QSAR Study and Molecular Docking Assisted Design of Novel Cyclohexane- 1,3-Dione Derivatives as Anticancer Agents for Non-Small Cell Lung Cancer

Khaoula Mkhayar [1,*] [ID], **Ossama Daoui** [1] [ID], **Souad Elkhattabi** [1,*], **Samir Chtita** [2] [ID], **Rachida Elkhalabi** [3]

[1] Laboratory of Engineering, Systems and Applications, National School of Applied Sciences, Sidi Mohamed Ben Abdellah-Fez University, Fez, Morocco; khaoula.mkhayar@usmba.ac.ma (K.M.); ossama.daoui@usmba.ac.ma (O.D.); souad.elkhattabi@usmba.ac.ma ((S.E.);

[2] Laboratory of Analytical and Molecular Chemistry, Faculty of Sciences Ben M'Sik, Hassan II University of Casablanca, B.P 7955, Casablanca, Morocco; samirchtita@gmail.com (S.C.);

[3] Laboratory of Applied Organic Chemistry, Faculty of Sciences and Technologies, Sidi Mohamed Ben Abdellah-Fez University, Fez, Morocco; rachida.elkhalabi@usmba.ac.ma (R.E.);

* Correspondence: Khaoula.mkhayar@usmba.ac.ma (K.M.); Souad.elkhattabi@usmba.ac.ma (S.E.);

Abstract: One of the deadly diseases affecting the global public is cancer. Nearly 85% of lung cancer cases are non-small cell lung cancer (NSCLC). As part of this study, we examined the quantitative relationship between the biological activity against NSCLC and the molecular structure of a series of 38 cyclohexane-1,3-dione-dimedone derivatives. For this purpose, molecular descriptors were calculated by DFT-B3LYP/6-31G and topological and physicochemical analysis. This work shows the results of the QSAR models' evaluations. Using MLR and MNLR techniques, these models show high predictive power for the linear model ($R^2 = 0.913$; $R^2_{CV} = 0.85$, $R^2_{test} = 0.934$), as well as for the non-linear model ($R^2 = 0.991$; $R^2_{CV} = 0.82$; $R^2_{test} = 0.997$). Using the QSAR model predictions, new molecular structures are designed, and their possible interaction mode with the c-Met receptor is analyzed by molecular docking. QSAR model predictions, molecular docking, and ADMET in silico property assessment suggested that only one of the 3 newly designed molecules can be recommended as anti-lung cancer (NSCLC) drug after further in vivo and in vitro investigations before clinical trials.

Keywords: QSAR; ADMET; molecular docking; NSCLC.

1. Introduction

According to World Health Organization reports, cancer remains one of the leading causes of death worldwide [1]. Indeed, the pharmaceutical industry is in continuous research to develop effective anticancer drugs [1]. It is moving towards new techniques capable of predicting the activity of molecules before their synthesis, such as molecular modeling, including the quantitative activity structure relationship (QSAR) and molecular docking. This technique (QSAR) reduces the number of clinical trials and reduces the margin of error between experimental and theoretical results. The QSAR technique is more used in modern medicinal chemistry, which makes it possible to quantitatively correlate biological activity and molecular descriptors (1D, 2D, 3D) by a mathematical model.

To develop new c-Met inhibiting compounds, we carried out a molecular modeling study of pyrazole, isoxazole, and thiophene molecules derived from dimedone synthesized by

Mohareb *et al*., which showed high biological activity as well as an important ability to inhibit the c-Met protein. Protein kinases and enzymes participate in many signal transduction pathways. They regulate all aspects of cellular function. There are more than 500 coding genes [2]. The tyrosine kinase of the c-Met receptor represents an interesting anticancer target [3]. In vivo and vitro tests [4-5], Foretinib (GSK1363089) (Figure 1), has shown that they have an essential role in cell signaling, importance in inhibiting Met multikinase, inhibiting the growth of cancer cell lines by blocking receptor tyrosine kinases.
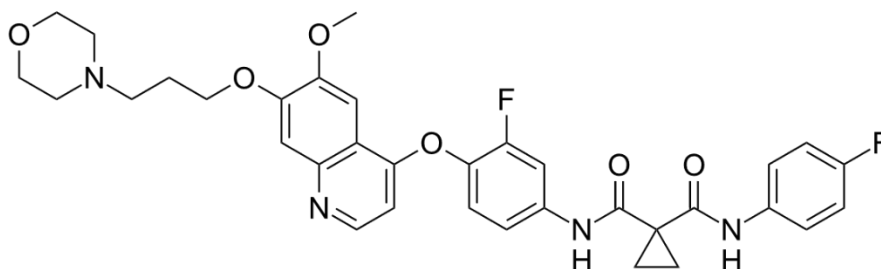

Figure 1. The structure of orally bioavailable Foretinib.

In this work, we developed the models (QSAR) based on multiple linear regression (MLR), multiple non-linear regression (MNLR), and artificial neural network (ANN) techniques for a series of 38 dimedon-derived molecules that have biological activity against NSCLC non–small cell lung cancer [6]. The electronic properties of the molecules studied are described using density function theory (DFT). Thanks to the interesting results of the QSAR models developed that are validated by the domain of applicability and the randomization test, we have designed new molecules based on the structural modification of the reference molecule (the molecule with the highest biological activity against NSCLC). Indeed, we evaluated the properties of ADMET (Adsorption, Distribution, Metabolism, Excretion, and Toxicity) in silico of the molecules of the series and designed a molecular docking study of the compounds designed that showed good pharmaceutical properties.

This work is organized as follows: the second part deals with present methods and materials. The third part presents the results of the simulation and the discussions. Finally, the flowchart (Figure 2) summarizes the main steps followed in this work.
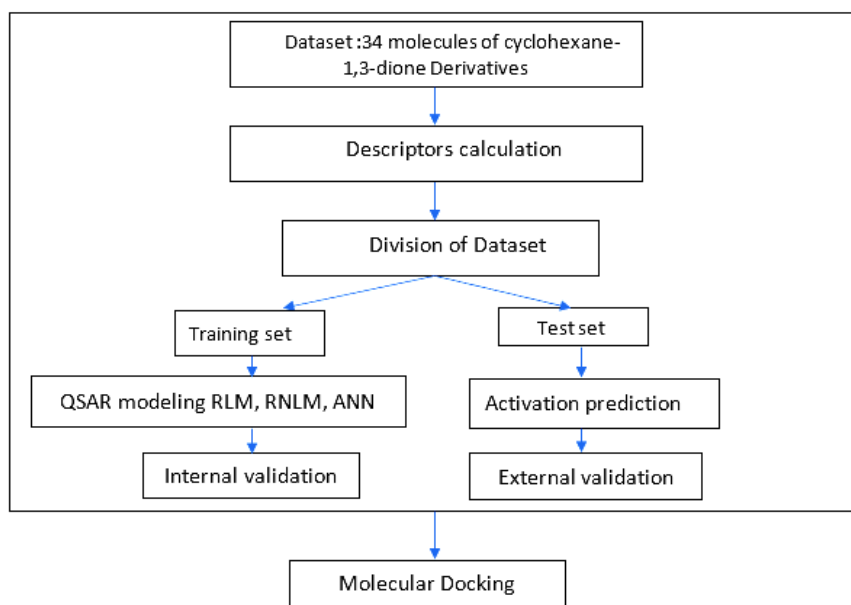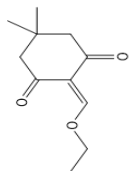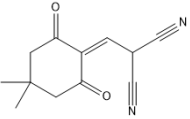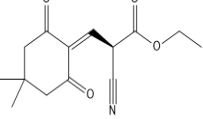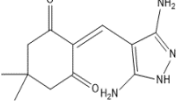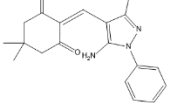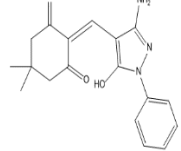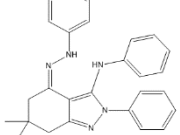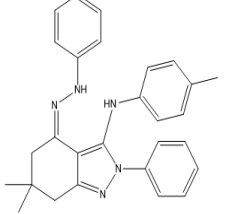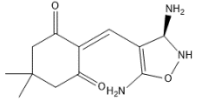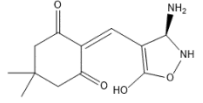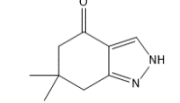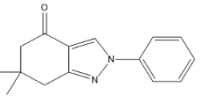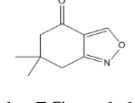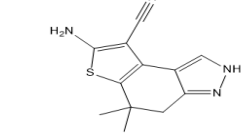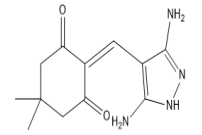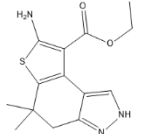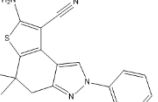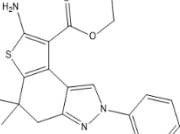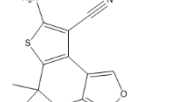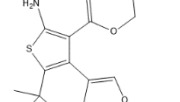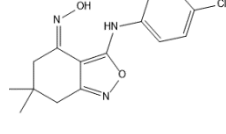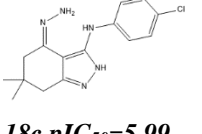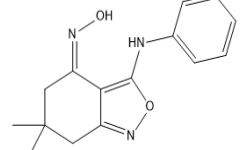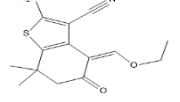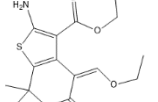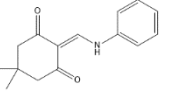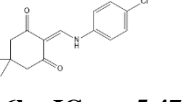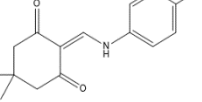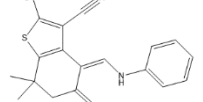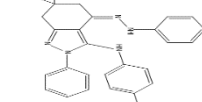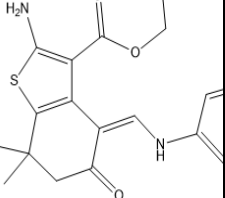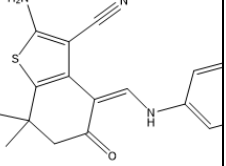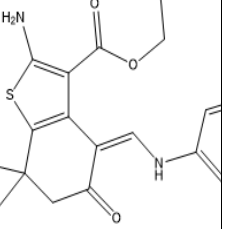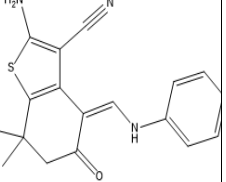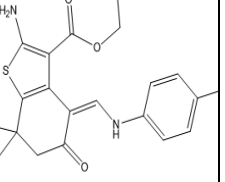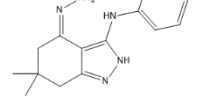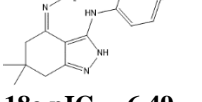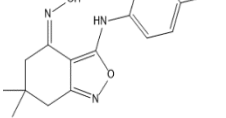

Figure 2. Flowchart of the methodology used in this work.

