World Journal of Advanced Research and Reviews
eISSN: 2581-9615 CODEN (USA): WJARAI
Cross Ref DOI: 10.30574/wjarr
Journal homepage: https://wjarr.com/

(RESEARCH ARTICLE)

Check for updates

QSAR Modeling of phthalocyanine Derivatives as Potential Anticancer Agents against Hepatocellular Carcinoma

Francis KOUAME [1], Mawa KONE [1, 5], Lawson Ekossias Digre BEKE [2], Camille Medy NONGBE [3, 5] and Georges Stephane DEMBELE [4, *]

[1] *Laboratory of Constitution and Reaction of Matter, University of Cocody (Now Felix Houphouët Boigny), Côte d'Ivoire.*
[2] *Environment, Climate, Health, Engineering, and Sustainable Development Laboratory (LECI2D), Universite PELEFORO GON COULIBALY DE KORHOGO.*
[3] *Laboratory of Environmental Sciences and Technologies, Jean Lorougnon Guédé University, Daloa, Côte d'Ivoire.*
[4] *Laboratory of Environmental Thermodynamics and Physical Chemistry, UFR SFA, UNIVERSITE NANGUI ABROGOUA 02 BP 801 Abidjan 02, Côte-d'Ivoire.*
[5] *National Laboratory for Quality Testing, Metrology and Analyses, Abidjan, Côte d'Ivoire.*

World Journal of Advanced Research and Reviews, 2025, 28(02), 806-818

Publication history: Received on 27 September 2025; revised on 05 November 2025; accepted on 08 November 2025

Article DOI: https://doi.org/10.30574/wjarr.2025.28.2.3767

Abstract

Liver cancer is one of the major public health challenges worldwide, characterized by high molecular heterogeneity and a poor prognosis. In this study, a QSAR (Quantitative Structure-Activity Relationship) analysis was conducted on a series of phthalocyanine derivatives to evaluate and predict their anticancer activity. The molecular descriptors logP (hydrophobicity) and TS (surface tension) were calculated using density functional theory (DFT/ **B3LYP/LanL2DZ**) with the Gaussian 09 software. The model, constructed using multiple linear regression, demonstrates excellent statistical performance (R^2 = 0.9772; S = 0.0602; F = 21.4383). External validation according to Roy's criteria confirms the robustness and predictive capability of the model. Contribution analysis reveals that logP is the key descriptor for anticancer activity, while TS plays a complementary role. The applicability domain demonstrates the model's reliability for predicting new molecules. These results offer promising prospects for the rational design of phthalocyanine derivatives with anticancer potential.

Keywords: QSAR; Phthalocyanine; Liver Cancer; DFT; Logp; Molecular Modeling.

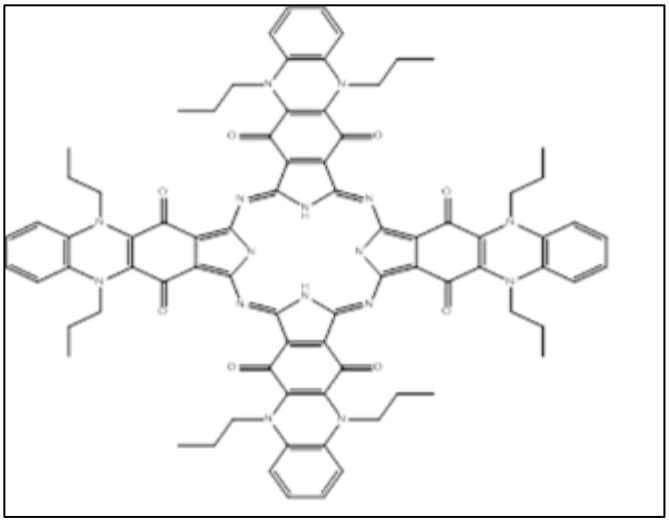
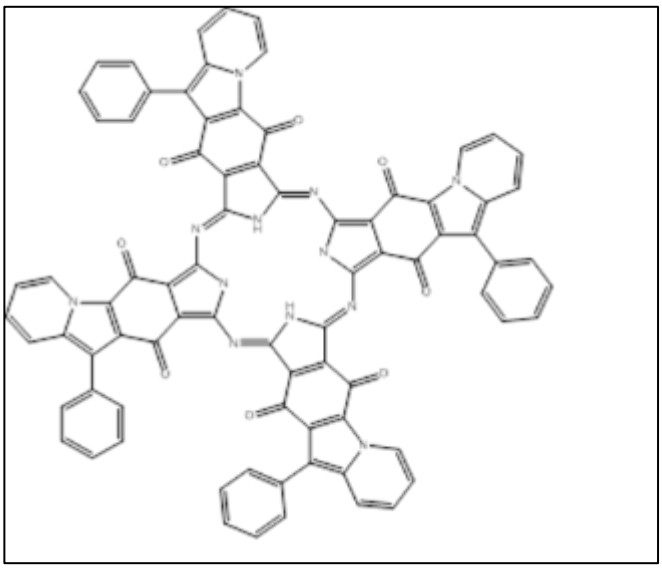
1. Introduction

