Article

Design and Optimization of Quinazoline Derivatives as Potent EGFR Inhibitors for Lung Cancer Treatment: A Comprehensive QSAR, ADMET, and Molecular Modeling Investigation

Mohamed Moussaoui,*,# Soukayna Baammi,# Hatim Soufi, Mouna Baassi, Mohammed Salah, Achraf EL Allali, Belghiti Elalaoui Mohammed, Rachid Daoud,* and Said Belaaouad

Read Online

ACCESS | Metrics & More | Article Recommendations | SI Supporting Information

ABSTRACT: The epidermal growth factor receptor (EGFR) is part of a protein family that controls cell growth and development. Due to its importance, EGFR has been identified as a suitable target for creating novel drugs. For this research, we conducted a 2D-QSAR analysis on a set of 31 molecules derived from quinazoline, which exhibited inhibitory activity against human lung cancer. This investigation incorporated principal component analysis (PCA) and multiple linear regression (MLR), leading to the development of QSAR models with strong predictive capabilities (R^2 = 0.745, R^2_adj = 0.723, MSE = 0.061, R^2_test = 0.941, and Q^2_cv = 0.669). The reliability of these models was confirmed through internal, external, Y-randomization, and applicability domain validations. Leveraging the predictions from the QSAR model, we designed 18 new molecules based on the modifications at the N-3 and C-6 positions of the quinazoline ring, with electronegative substituents at these positions fostering optimal polar interactions and hydrophobic contacts within the ATP-binding site of EGFR, significantly enhancing the inhibitory activity against the lung cancer cell line. Subsequently, ADMET predictions were conducted for these 18 compounds, revealing outstanding ADMET profiles. Molecular docking analyses were performed to investigate the interactions between the newly designed molecules—Pred15, Pred17, Pred20, Pred21—and the EGFR protein, indicating high affinity of these proposed compounds to EGFR. Furthermore, molecular dynamics (MD) simulations were utilized to assess the stability and binding modes of compounds Pred17, Pred20, and Pred21, reinforcing their potential as novel inhibitors against human lung cancer. Overall, our findings suggest that these investigated compounds can serve as effective inhibitors, showcasing the utility of our analytical and design approach in the identification of promising therapeutic agents.
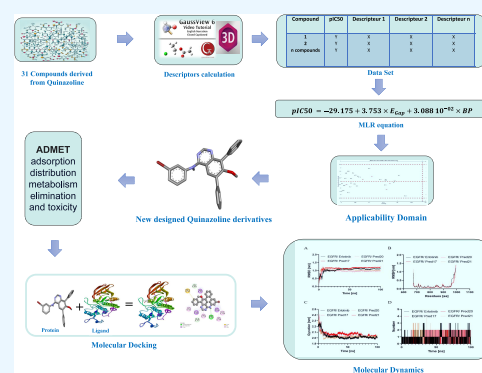
1. INTRODUCTION

Lung cancer is the leading cause of cancer-related deaths across both sexes, responsible for roughly 18.4% of all cancer fatalities, outstripping the death rates from colon, prostate, and breast cancers combined. Manifesting primarily in two forms, lung cancer is categorized as small-cell lung carcinoma (SCLC), accounting for about 13% of cases, and nonsmall cell lung cancer (NSCLC), which comprises 84% of cases.[1−4] Despite noteworthy progress in diagnostics and drug therapies,[5] the survival rates for lung cancer patients remain disappointingly low. The current treatment modalities include chemotherapy,[6] radiation therapy,[7] systemic therapy,[8] targeted therapy,[9] and immunotherapy.[10] Nevertheless, for those with metastatic lung cancer—a condition marked by the spread of primary lung tumors to other parts of the body through the blood and lymphatic systems, commonly impacting the adrenal glands,[11] liver,[12] brain,[13] bones,[14] and lymph nodes[15]—the options for effective treatment are severely limited. Lung transplantation is acknowledged as an effective measure to prolong the life of end-stage lung disease patients.[16] Recent

research underscores the potential of pulmonary drug delivery systems, especially those utilizing nanoparticles,[17] to directly target the lungs[18] and significantly improve the effectiveness of anticancer treatments.[19−22] These nanoparticle-based delivery platforms are becoming recognized as a highly promising approach to enhance lung cancer therapy, offering a new and potentially more effective strategy for battling this deadly disease.[23] Currently, EGFR inhibitors have emerged as a pivotal category of anticancer medications. The discovery of highly potent EGFR inhibitors holds immense significance in combating various tumors, such as lung cancer, cervical cancer, and breast cancer. Currently, a plethora of small-molecule

tyrosine kinase inhibitors (TKIs) targeting EGFR have been developed. Gefitinib (with an IC50 range of 23−79 nM) and erlotinib (with an IC50 of 80 nM) stand as first-generation EGFR-TKIs, surpassing standard chemotherapy in clinical usage. These medications function as reversible competitive inhibitors of ATP, effectively impeding EGFR autophosphorylation. In the realm of breast cancer treatment, afatinib, a second-generation EGFR-TKI with an IC50 of 0.5 nM, has garnered significant attention.

Rational drug design can be understood as a collection of innovative methods aimed at discovering new medications based on the knowledge of biological targets. This process often relies on computer modeling techniques, commonly referred to as "in silico" methods.[24] In this study, we focus on developing novel EGFR inhibitors to explore new therapeutic avenues for treating human lung cancer, utilizing computer-aided drug design techniques. Specifically, we employ the "ligand-based drug design" approach,[25] which includes quantitative structure−activity relationships (QSAR). This method establishes a correlation between the calculated properties of molecules and their experimentally determined biological activities. Additionally, we utilize a second strategy that targets the molecular structure, known as "structure-based drug design" (SBDD).[26] This approach is a successful, appealing, and widely adopted strategy in both academic and pharmaceutical research domains. It facilitates the examination of the drug's target structure, providing detailed molecular insights into the mechanism of action and inhibition of the target.

In recent years, in silico methods have emerged as essential tools in the field of drug discovery, particularly for understanding and predicting activity against cancer-related targets such as kinases. Multitarget quantitative structure−activity relationship (QSAR) models have played a pivotal role in these advancements. These models enable researchers to simultaneously analyze multiple targets, offering a comprehensive overview that is crucial for the development of broad-spectrum cancer therapeutics. A notable advancement in this area is the development of predictive therapeutic modeling for lung cancer (PTML), which facilitates the design of inhibitors that target multiple proteins and cells implicated in pancreatic cancer.[27]

Furthermore, the integration of fragment-based topological design into QSAR and machine learning models enhances the physicochemical and structural interpretation of molecular interactions. This methodology supports the rational design of novel molecules, as it allows for the intricate dissection of molecular fragments that contribute significantly to the activity against specific targets. Noteworthy applications of this approach include the design of dual inhibitors of CDK4 and HER2[28] and BET bromodomain inhibitors using multitarget QSAR models.[29]

Additionally, combining ensemble learning with fragment-based topological approaches has led to significant innovations in drug discovery. This combined method has been employed in the in silico design of Hsp90 inhibitors, generating new molecular diversity that could potentially lead to more effective cancer treatments.[30,31] The TOPS-MODE approach further exemplifies the utility of these advanced QSAR models in developing inhibitors for a range of tyrosine kinases, which are critical in the regulation of cancer cell growth and survival.[32]

These methodologies not only enhance the accuracy of drug design but also streamline the discovery process by predicting the efficacy and safety of multitarget inhibitors against complex diseases like cancer. By leveraging these advanced in silico techniques, researchers can significantly reduce the time and cost associated with traditional drug discovery processes.

In this research, 2D-QSAR techniques were used to develop and predict the anticancer potential of novel molecules based on a set of 31 analogs of quinazoline derivatives. Molecular docking assessment and simulations revealed significant enhancement in binding affinity and inhibitory effects against EGFR, a pivotal receptor in lung cancer, in comparison to erlotinib, the current standard lung cancer medication. These findings lay the groundwork for the synthesis of novel quinazoline analogs with enhanced efficacy against lung cancer.

2. RESEARCH METHOD

2.1. Data Sources. In order to establish a relationship between the structure and activity of various compounds quantitatively, we analyzed a collection of 31 quinazoline derivatives. These compounds, which act as tyrosine kinase (EGFR) inhibitors, were sourced from a study conducted by Bridges et al.[33] In the calculation of pIC50, the constant "6" is subtracted from the negative logarithm of the IC50 value, which is calculated as follows: pIC50 = 6 − $\log_{10}$ (IC50).

This adjustment allows for the direct comparison of binding affinities across a range of concentrations, simplifying the interpretation of lower IC50 values as higher potencies.

The collected data was randomly segmented into two sets: a training set and a test set. The training set comprised 26 compounds and was utilized to develop a 2D-QSAR (quantitative structure−activity relationship) model. Meanwhile, the test set consisted of five compounds, which were used to assess and validate the accuracy and quality of the developed model.

2.2. Molecular Descriptors. In our endeavor to develop a robust QSAR model, we meticulously calculated over 26 molecular descriptors from the studied series. These descriptors, covering physicochemical, topological, geometrical, and constitutional aspects, were computed using a variety of software tools, including ChemOffice[34] and Gaussian 09 Software,[35] as detailed in Tables S1 and S2. For optimization purposes, we employed the MM2 method. We also refined the geometries of the 31 quinazoline derivatives using the density functional theory (DFT) approach, incorporating the B3LYP functional[36] and the 6-31G(p, d) basis set.[37] Following the quantum computations,[38−40] we extracted several critical structural parameters. These parameters included the highest occupied molecular orbital energy (E_{HOMO}), lowest unoccupied molecular orbital energy (E_{LUMO}), dipole moment (μm), total energy (TE), absolute hardness (η), absolute electronegativity (χ), and reactivity index (ω). Specific equations were utilized to determine the values of η, χ, and ω.

$$\eta = \frac{E_{\text{LUMO}} - E_{\text{HOMO}}}{2} \tag{1}$$

$$\chi = \frac{E_{\text{LUMO}} + E_{\text{HOMO}}}{2} \tag{2}$$

$$\omega = \frac{\mu^2}{2\eta} \tag{3}$$

The topological descriptors were calculated using ChemOffice software.[34] Various parameters were evaluated in this process, such as the molecular weight (MW), the number of

hydrogen bond acceptors (NHA), and the number of hydrogen bond donors (NHD). Other considerations included the octanol−water partition coefficient (LogP) and water solubility (LogS). Furthermore, several indices and coefficients were also assessed, including the Balaban Index (J), the molecular topological index (MTI), the polar surface area (PSA), the radius (RDWV), the shape coefficient (I), the sum of valence degrees (SVD), the boiling point (BP), the Wiener index (WI), and the number of rotatable bonds (N_{ROT}).

2.3. Statistical Analysis.
2.3.1. Principal Component Analysis. Principal component analysis (PCA) is a data analysis technique used to reduce the dimensionality of large data sets, simplifying information and minimizing redundancy.[41] The results obtained from PCA are crucial in determining the relationships among descriptors, aiding in the identification of key input variables necessary for the QSAR model.

2.3.2. Linear Regression. In linear regression, the metrics employed to gauge the relationship between dependent and independent variables comprise the coefficient of determination (R^2), the root mean square error (RMSE), and the Fisher ratio's value (*F*-test). The equation for multiple linear regression (MLR) is established as follows:

$$Y = b_0 + \sum_{i=1}^{n} biXi \tag{4}$$

In this model, Y signifies the predicted value of biological activity. Xi corresponds to the molecular descriptors, which serve as the independent variables. The term n represents the count of molecular descriptors, and b_0 is identified as the constant within the equation.