Table 1. Structure of compounds and their pIC50 activities

3 pIC₅₀=5,16	**5a pIC₅₀=5,37**	**5b pIC₅₀=5,66**	**7b pIC₅₀= 6,42**	**7c pIC₅₀= 5,79**	**7d pIC₅₀=6,60**	**18b pIC₅₀=5,27**	**18 pIC₅₀=5,37**
9a pIC₅₀=5,94	**9b pIC₅₀=6,25**	**10a pIC₅₀=5,47**	**10b pIC₅₀=5,28**	**11 pIC₅₀=6,60**	**12a pIC₅₀=5,84**	**7a pIC₅₀=5,20**	
12b pIC₅₀=5,35	**12c pIC₅₀=6,37**	**12d pIC₅₀=6,20**	**13a pIC₅₀=5,32**	**13b pIC₅₀=5,89**	**19b pIC₅₀=5,02**	**18c pIC₅₀=5,99**	**19a pIC₅₀=5,2**
14a pIC₅₀= 5,79	**14b pIC50= 6,04**	**16a pIC₅₀= 5,35**	**16b pIC₅₀= 5,47**	**16c pIC₅₀= 5,85**	**17a pIC₅₀= 5,07**	**18d pIC₅₀=6,55**	
17b pIC₅₀= 5,02	**17c pIC₅₀= 6,38**	**17d pIC₅₀= 6,49**	**17e pIC₅₀= 5,48**	**17f pIC₅₀= 5,2**	**18a pIC₅₀= 5,34**	**18e pIC₅₀=6,49**	**19b pIC₅₀=5,37**

2. Materials and Methods

In this work, we aim to study a series of 38 molecules of derivatives of cyclohexane-1,3-dione (dimedone) [6]. The observed IC50 (μM) activities are converted to logarithmic pIC_{50} via this equation (pIC_{50}= $-logIC_{50}$). They are presented in Table 1. Following our study, we calculated 28 molecular descriptors for our series, which are of different types (1D, 2D, 3D), topological, quantum, physicochemical, and electronic molecular descriptors calculated by the functional density theory DFT-B3LYP/6-31G.

The software (ChemSketh [7], Chem3D [8], and Gaussian09 [9]) mainly aimed to draw and calculate the topological, quantum, physicochemical, and electronic descriptors of molecules to quantitatively link the biological activity and descriptors based on statistical techniques (MLR, MLNR, ANN) in order to develop a QSAR model. The calculated descriptors and observed activities are presented in the supplementary document (table 1).

2.1. Principal component analysis.

PCA [10] is a very effective quantitative data analysis method used to reduce the size of the data representation space. The initial variables are replaced by new ones, called principal components, two to two uncorrelated, and such that the projections of the data on these components are of maximum variance. This PCA technique is generally used to quickly analyze and visualize the correlations between the variables and the observations initially described by the variables on a two- or three-dimensional graph, constructed in such a way that the dispersion between the data is as well preserved as possible. In our work, we have applied the PCA method for 28 descriptors presented in Table 2.

Table 2. List of descriptors used in this work.

Descriptors	Type	Symbol
The number of H-binding donors	Physicochemical	HBA
The number of binding acceptors of H		HBD
Electronegativity		X
Parachlor		PC (cm^3)
The surface area of the polar surface	Topological	PSA
Molecular weight		MW
The Balaban Index		BI
The molecular topological index		MTI
Molecular refractivity		MR
The number of rotational links		NRB
The coefficient of the shape		HIS
The sum of degree		SD
The sum of the degrees of valence		SDV
Topological diameter		TD
Total connectivity		Tcon
Total valence connectivity		TV
Wiener index		WI
The Octanal/water partition coefficient		LogP
Solubility coefficient	Quantum	LogS
Total energy		And
HOMO energy		E_{HOMO} (ev)
LUMO energy		E_{LUMO} (ev)
The energy gap		E_{gap} (ev)
Chemical Hardness		C (ev)
Chemical Softness		S(ev)
Electrophilicity index		Ш(ev)
The dipole moment		DM

2.2. The K-means method.

The k-means method is a widely used tool for dividing the database into two sets [4], the learning set of 80% and the test set of 20%. In our work, the learning set contains 30 compounds, and the test set contains 8 compounds. Each set has a specific role, the learning set is used to develop the QSAR models, and the test set is damaged to estimate the effectiveness of each model developed.

2.3. The relationship structure-activity.

To study the relationship between structure and biological activity, we need, on the one hand, the statistical methods PCA [11], K-means, MLR [12], MNLR [12], and ANN [13] which are widely known in the domain of molecular modeling, and on the other hand to carry out the domain of applicability and the randomization test for validation. The parameters calculated during the development of QSAR models are the coefficient of determination R^2, the adjusted coefficient R^2_{aj}, the mean square error MSE, the Fisher test F, and the p-value2.4.MLR multiple linear regression.

MLR multiple linear regression is a necessary step in the procedure for studying QSAR due to its robustness [12]. It is based on the search for a linear equation linking the dependent variable y (biological activity) to independent variables X_i (molecular descriptors) and their coefficients ai is a_i constant a_0, it is expressed by equation 1:

$$Y = a_0 + \sum_{i=1}^{n} a_i x_i \qquad \text{Equation 1}$$

2.5. MNLR multiple non-linear regression.

MNLR multiple non-linear regression is a non-linear analysis technique (logarithm, exponential...) [14]. It has the role of predicting a mathematical model capable of explaining the non-linearly of a biological activity Y in relation to molecular descriptors. The descriptors used in the MLNRmethod are the same as the descriptors used in the MLR method, whose non-linear model has high descriptors X_i at power 2 expressed by equation 2. With a_0 the constant and a_i and b_i the coefficients of the descriptors.

$$Y = a_0 + \sum_{i=0}^{n} a_i x_i + b_i x_i^2 \qquad \text{Equation 2}$$

2.6. ANN Artificial Neural Network.

An artificial neural network (ANN) is a powerful technique [15]. It is used to approximate complex systems, including systems that are difficult to model by statistical techniques. The QSAR model obtained by the RNA technique arrived to confirm the accuracy of the molecular descriptors chosen by the MLR technique. The ANN model that we develop in our work is of the feed-forward type [16]. Statistical tests verify the validation of the developed QSAR models. Our study is based on internal validation and external validation of more than one Y-randomization test followed by the MLR domain of applicability. The parameter plays a major role in determining the best artificial neural network architecture [17]. It is defined as follows:

$$\rho = \frac{number\ of\ weights}{number\ of\ connections}$$

It is recommended that the value should be between 1.00 and 3.00; if < 1, the network stores the data, while if > 3, the network is not able to generalize [15].

2.7. Internal validation.

Cross-validation (CV) [18] is the most widely used technique to validate the models obtained bythe different methods (MLR, MNLR, and ANN). R^2_{cv} is the calculated parameter used to determine whether the models obtained are robust. The internal validation is based on the calculation of the R^2_{cv} parameter by equation 3. The value of the parameter must be greater than 0.6.

$$R^2_{cv} = 1 - \frac{\sum(Y_{obs}(train) - Y_{cal}(train))^2}{\sum(Y_{obs}(train) - \bar{Y}_{cal}(train))^2}$$ Equation 3

Y_{obs}(train): the value of the observed response,

Y_{calc}(train): the value of the response predicted by Loo-cv,

$\bar{Y}_{calc}$(train): the mean value of the observed/predicted responses.2.8. External validation.

External validation [19] is a method that is interested in predicting the activities of molecules thatbelong to the test series. This method is characterized by the R^2_{test} parameter. This parameter is known as external validation criteria or Trophsa criteria [20].

2.8.1. Y-Randomization Test.

The y-randomization test [21] is a test used in the validation of QSAR/QSPR models. It is used in the model obtained by the MLR method to reduce the possibility of a random correlation between biological activities and descriptors in the model initially obtained by the MLR technique. The QSAR model is acceptable if the correlation coefficient (R^2) of the original model is greater than the mean random correlation coefficient (R_r^2) of randomly constructed models.

2.8.2. Domain of Applicability (AD).

A model is valid only within its training domain, and new molecules must be considered as belonging to the domain before the model has applied Organization for (Economic Co-operation and Development) (OECD) [22]. AD is used to determine the molecules that are outside the applicability of the QSAR model built. Different methods are used to determine the domain of applicability; among them is the " leverage " method [23]. If a compound has a residual and leverage that exceeds the critical value h*, this compound isconsidered outside the scope of applicability of the model developed [23].

2.8.3. The properties of ADMET.

The evaluation of the properties of ADMET (Absorption, Distribution, Metabolism, Excretion, and Toxicity) [24] is a necessary step in drug discovery. Drug-likeness is considered to be the most useful method for identifying compounds not recommended as drugs, including compliance with the three interesting rules that are rules: Lipinski [25], Veber [26], and Igan [27].

In our work, we evaluate the ADMET properties in silico of the molecules in our series, using the online servers SwissADME [28] and pkCSM [29], respectively.

2.8.4. Molecular docking.

Molecular docking [30] is a key tool in structural molecular biology and computer-assisted drug design. The goal of ligand-protein docking is to predict the predominant binding mode of a ligand with a protein of known three-dimensional structure [31]. Using virtual screening of small molecule databases, it is possible to identify new potential inhibitors against a target of interest.

In our work, the protein c-met (Foretinib) is the receptor (code pdb:6SD9). We used the Autodocktools software to prepare the ligand and the receptor for docking, the 'gridbox' dimensions (104, 110, 108).

3. Resultats and Discussions

To evaluate the correlation between the calculated descriptors and the studied activity (pIC$_{50}$), we used several statistical methods to develop the QSAR models as well as for their validation: principal component analysis (PCA), multiple linear regression (MLR), multiple non-linear regression (MNLR), artificial neural network (ANN), the Y-randomization test, and the application domain of the QSAR model. Among the 38 compounds selected for this study, 30 were chosen as the model building (Training) set and the remaining eight (8) as the external confirmation (test) set.

