

In Silico Screening for Novel Tyrosine Kinase Inhibitors With Oxindole Scaffold as Anticancer Agents: Design, QSAR Analysis, Molecular Docking and ADMET Studies

Jalalaldin Zangeneh

Isfahan University of Medical Sciences

Pouria Shirvani

Isfahan University of Medical Sciences

Mahmoud Etebari

Isfahan University of Medical Sciences

Iofollah saghaie (✉ I.saghaie2021@gmail.com)

Isfahan University of Medical Sciences <https://orcid.org/0000-0002-6727-2962>

Research Article

Keywords: In Silico Screening, TKIs, QSAR, oxindole derivatives, tyrosine kinase, Molecular docking, ADMET analysis

Posted Date: December 6th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1121857/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Recently, Anti-cancer targeting drugs are directed against specific molecules and signaling pathways. These targeting agents have reasonable specificity, efficacy, and less side effects. Tyrosine kinases, which play an essential role in growth factor signaling regulation, are significant targets in this type of therapy. Synthesized numerous tyrosine kinase inhibitors (TKIs), such as substituted indolin-2-ones, are effective as anti-tumor and anti-leukemia agents.

In this study, a series of novel substituted indolin-2-ones were studied as kinase inhibitor analogs through quantitative structure-activity relationship (QSAR) analysis. Two chemometrics methods, such as multiple linear regression (MLR) and partial least squares combined with genetic algorithm for variable selection (GA-PLS), were employed to establish relationships between structural characteristics and kinase inhibitory activity of oxindole analogs. The GA-PLS was developed as the best predictor and validated QSAR model. The data set compounds were also studied by molecular docking to investigate their binding mechanism in the active site of tyrosine kinase enzymes. According to the information obtained from QSAR models and molecular docking analysis, 44 new potent lead compounds with novel structural features were introduced. Molecular docking, drug-likeness rules, ADMET analysis, bioavailability, toxicity prediction, and target identification were carried out on the newly designed oxindoles to elucidate the fundamental structural properties that affect their inhibitory activity. The results of our study could provide significant insight for future design and development of novel tyrosine kinase inhibitors.

1. Introduction

Targeted anticancer therapy has been one of the top developments in medicine in recent years [1]. Today, protein kinases are one of the significant drug targets in the field of chemotherapeutic antitumor drugs [2]. In recent years, more than 20 kinase-targeted drugs have been applied in chemotherapy, and hundreds of drug candidates are in pre-clinical and clinical trials [3]. Tyrosine kinase receptors play an essential role in cellular signal transduction pathways, modulating cellular responses to an external stimuli and affecting a variety of cellular processes including cell cycle progression, migration, cell metabolism, survival, proliferation and cell differentiation [4]. Due to their high expression in most cancer cells and their important role in modulating growth factor signaling in the cancer cell cycle, tyrosine kinases have been considered as one of the potential targets in the targeted therapy of cancers [2, 5]. Two tyrosine kinase receptors, vascular endothelial growth factor receptor1 (VEGFR1) and vascular endothelial growth factor receptor 2 (VEGFR2 or KDR), regulate the function of endothelial cells [6, 7]. Small molecule tyrosine kinase inhibitors (TKIs) have been successfully entering the drug market as selective anticancer agents with good potency, efficacy and low side effects [8]. Additionally, TKIs have good safety profiles and can be easily combined with other forms of chemotherapy or radiation therapy [9]. In recent years, several small molecules acting on tyrosine kinases such as Sorafenib, Sunitinib, Axitinib, and Regorafenib have been developed and approved for clinical use [10-14]. Many oxindoles or oxindole-like compounds have been documented in the literature as receptor tyrosine kinase (RTK) inhibitors. Among them, the scaffold of sunitinib has been extensively studied and the structure-activity relationships (SAR) of many of the indolin-2-one derivatives have been described [15-17]. As shown in Figure 1, Sunitinib structure does not have the phenyl group on its alkene substitution and because of that Sunitinib is inactive against FGFR1 [18].

Regarding such significant inhibitory activities of sunitinib on KDR and FGFR in the presence and absence of the phenyl ring substitution, there have been many efforts for developing SAR of sunitinib and its similar structures by changing the substituents of rings A, B, and C. From published SAR studies on indolin-2-ones, modifications at the C-5 position of C-ring resulted compounds with different kinase inhibition profiles against VEGFR2 enzyme [18, 19]. Therefore, substitution at this position can be used to improve the pharmacological characteristics of indolin-2-one derivatives. In this study, various substituents have been applied at different positions of rings A, B, and C of indolin-2-ones derivatives to improve their kinase inhibition spectrum.

In silico procedures have become a crucial part of drug discovery processes such as identify, re-design, and evaluate new candidate inhibitors [20]. Molecular modeling studies such as quantitative structure-activity relationship studies (QSARs) and molecular docking have been traditionally used as lead optimization approaches in the drug design field. Recently, some novel oxindole derivatives have been reported as selective inhibitors of VEGFR2 with an excellent binding affinity [18, 19]. Accordingly, this encouraged us to systematically probe the relationship between their structures and their inhibitory activities.

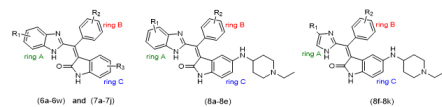
In this research, a novel series of oxindole analogs with broad spectrum kinase inhibitory activity was selected to perform a QSAR analysis [18]. Two different methods were applied to model the relationship between the structural properties and biological activity of the studied compounds: (1) multiple linear regression (MLR), and (2) partial least squared combined with genetic algorithm for variable selection (GA-PLS). The best model was obtained from GA-PLS, which is a linear six-parameter model. Additionally, the binding affinity, ligand orientation, and protein-ligand interactions of the designed compounds were revealed using the molecular docking technique. Drug-likeness rules like Lipinski's rule of five, bioavailability, toxicology prediction, and target identification can provide useful guidelines in early stage drug discovery. Accordingly, in this study, the useful drug-likeness descriptors were also calculated for selected ligands. Finally, from the final selected compounds, some novel potent hit compounds were chosen for future study. Hence, the results of the present study might provide helpful guidelines for the future design of potent tyrosine kinase inhibitors as potential anticancer agents.

2. Materials And Computational Methods

2.1. Data set preparation

In this study, a data set of 44 oxindole derivatives was collected from the literature (Kim MH et al.) to perform this study [18]. The chemical structures and kinase inhibitory activity (IC_{50}) of these 44 molecules against KDR are presented in Table 1. The IC_{50} values of the compounds were converted into the corresponding pIC_{50} ($-\log IC_{50}$) values and were used to develop QSAR models. To perform the QSAR studies, the whole data set (44 molecules) was randomly divided into a training set (35 molecules, 80%) for model generation and a test set (9 molecules, 20%) for external evaluation of the generated models.

Table 1: Chemical structure of oxindole derivatives used in this study and their inhibitory activity data



