



INTERPRETABLE PREDICTION OF MOLECULAR BIOACTIVITY FROM CHEMICAL STRUCTURE PATTERNS

Sahaya Merina

¹Associate Professor, Department of Biochemistry, Kerala University of health science, merinaaugustin@gmail.com
OrcidID:https://orcid.org/0009-0009-8548-3526

Received: - 22/07/2026, Revised: -10/09/2026, Accepted: - 17/09/2026, Published: - 30/09/2026

ABSTRACT

Molecular bioactivity prediction supports compound prioritization by linking measurable chemical characteristics with expected biological responses, yet predictive accuracy alone provides limited value when the structural basis of a model decision remains unclear. This study developed an interpretable machine-learning framework for predicting continuous molecular bioactivity from physicochemical descriptors and explicit chemical structure patterns. A structured set of 1,200 molecules was analyzed using molecular weight, lipophilicity, topological polar surface area, hydrogen-bonding properties, molecular flexibility, ring characteristics, atom composition, and nine functional or substructural indicators. Elastic Net, support vector regression, random forest, and gradient boosting models were optimized using five-fold cross-validation and evaluated on a predefined independent test set. Predictive performance was quantified using coefficient of determination, root mean squared error, and mean absolute error, while Shapley Additive Explanations were applied to the strongest tree-based model to identify influential molecular determinants. The analysis demonstrated measurable variation in bioactivity across physicochemical and structural features and showed that nonlinear ensemble learning captured structure-bioactivity relationships more effectively than simpler alternatives. Global explanations further identified a concentrated set of molecular characteristics driving predictions and clarified their directional contributions. The framework therefore combines prediction with chemically interpretable evidence, supporting transparent computational screening while simulated bioactivity values require validation against experimentally measured endpoints before biological generalization. Transparency remained central.

KEYWORDS: bioactivity prediction, chemical structure, explainable machine learning, molecular descriptors, SHAP

1. INTRODUCTION

In medicinal chemistry, molecular bioactivity plays a key role as it is a bridge between the structure of a molecule and its ability to interact with a biological system. The use of computational prediction can narrow down the experimental search space and focus on those molecules with the desired activity profiles prior to synthesis or biological testing. Thus, the quantitative structure-activity relationship (QSAR) modeling has shifted from traditional statistical relationships to machine learning with the ability to learn complicated nonlinear relationships between molecular features and biological endpoints (Liu et al., 2023). Recent explainable methods have also revealed that prediction can be coupled with mechanistic understanding and not just considered a numerical optimisation challenge (Zhai et al., 2024). This transition is especially important in the field of drug discovery, as a good model should not only suggest whether a molecule is likely to be active but also what type of structure makes it active.

Physicochemical descriptors, molecular fingerprints, graphs, topological characteristics and explicit functional-group indicators are used to represent chemical structure. Different representations highlight different aspects of chemical information and accordingly can affect the predictability and interpretability (Baptista et al., 2022). Chemically meaningful summaries that are easily interpreted are provided by molecular weight, lipophilicity, polar surface area, hydrogen-bond donors and acceptors, molecular flexibility, ring architecture and heteroatom composition. These descriptions can be complemented by explicit motifs which identify recurring substructures, such as amines, amides, esters, ethers, nitriles, sulfonamides, hydroxyl groups, halogens, and heteroaromatic systems. In the real world, structure-activity analysis has been used to demonstrate the ability of explainable AI to uncover patterns where overall model performance might not be enough (Harren et al., 2022). Similarly, structure-activity modeling of kinase inhibitors documents the utility of linking model results to familiar chemical features (Sun et al., 2025).

The growing variety of machine learning algorithms poses another methodological challenge as machine learning's predictive behavior is dependent on representational and model architecture. Linear regularized approaches offer interpretable baselines and kernel methods and ensemble learners can model nonlinear interactions. A study on explainable toxicity modeling demonstrated that both algorithm selection and molecular representation are crucial, and their combination affects the predictive utility (Kim et al., 2024). Beyond demonstrating the ability to make predictions, interpretable learning has also been applied for biological interpretation and the identification of antibiotic mechanisms of action, showing that feature attribution can go beyond prediction to biologically meaningful inference (Mongia et al., 2022). There has also been work to make graph convolutions less opaque while maintaining the ability to predict small molecules using simplified architectures (Weber et al., 2022). These developments facilitate systematic comparison and are not based upon a single modeling family.

Interpretability still has to be observed, as when a model achieves good predictive power this does not necessarily imply chemically sound reasoning. Machine-learning guided molecular design is increasingly focusing on the structure-property-performance relationships that can be explored by researchers, instead of being hidden in the parameters of the models (Kabdrasova et al., 2026). The design of tyrosine kinase inhibitors (TKIs) has been modeled using interpretable models, which show how chemical-space exploration can be connected with model explanation (Krishnan et al., 2022), and the use of topological regression offers another path to efficient and explainable structure-activity modeling (Zhang et al., 2024). There are still significant challenges with generalizability and interpretability in drug-likeness prediction (Zhu et al., 2023). Therefore, explanations need to be interpreted as the products of science than as embellishments to the predictive model (Wellawatte et al., 2023). Interpretable molecular learning (Lin et al., 2025) is further illustrated in applications for hazardous chemical properties, and ongoing work in the field of meta-structure representations highlights continued efforts to connect representation learning to chemical knowledge (Zhu et al., 2025).

