

Características moleculares de inhibidores de corrosión: Enfoque estadístico y computacional

Molecular characteristics of corrosion inhibitors: Statistical and computational approach

U. J. Rangel-Peña ^a, J. M. Vásquez-Pérez ^b, R. L. Camacho-Mendoza ^a, L. A. Zárate-Hernández ^a,
S. González-Montiel ^a, J. Cruz-Borbolla ^{a,*}

^a Área Académica de Química, Universidad Autónoma del Estado de Hidalgo, 42184, Mineral de la Reforma, Hidalgo, México.

^b Investigador SECIHTI, Universidad Autónoma del Estado de Hidalgo, 42184, Mineral de la Reforma, Hidalgo, México.

Resumen

En el presente estudio se desarrolló un modelo de clasificación basado en aprendizaje automático para discriminar compuestos orgánicos según su eficiencia inhibitoria contra la corrosión de acero en medio ácido. Una base de datos de 117 inhibidores fue construida, de la cual se emplearon 73 compuestos nitrogenados caracterizados mediante descriptores moleculares 2D y parámetros termodinámicos calculados por DFT (B3LYP/6-311++G**). Del análisis estadístico no paramétrico se identificaron diferencias significativas entre buenos ($\theta \geq 90\%$) y malos inhibidores, lo cual reveló once descriptores discriminantes relacionados con tamaño molecular, complejidad estructural y flexibilidad conformacional. Mediante la metodología de selección "step forward" se identificó al poder electrodonador y la entalpía como los descriptores óptimos para la clasificación. El modelo Random Forest logró una exactitud del 73 % en la clasificación binaria de la eficiencia inhibitoria, demostrando la capacidad del enfoque para capturar relaciones no lineales complejas. Este método computacional proporciona una herramienta valiosa para la discriminación y selección racional de inhibidores de corrosión en medio ácido.

Palabras Clave: Corrosión, Aprendizaje Automático, QSAR, Random Forest, Clasificación.

Abstract

In this study, a machine learning-based classification model was developed to discriminate organic compounds according to their inhibitory efficiency against steel corrosion in acidic medium. A database of 117 inhibitors was constructed, from which 73 nitrogen-containing compounds were employed, characterized by 2D molecular descriptors and thermodynamic parameters calculated by DFT (B3LYP/6-311++G**). From the non-parametric statistical analysis, significant differences were identified between good ($\theta \geq 90\%$) and poor inhibitors, which revealed eleven discriminant descriptors related to molecular size, structural complexity, and conformational flexibility. Through the "step forward" selection methodology, the electrodonating power and the enthalpy were identified as the optimal descriptors for classification. The Random Forest model achieved 73% accuracy in the binary classification of inhibitory efficiency, demonstrating the approach's capability to capture complex non-linear relationships. This computational method provides a valuable tool for the discrimination and rational selection of corrosion inhibitors in acidic medium.

Keywords: Corrosion, Machine Learning, QSAR, Random Forest, Classification.

*Autor para la correspondencia: jrcruz@uaeh.edu.mx

Correo electrónico: ra351708@uaeh.edu.mx (Uriel Josafat Rangel-Peña), josemanuel_vasquez@uaeh.edu.mx (José Manuel Vásquez-Pérez), rosa_camacho@uaeh.edu.mx (Rosa Luz Camacho-Mendoza), luis_zarate@uaeh.edu.mx (Luis Ángel Zárate-Hernández), gmontiel@uaeh.edu.mx (Simplicio González-Montiel), jrcruz@uaeh.edu.mx (Julián Cruz-Borbolla)

Historial del manuscrito: recibido el 09/07/2025, última versión-revisada recibida el 03/11/2025, aceptado el 09/11/2025, publicado el 12/12/2025. DOI: <https://doi.org/10.29057/icbi.v13iEspecial4.15476>



1. Introducción

En México, la industria del hierro y el acero ha desempeñado un papel fundamental en la economía del país, contribuyendo, al cuarto trimestre de 2023, con aproximadamente el 20 % del Producto Interno Bruto (PIB), lo que refleja la influencia del sector en el desarrollo económico del país (Instituto Nacional de Estadística y Geografía (INEGI), 2023; Makrini *et al.*, 2016; Secretaría de Economía, 2025). Sin embargo, la industria del acero se enfrenta a desafíos considerables, siendo uno de los más destacados el impacto de la corrosión en la durabilidad de estructuras metálicas utilizadas en diversas aplicaciones industriales (Contreras, Quej, Liu, Sosa, & Alamilla, 2020; Landolfo, Cascini, & Portioli, 2010).

La corrosión en ambientes ácidos presenta desafíos particulares, debido a que puede provocar un deterioro considerable de los materiales metálicos. Se estima que la implementación de tecnologías para prevenir o mitigar la corrosión podría reducir entre un 25 % y un 30 % las pérdidas económicas anuales ocasionadas por este fenómeno (Contreras *et al.*, 2020; Khan *et al.*, 2015).

En respuesta a este desafío, han surgido los inhibidores de la corrosión como una estrategia eficaz y económicamente viable para proteger los metales y aleaciones, particularmente en entornos ácidos (Ayati *et al.*, 2011). Los inhibidores, ya sean de origen natural o sintético, presentan una variedad de grupos funcionales, como enlaces dobles o triples y heteroátomos como nitrógeno, oxígeno o azufre. La eficiencia de los inhibidores se encuentra estrechamente relacionada con su estructura molecular y su capacidad para interactuar con la superficie metálica. En este sentido, destacan compuestos como azoles, diazoles y triazoles con grupos hidrofóbicos terminales, donde la deslocalización electrónica (anillos conjugados de cinco o seis miembros) mejora la interacción metal-inhibidor (Obayes *et al.*, 2017; Tansuğ *et al.*, 2014).