ID	R ₁	R ₂	R ₃	Exp. pIC ₅₀	Pred. pIC ₅₀	leverage	Docking binding energy (kcal/mol)
6a	H	<i>o</i> -Me	H	7.014	6.2326	0.19179	-9.21
6b	H	<i>m</i> -Me	H	6.925	6.82479	0.15652	-9.38
6c	H	<i>p</i> -Me	H	7.056	6.81717	0.07631	-9.53
6d	H	<i>p</i> -F	H	6.199	6.49173	0.06798	-9.18
6e	H	<i>m</i> -F	H	6.904	6.82258	0.04510	-9.07
6f	H	3,5-DiF	H	6.975	7.18270	0.50859	-9.01
6g	H	2,4-DiF	H	6.538	6.69447	0.18906	-9.05
6h	H	<i>p</i> -Cl	H	4.702	5.70999	0.12388	-8.32
6i	H	2,4-DiCl	H	5.662	5.99193	0.10963	-8.68
6j	H	<i>p</i> -OMe	H	6.486	6.14099	0.26971	-9.29
6k	H	<i>p</i> -OEt	H	5.006	6.59108	0.17548	-8.89
6l	H	<i>p</i> -NH ₂	H	7.168	6.76266	0.28827	-9.72
6m	H	H	5-Br	7.293	6.53302	0.07725	-9.86
6n	H	H	6-Br	5.312	5.05315	0.16712	-8.43
6o	H	H	7-Br	5.312	5.52642	0.13465	-8.59
6p	H	H	5-COOH	8.398	7.17493	0.19554	-10.47
6q	H	H	6-COOH	6.672	6.49197	0.21199	-9.46
6r	H	H	7-COOH	4.627	5.81252	0.20991	-7.34
6s	H	H	5-NH ₂	7.495	7.28627	0.12562	-9.23
6t	H	<i>p</i> -Me	5-COOH	8.301	7.78710	0.21304	-10.53
6u	H	<i>p</i> -Me	5-NH ₂	7.087	7.95460	0.30791	-9.15
6v	5- <i>o</i> -Me	H	5-COOH	8.301	7.40618	0.21851	-9.86
6w	5- <i>o</i> -Me	H	5-NH ₂	7.222	7.64815	0.23671	-9.45
7a	H	H		7.824	7.68209	0.32447	-10.36
7b	H	H		7.162	7.52954	0.23191	-9.87
7c	H	H		6.805	7.38354	0.09379	-9.89
7d	H	H		5.946	5.96758	0.22812	-9.29
7e	H	H		6.181	5.87699	0.06622	-9.85
7f	H	H		5.71	5.79548	0.09472	-9.58
7g	5- <i>o</i> -Me	H		7.638	7.02492	0.11105	-10.77
7h	5- <i>o</i> -Me	H		6.555	7.01229	0.05338	-10.32
7i	5- <i>o</i> -Me	H		8	7.32723	0.11230	-10.72
7j	H	H		8.046	7.56491	0.12181	-10.75
8a	H	H		8.097	7.14915	0.12339	-11.05
8b	H	<i>p</i> -Me		7.887	7.38248	0.13627	-10.98
8c	H	<i>p</i> -OMe		8.523	7.14691	0.19179	-10.46
8d	H	<i>m</i> -F		4.293	7.47357	0.15652	-7.53
8e	H	3,5-DiF		8.398	7.61822	0.07631	-10.91
8f	H	H		7.796	8.32140	0.06798	-9.63
8g	H	<i>m</i> -F		8.154	8.62573	0.04510	-9.88
8h	H	<i>p</i> -Cl		8.398	8.01331	0.50859	-10.23
8i	Me	<i>m</i> -F		8.398	8.66632	0.18906	-10.12
8j	Me	3,5-DiF		8.523	8.69498	0.12388	-10.21
8k	Me	<i>o</i> -F		8.523	8.33769	0.10963	-10.46

2.2. Descriptor preparation and Optimization methods

The 2D structures of all studied compounds were sketched employing ChemBio Draw Ultra 12.0 [20]. The energy minimization was carried out using molecular mechanics (MM⁺) followed by quantum-based semi-empirical (AM1) method employing Hyperchem 8.0 software [21]. The software Gaussian 98, Hyperchem 8.0, and Dragon (version 5.5) were used to calculate molecular descriptors [22]. For example, the highest occupied molecular orbital energy (E_{HOMO}), the lowest unoccupied molecular orbital energy (E_{LUMO}), as well as the total dipole moment (μ) were calculated by Gaussian 98. Furthermore, some chemical parameters including

molecular volume (V), hydrophobicity (logP), molecular surface area (SA), molecular polarizability (MP), hydration energy (HE) were computed employing Hyperchem 8.0. Dragon software provides a large domain of various descriptors such as constitutional, topological, empirical, 2D autocorrelation, geometrical, aromaticity index, atom-centered fragments, functional group counts, topological charge index, molecular properties and charge descriptors for each molecule [22]. Moreover, according to the equations proposed by Thanikaivelan et al., hardness, softness, electronegativity, and electrophilicity were calculated [23]. Some of these descriptors are depicted in Table S1 in the supplementary file.

2.3. Molecular docking procedure

Molecular docking analysis was carried out using an *in-house* batch script (DOCKFACE) to run AutoDock 4.2 and Vina in parallel mode automatically [24]. MGLTools 1.5.6 was used to convert the output structures to PDBQT [25]. The 3D crystal structure of EphB4, a surrogate receptor tyrosine kinase (PDB ID:4AW5), were obtained from the protein data bank (<http://www.rcsb.org/>) [12]. Regarding the preparation of the protein, co-crystallized ligand and water molecules were removed from the crystal structure using Accelrys DS Visualizer 4. Then, missing hydrogens were added, nonpolar hydrogen atoms were merged, Kollman charges, and AD4 atom types were assigned using AutoDock Tools [26]. Subsequently, the ligand was prepared by addition of Gasteiger partial charges and merging the nonpolar hydrogen atoms. The grid box of dimensions 45×45×45 Å³ with a grid spacing of 0.375 Å was generated by the auxiliary program AutoGrid 4.0 using the x, y, and z coordinates of the active site. The Lamarckian Genetic Algorithm (LGA) was applied to perform the docking calculations [27]. For the LGA method, the number of runs was set to 100 and the number of population was set to 150. The binding visualization and further ligand-receptor interactions were performed using PyMOL, Autodock tool program and VMD software [28, 29].

The internal validation of the docking protocol was performed by self-docking of the crystallized ligand to the binding pocket of the receptor. The root-mean-square deviation (RMSD) value was calculated between the docked and crystalized ligand to evaluate the accuracy of the docking protocol. The RMSD value between the crystalized and redocked conformations was less than 2 Å.

2.4. QSAR model development

Two different regression methods were developed and used as QSAR models: (1) simple multiple linear regression with stepwise variable selection (MLR) and (2) Genetic algorithm–partial least squares (GA-PLS). These well-known methods are commonly used in QSAR analysis [30]. In QSAR procedure, various approaches can be used to obtain many descriptors, but selecting the best of them is crucial because these descriptors must contain all the essential structural details of compounds. Here, using the SPSS statistics 18.0, stepwise selection and variable elimination were used to obtain a final descriptor set for QSAR modeling.

2.5. Model validation

Model validation is a critical stage in the development of a reliable model. Leave-one-out (LOO) cross-validation (CV) is a powerful tool in the GA-PLS and MLR methods that consider the potential and predictive ability of the model. If the following conditions are satisfied, the QSAR model is called predictive: $R^2 > 0.6$, $Q^2 > 0.6$ and $R^2_{pred} > 0.5$ (30). The higher R^2 and Q^2 values indicate the better prediction ability and more robustness of the model [31]. In this study, both GA-PLS and MLR methods were validated by the Leave-one-out cross-validation approach to check their predictive ability using MATLAB 2013 software.

2.6. Applicability domain

The applicability domain (AD) is a principle key in creating the QSAR model, and it must be quantified before predicting a collection of molecules. Therefore, the predictions of only those compounds that fall into the applicability domain of the model might be considered reliable. The leverage values for each compound were used to evaluate the applicability domain. Furthermore, the Williams plot was used to show standardized residuals versus leverage values (h). According to the Williams plot, the obtained GA-PLS and MLR models were reliable if the leverage value (h) of each compound in the data set was lower than the warning leverage value (h^*). In other words, prediction must be considered unreliable for compounds with a high leverage value ($h > h^*$). The warning leverage value (h^*) is generally fixed at $3(k + 1) / n$, where k is the number of the predictor variables and n is the number of training set compounds) [32–34].

2.7. *In Silico* Drug-Likeness, Bioavailability, Toxicology Prediction and Target identification

Nowadays, the assessments of pharmacokinetics and bioavailability become more important in the drug discovery process. Hence, in this work, the ADME (absorption, distribution, metabolism, and excretion) properties, drug-likeness (Lipinski's rule of five), and oral bioavailability of the 10 best-predicted compounds were calculated and compared with those of sunitinib as a control drug. The drug-likeness properties and ADME profiles were calculated through the SwissADME web tool (<http://www.swissadme.ch/>) [35]. The human intestinal absorption (HIA) and blood–brain barrier penetration (BBB) properties of ADME were predicted using BOILED-Egg graphical model in the SwissADME web tool [35]. Additionally, the toxicity which is one of the key ingredients in medicine was predicted by the admetSAR online tool that is available at <http://lmmd.ecust.edu.cn/admetsar2> [36, 37].

Target identification is an essential step in screening the best and effective compounds in the modern drug design process. We used SwissADME online tool for target prediction of selected compounds which is accessible through the <http://www.swisstargetprediction.ch/> server [38].