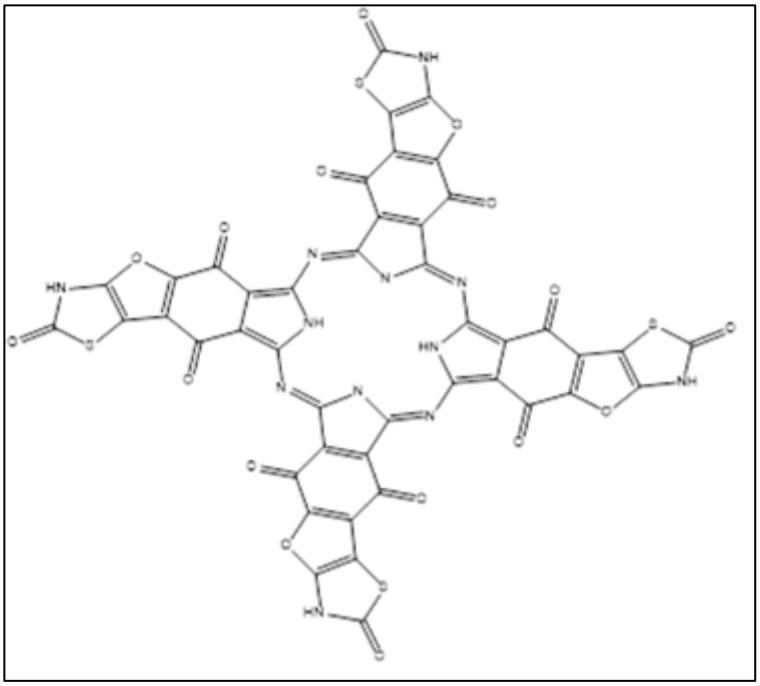
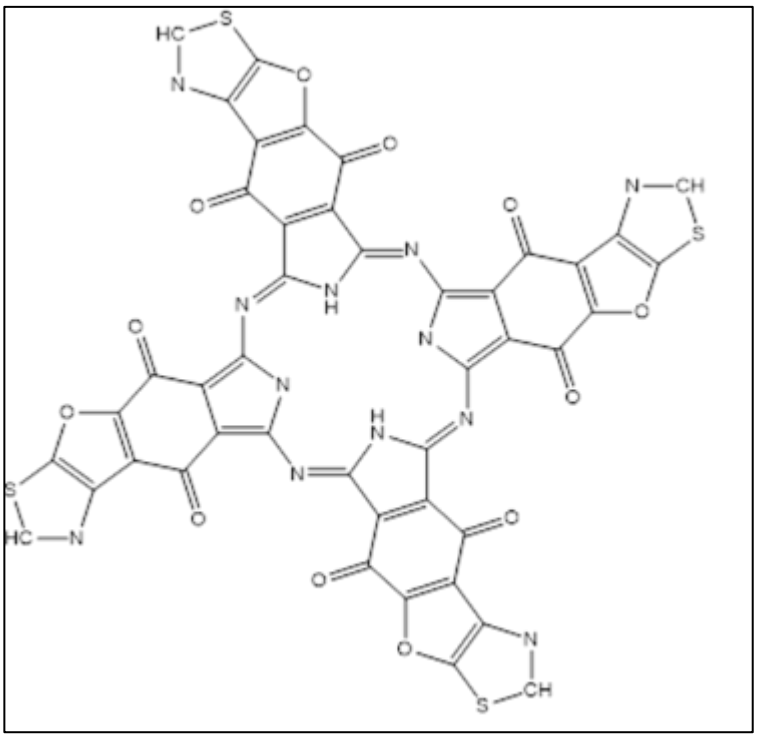
Considered one of the most lethal forms of cancer, liver cancer is a major cause of death worldwide and represents a significant burden on public health systems. Globally, liver cancer is among the most serious cancers, ranking third in cancer-related deaths and sixth in the most diagnosed cancers, with a slightly higher prevalence in men (5th) than in women (7th). Available data indicate that at least 800,000 people worldwide are affected by this type of cancer each year. Its distinctive features include significant molecular heterogeneity, poor prognosis, high metastatic potential, strong tendency for recurrence, and a lack of effective therapeutic strategies. This diversity of causes results in biological and genetic variability among tumors, making the disease complex to understand and treat [6, 9, 10]. Therapeutic options for liver cancer are tailored to each patient and vary depending on the stage of the disease [11]. At an early stage, liver cancer can be cured by complete surgical resection. In intermediate stages, locoregional interventions, including transarterial chemoembolization, ablation, and selective internal radiation therapy, are the main therapeutic options [12]. For advanced stages, available treatment options are primarily systemic therapies such as chemotherapy and immunotherapy [11, 13]. However, these treatments remain limited and minimally effective [7]. Thus, the

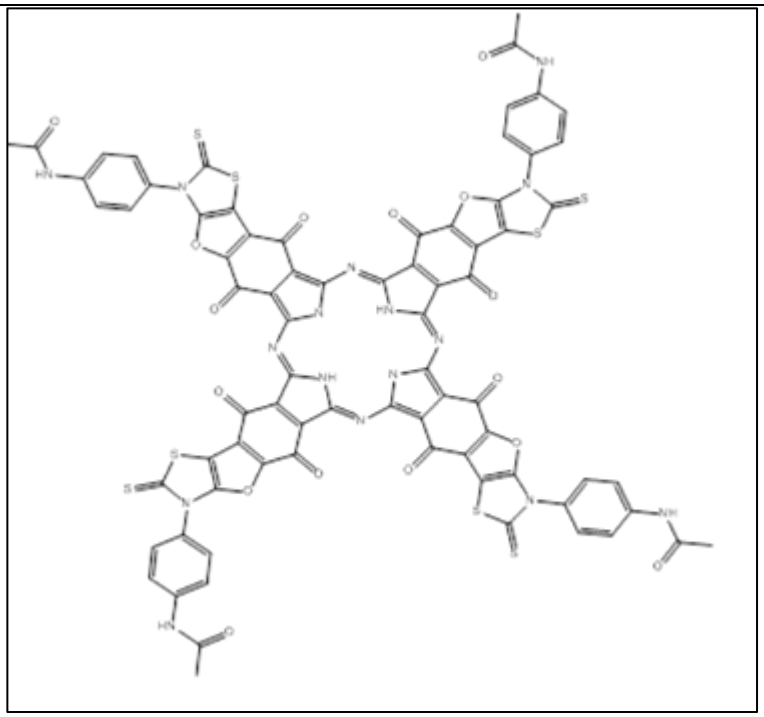
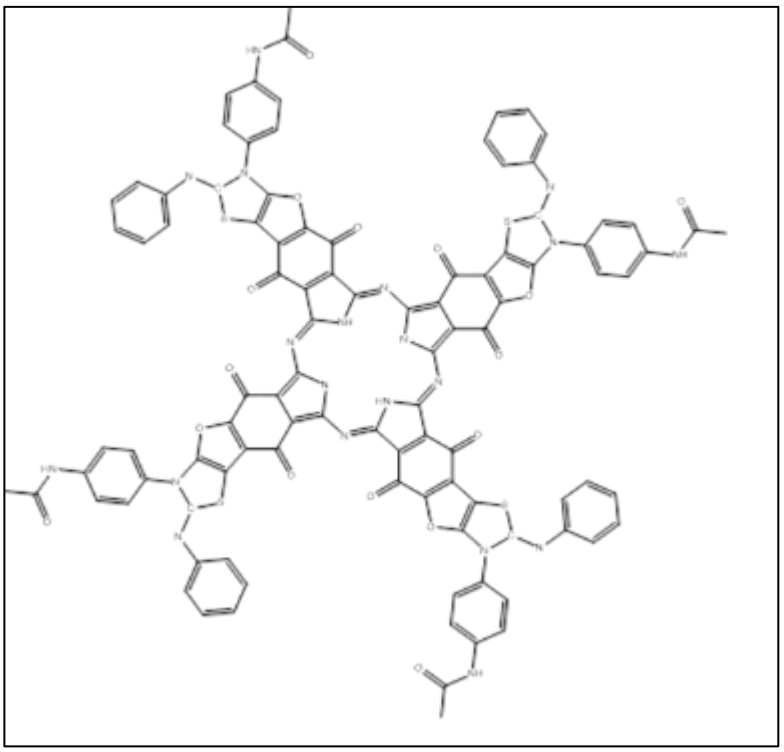
* Corresponding author: Georges Stephane DEMBELE

