STM JOURNALS

International Journal of
Bioinformatics and Computational
Biology

ISSN: 2583-9977
Volume 3, Issue 2, 2025
July—December
DOI (Journal): 10.37591/IJBCB

https://journals.stmjournals.com/ijbcb

Exploring Antimalarial Activity of Chalcone Derivatives Through QSAR

Rahila Khan[1], Neerja Shukla[2,*], Anita Singh[1], Anushka Shukla[2], Kumari Suman[2], Bechan Sharma[3], Shashikant Verma[4]

Abstract

Background: *The core structure of chalcones contains a reactive α, β-unsaturated system within the aromatic rings, which plays a key role in mediating various biological effects. These effects include enzyme inhibition, anticancer activity, anti-inflammatory properties, as well as antibacterial, antifungal, anti-malarial, antiprotozoal, and anti-filarial actions. Modifying the structure by introducing substituent groups to the aromatic ring can enhance potency, reduce toxicity, and expand their range of pharmacological actions.* ***Methods:*** *A total of twenty-seven chalcone derivatives were analyzed for their physicochemical and molecular properties. Chalcones exhibit a broad spectrum of pharmacological properties, encompassing antibacterial, antifungal, anti-inflammatory, and anticancer activities, making them promising candidates for therapeutic development, and molecular connectivity (χ), were computed using E-Dragon software. The software primarily accepts molecular file formats, such as SMILES notations, which were generated online using Babel software. These 2D structures were then converted into 3D optimized structures using CORINA, an online tool provided by Molecular Networks GmbH. Furthermore, the 3D structures of the compounds were constructed using Gauss View software to calculate various quantum chemical descriptors based on density functional theory (DFT). The analysis included key values, such as total energy, softness, hardness, chemical potential, and the energy levels of the molecule's outermost electrons – specifically, the highest occupied (HOMO) and lowest unoccupied (LUMO) orbitals. All the compounds were completely optimized in the gas phase using the 6-31+G* basis set for accurate structural calculations. To ensure that the optimized structures represented true energy minima, harmonic vibrational frequency calculations were carried out—confirmed by the presence of only positive vibrational frequencies.* ***Results:*** *A comprehensive array of molecular descriptors – including D, ID, IOR, Log P, Mr, Mv, Mw, Pc, BAC, Pz, St, W, Wa, 0c, 1c, 2c, 3c, 4c, 5c, and Xeq – has been recognized as crucial for effectively modeling the biological activity of the investigated compounds. These descriptors serve as fundamental parameters in capturing the structural and physico-chemical complexities necessary for accurate predictive analysis. QSAR Model 2, mathematically represented by the following equation:* $pIC50 = -20.605\ (\pm 6.600)\ IOR - 0.747\ (\pm 0.454)\ I1 - 5.083\ (\pm 3.478)\ Xeq + 51.647$ *This model was systematically refined using empirical parameters to enhance its predictive accuracy. Each coefficient in the equation signifies the relative influence of the corresponding molecular descriptor on the predicted biological activity. The negative values associated with IOR, I1, and Xeq indicate an inverse correlation between these descriptors and the pIC50 values, suggesting their*

***Author for Correspondence**
Neerja Shukla
E-mail: drneerjashukla28@gmail.com

[1]Assistant Professor, Department of Chemistry, Nehru Gram Bharati University (Deemed to be University), Prayagraj, Uttar Pradesh, India
[2]Research Scholar, Department of Chemistry, Nari Shiksha Niketan PG College, Mona Chandrawati Gupta Marg, Kaiserbagh, Lucknow, Uttar Pradesh, India
[3]Professor, Department of Biochemistry, Faculty of Science, University of Allahabad, Prayagraj, Uttar Pradesh, India
[4]Research Scholar, Department of Chemistry, Nari Shiksha Niketan PG College, Mona Chandrawati Gupta Marg, Kaiserbagh, Lucknow, Uttar Pradesh, India

Received Date: February 15, 2025
Accepted Date: May 03, 2025
Published Date: May 03, 2025

Citation: Rahila Khan, Neerja Shukla, Anita Singh, Anushka Shukla, Kumari Suman, Bechan Sharma, Shashikant Verma. Exploring Antimalarial Activity of Chalcone Derivatives Through QSAR. International Journal of Bioinformatics and Computational Biology. 2025; 3(2): 1–6p.

