

Ligand Based 3D-QSAR Pharmacophore, Molecular Docking and Adme to Identify Potential Fibroblast Growth Factor Receptor 1 Inhibitors

Zizhong Tang

Sichuan Agricultural University

Lu Huang

Sichuan Agricultural University

Xiaoli Fu

Sichuan Agricultural University

Haoxiang Wang

Sichuan Agricultural University

Biao Tang

Sichuan Agricultural University

Yirong Xiao

Sichuan Agricultural University Hospital

Caixia Zhou

Sichuan Agricultural University

Zhiqiao Zhao

Sichuan Agricultural University

Yujun Wan

Sichuan food fermentation industry research and design institute

Hui Chen (✉ 14126@sicau.edu.cn)

College of Life Sciences, Sichuan Agricultural University, Ya'an 625014, China

Huipeng Yao

Sichuan Agricultural University

Zhi Shan

sichuan agricultural university

Tongliang Bu

sichuan agricultural university

Xulong Wu

chengdu agricultural college

Keywords: FGFR1, Inhibitor, Molecular docking, ADME, Pharmacophore, ADME

Posted Date: September 3rd, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-69980/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

The FGF/FGFR system may affect tumor cells and stromal microenvironment through autocrine and paracrine stimulation, thereby significantly promoting oncogene transformation and tumor growth. Abnormal expression of FGFR1 in cells is considered to be the main cause of tumorigenesis and a potential target for the treatment of cancer. In this study, a combination of structure-based drug carriers and molecular docking-based virtual screening was used to screen new potential FGFR1 inhibitors. Twenty-one known inhibitors were collected as training sets to establish a 3D-QSAR pharmacophore model, and cost analysis, test set validation, and Fischer randomization test were used to validate the efficiency of the pharmacophore model. In Accelrys Discovery Studio 2016, the zinc database was filtered by Lipinski's Rule of Five and SMART's filtration. Then, Hypo01 was used for virtual screening of ZINC database. Compounds with predicted activity values less than 1 μM were molecularly docked with FGFR1 protein crystals, the docking results were observed, and the interaction between compounds and targets was studied. The absorption, distribution, metabolism and excretion (ADME) and toxicity of potential inhibitors were studied, and a compound with new structural scaffolds were obtained. It could be further studied to explore their better therapeutic effects.

Introduction

Cancer is the second leading disease in morbidity and mortality after cardiovascular disease. According to the Global Cancer Report issued by the World Health Organization (WHO), there were 18.1 million new cancer patients worldwide in 2018, and about 9.6 million patients died of cancer [1]. In fact, unhealthy lifestyle and dietary habits are one of the main factors inducing cancer [2]. In order to treat cancer and reduce cancer mortality, human beings have invested a lot of money and energy to find effective treatment methods. However, due to the diversity of cancer pathogenesis and different treatment mechanisms, treatment lacks selectivity and cannot directly act on cancer cells and avoid normal tissues [3], leading to difficulties in the development of new drugs. Therefore, it is necessary to find effective therapeutic targets.

The FGFR pathway was initially thought to promote angiogenesis and synergize with the vascular endothelial growth pathway [4], but recent studies have shown that this pathway is also a driving oncogene in some tumors [5]. In recent years, more and more researchers have paid attention to the FGF/FGFR system as an effective target for the development of new anticancer drugs. Fibroblast growth factor receptors (FGFRs) are a subfamily of tyrosine kinase receptors with four subtypes, including FGFR1, FGFR2, FGFR3 and FGFR4 [6]. Fibroblast growth factors bind to FGFRs with the assistance of heparan sulfate glycosaminoglycans, causing the dimerization of FGFRs, leading to the activation of multiple tyrosine residues in their intracellular tyrosine kinase regions by autophosphorylation. Activated FGFRs activate their substrates PLC γ and signal adaptor protein FRS2 by phosphorylation, and their substrates reactivate downstream signaling pathways such as PI3K/AKT, MEK/MAPK, and PKC [7–9]. Misexpression of fibroblast growth factor induces tumorigenesis and growth, so FGFR is considered an important target for cancer treatment. Amplification of the FGFR1 gene, the most common FGFR

amplification in cancer, has been reported in different types of tumors. Weiss et al. have shown that the frequent and local amplification of FGFR1 in squamous cell lung cancer is related to the therapeutic FGFR1 dependence [10]. Fischbach et al. have shown that the gene amplification of fibroblast growth factor is closely related to the occurrence of bladder cancer [11]. Turner et al demonstrated that the FGFR1 gene is a therapeutic target for breast cancer [12].

Researchers have designed and synthesized a large number of FGFR1 inhibitors [13], and some drugs have been put into clinical treatment, but the treatment failure rate is high. Therefore, it is still urgent to develop new FGFR1 inhibitors. Advances in science and technology have promoted the development of virtual screening technology, which has greatly improved the efficiency of drug research and development and reduced the cost of research. The aim of this study was to discover a novel inhibitor of FGFR1. Using 3D-QSAR pharmacophore technology, we constructed pharmacophores by analyzing the structural characteristics of 21 known active inhibitors with different molecular structure patterns. The selected pharmacophore was used as a 3D query to screen potential compounds in a Zinc database containing 86976 molecules. The screened molecules were docked with the receptor proteins, and six hit compounds were selected by visual inspection. The ADME (absorption, distribution, metabolism, excretion) and toxicity of the obtained screened compounds were analyzed to examine the properties of the drugs. Finally, one molecules were screened as potential inhibitors of FGFR1(Fig. 1).

Materials And Methods

Database Preparation

From the RCSB protein database (www.rcsb.org) Download the X-ray crystal structure of the inhibitor-containing complex FGFR1 (PDB ID: 4ZSA). 86956 compounds were downloaded from Zinc database for screening potential inhibitors of FGFR1. Discovery Studio was used to define the physical properties of the compound library so that its molecular mass does not exceed 500, the lipid-water partition coefficient (ClogP) is less than 5, the number of hydrogen bond donors is not more than five, the number of hydrogen bond acceptors is not more than 10, and the number of rotatable bonds is not more than 10 [14, 15]. The SMART's filtration is then used to filter out molecules containing unwanted functional groups .

Generation of Pharmacophore Models

Compound Preparations

Thirty-two FGFR1 receptor inhibitors were collected from previously published literature [16–18]. These compounds obtained the experimental data using the same experimental method, with good biological activity and effective binding to the active site of the FGFR1 receptor. The IC₅₀ values cover four orders of magnitude, including the highest activity, higher activity, intermediate activity, and lower activity. The IC₅₀ values of the most active molecule are less than 0.1 μM, those of the higher active molecule are between

0.1 μM to 1 μM , The IC_{50} values of moderately active molecules range from 1 μM to 10 μM , and those of the lower active molecule are higher than 10 μM [19].

ChemDraw was used to map the structure of the selected FGFR1 receptor inhibition, which was saved as a .sdf file [20]. Discovery Studio was used to convert it into a 3D form with energy minimization, optimized with the MMFF force field, and constructed up to 255 different conformations for each compound in the energy range of 20 kcal mol^{-1} . These conformations were used to construct pharmacophores and to predict the activity of new discovered compounds [21]. Twenty-one compounds were randomly selected as training sets (Fig. 2) and the rest as test sets.

Construction of 3D-QSAR pharmacophore model

The pharmacophore characteristics of the two compounds with the best activity were analyzed using the Feature Mapping module, and it was found that the hydrogen bond donor (HBD), hydrogen bond acceptor (HBA), negative ion (NI), hydrophobicity (HY) and ring aromaticity (RA) features were well mapped on the training set molecules [22]. Additional PI and RA were added to improve the types of pharmacodynamic characteristics. Using the Conformation module in Discovery Studio, 21 compounds in the training set generate multi-conformations of compounds in Best mode. Each compound generates up to 255 conformations with an energy constraint of 10 kcal mol^{-1} , a minimum interference distance of 1.5, an IC_{50} for activity data selection, and an activity uncertainty factor of 3.0, resulting in 10 pharmacophore models [23, 24].

Verification of Pharmacophore and Database Screening

In the process of pharmacophore evaluation, the cost value is an evaluation of the deviation between the predicted activity value and the real activity value of the pharmacophore model and the complexity of the model, and is an important parameter to evaluate the quality of the model. The smaller the cost value, the stronger the predictive activity of the pharmacophore [25–27]. Fixed cost refers to the situation where the activity value predicted by the generated model is exactly the same as the experimental activity value. Null cost is an extremely unsatisfactory situation, that is, the activity values predicted by the generated model are completely different from the experimental activity values. The cost value of an ideal pharmacophore model should be as close as possible to Fixed cost and away from Null cost. Cost is the difference between cost value and Null cost value. If the cost of the pharmacophore model is greater than 40, it indicates that 75%~90% of the pharmacophore model may reflect the objective situation in a statistical sense. If the cost of the pharmacophore model is greater than 60, it indicates that more than 95% of the pharmacophore model in a statistical sense may reflect the objective situation. In addition, the closer the correlation value is to 1, the more credible the model is [28].