Table 2 presents the correlation matrix between 28 descriptors obtained from the PCA analyses. Based on the values of the correlation coefficients between the descriptors. We found that there is a strong correlation between several descriptors, as Figure 3 presents:
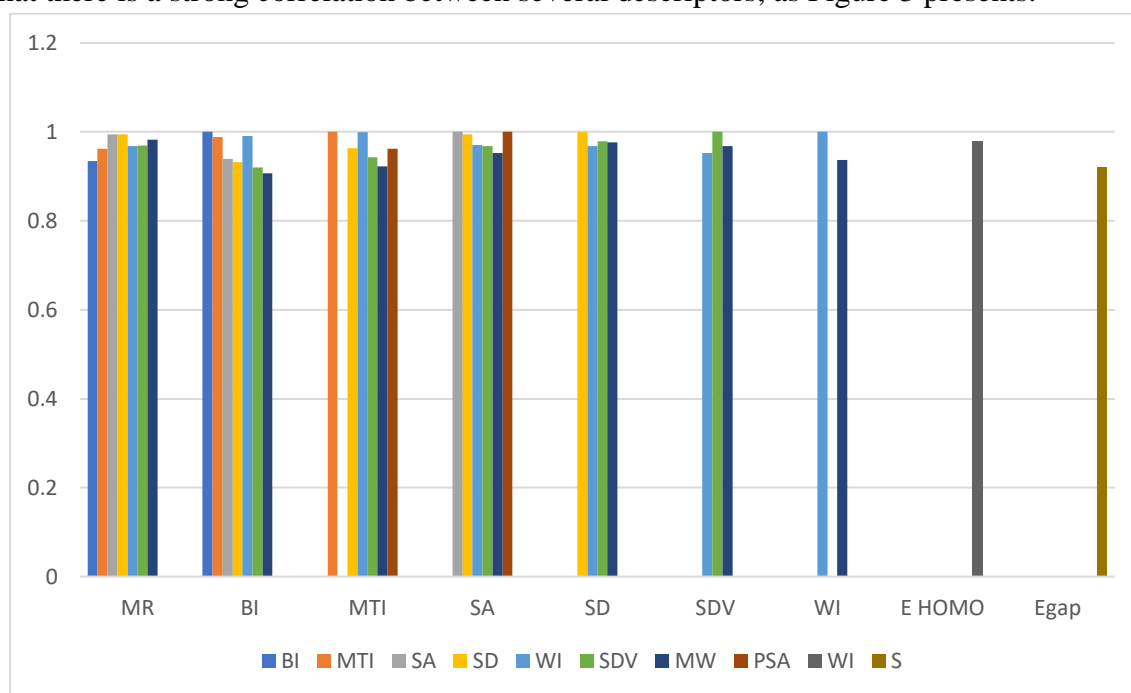
Figure 3. A strong correlation between descriptors.

On the other hand, we proceeded to eliminate one of the two highly correlated descriptors (MR, SA, WI, BI, SD, MTI, SDV, WI). Therefore, we used the PCA technique to reduce the number of remaining descriptors.
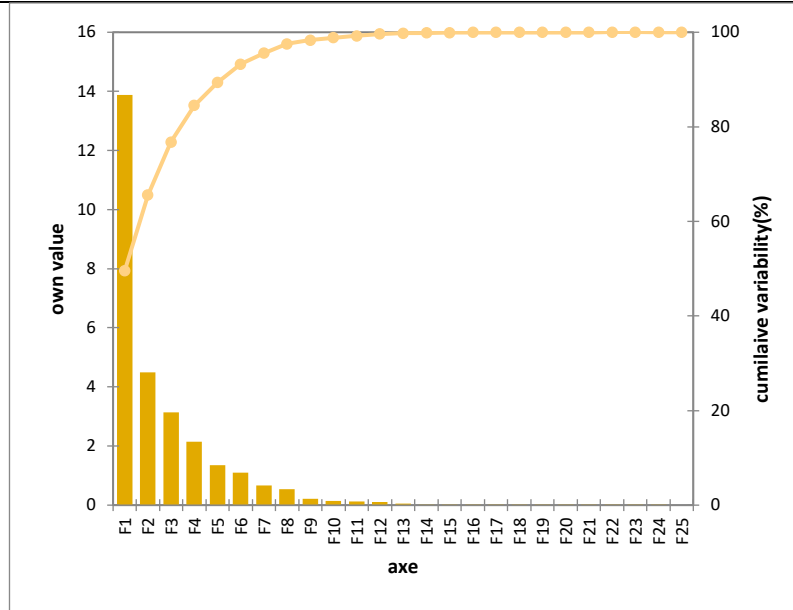
Figure 4. New variables by the PCA method.

The new variables created by the PCA show that F1 and F2 (Figure 4) are the two variables that give more information on pIC$_{50}$activity, with a percentage of 49.551% and 16.013%, respectively.
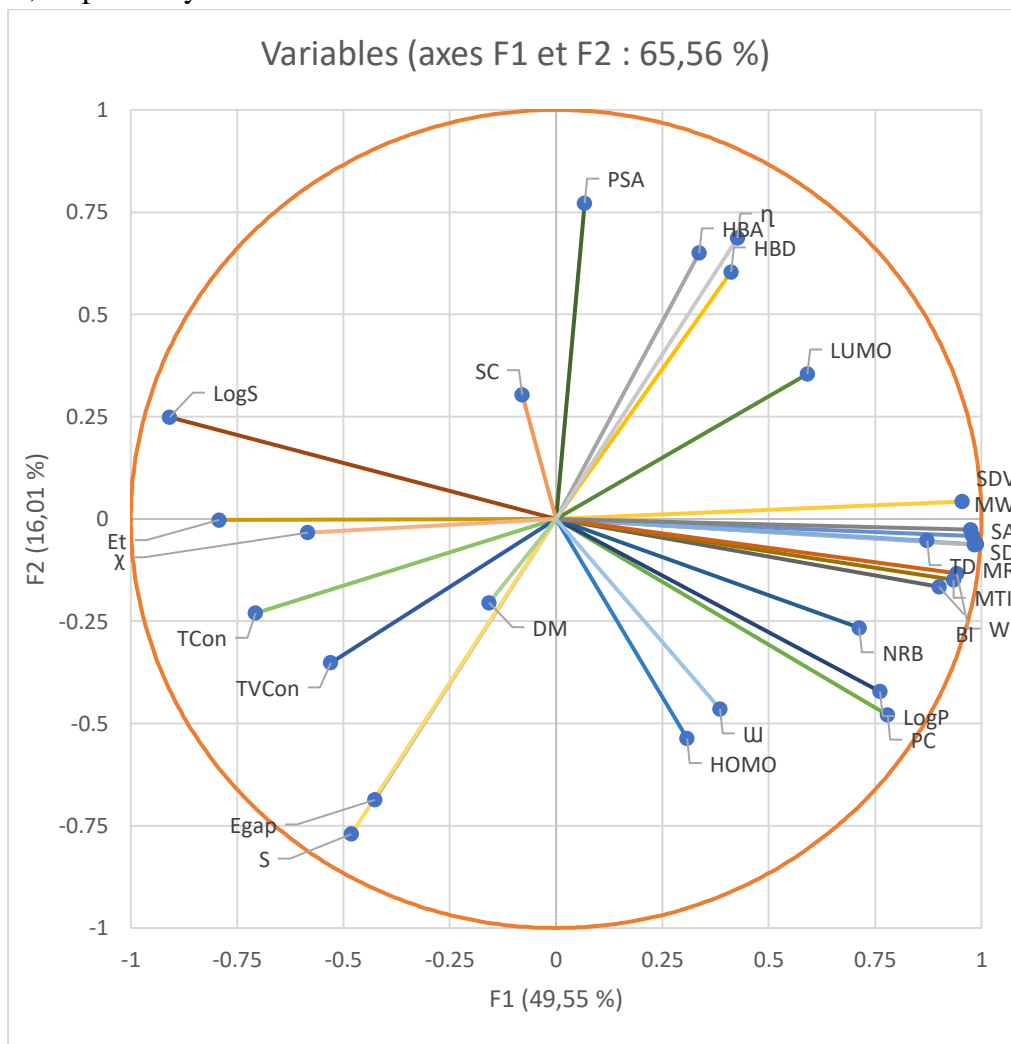
Figure 5. Corrélation cercle.

From the new variables F1 and F2 (Figure 5) we can see that:

A strong positive correlation: between the variables furthest at the edge of the circle (MW, SA, TD, SD, MR, and MR) and (WI, MTI and BI) represented by F2.

A strong negative correlation: between the variables at the edge of the circle (Et, X) represented by F2

No correlation: between variables that are almost straight to each other (angle ≈ 900), (PSA and LogS), (SC and TVcon), and (HDA and Ш).

Then we used multiple linear regression (MLR) to extract the original data from the new uncorrelated variables called latent variables that are none other than principal components of the original variables. Therefore, we obtained four descriptors (LogP: The partition coefficient, E Homo: The highest energy, ELumo: The lowest energy, and DM: The dipole moment) (Figure 6) that we will use in our study as inputs during the development of QSAR models explaining the relationships of the biological activities of dimedone derivatives with kinase inhibitory effects.
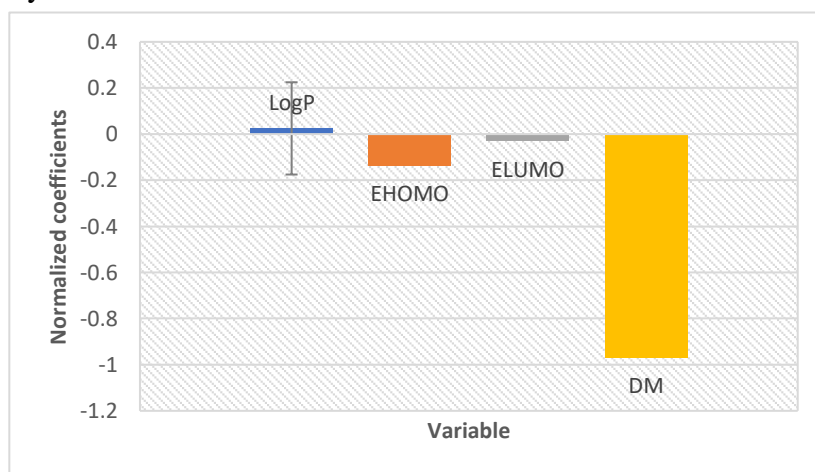
Figure 6. pIC_{50} / Normalized coefficients.