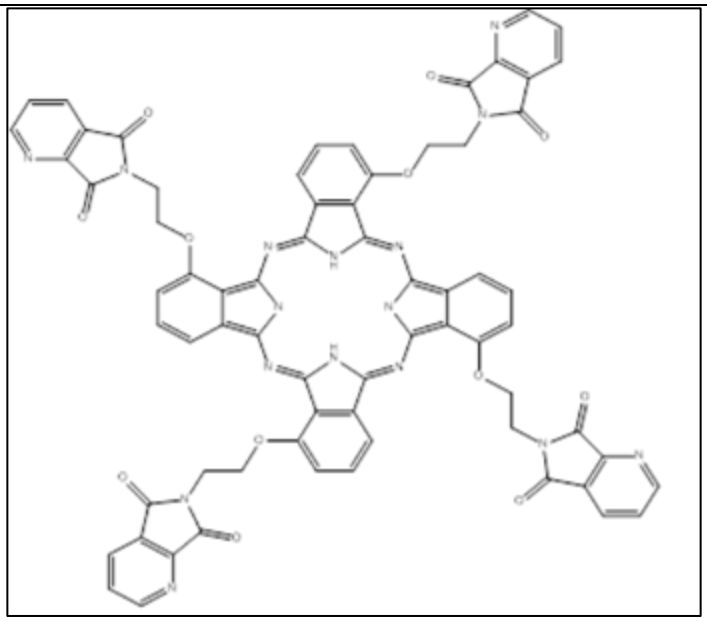
pharmaceutical industry is heavily investing in the discovery of new molecules with innovative mechanisms of action and increased selectivity. The distinct structure of macrocyclic compounds and their ability to interact with numerous biological targets have generated strong interest in this quest [14]. Within this category, phthalocyanines (Pcs) stand out. Their porphyrin-like structure gives them intense absorption bands in the therapeutic range (600-800 nm), notable efficiency in singlet oxygen production, and minimal toxicity in the absence of light irradiation, making them particularly suitable agents for photodynamic therapy (PDT) [15]. Beyond their role in PDT, certain phthalocyanine derivatives have shown, in recent studies, intrinsic anticancer potential [16, 17]. The development of a drug molecule generally requires about fifteen years of research [18]. In this context, new research approaches based on predictive methods for the activities and properties of molecules have emerged, particularly QSAR (Quantitative Structure-Activity Relationship) methods. These predictive methods have significantly reduced biological testing and facilitated the design of new therapeutic compounds [19]. In this study, we apply this method, QSAR (Quantitative Structure-Activity Relationship), with the aim of designing phthalocyanine derivatives with optimized anticancer activity. This work aims to develop robust models capable of interpreting and predicting anticancer activity, measured by the median inhibitory concentration (IC50, expressed in μg/ml), of a series of phthalocyanine derivatives. The structures of the molecules studied are presented in the following table.

Table 1 Structures, codes, and median inhibitory concentrations (IC50) of the studied molecules

codes	Structures	pIC_{50}(HepG2)
mol1		5.1169
mol2		5.0516

mol3		5.0318
mol4		4.9306
mol5		5.2440

		
mol6		5.2920
mol7		4.7058



2. Materials and methods

2.1. Theoretical framework of computational simulation

In order to evaluate and predict the anticancer activity of phthalocyanine derivatives, quantum chemistry calculations were carried out using the Gaussian 09 software [20]. Calculations based on density functional theory are known for their ability to generate a rich and reliable set of molecular properties, which justifies their use in QSAR studies [21, 22]. These approaches not only increase the accuracy of QSAR models but also reduce the time required for calculations and the costs associated with the rational design of new drug candidates [23, 24]. The determination of molecular descriptors was performed based on density functional theory, using the B3LYP/LanL2DZ level of theory. The median inhibitory concentrations of the seven phthalocyanine derivative molecules considered in this study are relatively very low. Among the parameters considered, the median inhibitory concentration (IC50) is of particular importance, as it serves as a quantitative indicator of a compound's potency in inhibiting a specific biological or biochemical target. To optimize the statistical robustness of the models, biological activity data are commonly transformed into the negative decimal logarithm (-log10(C)), which provides more homogeneous and usable values for biologically active compounds [25, 26]. The anticancer efficacy of the compounds is expressed as pIC50, known as the inhibitory concentration potential, a parameter obtained according to the following equation :

$$\mathrm{pIC}_{50} = -\log_{10}(\mathrm{IC}_{50} * 10^{-6})$$

Where IC50 is the inhibitory concentration in µM.

To build the model, three statistical methods were used. Among them, multiple linear regression (MLR) was applied using the tools provided by Excel [27] and XLSTAT [28].

2.2. Descriptors used

For the development of our QSAR model, we identified certain theoretical descriptors. The physico-chemical descriptors identified are : LogP, the logarithm of the Octanol/Water partition coefficient, which measures the hydrophobicity of a molecule, that is, its ability to dissolve in a lipophilic solvent (Octanol) compared to an aqueous solvent, and TS, the Surface Tension, used as a descriptor to predict interactions and physico-chemical properties.

