Journal of Molecular Structure 1220 (2020) 128617

Contents lists available at ScienceDirect

Journal of Molecular Structure

journal homepage: http://www.elsevier.com/locate/molstruc

In silico design novel dihydropyrimio[4, 5-d]pyrimidine derivatives as inhibitors for colony-stimulating factor-1 receptor based on 3D-QSAR, molecular docking and molecular dynamics simulation

Han Chu [a, b, 1], Qing-xiu He [a, b, 1], Juan Wang [a, b], Yong Hu [a, b], Yuan-qiang Wang [a, b, **], Zhi-hua Lin [a, b, *]

[a] Department of Pharmacy and Bioengineering, Chongqing University of Technology, Chongqing, 400054, PR China
[b] Key Laboratory of Screening and Activity Evaluation of Targeted Drugs, Chongqing, 400054, PR China

ARTICLE INFO

Article history:
Received 23 March 2020
Received in revised form
1 May 2020
Accepted 3 June 2020
Available online 12 June 2020

Keywords:
Colony-stimulating factor-1 receptor
Inhibitor
dihydropyrimio[4
5-D] pyrimidine derivatives
3D-QSAR
Molecular docking
Molecular dynamics simulation

ABSTRACT

Colony-stimulating factor-1 receptor (CSF-1R) can affect the further development of tumors by regulating the generation and biological activity of mononuclear-phagocytic lineage cells, and is an important potential target for cancer treatment. In this study, we selected 48 dihydropyrimio [4, 5-d]pyrimidine derivatives and performed 3D-QSAR studies on them using CoMFA and CoMSIA model. The results showed that CoMFA ($n = 10$; $q^2 = 0.626$; $r^2 = 0.998$) and CoMSIA ($n = 10$; $q^2 = 0.684$; $r^2 = 0.997$) had good stability and predictability. The relationship between the activity and structure of the compounds was analyzed by counter maps of the steric, electrostatic and hydrophobic fields. Subsequently, molecular docking was used to explore key amino acids and docking modes at the active site. Based on these results, we designed 9 new dihydropyrimio [4, 5-d] pyrimidine derivatives and predicted their activities. The results indicated that these compounds had good predictive activities and ADME/T properties. MD simulation analysis confirmed that the residues Leu588 and Cys666 in the active site played a key role; at the same time, these compounds mainly bind to CSF-1R through hydrophobic interactions and hydrogen bonding. These results provided important reference for the discovery and design of new CSF-1R inhibitors.

1. Introduction

Colony stimulating factor 1 receptor (CSF-1R) is a single chain transmembrane glycoprotein composed of 972 amino acid residues, which belongs to a class III receptor tyrosine kinase (PTK) [1]. It is one of the most important regulatory factors in the immune system. It is encoded by the proto-oncogene c-fms and belongs to the platelet-derived growth factor (PDGF) family [2,3]. CSF-1R is structurally different from other family members. There are five immunoglobulin-like extracellular domains and a kinase insert dividing the intracellular kinase domain into two lobes [4,5]. CSF-1R affects the further development of tumors by regulating the generation and biological activity of monocyte-phagocytic lineage cells. Tumor cells can attract more macrophages into tumor microenvironment by secreting colony stimulating factor-1 (CSF-1), and CSF-1R can induce these tumor-associated macrophages (TAMs) to further promote the development and metastasis of tumors [6,7], such as breast cancer [8], ovarian cancer, colorectal cancer [9], pancreatic cancer and Hodgkin's lymphoma [10]. TAMs play an important role not only in the development and metastasis of tumors, but also in resistance to chemotherapy and radiotherapy [11,12]. Preclinical models have elucidated the critical role of TAMs in cancer development, development and metastasis [13–15]. In addition, blocking the CSF-1/CSF-1R signal can not only directly target tumor cells expressing CSF-1R, but also improve tumor microenvironment. Therefore, the inhibition of TAMs survival/activation by inhibiting CSF-1/CSF-1R signaling has become a very attractive cancer treatment strategy.

At present, it has been reported that several selective or non-

* Corresponding author. Department of Pharmacy and Bioengineering, Chongqing University of Technology, Chongqing, China.
** Corresponding author. Key Laboratory of Screening and activity evaluation of targeted drugs, Chongqing, 400054, PR China.
E-mail addresses: wangyqnn@cqut.edu.cn (Y.-q. Wang), zhlin@cqut.edu.cn (Z.-h. Lin).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.molstruc.2020.128617
0022-2860/

selective small molecule CSF-1R inhibitors have entered different stages of clinical research [16–18]. For example, PLX3397 is an oral selective CSF-1R/c-kit inhibitor for the treatment of patients with recurrent glioblastoma, and is undergoing phase III clinical studies [19,20]. BLZ945 is also a selective small molecule inhibitor of CSF-1R, which has been studied in the treatment of solid tumors [21,22]. However, whether these inhibitors are safe or clinically active, their effectiveness is limited. Therefore, it is still necessary to find new small molecule CSF-1R inhibitors. In this paper, 48 potential 2-Oxo-3,4-dihydropyrimido [4, 5-*d*]pyrimidines were collected as CSF-1R inhibitors, and their good 3D-QSAR models were obtained by CoMFA and CoMSIA method. Then the interaction between CSF-1R and 2-Oxo-3,4-dihydropyrimido [4,5-*d*]pyrimidines was analyzed by molecular docking. Finally, we designed 9 new 2-Oxo-3,4-dihydropyrimido [4, 5-*d*]pyrimidines, and theoretically verified their correct binding to CSF-1R through molecular dynamics simulations. It provides a theoretical basis for the research and development of new inhibitors targeting CSF-1R.

2. Materials and methods

2.1. Preparation of data set

In the present work, 48 Dihydropyrimido [4,5-d] pyrimidine were collected from literature [23] which were divided into training (37 compounds) and test dataset (11 compounds, represented by the symbol "*" superscript) randomly. The bioactivity values of these compounds in vitro were reported as IC_{50} ($Mol{\bullet}L^{-1}$), which were converted to the corresponding pIC_{50} ($pIC_{50} = -\log IC_{50}$). The structure and activity values of the compounds were listed in Table 1.

2.2. Structural optimization and alignment

The structures of all compounds were constructed in sybyl 2.0, and the compounds were optimized using Gasteiger-Huckel charge, Tripos force field and Powll energy gradient method. The maximum optimization limit was 10,000 times, the convergence criterion was 0.005 kJ mol^{-1}, and the rest of the parameters were default. We used the small molecule conformation obtained in the above method as the active conformation for further research. At the same time, we chose the compound 26 with the highest inhibitory activity as the aligned template. After selecting the appropriate Skeleton (Fig. 1), the alignment of all compounds was shown in Fig. 2.

2.3. 3D-QSAR modeling and validation

In this study, we used CoMFA and CoMSIA [24,25] to conduct 3D-QSAR research on 37 compounds in the training set, and constructed a prediction model. In SYBYL 2.0, the field properties of the models were calculated using 3D cubic lattices with a spaced grid of 2.0 Å. For the CoMFA method, it was assumed that ligand–receptor binding was non-covalent. An sp^3 hybridized carbon atom with +1 charge as the probe was used to detect the size and distribution of the steric and electrostatic fields. In the CoMSIA model, besides the steric and electrostatic field, the hydrophobic, H-bond acceptor and donor fields were also introduced. Further, similarity indices were calculated between a probe and each atom of the molecule based on a Gaussian distance function in the CoMSIA method [26]. The optimum number of components (n) and the highest cross-validation correlation coefficient (q^2) for the correlation models were calculated by the leave-one-out (LOO) cross-validation procedure. The conventional multiple correlation coefficients (r^2), standard error of estimate (SEE), and the Fisher test (F) values was obtained by non-cross-validated analysis [27]. The counter maps of the 3D-QSAR model was used to study the effect of molecular structure on biological activity, and guided the design and modification of the structure to improve the activity of compounds.