The statistic of the pharmacophore model was obtained by Fischer randomization test. The activity values of the training set molecules were randomly assigned to each molecular, and then the pharmacophore model was constructed using the same method. If the resulting pharmacophore model has similar or better Cost value, RMSD value and correlation coefficient with Hypo01, the evaluated model

is considered to be accidental and not statistically significant. The Fischer randomization test was set at a 95% confidence level to evaluate the ten pharmacophore models generated.

In order to further verify the reliability of the best pharmacophore, the pharmacophore will be evaluated on the test set molecule, and the actual and predicted activity values of the compounds will be compared. Discovery Studio was used to construct a database of small molecules to be screened, map pharmacophores on compounds, select molecules with pharmacodynamic characteristics, and retain molecules with predicted activity values less than 1 μM .

Molecular Docking

Molecular docking is performed based on the LibDock and CDOCKEER modules in Discovery studio 2016 client. LibDock is a fast rigid docking method [29], while CDOCKER is based on the CHARMM energy force field to achieve docking, keeping the receptor rigid while the ligand allows bending during docking [30]. The binding mode, interaction between molecule and active site, binding energy and docking score generated by the co-execution of LibDock and CDOCKER can facilitate us to screen the optimal docking posture.

E-3810, also known as Lucitanib, is a novel small molecule drug that selectively binds RTK to the ATP-binding capsule of VEGFR1-3, PDGFR α/β , and FGFR1-3 receptor tyrosine kinases as competitive inhibitors, thereby impeding receptor dimerization and downstream signaling [31, 32]. E-3810 was added with polar hydrogen and stored after energy optimization using the CHARMM force field.

After removing the B chain and water molecule, adding polar hydrogen, and energy optimization with CHARMM force field, the interacting amino acid residues were inspected, and the original ligands were removed. In chain A, the active sites were defined as Leu484, Lys514, Glu531, Val561, Glu562, Ala564, Leu630, and the radius was set to 10 so that the compounds could accurately dock with the molecules. The compound obtained after pharmacophore-based screening was used as a pair acceptor with E-3810, and the docking of the hit compound with the training set compound was at the protein active site where the crystalline ligand was located [33, 34]. LibDock Score and CDOCKER Energy were used as scoring functions, respectively, and the higher LibDock Score and the lower CDOCKER Energy were the more advantageous combinations. In addition, intermolecular interactions and hydrogen bonding interactions between key amino acid residues of proteins and compounds should also be considered.

ADMET and Toxicity Prediction

The ADMET and TOPKAT modules of DS are used to calculate the ADMET properties of compounds, such as their water solubility, BBB permeability, cytochrome P4502D6 (CYP2D6) inhibition, hepatotoxicity, human intestinal absorption, plasma protein binding (PPB) levels, rodent carcinogenicity, Ames mutagenicity and developmental toxicity potential [26, 35]. According to these properties, the lead compounds can be selected and optimized in the early stage of drug research and development, which is of great significance to improve the success rate of drug development and reduce the waste of funds in the process of drug research and development.

Results

Pharmacophore Model Generation

Cost analysis

21 compounds with different activities were included in the training set, and their IC₅₀ values covered four orders of magnitude to generate 3D-QSAR pharmacophore model. The null cost of the generated pharmacophore was 284.43 and the fixed cost was 65.5657. Since the cost value of the ideal pharmacophore model should be as close as possible to the fixed cost, the larger the Δ cost, the better. Table 1 lists the cost values, Δ cost, RMSD and correlation of the ten pharmacophore models generated. It is not difficult to see that among the 10 pharmacophores generated, hypo01 is the closest to fixed cost, Δ cost value is 195.736 (more than 60). Compared with other pharmacophore models, hypo01 has a certain predictive ability and low complexity. Also, from the analysis of other pharmacophore indexes, hypo01 has the largest correlation coefficient (0.950852) and the minimum RMSD (1.46439), which means that the pharmacophore is better than other models and can better predict the activity value of compounds.

Table 1
Results of Hypogen model generation with the training set data.

HYP0	Total Cost	Δ Cost ¹	RMSD	Correlation	Maximum Fit	Features ²
01	88.694	195.736	1.46439	0.950852	6.66393	HBA,HBA,HYA,HYD
02	90.0072	194.4228	1.49407	0.948795	6.28908	HBA,HBA,HYA,HYD
03	90.8522	193.5778	1.53574	0.9485	5.07409	HBA,HBA,HYA
04	91.346	193.084	1.55899	0.944088	5.34599	HBA,HBA,HYA
05	103.217	181.213	1.88939	0.916699	6.52523	HBA,HYA,HYD
06	110.053	174.377	2.04936	0.9012	5.2013	HBA,HYA,HYD
07	111.188	173.242	1.99687	0.906521	4.69033	HBA,HBA,HYA,HYD
08	112.159	172.271	2.09768	0.896214	5.19966	HBA,HYA,HYD
09	112.685	171.745	2.05	0.901195	5.04447	HBA,HBA,HBD,HYD
10	113.703	170.727	2.13622	0.892124	5.39697	HBA,HBA,HYA

¹ Null cost = 284.43, Fixed cost = 65.5657, Δ Cost = Null cost-Total Cost.

²HBA-Hydrogen bond acceptor, HYA-Hydrophobic aromatic, HYD- Hydrophobic, HBD- Hydrogen bond donor.

The pharmacophore characteristics of hypo01 including four hydrogen bond receptor characteristics (HBA), one hydrophobic characteristic (HYD) and hydrophobic aromatic characteristic (HYA) (Fig. 3). 3D-QSAR pharmacophores can be predicted by the degree of overlap between molecules and models. The activity prediction value (Estimate), matching value (Fit Value) and grade classification of 21 training set compounds are listed in Table 2. The activity levels are divided into four orders of magnitude: the most active molecule ($IC_{50} \leq 0.1 \mu M$, ++++), and the more active molecule ($0.1 \mu M < IC_{50} \leq 1 \mu M$, +++), the moderate activity molecules ($1 \mu m < IC_{50} \leq 10 \mu M$, ++), and lowest activity molecules ($IC_{50} > 10 \mu M$, +). The most active molecules can be well mapped on the pharmacophore model, while other molecules have some pharmacodynamic characteristics that are not mapped.

Table 2
Comparison between the experiment biological (IC_{50}) and the estimated activity (IC_{50}) of the Hypo01 model

NO	IC50 Value(μ M)		Fit Valve	Error ¹	Activity scale ²	
	Experimental	Estimated			Experimental	Estimated
1	77	56	2.95	-1.4	+	+
2	90	34	3.16	-2.7	+	+
3	2	2.9	4.23	1.4	++	++
4	12	28	3.25	2.3	+	+
5	9.7	18	3.44	1.8	++	+
6	14	28	3.25	2	+	+
7	0.03	0.034	6.16	1.1	++++	++++
8	0.08	0.1	5.69	1.3	++++	++++
9	0.28	0.71	4.84	2.5	+++	+++
10	0.46	0.24	5.3	-1.9	+++	+++
11	0.77	0.73	4.83	-1.1	+++	+++
12	1.2	0.85	4.76	-1.4	++	+++
13	1.4	2.6	4.27	1.9	++	++
14	1.5	2.8	4.24	1.8	++	++
15	2.1	2.6	4.27	1.2	++	++
16	3	3.5	4.14	1.2	++	++
17	5.4	2.3	4.33	-2.3	++	++
18	5.8	2.8	4.24	-2	++	++
19	7.1	4.9	4	-1.5	++	++
20	7.3	2.7	4.27	-2.8	++	++
21	3.7	3.4	4.16	-1.1	++	++

¹ The error coefficient is calculated as the ratio of measured activity to estimated activity; + indicates that the estimated IC_{50} is higher than the experimental IC_{50} ; - indicates that the estimated IC_{50} is lower than the experimental IC_{50} .

² The activity scale is the same as that defined in the 2.1. session.

Fischer randomization test

Fischer randomization test method can verify whether the pharmacophore model is statistically significant. By randomly corresponding the activity values to different molecular structures and comparing Hypo01 with 10 other randomly generated models, it is obvious that Hypo01 has a lower cost and a higher correlation coefficient (Fig. 4), which proves that the model we constructed is not a result of chance.