2.4. QSAR Model validation.
The QSAR model was developed utilizing anticancer activity data obtained from a study by Bridges et al.[33] A robust QSAR model is capable of making reliable predictions concerning a set of query compounds. To ascertain our model's reliability, we conducted both internal (represented by R^2) and external (denoted by R^2_{pred}) cross-validations. The cross-validated correlation coefficient (R^2) and the predictive squared correlation coefficient (R^2_pred) are key indicators in the conventional validation of QSAR models employing MLR, detailed in eqs 5−7. A principal method applied in this validation procedure is the leave-one-out (LOO) cross-validation, explicitly presented in eq 6.

$$Q^2 = 1 - \frac{\sum \left(Y_{obs(train)} - Y_{pred(train)}\right)^2}{\sum \left(Y_{obs(train)} - \overline{Y}_{training}\right)^2} \tag{5}$$

$$R^2 = 1 - \frac{\sum \left(Y_{obs(train)} - Y_{calc(train)}\right)^2}{\sum \left(Y_{obs(train)} - \overline{Y}_{obs(train)}\right)^2} \tag{6}$$

$$R^2_{pred} = 1 - \frac{\sum \left(Y_{obs(test)} - Y_{pred(test)}\right)^2}{\sum \left(Y_{obs(test)} - \overline{Y}_{training}\right)^2} \tag{7}$$

In eqs 5 and 6, $Y_{obs(train)}$ denotes the observed activity for the training set, $Y_{pred(train)}$ indicates the predicted activity for the molecules in the training set, derived using the LOO technique, $Y_{calc(train)}$ refers to the response calculated by the model for the training set, and $Y_{obs(train)}$ represents the average of the observed response values for the training set.

Additionally, the Golbraikh and Tropsha-proposed parameters for calculating a QSAR model's external predictability were derived.[41]

2.4.1. Y-Randomization. Y-randomization is an auxiliary approach for evaluating the robustness of a model. By shuffling the values of the dependent variable (Y) and preserving the integrity of the descriptor values, this method facilitates the creation of new MLR models. This process is repeated n times, generating n different models. The R^2 and Q^2 values of each new model are then compared with those of the original models. A model is considered robust if these values are as low as possible.

2.4.2. Domain of Applicability. The applicability domain of the model refers to the theoretical space occupied by compounds, as defined by their descriptors and modeled activity, where a QSAR model can make reliable predictions.[42] This concept is instrumental in identifying both structural and response outliers within the training and test sets. A leverage approach was utilized to assess the chemical space of the QSAR model, employing a specific plot known as the Williams plot. This plot juxtaposes standardized residuals against leverage values (h). Compounds are deemed within the model's applicability domain if their leverage scores fall below a certain threshold ($h < h^*$), and their standardized residual scores reside within $\pm 3.0\sigma$ (standard deviation units).[43] The warning leverage (h^*) is ascertained through a specific calculation:

$$h^* = 3 \times \frac{(d + 1)}{N} \tag{8}$$

Here, d refers to the quantity of descriptors used in the model while N indicates the total number of compounds forming the training set.

2.5. Drug-Likeness and ADMET Properties.
Today's drug research and development are characterized by substantial costs and extensive time commitments. These challenges are exacerbated by the complexities associated with emerging diseases and the unpredictability of research outcomes.[44] Computational strategies have significantly transformed this domain, facilitating rapid optimization of pharmacokinetic properties and the comprehensive evaluation of drug candidate toxicity. Fundamental to the development of new drug candidates is an exhaustive examination of Absorption, Distribution, Metabolism, Excretion, and Toxicity (ADMET) properties. A thorough assessment of these ADMET characteristics is crucial for reducing failure risks during the drug discovery process, thereby ensuring the efficacy and reliability of the resulting pharmaceutical products.[45]

To assess the pharmacokinetic and pharmacodynamic properties of molecules, their formats are converted from.cdx to SMILES using the SwissADME server. This conversion facilitates the evaluation of various attributes, such as synthetic accessibility and compliance with Lipinski's, Veber's, and Egan's rules. The molecules, now represented in SMILES format, are submitted to the pkCSM server (http://biosig. unimelb.edu.au/pkCSM/)[46] for property calculations. This includes determining the MW, NHA, NHD, and N_{Rot}. According to Lipinski's rule of five, a molecule is likely to be orally active if it satisfies the following criteria: MW less than 500 g/mol, LogP 5 or less, NHD fewer than 5, NHA fewer than 10, and N_{Rot} 10 or fewer.

The pkCSM servers[47] provide extensive information on selected molecules, covering various absorption properties

Table 1. pIC50 Values and Chemical Descriptor Analysis for Quinazoline Derivatives against EGFR: Observed, Predicted Activities, and MLR Model Residuals
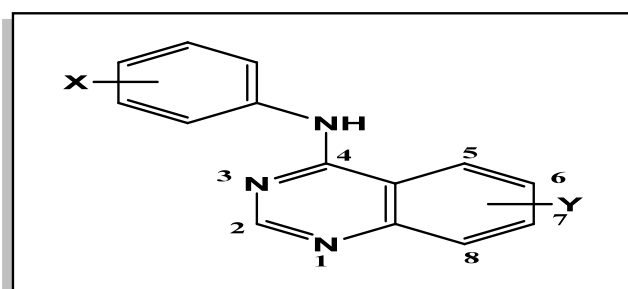
no.	substituent		E_{Gap}	boiling Point	observed pIC50	predicted pIC50	residuals
	X	Y					
1	H	H	4.204	647.27	6.463	6.592	−0.129
2	3-F	H	4.244	644.61	7.252	6.662	0.590
3*	3-Cl	H	4.269	665.98	7.638	7.417	0.221
4	3-Br	H	4.265	677.99	7.569	7.770	−0.201
5*	3-CF3	H	4.305	650.70	6.239	7.078	−0.840
6	H	6-OMe	4.036	670.94	7.260	6.693	0.567
7	3-Br	6-OMe	4.076	701.64	7.523	7.791	−0.268
8	H	6-NH2	4.003	689.89	6.113	7.152	−1.039
9	3-CF3	6-NH2	4.033	693.27	6.241	7.371	−1.130
10	3-Br	6-NH2	4.023	720.56	9.108	8.179	0.929
11	H	6-NO2	3.358	702.97	5.301	5.138	0.163
12	3-Br	6-NO2	3.452	733.67	6.046	6.438	−0.393
13	H	7-OMe	4.237	670.94	6.921	7.449	−0.528
14	3-Br	7-OMe	4.302	701.64	8.000	8.638	−0.638
15	H	7-NH2	4.306	689.86	7.000	8.290	−1.290
16	3-F	7-NH2	4.350	687.18	8.699	8.375	0.324
17	3-Cl	7-NH2	4.376	708.55	9.602	9.132	0.470
18*	3-Br	7-NH2	4.373	720.56	10.000	9.490	0.509
19	3-CF3	7-NH2	4.410	693.27	8.481	8.789	−0.307
20	H	7-NO2	3.294	702.97	4.921	4.897	0.024
21	3-F	7-NO2	3.368	700.29	5.215	5.094	0.120
22	3-Cl	7-NO2	3.408	721.66	6.091	5.904	0.187
23*	3-Br	7-NO2	3.397	733.667	6.000	6.234	−0.234
24	H	6,7-Di-OMe	4.195	694.59	7.538	8.019	−0.482
25	3-F	6,7-Di-OMe	4.233	691.92	8.420	8.082	0.338
26	3-Cl	6,7-Di-OMe	4.256	713.29	9.509	8.828	0.681
27	3-Br	6,7-Di-OMe	4.252	725.29	10.602	9.184	1.418
28	3-CF3	6,7-Di-OMe	4.287	698.00	9.620	8.473	1.146
29	3-Br	6-NHMe	3.911	711.28	8.398	7.470	0.928
30*	3-Br	6-NMe2	3.894	712.47	7.076	7.445	−0.369
31	3-Br	6-NHCOOMe	4.114	749.06	7.921	9.400	−1.480

such as water solubility, intestinal absorption, and Caco-2 permeability.[48] Additionally, they offer insights into distribution attributes like permeability to the central nervous system (CNS)[49] and the blood−brain barrier (BBB).[50] The servers also detail metabolic aspects, including the inhibition of CYP1A2, CYP2C19, CYP2C9, CYP2D6, and CYP3A4, as well as overall clearance. Information regarding AMES toxicity and skin sensitization is also made available.[51]

2.6. Molecular Docking. Eighteen compounds, or ligands, were analyzed through in silico docking experiments using AutoDock Vina[52] and AutoDock Tools[53] on Windows 10. The two-dimensional structures of the ligands were crafted in ChemDraw and converted to 3D using ChemOffice.[34] AutoDock Tools facilitated the preparation of the ligands for docking, converting them from MOL file format to PDBQT files. Data, including the cocrystallized form of EGFR, were

sourced from the RCSB Protein Data Bank (https://www.rcsb.org). After evaluating multiple submissions, a ligand-bound complex (PDB entry: 1M17)[54] was selected based on its high resolution and the structural similarity of the cocrystallized ligand to the test compounds. Noninteracting water molecules and ions were removed from the EGFR PDB file using Discovery Studio 2021.[55] AutoDock Tools was used to prepare the proteins and generate PDBQT files from PDB files, adding every polar hydrogen atom to the macromolecule to enhance docking. Grid settings were configured using AutoDock Tools, and details were compiled in a "conf.txt" file. The grid dimensions were set at 40 × 40 × 40 XYZ points and centered at coordinates 22.014, 0.253, and 52.794. AutoDock's scoring algorithms evaluated the nine best poses for each ligand, recording the binding free energy of each ligand in a log file. Poses with the lowest binding free energy (ΔG) were selected,

indicating more robust and stable ligand-protein binding. Discovery Studio was used for a detailed analysis of the docked poses, exploring the hydrogen bonds and hydrophobic interactions between the ligands and proteins.

2.7. Molecular Dynamics (MD) Simulations. MD simulations were conducted utilizing GROMACS software, specifically version 2019.3, to investigate the conformational dynamics of the most favorable docking complexes with the utilization of the CHARMM 27 force field.[56] The protein topology was constructed using the GROMACS pdb 2gmx modules with the Chemistry at Harvard Macromolecular Mechanics force field (CHARMM ff),[57] and the ligand topology was generated via the SwissParam server.[58] The docked structures were immersed in a simulation box with dimensions of 9.6 nm on each side and solvated using the three-point transferable intermolecular potential (TIP3P) solvent.[59] To ensure system neutrality, chloride (Cl^-) and sodium (Na^+) ions were introduced when needed. The system was then subjected to energy minimization using the steepest descent algorithm, with a maximum force threshold of 1000 kJ/mol/nm. The pressure and temperature were then set to 1 bar and 300 K using the Nosé–Hoover thermostat and isotropic Parrinello–Rahman barostat.[60] Finally, a 100 ns simulation was carried out for each docked complex. We employed custom scripts derived from the results of MD simulations to compute several metrics, such as RMSD (utilizing "gmx rms"), root-mean-square fluctuation of residues (with "gmx rmsf"), solvent-accessible surface area (SASA), the radius of gyration (Rg), and the count of hydrogen bonds (utilizing "gmx hbond").

2.8. Binding Energy Calculation by MM-PBSA. To calculate binding free energies of the screened complexes, the molecular mechanics Poisson–Boltzmann surface area (MM-PBSA) was used.[61] The binding free energy (E binding) is obtained by the following equations:

$$\Delta E_{MM\text{-}PBSA} = E_{complex} - (E_{protein} + E_{ligand}) \quad (9)$$

Equation 9 is the total MM-PBSA energy of the protein–ligand complex; $E_{protein}$ and E_{ligand} are the isolated proteins' and ligands' total free energies in solution, respectively.

$$\Delta G_{MM\text{-}PBSA} = \Delta E_{vdw} + \Delta E_{Elec} + \Delta E_{pol} \quad (10)$$

Equation 10 is the sum MM-PBSA of the following energies: electrostatic (E_{Elec}), van der Waals (E_{vdw}), polar (E_{polar}), and nonpolar (E_{Apolar}).