3. Results And Discussion

3.1. QSAR analysis

3.1.1. MLR modeling

First, various descriptors were used in stepwise multiple linear regression (MLR) analyses, then the MLR equation was obtained from the calculated descriptors. As shown in Table 2, the obtained model provided an acceptable predictive performance as confirmed by statistical parameters such as R^2 value of 0.73, R^2_p value of 0.67 for the test set, standard error of regression (SE) value of 0.59, variance ratio (F) value of 14.7 at specified degrees of freedom, Q^2 value of 0.67, Cvcv value of 10.08 and root mean square error of cross-validation (RMScv) value of 0.7. According to the MLR model, the connectivity indices (X5AV, X3A), 2D autocorrelation (MATS6m, MATS1e), geometrical (G(F.i .F)), geometrical descriptor (SPAM), atom-centered fragments (H-046) and constitutional indices (nN) descriptors influenced the kinase inhibitory activity of the studied structures. The obtained MLR equation and its results are shown in Table 2.

Table 2. MLR analyses results

PIC ₅₀ = 469.13X _{3A} (±74.433)-23.69MATS _{1e} (±4.791)-37.827SPAM(±12.547) +0.47G(F..F)(±0.098) -1.525nN(±0.313)-96.559MATs6m (±20.569) -0.285H 046(±0.119) +240.341X _{5A} v(±103.047) +36.153 (±21.905)						
R ² _p =0.67	Q ² =0.6774	R _m scv=0.7091	C _v cv=10.0813	R ² =0.73	F=14.741	Std.Error=0.598437

3.1.2. GA-PLS method

In the GA-PLS method, unlike MLR analysis, PLS analysis ignores the multicollinearity problem in the descriptors. In GA-PLS modeling, a genetic algorithm was used to find the most helpful set of descriptors. It is an essential step in QSAR analysis because the selection of relevant descriptors directly relates to the predicted inhibitory activity values. In our work, the GA-PLS method was able to select the best descriptors for building the QSAR model. The details of the descriptors and the statistical results of the model obtained from GA-PLS method are shown in Table 3. As can be seen, a combination of 2D autocorrelation (ATS1m, GATS5v), information content index (IC1), topological index (MAXDN), geometrical descriptor (SPAM) and quantum chemical descriptor (softness), selected by GA-PLS approach, were responsible for the kinase inhibitory activity of oxindole analogs.

The obtained GA-PLS model showed very high statistical quality parameters (i.e., R² value of 0.8284 and Q² value of 0.7175). The predictive ability of the obtained model was evaluated by five external test set molecules. The predictive R² value for the test set was found to be 0.79, which indicated the great predictive ability of the model. A detailed statistical results of the two generated models by GA-PLS and MLR methods are shown in Table S2. Comparison of statistical results of these two models indicates the higher predictability and accuracy of the GA-PLS model to the MLR model.

Table 3: GA-PLS results

PIC ₅₀ = -10.117ATS1m(±1.741)+4.526IC1(±0.813)-0.590MAXDN(±0.112)-3.604GATS5v(±1.440)-21.621SPAM(+7.366)+1.108softness(+0.422)+23.451(+5.151)						
R ² _p = 0.7956	Q ² = 0.7175	R _m scv = 0.5662	C _v cv = 7.9303	R ² = 0.82	F = 22.528	Std.Error = 0.4827

3.2. Application domain of the selected GA-PLS model

The William plot for the applicability domain of GA-PLS model is displayed in Figure 3. This plot was acquired using the leverage approach to define the applicability domain of the model. As shown in Figure 3, all values of the training set and test set compounds were lower than the warning h* value of 0.6, indicating the good predicting ability of the final model. Hence, new compounds with enhanced tyrosine kinase inhibitory activity could be designed using GA-PLS model.

3.3. Docking Studies

With the aim of exploring the possible binding modes of TKIs to the tyrosine kinase active site and obtaining their crucial binding interactions, 44 TKIs were docked into the active site of the tyrosine kinase enzyme (PDB ID: 4AW5). Before docking, the reliability of the docking method and parameters was evaluated by redocking the original co-crystallized ligand into the binding pocket.

As shown in Figure 4A, the best redocked pose and the original structure of the co-crystallized ligand were superimposed well with each other. Additionally, the RMSD value between them was 0.33 Å, confirming the reliability and accuracy of the docking method employed in this study. An H-bond was found between the oxindole carbonyl group of co-crystallized ligand and Gly98 residue (Figure 4B). Hydrogen bond and electrostatic interactions were also found between the piperidine ring and Glu140 and Asp234, respectively.

All TKIs were then docked into the active site applying the same docking parameters. The ranking order of the inhibitory activity (pIC_{50} values) were found to be in accordance with the docking scores (Figure 5), which further indicated the reliability of the employed docking algorithm. The estimated free binding energy values (ΔG_{bind} ; kcal/mol) and the possible binding mode of this compound with the key amino acid residues at the active site of the enzyme are summarized in Tables 1, S5 and S6.

As can be seen from Table 1, the studied molecules were classified into four different classes based on their structures. Accordingly, their docking results have changed with these structural changes. The binding energy values (ΔG_{bind}) for the best-docked poses of the class 1 (6a-6w), class 2 (7a-7j), class 3 (8a-8e) and class 4 (8f-8k) compounds were within the range of -7.34 to -10.53 kcal/mol, -9.29 to -10.77 kcal/mol, -7.53 to -11.05 kcal/mol and -9.63 to -10.46 kcal/mol, respectively. Based on the results, class 2 and 3 compounds possessed the lowest docking binding energies. Therefore, the docking results of compounds **7g**, **7j**, **8b**, and **8e** with the lowest binding energy among the 44 compounds were chosen for further binding mode analysis between protein and ligands.

It was noted that the nitrogen and oxygen atoms of oxindole core and the nitrogen atom of benzimidazole group of all four compounds formed hydrogen bonding contacts with Glu170, Phe171 and Met172, respectively, which indicated the remarkable contribution of these interactions for the binding affinity (Figures 6 and S1). For compounds **7g** and **7j**, other CH—O interactions were observed between the sidechain group on the nitrogen atom of piperidine ring and the oxygen atom of Arg220, Asn221 and Asp234. Additionally, the piperidine ring in the three studied compounds **7g**, **7j**, and **8b** formed CH—O interactions with Ser233.

3.3. Design of novel potent compounds

According to the information obtained from the developed QSAR models (MLR and GA-PLS) and the binding mechanisms of molecular docking, 44 compounds were proposed as novel tyrosine kinase inhibitors. The main strategy in the design of these novel inhibitors was the skeleton modification of the most potent compounds (8a-8e) and (8f-8k) *via* bioisosterism principle. The inhibitory activities of the newly designed compounds were predicted by the derived MLR and GA-PLS models and the results are listed in Table S3. From the designed compounds, molecules **9i**, **9j**, **9k**, **9m**, **10j**, **11b**, **11e**, **11h**, **11k**, and **11m** showed equal to or higher predicted pIC_{50} values than compounds (8a-8e) and (8f-8k) (Figure 7).

To explore the binding modes of the novel designed compounds, they were docked into the binding site of tyrosine kinase enzyme. The docking score of all newly proposed compounds is shown in Table S3 as well as the binding mode details of 12 of them are listed in Table S4. Based on their structures, the designed compounds can be divided into three groups (9a-9p), (10a-10k), and (11a-11m). Hence, the docking results were affected by structural changes in these compounds. As expected, compounds 11a-11m showed better docking scores than those of other designed compounds. Moreover, all proposed compounds of group 11a-11m showed

lower binding energy values than that of sunitinib as the control drug and some of them displayed equal and even better binding energy values than that of co-crystallized ligand.

The binding interactions of Sunitinib and new compounds 9i, 9j, 9m, 11b, 11e, 11j, and 11k in the active pocket of tyrosine kinase are shown in Figures 8 and S2-S4. The docking results suggested that most of the designed compounds had a common binding mode similar to that of the most potent compounds (8a-8e) and (8f-8k). As shown in Figure 8A, there is a CH—O interaction between Sunitinib and Glu170, further indicating the crucial role of Glu170 in stabilizing ligands in the active site of enzyme. The designed compound 9i formed two H-bond interactions with Glu170 and Met172. Similar to compounds 7a-7j, the piperidine ring and the side chain group on the nitrogen atom of the piperidine ring of compound 9i participated in CH—O interactions with Arg220 and Asn221. The docking findings also indicated that, in addition to H-bonding, hydrophobic interactions with nonpolar amino acids and arene-H interactions play a significant role in the stability of the ligand-protein complex.

To investigate the specificity of the newly proposed compounds toward the tyrosine kinase enzyme, the target identification was carried out using SwissADME online tool. Table S5 summarizes the results of this prediction. Among these new structures, compounds 9j, 9k, 9m, 10j, 11b, and 11e demonstrated the highest probability to act as kinase inhibitors with similarity of 93.3%, 66.7%, 80%, 60%, 66.7% and 60% to previously reported kinase inhibitors, respectively. It was notable that SwissADME showed 80% similarity with sunitinib as a control drug, which could indicate the acceptable specificity of our proposed compounds.

