=O)O)C	1000000 Train

Anexo 2

Modelo de clasificación

No	Nombre	Código CAS	InChIKey	Código SMILES	IC50 (nM)	Clase
1	BDBM50149468		MTRYCRPNWRHLE-CPWLGJMPA-M	C1COCC2=NC=C(N21)/C=C/3\C@ @H]4N(C3=O)C(=CS4)C(=O)[O-]	0,4	1
2	BDBM50149466		YJSNNWLWFJHTLH-DJTUGOFGSA-M	C1CC2=CC(=NN2C1)/C=C/3\C@ @H]4N(C3=O)C(=CS4)C(=O)[O-]	1	1
3	BDBM50149467		XFARUGBBAMTOCU-IBSXUBTNSA-M	C1CC2=C(C1)SC3=NC(=CN23)/C=C/4\C@ @H]5N(C4=O)C(=CS5)C(=O)[O-]	1	1
4	BDBM50191390		IBSSBUWAVGFIIX-TWLAMCEISA-M	C1CS(=O)(=O)CC2=CC(=NN21)/C=C/3\C@ @H]4N(C3=O)C(=CS4)C(=O)[O-]	1	1
5	BDBM467003		XMFXJVBXAWNAOLU-ZLUOBGJFSA-N	C1C[C@H](N2C[C@H]1N(C2=O)O[C@ @H](C(=O)O)F)C(=O)N	1,3	1
6	BDBM466965		HBXJCJGYADJSFE-ACZMJJKPASA-N	CC1=C[C@H]2CN([C@ @H]1C(=O)N)C(=O)N2O[C@ @H](C(=O)O)F	1,35	1
7	BDBM178215		HXKNILWLBVUKER-SUVRPQBXSAN-N	C1[C@ @H]2N(C1=O)C/C(=C(\C/C=C/CN(CC3=CC=CC=C3)C(=O)OCC4=CC=CC=C4)/C(=O)O)/O2)C(=O)OCC5=CC=CC=C5	1,4	1
8	BDBM466967		QYPIGCMIMCHYCS-WDSKDSINSA-N	CC1=C[C@H]2CN([C@ @H]1C(=O)N)C(=O)N2OC(C(=O)O)F	1,7	1
9	BDBM178170		OUWZKDIHVIMRZ-ZKAZXBEYSA-N	CC1C/C(=C)2C(N3[C@H](O2)CC3=O)C(=O)O/C(=O)O1	1,8	1
10	BDBM50339145	1192500-31-4		C1C[C@H](N2C[C@ @H]1N(C2=O)OS(=O)(=O)O)C(=O)N	1,9	1
11	BDBM50191381		ZKKAQLLQVNCVQU-VYBYTRPSA-M	COC1=CC=C(C=C1)CN2CCC3=C(C2)C=C(S3)/C=C/4[C@ @H]5N(C4=O)C(=CS5)C(=O)[O-]	2	1
12	BDBM50284790		IKGZFJCMJFERPR-JSCIZMSRSA-M	CC(=O)C=C/1\C@ @H]2N(C1=O)[C@H](C(S2)(C)C)C(=O)[O-]	3	1
13	BDBM466972		CZCMDCKIUIXEQC-XVMARJQXSA-N	CC1=C[C@H](N2C[C@H]1N(C2=O)O[C@ @H](C(=O)O)F)C(=O)N	3,2	1
14	BDBM178165		VZTMHDMZOMWRDY-BNGSUFQRSA-N	COC(=O)C1/C(=C)2/CC(OC2=O)C1)O[C@H]3N1C(=O)C3	3,5	1
15	BDBM466974		ZEICLMMMMGQHIK-XVYDVKMFSA-N	C1CC1C2=C[C@H]3CN([C@ @H]2C(=O)N)C(=O)N3O[C@ @H](C(=O)O)F	3,5	1
16	BDBM50191387		MRXVYOLTKNEJHK-VVOKDJNLSA-M	CN1CCN2C=C(N=C2C1)/C=C/3\C@ @H]4N(C3=O)C(=CS4)C(=O)[O-]	4	1
17	BDBM178171		OUWZKDIHVIMRZ-ZKAZXBEYSA-N	CC1C/C(=C)2C(N3[C@H](O2)CC3=O)C(=O)O/C(=O)O1	4,2	1
18	BDBM466966		HBXJCJGYADJSFE-WABBHOIFSA-N	CC1=C[C@H]2CN([C@ @H]1C(=O)N)C(=O)N2OC(C(=O)O)F	4,4	1
19	BDBM50076683		XPRZVSWHOFCHGP-SZVQBCOZSA-M	C[C@ @]1([C@ @H](N2[C@H](S1(=O)=O)[C@ @H](C2=O)CO)C(=O)[O-])CN3C=CN=N3	6	1
20	BDBM50076684		ODGTYAIFNCCHEW-JCGDXUMPSA-N	CC1([C@ @H](N2[C@H](S1(=O)=O)[C@ @H](C2=O)CO)C(=O)O)C	8	1
21	BDBM50076685		FVJOWYDVEXFNKI-BCMBZQKMSA-N	C1[C@ @H]2N(C1=O)C/C(=C(\C/C=C/C3=CC=CC=C3)/C(=O)O)/O2)C(=O)OCC4=CC=CC=C4	9,2	1
22	BDBM50076686		HTRXHLGQLRRIRG-UHFFFAOYSA-N	CCC(C)CCCCCCCCCCC1=C(C(=CC=C1)O)C(=O)O	10	1
23	BDBM50076687		JTXLOXAHJKZYOH-ZUYGLVRRSA-M	C1CN(CC2=C1SC(=C2)/C=C/3\C@ @H]4N(C3=O)C(=CS4)C(=O)[O-])CC5=CN=CC=C5	10	1
24	BDBM50076688		QOMLCKUMUUKMOX-XUJVJEKNSA-N	C[C@ @]1([C@ @H](N2[C@H](S1(=O)=O)CC2=O)C(=O)O)CSC3=NC4=CC=CC=C4O3	10,5	1
25	BDBM50076689		WUZOFCVGRXSCHI-VLCNFMPSA-N	COC(=O)C1/C(=C(\CCCC2=CC=CC=C2)/C(=O)O)O[C@H]3N1C(=O)C3	10,6	1
26	BDBM50076690	148529-38-8		CCCCCCCCCCC1=C(C(=CC=C1)O)C(=O)O	12	1
27	BDBM50076691		ZZWPJCNMJSBJDZ-UHEBKQLESA-M	CCCCO[C@ @H]1CCC[C@ @H]2C1=C(N3[C@H]2/C(=C)C3=O)C(=O)[O-]	12	1
28	BDBM50076692	26631-90-3		CC1([C@ @H](N2[C@H](S1)[C@ @H](C2=O)Br)C(=O)O)C	14,3	1