3. RESULTS AND DISCUSSION

3.1. Data Set for Analysis. A QSAR analysis was performed on 31 molecules, utilizing calculated descriptors, the values of which are outlined in Table S2. The K-means method was employed to categorize these molecules into two groups: a training set comprising 26 molecules and a test set of 5 molecules. Through iterative processes using MLR, a reliable model was formulated. This involved discarding descriptors with high correlations and retaining those with low correlations. The refined descriptors, obtained through the MLR method, along with their predicted pIC50 values, are presented in Table 1.

3.2. Physicochemical and Structural Interpretation of QSAR Models. In our study, we have utilized a series of quinazoline derivatives, each featuring modifications at specific positions on the quinazoline ring system. These structural variations include substitutions with different functional groups such as fluorine, chlorine, bromine, and various alkyl and aryl groups. These substitutions are strategically placed to explore their electronic and steric impacts on binding affinity toward the epidermal growth factor receptor (EGFR). The quinazoline derivatives were selected based on a range of physicochemical properties that are known to influence drug behavior, including lipophilicity (LogP), molecular weight, polar surface area, hydrogen bond donors, and acceptors. These properties are critical in determining the molecules' pharmacokinetic profiles, such as ADMET. For instance, the LogP values of these compounds help in understanding their ability to permeate cell membranes, which is crucial for their efficacy as intracellular inhibitors.

Our QSAR model incorporates several molecular descriptors that are directly linked to the physicochemical attributes of the quinazoline derivatives. Key descriptors include the following:

1. Hydrophobicity (LogP): correlates with the compounds' cell membrane permeability.
2. Molecular weight and topological polar surface area (TPSA): influence the oral bioavailability and ability to cross BBBs.
3. Electronic properties: such as the highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) energies, which impact the reactivity of the compounds with the EGFR kinase domain.

In our analysis, we discuss how the structural modifications of the quinazoline ring affect their interaction with the ATP-binding site of EGFR. For example, the presence of electron-donating or electron-withdrawing groups at specific positions on the ring influences the electronic distribution within the molecule, which in turn can enhance or reduce binding affinity through alterations in molecular dipole moments and potential electronic interactions with the kinase domain.

3.3. PCA. Twenty-six descriptors were calculated for PCA, as detailed in Table S3. Descriptors (independent variables) with correlation coefficients greater than 0.90 were omitted. Significant correlations were noted among various descriptors, including the following:

- E_{LUMO}, correlated with μ ($r = 0.98$), S ($r = 0.92$), and X ($r = 0.98$), resulting in the exclusion of E_{LUMO}.
- μ, correlated with X ($r = 1$) and ω ($r = 0.93$), leading to the removal of μ.
- E_{HOMO}, correlated with ω ($r = 0.99$) and X ($r = 0.93$), causing the elimination of E_{HOMO}.

The PCA correlation tables are available in Table S3. Following the removal of highly correlated descriptors, linear regression was executed to develop a QSAR model.

3.4. MLR. MLRs are predominantly utilized in 2D-QSAR analysis due to their straightforwardness, representativeness, and traceability. This approach relies on three critical parameters: the coefficient of determination (R^2), Fisher's ratio (F), and RMSE. The analysis was conducted using the XLSTAT 2019 software.[62] The derived QSAR model for the training class is depicted by eq 11, and the associated normalized descriptor coefficients are illustrated in Figure S1.

$$pIC50 = -29.175 + 3.753 \times E_{gap} + 3.08810^{-02} \times BP \quad (11)$$

The outcomes of the MLR model were assessed in comparison with Golbraikh and Tropsha's parameters.[63]

Table 2. Comparison of Model Parameters (MLR) with Golbraikh and Tropsha Criteria

	parameter	expression	model score	threshold	comment
fitting criteria	R^2	$R^2 = 1 - \dfrac{\sum (Y_{obs} - Y_{pred})^2}{\sum (Y_{pred} - \overline{Y}_{obs})^2}$	0.745	>0.6	passed
	R^2_{adj}	$R^2_{adj} = \dfrac{(N-1)R^2 - P}{N - P - 1}$	0.723	>0.6	passed
	MSE	$MSE = \dfrac{\sum (Y_{obs} - Y_{pred})}{N}$	0.061	a low value	passed
	F_{test}	$F_{test} = \dfrac{\sum (Y_{pred} - \overline{Y}_{pred})^2}{\sum (Y_{obs} - Y_{pred})^2} \times \dfrac{N - p - 1}{p}$	33.601	a high value	passed
internal validation	Q^2_{CV}	$Q^2_{CV} = 1 - \dfrac{\sum (Y_{obs} - Y'_{pred})^2}{\sum (Y'_{pred} - \overline{Y}_{obs})^2}$	0.669	>0.5	passed
	R_{Rand}	average of the 100 $R_{Rand}(i)$	0.253	<R	passed
	R^2_{Rand}	average of the 100 $R^2_{Rand}(i)$	0.080	<R^2	passed
	$Q^2_{CVLOORand}$	average of the 100 $Q^2_{CVLOO(Rand)}(i)$	−0.176	<Q^2_{CV}	passed
	cR^2_P	$cR^2_P = R^* \sqrt{R^2 - (average R_{Rand})^2}$	0.712	>0.5	passed
	R^2_{test}	$R^2_{test} = 1 - \dfrac{\sum (Y_{pred(test)} - Y_{obs(test)})^2}{\sum (Y_{obs(test)} - \overline{Y}_{pred(train)})^2}$	0.941	>0.6	passed
	$\overline{R}^2_{m(test)}$	$\dfrac{|R^2_m - R'^2_m|}{2}$	0.722	>0.5	passed
	ΔR^2_{test}	$|R^2_m - R'^2_m|$	0.012	<0.2	passed
	$\Delta R^2_{0(test)}$	$|R^2_0 - R'^2_0|$	0.006	<0.3	passed
external validation	$\dfrac{R^2 - R^2_0}{R^2}$	$\dfrac{R^2 - R^2_0}{R^2}$	0.054	<0.1	passed
	$\dfrac{R^2 - R'^2_0}{R^2}$	$\dfrac{R^2 - R'^2_0}{R^2}$	0.060	<0.1	passed
	k	$k = \dfrac{\sum Y_{obs}Y_{pred}}{\sum Y^2_{pred}}$	0.987	$0.85 \leq k \leq 1.15$	passed
	k'	$k' = \dfrac{\sum Y_{obs}Y_{pred}}{\sum Y^2_{obs}}$	1.009	$0.85 \leq k' \leq 1.15$	passed

Table 2 illustrates that the MLR model consistently aligns with the benchmarks set by Golbraikh and Tropsha, showcasing its dependability. The suggested model exhibits strong predictive accuracy for novel compounds.

In our QSAR analysis, we found that the boiling point serves as a significant descriptor correlated with the biological activity of quinazoline derivatives. Boiling point, reflecting the thermal stability and molecular size of a compound, may indicate a greater propensity for strong interactions at biological binding sites. Compounds with higher boiling points could demonstrate stronger intermolecular interactions, leading to enhanced biological activity against EGFR receptors.

Our analysis indicates that the structural properties of quinazoline derivatives are crucial in determining their biological activity. For instance, compounds featuring nitrogen or fluorine groups at specific positions on the quinazoline ring exhibit a higher affinity for binding. This suggests that such functional groups may form key interactions with the active sites of EGFR, thereby enhancing inhibitory effectiveness. Moreover, detailing how these structural modifications contribute to activity could include a discussion on electron-donating or withdrawing effects and their impact on molecular stability and binding affinity. By examining the most active compounds against less active ones, it becomes evident which structural elements are vital for enhanced efficacy, enabling us to propose targeted modifications for future compound design.

The experimentally determined and theoretical pIC50 values displayed in Figure 1 exhibit a strong correlation. The QSAR model obtained is based on two descriptors: energy gap and boiling point. According to the coefficient normalization diagram (Figure S1), in the QSAR model, E_{Gap} emerges as the most significant coefficient, where an increase in its value correlates with enhanced activity of the compound.

The variance inflation factor (VIF) is calculated using the formula $VIF = 1/(1 - R^2)$, where R represents the multiple correlation coefficient of an independent variable relative to all other descriptors in the model. A VIF greater than 10 indicates instability in the model,[64] necessitating its exclusion, while a VIF between 1 and 9 is considered acceptable. Table 3 reveals that the VIF values of all two descriptors are below 10. This suggests an absence of collinearity among the chosen descriptors, affirming the model's stability.

3.5. Y-Randomization. A randomization test was performed on the anticancer activity values (pIC50) to validate the developed QSAR model and ensure there were no random correlations.[65] In this process, the pIC50 values were randomly distributed a hundred times while keeping the two descriptors constant. The results of this test are compiled
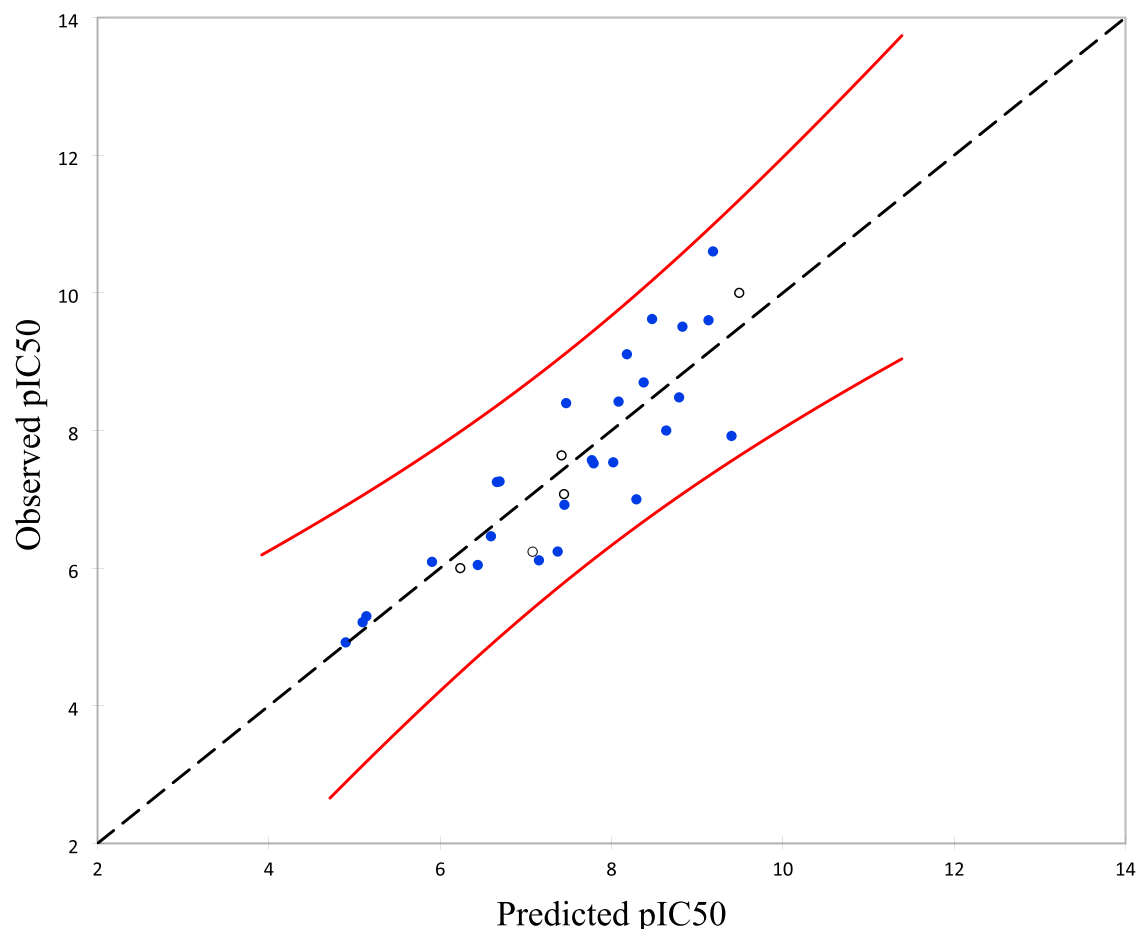
Figure 1. Relationship between observed and predicted activities.

