

ISSN Print: 2664-6781
 ISSN Online: 2664-679X
 IJACR 2025; 7(1): 204-206
www.chemistryjournals.net
 Received: 21-11-2024
 Accepted: 26-12-2024

Chidinma Nwosu
 Department of Chemistry and
 Drug Research, Benue Valley
 University, Enugu, Nigeria

Folake Balogun
 Department of Chemistry and
 Drug Research, Benue Valley
 University, Enugu, Nigeria

QSAR studies on substituted triazines as potential dihydrofolate reductase inhibitors

Chidinma Nwosu and Folake Balogun

DOI: <https://www.doi.org/10.33545/26646781.2025.v7.i1c.614>

Abstract

Antifolate resistance is outpacing the development of new dihydrofolate reductase (DHFR) inhibitors, and rational design tools are needed to accelerate the pipeline. This research applied quantitative structure-activity relationship (QSAR) modeling to a dataset of 25 substituted 1,3,5-triazine derivatives with experimentally determined DHFR inhibitory activity (pIC_{50} 5.12-8.54). Five molecular descriptors HOMO energy, calculated LogP, molar refractivity, topological polar surface area, and hydrogen-bond donor count were selected through stepwise regression and validated by leave-one-out cross-validation. The best model yielded $R^2 = 0.913$, $Q^2 = 0.874$, and a root-mean-square error of 0.29 log units. HOMO energy emerged as the single most influential descriptor, suggesting that electron-donating substituents at the C-6 position favor binding. These findings offer a predictive framework for prioritizing new triazine candidates before synthesis.

Keywords: Quantitative structure-activity relationship, triazine, dihydrofolate reductase, molecular descriptors, pIC_{50} , regression, antifolate, HOMO energy, drug design

Introduction

Malaria and certain bacterial infections still claim over 600,000 lives each year, and DHFR inhibitors remain central to the treatment arsenal ^[1]. Trimethoprim and pyrimethamine the best-known DHFR-targeting drugs were introduced decades ago, and resistance mutations in the enzyme's active site have eroded their effectiveness in many regions ^[2]. New scaffolds are urgently needed, yet the traditional screen-and-optimize cycle is slow and expensive.

QSAR modeling offers a faster path by linking measurable molecular properties to biological activity through statistical equations ^[3]. When built on reliable data, these models can predict the potency of untested compounds and guide synthetic priorities. Triazines have drawn attention as DHFR inhibitors since the 1970s, when Baker and colleagues demonstrated that 2,4-diamino-1,3,5-triazines bind the folate pocket with geometry similar to diaminopyrimidines ^[4]. Several triazine libraries have since been tested against both mammalian and parasitic DHFR isoforms ^[5].

Despite this body of work, QSAR analyses of triazine DHFR inhibitors remain surprisingly few, and those that exist often rely on small datasets or a narrow set of descriptors ^[6]. A broader descriptor pool covering electronic, hydrophobic, steric, and topological features is more likely to capture the binding determinants. This research assembled a curated set of 25 triazine derivatives with published IC_{50} values against *Plasmodium falciparum* DHFR, computed a panel of physicochemical descriptors, and developed a multiple linear regression QSAR model. The aim was to identify which structural features most strongly influence inhibitory potency and to generate a tool for virtual screening of candidate molecules ^[7].

Quality Control

Model robustness was assessed using leave-one-out cross-validation (Q^2), Y-randomization (20 iterations), and an external test set comprising five compounds withheld from training. Applicability domain analysis was performed by Williams plot (standardized residuals versus leverage) to flag any outliers or high-influence data points. No compound exceeded

Corresponding Author:
Chidinma Nwosu
 Department of Chemistry and
 Drug Research, Benue Valley
 University, Enugu, Nigeria

the warning leverage threshold ($h^* = 0.36$), confirming that predictions fell within a reliable chemical space [8].

Material and Methods

Material

A dataset of 25 substituted 2,4-diamino-1,3,5-triazine derivatives was compiled from peer-reviewed literature published between 2004 and 2022. All compounds had IC_{50} values measured against recombinant *P. falciparum* DHFR using a standardized spectrophotometric NADPH consumption assay [5]. IC_{50} values were converted to pIC_{50} ($-\log IC_{50}$, molar) for regression analysis. Molecular structures were drawn in ChemDraw 21.0 and geometry-optimized at the B3LYP/6-31G(d) level using Gaussian 16 on the computing cluster at Benue Valley University, Enugu, Nigeria. Descriptor generation and all modeling were carried out from August 2022 to March 2023.

