Research Article

3D-QSAR Study of the Chalcone Derivatives as Anticancer Agents

Larbi ElMchichi,1 Assia Belhassan,1 Tahar Lakhlifi,1 and Mohammed Bouachrine1,2

1Natural Substances and Molecular Chemistry Laboratory (SOOM), Faculty of Science, Moulay Ismail University of Meknes, Meknes, Morocco
2EST Khenifra, Sultan Moulay Slimane University, Beni Mellal, Morocco

Correspondence should be addressed to Mohammed Bouachrine; m.bouachrine@est-umi.ac.ma

Received 1 March 2020; Revised 15 June 2020; Accepted 3 July 2020; Published 27 July 2020

Academic Editor: Maria P. Robalo

Copyright © 2020 Larbi Elmchichi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

For their biological properties and particularly for their anticancer activities, chalcones are widely studied. In this work, we have submitted diverse sets of chalcone derivatives to the 3D-QSAR (3-dimensional quantitative structural-activity relationship) to study their anticancer activities against HTC116 (human colon cancer), relying on the 3-dimensional descriptors: steric and electrostatic descriptors for the CoMFA (comparative molecular field analysis) method and steric, electrostatic, hydrophobic, H-bond donor, and H-bond acceptor descriptors for the CoMSIA method. CoMFA as well as the CoMSIA model have encouraging values of the cross-validation coefficient ($Q^2$) of 0.608 and 0.806 and conventional correlation coefficient ($R^2$) of 0.960 and 0.934, respectively. Furthermore, values of $R^2_{\text{test}}$ have been obtained as 0.75 and 0.90, respectively. Besides, $y$-randomization test was also performed to validate our 3D-QSAR models. Based on these satisfactory results, ten new compounds have been designed and predicted by in silico ADMET method. This study could expand the understanding of chalcone derivatives as anticancer agents and would be of great help in lead optimization for early drug discovery of highly potent anticancer activity.

1. Introduction

Cancer is one of the most serious health problems in the world. Over the past three decades, the number of people with different types of cancer has increased, and the problem is expected to worsen in the next few years if new methods of treatment are not discovered [1, 2].

Chalcone precursors, called α-b-unsaturated ketones (Figure 1), play an important role in synthetic manipulations and also represent a major component of natural products. Chalcones as well as their synthetic analogues (Figure 1) display enormous number of biological activities [3]. The presence of chalcone derivatives in the biologically active compounds as a substituent, or main component, has prompted synthetic organic chemists to synthesize novel compounds bearing this moiety. Current research shows that molecules containing more than one pharmacophore are useful for the treatment of cancer [4, 5].

The biological activities of chalcones due to the presence of double bonds in conjunction with carbonyl functionality, as removal of this functionality, make them inactive [6].

In the modern drug design and discovery, quantitative structure-activity relationship (QSAR) methodology generally plays a crucial role [7]. QSAR studies are performed by correlation structural descriptors from a set of chemicals to their known biological activities and predict the activity values of nonexamined compounds lying in the applicability domain of the model [8, 9]. 3D-QSAR studies are successfully applied to guide the development and design of new novel molecules with an improved activity [10]. Steric and electrostatic fields have been used as an independent variable against the biological activity of compounds [11]. One of the types of 3D-QSAR technique is the CoMFA method. This method has an extension called CoMSIA with a difference only in the implementation of the fields [12, 13]. The correlation between the 3D structures of the studied molecules and their activities can be better interpreted relying on the calculation of hydrophobic, H-bond donor, and H-bond acceptor similarity fields, as well as the steric and electrostatic fields in the CoMSIA approach.

The objective behind this work is developing some 3D-QSAR techniques. This development includes CoMFA and CoMSIA methods for modeling and prediction of the
2. Materials and Methods

2.1. Computer Simulations. The study done in these research procedures relies on two main components (Figure 2): data preparation and ligand-based study (3D-QSAR model).

2.1.1. Data Set. A database of 19 compounds consisting of chalcone derivatives as anticancer agents against HTC116 (human colon cancer) is collected from the literature [14]. All compounds are chemically synthesized and experimentally evaluated in biochemical assays by Wang et al. [14]. The selection of the compounds used for our 3D-QSAR study was done by considering the fact that compounds represent structural diversity and a range of biological activities, as well as the chemistry of chalcones remains a fascination among researchers in the 21st century due to the large number of replaceable hydrogen molecules that allows a large number of derivatives and a variety of promising biological activities to be generated [15]. The structures and biological activities of both training and test sets are given in Table 1 and Figure 3. For the 3D-QSAR study, the reported IC$_{50}$ values of compounds have been converted into corresponding $p$IC$_{50}$ (−log IC$_{50}$) and subsequently used as the dependent variable for 3D-QSAR model development [16]. The total set of molecules has been randomly split into a training set (14 compounds) for generating the QSAR model and a test set (5 compounds) for validating the predictive ability of the model. The selection of training and test set molecules has been carried out manually such that the test set contains a range of low-, moderate-, and high-activity compounds similar to the training set.