Table 3. Variance Inflation Factors (VIF) of Descriptors in the QSAR Model

descriptors	E_{Gap}	boiling point
tolerance[a]	0.897	0.897
VIF	1.115	1.115

[a]VIF is the inverse of tolerance.

in Table S4. Given that the obtained value of cR^{2}p is 0.712 (greater than 0.5), it is concluded that no randomized test could match the original model. This affirms the robustness of the developed QSAR models and confirms the absence of any random correlation.

3.6. Application Domain. The Williams plot is generated by plotting standardized residual values along the Y-axis against leverage values on the X-axis, as illustrated in Figure 2. With the training set comprising two descriptors and 31 compounds, and the limits for normalized residuals defined as $r \pm 3$, the critical leverage value (h^*) is determined to be $[3(2 + 1)]/31 = 0.29$. The objective of defining the applicability domain (AD) is to identify compounds that deviate from the AD of the constructed QSAR model. Compounds that surpass the critical leverage value (h^*) of 0.29 are deemed unacceptable due to the inaccurate prediction of their activity. Using MATLAB V2021a software,[66] the AD was delineated based on the Williams plot in this study. According to the critical leverage

values ($h^* > 0.29$), no compounds were found to be outside the AD of the developed model.

3.7. Design of New Compounds. The proposed model demonstrates that by modifying the descriptors present in the QSAR model, the prediction of activity can be enhanced. The descriptors from the finalized model facilitate data interpretation, allowing for a clearer understanding of the relationship between descriptor values and the structural characteristics of the examined compounds.

In pursuit of discovering new compounds, specific structural modifications were made to the compounds under study. The influence of descriptors on biological activity, coupled with the structural features of the most active compounds, guided the suggestion of several new entities. ChemOffice software was employed for sketching and optimizing the structures of these proposed compounds, while ChemOffice and Gaussian software were utilized to ascertain the relevant descriptors.

Table S5 consolidates the structures of the proposed compounds and their corresponding predicted pIC50 values. An inference drawn from Table S5 is that an increase in E_{Gap} and boiling point values enhances the anticancer activity of the compounds in the data set. Armed with this knowledge, modifications were carried out around the quinazoline ring of the template (compound **27**).

During this stage, 18 compounds were conceptualized and their leverage values were computed to identify any outliers.
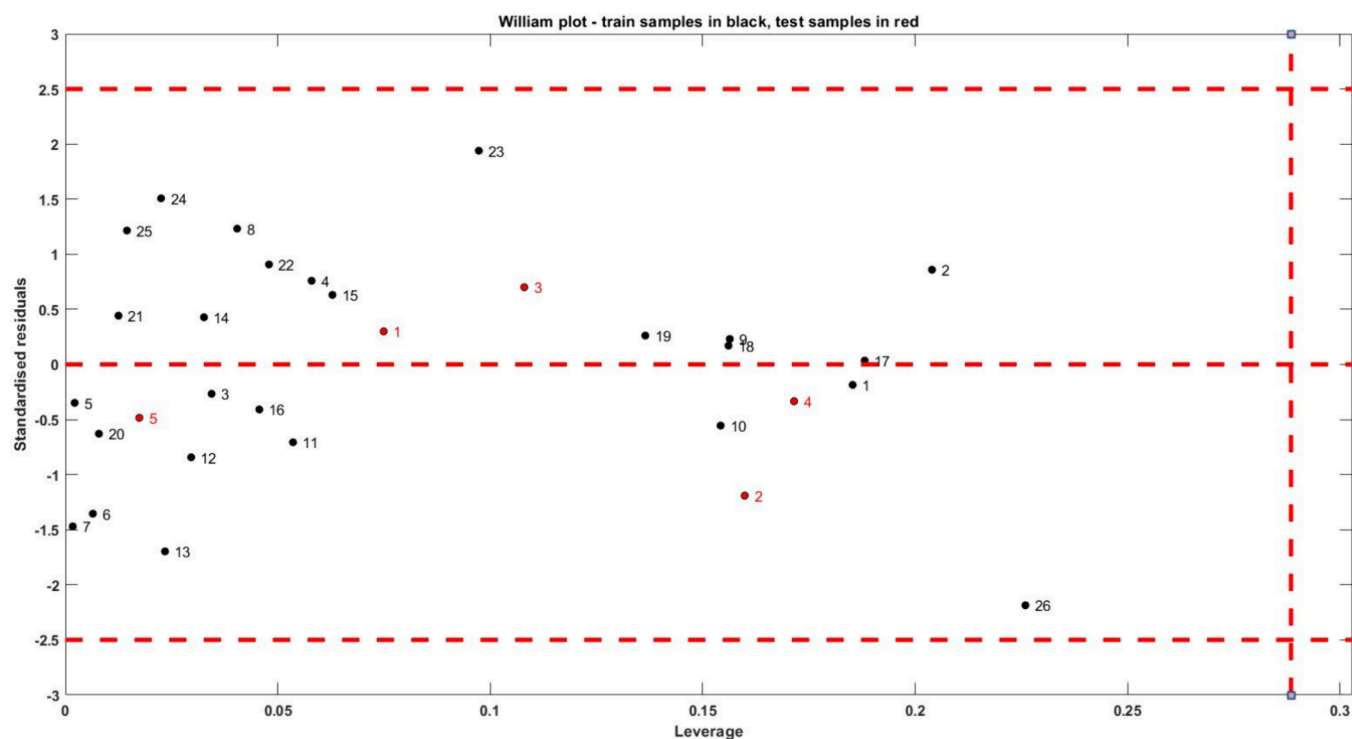
Figure 2. Williams plot analysis: standardized residuals versus leverage in the MLR model ($h* = 0.288$ and residual limits = ± 3).

Table 4. Chemical and Physical Attributes of the 18 Ligands

ligand	molecular weight	rotatable bonds	H-bond acceptors	H-bond donors	LogP	Lipinski violations
Pred1	390.23	5	6	1	3.28341	0
Pred2	378.20	4	6	1	3.54219	0
Pred3	394.65	4	5	1	4.00291	0
Pred5	420.26	6	7	1	2.92745	0
Pred6	436.31	5	5	1	5.32086	1
Pred7	375.23	4	6	2	2.56403	0
Pred8	390.24	4	7	3	1.33122	0
Pred9	388.26	6	5	1	4.23672	0
Pred10	374.24	5	5	1	3.80883	0
Pred12	422.28	5	5	1	4.96677	0
Pred14	350.22	2	3	1	5.10124	1
Pred15	394.27	4	4	1	5.09045	1
Pred16	436.31	5	5	1	5.32086	1
Pred17	482.38	5	4	1	4.61675	0
Pred18	450.34	6	5	1	5.74875	1
Pred19	462.39	5	4	1	5.28979	1
Pred20	498.38	6	5	1	4.90669	0
Pred21	434.29	3	5	1	4.38704	0

Equation 11 was used for determining the pIC50 values, and Table S5 displays the structures, pIC50 values, and leverage values of the compounds.

3.8. Lipinski Rules. Lipinski's "Rule of Five," introduced in 1997, serves as a fundamental framework for evaluating the drug-likeness or pharmacokinetic characteristics of chemical substances.[67] It holds significant importance within the pharmaceutical sector, assisting in the determination of a compound's viability as an orally active pharmaceutical.[68]

Lipinski's Rule involves the following parameters:

- Molecular weight (MW): should be less than 500 g/mol.
- Number of hbond donors (NHD): should not exceed 5, commonly involving OH and NH groups.

- Number of hbond acceptors (NHA): should not exceed 10, typically involving O and N atoms.
- Lipophilicity (LogP): the LogP value should be less than 5.
- Number rotatable bonds (N_{Rot}): should be less than 10.

We utilized the SwissADME server (http://www. SwissADME.ch/)[69] to evaluate several pharmacokinetic parameters, such as molecular weight, LogP partition coefficients, hydrogen donors and acceptors, rotatable bonds, and adherence to Lipinski's rules. The findings, outlined in Table 4, assist in exploring the pharmacokinetic characteristics of new drugs.

Table 5. Predicted Properties ADMET In Silico for the New Designed Quinazoline Derivatives

property	Pred1	Pred2	Pred3	Pred5	Pred6	Pred7	Pred8	Pred9	Pred10	Pred12	Pred14	Pred15	Pred16	Pred17	Pred18	Pred19	Pred20	Pred21
absorption																		
water solubility	-4.635	-4.572	-4.682	-4.785	-5.129	-3.78	-3.419	-4.605	-4.58	-5.056	-5.348	-4.703	-5.129	-3.47	-5.245	-4.122	-3.464	-4.362
human intestinal absorption	92.546	91.846	91.33	93.253	95.297	92.354	83.769	91.861	94.069	94.895	94.042	90.68	95.297	95.975	94.878	92.756	94.775	96.615
Caco-2 permeability	1.21	1.366	1.389	1.312	1.18	0.927	0.266	1.436	1.332	1.228	1.555	1.423	1.18	1.382	1.168	1.479	1.429	1.046
distribution																		
BBB permeability (log BB)	0.197	0.305	0.232	-0.406	-0.308	-0.057	-1.499	0.233	-0.277	-0.328	0.507	0.178	-0.308	0.044	-0.324	0.021	0.18	0.103
CNS permeability (log PS)	-2.208	-2.088	-1.974	-3.214	-1.761	-2.327	-2.456	-2.108	-2.055	-1.745	-0.834	-1.091	-1.761	-0.364	-1.77	-0.452	-1.471	-1.664
metabolism, cytochrome P450																		
substrate																		
2D6	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
3A4	no	no	no	no	yes	no	no	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
inhibitor																		
1A2	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	no	yes
2C19	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
2C9	yes	yes	yes	yes	yes	yes	yes	yes	no	yes	yes	no	yes	yes	yes	yes	yes	yes
2D6	no	no	no	no	no	no	no	yes	no	no	no	no	no	no	no	yes	no	no
3A4	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
excretion																		
total clearance	0.675	0.491	0.577	0.766	0.465	0.491	0.322	0.291	0.335	0.151	-0.003	0.168	0.465	0.346	0.418	0.285	0.25	0.214
toxicity																		
AMES toxicity	no	no	no	no	yes	no	no	no	no	yes	yes	no	yes	no	yes	yes	no	no
skin sensitization	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no

3.9. In Silico Pharmacokinetics ADMET Properties. In pharmacokinetics, the ADMET properties of a substance are paramount. These properties are analyzed using the pkCSM online tool[47] (Table 5). Water solubility is categorized based on its log (mol/L) values, ranging from "insoluble" to "very soluble" (insoluble ≤10 < poorly soluble ≤6 < moderately soluble ≤4 < soluble ≤2 < very soluble <0 < very soluble).[41] All studied compounds exhibited high solubility in water. For intestinal absorbance, values above 90% signify excellent absorbance. Most selected compounds exhibited values exceeding this threshold.