Furthermore, the Y-randomization of response test [28] was utilized to ensure the robustness of constructed models. In this test, we repeated the 3D-QSAR calculation procedure several times with randomly shuffling activities of the training set. The relatively low Q2 and R2 values built with randomized activities of the training sets, indicating that the constructed models using current modeling method for collected data set were acceptable and reliable [29].

2.4. Molecular docking

Molecular docking is an important method for predicting the optimal binding mode between small molecules and proteins based on the structure of the receptor. The crystal structure of CSF-1R (PDB: 3DPK) was derived from the Protein Data Bank (PDB). Sybyl 2.0 software was used to hydrogenate and charge the CSF-1R crystal structure [30]. At the same time, the original ligand structure (Q8C5) of the protein structure was removed. The amino acid residues within 0.5 nm of the original ligand were used as the docking binding pocket region, and the ligand was used as the reference molecule (Fig. 3). Then, the homologous ligand Q8C5 was redocked in the binding pocket to verify the reliability of the docking method. After that, all compounds were docked into the binding pocket using the same parameters. The binding visualization was performed utilizing PyMOL software.

2.5. ADME/T properties prediction

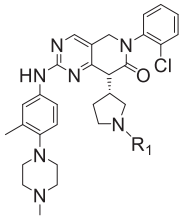
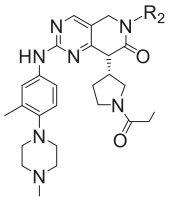
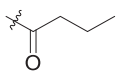
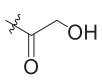
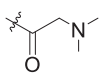
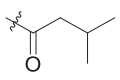
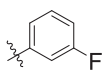
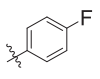
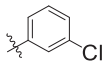
In this study, using the pkCSM online server (http://biosig.unimelb.edu.au/pkcsm/prediction), the ADME/T characteristics of the designed CSF-1R inhibitor could be predicted [30]. For a given compound, the web server predicted the percentage that will be absorbed through the human intestine. Molecules with an absorbance of less than 30% were considered to be poorly absorbed. The higher the volume of distribution (VD), the more the drug was distributed in the tissue rather than in the plasma. If it was lower than 0.71 L kg^{-1} (logVDss < -0.15), the volume of distribution at steady state (VDss) was considered low; if it was higher than 2.81 L kg^{-1}, VDss was considered high (logVDss> 0.45). Predictors evaluated the designed molecule to determine whether it might be metabolized by P450 or might be a cytochrome P450 inhibitor. The total clearance (log mL • (min·kg) $^{-1}$) was used to determine the dose rate to reach steady-state concentration, which was related to bioavailability. Predictive evaluation of AMES toxicity/hepatotoxicity/skin sensitization to predict whether the designed new molecules are toxic.

The synthetic accessibility of the new derivatives was evaluated by using the online tool SwissADME (http://www.swissadme.ch/index.php). The scale of the tool indicates that a score closes to 10 had high structural complexity, and therefore it was difficult to synthesize, and a score of 1 indicates that the synthesis route was relatively easy.

2.6. Molecular dynamics (MD) simulation and free energy analysis

The CSF-1R complexed with drug candidates were set up for MD simulations. Before simulation, molecular mechanics methods are usually used to optimize biological macromolecule systems. The GAFF force field is used for small molecules and the ff99 force field is used for proteins. Then, for example, the complexes were put into 0.15 M NaCl solution within a cubic water box, which included

Table 1
Experimental and predicted activity (pIC_{50}) values of compounds.

No.	Formula	R_1	R_2	Obs.	CoMFA	Err (%)	CoMSIA	Err (%)	score
	 I	 II							
1*	I		–	5.17	4.80	7.16	5.03	2.71	11.19
2	I		–	5.30	5.24	1.13	5.29	0.19	11.31
3	I		–	5.26	5.26	0.00	5.29	0.57	10.23
4*	I		–	5.00	5.04	0.80	4.98	0.40	12.18
5*	I		–	4.70	4.93	4.89	4.92	4.68	10.18
6	I		–	4.68	4.68	0.00	4.67	0.21	10.15
7	I		–	4.55	4.59	0.88	4.62	1.54	10.61
8	I		–	4.51	4.51	0.00	4.47	0.89	10.70
9*	I		–	4.70	5.05	7.45	5.05	7.45	10.39
10	I		–	4.64	4.64	0.00	4.63	0.22	11.76
11*	I		–	4.99	4.76	4.61	4.67	6.41	10.67
12	II	–		5.28	5.32	0.76	5.25	0.57	10.78
13*	II	–		5.25	5.10	2.86	5.15	1.90	11.72
14	II	–		4.82	4.83	0.21	4.79	0.62	10.88
15	II	–		4.74	4.73	0.21	4.79	1.05	10.68
16*	II	–		4.88	4.80	1.64	4.52	7.38	10.12

(*continued on next page*)

Table 1 (*continued*)

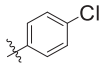
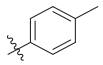
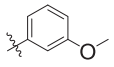
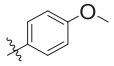
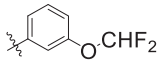
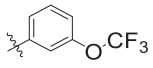
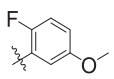
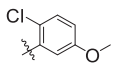
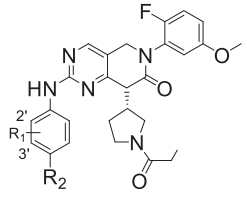
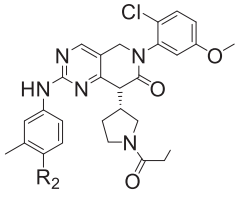
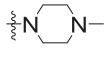
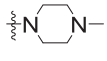
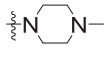
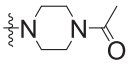
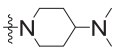
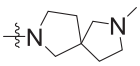
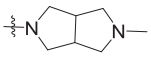
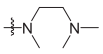
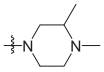
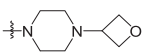
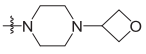
No.	Formula	R_1	R_2	Obs.	CoMFA	Err (%)	CoMSIA	Err (%)	score
17	II	–		4.83	4.83	0.00	4.84	0.21	10.15
18	II	–		4.79	4.79	0.00	4.76	0.63	10.61
19*	II	–		4.69	5.08	8.32	5.06	7.89	11.26
20	II	–		5.29	5.28	0.19	5.30	0.19	12.42
21	II	–		5.35	5.36	0.19	5.36	0.19	10.86
22	II	–		5.24	5.25	0.19	5.23	0.19	9.98
23	II	–		4.55	4.53	0.44	4.56	0.22	10.64
24	II	–		3.82	3.82	0.00	3.82	0.00	10.10
25	II	–		5.35	5.35	0.00	5.35	0.00	10.59
26*	II	–		5.70	5.70	0.00	5.69	0.18	11.37
27*	II	–		5.36	5.54	3.36	5.74	7.09	9.35
28	II	–		5.68	5.67	0.18	5.68	0.00	11.09
	 III		 IV						
29	III	2′-CH3		4.69	4.69	0.00	4.69	0.00	11.68
30	III	2′-OCH3		4.52	4.52	0.00	4.51	0.22	10.38
31	III	3′-OCH3		5.24	5.23	0.19	5.23	0.19	11.72
32	III	3′-F		4.77	4.77	0.00	4.80	0.63	11.96
33	III	3′-Cl		4.83	4.81	0.41	4.82	0.21	10.87
34	III	3′-CF3		4.50	4.51	0.22	4.50	0.00	11.16
35	III	3′-ethyl		5.00	5.00	0.00	5.00	0.00	10.78

Table 1 (*continued*)