Test Set Analysis

To further verify the predictive ability of the pharmacophore, the pharmacophore was used to predict the activity of the test set, which contained 11 compounds with different activities. The difference between the predicted activity value and the experimental activity value was compared. The regression analysis (Fig. 5) demonstrated that the pharmacophore had a good predictive ability.

Database screening

The Zinc database containing 86 976 molecules was obtained by Lipinski's Rule of Five, and then SMART's filtration was used to reduce the number of molecules to 82 663. Using the validated hypo01 model as a 3D query, the processed database was screened, and 16181 molecules could be mapped on hypo01, of which 2763 molecules had predicted activity values less than 1 μM .

Molecular Docking

The FGFR1 protein complex (PDB ID: 4ZSA) contains the entire protein and its original ligand, which is immobilized at the active site of the FGFR1 protein, and key amino acid residues are determined by non-bonding action to define the active site. The ligand was extracted and then docked with the FGFR1 protein, with a high degree of molecular overlap and a RMSD value of 1.7. It can be concluded that the LibDock docking has a high reliability and can be used for the next virtual screening.

The well-known inhibitor of FGFR, E-3810, can effectively inhibit the activity of FGFR1, which is used as a reference compound. E-3810 inhibits the activity of FGFR1 by forming hydrogen bonds between methylamino group, phenoxy group and key residues such as Gly485, Asp641, Glu562, Glu531 on the phenyl ring (Figure 6). The docking score of E-3810 and FGFR1 is 145.145. A series of compounds with docking score greater than this value were screened using LibDock.

In LibDock, 659 compounds were successfully docked, and these compounds were docked on the receptor protein using the CDOCKER module. We selected the molecules with the highest LibDock score and -CDOCKER energy score for visual observation, and found 6 small molecules, and the docking results of 6 small molecules are shown in Table 3.

Table 3
The docking results of 7 potential FGFR1 inhibitors.

Name	2D Structure	LibDock Score	-CDOCKER Energe	Estimated IC ₅₀
ZINC000038579824		158.868	44.3841	0.110979
ZINC000230274254		151.735	12.6032	0.143403
ZINC000011878462		146.958	30.956	0.181035
ZINC000219521723		148.581	31.1579	0.511363
ZINC000072436123		148.899	18.813	0.877224
ZINC000097657383		150.456	20.7042	0.88008
E-3810		145.145	3.35382	1.63942

All the selected molecules can be linked to the key residues of the receptor protein, thus inhibiting the activity of the protein. The hydrogen bond interaction between the six molecules and the protein is shown in Table 4.

Table 4
Hydrogen bond interaction parameters of each compound

Name	Donor Atom	Receptor Atom	H-Bond Distance (Å)
ZINC000038579824	Asp641:HN	ZINC000038579824	3.04
	Phf642:HN	O2	2.39
	Asn568:HD21	N25	2.20
ZINC000230274254	Phe489:O	H43	2.79
	Asp641:HN	O15	2.96
	Ala564:HN	N11	2.60
	Ala564:O	H28	1.57
ZINC000011878462	Ala564:HN	O6	2.70
	Ala564:O	H31	2.45
	Ala564:O	H33	2.86
ZINC000219521723	Asp641:HN	O24	2.01
	Ala564:O	H28	2.74
	Tyr563:HA	O3	2.89
ZINC000072436123	Phe642:HN	N23	2.09
	Asp641:HA	N24	2.15
	Ala:564:O	H41	2.76
ZINC000097657383	Ala564:HN	O10	2.53
	Gly643:HN	N27	2.41
	Asp641:OD1	H46	2.92

The docking score was 158.868 for ZINC000038579824 (2-(1,3-dimethyl-2,6-dioxo-1,2,3,6-tetrahydro-7H-purin-7-yl)-N-(2-((6-methoxy-3,4-dihydroisoquinolin-2(1H)-yl)sulfonyl)ethyl)acetamide) has - 44.3841 CDOCKER energy, and its predicted activity value is only 0.110979, which can be well mapped on the pharmacophore (Fig. 7A). The oxygen atom at the end and the nitrogen atom on the imidazole group form hydrogen bonds with Phf64 2 and Asn5688, respectively (Fig. 8A). ZINC000230274254 ((R)-3-(2,5-dimethyl-7-oxo-4,7-dihydro-[1, 2, 4]triazolo[1,5-a]pyrimidin-6-yl)-N-(1-hydroxy-3-phenylpropan-2-yl)propanamide) can form four hydrogen bonds with Phe489, Ala564, Asp641 (Fig. 8B), and the pharmacophore predicted activity is 0.143403 (Fig. 7B). ZINC000011878462 ((R)-2-(3-oxo-1-phenethylpiperazin-2-yl)-N-(2-(2-oxooxazolidin-3-yl) ethyl) acetamide) had higher docking score and lower CDOCKER energy, formed three hydrogen bonds on Ala564 (Fig. 8C), and could bind stably on the pharmacophore (Fig. 7C). ZINC000219521723 ((R)-2-(5-phenyl-2H-tetrazol-2-yl)-N-(2-(5-(tetrahydrofuran-2-

yl)-1,2,4-oxadiazol-3-yl)ethyl)acetamide) with docking score of 148.581 formed three hydrogen bonds with Asp641, Ala564, Tyr563, Oxadiazole and Phf489 formed a pi-pi conjugate system (Fig. 8D), and tetrazolium group mapped on HBA (Fig. 7D). ZINC000072436123 ((R)-2-(4-(2-(3-(1H-tetrazol-1-yl)phenoxy) acetyl) morpholin-3-yl)-N, N-dimethylacetamide) formed three hydrogen bond interactions with Ala564, Asp641, Phe642 (Fig. 8E), and the tetrazolium group was mapped on HBA (Fig. 7E). ZINC000097657383 ((R)-N-(3-(1H-imidazol-1-yl)propyl)-2,7-dioxo-N-(pyridin-3-ylmethyl)-1,3-diazepane-4-carboxamide) with docking score of 150.456 formed three hydrogen bonds with Ala564, Gly643, Asp641, and benzene ring formed a pi-pi conjugate system with Phf642 (Fig. 8F), which could be mapped on the pharmacophore (Fig. 7F).

ADMET and Toxicity Prediction

The ADMET module of DS2016 was used to predict the absorption, distribution, metabolism, excretion level of all selected small molecules and E-3810. The results are shown in Table 5. The solubility level (in water at 25 °C) indicates that all compounds are soluble in water. For human intestinal absorption, except ZINC00000385579824, it has good absorption level. CYP2 D6 is one of the important enzymes involved in drug metabolism. All compounds are non-inhibitors of cytochrome P4502D6 (CYP2D6). For hepatotoxicity, compared with E-3810 (toxic), ZINC0002219521723 was predicted to be toxic, and the other compounds were not hepatotoxic (Fig. 9).

Table 5
ADME prediction

Name	Solubility Level ¹	BBB Level ²	Absorption Level ³	CYP2D6 ⁴	Hepatotoxicity ⁵	PPB Level ⁶
ZINC000038579824	3	4	1	0	0	0
ZINC000230274254	3	3	0	0	0	0
ZINC000011878462	4	4	0	0	0	0
ZINC000219521723	3	4	0	0	0	1
ZINC000072436123	4	3	0	0	0	0
ZINC000097657383	4	4	0	0	1	0
E-3810	2	3	0	0	1	1

¹ footer. a. Solubility level: 0 (extremely low); 1 (very low, but possible); 2 (low); 3 (good). ² BBB(Blood Brain Barrier level)Level: 0(Very High Penetrant); 1(High); 2(Middle); 3(Low); 4(undefined). ³ Absorption (Human Intestinal Absorption) Level: 0(good); 1(moderate); 2(Poor); 3(Very Poor). ⁴ CYP2D6(Cytochrome P4502D6) level: 0(non-inhibitor); 1(inhibitor). ⁵ Hepatotoxicity: 0 (Nontoxic); 1 (Toxic). ⁶ PPB (Plasma Protein Binding): 0 (Binding is < 90%); 1 (Binding is > 90%); 2 (Binding is > 95%).

To check the safety of the compound, the toxicity of all selected small molecules and E-3810 was predicted using the TOPKAT module, as shown in Table 6. The results showed that all compounds except ZINC0000038579824 were predicted to be non-mutagens. All of the three compounds have good aerobic biodegradation performance, four compounds have no potential developmental toxicity, and one is a non-carcinogen.