2.3. Estimation of the predictive capacity of a QSAR model

The validity of a model is assessed using several statistical criteria : the coefficient of determination R^2, standard deviation S, cross-validation correlation Q^2_{CV}, and Fisher F test. R^2, S, and F provide information on the fit between the experimental values and those predicted by the model. These parameters indicate the extent to which the model can correctly predict outcomes and are used to estimate the reliability of the values obtained for the test set [29, 30]. The cross-validation coefficient Q^2_{CV} assesses the model's ability to generalize its predictions beyond the data used to construct it. This predictive power is referred to as internal since it is obtained from the same compounds that were

used to build the model. The coefficient of determination R^2 indicates how close the calculated values are to the experimental values. The accuracy of the model increases when the points are close to the regression line [31], which can be assessed using the coefficient of determination. R^2 is expressed as follows :

$$\boldsymbol{R}^2 = 1 - \frac{\sum\left(y_{i,exp} - \hat{y}_{i,théo}\right)^2}{\sum\left(y_{i,exp} - \bar{y}_{i,exp}\right)^2}$$

Where :

$\bar{y}_{i,exp}$: Average value of the experimental anticancer activity;
$y_{i,exp}$: Experimental value of the anticancer activity;
$\hat{y}_{i,théo}$: Theoretical value of the anticancer activity.
The closer R^2 is to 1, the more accurately the model reproduces the experimental values.

The Fisher F test is used to determine the statistical validity of the model and the adequacy of the descriptors used. F is defined by the following expression :

$$\mathbf{F} = \frac{\sum\left(y_{i,théo} - y_{i,exp}\right)^2}{\sum\left(y_{i,exp} - y_{i,théo}\right)^2} * \frac{n-k-1}{k}$$

Where :
n : the total number of data points or observations used to build the model
k : the number of independent descriptors, explanatory variables used in the model.

The cross-validation coefficient is used to measure the reliability of predictions on the training set and is determined using the following relationship :

$$\boldsymbol{Q}^2_{\boldsymbol{cv}} = \frac{\sum\left(y_{i,théo} - \bar{y}_{i,exp}\right)^2 - \sum\left(y_{i,théo} - y_{i,exp}\right)^2}{\sum\left(y_{i,théo} - \bar{y}_{i,exp}\right)^2}$$

2.4. Acceptance settings of a model

According to the criteria proposed by Eriksson et al. [32], a model is satisfactory when $\mathbf{Q^2_{CV}} > \mathbf{0.5}$ and excellent when $\mathbf{Q^2_{CV}} < \mathbf{0.9}$. For a given test set, the condition $R^2 - Q^2_{cv} < 0.3$ must also be met to assess its performance.

Finally, Roy and Roy [33] enhanced the tools for evaluating QSAR models by introducing the index $\boldsymbol{r^2_m}$ and $\Delta\boldsymbol{r^2_m}$, referred to as metric values. The $\boldsymbol{r^2_m}$ index indicates the agreement between observed and predicted activities. These metrics are calculated from experimental and theoretical data and can be determined for both the training set and the test set. According to them, the acceptance of a model is based on meeting these two criteria :

$\overline{r^2_m} = \frac{(r^2_m + r'^2_m)}{2} > 0.5$ et $\Delta r^2_m = |r^2_m - r'^2_m| < 0.2$

Où : $r^2_m = r^2 * \left(1 - \sqrt{(r^2 - r^2_0)}\right)$ et $r'^2_m = r^2 * \left(1 - \sqrt{(r^2 - r'^2_0)}\right)$

2.5. Statistical Analysis: Multiple Linear Regressions (MLR)

Multiple linear regression is a statistical approach aimed at establishing an equation linking a property to a set of descriptors. In the context of QSAR, it facilitates the development of mathematical models capable of predicting the biological activity of compounds that have not been experimentally tested. The objective of this method is to minimize the difference between experimental data and calculated values. It represents the benchmark approach for handling multidimensional data. In this work, it was carried out using the pre-programmed tools of XLSTAT :

$$y = a + (bx_1 + cx_2 + dx_3 + ex_4) + (fx_{12} + gx_{22} + hx_{32} + ix_{42})$$

Where : x_1, x_2, x_3, x_4,... represent the variables and a, b, c, d,... represent the parameters

2.6. Model relevance domain

The applicability domain of a QSAR model is the set of physicochemical, structural, and biological limits within which the model can reliably predict the properties of new molecules [34]. It represents the perimeter of the chemical space that includes the compounds of the training set and those homologous to them, similar in their structural and physicochemical characteristics [35]. Being built on a limited set of compounds and descriptors chosen from a wide range, the model cannot provide universal and reliable predictions for all molecules. Therefore, identifying the domain of applicability is a fundamental requirement for any QSAR model, in accordance with the recommendations of the Organization for Economic Co-operation and Development (OECD) [36]. The determination of a model's applicability domain can be carried out using various methods [34]. In this study, we use the so-called leverage approach. The principle of this method is to examine the variation of standardized residuals as a function of the distance of the descriptor values from their mean, called leverage [37]. The hii represent the diagonal elements of the matrix H. H is the projection matrix that associates the experimental observations Yexp with their estimates Ypred in the regression space, and is defined by the expression :

$$Y_{préd} = HY_{expé}$$

H is defined by the expression :

$$H = X(X^tX)^{-1}X^t$$

The domain of applicability is defined by a leverage threshold, denoted h*. This threshold is generally calculated according to the relation h* = $\frac{3(p+1)}{n}$, where n represents the number of compounds in the training set and p the number of descriptors included in the model [38, 39]. Standardized residuals are generally considered acceptable when they fall within the range of ±3σ, with σ representing the standard deviation of the experimental measurements of the response variable [40]; this principle is commonly referred to as "the three-sigma rule" [41].

3. Results and discussion

The modeling of anticancer activity based on structure relies on the following descriptors: LogP, the logarithm of the partition coefficient (hydrophobicity), and TS, Surface Tension. For each molecule studied, the descriptors and the values of the minimum inhibitory concentration potential were summarized in the following table.