3.4. *In Silico* Drug-Likeness, Bioavailability, Toxicology Prediction and, Synthetic accessibility

The druggability of the newly designed compounds was analyzed by assessing their physiochemical properties and by applying Lipinski's rule of five using SwissADME web tool and compared with that of sunitinib (Table 4). It could be seen that most of the selected compounds successfully passed Lipinski's five rules. Compounds 9i, 11h, 11k, and 11m had one violation of Lipinski's rule of five, and that was their molecular weight.

Table 4
In silico Lipinski's rule of five results for 10 selected designed compounds

Compound	Molecular weight ($\leq 500\text{Da}$)	Oral Bioavailability: TPSA($\leq 140\text{\AA}$)	H- Bond Donor (≤ 5)	H- Bond Acceptor (≤ 10)	logP (≤ 5)	Rotatable Bond (≤ 10)	Rule of 5 violation
Sunitinib	398.47	77.23	3	4	3.55	8	0
9i	536.62	96.55	2	6	4.01	8	1
9j	463.57	73.05	3	3	3.47	5	0
9k	480.51	96.55	2	5	3.26	5	0
9m	473.53	102.93	4	5	2.16	6	0
10j	434.56	114.18	3	4	2.85	5	0
11b	446.52	70.25	2	5	3.44	5	0
11e	476.52	78.09	2	6	2.86	5	0
11h	538.59	78.09	2	6	3.32	6	1
11k	587.50	98.32	3	5	3.39	6	1
11m	504.58	103.87	2	6	2.99	6	1

Table 5 shows the physiochemical properties of the selected newly designed compounds. As shown, most of the compounds have moderate water solubility and some of them are poorly soluble in water. Although all compounds except compound 11k displayed high gastrointestinal absorption, most did not show brain penetration. Half of the compounds depicted that they could act as cytochrome CYP2D6 inhibitors like the control drug, indicating that these compounds might not easily undergo oxidation and hydroxylation in the phase-1 metabolism. The synthetic accessibility of all new compounds were lower than 5, suggesting that they were relatively easy to be synthesized.

Table 5
Pharmacokinetic parameters of the best selected compounds

Compounds	Water solubility ¹	GI absorption	BBB permeant	CYP2D6 inhibitor	P-gp substrate	Lipophilicity Log $P_{o/w}$	Synthetic accessibility ²
Sunitinib	Soluble	High	Yes	Yes	Yes	3.55	3.58
9i	Moderately	High	No	Yes	Yes	4.01	4.08
9j	Poorly	High	Yes	Yes	Yes	3.47	4.21
9k	Moderately	High	No	Yes	Yes	3.26	3.64
9m	Poorly	High	No	No	No	2.16	3.61
10j	Moderately	High	No	Yes	Yes	2.85	3.83
11b	Moderately	High	Yes	Yes	Yes	3.44	3.71
11e	Moderately	High	No	No	Yes	2.86	3.74
11h	Poorly	High	No	No	No	3.32	4.11
11k	Poorly	Low	No	No	No	3.39	4.15
11m	Moderately	High	No	No	Yes	2.99	3.98
1-log S scale: insoluble < -10 < poorly < -6 < poorly < moderately < -4 < soluble < -2 very < 0 < highly							
2-range: 1= very easy, 10= very hard							

Bioavailability radar plots depicted the oral bioavailability of the selected compounds (Figure 9). Orally bioavailable compounds were predicted to be 9i, 9j, 10j, 11b, 11e, 11h, and 11m. compounds 9k and 9m showed one off-shoot relative to insaturation (INSATU), and compound 11k showed four off-shoot relatives to insaturation (INSATU), insolubility (INSOLU), size and Lipophilicity (LIPO) indicated that they might have poor physicochemical properties for oral bioavailability.

The toxicity risk calculation is critical in the early drug design phase for screening the safest and most effective drugs or compounds. The toxicity of 10 selected compounds was predicted using the admetSAR online tool [37]. Table S6 shows the toxicity prediction results. None of the compounds showed carcinogenicity, eye corrosion, or eye irritation. Compound 9m was predicted to cause Ames mutagenesis. All Compounds and control drug (Sunitinib) were determined to be hepatotoxic, however, the toxicity range of the selected compounds was better than that of Sunitinib. According to the US EPA classification, all compounds were predicted to have class III acute oral toxicity, which means that their LD₅₀ values are greater than 500 mg/kg but less than 5000 mg/kg. These computational pharmacokinetics and bioavailability information could improve the success rate of newly proposed compounds.

Conclusion

In this work, a molecular modeling study on a series of oxindole derivatives has been carried out employing 2D-QSAR, molecular docking and ADMET property analysis to investigate the structural requirements for inhibition of tyrosine kinases. Two different QSAR modeling methods, MLR and GA-PLS, were used to develop

QSAR models by 35 molecules as training set and 9 molecules as the test set to validate these models. The QSAR model obtained from GA-PLS approach was superior to the MLR version and had more powerful predictive ability. From the molecular docking, the H-bond interactions with residues Glu170, Phe171, and Met172 contributed mainly to the inhibitory activity of the studied compounds. Furthermore, residues Arg220, Asn221 played an important role in stabilizing inhibitors in the tyrosine kinase active site. Based on information derived from QSAR models and molecular docking studies, 44 novel potential inhibitors with improved theoretical inhibitory activity were designed by modification of the oxindole scaffold. According to their predicted pIC₅₀ values and docking scores, ten newly designed compounds 9i, 9j, 9k, 9m, 10j, 11b, 11e, 11h, 11k, and, 11m were selected and studied through molecular docking and their ADMET properties were investigated comprehensively. The docking results further verified the crucial role of Glu170, Phe171, Met172, Arg220, and Asn221 in binding interactions. Furthermore, ADMET studies suggested that the six proposed compounds 9i, 9j, 9m, 11b, 11e, and 11h had favorable drug-like properties, pharmacokinetic parameters and toxicological profiles as novel tyrosine kinase inhibitors. These selected compounds, in particular, proved to be a helpful starting point for the development of new synthetic oxindole derivatives with improved inhibitory activity. Overall, these results might provide and insight into some instructions for the rational design of more potent tyrosine kinase inhibitors.

Declarations

Acknowledgments

This work was supported by the Isfahan University of medical sciences (grant number: 399739). Collaboration of medicinal chemistry, Faculty of Pharmacy, Isfahan University of Medical sciences (G 399739), in providing the required facilities for this work is greatly acknowledged.

References

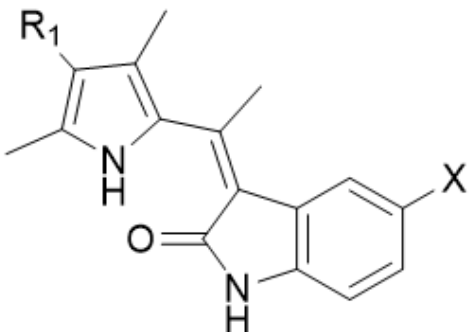
1. Arora A, Scholar EM. Role of tyrosine kinase inhibitors in cancer therapy. *Journal of Pharmacology and Experimental Therapeutics*. 2005;315(3):971-9.
2. Brunelleschi S, Penengo L, Santoro MM, Gaudino G. Receptor tyrosine kinases as target for anti-cancer therapy. *Current pharmaceutical design*. 2002;8(22):1959-72.
3. Cohen P, Alessi DR. Kinase drug discovery—what's next in the field? *ACS chemical biology*. 2013;8(1):96-104.
4. Schlessinger J. Cell signaling by receptor tyrosine kinases. *Cell*. 2000;103(2):211-25.
5. Heldin C-H, Östman A, Rönstrand L. Signal transduction via platelet-derived growth factor receptors. *Biochimica et Biophysica Acta (BBA)-reviews on cancer*. 1998;1378(1):F79-F113.
6. Ferrara N, Gerber H-P, LeCouter J. The biology of VEGF and its receptors. *Nature medicine*. 2003;9(6):669-76.
7. Takahashi H, Shibuya M. The vascular endothelial growth factor (VEGF)/VEGF receptor system and its role under physiological and pathological conditions. *Clinical science*. 2005;109(3):227-41.
8. Zhang J, Yang PL, Gray NS. Targeting cancer with small molecule kinase inhibitors. *Nature reviews cancer*. 2009;9(1):28-39.

