Journal of Saudi Chemical Society (2023) 27, 101728

King Saud University

Journal of Saudi Chemical Society

www.ksu.edu.sa
www.sciencedirect.com

ORIGINAL ARTICLE

QSAR, DFT studies, docking molecular and simulation dynamic molecular of 2-styrylquinoline derivatives through their anticancer activity

Sara Zarougui [a], Mohammed Er-rajy [a], Abdelmoujoud Faris [a], Hamada Imtara [b], Mohamed El fadili [a], Omkulthom Al kamaly [c,*], Samar Zuhair Alshawwa [c], Fahd A. Nasr [d], Mourad Aloui [a], Menana Elhallaoui [a]

[a] *LIMAS Laboratory, Faculty of Sciences Dhar El Mahraz, Sidi Mohamed Ben Abdellah University, Fez, Morocco*
[b] *Faculty of Sciences, Arab American University Palestine, P.O. Box 240, Jenin 44862, Palestine*
[c] *Department of Pharmaceutical Sciences, College of Pharmacy, Princess Nourah bint Abdulrahman University, P.O Box 84428, Riyadh 11671, Saudi Arabia*
[d] *Biology Department, College of Science, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh 11623, Saudi Arabia*

Received 26 March 2023; revised 24 August 2023; accepted 10 September 2023
Available online 15 September 2023

KEYWORDS

Anti-cancer;
2-Styrylquinoline derivatives;
2D-QSAR;
DFT reactivity;
Molecular Docking;
Simulation dynamic molecular

Abstract In this study, a 2D-QSAR (quantitative structure–activity relationship) was performed on 54 new 2-Styrylquinoline derivatives as anticancer substances capable of inhibiting the p53 protein in the cell HCT116^{++}. The 54 2-Styrylquinoline derivatives was calculated applying DFT 6-31G basis to calculate Quantum descriptors, using MM2 for: Topological, Physico-chemical, Geometrical and Constitutional. The study was carried out by performing multiple linear regression (R^2 = 0.90), the QSAR model achieved was tested by artificial neural networks method, which is showed high predictability (R^2_{ANN} = 0.89). A DFT study was performed to determine the reactivity of the 2-Styrylquinoline derivatives using frontier molecular orbital analysis and analysis of the molecular electrostatic potential (MEP). Derivatives of 2–4 Styrylquinoline are studied for their synthetic accessibility and their similarity to drug. The obtained results show that all the evaluated compounds have similar properties to drug and are accessible to synthesize.

A molecular docking analysis was performed for three compounds: 14, 34, and 54, having various reactivities against the p53 HCT116^{++} protein (identified by PDB ID: 2GEQ). The results showed strong interactions between the three ligands and the 2GEQ protein, the amino acids HIS 176, SER A180, PRO A188 and ARG A178 are the most active sites of the 2GEQ protein, and based on these result we performed a molecular dynamics simulation to evaluate the stability of our complexes. The MD demonstrates the thermodynamic stability of select compounds during 40 and 100 ns, with all three complexes showing a high level of structural stability.

* Corresponding author.
E-mail address: omalkmali@pnu.edu.sa (O. Al kamaly).

https://doi.org/10.1016/j.jscs.2023.101728

1. Introduction

In the early 2000 s, almost 6.72 million people died worldwide from malignant tumors, and almost 10 million people, both men and women, contracted a serious tumor. According to statistics from the World Health Organization, cancer is one of the leading causes of death in the world [1]. Cancer rates have increased by 50% to reach 15 million new cases in 2020, according to the World Cancer Report. In the case of colon cancer, more than 940,000 cases are reported each year worldwide, and nearly 500,000 people die from it each year [1]. Cancer is characterized by the rapid, uncontrolled, and pathological multiplication of abuses cells. The p53 mutation forms a disjunctive advantage for tumor cells, given its essential activities for the cells, which expresses its important role in transforming a benign tumor into cancer [2]. Moreover, p53 gene mutations are responsible for 50% of tumors (cancer cells) [3]. Currently, most of the treatments used to beat the cancer cells agree with chemotherapy, radiotherapy, surgery, or even hormone therapy. Despite the widespread use of these drugs in the treatment of different types of cancer, they have several disadvantages, such as non-typical targeting of cancer cells, the toxicity of treatments … etc [4].

Due to their basic structures, heterocyclic molecules have great exploitation in pharmacology, biology, etc [5]. 2-Styrylquinoline derivatives are bioactive aromatic compounds that show activity and antiviral properties in HIV-1 EMC cell lines [6]. Moreover, 2-Styrylquinoline derivatives present a poly-activity in pharmacology by possessing a multitude of activities such as, antimicrobial [7], anti-parasitic [8], anti-fungal [9], etc.

To discover new compounds that specifically target the cancer cell, which requires an extensive investment, besides saving a lot of energy and time, there are several methods that enter the process of computerized drug design, such as quantitative structure-Activity relationship (QSAR), molecular docking, and molecular dynamic simulation. We simulated the anti-carcinogenic activity of 54 2-Styrylquinoline derivatives [10] evaluated according to their biological activities on wild-type p53 HCT116^{++} (exhibiting a p53 deletion), using the methods: QSAR (multiple linear regression (MLR), artificial neural network (ANN), applicability domain, Y-randomization); DFT study: to study the reactivity of 2-styrylquinoline derivatives in order to predict the minimum stability of compounds [11]. Lipinski rules and a synthetic accessibility study were used to predict the pharmacokinetic characteristics of 2-styrylquinoline derivatives proposed as the studied colon cancer inhibitors [12,13].

The molecular docking method helps for a better understanding of the interactions between the proposed inhibitor compound and the recipient protein 53. In this paper, three molecules were docked to determine the most stable molecules using molecular dynamic simulation (MD) [14,15].

2. Materials and methods

2.1. Experimental data set

We simulated 54 of 2-Styrylquinoline derivatives (Fig. 1) synthesized by Anna Mrozek-Wilczkiewicz, with various substitutions present in Table 1, divided into 2 groups; 11 molecules were tested, and a set of 43 derivatives was used to build the model. The conversion of all experimental activity values IC_{50} en pIC_{50}(μm) was according to the following equation:

$$pIC_{50} = 6 - log_{10}(IC_{50})$$

2.2. Calculations of descriptors

To establish a reliable QSAR model, we calculated more than 40 molecular descriptors of the studied series, grouped by class (1D, 2D, 3D). Table 2 shows the types of descriptors calculated using the software: ACD /ChemSketch [16], ChemOffice [17], Gaussian 09 Software [18]. The descriptors of physicochemical, topological, geometrical, and constitutional type are optimized using the MM2 method, and the optimized electronic descriptors are obtained by the Gaussian software using the DFT/ 6-31G/ B3LYP basis, including HOMO energies (highest occupied molecular orbital), LUMO energies (lowest unoccupied molecular orbital), repulsion energy E_r, dipole moment MD, total energy E_t. The selected descriptors and their classes are represented in Table 2.

2.3. Statistical method

The nine studied 2-Styrylquinoline derivatives synthesized in vitro show potent inhibitory activity (pIC_{50} greater than 4.881) on colon cancer cells of the HCT116^{++}. To establish the QSAR models we used the statistical approaches described below (see Table 3).

2.3.1. K-means analysis

The K-means method is a non-hierarchical method, and the classification of data according to this method is often divided according to groups and incidentally [19]. After determining the most interesting molecular descriptors, we divided the data into 2 sets (training and testing); these sets contain respectively 80% and 20% of the total molecular descriptors [19]. The method of dividing the series into two sets is performed according to the K-means 50 classifications; each of the two clusters includes a set of molecules that are arbitrarily chosen in order to build a set of test compounds, while the other molecules will be classified as a training set. After this classification, we find eleven compounds in the test cluster, forty-three compounds in the training set.

2.3.2. Construction of model

The statistical methods MLR and ANN are tools used to elaborate the QSAR model in order to illustrate and explain the structure–activity relationship of 20 descriptors calculated for the 54 molecules of 2-Styrylquinoline derivatives.

2.3.3. The linear regression

The criteria for linear regression to identify the relationship between the dependent and independent variables are: The determination correlation R^2, the mean square error (RMSE), the value of Fisher ratio (F_{Test}). The MLR equation can be written in the following form:

$$Y = b_0 + \sum_{i=1}^{n} biXi \quad (1)$$

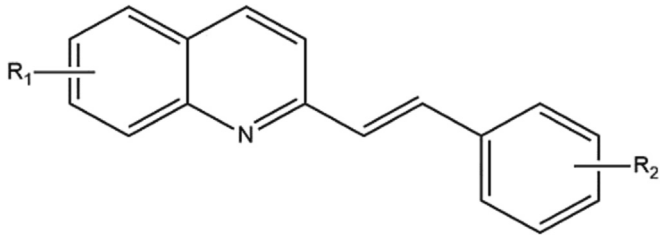

Fig. 1 The general structure of quinoline derivatives.

Table 1 Chemical structure and their experimental biological activity for the 54 derivatives.

HCT116 ($P53^{+/+}$)									
N°	R_1	R_2	IC_{50}	pIC_{50}	N°	R_1	R_2	IC_{50}	pIC_{50}
1	8-OAc	2-OC H_3	13.13	4.881	28	5,7-Cl-8-OAc	2,6-F-3-Cl	7.26	5.139
2	8-OAc	3-OC H_3	9.70	5.013	29	5,7-Cl-8-OAc	2,6-F-4-Cl	20.90	4.679
3	8-OAc	4-OC H_3	11.65	4.933	30	5,7-Cl-8-OAc	2-CN	0.93	6.031
4	8-OAc	3,4-OC H_3	16.84	4.773	31	5,7-Cl-8-OAc	3-CN	1.39	5.856
5	8-OAc	2-OEt	20.26	4.693	32	5,7-Cl-8-OAc	4-CN	1.78	5.749
6	8-OAc	4-OEt	20.87	4.680	33*	5,7-Cl-8-OAc	2-OAc-5-N O_2	2.71	5.322
7	8-OAc	3-OAc	5.40	5.267	34*	5,7-Cl-8-OAc	2-OAc-3-N O_2	4.76	5.567
8	8-OAc	2-Cl	10.63	4.973	35	5,7-Cl-8-OAc	3-NO2-4-OAc	4.22	5.374
9	8-OAc	3-Cl	14.47	4.839	36	5,7-Cl-8-OAc	2-N O_2	0.85	6.070
10	8-OAc	2-NO_2	4.60	5.337	37	5,7-Cl-8-OAc	3-N O_2	1.74	5.759
11*	8-OH	4-OH	13.53	4.868	38	5,7-Cl-8-OAc	4- N O_2	2.31	5.636
12	8-OH	2-OAc	7.32	5.135	39	5,7-Cl-8-OAc	2,4-N O_2	0.28	6.552
13	8-OH	4-OAc	12.36	4.907	40	5,7-Cl-8-OH	2-OH	10.48	4.979
14	8-OH	2-Ac-4Cl	19.92	4.70	41	5,7-Cl-8-OH	2-OAc-3,5-Cl	4.09	5.388
15*	8-OH	2-N O_2	11.72	4.931	42	5,7-Cl-8-OH	2-OH-3-Br-5-Cl	3.28	5.484
16*	5,7-Cl-8-OAc	2-OAc	11.20	4.950	43	5,7-Cl-8-OH	2-Cl	1.88	5.725
17*	5,7-Cl-8-OAc	2-OAc-5-Cl	13.85	4.858	44	5,7-Cl-8-OH	2,3-Cl	5.27	5.278
18	5,7-Cl-8-OAc	2-OAc-3,5-Cl	3.65	5.437	45	5,7-Cl-8-OH	2-Cl-6-F	7.30	5.136
19	5,7-Cl-8-OAc	2-OAc-3-Br-5-Cl	3.83	5.416	46	5,7-Cl-8-OH	2-I	3.29	5.482
20	5,7-Cl-8-OAc	2-OAc-3,5-I	5.13	5.289	47	5,7-Cl-8-OH	2-CN	0.86	6.065
21	5,7-Cl-8-OAc	2-I	1.20	5.920	48*	5,7-Cl-8-OH	3-CN	8.51	5.070
22	5,7-Cl-8-OAc	2-Cl	2.22	5.653	49	5,7-Cl-8-OH	4-CN	3.10	5.508
23*	5,7-Cl-8-OAc	2,6-Cl	5.93	5.226	50	5,7-Cl-8-OH	3-N O_2-4-OAc	3.63	5.440
24	5,7-Cl-8-OAc	2-Cl-6-F	3.41	5.467	51	5,7-Cl-8-OH	2-N O_2	0.73	6.136
25*	5,7-Cl-8-OAc	2-Br-6-F	3.70	5.431	52*	5,7-Cl-8-OH	3-N O_2	2.77	5.557
26	5,7-Cl-8-OAc	2,5-F	2.57	5.590	53	5,7-Cl-8-OH	4-N O_2	6.48	5.188
27	5,7-Cl-8-OAc	2,6-F	4.93	5.307	54*	5,7-Cl-8-OH	2,4-N O_2	0.75	6.124

*Test set.

