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

In sillico study of a series of molecules derived from thioureas with breast anticancer activities of the MCF-7 strain

Siriki DOUMBIA [1], Niameke Jean Baptiste Kangah [1], Georges Stéphane DEMBELE [1, 2, *], Nanou Tieba TUO [2, 3], Bibata Konaté [1, 2] and Charles Guillaume Kodjo [1]

[1] *Laboratory of Thermodynamics and Physico-Chemistry of the Environment, NANGUI ABROGOUA University, Abidjan, Ivory Coast.*
[2] *Ivorian Disease Modeling Research Group (GIR2M), NANGUI ABROGOUA University, Abidjan, Ivory Coast.*
[3] *Department of Training and Research in Chemical and Agro-Food Engineering, Félix HOUPHOUET BOIGNY National Polytechnic Institute of Yamoussoukro, Ivory Coast.*

World Journal of Advanced Research and Reviews, 2024, 24(03), 1085–1097

Publication history: Received on 29 October 2024; revised on 09 December 2024; accepted on 11 December 2024

Article DOI: https://doi.org/10.30574/wjarr.2024.24.3.3774

Abstract

In the present study, a Quantitative Structure Activity Relationship (QSAR) against breast cancer of the MCF-7 strain was carried out using a series of twenty-five (25) molecules derived from Thio-ureas. The molecular descriptors were obtained after optimization of all these molecules at the B3LYP/6-31+ G (d, p) calculation level. The multiple linear regression (MLR) method was used to carry out this study. The use of this method thus made it possible to obtain a model from the molecular descriptors which are lipophilicity LogP, the bond lengths d(C=N1) and d(N2-Cphen1), the bond angle A(O-C - Cphen2) and the number of fluorine atoms F in the molecule. Furthermore, the stability, exploratory and predictive power of the model obtained was achieved by internal and external validation methods and their area of applicability was verified by the levers approach. All compounds had their leverage value lower than the critical value (h*=1.06), so all belong to the applicability domain. The results from the statistical indicators obtained from the model (R^2 =0.976; RMCE=0.073; F= 89.502), made it possible to affirm that this model is acceptable, robust with good predictive power. Lipophilicity (LogP) turned out to be the priority descriptor in predicting the breast anticancer activity of the MCF-7 strain of our QSAR model.

Keywords: QSAR; MLR; Thiourea Derivatives; Lipophilicity; Area of applicability

1. Introduction

Breast cancer results from an uncontrolled proliferation of cells in the breast. In 2020, there were 2.3 million women worldwide with breast cancer, including 685,000 cases of death and 0.5 to 1% of these breast cancers affecting men [1]. Chemotherapy remains the treatment of choice for patients diagnosed with locally advanced or metastatic cancer. The challenge is to administer an amount of medication that will maximize the effectiveness and minimize the toxicity of the treatment [2]. To face these numerous difficulties, of the discovery of a real drug molecule [3], new areas of research based on predictive methods of the activities and properties of molecules have emerged, in particular QSAR (Quantitative Structure Activity Relationship) methods. QSAR methods are widely used in modern medical chemistry, and make it possible to establish, if possible, a quantitative correlation between biological activity and molecular descriptors using a mathematical model. They make it possible to reduce the number of clinical trials and the margins of error between experimental and theoretical results [4]. Thiourea derivatives are one of the most the antitumor compounds have exhibited potent anticancer properties [5]. The general formula of thioureas and their derivatives is (R1R2N)(R3R4N)C=S and they have a lot of applications in many fields such as organic syntheses and pharmaceutical

* Corresponding author: Georges Stéphane

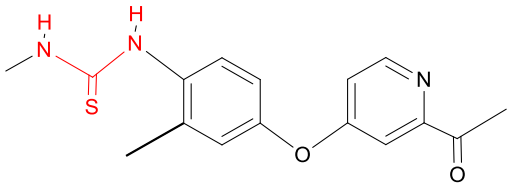
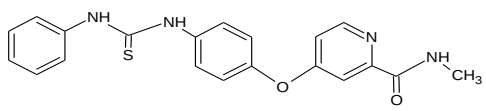
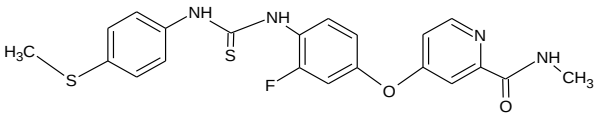
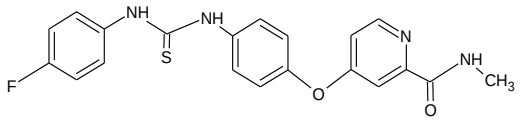
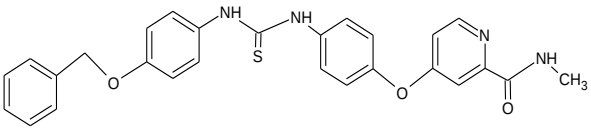
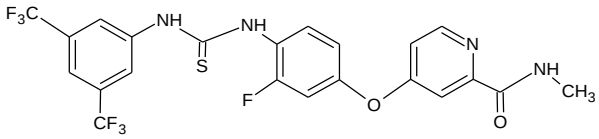
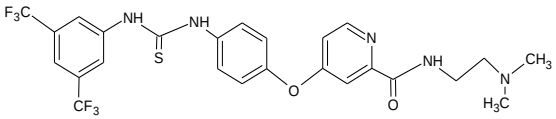
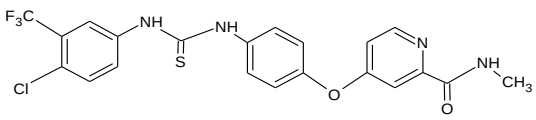
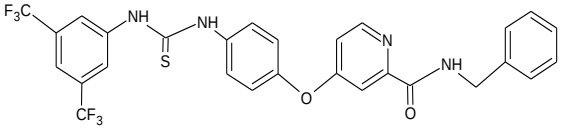
industries [6]. In the present work, a QSAR study was carried out on a series of thiourea derivatives active as potential breast anticancer agents using Density Functional Theory (DFT).