2.1.2. Molecular Modeling. The molecular modeling studies of 19 compounds have been performed utilizing SYBYL-X 2.0 software package (Tripos Inc., St. Louis, USA), working on Windows 7.0 and 64 bit workstation. After the construction of the 3D structures of the molecules, energy minimization has been performed using the Tripos force field [17], with Powell conjugate gradient methods under the conditions of a gradient convergence of 0.01 kcal/(Å mol) in Gasteiger–Huckel charge [18], using 5000 iterations.

2.1.3. Molecular Alignment. The data set has been aligned on the common core by distill rigid alignment technique available in SYBYL [19], using the most active compound (compound N’17) as the template (Figure 4). The structural alignment of the molecules is the most critical routine in building a reliable 3D-QSAR model. The accuracy of the prediction and the statistic quality of 3D-QSAR models depend strongly on the alignment [20]. Figure 4 describes the proposed alignment and common substructure for the alignment.

2.2. CoMFA and CoMSIA Studies. In the generation of CoMFA models [21], the aligned molecules have been imported in a 3D cubic lattice generated automatically with a grid spacing of 2 Å. In the CoMFA studies, Tripos force field takes a sp3 carbon probe atom with a Van der Waals radius of 1.52 Å and positive charge +1 to generate steric (Lennard–Jones 6-12 potential) field energies and electrostatic (Coulombic potential) fields with a distance-dependent dielectric at each lattice point. To reduce noise and speed up the calculation of potentials, the column filtering value has been set to 2.0 Kcal/mol (default parameters).

Table 1: Chemical structures and anticancer activities of chalcone derivatives.

| N° | Structure  | pIC$_{50}$ |
|----|------------|------------|
| 1  | 2-Cl       | 5.89       |
| 2  | 4-Cl       | 5.65       |
| 3  | 2-Br       | 5.77       |
| 4  | 3-Br       | 5.80       |
| 5  | 4-Br       | 5.65       |
| 6  | 2-CH$_3$   | 5.72       |
| 7  | 3-CH$_3$   | 5.84       |
| 8  | 4-CH$_3$   | 5.77       |
| 9  | 3-OCH$_3$  | 5.43       |
| 10 | 4-OCH$_3$  | 5.31       |
| 11 | 3,5-(Cl)$_2$ | 5.78    |
| 12 | 3,5-(CH$_3$)$_2$ | 5.80 |
| 13 | 2-Cl       | 6.00       |
| 14 | 3-Cl       | 6.13       |
| 15 | 4-Cl       | 5.70       |
| 16 | 2-CH$_3$   | 6.02       |
| 18 | 3-CH$_3$   | 6.22       |
| 19 | 3,4-(CH$_3$)$_2$ | 5.70   |

*Test.
In CoMSIA, steric (S), electrostatic (E), hydrophobic (H), hydrogen bond donor (D), and hydrogen bond acceptor (A) descriptors have been calculated using the sp3 atom with charge +1 and radius 1 Å to develop different CoMSIA models. Attenuation factor $\alpha$ has been set by default at 0.3 [22]. The statistical evaluation of CoMSIA and CoMFA has been performed in the same way.

### 2.3. Partial Least Square Analysis

PLS regression analysis [23], together with cross-validation test, has been used for the generation and internal validation of the CoMFA and CoMSIA models. PLS is a powerful tool to derive multilinear relationships between independent and dependent variables. The self-consistency of models generated by PLS can be then evaluated by cross-validation test. In the current study, CoMFA and CoMSIA descriptors (explanatory properties) are used as independent variables, and $p\text{IC}_{50}$ (target properties) is used as a dependent variable. The first run of PLS analysis has been performed using the full cross-validated leave-one-out (LOO) method, which gave the cross-validated correction coefficient ($Q^2$) and the optimum number of components. In this step, sample-distance partial least square (SAMPLS) method has been used to increase the speed of cross-validation calculation [24]. In the second run of PLS, the models have been calculated by the non-cross-validation method with the optimum number of components. The models in this step have been accessed by statistical indices such as squared correlation coefficient ($R^2$), standard error of estimation (SEE), and $F$-test value [25].

### 2.4. Validation of the Models

In order to be reliable and predictive, 3D-QSAR models should (1) be statistically significant and robust, (2) be validated by making accurate predictions for external data sets that have not been used in the model development, and (3) have their application boundaries defined. In this study, five compounds have been used as a testing set [26]. The molecules used as test data are also aligned by the same methods, before determining their predictive anticancer activities, using CoMFA and CoMSIA models established by the training set. The coefficient of determination for the test set ($R^2_{\text{test}}$) is calculated by the following equation [27]:

$$R^2_{\text{test}} = 1 - \frac{\sum_{i=1}^{\text{test}} (Y_i^{\text{Obs(test)}} - Y_i^{\text{Pred(test)}})^2}{\sum_{i=1}^{\text{test}} (Y_i^{\text{Obs(test)}} - \bar{Y}_{\text{training}})^2} \quad (1)$$

In the above equation, $Y_i^{\text{Obs(test)}}$ and $Y_i^{\text{Pred(test)}}$ indicate observed and predicted values, respectively, of the test set compounds, and $\bar{Y}_{\text{training}}$ indicates the mean activity value of the training set.