29	BDBM50076693		FEKDJDOMDQNFAM-GUKKNSHRSA-N	C[C@ @]1([C@ @H](N2[C@H](S1)[C@ @H](C2=O)Br)C(=O)O)CSC3=NC=CN3C	15,3	1
30	BDBM50076694	32099803		C[C@ @]1([C@ @H](N2[C@H](S1(=O)=O)CC2=O)C(=O)O)CN3C=CN=N3	17	1
31	BDBM50076695	32099803		C[C@ @]1([C@ @H](N2[C@H](S1(=O)=O)CC2=O)C(=O)O)CN3C=CN=N3	17,2	1
32	BDBM50076696		XQHZBWNBBUIYCN-FWWHASMVSA-N	C[C@ @]1([C@ @H](N2[C@H](S1(=O)=O)CC2=O)C(=O)O)CSC3=NN=NN3C	17,3	1
33	BDBM50076697		FUPMZGCETKRXT-FFKCEWRFSAN-M	C/C=C/1\C@H]2[C@H]3CCC[C@ @H](C3=C(N2C1=O)C(=O)[O-])OCCC#N	18	1
34	BDBM50076698		NZNHTQCXWOAKDT-QORJOZBASA-M	C[C@ @]1([C@ @H](N2[C@H](S1(=O)=O)[C@ @H](C2=O)CO)C(=O)[O-])/C=C/C#N	19	1
35	BDBM50076699		HBXJCJGYADJSFE-LYFYHCNISA-N	CC1=C[C@H]2CN([C@ @H]1C(=O)N)C(=O)N2O[C@H](C(=O)O)F	19	1
36	BDBM50076700	32099803		C[C@ @]1([C@ @H](N2[C@H](S1(=O)=O)CC2=O)C(=O)O)CN3C=CN=N3	20	1
37	BDBM50076701		OVEFSUBTHUNTRK-JCKKZLPESA-M	C/C=C/1\C@H]2[C@H]3CCC[C@ @H](C3=C(N2C1=O)C(=O)[O-])OCCF	21	1
38	BDBM50076702		DVGGIJCNIHENJ-ASTPYSOASA-N	C[C@ @]1([C@ @H](N2[C@H](S1)[C@ @H](C2=O)Br)C(=O)O)CSC3=NC=NN3	21,1	1
39	BDBM50076703		QEYQIONWPOMBON-CTJIYQFSA-N	COC(=O)C1/C(=C(\C/C=C/C2=CC=CC=C2)/C(=O)O)/O[C@H]3N1C(=O)C3	21,7	1
40	BDBM50076704		KYVATCVZAZHVKG-WOCUHLHXSA-M	CCO[C@ @H]1CCC[C@ @H]2C1=C(N3[C@H]2/C(=C)C3=O)C(=O)[O-]	22	1
41	BDBM50076705		MXCGPVSQZRZV-FYBVGQRMSA-N	C[C@ @]1([C@ @H](N2[C@H](S1(=O)=O)CC2=O)C(=O)O)CSC3=NCCS3	22,8	1

42	BDBM50076706	YHANUKIZKPOWPW-WABBHOIFSA-N	CC1=C[C@H]2CN([C@@H]1C3=NN=C(O)C(=O)N(C(=O)N2OC(C(=O)O)F	24	1
43	BDBM50076707	ZEICLMMMMGQHJK-CNUIFLNQSA-N	C1CC1C2=C[C@H]3CN([C@@H]2C(=O)N)C(=O)N3O[C@H](C(=O)O)F	24	1
44	BDBM50076708	OPBQWQXRAFJUEL-GARJFASQSA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NOCC3CC3)C(=O)N2O[C@H](C(=O)O)F	26	1
45	BDBM50076709	CZCMDCKIUIXEQC-RRKCRQDMSA-N	CC1=C[C@H](N2C[C@H]1N(C2=O)O[C@H](C(=O)O)F)C(=O)N	26	1
46	BDBM50076710	XQZUNXLRBIFXDD-XRWPXBKOSA-N	CC(C)(C)C(=O)OCOC(=O)C1/C(=C(\C/C=C/C2=CC=CC=C2)/C=O)/O[C@H]3N1C(=O)C3	27,1	1
47	BDBM50076711	RMNUWMLNSDZJHP-QRHSGQBVSA-N	C1C[C@H](N2C[C@H]1N(C2=O)OC(C(=O)O)F)C(=O)NC3CCNCC3	30	1
48	BDBM50076712	XJAMHNGKVYJQKN-HMDURAKOSA-N	C[C@@]1([C@H](N2[C@H](S1)[C@@H](C2=O)Br)C(=O)O)CSC3=NC4=CC=CC=C4O3	30,3	1
49	BDBM50076713	RBWUPNXZJVLGSA-OPGGXJNDSA-N	C=CC/C(=C/1\C(N2[C@H](O1)CC2=O)C(=O)OCC3=CC=CC=C3)/C(=O)O	30,6	1
50	BDBM50076714	ZZWPJCNMJSBJDZ-GNURROEBSA-M	CCCCO[C@H]1CCC[C@H]2C1=C(N3[C@H]2/C(=C\C)/C3=O)C(=O)[O-]	32	1
51	BDBM50076715	FCBUHCWXTMAQJG-BIIVOSGPSA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NCNS(=O)(=O)N)C(=O)N2O[C@H](C(=O)O)F	33	1
52	BDBM50076716	BVKNUAKQEQUMSY-RNJXMRFFSA-N	CC1=C[C@H](N2C[C@H]1N(C2=O)O[C@@H](C(=O)O)F)C#N	35	1
53	BDBM50076717	HTXPDRHSGPNWKA-VRGHXPGPSA-N	C[C@@H]1[C@@H]2[C@H](C(=O)N2C(=C1S[C@@H]3CCO[C@@H]3CN)C(=O)O)[C@@H](C)O	35	1
54	BDBM50076718	WOFAABWEQYFPBT-HMDURAKOSA-N	C[C@@]1([C@H](N2[C@H](S1)[C@@H](C2=O)Br)C(=O)O)CSC3=NC4=CC=CC=C4S3	35,9	1
55	BDBM50076719	IOZBBLWTTHEQCC-CPWLGJMPSA-M	CC(=O)OCC1=C(N2[C@H](C(=O)OC)/C2=O)S(=O)(=O)C1(C(=O)[O-]	37	1
56	BDBM50076720	VDZDNXKYYQECM-UIXYVRDFSAM	C/C=C/1\C@H]2[C@H]3CCC[C@@H](C3=C(N2C1=O)C(=O)[O-])OC(C)C	39	1