We divided the database into two sets by the k-means method, learning set (5a,10a, 11, 12a, 13a, 14a, 17a, 7b, 9a, 9b, 10b, 12b, 13b, 14a, 16a, 18a, 19a,7d, 12d, 17c, 17e, 16b, 18c, 18e, 19b, 19c,17d, 17f, 18f) and test set (3, 5b, 7a, 7c, 12c, 17b,18b,18d) (table 3).

Table 3. The results of k-means.

Classe	I	II	III	IV	V	VI	VII	VIII
Observations	3	5b	7a	7c	12c	17b	18b	18d
	5a	14b	7b	7d	16b	17d		18f
	10a	17a	9a	12d	16c	17f		
	11		9b	17c	18c			
	12a		10b	17e	18e			
	13a		12b		19b			
			13b		19c			
			14a					
			16a					
			18a					
			19a					

3.1. Linear multiple regression.

The QSAR model obtained by RLM is carried out using the values of the selected chemical descriptors and values of the inhibitory activity of c-Met (pIC_{50}). The constructed QSAR model is represented by the following linear equation 4:

pIC_{50} = 5.63+6.99E-03***LogP**-4.30E 02***E_{HOMO}**-3.325E-02***E_{LUMO}**-0.17***DM** (Equation 4)

According to the QSAR model equation obtained, the values of biological activity are linearly correlated with four descriptors. Different parameters are calculated for the evaluation of the performance of QSAR models, whichare: N, R, R^2, Adjusted R^2, MSE, P, and F.

N=30; R=0.9555; R^2 = 0.913; Adjusted R^2 = 0.899; R^2_{cv} =0.85; MSE = 0.004; P < 0.0001; F =65.255

A higher value of R^2 and R^2_{adj} and a low quadratic error value claim that the proposed model is reliable and predictive.

The model obtained by RLM is validated internally by the leave-one-out cross-validation technique, such that the cross-validation coefficient is R^2_{cv} greater than 0.6, which confirms the good predictability of our QSAR model.

The P of the model is less than 0.0001; this means that the linear regression equation has statistical significance. Therefore, we can say that the selected variables carry a significant amount of information.

The predictive power of the QSAR model obtained by MLR of the compounds of the test set(8 remaining compounds) is tested by external validation. The value of R^2_{test}=0.934 is greaterthan 0.6, which confirms the robustness of the model developed to predict the activities of themolecules of the test set.

Table 3 presents the values of the activities calculated fromequation (4) of the QSAR model obtained by MLR and the values of the observed activities. Figure 7 shows the correlations between observed and calculated activities.

In our model, the partition coefficient (LogP) positively influences homo activities and energy, LUMO energy and dipole moment (MD) negatively influence activities.

In equation (4), LogP has a positive sign which means that biological activity and LogP increases in parallel. To increase the activity value, we can increase the concentration of the organic phase and decrease the concentration of the aqueous phase.

The HOMO energy (negative values) has a negative sign in the model, which leads to the higher the HOMO, the easier the molecule will give up electrons, which shows the ease of feasibility of the nucleophilic reaction and the high activity value.

The LUMO energy (negative values) has a negative sign in equation (4) which leads that thelower the energy of this OM is, the easier the molecule will accept electrons.

The dipole moment (DM) has a negative sign in the model which means that biological activity decreases with the increase in MD.
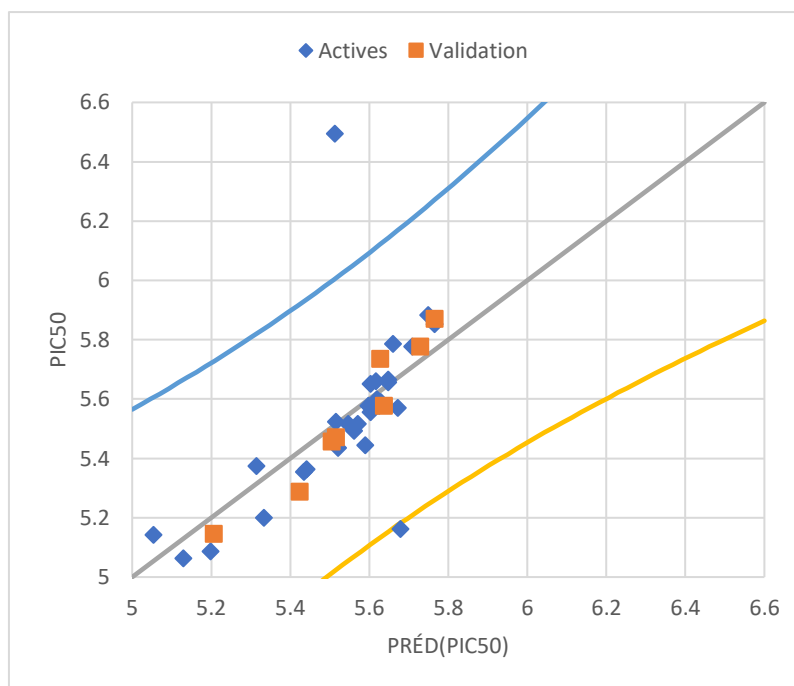
Figure 7. The relationship between the values of the activities observed and predicted by the MLR method.

In conclusion, the results indicate that to improve the biological activity of the cyclohexane-1,3-dione-dimidone-derived molecule structures examined in this work, we must endorse molecular structures with promoter (LogP) and nucleophilic groups that reduce the dipole and electrophilic moment.

3.2. Y-randomization test.

The Y-randomization test is performed on the original QSAR model obtained by the RLM method. The Y-randomization test is implemented according to the random distribution of pIC_{50} activity values fifty times without changing the four descriptors in equation (1). In order to confirm the quality of the developed QSAR model, to predict the biological inhibitory activity of c- Met from the four obtained molecular descriptors.

This random distribution of biological activities (Y) to descriptors (X) generates fifty new QSAR models; each of these models is characterized by values (R, R^2, R^2_{cv}). According to the results obtained, it can be noticed that the values of the random distribution are lower than the values of the original model, which affirms the strong correlation between the activity and the four descriptors and the robustness of our original model.

3.3. Domain of applicability.

We based on the obtained MLR model to realize the applicability domain. The applicability domain is examined by leverage. Figure 7 shows William's graph for the threshold value of leverage h*=0.416 with the interval ± x (x = 2.5).

From Figure 8, we notice that one molecule in the test set (5b) is outside the applicability range due to the activity value not being predicted correctly; the erroneous experimental data can justify this. For this reason, we will exclude molecule (5b) for the remainder of this study.
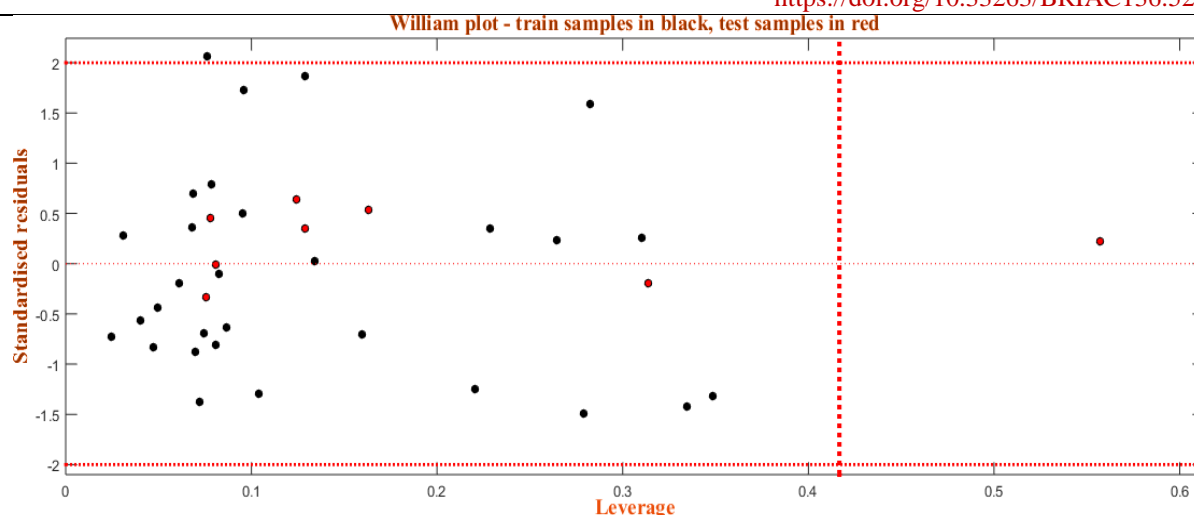
Figure 8. Williams graph of the leverage-normalized residue for the pIC50 MLR model.

3.4. Non-linear multiple regression.

We used the non-linear regression technique to improve the effectiveness of different descriptors on C-Met inhibitory activity. The descriptors used for this technique are the same descriptors proposed by MLR. To select the best performance, we calculated the coefficients, R and R^2. Equation (5) showsthe non-linear QSAR model obtained by the MNLR method.

$$\textbf{pIC}_{\textbf{50}}=6.805+2.041E^{03}*\textbf{LogP}+2.1674E^{04}*\textbf{LogP}^{\wedge 2}+0.126*\textbf{E}_{\textbf{HOMO}}+4.971E^{02}*\textbf{E}_{\textbf{HOMO}}{}^{\wedge 2}$$
$$+\textbf{0.23}*\textbf{E}_{\textbf{LUMO}}+2.028E^{-02}*\textbf{E}_{\textbf{LUMO}}{}^{\wedge 2}+ 0.229*\textbf{DM}+1.31E\text{-}02*\textbf{DM}^{\wedge 2}\qquad (\text{Equation 5})$$

Different parameters are calculated for the evaluation of the performance of QSAR models, which are: N, R, R^2, MCE, and R^2cv.

N=30; R=0.9954; R^2=0.991; MCE= 0.001; R^2_{cv}=0.83

These values confirm that the model is statistically acceptable. The values of the activities predicted by this model are quite similar to those observed, whichshows the uniform distribution of the observed experimental values of pIC$_{50}$ compared to thepredicted values presented in Figure 9.

The model obtained by RNLM is validated internally by the leave-one-out cross-validation technique, as the cross-validation coefficient is R^2_{cv} =0.829509; therefore, greater than 0.6, which confirms the good predictability of our QSAR model.

The predictive power of the QSAR model obtained by MNLR of the compounds of the test set (8 remaining compounds) is tested by an external validation as the value of R^2test= 0.997 is greater than 0.6, which confirms the robustness of the model developed to predict the activities of the molecules of the test set.