role in modulating compound potency. The optimization process involved fine-tuning these parameters to achieve the best fit with the experimental data, ensuring a robust and reliable depiction of the structure-activity relationship. ***Conclusions:*** *The QSAR model suggests that substituents with lower refractive indices and less electronegative groups improve activity. In contrast, indomethacin derivatives containing a* $-CH_2CH_2NHCONH(CH_2)_3ONO_2$ *group at the R1 position negatively impact activity. The findings are critically analyzed using regression data and cross-validation techniques. The Poglani factor (Q) and results from the LOO (leave-one-out) method validate the reliability and predictive accuracy of the proposed models, offering valuable insights for designing new analogues with improved potency in the future.*

Keywords: Indomethacin derivatives, Poglani Factor (Q), LOO (Leave-One-Out), anti-inflammatory activity, potency

INTRODUCTION

Malaria is a potentially fatal disease that is mostly found in tropical areas. While it is preventable and treatable, without prompt diagnosis and proper treatment, malaria can develop into a severe and often deadly form. The disease is not contagious and does not spread directly between people; instead, it is transmitted through the bites of female Anopheles mosquitoes. Malaria is a serious and potentially life-threatening illness caused by five different species of Plasmodium parasites that can infect humans. Among these species, Plasmodium falciparum and Plasmodium vivax are the most harmful, presenting major risks to global public health. *Plasmodium falciparum* is the primary cause of severe and fatal malaria cases worldwide, while *Plasmodium vivax* is particularly concerned due to its ability to remain dormant in the liver, leading to recurrent infections. Despite considerable progress in prevention, treatment, and control measures, malaria remains a pressing global health concern. It continues to contribute to high mortality rates, particularly in tropical and subtropical regions where transmission is widespread and persistent. Reports of parasite resistance to both chloroquine and artemisinin [1, 2] highlight the urgent need to develop new, effective antimalarial [3–7] drug candidates.

Chalcones have exhibited a wide range of pharmacological effects, such as anticancer, anti-inflammatory, and anti-HIV activities. Additionally, chalcone derivatives with arenyl and allyl substituents have shown promising antimalarial activity. This manuscript focuses on identifying the most efficient antimalarial drugs using in silico methods, aiming to reduce the cost, time, and energy involved, ultimately benefiting society.

MATERIALS AND METHODS

The 2D structures of twenty-seven chalcone derivatives were meticulously designed using ACD Lab Chem Sketch 3-8 (version 10) software. A wide variety of physicochemical and hydrophobic properties – such as molecular weight (Mw), molar volume (Mv), molar refractivity (Mr), parachor (Pc), density (D), refractive index (IOR), surface tension (St), polarizability (Pz), and partition coefficient (Log P) – were systematically computed using dedicated computational tools. To further explore the molecular topology, key descriptors, such as the Wiener index (W), mean Wiener index (Wa), Balaban index (J), Balaban centric index (BAC), and molecular connectivity (χ) were calculated using the E-Dragon software. Additionally, the SMILES notation, a widely used molecular representation format compatible with E-Dragon, was generated efficiently using Babel software. The transition from 2D to 3D molecular representations was achieved using CORINA, an advanced online tool provided by Molecular Networks GMBH, to generate optimized three-dimensional structures of the chalcone derivatives. Furthermore, Gauss View software was utilized to construct 3D molecular geometries and facilitate the computation of quantum chemical descriptors based on density functional theory (DFT). These quantum parameters included total energy (TE), softness (S), hardness (η), chemical potential (μ), as well as the energies of the highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) [8–13]. All molecular structures were fully optimized in the gas phase using the 6-31+G* basis set. To validate the accuracy of the optimized geometries, harmonic vibrational frequency calculations were performed, confirming that each structure corresponded to an energy minimum, as indicated by the presence of positive vibrational frequencies. Multiparametric regression analysis, based on structural parameters and inhibitory activity, was performed using Hansch analysis [14] and executed with SPSS (7.5) software. The inhibitory activity values and different analogues of the parent structure for the present QSAR studies are provided in Figure 1 and Table 1.

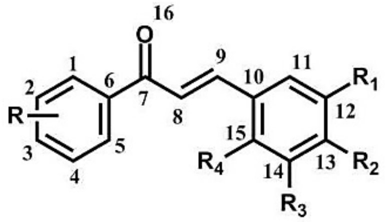

Figure 1. Parent structure.

Table 1. Structural details with biological activity and indicator parameters for the series of Indomethacin derivatives.