57	BDBM50076721	SIVNUZUWNGEGKW-AGNDVQSCSA-N	C[C@@]1([C@H](N2[C@H](S1)[C@@H](C2=O)Br)C(=O)O)CSC3=NCCS3	40,3	1
58	BDBM50076722	JYGFUXUBGZBBAC-RCRIAUEHSA-M	C[C@@]1([C@H](N2[C@H](S1(=O)=O)CC2=O)C(=O)[O-])/C=C/C#N	42	1
59	BDBM50076723	ZXZLCZGIXSRHMF-PJTZRWQUSA-N	CC(C)(C)C(=O)OCOC(=O)C1/C(=C(\C/C=C/C2=CC=CC=C2)/C(=O)O)/O[C@H]3N1C(=O)C3	44,2	1
60	BDBM50076724	BLOSWWBHNLGJRP-JVIMKECRSA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NNC(=O)C)C(=O)N2OC(C(=O)O)F	47	1
61	BDBM50076725	DRBXTSRICKBKWDR-GPCLJDEGSA-M	C[C@@]1([C@H](C2C(S1(=O)=O)CC2=O)C(=O)[O-])CN3C=C(N=N3)C(=O)OC	49	1
62	BDBM50076726	FXRXKUVQNAIPQC-LPEHRKFASA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NCNC(=O)C)C(=O)N2O[C@H](C(=O)O)F	54	1
63	BDBM50076727	SUJDDWICKSOENB-PTNRUYDSSA-M	CCOC(=O)C1=CN(N=N1)C[C@]2([C@@H](C3C(S2(=O)=O)CC3=O)C(=O)[O-])C	54	1
64	BDBM50076728	ACQFJWSLWQLLIA-PLWQYUFKSA-M	C/C=C/1\C@H]2[C@H]3CCC[C@@H](C3=C(N2C1=O)C(=O)[O-])OC	55	1
65	BDBM50076729	YFHXDSIOXRFREU-XHNCKOQMSA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NCCS(=O)(=O)N)C(=O)N2O[C@H](C(=O)O)F	57	1
66	BDBM50076730	KYVATCVZAZHV GK-SCMGXACRSA-M	CCO[C@H]1CCC[C@H]2C1=C(N3[C@H]2/C(=C\C)/C3=O)C(=O)[O-]	57	1
67	BDBM50076731	HZZVJAQRINQKSD-PBIFISZAISA-M	C1[C@@H]2N(C1=O)[C@H](C(=C/CO)/O2)C(=O)[O-]	60	1
68	BDBM50076732	32099803	C[C@@]1([C@H](N2[C@H](S1(=O)=O)CC2=O)C(=O)O)CN3C=CN=N3	60	1
69	BDBM50076733	FQSPFOVCHGZCIK-YOMQDVOZSA-M	C[C@@]1([C@H](N2[C@H](S1(=O)=O)[C@@H](C2=O)CO)C(=O)[O-])/C=N/OC	64	1
70	BDBM50076734	VRAFEQHIWHNGLC-LPEHRKFASA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NCCNS(=O)(=O)C)C(=O)N2O[C@H](C(=O)O)F	64	1
71	BDBM50076735	HLBYYVPQDVZIQ-XETXAFAXSA-N	COC(=O)C1/C(=C(\CC=C)/C=O)/O[C@H]2N1C(=O)C2	65,2	1
72	BDBM50076736	LSWFYHUYIAXPOB-VLCNFMYPYSA-N	C=CC/C(=C/1\C(N2[C@H](O1)CC2=O)C(=O)OCC3=CC=CC=C3)/C=O	66,6	1
73	BDBM50076737	QDJWMEGFGPTSBD-FYBVGQRMSA-N	C[C@@]1([C@H](N2[C@H](S1(=O)=O)CC2=O)C(=O)O)CSC3=NN=CN3C	72,2	1
74	BDBM50076738	XSDSQTYLBBGLJR-OPGGXJNDSA-N	COC(=O)C1/C(=C(\CCCC2=CC=CC=C2)/C(=O)O)/O[C@H]3N1C(=O)C3	72,8	1
75	BDBM50076739	NZNHTQCXWOAKDT-SURRTWTPSA-M	C[C@@]1([C@H](N2[C@H](S1(=O)=O)[C@@H](C2=O)CO)C(=O)[O-])/C=C/C#N	74	1
76	BDBM50076740	FCBPOMQVJRBNE-NFSZOTQKASA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NOC[C@@H]3CCC(=O)N3)C(=O)N2O[C@H](C(=O)O)F	77	1
77	BDBM50076741	PIPGAOFSHNWTAH-QGNKKJBSA-N	C1[C@H]2N(C1=O)C/C(=C(\C/C=C/CO)/C=O)/O2)C(=O)OCC3=CC=CC=C3	78,4	1
78	BDBM50076742	SFPTWFMBZJRMKD-NDQFZYFBSA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NCC3=NC=CN=C3)C(=O)N2OC(C(=O)O)F	81	1
79	BDBM50076743	ABPLMHATNQPHNC-XHNCKOQMSA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NCCNS(=O)(=O)N)C(=O)N2O[C@H](C(=O)O)F	83	1
80	BDBM50076744	OOBLDYJFUBJOPF-XHNCKOQMSA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NCC#N)C(=O)N2O[C@H](C(=O)O)F	85	1
81	BDBM50076745	NMQXSANBHIOUJJ-TYJSTWMOSA-N	C1[C@@H]2N(C1=O)C/C(=C(\C/C=C/C3=CC=CC=C3)/C=O)/O2)C(=O)OCC4=CC=CC=C4	89	1
82	BDBM50076746	UNYKCLXPBPYASU-UHFFFAOYSA-N	C[C@@]1([C@H](N2[C@H](S1(=O)=O)CC2=O)C(=O)O)CN3C=CN=N3	90	1