Es importante señalar que en la literatura los datos reportados como son: concentración del medio corrosivo, dosificación de inhibidores y temperatura, varían significativamente entre los estudios publicados, lo que limita comparaciones directas (Esmailzadeh, Aliofkhazraei, & Sarlak, 2018; Kutz, 2018; Rammelt & Reinhard, 1992). Por otra parte, técnicas como la gravimetría, polarización potenciodinámica y espectroscopia de impedancia electroquímica (EIS, por sus siglas en inglés) son ampliamente utilizadas, sin embargo, no existe un protocolo estandarizado que unifique los criterios experimentales.

La química computacional ofrece herramientas que pueden superar las limitaciones derivadas de las diversas técnicas experimentales para analizar los diversos inhibidores de la corrosión. Los modelos basados en inteligencia artificial (machine learning, deep learning), permiten correlacionar descriptores moleculares con la propiedad específica, como la toxicidad (Rangel-Peña *et al.*, 2023). Los descriptores moleculares se clasifican tradicionalmente en tres categorías principales, según su nivel de complejidad y representación:

- Descriptores 1D: son los más simples y se derivan directamente del conteo de átomos, considerando únicamente la fórmula molecular de la sustancia.
- Descriptores 2D: se basan en algoritmos aplicados a representaciones topológicas, estos descriptores analizan los fragmentos estructurales presentes en la molécula.
- Descriptores 3D: son más complejos que los anteriores, se obtienen a partir de representaciones geométricas que incorporan información sobre la configuración espacial de la molécula.

Entre las herramientas más utilizadas para generar estos descriptores destaca el software Dragon, reconocido por su capacidad para producir descriptores moleculares a partir de conjuntos de estructuras moleculares (Mauri, Consonni, Pavan, Todeschini, & others, 2006; Morell, Grand, & Toro-Labbé, 2005; Todeschini, Lasagni, & Marengo, 1994).

La variabilidad en los datos experimentales de corrosión, derivada de diferencias metodológicas, fluctuaciones ambientales y la diversidad estructural de los inhibidores, justifica el uso de enfoques estadísticos robustos que superen las limitaciones de los métodos paramétricos tradicionales (Kaur, Stoltzfus, & Yellapu, 2018; Kubisztal, Kubisztal, & Haneczok, 2020; Nick, 2007; Peddeferri, 2018). Las pruebas no paramétricas y técnicas de remuestreo (bootstrapping) son empleadas como alternativas eficaces para analizar datos con distribuciones no normales, tamaños de muestra pequeños o heterocedasticidad (Hesterberg, 2011). Por ejemplo, la prueba de Wilcoxon facilita la comparación de la eficiencia entre inhibidores de distintas familias químicas en medios ácidos, considerando datos asimétricos o con valores atípicos (Divine, Norton, Hunt, & Dienemann, 2013). El bootstrapping permite estimar intervalos de confianza para parámetros críticos sin depender de supuestos sobre la distribución de los datos, una ventaja clave al integrar resultados de múltiples estudios. Estas herramientas no solo cuantifican la incertidumbre en mediciones como las obtenidas por espectroscopía de impedancia electroquímica (EIS) o gravimetría, sino que también revelan correlaciones significativas entre descriptores moleculares (como el conteo de heteroátomos o la topología 2D) y la actividad inhibitoria. Al utilizar métodos que evitan supuestos restrictivos, este enfoque garantiza la consistencia de las tendencias observadas y establece un marco reproducible para interpretar datos complejos y dispersos, característicos de los estudios de corrosión metálica (Boada, Leon, & Ortiz, 2025; Diccio & Romano, 1988; Sultan, Mashrei, & Washer, 2020).

En este trabajo se presenta, un enfoque integrado que combina estadística no paramétrica y modelado computacional para identificar qué factores estructurales son clave en el clasificar inhibidores de corrosión para hierro y acero, implementado una metodología en etapas: en primer lugar, se construye una base de datos estandarizada mediante la recopilación de datos experimentales publicados en bases de datos, integrando mediciones de técnicas como EIS. Posteriormente, se aplican pruebas de Wilcoxon y técnicas de bootstrapping para identificar descriptores moleculares con capacidad discriminativa entre inhibidores eficientes e ineficientes, estableciendo umbrales cuantitativos específicos para distintas familias químicas. Finalmente, se determinaron las propiedades estructurales y electrónicas, mediante cálculos

