*Research Journal of Forensic Sciences* ISSN 2321–1792
Vol. **14(1),** 7-11, January **(2026)** *Res. J. Forensic Sci.*

# Computational Modelling of Antimicrobial Efficacy against *Staphylococcus aureus* Using Structure-Activity Analysis

**Thanvi Megha*, Shekhawat Divya, Chaudhary Vibha and Bhojak N.**
GCRC, P.G. Department of Chemistry, Govt Dungar College (NAAC 'A' Grade), MGS University, Bikaner (Raj.) 334001 India
thanvi.megha@gmail.com

**Available online at: www.isca.in, www.isca.me**
Received 10th August 2025, revised 13th November 2025, accepted 28th December 2025

## Abstract

*The field of computational drug discovery employs Quantitative Structure-Activity Relationship (QSAR) methodologies to construct mathematical models that correlate molecular characteristics with biological effectiveness. These computational approaches enable researchers to understand how specific chemical features influence therapeutic outcomes, thereby facilitating rational drug development strategies. This investigation focused on constructing a predictive framework to estimate the antibacterial potency of 24 chemical entities against Staphylococcus aureus. The antimicrobial effectiveness was quantified through Minimum Inhibitory Concentration (MIC) measurements, while computational molecular parameters were utilized to establish the predictive framework. The entire QSAR analysis was conducted using ChemMaster 1.2 computational platform.*

**Keywords:** QSAR modelling, Predictive model, Drug design, Chemical structure, Cross validation, Transition Metals.

## Introduction

Quantitative Structure-Activity Relationship (QSAR) modelling serves as a powerful computational tool for predicting biological activities and toxicological effects based on molecular structure. QSAR creates mathematical correlations linking molecular architecture to biological function, demonstrating utility in discovering novel inhibitors and enhancing lead compounds for various conditions. The integration of computational strategies with conventional medicinal chemistry enhances hit discovery, lead enhancement, and ADME/toxicity forecasting[1].

QSAR modelling involves three primary components: biological data, physico-chemical or structural descriptors, and statistical methodologies. Superior, well-characterized biological data serves as the foundation for dependable models, requiring suitable and mechanistically interpretable descriptors. The criticality of external validation has been emphasized, with studies showing that leave-one-out cross-validated $R^2$ (LOO $q^2$) alone is insufficient for ensuring predictive capability. External validation utilizing independent test sets comprising at least five compounds spanning the activity and structural range is crucial for evaluating genuine predictive power[2,3].

For limited datasets, robust modelling strategies have been developed. Research on nitroimidazole sulfonamide analogues demonstrated the effectiveness of exhaustive double cross-validation without dividing limited datasets into training and test sets. Genetic algorithm-multiple linear regression and partial least squares regression were utilized with consensus modelling to enhance reliability, showing that computational QSAR approaches can reliably predict novel compound activity when experimental data are limited[4].

Multiple studies have demonstrated QSAR's effectiveness in antimicrobial research. Isatin-derived Schiff base compounds showed superior antibacterial activity compared to standard antibiotics, with QSAR analysis identifying key molecular descriptors including lipophilicity and polar surface area correlating with antimicrobial activity[5]. Metal complexes, particularly sulfonylhydrazine and sulfonylhydrazone complexes of Ni(II), Pt(II), Pd(II), and Cu(II), demonstrated varying antimicrobial activities, with 2D-QSAR analysis determining nucleophilic reaction indices and HOMO-LUMO energy gaps as key determinants[6].

8-Hydroxyquinoline derivatives analysis revealed that chloro substituents at the 5-position enhance antibacterial efficacy, with less polarized lipophilic chloro groups contributing to superior antibacterial activity[7]. Curcumin derivatives studies using Kernel-based Partial Least Squares method identified β-diketone moiety and 1,6-heptadiene as consistently enhancing antibacterial activity, while hydroxyl groups contributed positively and methoxy groups had unfavorable effects[8].

QSAR modelling has proven effective in catalyst design. Two-dimensional QSAR modelling of 58 transition metal complexes for ethylene oligomerization demonstrated excellent predictive capability ($R^2 = 0.913$ for training set, $R^2t = 0.971$ for test set), with effective net charge and maximum partial charge for nitrogen atoms identified as significant descriptors[9]. The approach successfully bridged computational predictions with experimental catalyst development.