### 2.5. Y-Randomization Test

The generated model was further validated by the Y-randomization method [28]. The activities of the studied molecules ($p\text{IC}_{50}$) are randomly shuffled many times, and after every iteration, a new QSAR model is developed. The new QSAR models are expected to have lower $Q^2$ and $R^2$ values than those of the original model. This technique is performed to eliminate the possibility of chance correlation. If higher values of $Q^2$ and $R^2$ are obtained, it means that an acceptable 3D-QSAR cannot be generated for this data set because of structural redundancy and chance correlation.

### 2.6. Model Acceptability Criteria

According to Alexander Tropsha and Alexander Golbraikh, if methods give a $Q^2$ value $>0.5$ and $R^2 > 0.6$, models are termed as acceptable [29].
2.7. Lipinski’s Rule and ADMET Prediction. Lipinski’s rule and ADMET [30] parameters of the newly designed compounds and the most potent inhibitor in the dataset (M17) were calculated using pkCSM [31] and SwissADME [32] web servers. Lipinski’s rule including log P, number of hydrogen-bond acceptors, number of hydrogen-bond donors, number of rotatable bonds, and molecular weight was determinate. Molecules violating more than one of these parameters may have problems with bioavailability and a high probability of failure to display drug-likeness [33].

3. Results and Discussion

3.1. CoMFA Statistical Results. In CoMFA, the obtained independent variables of the training set have been subjected to cross-validated PLS analysis to identify the value of $Q^2$ (0.608 for 3 components), and then non-cross-validated PLS analysis has been performed to obtain $SEE = 0.058$, $R^2 = 0.960$, and $F$ values $= 79.404$. For the CoMFA model, only steric and electrostatic field contributions have been calculated, and their values have been 0.717 and 0.283, respectively. The possible CoMFA results are shown in Table 2.

3.2. CoMSIA Statistical Results. In the CoMSIA study, the analysis evaluation of all 30 possible models with different field combinations represented in Table 3 can be based on five different fields including steric (S), electrostatic (E), hydrophobic (H), hydrogen-bond donor (D), and hydrogen-bond acceptor (A) [34]. Among the different field combinations, SHA model derived the highest $Q^2$ value of 0.806 with 3 components: relatively higher $R^2$ of 0.934, $F$ value of 47.29, and $SEE$ value of 0.074. Three fields have been included in the best CoMSIA model. The contributions of the steric, hydrophobic, and hydrogen-bond acceptor fields have been 0.211, 0.317 and 0.472, respectively. Hydrogen-bond acceptor and hydrophobic descriptor fields played the dominant roles. The final results of CoMFA and CoMSIA models are listed in Table 4.

Graphs of observed versus predicted biological activity for the training set and test set molecules exhibited satisfactory linear correlation, as shown in Figure 5. Therefore, the constructed CoMFA and CoMSIA models are robust and can be used to design new compounds and predict the activities of new compounds in the future.

3.3. Analysis of CoMFA and CoMSIA Contour Maps. The greatest advantage of CoMFA and CoMSIA methodologies is that the field effect on the target property can be viewed with the contour map. The contour maps of CoMFA and CoMSIA models provide information that can be used to identify important regions in the 3D space around the molecules where any change in the steric and electrostatic fields may significantly affect the biological activity. Therefore, the contour maps around compound 17 (Figure 6) that has the highest activity value are presented in pictures.

3.3.1. CoMFA Contour Map. Figure 7 shows the steric as well as electrostatic contours of the CoMFA model. Green and yellow contours denote the steric interactions. The red and blue contours denote the electrostatic ones. The steric and electrostatic fields have fractions of 71.7% and 28.3%, respectively.

The steric and electrostatic contour maps of the CoMFA study can provide information to alert which area can increase or decrease biological activity. In steric contour maps, the green contour indicated the region where bulky groups favored activity, and the yellow contour indicated the region where bulky groups have been unfavorable to activity. For electrostatic fields, blue contours indicate regions where electron-donating groups increase activity, while red contours indicate regions where electron-withdrawing groups increase activity.