Table 6
Toxicity prediction

Name	Ames ¹	Mouse NTP ²		DTP ³	Aerobic Biodegradability ⁴
		Male	Female		
ZINC000038579824	1	1	1	1	0
ZINC000230274254	0	0	0	0	1
ZINC000011878462	0	1	0	0	1
ZINC000219521723	0	1	0	0	0
ZINC000072436123	0	0	1	1	0
ZINC000097657383	0	1	1	0	1
E-3810	0	1	1	1	1

¹ Mutagenicity (Ames text): 0 (non-toxic), 1 (toxic); ² Rodent Carcinogenicity: 0 (non-carcinogenicity), 1 (carcinogenicity); ³ Developmental Toxicity Potential: 0 (non-toxic), 1 (toxic); ⁴ Aerobic Biodegradability: 0 (non-aerobic biodegradability), 1 (aerobic biodegradability).

Therefore, ZINC000230274254 can not inhibit CYP2D6, has no hepatotoxicity, mutagenicity and potential developmental toxicity, and has good aerobic biodegradation performance and absorption level. It is predicted to be a safe candidate drug and selected for further research.

Discussion

Abnormal expression of FGFR1 is one of the causes of tumorigenesis and provides a new direction for cancer treatment. The designed drugs need to act selectively on the target and have high inhibitory activity. Therefore, based on some known inhibitors of FGFR1, we attempted to construct a 3D-QSAR pharmacophore model. Among the ten pharmacophores constructed, Hypo01 includes four pharmacodynamic characteristics: HBA, HBA, HYA and HYD, with significant cost differences, higher correlation coefficients and smaller root mean square deviations. Through cost analysis, Fisher random validation, and test set validation, the ability of this Hypo01 to predict activity was confirmed. We downloaded 86 976 molecules from the ZINC database, and Lipinski's Rule of Five and SMART's filter preliminarily filtered the compounds, using pharmacophores to screen out compounds with a predicted activity value of less than 1 μ M. To determine the docking conformation of the molecule to the receptor, the obtained compounds were molecularly docked, and a series of compounds were obtained by

comparing with the known inhibitor of FGFR1, E-3810. Among them, the molecules with the best conformation are ZINC000038579824, ZINC000230274254, ZINC000011878462, ZINC000219521723, ZINC000072436123, ZINC000097657383, which are more capable of interacting with the hydrogen bond of the key amino acid residues of the receptor and can be well mapped on the pharmacophore. Through ADME and toxicity analysis, it was found that ZINC000230274254, a non-toxic compound, could produce better absorption and metabolism levels, and had strong inhibitory activity. We believe that the selected compounds may be safe candidates for FGFR1, which has certain reference significance for the future research and development of FGFR1-related inhibitors.

Conclusion

ZINC000230274254, a non-toxic compound with good bioavailability, can selectively combine with FGFR1 to produce inhibitory effect, and has good inhibitory activity. The research and development of antitumor drugs has a good reference significance.

Abbreviations

ADME

absorption distribution metabolism excretion;

FGF

fibroblast growth factor;

FGFR

fibroblast growth factor receptor;

HBA

hydrogen bond acceptor;

HYA

hydrophobic aromatic;

HYD

hydrophobic;

HBD

hydrogen bond donor;

QSAR

quantitative structure - activity relationship.

Declarations

Availability of data and materials

The datasets and materials used during the current study are available from the corresponding author on reasonable request.

Competing Interests

The authors declare no conflict of interest.

Funding

This work was supported by Applied Basic Research Program of Sichuan Province Sichuan (No. 2019YJ0549), Sichuan Science and Technology Department Seedling engineering program (No. 201958), Sichuan Science and Technology program (No. 2019YFG0154).

Authors contributions

LH and ZZT conceived and designed the experiments; LH, XLF, HXW and BT performed the experiments; LH, YRX, ZQZ and CXZ analyzed the data; YJW, HC, XLW, HPY, ZS and TLB contributed analysis tools; LH and ZZT wrote the paper. ZZT are the leaders of the project, guiding the experimental design, data analysis, paper writing and revision. All authors read and approved the final manuscript.

Acknowledgements

Not applicable.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A (2018) Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Cancer J Clin* 68:394–424. doi:10.3322/caac.21492
2. Grosso G, Bella F, Godos J, Sciacca S, Del Rio D, Ray S, Galvano F, Giovannucci EL (2017) Possible role of diet in cancer: systematic review and multiple meta-analyses of dietary patterns, lifestyle factors, and cancer risk. *Nutrition reviews* 75:405–419. doi:10.1093/nutrit/nux012
3. Kumar B, Singh S, Skvortsova I, Kumar V (2017) Promising Targets in Anti-cancer Drug Development: Recent Updates. *Curr Med Chem* 24:4729–4752. doi:10.2174/0929867324666170331123648
4. Yu P, Wilhelm K, Dubrac A, Tung JK, Alves TC, Fang JS, Xie Y, Zhu J, Chen Z, De Smet F, Zhang J, Jin SW, Sun L, Sun H, Kibbey RG, Hirschi KK, Hay N, Carmeliet P, Chittenden TW, Eichmann A, Potente M, Simons M (2017) FGF-dependent metabolic control of vascular development. *Nature* 545:224–228. doi:10.1038/nature22322
5. Xue WJ, Li MT, Chen L, Sun LP, Li YY (2018) Recent developments and advances of FGFR as a potential target in cancer. *Future medicinal chemistry* 10:2109–2126. doi:10.4155/fmc-2018-0103
6. Kato H, Nakagama H (2014) FGF receptors: cancer biology and therapeutics. *Medicinal research reviews* 34:280–300. doi:10.1002/med.21288
7. Bockorny B, Rusan M, Chen W, Liao RG, Li Y, Piccioni F, Wang J, Tan L, Thorner AR, Li T, Zhang Y, Miao C, Ovesen T, Shapiro GI, Kwiatkowski DJ, Gray NS, Meyerson M, Hammerman PS, Bass AJ

- (2018) RAS-MAPK Reactivation Facilitates Acquired Resistance in FGFR1-Amplified Lung Cancer and Underlies a Rationale for Upfront FGFR-MEK Blockade. *Mol Cancer Ther* 17:1526–1539. doi:10.1158/1535-7163.mct-17-0464
8. Lau WM, Teng E, Huang KK, Tan JW, Das K, Zang Z, Chia T, Teh M, Kono K, Yong WP, Shabbir A, Tay A, Phua NS, Tan P, Chan SL, So JBY (2018) Acquired Resistance to FGFR Inhibitor in Diffuse-Type Gastric Cancer through an AKT-Independent PKC-Mediated Phosphorylation of GSK3 β . *Mol Cancer Ther* 17:232–242. doi:10.1158/1535-7163.mct-17-0367
 9. Xie X, Lin J, Zhong Y, Fu M, Tang A (2019) FGFR(3S249C) mutation promotes chemoresistance by activating Akt signaling in bladder cancer cells. *Experimental therapeutic medicine* 18:1226–1234. doi:10.3892/etm.2019.7672
 10. Weiss J, Sos ML, Seidel D, Peifer M, Zander T, Heuckmann JM, Ullrich RT, Menon R, Maier S, Soltermann A, Moch H, Wagener P, Fischer F, Heynck S, Koker M, Schöttle J, Leenders F, Gabler F, Dabow I, Querings S, Heukamp LC, Balke-Want H, Ansén S, Rauh D, Baessmann I, Altmüller J, Wainer Z, Conron M, Wright G, Russell P, Solomon B, Brambilla E, Brambilla C, Lorimier P, Sollberg S, Brustugun OT, Engel-Riedel W, Ludwig C, Petersen I, Sängler J, Clement J, Groen H, Timens W, Sietsma H, Thunnissen E, Smit E, Heideman D, Cappuzzo F, Ligorio C, Damiani S, Hallek M, Beroukhim R, Pao W, Klebl B, Baumann M, Buettner R, Ernestus K, Stoelben E, Wolf J, Nürnberg P, Perner S, Thomas RK (2010) Frequent and focal FGFR1 amplification associates with therapeutically tractable FGFR1 dependency in squamous cell lung cancer. *Science translational medicine* 2:62ra93. doi:10.1126/scitranslmed.3001451
 11. Fischbach A, Rogler A, Erber R, Stoehr R, Poulosom R, Heidenreich A, Schneevoigt BS, Hauke S, Hartmann A, Knuechel R, Veeck J, Gaisa NT (2015) Fibroblast growth factor receptor (FGFR) gene amplifications are rare events in bladder cancer. *Histopathology* 66:639–649. doi:10.1111/his.12473
 12. Turner N, Pearson A, Sharpe R, Lambros M, Geyer F, Lopez-Garcia MA, Natrajan R, Marchio C, Iorns E, Mackay A, Gillett C, Grigoriadis A, Tutt A, Reis-Filho JS, Ashworth A (2010) FGFR1 amplification drives endocrine therapy resistance and is a therapeutic target in breast cancer. *Cancer research* 70:2085–2094. doi:10.1158/0008-5472.can-09-3746
 13. Katoh M (2016) Therapeutics Targeting FGF Signaling Network in Human Diseases. *Trends Pharmacol Sci* 37:1081–1096. doi:10.1016/j.tips.2016.10.003
 14. Chagas CM, Moss S, Alisaraie L (2018) Drug metabolites and their effects on the development of adverse reactions: Revisiting Lipinski's Rule of Five. *International journal of pharmaceutics* 549:133–149. doi:10.1016/j.ijpharm.2018.07.046
 15. Shahin R, Swellmeen L, Shaheen O, Aboalhaja N, Habash M (2016) Identification of novel inhibitors for Pim-1 kinase using pharmacophore modeling based on a novel method for selecting pharmacophore generation subsets. *J Comput Aided Mol Des* 30:39–68. doi:10.1007/s10822-015-9887-7
 16. Çoban G, Aydın F Köse (2019) Synthesis, biological evaluations and molecular modelling studies of novel indolin-2-ones designing as FGFR inhibitors. *Saudi Pharm J* 27:952–967.