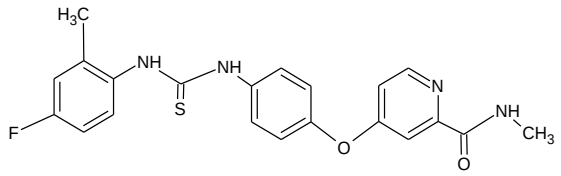
2. Material and methods

2.1. Materials

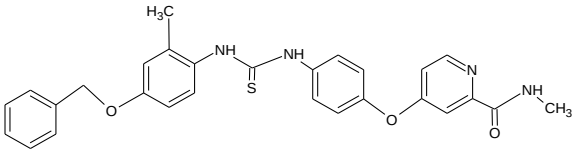
In vitro tests of a series of molecules from the thiourea derivative family on cancer cells including the MCF-7 strain (cell line) were carried out by W. Bai, et al [7]. They evaluated the percentages of inhibition at 10µM of the breast anticancer activities of this series of molecules. For our QSAR study, twenty-five (25) molecules from this series of Thio-urea derivatives were retained and coded from THS1 to THS25. Their designation codes, numbers and PI inhibition percentages are grouped in table 1.

Table 1 2D structures and designation codes of the twenty-five (25) molecules of the proposed thiourea derivatives:

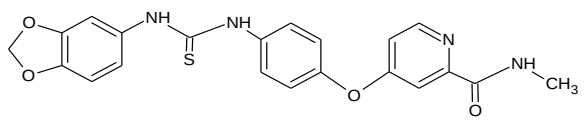
2D structure and designation codes	
	
 THS1	 THS14
 THS2	 THS15
 THS3	 THS16
 THS4	 THS17



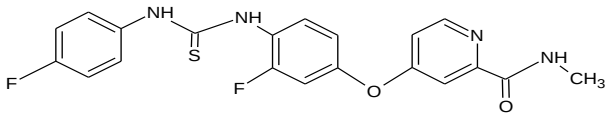
THS5



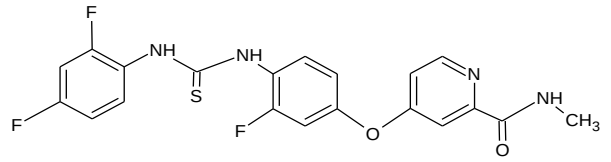
THS6



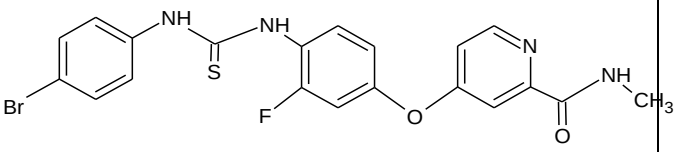
THS7



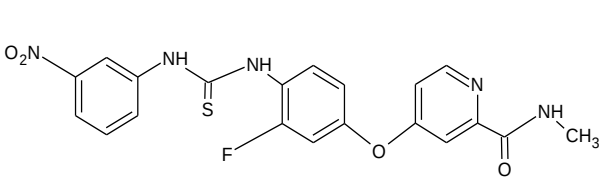
THS8



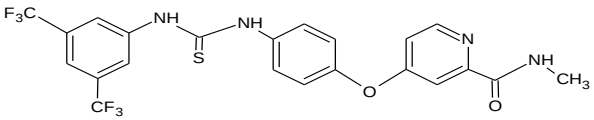
THS9



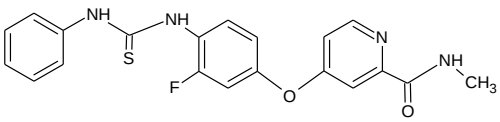
THS10;