9. Kasture VS, Musmade DS, Aher SJ, Patil PP. Tyrosine Kinases: promising targets for cancer chemotherapy. *Med Chem Drug Discov.* 2012;2(2):37-51.
10. Wilhelm S, Carter C, Lynch M, Lowinger T, Dumas J, Smith RA, et al. Discovery and development of sorafenib: a multikinase inhibitor for treating cancer. *Nature reviews Drug discovery.* 2006;5(10):835-44.
11. Atkins M, Jones CA, Kirkpatrick P. Sunitinib maleate. Nature Publishing Group; 2006.
12. Yang LP, McKeage K. Axitinib. *Drugs.* 2012;72(18):2375-84.
13. Wilhelm S, Dumas J, Adnane L, Lynch M, Carter C, Schütz G, et al. A new oral multikinase inhibitor of angiogenic, stromal and oncogenic receptor tyrosine kinases with potent preclinical antitumor activity. *Int J Cancer.* 2011;129(1):245-55.
14. Ivy SP, Wick JY, Kaufman BM. An overview of small-molecule inhibitors of VEGFR signaling. *Nature reviews Clinical oncology.* 2009;6(10):569-79.
15. Musumeci F, Radi M, Brullo C, Schenone S. Vascular endothelial growth factor (VEGF) receptors: drugs and new inhibitors. *Journal of medicinal chemistry.* 2012;55(24):10797-822.
16. Sun L, Liang C, Shirazian S, Zhou Y, Miller T, Cui J, et al. Discovery of 5-[5-fluoro-2-oxo-1, 2-dihydroindol-(3 Z)-ylidenemethyl]-2, 4-dimethyl-1 H-pyrrole-3-carboxylic acid (2-diethylaminoethyl) amide, a novel tyrosine kinase inhibitor targeting vascular endothelial and platelet-derived growth factor receptor tyrosine kinase. *Journal of medicinal chemistry.* 2003;46(7):1116-9.
17. Mendel DB, Laird AD, Smolich BD, Blake RA, Liang C, Hannah AL, et al. Development of SU5416, a selective small molecule inhibitor of VEGF receptor tyrosine kinase activity, as an anti-angiogenesis agent. *Anti-cancer drug design.* 2000;15(1):29-41.
18. Kim MH, Tsuhako AL, Co EW, Aftab DT, Bentzien F, Chen J, et al. The design, synthesis, and biological evaluation of potent receptor tyrosine kinase inhibitors. *Bioorganic & medicinal chemistry letters.* 2012;22(15):4979-85.
19. Ding L, Tang F, Huang W, Jin Q, Shen H, Wei P. Design, synthesis, and biological evaluation of novel 3-pyrrolo [b] cyclohexylene-2-dihydroindolinone derivatives as potent receptor tyrosine kinase inhibitors. *Bioorganic & medicinal chemistry letters.* 2013;23(20):5630-3.
20. Ultra C. 12.0, Cambridge Soft, Cambridge, MA 2010. Received: May. 2013;10.
21. Zare S, Fereidoonzhad M, Afshar D, Ramezani Z. A comparative QSAR analysis and molecular docking studies of phenyl piperidine derivatives as potent dual NK1R antagonists/serotonin transporter (SERT) inhibitors. *Computational biology and chemistry.* 2017;67:22-37.
22. Mauri A, Consonni V, Pavan M, Todeschini R. Dragon software: An easy approach to molecular descriptor calculations. *Match.* 2006;56(2):237-48.
23. Thanikaivelan P, Subramanian V, Raghava Rao J, Unni Nair B. Application of quantum chemical descriptor in quantitative structure activity and structure property relationship. *Chem Phys Lett.* 2000;323(1-2):59-70.
24. Fereidoonzhad M, Faghieh Z, Mojaddami A, Tabaei S, Rezaei Z. Novel Approach Synthesis, Molecular Docking and Cytotoxic Activity Evaluation of N-phenyl-2, 2-dichloroacetamide Derivatives as Anticancer Agents. *Journal of Sciences, Islamic Republic of Iran.* 2016;27(1):39-49.
25. Morris GM, Huey R, Olson AJ. Using AutoDock for Ligand-Receptor Docking. *Current Protocols in Bioinformatics: John Wiley & Sons, Inc.;* 2002.

26. Weiner SJ, Kollman PA, Case DA, Singh UC, Ghio C, Alagona G, et al. A new force field for molecular mechanical simulation of nucleic acids and proteins. *Journal of the American Chemical Society*. 1984;106(3):765-84.
27. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, et al. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *Journal of computational chemistry*. 1998;19(14):1639-62.
28. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *Journal of molecular graphics*. 1996;14(1):33-8, 27-8.
29. DeLano WL. The PyMOL molecular graphics system. <http://www.pymol.org>. 2002.
30. Khoshneviszadeh M, Faghieh Z, Jokar E, Mojaddami A, Rezaei Z. QSAR, Molecular Docking and protein ligand interaction fingerprint studies of N-phenyl dichloroacetamide derivatives as anticancer agents. *Trends in Pharmaceutical Sciences*. 2016;2(2).
31. Roy K, Roy PP. Comparative chemometric modeling of cytochrome 3A4 inhibitory activity of structurally diverse compounds using stepwise MLR, FA-MLR, PLS, GFA, G/PLS and ANN techniques. *European journal of medicinal chemistry*. 2009;44(7):2913-22.
32. Roy K, Kar S, Das RN. Chapter 7 - Validation of QSAR Models. *Understanding the Basics of QSAR for Applications in Pharmaceutical Sciences and Risk Assessment*. Boston: Academic Press; 2015. p. 231-89.
33. Weaver S, Gleeson MP. The importance of the domain of applicability in QSAR modeling. *J Mol Graphics Modell*. 2008;26(8):1315-26.
34. Cheng Z, Yang B, Chen Q, Gao X, Tan Y, Yuan T, et al. Quantitative-Structure-Activity-Relationship (QSAR) models for the reaction rate and temperature of nitrogenous organic compounds in supercritical water oxidation (SCWO). *Chemical Engineering Journal*. 2018;354:12-20.
35. Daina A, Michielin O, Zoete V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Scientific reports*. 2017;7:42717.
36. Egan WJ, Merz KM, Baldwin JJ. Prediction of drug absorption using multivariate statistics. *Journal of medicinal chemistry*. 2000;43(21):3867-77.
37. Cheng F, Li W, Zhou Y, Shen J, Wu Z, Liu G, et al. admetSAR: a comprehensive source and free tool for assessment of chemical ADMET properties. *ACS Publications*; 2012.
38. Gfeller D, Grosdidier A, Wirth M, Daina A, Michielin O, Zoete V. SwissTargetPrediction: a web server for target prediction of bioactive small molecules. *Nucleic acids research*. 2014;42(W1):W32-W8.

Figures

A



Sunitinib: $R_1 = \text{CONH}(\text{CH}_2)\text{N}(\text{C}_2\text{H}_5)$, $X = \text{F}$

B

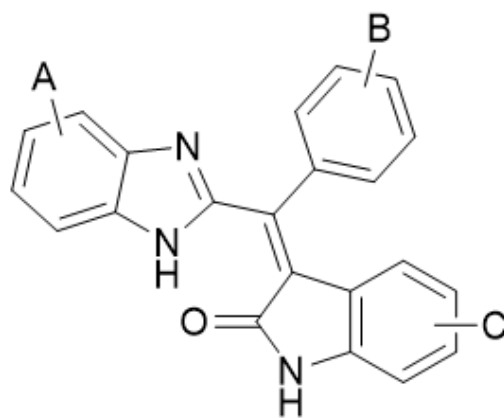


Figure 1

A) Sunitinib structure. B) Positions of rings A, B, and C in the indolin-2-ones structure.