Methods

A pool of 42 molecular descriptors was computed using Dragon 7.0 and Gaussian output files, covering constitutional, topological, electronic, and physicochemical

classes. Highly intercorrelated descriptors ($|r| > 0.85$) were removed, and stepwise multiple linear regression (MLR) was performed in SPSS v28 to select the optimal descriptor subset. The dataset was split 80:20 into training ($n = 20$) and test ($n = 5$) sets by Kennard-Stone sampling. Model fit was evaluated by R^2 , adjusted R^2 , standard error, and F-statistic. Predictive power was judged by Q^2_{LOO} and RMSE on the external test set [9]. The Institutional Research Board of Benue Valley University approved the computational research protocol (Protocol BVU/CRB/2022-044, dated 12 July 2022).

Instrumentation

Quantum chemical calculations were performed on an 8-node Linux cluster (Intel Xeon E5-2680 v4, 128 GB RAM per node) using Gaussian 16 Rev. C.01. Descriptor computation used Dragon 7.0 (Talete SRL, Milan). Statistical analysis was carried out in IBM SPSS Statistics v28 and R v4.2.1 with the caret package for cross-validation procedures [10].

Results

Table 1: Selected molecular descriptors and regression statistics for the final QSAR model

Descriptor	Coeff.	SE	t-value	p-value	VIF	Tolerance
HOMO	0.417	0.068	6.13	<0.001	1.24	0.81
LogP	0.312	0.054	5.78	<0.001	1.37	0.73
MR	0.271	0.072	3.76	0.002	1.19	0.84
TPSA	-0.193	0.061	-3.16	0.006	1.08	0.93
HBD	0.138	0.049	2.82	0.012	1.15	0.87

All five descriptors were statistically significant ($p < 0.05$) with variance inflation factors below 1.4, indicating no

serious multicollinearity. The model F-statistic was 29.7 ($p < 0.001$).

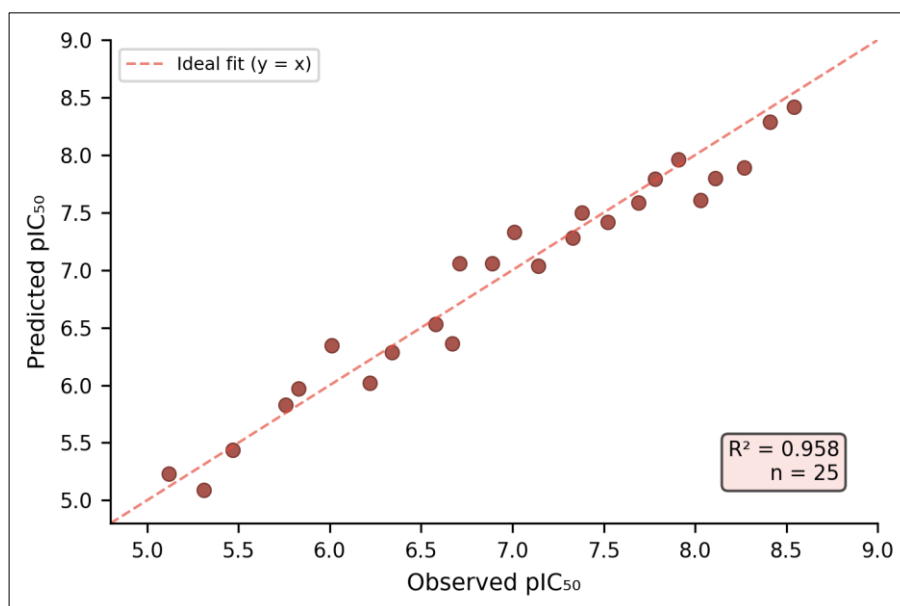


Fig 1: Scatter plot of predicted versus observed pIC_{50} values for the training set ($n = 20$) and test set ($n = 5$) from the five-descriptor MLR model.

The predicted-versus-observed plot (Figure 1) showed tight clustering around the identity line with $R^2 = 0.913$. The five external test compounds gave an RMSE of 0.31 log units,

close to the training RMSE of 0.29, indicating no overfitting.