Figure 8 shows good correlations between observed and calculated activities.

The values of the activities calculated from equation (5) and the values of the observed activities for the learning set and the test set show good correlations.

In our model, the partition coefficient (LogP) positively influences homo activities and energy,and LUMO energy and dipole moment (MD) negatively influence activities.
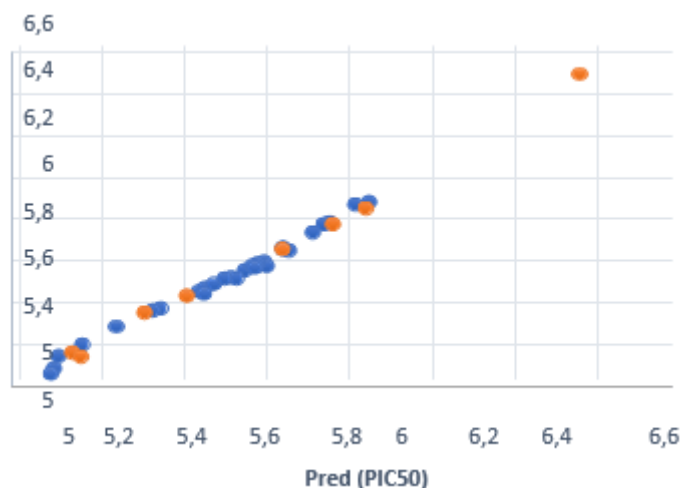
Figure 9. The relationship between the values of the activities observed and predicted by the MNLR method.

3.5. Artificial Neural Networks (ANN).

In our work, four descriptors are used as an input and the inhibitory activity of c-Met as anoutput. The number of hidden neurons is 4 to have the following architecture: 4-4-1 (Figure 10).
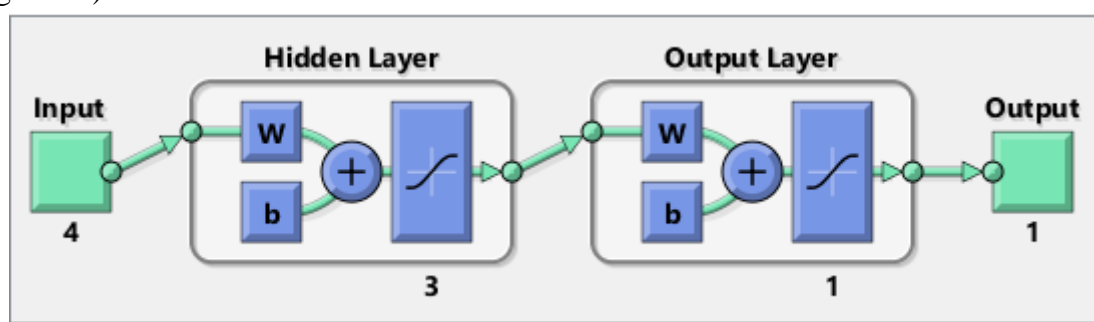
Figure 10. ANN Architecture.

Different parameters are calculated to evaluate the performance of the ANN model. $N=30$; $R= 0.994$; $R^2=0.99$; MSE$= 0.00047$; $R^2_{cv} = 0.989$; $p< 0.0001$

A low value of the quadratic error MSE and a high value of R2 ensures that the ANN model is reliable and predictive.

From the value $R^2_{cv} = 0.989$, which is lower than the value R2 = 0.99, it is clear that the ANN model is efficient.

According to the p-value obtained, which is between 1 and 3, it is attained that the RNA model is statistically acceptable, along with the predictions realized thanks to this model, are made in such a way as to ensure the contribution of all the compounds of the database used.

This model follows an external validation on the test series (8 compounds) from where the value of R2test is 0.954; this value is greater than 0.6, which confirms the predictive power.

From these results, we can say that the QSAR model is reliable for predicting the anticancer biological activity of the compounds studied. According to Figure 10, the distribution of pIC50 values candidates in learning is uniform with the test sets, which states that the pIC50 values obtained by the predictions of the ANN model are very close to the values observed experimentally (Table 6).
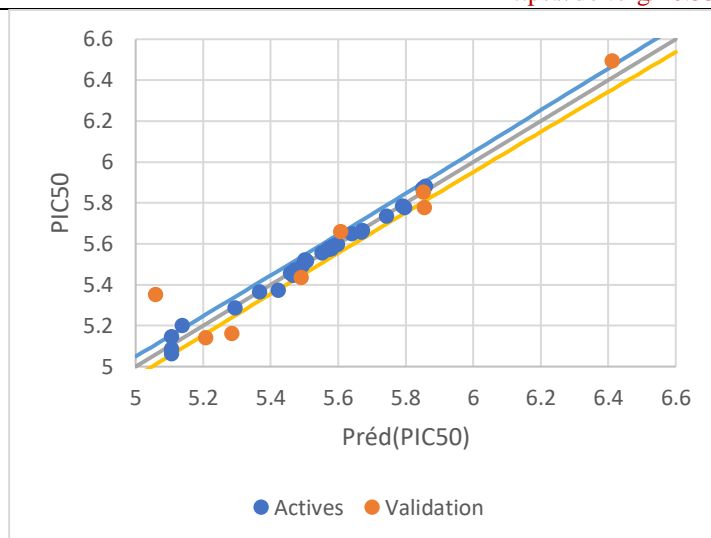
Figure 11. Correlations between observed activity values and predicted activity values via the ANN model.

3.6. Proposal of new anticancer molecules.

We used model 7d to design new molecules effective in inhibiting the non-small cell lung cancer (NSCLC) cell line. Molecule 7d exhibits the highest biological activity (pIC50=6.60). We designed 3 molecules by structural modifications (addition or deletion of nucleophilic and electrophilic groups) of molecule 7d. The prediction of the C-Met inhibitory activities of the designed molecules is made by the QSAR model equation A obtained by RLM (Table 4).

We chose molecules 1 and 2 to do the molecular docking because of their high activities according to the QSAR model via RLM.

Table 4. Molecules designed and their predicted biological activities.

The model molecule

7d	$pIC_{50} = 6.60$

The molecules designed

1	$pIC_{50}=7.06168844$	2	$pIC_{50}= 7.11784795$	3	$pIC_{50}=5.54971517$

3.7. Evaluation of the drug-likeness properties of the molecules designed.

Druglikness is a necessary concept after the design of molecules to determine if these designed molecules have the properties of a drug and to check Lipinski's rules simultaneously. This study is carried out using a SwissADME online server. Table 5 presents the results obtained from evaluating the drug-like properties of the molecules designed.

Table 5. Drug-likeness properties of the molecules designed.

Molecule	S. Has	MW	NROTB	BS	TPSA(^{A2})	HBD	HBA	Log P	Lipinski's violations	Violations de Veber	Egan's violations
	0<SA<10	<500	-	-	-	<5	<10	≤5	≤1	≤1	≤1
1	4,2	363,41	2	High	109,07	3	4	2,72	0	0	0
2	4,22	362,43	2	High	83,05	2	4	1,75	0	0	0
3	4,13	336,45	4	High	113,71	4	5	2,16	0	0	0
7d	3,47	387,46	5	High	121,05	1	4	1,08	0	0	0

From Table 5, we notice that all the molecules designed comply with the rules of Lipinski, Veber, and Egan, so these molecules do not have problems with oral bioavailability besides having properties similar to those of a drug. The assessment of the (S.A) of the molecule ranges from 1 (easy *to* synthesize) to 10 (very difficult to synthesize) [32]. We observe that all the values of S.A are between 3.74<SA<4.22 while these values are far from 10, which means they are easy to synthesize. The molecules are flexible and interact with the protein because of the number of rotational bonds NROTB <10 and the area of the polar surface TPSA<140 [33]. We select molecules 2 and 13 that have higher biological activities than the observed biological activities. Thereafter, we perform a molecular docking test of the new molecules 2 and 13 on the C-met receptor to find potential modes of interaction between the active residue sites in the pocket of the c-Met receptor and the structures of the molecules designed. As a result, we can predict the binding energies of the c-Met receptor, the proposed new ligands, and the Foretinib ligand. This may allow us to anticipate new candidate structures to replace Foretinib in cancer treatment.

3.8. Molecular docking.

First, we identified the active sites of the crystal complex (PDB code: 6SD9) with which the co-crystallized ligand Foretinib interacts. We used the online software ProteinsPlus [34] to identify the active sites (Asp1222A, Lys1110A, Phe1223A, and Met1160A) of the Foretinib inhibitor (Figure 12).
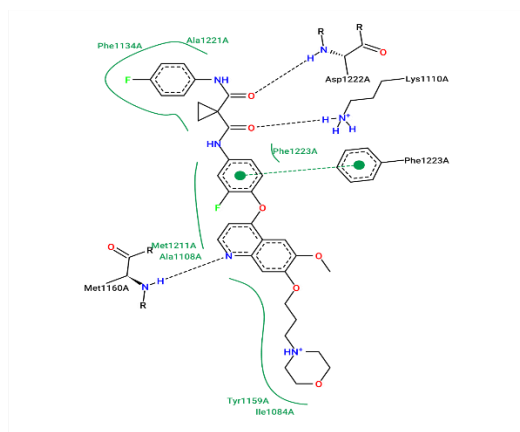
Figure 12. The 2D diagram of the interactions between the active sites in the 6SD9 crystalline complex and the co-crystallized ligand (Foretinib).