83	BDBM50076747	STUCJEURXSAYOH-OOHKFKTASA-N	C[C@@H]1C2[C@H](C(=O)N2C(=C1S[C@@H]3CCO[C@@H]3CN[C@@H](C(C)C)C(=O)O)C(=O)O)[C@@H](C)O	91	1
84	BDBM50076748	WYRIPYFYOUAORN-XHNCKOQMSA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NOCCS(=O)(=O)N)C(=O)N2O[C@H](C(=O)O)F	93	1
85	BDBM50076749	ALXADSUETYDFDB-RFFQBLADSA-N	CCC/C(=C/1\C(N2[C@H](O1)CC2=O)C(=O)O)C(=O)O	97	1

86	BDBM50076750	GCUVODSDAPCLNV-ASTPYSOASA-N	C[C@@@]1([C@@@H](N2[C@H](S1)[C@@H](C2=O)Br)C(=O)O)CSC3=NN=NN3C	101	1
87	BDBM50076751	WCRPDOYIEXPPSD-UJNFCWOMSA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NCCC(=O)N)C(=O)N2O[C@H](C(=O)O)F	110	1
88	BDBM50076752	QEEMJZXDEIFDDR-XWLWVQCSSA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NOC[C@H]3CC(CN3)(F)F)C(=O)N2O[C@H](C(=O)O)F	110	1
89	BDBM50076753	BFVDGHQWCMHZCR-YRNGGQEVSA-L	CC1([C@@H](N2[C@H](S1(=O)=O)/C(=C\C(=O)[O-])/C2=O)C(=O)[O-])C	120	1
90	BDBM50076754	DMUDPEVSVHCUNB-RFFQBLADSA-N	COC(=O)C1/C(=C(\CC=C)/C(=O)O)/O[C@H]2N1C(=O)C2	122	1
91	BDBM50076755	XUUNBOUUAUVVDD-RKEPMNIXSA-M	C[C@@H]1C2[C@H](C(=O)N2C(=C1OC)C(=O)[O-])[C@@H](C)O	130	1
92	BDBM50076756	OWOUVVRZHCNSFW-RWTFEVOBSA-M	CC1([C@@H](N2[C@H](S1(=O)=O)/C(=C\C(=O)OC(C)(C)C)/C2=O)C(=O)[O-])C	136	1
93	BDBM50076757	OUNQKFHIKDLFI-FTVIDMELSA-N	CC(=O)OC/C=C/C/C(=C/1\C(N2[C@H](O1)CC2=O)C(=O)OCC3=CC=CC=C3)/C=O	140	1
94	BDBM50076758	DIRLPWJRWGPCS-BIIVOSGPPSA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NCO)C(=O)N2O[C@H](C(=O)O)F	150	1
95	BDBM50076759	TYGMQYXAGAWAHU-NKWVEPMSA-N	CC1=C[C@H](N2C[C@H]1N(C2=O)OCC(=O)O)C(=O)N	150	1
96	BDBM50076760	SFPTWFMBZJRMKD-SDDRHHMPSA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NCC3=NC=CN=C3)C(=O)N2O[C@H](C(=O)O)F	180	1
97	BDBM50076761	SWIHLLNXGOJYRM-FIQHERPVSA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NOC[C@@H]3CCCN3)C(=O)N2O[C@H](C(=O)O)F	210	1
98	BDBM50076762	XUUNBOUUAUVVDD-CUWOBHIPSA-M	C[C@H]1C2[C@H](C(=O)N2C(=C1OC)C(=O)[O-])[C@@H](C)O	230	1
99	BDBM50076763	SOQURWNUDXDSTI-YKFGJVJNSA-N	CC1=C[C@H]2CN([C@@H]1C(=O)N)C(=O)N2OC(C)(C(=O)O)F	230	1
100	BDBM50076764	QAABSWVJCGWPMP-ITHRYJQJSA-N	CCC/C(=C/1\C(N2[C@H](O1)CC2=O)C(=O)O)/C(=O)O	254	1
101	BDBM50076765	KMIMTPSCZSJCPM-LPEHRKFASA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NCCCS(=O)(=O)N)C(=O)N2O[C@H](C(=O)O)F	280	1
102	BDBM50076766	RROHCWIBSPHGHX-PTRXPTGYSA-N	C[C@@@]1([C@@@H](N2[C@H](S1(=O)=O)CC2=O)C(=O)O)CSC3=NC=CN3C	289	1
103	BDBM50076767	MWZMTCLBFOZXKV-GPCLJDEGSA-M	C[C@@@]1([C@@@H](C2C(S1(=O)=O)CC2=O)C(=O)[O-])CN3C(=CN=N3)C(=O)OC	300	2
104	BDBM50076768	CJHINZCBCHTSOE-PRFJFZKWSA-M	C[C@@@]1([C@@@H](C2C(S1(=O)=O)CC2=O)C(=O)[O-])CN3C(=C(N=N3)C(=O)OC)C(=O)OC	300	2
105	BDBM50076769	LTSHHIPSCFMNJ-LZUCFRIBSA-M	CC1([C@@H](N2[C@H](S1(=O)=O)/C(=C\C(=O)OC)/C2=O)C(=O)[O-])C	308	2
106	BDBM50076770	CBBUHGXUJWGHXR-INTQDDNPSA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NCC3=NC=CO3)C(=O)N2O[C@H](C(=O)O)F	310	2
107	BDBM50076771	AGRPURZKBDMSSM-PTNRUYDSSA-M	CCOC(=O)C1=CN=NN1C[C@]2([C@@H](C3C(S2(=O)=O)CC3=O)C(=O)[O-])C	340	2
108	BDBM50076772	OWSYJEORPQIBAC-IGVIJBNSA-L	CC(=O)OCC1=C(N2[C@@H](/C(=C\C(=O)[O-])/C2=O)S(=O)(=O)C1)C(=O)[O-]	400	2