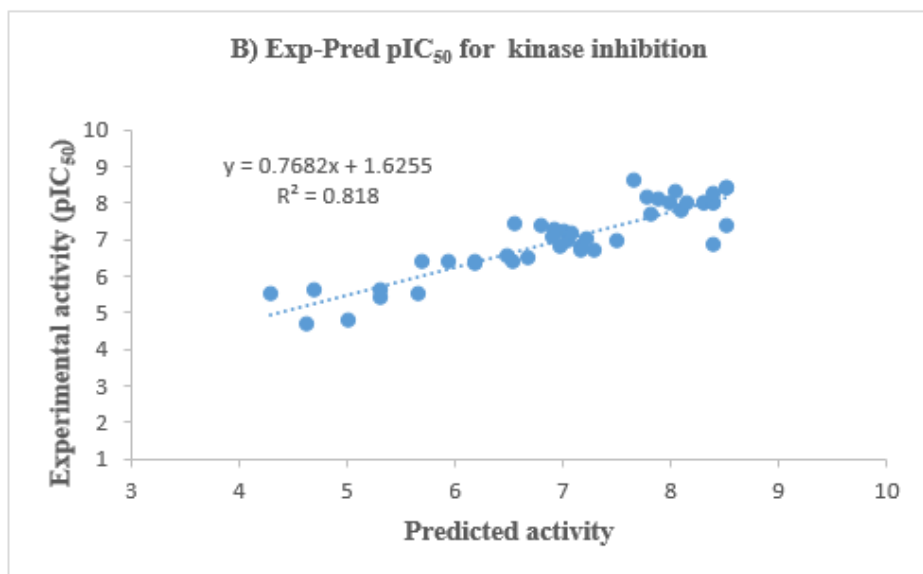
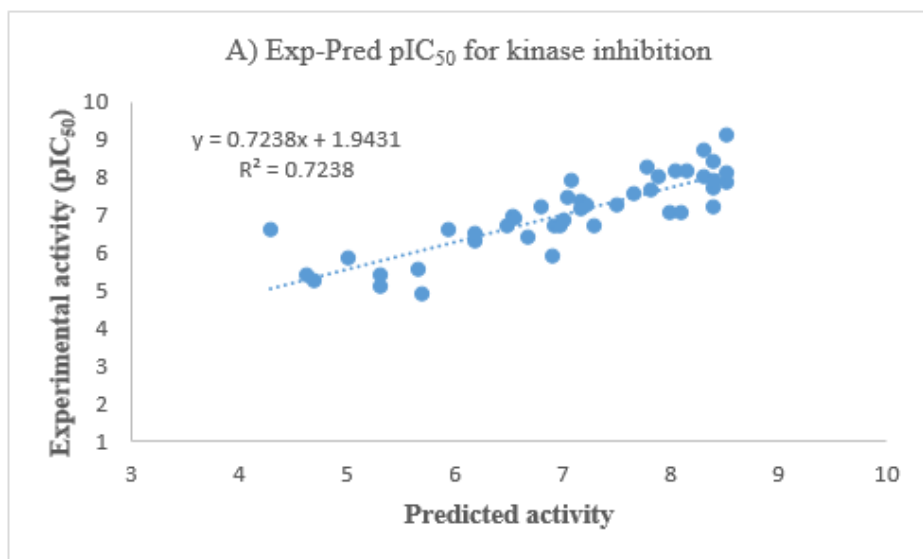


Figure 2

Plots of predicted pIC50 versus experimental pIC50 values of the training and test set compounds retrieved by the MLR (A) method and (B) GA-PLS method.

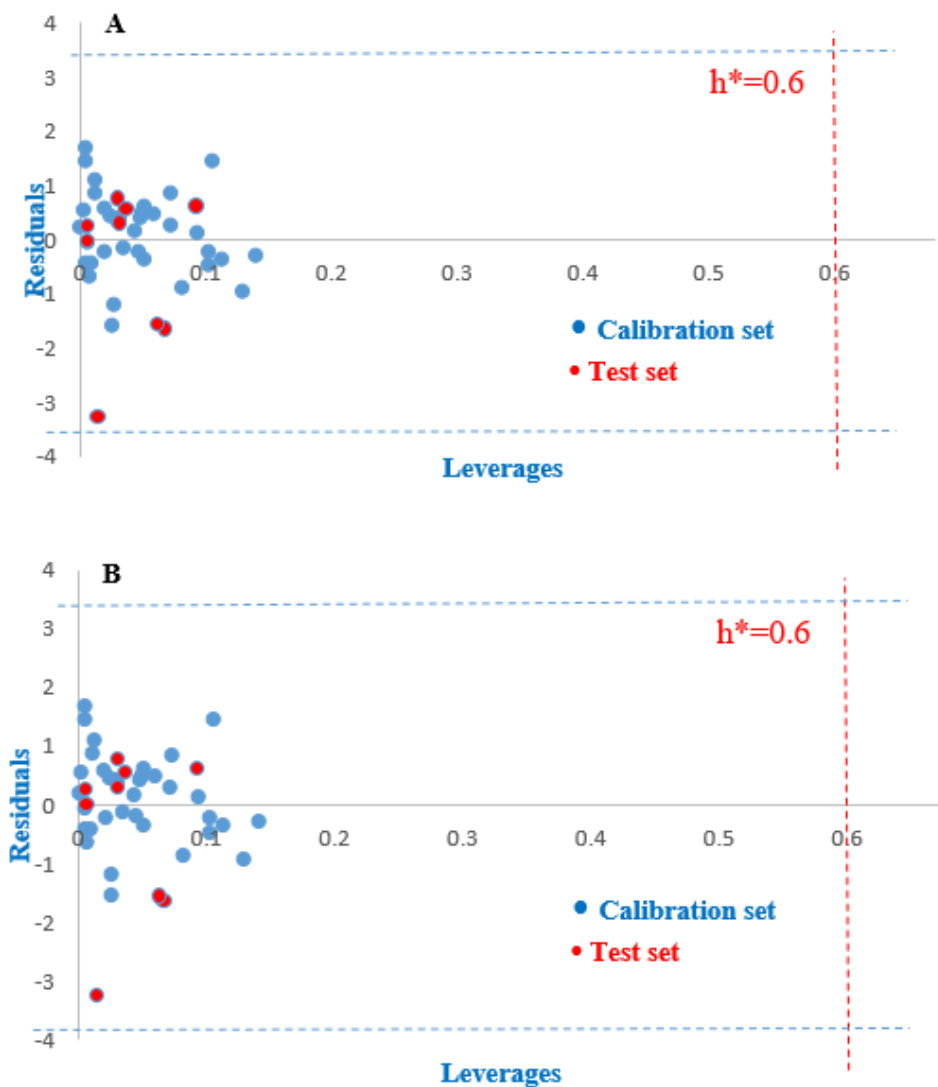


Figure 3

The Williams plots for the calibration set and external prediction set for kinase inhibition A) GA-PLS model B) MLR model.

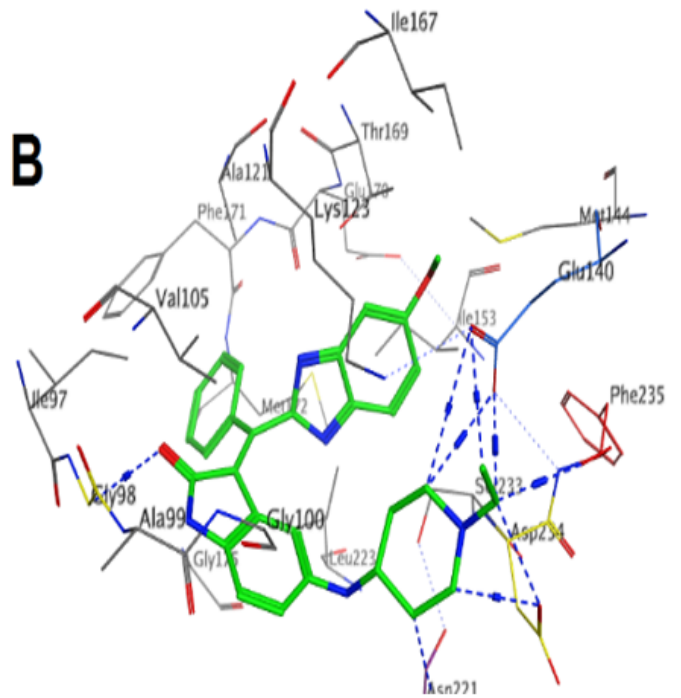
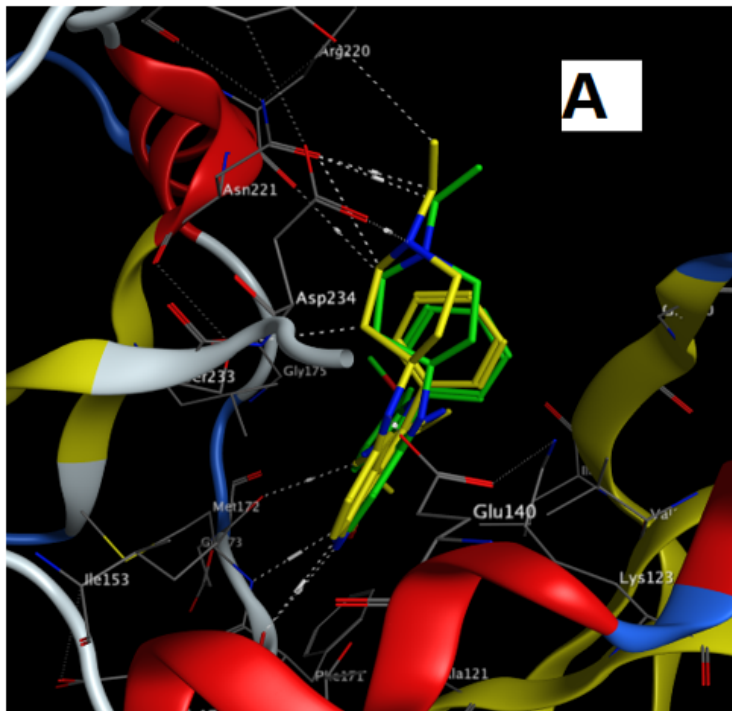


Figure 4

The redocked results of co-crystallized ligand in the binding site of tyrosine kinase enzyme (PDB ID = 4AW5). (A) Superimposition of co-crystallized ligand conformation (green) with the best redocked conformation (yellow). (B) Binding modality of the co-crystal with the key residues in the active site of 4AW5.