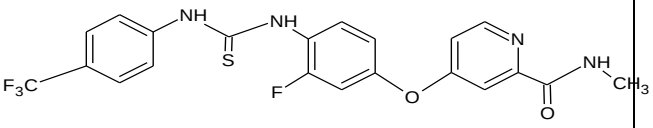
THS11



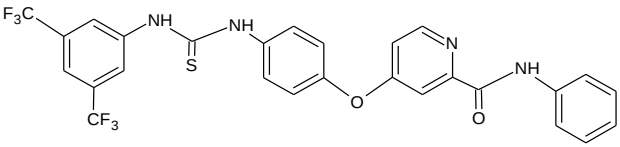
THS18



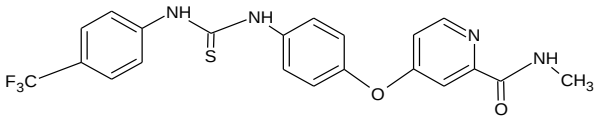
THS19



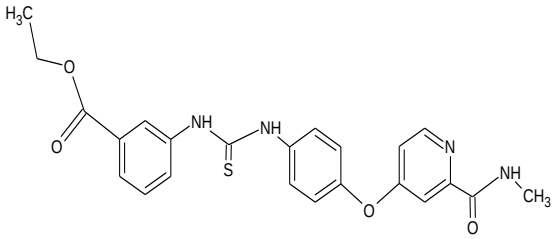
THS20



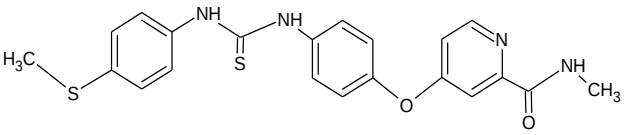
THS21



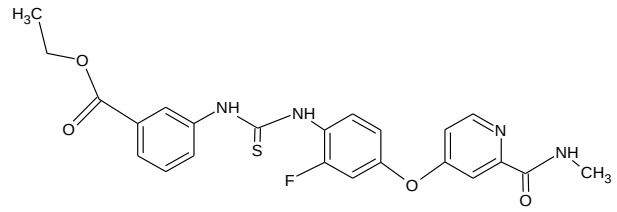
THS22



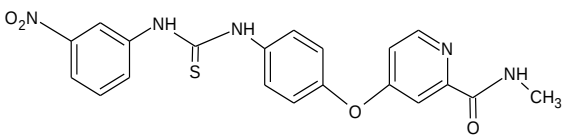
THS23



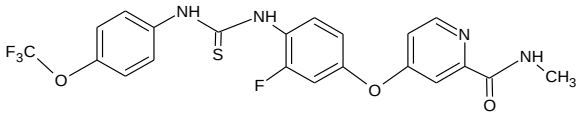
THS24



THS12



THS25



THS13;

2.2. Methods

2.2.1. Level of Computational Theory

In the present study of QSAR prediction of the breast anticancer activity of the MCF-7 strain, thiourea derivatives, the Gaussian 09 software [8] with its GaussView 05 graphical interface, allowed us to carry out the calculations of the quantum chemistry. To do this, all the molecules were optimized and the determination of the molecular descriptors was carried out at the B3LYP/6-31+G (d, p) level of theory. The modeling was then developed using the statistical method of Linear Multiple Regression (MLR) implemented in Excel [9] and XLSTAT [10] spreadsheets.

2.2.2. Anti-breast cancer activity

The biological activity (breast anticancer activity) will be expressed as a decimal logarithm logAA as a function of the inhibition potential PI defined by equation (1) [11]:

$\log AA = \log it - \log C$

with $\log it = \log \frac{PI}{(100-PI)}$

$$\log AA = \log \frac{PI}{(100-PI)} - \log C \qquad (1)$$

PI: Inhibition percentage

C: Molar volume concentration

2.2.3. Molecular Descriptors used

Five (05) theoretical molecular descriptors were determined for the development of our QSAR model. These are the molecular lipophilicity LogP, the bond lengths d(C=N1) and d(N2-Cphen1), the bond angle A(O-C-Cphen2) and the number of fluorine atoms F in the molecule.

- Molecular lipophilicity logP is conveniently evaluated from the value of the decimal logarithm of the partition coefficient P and which is equal to the logarithm of the ratio of the concentrations of the substance s tudied in octan-1-ol and in water:

$P = \frac{[oct]}{[water]}$

From where $\log P = \log \frac{[oct]}{[water]}$ (2)

[oct] and [water] are the molar volume concentrations of the substance in octan-1-ol and in water, respectively. In this work, the Chemsketch software [12] allowed us to determine the logP values.

- Bond lengths and angles are geometric descriptors that can be determined from the relative positions of a molecule in space and require knowledge of the 3D structure of the molecule. The lengths of d(C-N1) and d(N2-Cphen1), the bond angle A (O-C- Cphen2) used in our work were obtained using the Gaussian 09 software [8] with its graphical interface GaussView 05 and are illustrated in Figure 1 below.

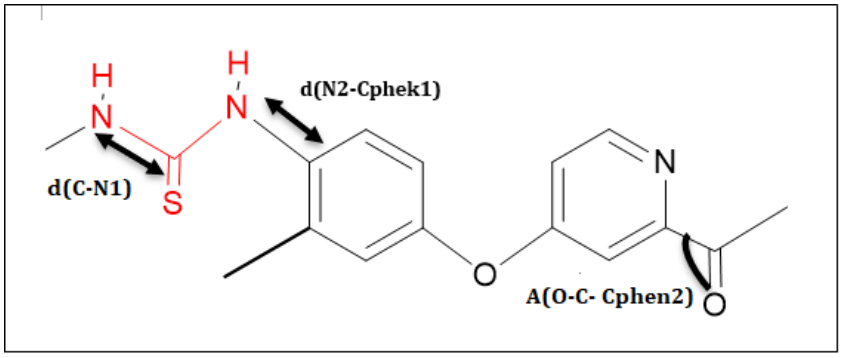


Figure 1 Bond lengths d(C-N1) and d(N2-Cphen1) and bond angle A(O-C- Cphen2)

2.2.4. QSAR model validation methods

Once a QSAR model is developed, its quality is linked to its reliability, its robustness and its predictive nature must be established using statistical analysis criteria. These criteria include internal and external validations.