We used Discovery Studio 2021 software to visualize the types of interactions that were created between the co-crystallized ligand Foretinib and the receptor 6SD9. Three conventional hydrogen bonds, a hydrogen bond Met1160 with the benzene ring at a distance of 3,26A°, a hydrogen bond Asp1222 with the nitrogen attached to the benzene ring at a distance of 2,76 A°, a hydrogen bond Lys 1110 with the ketone at a distance 3,07 A°. Two hydrogen carbon bonds, Pro1110 and His1201, with two different benzene rings at a distance, respectively (3,8A° and 3,43A°). A halogen bond (fluorine) Val1220 with the fluorine attached to the benzene ring at a distance of 3,63A°. A pi-sigma bond Met1211 with the benzene ring at a distance of 3,75A°. A pi-sulfur bond Met1131 with a benzene ring at a distance of 5,57A°. Two stacked pi-pi bonds, Phe1223 and Tyr1159, with a different benzene ring at a distance respectively (3,25A°, 4,3A°) and 4 alkyl and pi-alkyl bonds Ala1221-Leu1157-Ala1108-Ile1084. The binding energy value of ligand 2 with the C-Met kinase receptor is equal to -9,41 kcal / mol. Figure 13 shows the 2D and 3D interactions between Foretinib and the receptor.
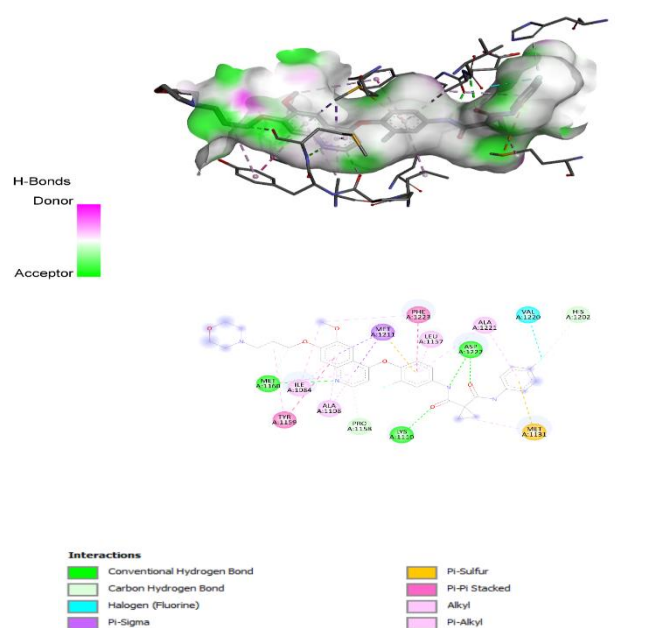
Figure 13. 2D and 3D of interactions between the receptor 6SD9 and foretinib.

To validate the simulation results, the template and the control drug, the co-crystallized ligand, were also re-docked into the binding pocket of the C-met receptor (Figure 14). The root-mean-square deviation (RMSD) was used to evaluate how different the obtained orientation is from the corresponding co-crystallized pose of the same ligand molecule [39]. The lower the value of RMSD, the more the alignment is perfect. The RMSD value obtained is 0, 9132A° is less than 2 A°, which means a good overlay.
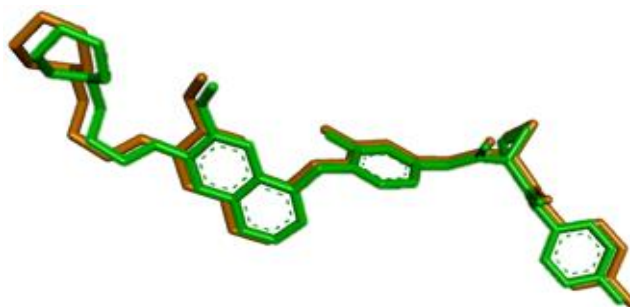
Figure 14. Original ligand redocking (green) and docké (orange) with RMSD 0,9132A.

Ligand 7d is bound with eight amino acid residues and two hydrogen bonds, Lys1110 with the hydrogen of the atom attached to the pyrimidine at a distance of 2.97 A° and Glu1127 with the hydrogen of the pyrimidine ring at a distance 2,79A°. A pi-sulfur interaction for Met1211 with the benzene ring at a distance of 5,24A°. A pi-lone pair interaction for Asp 1222 with the pyrimidine ring at a distance of 2.97 A°. Five pi-alkyl interactions are observed with Ala 1221-Leu1140-Leu1157-Phe1200-Met1131. The binding energy value of ligand 2 with the C-Met kinase receptor is equal to -9,46 kcal / mol. The 3D and 2D interactions of ligand 7d with the receptors are shown in Figure 15, respectively.
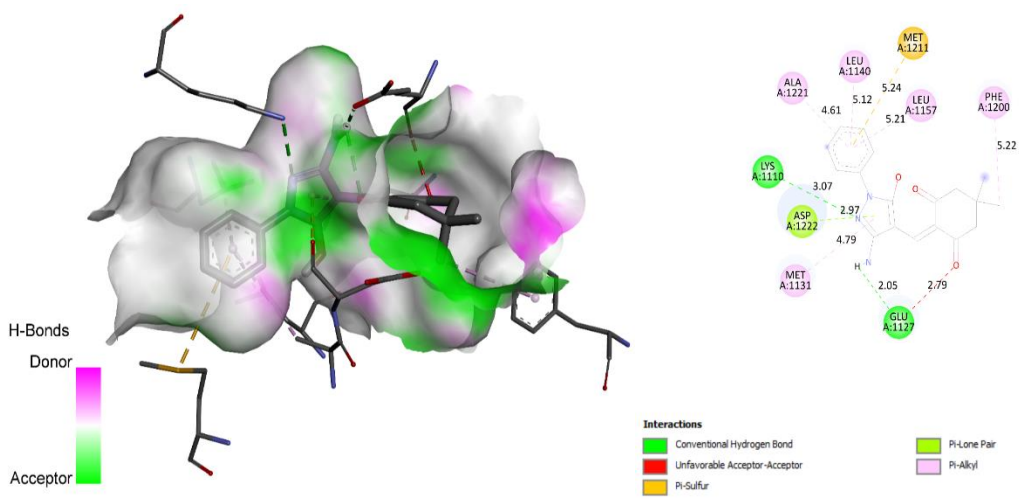
Figure 15. 3D and 2D Interactions of L7d with 6SD9 receptor.
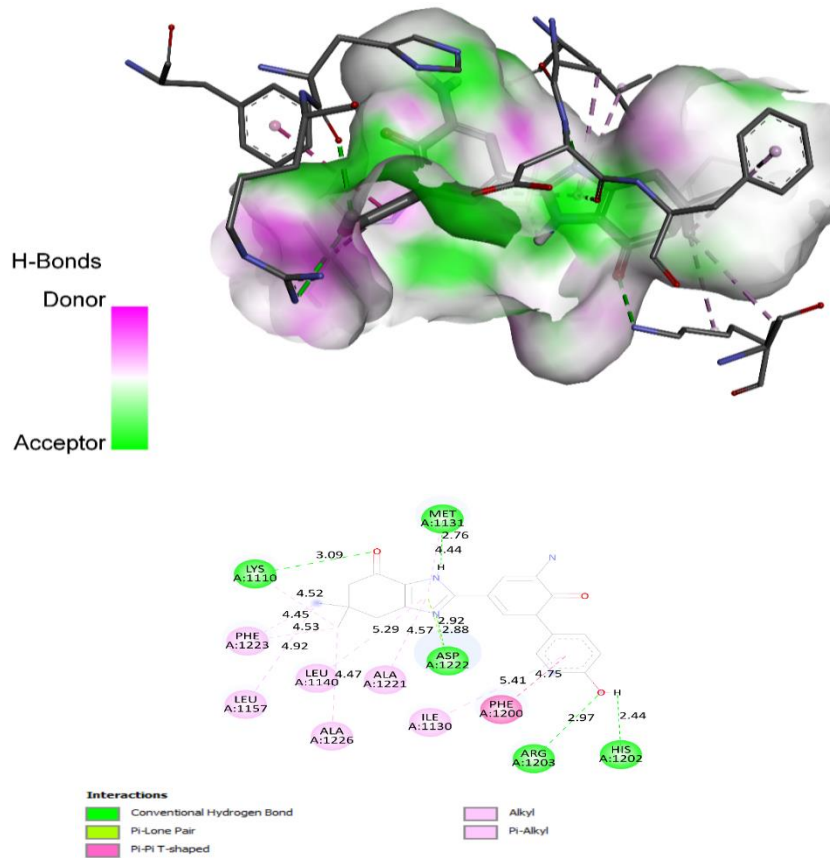
Figure16. 3D and 2D Interactions of L2 with 6SD9 receptor.

Figure 16 shows the 2D and 3D interactions of the designed ligand 2 with the 6SD9 receptor. Ligand 1 is bonded to the receptor by five conventional hydrogen bonds, Arg 1203 and His 1202, with the hydrogen atom of the hydroxyl group attached to the benzene at a distance of 2,44A°. Two hydrogen bonds with Met1131 and Asp1222 with the nitrogen of the pyrimidine ring at a distance respectively (2,76 A° and 2,88A°) and one hydrogen bond Lys 1110 with the ketone attached to cyclohexane at a distance of 3,09A°. Six pi-Alkyl and Alkyl bonds are observed with Phe1223-Leu1157-Leu1140-Ala1221-Ala1226-Ile1130 and one pipi-shaped Phe 1200 bond with the benzene ring at a distance of 4,75 A°. The binding energy value of ligand 2 with the C-Met kinase receptor is equal to -9,12 kcal / mol.

Ligand 2 is bound with the receptor by a single hydrogen bond, one pi-anion bond, one pi-pi-shaped bond, and 10 pi-alkyl and alkyl bonds.

The hydrogen bond His1202 with the benzene ring at a distance of 1,74 A°, a pi-anion bond Asp1222 with the benzene ring at a distance of 3,83A°, ten pi-alkyl, and alkyl bonds are observed with Ala1226-Val1092-Lys1110-Phe1223-Leu1157-Ala1221-Met1131-Phe1134-Val1139-Leu1140. A pi-pi T-shaped bond Phe1200 with the benzene ring at a distance of 485A°. The binding energy value of ligand 2 with the C-Met kinase receptor is equal to -9,43 kcal / mol.

Figure 17 shows the 2D and 3D interactions of ligand 13 with the 6SD9 receptor.

Figure 17. 3D and 2D Interactions of L13 with 6SD9 receptor.