Environmental applications include disinfection byproduct research, where QSAR models effectively link chemical activities with molecular structures for trend analysis and prediction[10]. Quantitative Ion Character-Activity Relationships (QICARs) successfully predicted metal toxicity using metal-ligand binding characteristics, providing mechanistic foundations for predicting metal bioactivity based on ion characteristics[11].

Machine learning-driven QSAR models for engineered nanoparticle mixture toxicity using Support Vector Machine and Neural Network techniques demonstrated superior accuracy, with the best model achieving $R^2$ test = 0.911, offering improved methods for environmental toxicology and risk assessment[12].

Various synthetic applications have been explored, including cinnamic acid derivatives where QSAR investigation using multiple linear regressions successfully correlated physicochemical parameters with biological activity[13]. Diarylpyrimidineanalogs research combined computational modelling with laboratory synthesis, establishing clear structure-activity correlations for antimicrobial therapeutics[14].

N-nitroso-2,6-diarylpiperidin-4-one derivatives investigation employed semiempirical computational methods, identifying ionization potential, dipole moment, and COSMO surface area as significant predictive factors[15]. Diorganotin (IV) coordination complexes studies demonstrated enhanced antimicrobial potency, with QSAR analysis correlating activity with molar refractivity indices[16,17].

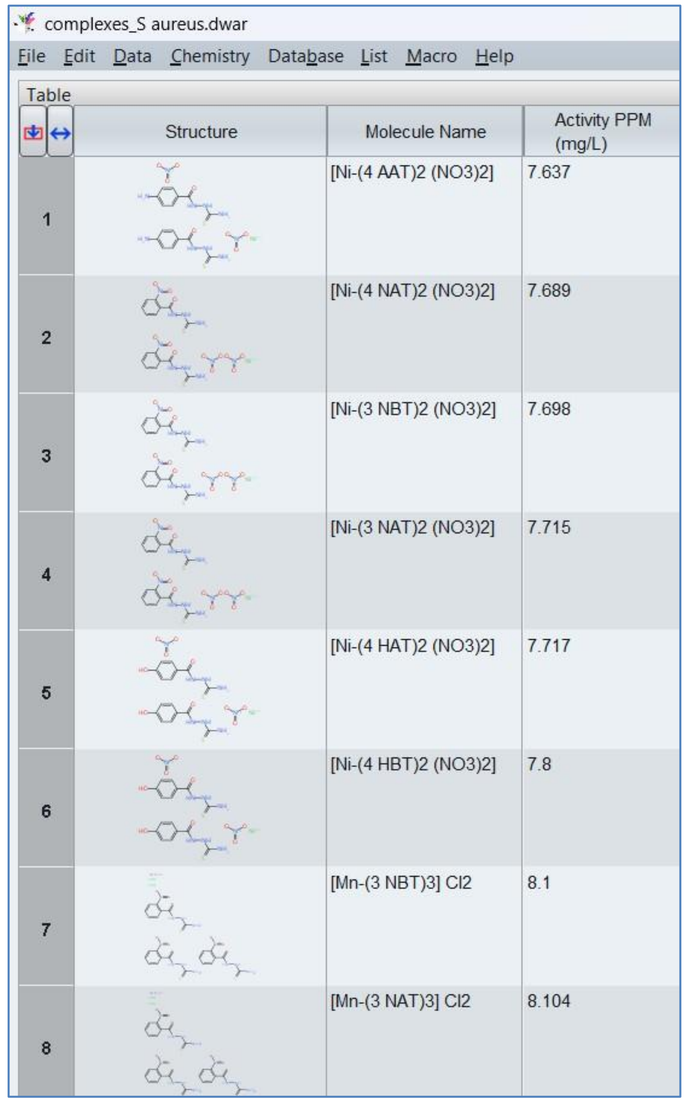
The comprehensive applications of QSAR methodologies across drug discovery, antimicrobial research, catalysis, and environmental science demonstrate the versatility and effectiveness of this computational approach. Studies collectively highlight the importance of rigorous model validation, high-quality data requirements, and the potential of computational approaches to accelerate research and development. While challenges remain in model reliability and validation, continued advancement of QSAR techniques offers promising opportunities for rational design and optimization of bioactive compounds and catalysts across diverse scientific disciplines.