No.	Formula	R_1	R_2	Obs.	CoMFA	Err (%)	CoMSIA	Err (%)	score
36*	III	3′-CH3		4.60	4.98	8.26	5.05	9.78	10.36
37	III	3′-CH3		4.59	4.59	0.00	4.59	0.00	11.46
38	III	3′-CH3		5.31	5.31	0.00	5.32	0.19	11.89
39	III	3′-CH3		5.33	5.33	0.00	5.33	0.00	11.59
40*	III	3′-CH3		5.17	5.28	2.13	5.46	5.61	11.95
41	III	3′-CH3		5.21	5.21	0.00	5.20	0.19	11.70
42	III	3′-CH3		5.19	5.19	0.00	5.19	0.00	11.27
43	III	3′-CH3		5.15	5.16	0.19	5.14	0.19	11.57
44	III	3′-CH3		4.99	4.99	0.00	4.98	0.20	11.86
45	III	3′-CH3		5.14	5.14	0.00	5.13	0.19	12.00
46	III	3′-CH3		5.11	5.11	0.00	5.13	0.39	11.56
47	III	3′-CH3		5.47	5.49	0.37	5.47	0.00	11.86
48	IV	–		5.52	5.52	0.00	5.52	0.00	11.68

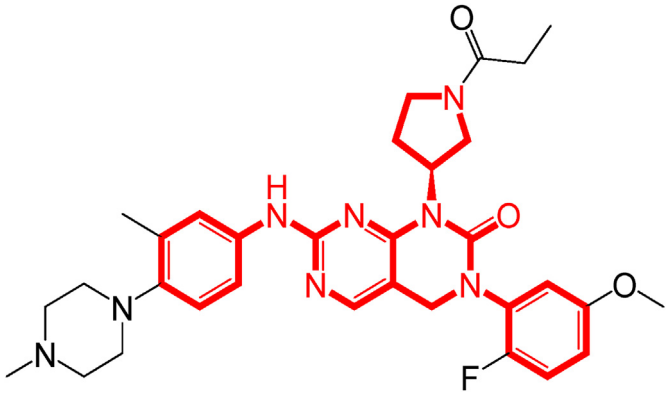


Fig. 1. Skeleton of compound 26.

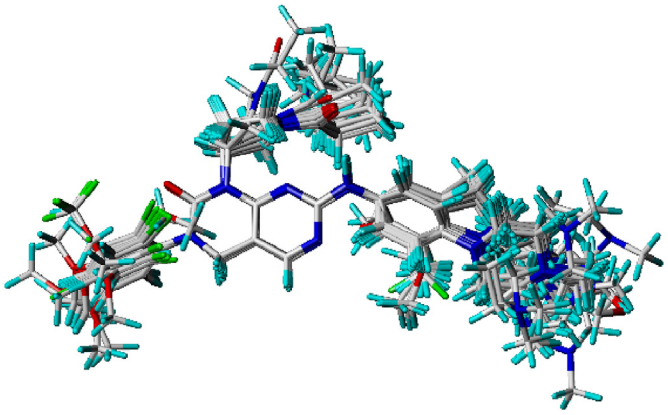


Fig. 2. Alignment of all compounds.

15,481 water molecules, 40 Na^+ ions, and 42 Cl^- ions. The initial configurations of the receptors and ligands were taken from docking studies. The sizes of the initial simulation boxes were ~10^3 Å × 10^3 Å × 10^3 Å. The other systems were set up with the same protocol. MD simulation was performed using the PMEMD. mpi and PMEMD. cuda modules in the AMBER16 software package [31,32]. First, five minimization steps were performed on the system to avoid possible space crashes. Each system was then gradually heated from 0 K to 300 K during the heating phase and maintained at 300 K during the subsequent equilibrium, and production phases. A 2 fs time step was used during the heating phase,

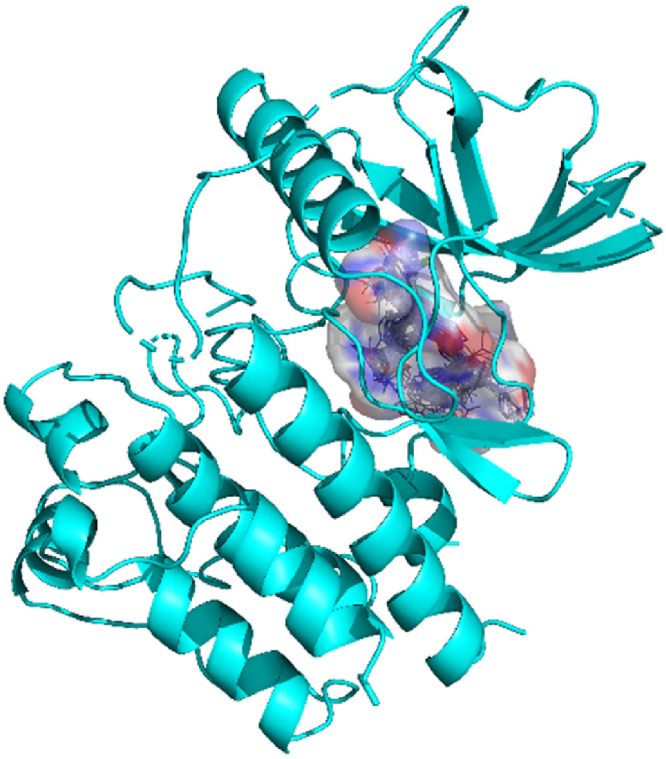


Fig. 3. The binding site of CSF-1R associated all compounds (PDB ID: 3dpk). The surfaces of the potential pocket in CSF-1R are presented in red and blue.

the equilibration phase and the entire production phase. Periodic boundary conditions were used to maintain a constant temperature and pressure set. The pressure was set to 1 atm and controlled by an anisotropic (x-, y-, z-) pressure scaling protocol with a pressure relaxation time of 1 ps? The temperature was adjusted using Langevin dynamics and the collision frequency was 2 ps^{-1} [33,34]. The particle grid Ewald (PME) method was used to deal with long-distance static electricity, and a 10 Å cutoff value was set to deal with real-space interactions [35,36]. All covalent bonds involving hydrogen atoms were constrained using the SHAKE algorithm [37]. Each system undergoes an MD simulation of 100 ns, and the trajectory of the simulated system was saved every 100 ps?

For the saved MD simulation trajectories, the MM/GBSA and MM/PBSA methods were used to calculate the binding energy of receptors treated with different ligands [38–40]. A total of 200 snapshots were taken from 80 to 100 ns to calculate the average binding energy. Calculated as follows:

$$\Delta E_{bind} = \Delta E_{MM} + \Delta E_{SO}L = \Delta E_{MM} + \Delta E_{GB} + \Delta E_{SA}$$

$$\Delta G_{bind} = G_{complex} - G_{protein} - G_{ligand} = \Delta H - T\Delta S \approx \Delta G_{gas} + \Delta G_{sol} - T\Delta S$$

$$\Delta G_{gas} = \Delta E_{ele} + \Delta G_{vdw};\ \Delta G_{sol} = \Delta_{GPB/GB} + \Delta G_{SA}$$

In the above formula, ΔG_{bind} was the final binding free energy, and $G_{complex}$, $G_{protein}$, and G_{ligand} were the free energy of the complex, CSF-1R protein, and ligand, respectively. ΔG_{gas} was the gas-phase interaction energy between a protein and a ligand, and it consisted of ΔE_{ele} (electrostatic energy) and ΔE_{vdw} (van der Waals force). ΔG_{sol} was the solvation free energy, which was the sum of the electrostatic solvation energy $\Delta G_{PB/GB}$ (polar contribution) and the non-static solvation energy ΔG_{SA} (non-polar contribution). In addition, the energy of each residue was broken down into the main chain and side chain atoms, and the energy decomposition could be analyzed to determine the contribution of key residues to binding [41].

3. Results and analysis

3.1. Analysis of CoMFA and CoMSIA models

With 37 Dihydropyrimido [4,5-d] pyrimidines as the training set, the CoMFA and CoMSIA modeling results were shown in Table 2 (31 possible combinations of COMSIA molecular fields were shown in Supplementary Material Table S1). In the 3D-QSAR study, if the model has a q^2 greater than 0.5 and an r^2 greater than 0.8, it indicates that it has better internal prediction capabilities.