---

### Table 2: Possible CoMFA results.

| Descriptors | $Q^2$ | $R^2$ | $N$ | $F$ | $SEE$ |
|-------------|------|------|----|----|-------|
| Steric (S)  | 0.637| 0.943| 3  | 55.27| 0.068 |
| Electrostatic (E) | 0.543| 0.901| 3  | 56.98| 0.090 |
| Hydrophobic (H) | 0.508| 0.932| 3  | 45.35| 0.075 |
| H-bond donor (D) | 0.406| 0.650| 3  | 06.18| 0.170 |
| H-bond acceptor (A) | 0.745| 0.872| 3  | 22.66| 0.103 |
| EA          | 0.724| 0.898| 3  | 29.29| 0.092 |
| ED          | 0.430| 0.834| 3  | 16.75| 0.117 |
| EH          | 0.487| 0.907| 3  | 32.64| 0.087 |
| ES          | 0.484| 0.931| 3  | 44.98| 0.075 |
| SA          | 0.833| 0.919| 3  | 37.57| 0.082 |
| SD          | 0.568| 0.832| 3  | 16.53| 0.118 |
| SH          | 0.584| 0.958| 3  | 75.26| 0.059 |
| AH          | 0.751| 0.911| 3  | 33.95| 0.086 |
| AD          | 0.637| 0.832| 3  | 16.86| 0.117 |
| DH          | 0.547| 0.819| 3  | 15.07| 0.122 |
| SHE         | 0.550| 0.925| 3  | 41.04| 0.079 |
| SED         | 0.501| 0.881| 3  | 24.65| 0.099 |
| SEA         | 0.779| 0.916| 3  | 36.19| 0.083 |

### Table 3: Possible CoMSIA results.

| Descriptors | $Q^2$ | $R^2$ | $N$ | $F$ | $SEE$ |
|-------------|------|------|----|----|-------|
| Steric (S)  | 0.637| 0.943| 3  | 55.27| 0.068 |
| Electrostatic (E) | 0.543| 0.901| 3  | 56.98| 0.090 |
| Hydrophobic (H) | 0.508| 0.932| 3  | 45.35| 0.075 |
| H-bond donor (D) | 0.406| 0.650| 3  | 06.18| 0.170 |
| H-bond acceptor (A) | 0.745| 0.872| 3  | 22.66| 0.103 |
| EA          | 0.724| 0.898| 3  | 29.29| 0.092 |
| ED          | 0.430| 0.834| 3  | 16.75| 0.117 |
| EH          | 0.487| 0.907| 3  | 32.64| 0.087 |
| ES          | 0.484| 0.931| 3  | 44.98| 0.075 |
| SA          | 0.833| 0.919| 3  | 37.57| 0.082 |
| SD          | 0.568| 0.832| 3  | 16.53| 0.118 |
| SH          | 0.584| 0.958| 3  | 75.26| 0.059 |
| AH          | 0.751| 0.911| 3  | 33.95| 0.086 |
| AD          | 0.637| 0.832| 3  | 16.86| 0.117 |
| DH          | 0.547| 0.819| 3  | 15.07| 0.122 |
| SHE         | 0.550| 0.925| 3  | 41.04| 0.079 |
| SED         | 0.501| 0.881| 3  | 24.65| 0.099 |
| SEA         | 0.779| 0.916| 3  | 36.19| 0.083 |

---
The most active molecule in the series (molecule 17) is displayed superimposed with CoMFA steric and electrostatic contour maps in Figure 7, respectively. In the CoMFA steric contour map (Figure 7(a)), two green contour maps are located near position 5 in the benzene ring, of compound 17, suggesting that bulky groups at this position would be favorable to increase the inhibitory activity, whereas two large yellow regions near the 2, 3, and 4 positions in the benzene ring of the same compound suggest that a bulky group at this position could decrease the activity.

Based on these observations, it is possible to explain why the activity of compounds 13 (pIC50 = 6.00), 14 (pIC50 = 6.13), 16 (pIC50 = 6.02), and 17 (pIC50 = 6.22) are higher than that of compounds 8 (pIC50 = 5.77) and 9 (pIC50 = 5.43). However, these compounds have less steric groups (Cl, Br, and CH3) in the 2, 3, and 4 positions of the benzene ring. Compounds 9 and 10, with low activities, have the bulk groups (OCH3) in the same position of the benzene ring. On the contrary, compound 1 (pIC50 = 5.89) with less steric group (Cl) in position 2 of the benzene ring is more active than compound 6 (pIC50 = 5.72), with bulk methyl group in the same position.

In the CoMFA electrostatic contour map (Figure 7(b)), the blue contour near position 3 on the benzene ring indicates the region where the generation of more negative electrostatic potential decreases the anticancer activity (pIC50). This can explain why compound 9 (pIC50 = 5.43) with methoxy in position 3 is less active than compound 4 (pIC50 = 5.80) with Br and 7 (pIC50 = 5.84) with methyl in the same position. Compound 14 (pIC50 = 6.13) with the Cl group in position 3 of the benzene ring is less active than compound 17 (pIC50 = 6.22) with the methyl group in the same position. The location of a red contour map near position 2 of the benzene ring suggests that electronegative groups at this position will increase the anticancer activity. This may explain why the activity of compounds 1 (pIC50 = 5.89) and 3 (pIC50 = 5.77), with Cl and Br groups in position 2, respectively, is greater than that of compound 7 (pIC50 = 5.72) with the methyl group in the same position.