Table 2 Experimental physicochemical characteristics and pIC50 used for the training and validation datasets**.**

Observation	logP	TS	pIC50(HepG2)
Training Set			
mol2	12.1600	82.6000	5.0516
mol5	12.8000	183.7000	5.2440
mol7	3.4600	109.1000	4.7058
mol1	11.2700	100.8000	5.1169
Test Set			
mol4	5.2900	236.0000	4.9306
mol6	15.6400	87.1000	5.2920
mol3	2.4700	269.4000	5.0318

We observe a dispersion of logP values. reflecting the diversity of hydrophobic behaviors of the molecules studied. Notable differences also appear in the pIC50 values. which can be correlated with the chemical structure of the compounds. For example. the molecule mol6. having the highest logP (***15.6400***). has a pIC50 of (***5.2920***). illustrating both significant inhibitory activity and the key role of hydrophobicity in its biological effect.

3.1. Internal validation

The correlation matrix of the descriptors used in the model is presented in the following table. It reveals the relationships between the descriptors.

Table 3 Matrix illustrating the correlation relationships between different physicochemical descriptors**.**

Variables	**logP**	**TS**
logP	1.0000	0.2481
TS	0.2481	1.0000

With a correlation of 0.2481 between logP and TS. these two descriptors seem to have distinct effects on activity. Integrating additional descriptors could enrich the model and optimize the prediction.

The QSAR (HepG2) model of this study is represented by the following equation.

$$\boldsymbol{pIC50(HepG2)^{prédicted} = 4.40321 + 0.04722 * logP + 0.00133 * TS}$$

The coefficients of the equation show that logP has a positive influence on pIC50. while TS has a negligible effect. The positive effect of lipophilicity indicates that the higher the logP. the greater the pIC50. reflecting enhanced biological activity. With a high logP. molecules are more lipophilic. which promotes their cellular absorption and their interaction with biological targets. This hydrophobicity also affects protein binding and tissue distribution.

For surface tension. a positive value of the coefficient indicates that even a small increase leads to a slight increase in pIC50. suggesting a moderate role in biological activity. Measuring the energy required to increase the surface of a liquid. surface tension influences how biological molecules interact with cell membranes. A lower surface tension can facilitate cellular diffusion and improve the intracellular bioavailability of compounds.

Even though its impact is limited compared to logP. surface tension could influence the ability of molecules to penetrate complex biological environments and interact with membranes. thereby contributing to their anticancer activity. In summary. this equation allows for predicting the activity of new molecules. helping to optimize drug discovery. The following table summarizes the statistical parameters of the developed model.

Table 4 Statistical analysis of the RML model

$\boldsymbol{R^2}$	$\boldsymbol{Q^2_{cv}}$	F	S
0.9772	0.9772	21.4383	0.0602

The coefficient of determination R^2 shows to what extent the model explains the variance in the data. With an R^2 of 0.9772. 97.72% of the variations in pIC50 are explained by logP and TS. The high R^2 value attests to the accuracy of the model and highlights the importance of physicochemical properties. such as hydrophobicity and surface tension. in predicting the biological activity of compounds. The $\boldsymbol{Q^2_{cv}}$ coefficient is used to test the model's ability to predict external data. A $\boldsymbol{Q^2_{cv}}$ equivalent to R^2 indicates that the model is robust and well-generalized. The relationships between descriptors and biological activity are reliable. ensuring that the model can be used for new compounds in drug discovery.

The F test allows measuring the overall significance of the model. Here. an F value of 21.4383 indicates that the model is superior to a model without explanatory variables and that logP and TS play a significant role in predicting biological activity. validating their importance in biological interactions. S represents the dispersion of the predicted pIC50 values around the experimental values. A low value (0.0602. here) indicates that the model is reliable and that the chemical properties of the molecules determine their biological activity in a measurable way. The following figure shows the regression line of the obtained model.

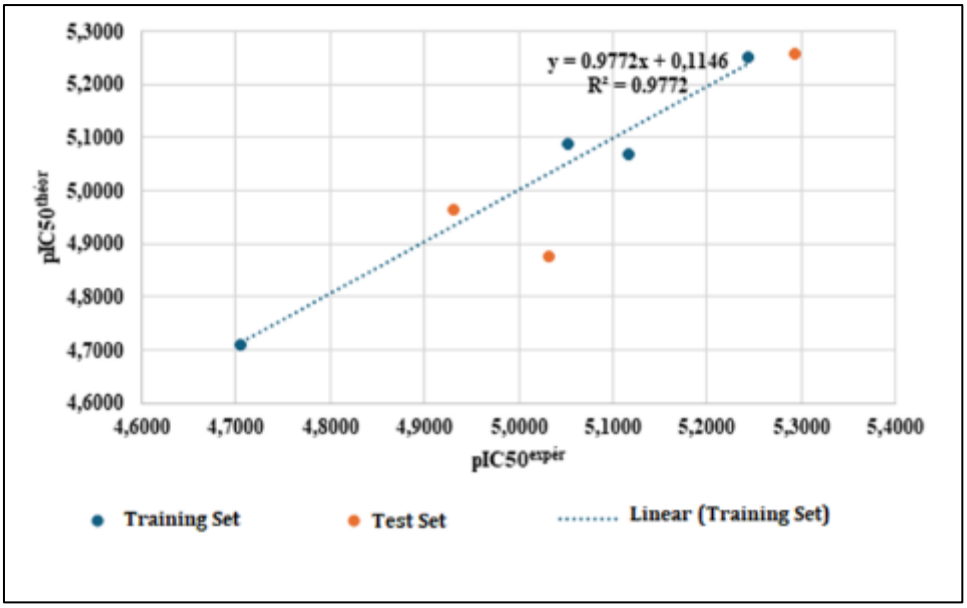


Figure 1 The regression line of the RML model

Figure 1 highlights the excellent fit of the QSAR model. with an R^2 and a slope close to unity. confirming the accuracy of the biological activity predictions. This underscores the role of hydrophobicity and surface tension descriptors in the activity of the compounds and their relevance for drug design. This observation is reinforced by Figure 2. where a near-perfect match between the experimental and predicted curves is observed. reflecting the excellent performance of the model.