In the CoMFA model, the combination of electrostatic and stereo fields (q2 = 0.626, r2 = 0.998) had a high Ficsher ratio. The CoMFA model was reliable. In the CoMSIA model, the combination of the steric field, electrostatic field, hydrophobic field, and hydrogen-bond acceptor field performed well among the 31 possible field combinations (q2 = 0.684, r2 = 0.997). At the same time, the rest of its parameters also proved that (F = 1001, SEE = 0.024). The steric field of CoMFA, the electrostatic field and the hydrophobic field of CoMSIA contributed more to our model, and were used to guide the modification of these CSF-1R inhibitors. The linear correlation analysis between the training set and the test set of the 3D-QSAR model was shown in Fig. 4. In the CoMFA and CoMSIA models, most of the compounds in the training and test sets were located at or near the trend line, which proved that the actual activity value of the compound and the predicted activity value (represented by pIC 50) had a good fit. Moreover , The low q^2 and r^2 values obtained using Y-randomization test indicated that the good results of our original models were not due to the accidental correlationor structural dependence of the training set (Supplementary material Table S5).

3.2. Counter maps analysis of CoMFA and CoMSIA

To explore the relationship between structure and activity of compounds, the most active compound 26 was used as a template for counter maps analysis of CoMFA and CoMSIA. The contour maps of the CoMFA steric field, COMSIA electrostatic field and hydrophobic field, which had a higher contribution in the 3D-QSAR model, were selected, as shown in Fig. 5.

Fig. 5A was a steric field contour map of CoMFA. The combination of green and yellow regions indicated that at the methyl group on the piperazine ring, we should replace it with a larger volume of groups. These have been confirmed by certain compounds in the model. In the case of compounds 37 and 47, compound 47 (pIC_{50} = 5.47) was substituted with a larger oxetanyl group, which could better fill the binding pocket and made the binding more stable. And its activity was significantly higher than that of compound 37 substituted with acetaldehyde (pIC_{50} = 4.59). Fig. 5B was an electrostatic field counter map of CoMSIA. The blue area

Table 2
Summary of 3D-QSAR results.

Method	LOO		PLS			Contribution of force fields				
	n	q^2	r^2	SEE	*F*	S	E	H	D	A
CoMFA	10	0.626	0.998	0.019	1562	0.595	0.405	–	–	–
CoMSIA	10	0.684	0.997	0.024	1001	0.113	0.386	0.387	–	0.113

Note: *n*: the optimum number of components; S: Steric; E: Electrostatic; H: Hydrophobic; D: Hydrogen bond donor; A: Hydrogen bond acceptor; F: the fischer ratio; SEE: the standard deviation of the regression.

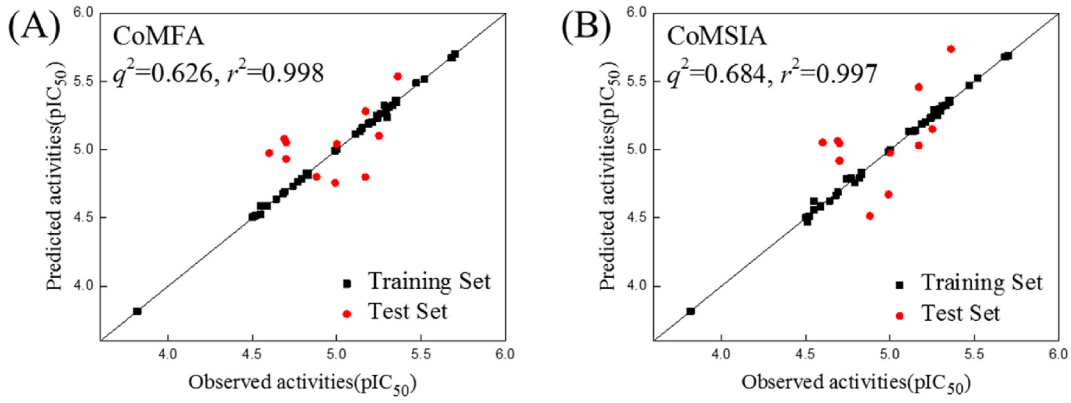


Fig. 4. Correlation between training set and test set observed value data and predicted value data. (A) CoMFA model results; (B) CoMSIA model results.

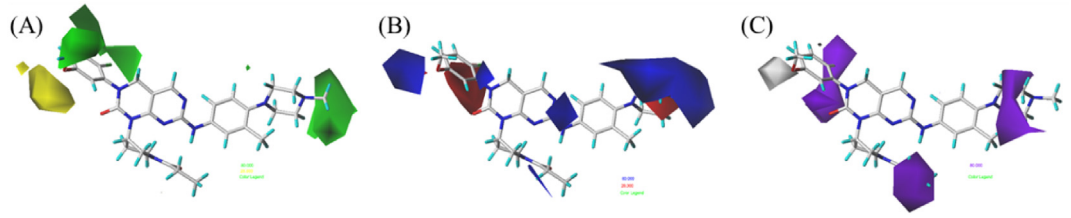


Fig. 5. Three-dimensional equipotential map of CoMFA and CoMSIA of 26 compound. (A): CoMFA Steric field; (B): CoMSIA Electrostatic field; (C) CoMSIA hydrophobic field.

indicated that increasing the positive charge here was beneficial to increase the activity of the compound, and the red area indicated that increasing the electronegativity was beneficial to increase the activity. This combination indicated that substitution of the 2-position of the benzene ring in the upper left corner with a negatively charged group was favorable for the activity. Compound 28 (pIC_{50} = 5.68) was more negatively charged with −F substitution, and had better activity than unmodified compound 21 (pIC_{50} = 5.35). The hydrophobic field of CoMSIA was shown in Fig. 5C. The purple area near the propionaldehyde substitution on the pyrrolidine ring indicated that increasing the hydrophobicity of the substituent there was conducive to the improvement of the activity. Compounds 2 and 9 verified this. The hydroxyacetaldehyde group of compound 9 (pIC_{50} = 4.70) was a more typical hydrophilic group, and its activity was significantly reduced. Compound 2 (pIC_{50} = 5.30) was substituted with propionaldehyde, which was more hydrophobic and more conducive to activity.

3.3. Analysis of molecular docking

In order to explore the binding mode between CSF-1R and ligands, the Surflex-Dock module of SYBYL software was used for molecular docking. Before docking all compounds, the co-crystallizing ligand Q8C5 of CSF-1R was redocked into the binding pocket to verify whether the methods and parameters we used were reliable and feasible. As shown in Figure S1(Supplementary Material), the redocked configuration of Q8C5 was almost the same as its co-crystallized ligand configuration and the RMSD of two conformations was 0.4 Å (<2.0 Å). The NH linker and pyrimidine ring in the compound formed a hydrogen bond with the CSF-1R residue Cys666. This characteristic interaction was consistent with previous studies. This showed that the generated docking protocol was reliable and could be used for subsequent research. Dock all the compounds into CSF-1R, as shown in Figure S3(Supplementary Material). The docking score of all the compounds are listed in Table 1.

We selected compound 26 as a template for analyzing binding mode and displayed it in two (Supplementary Material Figure S2) and three-dimensional plots (Fig. 6). First, compound 26 formed a hydrogen bond with the oxygen atom in the main chain of residue Cys666. Interestingly, Cys666 formed another hydrogen bond with the N atom in the pyrimidine ring of compound 26. This combination was exceptionally strong. Moreover, there was a rich hydrophobic interaction between compound 26 and CSF-1R. The residue Leu588 had a hydrophobic interaction with the benzene ring, pyrimidine ring and pyrrolidine ring of the ligand, which was extremely advantageous for binding. In addition, residues Phe797, Val596, and Leu785 all had hydrophobic interactions with CSF-1R. Rich hydrophobic interactions and hydrogen bonding played a key role in their binding, and they could make the binding of compounds to receptors more stable.