Table 2 The list of the calculated descriptors.

Descripteurs	Classe
E_{homo},E_{lumo} dipôle moment (MD), E_r,E_t.	Quantum
Balaban index (BI), Winer index(WI), Polar surface area(PSA), molecular topology(MT)	Topological
Number of H-bond donor (NHD), number of H-bond acceptor (NHA), Boiling point (BP), LogP....	Physico-chemical
Energy of Van der Waals(VDW), molecular Volume(MV)	Geometrical
Molecular weight (MW)	Constitutional

Table 3 The classification K-means.

K-means classification results	
Classe 1	1 3 8 9 11 12 13 14 15 40 42 43 44 45 46 47 48 52
Classe 2	2
Classe 3	4 7
Classe 4	5 6 23 24 25 26 27 32 33
Classe 5	10 17 21 22
Classe 6	16 18 19 20 34 35 36 37 41 50 54
Classe 7	28 29
Classe 8	30
Classe 9	31
Classe 10	51
Classe 11	53

With Y, the predicted value of the biological activity, X_i the molecular descriptors (independent variables), n: the number of molecular descriptors and b_0 the constant of the equation.

2.3.4. Neural networks artificial ANN

The neural network is a non-linear method linking the descriptors of the MLR method with the predicted biological activity of the ANN technique. We elaborate the neural network based on the model found by the multiple linear regressions in order to validate the accuracy of the descriptors obtained [20]. This method uses a sigmoidal transport method in the hidden layer and a linear transfer method in the output layer, and in this framework, the ANN architecture considered consists of 3 layers of neurons: the input layer passing through the hidden layer, and finally the output layer, as described in Fig. 2. The input layer includes neurons with the same numbers as the descriptors found by the RLM method, and the output layer includes the biological activities predicted by the ANN method. The neural network was realized with JMP Pro14 [21].

2.3.5. Y-randomization

The Y-randomization method is a test used to eliminate the arbitrary correlation of the descriptors obtained by the linear regression method, the neural network, and their biological activities in order to confirm the reliability and efficiency of the ANN and MLR tests by determining the random correlation between the descriptors and the biological activity. The models obtained by the Y-randomization test are the results of the random distribution of the experimental activity results to the descriptor origins of the model, and therefore the models appear. Moreover, we can confirm the QSAR model if the correlation coefficient R^2 found by the RLM method is higher than the R^2 found by the Y-Randomization test. The Y-Randomization test is performed using Matlab Software [22].

2.3.6. Domain of applicability

The applicability domain is one of the most used methods in QSAR; it is defined as a response area of the MLR model (pIC_{50}) and all variables (molecules) test and training [23]. Among the most used method to establish the DA of QSAR model, the Leverage effect approach characterized by the critical value h*, this value is calculated by the following relation: 3*(d + 1)/n, with n: the number of molecules training and d: the number of descriptors. A molecule (i) is estimated outside the applicability domain when the Lever effect criterion (h_i) is higher than h* (critical value). The DA approach was performed using Minitab19 [24].

2.4. Docking molecular

Molecular docking is one of the most widely used methods in QSAR studies, based on the simulation of interactions between the proposed compound and the receptor protein. We obtained the receptor protein through the protein database (PDB: code 2EGQ) [25] of https:// www. rcsb. org/, whose PDB form of the selected ligands was converted from Open Babel GUI software [26]. The preparation of the protein for the docking step such as: removal of water molecules, addition of charge, and hydrogen [27,28]. Then we do perform the docking operation in order to predict the ligand–protein interactions are done from AutoDock Tools software [29]. We performed molecular docking on compounds with interesting reactivity according to DFT study, and for visualization of docked complexes, we used Discovery studio 2020 software [30] for visualization of ligands and their interactions with proteins.

2.5. Synthetic accessibility and Lipinski rules

Numerical simulation methods offer an integrated, fast and affordable approach to determine the studied molecules capable of acting as drug agents if they have met certain overriding rules: The rules of Lipinski [31], the rules of Veber and Johnson rules [32], Egan and Merz rules [33], Ghose rule [34], synthetic accessibility [35].

In this work, we have determined the characteristics of similarities of some compounds that have experimentally demonstrated an important activity so that these molecules enter the application field of the study. These rules are interesting to describe the 2D chemical structure of proposed molecules and on the bioactivity of these compounds by ingestion through the route [28]. The Lipinski, Ghose, Egan, and Veber rules are established to evaluate the drug-like properties of small molecule compounds. These rules are widely used in drug discovery and development to predict the likelihood of a compound's success as a drug candidate. Compounds with physical and chemical characteristics that do not meet at least two of the Lipinski, Ghose, and Egan and Veber rules are responsible for a multitude of disorders in their pharmacokinetic characteristics. Moreover, more than 90% of drugs that reach the stage of clinical trials meet these rules. In this approach, we study the pharmacokinetic characteristics of compounds that will be selected as inhibitors of the activity of the protein p53HCT116++ using the SwissADME server [36].

2.6. Molecular dynamic simulation

MD simulation is a useful procedure to reveal the thermodynamic stability of the ligand in the macromolecular site and the contribution of key amino acids in the ligand–protein interaction. We have chosen based on molecular docking results, the ligands that will be subject the dynamic simulation. These ligands present an interesting affinity and more efficient interactions with the receptor (protein); the dynamic simula-

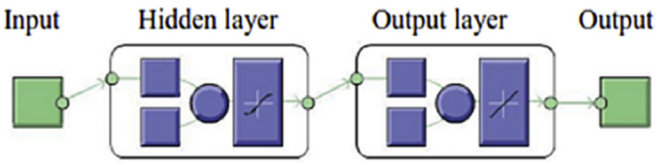

Fig. 2 The architecture of model neural network.

tion performed using the CHARMM-GUI web server with the CHARMM36 force field [37]. The Swiss PARAM web server generated the ligand topology [38]. The system was solvated in a box of TIP3P water molecules) with a concentration of 0.15 M NaCl using the default box size. The temperature was set to 300 K and the pressure was maintained at one atm using the Parrinello-Rahman barostat with a relaxation time of two ps. The initial configuration was generated using the CHARMM-GUI Membrane Builder module, which embeds the ligand in a pre-equilibrated lipid bilayer [39]. The system was then equilibrated for five ns using the NVT ensemble with a time step of 1 fs, followed by 10 ns of NPT equilibration with a time step of 2 fs. The production run was performed using the GROMACS software [40] with the CHARMM36 force field. The simulations were carried out for 100 ns with a time step of 2 fs using the leapfrog integrator. The temperature and pressure were maintained at 300 K and 1 atm, respectively, using the v-rescale thermostat [41] and the Parrinello-Rahman barostat with a relaxation time of two ps.

3. Results and discussion

3.1. Dataset for analysis

The QSAR analysis was carried out according to several trials performed on 54 molecules, based on the calculation of descriptors whose values are presented in Table 4. Subsequently we divided the 54 molecules analyzed into two classes using the K-means method, the training class contains 43 molecules, and 11 molecules for the test set. After successive attempts on the data using multiple linear regression several times until we obtained a reliable model, eliminating the descriptors with high correlation and retaining those with low correlation, whose descriptors obtained by the MLR method are used as input factors in the artificial neural network method. The pIC_{50} predicted by the MLR and ANN methods are presented in Table 5.

3.2. Multiple linear regressions RLM

Given its ease, representativeness and traceability, multiple linear regressions are the most widely used method in 2D-QSAR analysis. This method is based on three parameters: coefficient of determination (R square), Fisher ratio (F), and root mean square error (RMSE). we performed the multiple linear regression analysis using XLSTAT 2019 software [42].

The obtained QSAR model of the training class is presented by Eq. (2), whose corresponding normalized descriptor coefficients (Fig. 4).

$$pIC_{50} = -1,58 - 17,53E_{homo} - 18,86E_{lumo} - 9,51DM + 5,95SB - 3,11BI + 2,56PSA + 1,10BP \quad (2)$$

The statistical parameters of the equation:

N = 42 R^2 = 0.78 $R_{ajusté}$ = 0.73 F_{test} = 16.70 RMSE = 0.24

The QSAR model obtained is based on seven descriptors: Energy Homo, Energy LUMO, Dipole moment, Stretch bend, Balaban index, Polar surface area, Boiling point. According to the coefficient normalization diagram (Fig. 3), the QSAR model is constructed of five most important descriptors correlated with pIC_{50}: E_{HOMO}, BI, PSA, BP, SB, with a high correlation coefficient 0.76. The most influential coefficient in the QSAR model is PSA, which the larger the coefficient value, the higher the activity of the compound.

3.3. Artificial neural networks

The progress of the QSAR model requires the use of the non-linear method between the biological anti-cancer activity obtained and the seven descriptors obtained by MLR, whose ANN architecture is 7–3-1 presented in Fig. 4:

The results show a strong correlation between the observed biological activities and the activities predicted by ANN are presented in (Fig. 5), the coefficient of determination R^2 = 0.89 and RMSE = 0.17, which confirms well the performance of static results obtained by the MLR method (see Fig. 6).

3.4. Y randomization

The new random QSAR models correspond to the new values of Q^2 and R for each new model. The results found are shown in Table 6. We found that the correlation coefficient of the randomized models (R = 0.39) is lower than that found by the multiple linear regression (non-randomized models) R = 0.87 which confirms that the model is acceptable, moreover Rp^2 = 0.67 greater than 0.5 which confirms the robustness of the model developed in the study. Therefore the predicted biological activity values pIC_{50} obtained from the MLR method based on the seven descriptors are reliable which explains that the interdependence between the seven descriptors and the predicted activity is ensured; therefore the biological activity is not randomized.

3.5. Applicability domain

The domain of applicability is based on the QSAR model found using the MLR method, i.e., determining the substances outside the domain of applicability. The latter is based on the Williams diagram, which is represented by a graph whose X axis is determined by the normalized residual values and the Y abscissa represents the leverage values Fig. 7. The limit value of lever efficiency ish* = 0.55, all compounds that exceeded on the right the lever effect 0.55 are not acceptable, i.e., their activity pIC_{50} is incorrectly predicted.