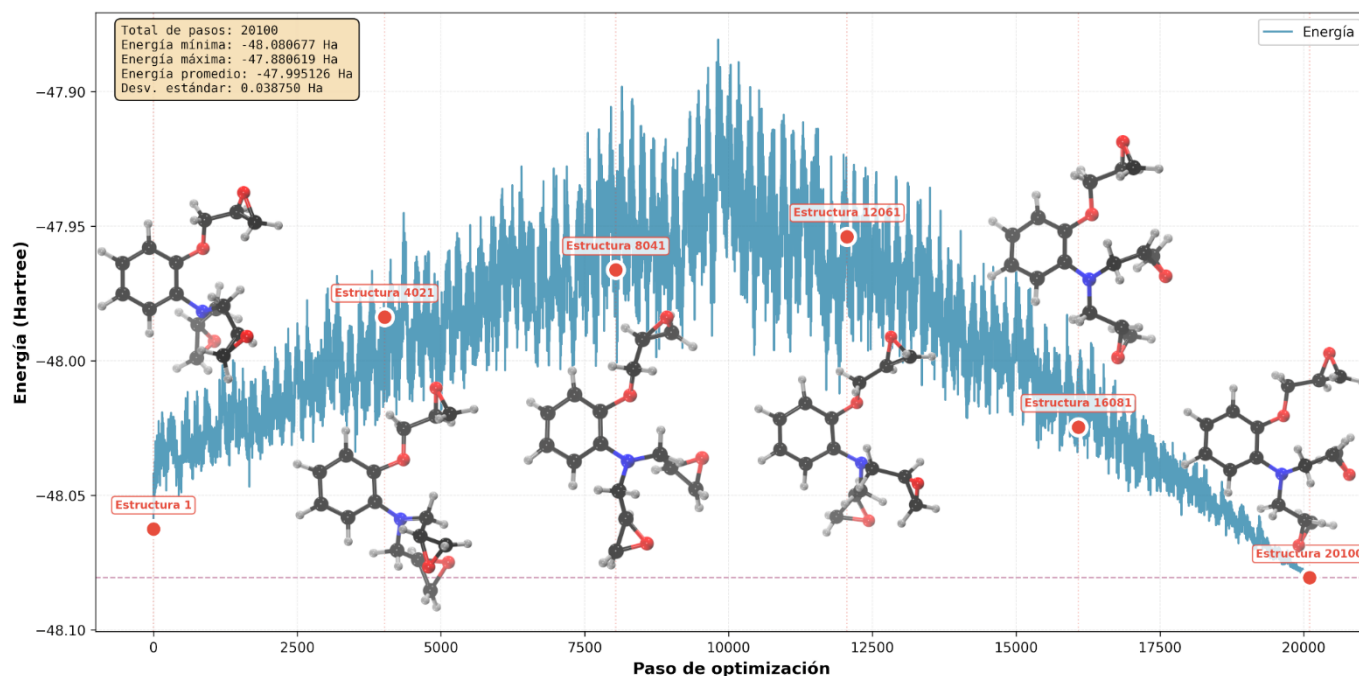


Figura 1. Perfil energético de recocido molecular para la molécula 070.

DFT a nivel B3LYP/6-311++G** + LANL2DZ para los inhibidores de corrosión para hierro y acero.

2. Metodología

2.1. Selección y construcción de la base de datos

Se realizó una búsqueda sistemática de publicaciones científicas sobre inhibidores de la corrosión para hierro y acero en medio ácido, enfocándose en soluciones de HCl (0.5 – 2.0 M) y H₂SO₄ (0.1 – 1.0 M) como los entornos corrosivos más representativos de aplicación industrial. La revisión bibliográfica mostró un desafío crítico debido a la heterogeneidad en los datos experimentales reportados. Considerando que la eficiencia inhibitoria se reporta habitualmente como porcentaje de inhibición (θ %) determinado principalmente por espectroscopía de impedancia electroquímica (EIS), en este trabajo se priorizaron los datos obtenidos mediante esta técnica, suplementados con métodos alternativos cuando la información por EIS no estaba disponible. Las variables experimentales documentadas incluyen la concentración del medio corrosivo, concentración del inhibidor, temperatura y composición porcentual de la superficie metálica.

2.2. Clasificación estructural de los inhibidores

Para agrupar los inhibidores según sus características estructurales, se implementó una clasificación sistemática en cuatro categorías: 1) Heterociclos Nitrogenados (azoles, pirimidinas, quinolinas, quinoxalinas, indoles, benzimidazoles), 2) Heterociclos con Azufre u Oxígeno (furanos, oxazoles, tiazoles), 3) Derivados Nitrogenados Acíclicos (iminas, hidrazonas, hidrazinas tioureas, tioamidas), y 4) Compuestos con Grupos Funcionales Oxigenados (ácidos carboxílicos, alcoholes, ésteres). Para moléculas

multifuncionales se aplicó un algoritmo de decisión jerárquico basado en dos reglas: prioridad estructural (Heterociclos > Grupos Nitrogenados > Grupos Oxigenados) y complejidad (la subestructura más compleja tiene mayor prioridad).

2.3. Análisis conformacional y optimización estructural

La diversidad conformacional se exploró mediante simulaciones de dinámica molecular aplicando el algoritmo de recocido simulado en DFTB+ con el método DFTB-SCC (aproximación de tercer orden) y el conjunto de parámetros 3ob (Hourahine *et al.*, 2020). El perfil térmico consistió en: etapa inicial a 300 K (100 pasos), calentamiento lineal hasta 1000K (10,000 pasos), y enfriamiento hasta 10 K (10,000 pasos), como se muestra para la molécula 070 en la Figura 1. Los conformeros de mínima energía fueron optimizados mediante cálculos B3LYP/6-311++G**, utilizando el pseudopotencial LANL2DZ para los compuestos con yodo (045 y 090) (Curtiss, Redfern, Rassolov, Kedziora, & Pople, 2001; Hay & Wadt, 1985; Lee, Yang, & Parr, 1988). A partir de las estructuras optimizadas se calcularon los descriptores globales de reactividad (potencial químico μ , dureza η , blandura S , electrofilicidad ω , electronegatividad χ , poder electrodonador y electroatractor), asimismo, se realizó el análisis termodinámico a 298 K determinando: entalpía, entropía y energía libre.

2.4. Cálculo de descriptores moleculares y análisis estadístico

Los conformeros optimizados fueron utilizados para calcular los descriptores moleculares 2D mediante el software Dragon, incluyendo parámetros topológicos, constitucionales y geométricos. Los compuestos se categorizaron según su eficiencia inhibitoria: “buenos inhibidores” ($\theta \geq 0.90$) y “malos inhibidores” ($\theta < 0.9$). Considerando la heterogeneidad de la base de datos, se implementó un enfoque no paramétrico

Tabla 1. Descriptores con capacidad discriminante significativa para compuestos nitrogenados.

Descriptor	Umbral	Exactitud	Sensibilidad	Especificidad
Enlaces sin hidrógeno	< 13.5	0.34	0.71	0.61
Número de átomos	> 20.0	0.31	0.78	0.58
Número de átomos sin H	> 12.5	0.32	0.73	0.61
Número de carbonos	> 9.5	0.34	0.71	0.61
Número de enlaces	> 20.5	0.3	0.78	0.61
Número de enlaces rotables	> 1.5	0.34	0.71	0.61
Número de hidrógenos	> 7.5	0.32	0.78	0.56
Perímetro del anillo	> 13.0	0.39	0.37	0.89
Peso molecular	> 161.7	0.29	0.83	0.58
Tamaño total del anillo	> 13.5	0.39	0.37	0.89
Volumen de Van der Waals	> 12.9	0.29	0.80	0.61

basado en la prueba de Wilcoxon ($p < 0.05$) para identificar descriptores con diferencias estadísticamente significativas entre grupos. Para los descriptores se determinaron los valores de umbral óptimo mediante bootstrapping, evaluando su robustez como predictores a través de matrices de confusión.