doi:10.1016/j.jsps.2019.07.004

17. Sun L, Tran N, Liang C, Tang F, Rice A, Schreck R, Waltz K, Shawver LK, McMahon G, Tang C (1999) Design, synthesis, and evaluations of substituted 3-[(3- or 4-carboxyethylpyrrol-2-yl)methylidene]indolin-2-ones as inhibitors of VEGF, FGF, and PDGF receptor tyrosine kinases. *Journal of medicinal chemistry* 42:5120–5130. doi:10.1021/jm9904295
18. Turner LD, Summers AJ, Johnson LO, Knowles MA, Fishwick CWG (2017) Identification of an Indazole-Based Pharmacophore for the Inhibition of FGFR Kinases Using Fragment-Led de Novo Design. *ACS medicinal chemistry letters* 8:1264–1268. doi:10.1021/acsmchemlett.7b00349
19. Zhou Y, Di B, Niu MM (2019) Structure-Based Pharmacophore Design and Virtual Screening for Novel Tubulin Inhibitors with Potential Anticancer Activity. *Molecules* 24: doi:10.3390/molecules24173181
20. Sastry GM, Adzhigirey M, Day T, Annabhimoju R, Sherman W (2013) Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J Comput Aided Mol Des* 27:221–234. doi:10.1007/s10822-013-9644-8
21. Pal S, Kumar V, Kundu B, Bhattacharya D, Preethy N, Reddy MP, Talukdar A (2019) Ligand-based Pharmacophore Modeling, Virtual Screening and Molecular Docking Studies for Discovery of Potential Topoisomerase I Inhibitors. *Comput Struct Biotechnol J* 17:291–310. doi:10.1016/j.csbj.2019.02.006
22. Vuorinen A, Schuster D (2015) Methods for generating and applying pharmacophore models as virtual screening filters and for bioactivity profiling. *Methods* 71:113–134. doi:10.1016/j.ymeth.2014.10.013
23. Atatreh N, Ghattas MA, Bardaweel SK, Rawashdeh SA, Sorkhy MA (2018) Identification of new inhibitors of Mdm2-p53 interaction via pharmacophore and structure-based virtual screening. *Drug Des Devel Ther* 12:3741–3752. doi:10.2147/DDDT.S182444
24. John S, Thangapandian S, Arooj M, Hong JC, Kim KD, Lee KW (2011) Development, evaluation and application of 3D QSAR Pharmacophore model in the discovery of potential human renin inhibitors. *BMC Bioinform* 12(Suppl 14):S4. doi:10.1186/1471-2105-12-s14-s4
25. Du S, Yang B, Wang X, Li WY, Lu XH, Zheng ZH, Ma Y, Wang RL (2019) Identification of potential leukocyte antigen-related protein (PTP-LAR) inhibitors through 3D QSAR pharmacophore-based virtual screening and molecular dynamics simulation. *J Biomol Struct Dyn*:1–14 doi:10.1080/07391102.2019.1676825
26. Ghayas S, Ali Masood M, Parveen R, Aquib M, Farooq MA, Banerjee P, Sambhare S, Bavi R (2019) 3D QSAR pharmacophore-based virtual screening for the identification of potential inhibitors of tyrosinase. *J Biomol Struct Dyn*:1–12 doi:10.1080/07391102.2019.1647287
27. Vrontaki E, Kolocouris A (2018) Pharmacophore Generation and 3D-QSAR Model Development Using PHASE. *Methods in molecular biology*. (Clifton NJ) 1824:387–401. doi:10.1007/978-1-4939-8630-9_23
28. Markt P, Feldmann C, Rollinger JM, Raduner S, Schuster D, Kirchmair J, Distinto S, Spitzer GM, Wolber G, Laggner C, Altmann KH, Langer T, Gertsch J (2009) Discovery of novel CB2 receptor ligands by a

- pharmacophore-based virtual screening workflow. *Journal of medicinal chemistry* 52:369–378. doi:10.1021/jm801044g
29. Xu W, Lucke AJ, Fairlie DP (2015) Comparing sixteen scoring functions for predicting biological activities of ligands for protein targets. *J Mol Graph Model* 57:76–88. doi:10.1016/j.jmgm.2015.01.009
30. Gagnon JK, Law SM, Brooks CL 3rd (2016) Flexible CDOCKER: Development and application of a pseudo-explicit structure-based docking method within CHARMM. *J Comput Chem* 37:753–762. doi:10.1002/jcc.24259
31. Ghedini GC, Ronca R, Presta M, Giacomini A (2018) Future applications of FGF/FGFR inhibitors in cancer. *Expert Rev Anticancer Ther* 18:861–872. doi:10.1080/14737140.2018.1491795
32. 10.1016/j.neo.2016.11.008
Guffanti F, Chilà R, Bello E, Zucchetti M, Zangarini M, Ceriani L, Ferrari M, Lupi M, Jacquet-Bescond A, Burbridge MF, Pierrat MJ, Damia G (2017) In Vitro and In Vivo Activity of Lucitanib in FGFR1/2 Amplified or Mutated Cancer Models. *Neoplasia (New York, N.Y.)* 19:35–42. doi:10.1016/j.neo.2016.11.008
33. Fei J, Zhou L, Liu T, Tang XY (2013) Pharmacophore modeling, virtual screening, and molecular docking studies for discovery of novel Akt2 inhibitors. *International journal of medical sciences* 10:265–275. doi:10.7150/ijms.5344
34. Zhou X, Yu S, Su J, Sun L (2016) Computational Study on New Natural Compound Inhibitors of Pyruvate Dehydrogenase Kinases. *Int J Mol Sci* 17:340. doi:10.3390/ijms17030340
35. Zekri A, Harkati D, Kenouche S, Saleh BA (2020) QSAR modeling, docking, ADME and reactivity of indazole derivatives as antagonizes of estrogen receptor alpha (ER- α) positive in breast cancer. *J Mol Struct* 1217:128442. doi:10.1016/j.molstruc.2020.128442