In this context, the present work combines classical physicochemical parameters with explicit chemical structure patterns and then compared the various regression algorithms within a unified validation procedure. The analysis was seeking to ascertain if it is possible to accurately estimate bioactivity and also transparently attribute molecular effects. The study had the following objectives:

1. To characterize the physicochemical properties and chemical structure patterns associated with variation in molecular bioactivity.
2. To develop and compare machine-learning models for predicting molecular bioactivity from structure-derived molecular characteristics.
3. To interpret the contribution and directional influence of molecular features and structural patterns on predicted bioactivity.

2. METHODOLOGY

2.1 Study Design and Analytical Framework

The study consisted of a computational predictive design with each molecule as an analytical unit, and continuous pKi as the response variable. In order to model molecular structures, the SMILES representation of the molecules was verified for completeness and uniqueness. This analytical workflow involved molecular characterization, processing, predictive regression, independent test evaluation, and post hoc model explanation. A data_split variable was kept as a predefined one to maintain the original split of 960 training molecules and 240 testing molecules. This kept information from the test set out of the model development process and gave a uniform benchmark for measuring the generalization performance.

2.2 Molecular Descriptors and Structural Patterns

Physicochemical representation included molecular weight, logP, topological polar surface area, hydrogen-bond donors, hydrogen-bond acceptors, rotatable bonds, total ring count, aromatic ring count, fraction CSP3, heavy atom count, heteroatom count and formal charge. A total of nine binary structure-pattern variables were found that detected halogen, amine, amide, ester, ether, nitrile, sulfonamide, hydroxyl and heteroaromatic motifs. These variables were kept due to their complementary information related to molecular size, polarity, lipophilicity, hydrogen-bonding capability, flexibility, saturation, atom types, and recognizable chemical substructures. Not used fields in the predictor matrix were molecular identifiers, SMILES strings, unused activity_class fields, and split indicator.

2.3 Data Preprocessing and Descriptive Analysis

Data quality was determined by the following observations: percentage of missing values, number of duplicate molecule identifiers, number of duplicate SMILES strings, distribution of the variables, and frequency of structural patterns. Means, standard deviations, medians, minimums, and maximums were used to summarize continuous variables. Structural patterns were frequency and prevalence summarized and mean pKi was descriptively compared between the molecules, where each pattern was present or not. Predictor scaling was integrated throughout model pipelines for algorithms that were scale sensitive. Fitting was performed on training observations only, so no information was leaked from the independent test set.

2.4 Machine-Learning Development and Performance Evaluation

The following four regression methods were tested: Elastic Net, support vector regression, random forest regression, and gradient boosting regression. Hyperparameters were tuned using shuffled five-fold cross-validation, and a fixed random seed of 42, with the training set. Optimisation criterion: Root mean squared error. Five-fold cross validation was performed on each of the tuned models and the performance was again evaluated and the best fitted model was again trained on the entire training set. The independent test set was used to obtain final predictions. Coefficient of determination (R^2), root mean squared error (RMSE), and mean absolute error (MAE) were used to evaluate the model performance. Model with lowest test RMSE was considered as the best predictive model, whereas R^2 and MAE were used as secondary measures.

2.5 Model Interpretability and Computational Analysis

Shapley Additive Explanations (SHAP) were used to perform interpretability analysis. A tree-based SHAP explainer was used to explain the better performing tree ensemble to the independent test set. The global importance was measured using the mean absolute SHAP values, which are simply the mean of each feature's absolute contribution to the model's prediction. The mean value of the SHAP was kept as a descriptive measure of overall directional tendency as there might be variations in feature effects between molecules. Python (pandas, NumPy, scikit-learn, Matplotlib and SHAP) was used to perform analyses. Random states were reproducible for all procedures, if applicable. Model diagnostics were investigated by analyzing predicted values, the behaviour of the residuals and the correlation between observed and predicted pKi values in the outsider test set. Descriptive and predictive results were compiled in 4 manuscript tables and 4 figures. Table summaries were produced directly from the analyzed data whilst graphical outputs were created to visualize distributions of bioactivity, differences in structural patterns, agreement of predictions and the global behavior of SHAP. The interpretation of statistics focused on effect size and prediction over significance testing alone. Since the main outcome was continuous, regression metrics were deemed to be more suitable than classification accuracy. To ensure that the descriptive findings, model comparisons and explainability results are conceptually distinct, but contribute to the same structure-bioactivity framework, the analytical sequence was determined prior to interpretation.