Caco-2 permeability is a prevalent metric for predicting the absorption of orally administered drugs.[70] A compound is deemed highly permeable if its Caco-2 value exceeds 0.90.[71] Nearly all selected compounds, except Pred8, met this criterion.

BBB permeability is crucial, determining whether a compound can impact the brain.[72] A value above 0.3 is considered favorable for BBB permeability.[48] Among the compounds studied, only compounds Pred2 and Pred15 demonstrated satisfactory BBB permeability. For a compound to penetrate the CNS, its LogPS value should exceed −2.[73] The cytochrome P450 isoenzymes, specifically CYP2D6 and CYP3A4,[74] are instrumental in drug metabolism. All tested compounds were identified as substrates of CYP3A4, but no compound becomes a substrate of CYP2D6, and a majority were inhibitors of CYP1A2, CYP2C19, and CYP2C9. In terms of drug clearance, a lower value indicates an extended drug half-life.[75] All compounds presented low clearance values, suggesting prolonged half-lives. Toxicity was assessed using the AMES Toxicity Test and the Skin Sensitization Test. A majority of compounds were deemed nontoxic, with exceptions such as Pred5, Pred12, Pred14, Pred16, Pred18, and Pred19.

In conclusion, compounds Pred15, Pred17, Pred20, and Pred21 are recommended as potential anticancer inhibitors due to their favorable ADMET properties, including effective absorption, distribution, metabolism, reduced clearance, and lack of toxic attributes.

3.10. Molecular Studies. This study involves a docking process aimed at identifying the optimal position of the ligand, quinazoline derivatives as tyrosine kinase (EGFR), within the receptor-binding site (1M17). The latter has been obtained from the PDB databank with a resolution of 2.60 Å.[54] Before being utilized by AutoDock, the complex is dissociated from its ligand, thereby freeing the active site. This is followed by the removal of water molecules, and the prepared files are subsequently saved in the pdbqt format.

AutoDock facilitates a quicker energy evaluation of the system by employing a three-dimensional grid, which extensively covers the active site of the 1M17 protein. This allows for unrestricted rotation of the ligand within the site. In this study, the center of the grid box is established at coordinates $X = 22.014$, $Y = 0.253$, and $Z = 52.794$, with dimensions of $40 \times 40 \times 40$ A^3, ensuring that the box is centered around the active site and that its dimensions align proportionally with the size of the studied ligands.

After generating the protein and ligand files, the docking process is initiated using the Genetic Algorithm (GA) with standard settings. The results can be visualized through the Discovery Studio 2021 Client software. The most favorable docking outcome is identified by the lowest energy conformation, which in this case is −10.4 kcal/mol.

Visual analysis is crucial for assessing the software's accuracy. Figure 3 illustrates that the ligand model produced by
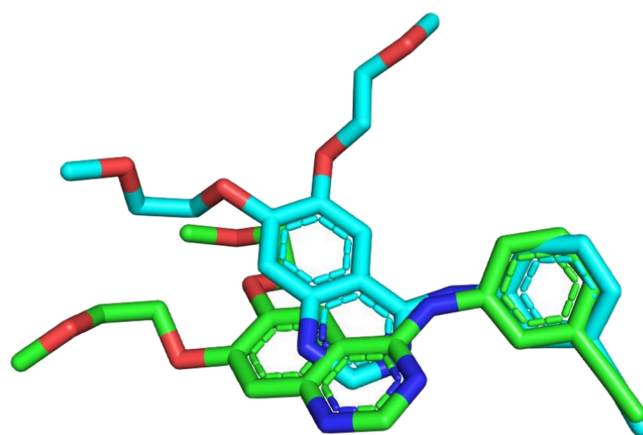

Figure 3. Redocking configuration displaying an RMSD of 1.223 Å (native ligand in green, docked ligand in blue).

AutoDock is precisely positioned in the active site of 1M17, displaying a spatial conformation that is very similar to, or even indistinguishable from, the one experimentally determined through crystallography, as recorded in the PDB. The cocrystallized ligand in the receptor was utilized as the reference drug.[54] The best obtained poses were examined by visualizing the most stable complex created using Discovery Studio.

Informed by ADMET analyses, molecular docking was subsequently conducted specifically on the designed molecules Pred15, Pred17, Pred20, and Pred21. These interactions are visually represented and analyzed using Discovery Studio 2021 software, as shown in Figure 4. Optimal results, corresponding to the minimal energies, have been identified for compounds Pred15 (−9.3 kcal/mol), Pred17 (−10.4 kcal/mol), Pred20 (−10.2 kcal/mol), and Pred21 (−10.2 kcal/mol).

Using Discovery Studio 2021 software, we visualized the interactions between the cocrystallized ligand and protein 1M17 (refer to Figure 4). The ligand formed two conventional hydrogen bonds with Cys 773 and one carbon hydrogen bond with Phe 771. It also established one pi−cation bond with Lys 721, one pi−sigma bond with Leu 820, and one pi−sulfur bond with Met 742. Hydrophobic interactions were also observed with several residues, including Lys 721, Leu 764, Val 702, Ala 719, Leu 820, and Cys 773.

Based on Figure 4 and Table S6. Compound Pred15 interacted through various bonds: one carbon hydrogen bond with Gln 767, two pi−anion bonds with Asp 831, and two pi−sigma bonds with Val 702. It also formed one pi−pi stacked bond with Phe 699 and exhibited hydrophobic interactions with residues such as Ala 719, Lys 721, Leu 764, Arg 817, Met 769, and Leu 820.

Compound Pred17 established one conventional hydrogen bond with Met 769 and two carbon hydrogen bonds with Gln 767 and Asp 831. It also formed three pi−sigma bonds with Leu 694, Val 702, and Leu 820, and one pi−sulfur bond with Met 742. Furthermore, hydrophobic interactions were noticed with residues including Leu 694, Lys 704, Leu 768, Ala 719, Leu 820, Val 702, and Lys 721.

Compound Pred20 exhibited one conventional hydrogen bond with Met 769, two carbon hydrogen bonds with Gln 767 and Asp 831, and three pi−sigma bonds with Leu 694, Val 702,
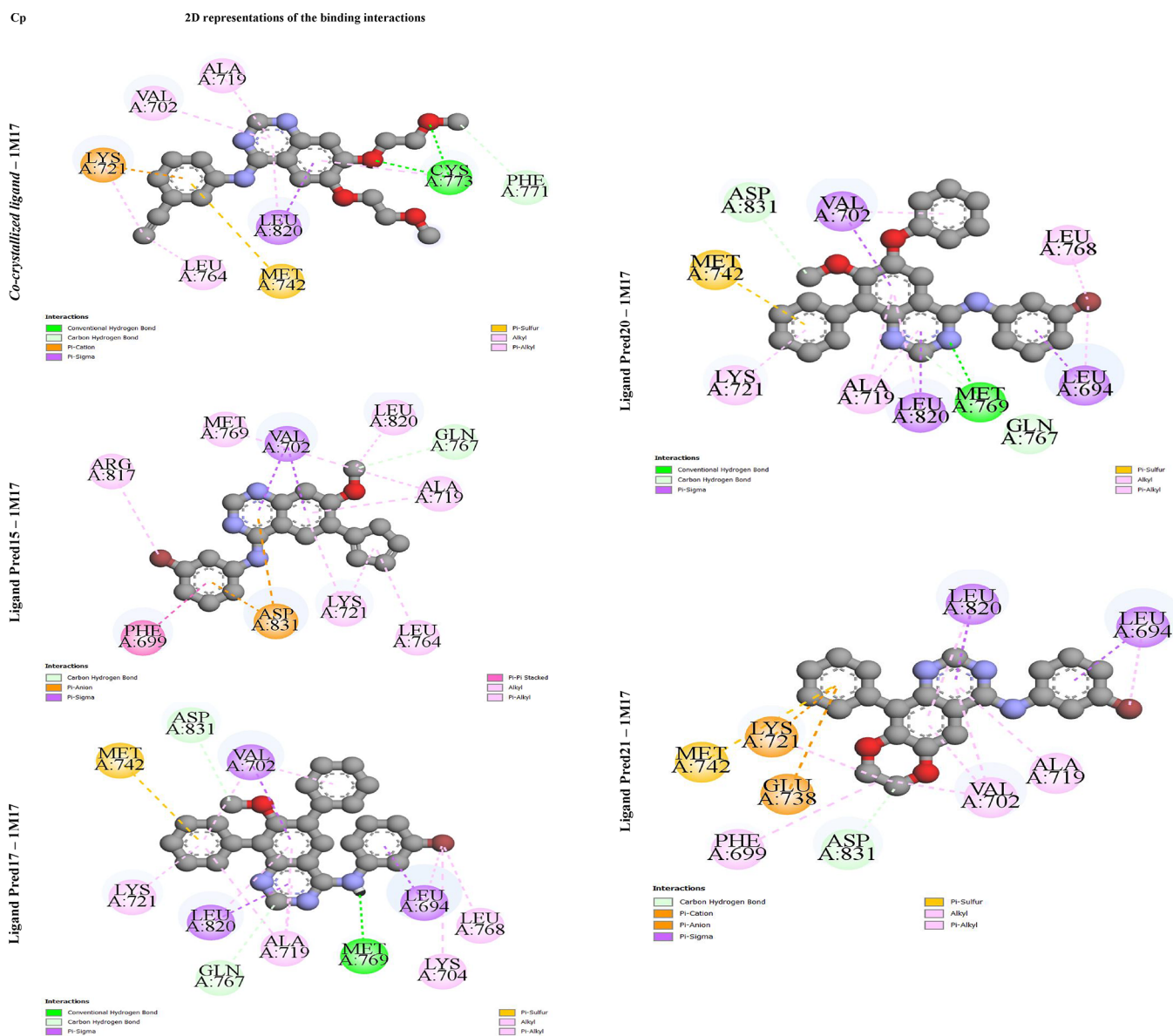
Figure 4. Two-dimensional representations of binding interactions in five complex compounds.

and Leu 820. Additionally, hydrophobic interactions were noted with residues such as Leu 694, Leu 768, Ala 719, Leu 820, Lys 721, and Val 702.

Compound Pred21 displayed various interactions: one carbon hydrogen bond with Asp 831, one pi—cation bond with Lys 721, one pi—anion bond with Glu 738, two pi—sigma bonds with Leu 694 and Leu 820, and one pi—sulfur bond with Met 742. Hydrophobic interactions were also observed with residues such as Val 702, Lys 721, Leu 694, Phe 699, Ala 719, and Leu 820.

3.11. Molecular Dynamics Simulation. Throughout this research, a molecular dynamics simulation (MDS) was performed over 100 ns to explore the binding mechanism, structural dynamics, and flexibility exhibited by EGFR when engaged with the three leading hits (Pred17, Pred20, and Pred21). Critical parameters such as RMSD, root mean square fluctuation (RMSF), and radius of gyration (Rg) were derived from the trajectory data of the 100 ns dynamics.[76]

RMSD Calculation. The stability of a biomolecular system is typically evaluated using its RMSD. Within computer-aided

drug design, systems with lower RMSD values are usually considered more stable whereas systems with higher RMSD values are seen as less stable.[77] The examination of the RMSD plot shown in Figure 5A reveals that all protein—ligand complexes demonstrate decreased variability in their spectrum, indicating a minimal disturbance in their conformational behavior during the simulation. The average RMSD values for Pred17, Pred20, Pred21, and erlotinib (with their receptors) are 1.13, 0.95, 1.02, and 1.03 nm, respectively. This result shows that Pred17, Pred20, and Pred21 are most like the standard drug erlotinib and may be the most stable drug candidates among the promising inhibitors.