Figure 5

Plot of correlation between experimental pIC50 values and docking binding energy

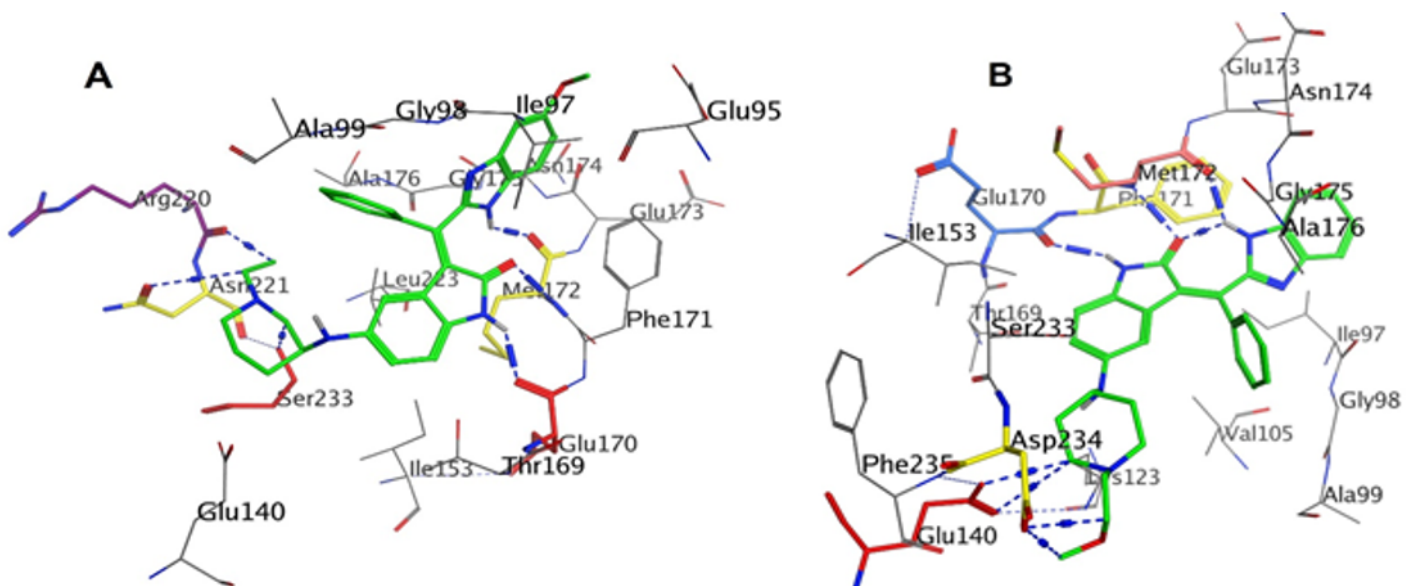


Figure 6

Binding modes of compounds 7g (A) and 7j (B) in the active pocket of tyrosine kinase enzyme

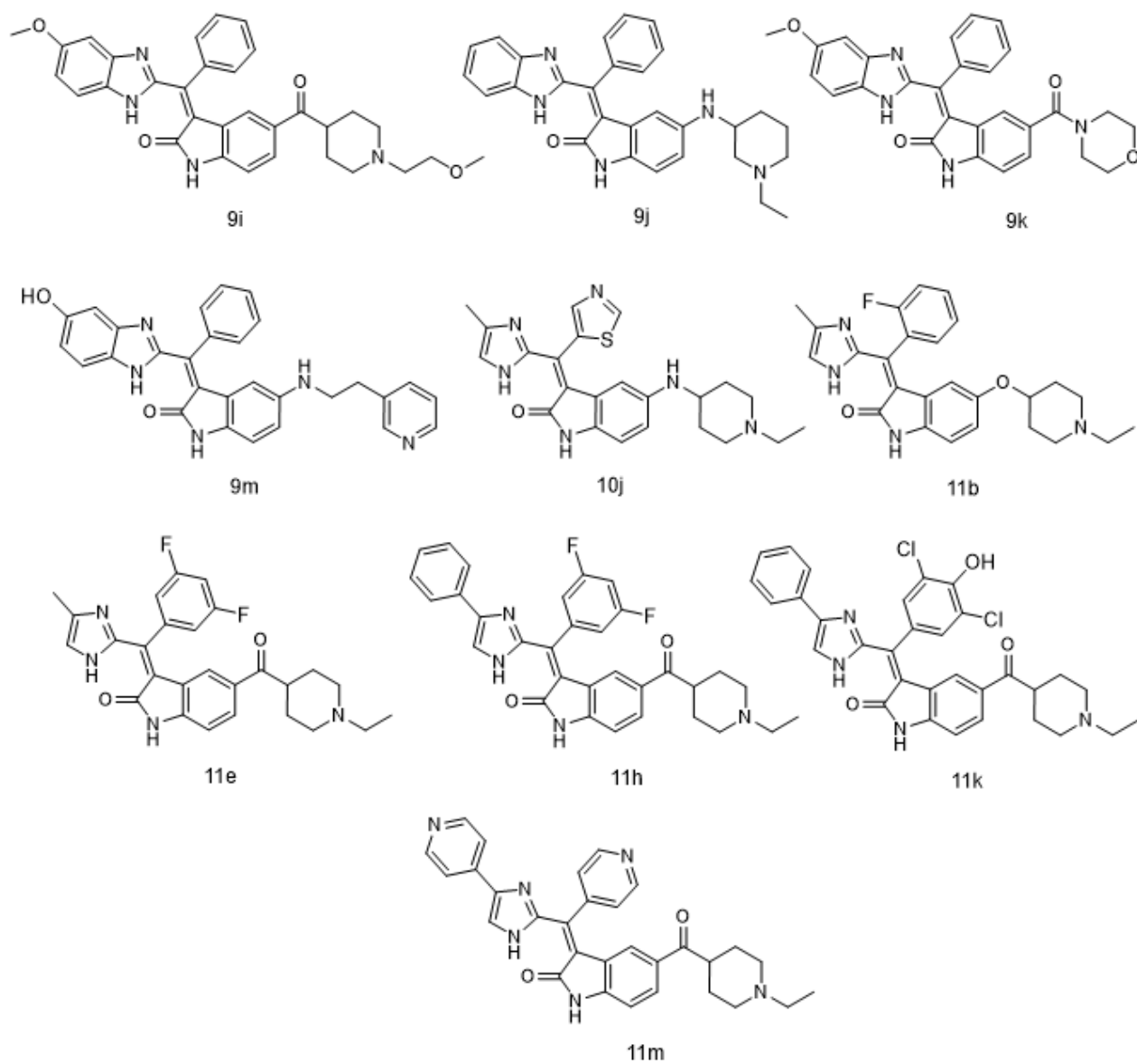


Figure 7

Structures of the compounds 9i, 9j, 9k, 9m, 10j, 11b, 11e, 11h, 11k, and 11m.

Figure 8

Docking results of control drug Sunitinib (A) and newly designed compound 9i (B) in the active site of tyrosine kinase enzyme.