Comp.	**R**	**R1**	**R2**	**R3**	**R4**	**pIC$_{50}$**
1	4-OH	Allyl	OH	OCH_3	H	4.414
2	4-Cl	Allyl	OH	OCH_3	H	4.552
3	4-Cl	Allyl	OCH_3	OCH_3	H	5.408
4	4-OCH_3	Allyl	OCH_3	OCH_3	H	5.366
5	3,4-OCH_2O–	Allyl	OCH_3	OCH_3	H	5.327
6	4-Br	Allyl	OCH_3	OCH_3	H	5.275
7	4-NO_2	Allyl	OCH_3	OCH_3	H	4.903
8	4-Cl	Allyl	O-Allyl	OCH_3	H	5.107
9	3,4-diCl	H	O-Allyl	OCH_3	H	5.000
10	4-Br	H	O-Allyl	OCH_3	H	5.091
11	4-F	H	O-Allyl	OCH_3	H	4.638
12	3-Cl	H	O-Allyl	OCH_3	H	5.275
13	4-OCH_3	H	O-Allyl	OCH_3	H	4.647
14	3,4-OCH_2O-	H	O-Allyl	OCH_3	H	4.414
15	4-NH_2	H	O-Allyl	OCH_3	H	4.443
16	H	H	O-Allyl	OCH_3	H	4.420
17	4-O-Allyl	H	O-Allyl	OCH_3	H	5.130
18	4-Cl	H	H	OCH_3	O-Allyl	4.903
19	4-Cl	H	OCH_3	O-Allyl	H	4.744
20	4-Cl	H	OH	OCH_3	H	4.552
21	4-Cl	H	O-Butyl	OCH_3	H	4.829
22	4-Cl	H	O-CH_2-Ph-4-Br	OCH_3	H	5.096
23	4-Cl	O–CH_3	O-Allyl	OCH_3	H	5.420
24	4-Cl	H	O-Allyl	H	H	4.366
25	4-Cl	H	H	H	O-Allyl	5.045
26	4-Cl	H	O-Allyl	O-C_2H_5	H	4.552
27	3-O-Allyl4-OCH_3	H	Cl	H	H	5.065

In multiparametric modelling [15], ensuring the development of a statistically significant model [16] requires first examining whether collinearity exists among the parameters used. Cross-validation analysis was conducted for all models using the leave-one-out (LOO) method [17–19], where one compound is excluded from the dataset, and the activity is predicted based on the remaining data.

RESULTS AND DISCUSSION

Multiparametric Modeling Using Indicator Parameters

Different combinations of parameters having no autocorrelation together with indicator parameters were used and the statistically significant models obtained are presented in the form of equations.

$$pIC50 = 0.0125 \times (V) - 0.147 \times (MR) + 0{,}0115 \times (MW) + 1.5481 \times (qC_{12}) + 35.1286 \times (qC_{10}) +- 0{,}6037$$

$$pIC50 = 0.0082 \times (SA) + 0.0122 \times (MW)\ 0.0709 \times (MR) + 22.3351 \times (E\_HOMO) + 50.389 \times (qC10) + 1.4152 \times (qC12) + 3.1403$$

The combination of various descriptors, such as MR, MV, Mw, and V, has been identified as significant for modeling the activity of chalcone derivatives. All the models produced satisfactory results, with their significance and quality assessed using statistical metrics, including "R," "R^2," "R^2A," quality factor "Q," and the standard error of estimate (SE) [20]. Among the statistically significant models, Models 1 and 2 demonstrated superior statistical parameters. To confirm that models with excellent statistics also possess strong predictive capabilities, the quality factor [18]. (Q) was evaluated. The Pogliani Q parameter further validated the exceptional predictive power of these models. Additionally, the low standard error of estimate underscores the statistical significance and outstanding predictive performance of these equations.

The cross-validated R2R^2 values for each model were found to be very similar to the R2R^2 values for the full dataset, indicating that the models are statistically significant, as shown in Table 2. Cross-validation [21]. provides key metrics, including PRESS, SSY, PSE, RCV2R^2_{CV}, and RA2R^2_A, which allow for the evaluation of the predictive accuracy of Models 1 and 2.

Table 2. QSAR models and statistical properties.

Model	N	R	R^2	SE	F	Fcal/Ftheo	PRESS
1	27	0.912	0.831	0.170	15.839	5.342	0.495
2	27	0.935	0.873	0.152	17.331	6.075	0.349

In this context, "n" refers to the total number of data points considered, "R" represents the correlation coefficient, which measures the strength of the relationship between variables, and "R^2" denotes the coefficient of determination, quantifying the proportion of variance in the dependent variable that is accounted for by the model. The "F" value indicates the variance ratio between observed and predicted activities, with confidence intervals at the 95% level shown in parentheses. Monoparametric models were deemed inadequate for activity prediction due to their poor statistical performance, leading to the adoption of multiparametric models. When developing these multiparametric models, parameters with a correlation coefficient (R) greater than 0.5 in the correlation matrix were excluded to prevent collinearity. Only parameters without autocorrelation are considered for regression studies. While many parameters lack collinearity, this study focuses solely on reporting statistically significant multiparametric equations.