109	BDBM50076773	FJEAGJSMLOQFME-MJSMECKXSA-N	COC(=O)C1/C(=C/C(=O)O)/O[C@H]2N1C(=O)C2/C=C/C3=CC=CC=C3	408	2
110	BDBM50076774	TXPJIXCYUFODBL-WWTCDDSASA-N	C1[C@@H]2N(C1=O)C(/C(=C/C(=O)O)/O2)(C/C=C/C3=CC=CC=C3)C(=O)OCC4=CC=CC=C4	426	2
111	BDBM50076775	RUDGJWAKUPBPKU-XPUUQOCRSA-N	CC1=C[C@H]2CN([C@@H]1C(=O)N)C(=O)N2OCC(=O)O	430	2
112	BDBM50076776	FKENQMMABCRJMK-RITPCOANSA-N	CC1([C@@H](N2[C@H](S1(=O)=O)CC2=O)C(=O)O)C	557	2
113	BDBM50076777	GIFHLGGIPIAVNV-RMVUORRVSAL	C[C@@@]1([C@@@H](C2C(S1(=O)=O)CC2=O)C(=O)[O-])CN3C=C(N=N3)C(=O)[O-]	600	2
114	BDBM50076778	QEYUTORLTFVDOI-VNHYZAJKSA-N	CC1([C@@H](N2[C@H](S1(=O)=O)[C@@H](C2=O)CC3=CN(N=N3)C4=CC=C(C=C4)N)C(=O)O)C	600	2
115	BDBM50076779	CRIIDFLOBDCDK-IESOIDBDSA-N	COC(=O)C1/C(=C/C=C/C2=CC=CC=C2)\C(=O)/O[C@H]3N1C(=O)C3	639	2
116	BDBM50076780	ODGTYAIFNCCHEW-JCGDXUMPSA-N	CC1([C@@H](N2[C@H](S1(=O)=O)[C@@H](C2=O)CO)C(=O)O)	650	2
117	BDBM50076781	XXDOEORVWMNHLY-LPEHRKFASA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NC3COC3)C(=O)N2O[C@H](C(=O)O)F	660	2
118	BDBM50076782	AVPVJZPIBNHZQN-CAUGXEKISA-M	C[C@@@]1([C@@@H](C2C(S1(=O)=O)CC2=O)C(=O)[O-])CN3C=CN=N3	690	2
119	BDBM50076783	KPAZRDMJBFRIBA-VNHYZAJKSA-N	CC1([C@@H](N2[C@H](S1(=O)=O)[C@@H](C2=O)CC3=CN(N=N3)C4=CC=C(C=C4)O)C(=O)O)C	700	2
120	BDBM50076784	LDDMNOXUAMVHDZ-WMUAKXSPSA-N	C[C@@@]1([C@@@H](C2C(S1(=O)=O)CC2=O)C(=O)O)CN3C=C(N=N3)N	700	2
121	BDBM50076785	XSJFKEVFSFGNJZ-XRRVDJEJSA-N	CC1=C[C@H]2CN([C@@H]1C(=O)NCC[C@@H]3CCC(=O)N3)C(=O)N2O[C@H](C(=O)O)F	730	2
122	BDBM50076787	ITFVZKLBKASFHP-NYXMJBRJSA-M	CC(=O)OCC1=C(N2[C@@H](/C(=C\C3=CC=CC=N3)/C2=O)S(=O)(=O)C1)C(=O)[O-]	800	2
123	BDBM50076788	SUKFPPMYJISQEB-SNVBAGLBSA-M	CC(=O)OCC1=C(N2[C@H](C(=C)C2=O)S(=O)(=O)C1)C(=O)[O-]	800	2
124	BDBM50076789	PODACQWVDNEGKW-WMYCVDESA-N	C1[C@@H]2N(C1=O)C(/C(=C/C(=O)O2)(C/C=C/C)O)C(=O)OCC3=CC=CC=C3	828	2
125	BDBM50076790	RTZMBJMWCNIEYEE-HUCLUGSFSA-N	CC(C)(C)C(=O)OCC(=O)C1(/C(=C/C(=O)O)/O[C@H]2N1C(=O)C2)C/C=C/C3=CC=CC=C3	871	2
126	BDBM50076791	NHCBGBXSGNHZFN-HBNTYKKEA-N	CC1=C[C@H]2CN([C@@H]1CNC(=O)OC(C)(C)C)C(=O)N2O[C@H](C(=O)O)F	940	2
127	BDBM50076792	SPTMQRZXEWVYFQ-LGJWTYRLSA-N	C=CCC1(/C(=C\C(=O)O)/O[C@H]2N1C(=O)C2)C(=O)OCC3=CC=CC=C3	945	2
128	BDBM50076793	JLLMLHDLOMWYNZ-VNHYZAJKSA-N	CC1([C@@H](N2[C@H](S1(=O)=O)[C@@H](C2=O)CC3=CN(N=N3)C4=CC=CC=C4)C(=O)O)C	1000	2
129	BDBM50076794	JYGFUXUBGZBBAC-RCRIAUEHSA-M	C[C@@@]1([C@@@H](N2[C@H](S1(=O)=O)CC2=O)C(=O)[O-])/C=C/C#N	1070	2

130	BDBM50076795	FJEAGJSMLOQFME-MJSMECKXSA-N	COC(=O)C1(/C(=C/C(=O)O)/O[C@H]2N1C(=O)C2)C/C=C/C3=CC=CC=C3	1560	2
131	BDBM50076796	QSGHUBRFRC SOPB-VNHYZAJKSA-N	CC1([C@@H](N2[C@H](S1(=O)=O)CC2=O)C@H](C2=O)CC3=CN(N=N3)C4=CC=C(C=C4)F)C(=O)O)C	1600	2
132	BDBM50076797	DZCHPZBOQLTTII-RMVUORRVSAL	C[C@@]1([C@@H](C2C(S1(=O)=O)CC2=O)C(=O)[O-])CN3C(=CN=N3)C(=O)[O-]	1700	2
133	BDBM50076798	ZUIFBYULLQUXTF-VLCNFMPIYSA-N	C=CC/C(=C/1\C(N2[C@H](O1)CC2=O)C(=O)OCC3=CC=CC=C3)/CO	1950	2
134	BDBM50076799	KBUJSCWFZMYQLR-QQKZXBOMSA-N	C1[C@H]2N(C1=O)C(/C(=C/C(=O)O2)(C/C=C/C3=CC=CC=C3)C(=O)OCC4=CC=CC=C4	2,79E+03	2