2.5. Desarrollo y validación de modelos de clasificación

Se desarrollaron modelos de clasificación utilizando cinco algoritmos de aprendizaje automático (CART, Random Forest, AdaBoost, KNN y Ridge) implementados en el paquete *caret* de R (Kuhn, 2008). La selección de variables empleó un método de exploración incremental “step forward”, iniciando con combinaciones de dos variables y agregando progresivamente nuevas variables únicamente cuando se observaba mejora en el rendimiento. La evaluación se realizó mediante exactitud, precisión, recuperación, F1-score y AUC-ROC, implementando validación cruzada k-fold ($k = 10$) con 10 repeticiones independientes y división de datos 80 %/20 % para validación interna/externa, minimizando el sobreajuste y garantizando la robustez estadística.

3. Resultados y discusión

3.1. Caracterización y clasificación estructural

La búsqueda sistemática de los inhibidores de corrosión permitió construir una base de datos de 117 moléculas para sistemas hierro/acero en medio ácido, generando, de manera general, 4 grupos de familias. Las estructuras se muestran en la Figura 2 y en las figuras del apéndice. La clasificación estructural (Figura 3) reveló predominancia de heterociclos y derivados nitrogenados acíclicos, proporcionando una base robusta para el análisis estructura-actividad.

3.2. Análisis estadístico discriminante

3.2.1. Compuestos con azufre y oxígeno

Los compuestos con azufre/oxígeno no mostraron diferencias estadísticamente significativas entre buenos ($\theta\% \geq$

0.90) y malos inhibidores ($\theta\% < 0.90$), sugiriendo que su eficiencia inhibitoria depende de factores estructurales más complejos, interacciones específicas superficie-inhibidor, o mecanismos dependientes de la química superficial no capturados por descriptores 1D-2D convencionales.

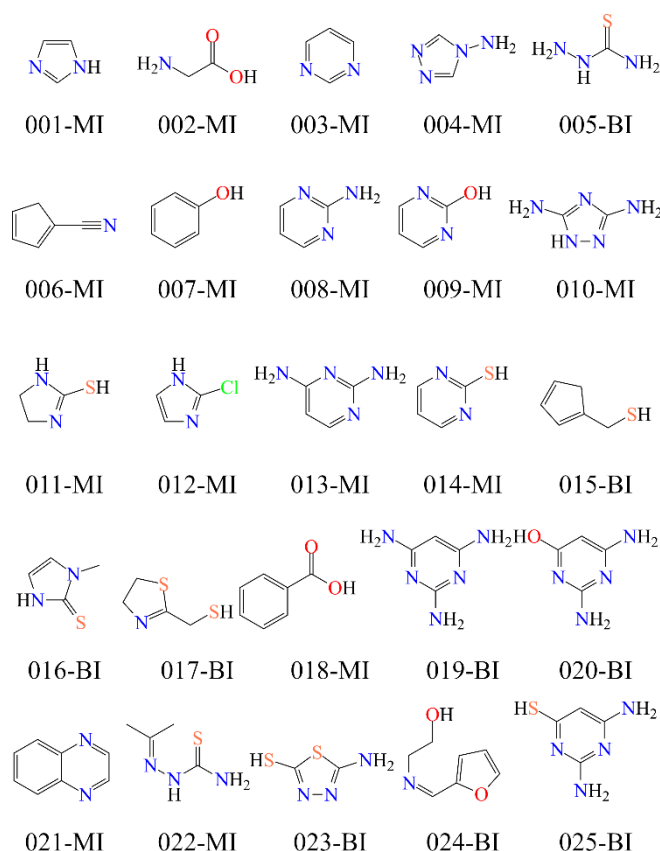


Figura 2. Estructura en 2D de las primeras 25 moléculas estudiadas. MI, Mal Inhibidor; BI, Buen Inhibidor.

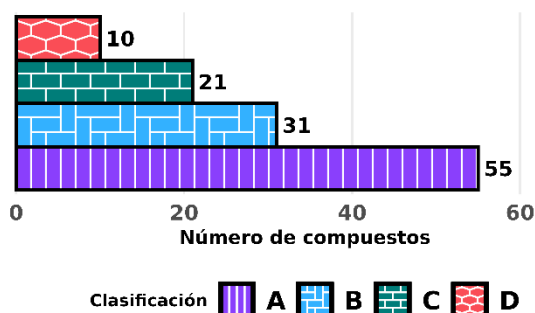


Figura 3. Clasificación de las moléculas estudiadas. A, heterociclos nitrogenados; B, derivados nitrogenados acíclicos; C, heterociclos con azufre u oxígeno y D, compuestos con grupos funcionales oxigenados.

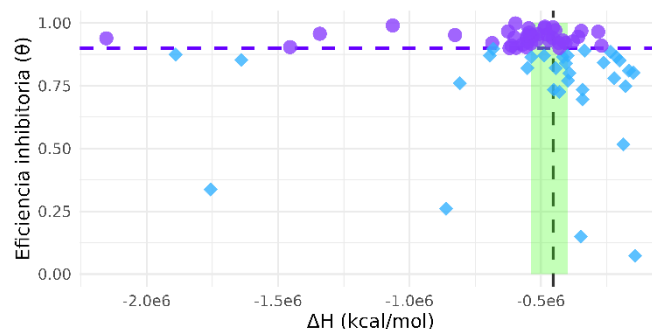


Figura 5. Correlación entre la entalpía y la eficiencia inhibitoria de las moléculas estudiadas. En color morado buenos inhibidores (BI) y en azul los malos inhibidores (MI).

3.2.2. Compuestos nitrogenados

Del análisis de 86 compuestos nitrogenados se identifican once descriptores discriminantes significativos (Tabla 1). Los buenos inhibidores se caracterizan por mayor tamaño molecular (>20 átomos, >161.74 g/mol), complejidad estructural moderada-alta (>9.5 carbonos, >7.5 hidrógenos), flexibilidad conformacional (>1.5 enlaces rotables) y sistemas de anillos extendidos (perímetro >13). El descriptor "Enlaces sin hidrógeno" muestra un umbral con dirección inversa (<13.5), lo que podría indicar que existe un límite superior para la complejidad estructural de los buenos inhibidores. Estos parámetros sugieren que la eficiencia inhibitoria requiere: (i) múltiples sitios de interacción molecular, (ii) estabilidad termodinámica en ambiente corrosivo, y (iii) capacidad conformacional para optimizar la función inhibitoria.