Interestingly, these findings align with the docking simulation, which revealed that these compounds had the best binding affinities (see Figure 5).

RMSF Calculation. To investigate the flexibility of specific residues that might have influenced the system's overall variations, the RMSF of the complex was determined.[78] All proteins' backbone RMSFs were estimated using the GROMACS "gmx rmsf" command line. For biomolecular
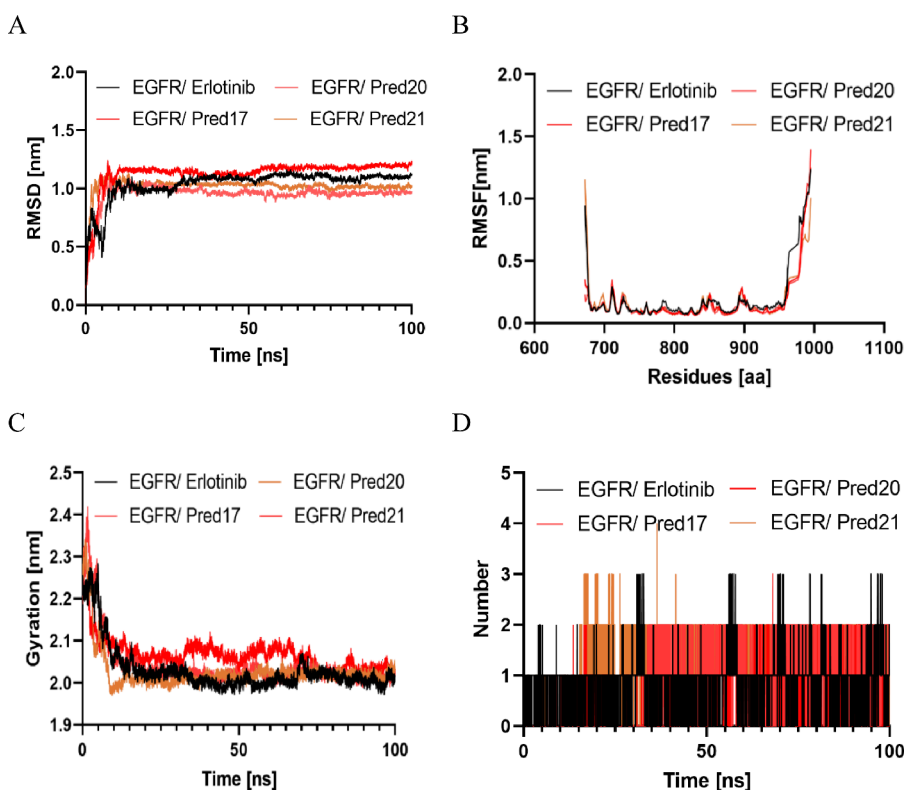
Figure 5. Results of the molecular dynamics study. (A) Time evolution of the backbone of three best hits inhibitors and the standard (Erlotinib). (B) RMSF spectra of three promising inhibitors and the standard (Erlotinib). (C) Comparative radius of gyration values for the target protein with three best hits inhibitors and the standard (Erlotinib). (D) Comparative hydrogen bonds for the target protein with the reference drug (Erlotinib) and three best hits.

Table 6. MM-PBSA Calculations of Binding Free Energy for All the Complexes

complex	binding energy (kJ/mol)	SASA energy (kJ/mol)	polar solvation energy (kJ/mol)	electrostatic energy (kJ/mol)	van der Waals energy (kJ/mol)
EGFR−Pred17	−101.007 ± 15.086	−16.719 ± 1.923	90.162 ± 25.842	−30.487 ± 8.109	−143.963 ± 17.430
EGFR−Pred20	−99.311 ± 19.158	−16.947 ± 2.246	90.765 ± 32.120	−25.374 ± 11.675	−147.755 ± 22.875
EGFR−Pred21	−99.467 ± 22.352	−17.985 ± 1.858	110.260 ± 26.860	−27.457 ± 12.246	−164.285 ± 21.439
EGFR−Erlotinib	−76.570 ± 16.232	−20.616 ± 1.238	167.175 ± 22.949	−39.915 ± 12.776	−183.214 ± 12.605

systems, larger RMSF values indicate less residual stability, similar to the RMSD measure, whereas lower values indicate more stability.[79,80] Figure 5B displays the results showing that all the proposed compounds have an RMSF pattern that is quite close to the standard medicine. The average RMSF values for Pred17, Pred20 and Pred21, and Erlotinib are 0.16, 0.15, 0.16, and 0.18 nm, respectively. Once again, these proposed compounds with the lowest average RMSF values stands out as the most effective inhibitor.

Radius of Gyration Analysis. To monitor variations in compactness and overall dimensions of the structures throughout the MD simulation, the Rg values of the protein−ligand complexes were evaluated. A consistent Rg value typically signifies a stable, folded structure, whereas fluctuations in the Rg value during the simulation indicate an unfolding of the structure.[81] Figure 5C presents the Rg spectra of the complexes, showcasing the similarity between the most promising proposed compounds and erlotinib throughout the entire simulation. This similarity is consistent with the findings from previous analyses, including docking, RMSD, and RMSF. While other compounds also exhibit comparable patterns in their graphs, Pred20 and Pred21 demonstrate the closest

resemblance. The average Rg values for Pred17, Pred20, Pred21, and Erlotinib are 2.06, 2.04, 2.02, and 2.03 nm, respectively. Notably, the Pred20−EGFR complex displays lower Rg values compared to all other small-molecule compounds, suggesting that it is the most compact biomolecular system.

H-Bond Analysis. In a water environment, hydrogen bonds and their respective strengths are critical in the interaction between proteins and ligands, particularly in scenarios where the mechanism of action involves hydrolysis, in which water plays an important role in chemical breakdown.[82] The process of hydrogen bond formation involves the interaction of two electronegative atoms, one from a hydrogen bond acceptor and one from a hydrogen bond donor.[83]

In this study, the intermolecular hydrogen bonds formed between EGFR and the compounds (Pred17, Pred20 and Pred21, and Erlotinib) were investigated. The results are presented in Figure 5D. Based on our observations, Pred20, Pred21, and Erlotinib exhibit the most pronounced and extensive spectrum of hydrogen bonds among all simulated entities. These compounds have an average of 0.83, 1.48, and 1.03 hydrogen bonds, respectively. However, when compared

to Pred17, it is evident that this compound forms fewer intermolecular hydrogen bonds with their receptors, averaging only 0.53 hydrogen bond, respectively. It is worth noting that the previous analysis, which utilized RMSD and RMSF metrics, identified Pred20 as the most effective compound. Interestingly, the results from the hydrogen bonding analysis further validate and reinforce this conclusion.

3.12. MM-PBSA Binding Free Energy. The MM-PBSA approach was employed to calculate the binding free energy (ΔE) between the EGFR−Erlotinib, EGFR−Pred17, EGFR−Pred20, and EGFR−Pred21 complexes using the MmPbStat.py script for whole trajectories.[61,65] The cumulative energies resulting from nonpolar, polar, and nonbonded interactions, including electrostatic interactions and van der Waals forces, were computed for each complex and are presented in Table 6. The molecules Erlotinib, Pred17, Pred20, and Pred21 all exhibit binding to EGFR with free energies of −76.570, −101.007, −99.311, and −99.467 kJ/mol, respectively. This provides evidence for the accuracy and reliability of the MDS model employed in this investigation. The binding energy of the four of these systems was influenced by the nonpolar solvation free energy (Enon polar), electrostatic energy (Eele), and van der Waals energy (Evdw). However, the polar energy (E polar) was unfavorable, highlighting the importance of the intermolecular van der Waals contribution. The observation of a significant interaction between the ligand and the hydrophobic binding pocket is in line with the findings from the docking studies and MD simulation interactions.

4. CONCLUSIONS

This study presents a QSAR analysis of 31 quinazoline derivatives and their effectiveness against EGFR in lung cancer cell lines. The developed models, RLM, are both internally and externally validated, demonstrating good predictive capabilities and significant statistical validity. The descriptors EGap and boiling point were found to correlate significantly with the inhibitory activity, where higher values are associated with increased effectiveness of the quinazoline derivatives as EGFR inhibitors. Furthermore, ADMET predictions, molecular docking, and dynamics analyses were conducted for four newly designed molecules based on structural modifications of the most active synthesized molecule. The results revealed that three compounds, identified as Pred17, Pred20, and Pred21, adhere to all ADMET criteria and exhibit notable interactions with the protein (PDB code: 1M17). MD simulations were performed on these designed compounds, indicating their promise as prospective candidates for continued investigation in drug development targeted at lung cancer.

ASSOCIATED CONTENT

Data Availability Statement

All available data can be found within the article and in the supplementary file. Further inquiries can be directed to the corresponding author.

Ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.4c04639.

Descriptors selected and software packages used in the calculation of descriptors; values of the 26 chemical descriptors; matrix illustrating the correlations between various derived descriptors; contribution of the descriptors in the model; results of Y randomization

tests for the model; descriptor values, predicted anticancer activity, and leverages (h) for newly designed quinazoline derivatives; and interaction table between the compounds and 1M17-protein cancer (PDF)

AUTHOR INFORMATION

Corresponding Authors

Mohamed Moussaoui − *Laboratory of Physical Chemistry of Materials, Faculty of Sciences Ben M'Sick, Hassan II University of Casablanca, Casablanca 20670, Morocco;* ⓞ orcid.org/0000-0002-3363-0864; Email: moussaouimohamed143@gmail.com

Rachid Daoud − *Chemical and Biochemical Sciences-Green Processing Engineering, Mohammed VI Polytechnic University, Ben Guerir 43150, Morocco;* Email: rachid.daoud@um6p.ma

Authors

Soukayna Baammi − *Bioinformatics Laboratory, College of Computing, Mohammed VI Polytechnic University, Ben Guerir 20670, Morocco*

Hatim Soufi − *Laboratory of Physical Chemistry of Materials, Faculty of Sciences Ben M'Sick, Hassan II University of Casablanca, Casablanca 20670, Morocco*

Mouna Baassi − *Laboratory of Physical Chemistry of Materials, Faculty of Sciences Ben M'Sick, Hassan II University of Casablanca, Casablanca 20670, Morocco*

Mohammed Salah − *Team of Chemoinformatics Research and Spectroscopy and Quantum Chemistry, Department of Chemistry, Faculty of Science, University Chouaib Doukkali, El Jadida 2300, Morocco*

Achraf EL Allali − *Bioinformatics Laboratory, College of Computing, Mohammed VI Polytechnic University, Ben Guerir 20670, Morocco*

Belghiti Elalaoui Mohammed − *Laboratory of Physical Chemistry of Materials, Faculty of Sciences Ben M'Sick, Hassan II University of Casablanca, Casablanca 20670, Morocco; Laboratory of Nernest Technology, Sherbrook, Quebec J1H5C7, Canada*

Said Belaaouad − *Laboratory of Physical Chemistry of Materials, Faculty of Sciences Ben M'Sick, Hassan II University of Casablanca, Casablanca 20670, Morocco*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acsomega.4c04639

Author Contributions

#M.M. and S.B. contributed equally to this work.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Li, X.; Zhang, Q.; Yang, Z. Knockdown of hsa_circ_0058124 inhibits the proliferation of human lung cancer cells by up-regulation of miR-1297. *Artif. Cells, Nanomedicine Biotechnol.* **2020**, *48*, 584−593.

(2) Xu, W. T.; Li, T. Z.; Li, S. M.; Wang, C.; Wang, H.; Luo, Y. H.; Piao, X. J.; Wang, J. R.; Zhang, Y.; Zhang, T.; Xue, H.; Cao, L. K.; Jin, C. H. Cytisine exerts anti-tumour effects on lung cancer cells by modulating reactive oxygen species-mediated signalling pathways. *Artif. Cells, Nanomedicine Biotechnol.* **2020**, *48*, 84−95.