According to the preceding analysis, the presence of electropositive groups in position 3 of the benzene ring is favorable for increased activity.

### 3.3.2. CoMSIA Contour Map

In the CoMSIA model, three different fields named as steric, hydrophobic, and hydrogen-bond acceptor fields have been analyzed. The three contour maps are displayed in Figure 8. CoMSIA steric (a), hydrophobic (b), and hydrogen-bond acceptor (c) fields give the contribution accounting for 0.211, 0.317, and 0.472, respectively.

### Table 4: Calculated data for the 3D-QSAR model.

| Model     | Q²   | R²     | SEE | F-test | N | R² test | Ster | Elec | Acc | Don | Hyd |
|-----------|------|--------|-----|--------|---|---------|------|------|-----|-----|-----|
| CoMFA     | 0.608| 0.960  | 0.058| 79.404 | 3 | 0.75    | 0.717| 0.283|     |     |     |
| CoMSIA    | 0.806| 0.934  | 0.074| 47.29  | 3 | 0.90    | 0.211|     | 0.472|     | 0.317|

![Figure 5](image-url) Scatter plot of the predicted values versus the actual pIC50 values based on CoMFA and CoMSIA models. (a) CoMFA/ES model. (b) CoMSIA/SHA model.

![Figure 6](image-url) The structure of the most active molecule (17) used in the contour analyses.
The steric contour maps in the CoMSIA model (Figure 8(a)) are more or less similar to those in the CoMFA model discussed in the CoMFA section. Therefore, the following discussion will focus precisely on the hydrophobic field and hydrogen-bond acceptor field.

Figure 8(b) shows the CoMSIA model for hydrophobic and hydrophilic groups which have been illustrated as yellow and gray contour maps, respectively. The yellow contour covering positions 3 and 4 of the benzene ring indicates that the presence of hydrophobic groups at these positions is favorable for increasing the activity, while a gray contour map covering position 2 of the benzene ring in compound 17 indicates that hydrophilic groups are favored for increasing the activity in this region. This can be interpreted by the fact that compounds 1 (pIC\textsubscript{50} = 5.89), 3 (pIC\textsubscript{50} = 5.77), and 13 (pIC\textsubscript{50} = 6.13) with huge hydrophilic group (Cl or Br) in position 2 of the benzene ring show higher activity than compound 6 (pIC\textsubscript{50} = 5.72) with the methyl group at the same position.

On the contrary, compound 19 (pIC\textsubscript{50} = 5.91) with the methyl group in position 4 of the benzene ring shows higher activity than compound 10 (pIC\textsubscript{50} = 5.43) with the methoxy group in the same position of benzene. Compound 8 (pIC\textsubscript{50} = 5.77) with the hydrophobic group in position 4 shows higher activity than compounds 2 (pIC\textsubscript{50} = 5.65), 5 (pIC\textsubscript{50} = 5.65), and 15 (pIC\textsubscript{50} = 5.70) with the hydrophilic group (Br or Cl) in the same position of the benzene ring.
For the H-bond acceptor contour maps (Figure 8(c)), the large red region converting position 2 of the benzene ring and sulfone group indicates that the presence of the hydrogen-bond acceptor is unfavored for improving the activity in this region. Conversely, the presence of H-bond donor groups in this region would increase the activity. This can be interpreted by the fact that compounds 13 (pIC\textsubscript{50} = 6.00) and 16 (pIC\textsubscript{50} = 6.02) with sulfone as the H-bond donor group show higher activity than compounds 1 (pIC\textsubscript{50} = 5.89) and 6 (pIC\textsubscript{50} = 5.72) with no sulfonic group.

According to the above analysis of CoMFA and CoMSIA contour maps, the deduced optimization strategy for improving the inhibitory activities of the compounds against HTC116 (human colon cancer) is summarized and presented in Figure 9. (Table 5)

3.4. Y-Randomization Test. Y-randomization method was carried out to validate the 3D-QSAR model. Several random scrambles of the dependent variable (pIC\textsubscript{50}) were conducted; then, after every shuffle, a 3D-QSAR was developed, and the obtained results are shown in Table 6. The low $Q^2$ and $R^2$ values obtained after every shuffle indicate that the good result in our original 3D-QSAR model is not due to a chance correlation of the training set.

3.5. Design for New Chalcone as Anticancer Agents. The foregoing study is aimed at designing of novel molecules as anticancer agents by extracting different structural features from the proposed 3D-QSAR (CoMSIA) as the best model in this study and by exploring the main structure-activity relationships derived from this study (Figure 9). It is observed that sterics and H-bond acceptor play an important role in the activity of studied compounds. This study proposes nine new molecules, which are optimized and aligned in the same way as those of the dataset compounds. The chemical structure and predicted activity values of new compounds are shown in Table 7. The newly designed compounds show good predicted activity compared with the most potent compound (molecule 17) in the dataset.