3.6. DFT calculation studies

In quantum chemistry, one of the most reliable and efficient methods is the DFT method (Density Function Theory). To properly analyze the chemical structure and reactivity of the new 2-styrylquinoline derivatives as well as their stability, we performed a DFT calculation based on the Becke-3-parameter-Lee-Yang-Parr (B3LYP-631G) theory for 54 derivatives. We optimized the geometry of the studied molecules with Gaussian software 09 W and visualized them with GaussView 6 using the B3LYP-6-31G basis [43,44].

Table 4 The descriptors calculated for the 54 molecules.

N°	E^{a}_{homo}	E^{b}_{lumo}	DM^{c}	SB^{d}	BI^{e}	PSA^{f}	BP^{j}	E^{i}_{t}	$LogP^{h}$	pIC_{50}
1	−0.205	−0.071	2.888	0.139	455,731	68.12	760.348	−1088.3	4.514	4.881
2	−0.196	−0.066	3.338	0.129	465,201	68.12	760.348	−1088.3	4.514	5.013
3	−0.206	−0.069	2.448	0.121	455,731	68.12	760.348	−1088.3	4.514	4.933
4	−0.199	−0.069	4.493	0.121	691,287	77.35	784.001	−1202.8	4.388	4.773
5	−0.204	−0.07	2.919	0.163	559,981	68.12	771.952	−1127.6	4.852	4.693
6	−0.205	−0.068	2.262	0.145	559,981	68.12	771.952	−1237.5	4.852	4.68
7	−0.201	−0.074	1.968	0.057	699,666	105.42	826.455	−1433.4	4.549	5.267
8	−0.217	−0.078	2.173	0.084	369,542	58.89	755.39	−1433.4	5.198	4.973
9	−0.214	−0.077	3.504	0.082	373,887	58.89	755.39	−1178.3	5.198	4.839
10	−0.225	−0.104	4.917	0.104	553,058	110.7	0	−1049.0	4.37	5.337
11*	−0.208	−0.07	2.746	0.046	369,542	79.12	771.531	−1049.1	4.251	4.868
12	−0.207	−0.076	2.513	0.057	383,575	79.12	771.531	−1049.1	4.251	5.135
13	−0.207	−0.076	3.244	0.051	459,685	79.12	771.531	−1508.6	4.251	4.907
14	−0.212	−0.084	4.889	0.065	459,685	79.12	790.225	−989.84	4.809	4.7
15*	−0.216	−0.108	3.398	0.096	305,070	84.4	0	−2156.7	4.345	4.931
16*	−0.229	−0.093	3.628	0.11	926,928	105.42	863.844	−2352.6	5.665	4.95
17*	−0.234	−0.095	6.373	0.21	519,357	58.89	792.779	−3075.8	6.315	4.858
18	−0.238	−0.102	4.739	0.223	1,249,742	105.42	901.232	−3075.85	6.781	5.437
19	−0.236	−0.101	4.98	0.239	1,249,742	105.42	913.239	−5187.2	7.052	5.416
20	−0.235	−0.108	5.475	0.301	1,249,742	105.42	961.733	−1287.9	8.38	5.289
21	−0.236	−0.103	2.178	0.14	525,199	58.89	823.029	−1013.4	7.114	5.92
22	−0.234	−0.097	1.594	0.131	525,199	58.89	792.779	−2353.6	6.315	5.653
23*	−0.244	−0.099	3.42	0.241	614,698	58.89	807.013	−2812.1	6.857	5.226
24	−0.24	−0.098	3.608	0.155	614,698	58.89	790.747	−2451.8	6.457	5.467
25*	−0.239	−0.097	4.056	0.155	614,698	58.89	802.753	−4563.2	6.728	5.431
26	−0.238	−0.1	4.225	0.146	625,683	58.89	768.737	−2091.4	6.073	5.59
27	−0.238	−0.097	3.979	0.14	614,698	58.89	769.379	−2091.4	6.057	5.307
28	−0.239	−0.102	2.301	0.152	739,946	58.89	788.074	−2551.0	6.615	5.139
29	−0.243	−0.103	4.062	0.148	733,230	58.89	788.074	−2551.0	6.615	4.679
30	−0.239	−0.1	1.454	0.121	638,922	82.68	832.242	−1985.2	5.79	6.031
31	−0.239	−0.105	3.466	0.119	651,290	82.68	832.242	−1985.2	5.79	5.856
32	−0.239	−0.101	6.299	0.116	638,922	82.68	832.242	−1985.2	5.79	5.749
33*	−0.251	−0.126	4.474	0.116	638,922	82.68	832.242	−2361.1	5.79	5.567
34*	−0.241	−0.132	3.54	0.284	1,438,564	157.23	0	−2361.1	5.244	4.769
35	−0.251	−0.126	4.438	0.094	1,481,608	157.23	0	−2361.1	5.244	4.226
36	−0.243	−0.112	1.936	0.151	766,274	110.7	0	−2097.4	5.608	6.07
37	−0.246	−0.121	4.491	0.143	786,411	110.7	0	−2097.4	5.608	5.759
38	−0.243	−0.113	6.778	0.165	766,274	110.7	0	−2097.4	5.608	5.636
39	−0.256	−0.135	4.609	0.149	1,239,487	162.51	0	−2301.8	5.568	6.552
40	−0.221	−0.091	4.791	0.061	294,384	52.82	753.995	−1774.7	5.069	4.979
41	−0.228	−0.102	5.61	0.157	772,614	79.12	846.308	−2887.3	6.484	5.388
42	−0.227	−0.1	5.845	0.138	439,214	52.82	803.391	−4810.2	6.456	5.484
43	−0.227	−0.096	2.767	0.1	294,384	32.59	737.855	−2164.1	6.017	5.725
44	−0.23	−0.101	1.698	0.129	365,990	32.59	756.549	−2623.7	6.575	5.278
45	−0.226	−0.097	4.324	0.101	351,202	32.59	735.823	−2263.3	6.159	5.136
46	−0.231	−0.103	3.567	0.11	294,384	32.59	768.105	−824.94	6.816	5.482
47	−0.23	−0.1	1.059	0.09	367,025	56.38	777.318	−1796.7	5.492	6.065
48*	−0.231	−0.105	3.648	0.089	375,202	56.38	777.318	−1796.7	5.492	5.07
49	−0.23	−0.1	7.247	0.089	367,025	56.38	777.318	−1796.7	5.492	5.508
50	−0.237	−0.105	4.289	0.075	930,998	130.93	0	−2172.6	5.219	5.44
51	−0.232	−0.115	0.985	0.12	449,534	84.4	0	−1908.9	5.583	6.136
52*	−0.235	−0.122	4.357	0.111	462,967	84.4	0	−1908.9	5.583	5.55
53	−0.231	−0.115	7.634	0.117	449,534	84.4	0	−1908.9	5.583	5.188
54*	−0.239	−0.139	4.435	0.159	765,301	136.21	0	−2113.3	5.543	6.124

*test set, [a] Energy HOMO, [b] Energy LUMO, [c] Dipole moment, [d] Strech Bend, [e] Balaban Index, [f] Polar Surface Area, [j] Boiling point, [i] Energy total, [h]Water/Octanol partition coefficient.

3.7. Frontier molecular orbital analysis

The analysis of the molecular orbitals of border presents a qualitative study on the excited states of the electrons in a molecule, these states are characterized by the energy of the molecular orbitals, and one finds two types: HOMO energies (the energies of the highest occupied molecular orbitals), which designate the ability to release electrons, are the nucleophilic-

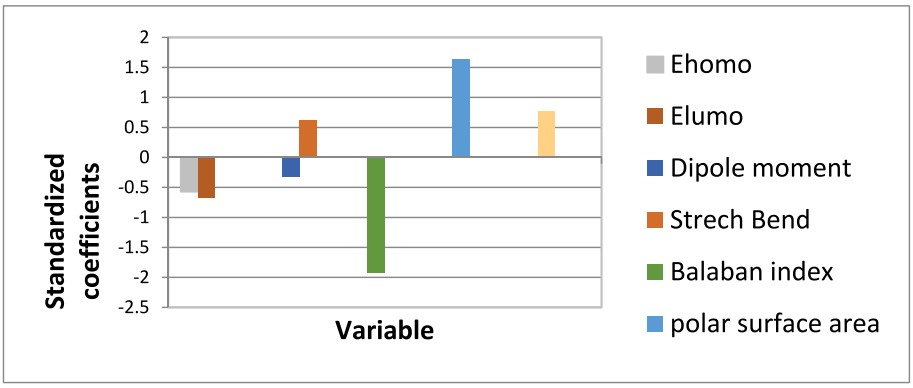

Fig. 3 The corresponding standardized descriptor coefficients.

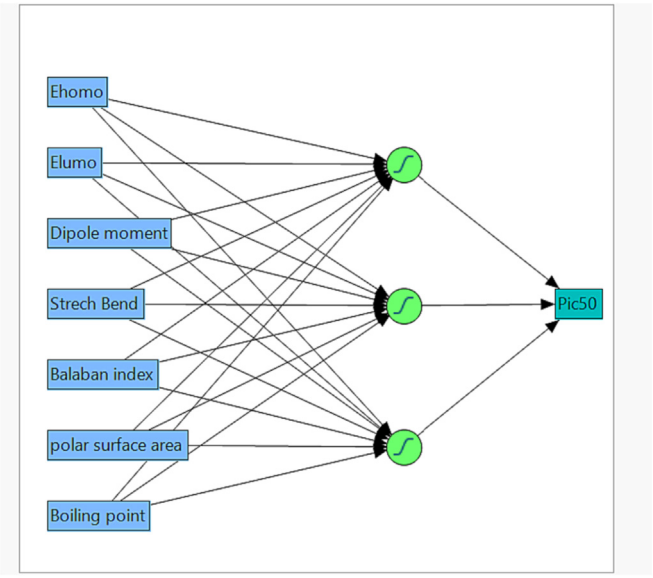

Fig. 4 The ANN architecture of the QSAR model.

ity; LUMO energies (the lowest unoccupied molecular orbital energies), which designate the ability to attract electrons, are the electrophilicity[36,37].

The reactivity is designated by the energy gap. The latter is represented by an energy gap between the HOMO energy and the LUMO energy, this gap reflects a minimum kinetic stability[45,46]. The small HOMO-LUMO energy gap allows for large HOMO-LUMO electronic charge exchange, which allows the molecule a high polarizability, moreover the more chemical compounds have a high electronic margin the more stable they have become and therefore a low reactivity compared to less strong energy gaps [37].

The bio-activity of chemical compounds based on the study of HOMO/LUMO energies to analyze the molecular reactivity, studied the interaction between molecules and intermolecular and electronic exchanges. To understand the chemical reactivity we decided to study some parameters of chemical reactivity: E_{homo}, E_{lumo}, energy gap ΔE, Electronic affinity A, chemical softness S, Electronic chemical potential μ, Ionization energy I, Global electrophilicity index ω, chemical hardness η, electronegativity X. The parameters are shown in Table 7. These parameters are linked by the following relationships:

$I = -E_{homo}$, $A = -E_{lumo}$, $\eta = \frac{1}{2}(E_{lumo}-E_{homo})$, $X = \frac{1}{2}(I + A)$, $S = \frac{1}{2}\eta$, $\mu = -\frac{1}{2}(I + A)$

$\omega = \frac{\mu^2}{2}\eta$, avec $\Delta E = E_{lumo}-E_{homo}$.