3.4. Molecular design and activity prediction

By analyzing the counter map of 3D-QSAR and the results of molecular docking, we determined the modified regions of this class of compounds. Taking compound 26 as a lead compound, a larger volume group was introduced near the piperazine ring, and a hydrophobic group was introduced near the pyrrolidine ring to seek more hydrophobic interaction with the receptor. In this study, we modified and designed 9 new compounds (Slimes of these compounds are listed in Supplementary Material Table S3). First, we predicted their activity through the established 3D-QSAR model. Second, we docked and scored them. The structure predicted activity and docking score of the compounds were shown in Table 3. The results showed that all new compounds had high inhibitory activity and high docking scores. Both of them confirm each other, which showed the reliability of this research.

3.5. Molecular dynamics simulations and binding free energy analysis

We selected compound a4 with a higher docking score and compound 26 with the highest actual activity value. Using the

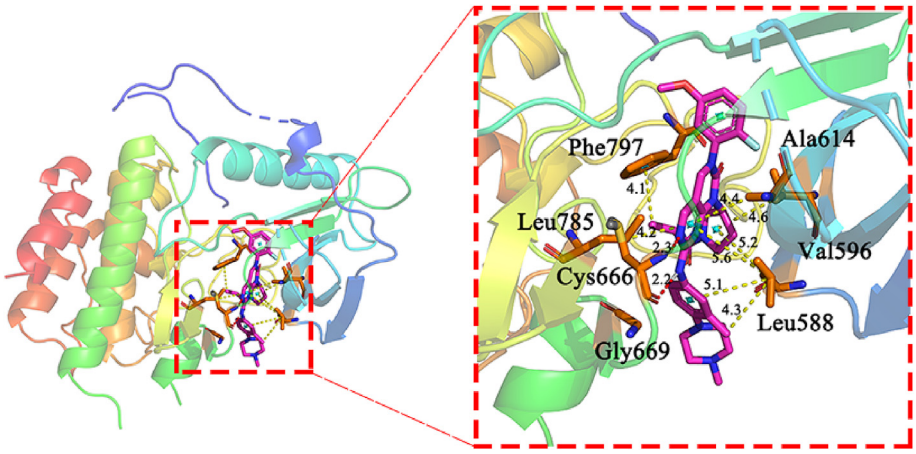
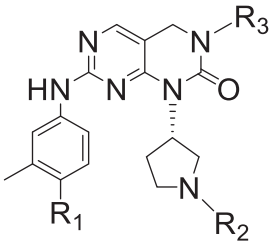


Fig. 6. Docking model of CSF-1R with compound 26. The hydrogen bonds were shown as red dashed lines and Hydrophobic bonds were shown as yellow dashed lines. Key residues are shown in orange and compound 26 is shown in purple.

Table 3
Structures and predicted activities of new designed compounds.



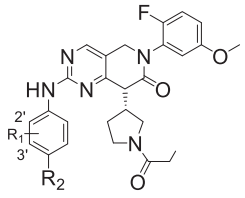
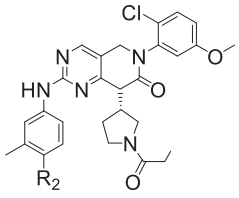
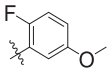
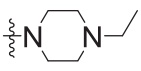
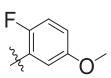
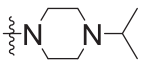
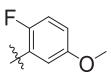
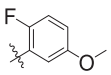
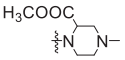
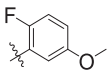
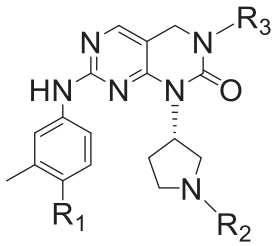
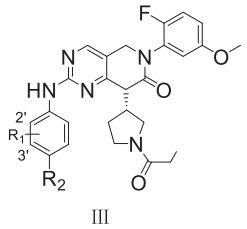
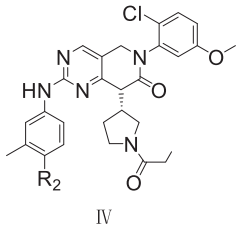
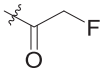
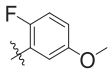
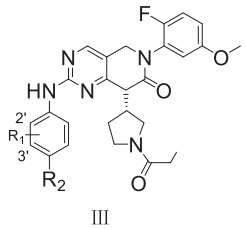
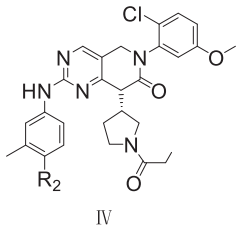
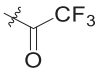
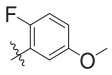
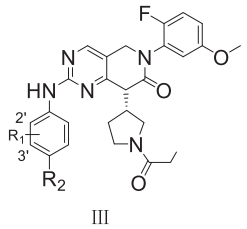
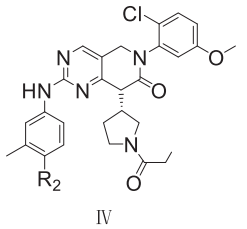
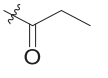
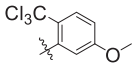
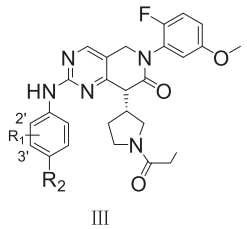
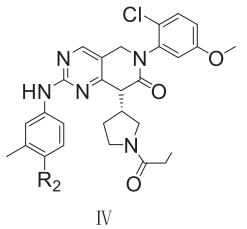
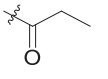
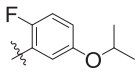
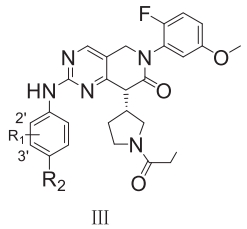
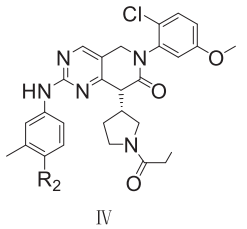
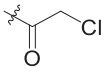
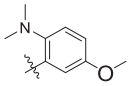
No.	R_1	R_2	R_3	Predicted activity (pIC_{50})		score	Synthetic accessibility
				CoMFA	CoMSIA		
26	 III  IV			5.701	5.69	11.469	5.24
a1				5.746	5.715	11.213	5.36
a2				5.710	5.724	11.393	5.49
a3				5.887	6.022	9.670	5.57
a4				6.014	5.746	12.150	5.77

Table 3 (*continued*)



No.	R_1	R_2	R_3	Predicted activity (pIC_{50})		score	Synthetic accessibility
				CoMFA	CoMSIA		
a5	 III  IV			5.841	5.815	11.179	5.15
a6	 III  IV			5.717	5.886	11.423	5.15
a7	 III  IV			5.795	5.757	6.723	5.32
a8	 III  IV			5.812	5.718	6.936	5.50
a9	 III  IV			5.768	5.817	9.599	5.46

complex formed by docking them with CSF-1R as the initial conformation, a 100ns MD simulation was performed to further understand the detailed dynamic binding mode. The overall convergence and stability of the MD simulation system were monitored by the root mean square deviation (RMSD) of the skeleton atoms (C, Ca, N, and O) relative to the initial docking structure (Fig. 7). As shown in Fig. 7A, after 50 ns, the RMSD of the complex formed by the receptor and compound 26 fluctuates between ~2.5 Å and ~3.3 Å. The RMSD of CSF-1R and compound a4 complex (Fig. 7B) was stable between ~ 1.8 Å and ~3.7 Å after 50 ns?