### Table 5: Anticancer activity (pIC\textsubscript{50}) of compounds in the training set and the test set for the final CoMFA and CoMSIA models.

| No | pIC\textsubscript{50} (obs) | pIC\textsubscript{50} (pred) |
|----|---------------------------|---------------------------|
|    | CoMFA-ES Residu | CoMSIA-SHA Residu | CoMFA-ES Residu | CoMSIA-SHA Residu |
| 1  | 5.89 | 5.79 | 0.10 | 5.76 | 0.13 |
| 2  | 5.65 | 5.71 | 0.06 | 5.61 | 0.04 |
| 3  | 5.77 | 5.80 | 0.03 | 5.77 | 0.00 |
| 4  | 5.80 | 5.79 | 0.01 | 5.80 | 0.00 |
| 5  | 5.65 | 5.65 | 0.00 | 5.70 | 0.05 |
| 6  | 5.72 | 5.77 | 0.05 | 5.76 | 0.04 |
| 7  | 5.84 | 5.65 | 0.19 | 5.52 | 0.32 |
| 8  | 5.77 | 5.79 | 0.02 | 5.78 | 0.01 |
| 9  | 5.43 | 5.41 | 0.02 | 5.44 | 0.01 |
| 10 | 5.31 | 5.31 | 0.00 | 5.30 | 0.01 |
| 11* | 5.78 | 5.79 | 0.01 | 5.81 | 0.03 |
| 12* | 5.80 | 5.78 | 0.02 | 5.80 | 0.00 |
| 13 | 6.00 | 6.05 | 0.05 | 6.05 | 0.05 |
| 14 | 6.13 | 6.18 | 0.05 | 6.11 | 0.02 |
| 15* | 5.70 | 5.91 | 0.21 | 5.99 | 0.29 |
| 16 | 6.02 | 5.94 | 0.08 | 6.03 | 0.01 |
| 17 | 6.22 | 6.17 | 0.05 | 6.09 | 0.15 |
| 18* | 5.70 | 5.81 | 0.11 | 6.01 | 0.31 |
| 19 | 5.91 | 5.88 | 0.03 | 6.02 | 0.11 |

*Test.

3.6. Lipinski’s Rule and ADMET Prediction. Lipinski’s rule and ADMET parameters of the newly designed compounds and the most potent inhibitor in the dataset (M17) were calculated using pkCSM and SwissADME web servers. Lipinski’s rule including log P, number of hydrogen-bond acceptors, number hydrogen-bond donors, number of rotatable bonds, and molecular weight is shown in Table 8.

All the compounds respect the conditions mentioned in Lipinski’s rule, except two compounds 4 and 8 with two Lipinski violations. The ADMET prediction was used in this study to calculate the pharmacokinetic parameters (PK) of compounds (Table 9).

The results of ADMET analysis show that all the newly designed compounds have low values for the blood-brain partition coefficient, indicating that they will have a very low potential to cross the blood-brain barrier thereby eliminating the possibility of CNS-related toxicity. Additionally, other pharmacokinetic parameters such as...
Table 6: $Q^2$ and $R^2$ values after several Y-randomization tests.

| Iteration | CoMFA | CoMSIA |
|-----------|-------|--------|
| $Q^2$     | $R^2$ | $Q^2$  | $R^2$ |
| 1         | 0.148 | 0.774  | -0.060 | 0.750 |
| 2         | -0.194| 0.819  | 0.076  | 0.849 |
| 3         | -0.157| 0.709  | -0.314 | 0.724 |
| 4         | 0.302 | 0.822  | 0.111  | 0.765 |
| 5         | -0.259| 0.699  | -0.217 | 0.698 |
| 6         | 0.120 | 0.736  | 0.001  | 0.732 |
| 7         | 0.131 | 0.718  | 0.037  | 0.709 |

Table 7: Structures and their predicted $pIC_{50}$ based on the 3D-QSAR (COMSIA) model.

| N° | Structure | $pIC_{50}$ (pred) CoMSIA/SHA |
|----|-----------|------------------------------|
| 1  | ![Structure 1](image1.png) | 6.568                        |
| 2  | ![Structure 2](image2.png) | 6.567                        |
| 3  | ![Structure 3](image3.png) | 6.547                        |
| N° | Structure | $pIC_{50}$ (pred) | CoMSIA/SHA |
|----|-----------|-------------------|------------|
| 4  | ![Structure Image](image1) | 6.547            |            |
| 5  | ![Structure Image](image2) | 6.521            |            |
| 6  | ![Structure Image](image3) | 6.515            |            |
| 7  | ![Structure Image](image4) | 6.506            |            |
Table 7: Continued.