According to Table 7 we noticed two molecules; 34 and 54 having a low value of energy gap ΔE: 2.966 and 2.721 ev respectively, of which molecule 54 has a higher ionization energy 6.503 ev and shows the most marked value of electronic affinity 3.78 ev. For the hardness, the two molecules noticed present the lowest values which show that they have an important reactivity especially that these compounds have high values for the parameter of softness and therefore they possess an intense reactivity and a weak stability. If we compare the results of these molecules with the results of a molecule with a high energy gap, the molecule 14 for example, presents a high energy gap 3.483 ev and a high ionization rate 5.768 ev, we can conclude the stable effect of this molecule compared to the two molecules.

Other parameters are represented on Table 7. We find S the electronegativity shows an ability to capture electrons in order to build a bond, we notice that the 2 molecules with the smallest energy gap of 2-Styrylquinoline derivatives have an intense

Table 5 The predicted activities found by MLR and ANN methods.

N^a	$pIC_{50}Obs^b$	MLR^c	Residues (MLR)	ANN^e	Residues (ANN)	N^a	$pIC_{50}Obs^b$	MLR^c	Residues (MLR)	ANN^e	Residues (ANN)
1	4.881	5.067	−0.186	4.703	0.178	28	5.139	5.359	−0.22	5.501	−0.362
2	5.013	4.683	0.33	4.923	0.09	29	6.031	6.121	−0.09	5.978	0.053
3	4.933	4.982	−0.049	4.828	0.105	30	5.856	5.973	−0.117	5.807	0.049
4	4.773	4.532	0.241	4.733	0.04	31	5.749	5.649	0.1	5.883	−0.134
5	4.693	4.858	−0.165	4.642	0.051	32	5.59	5.350	0.24	5.511	0.079
6	4.68	4.793	−0.113	4.789	−0.109	33*	5.567	6.504	−0.937	5.400	0.167
7	5.267	4.921	0.346	5.221	0.046	34*	5.322	6.033	−0.711	4.411	0.911
8	4.973	5.177	−0.204	5.105	−0.132	35	4.226	4.744	−0.518	6.181	−1.955
9	4.839	4.953	−0.114	4.868	−0.029	36	6.07	5.955	0.115	5.776	0.294
10	5.337	5.590	−0.253	5.338	−0.001	37	5.759	5.824	−0.065	5.600	0.159
11*	4.868	5.123	−0.255	5.151	−0.283	38	5.636	5.597	0.039	6.465	−0.829
12	5.135	5.263	−0.128	5.100	0.035	39	6.552	6.203	0.349	5.148	1.404
13	4.907	4.920	−0.013	5.008	−0.101	40	4.979	5.183	−0.204	5.284	−0.305
14	4.7	5.107	−0.407	4.644	0.056	41	5.388	5.292	0.096	5.231	0.157
15*	4.931	5.703	−0.772	5.035	−0.104	42	5.484	5.419	0.065	5.484	0
16*	4.95	5.26	−0.31	4.806	0.144	43	5.725	5.271	0.454	5.619	0.106
17*	4.858	5.72	−0.862	5.214	−0.356	44	5.278	5.490	−0.212	5.144	0.134
18	5.437	5.190	0.247	5.467	−0.03	45	5.136	4.951	0.185	5.360	−0.224
19	5.416	5.221	0.195	5.389	0.027	46	5.482	5.490	−0.008	5.863	−0.381
20	5.289	5.712	−0.423	5.491	−0.202	47	6.065	5.929	0.136	5.505	0.56
21	5.92	5.904	0.016	5.558	0.362	48	5.07	5.763	−0.693	5.484	−0.414
22	5.653	5.724	−0.071	5.736	−0.083	49	5.508	5.334	0.174	5.125	0.383
23*	5.226	6.155	−0.929	5.320	−0.094	50	5.44	5.046	0.394	6.057	−0.617
24	5.467	5.518	−0.051	5.428	0.039	51	6.136	6.038	0.098	5.596	0.54
25	5.431	5.523	−0.092	5.378	0.053	52	5.557	5.806	−0.249	5.291	0.266
26	5.59	5.350	0.24	5.397	0.193	53	5.188	5.370	−0.182	5.879	−0.691
27	5.307	5.333	−0.026	5.448	−0.141	54	6.124	6.861	−0.737	4.703	1.421

[a]Number of compounds, [b] Observed activity, [c] activities observed by MLR method,[e] activities observed by ANN.

value of electronegativity which shows the reactivity of the compounds. Another criterion, the electrophilicity value ω determines the energy change related to the electron shift from HOMO to LUMO. The compounds 34, 54 present a higher electron transfer rate of the other synthesized compounds (Fig. 8).

3.8. Analysis of the molecular electrostatic potential (MEP)

The MEP is a method based on DFT calculation using the B3LYP-6-31G base in order to visually present the obtained results [47–49]. The importance of the visual representations of the MEP method is to show the chemical reactivity of the atoms and the active centers. The MEP analysis gives an idea about the biological reactivity at the nucleophilic and electrophilic sites of the chemical compounds, plus the hydrogen bonds. The MEP diagram represents the variation of electrostatic potentials displayed according to different colors: red, orange, yellow, green and blue follow the directions below: Red < orange < yellow < green < blue, besides the value of the chemical potential increases according to this direction, the red and orange colors represent the electrophilic potential, and the blue color represents the nucleophilicity Fig. 9 [38,39]:

Fig. 9 shows the disappearance of the active sites on the 2,4-styrylquinoline derivatives, the electrophilic sites are characterized by the red color (electrophilic reactivity), and the nucleophilic sites are characterized by the blue color (nucleophilic reactivity). One notices strong red zones are located on the attracting group $–NO_2$, are also located on the function –OH in particular the compounds, 14 and 34, which comprise red zones on the groups $–NO_2$ and O-Acetyl. Moreover, areas are characterized by a strong blue color around the aromatic hydrogens of 2-styrylquinoline which is noticed on the mapping of molecules 54 and 34. The schematization based on the MEP method is essential for analyzing the interactions between molecules, such as intermolecular interactions, and ligand–protein interactions [50].

3.9. Synthetic accessibility and Lipinski rules

The Lipinski rule of five, also known as the "Rule of Five," is a set of criteria used to evaluate the drug-likeness or pharmacokinetic properties of a chemical compound [51]. The rule was developed by Christopher Lipinski in 1997, and it is widely used in the pharmaceutical industry to predict whether a compound has the potential to become an orally active drug [31]. On the other hand, synthetic accessibility refers to the ease with which a compound can be synthesized or prepared. It is an important consideration in drug discovery because compounds that are difficult or expensive to synthesize may not be practical to develop into drugs, even if they meet the Lipinski rules [52].

The Lipinski parameters are:

- Molecular weight (MW): The molecular weight of the compound should be less than 500 Dalton (Da).

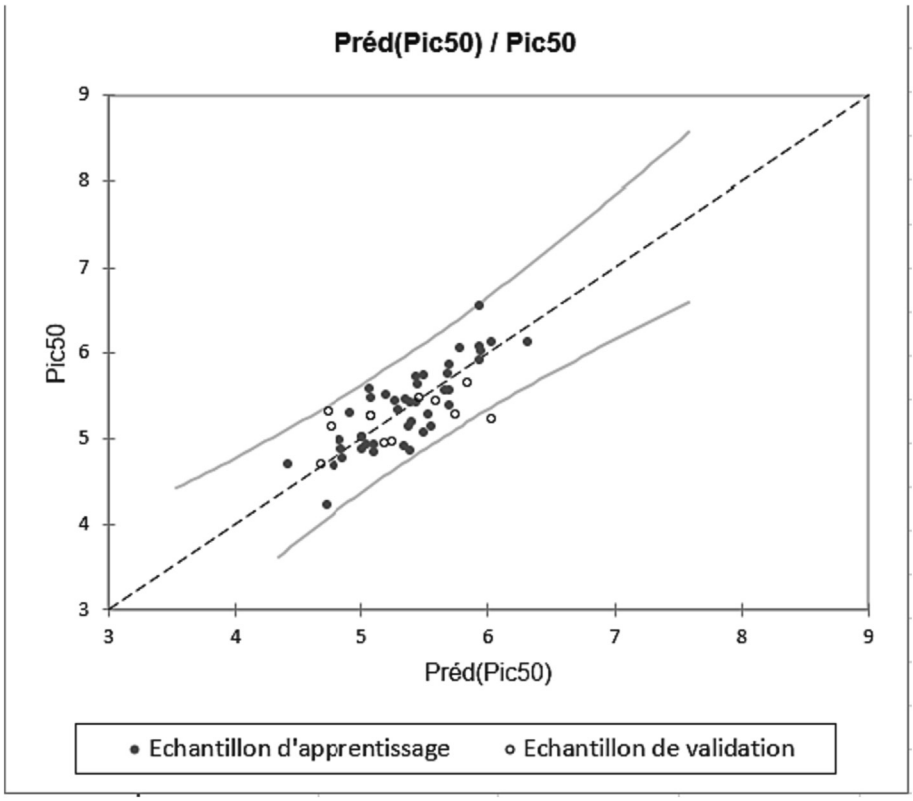

Fig. 5 The correlation between the observed biological activity and the biological activity predicted via the MLR method.

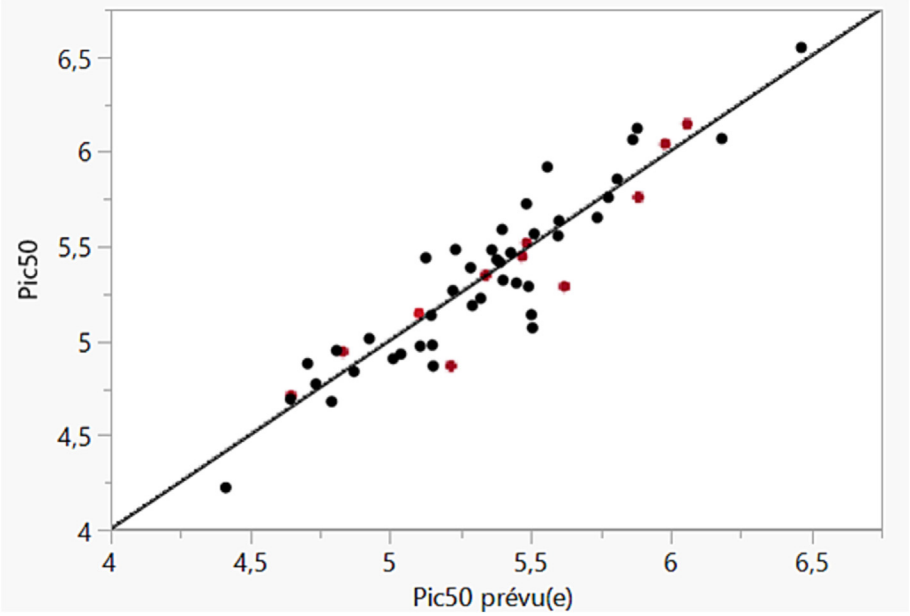

Fig. 6 Correlation between the observed activities and the activities predicted by the ANN method.

- Hydrogen bond donors (HBD): The compound should have no more than five hydrogen bond donors, which are typically OH and NH groups
- Hydrogen bond acceptors (HBA): The compound should have no more than ten hydrogen bond acceptors, which are typically O and N atoms
- Lipophilicity (Log (P)): The compound should have a logP (octanol–water partition coefficient) value of less than 5.
- Topological Polar Surface Area (TPSA) The compound should have a TPSA value of less than 140 $Å^2$.