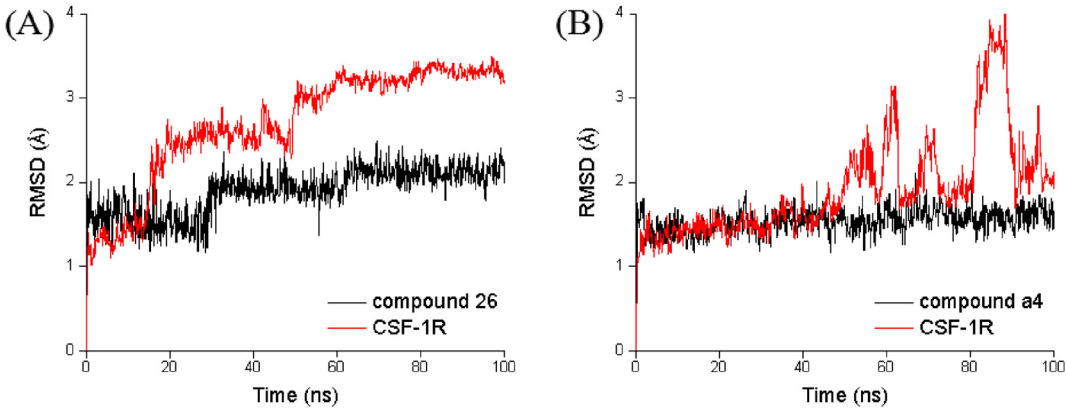


Fig. 7. RMSDs of complex CSF-1R/compounds (26, a4) for 100 ns MD simulations.

Besides, the root-mean-square fluctuations (RMSFs) versus the residue number of all the compounds are shown in Figure S4(Supplementary Material). These indicated that they were already in a metastable state and could be used for subsequent analysis.

The stable conformation after MD simulation of the two complexes was shown in Fig. 8. Fig. 8A showed that the N atom and NH linker on the pyrimidine ring of compound 26 formed a hydrogen bond with Cys666 at a distance of 3.0 Å and 3.1 Å, respectively. At the same time, compound 26 could have a rich hydrophobic interaction with CSF-1R (Leu588, Phe797, Val596, ALA614, and Leu785). Especially Leu588 and Val596, they had hydrophobic bonds with the benzene ring and pyrrolidine ring of compound 26, which can increase the binding affinity of compound 26 to CSF-1R. Fig. 8B showed that the binding mode of compound a4 was basically similar to that of compound 26. The pyrimidine ring in the molecule formed a hydrogen bond with Cys666, and the oxygen atom on the dihydropyrimidine ring formed a hydrogen bond with Lys 616. In addition, the benzene ring, pyrimidine ring, and pyrrole of the molecule Alkanes could form rich hydrophobic interactions with CSF-1R (Leu588, Phe797, Val596, ALA614, and Leu785). Especially Leu588, Leu785 and Val596, they had multiple

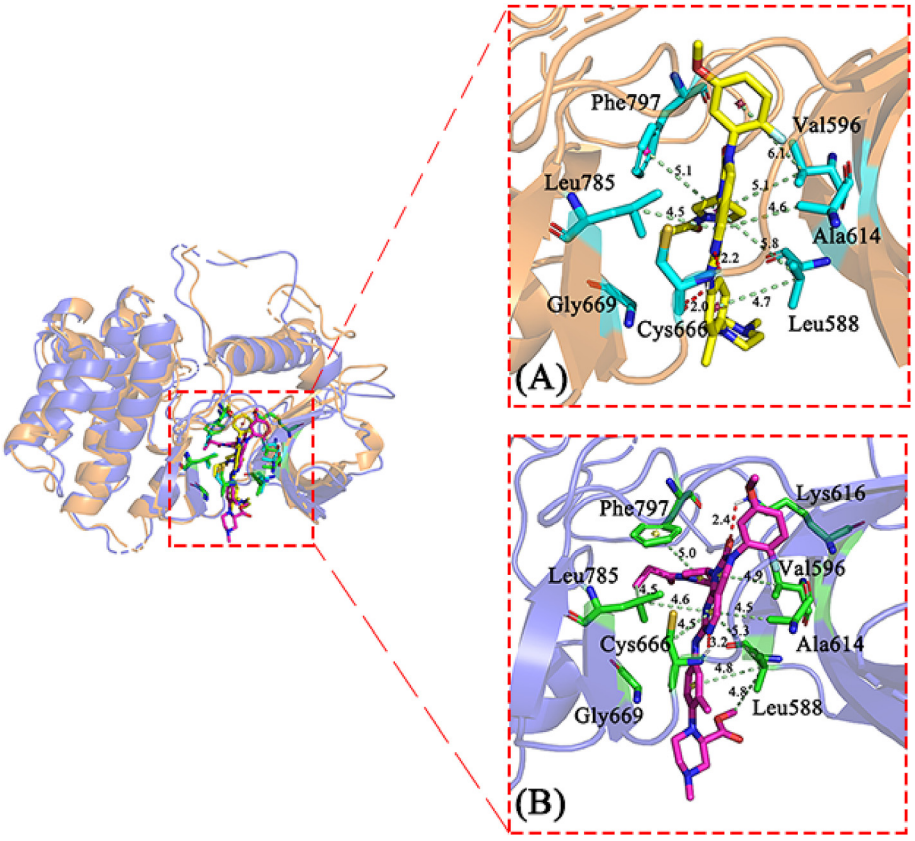


Fig. 8. Analysis of the binding mode of ligand molecules to CSF-1R. (A) Interaction between compound 26 and CSF-1R after 100 ns MD. Key residues are presented in cyan and compound 26 is presented in yellow. (B) Interaction between compound a4 and CSF-1R after 100 ns MD. Key residues are presented in green and compound a4 is presented in purple. The hydrogen bonds were shown as red dashed lines and Hydrophobic bonds were shown as palegreen dashed lines.

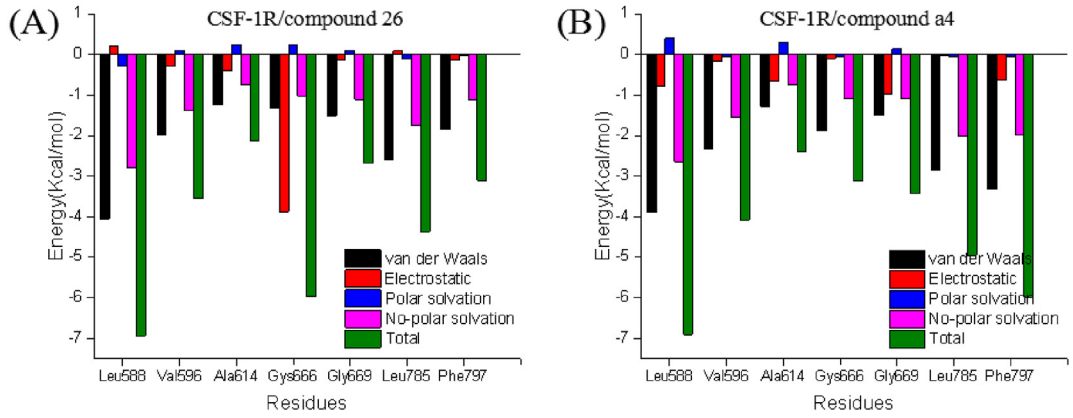


Fig. 9. Contributions of free energy calculated by MM/GBSA method for key residues of CSF-1R.

Table 4
The binding free energies and the individual energy components (kcal · mol^{-1}).

Method	Compound	ΔG_{vdw}	ΔG_{ele}	ΔG_{pol}	ΔG_{nonpol}	ΔG_{bind}
MM/PBSA	26	−36.137	−15.580	26.276	0	−25.439
	a4	−41.507	−18.710	30.985	0	−29.230

hydrophobic interactions with the benzene ring, pyrimidine ring and pyrrolidine ring of the compound. These results indicated that our compounds bind to CSF-1R mainly through hydrophobic interactions and hydrogen bonding.