| N° | Structure | $pIC_{50}$ (pred) | CoMSIA/SHA |
|----|-----------|------------------|------------|
| 8  | ![Structure](image1.png) | 6.495 | |
| 9  | ![Structure](image2.png) | 6.488 | |
| 10 | ![Structure](image3.png) | 6.468 | |

Table 8: Lipinski’s rule of newly designed inhibitors and the most potent inhibitor in the dataset.

| Compounds | Log $P$ | H-bond acceptor | H-bond donor | Rotatable bonds | Molecular weight (g/mol) | Lipinski violations |
|-----------|---------|-----------------|--------------|-----------------|--------------------------|-------------------|
| Rule      | <5      | ≤10             | <5           | <10             | ≤500                     | ≤1                |
| M17       | 4.57    | 4               | 2            | 6               | 417.49                   | 0                 |
| 1         | 3.23    | 8               | 5            | 9               | 639.30                   | 1                 |
| 2         | 3.89    | 8               | 5            | 9               | 673.97                   | 1                 |
| 3         | 3.13    | 8               | 5            | 9               | 595.08                   | 1                 |
| 4         | 3.90    | 9               | 5            | 10              | 733.44                   | 2                 |
| 5         | 3.23    | 8               | 5            | 9               | 639.53                   | 1                 |
| 6         | 4.77    | 8               | 5            | 9               | 811.35                   | 1                 |
| 7         | 4.77    | 8               | 5            | 9               | 811.35                   | 1                 |
| 8         | 4.26    | 8               | 5            | 10              | 732.45                   | 2                 |
| 9         | 4.95    | 8               | 5            | 9               | 712.44                   | 1                 |
| 10        | 2.47    | 8               | 5            | 9               | 560.63                   | 1                 |
human intestinal absorption (HIA), water solubility (log mol/L), metabolism, AMES toxicity, and synthetic accessibility are all acceptable.

4. Conclusions

The construction of the 3D-QSAR model has been the objective behind the collection, optimization, and calculation of a series of chalcone derivatives acting as anticancer agents against HTC116 (human colon cancer). Thus, the internal and external predictive powers of the obtained CoMFA and CoMSIA models are satisfactory ($Q^2$: 0.608, 0.806, respectively, and $R^2$: 0.960, 0.934, respectively). The correlation between different fields and activity is well previewed by the contour maps of CoMFA and CoMSIA models. This has been a reference for the design of new compounds. Thus, compounds that obtained key structural factors have been used to design ten inhibitors by modifying the chalcone scaffold. The overall result of this generated work gives the possibility to design and synthesize new chalcone derivatives with potential inhibitory activity to treat HTC116 (human colon cancer) and then obtain a better result in terms of biological activity, in silico ADME, and toxicity.

Data Availability

All the data are included within the manuscript.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the Natural Substances and Molecular Chemistry Laboratory (SOOM), Department of Chemistry, Faculty of Science, Moulay Ismail University of Meknes, Morocco, and Materials, Environment and Modeling (MEM), LASMAR Laboratory, ESTM, Moulay Ismail University of Meknes, Morocco.

References

[1] H. Varmus, “The new era in cancer research,” Science, vol. 312, no. 5777, pp. 1162–1165, 2006.
[2] S. Reinberg, Health Day News, World Health Organization, Geneva, Switzerland, 2008.
[3] G. Carlo Di, N. Mascolo, A. A. Izzo, and F. Capasso, “Flavonoids: old and new aspects of a class of natural therapeutic drugs,” Life Sciences, vol. 65, no. 4, pp. 337–353, 1999.
[4] C. Cells, Y. C. Mayur, G. J. Peters, V. V. S. R. Prasad, C. Lemos, and N. K. Sathish, “Design of new drug molecules to be used in reversing multidrug resistance in cancer cells,” Current Cancer Drug Targets, vol. 9, no. 3, pp. 298–306, 2009.
[5] V. R. Solomon, C. Hu, and H. Lee, “Hybrid pharmacophore design and synthesis of isatin-benzothiazole analogs for their anti-breast cancer activity,” Bioorganic & Medicinal Chemistry, vol. 17, no. 21, pp. 7585–7592, 2009.
[6] P. Singh, A. Anand, and V. Kumar, “Recent developments in biological activities of chalcones: a mini review,” European Journal of Medicinal Chemistry, vol. 85, pp. 758–777, 2014.
[7] H. Yi, K. Kaan, J. Weiss et al., “Structure-activity relationship and multidrug resistance study of new S-trityl-L-cysteine derivatives as inhibitors of eg5,” Journal of Medicinal Chemistry, vol. 54, pp. 1576–1586, 2011.
[8] S. Nagarajan, D. A. Skoufias, F. Kozielski, and A. N. Pae, “Receptor-ligand interaction-based virtual screening for novel eg5/kinesin spindle protein inhibitors,” Journal of Medicinal Chemistry, vol. 55, no. 6, pp. 2561–2573, 2012.
[9] C. Jiang, Y. Chen, and X. Q. Wang, “Docking studies on kinesin spindle protein inhibitors: an important cooperative "minor binding pocket" which increases the binding affinity significantly,” Journal of Molecular Modeling, vol. 13, no. 9, pp. 987–992, 2007.
[10] H. You, 3D QSAR in Drug Design: Theory Methods and Application, Springer Science and Business Media, Berlin, Germany, 1993.
[11] R. D. Cramer, J. D. Bunce, D. E. Patterson, and I. E. Frank, “Crossvalidation, bootstrapping, and partial least squares compared with multiple regression in conventional QSAR studies,” Quantitative Structure-Activity Relationships, vol. 7, no. 1, pp. 18–25, 1988.
[12] G. Klebe and U. Abraham, “Comparative molecular similarity index analysis (CoMSIA) to study hydrogen-bonding properties and to score combinatorial libraries,” *Journal of Computer-Aided Molecular Design*, vol. 13, no. 1, pp. 1–10, 1999.