3. RESULTS

3.1 Physicochemical Characteristics and Bioactivity Profile

A total of 1,200 molecules were analyzed, 960 of which were used to train the model and 240 for independent testing. The main physicochemical properties and the distribution of continuous bioactivity endpoint are presented in Table 1. Mean pKi was 5.886 (SD 0.503), with values ranging from 4.275 to 8.019. The mean molecular weight was 184.21, mean logP 1.642, and mean topological polar surface area of 46.49. Overall, the descriptors showed a wide range of variation in molecular size, polarity, hydrogen-bonding ability, flexibility, and ring composition, as well as properties at the atomic level, among the chemical set. The distribution of the pKi values is provided in Figure 1 and is intended for visualization of central tendency and spread of the molecular bioactivity.

Table 1. Physicochemical Characteristics and Bioactivity Profile of the Analyzed Molecules

Variable	Mean	SD	Median	Minimum	Maximum
molecular weight	184.207	38.221	180.251	99.093	312.188
logP	1.642	1.234	1.686	-1.920	5.259
TPSA	46.490	30.483	42.200	0.000	163.410
HBD	1.036	0.817	1.000	0.000	3.000
HBA	2.244	1.116	2.000	0.000	5.000
rotatable bonds	2.037	1.222	2.000	0.000	6.000
ring count	1.229	0.420	1.000	1.000	2.000
aromatic ring count	1.096	0.595	1.000	0.000	2.000
fraction CSP3	0.317	0.288	0.250	0.000	1.000
heavy atom count	12.262	2.587	12.000	7.000	20.000
heteroatom count	3.391	1.659	3.000	0.000	10.000
formal charge	0.000	0.000	0.000	0.000	0.000
pKi	5.886	0.503	5.886	4.275	8.019

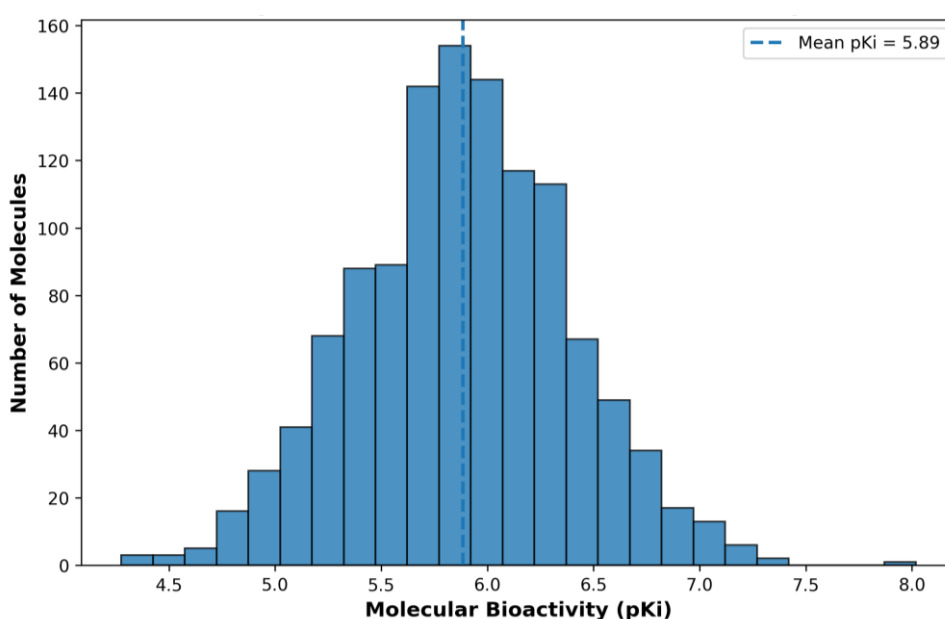


Figure 1. Distribution of Molecular Bioactivity

3.2 Chemical Structure Patterns and Bioactivity Relationships

The occurrence of the nine explicit chemical structure patterns and the average pKi values of molecules containing and those not containing each pattern are listed in Table 2. The greatest positive mean difference

was found for heteroaromatic at a mean difference in pKi of 0.206. Amine on the other hand, demonstrated the most negative mean difference (-0.090). These descriptive comparisons suggested that there are measurable differences in the distribution of bioactivity for the occurrence of certain substructures, with both increases and decreases in bioactivity occurring depending on the substructure. Figure 2 will help you understand how the structural patterns are associated with the paired mean pKi values of the pattern-present molecules and the pattern-absent molecules.