(3) de Koning, H. J.; van der Aalst, C. M.; de Jong, P. A.; Scholten, E. T.; Nackaerts, K.; Heuvelmans, M. A.; Lammers, J. W. J.; Weenink, C.; Yousaf-Khan, U.; Horeweg, N.; van 't Westeinde, S.; Prokop, M.; Mali, W. P.; Mohamed Hoesein, F. A. A.; van Ooijen, P. M. A.; Aerts, J. G. J. V.; den Bakker, M. A.; Thunnissen, E.; Verschakelen, J.; Vliegenthart, R.; Walter, J. E.; ten Haaf, K.; Groen, H. J. M.; Oudkerk, M. Reduced Lung-Cancer Mortality with Volume CT Screening in a Randomized Trial. *N. Engl. J. Med.* **2020**, *382*, 503−513.

(4) Goebel, C.; Louden, C. L.; McKenna, R.; Onugha, O.; Wachtel, A.; Long, T. Diagnosis of Non-small Cell Lung Cancer for Early Stage Asymptomatic Patients. *Cancer Genomics and Proteomics.* **2019**, *16*, 229−244.

(5) Wu, H.; Zhou, C. Long non-coding RNA UCA1 promotes lung cancer cell proliferation and migration via microRNA-193a/HMGB1 axis. *Biochem. Biophys. Res. Commun.* **2018**, *496*, 738−745.

(6) Zhang, C.; Wang, X.; Wang, H.; Wu, X.; Shen, J. A positive-negative alternate adsorption effect for capacitive deionization in nano-porous carbon aerogel electrodes to enhance desalination capacity. *Desalination.* **2019**, *458*, 45−53.

(7) Cheng, M.; Jolly, S.; Quarshie, W. O.; Kapadia, N.; Vigneau, F. D.; Kong, F. M. Modern radiation further improves survival in non-small cell lung cancer: An analysis of 288,670 patients. *J. Cancer.* **2019**, *10*, 168−177.

(8) Arbour, K. C.; Riely, G. J. Systemic therapy for locally advanced and metastatic non-small cell lung cancer: A review. *JAMA - J. Am. Med. Assoc.* **2019**, *322*, 764−774.

(9) Dong, J.; Li, B.; Lin, D.; Zhou, Q.; Huang, D. Advances in Targeted Therapy and Immunotherapy for Non-small Cell Lung Cancer Based on Accurate Molecular Typing. *Front. Pharmacol.* **2019**, *10*, 1−10.

(10) Doroshow, D. B.; Sanmamed, M. F.; Hastings, K.; Politi, K.; Rimm, D. L.; Chen, L.; Melero, I.; Schalper, K. A.; Herbst, R. S. Immunotherapy in non-small cell lung cancer: Facts and hopes. *Clin. Cancer Res.* **2019**, *25*, 4592−4602.

(11) Da Dalt, G.; Friziero, A.; Grego, A.; Serafini, S.; Fassina, A.; Blandamura, S.; Sperti, C. Adrenal metastasis from endometrial cancer: A case report. *World J. Clin. Cases.* **2019**, *7*, 1844−1849.

(12) Funazo, T.; Nomizo, T.; Kim, Y. H. Liver Metastasis Is Associated with Poor Progression-Free Survival in Patients with Non−Small Cell Lung Cancer Treated with Nivolumab. *J. Thorac. Oncol.* **2017**, *12*, e140−e141.

(13) Schulz, M.; Salamero-Boix, A.; Niesel, K.; Alekseeva, T.; Sevenich, L. Microenvironmental Regulation of Tumor Progression and Therapeutic Response in Brain Metastasis. *Front. Immunol.* **2019**, *10*, 1713.

(14) Hiraga, T. Bone metastasis: Interaction between cancer cells and bone microenvironment. *J. Oral Biosci.* **2019**, *61*, 95−98.

(15) Liu, Z.; Liang, H.; Lin, J.; Cai, X.; Pan, Z.; Liu, J.; Xie, X.; Li, C.; Cheng, B.; Zhao, Y.; He, J.; Liang, W. The incidence of lymph node metastasis in patients with different oncogenic driver mutations among T1 non-small-cell lung cancer. *Lung Cancer.* **2019**, *134*, 218−224.

(16) Dako, F.; Hota, P.; Kahn, M.; Kumaran, M.; Agosto, O. Post-lung transplantation abdominopelvic complications: the role of multimodal imaging. *Abdom. Radiol.* **2020**, *45*, 1202−1213.

(17) Skamrova, G. B.; Laponogov, I.; Buchelnikov, A. S.; Shckorbatov, Y. G.; Prylutska, S. V.; Ritter, U.; Prylutskyy, Y. I.; Evstigneev, M. P. Interceptor effect of c60 fullerene on the in vitro action of aromatic drug molecules. *Eur. Biophys. J.* **2014**, *43*, 265−276.

(18) Zhang, P.; Tao, H.; Yu, L.; Zhou, L.; Zhu, C. Developing protein arginine methyltransferase 1 (PRMT1) inhibitor TC-E-5003 as an antitumor drug using INEI drug delivery systems. *Drug Delivery* **2020**, *27*, 491−501.

(19) Mukherjee, A.; Paul, M.; Mukherjee, S. Recent Progress in the Theranostics Application of. *Cancers* **2019**, *11*, 597.

(20) Wang, X.; Chen, H.; Zeng, X.; Guo, W.; Jin, Y.; Wang, S.; Tian, R.; Han, Y.; Guo, L.; Han, J.; Wu, Y.; Mei, L. Efficient lung cancer-targeted drug delivery via a nanoparticle/MSC system. *Acta Pharm. Sin. B* **2019**, *9*, 167−176.

(21) Prylutskyy, Y.I.; Evstigneev, M.P.; Cherepanov, V. V.; Kyzyma, O.A.; Bulavin, L.A.; Davidenko, N.A.; Scharff, P. Structural organization of C60 fullerene, doxorubicin, and their complex in physiological solution as promising antitumor agents. *J. Nanoparticle Res.* **2015**, *17*, 45.

(22) Prylutska, S.; Politenkova, S.; Afanasieva, K.; Korolovych, V.; Bogutska, K.; Sivolob, A.; Skivka, L.; Evstigneev, M.; Kostjukov, V.; Prylutskyy, Y.; Ritter, U. A nanocomplex of C60 fullerene with cisplatin: Design, characterization and toxicity. *Beilstein J. Nanotechnol.* **2017**, *8*, 1494−1501.

(23) Sivarajakumar, R.; Mallukaraj, D.; Kadavakollu, M.; Neelakandan, N.; Chandran, S.; Bhojaraj, S.; Reddy Karri, V. V. S. Nanoparticles for the treatment of lung cancers. *J. Young Pharm.* **2018**, *10*, 276−281.

(24) COHEN, N. C. The Molecular Modeling Perspective in Drug Design. *Guideb. Mol. Model. Drug Des.* **1996**, 1−17.

(25) Chatterjee, A. *Computational methods and tools for sustainable and green approaches in drug discovery*; Elsevier Inc., 2020. DOI: 10.1016/B978-0-12-817592-7.00027-7.

(26) Hardianto, A.; Yusuf, M.; Liu, F.; Ranganathan, S. Structure-based drug design workflow. *Encycl. Bioinforma. Comput. Biol. ABC Bioinforma.* **2019**, *1−3*, 273−282.

(27) Kleandrova, V. V.; Speck-Planche, A. PTML Modeling for Pancreatic Cancer Research In Silico Design of Simultaneous Multi-Protein and Multi-Cell Inhibitors. *Biomedicines* **2022**, *10*, 491.

(28) Kleandrova, V. V.; Scotti, M. T.; Scotti, L.; Speck-Planche, A. Multi-target Drug Discovery via PTML Modeling: Applications to the Design of Virtual Dual Inhibitors of CDK4 and HER2. *Curr. Top. Med. Chem.* **2021**, *21*, 661−675.

(29) Speck-Planche, A.; Scotti, M. T. BET bromodomain inhibitors: fragment-based in silico design using multi-target QSAR models. *Mol. Divers.* **2019**, *23*, 555−572.

(30) Speck-Planche, A. Combining Ensemble Learning with a Fragment-Based Topological Approach to Generate New Molecular Diversity in Drug Discovery: In Silico Design of Hsp90 Inhibitors. *ACS Omega.* **2018**, *3*, 14704−14716.

(31) Speck-Planche, A.; Cordeiro, M. N. D. S. Fragment-based in silico modeling of multi-target inhibitors against breast cancer-related proteins. *Mol. Divers.* **2017**, *21*, 511−523.

(32) Marzaro, G.; Chilin, A.; Guiotto, A.; Uriarte, E.; Brun, P.; Castagliuolo, I.; Tonus, F.; González-Díaz, H. Using the TOPS-MODE approach to fit multi-target QSAR models for tyrosine kinases inhibitors. *Eur. J. Med. Chem.* **2011**, *46*, 2185−2192.

(33) Bridges, A. J.; Zhou, H.; Cody, D. R.; Rewcastle, G. W.; McMichael, A.; Showalter, H. D. H.; Fry, D. W.; Kraker, A. J.; Denny, W. A. Tyrosine kinase inhibitors. 8. An unusually steep structure-activity relationship for analogues of 4-(3-bromoanilino)-6,7-dime-thoxyquinazoline (PD 153035), a potent inhibitor of the epidermal growth factor receptor. *J. Med. Chem.* **1996**, *39*, 267−276.

(34) *ChemOffice*; PerkinElmer Informatics 2020., (n.d.). http://www.cambridgesoft.com.

(35) Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G.A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H.P.; Izmaylov, A.F.; Bloino, J.; Zheng, G.; Sonnenberg, J.L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J.A.; Peralta, J.E.; Ogliaro, F.; Bearpark, M.; Heyd, J.J.; Brothers, E.; Kudin, K.N.; Staroverov, V.N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J.C.; Iyengar, S.S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J.M.; Klene, M.; Knox, J.E.; Cross, J.B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R.E.; Yazyev, O.; Austin, A.J.; Cammi, R.; Pomelli, C.;

Ochterski, J.W.; Martin, R.L.; Morokuma, K.; Zakrzewski, V.G.; Voth, G.A.; Salvador, P.; Dannenberg, J.J.; Dapprich, S.; Daniels, A.D.; Farkas, J.B. Foresman; Ortiz, J. V.; Cioslowski, J.; Fox, D.J., *Gaussian 09*, Revision B.01, Gaussian 09, Revis. B.01; Gaussian, Inc.: Wallingford CT. (2009) 1−20. citeulike-article-id:9096580.

(36) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B. Condens. Matter.* **1988**, *37*, 785−789.

(37) Jollife, I.T.; Cadima, J. Principal component analysis: A review and recent developments. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **2016**, 45.

(38) Kerraj, S.; Salah, M.; Chtita, S.; El Idrissi, M.; Belaaoud, S.; Mohammed, M.; Acharjee, N.; Komiha, N. Theoretical study of photovoltaic performances of Ru, Rh and Ir half sandwich complexes containing N,N chelating ligands in Dye-Sensitized Solar Cells (DSSCs). DFT and TD-DFT investigation. *Comput. Theor. Chem.* **2022**, *1209*, No. 113630.

(39) Kerraj, S.; Salah, M.; Belaaouad, S.; Mohammed, M. Effects of chelate ligands containing NN, PN, and PP on the performance of half-sandwich ruthenium metal complexes as sensitizers in dye sensitized solar cells (DSSCs): Quantum chemical investigation. *Polyhedron.* **2023**, *230*, No. 116190.