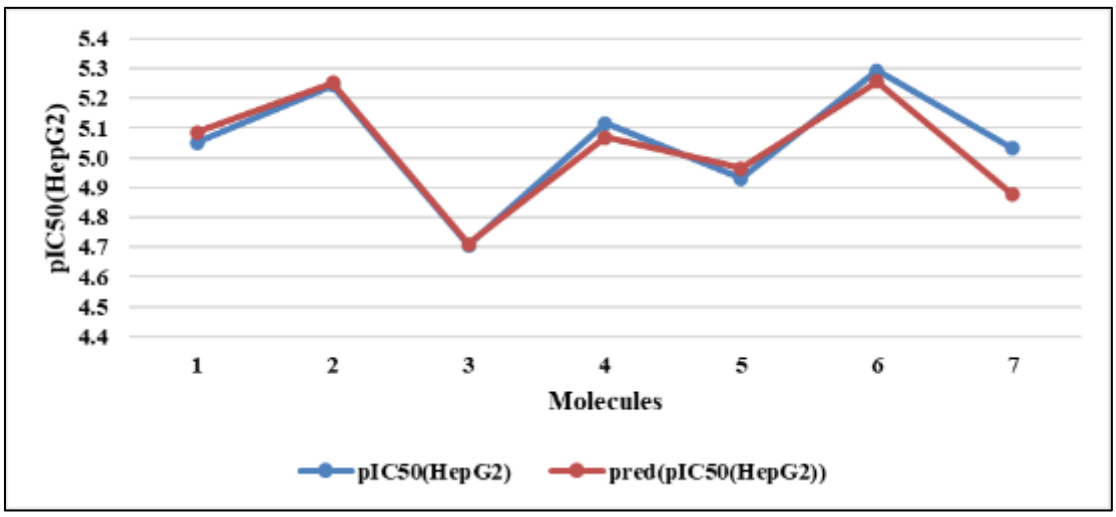


Figure 2 HepG2 Model Similarity Curve

The similarity curve presented shows the relationship between the experimental pIC50 values and the values predicted by the QSAR model for the compounds tested on HepG2 cells. Most of the data points overlap. indicating a good agreement between the predicted and experimental values. In addition to model validation. these results open up prospects for the design of new active molecules. suggesting that optimizing certain descriptors could increase their biological efficacy. However, although the overall performance is very encouraging. the study of individual deviations remains essential for refining the model and improving its accuracy.

3.2. **External validation**

In order to strengthen the reliability of the obtained QSAR model. external validation was carried out through the calculation of the statistical parameters proposed by Roy et al. These indicators allow for the assessment of the model's robustness and predictive power on independent data. The corresponding values are presented in the following table:

Table 5 Roy's Criteria

r_m^2	$r_m^{2\prime}$	$\Delta r_m^2 < 0.2$	$\overline{r_m^2} > 0.5$
0.731405	0.61592	0.115485	0.6736625

Table 4 compiles Roy's criteria. which serve as fundamental indicators for assessing the reliability and robustness of a QSAR model. They provide key information on the model's effectiveness. particularly regarding its ability to generalize to new compounds. Firstly. the parameter Δr_m^2. with a reference value below 0.2. gives an indication of the dispersion of residuals. The calculated value (0.115485). being below this threshold. results in limited dispersion. which attests to a well-fitted model and reliable predictions relative to the experimental data. Next. the parameter $\overline{r_m^2}$. used to evaluate the stability of the model. must exceed the threshold of 0.5. With an obtained value of 0.6736625. the model fully meets this criterion. This suggests that even when tested on various data subsets. it retains a good capacity to explain the observed variability. In practice. this reinforces confidence in the reliability of the logP and TS descriptors for accurately predicting the biological activity of compounds. The high values of $r_m^2(\mathbf{0.731405})$ and $r_m^{2\prime}$ ($\boldsymbol{0.61592}$). which are above the critical threshold of 0.5. confirm that the selected descriptors effectively capture the variance of these biological data. This reflects a good fit between the physicochemical properties and the measured activity. Thus. the analysis of Roy's criteria. presented in Table 4. reveals that the QSAR model is both robust and reliable. capable of generalizing its predictions to new molecules.

3.3. **Evaluation of the contribution of descriptors**

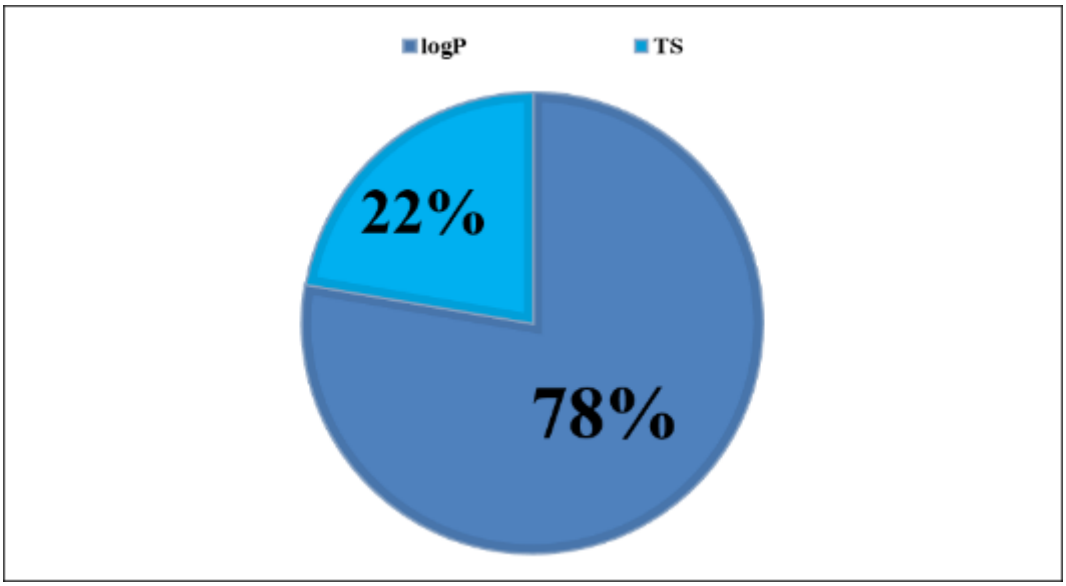


Figure 3 Contribution of descriptors in the RQSA model

The pie chart shown in Figure 3 highlights the proportional contribution of physicochemical descriptors within the established QSAR model. The analysis reveals that logP and TS are the main descriptors involved in predicting the biological activity of our different compounds. The majority contribution of logP (78%) underscores its dominant role in model construction. Although TS contributes (22%) to the model. its share remains significantly lower than that of logP. This indicates that TS plays a complementary role in explaining biological activity. without competing with the central importance of logP. The comparison of the contributions of logP and TS highlights the need to prioritize logP when designing new compounds.