Using cross-validation techniques, several modified datasets were generated by systematically excluding one (leave-one-out, LOO) or a small group (leave-some-out) of data points. For each dataset, an input-output model was developed using the chosen modeling technique. Each model's accuracy was assessed based on its ability to predict the responses of excluded data points. The number of models generated by the LOO procedure was equal to the total number of data points (n).

Cross-validation prioritizes a model's predictive accuracy over its data-fitting capability, making it a key determinant of a proposed model's predictive power. This is crucial, as a model with strong statistical parameters may not necessarily possess robust predictive potential. By employing this methodology, the robustness and reliability of the QSAR models are ensured.

The predicted residual sum of squares (PRESS) serves as a strong measure of a model's true predictive performance. When the PRESS value is lower than the total sum of squares (SSY), it indicates that the model's predictions exceed random chance, highlighting its statistical significance. Additionally, the PRESS-to-SSY ratio provides a practical approach for estimating confidence intervals, enhancing the reliability and applicability of the model in forecasting the activity of new compounds. For a QSAR model to be considered reasonable, the PRESS/SSY ratio should be below 0.4. Furthermore, by transforming the PRESS value into a dimensionless term relative to the initial sum of squares, we obtain R^2, which represents the proportion of explained variance versus total variance. Both PRESS and R^2 are known to have favorable properties.

In this study, Models 1 and 2 exhibited R^2 values ranging from 0.831 to 0.873, indicating strong predictive performance. However, for practical implementation, the predictive square error (PSE), derived as the square root of PRESS/n, serves as a more direct measure of predictive uncertainty. The PSE values further validate the reliability and robustness of our results. Moreover, the adjusted coefficient of determination (R^2A) offers a more precise assessment of the model's explanatory strength by factoring in both the sample size and the complexity introduced by multiple independent variables, ensuring a well-balanced evaluation of model effectiveness. Unlike R^2, which always increases when new variables are added, R^2A accounts for model complexity. Both models 1 and 2 meet all the criteria for an ideal QSAR model [22]. The findings from this study are summarized in a graphical abstract.

CONCLUSIONS

The main goal of this study was to discover chalcone derivatives that could function as more potent alternatives. The QSAR models presented indicate promising directions for malaria therapeutics. Specifically, the negative coefficient for Molar Refractivity (MR) suggests that decreasing its value could potentially improve the activity of chalcone derivatives. In contrast, the positive coefficients for Volume (V), Molecular Weight (Mw), and Surface Area (SA) suggest that increasing these parameters may enhance the compound's efficacy. QSAR model (2), which incorporates empirical parameters and demonstrates robust statistical performance ($R = 0.935$, $R^2 = 0.873$), highlights that substituents with lower molar refractivity and higher values of V, Mw, and SA are beneficial for optimizing activity. These findings play a pivotal role in guiding the strategic design and development of next-generation antimalarial drugs. By offering deeper insights into the molecular characteristics that influence biological activity, they support the rational creation of more potent and selective therapeutic agents. This knowledge can be effectively utilized to refine drug candidates, enhance their efficacy, and improve their pharmacokinetic and pharmacodynamic properties, ultimately contributing to the development of superior treatment options in the ongoing battle against malaria.

Acknowledgments

The authors NS, SKV, AS, KS are grateful to the Department of Chemistry, Nari Shiksha Niketan PG College, Mona Chandrawati Gupta Marg, Kaiserbagh, Lucknow, AS, RK are grateful to Department of Chemistry, Nehru Gram Bharati University (Deemed to be University), Prayagraj and BS is grateful to Department of Biochemistry, University of Allahabad, Prayagraj for providing facilities to carry out the present work.

REFERENCES

1. Ramos-Inza S, Encio I, Raza A, Sharma AK, Sanmartin C, Plano D. Design, synthesis and anticancer evaluation of novel Se-NSAID hybrid molecules: Identification of a Se-indomethacin analog as a potential therapeutic for breast cancer. Eur J Med Chem. 2022;244:114839.