## Materials and Methods

**Background and Objectives:** The field of computational drug discovery employs Quantitative Structure-Activity Relationship (QSAR) methodologies to construct mathematical models that correlate molecular characteristics with biological effectiveness. These computational approaches enable researchers to understand how specific chemical features influence therapeutic outcomes, thereby facilitating rational drug development strategies. This investigation focused on constructing a predictive framework to estimate the antibacterial potency of 24 chemical entities against *Staphylococcus aureus*. The antimicrobial effectiveness was quantified through Minimum Inhibitory Concentration (MIC) measurements, while computational molecular parameters were utilized to establish the predictive framework. The entire QSAR analysis was conducted using ChemMaster 1.2 computational platform.

## Data Processing and Preparation

**Dataset:** The 24 Transition metal complexes with the experimental antibacterial activity against S. aureus which were used for the development of QSAR predictive model.


complexes_S aureus.dwar
File Edit Data Chemistry Database List Macro Help
Table

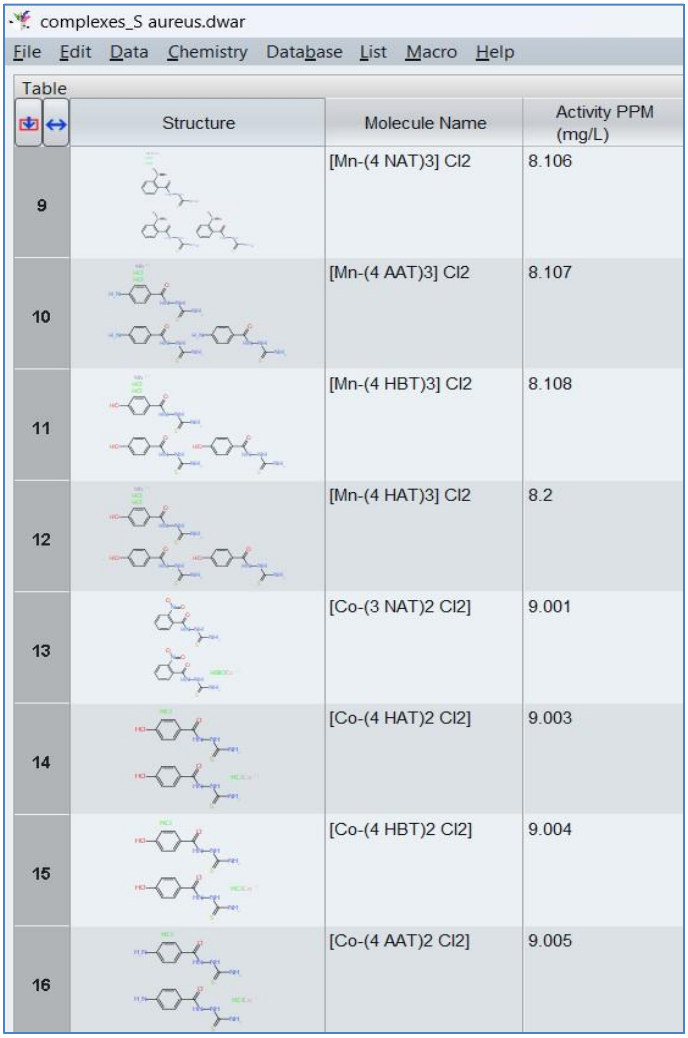
| | Structure | Molecule Name | Activity PPM (mg/L) |
|---|---|---|---|
| 1 | | [Ni-(4 AAT)2 (NO3)2] | 7.637 |
| 2 | | [Ni-(4 NAT)2 (NO3)2] | 7.689 |
| 3 | | [Ni-(3 NBT)2 (NO3)2] | 7.698 |
| 4 | | [Ni-(3 NAT)2 (NO3)2] | 7.715 |
| 5 | | [Ni-(4 HAT)2 (NO3)2] | 7.717 |
| 6 | | [Ni-(4 HBT)2 (NO3)2] | 7.8 |
| 7 | | [Mn-(3 NBT)3] Cl2 | 8.1 |
| 8 | | [Mn-(3 NAT)3] Cl2 | 8.104 |