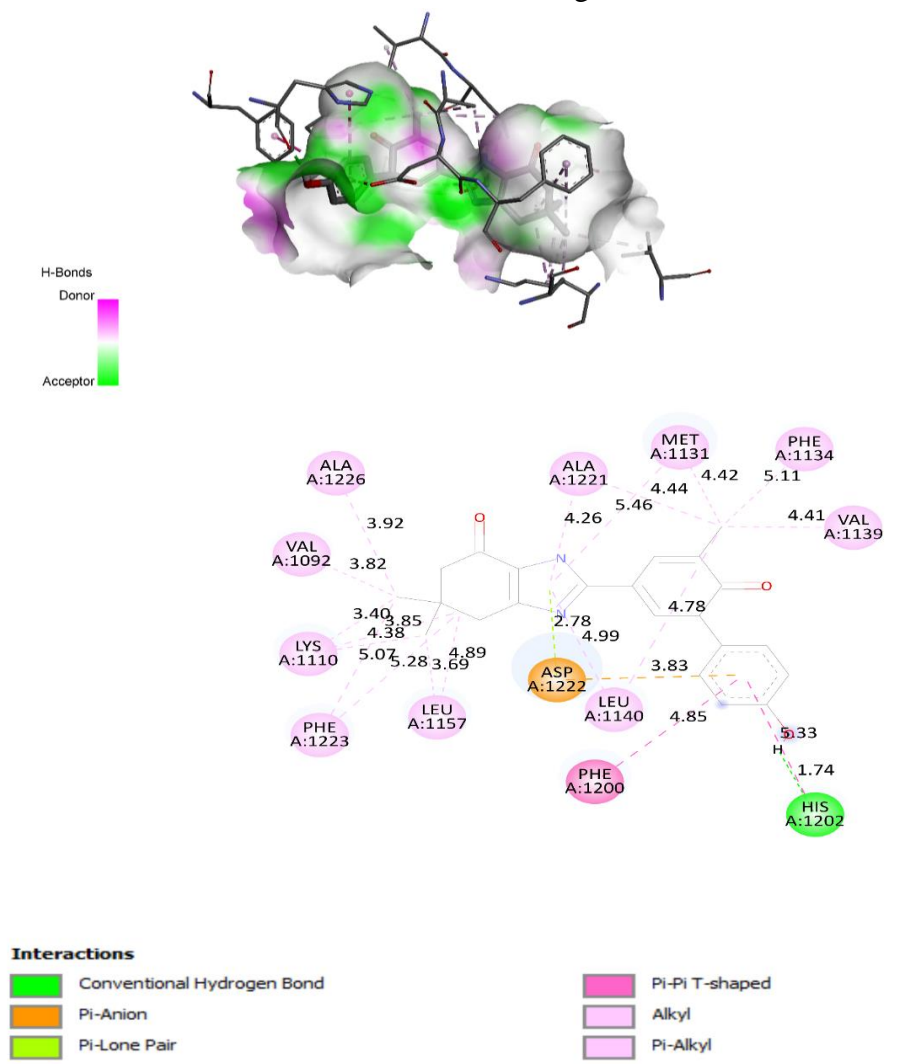The docking outcomes of the designed molecules suggested that Hydrogen bonds, electrostatic and hydrophobic (alkyl and Pi-Alkyl) interactions are the main driving forces that control the molecules' binding interactions and the receptor. Table 5 summarizes the interactions between ligands and C-met kinase with bondenergies.

To ensure that the designed molecules are the likely drug candidates, their ADMET and drug-likeness properties were examined. The predicted drug-likeness properties of the formulated compounds are expressed in Table 6, and their predicted ADMET (Absorption, Distribution, Metabolism, Excretion, Toxicity) properties are depicted in this table. We have realized it through the pkCSM online server [35].

Table 5. Docking results of L7d, L1, L2, and Lref ligands with C-met kinase sites.

Ligands	Hydrogen bond interactions	Interactions by hydrophobic bonds	Bond energies (Kcal\mol)
Foretinib	Met 1160, Asp1222, Lys1110 Conventional H-bond His1202, Pro1158 Carbon-Hydrogen bond	Phe1223, Tyr1159 pi-pi stacked Leu1157, Ala1108, Ala1121, Ile1084 alkyl and pi-alkyl Phe1223, Tyr1159 pi-pi stacked Met1211 pi-sigma Met1131 pi-sulfur Val1220 halogene(fluorin)	-9,41
Ligan7d	Lys1110A, Glu1127 Conventional H-bond	Ala1221, Leu1140, Leu1157, Phe1200, Met1131 pi -alkyl Asp1222 pi-lone paire Met1211 pi-sulfur	-9,46
Ligand 1	Met 1131, Lys1110, Asp1222, His1202 Carbon-Hydrogen bond	Phe1223, Leu1157, Ala1221, Arg1203, Leu1140, Ile1130, Ala 1226 , Phe1200 pi-pi stacked	-9.12
Ligand 2	His1202 Conventional H-bond	Ala 1221, Met1131, Lys1110, Phe1223, Leu1157, Ala1226, Phe1134, Val1092, Val1139, Leu1140 pi alkyl and alkyl Asp1222 pi-anion Phe1200 pi-pi Tshaped	-9.43

Table 6. The ADMET Properties predicted in silico for compounds 2, 13, and Foretinib.

N	Absorption	Distribution				Metabolism							Excretion	Toxicity
	Intestinal absorption (human)	VDss	fraction Unbound	BBB Permeability	CNS permeability	CYP							Total clearance	
						Substrate		Inhibitors						
						2D6	3A4	1A2	2C19	2C9	2D6	3A4		
1	78,80	0,54	0,227	-0,94	-2,374	No	No	No	No	No	No	No	0,779	No
2	93,82	0,64	0,118	-0,75	-0,051	Yes	Yes	Yes	No	No	Yes	Yes	1,002	No

The compounds present an interesting intestinal absorption (human) (78,8 and 93,82) which shows their good absorption by the human intestine because for a molecule to be badly absorbed, its value of absorbance must be at most 30%.

Greater than 0.3 to less than −1 log BB and greater than −2 to less than −3 log PS are the standard values for blood-brain barrier (BBB) and the central nervous system (CNS) permeability. Hence, for any compound, log BB less than −1 suggests a poor distribution of the drug to 26the brain, while a value of log BB greater than 0.3 recommends that the drug can pass over the blood-brain barrier.

A greater than -2 value for log PS, means that the drug candidate can penetrate the CNS [36], while a value less than −3 expresses the difficulty molecule to induce into the central nervous system [37].

Compound 2 indicated that it has a high potential to be distributed easily to the brain and penetrate the CNS.

Metabolism is the property that describes the biochemical modification of a drug by the body. Therefore, the metabolism of drug metabolism and drug interactions must be considered [38]. CYP enzymes are present in all tissues of the body [39]. The enzymes were identified as accountable for the biotransformation of more than 90% of the drugs that undergo phase-I metabolism [40]. The results in Table 6 imply that compound 2 is an inhibitor of the 2D6, 3A4, 1A2, 2C19, 2D6, and 3A4 enzymes. Clearance is important to determine drug dosing rations and their ejection rates. The compounds have a good value. Toxicity is an important property for choosing the best drugs. The designed compounds are not toxic. From Table 6, we notice that compound 2 meets all the properties of in-silico ADMET property evaluation. Therefore, compound 2 can be used as a drug to treat cancer by inhibiting the enzymatic activity of the c-Met protein.

4. Conclusions

In this work, we did a 2D QSAR study of a series of 38 molecules of Cyclohexane-1,3-dione derivatives (dimedone) that have c-Met inhibitory activity against NSCLC to develop QSAR models. The three QSAR models (MLR, MNLR, and ANN) have shown good statistical significance and predictive power during internal and external validation as well as y-randomization testing and domain of applicability. This study shows that 4 descriptors (LogP, EHOMO, ELUMO, and DM) effectively design new molecules based on Cyclohexane-1,3-dione derivatives (dimedone) and specifically on the 7d molecule. Among the 3 molecules designed, two compounds (molecules 1 and 2) have biological activities superior to the biological activities observed; for this reason, we did the molecular docking of these molecules (1 and 2) to determine the most stable ligand in the pocket of the c-met. Based on the molecular docking of compounds (1, 2) with the c-met protein, we found that molecule 2 is a good inhibitor of the enzymatic activity of the c-Met protein. Therefore, compound 2 makes 6 interactions with c-met sites; also, the pocket of the c-Met receptor is stable because of their binding energy. In addition, the in-silico evaluation of the ADMET properties of compound 2 indicates that it perfectly fulfills the pharmacokinetic properties. According to the results obtained in this work, the structure of derivatives of Cyclohexane-1,3-dione (dimedone) and specifically compound 2 can be used to design drugs for the treatment of lung cancer (NSCLC) after conducting additional investigations such as in vivo and in vitro synthesis before the procedure to clinical trials.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Mohana Roopan, S.; Sompalle, R. Synthetic Chemistry of Pyrimidines and Fused Pyrimidines: A Review. *Synthetic Communications* **2016**, *46*, 645–672, https://doi.org/10.1080/00397911.2016.1165254.
2. Wen, T.; Wang, J.; Shi, Y.; Qian, H.; Liu, P. Inhibitors Targeting Bruton's Tyrosine Kinase in Cancers: Drug Development Advances. *Leukemia* **2021**, *35*, 312–332, https://doi.org/10.1038/s41375-020-01072-6.
3. Schlam, I.; Swain, S.M. HER2-Positive Breast Cancer and Tyrosine Kinase Inhibitors: The Time Is Now. *npj Breast Cancer* **2021**, *7*, 1–12, https://doi.org/10.1038/s41523-021-00265-1.
4. Mathew, S.; Tess, D.; Burchett, W.; Chang, G.; Woody, N.; Keefer, C.; Orozco, C.; Lin, J.; Jordan, S.; Yamazaki, S.; *et al.* Evaluation of Prediction Accuracy for Volume of Distribution in Rat and Human Using In Vitro, In Vivo, PBPK and QSAR Methods. *Journal of Pharmaceutical Sciences* **2021**, *110*, 1799–1823, https://doi.org/10.1016/j.xphs.2020.12.005.
5. Zillhardt, M.; Park, S.-M.; Romero, I.L.; Sawada, K.; Montag, A.; Krausz, T.; Yamada, S.D.; Peter, M.E.; Lengyel, E. Foretinib (GSK1363089), an Orally Available Multikinase Inhibitor of c-Met and VEGFR-2, Blocks Proliferation, Induces Anoikis, and Impairs Ovarian Cancer Metastasis. *Clinical Cancer Research* **2011**, *17*, 4042–4051, https://doi.org/10.1158/1078-0432.ccr-10-3387.
6. Mohareb, R.M.; Manhi, F.M.; Mahmoud, M.A.A.; Abdelwahab, A. Uses of Dimedone to Synthesis Pyrazole, Isoxazole and Thiophene Derivatives with Antiproliferative, Tyrosine Kinase and Pim-1 Kinase Inhibitions. *Med Chem Res* **2020**, *29*, 1536–1551, https://doi.org/10.1007/s00044-020-02579-4.
7. Österberg, T.; Norinder, U. Prediction of Drug Transport Processes Using Simple Parameters and PLS Statistics The Use of ACD/LogP and ACD/ChemSketch Descriptors. *European Journal of Pharmaceutical Sciences* **2001**, *12*, 327–337, https://doi.org/10.1016/S0928-0987(00)00189-5.
8. Li, Z.; Wan, H.; Shi, Y.; Ouyang, P. Personal Experience with Four Kinds of Chemical Structure Drawing Software: Review on ChemDraw, ChemWindow, ISIS/Draw, and ChemSketch. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1886–1890, https://doi.org/10.1021/ci049794h.
9. Gaussian.Com | Expanding the Limits of Computational Chemistry Available online: https://gaussian.com/ (accessed on 13 January **2022**).
10. Hammoudan, I.; Matchi, S.; Bakhouch, M.; Belaidi, S.; Chtita, S. QSAR Modelling of Peptidomimetic Derivatives towards HKU4-CoV 3CLpro Inhibitors against MERS-CoV. *Chemistry* **2021**, *3*, 391–401, https://doi.org/10.3390/chemistry3010029.
11. Chtita, S.; Belhassan, A.; Bakhouch, M.; Taourati, A.I.; Aouidate, A.; Belaidi, S.; Moutaabbid, M.; Belaaouad, S.; Bouachrine, M.; Lakhlifi, T. QSAR Study of Unsymmetrical Aromatic Disulfides as Potent Avian SARS-CoV Main Protease Inhibitors Using Quantum Chemical Descriptors and Statistical Methods. *Chemometrics and Intelligent Laboratory Systems* **2021**, *210*, 104266, https://doi.org/10.1016/j.chemolab.2021.104266.
12. Hammoudan, I.; Chtita, S.; Bakhouch, M.; Temsamani, D.R. QSAR Study of a Series of Peptidomimetic Derivatives towards MERS-CoV Inhibitors. *Moroccan Journal of Chemistry* **2022**, *10*, 10–651, https://doi.org/10.48317/IMIST.PRSM/morjchem-v10i4.34288..
13. Salt, D.W.; Yildiz, N.; Livingstone, D.J.; Tinsley, C.J. The Use of Artificial Neural Networks in QSAR. *Pestic. Sci.* **1992**, *36*, 161–170, https://doi.org/10.1002/ps.2780360212.
14. David, C.C.; Jacobs, D.J. Principal Component Analysis: A Method for Determining the Essential Dynamics of Proteins. In *Protein Dynamics: Methods and Protocols*; Livesay, D.R., Ed.; Methods in Molecular Biology; Humana Press: Totowa, NJ, **2014**; pp. https://doi.org/10.1007%2F978-1-62703-658-0_11
15. Performance of Smoothly Clipped Absolute Deviation as a Variable Selection Method in the Artificial Neural Network-based QSAR Studies - Mozafari - **2021** - Journal of Chemometrics - Wiley Online Library https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/abs/10.1002/cem.3338.
16. Shen, Q.; Jiang, J.-H.; Jiao, C.-X.; Lin, W.-Q.; Shen, G.-L.; Yu, R.-Q. Hybridized Particle Swarm Algorithm for Adaptive Structure Training of Multilayer Feed-Forward Neural Network: QSAR Studies of Bioactivity of Organic Compounds. *J. Comput. Chem.* **2004**, *25*, 1726–1735, https://doi.org/10.1002/jcc.20094.