In the Internal validation are determinated The coefficient of determination R^2, the standard deviation (S) or the Root Mean Square Error (RMCE), the cross-validation correlation coefficients Q_{cv}^2 and Fischer F constitute the statistical analysis criteria of this internal validation and relate to the limits of the model, and make it possible to estimate the precision of the values calculated on the test set [13].

The coefficient of determination R^2 gives an evaluation of the dispersion of the theoretical values around the experimental values. The closer the value of R^2 is to 1, the more the theoretical and experimental values are correlated:

$$\boldsymbol{R^2} = 1 - \frac{\sum(y_{i,exp} - \hat{y}_{i,théo})^2}{\sum(y_{i,exp} - \bar{y}_{i,exp})^2} \qquad (3)$$

With:

$y_{i,exp}$: Experimental value of biological activity

$\hat{y}_{i,théo}$: Theoretical value of biological activity

$\bar{y}_{i,exp}$: Average value of experimental values of biological activity.

- Root Mean Square Errors RMCE is another statistical indicator used. It makes it possible to evaluate the reliability and precision of a model:

$$\mathbf{RMCE} = \sqrt{\frac{\sum(y_{i,exp} - y_{i,théo})^2}{n - k - 1}} \qquad (4)$$

- The Fisher F test is also used to measure the level of statistical significance of the model, that is to say the quality of the choice of descriptors constituting the model.

$$\mathbf{F} = \frac{\sum(y_{i,théo} - y_{i,exp})^2}{\sum(y_{i,exp} - y_{i,théo})^2} * \frac{n - k - 1}{k} \qquad (5)$$

- The Fisher F test is also used to measure the level of statistical significance of the model, that is to say the quality of the choice of descriptors constituting the model.

$$\boldsymbol{Q_{cv}^2} = \frac{\sum\left(y_{i,théo} - \bar{y}_{i,exp}\right)^2 - \sum\left(y_{i,théo} - y_{i,exp}\right)^2}{\sum\left(y_{i,théo} - \bar{y}_{i,exp}\right)^2} \qquad (6)$$

In the external validation we have used According to Eriksson et al [14] and Tropsha et al. [15,16] criterias, According to Eriksson et al [14], the performance of a mathematical model is characterized by a value of Q_{cv}^2 ; $Q_{cv}^2 > 0.5$ for a satisfactory model, and $Q_{cv}^2 > 0.9$ when for the excellent model. According to these authors, given a test set, a model will perform well if the acceptance criterion $R^2 - Q_{cv}^2 < 0.3$ is respected.

According to Tropsha et *al.* [15,16], for the external validation set, the prediction power of a model can be obtained from five criteria. These criteria, known as "external validation criteria" or often called "Tropsha criteria", are as follows:

Criteria 1 : $\boldsymbol{Q_{cv}^2 > 0.6}$ **,**

Criteria 2 : $\boldsymbol{R^2 > 0.7}$ **,**

Criteria 3: $\frac{|R^2 - R_0^2|}{R^2} < 0.1$ **et** $0.85 \leq k \leq 1.15$**,**

Criteria 4 : $\frac{|R^2 - R_0'^2|}{R^2} < 0.1$ **et** $0.85 \leq k' \leq 1.15$

Critère 5 : $\left|R_{Test}^2 - R_0^2\right| \leq 0.3$

With

$\boldsymbol{R^2}$**:** correlation coefficient for the molecules in the test series.

$\boldsymbol{R_0^2}$: correlation coefficient between predicted and experimental values for the test series.

$\boldsymbol{R_0'^2}$: correlation coefficient between experimental and predicted values for the test series.

k: constant of the correlation line (at the origin) (predicted values based on experimental values)

k': constant of the correlation line (at the origin) (experimental values as a function of predicted values).

2.2.5. Area of Applicability (DA)

One of the most used methods for determining the domain of applicability of a model is the so-called "leverage" method, which is based on the variation of the residuals. predictions standardized according to the values of the levers hi. These levers constitute the elements of a matrix H called the hat matrix, which is a projection of the experimental values of the variable explained into the space of values of the variable predicted as follows:

$$Y_{préd} = HY_{expé} \qquad (7)$$

is given by expression (8):

$$H = X(X^tX)^{-1}X^t \qquad (8)$$

Any compound that falls within the domain of applicability of the model has a value of its standardized residue included in the interval $[-3\sigma; +3\sigma]$., where σ is the standard deviation. The critical value of the lever h*, also called threshold lever, is given by:

$$h^* = \frac{3\,(k+1)}{n} \quad (9)$$

Where n is the number of test compounds used and k is the number of model descriptors. For a value du levier of a compond i:

If **hi < h*:** the measured and predicted values of the compound have a probability of agreement as high as that of the compounds in the database.

If **hi > h*:** the compound is considered outside the domain of applicability of the model developed [17].

3. Results and discussion

3.1. Descriptors and breast anticancer activities calculations

Our work focused on the QSAR study of a series of twenty-five (25) molecules derived from active thioureas as potential breast anticancer agents using DFT at the B3LYP/ 6-31+G (d, p) calculation level. Their breast anticancer activities of the MCF-7 strain were expressed by the decimal logarithm log AA MCF-7 calculated from their PI inhibition percentages at 10 µM. All thiourea derivatives have been grouped into two groups. Seventeen (17) were used for the training game and the other eight (08) for the validation game. Five descriptors which are the molecular lipophilicity LogP, the bond lengths d(C=N1) and d(N2-Cphen1), the bond angle A(O-C- Cphen2) and the number of fluorine atoms F in the molecule were used to carry out our modeling of breast anticancer activity. These different physicochemical descriptors PI inhibition potential and the experimental decimal logarithm LogAA MCF-7 of the training and validation sets are given in Table 2.