**Figure-1:** Transition metal complexes as they appear in the software.

**Activity Data Conversion:** The antimicrobial potency data, initially expressed as MIC values in molar units, underwent logarithmic conversion according to the following mathematical relationship:

pMIC = - Log(MIC (M))

This mathematical conversion serves to normalize the activity data distribution and establish a more linear correlation between dependent and independent variables, thereby optimizing the data for linear regression analysis.



complexes_S aureus.dwar

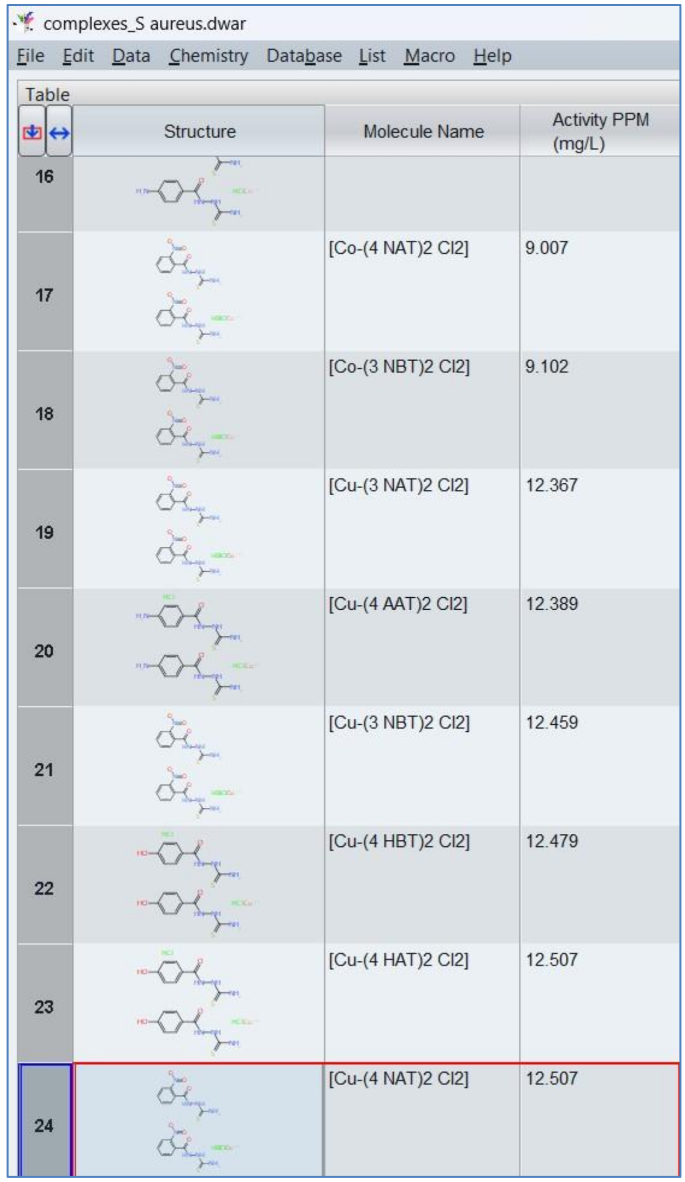
File Edit Data Chemistry Database List Macro Help

Table

| | Structure | Molecule Name | Activity PPM (mg/L) |
|---|---|---|---|
| 9 | | [Mn-(4 NAT)3] Cl2 | 8.106 |
| 10 | | [Mn-(4 AAT)3] Cl2 | 8.107 |
| 11 | | [Mn-(4 HBT)3] Cl2 | 8.108 |
| 12 | | [Mn-(4 HAT)3] Cl2 | 8.2 |
| 13 | | [Co-(3 NAT)2 Cl2] | 9.001 |
| 14 | | [Co-(4 HAT)2 Cl2] | 9.003 |
| 15 | | [Co-(4 HBT)2 Cl2] | 9.004 |
| 16 | | [Co-(4 AAT)2 Cl2] | 9.005 |