Table 6 The Random MLR model parameter.

R_P^2	0.67
Average R^2	0.39
Average R	0.17

The Lipinski studies allow exploiting the pharmacokinetic characteristics of new drugs; we used the SwissADMET server to evaluate pharmacokinetic parameters https://www.swissadme.ch/ [36], such as molecular weights, TPSA measurements, LogP partition coefficients, hydrogen donors, hydrogen acceptors and Linpinski's rules are presented in Tables 8.

According to the stated criteria, the studied molecules should not show violation of Lipinski Rules, which is validated in Table 8. The derivatives 2–4 Styrylquinolines (16, 34, 51, 52, 54) have a LogP rate that varies between 2.89 and 4.15; besides that, the HBD of all the studied compounds is inferior to 5, and the molecular mass always remained inferior to 500 Da (see Fig. 10).

The number of rotating bonds presents the flexibility of the molecule; therefore, the more the number of rotating bonds is important, the more the molecule is flexible and more it is able to give an important biological activity, of which we notice that all the compounds have high value between 3 and 5 bonds, which favors their interaction as inhibitors. Except the molecule 34 that undergoes a value of synthetic accessibility that exceeds the threshold by a margin of 0.28, which is considered a molecule that is hardly synthesized.

The TPSA is an important parameter that allows to test the cellular penetration of the candidate compound as a drug; this value must exceed 140 $Å^2$ to have excellent transmission. The TPSA values of 5 derivatives 2–4 Styrylquinoline varies between 78.94 and 124.76, thus showing some cellular transmission. According to Table 8, the 2–4 Styrylquinoline derivatives are acceptable according to the Lipinski, Veber, Egan, and Ghose rules, which favor the synthesis of the compounds..

3.10. Docking studies

As part of the studies of the interactions between the molecules and the receptor protein, we performed molecular docking of compounds 14, 34 and 54 to study their roles as inhibitors on wild-type columns p53HCT116^{++}, whose protein PDB code is: 2GEQ. The 2GEQ protein, shown in Fig. 11:

The docking study carried out establishes that the significant sites are PRO A188, HIS A176, ARG A178 and SER A180. The best results have been chosen which correspond to the minimum energies (-5.51 Kcal/mol) for compound 54 and (-6.20 Kcal/mol) for compound 34 and (-5.50 kcal/mol) for the compound 16. From the visualization, we see the strong presence of the hydrogen bonds of the three compounds due to the hydroxide and nitrous oxide groups.

Fig. 12 shows that the amino acids for the two compounds 34 and 54 are respectively: SER A180, GIN A189, GLY A182, PRO A188, PRO A187, GLY A177, HIS A176, and ARG A178. The first observation of the results indicates the presence of hydrogen bonds with the residue: GIN A189 and SER A180, a single Pi-sigma bond with the residue GLY A182 and the presence of two Pi-alkyl bonds with the residue Pro 188. In molecule 54, we also notice three hydrogen bonds with residues SER A180, HISA176 and GLU A177, three Pi-alkyl bonds with residues PRO A188, ARG A178. Visualization of compound 16 shows the presence of five hydrogen bonds with residues: ASP A180, CYS A179, HIS 176, and SER A180, a Pi-alkyl bond with residue PRO A188 and a Pi-cation bond with residue ARG A178. Therefore, the three compounds are addressed the same sites, which results that three molecules are inhibitors of the receptor (protein) docked. We can conclude that the amino acids HIS 176, SER A180, PRO A188 and ARG A178 are active sites of the 2GEQ protein.

3.11. MD simulation

Three compounds (14, 34, and 54) binding to the 2GEQ receptor were chosen for molecular dynamics simulation up to 100 ns. The MD paths were analyzed through the parameters RMSF, RMSD, and R (g).

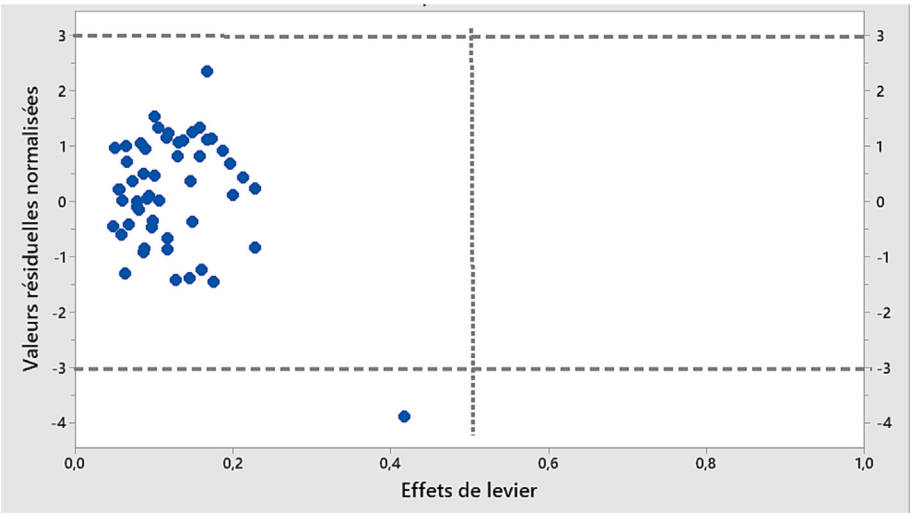

Fig. 7 Graphique de Williams du résidu standardisé en fonction de l'effet de levier pour le modèle MLR.

Table 7 Values obtained during the calculation of chemical reactivity parameters.

N°	E_{homo}	E_{lumo}	ΔE	I	A	X	S	η	μ	ω
1	−5.578	−1.932	3.646	5.578	1.932	3.755	0.9115	1.823	−3.755	3.8672
2	−5.333	−1.795	3.538	5.333	1.795	3.564	0.8845	1.769	−3.564	3.5901
3	−5.605	−1.877	3.728	5.605	1.877	3.741	0.932	1.864	−3.741	3.754
4	−5.415	−1.877	3.538	5.415	1.877	3.646	0.8845	1.769	−3.646	3.7572
5	−5.551	−1.904	3.647	5.551	1.904	3.7275	0.91175	1.8235	−3.727	3.8097
6	−5.578	−1.850	3.728	5.578	1.850	3.714	0.932	1.864	−3.714	3.7
7	−5.469	−2.013	3.456	5.469	2.013	3.741	0.864	1.728	−3.741	4.0495
8	−5.904	−2.122	3.782	5.904	2.122	4.013	0.9455	1.891	−4.013	4.2581
9	−5.904	−2.122	3.782	5.904	2.122	4.013	0.9455	1.891	−4.013	4.2581
10	−6.122	−2.829	3.293	6.122	2.829	4.4755	0.82325	1.6465	−4.475	6.0826
11	−5.659	−1.904	3.755	5.659	1.904	3.7815	0.93875	1.8775	−3.781	3.8081
12	−5.632	−2.068	3.564	5.632	2.068	3.85	0.891	1.782	−3.85	4.1589
13	−5.632	−2.068	3.564	5.632	2.068	3.85	0.891	1.782	−3.85	4.1589
14	−5.768	−2.285	3.483	5.768	2.285	4.0265	0.87075	1.7415	−4.026	4.6548
15	−5.877	−2.938	2.939	5.877	2.938	4.4075	0.73475	1.4695	−4.407	6.6097
16	−6.231	−2.530	3.701	6.231	2.530	4.3805	0.92525	1.8505	−4.380	5.1847
17	−6.367	−2.585	3.782	6.367	2.585	4.476	0.9455	1.891	−4.476	5.2973
18	−6.476	−2.775	3.701	6.476	2.775	4.6255	0.92525	1.8505	−4.625	5.7809
19	−6.421	−2.748	3.673	6.421	2.748	4.5845	0.91825	1.8365	−4.584	5.7221
20	−6.394	−2.802	3.592	6.394	2.802	4.598	0.898	1.796	−4.598	5.8857
21	−6.421	−2.802	3.619	6.421	2.802	4.6115	0.90475	1.8095	−4.611	5.8761
22	−6.367	−2.639	3.728	6.367	2.639	4.503	0.932	1.864	−4.503	5.4391
23	−6.639	−2.693	3.946	6.639	2.693	4.666	0.9865	1.973	−4.666	5.5173
24	−6.530	−2.666	3.864	6.530	2.666	4.598	0.966	1.932	−4.598	5.4714
25	−6.503	−2.639	3.864	6.503	2.639	4.571	0.966	1.932	−4.571	5.4073
26	−6.476	−2.721	3.755	6.476	2.721	4.5985	0.93875	1.8775	−4.598	5.6314
27	−6.503	−2.639	3.864	6.503	2.639	4.571	0.966	1.932	−4.571	5.4073
28	−6.612	−2.775	3.837	6.612	2.775	4.6935	0.95925	1.9185	−4.693	5.7411
29	−6.612	−2.802	3.81	6.612	2.802	4.707	0.9525	1.905	−4.707	5.8151
30	−6.503	−2.721	3.782	6.503	2.721	4.612	0.9455	1.891	−4.612	5.6241
31	−6.503	−2.857	3.646	6.503	2.857	4.68	0.9115	1.823	−4.68	6.0072
32	−6.503	−2.748	3.755	6.503	2.748	4.6255	0.93875	1.8775	−4.625	5.6978
33	−6.830	−3.428	3.402	6.830	3.428	5.129	0.8505	1.701	−5.129	7.7326
34	−6.557	−3.591	2.966	6.557	3.591	5.074	0.7415	1.483	−5.074	8.6802
35	−6.830	−3.428	3.402	6.830	3.428	5.129	0.8505	1.701	−5.129	7.7326
36	−6.612	−3.047	3.565	6.612	3.047	4.8295	0.89125	1.7825	−4.829	6.5425
37	−6.693	−3.292	3.401	6.693	3.292	4.9925	0.85025	1.7005	−4.992	7.3287
38	−6.612	−3.074	3.538	6.612	3.074	4.843	0.8845	1.769	−4.843	6.6293
39	−6.966	−3.673	3.293	6.966	3.673	5.3195	0.82325	1.6465	−5.319	8.5931
40	−6.013	−2.476	3.537	6.013	2.476	4.2445	0.88425	1.7685	−4.244	5.0935
41	−6.204	−2.775	3.429	6.204	2.775	4.4895	0.85725	1.7145	−4.489	5.8779
42	−6.176	−2.721	3.455	6.176	2.721	4.4485	0.86375	1.7275	−4.448	5.7276
43	−6.176	−2.612	3.564	6.176	2.612	4.394	0.891	1.782	−4.394	5.4172
44	−6.258	−2.748	3.51	6.258	2.748	4.503	0.8775	1.755	−4.503	5.7769
45	−6.149	−2.639	3.51	6.149	2.639	4.394	0.8775	1.755	−4.394	5.5006
46	−6.693	−2.802	3.891	6.693	2.802	4.7475	0.97275	1.9455	−4.747	5.7925
47	−6.258	−2.721	3.537	6.258	2.721	4.4895	0.88425	1.7685	−4.489	5.6985
48	−6.693	−2.857	3.836	6.693	2.857	4.775	0.959	1.918	−4.775	5.9438
49	−6.258	−2.721	3.537	6.258	2.721	4.4895	0.88425	1.7685	−4.489	5.6985
50	−6.449	−2.857	3.592	6.449	2.857	4.653	0.898	1.796	−4.653	6.0273
51	−6.313	−3.129	3.184	6.313	3.129	4.721	0.796	1.592	−4.721	6.9999
52	−6.395	−3.319	3.076	6.395	3.319	4.857	0.769	1.538	−4.857	7.6691
53	−6.693	−3.129	3.564	6.693	3.129	4.911	0.891	1.782	−4.911	6.767
54	−6.503	−3.782	2.721	6.503	3.782	5.1425	0.68025	1.3605	−5.142	9.7189

The energies are given in eV.