[13] G. Klebe, U. Abraham, and T. Mietzner, “Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity,” *Journal of Medicinal Chemistry*, vol. 37, no. 24, pp. 4130–4146, 1994.

[14] Y. Wang, S. Xue, R. Li, Z. Zheng, H. Yi, and Z. Li, “Synthesis and biological evaluation of novel synthetic chalcone,” *Bioorganic & Medicinal Chemistry*, vol. 26, no. 1, pp. 8–16, 2017.

[15] D. K. Mahapatra and V. S. K. Asati, “Chalcones and their therapeutictargetsforthe managementofdiabetes:structural and pharmacological perspectives,” *European Journal of Medicinal Chemistry*, vol. 92, pp. 839–865, 2015.

[16] L. Bharti, A. Belhassan, A. Aouidate, A. Ghaleb, T. Lakhlfii, and M. Bouachrine, “QSAR study of new compounds based on 1,2,4-triazole as potential anticancer agents,” *Physical Chemistry Research*, vol. 8, no. 1, pp. 125–137, 2020.

[17] M. Clark, R. D. Cramer, and N. Van Opdenbosch, “Validation of the general purpose tripos 5.2 force field,” *Journal of Computational Chemistry*, vol. 10, no. 8, pp. 982–1012, 1989.

[18] W. P. Purcell and J. A. Singer, “A brief review and table of semiempirical parameters used in the Hueckel molecular orbital method,” *Journal of Chemical & Engineering Data*, vol. 12, no. 2, pp. 235–246, 1967.

[19] Tripos Inc., St. Louis, MO, USA, SYBYL-X 2.0, (n.d.). http://www.tripois.com.

[20] M. D. M. Abdulhameed, A. Hamza, J. Liu, and C.-G. Zhan, “Combined 3D-QSAR modeling and molecular docking study on indolinone derivatives as inhibitors of 3-phosphoinositide-dependent protein kinase-1,” *Journal of Chemical Information and Modeling*, vol. 48, no. 9, pp. 1760–1772, 2008.

[21] R. D. Cramer, D. E. Patterson, and J. D. Bunce, “Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins,” *Journal of the American Chemical Society*, vol. 110, no. 18, pp. 5959–5967, 1988.

[22] J. Zheng, G. Xiao, J. Guo et al., “Exploring QSARs for 5-lipoxygenase (5-LO) inhibitory activity of 2-substituted 5-hydroxyindole-3-carboxylates by CoMFA and CoMSIA,” *Chemical Biology & Drug Design*, vol. 78, no. 2, pp. 314–321, 2011.

[23] S. Wold, “Validation of QSAR’s,” *Quantitative Structure-Activity Relationships*, vol. 10, no. 3, pp. 191–193, 1991.

[24] B. L. Bush and R. B. Nachbar, “Sample-distance partial least squares: PLS optimized for many variables, with application to CoMFA,” *Journal of Computer-Aided Molecular Design*, vol. 7, no. 5, pp. 587–619, 1993.

[25] L. El Mchichi, A. Aouidate, F. Z. Chokrafi, A. Ghaleb, T. Lakhlfii, and M. Bouachrine, “Prediction of biological activity of pyrazolo [3, 4-b] quinolinyl acitamide by qsar results,” *RHAZES: Green and Applied Chemistry*, vol. 3, pp. 79–93, 2018.

[26] A. Golbraikh and A. Tropsha, “Beware of q2!,” *Journal of Molecular Graphics and Modelling*, vol. 20, no. 4, pp. 269–276, 2002.

[27] K. Roy and S. Kar, *A Primer on QSAR/QSPR Modeling*, Springer, London, UK, 2015.

[28] C. Ru and G. Ru, “y-Randomization and its Variants in QSAR/QSAR,” *Journal of Chemical Information and Modeling*, vol. 47, no. 6, pp. 2345–2357, 2007.