**Figure-2:** Transition Metal Complexes as they appear in the software.

**Chemical Property Calculations:** A comprehensive suite of pharmaceutically relevant molecular parameters was computed for each chemical structure. These computational descriptors were selected based on their biological significance and their established roles in drug disposition and molecular recognition processes: i. Molecular Weight (MW): Complete atomic mass of the chemical entity, ii. Heavy Atom Molecular Weight (HAMW): Combined atomic masses excluding hydrogen atoms, iii. Hydrogen Bond Donors (HBD): Count of proton-donating functional groups, iv. Hydrogen Bond Acceptors (HBA): Count of electron pair-donating sites, v. LogP: Partition coefficient between octanol and water phases, indicating lipophilic character, vii. Rotatable Bonds (RB): Count of freely rotating single bond linkages.



complexes_S aureus.dwar

File Edit Data Chemistry Database List Macro Help

Table

| | Structure | Molecule Name | Activity PPM (mg/L) |
|---|---|---|---|
| 16 | | | |
| 17 | | [Co-(4 NAT)2 Cl2] | 9.007 |
| 18 | | [Co-(3 NBT)2 Cl2] | 9.102 |
| 19 | | [Cu-(3 NAT)2 Cl2] | 12.367 |
| 20 | | [Cu-(4 AAT)2 Cl2] | 12.389 |
| 21 | | [Cu-(3 NBT)2 Cl2] | 12.459 |
| 22 | | [Cu-(4 HBT)2 Cl2] | 12.479 |
| 23 | | [Cu-(4 HAT)2 Cl2] | 12.507 |
| 24 | | [Cu-(4 NAT)2 Cl2] | 12.507 |

**Figure-3:** Transition Metal Complexes as they appear in the software.

**Data Partitioning Strategy:** The complete dataset underwent random segregation into distinct subsets: i. Training subset (75% allocation, comprising 18 chemical structures)-Employed for model construction and parameter optimization, ii. Testing subset (25% allocation, comprising 6 chemical structures)-Reserved for independent model performance assessment.

**Computational Model Construction:** Multiple Linear Regression (MLR) methodology was selected as the primary modelling approach based on its computational transparency, interpretability advantages, and established utility in QSAR applications. The mathematical model was developed using the training dataset, with logarithmic MIC values serving as the response variable and molecular descriptors functioning as predictor variables.

**Model Performance Assessment:** The reliability and predictive capability of the developed QSAR framework were evaluated through comprehensive statistical analysis applied to both training and testing datasets. The evaluation encompassed the following performance indicators: i. Coefficient of Determination ($R^2$): Quantifies the fraction of response variable variance captured by the model, ii. Root Mean Squared Error (RMSE): Represents the typical magnitude of prediction deviations, iii. Mean Absolute Error (MAE): Quantifies the average absolute deviation between predicted and observed values, iv. Leave-One-Out Cross-Validation $Q^2$ (LOO-$Q^2$): Assesses internal predictive capacity through systematic single-point omission and prediction, applied exclusively to the training dataset as an internal validation approach.

## Results and Discussion

**Mathematical Model Expression:** The final QSAR equation derived through MLR analysis is presented as:

$$pMIC = -1.0482\ X\ Molecular\ Weight - 1.2052\ X\ Heavy\ Atoms\ Molecular\ Weight + 0.2398\ X\ LogP + 0.9993\ X\ H - Bond\ Acceptors + 1.7312\ X\ H - Bond\ Donors + 1.5837\ X\ Rotatable\ Bonds + 4.8452$$

The mathematical coefficients reveal important structure-activity insights. The positive regression coefficients associated with hydrogen bonding capabilities (both donors and acceptors) indicate their favorable contribution to antimicrobial potency, presumably through enhanced molecular recognition and binding interactions with bacterial targets. In contrast, the negative coefficients for both total molecular weight and heavy atom weight suggest that increased molecular size may diminish biological activity, potentially due to reduced membrane penetration or steric interference with target binding sites. The modest positive contribution of lipophilicity (LogP) indicates a secondary role in activity determination, while the substantial positive coefficient for rotatable bonds suggests that enhanced conformational flexibility facilitates improved target engagement.