Table 2 Physico-chemical descriptors, PI inhibition potential and the experimental decimal logarithm log AA MCF-7 of the training and validation sets.

Molecules	Log P	d(C=N1)	d(N2-Cphen1)	A(O-C- Cphen2)	F	PI MCF-7	logAA MCF-7
	Training Set						
THS1	3.420	1.369	1.418	121.322	0	10.1	4.051
THS2	3.580	1.370	1.417	121.084	1	11.7	4.122
THS3	3.740	1.367	1.422	121.000	2	16.3	4.289
THS4	4.900	1.371	1.422	120.937	3	50.7	5.012
THS5	4.070	1.364	1.419	121.006	1	10.4	4.065
THS6	5.520	1.367	1.423	121.655	0	26.8	4.564
THS7	3.200	1.365	1.418	121.020	0	6.2	3.820
THS8	3.740	1.365	1.420	120.867	2	8.1	3.945
THS9	3.900	1.365	1.421	120.877	3	8.5	3.968
THS10	4.410	1.365	1.419	120.869	1	22.1	4.453
THS11	3.470	1.370	1.423	120.826	1	50.4	5.007
THS12	3.740	1.365	1.420	120.901	1	16.2	4.286
THS13	5.110	1.367	1.420	120.861	4	17.7	4.333
THS14	4.020	1.365	1.419	120.877	1	15.7	4.270
THS15	5.190	1.362	1.418	120.911	1	18.2	4.347
THS16	5.270	1.373	1.423	121.045	6	30.4	4.640
THS17	7.000	1.373	1.423	120.975	6	57.1	5.124

	Test Set						
THS18	5.270	1.374	1.423	120.932	6	49.1	4.984
THS19	3.580	1.364	1.419	120.894	1	13.2	4.182
THS20	4.500	1.368	1.421	120.841	4	35.5	4.725
THS21	6.930	1.373	1.423	120.875	6	78.8	5.570
THS22	4.350	1.376	1.418	121.054	4	24.4	4.509
THS23	3.580	1.373	1.417	121.342	0	17.8	4.336
THS24	3.860	1.369	1.417	121.404	0	7.2	3.890
THS25	3.340	1.372	1.421	120.938	0	52.2	5.038

3.2. Interdependence of descriptors

The different descriptors used for the model must be independent of each other. This interdependence of these descriptors is measured using the partial correlation coefficients aij. The values of the partial correlation coefficients aij between the pairs of descriptors are recorded in Table 3.

Table 3 Correlation matrix between the different physicochemical descriptors.

Variables	**Log P**	**d(C=N1)**	**d(N2-Cphen1)**	**A(O-C- Cphen2)**	**F**
Log P	1.000	0.383	0.497	0.158	0.655
d(C=N1)	0.383	1.000	0.518	0.190	0.586
d(N2-Cphen1)	0.497	0.518	1.000	0.113	0.545
A(O-C- Cphen2)	0.158	0.190	0.113	1.000	-0.298
F	0.655	0.586	0.545	-0.298	1.000

The analysis of the data calculation of the partial correlation coefficient between each of the pairs of all the descriptors is less than 0.7 (aij < 0.7), Thus the partial correlation coefficients aij contained in table 3 between pairs of descriptors are all less than 0.7 (aij < 0.7). This demonstrates the independence of the descriptors used to develop our model.

3.3. Multiple Linear Regression (MLR)

The equation of the QSAR model is given by equation (10)

$$logAA\ MCF\ 7\ =\ -95.81000 + 0.35131 * Log\ P + 97.07246 * d(C = N1) + 90.15481 * d(N2 - Cphen1)$$

$$-1.33626 * A(O - C - Cphen2)\ - 0.22044 * F. \qquad (10)$$

The positive signs of the coefficients of lipophilicity LogP, the bond lengths d(C=N1) and d(N2-Cphen1) reflect that the fact that the anticancer activity of the breast is improved for large values of these three (03) different molecular descriptors. On the other hand, the anticancer activity of the breast, according to our model, decreases for large values of the bond angle A(O-C-Cphen2) and the number of fluorine atoms F in the molecule because of the positive sign of their respective coefficients. The statistical indicators of the model are given in Table 4.

Table 4 Statistical analysis report of the anti-breast cancer activity of thiourea derivatives of the RML model

Number of observations N	17
Coefficient of determination R^2	0.976
RMCE standard deviation	0.073

Fischer F test	89.502
Correlation coefficient of Q_{cv}^2 cross validation	0.885
Confidence level α	> 95%

The value of the coefficient of determination (R^2 =0.976) indicates that the estimated values of logAA MCF-7 contain practically 97.6% of the experimental values. The standard deviation (RMCE = 0.073) expresses the small variation of the predicted values compared to the experimental mean. The value of the Fisher test which is (F=89.502) is very high compared to the critical value Fcr =2.81. It shows

that the error committed is less than what the model explains. The cross-validation correlation coefficient is equal to Q_{cv}^2 = 0.8851 and R^2 - Q_{cv}^2=0.976 – 0.885= 0.091 < 0.3. All these statistical indicators clearly show that the model developed explains the anti-cancer activity of the breast in a statistically significant and satisfactory manner.