3.3. Análisis termodinámico

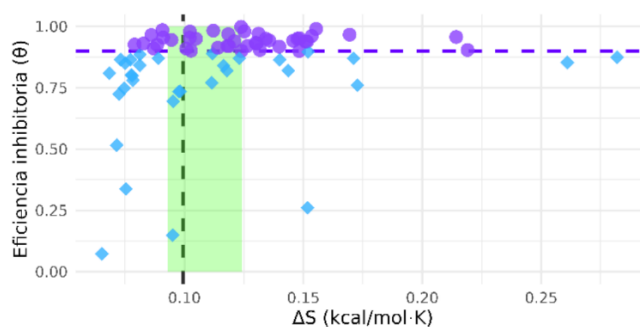


Figura 4. Correlación entre la entropía y la eficiencia inhibitoria de las moléculas estudiadas. En color morado buenos inhibidores (BI) y en azul los malos inhibidores (MI).

La optimización estructural B3LYP/6-311++G** + LANL2DZ reveló correlaciones significativas entre parámetros termodinámicos y eficiencia inhibitoria (Figuras 4-5). Los buenos inhibidores presentan entropías superiores a 0.10 kcal/mol-K y entalpías inferiores a -4.53×10^5 kcal/mol, sugiriendo un perfil termodinámico óptimo que combina estabilidad molecular (entalpías más negativas) con flexibilidad conformacional (entropías altas). Esta complementariedad indica que la efectividad inhibitoria requiere un equilibrio entre estabilidad estructural y adaptabilidad conformacional.

3.4. Modelos de clasificación

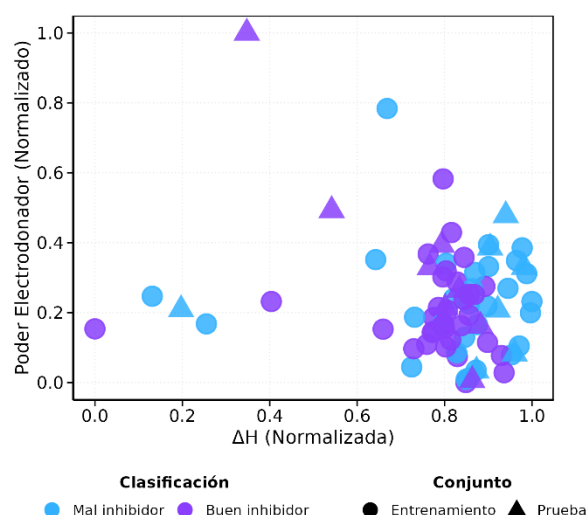


Figura 6. Gráfico de dispersión para los descriptores encontrados en el mejor modelo de clasificación. En color morado buenos inhibidores (BI) y en azul los malos inhibidores (MI). Los triángulos corresponden al conjunto de prueba y los círculos al de entrenamiento.

Los análisis previos se centraron en identificar descriptores individuales con capacidad discriminante. Sin embargo, la eficiencia inhibitoria puede depender de interacciones simultáneas entre múltiples propiedades moleculares, por lo que se implementó un enfoque multivariable basado en modelos de clasificación para evaluar estas interacciones complejas.

Se construyó una matriz de datos utilizando 73 compuestos nitrogenados caracterizados por 202 descriptores (combinando descriptores globales de reactividad calculados y descriptores generados por Dragon). El conjunto se dividió en 80 % de entrenamiento y 20 % de prueba, aplicando escalado min-max [0-1] para garantizar una contribución equilibrada de cada descriptor. Para optimizar el rendimiento, se implementó un filtrado en dos etapas: eliminación de descriptores con varianza cercana a cero y posteriormente un filtro de correlación para descartar descriptores altamente correlacionados ($r > 0.8$), generando 35 descriptores finales.

La metodología de selección "step forward" identificó al poder electrodonador y la entalpía como el conjunto óptimo de descriptores discriminantes. La relación entre ambos se observa en la