3.4. Scope of the model

The evaluation of the model's applicability domain was carried out using the leverage method. The result obtained is shown in the following figure.

Figure 4 Applicability domain of the QSAR model

Figure 4 illustrates the applicability domain of the QSAR model obtained using the leverage method. The analysis of the plot indicates that all the molecules are below the critical value. The leverage threshold h* for our QSAR model is defined as h* = 2.25. Furthermore. all the residuals fall within the range [-3σ. 3σ]. This result demonstrates the absence of outliers and confirms the robustness of the model. which can be used to predict the biological activity of new molecules within the applicability domain.

4. Conclusion

As part of this work. a QSAR analysis was performed on a series of phthalocyanine derivatives to evaluate their anticancer activity. The model. developed using multiple linear regression. is based on two molecular descriptors: LogP. the logarithm of the partition coefficient (hydrophobicity). and TS. the Surface Tension. The model demonstrates strong statistical performance (R^2 = 0.9772; S = 0.0602; F = 21.4383). reflecting its robustness. External validation using Roy's criteria confirms its predictive power. Furthermore. analysis of the descriptor contribution plot reveals that LogP is the most effective in explaining the anticancer activity of the studied series of phthalocyanine derivatives. Finally. the obtained applicability domain validates the reliability of the model and its potential for predicting new derived structures. These results provide a reliable basis for the rational design of new phthalocyanine derivatives aimed at hepatocellular targets.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] F. Bray. J. Ferlay. I. Soerjomataram. R.L. Siegel. L.A. Torre. A. Jemal. CA: Cancer J. Clin.. 2018. 68. 394–424. [CrossRef]. [Google Scholar]. [Publisher]

[2] Ferlay. J.; Soerjomataram. I.; Dikshit. R.; Eser. S.; Mathers. C.; Rebelo. M.; Parkin. D.M.; Forman. D.; Bray. F. Cancer incidence and mortality worldwide: Sources. methods and major patterns in GLOBOCAN 2012. International Journal of Cancer 2015. 136. E359-E386. https://doi.org/10.1002/ijc.29210.

[3] Ryerson. A.B.; Eheman. C.R.; Altekruse. S.F.; Ward. J.W.; Jemal. A.; Sherman. R.L.; Henley. S.J.; Holtzman. D.; Lake. A.; Noone. A.-M.; Anderson. R.N.; Ma. J.; Ly. K.N.; Cronin. K.A.; Penberthy. L.; Kohler. B.A. Annual Report to the Nation on the Status of Cancer. 1975-2012. featuring the increasing incidence of liver cancer. Cancer 2016. 122. 1312-1337. https://doi.org/10.1002/cncr.29936.

[4] Ferlay. J.; Colombet. M.; Soerjomataram. I.; Mathers. C.; Parkin. D.M.; Piñeros. M.; Znaor. A.; Bray. F. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. International Journal of Cancer 2019. 144. 1941-1953. https://doi.org/10.1002/ijc.31937.

[5] Thylur. R.P.; Roy. S.K.; Shrivastava. A.; LaVeist. T.A.; Shankar. S.; Srivastava. R.K. Assessment of risk factors. and racial and ethnic differences in hepatocellular carcinoma. JGH Open 2020. 4. 351-359. https://doi.org/10.1002/jgh3.12336.

[6] Forner. A.; Reig. M.; Bruix. J. Hepatocellular carcinoma. Lancet 2018. 391. 1301–1314. [CrossRef]

[7] Molina-Sánchez. P.; Lujambio. A. Experimental Models for Preclinical Research in Hepatocellular Carcinoma. In Hepatocellular Carcinoma: Translational Precision Medicine Approaches; Hoshida. Y.. Ed.; Humana Press: Cham. Switzerland. 2019; pp. 333–358.

[8] Hirschfield. H.; Bian. C.B.; Higashi. T.; Nakagawa. S.; Zeleke. T.Z.; Nair. V.D.; Fuchs. B.C.; Hoshida. Y. In vitro modeling of hepatocellular carcinoma molecular subtypes for anti-cancer drug assessment. Exp. Mol. Med. 2018. 50. e419. [CrossRef].

[9] Villanueva. A. Hepato-cellular carcinoma. N. Engl. J. Med. 2019. 380. 1450–1462. [CrossRef].

[10] Connor. F.; Rayner. T.F.; Aitken. S.J.; Feig. C.; Lukk. M.; Santoyo-Lopez. J.; Odom. D.T. Mutational landscape of a chemically induced mouse model of liver cancer. J. Hepatol. 2018. 69. 840–850. [CrossRef].

[11] Ghavimi. S.; Apfel. T.; Azimi. H.; Persaud. A.; Pyrsopoulos. N.T. Management and treatment of hepatocellular carcinoma with immunotherapy: A review of current and future options. J. Clin. Transl. Hepatol. 2020. 8. 168–176. [CrossRef].

[12] Neureiter. D.; Stintzing. S.; Kiesslich. T.; Ocker. M. Hepatocellular carcinoma: Therapeutic advances in signaling. epigenetic and immune targets. World J. Gastroenterol. 2019. 25. 3136–3150. [CrossRef].

[13] Piñero. F.; Silva. M.; Iavarone. M. Sequencing of systemic treatment for hepatocellular carcinoma: Second line competitors. World J. Gastroenterol. 2020. 26. 1888–1900. [CrossRef].

[14] Sessler. J. L.. & Gale. P. A. (Eds.). (2019). *Macrocyclic Chemistry: Current Trends and Future Perspectives*. Royal Society of Chemistry.