These different results are confirmed by the regression graph of the MLR model presenting the predicted anticancer activity as a function of the experimental activity shown in Figure 2.

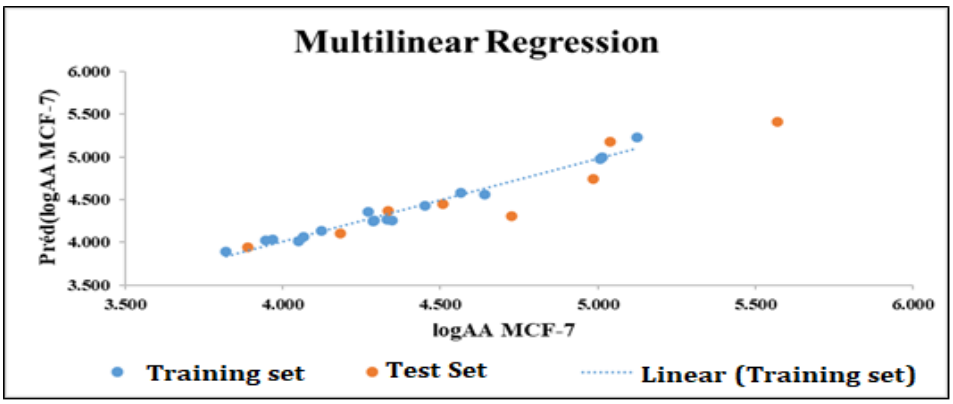


Figure 2 The regression line of the MLR model.

The regression line of the MLR model, after analysis, shows that all the points are around the regression line. This is confirmed by a low value of the difference in the standard deviation RMCE = 0.073 between the experimental values logAA MCF-7 and those predicted Pred(logAA MCF-7), therefore a good similarity at the level of these values. This similarity between these values is illustrated by the model similarity curve in Figure 3.

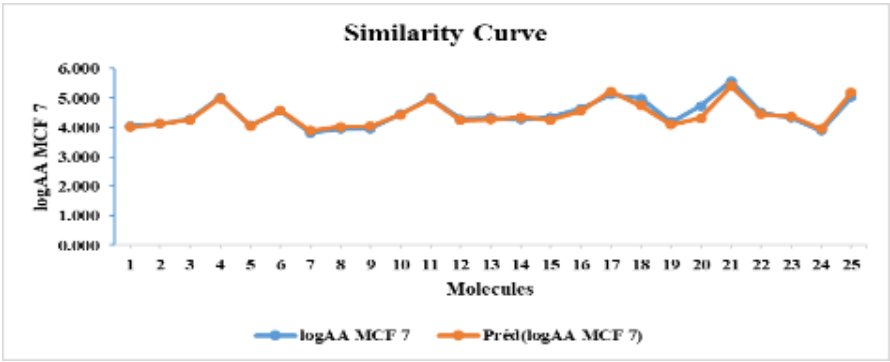


Figure 3 Similarity curve of experimental and predicted values of the MLR model.

Our QSAR model, in view of all these observations, presents good reliability with satisfactory predictive power on all the compounds in the training set. Our model is acceptable and suitable for predicting the breast anticancer activity of the MCF-7 strain from our series of thiourea derivatives.

3.4 External validation

3.4.1 Tropsha criteria

External validation of the model was carried out respectively with the thiourea derivatives of the validation set THS18 to TH25. The Tropsha criteria checks for these molecules in the validation set are presented in Table 5.

Table 5 Verifications of the Tropsha criteria of the external validation set of the RML model.

Statistical parameters	**Tropsha criteria[15,16],**	
R^2	> 0.7	**0.976**
Q_{cv}^2	> 0.6	**0.8570**
R_0^2		**0.8824**
$R_0'^2$		**0.8889**
$\lvert R^2 - R_0^2 \rvert$	≤ 0.3	**0.0073**
$\frac{\lvert R^2 - R_0^2 \rvert}{R^2}$	< 0.1	**0.0082**
k	$0.85 \leq k \leq 1.15$	**0.9795**
$\frac{\lvert R^2 - R_0'^2 \rvert}{R^2}$	< 0.1	**0.0008**
k'	$0.85 \leq k' \leq 1.15$	**1.0196**

The values in Table 5 show that all Tropsha criteria are respected. All these statistical indicators clearly show that the model developed explains the anticancer activity of thiourea derivatives in a statistically significant and satisfactory manner.

3.3.1. Statistical indicators from Roy et al.

Roy et al [18] developed quantities $\mathbf{r_m^2}$ and $\Delta\mathbf{r_m^2}$, called metric values in order to further refine the predictive capacity and verify the robustness of a QSAR model Thus, according to Roy et al, a QSAR model is acceptable if these two criteria are satisfied :

$\overline{r_m^2} = \frac{(r_m^2 + r'^2_m)}{2} > 0.5$

$\Delta r_m^2 = |r_m^2 - r'^2_m| < 0.2$

Avec $r_m^2 = R^2.(1 - \sqrt{(R^2 - R_0^2)}$ et $r'^2_m = r^2.(1 - \sqrt{(R^2 - R_0^2)}$

The indicators r_m^2, $r_m'^2$, Δr_m^2, and $\overline{r_m^2}$ were calculated and recorded in Table 6.