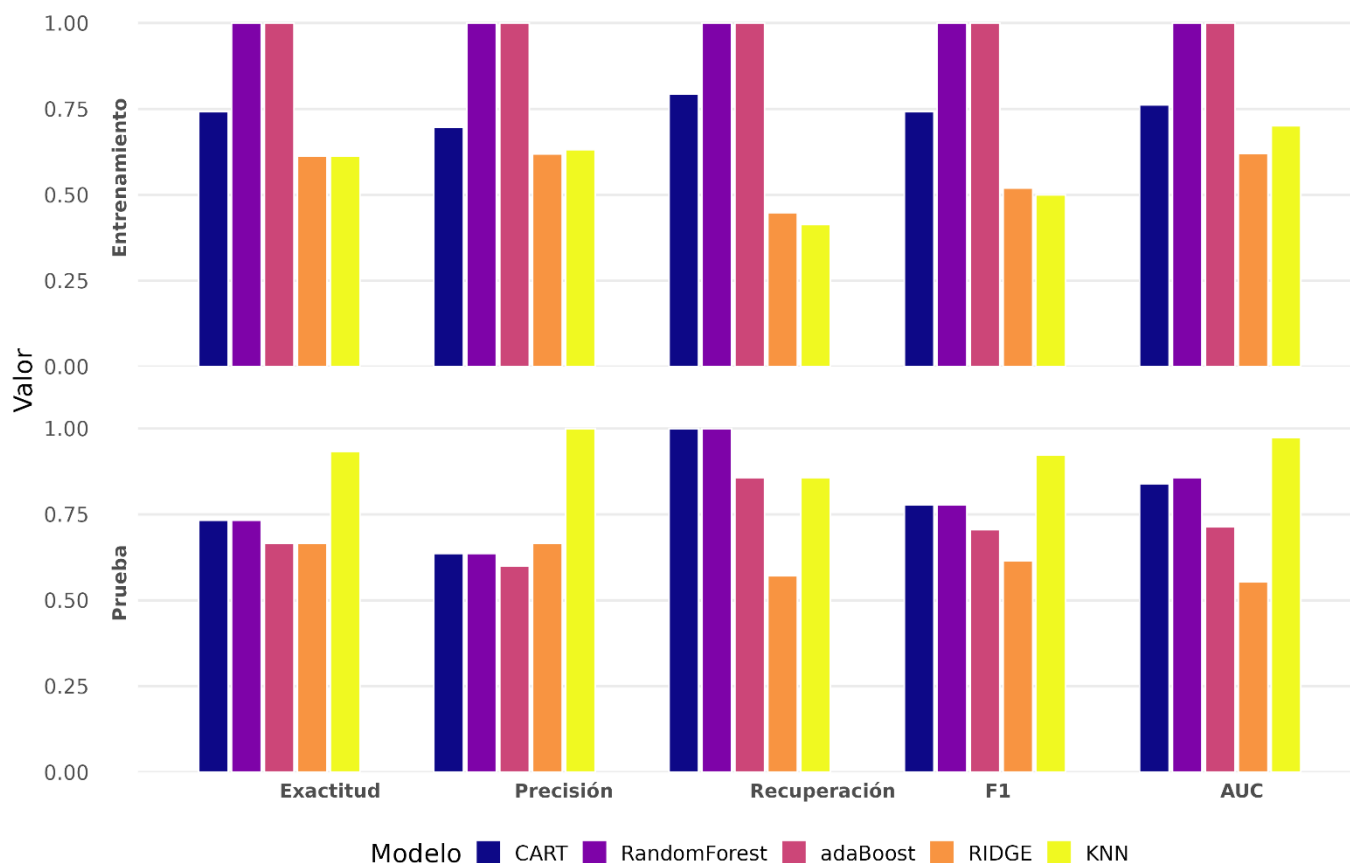


Figura 7. Comparación del desempeño de los diferentes algoritmos de clasificación.

referencia.6. Entre los algoritmos de clasificación evaluados, Random Forest y AdaBoost demostraron un desempeño superior cuando se utilizó este conjunto de descriptores. La robustez de estos modelos basados en ensamblaje sugiere que la relación entre los descriptores seleccionados y la eficiencia inhibitoria posee un carácter no lineal complejo, que estos algoritmos capturan efectivamente mediante la combinación de múltiples clasificadores débiles. La **¡Error! No se encuentra el origen de la referencia.7** presenta una comparación detallada de los principales indicadores de desempeño para los modelos evaluados.

4. Conclusiones

El presente estudio desarrolló un modelo de clasificación basado en aprendizaje automático para discriminar compuestos orgánicos nitrogenados según su eficiencia inhibitoria contra la corrosión del acero en medios ácidos.

A partir de la actividad inhibitoria reportada de 73 moléculas se identificaron dos descriptores moleculares clave: el poder electrodonador y la entalpía. La naturaleza no lineal de los modelos más efectivos sugiere que la discriminación entre buenos y malos inhibidores de corrosión implica interacciones complejas que no pueden capturarse mediante correlaciones simples o modelos lineales.

El modelo de Random Forest (RF) desarrollado mostró una exactitud del 73% en la clasificación binaria de inhibidores, demostrando un excelente rendimiento tanto en validación interna como externa, lo cual confirma la relevancia de los

descriptores seleccionados para discriminar entre diferentes niveles de actividad inhibitoria.

El análisis estadístico comparativo reveló diferencias significativas entre buenos y malos inhibidores para compuestos nitrogenados, identificando once descriptores moleculares con capacidad discriminante. En contraste, los heterociclos con azufre y oxígeno no mostraron correlaciones simples con la eficiencia inhibitoria, sugiriendo mecanismos de acción más complejos para esta familia.

La aplicación del modelo a la evaluación virtual de nuevos inhibidores demostró su utilidad para la discriminación y selección de compuestos prometedores, proporcionando una herramienta valiosa para la identificación racional de moléculas con potencial elevada eficiencia inhibitoria. Este enfoque computacional reduce significativamente el tiempo y costo asociados con la síntesis y evaluación experimental, contribuyendo al avance en el campo de la protección contra la corrosión.

Agradecimientos

A la Universidad Autónoma del Estado de Hidalgo por los recursos otorgados y el tiempo de cómputo.

U. J. Rangel-Peña agradece a SECIHTI por la beca para estudios de posgrado con número de apoyo 4023353.

Los autores agradecen a la SECIHTI por el apoyo brindado a través del proyecto CBF-2025-G-568 para el desarrollo de la investigación.

Los autores agradecen el apoyo otorgado a la SECIHTI a través del programa social SNII.