135	BDBM50076800	WHIIGHHQQWOKXKEP-KWLOKCHISA-N	CC1([C@@H](N2[C@H](S1(=O)=O)C(C2=O)C/C=C/C(=O)OC)C(=O)O)C	2800	2
136	BDBM50076801	SURSZFGYJRFQEX-IVZWLZJFSA-N	CC1=C[C@H]2CN([C@@H]1CNC(=O)C)C(=O)N2O[C@H](C(=O)O)F	3600	2
137	BDBM50076802	FAKBCYAIPOVOTK-TYJSTWMOSA-N	C1[C@H]2N(C1=O)C(/C(=C(\C/C=C/C3=CC=CC=C3)/CO)/O2)C(=O)OCC4=CC=CC=C4	4120	2
138	BDBM50076803	WKUUKIISZDMTQV-XLPZGREQSA-N	CC1=C[C@H]2CN([C@@H]1CNC(=O)C(F)(F)F)C(=O)N2O[C@H](C(=O)O)F	4200	2
139	BDBM50076804	GWGNBTHOPHGFDX-WXRZDVGUSA-N	C1[C@H]2N(C1=O)C(/C(=C\3/CC(OC3)CI)/O2)C(=O)OCC4=CC=CC=C4	4390	2
140	BDBM50076805	AHWXHWAYVLZNSQ-QHCPKHFHSA-N	C1[C@H]2N(C1=O)C(/C(=C/C(=O)O2)(C/C=C/C3=CC=CC=C3)C(=O)OCC4=CC=CC=C4	4920	2
141	BDBM50076806	DRYKFHDMYZDMRX-UHFFFAOYSA-N	C1=CC(=C(C(=C1)CI)CI)C2=CC(=C(C=C2)C(=O)O)CS	5500	2
142	BDBM50076807	PODACQWVDNEGKW-WMYCVDERSA-N	C1[C@H]2N(C1=O)C(/C(=C/C(=O)O2)(C/C=C\CO)C(=O)OCC3=CC=CC=C3	5640	2
143	BDBM50076808	BGQRDDSRQMAZIR-YDJAWIAZSA-N	CC(=O)OC/C=C\CC1(/C(=C/C(=O)O)[C@H]2N1C(=O)C2)C(=O)OCC3=CC=CC=C3	5900	2
144	BDBM50076809	KKSAGQLHTXTMBB-AOPMUKPMSA-N	CC(=O)C/C=C/C1[C@H]2N(C1=O)[C@H](C(S2(=O)=O)(C)C)C(=O)O	6500	2
145	BDBM50076810	MQNFPHMTROMTJB-JCGDXUMPSA-N	CC1([C@@H](N2[C@H](S1(=O)=O)C@H](C2=O)CS)C(=O)O)C	6800	2
146	BDBM50076811	AVPVJZPIBNHZQN-CAUGXEKISA-M	C[C@@]1([C@@H](C2C(S1(=O)=O)CC2=O)C(=O)[O-])CN3C=CN=N3	7000	2
147	BDBM50076813	BGQRDDSRQMAZIR-YDJAWIAZSA-N	CC(=O)OC/C=C\CC1(/C(=C/C(=O)O)[C@H]2N1C(=O)C2)C(=O)OCC3=CC=CC=C3	7310	2
148	BDBM50076814	ANCXGKPUKGCZNAI-OOPHUOBASA-N	CC1([C@@H](N2[C@H](S1(=O)=O)[C@H](C2=O)CC3=CN(N=N3)CC4CCCN4)C(=O)O)C	7900	2
149	BDBM50076815	ODGTYAIFNCCEHW-KZLJYQGOSAN	CC1([C@@H](N2[C@H](S1(=O)=O)[C@H](C2=O)CO)C(=O)O)C	8000	2
150	BDBM50076816	ZUIFBYULLQUXTF-VLCNFMPIYSA-N	C=CC/C(=C/1\C(N2[C@H](O1)CC2=O)C(=O)OCC3=CC=CC=C3)/CO	8020	2
151	BDBM50076817	ZLZSKFBDQKKT SB-SNVBAGLBSA-M	CC(=O)OCC1=C(N2[C@H](C(=C(CI)CI)C2=O)S(=O)(=O)C1)C(=O)[O-]	8300	2
152	BDBM50076818	RNEZBAWBDZGLDN-CGLKDYMESAN	C=CC/C(=C/1\C(N2[C@H](O1)CC2=O)(CC=C)C(=O)OCC3=CC=CC=C3)/C=O	8390	2
153	BDBM50076819	KKSAGQLHTXTMBB-XFDVOC DVSA-N	CC(=O)C/C=C/C1[C@H]2N(C1=O)[C@H](C(S2(=O)=O)(C)C)C(=O)O	8400	2
154	BDBM50076820	WHIIGHHQQWOKXKEP-ASQRRCTNSA-N	CC1([C@@H](N2[C@H](S1(=O)=O)C(C2=O)C/C=C/C(=O)OC)C(=O)O)C	8600	2
155	BDBM50076822	ZCVZNXFIILRURF-GOZTUDAPSA-M	C[C@H]([C@@H]1[C@H]2N(C1=O)C(=C(S2(=O)=O)C(C)C)C(=O)[O-])O	10000	2
156	BDBM50076823	NODZXNWCPAUXPG-IDLZWDITSA-N	CC1([C@@H](N2[C@H](S1=O)C(C2=O)C/C=C/C(=O)N(C)OC)C(=O)O)C	12000	2
157	BDBM50076824	APLLQHCD OOGSFW-OQAFJANISAN	CC1([C@@H](N2[C@H](S1(=O)=O)C(C2=O)C/C=C/C(=O)C(=O)O)C	12000	2
158	BDBM50076825	WEJZNOFAZVDYFB-FOXUPHSNSAN	COC(=O)C1(/C(=C\C(=O)O)/O[C@H]2N1C(=O)C2)CC=C	13300	2
159	BDBM50076826	83173-24-4	CC/C=C\C/C=C\C/C=C\C\CCCCCCCC1=C(C(=CC=C1)O)C(=O)O	13400	2
160	BDBM50076827	ODGTYAIFNCCEHW-KZLJYQGOSAN	CC1([C@@H](N2[C@H](S1(=O)=O)[C@H](C2=O)CO)C(=O)O)C	14600	2
161	BDBM50076828	NRHINWWLSKJJNI-NZHQHIMKSA-N	CC1([C@@H](N2[C@H](S1(=O)=O)[C@H](C2=O)CC3=CN(N=N3)CC(CC4=CC=C(C=C4)O)N)C(=O)O)C	15000	2
162	BDBM50076829	FPJUOUJTKUOLCV-ISDAKNICSA-M	CCCC1(/C(=C\CO)/O[C@H]2N1C(=O)C2)C(=O)[O-]	17800	2