Table 6 Verifications of the Roy criteria of the external validation set of the MLR model

Indicator	**Value**
r_m^2	0.813
$r_m'^2$	0.864
$\Delta r_m^2 = \lvert r_m^2 - r'^2_m \rvert$	0.051
$\overline{r_m^2} = \frac{(r_m^2 + r'^2_m)}{2}$	0.839

Analysis of this table shows that the different criteria were respected, r_m^2 > 0.5 and Δr_m^2< 0.2. We can therefore affirm that our QSAR model is robust and has good predictive power.

3.5. Contribution of model descriptors

It is important to determine the contribution of our five theoretical descriptors in the anticancer activity. Determining the contribution makes it possible to define an order of priority of the descriptors, thus facilitating arbitration at the level of the possible choice of parameters likely to be modified to achieve optimal activity. The values obtained are represented through the diagram in Figure 4.

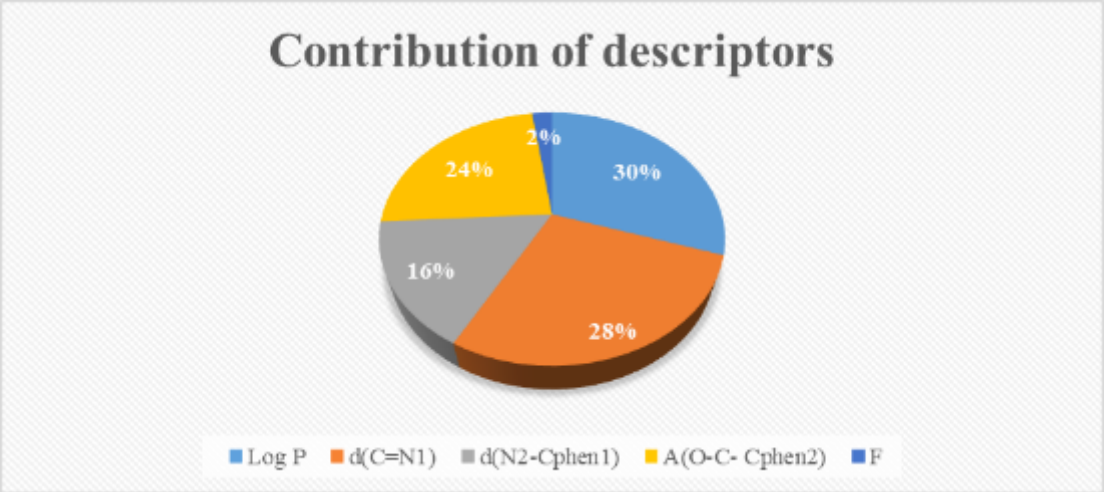


Figure 4 Contribution of descriptors to the breast anticancer activity model.

According to this graph, the importance of the weight of the descriptors involved in our model decreases in absolute value in the following order logP > d(C=N1)> A(O-C- Cphen2) > d(N2-Cphen1) >F. The lipophilicity, which has the greatest contribution compared to the others, is the priority descriptor in predicting the breast anticancer activity of our model.

3.6. Model applicability domain

The MINITAB software was used to calculate the levers of the molecules in the training set as well as those of the molecules in the validation set. This method involves analyzing through a diagram called a Williams diagram. Figure 6 presents the graph of the standardized residuals as a function of the levers hii of the compounds of the two games (learning and validation).

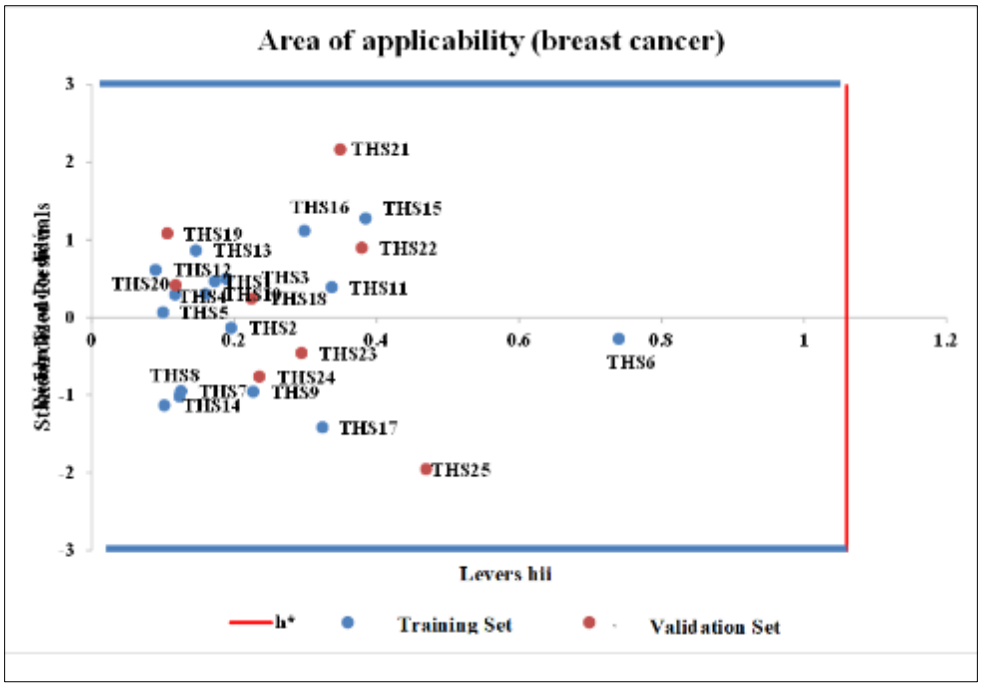


Figure 5 Williams diagram of standardized residuals according to the levers of the compounds used