[15] O'Connor. A. E.. Gallagher. W. M.. & Byrne. A. T. (2009). Porphyrin and phthalocyanine based photosensitizers for photodynamic therapy. *Photochemistry and Photobiology. 85*(5). 1053-1074.

[16] Zhang. Y.. Chen. J.. Feng. S.. & Li. Y. (2017). Anticancer activity of novel phthalocyanine derivatives in human cancer cells. *Journal of Porphyrins and Phthalocyanines. 21*(03n04). 209-216. (Note: Autre exemple générique à remplacer par une référence réelle sur l'activité anticancéreuse intrinsèque).

[17] Gunes. S.. Gumus. N. A.. & Ozturk. B. (2019). Recent advances in phthalocyanine derivatives as anticancer agents. *European Journal of Medicinal Chemistry. 182*. 111623.

[18] M. R. Yadav New drug discovery: Where are you heating to ?. J. Adv. Pharm. Tech. Res. 2020. 4 (1) 2.

[19] R. Hmamouchi. M. Bouachrine. T. Lakhlifi. Pratique de la Relation Quantitative Structure Activité/Propriété (RQSA / RQSP). Rev. Interdisc.. 2016. 1 (1).

[20] M. J. Frisch. G. W. Trucks. H. B. Schlegel et G. E. Scuseria. «Gaussian 09. Revision A.02.» Gaussian. Inc.. Wallingford CT. 2009.

[21] P. K. Chattaraj. A. Cedillo et R. G. Parr. J. Phys. Chem.. vol. 103. p. 7645. 1991.

[22] F. De Proft. J. M. L. Martin. P. Geerlings et :. :. Chem. Phys Let.. vol. 250. p. 393. 1996.

[23] C. Hansch. P. G. Sammes et J. B. Taylor. «in:Comprehensive Medicinal Chemistry.» Computers and the medicinal chemist. vol. 4. pp. 33-58. 1990.

[24] R. Franke. «Theoretical Drug Design Methods.» Elsevier. 1984.

[25] S. Chaltterjee. A. Hadi et B. Price. «Regression Analysis by Examples.» Wiley VCH: New York. 2000.

[26] H. Phuong. «Synthèse et étude des relations structure/activité quantitatives (QSAR/2D) d'analoguesBenzo[c]phénanthridiniques.» France. 2007.

[27] M. Excel. 2016.

[28] X. V. 2. C. Addinsoft. XLSTAT and Addinsoft are Registered Trademarks of Addinsoft.. 2014. pp. 1995-2014 .

[29] G. W. Snedecor et W. G. Cochran. « Methods. Statistical.» Oxford and IBH: New Delhi. India; . p. 381. 1967.

[30] N. J.-B. Kangah. M. G.-R. Koné. C. G. Kodjo. B. R. N'guessan. A. L. C. Kablan. S. A. Yéo et N. Ziao. «Antibacterial Activity of Schiff Bases Derived from Ortho Diaminocyclohexane. Meta-Phenylenediamine and 1.6-Diaminohexane: Qsar Study with Quantum Descriptors..» International Journal of Pharmaceutical Science Invention. vol. 6. n° %13. pp. 38-43. 2017.

[31] E. X. Esposito. A. J. Hopfinger et J. D. Madura. «Methods for Applying the Quantitative Structure-Activity Relationship Paradigm..» Methods in Molecular Biology. vol. 275. pp. 131-213.. 2004.

[32] L. Eriksson. J. Jaworska. A. Worth. M. D. Cronin. R. M. Mc Dowell et P. Gramatica. «Methods for Reliability and Uncertainty Assessment and for Applicability Evaluations of Classification- and Regression-Based QSARs.» Environmental Health Perspectives. vol. 111. n° %110. pp. 1361-1375. 2003.

[33] Roy P. P. et K. Roy. «on some aspects of variable selection for partial least squares regression models..» QSAR Comb Sci. vol. 27. pp. 302-313. 2008.

[34] N. N.-Jeliazkova et J. Jaworska. « An Approach to Determining Applicability Domains for QSAR Group Contribution Models: An Analysis of SRC KOWWIN.» ATLA 33. p. 461–470. 2005.

[35] F. Sahigara. K. Mansouri. D. Ballabio. A. Mauri et V. C. a. R. Todeschini. «Comparison of Different Approaches to Define the Applicability Domain of QSAR Models.» Molecules. vol. 17. pp. 4791-4810. 2012.

[36] K. Roy et e. al. «A Primer on QSAR/QSPR Modeling Chapter 2 Statistical Methods in QSAR/QSPR..» Springer Briefs in Molecular Science. pp. 37-59. 2015.

[37] J. Jaworska. N. N. Jeliazkova et T. Aldenberg. «QSAR Applicability Domain Estimation by Projection of the Training Set in Descriptor Space: A Review.» ATLA 33. p. 445–459. 2005.

[38] M. Ghamali. S. Chtita. M. Bouachrine et T. Lakhlifi. «Méthodologie générale d'une étude RQSA/RQSP. Revue Interdisciplinaire.» vol. 1. n° %11. 2016.

[39] S. Chtita. M. Ghamali. R. Hmamouchi. B. Elidrissi. M. Bourass. M. Larif. M. Bouachrine et T. Lakhlifi. «Investigation of Antileishmanial Activities of Acridines Derivatives against Promastigotes and Amastigotes Form of Parasites Using QSAR Analysis.» Advances in Physical Chemistry. pp. 1-16. 2016.

[40] T. Asadollahi. S. Dadfarnia. A. Shabani. J. Ghasemi et M. Sarkhosh. «QSAR Models for CXCR2 Receptor Antagonists Based on the Genetic Algorithm for Data Preprocessing Prior to Application of the PLS Linear Regression Method and Design of the New Compounds Using In Silico Virtual Screening.» Molecules. vol. 16. pp. 1928-1955. 2011

[41] S. Chtita. M. Larif. M. Ghamali. M. Bouachrine et T. Lakhlifi. «Quantitative structure–activity relationship studies of dibenzo[a.d]cycloalkenimine derivatives for non-competitive antagonists of N-methyl-D-aspartate based on density functional theory with electronic and topological descriptors.» Journal of Taibah University for Science . vol. 9. pp. 143-154. 2015.