**Statistical Performance Metrics:** The comprehensive validation statistics are summarized in Table-1.

**Table-1:** Statistical validation metrics for the developed QSAR model across training and testing datasets.

| Set | $R^2$ | RMSE | MAE | $Q^2$ |
|---|---|---|---|---|
| Training | 0.999 | 0.003 | 0.002 | 0.998 |
| Test | 0.999 | 0.005 | 0.005 | |

The constructed MLR framework exhibits exceptional predictive performance, achieving an $R^2$ value of 0.999 for the training dataset, demonstrating that virtually all variance (99.9%) in antimicrobial activity is captured by the selected molecular parameters. The testing dataset $R^2$ of 0.999 confirms robust predictive performance on independent data, indicating excellent model generalizability. Both RMSE and MAE metrics demonstrate minimal prediction errors across both datasets, with slightly elevated errors in the testing set reflecting typical model behaviour on unseen data.

The correlation plot comparing experimental versus computational pMIC predictions demonstrates the model's predictive accuracy. Data points clustering near the identity line indicate high prediction quality and close agreement with experimental measurements. Enhanced proximity to the diagonal correlation line corresponds to improved predictive accuracy. This relationship is visualized in Figure-1.

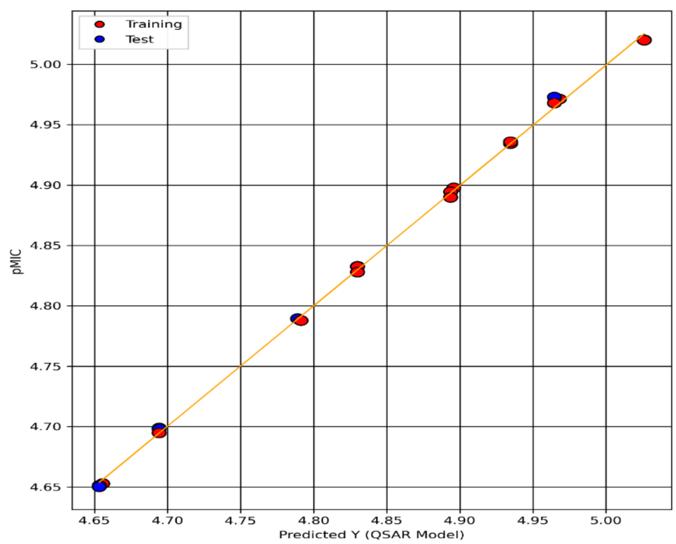



**Figure-1:** Correlation analysis between experimental antimicrobial activity (pMIC) and computational predictions from the QSAR model. Training dataset compounds are represented by red markers, while testing dataset compounds are shown as blue markers.

## Conclusion

The constructed QSAR framework demonstrates exceptional capability in predicting antimicrobial effectiveness against *S. aureus* using fundamental molecular characteristics. The model achieves outstanding statistical performance while maintaining excellent interpretability, establishing its value as a computational tool for early-phase drug discovery applications. Enhancement opportunities may include integration of additional molecular descriptors or implementation of non-linear modelling algorithms to further optimize predictive precision. This validated model provides a robust platform for virtual compound screening applications, enabling identification of promising new antimicrobial candidates against *S. aureus* infections.

**Terminology Reference**

| Abbreviation | Definition |
|---|---|
| QSAR | Quantitative Structure-Activity Relationship |
| MIC | Minimum Inhibitory Concentration |
| pMIC | Negative Logarithm of MIC |
| MW | Molecular Weight |
| HAMW | Heavy Atom Molecular Weight |
| HBD | Hydrogen Bond Donors |
| HBA | Hydrogen Bond Acceptors |
| LogP | Octanol-Water Partition Coefficient |
| RB | Rotatable Bonds |
| $R^2$ | Coefficient of Determination |
| RMSE | Root Mean Squared Error |
| MAE | Mean Absolute Error |
| $Q^2$ | Leave-One-Out Cross-Validation Coefficient |

## References

**1.** Shah, M., Patel, M., Shah, M., Patel, M., & Prajapati, M. (2024). Computational transformation in drug discovery: A comprehensive study on molecular docking and quantitative structure activity relationship (QSAR). *Intelligent Pharmacy,* 2, 589-595.

**2.** Cronin, M. T. D., & Schultz, T. W. (2003). Pitfalls in QSAR. *Journal of Molecular Structure: THEOCHEM*, 622, 39-51.