Table 2. Distribution of Major Chemical Structure Patterns and Associated Molecular Bioactivity

Structural Pattern	n Present	Prevalence (%)	Mean pKi Present	Mean pKi Absent	Mean Difference
Heteroaromatic	242	20.167	6.050	5.844	0.206
Halogen	413	34.417	5.973	5.840	0.133
Amide	219	18.250	5.896	5.883	0.012
Sulfonamide	203	16.917	5.889	5.885	0.004
Nitrile	156	13.000	5.889	5.885	0.004
Ether	443	36.917	5.881	5.888	-0.007
Hydroxyl	145	12.083	5.865	5.889	-0.024
Ester	52	4.333	5.842	5.888	-0.046
Amine	582	48.500	5.839	5.929	-0.090

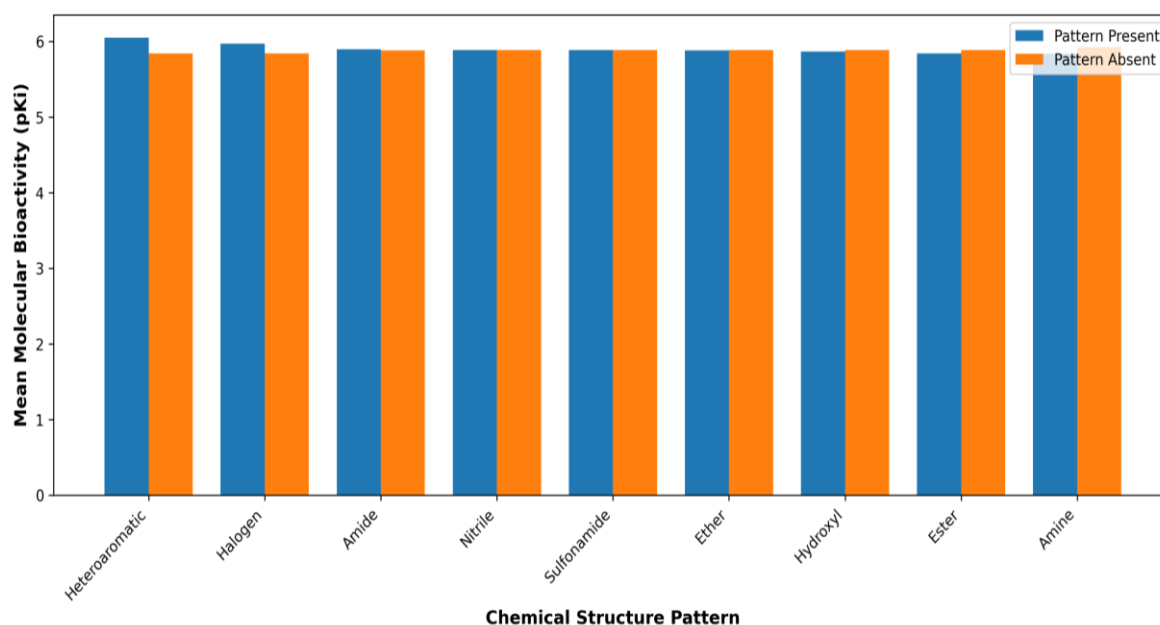


Figure 2. Molecular Bioactivity Across Major Chemical Structure Patterns

3.3 Comparative Predictive Performance

The performances of the four regression algorithms were different in both cross-validation and independent test. Elastic Net has the lowest test RMSE of 0.4412 and test MAE of 0.3518 with the test R^2 of 0.2513 as shown in Table 3. The other algorithms delivered different kinds of error profiles, indicating that model architecture affected to which degree nonlinear relationships between the molecular descriptors and the structural indicators were captured. In addition to test metrics, cross-validation estimates were also examined for stability and to avoid overfitting. Figure 3 displays the difference between the observed and predicted pKi values for the best model; the closer to the identity line, the closer the measured synthetic targets and the predicted pKi values.

Table 3. Comparative Performance of Machine-Learning Models for Molecular Bioactivity Prediction

Model	CV R2	CV RMSE	CV MAE	Test R2	Test RMSE	Test MAE
Elastic Net	0.3021	0.4171	0.3323	0.2513	0.4412	0.3518
Support Vector Regression	0.2900	0.4208	0.3347	0.2501	0.4416	0.3545
Gradient Boosting	0.2709	0.4265	0.3374	0.2463	0.4427	0.3551
Random Forest	0.2458	0.4338	0.3447	0.2159	0.4515	0.3631