Molecular dynamic simulation allows the study of the stability of ligand docking sites and amino acid participations, aims to test the stability of protein–ligand compounds (complexes) and identify the various possible ligand bonds [53,54]. A MD simulation was performed for the complexes: T_1 (ligand 34-protein) which is characterized by purple color and T_2 complex (ligand 54-protein) which is characterized by red color and T_3 (ligand14-protein) which is characterized by green color on the plots of results (Figs. 12, and 13).

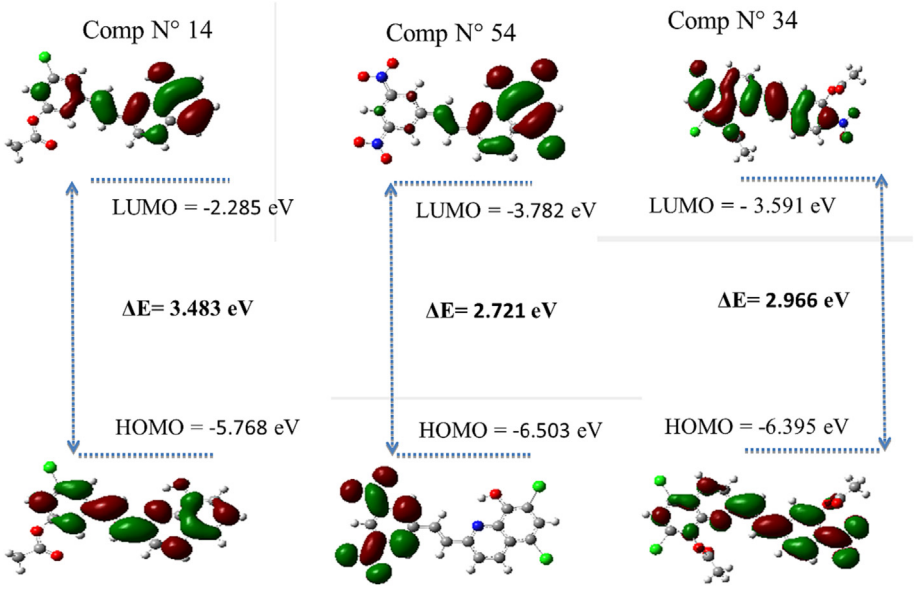

Fig. 8 The molecular boundary orbitals of 2-Styrylquinoline derivatives.

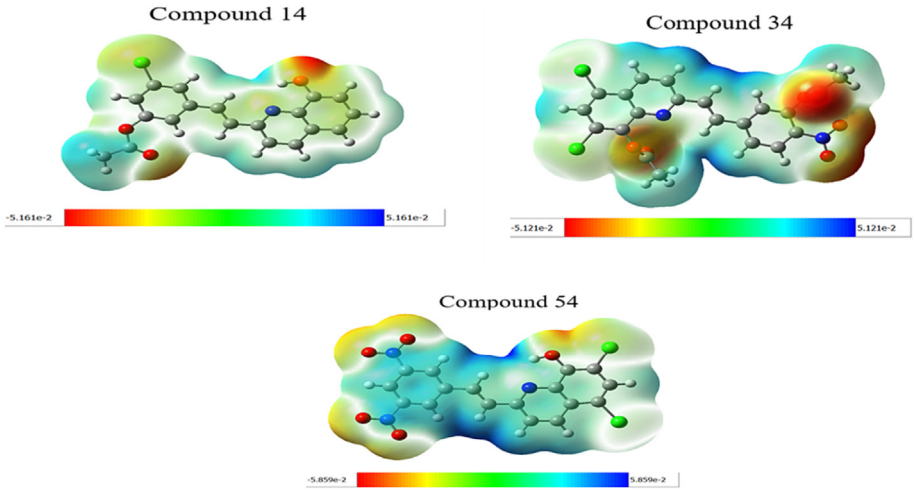

Fig. 9 The molecular electrostatic potential of 2-Styrylquinoline derivatives.

Table 8 The Summary of the parameters of synthetic accessibility.

N°	Properties								Synthetic accessibility
	M W^1	$LogP^2$	HBD^3	HBA^4	$TPSA^5$	Ghose	Egan	Veber	
14	339.77	3.37	1	4	59.42	Yes	Yes	Yes	2.59
34	461.25	4.15	0	7	111.31	Yes	Yes	Yes	3.28
51	419.21	3.83	1	6	105.24	Yes	Yes	Yes	3
52	361.18	3.87	1	4	78.94	Yes	Yes	Yes	2.65
54	406.18	3.10	1	6	124,76	Yes	Yes	Yes	2.89

M W^1 Molecular weight, $LogP^2$ Consensus of calculated lipophilicity, HBD^3 number of hydrogen bond acceptor, HBA^4 number of hydrogen bond Donor, MR^4 Molar refractivity, $TPSA^5$ Topological polar surface area.

As shown in Fig. 13, the RMSD value of the protein structure had a variation of 1.2 to 3.2 Å during the first 53 ns. On the other hand, the RMSD value of the protein structure continued its variation until it found stability around 2.7 Å at 60 ns. At 78 ns the T2 complex had a slight deviation of 3.8 Å which did not impact the average RMSD value of the

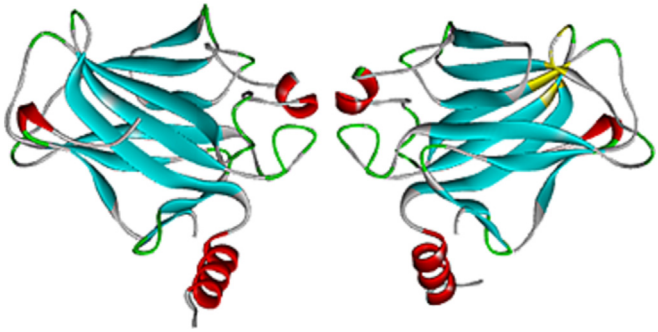

Fig. 10 The protein (2GEQ) presented with its two chains A and B.

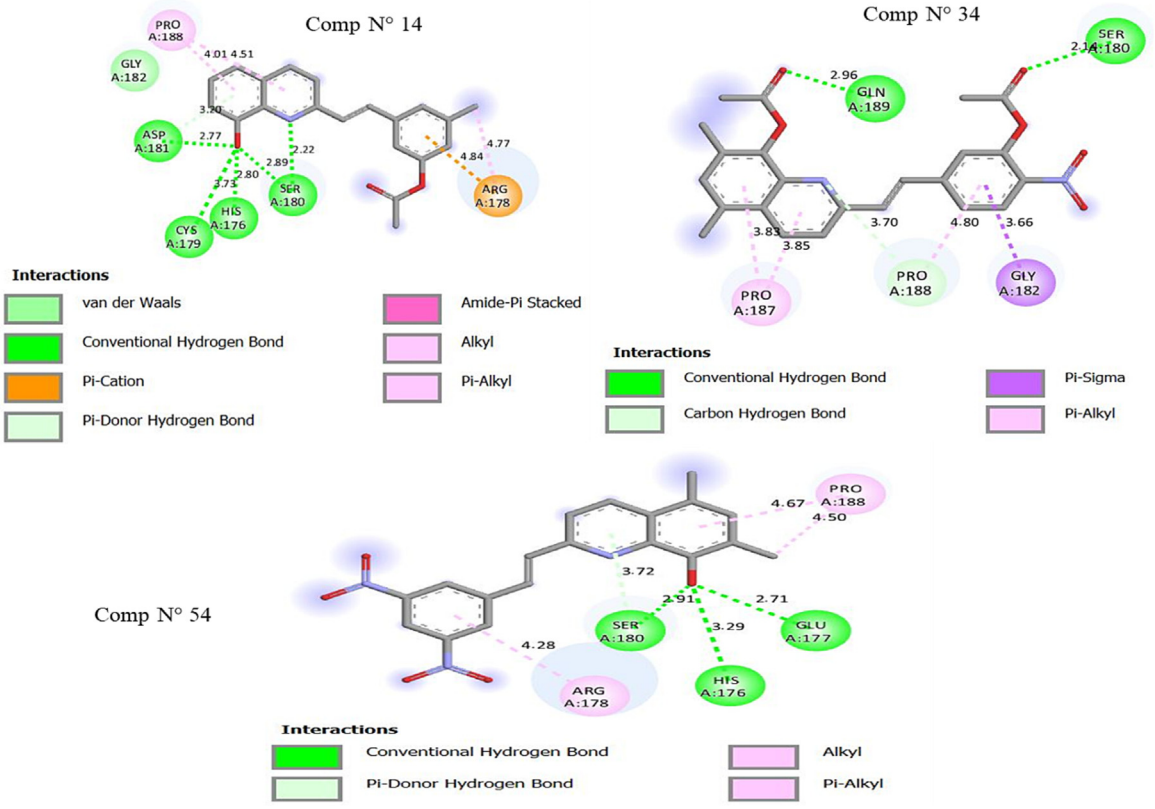

Fig. 11 The 2D representations of the interactions between the compounds and the 2EGQ protein.

T2 complex which equals 2.32 Å compared to the T1 complex which has an average RMSD value of 2.28 Å which shows that the T1 complex is more stable than the T_2 complex. The RMSD study shows that compounds 34 and 54 formed stable complexes with the 2GEQ receptor during the simulation. An RMSF simulation was also calculated (Fig. 12); a maximum variation of 5.7 Å was noted in the 187 ns residue loop region for T_1 and T_2 complexes.

The radius of gyration Rg (Fig. 12) corresponds to the change in density of the receptor structure (protein: 2GEQ) as a function of time. For the T_1 complex during the 34 ns, the simulation Rg fluctuated between 16.41 and 17.08 Å after this period and until the end of the simulation, the Rg values of the T_1 complex kept a reasonable stability between 16.7 and 17.1 Å, the average gyration variation Rg for the T_1 complex is 16.80.

For the T_2 complex during 53 ns, the Rg values varied between 16.3 and 17.2 Å after this time they stabilized until the end of the simulation except for a small deviation of 17.3 Å at 77 ns which did not influence the average value of Rg = 16.77.