**3.** Golbraikh A., & Tropsha, A. (2002). Beware of q2!. *Journal of Molecular Graphics & Modelling,* 20, 269-276.

**4.** De, P., & Roy, K. (2021). QSAR and QSAAR modeling of nitroimidazole sulfonamide radiosensitizers: Application of small dataset modeling. *Structural Chemistry,* 32, 631-642.

**5.** Mishra, R., Chaurasia, H., Singh, V. K., Naaz, F., & Singh, R. K. (2021). Molecular modeling, QSAR analysis and antimicrobial properties of Schiff base derivatives of isatin. *Journal of Molecular Structure,* 1243, 1-13.

**6.** Özbek, N., Alyar, S., Alyar, H., Sahin, E., & Karacan, N. (2013). Synthesis, characterization and anti-microbial evaluation of Cu(II), Ni(II), Pt(II) and Pd(II) sulfonylhydrazone complexes; 2D-QSAR analysis of Ni(II) complexes of sulfonylhydrazone derivatives. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy,* 108, 123-132.

**7.** Harsh, N. K., Beniwal, R. C., Sharma, S. D., Saran, R. K., & Bhojak, N. (2022). QSAR study of 8-Hydroxyquinoline and its Chloro derivatives as antimicrobial agents. *Bulletin of Environment, Pharmacology and Life Sciences, Special Issue* (1), 1627-1629.

**8.** Tebbaai, H., El Hamdani, N., Byadi, S., & Aboulmouhajir, A. (2024). Predicting antibacterial activity of some Curcumin compounds: Fingerprint-based 2D-QSAR models. *Scientific African,* 26, 1-17.

**9.** Yanga, W., Yia, J., Maa, Z., & Sun, W. H. (2017). 2D-QSAR modeling on the catalytic activities of 2-azacyclyl-6-aryliminopyridylmetal precatalysts in ethylene oligomerization. *Catalysis Communications*, 101, 40-43.

**10.** Chen, B., Zhang, T., Bond, T., & Gan, Y. (2015). Development of quantitative structure activity relationship (QSAR) model for disinfection byproduct (DBP) research: A review of methods and resources. *Journal of Hazardous Materials,* 299, 260-279.

**11.** Abramenko, N., Kustova, L., Metelytsia, L., Kovalishyn, V., Tetko, I., & Peijnenburg, W. (2020). A review of recent advances towards the development of QSAR models for toxicity assessment of ionic liquids. *Journal of Hazardous Materials,* 384, 1-14.

**12.** Zhang, F., Wang, Z., Peijnenburg, W. J. G. M., & Vijver, M. G. (2023). Machine learning-driven QSAR models for predicting the mixture toxicity of nanoparticles. *Environment International,* 177, 1-9.

**13.** Narasimhan, B., Belsare, D., Pharande, D., Mourya, V., & Dhake, A. (2004). Esters, amides and substituted derivatives of cinnamic acid: Synthesis, antimicrobial activity and QSAR investigations. *European Journal of Medicinal Chemistry,* 39, 827-834.

**14.** Singh, V. K., Chaurasia, H., Mishra, R., Srivastava, R., Naaz, F., Kumar, P., & Singh, R. K. (2022). Docking, ADMET prediction, DFT analysis, synthesis, cytotoxicity, antibacterial screening and QSAR analysis of diarylpyrimidine derivatives. *Journal of Molecular Structure*, 1247, 1-16.

**15.** Hemalatha, T., Imran, P. K. M., Gnanamani, A., & Nagarajan, S. (2008). Synthesis, antibacterial and antifungal activities of some N-nitroso-2,6-diarylpiperidin-4-one semicarbazones and QSAR analysis. *Nitric Oxide,* 19, 303-311.

**16.** Devi, J., Pachwania, S., Yadav, J., & Kumar, A. (2021). Pentacoordinateddiorganotin (IV) complexes resulting from tridentate (NOO) donor Schiff bases: Synthesis, characterization, in vitro antioxidant, antimicrobial activities, and QSAR studies. *Phosphorus, Sulfur, and Silicon and the Related Elements,* 196, 119-132.

**17.** Matveieva, M., & Polishchuk, P. (2021). Benchmarks for interpretation of QSAR models. *Journal of Cheminformatics,* 13, 1-20.