Referencias

- Ayati, N. S., Khandandel, S., Momeni, M., Moayed, M. H., Davoodi, A., & Rahimizadeh, M. (2011). Inhibitive effect of synthesized 2-(3-pyridyl)-3,4-dihydro-4-quinazolinone as a corrosion inhibitor for mild steel in hydrochloric acid. *Materials Chemistry and Physics*, 126(3), 873-879. doi: 10.1016/j.matchemphys.2010.12.023
- Boada, S. Z., Leon, G. J., & Ortiz, H. F. (2025). *Predicting Corrosion Rate in Oil Wells Using First Principles and Machine Learning*. MDPI.
- Contreras, D. A., Quej, L. M., Liu, H. B., Sosa, E., & Alamilla, J. L. (2020). Role of the Mexican clay soils on corrosiveness and stress corrosion cracking of low carbon pipeline steels: A case study. doi: 10.5006/3515
- Curtiss, L. A., Redfern, P. C., Rassolov, V., Kedziora, G., & Pople, J. A. (2001). Extension of Gaussian-3 theory to molecules containing third-row atoms K, Ca, Ga–Kr. *The Journal of Chemical Physics*, 114(21), 9287-9295. doi: 10.1063/1.1366337
- Diciccio, T. J., & Romano, J. P. (1988). A Review of Bootstrap Confidence Intervals. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 50(3), 338-354. doi: 10.1111/j.2517-6161.1988.tb01732.x
- Divine, G., Norton, H. J., Hunt, R., & Dienemann, J. (2013). A Review of Analysis and Sample Size Calculation Considerations for Wilcoxon Tests. *Anesthesia & Analgesia*, 117(3), 699-710. doi: 10.1213/ANE.0b013e31827f53d7
- Esmailzadeh, S., Aliofkhaei, M., & Sarlak, H. (2018). Interpretation of Cyclic Potentiodynamic Polarization Test Results for Study of Corrosion Behavior of Metals: A Review. *Protection of Metals and Physical Chemistry of Surfaces*, 54(5), 976-989. doi: 10.1134/S207020511805026X
- Hay, P. J., & Wadt, W. R. (1985). *Ab initio* effective core potentials for molecular calculations. Potentials for K to Au including the outermost core orbitals. *The Journal of Chemical Physics*, 82(1), 299-310. doi: 10.1063/1.448975
- Hesterberg, T. (2011). Bootstrap. *WIREs Computational Statistics*, 3(6), 497-526. doi: 10.1002/wics.182
- Hourahine, B., Aradi, B., Blum, V., Bonafé, F., Buccheri, A., Camacho, C., ... Frauenheim, T. (2020). DFTB+, a software package for efficient approximate density functional theory based atomistic simulations. *The Journal of Chemical Physics*, 152(12), 124101. doi: 10.1063/1.5143190
- Instituto Nacional de Estadística y Geografía (INEGI). (2023). Producto Interno Bruto (PIB). Recuperado de <https://www.inegi.org.mx/>
- Kaur, P., Stoltzfus, J., & Yellapu, V. (2018). Descriptive statistics. *International Journal of Academic Medicine*, 4(1), 60. doi: 10.4103/IJAM.IJAM_7_18
- Khan, G., Newaz, K. Md. S., Basirun, W. J., Ali, H. B. M., Faraj, F. L., & Khan, G. M. (2015). Application of Natural Product Extracts as Green Corrosion Inhibitors for Metals and Alloys in Acid Pickling Processes-A review. *International Journal of Electrochemical Science*, 10(8), 6120-6134. doi: 10.1016/S1452-3981(23)06707-X
- Kubisztal, J., Kubisztal, M., & Haneczok, G. (2020). Corrosion damage of 316L steel surface examined using statistical methods and artificial neural network. *Materials and Corrosion*, 71(11), 1842-1855. doi: 10.1002/maco.202011830
- Kuhn, M. (2008). Building Predictive Models in R Using the caret Package. *Journal of Statistical Software, Articles*, 28(5), 1-26. doi: 10.18637/jss.v028.i05
- Kutz, M. (2018). *Handbook of environmental degradation of materials* (3e ed). Oxford: William Andrew.
- Landolfo, R., Cascini, L., & Portioli, F. (2010). Modeling of Metal Structure Corrosion Damage: A State of the Art Report. *Sustainability*, 2(7), 2163-2175. doi: 10.3390/su2072163
- Lee, C., Yang, W., & Parr, R. G. (1988). Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Physical Review B*, 37(2), 785-789. doi: 10.1103/PhysRevB.37.785
- Makrini, B. E., Lgaz, H., Larouj, M., Salghi, R., Hasan, A. R., Belkhaouda, M., ... Oudda, H. (2016). The inhibition performance of sulfamerazine for corrosion of mild steel in HCl.
- Mauri, A., Consonni, V., Pavan, M., Todeschini, R., & others. (2006). Dragon software: An easy approach to molecular descriptor calculations. *Match*, 56(2), 237-248.
- Morell, C., Grand, A., & Toro-Labbé, A. (2005). New Dual Descriptor for Chemical Reactivity. *The Journal of Physical Chemistry A*, 109(1), 205-212. doi: 10.1021/jp046573a
- Nick, T. G. (2007). Descriptive Statistics. En W. T. Ambrosius (Ed.), *Topics in Biostatistics* (pp. 33-52). Totowa, NJ: Humana Press. doi: 10.1007/978-1-59745-530-5_3
- Obayes, H. R., Al-Amiery, A. A., Alwan, G. H., Abdullah, T. A., Kadhum, A. A. H., & Mohamad, A. B. (2017). Sulphonamides as corrosion inhibitor: Experimental and DFT studies. *Journal of Molecular Structure*, 1138, 27-34. doi: 10.1016/j.molstruc.2017.02.100
- Pedefferri, P. (2018). Statistical Analysis of Corrosion Data. En P. Pedefferri, *Corrosion Science and Engineering* (pp. 679-696). Cham: Springer International Publishing. doi: 10.1007/978-3-319-97625-9_30
- Rammelt, U., & Reinhard, G. (1992). Application of electrochemical impedance spectroscopy (EIS) for characterizing the corrosion-protective performance of organic coatings on metals. *Progress in Organic Coatings*, 21(2-3), 205-226. doi: 10.1016/0033-0655(92)87005-U
- Rangel-Peña, U. J., Zárate-Hernández, L. A., Camacho-Mendoza, R. L., Gómez-Castro, C. Z., González-Montiel, S., Pescador-Rojas, M., ... Cruz-Borbolla, J. (2023). Conceptual DFT, machine learning and molecular docking as tools for predicting LD50 toxicity of organothiophosphates. *Journal of Molecular Modeling*, 29(7), 217. doi: 10.1007/s00894-023-05630-4
- Secretaría de Economía. (2025). Basic Industry of Iron and Steel: Wages, production, investment, opportunities and complexity. Recuperado de Data México website: <https://www.economia.gob.mx/>
- Sultan, A. A., Mashrei, M. A., & Washer, G. A. (2020). Utilization of Wilcoxon-Mann-Whitney statistics in assessing the reliability of nondestructive evaluation technologies. *Structures*, 27, 780-787. doi: 10.1016/j.istruc.2020.06.024
- Tansuğ, G., Tüken, T., Giray, E. S., Fındıkkıran, G., Sığircık, G., Demirkol, O., & Erbil, M. (2014). A new corrosion inhibitor for copper protection. *Corrosion Science*, 84, 21-29. doi: 10.1016/j.corsci.2014.03.004
- Todeschini, R., Lasagni, M., & Marengo, E. (1994). New molecular descriptors for 2D and 3D structures. *Theory. Journal of Chemometrics*, 8(4), 263-272. doi: 10.1002/cem.1180080405