Figures

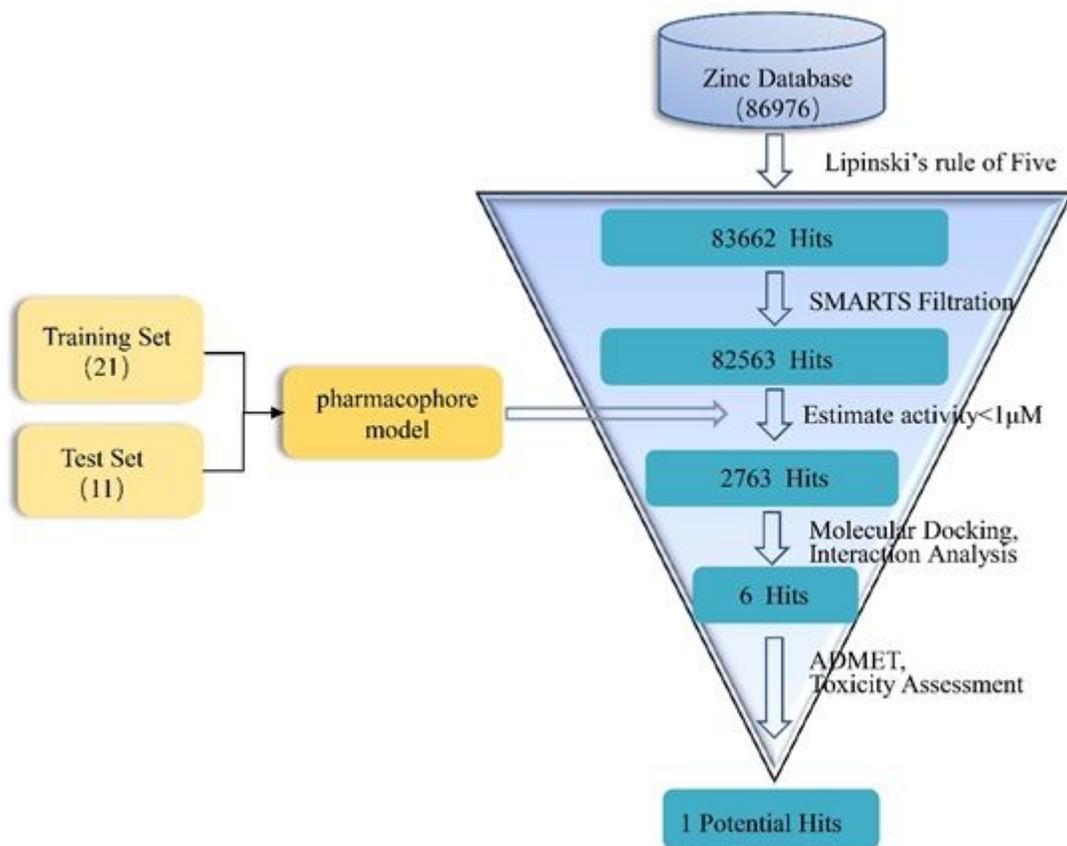


Figure 1

Schematic representation of the virtual screening process implemented to screen potential FGFR1 inhibitors.

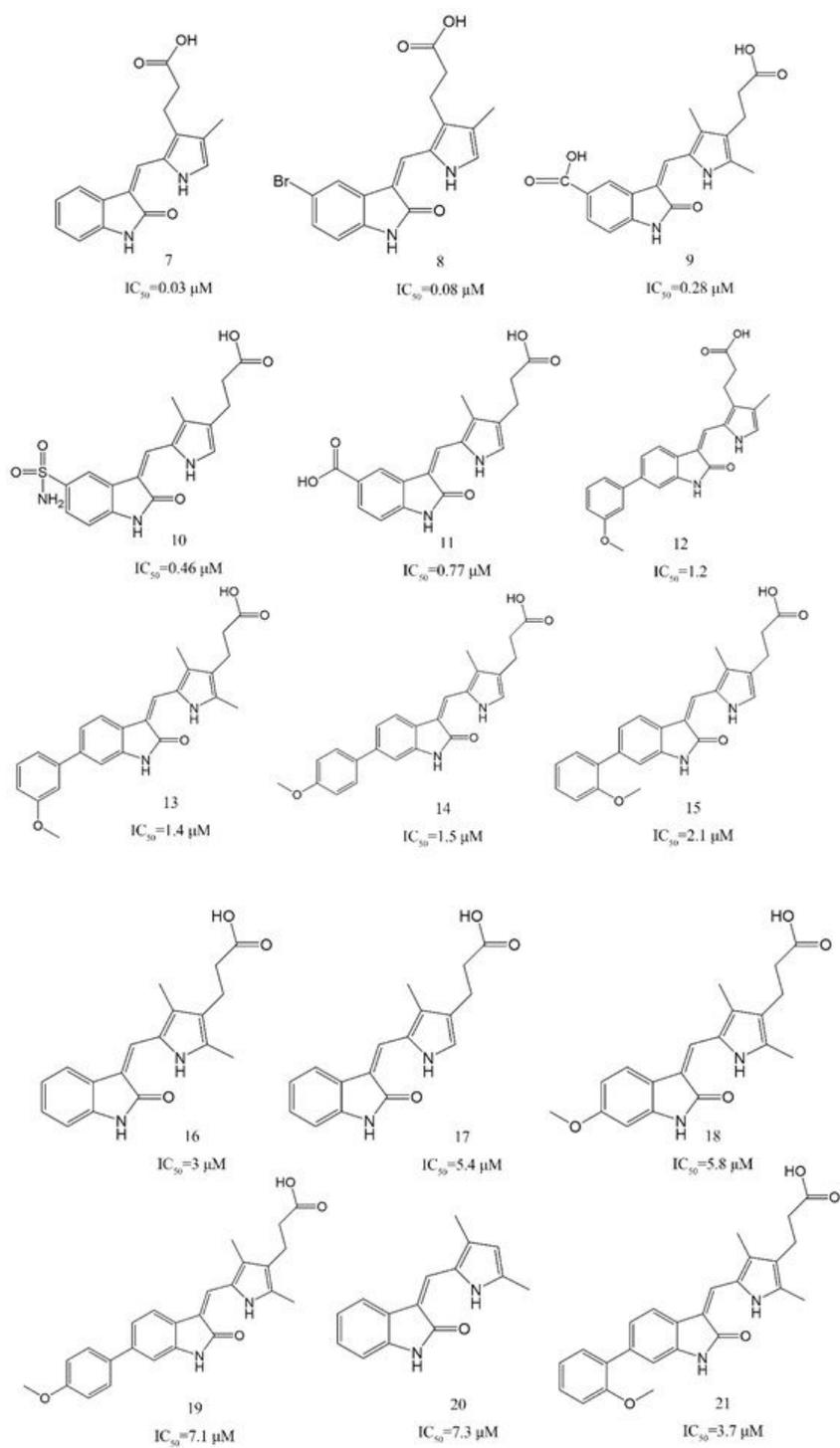


Figure 2

Training set molecules along with their corresponding IC_{50} values.

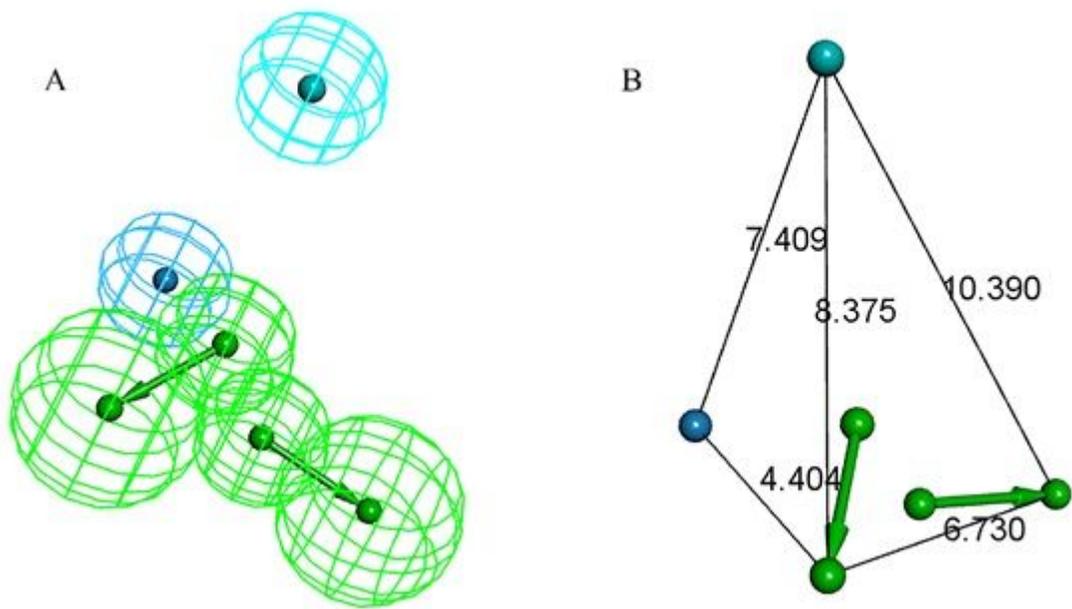


Figure 3

The best Hypogen Pharmacophore mode. (A) Features present in Hypo01 (B) Distance between features. (Light blue: hydrophobic, Green: hydrogen bond receptor, Dark blue: Hydrophobic aromatic.)