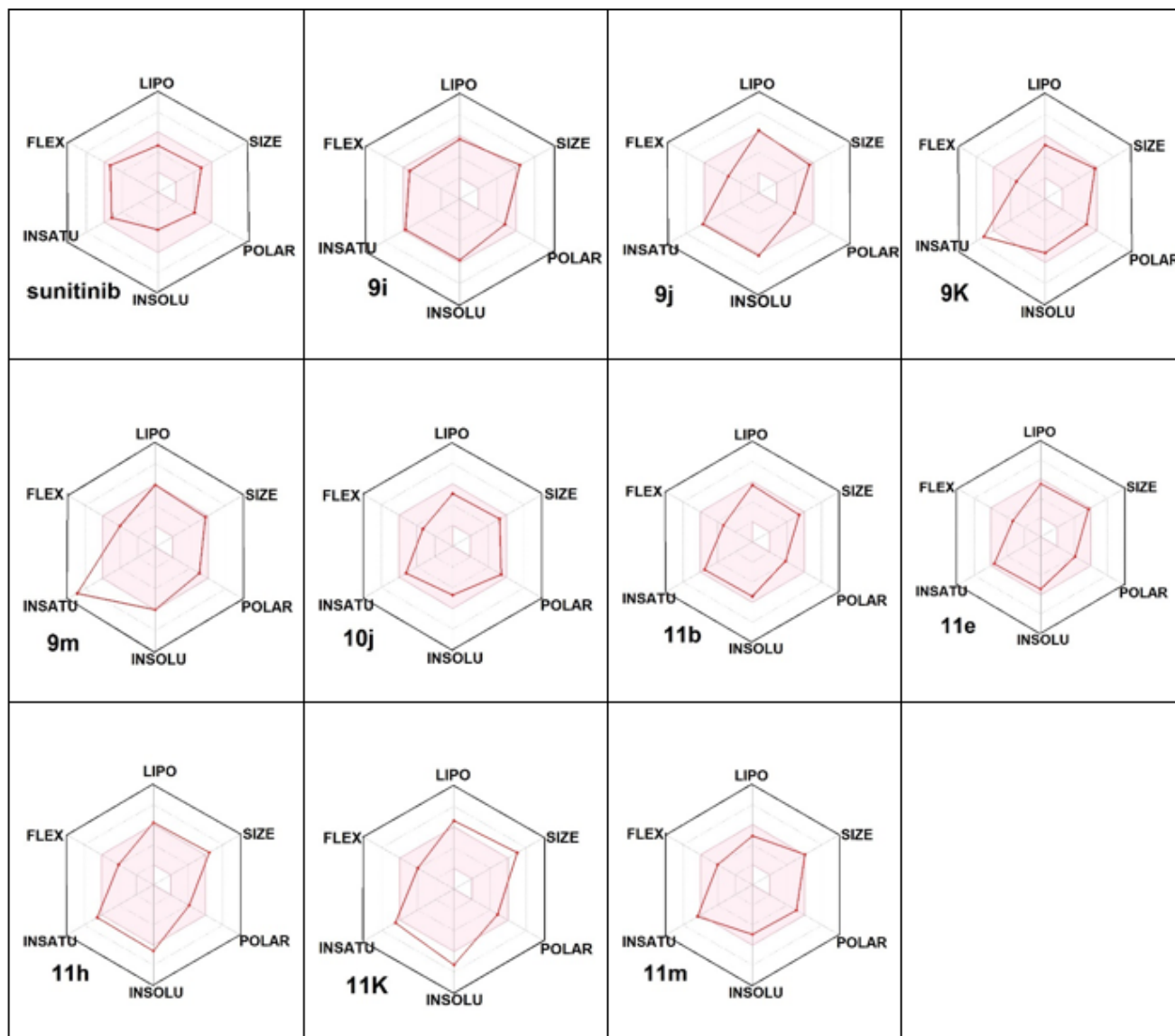


Figure 9

The bioavailability radar plots of compounds 1–10 and control drug Sunitinib. The pink area shows the ideal range for each parameter. Lipophilicity (LIPO): $-0.7 < XLOGP3 < 5.0$; SIZE: $150 \text{ g/mol} < MW < 500 \text{ g/mol}$; polarity (POLAR): $20 \text{ \AA}^2 < \text{topological polar surface area (TPSA)} < 130 \text{ \AA}^2$, and insolubility (INSOLU): $0 < \text{LogS} < 6$; INSATU (insaturation): $0.25 < \text{fraction of carbons in } sp^3 \text{ hybridization} < 1$; FLEX (flexibility): $0 < \text{number of rotatable bonds} < 9$. The radar plots were obtained with the support of the SwissADME web tool.

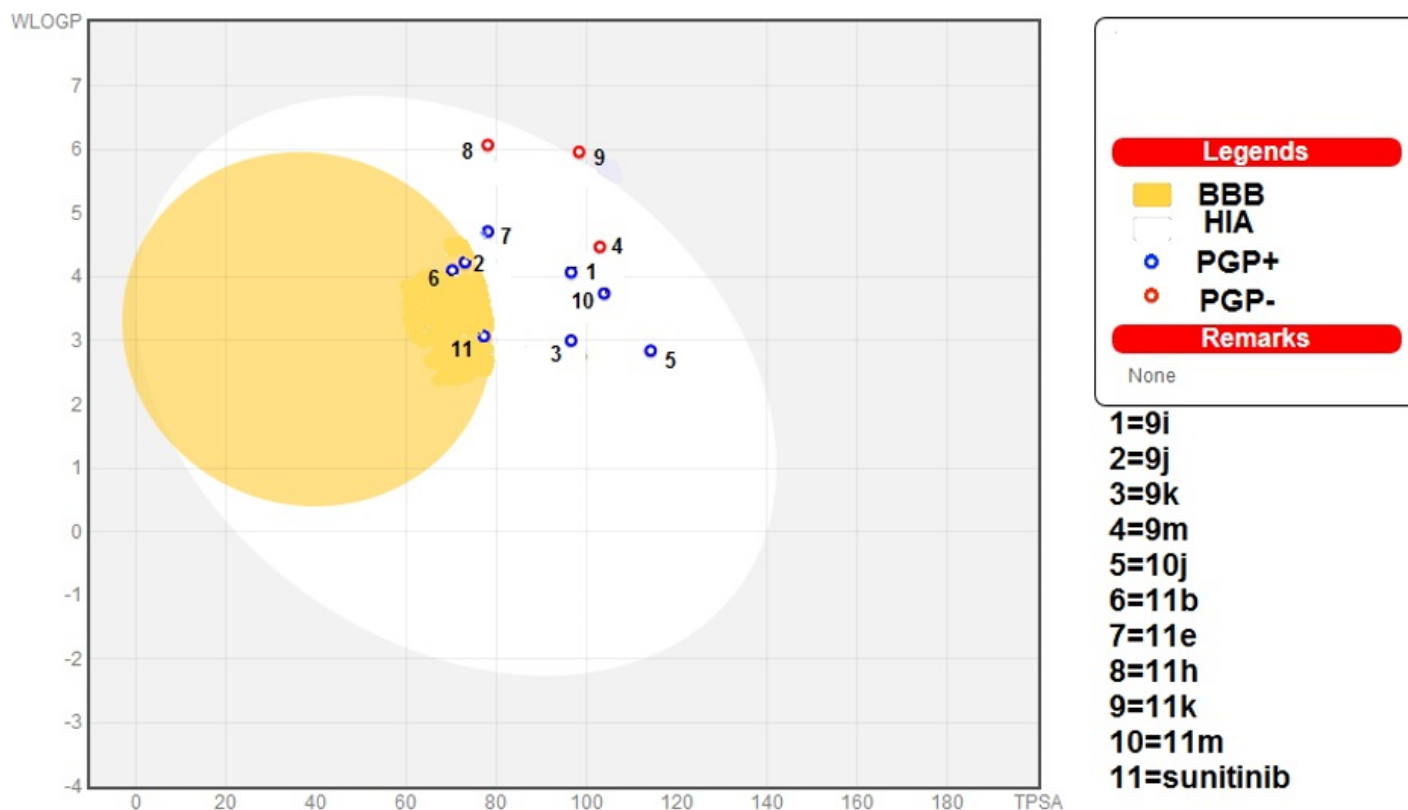


Figure 10

BOILED-Egg plot of 10 selected designed compounds and the control drug Sunitinib. Spots on the yellow area of the plot represent passive diffusion molecules through the BBB. In contrast, points in the white space of the plot are predicted to be passively absorbed by the gastrointestinal tract (HIA). Blue dots (PGP+) indicate the molecules expected to be effluated from the central nervous system (CNS) by P-glycoprotein. In contrast, the red ones (PGP-) show the molecules predicted not to be effluated from the CNS by P-glycoprotein [37].

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupportingInformation.docx](#)
- [Onlinefloatimage1.png](#)