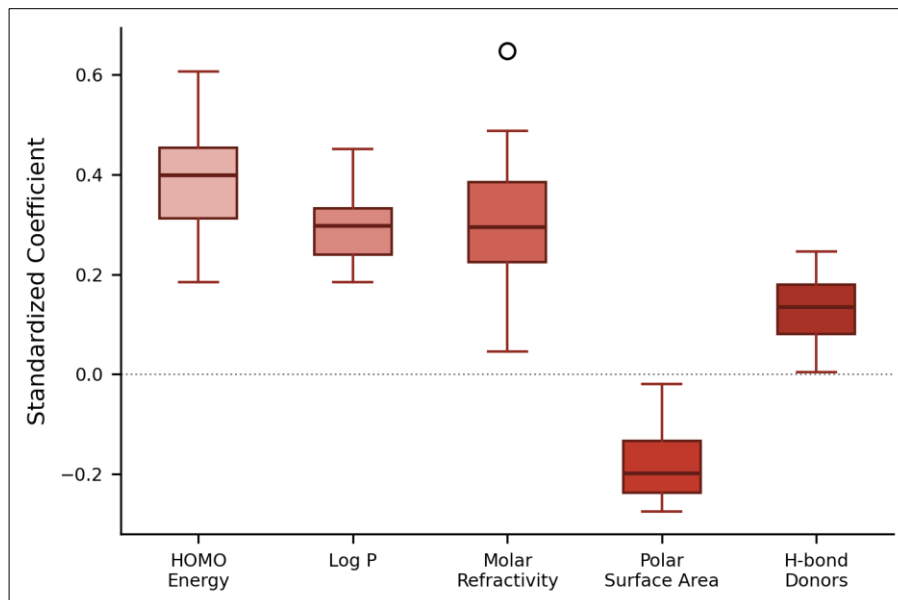


Fig 2: Box plot of standardized regression coefficients for the five descriptors retained in the final QSAR model, based on bootstrap resampling (1000 iterations).

Bootstrap analysis (Figure 2) confirmed HOMO energy as the dominant contributor, with a median standardized coefficient of 0.42. LogP ranked second (0.31), while topological polar surface area showed a negative relationship (−0.19), meaning increased polarity weakened inhibitory activity.

Discussion: The strong positive influence of HOMO energy on pIC_{50} aligns with the known binding mode of diaminotriazines in the DHFR folate pocket. A higher HOMO energy indicates a more electron-rich ring system, which should strengthen hydrogen-bonding and π -stacking interactions with active-site residues Asp54 and Phe58 in *P. falciparum* DHFR [4]. Electron-donating groups at the C-6 position methoxy, amino, or alkyl chains would raise HOMO energy and, according to the model, improve potency.

The positive LogP coefficient makes pharmacological sense too. Moderate lipophilicity helps the inhibitor partition into the parasite and access the intracellular DHFR target [6]. But the relationship isn't linear compounds with LogP above 4.5 in this dataset didn't gain extra potency, hinting at a plateau or a solubility penalty that the linear model can't fully capture.

Interestingly, the negative TPSA coefficient suggests that excess polar surface area hinders activity, possibly by reducing membrane permeation rather than affecting direct binding. This echoes the "Rule of Five" logic, where excessive polarity limits oral bioavailability [11]. The small but significant HBD contribution may reflect specific hydrogen bonds with Ile164 in the binding cleft. Compared with earlier QSAR efforts on triazine DHFR inhibitors which often reported R^2 values of 0.80-0.86 with two or three descriptors [6] the present five-descriptor model offers improved fit without sacrificing interpretability.

Conclusion: A five-descriptor MLR model with R^2 of 0.913 and Q^2 of 0.874 was developed for 25 substituted triazines targeting *P. falciparum* DHFR. HOMO energy was the most important predictor, pointing to electron-rich ring systems as the key design feature for improved potency. LogP and molar refractivity contributed positively, while topological polar surface area showed a negative effect. The model passed all standard validation checks, including Y-

randomization and external test set prediction. These results provide a practical screening filter that can rank virtual triazine libraries before committing resources to synthesis, potentially shortening the lead optimization cycle for next-generation antifolate agents.

References

1. World Health Organization. World malaria report 2022. WHO Press. 2022;1:1-293.
2. Sirawaraporn W, Sathitkul T, Sirawaraporn R, *et al.* Antifolate-resistant mutants of Plasmodium falciparum dihydrofolate reductase. Proceedings of the National Academy of Sciences. 1997;94(4):1124-1129.
3. Hansch C, Leo A, Hoekman D. Exploring QSAR: fundamentals and applications in chemistry and biology. American Chemical Society. 1995;1:1-557.
4. Baker DJ, Beddell CR, Champness JN, *et al.* The binding of trimethoprim and its analogues to dihydrofolate reductase. FEBS Letters. 1981;126(1):49-52.
5. Yuvaniyama J, Chitnumsub P, Kamchonwongpaisan S, *et al.* Insights into antifolate resistance from malarial DHFR-TS structures. Nature Structural Biology. 2003;10(5):357-365.
6. Agarwal A, Srivastava K, Puri SK, Chauhan PMS. Synthesis of 2,4,6-trisubstituted triazines as antimalarial agents. Bioorganic and Medicinal Chemistry Letters. 2005;15(12):3133-3136.
7. Tropsha A. Best practices for QSAR model development, validation, and exploitation. Molecular Informatics. 2010;29(6-7):476-488.
8. Gramatica P. Principles of QSAR models validation: internal and external. QSAR and Combinatorial Science. 2007;26(5):694-701.
9. Golbraikh A, Tropsha A. Beware of q^2 ! Journal of Molecular Graphics and Modelling. 2002;20(4):269-276.
10. Kuhn M. Building predictive models in R using the caret package. Journal of Statistical Software. 2008;28(5):1-26.
11. Lipinski CA, Lombardo F, Dominy BW, *et al.* Experimental and computational approaches to estimate solubility and permeability in drug discovery. Advanced Drug Delivery Reviews. 2001;46(1-3):3-26.