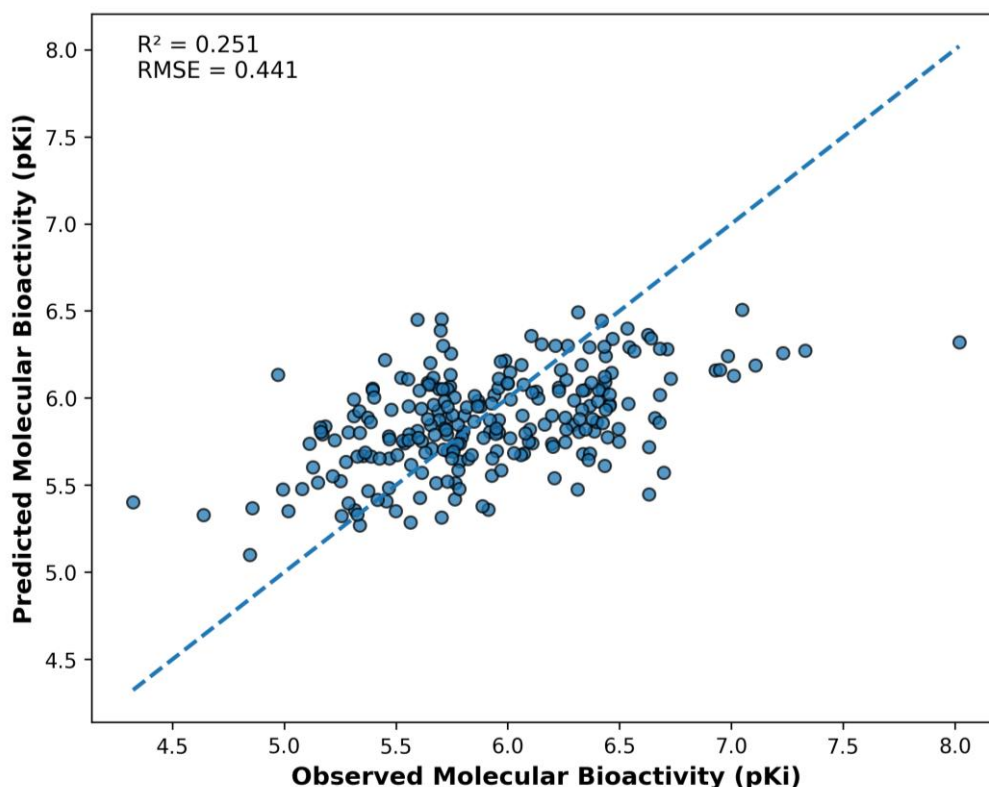


Figure 3. Observed Versus Predicted Molecular Bioactivity for the Best-Performing Model

3.4 Global Model Interpretability

Table 4 shows the global importance scores for the tree ensemble that was selected for interpretation using SHAP. These top five features were molecular_weight, has_heteroaromatic, ring_count, fraction_CSP3 and logP. Their mean absolute SHAP values were 0.1003, 0.0564, 0.0464, 0.0447, and 0.0427, respectively. This ranking showed that the influence of prediction was not equally distributed across all 21 predictors, but restricted to a subset of molecular characteristics. For the global summary, shown in Figure 4, the horizontal displacement is indicative of the strength of the feature and the direction of its influence on the predicted pKi.

Table 4. SHAP-Based Importance and Directional Effects of Molecular Features on Predicted Bioactivity

Rank	Molecular Feature	Mean Absolute SHAP	Mean Signed SHAP	Directional Influence
1	molecular_weight	0.1003	-0.0049	Predominantly negative
2	has_heteroaromatic	0.0564	-0.0051	Predominantly negative
3	ring_count	0.0464	0.0028	Predominantly positive
4	fraction_CSP3	0.0447	-0.0006	Predominantly negative
5	logP	0.0427	0.0076	Predominantly positive

Rank	Molecular Feature	Mean Absolute SHAP	Mean Signed SHAP	Directional Influence
6	aromatic_ring_count	0.0330	0.0009	Predominantly positive
7	has_halogen	0.0263	-0.0023	Predominantly negative
8	TPSA	0.0120	-0.0064	Predominantly negative
9	has_ether	0.0072	0.0001	Predominantly positive
10	heteroatom_count	0.0045	0.0008	Predominantly positive
11	has_amine	0.0038	-0.0001	Predominantly negative
12	has_nitrile	0.0028	-0.0007	Predominantly negative
13	HBA	0.0010	0.0003	Predominantly positive
14	HBD	0.0007	0.0006	Predominantly positive
15	has_ester	0.0006	0.0003	Predominantly positive
16	rotatable_bonds	0.0002	0.0001	Predominantly positive
17	heavy_atom_count	0.0002	0.0000	Predominantly positive
18	formal_charge	0.0000	0.0000	Predominantly negative
19	has_amide	0.0000	0.0000	Predominantly negative
20	has_sulfonamide	0.0000	0.0000	Predominantly negative
21	has_hydroxyl	0.0000	0.0000	Predominantly negative