(40) Kerraj, S.; Harbi, A.; El Mecherfi, K.; Moussaoui, M.; Salah, M.; Belaaouad, S.; Mohammed, M. Computational analysis of ligand design for Ru half-sandwich sensitizers in bulk heterojunction (BHJ) solar cells: Exploring the role of − NO2 group position and π-conjugation in optimizing efficiency. *J. Indian Chem. Soc.* **2024**, *101*, No. 101148.

(41) Chtita, S.; Belhassan, A.; Bakhouch, M.; Taourati, A.I.; Aouidate, A.; Belaidi, S.; Moutaabbid, M.; Belaaouad, S.; Bouachrine, M.; Lakhlifi, T., QSAR study of unsymmetrical aromatic disulfides as potent avian SARS-CoV main protease inhibitors using quantum chemical descriptors and statistical methods, *Chemom. Intell. Lab. Syst. 210* (**2021**). 104266.

(42) Yali, Z. P.; Fatemi, M. H. Developing quantitative structure−retention relationship model to prediction of retention factors of some alkyl-benzenes in nano-LC. *J. Iran. Chem. Soc.* **2019**, *16*, 1545−1551.

(43) Gramatica, P. Principles of QSAR models validation: Internal and external. *QSAR Comb. Sci.* **2007**, *26*, 694−701.

(44) Projan, S. J. Why is big Pharma getting out of antibacterial drug discovery? *Curr. Opin. Microbiol.* **2003**, *6*, 427−430.

(45) Zhang, Y.; Liu, H.; Jiao, Y.; Yuan, H.; Wang, F.; Lu, S.; Yao, S.; Ke, Z.; Tai, W.; Jiang, Y.; Chen, Y.; Lu, T. De novo design of N-(pyridin-4-ylmethyl)aniline derivatives as KDR inhibitors: 3D-QSAR, molecular fragment replacement, protein-ligand interaction fingerprint, and ADMET prediction. *Mol. Divers.* **2012**, *16*, 787−802.

(46) Pires, D. E. V.; Blundell, T. L.; Ascher, D. B. pkCSM: Predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. *J. Med. Chem.* **2015**, *58*, 4066−4072.

(47) *pkCSM*, (n.d.). https://biosig.lab.uq.edu.au/pkCSM/prediction (accessed November 2, 2023).

(48) Zhu, L.; Zhao, J.; Zhang, Y.; Zhou, W.; Yin, L.; Wang, Y.; Fan, Y.; Chen, Y.; Liu, H. ADME properties evaluation in drug discovery: in silico prediction of blood−brain partitioning. *Mol. Divers.* **2018**, *22*, 979−990.

(49) Summerfield, S. G.; Read, K.; Begley, D. J.; Obradovic, T.; Hidalgo, I. J.; Coggon, S.; Lewis, A. V.; Porter, R. A.; Jeffrey, P. Central nervous system drug disposition: The relationship between in situ brain permeability and brain free fraction. *J. Pharmacol. Exp. Ther.* **2007**, *322*, 205−213.

(50) Garg, P.; Dhakne, R.; Belekar, V. Role of breast cancer resistance protein (BCRP) as active efflux transporter on blood-brain barrier (BBB) permeability. *Mol. Divers.* **2015**, *19*, 163−172.

(51) Han, Y.; Zhang, J.; Hu, C. Q.; Zhang, X.; Ma, B.; Zhang, P. In silico ADME and toxicity prediction of ceftazidime and its impurities. *Front. Pharmacol.* **2019**, *10*, 1−12.

(52) Trott, O.; Olson, A.J., AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient

optimization, and multithreading, *J. Comput. Chem.* (**2010**) NA-NA. 31455.

(53) *Download AutoDock4 − AutoDock*, (n.d.). https://AutoDock.scripps.edu/download-autodock4/ (accessed December 30, 2022).

(54) Stamos, J.; Sliwkowski, M. X.; Eigenbrot, C. Structure of the epidermal growth factor receptor kinase domain alone and in complex with a 4-anilinoquinazoline inhibitor. *J. Biol. Chem.* **2002**, *277*, 46265−46272.

(55) *Free Download: BIOVIA Discovery Studio Visualizer - Dassault Systèmes*, (n.d.). https://discover.3ds.com/discovery-studio-visualizer-download (accessed April 9, 2023).

(56) Al-Khafaji, K.; Taskin Tok, T. Molecular dynamics simulation, free energy landscape and binding free energy computations in exploration the anti-invasive activity of amygdalin against metastasis. *Comput. Methods Programs Biomed.* **2020**, *195*, No. 105660.

(57) Chen, K.; Sun, M.; Chen, H.; Lee, C.; Chen, C.Y. Potential Smoothened Inhibitor from Traditional Chinese Medicine against the Disease of Diabetes, Obesity, and Cancer. *BioMed Res. Int.* **2014**, *2014*, No. 873010.

(58) Grosdidier, L.; Michielin, O.; Zoete, V.; Cuendet, M.A. SwissParam: A Fast Force Field Generation Tool for Small Organic. *Molecules* **2011**, 2359.

(59) Mark, P.; Nilsson, L. Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K. *J. Phys. Chem. A* **2001**, 9954−9960.

(60) Gomari, M. M.; Rostami, N.; Omidi-ardali, H. Insight into molecular characteristics of SARS- CoV-2 spike protein following D614G point mutation, a molecular dynamics study. *J. Biomol. Struct. Dyn.* **2021**, *0*, 1−9.

(61) Baammi, S.; Daoud, R.; El Allali, A. In silico protein engineering shows that novel mutations affecting NAD+ binding sites may improve phosphite dehydrogenase stability and activity. *Sci. Rep.* **2023**, *13*, 1−18.

(62) XLSTAT version 2019.1, XLSTAT, Your data analysis solution. https://www.xlstat.com/fr/articles/xlstat-version-2019-1. - Recherche Google, (n.d.).

(63) Moussaoui, M.; Baassi, M.; Baammi, S.; Soufi, H.; Salah, M.; Daoud, R.; Allali, A. EL; Belghiti, M.; Belaaouad, S. In silico design of novel CDK2 inhibitors through QSAR, ADMET, molecular docking and molecular dynamics simulation studies. *J. Biomol. Struct. Dyn.* **2023**, 13646.

(64) Thompson, C. G.; Kim, R. S.; Aloe, A. M.; Becker, B. J. Extracting the Variance Inflation Factor and Other Multicollinearity Diagnostics from Typical Regression Results. *Basic Appl. Soc. Psych.* **2017**, *39*, 81−90.

(65) Baammi, S.; El Allali, A.; Daoud, R. Potent VEGFR-2 inhibitors for resistant breast cancer: a comprehensive 3D-QSAR, ADMET, molecular docking and MMPBSA calculation on triazolopyrazine derivatives. *Front. Mol. Biosci.* **2023**, No. 1288652.

(66) *MATLAB*, (n.d.). https://www.mathworks.com/products/matlab.html (accessed November 2, 2023).

(67) Giménez, B. G.; Santos, M. S.; Ferrarini, M.; Dos Santos Fernandes, J. P. Evaluation of blockbuster drugs under the rule-of-five. *Pharmazie.* **2010**, *65*, 148−152.

(68) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Delivery Rev.* **2012**, *64*, 4−17.

(69) Daina, A.; Michielin, O.; Zoete, V. SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci. Rep.* **2017**, *7*, 1−13.

(70) Baassi, M.; Moussaoui, M.; Soufi, H.; Rajkhowa, S.; Sharma, A.; Sinha, S.; Belaaouad, S.; Ghosh, A. Towards designing of a potential new HIV-1 protease inhibitor using QSAR study in combination with Molecular docking and Molecular dynamics simulations. *PLoS One.* **2023**, *18*, No. e0284539.

(71) Soufi, H.; Moussaoui, M.; Baammi, S.; Baassi, M.; Salah, M.; Daoud, R.; El Allali, A.; Belghiti, M. E.; Moutaabbid, M.; Belaaouad, S. Multi-combined QSAR, molecular docking, molecular dynamics

simulation, and ADMET of Flavonoid derivatives as potent cholinesterase inhibitors. *J. Biomol. Struct. Dyn.* **2023**, *0*, 1−15.

(72) Clark, D. E. In silico prediction of blood-brain barrier permeation. *Drug Discovery Today.* **2003**, *8*, 927−933.

(73) Suenderhauf, C.; Hammann, F.; Huwyler, J. Computational prediction of blood-brain barrier permeability using decision tree induction. *Molecules.* **2012**, *17*, 10429−10445.

(74) Zanger, U. M.; Schwab, M. Cytochrome P450 enzymes in drug metabolism: Regulation of gene expression, enzyme activities, and impact of genetic variation. *Pharmacol. Ther.* **2013**, *138*, 103−141.

(75) Hammarlund-Udenaes, M. The use of microdialysis in CNS drug delivery studies: Pharmacokinetic perspectives and results with analgesics and antiepileptics. *Adv. Drug Delivery Rev.* **2000**, *45*, 283−294.

(76) Broni, E.; Kwofie, S. K.; Asiedu, S. O.; Miller, W. A.; Wilson, M. D. A molecular modeling approach to identify potential antileishmanial compounds against the cell division cycle (Cdc)-2-related kinase 12 (crk12) receptor of leishmania donovani. *Biomolecules* **2021**, *11*, 458.

(77) Yang, S.; Kar, S. Protracted molecular dynamics and secondary structure introspection to identify dual-target inhibitors of Nipah virus exerting approved small molecules repurposing. *Sci. Rep.* **2024**, *14*, 1−14.

(78) Wang, S.; Xu, X.; Pan, C.; Guo, Q.; Li, Q.; Wan, S.; Li, Z.; Zhang, J.; Wu, X., Identification of New EGFR Inhibitors by Structure-Based Virtual Screening and Biological Evaluation, *Int. J. Mol. Sci. 25* (**2024**). 1887.

(79) Sargsyan, K.; Grauffel, C.; Lim, C. How Molecular Size Impacts RMSD Applications in Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2017**, *13*, 1518−1524.

(80) Baammi, S.; El Allali, A.; Daoud, R. Unleashing Nature' s potential: a computational approach to discovering novel VEGFR-2 inhibitors from African natural compound using virtual screening, ADMET analysis, molecular dynamics, and MMPBSA calculations. *Front. Mol. Biosci.* **2023**, No. 1227643.

(81) Liao, K.H.; Chen, K.B.; Lee, W.Y.; Sun, M.F.; Lee, C.C.; Chen, C. Y. C., Ligand-based and structure-based investigation for Alzheimer's disease from traditional Chinese medicine, *Evidence-Based Complement. Altern. Med.* **2014** (*2014*). 1.

(82) Ashiru, M. A.; Ogunyemi, S. O.; Temionu, O. R.; Ajibare, A. C.; Cicero-Mfon, N. C.; Ihekuna, O. A.; Jagun, M. O.; Abdulmumin, L.; Adisa, Q. K.; Asibor, Y. E.; Okorie, C. J.; Lawal, M. O.; Babalola, M. O.; Abdulrasaq, I. T.; Salau, L. B.; Olatunji, I. O.; Bankole, M. A.; Daud, A. B.; Adeyemi, A. O. Identification of EGFR inhibitors as potential agents for cancer therapy: pharmacophore-based modeling, molecular docking, and molecular dynamics investigations. *J. Mol. Model.* **2023**, *29*, 128.

(83) Gao, Y.; Mei, Y.; Zhang, J. Z. H. Treatment of Hydrogen Bonds in Protein Simulations. *Adv. Mater. Renew. Hydrog. Prod. Storage Util.* **2015**, 1854.