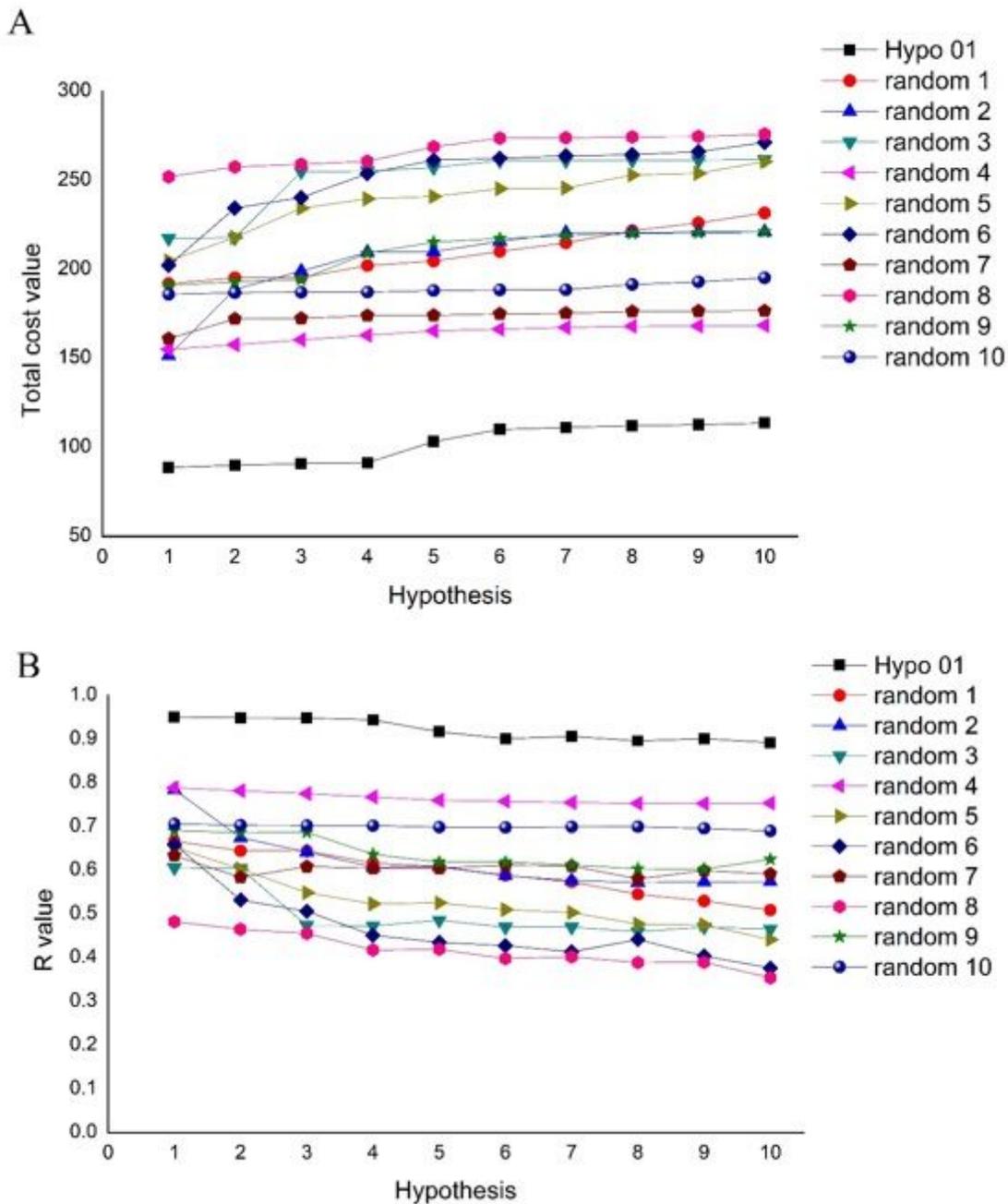


Figure 4

Fisher Validation of Total Cost Value and Correlation Coefficient. (A) Comparison of the total cost value of Hypo01 with the total cost value of 10 pharmacophores randomly generated; (B) Comparison of correlation coefficients of Hypo01 with those of 10 pharmacophores randomly generated.

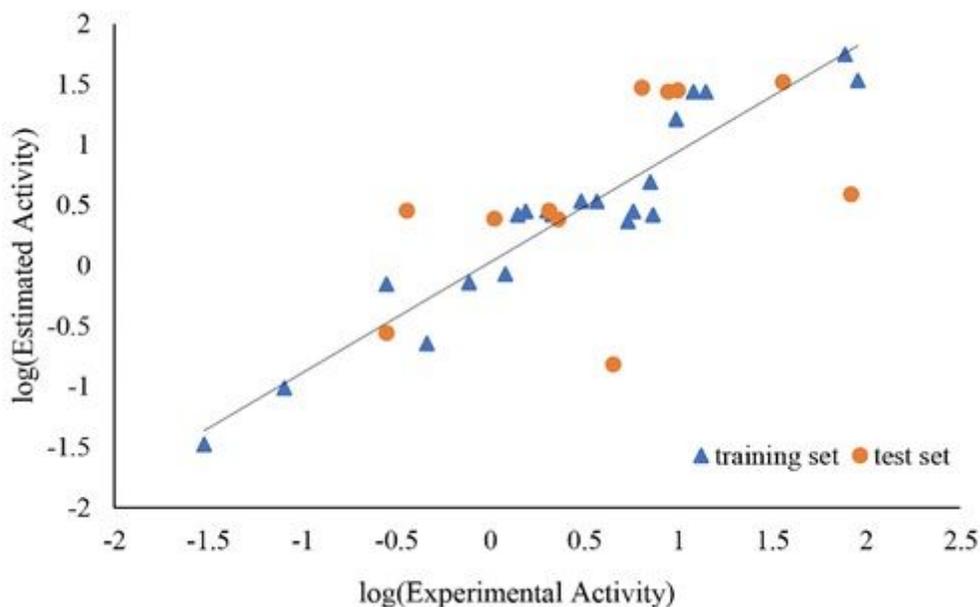


Figure 5

Linear correlation between experimental activity and estimated activity of Hypo01.

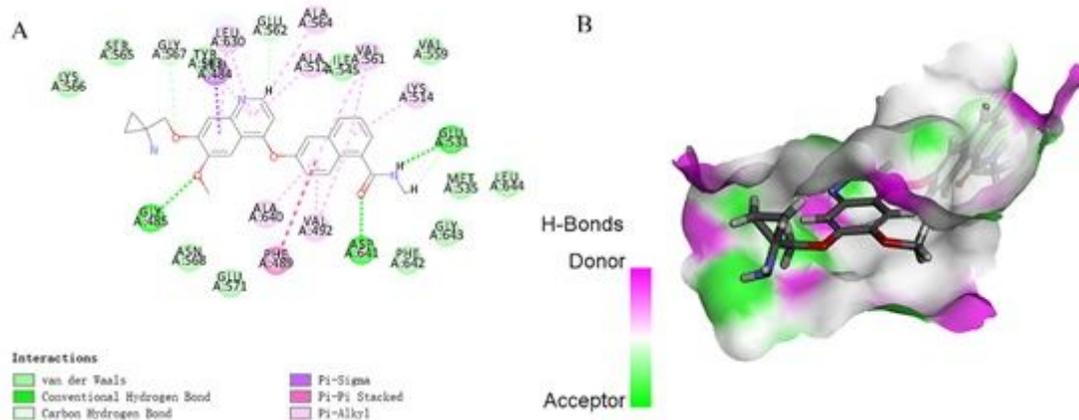


Figure 6

FGFR1 combines with E-3810: (A) 2D schematic diagram I; (B) Distribution of hydrogen bond in binding pocket.