As shown in Fig. 13, the RMSD of compound 14 had a stability that did not exceed 2 Å throughout the 40 ns simulation. During the first 20 ns, the RMSD value indicated a change in RMSD between 1.2 and 1.9 Å, after which the RMSD value of the protein structure is stabilized until we reach 40 ns. The RMSD is included between 1.2 and 1.9 Å, and the average RMSD value is 1.56 Å. An RMSF simulation was also evalu-

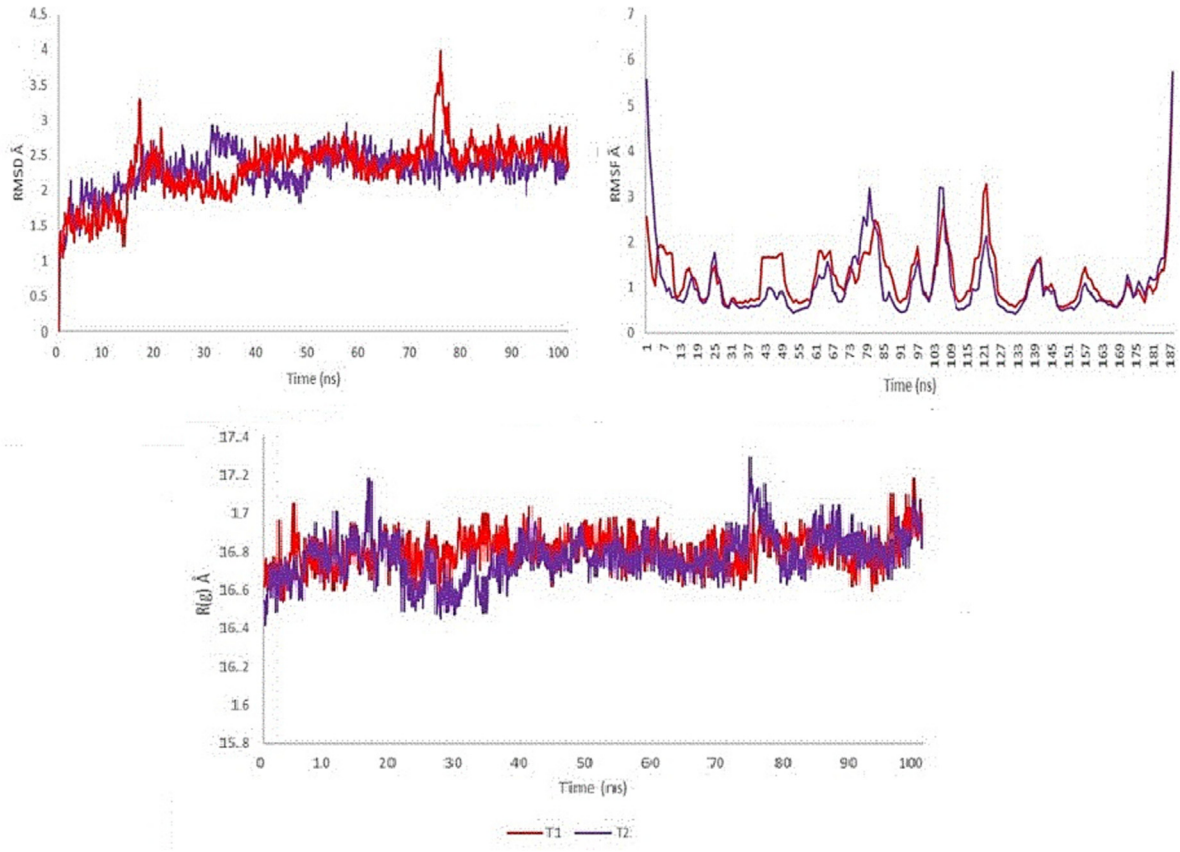

Fig. 12 RMSD, RMSF and Rg observed during MD simulation of 100 ns for the macromolecular complex of compound 34 and 54 against 2EGQ protein.

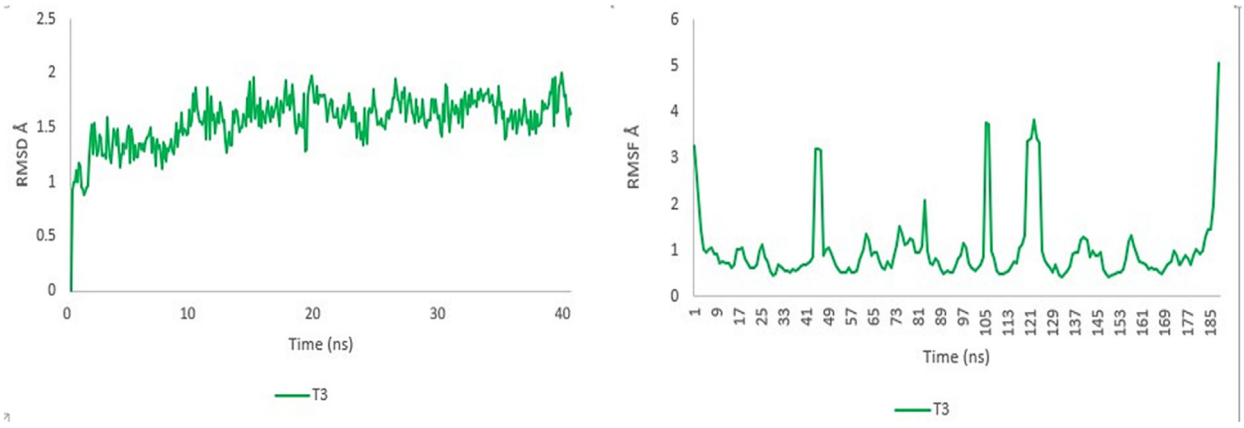

Fig. 13 RMSD, RMSF observed during MD simulation of 40 ns for the macromolecular complex of compound 14 against 2EGQ protein.

ated (Fig. 13); a maximum variation of 5.1 was noted in the area of the 185 ns residual loop for the T_3 complex.

4. Conclusion

In this study, we developed a mathematical QSAR model to determine the relationship between molecular structure parameters and biological activity for a series of 54 derivatives of 2-Styrylquinoline. The obtained (QSAR) MLR model is validated for its statistical content, the tests: ANN, Y-randomization, and domain of applicability are also performed to confirm the solidity of the MLR model. Using the DFT B3LYP method in the 6-31G theory, we found that the energy gap (HOMO-LUMO) is about 3.483 eV, 2.966 eV and 2.721 eV of the molecules N° 14, 34 and 54. The findings indicate that these three molecules have a high degree of chemical stability and reactivity. The descriptors: Ionization energy, electronic affinity, electronegativity, and other parameters that

influence the reactivity of a molecule are calculated and explained in a thorough way. The potential analysis (MEP) presents electrophilic zones around the –OH function and the –NO2 attracting group that form hydrogen bonds with the different residues of the protein (PDB: 2GEQ). The five compounds (N° 14, 34, 51, 52, and 54) all meet the Lipinski rules, and respect the synthetic accessibility criteria. The five compounds have respected all the rules of similarity among the drugs and the rules of bioavailability.

The molecular docking simulation showed the way 2-Styrylquinoline derived compounds move into active protein sites due to their interactions with residues (amino acids) that play a primary role in relation to their anticancer activity. Molecular docking studies have shown that hydrogen bonds with key amino acids of the targeted protein are essential for stabilizing ligands in the active sites of the 2GEQ Protein and increasing their inhibitory efficacy. A molecular dynamics simulation was performed on three docked compounds, whose thermodynamic stability was tested by running the MD simulation in 40 and 100 ns, all 3 compounds showed significant hydrogen bonding to key amino acids of the target protein throughout the molecular dynamics simulation period. This suggests that they are stable and likely to bind to the target protein in a biological setting.

Based on the findings of the study, the three molecules recommended are highly suggested as potential anti-cancer agents targeting $HCT116^{++}$ line.

Funding

This article publication charge was funded by the Princess Nourah bint Abdulrahman University, Researchers Supporting Project number (PNURSP2023R165), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors extend their appreciation to Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2023R165), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

References

[1] World Health Organization (WHO), (n.d.). https://www.who.int/ (accessed March 8, 2023).
[2] A selective p53 activator and anticancer agent to improve colorectal cancer therapy. - Abstract - Europe PMC, (n.d.). https://europepmc.org/article/med/33852837 (accessed March 8, 2023).
[3] J.P. Gudikote, T. Cascone, A. Poteete, P. Sitthideatphaiboon, Q. Wu, N. Morikawa, F. Zhang, S. Peng, P. Tong, L. Li, L. Shen, M. Nilsson, P. Jones, E.P. Sulman, J. Wang, J.-C. Bourdon, F.M. Johnson, J.V. Heymach, Inhibition of nonsense-mediated decay rescues p53β/γ isoform expression and activates the p53 pathway in MDM2-overexpressing and select p53-mutant cancers, J. Biol. Chem. 297 (2021), https://doi.org/10.1016/j.jbc.2021.101163.
[4] V. Khwaza, O.O. Oyedeji, B.A. Aderibigbe, Ursolic Acid-Based Derivatives as Potential Anti-Cancer Agents: An Update, Int. J. Mol. Sci. 21 (2020) 5920, https://doi.org/10.3390/ijms21165920.
[5] Quinoline Heterocycles: Synthesis and Bioactivity | IntechOpen, (n.d.). https://www.intechopen.com/chapters/64434 (accessed March 8, 2023).
[6] Quinoline-based HIV Integrase Inhibitors: Ingenta Connect, (n.d.). https://www.ingentaconnect.com/content/ben/cpd/2013/00000019/00000010/art00008 (accessed March 8, 2023).
[7] A. Kamal, A. Rahim, S. Riyaz, Y. Poornachandra, M. Balakrishna, C.G. Kumar, S.M.A. Hussaini, B. Sridhar, P.K. Machiraju, Regioselective synthesis, antimicrobial evaluation and theoretical studies of 2-styryl quinolines, Org. Biomol. Chem. 13 (2015) 1347–1357, https://doi.org/10.1039/C4OB02277G.
[8] Derivatives of 2-Styrylquinoline | Journal of Medicinal Chemistry, (n.d.). https://pubs.acs.org/doi/pdf/10.1021/jm50009a001 (accessed March 8, 2023).
[9] R. Musiol, J. Jampilek, V. Buchta, L. Silva, H. Niedbala, B. Podeszwa, A. Palka, K. Majerz-Maniecka, B. Oleksyn, J. Polanski, Antifungal properties of new series of quinoline derivatives, Bioorg. Med. Chem. 14 (2006) 3592–3598, https://doi.org/10.1016/j.bmc.2006.01.016.
[10] A. Mrozek-Wilczkiewicz, M. Kuczak, K. Malarz, W. Cieślik, E. Spaczyńska, R. Musiol, The synthesis and anticancer activity of 2-styrylquinoline derivatives. A p53 independent mechanism of action, Eur. J. Med. Chem. 177 (2019) 338–349, https://doi.org/10.1016/j.ejmech.2019.05.061.
[11] M. Er-rajy, M. El fadili, S. Mujwar, H. Imtara, O. Al kamaly, S. Zuhair Alshawwa, F.A. Nasr, S. Zarougui, M. Elhallaoui, Design of novel anti-cancer agents targeting COX-2 inhibitors based on computational studies, Arabian Journal of Chemistry. (2023) 105193. https://doi.org/10.1016/j.arabjc.2023.105193.
[12] M. Er-Rajy, M. El Fadili, S. Mujwar, S. Zarougui, M. Elhallaoui, Design of novel anti-cancer drugs targeting TRKs inhibitors based 3D QSAR, molecular docking and molecular dynamics simulation, Journalof Biomolecular Structure & Dynamics (2023) 1–14, https://doi.org/10.1080/07391102.2023.2170471.
[13] Life | Free Full-Text | 3D-QSAR Studies, Molecular Docking, Molecular Dynamic Simulation, and ADMET Proprieties of Novel Pteridinone Derivatives as PLK1 Inhibitors for the Treatment of Prostate Cancer, (n.d.). https://www.mdpi.com/2075-1729/13/1/127 (accessed March 11, 2023).
[14] QSAR, molecular docking, and molecular dynamics simulation–based design of novel anti-cancer drugs targeting thioredoxin reductase enzyme | SpringerLink, (n.d.). https://link.springer.com/article/10.1007/s11224-022-02111-x (accessed March 11, 2023).
[15] M. Er-rajy, M. El Fadili, H. Hadni, N.N. Mrabti, S. Zarougui, M. Elhallaoui, 2D-QSAR modeling, drug-likeness studies, ADMET prediction, and molecular docking for anti-lung cancer activity of 3-substituted-5-(phenylamino) indolone derivatives, Struct. Chem. 33 (2022) 973–986, https://doi.org/10.1007/s11224-022-01913-3.
[16] Free chemical structure drawing software | free ADME properties | TPSA, (n.d.). https://www.simulations-plus.com/software/medchem-designer/?gclid=Cj0KCQiAgaGgBhC8ARIsAAAyLfE2n2D_O5-xdflCV12JQhElBvU0Y_xtk4Vs755o_gCwVzjkoWThcgcaAiT4EALw_wcB (accessed March 8, 2023).
[17] C. Corporation, ChemOffice, (n.d.). https://www.cambridgesoft.com/Ensemble_for_Chemistry/details/Default.aspx?fid=16 (accessed March 8, 2023).