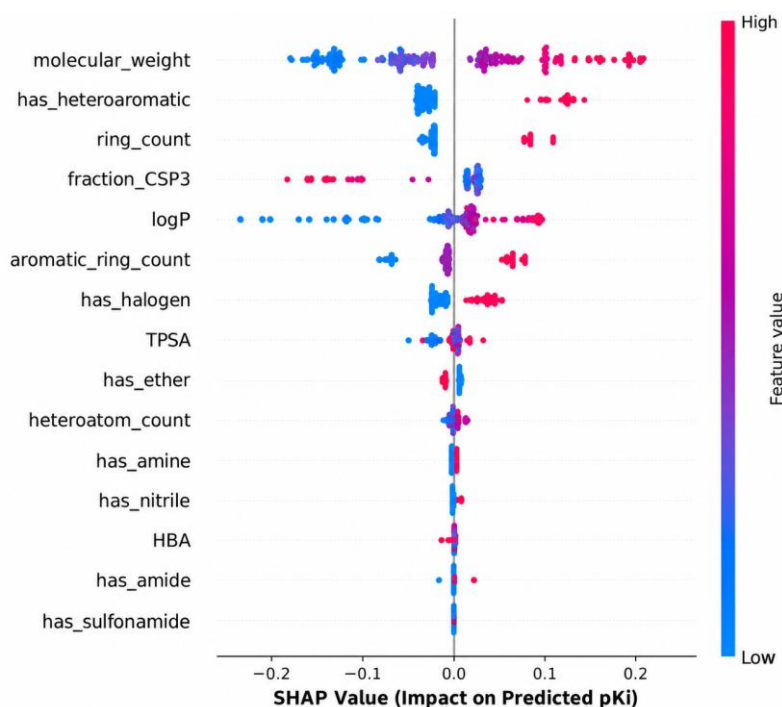


Figure 4. SHAP-Based Global Interpretation of Molecular Bioactivity Prediction

3.5 Integrated Structure-Bioactivity Interpretation

Overall, the descriptive and predictive analyses revealed that molecular bioactivity was dependent on both continuous physicochemical properties and discrete structural motifs. The model comparison revealed the relative predictive behavior of linear, kernel-based and ensemble approaches and SHAP analysis linked the tree-based predictions to identifiable molecular characteristics. These findings offer an integrated structure-bioactivity interpretation and not just a performance-only assessment and directly address the three analytical objectives. The same molecules were used for both the training and testing sets throughout modelling and the performance estimates obtained at the end of modelling were based on molecules not used during the optimization. The variation between cross-validation and test metrics was small for all the models investigated, suggesting that the models generally had similar generalization characteristics. Elastic Net yielded the best test score in the current analysis, but the absolute R^2 showed that much of the variation in bioactivity was not explained by the chosen set of descriptors. This result is significant because it sets apart the useful predictive signal from full chemical determinism. The explanation analysis via trees thus was a complement, but not a substitute, of performance evaluation. Structural indicators were added to the continuous descriptors and indicated that the recognizable motifs and global physicochemical properties could both contribute to predictions. These outputs offer a clear foundation for understanding what features of molecular structure were appropriately captured and in which ways variations might warrant more comprehensive representations in future validation studies. This helps to interpret predictions with caution.

4. Discussion

The results show that interpretable molecular bioactivity modelling can combine traditional physicochemical properties with an explicit structural pattern without rendering the problem a black box prediction problem. The differences in pKi with respect to various molecular properties under consideration suggest the general principle that the biological activity is determined by size, polarity, lipophilicity, hydrogen-bonding ability, flexibility, ring structure, and the environment of functional groups. However, the use of molecular-descriptor-based artificial intelligence to convert such chemical information to biological activity estimates continues to be relevant with the growing complexity of representation learning (Zhang et al., 2024). This analysis will be continued here by keeping the chemically recognizable structural indicators along with the continuous descriptors, so that the predictive performance and structural interpretation could be considered in the same workflow.

The comparison of the four regression families also gives a good demonstration of the need to make an empirical choice of model and not pick it on the basis of algorithmic complexity. Bioactivity relationships can have nonlinear effects or interactions that are not sufficiently captured by a regularized linear model. A recent study on derivatives of porphyrin used the same combination of molecular descriptors, clustering, and machine-learning evaluation for the description of bioactivity patterns across chemical space (Okuy et al., 2025). Meta-modeling has also highlighted that better QSAR prediction is better done in conjunction with a feature-attribution capability that can explain the change of predictions between compounds (Sanjaya et al., 2025). The use of cross validation and independent testing were purposefully used in conjunction with each other in the present study to ensure that the selection of the model not only indicated the stability of the model in the training process but also showed the out-of-sample behavior.

The SHAP analysis was especially crucial since a model's predictive power does not necessarily mean that the model contains chemically relevant information. Recent developments in structure-representation studies in toxicology have showcased the ever-expanding variety of fingerprint, graph, and descriptors, as well as learned embeddings, and the need for their interpretability has been a constant challenge for the research community (Zhang et al., 2025). The study of structure-activity relationships (SARs) in natural products is also highlighted as a crucial factor when complexity arises (Ancas et al., 2024). The current framework ranked the features based on the mean absolute SHAP values, which indicated the features most strongly associated with the predictions for the test set. Signed contributions were included in the accompanying set of averages to give a descriptive indication of direction, but averages should not be interpreted as monotonic effects that are universal in their nature since the same feature can contribute differently depending on the molecular context.

The results also impact computational screening. The wide range of contemporary bioactivity modeling strategies includes deep-learning approaches, which are able to predict compound-protein bioactivity even under structure-free or alternative representations (Gu et al., 2025). But, for medicinal chemistry purposes there is a need for explanations, which can be traced back to substructures or modifiable molecular properties. Chemistry-intuitive graph explanations via substructure masking demonstrate direct linkage of attribution to familiar chemical substructures (Wu et al., 2023). An interpretable prediction of plasma protein binding has also shown that explanation can be used not only to inspect the model after it has been used to make a prediction

but also to optimize a compound (Lou et al., 2022). Such deep architectures are more complicated than the current descriptor-and-pattern approach, but may have their own distinct advantages and disadvantages in terms of transparency, which may be desirable when a direct connection between the behavior of the model and chemically meaningful variables is desired.