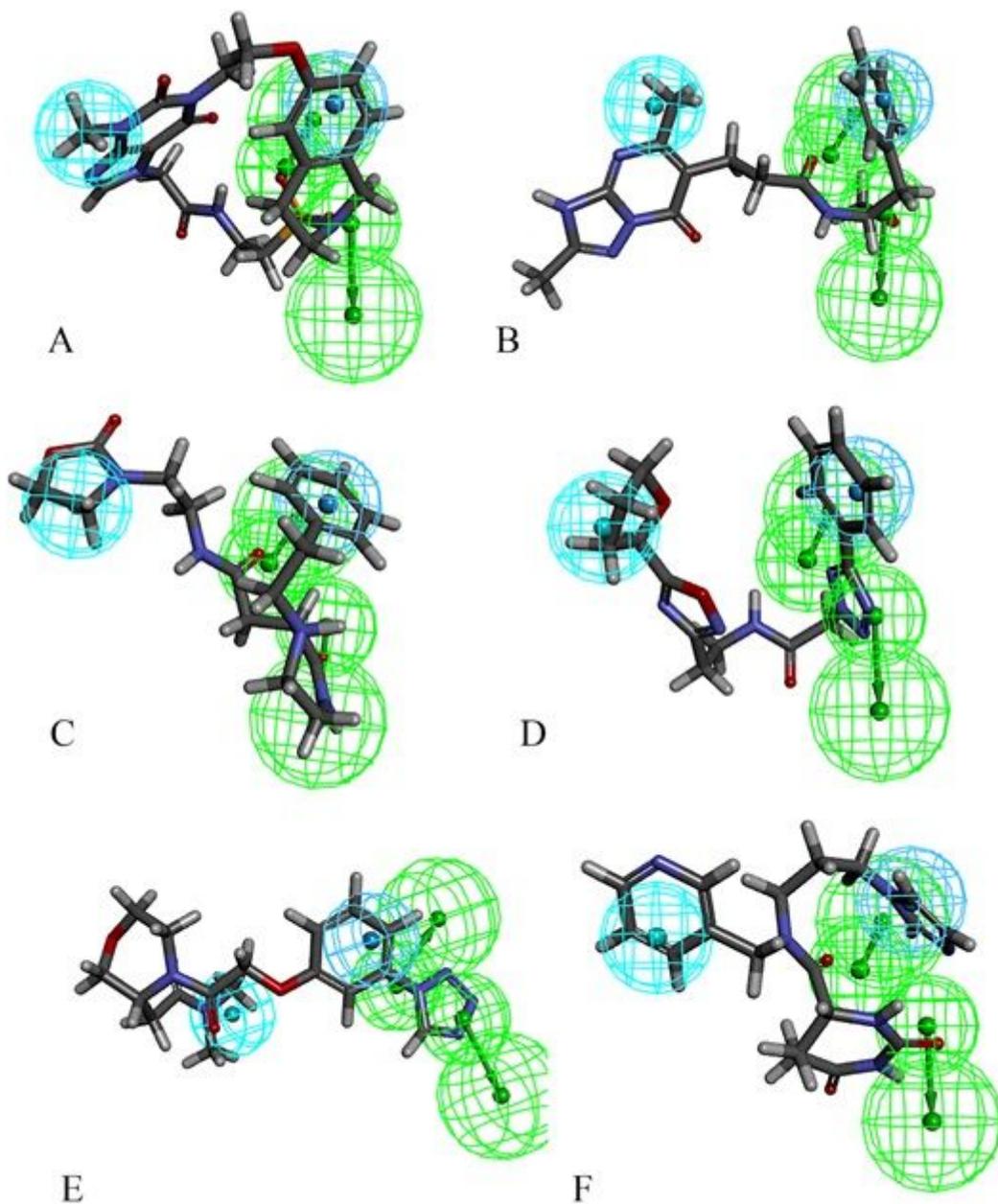


Figure 7

Overlay of lead molecules on the pharmacophore Hypo01. (A)ZINC000038579824; (B)ZINC000230274254; (C)ZINC000011878462; (D)ZINC000219521723; (E)ZINC000072436123; (F) ZINC000097657383.

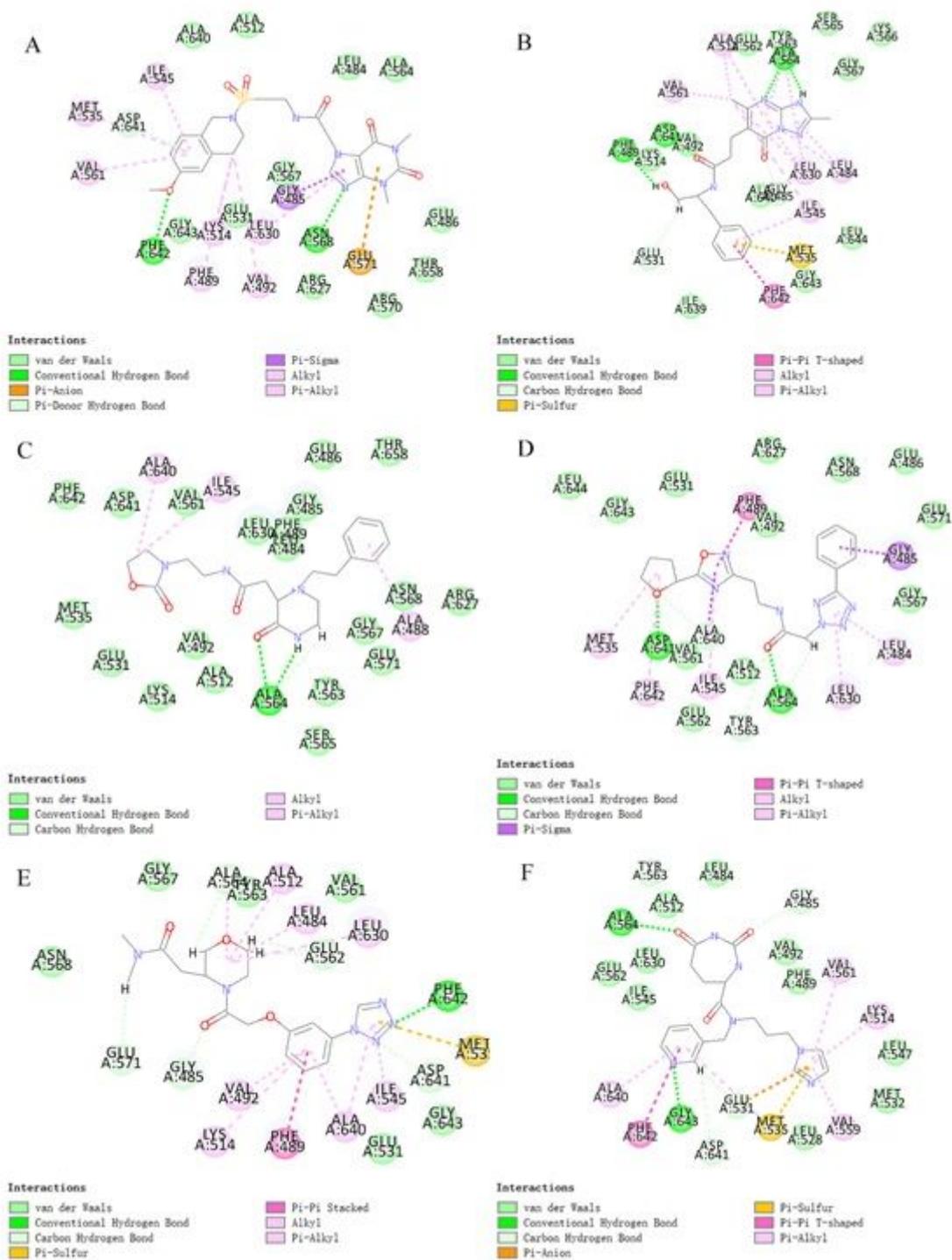


Figure 8

Interaction of Receptor Proteins and Molecules. (A)ZINC000038579824;(B)ZINC000230274254;
 (C)ZINC000011878462; (D)ZINC000219521723; (E)ZINC000072436123; (F) ZINC000097657383.

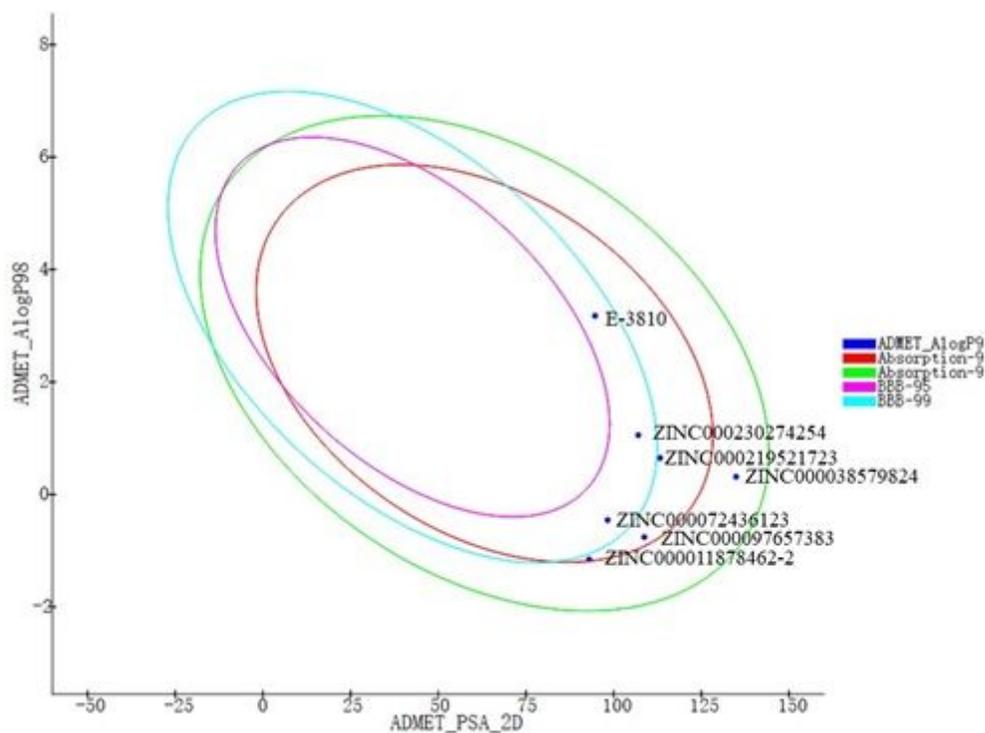


Figure 9

The two ovals in the diagram represent 95% and 99% confidence intervals for the blood-brain barrier permeability (BBB) model and 95% and 99% confidence intervals for the human intestinal absorptivity (HIA) model.