2. Bardia A, Olson JE, Vachon CM, Lazovich D, Vierkant RA, Wang AH, et al. Effect of aspirin and other NSAIDs on postmenopausal breast cancer incidence by hormone receptor status: Results from a prospective cohort study. Breast Cancer Res Treat. 2011;126:149–55.
3. Perontsis S, Geromichalou E, Perdih F, Hatzidimitriou AG, Geromichalos GD, Turel I, et al. Synthesis, structural determination, in vitro and in silico biological evaluation of divalent or trivalent cobalt complexes with indomethacin. J Inorg Biochem. 2020;212:111213.
4. Wang H, Chang Z, Cai GD, Yang P, Chen JH, Yang SS, et al. The novel indomethacin derivative CZ-212-3 exerts antitumor effects on castration-resistant prostate cancer by degrading androgen receptor and its variants. Acta Pharmacol Sin. 2022;43(4):1024–32.
5. Jain HK. QSAR analysis of indomethacin derivatives as selective COX-2 inhibitors. Internet Electron J Mol Des. 2006;5(4):224–36.
6. Obata T, Shiratani S, Nada T, Kasaya Y, Arisawa M, Shuto S, et al. Structure-activity relationship of indomethacin derivatives as IDO1 inhibitors. Anticancer Res. 2021;41(5):2287–96.
7. Sava A, Buron F, Routier S, Panainte A, Bibire N, Constantin SM, et al. Design, synthesis, in silico and in vitro studies for new nitric oxide-releasing indomethacin derivatives with 1,3,4-oxadiazole-2-thiol scaffold. Int J Mol Sci. 2021;22(13):7097.
8. Nabil M, El Raey MA, Abdo W, Abdelfattah MAO, El-Shazly AM, Sobeh M, et al. Gastro-protective effects of Albizia anthelmintica leaf extract on indomethacin-induced gastric ulcer in Wistar rats: In silico and in vivo studies. Antioxidants. 2021;10(2):176.
9. Srivastava AK, Shukla N, Pathak VK. Quantitative structure activity relationship (QSAR) studies on a series of off-target ion channel selective diltiazem sodium derivatives. J Indian Chem Soc. 2010;87:1–7.
10. Yagiela JA, Dowd FJ, Johnson B, Mariotti A, Neidle EA. Pharmacology and therapeutics for dentistry. In: Drugs for treating orofacial pain syndromes. 6th ed. St. Louis: Mosby; p. 384–5.
11. Srivastava AK, Shukla N, Pandey A, Srivastava A. QSAR based modeling on a series of alpha-hydroxy amides as a novel class of bradykinin B1 selective antagonists. J Saudi Chem Soc. 2011;15:215–20.
12. Srivastava AK, Pandey A, Srivastava A, Shukla N. QSAR based modeling of hepatitis C virus NS5B inhibitors. J Saudi Chem Soc. 2011;15:25–8.
13. Srivastava AK, Shukla N. Quantitative structure activity relationship (QSAR) studies on a series of imidazole derivatives as novel ORL1 receptor antagonists. J Saudi Chem Soc. 2013;17(3):321–8.
14. Sharma D, Shukla SD. Neuroprotection: It's phytoremediation (theoretical perspective). Int J Green Herb Chem. 2013;2:1118–31.
15. Shukla N, Srivastava AK. QSAR based modeling on a novel series of 2-chlorobenzamide derivatives as potent P2X7 receptor antagonist. Oxid Commun. 2014;37(1):290–300.
16. Srivastava AK, Shukla N, Pathak VK. Quantitative structure activity relationship (QSAR) studies on a series of carbamate appended N-alkylsulfonamides as inhibitors of peptide amyloid-β (Aβ). Oxid Commun. 2013;36(4):1090–101.
17. Shukla N, Srivastava AK. QSAR studies on a novel cyclohexanamine class of human serotonin transporter (hSERT) inhibitory activity. Oxid Commun. 2013;36(1):133–42.
18. Srivastava AK, Shukla N. QSAR-based modelling on a novel series of pyrimidine-4-carboxamides as antagonists of the human A1 receptor. Oxid Commun. 2012;35(2):423–37.
19. Diudea MV. QSPR/QSAR studies for molecular descriptors. Huntington, NY: Nova Science; 2000.
20. QSAR Methods & Applications. Thane, Mumbai: Gyandhara International Academy Publications; ISBN: 978-93-83006-83-0.
21. Sahu R, Yadav PK, Mishra P, Shukla N. Identification and characterization of a novel therapeutic drug target in Mycobacterium abscessus by in silico method. Int J Green Herb Chem. 2019;8:91–107.
22. Shukla N, Verma S. QSAR modelling of a novel series of methoxylated chalcones as antioxidant agents against Gram-positive bacteria Staphylococcus aureus. Int J Biochem Biomol Res. 2025;3(1).