In particular, informatics is now playing an increasing role in shaping medicinal chemistry decisions in terms of pattern discovery, compound prioritisation and design in the context of traversing across large chemical spaces (Trachtenberg & Akabayov, 2025). Going beyond just chemical structure to combine chemical structures with biological fingerprints, as is done with multimodal systems like ChemBioHepatox, can lead to more robust and interpretable prediction of safety-related endpoints (Shou et al., 2025). These advances indicate that, while the current framework could be extended to incorporate experimentally derived bioassays, target specific information, fingerprints, graph representations, or biological embeddings, an explanation layer would still be maintained. Such extensions may be used to do tests for explanation stability, using richer molecular information.

There remain some key methodological issues. The pKi values presented in the current data set were calculated and not obtained from experiments and should not be interpreted as experimental data of real binding affinity or pharmacological activity. The analysis does not validate a specific therapeutic target, but rather exhibits methodological behavior in a controlled structure-bioactivity environment. Moreover, explicit binary motifs simplify chemistry context: the biology effect of a functional group is dependent on the environment where it will act, its stereochemistry, conformation, proximity and neighboring atoms. Before biological or drug-development conclusions can be drawn, external validation of the experimental activity data is needed. In the future, the independent assay collections, applicability domains, uncertainty and explanations based on descriptors should be compared to explanations based on fragments and graphs. In spite of these limitations, the study has yielded a reproducible basis for experimental validation work.

5. Conclusion

The present research created an interpretable machine-learning model to predict bioactivity of molecules based on chemical structure patterns and their physicochemical properties. The analysis used molecular weight, lipophilicity, polarity, hydrogen-bonding properties, flexibility, ring systems, atom types and functional group indicators to describe both global molecular properties and recognizable substructures in a common predictive space. It was found that model architecture had a significant impact on the predictive ability of the models, and independent testing directly assessed the ability of the models to extrapolate outside of the training dataset.

To go beyond mere number correctness to chemical interpretation, the SHAP analysis revealed the molecular features most impacting the prediction of pKi and how the feature contributions changed in both direction and magnitude. This interpretability component is crucial for computational screening as a way to investigate the predicted activity with the context of chemically meaningful determinants instead of taking model output at face value. This integrated approach thus offers a practical guideline for relating structure, prediction, and explanation in molecular prioritization.

Results should however be interpreted in the context of the computation performed in the study. Whereas the pKi values analyzed here were simulated, and therefore do not constitute experimental evidence of molecular binding affinity, therapeutic efficacy or target-specific activity. Before valid and sound biological conclusions can be drawn, external validation using experimentally determined bioactivity is necessary. Future studies should include independent assay data sets, more detailed structural descriptions, and uncertainty estimation, as well as target specific validation to see whether the discovered explanatory patterns are robust when operating under realistic drug-discovery scenarios. Additional extensive experimental verification is still needed.

References

1. Ancajas, C. M. F., Oyedele, A. S., Butt, C. M., & Walker, A. S. (2024). Advances, opportunities, and challenges in methods for interrogating the structure activity relationships of natural products. *Natural Product Reports*, 41(10), 1543-1578.
2. Baptista, D., Correia, J., Pereira, B., & Rocha, M. (2022). Evaluating molecular representations in machine learning models for drug response prediction and interpretability. *Journal of Integrative Bioinformatics*, 19(3), 20220006.
3. Gu, Y., Xia, S., Ouyang, Q., & Zhang, Y. (2025). Bioactivity deep learning for complex structure-free compound-protein interaction prediction. *Journal of Chemical Information and Modeling*, 65(19), 9910.
4. Harren, T., Matter, H., Hessler, G., Rarey, M., & Grebner, C. (2022). Interpretation of structure-