An analysis of this Williams diagram shows that all the samples of the learning games all have levers lower than the threshold value h*=1.06. This result therefore reflects that all the molecules belong to the domain of applicability. The extreme values of the standardized residuals are between ±3σ and delimit the domain of applicability of the model. In this series, there is no aberrant or influential compound. Thus, all compounds that fall within the scope of the model all have the value of their standardized residue included in the interval of standardized residue limits −3 and +3 of their standard deviation.

4. Conclusion

At the end of this QSAR study allowed us to establish a relationship between the anticancer activity of the MCF-7 strain and their physicochemical properties characterized by molecular descriptors. Five physicochemical descriptors, namely LogP lipophilicity, bond lengths d(C=N1) and d(N2-Cphen1), bond angle A(O-C-Cphen2) and number of fluorine atoms in the molecule in the different compounds allowed us to explain the breast anticancer activity of the MCF-7 strain. Multiple linear regression (MLR) was used as a method to establish our mathematical model and provided the values of the statistical indicators of the model (R^2 = 0.976; RMCE = 0.073; F = 89.502). This MLR statistical model which verifies the different validation criteria has proven that it is acceptable, robust and has good predictive power. Lipophilicity was found to be the priority descriptor in the prediction of breast anticancer activity. Analysis of the applicability domain of this model shows that a prediction of the anticancer activity of new thiourea derivatives is acceptable when its leverage value is less than 1.06, otherwise the breast anticancer activity of this compound is not could be reliably predicted.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Breast Cancer; 12 July 2023 .www.who.int

[2] G. S. Dembele . N. T. Tuo F. Konate , D Soro. Konate B. and N. Ziao ,Quantitative Structure Activity Relationship(QSAR) Study of a series of Molecules derivated from Thiazoline and Thiazine Multithioether Having Activity against Antitumor Activity(A-549), Int. J. of Chem. and Life Sc. , 2022, 11(8), 2426-2435.

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

[4] K. Mkhayar , O. Daoui ,S. Elkhattabi , S. Chtita, R. Elkhattabi ,QSAR Study and Molecular Docking Assisted Design of Novel Cyclohexane-1,3-Dione Derivativesas Anticancer Agents for Non-Small Cell Lung Cancer ,2023, 13(6) , 524.

[5] G. Canudo-Barreras , L. Ortego , A. Izaga , I. Marzo , R. P. Herrera and M. C. Gimeno, Synthesis of New Thiourea-Metal Complexes with Promising Anticancer Properties , Molecules , 2021 , 26(22), 6891.

[6] S. Naz , M. Zahoor ,M. N. Umar ,S. Alghamdi , Muhammaad Umar, K , Sahibazada , W. Ulbari , Synthesis ,characterization ,and pharmacological evaluation of thiourea derivatives , Open Chem. , 2020 , 18 :764-777 .

[7] W. Bai, J. Ji , Q. Huang , Synthesis and evaluation of new thiourea derivatives as antitumor and antiangiogenic agents, Tetrahed . Lett. , 2020, 61, 15236 .

[8] Gaussian 09, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Gaussian, Inc., Wallingford CT, 2009 .

[9] Microsoft Excel, «(15.0.4420.1017) MSO (15.0.4420.1017) 64 Bits,» Microsoft Office Professionnel, 2016.

[10] XLSTAT, « XLSTAT and Addinsoft are Registered Trademarks of Addinsoft,» Copyright Addinsoft , 2014 .

[11] A. Nayyar, A. Malde, R. Jain, and E. Coutinho, 3D-QSAR study of ring-substituted quinoline class of anti-tuberculosis agents, Bio. & Med. Chem., 2006, 14 ,847-856.

[12] acdlabs, Advanced Chemistry Development/ Chemskecht, 1994-2010.

[13] N. J.-B. Kangah, M. G.-R. Kone, C. G. Kodjo, B. R. N'guessan, S. A. Kablan, S. Yeo et N. Ziao, «Antibacterial Activity of Schiff Bases Derived from Ortho Diaminocyclohexane, Meta-Phenylenediamine and 1,6-Diaminohexane: Qsar Study with Quantum Descriptors,» Int. J. of Pharm. Sci. Inv., 2017 . 6(113) , 38-43 .

[14] 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 .» , Environ. Health Persp. , 2003 , 111, (110), 1361-1375 .

[15] A. Golbraikh et A. Tropsha, «Beware of qsar,,» J. Mol. Graph. Model, 2002 ,vol. 20, pp. 269-276 .

[16] A. Tropsha, P. Gramatica et V. K. Gombar, «The importance of being earnest, validation is the absolute essential for successful application and interpretation of QSPR models,,» QSAR, Comb. Sci., 2003 , vol. 22, pp. 69-77 .

[17] F. En-nahli , Quantitative Structure Activity relation Study of ursolic acide and its derivates, End of cycle dissertation , Faculty of Science and Technology Sidi Mohamed Ben Abdellah University Fes,,Maroc , 2019, p 4 .

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