163	BDBM50076830	VW SIAIDFEKNQNE-SRNKDFPPSA-N	COC(=O)C1(/C(=C\C(=O)O)/O[C@H]2N1C(=O)C2)C/C=C/C3=CC=CC=C3	19600	2
164	BDBM50076831	YOCVKISIBJBEEES-ROCHPLJESAM	CC(=O)OCC1=C(N2[C@H](C(=C/C3=CC=CC=C3)/C2=O)S(=O)(=O)C1)C(=O)[O-]	20000	2
165	BDBM50076832	YOCVKISIBJBEEES-VXXVJXPCSAM	CC(=O)OCC1=C(N2[C@H](C(=C\C3=CC=CC=C3)/C2=O)S(=O)(=O)C1)C(=O)[O-]	20000	2
166	BDBM50076833	ZSPQUKSEVIUMSR-KNVNGTOESAM	CC(=O)/C=C/1\C@H]2N(C1=O)C(=C(CS2)COC(=O)C)C(=O)[O-]	20000	2
167	BDBM50076834	HFICAXNAWUOCGS-XJQCMFKW SAMA	CC(=O)OCC1=C(N2[C@H](C(=C=CC(C)(C)C)C2=O)S(=O)(=O)C1)C(=O)[O-]	20000	2
168	BDBM50076835	GGHHCDXUXTXKBM-OMPYUXFZSA-M	CC(=O)OCC1=C(N2[C@H](C(=C\C(C)(C)C)/C2=O)S(=O)(=O)C1)C(=O)[O-]	20000	2
169	BDBM50076836	KKHROROKABO CRS-XTXQVDMPSA-M	CC(=O)OCC1=C(N2[C@H](C(=C\C(=O)OC(C)C)/C2=O)SC1)C(=O)[O-]	20000	2
170	BDBM50076837	BIUVASJLHVNN EP-GPWUBDCCSAMA	CC(=O)OCC1=C(N2[C@H](C(=C\C(CO)/C2=O)S(=O)(=O)C1)C(=O)[O-]	20000	2
171	BDBM50076838	IAYUXXWVOBMOKN-SNVBAGLBSAMA	CC(=O)OCC1=C(N2[C@H](C(=C(Br)Br)C2=O)SC1)C(=O)[O-]	20000	2
172	BDBM50076839	VGEKZJB JPIFSJW-VXXVJXPCSAM	CC(=O)OCC1=C(N2[C@H](C(=C\C3=CC=CC=C3)/C2=O)SC1)C(=O)[O-]	20000	2
173	BDBM50076840	MDAINRPXGSCIOC-IGVIJBNESAL	CC(=O)OCC1=C(N2[C@H](C(=C\C(=O)O)/C2=O)SC1)C(=O)[O-]	20000	2

174	BDBM50076841	BQYSWXZGFUSYTK-RWTFEVOBSA-M	CC1([C@@H](N2[C@H](S1)/C=C\C(=O)OC(C)(C)C)/C2=O)C(=O)[O-]C	20000	2
175	BDBM50076842	ZLYQUOKXGNQRIC-LLVKDONJSA-M	CC(=O)OCC1=C(N2[C@H](C(=C=C)C2=O)S(=O)(=O)C1)C(=O)[O-]	20000	2
176	BDBM50076843	WMEHQPKGOVJLPV-LZUCFRIBSA-M	CC1([C@@H](N2[C@H](S1)/C=C\C(=O)OC)/C2=O)C(=O)[O-]C	20000	2
177	BDBM50076844	AOZHCLHLXLZFCC-CPWLGJMPSA-M	CC(=O)OCC1=C(N2[C@H](/C=C\C(=O)OC)/C2=O)SC1C(=O)[O-]	20000	2
178	BDBM50076845	BJXIZXYMGFDKBP-YRNGGQEVSA-L	CC1([C@@H](N2[C@H](S1)/C=C\C(=O)[O-])/C2=O)C(=O)[O-]C	20000	2
179	BDBM50076846	KKHROROKABO CRS-XTXQVDMPSA-M	CC(=O)OCC1=C(N2[C@H](/C=C\C(=O)OC(C)(C)C)/C2=O)SC1C(=O)[O-]	20000	2
180	BDBM50076847	IWSSJMHCOIGHBY-CYZOPGGISA-M	CC(=O)OCC1=C(N2[C@H](/C=C/C/Br)/C2=O)S(=O)(=O)C1)C(=O)[O-]	20000	2
181	BDBM50076848	PXOKVALSRAGFAH-SNVBAGLBSA-M	CC(=O)OCC1=C(N2[C@H](C(=C(C)Br)C2=O)S(=O)(=O)C1)C(=O)[O-]	20000	2
182	BDBM50076849	SPOYTMWGOMJDIX-SNVBAGLBSA-M	CC(=O)OCC1=C(N2[C@H](C(=C(C)Cl)C2=O)SC1)C(=O)[O-]	20000	2
183	BDBM50076850	LEADWAGQWAFRFE-CYZOPGGISA-M	CC(=O)OCC1=C(N2[C@H](/C=C/C/Br)/C2=O)SC1C(=O)[O-]	20000	2
184	BDBM50076851	GVESJWVLYJUUMT-OMPYUXFZSA-M	CC(=O)OCC1=C(N2[C@H](/C=C\C(C)(C)C)/C2=O)SC1C(=O)[O-]	20000	2
185	BDBM50076852	ADMWSBHYYMUDFV-LGJWTYRLSA-N	C=CCC1(/C=C\C(O)/O[C@H]2N1C(=O)C2)C(=O)OCC3=CC=CC=C3	23600	2
186	BDBM50076853	HXCNDHVLZREMRV-QYJMCNQRSA-N	CC1([C@@H](N2[C@H](S1(=O)=O)C(C2=O)C/C=C#N)C(=O)O)C	24000	2
187	BDBM50076854	DJIMYYWAZIOWRD-JCGDXUMPSA-M	CC1([C@@H](N2[C@H](S1)[C@@H](C2=O)CO)C(=O)[O-])C	31000	2
188	BDBM50076855	KYTCAZSHGDLTNM-UHFFFAOYSA-N	C1=CC(=CC=C1C2=CC(=C(C=C2)C(=O)O)CS)F	32800	2
189	BDBM50076856	TXPJXCYUFODBL-WWTCDDSSASA-N	C1[C@@H]2N(C1=O)C(/C=C/C(=O)O)/O2)/C(C=C/C3=CC=CC=C3)C(=O)OCC4=CC=CC=C4	45700	2
190	BDBM50076857	JWBKZXUHEFRKRV-FZGKIIMSASA-M	C[C@H]([C@@H]1[C@@H]2N(C1=O)[C@H](C(S2(=O)=O)(C)C)C(=O)[O-])O	47900	2