After MD simulation, we calculated the interaction energy between the ligand molecule and the key amino acids of CSF-1R through MM/GBSA (Fig. 9). The calculation results showed that Leu588, Val596, Ala614, Gys666, Gly669, Leu785, Phe797 in CSF-1R all had strong interactions with compound 26. In particular, Leu588 had interaction energy with the ligand of fewer than −6 kcal mol^{-1} (Fig. 9A), which was related to its ability to form multiple hydrophobic bonds with the ligand. Similarly, compound a4 had the lowest interaction energy with Leu588, which reduced the energy of the entire system by more than 6 kcal mol^{-1} (Fig. 9B); In addition, Val596, Ala614, Gys666, Gly669, Leu785 and Phe797 also had lower binding energies. This made the system more stable. Subsequently, we used the MM/PBSA method to accurately calculate the binding free energy of compound 26, a4 and CSF-1R (Table 4). The results showed that the binding free energy of compound 26 (−25.44 kcal mol^{-1}) was relatively high. Not surprisingly, compound a4 had a lower free binding energy (−29.23 kcal mol^{-1}), indicating that compound a4 compound a4 had a better inhibitory effect on CSF-1R.

3.6. ADMET profiles

We used the pkCSM online server to predict the ADME/T characteristics of the designed compounds (Table 5). The intestinal absorption rate of the compound was in the range of 88.896%–97.084% and shown very high absorption characteristics, especially for the compound a8. The logVDss of most compounds were higher than −0.15, which meant that they were more distributed in tissues than in plasma. For metabolism, all compounds were predicted to be substrates for CYP450 3A4 isoform, which meant that they might be metabolized by CYP 3A4. In addition, all compounds might not inhibit CYP450 1A2 isoform, but they might inhibit CYP450 3A4 isoform. Based on the predicted total clearance rate, liver and kidney tissue could be used to clear all compounds in combination. The expected toxicity indicated that all compounds might be harmful to the liver and that none of them caused skin irritation and mutagenicity. All design molecules with synthetic accessibility scores between 5.15 and 5.45 showed lower structural complexity, thus demonstrating synthetic feasibility (Table 3). Computational pharmacokinetic and toxicological studies and accessible synthetic methods indicated that virtually designed compounds could be used as lead compounds for further development.

4. Conclusion

CSF-1R has attracted much attention as a potential anti-cancer target. In this study, we combined 3D-QSAR, molecular docking, and MD simulation to study dihydropyrimido [4, 5-d] pyrimidine derivatives. The 3D-QSAR model established in this study has high

Table 5
ADMET properties of Novel designed molecules.

No	Absorption	Distribution	Metabolism CYP							Excretion	Toxicity		
	Intestinal Absorption (human) (%absorbed)	VDss (human) (logL/kg)	2D6	3A4	1A2	2C19	2C9	2D6	3A4	Total Clearance (log mL/min/kg)	AMES	Hepatotoxicity	Skin Sensitization
			Substrate		Inhibitor								
a1	96.272	1.211	YES	YES	NO	NO	YES	NO	YES	0.421	NO	YES	NO
a2	97.065	1.193	YES	YES	NO	YES	NO	NO	YES	0.438	NO	YES	NO
a3	93.210	0.998	NO	YES	NO	NO	YES	NO	YES	0.256	NO	YES	NO
a4	88.896	0.934	NO	YES	NO	NO	NO	NO	YES	0.268	NO	YES	NO
a5	94.821	0.977	NO	YES	NO	NO	NO	NO	YES	0.309	NO	YES	NO
a6	95.744	0.894	YES	YES	NO	NO	NO	NO	YES	−0.070	NO	YES	NO
a7	96.529	1.058	YES	YES	NO	YES	YES	NO	YES	0.209	NO	NO	NO
a8	97.084	0.933	NO	YES	NO	YES	YES	NO	YES	0.326	NO	YES	NO
a9	90.400	1.186	YES	YES	NO	NO	YES	NO	YES	0.483	NO	YES	NO

predictive power, revealing the key structural factors of CSF-1R inhibitors. Based on the 3D-QSAR and molecular docking results, we designed and predicted 9 new dihydropyrimido [4, 5-d] pyrimidine derivatives. The results indicated that these compounds had good predictive activity and a fairly good ADME/T curve. Molecular docking and MD simulation results showed that the binding of compounds to CSF-1R was mainly achieved through hydrophobic and hydrogen bonding. In the course of our research, we also found some key residues of the binding site (such as Leu588 and Gys666). Free energy decomposition showed that compound a4 had a better inhibitory effect on CSF-1R. These results provide a useful reference for the rational design of such compounds.

Disclosure statement

The authors declare no conflict of interest, financial or otherwise.

Declaration of competing interest

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

CRediT authorship contribution statement

Han Chu: Conceptualization, Writing - original draft. **Qing-xiu He:** Data curation, Visualization. **Juan Wang:** Software. **Yong Hu:** Project administration. **Yuan-qiang Wang:** Writing - review & editing. **Zhi-hua Lin:** Writing - review & editing, Funding acquisition.

Acknowledgments

This work was supported by the [national natural science foundation of China] under Grant [81171508, 31400667]; [Key project of Chongqing natural science foundation] under Grant [cstc2018jcyjAX0683, cstc2015jcyjBX0080]; and [Scientific and technological research project of Chongqing municipal education commission] under Grant [KJZD-K201801102, KJ1600908, KJQN201801132].

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.molstruc.2020.128617.