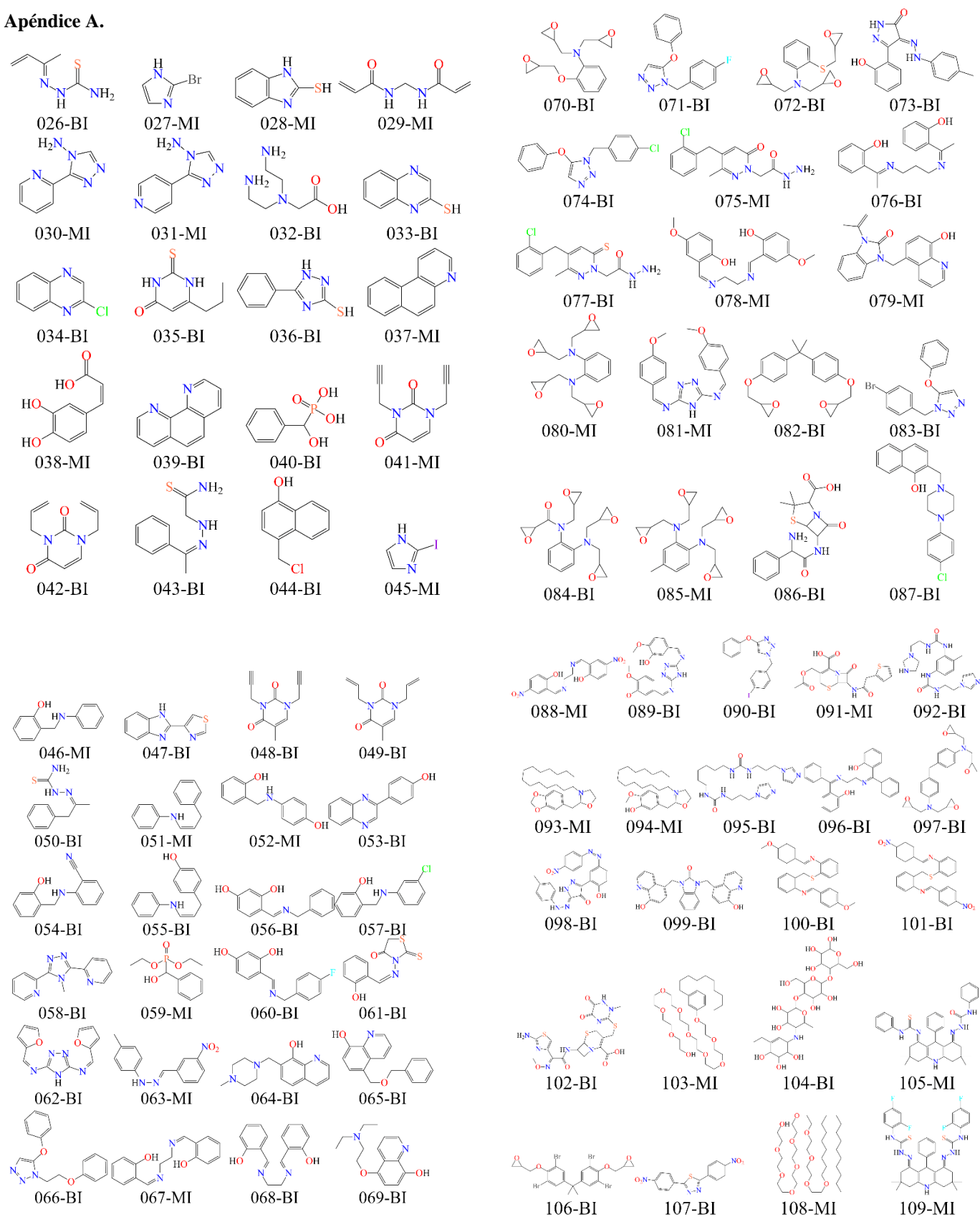
Apéndice A.

Figura 2-Continuación: Estructura en 2D de moléculas estudiadas. MI, Mal Inhibidor; BI, Buen Inhibidor.

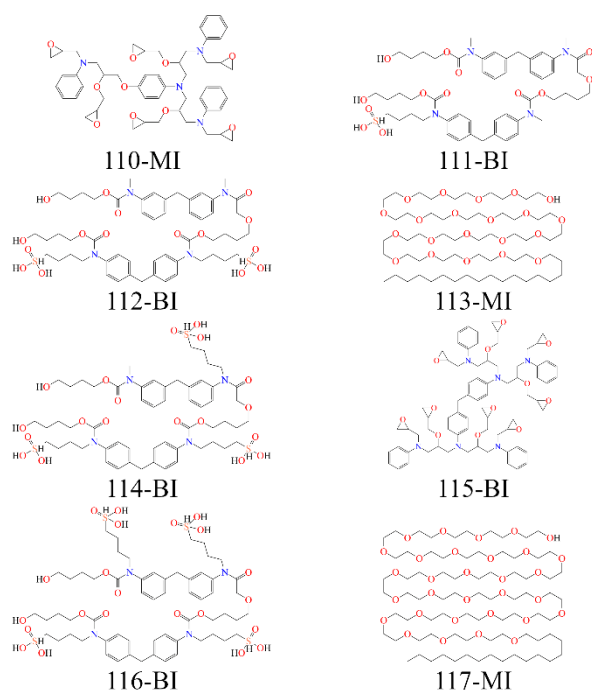


Figura 2-Continuación: Estructura en 2D de moléculas estudiadas. MI, Mal Inhibidor; BI, Buen Inhibidor.