191	BDBM50076858	YJWBOGUQTYCPBN-GOZTUDAPSA-M	C[C@H]([C@@H]1[C@@H]2N(C1=O)C(=C(S2)C(C)(C)C)C(=O)[O-])O	50000	2
192	BDBM50076859	MQNFPMTROMTJB-KZLJYQGOSA-N	CC1([C@@H](N2[C@H](S1(=O)=O)[C@H](C2=O)CS)C(=O)O)C	51700	2
193	BDBM50076860	DJIMYYWAZIOWRD-KZLJYQGOSA-M	CC1([C@@H](N2[C@H](S1)[C@H](C2=O)CO)C(=O)[O-])C	59000	2
194	BDBM50076861	AFPLSUQTYBHSDG-RVJQKOHUSA-N	CC1([C@@H](N2[C@H](S1(=O)=O)[C@H](C2=O)S)C(=O)O)C	69200	2
195	BDBM50076862	GSVZWBPOOFMGAJ-PISQCCPLSA-M	C[C@H]([C@@H]1[C@@H]2N(C1=O)C(=C(S2=O)C(C)(C)C)C(=O)[O-])O	90000	2
196	BDBM50076863	VNFWGXIXPWXBBS-XLPZGREQSA-N	CC1=C[C@H]2CN([C@@H]1CN)C(=O)N2O[C@H](C(=O)O)F	100000	2
197	BDBM50076864	MMKATEMYLYLBTG-RFFQBLADSA-M	CCC/C(=C/1\C(N2[C@H](O1)CC2=O)C(=O)[O-])/CO	118000	2
198	BDBM50076865	22916-47-8	C1=CC(=C(C=C1Cl)Cl)COC(CN2C=CN=C2)C3=C(C=C(C=C3)Cl)Cl	125000	2
199	BDBM50076866	JWBKZXUHEFRKRV-SKHQTKALSA-M	C[C@H]([C@H]1[C@@H]2N(C1=O)[C@H](C(S2(=O)=O)(C)C)C(=O)[O-])O	204000	2
200	BDBM50076867	DJIMYYWAZIOWRD-KZLJYQGOSA-N	CC1([C@@H](N2[C@H](S1)[C@H](C2=O)CO)C(=O)O)C	275000	2
201	BDBM50076869	JWBKZXUHEFRKRV-VUBPFVNUUSA-M	C[C@H]([C@@H]1[C@@H]2N(C1=O)[C@H](C(S2(=O)=O)(C)C)C(=O)[O-])O	334000	2
202	BDBM50076870	JQVSNQBXQNNKIW-WYDQCIBASA-M	C[C@H]([C@@H]1[C@@H]2CC(=C(N2C1=O)C(=O)[O-])OC)O	340000	2
203	BDBM50076871	RHBQBGRVEOOJQL-JCGDXUMPSA-N	CC1([C@@H](N2[C@H](S1)[C@H](C2=O)CS)C(=O)O)C	601000	2
204	BDBM50076872	RHBQBGRVEOOJQL-KZLJYQGOSA-N	CC1([C@@H](N2[C@H](S1)[C@H](C2=O)CS)C(=O)O)C	648000	2
205	BDBM50076873	DJIMYYWAZIOWRD-JCGDXUMPSA-M	CC1([C@@H](N2[C@H](S1)[C@H](C2=O)CO)C(=O)O)C	752000	2
206	BDBM50076874	WOYLBXUCZPJYOV-RVJQKOHUSA-N	CC1([C@@H](N2[C@H](S1)[C@H](C2=O)S)C(=O)O)C	988000	2
207	BDBM50076875	WOYLBXUCZPJYOV-ALEPSDHESA-N	CC1([C@@H](N2[C@H](S1)[C@H](C2=O)S)C(=O)O)C	1000000	2

Modelo in silico con actividad contra la betalactamasa TEM-1 presente en Escherichia coli.

Resumen

La resistencia bacteriana de Escherichia Coli está dada principalmente por betalactamasas, siendo una de las más importantes TEM-1. Para minimizar los obstáculos de métodos tradicionales, se ha propuesto el uso de métodos in silico como QSAR para la búsqueda de nuevos medicamentos. En este estudio, se trabajó sobre 207 moléculas recolectadas de BindingDB, cada una con 977 descriptores calculados mediante Dragon 7.0. Aplicando algoritmos genéticos se seleccionaron los mejores descriptores adaptados a métodos de clasificación (kNN, N3, BNN, CART) para lo que se crearon 2 clases de actividad cuyo umbral es la mediana de los valores de IC50 y se clasificaron 102 moléculas como “muy activas” y 105 moléculas como “menos activas”, con los que se obtuvieron modelos con buena capacidad predictiva. Los resultados demuestran que los modelos QSAR de clasificación (kNN, N3, BNN y CART) son capaces de distinguir entre moléculas activas y menos activas contra la betalactamasa TEM-1 de Escherichia coli, con un alto poder predictivo tanto interna como externamente.

Palabras clave: Escherichia Coli, TEM-1, betalactamasa, QSAR, resistencia bacteriana.