17. Roy, K.; Mitra, I. On Various Metrics Used for Validation of Predictive QSAR Models with Applications in Virtual Screening and Focused Library Design. *Combinatorial Chemistry & High Throughput Screening* **2011**, *14*, 450–474, https://doi.org/10.2174/138620711795767893.

18. Al, H.H. et Antiproliferative Activity: Discovery of New Benzoxanthenes Derivatives by Using Various Statistical Methods 2D/3D-QSAR and Molecular Docking. *RHAZES: Green and Applied Chemistry* **2021**, *12*, 40–59, https://doi.org/10.48419/IMIST.PRSM/rhazes-v10.23861.

19. Chalkha, M.; Akhazzane, M.; Moussaid, F.Z.; Daoui, O.; Nakkabi, A.; Bakhouch, M.; Chtita, S.; Elkhattabi, S.; Housseini, A.I.; Yazidi, M.E. Design, Synthesis, Characterization, in Vitro Screening, Molecular Docking, 3D-QSAR, and ADME-Tox Investigations of Novel Pyrazole Derivatives as Antimicrobial Agents. *New J. Chem.* **2022**, *46*, 2747–2760, https://doi.org/10.1039/D1NJ05621B..

20. Golbraikh, A.; Tropsha, A. Beware of Q2! *Journal of Molecular Graphics and Modelling* **2002**, *20*, 269–276, https://doi.org/10.1016/S1093-3263(01)00123-1.

21. Rücker, C.; Rücker, G.; Meringer, M. Y-Randomization and Its Variants in QSPR/QSAR. *J. Chem. Inf. Model.* **2007**, *47*, 2345–2357, https://doi.org/10.1021/ci700157b. .

22. Roy, K.; Kar, S.; Ambure, P. On a Simple Approach for Determining Applicability Domain of QSAR Models. *Chemometrics and Intelligent Laboratory Systems* **2015**, *145*, 22–29, https://doi.org/10.1016/j.chemolab.2015.04.013.

23. Chtita, S. Modélisation de molécules organiques hétérocycliques biologiquement actives par des méthodes QSAR/QSPR. Recherche de nouveaux médicaments. phdthesis, Université Moulay Ismaïl, Meknès, **2017**. https://tel.archives-ouvertes.fr/tel-01568788/.

24. Gurumallappa, G.; Jayaprakash, J.; Puttaswamy, A.A.; Sunderraj, J.H.; Mallu, P.; Beeregowda, K.N. 4-(Tert-Butyl)-N,N-Diethylbenzenesulfonamide: Structural, Absorption Distribution Metabolism Excretion Toxicity (ADMET) and Molecular Docking Studies, **2021**, http://dx.doi.org/10.5267/j.ccl.2021.3.004.

25. Mkhayar, K.; Daoui, O.; Elkhattabi, S.; Chtita, S.; Elkhalabi, R. In Silico Molecular Investigations of Derived Cyclohexane-1,3-Dione Compounds as Potential Inhibitors of Protein Tyrosine Kinase C-Met: 2D QSAR, Molecular Docking and ADMET. In Proceedings of the 2022 International Conference on Intelligent Systems and Computer Vision (ISCV); May **2022**; pp.1–8. https://doi.ieeecomputersociety.org/10.1109/ISCV54655.2022.9806058.

26. Antichagasic Evaluation, Molecular Docking and ADMET Properties of the Chalcone (2E)-3-(2-Fluorophenyl)-1-(2-Hydroxy- 3,4,6-Trimethoxyphenyl)Prop-2-En-1-One against Trypanosoma Cruzi: Journal of Biomolecular Structure and Dynamics: Vol 0, No 0 Available online: (accessed on 25 September **2022**). https://doi.org/10.1080/07391102.2022.2123394.   .

27. Rules-for-Drug-Discovery.Pdf. https://www.ddw-online.com/media/32/rules-for-drug-discovery.pdf

28. SwissADME: A Free Web Tool to Evaluate Pharmacokinetics, Drug-Likeness and Medicinal Chemistry Friendliness of Small Molecules | Scientific Reports https://www.nature.com/articles/srep42717. .

29. PkCSM: Predicting Small-Molecule Pharmacokinetic and Toxicity Properties Using Graph-Based Signatures | Journal of Medicinal Chemistry Available online: (accessed on 14 January **2022**). https://doi.org/10.1021/acs.jmedchem.5b00104..

30. Morris, G.M.; Lim-Wilby, M. Molecular Docking. In *Molecular Modeling of Proteins*; Kukol, A., Ed.; Methods Molecular Biology™; Humana Press: Totowa, NJ, **2008**; pp. 365–382 ISBN 978-1-59745-177-2. https://doi.org/10.1007/978-1-59745-177-2_19.

31. Vázquez-Jiménez, L.K.; Juárez-Saldivar, A.; Gómez-Escobedo, R.; Delgado-Maldonado, T.; Méndez-Álvarez, D.; Palos, I.; Bandyopadhyay, D.; Gaona-Lopez, C.; Ortiz-Pérez, E.; Nogueda-Torres, B.; *et al.* Ligand-Based Virtual Screening and Molecular Docking of Benzimidazoles as Potential Inhibitors of Triosephosphate Isomerase Identified New Trypanocidal Agents. *International Journal of Molecular Sciences* **2022**, *23*, 10047, https://doi.org/10.3390/ijms231710047.

32. Support Vector Regression-Based QSAR Models for Prediction of Antioxidant Activity of Phenolic Compounds | Scientific Reports https://www.nature.com/articles/s41598-021-88341-1.

33. PROTAC Targeted Protein Degraders: The Past Is Prologue | Nature Reviews Drug Discovery https://www.nature.com/articles/s41573-021-00371-6.

34. Zentrum Für Bioinformatik: Universität Hamburg - Proteins Plus Server Available online: (accessed on 28 April **2022**). https://proteins.plus/.

35. PkCSM: Predicting Small-Molecule Pharmacokinetic and Toxicity Properties Using Graph-Based Signatures | *Journal of Medicinal Chemistry* **2015**, *58*, 4066-4072, https://doi.org/10.1021/acs.jmedchem.5b00104.

36. Computational Pharmacokinetics Report, ADMET Study and Conceptual DFT-Based Estimation of the Chemical Reactivity Properties of Marine Cyclopeptides . https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8607802/.

37. El Rhabori, S.; El Aissouq, A.; Chtita, S.; Khalil, F. QSAR, Molecular Docking and ADMET Studies of Quinoline, Isoquinoline and Quinazoline Derivatives against Plasmodium Falciparum Malaria. *Struct Chem* **2022**, https://doi.org/10.1007/s11224-022-01988-y.

38. Drug Metabolism and Disposition: 50 (1). *Drug Metab Dispos* **2022**, *50*. https://dmd.aspetjournals.org/.

39. Daoui, O.; Elkhattabi, S.; Chtita, S.; Elkhalabi, R.; Zgou, H.; Benjelloun, A.T. QSAR, Molecular Docking and ADMET Properties in Silico Studies of Novel 4,5,6,7-Tetrahydrobenzo[D]-Thiazol-2-Yl Derivatives Derived from Dimedone as Potent Anti-Tumor Agents through Inhibition of C-Met Receptor Tyrosine Kinase. *Heliyon* **2021**, https://doi.org/10.1016/j.heliyon.2021.e07463.

40. Knoll, K.E.; van der Walt, M.M.; Loots, D.T. In Silico Drug Discovery Strategies Identified ADMET Properties of Decoquinate RMB041 and Its Potential Drug Targets against Mycobacterium Tuberculosis. *Microbiol Spectr* **2022**, https://doi.org/10.1128/spectrum.02315-21.