[18] Gaussian.com | Expanding the limits of computational chemistry, (n.d.). https://gaussian.com/ (accessed March 8, 2023).
[19] A. Lagunin, A. Zakharov, D. Filimonov, V. Poroikov, QSAR Modelling of Rat Acute Toxicity on the Basis of PASS Prediction, Mol Inform. 30 (2011) 241–250, https://doi.org/10.1002/minf.201000151.
[20] M. Er-rajy, M. El fadili, N.N. Mrabti, S. Zarougui, M. Elhallaoui, QSAR, molecular docking, ADMET properties in silico studies for a series of 7-propanamide benzoxaboroles as potent anti-cancer agents, Chinese Journal of Analytical Chemistry. 50 (2022) 100163. https://doi.org/10.1016/j.cjac.2022.100163.
[21] Statistical Software | JMP, (n.d.). https://www.jmp.com/en_us/home.html (accessed March 8, 2023).
[22] MATLAB - MathWorks, (n.d.). https://www.mathworks.com/products/matlab.html (accessed March 8, 2023).
[23] Developing quantitative structure–retention relationship model to prediction of retention factors of some alkyl-benzenes in nano-LC | SpringerLink, (n.d.). https://link.springer.com/article/10.1007/s13738-019-01624-3 (accessed March 8, 2023).
[24] Data Analysis, Statistical & Process Improvement Tools | Minitab, (n.d.). https://www.minitab.com/en-us/ (accessed March 8, 2023).
[25] RCSB PDB: Homepage, (n.d.). https://www.rcsb.org/ (accessed March 8, 2023).
[26] Open Babel, SourceForge. (2017). https://sourceforge.net/projects/openbabel/ (accessed March 8, 2023).
[27] K. Karrouchi, S.A. Brandán, Y. Sert, M.E. Karbane, S. Radi, M. Ferbinteanu, Y. Garcia, M. Ansar, Synthesis, structural, molecular docking and spectroscopic studies of (E)-N'-(4-methoxybenzylidene)-5-methyl-1H-pyrazole-3-carbohydrazide, J. Mol. Struct. 1225 (2021), https://doi.org/10.1016/j.molstruc.2020.129072 129072.
[28] M. Gümüş, Ş.N. Babacan, Y. Demir, Y. Sert, İ. Koca, İ. Gülçin, Discovery of sulfadrug–pyrrole conjugates as carbonic anhydrase and acetylcholinesterase inhibitors, Arch. Pharm. 355 (2022) 2100242, https://doi.org/10.1002/ardp.202100242.
[29] mgl-admin, Homepage, AutoDock. (n.d.). https://autodock.scripps.edu/ (accessed March 9, 2023).
[30] D. Systèmes, Free Download: BIOVIA Discovery Studio Visualizer, Dassault Systèmes. (2020). https://discover.3ds.com/discovery-studio-visualizer-download (accessed March 9, 2023).
[31] C.A. Lipinski, F. Lombardo, B.W. Dominy, P.J. Feeney, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, Adv. Drug Deliv. Rev. 64 (2012) 4–17, https://doi.org/10.1016/j.addr.2012.09.019.
[32] D.F. Veber, S.R. Johnson, H.-Y. Cheng, B.R. Smith, K.W. Ward, K.D. Kopple, Molecular Properties That Influence the Oral Bioavailability of Drug Candidates, J. Med. Chem. 45 (2002) 2615–2623, https://doi.org/10.1021/jm020017n.
[33] W.J. Egan, K.M. Merz, J.J. Baldwin, Prediction of Drug Absorption Using Multivariate Statistics, J. Med. Chem. 43 (2000) 3867–3877, https://doi.org/10.1021/jm000292e.
[34] A.K. Ghose, V.N. Viswanadhan, J.J. Wendoloski, A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases, J. Comb. Chem. 1 (1999) 55–68, https://doi.org/10.1021/cc9800071.
[35] Y.C. Martin, A Bioavailability Score, J. Med. Chem. 48 (2005) 3164–3170, https://doi.org/10.1021/jm0492002.
[36] SwissADME, (n.d.). http://www.swissadme.ch/ (accessed March 9, 2023).
[37] A. Montanari, G. Young, H.H.G. Savenije, D. Hughes, T. Wagener, L.L. Ren, D. Koutsoyiannis, C. Cudennec, E. Toth, S. Grimaldi, G. Blöschl, M. Sivapalan, K. Beven, H. Gupta, M. Hipsey, B. Schaefli, B. Arheimer, E. Boegh, S.J. Schymanski, G. Di Baldassarre, B. Yu, P. Hubert, Y. Huang, A. Schumann, D. A. Post, V. Srinivasan, C. Harman, S. Thompson, M. Rogger, A. Viglione, H. McMillan, G. Characklis, Z. Pang, V. Belyaev, "Panta Rhei—Everything Flows": Change in hydrology and society—The IAHS Scientific Decade 2013–2022, Hydrol. Sci. J. 58 (2013) 1256–1275, https://doi.org/10.1080/02626667.2013.809088.
[38] V. Zoete, M.A. Cuendet, A. Grosdidier, O. Michielin, SwissParam: A fast force field generation tool for small organic molecules, J. Comput. Chem. 32 (2011) 2359–2368, https://doi.org/10.1002/jcc.21816.
[39] M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method, J. Appl. Phys. 52 (1981) 7182–7190, https://doi.org/10.1063/1.328693.
[40] M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, E. Lindahl, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX. 1–2 (2015) 19–25, https://doi.org/10.1016/j.softx.2015.06.001.
[41] G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling, J. Chem. Phys. 126 (2007), https://doi.org/10.1063/1.2408420 014101.
[42] XLSTAT | Statistical Software for Excel, XLSTAT, Your Data Analysis Solution. (n.d.). https://www.xlstat.com/en/ (accessed March 9, 2023).
[43] F. Nimr Ajeel, A. Mohsin Khuodhair, S. Mahdi AbdulAlmohsin, Improvement of the Optoelectronic Properties of Organic Molecules for Nanoelectronics and Solar Cells Applications: via DFT-B3LYP Investigations, Current Physical, Chemistry 7 (2017) 39–46, https://doi.org/10.2174/1877946806666161115141959.
[44] A. Mermer, H. Bayrak, S. Alyar, M. Alagumuthu, Synthesis, DFT calculations, biological investigation, molecular docking studies of β-lactam derivatives, J. Mol. Struct. 1208 (2020), https://doi.org/10.1016/j.molstruc.2020.127891 127891.
[45] N. Dege, H. Gökce, O.E. Doğan, G. Alpaslan, T. Ağar, S. Muthu, Y. Sert, Quantum computational, spectroscopic investigations on N-(2-((2-chloro-4,5-dicyanophenyl)amino) ethyl)-4-methylbenzenesulfonamide by DFT/TD-DFT with different solvents, molecular docking and drug-likeness researches, Colloids Surf.A: Physicochemical and Engineering Aspects. 638 (2022), https://doi.org/10.1016/j.colsurfa.2022.128311 128311.
[46] A.A. Abdulridha, M.A. Albo Hay Allah, S.Q. Makki, Y. Sert, H.E. Salman, A.A. Balakit, Corrosion inhibition of carbon steel in 1 M H2SO4 using new Azo Schiff compound: Electrochemical, gravimetric, adsorption, surface and DFT studies, Journal of Molecular Liquids. 315 (2020) 113690. https://doi.org/10.1016/j.molliq.2020.113690.
[47] T.K. Kuruvilla, J.C. Prasana, S. Muthu, J. George, S.A. Mathew, Quantum mechanical and spectroscopic (FT-IR, FT-Raman) study, NBO analysis, HOMO-LUMO, first order hyperpolarizability and molecular docking study of methyl [(3R)-3-(2-methylphenoxy)-3-phenylpropyl]amine by density functional method, Spectrochim. Acta A Mol. Biomol. Spectrosc. 188 (2018) 382–393, https://doi.org/10.1016/j.saa.2017.07.029.
[48] L. Li, C. Wu, Z. Wang, L. Zhao, Z. Li, C. Sun, T. Sun, Density functional theory (DFT) and natural bond orbital (NBO) study of vibrational spectra and intramolecular hydrogen bond interaction of l-ornithine–l-aspartate, Spectrochim. Acta A Mol. Biomol. Spectrosc. 136 (2015) 338–346, https://doi.org/10.1016/j.saa.2014.08.153.
[49] G. Subhapriya, S. Kalyanaraman, N. Surumbarkuzhali, S. Vijayalakshmi, V. Krishnakumar, Investigation of intermolecular hydrogen bonding in 2,3,4,5,6

pentafluorobenzoic acid through molecular structure and vibrational analysis – A DFT approach, J. Mol. Struct. 1083 (2015) 48–56, https://doi.org/10.1016/j.molstruc.2014.11.033.

[50] S. Mujwar, A. Tripathi, Repurposing benzbromarone as antifolate to develop novel antifungal therapy for Candida albicans, J. Mol. Model. 28 (2022) 193, https://doi.org/10.1007/s00894-022-05185-w.

[51] B.G. Giménez, M.S. Santos, M. Ferrarini, J.P.S. Fernandes, J.P. S. Fernandes, Evaluation of blockbuster drugs under the Rule-of-five, Die Pharmazie - An International Journal of Pharmaceutical Sciences. 65 (2010) 148–152, https://doi.org/10.1691/ph.2010.9733.

[52] M.-Q. Zhang, B. Wilkinson, Drug discovery beyond the 'rule-of-five', Curr. Opin. Biotechnol. 18 (2007) 478–488, https://doi.org/10.1016/j.copbio.2007.10.005.

[53] M. El fadili, M. Er-rajy, H. Imtara, O.M. Noman, R.A. Mothana, S. Abdullah, S. Zerougui, M. Elhallaoui, QSAR, ADME-Tox, molecular docking and molecular dynamics simulations of novel selective glycine transporter type 1 inhibitors with memory enhancing properties, Heliyon. 9 (2023) e13706. https://doi.org/10.1016/j.heliyon.2023.e13706.

[54] M. El fadili, M. Er-rajy, H. Imtara, M. Kara, S. Zarougui, N. Altwaijry, O. Al kamaly, A. Al Sfouk, M. Elhallaoui, 3D-QSAR, ADME-Tox In Silico Prediction and Molecular Docking Studies for Modeling the Analgesic Activity against Neuropathic Pain of Novel NR2B-Selective NMDA Receptor Antagonists, Processes. 10 (2022) 1462. https://doi.org/10.3390/pr10081462.