References

[1] F. Peyraud, S. Cousin, A. Italiano, CSF-1R inhibitor development: current clinical status, Curr. Oncol. Rep. 19 (2017) 70.
[2] E.R. Stanley, V. Chitu, CSF-1 receptor signaling in myeloid cells, Cold Spring Harb Perspect Biol 6 (2014) a021857-a021857.
[3] C.J. Sherr, C.W. Rettenmier, R. Sacca, M.F. Roussel, A.T. Look, E.R. Stanley, The c-fms proto-oncogene product is related to the receptor for the mononuclear phagocyte growth factor, CSF 1, Cell 41 (1985) 665–676.
[4] K.A. Mouchemore, F.J. Pixley, CSF-1 signaling in macrophages: pleiotrophy through phosphotyrosine-based signaling pathways, Crit. Rev. Clin. Lab Sci. 49 (2012) 49–61.
[5] S.R. Hubbard, J.H. Till, Protein tyrosine kinase structure and function, Annu. Rev. Biochem. 69 (2000) 373–398.
[6] R. Ostuni, F. Kratochvill, P.J. Murray, G. Natoli, Macrophages and cancer: from mechanisms to therapeutic implications, Trends Immunol. 36 (2015) 229–239.
[7] F.J. Pixley, Macrophage migration and its regulation by CSF-1, Int. J. Cell Biol. (2012) 2012.
[8] S. Aharinejad, M. Salama, P. Paulus, K. Zins, A. Berger, C.F. Singer, Elevated CSF1 serum concentration predicts poor overall survival in women with early breast cancer, Endocr. Relat. Canc. 20 (2013) 777–783.
[9] B. Mroczko, M. Groblewska, U. Wereszczyńska-Siemiątkowska, B. Okulczyk, B. Kędra, W. Łaszewicz, A. Dąbrowski, M. Szmitkowski, Serum macrophage-colony stimulating factor levels in colorectal cancer patients correlate with lymph node metastasis and poor prognosis, Clin. Chim. Acta 380 (2007) 208–212.
[10] M. Kowalska, J. Tajer, M. Chechlinska, M. Fuksiewicz, B. Kotowicz, J. Kaminska, J. Walewski, Serum macrophage colony-stimulating factor (M-CSF) in patients with Hodgkin lymphoma, Med. Oncol. 29 (2012) 2143–2147.
[11] M. De Palma, C.E. Lewis, Macrophage regulation of tumor responses to anticancer therapies, Canc. Cell 23 (2013) 277–286.
[12] B.-Z. Qian, J.W. Pollard, Macrophage diversity enhances tumor progression and metastasis, Cell 141 (2010) 39–51.
[13] A. Mantovani, F. Marchesi, A. Malesci, L. Laghi, P. Allavena, Tumour-associated macrophages as treatment targets in oncology, Nat. Rev. Clin. Oncol. 14 (2017) 399.
[14] S.R. Nielsen, V. Quaranta, A. Linford, P. Emeagi, C. Rainer, A. Santos, L. Ireland, T. Sakai, K. Sakai, Y.-S. Kim, Corrigendum: macrophage-secreted granulin supports pancreatic cancer metastasis by inducing liver fibrosis, Nat. Cell Biol. 18 (2016) 822.
[15] M. Sluijter, T.C. van der Sluis, P.A. van der Velden, M. Versluis, B.L. West, S.H. van der Burg, T. van Hall, Inhibition of CSF-1R supports T-cell mediated melanoma therapy, PloS One 9 (2014).
[16] M.A. Cannarile, M. Weisser, W. Jacob, A.-M. Jegg, C.H. Ries, D. Rüttinger, Colony-stimulating factor 1 receptor (CSF1R) inhibitors in cancer therapy, J. Immunother. Canc. 5 (2017) 53.
[17] C.H. Ries, S. Hoves, M.A. Cannarile, D. Rüttinger, CSF-1/CSF-1R targeting agents in clinical development for cancer therapy, Curr. Opin. Pharmacol. 23 (2015) 45–51.
[18] D.A. Hume, K.P. MacDonald, Therapeutic applications of macrophage colony-stimulating factor-1 (CSF-1) and antagonists of CSF-1 receptor (CSF-1R) signaling, Blood, J. Am. Soc. Hematol. 119 (2012) 1810–1820.
[19] M.R. Elmore, A.R. Najafi, M.A. Koike, N.N. Dagher, E.E. Spangenberg, R.A. Rice, M. Kitazawa, B. Matusow, H. Nguyen, B.L. West, Colony-stimulating factor 1 receptor signaling is necessary for microglia viability, unmasking a microglia progenitor cell in the adult brain, Neuron 82 (2014) 380–397.
[20] N. Butowski, H. Colman, J.F. De Groot, A.M. Omuro, L. Nayak, P.Y. Wen, T.F. Cloughesy, A. Marimuthu, S. Haidar, A. Perry, Orally administered colony stimulating factor 1 receptor inhibitor PLX3397 in recurrent glioblastoma: an Ivy Foundation Early Phase Clinical Trials Consortium phase II study, Neuro Oncol. 18 (2015) 557–564.
[21] J.A. Krauser, Y. Jin, M. Walles, U. Pfaar, J. Sutton, M. Wiesmann, D. Graf, V. Pflimlin-Fritschy, T. Wolf, G. Camenisch, Phenotypic and metabolic investigation of a CSF-1R kinase receptor inhibitor (BLZ945) and its pharmacologically active metabolite, Xenobiotica 45 (2015) 107–123.
[22] M. Wiesmann, D.L. Daniel, N. Pryer, J. Sutton, J. Sim, BLZ945, a selective c-fms (CSF-1R) kinase inhibitor for the suppression of tumor-induced osteolytic lesions in bone, Canc. Res. 70 (2010), 3629-3629.
[23] Q. Xun, Z. Zhang, J. Luo, L. Tong, M. Huang, Z. Wang, J. Zou, Y. Liu, Y. Xu, H. Xie, Design, synthesis, and structure–activity relationship study of 2-oxo-3, 4-dihydropyrimido [4, 5-d] pyrimidines as new colony stimulating factor 1 receptor (CSF1R) kinase inhibitors, J. Med. Chem. 61 (2018) 2353–2371.
[24] R.D. Cramer, D.E. Patterson, J.D. Bunce, Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins, J. Am. Chem. Soc. 110 (1988) 5959–5967.
[25] G. Klebe, U. Abraham, T. Mietzner, Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity, J. Med. Chem. 37 (1994) 4130–4146.
[26] R. Abdizadeh, F. Hadizadeh, T. Abdizadeh, QSAR analysis of coumarin-based benzamides as histone deacetylase inhibitors using CoMFA, CoMSIA and HQSAR methods, J. Mol. Struct. 1199 (2020) 126961.
[27] H. Chu, Q.X. He, J.W. Wang, Y.T. Deng, J. Wang, Y. Hu, Y.Q. Wang, Z.H. Lin, 3D-QSAR, molecular docking, and molecular dynamics simulation of a novel thieno[3,4-d]pyrimidine inhibitor targeting human immunodeficiency virus type 1 reverse transcriptase, J Biomol Struct Dyn, 2019, pp. 1–12.
[28] E. Vrontaki, G. Melagraki, T. Mavromoustakos, A. Afantitis, Exploiting ChEMBL database to identify indole analogs as HCV replication inhibitors, Methods 71 (2015) 4–13.
[29] H. Guo, Y. Wang, Q. He, Y. Zhang, Y. Hu, Y. Wang, Z. Lin, In silico rational design and virtual screening of antixoidant tripeptides based on 3D-QSAR modeling, J. Mol. Struct. 1193 (2019) 223–230.
[30] Q. He, H. Chu, Y. Wang, H. Guo, Y. Wang, S. Wang, Z. Feng, X.-Q. Xie, Y. Hu, H. Liu, In silico design novel vibsanin B derivatives as inhibitor for heat shock protein 90 based on 3D-QSAR, molecular docking and molecular dynamics simulation, J. Biomol. Struct. Dyn. (2019) 1–12.
[31] A.W. Götz, M.J. Williamson, D. Xu, D. Poole, S. Le Grand, R.C. Walker, Routine microsecond molecular dynamics simulations with AMBER on GPUs. 1. Generalized born, J. Chem. Theor. Comput. 8 (2012) 1542–1555.
[32] R. Salomon-Ferrer, A.W. Götz, D. Poole, S. Le Grand, R.C. Walker, Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh Ewald, J. Chem. Theor. Comput. 9 (2013) 3878–3888.
[33] J.A. Izaguirre, D.P. Catarello, J.M. Wozniak, R.D. Skeel, Langevin stabilization of molecular dynamics, J. Chem. Phys. 114 (2001) 2090–2098.
[34] R.J. Loncharich, B.R. Brooks, R.W. Pastor, Langevin dynamics of peptides: the frictional dependence of isomerization rates of N-acetylalanyl-N′-methylamide, Biopolymers: Orig. Res. Biomol. 32 (1992) 523–535.

[35] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, L.G. Pedersen, A smooth particle mesh Ewald method, J. Chem. Phys. 103 (1995) 8577–8593.
[36] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: an N· log (N) method for Ewald sums in large systems, J. Chem. Phys. 98 (1993) 10089–10092.
[37] J.-P. Ryckaert, G. Ciccotti, H.J. Berendsen, Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, J. Comput. Phys. 23 (1977) 327–341.
[38] G.D. Hawkins, C.J. Cramer, D.G. Truhlar, Parametrized models of aqueous free energies of solvation based on pairwise descreening of solute atomic charges from a dielectric medium, J. Phys. Chem. 100 (1996) 19824–19839.
[39] P.A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models, Acc. Chem. Res. 33 (2000) 889–897.
[40] F. Chen, H. Liu, H. Sun, P. Pan, Y. Li, D. Li, T. Hou, Assessing the performance of the MM/PBSA and MM/GBSA methods. 6. Capability to predict protein–protein binding free energies and re-rank binding poses generated by protein–protein docking, Phys. Chem. Chem. Phys. 18 (2016) 22129–22139.
[41] Y.-q. Wang, W.-w. Lin, N. Wu, S.-y. Wang, M.-z. Chen, Z.-h. Lin, X.-Q. Xie, Z.-w. Feng, Structural insight into the serotonin (5-HT) receptor family by molecular docking, molecular dynamics simulation and systems pharmacology analysis, Acta Pharmacol. Sin. 40 (2019) 1138–1156.