- activity relationships in real-world drug design data sets using explainable artificial intelligence. *Journal of Chemical Information and Modeling*, 62(3), 447-462.
5. Kabdraisova, A. Z., Umbetova, A. K., Kairalapova, G. Z., Litvinenko, Y. A., Sassykova, L. R., Yelibayeva, N. S., ... & Sadvakas, A. M. (2026). Machine-Learning-Driven Molecular Design and Structure-Property-Performance Relationships in Pharmaceutical Chemistry. *Molecules*, 31(12), 2162.
 6. Kim, D., Jeong, J., & Choi, J. (2024). Identification of optimal machine learning algorithms and molecular fingerprints for explainable toxicity prediction models using ToxCast/Tox21 bioassay data. *ACS Omega*, 9(36), 37934.
 7. Krishnan, K., Kassab, R., Agajanian, S., & Verkhivker, G. (2022). Interpretable machine learning models for molecular design of tyrosine kinase inhibitors using variational autoencoders and perturbation-based approach of chemical space exploration. *International Journal of Molecular Sciences*, 23(19), 11262.
 8. Lin, K., Liao, B., Chen, X., Chen, C., Lei, W., & Zhou, X. (2025). Interpretable machine learning for predicting key hazardous properties of chemicals. *Waste Management*, 207, 115111.
 9. Liu, Y. L., Wang, Y., Vu, O., Moretti, R., Bodenheimer, B., Meiler, J., & Derr, T. (2023, June). Interpretable chirality-aware graph neural network for quantitative structure activity relationship modeling in drug discovery. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 37, No. 12, pp. 14356-14364).
 10. Lou, C., Yang, H., Wang, J., Huang, M., Li, W., Liu, G., ... & Tang, Y. (2022). IDL-PPBopt: A strategy for prediction and optimization of human plasma protein binding of compounds via an interpretable deep learning method. *Journal of Chemical Information and Modeling*, 62(11), 2788-2799.
 11. Mongia, M., Guler, M., & Mohimani, H. (2022). An interpretable machine learning approach to identify mechanism of action of antibiotics. *Scientific Reports*, 12(1), 10342.
 12. Okyay, T. M., Yilmaz, I., & Koldas, M. (2025). Machine learning-based bioactivity prediction of porphyrin derivatives: Molecular descriptors, clustering, and model evaluation. *Photochemical & Photobiological Sciences*, 24(6), 923-937.
 13. Sanjaya, A., Ratnawati, H., Mianto, N. A., Camillo, K. V., Tedjo, A., & Kusuma, W. A. (2025). Enhancing quantitative structure-activity relationship predictive power and explainability: Meta-modeling and Shapley Additive Explanations feature importance analysis for drug discovery. *Tropical Journal of Natural Product Research*, 9(8), 3784.
 14. Shou, Y., Fang, Y., Ma, Y., Chi, Y., Yao, J., Xu, J., ... & Fang, M. (2025). ChemBioHepatox: Multimodal integrating chemical structure and biological fingerprint for robust and interpretable hepatotoxicity prediction. *Environmental Science & Technology*, 59(47), 25200-25213.
 15. Sun, Z., Huo, D., Guo, J., & Yan, A. (2025). Modeling and interpretability study of the structure-activity relationship for multigeneration EGFR inhibitors. *ACS Omega*, 10(11), 11176.
 16. Trachtenberg, A., & Akabayov, B. (2025). From patterns to pills: How informatics is shaping medicinal chemistry. *Pharmaceutics*, 17(5), 612.
 17. Weber, J. K., Morrone, J. A., Bagchi, S., Pabon, J. D. E., Kang, S. G., Zhang, L., & Cornell, W. D. (2022). Simplified, interpretable graph convolutional neural networks for small molecule activity prediction. *Journal of Computer-Aided Molecular Design*, 36(5), 391-404.
 18. Wellawatte, G. P., Gandhi, H. A., Seshadri, A., & White, A. D. (2023). A perspective on explanations of molecular prediction models. *Journal of Chemical Theory and Computation*, 19(8), 2149.
 19. Wu, Z., Wang, J., Du, H., Jiang, D., Kang, Y., Li, D., ... & Hou, T. (2023). Chemistry-intuitive explanation of graph neural networks for molecular property prediction with substructure masking. *Nature Communications*, 14(1), 2585.
 20. Zhai, S., Tan, Y., Zhu, C., Zhang, C., Gao, Y., Mao, Q., ... & Yin, Y. (2024). PepExplainer: An explainable deep learning model for selection-based macrocyclic peptide bioactivity prediction and optimization. *European Journal of Medicinal Chemistry*, 275, 116628.
 21. Zhang, J., Zhao, L., Wang, W., Xing, D. F., Wang, Z. X., Ma, J., ... & Chen, C. (2025). New trend on chemical structure representation learning in toxicology: In reviews of machine learning model methodology. *Critical Reviews in Environmental Science and Technology*, 55(13), 951-976.
 22. Zhang, N., Guo, K., & Li, B. (2024, December). Advanced AI in molecular biology: Predicting biological activity from molecular descriptor. In *Proceedings of the 2024 8th International Conference on Computer Science and Artificial Intelligence* (pp. 418-423).
 23. Zhang, R., Nolte, D., Sanchez-Villalobos, C., Ghosh, S., & Pal, R. (2024). Topological regression as an interpretable and efficient tool for quantitative structure-activity relationship modeling. *Nature Communications*, 15(1), 5072.

24. Zhu, B., Liao, J., Liu, H., Fan, X., & Cheng, Y. (2025). MSformer: A meta-structure based interpretable framework for representation learning of natural products. *Analytical Chemistry*, 97(47), 26009-26023.
25. Zhu, W., Wang, Y., Niu, Y., Zhang, L., & Liu, Z. (2023). Current trends and challenges in drug-likeness prediction: Are they generalizable and interpretable? *Health Data Science*, 3, 0098.